

**A Two-Stage Mathematical Programming Algorithm for
Predicting
Secondary Structures of Proteins**

by

736749

Özlem Yilmaz

**A Thesis Submitted to the
Graduate School of Engineering
in Partial Fulfillment of the Requirements for
the Degree of**

**Master of Science
in
Industrial Engineering**

Koç University

October 2003

**YÜKSEK LİSANS TEZİ
KÜTÜPHANESİNE KAYITLI
T.C. YÖKSEKÖĞRETİM KURULU
BİLİMSEL ARAŞTIRMA VE TEKNOLOJİ YATIRIM
KURUMU**

Koç University
Graduate School of Sciences and Engineering

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

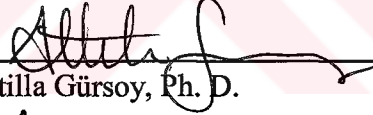
Özlem Yilmaz

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.

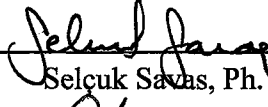
Committee Members:



Burak Erman, Ph. D.



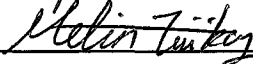
Atilla Gürsoy, Ph. D.



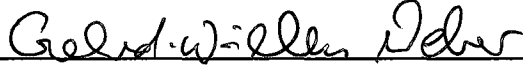
Selçuk Savas, Ph. D. (Advisor)



Uğur Sezerman, Ph. D.



Metin Türkay, Ph. D. (Advisor)



Gerhard Wilhelm Weber, Ph. D.

Date:

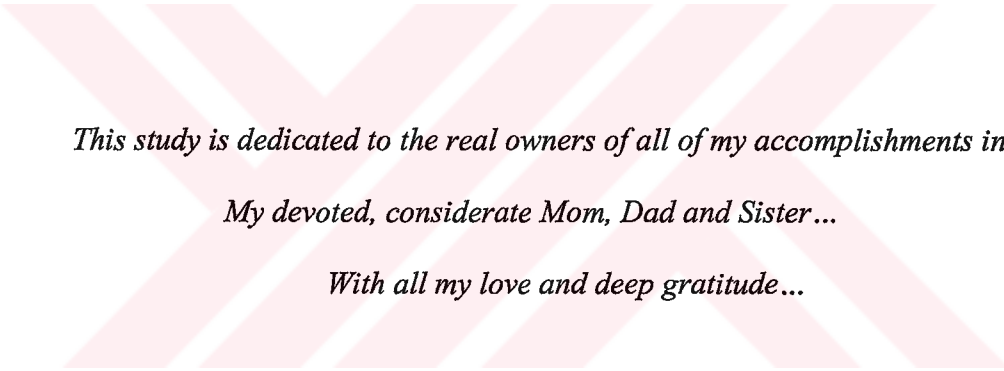
24.10.2003

ABSTRACT

Recognition of function of protein is one of the major steps in drug discovery process and it is known that functions of polypeptides are dictated by their structures. This study aims to predict secondary structures of proteins utilizing a two-stage probabilistic algorithmic approach.

First stage is determination of structural class (all- α , all- β , α/β or $\alpha+\beta$) of unknown protein by a Mixed Integer Linear Programming formulation. Next phase is searching 3 to 7-residues-long segments of unknown protein's residue sequence in database of structures of experimentally identified proteins, belonging to the same structural class determined for the unknown one. Source for data on structures of known proteins is Protein Data Bank; <http://www.pdb.org>. Three states (α -helix, β -sheet or loop) per residue predictions are obtained through a probabilistic approach utilizing outcomes of database search process.

We achieved 100% accuracy in folding type determination phase. Weights to put on probabilities obtained from 3 to 7-residues-long segments for each structural class are determined optimally for each structural class via Non-Linear Programming Formulations again utilizing structural information on experimentally identified proteins. The prediction method is tested on 419, 579, 707 and 601 known proteins from all- α , all- β , α/β and $\alpha+\beta$ classes, respectively, as if they were not known. 3-states-per-residue accuracy levels obtained for all- α , all- β , α/β and $\alpha+\beta$ classes are 80.5%, 72.4%, 71.9% and 75.5%, respectively.



This study is dedicated to the real owners of all of my accomplishments in life:

My devoted, considerate Mom, Dad and Sister...

With all my love and deep gratitude...

ACKNOWLEDGEMENTS

I would like to gratefully acknowledge supervision of my professors
Metin Türkay and Selçuk Savas and thank them with most sincere regards for
their trust, inspiration and support. I also would like to thank Sibel Bilge Sonuç,
Eda Kaya and Emrah Coskun for their help.

TABLE OF CONTENTS

	Page
Abstract	iii
Acknowledgments	v
Table of Contents	vi
List of Tables	xi
List of Figures	xiii
 Chapter 1: Introduction	 1
1.1 Drug Discovery Process	2
1.1.1 Genome Analysis	3
1.1.2 Target Selection	3
1.1.3 Protein Function (Structure)	4
1.1.4 Lead Discovery	4
1.1.5 Lead Optimization	5
1.1.6 Formulation Chemistry	5
1.1.7 Toxicity Analysis	6
1.1.8 Clinical Trials Phase I, II, III	7
1.1.8.1 Clinical Trial Phase I	7
1.1.8.2 Clinical Trial Phase II	8
1.1.8.3 Clinical Trial Phase III	9
1.2 Role of Protein Structure in Drug Discovery	10
1.3 Proteins: Molecules of Life	11
1.3.1 Introduction	11
1.3.2 Amino Acids	12

1.3.3	Structures of Proteins	14
1.3.3.1	Primary Structure	15
1.3.3.2	Secondary Structure	15
1.3.3.2.1	Types of Secondary Structures	16
1.3.3.2.1.1	α -helices	16
1.3.3.2.1.2	β -sheets	17
1.3.3.2.1.3	Connecting Structures (Loops, turns, coils, etc.)	17
1.3.3.2.2	Classification of Proteins According to Their Secondary Structures	18
1.3.3.3	Tertiary Structure	19
1.3.3.4	Quaternary Structure	19
1.4	Determination of Structures of Proteins	19
1.4.1	X-Ray crystallography	19
1.4.2	Nuclear Magnetic Resonance (NMR) Spectroscopy	20
1.5	Protein Secondary Structure Prediction	22
1.5.1	Protein Secondary Structure Prediction Methods	22
1.5.2	Performances of Prediction Methods	24
1.5.3	The State-of-the-art	27
1.5.4	Approach of This Study	28
Chapter 2:	Folding Type Prediction	30
2.1	Introduction	30
2.2	Mixed-Integer Programming Formulation	31
Chapter 3:	Secondary Structure Prediction	34
3.1	Introduction	34

3.2	Data Analysis and Formation of Databases	36
3.3	Folding Type Determination	48
3.4	Secondary Structure Prediction Algorithm	48
3.4.1	Algorithm for Searching Segments of Sequence in Database	48
3.4.2	Illustrative Example for Search Algorithm	54
3.4.3	Statistical Secondary Structure Prediction	56
3.4.3.1	Secondary Structure Prediction Approach	56
3.4.3.2	Determination of Optimal Weights for Probabilities	59
3.4.3.3	Illustrative Example for Statistical Secondary Structure Prediction.	62
Chapter 4:	Results	66
4.1	Determination of Secondary Structure	66
4.2	Optimal Weights of Residue Segment Length	70
4.3	Prediction Results	71
4.4	Discussion of Results	72
Chapter 5:	Conclusions and Recommendations	75
5.1	Major Conclusions	75
5.2	Contributions	76
5.2.1	Two-Stage Algorithm Approach	76
5.2.2	Folding Type Determination	76
5.2.3	Statistical Weights	77
5.2.4	Prediction of Secondary Structure	77
5.3	Recommendations	77

Appendices	81
Appendix A: Names of proteins in all- α class in SCOP	82
Appendix B: Names of proteins in all- β class in SCOP	87
Appendix C: Names of proteins in α/β class in SCOP	95
Appendix D: Names of proteins in $\alpha+\beta$ class in SCOP	103
Appendix E: List of proteins that could not be processed in code Converter.c due to problems in “chain problem” category	110
Appendix F: List of proteins that could not be processed in code Converter.c due to problems in “inconsistency” category	111
Appendix G: Names of proteins selected as test data for all- α secondary structural class	117
Appendix H: Names of proteins selected as test data for all- β secondary structural class	118
Appendix I: Names of proteins selected as test data for α/β secondary structural class	119
Appendix J: Names of proteins selected as test data for $\alpha+\beta$ secondary structural class	121
Appendix K: Names of proteins in database for all- α secondary structural class	122
Appendix L: Names of proteins in database for all- β secondary structural class	127
Appendix M: Names of proteins in database for α/β secondary structural class	132
Appendix N: Names of proteins in database for $\alpha+\beta$ secondary structural class	138
Appendix O: Prediction accuracies for each chain of each protein in all- α class in trial set	143
Appendix P: Prediction accuracies for each chain of each protein in all- β class in trial set	154

Appendix Q: Prediction accuracies for each chain of each protein in α/β class in trial set	172
Appendix R: Prediction accuracies for each chain of each protein in $\alpha+\beta$ class in trial set	190
References	206



LIST OF TABLES

Table	Page
1.1: List of amino acids, their three-letter and single-letter representations	14
1.2: Accuracies of Secondary Structure Prediction Methods (Rost, 2001)	25
1.3: Average three-state accuracy indices calculated for six prediction algorithms based on 396 proteins (%) (Cuff and Barton, 1999)	25
3.1: Number of proteins in SCOP's list of each structural class and the number of proteins with successfully processed structures through our program	45
3.2: Number of secondary structures in each structural class database of 80% of data, before and after exclusion of homologous structures via query in the database	47
3.3: Table of frequencies of occurrences of residues in chain A of protein 1QUU	52
3.4: Information on number of proteins in test set and database set	61
3.5: Frequency matrix of illustrative example	62
3.6: Hypothetical weights for illustrative example	63
3.7: Real structure matrix for illustrative example	64
3.8: Predictions of illustrative example	65
4.1: The list of 122 proteins that are used to train the mixed-integer programming model for determining folding types	67
4.2: The list of 61 proteins that are used to test the mixed-integer programming model for determining folding types	69
4.3: Optimal weights for each structural class for segments of 3 to 7 residues	71
4.4: Results of secondary structure prediction for trial data sets from each structural class	72
4.5.a: Accuracies of Secondary Structure Prediction Methods (Rost, 2001)	73

4.5.b: Accuracies of Secondary Structure Prediction Methods (Cuff and Barton, 1999) .	74
4.5.c: Accuracy of Two-Stage Secondary Structure Prediction Method of this thesis	74
5.1: Illustrative example for manual adjustments on secondary structure predictions of the method	79



LIST OF FIGURES

Figure	Page
1.1: Drug Discovery Process	2
1.2: A typical protein molecule	12
1.3: Chemical structure of amino acids	13
1.4: Chemical structure of protein molecule and peptide bond	13
1.5: Primary structure of proteins	15
1.6: α -helix	16
1.7: β -sheet	17
1.8: Connectors	18
2.1.a: Schematic representation of classification of data using hyper planes	31
2.1.b: Schematic representation of classification of data using boxes	31
3.1: Flow of secondary structure prediction algorithm	35
3.2: Sample format for residue sequence part of PDB file of protein 11AS	37
3.3.a: Sample format for part of PDB file explaining α -helical regions for protein 11AS	38
3.3.b: Sample format for part of PDB file explaining β -sheet regions for protein 11AS	39
3.4: Part of output file of C code showing structural information of protein 11AS, to be loaded to database	41
3.5: Flow diagram showing general execution steps of CONVERTER.C	42
3.6: Part of PDB file of protein 1CD3	43
3.7: Part of PDB file of protein 1A1R	45
3.8: Major steps in database establishment	47
3.9: Flowchart for residue-segment search algorithm	54
3.10: 3-residues-long segments searched in illustrative example for database search .	54
3.11: 4-residues-long segments searched in illustrative example for database search .	55

Chapter 1

INTRODUCTION

Proteins are the molecules of life that play a key role in realizing the functions of any biological organism. Discovery of the functions of proteins will enable us to understand the principles of life and working mechanisms of any organism. In the case of humans, this discovery will lead to the design of drugs that will regulate the functions of proteins improving the quality of life.

Functions of proteins are highly correlated to their three dimensional structure. There are experimental methods to determine the protein structure including X-ray diffraction and nuclear magnetic resonance (NMR). These experimental methods are very slow and require large amounts of resources. In order to overcome the shortcomings of experimental methods, researchers have developed a host of methods to predict the protein structures.

A two-stage method to predict the protein structure is presented in this study. This report starts with introduction of drug discovery process to emphasize the role of protein structure prediction in the emerging biotechnology field. Next, a review of the protein structure prediction methods are compared including their reported accuracies. Following chapters include the details of the developed method, analysis of results and prediction accuracy, conclusions and recommendations for future work.

1.1 Drug Discovery Process

Drugs are chemicals given to people in order to treat or prevent diseases. Discovery of a drug is a long process with many phases. The ultimate goal in drug discovery process is not to treat symptoms of disease but to find a complete cure for it. The drug discovery process takes about 15 years and cost more than \$500 million for a single drug on the average [Data from PhRMA Annual Report (2001)] (Robbins, 2000).

Most of the diseases in biological organisms are caused by failure of a protein (or a set of proteins) to perform its function. Biological and chemical activities of organisms can be regulated to prevent diseases by introducing a drug to the biological system that will target the protein that is responsible of the particular disease of interest. The regulation of the biological system is established when a drug molecule interacts with the active site of the target protein. The drug discovery process includes genome analysis, target selection, determination of protein structure and functions, lead discovery and optimization, formulation chemistry, toxicity analysis and clinical trials as shown in Figure 1.1.

Detailed discussion of steps of the drug discovery process are given in the following subsections:

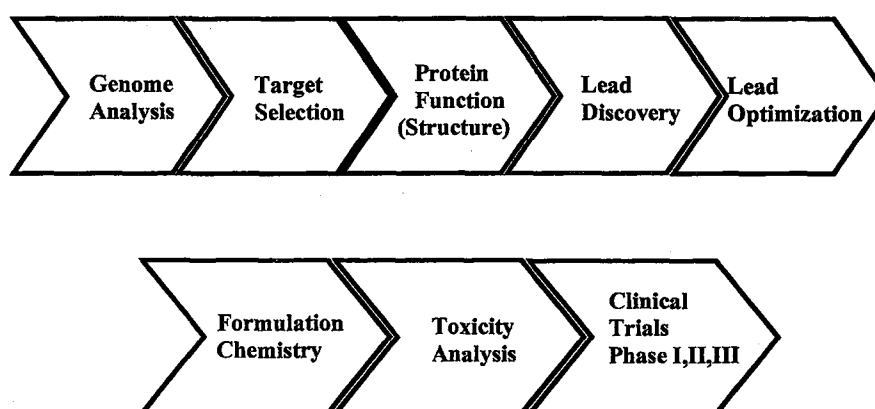


Figure 1.1: Drug Discovery Process

1.1.1 Genome Analysis

Major chemical and biological activities in organisms are initiated, catalyzed, carried out or terminated by biological entities like enzymes or hormones that are basically polypeptide chains. These biological entities have special sequences that determine their unique properties.

The term “genome” means all set of genes of a biological organism. Genes are part of chromosomes that are located in the nucleus of the cell carrying heredity information. Genes are like passwords containing all properties of the individual and each gene is responsible for coding at least one protein. Diseases in biological organisms are caused by the failure of a protein (or a set of proteins) to perform its function in general. Therefore, the gene that codes the particular protein that plays an important role in the disease should be analyzed. The genome sequencing is done using experimental methods, and the most popular two projects on analysis of genetic code are human genome project (Wolfsberg *et al.*, 2001), and mouse genome project (Powell *et al.*, 2002).

1.1.2 Target Selection

When a new drug discovery process is initiated, the “target” gene or protein that plays a central role on the onset and progress of the disease of interest is identified. When the underlying problem resulting with the disease and all related symptoms is determined, the aim can be to change, improve, inhibit or completely block concerned function of the target, depending on the way to prevent illness.

1.1.3 Protein Function (Structure)

When target gene is selected, the protein coded by it must be analyzed to investigate appropriate chemicals that can interact with it to cure the disease. Analysis of a protein does not only mean discovery of its chemical structure or sequence of substructures in its chains but it means discovery of its functions, as well. Functions of proteins and their interactions with other substances in an organism are determined by three-dimensional structures and configurations. Proteins function through their active sites (also called binding site), which are the specific regions playing central role in proteins' interactions with other molecules. The drugs that prevent the onset or progress of a particular disease perform their function by binding to the active site of target proteins. Therefore, understanding the functions and structures of proteins is indispensable in drug discovery process.

1.1.4 Lead Discovery

The next step in the drug discovery process is to identify drug candidates that will interact with the active site of the target protein. In general, a large number of chemical substances that may be suitable drugs are tested on the target protein to observe the effect for interaction. The chemical substances that show any level of interaction are categorized as a hit. The drug candidates interact with the active site to have therapeutic effect by blocking, accelerating, decelerating, reversing or initiating reactions depending on the way to cure the disease. This phase ends up with a list of drug candidates interacting successfully with the target in laboratory conditions.

1.1.5 Lead Optimization

Drug candidates determined in lead discovery phase are further refined using experimental and computational methods. The result is a list of best performing substances to be active pharmaceutical ingredient for the drug. Various factors including maximum therapeutic efficiency, minimum side reactions with other substances or structures that can exist in the environment of the target or other regions of the organism, are considered in the selection process. Sometimes, the best performing drug candidate can involve in reactions with other structures that can exhibit side effects and lead to even other diseases. Also, the drugs enter the blood after they are approved by liver; so it should be assured that the selected drugs do not severely affect or destroy this organ. These facts are analyzed through computations and experiments with relevant substances trying to imitate the real environment as much as possible. Result of this phase is not always the very best performing substance on the target but it is the most suitable drug candidates considering both therapeutic power and various other effects on all organs and systems of the body.

1.1.6 Formulation Chemistry

The next phase is testing the drug candidates in-vivo conditions. Various effects investigated in previous phase are utilized to finalize formulation of active pharmaceutical and to determine other filler substances that will be effective in various steps starting with the entry of the drug in the organism until it is disposed of from the body. The drug together with the fillers should safely reach the target passing through various systems of the body and it should be able to carry out necessary reactions. Another purpose of adding fillers is to minimize drug's probable side effects. The fillers can be ingredients to block or trigger some reactions as well as immune system strengtheners or other related formations.

Minimizing side effects and reactions while maintaining its therapeutic effect on the disease is the main concern at this step. The result is the complete drug to be tested in real organisms to see the performance and its unpredicted effects in the environment of a living organism.

1.1.7 Toxicity Analysis

After carrying out all analysis and experiments on target substances, cells, tissues or organs trying to create the environment of a living organism as much as possible, further tests are initiated on laboratory animals to examine performance of the drug, the extent of its predicted side effects and other probable unexpected effects on different organs and tissues of the body. These experiments demonstrate correct predictions and estimations of design phases and provide strong clues about unexpected effects and probable improvements on ingredients, dosage and method of application of the drug. It is possible to see whether the active pharmaceutical can effectively reach target passing through various systems of the body and whether it can carry out necessary reactions to cure the disease in environment of a real organism. While the extent of forecasted side effects are measured, other unexpected toxic effects can be observed and precautions can be taken by changing the dosage, periods of application or ingredients of the drug. The prepared pharmaceutical can no longer be considered as a drug after experiments of this phase rendering all of the previous steps useless. Although various systems, chemicals, enzymes, hormones and reactions act similarly in animals and humans, there are quite significant differences that prevent scientists from stating a drug successful at the end of these experiments as 'safe' for humans. There are various facts working different in systems and reactions taking place; while some viruses or bacteria cause an illness in humans it can even have no effect on animals or vice versa. Therefore, a drug having desired effects on

animals is not guaranteed to have the same performance and safety in humans' body. Toxicity analysis is the final stage where the drug is tested on living organisms before permission is granted for trials on real patients.

1.1.8 Clinical Trials Phase I, II, III

With proven success on laboratory animals with metabolic and hormonal systems as similar to humans' as possible, the drug candidates are tested on patients and volunteers in hospitals. It usually requires trials on thousands of people before exact design of the drug is completed and this particular phase is divided into three sub-phases each of which can only be initiated after approval of success at the end of previous sub-phase.

1.1.8.1 Clinical Trial Phase I

The most essential fact to be assured before taking any action about a new drug is its safety. For this reason, the first phase of trials is dedicated to investigation of effects of the drug on metabolism and other systems of the body. A group of patients or healthy volunteers are exposed to increasing dosages of drug to observe the level of side effects and to determine extent of toxic effects. Starting with the introduction of the drug into the organism, it is closely monitored until it is completely disposed of from the body. The systems and organs the drug passes through, chemical reactions it undergoes, its success to reach the target, processes that occur at the target and the route it leaves the body or the way it is destroyed is meticulously followed and investigated to make every aspect as clear as possible. This phase is completely dedicated to make safety of the new drug certain to initiate further trials to assess its effectivity on treatment of the relevant disease. It should be emphasized that side or toxic effects are not expected to be completely eliminated for

closing this phase successfully; they most probably remain to exist, but the dosage they emerge and the level they become completely toxic for the patient are determined. The extent of detrimental effects both known from previously held toxicity analysis and the unpredicted ones observed at this phase on humans are measured and the benefits and harms of the drug are evaluated. Thorough investigations at this stage may give feedback for formulation of the drug and previously held formulation chemistry phase can be revised for a better composition to diminish toxic effects or to strengthen body for these by adding new ingredients. This process may result with a complete failure terminating any further tests and returning back to lead optimization or even lead discovery phase for an entirely different new drug candidate. When the trials are successful, safety intervals of drug are settled and the treatment effects overwhelm side effects, scientists are permitted to initiate second phase clinical trials regarding drug's effectiveness and application.

1.1.8.2 Clinical Trial Phase II

Results obtained at phase I guide scientists for design of tests at phase II to assess effectiveness of drug on treatment of disease by investigating its effects on various patients suffering from different forms of the disease at initial phase, maturity phase, and final phase and by discovering most efficient doses and ways to deliver it in metabolism. Main target of the phase II clinical trials is to maximize the drug's effectiveness of treatment while maintaining safety. Dosages and periods of drug intake and the ways to deliver into the body are all tried to discover best composition and delivery mechanism. Groups of patients at different levels and severity of disease are selected and each group is exposed to a different dosage. These groups are composed of sub-ensembles with different periods of drug intake. Another issue of concern is selection of most effective way to make drug enter organism. This fact was also studied in formulation phase but here is the main medium it is

tested. Some people may take it orally while some have intravenous or intramuscular injections. While several groups of people are tested with varying dosages taken in different periods, drugs enter the bodies of these subjects through different routes. Efficiency and effectivity of the drug for treatment is assessed at the end of the period by evaluating its performance on each group of patients suffering from disease at different levels. Results of this stage may vary between a complete failure meaning no more consideration of this substance as a drug, returning back to drug discovery or optimization phase; feedback for improvement of its formulation for better achievements or diminished side effects, or approval of tests to start design of phase III trials. When phase II is terminated, effectiveness of drug for treatment of disease on various types of patients, its best dosage, frequency of intake and way of entrance in organism for each case are decided.

1.1.8.3 Clinical Trial Phase III

This is the final stage where the drug developed is tested on many patients to have supplementary statistical data in order to demonstrate results with the drug to get permission for its serial production and distribution. Amount of data required, range of acceptability and statistical factors to be closely watched change depending on the type of the drug, its probable side effects, duration of treatment and time required for emerge of possible side effects, severity and fatality of the disease and opportunity for patients to be treated with another drug. If patients are suffering from a strictly fatal disease and there is no hope for them other than this pharmaceutical, this final phase can be accelerated to have enough time for patients to have chance to try. Other than such emergency cases, this stage is quite long, procedural and detailed to be as confident about the new product as possible.

1.2 Role of Protein Structure in Drug Discovery

Drug discovery discussed in previous sections is a long and demanding process with an average duration of 15 years and estimated cost of more than \$500 million [Data from PhRMA Annual Report (2001)] (Robbins, 2000). Even a slight shortening in duration of any stage means faster discovery of a drug and quicker entrance to market saving more lives. Cost reduction in any step will mean more patients being able to afford to buy the drug, again ending up with a larger segment of the population having access to treatment of a disease.

Discovery of functions of protein(s) produced by target gene(s) is one of the key steps in drug discovery. As discussed in Section 1.1.3, functions of proteins are highly correlated with their structures. The determination of protein structure is possible by using sophisticated experimental techniques such as X-ray diffraction in synchrotron facilities or NMR. Besides, these methods are very slow requiring 1 year of experiments in order to determine the three dimensional structure of a protein. The demanding nature of the experimental methods adds significant expenses to the drug discovery process.

Developing drugs in shorter time using more efficient methods has been the main motivation behind the studies on protein structure prediction. Most of the prediction methods in literature utilize data on proteins identified up to date and try to estimate structure of the molecule to discover its functions. Methods are usually purely computational, enormously reducing time required to predict the structure and cost of the process. Results up to date have been quite satisfactory and many researchers are conducting innovative research to improve prediction accuracy.

Main drive behind prediction of proteins' structures with better accuracy is to have more information and better estimations on their functions that will accelerate an important

phase of drug discovery process enabling faster drug availability at an affordable price for larger segments of population suffering from a particular disease.

1.3 Proteins: Molecules of Life

1.3.1 Introduction

Proteins are large molecules indispensable for existence and proper functioning of biological organisms. Proteins are used in structure of cells, which are main constituents of larger formations like tissues and organs. Bones, muscles, skin and hair of organisms are made basically up of proteins. Besides their necessity for structure, they are also required for proper functioning and regulation of organisms such as enzymes, hormones, antibodies. Enzymes catalyze all biochemical reactions in the cells and systems of organisms; e.g., hemoglobin is the protein binding oxygen in blood to transport to cells for aerobic respiration (respiration using oxygen). Hormones are responsible for regulatory processes like adjustment of glucose level in blood and calcium level in bones and blood, stimulation of growth, expansion and contraction of blood vessels, increases or decreases in heart rate, reabsorption of water from kidneys. Antibodies are built by immune system to defeat invading bacteria, viruses and toxic substances. Understanding functions of proteins is crucial for discovery of drugs to treat various diseases and disorders.

A protein molecule is the chain(s) of amino acids (also called residues). A typical protein (Figure 1.2) contains 200 – 300 amino acids but this may increase up to 27,000 in a single chain.



Figure 1.2: A typical protein molecule

1.3.2 Amino Acids

Amino acids, subunits of proteins, are organic compounds that contain a basic amino (NH_2) group, an acidic carboxyl (COOH) group and a side chain attached to an alpha carbon atom (Figure 1.3). In order to build chains that form protein molecule, amino acids are attached to one another by dehydration synthesis. A water molecule (H_2O) is released with this process while the chemical C-N bond (called “peptide bond”) is formed between carboxyl group of one amino acid and amino group of the other. When consecutive peptide bonds attach one amino acid to another, a long chain is formed, one or more of which make up the ultimate protein molecule. Chemical structure of residue sequence and peptide bond are displayed in Figure 1.4.

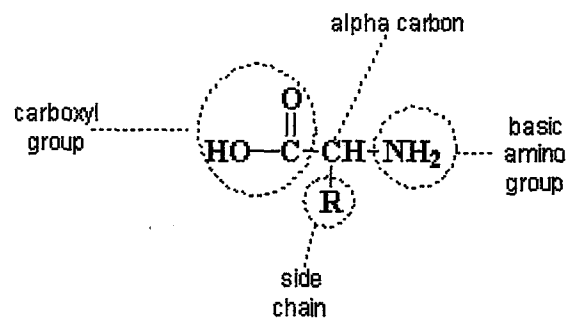


Figure 1.3: Chemical structure of amino acids

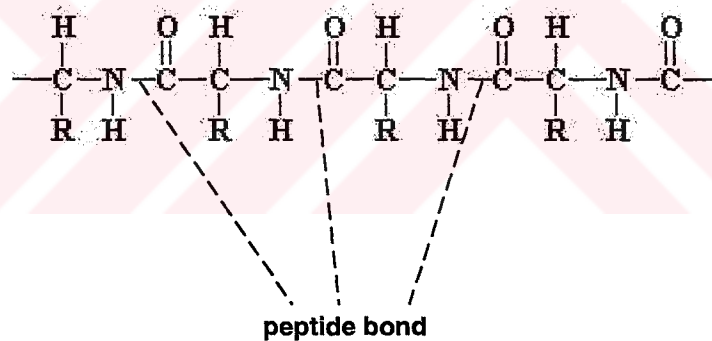


Figure 1.4: Chemical structure of protein molecule and peptide bond

Although carboxyl and basic amino groups of all amino acids are the same, their side chains differ providing diversity. There are 20 types of amino acids in nature and their names, 3-letter representations and single-letter representations are provided in Table 1.1.

Table 1.1: List of amino acids, their three-letter and single-letter representations

Amino Acid	Three-Letter Representation	Single-Letter Representation	Amino Acid	Three-Letter Representation	Single Letter Representation
alanine	ALA	A	leucine	LEU	L
arginine	ARG	R	lysine	LYS	K
asparagine	ASN	N	methionine	MET	M
aspartic acid	ASP	D	phenylalanine	PHE	F
cysteine	CYS	C	proline	PRO	P
glutamic acid	GLU	Q	serine	SER	S
glutamine	GLN	E	threonine	THR	T
glycine	GLY	G	tryptophan	TRP	W
histidine	HIS	H	tyrosine	TYR	Y
isoleucine	ILE	I	valine	VAL	V

Residue content and order in chain(s) is unique for each protein just like specificity of gene sequence. The sequence and types of side chains determine all properties of protein; its three-dimensional shape, and chemical and biological functions. Residue order is determined by the base sequence of nucleotides (building blocks of genes) in the gene coding the protein.

1.3.3 Structures of Proteins

Starting with the sequence of residues in the chain(s) making up protein, there are 4 basic structural phases: primary structure, secondary structure, tertiary structure and quaternary structure.

1.3.3.1 Primary Structure

The primary structure is the sequence of amino acids that the protein molecule consists of (Figure 1.5) like H G Y E P A R T Only the plain chain(s) formed by attachment of amino acids to each other is considered; higher dimensional shapes, arrangement of peptide bonds, angles between chemical bonds in amino acids and interactions between any parts of residues are all ignored. The amino acid content and order dictates the shape of the protein molecule, its spatial and biochemical properties.

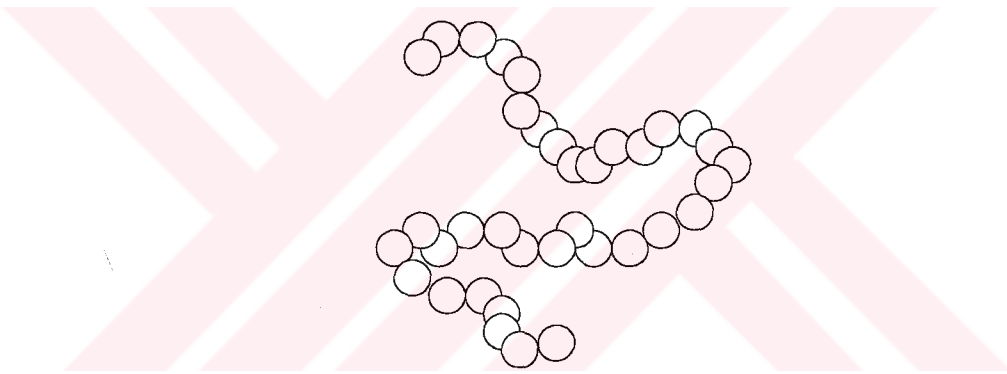


Figure 1.5: Primary structure of proteins

1.3.3.2 Secondary Structure

The secondary structure (folding type) of a segment of polypeptide chain is the local spatial arrangement of its main-chain atoms without regard to the conformation of its side chains or to its relationship with other segments (IUPAC-IUB, 1970). This is the shape formed by amino acid sequences due to interactions between different parts of molecules.

1.3.3.2.1 Types of Secondary Structures

There are mainly three types of secondary structural shapes: α -helices, β -sheets and other structures connecting these such as loops, turns or coils.

1.3.3.2.1.1 α -helices

Alpha-helices are spiral strings (Figure 1.6) formed by hydrogen bonds between CO and NH groups in residues (Brandon and Tooze, 1998). i^{th} amino acid in an α - helix bonds to $i+4^{\text{th}}$ amino acid. The repeat length of an α - helix is 3.6 residues and the rise is 5.6 Å per turn. Residues $i+3$ and $i+4$ are the nearest ones to residue i (Jarey and Hanley). Alpha-helices are very stable and common structures, accounting for 32-38% of all residues in a typical protein (Kabsch and Sander, 1983). Minimum number of residues to form an α - helix is 4 and the average length of an α - helix is 10 residues.



Figure 1.6: α -helix

1.3.3.2.1.2 β -sheets

Beta-sheets are plain strands (Figure 1.7) formed by stretched polypeptide backbone. When β -sheets come together, hydrogen bonds form between C=O and NH groups of residues of adjacent chains, keeping them together. This makes them leading either to their parallel arrangement in the same direction or antiparallel arrangement in opposite directions. The average length of a β -sheet is 6 residues where min number of residues is 2. Typically 20-28% of residues of a protein are placed in β -sheets (Kabsch and Sander, 1983).



Figure 1.7: β -sheet

1.3.3.2.1.3 Connecting Structures (Loops, turns, coils, etc.)

Connecting structures do not have regular shapes; they connect α -helices and β -sheets to each other (Figure 1.8). Turns enable parts of polypeptide chain to fold onto itself reversing the direction of the polypeptide chain to form its three-dimensional shape. Minimum number of residues required to form a connecting structure is 1. Loops

approximately contain 21% of residues with general length of 6 to 16 residues in an average protein where turns hold nearly 33%.

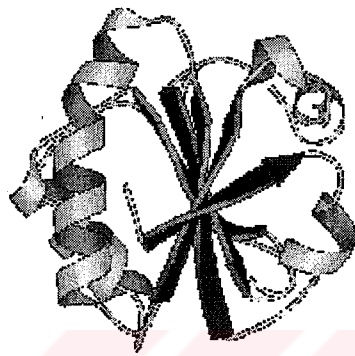


Figure 1.8: Connectors

1.3.3.2.2 Classification of Proteins According to Their Secondary Structures

Proteins are classified according to their secondary structure content, considering α -helices and β -sheets. Levitt and Chothia (1976) were the first to propose such a classification with four basic types (Mount, 2001). “All-alpha” proteins consist almost entirely (at least 90%) of α -helices. “All-beta” are the ones composed mostly of β -sheets (at least 90%) in their secondary structures. There are two intermediate classes which have mixed α -helices and β -sheets. “Alpha/beta” proteins have approximately alternating, mainly parallel segments of α -helices and β -sheets. The last class, “alpha+beta” has mixture of all alpha and all beta regions, mostly in an antiparallel fashion.

1.3.3.3 Tertiary Structure

Tertiary structure corresponds to the shape of a protein molecule in space formed by bends and folds of secondary structure of polypeptide chain (Brandon and Tooze, 1998). α -helices and β -sheets in secondary structure stabilize these regions of the chain and positions of them relative to each other. Angles among chemical bonds and interactions shape the most stable arrangement of the molecule in space. Structure is maintained by strong and weak chemical bonds called ionic, hydrogen, disulfide, Van der Waals bonds, and hydrophobic interactions. All repulsive and attractive forces are balanced at the stable arrangement of the polypeptide chain(s).

1.3.3.4 Quaternary Structure

Many proteins are very large molecules consisting of more than one polypeptide chain. Hydrophobic interactions, hydrogen bonds, and ionic bonds keep chains together either symmetrically or not. Precise arrangement of tertiary structures of chains is called “Quaternary Structure” (Brandon and Tooze, 1998). Hemoglobin has four chains that associate in quaternary structure to build up the final protein molecule.

1.4 Determination of Structures of Proteins

1.4.1 X-Ray crystallography

Determination of protein structure is possible using experimental methods. X-Ray crystallography technique was first utilized by John Kendrew and Max Perutz to build atomic model of myoglobin showing secondary and tertiary structures in 1959 (Brandon

and Tooze, 1998). Main principle of the method is placing proteins in crystal form in front of an X-Ray source and reflecting lights diffracted by atoms onto a film. Proteins are crystallized in special conditions showing diversity from molecule to molecule. It may sometimes be a challenging task to discover the best-crystallized form of the polypeptide since the technique depends on trial and error. When crystallized, molecules diffract ordered beams of light since they are packed regularly. Utilizing darkness, brightness, distribution, and pattern of dots marked by light beams on the film, electron density maps are extracted. Electrons revolve around nucleus of atoms and these maps provide information on regions where electron density is high. Therefore, coordinates of atoms are determined in two dimensions using these maps. Hydrogen atoms create a significant problem for this method since they have only one electron. Information on three-dimensions is drawn by rotating the molecule continuously during the process. After computerized analysis and interpretation of the collected data, the three-dimensional structure of the molecule is identified.

The X-ray diffraction experiments are carried out in special facilities called “synchrotrons” in general (Mitschler, 2002).

X-Ray crystallography method is rather expensive and very slow requiring continuous work of a synchrotron for months in order to identify three-dimensional structure of a single protein, with an average size.

1.4.2 Nuclear Magnetic Resonance (NMR) Spectroscopy

Another experimental method to determine structures of protein molecules is Nuclear Magnetic Resonance (NMR) spectroscopy. Atoms of a molecule are placed in a magnetic field and energy of interaction of charged nucleus of atoms is utilized for estimation of relative positions of pairs of atoms. These are called constraints and various constraints are

obtained between several pairs of bonded and non-bonded atoms. When enough data is obtained, all possible configurations that obey these constraints are generated. The models are then averaged and adjusted to assure that they satisfy conditions for normal bond lengths and angles. Often the positions of atoms in the different models are averaged, and the average model is then adjusted to obey normal bond distances and angles, which is called "restrained minimization". Most usually, comparisons between the models in the ensemble, with perhaps 10 - 50 members, provide some information on how well the protein conformation was determined by the NMR constraints. But in parts of the molecule where generated models display diversity, the results may be unsatisfactory. On the average, this method is not as accurate and giving detailed information as X-Ray method, but it is quite successful and satisfactory on hydrogen atoms where there is some trouble with the other method at these regions. Advantage of this technique is that crystallization of the molecule is not required.

The result of an NMR study is not as detailed and accurate as that is obtained crystallographically (Creighton, 1993). Regions where there is a large spread between NMR models (often observed in loops or at termini) may signify lack of sufficient information (larger uncertainty), or actual motion or disorder of the molecule. One powerful aspect of NMR is the ability to measure the dynamics of each residue, which helps to distinguish between possibilities. Portions of a chain with high mobility produce telltale signals in NMR. In contrast, by crystallography such portions would produce low and nearly uniform electron density and yield no information. Unlike X-ray crystallography, NMR data provides the relative positions of many (but not all) hydrogen atoms.

1.5. Protein Secondary Structure Prediction

1.5.1 Protein Secondary Structure Prediction Methods

Structures of only 12,500 of proteins were known although 86,000 proteins' residue sequences were identified in year 2001 (Mount, 2001). All structure prediction methods basically rely on the idea that there is correlation between residue sequence and structure. Most methods to predict protein structure from residue sequence utilize information on known protein structures. Databases are formed and examined for relationships between amino acid sequence and protein structure. First predictions were made in 1970s with a few dozen structures available. Currently structures of about 22,000 (as of October 2003) proteins are identified that means vast amount of data supporting more reliable predictions with better accuracy is available. Protein structures are stored in and accessible from Protein Data Bank; <http://www.pdb.org>.

In 30-years-time since protein structure prediction studies started, there have been various methods developed; the most successful ones include neural network models, database search tools, multiple sequence alignment, local sequence alignment, threading, hidden Markov model-based prediction, nearest neighbor methods, molecular dynamic simulation, and approaches combining different prediction methods.

Neural Network Models draw its inspiration from neurons in brain. Neural networks are parallel, distributed information processing structures and the method tries to solve the problem by training the network (Bohr *et al.*, 1990; Cai *et al.*, 2002; Rost, 2001). The most successful ones are Copenhagen (Petersen *et al.*, 2000) (PSI-BLAST [Altschul *et al.*, 1997] is utilized and up to 800 networks are averaged [Rost, 2001]), PROF, PSIPRED, SAM-T99sec, PHD (builds second-level structure to structure network combining different neural

networks [Rost, 2001; Rost and Sander, 1993]) and SSpro (uses advanced recursive neural network system with no correlation among training data [Baldi *et al.*, 1999]).

The most successful database search tool is stated (Rost, 2001) to be PSI- BLAST (Altschul *et al.*, 1997), considering its speed and accuracy.

The multiple sequence alignment method aligns each sequence such that one base in a sequence corresponds to bases in the other sequences to reveal the similarity of genetic code, evolutionary history, and common biological functions of the strings. Consensus is one of the latest approaches utilizing this method with significant performance (Cuff and Barton, 1999).

The local sequence alignment approach utilizes local pair-wise alignment of the sequence to be predicted and the most significant method developed with this approach is named PREDATOR (Frishman and Argos, 1997).

Threading maps the unknown structure to the most similar known sequence (Thiele *et al.*, 1999).

~~Hidden Markov Model-Based Prediction of Secondary Structure (HMMSTR)~~ considers similarity of unknown protein to segments of known structures (Bystroff *et al.*, 2000).

The nearest neighbor methods operate by matching segments of the sequence with segments within a database of known structures, and making a prediction based on the observed secondary structures of the best matches (nearest neighbors) (Sen, 2003; Westhead and Thornton, 1998; Yi and Lander, 1993). NSSP is a successful method utilizing this approach (Salamov and Soloveyev, 1997).

There are two significant approaches: combination of GOR Algorithm and Multiple Sequence Alignment Method and combination of Nearest-Neighbor Algorithms and Multiple Sequence Alignment Method.

The combination of GOR algorithm and multiple sequence alignment method (Kloczkowski *et al.*, 2002) starts with selection of a set of proteins (12 proteins) with well-

determined structures, none of which belonging to or having identity to any of proteins in databank of GOR program. Next, multiple sequence alignment of these proteins is carried out and the results are the inputs for GOR algorithm.

A scoring system considering sequence-similarity matrix, local structural environment scoring scheme and N and C-terminal positions of secondary structure types is utilized with a restricted database of a small subset of proteins that are similar in the combination of nearest-neighbor algorithms and multiple sequence alignment method. Nearest-neighbor algorithmic part is followed by multiple sequence alignments (Salamov and Soloveyev, 1995).

1.5.2 Performances of Prediction Methods

Most recent review of prediction methods in literature covers the ones up to year 2000 with evaluations of the techniques and their success (Rost, 2001). In this review, best methods predicting secondary structure for globular proteins in general (not considering a special class like membrane proteins) are considered, tested on 195 proteins and three state per residue performances (helix, sheet or other), Q_3 are measured by experiments of Rost. Accuracies claimed in original publications of these methods and the ones calculated after Rost's experiments are available in Table 1.2.

Copenhagen (Petersen *et al.*, 2000) appears to have the best prediction with three-state accuracy of 78% and PROF is the follower with 77%.

Various prediction methods with successful accuracies were tested utilizing CB396 database (Cuff and Barton, 1999) to try a newly proposed accuracy index (Zhang and Zhang, 2003). CB396 database contains 396 non-homologous proteins and secondary structures of these proteins have been predicted with 6 different algorithms. Table 1.3 shows the three-state accuracies (Q_3) of these algorithms on the CB396 database.

Table 1.2: Accuracies of Secondary Structure Prediction Methods (Rost, 2001)

Method	Q ₃ (%)	Q ₃ Claim (%)
PROF	77.0	
PSIPRED	76.6	76.5-78.3
SSpro	76.3	76
JPred2	75.2	76.4
PHDpsi	75.1	
PHD	71.9	71.6
Copenhagen	78	77.8

Table 1.3: Average three-state accuracy indices calculated for six prediction algorithms based on 396 proteins (%) (Cuff and Barton, 1999)

Method	Q ₃ (%)
PHD	71.950
NNSSP	71.400
DSC (King and Sternberg, 1996)	68.413
PREDATOR	68.602
ZPRED (Zvelebil, Barton, Taylor and Sternberg, 1987)	59.637
Consensus	72.707

After submission of review of Rost, intense studies on the area continued with various publications. Baello and coworkers proposed enhanced accuracy with hydrogen exchange fourier transform infrared spectroscopy (Baello *et al.*, 2000). The method studied only 23 proteins, 19 of which were treated as known and three-state accuracy of the method is not available.

Another strong method was support vector machine approach with the published Q_3 per-residue accuracy of 73.5% (Hua and Sun, 2001).

Severcan and coworkers worked on another neural network study, but Q_3 accuracy of the method was not published (Severcan *et al.*, 2001).

Andersen, Bohr and Brunak developed a method quite different from what had been proposed up to then: they predicted hydrogen bonds that stabilized secondary structure of the molecule (Andersen *et al.*, 2001). Q_3 indices for the technique are displayed and compared with other methods but the performance is below Copenhagen and PROF.

Kloczkowski and coworkers utilized GOR algorithm and multiple sequence alignment to make the basic predictions on 12 different proteins with 71.9% accuracy (Kloczkowski *et al.*, 2002). After short structures separating longer ones were corrected by hand and minimum difference values were set, total accuracy level climbed up to 77.0%.

Kilosanidze and coworkers' study of incorporating molecular mechanics for secondary structure prediction achieved impressive results with physical grounds of α -helix formation without utilizing any statistical data (Kilosanidze *et al.*, 2002). Since method mainly deals with α -helices, its accuracy is supplied in terms of percentage of residues correctly predicted as α -helix; $Q_{3\alpha}$ is 79%, which is a rather impressive result. But this method is only tested on 8 randomly selected proteins with resolutions less than 2.5 Å, no disulfide bridges, no ligands or ions interacting with α -helical segments. Although it achieves significant results among the methods using specially defined set of protein data, it should be tested on much more proteins to have more reliable accuracy values.

Artificial neural network model by Cai and coworkers had a broader spectrum considering parallel and anti-parallel β -sheet strands, β -bridges, and connecting structures separately as turns, bends and coils (Cai *et al.*, 2002). The method was tested with database of 52 proteins of Liu and Chou (Liu and Chou, 1999), which consists of independent proteins with no more than 52% of homology with the others or those in the training set. Competencies of predictions are supplied in terms of errors rather than per-residue accuracy of Q_3 .

One of the most recent applications of nearest neighbor method predicted α -helices and β -sheets on 100 protein sequences with reported average accuracy of 59.6% with the pure results of the algorithm (Sen, 2003). But after replacing odd assignments to follow the local trend around the residue, overall precision improved to 63.3% with maximum achieved among 100 sequences being 79%.

Prediction of α -helices and connecting structures via continuous wavelet transform was another original approach, but predictions could only propose relative positions and numbers of structures, not their length or specific locations in the sequence (Qui *et al.*, 2003).

Although secondary structure prediction of proteins is a new subject, there are quite significant contributions from scientists in a large variety of disciplines such as biology, chemistry, informatics, computer science, applied mathematics. The algorithms and methods developed and tested propose successful results with rather satisfactory prediction accuracy indices up to 75%.

1.5.3 The State-of-the-art

The Copenhagen method (Petersen *et al.*, 2000) seems to be the state-of-the-art with 78% three-state per residue accuracy. Copenhagen utilizes divergent PSI-BLAST to train

and test neural networks. Up to 800 networks are averaged and standard three structures (helix, strand and others) are replaced with nine units coding for the three secondary structure states of three adjacent residues. Prediction accuracy in the publication is estimated to be higher than 77% (Bystroff *et al.*, 2000).

Next follower after Copenhagen is PROF with 77%. These levels are for general prediction with utilization of sequences of random proteins. Better accuracies have been achieved with non-random protein selections. Solubility (Jureti *et al.*, 1995), proline content (Cid *et al.*, 1982), hydrophobicity (Cid *et al.*, 1982), selection of membrane-proteins (dominated by α -helices) or correlated sets of data can improve results significantly.

1.5.4 Approach of This Study

A two-stage algorithm for the secondary structure prediction of proteins has been developed. The algorithm has a probabilistic approach utilizing data on all structurally identified proteins having the same folding type with the new unknown protein. The first stage in the method is determination of class of unknown protein. This is accomplished by solving a mixed integer linear program (MILP) problem with 100% accuracy.

Next phase is searching sequential groups of 3 to 7 residues of unknown protein's amino acid chain, in the database of structures of known proteins having the same folding type with the unknown one. The folding types of proteins are given in SCOP (Structural Classification Of Proteins) database (<http://scop.mrc-lmb.cam.ac.uk>). A local database is formed for each folding type by extracting structural data from PDB files.

After matrix of frequency of occurrences of each sequential group of new residue chain is generated, probabilities of being in an α -helix, a β -sheet or a connecting structure are

calculated for each residue. The structure with maximum probability becomes the prediction of the method for that residue.

Following sections of the report continue with establishment of method for prediction of structural class of proteins, detailed explanation of two-stage algorithm for prediction of secondary structures of proteins, results obtained with the study, conclusions and recommendations for future work.



Chapter 2

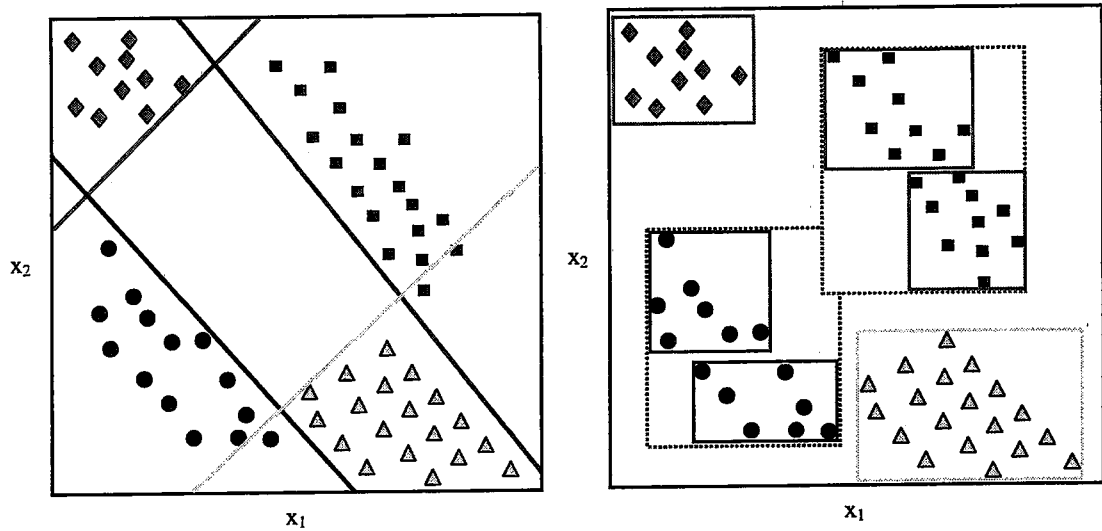
FOLDING TYPE PREDICTION

2.1 Introduction

The two-stage algorithm decomposes the secondary structure prediction problem into two steps: first the overall folding type of the protein is predicted, and then the secondary structure is predicted using the refined statistical data from the first step.

The overall folding type of a protein depends on amino acid composition (Nakashima *et. al.*, 1986). There have been several methods proposed to exploit this theory for predicting folding type of a protein (Chou, 1995; Bahar *et. al.*, 1997; Cai *et. al.*, 2001). These methods conduct a statistical analysis and separate multi-dimensional amino acid composition data into several folding types.

Classification of multi-dimensional data plays an important role in the decision determining main characteristics of a set. Traditional approaches to data classification use separating hyper planes as shown in Figure 2.1a. Although these methods can be efficient in classifying data into two sets, they are inaccurate and inefficient when the data needs to be classified into more than two sets. Mixed-integer programming allows the use of boxes for defining boundaries of the sets that include all or some of the points in that set as shown in Figure 2.1b. Therefore, the efficiency and accuracy of data classification can be improved dramatically compared to traditional methods. This chapter consists of details of the mixed-integer programming model and the results for predicting the folding type of the proteins given their amino acid composition.



(a) Data classification using hyper planes.

(b) Data classification using boxes.

Figure 2.1. Schematic representation of classification of data.

2.2 Mixed-Integer Programming Formulation

The mixed-integer programming model for determining the folding type of a protein is given below:

Indices

i	protein
j	subunit of the protein
k	the folding type of the protein (α , β , $\alpha+\beta$, α/β)
l	box that encloses a number of data points belonging to a type ($l=1,\dots,L$)
m	amino acid
n	bound ($n=\text{lower, upper}$)

Parameters

$\text{comp}_{i,j,m}$	composition of amino acid m in the subunit j of protein i
$\text{foldtype}_{i,j,k}$	folding type k of the subunit j of protein i
comp^U	a sufficiently large parameter (1 for this problem)

Variables

YB_l	binary variable to model the existence of box l
$YBC_{l,k}$	binary variable to model the assignment of folding type k to box l
$YPB_{i,j,l}$	binary variable to model the assignment of subunit j of protein i to box l
$YPC_{i,j,k}$	binary variable to model the assignment of subunit j of protein i to class k
$X_{l,m,n}$	continuous variable to define the bounds n for amino acid m in box l
$XD_{l,k,m,n}$	continuous variable to define bounds n for amino acid m in box l for class k
$XP1_{i,j,k}$	continuous variable to model misallocation of subunit j of protein i to class k
$XP2_{i,j,k}$	continuous variable to model misallocation of subunit j of protein i to class k
Z	objective function variable

$$\min Z = \sum_i \sum_j \sum_k (XP1_{ijk} + XP2_{ijk}) \quad (2.1)$$

subject to

$$XD_{lkmn} \leq \text{comp}_{ijm} * YPB_{ijl} \quad \forall ijklmn | n=\text{lower} \quad (2.2)$$

$$XD_{lkmn} \geq \text{comp}_{ijm} * YPB_{ijl} \quad \forall ijklmn | n=\text{upper} \quad (2.3)$$

$$XD_{lkmn} \leq \text{comp}^U * YBC_{lk} \quad \forall klmn \quad (2.4)$$

$$\sum_l YPB_{ijl} = 1 \quad \forall ij \quad (2.5)$$

$$\sum_k YPC_{ijk} = 1 \quad \forall ij \quad (2.6)$$

$$\sum_l YPB_{ijl} = \sum_k YPC_{ijk} \quad \forall ij \quad (2.7)$$

$$\sum_k YBC_{lk} \leq YB_l \quad \forall l \quad (2.8)$$

$$\sum_l YBC_{lk} \geq 1 \quad \forall k \quad (2.9)$$

$$YBC_{lk} - \sum_i \sum_j YPB_{ijl} \leq 0 \quad \forall lk \quad (2.10)$$

$$YBC_{lk} - \sum_i \sum_j YPC_{ijk} \leq 0 \quad \forall lk \quad (2.11)$$

$$XP1_{ijk} - YPC_{ijk} \leq 0 \quad \forall ijk \quad (2.12)$$

$$XP2_{ijk} + YPC_{ijk} \geq 1 \quad \forall ijk \quad (2.13)$$

$$YB_p, YBC_{i,k}, YPB_{i,j,p}, YPC_{i,j,k} \in \{0,1\} \quad (2.14)$$

$$X_{l,m,p}, XD_{i,k,m,p}, XP1_{i,j,k}, XP2_{i,j,k} \geq 0 \quad (2.15)$$

Chapter 3

SECONDARY STRUCTURE PREDICTION

3.1 Introduction

Discovering functions of proteins is a very important problem in understanding biological processes in organisms and developing drugs for uncured diseases. Significance of this problem justifies intensive research on structure prediction of proteins to estimate their functions. This issue, its urgency, probable consequences, major methods developed up to date, and their accuracies are all discussed in the first chapter of this thesis.

This thesis proposes a new approach to secondary structure prediction using a two-stage method with enhanced accuracy. When the sequence of a new protein is received, its folding type is determined (is discussed in detail in the method to predict the folding type in chapter 2) and segments of its residue chain are searched in database of structures of all experimentally identified proteins belonging to the same class. This comparative search is conducted in four different tables corresponding to four folding types in our local database. Flowchart of the proposed method is displayed in Figure 3.1. This chapter discusses major steps in database formation, a brief review of structural classification method, segment search algorithm's structure and code, and finally statistical secondary structure prediction approach with a relevant illustrative example.

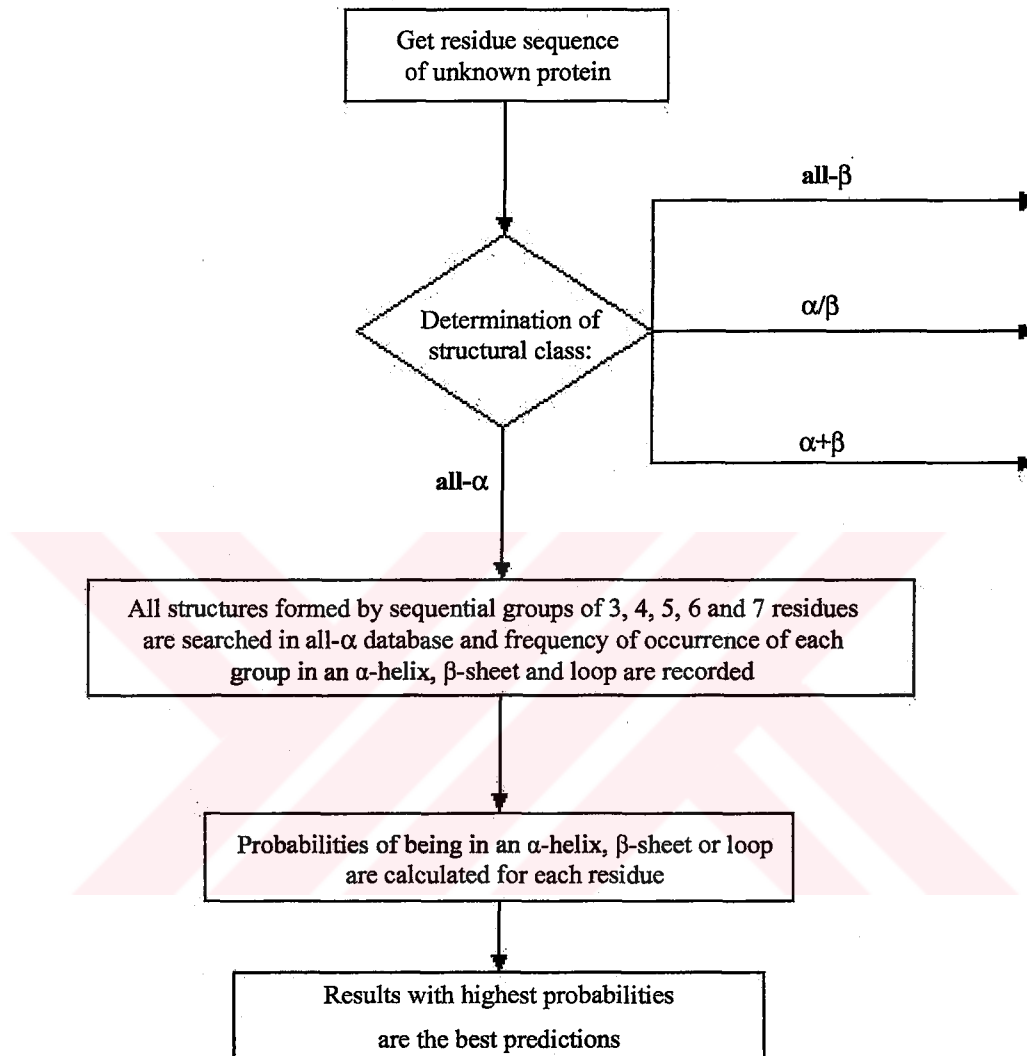


Figure 3.1: Flowchart of secondary structure prediction algorithm

3.2 Data Analysis and Formation of Databases

SCOP (Structural Classification of Proteins) is the database of protein names and chain letters belonging to each structural class (all- α , all- β , α/β or $\alpha+\beta$). Lists of all proteins experimentally identified up to date in each class are taken from SCOP; <http://scop.mrc-lmb.cam.ac.uk>. Names of all- α , all- β , α/β and $\alpha+\beta$ proteins are available in Appendices A, B, C and D respectively.

Experimental result files of proteins in each list are downloaded from PDB. The result files contain necessary information to identify structures and their residue sequences, but it should be processed for efficient use in a database. This is overcome by a program coded in C Programming Language. This code reads necessary data from PDB file and returns information to be loaded to the local database. Lines starting with the phrase “SEQRES” contain residue sequence for each chain in the protein. A sample is available in Figure 3.2 and the explanation for reading sequence from PDB file is supplied below the figure.

When there is the word “HELIX” or “SHEET” at the beginning of a line, information on helical structures and sheets are supplied in the rest of the line, respectively. Samples of these parts of PDB files and relevant explanations are provided in Figure 3.3.a and Figure 3.3.b.

The program loads residue sequences for each chain by reading lines starting with “SEQRES” and marks the ones that are part of an α -helix or a β -sheet by reading helical and sheet regions from lines starting with the words “HELIX” and “SHEET”, in PDB files. Remaining residues in sequence of each chain represent connecting regions like turns or loops.

Column #:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
SEQRES	1	A	330	MET	LYS	THR	ALA	TYR	ILE	ALA	LYS	GLN	ARG	GLN	ILE	SER	
SEQRES	2	A	330	PHE	VAL	LYS	SER	HIS	PHE	SER	ARG	GLN	LEU	GLU	GLU	ARG	
SEQRES	3	A	330	LEU	GLY	LEU	ILE	GLU	VAL	GLN	ALA	PRO	ILE	LEU	SER	ARG	
SEQRES	4	A	330	VAL	GLY	ASP	GLY	THR	GLN	ASP	ASN	LEU	SER	GLY	ALA	GLU	
SEQRES	5	A	330	LYS	ALA	VAL	GLN	VAL	LYS	VAL	LYS	ALA	LEU	PRO	ASP	ALA	
SEQRES	6	A	330	GLN	PHE	GLU	VAL	VAL	HIS	SER	LEU	ALA	LYS	TRP	LYS	ARG	
SEQRES	7	A	330	GLN	THR	LEU	GLY	GLN	HIS	ASP	PHE	SER	ALA	GLY	GLU	GLY	
SEQRES	8	A	330	LEU	TYR	THR	HIS	MET	LYS	ALA	LEU	ARG	PRO	ASP	GLU	ASP	
SEQRES	9	A	330	ARG	LEU	SER	PRO	LEU	HIS	SER	VAL	TYR	VAL	ASP	GLN	TRP	
SEQRES	10	A	330	ASP	TRP	GLU	ARG	VAL	MET	GLY	ASP	GLY	GLU	ARG	GLN	PHE	

Column #	Explanation
1	Lines starting with the phrase "SEQRES" provide information on residue sequences of chains
2	Line number
3	Chain identification letter (residues belonging to this chain are supplied in this line)
4	Total number of residues in the relevant chain
5 .. 17	3 letter representation of residues in the chain (each line contains at most 13 residues)

Chain A in protein 11AS has totally 330 residues and its residue sequence is methionine, lysine, threonine, alanine, tyrosine, isoleucine, alanine, lysine, glutamic acid, arginine, glutamic acid, isoleucine, serine, phenylalanine, valine, lysine, serine, histidine,

Figure 3.2: Sample format for residue sequence part of PDB file of protein 11AS

Column #	1	2	3	4	5	6	7	8	9	10	11
HELIX	1	1	TYR	A	5	LEU	A	27	1		23
HELIX	2	2	ALA	A	74	GLN	A	83	1		10
HELIX	3	3	PHE	A	130	GLU	A	154	1		25
HELIX	4	4	SER	A	170	ARG	A	176	1		7
HELIX	5	5	ALA	A	182	LEU	A	193	1		12

Column #	Explanation
1	Lines starting with the phrase "HELIX" provide information on residue sequences of α -helices
2	α -helix number
3	α -helix number
4	Starting residue of α -helix (represented in three letters)
5	Chain identification letter (Starting residue of α -helix belongs to this chain)
6	Position of starting residue of α -helix in residue chain of the protein
7	Ending residue of α -helix (represented in three letters)
8	Chain identification letter (Ending residue of α -helix belongs to this chain)
9	Position of ending residue in residue chain of the protein
10	Information in this column is irrelevant to this study
11	Total number of residues in the α -helix described in this line

- First α -helix in protein 11AS starts from 5th residue in amino acid sequence of chain A which is tyrosine, 23 residues long and ends at 27th residue which is leucine.
- Second α -helix in protein 11AS starts from 74th residue in amino acid sequence of chain A which is alanine, 10 residues long and ends at 83rd residue which is glycine.

Figure 3.3.a: Sample format for part of PDB file explaining α -helical regions for protein 11AS

Column #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
	SHEET	1	A	8	LEU	A	29	GLU	A	31	0								
	SHEET	2	A	8	GLY	A	91	LEU	A	99	1	N	GLY	A	91	O	ILE	A	30
	SHEET	3	A	8	TYR	A	113	VAL	A	122	-1	N	GLU	A	120	O	LEU	A	92
	SHEET	4	A	8	THR	A	290	GLY	A	296	-1	N	ILE	A	295	O	TRP	A	117
	SHEET	5	A	8	ASP	A	245	ILE	A	254	-1	N	GLY	A	253	O	GLY	A	292
	SHEET	6	A	8	ASN	A	233	ASN	A	240	-1	N	ASN	A	240	O	ASP	A	245

Column #	Explanation
1	Lines starting with the phrase "SHEET" provide information on residue sequences of β -sheets
2	β -sheet number
3	Information in this column is irrelevant to this study
4	Information in this column is irrelevant to this study
5	Starting residue of β -sheet (represented in three letters)
6	Chain identification letter (Starting residue of β -sheet belongs to this chain)
7	Position of starting residue of β -sheet in residue chain of the protein
8	Ending residue of β -sheet (represented in three letters)
9	Chain identification letter (Ending residue of β -sheet belongs to this chain)
10	Position of ending residue of β -sheet in residue chain of the protein
11...19	Information in these columns are irrelevant to this study
- First β-sheet in protein 11AS starts from 29th residue in amino acid sequence of chain A which is leucine and ends at 31st residue which is glutamine. - Second β-sheet in protein 11AS starts from 91st residue in amino acid sequence of chain A which is glycine and ends at 99th residue which is leucine.	

Figure 3.3.b: Sample format for part of PDB file explaining β -sheet regions for protein 11AS

After reading necessary information from the PDB file, the program returns all structures in the protein in a text file in a format for easy transfer to out local database. A sample part of output file of protein 11AS and relevant explanations are provided in Figure 3.4 as an example.

Names of PDB files have the format: “pdbprotein_id.ent” like ”pdb11AS.ent”. Order of execution in the program is seen in the flow diagram in Figure 3.5.

The input of the code is list of names of proteins to be processed by program, in a text file containing a single name with 4 letters in each line. The outputs are: A text file named “protein_id_out.txt” (like “11AS_out.txt”) for each protein in the list whose structure is successfully read and rearranged, a report file named "report.txt" giving information on success and failure of process for each protein in the list, a file named "inconsistent.txt" listing names of proteins with some disorder, complication or inconsistency in their experimental result files and a file named "chain_problem.txt" listing names of proteins with some formatting problems in their chain letters or other fields in their experimental result files.

Column #	1	2	3	4	5	6	7	8
	11AS	A	HELIX	1	5	27	23	YIAKQRQISFVKSHFSRQLEERL
	11AS	A	HELIX	2	74	83	10	AKWKRQTLGQ
	11AS	A	HELIX	3	130	154	25	FSTLKSTVEAIWAGIKATEAAVSEE
	11AS	A	HELIX	4	170	176	7	SQELLSR
	11AS	A	HELIX	5	182	193	12	AKGRERAIKDL
	11AS	A	HELIX	6	258	268	11	ADTLKHQLALT
	11AS	A	HELIX	7	271	275	5	EDRLE
	11AS	A	HELIX	8	277	283	7	EWHQALL

Column #	Explanation
1	Protein name (ID)
2	Chain identification letter (Starting residue structure belongs to this chain)
3	Name of the structure being described
4	Structure number
5	Position of starting residue of structure in residue chain of the protein
6	Position of ending residue of structure in residue chain of the protein
7	Total number of residues in the structure
8	Residue sequence of the structure

Figure 3.4: Part of output file of the program showing structural information of protein 11AS, to be loaded to database

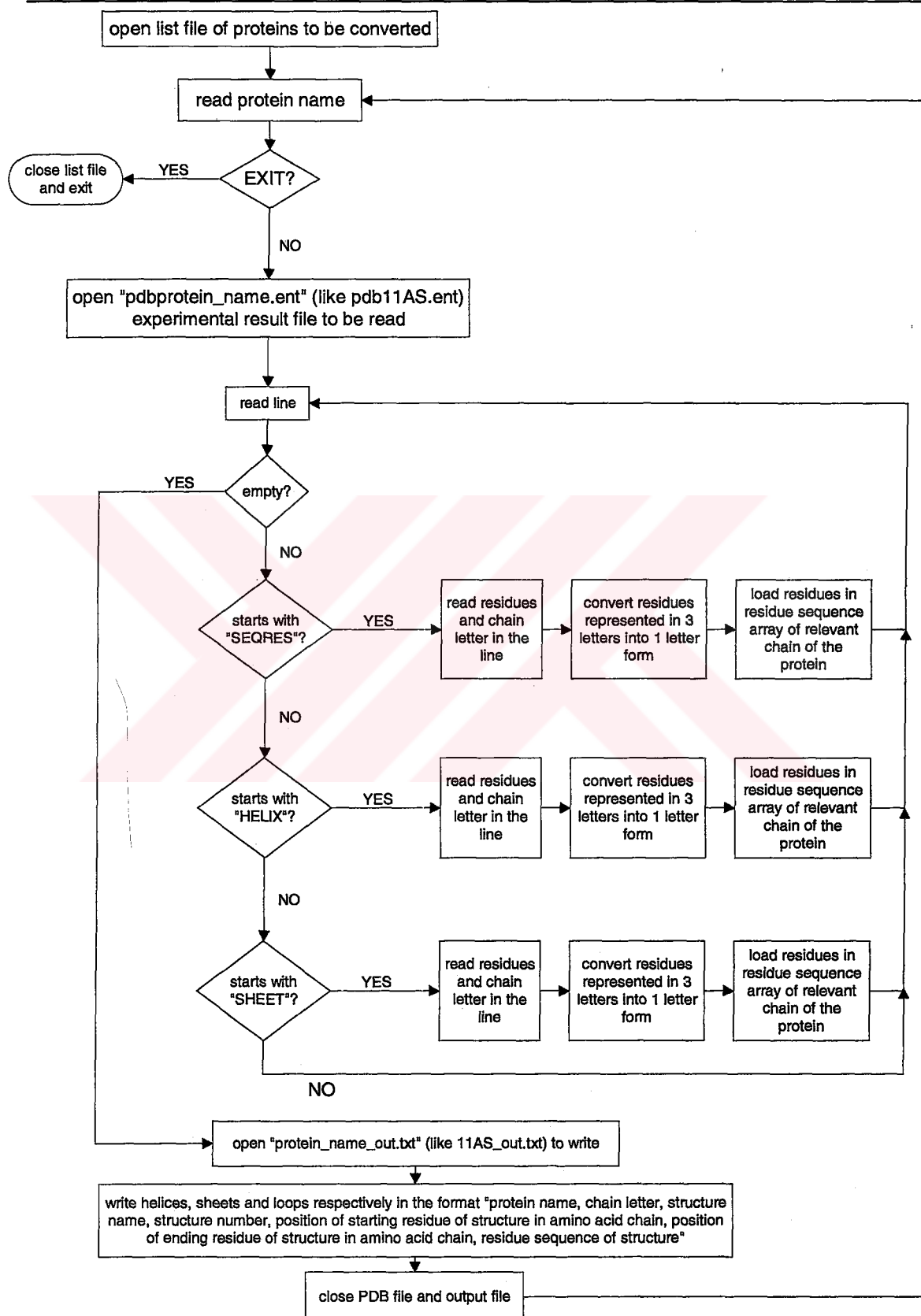


Figure 3.5: Flow diagram showing general execution steps of CONVERTER.C

When program “converter.c” was coded, the general format published by PDB for result files had been considered. But quite many files remained to be unprocessed after execution of program. There are several reasons behind this, most obvious ones being formatting differences and some disorders or complications leading to inconsistencies in files. There are some precautions in our program to clean out such files and to assure data reliability. First of all, if character (in the experimental result file) at the place of chain letter is not alphanumeric or is not space, the structure can not be read and name of the file is recorded in the report as chain letter problem, also recorded in chain_letter.txt and the code skips to process next protein finally with no more processing on this one. List of proteins in chain_letter.txt is available in Appendix E.

As an illustrative example for such a situation, a sample from experimental result file of protein 1CD3 is given in Figure 3.6. Chain letters are written in 3rd column of lines starting with the phrase “SEQRES”, but first a number, then a letter is written in this file violating data standardization. This protein is not processed when the problem is detected by the program; error is recorded in report file and name of the protein is written in chain_problem.txt file.

```

SEQRES   9 4   152  GLU ASN ILE ILE ASN GLY VAL GLU ARG PRO VAL LYS ALA
SEQRES  10 4   152  ALA GLU LEU PHE ALA PHE THR LEU ARG VAL ARG ALA GLY
SEQRES  11 4   152  ASN THR ASP VAL LEU THR ASP ALA GLU GLU ASN VAL ARG
SEQRES  12 4   152  GLN LYS LEU ARG ALA GLU GLY VAL MET
SEQRES   1 F  426  SER ASN ILE GLN THR GLY ALA GLU ARG MET PRO HIS ASP
SEQRES   2 F  426  LEU SER HIS LEU GLY PHE LEU ALA GLY GLN ILE GLY ARG
SEQRES   3 F  426  LEU ILE THR ILE SER THR THR PRO VAL ILE ALA GLY ASP
SEQRES   4 F  426  SER PHE GLU MET ASP ALA VAL GLY ALA LEU ARG LEU SER
SEQRES   5 F  426  PRO LEU ARG ARG GLY LEU ALA ILE ASP SER THR VAL ASP
SEQRES   6 F  426  ILE PHE THR PHE TYR VAL PRO HIS ARG HIS VAL TYR GLY

```

Figure 3.6: Part of PDB file of protein 1CD3

Another category of problem ending up with unprocessed files is due to some disorders or complications in PDB files leading to inconsistencies for the program. To detect these files, the program checks whether starting position of structure is greater than total number of residues in the chain, or there exists a negative number at the character when its starting position is written, or its finishing position is greater than total number of residues in the chain, or there exists a negative number at the character when its ending position is written. When such complications exist, the PDB file has some disorders or complications that can not be solved and the program skips the processing of the current PDB file and this protein is categorized as inconsistent PDB file; its name is recoded in report file as inconsistent and in inconsistent.txt file. Protein names in this file are available in Appendix F. An example for such an experimental result file is available in Figure 3.7. Third and fourth columns in the lines starting with the phrase "SEQRES" give information on chain letter and number of residues in that chain, respectively. So, number of residues in chain B for the protein 1A1R in Figure 3.7 is 198. Eighth and ninth columns in the lines starting with the phrase "HELIX" give information on chain letter and ending position of the structure in residue sequence. The last line in the figure supplies information on a helical structure in chain B starting at position 198, ending at position 204, although chain B was informed to have 198 residues previously.

Due to the reasons mentioned above, experimental results on all proteins could not be utilized and average success rate was around 70%. Table 3.1 displays the number of proteins in SCOP's list and the number successfully processed for each structural class.

The results of the program are written in text files and transferred to Microsoft Access database management software.

```

SEQRES 12 B 198 TYR LEU LYS GLY SER SER GLY GLY PRO LEU LEU CYS PRO
SEQRES 13 B 198 THR GLY HIS ALA VAL GLY LEU PHE ARG ALA ALA VAL CYS
SEQRES 14 B 198 THR ARG GLY VAL ALA LYS ALA VAL ASP PHE ILE PRO VAL
SEQRES 15 B 198 GLU ASN LEU GLU THR THR MET ARG GLY SER HIS HIS HIS
SEQRES 16 B 198 HIS HIS HIS
SEQRES 1 C 23 LYS LYS GLY SER VAL VAL ILE VAL GLY ARG ILE VAL LEU
SEQRES 2 C 23 SER GLY LYS PRO ALA ILE ILE PRO LYS LYS
SEQRES 1 D 23 LYS LYS GLY SER VAL VAL ILE VAL GLY ARG ILE VAL LEU
SEQRES 2 D 23 SER GLY LYS PRO ALA ILE ILE PRO LYS LYS
HET ZN A 901 1
HET ZN B 902 1
HETNAM ZN ZINC ION
FORMUL 5 ZN 2(ZN1 2+)
FORMUL 6 HOH *130(H2 O1)
HELIX 1 1 TYR A 82 ALA A 85 1 4
HELIX 2 2 ILE A 158 LEU A 161 1 4
HELIX 3 3 VAL A 198 THR A 204 1 7
HELIX 4 4 LEU B 39 THR B 48 1 10
HELIX 5 5 TYR B 82 ALA B 85 1 4
HELIX 6 6 ILE B 158 LEU B 161 1 4
HELIX 7 7 VAL B 198 THR B 204 1 7

```

Figure 3.7: Part of PDB file of protein 1A1R

Table 3.1: Number of proteins in SCOP's list of each structural class and the number of proteins with successfully processed structures through our program

	Number of proteins in SCOP	Number of proteins successfully processed by the program
all- α	3242	2419
all- β	5002	3345
α/β	5113	4082
$\alpha+\beta$	5667	3467

Some proteins in organisms are coded by the same gene and are very similar in terms of residue sequence, structure, and function. These molecules are called “homologous” and can be considered to contain the same information due to the very high level of similarity mentioned. While searching a segment of residue sequence of a protein in the database, occurrence of segment in homologous structures can bias the results. That is, although homologous structures contain the same information, all occurrences of segment in the database are being counted, which is similar to utilizing the same data more than once. So, segment’s frequency of occurrence in the structure type homologous structure belongs increases, and this raises probability of being part of the structure same as that of homologous part’s in a misleading manner for the residues in the segment. Therefore, it is necessary to exclude homologous structures from databases, which is carried out by queries in the database. The query searches database for records whose structure type, starting position, ending position, and residue sequences are the same. Duplicates are removed for those cases and database is left with only one record representing those structures. Four databases are established for this study with 80% of structurally identified proteins from each class (details and purpose is explained in Section 3.4.3.2); number of records in each of these databases before and after cleaning homologous structures are available in Table 3.2.

After sorting out homologous structures, all databases are ready for use by the algorithm that searches occurrences of residue segments. Search algorithm is also coded in C Programming Language as will be explained in next chapter. When the C code is integrated with MS Access, each search operation takes considerably long time, which leads to intolerable runtime for the whole process. This is overcome by carrying all databases into text files and for quick operation, leaving only necessary information in databases that will be needed for the algorithm which is the structure type and residue sequence. Major steps of database establishment process explained are shown in Figure 3.8.

Table 3.2: Number of secondary structures in each structural class database of 80% of data, before and after exclusion of homologous structures via query in the database.

Structural Type	Before exclusion of homologous structures	After exclusion of homologous structures
all- α	144299	41968
all- β	221368	67373
α/β	347150	100255
$\alpha+\beta$	197727	62935

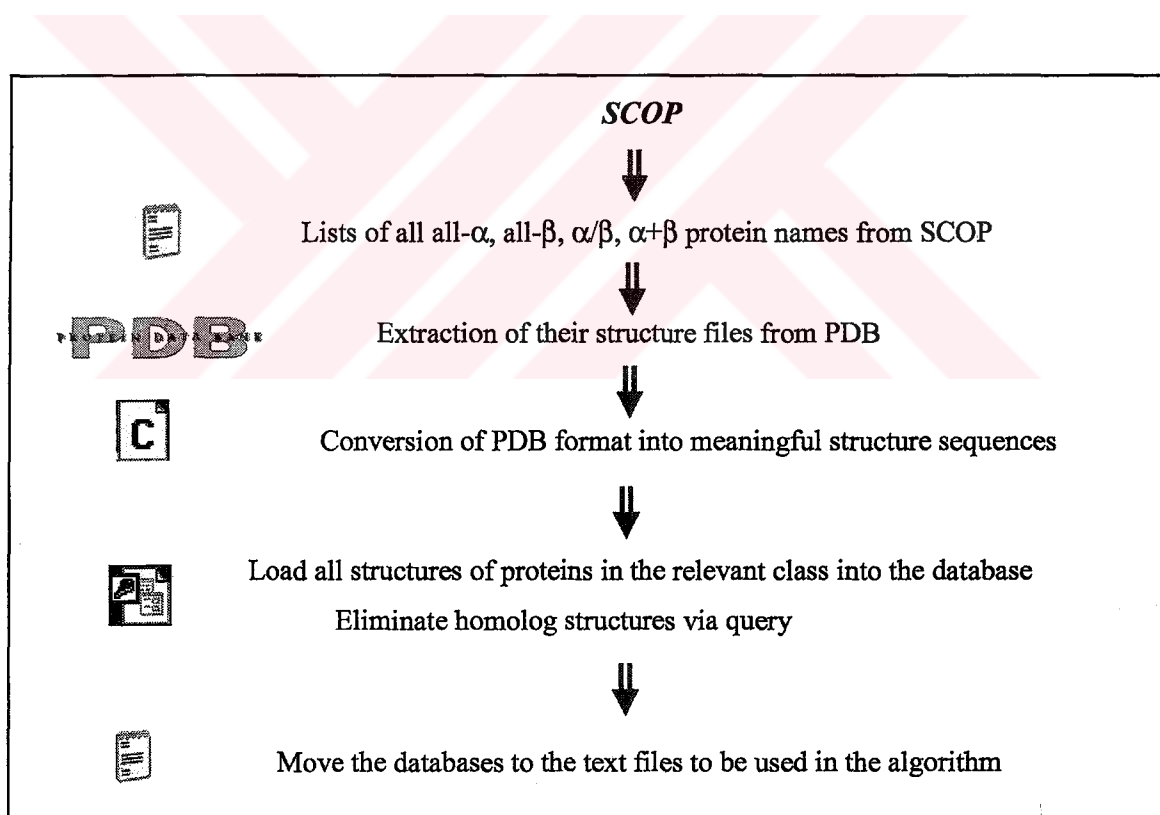


Figure 3.8: Major steps in database establishment

By carrying all databases into text files, database establishment step is completed. The next section summarizes the structural classification method and is followed by the search algorithm to find incidences of residue segments in non-homologous protein structures.

3.3 Folding Type Determination

The overall folding type of a protein depends on its amino acid composition (Nakashima *et. al.*, 1986) and basics of our method for folding type determination rely on utilization of amino acid content of protein and mixed-integer programming model enabling usage of boxes for defining boundaries of data sets. The method achieves 100% accuracy and its details are discussed in Chapter 2 of this thesis.

3.4 Secondary Structure Prediction Algorithm

After determination of structural class of a protein, its secondary structure is predicted. The method mainly consists of two steps: searching new sequence in the database of known proteins belonging to its structural class, and predicting its secondary structure utilizing results of this search. Here, these two major steps will be explained in detail with illustrative examples.

3.4.1 Algorithm for Searching Segments of Sequence in Database

Amino acids are attached to each other through peptide bonds in a protein, as discussed in Chapter 1. There are various forces like ionic, hydrogen, disulfide, Van der Waals bonds, hydrophobic interactions, repulsive, and attractive forces playing the major role for interactions between different parts of proteins ending up with structural arrangement of

sequence. If some regions of sequence are the same or similar, the forces mentioned may form between the same residues which means they will act similarly and shape the structure in a similar manner. This proposition has been the main logic for a major portion of structure prediction methods considering sequence similarity and set of known data as the basis for their forecasts. As discussed in detail in literature review section of introduction part, methods relying on this approach have produced very significant results and accuracy levels. Sequence similarity is at the core of this study, too.

The basis for the algorithm is searching segments of its residue sequence in pool of known protein structures and predicting structure for each residue on the basis of frequency of occurrence. To determine the number of residues in each segment to be searched, two facts are considered: the segment should be long enough to have a legitimate reason to search considering interactions and bonds formed between amino acids to shape their structures. Every structure is searched in the relevant database whose dimensions were stated in previous section.

Starting with 3-residue-groups is sensible because dealing with each residue one by one means trying to find out each residue's frequency of occurrence in an α -helix, a β -sheet, or a loop, ignoring any bonds and interactions between them which is the main factor for shaping structures. Considering 2-residue-groups in search would mean dealing with only a single bond or interaction which does not have a significant contribution itself to the overall structure. Moreover, probability drawn from their frequency of occurrence can mislead the result, because it is known that the more similar the residue sequences, the more association between their structures. Two residues do not reveal a significant similarity among whole sequence of hundreds of residues. Regarding all these factors, examining the databases for occurrences of various 2-residue-sequences and their reliability to give accurate predictions for some structurally known groups, we decided to start with 3-residues-long segments.

As was previously stated, more similarity means more information on the unknown. But when data on known proteins is examined in detail, it is observed that segments longer than 7 residues are not found in any structures other than their own. So, searching segments longer than 7 residues would not provide any significant enhancement in the prediction method.

With the decision of searching 3 to 7-residues-segments of new sequence in database, search operation starts with first 3 residues. After recording frequencies of occurrences of this segment in α -helices, β -sheets and loops, next comes search of segment of 2nd, 3rd and 4th residues, followed by 3rd, 4th, 5th, next 4th, 5th, 6th and so on. After (n-2)nd, (n-1)st and nth residue segment is searched in an n-residues-long chain, the algorithm proceeds to 4-residues-groups starting with 1st, 2nd, 3rd, 4th followed by 2nd, 3rd, 4th, 5th, finally ending with (n-3)rd, (n-2)nd, (n-1)st and nth residue segment. This operation is carried out in this manner up to 7-residues-long segments. A sample search of algorithm is explained on a hypothetical residue sequence in illustrative example in Section 3.4.2.

To sum up, when a new sequence of residues is received, 3 to 7 residues long segments of the chain are searched in the database of proteins of the same class as the new one. The ultimate goal of searching a segment of residues in the database is recording the number of occurrences of each residue in a 3 to 7 residues segment in an α -helix, a β -sheet and a loop. So, for each residue in the sequence, $(7-3+1)*3=15$ records exist:

- frequency of being in a 3-residues segment in an α -helix,
- frequency of being in a 3-residues segment in a β -sheet,
- frequency of being in a 3-residues segment in a loop.
- frequency of being in a 4-residues segment in an α -helix,
- frequency of being in a 4-residues segment in a β -sheet,
- frequency of being in a 4-residues segment in a loop,
- frequency of being in a 5-residues segment in an α -helix,

- frequency of being in a 5-residues segment in a β -sheet,
- frequency of being in a 5-residues segment in a loop,
- frequency of being in a 6-residues segment in an α -helix,
- frequency of being in a 6-residues segment in a β -sheet,
- frequency of being in a 6-residues segment in a loop,
- frequency of being in a 7-residues segment in an α -helix,
- frequency of being in a 7-residues segment in a β -sheet and
- frequency of being in a 7-residues segment in a loop.

Then, table of frequencies of occurrences is printed, a sample of which is available in Table 3.3.

General structure and execution of algorithm is displayed in a flowchart in Figure 3.9 for better demonstration.

This algorithm is coded in C Programming Language; input of program is residue sequence and database in a text file. Output matrix showing frequencies for residues in sequence is written in text file "frequency.txt". This file is used in the prediction stage.

Table 3.3: Table of frequencies of occurrences of residues in chain A of protein 1QUU

[illegible]

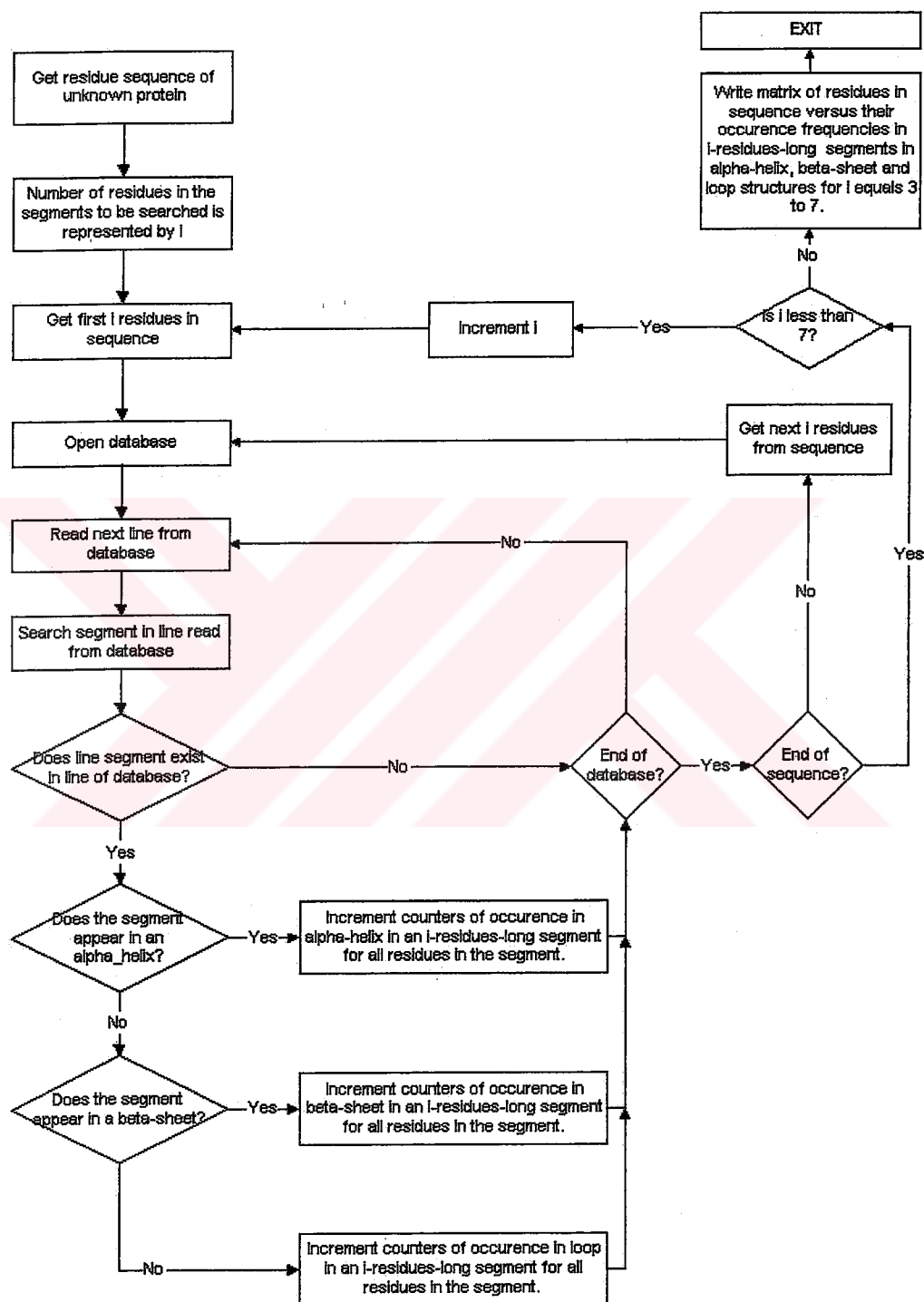


Figure 3.9: Flowchart for residue-segment search algorithm

3.4.2 Illustrative Example for Search Algorithm

Let the residue sequence of unknown protein be:

P Q I T L W Q R P L V T I K

The algorithm searches 3 to 7-residues-long segments of this chain in the database. First, 3-residues-long segments are searched; starting group is 'P Q I' and it goes on by skipping one residue at each step until all possible segments are traced as it is demonstrated in Figure 3.10.

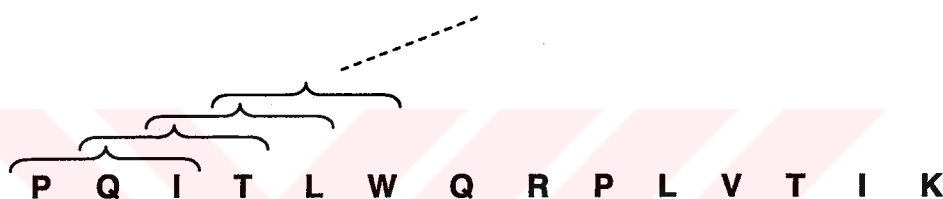


Figure 3.10: 3-residues-long segments searched in illustrative example for database search

Complete list of 3-residues-long segments in order of search is as follows:

- | | |
|----------|-----------|
| 1) P Q I | 7) Q R P |
| 2) Q I T | 8) R P L |
| 3) I T L | 9) P L V |
| 4) T L W | 10) L V T |
| 5) L W Q | 11) V T I |
| 6) W Q R | 12) T I K |

Then, 4-residues-long segments are searched; starting group is 'P Q I T' and again skipping one residue at a time as it is visualized in Figure 3.11.

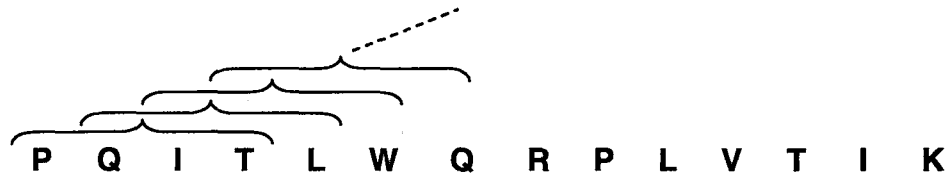


Figure 3.11: 4-residues-long segments searched in illustrative example for database search

List of the 4-residues-segments searched is as follows:

- | | |
|------------|-------------|
| 1) P Q I T | 7) Q R P L |
| 2) Q I T L | 8) R P L V |
| 3) I T L W | 9) P L V T |
| 4) T L W Q | 10) L V T I |
| 5) L W Q R | 11) V T I K |
| 6) W Q R P | |

Next comes 5-residues-long segments as follows:

- | | |
|--------------|---------------|
| 1) P Q I T L | 6) W Q R P L |
| 2) Q I T L W | 7) Q R P L V |
| 3) I T L W Q | 8) R P L V T |
| 4) T L W Q R | 9) P L V T I |
| 5) L W Q R P | 10) L V T I K |

List of 6-residues-long segments is as follows:

- | | |
|----------------|----------------|
| 1) P Q I T L W | 6) W Q R P L V |
| 2) Q I T L W Q | 7) Q R P L V T |
| 3) I T L W Q R | 8) R P L V T I |
| 4) T L W Q R P | 9) P L V T I K |
| 5) L W Q R P L | |

Finally, 7-residues-long segments are searched:

- | | |
|------------------|------------------|
| 1) P Q I T L W Q | 5) L W Q R P L V |
| 2) Q I T L W Q R | 6) W Q R P L V T |
| 3) I T L W Q R P | 7) Q R P L V T I |
| 4) T L W Q R P L | 8) R P L V T I K |

3.4.3 Statistical Secondary Structure Prediction

3.4.3.1 Secondary Structure Prediction Approach

After searching segments of residues in database, the frequency table is written as the output and next step is prediction of structure for each residue in sequence utilizing information on table of frequency of occurrences. A sample table is available in Table 3.3.

For each residue, (7-3+1) probabilities can be calculated for being in any type of structure by dividing frequency of occurrence in a certain type of structure in i-residues-long segment by all occurrences of residue in i-residues-long segments in structures in database. For the first residue in Table 3.3, if frequencies for 3-residues-segments are utilized;

$$P \{ \text{"residue a" is in an } \alpha\text{-helix utilizing 3-residues-segments} \} = \left(\frac{\text{Total number of occurrences of "residue a" in an } \alpha\text{-helix in 3-residue-groups in database}}{\text{Total number of occurrences of "residue a" in 3-residue-groups in database}} \right)$$

$$= 9 / (9+2+72) = \underline{0.108}$$

But, to figure out the ultimate probability for a residue's being in a certain type of structure, probability of being in that type of structure found with 3 to 7-residues-segments must be weighted. The ultimate probability for a residue to be in an α -helix, a β -sheet and a loop are calculated with equations (3.1), (3.2) and (3.3), respectively.

$$P_H = P \{ \text{"residue a" is in an } \alpha\text{-helix} \} = \sum_{i=3}^7 W_i * \left(\frac{\text{Total number of occurrences of "residue a" in an } \alpha\text{-helix in i-residue-groups in database}}{\text{Total number of occurrences of "residue a" in i-residue-groups in database}} \right) \quad (3.1)$$

$$P_S = P \{ \text{"residue a" is in a } \beta\text{-sheet} \} = \sum_{i=3}^7 W_i * \left(\frac{\text{Total number of occurrences of "residue a" in a } \beta\text{-sheet in i-residue-groups in database}}{\text{Total number of occurrences of "residue a" in i-residue-groups in database}} \right) \quad (3.2)$$

$$P_L = P \{ \text{"residue a" is in a loop} \} = \sum_{i=3}^7 W_i * \left(\frac{\text{Total number of occurrences of "residue a" in a loop in i-residue-groups in database}}{\text{Total number of occurrences of "residue a" in i-residue-groups in database}} \right) \quad (3.3)$$

The challenge is the determination of weights here, that is, to find out the most suitable numbers to multiply each probability. The question to be answered is what number of residue groups gives more reliable predictions for secondary structures. Is it better to put more weight on longer structures or shorter ones? It is a well-known perception that there is direct proportion between residue sequence similarity and structure type. So, it may seem reasonable to emphasize 6 and 7-residues-groups' probabilities while finding the final one. But when data is examined, it is observed that 6 or 7-residues-long segments match very rarely, which means very few data to calculate probability. So, considering long segments would lead to unreliable results calculated with only a few occurrences.

On the other hand, 3 and 4-residues-groups occur quite frequently, meaning more reliability for probabilities calculated using them. So, it is hard to draw a convincing decision on best weights without doing detailed analysis. In order to determine optimum weights for each structural class, its own pool of known proteins have been used and a non-linear programming model has been formulated to assure optimality. The set of known proteins for each structural class has been divided into two. While approximately 80% of data remained in database, other 20% was extracted and sequences of these proteins are used for testing. Lists of test data for each class are supplied in Appendices G, H, I and J. Sequences of test data have been extracted by a C code, they are entered in the search algorithm as input and their frequency matrices are received as the output. While this process was being carried out, the database was assured not to contain any data coming from a protein being predicted.

Probabilities for 3 to 7-residues-groups for each structural class and structure type have been calculated after text files including probability matrices were exported to Microsoft Excel and carrying out necessary calculations. The only lacking parameter set was the weights for each structural class and each number of residue groups. The aim is to select

the w_i values that will do the most accurate predictions for the real structures. The approach and the model for determination of weights is discussed in detail in the following section.

With optimal w_i values for each structural class, everything needed for prediction are completed. Final probabilities for each residue being in a certain structure are found by multiplying probabilities drawn with a number of residues in segment by its relevant weight by utilizing equations (3.1), (3.2) and (3.3). Structure with the maximum probability is the predicted structure for that residue. Since test data was treated as if it was unknown, real structures of these data are compared with predictions and matching residues are marked as '1' while others are marked '0'.

By averaging 0-1 scores for all residues in all proteins in each structural class, prediction accuracies are calculated. A sample matrix generated during calculation is available in Table 3.8. All prediction calculations are done in Microsoft Excel. The number of residues predicted in each class and prediction accuracies are displayed in Table 4.2.

3.4.3.2 Determination of Optimal Weights for Probabilities

After segments of residue sequence are searched in database, a set of probabilities are estimated for each residue and a balance should be obtained among these probability figures to assure best secondary structure prediction for the overall sequence. This requires calculation of optimal W_i ($i=3..7$) values for each structural class. This issue is accomplished by a Non-Linear Programming Model (NLP) formulated and run for each class.

Objective of the model is to minimize overall difference between real structures and predictions. The formulation is as follows:

SETS

s_real: Helix, Sheet, Turn (*real_structures*)

prob_structure: 3_H, 3_S, 3_T, 4_H, 4_S, 4_T, 5_H, 5_S, 5_T, 6_H, 6_S, 6_T, 7_H, 7_S, 7_T
(*probability_structure*)

residues: 1 .. total number of residues for each protein in the training data set

PARAMETERS

Realresidues_{s_real}: real structure for each residue in each protein in the training data set

Frequency_{residues, prob_structure}: frequency of occurrences in relevant structure for each residue in each protein in the training data set

VARIABLES

W_i: weight for probability of occurrence of a residue in an i-residue-structure (i: 3, 4,..., 7) (*nonnegative variables*)

CONSTRAINTS

$$\sum_{i=3}^7 W_i = 1$$

OBJECTIVE FUNCTION

$$\begin{aligned} \text{Min } z = \sum_{i=3}^7 \text{sqr}(W_i) & \left\{ \sum_{\text{residues}} \text{sqr}[\text{Real}_{r,\text{Helix}} - \text{Frequency}_{i,H} / (S^* + \text{Frequency}_{i,H} + \text{Frequency}_{i,S} + \text{Frequency}_{i,T})] \right\} \\ & + \left(\sum_{\text{residues}} \text{sqr}[\text{Real}_{r,\text{Sheet}} - \text{Frequency}_{i,S} / (S^* + \text{Frequency}_{i,H} + \text{Frequency}_{i,S} + \text{Frequency}_{i,T})] \right) \\ & + \left(\sum_{\text{residues}} \text{sqr}[\text{Real}_{r,\text{Turn}} - \text{Frequency}_{i,T} / (S^* + \text{Frequency}_{i,H} + \text{Frequency}_{i,S} + \text{Frequency}_{i,T})] \right) \end{aligned}$$

* S is a very small number to avoid probable division by 0 while not having a significant effect on the final probability figure.

To determine optimum weights for each structural class, the set of known proteins has been divided into two. While approximately 80% of data remained in database, remaining 20% was extracted and sequences of these proteins are used for testing. So, we obtained 8 sets of data, 4 sets containing 20% of pool for each class to be used as test data and 4 sets containing 80% of data for database of each class. The datasets are divided into two to ensure that databases where test data will be searched do not contain any structure from a test protein. List of proteins in test set are supplied in Appendices G, H, I, J and list of proteins in database set are in Appendices K, L, M, N for each class. Also, number of test proteins and number of the ones selected for database for each class are displayed in Table 3.4.

Table 3.4: Information on number of proteins in test set and database set

	Number of proteins in database	Number of structures in database	Number of proteins predicted	Number of residues predicted
all- α	2000	41139	419	203302
all- β	2766	67373	579	313844
α/β	3375	100255	707	505693
$\alpha+\beta$	2866	62935	601	279355

Sequences of test data have been entered in the search algorithm as the input and their frequency matrices are received as the output. While this process was being carried out, the database was assured not to contain any data coming from a protein being predicted.

To solve this NLP problem, real structures of residue sequences must be entered. This is accomplished by extracting real structures of these proteins from their PDB files in the

form of a matrix. A C Programming Language code writes real structures of proteins and major parts of NLP code to be solved in Gams optimization environment into text files.

3.4.3.3 Illustrative Example for Statistical Secondary Structure Prediction

(Illustrative example in section 3.4.2 is continued here for complete demonstration)

Let the residue sequence be the same as the one used in Section 3.4.2:

P Q I T L W Q R P L V T I K

Let this sequence be entered in search algorithm code SEARCH.C and Table 3.5 be the frequency matrix obtained as the output of the code through utilization of its relevant database.

Table 3.5: Frequency matrix of illustrative example

	3_H	3_S	3_T	4_H	4_S	4_T	5_H	5_S	5_T	6_H	6_S	6_T	7_H	7_S	7_T
P	14	30	57	1	0	3	0	0	2	0	0	2	0	0	2
Q	21	41	129	1	0	9	0	0	4	0	0	4	0	0	4
I	59	55	241	2	0	12	0	0	6	0	0	6	0	0	6
T	68	39	238	4	0	16	0	0	8	0	0	8	0	0	8
L	100	43	189	4	0	15	0	0	10	0	0	10	0	0	9
W	82	40	109	4	0	12	0	0	10	0	0	11	0	0	10
Q	84	45	98	8	1	11	2	0	9	0	0	10	0	0	11
R	76	49	131	11	1	9	2	0	8	0	0	9	0	0	10
P	94	54	138	10	1	8	2	0	7	0	0	8	0	0	9
L	74	41	116	10	2	7	2	0	6	0	0	7	0	0	8
V	45	25	100	6	1	7	2	0	5	0	0	6	0	0	7
T	8	9	64	1	1	6	0	0	5	0	0	6	0	0	7
I	3	7	62	1	1	7	0	0	5	0	0	6	0	0	7
K	11	11	36	2	1	6	0	0	5	0	0	6	0	0	7

Then, probabilities for each residue will be calculated by utilizing equations 3.1, 3.2 and 3.3. A sample calculation for first residue in sequence is as follows:

$$P_H = P \{ \text{"residue P" is in an } \alpha \text{-helix} \} = W_3*(14/(14+30+57)) + W_4*(1/(1+0+3)) + W_5*(0/(0+0+2)) \\ + W_6*(0/(0+0+2)) + W_7*(0/(0+0+2))$$

$$P_S = P \{ \text{"residue P" is in a } \beta \text{-sheet} \} = W_3*(30/(14+30+57)) + W_4*(0/(1+0+3)) + W_5*(0/(0+0+2)) \\ + W_6*(0/(0+0+2)) + W_7*(0/(0+0+2))$$

$$P_L = P \{ \text{"residue P" is in a loop} \} = W_3*(57/(14+30+57)) + W_4*(3/(1+0+3)) + W_5*(2/(0+0+2)) \\ + W_6*(2/(0+0+2)) + W_7*(2/(0+0+2))$$

Weights are needed for exact figures. Let the values in Table 3.6 be the W_i values obtained for structural class of this protein.

Table 3.6: Hypothetical weights for illustrative example

W_3	W_4	W_5	W_6	W_7
0.18	0.27	0.26	0.17	0.12

Then, exact probability figures can be calculated for each residue. Structures with maximum probabilities will be the predictions of the algorithm. To measure assessment of the method, real structure of the sequence should be obtained. Real structure matrices are extracted. Let the structure of this hypothetical sequence be the one displayed in Table 3.7. Fields marked with 1 are the real structures. Table 3.8 shows all prediction steps and assessment for this sequence with specified data. First 15 columns constitute frequency matrix for residues. Next 3 columns display real structure of the sequence. Probabilities of

being in each type of structure are calculated in following 3 columns by utilizing equations 3.1, 3.2 and 3.3 and weights in Table 3.6.

Table 3.7: Real structure matrix for illustrative example

	H	S	T
P	0	1	0
Q	0	1	0
I	0	0	1
T	0	0	1
L	0	0	1
W	0	0	1
Q	1	0	0
R	1	0	0
P	1	0	0
L	1	0	0
V	1	0	0
T	1	0	0
I	0	0	1
K	0	0	1

Structures with highest probabilities are marked as predictions and they are visualized in letters in successive columns. Finally, real structure of the residue is visualized via letters, too and residues where predicted and real structures match are marked as '1' in the last column.

Accuracy of prediction in this example can be calculated by averaging 0 – 1 values in the last column of Table 3.8, which is 42.86% for this case.

Table 3.8: Predictions of illustrative example

Frequency of occurrence of residues in 3 to 7-residues-long segments in database																	Real structures of residues			Probabilities of being a helix, sheet and loop			Prediction (Predicted structure is 1)					
Residue																												
	3_H	3_S	3_I	4_H	4_S	4_I	5_H	5_S	5_I	6_H	6_S	6_I	7_H	7_S	7_I	R_H	R_S	R_I	Prob (H)	Prob (S)	Prob (L)	Prediction (H)	Prediction (S)	Prediction (L)	Predicted Structure	Real Structure	Result	
P	14	30	57	1	0	3	0	0	2	0	0	2	0	0	2	0	1	0	0.09	0.05	0.85	0	0	1	S	T	0	
Q	21	41	129	1	0	9	0	0	4	0	0	4	0	0	4	0	1	0	0.05	0.04	0.91	0	0	1	S	T	0	
I	59	55	241	2	0	12	0	0	6	0	0	6	0	0	6	0	0	1	0.07	0.03	0.90	0	0	1	T	T	1	
T	68	39	238	4	0	16	0	0	8	0	0	8	0	0	8	0	0	1	0.09	0.02	0.89	0	0	1	T	T	1	
L	100	43	189	4	0	15	0	0	10	0	0	10	0	0	9	0	0	1	0.11	0.02	0.87	0	0	1	T	T	1	
W	82	40	109	4	0	12	0	0	10	0	0	11	0	0	10	0	0	1	0.13	0.03	0.84	0	0	1	T	T	1	
Q	84	45	98	8	1	11	2	0	9	0	0	10	0	0	11	1	0	0	0.22	0.05	0.73	0	0	1	H	T	0	
R	76	49	131	11	1	9	2	0	8	0	0	9	0	0	10	1	0	0	0.25	0.05	0.71	0	0	1	H	T	0	
P	94	54	138	10	1	8	2	0	7	0	0	8	0	0	9	1	0	0	0.26	0.05	0.69	0	0	1	H	T	0	
L	74	41	116	10	2	7	2	0	6	0	0	7	0	0	8	1	0	0	0.26	0.06	0.67	0	0	1	H	T	0	
V	45	25	100	6	1	7	2	0	5	0	0	6	0	0	7	1	0	0	0.24	0.05	0.72	0	0	1	H	T	0	
T	8	9	64	1	1	6	0	0	5	0	0	6	0	0	7	1	0	0	0.05	0.05	0.89	0	0	1	H	T	0	
I	3	7	62	1	1	7	0	0	5	0	0	6	0	0	7	0	0	1	0.04	0.05	0.91	0	0	1	T	T	1	
K	11	11	36	2	1	6	0	0	5	0	0	6	0	0	7	0	0	1	0.09	0.06	0.84	0	0	1	T	T	1	

Chapter 4

RESULTS

A two-stage algorithm has been developed in this study to predict secondary structures of protein molecules; the first stage is the determination of structural class of an unknown protein, the second stage is the prediction of secondary structure by considering structures of known proteins belonging to the same class. Details of methods of both phases are explained in Chapters 2 and 3.

This study yielded 3 main results after its thorough processes and calculations:

1. 100% accuracy in predicting fold type of an unknown protein; since result of first part is directly used in second phase, 100% accuracy is very important for overall accuracy of the prediction method.
2. The weights for each structural class for probabilities drawn from 3 to 7-residues-long segments
3. Prediction performance of the method tried on 20% of structurally identified proteins from each class as if they were unknown

These results are discussed in detail in the following sections.

4.1 Determination of Secondary Structure

The mixed-integer programming model was used to find the boundaries of four folding types in the 20-dimensional amino acid composition space. The testing is done on 122 subunits that mostly belong to different proteins as given in Table 4.1.

Table 4.1. The list of 122 proteins that are used to train the mixed-integer programming model for determining folding types.

PDB_ID	Chain	Folding Type	PDB_ID	Chain	Folding Type
1ABA	A	α/β	1TIM	A	α/β
1ACX	A	β	1TLK	A	β
1AVH	A	α	1TRE	A	α/β
1BAB	B	α	1TRO	A	α
1BKS	B	α/β	1ULA	A	α/β
1BRD	A	α	1UTG	A	α
1C5A	A	α	256B	B	α
1CD8	A	β	2AAK	A	$\alpha+\beta$
1CDT	A	β	2ACH	A	$\alpha+\beta$
1CID	A	β	2ACT	A	$\alpha+\beta$
1CIS	A	α/β	2ALP	A	β
1CPC	A	α	2AVI	A	β
1CPC	L	α	2AYH	A	β
1CSE	I	$\alpha+\beta$	2BPA	A	β
1CTF	A	$\alpha+\beta$	2BPA	B	β
1DBP	A	α/β	2CCY	A	α
1DFN	A	β	2CTC	A	α/β
1DHR	A	α/β	2FOX	A	α/β
1DNK	A	$\alpha+\beta$	2HAD	A	α/β
1EAF	A	α/β	2ILA	A	β
1ECO	A	α	2LAL	A	β
1ETU	A	α/β	2LH1	A	α
1FCS	A	α	2LHB	A	α
1FHA	A	α	2LIG	A	α
1FIA	B	α	2LIV	A	α/β
1FXI	A	$\alpha+\beta$	2MHB	A	α
1FXI	B	$\alpha+\beta$	2MHB	B	α
1FXI	C	$\alpha+\beta$	2OMF	A	β
1FXI	D	$\alpha+\beta$	2PGD	A	α/β
1GPB	A	α/β	2PHY	A	$\alpha+\beta$

PDB_ID	Chain	Folding Type	PDB_ID	Chain	Folding Type
1HBG	A	α	2PRF	A	$\alpha+\beta$
1HDD	C	α	2RNT	A	$\alpha+\beta$
1HIG	A	α	2SNS	A	$\alpha+\beta$
1HIL	A	β	2SNV	A	β
1HIV	A	β	2TMD	A	α/β
1HLE	B	β	2VAA	A	$\alpha+\beta$
1HSB	A	$\alpha+\beta$	2VAA	B	β
1LE4	A	α	2ZTA	A	α
1LTS	A	$\alpha+\beta$	3CD4	A	β
1LTS	C	β	3GBP	A	α/β
1LTS	D	β	3HHR	C	β
1MAM	H	β	3IL8	A	$\alpha+\beta$
1MBC	A	α	3MON	A	$\alpha+\beta$
1MBS	A	α	3RUB	S	$\alpha+\beta$
1NRC	A	$\alpha+\beta$	3SGB	I	$\alpha+\beta$
1OFV	A	α/β	3SIC	I	$\alpha+\beta$
1OVB	A	α/β	3SSI	A	$\alpha+\beta$
1PAZ	A	α/β	4BLM	A	$\alpha+\beta$
1PFK	A	α/β	4GCR	A	β
1PPN	A	$\alpha+\beta$	4MBA	A	α
1Q21	A	α/β	4MBN	A	α
1REI	A	β	4TMS	A	$\alpha+\beta$
1RND	A	$\alpha+\beta$	5CPA	A	α/β
1RPR	A	α	5P21	A	α/β
1S01	A	α/β	8ABP	A	α/β
1SBP	A	α/β	8ATC	A	α/β
1SBT	A	α/β	8CAT	A	$\alpha+\beta$
1SNC	A	β	8FAB	A	β
1TEN	A	β	8FAB	B	β
1TFG	A	$\alpha+\beta$	8IIB	A	β
1TGS	I	$\alpha+\beta$	9RSA	A	$\alpha+\beta$

The training of the mixed-integer programming has been completed with 100% accuracy. Then, the consistency of this method is tested by randomly selecting 61 subunits that mostly belong to different proteins as given in Table 4.2.

Table 4.2. The list of 61 proteins that are used to test the mixed-integer programming model for determining folding types.

PDBID	Chain	Folding	PDBID	Chain	Folding Type
2AAI	B	β	1SHA	A	$\alpha+\beta$
2ABH	A	α/β	1SHF	A	β
1ATX	A	β	1THO	A	α/β
1BBL	A	α	1TIE	A	β
1BBP	A	β	1TNF	A	β
1BKS	A	α/β	1XOB	A	α/β
1BW4	A	β	2AAA	A	β
1COB	A	β	2ABH	A	α/β
1DNK	A	$\alpha+\beta$	2CTX	A	β
1EGF	A	β	2MEV	1	β
1EST	A	β	2MS2	A	$\alpha+\beta$
1FOJ	A	$\alpha+\beta$	2PIA	A	α/β
1GD1	O	α/β	2PLV	1	β
1GLA	G	α/β	2SN3	A	$\alpha+\beta$
1GPB	A	α/β	2SOD	O	β
1GPS	A	α/β	2TAA	A	α/β
1HBB	A	α	2TMV	P	α
1HCC	A	β	1CYO	A	$\alpha+\beta$
1IFA	A	α	3COX	A	$\alpha+\beta$
1IXA	A	β	3RP2	A	β
1MDA	A	β	3SC2	A	α/β
2MIN	A	α/β	4CPV	A	α
1MRR	A	α	4ENL	A	α/β
1NIP	B	α/β	4ICD	A	α/β

PDBID	Chain	Folding	PDBID	Chain	Folding Type
1OVO	A	$\alpha+\beta$	4INS	B	$\alpha+\beta$
1POC	A	α	4SGB	I	β
1PPF	E	β	5NN9	A	β
1PRC	M	α	7AAT	A	α/β
1RIA	2	β	8TLN	E	$\alpha+\beta$
2SAS	A	α	9RUB	B	α/β
1SBP	A	α/β			

The testing also concluded that the mixed-integer programming model determines the folding type with 100% accuracy using the amino acid composition only.

4.2 Optimal Weights of Residue Segment Length

Prediction method in this study considers sequence similarity to estimate secondary structure. All sequential segments from 3 to 7-residues-long are searched in structures in database. Although structural similarity increases as segments get longer, frequency of occurrence in database decreases which means less reliable results due to reduced data. On the other hand, 3 or 4-residues-sequences were much more frequent but less reliable since sequence similarity is less. To achieve the balance between these two trade-offs (structural similarity and frequency of occurrence), optimal weights are found to multiply each probability figure calculated by using frequencies of occurrences in 3 to 7-residues-segments. Method of finding w_i 's is explained in Section 3.4.3.2. Optimal w_i values obtained as results of Non-Linear Programming formulations are displayed in Table 4.3.

Table 4.3: Optimal weights for each structural class for segments of 3 to 7 residues

	w_3	w_4	w_5	w_6	w_7
all-α	0,175	0,218	0,223	0,202	0,182
all-β	0,192	0,226	0,216	0,193	0,174
α/β	0,176	0,208	0,226	0,206	0,185
$\alpha+\beta$	0,174	0,215	0,223	0,204	0,183

4.3 Prediction Results

Databases belonging to each structural class are divided into two: 20% of data being tested and 80% making up the database to be searched. These were used to find out optimal weights in Section 3.4.3.2. After weights are determined, prediction of structures of test data is completed with the method presented in Section 3.4.3.1. Since test data was treated as if it was unknown, real structures of these data are compared with predictions and matching residues are marked as '1' while others are marked '0'. By averaging 0-1 scores for all residues in all proteins in each structural class, prediction accuracies are calculated. Number of residues predicted in each class and overall prediction accuracies are displayed in Table 4.4. The list of names of proteins predicted, their chain letters, number of residues and prediction accuracies are supplied in Appendices O, P, Q and R for all- α , all- β , α/β and $\alpha+\beta$ folding types, respectively.

Table 4.4: Results of secondary structure prediction for trial data sets from each structural class

	Number of proteins in database	Number of structures in database	Number of proteins predicted	Number of residues predicted	<u>Prediction Accuracy (%)</u>
all- α	2000	41139	419	203302	<u>80.5</u>
all- β	2766	67373	579	313844	<u>72.4</u>
α/β	3375	100255	707	505693	<u>71.9</u>
$\alpha+\beta$	2866	62935	601	279355	<u>75.5</u>
Total	11007	271702	2306	1302194	<u>74.1</u>

4.4 Discussion of Results

Mixed-integer programming model for determining folding types is trained with 122 subunits that mostly belong to different proteins. The method is tested on 61 subunits, again belonging mostly to different proteins, and overall accuracy achieved is 100% which means solution of secondary structure determination problem when amino acid composition of the protein is known. Second phase of the method in this thesis directly utilizes results of folding type determination by selecting relevant database to search residue segments.

Optimal weights calculated in this study resolve the trade-off between sequence similarity and data reliability. When results are examined, for all classes, weights of probabilities calculated utilizing 4 and 5-residues-segments are highest. This was the expected case since 6 and 7-residues-segments are very rarely observed, leading to very high probability in one structure and very low in others. And if the high probability belongs to a structure other than real structure of predicted residue, putting heavy weight on this

probability can lead to wrong estimations. That is the reason for low weights of 6 and 7 residues cases, 7-residues-groups being the least. Since 3-residues-long segments are very small to reveal any structural similarity and occur in diverse type of structures, they are more far away from providing a general picture on the structure of the unknown sequence, and this fact leads to their low weights among others.

Protein secondary structure prediction in order to discover its functions has been an important, challenging problem; research on the subject have been accelerated and emphasized in the last 30 years. Remaining still unsolved, very significant progress has been achieved with current state-of-the-art of 78% accuracy with Copenhagen method (Petersen *et al.*, 2000) on the basis of 3 states (α -helix, β -sheet or loop) per residue prediction. Introduction part of this report contains a review of methods developed in literature and their performances. Best methods up to date are displayed here again with their accuracy levels in Table 4.5.a and Table 4.5.b and performance of method developed in this study is available in Table 4.5.c for comparative purposes. It should be emphasized here that the accuracy levels in Table 4.5.c are calculated according to the raw prediction results of the algorithm.

Table 4.5.a: Accuracies of Secondary Structure Prediction Methods (Rost, 2001)

Method	Q ₃ (%)	Q ₃ Claim (%)
PROF	77.0	
PSIPRED	76.6	76.5-78.3
SSpro	76.3	76
JPred2	75.2	76.4
PHDpsi	75.1	
PHD	71.9	71.6
Copenhagen	78	77.8

Table 4.5.b: Accuracies of Secondary Structure Prediction Methods (Cuff and Barton, 1999)

Method	Q ₃ (%)
PHD	71.950
NNSSP	71.400
DSC (King and Sternberg, 1996)	68.413
PREDATOR	68.602
ZPRED (Zvelebil, Barton, Taylor and Sternberg, 1987)	59.637
Consensus	72.707

Table 4.5.c: Accuracy of Two-Stage Secondary Structure Prediction Method of this thesis

Folding Type	all- α	all- β	α/β	$\alpha+\beta$	Total
Q ₃ (%)	80.5	72.4	71.9	75.5	74.1

No manual correction has been done to adjust predictions according to pattern of estimated structures of preceding and succeeding residues at any point in the chain. When prediction results of method in this study is compared with others mentioned in literature review, prediction accuracies are quite successful in all cases and even better than current state-of-the-art for all- α case. Since method is tested on 419 all- α , 519 all- β , 707 α/β and 601 $\alpha+\beta$ proteins, accuracy values are much more reliable than those of other methods listed in Table 4.5.a and Table 4.5.b; they all have been tested on less data.

Chapter 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Major Conclusions

Discovery of functions of proteins is one of the most challenging problems in science and technology, impacts of which will be great contributions in understanding biological processes and speeding up drug discoveries for more efficient, direct hit drugs with less side effects. Functions and active sites of proteins are determined by their molecular structures and aim of this study was to predict secondary structures of unknown proteins to estimate their functions. Predictions with reliable accuracies can have an accelerating effect on all processes needing information of functions of proteins. The method developed assigns each residue in protein's amino acid sequence to be part of either an α -helix a β -sheet or a loop. A two-stage algorithm has been developed for this goal; first stage is assignment of the unknown protein into one of 4 structural classes, namely all- α , all- β , α/β and $\alpha+\beta$ and second stage is searching segments of new protein's residue sequence in structures of experimentally identified proteins belonging to the same structural class with the unknown one. The first stage is modeled and solved using Mixed Integer Linear Programming approach.

The second stage is accomplished by a set of tools and methods. Experimental result files are taken from PDB for known proteins to take place in database. They are processed by a program coded in C Programming Language. An optimization code is solved in GAMS to determine best weights for different segment lengths. Gams code is also written

by a C program. After completion of search operation, prediction is done by utilizing text files, Microsoft Access and Microsoft Excel.

The algorithm is tested on 419, 579, 707 and 601 different PDB files for prediction of all- α , all- β , α/β and $\alpha+\beta$ proteins, respectively. Results are quite satisfactory in terms of average prediction accuracy. Best accuracy is achieved with all- α proteins, 80.5%, which is better than the current state-of-the-art. Claim of this study was to predict secondary structures of unknown proteins by a two-stage algorithm relying on statistical data and with the results achieved, the claim is validated with a significant prediction performance, even with results of automated operations, without doing manual adjustments or corrections on the predictions. Such adjustments can be very effective in achieving better results and higher prediction accuracies. This approach will be discussed further in Section 5.3.

5.2 Contributions

5.2.1 Two-Stage Algorithm Approach

The method integrates two algorithms for prediction of secondary structures. First step is the prediction of structural class of new protein and second is prediction of its secondary structure on 3-states (α -helix, β -sheet or loop) per residue basis.

5.2.2 Folding Type Determination

A mixed-integer programming model is developed that is trained with subunits of known proteins and utilizes amino acid composition to determine folding type of unknown proteins with 100% accuracy.

5.2.3 Statistical Weights

The trade-off between sequence similarity and data availability is analyzed in this study by using a nonlinear programming formulation to determine optimal weights to emphasize search results of residue sequences of different lengths. In our method, probabilities are calculated for each residue in sequence for being in a certain type of structure (α -helix, β -sheet or loop) considering 3 to 7 residues in each segment being searched. But the approach can easily and effectively be adapted to longer or shorter segments.

5.2.4 Prediction of Secondary Structure

A prediction method relying on structural information on experimentally identified proteins' database of structures and a search routine extracting each and every possible residue sequence segment with amino acid content between minimum and maximum parameters selected are developed. Final prediction approach directly depends on statistical search results and optimal weights determined for each structure type and segment population. Search algorithm is coded in C Programming Language. Besides its quite satisfactory performance in prediction, runtime duration of the algorithm is short enough to assure computational efficiency.

5.3 Recommendations

The method described in this study is a completely automated approach; everything is executed by computer codes to be able to obtain more accurate results with increased data availability by doing more tests in a shorter time. Databases utilized consist of 41139, 67373, 100255 and 62935 structures to predict 419 all- α , 579 all- β , 707 α/β and 601 $\alpha+\beta$

proteins. Various precautions are included in code processing experimental result files of proteins to avoid unreliable data utilization. Formatting problems, inconsistencies and some other troubles in PDB files are detected by C code. This automation resulted in automated solutions, too which can be improved significantly by a set of manipulations.

Exclusion of homologous structures is carried out by a query in Microsoft Access. This query only compares the whole structures and if they are completely the same in segment's sequence and starting and ending positions, it eliminates duplicates of them. But there are quite many cases where starting and ending positions of segments differ only one residue or segments are completely same except a single residue inside it. Some amino acids have similar structural and chemical properties that do not tend to change the structure of the sequence when they are altered with each other. The source of this alteration is called "mutation". Mutation is the change in structure like loss of some part due to various possible factors including chemicals and radiation. The resulting residue after mutation has very similar chemical and physical properties as the original one and does not tend to change the structure of the polypeptide chain or functions of protein, which means creation of homology between original and new chains. Although sequence of amino acids is completely same except for a few mutated residues, computerized queries can not take these facts into account. Mutations in some residue types usually end up with some other residues belonging to a certain set of probable amino acids. Possible exchanges of residues are scored according to their tendencies for each of 20 amino acid types resulting in a table called "Substitution Matrix" (Dayhoff and Eck, 1968; Dayhoff, 1978; Henikoff and Henikoff, 1992). Homologies in sequences can be searched by utilizing this matrix and additional homologies created by probable replacements of residues can be marked in databases of structural classes. These issues may be traced manually in database, but it should be noted that this is a severely rigorous operation that either takes tremendous amount of time or drastically decreases quantity of data in database. This matrix could also

be utilized for finding most probable mutations for residues in sequence of new protein so that, mutated amino acids could be replaced with the associated ones and database could be searched again with the altered sequence. This could provide more logical, accurate predictions and opportunity for comparison with first prediction with the original residue sequence.

Another approach is changing predictions for some residues in accordance with predicted structures of their preceding and succeeding residues. A hypothetical residue sequence and their predictions are shown in Table 5.1 for illustrative purpose.

Table 5.1: Illustrative example for manual adjustments on secondary structure predictions of the method

* H: residue is predicted to be belonging to an α -helix

S: residue is predicted to be belonging to a β -sheet

L: residue is predicted to be belonging to a loop

Residue Sequence	A	R	N	D	C	Q	E	D	N	A	K	F	M	
Predicted Structure*	H	H	H	S	H	H	L	L	S	S	S	S	S	...

Structure of residue 'D' is predicted to be a β -sheet while it is obvious that it can not be actually. Its preceding 3 and succeeding 2 residues are assigned to be helices and 'D' is predicted to be sheet in the middle of them. A sheet can be formed with at least 2 residues and it is most probable that the complete set of 6 residues from 'A' to 'E' belong to a helix.

If all predictions could be traced for such cases, a significant improvement would be observed in accuracy.

Chou and Fasman (1978) and Garnier and coworkers (1978) searched structural database of proteins for occurrence of amino acids in each of the secondary structure types; α -helices, β -sheets and loops. It was observed that some amino acids are known to dictate or destroy especially one of three secondary structural shapes. For example the residue proline is not found in α -helices frequently; because its side chain is not suitable for being in a helical structure. Other similar cases can be investigated and prediction results of algorithm can be examined to assure that they satisfy these special conditions where mentioned special amino acids occur in sequence. This will again mean more reliable and accurate estimations.

Another approach could be testing algorithm with more specialized proteins. There are studies in literature focusing only on membrane proteins, which are all- α (Pilpel *et al.*, 1999; Pasquier *et al.*, 1999; Chou and Elrod, 1999). Membrane proteins or hydrophobic, hydrophilic, soluble, insoluble, polar, non-polar set of proteins could be selected and utilized with the method presented in this study. This could probably lead to better accuracy figures since proteins with similar properties tend to have similar structures.



APPENDICES

APPENDIX A

Names of proteins in all- α class in SCOP; <http://scop.mrc-lmb.cam.ac.uk>

0c3a	1aum	1bvp	1cz2	1e96	1fdh	1gkx	1hrm	1jcq	1ksu	1mun	1qtg	1zqn	3car
101m	1auw	1bvs	1czj	1e9h	1fdk	1gkz	1hro	1jcr	1ktp	1mux	1qu2	1zqo	3cep
102m	1av8	1bvy	1d1d	1e9x	1fe2	1gli	1hrs	1jcs	1ku1	1muy	1qu3	1zqp	3ccx
103m	1avc	1bw5	1d1l	1ea1	1fe5	1glm	1hry	1jdb	1ku2	1mwc	1quu	1zqq	3cln
104m	1avh	1bw6	1d1m	1ea3	1few	1gln	1hrz	1jdh	1ku3	1mwd	1rlb	1zqr	3cp4
105m	1avm	1bwo	1d1o	1ea4	1fex	1glp	1hs5	1jdi	1ku7	1mxl	1rlr	1zqs	3cpp
106m	1avo	1bx9	1d2d	1ea8	1ff1	1glq	1hs6	1jdl	1kv9	1myf	1r2a	1zqt	3crd
107m	1avr	1bxi	1d2t	1eak	1ffh	1gm0	1hs7	1jdo	1kvo	1myg	1r2f	1zug	3cro
108m	1avs	1bxm	1d2v	1eb4	1ffk	1gm5	1hsj	1jdr	1kvw	1myh	1r63	1zym	3crx
109m	1aw9	1bxr	1d2z	1eb7	1ffu	1gmz	1hsm	1jeb	1kvx	1myi	1r69	20gs	3csc
10gs	1awc	1by1	1d3u	1ebc	1ffv	1gn2	1hsn	1jei	1kvy	1myj	1rap	21gs	3csm
110m	1ax8	1bye	1d4c	1ebd	1ffy	1gn3	1hsr	1jeq	1kw2	1myk	1raq	22gs	3ctn
111m	1axd	1byt	1d4d	1ebe	1fg9	1gn4	1hss	1jfo	1kw4	1myl	1rcb	256b	3cyr
112m	1axf	1bz0	1d4e	1ebm	1fgj	1gn6	1hst	1jft	1kwf	1mym	1rcc	2a0b	3cyt
11gs	1axi	1bz1	1d4u	1ebt	1fgm	1gnc	1hsy	1jft	1kwj	1myp	1rcd	2a8v	3erd
12gs	1axn	1bz4	1d5l	1ebv	1fgo	1gne	1ht5	1jfa	1kx2	1myt	1rce	2abd	3ert
13gs	1ayg	1bz6	1d5n	1eca	1fgq	1gni	1ht8	1jfa	1kx7	1mzl	1rcg	2abk	3eza
14gs	1ayi	1bap	1d5v	1ecd	1fgr	1gnj	1ht9	1jfb	1kxm	1mzm	1rci	2af8	3ezb
155c	1ayp	1bap	1d5y	1eci	1fgt	1gnw	1hta	1jfc	1kxn	1n15	1rcp	2asr	3eze
16gs	1ayx	1bzz	1d7w	1ecm	1fha	1go3	1htj	1jfg	1kxp	1n50	1rcs	2atj	3fap
17gs	1azi	1c02	1d8b	1ecn	1fha	1go5	1hu3	1jfi	1kxu	1n90	1rec	2av8	3fis
18gs	1azs	1c03	1d8d	1eco	1fhe	1god	1hue	1jfi	1kyo	1nbb	1rep	2bbm	3fyg
19gs	1azt	1c07	1d8e	1ecw	1fhf	1got	1hul	1jfk	1kyw	1nbn	1res	2bbn	3gly
19hc	1azy	1c0w	1d8j	1ed1	1fhj	1gp2	1huo	1jfs	1kyz	1ncx	1ret	2bby	3gss
1a00	1b01	1c15	1d8k	1edi	1fi3	1gpj	1hus	1jft	1kz7	1ncy	1rfb	2bca	3gst
1a01	1b06	1c1j	1d8l	1edj	1fi5	1gqm	1huu	1jft	1kzg	1ncz	1rgp	2bcb	3gtu
1a03	1b0b	1c1k	1d8u	1edk	1fi6	1grj	1huw	1jgc	1l0o	1neq	1rh2	2bcc	3gwx
1a04	1b0n	1c20	1d9c	1edl	1fi7	1grl	1huz	1jgg	1l0v	1ner	1rhg	2bct	3had
1a0a	1b0x	1c26	1d9g	1edu	1fi9	1grn	1hv4	1jgs	1l1m	1new	1rib	2bmh	3hbi
1a0b	1b1b	1c2n	1daq	1ee4	1fia	1gsb	1hv6	1jhb	1l1y	1nfl	1rlr	2bp2	3hdd
1a0f	1b1g	1c2r	1dat	1ee5	1fin	1gsc	1hvd	1jhf	1l2a	1nfn	1rmi	2bpf	3hdd
1a0p	1b1u	1c30	1dav	1eem	1fio	1gsd	1hve	1jhg	1l2i	1nfo	1rmv	2bpg	3hbb
1a0r	1b22	1c3c	1db1	1eer	1fip	1gse	1hvf	1jhh	1l2j	1ngl	1rnl	2bpp	3hhr
1a0u	1b25	1c3d	1db4	1eex	1fiq	1gsf	1hvg	1ji5	1l2k	1ngr	1rmr	2btv	3hsf
1a0v	1b28	1c3o	1db5	1efl	1fj0	1gsq	1hw1	1ji7	1l3e	1nhm	1rom	2c2c	3hts
1a0w	1b33	1c3u	1db7	1ef4	1fjg	1gss	1hw2	1jia	1l3l	1nhn	1rop	2cbl	3icb
1a0x	1b3q	1c3y	1db8	1efa	1fjl	1gsu	1hw5	1jig	1l3o	1nih	1rpe	2ccp	3ink
1a0y	1b3u	1c3z	1db9	1efr	1fk0	1gsy	1hwg	1jin	1l5j	1nir	1rpo	2ccy	3lbd
1a0z	1b43	1c40	1dbc	1efu	1fk1	1gt8	1hwh	1jio	1l6e	1nk2	1rpr	2cdv	3lhb
1a17	1b47	1c52	1dbh	1efy	1fk2	1gta	1hx1	1jip	1l6j	1nk3	1ro	2cep	3ljr
1a1u	1b48	1c53	1dcc	1eg3	1fk3	1gtb	1hx8	1jj2	1l6n	1nkd	1rse	2cgp	3lyn

1a1w	1b4a	1c5a	1dce	1eg4	1fk4	1gte	1hxd	1jj6	1l7c	1nkl	1rss	2cmm	3mba
1a1z	1b4c	1c6o	1dcn	1egc	1fk5	1gth	1hxi	1jj8	1l8c	1nno	1rtp	2cnp	3mdd
1a22	1b4f	1c6r	1dcy	1egd	1fk6	1gti	1hy0	1jic	1l8r	1nol	1run	2cp4	3mde
1a23	1b4n	1c6s	1dd3	1ege	1fk7	1gto	1hyl	1jir	1l8w	1noo	1ruo	2cpg	3mds
1a24	1b4u	1c72	1dd4	1egm	1fkm	1gtu	1hy5	1jjs	1l8y	1nre	1rvw	2cpo	3ngl
1a26	1b4w	1c74	1ddc	1egu	1fli	1gu6	1hy8	1jju	1l8z	1ns1	1ryc	2cpp	3orc
1a28	1b59	1c75	1ddf	1egv	1flo	1gu9	1hyp	1jko	1l9b	1nsg	1ryt	2crk	3p2p
1a29	1b5l	1c7b	1ddn	1egy	1flp	1guh	1hz4	1jko	1l9j	1ntc	1rzt	2cro	3pal
1a2a	1b67	1c7c	1ddx	1eh2	1fm6	1guk	1hzi	1jko	1la3	1nub	1sae	2crx	3pat
1a2f	1b68	1c7d	1dc4	1eh6	1fm9	1gul	1hzu	1jkq	1la6	1obm	1saf	2csc	3pax
1a2g	1b6a	1c7m	1dee	1eh7	1fmi	1gum	1hzu	1jkr	1lb1	1occ	1sag	2csm	3pgh
1a2i	1b6q	1c7v	1deg	1eh8	1fn7	1gux	1i08	1jku	1lbd	1ocd	1sah	2cth	3pgt
1a2j	1b6w	1c7w	1der	1ehe	1fnn	1gv3	1i0a	1jku	1lbg	1oco	1sai	2ctn	3pre
1a2l	1b70	1c7y	1df0	1ehf	1fo2	1gvh	1i0h	1jkw	1lbg	1ocp	1saj	2cts	3prg
1a2m	1b71	1c82	1dfk	1ehg	1fo3	1gw5	1i11	1jld	1lbu	1ocr	1sak	2cxb	3psr
1a2s	1b72	1c8i	1dfi	1ehj	1fo4	1gwc	1i1g	1jld	1lcc	1oct	1sal	2cy3	3rlr
1a2x	1b79	1c8j	1dg3	1ei7	1foc	1gwg	1i1r	1jli	1lcd	1ocz	1san	2cyk	3sak
1a32	1b7t	1c9b	1dgj	1eia	1foe	1gwp	1i1s	1jlt	1ld7	1odd	1sav	2cym	3sdh
1a36	1b7v	1c9i	1dgn	1eij	1fok	1gwx	1i27	1jlv	1ld8	1oel	1scf	2cyp	3sdp
1a37	1b7y	1c9l	1dgp	1eiy	1fow	1gxd	1i2t	1jlv	1lde	1ofj	1sch	2dhh	3sqc
1a38	1b80	1ca1	1dgs	1ej3	1fox	1gxx	1i36	1jmw	1ldj	1ofk	1scm	2dtr	3t4
1a3d	1b82	1cb1	1dgt	1ej5	1foy	1gxh	1i37	1jmw	1ldk	1olg	1sct	2dvh	3vhh
1a3f	1b85	1cb8	1dgu	1eji	1fp1	1gxp	1i38	1jme	1le2	1olh	1sdk	2eia	3wrp
1a3n	1b86	1cb1	1dgv	1ejy	1fp2	1gxq	1i3d	1jmr	1le4	1om2	1sdl	2end	3ygs
1a3o	1b89	1cbm	1dii	1eku	1fp3	1gy3	1i3e	1jms	1le6	1omd	1ser	2erl	451c
1a43	1b8c	1cc5	1dii	1el4	1fpo	1gyg	1i3q	1jmw	1le7	1opc	1ses	2eza	4aah
1a4f	1b8d	1cca	1dio	1elk	1fpp	1gyo	1i4a	1jmx	1le8	1or2	1set	2ezb	4bp2
1a4o	1b8f	1ccb	1diq	1elr	1fpq	1gyz	1i4s	1jmx	1lea	1or3	1sfe	2ezc	4ccp
1a4p	1b8i	1ccc	1diy	1elw	1fps	1gza	1i4y	1jni	1leb	1orc	1sfv	2ezh	4ccx
1a52	1b8l	1ccd	1diz	1em1	1fpw	1gzb	1i4z	1jnr	1lfg	1ork	1sfw	2ezi	4cln
1a56	1b8r	1cce	1djl	1em9	1fpx	1gzx	1i50	1jnz	1lfb	1osa	1sgg	2ezk	4cox
1a59	1b8x	1ccg	1dj5	1emr	1fqi	1h0n	1i54	1jow	1lfn	1otp	1sig	2ezl	4cp4
1a5j	1b8z	1ceh	1dj8	1emu	1fqj	1h0o	1i55	1joy	1lga	1out	1skn	2ezx	4cpp
1a5t	1b9l	1cci	1djg	1emv	1fqk	1h19	1i5n	1jpb	1lh1	1ouu	1skt	2ezy	4cpv
1a62	1b9a	1ccj	1djh	1emy	1fqv	1h1j	1i5t	1jpb	1lh2	1oxa	1sky	2ezz	4cro
1a63	1b9m	1cek	1dji	1en4	1fr0	1h4i	1i6d	1jpb	1lh3	1oxy	1slm	2fal	4crx
1a6d	1b9n	1cel	1djw	1en5	1fs0	1h4j	1i6e	1jpb	1lh5	1p2p	1sly	2fam	4csc
1a6e	1b9x	1ccp	1dix	1en6	1fs1	1h4l	1i6h	1jpb	1lh6	1pa2	1smg	2fap	4csm
1a6g	1b9y	1ccr	1dix	1en7	1fs2	1h4r	1i6v	1jpl	1lh7	1pal	1smt	2ffh	4cts
1a6i	1ba5	1cd3	1djz	1enh	1fs7	1h54	1i6z	1jpn	1lht	1par	1spe	2fha	4fap
1a6k	1bab	1cd9	1dk1	1eni	1fs8	1h55	1i77	1jpp	1lht	1pax	1spg	2fhe	4fis
1a6m	1baj	1cdl	1dk2	1enj	1fs9	1h57	1i7g	1jpr	1li5	1pbv	1spy	2fok	4gss
1a6n	1bal	1cdm	1dk3	1enk	1fse	1h58	1i7i	1jpw	1li7	1pbw	1sqc	2fow	4gst
1a6s	1baq	1cdn	1dk5	1enw	1fsh	1h5a	1i7w	1jpw	1lia	1pbx	1sra	2frc	4gtu
1a75	1bax	1cdp	1dk8	1eo0	1fsl	1h5c	1i7x	1jqc	1lih	1pd2	1sry	2fus	4hbi
1a76	1bay	1ce8	1dke	1eog	1fsx	1h5d	1i7y	1jqc	1lin	1pdn	1sss	2gdm	4hbb
1a77	1baz	1ced	1dkf	1eoh	1ft1	1h5e	1i8o	1jqj	1lip	1per	1swm	2giw	4icb
1a7d	1bbb	1ceh	1dl2	1eoi	1ft2	1h5f	1i8p	1jrc	1lis	1pes	1sym	2glr	4lbd
1a7e	1bbc	1cei	1dli	1eoi	1ft5	1h5g	1i8q	1jrc	1ljp	1pet	1tin	2gmf	4mba

1a7m	1bbh	1cem	1dlj	1eq1	1ft6	1h5h	1i94	1jr8	1ljr	1pfr	1tad	2gsq	4mbn
1a7v	1bb1	1cf7	1dlw	1eqf	1ft9	1h5i	1i95	1jr9	1ljw	1pge	1taf	2gsr	4p2p
1a7w	1bbn	1cfa	1dly	1eqg	1fts	1h5j	1i96	1jra	1lk3	1pgf	1tag	2gss	4pal
1a8a	1bby	1cfc	1dm1	1eqh	1ftt	1h5k	1i97	1jro	1lki	1pgg	1tam	2gst	4pax
1a8b	1bc0	1cfd	1dm5	1eqz	1ftz	1h5l	1ial	1jrp	1lkv	1pgj	1taq	2gtu	4pgt
1a8c	1bc1	1cff	1dmo	1erc	1fu0	1h5m	1iap	1jrx	1lky	1pgn	1tau	2gwx	4prc
1a8h	1bc3	1cfp	1dn1	1erd	1fup	1h67	1iar	1jry	1lll	1pgo	1tba	2hbc	4prg
1a8o	1bc5	1cg5	1dnp	1ere	1fuq	1h6g	1ib1	1jrz	1lla	1pgp	1tbg	2hbd	4r1r
1a8v	1bc7	1cg7	1dnu	1erp	1fur	1h6k	1ib2	1js4	1lli	1pgq	1tc3	2hbe	4sdh
1a9w	1bc8	1cg8	1dnw	1err	1fv0	1h6o	1ib3	1js8	1llp	1pgr	1tcf	2hbf	4tf4
1a9x	1bc9	1cgm	1dny	1ery	1fvj	1h6p	1ibe	1jsa	1lm2	1pgt	1tco	2hbg	4tnc
1aa2	1bcc	1cgn	1do1	1etc	1fvk	1h7c	1ibr	1jsp	1lm3	1pha	1tes	2hbi	4vhh
1aa4	1bcf	1cgo	1do3	1etd	1fvv	1h7i	1ich	1jst	1lmb	1phb	1tf4	2hbs	5ccp
1aa7	1ben	1cgp	1do4	1ete	1fw1	1h7w	1id3	1jsu	1lmz	1phc	1tfb	2hco	5cox
1aab	1bew	1chl	1do7	1etk	1fw4	1h7x	1idq	1jsw	1lnh	1phd	1tfr	2hdc	5cp4
1ab3	1bcy	1ch2	1dof	1eto	1fw6	1h88	1idr	1jt0	1loh	1phe	1thb	2hdd	5cpp
1abm	1bcz	1ch3	1dog	1etp	1fwz	1h89	1ids	1jt6	1lot	1phf	1tmq	2hdh	5cpv
1abs	1bdc	1ch4	1doq	1etq	1fx0	1h8a	1idu	1jtb	1lox	1phg	1tn4	2hfh	5cro
1abv	1bdd	1ch5	1dov	1etv	1fx7	1h8b	1idy	1jtx	1lpe	1phn	1tnd	2hfb	5crx
1abw	1bdh	1ch7	1dow	1etw	1fx9	1h8e	1idz	1jty	1lpq	1pir	1tnp	2hhd	5csc
1aby	1bdi	1ch9	1dp3	1etx	1fxf	1h8h	1ie8	1jum	1lq1	1pis	1tnq	2hhe	5csm
1ac1	1bdj	1chh	1dp5	1ety	1fxk	1h96	1ie9	1juo	1lqc	1pmb	1tns	2hio	5cts
1ac4	1bdt	1chi	1dp7	1eum	1fyh	1h97	1ier	1jup	1lqs	1pmt	1tnt	2hmq	5cyt
1ac8	1bdv	1chj	1dpj	1eup	1fyj	1h99	1ies	1juq	1lre	1pnb	1tnw	2hmx	5eas
1aca	1bdx	1chu	1dpr	1ev4	1fyk	1h9e	1ifl	1jus	1lri	1pnr	1tnx	2hmr	5eat
1aci	1be2	1ci4	1dps	1ev9	1fyl	1h9f	1ifa	1jvr	1lrp	1poa	1top	2hoa	5eau
1acp	1be3	1cie	1dpu	1evs	1fym	1h9g	1ifw	1jw2	1lrv	1pob	1tpt	2hp8	5fwg
1acv	1bea	1cif	1dpy	1evy	1fyz	1h9t	1ify	1jw8	1lsh	1poc	1trf	2hpd	5gss
1ad6	1bed	1cig	1dqe	1evz	1fz0	1h9u	1ig5	1jwd	1ltw	1pod	1tro	2hts	5gst
1adr	1beg	1cih	1drg	1ew6	1fz1	1h9x	1ig6	1jwe	1lus	1poe	1trr	2il6	5hbi
1adt	1bej	1cik	1ds4	1ewa	1fz2	1h9y	1ig7	1jwf	1luv	1pog	1tx4	2ilk	5mba
1adu	1bek	1cio	1dsb	1ewq	1fz3	1h9z	1ign	1jwg	1luw	1pon	1ubv	2int	5mbn
1adv	1bem	1cip	1dse	1ewr	1fz4	1ha2	1igv	1jwl	1lva	1pou	1ubw	2irf	5p2p
1ae7	1beo	1cig	1dsg	1exa	1fz5	1ha7	1igx	1jwn	1lvf	1pp2	1ubx	2isd	5pal
1aeb	1bep	1cjk	1dsh	1exe	1fz6	1hab	1igz	1jy7	1lw1	1ppa	1uby	2lbd	5prc
1aed	1beq	1cjt	1dso	1exi	1fz7	1hac	1ih0	1jyg	1lxx	1ppr	1une	2lef	5r1r
1aee	1bes	1cju	1dsp	1exj	1fz8	1hak	1ihf	1jzk	1ly8	1pra	1unk	2lfb	5tnc
1aef	1bf5	1cju	1dt0	1exn	1fz9	1hb6	1ihg	1jzl	1ly9	1prb	1uor	2lh1	6atj
1aeg	1bfa	1ck6	1dt1	1exr	1fzh	1hb8	1iie	1jzm	1lyc	1prc	1utg	2lh2	6ccp
1aeh	1bfm	1ck7	1dt6	1exx	1fzi	1hba	1iio	1jzq	1lyk	1prg	1utr	2lh3	6cox
1aei	1bfr	1ckk	1dt7	1exz	1fzp	1hbb	1iip	1jzr	1lyn	1prh	1uwo	2lh5	6cp4
1aej	1bg0	1ckt	1dti	1ey1	1g03	1hbg	1ijl	1jzs	1mlk	1pru	1uxc	2lh6	6cpp
1aek	1bg1	1cl5	1dtl	1eyh	1g08	1hbh	1ijw	1k04	1mlp	1prv	1uxd	2lh7	6cro
1aem	1bg5	1cl6	1dtm	1eys	1g09	1hbi	1ijx	1k0a	1mlq	1ps1	1vap	2lhb	6csc
1aen	1bg6	1clb	1du0	1eyv	1g0a	1hbk	1ijy	1k0b	1mlr	1psh	1var	2lig	6cts
1aen	1bg7	1clc	1du2	1eyx	1g0b	1hbm	1ijz	1k0c	1m47	1psj	1vas	2lis	6gss
1aep	1bg8	1cll	1du6	1ez3	1g0v	1hbn	1ik0	1k0d	1m48	1psr	1vew	2lir	6gst
1aeq	1bgc	1clm	1du7	1eza	1g0w	1hbo	1ik3	1k0m	1m49	1pth	1vhh	2lyn	6gsu
1aes	1bgd	1clp	1dug	1ezb	1gle	1hbr	1ik7	1k0n	1m4a	1pue	1vii	2mb5	6gsv

1aet	1bge	1cls	1duk	1ezc	1glu	1hbs	1liku	1k0o	1m4b	1pva	1vin	2mbw	6gsw
1aeu	1bgf	1cm1	1dul	1ezd	1glx	1hbu	1il0	1k0y	1m4c	1pvb	1vip	2mga	6gsx
1aev	1bgp	1cm4	1duo	1ezf	1g2h	1hbx	1il6	1k1f	1m60	1pys	1vlk	2mgb	6gsy
1aew	1bgx	1cma	1duw	1ezt	1g2n	1hc1	1ile	1k1v	1m62	1qa6	1vls	2mgc	6hbi
1af7	1bgy	1cmb	1dux	1ezv	1g2y	1hc2	1ilg	1k2h	1m6c	1qag	1vlt	2mgd	6hbw
1af8	1bh2	1cmc	1dv0	1ezy	1g2z	1hc3	1ilh	1k2o	1m6m	1qas	1vnc	2mge	6pax
1afh	1bh8	1cmf	1dv5	1f0j	1g33	1hc4	1ilk	1k3g	1m6v	1qat	1vnd	2mgf	6prc
1afr	1bh9	1cmg	1dvb	1f0m	1g39	1hc5	1imp	1k3h	1m7k	1qaz	1vne	2mgg	6rlr
1afv	1bhd	1cmj	1dve	1f0y	1g3j	1hc6	1imq	1k40	1m9i	1qb2	1vnf	2mgh	7atj
1agm	1bi0	1cmn	1dvg	1f12	1g3n	1hc8	1in4	1k4w	1ma1	1qbj	1vng	2mgi	7ccp
1agr	1bil	1cmp	1dvh	1f14	1g3s	1hch	1in5	1k5h	1ma6	1qbk	1vnh	2mgj	7cei
1ags	1bi2	1cmq	1dvi	1f17	1g3t	1hci	1in6	1k62	1mba	1qbk	1vni	2mgk	7cpp
1ah7	1bi3	1cmt	1dvk	1f1c	1g3w	1hcm	1in7	1k6o	1mbc	1qc7	1vns	2mgl	7gss
1ahd	1bia	1cmu	1dvo	1fle	1g3y	1hco	1in8	1k6y	1mbd	1qck	1vol	2mgm	7hbi
1ahr	1bib	1cmw	1dvp	1flf	1g4d	1hcr	1inz	1k74	1mbe	1qcr	1vpi	2mhb	7ice
1aie	1bij	1cmy	1dvv	1flm	1g4i	1hcu	1io3	1k78	1mbf	1qdb	1vpw	2mhr	7icf
1aii	1bin	1cmz	1dw0	1flo	1g4u	1hcy	1io7	1k79	1mbg	1qdm	1vre	2mm1	7icg
1ail	1bip	1cn4	1dw1	1fls	1g4w	1hd6	1io8	1k7a	1mbh	1qgh	1vrf	2mta	7ich
1ain	1biq	1cno	1dw2	1flz	1g4y	1hda	1io9	1k7l	1mbi	1qgi	1vtm	2mya	7ici
1aip	1bj0	1cnp	1dw3	1f22	1g50	1hdb	1iok	1k7w	1mbj	1qgk	1vwt	2myb	7iej
1ais	1bj5	1cnt	1dw9	1f24	1g59	1hdd	1iop	1k7y	1mbk	1qgp	1vxa	2myc	7ick
1aj3	1bj9	1co6	1dwk	1f25	1g5n	1hdj	1iq0	1k83	1mbn	1qgr	1vxb	2myd	7icl
1aj4	1bja	1co8	1dwr	1f26	1g5y	1hdp	1li1	1k8a	1mbo	1qgt	1vxc	2mye	7icm
1aj5	1bje	1co9	1dws	1f2e	1g5z	1hds	1li3	1k8k	1mbs	1qgw	1vxd	2mys	7icn
1aj8	1bjf	1coh	1dwt	1f36	1g6n	1he1	1li5	1k8u	1mc2	1qh4	1vxe	2new	7ico
1aj9	1bjj	1coj	1dxc	1f3a	1g6w	1he8	1liq	1k90	1mcy	1qhb	1vxf	2ng1	7icp
1ajg	1bjy	1cok	1dxd	1f3b	1g6y	1he9	1liq	1k93	1md3	1qhd	1vxg	2ngr	7icq
1ajh	1bjz	1coo	1dxr	1f43	1g72	1hek	1liq	1k94	1md4	1qi8	1vxh	2not	7icr
1ak0	1bk5	1cop	1dxs	1f44	1g73	1hf0	1liqu	1k95	1mdn	1qi9	1wad	2occ	7ics
1ak4	1bk6	1cor	1dxt	1f45	1g7d	1hf8	1liq	1k96	1mdv	1qiv	1was	2or1	7ict
1ak8	1bk9	1cot	1dxu	1f4i	1g7o	1hfa	1lir	1k98	1mdy	1qiw	1wat	2orc	7icu
1akd	1bkd	1cp0	1dxv	1f4k	1g8e	1hfe	1lir	1k99	1mfr	1qj2	1wdc	2pac	7icv
1akh	1bke	1cp4	1dxx	1f4l	1g8i	1hg2	1lir	1k9k	1mgn	1qja	1wej	2pal	7prc
1akk	1bkr	1cp5	1dy7	1f4m	1g8q	1hg4	1lir	1k9m	1mgt	1qjb	1wer	2pas	7rlr
1al0	1bl0	1cpc	1dyv	1f4n	1g9l	1hg5	1lir	1k9p	1mhl	1qjd	1wet	2paw	8cpp
1al6	1bl9	1cpd	1dz4	1f4o	1g9v	1hga	1lrj	1ka3	1mho	1qjt	1wja	2pax	8gss
1ala	1blq	1cpe	1dz6	1f4q	1ga3	1hgb	1lrl	1ka8	1mhy	1qk1	1wjb	2pcb	8ica
1al1	1bm0	1cpf	1dz8	1f4t	1gab	1hgc	1lrm	1kb0	1mhz	1qkm	1wjc	2pcc	8icb
1alu	1bm9	1cpg	1dz9	1f4u	1gah	1hgu	1lrq	1kbh	1mj2	1qln	1wjd	2pdd	8icc
1alv	1bmf	1cpo	1e0e	1f59	1gai	1hh5	1lrv	1kbu	1mj1	1qkr	1wje	2pde	8ice
1alw	1bmo	1cpq	1e0h	1f5n	1gak	1hh7	1lrw	1kcf	1mjl	1qks	1wjf	2pgd	8icf
1am4	1bmt	1cpr	1e17	1f5o	1gax	1hh8	1lrx	1lky	1mjm	1qkt	1wla	2pgg	8icg
1am9	1bn5	1cpt	1e1q	1f5p	1gbu	1hho	1lis	1kd1	1mjo	1lku	1wq1	2pgt	8ich
1amz	1bno	1cpw	1e1r	1f5q	1gbv	1hig	1lisa	1kdh	1mjp	1ql3	1wrp	2phi	8ici
1an2	1bnp	1cq2	1e28	1f5t	1gc6	1hij	1lis	1kdx	1mj1	1ql4	1wrs	2pmt	8icj
1an4	1bo9	1cqe	1e2a	1f5x	1gc7	1hik	1lisc	1kea	1mks	1qla	1wrt	2prc	8ick
1ann	1boa	1cqt	1e2r	1f63	1gcj	1hio	1lisc	1kee	1mkt	1qlb	1wtu	2prg	8icl
1anv	1boc	1cqx	1e2x	1f65	1gcv	1hiw	1lit	1kej	1mku	1qlk	1xbl	2psr	8icm
1anw	1bod	1crc	1e39	1f66	1gcw	1hjl	1lit	1key	1mkv	1ql1	1xch	2pua	8icn

lanx	lbof	lcrg	le3g	lf68	lgdd	lhj3	lit3	lkf6	lmif	lqls	lxgm	2pub	8ico
lao6	lbou	lcrh	le3h	lf6f	lgdi	lhj4	litf	lkfr	lmig	lqm6	lxgn	2puc	8icp
laoa	lbp2	lcri	le3k	lf6h	lgdj	lhj5	lith	lkfu	lmih	lqmd	lxgo	2pud	8icq
laof	lbp3	lcrj	le3m	lf6v	lgdk	lhjp	liti	lkfx	lmij	lqmg	lxgs	2pue	8icr
laoi	lbp4	lcrk	le3o	lf70	lgdl	lhjt	litk	lkfy	lmik	lqmq	lxik	2puf	8ics
laok	lbpe	lcrx	le3p	lf71	lgds	lhks	litl	lkgm	lmll	lqmz	lxol	2pug	8ict
laom	lbpo	lcry	le3y	lf7c	lgdt	lhkt	litm	lkgp	lmim	lqn0	lxpa	2pvb	8ieu
laon	lbpq	lcs0	le41	lf7u	lgdv	lhlb	lity	lkgp	lmim	lqn1	lxsm	2r1r	8icv
laog	lbpq	lcs4	le52	lf7v	lgdy	lhlm	liuf	lkgs	lmlo	lqn2	lyca	2r2f	8icw
laor	lbpq	lcs4	le52	lf7v	lgdy	lhlm	liuf	lkgs	lmlo	lqn2	lyca	2r2f	8icw
laor	lbpq	lcs4	le52	lf7v	lgdy	lhlm	liuf	lkgs	lmlo	lqn2	lyca	2r2f	8icw
laos	lbpz	lcsq	le68	lf80	lgeb	lhhr	liuh	lkil	lmir	lqnm	lycc	2ran	8icy
laow	lbq0	lcsi	le6b	lf8h	lged	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
laoy	lbq7	lcsi	le6i	lf8n	lgei	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
lap4	lbqv	lcsn	le6j	lf93	lgej	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
lap5	lbqz	lcsr	le6v	lf99	lgek	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
lap6	lbr0	lcss	le6y	lf9d	lgem	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
lapc	lbr1	lcsu	le78	lf9g	lgfi	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
lapl	lbr4	lcsv	le79	lf9n	lgg2	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
lapx	lbrw	lcsw	le7a	lf9o	lgg3	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
laqe	lbs2	lcsx	le7b	lfad	lggq	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
laqt	lbs3	lcta	le7c	lfaf	lggw	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
laqv	lbsh	lctd	le7d	lfaf	lggz	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
laqw	lbsm	lctj	le7e	lfag	lgh0	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
laqx	lbsn	lctr	le7f	lfah	lgh1	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
lar4	lbsx	lcts	le7g	lfap	lgh4	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
lar5	lbt1	lcty	le7h	lfaw	lgh6	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
larp	lbt2	lctz	le7i	lfaz	lghc	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
larq	lbt3	lcuk	le7j	lfb2	lghq	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
larr	lbt6	lcuk	le7l	lfbo	lgia	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
laru	lbt8	lcun	le7p	lfbo	lgia	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
larv	lbu2	lcun	le7p	lfbo	lgia	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
larw	lbu3	lcun	le7p	lfbo	lgia	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
larx	lbu7	lcx2	le83	lfbv	lgjm	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
lary	lbuc	lcx8	le84	lfbw	lgjn	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
las0	lbug	lca	le85	lfby	lgjq	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
las2	lbun	lcc	le86	lfce	lgjs	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
las3	lbun	lcc	le86	lfce	lgjs	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
lash	lbuy	lcx2	le8e	lfcd	lgjv	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
latj	lbv2	lcy5	le8w	lfce	lgjy	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
lau1	lbva	lcy	le8x	lfch	lgk2	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
lau7	lbvb	lcyf	le8y	lfcs	lgk3	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
laua	lbvc	lcyi	le8z	lfex	lgkj	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
laue	lbvd	lcyj	le90	lfey	lgkm	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz
laur	lbvm	lcy	le91	lfcz	lgks	lhiv	liw7	lkib	lmis	lqnn	lycq	2rig	8icz

APPENDIX B

Names of proteins in all- β class in SCOP; <http://scop.mrc-lmb.cam.ac.uk>

117e	1b4l	1c8e	1dfx	1ele	1fpv	1gvi	1i4q	1jot	1lll	1ops	1rxg	1xkb	2sim
12ca	1b4m	1c8f	1dgl	1elf	1fq3	1gvk	1i4r	1jou	1llg	1ose	1rza	1xna	2sli
12e8	1b4r	1c8g	1dgg	1elg	1fq4	1gvl	1i4u	1jp5	1lli	1osp	1rzb	1xnb	2snm
15c8	1b4t	1c8h	1dgl	1elo	1fq5	1gvm	1i4v	1jpc	1lln	1osz	1rzc	1xnc	2sns
1a02	1b55	1c8m	1dgr	1elp	1fq6	1gvp	1i4x	1jpf	1llo	1otc	1rzd	1xnd	2snv
1a0h	1b56	1c8n	1dgs	1elt	1fq7	1gvt	1i50	1jpg	1l2e	1ovw	1rze	1xnt	2snw
1a0i	1b5f	1c8p	1dgt	1elv	1fq8	1gvu	1i53	1jps	1l3j	1oxy	1sac	1xso	2sob
1a0j	1b5g	1c8q	1dgw	1emt	1fq9	1gvv	1i5f	1jpt	1l3n	1p01	1sat	1xtc	2sod
1a0l	1b6d	1c9i	1dhk	1ena	1fq1	1gvw	1i5h	1jq6	1l3w	1p02	1sbb	1xuf	2spg
1a0n	1b6j	1c9l	1dhx	1enc	1fqm	1gvx	1i5i	1jq7	1l4i	1p03	1sbd	1xug	2src
1a0q	1b6k	1c9o	1dic	1enf	1fqn	1gw0	1i5l	1jqy	1l5b	1p04	1sbe	1xuh	2sta
1a0r	1b6l	1c9p	1dif	1enq	1fqr	1gw5	1i5p	1jqz	1l5c	1p05	1sbf	1xui	2stb
1a12	1b6m	1c9s	1dil	1enr	1fqt	1gwa	1i6c	1jr0	1l5d	1p06	1sbg	1xuj	2stv
1a14	1b6n	1c9t	1dim	1ens	1fqx	1gx8	1i6h	1jr7	1l5e	1p09	1sbs	1xuk	2taa
1a18	1b6o	1c9u	1dit	1ent	1fr3	1gx9	1i6t	1jrg	1l5g	1p10	1sbw	1xyf	2tbs
1a1d	1b6p	1ca0	1diw	1enx	1fr4	1gxa	1i75	1jrh	1l6f	1p11	1scr	1xyn	2tbv
1a1m	1b6r	1ca1	1dj7	1eo2	1fr7	1gxc	1i7a	1jri	1l6g	1p12	1scs	1xyo	2tdt
1a1n	1b6s	1ca2	1djg	1eo5	1frg	1gxd	1i7p	1jrq	1l6p	1p35	1sda	1xyp	2tep
1a1o	1b6u	1ca3	1djh	1eo7	1frn	1gxr	1i7r	1jrs	1l6x	1pam	1sdy	1yai	2tga
1a1q	1b70	1ca4	1dji	1eo8	1frq	1gy1	1i7t	1jrt	1l7f	1paz	1se2	1yaz	2tgd
1a1r	1b75	1ca8	1djr	1eo9	1frt	1gy2	1i7u	1js4	1l7g	1pbo	1se3	1ycp	2tgp
1a1x	1b7a	1ca9	1djs	1eoa	1fs0	1gy9	1i7z	1js8	1l7h	1pcl	1se4	1ycs	2tgt
1a21	1b7i	1cah	1djw	1eob	1fsk	1gya	1i81	1js9	1l7i	1pcr	1seb	1yda	2thf
1a22	1b7j	1cai	1dix	1eoc	1fsn	1gyc	1i82	1jsd	1l7j	1pcs	1sem	1ydb	2tio
1a25	1b7k	1caj	1dij	1eoj	1fso	1gyd	1i85	1jsg	1l7k	1pdk	1sfi	1ydc	2tld
1a2c	1b7t	1cak	1djz	1eol	1fsq	1gye	1i8a	1jsh	1l9b	1pdr	1sfp	1ydd	2tmp
1a2d	1b7x	1cal	1dkg	1eou	1fsr	1gyg	1i8d	1jsi	1l9j	1ped	1sft	1yec	2tnf
1a2t	1b7y	1cam	1dl7	1eov	1fst	1gyh	1i8f	1jsm	1l9m	1pex	1sgc	1yed	2tod
1a2u	1b88	1can	1dlc	1ep0	1ft0	1gyu	1i8g	1jsn	1l9n	1pfc	1sgf	1yee	2tpi
1a2v	1b8a	1cao	1dle	1ep1	1ft3	1gyv	1i8h	1jso	1lab	1pfs	1sgk	1yef	2trc
1a2y	1b8e	1cau	1dlf	1ep2	1ft9	1gyw	1i8i	1jsy	1lac	1pfx	1sgp	1yeg	2trh
1a30	1b8q	1cav	1dlh	1ep3	1ftp	1gz9	1i8k	1jt3	1lay	1pfz	1sgq	1yeh	2trm
1a33	1b90	1caw	1dlk	1epa	1fu1	1gzc	1i8l	1jt4	1lb1	1pgr	1sgr	1yei	2try
1a34	1b9k	1cax	1dll	1epb	1fui	1gzh	1i8m	1jt5	1lb4	1pgs	1sgt	1yej	2tsa
1a39	1b9m	1cay	1dlm	1epf	1fuj	1gzi	1i8q	1jt7	1lb5	1phm	1shc	1yek	2tsb
1a3b	1b9n	1caz	1dlp	1ep1	1fun	1gzp	1i8u	1jt8	1lb6	1phs	1shf	1yge	2tss
1a3e	1b9s	1cb4	1dlq	1epm	1fux	1gzq	1i8z	1jta	1lck	1pht	1shg	1ygs	2tun
1a3k	1b9t	1cb8	1dlr	1epn	1fv1	1h0g	1i90	1jtc	1lcl	1pif	1shs	1yha	2ubp
1a3l	1b9v	1cbi	1dm0	1epo	1fv2	1h0i	1i91	1jtd	1lcy	1pig	1sid	1yhb	2ull
1a3q	1b9x	1cbj	1dm4	1epp	1fv3	1h19	1i92	1jto	1ld9	1pin	1sie	1ykf	2upj
1a3r	1b9y	1cbq	1dmh	1epq	1fv9	1h4g	1i94	1jtp	1lde	1piv	1sip	1yna	2vaa

1a3t	1b9z	1cbr	1dmp	1epr	1fvc	1h4h	1i95	1jtr	1ldp	1pjp	1siv	1ypp	2vab
1a3u	1ba7	1cbs	1dmr	1ept	1fvd	1h4i	1i96	1jtt	1lds	1pkl	1sky	1yso	2vgc
1a3v	1ba8	1cbv	1dms	1epw	1fve	1h4j	1i97	1jtz	1ldt	1pkm	1sla	1yst	2vir
1a3w	1ba9	1cbw	1dmx	1epz	1fvi	1h4n	1i9b	1ju3	1ldy	1pkn	1slb	1ytf	2vis
1a3x	1baf	1cc0	1dmy	1eq9	1fwa	1h4r	1i9e	1ju4	1lec	1pks	1slc	1ytg	2vit
1a3y	1bag	1cc3	1dmz	1eqd	1fwb	1h4w	1i9h	1ju9	1led	1pkt	1slid	1yth	2viu
1a3z	1bai	1ccs	1dm0	1eqr	1fwc	1h54	1i9i	1juf	1leg	1pky	1sle	1yti	2vub
1a42	1bak	1cct	1dn2	1eqw	1fwd	1h5b	1i9j	1juh	1lek	1pla	1slf	1ytj	2wbc
1a44	1bar	1ccu	1dnl	1er8	1fwe	1h5n	1i9l	1jui	1lem	1plb	1slg	1yuh	2wea
1a45	1bas	1ccz	1dnv	1erb	1fwf	1h5p	1i9m	1jv0	1len	1pic	1sli	1yyz	2web
1a46	1baw	1cd0	1do5	1erj	1fwg	1h6e	1i9n	1jv2	1lep	1plg	1sll	1zag	2wec
1a47	1bb0	1cd1	1do6	1ern	1fwh	1h6f	1i9o	1jv4	1les	1pls	1slt	1zap	2wed
1a49	1bb9	1cd3	1doa	1erx	1fwi	1h6l	1i9p	1jv5	1lf7	1plu	1slu	1zia	2wpo
1a4a	1bbd	1cd8	1doj	1es0	1fwj	1h6q	1i9q	1jvk	1lfo	1pmi	1slv	1zib	2xat
1a4b	1bbj	1cd9	1dp0	1es6	1fwq	1h6t	1i9r	1jvl	1lg5	1pmp	1slw	1znc	2xbd
1a4c	1bbp	1cdb	1dp5	1esa	1fwu	1h6u	1i9w	1jvo	1lg6	1pmr	1slx	1zsa	2znc
1a4g	1bbr	1cdc	1dpj	1esb	1fwv	1h6w	1ia5	1jw6	1lgb	1pms	1sm3	1zsb	3lbi
1a4j	1bbs	1cdg	1dpo	1esf	1fwx	1h6x	1iai	1jw7	1lgc	1pmv	1sma	1zsc	32c2
1a4k	1bbt	1cdh	1dq0	1eso	1fwy	1h6y	1iak	1jwy	1lgd	1pnc	1smd	1zxq	35c8
1a4q	1bbu	1cdi	1dq1	1est	1fwz	1h7y	1iam	1jx2	1lgn	1pnd	1sme	1zzz	3aid
1a4w	1bbw	1cdj	1dq2	1et5	1fx0	1h7z	1iao	1jxd	1lgp	1pnf	1smf	2lbi	3ait
1a57	1bbz	1cd0	1dq3	1et6	1fx5	1h80	1iar	1jxf	1lgq	1png	1smp	25c8	3ame
1a58	1bcc	1cd1	1dq4	1et7	1fx7	1h85	1iau	1jxg	1lhc	1pnj	1smr	2a2g	3app
1a5d	1bcd	1cdy	1dq5	1et8	1fxj	1h8d	1iaz	1jxj	1lhd	1poi	1smv	2a2u	3apr
1a5f	1bei	1ce1	1dq6	1et9	1fxy	1h8e	1ib0	1jxk	1lhe	1po2	1snc	2a39	3azu
1a5g	1bck	1ce5	1dqd	1eta	1fxz	1h8g	1ib4	1jxm	1lhf	1poo	1snd	2a8v	3bcc
1a5h	1bcm	1ce6	1dqg	1etb	1fyl	1h8h	1ib5	1jxo	1lhg	1pov	1snm	2aaa	3bdo
1a5i	1bco	1ce7	1dqi	1eth	1fy3	1h8i	1ib8	1jxp	1lib	1ppb	1sno	2aac	3bjl
1a5k	1bcp	1ce1	1dqj	1etj	1fy4	1h8k	1ibb	1jyn	1lic	1ppe	1snp	2aai	3blg
1a5l	1bcu	1ces	1dqk	1etr	1fy5	1h8l	1ibd	1jyt	1lid	1ppe	1snq	2abl	3bta
1a5m	1bcx	1cfl	1dql	1ets	1fy8	1h8n	1ibf	1jyv	1lie	1ppf	1sos	2afg	3btd
1a5n	1bd0	1cf8	1dqm	1ett	1fyc	1h8o	1ibg	1jyw	1lif	1ppg	1sox	2afn	3bte
1a5o	1bd2	1cfb	1dqo	1etz	1fyh	1h8s	1ibh	1jyx	1lil	1pph	1spd	2ahj	3btf
1a5u	1bd7	1cfm	1dqq	1eu1	1fyn	1h8t	1ibq	1jyy	1lj7	1ppi	1spp	2aid	3btg
1a61	1bd9	1cfm	1dqt	1eu3	1fyt	1h8v	1iby	1jyz	1ljo	1ppk	1spu	2ait	3bth
1a62	1bda	1cfq	1dqv	1eu4	1fyu	1h91	1ibz	1jz0	1lk3	1ppl	1srd	2alp	3btk
1a63	1bdl	1cfs	1dr9	1eu5	1fzj	1h92	1ic0	1jz1	1lka	1ppm	1sre	2ame	3btm
1a64	1bdo	1cft	1dro	1euf	1fzk	1h95	1ic1	1jz2	1lkb	1prc	1srf	2amg	3bto
1a65	1bdq	1cfv	1drt	1euj	1fzm	1h9d	1ic4	1jz3	1lkt	1prl	1srg	2ans	3btq
1a66	1bdr	1cgh	1dry	1euo	1fzo	1h9j	1ic5	1jz4	1lli	1prm	1srh	2ap2	3btt
1a6a	1bdx	1cgi	1ds0	1euq	1fzz	1h9k	1ic7	1jz5	1lla	1pro	1sri	2apr	3btw
1a6c	1bdy	1cgj	1ds1	1eur	1g0d	1h9l	1icm	1jz6	1llr	1prp	1srj	2aps	3ca2
1a6p	1be3	1cgp	1ds2	1eus	1g0e	1h9m	1icn	1jz7	1llw	1prs	1srl	2ara	3obs
1a6t	1be9	1cgs	1ds6	1eut	1g0f	1h9n	1ict	1jz8	1llz	1prt	1srn	2arc	3cd4
1a6u	1beb	1cgt	1ds8	1euv	1g0v	1h9p	1id2	1jze	1lm1	1psa	1sro	2asi	3cel
1a6v	1bec	1cgu	1dsf	1euw	1g0x	1h9q	1id4	1jzf	1lm8	1pse	1srp	2avi	3cgt
1a6w	1bef	1cgv	1dsl	1euy	1g0y	1h9r	1id5	1jzg	1lmk	1psf	1sta	2axm	3chb
1a6x	1beh	1cgw	1dst	1ev1	1g13	1h9s	1ida	1jzh	1lmw	1psk	1stb	2ayh	3cms
1a6z	1bev	1cgx	1dsu	1ev2	1g1c	1h9v	1idb	1jzi	1lnh	1psn	1ste	2aza	3cna

1a7l	1bex	1cgy	1dsw	1evh	1gld	1h9w	1idj	1jzj	1lmu	1ps0	1stg	2azu	3cpu
1a72	1bey	1chg	1dsy	1evt	1g1k	1h9x	1idk	1jzq	1lo0	1pss	1sth	2b3i	3cyh
1a78	1bf2	1cho	1dto	1evu	1g1o	1h9y	1ie4	1jzs	1lo2	1pst	1stm	2bat	3cys
1a7b	1bf5	1chp	1dtu	1ew3	1g1y	1ha0	1ie5	1k0h	1lo3	1pto	1stn	2bb2	3dmr
1a7h	1bf8	1chq	1du3	1ewc	1g29	1ha3	1ie7	1k0r	1lo4	1pts	1stp	2bbk	3dpa
1a7n	1bfb	1ci0	1du5	1ewh	1g2b	1ha4	1iea	1k0s	1lo6	1pvc	1str	2bbv	3eng
1a7o	1bfc	1ci3	1duc	1ewi	1g2k	1ha5	1ieb	1k0x	1loa	1pvx	1sts	2bcc	3enr
1a7p	1bfe	1ci5	1dud	1ewn	1g2l	1hag	1iec	1k12	1lob	1pwt	1sty	2bdo	3er3
1a7q	1bff	1ci6	1dum	1eww	1g2m	1hah	1ied	1k1d	1loc	1pyp	1suh	2bfh	3er5
1a7r	1bfg	1cid	1dup	1ewy	1g30	1hai	1ief	1k1i	1lod	1pys	1sva	2bfv	3est
1a7s	1bfo	1cil	1dut	1ex3	1g31	1hao	1ieg	1k1j	1loe	1pyt	1svb	2blg	3ezm
1a8d	1bfs	1cim	1duy	1ex4	1g32	1hap	1ieh	1k1l	1lof	1pza	1svc	2bn2	3f58
1a8g	1bft	1cin	1duz	1exd	1g34	1hau	1if4	1k1m	1log	1pzb	1svp	2bos	3fct
1a8j	1bfv	1ciu	1dv1	1exf	1g35	1hav	1if5	1k1n	1loh	1pzc	1swa	2bpa	3fiv
1a8k	1bg1	1ciw	1dv2	1exg	1g36	1haw	1if6	1k1o	1lom	1qa0	1swb	2bpv	3fru
1a8m	1bg9	1ciy	1dv3	1exh	1g37	1hax	1if7	1k1p	1lop	1qa1	1swc	2bpw	3gal
1a8p	1bgl	1cj5	1dv6	1exm	1g3b	1hay	1if8	1k1t	1lox	1qa2	1swd	2bpx	3gbq
1a8v	1bgm	1cjd	1dv9	1exs	1g3c	1haz	1if9	1k1u	1lp3	1qa3	1swe	2bpy	3gch
1a8z	1bgx	1cjp	1dva	1exu	1g3d	1hb0	1ifb	1k1z	1lpa	1qa7	1swf	2bpz	3gct
1a94	1bgy	1cgy	1dvf	1ey0	1g3e	1hb1	1ifc	1k21	1lpb	1qa9	1swg	2bqp	3hat
1a9b	1bhe	1ck1	1dvq	1ey2	1g3g	1hb2	1ifg	1k22	1lpl	1qab	1swh	2bsp	3hfl
1a9e	1bhg	1ck7	1dvs	1ey4	1g3p	1hb3	1ifh	1k28	1lqb	1qac	1swj	2btc	3hfm
1a9m	1bhu	1cka	1dvt	1ey5	1g3s	1hb4	1ifr	1k2b	1lqs	1qaf	1swk	2btv	3hhr
1a9v	1bhx	1ckb	1dvv	1ey6	1g3t	1hbp	1ig0	1k2c	1lr5	1qak	1swl	2bvz	3hla
1aa6	1bi0	1ckm	1dvx	1ey7	1g3w	1hbq	1ig3	1k2f	1lrh	1qal	1swn	2bza	3hmg
1aac	1bi1	1ckn	1dvy	1ey8	1g3y	1hbt	1igc	1k2i	1ls0	1qas	1swo	2ca2	3hte
1aaj	1bi2	1cko	1dvz	1ey9	1g3z	1hbw	1igf	1k2m	1ls5	1qat	1swp	2cab	3hud
1aan	1bi3	1cl3	1dw6	1eya	1g43	1hc1	1igi	1k2n	1lt3	1qau	1swq	2cam	3hvp
1aaq	1bia	1cl7	1dwb	1eyb	1g45	1hc2	1igj	1k32	1lt4	1qav	1swr	2cas	3ifb
1ab0	1bib	1clc	1dwc	1eyc	1g46	1hc3	1igm	1k3b	1lt5	1qaw	1sws	2cau	3kbp
1ab9	1bic	1clh	1dwd	1eyd	1g48	1hc4	1igo	1k3i	1lt6	1qbl	1swt	2cav	3lpr
1abi	1bih	1clo	1dwe	1eyg	1g4j	1hc5	1igp	1k3j	1lta	1qb6	1swu	2cba	3lve
1abj	1bii	1clv	1dx5	1eyl	1g4m	1hc6	1igq	1k3n	1ltb	1qb9	1sxa	2cbb	3mcg
1abo	1bil	1clw	1dxm	1eys	1g4o	1hca	1igt	1k3q	1lte	1qba	1sxb	2cbc	3mef
1abq	1bim	1cly	1dyp	1eyz	1g4r	1hcb	1igu	1k3r	1ltg	1qbb	1sxc	2cbd	3msi
1abr	1bio	1clz	1dxr	1ez1	1g51	1hcd	1igy	1k3v	1lti	1qbi	1sxn	2cbe	3msp
1ac0	1bit	1cmo	1dxw	1ez6	1g52	1hce	1ihg	1k42	1ltr	1qbl	1sxs	2cbp	3ncm
1ac6	1bj1	1cms	1dy4	1ez8	1g53	1hcf	1ihj	1k45	1lts	1qbm	1sxt	2cbr	3nla
1acb	1bj7	1cmv	1dy7	1ezg	1g54	1hcg	1ihk	1k4c	1ltt	1qbn	1sxz	2cbs	3nn9
1acd	1bj8	1cn1	1dy8	1ezl	1g5a	1hcm	1ihm	1k4d	1lu1	1qbo	1syb	2cd0	3npl
1acx	1bji	1cn3	1dy9	1ezq	1g5v	1hcv	1ihs	1k4u	1lu2	1qbr	1syc	2cel	3nuc
1acy	1bjk	1cn4	1dyk	1ezs	1g5w	1hcx	1iht	1k4z	1lu1	1qbs	1syd	2cga	3ovw
1acz	1bjm	1cnb	1dym	1ezu	1g6e	1hcy	1ihv	1k5c	1luz	1qbt	1syf	2cgp	3paz
1ad0	1bjq	1cnc	1dyn	1ezv	1g6g	1hcz	1ihw	1k5d	1lv1	1qbu	1syf	2cgr	3pca
1ad5	1bju	1cne	1dyo	1ezx	1g6l	1hd3	1iia	1k5g	1lve	1qbv	1syg	2cha	3pcb
1ad8	1bjv	1cnf	1dyp	1f00	1g6n	1hd5	1iil	1k5j	1lvk	1qcb	1tab	2chb	3pcc
1ad9	1bk0	1cng	1dyq	1f02	1g6p	1hd8	1iip	1k5m	1lvo	1qcf	1tal	2ckb	3pcd
1adb	1bk1	1cnh	1dyu	1f01	1g6r	1hdf	1iiq	1k5r	1lvy	1qcn	1taw	2clr	3pce
1adc	1bk2	1cni	1dyw	1f0r	1g6v	1hdk	1iis	1k5u	1lwh	1qco	1tbg	2cna	3pcf

ladf	lbbk	lcnj	ldyz	lf0s	lg72	lhdm	liiu	lk5v	lhwj	lqcp	ltbq	l2nd	l3pg
ladg	lbbz	lcnk	ldz0	lf0t	lg76	lhdt	liix	lk5w	l1xa	lqcr	ltbr	l2cpl	l3pch
ladl	lbl9	lcnw	ldzb	lf0u	lg77	lhdx	liiy	lk6c	l1bk	lqcs	ltbz	l2cpu	l3pci
ladq	lbla	lcnx	ldzc	lf11	lg78	lhdy	lij8	lk6l	l1ya	lqcx	ltcm	l2ctv	l3pej
ladw	lblb	lcn y	ldzd	lf13	lg79	lh dz	lij9	lk6n	l1yb	lqd0	ltcr	l2cua	l3pck
lae2	lbld	lco1	ldzj	lf15	lg7c	l1he7	lija	lk6p	l1yl	lqdc	ltcw	l2cxg	l3pcl
lae3	lbli	lco b	ldzk	lf1s	lg7h	l1he8	lije	lk6q	l1yw	lqdm	ltcx	l2cyh	l3pcm
lae5	lbln	lcon	ldzl	lf20	lg7i	l1hea	lijf	lk6t	l1mlk	lqdn	lt dt	l2dbl	l3pcn
lae6	lblr	lcov	ldzm	lf2n	lg7j	l1heb	lijq	lk6v	l1mlx	lqdo	lteh	l2dij	l3pey
lae8	lblz	lcp i	ldzp	lf2q	lg7l	l1hec	lijs	lk6w	l1m3x	lqex	ltei	l2dl2	l3pdz
lael	lbm3	lcpm	ldzq	lf2s	lg7m	l1hed	lijt	lk70	l1m43	lqez	lt en	l2dlf	l3pep
laex	lbm5	lcpn	ldzr	lf2w	lg7p	l1hef	lik3	lk83	l1m4h	lqf3	ltet	l2dli	l3phm
laey	lbm7	lcpu	ldzt	lf2x	lg7q	l1heg	lik9	lk8a	l1m4v	lqfh	ltf4	l2dmr	l3phv
laf0	lbma	lcq1	le00	lf31	lg7r	l1heh	like	lk8d	l1m56	lqff	ltfh	l2dtr	l3prc
laf9	lbmf	lcq3	le02	lf34	lg7s	l1hej	likf	lk8g	l1m57	lqfk	ltfp	l2dyn	l3pro
lafc	lbmg	lcq9	le06	lf35	lg7t	l1hes	likj	lk8h	l1m6h	lqfm	ltfx	l2efg	l3psg
lafe	lbml	lcqf	le0f	lf39	lg7y	l1het	likn	lk8i	l1m6p	lqfo	ltgb	l2eif	l3ptb
lafq	lbmm	lcqk	le0l	lf3d	lg82	l1heu	liko	lk8k	l1m6w	lqfp	ltgc	l2eip	l3ptn
lafv	lbmn	lcqq	le0n	lf3g	lg83	l1hez	likp	lk8m	l1m9u	lqfs	ltgn	l2enb	l3rdn
lag0	lbmv	lcqv	le0o	lf3j	lg84	l1hf2	likq	lk8o	l1m9w	lqft	ltgs	l2eng	l3rp2
lag4	lbmw	lcqx	le0t	lf3r	lg85	l1hf3	lii1	lk9a	l1ma0	lqfu	ltgt	l2enr	l3rpb
lag6	lbmz	lcqy	le0u	lf3t	lg86	l1hfd	lii2	lk9m	l1ma b	lqfv	ltha	l2er0	l3seb
lagb	lbn1	lc r5	le14	lf3u	lg88	l1hfu	lii e	lk9o	l1mac	lqfw	lt hc	l2er6	l3sem
lagc	lbn3	lcr7	le15	lf3v	lg8f	l1hg7	liif	lk9q	l1mae	lqfy	lt hi	l2er7	l3sga
lagd	lbn4	lcr9	le18	lf3w	lg8g	l1hg8	liir	lk9r	l1ma f	lqfz	lt hj	l2er9	l3sgb
lage	lbn8	lcra	le1a	lf3x	lg8h	l1hgd	liis	lkaa	l1mai	lqq0	lt hp	l2est	l3sil
lagf	lbnc	lcrb	le1o	lf3z	lg8j	l1hge	liit	lkab	l1maj	lqq3	lt hr	l2ezm	l3sli
lagj	lbnk	lcr m	le1q	lf41	lg8k	l1hgf	liiu	lkac	l1mak	lqq5	lt hs	l2ezn	l3sod
lagn	lbnm	lcr u	le1r	lf42	lg8l	l1hgg	lim3	lkap	l1mam	lqga	lt hu	l2f19	l3tdt
lah1	lbnn	lcrz	le1t	lf45	lg8r	l1hgh	lim9	lkaw	l1max	lqgf	lt hv	l2f3g	l3tf4
lah9	lbnq	lcs6	le22	lf49	lg8w	l1hgi	limh	lk b0	l1may	lqgl	lt hw	l2f58	l3tgi
lahj	lbnt	lcsk	le24	lf4a	lg8z	l1hgj	limc	lk b3	l1mcb	lqgz	lt ie	l2fb4	l3tgj
lahk	lbnu	lcs o	le27	lf4h	lg94	l1hgt	limd	lk b5	l1mcc	lqh0	lt ii	l2fbj	l3tgk
lahm	lb nv	lcs p	le28	lf4r	lg95	l1hh2	lime	lkbb	l1mcd	lqh6	lt io	l2fcb	l3tlh
lahs	lb nw	lcs q	le2r	lf4w	lg97	l1hh4	linf	l1kbg	l1mce	lqh7	lt it	l2fgf	l3tpi
laht	lbog	lct0	le2v	lf4x	lg9f	l1hh6	ling	l1kbk	l1mcf	lqhd	ltiu	l2fgw	l3tss
lahw	lboq	lct1	le2w	lf4y	lg9h	l1hh9	linh	l1kbp	l1mch	lqho	ltl2	l2fiv	l3ubp
lai1	lbos	lct2	le2z	lf53	lg9i	l1hhg	lino	l1kbv	l1mci	lqhp	ltld	l2fmb	l3ull
lai8	lbov	lct4	le30	lf56	lg9m	l1hhh	linq	l1kbw	l1mcj	lqh q	ltlk	l2fnt	l3upj
laid	lboy	lct8	le32	lf58	lg9n	l1hhi	linv	l1kc5	l1mck	lqhr	ltlm	l2fnb	l3vgc
laif	lbp3	lctm	le34	lf5f	lg9o	l1hhj	linw	l1kca	l1mcl	lqhu	ltmb	l2g3p	l3vub
laig	lbpl	lctn	le35	lf5j	lgaf	l1hhk	linx	l1kcc	l1mcn	lqhv	ltmc	l2gal	l3znc
laij	lbpo	lcto	le36	lf5k	lgam	l1hhn	liny	l1kod	l1mco	lqi3	ltme	l2gbq	l41bi
laip	lbpv	lcu1	le37	lf5l	lgan	l1hhp	lio4	l1ckl	l1mcp	lqi4	ltmf	l2gch	l43o9
lair	lbq5	lcu4	le38	lf5r	lgaq	l1hi6	lio6	l1ckl	l1mcq	lqi5	ltmo	l2gct	l43ca
laiw	lbqe	lcuk	le3b	lf5w	lgaw	l1hia	lio a	l1kcl	l1mcr	lqil	ltmq	l2gfb	l4aah
laix	lbqh	lcuo	le3e	lf60	lgax	l1hib	liob	l1kcs	l1mcs	lqiq	ltmt	l2gmt	l4ait
laiz	lbqk	lcur	le3f	lf6a	lgba	l1hih	lipj	l1kcu	l1mct	lqiu	ltmu	l2gn5	l4ame
laj7	lbql	leva	le3i	lf6f	lgbb	l1hii	lipk	l1ke v	l1mcw	lqix	ltnf	l2gpr	l4ape

1ajk	1bqp	1cvb	1e3j	1f6n	1gbc	1hil	1ips	1kcw	1mda	1qjl	1tng	2gva	4apr
1ajo	1bqq	1cvc	1e3l	1f6o	1gbd	1him	1ipw	1kdl	1mdc	1qj6	1tnh	2gvb	4azu
1ajv	1bqr	1cvd	1e3p	1f77	1gbe	1hin	1iqa	1kd6	1mdt	1qj7	1tni	2h1p	4bcl
1ajw	1bqs	1cve	1e3x	1f7a	1gbf	1hiv	1iqd	1kda	1mec	1qje	1tnj	2h4n	4bjl
1ajx	1bqu	1cvf	1e3z	1f7d	1gbg	1hix	1iqw	1kdb	1mel	1qjf	1tnk	2hck	4ca2
1ak4	1bqy	1cvh	1e40	1f7k	1gbh	1hj3	1ira	1kdc	1mer	1qjo	1tnl	2hft	4cac
1akj	1br1	1cvm	1e42	1f7n	1gbi	1hj4	1irp	1kde	1mes	1qjs	1tnm	2hgt	4cel
1akl	1br2	1cvn	1e43	1f7o	1gbj	1hj5	1irs	1kdf	1met	1qju	1tnn	2hla	4cgt
1akp	1br4	1cvr	1e44	1f7p	1gbk	1hj8	1isn	1kdi	1meu	1qjv	1tnr	2hlc	4cha
1aks	1br9	1cvs	1e4h	1f7q	1gbl	1hj9	1isv	1kdj	1mf2	1qjx	1toc	2hmb	4cms
1al0	1bra	1cvw	1e4j	1f7r	1gbm	1hja	1isw	1kdk	1mfa	1qjy	1tol	2hmg	4cyh
1al2	1brb	1cwa	1e4k	1f7z	1gbq	1hjb	1isx	1kdm	1mfb	1qjz	1tom	2hmi	4dmr
1alb	1brc	1cwb	1e4w	1f86	1gbr	1hjc	1isy	1kef	1mfc	1qkq	1ton	2hnt	4dpv
1aly	1bre	1cwc	1e4x	1f8a	1gbt	1hjd	1isz	1kel	1mfd	1qks	1tox	2hpe	4eng
1am2	1brp	1cwf	1e50	1f8b	1gcl	1hjf	1it0	1kem	1mfe	1qkw	1tpa	2hpf	4er1
1am5	1brq	1cwh	1e57	1f8c	1gc6	1hjh	1itb	1ken	1mfh	1qkx	1tpo	2hpp	4er2
1am6	1bru	1cwi	1e5a	1f8d	1gc7	1hjp	1iu1	1keo	1mfm	1qkz	1tpp	2hpg	4er4
1ame	1bsh	1cwj	1e5b	1f8e	1gcd	1hkl	1iuz	1keq	1mfn	1ql7	1tps	2hrp	4est
1amh	1bsi	1cwk	1e5c	1f8n	1gcf	1hla	1ivb	1kev	1mhc	1ql8	1trm	2hrv	4fab
1amm	1bsn	1cwl	1e5h	1f8t	1gcp	1hlc	1ivc	1kff	1mhe	1ql9	1trn	2hsp	4fgf
1amo	1bso	1cwm	1e5i	1f8v	1gcq	1hld	1ivd	1kfu	1mik	1qlc	1try	2hvp	4fiv
1amx	1bsq	1cwo	1e5o	1f90	1gcs	1hlt	1ive	1kfx	1mim	1qle	1ts2	2hwb	4gal
1amy	1bsy	1cwp	1e5p	1f92	1gct	1hm0	1ivf	1kg0	1mjc	1qlf	1ts3	2hwc	4gbq
1an1	1bt7	1cww	1e5r	1f97	1gcy	1hm2	1ivg	1kgl	1mjw	1qlg	1ts4	2hwd	4gch
1an8	1bth	1cx1	1e5s	1f9g	1gd7	1hm3	1ivl	1kgq	1mjx	1qlh	1ts5	2hwe	4gcr
1anb	1btk	1cx4	1e5t	1f9k	1gdf	1hm8	1ivp	1kgt	1mjj	1qlj	1tsh	2hwh	4hck
1anc	1btn	1cxe	1e5u	1fa4	1gdn	1hm9	1ivq	1kgu	1mjz	1qlr	1tsp	2i1b	4hmg
1and	1bto	1cxh	1e5v	1fai	1gdq	1hmr	1ivt	1kgw	1mkw	1qly	1tsr	2iad	4htc
1ane	1btp	1cxh	1e5w	1faj	1gdu	1hms	1ivu	1kgx	1mkx	1qm6	1tta	2iif	4hvp
1anu	1btu	1cxi	1e5y	1fak	1gen	1hmt	1ivv	1kgy	1mlb	1qmc	1ttb	2iif	4i1b
1ao5	1btw	1cxk	1e5z	1fao	1gfc	1hmu	1ivw	1khc	1mlc	1qmd	1tte	2ig2	4kbp
1ao7	1btx	1cxl	1e60	1fat	1gfd	1hmw	1ivx	1kho	1mma	1qmj	1ttf	2igf	4lve
1aof	1bty	1cyg	1e61	1fax	1gff	1hnl	1iw2	1khr	1mmd	1qmo	1ttg	2ila	4msi
1aoh	1btz	1cyn	1e62	1fay	1gg2	1hn2	1iyu	1khu	1mmg	1qmu	1ttr	2imm	4nn9
1aoj	1bul	1cyw	1e63	1fb3	1gg3	1hne	1iyv	1khy	1mmn	1qng	1ttt	2imn	4np1
1aol	1bu8	1cyx	1e64	1fb7	1gg6	1hnf	1j4k	1ki1	1mnd	1qnh	1tuc	2irt	4ovw
1aom	1bui	1cz4	1e65	1fb8	1ggb	1hng	1j4l	1kig	1mne	1qni	1tud	2isd	4paz
1aon	1buv	1cz5	1e67	1fbi	1ggc	1hnw	1j4o	1kik	1mnu	1qnj	1tui	2iza	4pcy
1aoq	1bv3	1cz8	1e6a	1fbl	1ggd	1hnx	1j4p	1kip	1mpa	1qnu	1tul	2izb	4pep
1aoz	1bv7	1czf	1e6d	1fc1	1ggi	1hny	1j4q	1kiq	1mph	1qnw	1tup	2izc	4phv
1ap2	1bv8	1czg	1e6f	1fc2	1gg1	1hnz	1j4s	1kir	1mpp	1qny	1tvd	2izd	4prc
1apn	1bv9	1czi	1e6g	1fc5	1ggr	1hoc	1j4t	1kit	1mps	1qnz	1tyu	2ize	4pro
1apt	1bve	1czm	1e6h	1fc6	1ggt	1hoe	1j4u	1kiu	1mrc	1qo3	1tyr	2izf	4rcr
1apu	1bvg	1czs	1e6j	1fc7	1ggu	1hos	1j4v	1kix	1mrd	1qoh	1tyu	2izg	4rhv
1apv	1bvk	1czt	1e6n	1fc9	1ggy	1hpc	1j58	1kj1	1mre	1qoi	1tyv	2izh	4sbv
1apw	1bvl	1czv	1e6o	1fcc	1gh5	1hpg	1j59	1kj2	1mrf	1qok	1tyw	2izi	4sga
1aq7	1bvn	1czw	1e6p	1fcf	1gh7	1hpl	1j5c	1kj3	1msa	1qoo	1tyx	2izj	4sgb
1aq8	1bvo	1czy	1e6r	1fcg	1gha	1hpo	1j5d	1kj4	1msi	1qor	1ubp	2izk	4sli
1aqb	1bvp	1czz	1e6z	1fd7	1ghb	1hps	1j5e	1kj7	1msj	1qos	1ucb	2izl	4tf4

laqc	lbvs	ld00	le79	ldi	lghf	lhpv	lj5o	lkj8	lmsp	lqot	lucy	2jcw	4thn
laqd	lbvv	ld0a	le7n	ldl	lghj	lhpx	lj71	lkj9	lmtn	lqou	luea	2jel	4tpi
laqf	lbvz	ld0g	le7u	ldo	lghk	lhql	lj7v	lkjf	lmtr	lqov	luga	2jhb	4tss
laqh	lbw3	ld0h	le7v	ldp	lgho	lhqn	lj83	lkjg	lmts	lqp1	lugb	2jia	4tsv
laqk	lbw4	ld0j	le7z	ldq	lghv	lhqp	lj84	lkjh	lmtu	lqp2	lugc	2jxr	4ubp
laqm	lbw8	ld0x	le80	ldr	lghw	lhqr	lj86	lkji	lmtv	lqp3	lugd	2kai	4ull
laqt	lbwa	ld0y	le81	le3	lghx	lhr0	lj87	lkjj	lmtw	lqpk	luge	2kau	4upj
lar1	lbwb	ld0z	le82	le8	lghy	lhri	lj88	lkjq	lmua	lqpp	lugf	2lal	4vgc
lar2	lbwm	ld1a	le8k	fej	lghz	lhrm	lj89	lkjw	lmup	lqpx	lugg	2lpr	4vub
lar6	lbwn	ld1b	le8m	fel	lgi0	lhrt	lj8a	lkjz	lmvm	lqq0	lukr	2ltm	4wbc
lar7	lbwu	ld1c	le8n	fem	lgi1	lhrv	lj8h	lkk0	lmvp	lqq1	lulo	2lve	5adh
lar8	lbww	ld1i	le8t	fen	lgi2	lhs6	lj8k	lkk1	lmwe	lqq4	lulp	2mad	5apr
lar9	lbwy	ld1k	le8u	feu	lgi3	lhsa	lj8r	lkk2	ln15	lqqd	luma	2mcg	5azu
larb	lbx0	ld1n	le8v	ff0	lgi4	lhsb	lj8s	lkk3	ln50	lqqg	lumu	2mcm	5bca
larc	lbx1	ld1s	le8w	ff5	lgi5	lhsg	lj9q	lkk4	ln90	lqqh	luok	2mcp	5ca2
lark	lbx2	ld1t	le8x	fff	lgi6	lhsh	lj9r	lkk5	lnbc	lqqj	lupj	2mev	5cac
las6	lbxa	ld2b	le8y	ffi	lgi7	lhsi	lj9s	lkk6	lnbm	lqqk	lure	2mfn	5cel
las7	lbxb	ld2e	le8z	ffk	lgi8	lhso	lj9t	lkke	lnbv	lqq1	luri	2mha	5cgt
las8	lbxo	ld2o	le90	ffr	lgi9	lhsq	lj9w	lkl3	lnca	lqqp	lurt	2mib	5cha
lasj	lbxq	ld2p	le9g	fft	lgic	lhsz	lj9z	lkl4	lncb	lqqz	luvs	2mip	5cna
laso	lbxt	ld2q	le9o	ffy	lgig	lht0	lja0	lkl5	lncc	lqqu	luvt	2mll	5cyh
lasp	lbxu	ld2s	le9p	fg2	lgj4	lhtb	lja1	lkl9	lnod	lqr1	luvu	2mpa	5er1
lasq	lbxv	ld2u	le9q	fg6	lgj5	lhte	ljab	lklf	lncg	lqr3	lvac	2msi	5er2
lasy	lbxx	ld3b	le9y	fg8	lgj6	lhtf	ljac	lklg	lnch	lqr4	lvad	2msj	5est
lasz	lbxz	ld3c	le9z	fg9	lgj7	lhtg	ljac	lklt	lncl	lqrb	lval	2msp	5fiv
lat0	byh	ld3d	lea0	fga	lgj8	lhtl	ljb0	lklu	lncl	lqrc	lvam	2mta	5gal
lat3	bybm	ld3l	lea9	fgb	lgj9	lhtp	ljb3	lkmh	lnco	lqre	lvba	2mys	5gch
lau8	bybn	ld3p	leag	fgc	lgja	lhtr	ljb7	lkn2	lnct	lqrf	lvbb	2nap	5gds
lauj	bybo	ld3q	leah	fgl	lgjb	lhtr	ljb9	lkn3	lncl	lqrg	lvbc	2ncm	5hck
laun	bybp	ld3s	leai	fgm	lgjc	lhug	ljb9	lkn4	lncl	lqrk	lvbd	2nfr	5hmg
laut	bybt	ld3t	leaj	fgn	lgjd	lhuh	ljbh	lkn9	lncl	lqrl	lvbe	2nmb	5hvp
laui	bybz	ld4h	leal	fgo	lgji	lhuj	ljbj	lknb	lncl	lqrm	lvbs	2npl	5ilb
laui	bybz	ld4i	leap	fgp	lgjq	lhuk	ljbv	lkn1	lncl	lqrm	lvbt	2nrd	5lpr
lav4	lbz9	ld4j	lear	fgq	lgjr	lhut	ljb7	lknm	lncl	lqrp	lvca	2nuc	5lve
lava	lbzd	ld4k	leas	fgr	lgju	lhv0	ljch	lkn1	lncl	lqrs	lvcb	2nul	5msi
lavb	lbze	ld4l	leat	fgt	lgjw	lhv1	ljck	lkn0	lncl	lqrt	lvcp	2occ	5nn9
lavd	lbzm	ld4m	leau	fgu	lgjx	lhv7	ljcv	lkoa	lncl	lqrq	lvcp	2ohx	5nuc
lave	lbzo	ld4p	leaw	fgv	lgka	lhv9	ljcz	lkoi	lncl	lqrw	lvde	2ovw	5paz
lavf	lbzq	ld4s	leax	fgy	lgkb	lhva	ljd0	lkop	lncl	lqrx	lvfa	2oxi	5pcy
lavg	lbzw	ld4v	leaz	fgz	lgke	lhvc	ljda	lkoq	lncl	lqrz	lvfb	2pab	5pep
lavk	lbzx	ld4y	leb0	fth2	lgkh	lhvh	ljdc	lkl1	lncl	lqs8	lvge	2pcd	5prc
lavl	lc01	ld5b	leb1	fth5	lgko	lhvi	ljdd	lkl2	lncl	lqsc	lvge	2pcp	5ptp
lavn	lc08	ld5g	leb2	fthg	lgkp	lhvj	lje5	lkl5	lncl	lqse	lvhp	2pcy	5sga
lavu	lc0a	ld5i	leba	fthn	lgkq	lhvk	lje6	lklq	lncl	lqsf	lvie	2pdz	5tss
lavw	lc0m	ld5m	lebk	ftho	lgkr	lhvl	ljec	lklf	lncl	lqsv	lvif	2pec	5tsw
lavx	lc0w	ld5r	lebp	fthq	lgkt	lhvr	ljed	lklg	lncl	lqsz	lvij	2pel	5tr
lavz	lc12	ld5x	lebw	fthr	lglo	lhvs	ljee	lklr	lncl	lqtf	lvik	2phl	5upj
law7	lc16	ld5z	leby	fthw	lgll	lhvx	ljeg	lklq	lncl	lqtp	lvit	2pia	6adh
law8	lc1a	ld6e	lebz	fthx	lgll	lhv5	ljer	lklw	lncl	lqtq	lviv	2pka	6ame

lawe	lc1e	ld6k	lec0	lf12	lg15	lhwg	ljf1	lkqx	lnif	lqts	lvjs	2plt	6apr
lawf	lc1f	ld6r	lec1	lf18	lg1a	lhwh	ljf5	lkra	lnin	lqtt	lvkx	2plv	6ca2
lawh	lc1l	ld6u	lec2	lf1e	lg1b	lhwm	ljf6	lkrb	lnir	lqtu	lvln	2pni	6cel
lawj	lc1m	ld6v	lec3	lf1g	lg1c	lhwn	ljfd	lkrc	lniv	lqty	lvlx	2poo	6cgt
lawo	lc1n	ld6w	lec4	lf1v	lg1d	lhwo	ljfh	lkrh	lnkr	lqu0	lvmo	2prc	6cha
lawq	lc1o	ld6y	lec5	lf1w	lg1e	lhwp	ljfq	lkri	lnld	lqu2	lvom	2prd	6est
lawr	lc1p	ld6z	led3	lf1z	lg1h	lhwr	ljg9	lkrr	lnlo	lqu3	lvpn	2psg	6fab
laws	lc1q	ld7b	led7	lf1j	lg1l	lhx0	ljgi	lkrs	lnlp	lqu4	lvps	2ptc	6fiv
lawt	lc1r	ld7c	ledh	lf1g	lg1m	lhx5	ljgl	lkrt	lnlr	lqu5	lvqa	2ptk	6gch
lawu	lc1s	ld7d	ledq	lf1j	lg1m	lhx6	ljgu	lkru	lnls	lque	lvqb	2ptn	6ilb
lawv	lc1t	ld7f	ledy	lf1r	lg1m	lhxb	ljgv	lkrv	lnma	lquf	lvqc	2qil	6lpr
laww	lc1u	ld7k	lee2	lf1s	lg1m	lhxd	ljgw	lksa	lnmb	lqun	lvqd	2qwa	6msi
lawx	lc1v	ld7p	lee6	lf1n	lg1m	lhxe	ljgx	lksq	lnmc	lqup	lvqe	2qwb	6nn9
lax0	lc1w	ld7q	leed	lf10	lg1m	lhxf	ljgy	lksh	lnmf	lquq	lvqf	2qwc	6paz
lax1	lc28	ld8l	leef	lf1l	lg1m	lhxm	ljgz	lksi	lnmg	lqur	lvqg	2qwd	6pcy
lax2	lc2d	ld8t	leei	lf13	lg1m	lhxn	ljh0	lksj	lnn2	lqvc	lvqh	2qwe	6prc
lax3	lc2e	ld9i	leeq	lf15	lg1m	lhxr	ljh2	lksn	lnna	lqwe	lvqi	2qwf	6taa
laxa	lc2f	ld9k	leer	lf16	lg1m	lhxs	ljh5	lksr	lnnb	lqwf	lvqj	2qwg	6upj
laxe	lc2g	ldab	leeu	lf1c	lg1m	lhxx	ljhc	lksw	lnnc	lr08	lvrl	2qwh	7adh
laxg	lc2h	ldan	lef0	lf1e	lg1n	lhxy	ljhd	lkt2	lnno	lr09	lvrh	2qwi	7ame
laxi	lc2i	ldar	lef1	lf1g	lg1n	lhy1	ljhe	lktb	lnoa	lr1a	lvsc	2qwj	7ca2
laxj	lc2j	ldaz	lef2	lf1h	lg1n	lhyo	ljhf	lktc	lnob	lr24	lvub	2qwk	7cel
laxk	lc2k	ldb7	lef3	lf1j	lg1n	lhyr	ljhh	lktf	lnol	lram	lvwa	2r04	7cgt
laxm	lc2l	ldb8	lef4	lf1k	lg1n	lhys	ljhj	lktj	lnov	lrav	lvwb	2r06	7est
laxs	lc2m	ldb9	lef5	lf1l	lg1n	lhyx	ljhk	lktk	lnp1	lray	lvwc	2r07	7fab
laxt	lc39	ldb1	lef6	lf1m	lg1o	lhyy	ljhl	lktv	lnp4	lrax	lvwd	2rac	7gch
laxy	lc3g	ldb2	lef7	lf1r	lg1o	lhz9	ljhn	lktv	lnp1	lrbp	lvwe	2ram	7hvp
laxz	lc3k	dbc	lef8	lf1t	lg1o	lhz1	ljht	lktv	lnpm	lrcy	lvwf	2rcr	7ilb
lay1	lc3m	ldb3	lef9	lf1b	lg1o	lhz2	ljil	lku8	lnpo	lred	lvwg	2rcs	7kme
lay6	lc3n	dbh	lef0	lf1d	lg1o	lhz3	ljiz	lkuj	lnps	lree	lvwh	2ren	7lpr
lay9	lc48	dbj	lef1	lf1g	lg1o	lhz4	ljiz	lkul	lnqb	lref	lvwi	2rhe	7msi
laym	lc4q	dbk	lef2	lf1n	lg1o	lhz5	ljib	lkum	lnrm	lrei	lvwj	2rm2	7nn9
layn	lc4r	dbm	lef3	lf1m	lg1p	lhz6	ljim	lkv3	lnro	lrfr	lvwk	2rma	7odc
layo	lc4u	dbn	lef4	lf1t	lg1p	lhz7	ljir	lkv7	lnrp	lrfs	lvwl	2rmb	7paz
layr	lc4v	dbo	lef5	lf1u	lg1p	lhz8	ljiv	lkv9	lnrq	lrfs	lvwm	2rmc	7pcy
laz5	lc4y	dc9	lef6	lf1v	lg1p	lhz9	ljiz	lkvp	lnrr	lrhi	lvwn	2rmp	7prc
laz8	lc57	dca	lef7	lf1w	lg1p	lhz0	ljic	lkwa	lnrs	lrho	lvwo	2rmu	7taa
lazb	lc5b	dc1	lef8	lf1x	lg1p	lhz1	ljic	lkxa	lns3	lrir	lvwp	2rox	7upj
lacz	lc5c	dce	lef9	lf1z	lg1p	lhz2	ljij	lkxb	lnsb	lrin	lvwq	2roy	8adh
lazd	lc5d	dcl	leg0	lf10	lg1p	lhz3	ljju	lkxc	lnsc	lrir	lvwr	2rr1	8ame
laze	lc5e	dcs	leg1	lf14	lg1p	lhz4	ljic	lkxd	lnsd	lrkr	lvzv	2rs1	8ca2
lazg	lc5f	dcz	leg2	lf16	lg1p	lhz5	ljic	lkxe	lnsn	lriz	lvap	2rs3	8cgt
lazm	lc5h	ddl	leg3	lf18	lg1p	lhz6	ljil	lkxf	lntd	lrir	lvba	2rs5	8est
lazn	lc5i	dd2	leg4	lf1a	lg1p	lhz7	ljil	lkxg	lntp	lrir	lvbc	2rsp	8fab
lazr	lc5k	ddg	leg5	lf1b	lg1q	lhz8	ljil	lkxh	lnuc	lrir	lvbf	2rta	8gch
lazu	lc5l	ddh	leg6	lf1c	lg1q	lhz9	ljil	lkxl	lnuk	lrir	lvbl	2rtb	8hvp
lazv	lc5m	ddi	leg7	lf1d	lg1q	lhz0	ljil	lkxq	lnwo	lrir	lvew	2rtc	8ilb
lazz	lc5n	ddj	leg8	lf1e	lg1q	lhz1	ljil	lkxt	lnwp	lrir	lvgi	2rtid	8kme
lb07	lc5o	ddl	leg9	lf1f	lg1q	lhz2	ljil	lkxv	lnyf	lrir	lvgh	2rte	8lpr

1b09	1c5p	1ddm	1eii	1fng	1gqw	1i1f	1jmc	1ky6	1nyg	1rml	1whi	2rtf	8msi
1b0e	1c5q	1ddt	1eio	1fnh	1gr3	1i1j	1jmm	1ky7	1nzt	1rmu	1who	2rtg	8paz
1b0f	1c5r	1ddv	1eiy	1fni	1gr7	1i1r	1jmq	1ky9	1oac	1rne	1whp	2rth	8prk
1b0g	1c5s	1ddw	1ej8	1fnl	1gri	1i1y	1jmx	1kya	1oak	1rpl	1wio	2rti	9ame
1b0i	1c5t	1de4	1eja	1fnm	1grw	1i2d	1jnz	1kyd	1obp	1rrp	1wip	2rtj	9ca2
1b0o	1c5u	1de7	1eje	1fnp	1gs6	1i2h	1jnh	1kyf	1obw	1rst	1wiq	2rtk	9cgt
1b0r	1c5v	1ded	1ejf	1fnq	1gs7	1i2m	1jnh	1kyn	1oca	1rsu	1wit	2rtl	9est
1b0w	1c5w	1dee	1ejm	1fns	1gs8	1i3l	1jnj	1kyo	1occ	1rsy	1wiu	2rtm	9hvp
1b1l	1c5x	1deh	1ejn	1fnu	1gs9	1i3g	1jnl	1kyu	1oco	1rtf	1wkt	2rtn	9ilb
1b12	1c5y	1dev	1ejo	1fnv	1gs1	1i3h	1jnn	1kz7	1ocr	1rtg	1wpo	2рто	9lpr
1b1z	1c5z	1df0	1ejr	1fnw	1gsm	1i3p	1jnp	1kzg	1ocz	1ruc	1wtl	2rtp	9msi
1b23	1c6v	1df3	1ejs	1fny	1gso	1i3q	1jnw	1kzk	1odw	1rud	1wwa	2rtq	9pcy
1b2p	1c6x	1df8	1ejt	1fnz	1gtf	1i3r	1jny	1kzq	1odx	1rue	1wwb	2rtr	
1b2r	1c6y	1df9	1eju	1fo0	1gtm	1i3s	1jo8	1kzy	1ody	1ruf	1wwc	2sam	
1b2w	1c6z	1dfa	1ejv	1fod	1gtr	1i3u	1job	1kzz	1ohr	1rug	1www	2sba	
1b2y	1c70	1dfb	1ek3	1foe	1gts	1i3v	1jod	1l0a	1okl	1ruh	1wyk	2sbl	
1b34	1c7s	1dfc	1ekb	1fon	1gug	1i40	1jof	1l0p	1okm	1rui	1x1l	2seb	
1b35	1c7t	1dfk	1ekl	1for	1gun	1i4e	1joi	1l0s	1okn	1ruj	1xat	2sem	
1b3i	1c7y	1dfl	1ekm	1fp5	1guo	1i4f	1joj	1l0w	1ona	1run	1xbd	2sfa	
1b3j	1c82	1dfp	1ela	1fpc	1gus	1i4g	1jok	1l0x	1opa	1ruo	1xbr	2sfp	
1b3q	1c89	1dfq	1elb	1fph	1gut	1i4h	1joo	1l0y	1opb	1rux	1xca	2sga	
1b44	1c8a	1dfu	1elc	1fpn	1gv9	1i4k	1joq	1l0z	1opg	1rvf	1xdt	2sgp	
1b4j	1c8d	1dfv	1eld	1fpt	1gvh	1i4p	1jor	1l1c	1opm	1rxf	1xca	2sil	

APPENDIX C

Names of proteins in α/β class in SCOP; <http://scop.mrc-lmb.cam.ac.uk>

10gs	1b1a	1c86	1de6	1e9d	1f9v	1gef	1hnk	1j8x	1kbl	1lx7	1qmi	1vom	2sec
10mh	1b1c	1c87	1dea	1e9e	1f9w	1geg	1hnn	1j8y	1kbo	1lxt	1qml	1vot	2sfp
11gs	1b1h	1c88	1ded	1e9f	1fa2	1geh	1hno	1j90	1kbq	1ly1	1qmp	1vp3	2shk
121p	1b1x	1c8b	1deh	1e9i	1fa9	1geq	1hnu	1j92	1kbz	1ly3	1qmu	1vp9	2shp
12gs	1b1y	1c8k	1dej	1e9l	1faa	1ger	1hmv	1j93	1kc0	1ly4	1qmv	1vpe	2sic
13gs	1b23	1c8l	1dek	1e9r	1fb0	1ges	1hnw	1j94	1kc1	1lyv	1qnf	1vpt	2skc
13pk	1b26	1c8q	1del	1e9s	1fb3	1get	1hnx	1j96	1kc3	1lz0	1qnl	1vpw	2skd
14gs	1b27	1c8v	1deo	1e9y	1fb6	1geu	1hny	1j97	1kc6	1lz7	1qno	1vrt	2ske
16gs	1b2h	1c8x	1der	1e9z	1fba	1gew	1hnz	1j99	1kc7	1lzi	1qnp	1vru	2sni
16pk	1b2l	1c8y	1dex	1ea0	1fbn	1gex	1ho1	1j9b	1kef	1ljz	1qnq	1vsb	2ssp
17gs	1b2r	1c90	1df7	1ea5	1fbt	1gey	1ho3	1j9e	1kck	1m0d	1qnr	1vsd	2st1
18gs	1b2s	1c91	1dfg	1ea7	1fc0	1gfi	1ho4	1j9g	1kcl	1m0z	1qns	1vse	2taa
19gs	1b2u	1c92	1dfh	1ea9	1fc4	1gfs	1hon	1j9j	1kcz	1m10	1qnt	1vsf	2tec
1a04	1b2y	1c93	1dfi	1eaa	1fc5	1gfw	1hoo	1j9k	1kd0	1m1b	1qnv	1vsh	2thi
1a05	1b30	1c96	1dfj	1eab	1fc6	1gfy	1hop	1j9l	1kd1	1m1k	1qo0	1vsi	2tir
1a0c	1b31	1c97	1dfk	1eac	1fc7	1gfz	1hor	1j9y	1kdo	1m1m	1qo2	1vsj	2tmd
1a0d	1b32	1c9c	1dfi	1ead	1fc9	1gg0	1hot	1j9z	1kdp	1m1u	1qo5	1vsk	2tmg
1a0e	1b37	1c9d	1dfm	1eae	1fcb	1gg1	1hox	1ja0	1kdr	1m1x	1qo7	1vsl	2tmk
1a0f	1b3b	1c9e	1dfo	1eaf	1fcd	1gg2	1hoz	1ja1	1kdt	1m54	1qo8	1vsm	2tmy
1a0o	1b3e	1c9j	1dgl	1eam	1fcf	1gg4	1hp0	1ja9	1kdv	1m6b	1qo9	1vsr	2tod
1a0r	1b3f	1c9k	1dg3	1eav	1fcj	1gg5	1hp4	1jae	1kdy	1m6h	1qon	1vtr	2tpl
1a16	1b3g	1c9m	1dg5	1eay	1fck	1gg8	1hp5	1jag	1kdz	1m6i	1qop	1vxo	2tpr
1a19	1b3h	1c9n	1dg7	1eb3	1fcq	1gg9	1hpb	1jah	1ke1	1m6v	1qoq	1vxr	2tps
1a1s	1b3l	1c9w	1dg8	1eb8	1fcu	1gga	1hpl	1jai	1ke2	1m6w	1qor	1wab	2tpt
1a1v	1b3n	1c9y	1dg9	1eb9	1fcv	1gge	1hpm	1jak	1kee	1m7b	1qox	1waf	2trc
1a23	1b3o	1cb0	1dga	1ebb	1fdi	1ggf	1hpz	1jaw	1kek	1ma0	1qoz	1waj	2trs
1a24	1b3r	1cb2	1dgd	1ebd	1fdj	1ggg	1hq5	1jax	1kep	1maa	1qp0	1wal	2trx
1a27	1b3s	1cb6	1dge	1ebf	1fdo	1ggh	1hqa	1jay	1ker	1mab	1qp4	1wdn	2tsl
1a2b	1b3v	1cb7	1dgk	1ebg	1fdr	1ggi	1hqc	1jb1	1ket	1mah	1qp7	1wet	2tsy
1a2j	1b3w	1cbf	1dgm	1ebh	1fds	1gk	1hq	1jbe	1keu	1map	1qp8	1whs	2tys
1a2k	1b3x	1cbg	1dgq	1ebi	1fdt	1ggm	1hqe	1jbe	1kev	1maq	1qpb	1wht	2uag
1a2l	1b3y	1cbo	1dgt	1ebu	1fdu	1ggm	1hqf	1jbe	1kew	1MAR	1qpg	1wkd	2ubp
1a2m	1b3z	1cbu	1dhf	1ec7	1fdv	1ggo	1hqg	1jbq	1kez	1mas	1qpk	1wke	2udp
1a2o	1b40	1cbx	1dhi	1ec8	1fdw	1ggv	1hqh	1jbv	1kf0	1md3	1qpn	1wkf	2ukd
1a2q	1b41	1cc0	1dhj	1ec9	1fdy	1gh2	1hqk	1jbw	1kf6	1md4	1qpo	1wod	2ung
1a2z	1b42	1cc2	1dhk	1ecb	1fdz	1gho	1hqo	1jce	1kfd	1mdi	1qpq	1wq1	2vp3
1a3c	1b43	1cc4	1dhp	1ecc	1fe8	1ghr	1hqs	1jcf	1kfi	1mdj	1qpr	1wsa	2vtk
1a3h	1b46	1cc6	1dhr	1ece	1fea	1ghs	1hqt	1jcg	1kfl	1mdk	1qps	1xaa	2wsy
1a3w	1b48	1ccw	1dhs	1ecf	1feb	1ght	1hqu	1jce	1kfq	1mdl	1qpz	1xab	2xim
1a3x	1b4d	1cd2	1dht	1ecg	1fec	1gia	1hqx	1jcl	1kfs	1mdp	1qq2	1xac	2xin
1a40	1b4e	1cd5	1di0	1ecj	1feh	1gil	1hgy	1jem	1kfy	1mdq	1qq5	1xad	2xis

1a47	1b4h	1cdd	1di6	1ecp	1fez	1gim	1hr0	1jct	1kgd	1mdr	1qq6	1xan	2xyl
1a49	1b4k	1cde	1di7	1ecq	1ff9	1gin	1hrd	1jcx	1kgs	1mee	1qq7	1xas	2yas
1a4i	1b4q	1cdg	1dia	1ecv	1ffa	1git	1hrh	1jcy	1kgu	1mek	1qq9	1xel	2ypi
1a4l	1b4u	1cdo	1dib	1ecx	1ffb	1gjp	1hrk	1jda	1kgw	1mg1	1qqa	1xib	2ypn
1a4m	1b4v	1cdz	1did	1ed8	1ffc	1gjr	1hse	1jdb	1kgx	1mgt	1qqb	1xic	3a3h
1a4r	1b4x	1ce2	1die	1ed9	1ffd	1gju	1hsj	1jdc	1kh1	1mhl	1qqc	1xid	3aat
1a4s	1b4z	1ce8	1dig	1edb	1ffe	1gjw	1hsl	1jdd	1kh2	1mht	1qqm	1xie	3adk
1a4u	1b51	1cec	1dih	1edd	1ffg	1gk8	1hso	1jde	1kh4	1mio	1qqn	1xif	3aky
1a4x	1b52	1cee	1dik	1ede	1ffh	1gki	1hsz	1jdf	1kh5	1mjh	1qqo	1xig	3amv
1a4y	1b54	1cen	1din	1edg	1ffk	1gkk	1ht0	1jdi	1kh7	1mkp	1qqt	1xih	3at1
1a4z	1b57	1ceo	1dio	1edo	1ffr	1gkl	1ht1	1jdj	1kh9	1ml1	1qr2	1xii	3bam
1a50	1b58	1ceq	1dir	1edq	1ffs	1gkp	1ht2	1jdn	1khh	1mla	1qr6	1xij	3bdp
1a53	1b5h	1cer	1dis	1edt	1ffw	1gkq	1ht3	1jdp	1khe	1mld	1qr7	1xim	3bif
1a54	1b5i	1cet	1diu	1edz	1ffy	1gkr	1htb	1jds	1khf	1mma	1qra	1xin	3bto
1a55	1b5j	1cev	1dj2	1ee0	1fg3	1gku	1hti	1jdt	1khg	1mmd	1qrd	1xis	3cbh
1a5a	1b5q	1cex	1dj3	1ee1	1fg4	1gky	1htt	1jdu	1khh	1mmg	1qrh	1xjo	3cd2
1a5b	1b5s	1cey	1dj9	1ee2	1fg5	1gl3	1htw	1jdv	1khj	1mmn	1qri	1xkj	3cev
1a5c	1b5t	1cf2	1dje	1ee3	1fg7	1gl6	1hu4	1jdy	1khk	1mnd	1qrq	1xla	3cgt
1a5k	1b6g	1cf3	1djg	1ee9	1fgg	1gl7	1huq	1jdz	1khl	1mne	1qrr	1xlb	3chy
1a5l	1b6h	1cf4	1djh	1eea	1fgh	1gl8	1hur	1je0	1khn	1mns	1qrs	1xlc	3cla
1a5m	1b6r	1cf9	1dji	1eeh	1fgs	1gl9	1huv	1jel	1kil	1moq	1qrt	1xld	3cox
1a5n	1b6s	1cfj	1djl	1eej	1fgx	1gla	1hux	1jea	1ki2	1mor	1qru	1xle	3cpa
1a5o	1b6t	1cfr	1djm	1eem	1fh7	1glb	1hv8	1jec	1ki3	1mos	1qs0	1xlf	3cpu
1a5s	1b73	1cfz	1djn	1een	1fh8	1glc	1hv9	1jed	1ki4	1mpb	1qs4	1xlg	3csu
1a5t	1b74	1cg0	1djo	1eeo	1fh9	1gld	1hvx	1jee	1ki6	1mpc	1qsg	1xlh	3dap
1a5u	1b76	1cg1	1djp	1eep	1fhd	1gle	1hvu	1jeh	1ki7	1mpd	1qsl	1xli	3dbv
1a5v	1b78	1cg2	1djg	1ees	1fhe	1glf	1hvx	1jej	1ki8	1mpt	1qss	1xlj	3dfr
1a5w	1b7b	1cg3	1dju	1eet	1fhu	1glg	1hw6	1jeo	1kic	1mra	1qsy	1xlk	3dhe
1a5x	1b7e	1cg4	1djw	1eex	1fhv	1glj	1hwx	1jet	1kid	1mrp	1qtl	1xll	3dmr
1a5y	1b7g	1cg6	1djx	1ef2	1fiu	1gll	1hwy	1jeu	1kie	1mss	1qti	1xlm	3drc
1a5z	1b7h	1cgk	1djj	1ef3	1fiy	1ghn	1hwz	1jev	1kif	1muc	1qtm	1xol	3eca
1a69	1b7t	1cgt	1djz	1ef8	1fj2	1glp	1hx0	1jez	1kim	1mug	1qtn	1xoa	3enl
1a6d	1b7u	1cgu	1dk7	1ef9	1fj4	1glq	1hx1	1jff	1kj8	1myr	1qtq	1xob	3eug
1a6e	1b7z	1cgv	1dka	1efa	1fj8	1glv	1hx7	1jff	1kj9	1n2c	1qtr	1xva	3eza
1a71	1b8g	1cgw	1dkd	1efc	1fjg	1gm5	1hy3	1jff	1kji	1nah	1qtw	1xwl	3ezb
1a72	1b8j	1cgx	1dkg	1efd	1fjh	1gml	1hyb	1jff	1kjj	1nai	1qu2	1xya	3eze
1a76	1b8n	1cgy	1dkl	1efg	1fjx	1gmx	1hye	1jff	1kja	1nal	1qu3	1xyb	3fua
1a77	1b8o	1cgz	1dkm	1efh	1fk8	1gn0	1hyg	1jff	1kjq	1nar	1qu4	1xyc	3fx2
1a79	1b8p	1ch8	1dkn	1efk	1fk9	1gn1	1hyh	1jff	1kja	1nas	1que	1xyf	3fyg
1a7a	1b8s	1chd	1dko	1efl	1fko	1gn8	1hyq	1jff	1kja	1nat	1quf	1xyl	3gar
1a7j	1b8u	1chm	1dkp	1efm	1fkp	1gnd	1hys	1jff	1kja	1nba	1qui	1xym	3gbp
1a7k	1b8v	1chn	1dkq	1efp	1fkx	1gne	1hyu	1jff	1kja	1nbe	1quj	1xys	3gpb
1a7l	1b8x	1chr	1dkr	1efr	1fkx	1gnp	1hyv	1jff	1kka	1nbm	1quk	1xyz	3gpd
1a7u	1b90	1chu	1dku	1eft	1fl2	1gnq	1hyz	1jff	1kka	1nda	1qul	1xza	3grs
1a80	1b92	1chw	1dkw	1efu	1fl9	1gnr	1hzd	1jff	1kka	1ndh	1qum	1xzb	3grt
1a82	1b93	1ci1	1dl3	1efv	1fla	1gns	1hzj	1jff	1kka	1nel	1quw	1xzc	3grx
1a88	1b94	1cia	1dl5	1efz	1flc	1gnv	1hzm	1jff	1kka	1nfp	1qva	1xzd	3gsb
1a8e	1b95	1cib	1dla	1eg2	1fld	1gnw	1hzp	1jff	1kka	1nft	1qvb	1xze	3gss
1a8f	1b96	1cij	1dli	1eg5	1fln	1goa	1hzy	1jff	1kka	1ngl	1rlr	1xzf	3gst

1a8h	1b97	1cip	1dlj	1eg7	1flv	1gob	1hzz	1jg4	1kk1	1nga	1ra1	1xzg	3gtu
1a8i	1b9b	1ciu	1dlo	1ega	1flz	1goc	1i01	1jg6	1kkm	1ngb	1ra2	1xzh	3had
1a8l	1b9d	1civ	1dlr	1egh	1fmc	1goi	1i0b	1jg7	1kko	1ngc	1ra3	1xzi	3hdh
1a8p	1b9f	1cj0	1dls	1egm	1fmf	1goj	1i0d	1jg8	1kkp	1ngd	1ra8	1xzy	3hsc
1a8q	1b9h	1cj2	1dlu	1ego	1fmj	1gok	1i0i	1jg9	1kkp	1nge	1ra9	1xzk	3hud
1a8s	1b9i	1cj3	1dlv	1egq	1fml	1gom	1i0l	1jgi	1kku	1ngf	1raa	1xzl	3hvt
1a8u	1b9j	1cj4	1dm3	1egr	1fmt	1goo	1i0z	1jgm	1kl1	1ngg	1rab	1xzm	3icd
1a8y	1b9x	1cjb	1dmb	1egv	1fmv	1goq	1i10	1jgt	1kl2	1ngh	1rac	1yaa	3kar
1a95	1b9y	1cjc	1dmg	1egz	1fmw	1gor	1i13	1jht	1kl7	1ngi	1rad	1yac	3kin
1a96	1b9z	1cjk	1dmr	1eh3	1fmb	1gos	1i14	1jhu	1klm	1ngj	1rac	1yag	3ktq
1a97	1ba0	1cjm	1dms	1eh5	1fmc	1got	1i1h	1jha	1kln	1ngs	1raf	1yas	3lad
1a98	1ba1	1cjt	1dmu	1eh6	1fnd	1gow	1i1n	1jhb	1kly	1nhp	1rag	1ybv	3lip
1a99	1ba2	1cju	1dnc	1eh7	1fnn	1gox	1i1o	1jhd	1klz	1nhq	1rah	1ydv	3ljr
1a9n	1bag	1civ	1dnp	1eh8	1fnn	1gpl	1i1q	1jhm	1km0	1nhr	1rai	1yfo	3mag
1a9o	1bam	1civ	1do0	1eh9	1fno	1gp2	1i2d	1jho	1km1	1nhs	1ral	1ygp	3man
1a9p	1bap	1ck4	1do2	1eha	1fns	1gpa	1i2m	1jhp	1km2	1nht	1rb2	1yja	3mbp
1a9q	1bav	1cke	1do8	1ehc	1fo1	1gpb	1i2n	1jhq	1km3	1NIS	1rb3	1yjb	3mct
1a9r	1bay	1ckq	1doa	1ehi	1fo5	1gpd	1i2o	1jhr	1km4	1nip	1rba	1yjc	3mht
1a9s	1bc5	1cl0	1dob	1ehn	1fo8	1gpe	1i2p	1jhu	1km5	1nit	1rbl	1ykc	3min
1a9t	1bcm	1cl1	1doc	1ehy	1fo9	1gph	1i2q	1jhw	1km6	1nks	1rbo	1ylv	3muc
1a9x	1bco	1cl2	1dod	1ei9	1foa	1gpj	1i2r	1jhx	1kmc	1nnt	1rbr	1yme	3ng1
1a9y	1bcr	1cl8	1doe	1eib	1foe	1gpl	1i2z	1jhy	1kmh	1noc	1rbs	1ymu	3nll
1a9z	1bcs	1cla	1doh	1ein	1foh	1gpm	1i30	1jhz	1kmi	1noi	1rbt	1ymv	3pbg
1aa1	1bd0	1cle	1dor	1eiw	1fok	1gpu	1i32	1ji0	1kmj	1noj	1rbu	1yoo	3pfk
1aa6	1bd3	1clk	1dos	1eix	1fov	1gpw	1i33	1ji1	1kmk	1nok	1rbv	1ypi	3pfl
1aa8	1bd4	1clq	1dot	1eiz	1fp1	1gpy	1i36	1ji2	1kmm	1noy	1rc4	1ypn	3pga
1aa9	1bdb	1clu	1doz	1ej0	1fp2	1gq2	1i39	1jib	1kmn	1noz	1rcf	1ypt	3pgk
1aam	1bdg	1clv	1dp0	1ej2	1fp6	1gq6	1i3a	1jii	1kms	1npv	1rch	1yrg	3pgm
1aat	1bdh	1clx	1dp2	1ej7	1fp7	1gq7	1i3c	1jij	1kmv	1nsa	1rco	1ysc	3pgt
1aaw	1bdi	1cm5	1dp4	1ejb	1fpm	1gq9	1i3k	1jik	1kng	1nsf	1rcx	1ytn	3pjr
1aax	1bdj	1cm7	1dpb	1eji	1fpq	1gqc	1i3l	1jil	1knp	1nsj	1rd7	1yts	3pmg
1aaz	1bdl	1cml	1dpc	1ejj	1fpr	1gqf	1i3m	1jiq	1knq	1nst	1rda	1ytw	3pnp
1ab5	1be0	1cmw	1dpd	1ejr	1fpv	1gqr	1i3n	1jiu	1knr	1nsy	1rdb	1yub	3prk
1ab6	1be1	1cne	1dpe	1ejs	1fpz	1gqs	1i3o	1jiv	1knv	1ntr	1rdc	1yve	3ptd
1ab7	1be6	1cnf	1dpf	1ejt	1fq0	1gqt	1i3q	1jix	1ko1	1nul	1rdd	1yvn	3pvi
1aba	1be8	1cnm	1dpg	1eju	1fq1	1gr0	1i41	1jj2	1ko4	1nzy	1rdh	1zak	3r1r
1abb	1bed	1cnv	1dpi	1ejv	1fqa	1gr2	1i43	1jj7	1ko5	1oaa	1re7	1zay	3rab
1abe	1bee	1cnz	1dpm	1ek0	1fqb	1gr8	1i45	1jjb	1ko7	1oak	1rea	1zbd	3ran
1abf	1beu	1cow	1dpp	1ek1	1fqc	1gr9	1i48	1jjf	1ko8	1oas	1req	1zen	3rap
1abn	1bez	1coy	1dpz	1ek2	1fqd	1gra	1i4d	1jji	1kof	1oat	1rev	1zfi	3req
1acl	1bf2	1coz	1dqn	1ek4	1fqe	1grb	1i4e	1jjk	1kog	1obr	1rf7	1zin	3rla
1ac5	1bf3	1cp2	1dqp	1ek5	1fqf	1grc	1i4l	1jjv	1koh	1oce	1rg7	1zin	3rub
1acj	1bf6	1cp3	1dqr	1ek6	1fqj	1gre	1i4n	1jke	1koj	1oel	1rh3	1zio	3sc2
1acl	1bfd	1cp6	1dqu	1eke	1fqk	1grf	1i4o	1jki	1koo	1oen	1rhd	1zip	3sic
1acm	1bfk	1cp7	1dqw	1ekj	1fqo	1grg	1i4t	1jkh	1kor	1ofg	1rhs	1zon	3tat
1aco	1bfm	1cpb	1dqx	1ekk	1fqv	1grk	1i4w	1jki	1kp0	1ofv	1ril	1zoo	3tec
1acv	1bfu	1cps	1dqy	1eko	1fqw	1grl	1i50	1jjk	1kp2	1ohj	1rk2	1zop	3tgi
1ad1	1bg2	1cpu	1dqz	1ekq	1fr8	1grm	1i51	1jkm	1kp3	1ohk	1rka	1zpd	3thi
1ad3	1bg3	1cpx	1dr0	1ekx	1fr9	1gro	1i52	1jlx	1kp9	1oib	1rkd	1zrm	3tim

lad4	lbg4	lcpy	ldr1	lel3	lfrb	lgrp	li57	lj11	lkgp	loil	lrm	lzm	3tmk
ladb	lbg5	lcq6	ldr2	lel5	lfrn	lgrq	li5e	lj12	lkph	lojt	lrks	lzym	3tmy
ladc	lbg6	lcq7	ldr3	lel7	lfrq	lgrr	li5g	lj13	lkpi	lola	lrta	20gs	3trx
ladd	lbg9	lcq8	ldr4	lel8	lfrw	lgrt	li5o	lj15	lkqf	lola	lrta	21gs	3ts1
lade	lbg8	lcqg	ldr5	lel9	lfrz	lgrx	li5s	lj18	lkqg	lomp	lrld	221p	3uag
ladf	lbgb	lcqh	ldr6	leli	lfs0	lgs5	li60	lj1a	lkqp	lone	lrir	22gs	3ubp
ladg	lbgg	lcqi	ldr7	lelj	lfs2	lgsa	li69	lj1b	lkqs	lonr	lrnh	2a3h	3ukd
ladi	lbgj	lcqj	ldr8	lell	lfs4	lgsb	li6a	lj1c	lkqy	lopr	lrnl	2aaa	3vsb
ladj	lbg1	lcqp	ldr9	lelm	lfs5	lgsc	li6b	lj1e	lkqz	lorb	lrpl	2aat	3vbk
lado	lbgm	lcqw	ldr6	lelo	lfs6	lgsc	li6h	lj1f	lkr0	lord	lrpa	2abh	3xim
lads	lbgm	lcqx	ldre	lelq	lfsf	lgse	li6i	lj1g	lkr1	loro	lrpj	2ace	3xin
lady	lbg8	lcqz	ldr7	lels	lfsf	lgse	li6k	lj1j	lkr5	lort	lrpm	2ack	3xis
lae1	lbg1	lcr0	ldr8	lelu	lfsf	lgsg	li6l	lj1k	lkr6	lose	lrpt	2acq	3yas
lae4	lbg4	lcr1	ldrj	lelx	lfsf	lgsh	li6m	lj1l	lkrb	losi	lrpx	2acr	3ypi
laf2	lbgv	lcr2	ldrj	lely	lfsu	lgsi	li6n	lj1m	lkrb	losj	lrrf	2acs	421p
laf4	lbgx	lcr4	ldrj	lelz	lfsz	lgsj	li6o	lj1n	lkrh	loth	lrrg	2acu	4a3h
laf7	lbh2	lcr6	ldrj	lem6	lft7	lgsm	li6p	lj1q	lkrp	lotp	lrrp	2ada	4ake
lafs	lbh6	lcr1	ldrj	lemd	lft8	lgso	li6q	lj1r	lks2	lovb	lrsc	2adm	4ald
lafw	lbhg	lcrp	lds6	lemh	lftg	lgsg	li6w	lj1s	lks9	lovt	lrtd	2ae1	4atl
lag1	lbhj	lcrq	lds9	lemj	lftj	lgss	li74	lj1v	lksj	loxm	lrtd	2ae2	4bdp
lag8	lbh1	lcr	lds6	leno	lftk	lgss	li75	lj1w	lksj	loxm	lrtd	2ak2	4cd2
lag9	lbhm	lcrw	lds6	lenp	lftl	lgss	li7p	lj1m	lksj	loxp	lrtd	2ak3	4cev
lagu	lbho	lcrz	lds6	lenu	lftm	lg7	li7q	lj1m	lksj	loya	lrtd	2aky	4cgt
lagp	lbhq	lcs0	ldt3	leny	lftn	lg8	li7s	lj1m	lksj	loyb	lrtd	2ald	4cla
lagr	lbhs	lcs1	ldt5	lenz	lfto	lg9	li80	lj1n	lksu	loyc	lrtd	2alr	4cpa
laga	lbhw	lcs4	ldt9	leo1	lftq	lgt1	li86	lj1n	lksj	lp39	lrtd	2amg	4dbv
lagx	lbhy	lcs5	ldtd	leo3	lfts	lgtb	li88	lj1a	lktb	lpam	lrtd	2amv	4dcg
laga	lbi4	lct5	ldte	leo4	lftw	lgtc	li89	lj1c	lkte	lpau	lrtd	2anh	4dfr
lah0	lbi5	lct9	ldtg	leo5	lftj	lgtg	li8b	lj1d	lkte	lpbb	lrtd	2atl	4dmr
lah2	lbi9	lctn	ldtm	leo7	lftu	lgtg	li8j	lj1e	lkti	lpbc	lrtd	2atl	4eca
lah3	lbif	lctq	ldtq	leog	lftu	lgti	li8t	lj1f	lkti	lpbd	lrtd	2atc	4enl
lah4	lbis	lctt	ldts	leoh	lftu	lgtj	li94	lj1r	lkti	lpbe	lrtd	2axe	4eug
lah5	lbis	lctu	ldtt	leok	lftu	lgtl	li95	lj1x	lku0	lpbf	lrtd	2ay1	4fua
lahe	lbis	lcu1	ldtu	leom	lftu	lgtm	li96	lj1y	lkv5	lpbg	lrtd	2ay2	4fx2
lahf	lbiz	lcua	ldtw	leon	lftu	lgtu	li97	lj1z	lkv8	lpbn	lrtd	2ay3	4gpb
lahg	lbj4	lcub	ldty	leoo	lftu	lgtu	li9c	lj1a	lkva	lpbp	lrtd	2ay4	4gpd
lahh	lbjk	lcuc	ldtz	leop	lftu	lgtu	li9d	lj1l	lkvb	lpca	lrtd	2ay5	4gr1
lahi	lbjn	lcud	ldu4	lep1	lftu	lgtv	li9g	lj1m	lkvc	lpd2	lrtd	2ay6	4grt
lahn	lbjo	lcue	ldub	lep2	lftj	lgtx	li9s	lj1n	lkvd	lpda	lrtd	2ay7	4gsa
lahp	lbjr	lcuf	ldug	lep3	lftk	lgtz	li9t	lj1p	lkvr	lpdh	lrtd	2ay8	4gss
lahx	lbjw	lcug	ldui	lep4	lfto	lgtu	li9u	lj1p	lkvs	lpdo	lrtd	2ay9	4gst
lahy	lbjx	lcuh	ldus	lep7	lftv	lgtu	li9v	lj1p	lkvt	lpdy	lrtd	2ayq	4gtu
lai2	lbka	lcui	lduv	lep8	lftx	lgtu	li9w	lj1p	lkvu	lpdz	lrtd	2azo	4icd
lai3	lbkd	lcuj	ldv1	lep9	lftw	lgtu	li9x	lj1p	lkw1	lpda	lrtd	2bam	4ktq
lai9	lbkg	lcul	ldv2	lepx	lftw	lgtu	li9y	lj1p	lkw1	lpda	lrtd	2bce	4lip
laia	lbkh	lcus	ldv7	leq2	lftw	lgtu	li9z	lj1p	lkw1	lpda	lrtd	2bdp	4mbp
laib	lbks	lcuu	ldvj	leqa	lftw	lgtu	li9a	lj1p	lkw1	lpda	lrtd	2bgt	4mdh
laic	lbkw	lcuv	ldvr	leqb	lftw	lgtu	li9b	lj1p	lkw1	lpda	lrtd	2bgu	4mht
laip	lbky	lcuw	ldwa	leqc	lftw	lgtu	li9c	lj1p	lkw1	lpda	lrtd	2bif	4nll

1aiu	1bl3	lcux	ldwf	leqj	lfwc	lgv0	liay	ljq3	lkxp	lpgo	lsay	2bjx	4nul
1aiv	1bl5	lcuy	ldwg	leqp	lfwd	lgv1	lib0	ljqd	lkxq	lpgp	lsbc	2bnh	4pbg
1aj0	1ble	lcuz	ldwh	lequ	lfwe	lgv4	lib6	ljqe	lkxt	lpgq	lsbh	2btf	4pfk
1aj2	1blf	lcw2	ldwi	leqy	lfwf	lgv8	libc	ljqf	lkxv	lpgt	lsbi	2btm	4pga
1aja	1bli	lcvi	ldwj	leri	lfwg	lgvc	libj	ljag	lky2	lphh	lsbn	2bup	4pgm
1ajb	1bli	lcvi	ldwo	lert	lfwh	lgve	libr	ljqi	lky3	lphp	lsbp	2bvww	4pgt
1ajc	1bmd	lcvr	ldwp	leru	lfwi	lgvf	libs	ljql	lkyb	lphr	lsbt	2can	4pnp
1ajd	1bmf	lcw0	ldwq	lerv	lfwj	lgvh	lic6	ljrl	lkyg	lpif	lsca	2cbf	4ptd
1aje	1bmj	lcw1	ldx4	lerw	lfwn	lgvi	lice	ljr2	lkyh	lpig	lsch	2cd2	4q2l
1ajr	1bmt	lcw2	ldx6	les8	lfwr	lgw9	lici	ljr3	lkyi	lpil	lsch	2cev	4r1r
1ajs	1bn6	lcw3	ldx9	les9	lfws	lgwc	licp	ljr4	lkyq	lpjb	lsch	2che	4req
1ajz	1bn7	lcw4	ldxe	lesc	lfwt	lgwj	licq	ljr6	lkyv	lpjc	lsch	2chf	4rla
1ak1	1bnc	lcw7	ldxf	lesd	lfww	lgwn	lics	ljrx	lkyw	lpjr	lsch	2chr	4rub
1ak2	1bo5	lcwn	ldxh	lese	lfwy	lgwz	lid1	ljry	lkyx	lpkl	lsel	2chy	4rve
1ak5	1bo6	lcwr	ldxi	lesg	lfx0	lgyl	lid8	ljrz	lkyx	lpkm	lsch	2cla	4skn
1ak9	1bof	lcws	ldxl	lesj	lfx1	lgym	lidc	ljs1	lkyz	lpkn	lsch	2cmd	4tgl
1aka	1boh	lcwt	ldxo	lesm	lfx6	lgyp	lidd	ljs3	lkz1	lpky	lsft	2cmk	4thi
1akb	1boi	lcwu	ldxq	lesn	lfxj	lgyq	lide	ljs6	lkz4	lpjl	lsch	2cnd	4tim
1akc	1boo	lcwy	ldxy	lesq	lfxo	lgyt	lidf	ljsc	lkz6	lpik	lsib	2cpu	4tmk
1ake	1bot	lcx8	ldyh	lesv	lfxp	lgz4	lidm	ljsl	lkz7	lpil	lsjs	2cst	4tmy
1akm	1bou	lcx9	ldyi	lesw	lfxq	lgzd	lidn	ljsr	lkz9	lpmt	lsky	2ctb	4trx
1akn	1boz	lcxe	ldyj	lesz	lfxs	lgzg	lido	ljsx	lkzg	lpnr	lsma	2ctc	4tsl
1akq	1bp5	lcxf	ldyr	leth	lfxu	lgzh	lie3	ljsz	lkzh	lpnt	lsmd	2cut	4uag
1akr	1bpl	lcxh	ldys	letu	lfxw	lgzj	lie7	ljt2	lkzy	lpoi	lsom	2cxg	4ubp
1akt	1bpm	lcxi	ldyv	leul	lfxx	lgzu	liei	ljt9	llob	lpot	lson	2dap	4ukd
1aku	1bpn	lcxk	ldz3	leu8	lfy2	lh0g	liej	ljte	llop	lpow	lsoo	2dbv	4xia
1akv	1bpw	lcxl	ldzf	leua	lfy6	lh0i	lieq	ljtf	llov	lpox	lspa	2dfp	4xim
1akw	1bq3	lcxq	ldzi	leuc	lfy9	lh11	liev	ljtk	lila	lpoy	lspb	2dhc	4xis
1aky	1bq4	lcxu	ldzu	leud	lfya	lh11	liev	lju3	lile	lppi	lsrr	2dhd	4yas
1akz	1bq6	lcxz	ldzv	leug	lfyd	lh1n	liev	lju4	lilf	lprx	lsrv	2dhe	521p
1al3	1bq7	lcyd	ldzw	leuh	lfye	lh2e	lif2	lju9	lil1	lpsc	lsrx	2dhf	5a3h
1al7	1bqa	lcye	ldzx	leui	lfyf	lh2f	lifx	ljud	lils	lpsd	lssp	2dhq	5abp
1al8	1bqc	lcyg	ldzy	leun	lfyv	lh2h	lig0	ljuk	lil2	lpsz	lst2	2dij	5acn
1ald	1bqd	lczi	ldzz	leuq	lfyw	lh2j	lig3	ljul	lilr	lpta	lst3	2dik	5adh
1alh	1bqe	lcz3	le0a	leuy	lfyx	lh4k	lig8	ljuy	lilt	lptd	lsto	2dkb	5atl
1ali	1bqg	lcz7	le0c	leuz	lfzq	lh4m	lig9	ljv1	lilu	lptg	lsua	2dlld	5bca
1alj	1bqm	lcz9	le0d	lev4	lfzr	lh4o	ligb	ljv2	lilu	lptk	lsua	2dlld	5bca
1alk	1bqn	lcza	le0j	lev7	lfzt	lh4q	ligr	ljv3	lily	lptt	lsuc	2dmr	5cgt
1ahn	1br1	lczb	le0k	lev9	lfzw	lh4s	lign	ljvn	lila	lptu	lsud	2dor	5chy
1am4	1br2	lczc	le0r	leve	lg01	lh4t	ligw	ljvx	lilg	lptv	lsue	2dpg	5cpa
1ama	1br4	lcze	le0s	levi	lg0c	lh4v	lih7	ljvy	lilg	lpty	lsup	2dpm	5dfr
1amf	1brl	lczh	le0t	levj	lg0n	lh4x	lih8	ljw4	lilj	lpud	lsur	2drc	5eaa
1ami	1brm	lczk	le0u	levk	lg0o	lh4y	lih9	ljw5	lilj	lpvd	lsvn	2dri	5enl
1amj	1bro	lczi	le0v	levl	lg0r	lh4z	lihi	ljw9	lilv	lpvi	lszj	2dub	5eug
1amk	1brs	lczn	le0w	levq	lg15	lh50	lihn	ljwa	lilw	lpvu	lt7p	2ebn	5fwg
1amn	1brt	lczo	le0x	levy	lg16	lh5n	liho	ljwb	lilf	lpxa	ltad	2eck	5fx2
1amo	1brw	lczt	le15	levz	lg17	lh5q	lihp	ljwl	lilg	lpxb	ltag	2ecp	5gpb
1amp	1bs0	lczu	le18	lew2	lg18	lh5r	lihu	ljwx	lil6	lpxc	ltah	2eda	5grt
1amq	1bs1	ld02	le19	lew8	lg19	lh5s	lihx	ljwy	lil6	lpxt	ltaq	2edc	5gss

lamr	lbs2	ld03	lelc	lew9	lgl1	lh5t	lihy	ljx2	ll6w	lpyd	ltar	2efg	5gst
lams	lbs9	ld04	lele	lewd	lg1f	lh5u	lii0	ljx6	ll6y	lpyg	ltas	2erc	5icd
lamy	lbsi	ld07	lelf	lewe	lg1g	lh5v	lii2	ljxa	ll7m	lpym	ltat	2eug	5ktq
lan0	lbsl	ld09	lelk	lewg	lg1h	lh5y	lii5	ljxb	ll7n	lpyt	ltau	2euh	5ldh
lan9	lbsd	ld0b	lell	lewk	lg1l	lh60	lii6	ljxe	ll7o	lq21	ltc1	2exo	5lip
lanf	lbsu	ld0i	lelm	lewq	lg1m	lh61	lii8	ljxh	ll7p	lqam	ltc2	2eza	5mdh
lani	lbsd	ld0s	leln	lewr	lg1y	lh62	lii9	ljxi	ll7v	lqan	ltca	2ezb	5mht
lanj	lbt4	ld0v	lelq	lewt	lg20	lh63	liib	ljxj	ll8a	lqao	ltcb	2ezc	5nll
lank	lbt4	ld0x	lelr	lewv	lg21	lh65	liig	ljxk	ll8g	lqap	ltcc	2fax	5nul
lao0	lbtb	ld0y	lely	lewx	lg23	lh66	liih	ljxm	ll8k	lqaq	ltcd	2fer	5p2l
lao3	lbtj	ld0z	lelz	lewy	lg29	lh69	liim	ljxo	ll8p	lqas	ltcm	2fdx	5pgm
lao8	lbtj	ld1a	1E20	lex1	lg2i	lh6a	liin	ljxq	ll9f	lqat	ltde	2ffh	5pnt
laod	lbtm	ld1b	1E26	lex2	lg2o	lh6b	liir	ljxz	ll9x	lqb0	ltdf	2fhe	5ptd
laoe	lbt0	ld1c	le2b	lex4	lg2p	lh6c	liit	ljye	lla1	lqb4	ltdj	2flv	5r1r
laog	lbu5	ld1g	le2d	lex5	lg2q	lh6d	liiw	ljyf	lla2	lqb7	ltdr	2fnt	5req
laon	lbu6	ld1p	le2e	lex6	lg2u	lh6j	lijb	ljyk	llaf	lqb8	ltcc	2fok	5r1a
laov	lbu8	ld1q	le2f	lex7	lg2v	lh6t	lije	ljyl	llag	lqba	lteh	2fox	5rub
laox	lbua	ld1s	le2g	lex9	lg38	lh6u	lijf	ljyn	llah	lqbb	ltfa	2fsp	5sic
lapb	lbup	ld1t	le2h	lexb	lg3i	lh6v	lijh	ljyv	llam	lqbg	ltfd	2fua	5tgl
laq0	lbuz	ld2a	le2i	lexc	lg3l	lh6z	liji	ljyw	llan	lqbk	ltfr	2fvx	5tim
laq2	lbvh	ld2c	le2j	lexd	lg3q	lh75	lijj	ljyx	llap	lqc5	ltgl	2fx2	5tmp
laq6	lbvn	ld2e	le2k	lexm	lg3r	lh76	lijk	ljyy	llar	lqc9	ltgo	2gar	5ukd
laqh	lbvr	ld2f	le2l	lexn	lg3u	lh77	lik4	ljyz	llau	lqca	ltfh	2gbp	5ull
laqf	lbvu	ld2g	le2m	lexp	lg41	lh78	lik6	ljz0	llav	lqcc	ltgh	2gd1	5xia
laqi	lbvw	ld2h	le2n	lexq	lg42	lh79	lika	ljz1	llaw	lqcd	ltgm	2glr	5xim
laqj	lbvy	ld2i	le2o	lexv	lg4a	lh7a	likv	ljz2	llb1	lqcw	ltgo	2glt	5xin
laql	lbvz	ld2k	le2p	lexw	lg4b	lh7b	likw	ljz3	llb8	lqcz	ltgt	2gpa	5yas
laqm	lbw0	ld2m	le2q	ley3	lg4e	lh7e	likx	ljz4	llb9	lqde	ltgx	2gpb	621p
laqn	lbw9	ld2n	le2s	leye	lg4h	lh7f	liky	ljz5	llbb	lqdl	ltia	2gpn	6a3h
laqu	lbwc	ld2r	le2y	leyr	lg4p	lh7g	likz	ljz6	llbc	lqds	ltib	2grt	6abp
laqv	lbwf	ld3a	1E32	leyu	lg4s	lh7h	lii0	ljz7	llbf	lqdu	ltic	2gsa	6acn
laqw	lbwk	ld3c	1E33	leyw	lg4t	lh7i	liie	ljz8	llbg	lqe0	ltim	2gsq	6adh
laqx	lbwl	ld3g	1E39	leyy	lg4u	lh7n	liio	ljzq	llbh	lqe1	ltip	2gsr	6ald
laqy	lbwp	ld3h	le3c	leyz	lg4v	lh7o	liiv	ljzr	llbi	lqe3	ltka	2gss	6atl
larg	lbwq	ld3k	le3e	lez0	lg4w	lh7p	liiw	ljzs	llbk	lqe5	ltkb	2gst	6cgt
larh	lbwr	ld3v	le3i	lez1	lg4x	lh7q	liiy	ljzt	llbs	lqf4	ltkc	2gtu	6chy
lari	lbws	ld4a	le3j	lez2	lg55	lh7r	lim2	ljzw	llbt	lqf5	ltlf	2gyi	6cpa
lari	lbwv	ld4c	le3l	lez4	lg59	lh7t	lim5	lk02	llc0	lqf6	ltmh	2had	6dfr
larm	lbx0	ld4d	le3m	lez9	lg5a	lh7w	lim8	lk03	llc3	lqf7	ltmk	2hdh	6enl
lars	lbx1	ld4e	le3q	leza	lg5c	lh7x	limj	lk06	llc5	lqf9	ltml	2hgs	6gpb
lart	lbx3	ld4f	le3s	lezb	lg5f	lh81	limo	lk08	llc7	lqfe	ltmo	2his	6gss
larz	lbx4	ld4g	le3t	lezc	lg5h	lh82	lin1	lk0a	llc8	lqff	ltmq	2hlp	6gst
las0	lbx9	ld4n	le3w	lezd	lg5i	lh83	lin4	lk0b	llcf	lqfl	ltmy	2hmi	6gsu
las2	lbxb	ld4o	le3x	lezi	lg5p	lh84	lin5	lk0c	llco	lqfm	ltnd	2hmy	6gsv
las3	lbxc	ld4x	le3z	lezk	lg5q	lh85	lin6	lk0d	llcp	lqfn	ltoa	2hnp	6gsw
lasa	lbxc	ld4z	1E40	lezo	lg5r	lh86	lin7	lk0f	llct	lqfs	ltof	2hnq	6gsx
lasb	lbxc	ld5a	1E43	lezp	lg5t	lh8e	lin8	lk0i	llcu	lqfx	ltpb	2hpa	6gsy
lasc	lbxc	ld5c	1E46	lezz	lg5x	lh8h	lini	lk0j	lldb	lqfy	ltpe	2hsd	6icd
lasd	lbxr	ld5r	1E47	lezw	lg60	lh8l	linj	lk0l	lldc	lqfz	ltpd	2hvm	6ldh

lase	lbox	ld5t	1E48	lezz	lg63	lh93	linl	lk0m	lIde	lqg0	ltpc	2isd	6mht
lasf	lbxz	ld5w	1E49	lf05	lg64	lh94	lio2	lk0n	lIdg	lqg2	ltpf	2itg	6nul
lasg	lbya	ld6f	1e4a	lf06	lg66	lh9a	lioF	lk0o	lIdm	lqg4	ltpb	2kau	6pfk
lasl	lbyb	ld6h	1e4b	lf07	lg67	lh9b	lioI	lk0u	lIdn	lqg6	ltpI	2kfn	6ptd
lasm	lbyc	ld6i	1e4c	lf0k	lg69	lh9c	liok	lk1d	lIdy	lqg8	ltpT	2kFz	6q2l
lasn	lbyd	ld6j	1e4e	lf0n	lg6c	lha3	liol	lk1e	lIeh	lqga	ltpu	2ki5	6rlr
lass	lbye	ld6n	1e4f	lf0o	lg6h	lhbj	lion	lk20	lIfI	lqge	ltpv	2kin	6req
lasu	lbyi	ld6s	1e4g	lf0p	lg6o	lhc7	liov	lk23	lIfa	lqgn	ltpw	2ktq	6taa
lasv	lbyk	ld7a	1e4i	lf0y	lg6q	lhd2	liow	lk2v	lIfd	lqgo	ltrI	2kzm	6tim
lasw	lbyu	ld7f	1e4l	lf12	lg6w	lhdc	lioz	lk2y	lIfg	lqqq	ltrb	2kzz	6xia
lasx	lbzc	ld7k	1e4m	lf14	lg6y	lhde	lipa	lk30	lIfh	lqgs	ltrd	2lao	6xim
latl	lbzf	ld7l	1e4n	lf17	lg7c	lhdg	lipd	lk32	lIfi	lqgu	ltre	2lbp	6yas
latg	lbzh	ld7o	1e4o	lf1b	lg7e	lhdh	liph	lk35	lIgb	lqgv	ltrh	2ldb	721p
lati	lbzj	ld7r	1e4v	lf1j	lg7f	lhdi	lipi	lk3c	lIgy	lqgz	ltri	2ldx	7a3h
latn	lbzl	ld7s	1e4y	lf1z	lg7g	lhdo	liq0	lk3d	lIh0	lqh0	ltrl	2lip	7aat
latr	lbzy	ld7u	1E51	lf20	lg7o	lhdq	liq7	lk3f	lIi5	lqh1	ltrs	2liv	7abp
lats	lc0d	ld7v	1E55	lf21	lg7r	lhdr	liq8	lk3r	lIi7	lqh8	ltru	2ljr	7acn
latz	lc0e	ld7y	1E56	lf2d	lg7s	lhdu	liqp	lk3t	lIii	lqh9	ltrv	2man	7adh
lau9	lc0f	ld8a	1E58	lf2e	lg7t	lhdx	liqr	lk3u	lIij	lqha	ltrw	2mas	7atl
lauh	lc0g	ld8c	1E59	lf2j	lg7u	lhdy	liqu	lk4g	lIik	lqhe	ltsi	2min	7cgt
laur	lc0i	ld8r	1e5d	lf2o	lg7v	lhdz	lir1	lk4h	lIio	lqhf	ltti	2mjp	7cpa
laug	lc0k	ld8t	1e5e	lf2p	lg7w	lhel	lir2	lk4q	lIjl	lqhg	lttj	2mmr	7dfr
lauk	lc0l	ld8w	1e5f	lf2t	lg7x	lhe2	lir6	lk4v	lIjr	lqhh	lttp	2muc	7enl
lauo	lc0m	ld8y	1e5j	lf2u	lg8f	lhe3	liri	lk4y	lIju	lqhi	lttq	2mys	7gpb
laup	lc0n	ld9d	1e5k	lf2v	lg8g	lhe4	lirx	lk5d	lIk0	lqhl	lttt	2nac	7gss
lauq	lc0p	ld9e	1e5l	lf37	lg8h	lhe5	liso	lk5g	lIk5	lqhm	ltub	2nad	7icd
laur	lc0t	ld9f	1e5m	lf3a	lg8j	lhe8	liss	lk5h	lIk7	lqhn	ltui	2nap	7mdh
laus	lc0u	ld9t	1e5n	lf3b	lg8k	lhee	liss	lk6a	lIkc	lqho	ltux	2ncd	7mht
lauv	lc14	ld9v	1e5q	lf3e	lg8l	lhei	lisv	lk6d	lIkx	lqhp	ltvr	2ngl	7ode
laux	lc1a	ld9x	1e5t	lf3i	lg8m	lhet	lisw	lk6i	lIkz	lqhs	ltx4	2ngr	7ptd
lauz	lc1b	ld9y	1e5v	lf3l	lg8o	lheu	lisx	lk6j	lIl0	lqht	ltxx	2nip	7rlr
lav6	lc1c	ld9z	1e5x	lf3o	lg8p	lhex	lisy	lk6w	lIl2	lqhx	ltya	2npv	7req
lav7	lc1d	ldad	1E60	lf3p	lg8r	lhey	lisz	lk6x	lIl3	lqhy	ltyb	2nsy	7taa
lava	lc1h	ldae	1E61	lf3t	lg8y	lhf2	lit0	lk6y	lIlc	lqhz	ltyc	2oat	7tim
lavq	lc1x	ldaf	1E62	lf3w	lg93	lhf3	lit7	lk70	lIl d	lqi0	ltyd	2ohx	7xim
lavl	lc1y	ldag	1E63	lf3x	lg94	lhf6	lit8	lk75	lIl o	lqi1	ltyf	2olb	7yas
law1	lc25	ldah	1E64	lf48	lg95	lhfe	litg	lk77	lIl q	lqi2	ltyp	2one	821p
law2	lc29	ldai	1E66	lf49	lg97	lhfi	litq	lk7c	lIl w	lqi3	ltyt	2ora	8a3h
law5	lc2b	ldaj	1E69	lf4a	lg98	lhfk	litu	lk7e	lIl z	lqi4	luaa	2otc	8aat
law9	lc2o	ldak	1e6b	lf4h	lg99	lhfp	litx	lk7f	lIm1	lqi5	luag	2oxi	8abp
lax4	lc2t	ldam	1e6c	lf4l	lg9h	lhfq	liu8	lk7h	lIm q	lqi6	lubn	2pbg	8acn
lax9	lc2y	ldao	1e6e	lf4p	lg9r	lhfr	liud	lk7k	lIm y	lqid	lubp	2pda	8adh
laxd	lc30	ldap	1e6k	lf4v	lg9s	lhfv	lius	lk7s	lIl ol	lqie	lubs	2pfk	8atl
laxe	lc3e	ldar	1e6l	lf51	lg9t	lhfw	liut	lk7x	lIl on	lqif	lucw	2pfl	8atc
laxg	lc3f	ldb0	1e6m	lf57	lg9u	lhg0	liuu	lk7y	lIl oo	lqig	luda	2pgd	8cgt
laxr	lc3j	ldb3	1e6n	lf5n	lga1	lhgl	liuv	lk83	lIl oq	lqih	ludb	2pgi	8cpa
lay0	lc3l	ldbi	1e6p	lf5s	lga4	lhg3	liuw	lk86	lIl or	lqii	ludc	2pgt	8dfr
lay4	lc3o	ldbp	1e6q	lf5z	lga6	lhgw	liux	lk88	lIl os	lqij	ludg	2phh	8gpb
lay5	lc3p	ldbq	1e6r	lf60	lga8	lhgx	livr	lk89	lIl ot	lqik	ludh	2pia	8gss

lay7	lc3q	ldbr	le6s	lf6l	lgad	lhgy	livy	lk8a	lp6	lqim	ludi	2pjr	8icd
lay8	lc3r	ldbs	le6u	lf6b	lgae	lhh1	liwe	lk8c	lpa	lqir	ludr	2pkc	8ldh
laye	lc3s	ldbt	le6w	lf6d	lgaj	lhh4	lix6	lk8k	lpb	lqis	luh	2plc	8mht
layl	lc3x	ldbv	le6x	lf6k	lgai	lhi9	lix7	lk8q	lpf	lqit	luio	2pmt	8ohm
laz0	lc4l	ldbw	le6z	lf6m	lgaq	lhiz	lix8	lk8r	lpm	lqj3	luip	2pri	8ruc
lazl	lc47	ldby	lE70	lf6p	lgar	lhj6	lixg	lk8y	lpn	lqj4	luke	2prj	8tim
laz2	lc4a	ldc1	lE71	lf6w	lgaw	lhjk	lixh	lk8z	lpo	lqj5	luky	2prk	8xia
laz3	lc4c	ldc3	lE72	lf6y	lgax	lhjo	lixl	lk92	lpp	lqjc	lukz	2ptd	8xim
laz4	lc4g	ldc4	lE73	lf73	lgbn	lhjr	lj49	lk97	lps	lqjd	lula	2pth	9aat
laz9	lc4k	ldc5	lE77	lf74	lgc0	lhkb	lj4a	lk98	lqf	lqjm	lulb	2pua	9abp
lazl	lc4o	ldc6	lE79	lf75	lgc2	lhkc	lj4b	lk9m	lqt	lqjw	luok	2pub	9atc
lazo	lc4t	ldc7	le7m	lf7b	lgc3	lhld	lj4e	lk9s	lqu	lqk0	lupf	2puc	9cgt
lazz	lc4w	ldc8	le7p	lf7u	lgc4	lhlf	lj4x	lk9v	lrl	lqk2	lupu	2pud	9gss
lazi	lc4x	ldce	le7q	lf7v	lgc5	lhlg	lj5e	lkae	lrl	lqk3	lura	2pue	9icd
lazw	lc50	ldcf	le7r	lf8g	lgc8	lhlp	lj5o	lkag	lrl	lqk4	lurb	2puf	9ldb
lazy	lc5k	ldci	le7s	lf8i	lgc9	lhlu	lj5p	lkah	lss	lqk5	luro	2pug	9ldt
lb00	lc6v	ldck	le7w	lf8m	lgca	lhmo	lj5s	lkak	lst	lqk8	lurp	2pvi	9mht
lb05	lc72	ldcm	le7y	lf8r	lgcg	lhms	lj5t	lkam	lsu	lqka	luug	2q2l	9rub
lb0a	lc7e	ldct	lE89	lf8s	lgci	lhms	lj5v	lkao	ltd	lqkb	luwb	2qr2	9xia
lb0h	lc7f	ldd8	le8c	lf8u	lgck	lhmp	lj5x	lkaq	lth	lqki	lv39	2r1r	9xim
lb0i	lc7i	lddg	le8d	lf8w	lgco	lhmv	lj6n	lkar	luc	lqkj	lvdc	2rap	
lb0j	lc7j	lddi	le8m	lf8x	lgcu	lhmy	lj6z	lKAS	lvh	lqkk	lvdr	2reb	
lb0k	lc7n	lddo	le8n	lf8y	lgcx	lhnl	lj79	lkav	lvk	lqla	lvfd	2req	
lb0l	lc7o	lddr	lE92	lf9l	lgcy	lhns	lj7k	lkax	lvl	lqlb	lvfe	2rkm	
lb0m	lc7q	ldds	lE94	lf9a	lgdl	lhna	lj7x	lkay	lw6	lqlh	lvfn	2rla	
lb0p	lc7r	lddz	lE96	lf9b	lgd9	lhnb	lj8d	lkaz	lw7	lqlj	lvhr	2rn2	
lb0u	lc7s	lde0	lE98	lf9c	lgdd	lhnc	lj8f	lkb3	lwh	lqlv	lvld	2rsl	
lb0z	lc7t	lde1	lE99	lf9e	lgde	lhnd	lj8m	lkbb	lwi	lqlw	lviv	2rus	
lb14	lc83	lde2	le9a	lf9m	lgdh	lhnh	lj8q	lkbi	lwj	lqm5	lvjs	2rve	
lb15	lc84	lde4	le9b	lf9t	lgdr	lhni	lj8v	lkbj	lwn	lqmg	lvkl	2sbt	
lb16	lc85	lde5	le9c	lf9u	lgdt	lhnj	lj8w	lkbk	lwo	lqmh	lviz	2scu	

APPENDIX D

Names of proteins in $\alpha+\beta$ class in SCOP; <http://scop.mrc-lmb.cam.ac.uk>

0cdi	1ayw	1c0g	1dc4	1ek8	1fo1	1gre	1i95	1jpd	1102	1mdr	1qpv	1vad	2pil
0rst	1ayy	1c0i	1dc5	1ekg	1fo4	1grf	1i96	1jpf	1103	1mdt	1qqd	1vao	2pld
0tel	1ayz	1c0k	1dc6	1ekm	1fo6	1grg	1i97	1jpg	1104	1mee	1qqj	1var	2ple
1021	1az9	1c0l	1dch	1ekr	1fo7	1gri	1i9a	1jpm	1105	1meg	1qqq	1vcb	2pna
1031	1azf	1c0p	1dcj	1eks	1foh	1grj	1i9y	1jpo	1106	1mel	1qqr	1vcc	2pnb
1041	1azp	1c10	1dco	1ekt	1foi	1grl	1i9z	1jq4	1107	1mem	1qqy	1vde	2pol
1071	1azq	1c16	1dcp	1ekz	1foj	1grt	1ia8	1jqg	1108	1mff	1qr0	1vew	2prf
1081	1azs	1c1y	1dcq	1el0	1fol	1gsa	1ia9	1jqh	1109	1mfi	1qr1	1vfb	2pro
1091	1azy	1c21	1dd3	1el1	1foo	1gsh	1iaa	1jqj	110c	1mgs	1qr5	1vfr	2ptk
1101	1b02	1c22	1dd4	1el5	1fop	1gsn	1iab	1jq1	110o	1mhc	1qrk	1vhh	2ptl
1111	1b04	1c23	1dd5	1el6	1fot	1gso	1iac	1jqu	110v	1mhd	1qrm	1vhi	2pva
1121	1b06	1c24	1dd6	1el7	1fpu	1gsp	1iad	1jqw	110w	1mhe	1qs1	1vig	2pyp
1131	1b08	1c27	1dd7	1el8	1fpy	1gsv	1iae	1jr1	110x	1mi2	1qs2	1vih	2pyr
1141	1b0d	1c30	1ddh	1el9	1fq1	1gsw	1iag	1jr9	110y	1mif	1qs5	1vjw	2qil
1151	1b0g	1c3a	1ddk	1eli	1fqv	1gsx	1iah	1jrk	1110	1mil	1qs9	1vkl	2rat
1181	1b0p	1c3i	1ddo	1elo	1fr5	1gt8	1iaj	1jrm	1111	1mir	1qsa	1vmp	2rbi
1191	1b0r	1c3o	1ddt	1els	1frd	1gte	1iak	1jro	1112	1mit	1qsb	1vok	2reb
11as	1b0t	1c3t	1ddu	1em1	1frh	1gth	1ian	1jrp	1113	1mka	1qse	1vol	2rgf
11ba	1b0v	1c43	1de3	1em2	1fri	1gtp	1iao	1jrq	1114	1mkb	1qsf	1vr2	2rln
11bg	1b10	1c44	1de4	1em7	1frj	1gtq	1ias	1jru	1115	1mla	1qsm	1vza	2rnf
1201	1b1e	1c45	1de8	1ema	1frk	1gtt	1ib1	1jrx	1116	1mlc	1qsn	1vzb	2rns
1221	1b1i	1c46	1de9	1emb	1frl	1gua	1ib6	1jry	1117	1mld	1qso	1vzc	2rnt
1231	1b1j	1c47	1def	1emc	1frm	1gup	1ib7	1jrz	1118	1mli	1qsq	1vzd	2rus
1251	1b1z	1c4a	1dej	1emd	1fro	1guq	1ib8	1js0	1119	1mmb	1qsr	1vze	2sak
1261	1b20	1c4c	1der	1eme	1frf	1guw	1iba	1jsb	1120	1mmm	1qst	1wdb	2sar
1271	1b21	1c4f	1det	1emf	1frs	1guy	1ibt	1jse	1121	1mmp	1qsw	1wfc	2scu
1281	1b24	1c4g	1deu	1emg	1frt	1guz	1ibu	1jsf	1122	1mmq	1qt3	1wqm	2sdf
1291	1b25	1c4k	1dew	1emk	1frx	1gv0	1ibv	1jsk	1123	1mmr	1qt4	1wqn	2seb
12as	1b27	1c4p	1df0	1eml	1fs1	1gv1	1ibw	1jss	1124	1mms	1qt5	1wqo	2sec
1301	1b2k	1c4z	1df1	1emm	1fs2	1gv4	1ibx	1jst	1125	1mnc	1qt6	1wqp	2shp
1311	1b2m	1c54	1dfa	1ems	1fs3	1gv7	1ic4	1jsu	1126	1mng	1qt7	1wqq	2sic
1321	1b2s	1c60	1dfd	1emv	1fsi	1gw5	1ic5	1jsv	1127	1mnl	1qt8	1wqr	2sni
1331	1b2t	1c61	1dff	1emw	1fsk	1gwd	1ic7	1jtg	1128	1mmm	1qt9	1xan	2src
1341	1b2u	1c62	1dfj	1en4	1fsz	1gx1	1icc	1jtm	1129	1mns	1qtb	1xdt	2srt
1351	1b2v	1c63	1dga	1en5	1ft8	1gx3	1icf	1jtn	112n	1mol	1qtc	1xei	2std
1371	1b2x	1c64	1dgj	1en6	1ftc	1gxd	1icj	1jto	112w	1mom	1qtd	1xej	2sxl
1381	1b2z	1c65	1dgs	1en7	1fte	1gxj	1icr	1jtp	1130	1mpg	1qte	1xek	2tbd
1391	1b33	1c66	1dgt	1enf	1ftf	1gxk	1icu	1jtr	1131	1mpy	1qth	1xer	2tcl
1401	1b37	1c67	1dhn	1eo6	1ftg	1gxl	1icv	1jtr	1132	1mra	1qtk	1xgm	2tdd
1411	1b38	1c68	1dhn	1eod	1ftr	1gy3	1icw	1jtt	1133	1mrg	1qto	1xgn	2tdm
1421	1b39	1c69	1dhy	1eoe	1fu0	1gy5	1icx	1jtu	1134	1mrh	1qtp	1xgo	2tec
1431	1b3a	1c6a	1di2	1eof	1fu5	1gy6	1id0	1ju6	1135	1mri	1qts	1xgs	2tld

144l	1b3d	1c6b	1di3	1eos	1fu6	1gy7	1ids	1juf	1l36	1mrj	1qtv	1xps	2tli
145l	1b3j	1c6c	1di4	1eot	1fug	1gyb	1ie0	1jug	1l37	1mrk	1qtz	1xpt	2tlx
146l	1b3o	1c6d	1di5	1eov	1fus	1gyf	1ie3	1juj	1l38	1mro	1qu6	1xra	2tmn
147l	1b3q	1c6e	1di8	1eow	1fut	1gyp	1ie7	1jut	1l39	1msb	1qu9	1xrb	2toh
148l	1b3s	1c6f	1di9	1eps	1fuw	1gyq	1iea	1jv2	1l3a	1msc	1qua	1xrc	2tpr
149l	1b3t	1c6g	1dih	1epw	1fv1	1gzp	1ieb	1jvi	1l3f	1msd	1qud	1xte	2tpt
150l	1b47	1c6h	1dii	1epy	1fva	1gzq	1iee	1jvp	1l3k	1msg	1qug	1xxa	2tsc
151l	1b49	1c6i	1dik	1eq0	1fvq	1h0g	1iep	1jvt	1l3l	1msh	1quh	1xxb	2tsr
152l	1b4a	1c6j	1diq	1eq3	1fvi	1h0i	1iet	1jvu	1l3r	1mst	1quo	1xxc	2tss
153l	1b4b	1c6k	1div	1eq4	1fvq	1h19	1ieu	1jvv	1l40	1mst	1qup	1yal	2ula
154l	1b4n	1c6l	1dix	1eq5	1fvr	1h2h	1iez	1jvw	1l41	1muc	1qus	1yal	2u2f
155l	1b4o	1c6m	1diz	1eq6	1fvs	1h4l	1ifq	1jw3	1l42	1mut	1qut	1yam	2ubp
156l	1b4s	1c6n	1dj0	1eqe	1fvt	1h4q	1ifs	1jw9	1l43	1mva	1raa	1yan	2ucz
157l	1b4v	1c6p	1dk0	1eqk	1fvu	1h4r	1ift	1jwa	1l44	1mva	1rab	1yao	2ugi
158l	1b50	1c6q	1dkh	1eqm	1fvv	1h4s	1ifu	1jwb	1l45	1mwp	1rac	1yap	2upl
159l	1b53	1c6t	1dki	1eqo	1fvz	1h4t	1ifv	1jwh	1l46	1mxa	1rad	1yaq	2ush
160l	1b59	1c76	1dkj	1eqr	1fw6	1h4u	1igl	1jwi	1l47	1mxb	1rae	1yat	2usn
161l	1b5a	1c77	1dkk	1eqt	1fw9	1h4v	1ig4	1jwj	1l48	1mxc	1raf	1yes	2uug
162l	1b5b	1c78	1dks	1eqy	1fwa	1h52	1ige	1jwk	1l49	1myo	1rag	1ydr	2vaa
163l	1b5d	1c79	1dkt	1erk	1fwb	1h53	1igd	1jwo	1l50	1nap	1rah	1yds	2vab
164l	1b5e	1c7h	1dl5	1erz	1fwc	1h6a	1ihb	1jwr	1l51	1naw	1rai	1ydt	2vao
165l	1b5m	1c7k	1dlg	1es0	1fwd	1h6b	1ihr	1jww	1l52	1nbe	1rar	1yer	2vik
166l	1b5q	1c7p	1dlh	1esl	1fwe	1h6c	1ihx	1jx0	1l53	1ncl	1ras	1yes	2vil
167l	1b5u	1c7s	1dm0	1esf	1fwf	1h6d	1ihy	1jx1	1l54	1ncv	1rat	1yet	2yfp
168l	1b5v	1c7t	1dm2	1esl	1fwg	1h6h	1ii7	1jxa	1l55	1nda	1rax	1yfp	2ypn
169l	1b5w	1c7u	1dm6	1esp	1fwh	1h6k	1iic	1jxv	1l56	1ndc	1rba	1ygh	2znb
16vp	1b5x	1c8c	1dm7	1esr	1fwi	1h6m	1iid	1jyl	1l57	1nnd	1rbh	1ygw	3aig
170l	1b5y	1c8t	1dm8	1esv	1fwj	1h6r	1iiz	1jy5	1l58	1ndk	1rbc	1ypa	3aop
171l	1b5z	1c8u	1dm9	1et6	1fwk	1h6v	1ije	1jy8	1l59	1ndl	1rbd	1ypb	3atl
172l	1b62	1c8w	1dma	1et9	1fwl	1h6z	1ijf	1jya	1l5g	1ndo	1rbe	1ypc	3ayk
173l	1b63	1c8z	1dmi	1eu3	1fwp	1h70	1ijh	1jyh	1l5p	1ndp	1rbf	1ypn	3be2
174l	1b64	1c9b	1dmj	1eu4	1fx2	1h72	1ijk	1jym	1l60	1nec	1rbg	1ypr	3bcc
175l	1b65	1c9f	1dmk	1eub	1fx3	1h73	1ijr	1jyo	1l61	1ned	1rbh	1ytb	3bir
176l	1b66	1c9h	1dml	1euc	1fx4	1h74	1ikl	1jyq	1l62	1nel	1rbi	1ytf	3bta
177l	1b69	1c9v	1dmm	1eud	1fxa	1h7s	1ikm	1jyr	1l63	1nfi	1rbj	1ytt	3bu4
178l	1b6a	1c9x	1dmm	1eue	1fxd	1h7u	1ikn	1jyu	1l64	1nhb	1rbl	1yvn	3ci2
179l	1b6b	1ca5	1dmq	1eui	1fxh	1h7w	1ikp	1k0e	1l65	1nhk	1rbn	1yvs	3cox
180l	1b6c	1ca6	1dmt	1euv	1fxi	1h7x	1ikq	1k0g	1l66	1nhp	1rbo	1zag	3crx
181l	1b6e	1ca7	1dmw	1ev0	1fxl	1h81	1ikt	1k0i	1l67	1nhq	1rbw	1zdh	3dap
182l	1b6f	1caq	1dna	1ev5	1fxr	1h82	1il2	1k0j	1l68	1nhr	1rbx	1zdi	3dbv
183l	1b6i	1cb5	1dnc	1ev8	1fxt	1h83	1il3	1k0k	1l69	1nhs	1rca	1zdz	3dni
184l	1b6r	1cbk	1dnk	1evf	1fxv	1h84	1il4	1k0l	1l6j	1nja	1rck	1zdk	3eip
185l	1b6s	1cbo	1do9	1evg	1fy7	1h86	1il5	1k0r	1l6u	1njb	1rcd	1zer	3enl
186l	1b6v	1cby	1dob	1evi	1fyf	1h87	1il8	1k0t	1l6v	1njc	1rcm	1zjf	3erk
187l	1b6z	1cc2	1doc	1evj	1fyr	1h8c	1il9	1k0v	1l70	1njd	1rcn	1zfp	3eza
188l	1b70	1cc4	1dod	1evk	1fys	1h8f	1ilp	1k1a	1l71	1nje	1rco	1znb	3ezb
189l	1b76	1cc6	1doe	1evl	1fyt	1h8j	1ilq	1k1b	1l72	1nlk	1rcx	1zpr	3eze
190l	1b77	1cc7	1doi	1evu	1fza	1h8m	1im3	1k1c	1l73	1nmt	1rdi	200l	3fap
191l	1b7f	1cc8	1dok	1evw	1fzb	1h8u	1im9	1k1g	1l74	1noc	1rdj	201l	3fib

192l	1b7g	1cd1	1dol	1evx	1fzc	1h8x	1imu	1k25	1175	1nod	1rdk	205l	3fit
193l	1b7l	1cdk	1dom	1ew0	1fzd	1h93	1inn	1k26	1176	1nos	1rdl	206l	3fru
194l	1b7m	1cdw	1don	1ew4	1fze	1h94	1inq	1k28	1177	1nox	1rdm	207l	3gbl
195l	1b7n	1ce6	1dox	1ewc	1fzf	1h98	1inu	1k2a	1179	1npc	1rdn	208l	3gcb
196l	1b7o	1ce7	1doy	1ewf	1fzg	1h9a	1io5	1k2e	117y	1npk	1rdo	209l	3gcc
197l	1b7p	1ce8	1dp6	1ewj	1fzj	1h9b	1iod	1k2p	1180	1npv	1rds	210l	3geo
198l	1b7q	1ceq	1dp8	1ewp	1fzk	1h9o	1iok	1k3a	1181	1nrc	1re2	211l	3gpd
199l	1b7r	1cer	1dp9	1ewq	1fzm	1ha1	1ioo	1k3b	1182	1nsa	1rea	212l	3grs
1a06	1b7s	1cet	1dpg	1ewy	1fzo	1ha5	1ioq	1k3e	1183	1nse	1reg	213l	3grt
1a07	1b7y	1cew	1dpt	1ex8	1fzu	1ha6	1ior	1k3s	1184	1nsg	1rem	214l	3gsp
1a08	1b87	1cf0	1dpw	1exb	1fzy	1han	1ios	1k3t	1185	1nsi	1rex	215l	3hck
1a09	1b8a	1cf2	1dpx	1exi	1g02	1hbm	1iot	1k44	1186	1nsk	1rey	216l	3hfl
1a0i	1b8h	1cf3	1dq3	1exj	1g05	1hbn	1iov	1k46	1187	1nsp	1rez	217l	3hfm
1a0k	1b8p	1cf5	1dq8	1exu	1g06	1hbo	1iow	1k47	1188	1nsq	1rfa	218l	3hla
1a0o	1b8s	1cfe	1dq9	1eym	1g07	1hbu	1ip1	1k49	1189	1nue	1rfp	219l	3hoh
1a0p	1b8u	1cfy	1dqa	1eyn	1g0d	1hbx	1ip2	1k4i	118b	1oac	1rga	220l	3il8
1a15	1b8v	1cg2	1dqj	1eyp	1g0g	1hby	1ip3	1k4j	118j	1obs	1rgc	221l	3jdw
1a16	1b8y	1cge	1drg	1eyq	1g0j	1hc7	1ip4	1k4l	118p	1obt	1rge	222l	3kbp
1a1a	1b99	1cgf	1drm	1eyz	1g0k	1hcj	1ip5	1k4n	118t	1oel	1rgf	223l	3kmb
1a1b	1b9c	1cgl	1drq	1ez1	1g0l	1hck	1ip6	1k4o	118u	1ofg	1rgg	224l	3kvt
1a1c	1b9k	1cgq	1dru	1ez4	1g0m	1hcl	1ip7	1k4p	1190	1ojt	1rgh	225l	3lad
1a1e	1b9l	1ch0	1drv	1ezm	1g0p	1hcs	1ip9	1k4q	1191	1onc	1rgk	226l	3lek
1a1m	1b9o	1chk	1drw	1ezv	1g0q	1hct	1ipa	1k50	1192	1one	1rgl	227l	3lhm
1a1n	1b9r	1chm	1drz	1ezz	1g0s	1hd0	1ipb	1k51	1193	1opd	1rha	228l	3lym
1a1o	1ban	1chr	1ds5	1f00	1g0u	1hd1	1ipc	1k52	1194	1opy	1rhh	229l	3lyo
1a26	1bao	1chu	1ds7	1f02	1g0w	1hd7	1ipg	1k53	1195	1ord	1rhl	230l	3lyt
1a2k	1bb3	1ci7	1dss	1f03	1g10	1hdg	1ipp	1k58	1196	1osz	1rhp	231l	3lyz
1a2n	1bb4	1ciq	1dsx	1f04	1g11	1hdm	1iq0	1k59	1197	1otf	1ris	232l	3lz2
1a2p	1bb5	1cir	1dt0	1f06	1g12	1hdn	1iq4	1k5a	1198	1otg	1rl6	233l	3lzm
1a2v	1bb6	1civ	1dt4	1f08	1g18	1hdw	1iqo	1k5b	1199	1otp	1rlc	234l	3lzt
1a2w	1bb7	1ciz	1dt9	1f09	1g19	1he0	1iqq	1k5h	119a	1oua	1rld	235l	3mat
1a2y	1bb8	1cjl	1dth	1f0b	1g1b	1he6	1iqs	1k6o	119f	1oub	1rlf	236l	3mds
1a31	1bbu	1cj2	1dtj	1f0i	1g1q	1he8	1iqt	1k6z	119m	1ouc	1rls	237l	3mon
1a35	1bbw	1cj3	1dtn	1f0l	1g1r	1hel	1iqz	1k7j	119n	1oud	1rms	238l	3muc
1a36	1bbx	1cj4	1dtp	1f0q	1g1s	1hem	1ir0	1k7y	119v	1oue	1rnl	239l	3nod
1a3a	1bc2	1cj6	1duj	1f0v	1g1t	1hen	1ir1	1k83	11a2	1ouf	1rm4	240l	3nos
1a3s	1bc4	1cj7	1dur	1f0w	1g1v	1heo	1ir2	1k8a	11aa	1oug	1rmb	241l	3nse
1a41	1bc6	1cj8	1duy	1f0x	1g1w	1hep	1ir3	1k8d	11ba	1ouh	1rnc	242l	3nul
1a48	1bcc	1cj9	1duz	1f0z	1g1x	1heq	1ir7	1k8i	11bu	1oui	1rmd	243l	3pah
1a4h	1bcj	1cja	1dv1	1f10	1g24	1her	1ir8	1k8k	11c0	1ouj	1rmf	244l	3pax
1a4v	1bcj	1cje	1dv2	1f13	1g27	1hew	1ir9	1k8r	11c3	1oun	1rmm	245l	3pbh
1a4y	1bcp	1cjf	1dv8	1f1b	1g28	1hez	1irk	1k8w	11ca	1p1g	1rmn	246l	3phy
1a5e	1bd2	1cjk	1dvc	1f1h	1g2a	1hf4	1iru	1k92	11cb	1p32	1rmo	247l	3pmg
1a5k	1bd6	1cjl	1dvd	1f1w	1g2e	1hfe	1is0	1k97	11ce	1p38	1rmq	248l	3pro
1a5l	1bd8	1cjq	1dw9	1f28	1g2r	1hfe	1is7	1k98	11cj	1pad	1rmt	249l	3pva
1a5m	1bdf	1cjr	1dwm	1f29	1g2s	1hfg	1is8	1k9a	11ck	1paf	1rmu	250l	3pyp
1a5n	1bdl	1cjt	1dwm	1f2a	1g2t	1hfn	1isa	1k9i	11cn	1pag	1rmv	251l	3rat
1a5o	1bdu	1cju	1dwn	1f2b	1g3i	1hfo	1isb	1k9j	11co	1pah	1rmw	252l	3rhn
1a5p	1be3	1civ	1dvw	1f2c	1g3k	1hfs	1isc	1k9m	11d9	1pax	1rmx	253l	3rm3

1a5q	1be4	1cjlw	1dww	1f2f	1g3n	1hfx	1isk	1ka6	1ldb	1pba	1lmy	254l	3rnt
1a5r	1bel	1cjl	1dwx	1f2g	1g3o	1hfy	1isn	1ka7	1ldc	1pbb	1lmz	255l	3rsd
1a5z	1bf3	1ck1	1dwy	1f2h	1g49	1hfx	1it6	1kaf	1ldg	1pbc	1rob	256l	3rsk
1a67	1bf4	1ck2	1dwz	1f2k	1g4a	1hh2	1itp	1kap	1ldk	1pbd	1rod	257l	3rsp
1a68	1bf5	1ck5	1dx0	1f2l	1g4b	1hhg	1itx	1kbc	1ldm	1pbe	1roe	258l	3rub
1a6a	1bf1	1ck7	1dx1	1f2r	1g4k	1hhh	1iu4	1kbg	1ldn	1pbf	1rof	259l	3sdp
1a6d	1bfj	1ck8	1dxj	1f31	1g51	1hhi	1ius	1kbi	1ldp	1pbh	1rot	260l	3seb
1a6e	1bfp	1ck9	1dxk	1f32	1g57	1hhj	1iut	1kbl	1leg	1pbk	1rou	261l	3sic
1a6f	1bfx	1ckc	1dxl	1f34	1g58	1hhk	1iuu	1kbp	1lek	1pca	1rpf	262l	3srn
1a6j	1bg0	1ckd	1dy0	1f3c	1g5b	1hhl	1iuv	1kbu	1leo	1pcf	1rpg	2a5e	3ssi
1a6l	1bg1	1ckf	1dy1	1f3f	1g5h	1hhq	1iuw	1kc2	1lew	1pch	1rph	2aad	3std
1a6q	1bgi	1ckg	1dy2	1f3j	1g5i	1hi2	1iux	1kc7	1lez	1pci	1rra	2aae	3sxl
1a6r	1bgj	1ckh	1dy3	1f3m	1g5s	1hi3	1ivm	1kce	1lfd	1pcz	1rrb	2aai	3tec
1a6z	1bgn	1cki	1dy5	1f3p	1g5u	1hi4	1ivu	1kcg	1lg7	1pda	1rsc	2aak	3tli
1a70	1bgo	1ckj	1dya	1f3v	1g61	1hi5	1ivv	1kcm	1lgr	1pdh	1rsm	2aas	3tmn
1a73	1bgq	1ckm	1dyb	1f3y	1g62	1hiy	1ivw	1kcq	1lgu	1pdx	1rsn	2abl	3tms
1a74	1bgs	1ckn	1dyc	1f40	1g65	1hka	1ivx	1kcz	1lgw	1pdy	1rta	2acg	3tss
1a79	1bgy	1cko	1dyd	1f44	1g6b	1hla	1iw7	1kd0	1lgx	1pdz	1rtb	2act	3ubp
1a7g	1bh5	1ckp	1dye	1f46	1g6j	1hll	1ixm	1kdl	1lhh	1pe6	1rtc	2acy	3usn
1a7k	1bhf	1cks	1dyf	1f47	1g6r	1hlp	1ixx	1kdn	1lhi	1pfd	1rtm	2afp	3znb
1a7t	1bhh	1ckv	1dyg	1f4b	1g6s	1hhr	1iy3	1ke5	1lhj	1pfn	1rtu	2ahj	456c
1a7x	1bhn	1clf	1dyq	1f4c	1g6t	1hlu	1iy4	1ke6	1lhk	1pfl	1rto	2aig	4aig
1a81	1bhy	1cli	1dyt	1f4d	1g6z	1hlw	1j4g	1ke7	1lhl	1pfm	1rts	2aim	4aop
1a85	1bh2	1cm0	1dyu	1f4e	1g7c	1hmj	1j4r	1ke8	1lhm	1pfn	1rtu	2ang	4atl
1a86	1bi7	1cm2	1dz1	1f4f	1g7h	1hmk	1j4w	1ke9	1li1	1pga	1rus	2aop	4ayk
1a8r	1bi8	1cm3	1dz5	1f4g	1g7i	1hml	1j5e	1kee	1li2	1pgb	1ruv	2atl	4bir
1a8t	1bia	1cm8	1dza	1f51	1g7j	1hnl	1j5k	1keh	1li3	1pgx	1rxo	2atc	4bu4
1a90	1bib	1cm9	1dzb	1f52	1g7k	1hnw	1j5p	1kek	1li6	1phh	1ryp	2aw0	4crx
1a9b	1bid	1cmi	1dzf	1f5a	1g7l	1hnz	1j5u	1kf2	1lit	1phk	1sap	2axx	4dbv
1a9c	1bid	1cmk	1dzn	1f5b	1g7m	1hnz	1j5w	1kf3	1ljn	1phz	1sar	2ayk	4enl
1a9e	1bir	1cmx	1dzo	1f5c	1g7p	1ho5	1j5y	1kf4	1lk5	1pic	1sat	2baa	4erk
1a9n	1biw	1cn7	1dzs	1f5m	1g7q	1hoc	1j6r	1kf5	1lk7	1pil	1sbb	2bb8	4fap
1a9u	1bix	1cn8	1E01	1f5q	1g81	1hov	1j6v	1kf6	1lkk	1pin	1sbn	2bbq	4fit
1a9x	1bj3	1cn9	1E09	1f5v	1g83	1how	1j6w	1kf7	1lkl	1pip	1sce	2bc2	4fxc
1aa1	1bjg	1cns	1e0b	1f60	1g8t	1hpl	1j6x	1kf8	1lkr	1pkp	1scj	2bcc	4gep
1aa3	1bjp	1cnu	1e0g	1f6r	1g91	1hp4	1j6y	1kfi	1lks	1plf	1scu	2bef	4gpd
1aa8	1bk7	1coa	1e0y	1f6s	1g96	1hp5	1j74	1kfq	1lkz	1plq	1sdf	2bir	4grl
1aar	1bke	1cof	1e0z	1f6t	1g9m	1hpu	1j7a	1kfu	1llc	1plr	1se2	2bkj	4grt
1ab2	1bkf	1coj	1E10	1f77	1g9n	1hpw	1j7b	1kfx	1lld	1pma	1se3	2bmi	4gsp
1ab8	1bkj	1com	1E15	1f71	1g9q	1hq0	1j7c	1kfy	1llh	1pmd	1se4	2bop	4hoh
1abm	1bkj	1coy	1e1g	1f7s	1g9y	1hq6	1j7d	1kg0	1llw	1pme	1seb	2bow	4jdw
1abr	1bkl	1cp9	1elj	1f7t	1g9z	1hq8	1j7i	1kh0	1llz	1pne	1sei	2bqa	4kbp
1acb	1bkm	1cpj	1elo	1f7u	1ga7	1hqi	1j7l	1kh1	1lm1	1pnk	1ser	2bqb	4kmb
1acf	1bkn	1cpz	1elp	1f7v	1gad	1hqr	1j7n	1kh2	1lm5	1pnl	1ses	2bqc	4lym
1acm	1bko	1cq4	1eis	1f7w	1gae	1hgy	1j7u	1khm	1lm7	1pnm	1set	2bqd	4lyo
1ad5	1bkp	1cqa	1elt	1f7x	1gag	1hqz	1j7z	1kif	1lm8	1poh	1sgk	2bqe	4lyt
1adj	1bkw	1cqd	1elu	1f80	1gal	1hr0	1j80	1kij	1lma	1pop	1sha	2bqf	4lyz
1ady	1bkx	1cqi	1elv	1f82	1gao	1hr6	1j81	1kip	1lmc	1ppd	1shb	2bqg	4lzm
1ae9	1bl4	1cqj	1elw	1f83	1gaq	1hr7	1j82	1kiq	1lmh	1ppn	1shd	2bqh	4lzt

laec	lbl6	lcqm	le1x	lf89	lgav	lhr8	lj8c	lkir	llml	lppo	lsib	2bqi	4mat
laer	lbl7	lcqn	1E21	lf8a	lgay	lhr9	lj8h	lkj2	llmn	lppp	lskj	2bj	4mdh
laf0	lblj	lcqr	1E22	lf8q	lgaz	lhrj	lj8i	lkj3	llmo	lpre	slm	2bqk	4mon
laf5	lblk	lcqu	1E24	lf8r	lgb0	lhru	lj8t	lkj8	llmp	lprq	sln	2bql	4nos
lafa	lblu	lcqv	1E27	lf8s	lgb1	lhs6	lj8u	lkj9	llmq	lprt	lsly	2bqm	4nse
lafb	blv	lcr5	1E28	lf8w	lgb2	lhsa	lj8w	lkji	llmt	lpsd	lsml	2bqn	4ota
lafd	blx	lcrk	1e2t	lf93	lgb3	lhsb	lj8x	lkjj	lln1	lptf	lsnm	2bqo	4otb
lafi	lbn2	lcrw	1E32	lf95	lgb4	lhsk	lj91	lkjk	lln2	lpto	lsmp	2btf	4otc
lafj	lbn6	lcrx	1E39	lf96	lgb5	lhsw	lj92	lkjq	lln3	lpuc	lsox	2bu4	4pad
lafk	lbn8	lcs0	1e3a	lf98	lgb6	lhsx	lj94	lkjt	lln4	lput	lspb	2cb5	4pah
lafl	lbnb	lcs3	1e3h	lf9c	lgb7	lht1	lj98	lkke	llna	lpxa	lsph	2cbl	4pax
lafu	lbnc	lcs4	1e3m	lf9f	lgb8	lht2	lj9o	lkkh	llnb	lpxb	lspr	2chr	4pro
lag2	lbnd	lcs8	1e3n	lf9i	lgb9	lhtd	lja2	lkkl	llnc	lpxc	lsps	2chs	4rat
lagb	lbmk	lcsb	1e3p	lf9j	lgb0	lhtl	lja3	lkkm	llnd	lpya	lspu	2cht	4rhn
lagc	lbml	lcse	1e3r	lf9n	lgb5	lhtn	lja4	lkko	llne	lpys	lsrn	2ci2	4rnt
lagd	lbms	lcsn	1e3v	lf9z	lgbw	lhto	lja6	lkkp	llnf	lpyt	lsrp	2cjn	4rsd
lage	lbn5	lcsy	1E42	1fa0	lgbx	lhtq	lja7	lklg	llng	lqad	lsrs	2cjo	4rsk
lagf	lbnc	lcsz	1E44	1fa3	lgby	lhtt	lja8	lklf	llni	lqae	lsry	2ckb	4rub
lagi	lbne	lct9	1e4e	1fa5	lgbz	lhuc	ljak	lklq	llnu	lqaf	lssa	2clr	4srn
lagw	lbnf	lcte	1e5d	1fa6	lgcl	lhuf	ljam	lklu	llo7	lqak	lssb	2cmd	4std
lah5	lbng	lctf	1e5l	1fa7	lgc6	lhum	ljan	lkmb	llo8	lqal	lssc	2cpk	4tli
lah6	lbni	lctn	1e5q	1fa8	lgc7	lhun	ljao	lkmd	llo9	lqap	lssn	2crk	4tln
lah8	lbnj	lctp	1e5u	1fap	lgcb	lhup	ljap	lkmm	llou	lqax	lssu	2crx	4tmn
laha	lbnl	lctw	1e5w	1fb1	lgcc	lhuy	ljq	lkmn	llov	lqay	lsss	2csn	4tms
lahb	lbnr	lcu0	1e6e	1fb2	lgcu	lhv2	ljat	lkmy	llow	lqb3	lstc	2dap	4tss
lahc	lbns	lcu2	1e6n	1fb1	lgcz	lhv5	ljaw	lkmz	lloy	lqba	lstf	2dbv	4tsu
lahj	lbnz	lcu3	1e6p	1fbv	lgd0	lhvy	ljb0	lkn6	lloz	lqbb	lste	2def	4ubp
lahq	lbo0	lcu5	1e6r	1fbx	lgd1	lhw3	ljb2	lkna	llp4	lqbe	lstf	2dhn	4znb
lahu	lbo1	lcu6	1e6t	1fbz	lgd3	lhw4	ljb4	lknd	llpf	lqcf	lstu	2dik	53BP2
lahv	lbo4	lcu1	1e6v	1fca	lgd4	lhw7	ljb5	lkne	llpi	lqcg	lsvq	2dln	5aop
lahz	lbo7	lcup	1e6y	1fcb	lgd5	lhw8	ljb6	lknf	llpq	lqci	lsvr	2dnj	5at1
lai4	lbo8	lcuq	1e6z	1fcc	lgd6	lhw9	ljb1	lkni	llpu	lqcj	lsvy	2dpg	5bir
lai5	lboa	lcu0	1E77	1fcd	lgd8	lhw1	ljb2	lknj	llpy	lqcn	lsu6	2e2c	5bu4
lai6	lbob	lcv1	1e7d	1fcl	lgdg	lhwj	ljb3	lknk	llqb	lqco	lsxl	2efg	5crx
lai7	lbol	lcv3	1e7k	1fd2	lgdo	lhwk	ljb5	lknq	llqv	lqcg	lsxt	2emd	5enl
laih	lbow	lcv4	1e7l	1fd6	lgdw	lhw1	ljb7	lknr	llqw	lqcr	lsyn	2emn	5fd1
laim	lbox	lcv5	1e7m	1fd8	lgdx	lhw2	ljb4	lko9	llqy	lqcs	lszj	2emo	5fit
laip	lbp0	lcv6	1e7p	1fd9	lge0	lhw3	ljb5	lkoa	llr0	lqd1	ltld	2eot	5gcn
laiq	lbp1	lcv8	1e7u	1fda	lge1	lhw4	ljb9	lkob	llr4	lqd2	ltay	2eql	5gep
lais	lbp4	lcvj	1e7v	1fdb	lge2	lhw5	ljb1	lkoe	llra	lqd9	ltba	2erk	5grt
laj6	lbp6	lcvk	1e7x	1fdd	lge3	lhx3	ljb2	lkog	llrm	lqdd	ltbd	2fap	5gsp
lajm	lbp7	lcvz	1e7y	1fd1	lge4	lhx4	ljb3	lkoh	llru	lqdl	ltbe	2fd2	5hoh
lajn	lbpj	lcwd	1E87	1fdn	lge5	lhx5	ljb4	lkoo	llry	lqdn	ltbp	2fdn	5jdw
lajp	lbq1	lcwe	1e8f	1fe0	lge6	lhx6	ljb5	lkor	llsa	lqdq	ltby	2fgi	5ldh
lajq	lbq2	lcwv	1e8g	1fe4	lge7	lhx7	ljb6	lkot	llsb	lqdr	ltce	2fhi	5lym
lak6	lbqb	lcx0	1e8h	1fea	lge8	lhx8	ljb7	lkou	llsc	lqdt	ltco	2ffb	5lyt
lak7	lbqg	lcx6	1e8i	1feb	lge9	lhy7	ljb8	lkp0	llsd	lqdv	ltcs	2fit	5lyz
laki	lbqh	lcx7	1e8l	1fec	lgec	lhye	ljb9	lkp2	llse	lqdw	ltcy	2fke	5lzm
lakj	lbqi	lcxv	1e8o	1fee	lgeh	lhyf	ljb1	lkp3	llsf	lqe0	ltda	2fmr	5mdh

1akl	1bql	1cxy	1e8s	1feh	1ger	1hyg	1jdg	1kp5	1lsg	1qe6	1tdb	2gac	5msf
1ako	1bqo	1cyo	1e8w	1fer	1ges	1hyh	1jdw	1kp6	1lsm	1qf0	1tdc	2gaw	5nse
1alc	1bqq	1cyq	1e8x	1fes	1get	1hym	1jdx	1kpa	1lsn	1qf1	1tdj	2gbl	5pad
1aml	1bqx	1cyu	1e8y	1fev	1geu	1hyn	1jdy	1kpb	1lsp	1qf2	1tdu	2gcc	5pah
1am7	1br5	1cyv	1e8z	1ff2	1gev	1hyo	1je3	1kpc	1lsq	1qf6	1tdy	2gd1	5rat
1amw	1br6	1cz0	1E90	1ff3	1gez	1hyr	1je4	1kpe	1lsy	1qfc	1tec	2gep	5rhn
1an5	1brg	1cz4	1E94	1ff9	1gf0	1hyt	1je6	1kpf	1lsz	1qfr	1tew	2gls	5rnt
1an7	1brh	1cz5	1E97	1ffg	1gf3	1hyw	1jef	1kpp	1lt3	1qg1	1tfe	2glt	5rsa
1an8	1bri	1czd	1e9h	1ffk	1gf4	1hzi	1jeh	1kpq	1lt4	1qg7	1tgh	2gnk	5rub
1an9	1brj	1czp	1e9i	1ffi	1gf5	1hzi	1jem	1kps	1lta	1qgi	1the	2grt	5sic
1ang	1brk	1d06	1e9l	1ffi	1gf6	1hzi	1jen	1kpt	1ltb	1qgw	1thl	2gsp	5std
1ao0	1brm	1d09	1e9m	1ffs	1gf7	1hzi	1jep	1kq4	1ltf	1qh3	1thy	2hcc	5thi
1ao7	1brn	1d0c	1e9n	1ffu	1gf8	1hzi	1jex	1kqb	1ltg	1qh4	1tif	2hck	5thn
1aob	1brs	1d0k	1e9y	1ffv	1gf9	1hzi	1jfl	1kqc	1lth	1qh5	1tig	2hea	5tmn
1aog	1brw	1d0l	1e9z	1ffw	1gfa	1i00	1jff	1kqd	1lti	1qhc	1tii	2heb	5tss
1aon	1bry	1d0m	1ea0	1fg2	1gfe	1i08	1jfi	1kqf	1ltm	1qhk	1tin	2hec	6atl
1aop	1bs2	1d0n	1ea2	1fgi	1gfg	1i09	1jfm	1kqg	1lts	1qhw	1tis	2hed	6cnl
1aor	1bs3	1d0o	1ea6	1fgk	1gfh	1i0h	1jg0	1kqk	1ltt	1qi7	1tjs	2hee	6fd1
1aot	1bs4	1d1j	1eak	1fh0	1gfi	1i0v	1jg5	1kqs	1ltu	1qib	1tki	2hef	6fdr
1aou	1bs5	1d1r	1ear	1fh6	1gfk	1i0x	1jgt	1kr4	1ltv	1qil	1tla	2hgs	6fit
1ap0	1bs6	1d1v	1eay	1fhi	1gfl	1i0z	1jhl	1kra	1ltz	1qin	1tlc	2hid	6gep
1ap5	1bs7	1d1w	1eb0	1fhs	1gfr	1i10	1jh3	1krb	1lus	1qio	1tlg	2hla	6gsp
1ap6	1bs8	1d1x	1eb6	1fht	1gft	1i12	1jh6	1krc	1lur	1qip	1tli	2hlp	6jdw
1ap7	1bsa	1d1y	1ebd	1fhu	1gfu	1i17	1jh7	1krh	1luw	1qj2	1tlp	2hoh	6ldh
1ap8	1bsb	1d1z	1ebf	1fhv	1gfv	1i19	1jhl	1krl	1lv1	1qjd	1tls	2hpr	6lyt
1apa	1bsc	1d2k	1ebg	1fi4	1gg3	1i1a	1jht	1ks2	1lw6	1qjg	1tlx	2hqi	6lyz
1apg	1bsd	1d2w	1ebh	1fib	1gga	1i1d	1jhw	1ksi	1lwu	1qjh	1tmc	2iad	6lzm
1apm	1bse	1d2y	1ebm	1fic	1ggm	1i1e	1ji8	1ksk	1lwx	1qji	1tmn	2ifl	6msf
1aps	1bsf	1d3a	1ebu	1fid	1ggo	1i1f	1jic	1ksl	1lxd	1qjj	1tn3	2ife	6nse
1apy	1bsj	1d3f	1ec6	1fie	1lil	1jld	1jid	1kss	1lxt	1qkl	1tn9	2iff	6pad
1apz	1bsk	1d3j	1ec7	1fff	1ggt	1i1i	1jie	1ksu	1ly7	1qk9	1toh	2igd	6pah
1aq1	1bsm	1d3m	1ec8	1fih	1ggu	1i1q	1jif	1ksv	1lyd	1qkf	1tol	2igg	6rat
1aq3	1bsp	1d3n	1ec9	1fik	1ggx	1i1y	1jis	1ksw	1lye	1qkh	1tox	2igh	6rhn
1aq4	1bsr	1d3u	1ecb	1fil	1ggy	1i1z	1jit	1ksx	1lyf	1qki	1tpt	2ihl	6mt
1aqa	1bsw	1d3w	1ecc	1fim	1gh8	1i20	1jiw	1ksy	1lyg	1qkl	1trg	2i18	6rsa
1aqd	1bsz	1d3z	1ecf	1fin	1ghh	1i21	1jiy	1kt2	1lyh	1qkz	1trp	2jdw	6std
1aqp	1bt0	1d4b	1ecg	1fiq	1ghl	1i22	1jiz	1kt9	1lyi	1ql0	1trq	2jdx	6tli
1aqz	1bt8	1d4c	1ecj	1fit	1ghu	1i2e	1ji0	1ktd	1lyj	1ql6	1ts2	2jel	6tmn
1ar0	1btv	1d4d	1ecq	1fj3	1gif	1i2f	1ji1	1ktg	1lyl	1qla	1ts3	2kau	7at1
1ar4	1bu4	1d4e	1ecs	1fj7	1gih	1i2g	1ji2	1ktk	1lyo	1qlb	1ts4	2kce	7cei
1ar5	1bu9	1d4t	1ed3	1fjc	1gii	1i2g	1ji3	1kuf	1lys	1qlf	1ts5	2kmb	7enl
1aro	1bud	1d4w	1ed4	1fjd	1gij	1i33	1ji4	1kug	1lyy	1qlm	1tsd	2i78	7fd1
1arz	1buh	1d4x	1ed5	1fje	1gio	1i35	1ji9	1kuh	1lyz	1qlt	1tsg	2ldb	7fdr
1ask	1bui	1d5d	1ed6	1fjg	1gio	1i3f	1jic	1kui	1lzl	1qlu	1tsl	2ldx	7gep
1ast	1buj	1d5e	1edq	1fjm	1gfv	1i3i	1jje	1kuk	1lzl	1qlx	1tsm	2lgs	7gsp
1asy	1buo	1d5f	1eej	1fjo	1giz	1i3q	1jih	1kuo	1lzl	1qlz	1tsn	2lhm	7jdw
1asz	1buq	1d5h	1ef0	1fjq	1gk0	1i3r	1jit	1kut	1lzl	1qm0	1tsv	2lym	7lyz
1atl	1buu	1d5j	1ef1	1fjt	1gk1	1i42	1jiw	1kuu	1lzl	1qm1	1tsw	2lyo	7lzm
1at5	1buv	1d5m	1ef2	1fju	1gk8	1i44	1jk3	1kuv	1lzl	1qm2	1tsx	2lyz	7mdh

lat6	lbux	ld5n	lef5	lfjv	lgk9	li4f	ljk7	lkux	llz8	lqm3	ltsy	2lz2	7msf
lati	lbv1	ld5t	lef7	lfjw	lgkc	li4g	ljk8	lkuy	llz9	lqm4	ltsz	2lzh	7nse
latk	lbv4	ld5x	lef9	lfkb	lgkd	li4h	ljk9	lkvl	llza	lqm9	ltub	2lzm	7pck
latl	lbvi	ld5y	lefn	lfkc	lgkf	li4m	ljka	lkv2	llzb	lqma	ltum	2lzt	7rat
latn	lbvk	ld5z	lefu	lfkd	lgkv	li4p	ljkj	lkv3	llzc	lqme	ltvu	2mat	7rnt
latp	lbvq	ld6a	lefw	lfkf	lgkw	li4q	ljkc	lkvd	llzd	lqmf	ltvv	2mbr	7rsa
lau0	lbvt	ld6e	lefx	lfkg	lgkx	li4r	ljkd	lkve	llze	lqmh	ltvw	2mea	7std
lau2	lbvx	ld6o	lefy	lfkh	lgkz	li4x	ljkf	lkvk	llzg	lqmi	ltvx	2meb	7tli
lau3	lbwc	ld6p	leg9	lfki	lgl3	li50	ljkg	lkvn	llzh	lqmr	ltyp	2mec	7tln
lau4	lbwd	ld6q	lega	lfkj	lgl4	li56	ljki	lkvv	llzn	lqmt	ltys	2med	830c
laud	lbwe	ld6t	legg	lfkk	lglv	li58	ljkj	lkvz	llzr	lqmy	ltyt	2mee	8atl
laur	lbwh	ld6u	legi	lfkl	lgm7	li59	ljkk	lkwm	llzs	lqnz	ltze	2mef	8atc
laus	lbwi	ld6y	legl	lfkq	lgm8	li5a	ljkl	lkwt	llzt	lqn3	lu2f	2meg	8cho
lauv	lbwj	ld6z	legp	lfkr	lgm9	li5b	ljkn	lkwu	llzy	lqn4	lu9a	2meh	8gep
laux	lbwv	ld7e	legw	lfks	lgmp	li5c	ljks	lkww	lm0v	lqn5	lu9b	2mei	8jdw
lav4	lbwz	ld7h	leh1	lfkt	lgmq	li5d	ljkt	lkww	lm1j	lqn6	lu9c	2mha	8ldh
lav5	lbx2	ld7i	leh4	lfkv	lgmr	li5o	ljky	lkwx	lm1k	lqn7	lubi	2mll	8lyz
lavk	lbx6	ld7j	lehb	lflo	lgms	li5u	ljlo	lkwy	lm1x	lqn8	lubb	2mnr	8nse
lavl	lbxd	ld7l	lehi	lfhq	lgmu	li6h	ljli	lkwx	lm44	lqn9	lubb	2mob	8pch
lavm	lbxe	ld7x	lehn	lfis	lgmv	li6s	ljll	lkx0	lm4d	lqna	luch	2ms2	8rat
lavp	lbxi	ld7y	lehw	lfli	lgmw	li6u	ljlu	lkx1	lm4g	lqnb	luco	2msb	8rnt
lavv	lbxn	ld8f	leil	lfli	lgn2	li6v	ljm6	lkxr	lm4i	lqnc	lud7	2mss	8rsa
lavz	lbxr	ld8h	leib	lfly	lgn3	li70	ljmf	lkxw	lm4v	lqnd	ludi	2mst	8ruc
law0	lbxt	ld8i	leic	lfm0	lgn4	li72	ljmg	lkxx	lm55	lqne	luca	2muc	8tli
law3	lbxy	ld8m	leid	lfm1	lgn6	li73	ljmh	lkxy	lm5h	lqnk	luhg	2myo	8tln
law7	lby2	ld8v	leie	lfm2	lgn7	li76	ljmi	lky6	lm5k	lqnm	lugi	2nck	966c
lawc	lby8	ld8z	leig	lfm5	lgnk	li79	ljml	lky7	lm5s	lqnn	luia	2nef	9atc
lawd	lby9	ld9a	leih	lfm7	lgnu	li7b	ljmr	lkyd	lm6i	lqnx	luib	2nmt	9gaa
lawi	lbyf	ld9k	leik	lfm8	lgo4	li7c	ljmt	lkyf	lm6v	lqo3	luic	2nod	9gac
lawp	lbyg	ld9n	leil	lfma	lgoi	li7e	ljmu	lkyi	lm8a	lqo8	luid	2nos	9gaf
lawz	lbyl	ld9u	leim	lfmk	lgoi	li7f	ljn0	lkyo	lm8d	lqoa	luie	2npk	9jdw
laxc	lbyq	ld9w	leiq	lfmo	lgos	li7h	ljn1	lkyp	lm8e	lqob	luif	2nse	9ldb
laxq	lbyr	ldao	leir	lfmp	lgou	li7k	ljn5	lkyr	lm8h	lqof	luig	2nsi	9ldt
laxw	lbys	ldap	leiy	lfm5	lgov	li7m	ljn8	lkys	lm8i	lqog	luih	2one	9nse
laxx	lbyw	ldar	lej1	lfm7	lgoy	li7o	ljna	lkyu	lm9j	lqol	lums	2pad	9pap
lay2	lbz9	ldaw	lej4	lfm9	lgoz	li7q	ljnc	lkza	lm9k	lqom	lunt	2pah	9rat
lay7	lbzl	ldax	lej7	lfne	lgpd	li7r	ljnd	lkzb	lm9m	lqpc	luna	2paw	9rnt
laya	lbzq	lday	lej9	lfng	lgpe	li7s	ljne	lkzc	lm9q	lqpd	luox	2pax	9rsa
layb	lbzs	ldb0	lejc	lfnj	lgph	li7t	ljnk	lkzd	lm9r	lqpe	lup1	2pbh	9rub
layc	lc05	ldb1	lejd	lfnk	lgpj	li7u	ljnr	lkze	lm9s	lqpf	lurn	2pda	s034
layd	lc06	ldb2	lejh	lfnm	lgpm	li83	ljnu	lkzf	lm9t	lqpj	lush	2pbh	s039
laye	lc08	ldbu	lejr	lfno	lgpx	li87	ljnz	lkzi	lm9u	lqpl	lusn	2phk	s041
layf	lc0a	ldbv	lejs	lfnu	lgqv	li8c	ljoa	lkzj	lm9v	lqpj	lute	2phm	s042
layk	lc0b	ldbx	lejt	lfnv	lgr0	li8t	ljoe	lkzn	lm9w	lqpo	luug	2phy	s046
layu	lc0c	ldc2	leju	lfnw	lgra	li8v	ljow	li00	lmd1	lqpq	luxy	2pia	s047
layv	lc0f	ldc3	lejv	lfo0	lgrb	li94	ljpa	li01	lmdl	lqpr	lvac	2pii	s055

APPENDIX E

List of proteins that could not be processed in code Converter.c due to problems in
“chain problem” category.

1AFB	1BMV	1ETB	1HXS	1KMB	1NAL	1R1A	1RUG	2RMU	1VBA	4GPD
1AFD	1BP7	1EV1	1H4K	1KQS	1PD2	1RDI	1RUH	2RR1	1VBB	
1AL0	1BVP	1FFK	1HDN	1KZA	1PIV	1RDJ	1RUI	2RS1	1VBC	
1AL2	1C4U	1FMD	1ILR	1KZB	1PMA	1RDK	1RUJ	2RS3	1VBD	
1AR6	1C4V	1FOD	1IR2	1KZC	1PO1	1RDL	1RVF	2RS5	1VBE	
1AR7	1C4Y	1FPN	1IRU	1KZD	1PO2	1RDM	1RVV	3KMB	1VRH	
1AR8	1C8M	1G0U	1JD2	1KZE	1POV	1RDN	1RYP	3MCG	1YYY	
1AR9	1CD3	1GFF	1JDB	1LMB	1POY	1RDO	1SVA	3PGA	1G29	
1ASJ	1CNT	1GIX	1JEW	1LNR	1PVC	1RHI	2KMB	4KMB	1G65	
1AYM	1COV	1GPH	1JGO	1LOC	1QGC	1RMU	2MCG	4RHV	1G6Q	
1AYN	1D3E	1GT9	1JGP	1M11	1QJU	1RTM	2MEV	8KME	1GAV	
1BBT	1D3I	1GTG	1JGQ	1M1K	1QJX	1RTP	2PLV	1TME	2HWB	
1BCH	1D4M	1GTJ	1JJ2	1M8Q	1QJY	1RUC	2R04	1TMF	2HWC	
1BCJ	1DGI	1GTL	1K2I	1M90	1QQP	1RUD	2R06	1TPB	2HWD	
1BEV	1EAH	1HQZ	1K8A	1MDP	1R08	1RUE	2R07	1TPC	2HWE	
1BHO	1EEK	1HRI	1K9M	1MEC	1R09	1RUF	2RM2	1TPH	2HWF	

APPENDIX F

List of proteins that could not be processed in code Converter.c due to problems in “inconsistency” category.

15C8	1B3T	1C5C	1D9F	1E8U	1F8E	5ER2	1HRU	1J5V	1KEE	1LWU	1QFZ	2KIN	1WAT
16VP	1B41	1C5E	1D9K	1E8V	1F8H	5FX2	1HS7	1J6N	1KEJ	1LXK	1QG3	2KTQ	1WBR
1A02	1B47	1C5P	1DAN	1E8W	1F8T	6EST	1HST	1J75	1KEQ	1LY7	1QGA	2KZM	1WDC
1A07	1B4B	1C5Q	1DAW	1E8X	1F8V	7EST	1HT5	1J7P	1KFA	1LYA	1QGB	2KZZ	1WHS
1A08	1B4R	1C5R	1DAY	1E8Y	1F8X	8EST	1HT8	1J7Z	1KFF	1LYB	1QGE	2LBD	1WHT
1A09	1B4W	1C5S	1DAZ	1E8Z	1F8Y	9EST	1HTJ	1J80	1KFS	1LYV	1QGF	2LDB	1WQ1
1A0B	1B4X	1C5T	1DB0	1E90	1F90	9GAA	1HTL	1J81	1KFU	1LYW	1QGH	2LIG	1WWA
1A0H	1B59	1C5U	1DB1	1E9H	1F97	9GAC	1HTM	1J82	1KFX	1LZ0	1QGK	2MPR	1WWB
1A0J	1B5F	1C5V	1DB7	1E9L	1F9E	9GAF	1HTO	1J83	1KG0	1LZ7	1QGP	2MSS	1WWC
1A0N	1B5S	1C6V	1DB8	1E9R	1F9F	1GC1	1HTQ	1J84	1KGD	1LZI	1QGR	2MST	1WWW
1A0O	1B5T	1C6X	1DB9	1E9S	1F9G	1GC3	1HTV	1J8A	1KGE	1LZJ	1QGZ	2MTA	1WYK
1A0R	1B66	1C6Y	1DBC	1EAA	1F9J	1GC4	1HTZ	1J8F	1KGF	1M0D	1QH0	2NCD	1X11
1A0S	1B67	1C6Z	1DBH	1EAB	1F9M	1GCC	1HU3	1J8H	1KGG	1M10	1QHH	2NEF	1XAN
1A0T	1B6B	1C70	1DBQ	1EAC	1F9T	1GCI	1HU8	1J8Q	1KGY	1M14	1QHI	2NGR	1XBD
1A12	1B6C	1C75	1DCI	1EAD	1F9U	1GCJ	1HUL	1J8T	1KH5	1M17	1QHK	2NLL	1XBR
1A14	1B6E	1C7E	1DCN	1EAE	1F9V	1GCK	1HUP	1J8U	1KHC	1M1U	1QHV	2NMB	1XDT
1A1A	1B6J	1C7F	1DCQ	1EAF	1F9W	1GCP	1HUQ	1J8W	1KHH	1M20	1QIB	2NMT	1XKA
1A1B	1B6K	1C7S	1DCZ	1EAI	1F9X	1GCQ	1HUW	1J8X	1KHU	1M40	1QIR	2NOD	1XKB
1A1C	1B6L	1C7T	1DD1	1EAJ	1FAD	1GCT	1HUY	1J92	1KHX	1M52	1QIS	2NOS	1XPA
1A1E	1B6M	1C7U	1DD2	1EAK	1FAO	1GD2	1HV2	1J94	1KJ0	1M54	1QIT	2NOT	1XPB
1A1F	1B6N	1C7V	1DD7	1EAN	1FAP	1GD9	1HV5	1J97	1KI1	1M56	1QIU	2NSE	1XTC
1A1G	1B6O	1C7W	1DD9	1EAO	1FAQ	1GDE	1HV6	1J9E	1KI2	1M57	1QIX	2NSI	1XUF
1A1H	1B6P	1C82	1DDE	1EAQ	1FAR	1GDF	1HV7	1J9Y	1KI3	1M62	1QJC	2PAH	1XUG
1A1I	1B6Z	1C8B	1DDF	1EAW	1FAV	1GDN	1HVU	1J9Z	1KI4	1M6I	1QJV	2PAW	1XUH
1A1J	1B72	1C8E	1DDG	1EAX	1FAX	1GDQ	1HVV	1JA0	1KJ6	1M6V	1QJW	2PAX	1XUI
1A1K	1B76	1C8F	1DDI	1EAY	1FB0	1GDU	1HVV	1JA1	1KI7	1M7B	1QK0	2PCD	1XUJ
1A1L	1B79	1C8G	1DDJ	1EAZ	1FB1	1GEN	1HW8	1JA3	1KI8	1M7K	1QK2	2PCP	1XUK
1A1R	1B7E	1C8O	1DDM	1EB2	1FB2	1GFN	1HW9	1JA9	1KID	1M7S	1QKL	2PHK	1XWL
1A1U	1B7F	1C8R	1DDX	1EBA	1FB3	1GFW	1HWI	1JAD	1KIG	1M7X	1QKM	2PJR	1XXA
1A1V	1B7V	1C8S	1DDZ	1EBD	1FB6	1GG4	1HWJ	1JAG	1KJ1	1M8D	1QKN	2PKA	1XXB
1A22	1B86	1C8T	1DE4	1EBF	1FB8	1GG6	1HWK	1JAM	1KIK	1M8E	1QKR	2PLC	1XXC
1A25	1B88	1C8X	1DE7	1EBK	1FBL	1GGD	1HWL	1JAN	1KIL	1M8H	1QKT	2PNA	1XYA
1A26	1B89	1C8Y	1DE8	1EBM	1FBM	1GGM	1HWT	1JAO	1KIM	1M8I	1QKU	2PNB	1XYB
1A28	1B8C	1C8Z	1DE9	1EBP	1FBQ	1GGO	1HX1	1JAP	1KIT	1M8W	1QL0	2POO	1XYC
1A2F	1B8H	1C90	1DEE	1EBU	1FBS	1GGQ	1HXI	1JAQ	1KJ2	1M8X	1QL7	2PRG	1XYF
1A2G	1B8I	1C91	1DEG	1EBV	1FBU	1GGR	1HXN	1JB1	1KJW	1M8Y	1QL8	2PTC	1XYL
1A2V	1B8Y	1C92	1DEV	1EBW	1FBV	1GH2	1HXR	1JB6	1KJY	1M8Z	1QL9	2PTK	1XYM
1A31	1B91	1C93	1DEW	1EBY	1FBY	1GH6	1HY5	1JBK	1KKE	1M9J	1QLC	2PTN	1XYZ
1A35	1B92	1C9B	1DF1	1EBZ	1FBZ	1GHI	1HY7	1JBP	1KKL	1M9K	1QLD	2QWA	1YAT
1A36	1B9B	1C9C	1DF8	1EC0	1FC1	1GHI	1HY9	1JBU	1KKM	1M9M	1QLE	2QWB	1YCP

1A3H	1B9D	1C9E	1DF9	1EC1	1FC2	1GHP	1HYH	1JC2	1KKQ	1M9Q	1QLG	2QWC	1YCQ
1A3K	1B9F	1C9J	1DFK	1EC2	1FC3	1GHU	1HYL	1JCK	1KL3	1M9R	1QLI	2QWD	1YCR
1A3L	1B9K	1C9M	1DFL	1EC3	1FC4	1GHZ	1HYM	1JD4	1KL4	1M9T	1QLX	2QWE	1YCS
1A3Q	1B9S	1C9N	1DFP	1ECS	1FC6	1GI0	1HYO	1JD5	1KL5	1M9U	1QLZ	2QWF	1YFO
1A3R	1B9T	1C9P	1DFQ	1ED4	1FC7	1GI1	1HYR	1JD6	1KLG	1M9Z	1QM0	2QWG	1YFP
1A41	1B9V	1C9T	1DG4	1ED5	1FC9	1GI2	1HYV	1JDE	1KLI	1MAP	1QM1	2QWH	1YGH
1A43	1B9W	1C9Y	1DG6	1ED6	1FCC	1GI3	1HYZ	1JDH	1KLO	1MAQ	1QM2	2QWI	1YGS
1A49	1B9X	1CA0	1DGB	1ED7	1FCF	1GI4	1HZD	1JDX	1KLQ	1MAX	1QM3	2QWJ	1YOO
1A4E	1B9Y	1CA4	1DGC	1ED8	1FCH	1GI5	1HZG	1JEA	1KLT	1MAY	1QMB	2QWK	1YPA
1A4G	1BAJ	1CA9	1DGF	1ED9	1FCP	1GI6	1HZH	1JEK	1KLU	1MBE	1QMC	2R1R	1YPB
1A4L	1BAK	1CAG	1DGG	1EDM	1FCX	1GI6	1HZO	1JEN	1KMC	1MBF	1QME	2RAM	1YPC
1A4M	1BBP	1CAQ	1DGH	1EDO	1FCY	1GIH	1HZW	1JFI	1KMD	1MBG	1QMF	2RAP	1YPT
1A4Q	1BBY	1CB2	1DGP	1EDQ	1FCZ	1GJI	1HZY	1JFQ	1KMR	1MBH	1QMG	2RLN	1YRN
1A4R	1BC5	1CB8	1DGQ	1EE4	1FDP	1GJO	1I00	1JFZ	1KN3	1MBJ	1QMJ	2RMC	1YSA
1A52	1BC9	1CBU	1DGR	1EE5	1FDQ	1GJY	1I07	1JGG	1KN9	1MBK	1QMO	2SAK	1YTB
1A5U	1BCC	1CBW	1DGW	1EED	1FE8	1GK0	1I0B	1JGM	1KNA	1MBL	1QMV	2SEC	1YTF
1A5V	1BCM	1CC3	1DH3	1EEQ	1FEJ	1GK1	1I0C	1JGS	1KNB	1MC2	1QMY	2SEM	1YTN
1A5W	1BCO	1CC5	1DHS	1EEU	1FEV	1GK4	1I0D	1JH1	1KNE	1MCT	1QMZ	2SGA	1YTS
1A5X	1BCR	1CCE	1DII	1EF0	1FEZ	1GK6	1I19	1JHC	1KNG	1MDA	1QNG	2SGP	1YTT
1A5Z	1BCS	1CCF	1DI2	1EF1	1FF0	1GK7	1I1A	1JHE	1KNO	1MDY	1QNJ	2SIV	1YTW
1A64	1BCT	1CCG	1DIC	1EF5	1FF1	1GKC	1I1C	1JHG	1KO6	1MFA	1QNU	2SLI	1YUI
1A67	1BD8	1CCK	1DII	1EFD	1FF7	1GKD	1I1D	1JHN	1KOA	1MFB	1QNX	2SNV	1YUJ
1A68	1BDE	1CCL	1DIQ	1EFN	1FFF	1GKE	1I1K	1JHZ	1KOE	1MFC	1QNZ	2SNW	1YVE
1A6A	1BDG	1CDW	1DIW	1EFU	1FFG	1GKG	1I1L	1JIS	1KOG	1MFD	1QO3	2SRC	1ZAG
1A6P	1BDJ	1CE5	1DIY	1EFV	1FFI	1GKI	1I1M	1JI6	1KOH	1MFE	1QO7	2SRT	1ZBD
1A6R	1BDM	1CEM	1DJI	1EFY	1FFM	1GKK	1I27	1JI7	1KOO	1MFO	1QOH	2SSP	1ZDC
1A6U	1BDO	1CEQ	1DJ5	1EG3	1FFR	1GKL	1I2M	1JIA	1KOY	1MFQ	1QOJ	2STA	1ZDD
1A6V	1BE9	1CET	1DJA	1EG4	1FFS	1GKN	1I2S	1JIE	1KP5	1MG6	1QOL	2STB	1ZFD
1A6W	1BEC	1CF4	1DJB	1EG9	1FFW	1GL2	1I2T	1JIF	1KPM	1MH3	1QOM	2STT	1ZFF
1A6X	1BEJ	1CF7	1DJC	1EGG	1FG6	1GL4	1I2W	1JIR	1KPS	1MH4	1QOY	2STW	1ZME
1A6Y	1BEK	1CFB	1DJG	1EGI	1FG8	1GL5	1I31	1JJ4	1KQ5	1MHL	1QP9	2TGA	1ZNB
1A79	1BEM	1CG2	1DJH	1EGJ	1FGC	1GL6	1I37	1JJ6	1KQL	1MIU	1QPC	2TGD	1ZON
1A7G	1BEP	1CGD	1DJI	1EGP	1FGG	1GL7	1I38	1JJ7	1KQR	1MJ3	1QPD	2TGP	1ZOO
1A7H	1BEQ	1CGE	1DJL	1EGU	1FGI	1GL8	1I3G	1JJ8	1KRI	1MKC	1QPE	2TGT	1ZOP
1A7N	1BES	1CGF	1DJO	1EH5	1FGK	1GLU	1I3J	1JJ9	1KRL	1MKP	1QPF	2TIO	1ZQU
1A7O	1BET	1CGH	1DJP	1EHB	1FGU	1GLV	1I3O	1JJB	1KRN	1ML1	1QPJ	3KAR	1ZQV
1A7P	1BF5	1CGL	1DJS	1EHI	1FGX	1GM1	1I42	1JJF	1KRP	1MMB	1QPL	3KIN	1ZQW
1A7Q	1BF9	1CHD	1DJW	1EHK	1FGY	1GML	1I44	1JJH	1KRS	1MML	1QPM	3KTQ	1ZQX
1A7R	1BFE	1CI6	1DJX	1EHL	1FGZ	1GMN	1I4E	1JJO	1KRT	1MMO	1QPN	3LBD	1ZQY
1A81	1BFF	1CIA	1DJY	1EHN	1FH5	1GMO	1I4M	1JK1	1KSP	1MMP	1QPO	3LCK	1ZQZ
1A85	1BFS	1CII	1DJZ	1EI1	1FH6	1GNI	1I4N	1JK2	1KSW	1MMQ	1QPQ	3MDD	1ZZZ
1A86	1BFT	1CIQ	1DK1	1EI7	1FHQ	1GNF	1I4O	1JK3	1KSX	1MMR	1QPS	3MDE	2TLD
1A87	1BGW	1CIR	1DK4	1EI9	1FHR	1GNS	1I4S	1JGK	1KSY	1MNC	1QQ1	3MEF	2TOH
1A8O	1BH2	1CIS	1DK5	1EIA	1FHW	1GNV	1I4V	1JKM	1KT2	1MNM	1QQ8	3MRA	2TPI
1A8Z	1BH7	1CIT	1DK7	1EIB	1FHX	1GNY	1I51	1JKO	1KTP	1MOF	1QQF	3MUC	2TPS
1A90	1BH8	1CIY	1DK8	1ELZ	1FI0	1GO4	1I58	1JKP	1KTQ	1MOQ	1QQH	3NOD	2TRC
1A92	1BH9	1CIZ	1DKD	1EJO	1FI1	1GP0	1I59	1JKQ	1KTR	1MOR	1QQM	3NOS	2TRM
1A99	1BHD	1CJO	1DKF	1EJ1	1FI5	1GP3	1I5A	1JKR	1KTS	1MOS	1QQN	3NSE	2TSY
1A9N	1BHF	1CJ1	1DKW	1EJ4	1FI8	1GP9	1I5B	1JL4	1KTT	1MP6	1QQO	3PAH	2ULL

1A9X	1BHG	1CJK	1DKX	1EJ7	1FIB	1GPC	1ISC	1JLI	1KTZ	1MRC	1QQR	3PAX	2UP1
1AA0	1BHH	1CJQ	1DKY	1EJ8	1FIC	1GPD	1ISD	1JLM	1KU1	1MRD	1QQU	3PCA	2USN
1AA3	1BHL	1CJR	1DKZ	1EJ9	1FID	1GPL	1ISH	1JLN	1KU2	1MRE	1QQV	3PCB	2VGC
1AA7	1BHT	1CJT	1DL2	1EJA	1FIF	1GPP	1ISK	1JLT	1KU3	1MRF	1QR3	3PCC	2VIS
1AAM	1BHX	1CJU	1DL3	1EJH	1FIH	1GPZ	1ISQ	1JLU	1KU7	1MSE	1QRB	3PCD	2VPF
1AAT	1BHY	1CJV	1DLH	1EJL	1FIN	1GQ2	1ISU	1JM1	1KUL	1MSF	1QRC	3PCE	2VTK
1AAW	1BI4	1CJW	1DLK	1EJM	1FIO	1GQ4	1I69	1JM4	1KUM	1MSK	1QRH	3PCF	2XBD
1AAZ	1BII	1CK3	1DLL	1EJO	1FIQ	1GQF	1I6A	1JMC	1KUZ	1MSL	1QRI	3PCG	2ZNB
1AB4	1BIO	1CK4	1DLU	1EJY	1FJ1	1GQP	1I6J	1JMM	1KV7	1MTN	1QRJ	3PCH	3TPI
1AB5	1BIS	1CK7	1DLV	1EK3	1FJ2	1GQR	1I6V	1JMS	1KVD	1MTS	1QRQ	3PCI	3VGC
1AB6	1BIU	1CKA	1DM3	1EKB	1FJD	1GQS	1I6Z	1JMT	1KVE	1MTU	1QRZ	3PCJ	3VTK
1AB8	1BIW	1CKB	1DM6	1EKG	1FJS	1GR3	1I72	1JMU	1KVL	1MTV	1QS2	3PCK	3ZNB
1AB9	1BIX	1CKR	1DM7	1EKJ	1FKA	1GRN	1I73	1JMW	1KVM	1MTW	1QS4	3PCL	4TPI
1ABO	1BIZ	1CKT	1DM8	1EKM	1FKC	1GTO	1I76	1JN3	1KW2	1MTY	1QSC	3PCM	4TSV
1ABQ	1BJI	1CL3	1DMA	1ELA	1FKM	1GTQ	1I79	1JN5	1KWA	1MWE	1QSJ	3PCN	4VGC
1AC0	1BJ4	1CL5	1DMG	1ELB	1FKW	1GU6	1I7B	1JN8	1KWF	1MWP	1QSL	3PHM	4ZNB
1AC6	1BJ9	1CL7	1DMI	1ELC	1FKX	1GUA	1I7C	1JNA	1KWS	1MYP	1QSM	3PRG	5TSW
1ACZ	1BJI	1CLA	1DMJ	1ELD	1FL0	1GUX	1I7E	1JNC	1KWT	1NAP	1QSN	3PTB	1G16
1AD5	1BJJ	1CLD	1DMK	1ELE	1FL2	1GV4	1I7G	1JNL	1KWU	1NCA	1QSO	3PTN	1G17
1AD6	1BJK	1CLV	1DML	1ELF	1FL3	1GV8	1I7I	1JNN	1KWV	1NCB	1QSR	3R1R	1G1E
1ADQ	1BJT	1CLW	1DMT	1ELG	1FL5	1GV9	1I7M	1JNP	1KWW	1NCC	1QSS	3RAB	1G1I
1ADT	1BJU	1CLY	1DMW	1ELR	1FL6	1GVC	1I7P	1JNU	1KWX	1NCD	1QST	3RHN	1G1J
1ADU	1BJV	1CLZ	1DMX	1ELT	1FLE	1GVK	1I7Q	1JNV	1KWY	1NCS	1QSV	3RP2	1G1U
1ADV	1BJX	1CM0	1DMY	1ELV	1FLI	1GVL	1I7W	1JNX	1KWZ	1NCT	1QSY	3RPB	1G24
1AE7	1BK2	1CM1	1DMZ	1EM2	1FLK	1GVM	1I7X	1JNZ	1KX0	1NCU	1QSZ	3SC2	1G28
1AE9	1BK5	1CM4	1DN2	1EMH	1FLL	1GVY	1I7Y	1JOB	1KX1	1NDD	1QT1	3SEM	1G2C
1AER	1BK6	1CMF	1DNL	1EMJ	1FLT	1GVZ	1I7Z	1JOC	1KXA	1NDO	1QTM	3SGA	1G2D
1AF7	1BK9	1CMG	1DNU	1EMN	1FM2	1GW1	1I83	1JOD	1KXB	1NES	1QTN	3SGB	1G2E
1AF9	1BKB	1CMI	1DNW	1EMO	1FM6	1GWA	1I8I	1JON	1KXC	1NFI	1QTP	3SLI	1G2F
1AFO	1BKC	1CMO	1DO0	1EMR	1FM9	1GWQ	1I8K	1JOY	1KXD	1NFD	1QTS	3STD	1G2I
1AFQ	1BKD	1CMX	1DO2	1EMU	1FMG	1GWR	1I8Q	1JPA	1KXE	1NFI	1QTY	3SXL	1G2L
1AFR	1BKE	1CMZ	1DO5	1ENH	1FMI	1GWX	1I9F	1JPL	1KXG	1NFK	1QU0	3TAT	1G2M
1AFW	1BKH	1CN3	1DOQ	1ENT	1FMK	1GWZ	1I9H	1JPP	1KXM	1NFT	1QU1	3TGI	1G2N
1AG2	1BKN	1CO1	1DOV	1ENV	1FN6	1GXC	1I9R	1JPW	1KXN	1NGR	1QU5	3TGJ	1G33
1AGW	1BL3	1COA	1DOW	1EO2	1FN7	1GXJ	1I9Y	1JQ7	1KXP	1NK1	1QU7	3TGK	1G34
1AGY	1BL8	1COO	1DP6	1EO9	1FN8	1GXK	1I9Z	1JQH	1KXR	1NK2	1QUB	4KTQ	1G36
1AHE	1BLC	1COZ	1DP8	1EOA	1FNA	1GXL	1IA9	1JQI	1KY6	1NK3	1QUQ	4LBD	1G38
1AHF	1BLH	1CPT	1DP9	1EOB	1FNF	1GXP	1IAH	1JQ7	1KY7	1NMA	1QUS	4LVE	1G3B
1AHG	1BLN	1CQ4	1DPB	1EOC	1FNI	1GXQ	1IAJ	1JRR	1KYB	1NMC	1QUT	4NOS	1G3C
1AHS	1BLP	1CQ6	1DPC	1EOD	1FNS	1GXR	1IAK	1JRS	1KYD	1NMT	1QVC	4NSE	1G3D
1AHX	1BM2	1CQ7	1DPD	1EOE	1FNU	1GY3	1IAL	1JRT	1KYF	1NNA	1R1R	4PAH	1G3E
1AHY	1BM6	1CQ8	1DPM	1EOF	1FNV	1GYD	1IAP	1JS2	1KYN	1NNB	1RAM	4PAX	1G3F
1AII	1BM8	1CQF	1DPO	1EOQ	1FNW	1GYE	1IAS	1JS8	1KYO	1NNC	1RBC	4PRG	1G3J
1AIA	1BMA	1CQK	1DPU	1EOV	1FO0	1GYF	1IAU	1JSC	1KYU	1NNT	1RBD	4R1R	1G49
1AIB	1BMB	1CQP	1DQ8	1EP9	1FO1	1GYH	1IAV	1JSP	1KZ7	1NOB	1RBE	4RHN	1G4D
1AIC	1BMC	1CQR	1DQ9	1EPL	1FO2	1GYU	1IAX	1JST	1KZG	1NOC	1RBF	4SGA	1G4E
1AIE	1BMD	1CQT	1DQA	1EPM	1FO3	1GYV	1IAY	1JSU	1KZN	1NOD	1RBG	4SKN	1G4K
1AIH	1BML	1CQW	1DQJ	1EPN	1FO7	1GYW	1IB0	1JT2	1KZS	1NOM	1RBH	4SLI	1G4P
1AIK	1BMO	1CQY	1DQM	1EPO	1FO8	1GYZ	1IB2	1JTD	1KZT	1NOS	1RBI	4STD	1G4S

1AIN	1BMQ	1CR0	1DQQ	1EPP	1FO9	1GZA	1IBC	1JTG	1KZV	1NOV	1RBL	4THI	1G4T
1AIS	1BMT	1CR1	1DQV	1EPQ	1FOA	1GZE	1IBG	1JTH	1KZY	1NPM	1RCS	5KTQ	1G4U
1AJ5	1BN6	1CR2	1DQZ	1EPR	1FOE	1GZF	1IBI	1JTR	1KZZ	1NRE	1RCY	5NSE	1G4V
1AJ8	1BN7	1CR4	1DRG	1EPT	1FOF	1GZH	1IBT	1JTZ	1L0A	1NS3	1RDH	5PAH	1G4W
1AJW	1BNK	1CR5	1DRM	1EQ3	1FOI	1GZJ	1IBU	1JUN	1L0B	1NSB	1RFA	5PTP	1G4X
1AJY	1BOE	1CRX	1DRQ	1EQ6	1FOJ	1GZS	1IBV	1JUQ	1L0D	1NSC	1RFN	5R1R	1G4Y
1AK4	1BOO	1CRZ	1DS2	1EQ9	1FOL	1GZX	1IBW	1JV5	1L0E	1NSD	1RFS	5RHN	1G50
1AKA	1BOS	1CS0	1DS3	1EQF	1FOO	1H0P	1ICE	1JVA	1L0F	1NSE	1RGS	5SGA	1G55
1AKB	1BP3	1CS3	1DS4	1EQN	1FOP	1H1B	1ICF	1JVJ	1L0G	1NSF	1RHO	5STD	1G56
1AKC	1BPL	1CS4	1DS5	1ER8	1FOT	1H1P	1ICH	1JVW	1L0M	1NSG	1RIE	6NSE	1G5H
1AKH	1BPR	1CS6	1DS6	1ERE	1FP5	1H1Q	1ICJ	1JVZ	1L0X	1NSI	1RL2	6PAH	1G5I
1AKS	1BQA	1CSB	1DSE	1ERJ	1FPR	1H1R	1ID0	1JW0	1L0Y	1NSN	1RLD	6R1R	1G5J
1AKZ	1BQD	1CSO	1DSG	1ERM	1FPS	1H1S	1ID1	1JWE	1L1D	1NST	1RLF	6RHN	1G5M
1ALI	1BQE	1CT0	1DSJ	1ERN	1FPT	1H2R	1ID4	1JWJ	1L1O	1NTC	1RLW	6RLX	1G5Q
1ALL	1BQO	1CT2	1DSK	1ERO	1FPU	1H4L	1IDO	1JWK	1L2E	1NTP	1RNE	6RXN	1G5V
1ALQ	1BQP	1CT4	1DSL	1ERQ	1FQ3	1H4U	1IDY	1JWO	1L2I	1NUB	1RP1	6STD	1G5Y
1ALV	1BQQ	1CTF	1DSO	1ERR	1FQ9	1H4W	1IDZ	1JWP	1L2J	1NUK	1RPL	7NN9	1G5Z
1ALW	1BQY	1CTN	1DSP	1ES1	1FQG	1H59	1IE8	1JWV	1L2P	1NYF	1RPM	7NSE	1G64
1ALY	1BR0	1CU1	1DST	1ES6	1FQI	1H5P	1IE9	1JWY	1L2S	1NYG	1RRB	7PCY	1G67
1AM4	1BRA	1CUE	1DSU	1ES7	1FQJ	1H5X	1IEC	1JWZ	1L2W	1O0V	1RRP	7R1R	1G68
1AM9	1BRB	1CUL	1DSV	1ESA	1FQK	1H67	1IED	1JX2	1L3A	1OAK	1RSC	7STD	1G69
1AMA	1BRC	1CUN	1DSX	1ESB	1FQV	1H6E	1IEF	1JX6	1L3E	1OCT	1RST	8NSE	1G6A
1AMH	1BRU	1CUS	1DSY	1ESJ	1FR0	1H6F	1IEG	1JX7	1L4A	1ODD	1RSU	8OHM	1G6C
1AMO	1BRV	1CVD	1DSZ	1ESQ	1FR8	1H6G	1IEL	1JXE	1L4W	1OEF	1RSY	9NSE	1G6G
1AMQ	1BS4	1CVH	1DT4	1EST	1FS0	1H6K	1IEM	1JXM	1L5E	1OEG	1RTG	9PAI	1G6N
1AMR	1BS5	1CVM	1DT6	1ESU	1FS1	1H6L	1IEP	1JXO	1L6J	1OIS	1RXR	1TKI	1G6R
1AMS	1BS6	1CVS	1DT7	1ESZ	1FS2	1H6O	1IF0	1JXQ	1L6X	1OJT	1SAE	1TLD	1G6V
1AMX	1BS7	1CVU	1DTD	1ET0	1FS7	1H6P	1IFG	1JY1	1L7C	1OME	1SAF	1TLF	1G6W
1AN1	1BS8	1CVW	1DTO	1ETC	1FS8	1H6T	1IFR	1JY2	1L7F	1OPC	1SAG	1TMQ	1G6Y
1AN2	1BSG	1CWR	1DTS	1ETD	1FSG	1H6U	1IFY	1JY3	1L7G	1OPM	1SAH	1TMT	1G73
1AN4	1BSW	1CWS	1DU0	1ETF	1FSO	1H6W	1IG7	1JYM	1L7H	1OSP	1SAI	1TMU	1G7C
1AN7	1BSX	1CWT	1DU3	1ETG	1FST	1H73	1IGN	1JYO	1L7M	1OTH	1SAJ	1TN3	1G7D
1ANB	1BSZ	1CWU	1DUB	1EUB	1FSU	1H74	1IGQ	1JYQ	1L7N	1OVA	1SAK	1TND	1G7E
1ANC	1BT5	1CWV	1DUI	1EUF	1FSW	1H7Z	1IGT	1JYR	1L7O	1OVB	1SAL	1TNV	1G7W
1AND	1BTG	1CX0	1DUJ	1EUR	1FSX	1H80	1IGU	1JYU	1L7P	1OXO	1SBB	1TOH	1G7X
1ANE	1BTL	1CX4	1DUL	1EUS	1FSY	1H85	1IGY	1JZI	1L7W	1OXP	1SBW	1TON	1G82
1ANV	1BTP	1CX8	1DUS	1EUT	1FT0	1H88	1IH0	1JZJ	1L8B	1P32	1SCJ	1TPA	1G83
1AO5	1BTU	1CXP	1DV8	1EUU	1FT2	1H89	1IHJ	1JZR	1L8C	1PAA	1SEB	1TPL	1G88
1AOA	1BTW	1CXQ	1DVA	1EUV	1FT3	1H8A	1IHK	1K04	1L8D	1PAH	1SEK	1TPO	1G80
1AOD	1BTX	1CXU	1DVC	1EV0	1FT8	1H8F	1IHN	1K05	1L8P	1PAU	1SEM	1TPP	1G8Q
1AOI	1BTY	1CXV	1DVD	1EV2	1FTS	1H8G	1IHP	1K0A	1L8Q	1PAX	1SER	1TPS	1G93
1AOK	1BTZ	1CXX	1DVI	1EV4	1FU0	1H8X	1IHR	1K0B	1L8R	1PBK	1SFC	1TRM	1G9I
1AOL	1BU1	1CXZ	1DVJ	1EVJ	1FU1	1H8Y	1IHS	1K0C	1LA1	1PBV	1SFI	1TRN	1G9M
1AOP	1BU2	1CY9	1DVK	1EVK	1FUJ	1H8Z	1IHV	1K0D	1LAR	1PBW	1SGC	1TRP	1G9N
1AOW	1BUD	1CYN	1DVO	1EVL	1FUK	1H99	1IHW	1K0F	1LAT	1PCF	1SGF	1TRQ	1G9S
1AOX	1BUE	1CYQ	1DW6	1EVT	1FV0	1H9D	1I14	1K0Y	1LB1	1PCS	1SGP	1TRY	1G9T
1API	1BUI	1CYW	1DWA	1EW0	1FV2	1H9L	1I18	1K1A	1LB2	1PDR	1SGQ	1TS2	1G9V
1APO	1BUL	1CYX	1DWF	1EW3	1FV3	1H9R	1I1C	1K1B	1LB4	1PEH	1SGR	1TS3	1G9Y
1APQ	1BUO	1CYY	1DWG	1EWK	1FVG	1H9S	1I1D	1K1G	1LB5	1PEI	1SGT	1TS4	1G9Z

1APY	1BUQ	1CZ0	1DWH	1EWN	1FVO	1H9U	1IIE	1K1I	1LB6	1PES	1SHC	1TS5	1GA9
1APZ	1BUU	1CZ7	1DWI	1EWR	1FVR	1HA0	1IIG	1K1J	1LBD	1PET	1SHV	1TSP	1GAG
1AQ7	1BUV	1CZ8	1DWJ	1EWT	1FVU	1HA4	1IIH	1K1L	1LBL	1PEX	1SIG	1TSR	1GAM
1AQC	1BV4	1CZ9	1DWV	1EWV	1FVV	1HA5	1IIL	1K1M	1LBV	1PFX	1SKN	1TTI	1GAO
1AQG	1BVN	1CZB	1DWW	1EWZ	1FW4	1HA7	1IIS	1K1N	1LBW	1PHM	1SLI	1TTJ	1GBA
1ARD	1BVO	1CZC	1DWX	1EX4	1FWK	1HAA	1IIX	1K1O	1LBX	1PHN	1SLL	1TUB	1GBB
1ARE	1BVW	1CZE	1DWY	1EX6	1FWL	1HAJ	1IJB	1K1P	1LBY	1PIO	1SLN	1TVX	1GBC
1ARF	1BVY	1CZG	1DWZ	1EXA	1FWX	1HAX	1IJB	1K1T	1LBZ	1PJP	1SLU	1TVX	1GBD
1ARG	1BW8	1CZI	1DX0	1EXB	1FX2	1HAY	1IJE	1K1U	1LCJ	1PMS	1SLV	1TX4	1GBE
1ARH	1BW9	1CZW	1DX1	1EXQ	1FX4	1HAZ	1IJF	1K1V	1LCK	1POI	1SLW	1TYN	1GBF
1ARI	1BWM	1CZY	1DX5	1EXT	1FXK	1HB0	1IJH	1K20	1LCU	1PON	1SLX	1TYU	1GBH
1ARS	1BWV	1CZZ	1DXZ	1EXW	1FXL	1HBK	1IJJ	1K21	1LD7	1POO	1SMF	1TYV	1GBI
1ART	1BXD	1D00	1DY0	1EXX	1FXY	1HBM	1IJK	1K22	1LD8	1POT	1SPA	1TYW	1GBJ
1AS4	1BXG	1D06	1DY1	1EXY	1FY4	1HBN	1IJK	1K25	1LDB	1POU	1SPB	1TYX	1GBK
1ASA	1BXX	1D0A	1DY2	1EXZ	1FY5	1HBO	1IJT	1K2B	1LDD	1POW	1SPP	1TZE	1GBL
1ASB	1BXL	1D0B	1DY6	1EY3	1FY7	1HBU	1IJW	1K2C	1LDG	1POX	1SRA	1UBD	1GBM
1ASC	1BXU	1D0C	1DY8	1EYH	1FY8	1HBW	1IK7	1K2F	1LDJ	1PPF	1SRS	1UCB	1GBN
1ASD	1BXV	1D0E	1DY9	1EYW	1FY9	1HBX	1IK9	1K2M	1LDK	1PPG	1SRV	1UCY	1GBT
1ASE	1BXX	1D0G	1DYK	1EZ2	1FYA	1HC8	1IKN	1K2N	1LDT	1PRG	1SSP	1UEA	1K7A
1ASF	1BY4	1D0H	1DYL	1EZ3	1FYF	1HCF	1IKZ	1K2P	1LE2	1PRH	1SSS	1UGH	1K7B
1ASG	1BY6	1D0J	1DYP	1EZ4	1FYH	1HCG	1ILG	1K2V	1LE3	1PSJ	1STR	1UIO	1K7L
1ASL	1BY7	1D0K	1DZ1	1EZ5	1FYK	1HCI	1ILH	1K3A	1LE4	1PSZ	1STS	1UIP	1K7S
1ASM	1BY9	1D0L	1DZA	1EZL	1FYL	1HCS	1ILK	1K3B	1LE8	1PTA	1SUA	1UMS	1K7Y
1ASN	1BYG	1D0M	1DZB	1EZS	1FYM	1HCT	1ILY	1K3G	1LF8	1PTQ	1SUC	1UMT	2GCC
1ASU	1BYK	1D0N	1DZC	1EZU	1FYN	1HCX	1IM3	1K3H	1LFA	1PTR	1SUE	1UMU	2GCT
1ASV	1BYM	1D0O	1DZD	1EZV	1FYR	1HD0	1IMH	1K3I	1LFD	1PUE	1SVF	1UPF	2GEP
1ASW	1BYN	1D0W	1DZI	1EZX	1FYV	1HD1	1IMO	1K3J	1LG1	1PVD	1SVN	1UPU	2GLI
1ASY	1BYW	1D1E	1DZO	1F00	1FYW	1HE1	1IN1	1K3N	1LG2	1PVU	1SVP	1URT	2GPR
1ASZ	1BYX	1D1F	1E08	1F02	1FYX	1HE7	1INF	1K3Q	1LG7	1PYA	1SVQ	1USN	2GRT
1AT0	1BYZ	1D1I	1E0A	1F06	1FZA	1HE8	1ING	1K40	1LGB	1PYC	1SVR	1VCB	2GWX
1ATI	1BZ4	1D1K	1E0B	1F08	1FZB	1HE9	1INH	1K4E	1LGP	1PYT	1SVY	1VCP	2HIP
1ATT	1BZA	1D1N	1E0J	1F0C	1FZC	1HEH	1IO1	1K4F	1LGQ	1QA0	1SW6	1VCQ	2HAP
1ATU	1BZQ	1D1U	1E0K	1F0I	1FZD	1HEI	1IO4	1K4Q	1LHY	1QA1	1SWA	1VDF	2HCK
1ATZ	1BZS	1D1V	1E0N	1F0J	1FZE	1HEJ	1IO7	1K4U	1LI0	1QA2	1SWB	1VFX	2HGF
1AU7	1BZX	1D1W	1E0O	1F0T	1FZF	1HES	1IO8	1K4V	1LI9	1QA3	1SWC	1VGC	2HLC
1AU8	1C04	1D1X	1E0R	1F0U	1FZG	1HEZ	1IO9	1K4W	1LIA	1QA6	1SWD	1VHH	2HLP
1AUE	1C0E	1D1Y	1E12	1F0V	1FZR	1HFC	1IOD	1K4Y	1LIH	1QA9	1SWE	1VHI	2HNT
1AUG	1C0F	1D2E	1E17	1F11	1FZY	1HFD	1IP3	1K4Z	1LIU	1QAB	1SWF	1VIE	2HPP
1AUJ	1C0M	1D2J	1E1G	1F1C	1FZZ	1HFS	1IP7	1K54	1LIW	1QAC	1SWG	1VIF	2HPQ
1AUM	1C12	1D2K	1E1J	1F1J	1G01	1HG4	1IQ1	1K55	1LIX	1QAF	1SWH	1VIG	2HRP
1AUQ	1C16	1D2N	1E1P	1F1M	1G05	1HG8	1IQ5	1K56	1LIY	1QAG	1SWJ	1VIH	2HTS
1AUT	1C17	1D2O	1E1S	1F1S	1G0C	1HGT	1IQ7	1K57	1LIJ	1QAI	1SWK	1VII	2IFE
1AUV	1C1A	1D2P	1E1U	1F1W	1G0H	1HGW	1IQA	1K5M	1LIJ	1QAJ	1SWL	1VIN	2IRF
1AUX	1C1E	1D2Q	1E1W	1F20	1G0I	1HGY	1IQD	1K5W	1LKK	1QAK	1SWN	1VIP	2ISD
1AV1	1C1G	1D2S	1E25	1F2F	2CUA	1HH4	1IQT	1K60	1LKL	1QAL	1SWO	1VIT	2ITG
1AV6	1C1J	1D2V	1E2H	1F2I	2CUT	1HHN	1IR3	1K74	1LKV	1QAS	1SWP	1VIW	2IZA
1AVO	1C1M	1D2Z	1E2I	1F2S	2DGC	1HIA	1IRD	1K78	1LKY	1QAT	1SWQ	1VKX	2IZB
1AVP	1C1N	1D3A	1E2J	1F2T	2DL2	1HIO	1IRK	1K79	1LL1	1QAU	1SWR	1VLK	2IZC
1AVV	1C1O	1D3D	1E2K	1F2U	2DLI	1HIY	1IRS	1K81	1LL4	1QAV	1SWS	1VLS	2IZD

1AVW	1C1P	1D3G	1E2L	1F2X	2DUB	1HJ1	1IS0	1K86	1LLC	1QAX	1SWT	1VLT	2IZE
1AVX	1C1Q	1D3H	1E2M	1F39	2E2C	1HJ7	1IS7	1K88	1LLW	1QAY	1SWU	1VOL	2IZF
1AVY	1C1R	1D3P	1E2N	1F3G	2EBO	1HJ8	1IS8	1K8B	1LLZ	1QAZ	1T1D	1VPC	2IZG
1AVZ	1C1S	1D3Q	1E2O	1F3J	2ECP	1HJ9	1ISN	1K8R	1LM1	1QB0	1TAB	1VPF	2IZH
1AW5	1C1T	1D3T	1E2P	1F3L	2EFG	1HJA	1ISR	1K8T	1LM5	1QB1	1TAD	1VPN	2IZI
1AW7	1C1Y	1D3U	1E2V	1F3M	2EIA	1HJB	1ISS	1K90	1LM7	1QB2	1TAF	1VPP	2IZJ
1AW8	1C20	1D4H	1E34	1F3V	2ER0	1HJC	1ISV	1K93	1LM8	1QB3	1TAG	1VPS	2IZK
1AWC	1C25	1D4I	1E35	1F3Z	2ER6	1HJI	1ISW	1K94	1LML	1QB6	1TAL	1VPU	2IZL
1AWE	1C26	1D4J	1E36	1F44	2ER7	1HKS	1ISX	1K95	1LO6	1QB9	1TAR	1VR2	2JDX
1AWO	1C28	1D4P	1E37	1F46	2ER9	1HKT	1ISY	1K98	1LOH	1QBA	1TAS	1VSD	2JEL
1AXB	1C2B	1D4T	1E38	1F40	2EST	1HLE	1ISZ	1K9I	1LP3	1QBB	1TAT	1VSE	3GCB
1AXC	1C2D	1D4V	1E3A	1F4Q	2EZH	1HLG	1IT0	1K9J	1LP4	1QBH	1TAW	1VSF	3GCC
1AXM	1C2E	1D4W	1E3G	1F4R	2EZI	1HLK	1IT2	1K9O	1LPE	1QBJ	1TBA	1VSH	3GCT
1AXT	1C2F	1D5A	1E3K	1F4V	2EZK	1HMC	1IT3	1K9V	1LPL	1QBN	1TBG	1VSI	3GEO
1AY1	1C2G	1D5C	1E3U	1F4Z	2EZL	1HN4	1ITX	1KAC	1LPQ	1QBO	1TBN	1VSJ	3GRT
1AY4	1C2H	1D5D	1E3Y	1F50	2FAP	1HNN	1ITY	1KAF	1LPU	1QBQ	1TBO	1VSK	3GWX
1AY5	1C2I	1D5E	1E4I	1F51	2FGI	1HNR	1IU1	1KB2	1LQ1	1QBZ	1TC2	1VSL	3HFL
1AY8	1C2J	1D5F	1E42	1F5R	2FIB	1HNS	1IUZ	1KB4	1LQB	1QCS	1TC3	1VSM	3HHR
1AY9	1C2K	1D5H	1E4D	1F5S	2FX2	1HO1	1IV1	1KB6	1LQF	1QC7	1TCR	1VSR	3HIP
1AYF	1C2L	1D5J	1E4K	1F5W	2GAC	1HO4	1IV2	1KB9	1LQX	1QCA	1TEM	1VTK	3HTS
1AYY	1C2M	1D5L	1E4Q	1F60	2GAW	1HO5	1IV3	1KBC	1LR6	1QCF	1TF3	1VWA	4GEP
1AZ8	1C2O	1D5R	1E4S	1F66	3CTN	1HOW	1IV4	1KBE	1LRE	1QCN	1TFA	1VWB	4GRT
1AZG	1C30	1D5S	1E50	1F68	3ER3	1HP1	1IV6	1KBF	1LRM	1QCO	1TFB	1VWC	4HB1
1AZL	1C39	1D6N	1E53	1F6A	3ER5	1HPG	1IVB	1KBH	1LSH	1QCP	1TFE	1VWD	4HCK
1AZS	1C3A	1D6R	1E5B	1F6O	3ERD	1HPU	1IVC	1KBJ	1LTA	1QCW	1TFX	1VWE	5GEP
1AZT	1C3B	1D7E	1E5C	1F71	3ERT	1HPW	1IVD	1KC2	1LTB	1QD1	1TGB	1VWF	5GRT
1AZZ	1C3F	1D7M	1E5G	1F7C	3EST	1HQ0	1IVE	1KCA	1LTG	1QD6	1TGC	1VWG	5HCK
1B03	1C3G	1D7O	1E6F	1F7D	3EZB	1HQ6	1IVF	1KCG	1LTI	1QDB	1TGH	1VWH	5HVP
1B07	1C3I	1D7P	1E6I	1F7E	3EZE	1HQ8	1IVG	1KCQ	1LTM	1QDR	1TGN	1VWI	5JDW
1B08	1C3O	1D7W	1E6J	1F7K	3FAP	1HQM	1IVT	1KCW	1LTS	1QDT	1TGS	1VWJ	6GEP
1B0E	1C3Q	1D7X	1E7T	1F7M	3FIB	1HQO	1IX6	1KDH	1LTT	1QDU	1TGT	1VWK	6JDW
1B0F	1C3W	1D8B	1E7U	1F7N	3FX2	1HQR	1IX7	1KDK	1LU9	1QDV	1THR	1VWL	7GEP
1B0R	1C3X	1D8F	1E7V	1F7O	4CRX	1HR6	1IX8	1KDL	1LUA	1QDW	1THS	1VWM	7HSC
1B0V	1C48	1D8H	1E7Z	1F7P	4ER1	1HR7	1IXM	1KDM	1LUF	1QE5	1TIG	1VWN	7JDW
1B0X	1C4P	1D8I	1E81	1F7Q	4ER2	1HR8	1J4K	1KDS	1LVA	1QEY	1TII	1VWO	830C
1B10	1C4Q	1D8M	1E82	1F7Z	4ER4	1HR9	1J4L	1KDW	1LVY	1QFA	1TIO	1VWP	8GEP
1B12	1C4R	1D8S	1E87	1F89	4FAP	1HRE	1J4O	1KDX	1LW6	1QFH	2KAI	1VWQ	8JDW
1B28	1C4T	1D8Y	1E8I	1F8B	4FX2	1HRF	1J4P	1KE0	1LW7	1QFK	2KFN	1VWR	966C
1B3D	1C4Z	1D9C	1E8R	1F8C	5EAA	1HRH	1J4Q	1KE3	1LWH	1QFW	2KFZ	1VZV	9JDW
1B3Q	1C5B	1D9D	1E8T	1F8D	5ER1	1HRK	1J4V	1KE4	1LWJ	1QFY	2KI5	1WAS	

APPENDIX G

Names of proteins selected as test data for all- α secondary structural class

1E28	1AP6	1BS2	1CPE	1E2A	1FHB	1GJS	1HO8	1IRX	1KSO	1MNI	1QQB	1ZQB	3HSF
1O1M	1APX	1BSH	1CPO	1E2R	1FHE	1GKJ	1HQ1	1IT1	1KSS	1MRR	1QUU	1ZQK	3SDH
1O8M	1ARR	1BSM	1CPQ	1E3H	1FHF	1GKS	1HQB	1ITH	1KSU	1MTK	1R2F	2CPP	3SQC
1I2M	1ASH	1BUN	1CPR	1E3O	1F3	1GLP	1HRO	1J47	1KVV	1MZM	1R63	2CTS	4CTS
1A03	1AW9	1BUW	1CPW	1E7P	1FJG	1GNC	1HS5	1J77	1KW4	1N90	1RCE	2CY3	4GST
1A0A	1AX8	1BVB	1CSW	1ECA	1FK5	1GNW	1HSJ	1J7S	1KXU	1NBM	1RFB	2CYK	4P2P
1A0P	1B0B	1BXI	1CSX	1ECD	1FK6	1GO5	1HSS	1J7V	1KYZ	1NCY	1RMI	2CYM	4TF4
1A1W	1B0N	1BXR	1CTA	1EDI	1FNN	1GP2	1HUU	1J9I	1L1M	1NFO	1RMV	2DTR	5CPP
1A2S	1B1B	1BZ6	1CTD	1EEM	1FOX	1GSS	1HUZ	1JDL	1L1Y	1OCC	1ROM	2EZC	5CTS
1A32	1B3U	1BZP	1CTJ	1EFA	1FOY	1GTI	1HV4	1JFK	1LSJ	1OM2	1RSS	2FOW	5EAS
1A38	1B6W	1C3U	1CWW	1EGM	1FP1	1GUM	1HY1	1JFT	1LBG	1ORK	1SDK	2HBI	5EAU
1A5T	1B7Y	1C3Y	1D8K	1EGY	1FPO	1GWC	1HY8	1JJS	1LBU	1OSA	1SKT	2HMX	6CTS
1A63	1B82	1C40	1D9G	1EH2	1FPQ	1GWP	1HZV	1JP6	1LCC	1PGJ	1SPY	2LFB	6PAX
1A7E	1B85	1C52	1DCY	1EHF	1FT6	1GXD	1I54	1JPJ	1LE7	1PGT	1TIN	2LH1	7ICK
1ABM	1B8F	1C6R	1DIO	1EJ5	1FUR	1H0N	1I8O	1JS4	1LH6	1PIR	1TN4	2LYN	7ICV
1ABY	1B8R	1C7C	1DKE	1EMV	1FVK	1H54	1IBE	1JW2	1LH7	1PIS	1TNP	2MBW	7PRC
1AEB	1B8Z	1C7D	1DN1	1EN4	1FYZ	1H55	1IBR	1JW8	1LHS	1PMB	1UBY	2MM1	8ICE
1AEF	1BAL	1C8J	1DO1	1EN6	1FZP	1H5D	1IDQ	1JWG	1LJP	1PRU	1UNE	2MYC	8ICF
1AEH	1BC0	1C9L	1DO3	1ENJ	1G2Y	1H8E	1IDR	1K0N	1LJR	1PS1	1UXD	2PGH	8ICJ
1AEN	1BC1	1CB1	1DP5	1EQZ	1G3N	1H8H	1IF1	1K0O	1LLP	1QDM	1VNI	2PUE	8ICL
1AES	1BCY	1CCA	1DQE	1ERD	1G8E	1HAB	1IG5	1KB0	1LY9	1QHB	1VXA	2PUG	8ICN
1AEU	1BEA	1CFA	1DT1	1ETE	1GAX	1HD6	1IHF	1KBU	1MIQ	1QI8	1VXB	2SBL	8ICQ
1AH7	1BGC	1CG7	1DUX	1ETV	1GC7	1HDB	1IHG	1KIB	1M1R	1QI9	1VXD	2SPL	9ICC
1AHR	1B2	1CH1	1DVH	1ETW	1GDV	1HF8	1IJL	1KJS	1MFR	1QJB	1WJA	2SQC	9ICF
1AJ4	1BMF	1CH5	1DW0	1EWQ	1GEJ	1HJT	1IN4	1KKC	1MJM	1QL4	1WLA	2WRP	9ICG
1AJ9	1BOC	1CH9	1DW3	1EXI	1GEK	1HM3	1IN8	1KLP	1MLG	1QLK	1XGN	2YCC	9ICK
1AKD	1BOD	1CIP	1DWR	1F0M	1GGW	1HM6	1IOP	1KNR	1MLH	1QN2	1XGO	3CP4	9ICL
1ALU	1BOF	1CMY	1DXU	1F5O	1GGZ	1HMA	1IQU	1KQ7	1MLU	1QNN	1YMC	3CRD	9ICM
1AOQ	1BQ0	1COR	1DZ9	1FAG	1GIA	1HMD	1IQV	1KQV	1MMM	1QPI	1ZAC	3FYG	9ICV
1AP5	1BQ7	1CP4	1E0E	1FFY	1GJN	1HNX	1IRF	1KRF	1MNH	1QPW	1ZQA	3HHB	

APPENDIX H

Names of proteins selected as test data for all- β secondary structural class

1ald	1b2y	1bwn	1d6k	1edy	1finw	1gsl	1hyx	1ilu	1kb0	1mim	1qbs	1vom	2ubp
1alx	1b3i	1byo	1d6v	1eei	1for	1gts	1hzc	1im9	1kbb	1mlb	1qbt	1vqb	2wea
1a34	1b4j	1byt	1d6w	1eio	1fph	1gy2	1i18	1ind	1kbp	1mmd	1qcr	1vqf	2znc
1a39	1b4m	1bze	1d7f	1eiy	1fql	1gzi	1i1j	1iob	1kc5	1mpp	1qd0	1vqh	3cyh
1a3e	1b56	1c08	1d8t	1ejf	1fqm	1gzq	1i3r	1ipj	1kcc	1nbc	1qdc	1wbl	3enr
1a3y	1b6d	1c5m	1dar	1elo	1fqt	1h0i	1i3u	1ivl	1kck	1nch	1qg5	1wej	3ezm
1a42	1b7a	1c9i	1dc9	1ena	1fr4	1h8o	1i40	1ivv	1kd6	1ncj	1qhr	1xat	3gal
1a47	1b7k	1c9l	1df0	1ep3	1fsk	1h9w	1i4r	1j4s	1kdc	1neu	1qi4	1xyo	3hfm
1a5f	1b90	1cah	1dfa	1epa	1fve	1hao	1i50	1j58	1ken	1nib	1qjs	1ydb	3msp
1a5h	1b9m	1cai	1dhx	1epb	1fvi	1hav	1i6c	1j88	1kg1	1nic	1qlh	1yei	3tf4
1a7l	1baw	1cb4	1dl7	1epw	1fwa	1hb4	1i75	1j8k	1kj8	1nkr	1qni	1yej	3upj
1a8m	1bbr	1cbs	1dlf	1eqw	1fwj	1hbq	1i7u	1jbh	1kjg	1nlo	1qnw	1yge	3znc
1a9m	1bcd	1cd1	1dlp	1eso	1fxz	1hdf	1i81	1jbj	1kjj	1nlp	1qoi	1zsc	4fgf
1aaq	1bci	1cdi	1dlr	1et7	1fzo	1hdy	1i85	1jez	1kk3	1nmb	1qoo	2dlf	4gbq
1acy	1bck	1cel	1dms	1euj	1g0x	1hdz	1i8d	1jdd	1kn4	1nmg	1qpx	2eip	4npl
1adb	1bd9	1cgi	1dnv	1evh	1g1k	1hef	1i92	1jec	1kq2	1nnp	1qqd	2enb	4phv
1ae8	1bdr	1cgj	1dqk	1ex3	1g35	1hfu	1i94	1jer	1ksj	1nrq	1qs8	2ezm	4rcr
1ael	1bdx	1cgt	1dgo	1ey5	1g3s	1hge	1i95	1jff6	1ku8	1nrs	1qwf	2fgw	4sbv
1aex	1bdy	1cil	1dqt	1eyd	1g53	1hh2	1i9b	1jgx	1kuj	1oac	1r24	2hmg	4upj
1afv	1bf2	1cin	1duc	1eyl	1g6p	1hh6	1i9e	1jgy	1kvp	1odx	1rkr	2hrv	5cyh
1agb	1bf0	1ckm	1dud	1eys	1g8r	1hii	1i9o	1jgz	1kzk	1okm	1ruo	2ila	5gch
1agf	1bgy	1clo	1dup	1ezl	1g94	1hla	1i9q	1jhk	1l0s	1ose	1scr	2jcw	5gds
1aif	1bi3	1cms	1duy	1f0l	1g9f	1hm8	1iaz	1jht	1ley	1otc	1shf	2jxr	5msi
1ajk	1bic	1cnc	1dvs	1f45	1g9h	1hmr	1ib4	1ji2	1lek	1ovw	1sky	2kau	5nuc
1ajo	1bla	1cne	1dvz	1f4w	1g9o	1hmt	1ibf	1jje	1lem	1p03	1slc	2lal	5prc
1ajx	1blz	1cng	1dyw	1f58	1gbg	1hn2	1ibh	1jnj	1lf7	1pbo	1snp	2mha	5tss
1anu	1bmw	1cnj	1dz0	1f5l	1gcf	1hoc	1ic7	1jo8	1lg6	1pcr	1sre	2mip	6fab
1aom	1bn8	1cpi	1e00	1f8a	1gg3	1hpv	1icm	1jok	1lib	1pfc	1srf	2mys	6gch
1aoq	1bnn	1cr7	1e0f	1f92	1ggb	1hql	1ict	1joq	1lid	1pgr	1sro	2ohx	6prc
1aqd	1bnt	1crm	1e0l	1faj	1ggt	1hsa	1ida	1jpt	1lmk	1pgs	1ste	2pab	6taa
1ar2	1bpv	1cru	1e1r	1fak	1gh5	1hsb	1ie5	1jri	1lo4	1phs	1stg	2pia	7hvp
1as7	1bqk	1csk	1e2r	1fdl	1ghv	1htb	1iff	1jsn	1lox	1pin	1stm	2qil	7ilb
1ava	1bqs	1csq	1e32	1fem	1ghy	1hte	1ifh	1jtp	1lr5	1pkl	1sxb	2rcs	7odc
1avd	1bso	1cwk	1e3p	1fg2	1gic	1huh	1igf	1ju3	1lt4	1pkt	1tei	2ren	
1awj	1bsq	1cwo	1e57	1fgm	1gid	1huk	1igi	1jyt	1lul	1pky	1thv	2rhe	
1awr	1bsy	1cwp	1e67	1fgt	1gix	1hv9	1ihg	1jz1	1lv1	1pnd	1tiu	2roy	
1awt	1btn	1cxe	1e9o	1fgv	1gka	1hvc	1ike	1jz3	1lve	1png	1tng	2rta	
1awv	1bto	1cxh	1e9p	1fiv	1glh	1hvr	1iko	1jz4	1m4h	1psa	1ttf	2rto	
1azn	1bv8	1czm	1e9y	1fkn	1gm5	1hwm	1ikq	1jzf	1mab	1pse	1ubp	2seb	
1azu	1bvl	1d0z	1eb0	1flr	1gmm	1hxb	1ile	1k5j	1mcl	1pst	1ulp	2sil	
1b0w	1bvq	1d5g	1eb1	1fmm	1gog	1hxm	1ilf	1k6n	1mcr	1qa7	1vca	2spg	
1b1l	1bw4	1d5z	1ed3	1fmv	1grw	1hys	1ilt	1k9a	1mdc	1qaw	1vhp	2try	

APPENDIX I

Names of proteins selected as test data for α/β secondary structural class

1lgs	1b16	1c1h	1dd8	1e5t	1f0k	1gaj	1h86	1ieq	1jxa	1kyv	1ptg	1tpd	2wsy
12gs	1b1x	1c2t	1ddr	1e6l	1f2o	1gc5	1hdo	1ig3	1jyf	1kyy	1pxc	1tpt	2xyl
14gs	1b23	1c3j	1dds	1e6m	1f3a	1gc9	1hdq	1ihy	1jyl	1l0p	1pym	1ttt	2yas
1a05	1b2h	1c3p	1de1	1e6w	1f4a	1gcg	1hdy	1ii2	1jyv	1l0v	1qan	1tyc	3cox
1a0c	1b2y	1c3r	1deh	1E70	1f7b	1gcy	1het	1iiv	1jyx	1l1f	1qao	1tyf	3dap
1a0e	1b3y	1c4a	1del	1E72	1f8m	1gdd	1hex	1ilo	1jyy	1l5h	1qbk	1udb	3eza
1a4z	1b3z	1c7j	1deo	1E73	1f8r	1gdr	1hf6	1im5	1jz6	1l7v	1qcd	1ukz	3gtu
1a50	1b40	1c7r	1dfh	1e7p	1fa9	1gef	1hfj	1in5	1jz7	1lcp	1qf4	1ula	3hsc
1a53	1b46	1c8v	1dfo	1e7y	1fbn	1gey	1hfk	1j49	1jzq	1lii	1qf9	1vdc	3lir
1a5a	1b4h	1cbf	1dg3	1e8d	1fc5	1gfs	1hg0	1j4e	1k0i	1ljl	1qfl	1vfe	3man
1a5k	1b5i	1cc0	1dg7	1e9f	1fcj	1gg2	1hh1	1j6z	1k0j	1lkc	1qfs	1vfn	3mht
1a5y	1b6h	1cc4	1dgd	1eb3	1ffy	1gg8	1hjr	1j79	1k0m	1llq	1qgn	1vid	3pbg
1a71	1b6t	1cc6	1dig	1eb8	1fg4	1gin	1hm8	1j7x	1k3d	1lol	1qhf	1vom	3pgm
1a8e	1b74	1cer	1dik	1ebg	1fgh	1git	1hm9	1j8m	1k3u	1lon	1qhg	1vru	3ptd
1a8y	1b8s	1cex	1dio	1ecv	1fj4	1gjp	1hmv	1j90	1k4h	1lp6	1qi4	1vvr	3tec
1a95	1b93	1cgl	1dju	1ecx	1fj8	1giw	1hna	1j99	1k5h	1lpm	1qj4	1waf	3thi
1a97	1b9h	1cgu	1dkn	1eel	1fjx	1gl9	1hnb	1jae	1k7e	1lqt	1qj5	1wal	3trx
1a9z	1b9j	1cke	1dkp	1ef2	1fld	1glf	1hnc	1jah	1k7h	1lr1	1qka	1wkd	3xim
1aa6	1bd3	1cml	1dlj	1efc	1fmj	1glj	1hnc	1jb9	1k8q	1lst	1qmp	1wke	4cpa
1adc	1be6	1cpx	1dls	1efh	1fnw	1gnp	1hon	1jbe	1kar	1ltd	1qor	1wod	4dmr
1ade	1beu	1cqx	1dmr	1efm	1fnc	1gnq	1hop	1jcg	1kdy	1lth	1qqt	1xib	4grl
1ag1	1bf3	1cr6	1dms	1egm	1fnm	1goo	1hot	1jcl	1kdz	1luc	1qtq	1xih	4gst
1agp	1bfu	1crr	1doe	1eix	1fo5	1got	1hp4	1jct	1kek	1m6b	1qui	1xim	4mbp
1agx	1bg3	1ese	1dpf	1ej2	1fok	1gpj	1hq5	1jdt	1ker	1maa	1quk	1xlf	4mdh
1ah3	1bg5	1ctq	1dr4	1eju	1fp1	1gpm	1hqh	1jeo	1kew	1mdq	1qva	1xys	4pbg
1aip	1bga	1cub	1dr5	1ek6	1fqa	1gpw	1hqu	1jfs	1kfd	1mmn	1qvb	1xza	4pga
1aja	1bh6	1cuf	1dr7	1eke	1fqf	1gq6	1hqy	1jff	1kfi	1mpd	1rae	1xzd	4pgt
1ajc	1bhj	1cuy	1dre	1el9	1frn	1gra	1hso	1jfu	1kgs	1mpt	1rbt	1xzl	4q2l
1ak1	1bkg	1cyr	1drv	1eli	1fs4	1grn	1hsz	1jg3	1kgu	1mug	1rco	1ypn	4ukd
1al7	1ble	1cw0	1drw	1els	1fsf	1grx	1hur	1jg9	1kh2	1nas	1rev	1yub	5gpb
1ali	1blf	1cw4	1dss	1elu	1fu7	1gsj	1hvr	1jhm	1kh9	1nbm	1rka	1zrn	5nul
1alj	1bli	1cw7	1dt9	1eol	1fua	1gtm	1hvq	1jik	1kic	1ngf	1rrg	1zym	5tgi
1aln	1bll	1cyg	1dtn	1eok	1fvp	1gtr	1hwy	1jji	1kjz	1nhp	1rt3	2ctb	5tim
1amf	1bo5	1czh	1dtz	1ep3	1fwb	1gts	1hwz	1jjk	1kkp	1noi	1rtj	2dhc	5tmp
1an9	1bof	1d04	1du4	1epx	1fwi	1gtz	1hx0	1jjv	1km0	1noj	1rv5	2edc	5xia
1ao8	1bp5	1dlc	1dvl	1eqc	1fww	1guy	1hx7	1jlv	1km6	1ohj	1rvb	2gsq	6cpa
1aqj	1bpw	1dlg	1dxo	1erv	1fx0	1gvf	1i13	1joa	1kmh	1ola	1rx7	2gtu	6dff
1aqw	1bqg	1dls	1dxq	1es8	1fxx	1gwj	1i2d	1jol	1kmj	1omp	1rx8	2hgs	6gpb
1as0	1br1	1d2c	1dyh	1ese	1gl8	1gzv	1i2p	1jpi	1kms	1ose	1s01	2his	6xim
1ava	1bsu	1d3c	1dxx	1eth	1gli	1h11	1i3k	1jpv	1knv	1oxm	1sky	2hsd	7enl
1avt	1bt4	1d4a	1dzz	1eve	1g3q	1h11	1i43	1jqg	1kor	1oyb	1sma	2mys	7gss

1axd	1bu5	1d4x	1e0x	1ewq	1g4h	1h4z	1i52	1jr1	1kpi	1oyc	1som	2pfl	7req
1az0	1bvh	1d6s	1E26	1ewx	1g5p	1h5s	1i6h	1jr2	1kqf	1pbb	1sub	2pgd	8ruc
1az2	1bvr	1d9e	1e2b	1ewy	1g66	1h61	1i6k	1jtf	1kr1	1pbn	1sur	2pia	9rub
1az3	1bv0	1dag	1e3s	1eye	1g7t	1h77	1i6m	1jtk	1krc	1pdo	1tah	2pmt	
1az9	1bw0	1dai	1e3w	1ez0	1g7v	1h78	1i6n	1ju9	1ksu	1ped	1tcd	2puf	
1azw	1bwk	1dam	1E43	1ez9	1g8f	1h7f	1ia2	1jv1	1kte	1phh	1tde	2rkm	
1azy	1bxc	1dar	1E48	1eza	1g98	1h71	1idc	1jv3	1kvb	1pii	1tho	2sfp	
1b0h	1bye	1dc6	1e4o	1ezk	1g9u	1h7p	1idd	1jvx	1kw1	1pkl	1tim	2shp	
1b0j	1c14	1dce	1e5j	1ezo	1ga8	1h7w	1ie7	1jvy	1kxh	1pmt	1tip	2sic	
1b0z	1c1b	1dck	1e5l	1ezz	1gad	1h82	1iej	1jwa	1kyh	1psd	1toa	2tmk	

APPENDIX J

Names of proteins selected as test data for $\alpha+\beta$ secondary structural class

1081	1b2x	1c0g	1d9n	1eb6	1ff2	1gfr	1i4x	1jql	1l31	1mfi	1qn6	1tlc	2sdf
1271	1b37	1c0p	1d9w	1ec6	1ffv	1gft	1i6s	1jru	1l36	1mi2	1qn8	1tli	2sni
1301	1b3s	1c22	1dao	1ec8	1fh0	1ggy	1i6u	1jrz	1l43	1mkb	1qnk	1tsg	2tbd
1351	1b49	1c3t	1dax	1ed3	1fik	1gij	1i8v	1js0	1l44	1mms	1qnm	1tsx	2tli
1381	1b4n	1c54	1dbf	1ef7	1fim	1gkx	1i9a	1jwh	1l61	1mns	1qqy	1u2f	2ubp
1451	1b5d	1c62	1dch	1egw	1fit	1gmr	1iab	1jx0	1l65	1mom	1qr0	1ubp	2yfp
1491	1b5e	1c63	1dcj	1eiy	1fjo	1gmu	1iac	1jy8	1l67	1mra	1qr5	1uia	3eip
1541	1b6v	1c6b	1dcp	1ejv	1fjt	1gn2	1ian	1k0i	1l69	1mrk	1qrk	1uif	3fit
1591	1b7o	1c6j	1ddu	1ekr	1fjv	1gra	1iez	1k0j	1l76	1msh	1qrm	1uig	3gb1
1621	1b7y	1c6k	1dfi	1ema	1fjw	1gsh	1ifu	1k0k	1l8t	1mxb	1qsa	1up1	3hck
1731	1b8s	1c6t	1dgs	1emb	1fkt	1gtt	1igc	1k2e	1l91	1mxb	1qte	1var	3hfm
1751	1b9c	1c7h	1dhm	1emk	1fma	1guz	1ihy	1k3t	1l9n	1naw	1qtk	1vcc	3il8
1761	1b9o	1ca5	1div	1ems	1fn5	1gx1	1ii7	1k4p	1laa	1ndc	1qtz	1vew	3lzt
1771	1b9r	1ca6	1dkk	1en6	1fpy	1gy5	1ikt	1k50	1lba	1ndk	1quh	1vjw	3phy
1821	1bb3	1cc7	1dm2	1enf	1fwb	1h2h	1il2	1k52	1lbu	1ndp	1quo	1vkl	3rsd
1901	1bb7	1ce6	1dm9	1epw	1fwc	1h4t	1il9	1k53	1lek	1nhr	1rax	1wfc	3rub
1931	1bc2	1cf3	1dmn	1eqe	1fwd	1h6a	1ilq	1k5b	1leo	1njb	1rbj	1wqm	3sic
1a0i	1bf4	1civ	1dmq	1erk	1fxa	1h6r	1inq	1k6z	1lgw	1nsq	1reo	1xei	3tec
1a0p	1bfj	1cj9	1dnk	1esf	1fxv	1h81	1inu	1k7j	1lhj	1oac	1rex	1xra	3ubp
1a16	1bgq	1ckd	1do9	1esp	1fzo	1h83	1ip5	1k8d	1li1	1ouc	1rez	1ypr	4dbv
1a48	1bi8	1cmk	1doy	1ev5	1gl8	1h9b	1ipb	1k8w	1li6	1pad	1rgl	1yys	4hoh
1a5e	1bia	1cns	1ds7	1evx	1glq	1hby	1ir1	1kbi	1lk7	1pbh	1rha	1zdi	4jdw
1a6q	1bid	1cnu	1dth	1ewc	1glx	1hdm	1ir7	1kbl	1lld	1pfd	1rms	2dap	4lyo
1a7x	1bj3	1coy	1duz	1ewj	1g2s	1hel	1itp	1kbu	1lnf	1pfl	1mo	2dnj	4lzt
1a9e	1bkf	1cpz	1dv2	1eym	1g3i	1hew	1ius	1kd0	1lni	1pfm	1rnu	2glt	4rsk
1aa8	1bkj	1cqm	1dxl	1f0w	1g3o	1hfg	1ivw	1ke9	1lo9	1pme	1rnv	2hla	5crx
1acf	1bms	1cqn	1dyb	1flh	1g51	1hh2	1j6x	1keh	1low	1ppo	1rny	2hqi	5fd1
1ady	1bn5	1cqy	1dyg	1f29	1g5b	1hhj	1j9o	1kf4	1loz	1prt	1rod	2iad	5gcn
1afk	1bni	1ctp	1e0z	1f2c	1g5s	1hi2	1jaw	1kf7	1lpf	1pxa	1roe	2jdw	5gsp
1age	1bo0	1ctw	1E28	1f2g	1g5u	1hs6	1jc4	1khh	1lqw	1pys	1rsm	2lym	5lzm
1ah5	1bo4	1cu3	1e2t	1f2k	1g6t	1hsa	1jdf	1kmy	1lsy	1qd9	1rtm	2lzm	6msf
1ai5	1bo7	1cu5	1e3h	1f2l	1ga7	1ht1	1je3	1knd	1lv1	1qg7	1sce	2mef	6pad
1ako	1bp1	1cx6	1e3n	1f3f	1gb4	1htn	1jef	1knk	1lyg	1qh5	1sei	2meg	6tli
1alc	1bql	1cx7	1e5l	1f4c	1gbw	1hwm	1jfm	1ko9	1lyh	1qjd	1sgk	2mei	6tmn
1amw	1br6	1cz4	1e5w	1f4f	1gby	1hx3	1jid	1kp2	1lyi	1qjg	1sib	2mll	7msf
1aor	1brj	1cz5	1e6z	1f5v	1gc7	1hyg	1jis	1ksi	1lz1	1qjj	1sml	2mnr	7rat
1ar0	1bsc	1d09	1E77	1f77	1gd1	1hyn	1jk8	1ktd	1lz5	1qk9	1srp	2nck	7tli
1aud	1bsm	1dlj	1e7k	1f8r	1ge1	1hyt	1jkf	1ktg	1lzd	1qlm	1stc	2one	8ldh
1ay7	1buh	1d2w	1e7x	1f93	1ges	1hyw	1jkj	1kv3	1lzg	1qlt	1tbe	2pad	8lyz
1ayz	1bvz	1d4d	1e7y	1fa6	1geu	1iox	1jmg	1kvk	1m44	1qm9	1tby	2pii	8pch
1azq	1bwe	1d5x	1e8l	1fcd	1gf0	1i32	1jmh	1kxy	1mdt	1qmi	1tcy	2ple	8rsa
1b27	1by8	1d6o	1e9z	1fcl	1gf8	1i3i	1jnd	1kzj	1meg	1qmr	1tec	2qil	8thn
1b2t	1byf	1d7i	1ea6	1fd9	1gfe	1i4f	1jpo	1l17	1mff	1qn5	1tjs	2scu	

APPENDIX K

Names of proteins in database for all- α secondary structural class

0c3a	1aum	1bw6	1czj	1e90	1fch	1gfv	1hs5	1jeq	1kvo	1mwd	1quu	1zqt	3cpp
101m	1auw	1bwo	1d1d	1e91	1fcs	1gjy	1hs6	1jf0	1kvw	1mxl	1rlb	1zug	3crd
102m	1av8	1bx9	1d1l	1e96	1fcx	1gk2	1hs7	1jf2	1kvx	1myf	1rlr	1zym	3cro
103m	1avc	1bxi	1d1m	1e9h	1fcy	1gk3	1hsj	1jf3	1kvy	1myg	1r2a	20gs	3crx
104m	1avm	1bxm	1d1o	1e9x	1fcz	1gkj	1hsm	1jf4	1kw2	1myh	1r2f	21gs	3csc
105m	1avo	1bxr	1d2d	1ea1	1fdh	1gkm	1hsn	1jfa	1kw4	1myi	1r63	22gs	3csm
106m	1avs	1byl	1d2t	1ea3	1fdk	1gks	1hsr	1jfb	1kwf	1myj	1r69	256b	3ctn
107m	1aw9	1bye	1d2v	1ea4	1fe2	1gkx	1hss	1jfc	1kwj	1myk	1rap	2a0b	3cyr
108m	1awc	1byt	1d2z	1ea8	1fe5	1gkz	1hst	1jfg	1kx2	1myl	1raq	2a8v	3cyt
109m	1ax8	1bz0	1d3u	1eak	1few	1gli	1hsy	1jfi	1kx7	1mym	1rcb	2abd	3erd
10gs	1axd	1bz1	1d4c	1eb4	1fex	1glm	1ht5	1jfi	1kxm	1myp	1rec	2abk	3ert
110m	1axf	1bz4	1d4d	1eb7	1ff1	1gin	1ht8	1jfk	1kxn	1myt	1red	2af8	3eza
111m	1axi	1bz6	1d4e	1ebc	1ffh	1glp	1ht9	1jfs	1kxp	1mzl	1rce	2asr	3ezb
112m	1axn	1bzp	1d4u	1ebd	1ffk	1gliq	1hta	1jft	1kxu	1mzm	1reg	2atj	3eze
11gs	1ayg	1bzt	1d5l	1ebe	1ffu	1gm0	1htj	1jgz	1kyo	1n15	1rci	2av8	3fap
12gs	1ayi	1bzz	1d5n	1ebm	1ffv	1gm5	1hu3	1jgc	1kyw	1n50	1rcp	2bbm	3fis
13gs	1ayp	1c02	1d5v	1ebt	1ffy	1gmz	1hue	1jgg	1kyz	1n90	1rcs	2bbn	3fyg
14gs	1ayx	1c03	1d5y	1ebv	1fg9	1gn2	1hul	1jgs	1kz7	1nbb	1rep	2bby	3gly
155c	1azi	1c07	1d7w	1eca	1fgj	1gn3	1huo	1jh9	1kzg	1nbn	1res	2bca	3gss
16gs	1azs	1c0w	1d8b	1ecd	1fgm	1gn4	1hus	1jhf	1l0o	1ncx	1ret	2bcb	3gst
17gs	1azf	1c15	1d8d	1eci	1fgo	1gn6	1huu	1jhg	1l0v	1ncy	1rfb	2bcc	3gtu
18gs	1azy	1c1j	1d8e	1ecm	1fgq	1gnc	1huw	1jhh	1l1m	1ncz	1rgp	2bct	3gwx
19gs	1b01	1c1k	1d8j	1ecn	1fgr	1gne	1huz	1ji5	1l1y	1neq	1rh2	2bmh	3had
19hc	1b06	1c20	1d8k	1eco	1fgt	1gni	1hv4	1ji7	1l2a	1ner	1rhg	2bp2	3hbi
1a00	1b0b	1c26	1d8l	1ecw	1fha	1gnj	1hv6	1jia	1l2i	1new	1rmi	2bpf	3hdd
1a01	1b0n	1c2n	1d8u	1ed1	1fhb	1gnw	1hvd	1jig	1l2j	1nfl	1rmv	2bpg	3hdd
1a03	1b0x	1c2r	1d9c	1edi	1fhe	1go3	1hve	1jin	1l2k	1nfn	1rnl	2bpp	3hdb
1a04	1b1b	1c30	1d9g	1edj	1fhf	1go5	1hvf	1jio	1l3e	1nfo	1rmr	2btv	3hhr
1a0a	1b1g	1c3c	1daq	1edk	1fhj	1god	1hvg	1jip	1l3l	1ngl	1rom	2c2c	3hsf
1a0b	1b1u	1c3d	1dat	1edl	1fi3	1got	1hw1	1jj2	1l3o	1ngr	1rop	2cbl	3hts
1a0f	1b22	1c3o	1dav	1edu	1fi5	1gp2	1hw2	1jj6	1l5j	1nhm	1rpe	2ccp	3icb
1a0p	1b25	1c3u	1db1	1ee4	1fi6	1gpj	1hw5	1jj8	1l6e	1nhn	1rpo	2ccy	3ink
1a0r	1b28	1c3y	1db4	1ee5	1fi7	1gqm	1hwg	1jjc	1l6j	1nih	1rpr	2cdv	3lbd
1a0u	1b33	1c3z	1db5	1eem	1fi9	1grj	1hwh	1jir	1l6n	1nir	1rro	2cep	3lhb
1a0v	1b3q	1c40	1db7	1eer	1fia	1grl	1hx1	1jjs	1l7c	1nk2	1rse	2cgp	3lir
1a0w	1b3u	1c52	1db8	1eex	1fin	1grn	1hx8	1jju	1l8c	1nk3	1rss	2cmm	3lyn
1a0x	1b43	1c53	1db9	1efl	1fio	1gsd	1hxd	1jk0	1l8r	1nkd	1rtp	2cnp	3mba
1a0y	1b47	1c5a	1dbc	1ef4	1fip	1gse	1hxi	1jko	1l8w	1nkl	1rum	2cp4	3mdd
1a0z	1b48	1c60	1dbh	1efa	1fiq	1gsf	1hy0	1jkp	1l8y	1nno	1ruo	2cpg	3mde
1a17	1b4a	1c6r	1dcc	1efr	1fj0	1gsq	1hyl	1jkq	1l8z	1nol	1rvw	2cpo	3mds
1a1u	1b4c	1c6s	1dce	1efu	1fjl	1gss	1hy5	1jkr	1l9b	1noo	1ryc	2cpp	3ngl
1a1w	1b4f	1c72	1dcn	1efy	1fk0	1gsu	1hy8	1jku	1l9j	1nre	1ryt	2crk	3orc

1a1z	1b4n	1c74	1dcy	1eg3	1fk1	1gsy	1hyp	1jkv	1la3	1ns1	1rzt	2cro	3p2p
1a22	1b4u	1c75	1dd3	1eg4	1fk2	1gt8	1hz4	1jkw	1la6	1nsg	1sae	2crx	3pal
1a23	1b4w	1c7b	1dd4	1egc	1fk3	1gta	1hzi	1jl6	1lb1	1ntc	1saf	2csc	3pat
1a24	1b59	1c7c	1ddf	1egd	1fk4	1gtb	1hzu	1jl7	1lbd	1nub	1sag	2csm	3pax
1a26	1b5l	1c7d	1ddn	1ege	1fk5	1gte	1hzv	1jli	1lbg	1obm	1sah	2cth	3pgh
1a28	1b67	1c7m	1ddx	1egm	1fk6	1gth	1i08	1jlt	1lbk	1occ	1sai	2ctn	3pgt
1a29	1b68	1c7v	1de4	1egu	1fk7	1gti	1i0a	1jlv	1lbu	1ocd	1saj	2cts	3prc
1a2a	1b6a	1c7w	1dee	1egv	1fkm	1gto	1i0h	1jlw	1lcc	1oco	1sak	2cxb	3prg
1a2f	1b6q	1c7y	1deg	1egy	1fli	1gtu	1i11	1jlm4	1lcd	1ocp	1sal	2cy3	3psr
1a2g	1b6w	1c82	1der	1eh2	1flo	1gu6	1i1g	1jlm6	1ld7	1ocr	1san	2cyk	3r1r
1a2i	1b7l	1c8i	1df0	1eh6	1flp	1gu9	1i1r	1jme	1ld8	1oct	1sav	2cym	3sak
1a2j	1b72	1c8j	1dfk	1eh7	1fm6	1guh	1i1s	1jmr	1ldd	1ocz	1scf	2cyp	3sdh
1a2l	1b79	1c9b	1dfl	1eh8	1fm9	1guk	1i27	1jms	1ldj	1odd	1sch	2dhh	3sdp
1a2m	1b7t	1c9i	1dg3	1ehe	1fmi	1gul	1i2t	1jmw	1ldk	1oel	1scm	2dtr	3sqc
1a2s	1b7v	1c9l	1dgj	1ehf	1fn7	1gum	1i36	1jmx	1le2	1ofj	1sct	2dvh	3tf4
1a2x	1b80	1ca1	1dgn	1ehg	1fmn	1gux	1i37	1jnz	1le4	1ofk	1sdk	2eia	3vhb
1a32	1b82	1cb1	1dgp	1ehj	1fo2	1gv3	1i38	1jni	1le6	1olg	1sdl	2end	3wrp
1a36	1b85	1cb8	1dgs	1ei7	1fo3	1gvh	1i3d	1jnr	1le7	1olh	1ser	2erl	3ygs
1a37	1b86	1cbl	1dgt	1eia	1foc	1gw5	1i3e	1jnz	1le8	1om2	1sfe	2eza	451c
1a38	1b89	1cbm	1dgu	1eij	1foe	1gwc	1i4a	1jow	1lea	1omd	1sfv	2ezb	4aah
1a3d	1b8c	1cc5	1dgv	1eiy	1fok	1gwg	1i4s	1joy	1leb	1opc	1sfw	2ezc	4bp2
1a3f	1b8d	1cca	1dil	1ej3	1fow	1gwp	1i4y	1jp6	1lfb	1or2	1sgg	2ezh	4ccp
1a3n	1b8f	1ccb	1dii	1ej5	1fox	1gwx	1i4z	1jp8	1lfb	1or3	1sig	2ezi	4ccx
1a3o	1b8i	1ccc	1dio	1ejl	1foy	1gxd	1i54	1jp9	1lffm	1orc	1skn	2ezk	4cln
1a43	1b8l	1ccd	1diq	1ejy	1fp1	1gxx	1i55	1jpb	1lga	1ork	1skt	2ezl	4cox
1a4f	1b8r	1cce	1diy	1eku	1fp2	1gxh	1i5n	1jpi	1lh1	1osa	1sky	2ezx	4cp4
1a4o	1b8x	1ccg	1diz	1el4	1fp3	1gxp	1i5t	1jpl	1lh2	1otp	1slm	2ezy	4cpp
1a4p	1b8z	1cch	1djl	1elk	1fpo	1gxq	1i6d	1jpn	1lh3	1out	1sly	2ezz	4cpv
1a52	1b9l	1cci	1dj5	1elr	1fpp	1gy3	1i6e	1jpp	1lh5	1ouu	1smg	2fal	4cro
1a56	1b9a	1cej	1dj8	1elw	1fpq	1gyg	1i6v	1jpr	1lh6	1oxa	1smt	2fam	4crx
1a59	1b9m	1cck	1djg	1em1	1fps	1gyo	1i6z	1jpw	1lh7	1oxy	1spe	2fap	4csc
1a5j	1b9n	1ccl	1djh	1em9	1fpw	1gyz	1i77	1jpz	1lht	1p2p	1spg	2fhh	4cts
1a5t	1b9x	1ccp	1dji	1emr	1fpw	1gza	1i7g	1jqc	1lht	1pa2	1spy	2fha	4fap
1a62	1b9y	1ccr	1djw	1emu	1fqi	1gzb	1i7i	1jqj	1li5	1pal	1sqc	2fhe	4fis
1a63	1ba5	1cd3	1dix	1emv	1fqj	1gzx	1i7w	1jqj	1li7	1par	1sra	2fok	4gss
1a6d	1bab	1cd9	1dij	1emy	1fqk	1h0n	1i7x	1jr3	1lia	1pax	1sss	2fow	4gst
1a6e	1baj	1cdl	1djz	1en4	1fqv	1h0o	1i7y	1jr5	1lih	1pbv	1swm	2frc	4gtu
1a6g	1bal	1cdm	1dk1	1en5	1fr0	1h19	1i8o	1jr8	1lin	1pbw	1sym	2fus	4hbi
1a6i	1baq	1cdn	1dk2	1en6	1fs0	1h1j	1i8p	1jr9	1lip	1pbx	1tln	2gdm	4hbb
1a6k	1bax	1cdp	1dk3	1en7	1fs1	1h4i	1i8q	1jra	1lis	1pd2	1tad	2giw	4icb
1a6m	1bay	1ce8	1dk5	1enh	1fs2	1h4j	1ial	1jro	1ljp	1pdn	1taf	2gr	4lbd
1a6n	1baz	1ced	1dk8	1eni	1fs7	1h4l	1iap	1jrp	1ljp	1per	1tag	2gmf	4mba
1a6s	1bbb	1ceh	1dke	1enj	1fs8	1h4r	1iar	1jrx	1ljw	1pes	1tam	2gsq	4mbn
1a75	1bbc	1cei	1dkf	1enk	1fs9	1h54	1ib1	1jry	1lk3	1pet	1taq	2gsr	4p2p
1a76	1bbh	1cem	1dl2	1enw	1fse	1h55	1ib2	1jrz	1lki	1pfr	1tau	2gss	4pal
1a77	1bbi	1cf7	1dli	1eo0	1fsh	1h57	1ib3	1js4	1lkv	1pge	1tba	2gst	4pax
1a7d	1bbn	1cfa	1dlj	1eog	1fsl	1h58	1ibe	1js8	1lky	1pgf	1thg	2gtu	4pgt
1a7e	1bby	1cfe	1dlw	1eoh	1fsx	1h5a	1ich	1jsa	1lli	1pgg	1tc3	2gwx	4prc
1a7m	1bc0	1cfd	1dly	1eoi	1ft1	1h5c	1id3	1jsp	1lli	1pgj	1tcf	2hbc	4prg

1a7v	1bc1	1cfe	1dm1	1eoq	1f2	1h5d	1ldq	1jst	1llp	1pgn	1tco	2hbd	4r1r
1a7w	1bc3	1cfp	1dm5	1eq1	1f5	1h5e	1ldr	1jsu	1lm2	1pgo	1tes	2hbe	4sdh
1a8a	1bc5	1cg5	1dmo	1eqf	1f6	1h5f	1lds	1jsw	1lm3	1pgp	1tf4	2hbf	4tf4
1a8b	1bc7	1cg7	1dn1	1eqg	1f9	1h5g	1idu	1jt0	1lmb	1pgq	1tff	2hbg	4tnc
1a8c	1bc8	1cg8	1dnp	1eqh	1ft5	1h5h	1idy	1jt6	1lmz	1pgr	1tfr	2hbi	4vnb
1a8h	1bc9	1cgm	1dnu	1eqz	1ft	1h5i	1idz	1jtb	1lnh	1pgt	1thb	2hbs	5ccp
1a8o	1bcc	1cgn	1dnw	1erc	1ftz	1h5j	1ie8	1jtx	1loh	1pha	1tnq	2hco	5cox
1a8v	1bcf	1cgo	1dny	1erd	1fuo	1h5k	1ie9	1jty	1lot	1phb	1tn4	2hdc	5cp4
1a9w	1bcn	1cgp	1do1	1ere	1fup	1h5l	1ier	1jum	1lox	1phc	1tnd	2hdd	5cpp
1a9x	1bcw	1chl	1do3	1erp	1fuq	1h5m	1ies	1juo	1lpe	1phd	1tnp	2hdh	5cpv
1aa2	1bcy	1ch2	1do4	1err	1fur	1h67	1if1	1jup	1lpq	1phe	1tnq	2hfh	5cro
1aa4	1bcz	1ch3	1do7	1ery	1fv0	1h6g	1ifa	1juq	1lq1	1phf	1tus	2hbb	5crx
1aa7	1bdc	1ch4	1dof	1etc	1fvj	1h6k	1ifw	1jus	1lqc	1phg	1tat	2hhd	5csc
1aab	1bdd	1ch5	1dog	1etd	1fvk	1h6o	1ify	1jvr	1lqs	1phn	1tnw	2hhe	5csm
1ab3	1bdh	1ch7	1doq	1ete	1fvv	1h6p	1ig5	1jw2	1lre	1pir	1tnx	2hio	5cts
1abm	1bdi	1ch9	1dov	1etk	1fw1	1h7c	1ig6	1jw8	1lri	1pis	1top	2hmj	5cyt
1abs	1bdj	1chh	1dow	1eto	1fw4	1h7i	1ig7	1jwd	1lrp	1pmb	1tpt	2hmx	5eas
1abv	1bdt	1chi	1dp3	1etp	1fw6	1h7w	1ign	1jwe	1lrz	1pmt	1trf	2hmz	5eat
1abw	1bdv	1chj	1dp5	1etq	1fwz	1h7x	1igv	1jwf	1lsh	1pnb	1tro	2hoa	5eau
1aby	1be2	1chu	1dp7	1etv	1fx0	1h88	1igx	1jwg	1ltw	1pnr	1trr	2hp8	5fwg
1ac1	1be3	1ci4	1dpj	1etw	1fx7	1h89	1igz	1jwl	1lus	1poa	1tx4	2hpd	5gss
1ac4	1bea	1cie	1dpr	1etx	1fx9	1h8a	1ih0	1jwn	1luv	1pob	1ubv	2hts	5gst
1ac8	1bed	1cif	1dps	1ety	1fxf	1h8b	1ihf	1jy7	1luw	1poc	1ubw	2il6	5hbi
1aca	1beg	1cig	1dpu	1eum	1fxk	1h8e	1ihg	1jyg	1lva	1pod	1ubx	2ilk	5mba
1aci	1bej	1cih	1dpy	1eup	1fyh	1h8h	1iie	1jzk	1lvf	1poe	1uby	2int	5mbn
1acp	1bek	1cik	1dqe	1ev4	1fyj	1h96	1iio	1jzl	1lw1	1pog	1une	2irf	5p2p
1acv	1bem	1cio	1drg	1ev9	1fyk	1h97	1iip	1jzm	1lxx	1pon	1unk	2isd	5pal
1ad6	1beo	1cip	1ds4	1evs	1fyl	1h99	1ijl	1jzq	1ly8	1pou	1uor	2lbd	5prc
1adr	1bep	1cig	1dsb	1evy	1fym	1h9e	1ijw	1jzr	1ly9	1pp2	1utg	2lef	5r1r
1adt	1beq	1cjk	1dse	1evz	1fyz	1h9f	1ijx	1jzs	1lyc	1ppa	1utr	2lfb	5tnc
1adu	1bes	1cjt	1dsg	1ew6	1fz0	1h9g	1ijy	1k04	1lyk	1ppr	1uwo	2lh1	6atj
1adv	1bf5	1cju	1dsh	1ewa	1fz1	1h9t	1ijz	1k0a	1lyn	1pra	1uxc	2lh2	6ccp
1ae7	1bfa	1cju	1dso	1ewq	1fz2	1h9u	1ik0	1k0b	1lmk	1prg	1uxd	2lh3	6cox
1aeb	1bfm	1ck6	1dsp	1ewr	1fz3	1h9x	1ik3	1k0c	1mlp	1prh	1vap	2lh5	6cp4
1aed	1bfr	1ck7	1dt0	1exa	1fz4	1h9y	1ik7	1k0d	1mlq	1pru	1var	2lh6	6cpp
1aee	1bg0	1ckk	1dt1	1exe	1fz5	1h9z	1iku	1k0m	1mlr	1prv	1vas	2lh7	6cro
1aef	1bg1	1ckt	1dt6	1exi	1fz6	1ha2	1il0	1k0n	1m47	1psl	1vew	2lhb	6csc
1aeg	1bg5	1cl5	1dt7	1exj	1fz7	1ha7	1il6	1k0o	1m48	1psh	1vhh	2lig	6cts
1aeh	1bg6	1cl6	1dti	1exn	1fz8	1hab	1ile	1k0y	1m49	1psj	1vii	2lis	6gss
1aei	1bg7	1clb	1dtl	1exr	1fz9	1hac	1ilg	1klf	1m4a	1psr	1vin	2ljr	6gst
1aej	1bg8	1clc	1dtm	1exx	1fzi	1hak	1ilh	1klv	1m4b	1pth	1vip	2lyn	6gsu
1aek	1bgc	1cll	1du0	1exz	1fzi	1hb6	1ilk	1k2h	1m4c	1pue	1vlk	2mb5	6gsu
1aem	1bgd	1clm	1du2	1eyl	1fzp	1hb8	1imp	1k2o	1m60	1pva	1vls	2mbw	6gsu
1aen	1bge	1clp	1du6	1eyh	1g03	1hba	1imq	1k3g	1m62	1pvb	1vlt	2mga	6gsx
1afo	1bgf	1cls	1du7	1eys	1g08	1hbb	1in4	1k3h	1m6c	1pys	1vnc	2mgb	6gsy
1aep	1bgp	1cm1	1dug	1eyv	1g09	1hbg	1in5	1k40	1m6m	1qa6	1vnd	2mgc	6hbi
1aeq	1bgx	1cm4	1duk	1eyx	1g0a	1hbh	1im6	1k4w	1m6v	1qag	1vne	2mgd	6hbw
1aes	1bgy	1cma	1dul	1ez3	1g0b	1hbi	1in7	1k5h	1m7k	1qas	1vnf	2mge	6pax
1aet	1bh2	1cmb	1duo	1eza	1g0v	1hbk	1in8	1k62	1m9i	1qat	1vng	2mgf	6prc

laeu	lbh8	lmc	lduw	lezb	lg0w	lhbm	linz	lk6o	lma1	lqaz	lvnh	2mgg	6r1r
laev	lbh9	lcmf	ldux	lezc	lg1e	lhbn	lio3	lk6y	lmab	lqb2	lvni	2mgh	7atj
laew	lbhd	lcmg	ldv0	lezd	lg1u	lhbo	lio7	lk74	lmab	lqbj	lvns	2mgi	7ccp
laf7	lbi0	lcmj	ldv5	lezf	lg1x	lhbr	lio8	lk78	lmbc	lqbk	lvol	2mgj	7cei
laf8	lbi1	lcmn	ldvb	lezt	lg2h	lhbs	lio9	lk79	lmbd	lqbq	lvpi	2mgk	7cpp
lafh	lbi2	lcmp	ldve	lezv	lg2n	lhbu	liok	lk7a	lmbe	lqc7	lvpw	2mgl	7gss
lifr	lbi3	lcmq	ldvg	lezy	lg2y	lhbx	liop	lk7l	lmbf	lqck	lvre	2mgm	7hbi
lafv	lbja	lcmr	ldvh	lf0j	lg2z	lhch	liq0	lk7w	lmbg	lqdb	lvrf	2mhb	7ice
lagm	lbib	lcmu	ldvi	lf0m	lg33	lhch	liq1	lk7y	lmbh	lqdm	lvtn	2mhr	7icf
ligr	lbij	lcmw	ldvk	lf0y	lg39	lhci	liq3	lk83	lmbi	lqgh	lvwt	2mm1	7icg
lah7	lbin	lcmx	ldvo	lf12	lg3j	lhcm	liq5	lk8a	lmbj	lqgj	lvxa	2mta	7ich
lahd	lbip	lcmz	ldvp	lf14	lg3n	lhco	liqc	lk8k	lmbk	lqgk	lvxb	2mya	7ici
lahr	lbis	lcn4	ldvv	lf17	lg3s	lhcr	liqp	lk8u	lmbn	lqgp	lvxc	2myb	7icj
laie	lbj0	lcno	ldw0	lf1c	lg3t	lhcu	liqr	lk90	lmbp	lqgr	lvxd	2myc	7ick
laia	lbj5	lcnp	ldw1	lf1e	lg3w	lhde	liqu	lk93	lmbq	lqgt	lvxe	2myd	7icl
lail	lbj9	lcot	ldw2	lf1f	lg3y	lhda	liqv	lk94	lmc2	lqgw	lvxf	2mye	7icm
lain	lbja	lco6	ldw3	lf1m	lg4d	lhdb	lirb	lk95	lmc3	lqh4	lvxg	2mys	7icn
laip	lbje	lco8	ldw9	lf1s	lg4i	lhdd	lirc	lk96	lmd3	lqhb	lvxh	2new	7ico
lais	lbjf	lco9	ldwk	lf1z	lg4u	lhdi	lird	lk98	lmd4	lqhd	lvad	2ngl	7icp
laj3	lbjj	lcoh	ldwr	lf22	lg4w	lhdp	lirf	lk99	lmdn	lqi8	lvaw	2ngr	7icq
laj4	lbjy	lcoj	ldws	lf24	lg4y	lhds	lirg	lk9k	lmdv	lqi9	lvat	2not	7icr
laj5	lbjz	lcok	ldwt	lf25	lg50	lhel	lirj	lk9m	lmdy	lqiv	lvdc	2occ	7ics
laj8	lbk5	lcoo	ldxc	lf26	lg59	lhes	lirl	lk9p	lmfr	lqiw	lvew	2or1	7ict
laj9	lbk6	lcop	ldxd	lf2e	lg5n	lhec	lirm	lka3	lmgn	lqi2	lvew	2orc	7icu
lajg	lbk9	lcor	ldxr	lf36	lg5y	lhek	lirq	lka8	lmgt	lqja	lvja	2pac	7icv
lajh	lbkd	lcot	ldxs	lf3a	lg5z	lhfo	lirv	lkb0	lmhl	lqjb	lvjb	2pal	7pre
lak0	lbke	lcp0	ldxt	lf3b	lg6n	lhfs	lirw	lkbh	lmho	lqjd	lvjc	2pas	7r1r
lak4	lbkr	lcp4	ldxu	lf43	lg6w	lhfa	lirx	lkbu	lmhy	lqjt	lvjd	2paw	8cpp
lak8	lb10	lcp5	ldxv	lf44	lg6y	lhfe	lis2	lkcf	lmhz	lqk1	lvje	2pax	8gss
lakd	lb19	lcp6	ldxx	lf45	lg72	lhg2	lisa	lkcy	lmj2	lqkm	lvjf	2pdd	8ica
lakh	lb1q	lcpd	ldy7	lf4i	lg73	lhg4	lisb	lkdl	lmjk	lqkn	lvla	2pde	8icb
lakk	lbm0	lcp6	ldyv	lf4k	lg7d	lhg5	lisc	lkdh	lmjl	lqkr	lvql	2pgd	8icc
lal0	lbm9	lcpf	ldz4	lf4l	lg7o	lhga	lisn	lkdx	lmjm	lqks	lvwr	2pgh	8ice
lal6	lbmf	lcp7	ldz6	lf4m	lg8e	lhgb	lit1	lkea	lmjo	lqkt	lvws	2pgt	8icf
lala	lbmo	lcpo	ldz8	lf4n	lg8i	lhgc	lit2	lkee	lmjp	lqku	lvwt	2phi	8icg
lall	lbmt	lcpq	ldz9	lf4o	lg8q	lhgu	lit3	lkej	lmjq	lql3	lvwt	2pmt	8ich
lalu	lbm5	lcpv	le0e	lf4q	lg9l	lhhs	litf	lkey	lmks	lql4	lvbl	2pre	8ici
lalv	lbno	lcpt	le0h	lf4t	lg9v	lhhs	lith	lkf6	lmkt	lqla	lvch	2prg	8icj
lalw	lbnp	lcpw	le17	lf4u	lga3	lhhs	liti	lkfr	lmku	lqlb	lvxm	2psr	8ick
lam4	lbo9	lcq2	le1q	lf59	lgab	lhho	litk	lkfu	lmkv	lqlk	lvxn	2pua	8icl
lam9	lboa	lcqe	le1r	lf5n	lgah	lhig	litl	lkfx	lmkf	lqll	lvxo	2pub	8icm
lamz	lboc	lcqt	le28	lf5o	lgai	lhij	litm	lkfy	lmkg	lqls	lvxs	2puc	8icn
lan2	lbod	lcqx	le2a	lf5p	lgak	lhik	lity	lkgm	lmkh	lqm6	lvxi	2pud	8ico
lan4	lbof	lcrc	le2r	lf5q	lgax	lhio	liuf	lkgp	lmkj	lqmd	lvxl	2pue	8icp
lann	lbou	lcrd	le2x	lf5t	lgbu	lhiw	liv6	lkgp	lmkl	lqmg	lvxp	2puf	8icq
lanv	lbp2	lcrh	le39	lf5x	lgbv	lhj1	livh	lkgv	lmll	lqmq	lvxm	2pug	8icr
lao6	lbp3	lcri	le3g	lf63	lgc6	lhj3	liwi	lkho	lmml	lqmz	lvyc	2pvb	8ics
laoa	lbpo	lcrj	le3h	lf65	lgc7	lhj4	liwj	lkil	lmnn	lqn0	lvcb	2r1r	8ict
laof	lbpq	lcrk	le3k	lf66	lgcj	lhj5	liwk	lkib	lmlo	lqn1	lvcc	2r2f	8icu

laoi	lbpv	lcrx	le3m	lf68	lgey	lhjp	lj46	lkiy	lmq	lqn2	lycq	2r63	8icv
laok	lbpz	lcry	le3o	lf6f	lgcw	lhjt	lj47	lkiz	lmr	lqnf	lycr	2rig	8icw
laom	lbpz	lcs0	le3p	lf6h	lgdd	lhks	lj51	lkjs	lmls	lqnm	lyea	2rom	8icx
laon	lbq0	lcs4	le3y	lf6v	lgdi	lhkt	lj59	lkjy	lmlu	lqnn	lyeb	2sas	8icy
laoq	lbq7	lcs5	le41	lf70	lgdj	lhlb	lj5y	lkkc	lmmm	lqnt	lyfe	2sbl	8icz
laor	lbqv	lcs6	le52	lf71	lgdk	lhlm	lj75	lkkq	lmmo	lqo4	lyfm	2scp	9ant
laos	lbqz	lcs7	le5w	lf7c	lgdl	lhlo	lj77	lkkv	lmms	lqo8	lyge	2spc	9gss
laow	lbr0	lcsi	le68	lf7u	lgds	lhhr	lj78	lkkx	lmnl	lqoj	lyic	2spl	9ica
laoy	lbr1	lcsr	le6b	lf7v	lgdt	lhiv	lj7e	lklp	lmn2	lqp0	lyma	2spm	9icb
lap4	lbr4	lcss	le6i	lf7y	lgdv	lhmn	lj7k	lkmh	lmng	lqp4	lymb	2spn	9icc
lap5	lbrw	lcsu	le6j	lf80	lgdy	lhmn	lj7o	lkn5	lmnh	lqp7	lymc	2spo	9ice
lap6	lbs2	lcsv	le6v	lf8h	lgdz	lhmn	lj7p	lkn6	lmni	lqpa	lyog	2spz	9icf
lape	lbs3	lcsw	le6y	lf8n	lgeb	lhma	lj7q	lkn7	lmnj	lqpi	lyoh	2sqc	9icg
lapl	lbsh	lcsx	le78	lf93	lged	lhmc	lj7r	lkn8	lmnk	lqpm	lyoi	2stt	9ich
laqe	lbsd	lcta	le79	lf99	lgei	lhmd	lj7s	lko9	lmmn	lqpu	lyrn	2stw	9ici
laqt	lbsd	lctd	le7a	lf9d	lgej	lhme	lj7v	lkpm	lmmo	lqpw	lytc	2tct	9icj
laqv	lbsx	lctj	le7b	lf9g	lgek	lhmf	lj7w	lkps	lmnp	lqpz	lytf	2tdx	9ick
laqw	lbt1	lctr	le7c	lf9n	lgem	lhmo	lj7y	lkq7	lmnt	lqq3	lyve	2tmv	9icl
laqx	lbt2	lcts	le7d	lf9o	lgfi	lhmu	lj8m	lkq8	lmoa	lqq8	lyac	2tm4	9icm
lar4	lbt3	lcty	le7e	lfad	lgg2	lhmv	lj8y	lkqj	lmoa	lqq8	lyac	2tm4	9icm
lar5	lbt6	lctz	le7f	lfaf	lgg3	lhna	lj9i	lkqk	lmoc	lqqb	lyza	2trc	9ico
larp	lbt8	lcuk	le7g	lfaf	lggq	lhnr	lja8	lkqs	lmod	lqqe	lyzb	2trt	9icp
larq	lbu2	lcul	le7h	lfag	lggw	lhns	ljaf	lkqv	lmoh	lqqf	lyzc	2utg	9icq
larr	lbu3	lcun	le7i	lfah	lggz	lho8	ljb6	lkr7	lmpg	lqqi	lyzd	2vhh	9icr
laru	lbu7	lcvu	le7j	lfap	lgh0	lhon	ljba	lkre	lmr8	lqqt	lyze	2wrr	9ics
larv	lbug	lcww	le7l	lfaw	lgh1	lhps	ljbg	lkrf	lmro	lqqv	lyzf	2ycc	9ict
larw	lbun	lcx2	le7p	lfaz	lgh4	lhq1	ljc2	lkrj	lmrr	lqrj	lyzg	2ygs	9icu
larx	lbuv	lcx8	le7u	lfb2	lgh6	lhq3	ljci	lkrq	lmsd	lqrv	lyzh	351c	9icv
lary	lbuy	lcta	le7v	lfbo	lghc	lhqb	ljcq	lks9	lmse	lqry	lyzi	3atj	9icw
las0	lbv2	lcxc	le83	lfbq	lghq	lhqc	ljcr	lksm	lmsf	lqsa	lyzj	3bcc	9icx
las2	lbva	lcpv	le84	lfbs	lgia	lhqo	ljcs	lkso	lmti	lqsd	lyzk	3bct	9icy
las3	lbvb	lcxz	le85	lfbu	lgil	lhqv	ljdb	lkss	lmtj	lqsh	lyzl	3bp2	
lash	lbvc	lcy5	le86	lfbv	lgit	lhrr	ljdh	lksu	lmtk	lqsi	lyzm	3c2c	
latj	lbvd	lcy6	le8a	lfbw	lgiv	lhrc	ljdi	lktp	lmtv	lqsj	lyzn	3cao	
lau1	lbvm	lcyf	le8e	lfby	lgjm	lhrr	ljdl	lku1	lmud	lqsp	lyzo	3car	
lau7	lbvp	lcyi	le8w	lfce	lgjn	lhro	ljdo	lku2	lmun	lqte	lyzp	3ccp	
laur	lbvs	lcyj	le8x	lfcd	lgjq	lhrr	ljdr	lku3	lmux	lqtg	lyzq	3ccx	
laue	lbvy	lcyj	le8y	lfcd	lgjs	lhry	ljeb	lku7	lmuy	lqu2	lyzr	3cln	
laui	lbw5	lczz	le8z	lfce	lgjt	lhrr	ljei	lkv9	lmwc	lqu3	lyzs	3cp4	

APPENDIX L

Names of proteins in database for all- β secondary structural class

117e	1azd	1c7y	1ddt	1ehx	1frg	1gr7	1hvj	1jdc	1kj4	1mjy	1qkw	1toc	2pdz
12ca	1aze	1c89	1ddv	1ei5	1fm	1gri	1hvk	1je5	1kj7	1mjz	1qkx	1tol	2pec
12e8	1azm	1c8a	1ddw	1eif	1frq	1gs6	1hvl	1je6	1kj9	1mkw	1qkz	1tom	2pel
1a0i	1azr	1c8d	1ded	1eii	1frt	1gs7	1hvs	1jed	1kjf	1mkx	1qlf	1tox	2phl
1a0l	1azv	1c8h	1deh	1eje	1fsn	1gs8	1hvx	1jee	1kjh	1mlc	1qlj	1tsh	2plt
1a0q	1b09	1c8n	1df3	1ejn	1fsq	1gsg	1hw5	1jeg	1kji	1mma	1qlr	1tta	2pni
1a18	1b0g	1c8p	1dfb	1ejr	1fsr	1gsm	1hwg	1jfl	1kjq	1mmg	1qly	1ttb	2pre
1a1m	1b0i	1c8q	1dfc	1ejs	1ft9	1gso	1hwh	1jfs	1kjj	1mmn	1qm6	1ttc	2prd
1a1n	1b0o	1c9o	1dfu	1ejt	1ftp	1gtf	1hwn	1jfd	1kk0	1mnd	1qmd	1ttg	2psg
1a1o	1b1z	1c9s	1dfv	1eju	1fui	1gtm	1hwo	1jfh	1kk1	1mne	1qmu	1ttt	2rac
1a1q	1b23	1c9u	1dfx	1ejv	1fun	1gtr	1hwp	1jg9	1kk2	1mnu	1qnh	1ttt	2rcr
1a21	1b2p	1ca1	1dgl	1ekl	1fux	1gug	1hwr	1jgi	1kk4	1mpa	1qny	1tuc	2rma
1a2c	1b2r	1ca2	1dgl	1elp	1fv1	1gun	1hx0	1jgl	1kk5	1mph	1qok	1tud	2rmb
1a2d	1b2w	1ca3	1dgs	1emt	1fv9	1guo	1hx5	1jgu	1kk6	1mps	1qor	1tui	2rmp
1a2t	1b34	1ca8	1dgt	1enc	1fvc	1gus	1hx6	1jgv	1kl9	1msa	1qos	1tul	2rox
1a2u	1b35	1caj	1dhk	1enf	1fvd	1gut	1hxd	1jgw	1klf	1msi	1qot	1tvd	2rsp
1a2y	1b3j	1cak	1dif	1enq	1fwb	1gvh	1hxe	1jho	1kmh	1msj	1qou	1tyr	2rtb
1a30	1b44	1cal	1dil	1enr	1fwc	1gvi	1hxf	1jh2	1kn2	1msp	1qov	1uga	2rtc
1a33	1b4l	1cam	1dim	1ens	1fwd	1gvp	1hwx	1jh5	1knl	1mtr	1qp1	1ugb	2rtd
1a3b	1b4t	1can	1dit	1enx	1fwe	1gvt	1hxy	1jhd	1knm	1mua	1qp2	1uge	2rte
1a3t	1b55	1cao	1dj7	1eo5	1fwf	1gvu	1hyy	1jhf	1koi	1mup	1qp3	1ugd	2rtf
1a3u	1b5g	1cau	1djr	1eo7	1fwg	1gvv	1hzy	1jhh	1kop	1mvm	1qpk	1uge	2rtg
1a3v	1b6r	1cav	1dkg	1eo8	1fwh	1gvw	1hza	1jhj	1koq	1mvp	1qpp	1ugf	2rth
1a3w	1b6s	1caw	1dlc	1eoj	1fwi	1gvx	1hzb	1jhl	1kq1	1n15	1qq0	1ugg	2rti
1a3x	1b6u	1cax	1dle	1eol	1fwq	1gw0	1hze	1jil	1kqa	1n50	1qq4	1ukr	2rtj
1a3z	1b70	1cay	1dlm	1eou	1fwu	1gw5	1hzk	1jib	1kqf	1n90	1qqg	1ulo	2rtk
1a44	1b75	1caz	1dlq	1ep0	1fwv	1gx8	1hzi	1jim	1kqg	1nbm	1qqj	1uma	2rtl
1a45	1b7i	1cbi	1dm0	1ep1	1fwy	1gx9	1hzu	1jiw	1kqw	1nbv	1qqk	1uok	2rtm
1a46	1b7j	1cbj	1dm4	1ep2	1fwz	1gxa	1hzu	1jjg	1kqx	1ncg	1qq1	1upj	2rtn
1a4a	1b7t	1cbq	1dmh	1epf	1fx0	1gxd	1i04	1jjj	1kra	1nci	1qqs	1ure	2rtp
1a4b	1b7x	1cbr	1dmp	1epz	1fx5	1gy1	1i05	1jju	1krb	1nco	1qr1	1uri	2rtq
1a4c	1b7y	1cbv	1dmr	1eqd	1fx7	1gy9	1i06	1jks	1krc	1ndh	1qr4	1uvs	2rtr
1a4j	1b8a	1cc0	1dm0	1eqr	1fxj	1gya	1i0r	1jks	1krh	1ndr	1gre	1uvt	2sam
1a4k	1b8e	1ccs	1do6	1erb	1fy1	1gye	1i0s	1jls	1krr	1nds	1qrf	1uyu	2sba
1a4w	1b8q	1cct	1doa	1erx	1fy3	1gyg	1i16	1jld	1kru	1ndt	1qrg	1vac	2sbl
1a57	1b9n	1ccu	1doj	1es0	1fyc	1gz9	1i1b	1jlx	1krv	1neb	1qrk	1vad	2sfa
1a58	1b9z	1ccz	1dp0	1esf	1fyt	1gzc	1i1e	1jly	1ksa	1nfa	1qrl	1val	2sfp
1a5d	1ba7	1cd0	1dp5	1et5	1fyu	1gzp	1i1f	1jma	1ksg	1ngl	1qrm	1vam	2sim
1a5g	1ba8	1cd8	1dpj	1et6	1fzj	1h0g	1i1r	1jmq	1ksh	1ngp	1qrm	1vbs	2snm
1a5i	1ba9	1cd9	1dq0	1et8	1fzk	1h19	1i1y	1jmx	1ksi	1ngq	1qrp	1vbt	2sns
1a5k	1baf	1cdb	1dq1	1et9	1fzm	1h4g	1i2d	1jnz	1ksn	1nia	1qrs	1vde	2sob
1a5l	1bag	1cdc	1dq2	1eth	1g0d	1h4h	1i2h	1jns	1ksr	1nid	1qrt	1vfa	2sod
1a5m	1bai	1cdg	1dq3	1etj	1g0e	1h4i	1i3h	1jnh	1ktb	1nie	1qru	1vfb	2stv

1a5n	1bar	1cdh	1dq4	1etr	1g0f	1h4j	1i3p	1jnw	1ktc	1nif	1qrw	1vge	2taa
1a5o	1bas	1cdj	1dq5	1ets	1g0v	1h4n	1i3q	1jny	1ktd	1nin	1qrx	1vij	2tbs
1a61	1bb0	1cdo	1dq6	1ett	1g0y	1h4r	1i3s	1jof	1ktj	1nir	1qse	1vik	2tbv
1a62	1bb9	1cdu	1dqd	1etz	1g13	1h54	1i3v	1joi	1ktk	1niv	1qsf	1vjs	2tdt
1a63	1bbd	1cdy	1dqg	1eu1	1g1c	1h5b	1i4f	1joj	1kv3	1nld	1qtf	1vln	2tep
1a65	1bbj	1ce1	1dqi	1eu3	1g1d	1h5n	1i4g	1joo	1kv9	1nlr	1qtq	1vix	2thf
1a66	1bbs	1ce6	1dql	1eu4	1g1o	1h6q	1i4h	1jor	1kxf	1nls	1qtt	1vmo	2tmp
1a6c	1bbu	1ce7	1dr9	1eu5	1g1y	1h6x	1i4p	1jot	1kxh	1nmf	1qtu	1vqa	2tnf
1a6t	1bbw	1ces	1dro	1euo	1g2b	1h6y	1i4q	1jou	1kxl	1nn2	1qu2	1vqc	2tod
1a6z	1bbz	1cfl	1drt	1euq	1g2k	1h7y	1i4u	1jp5	1kxq	1nno	1qu3	1vqd	2trh
1a72	1bcp	1cf8	1dry	1euw	1g30	1h8d	1i4x	1jpc	1kxt	1noa	1qu4	1vqe	2tsa
1a78	1bcu	1cfm	1ds0	1euy	1g31	1h8e	1i53	1jpf	1kxv	1nol	1que	1vqg	2tsb
1a7b	1bcx	1cfn	1ds1	1evu	1g32	1h8h	1i5f	1jpg	1ky9	1npl	1quf	1vqi	2tss
1a7s	1bd0	1cfq	1ds8	1ewc	1g37	1h8i	1i5i	1jps	1kya	1np4	1qun	1vqj	2tun
1a8d	1bd2	1cfs	1dsf	1ewh	1g3g	1h8k	1i5l	1jq6	1kzq	1npl	1qup	1vrl	2upj
1a8g	1bd7	1cft	1dsw	1ewi	1g3p	1h8l	1i5p	1jqy	1l0p	1npo	1qur	1vsc	2vaa
1a8j	1bda	1cfv	1dtu	1eww	1g3t	1h8n	1i6h	1jqz	1l0w	1nps	1qwe	1vub	2vab
1a8k	1bd1	1cgp	1du5	1ewy	1g3w	1h8s	1i6t	1jr0	1l0z	1nqb	1rav	1wap	2vir
1a8p	1bdq	1cgs	1dun	1exd	1g3y	1h8t	1i7a	1jrg	1l1c	1nm	1ray	1wba	2vit
1a8v	1be3	1cgu	1dut	1exf	1g3z	1h8v	1i7r	1jrh	1l1g	1nro	1raz	1wbc	2viu
1a94	1beb	1cgv	1duz	1exg	1g43	1h91	1i7t	1jrj	1l1i	1nrr	1rbp	1wbf	2vub
1a9b	1bef	1cgw	1dv1	1exh	1g45	1h92	1i82	1js4	1l1n	1ntd	1red	1wgi	2wbc
1a9e	1beh	1cgx	1dv2	1exm	1g46	1h95	1i8a	1js9	1l3j	1nuc	1ree	1wgj	2web
1a9v	1bex	1cgy	1dv3	1exs	1g48	1h9j	1i8f	1jsd	1l3n	1nwo	1ref	1whi	2wec
1aa6	1bey	1chg	1dv6	1exu	1g4j	1h9k	1i8g	1jsg	1l3w	1nwp	1rei	1who	2wed
1aac	1bf8	1cho	1dv9	1ey0	1g4m	1h9m	1i8h	1jsh	1l4i	1nzz	1rin	1whp	2wpo
1aaj	1bfb	1chp	1dvf	1ey2	1g4o	1h9n	1i8l	1jsi	1l5b	1obp	1rip	1wio	2xat
1aan	1bfc	1chq	1dvq	1ey4	1g4r	1h9p	1i8m	1jsm	1l5c	1obw	1rlb	1wip	3cna
1ab0	1bfg	1ci0	1dvt	1ey6	1g51	1h9q	1i8u	1jso	1l5d	1oca	1rlp	1wiq	3cpu
1abi	1bfv	1ci3	1dvv	1ey7	1g52	1h9v	1i8z	1jsy	1l5g	1occ	1rlq	1wit	3cys
1abj	1bg1	1ci5	1dvx	1ey8	1g54	1h9x	1i90	1jt3	1l6f	1oco	1rmf	1wiu	3dmr
1abr	1bg9	1cic	1dvy	1ey9	1g5a	1h9y	1i91	1jt4	1l6g	1ocr	1rmg	1wkt	3dpa
1acb	1bgl	1cid	1dwb	1eya	1g5w	1ha3	1i96	1jt5	1l6p	1ocz	1rmh	1wpo	3eng
1acd	1bgm	1cim	1dwc	1eyb	1g6e	1hag	1i97	1jt7	1l7i	1odw	1rml	1wtl	3f58
1acx	1bgx	1ciw	1dwd	1eyc	1g6l	1hah	1i9i	1jt8	1l7j	1ody	1rtf	1xca	3fct
1ad0	1bhe	1ciw	1dwe	1eyg	1g72	1hai	1i9j	1jta	1l7k	1ohr	1run	1xna	3fiv
1ad8	1bhu	1cj5	1dxm	1eyz	1g76	1hap	1i9l	1jtc	1l9b	1okl	1rux	1xnb	3fru
1ad9	1bi0	1cjd	1dxx	1ez6	1g77	1hau	1i9m	1jto	1l9j	1okn	1rxf	1xnc	3gbq
1adc	1bi1	1cjp	1dxr	1ez8	1g78	1haw	1i9n	1jtt	1l9m	1ona	1rxg	1xnd	3gch
1adf	1bi2	1cjj	1dxw	1ezg	1g79	1hbl	1i9p	1ju4	1l9n	1opa	1rza	1xnt	3hat
1adg	1bia	1ck1	1dy4	1ezq	1g7h	1hb2	1i9w	1ju9	1lab	1opb	1rzb	1xso	3hla
1adl	1bib	1ckn	1dy7	1f0r	1g7i	1hb3	1ia5	1juf	1lac	1opg	1rzc	1xyn	3hmg
1adw	1bih	1cko	1dym	1f0s	1g7j	1hbp	1iai	1juh	1lay	1ops	1rzd	1xyp	3htc
1ae2	1bil	1clc	1dyn	1f13	1g7l	1hbt	1iam	1jui	1lcl	1osz	1rze	1yai	3hud
1ae3	1bim	1clh	1dyo	1f15	1g7m	1hbv	1iao	1jv0	1lde	1oxy	1sac	1yaz	3hvp
1ae5	1bit	1cmv	1dyq	1f2n	1g7p	1hc1	1iar	1jv2	1lde	1p01	1sat	1yda	3ifb
1ae6	1bj7	1cn1	1dyu	1f2q	1g7q	1hc2	1ib5	1jv4	1ldp	1p02	1sbd	1ydc	3kbp
1aey	1bj8	1cn4	1dyz	1f2w	1g7r	1hc3	1ib8	1jvk	1lds	1p04	1sbe	1ydd	3lpr
1af0	1bjm	1cnb	1dzj	1f31	1g7s	1hc4	1ibb	1jvl	1ldy	1p05	1sbf	1yec	3lve

1afc	1bjq	1cnf	1dzk	1f34	1g7t	1hc5	1ibd	1jvo	1lec	1p06	1sbg	1yed	3msi
1afe	1bk0	1cnh	1dzl	1f35	1g7y	1hc6	1ibq	1jw6	1led	1p09	1sbs	1yee	3ncm
1ag0	1bk1	1cni	1dzm	1f3d	1g84	1hca	1iby	1jw7	1leg	1p10	1scs	1yef	3nla
1ag4	1bkz	1cnk	1dzp	1f3r	1g85	1hcb	1ibz	1jxd	1len	1p11	1sda	1yeg	3nn9
1ag6	1bl9	1cnw	1dza	1f3t	1g86	1hcd	1ic0	1jxf	1lep	1p12	1sdy	1yeh	3npl
1agc	1blb	1cnx	1dza	1f3u	1g8f	1hce	1ic1	1jxg	1les	1p35	1se2	1yek	3nuc
1agd	1bld	1cny	1dzt	1f3w	1g8g	1hcm	1ic4	1jxj	1lfo	1pam	1se3	1yha	3ovw
1age	1bli	1cob	1e02	1f3x	1g8h	1hev	1ic5	1jxk	1lg5	1paz	1se4	1yhb	3paz
1agj	1blr	1con	1e06	1f41	1g8j	1hey	1icn	1jxp	1lge	1pcl	1sfp	1ykf	3pcy
1agn	1bm3	1cpm	1e0t	1f42	1g8k	1hcz	1id2	1jyn	1lge	1pdk	1sft	1yna	3pdz
1ahl	1bm5	1cpn	1e0u	1f49	1g8l	1hd3	1id5	1jyv	1lgn	1ped	1sgk	1ypp	3pep
1ah9	1bm7	1cpu	1e14	1f4a	1g8w	1hd5	1idb	1jyw	1lhc	1pfs	1shg	1yso	3phv
1ahj	1bmf	1cq1	1e15	1f4h	1g8z	1hd8	1idj	1jyx	1lhd	1pfz	1shs	1yst	3prc
1ahk	1bmj	1cq3	1e18	1f4x	1g95	1hdk	1idk	1jyy	1lhe	1pht	1sid	1ytg	3pro
1ahm	1bmm	1cq9	1e1a	1f4y	1g97	1hdm	1ie4	1jyz	1lhf	1pif	1sie	1yth	3psg
1aht	1bmn	1cq9	1e1o	1f53	1gaf	1hdt	1ie7	1jz0	1lhg	1pig	1sip	1yti	3rdn
1ahw	1bmz	1cq9	1e1q	1f56	1gan	1hdx	1iea	1jz2	1lic	1pkm	1siv	1ytj	3seb
1ai8	1bn1	1cq9	1e1t	1f5f	1gaq	1hea	1ieb	1jz5	1lie	1pkn	1sla	1yuh	3sil
1aid	1bn3	1cr9	1e22	1f5j	1gaw	1heb	1ieh	1jz6	1lif	1pks	1slb	1zap	3sod
1aig	1bn4	1cra	1e24	1f5k	1gax	1hec	1if4	1jz7	1lil	1pla	1sls	1zia	3tdt
1aij	1bnc	1crb	1e27	1f6f	1gbq	1hed	1if5	1jz8	1lij	1plb	1sle	1zib	3dlh
1aip	1bnm	1csp	1e28	1f6n	1gbr	1heg	1if6	1jze	1ljo	1plc	1slf	1znc	3tss
1air	1bnq	1ct1	1e2w	1f77	1gc6	1het	1if7	1jzg	1lk3	1plg	1slg	1zsa	3ubp
1aiw	1bnu	1ct8	1e2z	1f7a	1gc7	1heu	1if8	1jzh	1lka	1pls	1slt	1zsb	3ull
1aix	1bnv	1ctm	1e30	1f7r	1gcd	1hf2	1if9	1jzq	1lkb	1plu	1sm3	1zxq	3vub
1aiz	1bnw	1cto	1e3b	1f86	1gcs	1hf3	1ifc	1jzs	1lkt	1pmi	1sma	2cna	4cyh
1aj7	1bog	1cu4	1e3e	1f8n	1gcy	1hg7	1ig0	1k0h	1lla	1pmp	1smd	2cnd	4dmr
1ajv	1boq	1cuk	1e3f	1f9k	1gd7	1hgd	1ig3	1k0r	1llr	1pmr	1sme	2cpl	4dpv
1akj	1bov	1cuo	1e3i	1fa4	1gfc	1hgf	1igc	1k0s	1lmw	1pmv	1smv	2cpu	4eng
1akl	1boy	1cur	1e3j	1fai	1gfd	1hgg	1igj	1k0x	1lnh	1pnc	1smr	2ctv	4est
1akp	1bpo	1cva	1e3l	1fat	1gg2	1hgh	1igm	1k12	1lnu	1pnf	1smv	2cxg	4fab
1alb	1bq5	1cvb	1e3x	1fay	1ggc	1hgi	1igo	1k1d	1lo0	1pnj	1smc	2cyh	4fiv
1am2	1bqh	1cvc	1e3z	1fb7	1ggi	1hgi	1igp	1k1z	1lo2	1ppb	1snd	2dbl	4gal
1am5	1bql	1cve	1e40	1fbi	1gg1	1hh9	1ihm	1k28	1lo3	1ppc	1snm	2dij	4gch
1am6	1bqr	1cvf	1e43	1fc5	1gg2	1hhg	1iht	1k32	1loa	1ppe	1sno	2dmr	4gcr
1ame	1bqu	1cvn	1e44	1fcg	1ggy	1hhh	1iip	1k3r	1lob	1pph	1snq	2dtr	4hmg
1amm	1br1	1cwr	1e4h	1fd7	1gh7	1hhi	1iiq	1k3v	1lod	1ppi	1sos	2dyn	4htc
1amy	1br2	1cwa	1e4j	1fdi	1gha	1hhj	1iiu	1k42	1loe	1ppk	1sox	2eif	4hvp
1an8	1br4	1cwb	1e4w	1fdo	1ghb	1hhk	1iiy	1k45	1lof	1ppl	1spd	2eng	4ilb
1ao7	1br9	1cwc	1e4x	1fdr	1ghf	1hhp	1ij9	1k4c	1log	1ppm	1spu	2epn	4kbp
1aof	1bre	1cwf	1e5a	1fe3	1ghj	1hi6	1ija	1k4d	1lom	1prc	1srd	2ezn	4msi
1aoh	1brp	1cwh	1e5h	1fel	1ghk	1hib	1ijs	1k5c	1lop	1pri	1srg	2f19	4nn9
1aoj	1brq	1cwi	1e5i	1fen	1gho	1hih	1ik3	1k5d	1lpa	1prm	1srh	2f3g	4ovw
1aon	1bsh	1cwj	1e5o	1feu	1ghw	1hil	1ikf	1k5g	1lpb	1pro	1sri	2f58	4paz
1aoz	1bsi	1cwl	1e5p	1ff5	1ghx	1him	1ikj	1k5r	1lqs	1prp	1srj	2fb4	4pcy
1ap2	1bsn	1cwm	1e5r	1fft	1gi7	1hin	1ikp	1k5u	1lrh	1prs	1srl	2fbj	4pep
1apn	1bt7	1cx1	1e5s	1ffy	1gi8	1hiv	1il1	1k5v	1ls5	1prt	1srn	2fcb	4prc
1apt	1bth	1cxf	1e5t	1fg9	1gi9	1hix	1il2	1k6c	1lt3	1psf	1srp	2fgf	4pro
1apu	1btk	1cxi	1e5u	1fga	1gig	1hj3	1ils	1k6l	1lt5	1psk	1sta	2fiv	4sgb

lapv	1bu8	1cxk	1e5v	1fgb	1gj4	1hj4	1inc	1k6p	1lt6	1psn	1stb	2fmb	4tf4
lapw	1bv3	1cxl	1e5w	1fgl	1gj5	1hj5	1ine	1k6q	1lte	1pso	1sth	2fnt	4thn
laq8	1bv7	1cyg	1e5y	1fgn	1gj7	1hjd	1ino	1k6t	1ltr	1pss	1stn	2fmb	4tss
laqb	1bv9	1cz4	1e5z	1fgo	1gj8	1hjf	1inq	1k6v	1lu1	1pto	1stp	2g3p	4ubp
laqf	1bve	1cz5	1e60	1fgp	1gj9	1hjj	1inv	1k6w	1lu2	1pts	1sty	2gal	4ull
laqh	1bvg	1czf	1e61	1fgq	1gja	1hjp	1inw	1k70	1luz	1pvx	1suh	2gbq	4vub
laqk	1bvk	1czs	1e62	1fgr	1gjb	1hkl	1inx	1k83	1lvk	1pwt	1svb	2gch	4wbc
laqm	1bvs	1czt	1e63	1fh2	1gjc	1hlc	1iny	1k8d	1lvo	1pyp	1svc	2gfb	5cna
laqt	1bvz	1czv	1e64	1fhg	1gjq	1hld	1io6	1k8g	1lxa	1pys	1sxa	2gmt	5est
lar1	1bw3	1d0x	1e65	1fhm	1gjr	1hlt	1ioa	1k8h	1lyl	1pza	1sxc	2gn5	5fiv
larb	1bwa	1d0y	1e6a	1fho	1gju	1hm0	1ipk	1k8i	1m1x	1pzb	1sxn	2gva	5gal
larc	1bw6	1d1a	1e6d	1fi2	1gfw	1hm2	1ips	1k8k	1m3x	1pzc	1sxs	2gvb	5hmg
lark	1bwu	1d1b	1e6g	1fie	1gkb	1hm3	1ipw	1k8m	1m43	1qbi	1sxt	2h4n	5ilb
las6	1bww	1d1c	1e6h	1fig	1gkh	1hm9	1iqw	1k8o	1m4v	1qbl	1sxz	2hft	5lpr
las8	1bwy	1d1s	1e6n	1fiw	1gko	1hms	1ira	1k9q	1m6h	1qbm	1syb	2hgt	5lve
laso	1bx0	1d1t	1e6o	1fiz	1gkp	1hmu	1irp	1k9r	1m6p	1qbr	1sys	2hla	5nn9
lasp	1bx1	1d2b	1e6p	1fjg	1gkq	1hmv	1itb	1kaa	1m6w	1qbu	1syd	2hmb	5paz
lasq	1bx2	1d2u	1e6r	1fjj	1gkr	1hnl	1ivp	1kab	1m9w	1qbv	1sye	2hmi	5pcy
lat3	1bxa	1d3b	1e6z	1fjr	1gkt	1hne	1ivq	1kap	1ma0	1qc6	1syf	2hpe	5pep
laun	1bxh	1d3c	1e79	1fli	1gl0	1hnf	1ivu	1kaw	1mac	1qcs	1syg	2hpf	5trr
laui	1bxo	1d3l	1e7n	1flc	1gl1	1hng	1ivw	1kb3	1mae	1qcx	1tbq	2hsp	5upj
lauy	1bxq	1d3s	1e80	1flg	1gla	1hnw	1ivx	1kb5	1maf	1qdm	1tbr	2hvp	6fiv
lav4	1bxt	1d4k	1e8k	1flh	1glb	1hnx	1iw2	1kbq	1mai	1qdn	1tbz	2ilb	6ilb
lavb	1bxz	1d4l	1e8m	1flj	1glc	1hny	1iyu	1kbb	1maj	1qdo	1tcm	2iad	6lpr
lave	1byh	1d4s	1e8n	1flm	1gld	1hnz	1iyv	1kbv	1mak	1qex	1tcw	2ifb	6msi
lavf	1byp	1d4y	1e9g	1flmb	1gle	1hoe	1j4t	1kbw	1mam	1qez	1tcx	2iff	6nn9
lavg	1bz7	1d5b	1e9q	1fimt	1gm6	1hos	1j4u	1kcd	1mcb	1qf3	1tdt	2ig2	6paz
lavk	1bz8	1d5i	1e9z	1fmv	1gmc	1hpc	1j59	1kcl	1mcc	1qfj	1teh	2igf	6pcy
lavl	1bz9	1d5m	1ea9	1fmw	1gmd	1hpl	1j5c	1kcr	1mcd	1qfm	1ten	2imm	6upj
lavn	1bzd	1d5x	1eag	1fmz	1gme	1hpo	1j5d	1kcs	1mce	1qfo	1tet	2imn	7fab
lavu	1bzm	1d6e	1eal	1fn0	1gmh	1hps	1j5e	1kcu	1mcf	1qfp	1tf4	2irt	7gch
lawf	1bzo	1d6u	1eap	1fn4	1gmi	1hpx	1j5o	1kcv	1mch	1qfs	1tfh	2jhb	7kme
lawh	1bzw	1d6y	1ear	1fnb	1gmu	1hqh	1j71	1kda	1mci	1qft	1tft	2jia	7lpr
lawq	1c01	1d6z	1eas	1fnc	1gmw	1hqp	1j7v	1kdb	1mcj	1qfu	1tha	2lpr	7msi
laws	1c0a	1d7b	1eat	1fnd	1gmw	1hr0	1j86	1kde	1mck	1qfv	1thc	2ltm	7paz
lawu	1c0w	1d7c	1eau	1fne	1gnh	1hrn	1j87	1kdf	1mcn	1qg0	1thi	2lve	7prc
laww	1c1f	1d7d	1ecy	1fng	1gnm	1hrt	1j89	1kdi	1mco	1qgl	1thj	2mad	7taa
lawx	1c1l	1d7k	1ecz	1fnh	1gnn	1hs6	1j8r	1kdj	1mcp	1qb6	1thp	2mcm	7upj
lax0	1c1u	1d7q	1edh	1fnl	1gno	1hsg	1j8s	1kef	1mcq	1qh7	1thu	2mcp	8fab
lax1	1c1v	1d8l	1ee2	1fnm	1gnz	1hsh	1j9q	1kel	1mcs	1qhd	1thw	2mfn	8gch
lax2	1c1w	1d9i	1ee6	1fnp	1go3	1hsi	1j9r	1kem	1mcw	1qho	1tie	2mib	8hvp
lax3	1c3k	1dab	1eef	1fnq	1gof	1hso	1j9s	1keo	1mdt	1qhp	1tit	2mll	8ilb
laxa	1c3m	1dba	1eer	1fny	1goh	1hsq	1j9t	1kev	1mel	1qhq	1tf2	2mpa	8lpr
laxe	1c3n	1dbb	1ef2	1fnz	1goi	1hsz	1j9w	1kgq	1mer	1qhu	1tlk	2msi	8msi
laxg	1c57	1dbg	1efc	1fon	1got	1ht0	1jab	1kgt	1mes	1qi3	1tlm	2msj	8paz
laxi	1c5d	1dbj	1efg	1fpc	1goz	1htf	1jac	1kgu	1met	1qi5	1tmb	2msp	8prk
laxj	1c5f	1dbk	1efi	1fpv	1gp2	1htg	1jae	1kgw	1meu	1qil	1tmc	2nap	9hvp
laxk	1c5h	1dbm	1efq	1fq4	1gp4	1htp	1jb0	1kgx	1mf2	1qiq	1tmo	2ncm	9ilb
laxs	1c5i	1dbn	1efr	1fq5	1gp5	1htr	1jb3	1kho	1mfh	1qjl	1tnf	2nlr	9lpr

laxy	lc5k	ldbo	left	lfq6	lgp6	lhug	ljb7	lchr	lmfm	lqj6	ltmh	2npl	9msi
laxz	lc5l	ldca	lefw	lfq7	lmpi	lhuj	ljb9	lkip	lmfn	lqj7	ltmi	2nrd	9pcy
lay6	lc5n	ldcb	lefx	lfq8	lgpo	lhut	ljbc	lkiq	lmhc	lqje	ltmj	2nuc	
layo	lc5o	ldce	legl	lfqm	lgpr	lhv0	ljc7	lkir	lmhe	lqjf	ltmk	2nul	
layr	lc5w	ldcl	legn	lfqr	lgq8	lhv1	ljch	lkiu	lmik	lqjo	ltml	2occ	
laz5	lc5x	ldcs	legx	lfqx	lgqg	lhva	ljcv	lkix	lmjc	lqjz	ltmm	2ovw	
lazb	lc5y	lddh	leh9	lfr3	lgqh	lhvh	ljd0	lkjl	lmjw	lqkq	ltmn	2oxi	
lazc	lc5z	lddl	leha	lfr7	lgqw	lhvi	ljda	lkj3	lmjx	lqks	ltmr	2pcy	

APPENDIX M

Names of proteins in database for α/β secondary structural class

10gs	1b27	1cb6	1dge	1ea7	1f9c	1gga	1hno	1j5o	1kah	1lld	1qil	1trb	2pth
10mh	1b2l	1cb7	1dgk	1ea9	1fa2	1gge	1hnu	1j5p	1kak	1llo	1qi2	1trd	2pua
121p	1b2r	1cbg	1dgm	1eam	1faa	1ggf	1hnv	1j5s	1kam	1lnq	1qi3	1tre	2pub
13gs	1b2s	1cbo	1dgt	1eav	1fba	1ggg	1hnw	1j5t	1kao	1lny	1qi5	1trh	2puc
13pk	1b2u	1cbx	1dhf	1eb9	1fbt	1ggh	1hnx	1j5x	1kaq	1loo	1qi6	1tri	2pud
16gs	1b30	1cc2	1dhi	1ebb	1fc0	1ggj	1hny	1j7k	1KAS	1loq	1qid	1trk	2pue
16pk	1b31	1ccw	1dhj	1ebh	1fcb	1ggk	1hnz	1j8d	1kav	1lor	1qie	1trs	2pug
17gs	1b32	1cd2	1dhk	1ebi	1fcd	1ggm	1ho3	1j8v	1kax	1los	1qif	1tru	2pvi
18gs	1b37	1cd5	1dhp	1ec7	1fck	1ggv	1hoo	1j8y	1kay	1lot	1qig	1trv	2q21
19gs	1b3b	1cdd	1dhr	1ec8	1fcq	1gho	1hor	1j93	1kaz	1lpa	1qih	1trw	2qr2
1a04	1b3e	1cde	1dht	1ec9	1fcu	1ghr	1hox	1j96	1kb3	1lpb	1qii	1tsi	2reb
1a0d	1b3f	1cdg	1di0	1ecb	1fcv	1ghs	1hoz	1j9b	1kbb	1lpf	1qij	1ttp	2req
1a0f	1b3g	1cdo	1di6	1ecc	1fdi	1ght	1hp0	1j9g	1kbi	1lpn	1qik	1ttq	2rla
1a16	1b3h	1cdz	1di7	1ece	1fdj	1gia	1hp5	1j9j	1kbb	1lpo	1qim	1tui	2rn2
1a19	1b3l	1ce2	1dia	1ecf	1fdo	1gil	1hpb	1j9k	1kbl	1lpp	1qj3	1tux	2rsl
1a1s	1b3n	1ce8	1dib	1ecg	1fdr	1gim	1hpl	1j9l	1kbo	1lps	1qjd	1tvr	2rus
1a23	1b3o	1cec	1did	1ecj	1fds	1gjr	1hpm	1jai	1kbq	1lqu	1qjm	1txx	2rve
1a24	1b3r	1cee	1die	1ecp	1fdt	1gju	1hpz	1jak	1kbz	1lrj	1qk3	1tya	2sbt
1a27	1b3s	1cen	1dih	1ecq	1fdu	1gk8	1hqa	1jaw	1kc0	1lrk	1qk4	1tyb	2scu
1a2b	1b3v	1ceo	1din	1edb	1fdv	1gkp	1hqc	1jax	1kc1	1lss	1qk5	1tyd	2shk
1a2j	1b3w	1cev	1dir	1edd	1fdw	1gkq	1hqd	1jay	1kc3	1lsu	1qk8	1typ	2skc
1a2k	1b3x	1cey	1dis	1ede	1fdy	1gkr	1hqe	1jbq	1kc6	1lvh	1qkb	1tyt	2skd
1a2l	1b42	1cf2	1diu	1edg	1fdz	1gku	1hqf	1jbv	1kc7	1lvk	1qki	1uaa	2ske
1a2m	1b43	1cf3	1dj2	1edt	1fea	1gky	1hqg	1jbw	1kcf	1lvi	1qkj	1uag	2sni
1a2o	1b48	1cf9	1dj3	1edz	1feb	1gl3	1hqk	1jce	1kek	1lwi	1qkk	1ubn	2stl
1a2q	1b4d	1cfj	1dj9	1ee0	1fec	1gla	1hqs	1jcf	1kcl	1lwn	1qla	1ubp	2taa
1a2z	1b4e	1cfr	1dje	1ee2	1feh	1glb	1hqt	1jcy	1ke3	1lwo	1qlb	1ubs	2tec
1a3c	1b4k	1cfz	1djm	1ec3	1ff9	1glc	1hqx	1jcm	1kd0	1lx7	1qlh	1ucw	2thi
1a3w	1b4q	1cg0	1djm	1ee9	1ffa	1gld	1hr0	1jcx	1kdo	1lxt	1qlj	1uda	2tir
1a3x	1b4u	1cg3	1djg	1eea	1ffb	1gle	1hrd	1jcy	1kdp	1ly1	1qlv	1udc	2tmd
1a40	1b4v	1cg4	1dka	1eeh	1ffc	1glg	1hse	1jda	1kdr	1ly3	1qlw	1udg	2tmg
1a47	1b4z	1cg6	1dkg	1eej	1ffd	1gli	1hsj	1jdc	1kdt	1ly4	1qm5	1udh	2tmy
1a4i	1b51	1cgk	1dkl	1eem	1ffe	1gln	1hsl	1jdd	1kdv	1m0z	1qmh	1udi	2tod
1a4s	1b52	1cgt	1dkm	1een	1ffh	1glp	1ht0	1jdf	1ke1	1mlb	1qmi	1udr	2tpl
1a4u	1b54	1cgv	1dko	1eeo	1fg3	1glq	1ht1	1jdi	1ke2	1mlm	1qml	1uke	2tpr
1a4x	1b57	1cgw	1dkq	1eep	1fg5	1gm5	1ht2	1jdi	1kep	1mlx	1qmu	1uky	2tpt
1a4y	1b58	1cgx	1dkr	1ees	1fg7	1gmx	1ht3	1jdn	1ket	1m6h	1qnf	1ulb	2trs
1a54	1b5h	1cgy	1dku	1eet	1fgs	1gn0	1hti	1jdp	1keu	1m6w	1qnl	1uok	2trx
1a55	1b5j	1cgz	1dl5	1eex	1fh7	1gn8	1hti	1jds	1kev	1ma0	1qno	1ura	2ts1
1a5b	1b5q	1ch8	1dla	1ef3	1fh8	1gnd	1htt	1jdu	1kez	1mab	1qnp	1urb	2tys
1a5c	1b6g	1chm	1dli	1ef8	1fh9	1gne	1htw	1jdv	1kf0	1mah	1qnq	1uro	2uag

1a5l	1b6r	1chn	1dlo	1ef9	1fhd	1gnr	1hu4	1jdy	1kf6	1MAR	1qnr	1urp	2ubp
1a5m	1b6s	1chr	1dlr	1efa	1fhe	1gnw	1huv	1jdz	1kfl	1mas	1qns	1uug	2udp
1a5n	1b73	1chu	1dmb	1efg	1fhu	1goa	1hux	1je0	1kfq	1md3	1qnt	1uwb	2ukd
1a5o	1b78	1chw	1dmu	1efk	1fhv	1gob	1hv9	1je1	1kfy	1md4	1qnv	1v39	2uug
1a5s	1b7b	1cil	1dnc	1efl	1fiu	1goc	1hvx	1jec	1kgw	1mdj	1qo0	1vdr	2vp3
1a5t	1b7g	1cib	1dnp	1efp	1fiy	1goi	1hw6	1jed	1kgx	1mdj	1qo2	1vfd	2xim
1a69	1b7h	1cij	1do8	1efr	1fjg	1goj	1hwx	1jee	1kh1	1mdk	1qo5	1vhr	2xin
1a6d	1b7t	1cip	1doa	1eft	1fjh	1gok	1hy3	1jeh	1kh4	1mdl	1qo8	1vjs	2xis
1a6e	1b7u	1ciu	1dob	1efz	1fk8	1gom	1hyb	1jej	1kh7	1mdr	1qo9	1vkl	2ypi
1a72	1b7z	1civ	1doc	1eg2	1fk9	1goq	1hye	1jet	1khh	1mee	1qon	1vlz	2ypn
1a76	1b8g	1cj2	1dod	1eg5	1fko	1gor	1hyg	1jeu	1khe	1mek	1qop	1vot	3cpa
1a77	1b8j	1cj3	1doh	1eg7	1fkp	1gos	1hyq	1jev	1khf	1mg1	1qoq	1vp3	3cpu
1a7a	1b8n	1cj4	1dor	1ega	1f9	1gow	1hys	1jez	1khg	1mgt	1qox	1vp9	3csu
1a7j	1b8o	1cjb	1dos	1egh	1fla	1gox	1hyu	1jff6	1khj	1mh1	1qoz	1vpe	3dbv
1a7k	1b8p	1cjc	1dot	1ego	1flc	1gp1	1hzj	1jff7	1khh	1mht	1qp0	1vpt	3dff
1a7l	1b8u	1cjm	1doz	1egq	1fln	1gp2	1hzm	1jff8	1khl	1mio	1qp4	1vpw	3dhe
1a7u	1b8v	1cjj	1dp0	1egr	1flv	1gpa	1hzz	1jff9	1khn	1mjh	1qp7	1vrt	3dmr
1a80	1b8x	1ckq	1dp2	1egv	1flz	1gpb	1hzz	1jffh	1kie	1mla	1qp8	1vsb	3dre
1a82	1b90	1cl0	1dp4	1egz	1fmc	1gpe	1i01	1jffl	1kif	1mld	1qpb	1vxo	3eca
1a88	1b94	1cl1	1dpe	1eh3	1fmf	1gpu	1i0i	1jffr	1kj8	1mma	1qpg	1wab	3enl
1a8f	1b95	1cl2	1dpg	1eh6	1fml	1gpy	1i0l	1jffs	1kj9	1mmd	1qpk	1waj	3eug
1a8h	1b96	1cl8	1dpi	1eh7	1fmt	1gq7	1i0z	1jfft	1kji	1mmg	1qpr	1wdn	3fua
1a8i	1b97	1cle	1dpp	1eh8	1fmv	1gq9	1i10	1jffv	1kjj	1mnd	1qpz	1wet	3fyg
1a8l	1b9i	1clk	1dpz	1eh9	1fnb	1gqc	1i14	1jffx	1kjin	1mne	1qq2	1wkf	3gar
1a8p	1b9z	1clq	1dqn	1eha	1fnd	1gqt	1i1h	1jffl	1kjq	1mns	1qq5	1wsa	3gbp
1a8q	1ba0	1clu	1dqp	1ehc	1fnn	1gr0	1i1n	1jff2	1kix	1mpb	1qq6	1xaa	3gpb
1a8s	1ba1	1clx	1dqr	1ehy	1fno	1gr2	1i1o	1jff4	1kk0	1mpc	1qq7	1xab	3gpd
1a8u	1ba2	1cm5	1dqu	1ein	1foh	1gr8	1i1q	1jff6	1kk1	1mra	1qq9	1xac	3grs
1a96	1bag	1cm7	1dqw	1eiw	1fov	1gr9	1i2n	1jff7	1kk2	1mrp	1qqa	1xad	3grx
1a98	1bam	1cmw	1dqx	1ejb	1fp2	1grb	1i2o	1jff8	1kk3	1mss	1qqb	1xas	3gsb
1a9o	1bap	1cne	1dqy	1eji	1fp6	1gre	1i2q	1jffi	1kkb	1muc	1qqc	1xel	3gss
1a9p	1bav	1cnf	1dr0	1ejj	1fp7	1gre	1i2r	1jfft	1kkf	1myr	1qr2	1xic	3gst
1a9q	1bay	1cnm	1dr1	1ejr	1fpm	1grf	1i2z	1jff8	1kkj	1n2c	1qr6	1xid	3had
1a9r	1bd0	1cnv	1dr2	1ejs	1fpq	1grg	1i30	1jff9	1kko	1nah	1qr7	1xie	3hdd
1a9s	1bd4	1cnz	1dr3	1ejt	1fpz	1grk	1i32	1jffa	1kkz	1nai	1qra	1xif	3hud
1a9t	1bdb	1cow	1dr6	1ejv	1fpz	1grl	1i33	1jffb	1kku	1nar	1qrd	1xig	3hvt
1a9y	1bdh	1coy	1dr8	1ek0	1fq0	1gro	1i36	1jffd	1kl1	1nat	1qrr	1xii	3icd
1aa1	1bdi	1cp2	1dra	1ek1	1fq1	1grp	1i39	1jffh	1kl2	1nba	1qrs	1xij	3lad
1aa8	1be0	1cp3	1drb	1ek2	1fqb	1grq	1i3a	1jffp	1kl7	1nbe	1qrt	1xin	3lip
1aa9	1be1	1cp6	1drf	1ek4	1fqc	1grt	1i3c	1jffq	1klm	1nda	1qru	1xis	3mag
1aax	1be8	1cp7	1drh	1ek5	1fqd	1gs5	1i3l	1jffr	1kln	1ndh	1qs0	1xjo	3mbp
1aaz	1bed	1cpb	1drj	1ekk	1fqe	1gsa	1i3m	1jffu	1kly	1nel	1qsg	1xkj	3mct
1ab7	1bee	1cps	1drk	1eko	1fqo	1gsb	1i3n	1jffv	1klz	1nfp	1qti	1xla	3min
1aba	1bez	1cpu	1dru	1ekq	1fqw	1gsc	1i3q	1jffx	1kml	1ng1	1qtr	1xlb	3ngl
1abb	1bf2	1cpy	1ds9	1ekx	1fr9	1gsd	1i41	1jffy	1km2	1nga	1qtw	1xlc	3nll
1abe	1bf6	1cog	1dsb	1el3	1frb	1gse	1i45	1jff0	1km3	1ngb	1qu2	1xld	3pfk
1abf	1bfd	1cqh	1dsn	1el5	1frq	1gsf	1i48	1jff1	1km4	1ngc	1qu3	1xle	3pfl
1abn	1bfk	1cqi	1dt3	1el7	1frw	1gsg	1i4d	1jff2	1km5	1ngd	1qu4	1xlg	3pgk
1ac1	1bfm	1cqj	1dt5	1el8	1frz	1gsh	1i4l	1jffb	1kmi	1nge	1que	1xlh	3pgt

lac5	1bg2	1cqz	1dte	1elj	1fs5	1gsi	1i4t	1jii	1kmk	1ngg	1quf	1xli	3pjr
lacj	1bg4	1crl	1dtg	1ell	1fs6	1gsn	1i4w	1jjj	1kmm	1ngh	1quj	1xlj	3pmg
laci	1bg6	1crp	1dtq	1elm	1fsp	1gso	1i50	1jil	1kmm	1ngi	1qul	1xik	3pnp
lacm	1bg9	1crq	1dtt	1elo	1fss	1gsq	1i57	1jiq	1kmv	1ngj	1qum	1xll	3prk
laco	1bgb	1crw	1dtu	1elq	1fsz	1gss	1i5e	1jiu	1knp	1ngs	1quw	1xlm	3pvi
lacr	1bgg	1csl	1dtw	1elx	1ft7	1gsu	1i5g	1jiv	1knq	1nhq	1ral	1xol	3ran
lad1	1bgj	1ct5	1dty	1ely	1ftg	1gsy	1i5o	1jix	1knr	1nhr	1ra2	1xoa	3rap
lad3	1bgl	1ct9	1dug	1elz	1ftj	1gt7	1i5s	1jke	1ko1	1nhs	1ra3	1xob	3req
lad4	1bgm	1ctt	1duv	1em6	1ftk	1gt8	1i60	1jke	1ko4	1nht	1ra8	1xva	3rla
ladb	1bgn	1ctu	1dv2	1emd	1ftl	1gta	1i6b	1jkh	1ko5	1NIS	1ra9	1xzb	3rub
ladd	1bgs	1cua	1dv7	1eno	1ftm	1gtb	1i6i	1jki	1ko7	1nip	1raa	1xzc	3sic
ladf	1bgt	1cuc	1dvr	1enp	1ftn	1gte	1i6l	1jkj	1ko8	1nit	1rab	1xze	3tgl
ladg	1bgu	1cud	1dwo	1enu	1fto	1gth	1i6o	1jcx	1kof	1nks	1rac	1xzf	3tim
ladi	1bgv	1cug	1dwp	1eny	1ftq	1gti	1i6p	1jll	1koj	1nok	1rad	1xzg	3tmk
ladj	1bgx	1cuh	1dwq	1enz	1ftw	1gtu	1i6q	1jl2	1kp0	1noy	1raf	1xzh	3tmy
lado	1bhm	1cui	1dx4	1eo3	1fty	1gtv	1i6w	1jl3	1kp2	1noz	1rag	1xzi	3tsl
lads	1bhs	1cuj	1dx6	1eo4	1fu4	1gtx	1i74	1jl5	1kp3	1npx	1rah	1xzz	3uag
lady	1bhww	1cuu	1dx9	1eo5	1fu8	1gu0	1i75	1jl8	1kp9	1nsa	1rai	1xzk	3ubp
lae1	1bi5	1cuv	1dxe	1eo7	1fue	1gu1	1i7s	1jla	1kpg	1nsj	1ral	1xzm	3ukd
lae4	1bi9	1cuw	1dx6	1eog	1fui	1guh	1i80	1jlb	1kph	1nsy	1rb2	1yaa	3vsb
laf2	1bif	1cux	1dxh	1eoh	1fuu	1guk	1i86	1jlc	1kqg	1ntr	1rb3	1yac	3xin
laf4	1biy	1cuz	1dxi	1eom	1fuy	1gul	1i88	1jle	1kqp	1nul	1rba	1yag	3xis
lafs	1bjn	1cv2	1dxl	1eon	1fvj	1gum	1i89	1jlf	1kqy	1nzy	1rbo	1yas	3yas
lag8	1bjo	1cvi	1dxy	1eoo	1fvk	1guz	1i8b	1jlg	1kqz	1oaa	1rbr	1ybv	3ypi
lag9	1bjr	1cvl	1dyi	1eop	1fvx	1gv0	1i8j	1jll	1kr0	1oas	1rbs	1ydv	4dbv
lagn	1bjw	1cw1	1dyj	1lep1	1fw0	1gv1	1i8t	1jlk	1kr5	1oat	1rbu	1ygp	4dcg
lagn	1bka	1cw2	1dyr	1lep2	1fw1	1gve	1i94	1jll	1kra	1obr	1rbv	1yja	4dfr
lags	1bks	1cw3	1dys	1lep4	1fw6	1gvh	1i95	1jll	1krb	1oce	1rc4	1yjb	4eca
lah0	1bkw	1cwn	1dyv	1lep7	1fw8	1gvi	1i96	1jlr	1krh	1oel	1rcf	1yjc	4enl
lah2	1bky	1cwy	1dz3	1lep8	1fwa	1gw9	1i97	1jls	1ks2	1oen	1rch	1ykf	4eug
lah4	1bl5	1cx9	1dzf	1lep2	1fwc	1gwc	1i9c	1jlw	1ks9	1ofg	1rcx	1ylv	4fua
lah5	1bm6	1cxe	1dzu	1leqa	1fwd	1gwn	1i9d	1jmk	1ksg	1ofv	1rd7	1yme	4gpb
lahh	1bnc	1cxf	1dzv	1leqb	1fwe	1gyl	1i9g	1jmv	1ksh	1ohk	1rda	1ymu	4gsa
lahi	1bo6	1cxh	1dzw	1leqj	1fwf	1gym	1i9s	1jmy	1ksj	1oib	1rdb	1ymv	4gss
lahn	1boh	1cxi	1dzy	1leqp	1fwg	1gyp	1i9t	1jnd	1ksk	1oil	1rdc	1ypi	4gtu
lahp	1boi	1cxk	1e0c	1lequ	1fwh	1gyq	1ia1	1jnd	1ksk	1oil	1rdc	1ypi	4gtu
lai2	1bot	1cxl	1e0d	1leqy	1fwj	1gyt	1ia3	1jne	1ktb	1one	1re7	1ysc	4lip
lai3	1bou	1cyd	1e0s	1leri	1fwn	1gzd	1ia4	1jnf	1kte	1onr	1rea	1yvn	4mht
lai9	1boz	1cye	1e0t	1lert	1fwr	1gzg	1iaq	1jnr	1kti	1opr	1req	1zak	4nll
laiu	1bpm	1cz1	1e0u	1leru	1fws	1h0g	1iat	1jny	1ktn	1orb	1rf7	1zay	4nul
laiv	1bpn	1cz3	1e0v	1lerw	1fwt	1h0i	1iaw	1jom	1ku0	1ord	1rg7	1zen	4pfb
laj0	1bq3	1cza	1e0w	1les9	1fwy	1h1n	1ib6	1jp3	1kv5	1oro	1rh3	1zdf	4pgm
laj2	1bq4	1czk	1E15	1lesc	1fx1	1h2e	1ibj	1jp7	1kv8	1ort	1rhd	1zid	4pnp
lajb	1bq6	1czl	1E18	1lesd	1fx6	1h2f	1ibr	1jpd	1kva	1osi	1rht	1zin	4ptd
lajd	1bq7	1czn	1E19	1lesg	1fxj	1h2h	1ibs	1jph	1kvc	1osj	1ril	1zio	4req
laje	1bqc	1czo	1elc	1lesm	1fxo	1h2j	1ic6	1jpi	1kvq	1otp	1rk2	1zip	4rla
lajr	1bqm	1czt	1ele	1lesn	1fxp	1h4k	1ici	1jpk	1kvr	1ovt	1rkd	1zpd	4rub
lajs	1bqn	1czu	1elf	1lesv	1fxq	1h4m	1icp	1jpm	1kvs	1oya	1rkm	1zrm	4rve
lajz	1br2	1d02	1elk	1lesw	1fxs	1h4o	1icq	1jpn	1kvt	1p39	1rks	2cnd	4tgl

1ak2	1br4	1d03	1e1l	1etu	1fxu	1h4q	1ics	1jq3	1kvu	1pam	1rla	2cpu	4tim
1ak5	1brl	1d07	1elm	1eul	1fxw	1h4s	1id8	1jqd	1kwh	1pbc	1rlc	2cst	4tmk
1ak9	1brm	1d09	1eln	1eu8	1fy2	1h4t	1ide	1jqe	1kwm	1pbd	1rlr	2ctc	4tmy
lake	1bro	1d0i	1elq	1eua	1fy6	1h4v	1idf	1jqf	1kxj	1pbe	1rmh	2cxg	4trx
lakm	1brs	1d0s	1elr	1euc	1fyd	1h4x	1idm	1jqj	1kxq	1pbf	1rml	2dap	4tsl
laln	1brt	1d0v	1ely	1eud	1fye	1h4y	1ie3	1jq1	1kxt	1pbg	1rpa	2dbv	4uag
1akq	1brw	1d0x	1elz	1eug	1fzq	1h50	1iei	1jr2	1kxv	1pbp	1rpj	2dfp	4ubp
1akr	1bs0	1d0y	1E20	1euh	1fzt	1h5n	1iev	1jr3	1ky2	1pca	1rpt	2dhd	4xia
1akt	1bs1	1d0z	1e2d	1eui	1fzw	1h5q	1iew	1jr4	1ky3	1pda	1rpx	2dhe	4xim
1aku	1bs2	1d1a	1e2e	1eun	1g0n	1h5r	1iex	1jr6	1kyg	1pdh	1rrf	2dhf	4xis
1akv	1bs9	1d1b	1e2f	1euq	1g0o	1h5t	1if2	1jrx	1kyi	1pdy	1rt1	2dhq	4yas
1akw	1bsi	1d1p	1e2g	1euy	1g0r	1h5u	1ifx	1jry	1kyq	1pdz	1rt2	2dij	521p
1aky	1bsl	1d1q	1e2q	1euz	1g15	1h5v	1ig0	1js1	1kyw	1pea	1rt4	2dik	5cpa
1al3	1bss	1d1t	1e2s	1ev7	1g19	1h5y	1ig8	1js3	1kyx	1pek	1rt5	2dkb	5dfr
1al8	1bsv	1d2a	1e2y	1ev9	1g1a	1h60	1ig9	1js6	1kyz	1pfk	1rt6	2did	5enl
1ald	1bta	1d2f	1E32	1evi	1glf	1h62	1igb	1jsl	1kz1	1pgj	1rt7	2dlh	5eug
1alh	1btb	1d2g	1E33	1evq	1glg	1h63	1igr	1jsr	1kz4	1pgn	1rtd	2dmr	5fwg
1alk	1btc	1d2h	1E39	1evy	1gl1	1h65	1igs	1jsx	1kz6	1pgo	1rth	2dor	5gss
1ami	1btj	1d2i	1e3c	1evz	1glm	1h66	1igw	1jsz	1kz9	1pgp	1rti	2dpg	5gst
1amj	1btm	1d2m	1e3e	1ew2	1gl1	1h69	1ih7	1jt9	1kzh	1pgq	1rus	2dpm	5icd
1amk	1bto	1d2r	1e3i	1ew8	1g20	1h6a	1ih8	1jte	1lla	1pgt	1rva	2drc	5ldh
1amm	1bu6	1d3k	1e3j	1ew9	1g21	1h6b	1ihc	1ju3	1lle	1php	1rvc	2dri	5lip
1amp	1bu8	1d3v	1e3l	1ewd	1g23	1h6c	1ihi	1ju4	1ll1	1phr	1rvd	2ebn	5mdh
1amy	1bua	1d4c	1e3m	1ewe	1g2o	1h6d	1iho	1jud	1lls	1pif	1rve	2eck	5mht
1an0	1bup	1d4d	1e3q	1ewg	1g2p	1h6j	1ihu	1juk	1l2q	1pig	1rx1	2eda	5nll
1anf	1buz	1d4e	1e3t	1ex1	1g2q	1h6v	1ihx	1jul	1l2r	1pjb	1rx2	2erc	5p21
1ani	1bvz	1d4f	1e3x	1ex2	1g2u	1h6z	1iio	1juy	1l2t	1pje	1rx3	2eug	5pgm
1anj	1bwc	1d4g	1e3z	1ex5	1g2v	1h75	1iis	1jv2	1l2u	1pjr	1rx4	2euh	5pnt
1ank	1bwf	1d4n	1E40	1ex7	1g3i	1h76	1iie	1jvn	1l4u	1pkm	1rx5	2exo	5ptd
1ao0	1bwl	1d4o	1E46	1ex9	1g3l	1h79	1iis	1jw4	1l4y	1pkn	1rx6	2eza	5req
1ao3	1bwp	1d4z	1E47	1exc	1g3r	1h7a	1iib	1jw5	1l5a	1pky	1rx9	2ezb	5rla
1aoe	1bwq	1d5t	1E49	1exd	1g3u	1h7b	1iim	1jw9	1l5g	1plj	1rxo	2ezc	5rub
1aog	1bwr	1d5w	1e4a	1exm	1g41	1h7e	1iin	1jwb	1l5j	1plk	1s02	2fax	5sic
1aon	1bws	1d6f	1e4b	1exn	1g42	1h7g	1iir	1jwl	1l5t	1pll	1say	2fcr	5ukd
1aov	1bx0	1d6h	1e4c	1exp	1g4a	1h7h	1iit	1jwx	1l5v	1pnr	1sbc	2fdx	5ull
1apb	1bx1	1d6i	1e4e	1exv	1g4b	1h7n	1iji	1jxb	1l5w	1pnt	1sbh	2ffh	5xim
1aq0	1bx3	1d6j	1e4f	1eyr	1g59	1h7o	1ik4	1jxh	1l6f	1ppi	1sbi	2fhe	5xin
1aq2	1bx4	1d7a	1e4g	1eyu	1g5a	1h7q	1ik6	1jxi	1l6g	1prx	1sbn	2flv	5yas
1aq6	1bx9	1d7f	1e4i	1eyy	1g5c	1h7r	1ika	1jxj	1l6i	1psc	1sbp	2fmt	621p
1aqf	1bxb	1d7k	1e4l	1eyz	1g5f	1h7t	1ikv	1jxk	1l6s	1ptd	1sbt	2fok	6enl
1aqh	1bxn	1d7l	1e4m	1ez1	1g5r	1h7x	1ikw	1jxz	1l6w	1ptk	1sca	2fox	6gss
1aql	1bxr	1d7r	1e4n	1ezb	1g5t	1h81	1ikx	1jye	1l6y	1ptt	1scb	2fip	6gst
1aql	1bxs	1d7s	1e4v	1ezc	1g5x	1h83	1iky	1jyk	1l8a	1ptu	1scd	2flua	6gsu
1aqm	1bxz	1d7u	1e4y	1ezd	1g60	1h84	1iio	1jyn	1l8g	1ptv	1scn	2fvx	6gsv
1aqn	1bya	1d7v	1E51	1ezi	1g63	1h8e	1ile	1jyw	1l8k	1pty	1scu	2gar	6gsw
1aqu	1byb	1d7y	1E55	1ezp	1g6h	1h8h	1ilv	1jyz	1l9f	1pud	1sel	2gbp	6gsx
1aqv	1byc	1d8a	1E56	1ezr	1g6o	1h8l	1ilw	1jz0	1l9x	1pvi	1sep	2gd1	6gsy
1aqx	1byd	1d8c	1E58	1ezw	1g7f	1h93	1im2	1jz1	1la2	1pxa	1sfe	2glr	6icd
1aqy	1byi	1d8r	1E59	1f05	1g7g	1h94	1im8	1jz2	1laf	1pxb	1sft	2glt	6ldh

larl	lbyu	ld8t	le5d	lf07	lg7o	lh9a	limj	ljz3	llag	lpxt	lshk	2gpa	6mht
larm	lbzc	ld8w	le5e	lf0n	lg7r	lh9b	lin4	ljz4	llah	lpyd	lsib	2gpb	6nul
larz	lbzf	ld9t	le5f	lf0o	lg7s	lh9c	lin6	ljz5	llam	lpyg	lsjs	2gpn	6pfk
las2	lbzh	ld9v	le5k	lf0p	lg7u	lha3	lin7	ljz8	llan	lq2l	lsmd	2gsa	6ptd
las3	lbzj	ld9x	le5m	lf0y	lg8g	lhbj	lin8	ljzs	llap	lqam	lson	2gsr	6q2l
lass	lbzl	ld9y	le5n	lf12	lg8h	lhc7	lini	ljzt	llau	lqap	lsoo	2gss	6req
lasx	lbzy	ld9z	le5q	lf14	lg8j	lhd2	linj	ljzw	llav	lqaq	lsrr	2gst	6taa
latl	lc0d	ldad	le5v	lf17	lg8k	lhdc	linl	lk02	llaw	lqb4	lsrx	2gyi	6tim
latg	lc0g	ldae	le5x	lf1b	lg8l	lhde	lio2	lk03	llb8	lqb7	lst2	2had	6xia
latn	lc0i	ldaf	le60	lf1z	lg8m	lhdg	liof	lk06	llb9	lqb8	lst3	2hdh	6yas
latr	lc0k	ldah	le61	lf2l	lg8p	lhdh	lioi	lk08	llbb	lqbg	lsto	2hmi	72lp
lats	lc0l	ldaj	le62	lf2d	lg8r	lhdi	liok	lk0l	llbc	lqc9	lsud	2hmy	7cpa
lau9	lc0n	ldak	le63	lf2e	lg8y	lhdr	liol	lk0n	llbf	lqcc	lsup	2hnp	7dff
laur	lc0p	ldao	le64	lf2j	lg94	lhdv	lion	lk0o	llbg	lqcz	lszj	2hmq	7gpb
laur	lc0t	ldap	le66	lf2p	lg95	lhdx	liov	lk0u	llbh	lqde	lt7p	2hpa	7icd
laur	lc0u	ldb3	le69	lf2v	lg97	lhdx	liow	lk1d	llbi	lqdl	ltaq	2hvm	7mdh
laur	lc1c	ldbi	le6b	lf37	lg99	lhcz	lioz	lk1e	llbk	lqds	ltau	2kau	7mht
laur	lc1d	ldbp	le6c	lf3b	lg9h	lhcz	lipa	lk23	llbs	lqe0	ltcl	2lao	7odc
laur	lc1x	ldbr	le6e	lf3e	lg9r	lhcz	lipd	lk2y	llbt	lqe1	ltca	2lbp	7ptd
laur	lc29	ldbs	le6k	lf3i	lga1	lhcz	liph	lk30	llc0	lqe3	ltcb	2ldx	7taa
laur	lc2y	ldbt	le6n	lf3o	lga4	lhcz	lipi	lk32	llc3	lqf5	ltcc	2lip	7tim
laur	lc3e	ldbv	le6p	lf3p	lga6	lhcz	liq0	lk35	llc5	lqf6	ltcm	2liv	7xim
laur	lc3l	ldbw	le6q	lf3t	lga6	lhcz	liq8	lk3c	llc7	lqf7	ltdf	2ljr	7yas
law1	lc3s	ldby	le6r	lf3w	lga1	lhcz	liqp	lk3f	llc8	lqfe	ltcj	2man	82lp
law2	lc41	ldc1	le6s	lf3x	lgaq	lhcz	liqr	lk3r	llcf	lqff	ltcr	2mas	8cpa
law9	lc47	ldc3	le6u	lf48	lgar	lhcz	liqu	lk3t	llco	lqfm	ltec	2min	8dff
lax4	lc4c	ldc4	le6x	lf49	lgaw	lhcz	lir1	lk4g	llct	lqfn	lteh	2mjp	8gpb
lax9	lc4g	ldc5	le6z	lf4h	lgax	lhcz	lir6	lk5d	llcd	lqfx	ltfd	2mnr	8gss
laxe	lc4k	ldc7	le71	lf4l	lgc0	lhcz	liri	lk5g	llde	lqg0	ltfr	2muc	8icd
laxg	lc4o	ldc8	le77	lf4p	lgc2	lhcz	lirx	lk6a	lldm	lqg2	ltgl	2nac	8ldh
laxr	lc4w	ldcf	le79	lf57	lgc8	lhcz	liso	lk6d	lldn	lqg4	ltgo	2nad	8mht
lay0	lc4x	ldcm	le7m	lf5n	lgca	lhcz	lit7	lk6i	lldy	lqg6	lthf	2nap	8tim
lay7	lc50	ldct	le7q	lf5z	lgco	lhcz	lit8	lk6j	lleh	lqg8	lthg	2ngl	8xia
laye	lc5k	lddo	le7r	lf6l	lgcu	lhcz	litg	lk6w	llfl	lqgo	lthm	2nip	8xim
layl	lc72	lde0	le7s	lf6b	lgcx	lhcz	litq	lk6x	llfg	lqgq	ltht	2npx	9gss
laz1	lc7i	lde2	le7w	lf6d	lgdl	lhcz	litu	lk6y	llfh	lqgs	lthx	2nsy	9icd
laz4	lc7n	lde5	le89	lf6k	lgdh	lhcz	liu8	lk70	llfi	lqgu	ltia	2oat	9ldb
lazo	lc7o	lde6	le8c	lf6m	lgdt	lhcz	liud	lk75	llgy	lqgv	ltib	2ohx	9ldt
lb00	lc7q	ldea	le8m	lf6p	lgeg	lhcz	lijo	lk77	llh0	lqh1	ltic	2olf	9mht
lb05	lc83	lded	le8n	lf6w	lgeh	lhcz	liut	lk7c	lli5	lqh8	ltka	2one	9xia
lb0a	lc84	ldej	le92	lf6y	lgeq	lhcz	liuu	lk7f	lli7	lqh9	ltkb	2ora	9xim
lb0i	lc85	ldek	le94	lf73	lger	lhcz	liuv	lk7k	llij	lqha	ltkc	2otc	
lb0k	lc86	lder	le96	lf74	lges	lhcz	liuw	lk7x	llik	lqhe	ltmh	2oxi	
lb0l	lc87	ldex	le98	lf75	lget	lhcz	liux	lk83	llio	lqhl	ltmk	2pbg	
lb0m	lc88	ldf7	le99	lf7u	lgeu	lhcz	livr	lk89	lljr	lqhm	ltml	2pda	
lb0p	lc8k	ldfg	le9a	lf7v	lgew	lhcz	liuy	lk8c	llju	lqhn	ltmo	2pfk	
lb0u	lc8l	ldfi	le9b	lf8g	lgex	lhcz	liwe	lk8k	llk0	lqho	ltmy	2pgi	
lb14	lc8q	ldfj	le9c	lf8i	lgfi	lhcz	lixg	lk8y	llk5	lqhp	ltof	2pgt	
lb15	lc96	ldfm	le9d	lf8s	lgfy	lhcz	lixh	lk8z	llk7	lqhs	ltpe	2phh	

1b1a	1c97	1dgl	1e9e	1f8u	1gfb	1hn9	1ixi	1k92	1lkx	1qht	1tpf	2pke
1b1c	1c9d	1dg5	1e9i	1f8w	1gg0	1hnd	1j4a	1k97	1lkz	1qhx	1tpu	2pri
1b1h	1c9k	1dg8	1e9y	1f91	1gg1	1hnh	1j4b	1k9s	1ll0	1qhy	1tpv	2prj
1b1y	1c9w	1dg9	1e9z	1f9a	1gg5	1hni	1j4x	1kae	1ll2	1qhz	1tpw	2prk
1b26	1cb0	1dga	1ea5	1f9b	1gg9	1hnj	1j5e	1kag	1ll3	1qi0	1trl	2ptd

APPENDIX N

Names of proteins in database for $\alpha+\beta$ secondary structural class

102l	latn	lbsj	ld3j	leic	lfkj	lggg	li0z	ljff	lkqb	llo8	lplq	lse4	2lzh
103l	latp	lbsk	ld3m	leid	lfkk	lggx	li10	ljg0	lkqc	llou	lplr	lset	2lzt
104l	lau0	lbsp	ld3n	leie	lfkl	lgh8	li12	ljg5	lkqd	llov	lpmd	lset	2mat
107l	lau2	lbsr	ld3w	leig	lfkq	lghh	li17	ljgt	lkqf	lloy	lpne	lsha	2mbr
109l	lau3	lbt0	ld3z	leih	lfkr	lghl	li1e	ljh3	lkqg	lpi	lpnk	lshb	2mea
110l	lau4	lbt8	ld4b	leik	lfks	lgif	li1f	ljh6	lkqk	lpy	lpnl	lshd	2meb
111l	laur	lbtv	ld4c	leil	lfkv	lgih	li1g	ljh7	lkr4	llqv	lpnm	lskj	2mec
112l	laus	lbu4	ld4e	leim	lfl0	lgii	li1i	ljhl	lkra	llqy	lpoh	lslm	2med
113l	lav4	lbu9	ld4x	leiq	lflq	lgio	li1q	ljht	lkrb	llr0	lpop	lsly	2mee
114l	lav5	lbu9	ld5m	leir	lfls	lgjv	li1y	ljhw	lkrc	llr4	lppd	lsmn	2meh
115l	lavk	lbux	ld5n	lejc	lflu	lgjz	li1z	lji8	lkrh	llra	lppn	lsmp	2mha
118l	lavl	lbv1	ld5t	lejd	lflw	lgk8	li20	ljic	lks2	llru	lppp	lsox	2mob
119l	lavm	lbvi	ld5y	lejr	lfly	lgk9	li21	ljit	lksk	llry	lpre	lsph	2ms2
11as	law0	lbvk	ld5z	lejs	lfm0	lgkf	li22	ljiv	lksl	llsa	lprq	lspr	2msb
11ba	law3	lbvq	ld6a	lejt	lfm1	lgkv	li2e	ljiy	lkss	llsb	lpsd	lspg	2muc
11bg	lawd	lbvt	ld6e	leju	lfm5	lgkw	li2f	ljiz	lksu	llsc	lptf	lspu	2myo
120l	lawi	lbwc	ld6p	lek8	lfm7	lgkz	li2g	ljj0	lksv	llsd	lpto	lsrn	2npx
122l	lawp	lbwd	ld6q	leks	lfm8	lgl3	li33	ljj1	lkt9	llse	lpuc	lsry	2pbh
123l	lawz	lbwh	ld6t	lekt	lfm0	lgm7	li35	ljj3	lktk	llsf	lput	lssa	2pda
125l	laxq	lbwi	ld6u	lekz	lfm9	lgm8	li3f	ljic	lkuf	llsg	lpxb	lssb	2phh
126l	laxw	lbwj	ld6y	lel0	lfn9	lgm9	li3q	ljje	lkug	llsm	lpxc	lssc	2phm
128l	laxx	lbwz	ld6z	lel1	lfn0	lgmp	li3r	ljit	lkuh	llsn	lqad	lssn	2phy
129l	lay2	lax2	ld7h	lel5	lfn9	lgmq	li4g	ljiv	lkui	llsp	lqae	lss0	2pia
12as	laya	lax6	ld7j	lel6	lfnj	lgms	li4h	ljk7	lkuk	llsq	lqap	lstd	2pil
131l	layb	laxe	ld7l	lel7	lfnk	lgmv	li4p	ljk9	lkuo	llsz	lqbe	lste	2pld
132l	layc	lbxi	ld7y	lel8	lfnm	lgmw	li4q	ljka	lkut	llt3	lqcg	lstf	2pol
133l	layd	lbnx	ld8v	lel9	lfn0	lgn3	li4r	ljkb	lkuv	llt4	lqci	lstu	2prf
134l	laye	lbnx	ld8z	leli	lfo4	lgn4	li50	ljkc	lkuv	lltd	lqej	lsxl	2pro
137l	layk	lbxt	ld9a	lelo	lfo6	lgn6	li56	ljkd	lkux	llth	lqeq	lsxt	2ptl
139l	layu	lbxy	ld9u	lels	lfoh	lgnd	li5o	ljki	lkuy	lltu	lqcr	lsyn	2pva
140l	layv	lby2	ldap	lem1	lfq1	lgnk	li6h	ljkk	lkvl	lltv	lqcs	lszj	2pyp
141l	layw	lbyl	ldar	lem7	lfr5	lgnu	li70	ljkl	lkv2	lltz	lqd2	ltay	2pyr
142l	laz9	lbyq	ldbd	lemc	lfrd	lgoi	li7f	ljkn	lkvn	llus	lqdd	ltbd	2rat
143l	lazf	lbyr	ldbu	lemd	lfrh	lgol	li7h	ljks	lkvv	lluv	lqdl	ltbp	2rbi
144l	lazp	lbys	ldbv	leme	lfrl	lgos	li7k	ljkt	lkvv	lluv	lqdn	ltce	2reb
146l	laze	lbz9	ldbx	lemf	lfrj	lgou	li7o	ljky	lkwm	llwx	lqdq	ltco	2rgf
147l	lb02	lbzl	ldc2	lemg	lfrk	lgov	li7r	ljlo	lkxw	llxd	lqe0	ltcs	2rnf
148l	lb04	lc05	ldc3	lemh	lfrl	lgoy	li7s	ljll	lkxx	llxt	lqe6	ltda	2rns
150l	lb06	lc06	ldc4	lemm	lfrn	lgoz	li7t	ljm6	lkyi	llyd	lqf0	ltdb	2rnt
151l	lb0d	lc08	ldc5	lemv	lfr0	lgpe	li7u	ljmf	lkyp	llye	lqf1	ltde	2rus
152l	lb0g	lc0a	ldc6	lemw	lfrf	lgpj	li87	ljmi	lkyr	llyf	lqf2	ltdj	2sar
153l	lb0p	lc0b	ldco	len4	lfrs	lgpm	li8c	ljml	lkys	llyj	lqf6	ltdu	2seb
155l	lb0t	lc0c	ldd3	len5	lfrt	lgpx	li8t	ljn0	lkzf	llyl	lqfe	ltdy	2shp

156l	1b1e	1c0i	1dd4	1en7	1frx	1gqv	1i94	1jn1	1kzi	1lyo	1qfr	1tew	2sic
157l	1b1i	1c0k	1dd5	1eo6	1fs3	1gr0	1i95	1jne	1l00	1lys	1qgl	1the	2std
158l	1b1j	1c0l	1dd6	1eos	1fsi	1grb	1i96	1jnk	1l01	1lyy	1qgi	1thl	2sxl
160l	1b1z	1c10	1ddh	1eot	1fsk	1gre	1i97	1jnr	1l02	1lyz	1qgw	1thy	2tel
161l	1b20	1c21	1ddk	1eow	1fsz	1grf	1ia8	1joa	1l03	1lz2	1qh3	1tif	2td
163l	1b21	1c23	1ddo	1eps	1fte	1grg	1iaa	1joe	1l04	1lz3	1qb4	1tin	2tdm
164l	1b24	1c24	1ddt	1epy	1fte	1gri	1iad	1jow	1l05	1lz4	1qhc	1tis	2tec
165l	1b25	1c27	1de3	1eq0	1ftf	1grj	1iae	1jpd	1l06	1lz6	1qhw	1tla	2tlx
166l	1b2k	1c43	1def	1eq4	1fth	1grl	1iag	1jpf	1l07	1lz8	1qi7	1tig	2tmn
167l	1b2m	1c44	1dej	1eq5	1ftr	1grt	1iao	1jpg	1l08	1lz9	1qil	1tlp	2tpr
168l	1b2s	1c45	1der	1eqk	1fu5	1gsa	1ib1	1jpm	1l09	1lza	1qin	1tls	2tpt
169l	1b2u	1c46	1det	1eqm	1fu6	1gsn	1ib6	1jq4	1l0c	1lzb	1qio	1tlx	2tsc
170l	1b2v	1c47	1deu	1eqo	1fug	1gso	1ib7	1jqg	1l0o	1lzc	1qip	1tmc	2tsr
171l	1b2z	1c4a	1df0	1eqr	1fus	1gsp	1ib8	1jqj	1l0v	1lze	1qj2	1tmn	2tss
172l	1b33	1c4c	1dfa	1eqt	1fut	1gsv	1iba	1jqu	1l0w	1lzh	1qjh	1tn9	2u1a
174l	1b38	1c4f	1dff	1eqy	1fuw	1gsw	1ibx	1jqw	1l10	1lzn	1qji	1tol	2u2f
178l	1b39	1c4g	1dfj	1erz	1fv1	1gsx	1ic4	1jrl	1l11	1lzt	1qk1	1tox	2ucz
179l	1b3a	1c4k	1dga	1es0	1fva	1gt8	1ic5	1jr9	1l12	1lzs	1qkf	1tpt	2ugi
180l	1b3j	1c60	1dgj	1esl	1fvi	1gte	1ic7	1jrk	1l13	1lzt	1qkh	1trg	2ush
181l	1b3o	1c61	1dgt	1esr	1fvq	1gth	1icc	1jrm	1l14	1lzy	1qki	1tsd	2ung
183l	1b4a	1c64	1dhn	1esv	1fvs	1gtp	1icr	1jro	1l15	1m0v	1qkz	1tsl	2vaa
184l	1b4o	1c65	1dhy	1et6	1fvt	1gup	1icu	1jrp	1l16	1mlj	1ql6	1tsm	2vab
185l	1b4s	1c66	1di3	1et9	1fvz	1guq	1icv	1jrj	1l18	1mlx	1qla	1tsn	2vao
186l	1b4v	1c67	1di4	1eu3	1fw6	1guw	1icw	1jrx	1l19	1m4d	1qlb	1tsv	2vik
187l	1b50	1c68	1di5	1eu4	1fw9	1guy	1icx	1jry	1l20	1m4g	1qlf	1tsw	2vil
188l	1b53	1c69	1di8	1euc	1fwa	1gv0	1ids	1jsb	1l21	1m4i	1qlu	1tsy	2ypn
189l	1b5a	1c6a	1di9	1eud	1fwe	1gv1	1ie0	1jse	1l22	1m4v	1qm4	1tsz	3cox
1914	1b5b	1c6c	1dih	1eue	1fwf	1gv7	1ie3	1jsf	1l23	1m55	1qma	1tum	3crx
191l	1b5m	1c6d	1dik	1eui	1fwg	1gw5	1ie7	1jsk	1l24	1m5h	1qmh	1tvu	3dap
192l	1b5q	1c6e	1dix	1ev8	1fwh	1gwd	1iea	1jss	1l25	1m5k	1qmt	1tvv	3dbv
194l	1b5u	1c6f	1diz	1evf	1fwi	1gx3	1ieb	1jsv	1l26	1m5s	1qn3	1tvw	3dni
195l	1b5v	1c6g	1dj0	1evg	1fwj	1gxd	1iee	1jtm	1l27	1m8a	1qn4	1typ	3enl
196l	1b5w	1c6h	1dk0	1evi	1fwp	1gy6	1iet	1jtn	1l28	1ma1	1qn7	1tys	3erk
197l	1b5x	1c6i	1dkh	1evu	1fx3	1gy7	1ieu	1jto	1l29	1mat	1qn9	1tyt	3eza
198l	1b5y	1c6l	1dki	1evw	1fxd	1gyb	1ifq	1jtp	1l2n	1mb1	1qna	1u9a	3fru
199l	1b5z	1c6m	1dkj	1ew4	1fxh	1gyp	1ifs	1jtg	1l30	1mbb	1qnb	1u9b	3gpd
1a06	1b62	1c6n	1dks	1ewf	1fxi	1gyq	1ift	1jtt	1l32	1mbt	1qnc	1uae	3grs
1a0k	1b63	1c6p	1dkt	1ewp	1fxr	1gzp	1ifv	1jtu	1l33	1md1	1qnd	1ubi	3gsp
1a15	1b64	1c6q	1dl5	1ewq	1fxt	1gzq	1igl	1ju6	1l34	1mdl	1qne	1ubq	3hla
1a1m	1b65	1c76	1dlg	1ewy	1fys	1h0g	1ig4	1juf	1l35	1mdr	1qnn	1uch	3hoh
1a1n	1b69	1c77	1dm0	1ex8	1fyt	1h0i	1igd	1jug	1l37	1mee	1qo8	1uco	3jdw
1a1o	1b6a	1c78	1dmm	1exi	1fzj	1h19	1ihb	1juj	1l38	1mel	1qoa	1ud7	3kbp
1a2k	1b6f	1c79	1dna	1exj	1fzk	1h4q	1ihx	1jut	1l39	1mem	1qob	1udi	3kvt
1a2n	1b6i	1c7k	1dnc	1exu	1fzm	1h4r	1iiz	1jv2	1l3f	1mgs	1qof	1ugi	3lad
1a2p	1b6r	1c7p	1dob	1eyn	1fzu	1h4s	1ijr	1jvi	1l3k	1mhc	1qog	1uib	3lhm
1a2w	1b6s	1c8c	1doc	1eyp	1g02	1h4v	1ikl	1jvp	1l3l	1mhd	1qpr	1uic	3lym
1a2y	1b70	1c8u	1dod	1eyq	1g06	1h52	1ikm	1jvt	1l3r	1mhe	1qpz	1uid	3lyo
1a3a	1b77	1c8w	1doe	1eyz	1g07	1h53	1ikp	1jvu	1l40	1mif	1qqd	1uie	3lyt
1a3s	1b7g	1c9f	1doi	1ez1	1g0d	1h6b	1ikq	1jvv	1l41	1mil	1qqj	1uih	3lyz

1a4h	1b7l	1c9h	1d0k	1e2m	1g0g	1h6c	1i13	1jw3	1l42	1mir	1qqq	1una	3lz2
1a4v	1b7m	1c9v	1d0l	1ezz	1g0j	1h6d	1i14	1jw9	1l45	1mit	1qr1	1uox	3lzm
1a4y	1b7n	1c9x	1d0m	1f03	1g0k	1h6h	1i15	1jwa	1l46	1mka	1qs1	1urn	3mat
1a5k	1b7p	1ca7	1d0n	1f04	1g0l	1h6m	1i18	1jwb	1l47	1mla	1qs5	1ush	3mds
1a5l	1b7q	1cb5	1d0x	1f09	1g0m	1h6v	1i1p	1jwi	1l48	1mlc	1qs9	1ute	3mon
1a5m	1b7r	1cbk	1dpg	1f0b	1g0p	1h6z	1im9	1jwr	1l49	1mld	1qsb	1uug	3nul
1a5n	1b7s	1cbo	1dpt	1f0l	1g0q	1h70	1imu	1jww	1l50	1mli	1qse	1uxy	3pbh
1a5o	1b87	1cby	1dpw	1f0q	1g0s	1h72	1inn	1jx1	1l51	1mmm	1qsf	1vac	3pmg
1a5p	1b8a	1cc2	1dpx	1f0x	1g0w	1h7s	1io5	1jxa	1l52	1mng	1qsq	1vad	3pro
1a5q	1b8p	1cc4	1dq3	1f0z	1g10	1h7u	1iok	1jxv	1l53	1mnl	1qsw	1vao	3pva
1a5r	1b8u	1cc6	1dru	1f10	1g11	1h7w	1ioo	1jy5	1l54	1mol	1qt3	1vde	3pyp
1a6d	1b8v	1cc8	1drv	1f13	1g12	1h7x	1ioq	1jya	1l55	1mpg	1qt4	1vfb	3rat
1a6e	1b99	1cd1	1drw	1f1b	1g19	1h82	1ior	1jyh	1l56	1mpy	1qt5	1vfr	3rn3
1a6f	1b9l	1cdk	1drz	1f28	1g1b	1h84	1ios	1k0e	1l57	1mrg	1qt6	1vmp	3rnt
1a6j	1ban	1ce7	1dss	1f2a	1g1r	1h86	1iot	1k0g	1l58	1mrh	1qt7	1vok	3rsk
1a6l	1bao	1ce8	1dt0	1f2b	1g1s	1h87	1iov	1k0l	1l59	1mri	1qt8	1vza	3rsp
1a6z	1bb4	1cer	1dt9	1f2h	1g1t	1h8c	1iow	1k0r	1l5g	1mrj	1qt9	1vzb	3sdp
1a70	1bb5	1cew	1dtj	1f2r	1g1v	1h8j	1ip1	1k0t	1l5p	1mro	1qtb	1vzc	3seb
1a73	1bb6	1cf0	1dtn	1f31	1g1w	1h8m	1ip2	1k0v	1l60	1msb	1qtc	1vzd	3srn
1a74	1bb8	1cf2	1dtp	1f32	1g27	1h8u	1ip4	1k1c	1l62	1msc	1qtd	1vze	3ssi
1a7k	1bbu	1cf5	1dur	1f34	1g2a	1h93	1ip6	1k26	1l63	1msd	1qth	1wdb	3tli
1a7t	1bbw	1cfe	1duy	1f3c	1g2r	1h94	1ip9	1k28	1l64	1msg	1qto	1wqn	3tmn
1a8r	1bbx	1cfy	1dv1	1f3p	1g2t	1h98	1ipa	1k2a	1l66	1mst	1qtv	1wqo	3tms
1a8t	1bc4	1cgq	1dw9	1f3y	1g3k	1h9a	1ipc	1k3e	1l68	1muc	1qu6	1wqp	3tss
1a9b	1bc6	1ch0	1dwm	1f40	1g3n	1h9o	1ipg	1k3s	1l6u	1mut	1qu9	1wqq	3usn
1a9c	1bcp	1chk	1dwm	1f47	1g4a	1ha1	1ipp	1k44	1l6v	1mva	1qua	1wqr	4enl
1a9u	1bd2	1chm	1dwn	1f4b	1g4b	1ha6	1iq0	1k46	1l70	1mxs	1qud	1xej	4erk
1aa1	1bd6	1chr	1dxj	1f4d	1g57	1han	1iq4	1k47	1l71	1mxc	1qug	1xek	4fit
1aar	1bdf	1chu	1dxk	1f4e	1g58	1hc7	1iqo	1k49	1l72	1myo	1qup	1xer	4fxc
1ab2	1bdu	1ci7	1dy3	1f4g	1g61	1hcj	1iqq	1k4i	1l73	1nbe	1raa	1xgm	4gr1
1abm	1be3	1cj2	1dy5	1f52	1g62	1hck	1iqs	1k4j	1l74	1nel	1rab	1xgn	4gsp
1abr	1be4	1cj3	1dya	1f5a	1g6b	1hcl	1iqz	1k4l	1l75	1ncv	1rac	1xgo	4kbp
1acb	1bel	1cj4	1dyc	1f5b	1g6j	1hd7	1ir0	1k4n	1l77	1nda	1rad	1xgs	4lym
1acm	1bf3	1cj6	1dyd	1f5c	1g6s	1hdg	1ir8	1k4o	1l79	1ndl	1rae	1xps	4lyt
1adj	1bfi	1cj7	1dye	1f5m	1g6z	1hdm	1ir9	1k51	1l7y	1nec	1raf	1xpt	4lyz
1aec	1bfp	1cj8	1dyf	1f5q	1g7h	1hdw	1isa	1k58	1l80	1ned	1rag	1xrb	4lzm
1af0	1bfx	1cja	1dyq	1f6r	1g7i	1he0	1isb	1k59	1l81	1nel	1rah	1xrc	4mat
1af5	1bg0	1cje	1dyt	1f6s	1g7j	1he6	1isc	1k5a	1l82	1nhb	1rai	1yag	4mdh
1afi	1bg1	1cjf	1dyu	1f6t	1g7k	1hem	1isk	1k5h	1l83	1nhk	1rar	1yal	4mon
1afj	1bgi	1cjl	1dz5	1f7l	1g7l	1hen	1it6	1k83	1l84	1nhp	1ras	1yam	4ota
1afl	1bgj	1cxj	1dzf	1f7s	1g7m	1heo	1iu4	1k8i	1l85	1nhq	1rat	1yan	4otb
1afu	1bgn	1ck1	1dzn	1f7t	1g7p	1hep	1iut	1k8k	1l86	1nhs	1rba	1yao	4otc
1agb	1bgo	1ck2	1dzs	1f7u	1g7q	1heq	1iuu	1k92	1l87	1nja	1rbh	1yap	4pad
1agc	1bgs	1ck5	1E01	1f7v	1g81	1her	1iuv	1k97	1l88	1njc	1rbn	1yaq	4pro
1agd	1bgy	1ck8	1E09	1f7w	1g8t	1hfd	1iuw	1k9a	1l89	1njd	1rbo	1ydr	4rat
1agf	1bh5	1ck9	1e0g	1f7x	1g91	1hfe	1iux	1ka6	1l8j	1nje	1rbw	1yds	4rnt
1agi	1bhn	1ckc	1e0y	1f80	1g96	1hfn	1ivm	1ka7	1l8u	1nlk	1rbx	1ydt	4rsd
1ah6	1bbz	1ckf	1E10	1f82	1g9q	1hfo	1ivu	1kap	1l90	1nox	1rca	1yer	4rub
1ah8	1bi7	1ckg	1E15	1f83	1gad	1hfx	1ivv	1kbg	1l92	1npc	1rck	1yes	4srn

1aha	1bib	1ckh	1elo	1f8a	1gae	1hfy	1ivx	1kbp	1193	1npk	1rci	1yet	4tli
1ahb	1bir	1cki	1elt	1f8q	1gal	1hgz	1iw7	1kc7	1194	1npv	1rcm	1ygw	4tln
1ahc	1bjg	1ckj	1elv	1f8s	1gaq	1hhg	1ixx	1kce	1195	1nrc	1rcn	1ypn	4tmn
1ahj	1bjp	1ckm	1elx	1f8w	1gay	1hhh	1iy3	1kcm	1196	1nsa	1rds	1yvn	4tms
1ahq	1bk7	1ckn	1E21	1f95	1gaz	1hhi	1iy4	1ke2	1197	1nsk	1re2	1zdj	4tss
1ahu	1bkl	1cko	1E22	1f96	1gb0	1hhk	1j4g	1kdn	1198	1nsp	1rea	1zdk	4tsu
1ahv	1bkm	1ckp	1E24	1f98	1gb1	1hhl	1j4r	1ke5	1199	1nue	1reg	1zer	4ubp
1ahz	1bko	1cks	1E27	1f9c	1gb2	1hhq	1j4w	1ke6	119a	1obs	1rem	1zer	5enl
1ai4	1bkp	1ckv	1E32	1f9i	1gb3	1hi3	1j5e	1ke7	119f	1obt	1rex	1zjf	5fit
1ai6	1bkw	1clf	1E39	1f9n	1gb5	1hi4	1j5k	1ke8	119m	1oel	1rey	1zpr	5hoh
1ai7	1bkx	1cli	1e3m	1f9z	1gb6	1hi5	1j5p	1kek	119v	1ofg	1rfp	2cpk	5ldh
1aim	1bl4	1cm2	1e3p	1fa0	1gb7	1hka	1j5u	1kf2	11a2	1onc	1rga	2crk	5lym
1aip	1bl6	1cm3	1e3r	1fa3	1gb8	1hla	1j5w	1kf3	11c0	1one	1rgc	2crx	5lyt
1aiq	1bl7	1cm8	1e3v	1fa5	1gb9	1hlp	1j5y	1kf5	11c3	1opd	1rge	2csn	5lyz
1aj6	1blj	1cm9	1E44	1fa7	1gbo	1hhr	1j6r	1kf6	11ca	1opy	1rgf	2dbv	5mdh
1ajm	1blk	1cn7	1e4e	1fa8	1gbs	1hlu	1j6v	1kf8	11cb	1ord	1rgg	2def	5msf
1ajn	1blu	1cn8	1e5d	1fbi	1gbx	1hlw	1j6w	1kfi	11ce	1osz	1rgh	2dhn	5pad
1ajp	1blv	1cn9	1e5q	1fbx	1gbz	1hmj	1j6y	1kfq	11cn	1otf	1rgk	2dik	5rat
1ajq	1blx	1cof	1e5u	1fca	1gc6	1hmk	1j74	1kfy	11co	1otg	1rhh	2dln	5rnt
1ak6	1bmk	1coj	1e6e	1fcb	1gcb	1hml	1j7a	1kh0	11d9	1otp	1rhl	2dpg	5rsa
1ak7	1bnc	1com	1e6n	1fd2	1gcu	1hml	1j7b	1kh1	11dc	1oua	1rhp	2emd	5rub
1aki	1bne	1cp9	1e6p	1fd6	1gc2	1hnw	1j7c	1kh2	11dm	1oub	1ris	2emn	5sic
1akj	1bnf	1cpj	1e6r	1fd8	1gd0	1hnz	1j7d	1khm	11dn	1oud	1rl6	2emo	5tli
1akl	1bng	1cqa	1e6t	1fda	1gd3	1hnc	1j7i	1kif	11dp	1oue	1rlc	2eot	5tln
1aml	1bnj	1cqd	1e6v	1fdb	1gd4	1hoc	1j7l	1kip	11eg	1ouf	1rls	2eql	5tmn
1am7	1bni	1cqi	1e6y	1fdd	1gd5	1hov	1j7n	1kiq	11ew	1oug	1rnl	2erk	5tss
1an5	1bnr	1cqj	1e7d	1fdl	1gd6	1hp4	1j7u	1kir	11ez	1ouh	1rm4	2fd2	6enl
1an8	1bns	1cqu	1e7l	1fdn	1gd8	1hp5	1j8c	1kj3	11gr	1oui	1rnb	2fdn	6fdl
1an9	1bnz	1crk	1e7m	1fe0	1gdg	1hqi	1j8i	1kj8	11gu	1ouj	1rnc	2fhi	6fdr
1ang	1bo1	1crw	1e7p	1fe4	1gdo	1hgy	1j91	1kj9	11gx	1oun	1rmd	2fit	6fit
1ao0	1bo8	1cs8	1e8f	1fea	1gdw	1hr0	1j98	1kji	11hh	1p1g	1rnf	2fke	6gsp
1ao7	1boa	1cse	1e8g	1feb	1gdx	1hrj	1ja2	1kjj	11hi	1p38	1rnm	2fmr	6ldh
1aob	1bob	1csn	1e8h	1fec	1ge0	1hsb	1ja4	1kjk	11hk	1paf	1rnn	2gbl	6lyt
1aog	1bol	1csy	1e8o	1fee	1ge2	1hsk	1ja6	1kjq	11hl	1pag	1rnq	2gdl	6lyz
1aon	1bow	1csz	1e8s	1feh	1ge3	1hsw	1ja7	1kjt	11hm	1pba	1rnt	2gls	6lzm
1aot	1box	1ct9	1E94	1fer	1ge4	1hsx	1ja8	1kkc	11i2	1pbb	1rnw	2gnk	6rat
1aou	1bp0	1cte	1E97	1fes	1ge5	1ht2	1jak	1kko	11i3	1pbc	1rnz	2gsp	6rnt
1ap0	1bp4	1cu0	1e9i	1ff3	1ge6	1htd	1jat	1kkr	11it	1pbd	1rmz	2hcc	6rsa
1ap5	1bp6	1cu2	1e9m	1ff9	1ge7	1htt	1jb0	1kll	11ja	1pbe	1rob	2hea	7enl
1ap6	1bpj	1cu6	1e9n	1ffl	1ge8	1huc	1jb2	1kmm	11k5	1pbf	1rof	2heb	7fdl
1ap7	1bq1	1cup	1e9y	1ffu	1ge9	1huf	1jb4	1kmm	11kr	1pca	1rot	2hec	7fdr
1ap8	1bq2	1cuq	1ea2	1fg2	1gec	1hum	1jb5	1kmz	11ks	1pch	1rou	2hed	7gsp
1apa	1bqb	1cv0	1ear	1fhi	1geh	1hun	1jbb	1kn6	11kz	1pci	1rpf	2hee	7lyz
1apg	1bqg	1cv1	1eb0	1fhs	1ger	1hw3	1jbi	1knf	11lh	1pcz	1rpg	2hef	7lzm
1apm	1bqh	1cv3	1ebg	1fit	1get	1hw4	1jbr	1kni	11ma	1pda	1rph	2hgs	7mdh
1aps	1bqi	1cv4	1ebh	1fhu	1gev	1hw7	1jbs	1knj	11mc	1pdh	1rra	2hid	7pck
1aq1	1bqx	1cv5	1ec7	1fhv	1gez	1hwn	1jbt	1knp	11mh	1pdx	1rsn	2hoh	7rnt
1aq3	1br5	1cv6	1ec9	1fi4	1gf3	1hwo	1jc5	1knr	11mn	1pdy	1rta	2hpr	7rsa
1aq4	1brg	1cv8	1ecb	1fie	1gf4	1hwp	1jc9	1kob	11mo	1pdz	1rtb	2ifl	7thn

1aqa	1brh	1cvj	1ecc	1fil	1gf5	1hxd	1jch	1kor	1lmp	1pe6	1rtc	2iff	8rat
1aqd	1bri	1cvk	1ecf	1fj3	1gf6	1hxp	1jct	1kot	1lmq	1pfi	1rto	2igd	8rnt
1aqp	1brk	1cvz	1ecg	1fj7	1gf7	1hxq	1jcu	1kou	1lmt	1pfn	1rts	2igg	8ruc
1aqz	1brm	1c wd	1ecj	1fjc	1gf9	1h xv	1jd1	1kp0	1ln1	1pga	1rtu	2igh	8tli
1ar4	1brn	1cwe	1ecq	1fje	1gfa	1hxy	1jd3	1kp3	1ln2	1pgb	1rus	2ihl	9ldb
1ar5	1brs	1cxy	1eej	1fjg	1gfg	1hye	1jdq	1kp6	1ln3	1pgx	1ruv	2i18	9ldt
1aro	1brw	1cyo	1ef2	1fjm	1gfh	1hyf	1jdw	1kpa	1ln4	1phh	1rxo	2kau	9pap
1arz	1bry	1cyu	1efg	1fjq	1gfj	1hz1	1jdy	1kpb	1lna	1phk	1sap	2kce	9rat
1ask	1bs2	1cyv	1efw	1fju	1gfk	1hz5	1je4	1kpc	1lnb	1phz	1sar	2178	9rnt
1ast	1bs3	1czd	1efx	1fkb	1gfl	1hz6	1je6	1kpe	1lnc	1pic	1sat	2ldx	9rsa
1atl	1bsa	1czp	1ega	1fkd	1gfu	1hzt	1jeh	1kpf	1lnd	1pil	1sbn	2lgs	9rub
1at5	1bsb	1dlr	1egl	1fkf	1gfv	1i08	1jem	1kpp	1lne	1pin	1scu	2lhm	
1at6	1bsd	1dlz	1eh1	1fkg	1gg3	1i09	1jep	1kpq	1lng	1pip	1sdf	2lyo	
1atk	1bse	1d2y	1eh4	1fkh	1gga	1i0h	1jex	1kpt	1lnu	1pkp	1se2	2lyz	
1atl	1bsf	1d3f	1ehw	1fki	1ggt	1i0v	1jfl	1kq4	1lo7	1plf	1se3	2lz2	

APPENDIX O

Prediction accuracies for each chain of each protein in all- α class in trial set

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.431	1E28	A	276
0.273	1E28	B	99
0.750	1E28	C	8
0.903	101M	A	154
0.903	108M	A	154
0.903	112M	A	154
0.744	1A03	A	90
0.744	1A03	B	90
0.714	1A0A	A	63
0.683	1A0A	B	63
0.617	1A0P	A	290
0.857	1A1W	A	91
0.888	1A2S	A	89
0.761	1A32	A	88
0.906	1A38	A	245
0.906	1A38	B	245
0.450	1A38	P	20
0.450	1A38	Q	20
0.811	1A5T	A	334
0.762	1A63	A	130
0.907	1A7E	A	118
0.929	1ABM	A	198
0.929	1ABM	B	198
0.908	1ABY	A	283
0.842	1ABY	B	146
0.870	1ABY	D	146
0.857	1AEB	A	294
0.857	1AEF	A	294
0.857	1AEH	A	294
0.857	1AEN	A	294
0.857	1AES	A	294
0.857	1AEU	A	294
0.588	1AH7	A	245
0.815	1AHR	A	146
0.807	1AJ4	A	161
0.887	1AJ9	A	141

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.967	1HJT	A	153
0.863	1HM3	A	700
0.751	1HM6	A	346
0.769	1HM6	B	346
0.918	1HMA	A	73
0.956	1HMD	A	113
0.956	1HMD	B	113
0.956	1HMD	C	113
0.956	1HMD	D	113
0.773	1HNX	B	256
0.833	1HNX	C	239
0.909	1HNX	D	209
0.821	1HNX	E	162
0.644	1HNX	F	101
0.885	1HNX	G	156
0.891	1HNX	H	138
0.867	1HNX	I	128
0.819	1HNX	J	105
0.860	1HNX	K	129
0.852	1HNX	L	135
0.905	1HNX	M	126
0.557	1HNX	N	61
0.955	1HNX	O	89
0.864	1HNX	P	88
0.848	1HNX	Q	105
0.886	1HNX	R	88
0.796	1HNX	S	93
0.972	1HNX	T	106
0.923	1HNX	V	26
0.698	1HO8	A	480
0.543	1HQ1	A	105
0.988	1HQB	A	80
0.745	1HRO	A	106
0.764	1HRO	B	106
0.676	1HS5	A	34
0.676	1HS5	B	34

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.863	1AJ9	B	146
0.862	1AKD	A	414
0.866	1ALU	A	186
0.866	1AOQ	A	567
0.847	1AOQ	B	567
0.929	1AP5	A	198
0.929	1AP5	B	198
0.929	1AP6	A	198
0.929	1AP6	B	198
0.715	1APX	A	249
0.378	1APX	B	249
0.378	1APX	C	249
0.378	1APX	D	249
0.774	1ARR	A	53
0.774	1ARR	B	53
0.660	1ASH	A	150
0.713	1AW9	A	216
0.664	1AX8	A	146
0.908	1B0B	A	142
0.694	1B0N	A	111
0.667	1B0N	B	57
0.857	1B1B	A	140
0.738	1B3U	A	588
0.723	1B3U	B	588
0.884	1B6W	A	69
0.834	1B7Y	A	350
0.304	1B7Y	B	785
0.746	1B82	A	351
0.741	1B82	B	351
0.749	1B85	A	351
0.749	1B85	B	351
0.825	1B8F	A	509
0.880	1B8R	A	108
0.644	1B8Z	A	90
0.633	1B8Z	B	90
0.922	1BAL	A	51
0.878	1BC0	A	319
0.881	1BC1	A	319
0.890	1BCY	A	319
0.654	1BEA	A	127
0.764	1BGC	A	174

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.573	1HSJ	A	487
0.581	1HSJ	B	487
0.710	1HSS	A	124
0.710	1HSS	B	124
0.710	1HSS	C	124
0.710	1HSS	D	124
0.667	1HUU	A	90
0.656	1HUU	B	90
0.678	1HUU	C	90
0.845	1HUZ	A	335
0.893	1HUZ	B	335
0.901	1HV4	A	141
0.918	1HV4	B	146
0.901	1HV4	C	141
0.918	1HV4	D	146
0.908	1HV4	E	141
0.918	1HV4	F	146
0.901	1HV4	G	141
0.918	1HV4	H	146
0.874	1HY1	A	468
0.874	1HY1	B	468
0.878	1HY1	C	468
0.880	1HY1	D	468
0.697	1HY8	A	76
0.869	1HZV	A	543
0.825	1I54	A	103
0.825	1I54	B	103
0.965	1I8O	A	114
0.901	1IBE	A	141
0.842	1IBE	B	146
0.769	1IBR	A	216
0.823	1IBR	B	462
0.769	1IBR	C	216
0.827	1IBR	D	462
0.905	1IDQ	A	609
0.721	1IDR	A	136
0.713	1IDR	B	136
0.478	1IF1	A	113
0.469	1IF1	B	113
0.867	1IG5	A	75
0.545	1IHF	A	99

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.726	1BI2	A	226
0.889	1BI2	B	226
0.831	1BMF	A	510
0.847	1BMF	B	510
0.849	1BMF	C	510
0.766	1BMF	D	482
0.790	1BMF	E	482
0.753	1BMF	F	482
0.960	1BMF	G	272
0.184	1BOC	A	76
0.176	1BOD	A	74
0.864	1BOF	A	353
0.961	1BQ0	A	103
0.815	1BQ7	A	189
0.820	1BQ7	B	189
0.825	1BQ7	C	189
0.825	1BQ7	D	189
0.820	1BQ7	E	189
0.820	1BQ7	F	189
0.779	1BS2	A	607
0.616	1BSH	A	138
0.950	1BSM	A	201
0.945	1BSM	B	201
0.583	1BUN	A	120
0.508	1BUN	B	61
0.936	1BUW	A	141
0.849	1BUW	B	146
0.915	1BUW	C	141
0.870	1BUW	D	146
0.848	1BVB	A	211
0.860	1BXI	A	86
0.560	1BXI	B	134
0.815	1BXR	A	1073
0.785	1BXR	B	382
0.809	1BXR	C	1073
0.785	1BXR	D	382
0.801	1BXR	E	1073
0.783	1BXR	F	382
0.808	1BXR	G	1073
0.783	1BXR	H	382
0.954	1BZ6	A	153

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.574	1IHF	B	94
0.741	1IHG	A	370
0.537	1IJL	A	123
0.520	1IJL	B	123
0.877	1IN4	A	334
0.868	1IN8	A	334
0.954	1IOP	A	153
0.900	1IQU	A	420
0.564	1IQV	A	218
0.875	1IRF	A	112
0.639	1IRX	A	523
0.648	1IRX	B	523
0.701	1IT1	A	107
0.617	1ITH	A	141
0.617	1ITH	B	141
0.976	1J47	A	85
0.694	1J77	A	209
0.879	1J7S	A	141
0.849	1J7S	B	146
0.879	1J7S	C	141
0.849	1J7S	D	146
0.856	1J7V	L	160
0.780	1J7V	R	214
0.574	1J9I	A	68
0.456	1J9I	B	68
0.785	1JDL	A	121
0.828	1JFK	A	134
0.824	1JFT	A	340
0.740	1JJS	A	50
0.941	1JP6	A	153
0.841	1JPJ	A	296
0.912	1JS4	A	605
0.912	1JS4	B	605
0.722	1JW2	A	72
0.942	1JW8	A	154
0.925	1JWG	A	147
0.932	1JWG	B	147
0.308	1JWG	C	13
0.308	1JWG	D	13
0.913	1K0N	A	241
0.909	1K0N	B	241

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.954	1BZP	A	153
0.889	1C3U	A	431
0.872	1C3U	B	431
0.972	1C3Y	A	108
0.901	1C40	A	141
0.890	1C40	B	146
0.687	1C52	A	131
0.933	1C6R	A	89
0.887	1C7C	A	283
0.836	1C7C	B	146
0.842	1C7C	D	146
0.884	1C7D	A	284
0.836	1C7D	B	146
0.836	1C7D	D	146
0.836	1C8J	A	414
0.829	1C8J	B	414
0.594	1C9L	A	357
0.583	1C9L	B	357
0.125	1C9L	C	8
0.125	1C9L	D	8
0.410	1CB1	A	78
0.872	1CCA	A	297
0.718	1CFA	A	71
0.750	1CFA	B	4
0.538	1CG7	A	93
0.909	1CH1	A	154
0.916	1CH5	A	154
0.916	1CH9	A	154
0.836	1CIP	A	353
0.922	1CMY	A	141
0.842	1CMY	B	146
0.922	1CMY	C	141
0.842	1CMY	D	146
0.695	1COR	A	82
0.812	1CP4	A	414
0.912	1CPE	A	296
0.906	1CPO	A	299
0.930	1CPQ	A	129
0.891	1CPR	A	129
0.903	1CPW	A	154
0.787	1CSW	A	108

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.909	1K0O	A	241
0.880	1K0O	B	241
0.486	1KB0	A	677
0.622	1KBU	A	349
0.622	1KBU	B	349
0.888	1KIB	A	89
0.888	1KIB	B	89
0.888	1KIB	C	89
0.888	1KIB	D	89
0.888	1KIB	E	89
0.888	1KIB	F	89
0.888	1KIB	G	89
0.888	1KIB	H	89
0.716	1KJS	A	74
0.692	1KKC	A	221
0.692	1KKC	B	221
0.692	1KKC	X	221
0.692	1KKC	Y	221
0.591	1KLP	A	115
0.850	1KNR	A	540
0.874	1KQ7	A	467
0.884	1KQ7	B	467
0.696	1KQV	A	79
0.885	1KRF	A	511
0.894	1KRF	B	511
0.772	1KSO	A	101
0.723	1KSO	B	101
0.769	1KSS	A	571
0.781	1KSU	A	571
0.793	1KSU	B	571
0.846	1KVW	A	123
0.708	1KW4	A	89
0.568	1KXU	A	333
0.855	1KYZ	A	365
0.882	1KYZ	C	365
0.907	1KYZ	E	365
0.468	1L1M	A	62
0.484	1L1M	B	62
0.892	1L1Y	A	678
0.872	1L1Y	B	678
0.881	1L1Y	C	678

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.806	1CSX	A	108
0.333	1CTA	A	36
0.333	1CTA	B	36
0.333	1CTD	A	36
0.333	1CTD	B	36
0.910	1CTJ	A	89
0.775	1CWW	A	102
0.889	1D8K	A	81
0.851	1D9G	A	121
0.810	1D9G	B	121
0.855	1DCY	A	124
0.700	1DIO	A	554
0.835	1DIO	B	224
0.888	1DIO	E	224
0.844	1DIO	G	173
0.706	1DIO	L	554
0.798	1DIO	M	173
0.879	1DKE	A	141
0.884	1DKE	B	146
0.887	1DKE	C	141
0.870	1DKE	D	146
0.556	1DN1	A	594
0.689	1DN1	B	267
0.935	1DO1	A	154
0.935	1DO3	A	154
0.593	1DP5	A	329
0.882	1DP5	B	68
0.745	1DQE	A	137
0.752	1DQE	B	137
0.628	1DT1	A	129
0.638	1DUX	C	94
0.681	1DUX	F	94
0.937	1DVH	A	79
0.946	1DW0	A	112
0.946	1DW0	B	112
0.946	1DW0	C	112
0.938	1DW3	A	112
0.938	1DW3	B	112
0.946	1DW3	C	112
0.882	1DWR	A	153
0.908	1DXU	A	141

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.894	1L1Y	D	678
0.883	1L1Y	E	678
0.886	1L1Y	F	678
0.575	1L5J	A	865
0.570	1L5J	B	865
0.686	1LBG	A	360
0.686	1LBG	B	360
0.686	1LBG	C	360
0.686	1LBG	D	360
0.643	1LBU	A	213
0.784	1LCC	A	51
0.894	1LE7	A	123
0.894	1LE7	B	123
0.915	1LH6	A	153
0.915	1LH7	A	153
0.954	1LHS	A	153
0.827	1LJP	A	98
0.796	1LJP	B	98
0.865	1LJR	A	244
0.861	1LJR	B	244
0.825	1LLP	A	343
0.831	1LY9	A	343
0.828	1LY9	B	343
0.879	1M1Q	A	91
0.879	1M1R	A	91
0.858	1MFR	A	176
0.858	1MFR	B	176
0.858	1MFR	C	176
0.858	1MFR	D	176
0.858	1MFR	E	176
0.858	1MFR	F	176
0.858	1MFR	G	176
0.858	1MFR	H	176
0.858	1MFR	I	176
0.858	1MFR	J	176
0.858	1MFR	K	176
0.858	1MFR	L	176
0.858	1MFR	M	176
0.858	1MFR	N	176
0.858	1MFR	O	176
0.858	1MFR	P	176

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.842	1DXU	B	146
0.908	1DXU	C	141
0.842	1DXU	D	146
0.848	1DZ9	A	414
0.855	1DZ9	B	414
0.727	1E0E	A	55
0.655	1E0E	B	55
0.724	1E2A	A	105
0.733	1E2A	B	105
0.733	1E2A	C	105
0.840	1E2R	A	567
0.825	1E2R	B	567
0.835	1E3H	A	757
0.900	1E3O	C	160
0.857	1E7P	A	656
0.766	1E7P	B	239
0.949	1E7P	C	256
0.857	1E7P	D	656
0.766	1E7P	E	239
0.949	1E7P	F	256
0.856	1E7P	G	655
0.766	1E7P	H	239
0.949	1E7P	I	254
0.856	1E7P	J	655
0.766	1E7P	K	239
0.949	1E7P	L	254
0.949	1ECA	A	136
0.949	1ECD	A	136
0.821	1EDI	A	56
0.664	1EEM	A	241
0.754	1EFA	A	333
0.766	1EFA	B	333
0.796	1EFA	C	333
0.857	1EGM	A	554
0.938	1EGM	B	224
0.879	1EGM	E	224
0.896	1EGM	G	173
0.852	1EGM	L	554
0.890	1EGM	M	173
0.854	1EGY	A	403

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.858	1MFR	Q	176
0.858	1MFR	R	176
0.858	1MFR	S	176
0.858	1MFR	T	176
0.858	1MFR	U	176
0.858	1MFR	V	176
0.858	1MFR	W	176
0.858	1MFR	X	176
0.971	1MJM	A	104
0.981	1MJM	B	104
0.909	1MLG	A	154
0.909	1MLH	A	154
0.909	1MLU	A	154
0.824	1MMM	A	205
0.839	1MMM	B	205
0.882	1MNH	A	153
0.882	1MNI	A	153
0.882	1MNI	B	153
0.971	1MRR	A	375
0.971	1MRR	B	375
0.909	1MTK	A	154
0.935	1MZM	A	93
0.890	1N90	A	543
0.888	1N90	B	543
0.837	1NBM	A	510
0.837	1NBM	B	510
0.825	1NBM	C	510
0.783	1NBM	D	480
0.783	1NBM	E	480
0.796	1NBM	F	480
0.960	1NBM	G	272
0.914	1NCY	A	162
0.953	1NFO	A	191
0.932	1OCC	A	514
0.890	1OCC	B	227
0.989	1OCC	C	261
0.905	1OCC	D	147
0.945	1OCC	E	109
0.857	1OCC	F	98
0.976	1OCC	G	84

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.632	1EH2	A	106
0.856	1EHF	A	403
0.598	1EJ5	A	107
0.860	1EMV	A	86
0.515	1EMV	B	134
0.849	1EN4	A	205
0.849	1EN4	B	205
0.893	1EN4	C	205
0.863	1EN4	D	205
0.859	1EN6	A	205
0.873	1EN6	B	205
0.888	1EN6	C	205
0.893	1EN6	D	205
0.971	1ENJ	A	138
0.907	1EQZ	A	129
0.952	1EQZ	B	126
0.949	1EQZ	C	136
0.864	1EQZ	D	103
0.907	1EQZ	E	129
0.952	1EQZ	F	126
0.949	1EQZ	G	136
0.864	1EQZ	H	103
0.650	1ERD	A	40
0.552	1ETE	A	134
0.552	1ETE	B	134
0.545	1ETE	C	134
0.537	1ETE	D	134
0.908	1ETV	A	98
0.898	1ETV	B	98
0.898	1ETW	A	98
0.806	1ETW	B	98
0.822	1EWQ	A	765
0.830	1EWQ	B	765
0.824	1EXI	A	278
0.915	1F0M	A	82
0.906	1F5O	A	149
0.906	1F5O	B	149
0.906	1F5O	C	149
0.906	1F5O	D	149
0.906	1F5O	E	149
0.919	1F5O	F	149

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.965	1OCC	H	85
0.986	1OCC	I	73
0.949	1OCC	J	59
1.000	1OCC	K	56
0.936	1OCC	L	47
0.935	1OCC	M	46
0.932	1OCC	N	514
0.890	1OCC	O	227
0.989	1OCC	P	261
0.905	1OCC	Q	147
0.945	1OCC	R	109
0.857	1OCC	S	98
0.976	1OCC	T	84
0.965	1OCC	U	85
0.986	1OCC	V	73
0.949	1OCC	W	59
1.000	1OCC	X	56
0.936	1OCC	Y	47
0.935	1OCC	Z	46
0.621	1OM2	A	95
0.636	1OM2	B	11
0.894	1ORK	A	207
0.824	1OSA	A	148
0.586	1PGJ	A	478
0.596	1PGJ	B	478
0.805	1PGT	A	210
0.819	1PGT	B	210
0.839	1PIR	A	124
0.847	1PIS	A	124
0.882	1PMB	A	153
0.882	1PMB	B	153
0.750	1PRU	A	56
0.662	1PS1	A	337
0.650	1PS1	B	337
0.404	1QDM	A	478
0.385	1QDM	B	478
0.383	1QDM	C	478
0.604	1QHB	A	596
0.599	1QHB	B	596
0.601	1QHB	C	596
0.589	1QHB	D	596

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.805	1FAG	A	471
0.811	1FAG	B	471
0.817	1FAG	C	471
0.809	1FAG	D	471
0.887	1FFY	A	917
0.778	1FHB	A	108
0.724	1FHE	A	217
0.734	1FHF	A	304
0.734	1FHF	B	304
0.734	1FHF	C	304
0.817	1FI3	A	82
0.824	1FJG	B	256
0.824	1FJG	C	239
0.828	1FJG	D	209
0.815	1FJG	E	162
0.832	1FJG	F	101
0.821	1FJG	G	156
0.862	1FJG	H	138
0.906	1FJG	I	128
0.810	1FJG	J	105
0.829	1FJG	K	129
0.837	1FJG	L	135
0.897	1FJG	M	126
0.754	1FJG	N	61
0.921	1FJG	O	89
0.818	1FJG	P	88
0.848	1FJG	Q	105
0.898	1FJG	R	88
0.828	1FJG	S	93
0.943	1FJG	T	106
0.923	1FJG	V	26
0.935	1FK5	A	93
0.935	1FK6	A	93
0.617	1FNN	A	389
0.622	1FNN	B	389
0.882	1FOX	A	76
0.855	1FOY	A	76
0.511	1FP1	D	372
0.690	1FPO	A	171
0.737	1FPO	B	171
0.760	1FPO	C	171

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.601	1QHB	E	596
0.591	1QHB	F	596
0.922	1QI8	A	141
0.842	1QI8	B	146
0.922	1QI8	C	141
0.842	1QI8	D	146
0.629	1QI9	A	556
0.640	1QI9	B	556
0.869	1QJB	A	245
0.849	1QJB	B	245
0.375	1QJB	Q	8
0.429	1QJB	S	7
0.970	1QL4	A	99
0.939	1QL4	B	99
0.939	1QL4	C	99
0.970	1QL4	D	99
0.902	1QLK	A	92
0.902	1QLK	B	92
0.720	1QN2	A	100
0.640	1QN2	B	100
0.650	1QN2	C	100
0.618	1QNN	A	191
0.613	1QNN	B	191
0.623	1QNN	C	191
0.618	1QNN	D	191
0.877	1QPI	A	203
0.801	1QPW	A	141
0.856	1QPW	B	146
0.830	1QPW	C	141
0.849	1QPW	D	146
0.812	1QQB	A	340
0.768	1QUU	A	250
0.928	1R2F	A	319
0.903	1R2F	B	319
0.698	1R63	A	63
0.965	1RCE	A	173
0.714	1RFB	A	119
0.714	1RFB	B	119
0.813	1RMI	A	160
0.580	1RMV	A	157
0.811	1ROM	A	403

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.524	1FPQ	A	372
0.938	1FT6	A	211
0.938	1FUR	A	467
0.949	1FUR	B	467
0.730	1FVK	A	189
0.730	1FVK	B	189
0.926	1FYZ	A	527
0.924	1FYZ	B	527
0.933	1FYZ	C	389
0.933	1FYZ	D	389
0.953	1FYZ	E	170
0.953	1FYZ	F	170
0.634	1FZP	B	123
0.659	1FZP	D	123
0.875	1G2Y	A	32
0.844	1G2Y	B	32
0.875	1G2Y	C	32
0.844	1G2Y	D	32
0.656	1G3N	A	326
0.768	1G3N	B	168
0.673	1G3N	C	257
0.672	1G3N	E	326
0.768	1G3N	F	168
0.665	1G3N	G	257
0.509	1G8E	A	116
0.509	1G8E	B	116
0.574	1GAX	A	862
0.580	1GAX	B	862
0.879	1GC7	A	297
0.753	1GDV	A	85
0.861	1GEJ	A	403
0.817	1GEK	A	415
0.629	1GGW	A	140
0.770	1GGZ	A	148
0.875	1GIA	A	353
0.882	1GJN	A	153
0.923	1GJS	A	65
0.874	1GKJ	A	509
0.679	1GKS	A	78
0.837	1GLP	A	209
0.837	1GLP	B	209

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.642	1RSS	A	151
0.901	1SDK	A	141
0.897	1SDK	B	146
0.957	1SDK	C	141
0.890	1SDK	D	146
0.889	1SKT	A	90
0.809	1SPY	A	89
0.930	1T1N	A	143
0.870	1T1N	B	146
0.925	1TN4	A	159
0.900	1TNP	A	90
0.956	1UBY	A	367
0.813	1UNE	A	123
0.908	1UXD	A	65
0.933	1VNI	A	609
0.954	1VXA	A	153
0.954	1VXB	A	153
0.954	1VXD	A	153
0.766	1WJA	A	47
0.766	1WJA	B	47
0.869	1WLA	A	153
0.847	1XGN	A	295
0.851	1XGN	B	295
0.831	1XGO	A	295
0.869	1YMC	A	153
0.833	1ZAC	A	90
0.878	1ZQA	A	335
0.878	1ZQB	A	335
0.878	1ZQK	A	335
0.812	2CPP	A	414
0.911	2CTS	A	437
0.568	2CY3	A	118
0.922	2CYK	A	129
0.813	2CYM	A	107
0.721	2DTR	A	226
0.869	2EZC	A	259
0.842	2FOW	A	76
0.870	2HBI	A	146
0.870	2HBI	B	146
0.797	2HMX	A	133
0.940	2LFB	A	100

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
1.000	1GLP	I	1
1.000	1GLP	J	1
0.927	1GNC	A	178
0.773	1GNW	A	211
0.773	1GNW	B	211
0.841	1GO5	A	69
0.929	1GP2	A	353
0.885	1GP2	B	340
0.915	1GP2	G	71
0.780	1GSS	A	209
0.780	1GSS	B	209
0.667	1GSS	I	3
0.667	1GSS	J	3
0.856	1GTI	A	209
0.842	1GTI	B	209
0.861	1GTI	C	209
0.861	1GTI	D	209
0.856	1GTI	E	209
0.856	1GTI	F	209
0.860	1GUM	A	222
0.860	1GUM	B	222
0.860	1GUM	C	222
0.860	1GUM	D	222
0.860	1GUM	E	222
0.860	1GUM	F	222
0.860	1GUM	G	222
0.860	1GUM	H	222
0.683	1GWC	A	230
0.678	1GWC	B	230
0.687	1GWC	C	230
0.841	1GWP	A	151
0.517	1GXD	A	631
0.520	1GXD	B	631
0.423	1GXD	C	194
0.387	1GXD	D	194
0.941	1H0N	A	390
0.491	1H54	A	754
0.504	1H54	B	754
0.932	1H55	A	308
0.938	1H5D	A	308
0.827	1H8E	A	510

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.915	2LH1	A	153
0.941	2LYN	A	136
0.949	2LYN	B	136
0.919	2LYN	C	136
0.912	2LYN	D	136
0.916	2MBW	A	154
0.327	2MM1	A	153
0.954	2MYC	A	153
0.851	2PGH	A	141
0.842	2PGH	B	146
0.851	2PGH	C	141
0.842	2PGH	D	146
0.809	2PUE	A	340
0.803	2PUG	A	340
0.837	2SBL	B	839
0.909	2SPL	A	154
0.914	2SQC	A	631
0.914	2SQC	B	631
0.738	2WRP	R	107
0.000	2WRP	W	1
0.682	2YCC	A	107
0.812	3CP4	A	414
0.690	3CRD	A	100
0.853	3FYG	A	217
0.829	3FYG	B	217
0.901	3HHB	A	141
0.842	3HHB	B	146
0.620	3HSF	A	92
0.863	3SDH	A	146
0.863	3SDH	B	146
0.919	3SQC	A	631
0.918	3SQC	B	631
0.919	3SQC	C	631
0.915	4CTS	A	437
0.915	4CTS	B	437
0.912	4GST	A	217
0.931	4GST	B	217
0.718	4P2P	A	124
0.912	4TF4	A	605
0.919	4TF4	B	605
0.812	5CPP	A	414

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.827	1H8E	B	510
0.833	1H8E	C	510
0.739	1H8E	D	482
0.755	1H8E	E	482
0.755	1H8E	F	482
0.680	1H8E	G	272
0.459	1H8E	H	146
0.960	1H8E	I	50
0.812	1H8H	A	510
0.812	1H8H	B	510
0.835	1H8H	C	510
0.770	1H8H	D	482
0.747	1H8H	E	482
0.761	1H8H	F	482
0.967	1H8H	G	272
0.957	1HAB	A	141
0.897	1HAB	B	146
0.957	1HAB	C	141
0.918	1HAB	D	146
0.730	1HD6	A	37
0.943	1HDB	A	141
0.918	1HDB	B	146
0.936	1HDB	C	141
0.918	1HDB	D	146
0.934	1HF8	A	289

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.924	5CTS	A	433
0.923	5EAS	A	548
0.912	5EAU	A	548
0.924	6CTS	A	433
0.707	6PAX	A	133
0.875	7ICK	A	335
0.875	7ICV	A	335
0.949	7PRC	C	336
0.833	7PRC	H	258
0.963	7PRC	L	273
0.960	7PRC	M	323
0.878	8ICE	A	335
0.878	8ICF	A	335
0.878	8ICJ	A	335
0.878	8ICL	A	335
0.878	8ICN	A	335
0.878	8ICQ	A	335
0.878	9ICC	A	335
0.878	9ICF	A	335
0.878	9ICG	A	335
0.878	9ICK	A	335
0.878	9ICL	A	335
0.878	9ICM	A	335
0.878	9ICV	A	335

APPENDIX P

Prediction accuracies for each chain of each protein in all- β class in trial set

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.658	1a1d	A	146
0.880	1a1x	A	108
0.610	1a34	A	159
1.000	1a34	B	10
1.000	1a34	C	10
1.000	1a34	D	1
0.826	1a39	A	402
0.676	1a3e	H	259
1.000	1a3e	I	19
0.889	1a3e	L	36
0.416	1a3y	A	149
0.430	1a3y	B	149
0.664	1a42	A	259
0.823	1a47	A	683
0.659	1a5f	H	217
0.564	1a5f	L	220
0.631	1a5h	A	252
0.639	1a5h	B	252
1.000	1a5h	C	7
1.000	1a5h	D	7
0.781	1a71	A	374
0.775	1a71	B	374
0.911	1a8m	A	157
0.911	1a8m	B	157
0.911	1a8m	C	157
0.586	1a9m	A	99
0.616	1a9m	B	99
0.444	1aaq	A	99
0.444	1aaq	B	99
0.851	1acy	H	221
0.963	1acy	L	215
1.000	1acy	P	24
0.719	1adb	A	374
0.719	1adb	B	374
0.668	1ae8	H	259
1.000	1ae8	I	10

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
1.000	1i50	J	70
0.925	1i50	K	120
1.000	1i50	L	70
0.923	1i6c	A	39
0.735	1i75	A	686
0.736	1i75	B	686
0.782	1i7u	A	275
0.770	1i7u	B	100
0.778	1i7u	C	9
0.782	1i7u	D	275
0.770	1i7u	E	100
0.778	1i7u	F	9
0.795	1i81	A	83
0.795	1i81	B	83
0.783	1i81	C	83
0.795	1i81	D	83
0.795	1i81	E	83
0.783	1i81	F	83
0.795	1i81	G	83
0.600	1i85	A	110
0.582	1i85	B	110
0.762	1i85	C	126
0.730	1i85	D	126
0.624	1i8d	A	213
0.620	1i8d	B	213
0.662	1i8d	C	213
0.571	1i92	A	91
1.000	1i94	A	1287
0.831	1i94	B	255
0.756	1i94	C	238
0.750	1i94	D	208
0.702	1i94	E	161
0.584	1i94	F	101
0.845	1i94	G	155
0.486	1i94	H	138
0.641	1i94	I	128

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.889	1ae8	L	36
0.748	1ael	A	131
0.611	1aex	A	149
0.404	1afv	A	151
0.404	1afv	B	151
0.536	1afv	H	220
0.568	1afv	K	220
0.493	1afv	L	217
0.512	1afv	M	217
0.866	1agb	A	276
0.848	1agb	B	99
1.000	1agb	C	8
0.848	1agf	A	276
0.848	1agf	B	99
0.500	1agf	C	8
1.000	1aif	A	215
0.950	1aif	B	218
0.950	1aif	H	218
1.000	1aif	L	215
0.897	1ajk	A	214
0.893	1ajk	B	214
0.874	1ajo	A	214
0.874	1ajo	B	214
0.626	1ajx	A	99
0.636	1ajx	B	99
0.667	1anu	A	138
0.864	1aom	A	567
0.864	1aom	B	567
0.885	1aoq	A	567
0.873	1aoq	B	567
0.854	1aqd	A	192
0.783	1aqd	B	198
0.067	1aqd	C	15
0.865	1aqd	D	192
0.788	1aqd	E	198
0.067	1aqd	F	15
0.865	1aqd	G	192
0.783	1aqd	H	198
0.067	1aqd	I	15
0.865	1aqd	J	192
0.783	1aqd	K	198

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.692	1i94	J	104
0.711	1i94	K	128
0.786	1i94	L	131
0.832	1i94	M	125
0.650	1i94	N	60
0.830	1i94	O	88
0.750	1i94	P	88
0.663	1i94	Q	104
0.862	1i94	R	87
0.652	1i94	S	92
0.724	1i94	T	105
0.808	1i94	U	26
1.000	1i95	A	1287
0.706	1i95	B	255
0.794	1i95	C	238
0.726	1i95	D	208
0.745	1i95	E	161
0.782	1i95	F	101
0.568	1i95	G	155
0.949	1i95	H	138
0.898	1i95	I	128
0.808	1i95	J	104
0.789	1i95	K	128
0.962	1i95	L	131
0.680	1i95	M	125
0.900	1i95	N	60
0.352	1i95	O	88
0.864	1i95	P	88
0.788	1i95	Q	104
0.851	1i95	R	87
0.924	1i95	S	92
0.533	1i95	T	105
0.923	1i95	U	26
0.410	1i9b	A	217
0.410	1i9b	B	217
0.410	1i9b	C	217
0.410	1i9b	D	217
0.410	1i9b	E	217
0.600	1i9e	A	115
0.629	1i9o	A	259
0.629	1i9q	A	259

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.067	1aqd	L	15
0.679	1ar2	A	109
0.688	1as7	A	343
0.682	1as7	B	343
0.691	1as7	C	343
0.844	1ava	A	403
0.844	1ava	B	403
0.718	1ava	C	181
0.696	1ava	D	181
0.914	1avd	A	128
0.914	1avd	B	128
0.351	1awj	A	77
0.854	1awr	A	164
0.835	1awr	B	164
0.854	1awr	C	164
0.854	1awr	D	164
0.854	1awr	E	164
0.835	1awr	F	164
1.000	1awr	G	6
1.000	1awr	H	6
1.000	1awr	I	6
1.000	1awr	J	6
1.000	1awr	K	6
1.000	1awr	L	6
0.866	1awt	A	164
0.841	1awt	B	164
0.866	1awt	C	164
0.835	1awt	D	164
0.872	1awt	E	164
0.866	1awt	F	164
1.000	1awt	G	6
1.000	1awt	H	6
1.000	1awt	I	6
1.000	1awt	J	6
1.000	1awt	K	6
1.000	1awt	L	6
0.866	1awv	A	164
0.835	1awv	B	164
0.854	1awv	C	164
0.854	1awv	D	164
0.854	1awv	E	164

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.464	1iaz	A	179
0.469	1iaz	B	179
0.755	1ib4	A	339
0.749	1ib4	B	339
0.742	1ibf	A	151
0.788	1ibh	A	151
0.640	1ic7	H	114
0.542	1ic7	L	107
0.589	1ic7	Y	129
0.824	1icm	A	131
0.780	1ict	A	127
0.764	1ict	B	127
0.780	1ict	C	127
0.787	1ict	D	127
0.780	1ict	E	127
0.787	1ict	F	127
0.772	1ict	G	127
0.780	1ict	H	127
0.859	1ida	A	99
0.859	1ida	B	99
1.000	1ida	I	5
0.449	1le5	A	107
0.771	1lfb	A	131
0.945	1lfh	H	220
1.000	1lfh	L	218
1.000	1lfh	P	7
0.955	1lgi	H	221
0.955	1lgi	J	221
0.995	1lgi	L	219
0.995	1lgi	M	219
0.927	1lgi	H	218
1.000	1lgi	L	219
0.759	1lhg	A	370
0.886	1lke	A	184
0.584	1lko	P	178
0.855	1lkq	A	613
0.825	1lle	A	821
0.794	1lif	A	141
0.401	1lit	A	152
0.401	1lit	B	152
0.727	1llu	A	128

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.841	1awv	F	164
1.000	1awv	G	6
1.000	1awv	H	6
1.000	1awv	I	6
1.000	1awv	J	6
1.000	1awv	K	6
1.000	1awv	L	6
0.828	1azn	A	128
0.828	1azn	B	128
0.828	1azn	C	128
0.828	1azn	D	128
0.828	1azu	A	128
0.593	1b0w	A	108
0.593	1b0w	B	108
0.593	1b0w	C	108
0.735	1b11	A	113
0.891	1b2y	A	496
0.629	1b3i	A	97
0.686	1b4j	H	220
0.659	1b4j	L	214
0.694	1b4m	A	134
0.874	1b56	A	135
0.566	1b6d	A	212
0.571	1b6d	B	212
0.871	1b7a	A	186
0.844	1b7a	B	186
0.667	1b7k	A	66
0.855	1b90	A	516
0.789	1b9m	A	265
0.815	1b9m	B	265
0.714	1baw	A	105
0.714	1baw	B	105
0.714	1baw	C	105
0.917	1bbr	E	109
1.000	1bbr	F	10
1.000	1bbr	G	10
0.980	1bbr	H	150
1.000	1bbr	I	10
1.000	1bbr	J	49
0.954	1bbr	K	259
1.000	1bbr	L	49

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.656	1ilu	B	128
0.680	1ilu	C	128
0.680	1ilu	D	128
0.680	1ilu	E	128
0.680	1ilu	F	128
0.727	1ilu	G	128
0.680	1ilu	H	128
0.656	1ilu	I	128
0.680	1ilu	K	128
0.680	1ilu	L	128
0.680	1ilu	M	128
0.768	1im9	A	276
0.770	1im9	B	100
0.667	1im9	C	9
0.724	1im9	D	225
0.775	1im9	E	276
0.760	1im9	F	100
0.667	1im9	G	9
0.951	1ind	H	226
0.930	1ind	L	215
0.712	1iob	A	153
0.736	1ipj	A	416
0.736	1ipj	B	416
0.726	1ipj	C	416
0.888	1ivl	A	107
0.888	1ivl	B	107
0.867	1ivv	A	638
0.857	1ivv	B	638
0.799	1j4s	A	149
0.846	1j4s	B	149
0.846	1j4s	C	149
0.846	1j4s	D	149
0.826	1j58	A	385
0.773	1j88	A	172
0.773	1j88	B	172
0.773	1j88	C	172
0.773	1j88	D	172
0.773	1j88	E	172
0.617	1j8k	A	94
0.770	1jbh	A	135
0.667	1bj	A	186

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
1.000	1bbr	M	49
0.954	1bbr	N	259
0.653	1bcd	A	259
0.855	1bci	A	138
0.861	1bck	A	165
1.000	1bck	C	11
0.925	1bd9	A	187
0.930	1bd9	B	187
0.727	1bdr	A	99
0.909	1bdr	B	99
0.296	1bdx	A	203
0.296	1bdx	B	203
0.296	1bdx	C	203
0.296	1bdx	D	203
1.000	1bdx	J	16
1.000	1bdx	K	16
1.000	1bdx	L	16
1.000	1bdx	M	16
0.520	1bdy	A	123
0.520	1bdy	B	123
0.641	1bf2	A	750
0.603	1bfo	A	214
0.597	1bfo	B	216
0.575	1bfo	C	214
0.588	1bfo	D	216
0.575	1bfo	E	214
0.588	1bfo	F	216
0.589	1bfo	G	214
0.602	1bfo	H	216
0.897	1bgy	A	446
0.907	1bgy	B	439
0.918	1bgy	C	379
0.925	1bgy	D	241
0.816	1bgy	E	196
0.882	1bgy	F	110
0.975	1bgy	G	81
0.833	1bgy	H	78
0.936	1bgy	I	78
0.935	1bgy	J	62
1.000	1bgy	K	56
0.908	1bgy	M	446

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.871	1jcz	A	263
0.875	1jcz	B	263
0.869	1jdd	A	429
0.835	1jec	A	510
0.587	1jer	A	138
0.805	1jf6	A	585
0.791	1jf6	B	585
0.658	1jgx	H	260
0.726	1jgx	L	281
0.792	1jgx	M	307
0.665	1jgy	H	260
0.737	1jgy	L	281
0.785	1jgy	M	307
0.658	1jgz	H	260
0.730	1jgz	L	281
0.785	1jgz	M	307
0.537	1jhk	H	218
0.453	1jhk	L	214
0.804	1jht	A	275
0.770	1jht	B	100
0.778	1jht	C	9
0.802	1ji2	A	585
0.798	1ji2	B	585
0.840	1jjc	A	350
0.563	1jjc	B	785
0.750	1jnj	A	100
0.569	1jo8	A	58
0.530	1jok	A	149
0.550	1joq	A	149
0.511	1jpt	H	225
0.523	1jpt	L	214
0.447	1jri	A	85
0.459	1jri	B	85
0.482	1jri	C	85
0.471	1jri	D	85
0.471	1jri	E	85
0.459	1jri	F	85
0.482	1jri	G	85
0.471	1jri	H	85
0.482	1jri	I	85
0.506	1jri	J	85

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.923	1bgy	N	439
0.910	1bgy	O	379
0.925	1bgy	P	241
0.842	1bgy	Q	196
0.909	1bgy	R	110
0.975	1bgy	S	81
0.833	1bgy	T	78
0.936	1bgy	U	78
0.935	1bgy	V	62
1.000	1bgy	W	56
0.743	1bl3	A	226
0.876	1bl3	B	226
0.653	1bic	A	259
0.748	1bla	A	155
0.921	1blz	A	331
0.729	1bmw	A	96
0.526	1bn8	A	420
0.672	1bnn	A	259
0.660	1bnt	A	259
0.625	1bpv	A	112
0.911	1bqk	A	124
0.550	1bqs	A	209
0.877	1bso	A	162
0.840	1bsq	A	162
0.846	1bsy	A	162
0.830	1btn	A	106
0.762	1bto	A	374
0.762	1bto	B	374
0.757	1bto	C	374
0.778	1bto	D	374
0.667	1bv8	A	138
0.726	1bvl	A	117
0.546	1bvl	B	108
0.641	1bvl	C	117
0.583	1bvl	D	108
0.919	1bvz	A	185
0.920	1bw4	A	125
0.858	1bwn	A	169
0.858	1bwn	B	169
0.737	1byo	A	99
0.677	1byo	B	99

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.518	1jri	K	85
0.482	1jri	L	85
0.459	1jri	M	85
0.471	1jri	N	85
0.748	1jsn	A	325
0.915	1jsn	B	176
0.696	1jtp	A	148
0.709	1jtp	B	148
0.535	1jtp	L	129
0.550	1jtp	M	129
0.928	1ju3	A	583
0.742	1jyt	A	163
0.758	1jz1	I	1023
0.757	1jz1	J	1023
0.758	1jz1	K	1023
0.755	1jz1	L	1023
0.758	1jz1	M	1023
0.757	1jz1	N	1023
0.758	1jz1	O	1023
0.757	1jz1	P	1023
0.742	1jz3	A	1023
0.741	1jz3	B	1023
0.742	1jz3	C	1023
0.740	1jz3	D	1023
0.742	1jz4	A	1023
0.737	1jz4	B	1023
0.741	1jz4	C	1023
0.733	1jz4	D	1023
0.602	1jzf	A	128
0.629	1k5j	A	124
0.589	1k5j	B	124
0.621	1k5j	C	124
0.613	1k5j	D	124
0.613	1k5j	E	124
0.638	1k6n	H	260
0.747	1k6n	L	281
0.796	1k6n	M	314
0.587	1k9a	A	450
0.596	1k9a	B	450
0.580	1k9a	C	450
0.569	1k9a	D	450

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.945	1byt	A	857
0.780	1bze	A	127
0.772	1bze	B	127
0.542	1c08	A	107
0.640	1c08	B	114
0.597	1c08	C	129
0.890	1c5m	D	255
0.771	1c5m	F	96
0.841	1c9i	A	359
0.833	1c9i	B	359
1.000	1c9i	C	9
1.000	1c9i	D	9
0.625	1c9i	A	357
0.602	1c9i	B	357
0.875	1c9i	C	8
0.875	1c9i	D	8
0.653	1cah	A	259
0.649	1cal	A	259
0.894	1cb4	A	151
0.881	1cb4	B	151
0.832	1cbs	A	137
0.448	1cd1	A	315
0.808	1cd1	B	99
0.438	1cd1	C	315
0.808	1cd1	D	99
0.358	1cdi	A	179
0.652	1cel	A	434
0.652	1cel	B	434
0.571	1cgl	E	245
0.446	1cgl	I	56
0.571	1cgj	E	245
0.446	1cgj	I	56
0.829	1cgt	A	684
0.892	1cil	A	259
0.892	1cin	A	259
0.933	1ckm	A	330
0.921	1ckm	B	330
0.464	1clo	H	222
0.465	1clo	L	213
0.533	1cms	A	323
0.703	1cnc	A	259

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.573	1k9a	E	450
0.609	1k9a	F	450
0.616	1kb0	A	677
0.827	1kbb	A	496
0.870	1kbp	A	432
0.870	1kbp	B	432
0.870	1kbp	C	432
0.870	1kbp	D	432
0.581	1kc5	H	217
0.505	1kc5	L	214
1.000	1kc5	P	15
0.839	1kcc	A	335
0.697	1kck	A	686
0.575	1kd6	A	179
0.946	1kdc	A	149
0.698	1ken	A	328
0.823	1ken	B	175
0.710	1ken	C	328
0.823	1ken	D	175
0.720	1ken	E	328
0.823	1ken	F	175
0.692	1ken	H	221
0.681	1ken	L	213
0.670	1ken	T	221
0.681	1ken	U	213
0.800	1kgl	A	135
0.801	1kj8	A	391
0.808	1kj8	B	391
0.556	1kjg	A	99
0.545	1kjg	B	99
0.600	1kjg	P	10
0.801	1kjj	A	391
0.808	1kjj	B	391
0.893	1kk3	A	410
0.532	1kn4	H	222
0.470	1kn4	L	219
0.896	1kq2	A	77
0.896	1kq2	B	77
0.883	1kq2	H	77
0.870	1kq2	I	77
0.909	1kq2	K	77

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.915	1cne	A	270
0.680	1cng	A	259
0.683	1cnj	A	259
0.465	1cpi	A	99
0.465	1cpi	B	99
0.778	1cpi	C	9
0.890	1cr7	A	236
0.890	1cr7	B	236
0.890	1cr7	C	236
0.881	1cr7	D	236
0.890	1cr7	E	236
0.890	1cr7	F	236
0.894	1cr7	G	236
0.890	1cr7	H	236
0.804	1crm	A	260
0.872	1cru	A	454
0.866	1cru	B	454
0.732	1csk	A	71
0.732	1csk	B	71
0.732	1csk	C	71
0.732	1csk	D	71
0.866	1csq	A	67
0.861	1cwk	A	165
1.000	1cwk	C	11
0.861	1cwo	A	165
1.000	1cwo	C	11
0.679	1cwp	A	190
0.684	1cwp	B	190
0.658	1cwp	C	190
1.000	1cwp	D	4
1.000	1cwp	E	2
1.000	1cwp	F	4
0.749	1cxe	A	686
0.754	1cxh	A	686
0.796	1czm	A	260
0.816	1d0z	A	761
0.823	1d5g	A	96
0.600	1d5g	B	15
0.867	1d5z	A	181
0.781	1d5z	B	192
0.895	1d5z	C	239

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.896	1kq2	M	77
1.000	1kq2	R	7
0.747	1ksj	A	186
0.829	1ksj	B	152
0.842	1ku8	A	133
0.500	1ku8	B	18
0.842	1ku8	C	133
0.556	1ku8	D	18
0.842	1ku8	E	133
0.500	1ku8	F	18
0.850	1ku8	G	133
0.500	1ku8	H	18
0.842	1kuj	A	133
0.500	1kuj	B	18
0.842	1kuj	C	133
0.556	1kuj	D	18
0.842	1kuj	E	133
0.500	1kuj	F	18
0.850	1kuj	G	133
0.500	1kuj	H	18
0.736	1kvp	A	497
0.576	1kzk	A	99
0.525	1kzk	B	99
0.944	1l0s	A	90
0.944	1l0s	B	90
0.944	1l0s	C	90
0.944	1l0s	D	90
0.566	1lcy	A	325
0.858	1lek	A	274
0.737	1lek	B	99
1.000	1lek	P	8
0.630	1lem	A	181
0.673	1lem	B	52
0.912	1lf7	A	182
0.627	1lg6	A	260
0.885	1lib	A	131
0.885	1lid	A	131
0.878	1lmk	A	238
0.878	1lmk	C	238
0.878	1lmk	E	238
0.878	1lmk	G	238

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.833	1d5z	D	6
0.809	1d6k	A	94
1.000	1d6k	B	37
0.507	1d6v	H	221
0.573	1d6v	L	211
0.585	1d6w	A	287
1.000	1d6w	I	10
0.730	1d7f	A	686
0.743	1d7f	B	686
0.768	1d8t	A	393
0.774	1d8t	B	393
0.630	1dar	A	691
0.779	1dc9	A	131
0.504	1df0	A	700
0.489	1df0	B	184
0.778	1dfa	A	454
0.753	1dhx	A	967
0.625	1dl7	H	112
0.505	1dl7	L	109
0.508	1dlf	H	124
0.531	1dlf	L	113
0.695	1dlp	A	236
0.742	1dlp	B	236
0.678	1dlp	C	236
0.720	1dlp	D	236
0.669	1dlp	E	236
0.725	1dlp	F	236
0.926	1dlr	A	311
0.926	1dlr	B	311
0.513	1dms	A	781
0.787	1dnv	A	437
0.919	1dqk	A	124
0.919	1dqk	B	124
0.919	1dqk	C	124
0.919	1dqk	D	124
0.711	1dqp	A	135
0.607	1dqt	A	117
0.624	1dqt	B	117
0.607	1dqt	C	117
0.607	1dqt	D	117
0.896	1duc	A	134

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.532	1lo4	H	220
0.415	1lo4	L	217
0.461	1lox	A	662
0.859	1lr5	A	163
0.847	1lr5	B	163
0.840	1lr5	C	163
0.840	1lr5	D	163
0.777	1lt4	A	247
0.854	1lt4	D	103
0.864	1lt4	E	103
0.864	1lt4	F	103
0.864	1lt4	G	103
0.864	1lt4	H	103
0.810	1lul	A	253
0.842	1lul	B	253
0.842	1lul	C	253
0.810	1lul	D	253
0.842	1lul	E	253
0.818	1lul	F	253
0.690	1lv1	A	203
0.656	1lve	A	122
0.381	1m4h	A	391
0.384	1m4h	B	391
0.375	1m4h	C	8
0.375	1m4h	D	8
0.814	1mab	A	510
0.647	1mab	B	479
0.841	1mab	G	270
0.384	1mcl	A	216
0.421	1mcl	B	216
1.000	1mcl	P	2
0.384	1mcr	A	216
0.421	1mcr	B	216
1.000	1mcr	P	2
0.481	1mdc	A	131
0.567	1mim	H	215
0.562	1mim	L	210
0.519	1mlb	A	214
0.670	1mlb	B	218
0.850	1mmd	A	762
0.374	1mpp	A	361

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.770	1dud	A	152
0.770	1dup	A	152
0.778	1duy	A	275
0.770	1duy	B	100
1.000	1duy	C	8
0.782	1duy	D	275
0.770	1duy	E	100
1.000	1duy	F	8
0.758	1dvs	A	124
0.758	1dvs	B	124
0.766	1dvz	A	124
0.766	1dvz	B	124
0.936	1dyw	A	173
0.845	1dz0	A	129
0.815	1e00	A	157
0.809	1e00	B	157
0.833	1e0f	A	36
0.833	1e0f	B	36
0.829	1e0f	C	35
0.653	1e0f	D	259
0.629	1e0f	E	259
0.633	1e0f	F	259
0.702	1e0f	I	57
0.636	1e0f	J	55
0.632	1e0f	K	57
0.541	1e0l	A	37
0.837	1e1r	A	510
0.837	1e1r	B	510
0.841	1e1r	C	510
0.795	1e1r	D	482
0.801	1e1r	E	482
0.826	1e1r	F	482
0.978	1e1r	G	272
0.869	1e2r	A	567
0.854	1e2r	B	567
0.469	1e32	A	458
0.387	1e3p	A	757
0.782	1e57	A	188
0.856	1e57	B	188
0.798	1e57	C	188
0.664	1e67	A	128

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.729	1nbc	A	155
0.729	1nbc	B	155
0.873	1nch	A	110
0.836	1nch	B	110
0.726	1ncj	A	215
0.500	1neu	A	124
0.794	1nib	A	340
0.803	1nib	B	340
0.800	1nib	C	340
0.818	1nic	A	340
0.701	1nkr	A	201
0.719	1nlo	C	64
1.000	1nlo	N	10
0.750	1nlp	C	64
0.889	1nlp	N	9
0.475	1nmb	H	122
0.578	1nmb	L	109
0.581	1nmb	N	470
0.866	1nmg	A	67
0.587	1nnp	H	259
1.000	1nnp	L	36
0.957	1nnp	R	23
0.571	1nrq	H	259
1.000	1nrq	L	36
1.000	1nrq	R	22
0.587	1nrs	H	259
1.000	1nrs	I	12
1.000	1nrs	L	36
1.000	1nrs	R	4
0.809	1oac	A	727
0.809	1oac	B	727
0.525	1odx	A	99
0.556	1odx	B	99
1.000	1odx	C	5
0.676	1okm	A	259
0.710	1ose	A	496
0.826	1otc	A	495
0.792	1otc	B	260
1.000	1otc	D	12
0.872	1ovw	A	398
0.852	1ovw	B	398

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.609	1e67	B	128
0.625	1e67	C	128
0.625	1e67	D	128
0.796	1e9o	A	152
0.757	1e9o	B	152
0.868	1e9p	A	151
0.854	1e9p	B	151
0.937	1e9y	A	238
0.837	1e9y	B	569
0.823	1eb0	A	147
1.000	1eb1	A	10
0.669	1eb1	H	257
0.778	1eb1	L	27
0.800	1ed3	A	275
0.687	1ed3	B	99
0.769	1ed3	C	13
0.800	1ed3	D	275
0.727	1ed3	E	99
0.769	1ed3	F	13
0.522	1edy	A	134
0.515	1edy	B	134
0.864	1eei	D	103
0.864	1eei	E	103
0.874	1eei	F	103
0.874	1eei	G	103
0.883	1eei	H	103
0.772	1eio	A	127
0.649	1eiy	A	350
0.518	1eiy	B	785
1.000	1eiy	C	76
0.552	1ejf	A	125
0.552	1ejf	B	125
0.616	1elo	A	691
0.511	1ena	A	135
0.814	1ep3	A	311
0.740	1ep3	B	262
0.427	1epa	A	164
0.433	1epa	B	164
0.427	1epb	A	164
0.433	1epb	B	164
0.840	1epw	A	1287

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.862	1ovw	C	398
0.862	1ovw	D	398
1.000	1p03	A	198
1.000	1p03	P	4
0.862	1pbo	A	159
0.899	1pbo	B	159
0.835	1pcr	H	260
0.822	1pcr	L	281
0.893	1pcr	M	307
0.823	1pfc	A	113
0.994	1pgr	A	175
0.823	1pgr	B	215
0.989	1pgr	C	175
0.870	1pgr	D	215
0.994	1pgr	E	175
0.823	1pgr	F	215
0.989	1pgr	G	175
0.870	1pgr	H	215
0.522	1pgs	A	314
0.476	1phs	A	397
0.663	1pin	A	163
1.000	1pin	B	2
0.547	1pkl	A	499
0.555	1pkl	B	499
0.553	1pkl	C	499
0.553	1pkl	D	499
0.545	1pkl	E	499
0.539	1pkl	F	499
0.551	1pkl	G	499
0.545	1pkl	H	499
0.646	1pkt	A	79
0.843	1pky	A	470
0.830	1pky	B	470
0.823	1pky	C	470
0.838	1pky	D	470
0.949	1pnd	A	99
0.939	1png	A	314
0.558	1psa	A	326
0.558	1psa	B	326
0.750	1psa	I	4
0.750	1psa	J	4

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.468	1eqw	A	156
0.474	1eqw	B	156
0.487	1eqw	C	156
0.474	1eqw	D	156
0.487	1eso	A	154
0.674	1et7	A	341
0.602	1euj	A	166
0.633	1euj	B	166
0.696	1evh	A	112
1.000	1evh	B	6
0.576	1ex3	A	245
0.490	1ey5	A	149
0.497	1eyd	A	149
0.892	1eyl	A	186
0.518	1eys	C	382
0.591	1eys	H	259
0.504	1eys	L	280
0.577	1eys	M	324
0.798	1ez1	A	392
0.801	1ez1	B	392
0.787	1f0l	A	535
0.776	1f0l	B	535
0.781	1f45	A	306
0.523	1f45	B	197
0.458	1f4w	H	216
0.457	1f4w	L	210
0.570	1f58	H	228
0.486	1f58	L	216
1.000	1f58	P	23
0.656	1f5l	A	253
0.623	1f8a	B	167
0.571	1f8a	C	7
0.668	1f92	A	253
0.714	1faj	A	175
0.520	1fak	H	254
0.800	1fak	I	55
0.651	1fak	L	152
0.515	1fak	T	206
0.927	1fdl	H	218
1.000	1fdl	L	214
0.853	1fdl	Y	129

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
1.000	1pse	A	69
0.852	1pst	H	237
0.492	1pst	L	266
0.449	1pst	M	296
0.452	1qa7	A	217
0.465	1qa7	B	217
0.456	1qa7	C	217
0.419	1qa7	D	217
0.811	1qaw	A	74
0.797	1qaw	B	74
0.811	1qaw	C	74
0.811	1qaw	D	74
0.811	1qaw	E	74
0.811	1qaw	F	74
0.811	1qaw	G	74
0.811	1qaw	H	74
0.824	1qaw	I	74
0.811	1qaw	J	74
0.811	1qaw	K	74
0.576	1qbs	A	99
0.636	1qbs	B	99
0.626	1qbt	A	99
0.626	1qbt	B	99
0.413	1qcr	A	446
0.407	1qcr	B	423
0.296	1qcr	C	378
0.320	1qcr	D	75
0.628	1qcr	E	196
0.320	1qcr	F	103
0.386	1qcr	G	70
0.183	1qcr	H	60
1.000	1qcr	I	28
0.271	1qcr	J	59
0.578	1qcr	K	45
0.570	1qd0	A	128
0.675	1qdc	A	237
0.675	1qdc	B	237
0.675	1qdc	C	237
0.675	1qdc	D	237
0.827	1qg5	A	162
0.889	1qhr	A	36

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.678	1fem	A	183
0.851	1fg2	A	281
0.770	1fg2	B	100
0.778	1fg2	C	9
0.843	1fg2	D	281
0.770	1fg2	E	100
0.778	1fg2	F	9
0.851	1fg2	G	281
0.770	1fg2	H	100
0.778	1fg2	I	9
0.851	1fg2	J	281
0.770	1fg2	K	100
0.778	1fg2	L	9
0.837	1fgm	A	839
0.837	1fgt	A	839
0.476	1fgv	H	124
0.477	1fgv	L	109
0.735	1fiv	A	113
0.833	1fiv	B	6
0.381	1fkn	A	391
0.384	1fkn	B	391
0.375	1fkn	C	8
0.375	1fkn	D	8
0.521	1flr	H	219
0.507	1flr	L	219
0.758	1fmm	S	132
0.819	1fmv	A	761
0.834	1fmw	A	761
0.443	1for	H	219
0.329	1for	L	210
1.000	1fph	F	11
0.903	1fph	H	259
1.000	1fph	I	12
0.972	1fph	L	36
0.623	1fqj	A	260
0.635	1fqn	A	260
0.509	1fmt	A	112
0.473	1fmt	B	112
0.635	1fr4	A	260
0.472	1fsk	A	159
0.514	1fsk	B	214

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.618	1qhr	B	259
1.000	1qhr	I	10
0.872	1ql4	A	429
0.737	1qjs	A	460
0.737	1qjs	B	460
0.770	1qlh	A	374
0.530	1qni	A	581
0.527	1qni	B	581
0.528	1qni	C	581
0.535	1qni	D	581
0.528	1qni	E	581
0.530	1qni	F	581
0.909	1qmw	A	242
0.909	1qmw	B	242
0.905	1qmw	C	242
0.905	1qmw	D	242
0.605	1qoi	A	177
0.893	1qoo	A	242
0.872	1qoo	B	242
0.893	1qoo	C	242
0.893	1qoo	D	242
0.706	1qpx	A	218
0.706	1qpx	B	218
0.744	1qqd	A	273
0.788	1qqd	B	99
0.667	1qqd	C	9
0.550	1qs8	A	329
0.550	1qs8	B	329
1.000	1qs8	C	6
1.000	1qs8	D	6
0.766	1qwf	A	64
0.583	1qwf	B	12
0.636	1r24	A	206
0.571	1r24	B	217
0.660	1r24	C	206
0.567	1r24	D	217
0.667	1rkr	A	129
0.651	1rkr	B	129
0.667	1rkr	C	129
0.667	1rkr	D	129
0.789	1ruo	A	209

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.427	1fsk	C	220
0.472	1fsk	D	159
0.514	1fsk	E	214
0.432	1fsk	F	220
0.472	1fsk	G	159
0.514	1fsk	H	214
0.432	1fsk	I	220
0.472	1fsk	J	159
0.514	1fsk	K	214
0.427	1fsk	L	220
0.458	1fve	A	214
0.547	1fve	B	223
0.449	1fve	C	214
0.547	1fve	D	223
0.566	1fvi	A	297
0.910	1fwa	A	100
0.887	1fwa	B	106
0.817	1fwa	C	567
0.900	1fwj	A	100
0.887	1fwj	B	106
0.799	1fwj	C	567
0.574	1fxz	A	476
0.590	1fxz	B	476
0.597	1fxz	C	476
0.869	1fzo	A	274
0.737	1fzo	B	99
1.000	1fzo	P	9
0.619	1g0x	A	197
0.455	1g1k	A	143
0.455	1g1k	B	143
0.444	1g35	A	99
0.485	1g35	B	99
0.796	1g3s	A	226
0.618	1g53	A	259
0.667	1g6p	A	66
0.689	1g8r	A	411
0.693	1g8r	B	411
0.819	1g94	A	448
0.858	1g9f	A	253
0.835	1g9h	A	448
0.516	1g9o	A	91

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.861	1ruo	B	209
1.000	1ruo	C	14
1.000	1ruo	D	17
1.000	1ruo	E	14
1.000	1ruo	F	17
0.696	1scr	A	237
0.831	1shf	A	59
0.831	1shf	B	59
0.584	1sky	B	502
0.590	1sky	E	473
0.351	1slc	A	134
0.351	1slc	B	134
0.351	1slc	C	134
0.351	1slc	D	134
0.604	1snp	A	149
1.000	1sre	A	121
1.000	1sre	B	121
1.000	1srf	A	121
1.000	1srf	B	121
0.500	1sro	A	76
0.795	1ste	A	239
0.946	1stg	A	149
0.548	1stm	A	157
0.548	1stm	B	157
0.548	1stm	C	157
0.548	1stm	D	157
0.548	1stm	E	157
0.894	1sxb	A	151
0.874	1sxb	B	151
0.671	1tei	A	237
0.671	1tei	B	237
0.675	1tei	C	237
0.675	1tei	D	237
0.675	1tei	E	237
0.650	1tei	F	237
0.675	1tei	G	237
0.650	1tei	H	237
1.000	1thv	A	207
0.857	1tiu	A	98
1.000	1tng	A	229
0.936	1tff	A	94

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.752	1gbg	A	214
0.872	1gcf	A	109
0.545	1gg3	A	279
0.523	1gg3	B	279
0.541	1gg3	C	279
0.953	1ggb	H	215
1.000	1ggb	L	215
0.844	1ggt	A	731
0.844	1ggt	B	731
0.782	1gh5	A	87
0.661	1ghv	H	257
1.000	1ghv	I	11
0.833	1ghv	L	36
0.677	1ghy	H	257
1.000	1ghy	I	11
0.833	1ghy	L	36
0.658	1gic	A	237
0.696	1gic	B	237
0.739	1gjd	A	23
0.751	1gjd	B	253
0.457	1gix	A	81
0.872	1gka	A	180
0.592	1gka	B	174
0.879	1glh	A	214
0.533	1gm5	A	780
1.000	1gm5	X	20
1.000	1gm5	Y	20
1.000	1gm5	Z	9
0.526	1gmm	A	133
0.862	1gog	A	639
0.651	1grw	A	126
0.619	1grw	B	126
0.675	1grw	C	126
0.643	1grw	D	126
0.733	1gsl	A	243
0.810	1gts	A	553
1.000	1gts	B	74
0.819	1gy2	A	155
0.826	1gy2	B	155
0.631	1gzi	A	65
0.860	1gzq	A	300

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.802	1ubp	A	101
0.607	1ubp	B	122
0.721	1ubp	C	570
0.934	1ulp	A	152
0.851	1vca	A	202
0.851	1vca	B	202
0.641	1vhp	A	117
0.841	1vom	A	762
0.644	1vqb	A	87
0.644	1vqf	A	87
0.621	1vqh	A	87
0.813	1wbl	A	241
0.813	1wbl	B	241
0.813	1wbl	C	241
0.813	1wbl	D	241
0.577	1wej	F	104
0.583	1wej	H	223
0.472	1wej	L	214
0.840	1xat	A	212
0.832	1xyo	A	190
0.832	1xyo	B	190
0.718	1ydb	A	259
0.559	1yei	H	222
0.507	1yei	L	219
0.554	1yej	H	222
0.498	1yej	L	219
0.844	1yge	A	839
0.703	1zsc	A	259
0.435	2dlf	H	124
0.469	2dlf	L	113
0.737	2eip	A	175
0.709	2eip	B	175
0.511	2enb	A	135
0.693	2ezn	A	101
0.453	2fgw	H	232
0.537	2fgw	L	214
0.848	2hmg	A	328
0.914	2hmg	B	175
0.848	2hmg	C	328
0.914	2hmg	D	175
0.848	2hmg	E	328

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.770	1gzq	B	100
0.874	1h0i	A	499
0.880	1h0i	B	499
0.000	1h0i	C	5
0.000	1h0i	D	5
0.599	1h8o	A	252
0.579	1h8o	B	252
0.738	1h9w	A	237
0.709	1h9w	B	237
1.000	1hao	D	15
0.734	1hao	H	259
0.972	1hao	L	36
1.000	1hao	P	4
0.433	1hav	A	217
0.465	1hav	B	217
0.906	1hb4	A	331
0.781	1hbq	A	183
0.569	1hdf	A	102
0.627	1hdf	B	102
0.802	1hdy	A	374
0.802	1hdy	B	374
0.802	1hdz	A	374
0.802	1hdz	B	374
0.444	1hef	E	99
0.625	1hef	I	8
0.773	1hfu	A	503
0.848	1hge	A	328
0.914	1hge	B	175
0.848	1hge	C	328
0.914	1hge	D	175
0.848	1hge	E	328
0.914	1hge	F	175
0.453	1hh2	P	344
0.463	1hh6	A	214
0.610	1hh6	B	213
1.000	1hh6	C	11
0.626	1hil	A	99
0.616	1hil	B	99
0.381	1hla	A	270
0.670	1hla	M	91
0.536	1hm8	A	468

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.914	2hmg	F	175
0.613	2hrv	A	142
0.592	2hrv	B	142
0.684	2ila	A	155
0.876	2jcw	A	153
0.672	2jxr	A	329
1.000	2jxr	I	5
0.910	2kau	A	100
0.858	2kau	B	106
0.811	2kau	C	567
0.635	2lal	A	181
0.673	2lal	B	52
0.635	2lal	C	181
0.673	2lal	D	52
0.870	2mha	A	270
0.768	2mha	B	99
0.870	2mha	C	270
0.768	2mha	D	99
1.000	2mha	E	8
1.000	2mha	F	8
0.859	2mip	A	99
0.859	2mip	B	99
0.859	2mip	C	99
0.859	2mip	D	99
0.000	2mip	E	7
0.000	2mip	F	7
0.000	2mip	G	7
0.000	2mip	H	7
0.641	2mys	A	843
0.657	2mys	B	166
0.631	2mys	C	149
0.684	2ohx	A	374
0.684	2ohx	B	374
0.787	2pab	A	127
0.787	2pab	B	127
0.607	2pia	A	321
0.820	2qil	A	194
0.835	2qil	B	194
0.835	2qil	C	194
0.553	2rcs	H	217
0.575	2rcs	L	214

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.536	1hm8	B	468
0.833	1hmr	A	132
0.833	1hmt	A	132
0.855	1hn2	A	159
0.818	1hn2	B	159
0.754	1hoc	A	272
0.677	1hoc	B	99
0.667	1hoc	C	9
0.556	1hvp	A	99
0.566	1hvp	B	99
0.829	1hql	A	257
0.833	1hql	B	257
0.801	1hsa	A	276
0.788	1hsa	B	99
0.778	1hsa	C	9
0.801	1hsa	D	276
0.788	1hsa	E	99
0.778	1hsa	F	9
0.837	1hsb	A	270
0.788	1hsb	B	99
0.000	1hsb	C	3
0.000	1hsb	D	2
0.832	1htb	A	374
0.824	1htb	B	374
0.515	1hte	A	99
0.515	1hte	B	99
0.000	1hte	C	4
0.815	1huh	A	260
0.840	1huk	A	281
0.875	1huk	B	281
0.759	1hv9	A	456
0.741	1hv9	B	456
0.911	1hvc	A	203
0.374	1hvr	A	99
0.374	1hvr	B	99
0.862	1hwm	A	254
0.590	1hwm	B	266
0.566	1hxb	A	99
0.556	1hxb	B	99
0.800	1hxb	I	5
0.537	1hxm	A	229

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.515	2ren	A	340
0.605	2rhe	A	114
0.787	2roy	A	127
0.787	2roy	B	127
0.474	2rta	A	135
0.437	2rto	B	135
0.474	2rto	D	135
0.856	2seb	A	181
0.771	2seb	B	192
0.862	2seb	D	239
0.917	2seb	E	12
0.824	2sil	A	381
0.667	2spg	A	66
0.787	2try	A	127
0.803	2try	B	127
0.870	2ubp	A	100
0.607	2ubp	B	122
0.719	2ubp	C	570
0.502	2wea	A	323
0.554	2znc	A	258
0.848	3cyh	A	164
1.000	3cyh	B	2
0.654	3enr	A	237
0.654	3enr	B	237
0.634	3ezm	A	101
0.859	3gal	A	135
0.837	3gal	B	135
0.814	3hfm	H	215
0.953	3hfm	L	214
0.853	3hfm	Y	129
0.841	3msp	A	126
0.841	3msp	B	126
0.917	3tf4	A	605
0.921	3tf4	B	605
0.737	3upj	A	99
0.758	3upj	B	99
0.543	3znc	A	258
0.575	4fgf	A	146
0.649	4gbq	A	74
1.000	4gbq	B	11
0.886	4np1	A	184

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.533	1hxm	B	242
0.568	1hxm	C	229
0.541	1hxm	D	242
0.550	1hxm	E	229
0.517	1hxm	F	242
0.533	1hxm	G	229
0.541	1hxm	H	242
0.761	1hys	A	553
0.692	1hys	B	425
0.523	1hys	C	214
0.668	1hys	D	220
1.000	1hys	E	23
1.000	1hys	F	22
0.509	1hyx	H	224
0.523	1hyx	L	216
0.924	1hzc	A	66
0.924	1hzc	B	66
0.835	1i18	A	97
0.835	1i18	B	97
0.713	1i1j	A	108
0.667	1i1j	B	108
0.859	1i3r	A	192
0.496	1i3r	B	228
0.849	1i3r	C	192
0.504	1i3r	D	228
0.865	1i3r	E	192
0.500	1i3r	F	228
0.854	1i3r	G	192
0.491	1i3r	H	228
0.630	1i3u	A	127
0.691	1i40	A	175
0.941	1i4r	A	239
0.399	1i50	A	1287
0.511	1i50	B	1224
0.925	1i50	C	318
0.981	1i50	E	215
0.935	1i50	F	155
1.000	1i50	H	146
1.000	1i50	I	122

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.875	4np1	B	184
0.919	4phv	A	99
0.919	4phv	B	99
0.865	4rcr	H	260
0.520	4rcr	L	281
0.469	4rcr	M	307
0.704	4sbv	A	260
0.704	4sbv	B	260
0.704	4sbv	C	260
0.687	4upj	A	99
0.677	4upj	B	99
0.848	5cyh	A	164
1.000	5cyh	B	2
0.567	5gch	A	245
0.668	5gds	H	259
1.000	5gds	I	26
0.889	5gds	L	36
0.621	5msi	A	66
0.570	5nuc	A	149
0.875	5prc	C	336
0.694	5prc	H	258
0.864	5prc	L	273
0.898	5prc	M	323
0.851	5tss	A	194
0.825	5tss	B	194
0.905	6fab	H	222
0.893	6fab	L	214
0.567	6gch	A	245
0.875	6prc	C	336
0.694	6prc	H	258
0.864	6prc	L	273
0.898	6prc	M	323
0.592	6taa	A	478
0.465	7hvp	A	99
0.465	7hvp	B	99
1.000	7hvp	C	9
0.732	7i1b	A	153
0.535	7odc	A	424

APPENDIX Q

Prediction accuracies for each chain of each protein in α/β class in trial set

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.838	11gs	A	210
0.838	11gs	B	210
0.810	12gs	A	210
0.829	12gs	B	210
0.667	12gs	C	3
0.667	12gs	D	3
0.819	14gs	A	210
0.819	14gs	B	210
0.626	1a05	A	358
0.626	1a05	B	358
0.733	1a0c	A	438
0.733	1a0c	B	438
0.733	1a0c	C	438
0.733	1a0c	D	438
0.779	1a0e	A	443
0.779	1a0e	D	443
0.741	1a4z	A	499
0.739	1a4z	B	499
0.737	1a4z	C	499
0.741	1a4z	D	499
0.765	1a50	A	268
0.811	1a50	B	396
0.765	1a53	A	247
0.810	1a5a	A	268
0.801	1a5a	B	397
0.860	1a5k	A	100
0.881	1a5k	B	101
0.761	1a5k	C	566
0.848	1a5y	A	330
0.754	1a71	A	374
0.749	1a71	B	374
0.760	1a8e	A	329
0.556	1a8y	A	367
0.776	1a95	A	152
0.783	1a95	B	152

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.859	1hg0	B	327
0.869	1hg0	C	327
0.865	1hg0	D	327
0.608	1hh1	A	143
0.513	1hjr	A	158
0.513	1hjr	B	158
0.513	1hjr	C	158
0.513	1hjr	D	158
0.476	1hm8	A	468
0.476	1hm8	B	468
0.468	1hm9	A	468
0.468	1hm9	B	468
0.807	1hmv	A	560
0.798	1hmv	B	440
0.442	1hna	A	217
0.442	1hnb	A	217
0.442	1hnb	B	217
0.442	1hnc	A	217
0.442	1hnc	B	217
0.442	1hnc	C	217
0.442	1hnc	D	217
0.792	1hnk	A	317
0.868	1hon	A	431
0.868	1hon	B	431
0.868	1hop	A	431
0.868	1hop	B	431
0.774	1hot	A	266
0.793	1hot	B	266
0.840	1hp4	A	512
0.889	1hq5	A	323
0.889	1hq5	B	323
0.898	1hqh	A	323
0.898	1hqh	B	323
0.898	1hqh	C	323
0.807	1hqu	A	560

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.737	1a95	C	152
0.796	1a95	D	152
0.608	1a97	A	148
0.581	1a97	B	148
0.588	1a97	C	148
0.696	1a97	D	148
0.870	1a9z	A	338
0.843	1aa6	A	715
0.695	1adc	A	374
0.695	1adc	B	374
0.868	1ade	A	431
0.868	1ade	B	431
0.488	1ag1	O	250
0.508	1ag1	T	250
0.651	1agp	A	166
0.414	1agx	A	331
0.825	1ah3	A	315
0.805	1aip	A	405
0.798	1aip	B	405
0.643	1aip	C	196
0.643	1aip	D	196
0.798	1aip	E	405
0.810	1aip	F	405
0.648	1aip	G	196
0.597	1aip	H	196
0.775	1aja	A	449
0.775	1aja	B	449
0.775	1ajc	A	449
0.775	1ajc	B	449
0.803	1ak1	A	310
0.735	1al7	A	359
0.793	1ali	A	449
0.786	1ali	B	449
0.795	1alj	A	449
0.797	1alj	B	449
0.881	1aln	A	294
0.614	1amf	A	233
0.818	1an9	A	340
0.806	1an9	B	340
0.895	1ao8	A	162
0.836	1aqj	A	421

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.774	1hqu	B	430
0.657	1hqy	A	175
0.691	1hqy	B	175
0.680	1hqy	C	175
0.771	1hqy	D	175
0.742	1hqy	E	449
0.706	1hqy	F	449
0.743	1hso	A	374
0.759	1hso	B	374
0.786	1hsz	A	374
0.791	1hsz	B	374
0.678	1hur	A	180
0.678	1hur	B	180
0.529	1hv8	A	367
0.542	1hv8	B	367
0.425	1hvq	A	273
0.886	1hwy	A	501
0.886	1hwy	B	501
0.886	1hwy	C	501
0.886	1hwy	D	501
0.886	1hwy	E	501
0.886	1hwy	F	501
0.804	1hwz	A	501
0.804	1hwz	B	501
0.804	1hwz	C	501
0.804	1hwz	D	501
0.804	1hwz	E	501
0.804	1hwz	F	501
0.817	1hx0	A	496
0.762	1hx7	A	105
0.846	1i13	A	221
0.846	1i13	B	221
0.609	1i2d	A	573
0.607	1i2d	B	573
0.606	1i2d	C	573
0.835	1i2p	A	316
0.835	1i2p	B	316
0.856	1i3k	A	348
0.859	1i3k	B	348
0.870	1i43	A	445
0.845	1i43	B	445

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.824	1aqj	B	421
0.837	1aqw	A	209
0.837	1aqw	B	209
0.813	1aqw	C	209
0.847	1aqw	D	209
0.667	1aqw	E	3
0.667	1aqw	F	3
0.667	1aqw	G	3
0.667	1aqw	H	3
0.873	1as0	A	353
0.836	1ava	A	403
0.836	1ava	B	403
0.448	1ava	C	181
0.414	1ava	D	181
0.719	1avt	A	274
0.651	1axd	A	209
0.660	1axd	B	209
0.500	1axd	C	4
0.500	1axd	D	4
0.701	1az0	A	244
0.697	1az0	B	244
1.000	1az0	C	11
1.000	1az0	D	11
0.787	1az2	A	315
0.779	1az3	A	244
0.791	1az3	B	244
0.468	1az9	A	457
0.671	1azw	A	313
0.661	1azw	B	313
0.884	1azy	A	440
0.884	1azy	B	440
0.857	1b0h	A	517
1.000	1b0h	B	3
0.826	1b0j	A	754
0.856	1b0z	A	445
0.839	1b16	A	254
0.827	1b16	B	254
0.739	1b1x	A	689
0.753	1b23	P	405
1.000	1b23	R	74
0.857	1b2h	A	517

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.867	1i43	C	445
0.858	1i43	D	445
0.874	1i43	E	445
0.845	1i43	F	445
0.861	1i43	G	445
0.843	1i43	H	445
0.854	1i43	I	445
0.847	1i43	J	445
0.854	1i43	K	445
0.847	1i43	L	445
0.826	1i52	A	236
0.436	1i6h	A	1287
0.561	1i6h	B	1224
0.896	1i6h	C	318
1.000	1i6h	D	13
0.856	1i6h	E	215
0.871	1i6h	F	155
1.000	1i6h	H	146
1.000	1i6h	I	122
0.914	1i6h	J	70
0.850	1i6h	K	120
1.000	1i6h	L	70
1.000	1i6h	R	9
0.905	1i6k	A	328
0.909	1i6m	A	328
0.874	1i6n	A	278
0.797	1ia2	A	192
0.813	1ia2	B	192
0.858	1idc	A	416
0.858	1idd	A	416
0.900	1ie7	A	100
0.675	1ie7	B	126
0.653	1ie7	C	570
0.744	1iej	A	332
0.838	1ieq	A	605
0.593	1ig3	A	263
0.570	1ig3	B	263
0.766	1ihy	A	333
0.766	1ihy	B	333
0.766	1ihy	C	333
0.766	1ihy	D	333

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
1.000	1b2h	B	3
0.849	1b2y	A	496
0.868	1b3y	A	302
0.868	1b3z	A	302
0.867	1b40	A	517
1.000	1b40	B	3
0.859	1b46	A	517
1.000	1b46	B	3
0.861	1b4h	A	517
1.000	1b4h	B	3
0.857	1b5i	A	517
1.000	1b5i	B	3
0.857	1b6h	A	517
1.000	1b6h	B	3
0.849	1b6t	A	159
0.830	1b6t	B	159
0.772	1b74	A	254
0.835	1b8s	A	504
0.664	1b93	A	152
0.697	1b93	B	152
0.638	1b93	C	152
0.840	1b9h	A	388
0.859	1b9j	A	517
0.000	1b9j	B	3
0.848	1bd3	A	243
0.844	1bd3	B	243
0.852	1bd3	C	243
0.852	1bd3	D	243
0.723	1be6	A	274
0.795	1beu	A	268
0.824	1beu	B	397
0.807	1bf3	A	394
0.723	1bfu	A	274
0.781	1bg3	A	919
0.783	1bg3	B	919
0.713	1bg5	A	254
0.794	1bga	A	447
0.794	1bga	B	447
0.796	1bga	C	447
0.805	1bga	D	447
0.745	1bh6	A	274

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.546	1li2	A	524
0.540	1li2	B	524
0.793	1liw	A	232
0.597	1ilo	A	77
0.906	1im5	A	180
0.859	1in5	A	334
0.841	1j49	A	333
0.814	1j49	B	333
0.898	1j4e	A	363
0.901	1j4e	B	363
0.898	1j4e	C	363
0.898	1j4e	D	363
0.747	1j6z	A	375
0.476	1j79	A	347
0.481	1j79	B	347
0.599	1j7x	A	302
0.916	1j8m	F	297
0.635	1j90	A	230
0.630	1j90	B	230
0.413	1j99	A	293
0.611	1jae	A	471
0.602	1jah	A	166
0.649	1jb9	A	316
0.773	1jbe	A	128
0.840	1jcg	A	344
0.846	1jcl	A	260
0.846	1jcl	B	260
0.872	1jct	A	446
0.874	1jct	B	446
0.903	1jdt	A	236
0.907	1jdt	B	236
0.890	1jdt	C	236
0.622	1jeo	A	180
0.788	1jf5	A	585
0.781	1jf5	B	585
0.585	1jff	A	451
0.483	1jff	B	445
0.543	1jfu	A	186
0.548	1jfu	B	186
0.881	1jg3	A	235
0.902	1jg3	B	235

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.839	1bhj	A	292
0.829	1bhj	B	292
0.831	1bkg	A	385
0.829	1bkg	B	385
0.826	1bkg	C	385
0.829	1bkg	D	385
0.613	1ble	A	163
0.816	1blf	A	689
0.847	1bli	A	483
0.731	1bli	E	487
0.400	1bli	I	5
0.814	1bo5	O	501
0.804	1bo5	Z	501
0.881	1bof	A	353
0.783	1bp5	A	337
0.792	1bp5	B	337
0.786	1bp5	C	337
0.798	1bp5	D	337
0.803	1bpw	A	503
0.803	1bpw	B	503
0.803	1bpw	C	503
0.803	1bpw	D	503
0.703	1bqg	A	451
0.876	1br1	A	820
0.913	1br1	B	150
0.867	1br1	C	820
0.913	1br1	D	150
0.866	1br1	E	820
0.893	1br1	F	150
0.871	1br1	G	820
0.920	1br1	H	150
0.700	1bsu	A	243
0.716	1bsu	B	243
1.000	1bsu	C	10
1.000	1bsu	D	11
0.566	1bt4	A	362
0.728	1bu5	A	147
0.701	1bu5	B	147
0.783	1bvh	A	157
0.813	1bvr	A	268
0.810	1bvr	B	268

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
1.000	1jg3	C	6
1.000	1jg3	D	6
0.815	1jg9	A	628
0.888	1jhm	A	356
0.905	1jjk	A	420
0.579	1jji	A	311
0.582	1jji	B	311
0.579	1jji	C	311
0.572	1jji	D	311
0.767	1jjk	A	245
0.767	1jjk	B	245
0.767	1jjk	C	245
0.767	1jjk	D	245
0.767	1jjk	E	245
0.767	1jjk	F	245
0.767	1jjk	G	245
0.767	1jjk	H	245
0.767	1jjk	I	245
0.767	1jjk	J	245
0.767	1jjk	K	245
0.767	1jjk	L	245
0.767	1jjk	M	245
0.767	1jjk	N	245
0.767	1jjk	O	245
0.767	1jjk	P	245
0.660	1jlv	A	206
0.675	1jlv	A	209
0.703	1jlv	B	209
0.689	1jlv	C	209
0.684	1jlv	D	209
0.699	1jlv	E	209
0.718	1jlv	F	209
0.794	1joa	A	447
0.692	1jol	A	159
0.723	1jol	B	159
0.861	1jpl	A	296
0.898	1jpv	A	236
0.907	1jpv	B	236
0.903	1jpv	C	236
0.499	1jqg	A	433
0.755	1jr1	A	514

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.810	1bvr	C	268
0.769	1bvr	D	268
0.720	1bvr	E	268
0.813	1bvr	F	268
0.715	1bvu	A	418
0.722	1bvu	B	418
0.713	1bvu	C	418
0.715	1bvu	D	418
0.715	1bvu	E	418
0.720	1bvu	F	418
0.596	1bw0	A	416
0.591	1bw0	B	416
0.830	1bwk	A	399
0.783	1bxc	A	387
0.773	1bxc	B	387
0.778	1bxc	C	387
0.773	1bxc	D	387
0.657	1bye	A	213
0.638	1bye	B	213
0.648	1bye	C	213
0.563	1bye	D	213
0.786	1c14	A	262
0.798	1c14	B	262
0.745	1c1b	A	560
0.757	1c1b	B	440
0.813	1c1h	A	310
0.821	1c2t	A	212
0.816	1c2t	B	212
0.843	1c3j	A	351
0.885	1c3p	A	375
0.877	1c3r	A	375
0.877	1c3r	B	375
0.779	1c4a	A	574
0.859	1c7j	A	489
0.402	1c7r	A	445
0.817	1c8v	A	268
0.816	1c8v	B	397
0.558	1cbf	A	285
0.837	1cc0	A	190
0.837	1cc0	C	190
0.623	1cc0	E	204

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.782	1jr1	B	514
0.946	1jrz	A	571
0.946	1jrz	B	571
0.700	1jtf	A	307
0.522	1jtk	A	136
0.522	1jtk	B	136
0.709	1ju9	A	374
0.717	1ju9	B	374
0.566	1jv1	A	505
0.539	1jv1	B	505
0.564	1jv3	A	505
0.543	1jv3	B	505
0.659	1jvx	A	372
0.626	1jvy	A	372
0.855	1jwa	B	249
0.778	1jwa	D	81
0.548	1jxa	A	608
0.576	1jxa	B	608
0.502	1jxa	C	608
0.814	1jyf	A	349
0.720	1jyl	A	254
0.701	1jyl	B	254
0.720	1jyl	C	254
0.713	1jyl	D	254
0.702	1jyv	A	1023
0.702	1jyv	B	1023
0.693	1jyv	C	1023
0.706	1jyv	D	1023
0.695	1jyx	A	1023
0.701	1jyx	B	1023
0.700	1jyx	C	1023
0.701	1jyx	D	1023
0.713	1jyy	A	1023
0.712	1jyy	B	1023
0.713	1jyy	C	1023
0.716	1jyy	D	1023
0.713	1jyy	E	1023
0.712	1jyy	F	1023
0.713	1jyy	G	1023
0.716	1jyy	H	1023
0.701	1jz6	A	1023

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.623	1cc0	F	204
0.820	1cc4	A	394
0.820	1cc6	A	394
0.589	1cer	O	331
0.580	1cer	P	331
0.586	1cer	Q	331
0.586	1cer	R	331
0.799	1cex	A	214
0.838	1cg1	A	431
0.819	1cgu	A	684
0.758	1cke	A	227
0.794	1cm1	A	389
0.769	1cpx	A	307
0.593	1cqx	A	403
0.598	1cqx	B	403
0.787	1cr6	A	554
0.767	1cr6	B	554
0.675	1crr	A	166
0.723	1cse	E	274
0.586	1cse	I	70
0.651	1ctq	A	166
0.790	1cub	A	214
0.790	1cuf	A	214
0.794	1cuy	A	214
0.483	1cvr	A	435
0.000	1cvr	I	3
0.574	1cw0	A	155
1.000	1cw0	M	12
1.000	1cw0	N	3
1.000	1cw0	O	9
0.829	1cw4	A	416
0.856	1cw7	A	416
0.672	1cyg	A	680
0.740	1czh	A	169
0.763	1d04	A	169
0.845	1d1c	A	761
0.833	1d1g	A	168
0.833	1d1g	B	168
0.812	1d1s	A	373
0.818	1d1s	B	373
0.810	1d1s	C	373

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.698	1jz6	B	1023
0.706	1jz6	C	1023
0.706	1jz6	D	1023
0.703	1jz7	A	1023
0.704	1jz7	B	1023
0.704	1jz7	C	1023
0.705	1jz7	D	1023
0.875	1jzq	A	821
0.830	1k0l	A	394
0.820	1k0j	A	394
0.867	1k0m	A	241
0.871	1k0m	B	241
0.830	1k3d	A	540
0.810	1k3u	A	268
0.803	1k3u	B	396
0.868	1k4h	A	386
0.530	1k5h	A	398
0.523	1k5h	B	398
0.535	1k5h	C	398
0.813	1k7e	A	268
0.788	1k7e	B	396
0.563	1k7h	A	476
0.569	1k7h	B	476
0.552	1k8q	A	377
0.554	1k8q	B	377
0.857	1kar	A	434
0.857	1kar	B	434
0.854	1kdy	A	370
0.000	1kdy	B	3
0.849	1kdz	A	370
0.250	1kdz	B	4
0.850	1kek	A	1231
0.844	1kek	B	1231
0.874	1ker	A	348
0.871	1ker	B	348
0.848	1kew	A	361
0.848	1kew	B	361
0.669	1kfd	A	605
0.834	1kfi	A	572
0.841	1kfi	B	572
0.551	1kgs	A	225

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.820	1d1s	D	373
0.877	1d2c	A	292
0.884	1d2c	B	292
0.703	1d3c	A	686
0.780	1d4a	A	273
0.784	1d4a	B	273
0.799	1d4a	C	273
0.799	1d4a	D	273
0.757	1d4x	A	375
0.857	1d4x	G	126
0.798	1d6s	A	322
0.798	1d6s	B	322
0.813	1d9e	A	284
0.845	1d9e	B	284
0.838	1d9e	C	284
0.827	1d9e	D	284
0.839	1dag	A	224
0.813	1dai	A	224
0.795	1dam	A	224
0.689	1dar	A	691
0.682	1dc6	A	330
0.688	1dc6	B	330
0.545	1dce	A	567
0.677	1dce	B	331
0.545	1dce	C	567
0.692	1dce	D	331
0.825	1dck	A	126
0.810	1dck	B	126
0.697	1dd8	A	406
0.695	1dd8	B	406
0.700	1dd8	C	406
0.695	1dd8	D	406
0.686	1ddr	A	159
0.698	1ddr	B	159
0.698	1dds	A	159
0.711	1dds	B	159
0.494	1de1	A	87
0.807	1deh	A	374
0.821	1deh	B	374
0.851	1del	A	241
0.859	1del	B	241

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.829	1kgu	A	496
0.885	1kh2	A	400
0.893	1kh2	B	400
0.888	1kh2	C	400
0.898	1kh2	D	400
0.744	1kh9	A	449
0.731	1kh9	B	449
0.876	1kic	A	339
0.873	1kic	B	339
0.871	1kjj	A	410
0.864	1kjp	A	419
0.692	1km0	A	247
0.765	1km0	B	247
0.709	1km0	C	247
0.757	1km0	D	247
0.798	1km6	A	247
0.527	1kmh	A	507
0.590	1kmh	B	498
0.798	1kmj	A	406
0.640	1kms	A	186
0.611	1knv	A	293
0.601	1knv	B	293
0.895	1kor	A	400
0.870	1kor	B	400
0.865	1kor	C	400
0.888	1kor	D	400
0.619	1kpi	A	302
0.837	1kqf	A	1015
0.840	1kqf	B	294
0.917	1kqf	C	217
0.846	1kr1	A	273
0.860	1krc	A	100
0.858	1krc	B	106
0.772	1krc	C	567
0.788	1ksu	A	571
0.797	1ksu	B	571
0.733	1kte	A	105
0.858	1kvb	A	155
0.829	1kw1	A	216
0.819	1kw1	B	216
0.801	1kxh	A	448

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.910	1deo	A	233
0.801	1dfh	A	261
0.789	1dfh	B	261
0.847	1dfo	A	417
0.859	1dfo	B	417
0.851	1dfo	C	417
0.861	1dfo	D	417
0.919	1dg3	A	592
0.723	1dg7	A	159
0.887	1dgd	A	432
0.908	1dig	A	306
0.886	1dig	B	306
0.787	1dik	A	874
0.516	1dio	A	554
0.469	1dio	B	224
0.513	1dio	E	224
0.630	1dio	G	173
0.545	1dio	L	554
0.618	1dio	M	173
0.570	1dju	A	388
0.567	1dju	B	388
0.905	1dkn	A	410
0.883	1dkp	A	410
0.856	1dlj	A	402
0.710	1dls	A	186
0.807	1dmr	A	823
0.510	1dms	A	781
0.845	1doe	A	394
0.794	1dpf	A	180
0.804	1dr4	A	189
0.804	1dr5	A	189
0.804	1dr7	A	189
0.730	1dre	A	159
0.795	1drv	A	273
0.810	1drw	A	273
0.772	1dss	G	333
0.703	1dss	R	333
0.540	1dt9	A	437
0.766	1dtn	A	359
0.798	1dtz	A	689
0.610	1du4	A	269

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.616	1kyh	A	276
0.893	1kyv	A	159
0.906	1kyv	B	159
0.912	1kyv	C	159
0.906	1kyv	D	159
0.899	1kyv	E	159
0.906	1kyy	A	159
0.906	1kyy	B	159
0.925	1kyy	C	159
0.906	1kyy	D	159
0.912	1kyy	E	159
0.804	1l0p	A	448
0.869	1l0v	A	602
0.827	1l0v	B	243
0.931	1l0v	C	130
0.975	1l0v	D	119
0.772	1l0v	M	602
0.831	1l0v	N	243
0.908	1l0v	O	130
0.966	1l0v	P	119
0.863	1l1f	A	505
0.863	1l1f	B	505
0.863	1l1f	C	505
0.863	1l1f	D	505
0.863	1l1f	E	505
0.863	1l1f	F	505
0.778	1l5h	A	491
0.826	1l5h	B	522
0.653	1l7v	A	326
0.641	1l7v	B	326
0.530	1l7v	C	249
0.530	1l7v	D	249
0.715	1lcp	A	484
0.715	1lcp	B	484
0.873	1lil	A	363
0.870	1ljl	A	131
0.854	1lkc	A	364
0.565	1llq	A	605
0.557	1llq	B	605
0.751	1lol	A	229
0.773	1lol	B	229

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.606	1du4	B	269
0.602	1du4	C	269
0.625	1du4	D	269
0.746	1dv1	A	449
0.753	1dv1	B	449
0.780	1dxo	A	273
0.766	1dxo	B	273
0.784	1dxo	C	273
0.780	1dxo	D	273
0.744	1dxq	A	273
0.744	1dxq	B	273
0.744	1dxq	C	273
0.744	1dxq	D	273
0.698	1dyh	A	159
0.698	1dyh	B	159
0.823	1dzz	P	215
0.823	1dzz	P	215
0.840	1e0x	A	313
0.840	1e0x	B	313
0.850	1E26	A	206
0.726	1e2b	A	106
0.575	1e3s	A	261
0.548	1e3s	B	261
0.571	1e3s	C	261
0.563	1e3s	D	261
0.582	1e3w	A	261
0.548	1e3w	B	261
0.567	1e3w	C	261
0.590	1e3w	D	261
0.857	1E43	A	483
0.833	1E48	P	215
0.837	1e4o	A	796
0.837	1e4o	B	796
0.898	1e5j	A	305
0.836	1e5l	A	450
0.856	1e5l	B	450
0.839	1e5t	A	710
0.638	1e6l	A	127
0.742	1e6m	A	128
0.542	1e6w	A	260
0.531	1e6w	B	260

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.821	1lon	A	457
0.751	1lp6	A	229
0.738	1lp6	B	229
0.592	1lpm	A	549
0.853	1lqt	A	456
0.851	1lqt	B	456
0.852	1lrl	A	338
0.707	1lst	A	239
0.524	1ltd	A	506
0.621	1ltd	B	506
0.806	1lth	R	319
0.803	1lth	T	319
0.434	1luc	A	355
0.722	1luc	B	324
0.478	1m6b	A	621
0.438	1m6b	B	621
0.837	1maa	A	547
0.841	1maa	B	547
0.841	1maa	C	547
0.837	1maa	D	547
0.431	1mdq	A	371
0.849	1mmn	A	762
0.743	1mpd	A	370
0.717	1mpt	A	269
0.560	1mug	A	168
0.822	1nas	A	259
0.843	1nbm	A	510
0.839	1nbm	B	510
0.827	1nbm	C	510
0.783	1nbm	D	480
0.773	1nbm	E	480
0.788	1nbm	F	480
0.952	1nbm	G	272
0.744	1ngf	A	386
0.794	1nhp	A	447
0.739	1noi	A	842
0.749	1noi	B	842
0.749	1noi	C	842
0.740	1noi	D	842
0.745	1noj	A	842
0.715	1ohj	A	186

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.546	1e6w	C	260
0.558	1e6w	D	260
0.904	1E70	M	501
0.904	1E72	M	501
0.904	1E73	M	501
0.852	1e7p	A	656
0.808	1e7p	B	239
0.941	1e7p	C	256
0.852	1e7p	D	656
0.808	1e7p	E	239
0.941	1e7p	F	256
0.852	1e7p	G	655
0.808	1e7p	H	239
0.941	1e7p	I	254
0.852	1e7p	J	655
0.808	1e7p	K	239
0.941	1e7p	L	254
0.889	1e7y	A	485
0.790	1e8d	A	262
0.798	1e8d	B	262
0.806	1e9f	A	217
0.850	1eb3	A	340
0.794	1eb8	A	262
0.802	1eb8	B	262
0.890	1ebg	A	436
0.890	1ebg	B	436
0.849	1ecv	A	298
0.583	1ecx	A	384
0.576	1ecx	B	384
0.886	1ee1	A	271
0.897	1ee1	B	271
0.749	1ef2	A	566
0.762	1ef2	B	101
0.940	1ef2	C	100
0.687	1efc	A	393
0.669	1efc	B	393
0.503	1efh	A	292
0.483	1efh	B	292
0.590	1efm	A	393
0.516	1egm	A	554
0.522	1egm	B	224

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.789	1ola	A	517
1.000	1ola	B	4
0.432	1omp	A	370
0.718	1ose	A	496
0.794	1oxm	A	214
0.799	1oxm	B	214
0.488	1oyb	A	400
0.488	1oyc	A	400
0.812	1pbb	A	394
0.848	1pbn	A	289
0.630	1pdo	A	135
0.866	1ped	A	351
0.866	1ped	B	351
0.866	1ped	C	351
0.866	1ped	D	351
0.782	1phh	A	394
0.606	1pil	A	452
0.603	1pki	A	499
0.607	1pki	B	499
0.603	1pki	C	499
0.607	1pki	D	499
0.595	1pki	E	499
0.601	1pki	F	499
0.595	1pki	G	499
0.607	1pki	H	499
0.601	1pmt	A	203
0.555	1psd	A	409
0.555	1psd	B	409
0.000	1psd	C	1
0.000	1psd	D	1
0.872	1ptg	A	298
0.827	1pxc	A	394
0.776	1pym	A	295
0.783	1pym	B	295
0.787	1qan	A	244
0.803	1qao	A	244
0.548	1qbk	B	890
0.713	1qbk	C	216
0.801	1qcd	A	236
0.826	1qf4	A	431
0.871	1qf9	A	194

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.531	1egm	E	224
0.659	1egm	G	173
0.529	1egm	L	554
0.642	1egm	M	173
0.824	1eix	A	245
0.849	1eix	B	245
0.861	1eix	C	245
0.841	1eix	D	245
0.906	1ej2	A	181
0.950	1eju	A	100
0.762	1eju	B	101
0.785	1eju	C	567
0.845	1ek6	A	348
0.848	1ek6	B	348
0.613	1eke	A	230
0.604	1eke	B	230
0.817	1el9	A	389
0.825	1el9	B	389
0.812	1eli	A	389
0.825	1eli	B	389
0.890	1els	A	436
0.851	1elu	A	390
0.846	1elu	B	390
0.597	1eo1	A	124
0.886	1eok	A	290
0.785	1ep3	A	311
0.683	1ep3	B	262
0.727	1epx	A	370
0.727	1epx	B	370
0.727	1epx	C	370
0.727	1epx	D	370
0.909	1eqc	A	394
0.467	1erv	A	105
0.776	1es8	A	223
0.941	1ese	A	306
0.676	1eth	A	448
0.632	1eth	B	95
0.676	1eth	C	448
0.611	1eth	D	95
0.847	1eve	A	543
0.850	1ewq	A	765

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.517	1qfl	A	389
0.527	1qfl	B	389
0.522	1qfl	C	389
0.530	1qfl	D	389
0.865	1qfs	A	710
0.818	1qgn	A	445
0.816	1qgn	B	445
0.818	1qgn	C	445
0.820	1qgn	D	445
0.818	1qgn	E	445
0.809	1qgn	F	445
0.820	1qgn	G	445
0.796	1qgn	H	445
0.850	1qhf	A	240
0.846	1qhf	B	240
0.815	1qhg	A	724
0.858	1qj4	A	429
0.794	1qj4	A	257
0.811	1qj5	A	429
0.825	1qj5	B	429
0.867	1qka	A	517
1.000	1qka	B	3
0.885	1qmp	A	130
0.869	1qmp	B	130
0.885	1qmp	C	130
0.869	1qmp	D	130
0.523	1qor	A	327
0.523	1qor	B	327
0.831	1qqt	A	551
0.814	1qtq	A	553
1.000	1qtq	B	75
0.679	1qui	A	321
0.676	1quk	A	321
0.601	1qva	A	223
0.613	1qvb	A	481
0.613	1qvb	B	481
0.819	1rae	A	310
0.797	1rae	B	153
0.819	1rae	C	310
0.797	1rae	D	153
0.903	1rbt	A	155

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.859	1ewq	B	765
1.000	1ewq	C	23
1.000	1ewq	D	22
0.724	1ewx	A	145
0.828	1ewy	A	303
0.848	1ewy	B	303
0.469	1ewy	C	98
0.711	1eye	A	280
0.894	1ez0	A	510
0.894	1ez0	B	510
0.894	1ez0	C	510
0.894	1ez0	D	510
0.722	1ez9	A	370
0.703	1ez9	B	370
0.892	1eza	A	259
0.699	1ezk	A	153
0.754	1ezo	A	370
0.874	1ezz	A	310
0.732	1ezz	B	153
0.835	1ezz	C	310
0.745	1ezz	D	153
0.566	1f0k	A	364
0.577	1f0k	B	364
0.863	1f2o	A	284
0.833	1f3a	A	222
0.829	1f3a	B	222
0.676	1f4a	A	1021
0.676	1f4a	B	1021
0.676	1f4a	C	1021
0.676	1f4a	D	1021
0.846	1f7b	A	293
0.877	1f7b	C	293
0.830	1f8m	A	429
0.830	1f8m	B	429
0.832	1f8m	C	429
0.839	1f8m	D	429
0.817	1f8r	A	498
0.817	1f8r	B	498
0.817	1f8r	C	498
0.817	1f8r	D	498
0.786	1fa9	A	846

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.829	1rco	B	475
0.919	1rco	C	123
0.829	1rco	E	475
0.919	1rco	F	123
0.829	1rco	H	475
0.919	1rco	I	123
0.829	1rco	K	475
0.829	1rco	L	475
0.919	1rco	M	123
0.829	1rco	O	475
0.919	1rco	P	123
0.829	1rco	R	475
0.919	1rco	S	123
0.919	1rco	T	123
0.829	1rco	V	475
0.919	1rco	W	123
0.821	1rev	A	560
0.768	1rev	B	440
0.735	1rka	A	309
0.790	1rrg	A	181
0.729	1rrg	B	181
0.779	1rt3	A	560
0.766	1rt3	B	440
0.786	1rtj	A	560
0.786	1rtj	B	440
0.701	1rv5	A	244
0.660	1rv5	B	244
1.000	1rv5	C	11
1.000	1rv5	D	11
0.770	1rvb	A	244
0.770	1rvb	B	244
1.000	1rvb	C	11
1.000	1rvb	D	11
0.730	1rx7	A	159
0.748	1rx8	A	159
0.705	1s01	A	275
0.606	1sky	B	502
0.567	1sky	E	473
0.791	1sma	A	588
0.791	1sma	B	588
0.891	1som	A	543

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.457	1fbn	A	230
0.757	1fc5	A	411
0.686	1fc5	B	411
0.798	1fcj	A	322
0.798	1fcj	B	322
0.801	1fcj	C	322
0.804	1fcj	D	322
0.879	1ffy	A	917
1.000	1ffy	T	75
0.711	1fg4	A	149
0.718	1fg4	B	149
0.842	1fgh	A	754
0.690	1fj4	A	406
0.690	1fj4	B	406
0.690	1fj4	C	406
0.697	1fj4	D	406
0.697	1fj8	A	406
0.680	1fj8	B	406
0.687	1fj8	C	406
0.687	1fj8	D	406
0.789	1fjx	A	327
1.000	1fjx	C	13
1.000	1fjx	D	13
0.877	1fid	A	138
0.826	1fmj	A	351
0.826	1fmj	B	351
0.846	1fmw	A	761
0.892	1fnc	A	314
0.598	1fnm	A	691
0.553	1fo5	A	85
0.511	1fok	A	568
1.000	1fok	B	20
1.000	1fok	C	20
0.844	1fp1	D	372
0.703	1fqa	A	370
0.798	1qf	A	331
0.895	1frn	A	314
0.759	1fs4	A	842
0.793	1fsf	A	266
0.759	1fu7	A	842
0.823	1fua	A	215

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.796	1sub	A	275
0.516	1sur	A	215
0.720	1tah	A	318
0.720	1tah	B	318
0.720	1tah	C	318
0.720	1tah	D	318
0.663	1tcd	A	249
0.651	1tcd	B	249
0.823	1tde	A	316
0.706	1tho	A	109
0.506	1tim	A	247
0.506	1tim	B	247
0.838	1tip	A	191
0.822	1tip	B	191
0.470	1toa	A	313
0.476	1toa	B	313
0.952	1tpd	A	250
0.952	1tpd	B	250
0.834	1tpt	A	440
0.748	1ttt	A	405
0.753	1ttt	B	405
0.741	1ttt	C	405
1.000	1ttt	D	77
1.000	1ttt	E	77
1.000	1ttt	F	77
0.875	1tyc	A	319
0.539	1tyf	A	193
0.544	1tyf	B	193
0.549	1tyf	C	193
0.539	1tyf	D	193
0.539	1tyf	E	193
0.554	1tyf	F	193
0.554	1tyf	G	193
0.549	1tyf	H	193
0.549	1tyf	I	193
0.560	1tyf	J	193
0.544	1tyf	K	193
0.554	1tyf	L	193
0.549	1tyf	M	193
0.549	1tyf	N	193
0.867	1udb	A	338

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.519	1fvp	A	231
0.519	1fvp	B	231
0.860	1fwb	A	100
0.877	1fwb	B	106
0.790	1fwb	C	567
0.870	1fwi	A	100
0.887	1fwi	B	106
0.783	1fwi	C	567
0.944	1fww	A	267
0.936	1fww	B	267
0.529	1fx0	A	507
0.578	1fx0	B	498
0.587	1fix	A	482
0.823	1g18	A	350
0.849	1g1h	A	298
1.000	1g1h	B	13
0.857	1g3q	A	237
0.861	1g4h	A	296
0.699	1g5p	A	289
0.751	1g5p	B	289
0.831	1g66	A	207
0.791	1g7t	A	594
0.796	1g7v	A	284
0.847	1g8f	A	511
0.853	1g98	A	558
0.871	1g98	B	558
0.874	1g9u	A	454
0.839	1ga8	A	311
0.730	1gad	O	330
0.730	1gad	P	330
0.809	1gaj	A	257
0.555	1gc5	A	467
0.791	1gc9	A	345
0.718	1gcy	A	309
0.805	1gcy	A	527
0.864	1gdd	A	353
0.429	1gdr	A	140
0.724	1gef	A	123
0.821	1gef	B	123
0.772	1gef	D	123
0.821	1gef	E	123

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.931	1ukz	A	203
0.910	1ula	A	289
0.384	1vdc	A	333
0.826	1vfe	A	333
0.616	1vfn	A	281
0.864	1vid	A	221
0.828	1vom	A	762
0.800	1vru	A	560
0.780	1vru	B	440
0.890	1vrx	A	537
0.879	1waf	A	903
0.883	1waf	B	903
0.829	1wal	A	345
0.837	1wkd	A	386
0.839	1wke	A	386
0.627	1wod	A	233
0.577	1xib	A	388
0.577	1xih	A	388
0.784	1xlm	A	393
0.784	1xlm	B	393
0.784	1xlm	C	393
0.784	1xlm	D	393
0.827	1xlf	A	394
0.827	1xlf	B	394
0.501	1xys	A	347
0.822	1xza	A	214
0.804	1xzd	A	214
0.804	1xzl	A	214
0.866	1ypn	A	313
0.608	1yub	A	245
0.823	1zrn	A	232
0.795	1zym	A	258
0.795	1zym	B	258
0.805	2ctb	A	307
0.913	2dhe	A	310
0.913	2edc	A	310
0.817	2gsq	A	202
0.829	2gtu	A	217
0.820	2gtu	B	217
0.561	2hgs	A	474
0.712	2hls	A	312

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.868	1gey	A	356
0.854	1gfs	A	321
0.941	1gg2	A	353
0.876	1gg2	B	340
0.930	1gg2	G	71
0.759	1gg8	A	842
0.868	1gin	A	431
0.805	1git	A	353
0.838	1gjp	A	340
0.840	1giw	A	637
0.788	1gl9	B	1054
0.767	1gl9	C	1054
0.806	1glf	O	501
0.806	1glf	X	501
0.780	1glf	Y	501
0.788	1glf	Z	501
0.824	1glj	O	501
0.816	1glj	Y	501
0.657	1gnp	A	166
0.651	1gnq	A	166
0.861	1goo	A	303
0.686	1got	A	350
0.762	1got	B	340
0.671	1got	G	73
0.634	1gpj	A	404
0.512	1gpm	A	525
0.507	1gpm	B	525
0.520	1gpm	C	525
0.507	1gpm	D	525
0.763	1gpw	A	253
0.697	1gpw	B	201
0.783	1gpw	C	253
0.726	1gpw	D	201
0.775	1gpw	E	253
0.692	1gpw	F	201
0.920	1gq6	A	313
0.920	1gq6	B	313
0.920	1gq6	C	313
0.835	1gra	A	478
0.876	1grr	A	178
0.765	1grx	A	85

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.957	2hsd	A	253
0.957	2hsd	B	253
0.957	2hsd	C	253
0.957	2hsd	D	253
0.686	2mys	A	843
0.410	2mys	B	166
0.383	2mys	C	149
0.856	2pfl	A	759
0.856	2pfl	B	759
0.932	2pgd	A	482
0.461	2pia	A	321
0.591	2pmt	A	203
0.591	2pmt	B	203
0.591	2pmt	C	203
0.591	2pmt	D	203
0.815	2puf	A	340
1.000	2puf	B	17
0.857	2rkm	A	517
1.000	2rkm	B	2
0.809	2sfp	A	388
0.825	2sfp	B	388
0.566	2shp	A	525
0.571	2shp	B	525
0.749	2sic	E	275
0.916	2sic	I	107
0.917	2tnk	A	216
0.912	2tnk	B	216
0.780	2wsy	A	268
0.763	2wsy	B	396
0.737	2xyl	A	312
0.825	2yas	A	257
0.844	3cox	A	507
0.847	3dap	A	320
0.856	3dap	B	320
0.863	3eza	A	249
0.471	3eza	B	85
0.811	3gtu	A	217
0.705	3gtu	B	224
0.811	3gtu	C	217
0.705	3gtu	D	224
0.547	3hsc	A	386

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.702	1gsj	A	258
0.742	1gtm	A	419
0.733	1gtm	B	419
0.733	1gtm	C	419
0.805	1gtr	A	553
1.000	1gtr	B	74
0.817	1gts	A	553
1.000	1gts	B	74
0.801	1gtz	A	156
0.801	1gtz	B	156
0.795	1gtz	C	156
0.801	1gtz	D	156
0.795	1gtz	E	156
0.814	1gtz	F	156
0.814	1gtz	G	156
0.808	1gtz	H	156
0.814	1gtz	I	156
0.801	1gtz	J	156
0.801	1gtz	K	156
0.801	1gtz	L	156
0.563	1guy	A	309
0.573	1guy	C	309
0.517	1gvf	A	286
0.531	1gvf	B	286
0.610	1gwj	A	377
0.410	1gzu	A	290
0.410	1gzu	B	290
0.410	1gzu	C	290
0.894	1h11	A	303
0.840	1h1l	A	481
0.829	1h1l	B	519
0.850	1h1l	C	481
0.813	1h1l	D	519
0.821	1h4z	A	117
0.870	1h5s	A	293
0.857	1h5s	B	293
0.860	1h5s	C	293
0.867	1h5s	D	293
0.871	1h6l	A	364
0.886	1h77	A	605
0.901	1h78	A	605

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.885	3ljr	A	244
0.885	3ljr	B	244
0.887	3man	A	302
0.835	3mht	A	327
1.000	3mht	C	12
1.000	3mht	D	13
0.850	3pbg	A	468
0.850	3pbg	B	468
0.801	3pgm	A	241
0.872	3ptd	A	298
0.860	3tec	E	279
0.886	3tec	I	70
0.396	3thi	A	371
0.857	3trx	A	105
0.786	3xim	A	393
0.786	3xim	B	393
0.786	3xim	C	393
0.786	3xim	D	393
0.779	4cpa	A	307
1.000	4cpa	G	1
0.316	4cpa	I	38
0.807	4dmr	A	823
0.835	4gr1	A	478
0.825	4gst	A	217
0.834	4gst	B	217
0.743	4mbp	A	370
0.793	4mdh	A	333
0.793	4mdh	B	333
0.489	4pbg	A	468
0.489	4pbg	B	468
0.588	4pga	A	337
0.573	4pga	B	337
0.810	4pgt	A	210
0.833	4pgt	B	210
0.656	4q21	A	189
0.861	4ukd	A	194
0.724	5gpb	A	842
0.870	5nul	A	138
0.996	5tgi	A	269
0.444	5tim	A	250
0.468	5tim	B	250

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.567	1h7f	A	245
0.559	1h7f	B	245
0.875	1h7l	A	255
0.842	1h7p	A	342
0.850	1h7w	A	1025
0.848	1h7w	B	1025
0.844	1h7w	C	1025
0.849	1h7w	D	1025
0.850	1h82	A	472
0.856	1h82	B	472
0.850	1h82	C	472
0.833	1h86	A	472
0.845	1h86	B	472
0.831	1h86	C	472
0.830	1hdo	A	206
0.730	1hdq	A	307
0.783	1hdy	A	374
0.783	1hdy	B	374
0.660	1het	A	374
0.660	1het	B	374
0.342	1hex	A	345
0.901	1hf6	A	303
0.862	1hfj	A	327
0.862	1hfj	C	327
0.869	1hfk	A	327
0.856	1hfk	C	327
0.869	1hg0	A	327

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.920	5tmp	A	213
0.822	5xia	A	393
0.822	5xia	B	393
0.769	6cpa	A	307
0.686	6dfr	A	159
0.724	6gpb	A	842
0.786	6xim	A	393
0.786	6xim	B	393
0.786	6xim	C	393
0.786	6xim	D	393
0.888	7enl	A	436
0.828	7gss	A	209
0.828	7gss	B	209
0.864	7req	A	727
0.863	7req	B	637
0.867	7req	C	727
0.865	7req	D	637
0.829	8ruc	A	475
0.829	8ruc	C	475
0.829	8ruc	E	475
0.829	8ruc	G	475
0.919	8ruc	I	123
0.919	8ruc	J	123
0.919	8ruc	K	123
0.919	8ruc	L	123
0.693	9rub	A	466
0.693	9rub	B	466

APPENDIX R

Prediction accuracies for each chain of each protein in $\alpha+\beta$ class in trial set

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.756	108l	A	164
0.750	127l	A	164
0.750	130l	A	164
0.473	135l	A	129
0.768	138l	A	164
0.756	145l	A	164
0.756	149l	A	164
0.865	154l	A	185
0.774	159l	A	164
0.768	162l	A	164
0.665	173l	A	164
0.707	175l	A	164
0.707	175l	B	164
0.738	176l	A	164
0.738	176l	B	164
0.726	177l	A	164
0.726	182l	A	164
0.671	190l	A	164
0.674	193l	A	129
0.560	1a0l	A	348
0.666	1a0p	A	290
0.461	1a16	A	440
0.000	1a16	L	2
0.482	1a48	A	305
0.910	1a5e	A	156
0.594	1a6q	A	382
0.916	1a7x	A	107
0.916	1a7x	B	107
0.859	1a9e	A	277
0.850	1a9e	B	100
0.333	1a9e	C	9
0.836	1aa8	A	347
0.862	1aa8	B	347
0.720	1acf	A	125
0.739	1ady	A	421
0.739	1ady	B	421

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.790	1hhj	E	100
1.000	1hhj	F	9
0.874	1hl2	A	135
0.840	1hs6	A	611
0.768	1hsa	A	276
0.768	1hsa	B	99
0.222	1hsa	C	9
0.768	1hsa	D	276
0.768	1hsa	E	99
0.222	1hsa	F	9
0.646	1ht1	A	175
0.703	1ht1	B	175
0.754	1ht1	C	175
0.674	1ht1	D	175
0.724	1ht1	E	449
0.699	1ht1	F	449
0.722	1ht1	G	449
0.722	1ht1	I	449
0.726	1ht1	V	175
0.691	1ht1	X	175
0.691	1ht1	Y	175
0.651	1ht1	Z	175
0.473	1htn	A	182
0.866	1hwm	A	254
0.598	1hwm	B	266
0.789	1hx3	A	190
0.795	1hx3	B	190
0.843	1hyg	A	313
0.840	1hyg	B	313
0.596	1hyn	P	379
0.583	1hyn	Q	379
0.578	1hyn	R	379
0.596	1hyn	S	379
0.826	1hyt	A	316
0.618	1hyw	A	68
0.731	1l0x	A	104

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.734	1ady	C	421
0.734	1ady	D	421
0.685	1afk	A	124
0.645	1afk	B	124
0.862	1age	A	276
0.828	1age	B	99
0.875	1age	C	8
0.898	1ah5	A	313
0.904	1ai5	A	209
0.916	1ai5	B	557
0.451	1ako	A	268
0.528	1alc	A	123
0.701	1amw	A	214
0.564	1aor	A	605
0.557	1aor	B	605
0.795	1ar0	A	127
0.803	1ar0	B	127
0.743	1aud	A	101
1.000	1aud	B	30
0.875	1ay7	A	96
0.551	1ay7	B	89
0.675	1ayz	A	169
0.675	1ayz	B	169
0.680	1ayz	C	169
0.455	1azq	A	66
1.000	1azq	B	8
1.000	1azq	C	8
0.782	1b27	A	110
0.773	1b27	B	110
0.736	1b27	C	110
0.489	1b27	D	90
0.589	1b27	E	90
0.478	1b27	F	90
0.403	1b2t	A	77
0.700	1b2x	A	110
0.700	1b2x	B	110
0.727	1b2x	C	110
0.862	1b37	A	472
0.841	1b37	B	472
0.854	1b37	C	472
0.773	1b3s	A	110

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.721	1i0x	B	104
0.673	1i0x	C	104
0.731	1i0x	D	104
0.828	1i32	A	360
0.828	1i32	B	360
0.825	1i32	C	360
0.833	1i32	D	360
0.831	1i32	E	360
0.822	1i32	F	360
0.721	1i3i	A	104
0.745	1i4f	A	275
0.790	1i4f	B	100
1.000	1i4f	C	10
0.912	1i4x	A	239
0.585	1i6s	A	164
0.577	1i6u	A	130
0.569	1i6u	B	130
1.000	1i6u	C	37
1.000	1i6u	D	37
0.833	1i8v	A	96
0.802	1i8v	B	96
0.731	1i9a	A	182
0.687	1i9a	B	182
0.910	1iab	A	200
0.910	1iac	A	200
0.481	1ian	A	366
0.705	1iez	A	217
0.856	1ifu	A	263
0.738	1igc	A	61
0.689	1igc	H	222
0.883	1igc	L	213
0.778	1ihy	A	333
0.778	1ihy	B	333
0.778	1ihy	C	333
0.778	1ihy	D	333
0.495	1ii7	A	333
0.495	1ii7	B	333
0.550	1ikt	A	120
0.854	1il2	A	590
0.861	1il2	B	590
1.000	1il2	C	75

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.736	1b3s	B	110
0.755	1b3s	C	110
0.567	1b3s	D	90
0.667	1b3s	E	90
0.556	1b3s	F	90
0.528	1b49	A	246
0.528	1b49	C	246
0.910	1b4n	A	619
0.926	1b4n	B	619
0.924	1b4n	C	619
0.922	1b4n	D	619
0.541	1b5d	A	246
0.545	1b5d	B	246
0.528	1b5e	A	246
0.520	1b5e	B	246
0.750	1b6v	A	124
0.750	1b6v	B	124
0.754	1b7o	A	130
0.909	1b7y	A	350
0.766	1b7y	B	785
0.849	1b8s	A	504
0.869	1b9c	A	236
0.864	1b9c	B	236
0.873	1b9c	C	236
0.856	1b9c	D	236
0.626	1b9o	A	123
0.429	1b9r	A	105
0.738	1bb3	A	130
0.746	1bb3	B	130
0.690	1bb7	A	129
0.850	1bc2	A	227
0.841	1bc2	B	227
0.381	1bf4	A	63
1.000	1bf4	B	8
1.000	1bf4	C	8
0.821	1bfj	A	112
0.676	1bgq	A	225
0.887	1bi8	A	326
0.855	1bi8	B	166
0.887	1bi8	C	326
0.855	1bi8	D	166

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
1.000	1li2	D	75
0.809	1li9	A	267
0.708	1liq	A	72
0.694	1liq	B	72
1.000	1liq	C	18
0.855	1lnq	A	275
0.808	1lnq	B	99
1.000	1lnq	C	9
0.785	1lnu	A	130
0.754	1ip5	A	130
0.935	1ipb	A	217
0.766	1ir1	A	475
0.764	1ir1	B	475
0.766	1ir1	C	475
0.764	1ir1	D	475
0.862	1ir1	S	123
0.862	1ir1	T	123
0.862	1ir1	U	123
0.862	1ir1	V	123
0.543	1ir7	A	129
0.455	1itp	A	77
0.865	1ius	A	394
0.851	1ivw	A	638
0.842	1ivw	B	638
0.594	1j6x	A	160
0.600	1j6x	B	160
0.892	1j9o	A	93
0.464	1jaw	A	440
0.750	1jc4	A	148
0.777	1jc4	B	148
0.770	1jc4	C	148
0.757	1jc4	D	148
0.886	1jdf	A	446
0.888	1jdf	B	446
0.888	1jdf	C	446
0.888	1jdf	D	446
0.629	1je3	A	97
0.682	1jef	A	129
0.362	1jfm	A	174
0.356	1jfm	B	174
0.391	1jfm	C	174

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.760	1bia	A	321
0.864	1bid	A	265
0.829	1bj3	A	129
0.935	1bj3	B	123
0.832	1bkf	A	107
0.467	1bkj	A	240
0.467	1bkj	B	240
0.822	1bms	A	129
0.845	1bms	B	129
0.837	1bms	C	129
0.925	1bn5	A	478
0.700	1bni	A	110
0.736	1bni	B	110
0.691	1bni	C	110
0.613	1bo0	A	80
0.512	1bo4	A	168
0.506	1bo4	B	168
0.851	1bo7	A	316
0.814	1bp1	A	456
0.512	1bql	H	215
0.557	1bql	L	212
0.698	1bql	Y	129
0.843	1br6	A	268
0.736	1brj	A	110
0.700	1brj	B	110
0.700	1brj	C	110
0.636	1bsc	A	110
0.636	1bsc	B	110
0.636	1bsc	C	110
0.970	1bsm	A	201
0.965	1bsm	B	201
0.805	1buh	A	298
0.684	1buh	B	79
0.659	1bvx	A	129
0.831	1bwe	A	77
0.761	1by8	A	314
0.936	1byf	A	125
0.936	1byf	B	125
0.808	1c0g	A	375
0.921	1c0g	S	127
0.871	1c0p	A	363

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.374	1jfm	D	174
0.397	1jfm	E	174
0.617	1jid	A	128
1.000	1jid	B	29
0.589	1jis	A	129
0.486	1jk8	A	181
0.458	1jk8	B	190
0.357	1jk8	C	14
0.841	1jkl	A	533
0.831	1jkl	B	533
0.649	1jkj	A	288
0.580	1jkj	B	388
0.646	1jkj	D	288
0.572	1jkj	E	388
0.829	1jmg	A	316
0.854	1jmh	A	316
0.838	1jnd	A	420
0.682	1jpo	A	129
0.467	1jql	A	366
0.800	1jql	B	140
0.506	1jru	A	89
0.912	1jrz	A	571
0.914	1jrz	B	571
0.669	1js0	A	124
0.669	1js0	B	124
0.669	1js0	C	124
0.591	1jwh	A	337
0.579	1jwh	B	337
0.623	1jwh	C	215
0.623	1jwh	D	215
0.842	1jx0	A	222
0.878	1jx0	B	222
0.805	1jy8	A	159
0.840	1k0i	A	394
0.830	1k0j	A	394
0.464	1k0k	A	125
0.872	1k2e	A	156
0.833	1k2e	B	156
0.763	1k3t	A	359
0.758	1k3t	B	359
0.755	1k3t	C	359

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.814	1c22	A	263
0.632	1c3t	A	76
0.792	1c54	A	96
0.579	1c62	A	164
0.585	1c63	A	164
0.585	1c6b	A	164
0.579	1c6j	A	164
0.585	1c6k	A	164
0.573	1c6t	A	164
0.809	1c7h	A	131
0.621	1ca5	A	66
1.000	1ca5	B	8
1.000	1ca5	C	8
0.561	1ca6	A	66
1.000	1ca6	B	8
1.000	1ca6	C	8
0.863	1cc7	A	73
0.891	1ce6	A	274
0.491	1ce6	B	108
1.000	1ce6	C	9
0.816	1cf3	A	583
0.494	1civ	A	385
0.746	1cj9	A	130
0.754	1ckd	A	130
0.877	1cmk	E	350
0.682	1cmk	I	22
0.695	1cns	A	243
0.695	1cns	B	243
0.949	1cnu	A	137
0.852	1coy	A	507
0.515	1cpz	A	68
0.584	1cqm	A	101
0.584	1cqm	B	101
0.574	1cqn	A	101
0.594	1cqn	B	101
0.912	1cqv	A	239
0.897	1ctp	E	350
0.700	1ctp	I	20
0.579	1ctw	A	164
0.585	1cu3	A	164
0.622	1cu5	A	164

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.774	1k3t	D	359
0.854	1k4p	A	233
0.460	1k50	A	63
0.444	1k50	B	63
0.524	1k50	C	63
0.444	1k50	D	63
0.653	1k52	A	72
0.653	1k52	B	72
0.639	1k53	A	72
0.639	1k53	B	72
0.817	1k5b	A	120
0.482	1k6z	A	141
0.496	1k6z	B	141
0.539	1k7j	A	206
0.843	1k8d	A	274
0.808	1k8d	B	99
0.556	1k8d	P	9
0.385	1k8w	A	327
1.000	1k8w	B	22
0.693	1kbi	A	511
0.802	1kbi	B	511
0.833	1kbl	A	873
0.602	1kbu	A	349
0.599	1kbu	B	349
1.000	1kbu	C	34
1.000	1kbu	D	34
0.792	1kd0	A	413
0.804	1kd0	B	413
0.772	1ke9	A	298
0.518	1keh	A	689
0.685	1kf4	A	124
0.685	1kf7	A	124
0.435	1kqh	A	317
0.872	1kmy	A	297
0.872	1knd	A	297
0.761	1knk	A	159
0.516	1ko9	A	345
0.908	1kp2	A	455
0.344	1ksi	A	642
0.340	1ksi	B	642
0.885	1ktd	A	182

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.646	1cx6	A	164
0.640	1cx7	A	164
0.389	1cz4	A	185
0.400	1cz5	A	185
0.813	1d09	A	310
0.732	1d09	B	153
0.784	1d09	C	310
0.771	1d09	D	153
0.525	1d1j	A	139
0.525	1d1j	B	139
0.525	1d1j	C	139
0.518	1d1j	D	139
0.591	1d2w	A	164
0.825	1d4d	A	572
0.845	1d5x	A	181
0.781	1d5x	B	192
0.870	1d5x	C	239
1.000	1d5x	D	5
0.879	1d6o	A	107
0.879	1d6o	B	107
0.879	1d7l	A	107
0.879	1d7l	B	107
0.893	1d9n	A	75
0.579	1d9w	A	164
0.853	1dao	A	347
0.850	1dao	B	347
0.844	1dao	C	347
0.853	1dao	D	347
0.850	1dao	E	347
0.859	1dao	F	347
0.859	1dao	G	347
0.850	1dao	H	347
0.641	1dax	A	64
0.795	1dbf	A	127
0.850	1dbf	B	127
0.803	1dbf	C	127
0.885	1dch	A	104
0.942	1dch	B	104
0.894	1dch	C	104
0.942	1dch	D	104
0.885	1dch	E	104

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.484	1ktd	B	215
0.885	1ktd	C	182
0.474	1ktd	D	215
0.833	1ktg	A	138
0.826	1ktg	B	138
0.473	1kv3	A	687
0.475	1kv3	B	687
0.475	1kv3	C	687
0.473	1kv3	D	687
0.477	1kv3	E	687
0.489	1kv3	F	687
0.430	1kvk	A	395
0.674	1kxy	A	129
0.871	1kzj	A	264
0.867	1kzj	B	264
0.864	1kzj	C	264
0.875	1kzj	D	264
0.879	1kzj	E	264
0.879	1kzj	F	264
0.646	1l17	A	164
0.646	1l31	A	164
0.652	1l36	A	164
0.652	1l43	A	164
0.616	1l44	A	164
0.671	1l61	A	164
0.652	1l65	A	164
0.646	1l67	A	164
0.665	1l69	A	164
0.652	1l76	A	164
0.810	1l8t	A	263
0.665	1l91	A	164
0.903	1l9n	A	692
0.899	1l9n	B	692
0.585	1laa	A	130
0.562	1lba	A	146
0.723	1lbu	A	213
0.872	1lek	A	274
0.778	1lek	B	99
1.000	1lek	P	8
0.600	1leo	A	150
0.573	1lgw	A	164

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.942	1dch	F	104
0.933	1dch	G	104
0.942	1dch	H	104
0.395	1dcj	A	81
0.962	1dcp	A	104
0.971	1dcp	B	104
0.952	1dcp	C	104
0.971	1dcp	D	104
0.952	1dcp	E	104
0.962	1dcp	F	104
0.971	1dcp	G	104
0.962	1dcp	H	104
0.845	1ddu	A	265
0.860	1ddu	B	265
0.438	1dfd	A	64
0.846	1dgs	A	667
0.861	1dgs	B	667
0.446	1dhm	A	83
0.446	1dhm	B	83
0.544	1div	A	149
0.674	1dkk	A	129
0.667	1dkk	B	129
0.789	1dm2	A	298
0.466	1dm9	A	133
0.481	1dm9	B	133
0.847	1dmn	A	131
0.847	1dmq	A	131
0.927	1dnk	A	260
1.000	1dnk	B	7
1.000	1dnk	C	8
0.521	1do9	A	94
0.854	1doy	A	96
0.885	1ds7	A	217
0.853	1ds7	B	217
0.837	1dth	A	203
0.842	1dth	B	203
0.735	1duz	A	275
0.790	1duz	B	100
1.000	1duz	C	9
0.742	1duz	D	275
0.790	1duz	E	100

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.677	1lhj	A	130
0.358	1li1	A	229
0.358	1li1	B	229
0.395	1li1	C	228
0.358	1li1	D	229
0.358	1li1	E	229
0.395	1li1	F	228
0.573	1li6	A	164
0.825	1lk7	A	229
0.812	1lk7	B	229
0.817	1lk7	C	229
0.821	1lk7	D	229
0.774	1lld	A	319
0.774	1lld	B	319
0.870	1lnf	E	316
1.000	1lnf	P	2
0.854	1lni	A	96
0.844	1lni	B	96
0.858	1lo9	A	141
0.721	1low	A	104
0.746	1loz	A	130
0.709	1lpf	A	477
0.709	1lpf	B	477
0.820	1lqw	A	183
0.820	1lqw	B	183
0.864	1lsy	A	147
0.526	1lvi	A	458
0.671	1lyg	A	164
0.671	1lyh	A	164
0.671	1lyi	A	164
0.708	1lz1	A	130
0.679	1lz5	A	134
0.698	1lzd	A	129
0.698	1lzg	A	129
0.829	1m44	A	181
0.829	1m44	B	181
0.886	1mdt	A	535
0.886	1mdt	B	535
0.583	1meg	A	216
0.623	1mff	A	114
0.623	1mff	B	114

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
1.000	1duz	F	9
0.626	1dv2	A	452
0.631	1dv2	B	452
0.589	1dxi	A	470
0.602	1dxi	B	470
0.602	1dxi	C	470
0.604	1dxi	D	470
0.622	1dyb	A	164
0.634	1dyg	A	164
0.914	1e0z	A	128
0.815	1E28	A	276
0.768	1E28	B	99
0.000	1E28	C	8
0.423	1e2t	A	284
0.408	1e2t	B	284
0.430	1e2t	C	284
0.440	1e2t	D	284
0.412	1e2t	E	284
0.433	1e2t	F	284
0.423	1e2t	G	284
0.423	1e2t	H	284
0.827	1e3h	A	757
0.893	1e3n	A	131
0.885	1e3n	B	131
0.856	1e5l	A	450
0.853	1e5l	B	450
0.720	1e5w	A	346
0.823	1e6z	A	498
0.827	1e6z	B	498
0.938	1E77	A	485
0.594	1e7k	A	128
0.602	1e7k	B	128
1.000	1e7k	C	22
1.000	1e7k	D	22
0.814	1e7x	A	129
0.860	1e7x	B	129
0.806	1e7x	C	129
1.000	1e7x	R	19
1.000	1e7x	S	19
0.921	1e7y	A	485
0.667	1e8l	A	129

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.623	1mff	C	114
0.623	1mfi	A	114
0.623	1mfi	B	114
0.623	1mfi	C	114
0.726	1mi2	A	73
0.726	1mi2	B	73
0.971	1mkb	A	171
0.988	1mkb	B	171
0.607	1mms	A	140
0.643	1mms	B	140
1.000	1mms	C	58
1.000	1mms	D	58
0.667	1mns	A	357
0.610	1mom	A	246
0.855	1mra	A	359
0.814	1mrk	A	247
0.792	1msh	A	72
0.736	1msh	B	72
0.791	1mvb	A	129
0.837	1mvb	B	129
0.798	1mvb	C	129
0.888	1mxb	A	383
0.783	1naw	A	419
0.795	1naw	B	419
0.323	1ndc	A	155
0.323	1ndk	A	155
0.323	1ndp	A	155
0.323	1ndp	B	155
0.812	1nhr	A	447
0.842	1njb	A	316
0.739	1nsq	A	153
0.556	1nsq	B	153
0.556	1nsq	C	153
0.799	1oac	A	727
0.799	1oac	B	727
0.769	1ouc	A	130
0.778	1pad	A	212
0.000	1pad	I	4
0.722	1pbh	A	317
0.563	1pfd	A	96
0.871	1pfi	A	139

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.912	1e9z	A	238
0.828	1e9z	B	569
0.904	1ea6	A	364
0.934	1ea6	B	364
0.571	1eb6	A	177
0.782	1ec6	A	87
0.793	1ec6	B	87
1.000	1ec6	C	20
1.000	1ec6	D	20
0.890	1ec8	A	446
0.883	1ec8	B	446
0.892	1ec8	C	446
0.883	1ec8	D	446
0.767	1ed3	A	275
0.677	1ed3	B	99
0.846	1ed3	C	13
0.775	1ed3	D	275
0.717	1ed3	E	99
0.846	1ed3	F	13
0.339	1ef7	A	242
0.351	1ef7	B	242
0.494	1egw	A	77
0.481	1egw	B	77
0.481	1egw	C	77
0.506	1egw	D	77
1.000	1egw	E	17
1.000	1egw	F	17
1.000	1egw	G	17
1.000	1egw	H	17
0.680	1ely	A	350
0.532	1ely	B	785
1.000	1ely	C	76
0.890	1ejv	A	100
0.782	1ejv	B	101
0.802	1ejv	C	567
0.870	1ekr	A	161
0.843	1ema	A	236
0.877	1emb	A	236
0.928	1emk	A	237
0.543	1ems	A	440
0.534	1ems	B	440

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.897	1pfm	A	68
0.897	1pfm	B	68
0.897	1pfm	C	68
0.897	1pfm	D	68
0.539	1pme	A	380
0.569	1ppo	A	216
0.620	1prt	A	234
0.791	1prt	B	196
0.765	1prt	C	196
0.973	1prt	D	110
0.973	1prt	E	110
0.847	1prt	F	98
0.620	1prt	G	234
0.791	1prt	H	196
0.765	1prt	I	196
0.973	1prt	J	110
0.973	1prt	K	110
0.847	1prt	L	98
0.853	1pxa	A	394
0.903	1pys	A	350
0.561	1pys	B	785
0.500	1qd9	A	124
0.508	1qd9	B	124
0.500	1qd9	C	124
0.896	1qg7	A	67
0.851	1qg7	B	67
0.538	1qh5	A	260
0.892	1qh5	B	260
0.830	1qjd	A	571
0.768	1qjg	A	125
0.768	1qjg	B	125
0.744	1qjg	C	125
0.760	1qjg	D	125
0.736	1qjg	E	125
0.768	1qjg	F	125
0.555	1qjj	A	200
0.750	1qjj	B	4
0.533	1qk9	A	92
0.453	1qlm	A	316
0.854	1qlt	A	560
0.850	1qlt	B	560

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.863	1en6	A	205
0.878	1en6	B	205
0.893	1en6	C	205
0.898	1en6	D	205
0.726	1enf	A	212
0.813	1epw	A	1287
0.731	1eqe	A	130
0.857	1erk	A	364
0.820	1esf	A	233
0.777	1esf	B	233
0.877	1esp	A	317
0.864	1ev5	A	264
0.833	1evx	A	162
0.833	1evx	B	162
0.734	1ewc	A	214
0.437	1ewj	A	126
0.437	1ewj	B	126
0.444	1ewj	C	126
0.468	1ewj	D	126
0.452	1ewj	E	126
0.437	1ewj	F	126
0.437	1ewj	G	126
0.444	1ewj	H	126
0.850	1eym	A	107
0.841	1eym	B	107
0.581	1f0w	A	129
0.784	1f1h	A	468
0.784	1f1h	B	468
0.784	1f1h	C	468
0.784	1f1h	D	468
0.784	1f1h	E	468
0.784	1f1h	F	468
0.784	1f1h	G	468
0.784	1f1h	H	468
0.784	1f1h	I	468
0.784	1f1h	J	468
0.784	1f1h	K	468
0.784	1f1h	L	468
0.828	1f29	A	215
0.828	1f29	B	215
0.828	1f29	C	215

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.561	1qm9	A	198
0.873	1qmi	A	347
0.867	1qmi	B	347
0.876	1qmi	C	347
0.853	1qmi	D	347
0.799	1qmr	A	159
0.775	1qn5	A	200
0.810	1qn5	B	200
1.000	1qn5	C	14
1.000	1qn5	D	14
1.000	1qn5	E	14
1.000	1qn5	F	14
0.775	1qn6	A	200
0.795	1qn6	B	200
1.000	1qn6	C	14
1.000	1qn6	D	14
1.000	1qn6	E	14
1.000	1qn6	F	14
0.775	1qn8	A	200
0.765	1qn8	B	200
1.000	1qn8	C	14
1.000	1qn8	D	14
1.000	1qn8	E	14
1.000	1qn8	F	14
1.000	1qnk	A	13
1.000	1qnk	B	13
0.874	1qnm	A	198
0.874	1qnm	B	198
0.831	1qqy	A	130
0.531	1qr0	A	228
0.670	1qr5	A	88
0.847	1qrk	A	731
0.840	1qrk	B	731
0.737	1qrm	A	274
0.790	1qrm	B	100
0.556	1qrm	C	9
0.735	1qrm	D	200
0.658	1qrm	E	243
0.819	1qsa	A	618
0.822	1qte	A	618
0.589	1qtk	A	129

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.847	1f2c	A	215
0.776	1f2g	A	58
0.536	1f2k	A	125
0.544	1f2k	B	125
0.395	1f2l	A	76
0.382	1f2l	B	76
0.382	1f2l	C	76
0.474	1f2l	D	76
0.826	1f3f	A	155
0.826	1f3f	B	155
0.826	1f3f	C	155
0.841	1f4c	A	264
0.848	1f4c	B	264
0.822	1f4f	A	264
0.875	1f4f	B	264
0.688	1f5v	A	240
0.688	1f5v	B	240
0.806	1f77	A	217
0.811	1f77	B	217
0.813	1f8r	A	498
0.813	1f8r	B	498
0.813	1f8r	C	498
0.817	1f8r	D	498
0.875	1f93	A	104
0.846	1f93	B	104
0.885	1f93	C	104
0.808	1f93	D	104
0.750	1f93	E	32
0.750	1f93	F	32
0.750	1f93	G	32
0.750	1f93	H	32
0.926	1fa6	A	135
0.933	1fa6	B	135
0.394	1fcd	A	401
0.394	1fcd	B	401
0.764	1fcd	C	174
0.764	1fcd	D	174
0.607	1fcl	A	56
0.488	1fd9	A	213
0.434	1ff2	A	106
0.865	1ffv	A	163

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.579	1qtz	A	164
0.556	1quh	A	162
0.568	1quo	A	162
0.913	1rax	A	115
0.734	1rbj	A	124
1.000	1rbj	B	4
0.834	1rco	B	475
0.984	1rco	C	123
0.834	1rco	E	475
0.984	1rco	F	123
0.834	1rco	H	475
0.984	1rco	I	123
0.834	1rco	K	475
0.834	1rco	L	475
0.984	1rco	M	123
0.834	1rco	O	475
0.984	1rco	P	123
0.834	1rco	R	475
0.984	1rco	S	123
0.984	1rco	T	123
0.834	1rco	V	475
0.984	1rco	W	123
0.829	1rcx	B	475
0.976	1rcx	C	123
0.829	1rcx	E	475
0.976	1rcx	F	123
0.829	1rcx	H	475
0.976	1rcx	I	123
0.829	1rcx	K	475
0.829	1rcx	L	475
0.976	1rcx	M	123
0.829	1rcx	O	475
0.976	1rcx	P	123
0.829	1rcx	R	475
0.976	1rcx	S	123
0.976	1rcx	T	123
0.829	1rcx	V	475
0.976	1rcx	W	123
0.738	1rez	A	130
0.654	1rgl	A	104
0.573	1rha	A	124

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.873	1ffv	B	803
0.829	1ffv	C	287
0.871	1ffv	D	163
0.880	1ffv	E	803
0.836	1ffv	F	287
0.525	1fh0	A	221
0.525	1fh0	B	221
0.333	1fh0	D	3
0.333	1fh0	E	3
0.871	1fik	A	139
0.614	1fim	A	114
0.816	1fit	A	147
0.826	1fjo	A	316
0.829	1fjt	A	316
0.826	1fjv	A	316
0.829	1fjw	A	316
0.897	1fkt	A	107
0.778	1fma	D	81
0.900	1fma	E	150
0.574	1fn5	A	129
0.799	1fpy	A	468
0.799	1fpy	B	468
0.799	1fpy	C	468
0.799	1fpy	D	468
0.799	1fpy	E	468
0.799	1fpy	F	468
0.799	1fpy	G	468
0.799	1fpy	H	468
0.799	1fpy	I	468
0.799	1fpy	J	468
0.799	1fpy	K	468
0.799	1fpy	L	468
0.860	1fwb	A	100
0.887	1fwb	B	106
0.802	1fwb	C	567
0.850	1fwc	A	100
0.887	1fwc	B	106
0.802	1fwc	C	567
0.850	1fwd	A	100
0.887	1fwd	B	106
0.794	1fwd	C	567

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.790	1rms	A	105
0.573	1mo	A	124
0.629	1mu	A	124
0.629	1mv	A	124
0.573	1my	A	124
0.792	1rod	A	72
0.792	1rod	B	72
0.753	1roe	A	97
0.573	1rsm	A	124
0.838	1rtn	A	68
0.838	1rtn	B	68
0.643	1sce	A	112
0.643	1sce	B	112
0.643	1sce	C	112
0.643	1sce	D	112
0.469	1sei	A	130
0.454	1sei	B	130
0.824	1sgk	A	535
0.876	1sib	E	275
0.543	1sib	I	70
0.546	1sml	A	269
0.862	1srp	A	471
0.857	1stc	E	350
0.700	1stc	I	20
0.868	1tbe	A	76
0.882	1tbe	B	76
0.585	1tby	A	130
0.585	1tcy	A	130
0.964	1tec	E	279
0.871	1tec	I	70
0.857	1tjs	A	265
0.857	1tlc	A	265
0.860	1tlc	B	265
0.886	1tli	A	316
0.551	1tsg	A	98
0.835	1tsx	A	316
0.433	1u2f	A	90
0.802	1ubp	A	101
0.582	1ubp	B	122
0.635	1ubp	C	570
0.638	1uia	A	127

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.786	1fxa	A	98
0.786	1fxa	B	98
0.947	1fxv	A	209
0.846	1fxv	B	557
0.869	1fzo	A	274
0.778	1fzo	B	99
1.000	1fzo	P	9
0.826	1g18	A	350
0.735	1g1q	A	162
0.741	1g1q	B	162
0.710	1g1q	C	162
0.753	1g1q	D	162
0.653	1g1x	A	98
0.875	1g1x	B	88
0.898	1g1x	C	88
1.000	1g1x	D	41
1.000	1g1x	E	44
0.622	1g1x	F	98
0.886	1g1x	G	88
0.920	1g1x	H	88
1.000	1g1x	I	41
1.000	1g1x	J	44
0.831	1g2s	A	71
0.786	1g3i	A	444
0.775	1g3i	B	444
0.770	1g3i	C	444
0.775	1g3i	D	444
0.777	1g3i	E	444
0.759	1g3i	F	444
0.730	1g3i	G	174
0.736	1g3i	H	174
0.753	1g3i	I	174
0.736	1g3i	J	174
0.736	1g3i	K	174
0.753	1g3i	L	174
0.655	1g3i	M	174
0.644	1g3i	N	174
0.644	1g3i	O	174
0.649	1g3i	P	174
0.649	1g3i	Q	174
0.649	1g3i	R	174

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.667	1ulf	A	129
0.674	1uig	A	129
0.659	1up1	A	182
0.960	1var	A	198
0.960	1var	B	198
0.584	1vcc	A	77
0.849	1vew	A	205
0.839	1vew	B	205
0.849	1vew	C	205
0.839	1vew	D	205
0.633	1vjw	A	60
0.838	1vki	A	561
0.852	1vki	B	561
0.516	1wfc	A	366
0.754	1wqm	A	130
0.581	1xel	A	129
0.906	1xra	A	383
0.480	1ypr	A	125
0.472	1ypr	B	125
0.755	1yvs	A	110
0.791	1zdi	A	129
0.837	1zdi	B	129
0.798	1zdi	C	129
1.000	1zdi	R	19
1.000	1zdi	S	19
0.856	2dap	A	320
0.927	2dnj	A	260
1.000	2dnj	B	8
1.000	2dnj	C	6
0.782	2glt	A	316
0.815	2hla	A	270
0.737	2hla	B	99
0.861	2hqi	A	72
0.722	2iad	A	194
0.580	2iad	B	205
0.811	2jdw	A	423
0.543	2lym	A	129
0.640	2lzm	A	164
0.738	2mef	A	130
0.746	2meg	A	130
0.738	2mei	A	130

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.725	1g3i	S	444
0.682	1g3i	T	444
0.705	1g3i	U	444
0.725	1g3i	V	444
0.736	1g3i	W	444
0.694	1g3i	X	444
0.462	1g3o	A	106
0.812	1g51	A	580
0.821	1g51	B	580
0.434	1g5b	A	221
0.448	1g5b	B	221
0.443	1g5b	C	221
0.822	1g5s	A	298
0.542	1g5u	A	131
0.542	1g5u	B	131
0.625	1g6t	A	427
0.876	1ga7	A	209
0.895	1ga7	B	209
0.421	1gb4	A	57
0.769	1gbw	A	130
0.792	1gby	A	130
0.879	1gc7	A	297
0.829	1gd1	O	334
0.829	1gd1	P	334
0.829	1gd1	Q	334
0.829	1gd1	R	334
0.723	1ge1	A	130
0.842	1ges	A	450
0.842	1ges	B	450
0.842	1geu	A	450
0.842	1geu	B	450
0.715	1gf0	A	130
0.738	1gf8	A	130
0.723	1gfe	A	130
0.746	1gfr	A	130
0.769	1gft	A	130
0.858	1ggy	A	731
0.854	1ggy	B	731
0.799	1gij	A	298
0.907	1gkx	A	388
0.896	1gmr	A	96

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.892	2mll	A	241
0.686	2mll	B	255
0.661	2mnr	A	357
0.896	2nck	L	144
0.896	2nck	R	144
0.885	2one	A	436
0.897	2one	B	436
0.778	2pad	A	212
1.000	2pad	I	1
0.652	2pli	A	112
0.867	2ple	A	105
1.000	2ple	B	12
0.856	2qil	A	194
0.871	2qil	B	194
0.871	2qil	C	194
0.684	2scu	A	288
0.611	2scu	B	388
0.688	2scu	D	288
0.621	2scu	E	388
0.940	2sdf	A	67
0.895	2sni	E	275
0.530	2sni	I	83
0.970	2tbd	A	134
0.892	2til	A	316
0.860	2ubp	A	100
0.582	2ubp	B	122
0.633	2ubp	C	570
0.873	2yfp	A	236
0.833	3elp	A	84
0.845	3elp	B	84
0.803	3fit	A	147
1.000	3fit	B	1
0.554	3gb1	A	56
0.495	3hck	A	107
0.712	3hfm	H	215
0.874	3hfm	L	214
0.845	3hfm	Y	129
0.806	3il8	A	72
0.674	3lzt	A	129
0.776	3phy	A	125
0.573	3rsd	A	124

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.896	1gmr	B	96
0.741	1gmu	A	143
0.748	1gmu	B	143
0.706	1gmu	C	143
0.741	1gmu	D	143
0.937	1gn2	A	207
0.937	1gn2	B	207
0.937	1gn2	C	207
0.937	1gn2	D	207
0.937	1gn2	E	207
0.937	1gn2	F	207
0.937	1gn2	G	207
0.937	1gn2	H	207
0.849	1gra	A	478
0.791	1gsh	A	316
0.830	1gtt	A	429
0.821	1gtt	B	429
0.837	1gtt	C	429
0.828	1gtt	D	429
0.794	1guz	A	310
0.829	1guz	B	310
0.787	1guz	C	310
0.803	1guz	D	310
0.756	1gx1	A	160
0.794	1gx1	B	160
0.788	1gx1	C	160
0.898	1gy5	A	127
0.882	1gy5	B	127
0.353	1h2h	A	241
0.841	1h4t	A	477
0.851	1h4t	B	477
0.860	1h4t	C	477
0.845	1h4t	D	477
0.905	1h6a	A	433
0.903	1h6a	B	433
0.831	1h6r	A	236
0.814	1h6r	B	236
0.835	1h6r	C	236
0.862	1h81	A	472
0.845	1h81	B	472
0.867	1h81	C	472

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.954	3rub	L	477
1.000	3rub	S	123
0.898	3sic	E	275
0.682	3sic	I	107
0.964	3tec	E	279
0.871	3tec	I	70
0.860	3ubp	A	100
0.865	3ubp	B	126
0.633	3ubp	C	570
0.853	4dbv	O	334
0.853	4dbv	P	334
0.853	4dbv	Q	334
0.856	4dbv	R	334
0.923	4hoh	A	104
0.923	4hoh	B	104
0.923	4hoh	C	104
0.923	4hoh	D	104
0.801	4jdw	A	423
0.000	4jdw	B	1
0.698	4lyo	A	129
0.682	4lzt	A	129
0.750	4rsk	A	124
0.867	5crx	A	345
0.875	5crx	B	345
1.000	5crx	C	35
1.000	5crx	D	35
1.000	5fd1	A	106
0.416	5gcn	A	166
0.654	5gsp	A	104
0.640	5lzm	A	164
0.791	6msf	A	129
0.837	6msf	B	129
0.822	6msf	C	129
1.000	6msf	R	14
1.000	6msf	S	10
0.778	6pad	A	212
0.000	6pad	I	3
0.889	6tli	A	316
0.864	6tmn	E	316
1.000	6tmn	I	4
0.791	7msf	A	129

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.871	1h83	A	472
0.864	1h83	B	472
0.869	1h83	C	472
0.940	1h9b	A	485
0.829	1hby	A	123
0.607	1hdm	A	201
0.611	1hdm	B	193
0.581	1hel	A	129
0.674	1hew	A	129
0.775	1hfg	A	71
0.401	1hh2	P	344
0.745	1hhj	A	275
0.790	1hhj	B	100
1.000	1hhj	C	9
0.745	1hhj	D	275

Prediction Accuracy	Protein Name	Chain	No of Residues Predicted
0.837	7msf	B	129
0.798	7msf	C	129
1.000	7msf	R	14
1.000	7msf	S	14
0.645	7rat	A	124
0.889	7tll	A	316
0.839	8ldh	A	329
0.698	8lyz	A	129
0.600	8pch	A	220
1.000	8pch	P	8
0.573	8rsa	A	124
0.573	8rsa	B	124
0.000	8tln	D	2
0.870	8tln	E	316

REFERENCES

Altschul, S., Madden, T., Shaffer, A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. (1997) Gapped Blast and PSI-Blast: A new generation of protein database search programs, *Nucleic Acids Res.* **25**, 3389-3402.

Andersen, C. A., Bohr, H. And Brunak, S. (2001), Protein secondary structure: category assignment and predictability, *FEBS Letters* **507**, 6-10.

Baello, B. I., Pancoska, P., and Keiderling, T. A. (2000), Enhanced prediction accuracy of protein secondary structure using hydrogen exchange fourier transform infrared spectroscopy, *Analytical Biochemistry* **280**, 46-57.

Bahar, I., Atilgan, A.R., Jernigan, R.L., and Erman, B. (1997), Understanding the Recognition of Protein Structural Classes by Amino Acid Composition, *Proteins: Structure, Function, and Genetics* **29**, 172-185.

Baldi, P., Brunak, S., Frasconi, P., Soda, G., and Pollastri, G. (1999) Exploiting the past and the future in protein secondary structure prediction, *Bioinformatics* **15**, 937-946.

Bohr, H., Bohr, J., Brunak, S., Cotterill, R. M J., Fredholm, H., Lautrup, B., and Petersen, S. B. (1990), A novel approach to prediction of the 3-dimensional structures of protein backbones by neural networks, *FEBS Letters* **261**, 43-46.

Brandon, C., and Tooze, J. (1998), Introduction to Protein Structure, Garland Publishing, Inc., New York.

Bystroff, C., Thorsson, V., and Baker, D. (2000), HMMSTR: A hidden Markov model for local sequence – structure correlations in proteins, *J. Mol. Biol.* **301**, 173-190.

Cai, Y. D., Liu, X. J., Xu, X. B., and Chou, K. C. (2002), Artificial neural network method for predicting protein secondary structure content, *Computers and Chemistry* **26**, 347-350.

Cai, Y.D., Liu, X.J., Xu, X.B., and Zhou, G.P. (2001), Support Vector Machines for predicting protein structural class, *BMC Bioinformatics* **2**, 3.

Chou, P. Y., and Fasman, G. (1978), Empirical predictions of protein conformation, *Ann. Rev. Biochem.* **47**, 251- 76.

Chou, K.C. (1995), Does the folding type of a protein depend on its amino acid composition?, *FEBS Letters* **363**, 127-131.

Chou, K. C., and Elrod, D. W. (1999), Prediction of membrane protein types and subcellular locations, *Proteins* **34**, 137-153.

Cid, H., Vargas, V., Bunster, M., and Bustos, S. (1982), Secondary structure prediction of human salivary proline-rich proteins, *FEBS Letters* **198**, 140-144.

Cid, H., Bunster, M., Arriagada, E., and Campos, M. (1982), Prediction of secondary structure of proteins by means of hydrophobicity profiles, *FEBS Letters* **150**, 247-254.

Creighton, T. E. (1993), *Proteins: Structures and Molecular Properties*, W H Freeman & Co., New York.

Cuff, J. A., Barton, G. (1999), Evaluation and improvement of multiple sequence methods for protein secondary structure prediction, *Proteins* **34**, 508-519.

Dayhoff, M. O., and Eck, R. V., eds. (1968), *Atlas of protein sequence and structure*, *Natl. Biomed. Res. Found.* **3**, 33.

Dayhoff, M., (1978), Atlas of protein sequence and structure, *Natl. Biomed. Res. Found.* **5**, Suppl. 3, 345-358.

Frishman, D., and Argos, P. (1997), Seventy-five percent accuracy in secondary structure prediction, *PROTEINS: Structure, Function and Genetics* **27**, 329-335.

Garnier, J., Osguthorpe, D. J., and Robson, B. (1978), Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins, *J. Mol. Biol.* **120**, 97-120.

Henikoff, S., and Henikoff, J. G. (1992), Amino acid substitution matrices from protein blocks, *Biochemistry* **89**, 10915-10919.

Hua, S., and Sun, Z. (2001), A novel method of protein secondary structure prediction with high segment overlap measure: Support vector machine approach, *J. Mol. Biol.* **308**, 397-407.

Jarey, J., and Hanley, V., Proteins - Biophysical society on-line text book

Jureti, D., Lu, B., and Trinajsti, N. (1995), Secondary structure prediction quality for naturally occurring amino acids in soluble proteins, *Journal of Molecular Structure: THEOCHEM* **338**, 43-50.

Kabsch, W., and Sander, C. (1983), A dictionary of protein secondary structure, *Biopolymers* **22**, 2577-2637.

Kilosanidze, G. T., Kutsenko, A. S., Esipova, N. G., and Tumanyan, V. G. (2002), Use of molecular mechanics for secondary structure prediction. Is it possible to reveal α -helix?, *FEBS Letters* **510**, 13-16.

King, R. D., and Sternberg, M. J. E. (1996), Identification and application of the concepts important for accurate and reliable protein secondary structure prediction, *Protein Science* **5**, 2298-2310.

Kloczkowski, A., Ting, K.-L., Jernigan, R.L., and Garnier, J. (2002), Protein secondary structure prediction based on the GOR algorithm incorporating multiple sequence alignment information, *Polymer* **43**, 441-449.

Levitt, M., and Chothia, C. (1976), Structural patterns in globular proteins, *Nature* **261**, 552-558.

Liu, W. M., and Chou, K. C. (1999), *Protein Engineering* **12**, 1041-1050.

Mount, D. W. (2001), *Bioinformatics: Sequence & Genome Analysis*, Cold Spring Harbor Laboratory Press, Woodbury, New York.

Mitschler, A. (2002), X-Ray Data collection at Synchrotrons: Best strategies and new challenges, Protein Structure Prediction Workshop, Antalya, Turkey.

Nakashima, H., Nishikawa, K., and Ooi, T. (1986), *J. Biochem.* **99**, 152-162.

Pasquier, C., Promponas, V. J., Palaios, G. A., Hamodrakas, J. S., and Hamodrakas, S. J. (1999), A novel method for predicting transmembrane segments in proteins based on a statistical analysis of the SwissProt database: The PRED-TMR algorithm, *Protein Eng.* **12**, 381-385.

Petersen, T. N., Lundegaard, C., Nielsen, M., Bohr, H., Bohr, J., Brunak, S., Gippert, G. P., and Lund, O. (2000), Prediction of protein secondary structure at 80% accuracy, *Proteins* **41**, 17-20.

Pilpel, Y., Ben-Tal N., and Lancet, D. (1999), kPROT: A knowledge-based scale for the propensity of residue orientation in transmembrane segments. Application to membrane protein structure prediction, *J. Mol. Biol.* **294**, 921–935.

Powell, K., Abbott, A., and Check, E. (2002), Mouse genome: The real deal, *Nature* **420**, 456.

Qui, J., Liang, R., Zou, X., and Mo, J. (2003), Prediction of protein secondary structure based on continuous wavelet transform, *Talanta* **61**, 285-293.

Robbins, R. C. (2000), From alchemy to IPO: The Business of Biotechnology, Persus Publishing.

Rost, B., and Sander, C. (1993), Prediction of protein secondary structure at better than 70% accuracy, *J. Mol. Biol.* **232**, 584-599.

Rost, B. (2001) Review: Protein secondary structure prediction continues to rise, *Journal of Structural Biology* **134**, 204-218.

Salamov, A. A., and Soloveyev, V. V. (1995), Prediction of Protein Secondary Structure by Combining Nearest-neighbor Algorithms and Multiple Sequence Alignments, *J. Mol. Biol.* **247**, 11-15.

Salamov, A. A., and Soloveyev, V. V. (1997) Protein secondary structure prediction using local alignments, *J. Mol. Biol.* **268**, 31-36.

Sen, S. (2003), Statistical analysis of pair-wise compatibility of spatially nearest neighbor and adjacent residues in α -helix and β -strands: Application to a minimal model for secondary structure prediction, *Biophysical Chemistry* **103**, 35-49.

Severcan, M., Severcan, F., and Haris, P. I. (2001), Estimation of protein secondary structure from FTIR spectra using neural networks, *Journal of Molecular Structure* **565-566**, 383-387.

Thiele, R., Zimmer, R., and Lengauer, T. (1999), Protein threading by recursive dynamic programming, *Journal of Molecular Biology* **290**, 757-779.

Westhead, D. R., and Thornton, J. M. (1998), Protein structure prediction, *Current Opinion in Biotechnology* **9**, 383-389.

Wolfsberg, T. G., Mcentyre, J, and Schuler, G. D. (2001), Guide to the draft human genome, *Nature* **409**, 824-826.

Yi, T. M., and Lander, E. S. (1993), Protein Secondary Structure Prediction Using Nearest-neighbor Methods, *Journal of Molecular Biology* **232**, 1117-1129.

Zhang, C. T., and Zhang, R. (2003), Q₉, a content-balancing accuracy index to evaluate algorithms of protein secondary structure prediction, *The International Journal of Biochemistry & Cell Biology* **35**, 1256-1262.

Zvelebil, M. J. J. M., Barton, G. J., Taylor, W. R., and Sternberg, M. J. E. (1987), Prediction of protein secondary structure and active sites using the alignment of homologous sequences, *Journal of Molecular Biology* **195**, 957-961.