Structural Investigations of the Kelch-like-ECH-Associated Protein 1 at Near-Physiological Conditions and Computational Analysis for New Therapeutics Discovery

by

Merve Yılmaz

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

Master of Science

in

Molecular Biology and Genetics

KOÇ ÜNİVERSİTESİ

September 17, 2024

Structural Investigations of the Kelch-like-ECH-Associated Protein 1 at Near-Physiological Conditions and Computational Analysis for New Therapeutics Discovery

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis by

Merve Yılmaz

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:

_____

Asst. Prof. Hasan Demirci

_____

Asst.Prof. Serkan Kır

_____

Assoc. Prof. Ayşe Özge Kürkçüoğlu Levitas

Date: _____

To my family...

ABSTRACT

Structural Investigations of the Kelch-like-ECH-Associated Protein 1 at Near-Physiological Conditions and Computational Analysis for New Therapeutics Discovery

Merve Yılmaz

Master of Science in Molecular Biology and Genetics

September 17, 2024

Kelch-like-ECH-associated protein 1 (Keap1) has a vital role in an important signal transduction pathway in mammalian cells, where it functions to segregate the nuclear factor erythroid 2-related factor 2 (Nrf2) protein transcription factor in the cytosol, targeting it for degradation by the 26S proteasome. Keap1 consists of an N-terminal region (NTR, residues 1–49), a Broad complex, Tramtrack, and Bric-a-Brac domain (BTB, residues 50–179) domain, an intervening region (IVR, residues 180–314), a Kelch domain/double glycine repeats (DGR, residues 315-599), and C-terminal region (CTR, residues 599-624).

Keap1 homodimerized through its BTB domain, and homodimerization is critical for Keap1 functionality. As a part of the E3-ubiquitin ligase complex, Keap1 interacts with Cullin3 (Cul3) via the BTB domain. The Kelch domain of Keap1 contains six repeated Kelch motifs, each about 45–55 residues in length, forming a β-propeller structure essential for dual binding Nrf2. Understanding Keap1-Nrf2 interaction is vital because Nrf2 protein regulation leads to a coordinated antioxidant and anti-inflammatory response in the cell. Dimethyl fumarate (DMF), one of the approved drugs for treating Multiple Sclerosis, surrounded Keap1 through its Kelch and BTB domain, releasing Nrf2 and facilitating its activation.

Conducting advanced structure determination technique 'X-ray crystallography' at Turkish Light Source, "Turkish Delight," the first ambient and dimeric Kelch domain structure, was solved. The conformational differences in the Nrf2 and DMF binding site within the Kelch domain, resulting from temperature shifts from cryogenic to ambient, were investigated. Furthermore, serial crystallography studies were performed further to elucidate the dynamic behavior of Keap1 under different conditions. Additionally, molecular docking studies were conducted using this new

structure to investigate various electrophilic irreversible inhibitors of Keap1/Nrf2, including DMF, as well as potential direct non-covalent inhibitors.

In conclusion, the comprehensive structural analysis of Keap1 releaved insights into its dynamics and interactions with associated proteins and drugs. The molecular docking studies reveal the potential of the novel compound, a hybrid of *L*-carnosine and *L*-histidyl hydrazide (CNN), as a Keap1/Nrf2 inhibitor. These findings open new avenues for research into targeted therapies that could modulate the Keap1-Nrf2 pathway, potentially offering benefits in treating oxidative stress and inflammation conditions.

ÖZETÇE

Kelch Benzeri ECH İlişkili Protein 1'in Fizyolojik Şartlara Yakın Koşullarda Yapısal İncelemeleri ve Yeni Terapötik Keşifler İçin Hesaplamalı Analizleri

Merve Yılmaz

Moleküler Biyoloji ve Genetik, Yüksek Lisans

17 Eylül 2024

Kelch-ECH-benzeri ilişkili protein 1 (Keap1), memeli hücrelerinde önemli bir sinyal iletim yolunda kritik bir rol oynar. Bu proteinin ana işlevi, nükleer faktör erythroid 2 ile ilişkili faktör 2 (Nrf2) adlı transkripsiyon faktörünü sitozolde tutmak ve onu 26S proteazom aracılığıyla parçalanmaya hedeflemektir. Keap1, N-terminal bölgesi (NTR, kalıntılar 1–49), Broad kompleks, Tramtrack ve Bric-a-Brac (BTB, kalıntılar 50–179) domaini, Intervening domaini (IVR, kalıntılar 180–314), Kelch domaini/çift glisin tekrarları (DGR, kalıntılar 315-599) ve C-terminal bölgesi (CTR, kalıntılar 599-624) olmak üzere birkaç farklı bölgeden oluşur. Keap1, BTB domaini aracılığıyla homodimerik bir yapı oluşturur.

Keap1 BTB domaini aracılığıyla homodimerize olur ve homodimerizasyon Keap1'in işlevi için kritik bir öneme sahiptir. Bir E3-ubikuitin ligaz kompleksinin parçası olarak, Keap1, BTB domaini aracılığıyla Cullin3 (Cul3) ile etkileşime girer. Keap1'in Kelch domaini, her biri yaklaşık 45-55 kalıntı uzunluğunda olan altı tekrarlayan dizilim motifinden oluşur ve Nrf2'nin çift bağlanması için gerekli olan beta-propeller yapısını oluşturur. Nrf2'nin aktivasyonu, koordine bir antioksidan ve anti-inflamatuar yanıt oluşturduğundan, Keap1'in Nrf2 aktivasyonunu baskılaması bu etkileşimin anlaşılmasını kritik hale getirir. Multiple skleroz tedavisi için onaylanmış bir ilaç olan Dimetil fumarat (DMF), Keap1 ile Kelch domaini üzerinden etkileşime girer, Nrf2'yi serbest bırakarak onun aktivasyonunu sağlar.

İleri bir yapı belirleme tekniği olan 'X-ışını kristalografisi' kullanılarak Türkiye'nin ilk ışık kaynağı olan "Turkish DeLight"ta, ortam koşullarında ve dimerik Kelch domaininin yapısı çözüldü. Kelch domaininin Nrf2 ve DMF bağlanma bölgelerinde, kriyojenik sıcaklıklardan ortam sıcaklıklarına geçiş nedeniyle meydana gelen konformasyonel değişiklikler incelendi. Ayrıca, Keap1'in farklı koşullar altındaki di-

namik davranışını daha iyi anlamak için seri kristalografi çalışmaları gerçekleştirildi. Ayrıca, DMF de dahil olmak üzere Keap1/Nrf2'nin çeşitli elektrofilik geri dönüşümsüz inhibitörlerinin yanı sıra potansiyel doğrudan kovalent olmayan inhibitörleri araştırmak için bu yeni yapı kullanılarak moleküler yerleştirme çalışmaları yapılmıştır.

Sonuç olarak, Keap1'in kapsamlı analizi, yapısal dinamikleri ve etkileşimleri hakkında önemli bulgular ortaya koydu. Moleküler yerleştirme çalışmaları, *L*-karnosin ve *L*-histidil hidrazidin bir melezi olan CNN bileşiğinin bir Keap1/Nrf2 inhibitörü olarak potansiyelini ortaya koymaktadır. Keap1'in ortam yapısının anlaşılması, onun konformasyonel esnekliği ve Nrf2 aktivitesini düzenlemedeki rolünü vurguladı. Bu bulgular, Keap1-Nrf2 yolunu hedefleyen tedaviler üzerine yeni araştırma yolları açarak, oksidatif stres ve inflamasyon koşullarının tedavisinde potansiyel faydalar sunabilir.

ACKNOWLEDGMENTS

I am very grateful to Asst. Prof. Hasan Demirci whose support and guidance have been crucial to my academic progress. His academic approach and guidance encouraged me to pursue excellence and overcome research challenges. I would also like to thank the members of my thesis committee, Asst. Prof. Serkan Kır and Assoc. Prof. Ayşe Özge Kürkçüoğlu Levitas. I would also like to thank the members of KUYBIIG-M for their cooperation, friendship, and support. I would especially like to thank my friends Alaleh Shafiei, Betül Ertem, Mehmet Gül, Nilüfer Baldır, Özüm Tosunoğlu, Serra Tavlı, and Ömür Güven for being such a great friend and always supporting me. Additionally, I would like to thank Prof. Türkan Haliloğlu, Asst Prof. Halil Ibrahim Çiftçi, Assoc. Prof. Belgin Sever and Yiğit Kutlu for their support of my research. I cannot express my gratefulness to my dear family, who have been my unfailing source of support and love throughout my academic journey. I extend my deepest gratitude to my mother, Ayşe; my father, Tahir; my little sisters, Rümeysa and Ecrin; and my brother, Abdullah. Their encouragement, sacrifices, and endless love have been the driving force behind my accomplishments.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ADT	AutodockTool
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic Acid
BTB	Broad complex, Tramtrack, and Bric-a-Brac domain
CTR	C-Terminal Region
Cul3	Cullin3
CNN	a hybrid of *L*-carnosine and *L*-histidyl hydrazide
Cys	Cysteine
DC	DGR and CTR
ddH_2O	Deionized distilled water
DGR	Double Glycine Repeat
DNA	Deoxyribonucleic acid
DMF	Dimethyl Fumarate
FPLC	Fast Protein Liquid Chromatography
FUM	Fumarate
Gln	Glutamine
Glu	Glutamic Acid
Gly	Glycine
HECT	Homologous to E6AP C-Terminus
His	Histidine
IVR	Intervening Region
Ile	Isoleucine
ITA	Itaconate

IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	Kilodalton
Keap1	Kelch-like-ECH-Associated Protein-1
L	Liter
LB	Lysogeny Broth
Leu	Leucine
MEF	Monoethyl Fumarate
MMF	Monomethyl Fumarate
ml	Milliliter
mM	Millimolar
MS	Multiple Sclerosis
Neh2	NRF2-ECH homology-2
Ni-NTA	Nickel-Nitrilotriacetic Acid
NMR2	Nuclear Magnetic Resonance
Nrf2	Nuclear Factor Erythroid 2–Related Factor 2
NTR	N-Terminal Region
OD	Optical Density
Phe	Phenylalanine
PTM	Post Translational Modifications
RBR	RING-between-RING
RING	Really Interesting New Gene
RMSD	Root Mean Square Deviation
RNA	Ribonucleic acid
ROS	Reactive Oxygen Species
rpm	Revolutions per Minute

SDS	Sodium Dodecylsulfate
SES	Size Exclusion Chromatography
SDS-PAGE	SDS-Polyacrylamide Gel Electrophoresis
Ser	Serine
Thr	Threonine
TLS	Translation/Libration/Screw
Tyr	Tyrosine
UV	Ultraviolet
Val	Valine
XRD	X-ray Diffractometer
μl	Microliter
Å	Angstrom

Chapter 1

INTRODUCTION

1.1 Ubiquitination

Post-translational modifications (PTMs) are essential in protein function, regulating their stability, localization, and function through various chemical changes. Ubiquitination is one of the fundamental PTMs in eukaryotic cells (Peng et al., 2023). Most PTMs, such as phosphorylation, glycosylation, methylation, or acetylation, involve additional small chemical groups. In contrast, ubiquitination is conducted with the small protein (8 kDa) known as ubiquitin, which attaches to target proteins and enables recognition and carrying of the substrate to the degradation by 26S proteasome (Morrow et al., 2015).

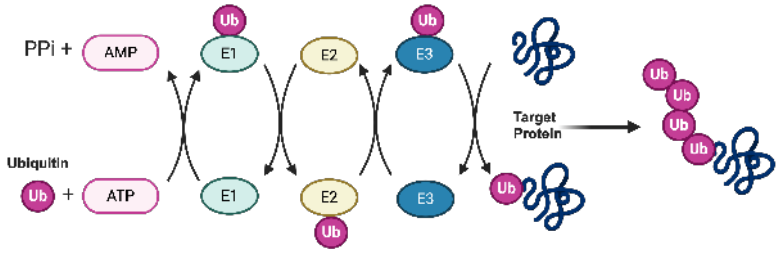

Figure 1.1: Schematic representation of the ubiquitination process

The ubiquitination process includes a series of enzymatic cascades containing three essential enzymes E1s, E2s, and E3 ubiquitin ligases. In the first step, E1 activated ubiquitin in an ATP-dependent manner by forming a thioester bond between its active cysteine site and the C terminus of ubiquitin. Afterward, activated ubiquitin is transferred to the E2 enzyme from E1 and binds to an active cysteine residue of E2, forming a thioester bond similarly. In the final step, the E3 ubiquitin

ligase recognizes the target protein and enables the transfer of ubiquitin to its specific lysine residues; thus, ubiquitin forms a covalent bond with the target protein and marks it for proteasomal degradation (Figure 1.1) (Li et al., 2016, Mevissen and Komande, 2017).

1.1.1 E3-Ubiquitin Ligase Complexes

E3 ubiquitin ligase complexes are enzymes that are required for the specificity of the ubiquitination process in recognizing target proteinsE3 ligases have substrate-binding domains that allow them to selectively identify proteins based on specific degradation signals, called degrons. In eukaryotic cells, highly diverse E3 ligase types have been identified, and this diversity ensures the precise regulation of cellular processes. E3 ligases might be classified into four types: Homologous to E6AP C-Terminus (HECT) type, U-box type, Really Interesting New Gene (RING)- finger type, and RING-between-RING (RBR) type (Yang et al., 2021). RING-type ligases are one of the most prominent and abundant members of E3-ubiquitin ligases. Unlike HECT-type ligases, RING-type E3 ligases directly transfer the ubiquitin to the target protein without intermediate ubiquitin-conjugated product on themselves. Instead, RING ligases bridge the E2 ubiquitin conjugation enzyme and the target protein; thus, ubiquitin can transfer to the target protein directly from E2 (Zheng and Shabek, 2017).

Structurally, RING complexes have a RING motif, a protein structure stabilized with zinc ions. The RING motif facilitates interaction with the E2 ubiquitin-conjugating enzyme. RING complexes can comprise single-unit (monomeric) or multiple proteins (multimeric). For instance, Cullin-RING ligases (CRLs) are huge complexes containing multiple subunits, allowing them to recognize and regulate a broad range of target proteins (Metzger et al., 2014). The Cul3-Keap1-RING complex coordinates the ubiquitination of the transcription factor Nuclear Factor Erythroid 2–Related Factor 2 (Nrf2) via its components: Kelch-like ECH-associated protein 1 (Keap1), Cullin-3 (Cul3), and a RING protein. Cul3 interacts with Keap1 and RING protein by acting as the backbone of the complex. Keap1 is the recog-

nition unit, mainly targeting the Nrf2 protein for ubiquitination under basal conditions. The RING protein enables the transferring of ubiquitin from the E2 ubiquitin-conjugating enzyme to Nrf2 (Wang et al., 2024). However, under oxidative stress conditions, the interaction of Keap1 and Nrf2 is inhibited by modifications in the Nrf2-binding domain of Keap1.

1.2 The Kelch-like-ECH-Associated Protein-1

Keap1 is a critical component of cellular defense mechanisms and a complex homodimeric protein that regulates multiple biological processes. Keap1 protects cells by maintaining homeostasis against oxidative stress (Baird and Yamamoto, 2020). The important role of Keap1 is to regulate the transcription factor, Nrf2 . Keap1 retains the Nrf2 protein in the cytoplasm and facilitates proteasomal degradation of Nrf2. Thus, Keap1 inhibits the translocation of Nrf2 to the nucleus and prevents the activation of antioxidant genes (Taguchi et al., 2012). However, when oxidative stress levels increase in the cell, Keap1 undergoes structural changes that enable Nrf2 to enter the nucleus. Nrf2 protein that is translocated to the nucleus can initiate the transcription of antioxidant genes . This process allows the cell to cope with oxidative stress and minimize cellular damage (Belezza et al., 2018). Keap1 consists of diverse functional domains, and the structure of Keap1 protein plays an important role in its functionality. Keap1 (also known as KLHL19) consists of 624 amino acids with a molecular mass of 70 kDa (Figure 1.2-a) .It exists as a homodimeric structure in the cell and consists of five distinct domains, each with different functions in ubiquitination. The conserved domains of Keap1 are N-terminal region (NTR), a Broad complex, Tramtrack, and Bric-a-Brac domain (BTB), intervening region (IVR), a Kelch domain/double glycine repeats (DGR) and C-terminal region (CTR), according to their placement in the polypeptide chain (Figure 1.2-b-c) (Canning et al., 2015).

First, the BTB domain located at the N-terminus of Keap1 facilitates Keap1 homodimerization and its interaction with the Cul3 ubiquitin ligase complex. This domain enables proteasomal degradation of the Nrf2 protein by ubiquitinating it

(Adamson et al., 2023). The IVR domain, located between the BTB and Kelch domains, modulates the Keap1-Nrf2 interaction. The Kelch or DGR domain mediates the degradation of Nrf2 by binding to it (Canning et al., 2015).

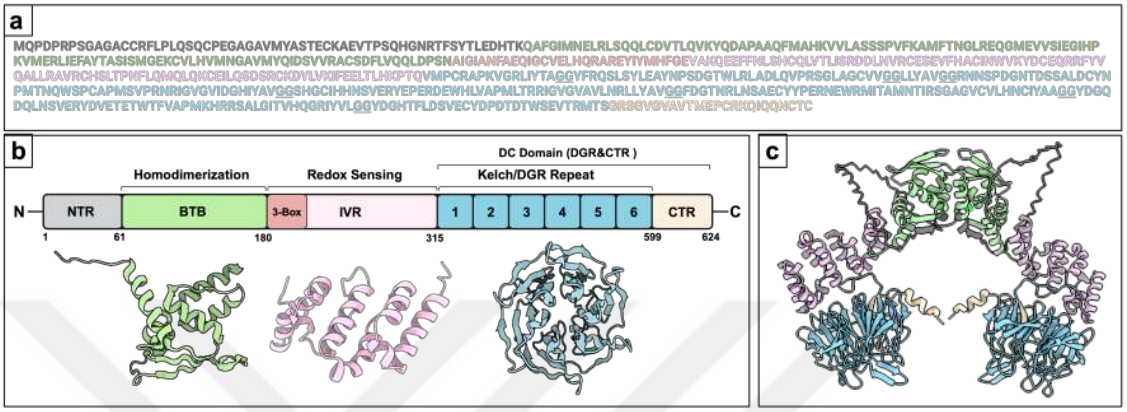

Figure 1.2: Structural representation of Keap1 (a) Sequence information of full-length Keap1, (b) Domain architecture of Keap (c) Illustration of full-length homodimer Keap1 structure by using *Alphafold2*.

Keap1 can also undergo various post-translational modifications, particularly at critical cysteine residues. Under conditions of oxidative stress, these cysteine residues can be modified, resulting in structural changes in Keap1 that regulate its interaction with Nrf2. Cellular oxidative stress is sensed through these cysteine residues (Song et al., 2024). The Keap1-Nrf2 pathway is important not only for protection against oxidative stress but also for other biological processes such as inflammation, apoptosis, and cellular proliferation. While overactivation of Nrf2 can lead to chemotherapy resistance in some cancers, loss-of-function mutations in Keap1 can disrupt its interaction with Nrf2, resulting in an excessive antioxidant response. Therefore, Keap1 function must be critically regulated to maintain cellular balance and ensure cell survival (Wu et al., 2019). Keap1 is also involved in several signaling pathways beyond its structural and functional properties related to Nrf2. In addition, modifications in Keap1 can affect signaling pathways in the cell by interacting with different proteins. These interactions demonstrate that Keap1 has a broad range of functions beyond its interaction with Nrf2.

1.2.1 The BTB Domain and Its Role in Keap1 Homodimerization

BTB proteins are known as Broad-complex, Tramtrack, and Bric-a-brac/poxvirus and zinc finger (BTB/POZ) family. The BTB/POZ proteins contain a zinc finger motif in their structures, which gives its name to the protein family. Zinc finger proteins have many functions, including protein-protein interactions, DNA/RNA binding, and apoptosis regulation. Broad-complex, Tramtrack (ttk), and Bric-a-brac (bab) are genes that encode zinc fingers (bab locus consists of bab1 and bab2 genes). The structure of zinc finger modules varies as much as their functions (Chaharbakshi and Jemc, 2016). The BTB domain of Keap1 is essential for establishing interactions with various proteins such as Cul3, including Keap1 itself; it promotes the homodimerization of the full-length Keap1 structure (Zipper and Mulcahy, 2002).

The crystal structure of BTB domain has six α-helices around three-stranded beta sheets that face each other. Each having a different role, the contribution of α1 helix of chain A, α2 and α3 helices of chain B to dimerization is to mediate connections between two BTB monomers (Chauhan et al., 2013) .

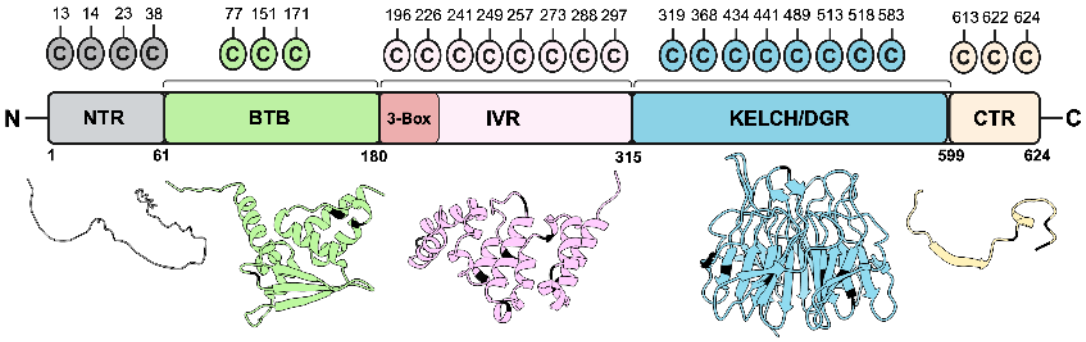

Figure 1.3: Illustration showing the specific locations of cysteine residues within the distinct domains of Keap1

The BTB domain of Keap1 contains three cysteine residues: Cys77, Cys151, and Cys171 (Qu et al., 2022) (Figure 1.3). The cysteine residues of the BTB domain, particularly Cys151, have a crucial role in sensing electrophiles and oxidants. This sensing capability enables a cysteine-based sensor for Keap1. Cys151 modifications by electrophiles lead to the loss of Keap1 repressor function due to Nrf2 activation

promotion. The post-translational modifications that Cys151 undergoes are significant for the functional variability of Keap1. For example, the single mutation of Cys151 to Serine does not display the slower migration similar to wild type Keap1 in SDS-PAGE. These results suggest that Cys151 is involved in modifications that affect the stability and function of Keap1 (Zhang et al., 2003, Saito et al.,2016).

The BTB domain is crucial for mediating the homodimerization of Keap1. *Zipper et al.,* show that wild-type Keap1 demonstrates dimeric form through BTB domain function. However, single mutation Serine 104 to Alanine in the BTB domain, Keap1, can not be dimerized. These results strongly emphasize the critical role of BTB in the homodimerization of Keap1 (Zipper and Mulcahy, 2002). Previous studies demonstrate that the BTB domain forms the stem-like region of the Keap1 and connects to short arm linkers that lead to the large spheres in the structure (Ogura et al.,2010). The BTB domain contributes to the overall architecture of Keap1 by forming a forked-stem structure, which is a required arrangement for the proper assembly of Keap1 homodimerization. The forked-stem structure is represented by two thin layers with a gap between them, demonstrating that each layer originates from a separate Keap1 monomer. The C-terminal part of the BTB domain serves as a short linker between the BTB domain and the spheres formed by the IVR and Kelch domains. This positioning suggests that the BTB domain plays a role in maintaining the spatial organization of Keap1's various functional domains (Ogura et al., 2010). As a part of the E3-Ubiqutin ligase complex, substrate adaptor protein Keap1 requires interaction with Cul3 through the BTB domain during basal conditions. The BTB domain of Keap1 has a conserved 3-box motif from the C-terminal side, which is found in many Cul3-dependent E3 ubiquitin ligase complexes (Figure 1.2-b) (Canning et al., 2013).

1.2.2 The IVR Domain of Keap1

The IVR domain (180 to 315) is located between the Kelch domain and the BTB domain of Keap1 as a short linker that connects the two domains. The flexibility of the IVR domain might contribute to asymmetry for the homodimerization of

Keap1. The IVR domain is attached to the outer surface of the Kelch domain rather than forming a long helical linker. This close attachment suggests that the IVR domain plays a critical role in maintaining the structural integrity of the Keap1 dimer by stabilizing the connection between the BTB and Kelch domains (Ogura et al., 2010). The crystal structure of the IVR domain of Keap1 has not been determined yet. The IVR domain of Keap1 comprises 8 cystein residues: Cys196, Cys226, Cys241, Cys249, Cys257, Cys273, Cys288 and Cys297, of which, Cys273 and Cys288 are well studied (Figure 1.3). When Cys273 and C288 are targeted with electrophiles or alkylating agents, Nrf2 protein is activated as a result of Keap1 inactivation (Dayalan and Dinkova-Kostova, 2020, Wu et al., 2010).

1.2.3 The Kelch Domain of Keap1

The first crystal structure of the human Keap1 Kelch domain, determined at a resolution of 1.85 Å, was reported (Li et al., 2004). To date, several crystal structures of both the human and murine Keap1 Kelch domains with different resolutions and in combination with compounds or short peptide sequences of the Neh2 domain of NRF2 have since been reported (Lo et al., 2006). The Kelch domain of Keap1 comprises six Kelch repeats that form a six-bladed β-propeller structure. Four-stranded antiparallel β-sheets form one blade, with the shortest β-sheet at the central core. The Kelch domain, also called the DGR domain, which includes conserved double glycine at the terminal end of the β-sheets. The difference between the kelch motif and other β -propeller secondary structure subtypes is that the kelch motif is a highly conserved sequence, including a double glycine following strand B, a tyrosine in strand C, and tryptophan in strand D (Figure 1.4) (Li et al., 2004).

Even though sequence identity differs between Kelch repeats, each Kelch repeat is relatively similar regarding secondary structure (Figure 1.4-c). As a result of the superposition of the six blades of the Kelch β -propeller, highly conserved residues were revealed in the Kelch motif. Seven conserved residues are located in six blades, two of which are tyrosine residues in strand C and tryptophan in strand D, which help define the Kelch motif. These highly conserved tyrosine and tryp-

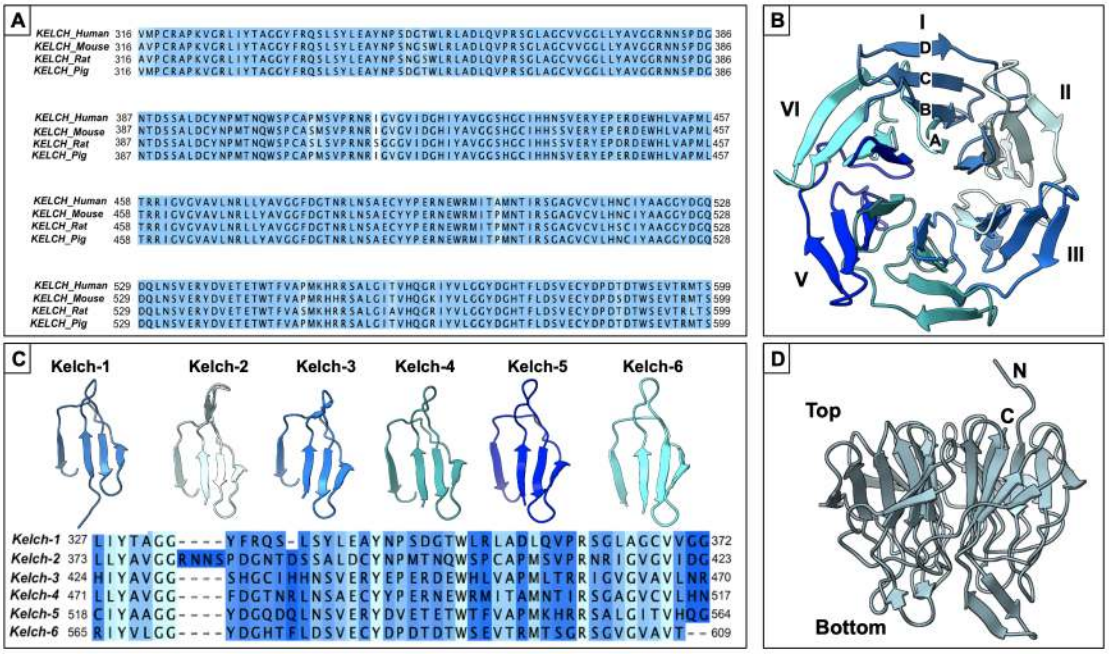

Figure 1.4: Structural representation and multiple sequence alignment of the Kelch domain of Keap1. (a) Multiple sequence alignment of the Kelch domain across different species: Human, Mouse, Rat, and Pig. (b) Illustration of the six Kelch repeats within the Kelch domain, highlighting the β-strands A-B-C-D. (c) Detailed view of each Kelch repeat along with their sequence alignment. (d) Top and bottom views of the Kelch domain structure based on the PDB ID: 8X34.

tophan residues' side chains contribute to forming the Kelch domain's hydrophobic core. They are also intimately involved in linking the blades of the β -propeller (Beamer et al., 2005). Overall, since the Kelch domain contains a DGR that terminates the β strand and a CTR that completes the β -propeller structure, both of which are key elements in maintaining the fold and function of the domain, the Kelch domain is also described as the DC (DGR and CTR) domain (Canning et al., 2015).

The Kelch domain has evolutionarily conserved nine cysteine residues which are highly reactive due to their sulfhydryl (thiol) groups: Cys319, Cys368, Cys395, Cys406, Cys434, Cys489, Cys513, Cys518, and Cys583 (Figure 1.3) (Shilovsky and Dibrova, 2023). The conservation of Cysteine residues in many species shows that they might have critical functions that have been maintained throughout evolution.

Cysteines flanked by basic residues (like Lys, Arg, and His) are more susceptible to post-translational modifications in response to oxidative stress (Dinkova-Kostova et al., 2017). Studies show that sulforaphane-modified Cysteine residues in purified recombinant human Keap1, mainly in the Kelch domain, are Cys489, Cys513, Cys518, and Cys583 (Hu et al., 2011).

The N-terminal Neh2 domain of NRF2 bind to the KEAP1 Kelch domain. Nrf2 protein has evolutionarily conserved DLG (Asp-Leu-Gly) and ETGE (Glu-Thr-Gly-Glu) motifs within the Neh2 domain which were responsible for interacting with Kelch domain of Keap1 (Figure 1.5) (Tong et al., 2006).

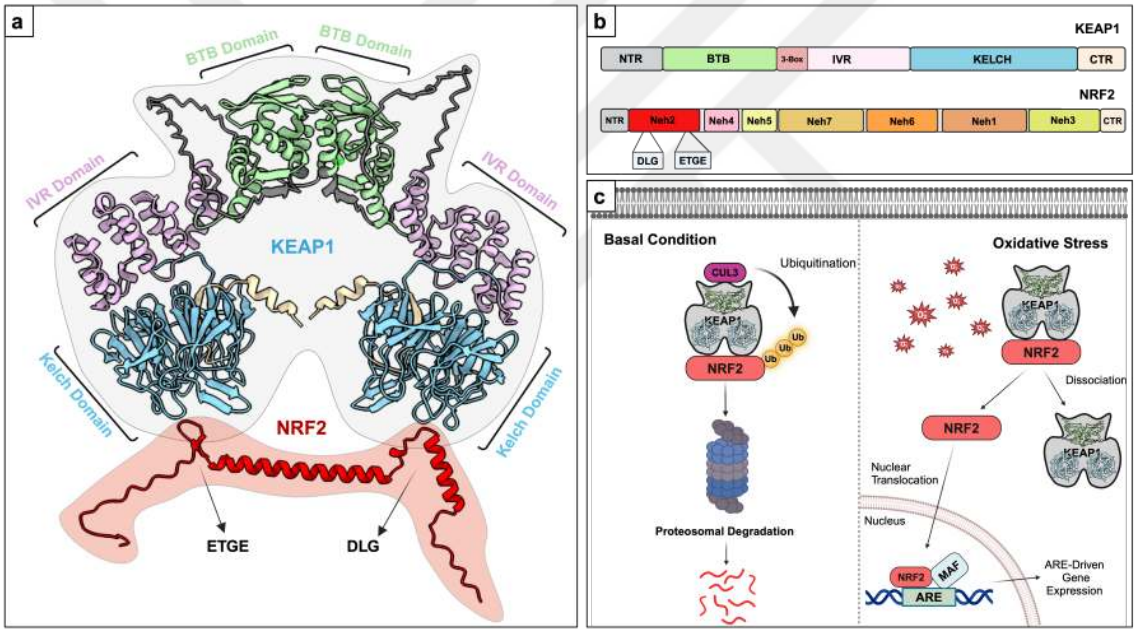

Figure 1.5: Detailed illustration of the Keap1-Nrf2 Pathway. (a) Visualization of the full-length dimeric Keap1 protein, predicted using *AlphaFold2*, with bound Nrf2 ETGE and DLG degrons. (b) Schematic representation of the domain architecture of Keap1 and Nrf2, highlighting key functional regions. (c) Overview of the Nrf2-Keap1 pathway under basal conditions and during oxidative stress

Keap1-Nrf2 full-length complete structure has not been determined; however, Keap1 Kelch domain and Nrf2 peptides, the ETGE and DLG motifs, complexes have been reported showing the crucial structural basis of the Nh2 domain and Kelch domain interaction (Crisman et al., 2023, Lo et al., 2006). The Kelch domain

of Keap1 β-propeller blades consists of four-stranded antiparallel β-sheets, which are linked loops in different lengths.The β-strands A-B and C-D are connected by loops of shorter length, which are exposed to the bottom face of the β-propeller. In contrast, the longer loops connecting β-strands D-A and B-C extend towards the top face, sculpting a shallow substrate-recognition pocket. This is the site where Nrf2 ETGE and DLG motifs anchor (Crisman et al., 2023).

The Kelch domain of Keap1 has been co-crystallized with the Nrf2-ETGE motif including peptide. It forms β -hairpin conformation in the binding site that contains residues Asp77, Glu78, Glu79, Thr80, Gly81, and Glu82. The Glu79 establishes essential hydrogen bonds with the Kelch domain residues of Arg415, Arg483, Ser508, and Ser555. Additional hydrogen bonds between Tyr80 and Ser602; Glu82 and Ser363, Arg380 and Asp382 formed, which, combined with van der Waals contacts and water-bridged interactions, further contribute to the high affinity of KEAP1 for the ETGE motif (Zhong et al., 2020, Crisman et al., 2023) (Figure 1.6-a) .

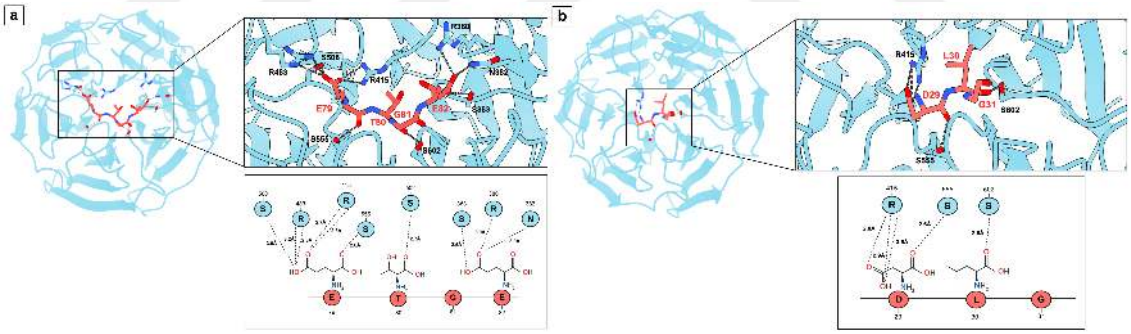

Figure 1.6: Interactions of the ETGE and DLG motifs of Nrf2 with the Kelch domain of Keap1 (a) Interaction between the ETGE peptide and the Kelch domain visualized using PDB ID: 5WFV, with interacting residues highlighted. (b) Interaction between the DLG peptide and the Kelch domain, visualized using PDB ID: 3WN7, with interacting residues highlighted.

Mouse Ke1ch domain and Nrf2-DLG complex showed that the DLG-containing peptide forms a U-shape with two short helixes (Ile28-Leu30, Arg34-Phe37) and a more extended N-terminal helix (Leu19-Arg25) instead of the -hairpin conformation of the ETGE motif. The overall conformation in ETGE-DLG and Kelch

domain complex are similar because some of the residues involved in the recognition of ETGE and DLG motifs partially overlap, even though the conformational differences between ETGE and DLG motifs. The Asp29 interacts with Arg415 and Ser555 similarly to the ETGE peptide. However, the Asp29-sidechain does not deeply include the binding pocket; therefore, the extensive hydrogen bond interaction is not observed like in the ETGE peptide. Rather, the carboxyl group of Asp29 establishes a strong salt bridge with Arg415 of Kelch domain. Particularly, Kelch domain residues ;Arg380, Arg415, Arg483, Ser555, and Ser602 interacts with Leu23, Arg25, Gln26, Asp29, and Leu30 carbonyl groups acting like hydrogen bond acceptors (Figure 1.6-b) (Fukutomi et al., 2014, Crisman et al., 2023).

The variations in the binding mechanisms of the ETGE and DLG motifs with the Kelch domain are explained with the Hinge-Latch model. While the ETGE motif is a "hinge" that remains attached to the Kelch domain, the DLGex motif acts as a "latch." In the first step of the Hinge-Latch model, the DLG motif dissociated from the Kelch domain, allowing Nrf2 to escape from Kelch domain inhibition, leading to its activation. In contrast to the DLG motif, the ETGE motif, which acts as a stable hinge, continues to bind to the Kelch domain. This dual binding mechanism is essential since it enables Nrf2 to remain partially linked to the Kelch domain of Keap1 which is necessary for the further steps of the activation process (Horie et al., 2012).

1.2.4 Interaction Between Keap1 and Dimethyl Fumarate

Dimethyl fumarate (DMF, Tecfidera) is an FDA-approved drug widely used to treat multiple sclerosis (MS) by regulating the Keap1-Nrf2 pathway to reduce oxidative stress (Scannevin et al., 2012). The production of antioxidant proteins that protect neurons from damage caused by reactive oxygen species (ROS) is increased by the activation of Nrf2 via DMF. This results in decreased immune-mediated damage to the central nervous system, which is critical to MS progress. Clinical trials demonstrate that DMF considerably mitigates disease progress and reduces the frequency of relapses and the number of new lesions in the brains of MS patients (Manai and

Amadio, 2022).

Structural studies show that DMF and monoethyl fumarate (MEF) target the BTB domain of Keap1, interacting with the highly reactive cysteine residue Cys151 (Figure 1.7-a). While DMF induces the thermal stability of the BTB domain, MEF reduces it (Qu et al., 2022). Even though the binding of fumarate derivatives does not considerably change the BTB domain structure, biochemical results of cysteine modifications are important since cysteines act as a sensor for Keap1.

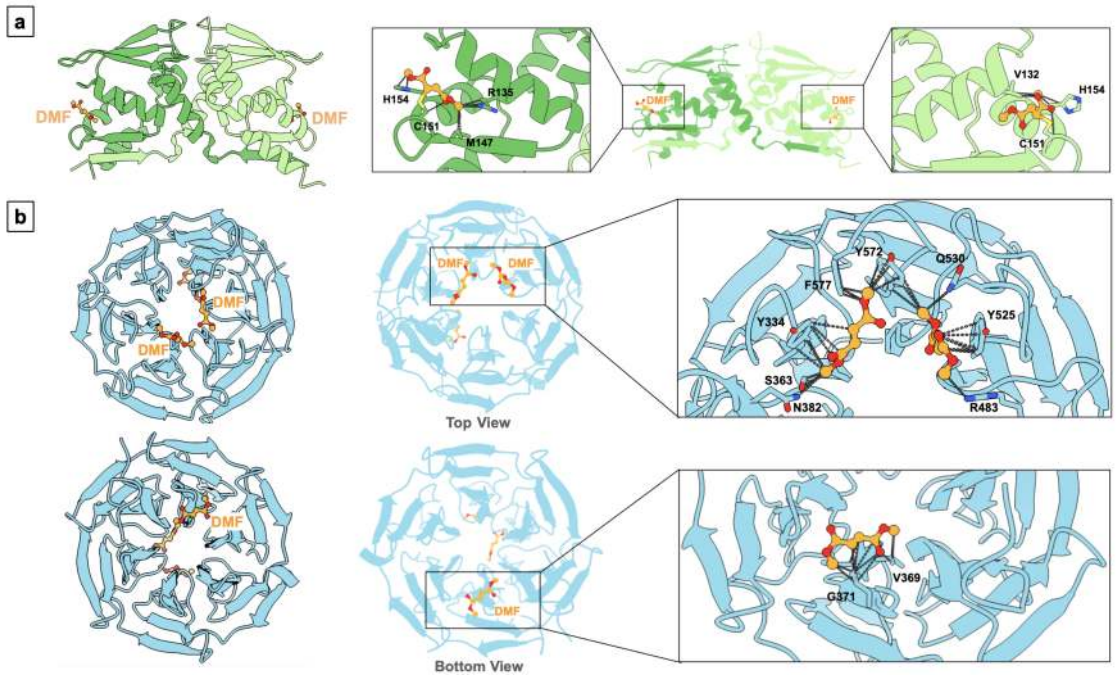

Figure 1.7: Representation of the BTB and Kelch Domain interaction with DMF (a) Interaction between the DMF molecule and the BTB domain visualized using PDB ID: 7X4W. (b) Interaction between the DMF molecule and the Kelch domain, visualized using PDB ID: 6LRZ.

DMF has also been shown to bind to the Kelch domain of Keap1, interacting with key residues crucial for the function and interaction of Nrf2. The DMF molecules bind to the three regions of the Kelch domain. The residues such as Try334 and Ser363 contribute to hydrophobic interactions, which are critical for anchoring DMF in the binding pocket. Ser602, Arg415 and Arg483 form water-mediated hydrogen bonds with DMF which enhance the specificity of the interaction. Glu530 and Ser555

form direct hydrogen bonds (Figure 1.7-b), allowing DMF to effectively displace the Nrf2 peptide in the competitive binding assays (Unni et al., 2021). Overall, these interactions are important for activating the Nrf2 pathway, confirming their significance in the therapeutic potential of DMF against neurodegenerative diseases.

1.3 X-Ray Crystallography

Many methods are used for protein structure determination, including X-ray crystallography, NMR spectroscopy, and electron microscopy. X-ray crystallography is the most commonly used technique for determining proteins and biological macromolecule structures at the atomic level. The determination of the protein structure by X-ray crystallography involves the crystal formation. The protein crystals exposed to X-rays result in diffraction patterns to generate electron density maps (Smyth and Martin, 2000).

An X-ray diffractometer is a fundamental crystal structure determination instrument that contains an X-ray source, sample holder, and detector to measure and analyze X-ray scattering from crystals. XRD is used to obtain high-resolution X-ray patterns. The atomic structure of crystals forms a characteristic diffraction pattern by scattering X-rays at specific angles. These diffraction patterns are converted into electron density maps and provide the detailed atomic structure of the crystal (Bunaciu et al., 2015).

Serial crystallography is a technique for acquiring structural data on the atomic level of a protein without large protein crystals. Instead, small diffraction patterns are collected on numerous tiny protein crystals.Serial crystallography is an eligible method for analyzing the protein structures at room temperature (Stellato et al., 2014).

Crystal quality is important for getting high-resolution protein structures with X-ray crystallography and depends on various parameters. These parameters include protein purity, concentration, temperature, and pH. The purity of the protein solution affects crystal ordering and size (McPherson and Gavira, 2014). Since the purity of the protein solution affects crystal order and size, proteins are typically

purified multiple times with techniques such as affinity, ion exchange, and size exclusion chromatography to achieve the highest possible purity (Giegé et al., 1995).

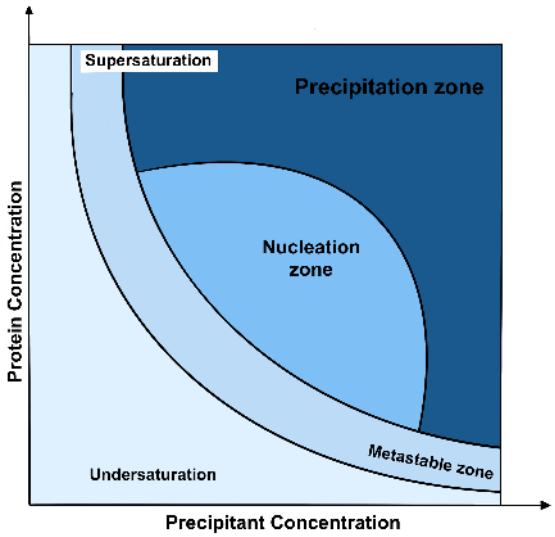

Figure 1.8: Protein crystallization Phase Diagram

Another critical parameter for getting high-quality crystals is the stability of the protein. Stable proteins in solution can quickly form crystals in their low-energy state. In contrast, an unstable protein will likely aggregate and precipitate in solution (McPherson and Gavira, 2014).Protein concentration is also a critical parameter that impacts crystal formation and diffraction quality. It is not possible to reach the supersaturated state of crystals at low concentrations, but too high concentrations also cause protein precipitation. Therefore, proteins should be driven to their supersaturated state to form crystals (Figure 1.8). A variety of crystallization screen solutions and techniques are used during crystallization to reach the supersaturated state. Crystallization screening solutions provide a wide range of environmental conditions, such as precipitant, pH, and salts, to increase the likelihood of crystal formation (Weber, 1991).

Several protein crystallization methods exist to obtain crystals from saturated protein solutions, such as vapor-diffusion crystallization, microbatch crystallization, and dialysis crystallization. Vapor-diffusion crystallization is the most common method involving a droplet of a saturated protein solution, the precipitant, and a buffer. In this method, crystals are formed as a result of a concentration gra-

dient between a mixture of protein and solvent and a reservoir solution. As the solvent evaporates from the protein solution, the protein concentration in the mixture increases and the protein becomes supersaturated. There are two main types of vapor-diffusion crystallization methods: a sitting drop and a hanging drop. The difference between the sitting drop and hanging drop methods is that in the sitting drop, the protein solution sits on a pedestal, whereas in the hanging drop, it hangs on a cover slip above the reservoir (McPherson, 2017).

The sitting drop vapor diffusion microbatch under the oil crystallization method is used in this study to screen all existing crystallization solutions (Ertem et al., 2022). A small drop of protein solution is mixed with the same amount of reservoir solution in one well of a Terasaki plate. The protein-reservoir solution mixture is then covered with paraffin oil to slow evaporation.

1.4 Molecular Docking

Molecular docking is a structure-based *in silico* method for predicting ligand-protein interactions. This technique requires a 3D representation of the target protein obtained from structure determination techniques such as X-ray crystallography, nuclear magnetic resonance, and cryo-electron microscopy (Agu et al., 2023). Autodock and AutodockTool (ADT) are widely used to facilitate ligand and receptor preparation steps by providing grid maps on the target protein. The docking algorithms search high-dimensional spaces to find optimal binding conformations and employ scoring functions to rank them based on their predicted binding affinities (Morris, 2008). Molecular docking is necessary for the virtual screening of large compound libraries and dramatically accelerates the drug discovery process. This allows researchers to prioritize which compounds to synthesize and test experimentally by predicting drug-target protein binding potential.

Chapter 2

MATERIALS AND METHODS

2.1 Materials and Equipment

2.1.1 Stock solutions and LB medium

Stock solutions of antibiotics, 0.4 M IPTG (Table 2.1) and LB medium (Table 2.2) were prepared for use in bacterial transformation and protein expression experiments.

Table 2.1: Preparation of Antibiotic and 0.4M IPTG Stock Solutions

Reagent	Amount (mg/ml)
Kanamycin	35
Chloramphenicol	35
IPTG	95.324

Table 2.2: Preparation of LB medium

Reagent	Amount (g/L)
Tryptone	10
NaCl	10
Yeast Extract	5
Total	25

2.1.2 Buffers

Lysis buffer (Table 2.3), HisA buffer (Table 2.4) , HisB buffer (Table 2.5), and SEC buffer (Table 2.6) were prepared for small-scale expression check ,protein purification, and size exclusion chromatography, all at pH 8.5.

Table 2.3: Preparation of Lysis Buffer

Reagent	Final concentration	Amount
NaCl	500 mM	29.22 g
Tris	50 mM	6.06 g
Imidazole	20 mM	1.362 g
Glycerol	5% (v/v)	50 ml
Triton X-100	0.1% (v/v)	5 ml
ddH_2O		up to 1 L
Total		1 L

Table 2.4: Preparation of HisA Wash Buffer

Reagent	Final concentration	Amount
NaCl	200 mM	11.688 g
Tris	20 mM	6.06 g
Imidazole	20 mM	1.362 g
ddH_2O		up to 1 L
Total		1 L

Table 2.5: Preparation of HisB Elution Buffer

Reagent	Final concentration	Amount
NaCl	200 mM	11.688 g
Tris	20 mM	6.06 g
Imidazole	250 mM	17.019 g
ddH_2O		up to 1 L
Total		1 L

Table 2.6: Preparation of SEC Buffer

Reagent	Final concentration	Amount
NaCl	200 mM	11.688 g
Tris	20 mM	6.06 g
ddH_2O		up to 1 L
Total		1 L

2.2 Bacterial Transformation

The Kelch domain of human Keap1 genes was cloned into an empty pET-28a(+) bacterial expression vector containing an N-terminal 6x-Histidine tag with thrombin protease cut site for affinity chromatography (Figure 2.1). The amino acid sequences of the Kelch domain constructs are given below.

The Kelch Domain of Human Keap1:

MGSSHHHHHHSSGLVPRGSAPKVGRLIYTAGGYFRQSLSYLEAYNPSDGTW
LRLADLQVPRSGLAGCVVGGLLYAVGGRNNSPDGNTDSSALDCYNPMTNQ
WSPCAPMSVPRNRIGVGVIDGHIYAVGGSHGCIHHNSVERYEPERDEWHLV
APMLTRRIGVGVAVLNRLLYAVGGFDGTNRLNSAECYYPERNEWRMITAM
NTIRSGAGVCVLHNCIYAAGGYDGQDQLNSVERYDVETETWTFVAPMKHR
RSALGITVHQGRIYVLGGYDGHTFLDSVECYDPDTDTWSEVTRMTSGRSG
VGVAVT

The cloned pET-28a(+) vector was transformed into a competent *Escherichia coli* BL21 Rosetta2^{TM}strain. The plasmid vector and competent cells were thawed on ice for 20 minutes. Then, 50 μl of the competent cells were combined with 2 μl of plasmid in a sterile Eppendorf tube, mixing gently by pipetting. The mixture was incubated on ice for an additional 20 minutes. Afterward, using a heat block, the bacterial cells underwent heat shock at 42°C for 45 seconds. Following the heat shock, the tube was placed back on ice for 2 minutes. Next, 500 μl of LB medium at room temperature was added to the mixture, and the cells were incubated at 37°C for 90 minutes. The tube was inverted several times every 15–20 minutes during

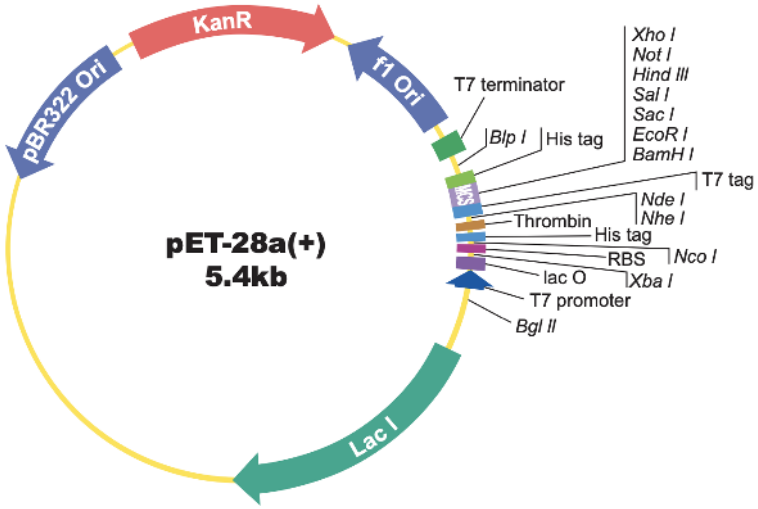

Figure 2.1: The plasmid map of the bacterial expression vector pET-28a(+). KanR stands for the kanamycin resistance gene. MCS is multiple cloning site (GenScript, USA).

this incubation.

The tube was centrifuged at 1500 rpm for 2 minutes, followed by incubation at 37°C, and the supernatant was removed.The remaining supernatant was resuspended with the pellet by pipetting and 50 μl of the cells were inoculated onto the LB agar plate (containing Kanamycin and Chloramphenicol) under a sterilized fume hood. The inoculation process was performed using spreader glass beads; the cells were spread on the LB agar plate and rocked back and forth. After spreading the cells on the plate, the beads were removed and the plate was incubated overnight at 37°C.

Since the pET-28a (+) plasmid contains the kanamycin antibiotic resistance gene, and the Rosetta2^{TM} *E. coli* strain contains chloramphenicol resistance, a double antibiotic selection was performed. For double antibiotic selection, 7 μl kanamycin and 7 μl chloramphenicol were added to the 7 ml LB medium (refer to Table 2.2). After overnight incubation, the colonies grown on the agar plate were inoculated into the LB medium with a micropipette and the cell culture was incubated at 37°C at 110 rpm for 5-6 hours. For glycerol stock preparation, after incubation,

750 μl sterilized 80% glycerol and 750 μl cell culture were mixed in a cryotube and stored at -80°C.

2.3 Mini Scale Protein Expression Check

For each temperature assay (18°C, 30°C, and 37°C) of the Keap1 Kelch domain, 10 ml of LB (30 g/L) medium was prepared and autoclaved at 121°C for 20 minutes. 6 μl kanamycin (35 mg/ml stock) and 6 μl chloramphenicol (35 mg/ml stock) antibiotics were mixed with 6 ml LB medium. Glycerol stocks containing the plasmid with the Kelch domain were inoculated into the prepared LB media.

Cell cultures were incubated overnight at 37°C with shaking at 110 rpm. Following overnight incubation, 1 ml before-induction sample was taken and centrifuged from each flask. The supernatant was discarded, and the 'before induction' samples were stored at -45°C until the SDS page electrophoresis.

The experiment was continued with the remaining 5 ml of cell culture in each flask. In order to induce protein expression, 5 μl of 0.4 M IPTG (isopropyl-β-D-1-thiogalactopyranoside) was added to each of the remaining cultures. The final IPTG concentration was 0.4 mM. The cultures were again incubated at 18°C, 30°C and 37°C for 8 hours at 110 rpm. After IPTG induction, the cells were collected in Eppendorf tubes by centrifugation at 3500 rpm.

The cell pellets, which were inducted at different temperatures, were resuspended in 1 ml lysis buffer (refer to Table 2.3) and then sonicated until they were completely dissolved. The cell lysates were then centrifuged, and the supernatants were transferred to another Eppendorf tube for a pull-down experiment. The supernatants were incubated with 40 μl Ni-NTA beads at 4°C for 4-5 hours. The samples were centrifuged at 2400 rpm for 2.5 minutes and the supernatants were gently discarded without touching the beads.

The samples were washed with 1 ml HisA buffer (refer to Table 2.4) and incubated for 20 minutes on ice. After incubation,the samples were centrifuged at 2400 rpm for 2.5 minutes, and the supernatants were removed carefully. For eluting the protein from the Ni-NTA beads, 200 μl of HisB elution buffer was then added to each tube

and incubated for 5 minutes on ice.

40 μl of samples were taken for each temperature at the before-induction, pellet, after-induction, and pull-down, and the samples were mixed with 10 μl of loading dye. SDS-PAGE electrophoresis was then performed to assess protein expression in these samples

2.4 Large Scale Protein Production and Purification

180 ml (for starter culture) and 6 x 1000 ml LB media were prepared and sterilized with an autoclave. To the sterilized 180 ml LB medium, 180 μl kanamycin stock solution and 180 μl chloramphenicol stock solution (1:1000 ratio) were added. The glycerol stock was inoculated into the 180 ml of LB medium and incubated overnight (18 to 24 hours) at 37 °C and 110 rpm.

Following overnight incubation, 1000 μl of kanamycin and 1000 μl of chloramphenicol were added to six flasks containing 1000 μl of LB medium to begin large-scale production. The starter culture was divided by these 6 flasks and incubated again at 37°C at 110 rpm. The OD600 value of the cultures was measured every hour with a NanoDrop instrument. When the OD600 reached approximately 0.8-1.2, the cell cultures were induced with 0.4 M IPTG, 1000 μl of IPTG was added to each flask. The cultures were incubated at 37°C at 110 rpm for 12 hours.

At the end of the incubation period, the cells were collected by centrifugation at 3500 rpm for 45 minutes and the supernatants were discarded. The majority of the cells were harvested from the centrifuge bottles with a spoon and transferred into falcon tubes. The remaining cells were harvested by dissolving in 20 ml LB, and Falcon tubes were centrifuged at 3500 rpm for 30 minutes, and the supernatants were discarded. Cell pellets were kept at -45°C until purification.

Cell pellets were placed on ice, and 45 ml of lysis buffer was added (refer to Table 2.3). The cells were sonicated for 1 second on and 1 second off for 30 seconds at 60% amplitude. This procedure was repeated until the pellets were dissolved entirely. Falcon tubes were kept on ice during sonication to prevent overheating of the samples.

Following sonication, the cell lysates were placed into ultracentrifuge tubes and balanced with the help of a weighing device. Cell lysates were ultracentrifuged with a Ti45 rotor in a Beckman Optima ^{TM} L-80XP Ultracentrifuge (Beckman, USA) at 35000 rpm for 1 hour 15 minutes at 4°C. When the ultracentrifugation was finished, the supernatants were collected and filtered using a filter paper to obtain a clean solution.

For protein purification, an AKTA-Go FPLC device was used. The supernatant was loaded to a Ni-NTA affinity column attached to AKTA go. In the first step, the Ni-NTA column was equilibrated with HisA wash buffer (refer to Table 2.4) until the UV absorbance stabilized. After the column was equilibrated, the entire supernatant was loaded into the column, and the flowthrough, which contains unbound proteins, was collected from the outlet valve. Afterward, the column was washed with His A buffer to eliminate the unbound and nonspecifically bound proteins, and the wash continued until the UV absorbance returned to a stable level. After the UV absorbance was stable, HisB elution buffer (refer to Table 2.5) was applied to the column to elute the Kelch domain of Keap1. The eluted protein was concentrated with a 10 kDa Amicon Ultra centrifugal filter tube.

2.5 Size Exclusion Chramotography

Size exclusion chromatography analysis was performed on an AKTA-Go FPLC device. A Superdex^{TM} 75 (10/300 GL) column was attached to the AKTA-Go and equilibrated with SEC buffer (refer to Table 2.6) until the UV absorbance was stable. A total of 500 μl of concentrated protein solution was manually injected using a 1 ml loop in the manual load mode of the AKTA-go system. Afterward, the system was switched to inject mode. After a 10-minute waiting period, the system was changed back to manual load mode. Using the fraction collector of the instrument, all of the fractions were collected in a 1 ml Eppendorf tube. SDS-PAGE gel electrophoresis was performed on the samples obtained from the fractions.

The fractions corresponding to the monomeric and dimer peaks were concentrated. Subsequently, 500 μl of each fraction (monomer and dimer) were subjected

to a second round of SEC using the same procedure.

2.6 Protein Crystallization

The crystallization method, sitting drop microbatch under oil, was performed for the Kelch domain of human Keap1 using commercially 3000 available sparse matrix crystallization screening sets. 0.83 μl protein solution and 0.83 μl crystallization solution were mixed in the wells of 72-well Terasaki plates. Each well was covered with approximately 16.6 μl of paraffin oil. Approximately 3500 crystallization conditions were screened two times and plates stored at +4°C.

The Keap1 protein crystals were acquired in crystallization solution Salt Rx-II #44 which include 4.0M ammonium acetate, 100mM sodium acetate trihydrate (pH 4.6) and Salt Rx-II #45 4.0M ammonium acetate, 100mM BIS TRIS propane (pH 7.0) and Salt Rx- II #46 which include 4.0 M ammonium acetate, 100mM Tris-HCl (pH 8.5) (Hampton Research, USA) within one month. In order to get a larger number of crystals, the crystallization experiments were repeated with same crystallization conditions in all the wells of a Terasaki plate.

2.7 Protein Crystal Optimization

The obtained crystals were initially optimized in a Terasaki plate with optimization screens. First, 0.83 μl of protein solution and 0.83 μl of Salt Rx-II #44 crystalization condition were mixed and screened with 0.50 μl of CryoPro screens, detergent screens, and additive screens, respectively, then covered with 16.6 μl of paraffin oil. Second, 0.83 μl of protein solution and 0.83 μl of Salt Rx-II #45 crystalization condition were mixed and screened with 0.50 μl of CryoPro screens, detergent screens, and additive screens, respectively, then covered with 16.6 μl of paraffin oil. Third, 0.83 ul of protein solution and 0.83 μl of Salt Rx-II #46 crystalization condition were mixed and screened with 0.50 μl of CryoPro screens, detergent screens, and additive screens, respectively, then covered with 16.6 μl of paraffin oil.

The best tiny crystals were observed with CryoPro #44, containing 5.0 M lithium acetate dihydrate, and Salt Rx-II #44, consisting of 4.0 M ammonium acetate and

100 mM sodium acetate trihydrate (pH 4.6).

Additionally, tiny protein crystals were grown in Eppendorf tubes. Protein, crystallization conditions, and optimization solutions were mixed in different drop ratios. The optimal crystal growth was obtained by mixing 20 μl of Salt Rx-II #44, 10 μl of CryoPro #44, and 20 μl of protein solution, followed by incubation at 4°C.

2.8 X-ray Data Collection and Processing

X-ray crystallographic data was collected using Rigaku XtaLAB Synergy Flow X-ray Diffractometer "Turkish Light Source" (University of Health Sciences, Istanbul, Türkiye). The Oxford Cryosystems Cryostream 800 Plus was adjusted to 300K for ambient temperature data collection. The Terasaki plate was mounted on the *XtalCheck-S* plate reader (Gul et al., 2023), and multiple crystals were screened in the Terasaki plate. After screening all the crystals, good diffracting crystals were selected. Diffraction data were collected from two crystals. The diffraction data were collected continuously for a period of approximately 3 hours at 90% attenuation. The detector distance was set to 120.0 mm. The scan width was set to 0.5 degrees and the exposure time was 5.0 seconds per frame.

The collected data were merged using the profit merge process with CrysAlisPro 1.171.42.59a software (Rigaku OxfordDiffraction, 2022) to generate an integrated reflection dataset (***.mtz**) file for further analysis. Each frame of the ambient temperature data was analyzed using peak hunting and masking, and only those with appropriate unit cells were selected for inclusion. In order to create a batch script, the command "**xx proffitbatch**" was entered in the command shell, and the batch script was run in the command shell. Each data reduction created an rrprof file that contained all the data that was unmerged and unscaled. The proffit merge tool integrated into CrysAlis Pro was used to merge all reduced data and export an mtz file.

2.9 Serial Synchrotron Crystallography Data Collection

Serial X-ray diffraction experiments were performed at room temperature at the European Molecular Biology Laboratory (EMBL) beamline P14-2 at the PETRA III synchrotron at DESY, Hamburg (TREXX). These experiments utilized a fixed target system with HARE chips (6x6 compartments).

First, the microbatches were centrifuged to increase the density of the crystals. Following centrifugation, the HARE chips were cleaned using Tergazyme or a 1 M HCl solution for 1 hour. After treatment, the chips were washed with deionized water. The chips were then glow discharged for the 60s at 35 mA.

200 μl slurry of crystals were loaded onto the chips using a pipette. The slurry was vacuumed using a laboratory vacuum pump, ensuring only the crystals remained in the chip wells. Once the chips were prepared, they were mounted into the system for data collection. The X-ray beam size was set to 15 $\times$ 10 m, and diffraction data were recorded using a Dectris Eiger 4M detector. Initial indexing of the X-ray diffraction images was performed using the XGANDALF algorithm and 696 cystals merged using CrystFEL's Monte Carlo method (Gevorkov et al., 2019, White, 2019).

2.10 Structure Determination and Refinement

The molecular replacement for both XRD and synchrotron structure was performed with the PHASER-MR program implemented in the PHENIX suite (Adams et al., 2010). For XRD data, the previously determined crystal crystal structure (PDB ID: 6ROG) was used as the initial search model. Our determined structure (PDB ID: 8X34) was used as a model for synchrotron data. In PHENIX, the rigid body and simulated annealing refinement, individual coordinates, and translation/libration/screw (TLS) parameters were refined. Potential positions of altered side chains and water molecules were examined, and the models were manually constructed/reconstructed using the COOT program (Emsley and Cowtan, 2004). The final structure refinement was performed using phenix.refine in PHENIX. The figures were generated using *PyMOL*, ChimeraX and BioRender.

2.11 Molecular Docking Analysis

The ZINC database (https://zinc.docking.org/) was used to find the appropriate compounds for the exploration of potential Keap1/Nrf2 inhibitors. The crystal structure of the Kelch domain of the Keap1 (PDB ID: 8X34) was obtained from the RCSB database. To prepare the raw protein for molecular docking, the Protein Preparation Wizard (Schrödinger) was used. Using Prime, the missing chains were automatically added, and PropKa calculated the protonation state at physiological pH. Sitemap was used to determine the top-ranked potential receptor binding sites of the protein.

The docking grid was established by grid generation, picking the identified binding site around the residues such as Asn382, Arg483, Gln530, Gly371, Phe577, Ser363, Tyr334, Tyr525, Tyr572, and Val369. In grid generation the size of ligand diameter midpoint box was determined as 10 Å x 10 Å x 10 Å for x,y and z dimensions. The generated grid was then utilized for docking experiments. In addition, compound sketches and cleanups were performed in the Maestro workspace. Using the LigPrep module, 300 compounds were prepared as ligands through energy minimization with the OPLS 2005 force field at a physiological pH of 7.0 ± 2.0. Without further modification, the best-minimized structures were used for docking experiments. The flexible ligand alignment tool was employed to align structurally similar ligands. A self-docking experiment was conducted with DMF to validate the docking protocol, and the optimum structure (lowest energy) was determined. After the ligand was processed with the Glide/SP docking protocol, the same docking procedures were applied to all designed compounds. (Schrödinger Release 2016-2: Schrödinger, LLC: New York, USA, Ciftci et al., 2022).

Chapter 3

RESULTS

3.1 Expression of Human Kelch Domain of Keap1

Bacterial transformation of the Kelch domain of Keap1 was successfully achieved. The transformation was confirmed by the growth of multiple colonies on LB agar plates with the antibiotics Kanamycin and Chloramphenicol, indicating that the *E. coli* BL21- Rosetta2^{TM} cells had been transformed.

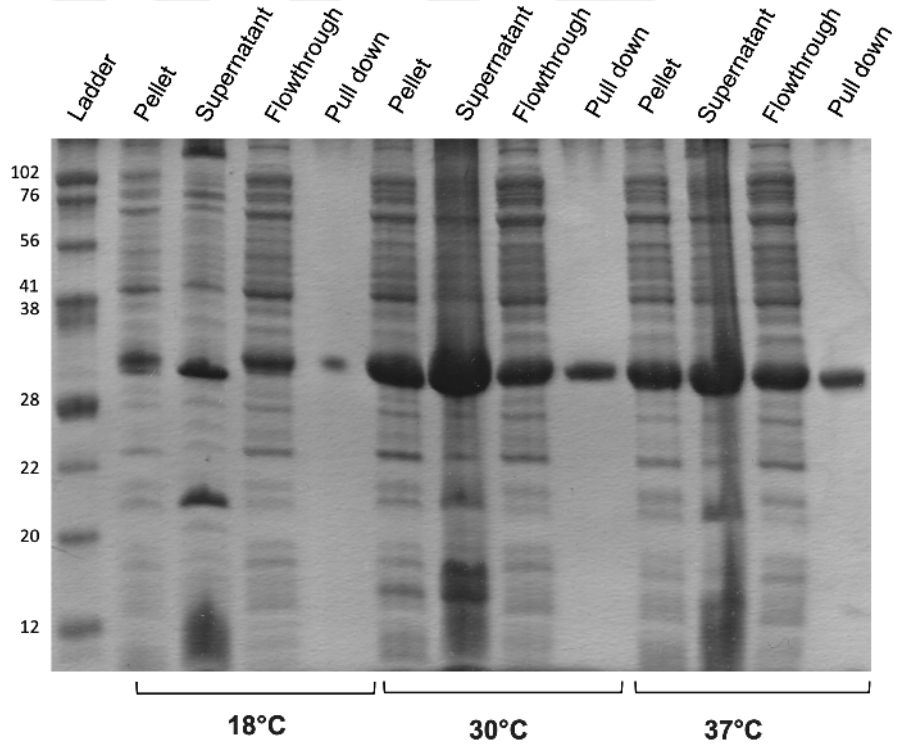

Figure 3.1: SDS -PAGE analysis of mini-scale expression of the Keap1 Kelch domain at different temperatures (18°C, 30°C, 37°C). The lanes represent samples from specific experiment steps, with the protein ladder providing molecular weight markers for size comparison.

To test that the Kelch domain of the Keap1 construct was expressed in the

transformed bacteria and to determine the optimal temperature for induction, a small-scale protein expression, and pulldown assay were carried out prior to large-scale protein production. The Kelch domain of Keap1 is 33.6 kDa. The results of the mini-scale protein expression showed that the construct was effectively overexpressed in *E. coli* Rosetta-2^{TM} cells with IPTG induction at 18°C, 30°C and 37°C (Figure 3.1). The "pull down" samples displayed the protein band around 33 kDa, indicating that the protein binds successfully to Ni-NTA beads, with the highest expression observed at IPTG induction at 37°C for 12 hours.

According to mini-scale expression check results, the Kelch domain of the Keap1 construct was expressed in 6 L of LB medium induced with IPTG at 37°C for 12 hours. We purified the Kelch domain of Keap1 with high purity and concentrated it to about 15 mg/ml (Figure 3.2).

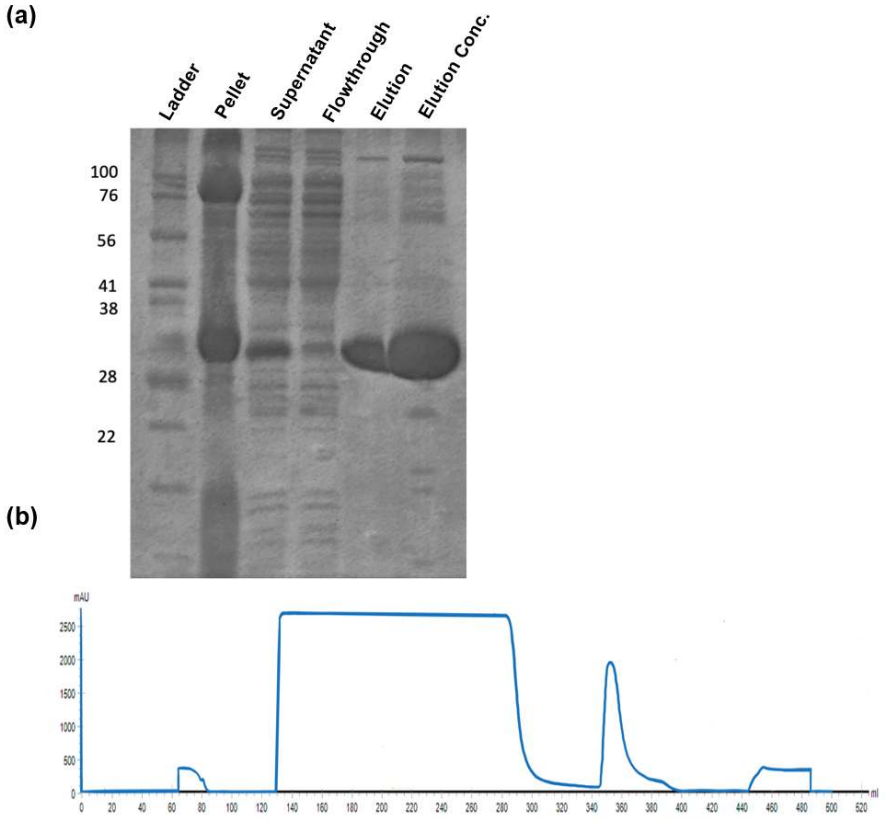

Figure 3.2: (a) SDS-PAGE analysis of large-scale protein expression of the Keap1 Kelch domain. (b) Chromatogram showing protein purification results.

3.2 Size Exclusion Chramotography Analysis of Human Kelch Domain of Keap1

Size exclusion chromatography was performed to analyze the monomeric and dimeric states of the Kelch domain. In the experiment, Superdex^{TM} 75 column was used because it is suitable for separating proteins with molecular weights ranging from 3 kDa to 70 kDa since the protein size is 33.6 kDa.

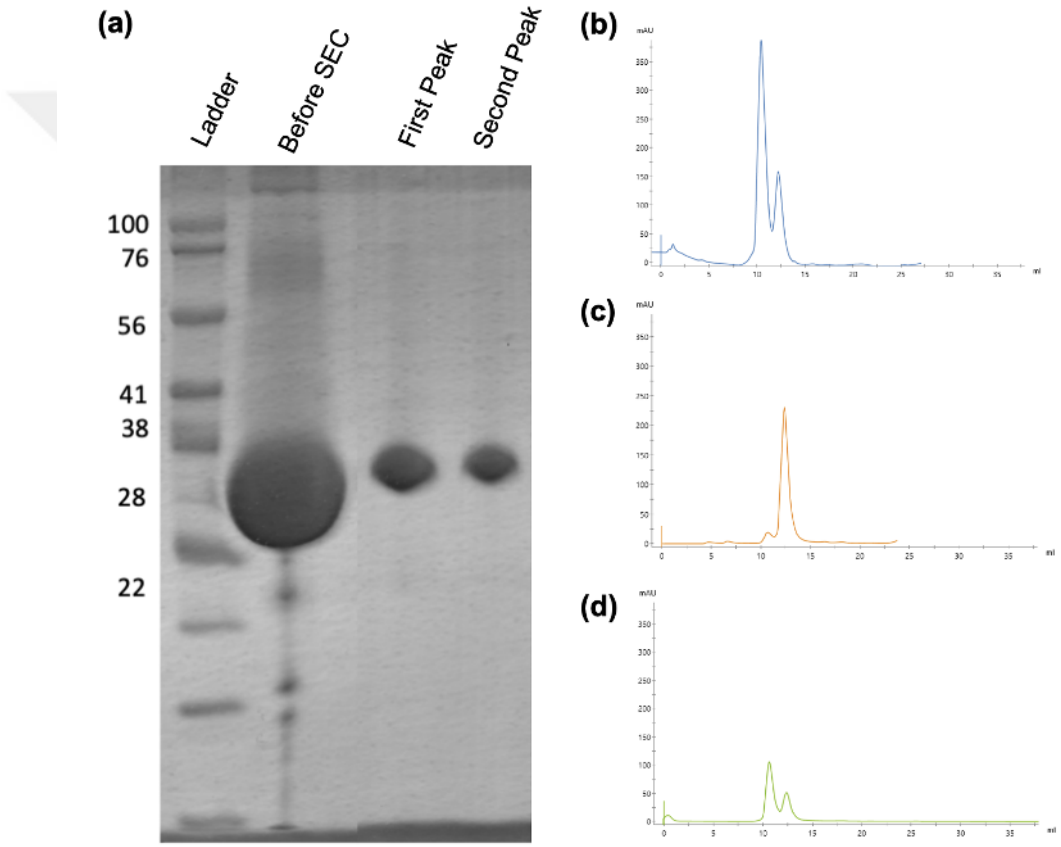

Figure 3.3: (a) SDS-PAGE results following size exclusion chromatography (b) Chromatogram of mixed protein soluiton (c)Chromatogram of dimeric protein soluiton (d)Chromatogram of dimeric protein soluiton.

Firstly, size exclusion chromatography was performed on directly eluted and concentrated protein samples. Two peaks were observed as a result: the first peak expected to correspond to the dimer and the second to the monomer(Figure 3.3-b). This is because larger proteins elute from the column before smaller proteins due to

the porous matrix of the column. Larger proteins are not able to pass through the pores as easily as smaller ones.

Afterward, the fractions collected from each peak ran individually and confirmed the dimeric and monomeric states of the protein solution (Figure 3.3-c,d). Size exclusion chromatography demonstrated that the protein solution mostly comprised a dimeric form of Keap1. In SDS-PAGE, the presence of the protein was confirmed. Although the dimeric and monomeric forms were identified, the results show similar sample sizes (Figure 3.3-a) because SDS denatures proteins and disrupts dimerization.

3.3 Crystallization and Crystal Optimization of Human Kelch Domain of Keap1

The Kelch domain of Keap1 was crystallized using the sitting-drop vapor-diffusion microbatch under the oil method. The protein was screened with approximately 3500 sparse matrix crystallization conditions twice. The crystallization was performed at room temperature, and the plates were stored at +4°C. The crystals were grown in Salt Rx-II #44, which contains 4.0M ammonium acetate, 100mM sodium acetate trihydrate (pH 4.6), and Salt Rx-II #45, which contains 4.0M ammonium acetate, 100 mM BIS TRIS propane (pH 7.0) and Salt Rx- II #46 which contain 4.0 M Ammonium Acetate, 100 mM tris-hCl (pH 8.5) (Hampton Research, USA)(Figure 3.4). Acetate ions were found to be present in all crystallization conditions where crystals were obtained. This indicates that acetate is essential for the crystallization of the Kelch domain of Keap1.

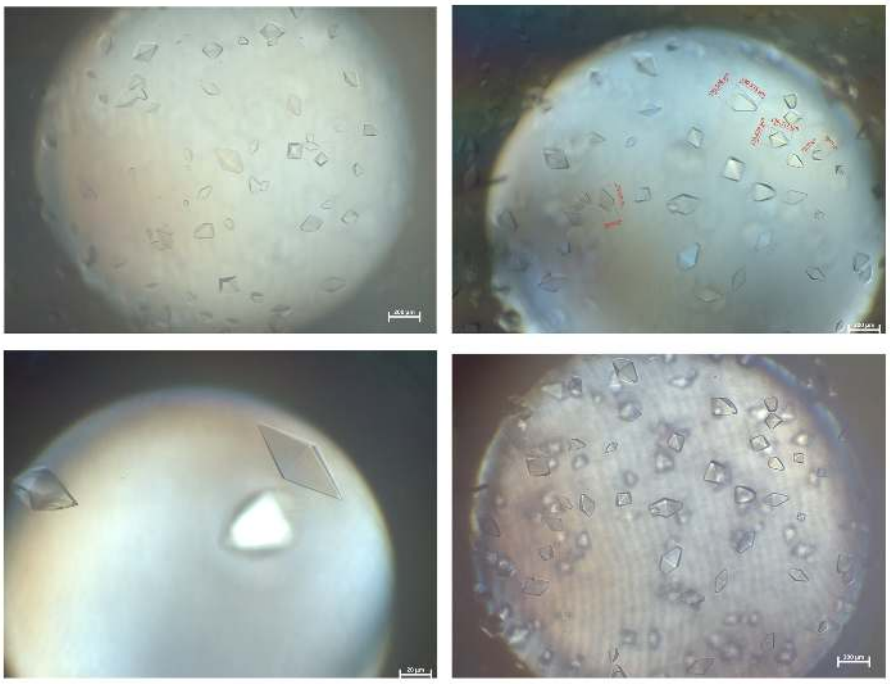

Figure 3.4: Light microscope images of Keap1 Kelch domain crystals, used for ambient data collection. The crystals display well-formed shapes suitable for X-ray diffraction.

The crystals obtained from Salt Rx-II #44, Salt Rx-II #45, and Salt-Rx #46 were optimized with optimization screens for acquiring well-diffracting and small crystals. Firstly, the crystal optimization was performed in the Terasaki plate, and the small crystals were obtained by mixing Cryo-Pro #44, which contains 5.0 M lithium acetate dihydrate, and Salt Rx-II #44, which includes 4.0M ammonium acetate, 100 mM sodium acetate trihydrate (pH 4.6) with 1:2 ratio(Figure 3.5-a,b).

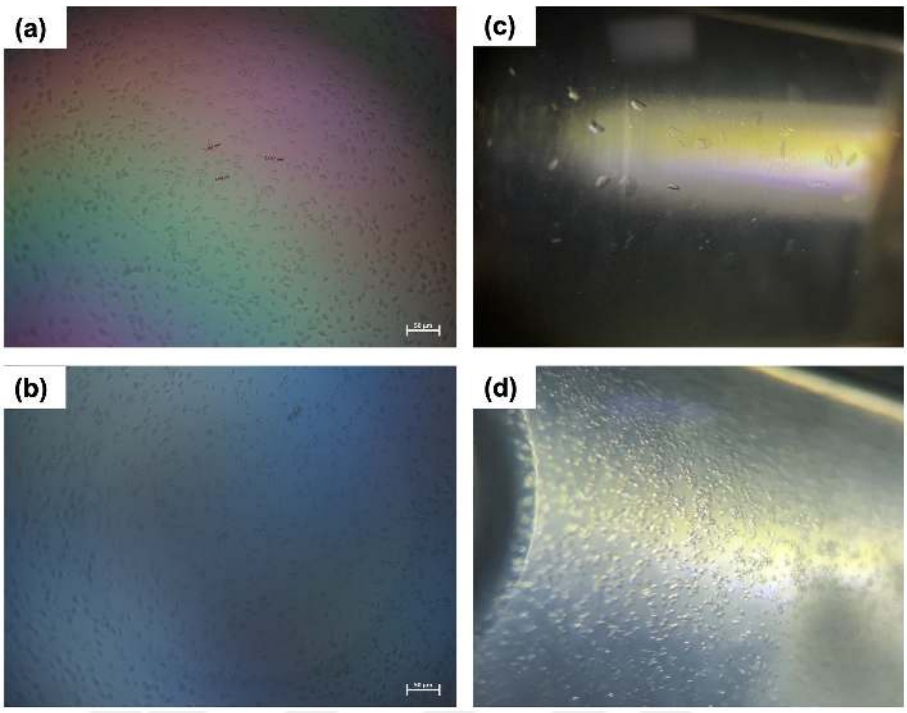

Figure 3.5: (a-b) Light microscope images of optimized small crystals grown in a Terasaki plate. (c) Larger microbatch crystals in an Eppendorf tube. (d) Tiny microbatch crystals in an Eppendorf tube were used for data collection at the TREXX beamline.

Additionally, the crystals were grown in Eppendorf tubes (microbatch). As a result of mixing 20 μl of Salt Rx-II #46 which contains 4.0 M ammonium acetate, 100 mM tris-hCl (pH 8.5) and 10 μl of protein solution, the big protein crystals were obtained in the Eppendorf tube (Figure 3.5-c). When the optimization was performed in eppendorf, mixing 20 μl of Salt Rx-II #44, 10 μl of CryoPro #44, and 20 μl of protein solution, the small crystals were obtained as well. According to the contents of optimization screening solutions, it has been shown that acetate concentration in the environment can affect protein size. The small crystals in the microbatch was used for serial synchrotron data collection (TREXX) (Figure 3.5-d).

3.4 Structural Analysis of Human Kelch domain of Keap1

3.4.1 Investigation of Dimeric Kelch Domain

The crystal structure of the apo Keap1 Kelch domain was determined at 3A resolution at ambient temperature by using conventional XRD (*Turkish DeLight*) (PDB ID: 8X34) and serial synchrotron crystallography (TREXX). The dimeric Kelch domain of Keap1 consists of 572 amino acids. The Keap1 crystal belongs to the P212121 orthorhombic space group with α= 90°, β= 90°, γ= 90°, a = 76.61 Å, b = 77.76 Å, and c = 213.29 Å. Table 3.1 demonstrates the data collection and refinement statistics.

Table 3.1: Data collection and refinement statistics. Keap1-XRD: Ambient structure obtained by conventional X-ray diffraction. Keap1-TREXX: Ambient structure obtained by serial synchrotron crystallography.

Dataset	**Keap1-XRD**	**Keap1-TREXX**
PDB ID	8X34	
Data collection		
X-ray source	Turkish DeLight	PETRA III
Temperature (K)	298	298
Space Group	P 21 21 21	P 21 21 21
a,b,c (Å)	76.61,76.76,213.29	75.600,75.800,218.650
α,β,γ (°)	90,90,90	90,90,90
Resolution	31.15-3.00	109.32-3.00
CC1/2	0.77	0.87
*CC**	0.93	0.98
$I/\sigma I$	2.07	
Completeness (%)	99.6	82.2
Refinement		
Resolution	3Å	3Å
No. reflections	25899	21358

Dataset	**Keap1-XRD**	**Keap1-TREXX**
Rwork / Rfree	0.275/0.329	0.366/0.314
No. atoms		
Protein	4387	4405
Ligand / Ion / Water		
***B*-factors**		
Protein		15.18
Ligand / Ion / Water		
Coordinate errors	0.71	0.85
Bond lengths (Å)	0.003	0.011
Bond angles (°)	1.363	1.559
Ramachandran plot		
Favored (%)	93.1	86.27
Allowed (%)	6.7	10.74
Disallowed (%)	-	-

The crystal structure of the Kelch domain shows the dimerization of the Kelch domain (Figure 3.6-a). Monomer alignment based on α carbons reveals a slight difference between chain A and chain B, with a root mean square deviation (RMSD) of 0.21 Å (Figure 3.6-c). The dimerization interface between the monomers is shown in Figure 3.6-b and the residues that mediate dimerization are presented in Figure 3.6-d.

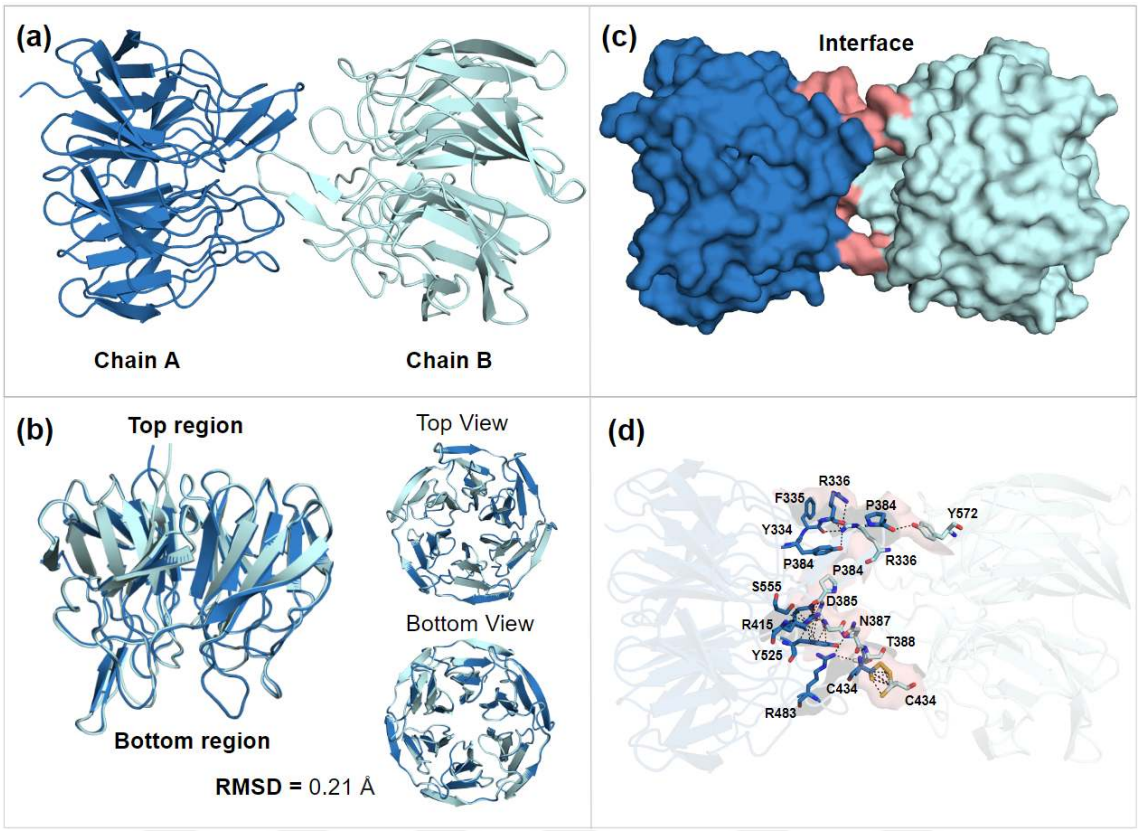

Figure 3.6: Dimerization of Kelch Domain and Alignment of monomers (a) The Ambient structure of Keap1 dimerizes through its Kelch domain, with Chain A colored in skyblue and Chain B colored in palecyan. (b) Alignment of Chain A and Chain B. The top and bottom views of the Keap1. (c) The protein interface between chains A and B is shown and colored in salmon. (d) Residues contributing to the protein interface and dimerization are highlighted and shown in the stick representation.

3.4.2 Comparison of Ambient and Cryogenic Temperature Structure of Kelch domain

Comparison of Kelch domain ambient structure at 3 Å and cryogenic structure at 2.1 Å shows a difference between the two structures with an RMSD score of 0.92 Å. To confirm that this difference is a result of temperature change, the cryogenic structure (PDB ID: 6ROG) was re-refined to the resolution range of the ambient structure (3 Å).The re-refined structure was then aligned with the PDB deposited structure. The resolution difference was eliminated from this perspective, clearly

indicating that the variation was caused by the temperature change.

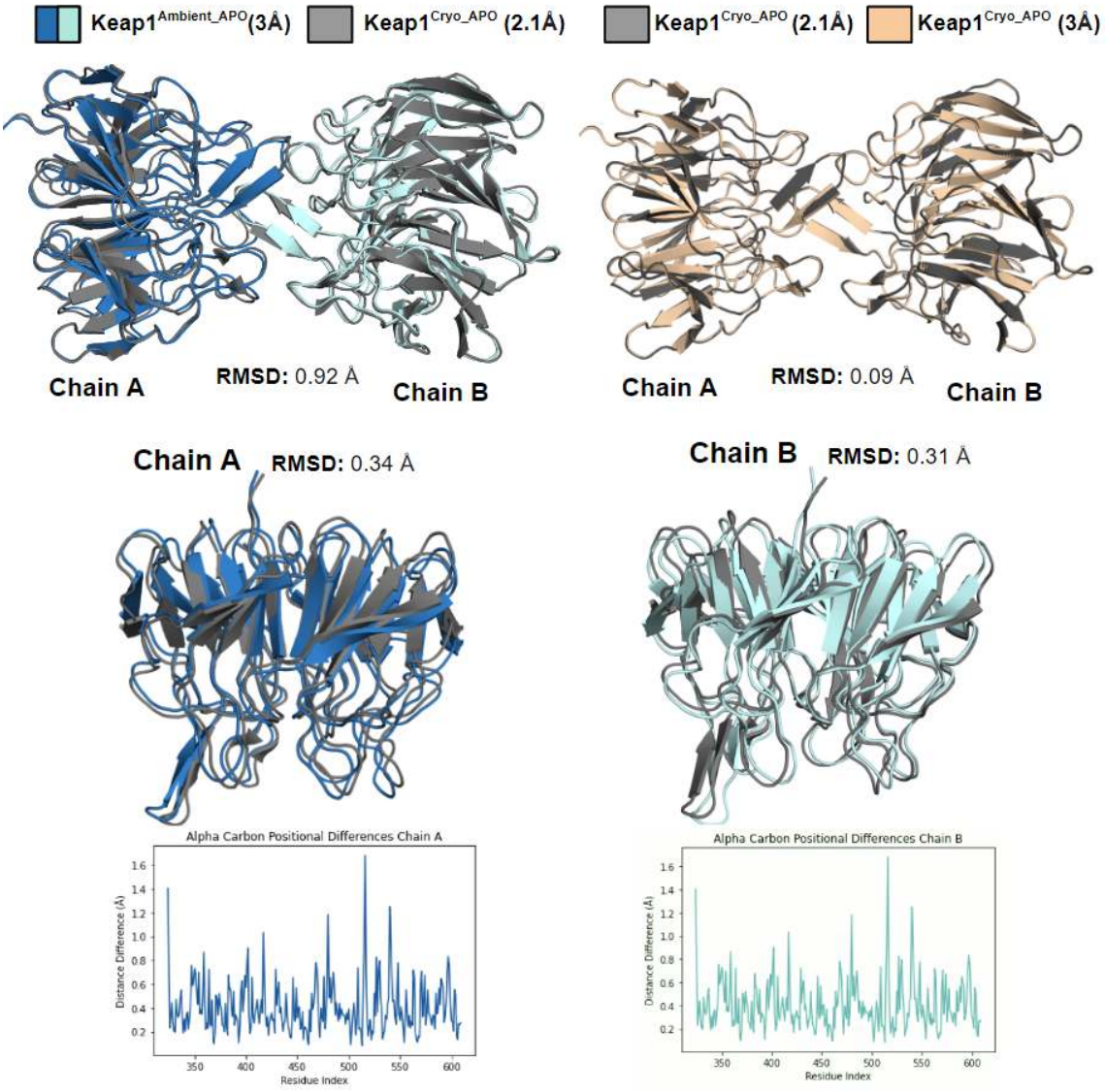

Figure 3.7: Superposition of 2.16 Å cryogenic Keap1 structure (colored in gray50) and 3 Å ambient Keap1 structure (colored in sky blue and palecyan according to monomers). Superposition of 2.16 Å cryogenic Keap1 structure and 3 Å cryogenic Keap1 structure (colored in wheat). Alignment of cryogenic and ambient Keap1 structure Chain A demonstrated by a pairwise distance plot. Alignment of cryogenic and ambient Keap1 structure Chain B shown by a pairwise distance plot.

The ambient structures obtained from serial synchrotron crystallography (TREXX) and conventional X-ray diffraction (XRD) were compared with the cryogenic struc-

ture. The flexibility of the ambient and cryogenic structures was compared by examining the distribution of the B-factor in each structure. The B-factor spectrum for the ambient structure ranged from a minimum of 31.43 to a maximum of 174.61. In the ambient structures, the thermal ellipsoids seem larger, ranging from red to yellow to green (Figure 3.8-b,c). In the cryogenic structure, they are smaller and appear to be more spherical, represented by dark blue (Figure 3.8-a).

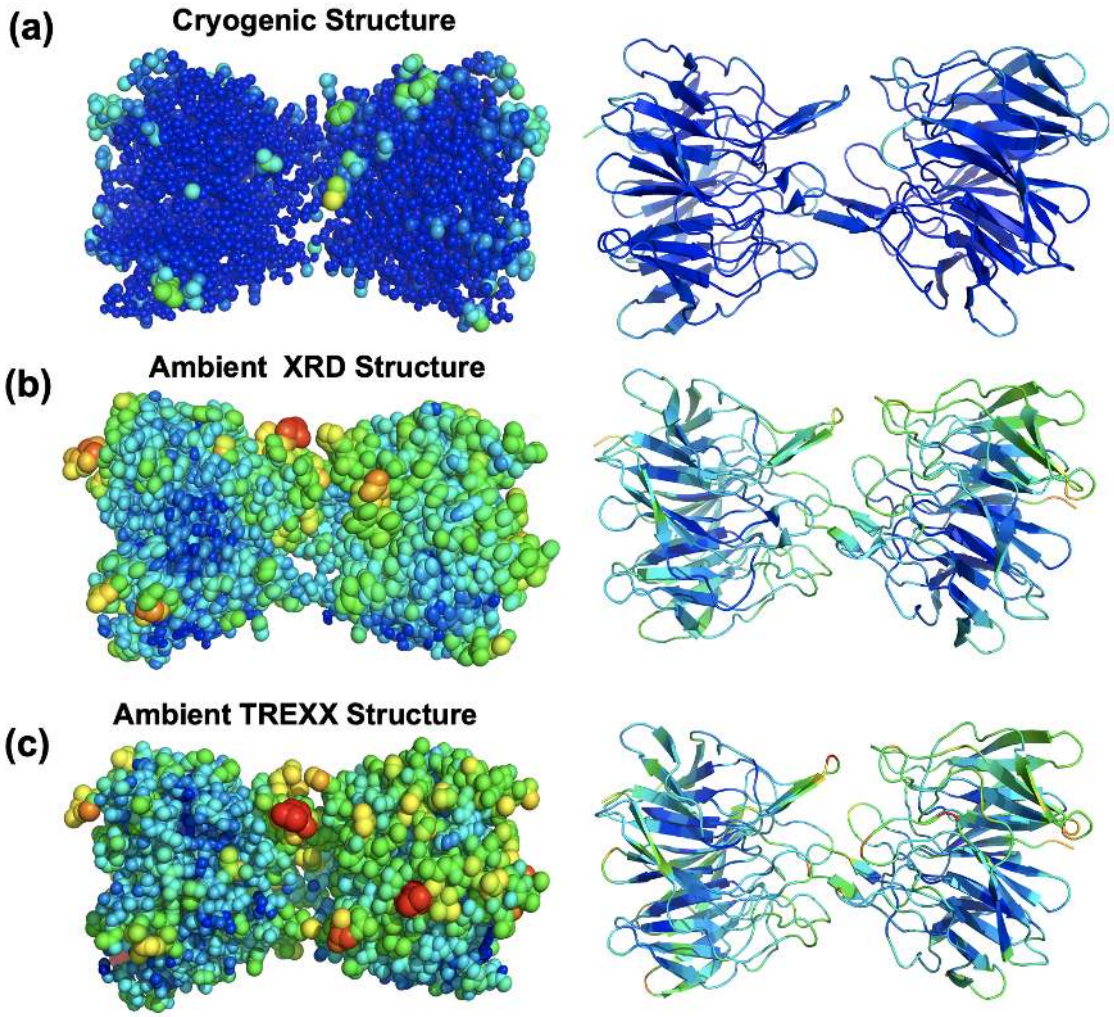

Figure 3.8: B-factor comparison of ambient and cryogenic structures (a) B-factor ellipsoid and cartoon representation of cryogenic structure (PDB ID:6ROG) (b) B-factor ellipsoid and cartoon representation of Ambient XRD structure (c) B-factor ellipsoid and cartoon representation of Ambient TREXX structure emphasizing differences in flexibility between the structures.

3.4.3 Structural Differences in DMF and Nrf2 Binding Residues of Kelch Domain

A comparison of the apo-ambient structure of the Kelch domain with the DMF-bound cryogenic structure of the Kelch domain revealed minor changes in the DMF-binding residues (Figure 3.9-a). These residues are found in the first kelch repeat (Tyr334, Ser363, Val369, and Gly371), second kelch repeat (Asn362), fourth kelch repeat (Arg483), fifth kelch repeat (Try525 and Glu530) and sixth kelch repeat (Phe577 and Try572). The structural alterations were observed on the top side of Keap1 residues Asp382, Tyr334, Phe577, Tyr572, Glu530, Tyr525, and Arg483, while at the bottom side, residues Val369, Gly371. Compared to the apo ambient structure and the apo cryogenic structure, conformational changes are similar to those of the DMF-bound structure (Figure 3.9-b).

A comparison of the apo ambient structure of the Kelch domain with the ETGE peptide-bound and DLG peptide-bound cryogenic structures of the Kelch domain indicated conformational changes in residues interacting with the Nrf2 protein as a result of the temperature shift (Figure 3.10-a). The conformational changes observed were at residues Arg483, Ser508, Ser555, Arg415, Ser602, Ser363, Arg380, and Asn382, which interact with the ETGE and DLG motifs. In comparison to the apo ambient structure and the apo cryogenic structure, the conformational changes are similar to those of the ETGE and DLG peptide-bound structure (Figure 3.10-b).

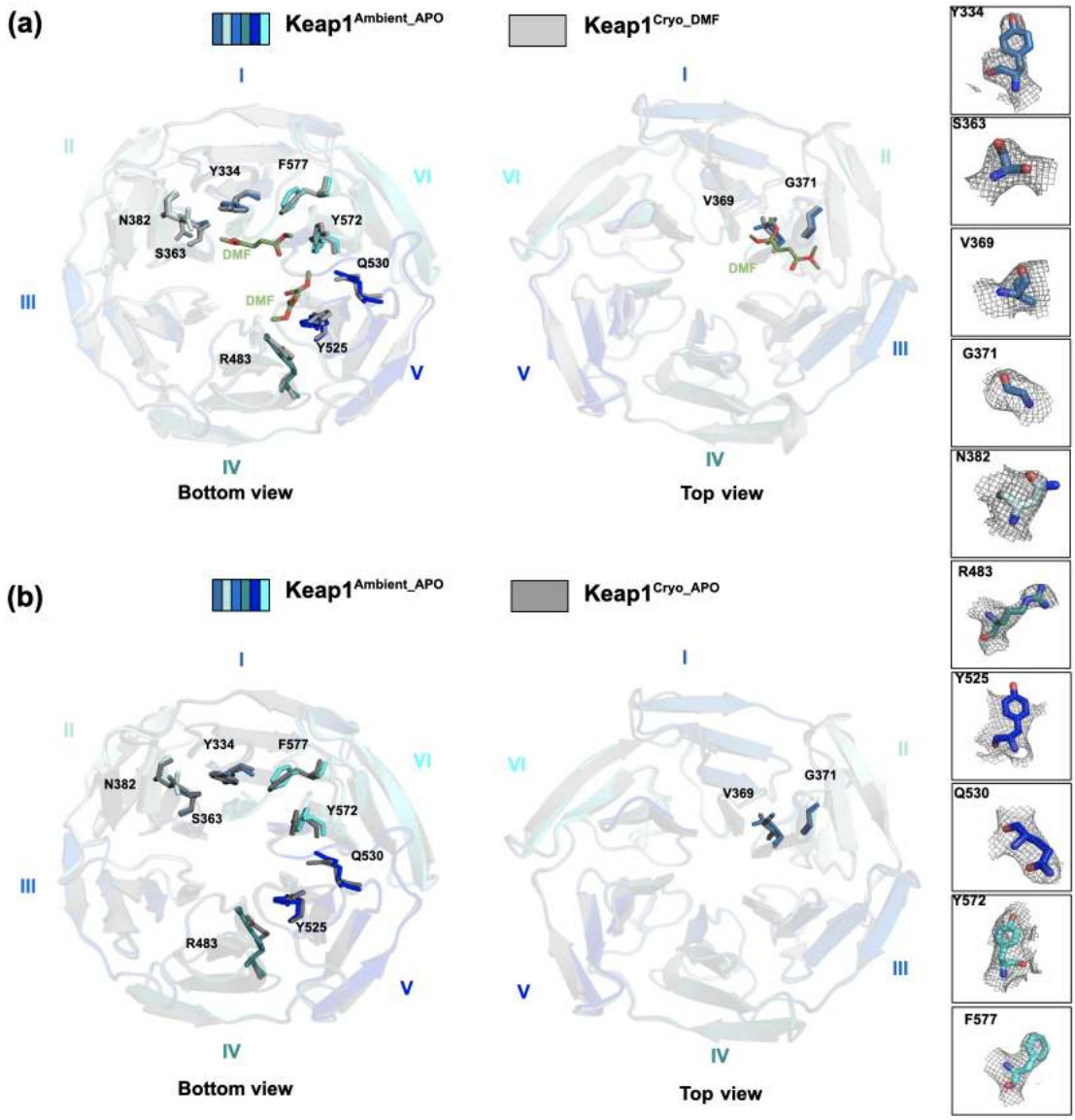

Figure 3.9: Investigation of residues interacting with DMF in cryogenic and ambient structures (a) The superposition of the apo ambient and cryogenic (bound with DMF, colored in gray80) structures is shown from both the bottom and top views, with DMF binding residues displayed in stick representation.(b) Apo ambient and cryogenic (colored in gray50) structures are superposed, with DMF interaction residues also shown in stick representation. DMF interaction residues are displayed with their respective electron density.

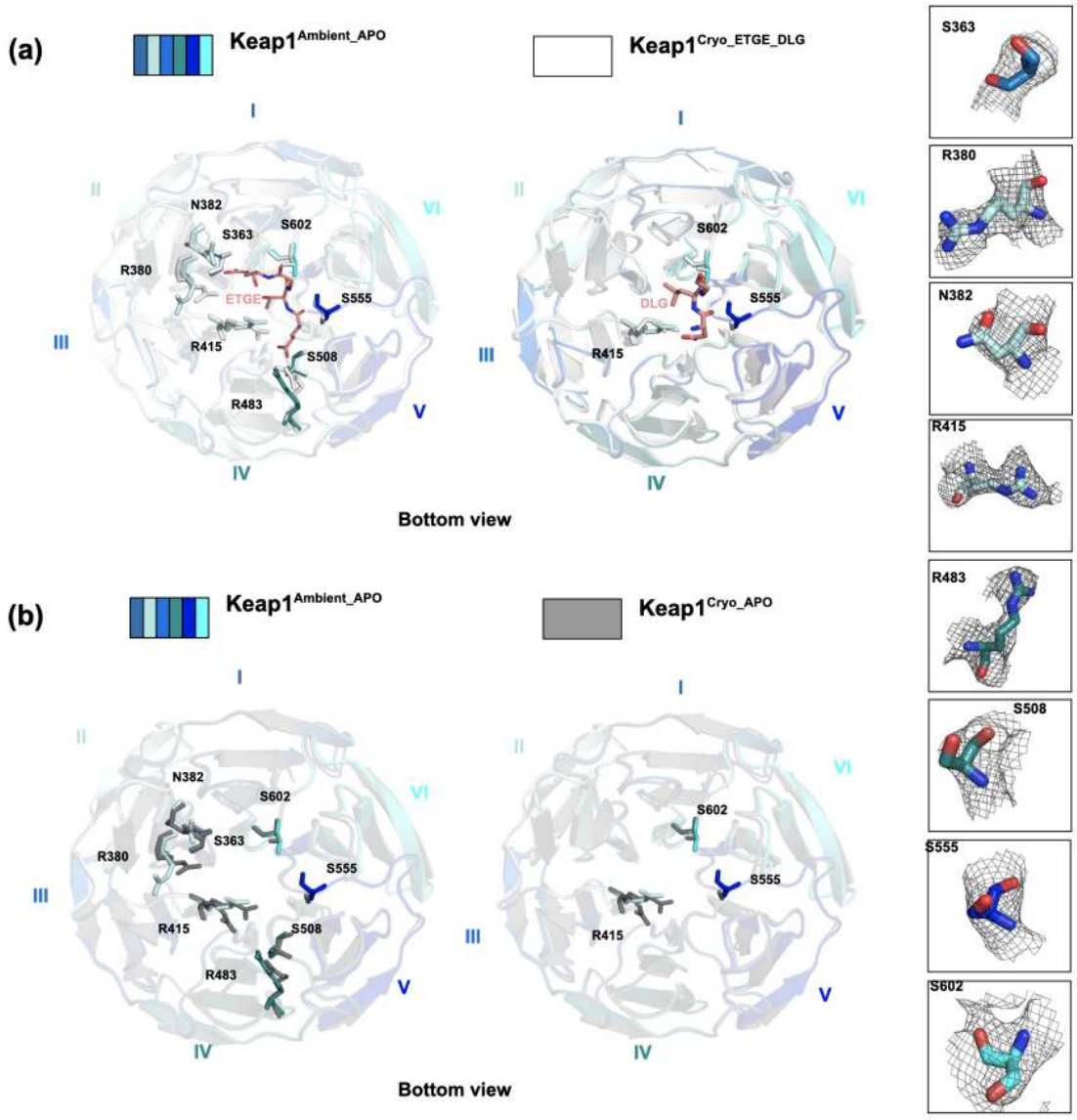

Figure 3.10: Investigation of residues interacting with Nrf2 DLG and ETGE peptides in cryogenic and ambient structures (a) The superposition of ambient structure of Keap1 and cryogenic structure bound with ETGE and DLG peptides (colored in white), with interacting residues displayed in stick representation. (b) Superposition of ambient structure of Keap1 and cryogenic apo structure (colored in gray50), with ETGE and DLG interacting residues also shown in stick representation. Nrf2 interacting residues are displayed with their respective electron density.

3.4.4 Molecular Docking Studies for Ambient Structure Kelch Domain

Molecular docking analysis was performed with 100 fumarate derivatives. The study shows 9 DMF derivatives have more significant affinity profiles and docking values than known fumarates such as DMF, MEF, fumarate (FUM), and monomethyl fumarate (MMF). The docking value of ZINC 12433145 was determined to be -9.130 kcal/mol, which was more significant compared to DMF, itaconate (ITA), MMF, MEF, and FUM docking values, which were -8.423, -7.533, -7.456, -7.146, and -6.934 kcal/mol, and respectively. Similar to DMF, ZINC 12433145 showed critical hydrogen bonding with the key residues Asn382 and Tyr334 (Figure 3.11).

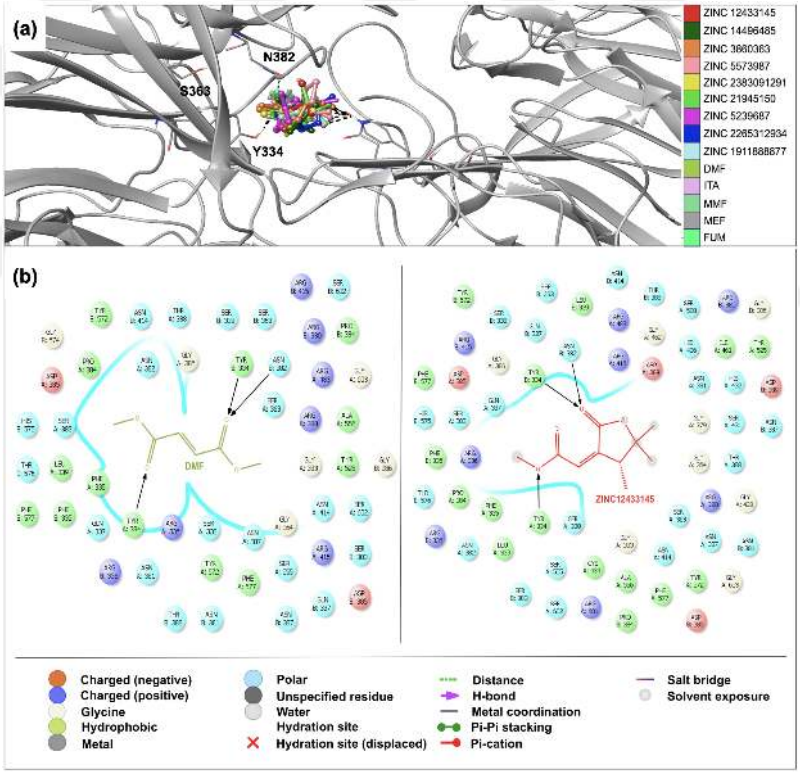

Figure 3.11: (a) Docking poses of fumarate derivatives and DMF, ITA, MMF, MEF, and FUM in the binding site of ambient Keap1 structure. The light gray ribbon diagram in a metallic preset was displayed for all residues. The key residues were represented as sticks and colored by atom type: (C gray, O red, N blue). H-bonding was outlined with dotted black lines. (b) Docking interactions of ZINC 12433145 and DMF in the binding site of ambient Keap1 structure.

The molecular docking analysis was also performed for carnosic acid and 100 carnosic acid derivatives, and compared to fumarate derivates binding affinity, carnosic acid and derivatives displayed a low binding profile in the binding site of the Kelch domain. Carnosic acid was the most prominent one with a docking value of -5.446 kcal/mol compared to its most significant derivative, ZINC 105508677 (-4.758 kcal/mol). Carnosic acid and ZINC 105508677 demonstrate specific binding profiles than fumarate and derivatives, forming hydrogen bonding and stacking with especially arginine residues (Figure 3.12).

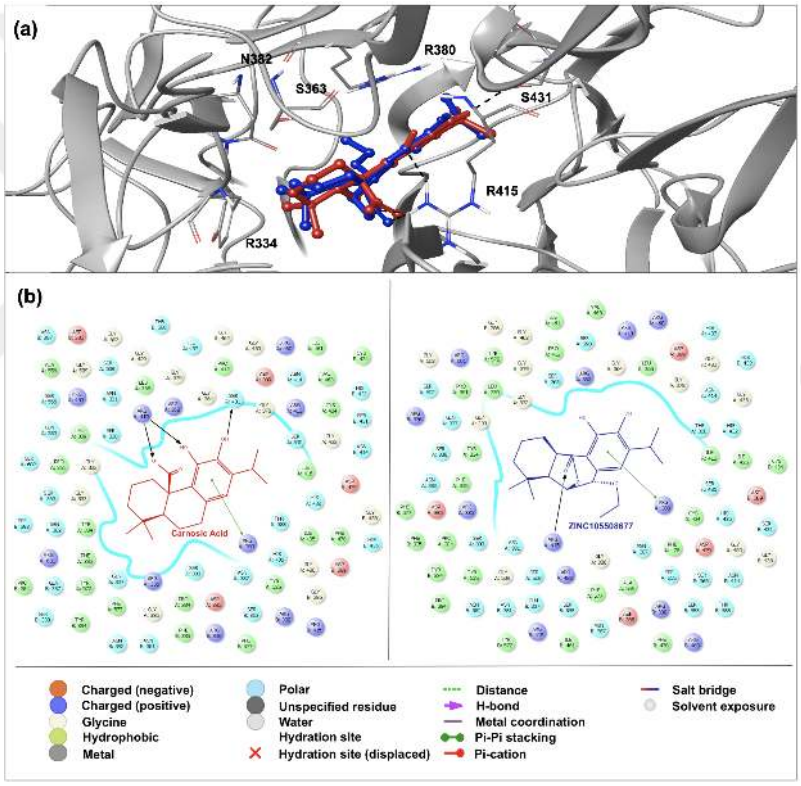

Figure 3.12: (a) Docking poses of carnosic acid and ZINC 105508677 in the binding site ambient structure of Keap1. The light gray ribbon diagram in a metallic preset was displayed for all residues. The key residues were represented as sticks and colored by atom type: (C gray, O red, N blue). H-bond and – stacking interactions were outlined with dotted black and green lines, respectively. (b) Docking interactions of carnosic acid and ZINC 105508677 (colored in red and blue, respectively) in the binding site of the ambient structure of Keap1.

Finally, molecular docking was performed with *L*-carnosine and its 100 derivatives. The analysis showed that *L*-carnosine and *L*-histidyl hydrazide (CNN) is the most potent drug candidate among the *L*-carnosine derivatives, with a value of -11.532 kcal/mol. *L*-carnosine derivative ZINC 450609947 (-10.529 kcal/mol) and *L*-carnosine (-9.835 kcal/mol) displayed the promising binding following CNN. The ionization state of CNN showed a strong hydrogen bond and cation interactions with Arg336, Asn382, Tyr334, and Tyr572. The imidazole ring and carbonyl groups contributed significantly to the high affinity for CNN and *L*-carnosine. Conversely, ZINC450609947 showed hydrogen bonding through pyrrolidinone and carboxamide moieties.

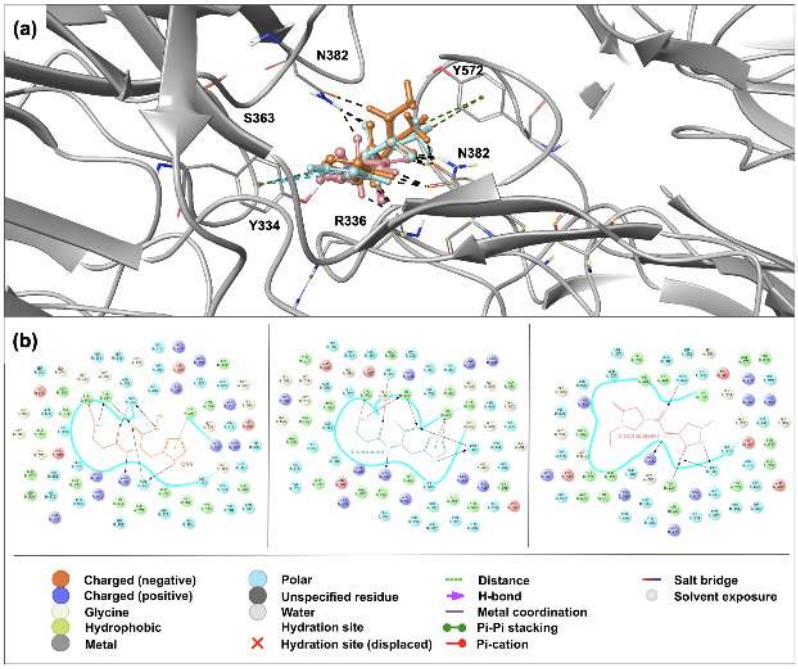

Figure 3.13: (a) Docking poses of *L*-carnosine, CNN, and ZINC 450609947 in the binding site of the ambient structure of Keap1. The light gray ribbon diagram in a metallic preset was displayed for all residues. The key residues were represented as sticks and colored by atom type: (C gray, O red, N blue). H-bond, – stacking and cation interactions were outlined with dotted black, green, and blue lines, respectively. (b) Docking interactions of CNN, *L*-carnosine, and ZINC 450609947 (colored in orange, turquoise, and pink, respectively) in the binding site of Keap1 ambient structure of Keap1.

Chapter 4

DISCUSSION AND CONCLUSION

The 3Å crystal structure of the apo-dimeric Kelch domain of Keap1, determined at ambient temperature by conventional XRD (PDB ID:8X34) and serial synchrotron crystallography, provides detailed structural insights into the dimeric Kelch domain. The Kelch domain of Keap1 dimerizes with a small difference between the monomers, with an RMSD score of 0.21 Å. In previous studies, the monomeric structure of the Kelch domain was shown at cryogenic temperatures. Only one structure of the dimeric Kelch domain has been published in the Protein Data Bank, PDB ID:6ROG. Nevertheless, the dimerization of the Kelch domain has not yet been investigated.

A comparative analysis of the apo ambient and cryogenic structures demonstrated an RMSD score of 0.92 Å, associated with temperature-dependent conformational differences. Specific residues involved in dimerization, such as Pro384, Asp385, and Gly386, showed correlated fluctuations across the chains. In previous studies, it has been demonstrated that full-length Keap1 exists in a dimeric form, with dimerization mediated by the BTB domain, and that the proximity of the Kelch domains might be critical for the Keap1-Nrf2 interaction (Ogura et al., 2010).The existence of the dimeric Kelch domain requires more investigation to unravel the enthalpy-entropy-based binding dynamics of the Keap1-Nrf2 interaction.

The alterations in critical residues such as Tyr334, Ser363, Val369, Gly371, Arg483, Tyr525, Glu530, Tyr572, and Phe577 were revealed by comparing ambient and cryogenic apo and DMF-bound structures. Specifically, the residues on the top side of Keap1, which included Asp382, Tyr334, Phe577, Tyr572, Glu530, Tyr525, and Arg483, and those on the bottom side, such as Val369 and Gly371 displayed conformational changes. These variations were found not only in the DMF-binding residues but also in those participating in the Nrf2 interaction, including Arg483,

Ser508, Ser555, Arg415, Ser602, Ser363, Arg380, and Asn382 when the ambient apo structure was compared with cryogenic ETGE and DLG peptide bound structures. The temperature-dependent conformational shifts of these residues may affect the binding dynamics and stability of Nrf2.

The majority of compounds that function on the Nrf2 pathway have electrophilic properties. These electrophilic compounds bind to Keap1 and allow the translocation of Nrf2 from the cytoplasm to the nucleus. Fumarates with an enone moiety that act as Michael acceptors or catechol-type compounds such as carnosic acid, which can acquire electrophilic potential through different mechanisms, can be cited as examples of these electrophilic compounds. (Satoh et al., 2008; Lu et al., 2016). The Keap1/Nrf2 complex is irreversibly inhibited by these electrophiles. Several potential direct non-covalent inhibitors of Keap1/Nrf2 interactions have been discovered. The reason is that the Kelch binding pocket is highly polar and can interact with carboxylic acid residues in the ETGE and DLG motifs.

L-carnosine is a dipeptide composed of β-alanine and *L*-histidine and is an endogenous antioxidant and free radical scavenger. It has metal chelating, anti-inflammatory, and neuroprotective effects, making it suitable for evaluating metabolic, cardiovascular, and neurodegenerative diseases. Since the *L*-carnosine has anti-inflammatory, antioxidant, antiglycation, and anti-carbonyl effects, it is considered that *L*-carnosine can be effective to the activation of the Nrf2 pathway. Nevertheless, the exact relationship between *L*-carnosine and Nrf2 expression is not well defined (Aldini et al., 2021, Zhou et al., 2021, Caruso et al., 2022;2023). In a previous study, the research group synthesized the hybrid of *L*-carnosine and*L*-histidyl hydrazide (CNN). It was previously reported that CNN exhibited significant 4-hydroxy-trans-2-nonenal (4-HNE) scavenging activity and reduced delayed neuronal death in the hippocampus (Noguchi et al., 2019).

The molecular docking studies were performed for fumarate derivatives, *L*-carnosine, CNN, and other *L*-carnosine derivatives, as well as carnosic acid and its derivatives in the Kelch domain of Keap1. CNN compound showed the most prominent docking results in the Keap1 Kelch domain. CNN established strong interactions with critical residues via both *L*-carnosine and *L*-histidyl hydrazide moieties. The self-

docking studies show that fumarate derivatives interact more effectively with the Kelch domain than well-known fumarates including DMF, ITA, MMF, MEF and FUM.

Specifically, the result of molecular docking of ZINC12433145 in the Kelch domain suggested that converting one of the alkyl esters to a lactone (cyclic ester) might increase the affinity of fumarate derivatives. The hydrogen bond distances between ZINC12433145 and Tyr334 were determined to be greater than that of DMF, although ZINC12433145 and DMF showed similar binding profile patterns. This results in increased occupation of the entire ligand in the Kelch domain, leading to higher affinity. On the other hand, carnosic acid and its derivatives, in spite of their prominent carboxylic acid and ester groups, respectively, showed a lower binding efficiency than all the tested compounds.

To conclude, the study provides an extensive analysis of the dimeric Kelch domain of Keap1, highlighting the differences in the structural dynamics between ambient and cryogenic temperatures and their implications for the interactions between Keap1 and Nrf2. The observed temperature-based conformational shifts, particularly in key residues involved in dimerization and Nrf2 and DMF binding, suggest potential implications for the binding dynamics and stability of the Keap1-Nrf2 complex. Notably, CNN, a hybrid of *L*-carnosine and *L*-histidyl hydrazide, exhibits a higher binding affinity and a higher potential to be an effective inhibitor when compared to fumarate and carnosic acid derivatives. These findings emphasize the importance of further investigation of the enthalpy-entropy-based dynamics of Keap1 to understand its interaction with Nrf2 better and to guide the development of more effective therapeutic agents that target the Keap1-Nrf2 signaling pathway.

BIBLIOGRAPHY

Adamson, R. J., Payne, N. C., Bartual, S. G., Mazitschek, R., Bullock, A. N. (2023). Structural and biochemical characterization establishes a detailed understanding of KEAP1-CUL3 complex assembly. Free Radical Biology and Medicine, 204, 215-225.

Agu, P. C., Afiukwa, C. A., Orji, O. U., Ezeh, E. M., Ofoke, I. H., Ogbu, C. O., Ugwuja, E. I., Aja, P. M. (2023). Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management. Scientific reports, 13(1), 13398. https://doi.org/10.1038/s41598-023-40160-2

Aldini, G., de Courten, B., Regazzoni, L., Gilardoni, E., Ferrario, G., Baron, G.,.. Carini, M. (2021). Understanding the antioxidant and carbonyl sequestering activity of carnosine: Direct and indirect mechanisms. Free radical research, 55(4), 321-330.

Baird, L., Yamamoto, M. (2020). The Molecular Mechanisms Regulating the KEAP1-NRF2 Pathway. Molecular and cellular biology,40(13),e00099-20.

Beamer, L. J., Li, X., Bottoms, C. A., Hannink, M. (2005). Conserved solvent and side-chain interactions in the 1.35 Å structure of the Kelch domain of Keap1. Acta Crystallographica Section D: Biological Crystallography, 61(10), 1335-1342.

Bellezza, I., Giambanco, I., Minelli, A., Donato, R. (2018). Nrf2-Keap1 signaling in oxidative and reductive stress. Biochimica et biophysica acta. Molecular cell research, 1865(5), 721–733.

Canning, P., Cooper, C. D. O., Krojer, T., Murray, J. W., Pike, A. C. W., Chaikuad, A., Keates, T., Thangaratnarajah, C., Hojzan, V., Marsden, B. D., Gileadi, O., Knapp, S., von Delft, F., Bullock, A. N. (2013). Structural basis for Cul3 protein assembly with the BTB-Kelch family of E3 ubiquitin ligases. The Journal of biological chemistry, 288(11), 7803–7814.

Canning, P., Sorrell, F. J., Bullock, A. N. (2015). Structural basis of Keap1

interactions with Nrf2. Free Radical Biology and Medicine, 88, 101-107.

Caruso, G., Di Pietro, L., Cardaci, V., Maugeri, S., Caraci, F. (2023). The therapeutic potential of carnosine: Focus on cellular and molecular mechanisms. Current Research in Pharmacology and Drug Discovery, 4, 100153.

Caruso, G., Privitera, A., Antunes, B. M., Lazzarino, G., Lunte, S. M., Aldini, G., Caraci, F. (2022). The therapeutic potential of carnosine as an antidote against drug-induced cardiotoxicity and neurotoxicity: Focus on nrf2 pathway. Molecules, 27(14), 4452.

Chauhan, N., Chaunsali, L., Deshmukh, P., Padmanabhan, B. (2013). Analysis of dimerization of BTB-IVR domains of Keap1 and its interaction with Cul3, by molecular modeling. Bioinformation, 9(9), 450–455.

Chaharbakhshi, E., Jemc, J. C. (2016). Broad-complex, tramtrack, and bric-à-brac (BTB) proteins: Critical regulators of development. Genesis (New York, N.Y. : 2000), 54(10), 505–518. https://doi.org/10.1002/dvg.22964

Ciftci H, Sever B, Ayan E, Can M, DeMirci H, Otsuka M, TuYuN AF, Tateishi H, Fujita M. Identification of New L-Heptanoylphosphatidyl Inositol Pentakisphosphate Derivatives Targeting the Interaction with HIV-1 Gag by Molecular Modelling Studies. Pharmaceuticals (Basel). 2022 Oct 12;15(10):1255. doi: 10.3390/ph15101255.

Crisman, E., Duarte, P., Dauden, E., Cuadrado, A., Rodríguez-Franco, M. I., López, M. G., León, R. (2023). KEAP1-NRF2 protein-protein interaction inhibitors: Design, pharmacological properties and therapeutic potential. Medicinal research reviews, 43(1), 237–287. https://doi.org/10.1002/med.21925

Dayalan Naidu, S., Dinkova-Kostova, A. T. (2020). KEAP1, a cysteine-based sensor and a drug target for the prevention and treatment of chronic disease. Open biology, 10(6), 200105.

Dinkova-Kostova, A. T., Kostov, R. V., Canning, P. (2017). Keap1, the cysteine-based mammalian intracellular sensor for electrophiles and oxidants. Archives of biochemistry and biophysics, 617, 84-93.

Emsley, P., Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta crystallographica. Section D, Biological crystallography, 60(Pt 12 Pt 1), 2126–2132. https://doi.org/10.1107/S0907444904019158

Ertem, F. B., Guven, O., Buyukdag, C., Gocenler, O., Ayan, E., Yuksel, B., ... DeMirci, H. (2022). Protocol for structure determination of SARS-CoV-2 main protease at near-physiological-temperature by serial femtosecond crystallography. STAR protocols, 3(1), 101158.

Fukutomi, T., Takagi, K., Mizushima, T., Ohuchi, N., Yamamoto, M. (2014). Kinetic, thermodynamic, and structural characterizations of the association between Nrf2-DLGex degron and Keap1. Molecular and cellular biology, 34(5), 832–846. https://doi.org/10.1128/MCB.01191-13 Gevorkov, Y., Yefanov, O., Barty, A., White, T. A., Mariani, V., Brehm, W., ... Chapman, H. N. (2019). XGANDALF–extended gradient descent algorithm for lattice finding. Acta Crystallographica Section A: Foundations and Advances, 75(5), 694-704.

Giegé, R., Lorber, B., Théobald-Dietrich, A. (1994). Crystallogenesis of biological macromolecules: facts and perspectives. Acta crystallographica. Section D, Biological crystallography, 50(Pt 4), 339–350.

Gul, M., Ayan, E., Destan, E., Johnson, J. A., Shafiei, A., Kepceoğlu, A., ... DeMirci, H. (2023). Rapid and efficient ambient temperature X-ray crystal structure determination at Turkish Light Source. Scientific reports, 13(1), 8123.

Horie, Y., Suzuki, T., Inoue, J., Iso, T., Wells, G., Moore, T. W., Mizushima, T., Dinkova-Kostova, A. T., Kasai, T., Kamei, T., Koshiba, S., Yamamoto, M. (2021). Molecular basis for the disruption of Keap1-Nrf2 interaction via Hinge Latch mechanism. Communications biology, 4(1), 576. https://doi.org/10.1038/s42003-021-02100-6

Hu, C., Eggler, A. L., Mesecar, A. D., Van Breemen, R. B. (2011). Modification of keap1 cysteine residues by sulforaphane. Chemical research in toxicology, 24(4), 515-521.

Li, J., Chai, Q. Y., Liu, C. H. (2016). The ubiquitin system: a critical regulator of innate immunity and pathogen–host interactions. Cellular molecular immunology, 13(5), 560-576.

Li, X., Zhang, D., Hannink, M., Beamer, L. J. (2004). Crystal structure of the Kelch domain of human Keap1. The Journal of biological chemistry, 279(52), 54750–54758. https://doi.org/10.1074/jbc.M410073200

Lo, S. C., Li, X., Henzl, M. T., Beamer, L. J., Hannink, M. (2006). Structure of the Keap1: Nrf2 interface provides mechanistic insight into Nrf2 signaling. The EMBO journal, 25(15), 3605-3617.

Lu, M. C., Ji, J. A., Jiang, Z. Y., You, Q. D. (2016). The Keap1–Nrf2–ARE pathway as a potential preventive and therapeutic target: an update. Medicinal research reviews, 36(5), 924-963.

Manai, F., Amadio, M. (2022). Dimethyl fumarate triggers the antioxidant defense system in human retinal endothelial cells through Nrf2 activation. Antioxidants, 11(10), 1924.

Mevissen, T. E., Komander, D. (2017). Mechanisms of deubiquitinase specificity and regulation. Annual review of biochemistry, 86(1), 159-192.

Metzger, M. B., Pruneda, J. N., Klevit, R. E., Weissman, A. M. (2014). RING-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 1843(1), 47-60.

McPherson A. (2017). Protein Crystallization. Methods in molecular biology (Clifton, N.J.), 1607, 17–50. https://doi.org/10.1007/978-1-4939-7000-1_2

McPherson, A., Gavira, J. A. (2014). Introduction to protein crystallization. Acta crystallographica. Section F, Structural biology communications, 70(Pt 1), 2–20. https://doi.org/10.1107/S2053230X13033141

Morris, G. M., Huey, R., Olson, A. J. (2008). Using AutoDock for ligand-receptor docking. Current protocols in bioinformatics, Chapter 8,

Morrow, J. K., Lin, H. K., Sun, S. C., Zhang, S. (2015). Targeting ubiquitination for cancer therapies. Future medicinal chemistry, 7(17), 2333–2350.

Noguchi K, Ali TFS, Miyoshi J, Orito K, Negoto T, Biswas T, Taira N, Koga R, Okamoto Y, Fujita M, Otsuka M, Morioka M. Neuroprotective effects of a novel carnosine-hydrazide derivative on hippocampal CA1 damage after transient cerebral ischemia. Eur J Med Chem. 2019 Feb 1;163:207-214.

Ogura, T., Tong, K. I., Mio, K., Maruyama, Y., Kurokawa, H., Sato, C., Yamamoto, M. (2010). Keap1 is a forked-stem dimer structure with two large spheres enclosing the intervening, double glycine repeat, and C-terminal domains. Proceed-

ings of the National Academy of Sciences, 107(7), 2842-2847.

Peng, Y., Liu, J., Inuzuka, H., Wei, W. (2023). Targeted protein posttranslational modifications by chemically induced proximity for cancer therapy. The Journal of biological chemistry, 299(4), 104572. https://doi.org/10.1016/j.jbc.2023.104572

Saito, R., Suzuki, T., Hiramoto, K., Asami, S., Naganuma, E., Suda, H., ... Yamamoto, M. (2016). Characterizations of three major cysteine sensors of Keap1 in stress response. Molecular and cellular biology.

Satoh, T., Kosaka, K., Itoh, K., Kobayashi, A., Yamamoto, M., Shimojo, Y., ... Lipton, S. A. (2008). Carnosic acid, a catechol-type electrophilic compound, protects neurons both in vitro and in vivo through activation of the Keap1/Nrf2 pathway via S-alkylation of targeted cysteines on Keap1. Journal of neurochemistry, 104(4), 1116-1131.

Scannevin, R. H., Chollate, S., Jung, M. Y., Shackett, M., Patel, H., Bista, P., ... Rhodes, K. J. (2012). Fumarates promote cytoprotection of central nervous system cells against oxidative stress via the nuclear factor (erythroid-derived 2)-like 2 pathway. Journal of Pharmacology and Experimental Therapeutics, 341(1), 274-284.

Shilovsky, G. A., Dibrova, D. V. (2023). Regulation of cell proliferation and nrf2-mediated antioxidant defense: conservation of keap1 cysteines and nrf2 binding site in the context of the evolution of KLHL family. Life, 13(4), 1045.

Smyth, M. S., Martin, J. H. (2000). x ray crystallography. Molecular pathology : MP, 53(1), 8–14. https://doi.org/10.1136/mp.53.1.8

Song, Y., Qu, Y., Mao, C., Zhang, R., Jiang, D., Sun, X. (2024). Post-translational modifications of Keap1: the state of the art. Frontiers in cell and developmental biology, 11, 1332049. https://doi.org/10.3389/fcell.2023.1332049

Stellato, F., Oberthür, D., Liang, M., Bean, R., Gati, C., Yefanov, O., ... Chapman, H. N. (2014). Room-temperature macromolecular serial crystallography using synchrotron radiation. IUCrJ, 1(4), 204-212.

Taguchi, K., Fujikawa, N., Komatsu, M., Ishii, T., Unno, M., Akaike, T., ... Yamamoto, M. (2012). Keap1 degradation by autophagy for the maintenance of redox homeostasis. Proceedings of the National Academy of Sciences, 109(34), 13561-

13566.

Tong, K. I., Katoh, Y., Kusunoki, H., Itoh, K., Tanaka, T., Yamamoto, M. (2006). Keap1 recruits Neh2 through binding to ETGE and DLG motifs: characterization of the two-site molecular recognition model. Molecular and cellular biology, 26(8), 2887–2900. https://doi.org/10.1128/MCB.26.8.2887-2900.2006

Unni, S., Deshmukh, P., Krishnappa, G., Kommu, P., Padmanabhan, B. (2021). Structural insights into the multiple binding modes of Dimethyl Fumarate (DMF) and its analogs to the Kelch domain of Keap1. The FEBS journal, 288(5), 1599-1613.

Yang, Q., Zhao, J., Chen, D., Wang, Y. (2021). E3 ubiquitin ligases: styles, structures and functions. Molecular biomedicine, 2(1), 23.

Zhang, D. D., Hannink, M. (2003). Distinct cysteine residues in Keap1 are required for Keap1-dependent ubiquitination of Nrf2 and for stabilization of Nrf2 by chemopreventive agents and oxidative stress. Molecular and cellular biology.

Zheng, N., Shabek, N. (2017). Ubiquitin Ligases: Structure, Function, and Regulation. Annual review of biochemistry, 86, 129–157. https://doi.org/10.1146/annurev-biochem-060815-014922

Zhong, M., Lynch, A., Muellers, S. N., Jehle, S., Luo, L., Hall, D. R., Iwase, R., Carolan, J. P., Egbert, M., Wakefield, A., Streu, K., Harvey, C. M., Ortet, P. C., Kozakov, D., Vajda, S., Allen, K. N., Whitty, A. (2020). Interaction Energetics and Druggability of the Protein-Protein Interaction between Kelch-like ECH-Associated Protein 1 (KEAP1) and Nuclear Factor Erythroid 2 Like 2 (Nrf2). Biochemistry, 59(4), 563–581. https://doi.org/10.1021/acs.biochem.9b00943

Zhou, J. Y., Lin, H. L., Qin, Y. C., Li, X. G., Gao, C. Q., Yan, H. C., Wang, X. Q. (2021). L-carnosine protects against deoxynivalenol-induced oxidative stress in intestinal stem cells by regulating the Keap1/Nrf2 signaling pathway. Molecular Nutrition Food Research, 65(17), 2100406.

Zipper, L. M., Mulcahy, R. T. (2002). The Keap1 BTB/POZ dimerization function is required to sequester Nrf2 in cytoplasm. Journal of Biological Chemistry, 277(39), 36544-36552.

Qu, L., Guo, M., Zhang, H., Chen, X., Wei, H., Jiang, L., ... Chen, Y. (2022). Characterization of the modification of Kelch-like ECH-associated protein

1 by different fumarates. Biochemical and Biophysical Research Communications, 605, 9-15.

Wang, W., Liang, L., Dai, Z., Zuo, P., Yu, S., Lu, Y., ... Yin, Y. (2024). A conserved N-terminal motif of CUL3 contributes to assembly and E3 ligase activity of CRL3KLHL22. Nature Communications, 15(1), 3789.

Weber P. C. (1991). Physical principles of protein crystallization. Advances in protein chemistry, 41, 1–36. https://doi.org/10.1016/s0065-3233(08)60196-5

Wu, J. H., Miao, W., Hu, L. G., Batist, G. (2010). Identification and characterization of novel Nrf2 inducers designed to target the intervening region of Keap1. Chemical biology drug design, 75(5), 475-480.

Wu, S., Lu, H., Bai, Y. (2019). Nrf2 in cancers: A double-edged sword. Cancer medicine, 8(5), 2252–2267. https://doi.org/10.1002/cam4.2101

Appendix A

CHEMICALS

Ampicillin	GOLDBIO
Chloramphenicol	GOLDBIO
Glycerol	ISOLAB
Imidazole	BioFroxx
IPTG	GOLDBIO
LB Agar	Caisson
NaCl	ISOLAB
HCl	ISOLAB
Ni-NTA	QIAGEN
Paraffin Oil	Tekkim Kimya
SDS	Sigma-Aldrich
Tris Base	Sigma-Aldrich
Triton X-100	Sigma-Aldrich
Tryptone	BD Biosciences
Yeast Extract	BD Biosciences

Appendix B

EQUIPMENT

AKTA go	Cytiva
Amicon Ultra Centrifugal filters	Merck Millipore
Beckman Allegra 15R Centrifuge	Beckman Coulter
Beckman Avanti J-26S	Beckman Coulter
Beckman Optima™ L-80 XP Ultracentrifuge	Beckman Coulter
Beckman Ti-45 Rotor	Beckman Coulter
Branson W250 Sonifier	Branson
Cryotubes	Wuxi NEST Biotechnology
Eppendorf Tubes	Eppendorf New Brunswick
Falcon Tubes	FIRATMED
Innova 44R Incubator Shaker	Eppendorf New Brunswick
Innova 4430R Incubator Shaker	Eppendorf New Brunswick
Kimtech Science KimWipes	Kimberly-Clark
NanoDrop™ 2000c	Thermo Scientific
Terasaki Plates	Greiner Bio-One
TGX-Mini Protean Gradient SDS-PAGE System	Bio-Rad Laboratories
Millipore Milli-Q Water Purification System	Merck Millipore
Oxford Cryosystems Cryostream 800 Plus	Oxford Cryosystems
Rigaku XtaLAB Synergy Flow XRD	Rigaku
Superdex™75 Increase Column	Cytiva

Appendix C

PROTEIN LADDER

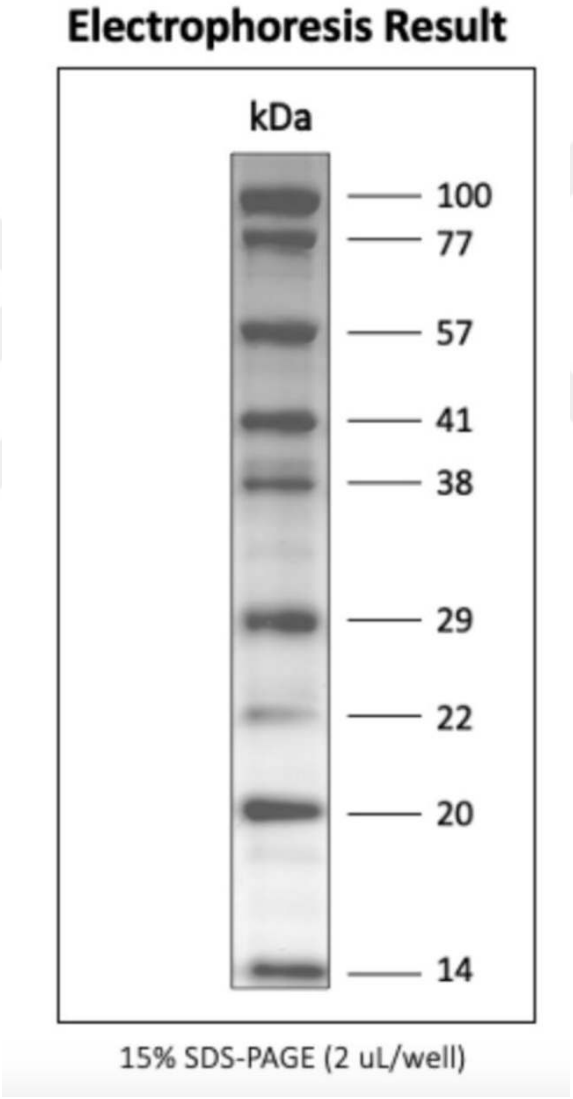

Figure C.1: KUYBIIG-M Protein Ladder.

Appendix D

CRYSTALLIZATION SCREENING SOLUTIONS

3D Structure Screen	Molecular Dimensions
Additive Screen	Hampton Research
Clear Strategy Screen	Molecular Dimensions
CryoPro	Hampton Research
Crystal Screen	Hampton Research
Crystal Screen Cryo	Hampton Research
Detergent Screen	Hampton Research
Grid Screen Ammonium Sulfate	Hampton Research
Grid Screen MPD	Hampton Research
Grid Screen PEG/LiCl	Hampton Research
Grid Screen PEG6000	Hampton Research
Grid Screen Sodium Chloride	Hampton Research
Helix	Molecular Dimensions
Index	Hampton Research
Ionic Liquid Screen	Hampton Research
JBScreen Nuc-Pro	Jena Bioscience
JCSG	Molecular Dimensions
MacroSol	Molecular Dimensions
MembFac	Hampton Research
MIDAS	Molecular Dimensions
Morpheus	Molecular Dimensions
MultiXtal	Molecular Dimensions
Natrix	Hampton Research
NeXtal Protein Complex Suite	NeXtal Biotechnologies
NR-LBD	Molecular Dimensions

PACT Premier	Molecular Dimensions
PEG/Ion	Hampton Research
PEGRx	Hampton Research
PGA Screen	Molecular Dimensions
ProPlex Screen	Molecular Dimensions
Quick Screen	Hampton Research
SaltRx	Hampton Research
Structure Screen	Molecular Dimensions
Stura FootPrint Screen	Molecular Dimensions
Wizard	Rigaku
Wizard Cryo	Rigaku
Wizard Precipitant Synergy	Rigaku