Empirical and Computational Inquest into Biosimilar and Next-Generation Insulin Analogs Manufacturing: Unveiling Structural, Analytical, and Dynamic Attributes through Recombinant Technological Approaches

by

Esra Ayan

A Dissertation Submitted to the
Graduate School of Sciences and
Engineering in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy

in

Molecular biology and genetics

KOÇ ÜNİVERSİTESİ

March 6, 2024

Empirical and Computational Inquest into Biosimilar and Next-Generation Insulin Analogs Manufacturing: Unveiling Structural, Analytical, and Dynamic Attributes through Recombinant Technological Approaches

Koc University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a doctoral dissertation by

Esra Ayan

and have found that it is complete and satisfactory in all respects, and that any and all revisions required by the final
examining committee have been made.

Committee Members:

`

Assist. Prof. Dr. Hasan Demirci (Advisor)

`

Assist. Prof. Dr. Ayşe Koca-Çaydaşı

`

Assist. Prof. Dr. Ahmet Katı

`

Prof. Dr. Türkan Haliloğlu

`

Prof. Dr. Takehiko Tosha

Date: 03/06/2024

Dedicated to my beloved mother.

ABSTRACT

Empirical and Computational Inquest into Biosimilar and Next-Generation Insulin Analogs Manufacturing:
Unveiling Structural, Analytical, and Dynamic Attributes through Recombinant Technological Approaches

Esra Ayan

Doctor of Philosophy in Molecular Biology and Genetics
March 6, 2024

The intricate molecular structure of insulin and its nuanced interactions within diverse tissues have carved out a domain within the scientific community dedicated to comprehending and harnessing the significance of this pivotal hormone. Investigative efforts on insulin encompass a broad spectrum of scientific disciplines, spanning cellular signaling, biochemical interactions, and genetic regulation. From this vantage point, insulin assumes a pivotal position at the nexus of modern biotechnology and fundamental science, constructing substantial contributions to the advancements in health and biological sciences. This thesis introduces a biotechnological product and method through the production and characterization of two new-generation domestic insulins but also offers the most comprehensive theoretical and analytical insights into the empirical and computational dynamics of these insulins and oligomeric commercial insulins. Introducing an atlas review encompassing all academic information about insulin, the thesis thoroughly explains the fundamental structural biology over this model protein. Subsequent chapters delve into the experimental and computational aspects of domestic new-generation rapid/ultra-rapid insulins. Furthermore, it presents revolutionary studies on the allosteric hierarchy of commercial long-acting oligomeric insulins. In this regard, this thesis can be regarded as a benchmark work in protein research.

ÖZETÇE

Biyobenzer ve Yeni Nesil İnsülin Analoglarının Üretimine Dair Deneysel ve Hesaplamalı Araştırma: Rekombinant Teknolojik Yaklaşımlarla Yapısal, Analitik ve Dinamik Özelliklerin Ortaya Çıkarılması

Esra Ayan

Doktora, Moleküler Biyoloji ve Genetik

Mart 6, 2024

İnsülinin karmaşık moleküler yapısı ve çeşitli dokulardaki nüanslı etkileşimleri, bu kilit hormonun önemini anlama ve kullanma amacıyla bilimsel bir topluluk içinde özel bir alan oluşturmuştur. İnsülin üzerindeki araştırmalar, hücresel sinyalizasyon, biyokimyasal etkileşimler ve genetik düzenleme gibi geniş bir bilim disiplinleri yelpazesini kapsar. Bu bakış açısından, insülin, modern biyoteknoloji ve temel bilimlerin kesişim noktasında önemli bir konumda bulunarak sağlık ve biyoloji bilimlerindeki ilerlemelere önemli katkılarda bulunur. Bu tez, iki yeni nesil yerli-milli insülinin üretimi ve karakterizasyonu yoluyla bir biyoteknolojik ürün ve yöntem sunmanın yanı sıra, bu insülinlerin ve oligomerik ticari insülinlerin deneysel ve hesaplamalı dinamikleri üzerine en kapsamlı teorik ve analitik içgörüleri de sunar. İnsülin hakkında tüm akademik bilgileri kapsayan bir atlas incelemesi tanıtarak, tez bu model proteinin temel yapısal biyolojisini ayrıntılı bir şekilde açıklar. Dahası, tez, yerli yeni nesil hızlı/ultra hızlı insülinlerin deneysel ve hesaplamalı yönlerine dalış yapar. Ayrıca, ticari uzun etkili oligomerik insülinlerin allosterik hiyerarşisi üzerine devrim niteliğinde çalışmalar sunar. Bu bağlamda, bu tez, protein araştırmalarında bir referans çalışma olarak kabul edilebilir

ACKNOWLEDGMENTS

Insulin production utilized the infrastructure of Koç University Structural Biology and Drug Development Center (KUYBIIGM). The *Turkish Delight* studies were supported by Validebağ Research Park (DETUAM). XFEL investigations received support under SACLA/XFEL proposal number 2023B8058. Serial synchrotron examinations were conducted at the European Synchrotron Radiation Facility (ESRF), while Hydrogen Deuterium Exchange Mass Spectrometry studies were performed at the Molecular Biology-Biotechnology and Genetics Research Center (MOBGAM). Gratitude is extended to all involved parties for their invaluable support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

DM	Diabetes Mellitus
T1D	Type-1 Diabetes
T2D	Type-2 Diabetes
rDNA	Recombinant DNA
GLUT-2	Glucose Transporter 2
PDX-1	Pancreatic Duodenal Homeobox Factor-1
E47/b42	β-cell E box transactivator 2
SRP	Signal Recognition Particles
RER	Rough Endoplasmic Reticulum
ER	Endoplasmic Reticulum
IAPP	islet amyloid polypeptide
PTB	phosphotyrosine binding
IRS	Insulin Receptor Substrates
SH2	Src Homology 2
PI3K	Phosphoinositide 3-Kinase
PIP2	phosphatidylinositol-(4,5)-bisphosphate
PIP3	phosphatidylinositol (3,4,5)-triphosphate
Akt/PKB	Protein Kinase B
PH	Pleckstrin Homology
SGK	Serine/threonine protein kinase
S6K	p70 ribosomal protein S6 Kinase
PDK1	Phosphoinositide Dependent Kinase-1
CAP	c-Cbl-Associated Protein
C3G	Rho family guanine nucleotide exchange factor
GLUT4	Glucose transporter type 4
GYS	glycogen synthase
INS	Insulin
CVD	Cardiovascular Disease
DR	diabetic retinopathy

aHCM	asymptomatic hypertrophic cardiomyopathy
IGF-1	insulin-like growth factor-1
JNK	c-Jun N-terminal kinase
IKK	nuclear factor-κB (IκB) kinase
TH2	T-helper type 2
Igs	immunoglobulins
CLS	crown-like structures
IRS-1	f insulin receptor substrate protein-1
NF-κB	Nuclear Factor-κB
FA	Fatty Acid
GRP120	omega-3 fatty acid receptors
PPAR	peroxisome proliferator-activated receptor
MHC-I	Major Histocompatibility Complex I
IGFR	Insulin-like growth factor receptor
CNS	central nervous system
AD	Alzheimer's Disease
NPH	Neutral Protamine Hagedorn
PPARG	Peroxisome Proliferator Activated Receptor Gamma
CDKAL1	CDK5 Regulatory Subunit Associated Protein 1 Like 1
IGF2BP2	Insulin Like Growth Factor 2 MRNA Binding Protein 2
ZnT8	Zinc transporter 8
INSL3	insulin-like peptide 3
RXFPP2	relaxin family peptide receptor 2
PPARGC1A	promoter of PPARG Coactivator 1 Alpha
IRS2	Insulin Receptor Substrate 2
AKT2	AKT Serine/Threonine Kinase 2
HNF4a	Hepatocyte Nuclear Factor 4 Alpha
IGFBP2	Insulin Like Growth Factor Binding Protein 2
NT5E	5'- Nucleotidase Ecto
PAK7	p21 protein (Cdc42/Rac)-activated kinase 7

TMED	Transmembrane emp24 domain
TSPAN33	Tetraspanin-33
FACS	fluorescence-activated cell sorting
FXYD2	FXYD Domain Containing Ion Transport Regulator 2
GPD2	Glycerol-3-Phosphate Dehydrogenase 2
m6A	mRNA methylation
ALKBH5	AlkB family of proteins
LRP1	lipoprotein receptorrelated protein 1
DCD	diabetic kidney disease
NXT	Naoxintong Capsule
SERT	serotonin transporter
TPMT	thiopurine methyltransferase
ECM	extracellular matrix
DTRI	destripeptide (B28-B30) insulin
GNM	Gaussian Network Model

Chapter 1

INTRODUCTION

This chapter largely excerpted from Ayan et.al. (2022) atlas-review with the copyright permissions of the journal. All the work has been done by myself.

Recombinant insulin is a critical life-saving molecule for the treatment of insulin-dependent diabetes mellitus (DM). Diabetic pathobiology is characterized by insulin therapy. Disorders that are candidates for insulin therapy are typically categorized as juvenile-onset type 1 diabetes (T1D) and adult-onset type 2 diabetes (T2D). Usage of insulin is more common in T2D compared to T1D [Warram et al., 1984]. Recombinant insulin was first produced in the 1970s and commercialized in the 1980s. This is a milestone in insulin history [Goeddel et al., 1979; Johnson, 1983]. Recombinant biosynthesis provides a more robust approach in terms of massive production. An unlimited amount of insulin can be obtained compared to the conventionally obtained native insulin methods, which only provide a limited amount. Along with the recombinant DNA (rDNA) technology, cost-effective insulin analogs have continued to improve the pharmacological aspect of insulin production [Lipska et al., 2010]. Insulin therapies lead to the production of sustained and rapid-acting analogs that represent the state-of-the-art therapy for millions of insulin-dependent diabetic patients [Zaykov et al., 2016]. Furthermore, therapeutic approaches and pharmacokinetic advances in normalizing blood glucose levels to date continue with reports of polypharmacy to ensure the regulation of optimized glucose and body weight management [Glendorf et al., 2011; Edgerton et al., 2014; Gough et al., 2015]. In addition, advances in materials science paved the way to control insulin release through real-time continuous blood glucose monitoring and a closed-loop system for the regulation of blood glucose levels [Mo et al., 2014; Hovorka, 2011]. However, even with such remarkable advancements, current treatments provide suboptimal results for most patients. Failure to develop a complete solution to diabetes caused insulin-oriented studies to remain on the agenda. It

has been repeatedly emphasized by studies, systematic reviews, and meta-analyses that diabetes, which has been considered a major health problem since 1922 [Banting et al., 1922; Macleod, 1922], will increase exponentially in the coming years [Lefever et al., 2021, Fralick and Zinman, 2021]. Innovative techniques in the fields of structural biology and structural proteomics enabled the high-resolution analysis of insulin and insulin-associated proteins. Each novel structure paves the way to new production approaches, optimization techniques, and clinical trials. Insulin is not only a molecule categorized with diabetes, but an important hormone targeted by a wide variety of diseases due to its vital role in the human proteome. Therefore, it will not be sufficient to consider only a single aspect of the multi-faceted insulin molecule. In this chapter, we aimed to evaluate all insulin-oriented details from a holistic perspective: from the history of insulin to its production techniques, from its structure to the signaling pathways, and from diagnosis therapy to omics approaches. Additionally, the literature is characterized by gaps in the knowledge of the structures and dynamics of insulin analogs. Holistically, we also desired to give the bird's eye information about the differences between the analog structures and dynamics.

1.1. Discovery and history of insulin

The discovery of insulin was a milestone in the main treatment for DM. After its discovery, it took another 30 years for recombinant insulin to be optimized as a therapeutic product in clinical use (Figure 1.1).

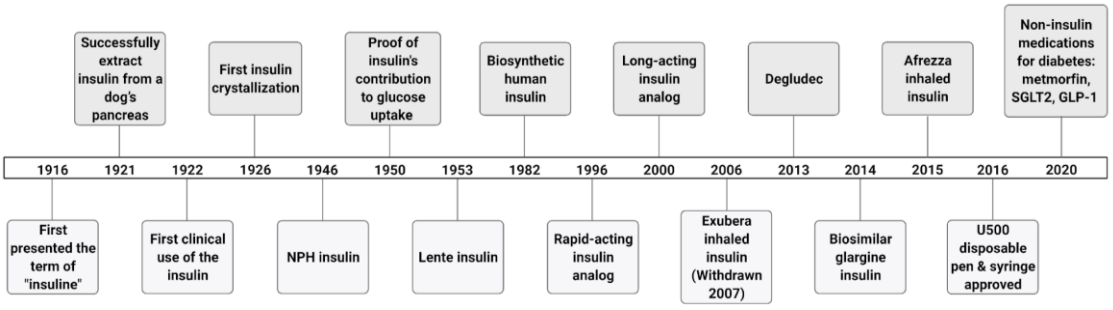


Figure 1.1 Discovery, history, and chronological milestones of insulin development

Insulin's discovery and DM treatment history is a notable story in Michael Bliss's notes [Bliss, 1982]. The name of insulin predates its discovery in 1889. Joseph von Mering and Oskar Minkowski discovered that some model organisms of

pancreatectomy are characterized by severe DM [Shah et al., 1997]. It caused speculation about a substance produced by the pancreas, which is responsible for metabolic regulation, and thought to be responsible for carbohydrate metabolism [Bliss, 1982]. In the following years, 1910 and 1916, Sir Edward Albert Sharpey-Schafer introduced pancreatic islands, regarding Langerhans islands, called "insulin" that could control glucose metabolism [Vecchio et al., 2018]. However, according to some scientists, the discovery of insulin was by Belgian Jean de Meyer. In 1920, Frederick Grant Banting, an orthopedic surgeon, focused on internal secretions and degenerate islets in dogs. First, Frederick Grant Banting and Charles Best worked together on account of John James Rickard Macleod, an expert in carbohydrate metabolism. Critical *in vivo* trials then began in May 1921 and continued until December [Joshi et al., 2007]. Bantings and Best observed that the extract they studied lowered blood sugar in diabetic dogs. James Bertram Collip, who was interested in pancreas studies at that time, also joined this team. Although their presentation about the extract on December 30 (1921) did not give satisfactory results due to their inexperience and haste, the team's further results were promising since the use of the relevant extract caused the restoration of glycogen mobilization in the liver. In January (1922), Banting and Best tried using a pancreatic extract on Leonard Thompson, who had severe DM [Joshi et al., 2007]. This test failed due to slight decreases in glycemia and glycosuria. In addition, this extract test did not show any positive effect on the patient nor his ketoacidosis and caused a sterile abscess formation. After this treatment, they started a new series of injections and this time, Thompson recovered. His ketonuria and glycosuria almost disappeared, and his blood sugar level turned out to be normal. It symbolized conclusive evidence that the extract that was thought to have been obtained by James Bertram Collip replaces the impaired function in DM. He successfully developed a robust method of extraction by optimizing the concentrations of acidic alcohol solutions in the beef pancreas. Eighteen months later, Banting and Macleod won the Nobel Prize in Physiology or Medicine, sharing this glory with Best and Collip [Bliss et al., 1982; Joshi et al., 2007]. Since the discovery of insulin, the focus was brought to biosimilar/biobetter/derivative production methods, including longand short-acting insulin analogs. Currently, the focus of DMoriented studies has been insulin-based therapies, and these therapies dominate the market.

1.2. A Comprehensive Overview

Insulin is a peptide hormone that has an essential regulatory effect. It has an anabolic role in the uptake of glucose into cells and provides carbon energy storage in the body. It is expressed in the β-cells of the pancreas and circulates through the blood. As a result, it is exported to the liver [Tokarz et al., 2018]. β-cells receive the glucose through Glucose Transporter 2 (GLUT-2) receptors after being triggered by internal and external stimuli such as glucose, sulfonylureas, and arginine. After the transport of glucose into the cell, glucokinase, acting as glucose sensor, oxidizes the glucose. If the concentration of glucose is below 90 mg/dL, a threshold known as "substimulatory glucose concentration", insulin release does not occur. This causes effluxing of K+ ions through KATP channels, which are characterized by the negative potential of the β-cell membrane, followed by the closing of the voltage-gated calcium channels. However, if plasma glucose rises above 90 mg/dL, there is an increase in glucose uptake followed by an increase in metabolism, resulting in an increase in ATP concentration and the closing of KATP channels. This means triggering membrane depolarization followed by the opening of voltage-gated calcium channels. Therefore, the calcium influx causes an increase in the intracellular calcium ion concentration, resulting in exocytosis of insulin through microvesicles [Bogardus et al., 1984] (Figure 1.2). On the other hand, once glucose enters β-cells, it triggers various regulators of the insulin promoter, such as pancreatic duodenal homeobox factor-1 (PDX-1), β-cell E box transactivator 2 (E47/b42), and basic leucine zipper protein MafA to stimulate transcription of INS gene [Howell and Taylor, 1968] and some protooncogenes such as c-Abl to inhibit GLUT-2 production [Brissova et al., 2002], causing activation of the feedback mechanism [Koranyi et al., 1992] (Figure 1.2).

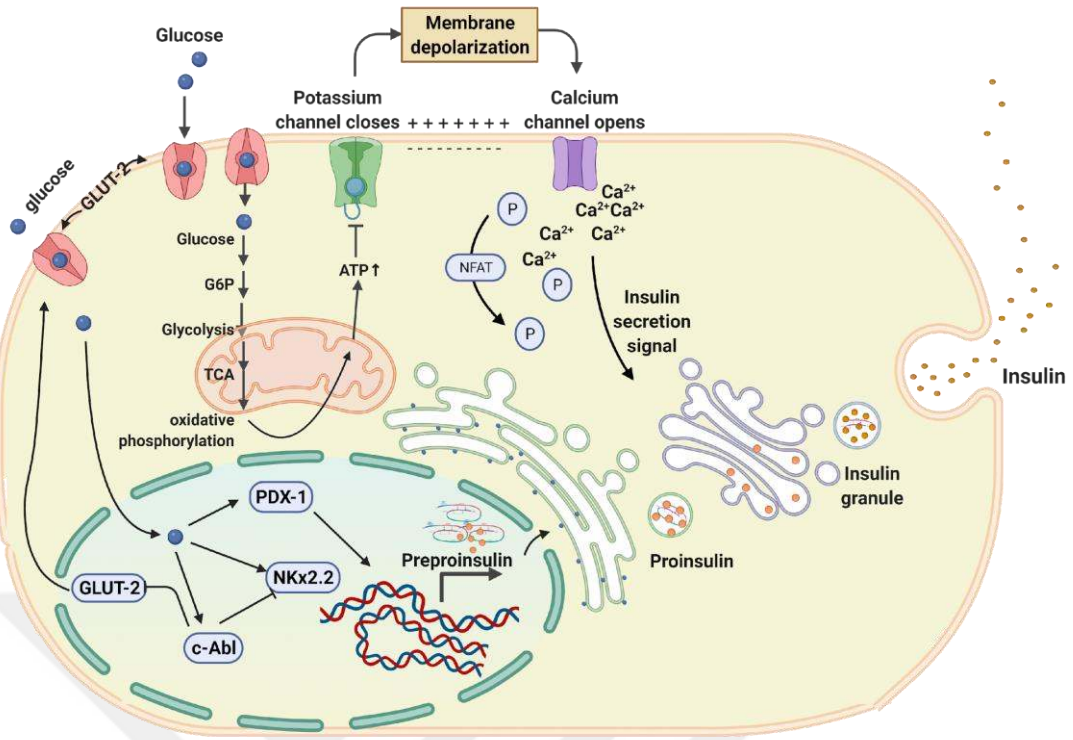


Figure 1.2 Endogenous insulin production pathway in β-cells

Demonstration of insulin secretion through both Ca+2-mediated membrane depolarization and promoter-mediated feedback mechanisms.

The production of insulin synthesized as pre-proinsulin from a 110 amino acid precursor is performed by the transcription and translation of the INS gene on chromosome 11 [Harper et al., 1981]. Pre-proinsulin has a 24-amino-acid-long N-terminal signal peptide recognized by cytoplasmic Signal Recognition Particles (SRP) to be localized to the Rough Endoplasmic Reticulum (RER) membrane. The processes in the RER are characterized by cleavage of this SRP sequence by the signal peptidase, folding of proinsulin, and formation of three disulfide bonds. These processes need a wide range of Endoplasmic Reticulum (ER) oxidoreductases and chaperones, such as protein-thiol reductase. After the ER-mediated oxidative folding process, proinsulin is transported to the Golgi apparatus by microvesicles for further processing, which includes cleavage to mature insulin and C-peptide and the formation of pale secretory "progranules," containing proinsulin-zinc hexamers. This results in the formation of mature insulin [Lemaire et al., 2009]. After C-peptide and insulin are stored together with some cell secretion products and islet amyloid polypeptide (amylin or IAPP) [Fu et al., 2013], proinsulin is converted to mature insulin using prohormone convertases 2 & 3 and carboxypeptidase E. This results in the secretion of mature granules that form

water-insoluble insulin crystals from the β-cell into the bloodstream through microtubules and microfilament dependent transport [Joshi et al., 2007, Lemaire et al., 2009] (Figure 1.3).

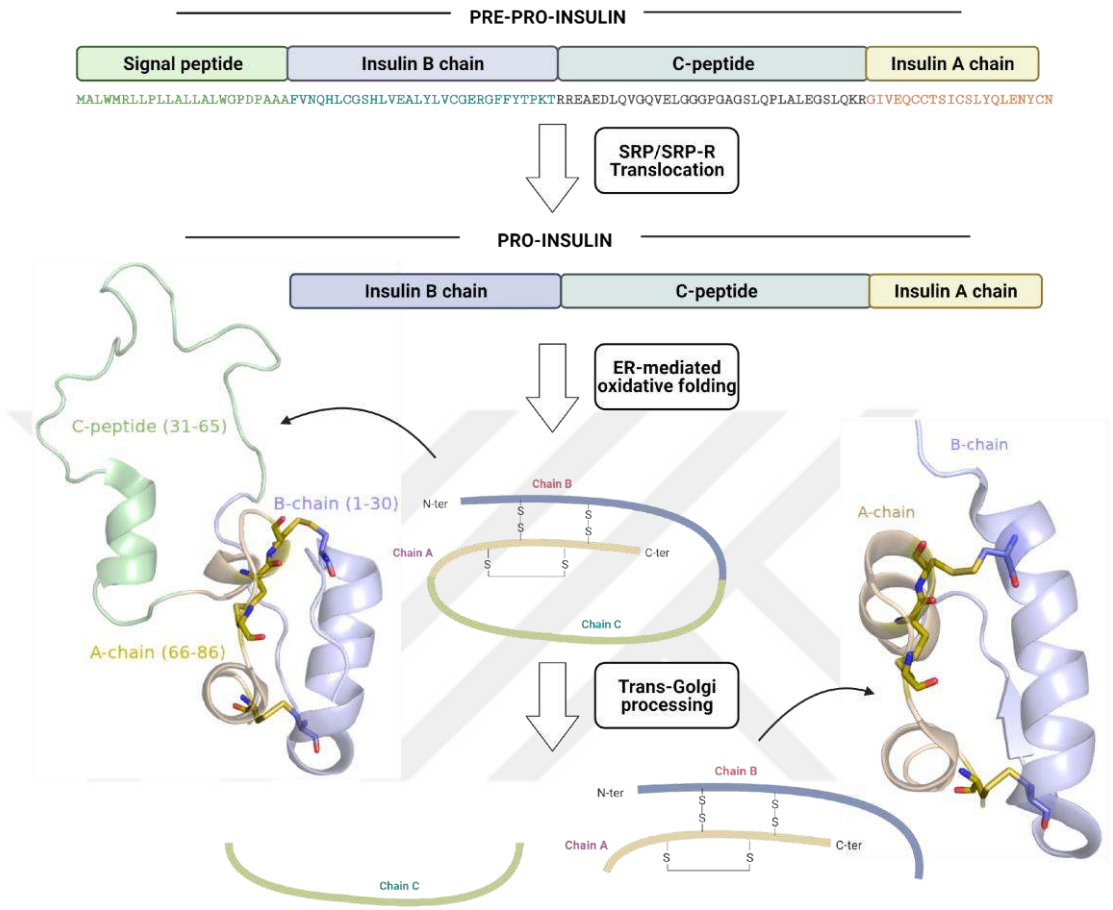


Figure 1.3 Representation of the process from pre-proinsulin to mature insulin on a structural and sequence basis.

Some studies suggest that although crystallization enhances the ability of soluble proinsulin to mature into insoluble insulin, noncrystallization of insulin does not adversely affect proinsulin processing [Frederickson et al., 2005; Carroll et al., 1988; Emdin et al., 1980; Smith, 1966]. After expression, insulin is secreted from the pancreas to the liver through the portal circulation. At the first pass, 50% of insulin is cleared from circulation in the liver hepatocytes to regulate the homeostatic insulin level. In contrast, the rest of the insulin is exported from the liver through the hepatic vein. Then, insulin is distributed to the body through the arterial circulation and promotes vasodilation, a mechanism that causes blood vessels to dilate. Insulin also returns to the liver, where it exerts the metabolic effect, and further clearance occurs with a second

pass [Tiran et al., 1979; Meier et al., 2005] (Figure 1.4).

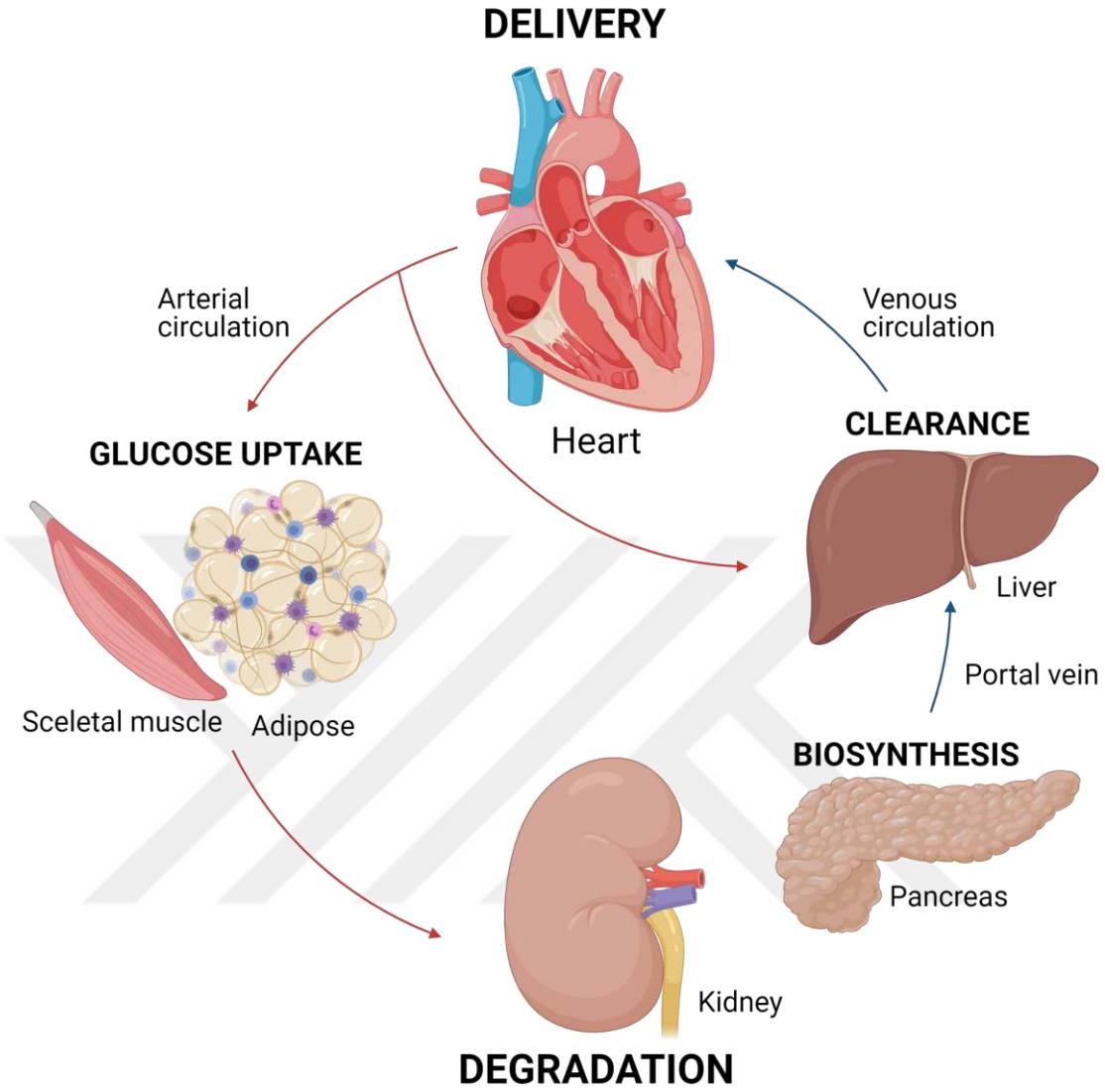


Figure 1.4 Representation of circulation, clearance and degradation of insulin. (Inspired by Tokarz et al. (2018)).

During circulation, hexamer insulin disperses into an active monomer to enter cells [Li, 2002]. After the monomer binds to the insulin Receptor Tyrosine Kinase (RTK), it triggers the autophosphorylation of tyrosine residues (Figure 1.5).

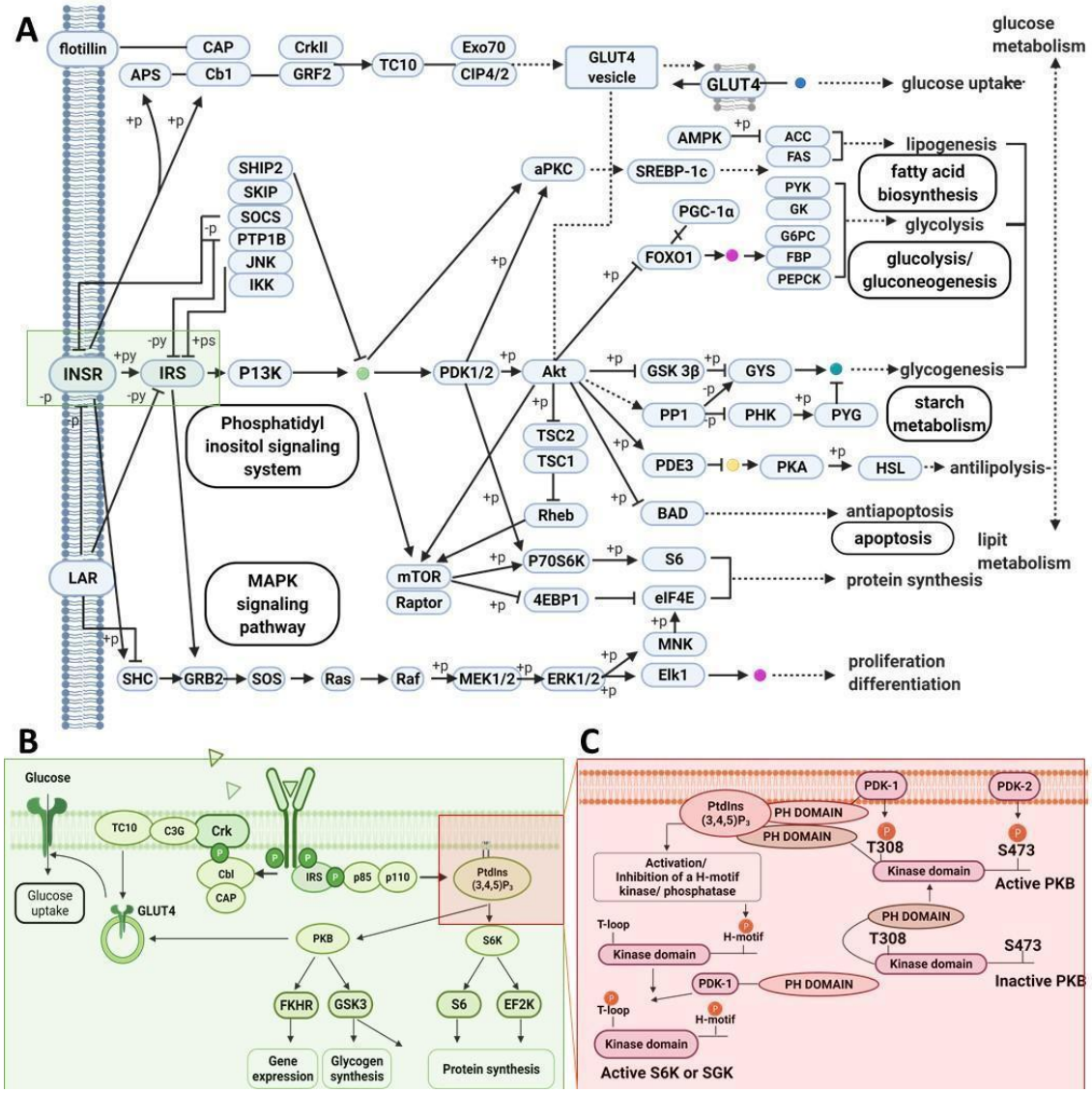


Figure 1.5 Representation of insulin-mediated signaling pathways

(A) Overall demonstration of insulin signalling pathway. The interaction of insulin and insulin receptor (IR) triggers phosphorylation of insulin receptor substrates (IRS) through insulin receptor tyrosine kinase (INSR), resulting in the interaction of the regulatory subunit of phosphoinositide 3-kinase (PI3K) and IRS. PI3K activation triggers 3- phosphoinositide-dependent protein kinase 1 (PDK1) and Akt (a serine kinase); this triggers glycogen synthase kinase 3 (GSK-3) inactivation, leading to glycogen synthesis through glycogen synthase (GYS) activation. Akt activation also triggers translocation of GLUT4 vesicles to the plasma membrane, thereby triggering glucose uptake into the cell. Akt activation also promotes mTOR-mediated protein synthesis through p70S6K and eIF4. GLUT4 translocation can also occur through CAP/Cbl/TC10 signaling by phosphorylating via Cbl INSR. Thus, other signal transduction pathways, especially GRB2, which is part of the RAS, SOS, MEK, and RAF cascade, interact with the IRS. SHC, which is another substrate of INSR, associates with GRB2 through tyrosine phosphorylation, thus enabling the RAS/MAPK pathway to be activated independently of IRS-1. Green sphere: Fosfatidilinositol 3,4,5-Trifosfat (PIP3); blue sphere: glucose; magenta sphere: DNA; cyan sphere: glycogen; yellow sphere: cAMP. Arrows, →: molecular interaction/relation; : unknown reaction/indirect

connection; ⊣: molecular inhibition; ⊀: missing interaction. +p: phosphorylation; -p: dephosphorylation. (B) A close look at the relationship between insulin and major metabolic responses in cells. Representation of glucose uptake through GLUT4 translocation after a series of signal transduction. (C) PtdIns(3,4,5)P3 mediates activation of PKB, S6K, and SGK through their T-cycles, and is upregulated by insulin and other agonists in contrast to PDK1 indirect activity.

Some of these residues are recognized by the phosphotyrosine binding (PTB) domain of Insulin Receptor Substrates (IRS), phosphorylating many tyrosine residues. The other tyrosine residues are recognized by the Src Homology 2 (SH2) domain, which belongs to the Phosphoinositide 3-Kinase (PI3K) [Lizcano and Alessi, 2002; Boura-Halfon and Zick, 2009]. Accordingly, the catalytic subunit p110 of PI3K is taken into the plasma membrane to phosphorylate the inositol ring-D3 position of phosphatidylinositol-(4,5)-bisphosphate (PIP2). This triggers the formation and increases the concentration of phosphatidylinositol (3,4,5)-triphosphate (PIP3) (Figure 1.5b). Studies have shown that PIP3 is the second messenger, and inactivation of PIP3 causes inhibition of critical cellular responses associated with insulin, including glucose transport and stimulation of insulin synthesis. Additionally, a close interaction of PIP3 with Protein Kinase B (Akt/PKB) in insulin signaling has been demonstrated by studies [Shisheva, 2008; Mackenzie and Elliott, 2014]. Accordingly, PKB interacts with PIP3 through the Pleckstrin Homology (PH) domain located at its amino terminus (Figure 1.5c). Thus, PKB has recruited the plasma membrane where PIP3 is localized. Here, PKB and PIP3 interaction does not directly affect PKB but phosphorylate the Ser473 and Thr308 residues, resulting in PKB being active [Feng et al., 2004]. Thus, PIP3 regulates PKB, Serine/threonine protein kinase (SGK), and p70 ribosomal protein S6 Kinase (S6K) activation. Therefore, PKB becomes a PhosphoinositideDependent Kinase-1 (PDK1) substrate. Accordingly, it is important that SGK and S6K also have PH domains for closely related interaction with PIP3 [Biondi et al., 2004] (Figure 1.5b-c). Additionally, the binding of insulin to the receptor triggers the phosphorylation of proto-oncogene Cbl (Casitas Blineage Lymphoma) in the c-Cbl-Associated Protein (CAP)- Cbl complex, resulting in the participation of the CAP-Cbl complex in the lipid raft (Figure 1.5b). Thus, Cbl interacts with the Crk adaptor protein, which is closely related to the Rhofamily guanine nucleotide exchange factor (C3G). This causes the activation of TC10, which belongs to the GTPbinding protein family, to stimulate Glucose transporter type 4 (GLUT4) translocation [Chiang et al., 2001]. Finally, insulin interacts with INSR in a monomeric form. It is a multifunctional and versatile hormone;

it regulates glycogen synthesis by triggering glucose uptake [Cohen et al., 1978; Shen et al., 1970] and stimulating housekeeping mechanisms such as lipogenesis [Gathercole et al., 2011], glycolysis [Beitner and Kalant,1971], apoptosis [Keku et al., 2005], and protein synthesis [Proud, 2006; Saltiel and Kahn, 2001] (Figure 1.5). Furthermore, it plays a vital role in many intra- and intercellular endocrine system reactions by regulating various homeostatic signaling pathways [Goldstein et al., 1998; Khan and Pessin, 2002; Di Camillo et al., 2016].

1.3. Pathology

Due to the function of insulin and its interaction with essential proteins, various insulin-related diseases have been observed and classified. Currently, 1099 diseases associated with the Insulin (INS) gene are available on the Open Targets Platform [OTP, 2022]. Although generally characterized as "pancreatic diseases," insulin-related disorders span a large spectrum and range from "genetic disorders" to "infectious diseases" (Figure 1.6).

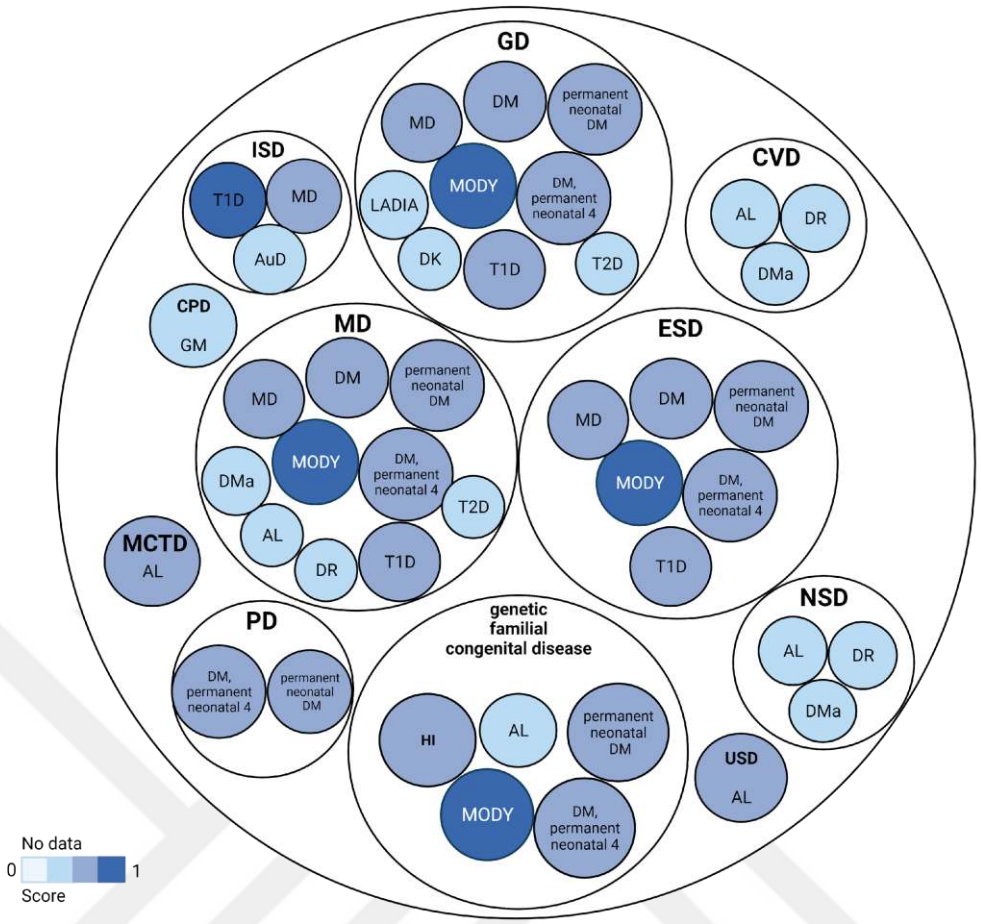


Figure 1.6 Bubbles diagram based on association score (As) representing main insulin-related diseases/ phenotypes related to INS gene

GD: Gastrointestinal Disease, MD: Monogenic Diabetes (As: 0.528), DM: Diabetes Mellitus (As: 0.406), Permanent Neonatal DM (As: 0.506), DM, permanent neonatal 4 (As: 0.715), MODY: Maturity Onset Diabetes of the Young (As: 0.722), LADIA: Latent Autoimmune Diabetes In Adults (As: 0.265), DK: Diabetic Ketoacidosis (As: 343), T1D: Type-I Diabetes (0.684), T2D: Type-II Diabetes (0.324); MD: Metabolic Disease, DMa: Diabetic Maculopathy (As: 0.376), AL: AL amyloidosis (As: 0.462), DR: Diabetic Retinopathy (As: 0.400); ESD: Endocrine System Disease; CVD: Cardiovascular Disease; NSD: Nervous System Disease; USD: Urinary System Disease; Genetic familial congenital disease; HI: Hyperinsulinemia (As: 0.665), PD: Pregnancy or Perinatal Disease; MCTD: Mixed Connective tissue Disease; CPD: Cell Proliferation Disorder; GM: Glioblastoma Multiforme (As: 0.290); ISD: Immune System Disease, AuD: Autoimmune Disease (As: 0.291).

1.3.1. Insulin and pancreatic disorders

Any disease affecting the pancreas, including pancreatic insufficiency, pancreatitis, neoplastic carcinoma, neuroendocrine neoplasms, cystadenomas, and lymphomas, is referred to as a pancreatic disease, and diabetes is the most prominent category among pancreatic diseases [OTP, 2022]. DM is an endocrine pancreatic

disease and is classified into a wide variety of diseases from "rare genetic diabetes mellitus" to "glucose metabolism disease". However, it is characterized by two most common forms: T1D and T2D (Figure 1.7).

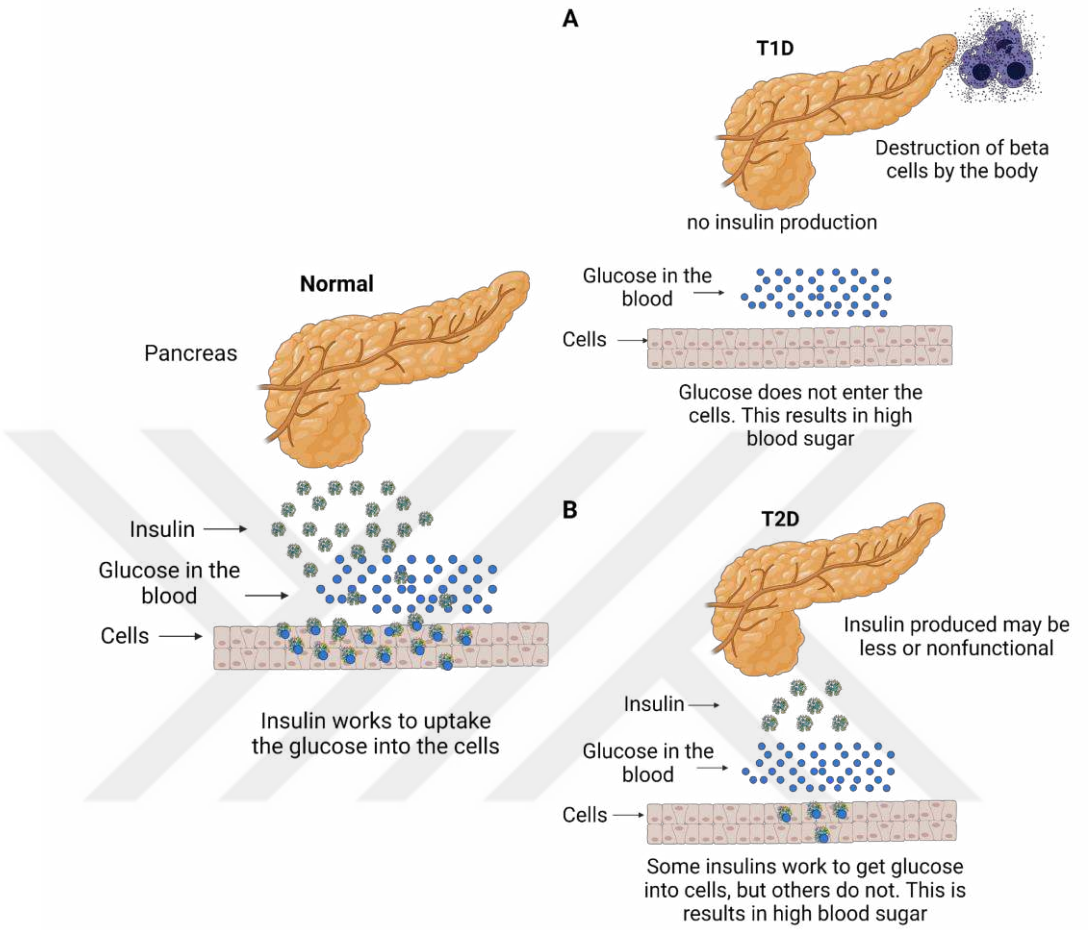


Figure 1.7 Demonstration of T1D (a) and T2D pathophysiology compared to normal physiology (b).

T1D is an autoimmune disease created by the destruction of insulin-producing β-cells, in which insulin cannot be produced, thus causing insulin deficiency (Figure 1.7a). Typically characterized by childhood and adolescence, T1D leads to severe hyperglycemia and diabetic ketoacidosis and may cause death without insulin therapy. Thousands of studies for the treatment of T1D have occurred in the literature, and research is still in progress as well as recent studies in T1D treatment are promising. Studies conducted in 2020 discovered that slowly progressive T1D correlates with islet antibody loss and a different β-cell phenotype [Hanna et al., 202]. In another study, it was discovered that autoreactive T-cells infiltrate the islets of Langerhans by destroying insulin-secreting βcells and early symptoms can characterize it, thus being preferred as

an early-stage marker in clinical studies [Ramirez et al., 2020]. T2D is a progressive metabolic disease that accounts for greater than 90% of the global diabetes epidemic (Figure 1.7b). It is not insulin-dependent besides being an insulin-sensitive disease. It is characterized by insulin resistance and hyperinsulinemia and is closely related to glucose intolerance and hyperglycemia. The overwhelming majority of studies on T2D are in progress, both clinically and exploratory approaches. Potential mechanisms of Agriophyllum oligosaccharide plants in T2D were investigated in an *in vivo* study conducted in 2020. Accordingly, the "ancient medicine" herb activates insulin by triggering the INS-R/IRS2/PI3K/AKT/ Glut4/PPAR-signaling pathway in T2D patients, proliferating hepatocytes, and decreasing blood glucose levels [Bao et al., 2020]. In a 2020 study, the remarkable effect of sleeve gastrectomy was investigated with a negligible risk of complications and a simple operational system in the control of glycometabolism. Accordingly, the downregulation of the SNHG5 gene and the expressional change of the TGR5 gene have alleviated the colorectal damage caused by high glucose. Furthermore, the mechanism under the positive effect of sleeve gastrectomy in T2D was indicated [Wei et al., 2020]. Eventually, insulin plays a vital role not only in diabetic diseases but also in a wide variety of pancreatic diseases such as hyperinsulinism [De Franco et al., 2020], pancreatic neuroendocrine tumor [Alshaikh et al., 2018], and malignant pancreatic neoplasm [Zheng et al., 2020].

1.3.2. Insulin and cardiovascular disorders

Any diseases related to the cardiovascular system are classified under cardiovascular diseases (CVD) [OTP, 2022; Fontbonne and Eschwege, 1991; Nabel, 2003]. In general, pathological conditions associated with CVD begin to appear at an exceedingly early age [Steinberger et al., 1995]. Obesity is associated with insulin resistance and is directly involved in the origin of CVD (Figure 1.8) [Steinberger et al., 1995; Ferreira et al., 2007; Ormazabal et al., 2018]. Insulin triggers the use of metabolic substrates in the heart as well as in other tissues. In cardiomyocytes, although insulin provides the intake of glucose and fatty acids, it prevents the use of fatty acids as an energy source [Reaven, 2012]. This causes insulin resistance, and the pancreas tries to increase insulin secretion, which leads to a phenomenon called hyperinsulinemia [Reaven, 2012]. Furthermore, glucose concentration, insulin resistance, and hyperinsulinemia contribute to the development of atherosclerosis and trigger

mechanisms such as inflammation, dyslipidemia, and hypertension [Giacco and Brownlee, 2010; Bornfeldt and Tabas, 2011]. Therefore, studies prove that the toxic effects of high glucose concentrations due to insulin resistance are strongly associated with CVD, even in people without diabetes [Gast et al., 2012; Sarwar et al., 2010; Sarwar et al., 2007] (Figure 1.8).

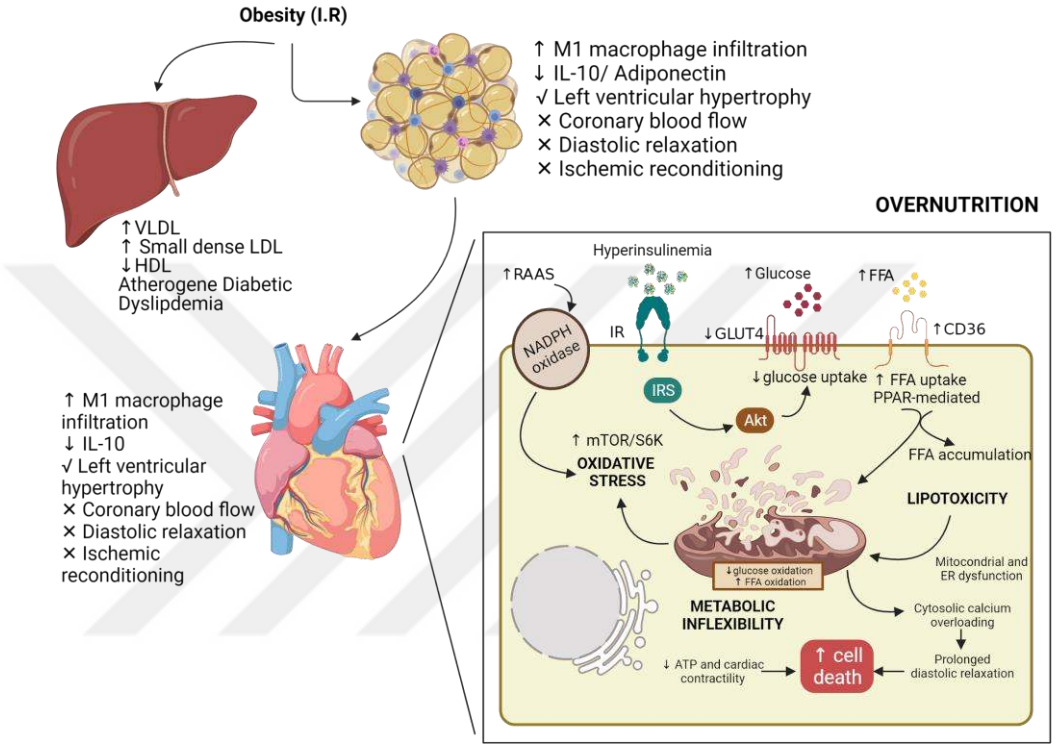


Figure 1.8 Demonstration of CVD-insulin-related pathophysiological and molecular mechanisms
In the case of obesity-related insulin resistance, metabolic dysregulation centered on low glucose oxidation and high lipid oxidation leads to abnormal Ca^{2+} utilization, resulting in low ATP production leading to cardiomyocyte death.

In a systematic review conducted in 2020, it was shown with ultrasonographic evidence that carotid disease increases the possibility of diabetic retinopathy (DR), highlighting a direct relationship between carotid disease and DR [Drinkwater et al., 2020]. Variable glucose grades due to impaired insulin levels play a key role in the etiology of CVD. Accordingly, it was demonstrated that serum insulin level plays a remarkable role in CVD pathophysiology, showing a strong relationship with the carotid plaque [Mujaj et al., 2020]. In addition, a study of cats with asymptomatic hypertrophic cardiomyopathy (aHCM) demonstrated that inflammation, insulin, and insulin-like growth factor-1 (IGF-1) had significant effects on aHCM and paved the way for clinical

trials [van Hoek et al., 2020]. Additionally, recent studies have proven that insulin plays a critical role in CVDs, particularly acute myocardial infarction [Chaudhuri et al., 2004], thrombotic disease [Kwon et al., 2020], and heart failure [Larsen et al., 2020, Okhuma et al., 2020]. This suggests that insulin triggers signaling pathways closely related to CVD.

1.3.3. Insulin and immune system disorders

The innate and adaptive immune system is a coordinated and critical system that maintains the organism's homeostasis against internal and external pathogenic agents [OTP, 2022; Medzhitov and Janeway, 1997]. Additionally, if metabolic dysfunction occurs in the organism, immune system components can be activated and cause pathology. Therefore, there is a close relationship between immune dysfunction and metabolic diseases. (Figure 1.9) [Eheim et al., 2014].

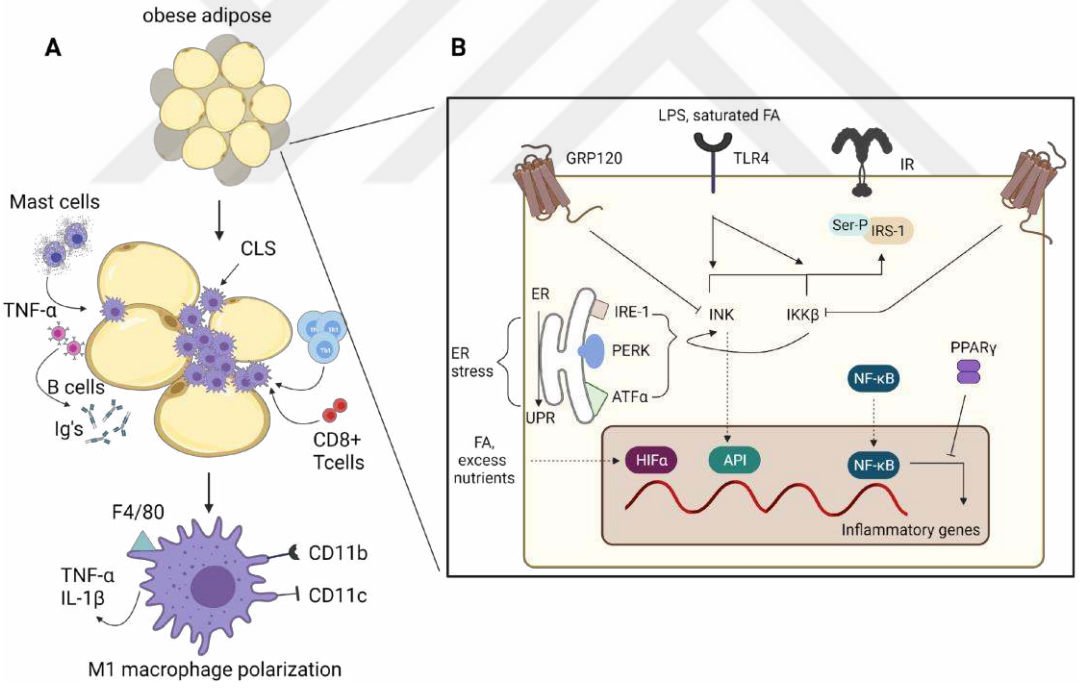


Figure 1.9 Demonstration of immune system-insulin-related pathophysiological and molecular mechanisms.

(A) In obese adipose tissue, the TH1-type cytokines interferon (IFN)-γ trigger a phenomenon called "M1 macrophage position". However, this condition is characterized by T-helper type 2 (TH2) cells and M2 macrophage polarization in normal adipose tissue. Thus, immunoglobulins (Igs), B-cells and mast cells, which are characterized by insulin resistance, increase in obese adipose tissue. CD8(+) T-cells increase in correlation with increased proinflammatory gene expression. In contrast to normal tissue, macrophages in obese adipose tissue are not homogeneously separated and form an accumulation called crown-like

structures (CLS). M1 macrophages are pro-inflammatory that secrete cytokines such as IL-1β and TNF-α and, unlike M2 macrophages, they also increase the expression of CD11c marker. (B) Activator and inhibitory mechanisms that trigger the inflammatory system in the case of insulin resistance. Arrows indicate inflammatory pathways that increase insulin resistance. The weakening of insulin action is related to the phosphorylation of insulin receptor substrate protein-1 (IRS-1) of the activator mechanism. The positive interaction between IKKβ and nuclear factor-κB (NF-κB) enables NF-κB to bind to DNA, thereby activating inflammatory mediators. The positive association between JNK1 and AP1 is also important in the regulation of these mediators. The increase in ovarian nutrition and fatty acid (FA) triggers the Endoplasmic Reticulum (ER) stress and thus the unfolded protein response (UPR), which is characterized by three main pathways (IRE-1, PERK and ATFα). Activation of omega-3 fatty acid receptors (GRP120) and PPARγ, which are promoters of insulin sensitivity, is characterized by the interference of NF-κB and AP1 signaling pathways.

In a study, macrophages play a critical role in triggering chronic inflammation in obesity, which is associated with insulin resistance [Weisberg et al., 2003; Morrison et al., 2015]. With the increase of fatty acids released from the adipose of the obese organism, the M1-polarized macrophages associated with inflammation are translocated into adipose tissue (Figure 1.9a). M1-macrophages, which lead to the activation of c-Jun N-terminal kinase (JNK) or inhibition of nuclear factor-κB (IκB) kinase (IKK) pathways, cause insulin resistance as they trigger inflammatory stress in the cell. Accordingly, inflammation is suppressed by the activation of peroxisome proliferator-activated receptor (PPAR) signals associated with M2-macrophages. However, when these are inhibited, the inflammation becomes chronic. This molecular regulation is closely related to the function of macrophages, lymphocytes, and mast cells, causing inflammation to increase gradually (Figure 1.9b). Therefore, the increase in fat mass in the tissues is characterized by insulin resistance, thereby triggering inflammation of adipose tissue [Patel et al., 2013; Lumeng et al., 2008; Miyachi et al., 2008]. A recent study showed that intestinal worms causing immunosuppressive and asymptomatic chronic infections prevent streptozotocin-induced diabetes through their trehalose. This shows therapeutic effects on streptozotocininduced diabetes by increasing the Ruminococcus spp and activating cytotoxic (CD8+) Treg cells [Miyachi et al., 2018]. As a result of a series of experiments conducted while considering the relationship of CD8+ T cells to Major Histocompatibility Complex I (MHC-I), it was shown that these cells gained an effector function during the clinical diagnosis period by presenting an antigen-experienced phenotype. Accordingly, evidence has been provided that these cells cause T1D [Yeo et al., 2020]. In addition, it has been proven in the latest studies

that insulin plays a role in various immunoreactive diseases, such as Shwachman-Diamond syndrome [Kamoda et al., 2005], type II hypersensitivity reaction disease [Murray et al., 2017], anemia [Lie et al., 2012], and celiac disease [Cronin and Shanahan, 1997]. All these suggest that insulin triggers signaling pathways closely related to immune system diseases.

1.3.4. Insulin and cancer

Insulin, with its well-known mitogenic effects on cells, is considered a growth factor [Vigneri et al., 2020]. It is known how crucial growth factors are to the development of cancer. By considering these situations, the effects of insulin on cancer cells are apparent. Insulin shows its effect by its interactions with the receptors on the cell surface that are composed of Insulin Receptor (IR) and Insulin-like growth factor receptor (IGFR) [Aaronson, 1991; Cai et al., 2017]. It has been observed many times that IGFR and IR have a key role in cancer development [Malaguarnera and Antonino, 2011]. Moreover, studies have shown that the relationship between higher circulating IGF families is positively correlated with the increased risk for various cancer types containing colorectal, ovarian cancer, prostate, and breast cancers (Figure 1.10) [Weroha and Haluska, 2012].

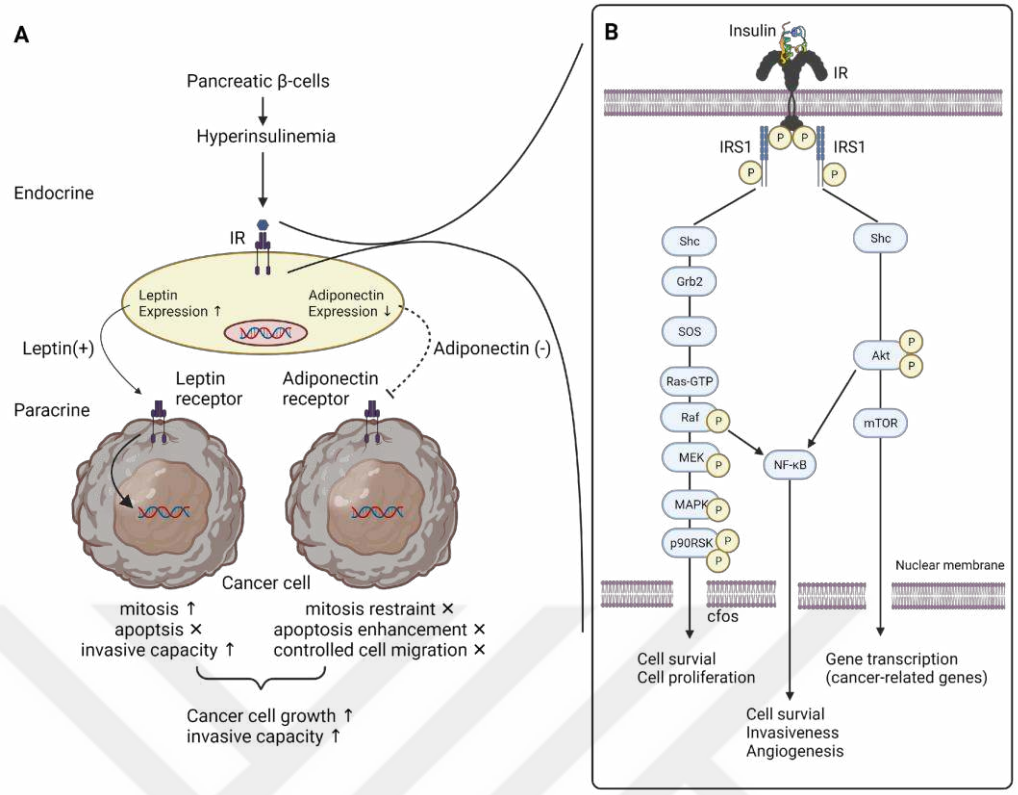


Figure 1.10 Demonstration of cancer-insulin-related pathophysiological and molecular mechanisms (A) Cancer cell growth characterized by paracrine stimulation is triggered by leptin invasion, adiponectin inhibition, and hyperinsulinemia. (B) Positive effect of PI3K/Akt/mTOR and The Ras/MAPK signaling pathway in cancer cell proliferation and anti-apoptosis through NF-κB nuclear translocation

Thus, it can be said that some types of cancer cells are more sensitive to growth by insulin when compared to normal cells. Therefore, understanding insulin and interactions with receptors from a holistic perspective can accelerate cancer treatment studies.

1.3.5. Insulin and nervous system disorders

Any neoplastic or non-neoplastic disorder, including the brain, spinal cord, or peripheral nerves, is classified as a nervous system disease [OTP, 2022]. Although there is no evidence that neurons or the central nervous system (CNS) are insulindependent, this system is insulin-responsive and has a vital role in glucose homeostasis (Figure 1.11) [Belfiore et al., 2009].

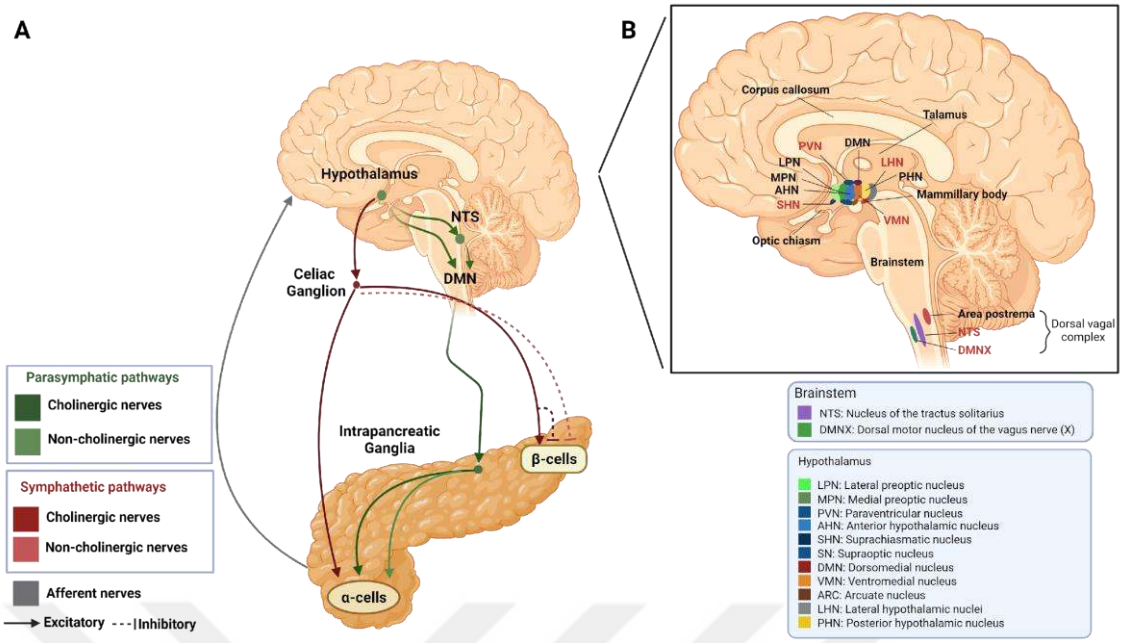


Figure 1.11 Schematic representation of the connection between the brain and the pancreas.
Sagittal representation of a human brain in terms of energy/glucose homeostasis. (A) The relationship of pancreatic α- and β-cells with postganglionic sympathetic and parasympathetic nerves are shown in green and red, respectively. Connections to the brain are shown in gray. (B) Representation of glucose homeostasis in relation to key hypothalamic nuclei.

Insulin and insulin-like growth factor type-1 (IGF-1) function as modulators that have an essential role in the central nervous system's metabolic function and neuronal growth, and their receptors are abundantly expressed in the CNS. Insulin and IGF-1 play vital roles in neurite growth, neuronal migration, and synapsis formation due to synthesizing related proteins and neuronal development, survival, and differentiation [Ye et al., 1996; Park, 2001; Kim and Feldman, 2012; O'Kusky et al., 2000]. Although information about the presence of insulin in the brain and its immunoreactive effects is insufficient, studies have shown that insulin can transport across the blood-brain barrier into the cerebrospinal fluid [Poduslo et al., 2001]. Accordingly, studies have proven that the brain can use both local and peripheral insulin for various functional requirements, including cognitive functions [De La Monte et al., 2005; De La Monte, 2009]. Therefore, possible dysfunctional disorders that may occur in the molecular mechanism of insulin will directly affect CNS neurons [Batista et al., 2018]. Numerous studies related to diabetes and nervous system diseases, especially Alzheimer's Disease (AD) [Batista et al., 2018], support this hypothesis (Figure 1.12).

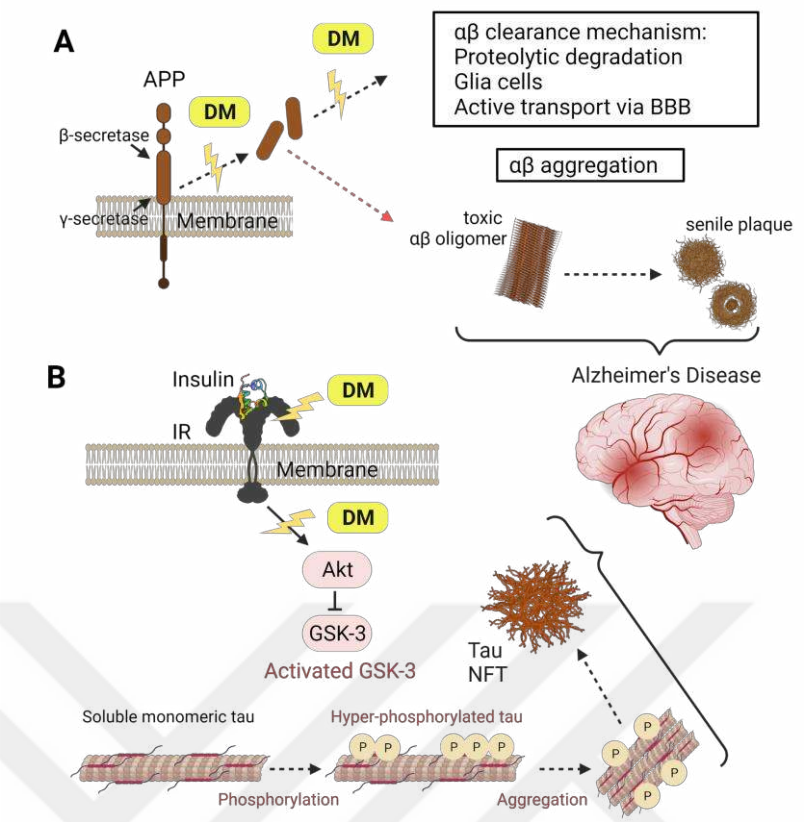


Figure 1.12 Demonstration of the effect of diabetes on Alzheimer's amyloid and tau pathology

(A) Insulin resistance caused by DM includes both senile plaque formation and amyloid pathology of the brain and (B) it triggers the activation of GSK-3 leading to hyperphosphorylated tau aggregation characterized by NFTs. DM diabetes mellitus, BBB blood-brain barrier, APP amyloid precursor protein, P phosphorylation, NFT neurofibrillary tangle, GSK-3 glycogen synthase kinase-3.

Accordingly, it is shown that the strong relationship between neuronal insulin resistance and Aβoligomers, driven by IRS inhibition and TNF-α activation, results in impaired synaptic plasticity, synaptic dysfunction, and synapse loss [Ferreira et al., 2018; De Felice et al., 2009; Townsend et al., 2007; Bomfim et al., 2012; Sebastião et al., 2014; Lourenco et al., 2013; Tai et al., 2018]. Collectively, recent studies have shown that insulin is closely related to nervous systemrelated diseases, including nicotine dependence [Faheem et al., 2020], stroke [Cerecedo-Lopez et al., 2020; Ryan et al., 2020], unipolar depression [Koslow et al., 1982], acromegaly [Corica et al., 2020; Guerreiro et al., 2020], epilepsy [Peng et al., 2020], social anxiety disorder [Ellis and Arsiwalla, 2020; Rodríguez-Rabassa et al., 2020], amyotrophic lateral sclerosis [McDonald et al., 2021], peripheral nerve injury [Toth et al., 2006], and proximal spinal muscular atrophy [Brener et al., 2020].

1.3.6. Insulin and genetic disorders

Diseases caused by ancestral genetic variations before or rarely at birth are classified as genetic, familial, or congenital diseases [OTP, 2022]. The close relationship between insulin and key signaling pathways represents an exceptionally large signal network from a metabolic perspective. Besides, recent studies have proven that insulin has been related to a wide variety of diseases such as familial hyperinsulinism (Figure 1.13), pancreatic neuroendocrine tumor [Alshaikh et al., 2018], congenital muscular dystrophy [Zambon and Muntoni, 2021], multiple endocrine neoplasia type 1 (ectopic insulinoma) [Lindhurst et al., 2012], MEHMO Syndrome [Mori et al., 2021], Complete Androgen Insensitivity Syndrome [Auer et al., 2022], cryptorchidism [Baker et al., 2002], 46, XX gonadal dysgenesis [Kanaka-Gantenbein et al., 2004; Rasio et al., 1976].

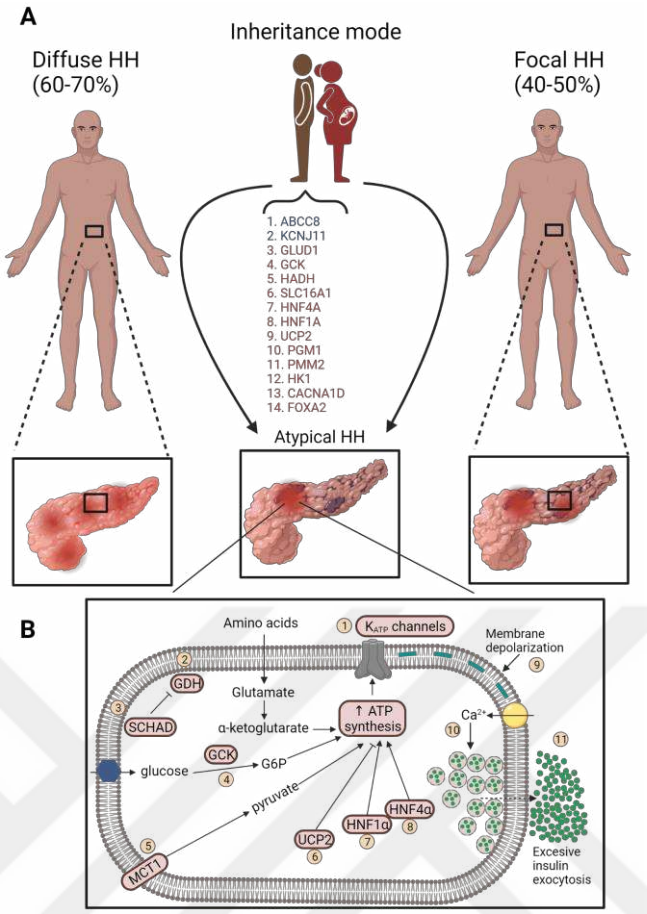


Figure 1.13 Schematic presentation of the genetic mechanisms of Congenital hyperinsulinism

(A) Diffuse subtype HH is characterized by hyperplasia and hyperchromatic β-cell enlargement. The Focal subtype HH is characterized by certain large and small focal lesions. Inheritance mode characterized by Atypical HH is known to be associated with both biallelic recessive and dominant mutations. (B) Genetic defects of CHI-associated genes play a critical role in dysregulated insulin secretion. The main of these defects involve pancreatic KATP channels (ABCC8 and KCNJ11). In addition, enzymatic defects (GLUD1, UCP2, HK1, SCHAD, GCK, etc.) and defects in transcription factors (HNF4A, HNF1A, etc.) are also associated with CHI. β-cell insulin secretion is dependent on a wide variety of metabolic signals, from genes to proteins, and defects in these pathways result in dysregulated membrane depolarization and calcium release caused by channels, resulting in HH. ATP, adenosine triphosphate; Ca^{2+}, calcium; G6P, glucose 6-phosphate; GDH, glutamate dehydrogenase; HNF1α, hepatocyte nuclear factor 1α; HNF4α, hepatocyte nuclear factor 4α; short-chain 3-hydroxyacyl CoA dehydrogenase (SCHAD), UCP2, mitochondrial uncoupling protein 2; monocarboxylate transporter subtype 1 (MCT-1); Uncoupling protein 2 (UCP2); HH, hyperinsulinaemic hypoglycaemia.

1.3.7. Insulin and Musculoskeletal Disorders

Diseases involving various musculoskeletal disorders, including connective tissue, have been represented as musculoskeletal or connective tissue diseases [OTP, 2022]. Studies have proven that insulin dysfunction/resistance is associated with muscle metabolic defects (Figure 1.14) [DeFronzo and Tripathy, 2009; Garneau and Aguer, 2019], and various musculoskeletal diseases such as panniculitis [Jabłońska et al., 2018; Avery, 1929], hypertrophic cardiomyopathy [Shigematsu et al., 2011], osteoporosis [Cannarella et al., 2019; Xia et al., 2012], Sanfilippo syndrome type B [Prill et al., 2019], pseudohypoparathyroidism [Muniyappa et al., 2013], and cholangiocarcinoma [Liu et al., 2019]. Collectively, insulin is closely related to various diseases including urinary system diseases [Liu et al., 2020; Devis-Jauregui et al., 2020; 149. Pugliese et al., 2020; Kukla et al., 2020], cell proliferation disorders [Ratnakumar et al., 2020; Gao et al., 2020; Zaman et al., 2020; Meerson et al., 2020], gastrointestinal diseases [Kern et al., 1999], breast diseases [Kaaks et al., 1996; Pellegrino et al., 2020; Rose et al., 2004; Bruning et al., 1992; Papa et al., 1990], hematologic system diseases [Barbieri et al., 2001], infectious diseases [Niklasson et al., 1998; Maeno et al., 2003], and psychiatric disorders [Bisschop et al., 2011; Amgarth-Duff et al., 2020].

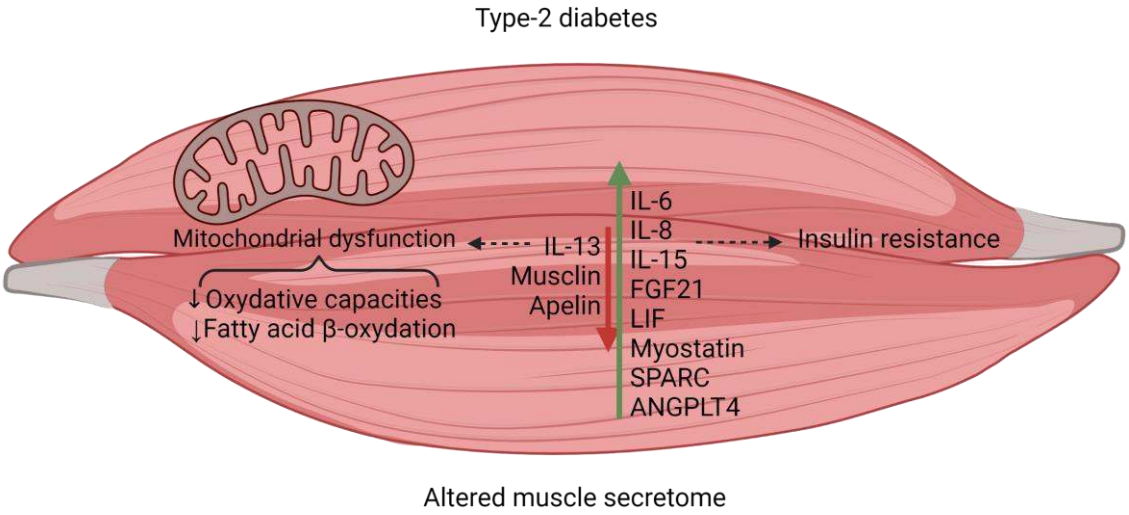


Figure 1.14 Demonstration of T2D-related muscle metabolic defects

The effect of myokines on the mechanism of the altered muscle secretome. Muscle metabolic defects and altered muscle secretome are strongly associated with T2D diabetes. Secretion of myokines, RNAs, mtDNA, and metabolites, and sedentary behaviors also play a role in altering this mechanism. An altered secretome is also characterized by skeletal muscle mitochondrial function and triggers the development of insulin resistance. IL-6, Interleukin 6; IL-8, Interleukin 8; IL-15, Interleukin 15; FGF21, Fibroblast growth factor 21; LIF, Leukemia inhibitory factor; SPARC, Secreted Protein Acidic And Cysteine Rich; ANGPLT4, Angiopoietin-like 4.

A wide variety of pathogenic processes are involved in the development of DM directly associated with insulin. Many different anomalies, from the destruction of pancreatic β-cells to insulin resistance, constitute critical points in the development of diabetes. The dysfunction caused by insulin insufficiency in target tissues forms the basis of abnormalities in protein, fat, and carbohydrate metabolism in diabetes. In DM, which correlates closely with hyperglycemia, hyperglycemia's prominent symptoms are characterized by weight loss, polyphagia, polydipsia, polyuria, and blurred vision. Additionally, susceptibility to certain infections and growth disturbances is also part of chronic hyperglycemia. Ultimately, uncontrolled diabetes will lead to life-threatening nonketotic hyperosmolar syndrome or hyperglycemia with ketoacidosis [Association, 2010]. Therefore, early diagnosis and screening are vital for people at risk and those with early-onset disease. In this context, the diagnostic limit for diabetes is mainly the fasting plasma glucose level and the glucose-containing hemoglobin level. These levels allow the patient to be diagnosed with diabetes or prediabetes [Inzucchi, 2012].

1.4. Chronological milestones guiding the historical narrative of insulin

The discovery of insulin in 1921 has caused an increase in what we know about insulin cumulatively each passing day (Figure 1.1) [Banting et al., 2007]. Not only is insulin knowledge widened, but also the diseases that are caused by the dysfunction of this versatile hormone are examined. As a result, therapeutic studies and biotechnological applications were conducted [Haeusler et al., 2018]. The studies on insulin itself, its function, its receptor on cells, and some pathways it triggered were performed between 1950 and 1985 [Haeusler et al., 2018]. Then, studies were performed on insulin's intact expression sequences and its receptor with cloning for the first time between 1985 and 1991 [Haeusler et al., 2018]. As insulin's function and structure began to be enlightened, its relationship with the disease gradually emerged. A closely related relationship between diabetes and insulin was revealed in a study conducted on dogs in 1921 [Haeusler et al., 2018]. The study found that insulin controls diabetes. Since then, therapeutic approaches to diabetes have always been based on the treatment of insulin or insulin analogs [Haeusler et al., 2018]. In the direction of finding insulin analogs, scientists have started to play with insulin since 1936 [Himsworth, 1936]. For this purpose, protamine was added to insulin for the first time in 1936, allowing insulin to remain in circulation for a long time [Hagedorn, 1936]. It was released by Novo Nordisk under the name "Neutral Protamine Hagedorn" (NPH) [Hagedorn, 1936]. In addition, around the 1950s, intermediate-acting Lente insulin was produced, which is now obsolete [Siddiqui and Scupola]. The discovery of the insulin receptor after 1971 was followed by the introduction of the use of U100 insulin [Haeusler et al., 2018]. In the years when U40 and U80 were used routinely, the U100 went down in history as a great convenience [Haeusler et al., 2018]. After the invention of the wearable insulin pump in 1976, recombinant human insulin production was realized with the first transportable insulin pump in 1978, and it was approved in 1982 [Haeusler et al., 2018]. Two years after the metformin approval, which was in 1996, fast-acting insulin lispro was introduced by Eli Lilly and his company for the first time [Haeusler et al., 2018]. Then, stimulating insulin secretion in the presence of glucose was another ground-breaking event in which a repaglinide drug was developed by Novo Nordisk [Haeusler et al., 2018]. Two years later, Sanofi Aventis in the US introduced long-acting insulin Lantus (glargine) as a basal insulin analog [Haeusler et al., 2018]. As the insulin-based drugs and their derivatives were released, an examination was needed

for the usage of other drugs with insulin. These examinations were performed in detail between 2005 and 2008 [Haeusler et al., 2018]. First, Exubera [Odegard and Capoccia, 2005] inhaled insulin, which was withdrawn in 2007 [Bailey and Barnett, 2007], then long-acting Degludec insulin was produced in 2013 [Monami and Manucci, 2013]. Today, various insulin modifications are still being tested and many patients are still active in this area. Although there are countless milestone studies on insulin, the most dynamic and disruptive field today is undoubtedly omics and structural-based protein discovery and analog inventions.

1.5. The multi-omics approach to insulin

Since the discovery of insulin, the most critical issue associated with many diseases, especially diabetes, is 'insulin resistance.' Understanding the molecular differences and the underlying mechanism of insulin resistance is important to take right strategic steps regarding the disease process, inhibit the progression of pathogenesis, and plan good therapeutics accordingly. In this context, the emergence of new "omics" approaches will provide invaluable opportunities to realize the prospective strategies mentioned in this phenomenon [De Jesus and Kulkarni, 2019].

1.5.1. Insulin and genomics & epigenomics

As known, the underlying mechanism of many insulinrelated diseases, especially T2D, correlates strongly with the interaction between genetics and the environment. In this context, some new insulin-related genes, including Transcription factor 7-like 2 (TCFL2) [Grant et al., 2006] and Peroxisome Proliferator Activated Receptor Gamma (PPARG) [Altshuler et al., 2000], were confirmed and identified with genomics-based studies, analyses, and prospective approaches conducted in the early 2000s. Of these, TCFL2 was determined to have a critical cell-autonomous role in controlling β-cell mass and function [Mitchell et al., 2015]. These studies were further advanced by nextgeneration sequencing and complex genomics analyses; following which a dozen new insulin-related genes were discovered, including the Solute Carrier Family 30 Member 8 (SLC30A8), CDK5 Regulatory Subunit Associated Protein 1 Like 1 (CDKAL1), and Insulin Like Growth Factor 2 MRNA Binding Protein 2 (IGF2BP2) genes [Scott et al., 2007; Consortium et al., 2007; saxena et al., 2007]. While Zinc

transporter 8 (ZnT8), a product of SLC30A8, provides the regulation of the pathways involved in the regulation of granulated Zinc intake required for insulin secretion; SLC30A8 haploinsufficiency has been proven by GWAS studies, combined with a meta-analysis, to lead to insulinrelated diseases [Davidson et al., 2014; Syring et al., 2016]. For example, Ayers et al., in their first genomics-based study on cryptorchidism, found a direct relationship between insulin-like peptide 3 (INSL3) and recessive variants of its receptor, relaxin family peptide receptor 2 (RXFP2), with this disease and showed that it was associated with testicular descent in humans [Ayers et al., 2019]. In the last years, studies have proven that the GWAS approach has been widely adopted in studies that analyze diseases, including their relationships with insulin, such as Simpson-GolabiBehmel syndrome [Cannon et al., 2019], cardiometabolic disease [Tekola-Ayele et al., 2019], and cryptorchidism [Harrison et al., 2019]. These have given large-scale genomics results. Additionally, combining GWAS studies with expression quantitative locus (eQTL) [Zhu et al., 2016] and DNA methylation [Wu et al.,2018; Xue et al., 2018] provides a more integrative analysis approach to clearly determine the causal effects of genetic variants of these diseases from a molecular perspective. More than 150 T2D variants have been identified in large-scale GWAS studies. This suggests the need for genomic-based approaches to understanding insulin-related pathways [Jaspers et al., 2001]. Epigenetics, which reveals the importance of the environment in the phenotype, refers to the expressional changes in the gene without changes in the nucleotide sequence. Doubtless, epigenetic changes are inherited phenomena that can dynamically modulate the genome and reshape the phenotype. Several epigenetic mechanisms, including histone and chromatin modification [Bhandare et al., 2010; Khetan et al., 2018; Rauell-Vila et al., 2018] DNA methylation modifications [Gao et al., 2014, Yang et al., 2011], and non-coding RNAs [La Pierre and Stoffel, 2017; Frost and Olson, 2011; Motterle et al., 2017; Singer et al., 2018], show critical effects on the islets and β-cell adaptation [Poy et al., 2009; De Jesus and Kulkarni, 2014]. Epigenomics studies conducted in this direction have shown that DNA methylation has great importance in T2D pathophysiology. According to these studies, the promoter of PPARG Coactivator 1 Alpha (PPARGC1A) [Ling et al., 2008], pancreatic and duodenal homeobox 1 (PDX1) [Sachdeva et al., 2009; Stoffers et al., 1997], and similar 254 genes, which play a role in the regulation of insulin released from human islets, is hypermethylated in T2D compared to healthy subjects [Volkmar et al., 2012]. This result is inversely correlated with expression results and β-cell

survival, indicating that insulin mRNA is downregulated in T2D patients as a result of the methylation of these promoters [Yan et al., 2011]. In a further study using the whole-genome bisulfite sequencing (WGBS) approach, they identified 25,820 methylation sites in the genome, 80% of which are CpGs, including genes important for the functions of β-cells, such as PDX1 [Volkov et al., 2017]. In epigenomics studies, blood-based DNA methylation biomarkers can become among the commonly used methods to evaluate β-cell survival [Willmer et al., 2018].

1.5.2. Insulin and transcriptomics & epitranscriptomics

Transcriptomics is another field of omics that can provide invaluable information from a molecular perspective that includes insulin-related details. Especially oligonucleotide microarray tools used in islets studies performed on patients and healthy subjects. They are among the most useful tools primarily preferred in transcriptomics studies. mRNA levels of many insulin-associated genes, such as Insulin Receptor (INSR), Insulin Receptor Substrate 2 (IRS2), AKT Serine/Threonine Kinase 2 (AKT2), and Hepatocyte Nuclear Factor 4 Alpha (HNF4a) [Gunton et al., 2005], and GWAS identified genes, such as Insulin Like Growth Factor Binding Protein 2 (IGFBP2) [Marselli et al., 2010], were downregulated in patients compared to controls that had been proven by transcriptomic based studies. Our knowledge about the expressional regulation of diseases has increased exponentially with transcriptome-based studies. In a study conducted on the subject, 83 islets were analyzed using genomics approaches, including tour de force transcriptomic analysis [Fadista et al., 2014]. This study is the most comprehensive study performed in human islets, revealing many variants associated with β-cells, mainly 5'- Nucleotidase Ecto (NT5E), p21 protein (Cdc42/Rac)-activated kinase 7 (PAK7), Transmembrane emp24 domain-containing protein (TMED), and Tetraspanin-33 (TSPAN33) [Bader et al., 2016]. Although this technique has limitations [Dorrell et al., 2016], this handicap has been overcome with fluorescence-activated cell sorting (FACS) strategies [Williams et al., 2019]. In another study on the subject, a systematic analysis of prospective islet progenitors with epigenetic and transcriptomic profiling techniques was presented, and clinical replacement therapies were criticized in this direction [Xin et al., 2016]. Another method is single-cell RNA sequencing (scRNA-seq) [Segerstolpe et al., 2016; Lawlor et al., 2017]. Using this technique, transcriptome-based genetic programs of

each cell type in islets were examined and affected genes, such as FXYD Domain Containing Ion Transport Regulator 2 (FXYD2) and Glycerol-3- Phosphate Dehydrogenase 2 (GPD2), were revealed in sick cells [Arystarkhova et al., 2013]. Accordingly, in an *in vivo* study, it was shown that the ablation of FXYD2 from β-cell caused an increase in β-cells and hyperinsulinemia [Desrosiers and Rottman, 1974; Lavi and Shatkin, 1975]. Transcription, the most crucial step of the central dogma, requires comprehensive regulation to ensure optimal gene expression. Damages that may occur during the genomic and epigenetic editing of mRNA copied from the genomic DNA locus may cause many diseases [Nachtergaele and He, 2018]. Necessary chemical modifications of mRNA (messenger RNA), tRNA (transfer RNA), and rRNAs (ribosomal RNA) in translational steps are required for proper translation efficiency [Nachtergaele and He, 2018; Desrosiers and Rottman, 1974; Lavi and Shatkin, 1975]. One of these, mRNA methylation (m6A), is a very dynamic phenomenon, and this has been demonstrated by the clear identification of α-ketoglutarate-dependent dioxygenase AlkB family of proteins (ALKBH5) [Zheng et al., 2013] and obesity-related proteins (FTO) [Jia et al., 2011]. Interestingly, studies show that obesity is closely related to insulin. In other words, m6A controls mRNA translation, translocation, splicing, and its decay, thereby modulating a variety of cellular functions involving the adaptation of β-cells associated with insulin resistance [Roignant and Soller, 2017].

1.5.3. Insulin and proteomics

All kinds of proteomic studies on human insulin, especially its measurement, are critical for all sorts of insulinrelated conditions such as diabetes, toxicology, hypoglycemia, and sports antidoping [Nedelkov et al., 2018]. Studies in which insulin and insulin analogs are detected and addressed are closely related not only to the treatment of diabetes and hypoglycemia but also to antidoping and toxicology [Marks and Teale, 1996; Brackenridge et al., 2006; Kristensen et al., 2012; Sonksen, 2001; Marks and Wark, 2013]. Undoubtedly, these studies' critical value is their contribution to immunological tests, including the detection of insulin and insulin analogs. However, these tests cannot show the desired success in determining insulin analogs [Parfitt et al., 2015]. Regarding the differences in the unique sequence and mass, the only technique that analogs can be measured with is proteomicsbased approaches, such as mass spectrometry [Blackburn, 2013]. These proteomics-based methods, in which

endogenous insulin and all kinds of analogs can be detected, are basic science results and are preferred for versatile biological applications [Sundsten and Ortsäter, 2009]. Today we know that a strategic experimental move has thousands of proteomic outputs and provides critical information, thanks to proteomics approaches that include the quantitative analysis of the entire proteome. For example, proteomic analysis of islets obtained from only eight mouse strains could yield diabetes-related metabolic profiles of the islets proteome of the genetic background and provide evidence of the vital link between dopamine and insulin secretion [Mitok et al., 2018]. Thanks to lipid droplet analysis with proteomic approaches, which have multi-process steps, more in-depth information was obtained about the storage and processing of lipids in cells, and the presence of proteins belonging to the mitochondrial category was observed at a high rate in cases of diabetes. Therefore, these results indicate the metabolic importance of β-cell mitochondria in insulin secretion [Larsson et al., 2012]. Additionally, proteomic techniques are not limited to discovery and invention-based studies. Still, they are an approach that provides excellent convenience to researchers in clinical studies and accelerates studies in this direction. In this context, a high-throughput proteomics study showed that proteins such as growth hormone receptors, insulin-like growth factor-binding protein 2, and aminoacylase-1 are closely related to diabetes according to the plasma proteome results obtained from diabetics, and new candidate proteins were determined [Elhadad et al., 2020]. In another similar high-throughput proteomics study, high efficiency results were obtained by determining the amount of insulin in plasma while simultaneously detecting all kinds of additional insulin isoforms and isoform sequences [Oran et al., 2011]. Proteomics approaches are also used as the choice of a fast and robust validation of previous expressional studies [Carlsson et al., 2020]. In a discovery-based study with proteomics techniques, a strong association of GLUT4 storage vesicles (GSVs) with low-density lipoprotein receptor-related protein 1 (LRP1) was determined, and secretion of this receptor required for GLUT4 storage vesicle function [Jedrychowski et al., 2010]. In conclusion, many insulin-related mechanisms, especially insulin secretion, signaling pathways, and complex functions of β-cells, have been elucidated by proteomics strategies.

1.5.4. Insulin and metabolomics

Metabolomics is the most common method of omics, as it is used to identify and quantify endogenous smallmolecule metabolites (<1,500Da) [Gibney et al., 2005]. A wide variety of metabolites in urine, blood, and tissues can be detected with nuclear magnetic resonance (NMR), gas or liquid chromatography-mass spectroscopy (GC-MS, LS-MS), and similar approaches, which form the backbone of omics methods [Lu et al., 2013]. Thanks to these techniques, whose sensitivity is characterized by minimal metabolic changes, biomarker detection is routinely performed in clinical [Gu et al., 2020; Bos et al., 2020] and *in vivo* studies [Salihovic et al., 2020; Yan et al., 2020]. A study conducted in this direction was discussed to find useful biomarkers that employ the metabolomic profile obtained from diabetic kidney disease (DCD) and its progression [Li et al., 2019]. Other studies have shown that the changing metabolomic states of obese individuals reflect metabolic disorders and shed light on the disease's underlying mechanism [Zhang et al., 2015; Kim et al., 2013]. In another study related to this, plasma metabolic profiles were examined to further elucidate the mechanisms underlying obesity associated with insulin resistance in children. It was determined that acylcarnitines, aromatic amino acids, and BCAAs could be prospective biomarkers [Lynch and Adams, 2014]. In a study investigating the effect of nutrients on insulin secretion, a detailed metabolomic analysis was performed on a β-cell line using GC-MS to reveal the critical role of a coordinated signaling cascade dependent on glucose metabolism [Zhao et al., 2016, Huang and Joseph, 2012]. In another study on T2D, the metabolomics and gut microbiome integrated mechanism of Naoxintong Capsule (NXT), a formulation of traditional Chinese medicine, was investigated and its effect on diabetic rats was revealed.

1.5.5. Insulin and glycomics

Following the overwhelming presence of approaches such as genomics and proteomics in studies, it has become clear that it is necessary to use various omics techniques, considering posttranslational modifications (PTM), to see the big picture [Parry et al., 2005]. When protein databases are sifted through, it will be easily seen that glycosylation (N-glycosylation) constitutes 70% of the PTMs occurring in proteins [Apweiler et al., 1999; Wormald et al., 2002]. The layer formed by the glycoproteins

surrounding the cell is the first point of contact with the extracellular matrix. It is here that oligosaccharides play a crucial role in providing epitopes to protein receptors, cell-matrix, and cell-cell interactions [Crocker et al., 2002; Esko et al., 2002; Rabinovich et al., 2002]. Considering the critical importance of N-glycans in protein folding and conformational maturation [Dwek, 1995; Helenius and Aebi, 2004; Wang et al., 1996], disruptions, such as dysfunction of glycoproteins and disorders in their secretion, will cause the onset of many diseases [Dennis et al., 1999]. One of such is insulin resistance, which is undoubtedly associated with glycosylation. Insulin resistance includes numerous metabolic syndromes such as cardiovascular disease, diabetes, obesity, hypertension, and dyslipidemia [Eckel et al., 2010; Lim et al., 2014]. Although the underlying mechanisms are not clear, studies show that dysfunctional mechanisms in the hexosamine biosynthetic pathway (HBP) lead to insulin resistance in peripheral tissues [Marshall and Traxinger, 1991; Patti et al., 1999]. Interestingly, the effect of the combined injection of insulin and glucosamine in rats was proved to be more effective than the result of insulin alone [Bosch et al., 2004], indicating a strong correlation of insulin with HBP. A direct relationship between N- and O-glycans and insulin has been shown in studies conducted up to this time. For example, it has been observed that changes in the activity of enzymes associated with glycosylation cause changes in fucose and sialic acid contents in patients with insulin resistance [Cohen-Forterre et al., 1990; Rellier et al., 1999; Wiese et al., 1997; Pickup et al., 1995]. Likewise, while insulin deficiency shows an inverse correlation with plasma sialic acid amount, chronic hyperinsulinemia shows an inverse correlation with plasma sialic acid, indicating a strong relationship between insulin and glycosylation [Biol et al., 1998]. Studies on weaning rats show that the regulation ofgalactosyltransferase and α-1,2-fucosyltransferase is directly related to insulin levels. This and numerous other glycomicsbased studies strongly suggest that complex glycans are associated with insulin resistance [Lenoir et al., 2000]. In addition, current hot topics about insulin-related subjects in the context of recent knowledge and the available literature is here; in a study observing the effect of glucose loading on serotonin transporter (SERT) availability from the brainstem in humans, the presence of SERT was reported to be negatively associated with BMI after glucose loading in humans. Accordingly, it has been stated that SERT may play a role in eating behavior with the direct effect of insulin [Pak et al., 2022]. Another study using various bioinformatic assays demonstrated that genes expressed by thiopurine methyltransferase (TPMT) in patients with colon cancer

function as extracellular matrix (ECM) structural component, insulinlike growth factor binding, cell adhesion molecule binding, and growth factor binding. As a result, it has been discussed that the mRNA level of TPMT may be a new prognostic biomarker for patients with colon cancer [Jia et al., 2022]. In a study showing promising effects of bile acids in drug encapsulation for oral administration, cloned pancreatic β-cell lines were encapsulated using bile acid-Eudragit NM30D® capsules. The relationships between the stability of the capsules and the cellular biological activities were observed to show improved cell viability, insulin, inflammatory profile and bioenergetic as well as thermal and chemical stability when compared to the control [Mooranian et al., 2022].

1.6. Insulin structure

The human insulin is structurally composed of four parts: a signal peptide, B-chain, C-peptide, and A-chain (Figure 1.15 and 1.16).

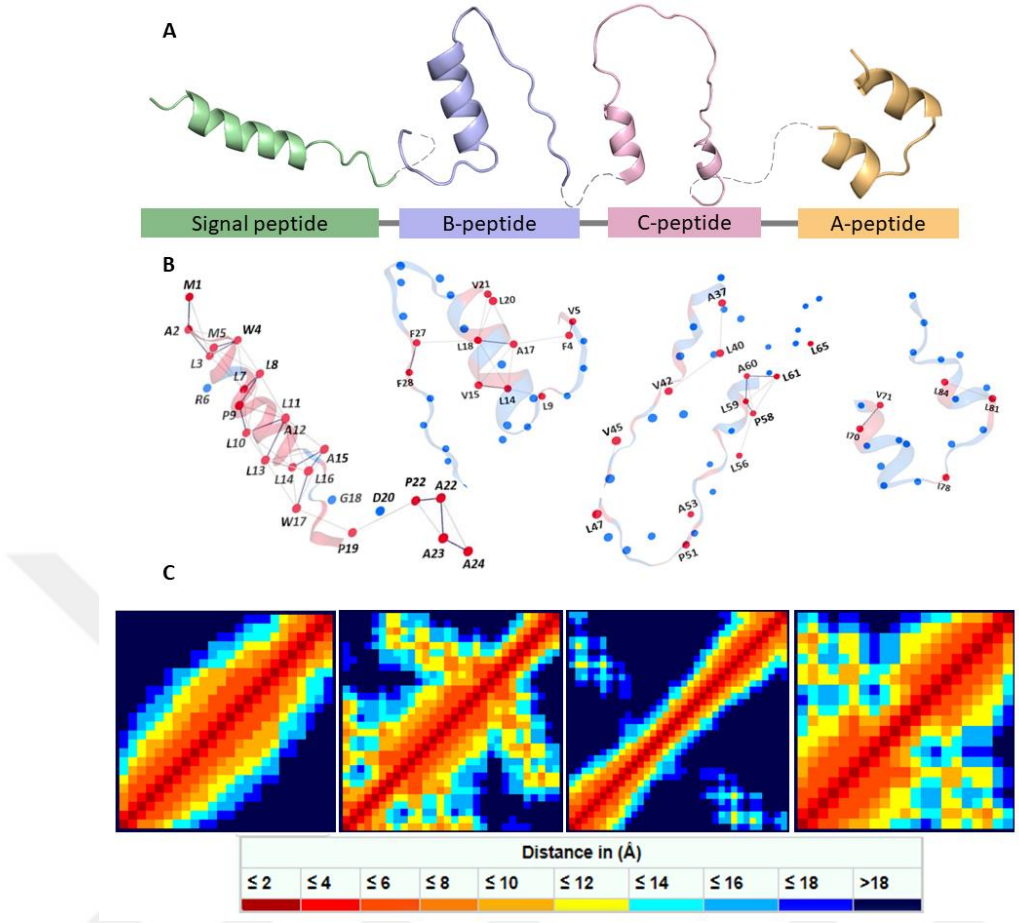


Figure 1.15 Demonstration of preproinsulin domains.

(A) Structural and schematic representation of the human preproinsulin (Uniprot ID: P01308) precursore using AlphaFold server. (B) 3D network analysis of human preproinsulin precursor for hydrophobic residues using the NAPS server. Hydrophobic residues are colored in red. (C) Residual distance matrix analysis of human preproinsulin domains using NAPS server.

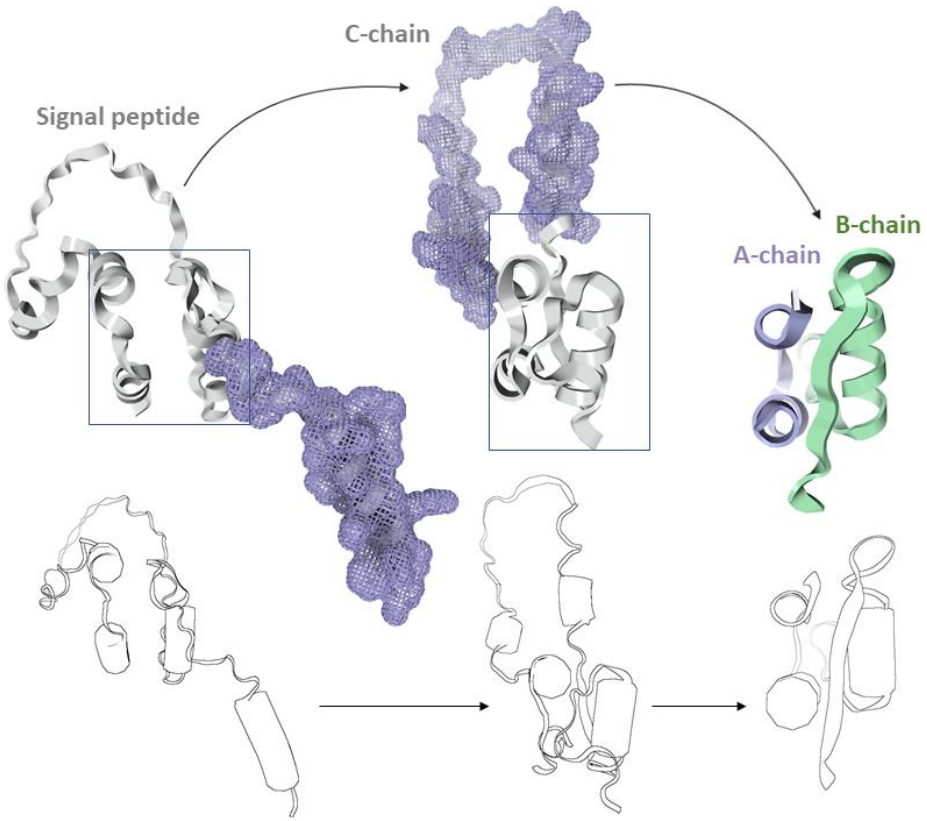


Figure 1.16 Demonstration of preproinsulin 3D conformation predicted

Symbolic representation of 3D preproinsulin structure predicted by AlphaFold server, towards mature conformation. Mature insulin is obtained by removing the twenty-four amino acid signal sequence and the C-peptide sequence (represented with colored double mesh in slate). *Visualization with Protein Imager*

β-cells in the islets of Langerhans in the pancreas, through the INS gene expression, synthesize the precursor insulin known as preproinsulin, which is 110 amino acids long. While the A-chain and B-chain are highly conserved sequences between vertebrae [Shu Jin Chan et al., 1990], the C-peptide has a highly variable amino acid length and composition sequence [Wang et al., 2012]. C-peptide was thought to be an inert byproduct in insulin synthesis and process, acting as a bridge between A and B chains [Wahren et al., 2000]. However, studies have shown that C-peptide is a peptide hormone that positively affects microvascular, kidney, and nervous system functions in diabetic model organisms [Hills and Brunskill, 2008; Wahren et al., 2012; Nordquist et al., 2009; Johansson et al., 2000; Samnegard et al., 2005]. Studies have proposed that showing the pentapeptide moiety (EGSLQ) in the C-terminal of the C-peptide may have the potential to represent a dynamic, active site for C-peptide itself. Furthermore, this pentapeptide, which includes the Glu1 and Gln6 residues, is thought to function evolutionarily as a hormone-like peptide [Johansson et al., 2002; Munte et al., 2005;

Shafqat et al., 2002]. The preproinsulin pre-part consists of a 24 amino acid long signal peptide that targets this secretory protein to the RER [Rhodes et al., 2004]. In RER, this signal peptide is removed by peptidase. The remaining polypeptide is known as proinsulin, and oxidative folding occurs in RER and insulin folds to form disulfide bonds [Liu et al., 2014]. At this point, misfolding due to mutational genetic predisposition, or the biosynthetic load of the ER, leads to the emergence of phenotypes closely related to diabetes [Shun et al., 2015]. Insulin is inactive in this state, and to take its biologically active form, the 31 amino acid (57-87) part, called C-peptide, must be removed from the polypeptide [Hills and Brunskill, 2009; Wahren et al., 2004; Wahren et al., 2007]. This process occurs in the Golgi body. In Golgi, 31 amino acids are removed by Prohormone convertase I&II and Carboxypeptidase E [Chevenne et al., 1999; Steiner, 2008; Docherty and Hutton, 1983]. After trans-Golgi processing, the remaining two short peptides are connected by disulfide bonds to form mature insulin [Orci et al., 1986]. Mature insulin consists of two peptides; a 21 amino acid long A chain corresponding to residues between 90-110 and a 30 amino acid long B chain corresponding to residues between 25 and 54. Three disulfide bridges stabilize these peptides [Brange and Langkjoer, 1993; Mayer et al., 2007]. The A-peptide consists of two short helices connected by a turn, while the B-peptide consists of a strand and a helix connected by a turn. In the B peptide, the N- and C-terminal ends show conformational flexibility [Smith et al., 2000]. Thus, the first eight residues of the B peptide can arrange either in a helical structure to form a relaxed (R) conformation or a taut (T) conformation or an open (O) conformation with an extended structure (Figure 1.17) [Smith et al., 2000; Ciszak and Smith, 1994; Yao et al., 1999; Hodgkin, 1971].

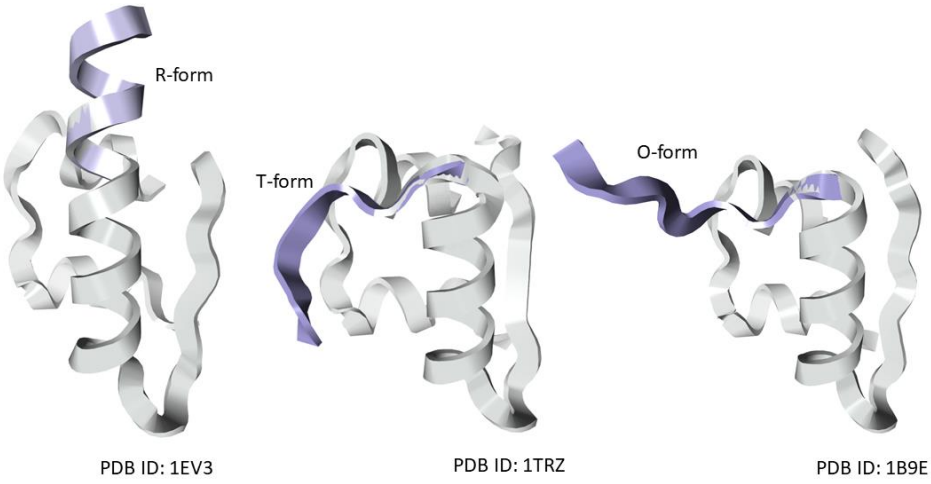


Figure 1.17 Conformational forms of insulin.

The B-peptide of insulin has three different forms known as the relaxed (R), tense (T) and open (O) form.

Another variant of the R form is that the first two residues of the B-peptide form a frayed relaxed conformation in the extended structure, along with the other two conformations [Smith et al., 2000; Ciszak and Smith, 1994]. As known, insulin in secretory granules is stored as hexameric microcrystals to prevent aggregation and degradation [307]. Monomer conformations of hexamers can be found in various conformations such as R6, T6, or T3R3 in the states concerned; these three types of hexamers differ both in their surface properties and in their hexamer conformation (Figure 1.18) [Weiss, 2009; Rahuel-Clermont, 1997].

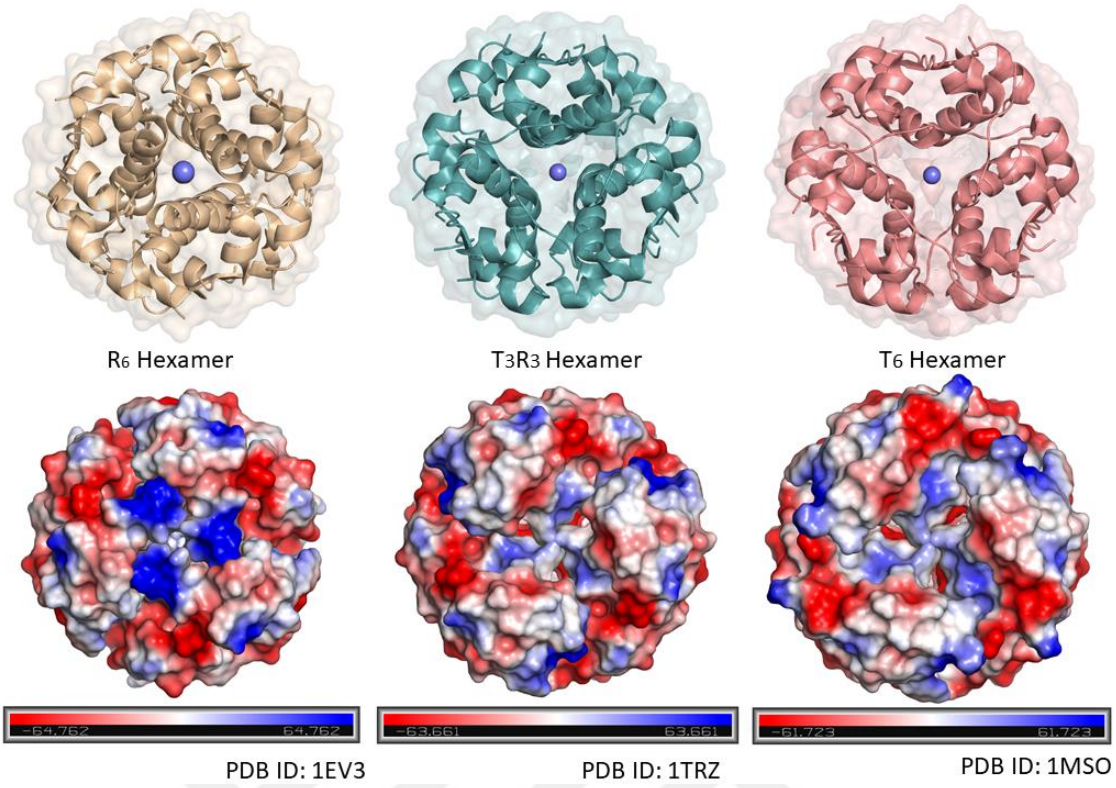


Figure 1.18 Representation of the surface mode and the electrostatic surface mode of hexamer insulin.

Hexamer insulins have different conformations such as R6, T6, or T3R3. Surface properties are different from each other.

Although they are different, they can co-exist equally in solution, and if phenol compounds such as phenol or m-cresol are bound to the hexamer, they facilitate hexameric formation such as R6 or T3R3; this is important in forming specific conformations in insulin formulations [Setter et al., 2000; Malaisse, 1979]. After being synthesized as a hexamer and given to the blood, it dissociates into dimers and monomers by diluting the blood before binding to its receptor (Figure 1.2) [Lacy et al., 1974; Smith and Ciszak, 1994]. Sometimes the dimer forms and the monomers are always small enough to diffuse through the adherent junctions of the microvasculature surrounding the tissues [Ciszak et al., 1995; Smith et al., 1996; Whittingham et al., 1998]. Hexameric inactivity and monomeric activity are characterized by the dissociation rate of the injected insulin in the body [Whittingham et al., 1998; Yao et al., 1999].

1.7. Chronological evolution of X-ray diffraction

In 1994, the T_3R_3 insulin hexamer structure was determined. T_3R_3 refers to alpha-helical and extended conformations of the B-chain's first 8 residues. According to this study, it was clearly evaluated the asymmetric unit of TR dimer, localization, and coordination of the zinc ions and their putative off-axial zinc-binding sites [Yao et al., 1999] (Protein Data Bank (PDB_ID: 1TRZ) (Figure 1.19a). In the same year, 4'-hydroxyacetanilide (Tylenol) and hexameric insulin complexes were studied. This nontoxic phenolic derivative was shown to induce the T→ R transition, forming a T_3R_3 hexamer conformation. The results of this study revealed the differences between therapeutic preparations and human insulin [Smith and Ciszak, 1994] (PDB_ID: 1TYL, 1TYM) (Figure 1.19b). In 1995, the fast-acting LysB28ProB29-human insulin was crystallized as a $T_3R^f_3$ hexamer. This structure revealed that B chains cause local conformational changes in the C terminus. It was stated that dimer stabilization was achieved by eliminating the two hydrophobic interactions and the loss of dimer interactions led to the weakening of hexamerization. This highlights the importance of the rapid-time effect of the hexameric structure in therapeutic formulations [Ciszak et al., 1995] (PDB_ID: 1LPH) (Figure 1.19c). In 1996, the complex of $T_3R^f_3$ insulin hexamer with 4-hydroxybenzamide was demonstrated. Whereas one phenol /m-cresol/resorcinol /4'- hydroxyacetanilide or methylparaben molecule is bound by each Rf-state monomer, two 4-hydroxybenzamide molecules are connected by each Rf-state monomer. The presence of two molecules causes a wide hydrogen-bonding network between T- and Rf-state trimmers. This network, organized by water molecules, gives stability to the hexamer formation [Smith et al., 1996] (PDB_ID: 1BEN) (Figure 1.19d). In 1997, the fatty acid acylated insulin structure was introduced. Substitution of B29 Lys with insulin prolongs the effect of the hormone by reversibly binding to circulating albumin [Whittingham et al., 1998] (PDB_ID: 1XDA) (Figure 1.19e). In 1998, the mutant insulin B28 Pro → Asp was introduced as a monomeric, fast-acting hormone. It was shown that this insulin could be stimulated to produce zinc hexamers via small phenolic derivatives such as phenol and m-cresol. This mutant insulin causes increased structural flexibility in the B chain through the loss of intermolecular van der Waals interactions. This demonstrates the importance of understanding the monomeric properties of the mutant insulin from the perspective of the monomermonomer interface [Whittingham et al., 1998] (PDB_ID: 1ZEG, 1ZEH) (Figure 1.19f). In the same year,

two crystalline forms of natural insulin structure were identified and refined at room temperature (PDB_ID: 1ZNJ, *unpublished*) (Figure 1.19g). In a 1999 study, Yao et al. showed that the insulin O state is different from the natural R and T forms. They demonstrated this with the B9 mutation (Ser-Glu) in human insulin. According to this study, the O state can be characterized by the natural form of insulin with a lower aggregation structure and is possibly related to insulin fibril formation [Yao et al., 1999] (Figure 1.19h). In the same year, B5 His → Tyr mutant insulin was synthesized to see if tyrosine induces the R-state of the insulin structure and if its side chain could mimic the phenol binding effect of the hexamer. This study showed that in the presence of resorcinol and phenol, B5 Tyr hexamers adopt an R6 conformation by destabilizing the T-state. This observation highlights the role of B5 His in regulating crystal packing interactions and in identifying the hexamer conformation [Tang et al., 1999] (PDB_ID: 1QIZ, 1QJ0) (Figure 1.19i).

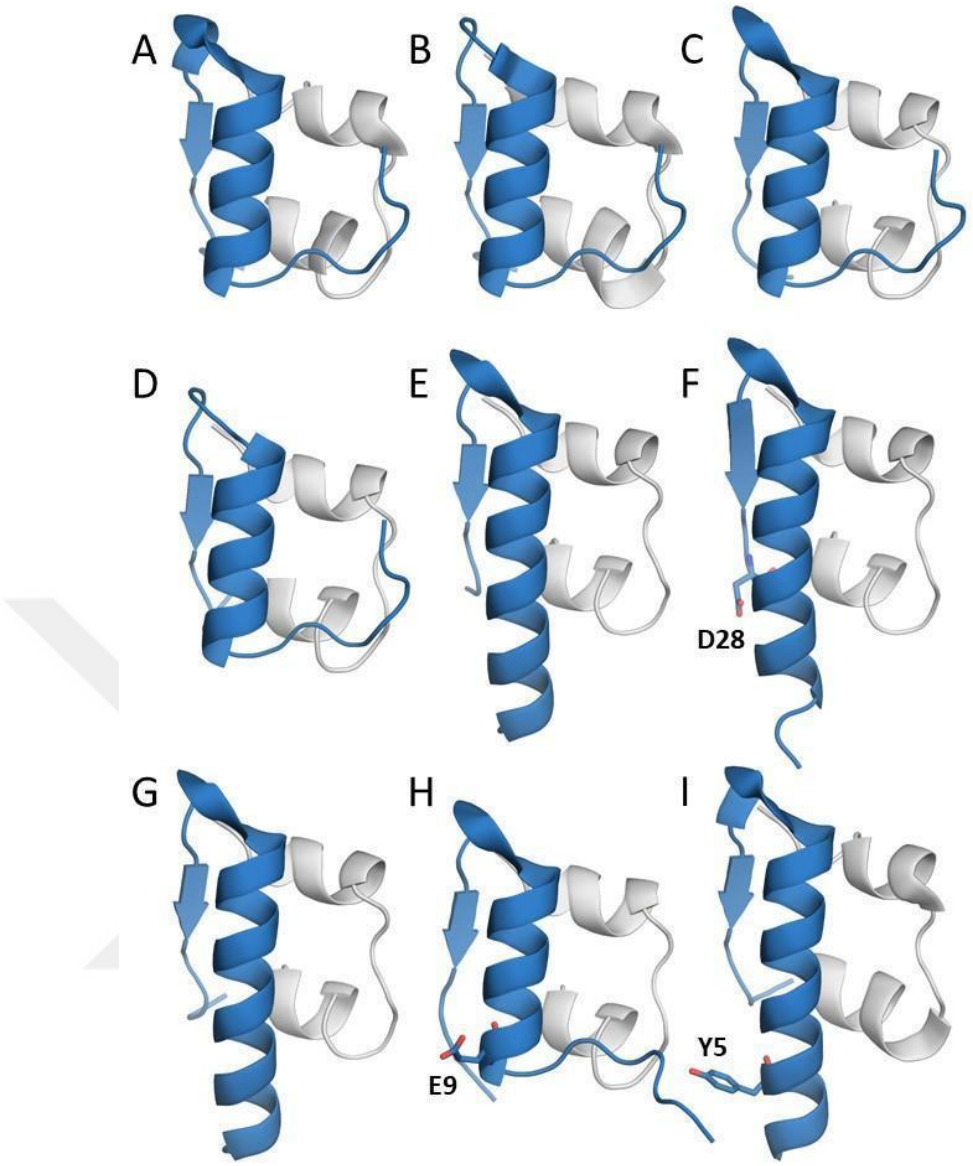


Figure 1.19 X-ray diffraction-oriented insulin structures in Protein Data Bank-I.

(A) 1TRZ, (B) 1TYL (C) 1LPH, (D) 1BEN, (E) 1XDA, (F) 1ZEG, (G) 1ZNJ, (H) 1B9E, (I) 1QIZ. Gray is A chain; Blue is B chain.

In 2000, the structure of three R_6 hexameric insulins complexed with resorcinol or m-cresol was determined. In this study, the relationship between T→R conformational transitions was determined by deprotonation of the GluB13 (EB13) side chains in R_6, $T_3R^f_3$, and T_6 hexamers [Smith et al., 2000] (PDB_ID: 1EV3, 1EV6, 1EVR) (Figure 1.20a). In the same year, a new $T_3R^f_3$ Zn-human insulin complex variant was identified by the synchrotron X-ray powder diffraction method. It had been suggested that this approach could be used to explain the structural differences of proteins associated with singlecrystal studies [Whittingham et al., 1998] (PDB_ID: 1FU2, 1FUB) (Figure 1.20b). In 2001, $T_3R^f_3$ hexamer insulin was shown to undergo a

phase transition when cryo-cooled. Based on this work, it was suggested that this transition is not related to temperature but that conditions, such as shrinking of the crystal itself, shrinks of the cryoprotectant drop, or shrinking of the thin layer of paraffin oil, are associated with pressure [Smith et al., 2001] (PDB_ID: 1G7A, 1G7B) (Figure 1.20c). In the same year, the structure of the destripeptide (B28-B30) insulin (DTRI) analog was determined by crystallizing with zinc ions at nearneutral pH. Because of the insulin B-chain C-terminus flexibility, in the native insulin dissociation process, the DTRI hexamer may be taken as a transition state. This can be considered as an advantage that can be achieved in a possible insulin hexamer degradation process due to the DTRI structure [Ye et al., 2001] (PDB_ID: 1HTV) (Figure 1.20d). Lee et al. demonstrated a three-dimensional DQ8-immunodominant peptide structure to explain the effect of the major histocompatibility complex HLA-DQ2 and HLA-DQ8 glycoproteins in nonobese diabetic (NOD) mice and humans. This study showed that diabetes may be caused by the same antigen-presenting event in NOD mice and humans. Some residues of the DQ8 structure associated with the P9 pocket are closely related to type 1 diabetes [Lee et al., 2001] (PDB_ID: 1JK8). In 2002, dab(A8)- insulin was introduced to show the correlation between stability and activity. Unfavorable helical C-cap, which refers to Thr(A8), was substituted with more favorable Arg (A8) and His (A8) amino acids. They also explore alternative mechanisms for A8 activity that suggest multiple substitutions. This study showed that Dab (A8) -insulin is preferable for hypothesis-driven protein design [Weiss et al., 2002] (PDB_ID: 1J73, 1JCA) (Figure 1.20e). In 2003, the T_6 hexameric structure of human insulin was demonstrated from a single air-dried crystal at room and cryogenic (100K) temperatures. In this study, some information about the space groups of insulin at two temperatures, their shapes, and the locations of zinc ions has been obtained. Furthermore, one of the different features compared to other structures is the absence of hydrogen bonding contacts between the Glu-B13 residue pairs. Such lack of contact causes disruption of the hydration structure in the hexamer, leading to the disorientation of the Glu-B13 side chains. These results showed that it is important that the hydrated structure is involved in stabilizing the structure of the B13 residues by regulating the charge repulsion between six different B13 glutamates [Smith and Blessing, 2003] (PDB_ID: 1OS3, 1OS4) (Figure 1.20f). In the same year, an inactive chiral analog of T_3R_3 zinc hexamer insulin structure was identified. This contains nonstandard substitution allo-Ile(A2). It is therefore called allo-Ile (A2)-insulin. Whereas there is a

high similarity between both native insulin and this variant, insulin, differences are observed near the variant region of the inactive chiral analog. Additionally, the T- and R-state conformation showed a response to chiral perturbation, suggesting their intrinsic plasticity. These variants retain their native-like protein structure, but this analog has a lower receptor binding activity [Wan et al., 2003] (PDB_ID: 1Q4V) (Figure 1.20g). In the same year, T_6 human insulin was determined at 120K at a 1.0 A resolution. In the cryofreezing condition, they observed that the B chain's first four residues undergo a conformational shift compared to the room-temperature condition. They also claimed that the disorder observed at room temperature was eliminated when frozen [Smith et al., 2003] (PDB_ID: 1MSO) (Figure 1.20h). In 2004, it was determined that the photo-cross-linking properties of Thr(A8) →His(A8) insulin (containing pazidophenylalanine) could be investigated, and the receptorbinding activity of insulin could be enhanced. This work suggests optimization capabilities involving increasing the binding interaction in nonconserved side regions of the insulin surface [Wan et al., 2004] (PDB_ID: 1RWE) (Figure 1.20i).

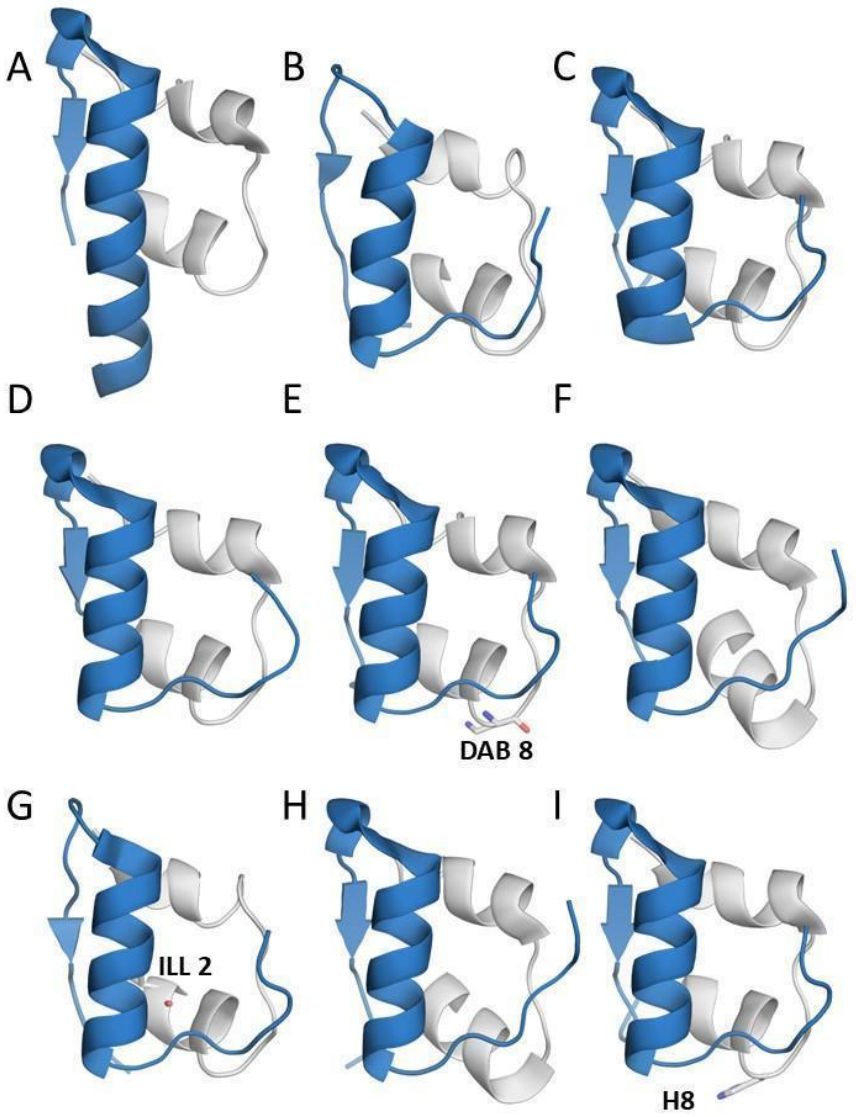


Figure 1.20 X-ray diffraction-oriented insulin structures in Protein Data Bank-II.

(A) 1EV3, (B) 1FU2, (C) 1G7B, (D) 1HTV, (E) 1J73, (F) 1OS3, (G) 1Q4V, (H) 1MSO, (I) 1RWE. Gray is A chain; Blue is B chain.

In the same year, N-lithocolyl insulin with long-acting human insulin was introduced. The addition of such hydrophobic groups increases the strong affinity for circulating serum albumin. Thus, it forms soluble macromolecular complexes at the injection sites and allows insulin analogs to be slowly absorbed into the bloodstream [Whittingham et al., 2004] (PDB_ID: 1UZ9) (Figure 1.21a). In addition, [TyrB25NMePheB26] insulin mutant function and structural features were introduced. This mutant insulin could solve a biological and structural transition problem from the metabolic, hormonal activity of insulin to the growth factor activity of the IGF-I insulin

analog. This structure is well compatible with an R_6 hexamer [Žáková et al., 2004] (PDB_ID: 1W8P) (Figure 21.1b). In 2005, substitution in this mutant insulin could disrupt receptor binding 500 times compared to other mutant insulins. To elucidate whether the variant side chain directly or indirectly regulates this reduced activity, they investigated the crystal structure of Leu (A3) - insulin through the photo-crosslinking properties of an A3 analog [Wan et. al., 2005] (PDB_ID: 1XW7) (Figure 1.21c). Even if visualized using an X-ray technique, the success of sample alignment along the synchrotron beam is limited due to the structure of insulin, i.e., its small size and several problems with mother liquor and optical microscopes. An XRD study in 2006 proposed using the endogenous fluorescence of aromatic amino acids to identify the crystal with a high success rate. Crystallographic data showed that the approach of UV laser-stimulated fluorescence did not cause any detectable structural damage [Vernede et al., 2006] (PDB_ID: 2C8R) (Figure 1.21d). In the same year, despentapeptide insulin (DPI; des-B26-B30) was crystallized in the 20% acetic acid pH 2 condition and this structure was analyzed as the space group I222. Accordingly, DPI was not capable of forming β-helical interactions to create physiological dimer matching. This structure remained as a monomeric conformation and produced an alternative lattice dimer [Whittingham et al., 2006] (PDB_ID: 2CEU) (Figure 1.21e). In the same year, it was studied in complex with the insulindegrading enzyme (IDE), insulin B chain, amyloid-β peptide (1-40), amylin, and glucagon. Accordingly, it elucidated novel structural details, such as the amino and carboxyterminal domains of the IDE, their wide contact with each other, and the repositioning of the IDE. Therefore, this study could help design IDE-based drugs to regulate blood sugar concentrations and therefore, cerebral amyloid-β [Shen et al., 2006] (PDB_ID: 2G54, 2G56). In 2007, two novel crystal forms of human insulin were discovered. In this study, a new hexamer-hexamer interaction was found in the novel insulin compared to previous insulin crystals. Additionally, different binding sites for urea have been found in the crystals. These results provide more insight into the role of urea in protein denaturation [Norrman and Schluckebier, 2007] (PDB_ID: 2OLY, 2OLZ, 2OM0, 2OM1) (Figure 1.21f). In the same year, the same team focused on the neutral protamine Hagedorn (NPH) because the insulinprotamine complex was unknown. They used three different systems to study protamine binding to insulin. They crystallized it with and without urea and co-crystallized the peptide consisting of 12 arginine residues to insulin for the clear electron density of the peptide. Accordingly, the results support the lack of well-defined conformation between insulin and protamine

[Norrman et al., 2007] (PDB_ID: 2OMG, 2OMH, 2OMI) (Figure 1.21g). In 2008, two zinc human arg (A0) insulin structures were introduced to elucidate the mechanism of the decrease in insulin activity. This study showed that changes in the charged surface and two-helix orientation in the A chain and decreased surface accessibility can cause a decrease in the activity [Sreekanth et al., 2008] (PDB_ID: 2QIU) (Figure 1.21h). In the same year, $1Rb^{1+}$, $2Mn^{2+}$, and $4Ni^{2+}$ human arg insulin were determined to explain the importance of metal ions in the insulin hexamer. These structures were compared to the $2Zn^{2+}$ structure. In this study, T&R states metal coordination and their metal-induced structural changes were explained [Sreekanth et al., 2009] (PDB_ID: 2R34, 2R35, 2R36) (Figure 1.21i)

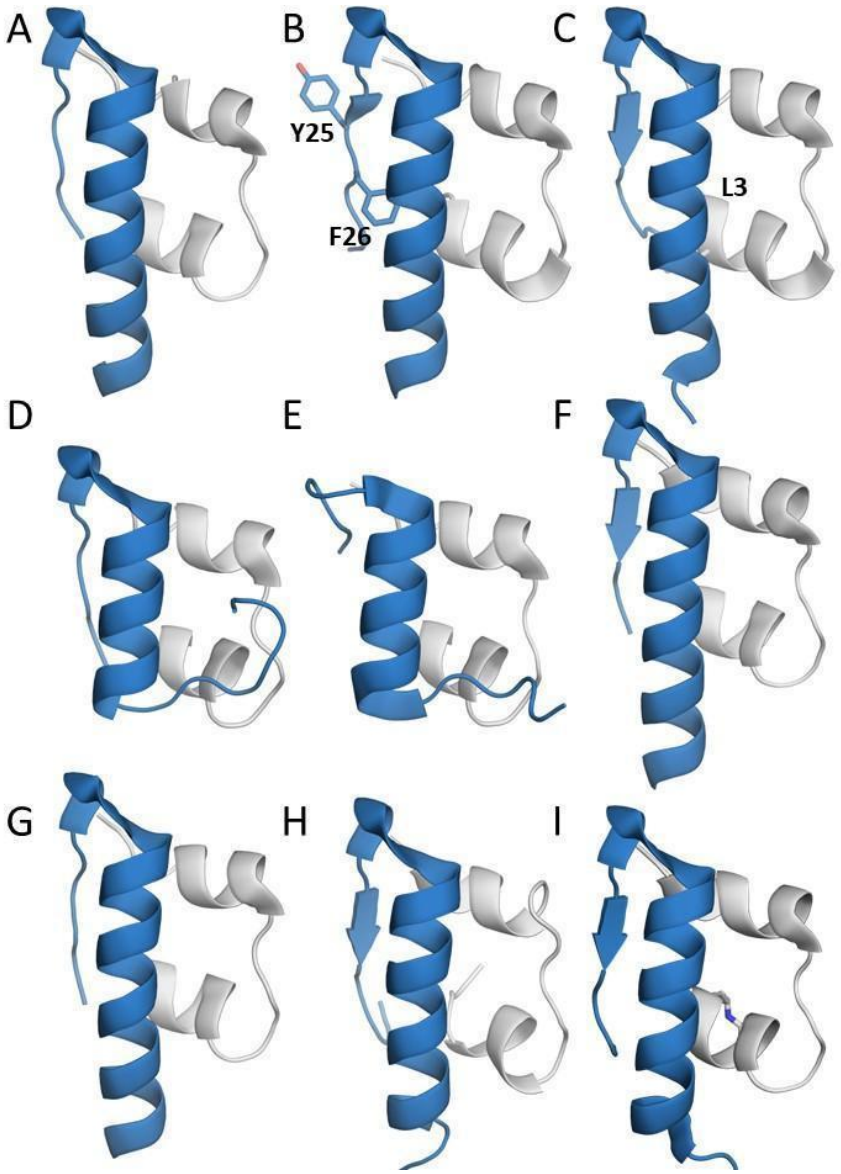


Figure 1.21 X-ray diffraction-oriented insulin structures in Protein Data Bank-III.

(A) 1UZ9, (B) 1W8P, (C) 1XW7, (D) 2C8R, (E) 2CEU, (F) 2OLY, (G) 2OMG, (H) 2QIU, (I) 2R34. A chain is gray, B chain is blue.

In the same year, the structure of the single-chain insulin analog (SCI-57), including a 6-residue linker (GGGPRR), was introduced. In this structure, the thermodynamic stability of SCI-57 is significantly higher. In addition, some of the connections found in the analog are consistent with the native insulin structure. These results also highlighted the intrinsic flexibility of an insulin monomer. This study suggests that such ultrastable SCIs may assist in enhancing the efficacy and safety of insulin replacement therapy in further studies [Hua et al., 2008] (PDB_ID: 3BXQ) (Figure 1.22a). In 2009, Ultralized insulin was studied. This study presented the crystal structures of ultralente insulin and their precursor microcrystals. According to this

study, ultralente crystals and their relationship with the protective agent methylparaben are explained. Thus, the longterm action mechanism of this insulin was revealed [Wagner et al., 2009] (PDB_ID: 2VJZ, 2VK0) (Figure 1.22b). In the same year, IDE was studied to elucidate its function on insulin clearance. They used Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) and electron capture dissociation (ECD) to investigate the composition and cleavage sites of insulin fragments produced by the IDE. This study revealed the relationship between IDE forms, their effects on the partially unfolded insulin molecule, and their affinity [Manolopoulou et al., 2009] (PDB_ID: 2WBY, 2WC0). In the same year, the structure of human insulin was re-evaluated by XRD at a resolution of 1.6 Å at cryogenic temperature (PDB_ID: 3E7Y, 3E7Z;) (Figure 1.22c). Likewise, in 2009, the structure of the T6 human nickel insulin derivative was investigated at a resolution of 1.35 Å at cryogenic temperature (PDB_ID: 3EXX;) (Figure 1.22d). In the same year, an insulin analog that binds to IR with higher affinity but has a lower affinity to insulin-like growth factor receptor (IGFR) was designed using a nonstandard mutagenesis technique. As a result of this data, it has been suggested that such insulin analogs could aid cancer studies in animal models of hyperinsulinemia [Zhao et al., 2009] (PDB_ID: 3FQ9) (Figure 1.22e). In 2010, the insulin- insulin receptor (IR) binding process was studied with the help of different insulin analogs, such as B17A insulin (site 2 mutated), Des A1-4 insulin (site 1 deleted), with and without a fluorescent probe attached. This result showed that insulin site 2 binds to the IR faster in the formation of the IR-insulin complex. Accordingly, this data demonstrates the importance of the A1-A4 terminal residues of the insulin Achain for the slow binding phase [Thorsøe et al., 2010] (PDB_ID: 2W44) (Figure 1.22f). In the same year, an insulin analog truncated at residue 26 of the B-chain (B26) was introduced. This data showed a structural convergence at the B(24)-B(26) position for active hormone analogs. It has been suggested that this novel β-turnover in residues B(24)-B(26) may be critical in the transition of insulin. These results represent structural details of prospective ligands for further rational design studies [Jiráček et al., 2010] (PDB_ID: 2WRU, 2WRV, 2WRW, 2WRX, 2WS0, 2WS1, 2WS6, 2WS7) (Figure 1.22g). In the same year, the recombinant human insulin-polysialic acid complex was investigated under microgravity and normal conditions. These results confirmed that these crystals are the same as cubic insulin crystals and that the cubic insulin did not contain polysialic acid or fragments thereof. In this study, the structure of the gene-engineered human insulin was compared with

different crystals [Timofeev et al., 2010] (PDB_ID: 3I3Z, 3I40) (Figure 1.22h). In the same year, the crystal structures of human insulin-Ni+2, Sr+2, and Cu+2 complexes were investigated to observe metal-induced 3D changes in human insulin. This data was obtained at 1.8 Å resolution at cryogenic temperature (PDB_ID: 3ILG, 3INC, 3IR0; *unpublished*) (Figure 1.22i)

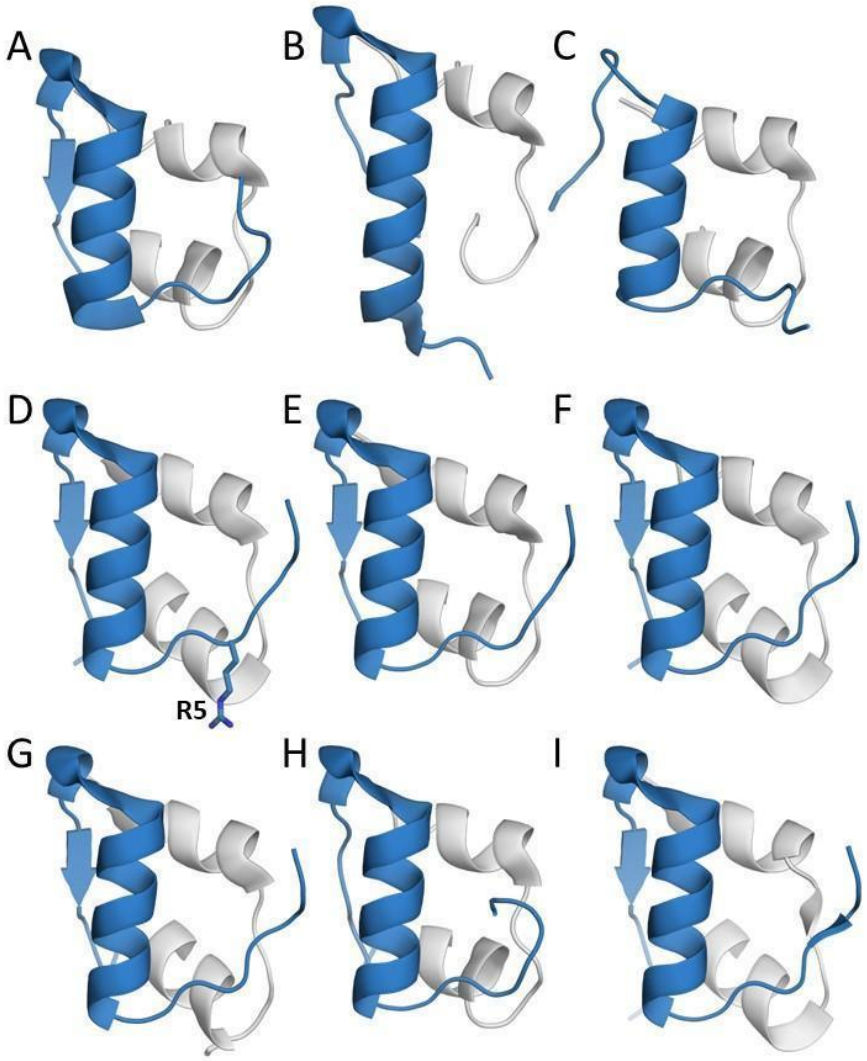


Figure 1.22 X-ray diffraction-oriented insulin structures in Protein Data Bank-IV.

(A) 2VJZ, (B) 2W44, (C) 2WRU, (D) 3BXQ, (E) 3E7Y, (F) 3EXX, (G) 3FQ9, (H) 3I3Z, (I) 3ILG. A chain is gray, B chain is blue.

In the same year, insulin biosynthesis and activity in neonatal DM were investigated structurally. This study was performed by XRD at room temperature at 2.5 Å resolution (PDB_ID: 3ROV, 3JSD; *unpublished*) (Figure 1.23a). In the same year, it

was shown that the pharmacological properties of insulin can be improved by "zinc staples" between hexamers. This structure was different from the previous one as it contained both new zinc ions and classical axial zinc ions at the hexamer-hexamer interfaces. Moreover, the function of the structure is similar to that of the long-acting formulation Lantus, although it has some analogous distinctions. This result suggests that it can provide a pharmacokinetic strategy to alter the biological properties of the insulin structure [Phillips et al., 2010] (PDB_ID: 3KQ6) (Figure 1.23b). In 2011, insulin fibrillation was studied with XRD at 2 Å resolution (PDB_ID: 3P2X, 3P33; *unpublished*) (Figure 1.23c). Additionally, the same mechanism was studied to explain insulin fibrillation at 2 Å resolution in 2012 (PDB_ID: 3V19, 3V1G; *unpublished*) (Figure 1.23d). In 2011, insulin and its synthetic receptor, known as the cucurbit [7] uryl (Q7) complex, were investigated. It was shown that there is a robust affinity between them compared to the larger protein-receptor complex. In this study, the molecular recognition of insulin has been investigated in detail [Chinai et al., 2011] (PDB_ID: 3Q6E) (Figure 1.23e). In the same year, Cu^{2+} human insulin derivative was investigated by XRD analysis at 1.12 Å resolution at cryogenic temperature (PDB_ID: 3TT8; *unpublished*) (Figure 1.23f). Additionally, in 2011, insulin analogs modified with B24, B25, or B26 were investigated to identify their structural contribution to the dimer interface. These analogs demonstrated impaired binding affinity to IR, highlighting the importance of the IR-interaction role for amide hydrogens. Together, the study revealed two new analog crystal structures [Antolíková et al., 2011] (PDB_ID: 3ZQR, 3ZS2) (Figure 1.23g). In the same year, Lys (B29) des (B30) human insulin structure was studied at 1.6A resolution at cryogenic temperature (PDB_ID: 3ZU1; *unpublished*) (Figure 1.23h). In 2012, a covalently linked insulin dimer was designed to evaluate the relationship between insulin stability, function, and insulin oligomerization. This dimer is the same as a human insulin dimer and can easily form a hexamer structure. Although this engineered dimer did not bind to the IR and did not cause any signal pathways, it was thermodynamically stable, and no amyloid fibril was formed under mechanical stress. This highlights the importance of oligomerization for insulin stability [Vinther et al., 2012] (PDB_ID: 3U4N) (Figure 1.23i).

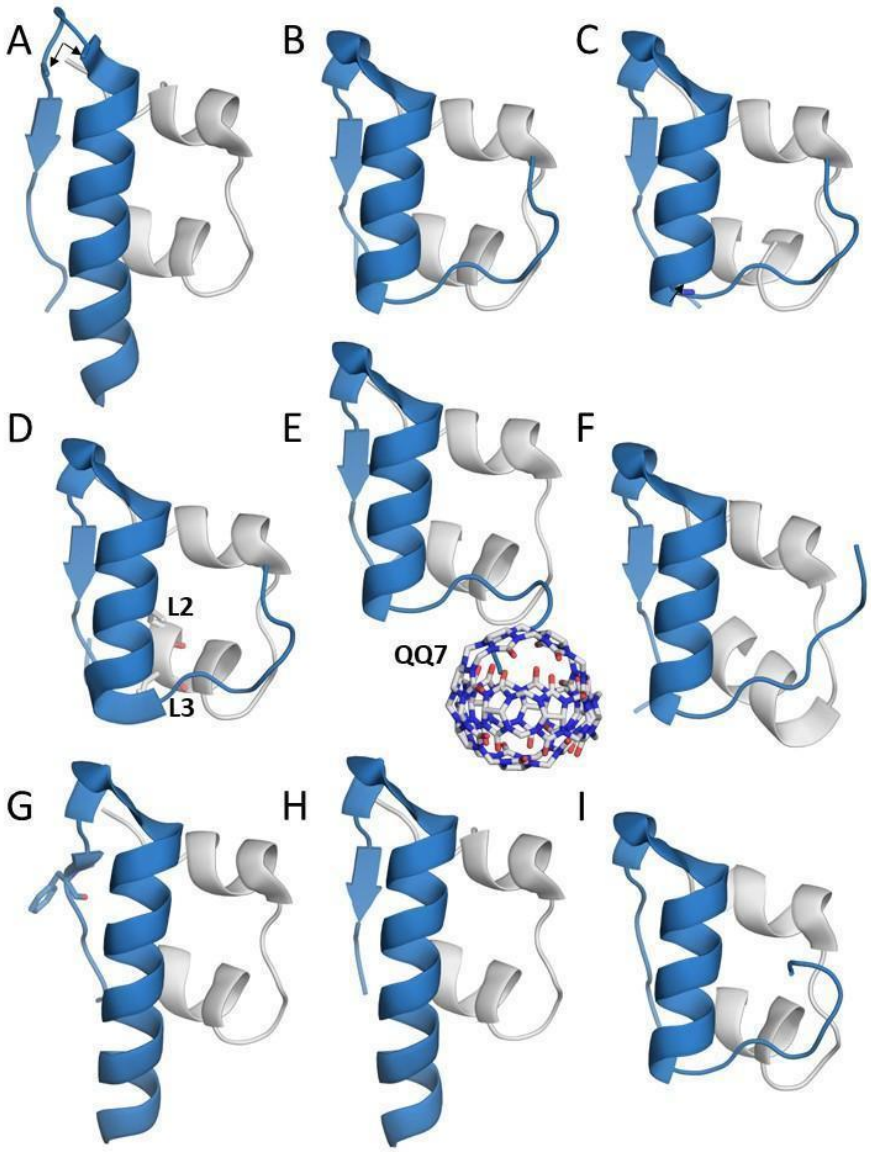


Figure 1.23 X-ray diffraction-oriented insulin structures in Protein Data Bank-V.

(A) 3ROV, (B) 3KQ6, (C) 3P2X, (D) 3V19, (E) 3Q6E, (F) 3TT8, (G) 3ZQR, (H) 3ZU1, (I) 3U4N. A chain is gray, B chain is blue. QQ7: curcumin.

In the same year, 1E6-HLAA*0201-peptide, a type of preproinsulin, and its association with CD8(+) T cells were investigated to explain the underlying mechanism of type 1 diabetes. This study suggests that such an association could potentially lead to CD8(+) T cellmediated autoreactivity and thymic escape [Bulek et al., 2012] (PDB_ID: 3UTQ, 3UTS, 3UTT). In the same year, the manganese derivative of insulin structure was investigated in high resolution in cryogenic structure [Prugovečki et al., 2012] (PDB_ID: 4FKA) (Figure 1.24a). In 2003, the primary binding sites of insulin and its receptor were crystallized to investigate the interaction of this complex in detail. These results can help design therapeutic insulin analogs through this interaction for further

studies [Menting et al., 2013] (PDB_ID: 3W11, 3W12, 3W13). In the same year, $2Zn^{+2}$ human insulin was studied at 0.92 resolution at cryogenic temperature (PDB_ID: 3W7Y; *unpublished*) (Figure 1.24b), while the same molecule was investigated at 1.15A resolution at room temperature (PDB_ID: 3W7Z; *unpublished*). In the same year, the dodecamer human insulin structure was studied at 1.4A resolution at room temperature (PDB_ID: 3W80; *unpublished*) (Figure 1.24c). In the same year, the C-terminus of the B-chain of insulin was crystallized to illuminate the Insulin-IR complex. For this, they characterized a series of B24-modified insulin analogs. This data indicates that the aromatic L-amino acid at the B24 site plays a critical role in binding insulin to its receptor [Žáková et al., 2013] PDB_ID: 3ZI3) (Figure 1.24d). In the same year, insulin degludec structure, self-assembly, and ligand binding properties were introduced using orthogonal structural methods. The folding properties of this insulin are similar to human insulin. Accordingly, this study gives us the R_6, T_3R_3 and T_6 states of information in detail. These results identified the connections between quaternary insulin structures of multihexamers, dihexamers, hexamers and their inherent propensity associated with their allosteric state [Steensgaard et al., 2013] (PDB_ID: 4AJX, 4AJZ, 4AK0, 4AKJ) (Figure 1.24e). In the same year, the importance of disulfide bonds was studied to investigate the biological activity of insulin. In this study, they designed insulin with an additional disulfide bond to investigate whether it could increase the structural stability of insulin for pharmaceutical uses. This analog had a higher affinity for IR and increased glucodynamic potency compared with human insulin. It also prevented the formation of insulin fibrils and the R- state of insulin, as well as having the ability to form hexamers. This design is the first analog of human insulin to which an additional interchain disulfide bond has been added [Vinther et al., 2013] (PDB_ID: 4EFX) (Figure 1.24f). In the same year, four different normal-acting wild-type human insulins were investigated using a combination of single-crystal protein crystallography (PX), nuclear magnetic resonance (NMR), small-angle X-ray scattering (SAXS), and mass spectrometry (MS) techniques. While all products showed similar oligomeric assembly in DLS and SAXS analysis, they showed a very close association with human insulin in NMR analysis. Together, a meta-analysis of 24 structures revealed close similarity between each other in some variables such as product batch, analytical approach, biological origin, and country origin of that product [Fávero-Retto et al., 2013] (PDB_ID: 4EWW, 4EWX, 4EWZ,4EX0,4EX1,4EXX, 4EY1, 4EY9, 4EYD, 4EYN, 4EYP, 4F0N, 4F0O, 4F1A, 4F1B, 4F1C, 4F1D, 4F1F, 4F1G, 4F4T, 4F4V,

4F51, 4F8F, 4FG3) (Figure 1.24g). In the same year, ester insulin was introduced for the efficient total synthesis of insulin. This insulin folded faster at physiological pH than proinsulin. Furthermore, this insulin was inactive in lower blood glucose conditions compared to synthetic insulin activity. Accordingly, the biological activity deficiency of ester insulin was discussed in detail in this study [Avital-Shmilovici et al., 2013] (PDB_ID: 4IUZ) (Figure 1.24h). In the same year, fast-acting insulin aspart, a human insulin variant B28Asp, was studied by small angle X-ray scattering measurements. Accordingly, this result showed a similar global behavior between aspart and normal human insulin. It has also shown that insulin adopts a T_3R_3 assembly. This data reveals in detail new structural information about aspart and its T- and R- state conformations [Palmieri et al., 2013] (PDB_ID: 4GBC, 4GBI, 4GBK, 4GBL, 4GBN) (Figure 1.24i).

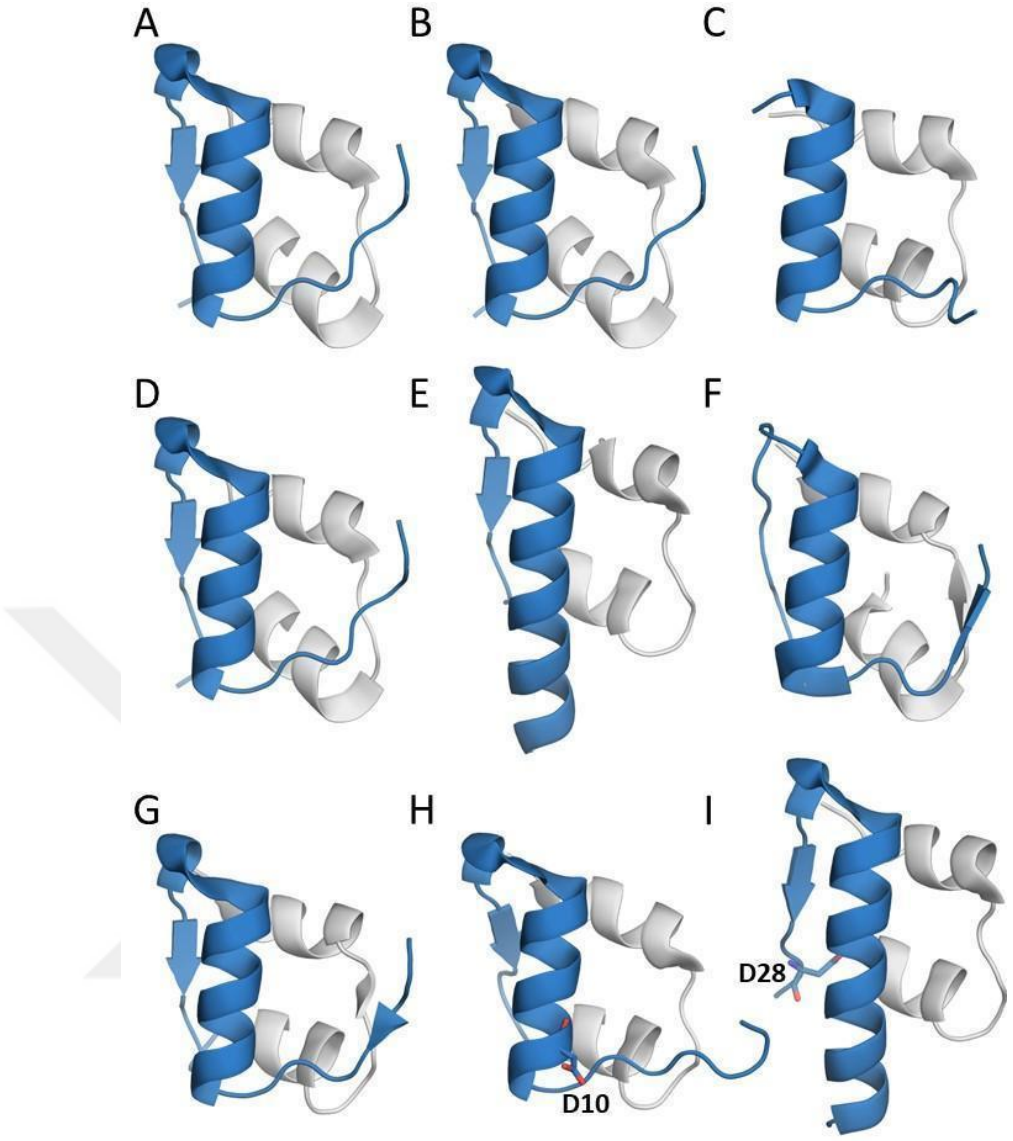


Figure 1.24 X-ray diffraction-oriented insulin structures in Protein Data Bank-VI

(A) 3W7Y, (B) 3W80, (C) 3ZI3, (D) 4FKA, (E) 4AJX, (F) 4EFX, (G) 4EWW, (H) 4IUZ, (I) 4GBC. A chain is gray, B chain is blue.

In 2014, structural and physicochemical analysis of biosimilar insulin glargine formulation was introduced at 1.6 Å resolution at cryotemperature (PDB_ID: 4IYD, 4IYF; *unpublished*) (Figure 1.25a). In the same year, the rational design, synthesis, and characterization of D-ProB8-insulin were introduced to explain the relationship between T-, R-states and the insulin receptor (IR) complex. These results revealed that a T-like state is more important for folding efficiency than an R-state compatible with an inactive form of insulin. Accordingly, the N-terminal of the B-chain must be distinct from the "conventional" T-state, and the B1-B8 segment flexibility is crucial for an insulin-IR interaction efficiency [Kosinova et al., 2014] (PDB_ID: 4CXL, 4CXN,

4CY7) (Figure 1.25b). In the same year, the B-chain's C-terminal segment of insulin was studied to explain the relationship between its receptor activation and its conformation. They proposed a transition between the receptor-bound and free conformation of insulin, resulting from the combination of function and biosynthesis and its evolutionary structural determinants [Menting et al., 2014] (PDB_ID: 4NIB, 4OGA) (Figure 1.25c). In the same year, F24 substitution by cyclohexanylalanine (Cha) of insulin was investigated to explain its importance for DM. These data revealed that B24 residue is required for the binding pocket of the IR because it provides an alternative anchor residue at the hormonereceptor interface. Accordingly, these results revealed not only the benefit of the residue presence but also the therapeutic utility of the nonstandard Cha analog side chain [Pandyarajan et al., 2014] (PDB_ID: 4P65) (Figure 1.25d). In the same year, the importance of the B21-B30 insulin site was investigated to explain the insulin-IR interaction. This study showed that native amino acids altered by the modified amino acids at the Tyr(B)26, Gly(B)26, or Asn(B)26 residues could affect the active form of this hormone. This data suggested that such analogs could provide the first structural insight into different insulin structural analyses [Zakova et al., 2014] (PDB_ID: 4UNE, 4UNG, 4UNH) (Figure 1.25g). In 2015, the insulin-β-cell chaperone sulfatide complex was cocrystallized at 1.5A resolution at cryogenic temperature (PDB_ID: 4XC4; *unpublished*) (Figure 1.25f). In the same year, two T cell receptor Mhc class II molecules containing proinsulin-derived peptides were studied to investigate the relationship between MHC molecules and TCR recognition at 2.5A resolution at a cryogenic temperature [Beringer et al., 2015] (PDB_ID: 4Y19, 4Y1A). In the same year, the relationship between insulin reactivity and a CD8+T cell clone known to induce T1D was studied to understand the disorder between the MHC class I binding channel and the insulin-derived peptide. This work proposed a new flexible peptide presentation model in the MHC peptide binding channel perspective [Motozono et al.,] (PDB_ID: 4WDI, 4Z76, 4Z77, 4Z78). In the same year, the cobalt human insulin derivative was crystallized for investigation at a resolution of 1.73A at cryo-temperature (PDB_ID: 4RXW; *unpublished*) (Figure 1.25g). In the same year, the crystal structure of human zinc insulin was studied at 1.7A resolution at a cryogenic temperature at pH 5.5 and pH 6.5 (PDB_ID: 5CNY, 5CO2, 5CO6, 5CO9; *unpublished*) (Figure 1.25h). In the same year, the dimer formation of human insulin was studied using twodimensional infrared (2D-IR) spectroscopy. This data revealed the hexameric structure of synthetic human insulin and synthetic ester insulin intermediate in the

absence of Zn metal ions and/or phenol derivatives. This result highlights the fractional crystallization of semi-racemic insulin mixtures with two synthetic proteins [Mandal et al., 2016] PDB_ID: 5EN9, 5ENA) (Figure 1.25i).

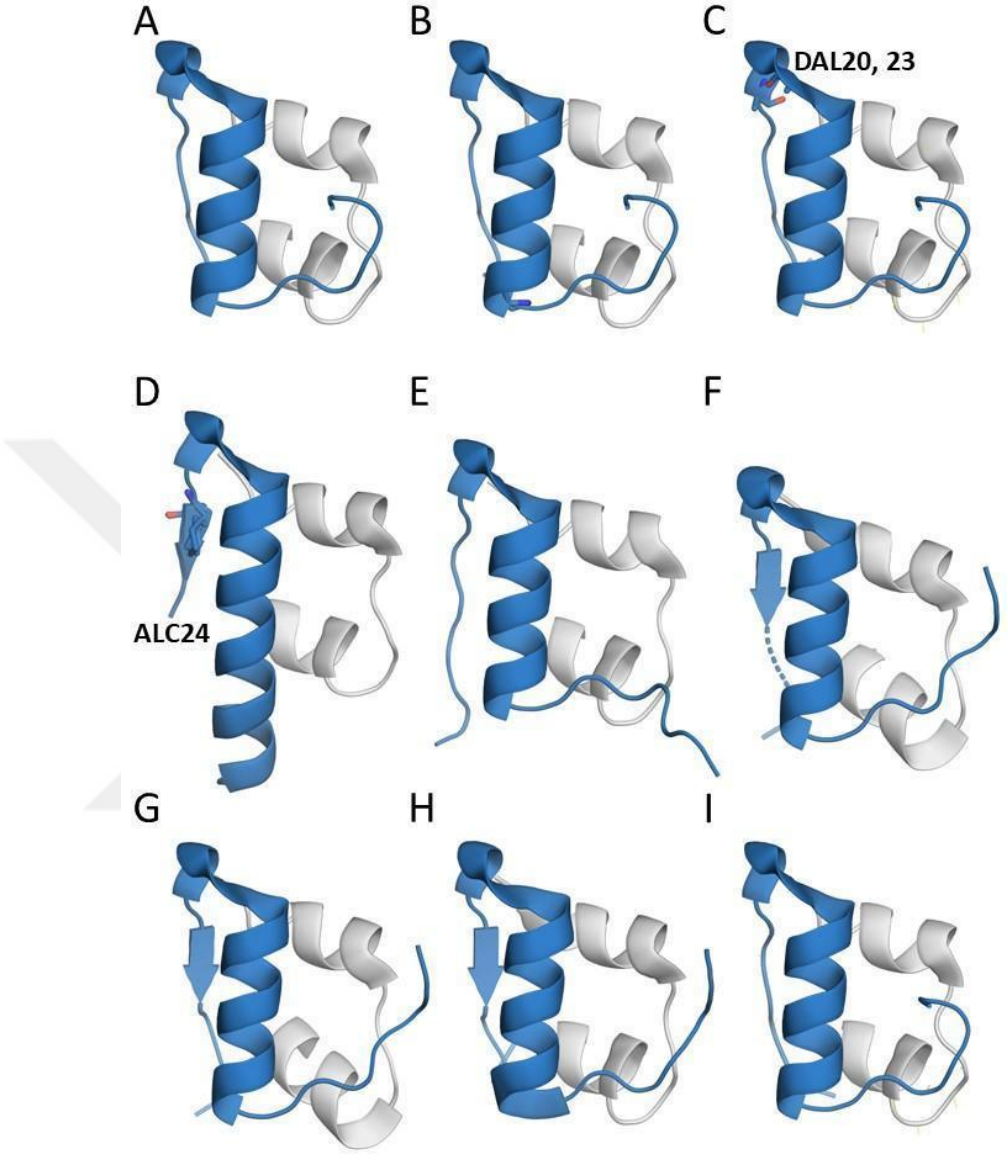


Figure 1.25 X-ray diffraction-oriented insulin structures in Protein Data Bank VII.

(A) 4IYD, (B) 4CXL, (C) 4NIB, (D) 4P65, (E) 4UNE, (F) 4XC4, (G) 4RXW, (H) 5CNY, (I) 5EN9. A chain is gray, B chain is blue.

In 2016, the B22-B30 region of insulin was studied to modulate and stabilize the receptor compatible structure of insulin. The structures and functions of 14 insulin analogs have been extensively characterized. One of the analogs has been shown to be highly active for binding to IR isoforms. These data demonstrated the importance of the B-chain C-terminal of insulin for the metabolic B-isoform of insulin receptor specificity [Vikova et al., 2016] (PDB_ID: 5BOQ, 5BPO, 5BQQ) (Figure 1.26a). In the same year,

insulin high molecular weight products (HMWPs) were characterized in terms of self-association, tertiary structure, fibrillation properties, and biological activity to explain the covalent insulin dimer state during storage of a marketed pharmaceutical formulation. These data shed light on the description of the characterization of a specific human insulin HMWP type that has been [Hjorth et al., 2016] (PDB_ID: 5BTS) (Figure 1.26b). In the same year, high-resolution structures of the 1E6 TCR bound to 7 modified peptide ligands were examined to explain the relationship between cross-reactivity of T cell peptides and autoimmunity development. These data showed how pathogen-derived antigens and T cell cross-reactivity can impair self-tolerance to stimulate autoimmune disease [Cole et al., 2016] (PDB_ID: 5C0D). In the same year, iodination of tyrosine preserved in insulin (Tyr (B26)) was studied to increase the critical properties of a fast-acting analog using molecular-mechanical and hybrid quantum methods. After a series of simulations of the 3-I-Tyr (B26)-insulin structure, the predictions were tested by X-ray crystallography. This result paved the way for the propagation of medicinal chemistry principles on proteins [El Hage et al., 2016] (PDB_ID: 5EMS) (Figure 1.26c). In the same year, the cross-reaction of an insulin-reactive T cell with pathogen-derived antigens was studied to understand the specificity of T cells and whether cross-reactivity profiles could be predicted [Standinski et al., 2016] (PDB_ID: 5HYJ). In the same year, the synthesis of insulin lispro by ester insulin was studied using the fmoc chemistry SPPs-based approach. Accordingly, the optimization of these pilot studies can provide an effective production of insulin lispro using Fmoc chemistry SPPS [Dhayalan et al., 2017] (PDB_ID: 5UDP) (Figure 1.26d). In 2017, two forms of human recombinant insulin structure (Insugen and Intergen) were observed at ultra-high resolution at 100K. In this study, the differences between electron density properties are examined in detail [Lisgarten et al., 2017] (PDB_ID: 5E7W) (Figure 1.26e). In the same year, amino acid mutagenesis was studied by replacing the proline residue at position 28 of the insulin B-chain (ProB28) with (4S)-hydroxyproline to generate more rapidly dissociating insulin. Accordingly, these data pave the way for engineering therapeutic insulin in terms of protein design and medicinal chemistry [Lieblich et al., 2017] (PDB_ID: 5HPR, 5HPU, 5HQI, 5HRQ) (Figure 1.26f). In the same year, whether endogenous ligands such as serotonin and dopamine stabilize insulin oligomers was investigated by protein crystallography (PC) and molecular dynamics (MD). Accordingly, it was observed that serotonin binds well to insulin hexamer and stabilizes it in the T3R3 structure. This result demonstrated that a T3R3 oligomer is a reasonable conformation for insulin

pancreatic storage via neurotransmitters, suggesting important implications for clinical insulin formulation [Palivec et al., 2017] (PDB_ID: 5MAM, 5MT3, 5MT9) (Figure 1.26g). In the same year, the substitution of the A6-A11 disulfide bond of insulin was studied to investigate its effects on insulin activity. Accordingly, this data sheds light on the allosteric role of the A6-A11 disulfide bond in regulating the transition of insulin to its active conformation [Van Lierop et al., 2017] (PDB_ID: 5T7R) (Figure 1.26h). In the same year, the X-Ray structure of Insulin Glargine was studied at 1.7 resolution at cryogenic temperature (PDB_ID: 5VIZ; *unpublished*) (Figure 1.26i).

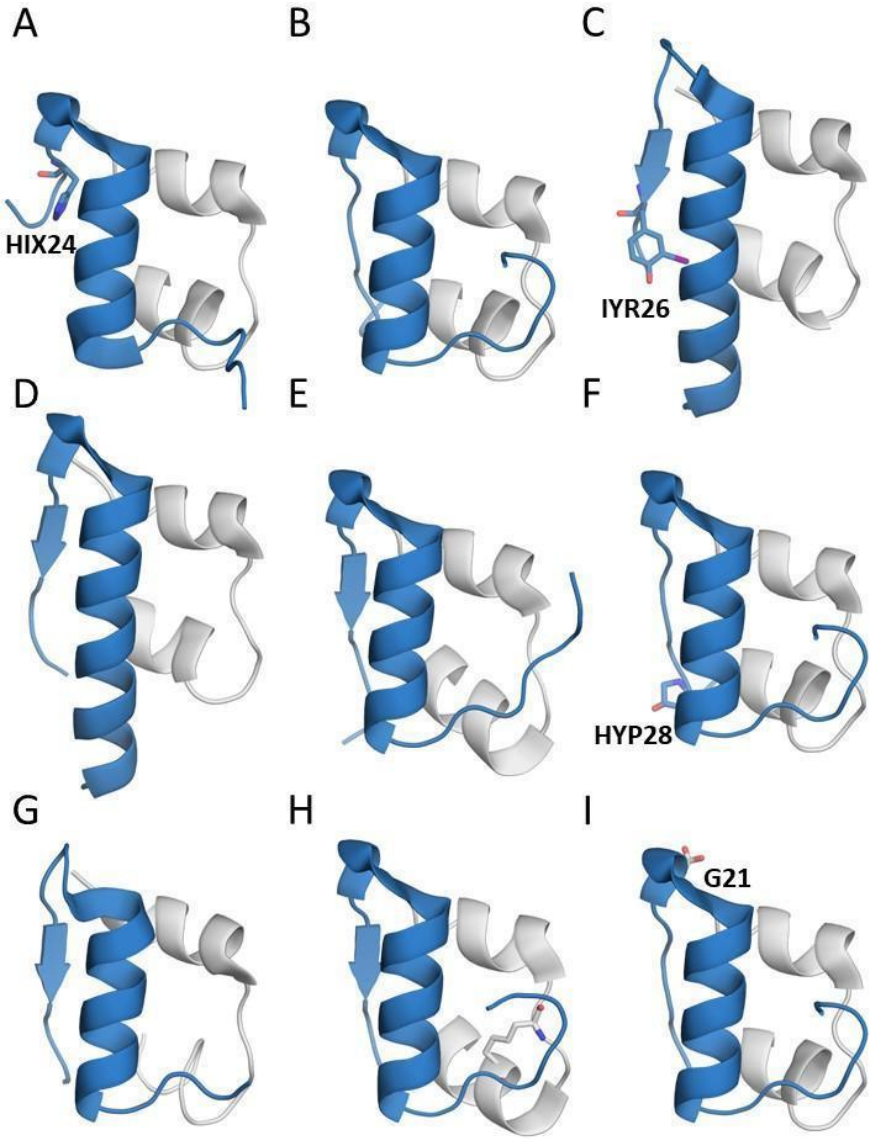


Figure 1.26 X-ray diffraction-oriented insulin structures in Protein Data Bank-VIII.

(A) 5BOQ, (B) 5BTS, (C) 5EMS, (D) 5UDP, (E) 5E7W, (F) 5HPR, (G) 5MAM, (H) 5T7R, (I) 5VIZ. A chain is gray, B chain is blue.

In the same year, the ultrastable single-chain insulin analog was studied to have

significant resistance to thermal inactivation. This data showed that the stability of this analog was higher than that of wild-type insulin. Accordingly, such an analog can provide a global therapeutic platform without worries about the cold chain [Glidden et al., 2018] (PDB_ID: 5WDM) (Figure 1.27a). In 2018, varied insulin-proline analogs at position B28 and their diverse T- & R- state conformation were studied at 1.17 resolution at cryogenic temperatures (PDB_ID: 5UOZ, 5UQA, 5URT, 5URU, 5USP, 5USS, 5USV, 5UU2, 5UU3, 5UU4; *unpublished*) (Figure 1.27b). In the same year, the insulin capture and degradation mechanisms were studied using X-ray crystallography, cryoEM, HDX-MS, and SAXS to investigate both insulin-dependent and apo- dimeric insulindegrading enzyme (IDE) states. Accordingly, this data not only explains the mechanism by which IDE disrupts amyloidogenic peptides but also suggests new structural insights for progressive IDE-based therapies [Zhang et al., 2018] (PDB_ID: 5WOB). In 2018, the Trp(B)26-insulin analog was studied to elucidate the insulin's dimer interface. In this study, it was suggested that residue substitution of Trp(B)26 and Tyr(B)26 can preserve the natural structure of insulin during the dimerization process. This is an exemplary study on the effects of protein stability-dependent conformation on protein assembly optimization [Rege et al., 2018] (PDB_ID: 6CK2) (Figure 1.27c). In the same year, the synthesis, structure, and properties of a gadoliniumcaged cobalt-III (Gd (III)), including the insulin-linked peptide cage, were studied. This study provides much information about the structural dynamics of insulin, including the stability of the hexamer [Taylor et al., 2018] (PDB_ID: 6P4Z) (Figure 1.27d). In 2019, human insulin was studied by forming a complex with meta-cresol at 2.2 Å at cryogenic temperature (PDB_ID: 6GNQ; *unpublished*) (Figure 1.27e). In the same year, the human seleno-insulin analog was studied to overcome the competition between disulfide aggregation and pairing. In this study, an internal disulfide bridge was replaced with diselenide to increase the stability and foldability of human insulin [Weil-Ktorza et al., 2019] (PDB_ID: 6H3M) (Figure 1.27f).

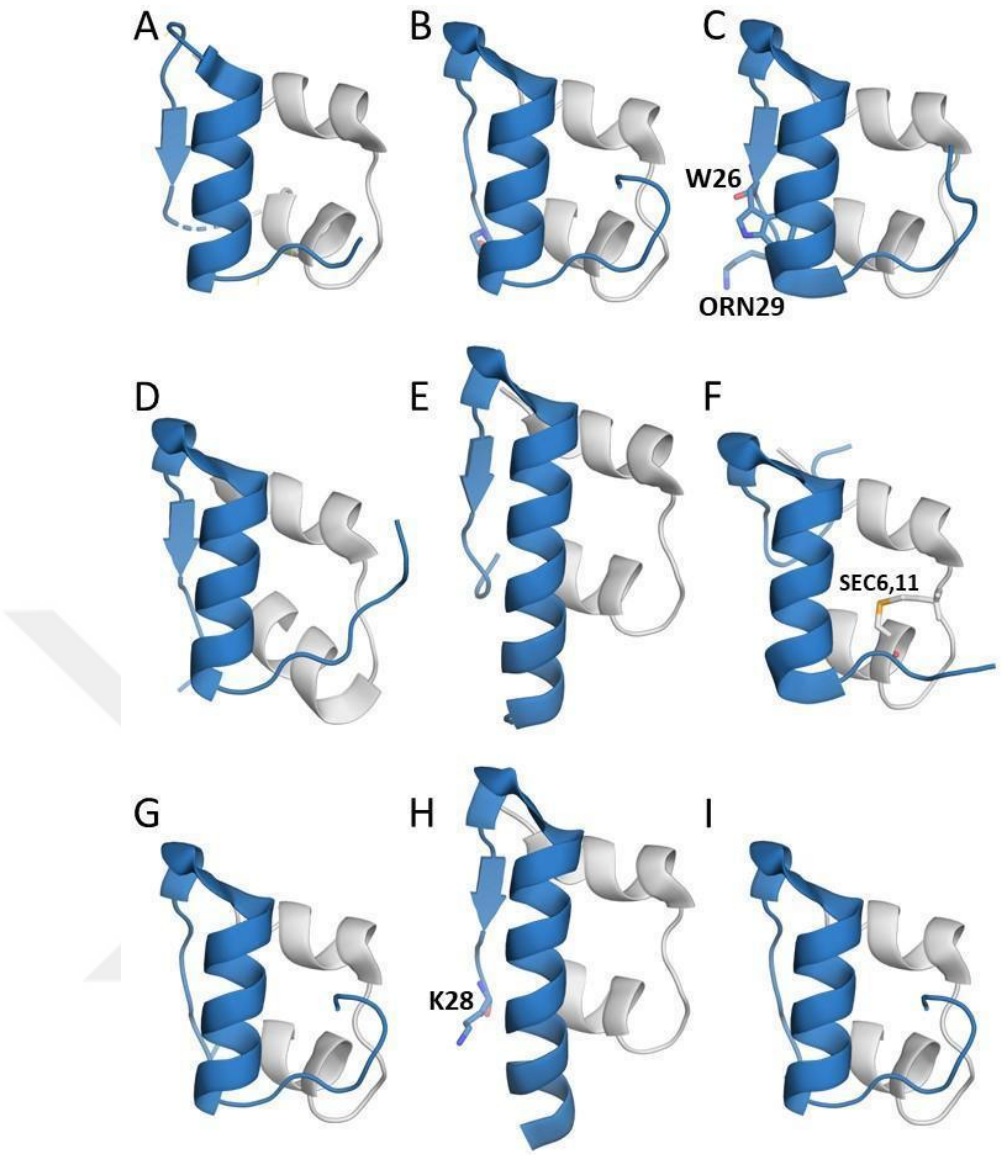


Figure 1.27 X-ray diffraction-oriented insulin structures in Protein Data Bank-IX.

(A) 5WDM, (B) 5UOZ, (C) 6CK2, (D) 6P4Z, (E) 6GNQ, (F) 6H3M, (G) 6S34, (H) 6NWV, (I) 6O17. A chain is gray, B chain is blue.

In the same year, zinc-free dimeric human insulin was studied at 1.35 Å resolution in the cryo-synchrotron (PDB_ID: 6S34; *unpublished*) (Figure 1.27g). In 2020, Insulin lispro was studied at 1.6 Å at cryogenic temperature (PDB_ID: 6NWV; *unpublished*) (Figure 1.27h). In the same year, recombinant human insulin was studied at 1.58 Å at cryogenic temperature (PDB_ID: 6O17; *unpublished*) (Figure 1.27i). In the same year, ultra-long oral insulin, glargine-like insulin, was studied to explain its pharmacological and biological properties. This analog is designed to exhibit ultra-long pharmacokinetic properties to eliminate variability in plasma exposure [Hubalek et al., 2020] (PDB_ID: 6S4I, 6S4J) (Figure 1.28a). In the same year, a four-disulfide insulin

analog, characterized by its high aggregation potency and stability, was studied to prevent the inactivation of insulin by fibrillation/aggregation. They added a 4th disulfide bond between a C-terminal of the Å chain and the Cterminal of the B chain. They showed that this analog was similar to native insulin in mice and displayed higher aggregation stability [Xiong et al., 2020] (PDB_ID: 6TYH) (Figure 1.28b). In the same year, Mini-Ins, an insulin-like peptide found in cone snail venom, was introduced to overcome self-assembly as dimers/hexamers of insulin and its analogs. This analog has insulin-like *in vivo* bioactivity, *in vitro* insulin signaling, and receptor binding affinity. This study paves the way for therapeutic insulin development [Xiong et al., 2020] (PDB_ID: 6VEP, 6VET) (Figure 1.29).

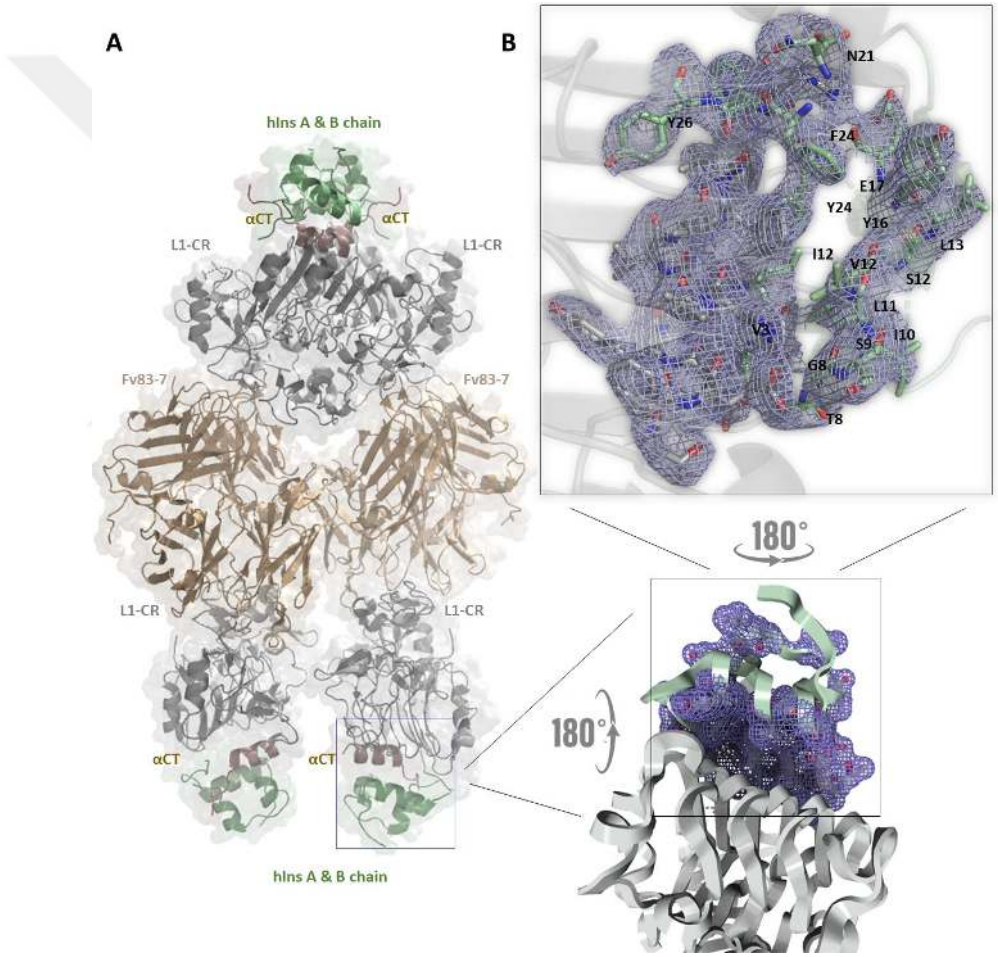


Figure 1.28 Representation of the human insulin and microreceptor complex.

(A) Overview of the microreceptor and insulin complex (PDB ID: 6VEP). Insulin is colored in pale green. (B) A closer look at the microreceptor and insulin complex. (B) Active residues in insulin & receptor interaction were labeled and demonstrated in mesh mode. Insulin is colored in pale green, and the receptor is colored in gray. Receptor's αCT, L1-CR, and Fv83-7 domains are different from each other. hIns A & hIns B: Human Insulin Chain A & Human Insulin Chain B.

In 2021, insulin glulisine was studied by combining the methods of biophysical characterization and X-ray crystallography to understand the function and structure of glulisine. This study revealed information about the pharmacodynamic and pharmacokinetic behavior of glulisine [Gillis et al., 2021] (PDB_ID: 6GV0) (Figure 1.28c).

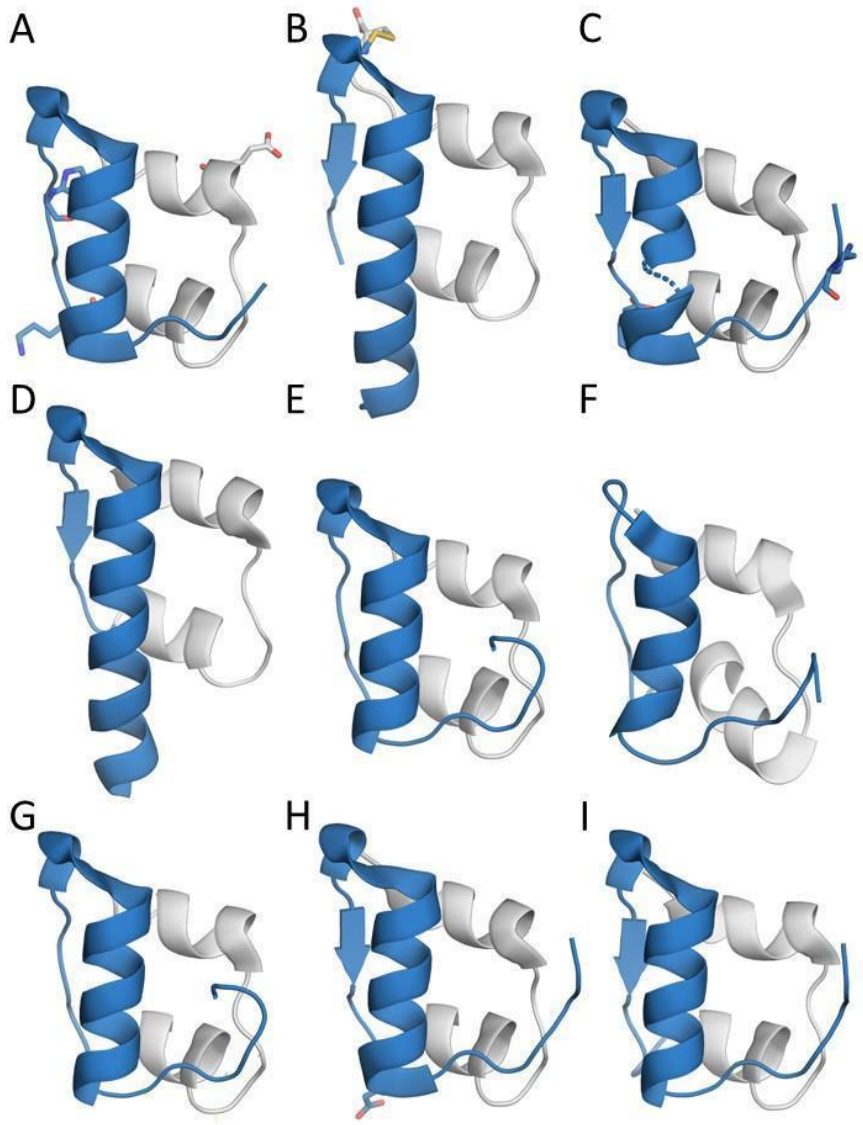


Figure 1.29 X-ray diffraction-oriented insulin structures in Protein Data Bank-X.

(A) 6S4I, (B) 6TYH, (C) 6GV0, (D) 7JP3, (E) 7NHU, (F) 7QAC, (G) 7RKD, (H) 7S4Y, (I) 8GSG. A chain is gray, B chain is blue.

In 2021, the hexamer Des-B29, B30-insulin variant was studied in detail (PDB_ID: 7JP3, *unpublished*) (Figure 1.28d). In 2021, research revealed that the accurate expression and folding of insulin were primarily determined by the inherent characteristics of the insulin's primary amino acid sequence, with eukaryotic expression

system features playing a relatively minor role. The *in vitro* transcription/translation reaction (IVTT) system effectively generated insulin analogs that matched the structure and insulin receptor affinity of those produced by yeast. In summary, the study demonstrated that the Protein synthesis Using Recombinant Elements (PURE) IVTT system is well-suited for expressing soluble molecules with advanced features and multiple disulphide bridges (PDB_ID: 7NHU) (Figure 1.28e). In 2023, researchers used X-ray powder diffraction to analyze the polymorphism of human insulin under varying pH conditions. They followed a crystallization protocol established for co-crystallization with phenolic derivatives. As the pH increased, two distinct rhombohedral (R3) polymorphs and one cubic (I213) polymorph were identified, corresponding to the T6, T3R3f, and T 2 conformations of insulin, respectively. The cubic T2 polymorph's structure was determined through multi-profile stereochemically restrained Rietveld refinement at 2.7 Å resolution. Remarkably, this represents the initial discovery of the structure of a cubic form of insulin derived from crystals cultivated in the presence of zinc ions, even though no zinc-binding was observed. The research extensively examined the unique characteristics of the polycrystalline variant in contrast to other cubic insulin structures and offered a thorough exploration of the pH-induced phase transitions (PDB_ID: 7QAC) (Figure 1.28f). In 2022, insulin glulisine has been determined to 1.25 Å resolution (PDB_ID: 7RKD, *unpublished*) (Figure 1.28g). In 2022, the initial SSX (Serial Synchrotron Crystallography) experiments employing viscous jets were conducted at ALBA beamline BL13-XALOC. Tiny insulin crystals were immersed in a lipidic cubic phase (LCP) and transported to the X-ray beam using a high-viscosity injector designed at Arizona State University. Complete datasets were successfully collected, leading to the determination of its high-resolution structures. The excellent diffraction data acquired from insulin, coupled with the lack of discernible radiation damage in the derived structures, validates that the existing capabilities of the beamline permit the precise determination of protein structures at atomic resolution, even from microcrystals as diminutive as 15 μm, utilizing viscous jets under room temperature conditions (PDB_ID: 7S4Y) (Figure 1.28h). In that same year, the T3R3 form of human insulin with a single zinc atom was determined at a resolution of 2.05 Å for the first time (PDB_ID: 8GSG) (Figure 1.28i).

1.8. Insulin analogs

In insulin replacement therapy, many types of insulin and insulin-like peptides are available. Typically, insulins are classified according to their action time and maximal effect [Konteatis et al., 2020]. Although there are many analogs available on the market, they can be listed according to their action time profile [Donner et al., 2019]. Typically, these analogs are categorized under three different headings: fast-acting analogs that are lispro, aspart, and glulisine; long-acting analogs that are glargine, detemir, degludec; and premixed analog formulations (25% lispro: 75% neutral protamine lispro; 70% neutral protamine aspart; 30% aspart, 50% lispro; 50% neutral protamine lispro) [Fullerton et al., 2018; Atkin et al., 2015]. Recombinant human insulin is a type of short-acting insulin used in the treatment of Type 2 and Type 1 Diabetes. Apart from insulin analogs, this is the same as endogenously produced insulin, using rDNA techniques [Mannucci et al., 2009; Fullerton et al., 2016]. Its marketed name is Humulin R or Novolin R, the action of this insulin begins within 30 minutes, and peak levels occur within 3-4 hours after injection [Daugherty et al., 2004]. This insulin is also known as "bolus insulin" because of its similarity with endogenous insulin in terms of its short-acting time [Daugherty et al., 2004]. Bolus insulin is often combined with long-acting insulin analogs such as insulin glargine, insulin detemir, and insulin degludec to restore the organism's overnight sugar balance [Goldman et al., 2017; Liebl et al., 2013]. This insulin has an inhalable version called Exubera, but this has been withdrawn from the market due to some adverse effects such as lung cancer development concerns. After Exubera, Afrezza was introduced. Despite similar concerns, it is still present in the US market [Oleck et al., 2016]. Typically, insulin molecules can self-assemble into the hexameric conformation, but these aggregates tend to dissociate into dimers and monomers over time [Gast et al., 2017; Nettleton et al., 2000]. The speed of this tendency is a critical step for insulin to enter the bloodstream through the interstitial fluid. Therefore, in analog formulations, the hexamer dissociation rate of insulin is often altered [Maikawa et al., 2020; Hartman et al., 2008].

In this atlas, we performed Gaussian Network Model (GNM) analysis to compare the dynamics of native insulin and marketing insulin analog structures obtained from the PDB as a bird's eye view. The GNM analysis is a Normal Mode Analysis method from which extensive information can be obtained about protein's dynamics and the natural states of the relevant proteins [Yang et al., 2006; Bahar et al.,

2005]. Thermal fluctuations obtained with normal modes in GNM analyses and those calculated from X-ray oriented structural analyses are parallel [Bahar et al., 2005]. Accordingly, cross-correlations between resi- due motions of Humulin structure were analyzed with mean square fluctuations characterized by the flexibility of the residues (Figure 1.30).

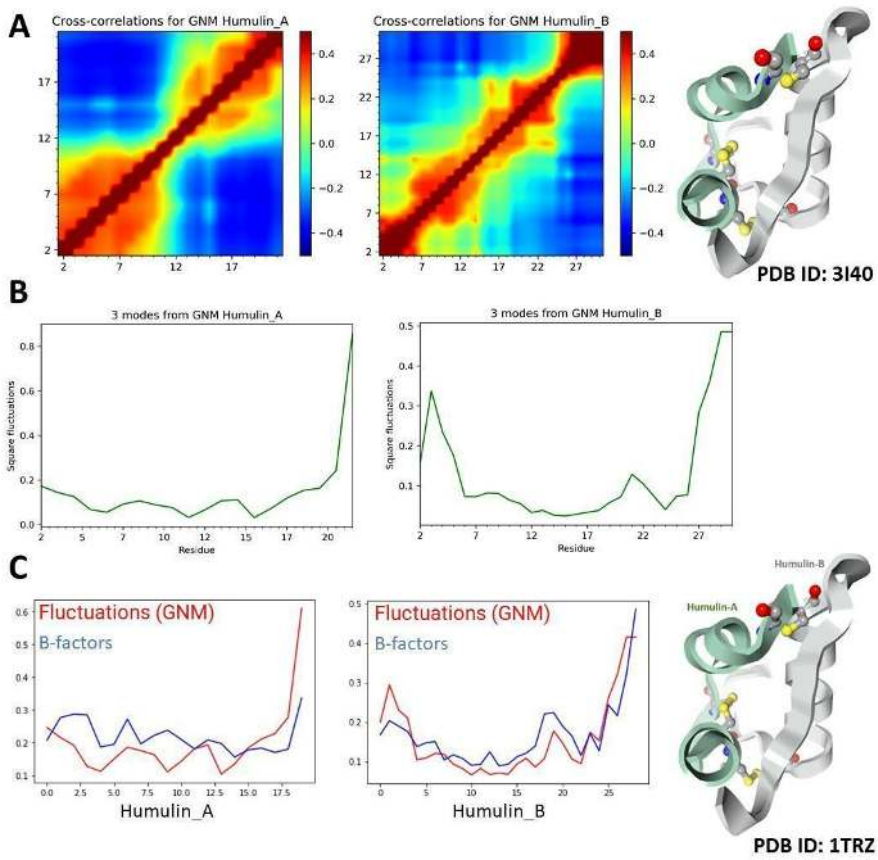


Figure 1.30 General analysis of regular insulin (Humulin).

Cross-correlation and squared fluctuations from Gaussian Network Model (GNM) analysis. In the cross-correlation result from the GNM analysis, the color range was determined as -0.5 and 0.5 to ensure compatibility between analogs' results (A). According to the squared fluctuations obtained from the weighted 3 slowest modes, global motion and residual fluctuations are relatively stable within the first 17 residues for chain A and between 7th-27th residues for chain B (B). While the graphs show the correlation between the theoretical (GNM results; fluctuations from GNM) and the experimental (B-factors), lines show only the residue fluctuations (squared fluctuations even) (C). Human insulin structures (3I40 and 1TRZ) are generated using Protein Imager server.

While the intrachain residues' crosscorrelation displayed that the first ten residues and 13-20th residues had highly correlated motions in themselves in chain A, a high correlation was observed between 2-6th, 9- 16th, 18-24th, and 27-30th residues in

themselves in chain B (Figure 1.30a). Squared fluctuations of the three modes, theoretical (fluctuation) and experimental (B-factor) results from GNM were inspected closer to further evaluate the intrinsic residue fluctuations of Humulin structure (Figure 1.30b-c). The theoretical fluctuations calculated from the three modes are aligned with the B factors in Figure 1.30c. Accordingly, the residues at position 17th elucidate the regional flexibility of the C-terminus of insulin chain B (Figure 1.30b-c). This supports previous studies on the critical aspect of the C-terminus that allows insulin to self-assemble, bind to its receptor [Mirmira et al., 1991; Dodson et al., 1983; Derewenda et al., 1991], and be convenient for analog modifications [Leyer et al., 1995]. We also performed a docking analysis to compare the interaction of regular insulin vs. its analogues with the insulin receptor. Accordingly, the insulin residues in chain A that bind to IR are GIVQTSLEYN, while in chain B, whichinteracts with IR, are HVEYLFFY. Docking of the insulins' active residues and αCT segment (TFEDYLHNVV FVPRPS) that is vital for insulin binding is shown in (Figure 1.31).

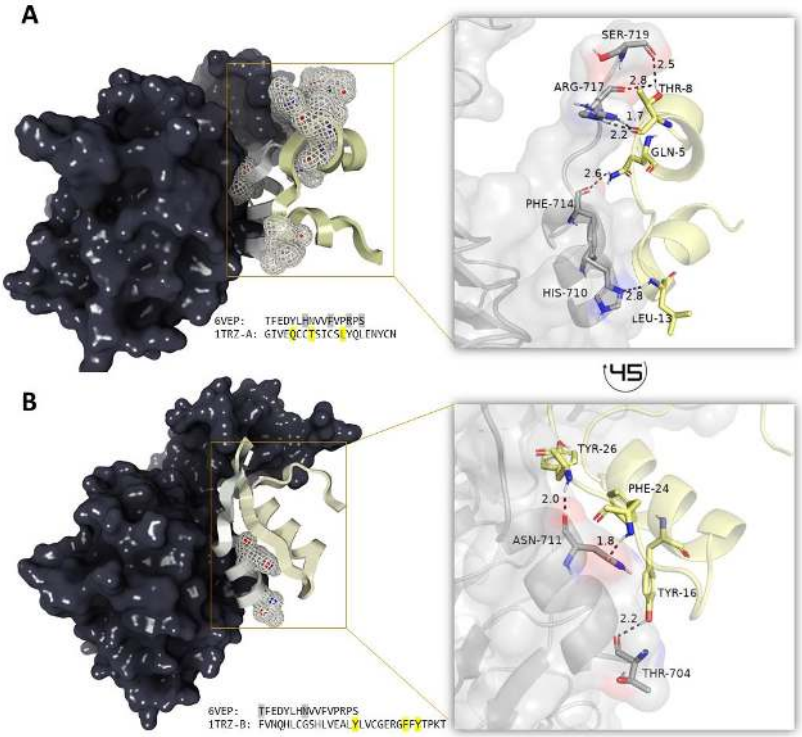


Figure 1.31 Regular insulin (Humulin) and insulin receptor interaction.

Docking the regular insulin (PDB_ID: 1TRZ) to the Apo-state IR structure (PDB_6VEP) containing αCT domain (colored in gray) that is critical for insulin binding and L1 domain (colored in black). Interaction with insulin-active residues reported in previous studies (Lukman et al., 2015) has succeeded. The insulin chain B (A) and chain A (B) are excluded for a better view. Interacted residues of αCT domain with the insulin have displayed the mesh mode in the Protein Imager server.

1.8.1. Rapid-acting analogs

This type of insulin analog has a shorter duration of activity than regular insulin due to a more rapid onset of its action. Accordingly, three fast-acting insulins are commercially available on the market: insulin lispro, insulin aspart, and insulin glulisine. Although its modifications are relatively minor, involving only a few residue substitutions, these alterations obtain the insulin to be absorbed faster, leading to a reduced tendency to self-associate into hexamers. These minor changes do not disrupt the relationship between insulin and its receptor; thus,, the analogs perform the same function as native insulin [Hartman et al., 2008].

Insulin lispro

Insulin lispro Humalog (Eli Lilly company) mimics the function of endogenously produced human insulin. Its effect begins within 15 minutes, its peak levels occur 30 to 90 minutes after injection, and the duration of action is approximately 5 hours. It is also known as "bolus insulin" as it provides high insulin levels in a fleeting time. Additionally, insulin lispro is equivalent to human insulin in terms of molar base. This analog was also produced using a nonpathogenic *Escherichia coli* strain using rDNA techniques and is the first insulin analog to be commercially marketed. The nomenclature LYS-PRO comes from LYS(B28) ⇆ PRO(B29) inversion, which is different from native insulin residues at position B28, where it is substituted by lysine, and lysine in B29 is modified by proline. Therefore, this change results in 300-fold lower self-assembly than native insulin and results in the dissolution of the dimer and monomer [Hartman et al., 2008]. Further, this modification eliminates some hydrophobic interactions by weakening some polar contacts that contribute to the insulin hexamer stability produced in the presence of zinc, phenol and m-cresol. Finally, this process causes the absorption of insulin more rapidly after the injection than regular insulin. Its absolute bioavailability ranges from 55% to 77% in the presence of doses between 0.1 and 0.2 units/kg. The serum concentration-time curve is 2390 pmol hours/L and 2360 pmol hour/L for HUMALOG U-100 and HUMALOG U-200, respectively. At the same time, the serum insulin concentration is 909 pmol/L and 795 pmol/L for HUMALOG U-100 and HUMALOG U-200, respectively. The median time to maximum concentration for both formulations is 1 hour. The mean volume of insulin

lispro distribution, corresponding to a bolus injection dose of 0.1 and 0.2 U/kg, was 1.55 and 0.72 L/kg, respectively. Its half-life after injection is t1/2 shorter than regular insulin (1: 1.5 hours) [Hartman et al., 2008]. We performed GNM analysis to compare the dynamics of regular insulin and insulin lispro structures obtained from the PDB as a bird's eye view. Accordingly, crosscorrelations between residue motions of the lispro structure were analyzed with mean square fluctuations characterized by the flexibility of the residues (Figure 1.32).

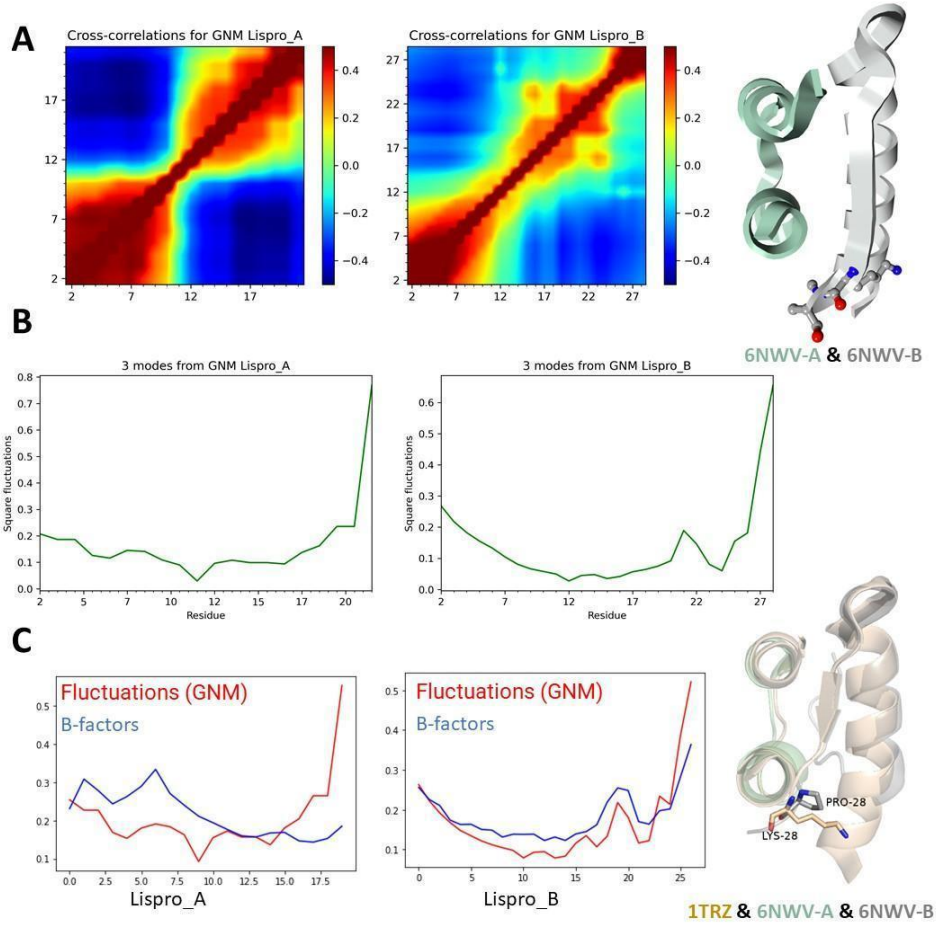


Figure 1.32 General analysis of lispro.

Cross-correlation and squared fluctuations from Gaussian Network Model (GNM) analysis. In the cross-correlation result from the GNM analysis, the color range was determined as -0.5 and 0.5 to ensure compatibility between analogs' results (A). According to the squared fluctuations obtained from the weighted 3 slowest modes, global motion and residual fluctuations are relatively stable (B). While the figures show the correlation between the theoretical (GNM results; fluctuations from GNM) and the experimental (B-factors), lines show only the residue fluctuations (squared fluctuations even) (C). Lispro structures (6NWV) are generated using the Protein Imager server. Superposition of the lispro (6NWV) and human insulin (1TRZ) was generated by PyMOL software (Root-mean-square deviation (RMSD): 0.39).

Similar to regular insulin, the intrachain residues' cross-correlation displayed that chain A had highly correlated motions with the first ten residues and 13-20th residues. In addition, chain B was observed with higher fluctuations at the 2-10th, 17-24th, and 27- 30th positions (Figure 1.32a). Three modes, theoretical (fluctuation), and experimental (B-factor) result from GNM, were inspected closer to further evaluate the intrinsic residue fluctuations of the lispro structure (Figure 1.32b-c). The theoretical fluctuations calculated from the three modes are aligned with the B factors in Fig. (32C). Accordingly, residues' position around the 15th elucidates the regional flexibility of the C-terminus of lispro chain B compared to regular insulin dynamics (Figure 1.32c). We also performed a docking analysis to compare the interaction of regular insulin vs. lispro with the insulin receptor (Figure 1.33).

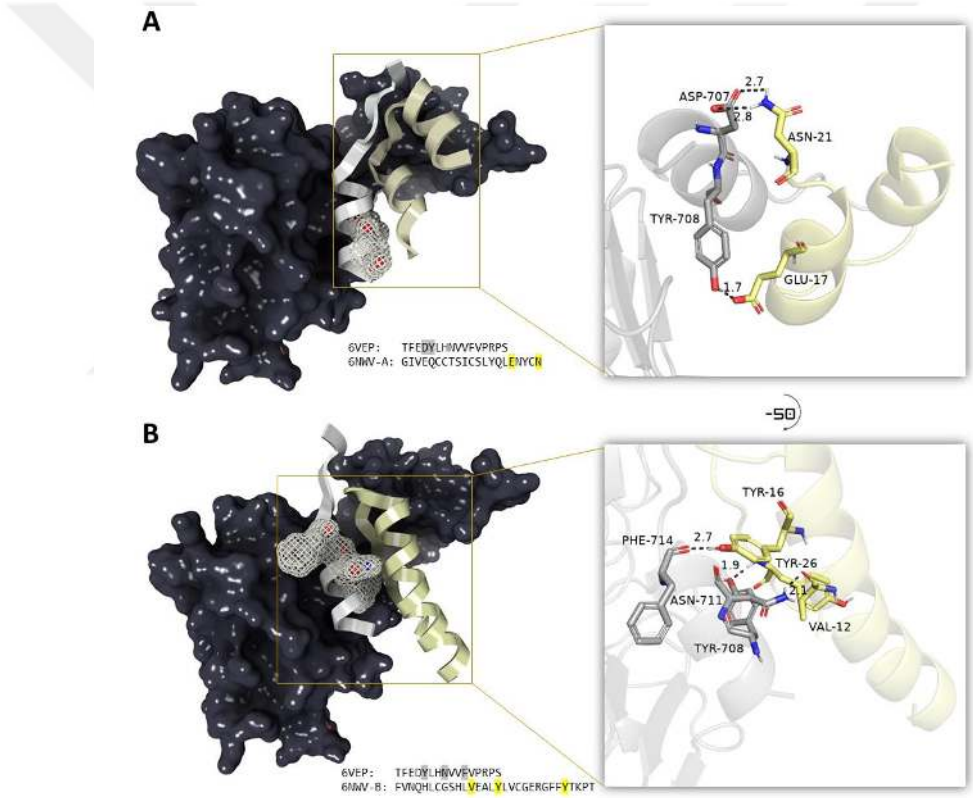


Figure 1.33 Insulin lispro and insulin receptor interaction.

Docking the lispro (PDB_ID: 6NWV) to the Apo-state IR structure (PDB_6VEP) containing the αCT domain (colored in gray) that is critical for insulin binding and L1 domain (colored in black). Interaction with insulinactive residues reported in previous studies (Lukman et al., 2015) has succeeded. The lispro chain B (A) and/or chain A (B) are excluded for a better view. Interacted residues of αCT domain with the lispro have displayed the mesh mode in the Protein Imager server.

Accordingly, Glu17, Asn21 (Figure 1.33a, pale-yellow), and Val12, Tyr16, Tyr26 (Figure 1.33b, pale-yellow) residues of lispro insulin have interacted with

Asp707, Tyr708 (Figure 1.33a, gray), and Tyr708, Asn11, Phe714 (Figure 1.33b, gray) residues of the IR-αCT segment; while Gln5, Thr8, Leu13 (Fig., 31A, pale-yellow) and Tyr16, Phe24, Tyr26 (Figure 1.31b, pale-yellow) residues of regular insulin have interacted with His710, Phe714, Arg717, Ser719 (Figure 1.31a, gray), and Thr704, Asn711 (Figure 1.31b, gray) residues of the IR-αCT segment.

Insulin aspart

Insulin aspart, NovoRapid (Novo Nordisk Company), mimics the function of endogenously produced human insulin and is similar to the lispro action. Its effect begins within 15 minutes, its peak levels occur 30 to 90 minutes after injection, and the duration of action is approximately 5 hours. It is also known as "bolus insulin" as it provides high insulin levels in a fleeting time. This analog has been produced as a biosynthetic and fast-acting insulin analog using recombinant techniques. This analog works to improve glycemic control in children and adults with diabetes. Compared to native insulin, it has a single residue modification in the position of proline (B28) replaced by aspartic acid. This change causes a higher dissociation effect of aspart than native insulin. It also has a lower binding affinity for plasma proteins, similar to regular insulin. The median time to maximum concentration for regular insulin and the aspart insulin analog is 50/40 minutes versus 120/80 minutes, respectively. Therefore, this analog is also a faster absorption feature than regular insulin after the injection. Typically, a 0.15 U/kg body weight dose for type 1 diabetes is characterized by a maximum concentration (Cmax) of 82 mU/L. Its half-life after injection is 81 minutes [Holleman et al., 2015; Chapman et al., 2002]. We also performed GNM analysis to compare the dynamics of regular insulin and insulin aspart structures obtained from the PDB as a bird's eye view. Accordingly, cross-correlations between residue motions of the aspart structure were analyzed with mean square fluctuations characterized by the flexibility of the residues (Figure 1.34).

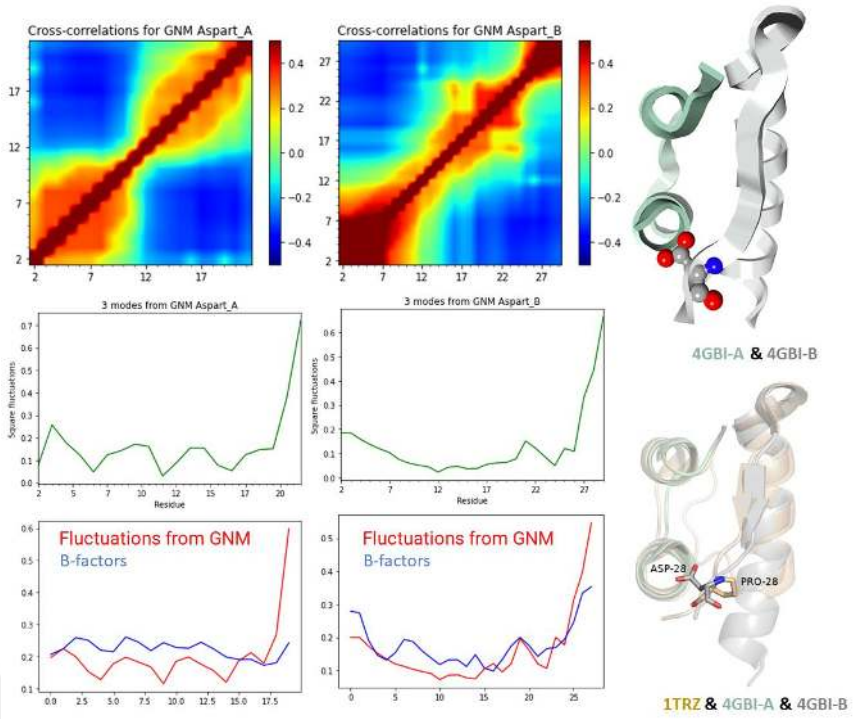


Figure 1.34 General analysis of aspart.

Cross-correlation and squared fluctuations from Gaussian Network Model (GNM) analysis. In the cross-correlation result from the GNM analysis, the color range was determined as -0.5 and 0.5 to ensure compatibility between analogs' results (A). According to the squared fluctuations obtained from the weighted 3 slowest modes, global motion and residual fluctuations are relatively stable (B). While the figures show the correlation between the theoretical (GNM results; fluctuations from GNM) and the experimental (B-factors), lines show only the residue fluctuations (squared fluctuations even) (C). Aspart structures (4GBI) are generated using the Protein Imager server. Superposition of the Aspart (4GBI) and human insulin (1TRZ) was generated by PyMOL software (Root-mean-square deviation (RMSD): 0.86).

Similarly, the intrachain residues' cross-correlation displayed that the first ten residues and 13-20th residues had highly correlated motions in chain A as well. In addition, chain B was observed with higher fluctuations at the 2-11st, 17-23rd, and 27-30th positions (Figure 1.34a). Three modes, theoretical (fluctuation), and experimental (B-factor) results from GNM were investigated closer to further evaluate the intrinsic residue fluctuations of the aspart structure (Figure 1.34b-c). The theoretical fluctuations calculated from the three modes are aligned to the B factors in Figure 1.34c. Accordingly, the residues around position 17th elucidate the regional flexibility of the C-terminus of aspart chain B compared to lispro dynamics (Figure 1.34c). We also performed a docking analysis to compare the interaction of regular insulin vs. aspart with the insulin receptor (Figure 1.35).

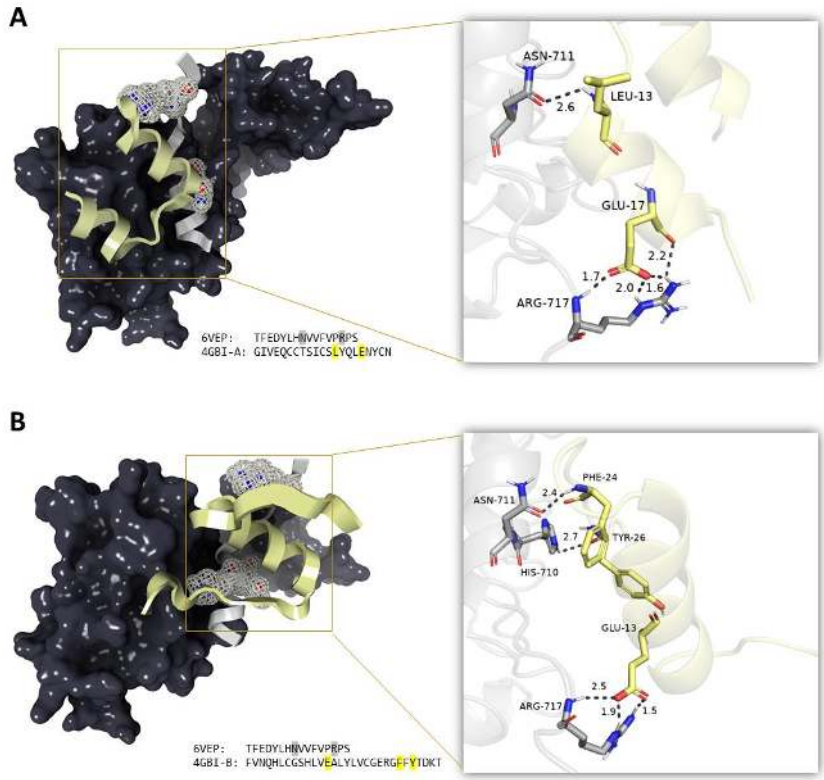


Figure 1.35 Insulin aspart and insulin receptor interaction.

Docking the aspart (PDB_ID: 4GBI) to the apo-state IR structure (PDB_6VEP) containing the αCT domain (colored in gray) that is critical for insulin binding and L1 domain (colored in black). Interaction with insulinactive residues reported in previous studies (Lukman et al., 2015) has succeeded. The Aspart chain B (A) and/or chain A (B) are excluded for a better view. Interacted residues of αCT domain with the aspart have displayed the mesh mode in the Protein Imager server.

Accordingly, Leu13, Glu17 (Figure 1.35a, pale-yellow), and Glu13, Phe24, Tyr26 (Figure 1.35b, paleyellow) residues of aspart insulin have interacted with Asn711, Arg717 (Figure 1.35a, gray), and His710, Asn711, Arg717 (Figure 1.35b, gray) residues of the IR-αCT segment.

Insulin glulisine

Insulin glulisine, Apidra (Sanofi-Aventis Company), mimics the function of endogenously produced human insulin and is similar to the lispro and aspart action. Its effect begins within 15 minutes, its peak levels occur 30 to 90 minutes after injection, and the duration of action is approximately 5 hours. It is also known as "bolus insulin" as it provides high insulin levels in a fleeting time. This analog has been produced as a biosynthetic and fast-acting insulin analog using recombinant techniques. This analog

works to improve glycemic control in children and adults with diabetes. Compared to native insulin, it has an asparagine residue modification in the position B3 replaced by lysine and B29 replaced by glutamic acid. This structural change causes a higher dissociation effect of glulisine than native insulin and stabilizes its monomer form. The action of insulin glulisine at the beginning is nearly 15 minutes. Its activity peaks occur 60 minutes after the injection and its duration of action is 2-4 hours. The median time to maximum concentration (Tmax) for regular insulin and aspart insulin analog is 60 minutes versus 120 minutes, respectively. As a result, this analogue has a faster rate of absorption than conventional insulin after injection. Typically, 0.15 U/kg body weight dose for type 1 diabetes is characterized by a maximum concentration (Cmax) of 83 mU/L versus 50 mU/L compared to regular insulin. Its absolute bioavailability is approximately 70%; this change depends on the injection area (thigh 68%, abdomen 73%, deltoid 71%), and its half-life after injection is 42 [Garnock-Jones and Plosker, 2009]. We also performed GNM analysis to compare the dynamics of regular insulin and insulin glulisine structures obtained from the PDB as a bird's eye view. Accordingly, cross-correlations between residue motions of the glulisine structure were analyzed with mean square fluctuations characterized by the flexibility of the residues (Figure 1.36).

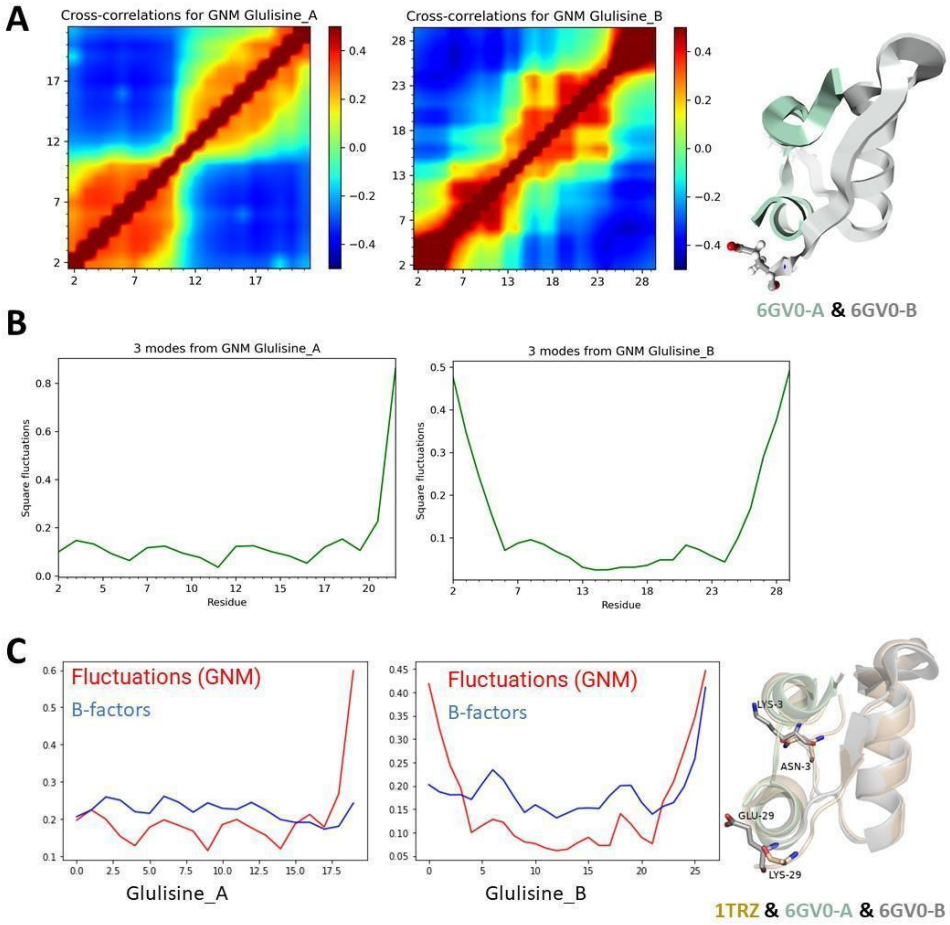


Figure 1.36 General analysis of glulisine.

Cross-correlation and squared fluctuations from Gaussian Network Model (GNM) analysis. In the cross-correlation result from the GNM analysis, the color range was determined as -0.5 and 0.5 to ensure compatibility between analogs' results (A). According to the squared fluctuations obtained from the weighted 3 slowest modes, global motion and residual fluctuations are relatively stable (B). While the figures show the correlation between the theoretical (GNM results; fluctuations from GNM) and the experimental (B-factors), lines show only the residue fluctuations (squared fluctuations even) (C). Glulisine (6GV0) is generated using the Protein Imager server. Superposition of the glulisine (6GV0) and human insulin (1TRZ) was generated by PyMOL software (Root-mean-square deviation (RMSD): 0.42).

Similarly, the intrachain residues' cross-correlation displayed that the first ten residues and 13-20th residues had highly correlated motions in chain A as well. In addition, chain B was observed with higher fluctuations at the 2-7th, 11-13rd, 17- 21st, 23-25th, and 27-30th positions (Figure 1.36a). Three modes, theoretical (fluctuation), and experimental (B-factor) results from GNM were investigated closer to further evaluate the intrinsic residue fluctuations of the aspart structure (Figure 1.36b-c). The theoretical fluctuations calculated from the three modes are aligned with the B factors in Figure 1.36c.

Accordingly, the residues around position 17th elucidate the regional flexibility of the C-terminus of aspart chain B compared to lispro dynamics (Figure 1.36c). We also performed a docking analysis to compare the interaction of regular insulin vs. aspart with the insulin receptor (Figure 1.37).

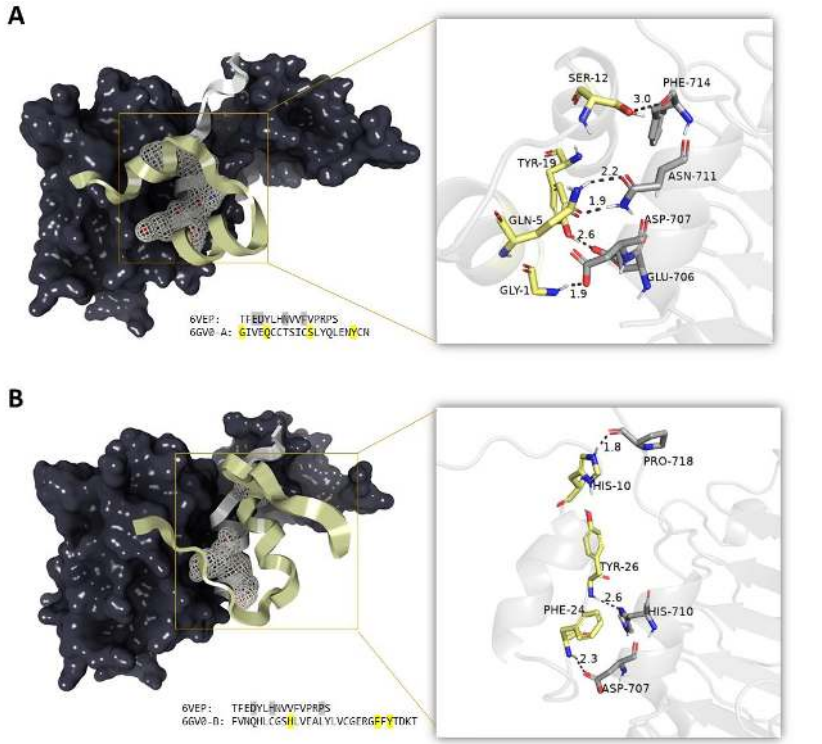


Figure 1.37 Insulin glulisine and insulin receptor interaction.

Docking the glulisine (PDB_ID: 6GV0) to the Apo-state IR structure (PDB_6VEP) containing the αCT domain (colored in gray) that is critical for insulin binding and L1 domain (colored in black). Interaction with insulin-active residues reported in previous studies (Lukman et al., 2015) has succeeded. The glulisine chain B (A) and/or chain A (B) are excluded for a better view. Interacted residues of αCT domain with the glulisine have displayed the mesh mode in the Protein Imager server.

Accordingly, Gly1, Gln5, Ser12, Tyr19 (Figure 1.37a, pale-yellow), and His10, Phe24, Tyr26 (Figure 1.37b, pale-yellow) residues of glulisine insulin have interacted with Glu706, Asp707, Asn711, Phe714 (Figure 1.37a, gray), and Asp707, His710, Pro718 (Figure 1.37b, gray) residues of the IR-αCT segment.

1.8.2. Long-acting analogs

There are several types of long-acting insulin formulations for commercial use, such as insulin glargine, insulin detemir, and insulin degludec. All analogs are designed to provide stable, relatively flat, and long-term basal insulin levels compared to the regular insulin formulation [Rosskamp and Park, 1999].

Insulin glargine

Insulin glargine Lantus (Sanofi-Aventis) mimics the function of endogenously produced human insulin (Figure 1.38a). This analog works to improve glycemic control in children and adults with diabetes. It has a duration of effect of up to 24 hours and is suitable for dosing once a day before going to bed. Glargine can be considered "basal insulin" in terms of its potency. Therefore, low insulin concentrations are sufficient, as it maintains blood sugar balance at night or between meals. Typically, long-acting analogs are combined with "bolus insulin" as they mimic the endogenous insulin function of the pancreas. This analog has been reformulated by Sanofi under the name Toujeo. It is more concentrated (300IU / ml) than Lantus, which contains 100IU / ml of product. Its effect starts within 6 hours and the duration of the effect is approximately 30 hours. This analog is produced by rDNA techniques. An asparagine characterizes the formulation at the 21st position replaced with glycine as well as the addition of two arginine residues at the 31st and B32nd positions. This product dissolves at pH 4.0, forming microprecipitates at physiological pH 7.4. It allows small amounts of insulin glargine to be released slowly, giving the product a long duration of action, thus allowing it to mimic basal insulin levels in the body. Its modifications allow the isoelectric point to shift to a neutral pH and stabilize it in acidic conditions relative to normal insulin. After injection, this analog is less soluble due to the body's neutral pH, allowing insulin to be released slowly as microprecipitates. This provides a relatively constant concentration over 24 hours compared to native insulin. It also contains small amounts of zinc to further delay absorption. The onset of action is 1.1 hours and its pharmacokinetic profiles for 0.4, 0.6, and 0.9 U/kg Toujeo doses in T1D are characterized by 12- and 16-hour serum insulin concentrations. Finally, this analog is metabolized as two active metabolites: 21a-Gly-des-30b-threonine insulin (M2) and 21a-Glyhuman insulin (M1) [Bolli and Owens, 2000; Dunn et al., 2003; McKeage and Goa, 2001]. Performed GNM analysis to compare the dynamics of regular insulin and insulin glargine structures obtained from the PDB as a bird's eye view. Accordingly, chain B of the glargine was observed with higher fluctuations at the 2- 7th, 19-23rd, and 27-30th positions (Figure 1.38a).

Insulin detemir

Insulin Detemir, Levemir (Novo Nordisk Company), mimics the endogenously produced human insulin (Figure 1.38b). This analog works to improve glycemic control in children and adults with diabetes. It has a duration of effect of up to 16-24 hours and is suitable for dosing once a day before going to bed. Because of its potency, detemir can be considered "basal insulin". Like glargine, low detemir concentrations are sufficient for diabetics as it maintains blood sugar balance at night or between meals. This analog is produced by yeast cells using rDNA technology. The formulation has a 14-C fatty acid called myristic acid attached to the lysine residue at the B29 position. The myristoyl moiety enhances albumin binding and self-assembly. Therefore, detemir can dissolve insulin slowly, so it provides a longacting effect. Its onset is 1 to 2 hours and has a 30% lower receptor binding affinity than native insulin. The Cmax is 6 to 8 hours after subcutaneous injection. The Cmax for 0.5 units/kg of this analog is 4.641 ± 2.299 pmol/L, and the apparent volume of distribution is about 0.1 L/kg. The median time to C max is 12 hours after injection and its absolute bioavailability is approximately 60%. More than 98% is bound to albumin and there is no clinically significant interaction between other protein-bound drugs or fatty acids and insulin detemir. The terminal half-life is 5 to 7 hours, depending on the dose after injection [Chapman and Perry, 2004]. Also performed GNM analysis to compare the dynamics of regular insulin and insulin detemir structures obtained from the PDB as a bird's eye view. Accordingly, chain B of the detemir was observed with higher fluctuations at the 2-10th, 17-20th, 22-24th, and 27-30th positions (Figure 1.38b).

Insulin degludec

Insulin Degludec, Tresiba (Novo Nordisk Company), mimics natural human insulin (Figure 1.38c). This analog works to improve glycemic control in elderly and 1-year-old patients with DM. It has an effective time of up to 42 hours and is suitable for dosing once a day before going to bed. This analog has an extra hexadecanedioic acid on the lysine residue at position B29, allowing multi-hexamers to form. After injection, this multi-hexamer conformation acts like a drug depot and releases the monomer rather slowly. This phenomenon offers a protracted-time action profile due to its delayed absorption from the tissue into the systemic circulation. Therefore, this analog provides

a lower peak than other long-acting insulin analogs. Cmax for 0.4 U/kg of this analog is 4472 pmol/L, and its Tmax is characterized by 9 hours. The average onset of appearance is about an hour and its concentration reaches steady-state levels after 3-4 days. More than 99% of insulin degludec is bound to plasma albumin and there is no clinically significant interaction between other protein-bound drugs and degludec. The terminal half-life is characterized by approximately 25 hours, regardless of the dose [Jonassen et al., 2012; Vora et al., 2015]. Also performed GNM analysis to compare the dynamics of regular insulin and insulin degludec structures obtained from the PDB as a bird's eye view. Accordingly, chain B of the degludec was observed with higher fluctuations at the 2-10th, 17-20th, 22-24th, and 27-30th positions (Figure 1.38c).

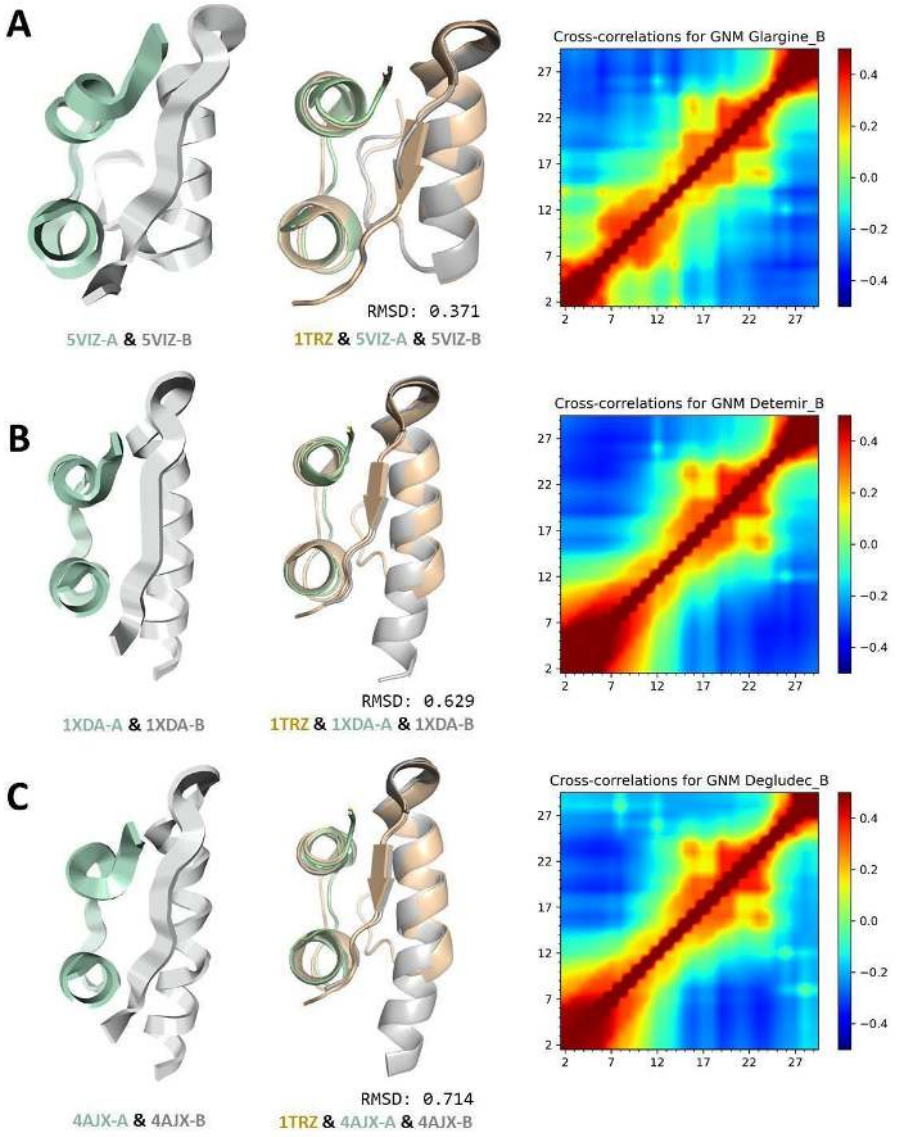


Figure 1.38 Detailed comparison of human insulin and long-acting analogs.

Insulin glargine (A) (PDB ID: 5VIZ), insulin detemir (B) (PDB ID: 1XDA), insulin degludec (C) (PDB ID: 4AJX). Long-acting analogs show similar cross-correlation with respect to the correlation between the monomers of glargine (A), detemir (B) and degludec (C). In the cross-correlation result from the GNM analysis, the color range was determined as -0.5 and 0.5 to ensure compatibility between analogs' results. All single structures are generated using the Protein Imager server. Superposition of the long-acting structures was generated by PyMOL software.

An in-depth GNM analysis was conducted to examine the intrinsic dynamics of regular insulin and commercially available insulin analogs obtained from the PDB (Figure 1.39). Accordingly, distinct correlations surfaced, particularly notable between residues

20-30 and 1-10 in insulin glulisine, showcasing a higher correlation relative to insulin lispro and aspart, similar to the behavior observed in human insulin. However, insulin lispro and aspart exhibited diminished fluctuation between these residues compared to insulin glulisine, implying a lesser stability and a more rapid-acting nature for insulin lispro and insulin aspart compared to insulin glulisine. Furthermore, an observation emerged between residues 46-50 and 1-5 in insulin lispro, revealing a higher correlation compared to insulin aspart, suggesting that insulin aspart is less stable and exhibits a more rapid-acting behavior compared to insulin lispro and insulin glulisine. Moreover, a collectivity analysis of slow modes was undertaken for each rapid-acting analog. The degree of collectivity in a given mode is a metric for evaluating the extent to which structural elements move cohesively. Notably, high collectivity indicates a highly cooperative mode involving a substantial portion of the structure, if not the entirety. In contrast, low collectivity is associated with modes affecting small or localized regions. Modes characterized by a high degree of collectivity are of particular interest due to their functional relevance and are typically found at the lower frequency end of the mode spectrum. In alignment with the cross-correlation analysis, insulin aspart, among the rapid-acting insulin analogs, exhibits lower collectivity and cooperativity, indicative of a less cohesive structural response compared to other rapid-acting analogs (Figure 1.39a). A parallel analysis was conducted for long-acting insulin analogs. Consequently, a relatively higher correlation was observed between residues 35-45 and 15-20, as well as residues 46-50 and 1-5, in insulin detemir and insulin degludec compared to insulin glargine and human insulin. However, insulin glargine exhibited intrinsically cooperative behavior in contrast to detemir and degludec, specifically between residues 20-30 and 1-10. This observation affirms that glargine is a sequence-based engineered analog, unlike detemir and degludec, which are formulated-based engineered analogs. On another note, collectivity analysis revealed that long-acting insulins demonstrate a higher level of collectivity and cooperativity compared to rapid-acting analogs, aligning with findings in the existing literature regarding the behavior of these analogs (Figure 1.39b)

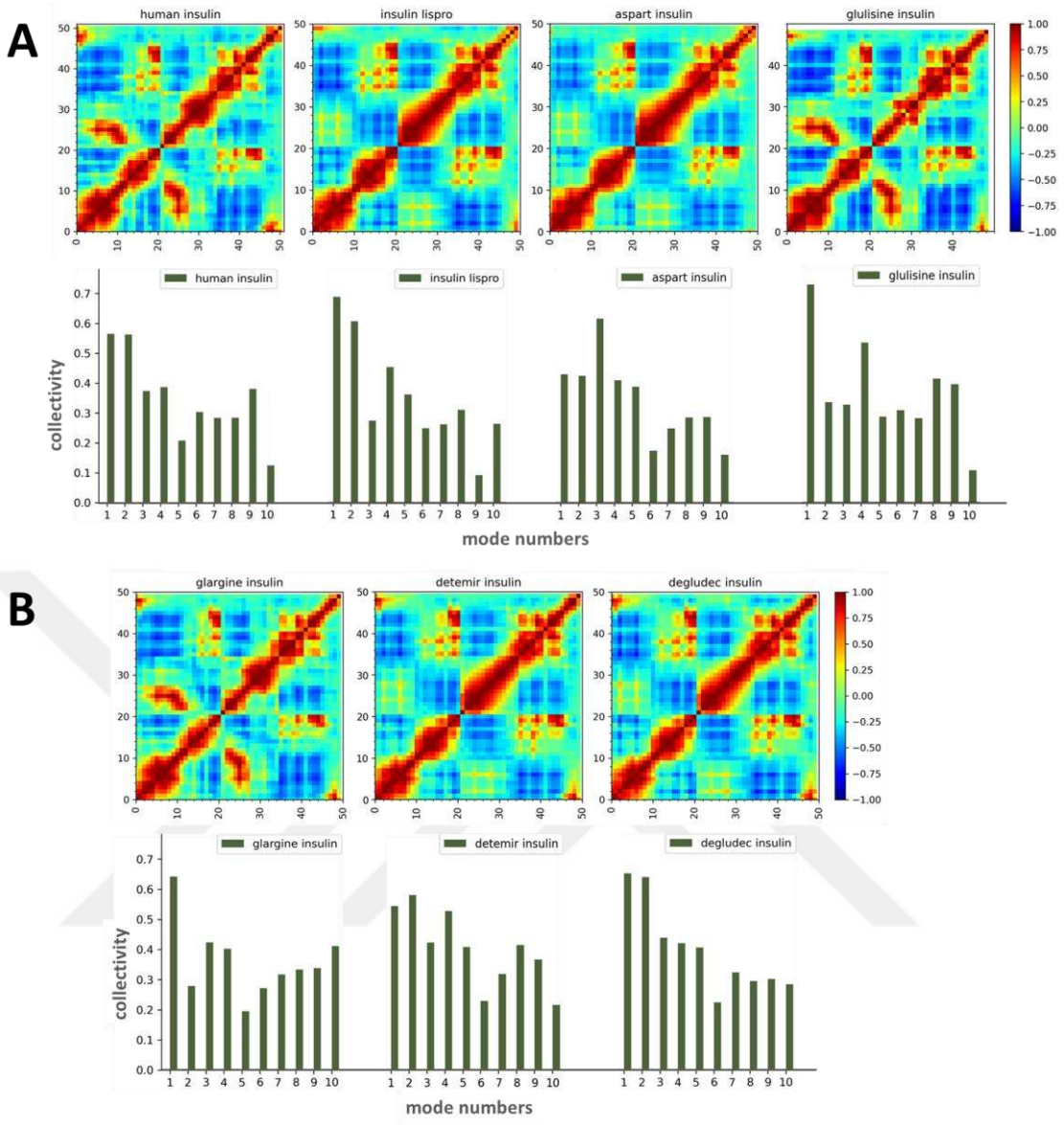


Figure 1.39 Detailed examination of the thorough intrinsic dynamic comparison using GNM analysis.

Panels (A) and (B) illustrate cross-correlation and collectivity analysis of slow modes for the rapid-acting and long-acting insulin analogs, respectively, compared with human insulin.

1.8.3. Premixed insulin analogs

There are three main types of prefixed insulin analog formulations: 50% insulin lispro with 50% insulin lispro protamine suspension, 25% insulin lispro with 75% insulin lispro protamine suspension, and 30% insulin aspart protamine suspension with 70% insulin aspart protamine. The purpose of these formulations is to minimize the errors of some mixed insulin combinations made by patients. Therefore, the insulin

regimen can be simplified and the number of injections per day can be reduced. Typically, the regular premixed insulin effect is approximately 0.5 to 2 hours, and the effect lasts up to 24 hours. However, the effect of premixed insulin analogs is about 15 minutes, and their biological activity is highest at 1 to 4 hours. Similar to regular human insulin premixes, premixed insulin analogs also last up to 24 hours. In addition, the pharmacokinetic profiles of pre-mixed insulin analogs are characterized by less individual variability compared to premixed regular human insulin [Wu et al., 2015].

1.9. Recombinant insulin production method

Typically, insulin production is characterized by three main approaches in terms of novel techniques (Figure 1.40).

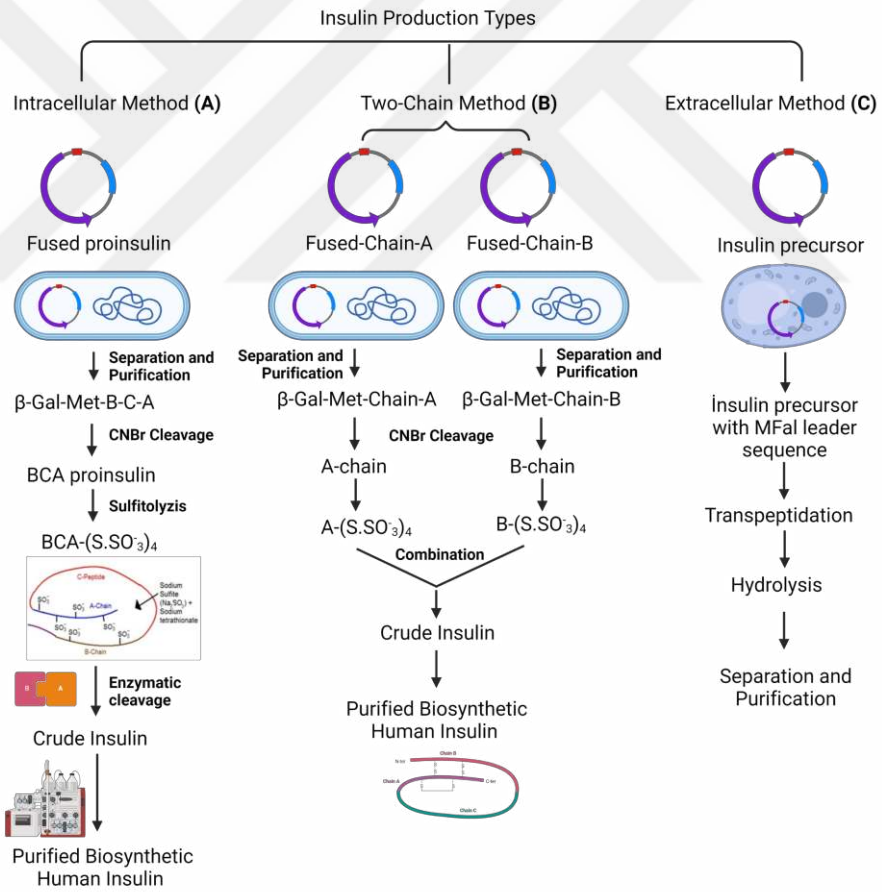


Figure 1.40 Comparison of main recombinant insulin production methods.

Two of these techniques involve using the organism *Escherichia coli* (Figure 1.40a-b) [Chance et al., 1981; Zimmerman et al., 2010], while the third method is to use

Saccharomyces cerevisiae (Figure 1.40c) [Chan et al., 1981] to get the insulin precursor in the medium. When the first two methods are preferred, either the precursor is produced using the large fusion protein in the cytoplasm [Thim et al., 1986] or the use of a signal peptide to release the product into the periplasmic space is preferred [Chan et al., 1981]. In the two-chain method, insulin is produced in a way that the A and B chains are separate. These strands are produced as a form of the fusion protein in *E. coli* culture containing a vector carrying the interested DNA. After cleavage of the fusion protein using cyanogen bromide (CNBr), the A and B chains are sulfonated by reacting along the sulfonated A chain with the sulfonated B chain. Finally, these products are purified to obtain insulin. Such a method has many disadvantages, such as requiring two fermentation processes and additional steps to prepare the sulphonated B chain and A chain. Hence, this results in low insulin yield [Chan et al., 1981]. In the intracellular proinsulin method, insulin is produced in *E. coli* as a proinsulin precursor in the form of the fusion protein. Similar to the two-chain method, proinsulin is cultured in *E. coli* containing a vector carrying DNA of interest. After cleavage of the fusion protein using the CNBr, proinsulin is obtained. After the precursor is separated in the form of sulfonated proinsulin, it is refolded to form native disulfide bonds using trypsin and carboxypeptidase B. Finally, the resultant product is purified to obtain insulin. However, in this method, the yield of proinsulin decreases sharply during the folding of the disulfide bonds. The causes may be misfolding of the protein or some polymerization problems during the process. Therefore, the additional laborious purification steps resulting from these problems can be cited as disadvantages of such a method [Chan et al., 1981]. The extracellular insulin production method is different from the others in terms of using Saccharomyces cerevisiae cells. This approach involves production, purification, reaction with enzymes, hydrolysis with acids, and purification of single-chain insulin. However, this process has an unacceptably low insulin production efficiency [Chan et al., 1981]. Until now, there are many new approaches that have been tried to produce clinically useful insulin. Some are patented and most continue to be patented, while others are presented as research articles. The remaining part of the article will discuss the production of both patented insulin and optimized insulin.

1.10. Academic survey of patented insulin analogs and production methods

In 1992 and 2001, various insulin analog formulations were extensively studied and patented [Chance et al., 1992; Brange et al., 1989; Chance et al., 1997; Ertl et al., 2001; Berchtold, 2007; Zimmerman et al., 2012; 2012; Chan et al., 2015; Li et al., 2015; Zimmerman et al., 2012; Loos et al., 2015; Ortigosa et al., 2015; Grant et al., 2016; Ortigosa et al., 2017; Borowicz et al., 2017; Qian et al., 2014]. In 2007, an insulin analog in which the asparagine (Asn) at position B3 of the B chain had been substituted to a basic amino acid residue and an amino acid residue at the B29 position of the B chain had been substituted to a neutral/acidic amino acid residue was crystallized by a novel approach. In this crystallization optimization, crystals with A = 81.5 Å and C = 33.3 Å cell units were found in the R3 space group. After the amorphous powder of this analog was dissolved in a suitable zinc-free liquid, this analog was crystallized for further processing involving their precipitation, isolation, and drying. The validity of this patent is active until 2024 [Berchtold, 2007]. In 2011, introduced the Aspart modified proinsulin and its production methods. In this approach, the sequence Ri-(Bi-B26)-B27- B28-B29-B3o-R2-3-X-R4-R5-(Ai-A2o)-A2i-R6 has been optimized and described in detail. Furthermore, in the method, *E.coli* was transformed by His Tagged Aspart proinsulin inserted pTrcHis2A(Kan) plasmid. After the inclusion bodies were dissolved in reducing and chaotropic agents at pH 11.8- 12, this analog was recovered for other processes through folding, enzymatic cleavage, purification, and elution [Zimmerman et al., 2012]. In 2012, the same team introduced the Glargine proinsulin and its production methods. In this approach, the sequence R1-(B1-B29)-B30-R2-R3-X-R4-R5-(A1-A20)-A21-R6 has been optimized and described in detail. In addition, in the method, *E.coli* was transformed by a His Tagged Glargine proinsulin inserted pTrcHis2A(Kan) plasmid. At a similar time, a long-acting insulin glargine analog (1 mg/ml to 20 mg/ml) was developed. This analog involves zinc, magnesium, copper, phenol, m-cresol, and calcium ions at pH from 3.5 to 4.5. After the dissolution of the inclusion bodies in chaotropic agents, this analog was recovered for other processes through folding, purification by metal affinity chromatography, protection of a Lys residue with protective compounds, enzymatic cleavage with an intermediate solution, followed by purification by chromatography column(s) [Zimmerman et al., 2012]. Likewise, around 2013, a long-acting insulin glargine analogue was patented as well [Chan et al., 2015]. In the same year, the purification method of detemir was patented.

Accordingly, this technique involves equilibrating the column with an acidic solution, adjusting the pH below 5, adding 0.5-1.0 M buffer solution to the column at pH: 2-4, and washing. After collection, the main peak is obtained so that insulin detemir is purified. This technique has some advantages, such as large sample loading, convenience of operation, and high recycling efficiency. Additionally, this method is suitable for industrial production, which is economical, environmentally friendly, and cost-saving [Li et al., 2015]. In 2015, the same team introduced the Lis-Pro modified proinsulin and its production methods. In this approach, the R1 -(B1-B27)-B28-B29-B30- R2-R3-X-R4-R5-(A1-A21)-R6 sequence was optimized and described in detail. Furthermore, in the method, E.coli was transformed with plasmid pTrcHis2A(Kan). After the inclusion bodies were dissolved in reducing and chaotropic agents at pH 11.8-12, this analog was recovered for other processes by folding, enzymatically cleaving, purification, and elution [Zimmerman et al., 2012]. In the same year, an aqueous pharmaceutical formulation of glulisine was introduced at concentrations of 200 U/mL to 1000 U/mL in conditions with or without zinc (20 g/mL). Additionally, adjustment of buffer concentrations was formulated with the amount of phosphate, glycerol, trometamol, nonionic surfactant, phenol, and/or m-cresol at pH 3.5 with 9.5 [Loos et al., 2015]. Moreover, in the same year, insulin and its analogs were purified by high-pressure liquid chromatography and reverse-phase chromatography. This approach was carried out in an acidic cation exchange medium in the presence of an water-miscible organic modifier at elevated temperatures [Ortigosa et al., 2015]. Likewise, the purification process of insulin and insulin analogs was developed by two tandem microfiltration (MF) methods. Such a method reduces both soluble and insoluble impurities as well as provides a suitable buffer for the downstream purification process of the insulin product [Grant et al., 2016; Ortigosa et al., 2017]. In about 2016, a method was developed for dissolving and refolding precursor insulin or its analogs from the inclusion body for use in the production of insulin and its various analogs [Ortigosa et al., 2017]. In the same year, insulin, insulin derivatives, and hybrid peptide production methods were introduced. After the Xn-B-Arg-Arg-A formula was optimized by some steps, some insulin and its derivatives, such as insulin, lispro, GR, insulin, SR, insulin, SK3R insulin, and a type of proinsulin, were characterized [Borowicz et al., 2017].

1.11. Overview of insulin analogs and production methods in the literature

In 2014, human proinsulin production in the milk of transgenic animals was optimized using the goat β-casein promoter. The efficiency of this proinsulin production was about 8.1 g/L. It was observed that there was no problem with the biological activity of insulin after the conversion of proinsulin obtained from milk into mature insulin. Accordingly, it can be suggested that insulin production can be made in dairy milk [Qian et al., 2014]. In 2015, large-scale refolding of human preproinsulin and a new refolding process to produce human insulin were optimized. It has been stated that a high protein concentration is achieved with this approach. This approach paves the way for the large-scale production of similar biologically active proteins [Kim et al., 2015]. In 2016, plantbased production of proinsulin analogs was optimized using Peanut seeds. Accordingly, it has been observed that some transgenic peanuts express proinsulin successfully. It was also revealed that peanut seeds act like insulin storage areas [Ling et al., 2016]. In 2018, the molecular chaperone aB-crystalline (aBCry) had been preferred as a fusion protein in insulin production. In this way, novel and sensitive insulin production and purification steps were optimized. In this study, in which the two-chain method was preferred, pET28b(+) was preferred as the vector, and *E.coli* was used as the host, providing high yield and proper folding of the enzyme [Akbarian et al., 2018]. In 2019, a new bacterial host and a compatible expression vector were optimized. Accordingly, the efficiency of this recombinant human insulin has been observed to be relatively high. In this study, an alternative method for producing recombinant human insulin is introduced [Zieliński et al., 2019]. In 2020, PCRbased insulin production was optimized by overexpressing *E.coli* at a lab scale. The yield of the product was 520.92 mg/L. After insulin detection by MALDI-TOF-MS and LCMS, this product was found to be 100% similar to human insulin. This study eliminated the use of affinity tags, body recovery steps, and enzymatic cleavage of the c-peptide processes [Govender et al., 2020].

1.12. Challenges and future directions

Insulin has been a molecule characterized by diabetes treatment for nearly a century. However, although therapeutic strategies appear to be evolving, the approaches are the same: diabetes-focused injectable therapies. Perspectives concentrated solely on

glucose management remain nonpragmatic conventional modalities, thus suppressing the risk of life-threatening hypoglycemia rather than treating it. These perspectives include insulin alternatives such as microreceptors, aptamers, non-insulin peptides, and antibodies and are found in the literature as approaches whose optimization is not fully established and requires validation. Diabetes is not a disease by itself, but rather a critical phenomenon that paves the way to the development of many insulin-related disorders. Diabetic chronicity only occurs in a variety of scales, from microvascular diseases to macrovascular diseases to the diabetic axis, while current treatments are only designed to minimize these symptoms. The last point reached in insulin production has progressed to supplying high purity insulin with state-of-the-art injection devices. However, insulin production and insulin therapy alone do not provide sufficient results in diabetes management. Medical advances have focused on advanced therapeutic approaches that take into account intermediate measures such as body weight, lipid and glucose, and disease mortality data from a holistic perspective rather than employing insulin-based improvements. Examples in this context include implantable mini-insulin pumps that control glucose with a closed-loop system that simulates pancreatic function, as well as personalized diabetes management approaches with artificial intelligence integrated sensors. In addition to these groundbreaking studies with ongoing optimization, the continuous intravenous insulin infusion method and basal bolus insulin strategies have already occurred in clinical studies as viable methods that begin to mimic holistic and personalized approaches, and such methods have given successful results so far. In addition, stem cell and tissue engineering technologies that pave the way to a nontechnical but biological axis for the same goal are among these approaches, including companies that play an active role in β-cell restoration. Such approaches do not only consist of therapies based on insulin/analogs but also make the treatment healthier by approaching the disease from a holistic and personal perspective. In many ways, adult-onset diabetes has a more complex system than juvenile-onset diabetes. This is due to the complicated symptoms that adult-onset patients have over a wider age range. The absence of a genetic predisposition in adults in the adult-onset condition indicates that the disease is triggered by environmental factors. The most critical example of this is obesity. Such environmental factors and diseases experienced by diabetic patients are closely related to the development of diabetes. This is the evidence of why the prevalence of insulin-related diseases mentioned earlier is so high. Therefore, when dealing with insulin, considering only diabetes is an inadequate

interpretation of the scenario and applying insulin therapy as the treatment alone is not sufficient. An example of this situation is that an obese diabetic patient who suffers from weight loss only with surgical intervention often encounters irreversible complications. Individuals who do not want to experience this situation will only turn to obesity-focused drugs. In both cases, since a holistic modality is not applied to the disease, the result will be to suppress symptoms of the disease with only temporary solutions. A similar situation is valid for nonclinical studies. Focusing on a particular point in the studies will lead to the evaluation of insulin, which is a complex system in itself, the proteins it is associated with, the pathways it triggers, and the diseases it causes in a narrow window. Science has a cumulative nature and progress must be maintained from the previous step. In this way, even these changes in injection techniques, from subcutaneous injections to oral insulin applications, from diabetes diaries to smart insulin pumps, have been developed by approaching insulin on the basis of holistic and personalized medicine. All holistic information about insulin obtained in structural, biological, omics, and clinical axes facilitate the path to be taken in treatment.

1.13. Conclusion

Insulin has played a central role in our understanding of the underlying mechanisms of metabolic diseases and in the advancement of peptide chemistry, biochemistry, structural biology, and omics strategies. The diabetes global epidemic has led to a dramatic increase in insulin demand. Over time, insulin-oriented optimizations have progressed in correlation with advances in rDNA biosynthesis and synthetic chemistry, and high purity insulin production has been successfully achieved. In this way, analogs that provide advanced basal glucose control have been designed, and robust optimization has been achieved in insulin therapy. Although insulin therapy is an established approach clinically, it has not been able to achieve complete integrity from the metabolic perspective. Therefore, it lacks a personalized strategy. As a result, this deficiency has remained closely related to the development of hypoglycemia with the usage of insulin. Even though the aim of insulin therapy is to eliminate diabetes, any treatment that lacks the axis of holistic and personalized medicine is destined to suppress symptoms of the disease rather than eliminate it. Therefore, more functional strategies are needed in the diagnosis, preventive treatment, and research studies to

eradicate insulin-dependent diseases. In this age, where technology has made significant progress compared to the past, making holistic and personalized medicine-based strategies a part of research and clinical studies will enable us to step into a new era where we can offer solutions from a thousand different perspectives.

`

Chapter 2

Exploring Domestically Produced Ultra-Acting Insulin

According to data from the World Health Organization, diabetes cases have exhibited a notable global increase over time. In 1980, the global diabetic population was 108 million, escalating to 422 million by 2014. While the worldwide prevalence of diabetes in adults aged 18 and above was 4.7% in 1980, it surged to 8.5% in 2019. The 10th edition of the IDF Diabetes Atlas underscores a sustained global ascent in diabetes prevalence, highlighting its continued significance as a worldwide concern for the health and well-being of individuals, families, and communities. Presently, one in every ten adults worldwide contends with diabetes. It's predicted that the number of people with diabetes will go up to 643 million by 2030 and 783 million by 2045. Three out of four adults with diabetes live in countries with lower or medium incomes. In 2021, someone in the world died from diabetes every five seconds. Over the last 15 years, dealing with diabetes cost at least $966 billion more, with healthcare spending going up by 316%. Also, 541 million adults have Impaired Glucose Tolerance (IGT), making them more likely to get type 2 diabetes. In the Middle East and North Africa, one out of every six adults, around 73 million people, has diabetes, and one out of every three adults with it hasn't been diagnosed. In 2021, 796,000 people around the world died from diabetes. As of 2019, Türkiye has about 6.6 million people dealing with diabetes, with 2.6 million still undiagnosed. Moreover, around 4 million individuals aged 20-79 fall into the Prediabetes category (Impaired Glucose Tolerance). In this regard, Türkiye spends around $1,404 per person annually for managing diabetes in the age group of 20-79. According to the provided estimates, Türkiye is anticipated to be among the top 10 countries globally with the largest diabetic population by 2035 [IDF, Diabetes Atlas, 2019]. Türkiye's pharmaceutical market provides a range of insulin formulations: rapid-acting, short-acting, intermediate-acting, long-acting, pre-mixed human insulins, and pre-prepared biosimilar insulins. According to an extensive statistical analysis, the total

sales revenue for these medications amounted to 554,515,557 Turkish Lira in 2014 [Dal and Sezer, 2017]. Despite the challenges posed by current insulin formulations in diabetes management, impeding insulin absorption for patients and resulting in elevated costs, there is a subsequent increase in the risk and occurrence of diverse diabetes-related complications [Watkins et al., 1990]. However, although official insulin production from scratch has not been introduced in Türkiye, the global expiration of patents on specific insulin formulations has eased the progression of biosimilar insulin development. Advancements in diabetes pharmacotherapy include rapid-acting insulin analogs and recombinant human insulin, mirroring the natural hormone and exhibiting short-acting or intermediate-acting characteristics based on the formulation. Rapid-acting insulins offer advantages over regular insulin, yet it is necessary to formulate quicker-acting insulins that closely emulate physiological insulin release. Ultra-rapid-acting insulin analogs mark a novel development in insulin therapy, introducing formulations aiming for an even swifter onset of insulin action to minimize the delay between subcutaneous insulin injection and the initiation and peak of insulin activity. This chapter aims to present innovative designer ultra-rapid-acting insulin analogs, *esrapid*, produced in hexameric and monomeric forms, respectively, employing the recombinant DNA technique. The comprehensive exploration will systematically elucidate: (i) the upstream process and optimization, (ii) the downstream process and optimization, (iii) crystallization, data collection, and data processing, and (iv) the experimental and computational **monomeric** kinetics associated with these analogs. Subsequent chapter (Chapter 3) will cover the hexameric-dihexameric kinetics of commercially available insulin analogs.

2.1. The upstream process and optimization of ultra-acting analogs: "esrapid"

First *esrapid* analog has been designed as a rapid-acting insulin analog, where the positively charged asparagine at position 21 in the peptide chain is substituted with the uncharged glycine. The formulation of the product includes 20 mM citric acid. In its monomeric state, the product adopts a conformation that is readily poised for direct binding to the receptor. The monomeric form is subjected to crystallization in a 20 mM citric acid solution with a pH of 3. The small side chain of glycine contributes to the formation of a stable structure under acidic conditions, thereby mitigating rapid degradation of the molecule during injection or upon contact with bodily fluids. Second

esrapid is engineered as an ultra-acting insulin analog [Hall, 1998]. At position 22 in the B chain, the rigid arginine is substituted with lysine, and at position 28, the uncharged proline is replaced with the charged aspartic acid. Following dissociation, the analog transforms into a monomer, where the alterations in amino acids contribute to its rapid efficacy in both the monomeric and hexameric states. The lysine side chain carries an amino group and a positive charge, akin to arginine. However, lysine exhibits limited hydrogen bonding capacity and a less reactive side chain compared to arginine. Furthermore, the more flexible structure of lysine, compared to arginine, is believed to influence the chemical stability of the *esrapid*^{2} analog [Kakade and Liener, 1969]. Taking these considerations into account, this *esrapid*^{2} is designed to be an ultra-acting insulin analog.

Various methodologies have been employed to express recombinant insulin in *E. coli*, encompassing diverse expression techniques, tags, and host strains (Khalilvand et al., 2022). A notably efficient approach involves the design of insulin constructs featuring a protease cleavage site, leveraging in-house proteases to facilitate the cost-effective synthesis of mature insulin [Akbarian et al., 2018]. Furthermore, host strains can be genetically engineered to augment insulin expression, exemplified by the utilization of Rosetta™ 2 host strains, BL21 derivatives optimized for the expression of eukaryotic proteins harboring rare codons [Tegel et al., 2010]. These strains bear a compatible chloramphenicol-resistant plasmid encoding tRNA genes for seven rare codons (AGA, AGG, AUA, CUA, GGA, CCC, and CGG) under their native promoters (Tegel et al., 2010). Induced by IPTG, such strains prove conducive to the high-yield production of insulin from target genes cloned into pET vectors [Pan and Malcolm, 2018].

2.1.1. Plasmid selection and design for "esrapid"

The overexpression of *esrapid* insulin analogs utilized pET24a (+) and pET28a (+) expression vectors, each incorporating a hexahistidine tag (6 x His) and a trypsin cleavage site (-GGR) (Figures 2.1 and Figure 2.2). Codon-optimized genes were synthetically manufactured and subsequently integrated into these vector plasmids through cloning. Genscript (USA) executed the gene cloning into the vectors, providing the amino acid sequences of these genes for construct design. The 6 x His tag,

thrombine cut site and trypsin recognition site within the amino acid sequences were highlighted in red, blue, and green respectively. (Mutated sections are highlighted by yellow color)

esrapid, pET24a (+):

MHHHHHHGGR|FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVEL GGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCG

esrapid[2], pET28a (+):

MGSSHHHHHHSSGLVPR|GSGGR|FVNQHLCGSHLVEALYLVCGEKGFFYTDKTRR EAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN

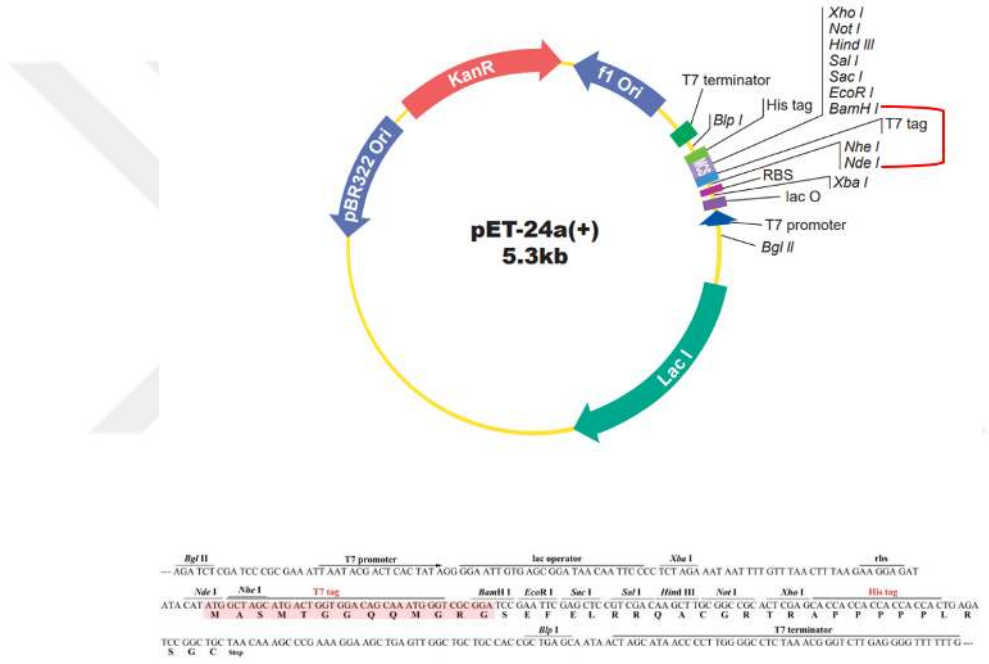


Figure 2.1 Vector map and sequence information of pET-24a (+) plasmid.

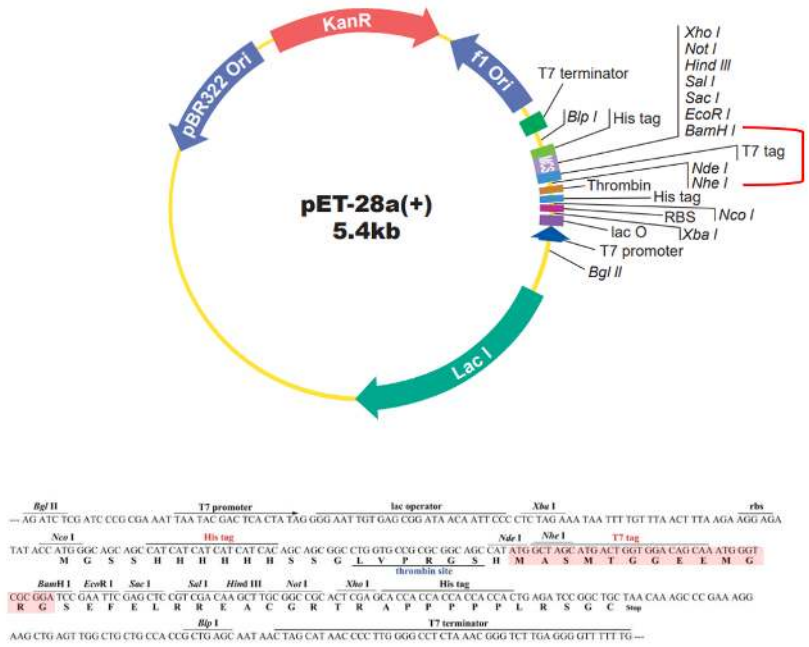


Figure 2.2 Vector map and sequence information of pET-28a (+) plasmid.

2.1.2. Transformation and preparation of bacterial stocks

The gene constructs, cloned onto the pET-28a(+) and pET-24a(+) vectors, underwent transformation using the heat shock method and were introduced into the *E. coli* BL21 Rosetta 2 bacterial strain. Subsequently, the transformed cells were cultured on LB agar plates supplemented with kanamycin and chloramphenicol antibiotics. The plates were incubated at 37°C overnight, and selected colonies were utilized to inoculate 150 mL LB-Miller broth media (30 g/L) containing antibiotics. The bacterial culture was grown at 37°C with agitation at 110 rpm in a New Brunswick Innova 4430R shaker/incubator overnight. For the prolonged preservation of the transformed bacterial culture, 750 µL of the cultured cells were mixed with an equal volume of 80% glycerol and stored in 2 mL Wuxi NEST Biotechnology cryovials at -80°C.

Small scale expression and optimization of "esrapid" analogs

Initially, samples of cells containing the standard human insulin construct (pET28b (+), IDT, USA) were used as a control in various medium concentrations (batch 1: 1g tripton, 0.5g yeast, 1g NaCl; batch 2: 1g tripton, 0.5g yeast, 1g NaCl, 1g glucose; batch 3: 1g tripton, 0.5g yeast, 1g NaCl, 2g glucose; batch 4: 1g tripton, 5.6g

yeast, 7g glucose), supplemented with 30 µg/mL kanamycin and µg/mL chloramphenicol. In contrast, "esrapid" analog constructs were introduced into 150 mL nutrient LB-Miller medium (30 g/L) supplemented with 30 µg/mL kanamycin and µg/mL chloramphenicol. The cultures underwent overnight incubation in the New Brunswick Innova 4430R shaker/incubator, maintained at 37°C and 110 rpm. Subsequently, induction was carried out by adding 0.4 mM IPTG, and the cultures were further incubated for 3 hours at 37°C and 110 rpm in the New Brunswick Innova 4430R incubator. Expression analysis of optimized standard native insulin and intact esrapid analogs smaller than 15 kDa was confirmed by 20% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (Figure 2.3).

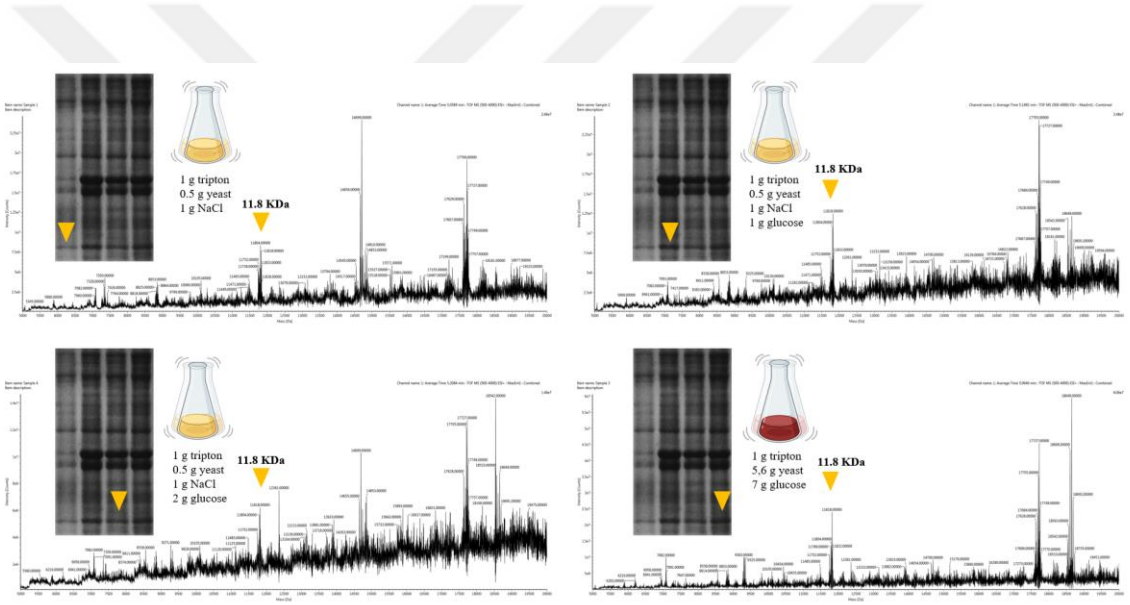


Figure 2.3 MALDI-TOF and SDS-PAGE results depict small-scale (150 mL) production optimization for native human insulins as controls.

Batch-1 (left top) exhibits weak expression, while batch-2 (right top) shows improved expression compared to batch-1. Batch-3 (left bottom) displays weak expression similar to batch-1, whereas batch-4 (right bottom) demonstrates better expression than the other batches in MALDI-TOF and SDS-PAGE analysis.

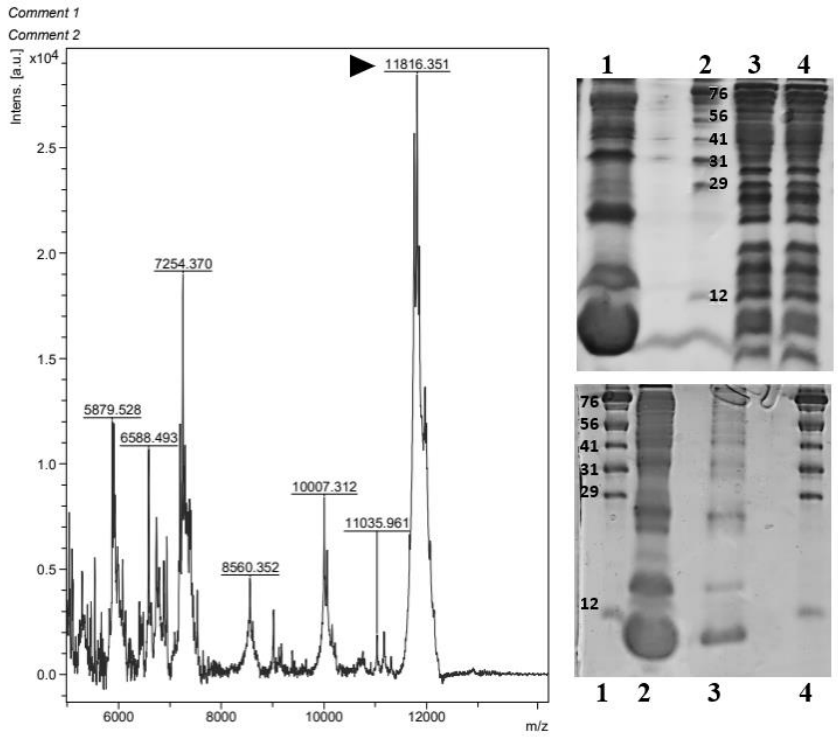


Figure 2.4 The outcomes of small-scale (150 mL) production for *esrapid* (1 & 2) insulin analogs are depicted through MALDI-TOF and SDS-PAGE results.

(left), *esrapid*[2] analog is indicated by the black arrow at 11.8 kDa. (right, top), *esrapid* analog is observed in the first well, with lane 2 serving as the marker, and lanes 3 and 4 representing conditions before the induction of *esrapid*. (right, bottom), *esrapid*[2] analog is displayed in the second well, where lane 1 functions as the marker, lane 3 represents conditions before the induction of *esrapid*[2], and lane 4 serves as the marker.

2.1.3. A pilot study for esrapid[2]

BioLector screening for esrapid production via response surface methodology

Throughout the process of recombinant protein production, the chemical and nutritional composition of the culture medium can significantly influence the growth of host cells [Shokri et al., 2003]. Critical factors such as carbon and nitrogen sources, metal ions, and medium pH are pivotal in shaping cell growth. Therefore, optimizing the culture composition becomes imperative to achieve high target protein yields [Kusuma et al., 2019; Nikerel et al., 2006]. However, the various factors influencing cell growth pose a challenge, and the One-Factor-At-A-Time (OFAT) approach proves time-consuming and labor-intensive in assessing their effects [Abu et al., 2017]. The OFAT approach must also consider potential dependencies and interactions between these factors [Abu et al., 2017]. In response to these limitations, the factorial approach emerges as a more comprehensive and efficient strategy for examining the collective impact of multiple factors on cell growth [Abu et al., 2017]. This approach, by simultaneously assessing all levels of all factors, facilitates the determination of their independent effects and interactions [Abu et al., 2017]. The Design of Experiment (DoE) serves as a statistical tool employing fractional factorial models, particularly Response Surface Methodology (RSM), to assess significant interactions among variables through a reduced number of experiments [Papaneophytou and Kontopidis, 2014; Sopyan et al., 2022; Dentener 2002]. Additionally, the BioLector XT high-throughput microbioreactor facilitates real-time monitoring of vital cultivation parameters for both aerobes and anaerobes, encompassing biomass, pH, dissolved oxygen in the liquid phase (DO), and fluorescence [Drummen; Osthege et al., 2022]. This tool offers expeditious and detailed insights into bioprocess development experiments, enabling more streamlined optimization of the culture medium and, ultimately, enhancing target protein yields [Osthege et al., 2022]. The objective of the this section was to optimize the culture media composition to enhance the protein expression of *esrapid* analog. We employed the Design-Expert (DoE) method, utilizing the BioLector XT system to support our investigations. By using these sophisticated approaches, we sought to identify a more comprehensive and efficient analysis of the various factors affecting the yield and growth of the *esrapid*; thus, we have gained real-time insights into crucial cultivation parameters, enabling us to identify the optimal

culture media composition to achieve maximum protein production of the *esrapid*. Accordingly, a screening experiment was conducted to assess the impact of various cultural components on the growth of Rosetta™ 2 (DE3) including *esrapid* gene, using the Plackett-Burman Design (PBD) and BioLector XT Microbioreactor. After that, the optimal culture media composition was determined through further experimental design utilizing the BioLector XT Microbioreactor. The consequential significant factors were subsequently optimized, employing the Response Surface Methodology (RSM) Central Composite Design (CCD) of DoE suite. After the optimization calculation was completed, the statistically relevant results were verified through experimental validation employing the BioLector XT Microbioreactor.

Media optimization

We employed the E. coli BL21 Rosetta 2 strain transformed with the pET28a(+) vector carrying the 'esrapid' insulin analog for growth and expression assessment. LB media, sourced from suppliers like Merck and Sigma, served as the foundational culture medium. Our internal resource, KUYBIIGM-Ladder, provided the protein weight marker. Designing and analyzing optimization experiments utilized the DoE 7.0.0 software package (Stat-Ease, Inc., Minneapolis, MN, USA). Screening experiments using the BioLector XT Microbioreactor (Brandd City, Country) with statistically determined parameters were conducted. Preliminary screening of various factors employed the PBD method, followed by the CCD method of RSM, to optimize significant variables. Experiments were executed in 2 mL volume wells within the Biolector XT Microbioreactor, each containing 1 mL of specific culture media. Cell growth occurred over 4 hours at 800 rpm and 37°C in 48-well FlowerPlates (mp2-labs, Baesweiler, Germany). Real-time monitoring of growth kinetics utilized scattered light intensities, and final insulin density, in grams per liter, was determined at 280 nm using a UV/Vis spectrophotometer. The result analysis included 20% SDS-PAGE.

Factorial design approach by Plackett-Burman Design

In the initial screening phase, we examined a total of eleven factors, including various nitrogen and carbon sources, pH levels, LB media concentration, as well as salts and metal ions (Figure 2.5) To organize the experiments, we utilized the Plackett-

Burman factorial design, involving sixteen experiments implemented through DoE software (Figure 2.5).

Std	Run	KCl (5-10mM)	$MgCl_2$ (0-15mM)	$MgSO_4$ (0-15mM)	Glycerol (0-5%)	Glucose (0-15mM)	IPTG (0.2-0.4mM)	pH (6.8-7)	LB (30-50g/L)	KH_2PO_4 (0-15mM)	Na_2HPO_4 (0-15mM)	Thiamine (0-10mM)	Insulin g/L
11	1	10	0	15	5	15	0,2	6,8	30	15	0	10	2,88
14	2	5	7,5	7,5	2,5	7,5	0,3	6,9	50	7,5	7,5	5	3,35
15	3	5	7,5	7,5	2,5	7,5	0,3	6,9	30	7,5	7,5	5	3,38
13	4	5	7,5	7,5	2,5	7,5	0,3	6,9	30	7,5	7,5	5	3,44
5	5	0	0	15	0	15	0,4	6,8	50	15	15	0	3,82
2	6	0	15	15	0	15	0,4	7	30	0	0	10	3,29
9	7	10	15	15	0	0	0,2	7	30	15	15	0	3,49
1	8	10	15	0	5	15	0,4	6,8	30	0	15	0	3,68
4	9	0	15	0	5	15	0,2	7	50	15	0	0	3,32
8	10	10	15	0	0	0	0,4	6,8	50	15	0	10	4,75
12	11	0	0	0	0	0	0,2	6,8	30	0	0	0	4,45
7	12	10	0	0	0	15	0,2	7	50	0	15	10	4,47
10	13	0	15	15	5	0	0,2	6,8	50	0	15	10	3,86
3	14	10	0	15	5	0	0,4	7	50	0	0	0	3,56
6	15	0	0	0	5	0	0,4	7	30	15	15	10	3,91
16	16	5	7,5	7,5	2,5	7,5	0,3	6,9	50	7,5	7,5	5	4,44

Figure 2.5 Utilized PBD for generating experimental configurations to evaluate factors and attained outcomes.

Following BioLector experiments, we conducted statistical analyses on the obtained responses. Results, including model validation parameters and significance values for the variables, were presented in ANOVA (Analysis of Variance) and fit statistic tables. Variables with significant effects (p-values < 0.05) were identified through the ANOVA table and complementary tools like the Half-normal plot, Pareto chart, and Perturbation chart of standardized effects. These analytical methods helped evaluate the significance and influence of each variable, aiding in the identification of key factors that notably impacted the observed responses in the experiment.

Optimizing factors using response surface methodology

Following the PBD results, we identified four highly impactful factors for further optimization through CCD in DoE software, resulting in thirty experimental runs (Figure 2.6).

std	run	KCl mM	MgCl2 mM	KH2PO4 mM	Tiamine mM	IPTG mM	pH	glucose (5-15 mM)	glycerol (1-5%)	LB broth (30-50 g/L)	$MgSO_4$ (5-15 mM)	Insulin g/L
6	1	10	15	15	10	0,25	6,8	15	1	50	5	6,4
26	2	10	15	15	10	0,25	6,8	10	3	40	10	9,5
19	3	10	15	15	10	0,25	6,8	10	0,171573	40	10	6,3
4	4	10	15	15	10	0,25	6,8	15	5	30	5	11
14	5	10	15	15	10	0,25	6,8	15	1	50	15	4,2
30	6	10	15	15	10	0,25	6,8	10	3	40	10	3,5
1	7	10	15	15	10	0,25	6,8	5	1	30	5	6
24	8	10	15	15	10	0,25	6,8	10	3	40	17,07107	5,7
17	9	10	15	15	10	0,25	6,8	2,9289322	3	40	10	7,7
12	10	10	15	15	10	0,25	6,8	15	5	30	15	5
7	11	10	15	15	10	0,25	6,8	5	5	50	5	12,5
15	12	10	15	15	10	0,25	6,8	5	5	50	15	7,4
5	13	10	15	15	10	0,25	6,8	5	1	50	5	9,9
28	14	10	15	15	10	0,25	6,8	10	3	40	10	5,8
29	15	10	15	15	10	0,25	6,8	10	3	40	10	9
27	16	10	15	15	10	0,25	6,8	10	3	40	10	3,9
11	17	10	15	15	10	0,25	6,8	5	5	30	15	-1,2
10	18	10	15	15	10	0,25	6,8	15	1	30	15	12,2
20	19	10	15	15	10	0,25	6,8	10	5,828427	40	10	9,2
23	20	10	15	15	10	0,25	6,8	10	3	40	2,928932	6,5
25	21	10	15	15	10	0,25	6,8	10	3	40	10	6,9
18	22	10	15	15	10	0,25	6,8	17,071068	3	40	10	2,7
22	23	10	15	15	10	0,25	6,8	10	3	54,1421356	10	5
9	24	10	15	15	10	0,25	6,8	5	1	30	15	5,1
2	25	10	15	15	10	0,25	6,8	15	1	30	5	11,5
16	26	10	15	15	10	0,25	6,8	15	5	50	15	5,1
3	27	10	15	15	10	0,25	6,8	5	5	30	5	4,9
13	28	10	15	15	10	0,25	6,8	5	1	50	15	4
21	29	10	15	15	10	0,25	6,8	10	3	25,8578644	10	5,7
8	30	10	15	15	10	0,25	6,8	15	5	50	5	4,9

Figure 2.6 CCD experimental runs.

CCD assessment of corresponding responses with consistent variables including KCl, MgCl2, KH2PO4, thiamine, IPTG, and pH.

The designated composition of model terms (selected factors) guided culture media preparation for each run, adhering to the designed points via BioLector. Additionally, constant values of other non-model media components were supplemented based on insulin results (mg/ml) from the PBD design and SDS-PAGE outcomes. Concentrations corresponding to the highest response for less significant variables and the -1 level of insignificant factors were provided. After executing experiments through BioLector, responses underwent scrutiny using different models. The optimal model was determined by assessing model validation parameters in the ANOVA table, fit statistic tables, and diagnostic analysis. The DoE software generated a perturbation chart, illustrating the graphical impact of each significant independent and dependent variable on the response through contour and 3D plots. Subsequently, the DoE software generated four scenarios outlining optimal points based on the acquired regression equation. Predicted design scenarios with the highest desirability were analyzed and

compared using the Biolector XT microbioreactor in triplicates. The suggested optimal media, defined as Scenario-3 (13 mg/ml) with a 3.3 deviation, is now poised for scale-up optimization.

Observations

The research evaluated the impact of eleven factors on bacterial growth through a series of experiments designed using DoE software, resulting in 16 trial configurations (Figure 2.6). The objective was to analyze the growth kinetics of the *esrapid* using the BioLector system. LB broth media at two concentrations, precisely 30 g/L or 50 g/L, were employed, and a comprehensive DoE-PBD with 16 parameters was conducted in triplicate on a single plate (Figure 2.5). The cells were inoculated for a 4-hour duration in each well, conducting experiments under constant agitation (800 rpm) at 37°C in 48-well with a working volume of 1000 µL. The induction with IPTG was carried out in each well at distinct final concentrations (0.2 mM, 0.3 mM, 0.4 mM), adhering to DoE-PBD parameters, and conducted overnight at 37°C. The final insulin density, quantified in grams per liter using a UV/Vis spectrophotometer (280 nm) and analyzed via 20% SDS-PAGE, served as the response variable, documented in the PBD table (Figure 2.7 and Figure 2.8).

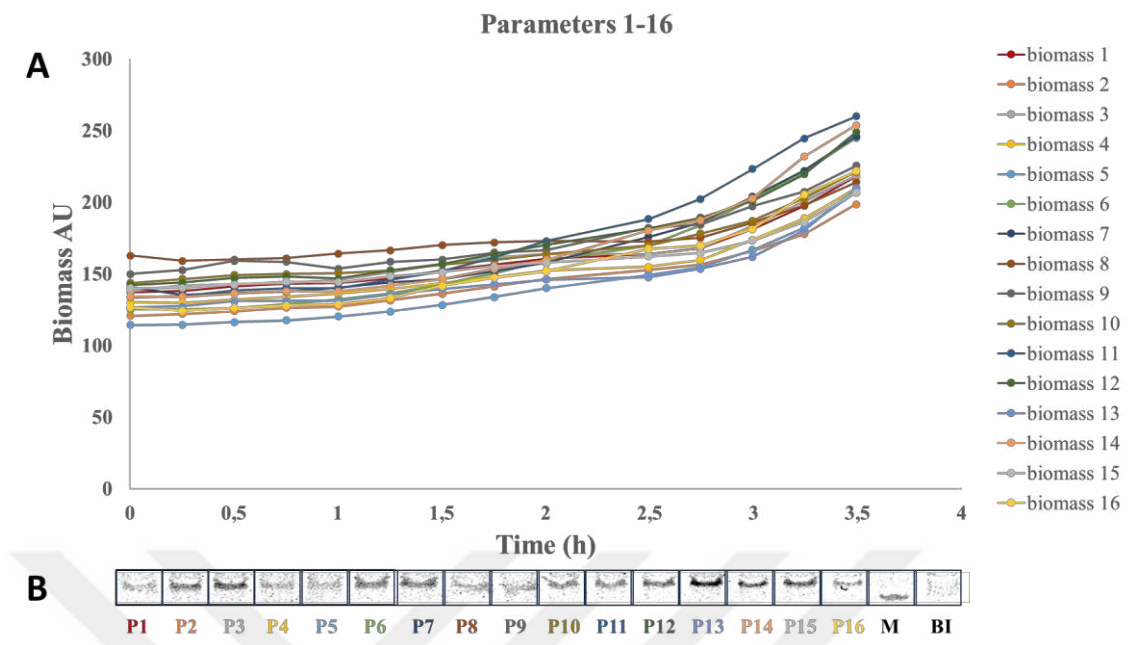


Figure 2.7 Production and characterization of modified proinsulin.

Growth Kinetics of 16 parameters using BioLector (A) and analysis of resulting modified proinsulins via 20% SDS-PAGE (B). Proinsulin size is 11.8 Kda. Each parameter is numerated with the character "P" synchronized with the biomass color. M: marker, BI: before IPTG induction of the insulin. The original SDS-PAGE result is mentioned in FigureX in below.

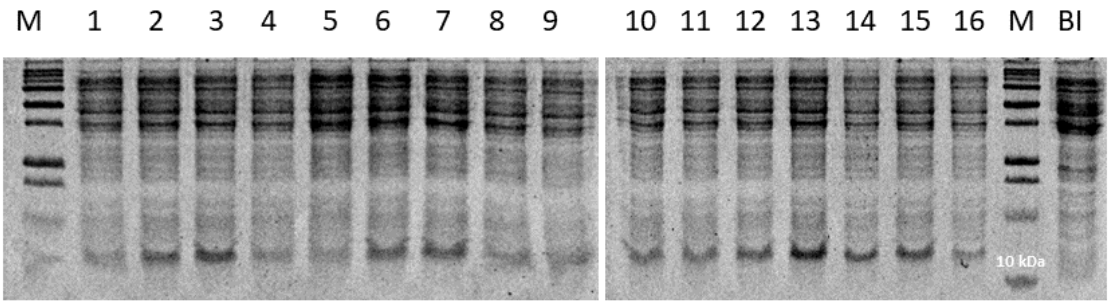


Figure 2.8 20% SDS-PAGE result of 16 Parameters.

M: markers, BI: before IPTG induction of insulins.

The statistical analysis of the data confirmed the significance of the considered model, with a p-value of 0.05 and an 85.37% coefficient of determination (R2) (Table 2.1). Notably, influential factors, including $MgSO_4$, glycerol, glucose, and LB broth concentration, were identified with p-values below the standard 0.05 threshold, indicating their substantial impact on the response variable. Additionally, the F-statistic demonstrated a stronger association with higher F-values for each term in relation to the

response variable.

Table 2.1 Presentation the ANOVA table for the chosen factorial model of the PBD

This table was generated to analyze the variance within the selected factorial model.

Source	Sum of squares	df	Mean square	F value	p-value Prob > F
Model	28155	4	7038,75	11,67	0,0020*
C-$MgSO_4$	11285,33	1	11285,33	18,72	0,0025*
D-Glycerol	7803	1	7803	12,94	0,0070*
E-Glucose	5461,33	1	5461,33	9,06	0,0168*
H-LB broth	3605,33	1	3605,33	5,98	0,0402*
Curvature	2674,71	1	2674,71	4,44	0,0683 ns
Residual	4823,5	8	602,94		
Lack of Fit	4783	7	683,29	16,87	0,1854 ns
Pure Error	40,5	1	40,5		
Cor Total	35653,21	13			

The half-normal plot visually represents the distribution of residuals in a regression model, serving as a tool to assess the normality assumption of the residuals [Zahn et al., 1975]. In Figure 2.9a, a linear pattern in the half-normal plot indicates that the residuals follow a normal distribution. Conversely, any departure from linearity in the plot suggests non-normality in residuals, meaning a violation of the normality assumption. The half-normal plot in Figure 2.9a illustrates the comparative significance of independent variables in the regression model, where the $MgSO_4$, glycerol, glucose, and LB broth concentration values exhibit higher percentage probabilities of standardized effects than other variables. A pareto chart is a graphical representation of the standardized effect exerted by each independent variable on the response variable [Kenett, 1991]. The evaluation of factors involves two statistical thresholds, the Benforrini limit and the t-value limit, determined at 3,898 and 2.306, respectively, for a significance level (α) of 0.05. The Pareto chart illustrates the magnitude of standardized effects through bars, with the most prominent values associated with $MgSO_4$, glycerol, glucose, and LB broth concentration. This underscores their significant influence on the

response, leading to their selection for optimization using the CCD approach, a widely used RSM technique (Figure 2.9b). The perturbation plot is a valuable tool for understanding and comparing the effects of multiple factors within a specific point in the design space [Bonnans and Shapiro, 2013]. It enables the visualization of the response by systematically varying a single factor across its entire range while keeping all other factors constant [Bonnans and Shapiro, 2013]. Perturbation plots provided a comprehensive understanding of independent variables' direct and interactive effects on insulin yield. The perturbation plot (Figure 2.9c) visually illustrates the immediate effects of variables C, D, E, and H on insulin yield.

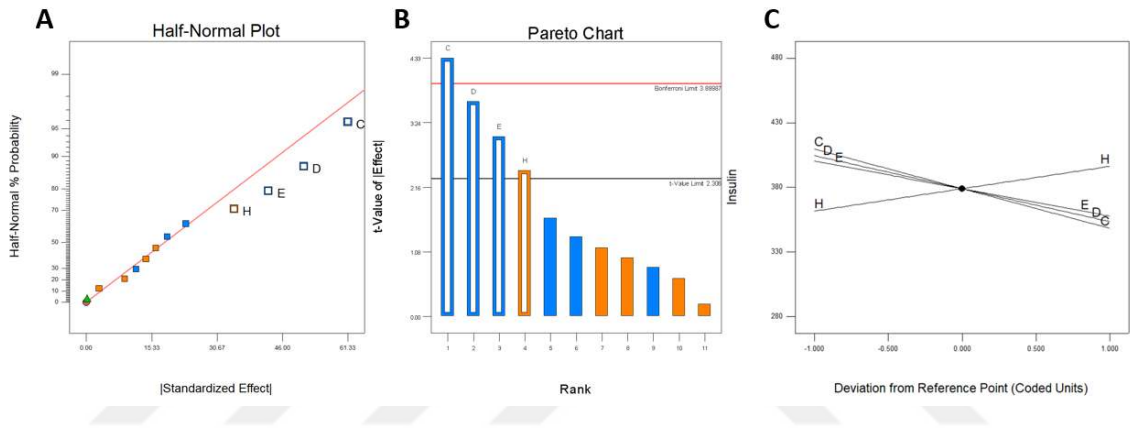


Figure 2.9 Statistical outputs of esrapid analog's expression yield in CCD perspective.

(A) Half-normal plot of Plackett-Burman Design. The graph illustrates the relative importance of various independent variables within a regression model, specifically highlighting the significant impact of $MgSO_4$, glycerol, glucose, and LB broth concentration values. These variables exhibit considerably higher percentage probabilities of standardized effects compared to other variables in the model, indicating their greater influence on the outcome. (B) Pareto chart of Placket-Burman Design. The effectiveness of factors was evaluated using two statistical thresholds, namely the Benforrini limit and t-value limit. These thresholds were determined to be 3,898 and 2.306, respectively, for a significance level (α) of 0.05. The Pareto chart visually represents the magnitude of standardized effects through bars. Notably, four bars stand out with the highest values, corresponding to $MgSO_4$, glycerol, glucose, and LB broth concentration. (C) Perturbation chart of Plackett-Burman Design. A comprehensive understanding of the impacts caused by the individual and combined effects of independent variables on insulin yield was achieved by employing perturbation plots. These plots visually depict the direct effects of variables C, D, E, and H on insulin yield, providing a clear elucidation of their respective influences.

The DoE package facilitated the creation of 30 experiments using the RSM to optimize specifically chosen model terms, including $MgSO_4$, glycerol, glucose, and a specific LB broth concentration (referred to as TB). The experimental runs were carried out with the

BioLector microfermenter in a 1 ml screening culture that maintained a constant concentration of 10 mM KCl, 15 mM $MgCl_2$, 15 mM KH_2PO_4, 10 mM thiamine, and 50 g/L 2X LB Miller media (Figure 2.6, Figure 2.10 and Figure 2.11).

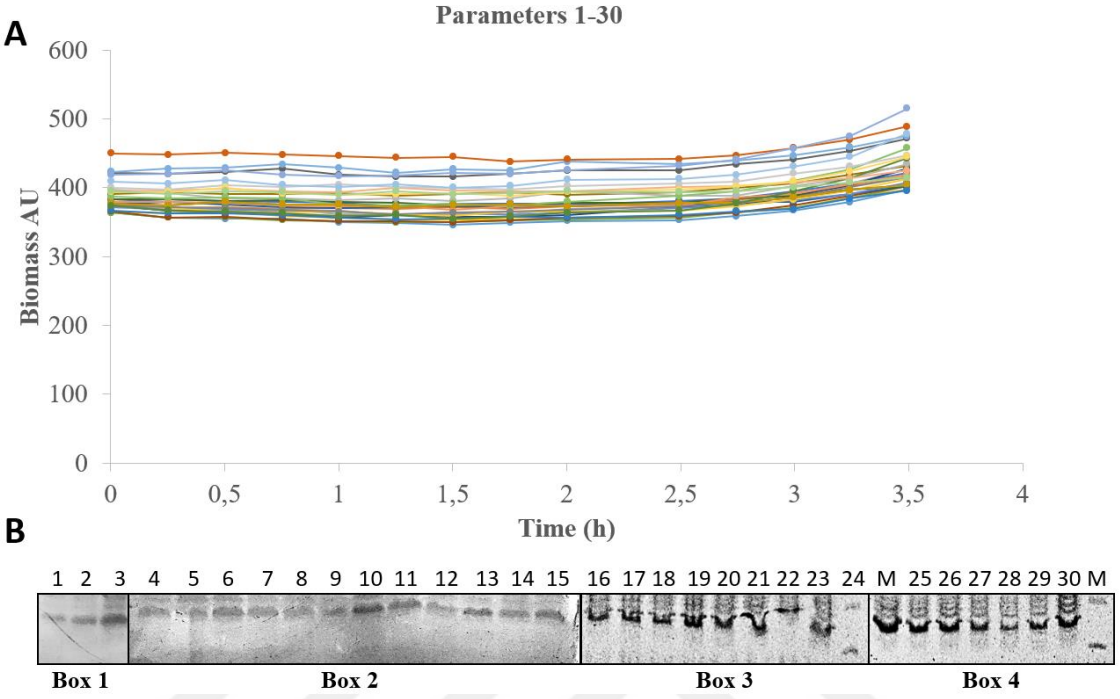


Figure 2.10 Illustration the production and characterization of esrapid in the second round of experimental validation.

It includes the growth kinetics of 30 parameters using BioLector to conduct CCD analysis, along with the analysis of resulting *esrapid* expressions via 20% SDS-PAGE. The original SDS-PAGE result is mentioned in Figure 2.11 in below.

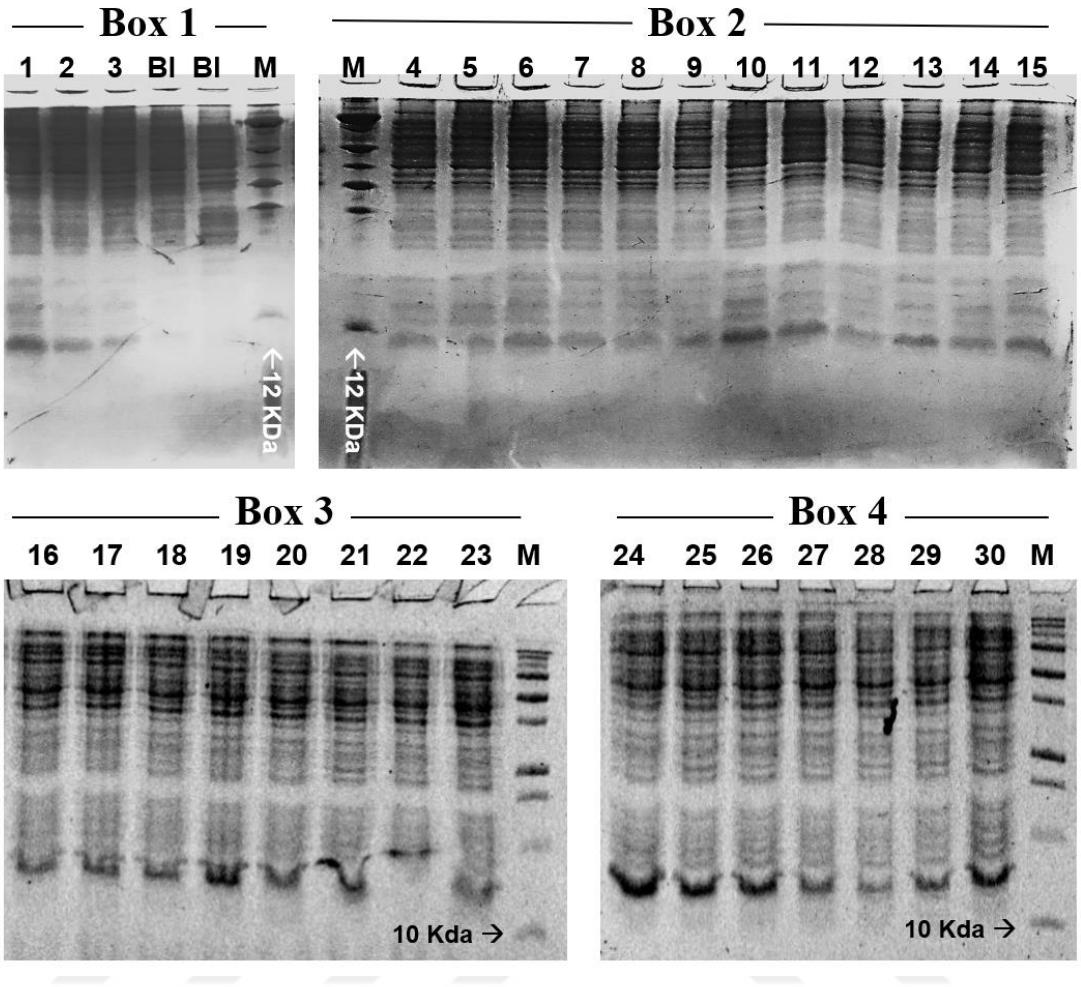


Figure 2.11 Original SDS-PAGE result of 30 Parameters.

Proinsulin size is 11.8 Kda. M: markers, 12 Kda and 10 KDa. BI is referred as before IPTG induction of the proinsulins.

Additionally, 0.25 mM IPTG was tailored to each specific design point. Following analyses using various models, the quadratic model emerged as a suitable candidate for predicting and validating optimal conditions. Significantly, the model exhibited a low p-value of 0.0037, providing a reliable prediction framework. Moreover, the lack of fit of the model displayed insignificance at 0.4479 relative to the pure error, suggesting the immateriality of the error in influencing the accuracy of the proposed model (Table 2.2).

Table 2.2 Presentation the ANOVA table used for the Response Surface Reduced Quadratic Model

Source	Sum of squares	df	Mean square	F value	p-value Prob > F
Model	201.91	8	25,24	5,31	0.0010*
A-Glucose	1,25	1	1,25	0,26	0.6139
B-Glycerol	93,10	1	93,10	19,58	0.0002*
C-TB	2,82	1	2,82	0,59	0.4502
D-$MgSO_4$	4,04	1	4,04	0,85	0.3671
AB	7,29	1	7,29	1,53	0.2293
BC	4,20	1	4,20	0,88	0.3579
BD	4,84	1	4,84	1,02	0.3245
B^2	84,38	1	84,38	17,74	0.0004
Residual	99,86	21	4,76		
Lack of Fit	78,33	16	4,90	1,14	0.4833 ns
Pure Error	21,53	5	4,31		
Cor Total	301,77	29			

The model's R2 value of 0.6866 indicates a reasonable fit to the experimental data. Moreover, the adequate precision value of 7.861 suggests that the signal is sufficient for scrutinizing the biological system. Perturbation analysis, factorial contour plots, and 3D response surface of the central composite design were employed to earn a deeper understanding of the influential effects and interactions of independent variables on insulin yield (Figure 2.12 and Figure 2.13). The two-dimensional contour and three-dimensional response surface plots effectively illustrated the intricate relationship between the independent variables and the resulting insulin yield (Figure 2.12 and Figure 2.13). The contour plots (Figure 2.12) provide valuable insights into the inherent characteristics and magnitude of interactions among various contributing factors.

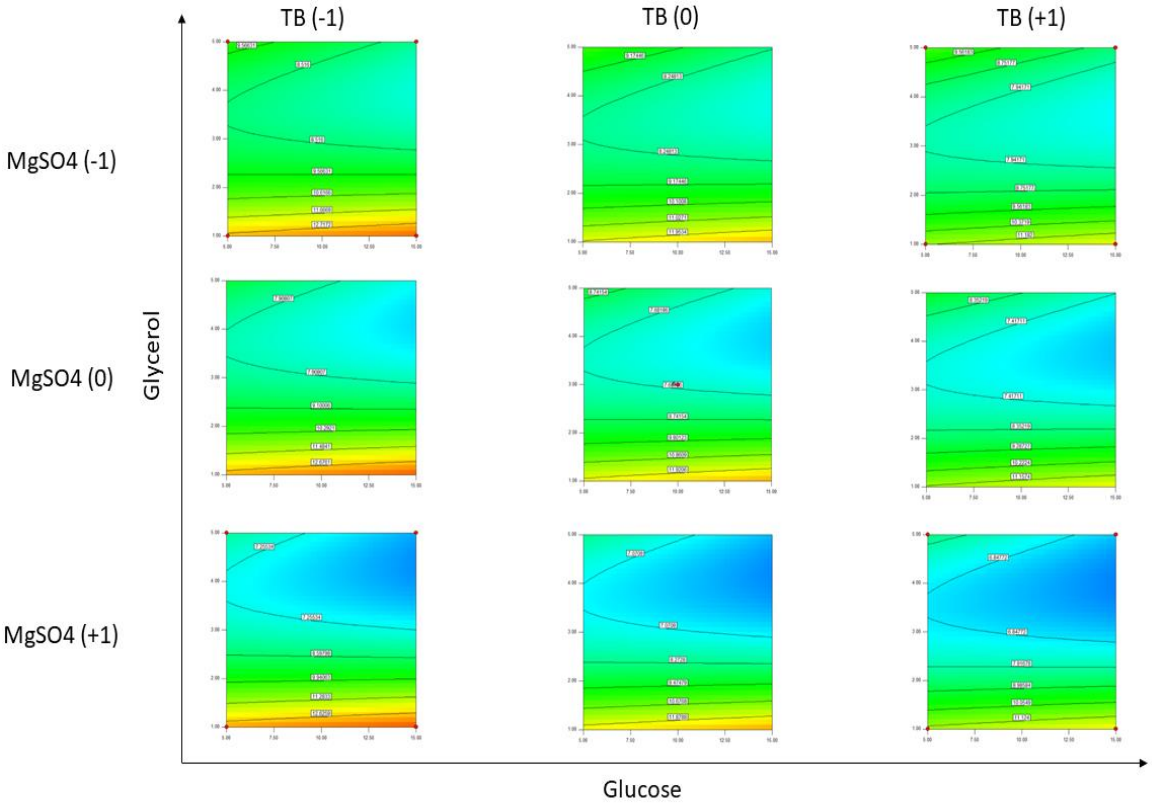


Figure 2.12 Factorial contour plots of CCD provide an effective visualization tool.

Illustration the complex relationship between independent variables and the resulting insulin yield. Systematic manipulation of these variables within predetermined ranges enables a comprehensive exploration of their collective impact on insulin production. In the plots, Glucose is represented on the X-axis, and Glycerol on the Y-axis.

Specifically, an increase in glycerol levels led to a substantial reduction in insulin yield, decreasing from 12.6259 mg/ml to 7.25534 mg/ml. Similarly, a simultaneous rise in glycerol and media content to elevated levels (exceeding 30 g/l) resulted in a decrease in insulin yield from 11.124 mg/ml to 6.84772 mg/ml. On the other hand, the presence or absence of $MgSO_4$ showed no discernible impact on the insulin yield. The interplay between dependent and independent variables was further elucidated by implementing 3D response surface plots (Figure 2.13).

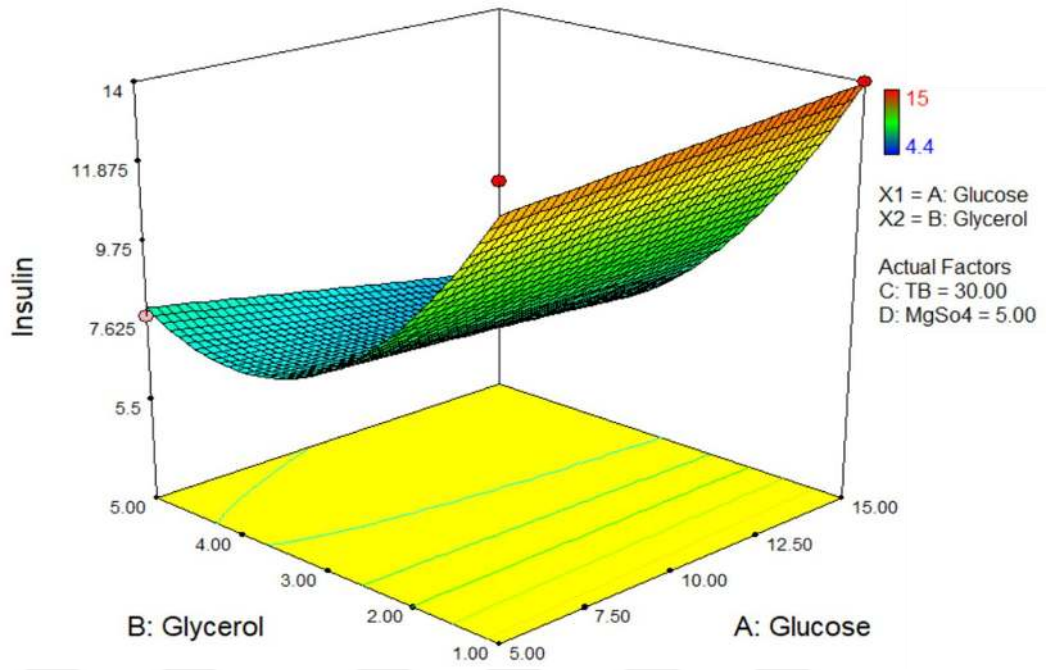


Figure 2.13 3D contour of CCD.

(A) Response surface plot demonstrating the interactive effects of glucose, glycerol, $MgSO_4$, and media content concentrations on insulin production efficiency, with constant temperature and pH.

These plots effectively confirmed the intricate and interactive influence of specific glucose, glycerol, $MgSO_4$ concentrations, and media content on insulin yield. Notably, when considering the actual factors of LB (30 g/l) and $MgSO_4$ (5mM), it was observed that a lower concentration of glycerol and the highest concentration of glucose displayed a positive correlation with the production of insulin at higher levels compared to the other variables. Consequently, based on these findings, four distinct scenarios were statistically derived and subsequently subjected to experimental validation. The experimentation was conducted using the BioLector micro-fermenter, with each scenario being replicated three times to ensure the reliability and accuracy of the results (Table 2.3).

Table 2.3 Projected optimal conditions for achieving peak insulin production.

Scenario	Glucose (mM)	Glycerol (1%)	TB (g/L)	$MgSO_4$ (mM)	Predicted Insulin (g/L)	Observed Insulin (g/L)	Deviation (%)
1	14.58	1	30.09	15	13.92	11.87	14.8
2	15	1	30	6.73	13.79	11.54	16.3
3	8.78	1	30	15	13.44	13.00	3.3
4	15	1	37.51	10.62	13.21	12.33	6.7

Despite the absence of substantial discrepancies among the scenarios, it is noteworthy that scenario III, characterized by a glucose concentration of 8.78 mM, glycerol concentration of 10 g/L, LB of 30 g/l, and $MgSO_4$ concentration of 15 mM, exhibits the slightest deviation and the most heightened insulin production in comparison to the additional scenarios (Table 2.3).

Concluding remarks

The upstream processing section of Chapter 2 introduces an innovative formulation for the high-throughput cultivation medium of the *esrapid*^{2} analog. This involves a combination of experimental approaches utilizing the BioLector microbioreactor and computational methods employing DoE software. The effectiveness of each stage was verified by using high-throughput PBD and CCD within the DoE software framework, complemented by the screening capabilities of the BioLector micro-bioreactor. This allows for comprehensive evaluations in a complex batch environment. The primary focus of the investigation was on Scenario III, which demonstrated superior attributes and highlighted its potential as a more streamlined and cost-effective medium for achieving high cell density fermentation of the *esrapid*^{2} analog in the *E. coli* Rosetta-2 strain.

Bacterial growth is intricately influenced by a multitude of parameters. Identifying and optimizing these factors pose significant financial and economic challenges [Packiam et al., 2020]. In addressing these challenges, combining experimental BioLector micro-bioreactors and computational DoE methodology provides a robust approach for both experimental and statistical optimization of bioprocesses. This

integrated approach enables superior outcomes while minimizing the expenditure of time and resources [Elibol, 2004]. Numerous studies have successfully utilized DoE methods to improve the yield of recombinant protein expression in *E. coli* by optimizing culture media [Sunitha et al., 1999; Shahbazmohammadi and Omidinia, 2017; Zare et al., 2019; Duan et al., 2020]. However, the existing literature is restricted in terms of investigations encompassing diverse protein types, design methodologies, factors under evaluation, and optimization outcomes [Sunitha et al., 1999; Shahbazmohammadi and Omidinia, 2017; Zare et al., 2019; Duan et al., 2020]. Examples include the optimization of culture media for producing L-Asparaginase, Phytase, Streptokinase, and Reteplase in *E. coli*, using DoE-based strategies to enhance production yields [Sunitha et al., 1999; Kenari et al., 2011; Ghoshoon et al., 2011]. A comprehensive literature review reveals numerous variables capable of influencing bacterial growth rate and biomass production, such as the nature and concentration of carbon and nitrogen sources, pH conditions, and the inclusion of trace elements [Sunitha et al., 1999; Shahbazmohammadi and Omidinia, 2017; Zare et al., 2019; Duan et al., 2020; Kenari et al., 2011; Ghoshoon et al., 2011].

Moreover, multiple studies have focused on employing the 48-well BioLector microbioreactor system [Osthege et al., 2022; Sparviero et al., 2023; Flitsch et al., 2016; Lennen et al., 2016]. For instance, this system has demonstrated effectiveness in real-time monitoring of lipid production and growth tracking in Y. lipolytica [Back et al., 2016]. The literature also highlights the transferability and comparability assessment between the BioLector and fully controlled bioreactor systems operating in fed-batch mode, especially at moderate to high cell densities [Toeroek et al., 2015]. Additionally, various investigations showcase the efficiency of the BioLector system in screening optimal growth conditions and engineered strains before scaling up [Kensy et al., 2009]. This robust cultivation protocol conducted on microtiter plates enables the screening of *E. coli* systems under conditions resembling lab-scale bioreactor cultivations [Kensy et al., 2009].

In the upstream part of the chapter 2, the Plackett-Burman Design (PBD) was employed to examine the impact of eleven factors on cellular growth. These factors encompassed various nitrogen (N) and carbon (C) sources, salts, metal ions, pH, and the buffering system. Notably, the lower concentration of LB medium (30 g/L) and, in contrast to prior studies, the levels of $MgSO_4$, glucose, and glycerol were identified as the most influential on cellular growth. Consequently, these factors were singled out for

subsequent optimization using RSM and CCD. Additionally, 0.25 mM IPTG demonstrated advantageous effects on cellular growth. During CCD experiments, model terms with non-significant p-values were intentionally excluded from the culture media. According to the RSM outcomes, glycerol concentration emerged as the most influential factor, demonstrating a direct relationship with increased cellular proliferation. These findings are supported by previous research highlighting the positive impact of glycerol on *E. coli* growth, particularly in anaerobic fermentation [Dharmadi et al., 2006]. When considering the coexistence of glycerol as a carbon source and IPTG as an inducer [Malakar and Venkatesh, 2012], our results indicate that the optimal glycerol concentration in the presence of IPTG is only 1% (v/v). This observation aligns with studies displaying a decline in cellular yield as glycerol concentration increases [Malakar and Venkatesh, 2012].

Furthermore, the fitting of the Hill equation has been validated to yield a prototypical Monod-type expression for growth on glycerol, both in the presence and absence of IPTG, as supported by previously published works [Malakar and Venkatesh, 2012]. On the other hand, glucose and $MgSO_4$ have been recognized in the existing literature as pivotal factors with a profound influence on *E. coli* growth [Izaki and Arima, 1965]. However, the positive effects of limited glycerol presence on growth may also be attributed to the detrimental consequences of glucose, acting as a carbon source, under conditions of limited nitrogen availability [Bren et al., 2016; Shiloach et al., 1996; Michaels et al., 1983]. In other words, in situations characterized by nitrogen limitation, the favorable impacts associated with a modest glycerol concentration become apparent when contrasted with the adverse effects of glucose utilization as a carbon source.

The gram-negative bacterium *E. coli* stands out as a favored host organism for heterologous protein expression due to its ease of cultivation, cost-effective culture media, and potential for high product titers [Kopp et al., 2017]. However, harsh induction strategies using IPTG can induce stress responses, leading to "metabolic" or "product burden" phenomena, marked by reduced growth rates and cellular lysis during prolonged induction. Alternative approaches, such as "gentle" or "modifiable" induction techniques using lactose as an alternative inducer, have been explored to address these challenges [Kopp et al., 2017]. This approach aims to alleviate the strain on the production host. Traditional induction methods using glucose as the primary carbon source alongside lactose can lead to catabolite repression effects on lactose uptake kinetics, ultimately compromising overall product yield.

In contrast, glycerol, as an alternative carbon source, has shown promising outcomes when combined with glucose and lactose in auto-induction systems [Kopp et al., 2017]. Glycerol has demonstrated favorable effects on recombinant protein production, exhibiting reduced signs of catabolite repression during co-cultivation with lactose. Moreover, an investigation into the mechanistic aspects of glycerol uptake in the presence of lactose as an inducer has highlighted significantly enhanced inducer uptake rates in product-producing strains compared to non-producing strains. These findings emphasize glycerol's practical and less disruptive nature regarding cellular viability and recombinant protein productivity compared to glucose [Kopp et al., 2017].

Additionally, the breakdown of the carbon source in metabolism leads to the accumulation of acidic by-products, mainly acetate, in the culture medium [Kusuma et al., 2019; Ukkonen et al., 2011]. These acidic conditions can significantly hinder cell growth and endanger the production of recombinant proteins [Kusuma et al., 2019; Ukkonen et al., 2011]. To address this issue, the addition of yeast extract and tryptone to the culture medium helps alleviate medium acidification, primarily by offsetting the increased ammonia levels produced during their metabolic utilization [Kusuma et al., 2019; Ukkonen et al., 2011]. Among the various conditions examined, Scenario III represents an optimized formulation of the culture medium, characterized by an ideal combination of yeast (0.5 g/L), tryptone (1 g/L), and salt (1 g/L) concentrations, along with comparatively lower glucose concentration (8.8 mM). This configuration may promote enhanced cell growth and delay entry into the death phase, indicating its suitability for achieving favorable outcomes regarding cellular dynamics and prolonged cell viability.

2.1.4. Large scale expression of "esrapid" analogs

Initially, stock cells containing *esrapid* analog genes were cultured overnight in 150 mL of LB medium (30 g/L), supplemented with 50 µg/mL kanamycin and 35 µg/mL chloramphenicol, at 37°C and 110 rpm using the New Brunswick Innova 4430R incubator. A 150 mL cell culture was transferred to a 6 L bioreactor (Braun Biotech Sartorius 6L Liter Jacketed Glass Vessel Bioreactor) with the same antibiotics and 4 mL of antifoam. The main cultivations were conducted at 37°C, and the batch culture conditions included an initial culture volume of 4 L, air flow of 15 L min^{-1}, and stirrer speed of 500 min−1. Thermal mass flow meters (Bronkhorst High-Tech B.V., Ruurlo, The Netherlands) were employed for mixing air and oxygen. The concentration of dissolved oxygen was analyzed using two polarographic electrodes (Ingold, Germany), and it was maintained at 40% of air saturation by adjusting stirrer speed and aeration rate. Pure oxygen was added to the inlet air if necessary. The pH was kept at 6.8 by adding aqueous ammonia (25% w/w), and all controls were managed by the bioreactor's process control system. Upon reaching an OD_{600} of 0.8, protein production was induced by introducing 0.25 mM IPTG to the cultured cells, followed by a subsequent 3-hour incubation at 37°C and 250 rpm. After the incubation, cell pellets were harvested by centrifugation at 3500 rpm for 45 minutes using the Beckman Allegra 15R centrifuge, and the collected pellets were stored in a -45°C freezer.

2.2. The downstream process and optimization of ultra-acting analogs: "esrapid"

2.2.1. Solubilization and sulfitolysis

To partially purify and fully solubilize inclusion bodies, 1-gram cell pellets containing *esrapid* proinsulin were individually suspended in 10 mL of lysis buffer, consisting of 50 mM Tris-HCl (pH 9.0) and 5 mM EDTA, and homogenized. Bacterial cells were lysed through sonication, and the resulting cell debris underwent low-speed centrifugation at 6000 × g for 45 minutes. Pellets were washed with a buffer comprising 50 mM Tris-HCl (pH 9.0), 5 mM EDTA, 0.1% (v/v) Triton X-100, 1 M urea, sonicated for 30 seconds, and centrifuged at 8000 × g, 4°C, for 45 minutes. The debris resuspended at 0.1 g/mL in binding buffer containing 50 mM Tris-HCl (pH 9.6) and 5

mM 2-Mercaptoethanol with 8 M urea was employed to collect inclusion bodies. During this step, pellet was homogenized completely. These bodies, containing *esrapid* proinsulin, were incubated at room temperature overnight. Proteins were retrieved from the debris through centrifugation at 10000 × g for 45 minutes at 4°C. Prior to sulfitolysis, the supernatants were filtered using a 0.2 μM pore diameter filter. Supernatant sample was treated with 0.2M sodium sulfite and 0.02M sodium tetrathionate and incubated at room temperature for 4-hour.

2.2.2. Purification and refolding

The sample underwent separation from contaminant E. coli proteins using Ni-NTA affinity resins in a manual gravity column, initially equilibrated in a solution of 50 mM Tris-HCl (pH 9.6), 5 mM 2-Mercaptoethanol, and 8 M urea. Elution was carried out with the same solvent enriched with 250 mM imidazole, and all fractions were collected in 5 mL Falcon tubes and subsequently analyzed using 20% SDS-PAGE. The refolding of esrapid followed a protocol outlined in previously published works [Kim et al., 2015]. In summary, purified proinsulin underwent dialysis against a solution of 0.1 M Gly/NaOH at pH 10.5, 0.5 M urea, 0.5 mM BME, and 0.5 mM cysteine and was stirred at 4 °C for 18 hours. The refolded proinsulin was exchanged with 25 mM Tris-HCl at pH 8.2 for tryptic digestion and went at 4 °C for 18 hours.

2.2.3. Tryptic digestion, pH precipitation and purification

The refolded proinsulin underwent treatment with trypsin (at a ratio of 1:1000) and was incubated at 30°C for 4 hours. Following digestion, mature insulin was precipitated at its isoelectric point (pI). The resulting mature insulin pellet was dissolved in 20 mM citric acid, reaching a total volume of 5 ml. For the purification of *esrapid*, gel filtration chromatography using Superdex 200 (GE Healthcare) with dimensions of 30 cm × 10 mm was employed. This purification process was performed at room temperature, employing a flow rate of 0.5 ml/min and a fraction size of 5 ml. Detection of eluted protein fractions was achieved by measuring absorbance at 276 nm. All fractions were concentrated through filtration using an Amicon® (Ultra Centrifugal Filter with a molecular weight cutoff (MWCO) of 10 kDa).

2.2.4. Crystallization

The mature form of *esrapid*, lacking zinc binding, was subjected to crystallization in 20 mM citric acid using approximately 3000 varied commercial sparse matrix and grid screen crystallization conditions. The transition of the *esrapid*^{2} sample from its monomeric state to the zinc-bound form involved a zinc precipitation step utilizing a solution containing 2.4 M NaCl, 100 mM citric acid, and 6 mM $ZnCl_2$. Subsequent solubilization was achieved with Tris-HCl at pH 8.0. The zinc-bound *esrapid*^{2} samples were then subjected to crystallization using the sitting drop vapor diffusion micro-batch under the oil technique. Approximately 3000 diverse commercial sparse matrix and grid screen crystallization conditions were employed. In a 72-well Terasaki plate (Cat#654,180, Greiner Bio-One, Austria), crystallization conditions were mixed in a 1:1 ratio with a 9 mg/ml *esrapid*^{2} solution (v/v, 0.83 µl). Each well was covered with 16.6 µl of paraffin oil (Cat#ZS.100510.5000, ZAG Kimya, Türkiye) and then incubated at 4 °C. Crystal formation in the Terasaki plate wells was observed using a compound light microscope.

2.2.5. A pilot study for esrapid^2 insulin analog

Analyzing esrapid^2 insulin kinetics through uniform HDX reactions

Recently, a new category of insulins, referred to as "ultra-rapid-acting" insulins, has emerged. Faster-acting insulin aspart, known as Fiasp and approved in 2017, falls into this category. Fiasp incorporates niacinamide and l-arginine, resulting in quicker initial absorption compared to insulin aspart (IAsp) [Avgerinos et al., 2021]. Subcutaneous administration of FIAsp has shown superior early glucose-lowering efficacy compared to IAsp. FIAsp is particularly suitable for insulin pump use, outperforming other short-acting insulin analogs. These ultra-rapid-acting insulins provide additional bolus insulin options for individuals with type 1 or type 2 diabetes, offering benefits such as improved glycemic control, compatibility with infusion pumps, and user-friendly pen delivery devices [Avgerinos et al., 2021]. They achieve rapid onset of action through accelerated insulin absorption and quicker attainment of maximum insulin concentration. However, our knowledge of the kinetics of monomers or oligomers in ultra/fast-acting insulins is limited. Exploring comparative dynamical mechanisms intrinsic to rapid-acting insulins is a key strategy, and a profound understanding of insulin's complex conformation is crucial for deciphering pivotal regions involved in insulin-receptor interactions [Avgerinos et al., 2021].

This pilot study of Chapter 2 aims to uncover the underlying mechanism driving dynamic changes in the monomer of a less stable insulin variant compared to commercially available insulin aspart. A mutation was introduced in the arginine residue on the B22 chain, replacing it with the more flexible lysine, and the impact on protein stability and electrostatic interactions was investigated. Insulin aspart served as the model system for this study. Initially, the surface arginine residue in the B22 chain of insulin aspart underwent maximum mutation to lysine, considering the protein's foldability. Subsequently, the stability of this variant monomer was evaluated using the differential scanning calorimetry technique. Hydrogen-deuterium exchange Mass Spectrometry (HDX-MS) analysis was employed to explore the intrinsic dynamics of the monomer, and a comparative study was conducted with the monomer of insulin aspart. The experimental findings were validated through computational analysis. Finally, differences in crystal structure between the hexamer *esrapid^2* and aspart were determined using the multi-well X-ray crystallography technique conducted at ambient

temperature. Lastly, we performed Ca^{2+} imaging to measure the esrapid^2 activity on neural cell.

Monomer sample preparation

Proinsulin was produced as described in section *2.2.1* and *2.2.2.* Briefly, proinsulin was subjected to tryptic cleavage using trypsin (1:1000) at 30°C for 4 hours. After cleavage, pH-induced precipitation was initiated by adding 20 mM citric acid to the protein sample. Following precipitation, *esrapid*^{2} underwent centrifugation at 3500 rpm for 5 minutes, and the resulting pellet was dissolved in 20 mM citric acid to halt trypsin activity and eliminate trypsin. Post-solubilization, *esrapid*^{2} passed through a 0.2 µM pore diameter filter and underwent size exclusion chromatography (SEC) using a 20 mM citric acid buffer. High-purity fractions were collected, and pH precipitation was carried out using 1 M Tris at pH 10. The *esrapid*^{2} pellet was then solubilized using a Tris-HCl buffer at pH 9.0, and the concentration was adjusted to 200 µg/mL for HDX-MS/MS analysis. pH-precipitation, followed by solubilization in Tris-HCl pH 10 buffer, was employed to convert insulin aspart (NovoRapid) from a hexameric to a monomeric state. A buffer exchange with Tris-HCl pH 9.0 was performed ten times using a 10K Amicon filter. The resulting monomeric NovoRapid form was then adjusted to a 200 µg/mL concentration for HDX-MS/MS analysis. Analytical size exclusion chromatography (SEC) was carried out to verify the monomeric state of both *esrapid*^{2} and NovoRapid.

Hydrogen-deuterium exchange mass spectrometry

HDX-MS analysis was executed at five different time points ranging from 12 seconds to 24 hours. The resulting data were used to generate a heat map, and pyMOL was employed to visualize structural features based on deuterium uptake rate, electrostatic potential, and hydrophobic potential. Both purified *esrapid*^{2} and NovoRapid were appropriately diluted in the resuspension buffer. To facilitate labeling, 2.5 µl (≥ 35 pmol) of each sample was incubated with 37.5 µl of Labeling Buffer (50 mM Tris; pD 9.4) at room temperature for varying incubation periods: 0, 0.2, 2, 20, 200, and 1440 minutes. Following labeling, samples were quenched with 40 µl of pre-chilled Quench Buffer (50 mM Tris, 3M GnHCl, 100 mM TCEP; pH 2.3), and the sample was

then incubated on ice to inhibit back-exchange. Triplicate analyses were performed for all labeling time points. Deuterated samples underwent digestion using The Enzymate BEH Pepsin Column (Waters Corp., Manchester, UK) at 20°C. After 3 minutes of trapping at Waters BEH C18 VanGuard pre-column (flow rate 200 L/min in 0.1% formic acid pH 2.5), proteins were subjected to online pepsin digestion and liquid chromatography/mass spectrometry (LC/MS) using the nanoACQUITY UPLC system and the Synapt G2 Si Q-Tof mass spectrometer from Waters Corp. (Milford, MA, USA). Online pepsin digestion occurred in 0.3% formic acid for 2 minutes at 0 °C. Peptide separation conditions involved a 5-minute gradient with 0.1% formic acid, 8–30% acetonitrile, maintained at 0 °C, and a 40 µl/min flow rate. MS data processing was conducted using DynamX software provided by Waters Corp.

Sample delivery and data collection

Data collection was conducted using the XtaLAB Synergy Flow X-ray diffraction (XRD) system from Rigaku (Figure), with control managed by CrysAlisPro software (Rigaku Oxford Diffraction, 2022), adhering to the procedure outlined in Gul et al. (2023). The airflow temperature of the Cryostream 800 Plus from Oxford Cryosystems was stabilized at 300 K (26.85 °C) and consistently maintained at ambient temperature throughout the data collection process. The 72-well Terasaki plate was affixed to the modified XtalCheck-S plate reader attachment on the goniometer omega stage. Initial screening involved two dozen crystals to assess diffraction quality. Adjustments were made to omega and theta angles, as well as X, Y, and Z coordinates to center crystals at the eucentric height of the X-ray focusing region. Subsequently, diffraction data were collected for each crystal. Crystals exhibiting favorable diffraction properties were selected for further data collection, with exposure times optimized to minimize radiation damage during overnight sessions. The most promising crystals (Figure 2.14) were grown in a buffer solution comprising 2.4M NaCI, 100 mM citric acid, 6 mM $ZnCl_2$, 20 (w/v) poly(ethylene glycol) PEG-8000.

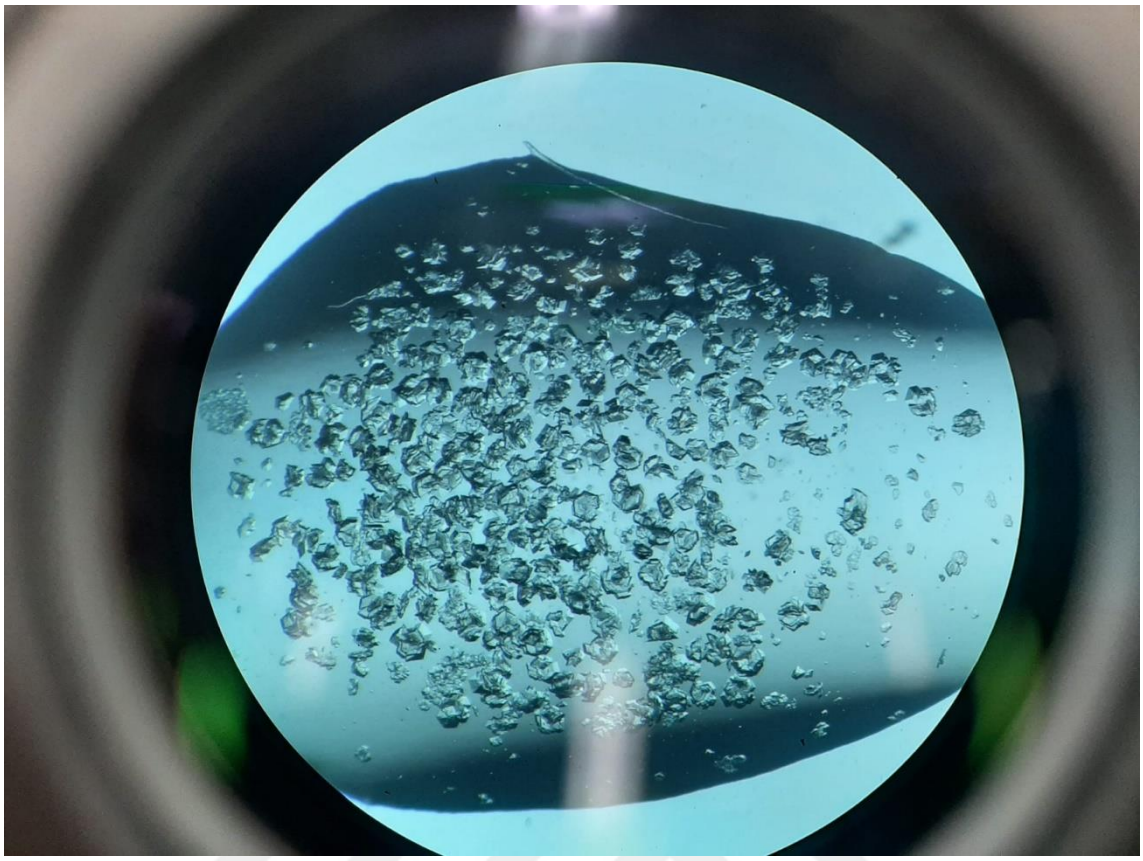


Figure 2.14 *esrapid*[2] crystals.

Crystallization buffer: 2.4M NaCI, 100 mM TRIS/HCI pH 7.4, 6 mM ZnCl2, 20 (w/v) poly(ethylene glycol) PEG-8000.

XtalCheck-S (Figure 2.15) was configured to oscillate to the maximum extent allowed by the detector distance to optimize crystal exposure oscillation angles throughout the data collection process. A total of 42 frames were captured, each lasting 3 minutes (5 seconds per frame) for all individual crystals. Parameters were set with the detector distance at 100.00 mm, a scan width of 0.5 degrees oscillation, and an exposure time of 10.00 seconds per image.

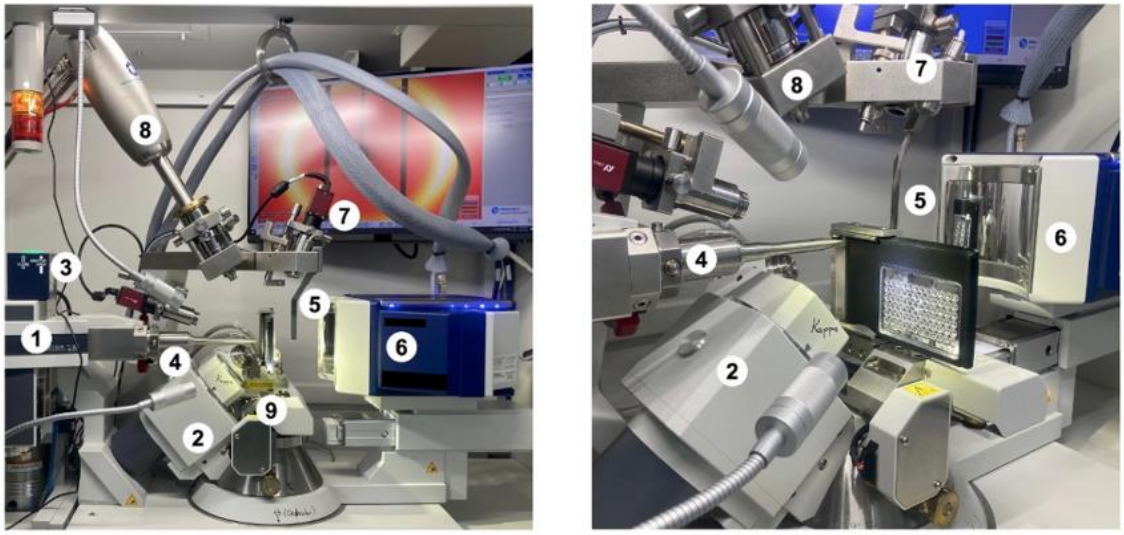


Figure 2.15 A visual depiction of the XtaLab Synergy-S instrument.

Comprehensive illustration of HyPix-Arc 150° detector and XtalCheck-S plate reader. Main hardware components of *Turkish DeLight*. (1) X-ray source, (2) four-circle Kappa goniometer, (3) shutter, (4) collimator, (5) beamstop, (6) X-ray detector, (7) video microscope, (8) low temperature insert, (9) XtalCheck module. (Atalay et al., 2023; Gul et al., 2023)

Data processing and structure determination

The plate screening parameters were optimized for all crystals, and a 42-degree data collecti (2014) on was performed for the selected crystal. All crystals were processed in CrysAlisPro . An optimal unit cell was determined, and peak identification and masking procedures were completed on the collected data. Subsequently, a batch script incorporating the “xx proffitbatch” command was executed for cumulative data collection. The batch data reduction was achieved using the CrysAlisPro Suite through the script command, producing files containing all integrated, unmerged, and unscaled data (*.rrpprof) for each dataset. To merge all datasets into a reflection data (.mtz) file, the proffit merge process in the “Data Reduction” section of the main window of CrysAlisPro was used. The reduced datasets (.rrpprof files) were merged using “proffit merge” option. The entirety of the data underwent refinement, integrating, and scaling processes through the aimless and pointless implementations within CCP4. Ultimately, the processed data were shipped into *.mtz formats. The crystal structure of *esrapid*[2] was determined under ambient temperature within the R3:H space group, operating the automated molecular replacement tool *PHASER* integrated into the *PHENIX* software package [McCoy et al., 2007]. Employing a previously published X-ray structure [PDB

ID: 4GBN; Palmieri et al., 2013] as the initial search model, its coordinates were used for the preliminary rigid-body refinement within *PHENIX*. Subsequent refinement steps contained simulated-annealing refinement, adjustment of individual coordinates, and Translation/Libration/Screw (TLS) parameters. Furthermore, a refinement technique known as composite omit map, embedded in *PHENIX*, was managed to identify potential locations of altered side chains and water molecules. The final model underwent meticulous examination and reconstruction in *COOT*, focusing on positions exhibiting notable difference density [Emsley and Cowtan, 2004]. Extraneous water molecules located beyond regions of significant electron density were manually excluded. The visualization of the X-ray crystal structure was created using *PyMOL* [Delano, 2002] and *COOT*. Data statistical of *esrapid*^{2} insulin analog has been indicated in Table 2.4.

Table 2.4 Data collection and refinement statistics of *esrapid*[2] insulin analog.

Data collection	***esrapid*[2]**
Instrument	*Turkish DeLight*
Resolution range	21.64 - 2.501 (2.59 - 2.501)
Space group	R 3: H
Unit cell	80.376 80.376 38.04 90 90 120
Redundancy	24.9 (12.9)
Completeness (%)	97.70 (91.02)
Mean I/sigma(I)	6.88 (1.72)
Wilson B-factor	37.28
R-merge	0.6092 (1.776)
R-meas	0.6197 (1.849)
R-pim	0.1113 (0.502)
CC1/2	0.952 (0.358)
CC*	0.988 (0.726)
R-work	0.3069 (0.4477)
R-free	0.3529 (0.3590)
CC (work)	0.846 (0.192)
CC (free)	0.689 (0.600)
Protein residues	102
RMS (bonds)	0.091
RMS (angles)	0.86
Ramachandran favored (%)	94.68
Ramachandran allowed (%)	4.26
Ramachandran outliers (%)	1.06
Average B-factor	55.83
macromolecules	56.44
ligands	38.47
solvent	45.06

Differential Scanning Calorimetry

Differential Scanning Calorimetry (DSC) was employed to evaluate the thermal stability of insulin aspart and *esrapid*[2] insulin samples at a concentration of 2 mg/ml. The experiments were conducted using a SETARAM Micro DSC-III calorimeter, covering the temperature range of 20 to 100 °C with a heating rate of 0.3 K/min. Precision balancing, accurate to ± 0.05 mg, was applied to the aspart and *esrapid*[2] samples, as well as the reference (a 20 mM citric acid buffer solution), to minimize corrections related to the heat capacity of the vessels. To correct the baseline, a secondary thermal scan of the denatured sample was performed. Analysis of parameters such as melting temperature (Tm), calorimetric enthalpy (ΔHcal), and change in entropy (ΔS) was carried out using Origin® 2021 software from OriginLab Corporation in Northampton, USA.

Gaussian Network Model analysis

We utilized *ProDy* for normal mode analysis, specifically applying the Gaussian Network Model (GNM) to examine the *esrapid*[2] structure obtained at ambient temperature and a previously known structure under cryogenic conditions (PDB ID: 4GBN). Contact maps were established using all Cα atoms with a cutoff distance of 8.0 Å. N−1 nonzero normal modes were computed, where N denotes the total Cα atom count. Squared fluctuations were then determined for both the weighted 10 slowest and 10 fastest modes for each structure, facilitating a comparison of their overall and local movements, respectively.

Primary Hippocampal Cell Culture and Ca^{2+} imaging

Newborn mice (aged 0-3 days) were euthanized following ethical guidelines. The decapitation procedure was conducted within a laminar flow cabinet, and the brains were carefully extracted and placed into a dissection medium. Subsequently, hippocampi were isolated under a stereo microscope, transferred to an enzyme medium, and incubated at 4°C for 45 minutes. Following this enzymatic treatment, the hippocampi were mechanically dissociated using progressively smaller glass Pasteur pipettes. The resulting homogenized solution was then moved to an enzyme inhibition

medium and held at 4°C for 15 minutes. Cells in the enzyme inhibition medium were subsequently centrifuged at 1000 rpm for 5 minutes, and the resulting cell pellet was supplemented with 700 µl of a sustaining medium. Careful pipetting was conducted to ensure a homogeneous distribution of cells within the medium. The cells were subsequently moved to six small Petri dishes at a density ranging from 15,000 to 20,000 cells/mm2, following a predetermined plan involving three applications of esrapid and three applications of commercial insulin. After seeding, the cells were incubated in a controlled environment at 37°C with 5% CO_2. On the second day of the cultures, a viral solution containing 10^{12} gc/ml was added to each culture dish. After that, half of the existing medium was replaced with freshly prepared sustaining medium every three days. Introducing a plasmid into the cellular milieu was imperative to facilitate GCaMP production in the cells. For this purpose, an adeno-associated virus (AAV) with a synapsin promoter known for its effectiveness in neurons was selected. GCaMP expression commenced one week after introducing this virus to the cultures. Experimental procedures were initiated on the 20th day, following the evident expression of GCaMP and the maturation of the neural network.

The electrical (calcium) activities within hippocampal neural networks were documented using a fluorescence microscope. Imaging procedures involved using 10x and 20x objectives, with photographs captured at a frequency ranging from 5 to 10 frames per second, contingent on the light intensity requirements. These images were recorded, and the video files were processed using Fiji software. Cell bodies identified in the photos were designated as Regions of Interest (ROIs). Assuming a uniform light change within each ROI, the luminance alterations in the selected areas for each frame were averaged. The resulting calculated data underwent further processing and analysis using MATLAB software. After the analysis, changes in light intensity within the identified cell bodies were graphically represented over time. Given variations in the intervals of light intensity changes across samples, amplitudes were normalized to a unit interval to ensure consistency in interpreting the results. The results were subjected to analysis using an unpaired t-test

Observations

To enhance the flexibility of insulin aspart, we substituted arginine (Arg) with lysine (Lys) at residue 22 in chain B (Figure 2.16).

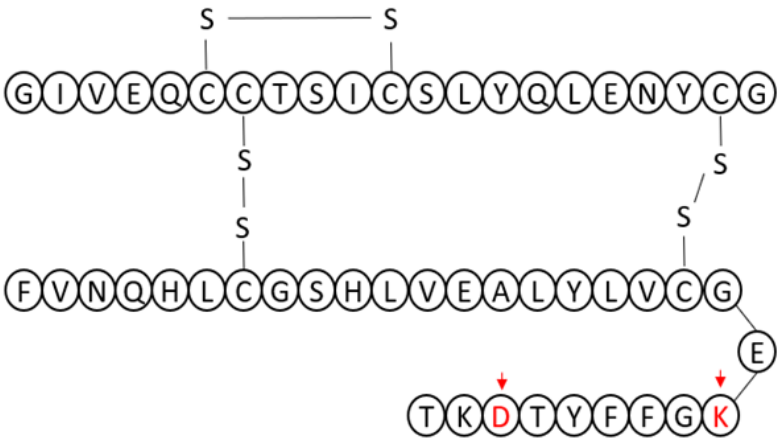


Figure 2.16 Primary sequence representation of *esrapid*^{2} insulin analog

This specific mutation was made to reduce the formation of undesired cleavage products, as trypsin has a higher affinity for arginine than lysine. Furthermore, this mutation enhances the flexibility of *esrapid*^{2} compared to commercially available aspart, resulting in *esrapid*^{2} being an ultra-acting insulin analog. *esrapid*^{2} represents the crystal structure of rhombohedral insulin, identified with the space group R3:H. The unit-cell parameter is determined to be a=b=80.3 Å, c=38 Å and the solvent content is measured at 45.06%. This structure has been successfully determined by molecular-replacement method and refined to a resolution of 2.5 Å, achieving an R factor of 30% (Table 2.4, Figure 2.17)

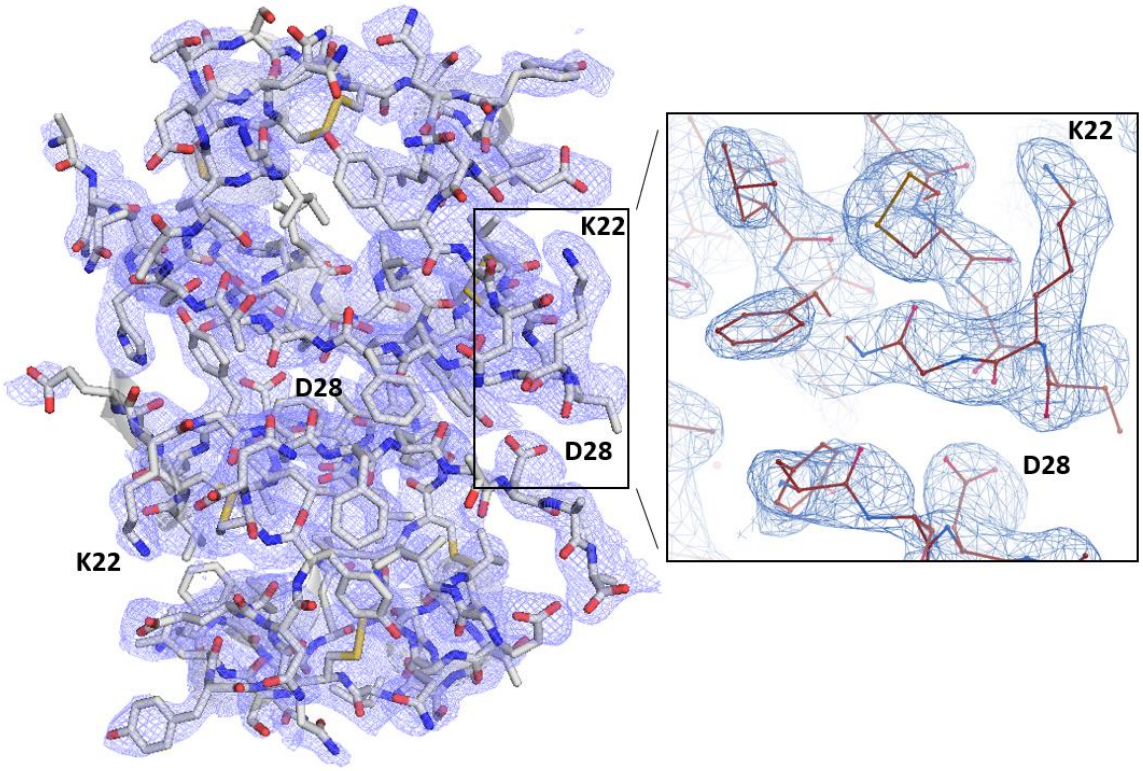


Figure 2.17 The *esrapid*^{2} insulin structure at a resolution of 2.5 Å.

The 2*Fo-Fc* simulated annealing-omit map, at the 1 sigma level, is presented in slate. The mutated residue R22 is indicated as K22, and the other mutated residue P28 is denoted as D28.

We conducted DSC measurements to explore the thermal stability of the *esrapid*^{2} (LysB22-AspB28) compared to commercially available insulin aspart (AspB28). Figure X illustrates heat flow changes relative to the melting temperature (Tm), where 50% of the protein denatures. Additionally, we calculated calorimetric enthalpy (ΔHcal) and determined conformational entropy change (ΔS) by the ratio of ΔHcal and Tm in Kelvin (K). DSC measurements unveiled significant differences in the thermodynamic properties of aspart and *esrapid*^{2}. The thermal unfolding of the aspart displayed a sharp endothermic denaturation with a Tm of 81.9 °C, ΔHcal of 35.46 KJ/mol, and ΔS of 0.099 KJ/mol·K. Conversely, *esrapid*^{2} exhibited a considerably lower Tm (77.6 °C), a significantly higher ΔHcal (49.96 KJ/mol), and a similar increasing tendency in ΔS (0.142 KJ/mol·K). The notably lower Tm suggests a less compact conformation for *esrapid*^{2}. Furthermore, the considerable increase in ΔHcal and ΔS aligns with the lower Tm, indicating a more flexible protein conformation (Figure 2.18).

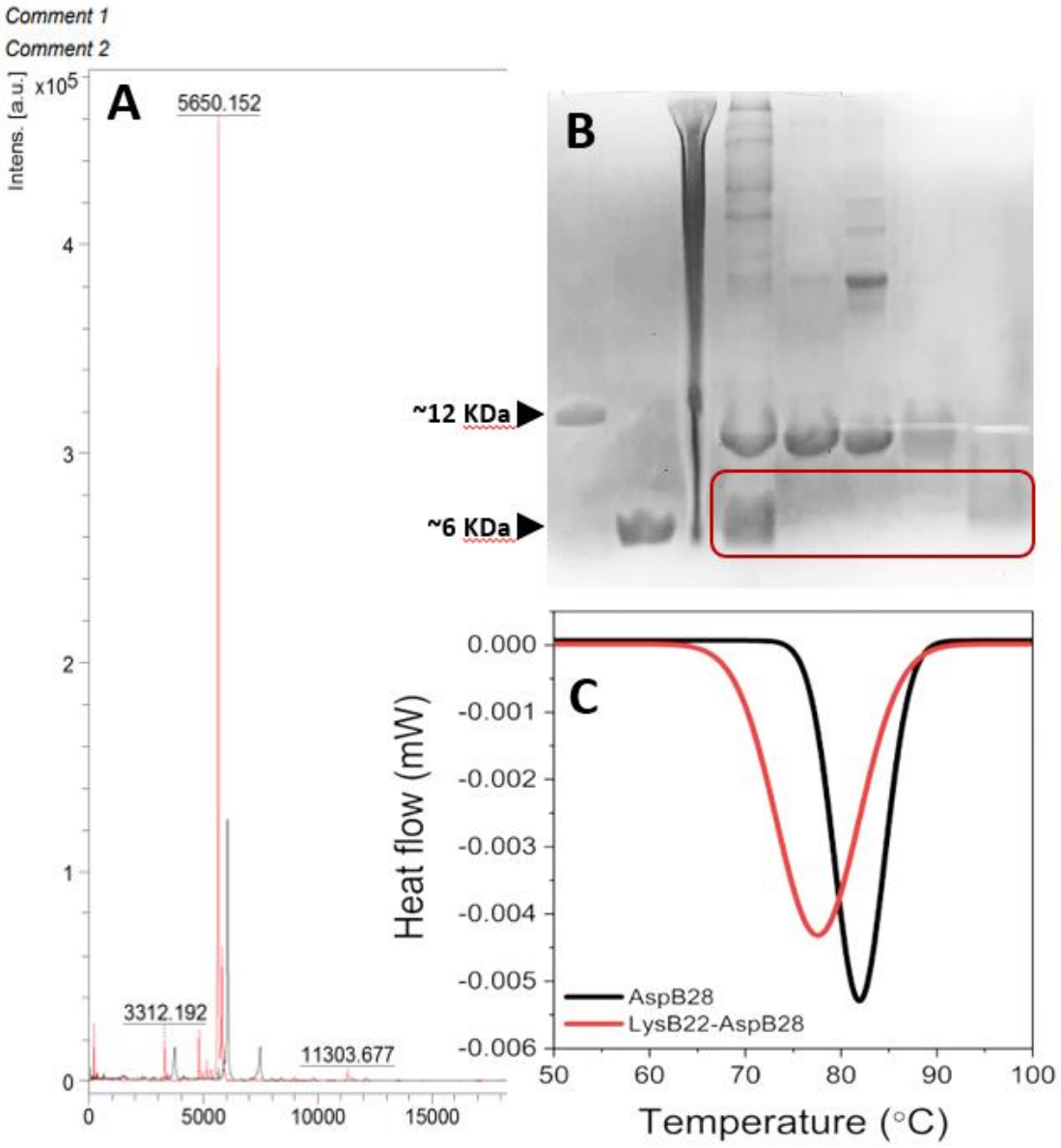


Figure 2.18 Representation the analytical characterization of the esrapid2 insulin analog.

(A) MALDI-TOF intact mass analysis reveals a molecular weight of 5.6 kDa, aligning with the size of commercially available analogs (~6 kDa). (B) 20% SDS-PAGE analysis depicts lane 1 as a marker for proinsulin at 12 kDa, lane 2 as a marker for cut insulin (*esrapid*^{2}) at 6 kDa, lane 3 as solubilized proinsulin and cut insulin, lanes 4 and 5 as purified and concentrated insulin, lane 6 as trypsin-treated proinsulin at zero time, and lane 7 as trypsin-treated cut insulin at the 4th time. (C) DSC analysis compares *esrapid*^{2} (LysB22-AspB28) and insulin aspart (AspB28).

The pilot study of this Chapter 2 also aims to delve into deciphering the structural intricacies of *esrapid*^{2} at ambient temperature, achieving a remarkable resolution of 2.5 Å from *Turkish DeLight* (Figure 2.19). The asymmetric unit cell unveiled two closely resembling monomer molecules, showcasing an RMSD of 0.81 Å. Compared with the

PDB structure of 4GBN (insulin aspart), disclosed RMSD values ranging from 0.3 Å to 0.83 Å for these akin monomer molecules within the asymmetric unit. Crystal structure determination of *esrapid*^{2} insulin at pH 7.5 was facilitated through molecular replacement involving a dimer in the asymmetric unit. Notably, amino acids 1 to 8 manifested a "T" conformation in one monomer and an "R" conformation in the other, elegantly delineated in the electron density (Figure 2.19c-d). In the structural composition of *esrapid*^{2}, the biological unit is formed by three symmetry-related "TR" dimers. The hexameric arrangement of *esrapid*^{2} is strategically organized around a crystallographic 3-fold axis aligned with the central channel. Diverging from the structure of 4GBN, a distinctive characteristic of the *esrapid*^{2} hexamer is the outward orientation of Asn3 in chain B, relative to the channel's interior. This spatial configuration results in a less stable electrostatic interaction with the two symmetry-related B3Asn (Asn 3 in chain B) residues (Figure 2.19a-b). Harmoniously, esrapid# and 4GBN crystal structures display congruence in both Cα atoms and side chains. While there are minor positional variations, notable differences are primarily observed in the side chains of 4GBN insulin compared to *esrapid*^{2}. Upon overlaying the current *esrapid*^{2}insulin structure onto 4GBN in the $T_3R_3^f$ conformation, dissimilarity emerges in the region spanning amino acids 1 to 8 in chain B. This dissimilarity is evident in backbone and side-chain conformations, proving that *esrapid*^{2} insulin adopts the T_3R_3 conformation with a pure TR-state dimer configuration. Conducting a pairwise global structural alignment between T_3R_3-esrapid and -4GBN insulin aspart structures emphasizes their similarity, with a global RMSD slightly exceeding 0.8 Å. This consistency signifies a coherent outcome attributable to the zinc-bound form.

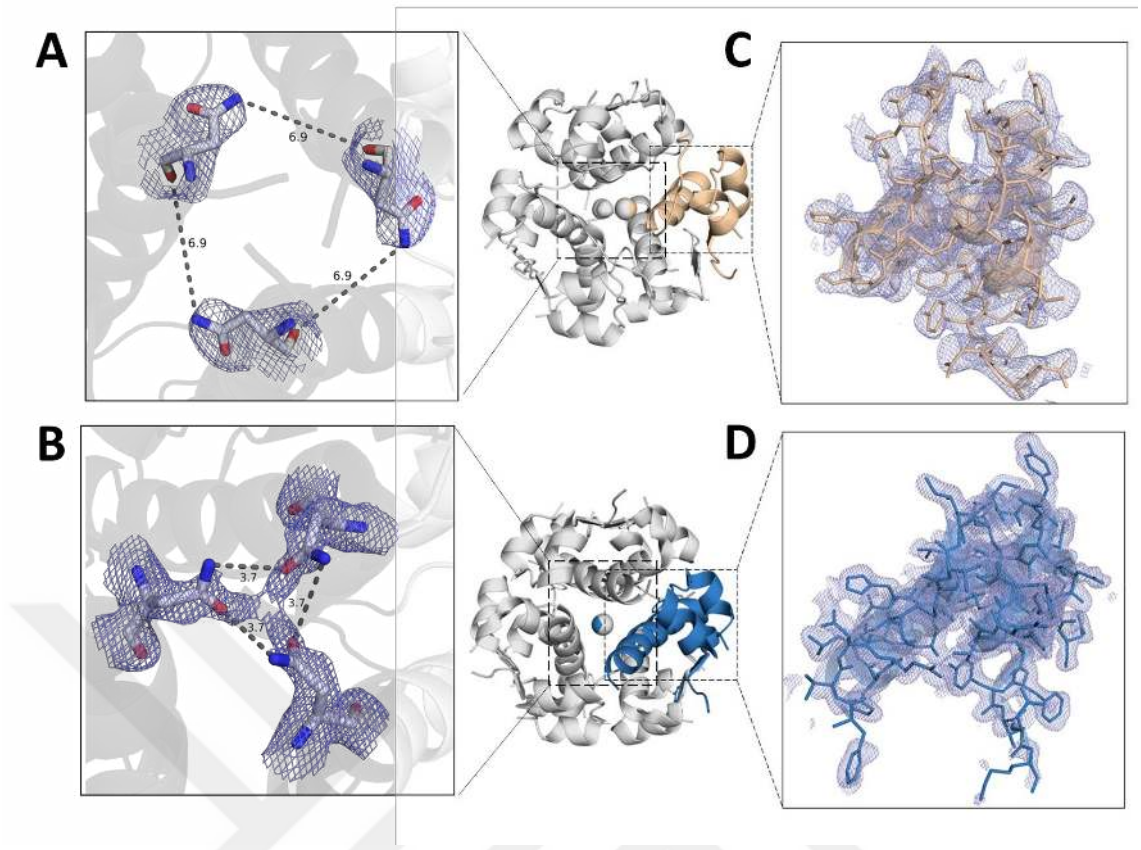


Figure 2.19 Detailed comparison the structural arrangements of esrapid2 insulin analog and insulin aspart (4GBN).

Examination of the spatial arrangement of B3As residues in the *esrapid*[2] insulin analog (A) and the 4GBN structure (B). Illustration of the monomer representation of *esrapid*[2] (C) and 4GBN (D) hexamer structures.

A distinct B-factor analysis was executed for the monomers within the structures of 4GBN and *esrapid*[2]. In examining the R-forms of both *esrapid*[2] and 4GBN, we observe slightly elevated fluctuations in the N-terminal segment (Asn21-Gln25) and Gly41-Glu42 of *esrapid*[2] B chain compared to their counterparts in the 4GBN structure. Similarly, within the T-forms of *esrapid*[2] and 4GBN, the regions spanning Cys40-Glu42 in *esrapid*[2] exhibit notable flexibility relative to the T-form of the 4GBN structure. This suggests that T-forms display increased fluctuation compared to R-forms. Moreover, substituting B22R with B22K may induce a flexible state despite the limited influence of $ZnCl_2$ on the hexameric form (Figure 2.20).

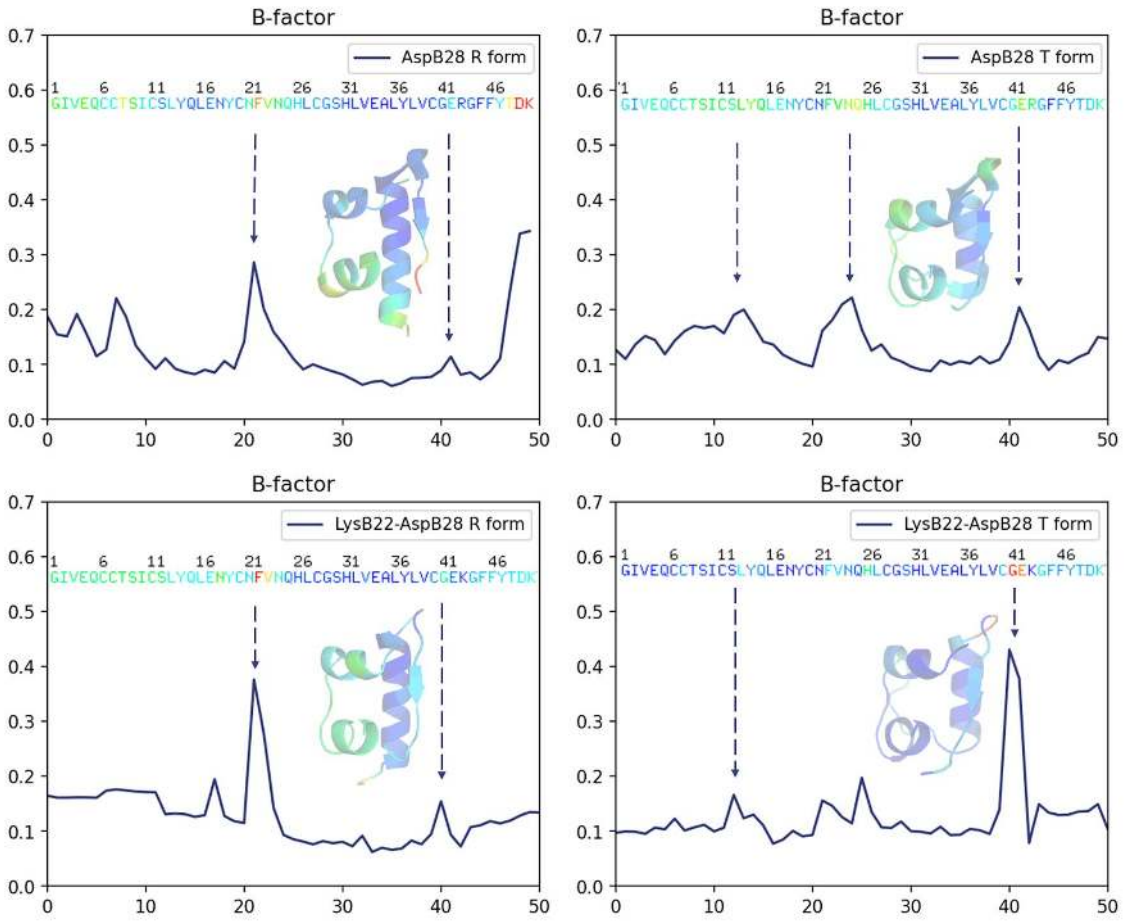


Figure 2.20 Illustration of a comparative depiction the experimental B-factor analysis for both esrapid2 (LysB22-AspB28) and insulin aspart (AspB28).

The R-forms of both insulin analogs demonstrate lower stability compared to the T-forms of the respective insulin analogs. Within the insulin dimers, the 20th and 42nd residues function as pivotal hinge sites.

Moreover, we measured the uptake of deuterium in various regions of *esrapid*^{2} and insulin aspart, identified areas of disparities, and clarified the differences in their local dynamics. In conducting HDX-MS reactions, we specifically chose the monomeric state of insulins to determine variations in monomer dynamics under uniform experimental conditions. Consequently, although significant dynamic differences were not observed in the mutated region of these insulins, noteworthy distinctions were detected in the dynamics of the remaining monomeric regions. We identified differences in electrostatic and hydrophobic properties consistent with the structural dynamics (Figure 2.21).

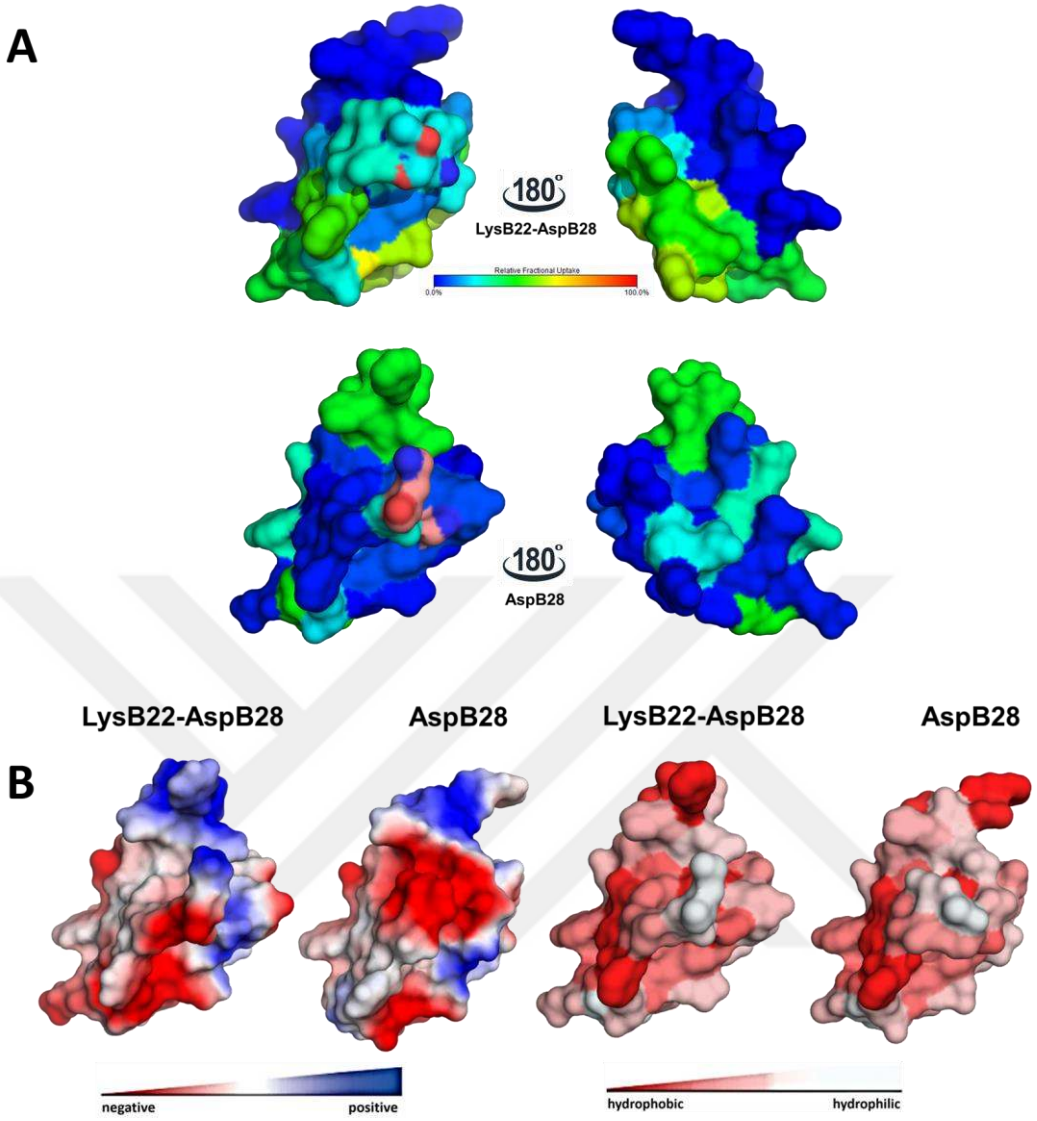


Figure 2.21 Illustrating the fluctuations in monomer dynamics.

(A) Presentation of deuterium uptake in different regions of *esrapid*[2] and insulin aspart. (B) Analysis of electrostatic potential and hydrophobicity in *esrapid*[2] and insulin aspart.

Aspart has been chosen for the standard reference for assessing deuterium incorporation, undergoing the HDX reaction, and subsequent analysis via LC/MS in the initial stage. To maintain consistent pH conditions throughout the exchange reaction, we chose Tris-HCl buffer at pH 9, given its capability to sustain insulin's *in vivo* monomeric state. The chronological progression of deuterium incorporation profiles by aspart is clarified in Figure 2.21, accompanied by a comparative evaluation of their hydrophobicity and electrostatic potential. During the HDX reaction involving the aspart monomer, specific residues, including Cys11-Leu13, Asn21-Cys28, Glu34-

Tyr37, Arg43-Gly44, and Asp49, display a notable level of relative flexibility compared to other residues (Figure 2.22a), aligning with outcomes derived from computational analyses (Figure 2.22b). Our normal mode analysis highlights Ile2, Cys6, Ile10, Gln15-Leu16, Asn18-Tyr19, Leu27-Cys28, Glu34-Leu36, Leu38-Val39, and Arg43-Gly44 residues as regions undergoing dynamic dissections, termed *kinetically hot regions* (Figure 2.22b). These identified hinge sites from our computational analysis are compatible with observations in the HDX-MS heatmap results. Ser12 and Gln15-Leu16 residues consistently emerge as pivotal hinge regions solely in GNM analysis.

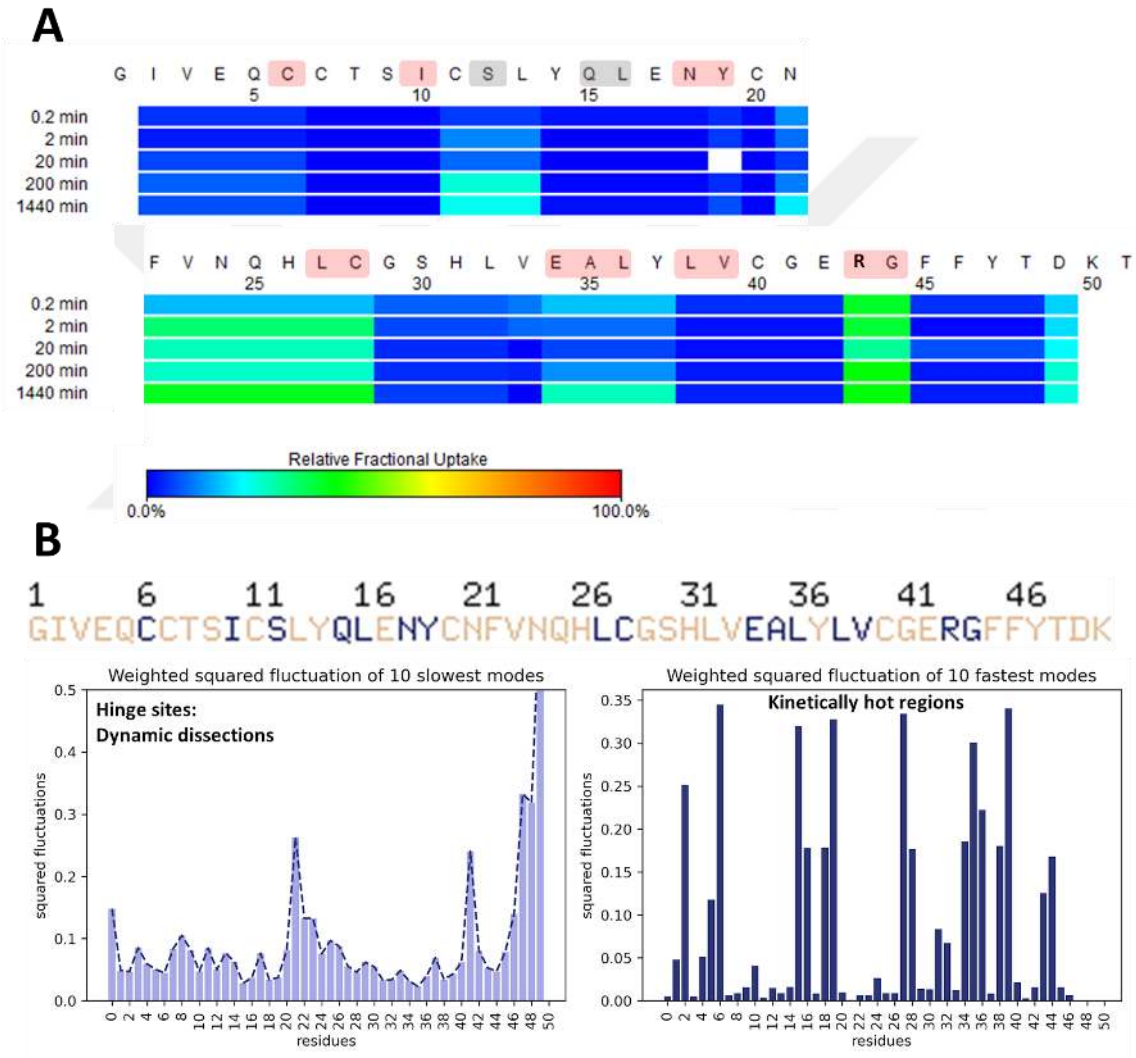


Figure 2.22 HDX reaction and normal mode analysis of insulin aspart.

(A) Illustration of deuteration levels monitored at 0.2, 2, 20, 200 minutes, and 24 hours. The H/D exchange of insulin provides insights into overall solvent accessibility. (B) Depiction of global protein motions from the slowest 10 modes and fastest 10 modes of GNMs. The normalized results of weighted squared fluctuations of insulin aspart are plotted. The residues highlighted in red from the relative fractional uptake analysis in panel A align with the hinge sites and kinetically hot regions identified in the normal mode analysis in panel B.

We compared HDX reactivity in *esrapid*^{2} insulin analogs and insulin aspart to assess potential distinctions. The *esrapid*^{2} analog exhibited elevated exchange reactivity compared to insulin aspart, suggesting less stability of *esrapid*^{2} compared to insulin aspart analog. Within the HDX reaction of the *esrapid*^{2} monomer, specific residues, notably Ile2-Cys6, Ser9-Asn18, Cys20-Asn21, and Val39-Asp49, displayed noticeable relative flexibility compared to other residues (Figure 2.23a), aligning with computational analyses (Figure 2.23b). Our normal mode analysis identified Glu4-Cys6, Ser9-Ile10, Asn18-Tyr19, His31-Leu32, and Val39 residues as regions undergoing dynamic dissections, called *kinetically hot regions*. Notably, the hinge sites identified from our computational analysis consistently aligned with observations in the HDX/MS heatmap results. Additionally, Ala35-Leu36 and Gly44 residues emerged as pivotal hinge regions solely in our GNM analysis (Figure 2.23a).

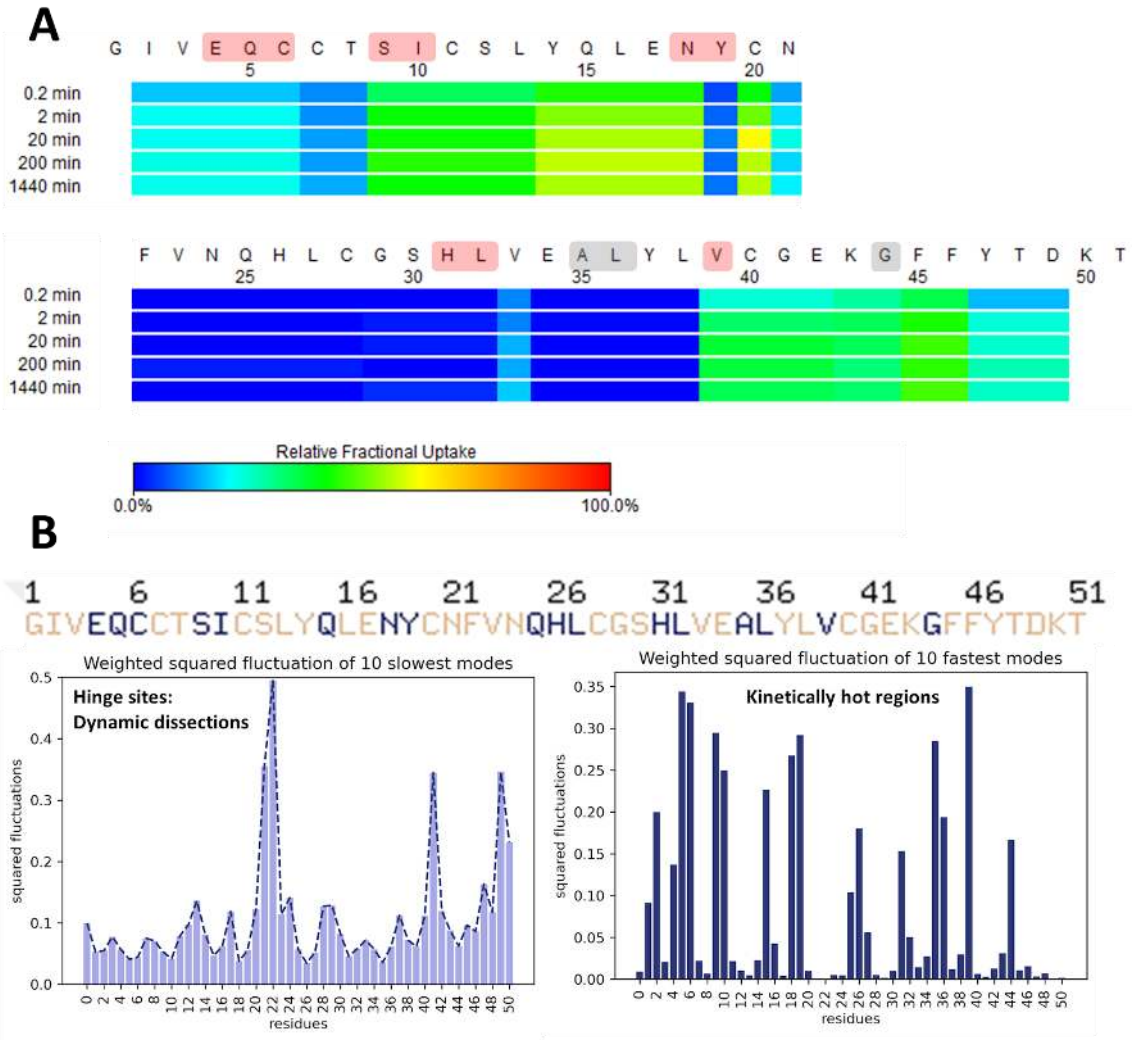


Figure 2.23 HDX reaction and normal mode analysis of *esrapid*[2].

(A) Illustration of deuteration levels monitored at 0.2, 2, 20, 200 minutes, and 24 hours. The H/D exchange of insulin provides insights into overall solvent accessibility. (B) Depiction of global protein motions from the slowest 10 modes and fastest 10 modes of GNMs. The normalized results of weighted squared fluctuations of *esrapid*[2] are plotted. The residues highlighted in red from the relative fractional uptake analysis in panel A align with the hinge sites and kinetically hot regions identified in the normal mode analysis in panel B.

The visuals in Figure 2.24 illustrate the deuteration levels for the respective fragments of *esrapid*[2] and insulin aspart throughout the 200-minute and 1440-minute time intervals. Dissimilar patterns were observed in the deuterium uptake profiles of most fragments between *esrapid*[2] and aspart. Notably, there was only a slight similarity in the fragment spanning His31-Asp49, encompassing the mutation site in *esrapid*[2].

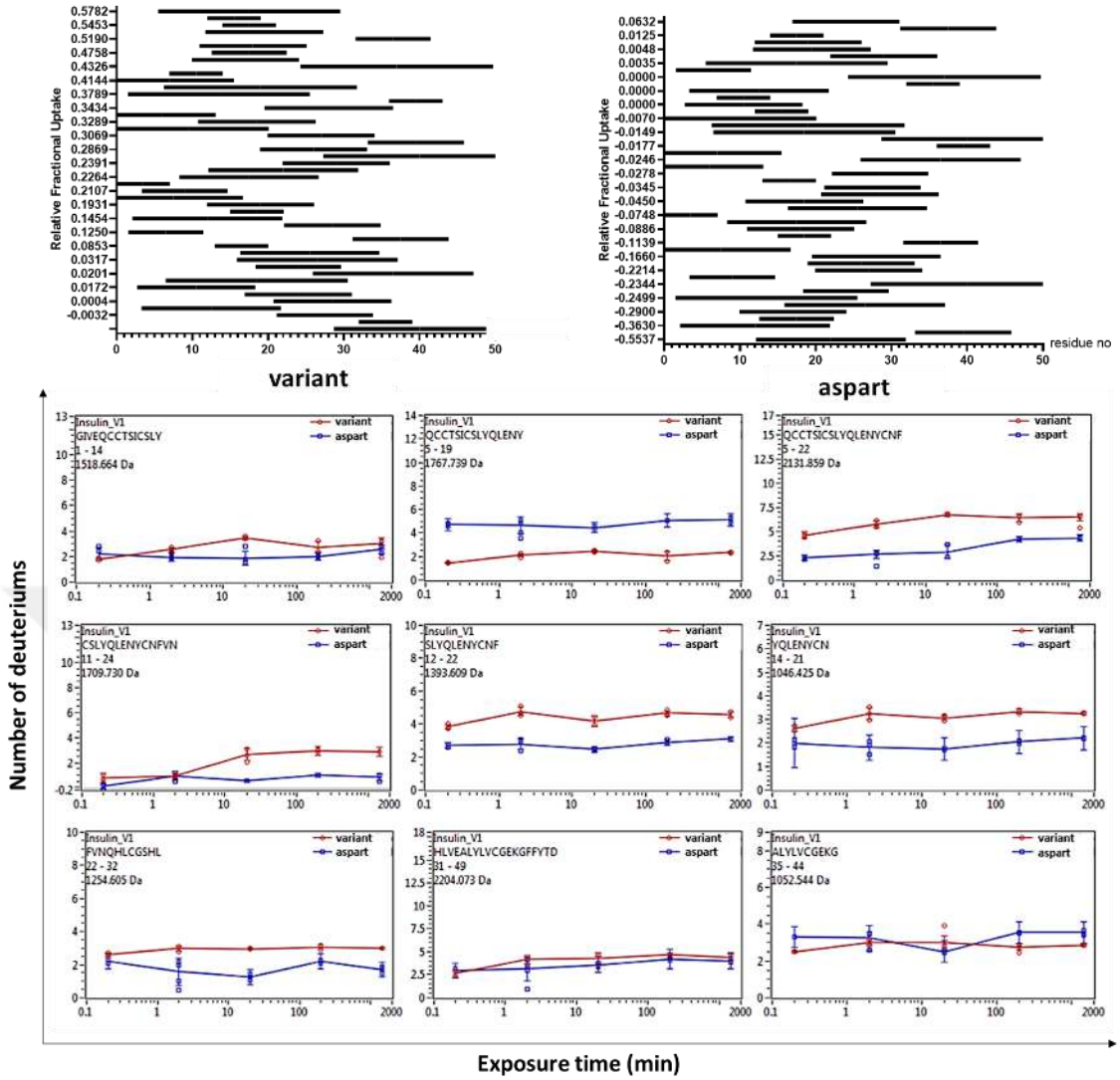


Figure 2.24 Deuteration profiles of the corresponding fragments in esrapid2 and insulin aspart over the 200-minutes (a) and 1440-minutes (b) durations.

Ca^{2+} imaging to observe esrapid^2 activation on neural cell

Calcium serves as a crucial intracellular messenger in neurons. At rest, the intracellular calcium concentration in most neurons is approximately 50–100 nM. This concentration can transiently surge to 10 to 100 times higher during electrical activity. Intracellular calcium signals play a regulatory role in processes spanning a range of periods, from neurotransmitter release occurring on the microsecond scale to gene transcription lasting minutes to hours. The influx of calcium triggers the release of neurotransmitter-carrying vesicles from the presynaptic terminals, while the transient elevation of calcium levels in the postsynaptic neuron induces activity-dependent synaptic plasticity. Calcium signals' temporal dynamics, amplitude, and localized extent are critical determinants of their functional impact [Bostan, 2020]. Consequently, techniques have been developed to visualize and quantitatively assess intracellular calcium signals.

The calcium imaging technique is one of the methodologies developed for real-time monitoring of neuronal network activity during active information processing. This method is grounded in the principle that neuronal activity induces swift alterations in the intracellular free calcium concentration. Genetically encoded calcium indicator proteins (GECI) are commonly employed. GCaMP, a representative of the GECI family, comprises enhanced green fluorescent protein (EGFP), calcium-binding protein calmodulin (CaM), and the CaM-interacting M13 peptide. The presence of calcium prompts conformational changes in CaM and M13, leading to an augmentation in the fluorescence intensity of the protein [Bostan, 2020].

In this study, we also administered AAV treatment to neural cells to assess the impact of insulin on cellular activity. The treated hippocampal cells were incubated with esrapid2 insulin and insulin aspart for half an hour. Subsequently, images and videos were captured using a fluorescence microscope. Our observations indicated that esrapid2 insulin demonstrated successful activity comparable to insulin aspart (Figure 2.25).

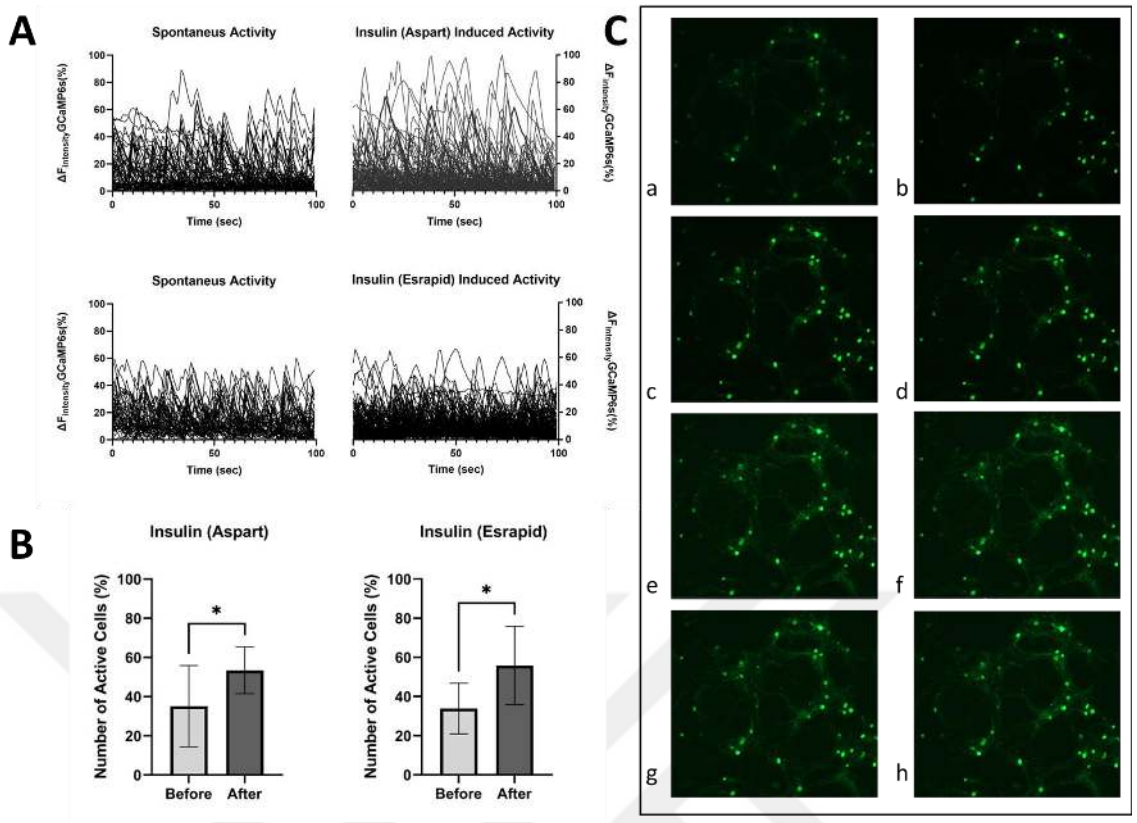


Figure 2.25 Comparison of Ca^{2+} imaging for esrapid2 and insulin aspart.

(A) Graph illustrating changes in light intensity within selected ROIs over time, derived from the recording of calcium activity. (B) The effects of esrapid2 and insulin aspart based on the percentage of active neural cells graph conducted using an unpaired t-test. *p*-value for aspart: 0,0358 (n=2); *p*-value for esrapid2: 0,0473 (n=2). (C) Step-by-step depiction the propagation of Ca^{2+} activity upon esrapid[2] stimulation. a) Network image before stimulation b-h) Dissemination of calcium activity throughout the entire network with corresponding annotations.

Concluding remarks

Crystallographic and NMR techniques offer intricate insights into the atomic-level structures controlling the oligomeric interactions of insulins. Meanwhile, HDX technique sheds light on monomeric insulin dynamics under neutral conditions in solution. In this study, we introduced an ultrafast/less stable insulin variant, *esrapid*[2]. The crystal structure of this hexameric insulin variant was determined at a resolution of 2.5 Å at ambient temperature and confirmed the distinctions with the 4GBN structure. An HDX comparison between insulin aspart and *esrapid*[2] revealed that *esrapid*[2] incorporates deuterium atoms more rapidly and extensively than intact insulin aspart, consistent with the diminished stability of our modified analog. GNM analysis further

supported these kinetic differences. The observed variations initiated the analysis of the regions responsible for the discrepant HDX reactivity. Identifying these regions revealed dynamic characteristics contributing to the observed differences in HDX reactivity, as supported by GNM analysis. These deuteration disparities and computational analysis imply that the R→K substitution induces distinct kinetic models, even under identical physiological conditions.

Zinc-bound hexamer insulins display characteristic structural features inherent in globular proteins [Chothia et al., 1983]. Serving as a pivotal model in the evolution and application of X-ray crystallographic techniques [Baker et al., 1988], the hexamer functions as a stable reservoir of the hormone within both the secretory granules of pancreatic β-cells and pharmaceutical formulations [Brange et al., 1987]. The integration of structure-based mutagenesis into the zinc insulin hexamer [Brange et al., 1988; Dodson and Steiner, 1998; Steiner, 1967] has significantly advanced the development of rapid-acting insulin analog formulations suitable for prandial injection and pump use [Zinman et al., 1997]. Despite their initial categorization as “monomeric” [Brange et al., 1988], these analogs are presently designed either as zinc insulin hexamers or as an equilibrium distribution of zinc-free oligomers [DeFelippis et al., 2001; Garg et al., 2005]. The overarching objective of this protein engineering is to optimize glycemic control while mitigating the risk of major hypoglycemic events per decrement in hemoglobin A1c [Renner et al., 1999; Melki et al., 1998; Hartemann-Heurtier et al., 2003]. Although insulin analogs represent a significant advancement in the pharmacotherapy of diabetes and rapid-acting insulins offer specific benefits compared to regular insulin, a requirement still exists to formulate ultrafast-acting insulin preparations closely mimicking the physiological release of insulin.

This pilot study aims to examine a modified and active (Figure2.25) insulin analog, *esrapid*[2], alongside the insulin aspart to illuminate the fundamental structure and dynamics of monomeric rapid-acting insulin in a solution. Insulin aspart incorporates a substitution, replacing proline with aspartic acid near the C-terminus of the B-chain. The substitution of arginine with lysine at position B22 and proline with aspartic acid at position B28 distinguishes esrapid2, which can be assumed to be a less stable/ultra-acting insulin analog. Lysine and arginine, both positively charged essential amino acids under physiological conditions, are predominantly exposed on the surfaces of proteins [Kumar et al., 2000; Yokota et al., 2006]. These two amino acids play pivotal roles in enhancing protein stability by engaging in ionic interactions and hydrogen bonding

within proteins and interacting with water molecules [Strickler et al., 2006; Barlow et al., 1983]. Despite their shared function as basic residues, arginine contributes more stability to protein structure than lysine, which is attributed to its distinctive geometric structure. The guanidinium group in arginine facilitates interactions in three possible directions over its three asymmetrical nitrogen atoms (Nη1, Nη2, Nε).

In contrast, the primary functional group of lysine allows interaction in only one direction [Borders et al., 1994; Donald et al., 2011]. This capability enables arginine to form more electrostatic interactions, including salt bridges and hydrogen bonds, than lysine. Presumably, this results in stronger interactions than those generated by lysine. These effects have been indirectly validated by assessing salt-bridge numbers produced by arginine and lysine in numerous proteins and by analyzing the arginine-to-lysine ratio and salt-bridges in mesophilic and thermophilic proteins [Musafia et al., 1995; Matsutani et al., 2011; Kumar et al., 1999; Strub et al., 2004; Chan et al., 2011]. Beyond geometric considerations, arginine's ionic interactions have demonstrated more excellent stability potential than lysine's [Turunen et al., 2002]. The solution structure of *esrapid*^{2} closely mirrors the crystallographic R state. However, unlike the insulin aspart structure, a distinct feature of the *esrapid*^{2} hexamer in the R form involves the outward orientation of Asn3 in chain B relative to the interior of the channel. This spatial configuration leads to a less stable electrostatic interaction with the two symmetry related B3Asn residues (Figure 2.19). Additionally, according to B-factor analysis, the unique R-state of the *esrapid*^{2} crystal structure is evident in a flexible *esrapid*^{2} monomer, which correlates with the HDX/MS findings for the *esrapid*^{2} monomer. Characterizing *esrapid*^{2} and insulin aspart as native-like monomers at pH 9.0 and without an organic co-solvent provides an opportunity for a quantitative assessment of amide-proton exchange in D_2O.

Consequently, to monitor deuteration levels and analyze the higher-order structure of *esrapid*^{2} compared to insulin aspart, HDX-MS experiments were carried out with five time points ranging from 12 seconds to 1440 minutes. HDX-MS data for *esrapid*^{2} indicated swift deuterium incorporation at the initial time points in most peptides, followed by local exchange after specific time points. In contrast, insulin aspart exhibited no deuterium incorporation in the first 35 residues, with only the Phe22-Cys28 and Arg43-Gly44 peptides showing deuterium uptake and local exchange at specific time points (Figure 2.22 and Figure 2.23).

The formation of insulin oligomers primarily relies on establishing noncovalent

bonds between interacting residues from adjacent monomers. However, alterations in pH can interrupt these interactions, causing residues in distant sites to cease their communication paths and transition into monomeric states. Modifying monomeric states results in substantial modifications to the customary residue communication network compared to oligomeric states. These alterations can also be elucidated through computational GNM analysis, analogous to our insights from experimental HDX reactions. NMA was conducted in this study to explain the distinctive dynamics exhibited by the esrapid2 monomer compared to the monomeric form of insulin aspart, aiming to support the accuracy of HDX reactions. Employed monomeric insulin aspart structure as the reference model, analogous to the methodology used in HDX reactions. NMA-based methodologies precisely assess global protein motions at lower frequencies and local motions at higher frequencies. In the esrapid2 monomer, pivotal hinge sites crucial for insulin activity are identified in residues Cys6, Ser9, Asn18-Tyr19, and Leu32, as observed through GNM and HDX/MS analyses (Figure 2.22 and Figure 2.23).

Conversely, in the aspart monomer, hinge sites critical for insulin activity are discerned in residues Cys6, Ile10, Asn18-Tyr19, Cys28, Glu34, Leu38, and Arg43-Gly44 (Figure 2.22). Remarkably, these residues exhibiting dynamic dissections are intricately linked to regions involved in receptor interactions. The changes we observe in the behavior of the zinc-bound hexamer, both through experiments and computer simulations, suggest that the hexamer remains relatively unaffected. However, substituting the amino acid arginine (R) with lysine (K) in a monomer causes noticeable alterations in how the insulin behaves over time. These changes in behavior depend on specific flexibilities, indicating that the R→K substitution in a monomer triggers different kinetic models or patterns of movement based on the inherent flexibility of the protein structure.

Due to the tendency of conformational fluctuations to accelerate the degradation of pharmaceutical formulations, implementing a paradigm called 'dynamic restructuring' for insulin may provide a strategic framework for the design of formulations characterized by ultrafast action, considering these kinetic disparities.

2.2.6. A pilot study for esrapid insulin analog

Experimental and computational analysis of esrapid analog

esrapid represents a novel designer insulin prototype designed to meet the criteria for both "ultra rapid-acting" and "intermedia-acting" insulin analogs. The strategic amino acid substitutions, involving the shift from asparagine to glycine in the B chain, maintain the isoelectric point (pI) of the molecules near that of native insulin. Preferably, the monomeric form of the esrapid analog demonstrates enhanced stability in a 20 mM citric acid environment. The monomeric esrapid promptly functions its action mechanism upon injection by directly binding to the insulin receptor. Introducing small quantities of zinc (30 pg/mL) into the monomeric esrapid formulation facilitates the preferential formation of hexamers, further supporting the molecule's stability and contributing to a delayed absorption time. Consequently, the formation of micro precipitates at the injection site results in a more gradual hexameric esrapid absorption than its monomeric esrapid into the circulation. Our understanding the kinetics of insulin monomers remains limited. Investigating the comparative dynamic mechanisms inherent to monomeric insulin appears to be a pivotal strategy to address this gap. A comprehensive understanding of insulin's monomer dynamics becomes critical for solving crucial regions involved in insulin-receptor interactions. This pilot study in Chapter 2 seeks to unveil the underlying proliferative mechanism of the esrapid monomer compared to insulin aspart. Dynamic changes in the cryogenic esrapid monomer structure at pH 3, obtained from the SPring-8 facility, are explored by contrasting it with previously published structures of porcine insulin at different pH environments. For proliferative effect comparison, insulin aspart served as the model molecule. In the context of pH-dependent dynamic comparisons, esrapid at pH 3 and previously published structures are docked with the insulin receptor, and pairwise RMSD analysis is performed between this structure and previously published structures spanning the pH range of 5-9.

In-vitro evaluation of esrapid insulin analog in human iPS cell culture

Human induced pluripotent stem (iPS) cell lines were developed from peripheral blood mononuclear cells (PBMCs) using an integration-free Sendai virus to introduce the Yamanaka factors, as detailed in a prior report [Yu et al., 2007]. These cell lines

underwent expansion on matrigel with mTeSR1 (StemCell Technologies) and were passaged using 0.5 mM EDTA. Subsequently, all cells were thawed and acclimatized to TeSR-E8 for at least 2 weeks before initiating experiments. *esrapid* protein, as detailed in Section 2.2, has been successfully produced. For insulin-treated group experiments, hiPS cells were cultured in a controlled incubator using an E8 medium supplemented with specific factors conducive to their growth. This medium comprised 100 U/ml penicillin, 100 µg/ml streptomycin, 20 mg/ml transferrin, 3 mM sodium selenite, 100 mg/ml ascorbic acid, 1 mg/ml FGF2-G3, 0.5 µg/ml TGF-Beta3, and 20 mg/ml of the novel designer *esrapid* insulin analog or commercially available insulin aspart (NovoRapid). *esrapid* protein, as detailed in Section 2.2, has been successfully produced. The cells were maintained by 5% CO2 at 37°C. The progression of the cell culture underwent meticulous monitoring at distinct time points, specifically on day 1, day 5, and day 7. To document and observe cellular developments, the Mateo TL Digital Transmitted Light Microscope was employed (Figure 2.26). Compared to insulin aspart, the *esrapid* insulin analog demonstrated a comparable growth-promoting effect on human iPS cells.

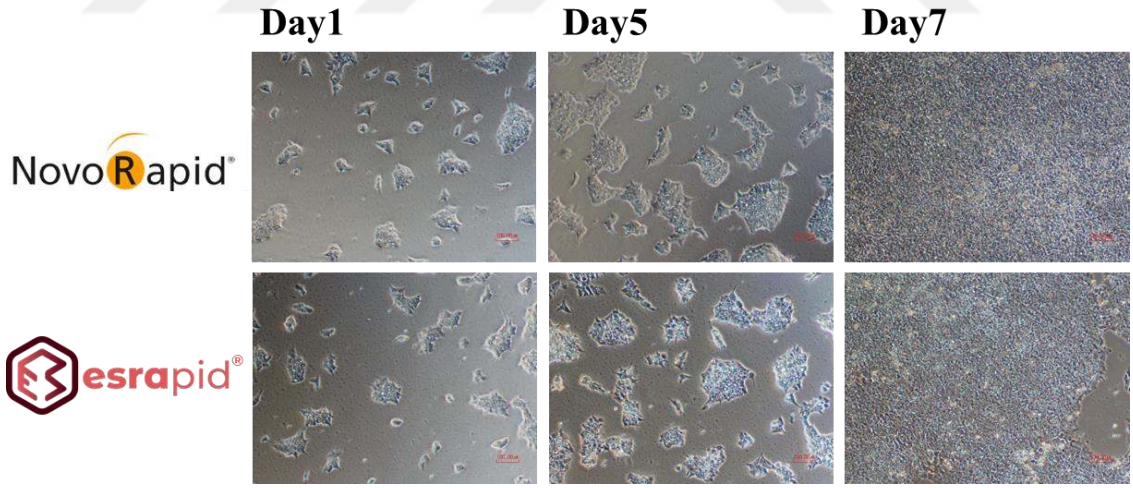


Figure 2.26 Representation comparing the effects of esrapid and insulin aspart on the hiPS cell line at different time points: day 1, day 5, and day 7.

hiPSCs can continuously expand and differentiate into various lineages within the human body. These distinctive features make hiPSCs a continuing cell source for developmental investigations, pharmaceutical screening, and applications in

regenerative medicine [Park et al., 2008]. Furthermore, undifferentiated hiPSCs can replicate characteristics of genetic diseases, mainly when the implicated cellular and molecular pathways are active within the pluripotent cell state [Karagiannis et al., 2019; Steichen et al., 2019; Aboul-Soud et al., 2021]. Since the initial characterization of human-induced pluripotent stem cells [Baghbaderani et al., 2016; Clavellina et al., 2023], there has been a burgeoning interest in determining the fundamental constituents of their culture milieu. As an essential component, insulin is a peptide hormone that governs a complex network of signaling pathways and cellular responses. Consequently, it has been recognized as crucial in cultivating human pluripotent stem cells. It has also been recognized as a pivotal ingredient in stimulating human pluripotent stem cells, essential to sustain their pluripotent state [Baghbaderani et al., 2016; Clavellina et al., 2023].

Conducting a comprehensive investigation into the effect of insulin on the seeding dynamics of hiPSC aggregates [Shahbazi et al., 2019; Li and Geng, 2010], colonies were isolated using EDTA and subsequently introduced to culture with or without insulin supplementation. The adhesion of hiPSC aggregates was assessed through microscopic analysis, employing crystal violet staining of plates conducted one day post-seeding. The outcome revealed a significant enhancement in the seeding efficacy of hiPSC aggregates in the presence of insulin. Cellular cultures were maintained with or without insulin for three days to elucidate the impact of insulin on the cell cycle. After this incubation, a cultivated flow cytometry-based cell cycle assay incorporating propidium iodide (PI) staining was executed. A noticeable reduction in the proportion of cells residing in the G1 phase was noted in the absence of insulin. This observation suggested a productive influence of insulin on both the initial seeding process and the subsequent proliferation of hiPSCs [Shahbazi et al., 2019; Li and Geng, 2010]. Conversely, in the absence of insulin, distinct alterations in the cellular cycle profile were apparent, accompanied by an elevated incidence of apoptosis in hiPSCs. Furthermore, the investigation verified the phosphorylation of critical components within the insulin signaling pathway when insulin was present, thereby highlighting the noticeable impact of insulin on directing the mRNA transcriptome of hiPSCs [Shahbazi et al., 2019; Li and Geng, 2010]. To summarize, insulin appeared as a critical regulatory factor influencing seeding efficiency and proliferation, phosphorylation events, and mRNA transcriptome modulation in hiPSCs.

Cryogenic atomic resolution of esrapid at the RIKEN SPring-8 synchrotron

Sample preparation

esrapid protein, as detailed in Section 2.2, has been successfully produced. The crystallization process involved utilizing sitting-drop microbatch screening under oil, employing 72-well Terasaki crystallization plates. The protein, maintained in a 20 mM citric acid buffer at pH 2.5 and a concentration of 5 mg/ml, was subjected to crystallization by combining it with approximately 3500 commercially available sparse matrix crystallization screening conditions at a 1:1 ratio (v/v) at room temperature. Optimal crystals were obtained in MembFac (#29), featuring a composition of 0.1 M ammonium sulfate, 0.1 M HEPES sodium at pH 7.5, 0.5 M Sodium phosphate dibasic dihydrate, and 0.5 M Potassium phosphate dibasic. These crystals were subsequently sealed with 100% paraffin oil (Figure 2.27).

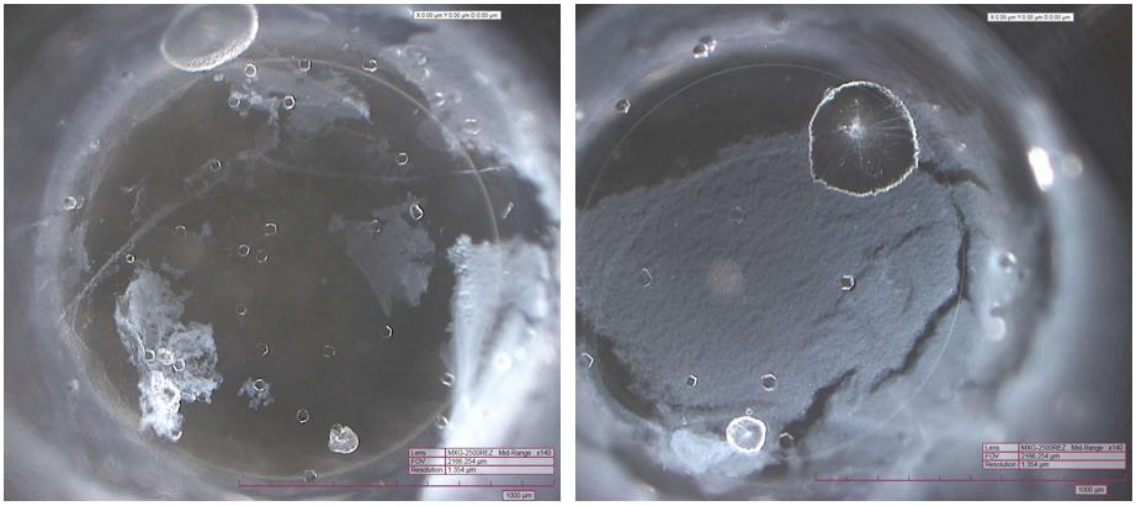


Figure 2.27 *esrapid* crystals.
Crystalization buffer: 0.1 M ammonium sulfate, 0.1 M HEPES sodium pH 7.5, 0.5 M sodium phosphate dibasic dihydrate,0.5 M potassium phosphate dibasic buffer.

Data collection and processing

Data was collected using the *ZOO* system [Hirata et al., 2013], which was designed for the BL32XU beamline at SPring-8. This microbeam beamline provided a variable beam size, ranging from 1 x 1 μm to 10 x 15 μm (horizontal x vertical), with a flux

density of 10^{10} photons μm-2 s-1. Crystal handling procedures involved the utilization of SPACE [Murakami et al., 2012], a robotic sample changer devised at SPring-8. SPACE could accommodate eight Unipucks (Crystal Positioning Systems, New York, USA) within a liquid-nitrogen dewar. Distinct identifiers were assigned to each Unipuck within the *SPACE* server, and these identifiers were subsequently cross-referenced with BSS and *ZooNavigator* datasets. A coaxial microscope from Union Optical (Tokyo, Japan) was utilized to observe micrometer-sized crystals in alignment with X-rays. The microscope image server captured crystal scenes upon receiving commands via a network socket. The high-precision crystal goniometer from *KOHZU* Precision (Kanagawa, Japan) facilitated the precision of crystal manipulation. This goniometer included 3D translation axes and an air-bearing spindle axis with a sphere of confusion smaller than 1 μm. The speed of each translation axis, powered by stepper motors, achieved a maximum of 500 μm s-1 with a resolution of 0.1 μm. The EIGER X 9M detector (Dectris, Switzerland) was employed for data collection. This detector featured a sizable area of 233.2 x 245.2 mm (width x height) and operated at a maximum frame rate of 238 Hz. The data collection process utilized the *'fast raster scan'* system and the rapid spot-finder program *SHIKA*. *HITO* autonomously implemented an appropriate data-collection strategy for each crystal by analyzing spatial relationships among crystals attached to the same loop. Simultaneously, the datasets underwent processing via the automated data-processing software *KAMO*, facilitating the organization of sets for subsequent merging. Heavy-atom site identification, phasing, and phase improvement procedures were executed using the *SHELXC*, *SHELXD*, and *SHELXE* programs.

Structure determination

The structural analysis of *esrapid* was conducted under ambient conditions, utilizing the I213 space group, and the automated molecular replacement tool *PHASER* integrated into the *PHENIX* software package [McCoy et al., 2007]. The initial search model was derived from a previously published X-ray structure (PDB ID: 5VIZ), and its coordinates were employed for the preliminary rigid-body refinement within *PHENIX* . Subsequent refinement steps included simulated-annealing refinement, individual coordinate adjustments, and the application of Translation/Libration/Screw (TLS) parameters. Additionally, the refinement process incorporated a composite omit map

embedded in *PHENIX* to identify potential locations of altered side chains and water molecules. The final model underwent thorough examination and reconstruction in *COOT* [Emsley and Cowtan, 2004], focusing on positions displaying significant difference density. Superfluous water molecules located outside regions of notable electron density were manually excluded. Visualization of the X-ray crystal structure was generated using *PyMOL* [Delano, 2002] and *COOT*. The statistical data for the *esrapid* insulin analog are presented in Table 2.5.

Table 2.5 Data collection and refinement statistics of esrapid.

Data collection	**esrapid**
Instrument	SPring-8
Resolution range	24.66 - 1.36 (1.409 - 1.36)
Space group	I 21 3
Unit cell	77.99 77.99 77.99 90 90 90
Multiplicty	2.0 (2.0)
Completeness (%)	99.57 (97.04)
Mean I/sigma(I)	49.24 (0.10)
R-merge	0.01123 (6.495)
R-meas	0.01589 (9.185)
R-pim	0.01123 (6.495)
CC1/2	1 (0.602)
CC*	1 (0.867)
R-work	0.2053 (0.3172)
R-free	0.2186 (0.3365)
CC (work)	0.952 (0.727)
CC (free)	0.950 (0.700)
Protein residues	50
RMS (bonds)	0.015
RMS (angles)	1.12
Ramachandran favored (%)	100.00
Ramachandran allowed (%)	0.00
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.00
macromolecules	28.77
ligands	47.48
solvent	41.54
macromolecules	28.77

Docking and pairwise studies for esrapid insulin analog

Structural coordinates

The *esrapid* insulin structure was collected at pH 3, and the coordinates of *esrapid* were utilized for comparison with the Diao (2003) study. The protein coordinates were sourced from the Protein Data Bank (PDB). The accession number for human insulin (hIns) at pH 5 is 1B17, at pH 6 is 1B2A, at pH 7 is 1B2F, and at pH 9 is 1B2G. And also 6VEP insulin receptor structure was used for insulin docking study [Diao, 2003].

Docking protocol and scoring of insulins with insulin receptor

The docking approach is divided into three key phases: (i) rigid-body docking, (ii) semi-flexible refinement, and (iii) final refinement in explicit solvent. Five insulin structures in the cubic phase and an insulin receptor structure were employed to facilitate comparison. We generated 200 structures for each insulin-insulin receptor combination from starting structure ensembles in the initial rigid-body docking stage. Each docking attempt was repeated ten times. The solution with the lowest *HADDOCK* [De Vries et al., 2010] score has been selected for further. To account for variations at different pH levels, an ensemble of one insulin receptor structure was utilized for each insulin, creating 1000 rigid-body. In the semi-flexible refinement phase, structures with the highest potential—comprising 36% at pH 3, 15% at pH 5, 36% at pH 6, 29% at pH 7, and 27% at pH 9—were selected from the initial rigid-body docking results based on their *HADDOCK* scores. Subsequently, these chosen structures underwent an extensive three-part refinement process. The concluding phase involves a subtle refinement for the pH 3 and pH 9 clusters, representing 82% of the water-refined models generated by *HADDOCK* [De Vries et al., 2010]. The pH 5 clusters represent 87% of the water-refined models, while the pH 6 clusters represent 89% of the water-refined models generated by *HADDOCK*. Lastly, the pH 7 clusters represent 83% of the water-refined models produced by *HADDOCK* [De Vries et al., 2010].

Pairwise Cα-Cα distance between esrapid and insulin structure

A detailed examination of the spatial relationships among the carbon alpha atoms in the cubic insulin structures was conducted. Employed the technique of atom-wise distances between corresponding atom groups, allowing for the computation of distances between groups with equivalent atom counts. Contrasted the distances between alpha-carbons of porcine insulin at varying pH levels and esrapid at pH 3 conformations. Utilized distances.dist to compute distances between atom groups with identical atom counts. To calculate the distance between Cα atoms, targeted the atoms labeled 'CA' in each structure. The resulting distances.dist output provided the residue numbers of both selections, with the option of adding an offset to facilitate comparison between them and compatibility with other file formats, utilizing the default offset of 0.

Observations

esrapid insulin analog is engineered by substituting asparagine in the A21 chain with glycine. This alteration results in the insulin analog mimicking regular insulin features by maintaining a consistent pH point (pI) and enhancing hydrophobicity. *esrapid* formulation comprises 20 mM citric acid, and this monomeric product assumes a conformation conducive to direct receptor binding, suggesting this analog functions like an ultra-acting analog. *esrapid* is crystallized in monomeric form under acidic conditions with 20 mM citric acid at pH 3. The reduced side chain of glycine plays a role in fostering a more stable structure in acidic environments. This modification is also believed to protect the molecule against degradation during injection and exposure to bodily fluids.

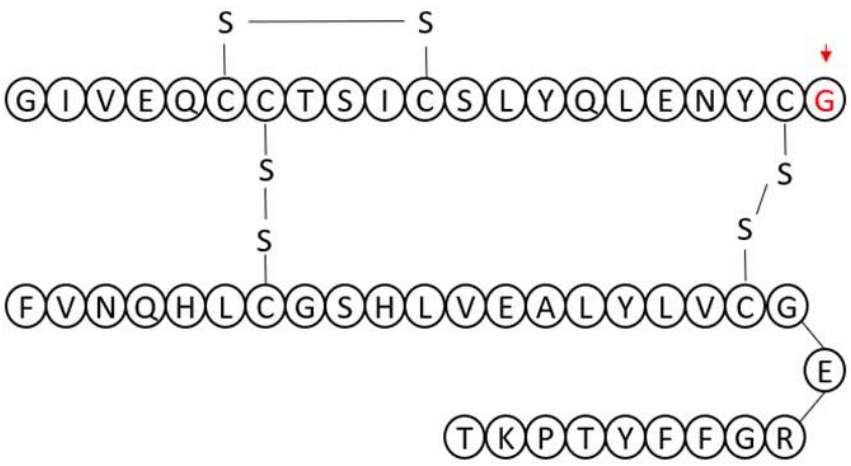


Figure 2.28 Primary sequence representation of *esrapid* insulin analog

esrapid represents the crystal structure of cubic insulin, identified with the space group I213. The the solvent content is measured at 41.54%. This structure has been successfully determined by molecular-replacement method and refined to a resolution of 1.36 Å, achieving an R factor of 21% (Table 2.5, Figure 2.28 and Figure 2.29).

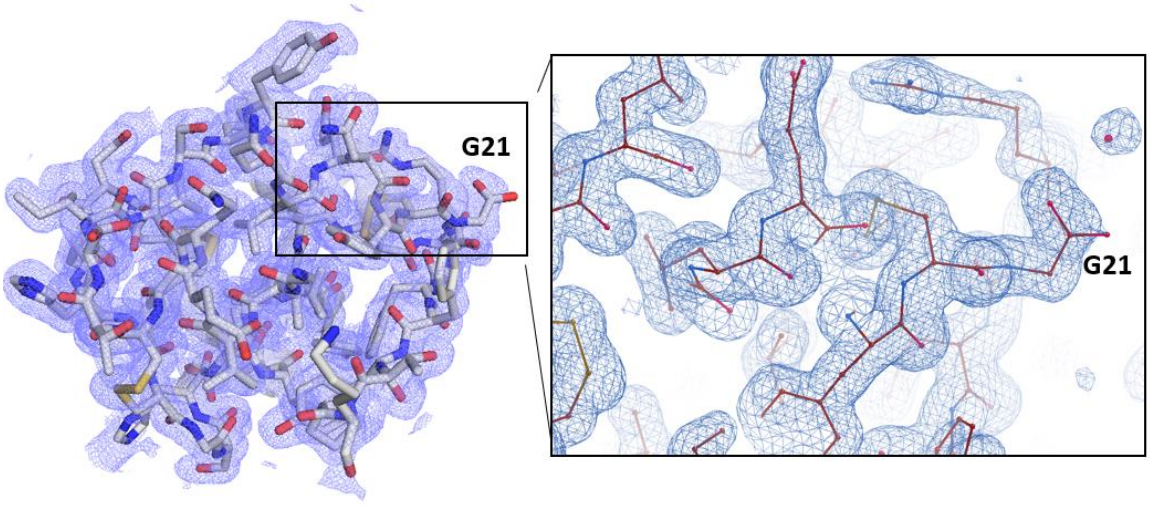


Figure 2.29 The esrapid insulin structure at 3D atomic resolution

The 2*Fo-Fc* simulated annealing-omit map at the 1 sigma level is shaded in slate. The mutated residue N21 is indicated in G21.

Investigations into cubic insulin crystals have predominantly centered on analyzing the impact of altering diverse parameters, including pH, ionic strength, ionic composition, and water activity. Various crystal structures of cubic insulin have undergone extensive comparative analyses under multiple conditions [Diao, 2003].

Remarkably, the structural integrity of the cubic insulin lattice survived with remarkable stability and minimal alteration across these diverse conditions. Even in instances where conformational alterations significantly displaced specific protein atoms, most of the insulin molecule and the crystal lattice remained undisturbed. In the investigation of the structural transition process, a prior study identified four structures within the pH range of 5.00 to 9.00. Utilizing crystallographic titration of cubic insulin crystals, the study explored for the presence of intermediates during the structural transition. This section in Chapter 2 attempts to expand on this investigation by introducing the *esrapid* insulin analog at pH 3. The primary objective is to determine potential conformational changes of *esrapid* analog through a detailed comparative analysis with the precisely determined four structures published in previous research.

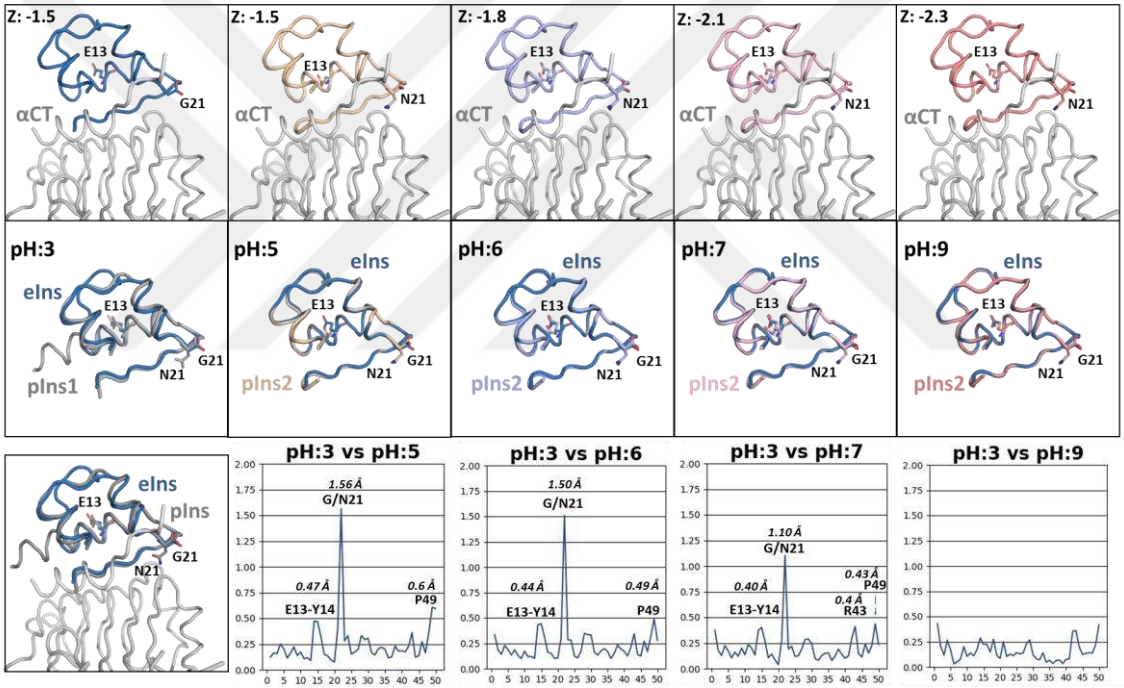


Figure 2.30 . Conducting docking and pairwise RMSD analyses on the structures of esrapid and insulin at varying pH conditions.

(Top) The docking analysis compares *esrapid* at pH 3 with porcine insulin structures at different pHs when interacting with the insulin receptor, displayed from left to right. Residue-based shifts are denoted as E13 and G21/N21, with the Z score of the docking result indicated accordingly. (Middle) Superposition is depicted between porcine R-state insulin (pIns1) and esrapid (eIns) at pH 3, porcine insulin (pIns2) at pH 5, and esrapid at pH 3, porcine insulin at pH 6 and esrapid at pH 3, porcine insulin at pH 7 and esrapid at pH 3, and porcine insulin at pH 9 and esrapid at pH 3. (Bottom) Pairwise RMSD analysis compares esrapid at pH 3 with porcine insulin structures at pH 5, pH 6, pH 7, and pH 9. Notably, conformational shifts in esrapid and porcine insulin exhibit similarities at both the lowest and highest pH values.

Docking analysis was conducted using the insulin-bound mini insulin receptor to assess the compatibility of insulins with the insulin receptor at different pH levels (Figure 2.30, top). The docking results exhibited a high concordance with the original insulin-bound insulin receptor. However, significant conformational changes were observed in residues B1-B8 of insulins compared to bound human insulin, including esrapid insulin analog. Notably, these insulins are in a monomeric T-stage, signifying their fully active insulin form. Furthermore, a resemblance was observed between porcine insulin and esrapid structures at different pH values (Figure 2.30, middle). The superposition of *esrapid* onto other structures proved satisfactory for all residues except for G/N21, FB1, E13, Y14, P49, and R43.

Remarkably, the residues G/N21, Y14, and P49 exhibited notable flexibility, a characteristic shared among various insulin structures. Significant shifts in multiple residues between insulin structures are evident in the pH range of 5 to 9, particularly noteworthy at pH 9, where numerous residues undergo significant changes. Between pH 6.00 and 7.00, a distinct rotation of approximately 120 degrees occurs in the side chain of E13 around the CA-CB bond, leading to a novel conformation where both configurations of E13 coexist within porcine insulin crystals. In contrast, at pH 9.00, E13 assumes a singular conformation which is similar to observed at pH 3. However, the pairwise RMSD analysis comparing *esrapid* at pH 3 and insulin from pH 5 to pH 7 reveals major shifts in G/N21 residues, with a prominent occurrence of double conformations at pH 6 and pH 7 (Figure 2.30, bottom).

Interestingly, a remarkable similarity is determined between the structural conformation of *esrapid* at pH 3 and porcine insulin at pH 9 in the pairwise RMSD analysis, indicating a close similarity in the structural conformation of insulin at both pH 3 and pH 9 environments.

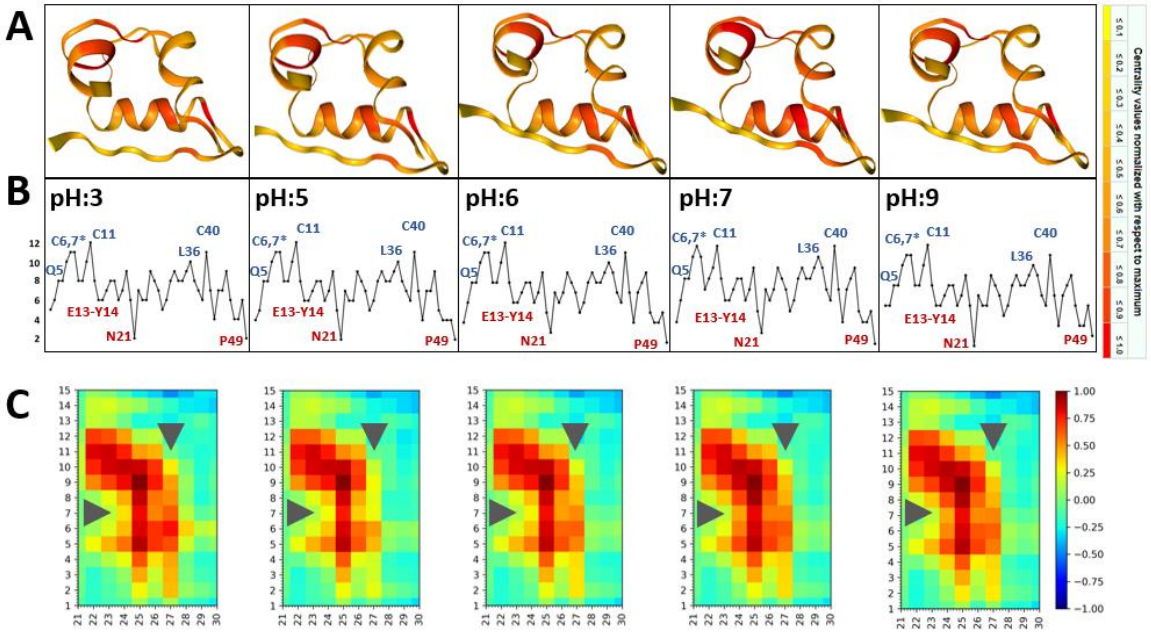


Figure 2.31 Node centrality and cross-correlation analyses on esrapid at pH 3 and four porcine insulin structures spanning the pH range of 5-9.

(A) Presentation of the node centrality analysis for five monomeric insulins. (B) Visualization of the node centrality plot for these five structures, with the highest and lowest residues highlighted accordingly. (C) Cross-correlation derived from 10 GNM modes for these five insulin structures, revealing noticeable correlation patterns involving residues S9 and Q5.

We also examined these structures about node centrality through network analysis of protein structures (NAPS) using default parameters on the Web server (Figure 2.31a). The centrality of a residue indicates its topological significance within the residue-residue contact network. Our subsequent topological analysis of *esrapid* and the four porcine insulin structures reveals that the residues with the lowest degree of centrality, that means the fewest contacts, in insulin molecules coincide with the regions identified by pairwise RMSD (residues E13, Y14, N21, P49) in monomeric insulin molecules (Figure 2.31b).

Moreover, we conducted cross-correlation analysis utilizing both *esrapid* and porcine insulin structures (Figure 2.31c). The cross-correlations among residue fluctuations were extracted from ten GNM modes, and we generated graphics with Matplotlib using the normalized outcomes. The HotRegion results of the structures were acquired to identify the regions on the interaction interfaces. We examined how the correlations between different residue fluctuations are modified, influencing the rewiring of residue communications among insulin structures identified at various pH values. The cross-correlation maps enable a comprehensive comparison among these

five monomeric insulin conformations across distinct pH states, highlighting subdomain regions introduced upon receptor binding. Notably, a robust correlation was observed between residues Q25 and S9, as well as between residues Q25 and Q5 across all pH conditions. Additionally, highly correlated residues, namely C11, Q5, C6-7, and I10, closely align with the findings of the node centrality analysis, suggesting a high degree centrality, precisely a substantial number of contacts, coinciding with the hinge regions in the monomeric insulin structures.

Interestingly, the interaction between residues C7 and N24 was observed to undergo pH-dependent changes. At pH 5 and 6, these residues exhibit a 25% correlation, while at pH 7 and 9, their correlation increases to 50%. However, at pH 3, these residues show a distinct cooperativity of at least 25% compared to the other structures. Moreover, it is hypothesized that L27 holds critical importance for monomer cooperativity in different pH environments, as there is a 50% correlation between I2-S9 and L27 at pH 3, whereas there is an above 25% correlation between them at pH 6-9 and at most a 25% correlation at pH 5. Residues at or near the receptor interacting region exhibit a stronger correlation with each other. This phenomenon leads to a more pronounced partitioning of monomeric insulin lobes at receptor interaction locations.

Concluding remarks

The esrapid insulin analog's structure was purified and crystallized in a 20 mM citric acid solution at pH 3, achieving a resolution of 1.36 Å. When at pH 3, this analog proves suitable for *in-vitro* analysis, revealing a proliferative effect. This signifies that the insulin correctly binds to the insulin receptor and functions as intended. Notably, the monomeric form of this insulin analog is well-suited for direct use without the need for hexamerization. This characteristic makes it highly compatible with pump utilization, minimizing potential obstructions and leakages in external infusion pump systems and tubing. Moreover, this analog can be readily transformed into its hexamer form with precise additives, allowing for formulation as an intermediate or regular insulin analog.

While ongoing analyses are being conducted, the structure of this insulin has been successfully determined (Table 2.5, Figure 2.28 and Figure 2.29). A comparative examination of esrapid and porcine insulin structures, along with assessments of observed and calculated structure factors, map correlation coefficients, and computational analyses, indicates a similarity in structural conformation between

esrapid at pH 3 and porcine insulin at pH 9. Notably, the structural dynamics of porcine insulin between pH 5 and 7 exhibit a gradual transformation with varying pH levels. Particularly, when superimposing pH 3 with pH 5, pH 6, and pH 7, critical residues such as E13, Y14, N21, and P49 show significant conformational changes. Nevertheless, the alignment of pH 3 and pH 9 insulin structures reveals no significant differences in conformation, suggesting that monomeric insulin at pH 3 behaves similarly to monomeric insulin at pH 9. Further analysis highlights the significance of residues Q5, C6-7, S9, N24, and L27 for monomeric insulin's stability and pH compatibility (Figure 2.30 and Figure 2.31). These residues are suggested to function as subdomains during the monomer form of insulin. In this comparative pilot study, the combined experimental and computational analyses of *esrapid* contribute significantly to our comprehensive understanding of monomeric insulin dynamics.

2.3. Chapter 2 overview

In this chapter, we investigate the domestic production and determination of the 3D structure of ultra-acting insulin analogs, namely *esrapid* and *esrapid*^{2}. The *esrapid* analog incorporates the substitution of A21N to A21G, ensuring its monomeric stability in a 20 mM citric acid environment. Our analysis of *esrapid* reveals its direct and accurate binding to its receptor, exhibiting proper growth function on hiPSCs. Moreover, we identified pH-dependent critical subdomains (Q5, C6-7, S9, N24, and L27) through protein-protein network analysis, docking, pairwise, and GNM analysis. These findings suggest the adaptability of this molecule to extreme pH conditions. Consequently, our structural and computational analyses indicate a striking similarity in the behavior of this insulin at pH 3 and the insulin structure at pH 9.

The *esrapid*^{2} analog is characterized by substituting B22R to B22K and B28P to B28D, which is aimed at providing a faster insulin analog. The lysine side chain, bearing an amino group and a positive charge similar to arginine, is chosen for its more flexible structure and perceived less impact on chemical stability compared to arginine. With these considerations, *esrapid*^{2} is strategically designed as an ultra-acting insulin analog. This analog exists in hexamer form, and its 3D structure bound to zinc has been determined at a resolution of 2.5 Å. Neural activity, experimental and computational analyses suggest that the production of this analog has been optimized, and it is demonstrated to be less stable than insulin aspart, supporting its classification as an ultra-acting insulin analog. A study utilizing a mini reactor and the Design Expert statistical tool indicates that the optimal conditions for maximum production of esrapid2 analog involve the presence of 8.78 mM glucose, 1% glycerol, 30 g/L LB, and 15 mM $MgSO_4$. Under these conditions, insulin has been produced at a concentration of 13 mg/ml. Following upstream characterization, the focus shifted to examining this insulin's monomeric and hexameric dynamics using HDX-MS analysis and conventional X-ray crystallography, respectively. The X-ray crystallography results revealed that the hexamer form of *esrapid*^{2} is less stable than insulin aspart, particularly in the absence of electrostatic interactions at BAsn3 in the hexamer. HDX-MS analysis highlighted the considerable flexibility of the *esrapid*^{2} monomer compared to the insulin aspart monomer. This experimental finding is corroborated by DSC analysis, indicating that the monomer insulin *esrapid*^{2} exhibits a lower melting point compared to the insulin aspart monomer. Moreover, the experimental HDX reaction output aligns with normal

mode analysis results, demonstrating a concordance between HDX reaction and NMA outcomes.

The findings presented in Chapter 2 suggest that active insulin *esrapid* and *esrapid*^{2} analogs may exhibit an ultra-acting effect on both T1D and T2D. However, comprehensive *in vitro*, *in vivo*, and clinical studies are imperative for a more thorough discussion. Further analyses are underway to complete our understanding. In this chapter, our emphasis has been on the dynamics of rapid/ultra-rapid-acting insulin. The subsequent chapter will delve into the extensive exploration of commercially available insulins and their oligomeric dynamics.

Chapter 3

Exploring Commercially Available Protracted Insulin Analogs

The administration of exogenous insulin to individuals with diabetes mellitus has significantly extended the lifespan of millions over nearly 80-year [Galloway et al., 1994]. However, achieving effective glycemic control remains challenging, leading to a substantial burden of morbidity and premature mortality associated with diabetes. This situation profoundly impacts the quality of life for those affected and their families. The primary limitation in diabetes treatment lies in our inability to restore or replicate the intricate physiological mechanisms governing nutrient regulation, particularly the dynamic insulin responses crucial for blood glucose homeostasis in a healthy state.

The most practical method for administering therapeutic insulin involves subcutaneous injection or infusion [Galloway et al., 1994]. However, insulin absorption into the circulation from a subcutaneous depot is an unphysiological approach to insulin replacement. In a healthy state, insulin secretion into the portal vein undergoes continuous modulation in response to various stimuli, including nutrients, to regulate hepatic glucose output and other metabolic processes such as peripheral glucose uptake and lipogenesis [Home and Mehta, 2021]. Insulin is synthesized in pancreatic beta cells, where it self-assembles at high concentrations of zinc ions into hexamers for efficient storage within vesicles. Dilution upon exocytosis leads to the immediate dissociation of insulin hexamers into biologically active monomers. The dynamic pattern of insulin secretion is characterized by a low and constant 'basal' output, particularly evident overnight [Ayan and DeMirci]. The clinical dilemma lies in replicating the physiological insulin profile through the subcutaneous absorption of exogenous insulin. The challenge originates from pharmaceutical formulations of human insulin, which need to be in a concentration where insulin is hexameric to be therapeutically effective, ensuring acceptable physical stability and tolerable injection volume [Home and Mehta, 2021; Gough and Narendran, 2017]. Upon injection, the dilution of the insulin depot

occurs at a relatively slow pace, resulting in a delayed dissociation into monomers. Self-associated insulin diffuses through the tissue and penetrates the capillary wall better slowly than monomeric insulin. Additionally, the entry rate into the systemic circulation depends on local environmental factors. Consequently, the plasma kinetic profile of injected human insulin does not resemble normal physiology's basal or prandial profiles [Home and Mehta, 2021; Gough and Narendran, 2017]. Moreover, standardized injections yield pharmacokinetic profiles that vary both between individuals. Optimally, the absorption characteristics of subcutaneously injected insulin should be varied to generate rapidly absorbed formulations to imitate endogenous prandial insulin and formulations with slow, steady absorption rates to simulate basal insulin secretion. Another objective is to achieve predictability in the absorption rate profile across injections [Home and Mehta, 2021; Gough and Narendran, 2017].

Actions to attain these objectives have experienced varying degrees of success. Enhanced insights into the insulin molecule's domains involved in receptor interaction and self-association, coupled with advances in recombinant DNA technology, allowed for modifying the molecular structure of insulin to create analogs with 'weak' self-association in the 1980s and 1990s [Home and Mehta, 2021]. Thus, prandial agents like insulin aspart and lispro were created, exhibiting increased dissociation rate constants compared to human insulin [Plank et al., 2002]. These rapidly dissociating analogs are absorbed more swiftly from the subcutaneous depot, enabling them to better mimic the prandial insulin secretion profiles observed in normal physiology [Plank et al., 2002]. While developing these analogs, it became evident that even minor alterations in the amino-acid sequence of insulin could modify the molecular ternary structure, influencing receptor interactions and metabolic and mitogenic activities. A prototype analog demonstrated a significantly elevated "IGF-I receptor": "insulin receptor" affinity ratio compared to human insulin and prolonged residence time at the insulin receptor [Home and Mehta, 2021; Plank et al., 2002]. This analog exhibited carcinogenic effects in female rats at high doses. This underscored the importance of thoroughly testing the complete pharmacodynamic profile of insulin analogs.

Attempts to develop insulin analogs with comprehensive and predictable impacts have been relatively incremental. Initial tries to modify insulin absorption properties primarily focused on delaying and prolonging absorption to reduce injection frequency. Formulations incorporating protamine or an excess of zinc were introduced to create suspensions that would slowly dissolve once upon injection. However, these

preformed precipitates necessitated thorough resuspension prior to injection for achieving the correct insulin concentration [Home and Mehta, 2021; Bradley et al., 2021]. Observed that incomplete resuspension by patients contributed significantly to variability. Moreover, the mean absorption profiles of traditional basal insulins are suboptimal. NPH insulin peaks in the pharmacokinetic profile approximately 4–6 hours post-injection, followed by a gradual decline, deviating from the desired low and constant absorption. This combination of a peak effect and the unpredictable timing and extent of the peak poses a risk of hypoglycemia [Bartal et al., 2020; Bradley et al., 2021]. Given that basal insulins are often administered in the evening, there is a real threat of nocturnal hypoglycemia. This limitation affects the tolerable dose and has prompted the quest for basal insulin analogs with enhanced pharmacokinetic and pharmacodynamic characteristics [Bartal et al., 2020; Home and Mehta, 2021]

To improve the properties of a basal insulin analog, the initial approach involved altering the amino-acid sequence to adjust the isoelectric point of the insulin analog toward neutrality, thereby reducing solubility at physiological pH levels. Formulating the resulting analog in an acidic medium can be injected as a solute, eliminating the need for resuspension. Upon encountering the neutral subcutaneous environment, the insulin precipitates and slowly dissolves the solid phase, resulting in prolonged absorption. However, the unpredictable nature of precipitate formation and redissolution might introduce additional variability [Rosenstock et al., 2023; Alhmoud et al., 2024]. Early prototypes based on this principle were not pursued clinically due to low bioavailability and action variability. Nevertheless, this principle was later successfully employed in the development of insulin glargine, where the two arginine residues addition to the C-terminus of the B-chain shifted the isoelectric point from pH 5.4 to 6.7 [Goykhman et al., 2009; Rosenstock et al., 2023].

Another approach for developing enhanced basal insulin involved engineering a stable hexamer by replacing the zinc ions located in the core of the hexameric structure with cobalt ions. This complex exhibited gradual absorption as a hexamer, subsequently undergoing enzymatic cleavage into monomers in the bloodstream. However, it did not supply any pharmacological benefits over NPH insulin [Kurtzhals and Ribel, 1995].

The latest approach applies acylating fatty acids to the insulin molecule to stabilize self-association and allow reversible insulin-albumin binding. This principle provides for creating a neutral liquid preparation that remains non-precipitating throughout the administration-absorption process, eliminating significant variability

sources. Insulin detemir is introduced by removing the amino acid threonine at the B30 locus, with a 14-carbon fatty acid being acylated to lysine at the B29 locus [Athanasiadou et al., 2022].

3.1. A pilot study of cryogenic insulin detemir
The unique mode of protracted insulin Levemir at 1.7 A resolution

This chapter largely excerpted from Ayan et.al. (2023) and completely re-written for copyright protection. All the work has been done by myself.

The protraction mechanism of insulin detemir has been monitored through physicochemical and pharmacological investigations [Havelund et al., 2004]. Differences in self-association profiles and albumin affinity were observed in size-exclusion chromatography experiments comparing insulin detemir with various acylated and non-acylated analogs. Upon injection, insulin detemir exists in a hexameric state; however, data suggest a subsequent dihexamer-hexamer equilibrium, with hexamers formed through interactions between myristic acid groups at the hexameric complex poles. This shift in the self-association state in the depot is believed to occur as pharmaceutical preservatives such as phenol and cresol equilibrate with physiological electrolytes. Size-exclusion chromatography further indicates that both self-associated dihexameric, hexameric, and dimeric and monomeric forms of insulin detemir can bind to albumin, presenting two potential mechanisms for the retention of insulin detemir in the injection depot [Whittingham et al., 1997].

To assess the relative contributions of self-association and albumin binding to depot protraction, studies on the disappearance rates of radiolabelled insulin analogs in porcine models were conducted. The T50% disappearance time for insulin detemir was found to be 10.2 h, surpassing the T50% values of an acylated monomer (2.9 h), an acylated weakly associating hexamer (6.9 h), and an acylated stable hexamer (8.8 h). This suggests that self-association significantly influences the slow absorption of insulin detemir [Havelund et al., 2004; Whittingham et al., 1997]. Conversely, the T50% of the albumin-binding acylated stable hexamer was considerably longer than that of the stable nonbinding hexamer, underscoring the role of albumin binding in delaying absorption. Hence, the probable mechanism for protracted absorption involves the self-association of insulin detemir, leading to its retention in the depot long enough to establish albumin binding, further delaying absorption [Havelund et al., 2004]. While insulin detemir also constitutes reversible bonds with albumin in the bloodstream, one could theorize that a further elongation of the pharmacodynamic effect might occur through maintained retention in the plasma compartment. When assessing plasma residence times compared

to other acylated analogs with varied albumin-binding affinities, it becomes evident that insulin detemir's retention in circulation has a degree of impact. However, this impact on the overall extension of action is markedly less than what is ascribed to depot retention [Kurtzhals et al., 1995].

Assorted studies have so far been conducted into detemir, aiming to comprehend its structure, functionality, and interaction with serum albumin [Hamdan et al., 2022; Havelund et al., 2022]. This pilot study of Chapter 3 reveals the high-resolution crystal structure of detemir, determined at 1.7 Å resolution under cryogenic conditions. The diffraction data were collected at the Turkish Light Source. Furthermore, we completed the computational analysis to explore variations in monomer-dimer equilibrium dynamics between our cryogenic structure and a previously published detemir structure (PDB ID: 1XDA). The crystallographic data presented herein indicate a distinct unit cell orientation in the asymmetric unit compared to the previously published detemir unit cell orientation. GNM analysis involved in the high-resolution structure discloses a hitherto unseen pattern of monomeric-dimeric-correlated motion in this long-acting insulin analog. Understanding the correlated motion between monomers and dimers in the detemir structure remains crucial for revealing this vital hormone's structural dynamics and pharmacokinetics. This insight lays the foundation for enhancing, modulating, and optimizing the protracted dissociation rate of detemir's hexameric form into its active monomeric form.

3.1.1. Preparation of Samples, Crystallization and crystal harvesting

The crystallization of Levemir® was conducted using the sitting-drop micro batch vapor diffusion screening technique with 72-well Terasaki crystallization plates, following the procedure outlined in Chapter 2. In a nutshell, a solution of detemir was combined with a 1:1 ratio (v/v) of approximately 3500 commercially available sparse matrix crystallization screening conditions at room temperature. Each well, containing 0.83 µL of the protein solution and crystallization condition, was sealed with 16.6 µL of paraffin oil (Cat#ZS.100510.5000, ZAG Kimya, Türkiye) and stored at room temperature until crystals formed. The Terasaki plates were analyzed under a compound light microscope to monitor crystal formation and growth. Large crystals were observed within 48 hours, and the most favorable crystals were obtained in a buffer comprising 0.2 M sodium acetate trihydrate and 0.1 M TRIS hydrochloride at pH 9.2. The recovery

of crystals from the Terasaki crystallization plates was conducted using MiTeGen and micro loop sample pins attached to a magnetic wand, executed under a compound light microscope. Subsequently, the gathered crystals were promptly flash-frozen by immersion in liquid nitrogen and introduced into a pre-cryo-cooled sample puck (Cat#M-CP-111-021, MiTeGen, Ithaca, NY, USA). Immersed in liquid nitrogen, the filled puck was then transferred to the Turkish DeLight dewar to collect diffraction data.

3.1.2. Data Collection and processing

The diffraction data were obtained using Rigaku's XtaLAB Synergy Flow XRD, which was equipped with CrysAlisPro software [Agilent, 2014], following the methodology outlined by Atalay et al. in 2022. Prior to data collection, the system temperature was reduced to 100 K using Oxford Cryosystems's Cryostream 800 Plus. Subsequently, the dewar was filled with liquid nitrogen, and a loaded sample puck was introduced into the cryo sample storage dewar at the Turkish DeLight facility. The MiTeGen sample pin, mounted within the sample storage dewar by the robotic auto sample changer (UR3), was then positioned by the six-axis robotic arm to the Intelligent Goniometer Head (IGH). The crystal alignment was achieved using the automatic centering feature of the CrysAlisPro software to the X-ray interaction position. For the data collection phase, the PhotonJet-R X-ray generator with a Cu X-ray source operated at full power (40 kV, 30 mA), and the beam intensity was set to 10% using the piezo slit system to minimize cross-fire and overlap of Bragg's reflections. The diffraction data, encompassing a total of 1800 frames, were collected at a resolution of 1.7 Å, a 45 mm detector distance, 0.2-degree scan width, and a 1 s exposure time. The automatic data reduction feature of CrysAlisPro processed the data. Upon completion of data reduction, the output files, comprising both unmerged and unscaled files in *.rrpprof* format, were converted to the structure factor file *(*.mtz format) through the *.hkl file using the *CCP4* crystallography suite, integrated into the CrysAlisPro software.

3.1.3. Structure determination and refinement

The crystal structure of detemir was determined at a high resolution of 1.7 Å under cryogenic conditions in space group R3:H. Molecular replacement was conducted using the automated program *PHASER* within the *PHENIX* software package [McCoy

et al., 2007], utilizing the previous detemir X-ray crystal structure at cryogenic temperature as a reference model (PDB ID: 1XDA). Initial rigid-body refinement was performed using the coordinates of 1XDA in the *PHENIX* software package. Following simulated-annealing refinement, further refinement involved adjusting individual coordinates and translation/libration/screw (TLS) parameters. *COOT* [Emsley and Cowtan, 2004] was employed to verify potential alterations in side chain positions and the presence of water molecules, with retention of positions exhibiting strong difference densities. Any missing water molecules were manually added, and those outside significant electron density were removed. The Ramachandran statistics for the detemir structure (most favored/additionally allowed/disallowed) were 100%/0.0%/0.0%, respectively. Detailed statistics of the structure refinement are presented in Table 1. Visualization of all X-ray crystal structure figures was generated using *PyMOL* [Delano, 2002].

3.1.4. GNM Analysis

Utilizing ProDy [Yang et al., 2006; Bahar et al., 2005], normal mode analysis was conducted on the 1XDA and 8HGZ structures. A cross-correlation map was established, encompassing all Cα atoms within residues 0–200 of the proteins, excluding myristic acids, phenols, zinc, and chloride atoms from the analysis. A cut-off distance of 8 Å was applied to assume pairwise interactions in both structures, and the default spring constant was set to 1.0 for both models. The analysis involved parsing 1883 and 1673 atoms, calculating 198 non-zero modes for 8HGZ [Ayan et al., 2023] and 1XDA [Whittingham et al., 1997] structures, respectively. Experimental B-factors were compared with theoretical fluctuations and orientational cross-correlations between residue fluctuations were determined across all GNM modes. Differences in cross-correlations between the 8HGZ and 1XDA structures were calculated for specific sections. This involved comparing intra-chain cross-correlations for each monomer in 8HGZ with intra-chain cross-correlations for each chain in the 1XDA reference model. Additionally, interchain cross-correlations between all chains were calculated similarly, and the results were presented as heat maps.

3.1.5. Normalization of B-Factor

The normalization of B-factors entails a transformation of experimental B-factors to establish their distribution about expected value and variance. This procedure facilitates the comparison of B-factors across different datasets by mitigating potential biases that could affect the data. The pdb library transfers the normalized data into a temporary record set, presenting B'-factors per residue. The initial algorithm for B-factor normalization, introduced by Karplus and Schulz, establishes a correlation between the experimental B-factor of a specific residue and the arithmetic mean of all B-factors in a given structure.

3.1.6. Observations

Insulin comprises two peptide chains, A and B, connected by two disulfide bonds, constituting a single 'monomer.' When two such monomers associate noncovalently, they form a dimer. In its inactive state, insulin is released as a hexamer outside the cells. Activation occurs when insulin dissociates into its monomeric form, allowing it to bind effectively to its receptor. Detemir insulin exhibits structural distinctiveness compared to other insulin analogs, characterized by two hexamers in its unit cell attributed to myristic acids (Figure 3.1).

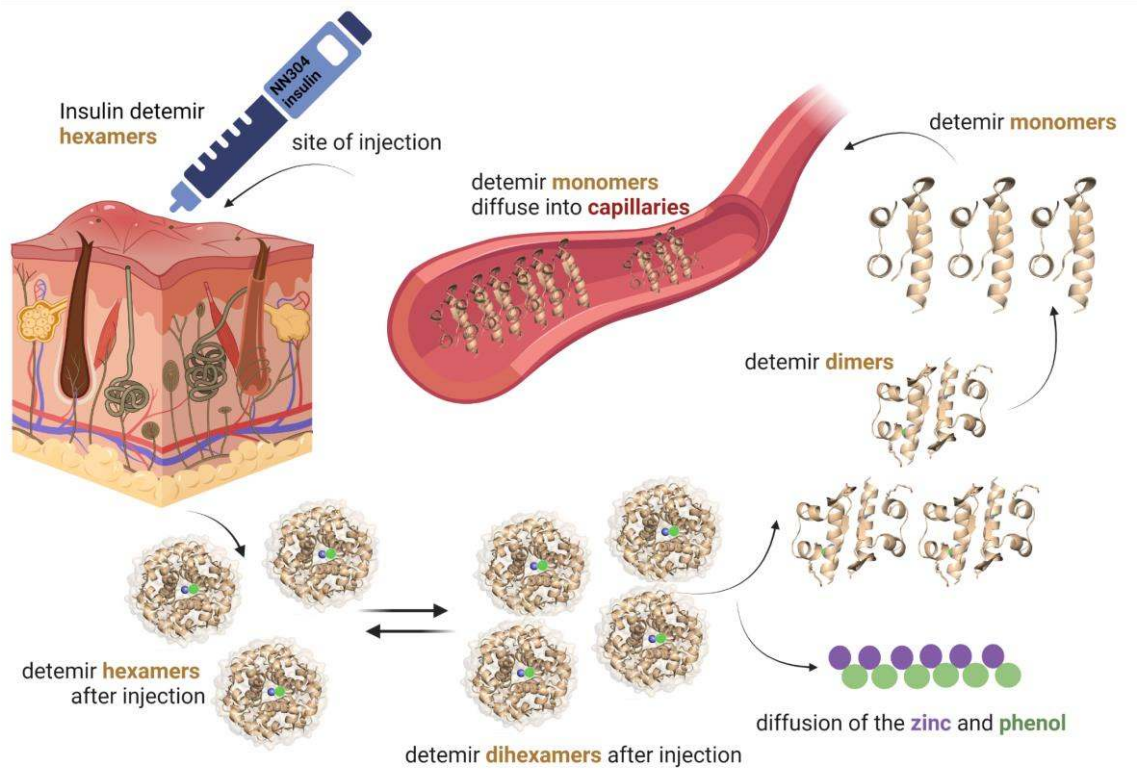
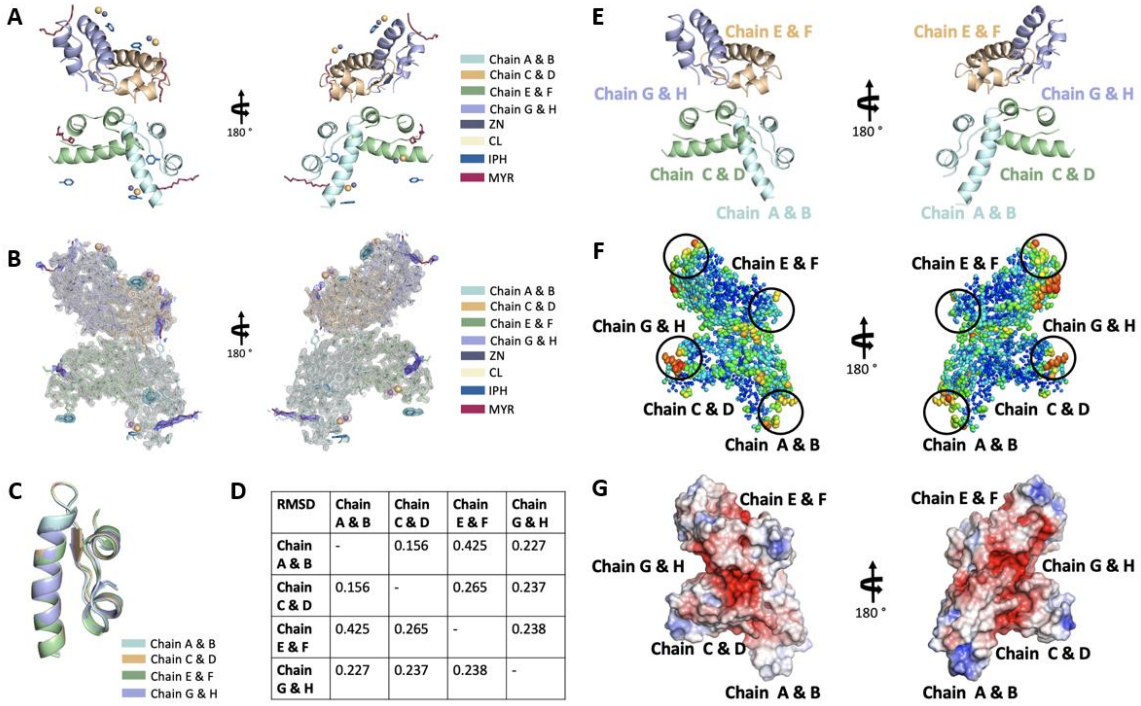


Figure 3.1 Overview the mechanism of action for insulin detemir.

Following subcutaneous injection, insulin detemir's hexamers or dihexamers disperse and gradually dissociate in the bloodstream. The dihexamers, hexamers, dimers, and monomers of insulin detemir are represented in a wheat color, while zinc and phenol are depicted as circles and colored in purple and green, respectively.

Consequently, detemir's asymmetric unit is also distinct, and its monomer conformation encompasses pairs of chains A and B (or C and D, or E and F, or G and H). In detemir, each dimer refers to the AB-CD and EF-GH pairs. In this investigation, the cryogenic structure of detemir was resolved with a resolution of 1.7 Å. The asymmetric unit cell consists of four nearly identical monomer molecules, each exhibiting RMSD values ranging from 0.156 to 0.425 (Figure 3.2, Table 3.1). Additionally, the unit cell contains four myristic acid residues, four phenol molecules, four zinc cations, four chloride anions, and 157 water molecules.



RMSD	Chain A & B	Chain C & D	Chain E & F	Chain G & H
Chain A & B	-	0.156	0.425	0.227
Chain C & D	0.156	-	0.265	0.237
Chain E & F	0.425	0.265	-	0.238
Chain G & H	0.227	0.237	0.238	-

Figure 3.2 Representation of the 8HGZ structure at cryogenic temperature.

(A) Presentation of the detemir structure in a cartoon view, with myristic acid (MYR), zinc (ZN), chloride (CL), and phenol (IPH) depicted in stick representation. (B) The refined *2Fo-Fc* electron density map encompassing detemir is portrayed in gray, where MYR is highlighted in blue, and Zn and Cl are shown in dark salmon, while IPH is represented in smudge. (C) Alignment of each detemir monomer in *PyMOL*, and (D) corresponding RMSD values are presented in the accompanying table. (E) Each detemir monomer is distinguished by a unique color. (F) Detemir is visualized with an ellipsoid presentation, accentuating stable and flexible regions displayed in warmer colors. Circles denote the binding site of myristic acid (MYR). (G) Electrostatics are generated to illustrate surface charges on detemir using the APBS electrostatic plug-in in *PyMOL*. (Adapted from Ayan et. al., 2023)

Table 3.1 Data collection and refinement statistics of cryogenic detemir.

Data collection	**detemir**
Instrument	Turkish DeLight
Resolution range	21.64–1.7 (1.761–1.7)
Space group	R3:H
Unit cell	78.885 78.885 79.272 90 90 120
Redundancy	5.4 (4.5)
Completeness (%)	96.94 (90.88)
Mean I/sigma(I)	23.33 (3.28)
R-merge	0.05 (0.39)
Total reflections	110,176 (9012)
Unique reflections	20,225 (1834)
CC1/2	0.99 (0.87)
CC*	1 (0.96)
R-work	0.21 (0.26)
R-free	0.24 (0.28)
Wilson B-factor	16.70
Protein residues	200
RMS (bonds)	0.003
RMS (angles)	0.37
Ramachandran favored (%)	100.00
Ramachandran allowed (%)	0.00
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.00
macromolecules	1601
ligands	103
solvent	188

Each disulfide-bonded monomer chain is identified as AB, CD, EF, and GH (Figure 3.2a-e). In the ellipsoid representation, each periphery of the monomer in the dimer form of 8HGZ exhibits more remarkable plasticity and flexibility compared to the body/center of the overall structure, as determined through TLS analysis on PyMOL (Figure 3.2f). Additionally, the 8HGZ structure reveals a more negative charge

distribution in the central region than its periphery (Figure 3.2g). Each monomer features a 14-carbon fatty acid chain (myristic acid) covalently linked to the B29 Lys residue. The myristic acids play a role in crystal contact among neighboring dimers, with aliphatic side chains extending from the crystallographic axis towards the periphery (Figure 3.3).

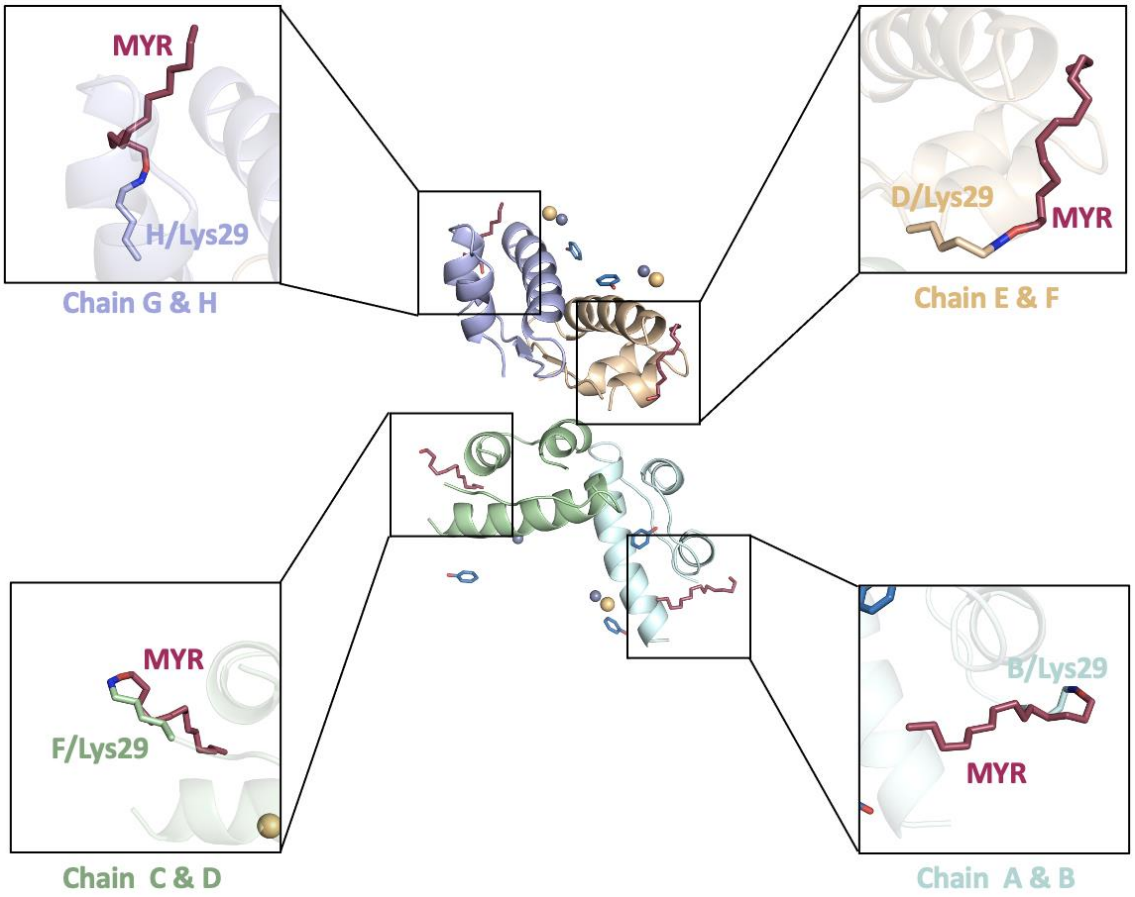


Figure 3.3 Depiction of the covalent bond between myristic acid (MYR) and detemir.
The stick representation illustrates the binding of each MYR to detemir. Residues are highlighted according to atom representation, indicating the covalent bond with MYR (Adapted from Ayan et. al.,2023).

The 8HGZ and 1XDA crystal structures adopt a "dimer of dimer" asymmetric form. At cryogenic temperature, our dimer of dimer crystal structures could not be aligned with the dimer of dimer structure (1XDA) due to distinct crystal packing arrangements in the detemir asymmetric unit cells (Figure 3.4). When aligned and compared to their hexamer forms, minor conformational changes are observed due to differences in crystal packing. Additionally, cross-correlation analysis using GNM supports the significant differences in biologically relevant structures between the two forms (Figure 3.5).

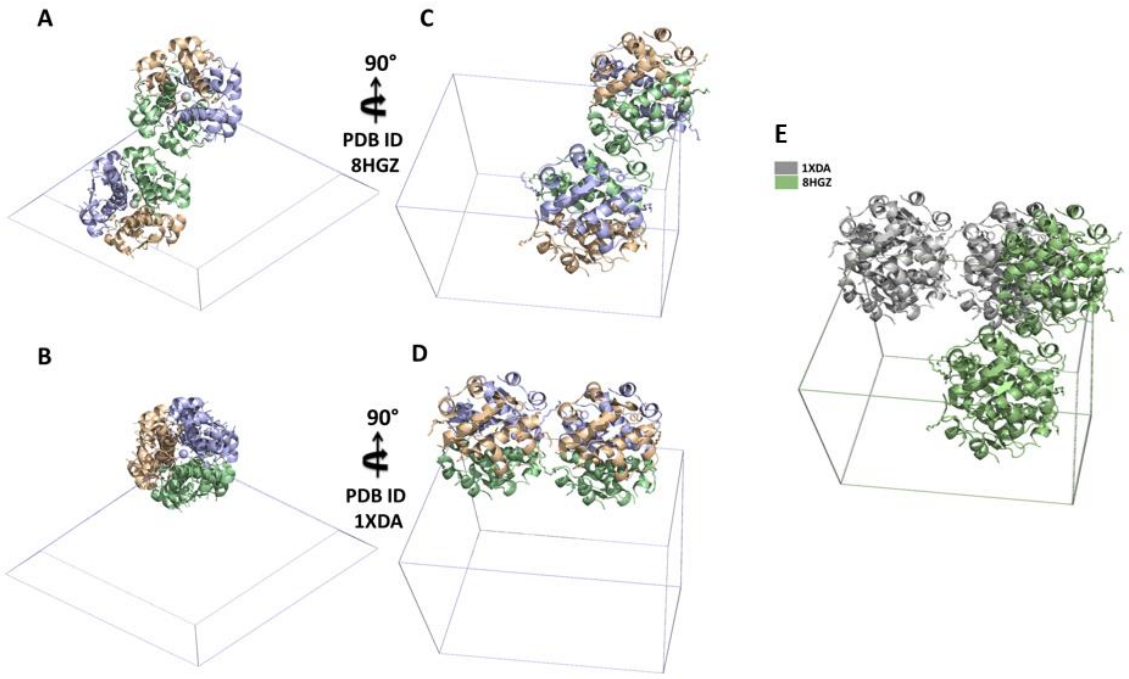


Figure 3.4 Comparison of variations in the unit cell origin between the cryogenic structures of 8HGZ and 1XDA.

(A,C) Presentation of the 8HGZ structure in unit cell mode, with an additional 90-degree rotation. (B,D) Display of the 1XDA reference model in unit cell mode, also rotated 90 degrees. (E) The two detemir structures are superimposed in unit cell mode, with the 8HGZ structure highlighted in green and the 1XDA structure in gray.

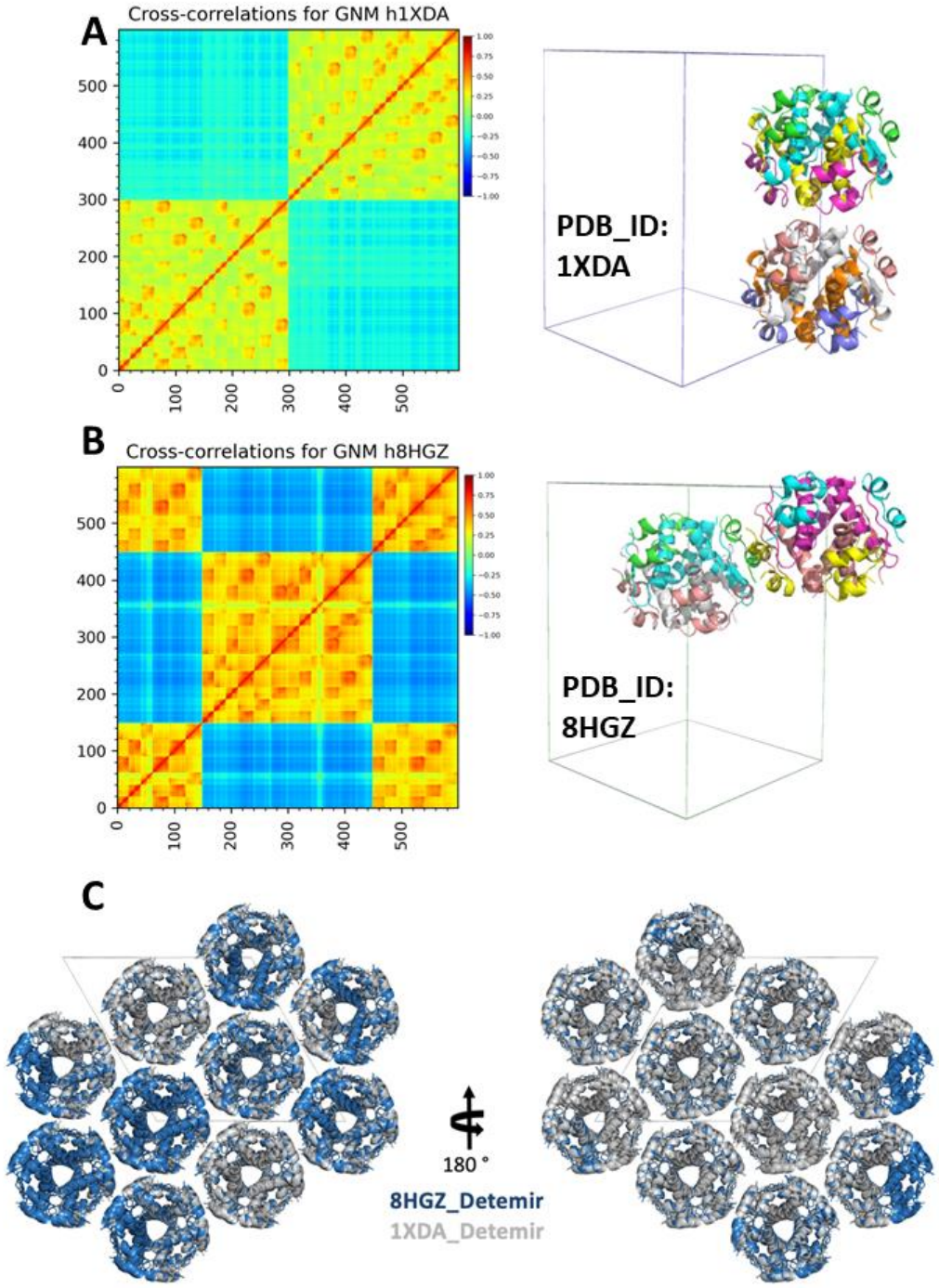


Figure 3.5 Depiction of the dihexamer conformation in the 8HGZ and 1XDA structures.

(A) GNM-based cross-correlation analysis of the dihexamer conformation in 1XDA and its cartoon representation in unit cell mode. (B) GNM-based cross-correlation analysis of the dihexamer conformation in 8HGZ and its cartoon representation in unit cell mode. A cut-off distance of 8 Å was applied in both structures to assume pairwise interactions. (C) Visualization of crystal packing for 1XDA (in gray) and 8HGZ (in sky blue) structures. (Adapted from Ayan et. al., 2023)

The GNM analysis serves as a simplified coarse-grained method designed to unveil a protein's dynamics. Experimental comprehension of a protein's equilibrium fluctuations is facilitated by assessing the Debye-Waller/B-factor of each atom through X-ray crystallography. Although interpreting crystallographic B-factors can be intricate, they encapsulate the cumulative impact of diverse sources of disorder. Additionally, these B-factors offer multiple analytical perspectives for decomposing molecular disorder into a hierarchical cluster of contributions, providing an intuitive framework for quantitative structural dynamics analysis. GNM is capable of predicting expected residue fluctuations that align favorably with experimentally measured fluctuations. Consequently, the relationship between expected residue fluctuations derived from GNM and thermal fluctuations/B-factors can be meticulously examined. The fluctuation profile and cross-correlations between pairs of nodes can also be deduced in alignment with experimental or adjustable parameters.

To explore potential differences in dynamics arising from the distinct unit cell arrangements in the asymmetric unit, a detailed analysis of the dimer of dimer form of the 1XDA and 8HGZ structures was conducted using GNM (Figure 3.6 and Figure 3.7). The examination involved scrutinizing orientational cross-correlations between residue motions and analyzing the flexibility of residues to describe mean squared fluctuations. A cross-correlation map was employed to investigate communication within the residue networks, where blue-colored regions denoted uncorrelated motions and red-colored regions indicated collective residue motions.

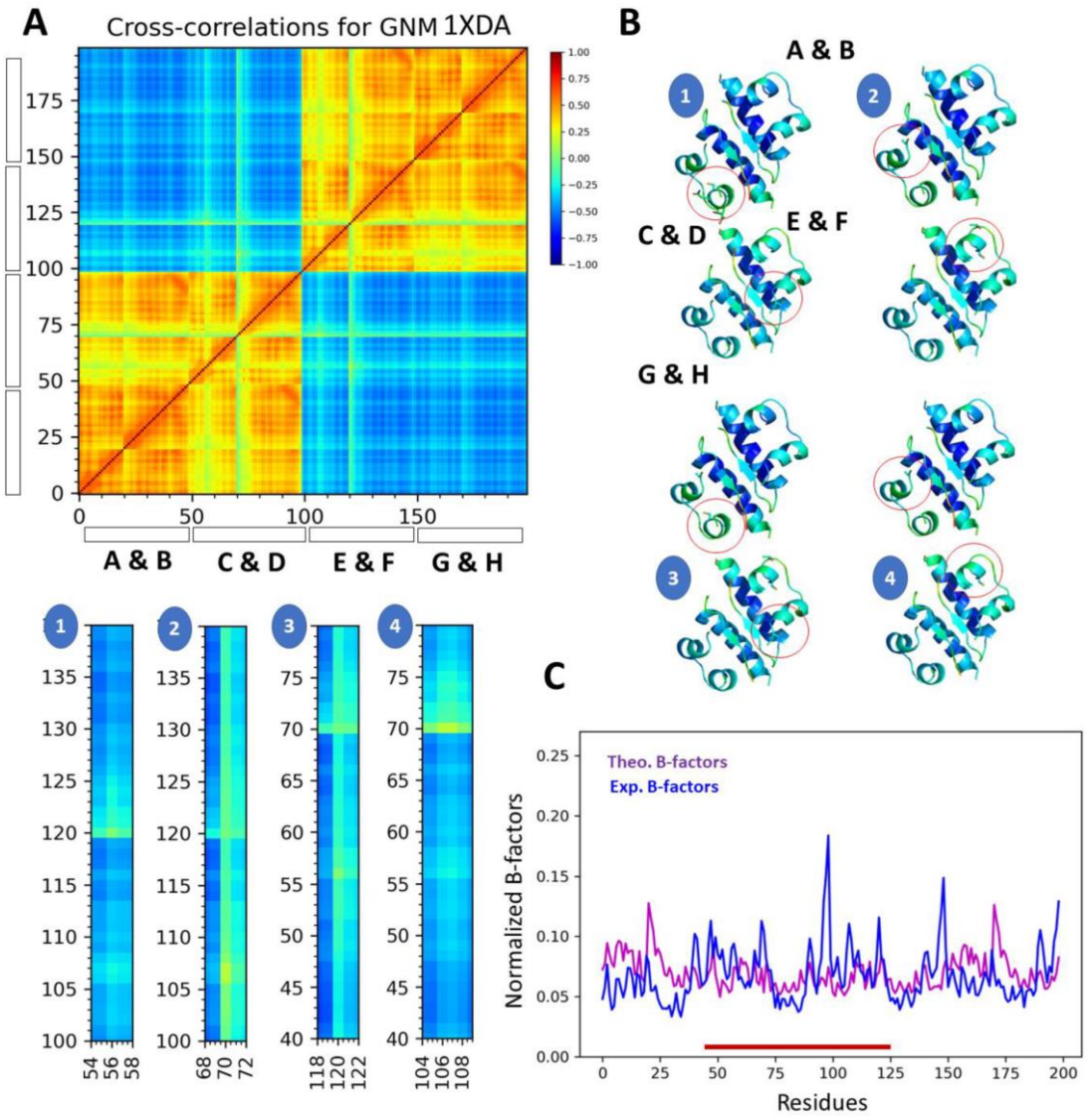


Figure 3.6 GNM Analysis results for the isolated 1XDA structure.

(A) Heat-map illustrating cross-correlation from overall GNM modes for the 1XDA insulin structure, highlighting ~25% correlated critical residues (54–58, 68–72, 104–108, and 118–122, along with 56, 70, 105–107, and 120). Residue pairs moving in the same direction are depicted in red regions (C_{ij}orient > 0), those moving in opposite directions in blue regions (C_{ij}orient < 0), and uncorrelated pairs in green (C_{ij}orient = 0.0; color-code bar on the right scale). (B) The cartoon representation of the 1XDA structure is color-coded by the B-factor. Critical residues, corresponding to the heat map, are numerically labeled and highlighted with red circles. (C) Normalized B-factors of 1XDA are represented in blue, and theoretical B-factors calculated by GNM are shown in magenta. Critical residues corresponding to the heat map are emphasized in the red line.

In the cross-correlation heat maps, the 1XDA structure revealed highly correlated motions for the AB-CD (1st dimer) and EF-GH (2nd dimer) pairs (Figure 3.6a). In

contrast, the 8HGZ detemir structure exhibited less apparent dimeric positive correlation due to the extroverted monomers of the 1st dimer packing in the asymmetric unit cell (Figure 3.7a).

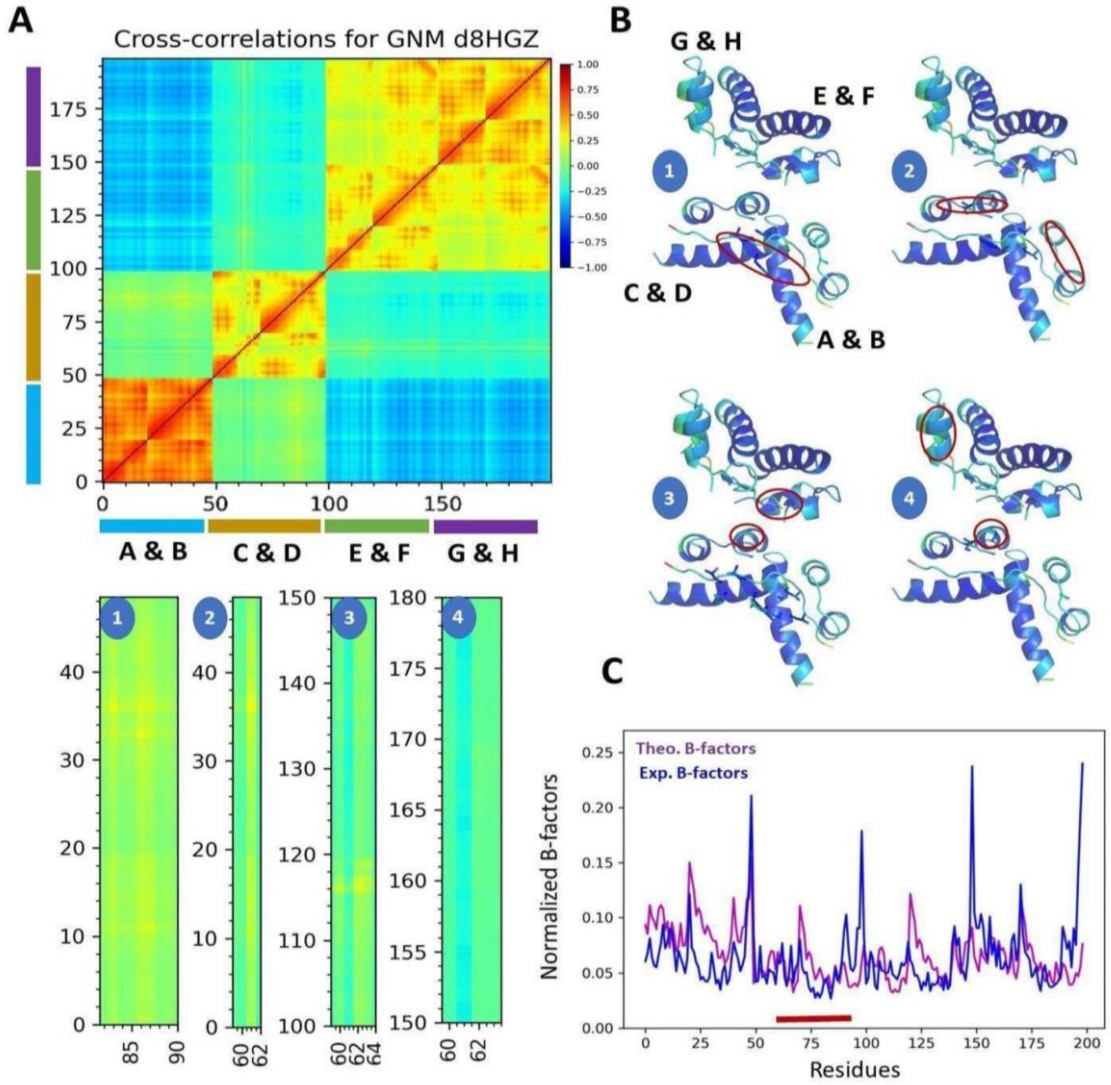


Figure 3.7 GNM Analysis Results for the isolated 8HGZ Structure.

(A) Heat map illustrating cross-correlation from overall GNM modes for our structure, highlighting ~25% correlated critical residues in the AB-CD dimer. Residue pairs moving in the same direction are shown in red regions (C_{ij}orient > 0), those moving in opposite directions in blue regions (C_{ij}orient < 0), and uncorrelated pairs in green (C_{ij}orient = 0.0; color-code bar on the right scale). (B) The cartoon representation of our structure is color-coded by the B-factor. Critical residues corresponding to the heat map are numerically labeled and highlighted with red circles. (C) Normalized B-factors of the structure are represented in blue, and theoretical B-factors calculated by GNM are shown in magenta. Critical residues corresponding to the heat map are emphasized in red lines.

Consequently, the 1st dimer of the 8HGZ structure (AB-CD) displayed a weak correlation with each other (Figure 3.7a, bottom; Figure 3.7b, top), while the 2nd dimer

(EF-GH) showed relatively higher correlated motions compared to the first dimer (Figure 3.8). In the 8HGZ crystal structure, residues between 1–40, 60–62, and 82–90 exhibited at most 25% correlation (Figure 3.7a bottom; Figure 3.7b). Moreover, interchain residues 60–64 demonstrated at most a 25% correlation with residues between 100–150 and 150–180. Conversely, the interchain residue cross-correlation map of the 1XDA structure indicated that residue 120 had ~25% correlation with residues 54–58 and residues 68–72, while residue 70 had ~25% correlation with residues 104–108 and residues 118–122 (Figure 3.6a bottom; Figure 3.7b). On the contrary, the first monomer (AB) exhibits a minimum of 75% intramolecular correlation, followed by the last monomer (GH), indicating a higher intramolecular correlation of at least 50% compared to monomer CD (~50%) and monomer EF (50%) (Figure 3.7a). At a cut-off distance of 8 Å, a distinct cross-correlation was evident in the GNM analysis of the 1XDA structure, with an overall correlation coefficient of 0.650 (Figure 3.7c). In contrast, the GNM analysis of the 8HGZ structure revealed a correlation coefficient of 0.501 (Figure 3.7c). Consequently, contrary to the theoretical B-factor calculation, the experimental B-factor calculation exhibits the highest fluctuation around the 50th, 100th, 150th, and 200th residues in the 8HGZ crystal structure (Figure 3.7c).

Nevertheless, the B-factor analysis of the 1XDA structure reveals increased fluctuation around the 100th residue compared to other regions (Figure 3.6c). Notably, critical residues within the 175–198 and 165–170 ranges are highly correlated with the 130–148 region in dimer EF-GH (Figure 3.8).

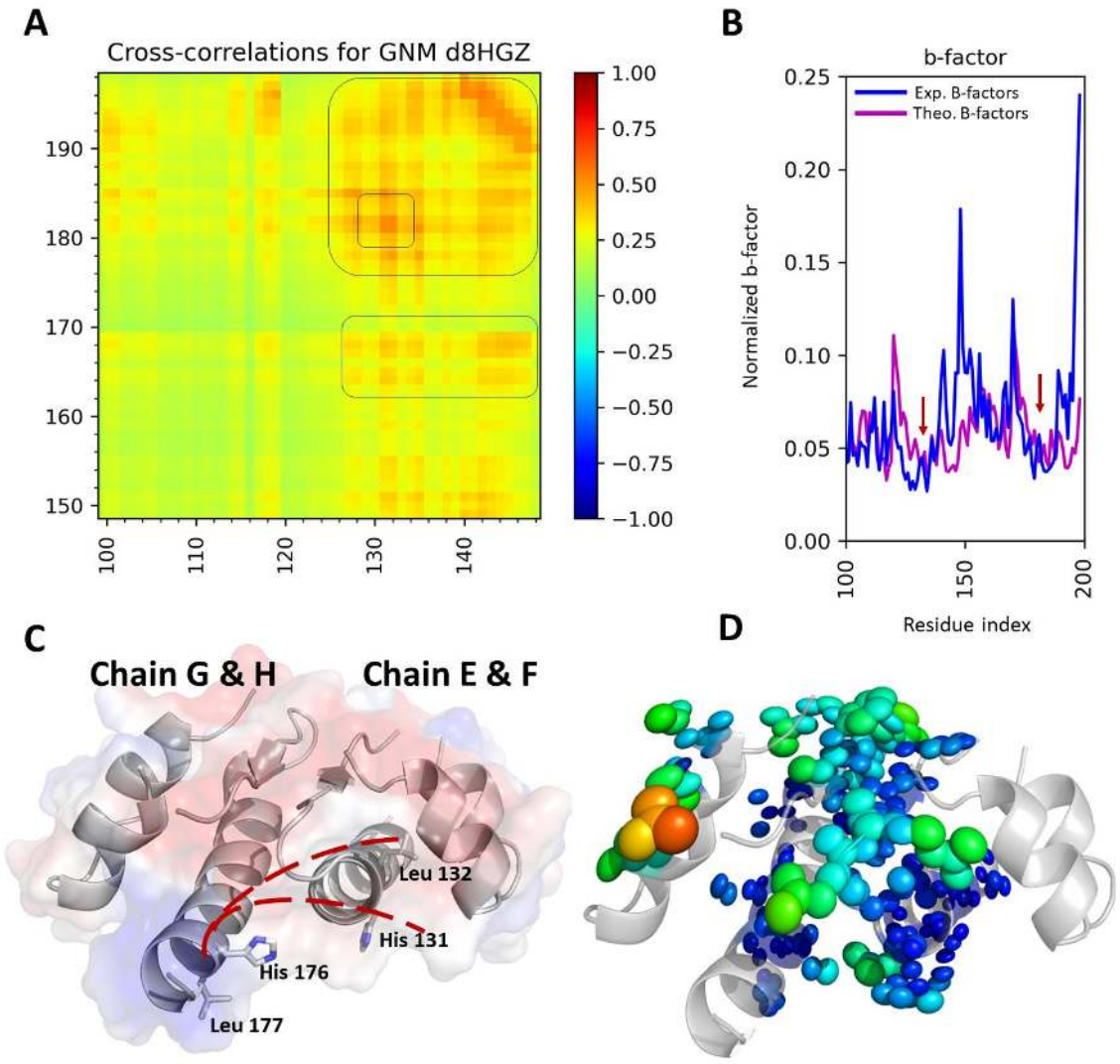


Figure 3.8 The GNM analysis outcomes pertain to the second dimer (EF-GH) within the isolated 8HGZ structure.

Notably, this second dimer displays a higher correlation than the first dimer (AB-CD; depicted in Fig. 4A, bottom). Utilizing GNM, a sliced model representation of the second dimer and the corresponding residues in the normalized b-factor plot were calculated (A, B). This analysis is further complemented by a visual presentation employing cartoon, electrostatic surface mode, and ellipsoid representations, illustrating the correlated motion of the dimer (C, D)

Specifically, His 131 and Leu 132 in EF monomer strongly correlate with His 176 and Leu 177 (Figure 3.8c), supported by normalized b-factor fluctuation. In comparing intra- and inter-monomer residue motions in the 8HGZ structure, the AB monomer exhibits at least 75% intramolecular correlation, making it the most correlated chain (Figure 3.9a-b).

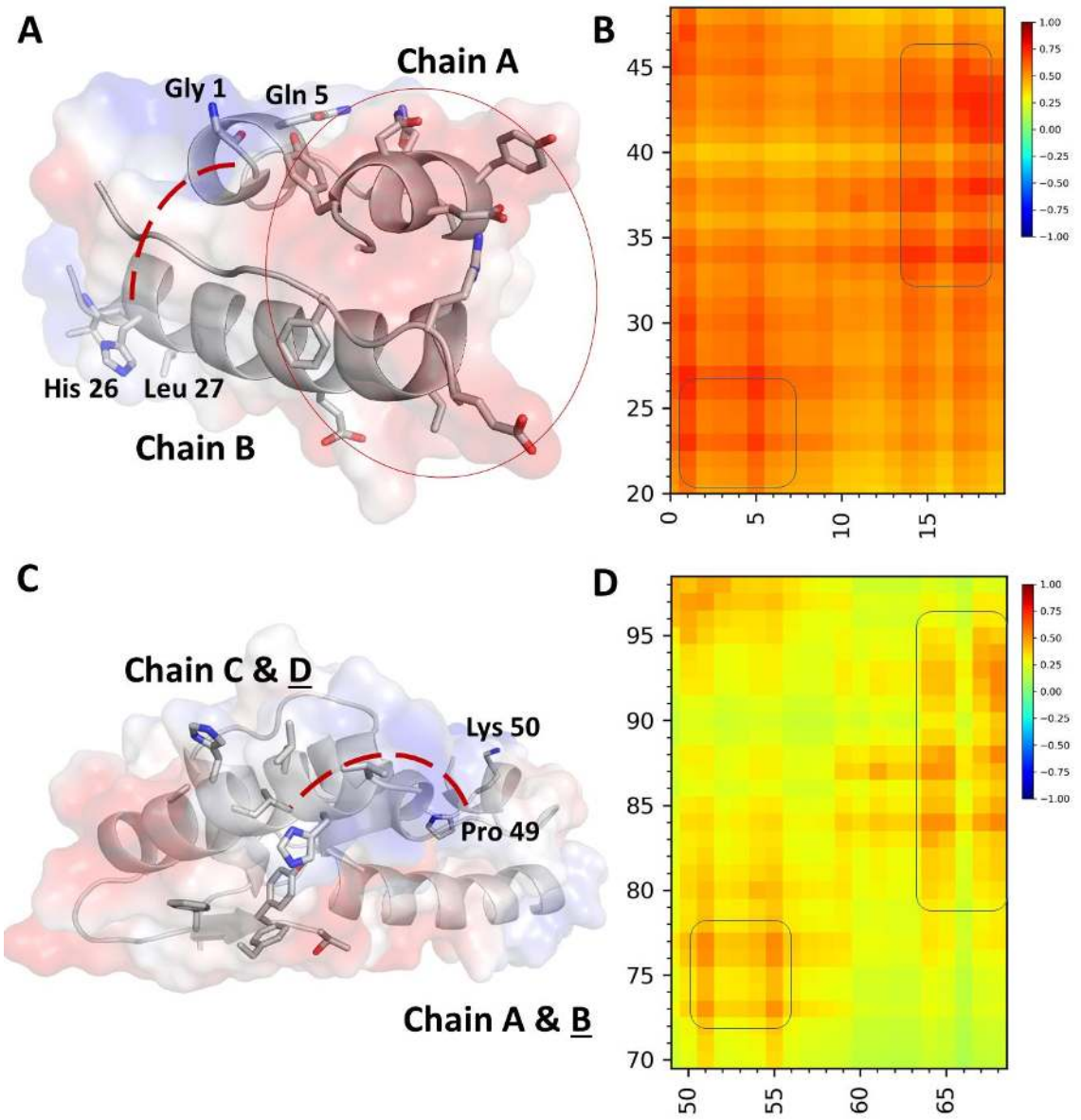


Figure 3.9 Representation of the GNM analysis outcomes, focusing on the intra- and intercorrelation motion within the AB-CD dimer.

The GNM-derived sliced model representations for chain A&B and chain C&D are provided (B, D), accompanied by a depiction using cartoon and electrostatic surface mode to highlight the correlated motion of these chains (A, C). Notably, the C-terminal residues in chain B significantly correlate with those in chain D, with the direction of movement emphasized by the red dashed line.

Intriguingly, C-terminal residues between 49–50 in chain B correlate significantly with critical residues 73, 76–77, 81–82, 85, and 92–98 in chain D (Figure 3.9c-d). Similarly, C-terminal residues 99 and 100 in chain D correlate highly with critical residues 123, 126–127, 130–131, 134, and 145–148 in chain F (Figure 3.10a-b).

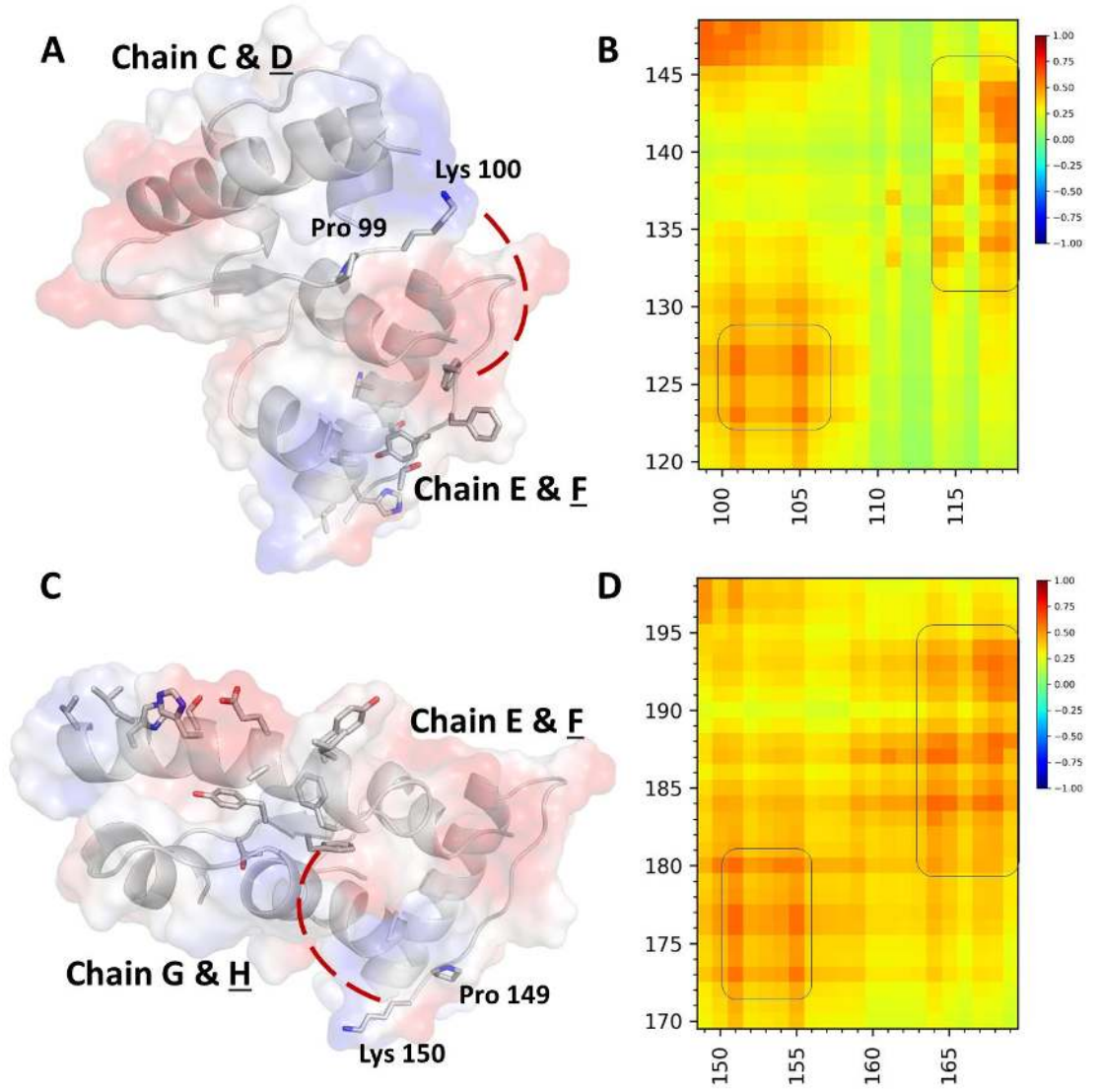


Figure 3.10 The GNM analysis results focus on the intra- and intercorrelation motion within CD, EF, and GH monomers.

The GNM-derived sliced model representations for EF and GH monomers are presented (B, D), accompanied by visual depictions using cartoon and electrostatic surface mode to illustrate the correlated motion of these chains (A, C). Noteworthy correlations include the C-terminal residues in chain D being highly correlated with chain F and the last residues of chain F demonstrating a strong correlation with chain H. The direction of movement is highlighted by the red dashed line.

Likewise, C-terminal residues 149 and 150 in chain F correlate firmly with critical residues 173, 176–177, 180–181, 184–185, 187–188, and 195–198 in chain H (Figure 3.10c-d). This indicates a consistent intramonomer correlation pattern across each chain (Figures 3.9 and Figure 3.10). A comparative cross-correlation analysis of intra- and inter-correlation motion in AB, CD, EF, and GH monomers of 1XDA and 8HGZ reveals a similar correlated pattern. However, in the 1XDA structure, intra- and inter-

molecular regions exhibit a higher and more comprehensive correlation compared to the corresponding regions in the 8HGZ structure (Figure 3.11).

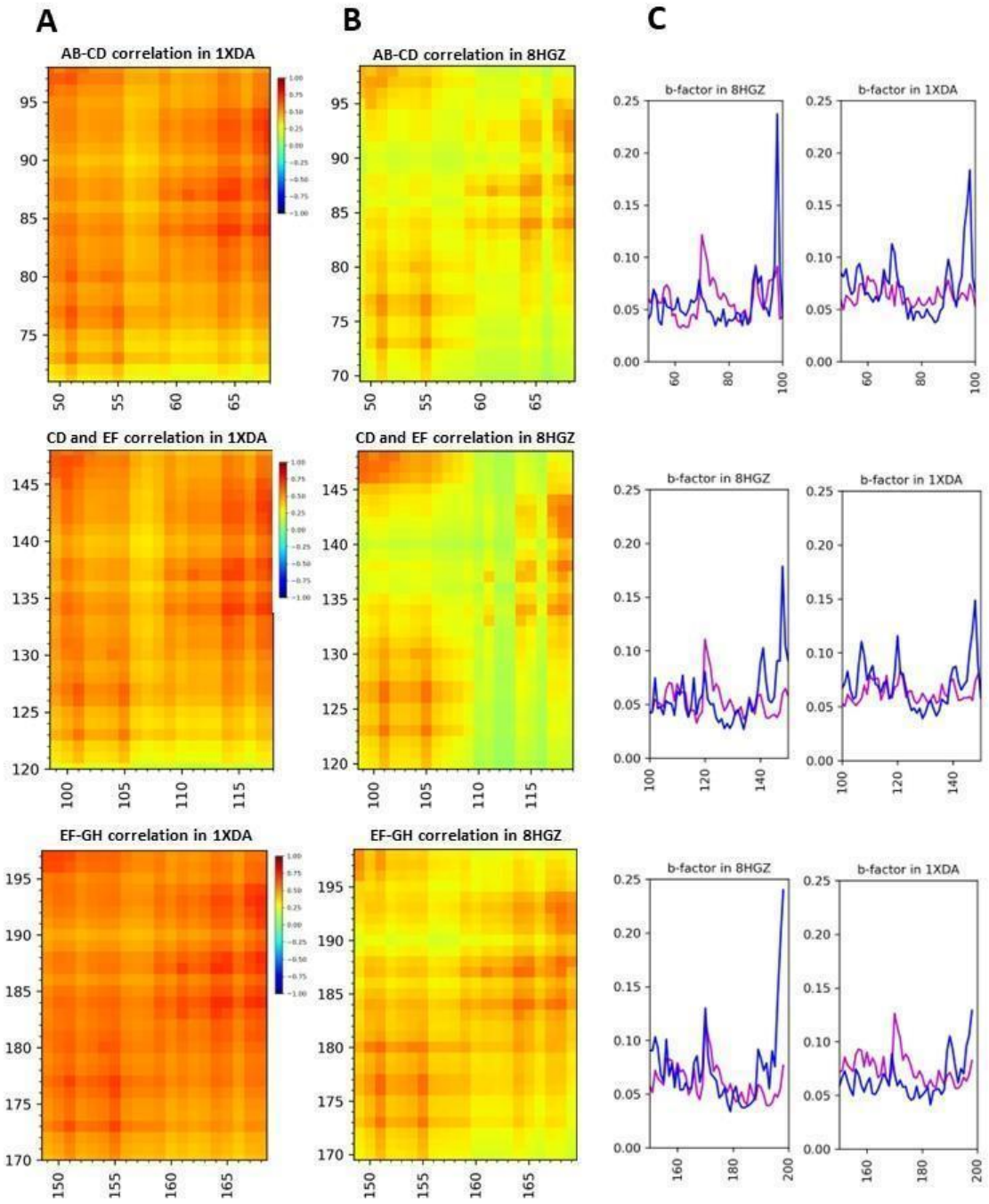


Figure 3.11 The comparative GNM analysis indicates the intra- and inter-correlation motion patterns in AB, CD, EF, and GH monomers of both 1XDA and 8HGZ structures.

(A) The monomers in 1XDA exhibit a notably high level of correlation in their intra- and intercorrelation motion compared to (B) the monomers in 8HGZ. (C) Comparative analysis of normalized b-factor values for corresponding residue indices in 1XDA and 8HGZ monomers is presented, with magenta representing theoretical b-factors and blue indicating experimental b-factors.

Furthermore, we employed the *RABDAM* program to calculate the radiation damage for two structures (Shelley et al., 2018). B damage values were determined using the total atomic isotropic B-factor values of the identified atoms. The results are illustrated in kernel density plots in Figure S11. Notably, the highest B damage value, reaching 2.05, was marked on the Asn 18 C atom (541) in chain C of the 1XDA structure. In contrast, within the 8HGZ structure, the highest B damage value, reaching 3.11, was noted on the Glu 154 O atom (1312) (Figure 3.12)

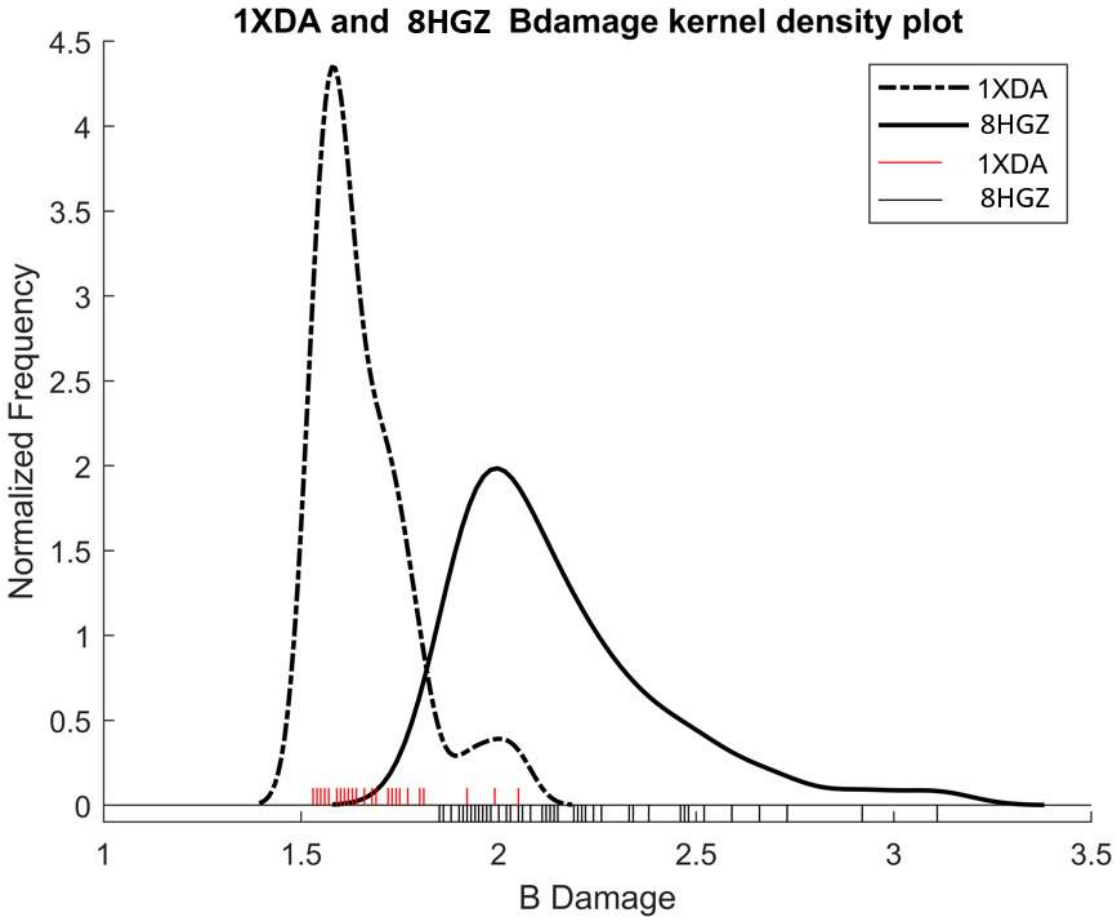


Figure 3.12 RABDAM software was employed to compute radiation damage values.

A comparison of the BDamage distribution plot between the 8HGZ and 1XDA structures indicates notable differences. The peak BDamage value of 3.11 was identified on the Glu 154 O atom (1312) within the 8HGZ structure. In contrast, the highest BDamage value (2.05) in the 1XDA structure was observed on the Asn 18 C atom (541) in chain C.

3.1.7. Concluding remarks

The protein structure of insulin detemir represents a significant advancement as it marks the first instance of a protein structure featuring an engineered fatty acid linkage/conjugation. While natural lipid modifications in proteins are common, engineering approaches provide insights into novel possibilities for oligomeric interactions, serving as a valuable biotechnological tool in drug therapy applications. Although many studies focus on insulin structures in dimer form, including a fatty acid chain in insulin, it allows for exploration in a two-dimer configuration, offering a unique perspective on the dynamic profile of the molecule from an oligomeric standpoint. The assembly of subunits, involving both covalent and noncovalent bonds, plays a crucial role in defining their structure–function relationships. Typically, phenolic ligands and zinc atoms maintain the hexameric form of insulin. The rationale behind the sustained action of the long-acting insulin detemir lies in its ability to stay in the dimer form of hexamers, determining it from other extended-action insulins. Following subcutaneous injection, phenolic ligands and fatty acid groups attached to insulin rapidly diffuse into the subcutaneous area, establishing a hexamer:dihexamer equilibrium. The accumulated forms gradually dissociate into dimers and monomers, and the myristic acid groups on insulin facilitate self-assembly, thereby prolonging its therapeutic effects (Figure 3.1).

The oligomeric configuration of insulin detemir provides valuable insights into understanding correlated motion between monomeric and dimeric states, exhibiting similarities to the 1XDA structure. In this study, we initially encountered a novel "dimer of dimer" configuration (Figure 3.2 and Figure 3.6), indicating that such oligomerization patterns may be present in the bloodstream, analogous to the conventional "dimer of dimer" form observed in insulin detemir (Figure 3.7). Additionally, we explored the interaction and correlation within each monomer and dimer state of insulin detemir, whose crystal structure was determined in a novel packing arrangement within the asymmetric unit cell (Figure 3.4). It is important to note that the structure presented here has a distinct unit cell origin compared to the 1XDA structure (Figure 3.4). In our structure, the two dimers of the 1XDA indicate an introverted state for each dimer (Figure 3.6b), while the 8HGZ exhibits an extroverted dimer form in the unit cell (Figure 3.7b). Precisely, the "dimer of dimer" form of 8HGZ

displays minor differences attributed to the crystal packing arrangement when compared to its dihexamer form (Figure 3.5).

Consequently, the GNM analysis revealed that the second dimer in the structure under consideration exhibits a more robust correlation among its components than the first dimer (Figure 3.8). Additionally, the C-terminal residues, specifically Pro and Lys, within chains B, D, F, and H, demonstrated correlated motions with the crucial regions in chains D, F, and H, respectively (Figure 5, Figures S8, and S9).

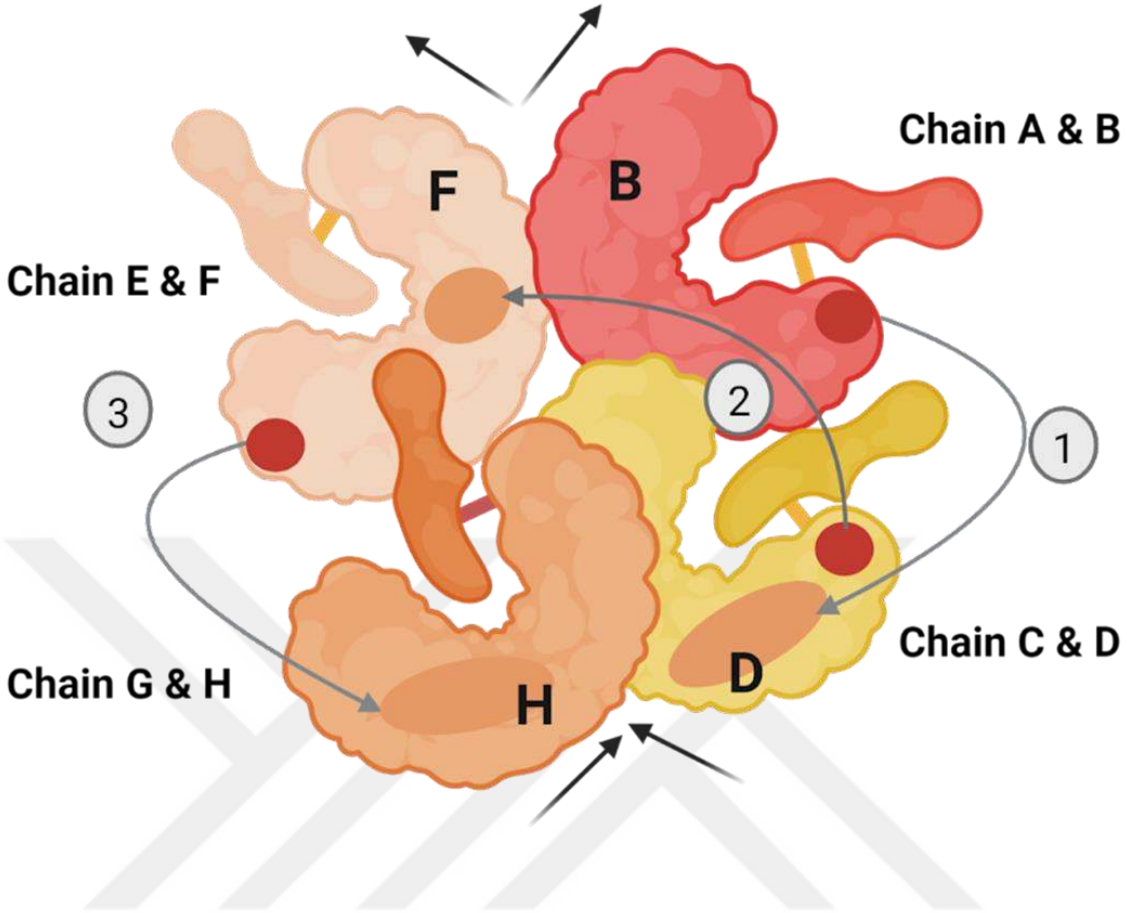


Figure 3.13 Illustrative representation of the monomeric–dimeric correlation in the cryogenic structure of 8HGZ reveals notable findings from the GNM analysis.

Cross-correlation results indicate that the C-terminal Pro and Lys residues of chain B exhibit a high correlation with chain D (1); similarly, the C-terminal Pro and Lys residues of chain D show a significant correlation with chain F (2), and the C-terminal Pro and Lys residues of chain F display correlation with chain H (3). The 8HGZ structure, consisting of two dimers or four monomers, differs from the 1XDA structure in that the AB-CD (1st dimer) adopts an extroverted form, while the EF-GH (2nd dimer) presents a slightly introverted form compared to the 1st dimer. Examining intrachain correlations, the AB monomer exhibits at least 75% correlation within itself (colored in red), the CD monomer and EF monomer display intrachain correlation up to 50% (colored in yellow and wheat), and the GH monomer demonstrates at least 50% correlation within itself (colored in orange).

Notably, the first insulin monomer exhibits a minimum intramonomer correlation of at least 75% with nearly all residues, with the final monomer displaying a higher intramolecular correlation than the other chains (Figure 3.7a). At neutral pH, insulin detemir assumes a dimer of dimer form in the asymmetric unit [Whittingham et al., 1997]. The estimated dissociation rate of an insulin dimer into two monomers in water is approximately 0.4 μs^{-1} [Acharya et al., 2021]. Our insulin detemir

crystallization occurred under a pH of 9.2. Considering that insulin dissociation is more effective at higher pH values (pH ≥ 8) and reaches its maximum around pH 10 [Fredericq, 1953], the minor distinct conformational arrangement of critical residues in the presented structure compared to the 1XDA structure could be attributed to the slow dissociation of the structure during crystallization. In a study examining conformational changes in insulin crystals across the pH range of 7–11, it was observed that these changes were closely linked to pH values, aligning with the minor conformational differences seen in the analysis of symmetry mates obtained from the 1XDA vs. 8HGZ dihexamer form [Gursky et al., 1992]. The correlated motion of each subunit depicted in Figure 5 may further confirm the causative interrelationship of oligomeric insulin's dissociation (Figures 3.9 and Figure 3.10). Furthermore, a comparative cross-correlation analysis of the 1XDA and 8HGZ structures reveals that the monomer-dimer correlation in 1XDA is consistently within the +75–100% range (Figure 3.11). Conversely, the corresponding regions in 8HGZ exhibit a weak correlation between monomers and dimers, implying a loosely coupled arrangement due to the gradual dissociation of the crystal structure induced by the elevated pH value. Additionally, the increased fluctuations observed in relevant Pro and Lys residues in normalized B-factors (Figure 3.7c) potentially validate the cooperative motion arising from the alternative conformation of the oligomeric insulin (Figure 3.13). Notably, insulin lispro, introduced as a rapid-acting insulin analog, involves reversing proline at position B28 and lysine at position B29. This substitution induces a conformational shift in the chain's C-terminal, hindering insulin monomers' ability to sterically form dimers. Consequently, the dissociation constant of the dimer is significantly reduced by 300-fold compared to regular insulin. Previous research suggests that proline exerts an inhibitory effect on the protein fibrillation process, promoting the retention of protein conformation. Analyzing the cross-correlation of C-terminal residues with neighboring monomers (Figures 3.9 and Figure 3.10), one could posit that detemir insulin exhibits oligomeric cooperativity facilitated by Pro and Lys residues (Pro 49, Lys 50; Pro 99, Lys 100; Pro 149, Lys 150). Conversely, the C-terminal lysine serves as a site for the covalent attachment of the structure with the fatty acid, contributing to the insulin's extended-release profile. When combined with the GNM findings, these observations suggest that the dissociation of the oligomeric insulin initiates with Brownian motions and fluctuations in the myristic acid.

To establish and confirm the precision of the correlated motion between monomers and dimers in prolonged insulins, coarse-grained normal mode analysis offers a robust dynamic profile for prospective investigations. Additionally, when subjected to global radiation damage, diffraction patterns may be disrupted, leading to an increase in both unit-cell volume and mosaicity. This augmentation in volume can result in non-isomorphism, posing challenges in determining the structure. Therefore, considering the radiation damage observed in the 8HGZ structure in comparison to the 1XDA structure (Figure S11), the subsequent course involves delving into the sophisticated details of this dynamic profile through time-resolved ambient temperature radiation-damage-free serial femtosecond X-ray (SFX) crystallographic studies, complemented by diffuse X-ray scattering.

3.2. A pilot study of detemir structures at ambient temperature

Oligomerization regulates the dynamics of detemir interaction with receptor.

Human insulin (hIns) predominantly exists in a balance between two oligomeric states, monomeric (5.8 kDa) and dimeric (11.6 kDa), in a slightly acidic aqueous environment [Derewenda et al., 1989]. In the presence of Zn^{2+} cations, Cl^{-} anions, and phenol, hIns tend to adopt a higher-order hexameric structural state (36 kDa) [Dunn, 2005]. Physiologically, hIns showcases various oligomeric states, introducing complexities in understanding its physicochemical properties [Dunn, 2005; Hassiepen et al., 1999]. The degree of self-assembly among insulin monomers is crucial in defining its pharmacological efficiency [Dunn, 2005]. Prioritizing the monomeric state is essential for designing hIns analogs with swift therapeutic effects, while favoring the oligomeric state leads to prolonged bioactivity [Manallack et al., 1985]. Levemir™, an engineered long-acting insulin analog, undergoes covalent modification with myristic acid at the Lys29 residue to boost hydrophobicity. This modification enables selective binding to a limited portion of serum albumin sites at therapeutic levels, reducing significant interactions with co-administered medications binding to albumin [Havelund et al., 2004; Ayan et al., 2023]. Comparative kinetic studies and euglycemic clamp analyses reveal that detemir demonstrates prolonged action, consistent pharmacokinetics, and superior temporal profiles compared to alternative long-lasting insulin analogs in humans [Brunner et al., 2000]. However, the intricate kinetics responsible for this prolonged effect, linked to the equilibrium of dimeric-monomeric dynamics, remain a puzzle.

This study presents two X-ray crystallographic detemir structures, specifically hexameric and di-hexameric configurations. These two forms were determined for the first time at ambient temperature with resolutions of 2.8 Å and 2.7 Å, respectively. The diffraction data were collected at the Turkish Light Source facility [Atalay et al., 2023; Gul et al., 2023]. Our main goal was to investigate the inherent conformational dynamics within and between molecules, focusing on understanding the distinctions between detemir's monomeric and dihexameric states. This exploration was conducted based on our structures obtained at temperatures close to physiological conditions. We employed an extensive analysis using the GNM and Molecular Dynamic (MD) simulations to investigate the local and global intrinsic fluctuations in detemir residues during their oligomerization process. Our analysis revealed clusters of residues

displaying correlated fluctuations, potentially indicating functional subdomains within the detemir structure. We delved into how the dynamics of assembly and disassembly of detemir monomers impact the flexibility of critical residues in receptor interactions.

3.2.1. Sample preparation and crystallization

The detemir, marketed under the brand name Levemir®, underwent crystallization using a sitting-drop micro batch vapor diffusion screening technique under oil, employing 72-well Terasaki crystallization plates. In a nutshell, the detemir solution was combined with a commercially available sparse matrix crystallization screening condition in a (1:1 ratio, v/v) at ambient temperature for crystallization, using approximately 3500 screening conditions. Each well, containing 0.83 µl of protein and crystallization condition, was sealed with 16.6 µl of paraffin oil (Cat#ZS.100510.5000, ZAG Kimya, Türkiye) and stored at room temperature until crystal harvesting. Crystal formation and growth were monitored by inspecting Terasaki plates under a compound light microscope. Large crystals were observed within 48 hours. The most optimal crystals were cultivated in a buffer comprising 0.2 M sodium acetate trihydrate and 0.1 M TRIS hydrochloride at pH 9.0.

3.2.2. Sample delivery and XtalCheck-S setup for data collection

Data collection was performed using Rigaku's XtaLAB Synergy Flow XRD system, which was controlled by CrysAlisPro 1.171.42.59a software (Rigaku Oxford Diffraction, 2022), following the procedure outlined in Gul et al. (2023). The airflow temperature of Oxford Cryosystems's Cryostream 800 Plus was set at 300 K (26.85 °C) and maintained consistently during data collection at ambient temperature. Instead of the intelligent goniometer head (IGH), the 72-well Terasaki plate was positioned on the modified adapter of the *XtalCheck-S* plate reader attachment mounted on the goniometer omega stage. For initial screening, two dozen crystals were utilized to assess diffraction quality. Omega and theta angles, followed by X, Y, and Z coordinates, were adjusted to center the crystals at the eucentric height of the X-ray focusing region. Subsequently, diffraction data were collected for each crystal. Crystals demonstrating good diffraction were chosen for further data collection, and the exposure time was optimized to minimize radiation damage. The best diffracting crystals were cultivated in

a buffer comprising 0.09 M HEPES–NaOH pH 7.5, 1.26 M sodium citrate tribasic dihydrate, 10% v/v glycerol (Crystal Screen Cryo, Cat#HR2-122). During data collection, *XtalCheck-S* was configured to oscillate as much as the detector distance allowed, maximizing crystal exposure oscillation angles. A total of 21 frames were collected, with each run lasting 1 minute and 45 seconds (5 seconds per frame) for all individual crystals. Thirteen crystals were utilized, with the detector distance set at 100.00 mm, the scan width at 1.00 degree oscillation, and the exposure time at 5.00 seconds per image.

3.2.3. Data processing

After optimizing plate screening parameters for all crystals, data was collected at 21 degrees for each pre-screened and selected crystal. All crystals were queued in *CrysAlisPro* for comprehensive data collection. An optimal unit cell was selected, and peak finding and masking were executed for the collected data. A batch script was generated using the 'xx proffitbatch' command for cumulative data collection. The batch data reduction was executed in *CrysAlisPro* Suite through the script command. This data reduction generated a file containing all integrated, unmerged, and unscaled data (*.rrpprof) for each dataset. To merge all datasets into a reflection data (.mtz) file, the proffit merge process from the Data Reduction section on the main window of *CrysAlisPro* was utilized. The reduced datasets (.rrpprof files) were subsequently merged again using proffit merge, as outlined. All data underwent refinement, merging, and scaling using the aimless and pointless implementations in *CCP4*. Finally, the processed data were exported to *.mtz formats.

3.2.4. Structure determination

The crystal structures of detemir were determined at ambient temperature in space group R3:H using the automated molecular replacement program *PHASER*, integrated into the *PHENIX* software package [McCoy et al., 2007]. An initial search model for molecular replacement was derived from a previously published X-ray crystal structure (PDB ID: 8HGZ) [Ayan et al., 2023]. The structural coordinates of 8HGZ were employed for the initial rigid-body refinement within *PHENIX*. Following simulated-annealing refinement, individual coordinates, as well as

Translation/Libration/Screw (TLS) parameters, underwent further refinement. Additionally, a composite omit map refinement, integrated into *PHENIX*, was conducted to identify potential positions of altered side chains and water molecules. The final model was meticulously examined and reconstructed in *COOT* [Emsley and Cowtan, 2004] version 0.8.9.2, retaining positions exhibiting strong difference density. Water molecules situated outside of significant electron density were manually excluded. All figures illustrating the X-ray crystal structures were generated using *PyMOL* [Delano, 2002] version 2.3 and *COOT*.

3.2.5. Gaussian Network Model with all Cα in the structure

We conducted normal mode analysis utilizing *ProDy* [Bahar et al., 2005]. GNM analysis was applied to two recently determined ambient temperature detemir structures and two previously identified cryogenic structures (PDB ID: 1XDA, PDB ID: 8HGZ) [Ayan et al., 2023; Whittingham et al., 1997]. Contact maps were defined using all Cα atoms, excluding myristic acids, phenols, zinc, and chloride atoms from the analysis. The cutoff distance was set at 8.0 Å, resulting in N − 1 nonzero normal modes (N = Cα). For each structure, squared fluctuations over the weighted ten slowest and ten fastest modes were calculated to compare their global and local motions. Cross-correlations between residue fluctuations were determined based on ten slowest GNM modes. Additionally, the structures were analyzed for node centrality using the network analysis of protein structures (NAPS) with default parameters on the web server [Chakrabarty and Parekh, 2016].

3.2.6. Molecular dynamics simulation

The Visual Molecular Dynamic (VMD) program was employed for preparing structure files, while molecular dynamics (MD) simulations were executed using the Nanoscale Molecular Dynamic (NAMD) program with all-atom additive *CHARMM36* force field parameters, which include a correction map for backbone atoms [Humphrey et al., 1996; Huang et al., 2017]. Protein structure files (PSF) were generated in VMD for both monomer and dimer structures, consisting of 405 and 806 atoms, respectively. Both detemir structures were placed in a water box with dimensions of 25 Å in each

direction using the solvate package in VMD. The resulting system comprised 32,152 atoms for the monomer structure and 59,795 atoms for the dimer structure.

Periodic boundary conditions for the entire protein-water system were set as x= 63.07, y= 72.99, z=73.87, and the center of the box was x= -10.89, y= 8.47, z= -29.17 for the monomer structure. For the dimer structure, periodic boundary conditions were x= 91.14, y= 75.29, z= 91.10, and the center of the box was x= -2.03, y= 12.17, z= 20.82. The TIP3P model was utilized for water molecules within the system during the simulation [Mark & Nilsson et al., 2001]. The isothermal isobaric (NpT) ensemble was employed to maintain pressure and temperature during MD simulation with the specified periodic boundary conditions. Long-range Coulombic interactions were calculated using the particle-mesh Ewald algorithm. Pressure and temperature were kept at 1 atm and 300 K, respectively, using Langevin pressure and temperature coupling. The time step for all MD simulations was set at 1 fs.

The systems underwent full energy minimization over 25,000,000 steps, with each step gradually heating from 0 K to 300 K in 2 ps. Subsequently, the systems were equilibrated under constant temperature and volume for 0.5 ns before initiating production runs. VMD was utilized for trajectory analysis and structure visualization. Root mean square displacements (RMSDs) for the backbone atoms of each protein were computed to assess their conformational stability in 50 ns MD simulations.

3.2.7. Observation

We determined the ambient-temperature structures of hexameric and dihexameric detemir at resolutions of 2.8 Å and 2.7 Å, respectively. The diffraction data were directly obtained at the Turkish DeLight facility from 72-well Terasaki micro batch screening plates, mounted directly to the *XtalCheck* plate reader adaptor module [Gul et al., 2023]. Initially, we observed detemir adopting a dihexameric conformation, while the subsequent variant exhibited a hexameric assembly within the crystal lattice (Figure 3.14). These structural arrangements display distinct 3-dimensional orientations within the asymmetric unit cell. The dihexameric structure forms an isolated "dimer of dimer" assembly, with unit cell dimensions denoted as a=, b=c=80 Å, and alpha=beta=90 gamma=120 (Figure 3.14). On the other hand, the hexameric structure adopts an isolated "dimeric" arrangement with unit cell dimensions specified as a=b=80 Å, and c=40 Å, and alpha=beta=90 gamma=120 (Figure 3.14). Both structural

configurations fall within the R3:H space group. A detailed analysis of RMSD values for each monomeric component in the two structures revealed a notable degree of conformational similarity within their asymmetric unit cells. Moreover, these structural forms encompass four myristic acid moieties, four phenol molecules, four Zn^{+2} cations, and four Cl^{-1} anions. The weak experimental unbiased *Fo-Fc* positive difference electron density map around the Lys29 residue did not definitively identify the precise conformation of covalently attached myristic acids. This weak *Fo-Fc* electron density map was attributed to the flexibility of the myristic acids [Whittingham et al., 1997] and the adverse effects of radiation-induced damage incurred during X-ray data collection (Figure 3.14).

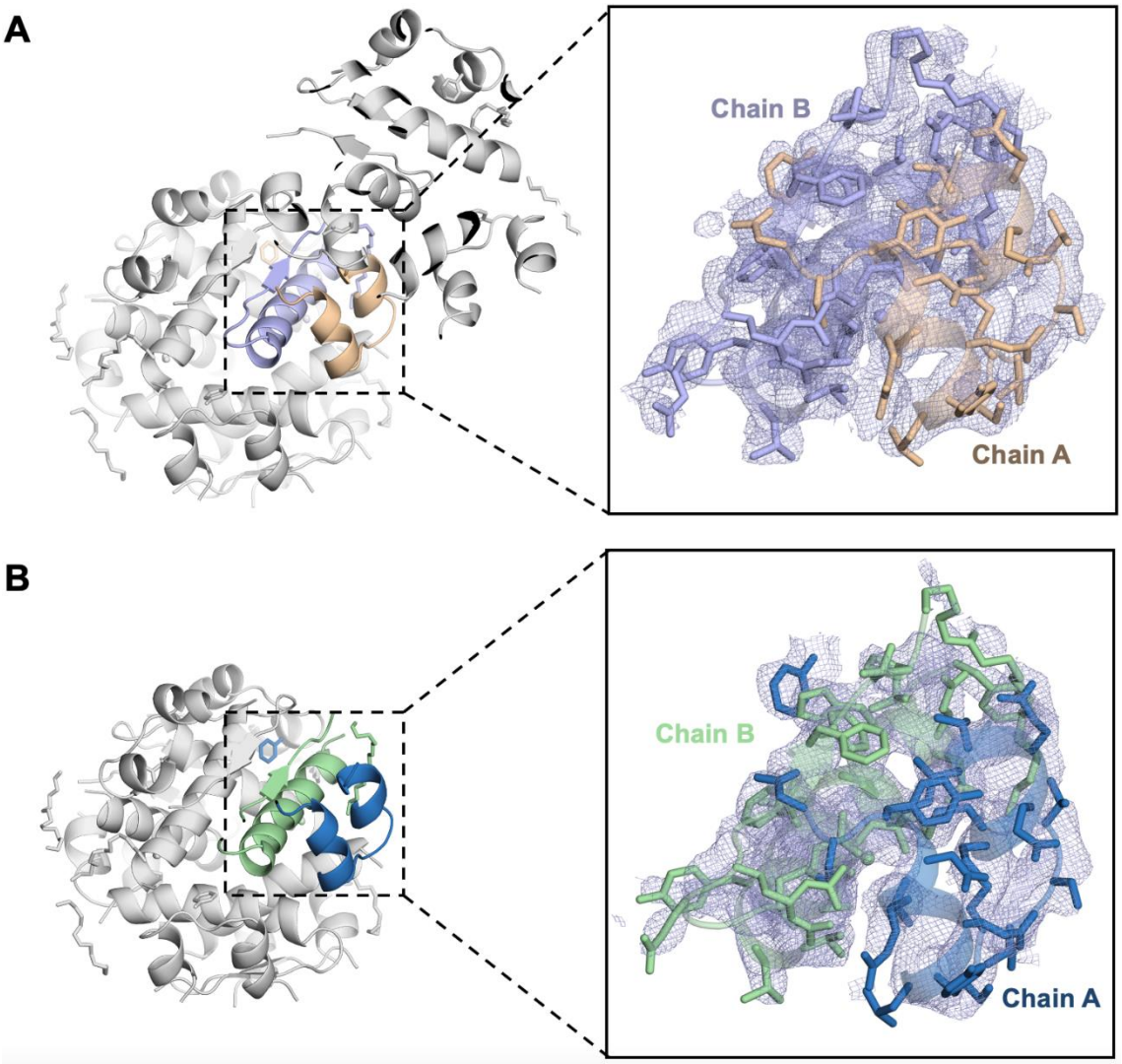


Figure 3.14 Illustration of hexameric and dihexameric detemir structures.

(A) The dihexameric structure is depicted in cartoon representation, with only the chains A and B of a single detemir monomer colored in wheat and light-blue, respectively. (B) The hexameric structure is presented in cartoon representation, featuring only the chains A and B of the detemir monomer colored in sky-blue and pale green, respectively. The *2Fo-Fc* electron density maps are delineated at the 1σ level and shaded in slate.

Table 3.2 Data collection and refinement statistics of ambient detemir

Data collection	**Detemir-dimer**	**Detemir-didimer**
Instrument	*Turkish DeLight*	*Turkish DeLight*
Resolution range	19.07 - 2.831 (2.932 - 2.831)	18.89 - 2.7 (2.797 - 2.7)
Space group	R 3: H	R 3: H
Unit cell	80.711 80.711 39.64 90 90 120	80.898 80.898 80.414 90 90 120
Redundancy	13.9 (10.2)	12.9 (11.1)
Completeness (%)	98.61 (93.21)	97.30 (93.40)
Mean I/sigma(I)	7.39 (2.39)	5.81 (2.57)
Wilson B-factor	53.57	34.78
R-merge	0.5189 (1.157)	0.5056 (0.9604)
R-meas	0.5377 (1.218)	0.5262 (1.007)
R-pim	0.1369 (0.3733)	0.1417 (0.293)
CC1/2	0.89 (0.345)	0.884 (0.737)
CC*	0.971 (0.716)	0.969 (0.921)
R-work	0.3162 (0.3762)	0.2734 (0.3511)
R-free	0.3761 (0.3547)	0.3259 (0.3337)
No. of reflections	31770 (2244)	69174 (5706)
Protein residues	100	200
RMS (bonds)	0.002	0.002
RMS (angles)	0.44	0.40
Ramachandran favored (%)	100.00	99.46
Ramachandran allowed (%)	0.00	0.54
Ramachandran outliers (%)	0.00	0.00
Average B-factor	65.67	46.98
macromolecules	63.15	41.49
ligands	108.43	55.12
solvent	39.90	99.60

The Gaussian network model (GNM) is a method based on normal mode analysis (NMA) that provides detailed insights into the overall dynamics of proteins in slow modes and localized spatial fluctuations in fast modes [Bahar et al., 1997]. Residues exhibiting significant spatial fluctuations in slow modes indicate increased

flexibility, while those with subtle fluctuations often occupy essential hinge sites crucial for protein function [Bahar et al., 1998]. In fast modes, residues that are clearly visible usually take on the role of "kinetically hot residues" [Bahar et al., 1998]. These residues are confined within the structural framework of the protein and play crucial roles in fundamental processes such as protein folding and intra- or intermolecular interactions [Bahar et al., 1998]. Protein-protein interactions are primarily established by creating new noncovalent bonds between interactive residues on the surface of two partnering proteins or subunits [Ozbek et al., 2013]. These interactions can impact residues situated at distant allosteric sites by disrupting preexisting inter-residue communication pathways and introducing new interaction possibilities [Tandon et al., 2021; Zhang et al., 2020; Atilgan and Atilgan, 2012]. This dynamic restructuring and configuration have the potential to significantly modify or modulate the conventional communication networks among residues compared to their isolated monomeric/dimeric states. The detection of such changes can be effectively unveiled through the robust analytical capabilities of GNM analysis [Bahar et al., 1997].

To gain insights into the unique dynamics of the extended detemir monomer and its oligomeric forms, we utilized NMA. The cross-correlation and B-factor analysis from the GNM provided valuable information on collective residue motions [Bahar et al., 1997]. To thoroughly understand changes in individual monomers during the formation of the detemir dihexamer, we focused on cross-correlation and B-factor analysis to obtain perspectives on modifications at the interface and distant regions. Initially, we generated residue-to-residue cross-correlation maps using the ten slowest modes from GNM. These maps served as a tool to investigate alterations in the connections between fluctuations of diverse residues within a spectrum of detemir conformations, including the monomer in the dihexamer, the monomer in the hexamer, the monomer in the "dimer of dimer," the monomer in the dimeric configuration, and the solitary monomeric state (Figure 3.15a). This study aimed to comprehensively understand the reconfiguration of residue interactions during the kinetic processes of intra- and intermolecular interactions between the dihexamer and monomer, especially during disassembly (Figure 3.15a and Figure 3.16).

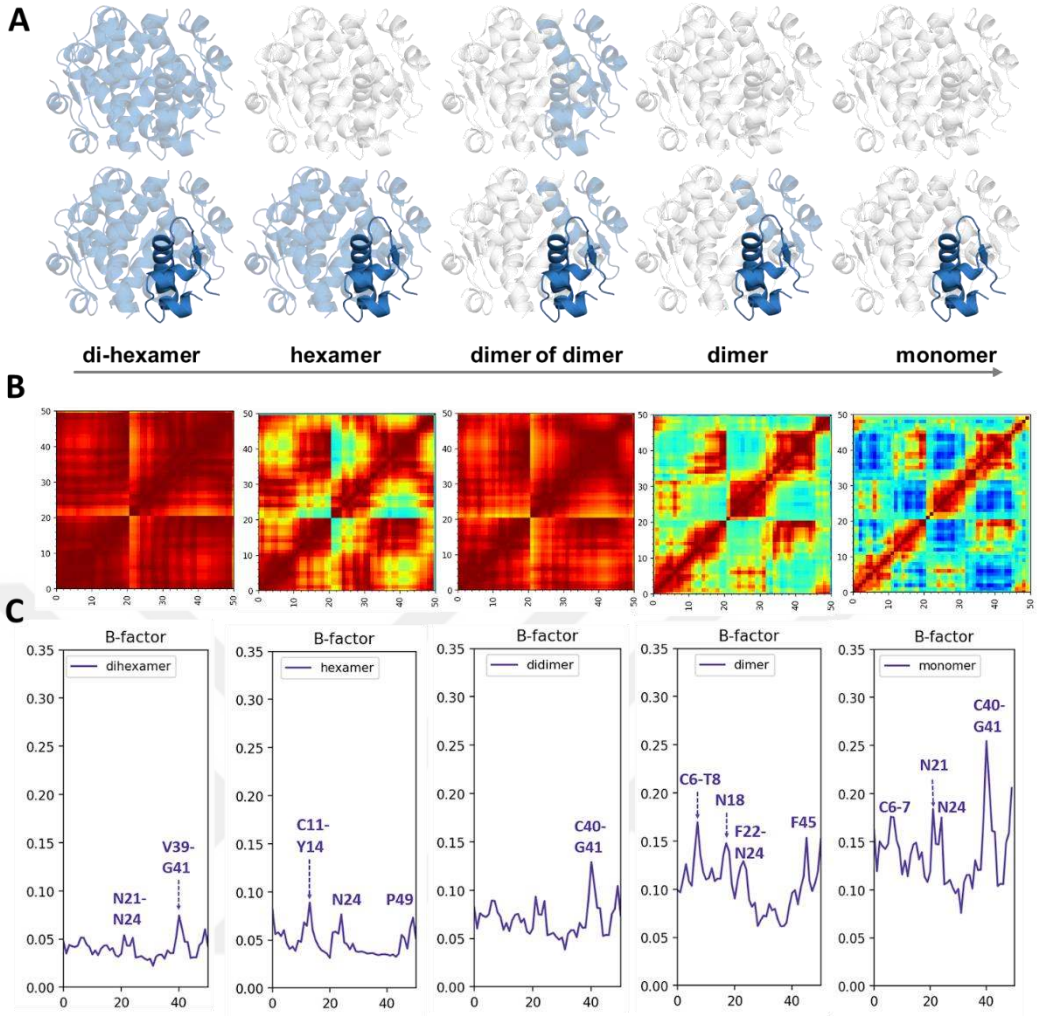


Figure 3.15 Comparative analysis of cross-correlation and B-factor evaluations between dihexamer and monomer states.

(A) Depiction of the dihexamer-to-monomer cartoon representation of detemir. (B) Heat-map illustrating the cross-correlation of the ten slowest modes from GNM for the dihexamer-to-monomer forms of detemir. (C) B-factor representation from GNM for the dihexamer-to-monomer forms of detemir.

The cross-correlation maps facilitated a meaningful comparative evaluation of the structural conformations of the monomer in its various states within complex oligomeric assemblies. This approach effectively revealed perturbations in specific subdomain regions resulting from intra- and intermolecular interactions. Interestingly, our observations indicated that residues within or close to the insulin receptor binding sites of the monomer in the dihexamer and in the "dimer of dimer" architecture exhibited significantly increased intra-correlations, surpassing the 75% threshold (Figure 3.16).

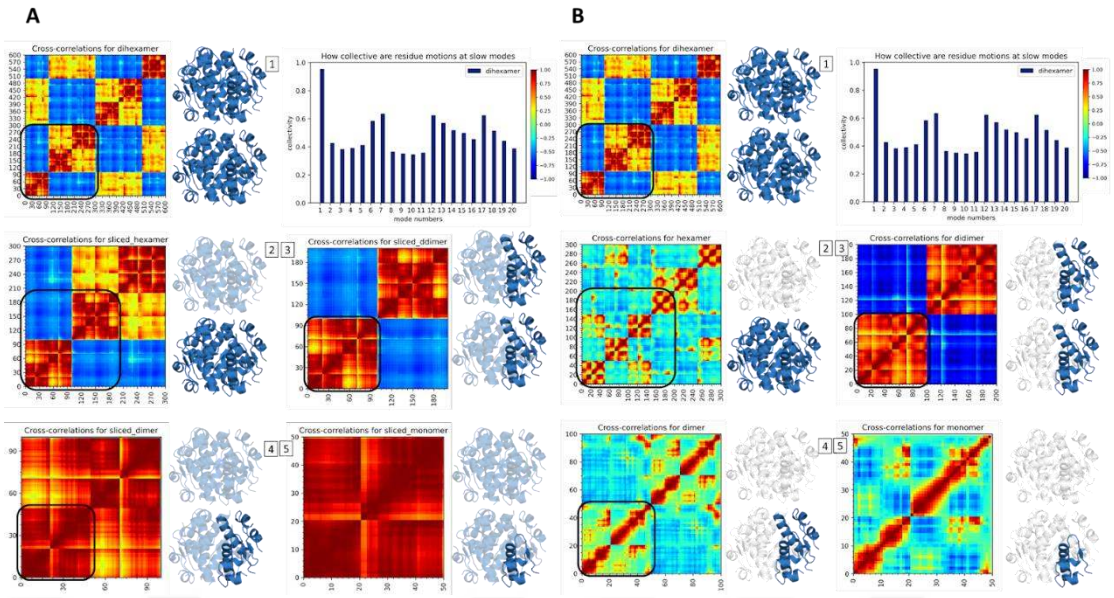


Figure 3.16 Comparative analysis of cross-correlation between dihexamer and monomer in both sliced and isolated forms.

(A) Visualization of dihexamer-to-monomer cross-correlation using the ten slowest modes from GNM analysis in a sliced version, presented in cartoon representation. (B) Heat maps depicting the cross-correlation of ten slowest modes from GNM analysis for dihexamer-to-monomer forms in the isolated version, utilizing dihexamer, 'dimer of dimer,' hexamer, dimer, and monomer structures, presented in cartoon representation. Plots illustrate collectivity analysis across up to 10 modes for the dihexamer form.

Conversely, residues positioned near the insulin receptor binding sites of the monomer in the hexamer, in the dimer, and in the isolated monomer displayed decreased intra-correlations, indicative of reduced intra- and interconnectivity (Figure 3.15b). Furthermore, in our empirical B-factor analysis, we noted that the monomer in the dihexamer exhibited an apparent reduction in structural flexibility, in contrast to other monomeric states, which displayed a gradual increase in structural flexibility (Figure 3.15c). In the other word, our observations indicate an apparent increase in the collectivity and cooperativity of insulin monomers through oligomerization (Figure 3.15a-b), consistent with the b-factor outputs that suggest a reduction in monomer flexibility (Figure 3.15c). Additionally, we noted that the N24 residue gradually modulates the flexibility of the monomer during oligomerization. Furthermore, residues C11-Y14 characterize a distinct mode in hexamerization flexibility, while residues C40-G41 are significant regions responsible for the flexibility observed in the "dimer of dimer" formation. Both dimer and monomer display more flexible regions than the oligomeric stages throughout the monomerization process (Figure 3.15c). Remarkably, both the dihexamer and "dimer of dimer" forms exhibit a similar cooperativity pattern

when compared to the cooperativity pattern observed in hexamer, despite the hexamer containing more monomers than the "dimer of dimer." This observation suggests that the critical point of cooperativity might be more closely associated with the "dimer of dimer" hierarchy than hexamerization.

We employed GNM to analyze slow modes, aiming to understand collective residue motions comprehensively. This approach allows for a detailed exploration of alterations in both inter- and intramolecular interactions during oligomerization, with a focus on the top slowest modes (Figure 3.17). The profile of mean square fluctuations from the weighted ten slowest modes revealed hinge sites characterized by local minima in the "dimer of dimer" form, extending up to 100 residues, which exhibits significant collectivity compared to the dimer form (Figure 3.17a). In the dimeric form of detemir, hinge sites were identified around residues N18-C20, L32-L36, R43-F46, N68, L82, L86, and R93-F96 (Figure 3.17a). Notably, the dimeric detemir form demonstrates increased flexibility in specific residues, including N21, F22, V23, N71, F72, and V73. In contrast, the "dimer of dimer" configuration displays increased flexibility in identical residues, specifically N21, F22, and V23 (Figure 3.17a).

In addition to exploring the overall conformational changes in detemir structures, our investigation delves into the detailed analysis of the ten fastest modes within both the "dimer of dimer" and dimer configurations of detemir (Figure 3.17b). We identified specific subdomain regions exhibiting significant spatial oscillations, indicating the presence of potential hot spots that play a crucial role in maintaining protein stability and functional efficacy. By calculating mean-square fluctuations of the fastest ten modes, we determined specific residues as kinetically hot regions, finding an overlap in both dimer and "dimer of dimer" structures. Consequently, the "dimer of dimer" configuration revealed a new set of potential hotspots, including Q15, L32, A35, L36, R43, G44, L82, A85, L86, R93, and G94 residues (Figure 3.17b). Notably, among these residues, A35, R43, G44, A85, L86, R93, and G94 are directly associated with insulin receptor binding [Pullen et al., 1976; De Meyts, 2016]. These findings were further corroborated through MD simulations conducted on both dimeric and monomeric detemir conformations (Figure 3.20). Conversely, we identified distinct amino acid residues, namely C6, Q15, Y19, I52, Q65, N68, and V89, along with the identical kinetically hot residues found in the "dimer of dimer" form (Figure 3.17b). Among them, Y19 and I52 establish a direct and unequivocal connection with the

molecular interaction between insulin and its receptor [Pullen et al., 1976; De Meyts, 2016].

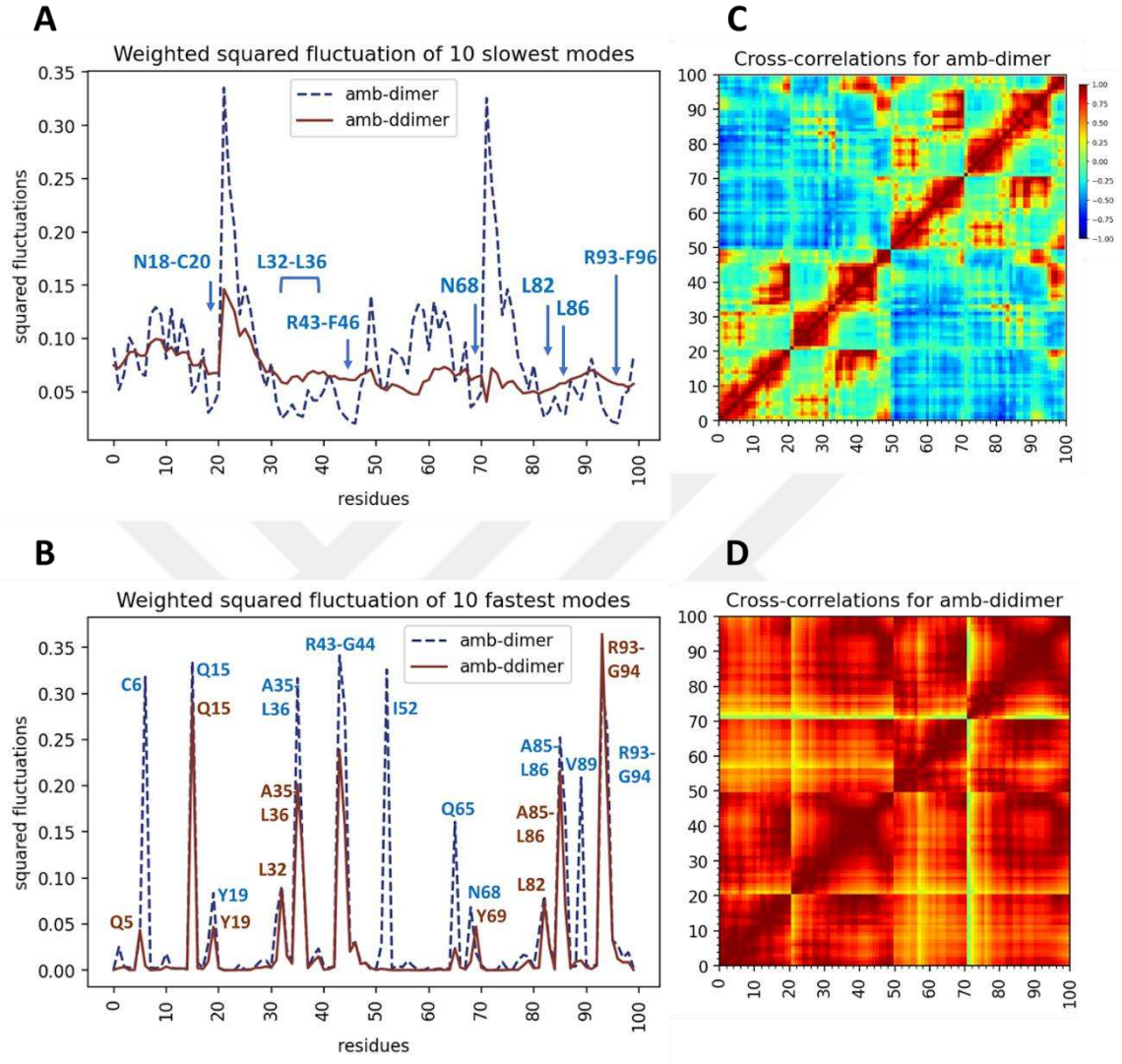


Figure 3.17 Unveiling the associations in residue fluctuations between the dimeric form and the 'dimer of dimer' configuration of detemir from GNM.

(A) Normalized weighted squared fluctuation of the ten slowest modes across dimer and 'dimer of dimer' structures. (B) Normalized weighted squared fluctuation of the ten fastest modes across dimer and 'dimer of dimer' structures. (C) Cross-correlation heat-map of the dimer and (D) 'dimer of dimer' structures spanning up to 100 residues (the 'dimer of dimer' form consists of 200 residues). Amb-dimer: dimer configuration at ambient temperature, amb-ddimer: 'dimer of dimer' configuration at ambient temperature.

A thorough comparison between the dimer or "dimer of dimer" and the hexamer or dihexamer has been elucidated in Figure 3.18 and Figure 3.19 This comparison utilizes measures of collectivity, weighted squared fluctuation of the ten fastest and ten

slowest modes, b-factor analyses, and evaluations about ten individual modes (modes 1-10) of dimer and "dimer of dimer" forms.

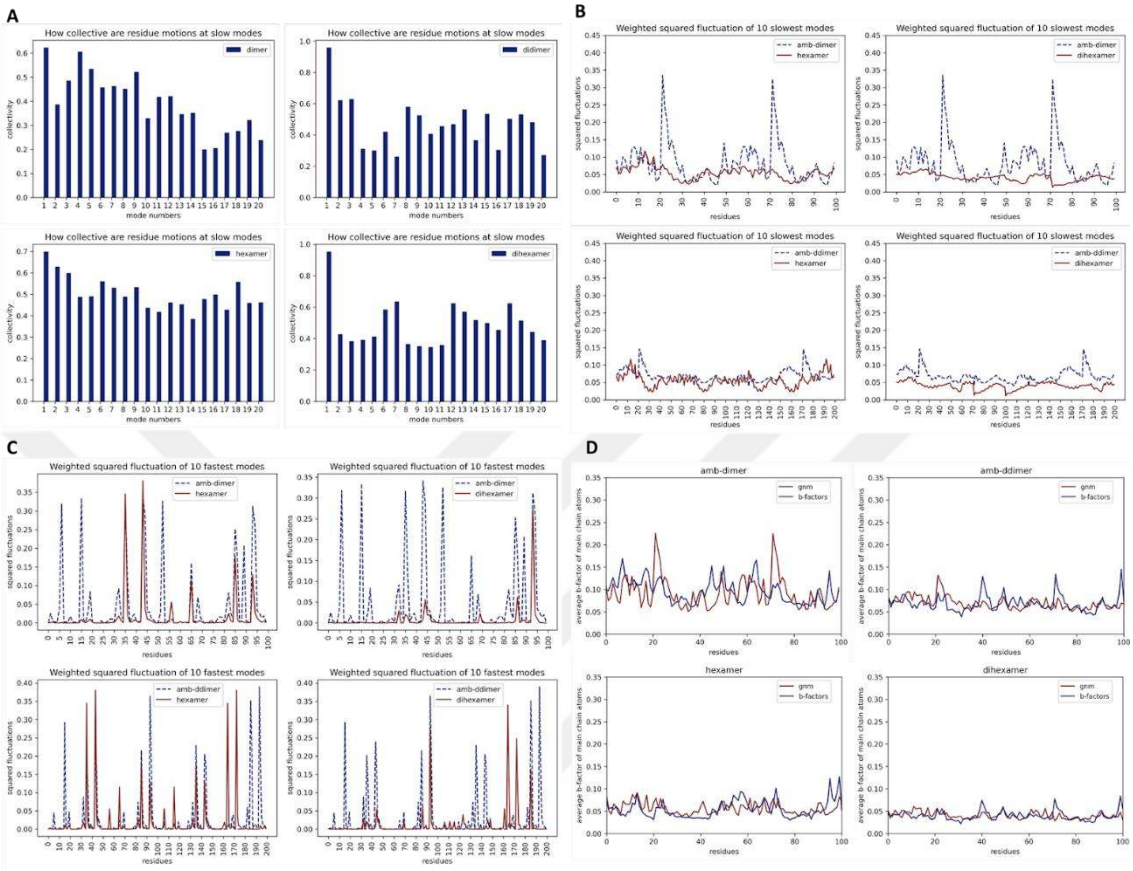


Figure 3.18 A thorough comparison was undertaken through normal mode analyses between the dimer or 'dimer of dimer' and hexamer or dihexamer structures.

Panel (A) illustrates the collectivity analysis for ten individual modes from GNM. Panels (B and C) portray the weighted squared fluctuation of the ten slowest and fastest modes from GNM. Panel (D) showcases the B-factor analysis obtained from GNM.

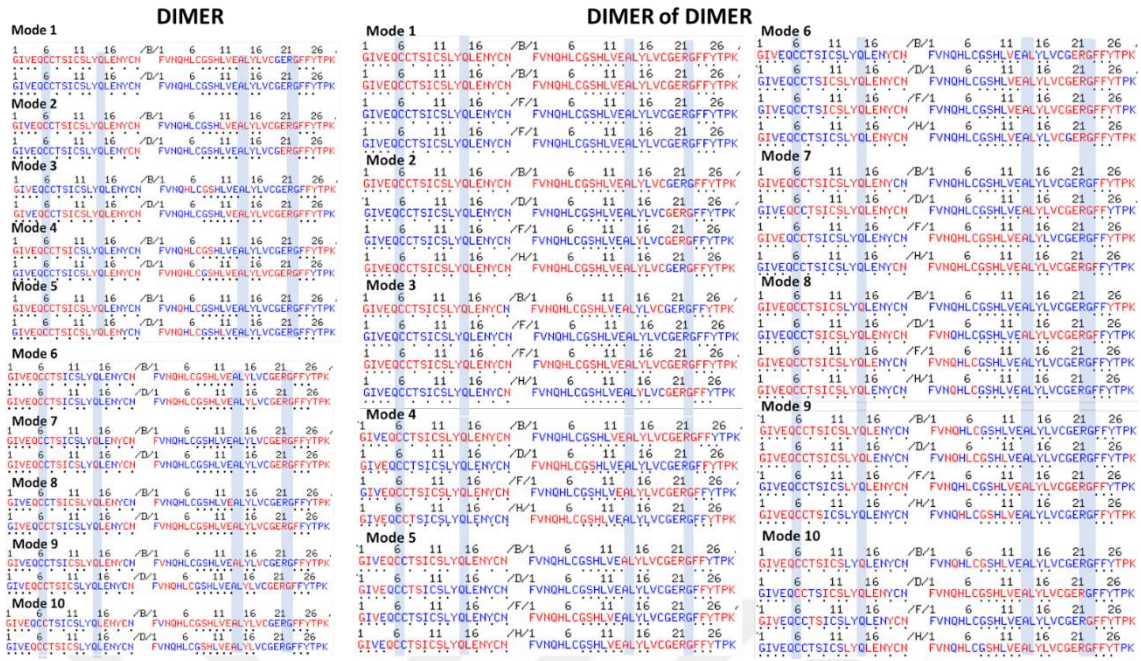


Figure 3.19 Illustration of the ten individual slow modes (1-10) of dimer and dihexamer forms arranged in a sequence-based.

The transition from blue to red sequences signifies dynamic variations. Receptor-binding residues are denoted by dots, and residues supported by the slowest and fastest modes from GNM are represented by highlighted blue sticks.

We constructed a cross-correlation heat map depicting residue fluctuations, utilizing data from the normal ten slowest modes from GNMs (Figure 3.17c-d). This analysis aimed to elucidate how the correlations in the fluctuations of different residues evolve during the transition from the dimeric form to the "dimer of dimer" configuration, providing insights into how residues communicate upon assembly. Consequently, the cross-correlation maps objectively compare the two distinct detemir forms. Consistent with the ten slowest modes from GNM, the dimer and "dimer of dimer" configurations exhibited distinct cooperative motion. Specifically, the dimeric form appeared to have lower correlation compared to the "dimer of dimer" form across up to 100 residues (Figure 3.17c). In other words, when contrasted with the dimeric form, the "dimer of dimer" configuration demonstrates a higher level of intra- and intermolecular cooperativity (Figure 3.17d). This comparative analysis underscores that detemir conformations represent and characterize subdomain regions introduced during the protein assembly process. Notably, within the "dimer of dimer" detemir structure, residues at or near the insulin receptor binding interfaces exhibit significantly stronger collectivity than those within the dimeric detemir structure.

We conducted MD simulations to gain deeper insights into the structural distinctions between the dimeric and monomeric configurations of detemir. Throughout the 50-nanosecond simulations of detemir's dimeric and monomeric structures, we observed numerous intra- and intersubunit molecular interactions. Specifically, in the dimeric form, substantial hydrogen bond interactions were identified in the insulin receptor binding sites (Figure 3.20), a finding supported by a cross-correlation heat map from GNM (Figure 3.21) and the slowest and fastest modes from GNM (Figure 3.17).

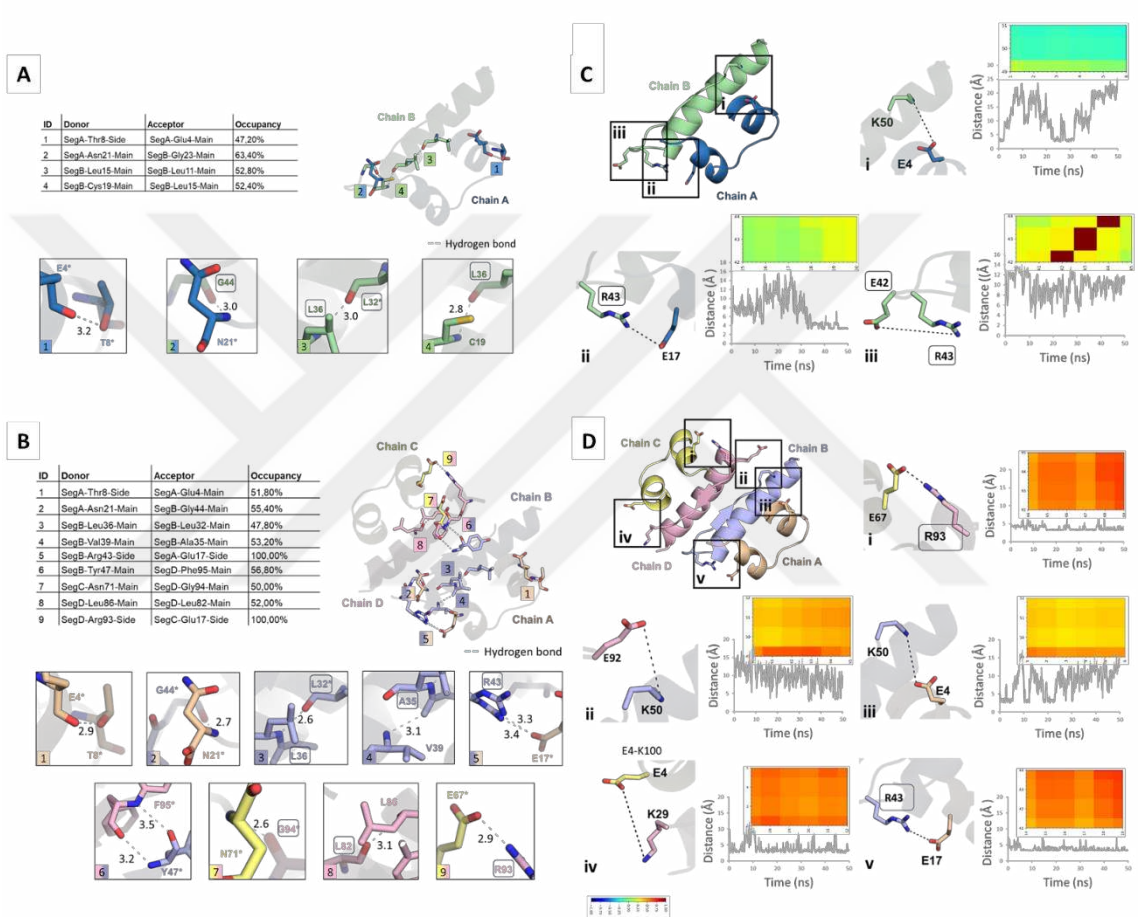


Figure 3.20 Molecular interactions in the monomer (from dimer) and dimer (from 'dimer of dimer') structures of detemir and distance calculation of salt bridges within a 50 ns MD simulation. (A) Key hydrogen bond interactions in the monomer structures with over 50% occupancy are presented in the table, and the interacting residues are visualized in PyMOL. (B) Major hydrogen bond interactions in the dimer structure with over 50% occupancy are listed in the table, and the residues are depicted in PyMOL. Each monomer is represented by a distinct color. Kinetically hot residues are marked with squares. The hydrogen bond distances (Å) between each residue are illustrated with dashed lines in pale cyan color. (C) Identified salt bridges in the monomer structure are displayed, and distances (Å) are calculated in VMD during the 50 ns MD simulation. (D) Detected salt bridges in the dimer structure are depicted, and distances (Å) are computed in VMD within the 50 ns MD simulation. Each monomer is distinguished by a different color. The corresponding residues are supported by cross-correlation analysis from GNM. Kinetically hot regions are highlighted with squares.

Notably, specific residues, including E4, T8, N21, L32, G44, Y47, E67, N71, G94, and F95, exhibited direct correlations with the receptor binding sites during the MD simulations (Figure 3.20b) [Pullen et al., 1976; De Meyts, 2016]. The analysis of the ten fastest weighted square fluctuations also reinforced the significance of residues L32, A35, L36, R43, L82, G94, and R93 (Figure 3.17b). Similarly, in the monomeric detemir, we identified notable hydrogen bond interactions in critical receptor binding residues, consistent with the observations from the slowest and fastest modes of GNM (Figure 3.20). During the MD simulation, interacting residues such as E4, T8, L11, N21, and G23 were directly associated with the active insulin receptor binding sites (Figure 3.17) [Pullen et al., 1976; De Meyts, 2016]

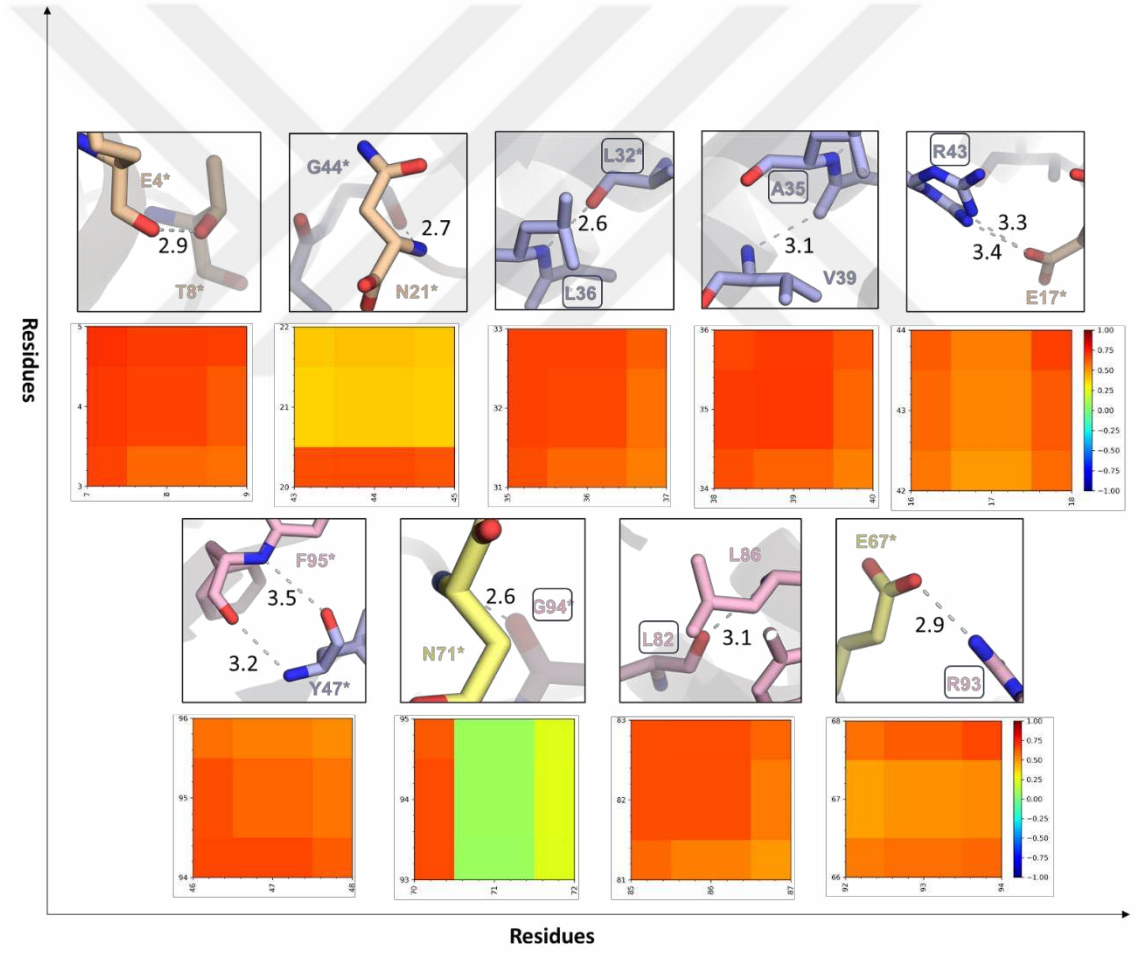


Figure 3.21 A cross-correlation heat map generated from GNM reinforces the molecular interactions observed in MD simulations.

The major hydrogen bond interactions presented in each panel exhibit correlations of at least 50%, with the exceptions of N71-G94 (~0%) and N21-G44 (~at least 25%)

Moreover, we conducted an interatomic distance analysis to identify salt bridges during a 50 ns simulation, predominantly observing interactions between Lys/Arg and Glu residues (Figure 3.20). These molecular interactions play a crucial role in both intra- and intermolecular interactions of detemir. Accordingly, we observed more considerable Cα distances between E4-K50, E17-R43, and E42-R43 residue pairs in the monomeric form, supported by GNM cross-correlation, revealing correlations of up to 25% between these residues (Figure 3.20a). On the other hand, the dimeric form of detemir exhibited lower fluctuations in Cα distances for E67-R93, E4-K29, and E17-R43 residue pairs Figure 3.20b), also supported by GNM cross-correlation analysis. This suggests strong correlations between residues E67-R93, E4-K29, and E17-R43 in the dimeric form. At the same time, K50-E92 and E4-K50 indicated less than 50% correlation with each other (Figure 3.20b). Additionally, the R43 and R93 residues identified from MD salt-bridge analysis were supported by the fastest mode in NMA (Figure 3.20b). A comprehensive illustration of monomer and dimer MD simulations is provided in Supplementary Figure 3.22, incorporating polar interaction, RMSD, and RMSF analyses.

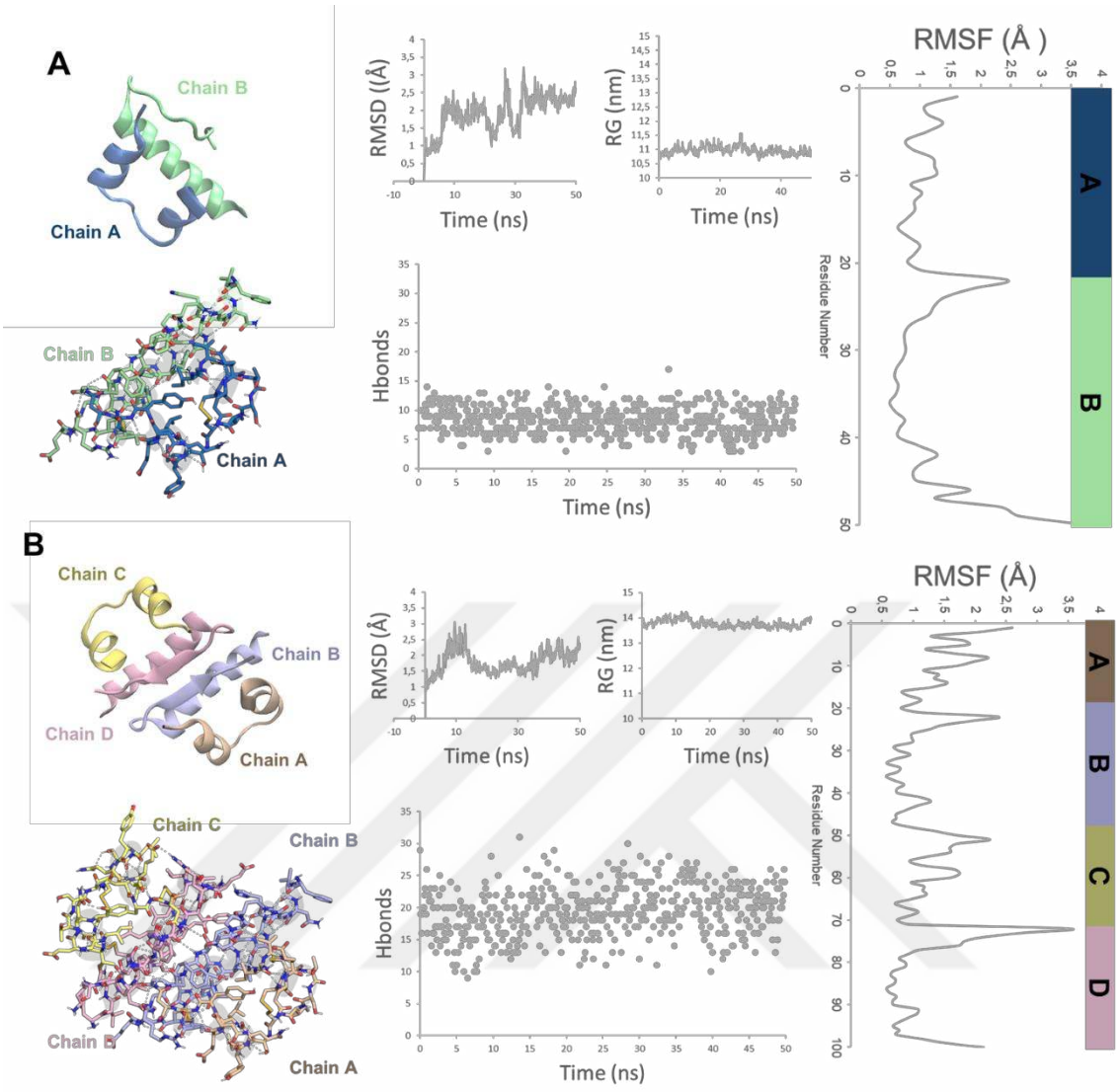


Figure 3.22 Structural dynamics of monomeric and dimeric structures.

(A, B) RMSD (root mean square deviation in Å) and RMSF (root mean square fluctuation in Å) were calculated to identify fluctuating residues during the 50 ns MD simulation for both monomeric and dimeric structures. The total count of hydrogen bonds formed throughout the 50 ns MD simulation is illustrated separately for monomeric and dimeric structures.

We performed cross-correlation analysis using GNM to validate whether the regions highly correlated in the 'dimer of dimer' form encompass residues involved in receptor interactions and newly identified subdomains. Consequently, we identified highly correlated regions in the "dimer of dimer" form, displaying at least a 75% correlation (Figure 3.23). These regions primarily include residues associated with insulin receptor binding sites. Specifically, within Chain A and B of the "dimer of dimer" form, residues I2, I10, S12, Y19, N24, G29, S30, H31, L32, V33, L36, L38, F45, F46, Y47, L82, V83, F95, and Y97 exhibit at least 75% correlation (Figure 3.23c-

d). Collectively, it is evident that the highly correlated residues identified constitute regions related to insulin receptor interaction sites and newly discovered interactive residues, findings supported by both MD and NMA (Figure 3.17b; Figure 3.20a-b; Figure 3.23). However, it's worth noting that these residues show weaker correlations, ranging up to 25%, or, in some cases, no correlation at all in the dimeric form. Interestingly, within the range of x=90-99 and y=40-49 residues, the residues involved in insulin interactions do not exhibit a significant correlation, unlike the situation observed in the 'dimer of dimer' form. Instead, in the dimer form, we keep that I2, C6, G91, G92, and G41 emerge as highly cooperative residues (Figure 3.23a-b).

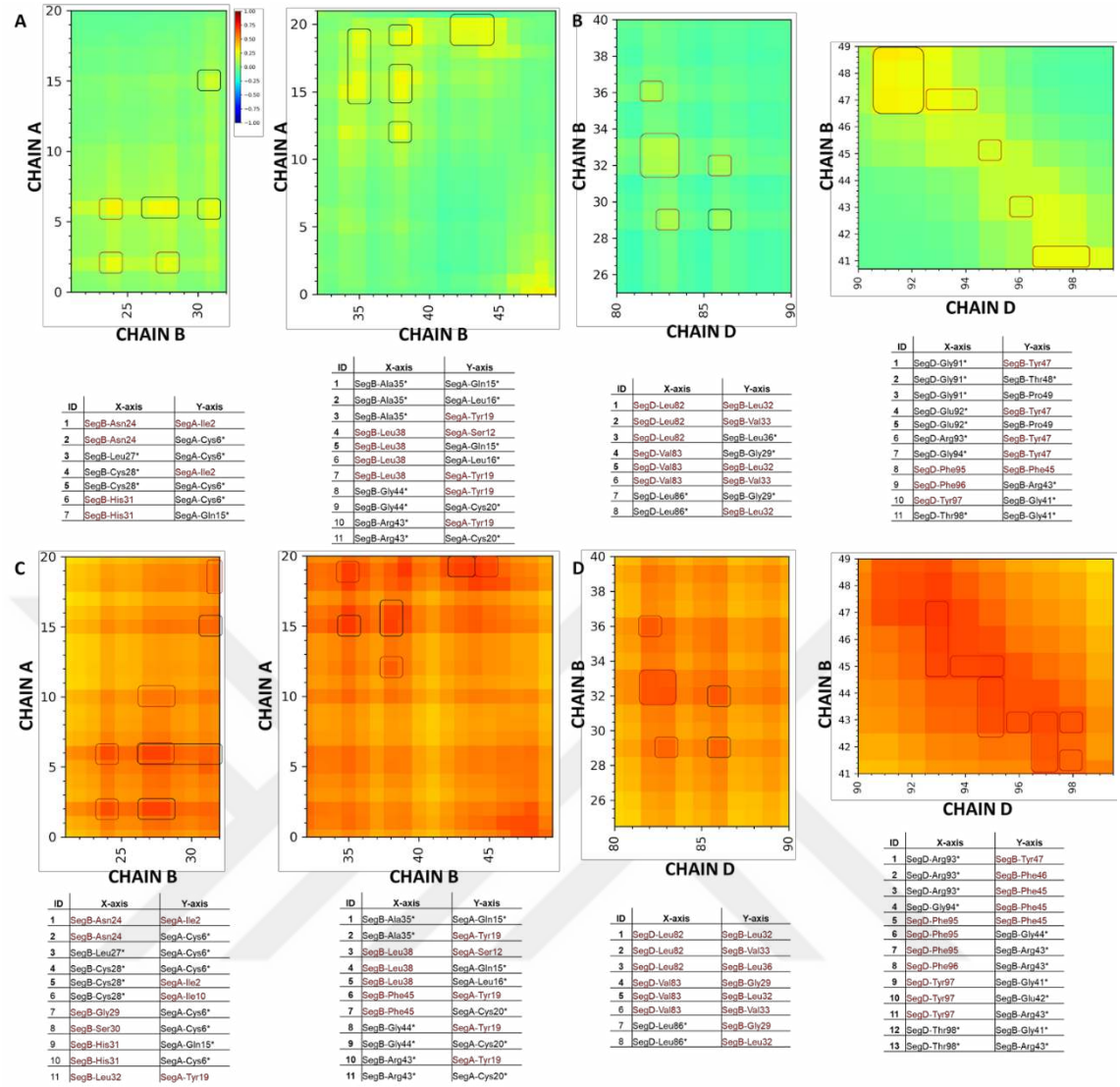


Figure 3.23 Comparative analysis of insulin receptor binding sites in dimeric and 'dimer of dimer' configurations using cross-correlation from GNM.

Cross-correlation analyses were conducted for residues within x=0-31, y=0-20 and x=0-48, y=0-20 regions for the dimeric form (A) and 'dimer of dimer' form (C). Similar analyses were performed for residues within x=80-90, y=25-40 and x=90-100, y=40-50 regions for the dimeric form (B) and 'dimer of dimer' form (D). Correlated residues are listed in the table, and insulin receptor interaction residues are highlighted in red. Significant residues supported by GNM are marked with an asterisk.

We conducted a molecular docking analysis to explore the dynamic interaction between the insulin receptor and significant residues identified through GNM and MD simulation. Our findings yielded two potential outcomes: the initial docking interaction of detemir with IR appeared sparser than the second docking interaction. Specifically, the analysis highlighted that the interaction of N18 and C20 in the A chain of detemir with IR buried respective totals of Q15 from the solvent, in contrast to C6 and Q5 for

the A chain. On an individual residue level, the most notable distinction between the complexes in Figure 3.24c-dwas observed in the occupation of the side chains (N24, Q25, H26, L27, R43, G44) of the potential hot loci in the B chain, primarily located on the periphery to influence B1-8 residues accordingly. R43-G44 may play a role in the engagement of the primary binding site, influencing the incorporation of F45-F46 with the L1 domain. However, the presented holo-receptor indicates that it is either not directly involved in primary binding-site engagement or requires secondary stimuli (A25 and L36) for primary binding-site interaction.

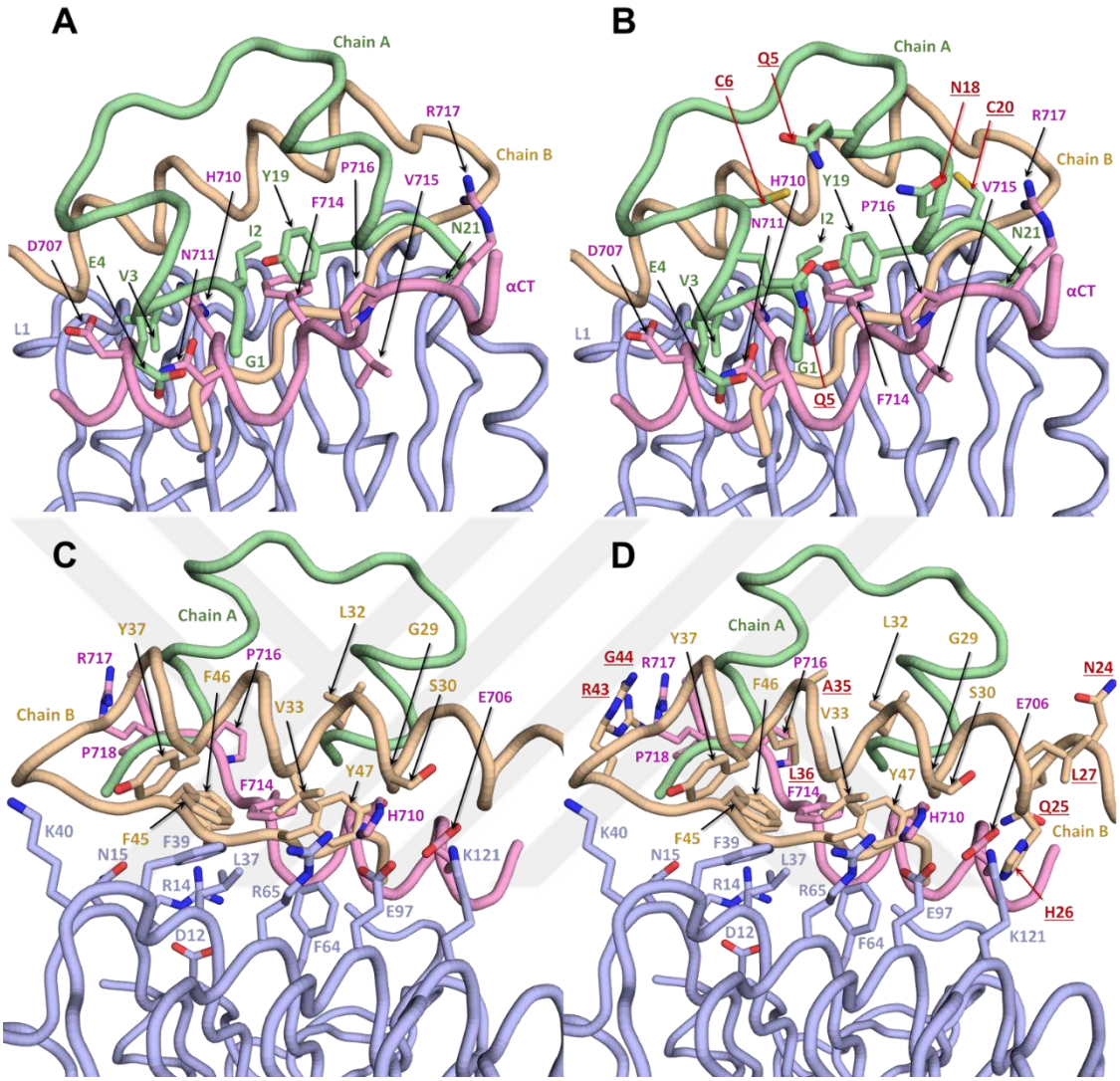


Figure 3.24 Illustration of the structural insights from the molecular docking analysis of the detemir monomer with the insulin receptor (IR).

(A) Depicts the interaction between the A chain of detemir and IR. (B) Shows the interaction involving newly identified residues of the detemir A chain with IR. (C) Illustrates the interaction of the B chain of detemir with IR. (D) Highlights the interaction of newly identified residues of the detemir B chain with IR. The A chain is represented in pale green, while the B chain is depicted in wheat. Additionally, the aCT domain of IR is shaded in pale-pink, and the L1 domain of IR is colored in slate-blue.

To evaluate the consistency of enhanced correlations within oligomeric structures, we conducted a comparative analysis involving cryogenic "dimer of dimer" structures (PDB ID: 1XDA and PDB ID: 8HGZ) and our ambient "dimer of dimer" and dimeric forms of detemir. Notably, we identified a distinct pattern of similarity within the 'dimer of dimer' structures based on their slowest and fastest modes from GNM,

contrasting with the behavior observed in the dimeric form (Figure 3.25a-h). Moreover, we observed reduced flexibility in all "dimer of dimer" structures compared to the dimeric form (Figure 3.25i-l). Additionally, we applied network analysis of protein structures (NAPS) to examine structures through node centrality, using default parameters on the web server. Node centrality, in this context, represents the topological importance of a residue within the residue-residue contact network. Further topological analysis of the four structures, including two cryogenic and two ambient forms, revealed that residues with the highest degree centrality in our ambient "dimer of dimer" align with the hinge regions in the dimeric form. These residues in the "dimer of dimer" configuration also display high betweenness centrality, indicating a significant influence on the information flow within the residue contact network (Figure 3.25m-o)

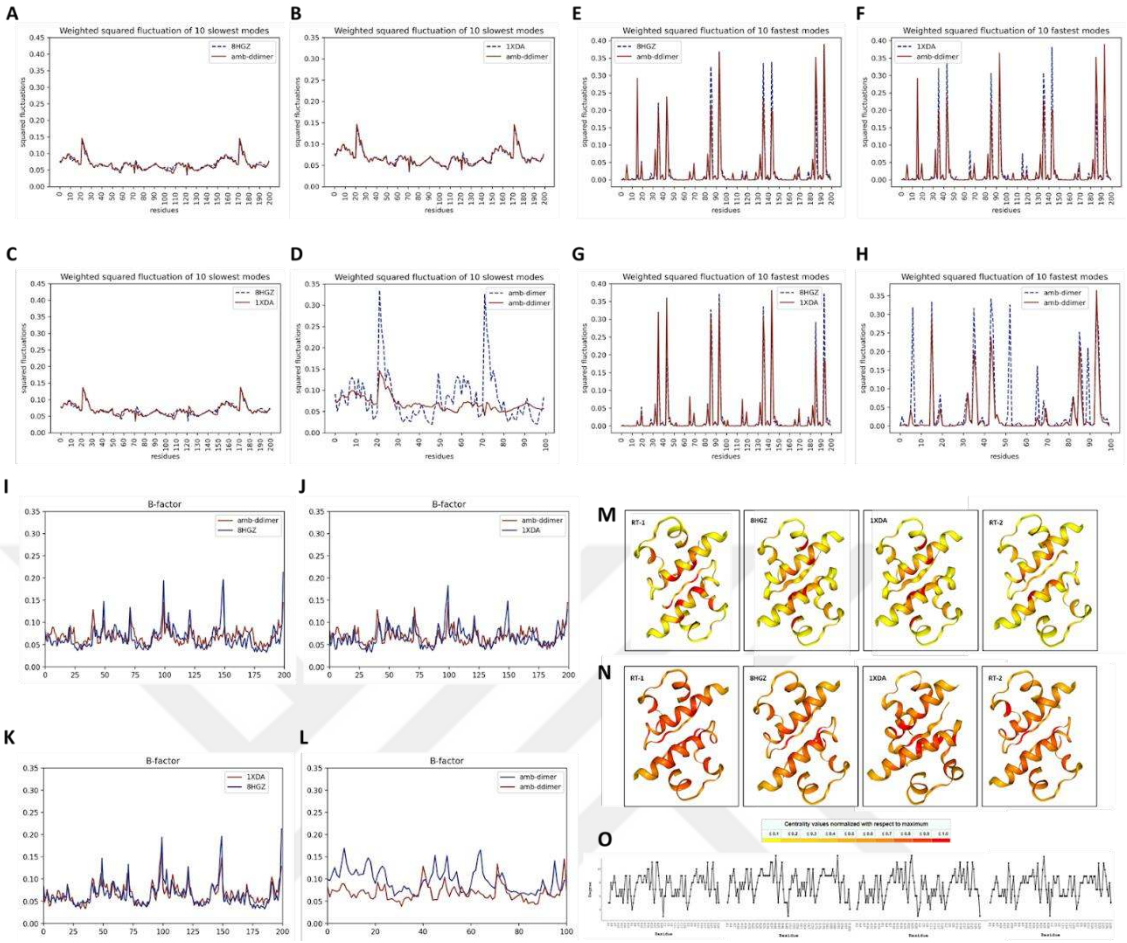


Figure 3.25 Providing an in-depth comparison of structures, 1XDA, 8HGZ, ambient 'dimer of dimer,' and ambient dimer, through analyses using GNM and NAPS.

(A-D) Illustrate the weighted squared fluctuation of the ten slowest modes from GNM, comparing ambient 'dimer of dimer' vs. 1XDA, ambient 'dimer of dimer' vs. 8HGZ, 1XDA vs. 8HGZ, and ambient 'dimer of dimer' vs. dimer forms, respectively. (E-H) Present the weighted squared fluctuation of the ten fastest modes from GNM for the same comparisons. (I-L) Display B-factor comparisons between theoretical and experimental analyses from GNM across ambient 'dimer of dimer,' 8HGZ, 1XDA, and ambient dimer forms, respectively. (M-O) Showcase comparisons of node and betweenness centrality from NAPS within ambient 'dimer of dimer,' 8HGZ, 1XDA, and ambient dimer forms, respectively.

3.2.8. Concluding remarks

The currently accepted model for insulin oligomerization involves a sequential progression through monomer-dimer-hexamer transitions [Derewenda et al., 2989]. However, it is crucial to recognize the presence of an additional di-hexameric intermediate as a vital element in this intricate and dynamic process [Whittingham et al.,

1997]. In both scenarios, the formation of di-hexameric detemir is intricately coordinated with specific Zn^{2+} ions and the presence or absence of myristic acid moieties [Whittingham et al., 1997]. These proposed mechanisms were validated through extensive macroscopic studies, establishing solid correlations between changes in macroscopic properties and the overall dynamics of oligomerization. These associations are further supported by fitting empirical data to models that, while not directly revealing di-hexameric dynamics, encompass the coexistence of dimeric and "dimer of dimer" configurations in the detemir hexamerization pathway [Whittingham et al., 1997]. We have determined that the dynamics inherent in detemir hexamers and dihexamers exhibit complexity beyond the current scope of prevailing models. Our focus has been on oligomeric structures positioned between monomers and dihexamers under ambient conditions. We have outlined the intricate transitions associated with monomeric and dimeric assembly and disassembly using GNM computational analysis. To date, neither experimental nor computational inquiries have unveiled the equilibrium dynamics controlling receptor interactions throughout the entirety of the pathway, spanning the phase transitions of multistep di-hexameric and monomeric dissociations. Our findings, supported by computational insights, reveal that oligomerization exerts a regulatory influence over detemir receptor binding sites across the spectrum of dihexameric-to-monomeric dynamics. Consequently, these dynamic interactions within the intermolecular hierarchy can introduce novel aspects of domain partitioning.

The analysis of the ten fastest and ten slowest modes from extensive GNM and molecular docking revealed dynamic similarities between the dimer and "dimer of dimer" structures of detemir determined at ambient temperature in *Turkish DeLight*. Consequently, certain residues (N24, Q25, H26, L27, L32, A35, L36, R43, and G44) located on the dimer-forming surface of Chain Bs and others (C6, Q5, Q15, N18, and C20) in Chain As can be considered as a second functional motif of detemir, playing a potentially crucial role in the opening process. Previous studies have established that the hydrophobic core of insulin consists of G8, L11, V12, L15, F24, and Y26, F24 and Y26 in chain B, while contributing significantly to the hydrophobic core, have distinct functions during the opening process [Pocker and Biswas et al., 1981; Papaioannou et al., 2015] Specifically, F24 is fully buried within the protein, forming hydrophobic solid interactions with the L15 side chain and not coming into contact with water [Papaioannou et al., 2015]. In contrast, Y26's side chain is exposed to water, displaying increased flexibility due to interactions with adjacent water molecules [Papaioannou et

al., 2015]. Moreover, T27, located adjacent to Y26, does not engage in stabilizing interactions with other residues [Papaioannou et al., 2015]. Consequently, the mobility of Y26 is less restricted, making it easier to disrupt hydrophobic interactions with other core residues and initiate the opening of the B-Chain-C-Terminal (BC-CT) [Papaioannou et al., 2015]. Therefore, our residues identified as kinetically significant (N24, Q25, H26, L27, L32, A35, L36, R43, and G44) are associated with the primary binding-site engagement of the B chain. This conclusion is primarily derived from our analysis of the zinc-bound dihexamer structure, which differs from the monomer structure typically associated with the opening process. Amongst the potentially significant residues, R43 stands out as a surface-exposed, positively charged amino acid positioned strategically, known for facilitating crucial electrostatic interactions that impact the stability of the insulin-insulin receptor complex during the insulin opening process [Jeevanandam et al., 2023]. Previous research has established the pivotal role of R-G-F-F-Y residues in governing the dynamic interplay with receptors [Taketomi and Iwatsuka, 1979]. Remarkably, these residues exhibit insulin-like attributes [Fujino et al., 1977]. In our analyses, we have observed that the ^{XX}L-A-R-G^{YY} residues emerge as kinetically hot elements within the dimeric and 'dimer of dimer' configurations, implying that the ^{XX}R-G-F-F-Y^{YY} sequence constitutes the secondary functional motif of detemir. This is supported by molecular interactions, including hydrogen bonds and salt bridges identified through MD simulation (Figure 3.20). These interactions suggest that all receptor-interacted and kinetically hot residues engage via hydrogen bonds, modulating the wide-opening-closing process of insulin. Specifically, Q15 and N18 have been identified within the dimeric assembly as kinetically significant residues, while within the 'dimer of dimer' configuration, C6 and Q15 exhibit kinetic prominence. This observation suggests the plausible significance of C6 in regulating the opening-closing dynamics of the inactive dihexameric state of detemir. Interestingly, C6 in Chain B, C20 in Chain A, and C28 in Chain B show high interactions with R43 and receptor-interacted residues in both forms of detemir. As observed in the 'dimer of dimer' form, Cys residues interacted 12 times with R43 and receptor-interacted residues, displaying a correlation of at least 50%, while in the dimeric form, Cys residues interacted 8 times with receptor-interacted residues, showing a correlation of at most 25% (Figure 3.19). It is well-documented in the literature that cysteine sulfhydryl (Cys-SH) residues in proteins exhibit pronounced hydrophobic characteristics [Guo and Feng et al., 2001]. Our investigation reveals an enhanced correlation between Cys residues,

R43, and insulin-interacted residues, particularly within the 'dimer of dimer' conformation compared to the dimeric state. This observation suggests a tight interrelation between oligomerization and the hydrophobic nature of these residues, which is crucial for maintaining the quiescent state of the dihexameric assembly. In monomeric insulin-receptor interaction, the initial insulin molecule demonstrates significantly superior binding affinity, approximately 100-fold greater than subsequent insulin acute events [Kristensen et al., 1995; Huang et al., 2004; Yunn et al., 2023; De Meyts et al., 2016]. This indicates a substantial decrease in affinity during these later interactions. Importantly, our analysis identifies a less positive correlation in the intramolecular cooperativity of monomeric insulin compared to oligomeric insulin (Figure 3.15 and Figure 3.16). Incorporating engineering strategies to enhance positive intramolecular correlations into insulin design can lead to insulin variants with elevated efficacy in engaging with their receptors. Such innovative insulin designs can potentially enable insulin use at economically feasible and globally sustainable concentrations.

Together, oligomerization dynamically modulates cooperativity at the insulin monomers' receptor interaction sites and the newly identified subunits. As oligomerization increases, there is a concurrent elevation in cooperation and collectivity. This collaborative behavior is facilitated through specific regions of the insulin receptor and the recently defined subunits (C6, Q5, Q15, N18, and C20 in chain A; N24, Q25, H26, L27, L32, A35, L36, R43, and G44 in chain B). The findings from both GNM and MD simulations consistently reinforce each other. Molecular docking analysis further reveals that the newly identified subunits are closely associated with the insulin receptor and its interaction domains. This emphasizes the proposition that these specific positions, intimately involved with residues crucial for insulin receptor binding, collectively constitute a secondary molecular domain of paramount significance in receptor binding (Figure 3.24). As a result, it is essential to experimentally examine the identified residues to uncover potential associations with a high-affinity site on the insulin molecule.

3.3. A pilot study of glargine structures at ESRF

Unconventional Ca^{2+} Binding Sites and the role of Ca^{2+} for oligomerization

Without divalent metal cations, the protein hormone insulin assumes a dynamic equilibrium within a solution comprising monomeric, dimeric, tetrameric, hexameric, and higher-order oligomer states The hexameric form, characterized by two high-affinity Zn^{2+} binding sites, undergoes a pronounced shift in distribution towards the hexameric state in the presence of Zn^{2+} [Hassiepen et al., 1999; Hill et al., 1991]. The versatility of various divalent metal cations to substitute for Zn^{2+} at these binding sites is noteworthy. As identified by Howell *et al.* (1975), insulin storage vesicles are acknowledged for their heightened concentrations of Ca^{2+} and Zn^{2+} ions, suggesting the *in vivo* uptake of Ca^{2+} by rhombohedral zinc insulin crystals [Howell *et al*., 1978]. From extant *in vivo* observations, inferring the significance of interactions involving Ca^{2+} with zinc proinsulin and zinc insulin hexamers is reasonable. Furthermore, empirical evidence posits a pivotal role for insulin in its crystalline state within mature storage vesicles. Research focusing on the stoichiometries of divalent metal cation binding in crystalline rhombohedral insulin hexamers provided the initial insight into a third high-affinity metal cation binding site within the insulin hexamer [Schlichtkrull, 1956]. In determining the structure of $(In)_6(Zn^{2+})_2$, isomorphous replacement studies involving Cd^{2+} and Pb^{2+}cations displayed the presence of a metal binding site near the center of the hexamer [Adams et al., 1967; Blundell et al., 1972]. Through X-ray diffraction analysis, it was followed that when crystals of $(In)_6(Zn^{2+})_2$ were absorbed in solutions with elevated Zn^{2+} concentrations, the crystals bound Zn^{2+} at various sites, including the EB13 carboxylates situated at the center of the hexamer [Hill et al., 1991].

The comprehensive ^{113}Cd FT NMR investigations proved the calcium-binding characteristics intrinsic in the two-zinc insulin hexamer [Sudmeier et al., 1981; Palmieri et al., 1988]. After this seminal revelation, rigorous physical-biochemical studies have been undertaken to provide a slight characterization of the calcium-binding site within $(In)_6(Zn^{2+})_2$. The findings prove that Ca^{2+} binds at the hexamer's central locus, forming a cage by the six Glu(B13) carboxylates. This binding event occurs with a stoichiometry of one Ca^{2+} cation per hexamer and a dissociation constant of 80 μM [Sudmeier et al., 1981; Storm and Dunn, 1985; Alameda et al., 1985; Kaarsholm and Dum, 1987; Palmieri et al., 1988; Coffman and Dunn, 1988]. The competitive binding of Cd^{2+} and Ca^{2+} cations at the B13 carboxylates, sharing identical stoichiometry, suggests

analogous coordination geometries for these two cations. This proposal aligns seamlessly with their closely matched ionic radii of 0.97 and 0.99 Å, supporting the harmony observed in the coordination preferences of Cd^{2+} and Ca^{2+} in both small-molecule chelators and other calcium-binding proteins [Hill et al.,1991; Cotton and Wilkinson et al., 1999; Norman, 2012]. The 3-fold symmetry axis of the hexamer crosses the center of a 10-Å diameter cavity constituted by the EB13 carboxylates [Blundell et al., 1972; Baker et al., 1988]. This cavity is strategically positioned midway between the two HB10 zinc sites. Within the crystal structure of the hexamer at pH 6.2, the six EB13 carboxylates adopt an arrangement characterized by hydrogen-bonded pairing. It is hypothesized that half of the EB13 side chains are protonated into the carboxylic acid form [Baker et al., 1988]. In the sequence of zinc-binding insulins, a consistent presence of a carboxylic acid residue (E or D) at position B13 is followed, except for hagfish insulin, where the B13 residue is N [Halldén et al., 1986; Peterson and Steiner et al., 1975; Cutfield et al., 1979]. This distinctive feature induces the insulin atypical, and its deviation from the typical sequence is associated with a non-conformance to hexameric formation [Peterson and Steiner et al., 1975; Cutfield et al., 1979].

Although the stoichiometry of one Ca2+ cation per hexamer suggests a singular Ca2+ locus within the cavity [Storm and Dunn, 1985], an examination of the three-dimensional structure of the hexamer reveals a substantial reorganization requirement for the hexamer to form a single site involving all six B13 carboxylates [Hill et al., 1991]. Moreover, initial isomorphous replacement data for Cd2+- and Pb2+-substituted hexamers each indicate the presence of three closely spaced metal ion loci in the EB13 cavity [Hill et al., 1991; Dodson et al., 1993]. Consequently, we have investigated an in-depth X-ray structure study on two hexameric insulin glargine analogs to further explore the binding of three calcium ions, particularly at the EB13 and HB10 sites.

3.3.1. Preparation of samples, crystallization and crystal harvesting

The Lantus® crystallization process employed the sitting-drop micro-batch vapor diffusion screening technique using 72-well Terasaki crystallization plates, following the outlined procedure in Chapter 2. In summary, a detemir solution was combined with a 1:1 ratio (v/v) of approximately 3500 commercially available sparse matrix crystallization screening conditions at room temperature. Each well, containing

0.83 µl of the protein solution and crystallization condition, was sealed with 16.6 µl of paraffin oil (Cat#ZS.100510.5000, ZAG Kimya, Türkiye) and stored at room temperature until crystals formed. Analysis of the Terasaki plates under a compound light microscope allowed for crystal formation and growth monitoring. Large crystals were observed within 24 hours, with the most favorable crystals obtained in a buffer comprising 20 mM calcium chloride, 100 mM TRIS/HCl pH 8.5, 30% (v/v) 2-methyl-2,4-pentanediol, 8% (w/v) PEG 8000. A notably large batch of 0.5 ml insulin crystals was directly mixed with an equal volume (1:1) of the same buffer. These crystals have been stored at room temperature as a stock until further use.

3.3.2. Sample delivery and data collection

Cryogenic data collection. The crystals underwent cryoprotection with 20% glycerol and were rapidly frozen in liquid nitrogen before conducting synchrotron data collection at 100K. Diffraction experiments were executed at the European Synchrotron Radiation Facility (ESRF) synchrotron (Grenoble, France) on beamline BM07. Glargine crystals exhibited diffraction up to 1.9 Å. The dataset experienced indexing and integration using *XDS* [Kabsch, 2010], and subsequent scaling and merging were performed with *AIMLESS* from the *CCP4* suite [Evans, 2014]. The structure of the hexameric glargine was determined through molecular replacement utilizing *MoRDa* [Vagin and Lebedev, 2015], employing the coordinates of lispro insulin from PDB (6NWV) as the search model (*unpublished*).

Serial data collection. Data collection performed to the protocol outlined by Grieco et al (2023). Briefly, microcrystals of the glargine hexamer, with a concentration of 10^8 crystals/ml, were introduced by dispensing 3.5 µl of a crystal slurry and entrapping it between two 13 µm mylar films on a downsized version of the sheet-on-sheet sandwich chip (SOSchip) with a 5 x 10 mm opening. The mylar-mylar sandwich encapsulating the crystal slurry was secured and airtight using the sample holder. Placed between the two halves of the chip holder, the mylar-sample-mylar sandwich underwent additional dispersion and thinning of the film through capillary action. Subsequently, the two frame halves were securely joined together using a single clamping screw. The SSX data of hexamer glargine were collected at the recently installed ID29 SMX beamline of the ESRF-EBS. This involved utilizing a wavelength of 1.072 Å and a pulse duration of 90 µs, accompanied by a repetition rate of 231.25 Hz. The X-ray beam

flux directed towards the sample was approximately 8 x 10^{14} ph/s and was focused on dimensions of 4 x 2 (H x V) µm (FWHM). Data collection has been facilitated by applying the recently developed MD3upSSX diffractometer and a JUNGFRAU 4M detector, with an integration time of 95 µs to comprehensively record the X-ray pulse. The configured sample-to-detector distance was 150 mm. Automatic corrections for pedestal and geometrical reconstruction have been implemented employing the LImA2 data acquisition library generated at the ESRF (https://limagroup.gitlab-pages.esrf.fr/lima2/).

Data processing was performed remotely using Virtual Infrastructure for Scientific Analysis (VISA, https://visa.esrf.fr). Initial hit identification utilized a GPU version of *NanoPeakCell*, employing the Peakfinder8 algorithm, with parameters set at an SNR of 5.0 and a threshold of 1200. Subsequent integration of Bragg reflections was executed through the *CrystFEL* software package [White et al. 2012; White 2019] version 0.10.1. Indexing attempts were made using *CrystFEL*'s indexamajig, employing the *XGANDALF* [Gevorkov et al. 2019], *MOSFLM* [Powell 1999], and ASDF algorithms. The detector geometry file underwent further optimization through the *geoptimizer* tool [Yefanov et al. 2015]. Phasing and refinement of MTZ files were generated using the *CTRUNCATE* program [Evans 2011] within the *CCP4* software package [Winn et al. 2011]. A subset comprising 5% of reflections was incorporated into the generated R_{free} set.

3.3.3. Structure determination

The SSX structure was determined at ambient temperature in space group $P12_11$ using the automated molecular replacement program *PHASER* integrated into the *PHENIX* software package [McCoy et al., 2007]. An initial molecular replacement search model was derived from a previously determined cryogenic glargine structure. The structural coordinates of the cryogenic glargine served as the starting point for the initial rigid-body refinement within *PHENIX*. Subsequent refinement steps included simulated-annealing refinement of individual coordinates and Translation/Libration/Screw (TLS) parameters. Additionally, integrated into *PHENIX*, a composite omit map refinement was performed to identify potential positions of altered side chains and water molecules. The final model experienced thorough examination and rebuilding in *COOT* [Emsley and Cowtan, 2004] version 0.8.9.2, focusing on

retaining positions showing strong difference density. Figures depicting the X-ray crystal structures were generated using *PyMOL* [Delano, 2002] version 2.3 and *COOT*.

3.3.4. Molecular dynamics simulation

MD simulation was executed using the *Amber* software, employing the all-atom additive *CHARMM36m* force field for hexamer glargine [Case et al., 2005]. The simulation was completed under periodic boundary conditions [Case et al., 2005] with a time step of 2 fs and a trajectory interval of 10 ps. Temperature control was held at 298.15 K using a Langevin thermostat, and pressure was kept at 1 bar with a *Berendsen* barostat [Miyamoto and Kollman, 1992; Ryckaert et al., 1977]. The shake method was applied, and short-range electrostatic and van der Waals forces were cut off at 12 Å. The simulation began with energy minimization, initially with restraint to the solute and no pressure control, followed by energy minimization without any restraint to eliminate unfavorable contacts. System temperature was gradually increased to 298.15 K over 100 ps, applying a 10.0 kcal/(mol $Å^2$) constraint to the solution through a harmonic restraint. The Langevin thermostat was employed for system heating, and an anisotropic Berendsen weak-coupling barostat was used to equilibrate pressure. After heating, a 5 ns equilibration phase was conducted with a more considerable skinny value to stabilize system dimensions and density. Ultimately, temperature control was accomplished using the Langevin thermostat in conjunction with a constant pressure periodic boundary, maintaining an average pressure of 1 atm, and a 500 ns MD simulation was completed.

3.3.5. Observations

We determined the structures of hexameric glargine at cryogenic and serial ambient temperatures, achieving resolutions of 1.95 Å and 2.3 Å, respectively. The diffraction data were collected directly at the ESRF (Grenoble, France). Notably, both hexamers revealed, for the first time, the existence of three closely spaced calcium cation binding sites within the EB13 cavity (Figure 3.26, Table 3.3).

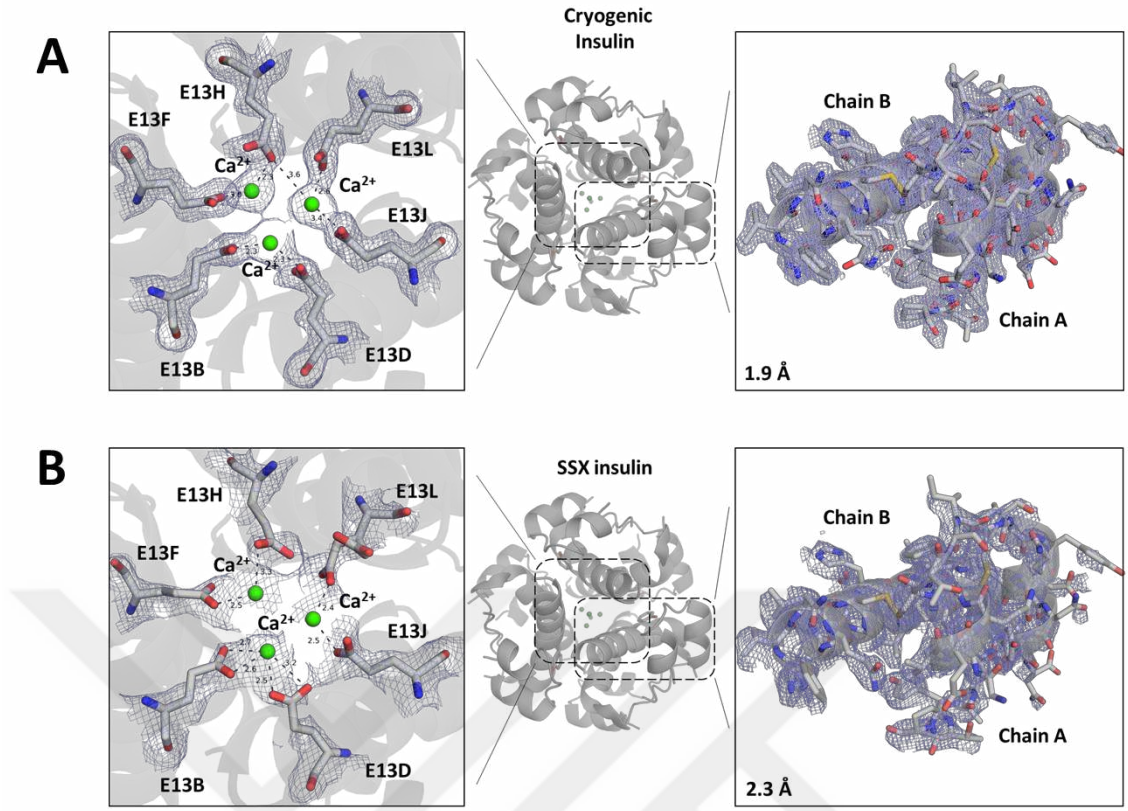


Figure 3.26 Demonstration of hexamer glargine structures determined at cryogenic and ambient temperatures.

(A) Depicting the cryogenic hexamer in monomer form (right) and a detailed view of all the E13 residues binding to calcium (left). (B) Presenting the serial-ambient hexamer in monomer form (right) and a detailed view of all the E13 residues binding to calcium (left). The *2Fo-Fc* simulated annealing-omit map at a 1 sigma level is shaded in slate.

Table 3.3 Data collection and refinement statistics of cryogenic and SSX glargine analogs.

Data collection	Glargine (Cryogenic)	Glargine (SSX)
Instrument	ESRF/BM07	ESRF/ID29
Resolution range	19.81 - 1.95 (2.02 - 1.95)	19.28 - 2.3 (2.382 - 2.3)
Space group	P 1 21 1	P 1 21 1
Unit cell	47.45 61.856 55.812 90 112.47 90	47.73 62.22 56.39 90 111.87 90
Multiplicity	2.0 (2.0)	1.9 (2.0)
Completeness (%)	96.04 (92.11)	93.40 (84.43)
Mean I/sigma(I)	11.78 (3.35)	5.71 (3.40)
Wilson B-factor	22.14	34.35
R-merge	0.04891 (0.274)	0.2238 (0.4219)
R-meas	0.06917 (0.3875)	0.3165 (0.5966)
R-pim	0.04891 (0.274)	0.2238 (0.4219)
CC1/2	0.996 (0.803)	0.825 (0.559)
CC*	0.999 (0.944)	0.951 (0.847)
R-work	0.1957 (0.2490)	0.2863 (0.3368)
R-free	0.2424 (0.3102)	0.3410 (0.4295)
No. of reflections	21030 (1995)	12841 (1160)
Unique reflection	1503 (148)	1113 (99)
Protein residues	297	297
RMS (bonds)	0.019	0.004
RMS (angles)	1.78	0.70
Ramachandran favored (%)	98.89	97.05
Ramachandran allowed (%)	1.11	2.21
Ramachandran outliers (%)	0.00	0.74
Average B-factor	29.18	42.80
macromolecules	28.63	42.73
ligands	23.19	39.14
solvent	35.76	46.09

We conducted a root-mean-square fluctuation (RMSF) analysis to observe how residues change during 5 sets of 500 ns MD simulations for each structure (Figure 3.27).

Residues Y14, F22, K50, Y64, F72, Y111, F121-Q124, E141, K149, Y163, F171-Q174, Y212, F220, Y261, F269, and K297 exhibited significant fluctuations at distances of at least 2.5 Å. Notably, residues E4, Y14, F22-H26, P49-I51, E54-Q55, 64Y, F72-V73, Q75-H76, K100, E104, I110, S112-Q115, E117-N118, F122-V123, H126, E142, P149-K150, E154-Q155, I160, Y164, N171-F172, E192, F196, T198, Y212, F220, E251, Y261, F269-N271, E289, F293, and T295-K297 also displayed significant fluctuations at distances of at least 2.5 Å. Notably, the first four residues of each monomer in the second chain showed the highest fluctuations compared to the others (Figure 3.27).

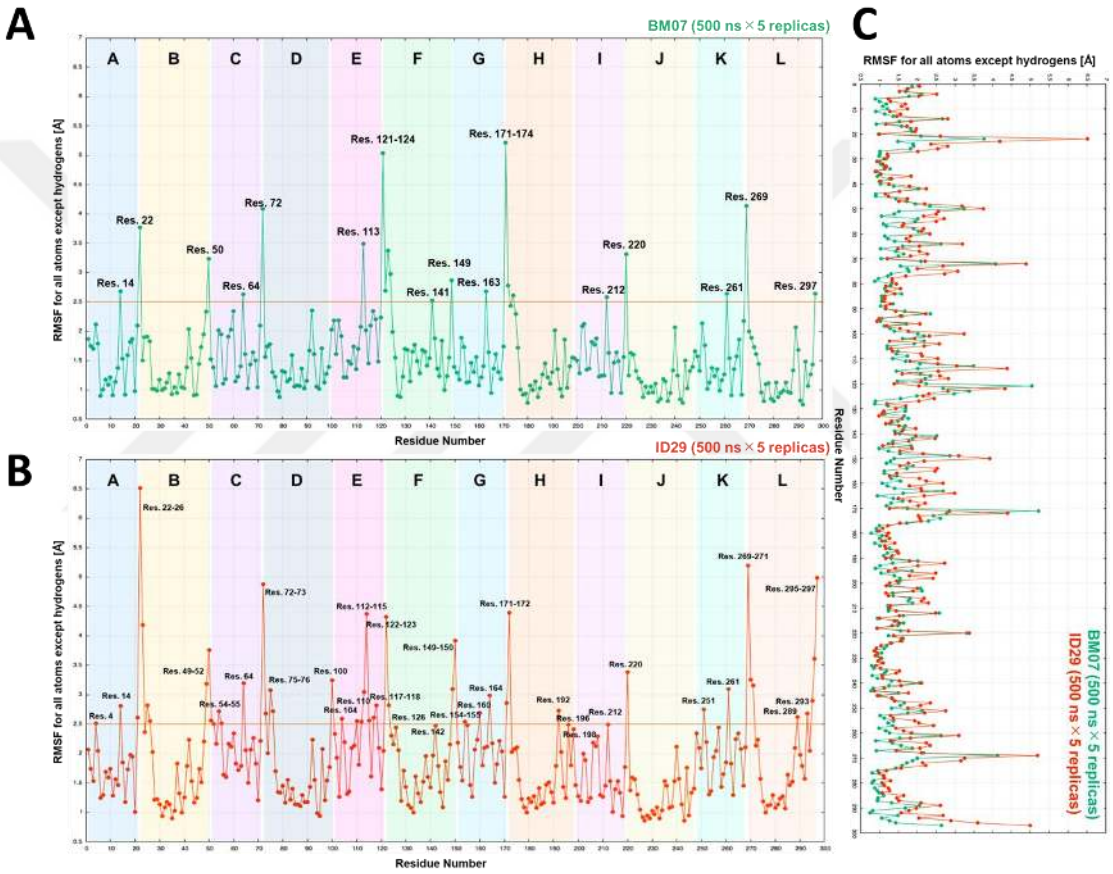


Figure 3.27 Calculated average-RMSF of cryogenic and SSX structures using 5 replicas for each system in 500 ns x 5 replicas.

(A) RMSF analysis of cryogenic glargine structure. (B) RMSF analysis of SSX glargine structure shows higher fluctuations than cryogenic structure. (C) RMSF of cryogenic and SSX structure for all atmos except hydrogens.

We conducted RMSD analysis through MD simulations to track the changes in EB13 for all atoms over time (Figure 3.28). The X-axis represents the time series of RMSD, calculated from the initial structure of the MD simulations. The analysis involved simultaneously computing the RMSD for all trajectories within each structure,

incorporating data from five independent 500 ns simulations for all E13 residues within each monomer.

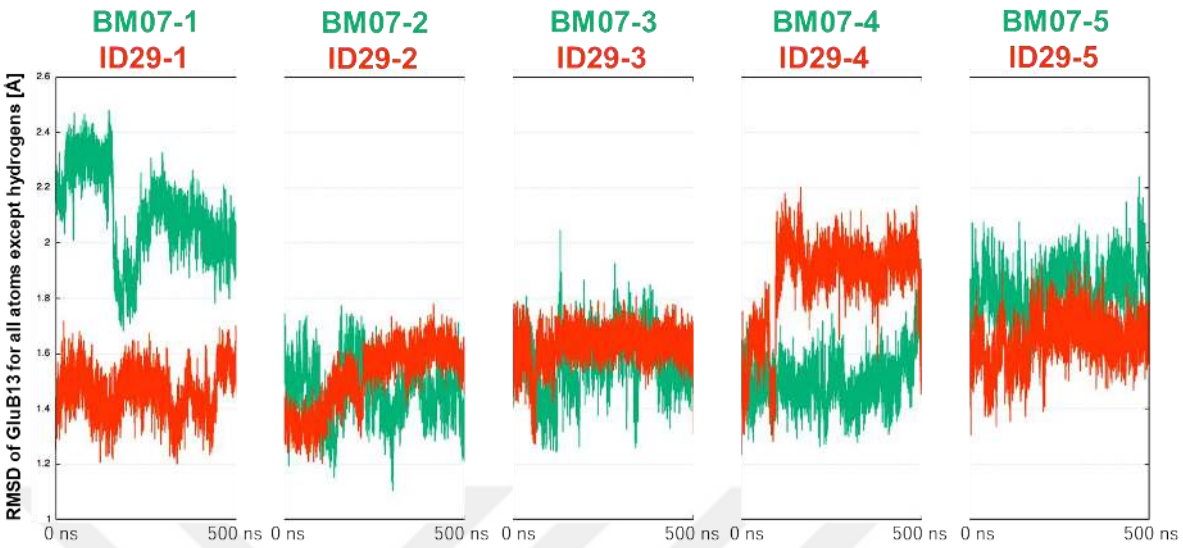


Figure 3.28 Time series of RMSD for E13 residues in cryogenic glargine and SSX glargine structures in 500 ns x 5 replicas.

RMSD was calculated for all atoms except hydrogen atoms in E13 residues, using the cryogenic glargine and SSX glargine structures as a reference structure.

We conducted a 500-nanosecond all-atom MD simulation utilizing both cryogenic and SSX hexameric glargine structures to analyze the interaction between E13 residues in each monomer and three calcium ions in detail. To investigate local conformational changes of E13 residues, the RMSD for all atoms, excluding hydrogen atoms, from the initial structure as a reference was calculated. The conformations of E13 residues exhibited significant changes during the MD simulation. In the BM07-1 system, the conformation of the E13 residue changed notably (Figure 3.29), while in the ID29-4 system, the conformation of the E13 residue displayed dynamic changes around 100 ns (Figure 3.30). Preceding the ID29-4 conformational change, the structural fluctuations of the E13 residue in the BM07-1 system increased.

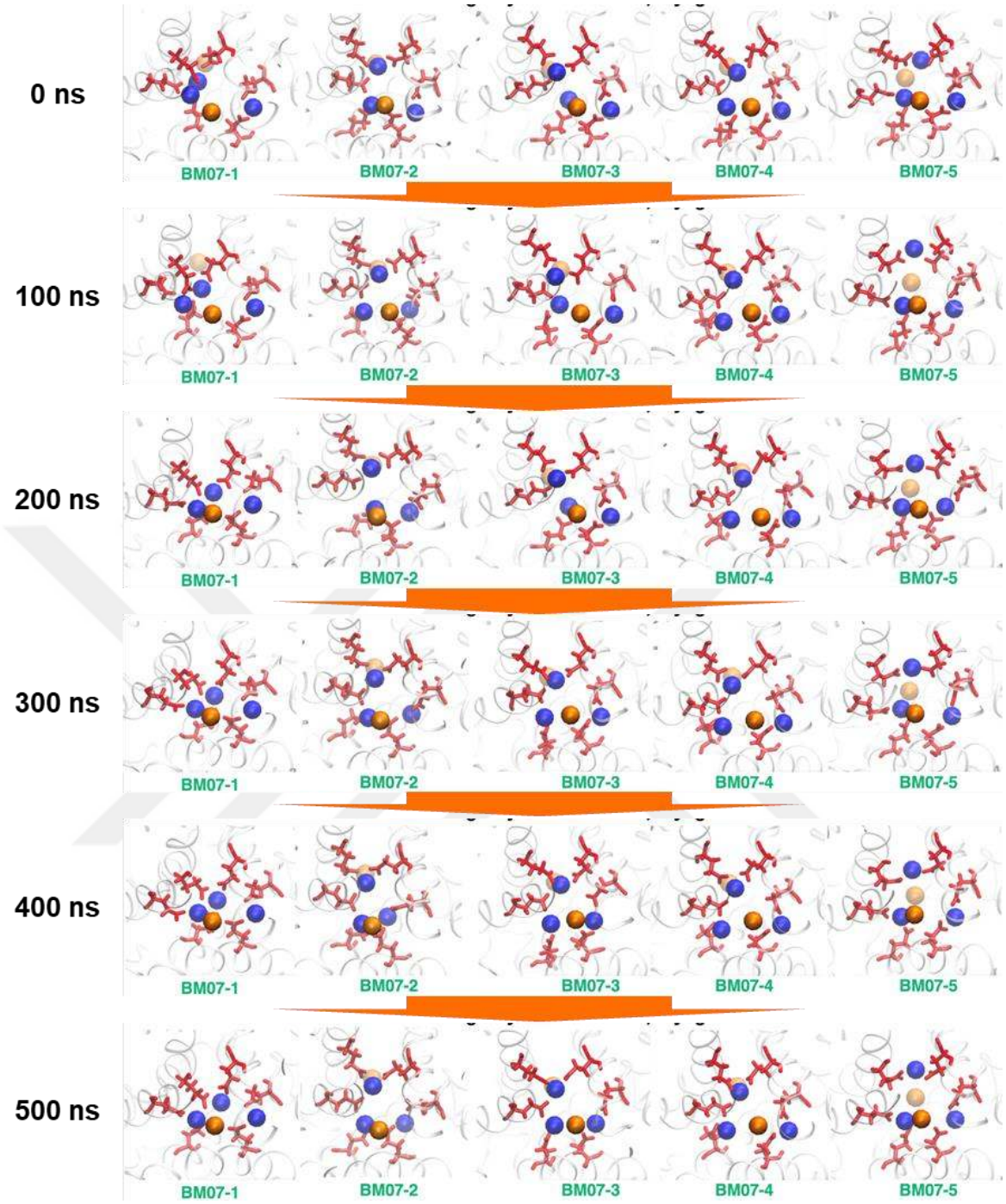


Figure 3.29 Representation of the time series of RMSD for E13 residues in cryogenic glargine across 5 replicas of 500 ns each.

Replicas 1 and 5 exhibit higher fluctuations compared to the other replicas.

In the BM07-1 replica, there was a substantial change in the conformation of E13. The conformation of the E13 residue in this replica differed from those observed in the other MD simulations, as illustrated in Figure 3.29.

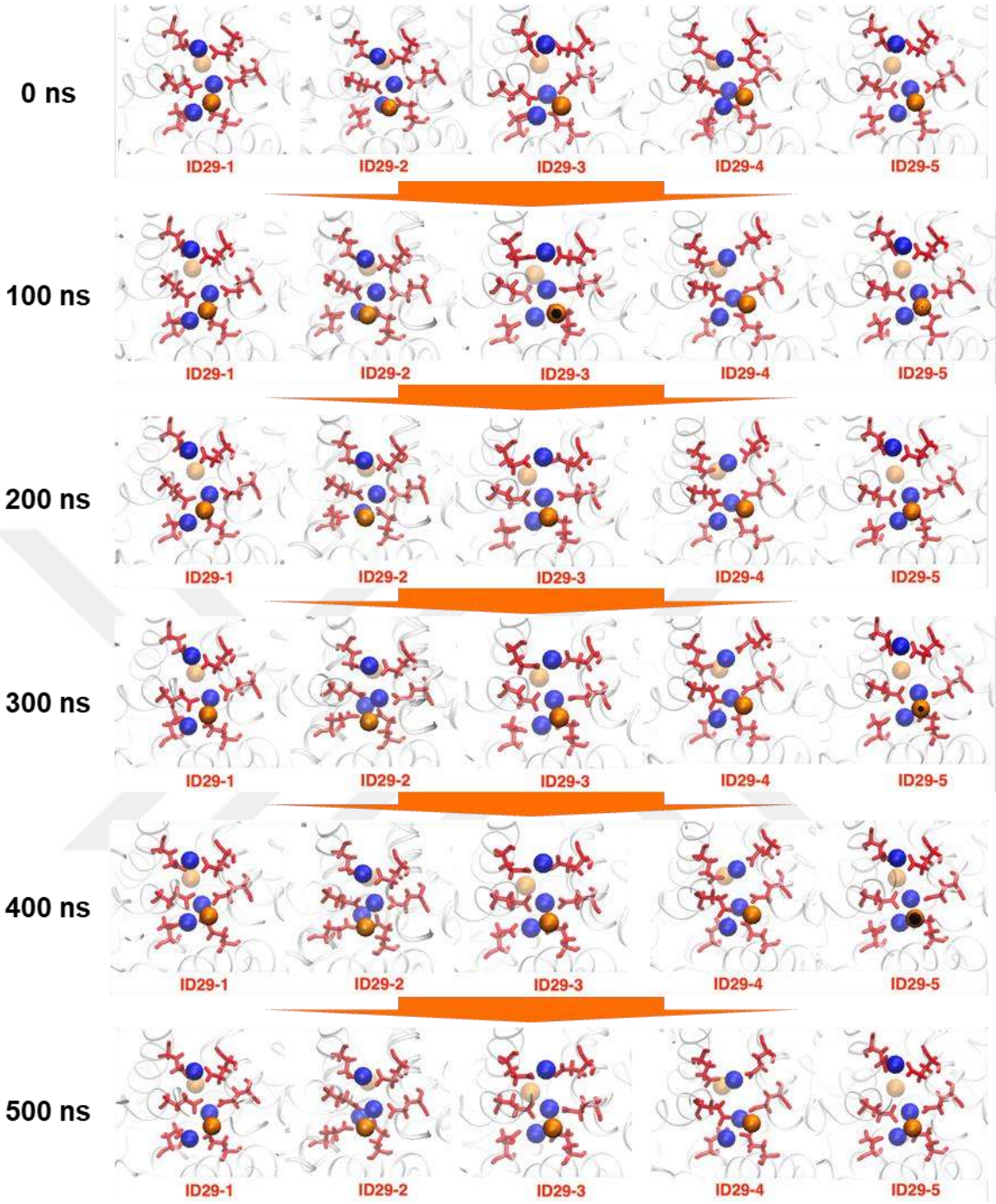


Figure 3.30 Representation of the time series of RMSD for E13 residues in SSX glargine across 5 replicas of 500 ns each.

Replicas 4 exhibit higher fluctuations compared to the other replicas.

In the ID29-4 replica, the conformation of the E13 residue exhibited dynamic changes, particularly around 100 ns. The timeframe of 0-500 ns captures the conformational alterations in the ID29-4, presenting the MD results from 0 ns to 200 ns. A noteworthy conformational change in the E13 residue was observed around the 100 ns interval (Figure 3.30).

3.3.6. Concluding remarks

The insulin hexamer's HB10 sites exhibit significantly higher Zn^{2+} binding affinity than Ca^{2+}. Nevertheless, substituting Zn^{2+} with Ca^{2+} demonstrates ease reminiscent of a small-molecule chelator, distinguishing it from many other zinc metalloproteins [Schlichtkrull, 1956; Dunn, 1975; Storm and Dunn, 1985]. In contrast, the EB13 sites display distinct specificity, with weak Zn^{2+} binding and strong Ca^{2+} binding [Blundell et al., 1972; Storm & Dunn, 1985; Dunn et al., 1987]. This specificity is influenced by the chemical nature of the ligands (carboxylate oxygens) and site constraints that limit the number of protein ligands to a pair of carboxylates [Hill et al., 1991]. Since the divalent metal ions, especially Ca^{2+}, are all spherical, there is no crystal field interaction favoring a specific ligand geometry. Thus, the coordination geometries at the H10 and E13 sites are shaped by ligand-ligand interactions, including both steric and electronic considerations, as well as the structural constraints imposed by the protein's three-dimensional arrangement [Hill et al., 1991]. According to X-ray diffraction studies, when Ca^{2+} binds within the EB13 cavity, the B13 side chains are disordered between two conformations. The intradimer hydrogen bonds between two sets of carboxylate-carboxylic acid pairs are disrupted, leading to ionizing one of the side chains in the carboxylic acid form. Two side chains are then rearranged to create an interdimer coordination site for the metal cation. As this new site extends across the dimer-dimer interface, the coordination of the metal cation introduces a novel interaction, likely facilitating assembly and contributing to the stabilization of the insulin hexamer [Hill et al., 1991]. Assorted NMR studies of Cd^{2+} and Ca^{2+} binding reveal that introducing these metal ions to the EB13 site results in perturbations in the NMR signals of distant side-chain residues [Palmieri et al., 1988]. Notably, aromatic side chains (F and Y) contributing to hydrophobic clusters along the monomer-monomer and dimer-dimer interfaces display significant outcomes. A Pb^{2+} site, identified near the carboxylate of EA17, offers a potential explanation for some of the perturbed chemical shifts marked in the NMR spectrum. The proximity of the metal ion to the aromatic rings of FB1 and YA14 could induce substantial perturbation even at low partial occupancy by Cd^{2+} or Ca^{2+} in solution. Palmieri et al. (1988) imply that these spectral perturbations primarily originated from motions of the EB13 site along

the subunit interfaces, inducing slight changes in the comparable positions of the aromatic rings. If this interpretation holds, these conformational changes may be too fine to be detected at the resolution of the structures reported in this study. The binding sites within Ca^{2+}-binding proteins, whose three-dimensional structures have been revealed, can be broadly classified into two or three primary classes. One prominent group features two closely related types distinguished by the helix-loop-helix motif, the "EF-hand." [Kretsinger, 1972] This motif is prevalent in proteins belonging to the calmodulin family [Norman et al., 2012; Sundaralingam et al., 1985]. Another class encompasses loop motifs identified in various Ca^{2+}-binding proteins, including trypsin, concanavalin A, phospholipase A2, and site 3 of thermolysin [Szebenyi and Moffat, 1986]. These distinctive structural motifs play a crucial role in easing the binding of calcium ions and represent recurring themes in the architecture of Ca^{2+}-responsive proteins. The B13 Ca^{2+} binding site within insulin exhibits a distinct feature, with a single metal ion dynamically transitioning between three symmetry-related sites and forming covalent coordination with separate protein molecules. This characteristic differentiates it from Ca^{2+} binding sites observed in other known protein structures, reflecting a unique biological function [Hill et al., 1991]. The exclusive protein ligands for Cd^{2+} at the center of the insulin hexamer are two B13 carboxylates originating from different dimers. Consequently, the Ca^{2+} binding site emerges due to dimer aggregation, forming the tetramer and hexamer [Hill et al., 1991]. Its physiological role involves stabilizing the hexamer during storage and facilitating hexamer formation from proinsulin. In contrast, the metal-coordinating groups in other Ca^{2+} binding proteins typically stem from the same molecule and are usually closely positioned in sequence [Hill et al., 1991].

This study presents the hexameric structure of glargine determined at two distinct temperatures. One configuration was determined at cryogenic temperatures with a resolution of 1.95 Å, while the other was determined at ambient temperature using the SSX technique with a resolution of 2.3 Å. Noteworthy that the binding of three calcium ions to the B13 cavity of glargine in hexameric form marks the first instance of such an observation. Despite the literature suggesting a stoichiometry of one Ca^{2+} cation per hexamer, implying a solitary Ca^{2+} locus within the cavity, scrutinizing the three-dimensional hexamer structures highlights a substantial need for reorganization to create a unified region encompassing all six B13 carboxylates [Hill et al., 1991]. Additionally, each set of preliminary isomorphous replacement data for Cd^{2+}- and Pb^{2+}-substituted

hexamers suggests the existence of three closely situated metal ion loci in the EB13 cavity [Dodson et al., 1993; Hill et al., 1991]. This study provides conclusive evidence for the presence of three Ca^{2+} cations at proximate locations within the three-dimensional hexameric structure, irrespective of both temperature conditions. Additionally, the results of the MD simulations we conducted indicate that the hexameric structure at ambient temperature is more fluctuating compared to the cryogenic structure. Specifically, simulations of E13 residues at both temperatures exhibit significant deviations in certain replicas.

3.4. A pilot study of glargine structures at SACLA/XFEL
Structures of hexamer insulin reaction states by pH-jumped TR-SFX

The shift towards structure-specific self-assemblies plays a crucial role in the functional pathways of various biomolecules, particularly in proteins [Levy and Teichmann, 2013]. The dynamic formation of these biomolecular entities into templates for receptor binding or storage modules is commonly stabilized by factors such as metal ions, hydrogen bonds, and the entropic advantages linked to resultant structural alterations [Rest et al., 2015]. These transformations are frequently influenced significantly by ambient conditions such as pH, temperature, ionic strength, and the presence of small molecules [Hua et al., 2011]. Since the revelation of insulin's biological function in the 1920s, extensive biophysical and structural characterizations have been conducted on this molecule. Notably, the crystallographic structure of insulin was among the pioneering protein structures to be elucidated. Clearly, numerical studies on insulin crystallization have predominantly focused on two key areas: 1) exploring the formation and dissolution dynamics of natural zinc-containing insulin crystals, particularly those with a rhombohedral structure [Rahuel-Clermont et al., 1997]; 2) developing insulin drugs with varying release profiles, categorized as rapid, intermediate, and long-acting formulations [Pillai and Panchagnula et al., 2001]. The first avenue aims to enhance our comprehension of the factors governing *in vivo* crystallization and dissolution of insulin, while the second addresses the therapeutic needs of individuals with insulin-related disorders. Insulin has been successfully crystallized in various space groups, with monoclinic, rhombohedral, cubic, and tetragonal crystal forms being the most common [Norrman and Schluckebier, 2007]. The type, size, and morphology of these crystals play a pivotal role in influencing the

rate of insulin release, highlighting the extensive research conducted on insulin crystallization [Norrman and Schluckebier, 2007].

As outlined in Chapter 1, the insulin molecule comprises two chains, A and B, with 21 and 30 residues, respectively. Chain A consists of two helical fragments separated by a brief loop connected to one of the helices by an intra-chain disulfide bond. Two more disulfide bonds connect chain A to the larger chain B. In its biologically active form, insulin exists as a monomer, where chain B incorporates a central helical region flanked by two elongated parts. Upon divalent cations like Zn^{2+}, the monomers aggregate into hexamers, where each central zinc cation coordinates with three histidine residues. In the hexameric state, chain B exhibits two allosteric states, denoted T and R. The R state features two allosteric binding sites: the phenolic binding site and the HisB10 anion site [Ayan and DeMirci, 2023]. The transition from T to R state and the two distinct B chain conformations, T6 and R6, have been elucidated through spectroscopic and crystallographic studies [Baker et al., 1988; Smith et al., 2003]. The T6 conformation, characterized by an extended form of residues 1–8 of chain B, is observed at low chloride concentrations and without phenol derivatives [Smith et al., 2003]. Phenolic derivatives, commonly used as preservatives in insulin pharmaceutical formulations, include phenol, meta-cresol, resorcinol, and methylparaben. The R conformation is adopted in the presence of these derivatives and at high chloride concentrations [Derewenda et al., 1989; Smith et al., 2000; Smitg et al., 1992; Whittingham et al., 1998]. In this state, the first eight residues of the B chain assume a helical conformation, resulting in a continuous helix comprising residues B1 to B19. This shift from an extended to an alpha-helical conformation induces a substantial ~30 Å displacement in the position of the first residue of chain B, PheB1 [Norrman and Schluckebier, 2007]. Like chloride, in the absence of phenolic derivatives and at elevated concentrations, these anions can trigger the R state in three out of six monomers within a hexamer [Whittingham et al., 1995; Ciszak et al., 1994]. The remaining three monomers maintain an extended conformation (T state) in the region encompassing residues B1 to B8. In the incomplete R state of the first three monomers, residues B4 to B8 adopt a helical conformation, while residues B1 to B3 remain in an extended conformation [Norrman and Schluckebier, 2007; Whittingham et al., 1995; Ciszak et al., 1994]. This hexamer configuration is labeled $T_3R^f_3$, with the 'f' signifying a frayed R conformation. The primary alteration during the T-to-R transition involves converting residues B1-B8 from an extended conformation (T state) to a helical

conformation (R state) [Norrman and Schluckebier, 2007; Whittingham et al., 1995; Ciszak et al., 1994]. Spectroscopic evidence from NMR circular dichroism and rapid kinetic studies suggest that the T and R states observed in crystals can be induced in solution by binding lyotropic anions or phenolic compounds [Kim and Shields, 1992; Roy et al., 1989]. For example, M-cresol enhances circular dichroism in the far-UV range, indicating an increase in helix content around 255 nm [Kim and Shields, 1992]. Despite extensive exploration of the T and R states of zinc or cobalt insulin in solution, pH-dependent oligomeric conformational changes must be reported. One potential explanation could be the occurrence of isoelectric precipitation within the neutral pH range [Kim and Shields, 1992].

Among the crystallographical techniques for protein further characterization, the rise of X-ray free-electron lasers (XFELs) marked a revolutionary phase in structural biology by providing an unprecedented ability to observe short-lived transtion states of biomolecules using ultra-short, high-intensity X-ray pulses [Sierra et al., 2019]. Techniques devised for macromolecular investigations at XFELs operate a distinct data collection method known as "diffract-before-destroy" [Shimazu et al., 2019]. Instead of exposing a single and sizeable macromolecular crystal to acquire a complete dataset, ultrafast pulses at XFELs diffract numerous microcrystals successively immediately prior to destruction [Shimazu et al., 2019]. Beyond static crystal structures, sample delivery methods at XFELs have been developed to dissect fleeting intermediates of biomacromolecules. Strategies like mix-and-inject XFEL serial crystallography have been designed to deliver novel routes for studying real-time inter- and intra-oligomers [Sierra et al., 2019]. Extending a new realm of Time-resolved Serial Femtosecond X-Ray crystallography (TR-SFX), the mix-and-inject technique has been operated to grab the conformation alterations of biomacromolecules [Sierra et al., 2019].

The investigation into the behavior of hexameric insulin structure is primarily driven by the necessity to develop modified long-acting insulins for diabetes treatment. Although the metabolic impact of insulin *in vivo* begins with the binding of monomeric insulin to specific receptors on the cell membrane, the protein is stored in pancreatic vesicles in a hexameric form, complexed with zinc cations [Li et al., 2014]. This hexameric structure, supported by phenolic agents (means R form of insulin), is relatively resistant to degradation and fibrillation [Berchtold and Hilgenfeld, 1999]. It constitutes the main component in all insulin pharmaceutical preparations, ensuring a consistent basal level of insulin release in diabetic individuals. Modifying intrinsic

factors, such as mutations, or extrinsic factors, like pH shifts, influencing the self-association and dissociation of insulin or its aggregates, is crucial in developing long-acting insulin formulations [Rosskamp and Park, 1999] . This optimization aims to enhance bioavailability during therapeutic action.

A profound comprehension of hexamer assembly, structural specificity, and stability in solution is essential to enhance therapeutically advanced long-acting insulin formulations. This pilot study in this Chapter 3 utilizes pH jumped time-resolved serial femtosecond crystallography (tr-SFX) to reveal insights into dissociation dynamics of oligomeric insulin structures at different pH stages at near-physiological temperature. This study demonstrates that a pH-jump triggers radical changes in space group and unit cell configurations when chaotropic agents starts to vanish, inducing allosteric transition followed by forming distinct crystal lattice. These variations are closely related to the charge state of insulin and, thus, the pH values involved.

3.4.1. Preparation of samples, crystallization, and crystal harvesting

The crystallization of commercially available glargine, one of the most common clinically-used and commercially known as Lantus®, was conducted using a sitting-drop micro-batch vapor diffusion screening technique under oil, employing 72-well Terasaki crystallization plates. In summary, the commercial glargine solution was mixed with a commercially available sparse matrix crystallization screening condition in a 1:1 ratio (v/v) at room temperature for crystallization, employing approximately 3500 screening conditions [Durdagi et al., 2021]. Each well containing 0.83 µL of protein and crystallization condition was sealed with 16.6 µL of paraffin oil (Cat#ZS.100510.5000, ZAG Kimya, Türkiye) and stored at room temperature until crystal harvesting. Crystal formation and growth were monitored by examining Terasaki plates under a compound light microscope, and tiny crystals were observed within 1-2 hours. The tiny crystal conditions were selected and batch cultivated in a buffer of 15% (v/v) reagent alcohol, 200 mM magnesium chloride, and 100 mM imidazole/HCl at pH 8.7. A substantially large batch of 12.5 mL insulin crystals was directly mixed with an equal volume (12.5 mL) of the same buffer, immediately observing tiny crystals. Final pH of this crystal solution is 8.40. These crystals have been stored at room temperature as a stock until they are used it.

3.4.2. Sample delivery and DAPHNIS setup for data collection

Data collection utilized a high-viscosity cartridge-type (HVC) injector, designed explicitly for serial femtosecond crystallography (SFX) at atmospheric pressure, developed at SPring-8 Angstrom Compact Free-Electron Laser (SACLA) [Shimazu et al. 2019]. The HVC injector is formulated to seamlessly integrate into the diffraction chamber of the Diverse Application Platform for Hard X-ray Diffraction in SACLA (DAPHNIS). The injector efficiently delivers microcrystals embedded in a highly viscous carrier through a capillary nozzle, assisted by a coaxial gas flow and a suction device [Shimazu et al. 2019]. The replaceable cartridge-type sample reservoir simplifies sample reloading or exchange. Positioned within a cooling jacket with temperature-regulated water flow, the reservoir prevents abrupt temperature changes during the data collection.

Preparation of grease-dispersed glargine microcrystals

A 25 ml solution with glargine microcrystals underwent centrifugation for 5 minutes at 1000 rpm, and the volume was then adjusted to a crystal density concentration of 10^8/ml. For pH 8, only 10 ul of crystals mixed with with 90 ul of Super Lube nuclear grade grease (No. 42150, Synco Chemical Co.) on a plastic plate, following a procedure previously reported by Sugahara et al. (2015). Using a spatula, the crystals were transferred directly to a 200 ul sample cartridge. Following the sample loading, the cartridge underwent centrifugation to eliminate bubbles from the mixture, and subsequently, the cartridge was placed into the HVC injector body (Figure 3.31). For samples prepared under alternative pH conditions (pH 7.27; pH: 6.36; and pH: 5.13), a combination of 10 ul glargine and 5 ul certain pH buffer was mixed with 85 ul grease, and for pH: 6.74, pH: 6.19, and pH 5.45 derived-conditions a combination of 10 ul glargine and 2 ul, 3ul, and 4 ul pH 5.13 buffer was mixed with 88 ul, 87 ul, and 86 ul grease, respectively, with the preparation completed 5 minutes prior to the initiation of data collection. For each sample prepared, final slurry pH values were measured directly by employing a micro pH meter electrode measurement.

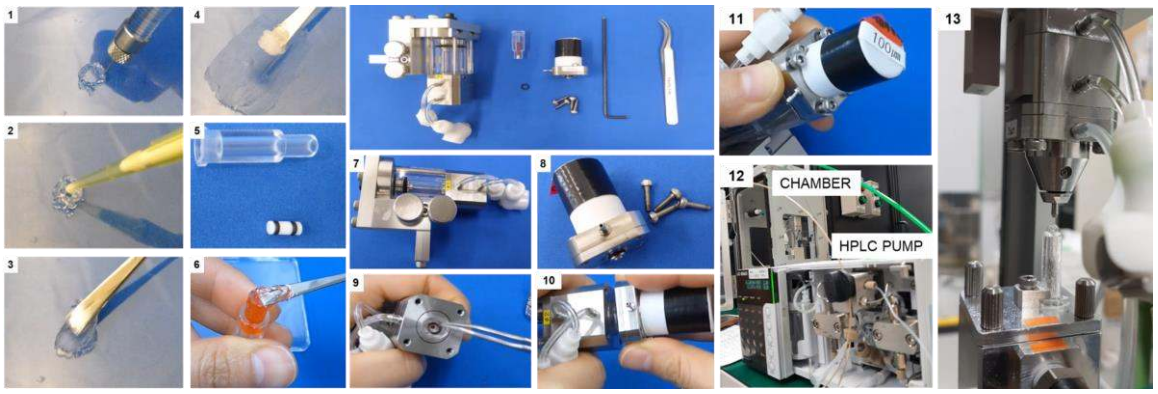


Figure 3.31 Preparing grease-dispersed glargine microcrystals, assembling the HVC injector, and setting up the DAPHNIS platform involved the following steps:

(1) Applying 85 ul of grease onto a slide, (2) Adding 10 ul of glargine + 5 ul of crystal buffer at various pH levels, (3) Thoroughly mixing the grease and samples, (4) Observing a homogeneous mixture, (5) Assembling the 200 ul cartridge, Teflon part, and O-rings, (6) Loading the sample into the cartridge, (7) Configuring the HVC, (8) Attaching the injector nozzle with a lid, (9) Inserting the cartridge and O-ring into the sample cartridge holder, (10) Placing the nozzle with assemblies of the sample cartridge holder, HPLC pump, and chiller connector, (11) Securing the nozzle with four screws, (12) Setting up the DAPHNIS chamber and HPLC pump, (13) Completing the final setup of the HVC injector in DAPHNIS.

pH jumped TR-SFX data collection

Data collection was conducted at BL2/EH3 of SACLA [Ishikawa et al., 2012; Tono et al., 2013] using X-rays with an energy of 10.0 keV, pulse duration less than 10 fs, and a repetition rate of 30 Hz. The XFEL beam was focused to a beam diameter of 1.5 um FWHM using two elliptical mirrors in the Kirkpatrick–Baez geometry. The beam was attenuated to approximately 10^{10} photons per pulse to prevent detector saturation for glargine crystals. The HVC injector was integrated into the helium-gas-filled diffraction chamber of the DAPHNIS instrument setup (Tono et al., 2015). While the sample was maintained at 293 K using the injector's cooling jacket, the chamber temperature was relatively higher (299–300 K). Positioned just below the injector, a suction device aspirated helium gas, and the aspirated gas was redirected from the vacuum pump's exhaust line back to the chamber. Diffraction patterns were captured using a multiport CCD detector (MPCDD) with eight sensor modules (Kameshima et al., 2014) at a sample-to-detector distance of 50 mm. Glargine crystals embedded in grease (~5 um in size) were ejected at a flow rate of 0.36 µl/min from a nozzle with an inner diameter of 200 µm. The crystals samples adjusted totreated with the specific pH values were subjected to data collectiondiffraction 5 minutes after being mix-injected

into the HVC injector. Subsequently, SFX data at pH 8.4, pH 7.27, pH 6.36, pH 5.13, pH 6.74, pH 6.19, and pH 5.45 were collected over 178, 196, 114, 153, 21, 45, and 71 minutes, at 293K. Detailed information was indicated in Table 3.4.

3.4.3. Data processing and structure determination

Diffraction images were processed to identify 'hit' images through the *CHEETAH* software package [Barty et al., 201], adapted for SFX data processing at SACLA [Nakane et al., 2016]. A 'hit' image contained a minimum of 20 Bragg peaks. The detector geometry refinement was conducted using geoptimiser in *CrystFEL* version 0.10.2 [White et al., 2012, 2019]. *CrystFEL* and *DirAx* carried out auto-indexing and integration [Duisenberg, 1992]. The resolution cutoffs were adjusted to be 2.45 Å for glargine at pH-8.4, 2.4 Å for glargine at pH-7.27, 2.48 Å for glargine at pH-6.74, 2.75 Å for glargine at pH-6.19, 2.88 Å for glargine at pH-6.36, 3.23 Å for glargine at pH-5.45, and 3.2 Å for glargine at pH-5.13, based on CC1/2 more than 30% and/or mean I/sigma>1.0. Structure determination for glargine employed molecular replacement with the *PHASER* program, part of the *PHENIX* software package [McCoy et al., 2007], using the 4AJX structure [Steensgaard et al., 2013] for the $P12_11$ space group and 4GBN [Palmieri et al., 2013] for the R3:H space group as the search models. All refinements were carried out using the phenix.refine [Afonine et al., 2012], followed by manual examination and rebuilding of the refined coordinates in *COOT* [Emsley et al., 2010]. Tables 3.4 summarizes data collection and refinement statistics. *PyMOL* (DeLano Scientific; http://www.pymol.org) was used to create all figures.

3.4.4. Observations

We determined seven ambient temperature glargine structures by employing Serial Femtosecond Crystallography (SFX) to record in a 5 mins time-resolved and pH-jumped manner. This method provided for the first time a direct high-resolution three-dimensional (3D) observation of pH-dependent allosteric changes in glargine at room temperature, with resolutions ranging between 2.4 Å and 3.23 Å at ambient temperature. Utilizing 10-femtosecond XFEL pulses at the SPring-8 Angstrom Compact Free Electron Laser (SACLA), our data collection strategy adhered to the "diffraction before destruction" principle, effectively minimizing the impact of X-ray-induced radiation

damage on the crystallographic structures.

We primarily focused on pH values of 8.4, 7.27, 6.36, and 5.13, as determined through our pH trials on crystals (Figure 3.32). Subsequently, pH values of 6.74, 6.19, and 5.45 were derived from the initial pH 5.13; thus, these values have not been applied to crystals for testing. It is noteworthy that following the pH 5.13 stage, the crystals gradually disappeared (Figure 3.32). Considering the pH-dependent conditions, we opted to choose data collection points at pH values of 8.4, 7.27, 6.74, 6.36, 6.19, 5.45, and 5.19.

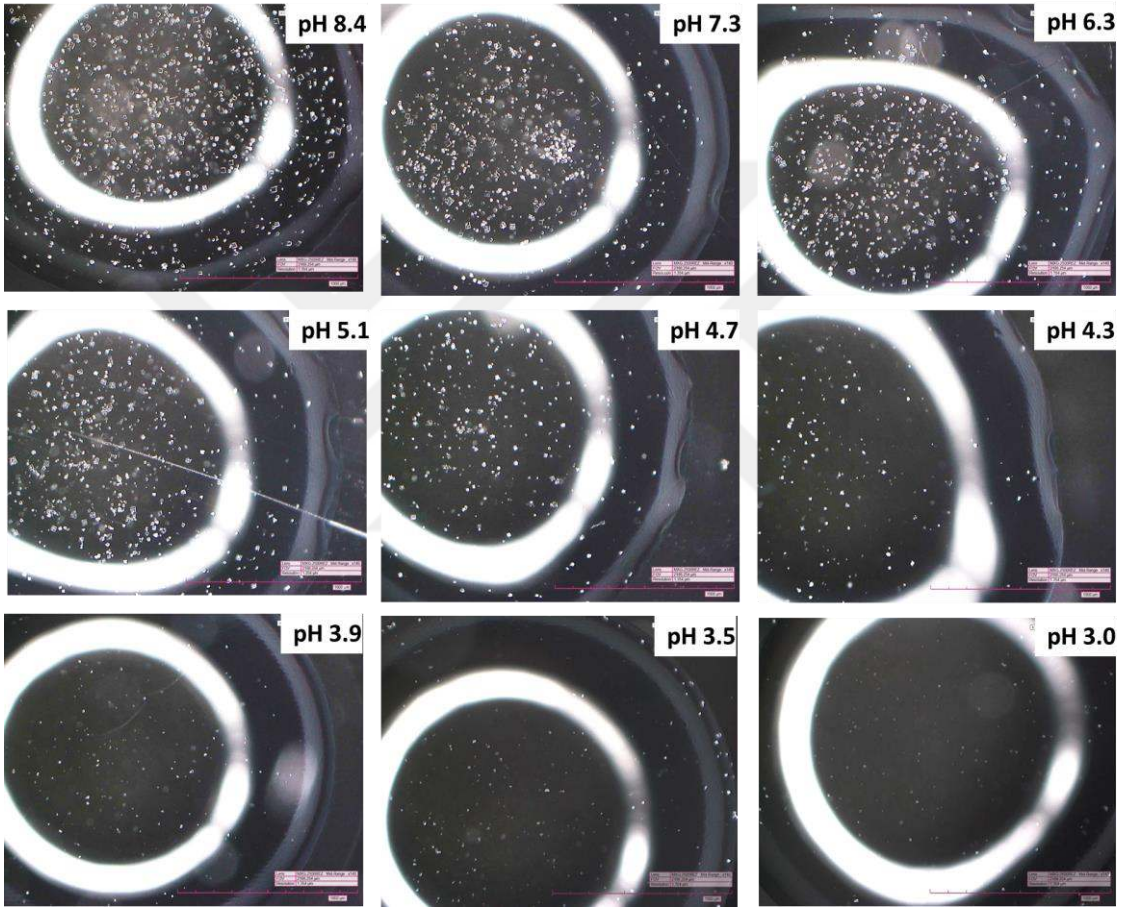


Figure 3.32 Crystal pH stability screening.

The first row indicates a range between pH 8.4 and 6.3, the second row displays a range between pH 5.1 and 4.3, and the last row demonstrates a range between pH 3.9 and 3.0. The first four pH degree have been selected for data collection, while the the other values were not utilized. Crystals disappear towards pH 3.0

Sample preparation was finalized more than five minutes before each data collection, as illustrated in Figure 3.33. First, a native crystal solution at pH 8.4 (10 ul) was mixed with 90 ul of grease, resulting in a final pH of 8.4. Second, a crystal solution

at pH 8.4 (10 ul) and a buffer at pH 5.75 (5 ul) were combined with 85 ul of grease, producing a final pH of 7.27 (Measured direclty). Likewise, crystal solutions at pH 8.4 (10 ul) were blended with buffers at pH 4.45 (5 ul), pH 4.00 (5 ul), pH 4.00 (4 ul), pH 4.00 (3 ul), and pH 4.00 (2 ul). The resulting mixtures, each combined with 85 ul, 85 ul, 86 ul, 87 ul, and 88 ul of grease, respectively, led to final pH valuesdegrees of 6.36, 5.13, 5.45, 6.19, and 6.74 (measured directly).

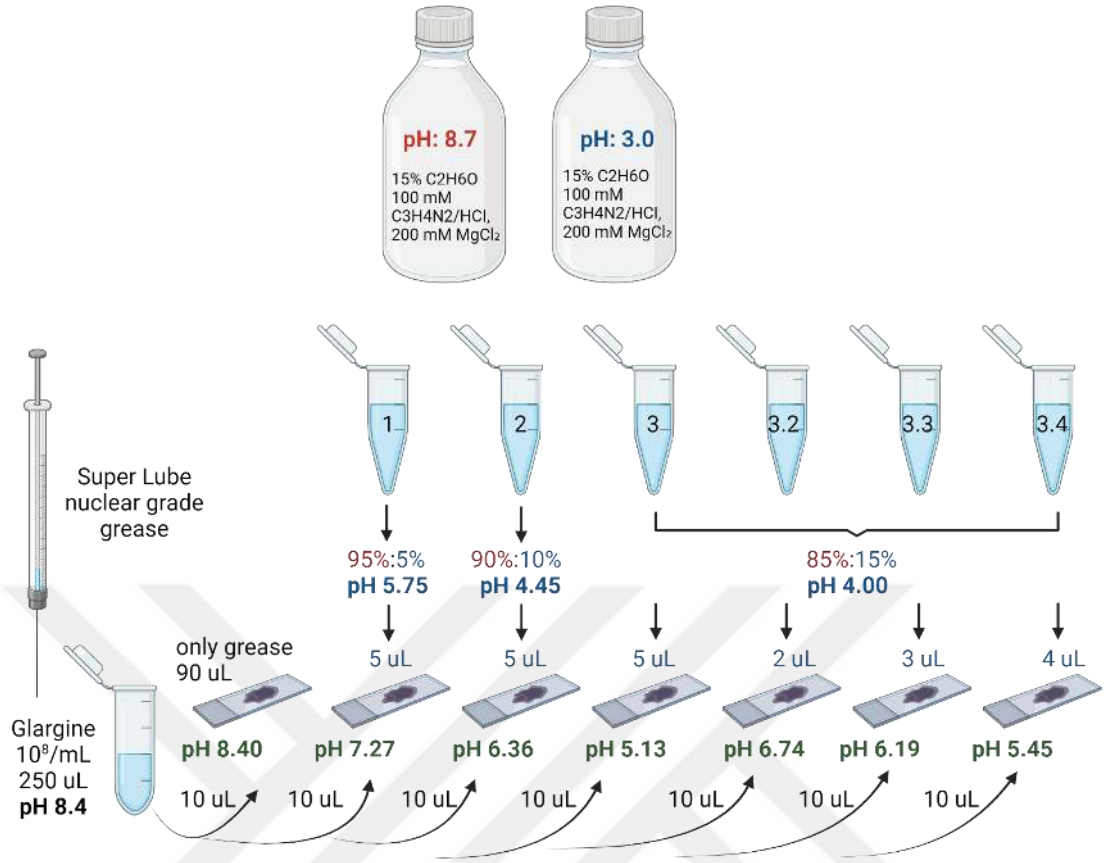


Figure 3.33 Diagram the sample preparation procedure for various pH levels.

Prior to 5 minutes before each data collection, sample preparation was completed accordingly. In the first step, a glargine crystal solution at pH 8.4 (10 ul) was mixed with 90 ul of grease, resulting in a final pH of 8.4. Subsequently, SFX data at pH 8.4 were collected over 178 minutes at 293 K. In the next experiment, a crystal solution at pH 8.4 (10 ul) and a buffer at pH 5.75 (5 ul) were combined with 85 ul of grease, yielding a final pH of 7.27. Then, SFX data at pH 7.27 were collected for 196 minutes at 293 K. Similarly, crystal solutions at pH 8.4 (10 ul) were mixed with buffers at pH 4.45 (5 ul), resulting in a final pH of 6.36. The subsequent SFX data at pH 6.36 were collected over 114 minutes at 293 K. In a different trial, crystal solutions at pH 8.4 (10 ul) were combined with a buffer at pH 4.00 (5 ul), leading to a final pH of 5.13. The corresponding SFX data at pH 5.13 were collected over 153 minutes at 293 K. Continuing with the variations, crystal solutions at pH 8.4 (10 ul) and a buffer at pH 4.00 (2 ul) were combined with 88 ul of grease, resulting in a final pH of 6.74. Subsequent SFX data at pH 6.74 were collected over 21 minutes at 293 K. In another case, crystal solutions at pH 8.4 (10 ul) and a buffer at pH 4.00 (3 ul) were mixed with 87 ul of grease, yielding a final pH of 6.19. The corresponding SFX data at pH 6.19 were collected over a 45-minute at 293 K. Finally, crystal solutions at pH 8.4 (10 ul) and a buffer at pH 4.00 (4 ul) were mixed with 86 ul of grease, resulting in a final pH of 5.45. Subsequently, SFX data at pH 5.45 were collected for 71 minutes at 293 K. Data collection period has been indicated in Table 3.4. *All the pH values have been directly measured by a calibrated microprob equipped pH meter.*

Data collection was performed through a 30-hour time slot (Table 3.4). During the glargine crystals at lower pHs, the Bragg diffraction is still observable as weaker, which is compatible with crystal pH screening testing (Figure 3.34 and Figure 3.35). We first performed SFX using only glargine crystals (10^8 ml^{-1}) with a viscous media mixing delay of more than 5 min (5-min-mix) to further demonstrate the feasibility of converting the glargine *in crystallo* to the final pH-dependent reaction state via allosteric conformational changes.

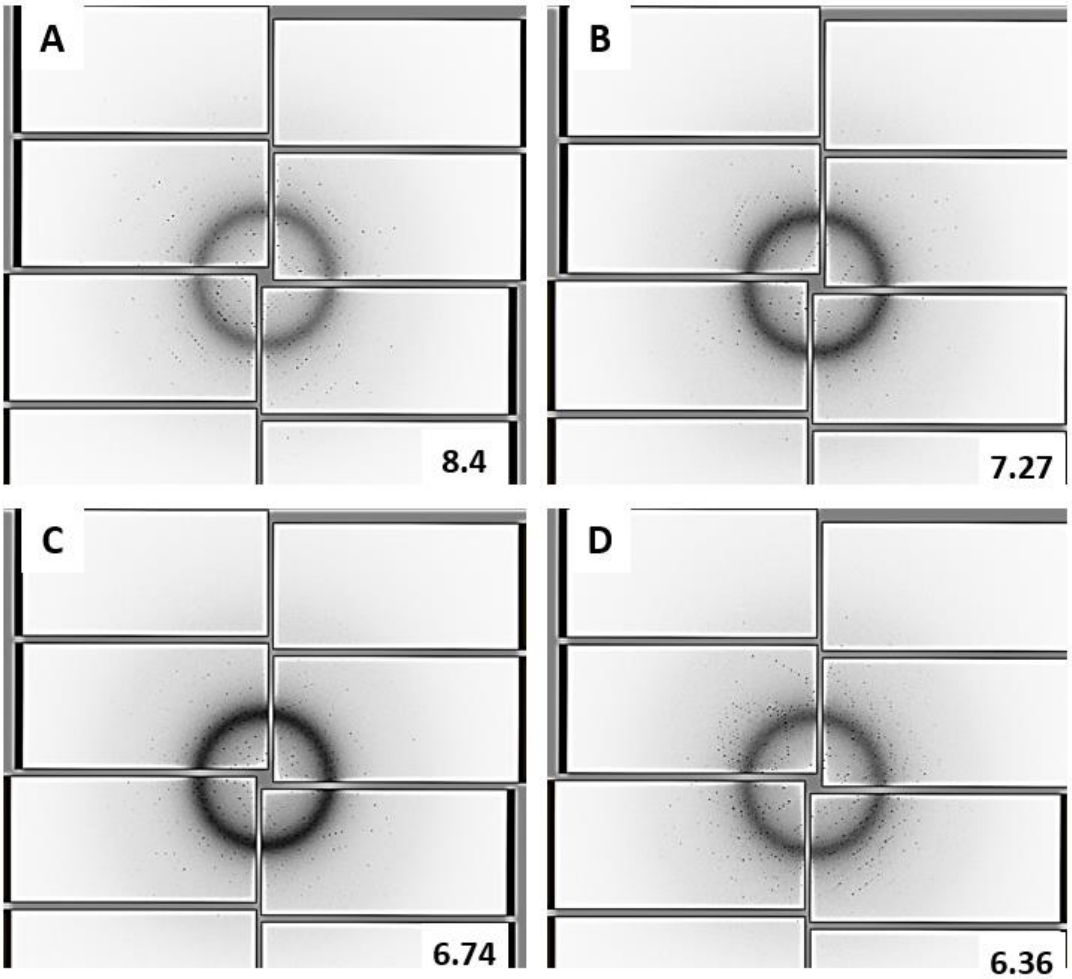


Figure 3.34 Representation of the diffraction pattern at pH 8.4, pH 7.27, pH 6.74, and pH 6.36.

Both pH 8.4 and pH 7.27 conformers were entirely hexamer state in asymmetric unit, resulting in the identical space group and unit cell parameters that even no matched a previously reported glargine structures at PDB (4IYD, 5VIZ). SFX data at pH 8.4 and pH 7.27, in the same unit cell $P12_11$, were collected over 178- and 198- minutes at 293 K.

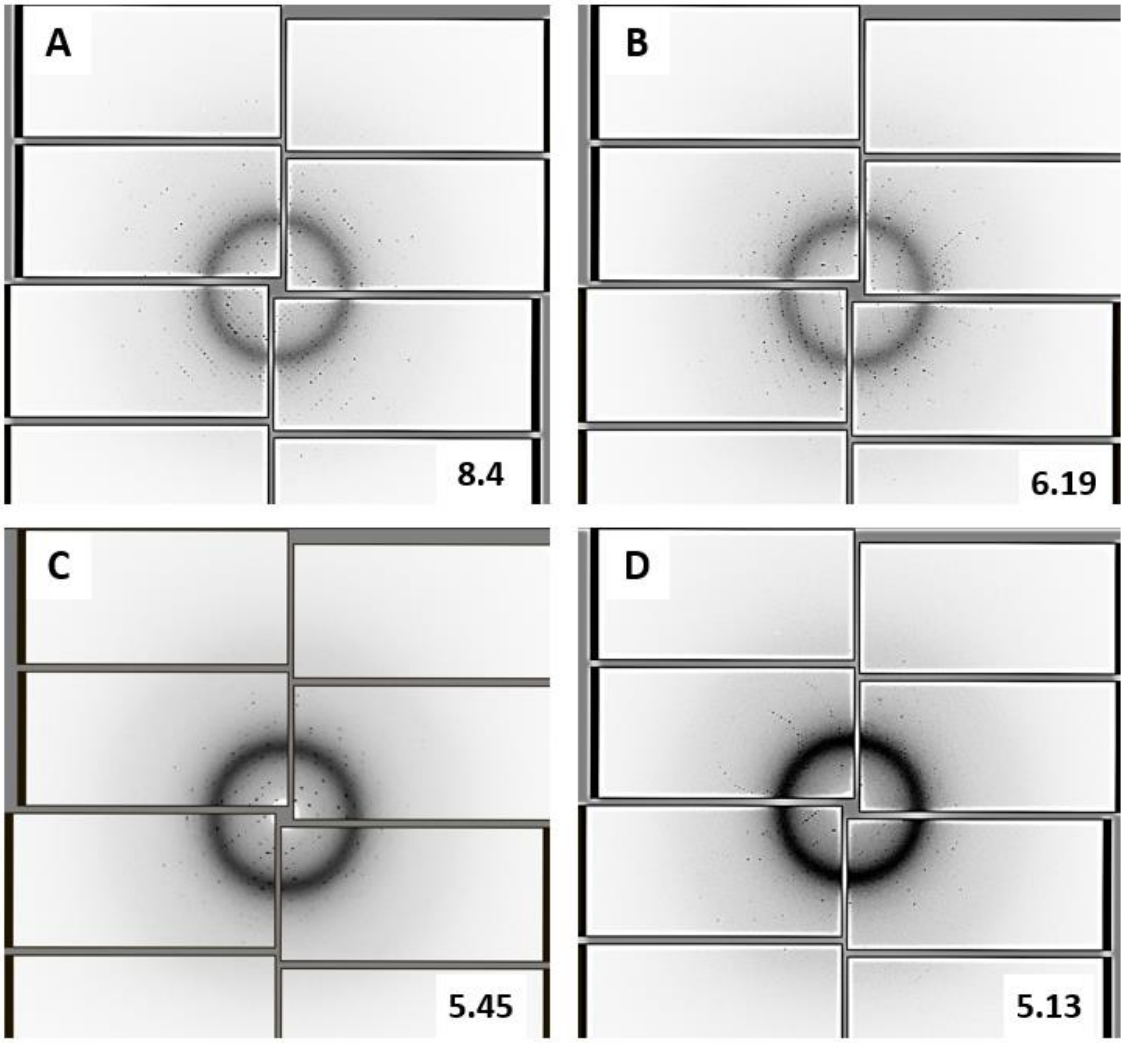


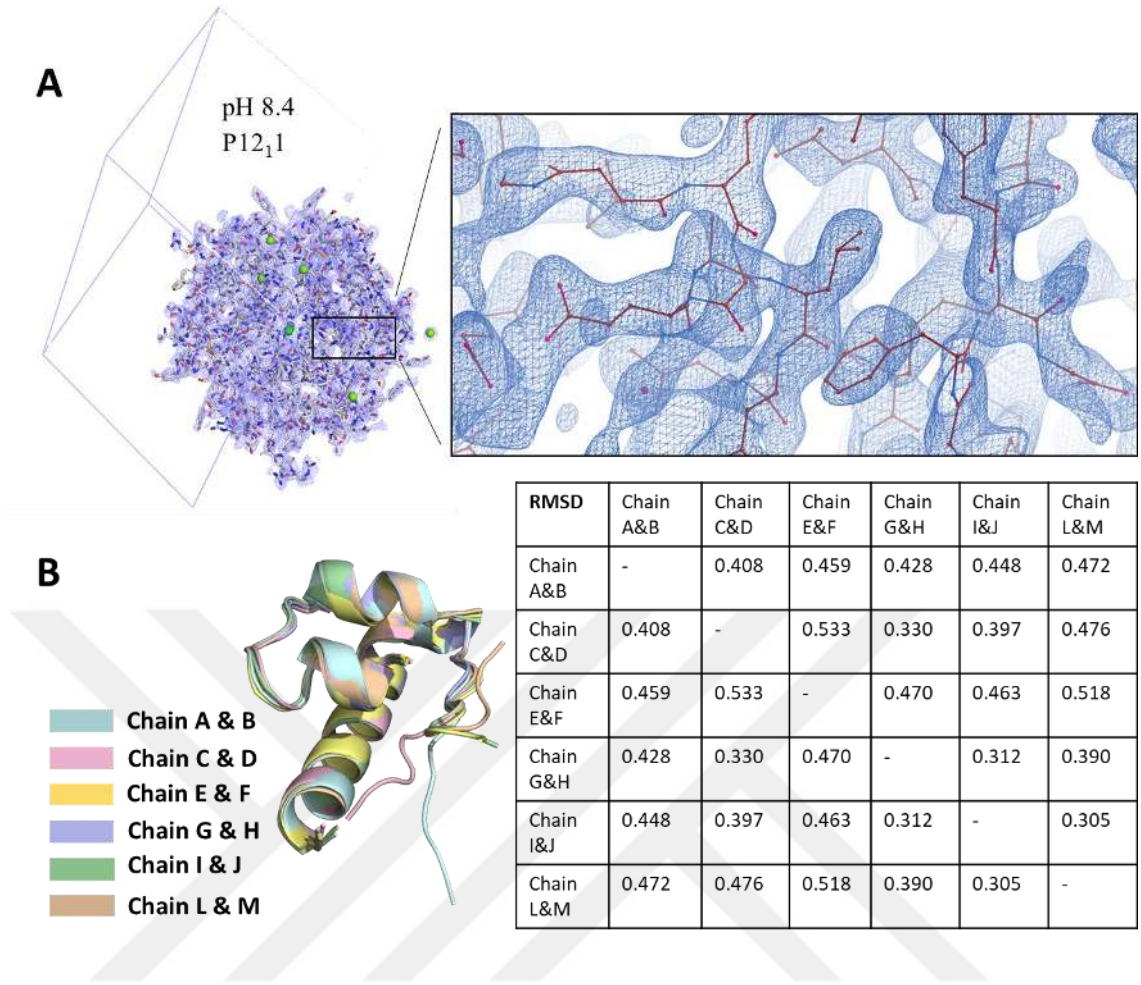
Figure 3.35 Representation of the diffraction pattern at pH 8.4, pH 6.19, pH 5.45, and pH 5.13.

Interestingly, we observed that pH 6.74, pH 6.36, pH 6.19, pH 5.45, and pH 5.13 conformers were dimer state in the asymmetric unit, but they are hexamer in the biological unit, resulting in the distinct space group and unit cell parameters compared to observing the previous $P12_11$ pH 8.4 and 7.27 structures. SFX data at pH 6.74, 6.36, 6.19, 5.45, and pH 5.13 were collected over 21-, 114-, 45-, 71, and 153- minutes at 293 K. Such a transition suggests that glargine molecules have undergone 'sufficiently ordered' allosteric alteration that is fully pH-dependent *in crystallo* that stimulated the formation of a new lattice. Such a transformation might be deemed implausible in macroscopic insulin crystals. Notably, reactions involving insulin structures that have not been studied thus far using time-resolved crystallography are those in which the minor changes are not large enough to induce a lattice transition.

Table 3.4 Data collection and data processing information of seven glargine structures.

sample	time (min)	total	processed	hits	hits accepted	indexed	index_hit	unit cell	R_{work}	R_{free}	SG	Res (Å)	Cmp %
glargine_pH_8.4_1		174902	174902	12363	7.1%	10295	83.3 %						
glargine_pH_8.4_2		66097	66068	3087	4.7%	2527	81.9 %						
glargine_pH_8.4_3	178	68661	68661	8946	13.0 %	6358	71.1 %	47.9 62.5 56.9 90 111.4 90	36%	41 %	P1211	2.40	98.1
glargine_ph_7.27_1		99253	99235	5680	5.7%	4286	75.5 %						
glargine_ph_7.27_2		131592	131592	9301	7.1%	7046	75.8 %						
glargine_ph_7.27_3	196	66709	66709	2838	4.3%	1677	59.1 %	48 62.4 57.0 90 111.4 90	34%	38 %	P1211	2.45	97.8
glargine_ph_6.74	21	38767	38767	6911	17.8 %	3843	55.6 %	81.7 81.7 39.7 90 90 120	30%	36 %	R3:H	2.48	77.2
glargine_ph_6.36_1		61761	61761	9686	15.7 %	4374	45.2 %						
glargine_ph_6.36_2	114	95000	95000	11647	12.3 %	6506	55.9 %	82.8 82.8 39.9 90 90 120	31%	40 %	R3:H	2.88	78.4
glargine_ph_6.19	45	75280	75280	6106	8.1%	3179	52.1 %	81.9 81.9 39.3 90 90 120	29%	36 %	R3:H	2.75	78.3
glargine_ph_5.45	71	120981	120981	2882	2.4%	1367	47.4 %	82.1 82.1 39.7 90 90 120	37%	49 %	R3:H	3.23	79.1
glargine_ph_5.13_1		134888	134888	2700	2.0%	1402	51.9 %						
glargine_ph_5.13_2		56839	56820	271	0.5%	152	56.1 %						
glargine_ph_5.13_3	153	67549	67549	199	0.3%	101	50.8 %	83.3 83.3 40.2 90 90 120	31%	37 %	R3:H	3.20	79.1

The structure of glargine obtained through SFX at pH 8.4 was resolved to a resolution of 2.40 Å in the $P12_11$ space group. Evaluation of the Root Mean Square Deviation (RMSD) values for each monomer reveals the presence of six closely similar monomer molecules within the asymmetric unit cell, exhibiting RMSD values ranging from 0.305 to 0.533 (Figure 3.36, Table 3.5). Additionally, the asymmetric unit comprises six phenol molecules, two zinc cations, two chloride anions, and 34 water molecules.



RMSD	Chain A&B	Chain C&D	Chain E&F	Chain G&H	Chain I&J	Chain L&M
Chain A&B	-	0.408	0.459	0.428	0.448	0.472
Chain C&D	0.408	-	0.533	0.330	0.397	0.476
Chain E&F	0.459	0.533	-	0.470	0.463	0.518
Chain G&H	0.428	0.330	0.470	-	0.312	0.390
Chain I&J	0.448	0.397	0.463	0.312	-	0.305
Chain L&M	0.472	0.476	0.518	0.390	0.305	-

Figure 3.36 The SFX glargine structure at pH 8.4.

(A) Hexamer form of glargine in the asymmetric unit at P12$_1$1 space group and close view of refined 2*Fo-Fc* simulated annealing-omit map at 1 sigma level is colored in slate. (B) Superposition the each chain of hexamer glargine structure with an overall RMSD between 0.305 and 0.533.

In close proximity to the R31, E4, N18, T27, G1, S9, F1, L17, and E21 residues of the largely B chains with noncanonical regions in the P12$_1$1 structure, there was a disordered electron density peak with a height of 1 σ in the 2*Fo-Fc* map and *mFo-DFc* difference omit map contoured at 0.23e/A^3. These corresponding residues in the P12$_1$1 structure were largely disordered or not present, probably due to flexibility that is originated from (elevated temperature; 100 K cryogenic vs 293 K) ambient temperature, to be interpreted similarly (Figure 3.37).

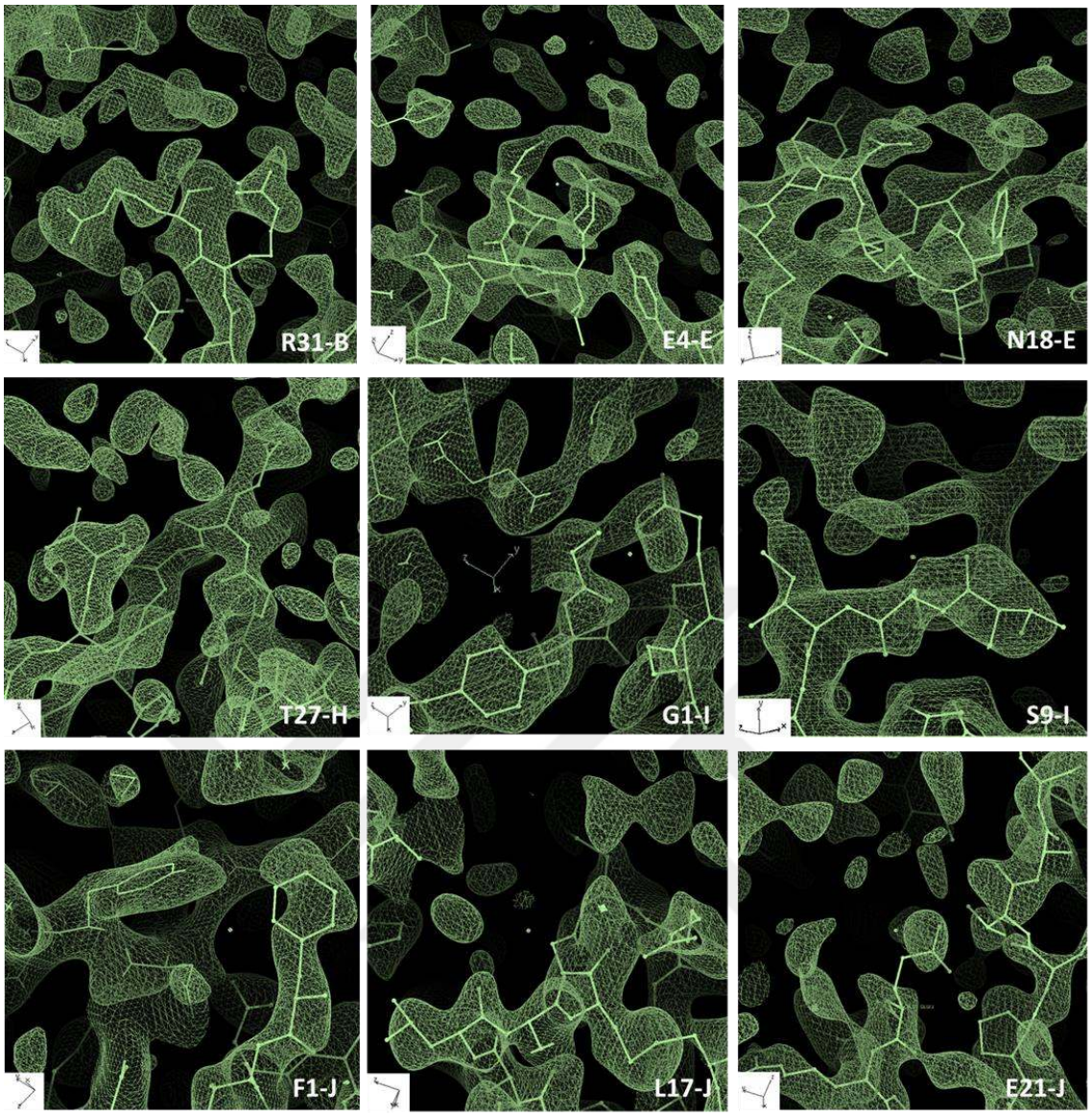
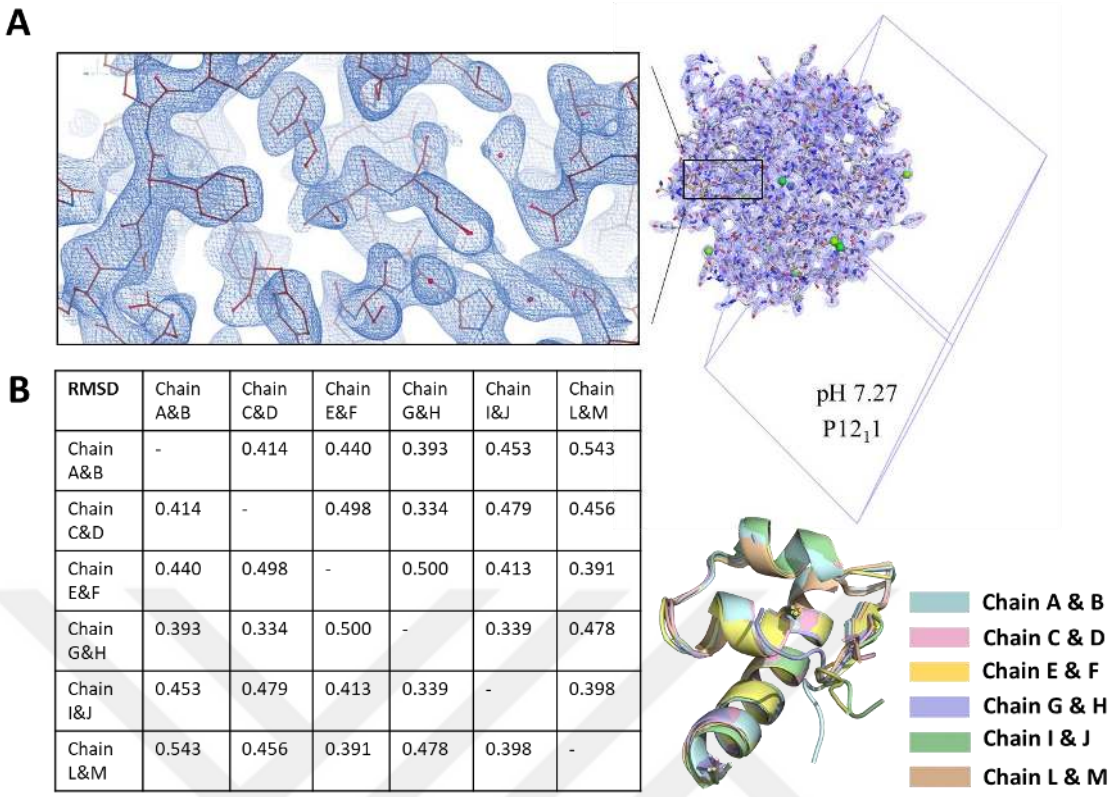


Figure 3.37 The newly created undefined and/or disordered electron density peaks.

Representation the density around the certain residues of the glargine structure at pH 8.4 have been represented in a 2*Fo-Fc* simulated annealing-omit map contoured at 1 σ level. The lower left corner indicates the estimated proximity of X, Y, and Z coordinates.

The configuration of glargine, as determined by SFX at pH 7.27, was determined with a precision of 2.45 Å within the $P12_11$ space group. Examination of the RMSD values for individual monomers uncovers the existence of six closely identical monomer molecules confined within the asymmetric unit cell, showcasing RMSD values spanning from 0.334 to 0.543 (Figure 3.38, Table 3.5). Furthermore, the asymmetric unit contains six phenol molecules, two zinc cations, two chloride anions, and 29 water molecules.



RMSD	Chain A&B	Chain C&D	Chain E&F	Chain G&H	Chain I&J	Chain L&M
Chain A&B	-	0.414	0.440	0.393	0.453	0.543
Chain C&D	0.414	-	0.498	0.334	0.479	0.456
Chain E&F	0.440	0.498	-	0.500	0.413	0.391
Chain G&H	0.393	0.334	0.500	-	0.339	0.478
Chain I&J	0.453	0.479	0.413	0.339	-	0.398
Chain L&M	0.543	0.456	0.391	0.478	0.398	-

Figure 3.38 The SFX glargine structure at pH 7.27.

(A) Hexamer form of glargine in the asymmetric unit at P12$_1$1 space group and close view of refined 2*Fo-Fc* simulated annealing-omit map at 1 sigma level is colored in slate. (B) Superposition the each chain of hexamer glargine structure with an overall RMSD between 0.334 and 0.543.

Table 3.5 Data collection and refinement statistics for glargine at 8.40 and 7.27 pH-values.

Data collection	**Glargine1 (pH:8.4)**	**Glargine2 (pH: 7.27)**
Instrument	SACLA/DAPHNIS	SACLA/DAPHNIS
Resolution range	19.82 - 2.4 (2.486 - 2.4)	19.37 - 2.45 (2.537 - 2.45)
Space group	P 1 21 1	P 1 21 1
Unit cell	47.9 62.6 56.9 90 111.4 90	48 62.4 57.1 90 111.4 90
Redundancy	14.7	107.19
Completeness (%)	98.11 (94.87)	97.81 (94.73)
Mean I/sigma(I)	3.04 (1.81)	2.56 (1.74)
Wilson B-factor	27.15	28.94
R-merge	0.2526 (0.3997)	0.3207 (0.4661)
R-meas	0.3573 (0.5652)	0.4535 (0.6591)
R-pim	0.2526 (0.3997)	0.3207 (0.4661)
CC1/2	0.775 (0.698)	0.658 (0.515)
CC*	0.935 (0.907)	0.891 (0.825)
R-work	0.3618 (0.4701)	0.3437 (0.4279)
R-free	0.4163 (0.4654)	0.3811 (0.4853)
No. of reflections	24032 (2448)	22617 (2285)
Unique reflection	12352 (1192)	11658 (1118)
Protein residues	310	313
RMS (bonds)	0.001	0.001
RMS (angles)	0.35	0.39
Ramachandran favored (%)	97.20	98.27
Ramachandran allowed (%)	2.80	1.04
Ramachandran outliers (%)	0.00	0.69
Average B-factor	36.50	41.43
macromolecules	36.63	41.66
ligands	34.29	34.80
solvent	33.98	33.88

Similarly, the identical residues R31, E4, N18, T27, G1, S9, F1, L17, and E21 within the glargine structure at pH 7.27 exhibit a noncanonical electron density pattern distinct

from the observed ones in the glargine structure at pH 8.4, which adheres to a canonical modeled structure. In the 2*Fo-Fc* map, characterized by the presence of the irrelevant electron density peaks with a height of 1 σ, and in the *mFo-DFc* difference omit map, contoured at ~1σ. These corresponding regions in the $P12_11$ structure are predominantly disordered or absent, likely due to Brownian motion arising from ambient temperature as well as pH differences, creating their interpretation similarly challenging (Figure 3.39). This output has been supported by ellipsoid-thermal displacement mentioned in Figure 3.48. It is noteworthy that L17-J and E21-J regions display striking differences between pH 8.4 and pH 7.27 structures; overall, a similar density pattern has been observed, though.

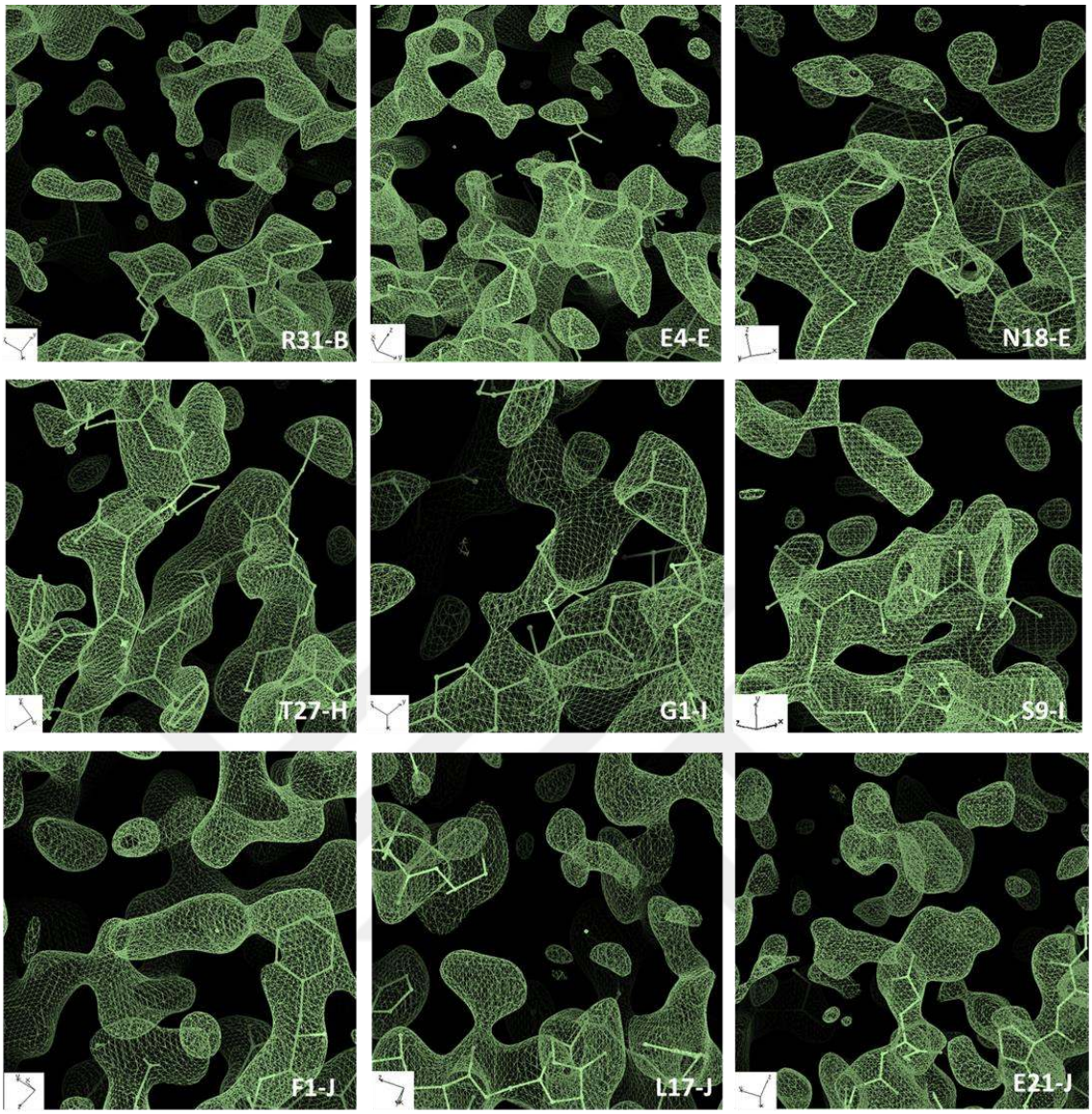


Figure 3.39 The newly created undefined and/or disordered electron density peaks.
Representation of the peaks around the certain residues of the glargine structure at pH 7.27 have been represented in a 2*Fo-Fc* simulated annealing-omit map contoured at 1 σ level. The lower left corner indicates the estimated proximity of X, Y, and Z coordinates.

The structures of glargine obtained through SFX at pH 6.74 and 6.36 were determined with resolutions of 2.48 Å and 2.88 Å, respectively, in the R3:H space group, triggered by pH-shifting. Assessment of the RMSD values for each monomer in both structures reveals the existence of two closely similar monomer molecules within the asymmetric unit cell. The RMSD values are determined as 0.414 for the glargine structure at pH 6.74 and 0.637 for the glargine structure at pH 6.36 (Figure 3.40, Table 3.6). One zinc cation and one chloride anion are present in the asymmetric unit. Intriguingly, there is a noticeable reduction in water molecules, and the phenolic agent's presence diminishes

markedly, attributed to the gradual alterations in the surrounding environment.

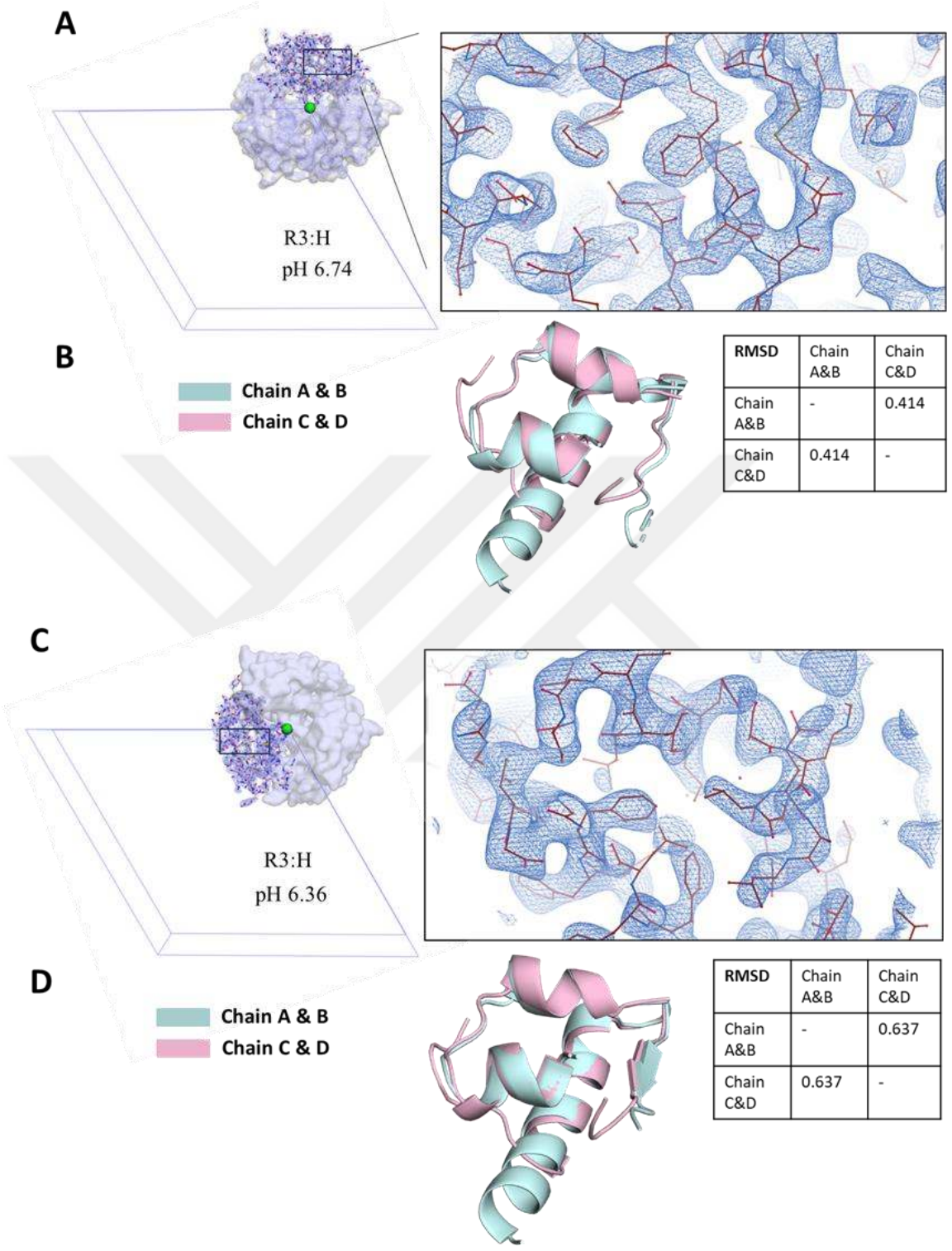


Figure 3.40 The SFX glargine structures at pH 6.74 and pH 6.36.

(A,C) Dimer forms of glargine in the asymmetric unit have changed sharply to the R3:H space group. The close view of the refined *2Fo-Fc* simulated annealing-omit map at 1 sigma level is colored in slate for both. (B,D) Superposition of each chain of hexamer glargine structures with an overall RMSD of 0.414 and 0.637, respectively. R3:H space group has dimer conformers in the asymmetric unit but hexamer form in the biological unit, represented by surface modes in the A, and C panels.

Table 3.6 Data collection and refinement statistics for glargine at 6.74 and 6.36 pH-values.

Data collection	**Glargine3.2 (pH: 6.74)**	**Glargine2 (pH: 6.36)**
Instrument	SACLA/DAPHNIS	SACLA/DAPHNIS
Resolution range	17.61 - 2.481 (2.569 - 2.481)	19.24 - 2.88 (2.983 - 2.88)
Space group	R 3 :H	R 3 :H
Unit cell	81.74 81.74 39.79 90 90 120	82.89 82.89 39.94 90 90 120
Redundancy	34.31	107.19
Completeness (%)	77.25 (72.16)	78.83 (77.93)
Mean I/sigma(I)	2.99 (1.6)	4.41 (2.62)
Wilson B-factor	49.70	52.74
R-split	0.29 (0.54)	0.17 (21.83)
CC1/2	0.89 (0.60)	0.96 (0.87)
CC*	0.97 (0.78)	0.99 (0.96)
R-work	0.3082 (0.4873)	0.2964 (0.3905)
R-free	0.3669 (0.5273)	0.3951 (0.3504)
No. of reflections	7829 (733)	8083 (814)
Unique reflection	2712 (254)	2025 (204)
Protein residues	103	101
RMS (bonds)	0.003	0.002
RMS (angles)	0.48	0.36
Ramachandran favored (%)	97.85	97.85
Ramachandran allowed (%)	2.15	2.15
Ramachandran outliers (%)	0.00	0.00
Average B-factor	73.90	38.22
macromolecules	74.06	38.24
ligands	48.30	34.22
solvent	58.68	35.19

Near the residues R31, T8, G21, F1, N3, C19, H10, P28, and K29 of glargine structures at pH 6.74 and pH 6.36, a noncanonical electron density pattern was also observed within the R3:H structures. Similarly, disordered electron density peak with a height of 1 σ was detected in the 2*Fo-Fc* map, and the *mFo-DFc* difference omit map, contoured at 0.23e/A^3. These residues in the R3:H structures were associated mainly with

receptor-interacting domains exhibiting highly disordered electron density, likely originating from newly adopted conformational changes triggered by **pH shifts**, resulting in **crystal packing rearrangement** (Figure 3.41 and Figure 3.42).

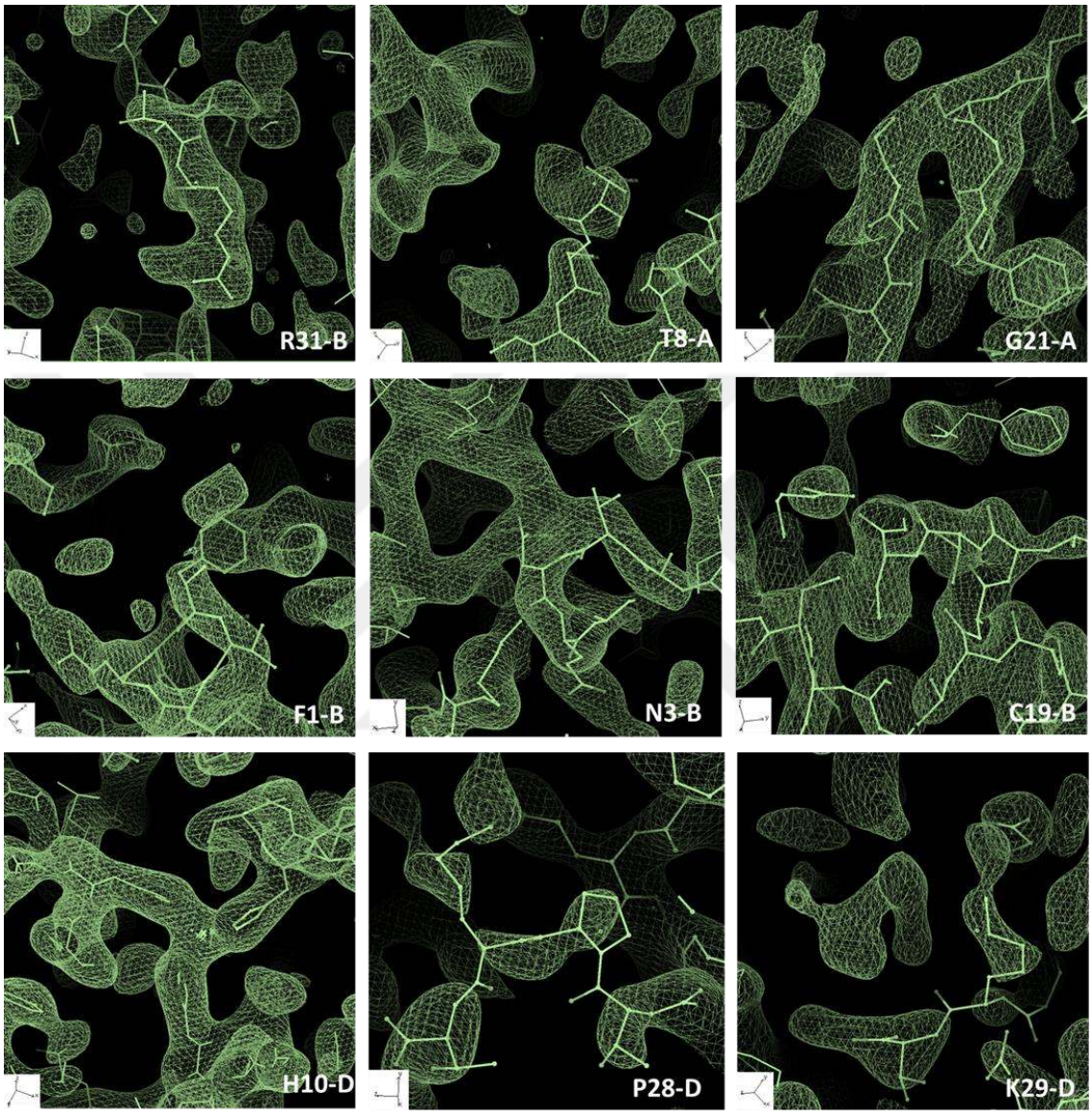


Figure 3.41 The newly created undefined and/or disordered electron density peaks.

Representation of the peaks around the certain residues of the glargine structure at pH 6.74 have been represented in a 2*Fo-Fc* simulated annealing-omit map contoured at 1 σ level. The lower left corner indicates the estimated proximity of X, Y, and Z coordinates.

Especially, F1-B and K29-D display relatively significant conformational change when comparing the structures at pH 6.74 and 6.36 (Fig. B and Fig. C), which is compatible with our thermal displacement/motion analysis showing distinct fluctuation between glargine pH 6.74 vs pH 6.36 (Figure 3.49). Additionally, T8-A, G21-A, N3-B, and K29-

D offer apparent newly formed undefined electron density, suggesting these regions of the map that are unable to be interpreted due to this electron mobility (Figure 3.41 and Figure 3.42). Finally, those disordered sections of the glargine structures frequently have essential functions.

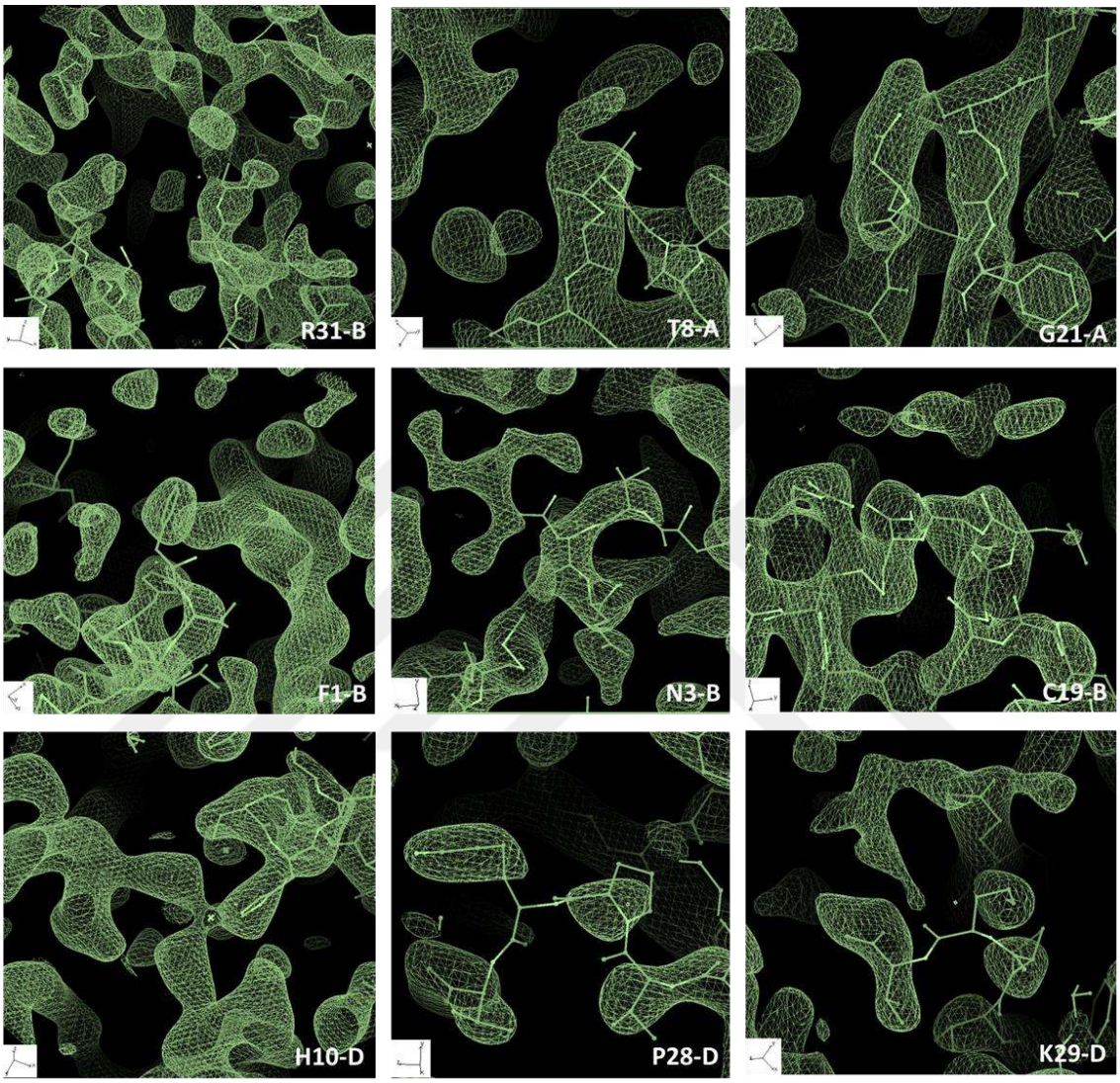
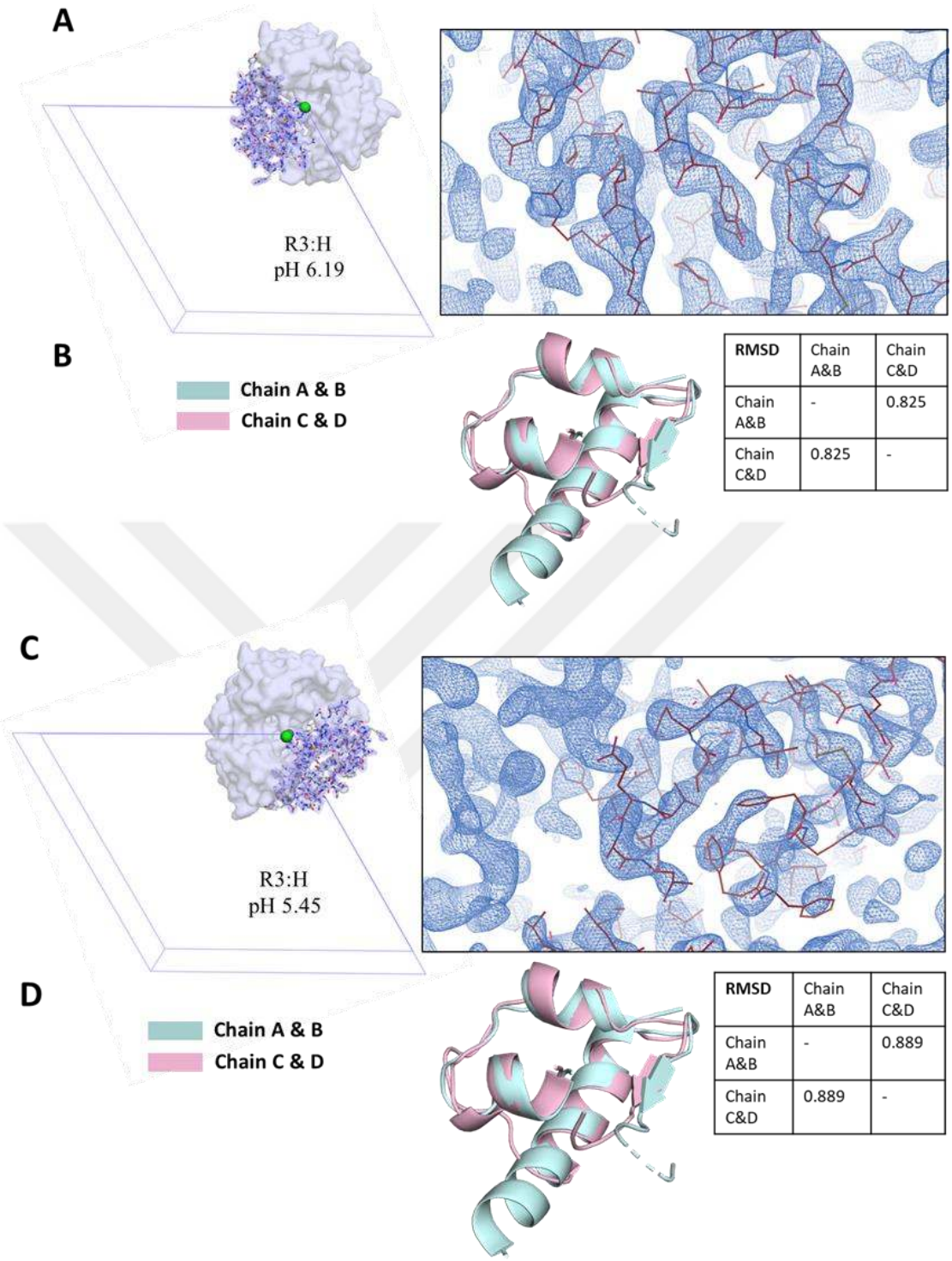


Figure 3.42 The newly created undefined and/or disordered electron density peaks.
Representation of the peaks around the certain residues of the glargine structure at pH 6.36 have been represented in a 2*Fo-Fc* simulated annealing-omit map contoured at 1 σ level. The lower left corner indicates the estimated proximity of X, Y, and Z coordinates.

The glargine structures obtained through SFX at pH 6.19 and 5.45 were determined with resolutions of 2.75 Å and 3.23 Å, respectively, within the R3:H space group, induced by pH-shifting. Evaluation of the RMSD values for each monomer in both structures reveals the presence of two closely analogs monomer molecules within the asymmetric unit cell. The RMSD values are computed as 0.825 for the glargine

structure at pH 6.19 and 0.889 for the glargine structure at pH 5.45 (Figure 3.43, Table 3.7). The asymmetric unit contains one zinc cation and one chloride anion. Notably, there is a conspicuous decrease in the number of water molecules, and the presence of the phenolic agent diminishes significantly, attributed to gradual alterations in the surrounding environment.



RMSD	Chain A&B	Chain C&D
Chain A&B	-	0.825
Chain C&D	0.825	-

RMSD	Chain A&B	Chain C&D
Chain A&B	-	0.889
Chain C&D	0.889	-

Figure 3.43 The SFX glargine structures at pH 6.19 and pH 5.45.

(A,C) Dimer forms of glargine in the asymmetric unit have the R3:H space group. The close view of the refined *2Fo-Fc* simulated annealing-omit map at 1 sigma level is colored in slate for both. (B,D) Superposition of each chain of hexamer glargine structures with an overall RMSD of 0.825 and 0.889, respectively. R3:H space group has dimer conformers in the asymmetric unit but hexamer form in the biological unit, represented by surface modes in the A, and C panels.

Table 3.7 Data collection and refinement statistics for glargine at 6.19 and 5.45 pH-values.

Data collection	**Glargine3.3 (pH: 6.19)**	**Glargine3.4 (pH: 5.45)**
Instrument	SACLA/DAPHNIS	SACLA/DAPHNIS
Resolution range	18.98 - 2.751 (2.849 - 2.751)	19.12 - 3.23 (3.345 - 3.23)
Space group	R 3 :H	R 3 :H
Unit cell	81.99 81.99 39.39 90 90 120	82.16 82.16 39.7 90 90 120
Redundancy	39.558	10.75
Completeness (%)	78.30 (73.98)	79.17 (77.85)
Mean I/sigma(I)	3.07 (1.59)	2.03 (1.1)
Wilson B-factor	52.20	39.74
R-split	0.27 (0.39)	0.51 (0.62)
CC1/2	0.91 (0.75)	0.68 (0.39)
CC*	0.97 (0.92)	0.90 (0.74)
R-work	0.2945 (0.3706)	0.3294 (0.4437)
R-free	0.3600 (0.5198)	0.3890 (0.5125)
No. of reflections	7805 (769)	7891 (733)
Unique reflection	2013 (199)	1271 (121)
Protein residues	103	102
RMS (bonds)	0.003	0.012
RMS (angles)	0.56	1.02
Ramachandran favored (%)	96.77	96.77
Ramachandran allowed (%)	3.23	1.08
Ramachandran outliers (%)	0.00	2.15
Average B-factor	57.18	47.11
macromolecules	57.37	47.26
ligands	39.59	18.38
solvent	42.56	-

Similarly, in the proximity of the residues R31, T8, G21, F1, N3, C19, H10, P28, and K29 in the glargine structures at pH 6.19 and pH 5.45, a noncanonical electron density pattern was also detected within the R3:H structures. These residues within the R3:H structures were primarily linked to receptor-interacting domains displaying significantly disordered electron density, still possibly stemming from electron mobility induced by

gradually lowered pH (from pH 6.19 → pH 5.45), continue to leading to rearrangement in crystal packing (Figure 3.44 and Figure 3.45).

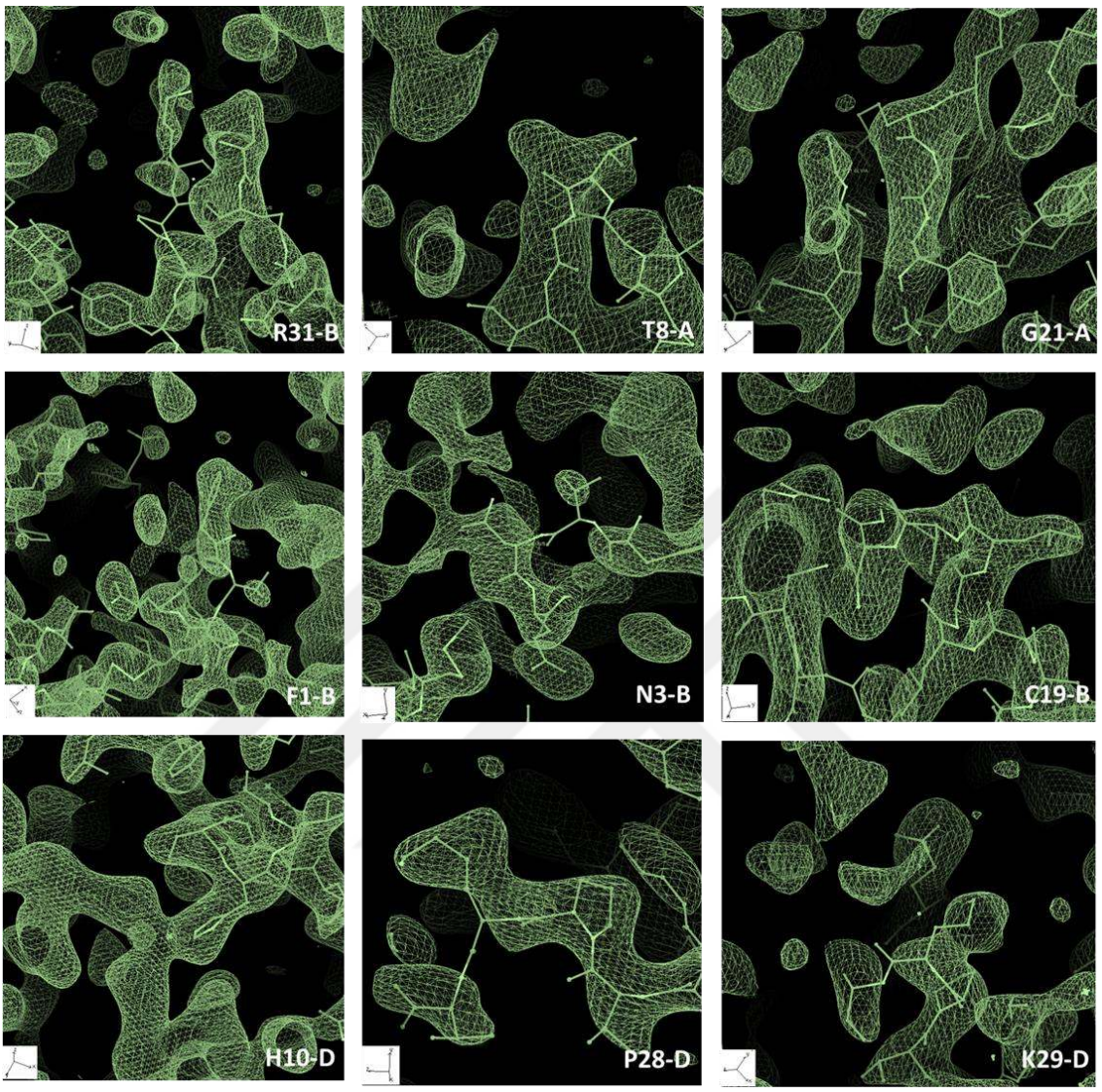


Figure 3.44 The newly created undefined and/or disordered electron density peaks.

Representation of the peaks around the certain residues of the glargine structure at pH 6.19 have been represented in a 2*Fo-Fc* simulated annealing-omit map contoured at 0.23e/A^3. The lower left corner indicates the estimated proximity of X, Y, and Z coordinates.

Additionally, R31-B, C19-B, and K29-D display the conformational changes between the structures at pH 6.19 and pH 5.45, probably due to thermal motion/fluctuations arising from ambient temperature and higher flexibility, which is supported by thermal displacement analysis. Likewise, T8-A, G21-A, F1-B, N3-B, and H10-D introduced newly assembled undefined electron density, deriving from gradually lowering pH values from pH 6.19 to pH 5.45. As expected, P28-D and K29-D are supported by electron density at pH 6.19, whereas they are characterized by fully disordered regions

at pH 5.45 (Figure 3.44 and Figure 3.45), suggesting these residues are highly flexible due to be the last residues in B-chain.

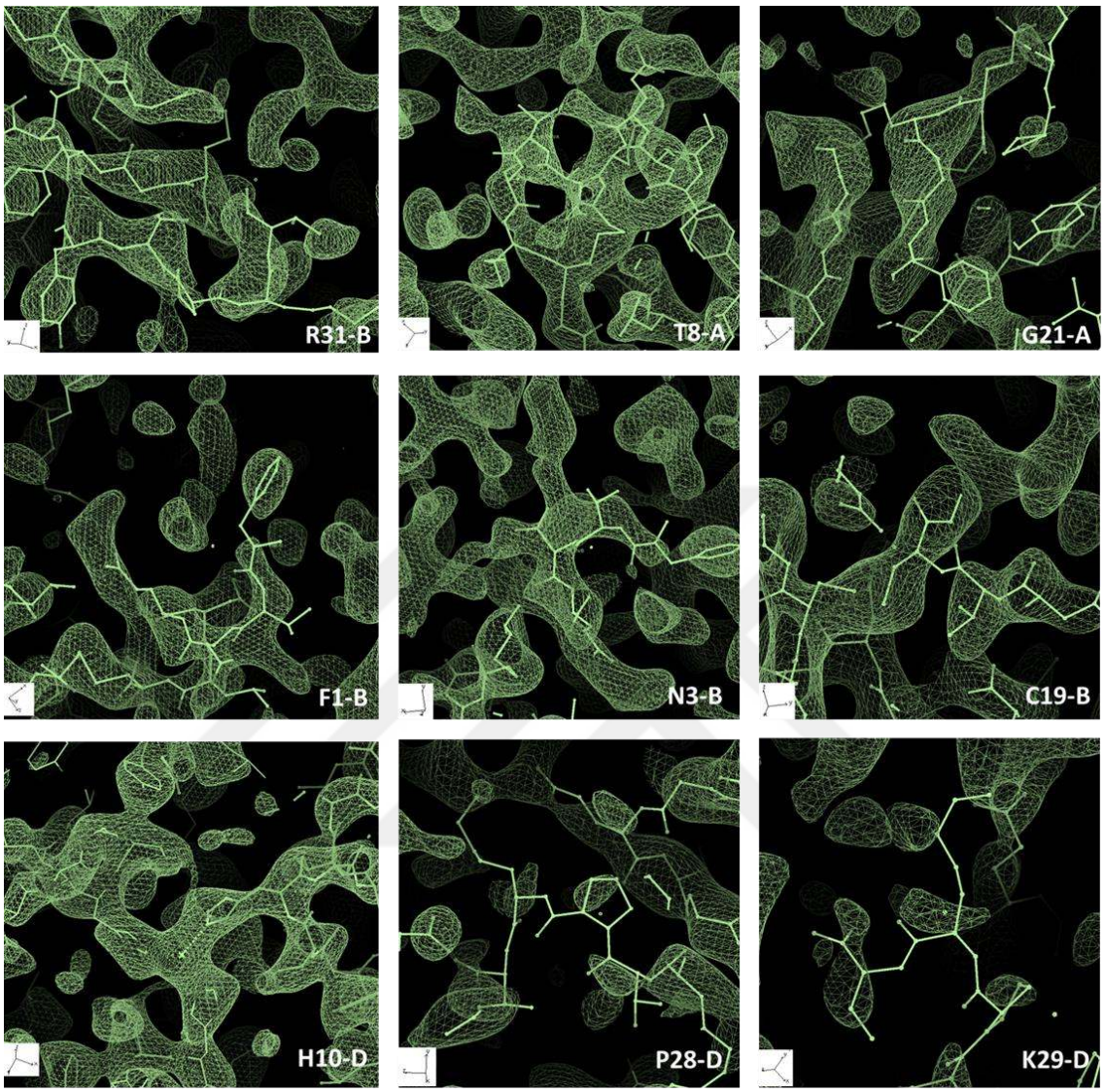
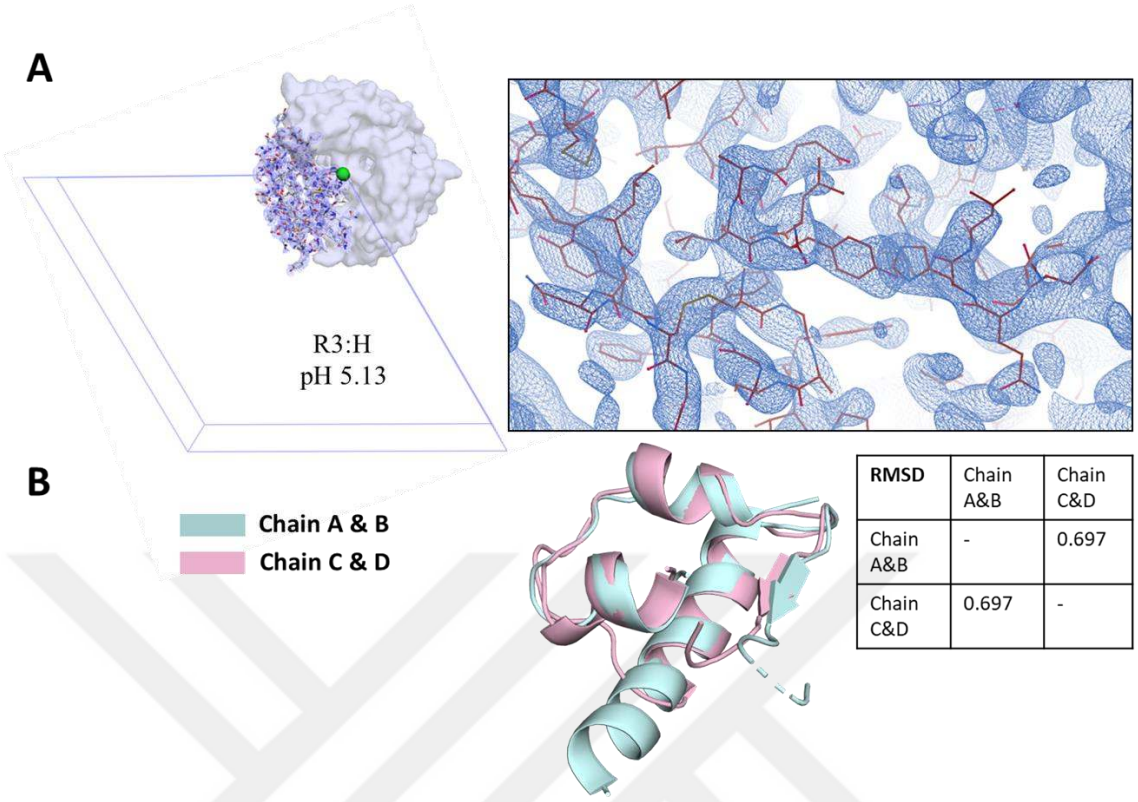


Figure 3.45 The newly created undefined and/or disordered electron density peaks.

Representation of the peaks around the certain residues of the glargine structure at pH 5.45 have been represented in a 2*Fo-Fc* simulated annealing-omit map contoured at 0.23e/A^3. The lower left corner indicates the estimated proximity of X, Y, and Z coordinates.

Finally, the glargine structure obtained through SFX at pH 5.13 was determined with resolutions of 3.20 Å within the R3:H space group, induced by pH-shifting. Analysis of the RMSD values for each monomer indicates the presence of two closely analogous monomer molecules within the asymmetric unit cell. The RMSD values are determined as 0.697 for the glargine structure at pH 5.13 (Figure 3.46, Table 3.8). The asymmetric unit comprises one zinc cation and one chloride anion. Notably, the absence of water molecules and the phenolic agent is distinctive, attributed to gradual alterations in the

surrounding environment.



RMSD	Chain A&B	Chain C&D
Chain A&B	-	0.697
Chain C&D	0.697	-

Figure 3.46 The SFX glargine structures at pH 5.13.

(A) Dimer form of glargine in the asymmetric unit have changed to the R3:H space group. The close view of the refined *2Fo-Fc* simulated annealing-omit map at 1 sigma level is colored in slate. (B) Superposition of each chain of dimer glargine structures with an overall RMSD of 0.697. R3:H space group has dimer conformers in the asymmetric unit but hexamer form in the biological unit, represented by surface modes in the A panel.

Table 3.8 Data collection and refinement statistics for glargine at 5.13.

Data collection	**Glargine3 (pH: 5.13)**
Instrument	SACLA/DAPHNIS
Resolution range	19.37 - 3.203 (3.317 - 3.203)
Space group	R 3 :H
Unit cell	83.32 83.32 40.21 90 90 120
Redundancy	13.92
Completeness (%)	79.13 (78.03)
Mean I/sigma(I)	2.19 (1.7)
Wilson B-factor	47.17
R-split	0.45 (0.46)
CC1/2	0.76 (0.59)
CC*	0.93 (0.86)
R-work	0.3176 (0.3976)
R-free	0.3799 (0.4289)
No. of reflections	8221 (799)
Unique reflection	1358 (132)
Protein residues	101
RMS (bonds)	0.002
RMS (angles)	0.37
Ramachandran favored (%)	97.85
Ramachandran allowed (%)	2.15
Ramachandran outliers (%)	0.00
Average B-factor	48.64
macromolecules	48.67
ligands	45.38
solvent	-

Lastly, a distinctive noncanonical electron density pattern was observed in the proximity of residues R31, T8, G21, F1, N3, C19, H10, P28, and K29 within the glargine structures at pH 5.13 in the R3:H space group. These residues were predominantly

associated with receptor-interacting domains, manifesting highly disordered or undefined electron density. These are likely a result of electron mobility induced by the gradual decrease in pH (transitioning from pH 5.45 to pH 5.13), persistently instigating rearrangement in crystal packing (Figure 3.47). Moreover, R31-B and K29-D exhibit noticeable conformational changes when comparing the pH 5.45 and 5.13 structures, likely attributed to thermal motion influenced by ambient temperature and increased flexibility, as corroborated by thermal displacement analysis (Figure 3.50). Similarly, F1-B, N3-B, C19-B, and K29-D introduce newly assembled & undefined electron density, a phenomenon resulting from the gradual decrease in pH values from pH 5.45 to pH 5.13 (Figure 3.47).

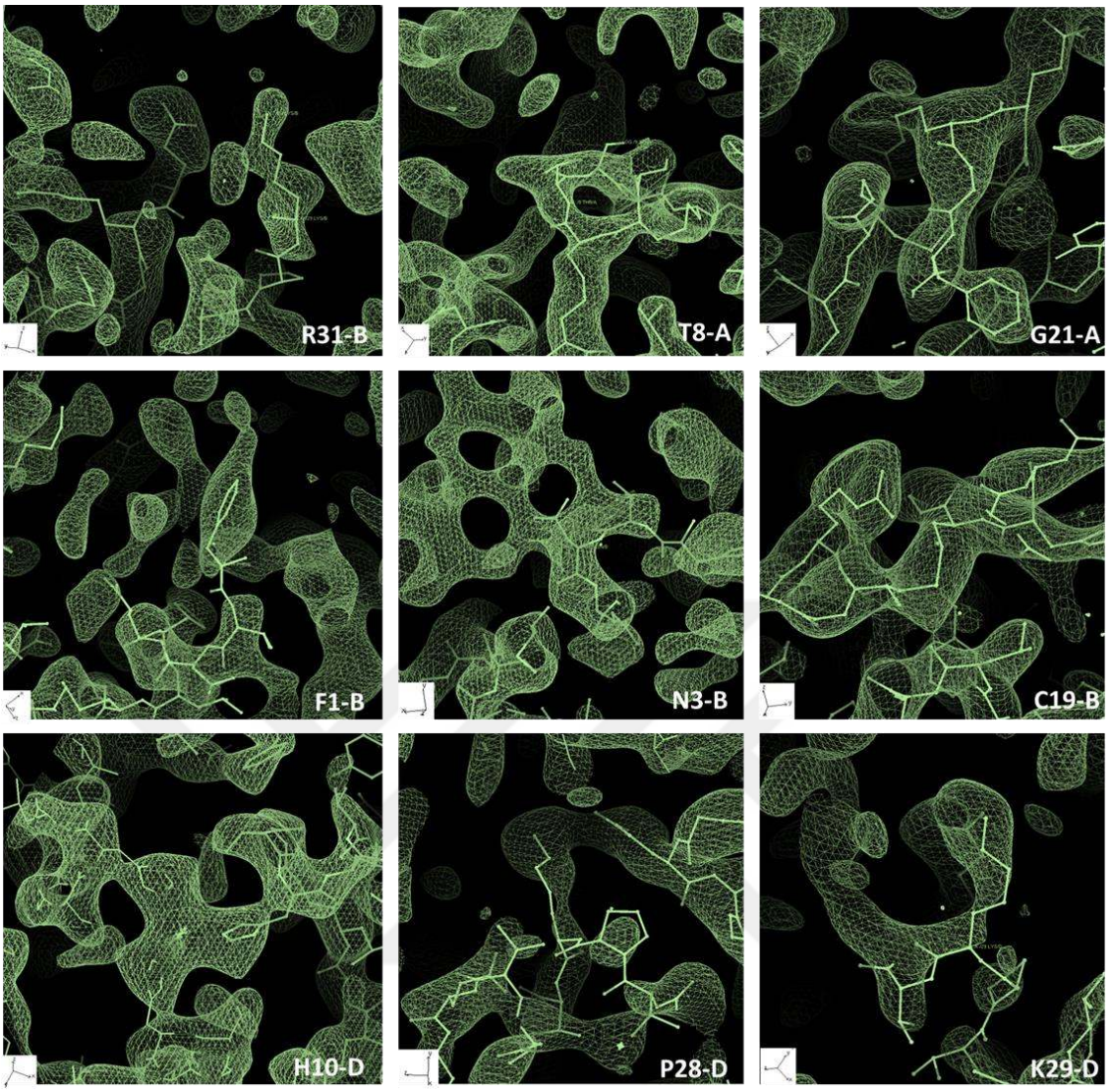


Figure 3.47 The newly created undefined and/or disordered electron density peaks.

Representation of the peaks around the certain residues of the glargine structure at pH 5.13 have been represented in a 2*Fo-Fc* simulated annealing-omit map contoured at 0.23e/A^3. The lower left corner indicates the estimated proximity of X, Y, and Z coordinates.

Conducted a thermal ellipsoid analysis and created a 2D b-factor plot utilizing the GNM [Bahar et al., 2005]. Experimental comprehension of protein equilibrium fluctuations interests the measurement of Debye-Waller/B-factors for each atom using X-ray crystallography [Bakan et al., 2011]. Despite the complexity in interpreting crystallographic B-factors, they denote the collective impact of numerous disorder sources. Diverse analytical perspectives are operated to break down molecular disorder into a hierarchical cluster of contributions, forming an intuitive foundation for quantitative structural dynamics analysis [Pearce and Gros, 2021]. As a result, the impact of pH-shift on thermal fluctuations/B-factors can be meticulously explored

through the application of thermal ellipsoid and b-factor analysis.

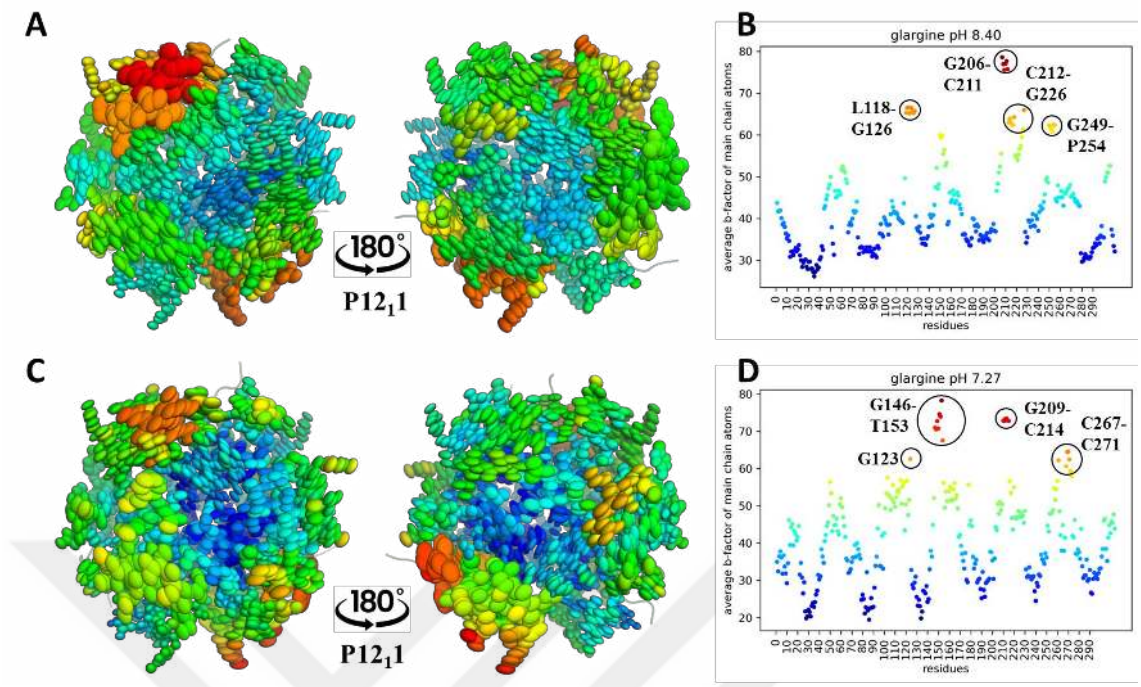


Figure 3.48 Thermal displacement analysis of glargine structures at pH 8.4 and pH 7.27.

Ellipsoid representation (A,C) and 2D trend of b-factor (B,D) have been colored in b-factor spectrum accordingly. The newly assembled undefined and/or disordered electron density peaks have been largely consistent with higher fluctuated residues at thermal displacement analysis.

Thermal ellipsoid and b-factor analysis of glargine structures at pH 8.4 represents that the residues L118-G126, G206-C211, C121-G226, and G249-P254 display higher flexibility than the other sections, suggesting glisin and cysteine have critical importance on oligomeric flexibility as dependent environmental changing and thermal fluctuation. Additionally these critical regions have been consistent with the newly created undefined and/or disordered electron density peaks which includes N18-E, T27-H, G1-I, S9-I residues (Figure 3.48a-b). Similarly, the residues G123, G146-T153, G209-C214, and C267-C271 in glargine structure at pH 7.27 show higher flexibility. Again, glisin and cysteine are critical residues for oligomeric mobility (Figure 3.48c-d). Furthermore, the electron density peak in the pH 7.27 structure has been consistent with thermal ellipsoid and b-factor analyses, including E4-E, N18-E, and G1-I residues (Figure 3.39).

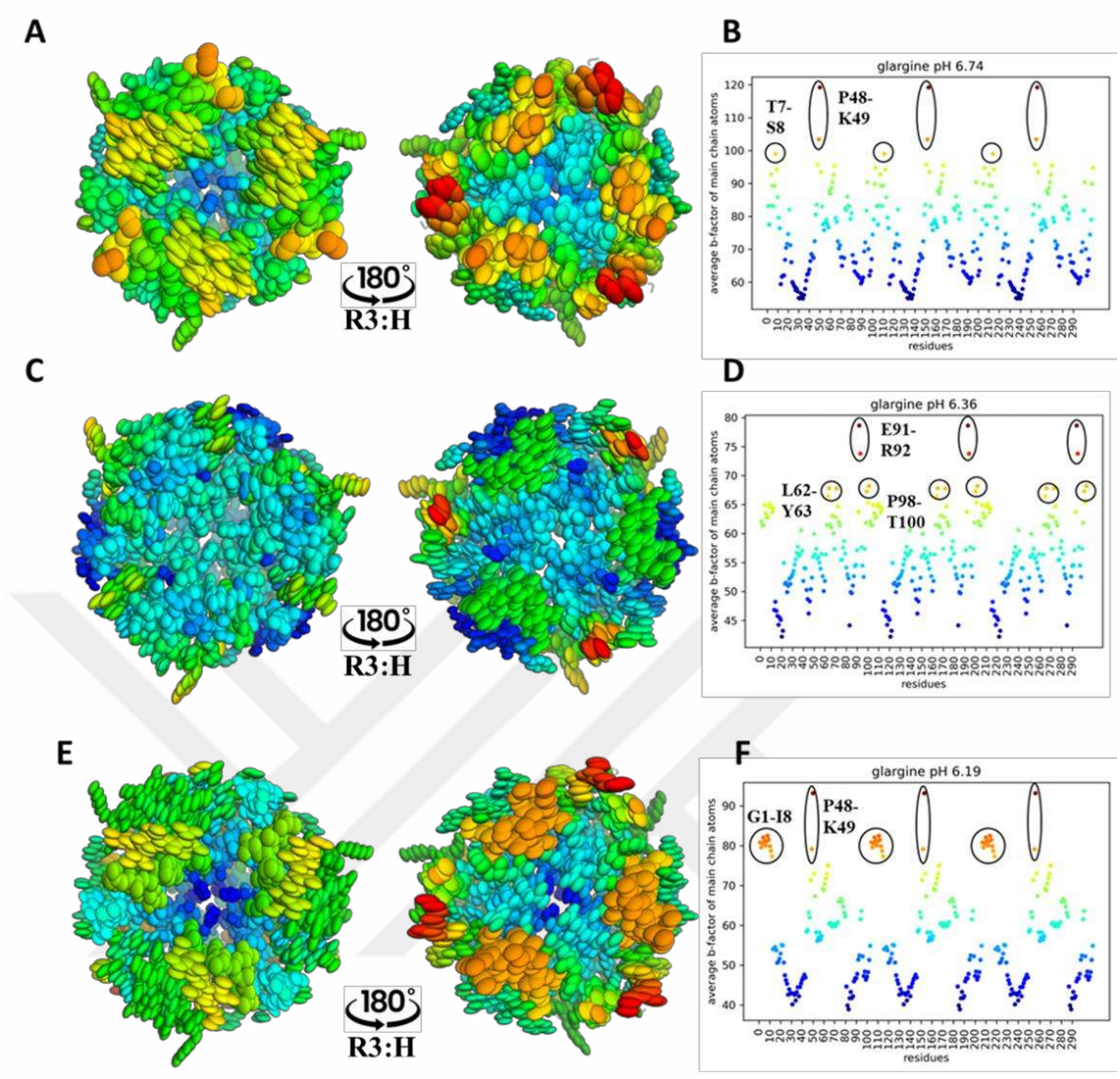


Figure 3.49 Thermal displacement analysis of glargine structures at pH 6.74, pH 6.36, and pH 6.19. Ellipsoid representation (A,C,E) and 2D trend of b-factor (B,D,F) have been colored in b-factor spectrum accordingly. The newly assembled undefined and/or disordered electron density peaks have been largely consistent with higher fluctuated residues at thermal displacement analysis.

The thermal ellipsoid and b-factor analysis of glargine structures at pH 6.74 reveal increased flexibility in specific regions, namely residues T7-S8, P48-K49, and their corresponding counterparts in other monomers (Figure 3.49a-b). Similarly, in the glargine structure at pH 6.36, higher flexibility is observed in residues L62-Y63, E91-R92, and P98-T100, along with identical residues in other monomers (Figure 3.49c-d). Finally, the pH 6.19 structure exhibits elevated flexibility in G1-I8, P48-K49, and corresponding residues in other monomers (Figure 3.49e-f). These findings align with the presence of undefined and disordered electron density peaks, including residues

R31-B, T8-A, G21-A, P28-D, and K29-D (Figure 41-42 and Figure 44)

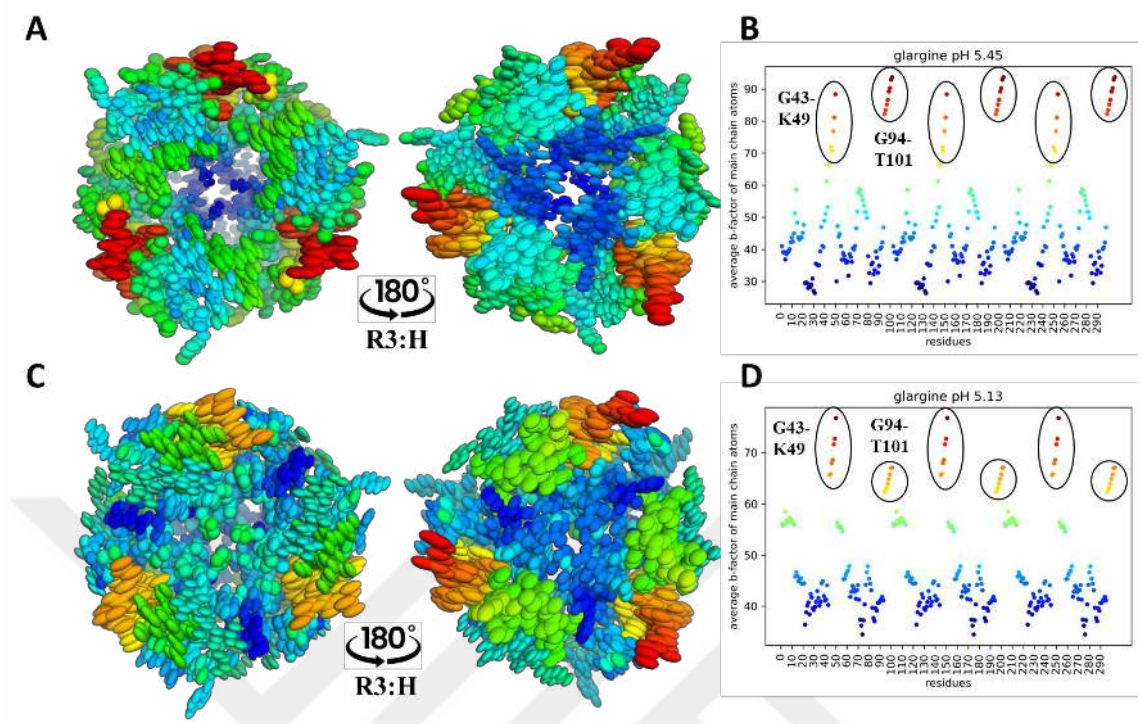


Figure 3.50 Thermal displacement analysis of glargine structures at pH 5.45, and pH 5.13.
Ellipsoid representation (A,C) and 2D trend of b-factor (B,D) have been colored in b-factor spectrum accordingly. The newly assembled undefined and/or disordered electron density peaks have been largely consistent with higher fluctuated residues at thermal displacement analysis.

The thermal ellipsoid and b-factor analyses of glargine structures at pH 5.45 and pH 5.13 indicate a heightened flexibility in specific regions, particularly residues G43-K49 and G94-T101, along with their corresponding counterparts in other monomers. At the lower pH levels, glycine, lysine, and threonine emerge as crucial elements for maintaining increased flexibility within the hexamer (Figure 3.50). These observations are consistent with the existence of undefined and disordered electron density peaks, with residues P28-D and K29-D exhibiting the highest peak (Figure 3.45 and Figure 3.47).

Collectively, we conducted viscous-media-adapted-mix-delay-and-inject SFX, incorporating a minimum 5-minute mixing delay to elucidate the structure of the intermediate states. At pH 8.4 and pH 7.27, the crystal lattice remained unaltered, suggesting no notable conformational changes within this time and pH range. Conversely, a substantial modification in the crystal lattice was observed at a lower pH, indicating significant conformational changes within this specific pH range (Figure

3.51a).

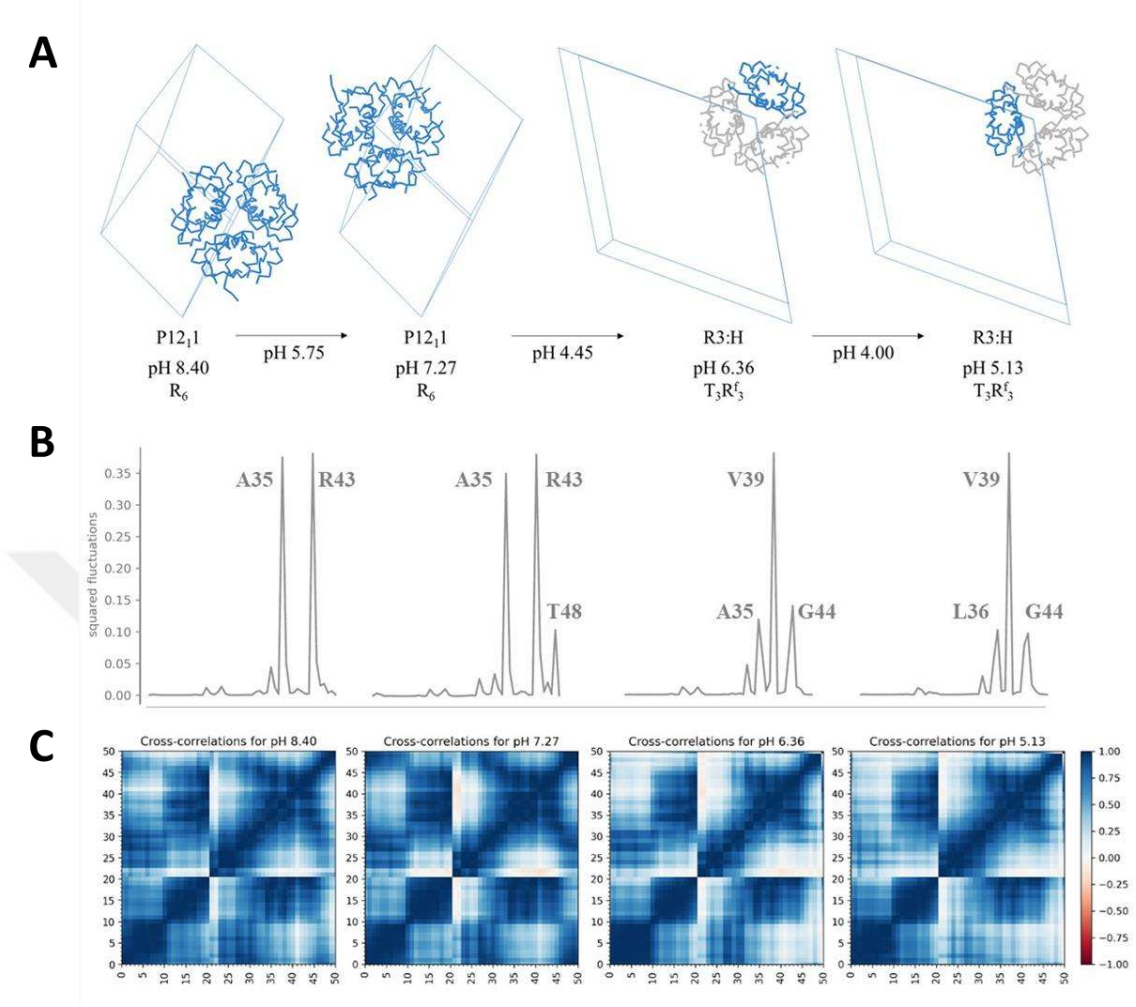


Figure 3.51 Comprehensive interpretation of time-resolved SFX for pH-jumped oligomeric glargine (A) The unit cells of the crystals of glargine structure at pH 8.4 and pH 7.27 ($P12_11$) and pH-jumped glargine structure at pH 6.36 and pH 5.13 (R3:H), the structure of which was generated *in crystallo* from the glargine structure at pH 8.4 after at least 5-min of mixing with the distinctive pH values. The structure at $P12_11$ group has hexamer conformer in the asymmetric unit which is characterized by R_6 allosteric state, whereas the crystal structure at R3:H group has dimer conformer in the asymmetric unit but hexamer in biological unit, which display T_3R_3 allosteric conformation of the insulin. (B) Residues identified as kinetically hot loci, based on the ten fastest modes from the GNM, are A35 and R43 at both pH 8.4 and pH 7.27. These residues undergo a significant replacement, being distinctly substituted by V39 and G44 residues. (C) Cross-correlation from GNM reveals a progressively diminishing correlation, particularly within receptor interaction regions, across decreasing pH values from pH 8.4 to pH 5.13. The calculation for the cross-correlation map utilized the first 50 residues of the hexamer.

Squared fluctuations over the weighted ten fastest modes were calculated for each structure to assess their local motion by comparing them. Residues exhibiting high fluctuations, identified as peak sites in the fastest modes, are considered potential hot

loci, characterized by playing a crucial role in the stability and functionality of the proteins. Notably, certain potential hotspots, such as A35 and R43, consistently appear in the conformational states at pH 8.4, pH 7.27, and pH 6.36 (Figure 3.51b).

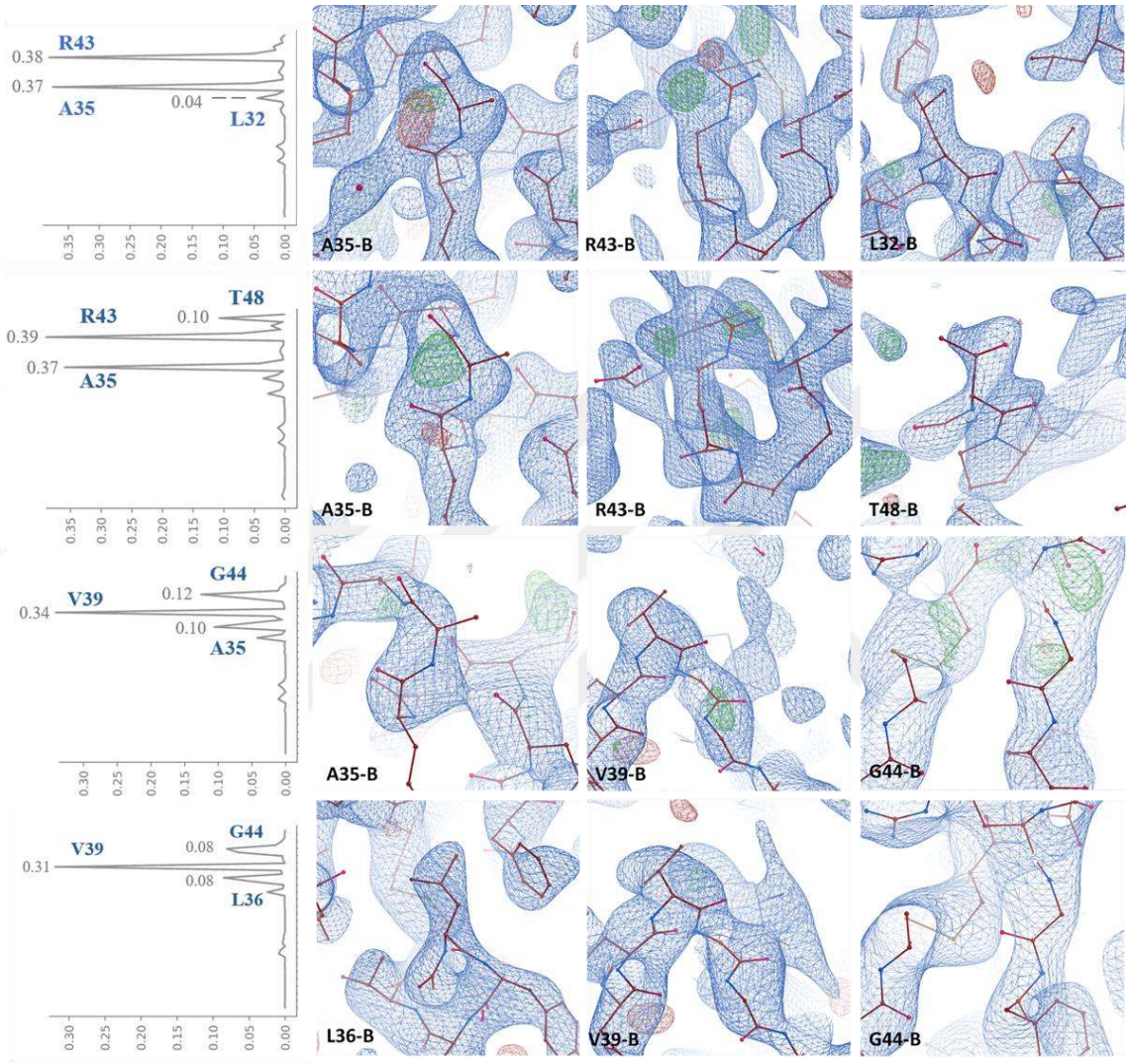


Figure 3.52 Residues identified as kinetically hot loci, based on the ten fastest modes.

The plot (illustrated as plot in left) has been supported by electron density peaks which have been represented in a 2*Fo-Fc* simulated annealing-omit map contoured at 0.23e/A^3.

Remarkably, these residues experience a substantial replacement under decreased pH conditions, with V39, G44, and L36 residues distinctly substituting them. This observation implies that the hotspots within the hexamer undergo a gradual change in response to pH shifts (Figure 3.51b). These kinetically hot regions have been supported by electron density peaks (Figure 3.52). We generated residue–residue cross-correlation maps, as illustrated in Figure 3.51c, utilizing the cumulative ten slowest modes obtained from GNMs. These maps enable us to examine alterations in correlations between

different residue fluctuations within hexamers, mainly focusing on how residue communications evolve during lowered pH intervals. The cross-correlation maps comprehensively compare the four hexamer conformations, specifically in the first 50 residues, highlighting subdomain regions influenced by protein binding. Notably, residues at or near the binding interface exhibit relatively lower correlations during lowered pH intervals, suggesting that the decrease in pH negatively impacts the collective motion of oligomeric insulin over the receptor interaction regions.

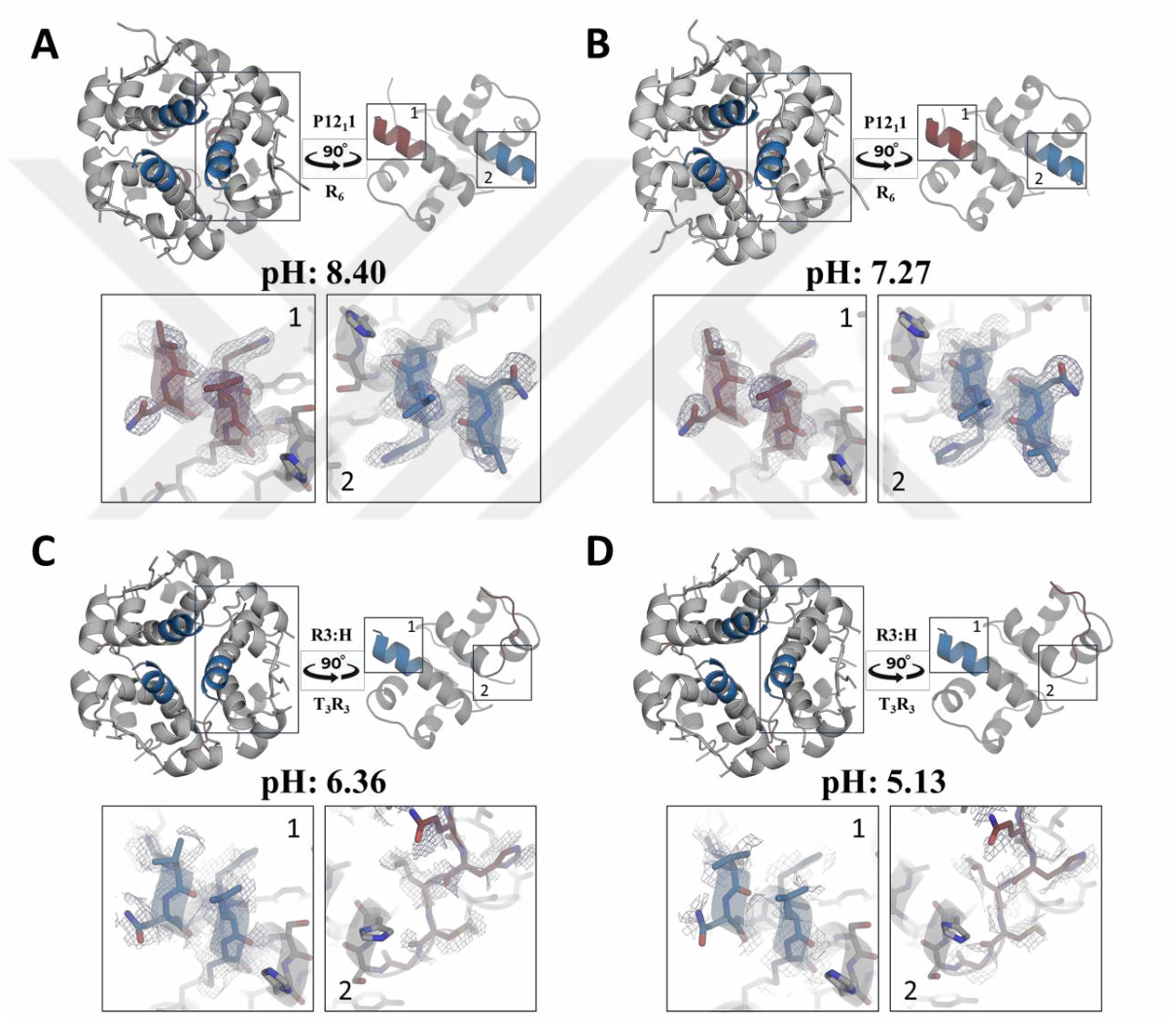


Figure 3.53 pH-dependent allosteric transition of hexamer insulin.

(A,B) The hexamer glargine structure in the $P12_11$ space group at pH 8.4 and pH 7.27 shows the R_6 conformer for each monomer. (C,D) pH-shift induced allosteric transition (T_3R_3) of hexamer glargine to the R3:H space group at pH 6.36 and pH 5.13. Allosteric conformers have been supported by electron density peaks that are contoured at 2*Fo-Fc* simulated annealing-omit map at 1 sigma level.

The time-resolved SFX study has yielded seven distinct glargine structures comprising five rhombohedral forms and two monoclinic forms of the insulin hexamer. These structures are T_3R_3 (formerly 4Zn) and R_6 (formerly 2Zn phenol-induced) species. The nomenclature T_3R_3 and R_6 are assigned to acknowledge their status as interconvertible allosteric states characterized by distinct spectroscopic and chemical properties and varying receptor binding affinities. The structures of the R_6 hexamer have been refined to 2.4 and 2.45 Å resolution, respectively, whereas the T_3R_3 structures have been refined to 2.88 and 3.2 Å resolution. A pH shift induces a change in the space group from $P12_11$ to R3:H, accompanied by an allosteric transition from R_6 to T_3R_3 conformers. The incremental decrease in pH intervals, coupled with the absence of phenol, significantly alters the conformation of the insulin hexamer *in crystallo*. Examination of the SFX structures of the Zn(II) hexamers reveals that in the absence of phenol due to the lower pH, residues B1-B8 in each monomer undergo a conversion from a helix state to an extended chain, causing certain residues to move by as much as 30 Å (Figure 3.53).

Furthermore, we performed cross-correlation analysis between residues B1-8 and the first 50 residues to observe changes in cooperativity during the allosteric transition from R_6 to T_3R_3 oligomeric conformers. Notice that reduced cooperativity is followed toward lower pH intervals. In pH 8.4 and 7.27, it was observed that B1-8 residues exhibit a high correlation with A1-10 and B46-48 residues, whereas a positive correlation was observed only between B1-8 and A1-10 at pH 6.36. However, relatively positive cooperativity was observed between B1-8, A2, and A8 residues, collectively suggesting that lower pH intervals have an impact on the collective motion of oligomeric insulin together with allosteric transitions (Figure 3.54).

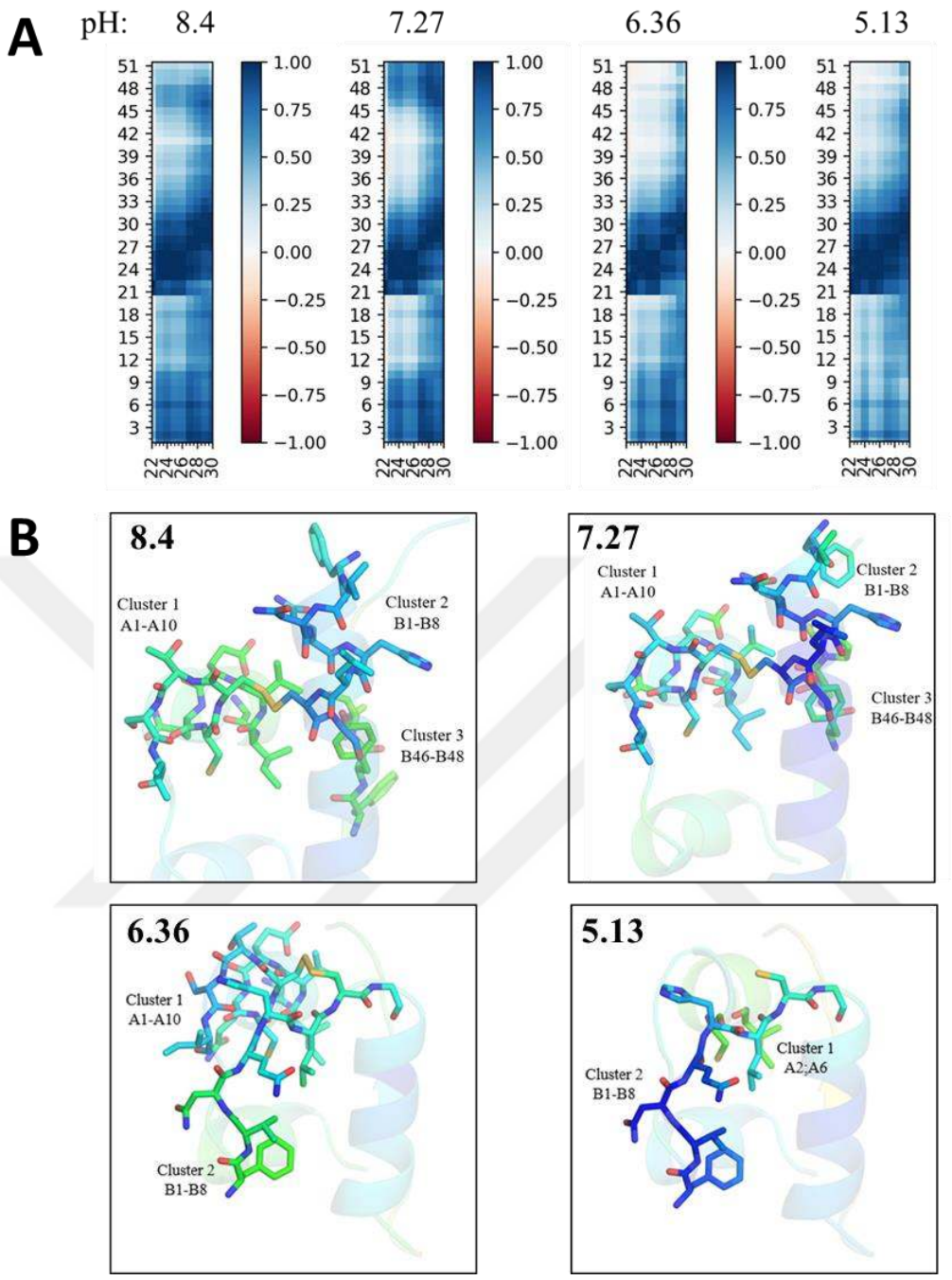


Figure 3.54 Cross-correlation analysis of pH change triggering the allosteric transition from R_6 to T_3R_3 conformers.

(A) Reduced collective motion toward lower pH intervals between B1-8 and other clusters. (B) Cartoon representation of the pH-dependent allosteric conformers has been colored b-factor.

3.4.5. Concluding remarks

Allosteric conformational transitions operate as a mechanism through which proteins convey biochemical information. Nevertheless, instances of obtaining high-

resolution crystal structures for the T- and R-states and their corresponding effector complexes still need to be discovered [Choi et al., 1993]. The insulin hexamer is a notable example within this category. The crystal structures of the Zn(II) hexamers reveal that in the presence of phenol, residues B1-B8 in each monomer are converted from an extended chain to a helix [Rahuel-Clermont et al., 1997]. This hexamer's topology modification establishes six equivalent phenol binding sites, one per monomer [Derewenda et al., 1989]. The binding of each phenol molecule in the monoclinic R_6 structure involves hydrogen bonds between the phenol's hydroxyl group and the carbonyl oxygen of CA6, as well as the amide nitrogen of CA11, along with favorable van der Waals contacts between the phenol ring and several side chain residues [Derewenda et al., 1989]. These hydrophobic pockets are sterically occluded in the T_6 conformation by the LB6 side chain. In the Zn(II)-T_6 hexamer, the two Zn^{2+} ions are coordinated in an octahedral arrangement by the three HB10 imidazole ring nitrogens and three water molecules [Derewenda et al., 1989; Rahuel-Clermont et al., 1997]. In the monoclinic Zn(II)-R_6 hexamer, the two Zn^{2+} ions are coordinated in a tetrahedral arrangement by the same HB10 nitrogens and a Cl^- ion [Derewenda et al., 1989; Brzovic et al., 1994]. The electron density maps of the rhombohedral Zn(II)-R_6 structure indicate the coordination of a phenolate ion with one of the two Zn^{2+} ions [Derewenda et al., 1989]. The acknowledgment of the insulin hexamer functioning as an allosteric protein and the determination of high-resolution crystal structures representing the three conformational states establishes the $T_6 \longleftrightarrow R_6$ transition as an attractive system for exploring the structural basis and mechanism of allosteric transitions [Derewenda et al., 1989; Rahuel-Clermont et al., 1997; Berchtold and Hilgenfeld, 1999].

In this study of this chapter, seven distinct glargine structures were determined at ambient temperature. The experiments applied recording data in a 5-minute time-resolved and pH-jumped method, supplying a high-resolution three-dimensional (3D) observation of pH-dependent allosteric changes ($R_6 \rightarrow T_3R_3$) in glargine at near-physiologic temperature. The resolutions ranged from 2.4 Å to 3.23 Å. Employing 10 fs XFEL pulses at the SACLA, the data collection strategy followed the "diffraction before destruction" principle, minimizing X-ray-induced radiation damage on crystallographic structures. The data collection spanned a 30-hour time slot. Both pH 8.4 and pH 7.27 conformers exhibited an entire hexamer state in the asymmetric unit, sharing identical space groups and unit cell parameters. Interestingly, at pH 6.74, pH 6.36, pH 6.19, pH 5.45, and pH 5.13, the conformers were in a dimer state in the asymmetric unit but

hexameric in the biological unit, resulting in distinct space group and unit cell parameters compared to the previously observed $P12_11$ pH 8.4 and 7.27 structures. Bragg diffraction was still observable but weaker when examining glargine crystals at lower pH levels, aligning with crystal pH screening tests. Data at pH 6.74, 6.36, 6.19, 5.45, and pH 5.13 shows that a transformation indicating sufficiently ordered allosteric alteration that is fully pH-dependent *in crystallo*, stimulating the formation of a new lattice.

In the proximity of residues R31, E4, N18, T27, G1, S9, F1, L17, and E21 of the B chains with noncanonical regions in the $P12_11$ structures, there were corresponding regions predominantly disordered or unidentified. This may be attributed to Brownian motion arising from ambient temperature and pH differences, challenging their interpretation. Near residues R31, T8, G21, F1, N3, C19, H10, P28, and K29 of glargine structures, a noncanonical electron density pattern was also observed within all R3:H structures. These residues in the R3:H structures were associated mainly with receptor-interacting domains exhibiting highly disordered/unidendified electron density, likely originating from newly adopted conformational alterations triggered by pH shifts, resulting in crystal packing rearrangement.

Collectively, this study that yielded seven distinct glargine structures includes five rhombohedral forms and two monoclinic forms of the insulin hexamer. These structures were identified as T_3R_3 and R_6 species, assigned to recognize them as inter-convertible allosteric forms characterized by distinct spectroscopic and chemical properties and altering receptor binding affinities. The R_6 hexamer structures were refined to resolutions of 2.4 and 2.45 Å, while the T_3R_3 structures were refined to resolutions of 2.88 to 3.2 Å. The gradual decrease in pH intervals, associated with the absence of phenol, significantly altered the conformation of the insulin hexamer *in crystallo*. Examination of the SFX structures of the Zn(II) hexamers displayed that, in the absence of phenol due to lower pH, residues B1-B8 in each monomer experienced a conversion from a helix state to an extended chain, causing certain residues to force up to 30 Å (Figure 3.53). Therefore, pH-jump method transformed the space group from P1211 to R3:H, simultaneous with an allosteric transition from R6 to T3R3 conformers. This observation is consistent with established literature, confirming that the P1211 space group corresponds to the R6 conformer, while the R3:H space group indicates the T3R3 conformer throughout crystal rearrangement (Table 3.9).

Table 3.9 Allosteric transition insights into hexameric insulin collection from the PDB.

	Space Group			Space Group			Space Group	
PDB_ID	**$P12_11$**	**R3: H**	**PDB_ID**	**$P12_11$**	**R3: H**	**PDB_ID**	**$P12_11$**	**R3: H**
1BEN	X	T_3R_3	**2QIU**	X	T_3R_3	**4AJX**	R_6	X
1EV6	R_6	X	**2R34**	X	T_3R_3	**4AJZ**	X	T_3R_3
1FU2	X	T_3R_3	**2R35**	X	T_3R_3	**4GBC**	X	T_3R_3
1G7A	X	T_3R_3	**2R36**	X	T_6	**4XC4**	X	T_6
1J73	X	T_3R_3	**2VJZ**	X	T_3R_3	**4P65**	R_6	X
1JCA	X	T_3R_3	**2VK0**	X	T_6	**4RXW**	X	T_6
1LPH	X	T_3R_3	**2WS6**	R_6	X	**5CNY**	X	T_6
1MSO	X	T_6	**2WS7**	R_6	X	**5EMS**	R_6	X
1OS3	X	T_6	**3BXQ**	X	T_3R_3	**5HRQ**	R_6	X
1OS4	X	T_6	**3E7Y**	X	T_6	**5MAM**	X	T_3R_3
1Q4V	X	T_3R_3	**3FQ9**	X	T_6	**5UDP**	R_6	X
1QIZ	R_6	X	**3ILG**	X	T_6	**5UQA**	R_6	X
1QJ0	R_6	X	**3JSD**	X	T_3R_3	**6CK2**	X	T_3R_3
1RWE	X	T_3R_3	**3KQ6**	X	T_3R_3	**6GNQ**	R_6	X
1TRZ	X	T_3R_3	**3P2X**	X	T_3R_3	**6GV0**	X	T_6
1TYL	X	T_3R_3	**3TT8**	X	T_6	**6NWV**	R_6	X
1TYM	X	T_3R_3	**3V19**	X	T_3R_3	**6P4Z**	X	T_6
1W8P	R_6	X	**3W7Y**	X	T_6	**6TYH**	R_6	X
1ZEI	R_6	X	**3WYZ**	X	T_6	**7RKD**	X	T_6
1ZNJ	R_6	X	**3ZQR**	R_6	X	**7S4Y**	X	T_6

`

Chapter 4

CONCLUSION

All studies are structured into two main chapters; Chapter 2 and Chapter 3. The initial chapter addresses structural, dynamic, computational, and *in-vitro* investigations of novel generation ultra-acting insulins designed by the author of this thesis. In Chapter 2, dynamic analyses have been executed using monomeric insulin as the focal point. Chapter 3 provides an in-depth exploration and discourse on the structural, dynamic, kinetic, and computational studies of commercially available multi-oligomeric long-acting insulins. Here, insulin dissociation and association dynamics are detailed both experimentally and computationally through the kinetics of hexameric and dihexameric insulins.

Chapter 2 involves comprehensive upstream, downstream, and advanced analytical characterizations of two highly potent analogs esrapid and esrapid2 insulin analogs. **(i)** To initiate the study, an upstream media optimization experiment for esrapid2 was executed utilizing a multibioreactor in conjunction with the statistical tool, Design Expert software. This study validated the efficiency of each phase by utilizing the high-throughput features of the BioLector micro bioreactor, integrating the DoE software. The effectiveness of this integration was further supported through the application of Placket-Burman Design (PBD) and Central Composite Design (CCD). The primary focus of the investigation revolved around Scenario III, highlighting the capacity of the esrapid2 analog to achieve high cell density fermentation in the *E. coli* Rosetta-2 strain. This highlighted a preference for a more efficient and cost-friendly medium. Accordingly, to enhance the production of recombinant insulin and explore new possibilities for diabetes treatment, it is crucial to develop innovative designer insulin analogs. A key strategy involves creating modified insulin analogs that can be cost-effectively produced as inclusion bodies (IBs) within the *Escherichia coli* BL21

(DE3) Rosetta-2 strain. This production method reduces production time and ensures high recovery rates. The primary objective of our study was to optimize the composition of the cultivation media, aiming for higher cell density fermentation of proinsulin fused with a cleavable hexa histidine purification tag. We systematically investigated various factors such as carbon and nitrogen sources, salts, metal ions, and pH-range using experimental screening with the BioLector multiwell bright plate. Additionally, we employed computational analysis with the Plackett-Burman Design in the Design Expert software to evaluate their impact on insulin concentration, serving as an indicator of insulin yield. From the screened variables, glucose, glycerol, $MgSO_4$, and lower Luria-Bertani mix concentration were identified as significantly influencing insulin production. Subsequently, we employed the Central Composite Design (CCD) approach to further assess and optimize the specific levels of these influential variables. This systematic approach led to the development of an optimized cultivation media formulation, resulting in a remarkable improvement in insulin production, reaching levels of up to 13 mg/ml in BioLector fermentation. The developed cultivation media proved effective in promoting high-density fermentation of the modified insulin, enabling the achievement of optimal proinsulin expression yields. The proinsulin obtained in this study underwent advanced stages of conversion into mature insulin and was subsequently subjected to the process of crystallization.

(ii) To assess the monomer kinetics of esrapid2, we conducted an analytical characterization study employing hydrogen-deuterium exchange mass spectrometry (HDX-MS), X-ray diffraction (XRD), and differential scanning calorimetry (DSC) techniques. This study seeks to uncover the mechanisms driving dynamic changes in the monomer of esrapid2, which is less stable than commercially available insulin aspart. A mutation replacing the arginine residue with the more flexible lysine in the B22 chain was introduced, and its impact on protein stability and electrostatic interactions was investigated, utilizing insulin aspart as the model system. The initial focus was on mutating the surface arginine residue in the B22 chain of insulin aspart to lysine, considering the protein's foldability. Subsequently, the stability of this variant monomer was assessed using the differential scanning calorimetry (DSC) technique. To explore the intrinsic dynamics of the monomer, hydrogen-deuterium exchange Mass Spectrometry (HDX-MS) analysis was employed, and a comparative study was

conducted with the monomer of insulin aspart. Experimental findings were validated through computational analysis. Finally, differences in crystal structures between the hexamer esrapid2 and aspart were determined using the multi-well X-ray crystallography technique conducted at ambient temperature. This research provides significant revelations about the structural characteristics of esrapid2, presenting a high-resolution crystal structure with a 2.5 Å resolution at ambient temperature. A distinct orientation of the NB3 conformation in the ambient temperature structure is disclosed through a comparative analysis with a previously available cryogenic aspart structure (PDB ID: 4GBN). To delve deeper into the molecular architecture of the aspart and esrapid2 insulin analogs, hydrogen-deuterium exchange mass spectrometry (HDX-MS) was applied. HDX dynamics indicate that esrapid2 incorporates deuterium atoms more rapidly and extensively than the intact aspart insulin analog, aligning consistently with the reduced stability observed in our modified analog. These findings are supported by differential scanning calorimetry (DSC) analysis, revealing that esrapid2 is less stable than insulin aspart. Augmented by computational analysis, these results provide unparalleled insights into the structural dynamics of esrapid2 and aspart insulin monomeric entities, along with their intricate intermolecular interactions. Lastly, to evaluate the activity of esrapid2 on cell culture, hippocampal neuron cultures derived from Balb-C mice at postnatal day 3 (newborn) were subjected to a 30-minute incubation with esrapid2 and insulin aspart, followed by Ca^{2+} imaging. Observations indicate that esrapid2 induces calcium stimulation in neurons, mirroring the effect of the commercial aspart analog.

Chapter 2 delves into the functionality, structural and dynamic characterization, and experimental and computational examinations of esrapid insulin analog. **(iii)** 17 mg/ml of the esrapid analog is derived from only 1-liter production, followed by completed upstream, downstream, and crystallization processes. Cryogenic data at a resolution of 1.36 Å were collected from rhombohedral small-scale crystals at the SPring8 facility (RIKEN, Japan). This structure survives in monomeric form in 20 mM citric acid, and it has been observed that the esrapid insulin monomer adopts the T-allosteric form at a pH of 2-3. The obtained esrapid structure was subjected to docking with the insulin receptor, like various pH-level porcine insulin monomers from the PDB. Docking analysis revealed interactions identical to the wt insulin, signifying the

capability of the esrapid analog to interact with the insulin receptor. Furthermore, studies conducted with hiPSC cultures also apparently demonstrate the activity of the esrapid analog on cells. Pairwise and cross-correlation analyses of this structure determined at pH 3 with porcine insulins at other pH values indicate that the insulin conformation at pH 3 is identical to that at pH 9. Notably, the residues critical in hexamerization, such as EB13, were observed to exhibit a similar conformation at pH 3 and pH 9. This observation suggests high and low-pH environments have an identical dissociation effect on insulin. Collectively, these conducted main studies in Chapter 2 illuminated the structural, dynamic and functional characterization of the esrapid insulin analogs.

Chapter 3 focuses on empirical and computational kinetics of the commercial long-acting oligomeric insulins. **(i)** Initially, the cryogenic structure of insulin detemir, predominantly dihexameric, was determined at 1.7 Å resolution at our Turkish Light Source (*Turkish DeLight*) homesource. Although the "dimer of dimer" formation in this structure differs technically from the detemir structure in PDB (1XDA), both share the same space group and unit cell in reciprocal space, generating them technically identical dihexameric structures. However, our "dimer of dimer" structure, processed with a distinct unit cell origin, can also manifest in real space. Therefore, dynamically exploring this distinct "dimer of dimer" conformation in the human body yields profound insights into the oligomeric insulin hierarchy. Consequently, we observed that the identified different conformation of didimer insulin is less collective and more flexible than the conventional didimer structure. We proposed that this conformation is induced by the pH value of its crystallization buffer.

Unraveling the intricate dynamics of the insulin association and dissociation process within the inherent "negative cooperativity" in receptor-insulin binding has posed a formidable challenge. We also **(ii)** presented two distinct X-ray crystallographic structures of the extended-duration human insulin analog in hexameric and dihexameric configurations. These structural conformations were shown for the first time with resolutions of 2.8 Å and 2.7 Å under near-physiological temperature conditions, and the diffraction data were collected at the Turkish Light Source facility.

By applying Normal Mode Analysis (NMA) and Molecular Dynamics (MD) simulations, we focused on these structures' intrinsic intra- and intermolecular conformational dynamics, aiming to uncover disparities between monomeric and oligomeric states. Three primary outcomes surfaced: 1) An examination of the detemir hierarchy using cross-correlation from GNM indicated a boost in dynamic domains during oligomerization. 2) The detemir hierarchy was characterized by heightened polar interactions, revealing a close correlation with regions involved in receptor interaction, dynamic domains, and kinetically significant regions identified through the ten slowest/fastest modes from GNM and MD simulation. 3) Emphasizing the proposition from molecular docking analysis, specific positions intricately engaged with residues essential for insulin receptor binding collectively constituted a secondary molecular domain of paramount significance in receptor binding. Crucially, residues such as C6, Q5, Q15, N18, and C20 in chain A, as well as N24, Q25, H26, L27, L32, A35, L36, R43, and G44 in chain B, were identified as particularly climacteric. These residues, marked by polar interactions, demonstrated pronounced positive cooperativity with residues implicated in insulin receptor binding. This observation prompts us to consider that these functional subdomains, although lacking a direct receptor-binding site, assume the role of secondary domains due to their proximity to receptor regions.

Once investigation the cryogenic and ambient detemir structures, **(iii)** the structures of glargine characterized by tiny crystals were determined at cryogenic and ambient temperatures at the ESRF. One configuration was established at cryogenic temperatures with a precision of 1.95 Å, while the other was determined at room temperature using the SSX technique with a precision of 2.3 Å. A noteworthy finding is the novel observation of three calcium cations binding to the B13 cavity of glargine in hexameric form. Although the existing literature suggests a one-to-one stoichiometry of Ca^{2+} ions per hexamer, implying a solitary Ca^{2+} locus within the cavity, a closer scrutiny of the three-dimensional hexamer structures highlights a significant need for reorganization to create a cohesive region encompassing all six B13 carboxylates. Furthermore, the preliminary isomorphous replacement data sets for Cd^{2+}- and Pb^{2+}-substituted hexamers hint at three closely situated metal ion loci in the EB13 cavity. This study conclusively proves the existence of three Ca^{2+} ions at proximate locations within the three-dimensional hexameric structure, irrespective of temperature

conditions. Additionally, our conducted MD simulations suggest that the hexameric structure at room temperature exhibits more pronounced fluctuations compared to the cryogenic structure. Notably, simulations of E13 residues at both temperatures display significant deviations in certain replicas.

(iv) Another comprehensive study in Chapter 3 focuses on the sophistication of hexamer assembly, structural specificity, and stability in solution, essential for advancing long-acting insulin formulations with therapeutic efficacy. An exploratory investigation employs pH-jumped-time-resolved serial femtosecond crystallography (tr-SFX) to solve the dissociation dynamics of oligomeric insulin structures at various pH stages and near-physiological temperatures. The study uncovers that a pH jump induces substantial changes in space group and unit cell configurations, prompting an allosteric transition that forms distinct crystal lattices. These changes are closely associated with the charge state of insulin and, consequently, the pH values involved. The research delivers seven unique glargine structures, including five rhombohedral forms and two monoclinic forms of the insulin hexamer. These structures, identified as T_3R_3 and R_6 species, represent interconvertible allosteric forms characterized by unique spectroscopic and chemical properties, influencing receptor binding affinities. The R_6 hexamer structures were refined at resolutions of 2.4 and 2.45 Å, while the T_3R_3 structures were refined at resolutions ranging from 2.88 to 3.2 Å. The gradual decrease in pH intervals, combined with the absence of phenol, significantly alters the conformation of the insulin hexamer in crystallo. Examination of the SFX structures of the Zn(II) hexamers reveals that, in the absence of phenol due to lower pH, residues B1-B8 in each monomer undergo a transition from a helix state to an extended chain, causing specific residues to shift up to 30 Å (Figure 3.53). Consequently, the pH-jump method transforms the space group from P1211 to R3:H concurrently with an allosteric transition from R_6 to T_3R_3 conformers. This comprehensive questioning sheds light on the dynamic complexities of insulin structures, providing valuable insights for developing enhanced insulin formulations. Collectively, these conducted main studies in Chapter 3 illuminated the empirical and dynamic characterization of the oligomeric long-acting insulin analogs.

Bibliography

Aaronson, S. A. (1991). Growth factors and cancer. *Science*. doi: 10.1126/science.1659742

Aboul-Soud, M. A. M., Alzahrani, A. J., & Mahmoud, A. (2021). Induced pluripotent stem cells (Ipscs)—roles in regenerative therapies, disease modelling and drug screening. In Cells. doi: 10.3390/cells10092319

Abu, M. L., Nooh, H. M., Oslan, S. N., & Salleh, A. B. (2017). Optimization of physical conditions for the production of thermostable T1 lipase in Pichia guilliermondii strain SO using response surface methodology. *BMC Biotechnology*. doi: 10.1186/s12896-017-0397-7

Acharya, S., Mondal, S., Mukherjee, S., & Bagchi, B. (2021). Rate of Insulin Dimer Dissociation: Interplay between Memory Effects and Higher Dimensionality. *Journal of Physical Chemistry B*. doi: 10.1021/acs.jpcb.1c03779

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., & Adams, P. D. (2012). Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallographica Section D: Biological Crystallography*. doi: 10.1107/S0907444912001308

Akbarian, M., & Yousefi, R. (2018). Human αb-crystallin as fusion protein and molecular chaperone increases the expression and folding efficiency of recombinant insulin. *PLoS ONE*. doi: 10.1371/journal.pone.0206169

Alhmoud, E. N., Saad, M. O., & Omar, N. E. (2023). Efficacy and safety of insulin glargine 300 units/mL vs insulin degludec in patients with type 1 and type 2 diabetes: a systematic review and meta-analysis. *Frontiers in Endocrinology*, *14*. doi: 10.3389/FENDO.2023.1285147

Alshaikh, O. M., Yoon, J. Y., Chan, B. A., Krzyzanowska, M. K., Butany, J., Asa, S. L., & Ezzat, S. (2018). Pancreatic Neuroendocrine Tumor Producing Insulin and Vasopressin. *Endocrine Pathology*. doi: 10.1007/s12022-017-9492-5

Amgarth-Duff, I., Hosie, A., Caplan, G., & Agar, M. (2020). A systematic review of the overlap of fluid biomarkers in delirium and advanced cancer-related syndromes. In BMC Psychiatry. doi: 10.1186/s12888-020-02584-2

Antolíková, E., Žáková, L., Turkenburg, J. P., Watson, C. J., Hančlová, I., Šanda, M., Cooper, A., Kraus, T., Brzozowski, A. M., & Jiráček, J. (2011). Non-equivalent role of inter- and intramolecular hydrogen bonds in the insulin dimer interface. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M111.265249

Apweiler, R., Hermjakob, H., & Sharon, N. (1999). On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica et Biophysica Acta - General Subjects*. doi: 10.1016/S0304-4165(99)00165-8

Arystarkhova, E., Liu, Y. B., Salazar, C., Stanojevic, V., Clifford, R. J., Kaplan, J. H., Kidder, G. M., & Sweadner, K. J. (2013). Hyperplasia of pancreatic beta cells and improved glucose tolerance in mice deficient in the FXYD2 subunit of Na,K-ATPase. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M112.401190

Association, A. D. (2010). Diagnosis and classification of diabetes mellitus. In Diabetes Care (Vol. 33, Issue SUPPL. 1, pp. S62–S69). American Diabetes Association. doi: 10.2337/dc10-S062

Atalay, N., Akcan, E. K., Gül, M., Ayan, E., Destan, E., Ertem, F. B., Tokay, N., Çakilkaya, B., Nergiz, Z., Karakadioğlu, G., Kepceoğlu, A., Yapici, İ., Tosun, B., Baldir, N., Yildirim, G., Johnson, J. A., Güven, Ö., Shafiei, A., Arslan, N. E., … Demirci, H. (2022). Cryogenic X-ray crystallographic studies of biomacromolecules at Turkish Light Source "Turkish DeLight." *BioRxiv*, *10*, 2022.09.03.506456. doi: 10.1101/2022.09.03.506456

Athanasiadou, K. I., Paschou, S. A., Stamatopoulos, T., Papakonstantinou, E., Haidich, A. B., & Goulis, D. G. (2022). Safety and efficacy of insulin detemir versus NPH in the treatment of diabetes during pregnancy: Systematic review and meta-analysis of randomized controlled trials. *Diabetes Research and Clinical Practice*. doi: 10.1016/j.diabres.2022.110020

Atilgan, A. R., & Atilgan, C. (2012). Local motifs in proteins combine to generate global functional moves. *Briefings in Functional Genomics*. doi: 10.1093/bfgp/els027

Atkin, S., Javed, Z., & Fulcher, G. (2015). Insulin degludec and insulin aspart: Novel insulins for the management of diabetes mellitus. In Therapeutic Advances in Chronic Disease. doi: 10.1177/2040622315608646

AU630912B2 - Insulin analogs - Google Patents. (n.d.). Retrieved from https://patents.google.com/patent/AU630912B2/fr

Auer, M. K., Birnbaum, W., Hartmann, M. F., Holterhus, P.-M., Kulle, A., Lux, A., Marshall, L., Rall, K., Richter-Unruh, A., & Werner, R. (2022). *Metabolic Effects of Estradiol Versus Testosterone in Complete Androgen Insensitivity Syndrome*.

Avery, H. (1929). Insulin fat atrophy a traumatic atrophic panniculitis. *British Medical Journal*. doi: 10.1136/bmj.1.3560.597-a

Avgerinos, I., Papanastasiou, G., Karagiannis, T., Michailidis, T., Liakos, A., Mainou, M., Matthews, D. R., Tsapas, A., & Bekiari, E. (2021). Ultra-rapid-acting insulins for adults with diabetes: A systematic review and meta-analysis. *Diabetes, Obesity and Metabolism*. doi: 10.1111/dom.14461

Avital-Shmilovici, M., Mandal, K., Gates, Z. P., Phillips, N. B., Weiss, M. A., & Kent, S. B. H. (2013). Fully convergent chemical synthesis of ester insulin: Determination of the high resolution X-ray structure by racemic protein crystallography. *Journal of the American Chemical Society*. doi: 10.1021/ja311408y

Ayan, E., & DeMirci, H. (2022). A Brief Atlas of Insulin. *Current Diabetes Reviews*, *19*. doi: 10.2174/1573399819666220610150342

Ayers, K., Kumar, R., Robevska, G., Bruell, S., Bell, K., Malik, M. A., Bathgate, R. A., & Sinclair, A. (2019). Familial bilateral cryptorchidism is caused by recessive variants in RXFP2. *Journal of Medical Genetics*. doi: 10.1136/jmedgenet-2019-106203

Back, A., Rossignol, T., Krier, F., Nicaud, J. M., & Dhulster, P. (2016). High-throughput fermentation screening for the yeast Yarrowia lipolytica with real-time monitoring of biomass and lipid production. *Microbial Cell Factories*. doi: 10.1186/s12934-016-0546-z

Bader, E., Migliorini, A., Gegg, M., Moruzzi, N., Gerdes, J., Roscioni, S. S., Bakhti, M.,

Brandl, E., Irmler, M., Beckers, J., Aichler, M., Feuchtinger, A., Leitzinger, C., Zischka, H., Sattler, R. W., Jastroch, M., Tschöp, M., Machicao, F., Staiger, H., … Lickert, H. (2016). Identification of proliferative and mature β-cells in the islets of langerhans. *Nature*. doi: 10.1038/nature18624

Baghbaderani, B. A., Syama, A., Sivapatham, R., Pei, Y., Mukherjee, O., Fellner, T., Zeng, X., & Rao, M. S. (2016). Detailed Characterization of Human Induced Pluripotent Stem Cells Manufactured for Therapeutic Applications. *Stem Cell Reviews and Reports*. doi: 10.1007/s12015-016-9662-8

Bahar, I., Atilgan, A. R., Demirel, M. C., & Erman, B. (1998). Vibrational dynamics of folded proteins: Significance of slow and fast motions in relation to function and stability. *Physical Review Letters*. doi: 10.1103/PhysRevLett.80.2733

Bahar, I., Atilgan, A. R., & Erman, B. (1997). Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Folding and Design*. doi: 10.1016/S1359-0278(97)00024-2

Bahar, I., & Rader, A. J. (2005). Coarse-grained normal mode analysis in structural biology. In Current Opinion in Structural Biology. doi: 10.1016/j.sbi.2005.08.007

Bailey, C. J., & Barnett, A. H. (2007). Why is Exubera being withdrawn? *BMJ*. doi: 10.1136/bmj.39409.507662.94

Bakan, A., Meireles, L. M., & Bahar, I. (2011). ProDy: Protein dynamics inferred from theory and experiments. *Bioinformatics*. doi: 10.1093/bioinformatics/btr168

Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M., Hubbard, R. E., Isaacs, N. W., & Reynolds, C. D. (1988). The structure of 2Zn pig insulin crystals at 1.5 A resolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *319*(1195), 369–456. doi: 10.1098/RSTB.1988.0058

Baker, L. A., Nef, S., Nguyen, M. T., Stapleton, R., Pohl, H., & Parada, L. F. (2002). The insulin-3 gene: Lack of a genetic basis for human cryptorchidism. *Journal of Urology*. doi: 10.1016/S0022-5347(05)65029-X

Banting, F. G., Best, C. H., Collip, J. B., Campbell, W. R., & Fletcher, A. A. (2007). Pancreatic extracts in the treatment of diabetes mellitus. 1922. *The Indian Journal*

of Medical Research. doi: 10.2337/diab.5.1.69

Banting, F. G., Best, C. H., Collip, J. B., Macleod, J. J. R., & Noble, E. C. (1922). THE EFFECT OF PANCREATIC EXTRACT (INSULIN) ON NORMAL RABBITS. *American Journal of Physiology-Legacy Content*. doi: 10.1152/ajplegacy.1922.62.1.162

Bao, S., Wu, Y. L., Wang, X., Han, S., Cho, S. B., Ao, W., & Nan, J. X. (2020). Agriophyllum oligosaccharides ameliorate hepatic injury in type 2 diabetic db/db mice targeting INS-R/IRS-2/PI3K/AKT/PPAR-γ/Glut4 signal pathway. *Journal of Ethnopharmacology*. doi: 10.1016/j.jep.2020.112863

Barbieri, M., Ragno, E., Benvenuti, E., Zito, G. A., Corsi, A., Ferrucci, L., & Paolisso, G. (2001). New aspects of the insulin resistance syndrome: Impact on haematological parameters. *Diabetologia*. doi: 10.1007/s001250100634

Barlow, D. J., & Thornton, J. M. (1983). Ion-pairs in proteins. *Journal of Molecular Biology*. doi: 10.1016/S0022-2836(83)80079-5

Barty, A., Kirian, R. A., Maia, F. R. N. C., Hantke, M., Yoon, C. H., White, T. A., & Chapman, H. (2014). Cheetah: Software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *Journal of Applied Crystallography*. doi: 10.1107/S1600576714007626

Batista, A. F., Forny-Germano, L., Clarke, J. R., Lyra e Silva, N. M., Brito-Moreira, J., Boehnke, S. E., Winterborn, A., Coe, B. C., Lablans, A., Vital, J. F., Marques, S. A., Martinez, A. M. B., Gralle, M., Holscher, C., Klein, W. L., Houzel, J. C., Ferreira, S. T., Munoz, D. P., & De Felice, F. G. (2018). The diabetes drug liraglutide reverses cognitive impairment in mice and attenuates insulin receptor and synaptic pathology in a non-human primate model of Alzheimer's disease. *Journal of Pathology*. doi: 10.1002/path.5056

Beitner, R., & Kalant, N. (1971). Stimulation of glycolysis by insulin. *Journal of Biological Chemistry*. doi: 10.1016/s0021-9258(18)62516-5

Belfiore, A., Frasca, F., Pandini, G., Sciacca, L., & Vigneri, R. (2009). Insulin receptor isoforms and insulin receptor/insulin-like growth factor receptor hybrids in physiology and disease. In Endocrine Reviews. doi: 10.1210/er.2008-0047

Berchtold, H., & Hilgenfeld, R. (1999). Binding of phenol to R6 insulin hexamers. *Biopolymers - Peptide Science Section*. doi: 10.1002/(SICI)1097-0282(1999)51:2<165::AID-BIP6>3.0.CO;2-X

Beringer, D. X., Kleijwegt, F. S., Wiede, F., Van Der Slik, A. R., Loh, K. L., Petersen, J., Dudek, N. L., Duinkerken, G., Laban, S., Joosten, A., Vivian, J. P., Chen, Z., Uldrich, A. P., Godfrey, D. I., McCluskey, J., Price, D. A., Radford, K. J., Purcell, A. W., Nikolic, T., … Rossjohn, J. (2015). T cell receptor reversed polarity recognition of a self-antigen major histocompatibility complex. *Nature Immunology*. doi: 10.1038/ni.3271

Bhandare, R., Schug, J., Lay, J. Le, Fox, A., Smirnova, O., Liu, C., Naji, A., & Kaestner, K. H. (2010). Genome-wide analysis of histone modifications in human pancreatic islets. *Genome Research*. doi: 10.1101/gr.102038.109

Biol, M. C., Lenoir, D., Greco, S., Galvain, D., Hugueny, I., & Louisot, P. (1998). Role of insulin and nutritional factors in intestinal glycoprotein fucosylation during postnatal development. *American Journal of Physiology - Gastrointestinal and Liver Physiology*. doi: 10.1152/ajpgi.1998.275.5.g936

Biological, W. K.-A. C. S. D., & 2010, undefined. (n.d.). xds. *Scripts.Iucr.Org*. Retrieved from https://scripts.iucr.org/cgi-bin/paper?dz5179

Biondi, R. M., Kieloch, A., Currie, R. A., Deak, M., & Alessi, D. R. (2001). The PIF-binding pocket in PDK1 is essential for activation of S6K and SGK, but not PKB. *EMBO Journal*. doi: 10.1093/emboj/20.16.4380

Bisschop, P. H., De Rooij, S. E., Zwinderman, A. H., Van Oosten, H. E., & Van Munster, B. C. (2011). Cortisol, insulin, and glucose and the risk of delirium in older adults with hip fracture. *Journal of the American Geriatrics Society*. doi: 10.1111/j.1532-5415.2011.03575.x

Blackburn, M. (2013). Advances in the quantitation of therapeutic insulin analogues by LC-MS/MS. In Bioanalysis. doi: 10.4155/bio.13.257

Bliss, M. (1982). Banting's, Best's, and Collip's accounts of the discovery of insulin. *Bulletin of the History of Medicine*.

Blundell, T., Dodson, G., Hodgkin, D., & Mercola, D. (1972). Insulin: The structure in

the crystal and its reflection in chemistry and biology. *Advances in Protein Chemistry*. doi: 10.1016/S0065-3233(08)60143-6

Bogardus, C., Lillioja, S., Howard, B. V., Reaven, G., & Mott, D. (1984). Relationships between insulin secretion, insulin action, and fasting plasma glucose concentration in nondiabetic and noninsulin-dependent diabetic subjects. *Journal of Clinical Investigation*. doi: 10.1172/JCI111533

Bolli, G. B., & Owens, D. R. (2000). Insulin glargine. In Lancet. doi: 10.1016/S0140-6736(00)02546-0

Bomfim, T. R., Forny-Germano, L., Sathler, L. B., Brito-Moreira, J., Houzel, J. C., Decker, H., Silverman, M. A., Kazi, H., Melo, H. M., McClean, P. L., Holscher, C., Arnold, S. E., Talbot, K., Klein, W. L., Munoz, D. P., Ferreira, S. T., & De Felice, F. G. (2012). An anti-diabetes agent protects the mouse brain from defective insulin signaling caused by Alzheimer's disease-associated Aβ oligomers. *Journal of Clinical Investigation*. doi: 10.1172/JCI57256

Bonnans, J. F., & Shapiro, A. (2000). Perturbation Analysis of Optimization Problems. In Perturbation Analysis of Optimization Problems. doi: 10.1007/978-1-4612-1394-9

Borders, C. L., Broadwater, J. A., Bekeny, P. A., Salmon, J. E., Lee, A. S., Eldridge, A. M., & Pett, V. B. (1994). A structural role for arginine in proteins: Multiple hydrogen bonds to backbone carbonyl oxygens. In Protein Science. doi: 10.1002/pro.5560030402

Bornfeldt, K. E., & Tabas, I. (2011). Insulin resistance, hyperglycemia, and atherosclerosis. In Cell Metabolism. doi: 10.1016/j.cmet.2011.07.015

Bos, M. M., Noordam, R., Bennett, K., Beekman, M., Mook-Kanamori, D. O., Willems van Dijk, K., Slagboom, P. E., Lundstedt, T., Surowiec, I., & van Heemst, D. (2020). Metabolomics analyses in non-diabetic middle-aged individuals reveal metabolites impacting early glucose disturbances and insulin sensitivity. *Metabolomics*. doi: 10.1007/s11306-020-01653-7

Bosch, R. R., Pouwels, M. J. J. M., Span, P. N., Olthaar, A. J., Tack, C. J., Hermus, A. R. M. M., & Sweep, C. G. J. (2004). Hexosamines are unlikely to function as a nutrient-sensor in 3T3-L1 adipocytes: A comparison of UDP-Hexosamine levels

after increased glucose flux and glucosamine treatment. *Endocrine*. doi: 10.1385/ENDO:23:1:17

Bostan, S. (2020). *Mikro elektrot düzeneği ve optogenetik araçlar kullanılarak hipokampal nöronlarda bellek oluşumunun incelenmesi*. Retrieved from https://acikerisim.medipol.edu.tr/xmlui/handle/20.500.12511/7964

Boura-Halfon, S., & Zick, Y. (2009). Phosphorylation of IRS proteins, insulin action, and insulin resistance. In American Journal of Physiology - Endocrinology and Metabolism. doi: 10.1152/ajpendo.90437.2008

Brackenridge, A., Wallbank, H., Lawrenson, R. A., & Russell-Jones, D. (2006). Emergency management of diabetes and hypoglycaemia. *Emergency Medicine Journal*. doi: 10.1136/emj.2005.026252

Bradley, M. C., Chillarige, Y., Lee, H., Wu, X., Parulekar, S., Muthuri, S., Wernecke, M., Macurdy, T. E., Kelman, J. A., & Graham, D. J. (2021). Severe Hypoglycemia Risk with Long-Acting Insulin Analogs vs Neutral Protamine Hagedorn Insulin. *JAMA Internal Medicine*. doi: 10.1001/jamainternmed.2020.9176

Brange, J. (2012). *Galenics of insulin: the physico-chemical and pharmaceutical aspects of insulin and insulin preparations*. Retrieved from https://books.google.com/books?hl=tr&lr=&id=EGnmCAAAQBAJ&oi=fnd&pg=PA1&dq=).+Galenics+of+Insulin:+The+Physico-Chemical+and+Pharmaceutical+Aspects+of+Insulin+and+Insulin+Preparations&ots=8UtJl23hgK&sig=IgD22mZwuOzYUnqteL0QlmunMLE

Brange, J., & Langkjoer, L. (1993). Insulin structure and stability. In Pharmaceutical biotechnology (Vol. 5, pp. 315–350). Springer, Boston, MA. doi: 10.1007/978-1-4899-1236-7_11

Brange, J., Ribel, U., Hansen, J. F., Dodson, G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Norris, K., Snel, L., Sørensen, A. R., & Voigt, H. O. (1988). Monomeric insulins obtained by protein engineering and their medical implications. *Nature*. doi: 10.1038/333679a0

Bren, A., Park, J. O., Towbin, B. D., Dekel, E., Rabinowitz, J. D., & Alon, U. (2016). Glucose becomes one of the worst carbon sources for E.coli on poor nitrogen sources due to suboptimal levels of cAMP. *Scientific Reports*. doi:

10.1038/srep24834

Brener, A., Sagi, L., Shtamler, A., Levy, S., Fattal-Valevski, A., & Lebenthal, Y. (2020). Insulin-like growth factor-1 status is associated with insulin resistance in young patients with spinal muscular atrophy. *Neuromuscular Disorders*. doi: 10.1016/j.nmd.2020.09.025

Brissova, M., Shiota, M., Nicholson, W. E., Gannon, M., Knobel, S. M., Piston, D. W., Wright, C. V. E., & Powers, A. C. (2002). Reduction in pancreatic transcription factor PDX-1 impairs glucose-stimulated insulin secretion. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M111272200

Bruning, P. F., Bonfrèr, J. M. G., van Noord, P. A. H., Hart, A. A. M., de Jong-Bakker, M., & Nooijen, W. J. (1992). Insulin resistance and breast-cancer risk. *International Journal of Cancer*. doi: 10.1002/ijc.2910520402

Brunner, G. A., Sendlhofer, G., Wutte, A., Ellmerer, M., Søgaard, B., Siebenhofer, A., Hirschberger, S., Krejs, G. J., & Pieber, T. R. (2000). Pharmacokinetic and pharmacodynamic properties of long-acting insulin analogue NN304 in comparison to NPH insulin in humans. *Experimental and Clinical Endocrinology and Diabetes*. doi: 10.1055/s-2000-5887

Brzović, P. S., Choi, W. E., Borchardt, D., Kaarsholm, N. C., & Dunn, M. F. (1994). Structural Asymmetry and Half-Site Reactivity in the T to R Allosteric Transition of the Insulin Hexamer. *Biochemistry*. doi: 10.1021/bi00248a015

Bulek, A. M., Cole, D. K., Skowera, A., Dolton, G., Gras, S., Madura, F., Fuller, A., Miles, J. J., Gostick, E., Price, D. A., Drijfhout, J. W., Knight, R. R., Huang, G. C., Lissin, N., Molloy, P. E., Wooldridge, L., Jakobsen, B. K., Rossjohn, J., Peakman, M., … Sewell, A. K. (2012). Structural basis for the killing of human beta cells by CD8 + T cells in type 1 diabetes. *Nature Immunology*. doi: 10.1038/ni.2206

Cai, W., Sakaguchi, M., Kleinridders, A., Gonzalez-Del Pino, G., Dreyfuss, J. M., O'Neill, B. T., Ramirez, A. K., Pan, H., Winnay, J. N., Boucher, J., Eck, M. J., & Kahn, C. R. (2017). Domain-dependent effects of insulin and IGF-1 receptors on signalling and gene expression. *Nature Communications*. doi: 10.1038/ncomms14892

Cannarella, R., Barbagallo, F., Condorelli, R. A., Aversa, A., La Vignera, S., &

Calogero, A. E. (2019). Osteoporosis from an endocrine perspective: The role of hormonal changes in the elderly. In Journal of Clinical Medicine. doi: 10.3390/jcm8101564

Cannon, M. E., Currin, K. W., Young, K. L., Perrin, H. J., Vadlamudi, S., Safi, A., Song, L., Wu, Y., Wabitsch, M., Laakso, M., Crawford, G. E., & Mohlke, K. L. (2019). Open chromatin profiling in adipose tissue marks genomic regions with functional roles in cardiometabolic traits. *G3: Genes, Genomes, Genetics*. doi: 10.1534/g3.119.400294

Carlsson, A. C., Nowak, C., Lind, L., Östgren, C. J., Nyström, F. H., Sundström, J., Carrero, J. J., Riserus, U., Ingelsson, E., Fall, T., & Ärnlöv, J. (2020). Growth differentiation factor 15 (GDF-15) is a potential biomarker of both diabetic kidney disease and future cardiovascular events in cohorts of individuals with type 2 diabetes: a proteomics approach. *Upsala Journal of Medical Sciences*. doi: 10.1080/03009734.2019.1696430

Carroll, R. J., Hammer, R. E., Chan, S. J., Swift, H. H., Rubenstein, A. H., & Steiner, D. F. (1988). A mutant human proinsulin is secreted from islets of Langerhans in increased amounts via an unregulated pathway. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.85.23.8943

Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., & Woods, R. J. (2005). The Amber biomolecular simulation programs. In Journal of Computational Chemistry. doi: 10.1002/jcc.20290

Cerecedo-Lopez, C. D., Cantu-Aldana, A., Patel, N. J., Aziz-Sultan, M. A., Frerichs, K. U., & Du, R. (2020). Insulin in the Management of Acute Ischemic Stroke: A Systematic Review and Meta-Analysis. *World Neurosurgery*. doi: 10.1016/j.wneu.2020.01.056

Chakrabarty, B., & Parekh, N. (2016). NAPS: Network analysis of protein structures. *Nucleic Acids Research*. doi: 10.1093/nar/gkw383

Chan, C. H., Yu, T. H., & Wong, K. B. (2011). Stabilizing salt-bridge enhances protein thermostability by reducing the heat capacity change of unfolding. *PLoS ONE*. doi:

10.1371/journal.pone.0021624

Chan, S. J., Weiss, J., Konrad, M., White, T., Bahl, C., Yu, S. D., Marks, D., & Steiner, D. F. (1981). Biosynthesis and periplasmic segregation of human proinsulin in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.78.9.5401

Chance, R. E., Kroeff, E. P., Hoffmann, J. A., & Frank, B. H. (1981). Chemical, physical, and biologic properties of biosynthetic human insulin. *Diabetes Care*. doi: 10.2337/diacare.4.2.147

Chapman, T. M., Noble, S., & Goa, K. L. (2002). Insulin aspart: A review of its use in the management of type 1 and 2 diabetes mellitus. In Drugs. doi: 10.2165/00003495-200262130-00014

Chapman, T. M., & Perry, C. M. (2004). Insulin detemir: A review of its use in the management of type 1 and 2 diabetes mellitus. In Drugs. doi: 10.2165/00003495-200464220-00008

Chaudhuri, A., Janicke, D., Wilson, M. F., Tripathy, D., Garg, R., Bandyopadhyay, A., Calieri, J., Hoffmeyer, D., Syed, T., Ghanim, H., Aljada, A., & Dandona, P. (2004). Anti-Inflammatory and Profibrinolytic Effect of Insulin in Acute ST-Segment-Elevation Myocardial Infarction. *Circulation*. doi: 10.1161/01.CIR.0000116762.77804.FC

Chevenne, D., Trivin, F., & Porquet, D. (1999). Insulin assays and reference values. In Diabetes and Metabolism.

Chiang, S. H., Baumann, C. A., Kanzaki, M., Thurmond, D. C., Watson, R. T., Neudauer, C. L., Macara, I. G., Pessin, J. E., & Saltiel, A. R. (2001). Insulin-stimulated GLUT4 translocation requires the CAP-dependent activation of TC10. *Nature*. doi: 10.1038/35073608

Chinai, J. M., Taylor, A. B., Ryno, L. M., Hargreaves, N. D., Morris, C. A., Hart, P. J., & Urbach, A. R. (2011). Molecular recognition of insulin by a synthetic receptor. *Journal of the American Chemical Society*. doi: 10.1021/ja201581x

Choi, W. E., Brader, M. L., Aguilar, V., Dunn, M. F., & Kaarsholm, N. C. (1993). The Allosteric Transition of the Insulin Hexamer Is Modulated by Homotropic and

Heterotropic Interactions. *Biochemistry*. doi: 10.1021/bi00094a021

Chothia, C., Lesk, A. M., Dodson, G. G., & Hodgkin, D. C. (1983). Transmission of conformational change in insulin. *Nature*. doi: 10.1038/302500a0

Ciszak, E., Beals, J. M., Frank, B. H., Baker, J. C., Carter, N. D., & Smith, G. D. (1995). Role of C-terminal B-chain residues in insulin assembly: the structure of hexameric LysB28ProB29-human insulin. *Structure*. doi: 10.1016/S0969-2126(01)00195-2

Ciszak, E., & Smith, G. D. (1994). Crystallographic Evidence for Dual Coordination around Zinc in the T3R3 Human Insulin Hexamer. *Biochemistry*. doi: 10.1021/bi00172a030

Clavellina, D., Balkan, W., & Hare, J. M. (2023). Stem cell therapy for acute myocardial infarction: Mesenchymal Stem Cells and induced Pluripotent Stem Cells. In Expert Opinion on Biological Therapy. doi: 10.1080/14712598.2023.2245329

CMVohl, A. Dh. Jnk. Ml. (2000). The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet*, *26*, 7680.

CN103145829B - Purification method of insulin detemir - Google Patents. (n.d.). Retrieved from https://patents.google.com/patent/CN103145829B/en

Coffman, F. D., & Dunn, M. F. (1988). Insulin—Metal Ion Interactions: The Binding of Divalent Cations to Insulin Hexamers and Tetrarners and the Assembly of Insulin Hexamers. *Biochemistry*. doi: 10.1021/bi00416a053

Cohen-Forterre, L., Andre, J., Mozere, G., Peyroux, J., & Sternberg, M. (1990). Kidney sialidase and sialyltransferase activities in spontaneously and experimentally diabetic rats. Influence of insulin and sorbinil treatments. *Biochemical Pharmacology*. doi: 10.1016/0006-2952(90)90549-Z

Cohen, P., Nimmo, H. G., & Proud, C. G. (1978). How does insulin stimulate glycogen synthesis? In Biochemical Society Symposia.

Cole, D. K., Bulek, A. M., Dolton, G., Schauenberg, A. J., Szomolay, B., Rittase, W., Trimby, A., Jothikumar, P., Fuller, A., Skowera, A., Rossjohn, J., Zhu, C., Miles, J. J., Peakman, M., Wooldridge, L., Rizkallah, P. J., & Sewell, A. K. (2016). Hotspot

autoimmune T cell receptor binding underlies pathogen and insulin peptide cross-reactivity. *Journal of Clinical Investigation*. doi: 10.1172/JCI85679

Consortium, W. T. C. C., Burton, P. R., Donnelly, P., Cardon, L. R., Clayton, D. G., Todd, J. A., Ouwehand, W. H., Craddock, N., Tobin, M. D., Samani, N. J., Nutland, S., Stevens, H. E., Walker, N. M., McCarthy, M. I., Newport, M., Sirugo, G., Kwiatkowski, D. P., Deloukas, P., Leung, H.-T., … Worthington, J. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*.

Corica, G., Ceraudo, M., Campana, C., Nista, F., Cocchiara, F., Zona, G., Criminelli, D., Ferone, D., & Gatto, F. (2020). Octreotide-resistant acromegaly: Challenges and solutions. In Therapeutics and Clinical Risk Management. doi: 10.2147/TCRM.S183360

Correction: Structure of a human insulin peptide–HLA-DQ8 complex and susceptibility to type 1 diabetes. (2001). *Nature Immunology*. doi: 10.1038/ni0901-889

Correction for De Felice et al., Protection of synapses against Alzheimer's-linked toxins: Insulin signaling prevents the pathogenic binding of Aβ oligomers. (2009). *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.0902058106

Cotton, F., Wilkinson, G., Murillo, C., & Bochmann, M. (1999). *Advanced inorganic chemistry*. Retrieved from https://books.google.com/books?hl=tr&lr=&id=NFHNEAAAQBAJ&oi=fnd&pg=PR11&dq=Cotton,+F.+A.,+Wilkinson,+G.,+Murillo,+C.+A.,+%26+Bochmann,+M.+(1999).+Advanced+inorganic+chemistry.+John+Wiley+%26+Sons.&ots=xjL-bdPrfH&sig=M3wz0d_WWhXvpuTbMEC7aNtkvB4

Crocker, P. R. (2002). Siglecs: Sialic-acid-binding immunoglobulin-like lectins in cell-cell interactions and signalling. In Current Opinion in Structural Biology. doi: 10.1016/S0959-440X(02)00375-5

Cronin, C. C., & Shanahan, F. (1997). Insulin-dependent diabetes mellitus and coeliac disease. In Lancet. doi: 10.1016/S0140-6736(96)09153-2

CrysAlisPro, V. (2014). CrysAlisPro. - Google Akademik. (n.d.). Retrieved from https://scholar.google.com/scholar?hl=tr&as_sdt=0%2C5&q=CrysAlisPro%2C+V.+%282014%29.+CrysAlisPro.&btnG=

Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Emdin, S. F., & Reynolds, C. D. (1979). Structure and biological activity of hagfish insulin. *Journal of Molecular Biology*. doi: 10.1016/0022-2836(79)90497-2

Dal, O. H., & Sezer, A. D. (2017). Türkiye'de İnsüline Ait Satış Verileri ve Referans İnsülin Aspart ile Biyobenzerinin Analitik Karşılaştırılması. Marmara Pharmaceutical Journal, 21(3).

Daugherty, K. K. (2004). Review of Insulin Therapy. In Journal of Pharmacy Practice. doi: 10.1177/0897190003261304

Davidson, H. W., Wenzlau, J. M., & O'Brien, R. M. (2014). Zinc transporter 8 (ZnT8) and β cell function. In Trends in Endocrinology and Metabolism. doi: 10.1016/j.tem.2014.03.008

De Franco, E., Saint-Martin, C., Brusgaard, K., Knight Johnson, A. E., Aguilar-Bryan, L., Bowman, P., Arnoux, J. B., Larsen, A. R., May, S., Greeley, S. A. W., Calzada-León, R., Harman, B., Houghton, J. A. L., Nishimura-Meguro, E., Laver, T. W., Ellard, S., del Gaudio, D., Christesen, H. T., Bellanné-Chantelot, C., & Flanagan, S. E. (2020). Update of variants identified in the pancreatic β-cell KATP channel genes KCNJ11 and ABCC8 in individuals with congenital hyperinsulinism and diabetes. *Human Mutation*. doi: 10.1002/humu.23995

De Jesus, D. F., & Kulkarni, R. N. (2014). Epigenetic modifiers of islet function and mass. In Trends in Endocrinology and Metabolism. doi: 10.1016/j.tem.2014.08.006

De Jesus, D. F., & Kulkarni, R. N. (2019). "Omics" and "epi-omics" underlying the β-cell adaptation to insulin resistance. In Molecular Metabolism. doi: 10.1016/j.molmet.2019.06.003

de la Monte, S. M. (2009). Insulin resistance and Alzheimer's disease. In BMB Reports (Vol. 42, Issue 8, pp. 475–481). The Biochemical Society of the Republic of Korea. doi: 10.5483/BMBRep.2009.42.8.475

De La Monte, S. M., & Wands, J. R. (2005). Review of insulin and insulin-like growth factor expression, signaling, and malfunction in the central nervous system: Relevance to Alzheimer's disease. In Journal of Alzheimer's Disease. doi: 10.3233/JAD-2005-7106

De Vries, S. J., Van Dijk, M., & Bonvin, A. M. J. J. (2010). The HADDOCK web server for data-driven biomolecular docking. *Nature Protocols*. doi: 10.1038/nprot.2010.32

DeFelippis, M. R., Chance, R. E., & Frank, B. H. (2001). Insulin self-association and the relationship to pharmacokinetics and pharmacodynamics. In Critical Reviews in Therapeutic Drug Carrier Systems. doi: 10.1615/critrevtherdrugcarriersyst.v18.i2.30

DeFronzo, R. A., & Tripathy, D. (2009). Skeletal muscle insulin resistance is the primary defect in type 2 diabetes. In Diabetes care. doi: 10.2337/dc09-s302

Delano, W. L. (2002). The PyMOL Molecular Graphics System. *CCP4 Newsletter on Protein Crystallography*.

Dennis, J. W., Granovsky, M., & Warren, C. E. (1999). Protein glycosylation in development and disease. In BioEssays. doi: 10.1002/(SICI)1521-1878(199905)21:5<412::AID-BIES8>3.0.CO;2-5

Dentener, A. (2002). *Design Expert DOEs it better*. Retrieved from www.statease.com

Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X., & Markussen, J. (1991). X-ray analysis of the single chain B29-A1 peptide-linked insulin molecule. A completely inactive analogue. *Journal of Molecular Biology*. doi: 10.1016/0022-2836(91)90022-X

Derewenda, U., Derewenda, Z., Dodson, G. G., Hubbard, R. E., & Korber, F. (1989). Molecular structure of insulin: The insulin monomer and its assembly. *British Medical Bulletin*. doi: 10.1093/oxfordjournals.bmb.a072320

Desrosiers, R., Friderici, K., & Rottman, F. (1974). Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.71.10.3971

Devis-Jauregui, L., Eritja, N., Davis, M. L., Matias-Guiu, X., & Llobet-Navàs, D. (2021). Autophagy in the physiological endometrium and cancer. In Autophagy. doi: 10.1080/15548627.2020.1752548

Dharmadi, Y., Murarka, A., & Gonzalez, R. (2006). Anaerobic fermentation of glycerol

by Escherichia coli: A new platform for metabolic engineering. *Biotechnology and Bioengineering*. doi: 10.1002/bit.21025

Dhayalan, B., Mandal, K., Rege, N., Weiss, M. A., Eitel, S. H., Meier, T., Schoenleber, R. O., & Kent, S. B. H. (2017). Scope and Limitations of Fmoc Chemistry SPPS-Based Approaches to the Total Synthesis of Insulin Lispro via Ester Insulin. *Chemistry - A European Journal*. doi: 10.1002/chem.201605578

Di Camillo, B., Carlon, A., Eduati, F., & Toffolo, G. M. (2016). A rule-based model of insulin signalling pathway. *BMC Systems Biology*. doi: 10.1186/s12918-016-0281-4

Diao, J. (2003). Crystallographic titration of cubic insulin crystals: pH affects GluB13 switching and sulfate binding. *Acta Crystallographica - Section D Biological Crystallography*. doi: 10.1107/S0907444903002208

Docherty, K., & Hutton, J. C. (1983). Carboxypeptidase activity in the insulin secretory granule. *FEBS Letters*. doi: 10.1016/0014-5793(83)81065-5

Dodson, E. J., Dodson, G. G., Hubbard, R. E., & Reynolds, C. D. (1983). Insulin's structural behavior and its relation to activity. *Biopolymers*. doi: 10.1002/bip.360220137

Dodson, G., & Steiner, D. (1998). The role of assembly in insulin's biosynthesis. *Current Opinion in Structural Biology*. doi: 10.1016/S0959-440X(98)80037-7

Donald, J. E., Kulp, D. W., & DeGrado, W. F. (2011). Salt bridges: Geometrically specific, designable interactions. *Proteins: Structure, Function and Bioinformatics*. doi: 10.1002/prot.22927

Donner, T., & Sarkar, S. (2019). Insulin – Pharmacology, Therapeutic Regimens And Principles Of Intensive Insulin Therapy. In Endotext.

Dorrell, C., Schug, J., Canaday, P. S., Russ, H. A., Tarlow, B. D., Grompe, M. T., Horton, T., Hebrok, M., Streeter, P. R., Kaestner, K. H., & Grompe, M. (2016). Human islets contain four distinct subtypes of β cells. *Nature Communications*. doi: 10.1038/ncomms11756

Drinkwater, J. J., Davis, T. M. E., & Davis, W. A. (2020). The relationship between carotid disease and retinopathy in diabetes: A systematic review. In Cardiovascular

Diabetology. doi: 10.1186/s12933-020-01023-6

Drummen, N. (n.d.). Aerobic cultivation of high-oxygen-demanding microorganisms in the BioLector XT microbioreactor. *Media.Beckman.ComN Drummenmedia.Beckman.Com*. Retrieved from https://media.beckman.com/-/media/app-note-support-items/anaerobic-cultivation-processes-in-the-biolector-xt/high-oxygen-appnote-bl-xt.pdf?la=en-us&hash=B170EF4879CCE7D0F2513E682CD5CC477146874B

Duan, M., Wang, Y., Yang, G., Li, J., Wan, Y., Deng, Y., & Mao, Y. (2020). High-level production of γ-cyclodextrin glycosyltransferase in recombinant Escherichia coli BL21 (DE3): culture medium optimization, enzymatic properties characterization, and product specificity analysis. *Annals of Microbiology*. doi: 10.1186/s13213-020-01610-8

Duisenberg, A. J. M. (1992). Indexing in single-crystal diffractometry with an obstinate list of reflections. *Journal of Applied Crystallography*. doi: 10.1107/S0021889891010634

Dunn, C. J., Plosker, G. L., Keating, G. M., McKeage, K., & Scott, L. J. (2003). Insulin glargine: An updated review of its use in the management of diabetes mellitus. In Drugs. doi: 10.2165/00003495-200363160-00007

Dunn, M. F. (2005). Zinc-ligand interactions modulate assembly and stability of the insulin hexamer - A review. *BioMetals*. doi: 10.1007/s10534-005-3685-y

Durdagi, S., Dag, C., Dogan, B., Yigin, M., Avsar, T., Buyukdag, C., Erol, I., Ertem, B., Calis, S., Yildirim, G., Orhan, M. D., Guven, O., Aksoydan, B., Destan, E., Sahin, K., Besler, S. O., Oktay, L., Shafiei, A., Tolu, I., … DeMirci, H. (2020). Near-Physiological-Temperature Serial Femtosecond X-ray Crystallography Reveals Novel Conformations of SARS-CoV-2 Main Protease Active Site for Improved Drug Repurposing. *BioRxiv*. doi: 10.1101/2020.09.09.287987

Dwek, R. A. (1995). Glycobiology: “Towards understanding the function of sugars.” *Biochemical Society Transactions*. doi: 10.1042/bst0230001

Eckel, R. H., Alberti, K. G. M. M., Grundy, S. M., & Zimmet, P. Z. (2010). The metabolic syndrome. In The Lancet. doi: 10.1016/S0140-6736(09)61794-3

Edgerton, D. S., Moore, M. C., Winnick, J. J., Scott, M., Farmer, B., Naver, H., Jeppesen, C. B., Madsen, P., Kjeldsen, T. B., Nishimura, E., Brand, C. L., & Cherrington, A. D. (2014). Changes in glucose and fat metabolism in response to the administration of a hepato-preferential insulin analog. *Diabetes*. doi: 10.2337/db14-0266

Eheim, A., Medrikova, D., & Herzig, S. (2014). Immune cells and metabolic dysfunction. In Seminars in Immunopathology. doi: 10.1007/s00281-013-0403-7

El Hage, K., Pandyarajan, V., Phillips, N. B., Smith, B. J., Menting, J. G., Whittaker, J., Lawrence, M. C., Meuwly, M., & Weiss, M. A. (2016). Extending halogen-based medicinal chemistry to proteins: Iodo-insulin as a case study. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M116.761015

Elhadad, M. A., Jonasson, C., Huth, C., Wilson, R., Gieger, C., Matias, P., Grallert, H., Graumann, J., Gailus-Durner, V., Rathmann, W., Toerne, C. von, Hauck, S. M., Koenig, W., Sinner, M. F., Oprea, T. I., Suhre, K., Thorand, B., Hveem, K., Peters, A., & Waldenberger, M. (2020). Deciphering the plasma proteome of type 2 diabetes. *Diabetes*. doi: 10.2337/db20-0296

Elibol, M. (2004). Optimization of medium composition for actinorhodin production by Streptomyces coelicolor A3(2) with response surface methodology. *Process Biochemistry*. doi: 10.1016/S0032-9592(03)00232-2

Ellis, B., & Arsiwalla, D. (2020). The Moderating Role of Anxiety in the Associations between Sleep and Insulin Resistance. *CSBS INSPIRE Student Research & Engagement Conference*. Retrieved from https://scholarworks.uni.edu/csbsresearchconf/2020/all/73

Emdin, S. O., Dodson, G. G., Cutfield, J. M., & Cutfield, S. M. (1980). Role of zinc in insulin biosynthesis. *Diabetologia*. doi: 10.1007/bf00275265

Emsley, P., & Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*. doi: 10.1107/S0907444904019158

Esko, J. D., & Selleck, S. B. (2002). Order out of chaos: Assembly of ligand binding sites in heparan sulfate. In Annual Review of Biochemistry. doi: 10.1146/annurev.biochem.71.110601.135458

Evans, P. (2014). CCP4 Tools for X-ray Integration and Data Processing. *Nihon Kessho Gakkaishi*. doi: 10.5940/jcrsj.56.s24

Evans, P. R. (2011). An introduction to data reduction: Space-group determination, scaling and intensity statistics. *Acta Crystallographica Section D: Biological Crystallography*. doi: 10.1107/S090744491003982X

Fadista, J., Vikman, P., Laakso, E. O., Mollet, I. G., Esguerra, J. Lou, Taneera, J., Storm, P., Osmark, P., Ladenvall, C., Prasad, R. B., Hansson, K. B., Finotello, F., Uvebrant, K., Ofori, J. K., Di Camillo, B., Krus, U., Cilio, C. M., Hansson, O., Eliasson, L., … Groop, L. (2014). Global genomic and transcriptomic analysis of human pancreatic islets reveals novel genes influencing glucose metabolism. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1402665111

Faheem, A., Rehman, K., Jabeen, K., & Akash, M. S. H. (2020). Nicotine-mediated upregulation of microRNA-141 expression determines adipokine-intervened insulin resistance. *Environmental Toxicology and Pharmacology*. doi: 10.1016/j.etap.2020.103506

Fávero-Retto, M. P., Palmieri, L. C., Souza, T. A. C. B., Almeida, F. C. L., & Lima, L. M. T. R. (2013). Structural meta-analysis of regular human insulin in pharmaceutical formulations. *European Journal of Pharmaceutics and Biopharmaceutics*. doi: 10.1016/j.ejpb.2013.05.005

Feng, J., Park, J., Cron, P., Hess, D., & Hemmings, B. A. (2004). Identification of a PKB/Akt hydrophobic motif Ser-473 kinase as DNA-dependent protein kinase. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M406731200

Ferreira, A. P., Oliveira, C. E. R., & França, N. M. (2007). Metabolic syndrome and risk factors for cardiovascular disease in obese children: the relationship with insulin resistance (HOMA-IR). *Jornal de Pediatria*. doi: 10.2223/jped.1562

Ferreira, L. S. S., Fernandes, C. S., Vieira, M. N. N., & De Felice, F. G. (2018). Insulin resistance in Alzheimer's disease. In Frontiers in Neuroscience. doi: 10.3389/fnins.2018.00830

Fishel Bartal, M., Ward, C., Refuerzo, J. S., Ashimi, S. S., Joycelyn, C. A., Chen, H. Y., Chauhan, S. P., & Sibai, B. M. (2020). Basal insulin analogs versus neutral

protamine hagedorn for type 2 diabetics. *American Journal of Perinatology*. doi: 10.1055/s-0039-1694733

Flitsch, D., Krabbe, S., Ladner, T., Beckers, M., Schilling, J., Mahr, S., Conrath, U., Schomburg, W. K., & Büchs, J. (2016). Respiration activity monitoring system for any individual well of a 48-well microtiter plate. *Journal of Biological Engineering*. doi: 10.1186/s13036-016-0034-3

Fontbonne, A. M., & Eschwege, E. M. (1991). Insulin and cardiovascular disease: Paris Prospective Study. In Diabetes Care. doi: 10.2337/diacare.14.6.461

Fralick, M., & Zinman, B. (2021). The discovery of insulin in Toronto: beginning a 100 year journey of research and clinical achievement. In Diabetologia. doi: 10.1007/s00125-020-05371-6

Frederickson, C. J., Koh, J. Y., & Bush, A. I. (2005). The neurobiology of zinc in health and disease. In Nature Reviews Neuroscience. doi: 10.1038/nrn1671

Fredericq, E. (1953). Reversible dissociation of insulin [10]. In Nature. doi: 10.1038/171570a0

Frost, R. J. A., & Olson, E. N. (2011). Control of glucose homeostasis and insulin sensitivity by the Let-7 family of microRNAs. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1118922109

Fu, Z., R Gilbert, E., & Liu, D. (2013). Regulation of insulin synthesis and secretion and pancreatic Beta-cell dysfunction in diabetes. *Current Diabetes Reviews*, *9*(1), 25–53.

Fujino, M., Wakimasu, M., Taketomi, S., & Iwatsuka, H. (1977). Insulin-like activities and insulin-potentiating actions of a modified insulin B21–26 fragment: β-Ala-Arg-Gly-Phe-Phe-Tyr-NH2. *Endocrinology*. doi: 10.1210/endo-101-2-360

Fullerton, B., Siebenhofer, A., Jeitler, K., Horvath, K., Semlitsch, T., Berghold, A., & Gerlach, F. M. (2018). Short-acting insulin analogues versus regular human insulin for adult, non-pregnant persons with type 2 diabetes mellitus. In Cochrane Database of Systematic Reviews. doi: 10.1002/14651858.CD013228

Fullerton, B., Siebenhofer, A., Jeitler, K., Horvath, K., Semlitsch, T., Berghold, A.,

Plank, J., Pieber, T. R., & Gerlach, F. M. (2016). Short-acting insulin analogues versus regular human insulin for adults with type 1 diabetes mellitus. In Cochrane Database of Systematic Reviews. doi: 10.1002/14651858.CD012161

Galloway, J. A., & Chance, R. E. (1994). Improving insulin therapy: Achievements and challenges. *Hormone and Metabolic Research*. doi: 10.1055/s-2007-1001766

Gao, T., McKenna, B., Li, C., Reichert, M., Nguyen, J., Singh, T., Yang, C., Pannikar, A., Doliba, N., Zhang, T., Stoffers, D. A., Edlund, H., Matschinsky, F., Stein, R., & Stanger, B. Z. (2014). Pdx1 maintains β cell identity and function by repressing an α cell program. *Cell Metabolism*. doi: 10.1016/j.cmet.2013.12.002

Gao, W., Guo, N., Zhao, S., Chen, Z., Zhang, W., Yan, F., Liao, H., & Chi, K. (2020). Carboxypeptidase A4 promotes cardiomyocyte hypertrophy through activating PI3K-AKT-mTOR signaling. *Bioscience Reports*. doi: 10.1042/BSR20200669

Garg, S. K., Ellis, S. L., & Ulrich, H. (2005). Insulin glulisine: A new rapid-acting insulin analogue for the treatment of diabetes. In Expert Opinion on Pharmacotherapy. doi: 10.1517/14656566.6.4.643

Garneau, L., & Aguer, C. (2019). Role of myokines in the development of skeletal muscle insulin resistance and related metabolic defects in type 2 diabetes. In Diabetes and Metabolism. doi: 10.1016/j.diabet.2019.02.006

Garnock-Jones, K. P., & Plosker, G. L. (2009). Insulin glulisine: A review of its use in the management of diabetes mellitus. In Drugs. doi: 10.2165/00003495-200969080-00006

Gast, K. B., Tjeerdema, N., Stijnen, T., Smit, J. W. A., & Dekkers, O. M. (2012). Insulin Resistance and Risk of Incident Cardiovascular Events in Adults without Diabetes: Meta-Analysis. *PLoS ONE*. doi: 10.1371/journal.pone.0052036

Gast, K., Schüler, A., Wolff, M., Thalhammer, A., Berchtold, H., Nagel, N., Lenherr, G., Hauck, G., & Seckler, R. (2017). Rapid-Acting and Human Insulins: Hexamer Dissociation Kinetics upon Dilution of the Pharmaceutical Formulation. *Pharmaceutical Research*. doi: 10.1007/s11095-017-2233-0

Gathercole, L. L., Morgan, S. A., Bujalska, I. J., Hauton, D., Stewart, P. M., & Tomlinson, J. W. (2011). Regulation of lipogenesis by glucocorticoids and insulin

in human adipose tissue. *PLoS ONE*. doi: 10.1371/journal.pone.0026223

Gevorkov, Y., Yefanov, O., Barty, A., White, T. A., Mariani, V., Brehm, W., Tolstikova, A., Grigat, R. R., & Chapman, H. N. (2019). XGANDALF - Extended gradient descent algorithm for lattice finding. *Acta Crystallographica Section A: Foundations and Advances*. doi: 10.1107/S2053273319010593

Ghoshoon, M. B., Berenjian, A., Hemmati, S., Dabbagh, F., Karimi, Z., Negahdaripour, M., & Ghasemi, Y. (2015). Extracellular Production of Recombinant l-Asparaginase II in Escherichia coli: Medium Optimization Using Response Surface Methodology. *International Journal of Peptide Research and Therapeutics*. doi: 10.1007/s10989-015-9476-6

Giacco, F., & Brownlee, M. (2010). Oxidative stress and diabetic complications. In Circulation Research. doi: 10.1161/CIRCRESAHA.110.223545

Gibney, M. J., Walsh, M., Brennan, L., Roche, H. M., German, B., & Van Ommen, B. (2005). Metabolomics in human nutrition: Opportunities and challenges. In American Journal of Clinical Nutrition. doi: 10.1093/ajcn.82.3.497

Gillis, R. B., Solomon, H. V., Govada, L., Oldham, N. J., Dinu, V., Jiwani, S. I., Gyasi-Antwi, P., Coffey, F., Meal, A., Morgan, P. S., Harding, S. E., Helliwell, J. R., Chayen, N. E., & Adams, G. G. (2021). Analysis of insulin glulisine at the molecular level by X-ray crystallography and biophysical techniques. *Scientific Reports*. doi: 10.1038/s41598-021-81251-2

Glendorf, T., Stidsen, C. E., Norrman, M., Nishimura, E., Sørensen, A. R., & Kjeldsen, T. (2011). Engineering of insulin receptor isoform-selective insulin analogues. *PLoS ONE*. doi: 10.1371/journal.pone.0020288

Glidden, M. D., Aldabbagh, K., Phillips, N. B., Carr, K., Chen, Y. S., Whittaker, J., Phillips, M., Wickramasinghe, N. P., Rege, N., Swain, M., Peng, Y., Yang, Y., Lawrence, M. C., Yee, V. C., Ismail-Beigi, F., & Weiss, M. A. (2018). An ultra-stable single-chain insulin analog resists thermal inactivation and exhibits biological signaling duration equivalent to the native protein. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M117.808626

Goeddel, D. V., Kleid, D. G., & Bolivar, F. (1979). Expression in Escherichia coli of chemically synthesized genes for human insulin. *Proceedings of the National*

Academy of Sciences of the United States of America. doi: 10.1073/pnas.76.1.106

Goldman, J., Kapitza, C., Pettus, J., & Heise, T. (2017). Understanding how pharmacokinetic and pharmacodynamic differences of basal analog insulins influence clinical practice. In Current Medical Research and Opinion. doi: 10.1080/03007995.2017.1335192

Goldstein, B. J., Ahmad, F., Ding, W., Li, P. M., & Zhang, W. R. (1998). Regulation of the insulin signalling pathway by cellular protein-tyrosine phosphatases. *Molecular and Cellular Biochemistry*. doi: 10.1023/A:1006812218502

Gough, S. C. L., Bode, B. W., Woo, V. C., Rodbard, H. W., Linjawi, S., Zacho, M., Reiter, P. D., & Buse, J. B. (2015). One-year efficacy and safety of a fixed combination of insulin degludec and liraglutide in patients with type 2 diabetes: Results of a 26-week extension to a 26-week main trial. *Diabetes, Obesity and Metabolism*. doi: 10.1111/dom.12498

Gough, S., & Narendran, P. (2017). Insulin and Insulin Treatment. *Textbook of Diabetes*, 399–413. doi: 10.1002/9781118924853.CH29

Govender, K., Naicker, T., Lin, J., Baijnath, S., Chuturgoon, A. A., Abdul, N. S., Docrat, T., Kruger, H. G., & Govender, T. (2020). A novel and more efficient biosynthesis approach for human insulin production in Escherichia coli (E. coli). *AMB Express*. doi: 10.1186/s13568-020-00969-w

Goykhman, S., Drincic, A., Desmangles, J. C., & Rendell, M. (2009). Insulin Glargine: A review 8 years after its introduction. In Expert Opinion on Pharmacotherapy. doi: 10.1517/14656560902775677

Grant, S. F. A., Thorleifsson, G., Reynisdottir, I., Benediktsson, R., Manolescu, A., Sainz, J., Helgason, A., Stefansson, H., Emilsson, V., Helgadottir, A., Styrkarsdottir, U., Magnusson, K. P., Walters, G. B., Palsdottir, E., Jonsdottir, T., Gudmundsdottir, T., Gylfason, A., Saemundsdottir, J., Wilensky, R. L., … Stefansson, K. (2006). Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nature Genetics*. doi: 10.1038/ng1732

Grieco, A., Boneta, S., Gavira, J. A., Pey, A. L., Basu, S., Orlans, J., De Sanctis, D., Medina, M., & Martin-Garcia, J. M. (n.d.). Structural dynamics and functional cooperativity of human NQO1 by ambient temperature serial crystallography and

simulations. *Biorxiv.OrgA Grieco, S Boneta, JA Gavira, AL Pey, S Basu, J Orlans, D de Sanctis, M MedinabioRxiv, 2023•biorxiv.Org*. doi: 10.1101/2023.12.21.572834

Gu, X., Al Dubayee, M., Alshahrani, A., Masood, A., Benabdelkamel, H., Zahra, M., Li, L., Abdel Rahman, A. M., & Aljada, A. (2020). Distinctive Metabolomics Patterns Associated With Insulin Resistance and Type 2 Diabetes Mellitus. *Frontiers in Molecular Biosciences*. doi: 10.3389/fmolb.2020.609806

Guerreiro, V., Bernardes, I., Pereira, J., Silva, R. P., Fernandes, S., Carvalho, D., & Freitas, P. (2020). Acromegaly with congenital generalized lipodystrophy - Two rare insulin resistance conditions in one patient: A case report. *Journal of Medical Case Reports*. doi: 10.1186/s13256-020-2352-9

Gul, M., Ayan, E., Destan, E., Johnson, J. A., Shafiei, A., Kepceoğlu, A., Yilmaz, M., Ertem, F. B., Yapici, İ., Tosun, B., Baldir, N., Tokay, N., Nergiz, Z., Karakadioğlu, G., Paydos, S. S., Kulakman, C., Ferah, C. K., Güven, Ö., Atalay, N., … DeMirci, H. (2023). Rapid and efficient ambient temperature X-ray crystal structure determination at Turkish Light Source. *Scientific Reports*. doi: 10.1038/s41598-023-33989-0

Gunton, J. E., Kulkarni, R. N., Yim, S. H., Okada, T., Hawthorne, W. J., Tseng, Y. H., Roberson, R. S., Ricordi, C., O'Connell, P. J., Gonzalez, F. J., & Kahn, C. R. (2005). Loss of ARNT/HIF1beta mediates altered gene expression and pancreatic-islet dysfunction in human type 2 diabetes. *Cell*, *122*(3), 337–349. doi: 10.1016/J.CELL.2005.05.027

Guo, Z. Y., & Feng, Y. M. (2001). Effects of cysteine to serine substitutions in the two inter-chain disulfide bonds of insulin. *Biological Chemistry*. doi: 10.1515/BC.2001.054

Gursky, O., Badger, J., Li, Y., & Caspar, D. L. (1992). Conformational changes in cubic insulin crystals in the pH range 7–11. *Biophysical Journal*. doi: 10.1016/S0006-3495(92)81697-1

Haeusler, R. A., McGraw, T. E., & Accili, D. (2018). Metabolic Signalling: Biochemical and cellular properties of insulin receptor signalling. In Nature Reviews Molecular Cell Biology. doi: 10.1038/nrm.2017.89

Hagedorn, H. C. (1937). Protamine Insulinate. *Journal of the Royal Society of Medicine*. doi: 10.1177/003591573703000643

Hall, J. C. (1998). Review: Glycine. *Journal of Parenteral and Enteral Nutrition*, *22*(6), 393–398. doi: 10.1177/0148607198022006393

Halldén, G., Gafvelin, G., Mutt, V., & Jörnvall, H. (1986). Characterization of cat insulin. *Archives of Biochemistry and Biophysics*. doi: 10.1016/0003-9861(86)90528-X

Hamdan, I. I., Farah, D. G. H., Khalil, E. A., Mansour, R. S. H., & Abdel-Halim, H. (2022). Interaction of esomeprazole with insulin detemir and human albumin: A potential cause of hypoglycemia. *Biophysical Chemistry*, *285*, 106809. doi: 10.1016/J.BPC.2022.106809

Hanna, S. J., Powell, W. E., Long, A. E., Robinson, E. J. S., Davies, J., Megson, C., Howell, A., Jones, T. J., Ladell, K., Price, D. A., Dayan, C. M., Williams, A. J. K., Gillespie, K. M., & Wong, F. S. (2020). Slow progressors to type 1 diabetes lose islet autoantibodies over time, have few islet antigen-specific CD8+ T cells and exhibit a distinct CD95hi B cell phenotype. *Diabetologia*. doi: 10.1007/s00125-020-05114-7

Harper, M. E., Ullrich, A., & Saunders, G. F. (1981). Localization of the human insulin gene to the distal end of the short arm of chromosome 11. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.78.7.4458

Harrison, S. M., Bush, N. C., Wang, Y., Mucher, Z. R., Lorenzo, A. J., Grimsby, G. M., Schlomer, B. J., Büllesbach, E. E., & Baker, L. A. (2019). Insulin-Like Peptide 3 (INSL3) Serum Concentration During Human Male Fetal Life. *Frontiers in Endocrinology*. doi: 10.3389/fendo.2019.00596

Hartemann-Heurtier, A., Sachon, C., Masseboeuf, N., Corset, E., & Grimaldi, A. (2003). Functional intensified insulin therapy with short-acting insulin analog: Effects on HbA1c and frequency of severe hypoglycemia: An observational cohort study. *Diabetes and Metabolism*. doi: 10.1016/S1262-3636(07)70007-1

Hartman, I. (2008). Insulin analogs: Impact on treatment success, satisfaction, quality of life, and adherence. In Clinical Medicine and Research. doi: 10.3121/cmr.2008.793

Hassiepen, U., Federwisch, M., Mülders, T., & Wollmer, A. (1999). The lifetime of insulin hexamers. *Biophysical Journal*. doi: 10.1016/S0006-3495(99)77012-8

Havelund, S., Plum, A., Ribel, U., Jonassen, I., Vølund, A., Markussen, J., & Kurtzhals, P. (2004). The mechanism of protraction of insulin detemir, a long-acting, acylated analog of human insulin. *Pharmaceutical Research*. doi: 10.1023/B:PHAM.0000036926.54824.37

Helenius, A., & Aebi, M. (2004). Roles of N-linked glycans in the endoplasmic reticulum. In Annual Review of Biochemistry. doi: 10.1146/annurev.biochem.73.011303.073752

Hill, C. P., Dodson, E. J., Dodson, G. G., Dauter, Z., Dauter, Z., & Dunn, M. F. (1991). X-ray Structure of an Unusual Ca2+ Site and the Roles of Zn2+ and Ca2+ in the Assembly, Stability, and Storage of the Insulin Hexamer. *Biochemistry*. doi: 10.1021/bi00218a006

Hills, C. E., & Brunskill, N. J. (2008). Intracellular signalling by C-peptide. In Experimental diabetes research. doi: 10.1155/2008/635158

Hills, C. E., & Brunskill, N. J. (2009). Cellular and physiological effects of C-peptide. In Clinical Science. doi: 10.1042/CS20080441

Himsworth, H. P. (2013). Diabetes mellitus: its differentiation into insulin-sensitive and insulin-insensitive types. 1936. *International Journal of Epidemiology*. doi: 10.1093/ije/dyt203

Hirata, K., Yamashita, K., Ueno, G., Kawano, Y., Hasegawa, K., Kumasaka, T., & Yamamoto, M. (2019). Zoo: An automatic data-collection system for high-throughput structure analysis in protein microcrystallography. *Acta Crystallographica Section D: Structural Biology*. doi: 10.1107/S2059798318017795

Hirsch, I. B., Juneja, R., Beals, J. M., Antalis, C. J., & Wright, E. E. (2021). The evolution of insulin and how it informs therapy and treatment choices. In Endocrine Reviews. doi: 10.1210/ENDREV/BNAA015

Hjorth, C. F., Norrman, M., Wahlund, P. O., Benie, A. J., Petersen, B. O., Jessen, C. M., Pedersen, T., Vestergaard, K., Steensgaard, D. B., Pedersen, J. S., Naver, H.,

Hubálek, F., Poulsen, C., & Otzen, D. (2016). Structure, Aggregation, and Activity of a Covalent Insulin Dimer Formed during Storage of Neutral Formulation of Human Insulin. *Journal of Pharmaceutical Sciences*. doi: 10.1016/j.xphs.2016.01.003

Hodgkin, D. C. (1971). X Rays and the Structure of Insulin. *British Medical Journal*. doi: 10.1136/bmj.4.5785.447

Holleman, F. (2015). Insulin lispro (revision number 20). In Diapedia. doi: 10.14496/dia.8104096170.20

Home, P. D., & Mehta, R. (2021). Insulin therapy development beyond 100 years. In The Lancet Diabetes and Endocrinology. doi: 10.1016/S2213-8587(21)00182-0

Hovorka, R. (2011). Closed-loop insulin delivery: From bench to clinical practice. In Nature Reviews Endocrinology. doi: 10.1038/nrendo.2011.32

Howell, S. L., Montague, W., & Tyhurst, M. (1975). Calcium distribution in islets of Langerhans: a study of calcium concentrations and of calcium accumulation in B cell organelles. *Journal of Cell Science*. doi: 10.1242/jcs.19.2.395

Howell, S. L., & Taylor, K. W. (1968). Potassium ions and the secretion of insulin by islets of Langerhans incubated in vitro. *The Biochemical Journal*. doi: 10.1042/bj1080017

Howell, Simon L., Tyhurst, M., Duvefelt, H., Andersson, A., & Hellerström, C. (1978). Role of zinc and calcium in the formation and storage of insulin in the pancreatic β-cell. *Cell and Tissue Research*. doi: 10.1007/BF00220518

Hua, Q. X., Nakagawa, S. H., Jia, W., Huang, K., Phillips, N. B., Hu, S. Q., & Weiss, M. A. (2008). Design of an active ultrastable single-chain insulin analog: Synthesis, structure, and therapeutic implications. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M800313200

Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., De Groot, B. L., Grubmüller, H., & MacKerell, A. D. (2016). CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nature Methods*. doi: 10.1038/nmeth.4067

Huang, K., Xu, B., Hu, S. Q., Chu, Y. C., Hua, Q. X., Qu, Y., Li, B., Wang, S., Wang,

R. Y., Nakagawa, S. H., Theede, A. M., Whittaker, J., De Meyts, P., Katsoyannis, P. G., & Weiss, M. A. (2004). How insulin binds: The B-chain α-helix contacts the L1 β-helix of the insulin receptor. *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2004.05.023

Huang, M., & Joseph, J. W. (2012). Metabolomic analysis of pancreatic β-cell insulin release in response to glucose. *Islets*. doi: 10.4161/isl.20141

Hubálek, F., Refsgaard, H. H. F., Gram-Nielsen, S., Madsen, P., Nishimura, E., Münzel, M., Brand, C. L., Stidsen, C. E., Claussen, C. H., Wulff, E. M., Pridal, L., Ribel, U., Kildegaard, J., Porsgaard, T., Johansson, E., Steensgaard, D. B., Hovgaard, L., Glendorf, T., Hansen, B. F., … Kjeldsen, T. (2020). Molecular engineering of safe and efficacious oral basal insulin. *Nature Communications*. doi: 10.1038/s41467-020-17487-9

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*. doi: 10.1016/0263-7855(96)00018-5

IDF,2019,https://diabetesatlas.org/upload/resources/material/20200302_133351_IDFATLAS9e-final-web.pdf

Insulin assembly: its modification by protein engineering and ligand binding. (1993). *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*. doi: 10.1098/rsta.1993.0126

Inzucchi, S. E. (2012). Clinical practice. Diagnosis of diabetes. *The New England Journal of Medicine*, *367*(6), 542–550. doi: 10.1056/NEJMCP1103643

Ishikawa, T., Aoyagi, H., Asaka, T., Asano, Y., Azumi, N., Bizen, T., Ego, H., Fukami, K., Fukui, T., Furukawa, Y., Goto, S., Hanaki, H., Hara, T., Hasegawa, T., Hatsui, T., Higashiya, A., Hirono, T., Hosoda, N., Ishii, M., … Kumagai, N. (2012). A compact X-ray free-electron laser emitting in the sub-ångström region. *Nature Photonics*. doi: 10.1038/nphoton.2012.141

Izaki, K., Bacteriology, K. A.-J. of, & 1965, undefined. (1965). Effect of Various Conditions on Accumulation of Oxytetracycline in Escherichia coli. *Am Soc MicrobiolK Izaki, K ArimaJournal of Bacteriology, 1965•Am Soc Microbiol*. Retrieved from https://journals.asm.org/doi/abs/10.1128/jb.89.5.1335-1339.1965

Jabłońska, K., Molęda, P., Safranow, K., & Majkowska, L. (2018). Rapid-acting and Regular Insulin are Equal for High Fat-Protein Meal in Individuals with Type 1 Diabetes Treated with Multiple Daily Injections. *Diabetes Therapy*. doi: 10.1007/s13300-017-0364-2

Jaspers, S., Lok, S., Lofton-Day, C. E., Haldeman, B. A., Whitmore, T. E., Foley, K. P., & Conklin, D. (2001). The genomics of Insulin 5. In Relaxin 2000 (pp. 363–369). Springer.

Jedrychowski, M. P., Gartner, C. A., Gygi, S. P., Zhou, L., Herz, J., Kandror, K. V., & Pilch, P. F. (2010). Proteomic analysis of GLUT4 storage vesicles reveals LRP1 to be an important vesicle component and target of insulin signaling. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M109.040428

Jeevanandam, J., Paramasivam, E., & Saraswathi, N. T. (2023). Glycation restrains open-closed conformation of Insulin. *Computational Biology and Chemistry*. doi: 10.1016/j.compbiolchem.2022.107803

Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y. G., & He, C. (2011). N6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nature Chemical Biology*. doi: 10.1038/nchembio.687

Jia, W., He, Y.-F., Qian, X.-J., & Chen, J. (2022). TPMT mRNA Expression: A Novel Prognostic Biomarker for Patients with Colon Cancer by Bioinformatics Analysis. *International Journal of General Medicine*, *15*, 151.

Jiráček, J., Žáková, L., Antolíková, E., Watson, C. J., Turkenburg, J. P., Dodson, G. G., & Brzozowski, A. M. (2010). Implications for the active form of human insulin based on the structural convergence of highly active hormone analogues. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0911785107

Johansson, B. L., Borg, K., Fernqvist-Forbes, E., Kernell, A., Odergren, T., & Wahren, J. (2000). Beneficial effects of C-peptide on incipient nephropathy and neuropathy in patients with Type 1 diabetes mellitus. *Diabetic Medicine*. doi: 10.1046/j.1464-5491.2000.00274.x

Johansson, J., Ekberg, K., Shafqat, J., Henriksson, M., Chibalin, A., Wahren, J., &

Jörnvall, H. (2002). Molecular effects of proinsulin C-peptide. *Biochemical and Biophysical Research Communications*. doi: 10.1016/S0006-291X(02)00721-0

Johnson, I. S. (1983). Human insulin from recombinant DNA technology. In Science. doi: 10.1126/science.6337396

Jonassen, I., Havelund, S., Hoeg-Jensen, T., Steensgaard, D. B., Wahlund, P. O., & Ribel, U. (2012). Design of the novel protraction mechanism of insulin degludec, an ultra-long-acting basal insulin. *Pharmaceutical Research*. doi: 10.1007/s11095-012-0739-z

Joshi, S. R., Parikh, R. M., & Das, A. K. (2007). Insulin--history, biochemistry, physiology and pharmacology. In The Journal of the Association of Physicians of India.

Joslin's Diabetes Mellitus: Edited by C. Ronald Kahn ... [et Al.]. - Google Kitaplar. (n.d.). Retrieved from https://books.google.com.tr/books?hl=tr&lr=&id=ohgjG0qAvfgC&oi=fnd&pg=PA1&dq=Warram,+J.+%26+Krolewski,+A.+Joslin's+Diabetes+Mellitus:+Edited+by+C.+Ronald+Kahn+...+%5Bet+Al.%5D.+in+Epidemiology+of+diabetes+mellitus,+2005.&ots=yLEj2IH7Ex&sig=I-LQuj2kRjSR8SoHtvC-qs6x-98&redir_esc=y#v=onepage&q=Warram%2C J. %26 Krolewski%2C A. Joslin's Diabetes Mellitus%3A Edited by C. Ronald Kahn ... %5Bet Al.%5D. in Epidemiology of diabetes mellitus%2C 2005.&f=false

Kaaks, R. (1996). Nutrition, hormones, and breast cancer: Is insulin the missing link? In Cancer Causes and Control. doi: 10.1007/BF00051703

Kaarsholm, N. C., & Dunn, M. F. (1987). Effects of Calcium Ion on Ternary Complexes Formed between 4-(2-Pyridylazo)resorcinol and the Two-Zinc Insulin Hexamer. *Biochemistry*. doi: 10.1021/bi00377a032

Kakade, M. L., & Liener, I. E. (1969). Determination of available lysine in proteins. *Analytical Biochemistry*. doi: 10.1016/0003-2697(69)90032-3

Kameshima, T., Ono, S., Kudo, T., Ozaki, K., Kirihara, Y., Kobayashi, K., Inubushi, Y., Yabashi, M., Horigome, T., Holland, A., Holland, K., Burt, D., Murao, H., & Hatsui, T. (2014). Development of an X-ray pixel detector with multi-port charge-coupled device for X-ray free-electron laser experiments. *Review of Scientific*

Instruments. doi: 10.1063/1.4867668

Kamoda, T., Saito, T., Kinugasa, H., Iwasaki, N., Sumazaki, R., Mouri, Y., Izumi, I., Hirano, T., & Matsui, A. (2005). A case of Shwachman-Diamond syndrome presenting with diabetes from early infancy. *Diabetes Care*. doi: 10.2337/diacare.28.6.1508

Kanaka-Gantenbein, C., Kitsiou, S., Mavrou, A., Stamoyannou, L., Kolialexi, A., Kekou, K., Liakopoulou, M., & Chrousos, G. (2004). Tall stature, insulin resistance, and disturbed behavior in a girl with the triple X syndrome harboring three SHOX genes: Offspring of a father with mosaic Klinefelter syndrome but with two maternal X chromosomes. *Hormone Research*. doi: 10.1159/000076532

Karagiannis, P., Takahashi, K., Saito, M., Yoshida, Y., Okita, K., Watanabe, A., Inoue, H., Yamashita, J. K., Todani, M., Nakagawa, M., Osawa, M., Yashiro, Y., Yamanaka, S., & Osafune, K. (2019). Induced pluripotent stem cells and their use in human models of disease and development. In Physiological Reviews. doi: 10.1152/physrev.00039.2017

Keku, T. O., Lund, P. K., Galanko, J., Simmons, J. G., Woosley, J. T., & Sandler, R. S. (2005). Insulin resistance, apoptosis, and colorectal adenoma risk. *Cancer Epidemiology Biomarkers and Prevention*. doi: 10.1158/1055-9965.EPI-05-0239

Kenari, S. L. D., Alemzadeh, I., & Maghsodi, V. (2011). Production of l-asparaginase from Escherichia coli ATCC 11303: Optimization by response surface methodology. *Food and Bioproducts Processing*. doi: 10.1016/j.fbp.2010.11.002

Kenett, R. S. (1991). Two Methods for Comparing Pareto Charts. *Journal of Quality Technology*. doi: 10.1080/00224065.1991.11979280

Kensy, F., Engelbrecht, C., & Büchs, J. (2009). Scale-up from microtiter plate to laboratory fermenter: Evaluation by online monitoring techniques of growth and protein expression in Escherichia coli and Hansenula polymorpha fermentations. *Microbial Cell Factories*. doi: 10.1186/1475-2859-8-68

Kern, W., Stange, E. F., Fehm, H. L., & Klein, H. H. (1999). Glucocorticoid-induced diabetes mellitus in gastrointestinal diseases TT - Glucocorticoid-induzierter Diabetes mellitus bei gastroenterologischen Erkrankungen. *Zeitschrift Fur Gastroenterologie*.

Khan, A., & Pessin, J. (2002). Insulin regulation of glucose uptake: A complex interplay of intracellular signalling pathways. In Diabetologia. doi: 10.1007/s00125-002-0974-7

Khetan, S., Kursawe, R., Youn, A., Lawlor, N., Jillette, A., Marquez, E. J., Ucar, D., & Stitzel, M. L. (2018). Type 2 diabetes-associated genetic variants regulate chromatin accessibility in human islets. *Diabetes*. doi: 10.2337/db18-0393

Kim, B., & Feldman, E. L. (2012). Insulin resistance in the nervous system. In Trends in Endocrinology and Metabolism. doi: 10.1016/j.tem.2011.12.004

Kim, C. K., Lee, S. B., & Son, Y. J. (2015). Large-scale refolding and enzyme reaction of human preproinsulin for production of human insulin. *Journal of Microbiology and Biotechnology*. doi: 10.4014/jmb.1504.04062

Kim, O. Y., Lee, J. H., & Sweeney, G. (2013). Metabolomic profiling as a useful tool for diagnosis and treatment of chronic disease: Focus on obesity, diabetes and cardiovascular diseases. In Expert Review of Cardiovascular Therapy. doi: 10.1586/erc.12.121

Kim, Y., & Shields, J. E. (1992). pH dependent conformational changes in the T-and R-states of insulin in solution: Circular dichroic studies in the pH range of 6 to 10. *Biochemical and Biophysical Research Communications*. doi: 10.1016/0006-291X(92)90862-F

Konteatis, Z., Artin, E., Nicolay, B., Straley, K., Padyana, A. K., Jin, L., Chen, Y., Narayaraswamy, R., Tong, S., Wang, F., Zhou, D., Cui, D., Cai, Z., Luo, Z., Fang, C., Tang, H., Lv, X., Nagaraja, R., Yang, H., … Biller, S. A. (2020). Vorasidenib (AG-881): A First-in-Class, Brain-Penetrant Dual Inhibitor of Mutant IDH1 and 2 for Treatment of Glioma. *ACS Medicinal Chemistry Letters*. doi: 10.1021/acsmedchemlett.9b00509

Kopp, J., Slouka, C., Ulonska, S., Kager, J., Fricke, J., Spadiut, O., & Herwig, C. (2018). Impact of glycerol as carbon source onto specific sugar and inducer uptake rates and inclusion body productivity in E. Coli BL21(DE3). *Bioengineering*. doi: 10.3390/bioengineering5010001

Koranyi, L., James, D. E., Kraegen, E. W., & Permutt, M. A. (1992). Feedback inhibition of insulin gene expression by insulin. *Journal of Clinical Investigation*.

doi: 10.1172/JCI115602

Kosinová, L., Veverka, V., Novotná, P., Collinsová, M., Urbanová, M., Moody, N. R., Turkenburg, J. P., Jiráček, J., Brzozowski, A. M., & Žáková, L. (2014). Insight into the structural and biological relevance of the T/R transition of the N-terminus of the B-chain in human insulin. *Biochemistry*. doi: 10.1021/bi500073z

Koslow, S. H., Stokes, P. E., Mendels, J., Ramsey, A., & Casper, R. (1982). Insulin Tolerance Test: human growth hormone response and insulin resistance in primary unipolar depressed, bipolar depressed and control subjects. *Psychological Medicine*, *12*(1), 45–55. doi: 10.1017/S0033291700043270

Kretsinger, R. H. (1972). Gene triplication deduced from the tertiary structure of a muscle calcium binding protein. *Nature New Biology*. doi: 10.1038/newbio240085a0

Kristensen, C., Andersen, A. S., Hach, M., Wiberg, F. C., Schaffer, L., & Kjeldsen, T. (1995). A single-chain insulin-like growth factor I/insulin hybrid binds with high affinity to the insulin receptor. *Biochemical Journal*. doi: 10.1042/bj3050981

Kristensen, P. L., Hansen, L. S., Jespersen, M. J., Pedersen-Bjergaard, U., Beck-Nielsen, H., Christiansen, J. S., Nørgaard, K., Perrild, H., Parving, H. H., Thorsteinsson, B., & Tarnow, L. (2012). Insulin analogues and severe hypoglycaemia in type 1 diabetes. *Diabetes Research and Clinical Practice*. doi: 10.1016/j.diabres.2011.10.046

Kukla, A., Hill, J., Merzkani, M., Bentall, A., Lorenz, E. C., Park, W. D., D'Costa, M., Kudva, Y. C., Stegall, M. D., & Shah, P. (2020). The Use of GLP1R Agonists for the Treatment of Type 2 Diabetes in Kidney Transplant Recipients. *Transplantation Direct*. doi: 10.1097/TXD.0000000000000971

Kumar, S., & Nussinov, R. (1999). Salt bridge stability in monomeric proteins. *Journal of Molecular Biology*. doi: 10.1006/jmbi.1999.3218

Kumar, S., Tsai, C. J., & Nussinov, R. (2000). Factors enhancing protein thermostability. *Protein Engineering*. doi: 10.1093/protein/13.3.179

Kurtzhals, P., Diabetes, U. R.-, & 1995, undefined. (n.d.). Action profile of cobalt (III)-insulin: a novel principle of protraction of potential use for basal insulin delivery.

Am Diabetes AssocP Kurtzhals, U RibelDiabetes, 1995•Am Diabetes Assoc. Retrieved from https://diabetesjournals.org/diabetes/article-abstract/44/12/1381/10315

Kurtzhals, P., Havelund, S., Jonassen, I., Kiehr, B., Larsen, U. D., Ribel, U., & Markussen, J. (1995). Albumin binding of insulins acylated with fatty acids: Characterization of the ligand-protein interaction and correlation between binding affinity and timing of the insulin effect in vivo. *Biochemical Journal*. doi: 10.1042/bj3120725

Kusuma, S. A. F., Parwati, I., Rostinawati, T., Yusuf, M., Fadhlillah, M., Ahyudanari, R. R., Rukayadi, Y., & Subroto, T. (2019). Optimization of culture conditions for Mpt64 synthetic gene expression in Escherichia coli BL21 (DE3) using surface response methodology. *Heliyon*. doi: 10.1016/j.heliyon.2019.e02741

Kwon, T. G., Jang, A. Y., Kim, S. W., Hong, Y. J., Bae, J. H., Lee, S. Y., Kim, S. H., & Han, S. H. (2020). Design and rationale of a randomized control trial testing the effectiveness of combined therapy with STAtin plus FENOfibrate and statin alone in non-diabetic, combined dyslipidemia patients with non-intervened intermediate coronary artery disease - STAFE. *Trials*. doi: 10.1186/s13063-020-04291-5

Lacy, P. E. (1974). Structure and function of the endocrine cell types of the islets. In Advances in metabolic disorders. doi: 10.1016/b978-0-12-027307-2.50013-x

LaPierre, M. P., & Stoffel, M. (2017). MicroRNAs as stress regulators in pancreatic beta cells and diabetes. In Molecular Metabolism. doi: 10.1016/j.molmet.2017.06.020

Larsen, A. H., Jessen, N., Nørrelund, H., Tolbod, L. P., Harms, H. J., Feddersen, S., Nielsen, F., Brøsen, K., Hansson, N. H., Frøkiær, J., Poulsen, S. H., Sörensen, J., & Wiggers, H. (2020). A randomised, double-blind, placebo-controlled trial of metformin on myocardial efficiency in insulin-resistant chronic heart failure patients without diabetes. *European Journal of Heart Failure*. doi: 10.1002/ejhf.1656

Larsson, S., Resjö, S., Gomez, M. F., James, P., & Holm, C. (2012). Characterization of the lipid droplet proteome of a clonal insulin-producing β-cell line (INS-1 832/13). *Journal of Proteome Research*. doi: 10.1021/pr200957p

Lavi, S., & Shatkin, A. J. (1975). Methylated simian virus 40 specific RNA from nuclei and cytoplasm of infected BSC 1 cells. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.72.6.2012

Lawlor, N., George, J., Bolisetty, M., Kursawe, R., Sun, L., Sivakamasundari, V., Kycia, I., Robson, P., & Stitzel, M. L. (2017). Single-cell transcriptomes identify human islet cell signatures and reveal cell-type-specific expression changes in type 2 diabetes. *Genome Research*. doi: 10.1101/gr.212720.116

Lefever, E., Vliebergh, J., & Mathieu, C. (2021). Improving the treatment of patients with diabetes using insulin analogues: current findings and future directions. In Expert Opinion on Drug Safety. doi: 10.1080/14740338.2021.1856813

Lemaire, K., Ravier, M. A., Schraenen, A., Creemers, J. W. M., Van De Plas, R., Granvik, M., Van Lommel, L., Waelkens, E., Chimienti, F., Rutter, G. A., Gilon, P., In't Veld, P. A., & Schuit, F. C. (2009). Insulin crystallization depends on zinc transporter ZnT8 expression, but is not required for normal glucose homeostasis in mice. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0906587106

Lennen, R. M., Nilsson Wallin, A. I., Pedersen, M., Bonde, M., Luo, H., Herrgård, M. J., & Sommer, M. O. A. (2015). Transient overexpression of DNA adenine methylase enables efficient and mobile genome engineering with reduced off-target effects. *Nucleic Acids Research*. doi: 10.1093/nar/gkv1090

Lenoir, D., Gréco, S., Louisot, P., & Biol, M. C. (2000). Implication of insulin and nutritional factors in the regulation of intestinal galactosyltransferase activity during postnatal development. *Metabolism: Clinical and Experimental*. doi: 10.1016/S0026-0495(00)80020-7

LeRoith, D., & Roberts, C. T. (2003). The insulin-like growth factor system and cancer. In Cancer Letters. doi: 10.1016/S0304-3835(03)00159-9

Levy, E. D., & Teichmann, S. (2013). Structural, evolutionary, and assembly principles of protein oligomerization. In Progress in Molecular Biology and Translational Science. doi: 10.1016/B978-0-12-386931-9.00002-7

LEYER, S., GATTNER, H. -G, LEITHÄUSER, M., BRANDENBURG, D., WOLLMER, A., & HÖCKER, H. (1995). The role of the C-terminus of the insulin

B-chain in modulating structural and functional properties of the hormone. *International Journal of Peptide and Protein Research*. doi: 10.1111/j.1399-3011.1995.tb01074.x

Li, J., Sipple, J., Maynard, S., Mehta, P. A., Rose, S. R., Davies, S. M., & Pang, Q. (2012). Fanconi anemia links reactive oxygen species to insulin resistance and obesity. *Antioxidants and Redox Signaling*. doi: 10.1089/ars.2011.4417

Li, L., Krznar, P., Erban, A., Agazzi, A., Martin-Levilain, J., Supale, S., Kopka, J., Zamboni, N., & Maechler, P. (2019). Metabolomics identifies a biomarker revealing in vivo loss of functional β-cell mass before diabetes onset. *Diabetes*. doi: 10.2337/db19-0131

Li, Yang V. (2014). Zinc and insulin in pancreatic beta-cells. In Endocrine. doi: 10.1007/s12020-013-0032-x

Li, Yangxin, & Geng, Y. J. (2010). A potential role for insulin-like growth factor signaling in induction of pluripotent stem cell formation. In Growth Hormone and IGF Research. doi: 10.1016/j.ghir.2010.09.005

Liebl, A., Davidson, J., Mersebach, H., Dykiel, P., Tack, C. J., & Heise, T. (2013). A novel insulin combination of insulin degludec and insulin aspart achieves a more stable overnight glucose profile than insulin glargine: Results from continuous glucose monitoring in a proof-of-concept trial. *Journal of Diabetes Science and Technology*. doi: 10.1177/193229681300700524

Lieblich, S. A., Fang, K. Y., Cahn, J. K. B., Rawson, J., Lebon, J., Ku, H. T., & Tirrell, D. A. (2017). 4S-Hydroxylation of Insulin at ProB28 Accelerates Hexamer Dissociation and Delays Fibrillation. *Journal of the American Chemical Society*. doi: 10.1021/jacs.7b00794

Lim, J.-M., Wollaston-Hayden, E. E., Teo, C. F., Hausman, D., & Wells, L. (2014). Quantitative secretome and glycome of primary human adipocytes during insulin resistance. *Clinical Proteomics*, *11*(1), 1–23.

Lindhurst, M. J., Parker, V. E. R., Payne, F., Sapp, J. C., Rudge, S., Harris, J., Witkowski, A. M., Zhang, Q., Groeneveld, M. P., Scott, C. E., Daly, A., Huson, S. M., Tosi, L. L., Cunningham, M. L., Darling, T. N., Geer, J., Gucev, Z., Sutton, V. R., Tziotzios, C., … Semple, R. K. (2012). Mosaic overgrowth with fibroadipose

hyperplasia is caused by somatic activating mutations in PIK3CA. *Nature Genetics*. doi: 10.1038/ng.2332

Ling, C., Del Guerra, S., Lupi, R., Rönn, T., Granhall, C., Luthman, H., Masiello, P., Marchetti, P., Groop, L., & Del Prato, S. (2008). Epigenetic regulation of PPARGC1A in human type 2 diabetic islets and effect on insulin secretion. *Diabetologia*. doi: 10.1007/s00125-007-0916-5

Ling, Z., Qi-Qing, J., Yu, W., Fei, B., Shu-Jie, Q., Shu-Bo, W., Zhen-Ying, P., & Yu-Ping, B. (2016). Transgenic Expression and Identification of Recombinant Human Proinsulin in Peanut. *Brazilian Archives of Biology and Technology*. doi: 10.1590/1678-4324-2016150131

Lipska, K. J., Ross, J. S., Van Houten, H. K., Beran, D., Yudkin, J. S., & Shah, N. D. (2014). Use and out-of-pocket costs of insulin for type 2 diabetes mellitus from 2000 through 2010. *JAMA - Journal of the American Medical Association*. doi: 10.1001/jama.2014.6316

Lisgarten, D. R., Palmer, R. A., Lobley, C. M. C., Naylor, C. E., Chowdhry, B. Z., Al-Kurdi, Z. I., Badwan, A. A., Howlin, B. J., Gibbons, N. C. J., Saldanha, J. W., Lisgarten, J. N., & Basak, A. K. (2017). Ultra-high resolution X-ray structures of two forms of human recombinant insulin at 100 K. *Chemistry Central Journal*. doi: 10.1186/s13065-017-0296-y

Liu, G., Liu, S., Xing, G., & Wang, F. (2020). lncRNA PVT1/MicroRNA-17-5p/PTEN Axis Regulates Secretion of E2 and P4, Proliferation, and Apoptosis of Ovarian Granulosa Cells in PCOS. *Molecular Therapy - Nucleic Acids*. doi: 10.1016/j.omtn.2020.02.007

Liu, J., Liu, W., Li, H., Deng, Q., Yang, M., Li, X., & Liang, Z. (2019). Identification of key genes and pathways associated with cholangiocarcinoma development based on weighted gene correlation network analysis. *PeerJ*. doi: 10.7717/peerj.7968

Liu, M., Wright, J., Guo, H., Xiong, Y., & Arvan, P. (2014). Proinsulin entry and transit through the endoplasmic reticulum in pancreatic beta cells. In Vitamins and Hormones. doi: 10.1016/B978-0-12-800174-5.00002-8

Lizcano, J. M., & Alessi, D. R. (2002). The insulin signalling pathway. In Current Biology. doi: 10.1016/S0960-9822(02)00777-7

Lourenco, M. V., Clarke, J. R., Frozza, R. L., Bomfim, T. R., Forny-Germano, L., Batista, A. F., Sathler, L. B., Brito-Moreira, J., Amaral, O. B., Silva, C. A., Freitas-Correa, L., Espírito-Santo, S., Campello-Costa, P., Houzel, J. C., Klein, W. L., Holscher, C., Carvalheira, J. B., Silva, A. M., Velloso, L. A., … De Felice, F. G. (2013). TNF-α mediates PKR-dependent memory impairment and brain IRS-1 inhibition induced by Alzheimer's β-amyloid oligomers in mice and monkeys. *Cell Metabolism*. doi: 10.1016/j.cmet.2013.11.002

Lu, J., Xie, G., Jia, W., & Jia, W. (2013). Metabolomics in human type 2 diabetes research. In Frontiers of Medicine in China. doi: 10.1007/s11684-013-0248-4

Lumeng, C. N., Delproposto, J. B., Westcott, D. J., & Saltiel, A. R. (2008). Phenotypic switching of adipose tissue macrophages with obesity is generated by spatiotemporal differences in macrophage subtypes. *Diabetes*. doi: 10.2337/db08-0872

Lynch, C. J., & Adams, S. H. (2014). Branched-chain amino acids in metabolic signalling and insulin resistance. In Nature Reviews Endocrinology. doi: 10.1038/nrendo.2014.171

Mackenzie, R. W. A., & Elliott, B. T. (2014). Akt/PKB activation and insulin signaling: A novel insulin signaling pathway in the treatment of type 2 diabetes. In Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy. doi: 10.2147/DMSO.S48260

Macleod, J. J. R. (1922). INSULIN AND DIABETES: A GENERAL STATEMENT OF THE PHYSIOLOGICAL AND THERAPEUTIC EFFECTS OF INSULIN. *BMJ*. doi: 10.1136/bmj.2.3227.833

Maeno, T., Okumura, A., Ishikawa, T., Kato, K., Sakakibara, F., Sato, K., Ayada, M., Hotta, N., Tagaya, T., Fukuzawa, Y., & Kakumu, S. (2003). Mechanisms of increased insulin resistance in non-cirrhotic patients with chronic hepatitis C virus infection. *Journal of Gastroenterology and Hepatology (Australia)*. doi: 10.1046/j.1440-1746.2003.03179.x

Maikawa, C. L., Smith, A. A. A., Zou, L., Meis, C. M., Mann, J. L., Webber, M. J., & Appel, E. A. (2020). Stable Monomeric Insulin Formulations Enabled by Supramolecular PEGylation of Insulin Analogues. *Advanced Therapeutics*. doi:

10.1002/adtp.201900094

Malaguarnera, R., & Belfiore, A. (2011). The insulin receptor: A new target for cancer therapy. In Frontiers in Endocrinology. doi: 10.3389/fendo.2011.00093

Malaisse, W. J., Hutton, J. C., Kawazu, S., Herchuelz, A., Valverde, I., & Sener, A. (1979). The stimulus-secretion coupling of glucose-induced insulin release - XXXV. The links between metabolic and cationic events. *Diabetologia*. doi: 10.1007/BF01223623

Malakar, P., & Venkatesh, K. V. (2012). Effect of substrate and IPTG concentrations on the burden to growth of Escherichia coli on glycerol due to the expression of Lac proteins. *Applied Microbiology and Biotechnology*. doi: 10.1007/s00253-011-3642-3

Manallack, D. T., Andrews, P. R., & Woods, E. F. (1985). Design. Synthesis, and Testing of Insulin Hexamer-Stabilizing Agents. *Journal of Medicinal Chemistry*. doi: 10.1021/jm00148a025

Mandal, K., Dhayalan, B., Avital-Shmilovici, M., Tokmakoff, A., & Kent, S. B. H. (2016). Crystallization of Enantiomerically Pure Proteins from Quasi-Racemic Mixtures: Structure Determination by X-Ray Diffraction of Isotope-Labeled Ester Insulin and Human Insulin. *ChemBioChem*. doi: 10.1002/cbic.201500600

Mannucci, E., Monami, M., & Marchionni, N. (2009). Short-acting insulin analogues vs. regular human insulin in type 2 diabetes: A meta-analysis. *Diabetes, Obesity and Metabolism*. doi: 10.1111/j.1463-1326.2008.00934.x

Manolopoulou, M., Guo, Q., Malito, E., Schilling, A. B., & Tang, W. J. (2009). Molecular basis of catalytic chamber-assisted unfolding and cleavage of human insulin by human insulin-degrading enzyme. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M900068200

Mark, P., & Nilsson, L. (2001). Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *Journal of Physical Chemistry A*. doi: 10.1021/jp003020w

Marks, V., & Teale, J. D. (1996). Investigation of hypoglycaemia. In Clinical Endocrinology. doi: 10.1046/j.1365-2265.1996.659478.x

Marks, V., & Wark, G. (2013). Forensic aspects of insulin. In Diabetes Research and

Clinical Practice. doi: 10.1016/j.diabres.2013.05.002

Marselli, L., Thorne, J., Dahiya, S., Sgroi, D. C., Sharma, A., Bonner-Weir, S., Marchetti, P., & Weir, G. C. (2010). Gene expression profiles of beta-cell enriched tissue obtained by laser capture microdissection from subjects with type 2 diabetes. *PLoS ONE*. doi: 10.1371/journal.pone.0011499

Marshall S, Bacote V, & Traxinger RR. (1991). Discovery of a metabolic pathway mediating glucose-induced desensitization of the glucose transport system. Role of hexosamine biosynthesis in the induction of insulin resistance. *Journal of Biological Chemistry*.

Mayer, J. P., Zhang, F., & DiMarchi, R. D. (2007). Insulin structure and function. In Biopolymers - Peptide Science Section. doi: 10.1002/bip.20734

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography*. doi: 10.1107/S0021889807021206

McDonald, T. S., Kumar, V., Fung, J. N., Woodruff, T. M., & Lee, J. D. (2021). Glucose clearance and uptake is increased in the SOD1G93A mouse model of amyotrophic lateral sclerosis through an insulin-independent mechanism. *FASEB Journal*. doi: 10.1096/fj.202002450R

McKeage, K., & Goa, K. L. (2001). Insulin glargine: A review of its therapeutic use as a long-acting agent for the management of type 1 and 2 diabetes mellitus. In Drugs. doi: 10.2165/00003495-200161110-00007

Medzhitov, R., & Janeway, C. A. (1997). Innate immunity: Impact on the adaptive immune response. *Current Opinion in Immunology*. doi: 10.1016/S0952-7915(97)80152-5

Meerson, A. (2020). Leptin-responsive MiR-4443 is a small regulatory RNA independent of the canonic microRNA biogenesis pathway. *Biomolecules*. doi: 10.3390/biom10020293

Meier, J. J., Veldhuis, J. D., & Butler, P. C. (2005). Pulsatile insulin secretion dictates systemic insulin delivery by regulating hepatic insulin extraction in humans. *Diabetes*. doi: 10.2337/diabetes.54.6.1649

Melki, V., Renard, E., Lassmann-Vague, V., … S. B.-D., & 1998, undefined. (1998). Improvement of HbA1c and Blood Glucose Stability in IDDM Patients Treated With Lispro Insulin Analog in External Pumps. *Am Diabetes AssocV Melki, E Renard, V Lassmann-Vague, S Boivin, B Guerci, H Hanaire-Broutin, J BringerDiabetes Care, 1998•Am Diabetes Assoc*. doi: 10.2337/diacare.21.6.977

Menting, J. G., Whittaker, J., Margetts, M. B., Whittaker, L. J., Kong, G. K. W., Smith, B. J., Watson, C. J., Žáková, L., Kletvíková, E., Jiráček, J., Chan, S. J., Steiner, D. F., Dodson, G. G., Brzozowski, A. M., Weiss, M. A., Ward, C. W., & Lawrence, M. C. (2013). How insulin engages its primary binding site on the insulin receptor. *Nature*. doi: 10.1038/nature11781

Menting, J. G., Yang, Y., Chan, S. J., Phillips, N. B., Smith, B. J., Whittaker, J., Wickramasinghe, N. P., Whittaker, L. J., Pandyarajan, V., Wan, Z. L., Yadav, S. P., Carroll, J. M., Strokes, N., Roberts, C. T., Ismail-Beigi, F., Milewski, W., Steiner, D. F., Chauhan, V. S., Ward, C. W., … Lawrence, M. C. (2014). Protective hinge in insulin opens to enable its receptor engagement. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1412897111

Michaels, E. K., Chmiel, J. S., Plotkin, B. J., & Schaeffer, A. J. (1983). Effect of D-mannose and D-glucose on Escherichia coli bacteriuria in rats. *Urological Research*. doi: 10.1007/BF00256954

Mirmira, R. G., Nakagawa, S. H., & Tager, H. S. (1991). Importance of the character and configuration of residues B24, B25, and B26 in insulin-receptor interactions. *Journal of Biological Chemistry*. doi: 10.1016/s0021-9258(18)52312-7

Mitchell, R. K., Mondragon, A., Chen, L., Mcginty, J. A., French, P. M., Ferrer, J., Thorens, B., Hodson, D. J., Rutter, G. A., & Da Silva Xavier, G. (2015). Selective disruption of Tcf7l2 in the pancreatic β cell impairs secretory function and lowers β cell mass. *Human Molecular Genetics*. doi: 10.1093/hmg/ddu553

Mitok, K. A., Freiberger, E. C., Schueler, K. L., Rabaglia, M. E., Stapleton, D. S., Kwiecien, N. W., Malec, P. A., Hebert, A. S., Broman, A. T., Kennedy, R. T., Keller, M. P., Coon, J. J., & Attie, A. D. (2018). Islet proteomics reveals genetic variation in dopamine production resulting in altered insulin secretion. *Journal of*

Biological Chemistry. doi: 10.1074/jbc.RA117.001102

Miyachi, Y., Tsuchiya, K., Shiba, K., Mori, K., Komiya, C., Ogasawara, N., & Ogawa, Y. (2018). A reduced M1-like/M2-like ratio of macrophages in healthy adipose tissue expansion during SGLT2 inhibition. *Scientific Reports*. doi: 10.1038/s41598-018-34305-x

Miyamoto, S., & Kollman, P. A. (1992). Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry*. doi: 10.1002/jcc.540130805

Mo, R., Jiang, T., Di, J., Tai, W., & Gu, Z. (2014). Emerging micro- and nanotechnology based synthetic approaches for insulin delivery. In Chemical Society Reviews. doi: 10.1039/c3cs60436e

Monami, M., & Mannucci, E. (2013). Efficacy and safety of degludec insulin: A meta-analysis of randomised trials. *Current Medical Research and Opinion*. doi: 10.1185/03007995.2013.772507

Mooranian, A., Ionescu, C. M., Walker, D., Jones, M., Wagle, S. R., Kovacevic, B., Chester, J., Foster, T., Johnston, E., & Kuthubutheen, J. (2022). Single-Cellular Biological Effects of Cholesterol-Catabolic Bile Acid-Based Nano/Micro Capsules as Anti-Inflammatory Cell Protective Systems. *Biomolecules*, *12*(1), 73.

Mori, M., Kumada, T., Inoue, K., Nozaki, F., Matsui, K., Maruo, Y., Yamada, M., Suzuki, H., Kosaki, K., & Shibata, M. (2021). Ketogenic diet for refractory epilepsy with MEHMO syndrome: Caution for acute necrotizing pancreatitis. *Brain and Development*. doi: 10.1016/j.braindev.2021.02.002

Morrison, M. C., & Kleemann, R. (2015). Role of macrophage migration inhibitory factor in obesity, insulin resistance, type 2 diabetes, and associated hepatic co-morbidities: A comprehensive review of human and rodent studies. In Frontiers in Immunology. doi: 10.3389/fimmu.2015.00308

Motozono, C., Pearson, J. A., De Leenheer, E., Rizkallah, P. J., Beck, K., Trimby, A., Sewell, A. K., Wong, F. S., & Cole, D. K. (2015). Distortion of the major histocompatibility complex class i binding groove to accommodate an insulin-derived 10-mer peptide. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M114.622522

Motterle, A., Gattesco, S., Peyot, M. L., Esguerra, J. L. S., Gomez-Ruiz, A., Laybutt, D. R., Gilon, P., Burdet, F., Ibberson, M., Eliasson, L., Prentki, M., & Regazzi, R. (2017). Identification of islet-enriched long non-coding RNAs contributing to β-cell failure in type 2 diabetes. *Molecular Metabolism*. doi: 10.1016/j.molmet.2017.08.005

Mujaj, B., Bos, D., Kavousi, M., Lugt, A. Van Der, Staessen, J. A., Franco, O. H., & Vernooij, M. W. (2020). Serum insulin levels are associated with vulnerable plaque components in the carotid artery: The Rotterdam Study. *European Journal of Endocrinology*. doi: 10.1530/EJE-19-0620

Muniyappa, R., Warren, M. A., Zhao, X., Aney, S. C., Courville, A. B., Chen, K. Y., Brychta, R. J., Germain-Lee, E. L., Weinstein, L. S., & Skarulis, M. C. (2013). Reduced insulin sensitivity in adults with pseudohypoparathyroidism type 1a. *Journal of Clinical Endocrinology and Metabolism*. doi: 10.1210/jc.2013-1594

Munte, C. E., Vilela, L., Kalbitzer, H. R., & Garratt, R. C. (2005). Solution structure of human proinsulin C-peptide. *FEBS Journal*. doi: 10.1111/j.1742-4658.2005.04843.x

Murakami, H., Ueno, G., Shimizu, N., Kumasaka, T., & Yamamoto, M. (2012). Upgrade of automated sample exchanger SPACE. *Journal of Applied Crystallography*. doi: 10.1107/S0021889812003585

Murray, B. R., Jewell, J. R., Jackson, K. J., Agboola, O., Alexander, B. R., & Sharma, P. (2017). Type III Hypersensitivity Reaction to Subcutaneous Insulin Preparations in a Type 1 Diabetic. *Journal of General Internal Medicine*. doi: 10.1007/s11606-017-4037-7

Musafia, B., Buchner, V., & Arad, D. (1995). Complex salt bridges in proteins: Statistical analysis of structure and function. *Journal of Molecular Biology*. doi: 10.1006/jmbi.1995.0653

Nabel, E. G. (2003). Cardiovascular Disease. *New England Journal of Medicine*, *349*(1), 60–72. doi: 10.1056/NEJMra035098

Nachtergaele, S., & He, C. (2018). Chemical modifications in the life of an mRNA transcript. In Annual Review of Genetics. doi: 10.1146/annurev-genet-120417-031522

Nakane, T., Joti, Y., Tono, K., Yabashi, M., Nango, E., Iwata, S., Ishitani, R., & Nureki, O. (2016). Data processing pipeline for serial femtosecond crystallography at SACLA. *Journal of Applied Crystallography*. doi: 10.1107/S1600576716005720

Nedelkov, D., Niederkofler, E. E., Oran, P. E., Peterman, S., & Nelson, R. W. (2018). Top-down mass spectrometric immunoassay for human insulin and its therapeutic analogs. In Journal of Proteomics. doi: 10.1016/j.jprot.2017.08.001

Nettleton, E. J., Tito, P., Sunde, M., Bouchard, M., Dobson, C. M., & Robinson, C. V. (2000). Characterization of the oligomeric states of insulin in self-assembly and amyloid fibril formation by mass spectrometry. *Biophysical Journal*. doi: 10.1016/S0006-3495(00)76359-4

Nikerel, I. E., Öner, E. T., Kirdar, B., & Yildirim, R. (2006). Optimization of medium composition for biomass production of recombinant Escherichia coli cells using response surface methodology. *Biochemical Engineering Journal*. doi: 10.1016/j.bej.2006.08.009

Niklasson, B., Hörnfeldt, B., & Lundman, B. (1998). Could myocarditis, insulin-dependent diabetes mellitus, and Guillain- Barre syndrome be caused by one or more infectious agents carried by rodents? In Emerging Infectious Diseases. doi: 10.3201/eid0402.980206

Nordquist, L., Brown, R., Fasching, A., Persson, P., & Palm, F. (2009). Proinsulin C-peptide reduces diabetes-induced glomerular hyperfiltration via efferent arteriole dilation and inhibition of tubular sodium reabsorption. *American Journal of Physiology - Renal Physiology*. doi: 10.1152/ajprenal.00228.2009

Norman, A. (2012). *Calcium-binding proteins in health and disease*. Retrieved from https://books.google.com/books?hl=tr&lr=&id=dzGfzDDYhhsC&oi=fnd&pg=PR1&dq=Norman,+A.+(Ed.).+(2012).+Calcium-binding+proteins+in+health+and+disease.+Elsevier&ots=cF5wWnxKrR&sig=CLHiAjf8AuvCSc3ncexBrW3zSS4

Norrman, M., Hubálek, F., & Schluckebier, G. (2007). Structural characterization of insulin NPH formulations. *European Journal of Pharmaceutical Sciences*. doi: 10.1016/j.ejps.2007.01.003

Norrman, M., & Schluckebier, G. (2007). Crystallographic characterization of two

novel crystal forms of human insulin induced by chaotropic agents and a shift in pH. *BMC Structural Biology*. doi: 10.1186/1472-6807-7-83

O'Kusky, J. R., Ye, P., & D'Ercole, A. J. (2000). Insulin-like growth factor-I promotes neurogenesis and synaptogenesis in the hippocampal dentate gyrus during postnatal development. *Journal of Neuroscience*. doi: 10.1523/jneurosci.20-22-08435.2000

Odegard, P. S., & Capoccia, K. L. (2005). Inhaled insulin: Exubera. In Annals of Pharmacotherapy. doi: 10.1345/aph.1E522

Ohkuma, T., Van Gaal, L., Shaw, W., Mahaffey, K. W., de Zeeuw, D., Matthews, D. R., Perkovic, V., & Neal, B. (2020). Clinical outcomes with canagliflozin according to baseline body mass index: results from post hoc analyses of the CANVAS Program. *Diabetes, Obesity and Metabolism*. doi: 10.1111/dom.13920

Oleck, J., Kassam, S., & Goldman, J. D. (2016). Commentary: Why was inhaled insulin a failure in the market? In Diabetes Spectrum. doi: 10.2337/diaspect.29.3.180

Oran, P. E., Jarvis, J. W., Borges, C. R., Sherma, N. D., & Nelson, R. W. (2011). Mass spectrometric immunoassay of intact insulin and related variants for population proteomics studies. *Proteomics - Clinical Applications*. doi: 10.1002/prca.201000112

Orci, L., Ravazzola, M., Amherdt, M., Madsen, O., Perrelet, A., Vassalli, J. D., & Anderson, R. G. W. (1986). Conversion of proinsulin to insulin occurs coordinately with acidification of maturing secretory vesicles. *Journal of Cell Biology*. doi: 10.1083/jcb.103.6.2273

Ormazabal, V., Nair, S., Elfeky, O., Aguayo, C., Salomon, C., & Zuñiga, F. A. (2018). Association between insulin resistance and the development of cardiovascular disease. In Cardiovascular Diabetology. doi: 10.1186/s12933-018-0762-4

Osthege, M., Tenhaef, N., Zyla, R., Müller, C., Hemmerich, J., Wiechert, W., Noack, S., & Oldiges, M. (2022). bletl - A Python package for integrating BioLector microcultivation devices in the Design-Build-Test-Learn cycle. *Engineering in Life Sciences*. doi: 10.1002/elsc.202100108

OTP, Diseases associated with INS | Open Targets Platform. Available from:

https://platform.opentargets.org/target/ENSG00000254647/associat ions?view=t:bubbles

OTP, EMBL-EBI. Available from: https://www.ebi.ac.uk/ols/ontologies/efo/terms/graph?iri= http://www.ebi.ac.uk/efo/EFO_0009605

OTP, https://platform.opentargets.org/evidence/ENSG00000254647/EFO_0000319

OTP, https://platform.opentargets.org/disease/OTAR_0000018

Ozbek, P., Soner, S., & Haliloglu, T. (2013). Hot spots in a network of functional sites. *PloS One*. doi: 10.1371/journal.pone.0074320

Packiam, K. A. R., Ramanan, R. N., Ooi, C. W., Krishnaswamy, L., & Tey, B. T. (2020). Stepwise optimization of recombinant protein production in Escherichia coli utilizing computational and experimental approaches. In Applied Microbiology and Biotechnology. doi: 10.1007/s00253-020-10454-w

Pak, K., Kim, K., Seo, S., Lee, M. J., & Kim, I. J. (2022). Serotonin transporter is negatively associated with body mass index after glucose loading in humans. *Brain Imaging and Behavior*. doi: 10.1007/S11682-021-00600-X

Palivec, V., Viola, C. M., Kozak, M., Ganderton, T. R., Křížková, K., Turkenburg, J. P., Halušková, P., Žáková, L., Jiráček, J., & Jungwirth, P. (2017). Computational and structural evidence for neurotransmitter-mediated modulation of the oligomeric states of human insulin in storage granules. *Journal of Biological Chemistry*, *292*(20), 8342–8355.

Palmieri, L. C., Fávero-Retto, M. P., Lourenço, D., & Lima, L. M. T. R. (2013). A T3R3 hexamer of the human insulin variant B28Asp. *Biophysical Chemistry*. doi: 10.1016/j.bpc.2013.01.003

Palmieri, R., Lee, R. W., & Dunn, M. F. (1988). 1H Fourier Transform NMR Studies of Insulin: Coordination of Ca2+to the Glu(B13) Site Drives Hexamer Assembly and Induces a Conformation Change. *Biochemistry*, *27*(9), 3387–3397. doi: 10.1021/BI00409A040

Pan, S. H., & Malcolm, B. A. (2018). Reduced Background Expression and Improved Plasmid Stability with pET Vectors in BL21 (DE3). *Https://Doi.Org/10.2144/00296st03*, *29*(6), 1234–1238. doi: 10.2144/00296ST03

Pandyarajan, V., Smith, B. J., Phillips, N. B., Whittaker, L., Cox, G. P., Wickramasinghe, N., Menting, J. G., Wan, Z. L., Whittaker, J., Ismail-Beigi, F., Lawrence, M. C., & Weiss, M. A. (2014). Aromatic anchor at an invariant hormone-receptor interface: Function of insulin residue B24 with application to protein design. In Journal of Biological Chemistry. doi: 10.1074/jbc.M114.608562

Papa, V., Pezzino, V., Costantino, A., Belfiore, A., Giuffrida, D., Frittitta, L., Vannelli, G. B., Brand, R., Goldfine, I. D., & Vigneri, R. (1990). Elevated insulin receptor content in human breast cancer. *Journal of Clinical Investigation*. doi: 10.1172/JCI114868

Papaioannou, A., Kuyucak, S., & Kuncic, Z. (2015). Molecular dynamics simulations of insulin: Elucidating the conformational changes that enable its binding. *PLoS ONE*. doi: 10.1371/journal.pone.0144058

Papaneophytou, C. P., & Kontopidis, G. (2014). Statistical approaches to maximize recombinant protein expression in Escherichia coli: A general review. In Protein Expression and Purification. doi: 10.1016/j.pep.2013.10.016

Parfitt, C., Church, D., Armston, A., Couchman, L., Evans, C., Wark, G., & McDonald, T. J. (2015). Commercial insulin immunoassays fail to detect commonly prescribed insulin analogues. *Clinical Biochemistry*. doi: 10.1016/j.clinbiochem.2015.07.017

Park, C. R. (2001). Cognitive effects of insulin in the central nervous system. In Neuroscience and Biobehavioral Reviews. doi: 10.1016/S0149-7634(01)00016-1

Park, I. H., Lerou, P. H., Zhao, R., Huo, H., & Daley, G. Q. (2008). Generation of human-induced pluripotent stem cells. *Nature Protocols*. doi: 10.1038/nprot.2008.92

Parry, S., Hadaschik, D., Blancher, C., Kumaran, M. K., Bochkina, N., Morris, H. R., Richardson, S., Aitman, T. J., Gauguier, D., Siddle, K., Scott, J., & Dell, A. (2006). Glycomics investigation into insulin action. *Biochimica et Biophysica Acta - General Subjects*. doi: 10.1016/j.bbagen.2005.12.013

Patel, P. S., Buras, E. D., & Balasubramanyam, A. (2013). The role of the immune system in obesity and insulin resistance. In Journal of Obesity. doi: 10.1155/2013/616193

Patti, M. E., Virkamäki, A., Landaker, E. J., Kahn, C. R., & Yki-Järvinen, H. (1999). Activation of the hexosamine pathway by glucosamine in vivo induces insulin resistance of early postreceptor insulin signaling events in skeletal muscle. *Diabetes*. doi: 10.2337/diabetes.48.8.1562

Pearce, N. M., & Gros, P. (2021). A method for intuitively extracting macromolecular dynamics from structural disorder. *Nature Communications*. doi: 10.1038/s41467-021-25814-x

Pellegrino, M., Traversi, G., Arena, A., Cappa, M., Rosado, M. M., Andreani, M., Delfino, D. V., Moretti, F., & Fierabracci, A. (2020). Effect of p53 activation through targeting MDM2/MDM4 heterodimer on T regulatory and effector cells in the peripheral blood of Type 1 diabetes patients. *PLoS ONE*. doi: 10.1371/journal.pone.0228296

Peng, S., Yang, J., Wang, Y., Fan, Y., Tang, F., Hou, C., Yu, J., Wang, X., & Jiang, G. (2020). Low-dose intranasal insulin improves cognitive function and suppresses the development of epilepsy. *Brain Research*. doi: 10.1016/j.brainres.2019.146474

Peterson, J. D., Steiner, D. F., Emdin, S. O., & Falkmer, S. (1975). The amino acid sequence of the insulin from a primitive vertebrate, the Atlantic hagfish (Myxine glutinosa). *Journal of Biological Chemistry*. doi: 10.1016/s0021-9258(19)41294-5

Phillips, N. B., Wan, Z. L., Whittaker, L., Hu, S. Q., Huang, K., Hua, Q. X., Whittaker, J., Ismail-Beigi, F., & Weiss, M. A. (2010). Supramolecular protein engineering: Design of zinc-stapled insulin hexamers as a long acting depot. *Journal of Biological Chemistry*. doi: 10.1074/jbc.C110.105825

Pickup, J. C., Day, C., Bailey, C. J., Samuel, A., Chusney, G. D., Garland, H. O., Hamilton, K., & Balment, R. J. (1995). Plasma sialic acid in animal models of diabetes mellitus: Evidence for modulation of sialic acid concentrations by insulin deficiency. *Life Sciences*. doi: 10.1016/0024-3205(95)02096-2

Pillai, O., & Panchagnula, R. (2001). Insulin therapies - Past, present and future. In Drug Discovery Today. doi: 10.1016/S1359-6446(01)01962-6

Plank, J., Wutte, A., Brunner, G., Siebenhofer, A., Semlitsch, B., Sommer, R., Hirschberger, S., & Pieber, T. R. (2002). A direct comparison of insulin aspart and insulin lispro in patients with type 1 diabetes. *Diabetes Care*. doi: 10.2337/diacare.25.11.2053

Pocker, Y., & Biswas, S. B. (1981). Self-Association of Insulin and the Role of Hydrophobic Bonding: A Thermodynamic Model of Insulin Dimerization. *Biochemistry*. doi: 10.1021/bi00518a019

Poduslo, J. F., Curran, G. L., Wengenack, T. M., Malester, B., & Duff, K. (2001). Permeability of proteins at the blood-brain barrier in the normal adult mouse and double transgenic mouse model of Alzheimer's disease. *Neurobiology of Disease*. doi: 10.1006/nbdi.2001.0402

Powell, H. R. (1999). The Rossmann Fourier autoindexing algorithm in MOSFLM. *Acta Crystallographica Section D: Biological Crystallography*. doi: 10.1107/S0907444999009506

Poy, M. N., Hausser, J., Trajkovski, M., Braun, M., Collins, S., Rorsman, P., Zavolan, M., & Stoffel, M. (2009). miR-375 maintains normal pancreatic alpha-and beta-cell mass. *Proceedings of the National Academy of Sciences of the United States of America*.

Prill, H., Luu, A., Yip, B., Holtzinger, J., Lo, M. J., Christianson, T. M., Yogalingam, G., Aoyagi-Scharber, M., LeBowitz, J. H., Crawford, B. E., & Lawrence, R. (2019). Differential Uptake of NAGLU-IGF2 and Unmodified NAGLU in Cellular Models of Sanfilippo Syndrome Type B. *Molecular Therapy - Methods and Clinical Development*. doi: 10.1016/j.omtm.2019.05.008

Proud, C. G. (2006). Regulation of protein synthesis by insulin. *Biochemical Society Transactions*. doi: 10.1042/bst0340213

Prugovečki, B., Pulić, I., Toth, M., & Matković-Čalogović, D. (2012). High resolution structure of the manganese derivative of insulin. *Croatica Chemica Acta*. doi: 10.5562/cca2108

Pugliese, G., Penno, G., Natali, A., Barutta, F., Di Paolo, S., Reboldi, G., Gesualdo, L., & De Nicola, L. (2020). Diabetic kidney disease: new clinical and therapeutic issues. Joint position statement of the Italian Diabetes Society and the Italian

Society of Nephrology on “The natural history of diabetic kidney disease and treatment of hyperglycemia in patients with. In Journal of Nephrology. doi: 10.1007/s40620-019-00650-x

Pullen, R. A., Lindsay, D. G., Wood, S. P., Tickle, I. J., Blundell, T. L., Wollmer, A., Krail, G., Brandenburg, D., Zahn, H., Gliemann, J., & Gammeltoft, S. (1976). Receptor-binding region of insulin. *Nature*. doi: 10.1038/259369a0

Qian, X., Kraft, J., Ni, Y., & Zhao, F. Q. (2014). Production of recombinant human proinsulin in the milk of transgenic mice. *Scientific Reports*. doi: 10.1038/srep06465

Rabinovich, G. A., Baum, L. G., Tinari, N., Paganelli, R., Natoli, C., Liu, F. T., & Iacobelli, S. (2002). Galectins and their ligands: Amplifiers, silencers or tuners of the inflammatory response? In Trends in Immunology. doi: 10.1016/S1471-4906(02)02232-9

Rahuel-Clermont, S., French, C. A., Kaarsholm, N. C., & Dunn, M. F. (1997). Mechanisms of stabilization of the insulin hexamer through allosteric ligand interactions. *Biochemistry*. doi: 10.1021/bi963038q

Ramirez, D. G., Abenojar, E., Hernandez, C., Lorberbaum, D. S., Papazian, L. A., Passman, S., Pham, V., Exner, A. A., & Benninger, R. K. P. (2020). Contrast-enhanced ultrasound with sub-micron sized contrast agents detects insulitis in mouse models of type1 diabetes. *Nature Communications*. doi: 10.1038/s41467-020-15957-8

Rasio, E., Antaki, A., & Van Campenhout, J. (1976). Diabetes Mellitus in Gonadal Dysgenesis: Studies of Insulin and Growth Hormone Secretion. *European Journal of Clinical Investigation*. doi: 10.1111/j.1365-2362.1976.tb00494.x

Ratnakumar, A., Weinhold, N., Mar, J. C., & Riaz, N. (2020). Protein-Protein interactions uncover candidate core genes’ within omnigenic disease networks. *PLoS Genetics*. doi: 10.1371/journal.pgen.1008903

Raurell-Vila, H., Ramos-Rodríguez, M., & Pasquali, L. (2018). Assay for transposase accessible chromatin (ATAC-Seq) to chart the open chromatin landscape of human Pancreatic Islets. In Methods in Molecular Biology. doi: 10.1007/978-1-4939-7768-0_11

Reaven, G. (2012). Insulin resistance and coronary heart disease in nondiabetic individuals. *Arteriosclerosis, Thrombosis, and Vascular Biology*. doi: 10.1161/ATVBAHA.111.241885

Rege, N. K., Wickramasinghe, N. P., Tustan, A. N., Phillips, N. F. B., Yee, V. C., Ismail-Beigi, F., & Weiss, M. A. (2018). Structure-based stabilization of insulin as a therapeutic protein assembly via enhanced aromatic–aromatic interactions. *Journal of Biological Chemistry*. doi: 10.1074/jbc.RA118.003650

Rellier, N., Ruggiero-Lopez, D., Lecomte, M., Lagarde, M., & Wiernsperger, N. (1999). In vitro and in vivo alterations of enzymatic glycosylation in diabetes. *Life Sciences*. doi: 10.1016/S0024-3205(99)00094-6

Renner, R., Pfützner, A., Trautmann, M., Harzer, O., Sauter, K., & Landgraf, R. (1999). Use of insulin lispro in continuous subcutaneous insulin infusion treatment: Results of a multicenter trial. *Diabetes Care*. doi: 10.2337/diacare.22.5.784

Rest, C., Kandanelli, R., & Fernández, G. (2015). Strategies to create hierarchical self-assembled structures via cooperative non-covalent interactions. In Chemical Society Reviews. doi: 10.1039/c4cs00497c

Rhodes, C. J. (2004). Processing of the Insulin Molecule. In Diabetes Mellitus: A Fundamental and Clinical Text.

Rodríguez-rabassa, M., López, P., Sánchez, R., Hernández, C., Rodríguez, C., Rodríguez-santiago, R. E., Orengo, J. C., Green, V., Yamamura, Y., & Rivera-amill, V. (2020). Inflammatory biomarkers, microbiome, depression, and executive dysfunction in alcohol users. *International Journal of Environmental Research and Public Health*. doi: 10.3390/ijerph17030689

Roignant, J. Y., & Soller, M. (2017). m6A in mRNA: An Ancient Mechanism for Fine-Tuning Gene Expression. In Trends in Genetics. doi: 10.1016/j.tig.2017.04.003

Rose, D. P., Komninou, D., & Stephenson, G. D. (2004). Obesity, adipocytokines, and insulin resistance in breast cancer. In Obesity Reviews. doi: 10.1111/j.1467-789X.2004.00142.x

Rosenstock, J., Bain, S. C., Gowda, A., Jódar, E., Liang, B., Lingvay, I., Nishida, T., Trevisan, R., & Mosenzon, O. (2023). Weekly Icodec versus Daily Glargine U100

in Type 2 Diabetes without Previous Insulin. *New England Journal of Medicine*. doi: 10.1056/nejmoa2303208

Rosskamp, R. H., & Park, G. (1999). Long-acting insulin analogs. *Diabetes Care*.

Roy, M., Brader, M. L., Lee, R. W. K., Kaarsholm, N. C., Hansen, J. F., & Dunn, M. F. (1989). Spectroscopic signatures of the T to R conformational transition in the insulin hexamer. *Journal of Biological Chemistry*. doi: 10.1016/s0021-9258(19)47269-4

Ryan, A. S., Hafer-Macko, C., & Ortmeyer, H. K. (2021). Insulin resistance in skeletal muscle of chronic stroke. *Brain Sciences*. doi: 10.3390/brainsci11010020

Ryckaert, J. P., Ciccotti, G., & Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*. doi: 10.1016/0021-9991(77)90098-5

Sachdeva, M. M., Claiborn, K. C., Khoo, C., Yang, J., Groff, D. N., Mirmira, R. G., & Stoffers, D. A. (2009). Pdx1 (MODY4) regulates pancreatic beta cell susceptibility to ER stress. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0904849106

Salihovic, S., Broeckling, C. D., Ganna, A., Prenni, J. E., Sundström, J., Berne, C., Lind, L., Ingelsson, E., Fall, T., Ärnlöv, J., & Nowak, C. (2020). Non-targeted urine metabolomics and associations with prevalent and incident type 2 diabetes. *Scientific Reports*. doi: 10.1038/s41598-020-72456-y

Saltiel, A. R., & Kahn, C. R. (2001). Insulin signalling and the regulation of glucose and lipid metabolism. In Nature. doi: 10.1038/414799a

Samnegård, B., Jacobson, S. H., Jaremko, G., Johansson, B. L., Ekberg, K., Isaksson, B., Eriksson, L., Wahren, J., & Sjöquist, M. (2005). C-peptide prevents glomerular hypertrophy and mesangial matrix expansion in diabetic rats. *Nephrology Dialysis Transplantation*. doi: 10.1093/ndt/gfh683

Sarwar, N., Gao, P., Kondapally Seshasai, S. R., Gobin, R., Kaptoge, S., Di Angelantonio, E., Ingelsson, E., Lawlor, D. A., Selvin, E., Stampfer, M., Stehouwer, C. D. A., Lewington, S., Pennells, L., Thompson, A., Sattar, N., White, I. R., Ray, K. K., Danesh, J., Tipping, R. W., … Wormser, D. (2010). Diabetes

mellitus, fasting blood glucose concentration, and risk of vascular disease: A collaborative meta-analysis of 102 prospective studies. *The Lancet*. doi: 10.1016/S0140-6736(10)60484-9

Sarwar, Nadeem, Sattar, N., Gudnason, V., & Danesh, J. (2007). Circulating concentrations of insulin markers and coronary heart disease: A quantitative review of 19 Western prospective studies. In European Heart Journal. doi: 10.1093/eurheartj/ehm115

Saxena, R., Voight, B. F., Lyssenko, V., Burtt, N. P., De Bakker, P. I. W., Chen, H., Roix, J. J., Kathiresan, S., Hirschhorn, J. N., Daly, M. J., Hughes, T. E., Groop, L., & Altshuler, D. (2007). Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science*. doi: 10.1126/science.1142358

Schlichtkrull, J., Årtun, T., Waterman, H. I., Eliasson, N. A., & Thorell, B. (1956). Insulin Crystals. I. The Minimum Mole-Fraction of Metal in Insulin Crystals Prepared with Zn++, Cd++, Co++, Ni++, Cu++, Mn++, or Fe++. *Acta Chemica Scandinavica*. doi: 10.3891/acta.chem.scand.10-1455

Scott, L. J., Mohlke, K. L., Bonnycastle, L. L., Willer, C. J., Li, Y., Duren, W. L., Erdos, M. R., Stringham, H. M., Chines, P. S., Jackson, A. U., Prokunina-Olsson, L., Ding, C. J., Swift, A. J., Narisu, Hu, T., Pruim, R., Xiao, R., Li, X. Y., Conneely, K. N., … Boehnke, M. (2007). A genome-wide association study of type 2 diabetes in finns detects multiple susceptibility variants. *Science*. doi: 10.1126/science.1142382

Sebastião, I., Candeias, E., Santos, M. S., de Oliveira, C. R., Moreira, P. I., & Duarte, A. I. (2014). Insulin as a bridge between type 2 diabetes and alzheimer disease - how anti-diabetics could be a solution for dementia. In Frontiers in Endocrinology. doi: 10.3389/fendo.2014.00110

Segerstolpe, Å., Palasantza, A., Eliasson, P., Andersson, E. M., Andréasson, A. C., Sun, X., Picelli, S., Sabirsh, A., Clausen, M., Bjursell, M. K., Smith, D. M., Kasper, M., Ämmälä, C., & Sandberg, R. (2016). Single-Cell Transcriptome Profiling of Human Pancreatic Islets in Health and Type 2 Diabetes. *Cell Metabolism*. doi: 10.1016/j.cmet.2016.08.020

Setter, S. M., Corbett, C. F., Campbell, R. K., & White, J. R. (2000). Insulin aspart: A

new rapid-acting insulin analog. In Annals of Pharmacotherapy. doi: 10.1345/aph.19414

Shafqat, J., Juntti-Berggren, L., Zhong, Z., Ekberg, K., Köhler, M., Berggren, P. O., Johansson, J., Wahren, J., & Jörnvall, H. (2002). Proinsulin C-peptide and its analogues induce intracellular Ca2+ increases in human renal tubular cells. *Cellular and Molecular Life Sciences*. doi: 10.1007/s00018-002-8496-5

Shah, S. N., Joshi, S. R., & Parmar, D. V. (1997). History of insulin. *The Journal of the Association of Physicians of India*, 4–9.

Shahbazi, M., Cundiff, P., Zhou, W., Lee, P., Patel, A., D'Souza, S. L., Abbasi, F., Quertermous, T., & Knowles, J. W. (2019). The role of insulin as a key regulator of seeding, proliferation, and mRNA transcription of human pluripotent stem cells. *Stem Cell Research and Therapy*. doi: 10.1186/s13287-019-1319-5

Shahbazmohammadi, H., & Omidinia, E. (2017). Medium Optimization for Improved Production of Dihydrolipohyl Dehydrogenase from Bacillus sphaericus PAD-91 in Escherichia coli. *Molecular Biotechnology*. doi: 10.1007/s12033-017-0013-z

Shelley, K. L., Dixon, T. P. E., Brooks-Bartlett, J. C., & Garman, E. F. (2018). RABDAM: quantifying specific radiation damage in individual protein crystal structures. *Journal of Applied Crystallography*. doi: 10.1107/S1600576718002509

Shen, S. W., Reaven, G. M., & Farquhar, J. W. (1970). Comparison of impedance to insulin-mediated glucose uptake in normal subjects and in subjects with latent diabetes. *The Journal of Clinical Investigation*. doi: 10.1172/JCI106433

Shen, Y., Joachimiak, A., Rich Rosner, M., & Tang, W. J. (2006). Structures of human insulin-degrading enzyme reveal a new substrate recognition mechanism. *Nature*. doi: 10.1038/nature05143

Shigematsu, Y., Hamada, M., Nagai, T., Nishimura, K., Inoue, K., Suzuki, J., Ogimoto, A., & Higaki, J. (2011). Risk for atrial fibrillation in patients with hypertrophic cardiomyopathy: Association with insulin resistance. *Journal of Cardiology*. doi: 10.1016/j.jjcc.2011.03.001

Shiloach, J., Kaufman, J., Guillard, A. S., & Fass, R. (1996). Effect of glucose supply strategy on acetate accumulation, growth, and recombinant protein production by

Escherichia coli BL21 (λDE3) and Escherichia coli JM109. *Biotechnology and Bioengineering*. doi: 10.1002/(SICI)1097-0290(19960220)49:4<421::AID-BIT9>3.0.CO;2-R

Shimazu, Y., Tono, K., Tanaka, T., Yamanaka, Y., Nakane, T., Mori, C., Kimura, K. T., Fujiwara, T., Sugahara, M., Tanaka, R., Doak, R. B., Shimamura, T., Iwata, S., Nango, E., & Yabashi, M. (2019). High-viscosity sample-injection device for serial femtosecond crystallography at atmospheric pressure. *Journal of Applied Crystallography*. doi: 10.1107/S1600576719012846

Shimokawa, C., Kato, T., Takeuchi, T., Ohshima, N., Furuki, T., Ohtsu, Y., Suzue, K., Imai, T., Obi, S., Olia, A., Izumi, T., Sakurai, M., Arakawa, H., Ohno, H., & Hisaeda, H. (2020). CD8+ regulatory T cells are critical in prevention of autoimmune-mediated diabetes. *Nature Communications*. doi: 10.1038/s41467-020-15857-x

Shisheva, A. (2008). Phosphoinositides in insulin action on GLUT4 dynamics: Not just PtdIns(3,4,5)P3. In American Journal of Physiology - Endocrinology and Metabolism. doi: 10.1152/ajpendo.90353.2008

Shokri, A., Sandén, A. M., & Larsson, G. (2003). Cell and process design for targeting of recombinant protein into the culture medium of Escherichia coli. In Applied Microbiology and Biotechnology. doi: 10.1007/s00253-002-1156-8

Shu Jin Chan, Cao, Q. P., & Steiner, D. F. (1990). Evolution of the insulin superfamily: Cloning of a hybrid insulin/insulin-like growth factor cDNA from amphioxus. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.87.23.9319

Siddiqui, H., & Scupola, A. (n.d.). *User Involvement in R&D at Novo Nordisk Diabetes Treatment Development*.

Sierra, R. G., Batyuk, A., Sun, Z., Aquila, A., Hunter, M. S., Lane, T. J., Liang, M., Yoon, C. H., Alonso-Mori, R., Armenta, R., Castagna, J. C., Hollenbeck, M., Osier, T. O., Hayes, M., Aldrich, J., Curtis, R., Koglin, J. E., Rendahl, T., Rodriguez, E., … Boutet, S. (2019). The macromolecular femtosecond crystallography instrument at the linac coherent light source. *Journal of Synchrotron Radiation*. doi: 10.1107/S1600577519001577

Singer, R. A., & Sussel, L. (2018). Islet long noncoding RNAs: A playbook for discovery and characterization. *Diabetes*. doi: 10.2337/dbi18-0001

Smith, G. D., & Ciszak, E. (1994). The structure of a complex of hexameric insulin and 4'-hydroxyacetanilide. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.91.19.8851

Smith, G. D., Ciszak, E., Magrum, L. A., Pangborn, W. A., & Blessing, R. H. (2000). R6 hexameric insulin complexed with m-cresol or resorcinol. *Acta Crystallographica Section D: Biological Crystallography*. doi: 10.1107/S0907444900012749

Smith, G. D., Pangborn, W. A., & Blessing, R. H. (2001). Phase changes in T3R3f human insulin: Temperature or pressure induced? *Acta Crystallographica Section D: Biological Crystallography*, *57*(8), 1091–1100. doi: 10.1107/S0907444901007685

Smith, G. David, & Blessing, R. H. (2003). Lessons from an aged, dried crystal of T6 human insulin. *Acta Crystallographica - Section D Biological Crystallography*, *59*(8), 1384–1394. doi: 10.1107/S090744490301165X

Smith, G. David, Ciszak, E., & Pangborn, W. (1996). A novel complex of a phenolic derivative with insulin: Structural features related to the T → R transition. *Protein Science*. doi: 10.1002/pro.5560050806

Smith, G. David, Pangborn, W. A., & Blessing, R. H. (2003). The structure of T6 human insulin at 1.0 Å resolution. *Acta Crystallographica - Section D Biological Crystallography*. doi: 10.1107/S0907444902023685

Smith, L. F. (1966). Species variation in the amino acid sequence of insulin. *The American Journal of Medicine*. doi: 10.1016/0002-9343(66)90145-8

Sonksen, P. H. (2001). Insulin, growth hormone and sport. *Journal of Endocrinology*. doi: 10.1677/joe.0.1700013

Sopyan, I., Gozali, D., Sriwidodo, & Guntina, R. K. (2022). DESIGN-EXPERT SOFTWARE (DOE): AN APPLICATION TOOL FOR OPTIMIZATION IN PHARMACEUTICAL PREPARATIONS FORMULATION. In International Journal of Applied Pharmaceutics. doi: 10.22159/ijap.2022v14i4.45144

Sparviero, S., Barth, L., Keil, T., Dinter, C., Berg, C., Lattermann, C., & Büchs, J. (2023). Black glucose-releasing silicon elastomer rings for fed-batch operation allow measurement of the oxygen transfer rate from the top and optical signals from the bottom for each well of a microtiter plate. *BMC Biotechnology*. doi: 10.1186/s12896-023-00775-9

Sreekanth, R., Pattabhi, V., & Rajan, S. S. (2008). Structural interpretation of reduced insulin activity as seen in the crystal structure of human Arg-insulin. *Biochimie*. doi: 10.1016/j.biochi.2007.09.012

Sreekanth, R., Pattabhi, V., & Rajan, S. S. (2009). Metal induced structural changes observed in hexameric insulin. *International Journal of Biological Macromolecules*. doi: 10.1016/j.ijbiomac.2008.09.019

Stadinski, B. D., Obst, R., & Huseby, E. S. (2016). A “hotspot” for autoimmune T cells in type 1 diabetes. In Journal of Clinical Investigation. doi: 10.1172/JCI88165

Steensgaard, D. B., Schluckebier, G., Strauss, H. M., Norrman, M., Thomsen, J. K., Friderichsen, A. V., Havelund, S., & Jonassen, I. (2013). Ligand-controlled assembly of hexamers, dihexamers, and linear multihexamer structures by the engineered acylated insulin degludec. *Biochemistry*. doi: 10.1021/bi3008609

Steichen, C., Hannoun, Z., Luce, E., Hauet, T., & Dubart-Kupperschmitt, A. (2019). Genomic integrity of human induced pluripotent stem cells: Reprogramming, differentiation and applications. *World Journal of Stem Cells*. doi: 10.4252/wjsc.v11.i10.729

Steinberger, J., Moorehead, C., Katch, V., & Rocchini, A. P. (1995). Relationship between insulin resistance and abnormal lipid profile in obese adolescents. *The Journal of Pediatrics*. doi: 10.1016/S0022-3476(95)70394-2

Steiner, D. F. (1967). Evidence for a precursor in the biosynthesis of insulin. *Transactions of the New York Academy of Sciences*. doi: 10.1111/j.2164-0947.1967.tb02452.x

Steiner, Donald F. (2008). The biosynthesis of insulin. In Pancreatic Beta Cell in Health and Disease (pp. 31–49). Springer Japan. doi: 10.1007/978-4-431-75452-7_3

Stoffers, D. A., Zinkin, N. T., Stanojevic, V., Clarke, W. L., & Habener, J. F. (1997).

Pancreatic agenesis attributable to a single nucleotide deletion in the human IPF1 gene coding sequence. *Nature Genetics*. doi: 10.1038/ng0197-106

Strickler, S. S., Gribenko, A. V., Gribenko, A. V., Keiffer, T. R., Tomlinson, J., Reihle, T., Loladze, V. V., & Makhatadze, G. I. (2006). Protein stability and surface electrostatics: A charged relationship. *Biochemistry*. doi: 10.1021/bi0600143

Strub, C., Alies, C., Lougarre, A., Ladurantie, C., Czaplicki, J., & Fournier, D. (2004). Mutation of exposed hydrophobic amino acids to arginine to increase protein stability. *BMC Biochemistry*. doi: 10.1186/1471-2091-5-9

Sugahara, M., Mizohata, E., Nango, E., Suzuki, M., Tanaka, T., Masuda, T., Tanaka, R., Shimamura, T., Tanaka, Y., Suno, C., Ihara, K., Pan, D., Kakinouchi, K., Sugiyama, S., Murata, M., Inoue, T., Tono, K., Song, C., Park, J., … Iwata, S. (2014). Grease matrix as a versatile carrier of proteins for serial crystallography. *Nature Methods*. doi: 10.1038/nmeth.3172

Sun, J., Cui, J., He, Q., Chen, Z., Arvan, P., & Liu, M. (2015). Proinsulin misfolding and endoplasmic reticulum stress during the development and progression of diabetes. In Molecular Aspects of Medicine. doi: 10.1016/j.mam.2015.01.001

Sundaralingam, M., Bergstrom, R., Strasburg, G., Rao, S. T., Roychowdhury, P., Greaser, M., & Wang, B. C. (1985). Molecular structure of troponin C from chicken skeletal muscle at 3-angstrom resolution. *Science*. doi: 10.1126/science.3969570

Sundsten, T., & Ortsäter, H. (2009). Proteomics in diabetes research. *Molecular and Cellular Endocrinology*. doi: 10.1016/j.mce.2008.06.018

Syring, K. E., Boortz, K. A., Oeser, J. K., Ustione, A., Platt, K. A., Shadoan, M. K., McGuinness, O. P., Piston, D. W., Powell, D. R., & O'Brien, R. M. (2016). Combined deletion of Slc30a7 and Slc30a8 unmasks a critical role for ZnT8 in glucose-stimulated insulin secretion. *Endocrinology*. doi: 10.1210/en.2016-1573

Szebenyi, D. M. E., & Moffat, K. (1986). The refined structure of vitamin D-dependent calcium-binding protein from bovine intestine. Molecular details, ion binding, and implications for the structure of other calcium-binding proteins. *Journal of Biological Chemistry*. doi: 10.1016/s0021-9258(19)84447-2

Tai, J., Liu, W., Li, Y., Li, L., & Hölscher, C. (2018). Neuroprotective effects of a triple GLP-1/GIP/glucagon receptor agonist in the APP/PS1 transgenic mouse model of Alzheimer's disease. *Brain Research*. doi: 10.1016/j.brainres.2017.10.012

Taketomi, S., & Iwatsuka, H. (1979). Effects of a synthetic hexapeptide (β-Ala-Arg-Gly-Phe-Phe-Tyr-NH2) on insulin binding and glucose metabolism of rat adipocytes. *Hormone and Metabolic Research*. doi: 10.1055/s-0028-1092769

Tandon, H., de Brevern, A. G., & Srinivasan, N. (2021). Transient association between proteins elicits alteration of dynamics at sites far away from interfaces. *Structure*. doi: 10.1016/j.str.2020.11.015

Tang, L., Whittingham, J. L., Verma, C. S., Caves, L. S. D., & Dodson, G. G. (1999). Structural consequences of the B5 histidine → tyrosine mutation in human insulin characterized by X-ray crystallography and conformational analysis. *Biochemistry*. doi: 10.1021/bi990700k

Taylor, S. K., Tran, T. H., Liu, M. Z., Harris, P. E., Sun, Y., Jambawalikar, S. R., Tong, L., & Stojanovic, M. N. (2018). Insulin Hexamer-Caged Gadolinium Ion as MRI Contrast-o-phore. *Chemistry - A European Journal*. doi: 10.1002/chem.201801388

Tegel, H., Tourle, S., Ottosson, J., & Persson, A. (2010). Increased levels of recombinant human proteins with the Escherichia coli strain Rosetta(DE3). *Protein Expression and Purification*. doi: 10.1016/j.pep.2009.08.017

Tekola-Ayele, F., Lee, A., Workalemahu, T., Zhang, W., Shrestha, D., Amare, A. T., & Ouidir, M. (2019). Genetic overlap between birthweight and adult cardiometabolic diseases has implications for genomic medicine. *Scientific Reports*. doi: 10.1038/s41598-019-40834-w

Thim, L., Hansen, M. T., Norris, K., Hoegh, I., Boel, E., Forstrom, J., Ammerer, G., & Fiil, N. P. (1986). Secretion and processing of insulin precursors in yeast. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.83.18.6766

Thorsøe, K. S., Schlein, M., Steensgaard, D. B., Brandt, J., Schluckebier, G., & Naver, H. (2010). Kinetic evidence for the sequential association of insulin binding sites 1 and 2 to the insulin receptor and the influence of receptor isoform. *Biochemistry*. doi: 10.1021/bi1000118

Timofeev, V. I., Chuprov-Netochin, R. N., Samigina, V. R., Bezuglov, V. V., Miroshnikov, K. A., & Kuranova, I. P. (2010). X-ray investigation of gene-engineered human insulin crystallized from a solution containing polysialic acid. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*. doi: 10.1107/S1744309110000461

Tiran, J., Avruch, L. I., & Albisser, A. M. (1979). A circulation and organs model for insulin dynamics. *American Journal of Physiology Endocrinology Metabolism and Gastrointestinal Physiology*. doi: 10.1152/ajpendo.1979.237.4.e331

Toeroek, C., Cserjan-Puschmann, M., Bayer, K., & Striedner, G. (2015). Fed-batch like cultivation in a micro-bioreactor: screening conditions relevant for Escherichia coli based production processes. *SpringerPlus*. doi: 10.1186/s40064-015-1313-z

Tokarz, V. L., MacDonald, P. E., & Klip, A. (2018). The cell biology of systemic insulin function. In Journal of Cell Biology. doi: 10.1083/jcb.201802095

Tono, K., Togashi, T., Inubushi, Y., Sato, T., Katayama, T., Ogawa, K., Ohashi, H., Kimura, H., Takahashi, S., Takeshita, K., Tomizawa, H., Goto, S., Ishikawa, T., & Yabashi, M. (2013). Beamline, experimental stations and photon beam diagnostics for the hard x-ray free electron laser of SACLA. *New Journal of Physics*. doi: 10.1088/1367-2630/15/8/083035

Tono, Kensuke, Nango, E., Sugahara, M., Song, C., Park, J., Tanaka, T., Tanaka, R., Joti, Y., Kameshima, T., Ono, S., Hatsui, T., Mizohata, E., Suzuki, M., Shimamura, T., Tanaka, Y., Iwata, S., & Yabashi, M. (2015). Diverse application platform for hard X-ray diffraction in SACLA (DAPHNIS): Application to serial protein crystallography using an X-ray free-electron laser. *Journal of Synchrotron Radiation*. doi: 10.1107/S1600577515004464

Toth, C., Brussee, V., Martinez, J. A., McDonald, D., Cunningham, F. A., & Zochodne, D. W. (2006). Rescue and regeneration of injured peripheral nerve axons by intrathecal insulin. *Neuroscience*. doi: 10.1016/j.neuroscience.2005.11.065

Townsend, M., Mehta, T., & Selkoe, D. J. (2007). Soluble Aβ inhibits specific signal transduction cascades common to the insulin receptor pathway. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M610390200

Turunen, O., Vuorio, M., Fenel, F., engineering, M. L.-P., & 2002, undefined. (2002).

Engineering of multiple arginines into the Ser/Thr surface of Trichoderma reesei endo-1,4-β-xylanase II increases the thermotolerance and shifts the pH optimum. *Academic.Oup.ComO Turunen, M Vuorio, F Fenel, M LeisolaProtein Engineering, 2002•academic.Oup.Com*, *15*(2), 141–145. Retrieved from https://academic.oup.com/peds/article-abstract/15/2/141/1501278

Ukkonen, K., Vasala, A., Ojamo, H., & Neubauer, P. (2011). High-yield production of biologically active recombinant protein in shake flask culture by combination of enzyme-based glucose delivery and increased oxygen transfer. *Microbial Cell Factories*. doi: 10.1186/1475-2859-10-107

US20120214199A1 - Lis-pro proinsulin compositions and methods of producing lis-pro insulin analogs therefrom - Google Patents. (n.d.). Retrieved from https://patents.google.com/patent/US20120214199A1/en

US7790677B2 - Insulin production methods and pro-insulin constructs - Google Patents. https://patents.google.com/patent/US7790677B2/en.

US20120214965A1 - Glargine proinsulin and methods of producing glargine insulin analogs therefrom - Google Patents. (n.d.). Retrieved from https://patents.google.com/patent/US20120214965A1/en?oq=US20120214965A1

US5700662A - Process for preparing insulin analogs - Google Patents. (n.d.). Retrieved from https://patents.google.com/patent/US5700662A/en

US6221633B1 - Insulin derivatives having a rapid onset of action - Google Patents. (n.d.). Retrieved from https://patents.google.com/patent/US6221633B1/en

US7193035B2 - Crystals of insulin analogs and processes for their preparation - Google Patents. (n.d.). Retrieved from https://patents.google.com/patent/US7193035B2/en

US7790677B2 - Insulin production methods and pro-insulin constructs - Google Patents. (n.d.). Retrieved from https://patents.google.com/patent/US7790677B2/en

Vagin, A., & Lebedev, A. (2015). MoRDa , an automatic molecular replacement pipeline . *Acta Crystallographica Section A Foundations and Advances*. doi: 10.1107/s2053273315099672

van Hoek, I., Hodgkiss-Geere, H., Bode, E. F., Hamilton-Elliott, J., Mõtsküla, P.,

Palermo, V., Pereira, Y. M., Culshaw, G. J., Ivanova, A., & Dukes-McEwan, J. (2020). Associations among echocardiography, cardiac biomarkers, insulin metabolism, morphology, and inflammation in cats with asymptomatic hypertrophic cardiomyopathy. *Journal of Veterinary Internal Medicine*. doi: 10.1111/jvim.15730

Van Lierop, B., Ong, S. C., Belgi, A., Delaine, C., Andrikopoulos, S., Haworth, N. L., Menting, J. G., Lawrence, M. C., Robinson, A. J., & Forbes, B. E. (2017). Insulin in motion: The A6-A11 disulfide bond allosterically modulates structural transitions required for insulin activity. *Scientific Reports*. doi: 10.1038/s41598-017-16876-3

Vecchio, I., Tornali, C., Bragazzi, N. L., & Martini, M. (2018). The discovery of insulin: An important milestone in the history of medicine. In Frontiers in Endocrinology. doi: 10.3389/fendo.2018.00613

Vernede, X., Lavault, B., Ohana, J., Nurizzo, D., Joly, J., Jacquamet, L., Felisaz, F., Cipriani, F., & Bourgeois, D. (2006). UV laser-excited fluorescence as a tool for the visualization of protein crystals mounted in loops. *Acta Crystallographica Section D: Biological Crystallography*. doi: 10.1107/S0907444905041429

Vigneri, R., Sciacca, L., & Vigneri, P. (2020). Rethinking the Relationship between Insulin and Cancer. In Trends in Endocrinology and Metabolism. doi: 10.1016/j.tem.2020.05.004

Viková, J., Collinsová, M., Kletvíková, E., Budšínský, M., Kaplan, V., Záková, L., Veverka, V., Hexnerová, R., Aviñó, R. J. T., Straková, J., Selicharová, I., Vank, V., Wright, D. W., Watson, C. J., Turkenburg, J. P., Brzozowski, A. M., & Jiráek, J. (2016). Rational steering of insulin binding specificity by intra-chain chemical crosslinking. *Scientific Reports*. doi: 10.1038/srep19431

Vinther, T. N., Norrman, M., Ribel, U., Huus, K., Schlein, M., Steensgaard, D. B., Pedersen, T. Å., Pettersson, I., Ludvigsen, S., Kjeldsen, T., Jensen, K. J., & Hubálek, F. (2013). Insulin analog with additional disulfide bond has increased stability and preserved activity. *Protein Science*. doi: 10.1002/pro.2211

Vinther, T. N., Norrman, M., Strauss, H. M., Huus, K., Schlein, M., Pedersen, T. A., Kjeldsen, T., Jensen, K. J., & Hubálek, F. (2012). Novel covalently linked insulin

dimer engineered to investigate the function of insulin dimerization. *PLoS ONE*. doi: 10.1371/journal.pone.0030882

Volkmar, M., Dedeurwaerder, S., Cunha, D. A., Ndlovu, M. N., Defrance, M., Deplus, R., Calonne, E., Volkmar, U., Igoillo-Esteve, M., Naamane, N., Del Guerra, S., Masini, M., Bugliani, M., Marchetti, P., Cnop, M., Eizirik, D. L., & Fuks, F. (2012). DNA methylation profiling identifies epigenetic dysregulation in pancreatic islets from type 2 diabetic patients. *EMBO Journal*. doi: 10.1038/emboj.2011.503

Volkov, P., Bacos, K., Ofori, J. K., Esguerra, J. L. S., Eliasson, L., Rönn, T., & Ling, C. (2017). Whole-Genome bisulfite sequencing of human pancreatic islets reveals novel differentially methylated regions in type 2 diabetes pathogenesis. *Diabetes*. doi: 10.2337/db16-0996

Von Dreele, R. B., Stephens, P. W., Smith, G. D., & Blessing, R. H. (2000). The first protein crystal structure determined from high-resolution X-ray powder diffraction data: A variant of T3R3 human insulin-zinc complex produced by grinding. *Acta Crystallographica Section D: Biological Crystallography*. doi: 10.1107/S0907444900013901

Vora, J., Cariou, B., Evans, M., Gross, J. L., Harris, S., Landstedt-Hallin, L., Mithal, A., Rodriguez, M. R., & Meneghini, L. (2015). Clinical use of insulin degludec. In Diabetes Research and Clinical Practice. doi: 10.1016/j.diabres.2015.04.002

Wagner, A., Diez, J., Schulze-Briese, C., & Schluckebier, G. (2009). Crystal structure of Ultralente - A microcrystalline insulin suspension. *Proteins: Structure, Function and Bioinformatics*. doi: 10.1002/prot.22213

Wahren, J., Ekberg, K., & Jörnvall, H. (2007). C-peptide is a bioactive peptide. In Diabetologia. doi: 10.1007/s00125-006-0559-y

Wahren, John, Ekberg, K., Johansson, J., Henriksson, M., Pramanik, A., Johansson, B. L., Rigler, R., & Jörnvall, H. (2000). Role of C-peptide in human physiology. In American Journal of Physiology - Endocrinology and Metabolism. doi: 10.1152/ajpendo.2000.278.5.e759

Wahren, John, Kallas, Å., & Sima, A. A. F. (2012). The clinical potential of C-peptide replacement in type 1 diabetes. In Diabetes. doi: 10.2337/db11-1423

Wahren, John, Shafqat, J., Johansson, J., Chibalin, A., Ekberg, K., & Jörnvall, H. (2004). Molecular and cellular effects of C-peptide - New perspectives on an old peptide. In Experimental Diabesity Research. doi: 10.1080/15438600490424479

Wan, Z. L., Huang, K., Xu, B., Hu, S. Q., Wang, S., Chu, Y. C., Katsoyannis, P. G., & Weiss, M. A. (2005). Diabetes-associated mutations in human insulin: Crystal structure and photo-cross-linking studies of A-chain variant insulin Wakayama. *Biochemistry*. doi: 10.1021/bi047585k

Wan, Z. L., Xu, B., Chu, Y. C., Katsoyannis, P. G., & Weiss, M. A. (2003). Crystal Structure of allo-IleA2-Insulin, an Inactive Chiral Analogue: Implications for the Mechanism of Receptor Binding. *Biochemistry*. doi: 10.1021/bi034430o

Wan, Z., Xu, B., Huang, K., Chu, Y. C., Li, B., Nakagawa, S. H., Qu, Y., Hu, S. Q., Katsoyannis, P. G., & Weiss, M. A. (2004). Enhancing the activity of insulin at the receptor interface: Crystal structure and photo-cross-linking of A8 analogues. *Biochemistry*. doi: 10.1021/bi048223f

Wang, C., Eufemi, M., Turano, C., & Giartosio, A. (1996). Influence of the carbohydrate moiety on the stability of glycoproteins. *Biochemistry*. doi: 10.1021/bi9517704

Wang, S., Wei, W., Zheng, Y., Hou, J., Dou, Y., Zhang, S., Luo, X., & Cai, X. (2012). The Role of Insulin C-Peptide in the Coevolution Analyses of the Insulin Signaling Pathway: A Hint for Its Functions. *PLoS ONE*. doi: 10.1371/journal.pone.0052847

Watkins, P. J., Amiel, S. A., Howell, S. L., & Turner, E. (2003). *Diabetes and its Management*. doi: 10.1002/9780470760277

Wei, W., Tian, H., Fu, X., Yao, R., & Su, D. (2020). Participates in vertical sleeve gastrectomy for type ii diabetes mellitus by regulating TGR5. *Medical Science Monitor*. doi: 10.12659/MSM.920628

Weil-Ktorza, O., Rege, N., Lansky, S., Shalev, D. E., Shoham, G., Weiss, M. A., & Metanis, N. (2019). Substitution of an Internal Disulfide Bridge with a Diselenide Enhances both Foldability and Stability of Human Insulin. *Chemistry (Weinheim an Der Bergstrasse, Germany)*. doi: 10.1002/chem.201900892

Weisberg, S. P., McCann, D., Desai, M., Rosenbaum, M., Leibel, R. L., & Ferrante, A.

W. (2003). Obesity is associated with macrophage accumulation in adipose tissue. *Journal of Clinical Investigation*. doi: 10.1172/jci19246

Weiss, M. A. (2009). Chapter 2 The Structure and Function of Insulin. Decoding the TR Transition. In Vitamins and Hormones. doi: 10.1016/S0083-6729(08)00602-X

Weiss, M. A., Wan, Z., Zhao, M., Chu, Y. C., Nakagawa, S. H., Burke, G. T., Jia, W., Hellmich, R., & Katsoyannis, P. G. (2002). Non-standard insulin design: Structure-activity relationships at the periphery of the insulin receptor. *Journal of Molecular Biology*. doi: 10.1006/jmbi.2001.5224

White, T. A. (2019). Processing serial crystallography data with crystFEL: A step-by-step guide. *Acta Crystallographica Section D: Structural Biology*. doi: 10.1107/S205979831801238X

White, T. A., Kirian, R. A., Martin, A. V., Aquila, A., Nass, K., Barty, A., & Chapman, H. N. (2012). CrystFEL: A software suite for snapshot serial crystallography. *Journal of Applied Crystallography*. doi: 10.1107/S0021889812002312

Whittingham, J. L., Edwards, D. J., Antson, A. A., Clarkson, J. M., & Dodson, G. G. (1998). Interactions of phenol and m-cresol in the insulin hexamer, and their effect on the association properties of B28 Pro → Asp insulin analogues. *Biochemistry*. doi: 10.1021/bi980807s

Whittingham, J. L., Havelund, S., & Jonassen, I. (1997). Crystal structure of a prolonged-acting insulin with albumin-binding properties. *Biochemistry*. doi: 10.1021/bi9625105

Whittingham, J. L., Jonassen, I., Havelund, S., Roberts, S. M., Dodson, E. J., Verma, C. S., Wilkinson, A. J., & Dodson, G. G. (2004). Crystallographic and solution studies of N-lithocholyl insulin: A new generation of prolonged-acting human insulins. *Biochemistry*. doi: 10.1021/bi036163s

Whittingham, J. L., Youshang, Z., Žáková, L., Dodson, E. J., Turkenburg, J. P., Brange, J., & Dodson, G. G. (2006). I222 crystal form of despentapeptide (B26-B30) insulin provides new insights into the properties of monomeric insulin. *Acta Crystallographica Section D: Biological Crystallography*. doi: 10.1107/S0907444906006871

Wiese, T. J., Dunlap, J. A., & Yorek, M. A. (1997). Effect of L-fucose and D-glucose concentration on L-fucoprotein metabolism in human Hep G2 cells and changes in fucosyltransferase and α-L-fucosidase activity in liver of diabetic rats. *Biochimica et Biophysica Acta - General Subjects*. doi: 10.1016/S0304-4165(96)00123-7

Williams, M. D., Joglekar, M. V., Satoor, S. N., Wong, W., Keramidaris, E., Rixon, A., O'Connell, P., Hawthorne, W. J., Mitchell, G. M., & Hardikar, A. A. (2019). Epigenetic and Transcriptome Profiling Identifies a Population of Visceral Adipose-Derived Progenitor Cells with the Potential to Differentiate into an Endocrine Pancreatic Lineage. *Cell Transplantation*. doi: 10.1177/0963689718808472

Willmer, T., Johnson, R., Louw, J., & Pheiffer, C. (2018). Blood-Based DNA Methylation Biomarkers for Type 2 Diabetes: Potential for Clinical Applications. In Frontiers in Endocrinology. doi: 10.3389/fendo.2018.00744

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., & Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. In Acta Crystallographica Section D: Biological Crystallography. doi: 10.1107/S0907444910045749

WO2012115637A1 - Compositions de pro-insuline aspart et procédés de production d'analogues de l'insuline aspart - Google Patents. (n.d.). Retrieved from https://patents.google.com/patent/WO2012115637A1/fr

WO2015059302A1 - Stable formulation of insulin glulisine - Google Patents. (n.d.). Retrieved from https://patents.google.com/patent/WO2015059302A1/en

WO2015071368A1 - Long-acting insulin glargine analogue - Google Patents. (n.d.). Retrieved from https://patents.google.com/patent/WO2015071368A1/tr

WO2015138548A1 - Purifying insulin using cation exchange and reverse phase chromatography in the presence of an organic modifier and elevated temperature - Google Patents. (n.d.). Retrieved from https://patents.google.com/patent/WO2015138548A1/en

WO2016144658A1 - Process for preparing recombinant insulin using microfiltration -

Google Patents. (n.d.). Retrieved from https://patents.google.com/patent/WO2016144658A1/en

WO2017040363A1 - A process for obtaining insulin with correctly formed disulfide bonds - Google Patents. (n.d.). Retrieved from https://patents.google.com/patent/WO2017040363A1/en

WO2017126984A1 - A method for producing insulin and insulin derivatives, and hybrid peptide used in this method - Google Patents. (n.d.). Retrieved from https://patents.google.com/patent/WO2017126984A1/en

Wormald, M. R., Petrescu, A. J., Pao, Y. L., Glithero, A., Elliott, T., & Dwek, R. A. (2002). Conformational studies of oligosaccharides and glycopeptides: Complementarity of NMR, X-ray crystallography, and molecular modelling. *Chemical Reviews*. doi: 10.1021/cr990368i

Wu, T., Betty, B., Downie, M., Khanolkar, M., Kilov, G., Orr-Walker, B., Senator, G., & Fulcher, G. (2015). Practical Guidance on the Use of Premix Insulin Analogs in Initiating, Intensifying, or Switching Insulin Regimens in Type 2 Diabetes. *Diabetes Therapy*. doi: 10.1007/s13300-015-0116-0

Wu, Y., Zeng, J., Zhang, F., Zhu, Z., Qi, T., Zheng, Z., Lloyd-Jones, L. R., Marioni, R. E., Martin, N. G., Montgomery, G. W., Deary, I. J., Wray, N. R., Visscher, P. M., McRae, A. F., & Yang, J. (2018). Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nature Communications*. doi: 10.1038/s41467-018-03371-0

Xia, C. Q., Zhang, P., Li, S., Yuan, L., Xia, T., Xie, C., & Clare-Salzler, M. J. (2014). C-Abl inhibitor imatinib enhances insulin production by β cells: C-Abl negatively regulates insulin production via interfering with the expression of NKx2.2 and GLUT-2. *PLoS ONE*. doi: 10.1371/journal.pone.0097694

Xia, J., Zhong, Y., Huang, G., Chen, Y., Shi, H., & Zhang, Z. (2012). The relationship between insulin resistance and osteoporosis in elderly male type 2 diabetes mellitus and diabetic nephropathy. *Annales d'Endocrinologie*. doi: 10.1016/j.ando.2012.09.009

Xin, Y., Kim, J., Okamoto, H., Ni, M., Wei, Y., Adler, C., Murphy, A. J., Yancopoulos, G. D., Lin, C., & Gromada, J. (2016). RNA Sequencing of Single Human Islet

Cells Reveals Type 2 Diabetes Genes. *Cell Metabolism*. doi: 10.1016/j.cmet.2016.08.018

Xiong, X., Blakely, A., Karra, P., Vandenberg, M. A., Ghabash, G., Whitby, F., Zhang, Y. W., Webber, M. J., Holland, W. L., Hill, C. P., & Chou, D. H. C. (2020). Novel four-disulfide insulin analog with high aggregation stability and potency. *Chemical Science*. doi: 10.1039/c9sc04555d

Xiong, X., Menting, J. G., Disotuar, M. M., Smith, N. A., Delaine, C. A., Ghabash, G., Agrawal, R., Wang, X., He, X., Fisher, S. J., MacRaild, C. A., Norton, R. S., Gajewiak, J., Forbes, B. E., Smith, B. J., Safavi-Hemami, H., Olivera, B., Lawrence, M. C., & Chou, D. H. C. (2020). A structurally minimized yet fully active insulin based on cone-snail venom insulin principles. *Nature Structural and Molecular Biology*. doi: 10.1038/s41594-020-0430-8

Xue, A., Wu, Y., Zhu, Z., Zhang, F., Kemper, K. E., Zheng, Z., Yengo, L., Lloyd-Jones, L. R., Sidorenko, J., Wu, Y., Agbessi, M., Ahsan, H., Alves, I., Andiappan, A., Awadalla, P., Battle, A., Beutner, F., Bonder, M. J. J., Boomsma, D., … Yang, J. (2018). Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. *Nature Communications*. doi: 10.1038/s41467-018-04951-w

Yan, Z., Wu, H., Zhou, H., Chen, S., He, Y., Zhang, W., Chen, T., Yao, H., & Su, W. (2020). Integrated metabolomics and gut microbiome to the effects and mechanisms of naoxintong capsule on type 2 diabetes in rats. *Scientific Reports*. doi: 10.1038/s41598-020-67362-2

Yang, B. T., Dayeh, T. A., Kirkpatrick, C. L., Taneera, J., Kumar, R., Groop, L., Wollheim, C. B., Nitert, M. D., & Ling, C. (2011a). Insulin promoter DNA methylation correlates negatively with insulin gene expression and positively with HbA1c levels in human pancreatic islets. *Diabetologia*, *54*(2), 360–367. doi: 10.1007/S00125-010-1967-6/FIGURES/5

Yang, B. T., Dayeh, T. A., Kirkpatrick, C. L., Taneera, J., Kumar, R., Groop, L., Wollheim, C. B., Nitert, M. D., & Ling, C. (2011b). Insulin promoter DNA methylation correlates negatively with insulin gene expression and positively with HbA1c levels in human pancreatic islets. *Diabetologia*. doi: 10.1007/s00125-010-

1967-6

Yang, L. W., Rader, A. J., Liu, X., Jursa, C. J., Chen, S. C., Karimi, H. A., & Bahar, I. (2006). oGNM: Online computation of structural dynamics using the Gaussian Network Model. *Nucleic Acids Research*. doi: 10.1093/nar/gkl084

Yao, Z. P., Zeng, Z. H., Li, H. M., Zhang, Y., Feng, Y. M., & Wang, D. C. (1999). Structure of an insulin dimer in an orthorhombic crystal: The structure analysis of a human insulin mutant (B9 Ser→Glu). *Acta Crystallographica Section D: Biological Crystallography*. doi: 10.1107/S0907444999008562

Ye, J., Chang, W., & Liang, D. (2001). Crystal structure of destripeptide (B28-B30) insulin: Implications for insulin dissociation. *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*. doi: 10.1016/S0167-4838(01)00160-1

Ye, P., Xing, Y., Dai, Z., & D'Ercole, A. J. (1996). In vivo actions of insulin-like growth factor-I (IGF-I) on cerebellum development in transgenic mice: Evidence that IGF-I increases proliferation of granule cell progenitors. *Developmental Brain Research*. doi: 10.1016/0165-3806(96)00492-0

Yefanov, O., Mariani, V., Gati, C., White, T. A., Chapman, H. N., & Barty, A. (2015). Accurate determination of segmented X-ray detector geometry. *Optics Express*. doi: 10.1364/oe.23.028459

Yeo, L., Pujol-Autonell, I., Baptista, R., Eichmann, M., Kronenberg-Versteeg, D., Heck, S., Dolton, G., Sewell, A. K., Härkönen, T., Mikk, M. L., Toppari, J., Veijola, R., Knip, M., Ilonen, J., & Peakman, M. (2020). Circulating β cell-specific CD8+ T cells restricted by high-risk HLA class I molecules show antigen experience in children with and at risk of type 1 diabetes. *Clinical and Experimental Immunology*. doi: 10.1111/cei.13391

Yokota, K., Satou, K., & Ohki, S. ya. (2006). Comparative analysis of protein thermostability: Differences in amino acid content and substitution at the surfaces and in the core regions of thermophilic and mesophilic proteins. *Science and Technology of Advanced Materials*. doi: 10.1016/j.stam.2006.03.003

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R., Slukvin, I. I., & Thomson, J.

A. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science*. doi: 10.1126/science.1151526

Yunn, N. O., Kim, J., Ryu, S. H., & Cho, Y. (2023). A stepwise activation model for the insulin receptor. In Experimental and Molecular Medicine. doi: 10.1038/s12276-023-01101-1

Zahn, D. A. (1975). An empirical study of the half-normal plot. *Technometrics*. doi: 10.1080/00401706.1975.10489303

Žáková, L., Brynda, J., Au-Alvarez, O., Dodson, E. J., Dodson, G. G., Whittingham, J. L., & Brzozowski, A. M. (2004). Toward the Insulin−IGF-I Intermediate Structures: Functional and Structural Properties of the [Tyr ^{B25} NMePhe ^{B26}] Insulin Mutant †. *Biochemistry*, *43*(51). doi: 10.1021/bi048856u

Žáková, L., Kletvíková, E., Lepšík, M., Collinsová, M., Watson, C. J., Turkenburg, J. P., Jiráček, J., & Brzozowski, A. M. (2014). Human insulin analogues modified at the B26 site reveal a hormone conformation that is undetected in the receptor complex. *Acta Crystallographica Section D: Biological Crystallography*. doi: 10.1107/S1399004714017775

Žáková, L., Kletvíková, E., Veverka, V., Lepšík, M., Watson, C. J., Turkenburg, J. P., Jiráček, J., & Brzozowski, A. M. (2013). Structural integrity of the B24 site in human insulin is important for hormone functionality. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M112.448050

Zaman, S., Patel, B., Glynne, P., Vanderpump, M., Alsafi, A., Khan, S., Flora, R., Palazzo, F., & Wernig, F. (2020). A rare cause of severe cushing's syndrome. *Endocrinology, Diabetes and Metabolism Case Reports*. doi: 10.1530/EDM-20-0011

Zambon, A. A., & Muntoni, F. (2021). Congenital muscular dystrophies: What is new? In Neuromuscular Disorders. doi: 10.1016/j.nmd.2021.07.009

Zare, H., Sadeghi, H. M. M., & Akbari, V. (2019). Optimization of Fermentation Conditions for Reteplase Expression by Escherichia coli Using Response Surface Methodology. *Avicenna Journal of Medical Biotechnology*, *11*(2), 162. Retrieved from /pmc/articles/PMC6490408/

Zaykov, A. N., Mayer, J. P., & Dimarchi, R. D. (2016). Pursuit of a perfect insulin. In Nature Reviews Drug Discovery. doi: 10.1038/nrd.2015.36

Zhang, Yan, Doruker, P., Kaynak, B., Zhang, S., Krieger, J., Li, H., & Bahar, I. (2020). Intrinsic dynamics is evolutionarily optimized to enable allosteric behavior. In Current Opinion in Structural Biology. doi: 10.1016/j.sbi.2019.11.002

Zhang, Yumin, Zhang, S., & Wang, G. (2015). Metabolomic biomarkers in diabetic kidney diseases - A systematic review. In Journal of Diabetes and its Complications. doi: 10.1016/j.jdiacomp.2015.06.016

Zhang, Z., Liang, W. G., Bailey, L. J., Tan, Y. Z., Wei, H., Wang, A., Farcasanu, M., Woods, V. A., McCord, L. A., Lee, D., Shang, W., Deprez-Poulain, R., Deprez, B., Liu, D. R., Koide, A., Koide, S., Kossiakoff, A. A., Li, S., Carragher, B., … Tang, W. J. (2018). Ensemble cryoEM elucidates the mechanism of insulin capture and degradation by human insulin degrading enzyme. *ELife*. doi: 10.7554/eLife.33572

Zhao, M., Wan, Z. I., Whittaker, L., Xu, B., Phillips, N. B., Katsoyannis, P. G., Ismail-Beigi, F., Whittaker, J., & Weiss, M. A. (2009). Design of an insulin analog with enhanced receptor binding selectivity: Rationale, structure, and therapeutic implications. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M109.028399

Zhao, X., Gang, X., Liu, Y., Sun, C., Han, Q., & Wang, G. (2016). Using Metabolomic Profiles as Biomarkers for Insulin Resistance in Childhood Obesity: A Systematic Review. In Journal of Diabetes Research. doi: 10.1155/2016/8160545

Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C. M., Li, C. J., Vågbø, C. B., Shi, Y., Wang, W. L., Song, S. H., Lu, Z., Bosmans, R. P. G., Dai, Q., Hao, Y. J., Yang, X., Zhao, W. M., Tong, W. M., Wang, X. J., Bogdan, F., … He, C. (2013). ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility. *Molecular Cell*. doi: 10.1016/j.molcel.2012.10.015

Zheng, Y., Wu, C., Yang, J., Zhao, Y., Jia, H., Xue, M., Xu, D., Yang, F., Fu, D., Wang, C., Hu, B., Zhang, Z., Li, T., Yan, S., Wang, X., Nelson, P. J., Bruns, C., Qin, L., & Dong, Q. (2020). Insulin-like growth factor 1-induced enolase 2 deacetylation by HDAC3 promotes metastasis of pancreatic cancer. *Signal Transduction and Targeted Therapy*. doi: 10.1038/s41392-020-0146-6

Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M. R., Powell, J. E., Montgomery, G.

W., Goddard, M. E., Wray, N. R., Visscher, P. M., & Yang, J. (2016). Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nature Genetics*. doi: 10.1038/ng.3538

Zieliński, M., Romanik-Chruścielewska, A., Mikiewicz, D., Łukasiewicz, N., Sokołowska, I., Antosik, J., Sobolewska-Ruta, A., Bierczyńska-Krzysik, A., Zaleski, P., & Płucienniczak, A. (2019). Expression and purification of recombinant human insulin from E. coli 20 strain. *Protein Expression and Purification*. doi: 10.1016/j.pep.2019.02.002

Zinman, B., Tildesley, H., Chiasson, J. L., Tsui, E., & Strack, T. (1997). Insulin lispro in CSII: Results of a double-blind crossover study. *Diabetes*. doi: 10.2337/diabetes.46.3.440