

**Peptide Based Drug Design and Pathway Analysis  
for Caspase-1**

by

**Cemre Kocahakimođlu**

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

**Master of Science in  
Chemical and Biological Engineering**

**Koç University**

**July 2011**



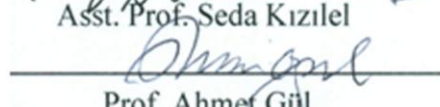
Koç University  
Graduate School of Sciences and Engineering

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

Cemre Kocahakimoğlu

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:

  
Prof. Burak Erman (Advisor)  
  
Asst. Prof. Seda Kızılel  
  
Prof. Ahmet Gül

Date: 11.07.2011

*To Ceysu (my sister), Tümay (my mother), Celal (my father) Kocahakimođlu*

## ABSTRACT

As one of the mostly studied protein in the literature, caspase-1 (ICE) attracts the attention of many scientists due to its crucial roles in inflammatory responses. It has other roles in the apoptotic path, for example, because of having more than 40 substrates. Increased expression of its substrates such as pro-IL-1 $\beta$  results in inflammatory disorders. Consequently, inhibition and pathway studies related to caspase-1 have gained importance.

Peptide based drug design for caspase-1 is performed and potent inhibitors are determined computationally. Bicyclic (a molecule that contains two fused rings) and ketone structures with Asp and D-enantiomeric aminoacids are obtained as good inhibitors in accordance with previous experimental work. Moreover, multi-target drug determination in the caspase-1 pathway is made. Conformational factors in tripeptides are also taken into consideration with Viterbi Algorithm, which indicates whether a peptide can change its conformation from minimized state to bound state.

Knockout analysis on the ICE pathway by Gaussian Network Model (GNM) shows knockouts of NLRP3, ASC, caspase-1, NF- $\kappa$ B, pro-IL-33, TLR4, TLR2, TRIF and MyD88 are effective.

## ÖZET

Literatürde çok çalışılan proteinlerden biri olan kaspaz-1 (ICE), inflamatuvar tepkilerdeki önemli rolleri yüzünden birçok bilim adamının ilgisini çekmektedir. Ayrıca, 40'dan fazla substratı olduğu için apoptotik yoldaki fonksiyonları gibi başka rolleri vardır. Pro-IL-1 $\beta$  gibi substratlarının fazla salınımı bazı inflamatuvar rahatsızlıklara yol açar. Bu yüzden, kaspaz-1 ile ilgili olan inhibisyon ve metabolik yol çalışmaları önem kazanmıştır.

Kaspaz-1 için peptit bazlı ilaç tasarımı çalışıldı ve potent inhibitörler hesaplamalarla belirlendi. Asp ve D-enantiomerik aminoasitlerle bisiklik (kaynaşmış olan iki tane halkalı yapıyı barındıran molekül) ve keton yapıların, önceki deneysel çalışmalarla uyumlu bir şekilde iyi inhibitörler olduğu görüldü. Bununla beraber, kaspaz-1 metabolik yolunda çok hedef belirlenmesi yapılmıştır. Bir peptidin konformasyonunu minimize durumundan bağlanmış durumuna geçerken değiştirip değiştiremeyeceğini gösteren Viterbi Algoritması ile de tripeptitlerdeki konformasyonel faktörler dikkate alınmıştır.

Kaspaz-1 metabolik yolunda Ağ Yapı Modeli ile yapılan nakavt analizi; NLRP3, ASC, kaspaz-1, NF- $\kappa$ B, pro-IL-33, TLR4, TLR2, TRIF ve MyD88 nakavtlarının etkili olduğunu göstermiştir.

## ACKNOWLEDGEMENTS

Firstly, I would like to express my deepest gratitude to my advisor Prof. Dr. Burak Erman for his supervision and continuous support he provided me during two years in Koç University. It is an honor and great pleasure for me to meet and work with him.

I specially express thanks to my thesis committee Asst. Prof. Dr. Seda Kızılel and Prof. Dr. Ahmet Gül for sparing their precious time and their support during my thesis study. I would also like to thank Assoc. Prof. Dr. Halil Kavaklı for his helps in my studies. I would like to extend my sincere thanks to everyone who had contributed to my education throughout my life.

I would like to thank dear departed Mr. Vehbi Koç and Koç family for giving the opportunity of master education. I also thank to Koç University and The Turkish Academy of Sciences for their funds throughout my master education.

I am grateful to Evrim Besray Ünal, Bilal Çakır, Burcu Mesut, Beytullah Özgür, Çiğdem Sevim Bayrak and İbrahim Hocaoğlu for their endless help in my thesis work. My special thanks go to my friends Ece Bulut and Ümmühan Canan for their endless and valuable friendship. Moreover, I am very grateful to my officemates, Cemal Erdem, Mehmet Ali Öztürk, Mümin Öztürk, Bilgen Bıçak, Mert Gür and Güzde Eskici. Emine Güven, Ömer An, Selimcan Azizoğlu, Meriç Ataman, Halil Peynirci, Zeynep Ülker, Ramazan Oğuz Canıaz, Bengisu Seferoğlu, Caner Nazlı, Cahit Dalgıçdır and other friends in Koç University have also made available their precious friendship. Dilek Turan has made available her support during my education as a good and trustworthy friend.

Last but definitely not the least, I wish to thank my mother, my father and my sister Ceysu for the endless support they have given. Without them, I could not be what I am today. I dedicate this work to my family.

## TABLE OF CONTENTS

ABSTRACT.....	iii
ÖZET .....	iv
ACKNOWLEDGEMENTS.....	v
TABLE OF CONTENTS.....	vi
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
1. INTRODUCTION .....	1
2. LITERATURE REVIEW.....	5
2.1 Caspase-1 (Interleukin converting enzyme-ICE).....	5
2.1.1. The structural and biological properties .....	5
2.1.2. Caspase-1 Pathway.....	12
2.1.3. Diseases related to caspase-1 .....	28
2.2 Drugs .....	29
2.2.1. Desired features of drugs .....	29
2.2.2. Peptide drug properties .....	31
2.2.3. Drugs for caspase-1 in literature .....	34
2.3 Docking.....	41
2.3.1. AutoDock.....	45
2.3.2. GOLD .....	48
2.4 Genetic Algorithm.....	49
2.4.1. The theory .....	49
2.4.2. Usage for ligand search .....	51
2.5 Pharmacophore analysis .....	52
2.6 GNM (Gaussian Network Model).....	54
3. METHODS.....	58
3.1 Caspase-1 .....	58

3.2	Computational methods for ICE drug design .....	59
3.2.1.	Docking .....	59
3.2.2.	Computational ligand analysis tools .....	61
3.2.3.	Genetic Algorithm.....	63
3.2.4.	Complete Enumeration.....	65
3.3	Conformational Factors in the Design of Peptide Drugs.....	66
3.3.1.	Determination of States .....	66
3.3.2.	Equilibrium and Kinetic Factors .....	70
3.3.3.	Determination of Path Followed .....	72
3.4	The Interaction Matrix for the Caspase-1 Pathway and Correlations .....	76
4.	RESULTS & DISCUSSIONS.....	79
4.1	Peptide-based drug design .....	79
4.1.1.	Determination of docking place.....	79
4.1.2.	Previously designed inhibitors for caspase-1.....	85
4.1.3.	Drug determination for caspase-1 inhibition .....	86
4.1.4.	Multi-target drug determination considering the ICE pathway .....	104
4.2	Conformational Factors in the Design of Peptide Drugs: Application to the Design of Caspase-1 Inhibitors.....	107
4.2.1.	Equilibrium and Kinetic Factor Analysis of Tripeptides .....	107
4.2.2.	Determination of Path Followed by Tripeptides .....	107
4.3	Knockout Analysis on the ICE pathway.....	111
5.	CONCLUSION .....	117
6.	APPENDIX-I.....	119
7.	APPENDIX-II .....	124
	BIBLIOGRAPHY .....	131
	VITA.....	147



## LIST OF TABLES

Table 2.1 Pathogens and their target NLRs. ....	25
Table 2.2 The peptide based drugs in market and clinical trials .....	32
Table 3.1 Definitions of 21 states.....	69
Table 3.2 The list of used proteins in the ICE pathway. ....	78
Table 4.1 Probabilities residues having different states in tripeptides at docking place candidates.....	81
Table 4.2 Docking results of previously designed inhibitors. ....	86
Table 4.3 Aminoacids GOLD score results. ....	87
Table 4.4 D-aminoacid+ 1 x iLib Diverse results. ....	88
Table 4.5 D-aminoacid+ 2 x iLib Diverse results. ....	91
Table 4.6 Docking results of Asp/Malonate + 1 x iLib Diverse. ....	96
Table 4.7 Docking results for 10 structures from 2 x iLib Diverse.....	99
Table 4.8 Docking scores of converged tripeptides. ....	100
Table 4.9 Drugs for procaspase-1.....	106
Table 4.10 Energy change and probabilities of residues having different states. ....	107
Table 4.11 Results of Knockout Analysis. ....	113
Table 7.1 Energy change and probabilities of residues having different states. ....	124

## LIST OF FIGURES

Figure 2.1 Relationship between human caspases. ....	7
Figure 2.2 Structural comparison of caspase-1, -3, -7, -8 and -9. ....	7
Figure 2.3 Structure of caspase-1. ....	8
Figure 2.4 The active site and substrate interaction in ICE. ....	10
Figure 2.5 Processed caspase. ....	10
Figure 2.6 Caspase signaling pathways. ....	11
Figure 2.7 A proposed schematic of ICE interactions in inflammasomes. ....	13
Figure 2.8 Caspase cleavage site nomenclature. ....	34
Figure 2.9 Caspase-3 and its tetrapeptide inhibitor (Ac-DVAD-FMK). ....	35
Figure 2.10 The interaction between 27c and ICE. ....	36
Figure 2.11 The structures in study of Wagner et al. (2006). ....	37
Figure 2.12 Crystal structure of compound 11. ....	39
Figure 2.13 Structure 12. ....	40
Figure 2.14 The binding of inhibitor Dmp323 to HIV protease. ....	42
Figure 2.15 a) Crossover and b) Mutation. ....	50
Figure 2.16 Definition of fluctuations. ....	56
Figure 3.1 Caspase-1. ....	58
Figure 3.2 A sample demonstration of a ligand in LigandScout. ....	62
Figure 3.3 Genetic algorithm flowchart. ....	65
Figure 3.4 The definitions of torsion angles. ....	67
Figure 3.5 The 21 states in Ramachandran map. ....	70
Figure 3.6 A schematic representation of energies. ....	71
Figure 3.7 Schematic representation of backtracking in Viterbi algorithm. ....	75
Figure 4.1 Interactions taken from GNM. ....	80
Figure 4.2 Behavior of Phe residue in FYA in a) bound and b) minimized conformations. ....	83
Figure 4.3 VitAL results for AWG. ....	84
Figure 4.4 VitAL results for YWG. ....	85
Figure 4.5 Docked structures of <i>d-tyr+bicyclic(22)</i> (left) and <i>d-phe+bicyclic(4)</i> (right). ....	90
Figure 4.6 Pharmacophores of <i>d-tyr+bicyclic(22)</i> (left) and <i>d-phe+bicyclic(4)</i> (right). ....	90
Figure 4.7 Docked structures of <i>d-Tyr-ester(5)</i> (top left), <i>d-Phe-monocyclic(27)</i> (top right), <i>d-Phe-monocyclic(20)</i> (middle left), <i>d-Tyr+pharma(32)</i> (middle right), <i>d-Phe-L-Pro</i> (bottom right) and <i>d-Phe-L-Gly</i> (bottom left). ....	94
Figure 4.8 Pharmacophores of <i>d-Tyr-ester(5)</i> (top left), <i>d-Phe-monocyclic(27)</i> (top middle), <i>d-Phe-monocyclic(20)</i> (top right), <i>d-Tyr+pharma(32)</i> (bottom right), <i>d-Phe-L-Pro</i> (bottom middle) and <i>d-Phe-L-Gly</i> (bottom left). ....	95
Figure 4.9 Docked structures of <i>Bicyclic-asp(13)</i> (left) and <i>Benzene-asp(12)</i> (right). ....	97

Figure 4.10 Pharmacophores of <i>Bicyclic-asp(13)</i> (left) and <i>Benzene-asp(12)</i> (right). .....	98
Figure 4.11 Pharmacophore (left) and docked conformation (right) of <i>L-Gly-a1(11)</i> . .....	99
Figure 4.12 MW of literature (red) and found drugs (green) with Lipinski limits. ....	102
Figure 4.13 cLogP of literature (red) and found drugs (green) with Lipinski limits. ....	102
Figure 4.14 TPSA of literature (red) and found drugs (green). .....	103
Figure 4.15 Hydrogen acceptor of literature (red) and found drugs (green) with Lipinski limits. ....	103
Figure 4.16 Hydrogen donor of literature (red) and found drugs (green) with Lipinski limits. ....	104
Figure 4.17 VitAL results for DWG. ....	108
Figure 4.18 VitAL results for AWG. ....	109
Figure 4.19 VitAL results for GWG. ....	110
Figure 4.20 Knockout analysis. ....	112
Figure 6.1 The molecular structures of D-aminoacid+ 1 x iLib Diverse. ....	119
Figure 6.2 The molecular structures of D-aminoacid+ 2 x iLib Diverse. ....	120
Figure 6.3 The molecular structures of D-aminoacid+ 2 x iLib Diverse. ....	121
Figure 6.4 The molecular structures of 9 molecules from Asp/Malonate + 1 x iLib Diverse. .....	122
Figure 6.5 The molecular structures of 10 molecules from Asp/Malonate + 2 x iLib Diverse. ....	123
Figure 7.1 VitAL results for AWE. ....	125
Figure 7.2 VitAL results for DFF. ....	126
Figure 7.3 VitAL results for DYF. ....	128
Figure 7.4 VitAL results for SFF. ....	130

## Chapter 1

### 1. INTRODUCTION

Caspase-1 (Interleukin converting enzyme, ICE) is an endopeptidase that mainly functions in processing pro-inflammatory cytokines of IL-1, IL-18 and IL-33. Increased expression of pro-inflammatory cytokines can result inflammatory diseases including hereditary fever syndromes [1]. Consequently, inhibition of caspase-1 has been studied extensively for treatment of inflammatory diseases and several candidate drugs have been found. There are problems related to ICE inhibitors such as penetration. First, candidate peptide based drugs are found in this study with new approaches. After that, analyses related to their drug-likeness and their potential for multi-target inhibition are performed.

Peptides form an important subgroup of drug molecules for protein inhibition [2]. One distinguishing feature of peptide drugs is that unlike small, relatively stiff organic molecules, peptides exhibit large conformational changes at physiological temperatures. This confers the peptides with an unusual capability of adopting the shape of the binding region on the protein. However, all conformations of peptides are not equally probable, and the probabilities may change significantly depending on conformation. These differences in probability result from the hopping of the two dihedral angles,  $\varphi$  and  $\psi$  over the two peptide backbone bonds flanking the alpha carbon. A survey of dihedral angles that are accessible to a peptide shows that there are more than ten distinct regions in the  $\varphi-\psi$  plane, or the so-called Ramachandran map. Each of these regions corresponds to a state with a certain probability. There are  $20^3$  tripeptides as drug candidates. In each tripeptide,

the eight rotatable peptide bonds lead to  $10^8$  different conformations or states. In solution, the free peptide spends most of its time in the high probability states. Upon binding, the peptide is forced to take a conformation that is compatible with the shape of the protein surface at the binding region. The accessibility and the possibility of the bound state depend on two thermokinetic conditions: First, the free energy difference between the bound state and the highly probable states of the free peptide should not be excessive. Secondly, the conformation obtained at the bound state should be accessible from the manifold of the high probability states of the free peptide. The second condition dictates that the pathways from the high probability states to the bound state should not pass over high energy barriers. The present study also deals mainly with the fulfillment of these two conditions as part of a peptide design process. To our knowledge, this important criterion in the peptide drug design has not been considered explicitly until now. The problem is necessarily an NP-complete problem of high complexity. One way to overcome the complexity is to design peptides in the absence of these two conditions using known design software such as Autodock, Gold, Openeye etc., and then eliminate the peptides that violate these criteria. This posterior-filtering strategy is adopted in the present study. The software mentioned are programs that order the different peptides according to their binding energies, and in none of them the two conditions mentioned above are taken into account.

Upstream and downstream interactions of caspase-1 have also been considered as important research topics. Since recognition of pathogens takes place mainly by two components which are important components in the ICE pathway. These two components are : (i) on the cell surface and in the endosomal compartments by toll-like receptors (TLRs) (ii) in the cytosol by nucleotide-binding oligomerization domain (NOD)-like receptors (NLRs) [3]. In this study, caspase-1 pathway studies in the literature will be provided with considering TLRs and NLRs. Interactions between proteins and triggers will also be given. All interactions in the ICE pathway taken from literature data are organized

in a connectivity Kirchhoff matrix,  $\mathbf{K}$ , and the properties of the resulting graph are analyzed using the Gaussian Network Model (GNM). A knockout analysis on the caspase-1 pathway is performed using GNM. State fluctuations between  $i^{\text{th}}$  and  $j^{\text{th}}$  proteins ( $\langle \delta x_i \delta x_j \rangle$ ) are considered in the knockout analysis. In previous studies related to GNM, residue based interactions in a protein were considered. However, protein level interactions through state fluctuations in a pathway have not been analyzed before by GNM. Knockout analysis for ICE pathway is done by GNM in order to identify defects in the pathway. In previous studies, knockout animals such as mice [4-5], genes [6] or cells [7] were used in an experimental manner. The present study extends the use of GNM to the search of knockout and defect relations in the ICE pathway.

In Chapter 2, a detailed literature review is given for caspase-1 (its structural and biological properties, its pathway and diseases related to caspase-1), drug (desired features, peptide drug properties, ICE drugs in literature), docking (its methods, AutoDock, GOLD), genetic algorithm (its theory, literature applications for ligand search), pharmacophore analysis by LigandScout, literature applications of Viterbi Algorithm and Gaussian Network Model (GNM).

In Chapter 3, methods are given. Caspase-1 PDB properties and computational methods for ICE drug design (AutoDock, GOLD, computational ligand analysis tools such as LigandScout), genetic algorithm, complete enumeration, determination of states, equilibrium and kinetic factors in a tripeptide, determination of a path followed by a tripeptide with Viterbi Algorithm and GNM (Gaussian Network Model) are given in this section.

In Chapter 4, peptide-based drug design for ICE is given initially comparing with previously designed inhibitors for caspase-1. Furthermore, multi-target drug determination is performed taking into account the ICE pathway. Thus, a more detailed analysis of

tripeptide inhibitors is performed by conformational factors using Viterbi algorithm. In pathway analysis part for the ICE pathway, a knockout study is performed.

In Chapter 5, conclusion of this study is provided with underlining important facts related to this study.

## Chapter 2

### 2. LITERATURE REVIEW

#### 2.1 Caspase-1 (Interleukin converting enzyme-ICE)

Caspases belong to dimeric thiol proteases that have a place in cellular processes like apoptosis and inflammation. Caspase-1 is the first member of this family that was studied. It is responsible for many diseases and studies throughout the world are being conducted to find a cure for these disorders by inhibition of ICE [8].

##### 2.1.1. The structural and biological properties

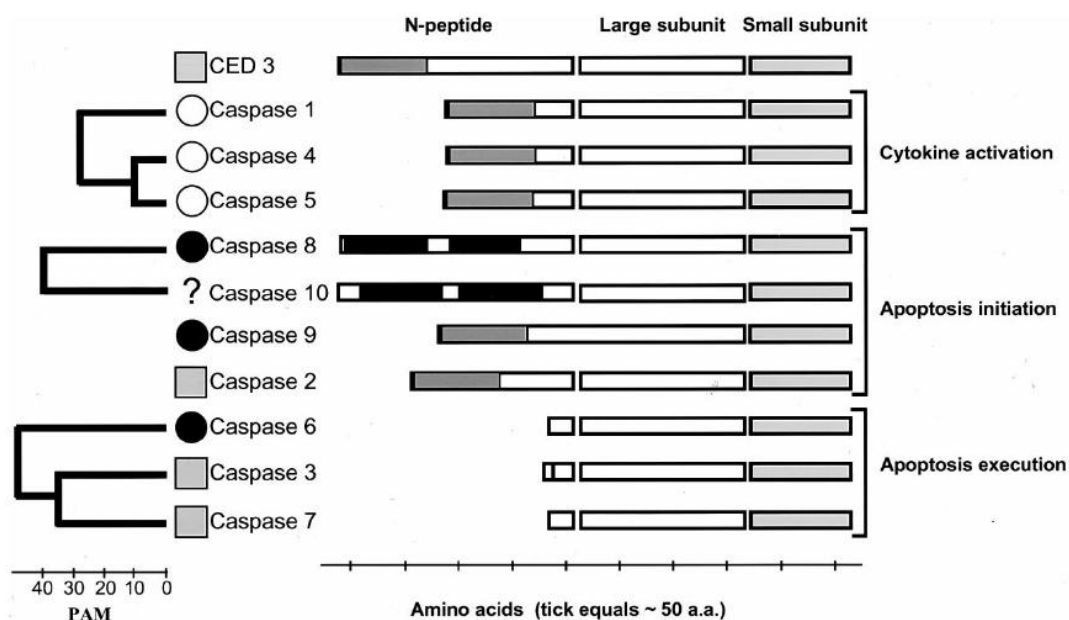
The enzymes, which cleave peptide bonds, have higher energy barriers to overcome that are peptidases or proteases. These enzymes contain endopeptidases and exopeptidases. Most proteases force the trigonal planar peptide bond into a tetrahedral geometry which is certainly required for hydrolysis. There are five mechanisms for this hydrolysis but formation of a tetrahedral intermediate by promotion of cysteine is a promising step for caspases [9]. Caspases are endopeptidases which have 11 types in humans. Caspase is an abbreviation of cysteine-dependent aspartate-specific protease. There are two types of caspases according to their functions that are pro-inflammatory (e.g. caspase-1) and apoptotic (e.g. caspase-3) cytokines [10]. Human caspases exist in certain pH values that are between 6.8 and 7.2 [11].

The caspase enzymes attract attention for drug design. The caspase family is composed of three subfamilies (group I, group II and group III) that are dependent on their



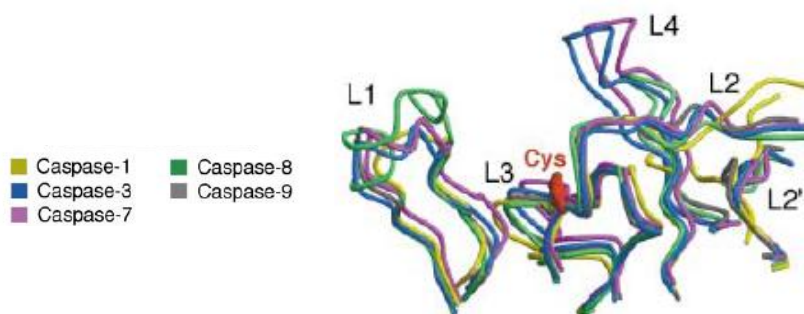
predominant functional roles and substrate specificities [12]. Caspases need an aspartic acid residue in the P1 position for their substrate and are grouped into three categories according to their substrate preferences. Caspase-1, -4 and -5 belong to group I that need a hydrophobic residue at P4 place. Caspase-2, -3 and -5 (group II) need an Aspartic acid residue in P4 position while caspase-6, -8, -9 and -10 (group III) do not prefer a specific amino acid in P4 position [13].

There are three main biological activities of caspases. A classification of these activities can be seen in Figure 2.1. Prediction analysis of microarray (PAM) classification of different caspases was also given with a cut-off of 50 aminoacids. Nematode death gene CED 3 resembles caspase-1 and can also be seen in Figure 2.1. The caspase family can be classified phylogenetically. Effector caspases like caspase-3 have no prodomain. On the other hand, other caspases have either a death effector domain (DED) or a caspase recruitment domain (CARD). Caspase-1 is one of the CARD-containing caspases. Moreover, all caspases have two domains: amino-terminal prodomain and a catalytic domain, which is composed of 20-kDa and 10-kDa units. The P4 preferences are also demonstrated with ○ for (W/Y/L) EHD, ● for (I/L/V) EXD and ■ for DEXD. In N-peptide parts, CARD (Caspase activation and recruitment domain) and DED (Death effector domain) domains are shown in the Figure 2.1 by grey and black colors, respectively [14-15].



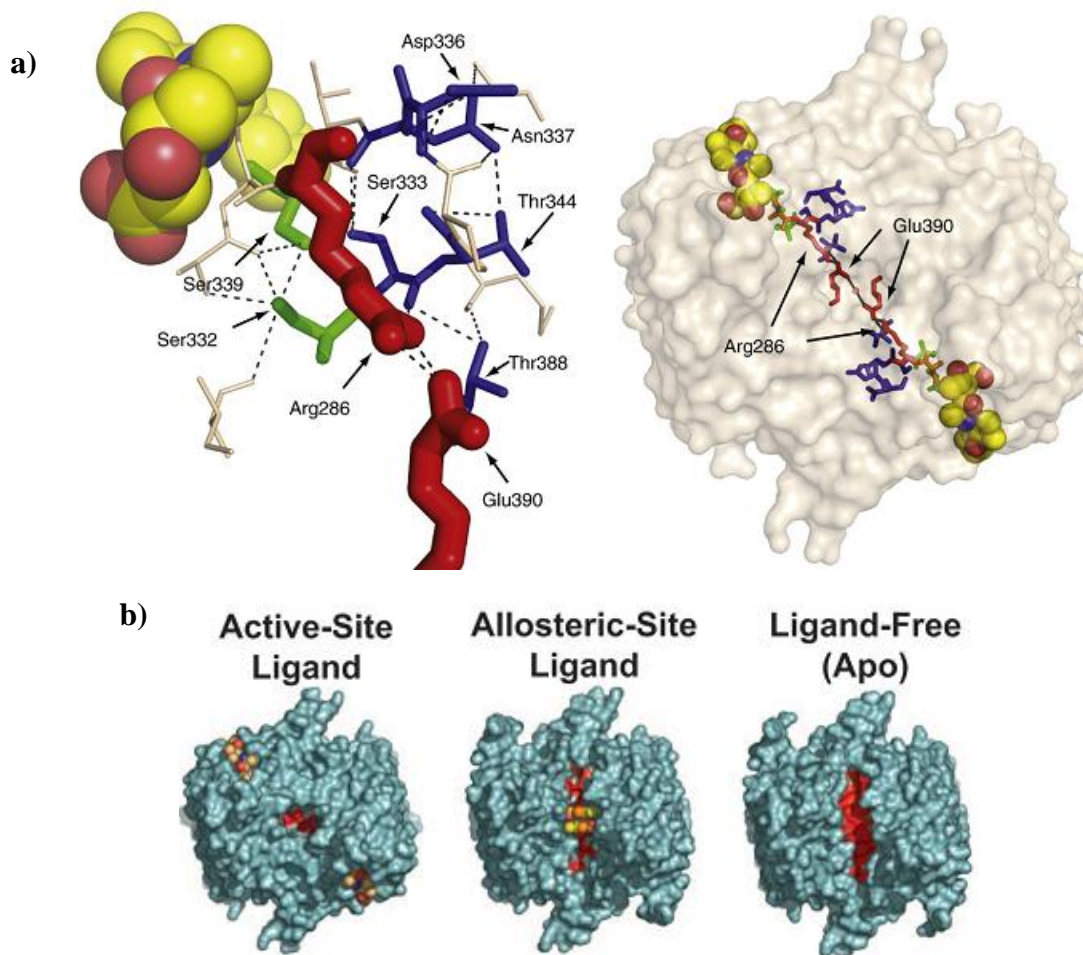
**Figure 2.1** Relationship between human caspases [14].

In all caspases, active site is conserved. In Figure 2.2, the catalytic residue Cys285 is shown by a red color and the conformations of loops according to Cys285 are also demonstrated. The L4 loop in caspases determines the P4 specificity of substrate. For instance, caspase-1 has short L4 loop, which results in bulkier and hydrophobic residues and caspase-3 has long L4 loop that is responsible for preference of Asp residue [16].



**Figure 2.2** Structural comparison of caspase-1, -3, -7, -8 and -9 [16].

In 1992, ICE was first purified and its DNA was cloned as well as sequenced. Caspase-1 is processed as a dormant 45 kDa proenzyme in cell cytoplasm. The active form contains two subunits: 20 kDa and 10 kDa parts. Activation of caspase-1 can happen autolytically or by other family members due to excision at the Asp site. [9].

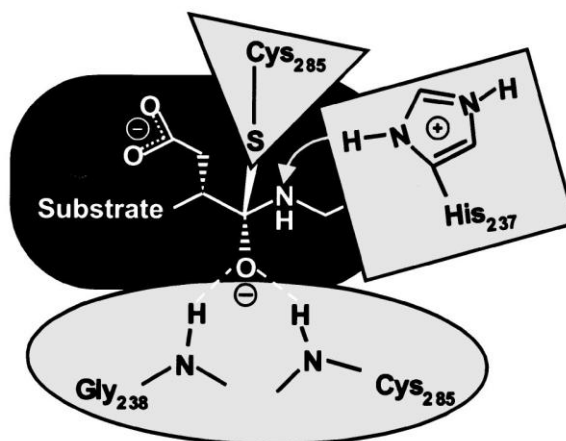


**Figure 2.3** Structure of caspase-1.

**a)** Wild type caspase-1 (Red, green and blue show a >100-fold, >2-fold and <2-fold reduction when replace with alanine). The yellow spheres represents the z-VAD-FMK inhibitor (PDB ID=2HBQ) [8]. **b)** Comparison of the surface of active-site ligand (z-VAD-FMK), allosteric site ligand (Compound 34) and ligand-free caspases (PDB ID=1SC1) [17].

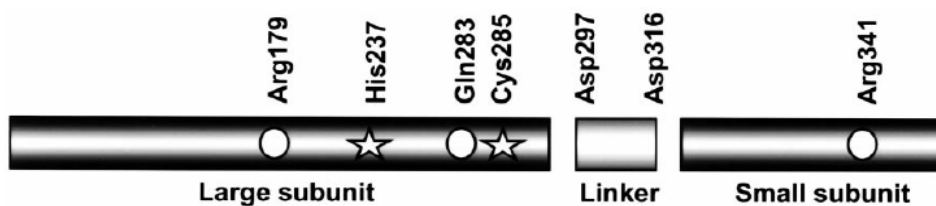
Datta et al. (2008) studied the effect of mutation along the active site of caspase-1 and determined that there is a line of residues which form an H-bonding network and a salt bridge by connecting the active and allosteric ligand sites of caspase-1. This line of residues was found to be crucial for wild type caspase-1 activity. In Figure 2.3, the impact of alanine replacement can be seen. Additionally, there is a salt bridge between Arg286-Glu390 with connecting two active sites with allosteric site. Datta et al. (2008) also determined Cys285 was the active site [8]. The 20-kDa chain contains active cysteine residue [18]. This finding was also verified by Scheer et al. (2006) by binding the inhibitor (z-VAD-FMK) on a sulphur atom of Cys285. However, they also determined that Cys285 would be blocked if Cys331 was used as docking place due to being at allosteric site as seen in Figure 2.3. As a result, both Cys285 and Cys331 can be used as binding sites [17].

Caspase-1 has a large five-stranded parallel  $\beta$ -sheets and a single anti-parallel strand near the C terminus. This antiparallel strand relies on the formation of the tetramer complex. In the catalytic dyad, Cys285 and His237 act as a nucleophile and general base that takes proton from Cys, respectively. Arg-179, Gln283 and Arg341 form the prominent binding site that is a hole in the active site. The backbone amides of Gly238 and Cys285 donate H-bonds to the carbonyl oxygen. Before nucleophilic attack of catalytic Cys285 on the carbonyl carbon, the thiol group of Cys285 donates its proton to His237. As a result, His237 becomes a catalytic acid and allows polarization of water molecules during deacylation of the enzyme as seen in Figure 2.4 [11, 14].



**Figure 2.4** The active site and substrate interaction in ICE [11].

All caspase types are produced as a single chain zymogen, which contains small and large subunits. Two large and two small subunits are the parts of the active hetero-tetramer. In the hetero-activation, that is the mechanism of a single chain caspase zymogen for producing an active protease, of caspase-1, Asp297 residue is the main site for activation because it is conserved. A schematic of processed caspase is given in Figure 2.5. Cleavage between small and large subunits forms the active caspase and sometimes results in a linker [13-14].

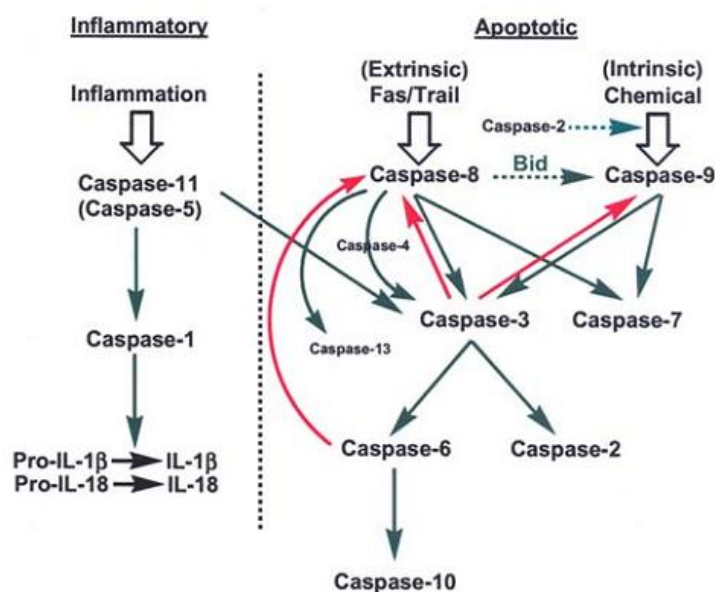


**Figure 2.5** Processed caspase.

Starred and circled residues demonstrate the residues in catalytic dyad and S1 subsite with caspase-1 numbering [13-14].

The inflammatory and apoptotic signaling pathway of the caspase family is also illustrated in Figure 2.6. The inflammatory caspases, caspase-1, -5 and -11, have a place in the activation of pro-IL-1 $\beta$  and IL-18. For caspase-1 activation, caspase-11 whose potential

homolog can be caspase-5 is needed. Caspase-11 activates caspase-1 after the activation of pro-caspase-11. When caspase-1 is activated, the cleavage, in other words, the activation of pro-IL-1 $\beta$  occurs. Moreover, caspase-4 is thought to mediate in the activation of caspase-1 since expression of both caspases occurs at the same time [19-21].



**Figure 2.6** Caspase signaling pathways [13].

Caspases were initially determined to have a part in apoptosis by genetic analysis in the nematode *C.elegans*. The change (with deletion or mutation) in one gene (CED-3) resulted in the cancelation of 131 programmed cell deaths. CED-3 was found to have a relation with mammalian caspase-1 protein, which was also validated with other studies. Caspase family members were identified to have a crucial role in apoptosis in mammals by cleaving key proteins at the onset of apoptosis. Additionally, peptide-based and non-peptide drugs inhibited the caspase family by preventing the apoptosis *in vitro*, in whole cells and in *in vivo* models. Caspases have a certain role in apoptotic processes but initial studies related to ICE demonstrated that caspase-1 did not have a crucial role in apoptosis in addition to inflammatory responses. However, recent studies show ICE involves both

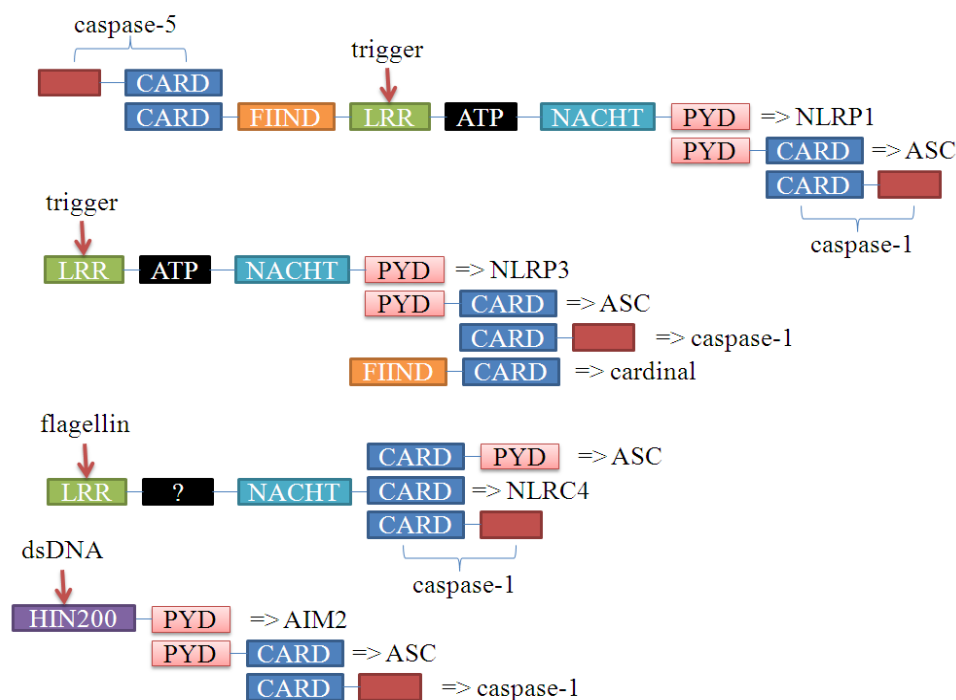
pro- and antiapoptotic roles. Moreover, these two processes may not be linked and occur in inflammatory conditions. That is why inhibition of ICE has gained importance in overcoming many inflammatory disorders [9, 22].

The importance of research for caspase-1 inhibition also depends on ILs. Proinflammatory interleukins (ILs) share a role with TNF- $\alpha$  in chronic and acute inflammation. IL-1 $\beta$ , IL-6 and IL-18 are proinflammatory interleukins for which development of antagonists or inhibitors for interleukin action has a valid place in scientific studies. As a consequence, drugs for IL-1 $\beta$  converting enzyme (ICE or caspase-1) are examined since caspase-1 intracellular protease for forming active IL-1 $\beta$ . It cleaves inactive 31-kDa pro-IL-1 $\beta$  (precursor of IL-1 $\beta$ ) between Asp116 and Ala117 for forming mature IL-1 $\beta$  that is in 17-kDa proinflammatory form. IL-1 $\beta$  levels in human macrophages can be decreased with inhibition of caspase-1. The Cys285 residue of caspase-1 which is catalytically active and the carbonyl group of Asp116 in human pro-IL-1 $\beta$  form a covalent bond. This covalent bond makes the hydrolysis of pro-IL-1 $\beta$  [23]. There is little knowledge related to initial events that regulate conversion of pro-IL-1 $\beta$  by caspase-1. Swaan et al. (2001) studied the sequential processing at ICE cut sites by combining in vitro analysis and molecular dynamics (MD) to investigate the molecular events in pro-IL-1 $\beta$  by a homology model with  $\alpha$ 1-antitrypsin. As determined from simulations, ICE cleaves the pro-IL-1 $\beta$  from two places. Asp27-Gly28 and Asp116-Ala117 are the cut places but only cut site 2 is needed for generating active IL-1 $\beta$  [24].

### **2.1.2. Caspase-1 Pathway**

Caspase-1 can be activated by protein complexes called inflammasomes. The most studied inflammasomes in literature are NLRP1 (NACHT, LRR and PYD-containing protein 1 which is also known as NALP1, DEFCAP, NAC, CARD7, CLR17.1), NLRP2

(also known as NALP2, PYPAF2, NBS1, PAN1, CLR19.9.), NLRP3 (which is also known as CIAS1, PYPAF1, Cryopyrin, CLR1.1, NALP3), AIM2 (Absent in melanoma 2) and NLRC4 (CARD, LRR, and NACHT-containing protein which is also known as CARD12, CLAN, CLR2.1, IPAF) [25-28]. In these inflammasomes, there are other proteins such as PYCARD (PYD and CARD containing protein which is also known as ASC, CARD5 and TMS1) [25-26]. Domain based interactions of the proteins in inflammasome result caspase-1 activation as shown in Figure 2.7.



**Figure 2.7** A proposed schematic of ICE interactions in inflammasomes [29-30].

### 2.1.2.1. NLRP1 inflammasome

NLRP1 is composed of PYD (pyrin domain) - NACHT (NTPase domain) - NAD (Nicotinamide Adenine Dinucleotide domain) - LRR (Leucine Rich Repeats domain) – FIIND -CARD (Caspase Recruitment Domain) domain organization. It is a basic



component of NLRP1 inflammasome [25-26]. NLRP1 inflammasome also contains ASC which can be regulated by TP53 (Cellular tumor antigen TP53) in small quantities [31-32], procaspase-1 and procaspase-5 (in humans). Caspase-1 mRNA levels increase upon overexpression of TP53. NLRP1 has interactions with ASC [15, 21, 25, 27, 29, 33-39], procaspase-1 [15, 25, 27-28, 36, 38-41] and caspase-5 [15, 29, 33, 35, 37, 42-43], which is activated by NLRP1 [44]. An adaptor protein CARDINAL interacts with caspase-5 and recruits NLRP1 (CARDINAL also interacts with NLRP1) for caspase-1 activation [45]. Furthermore, it is determined that CARDINAL can have direct interaction with active caspase-1 [46]. In mice, caspase-11 is the homolog of human caspase-5 (caspase-5 is found to be closer to caspase-11 than caspase-4.) and associates with NLRP1, ASC and procaspase-1 similar to caspase-5 [15, 19, 28, 34, 47-49]. It is formed by cleavage of procaspase-11 [50]. LPS (lipopolysaccharide) [51-52] and lysosome damage [20, 50] can trigger procaspase-11 and also its human homolog caspase-5.

The X-linked inhibitor of apoptosis protein (XIAP) is also a member of NLRP1 inflammasome due to its interaction with NLRP1, procaspase-1 and ASC. Additionally, it has an interaction with caspase-11, which is thought to secrete a component for ICE activation [34]. ASC can be stimulated with LPS [52-54] and pore forming toxins [55-56]. Additionally, lysosome damage can affect procaspase-1 [50].

NOD2 (abbreviation of nucleotide-binding oligomerization domain-containing protein 2 with other names NLRC2, CARD15, CD, BLAU, IBD1, PSORAS1, CLR16.3 [26]) is also found to have interaction with NLRP1 and is thought to be one of the possible components of NLRP1 inflammasome [25, 57-58]. Procaspase-1 also interacts with NOD2 [40, 59]. NOD2 can be influenced from LPS [60], *Listeria* which can produce Listerolysin O [56], anthrax lethal toxin (LT) [25] and a bacterial cofactor mutamyl dipeptide (MDP) which is formed from PGN (peptidoglycan) [25, 29, 33, 40, 61-64].

NLRP1 oligomerization can be started by ATP. This activation of ATP can be inhibited by Bcl-2 and Bcl-xL proteins which can directly affect NLRP1. In addition to ATP, microbial ligands which can activate NLRP1 are muramyl dipeptide (MDP) which is a component of peptidoglycan produced by Gram-positive and Gram-negative bacteria [25, 28, 33, 65-66]. MDP results in conformational changes in NLRP1. MDP binds with LRR of NLRP1 [66]. As a result of  $K^+$  flux due to ATP triggering the P2X7 receptor which binds with pannexin-1, NLRP1 can also be activated [34]. The P2X7 receptor of reversible permeable membrane can be stimulated by ATP that can tolerate hydrophilic solutes of molecular mass up to 900 Da [67]. Moreover, low intracellular  $K^+$  flux which can be mediated by LPC [68] is certainly needed for *Bacillus anthracis* which results in lethal toxin and activates NLRP1 complex and can be responsible for pore formation [29, 33, 45, 53, 69-71].  $Ca^{+2}$  flux can also affect the NLRP1 with pore formation [25].  $Ca^{+2}$  flux can result from P2X7 receptor [72-74], LPC [68] and pore forming toxins [25, 75].

Iceberg and COP are endogenous antagonists that bind the CARD domain of the zymogen (procaspase-1) and decreases the activation of caspase-1 [66]. POP1 also deactivates NLRP1, shows 64% identity to the PYD of ASC and interacts with ASC in a PYD dependent manner [76-77]. MEFV pyrin can also bind to several NLRPs including for example, NLRP1, NLRP3, with the help of its SPRY domain [77-78]. MEFV Pyrin can have a relation with ASC through PYD domains [79]. Furthermore, pyrin binds to PSTPIP1 which interacts with ASC through its BBCC domain [78]. Serum amyloid A protein (SAA) is known to have a role in activation of the MEFV pyrin [80]. Thus, some studies claim that pyrin can directly bind onto caspase-1 [78, 81-82]. Chae et al. (2008) claimed that MEFV pyrin was cleft by caspase-1 from Asp330 [83]. Its interactions with all proteins in the ICE pathway demonstrate not only activation but also inhibition features [27, 39, 78-79, 82, 84-86]. SPRY domain of pyrin can also interact with pro-IL-1 $\beta$  [78]. Consequently, its role in caspase-1 pathway is open to many questions. INCA (Caspase recruitment domain-

containing protein 17 or CARD17) which interacts with ASC [84] and procaspase-1 [54] is also a decoy protein for caspase-1. For ASC, caspase-12 is another decoy protein [70].

#### **2.1.2.2. NLRP2 inflammasome**

NLRP2 has a PYD-NACHT-NAD-LRR domain structure and is the main protein of NLRP2 inflammasome [25-26]. This inflammasome consists of ASC, CARDINAL and procaspase-1. NLRP2 interacts with CARDINAL and ASC directly. On the other hand, NLRP2 makes its procaspase-1 interaction via ASC by the CARD domain of ASC [1, 25, 27, 33, 37, 46, 84, 87-89]. There are many unanswered questions related to this inflammasome in caspase-1 activation.

NLRP2 can be triggered by LPS [27] and MDP [90]. NF- $\kappa$ B and interferons (IFN- $\gamma$  produced by IL-18 [1, 91] and IFN- $\beta$ ) proteins were found to have a relationship to NLRP2 [27]. Furthermore, SGT-1 and HSP90 need to separate from the NLRP2 inflammasome. If this does not happen, the NLRP2 cannot be active [51]. SGT-1 and HSP90 also exist in NLRP3 [41, 92-96], NOD1 [33, 41, 92, 94, 97], NLRP12 (also known as NALP12, PYPAF7) [51], NLRC4 [41, 84, 92, 96] and NOD2 [41, 97]. SGT-1 and HSP90 should be separated from NLRP2, NLRP3, NOD1, NLRP12, NLRC4 and NOD2 for inflammasome activation. Moreover, HSP90 and SGT-1 interacts with each other since SGT-1 cannot bind LRR domains of NLRPs without HSP90 [41, 92-93]. HSP90 can be stimulated by LPS directly [98].

#### **2.1.2.3. NLRP3 inflammasome**

NLRP3 is the most studied inflammasome that contains NLRP3, ASC, CARDINAL and procaspase-1. Interactions of NLRP3-ASC through PYD domains [29, 39, 99], NLRP3-CARDINAL [33, 53, 84], NLRP3-procaspase-1 [41, 72, 96, 100], ASC-CARDINAL [1, 87-88], ASC-procaspase-1 through CARD domains [1, 87-88] and

procaspase-1-CARDINAL [33, 46] exist. NOD2 was also found to bind NLRP3 [53]. NOD2 can also interact with procaspase-4 [57], which can be inhibited by COP [101].

TNFR-II (Tumor Necrosis Factor Receptor 2) can trigger NLRP3 protein for ICE activation with ATP stimulation in the absence of a microbial effect. TNF- $\alpha$  (Tumor Necrosis Factor  $\alpha$ ) stimulates the TNFR-II receptor [102]. Other stimulations of NLRP3 activation are:

- TXNIP (Thioredoxin-interacting protein) [30, 103]
- poly(I:C) (polyinosinic:polycytidylic acid) [104]
- dsRNA [104], ssRNA [70]
- lipoteichoic acid, Pam3CSK4 [42]
- imidazolequinoline compounds which are R837 and R848 [38, 70]
- ESX-1 secretion system which is secreted by *mycobacteria* [25, 105]
- bacterial RNA [33, 38, 61, 106]
- Ca<sup>+2</sup> and/or K<sup>+</sup> flux [25, 28, 53, 69, 96]
- ROS (reactive oxygen species) [96, 107-108]
- LPS [25, 109]
- pore forming toxins which are nigericin, maitotoxin, aerolysin, listerolysin O (LLO), saponin, streptolysin and hemolysinscan [30, 55-56]
- crystals (MSU (Monosodium urate), CPPD (Calcium Pyrophosphate Dehydrate), silica, asbestos, amyloid-beta, aluminum salts, skin irritant trinitrochlorobenzene, hemozoin crystals, cholesterol crystals ) [25, 41, 53, 84, 89, 103, 108, 110]
- cathepsin B which is formed due to lysosome damage [25, 111]
- MDP [61-62]
- ATP [42, 84]
- UVB (Ultraviolet B or medium wave) ultraviolet light [112].

There are many known mechanisms for activating NLRP3. One of them happens by lipid rafts which can trigger LPC (Lysophosphatidylcholine) and LPC results in NADPH oxidase activity. This activity is responsible for the ROS production which activates NLRP3 inflammasome [113]. ROS activation can also result from ATP [106, 111], P2X7 receptor [25, 106], lysosome damage [108], pannexin-1 [28], MSU [25], silica [114] and asbestos [100]. As one of the reasons for ROS production, lysosome damage can result from MSU [96], amyloid- $\beta$ , silica, aluminum salts [25, 28, 38], cholesterol crystals [108] and asbestos [96]. ATP mediated ROS can also be responsible for PI3K production [106] and ERK1/2 activation [106], which regulates Bcl-xL and Bcl-2 activation [115]. Thus, MSU and CPPD are found to affect NLRP3 protein through microtubules [53]. MSU is associated with Gout while CPPD is linked with pseudo-gout. Moreover, Amyloid- $\beta$  is related with Alzheimer's disease and cholesterol crystals are associated with atherosclerosis [108].

Another mechanism that is widely studied the P2X7 receptor, which is activated by ATP and pannexin-1 binds to P2X7 receptor. As a result of these, the P2X7 receptor and pannexin-1 complex results in a decrease of  $K^+$  flux in the cellular environment that leads to NLRP3 activation [25, 34, 61-62, 116-117]. Pannexin-1 was found to be influenced from LPS [1, 35] and MDP [28]. Furthermore, TLR agonists like LPS and/or MDP are needed for activation of caspase-1 by P2X7R [52]. Pore forming toxins [25, 28, 36], crystals [25, 93] and anthrax lethal toxin (LT) [71] are other stimulates of  $K^+$  flux decrease. Among pore forming toxins, only LLO can affect NLRP3 and ASC directly. The others use  $K^+$  flux decrease for NLRP3 activation [30, 55-56]. Biglycan can also bind onto P2X7R, TLR2 (Toll like receptor 2) and TLR4. In other words, biglycan can induce not only NLRP3 but also pro-IL-1 $\beta$  production through NF- $\kappa$ B production [95].

Iceberg and COP bind to CARDINAL [46] and procaspase-1 [46, 66, 76, 107] for inhibiting caspase-1 production. On the other hand, POP directly inhibits NLRP3 protein

[76, 84]. PYPAF3 can also inhibit procaspase-1 and pro-IL-1 $\beta$  [118]. Bcl-xL, which is also a decoy protein for caspase-1 [119], and Bcl-2 are other decoy proteins for NLRP3 [84]. Bcl-xL can interact with Alzheimer's presenilin-1 (PS1) and presenilin-2 (PS2) which demonstrates the relation of Alzheimer disease with apoptotic pathways [120]. PS1 can be cleaved by caspase-11 and caspase-8 whereas PS2 can be activated by ICE and caspase-8 [121]. MVK (Mevalonate kinase) is also associated with NLRP3, MEFV pyrin, PSTPIP-1, NOD2, caspase-1 and NF- $\kappa$ B [32].

#### **2.1.2.4. AIM2 inflammasome and RIG-I protein**

AIM2 inflammasome consists of AIM2, ASC and procaspase-1. AIM2 interacts with ASC by its PYD domain [28, 33, 96, 107, 122-123] and procaspase-1 [28, 96]. *Listeria* which can produce Listeriolysin O [56] and dsDNA with its oligonucleotide/oligosaccharidebinding domain [56, 96, 107, 123-124] are causes of activation of this inflammasome.

In addition to its role in NF- $\kappa$ B and IFN pathways [125], RIG-I also plays a role in caspase-1 activation in an ASC dependent way. Moreover, it may be an independent inflammasome or dependent on unknown inflammasome [96]. It senses dsRNA [64, 88, 104] and has a role in the triggering of IRF3 (Interferon regulatory factor 3), IRF7 and NF- $\kappa$ B [88]. IRF3, which is linked to TRIF (TIR domain-containing adapter molecule 1 and also known as TICAM1 PRVTIRB), and IRF7 can produce type 1 IFNs [64, 88, 126-127]. PYPAF7 (NLRP12, NALP12, RNO) and PYPAF5 (NLRP6, AVR, NALP6) also activate procaspase-1 by interacting with ASC. They also participate in the activation of NF- $\kappa$ B with ASC [21, 128].

### 2.1.2.5. NLRC4 inflammasome

NLRC4 contains an N-terminal CARD, a central nucleotide-binding site (NBS), and LRR [129]. Unlike NLRPs, NLRC4 does not have a PYD domain. On the other hand, it can directly bind onto ICE by the help of its CARD domain [42]. This inflammasome contains NLRC4, ASC and procaspase-1 in which NLRC4 interacts with ASC [21, 25, 29, 33, 36, 38, 53, 62, 129-130] and procaspase-1 [21, 27-28, 36, 40-41, 53, 60]. NLRC4 can be triggered by ATP [84], flagellin [25, 29, 33, 38, 41, 45, 70, 84, 124, 131] which is related with  $K^+$  and  $Ca^{+2}$  flux that will result in pore formation [25, 131], pore forming toxins [25], T3SS (Type 3 secretion system) [25, 132-133] and T4SS (Type 4 secretion system) [25].

There is no direct interaction between NLRC4 and flagellin. Despite this fact, NLRC4 recognizes flagellin but how the molecular mechanism flagellin is sensed by NLRC4 is not clear [103, 123]. Furthermore, NLRP3 and NLRC4 can be triggered at the same time with  $K^+$  flux by aerolysin [36]. NLRC4 interacts with NLRP3 [40, 56], AIM2[56], NOD2 [40, 59], and NOD1 [40, 59]. Thus, NF- $\kappa$ B can be activated by coexpression of NLRC4 and ASC [130]. NAIP5 is a protein found in mice. It is triggered by flagellin and directly interacts with NLRC4 [28-29, 53, 131, 134-136]. T4SS can be effective on NLRC4 by triggering NAIP5 [53, 136]. In similar manner, T3SS [25, 131] and T4SS [53, 134] can use flagellin interaction without having a direct interaction with NLRC4.

POP [76], INCA, COP and Iceberg [84] can inhibit the NLRC4 inflammasome by direct interaction with NLRC4. Moreover, COP was found to interact with INCA [54]. PYNOD (NACHT, NAD and PYD domains-containing protein 10 and also known as NALP10, NOD8 and NLRP10) interacts with ASC, caspase-1 and IL-1 $\beta$ . Consequently, it inhibits the functions of these proteins [137].

### 2.1.2.6. The role of NF- $\kappa$ B in IL-1 $\beta$ production

There are two steps for IL-1 $\beta$  production which are production of pro-IL-1 $\beta$  and cleavage of pro-IL-1 $\beta$  by ICE in order to become active form that is IL-1 $\beta$ . Caspase-1 cleaves pro-IL-1 $\beta$  [138] for forming the IL-1 $\beta$  [1] which is the second step for IL-1 $\beta$  production. The other one is the production of pro-IL-1 $\beta$  in nucleus by transcription, which is triggered by a mechanism that contains NF- $\kappa$ B [25, 28, 53, 59, 72, 126] and p38 MAPK [28, 86, 126]. In addition, NF- $\kappa$ B activation needs p38 MAPK [59, 86, 139]. The NF- $\kappa$ B [21, 40, 59-61, 140-142] and p38 MAPK [44, 61] can be induced by direct interaction with RIP2 protein which is receptor-interacting serine/threonine-protein kinase 2 and also known as CARDIAK, RICK and RIPK2 [143]. p38 MAPK can also be triggered by ATP [106].

RIP2 can be triggered by multiple ways. One of them is related to TRADD (Tumor necrosis factor receptor type 1-associated DEATH domain protein) and TRAF-2 (TRAF family member-associated NF-kappa-B activator) complex. When these two proteins form a complex because of the stimulation of TNFR-1 (Tumor necrosis factor receptor superfamily member 1) which is activated by TNF- $\alpha$  [101, 144-145], they stimulate the RIP2 protein [101]. TNFR-I can be turned into lipid rafts in order to form a complex. In this complex, TNFR-I and RIP2 are ubiquitylated by TRADD and TRAF2 complex [144]. NOD1 [21, 40, 60, 129, 146] and NOD2 [59, 61, 140, 142] have interaction on RIP2 for its activation. Triggered RIP2 can also interact with procaspase-1 for NF- $\kappa$ B production [27, 40, 76, 81, 101, 146-147]. As a result, adapter protein RIP2 binds not only the IKK $\gamma$  subunit of the NF- $\kappa$ B-activating IKK complex, but also CARD domain of pro-caspase-1 [40]. There is a competition between ASC and RIP2. ASC forces procaspase-1 to take place in NLRPs or NLRC4 inflammasome pathways whereas RIP2 forces procaspase-1 for becoming a part in activation of proteins such as NF- $\kappa$ B [76, 148]. COP and Iceberg are



also decoy proteins for RIP2 similar to other NALP containing proteins such as NLRP3 [54, 77, 101, 107, 141]. Caspase-8 activation is also needed for RIP2 [149].

PYPAF4 can deactivate NF- $\kappa$ B [110] and NF- $\kappa$ B can cooperate with IRF3 and IRF7 for secretion of type 1 IFNs [64, 88]. Besides, caspase-1 activation by the P2X7 receptor is dependent on TLR signals and needs of the NF- $\kappa$ B-driven protein synthesis [52]. Moreover, LPC [150], MEFV pyrin [83] and ASC [21, 81, 130, 132] can induce the NF- $\kappa$ B production whereas CARDINAL decreases NF- $\kappa$ B activation by its interaction with IKK $\gamma$  [142].

NOD1 can be activated by LPS [60, 98, 146], iE-DAP ( $\gamma$ -D-glutamyl-meso-diaminopimelic acid) which is formed from PGN [64, 84, 88] and *Listeria* which can produce Listerolysin O [56]. NOD1 has a role in pro-IL-1 $\beta$  processing through the interaction with procaspase-1 [146] and activation of p38 MAPK [59]. NOD1 interacts with several procaspases containing long prodomains such as caspase-1, caspase-2, caspase-4, caspase-8, and caspase-9 [151]. The active form of procaspase-4 [151], caspase-4 (ICH-2), can cleave pro-IL-1 $\beta$  slower than caspase-1. Caspase-4 can also convert pro-IL-18 (24 kDa) into IL-18 (18 kDa) [152]. In addition, NOD1 achieves its interaction with RIP2 and procaspase-1 by its CARD domain [146].

TLRs (Toll like receptors) that are separated into two subgroups with respect to their localization are important in NF- $\kappa$ B production. TLR1, TLR2, TLR4, TLR5 and TLR6 are the extracellular TLRs due to existing at the cell surface while TLR7, TLR8 and TLR9 are the intracellular TLRs [124]. TLR4 is triggered by LPS with MD2 (myeloid differentiation protein) [124, 153] and MSU [89]. On the other hand, *Listeria* can be detected by TLR2 with lipoteichoic acid, which is secreted by *Listeria* [56]. TLR2 also detects Pam3CSK4 [25, 154-155], MSU [89] and PGN [156]. CpG dinucleotides [25, 70, 154] and hemozoin [64] are sensed by TLR9. Additionally, TLR7 can detect

imidazolequinoline compounds (R837, R848) [25, 124, 154]. TLR3 sensors dsRNA [70, 124, 153] and poly(I:C) [88].

There are two different adaptor proteins for TLRs which are MyD88, which has an indirect relation with PI3K, and TRIF [124]. Flagellin is sensed by TLR5 which needs MyD88 and Mal adaptor proteins for NF- $\kappa$ B and p38 MAPK production. Additionally, TLR5 triggers the IRF5 protein, that mediates pro-IL-1 $\beta$  production [124]. TRIF can take place in the production of IRF3 and NF- $\kappa$ B after being stimulated with TLR4 or TLR3. It also forms a multiprotein complex with TRADD. Unlike the other TLRs, TLR4 needs both MyD88 and TRIF dependent pathways for cytokine production [38, 44, 64, 124]. On the other hand, TLR7, TLR9 and TLR2 need MyD88 as adaptor protein [64, 124]. TLR2 and TLR4 also use Mal in addition to MyD88 in order to trigger NF- $\kappa$ B production [127, 157]. TNFRSF1A (Tumor necrosis factor receptor superfamily member 1A) interacts with not only NF- $\kappa$ B but also PARP (Poly [ADP-ribose] polymerase), caspase-10, XIAP, STAT1 (Signal transducer and activator of transcription 1), TP53, TRADD, TRAF2, RIP2, p38 MAPK, IL-1 $\beta$ , BID (BH3-interacting domain death agonist) and caspase-8 [32]. TNFRSF1A gene encodes the TNFR-1 [158].

#### **2.1.2.7. Other important facts for ICE pathway**

Pro-IL-1 $\beta$  can be produced by iPLA2 (Ca independent phospholipase A2) which is triggered by K<sup>+</sup> efflux [117, 159]. iPLA2 is also responsible for the formation of active caspase-1 [74, 117]. Moreover phospholipase C, which results Ca<sup>+2</sup> efflux, can be triggered by K<sup>+</sup> efflux. The change in Ca<sup>+2</sup> efflux will result in the production of cPLA2 (Ca dependent phospholipase A2) [67, 74, 160] which can be cleft by caspase-1, caspase-4 as well as caspase-8 [161]. Both phospholipase C and cPLA2 are needed for lysosome damage [74] which mediate pro-IL-1 $\beta$  production [1]. Caspase-1 also needs

phosphorylation on Ser376 by Pak1 (the p21-activated kinase 1), which is activated by Rac1 (Ras-related C3 botulinum toxin substrate 1) and PI3K [86].

Caspase-8 also associates with TRADD for apoptotic pathways [145] and cleaves procaspase-4 [162]. Furthermore, caspase-14 can cleave procaspase-1, procaspase-4, procaspase-8 as well as procaspase-10 [163]. Caspase-10 and caspase-8 are other caspases that can cleave procaspase-1 and procaspase-4 [101, 162]. These two caspases (8 and 10) also relate with TRADD activity [101] that can interact with each other [145].

Inhibitors for the ICE pathway are grouped into two categories. One of them is exogenous inhibitors such as YopP, YopE, YopT, YopM, ESX-1, ExoU, Zmp1, mviN, ripA, Serp2, CrmA, B13R, M013, gp013L, p35 and NS1. These mainly inhibit caspase-1, ASC and NF- $\kappa$ B. Other group is endogenous inhibitors such as serpinB9, COP and POP [164]. SerpinB9 (Cytoplasmic antiproteinase 3) inhibits caspase-1. In a less amount than caspase-1, it can also inhibit caspase-4 as well as caspase-8 [165].

There is also importance of pathogens in ICE pathway since all pathogens do not follow the same path for caspase-1 activation. Each pathogen has its own path and can prefer affecting multiple NLRs as given in Table 2.1.

**Table 2.1** Pathogens and their target NLRs [25, 30, 42, 55-57, 84, 88, 132-133, 154, 166].

<b>PATHOGEN</b>	<b>VIRULENCE FACTORS, PAMPs</b>	<b>INFLAMMASOME (NLRs)</b>
<i>Salmonella typhimurium</i>	T3SS*, flagellin	NLRC4, ASC
<i>Pseudomonas aeruginosa</i>	T3SS*, flagellin	NLRC4, ASC
<i>Pseudomonas aeruginosa</i>	ExoU**	
<i>Legionella pneumophila</i>	T4SS***, flagellin	NLRC4
<i>Legionella pneumophila</i>	T4SS***, flagellin	NAIP5
<i>Shigella flexneri</i>	T3SS*	NLRC4, ASC
<i>Mycobacterium marinum</i>	ESX-1 ♦	NLRP3
<i>Mycobacterium bovis</i>	Zmp1**	
<i>Francisella tularensis</i>	FPI (T6SS?) ♦	Unknown
<i>Francisella tularensis</i>	FTT0748**, FTT0584**	
<i>Bacillus anthracis</i>	Lethal toxin	NLRP1, NOD2
<i>Streptococcus pneumoniae</i>		NOD2
<i>Mycobacteria</i>		NOD2
<i>Staphylococcus aureus</i>	α-Toxin	NLRP3
<i>Listeria monocytogenes</i>	Listeriolysin ♦, flagellin	NLRC4
<i>Listeria monocytogenes</i>	Listeriolysin ♦	NLRP3
<i>Listeria monocytogenes</i>	Listeriolysin ♦	AIM2
<i>Listeria monocytogenes</i>	Listeriolysin ♦	ASC
<i>Poxvirus</i>	M13L**	Unknown
<i>Streptomyces hygroscopicus</i>	Nigericin	NLRP3
<i>Aeromonas hydrophila</i>	Aerolysin	NLRP3
<i>Marina dinoflagellates</i>	Maitotoxin	NLRP3
<i>Bacillus brevis</i>	Gramicidin	NLRP3
<i>Escherichia coli</i>	Toxin?	NLRP3
<i>Escherichia coli</i>	T3SS*	NLRC4
<i>Sendai virus</i>	?	NLRP3
<i>Influenza virus</i>	?	NLRP3
<i>Burkholderia pseudomallei</i>	T3SS*	NLRC4
<i>Francisella tularensis</i>		AIM2
<i>Anaplasma phagocytophilum</i>		NLRC4, ASC
<i>Gambierdiscus toxicus</i>	Maitotoxin => K <sup>+</sup> efflux	NLRP3
<i>Streptomyces hygroscopicus</i>	Nigericin => K <sup>+</sup> efflux	NLRP3
<i>Neisseria gonorrhoe</i>		NLRP3
<i>Candida albicans</i>		NLRP3
<i>Plasmodium spp</i>	Hemozoin	NLRP3

\* Type III secretion system that is a protein appendage in gram negative bacteria. \*\*Virus factors that deactivate inflammasome. \*\*\* Type IV secretion system that is transporter of gram negative bacteria. \*\*\*\* ♦ Activators of inflammasomes.

### 2.1.2.8. Roles of caspase-1

Caspase-1 cleaves pro-IL-1 $\beta$ , pro-IL-18 and pro-IL-33 in order to activate forms which are IL-1 $\beta$ , IL-18 and IL-33, respectively [1]. As a consequence, ICE takes a part of pyroptosis activities in the cell [126]. ATP suppresses activation of IL-1 $\beta$  in mouse microglial cells using a Rho-dependent pathway, which is using prokaryotic protein in termination of transcription [68]. IL-1 $\beta$ , which also triggers the NF- $\kappa$ B production [167], can inhibit SAA (Serum amyloid A protein) [168].

ICE has also a role in the production of SREBPs (Sterol Regulatory Element Binding Proteins) [25, 36]. Secretion of pro-IL-1 $\alpha$  also needs caspase-1 activity [169]. In addition, pro-IL-1F7 can be cleft by ICE [1, 170]. It can also be cleft by caspase-4 but in a less efficient way [170].

BID (B cell lymphoma-2-interacting domain) protein, which is a pro-apoptotic BH3-only protein and connects the extrinsic and core intrinsic apoptotic pathways, can be cleft by caspase-1 [171], caspase-8 [172-173] and caspase-10 [174].

Mal [157], ataxin-3 (ATXN-3), atrophin-1 (ATN1) [175] and E3 ubiquitin-protein ligase NEDD4 [176] can also be activated by ICE cleavage. ATN1 can also be activated by caspase-8 [175]. Calpastatin was found to be cleft from three different sites (Asp137, Asp203 and Asp404) by ICE [177]. Furthermore, poly(ADP-ribose) polymerase (PARP) can be cleft by ICE with requiring higher ICE concentrations than pro-IL-1 $\beta$  case. Caspase-4 can also cleave this enzyme [178]. Thus, caspase-1 can associate with Alzheimer's disease due to cleaving the PS2 protein [121]. Caspase-1 also cleaves Huntingtin which is related with Huntington disease [175]

FGF-2 (Heparin-binding growth factor 2) needs caspase-1 activity despite not being a caspase-1 substrate [169]. Epidermal growth factor receptor (EGFR) is determined to be cleft by ICE [179]. EGFR can also interact with Pak1 [180] and Huntingtin as an adaptor

protein for EGFR [181]. Androgen receptor (AR) is another protein that can be activated by ICE as well as caspase-8 [175]. Moreover, AR has an interaction with EGFR [182].

Cell division protein kinase 11A (CDK11A, CDC2L2, CDC2L3, PITSLREB) and cell division protein kinase 11B (CDK11B, CDC2L1, CDK11, PITSLREA, PK58) are cleaved by caspase-1 and caspase-8 [183-184]. Pak1 can be inhibited with these PITSLRE kinases [185]. Caspase-1 and caspase-8 can also be related with sporadic Parkinson's disease since they can cleave parkin [186]. These two caspases can also activate microtubule-associated protein tau (MAPT, MAPTL, MTBT1, TAU) [187]. PS1 regulates phosphorylation of the MAPT and mutations related to Alzheimer's disease in PS1 are responsible for amyloid- $\beta$  in mice [188]. MAPT and parkin also interact with each other [189].

B-cell receptor-associated protein 31 (BCAP31, BAP31, DXS1357E) has two caspase recognition sites. Caspase-1 and caspase-8 can achieve BCAP31 cleavage. In procaspase-8 production, BCAP31 can take place with Bcl-2 or Bcl-xL as a member in a complex of the endoplasmic reticulum [190-191].

Actin, aldolase, TIM (triosephosphate-isomerase), GAPDH (glyceraldehyde-3-phosphate dehydrogenase),  $\alpha$ -enolase and pyruvate kinase can be cleaved by ICE which indicates the role of caspase-1 in multiple cellular pathways such as glycolysis [192]. Sptan-1, TFAP2A (Transcription factor AP-2-alpha) and PPAR- $\gamma$  (Peroxisome proliferator-activated receptor gamma) are other known substrates for caspase-1 [193]. Sptan-1 also interacts with actin [32, 194]. Caspase-1 can also be an important stimulator of TP53 dependent apoptosis [32, 195]. Furthermore, TP53 can induce NLRC4 transcription, which can have an effect on TP53 dependent apoptosis [32, 196]. TP53 also interacts with Stat-1 (signal transducer and activator of transcription 1) for DNA damage-induced apoptosis [197]. Stat-1 can be activated by IL-1 $\beta$  [32, 198] and has a relation with IL-18 [199].

### 2.1.3. Diseases related to caspase-1

Several autoimmune disorders, ischemia, and various cancers result from an increased amount of IL-1 production in which IL-1 $\beta$  is produced by ICE [200]. Furthermore, the pathogenesis of a number of peripheral inflammatory disorders is related to increased IL-1 $\beta$  production such as inflammatory bowel disease. Caspase-1 inhibitors can be a solution for brain damage in stroke, myocardial ischemia, traumatic brain injury, sepsis, Crohn's disease, intestinal inflammation, MS (multiple sclerosis) and fulminant liver disease since IL-1 $\beta$  production will also be inhibited by ICE inhibition [12-13, 201-203]. Acute life-threatening conditions such as autoimmuno-aggressive disease (for instance, rheumatoid arthritis) are related to caspase-1 and ICE has a role in proinflammatory systems during experimentally induced pancreatitis. Alzheimer's disease and Parkinson's disease also related with ICE. Additionally, overexpression of caspase-1 is found to have an effect on Huntington's disease with a rise in the length of disease progression [13, 204]. ICE can also participate in inflammation resulting from cowpox virus and in cell death induced by deprivation of nerve growth factor [205]. Furthermore, infection with *Shigella*, *Entamoeba histolytica* or *Salmenolla typhimurium* depends on caspase-1 [203].

In order to identify the problems related to ICE and IL-1 $\beta$  separately, there have been some experimental studies especially on mice. For instance, Talanian et al. (1997) found caspase-1 deficient mice were determined to be deficient in IL-1 $\beta$  maturation as well as resistant to endotoxic shock [206]. In another study by Fantuzzi et al. (1999), IL-1 $\beta$  or IL-18 is not released by ICE-deficient mice. ICE-deficient mice are resistant to lethal endotoxemia whereas IL-1 $\beta$ -deficient mice are not, since ICE-deficient mice cannot also produce IL-18. Furthermore, ICE-deficient mice are not protected against urate-induced inflammation and there is a reduction of circulating IFN- $\gamma$  level [18]. Sifringer et al. (2007)

found there was a certain role of ICE in the onset of trauma-triggered apoptosis due to increased expression of ICE by experimental work on mice. There were also previous studies related to brain injury and caspase-1. ICE-deficient mice were observed to be resistant to hypoxia–ischemia, which occurs before brain injury. By inhibiting ICE, caspase-1 dependent interleukins (IL-1 $\beta$  and IL-18) can also be controlled which can cure brain problems [207]. Fink et al. (1999) also determined caspase-1 inhibition reduces post-traumatic brain injury [208]. Tringali et al. (1999) found that depolarization and a NO-donor stimulates IL-1 $\beta$  production in which caspase-1 has a role in rat hypothalamus. The cell permeable form of caspase-1 inhibitor was used by Tringali et al. (1999) and was not found effective on basal release while it was effective on induced mechanism by K<sup>+</sup> and sodium nitroprusside [209]. Overexpression of ICE in the skin of mice was determined to be effective on atopic dermatitis-like lesions and excessive scratching. [210]. Hasegawa et al. (2009) stated that when the brain is injured, overexpression of IL-1 $\beta$  happens which leads to harmful situations for the tissue [211].

## 2.2 Drugs

### 2.2.1. Desired features of drugs

There are many rules to determine a compound for its drug-likeness. In order to define the suitable drug for a system, there are numerous rules that have been determined by scientists. Each group of a rule actually defines a filter property. The filters are mainly classified into filters for high drug-likeness, filters for orally bioavailable drugs and filters for blood brain barrier permeable drugs. There are also other types of filters for different purposes.

Ghose et al. (1999) analyzed the CMC (Comprehensive Medicinal Chemistry) database and found that a drug-like molecule should satisfy the following criteria: (1) an organic compound having a calculated log P (LOGP\_ a method for measuring of



hydrophobicity) between -0.4 and 5.6, a molar refractivity (AMR) between 40 and 130, a molecular weight between 160 and 480, and the total number of atoms between 20 and 70, (2) containing some of the following groups: a benzene ring, a heterocyclic ring (both aliphatic and aromatic), an aliphatic amine (preferably tertiary), a carboxamide group, an alcoholic hydroxyl group, a carboxy ester, and a keto group, (3) being stable in the physiological buffer without a reactive functional group or structural part [212]. Mozziconacci et al. stated that drug likeliness constraints are having less than six rings and seven halogens as well as having more than one oxygen and one nitrogen [213]. There are other filters for analyzing high drug-likeness. For instance, Lee et al. (2001) looked at the mean molecular weight and logP which were 356 and 2.1, respectively. Moreover, Oprea [0<Hydrogen bond donor (HDO) <2, 2< Hydrogen bond acceptor (HAC) <9 and 2<Rigid Bond (RB) <8] and Walters & Murcko [MW=200-500, HDO=0-5, HAC=0-10, RB=0-8, Number of heavy atoms=20-70, Charge=-2 to 2+] filters can also be used for analyzing whether a compound satisfies high drug-likeness criteria [214-216].

Filters for orally bio-available drugs are Lipinski, Fichert, Palm, Enhanced Palm and Veber. Lipinski et al. (2001) determined the 'rule of 5' that claims poor adsorption and permeation of a drug happens if there are more than 5 H-bond donors (that is the sum of OHs and NHs) and 10 H-bond acceptors (that is the sum of Ns and Os). Moreover, the molecular weight of the vaccine should not exceed 500 and LogP should be not over 5 (MLogP\_Moriguchi LogP\_ is over 4.15). This rule is not valid for the compound classes that are substrates for biological transporters. Fichert et al. (2003) used the constraints of molecular weight and logD (an apparent partition coefficient for any pH value) which should be less than 500 and between 0 and 3, respectively. Palm filter only considers the polar surface area (PSA) and drugs with  $PSA > 139 \text{ \AA}^2$  are only 10% absorbed whereas drugs with  $63 \text{ \AA}^2$  are completely absorbed. Veber et al. (2002) also determined the same value of (140) PSA and RB of 12 maximum for a drug [217-220].

Murcko et al. (1999) found that MW=200-450, logP=0-5.2, HAC= max.4, HDO=max.3 and RB=max.7 for blood brain permeable drugs. Van de Waterbeemd et al. (1998) determined molecular weight and PSA should be less than 450 and 90 Å<sup>2</sup> [221-222].

### **2.2.2. Peptide drug properties**

Usage of peptides in vaccine development is attracting more and more attention because of their relatively specific mode of action which reduces the risk created by other small molecular drugs or larger charged molecules by using low doses. There are many benefits of peptide based drugs such as having no infectious material. Furthermore, these vaccines can be analyzed easily due to well established methods. They are also economically available and their storage, transport as well as distribution do not need a cold-chain facility due to their freeze-dried preparation. In addition, there is no risk of genetic integration or recombination. The peptide usage in vaccine search will lead to more flexible and simple ways to fight diseases. In addition, there are many companies that support clinical trials related to peptide based vaccines. For instance, multi-epitope (antigenic determinant) usage on the treatment of hepatitis C viral infections was found to be both succeeding and proceeding in addition to many reported peptide drugs for cancer cure [223-224]. Molecular features determined experimentally and theoretically such as hydrogen bonding potential and molecular size are commonly used for finding how suitable a peptide or a conventional drug is in terms of membrane permeability. Especially, hydrogen bonding capacity is a valid parameter for membrane permeability of peptides [225].

Unal et al. (2009) studied structure-function relations of peptides due to their usage in blocking the activities of proteins which are responsible for some diseases. They stated that there is a need for the right type of sequence search in peptide design [226]. Vives et al. (2008) also stated that there were peptides evidenced to target specific tissues or cell

types. For example, TSPLNIHNGQKL targeted the tissue of the human head and neck solid tumours and LTVSPWY worked on breast carcinoma. In addition, there were tripeptides such as RGD and NGR that had tissue targets of integrin receptors and tumour neovasculatures, respectively. These peptides were also called cell-targeting peptides. Moreover, there are also cell-penetrating peptides which are efficient in delivering many molecules. As an example, one is known as *Antennapedia peptide* and marketed as *Penetratin*. However, there is some debate about the entry mechanism of the cell-penetrating cell (CPP). There are also some problems and limitations of CPPs for drug delivery *in vivo* [227]. There are many peptide based drugs on the market and clinical trials are being conducted for many diseases. Some of these peptide based drugs are given in Table 2.2 with their names and the range of aminoacid numbers that they contain.

**Table 2.2** The peptide based drugs in market and clinical trials [228]

Disease	Peptide based drug	Range of aminoacid number
Diabetes	GLP-1, Exenatide, Liraglutide, ZP10, Pramlinitide, PYY, Glucagon	29-39
Gastrointestinal	Teduglutide, Delmitide	10-33
Oncology	Bortezomib, Cilengitide, Leuprorelin, Histelin, Goserelin, Stimuvax, GV1001	2-25
Cardiovascular	Nesiritide, Eptifibatide, Bivalirudin, Icatibant, Rotigaptide	6-32
Immunotherapy	Cyclosporin, MPB8298	11-17
Acromegaly	Octreotide, Lanreotide	8
Enuresis	Desmopressin, Lypressin, Terlipressin	6-9
Antiviral	Enfuvirtide, Thymalfasin	28-36
Antibacterial	Daptamycin, Bacitracin, Gramidicin, Colistin, Pexiganan, Omiganan	10-22
Antifungal	Caspofungin, Micafungin, Anidulafungin, Histatin, Lactoferrin	6-12
Central Nervous System	Conotoxin, Nemifitide	5-25

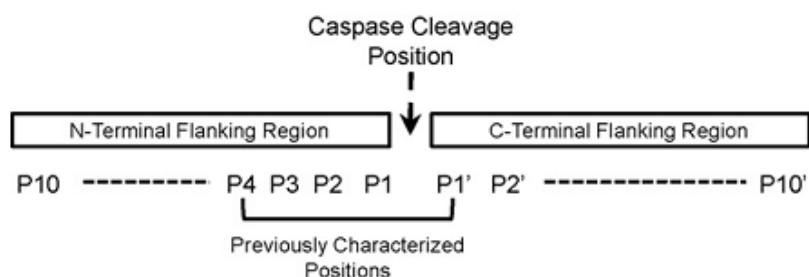
There are other examples for peptide based drugs in addition in Table 2.2. For example, Dao et al. (2008) studied peptide drugs for myeloid leukaemias in clinical trials.

They determined that these early clinical results were encouraging for future research because peptide based vaccines induce immune responses that are connected to a clinical advantage [229]. In cancer treatments, usage of peptides immunotherapeutic cancer drugs is important but there is a certain need for research related to interactions between peptide and major histocompatibility complex (MHC). Since high-affinity peptides which are poorly immunogenic, low-affinity peptides can be an alternative way for peptide based cancer vaccines which will create alternative approaches [230]. There are also many peptide drugs related to the neuroprotection such as Gonadotropin releasing hormone (GnRH, LH-RH), insulin and vasopressin which are also used clinically. Moreover, such peptide based drugs are useful in terms of doing the same work as large protein factors. As a result, the delivery to the brain will be easier than using larger molecules [231]. Beck et al. (2007) studied peptides in three different cases by using clinical techniques. In one of the cases, the Respiratory Syntycial Virus (RSV) subunit vaccine studied was Peptide G20 that had 69 amino acid residues and was found to induce pulmonary pathology in mice. Moreover, it was successfully synthesized in a practical manner. In addition, clinical trials have shown promising developments in the peptide-based melanoma vaccines and therapeutic monoclonal antibodies [232]. Iwai et al. (2006) studied lipopolysaccharide-induced IL-1 $\beta$  production in cultured rat astrocytes and found that glucagon-like peptide-1 was useful in terms of deactivating IL-1 $\beta$  production [233].

Using L or D-forms of aminoacid can also affect the physiological properties considerably. For instance, Shaji et al. (2009) stated that natural vasopressin, that contains L-arginine, is active in the water that is injected into a rat in high doses while desmopressin, that contains D-arginine, is twice as active at the 75<sup>th</sup> fraction of the dose used in vasopressin [234].

### 2.2.3. Drugs for caspase-1 in literature

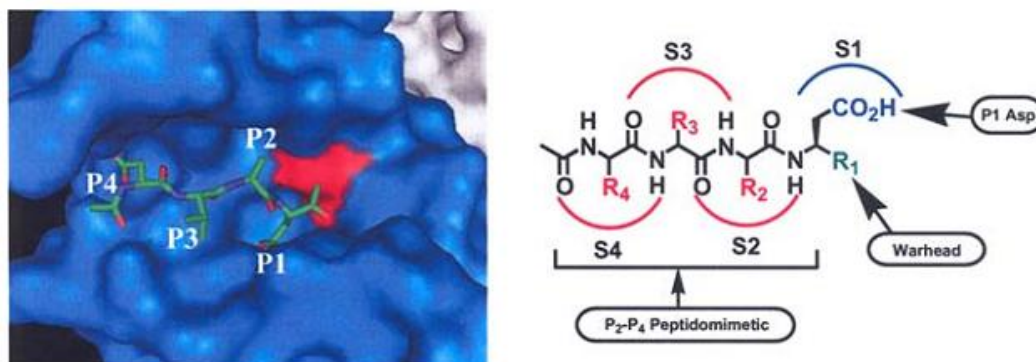
There have been a numerous studies for peptide inhibitor designs for ICE just like for other proteins. The initial potent inhibitors of ICE were C-terminal peptide based on the pro-IL-1 $\beta$  cleavage site [23]. According to the substrates, the significant aminoacids for each position of the binding site were also determined by Shen et al. (2010). The position schematic on binding site is given in Figure 2.8 [193].



**Figure 2.8** Caspase cleavage site nomenclature [193].

Romanowski studied the conformational changes of ligand-free and ligand-bound structures of caspase-1 by using both Asp based inhibitor and malonate. They also focused on the effect of mutation of Cys285 to Ala. Additionally, sodium malonate was determined to be sufficient like aspartic acid in terms of inhibiting caspase-1 [22]. However; O'Brien et al. (2004) stated that a typical caspase inhibitor should be a small tetrapeptide-based inhibitor which is composed of three structural parts. These parts are the warhead, the P<sub>1</sub> aspartic acid and the P<sub>2</sub>-P<sub>4</sub> peptidomimetic just like the caspase-3 inhibitor given in Figure 2.9. The warhead contains an electrophile that has a reaction to the nucleophilic cysteine residue of the active site. It can be irreversible or reversible from which reversible ones are mostly selected because they do not bind to the non-target proteins for caspase-1. In the P<sub>1</sub> position, caspases have a strong preference for Aspartic acid residue and non-peptide substitutions for P<sub>2</sub>-P<sub>4</sub> regions are suitable for developing the pharmacokinetic features of

molecules. Despite having some conflicts related which apoptotic caspase should be targeted, the inhibition of caspase-1 will be sufficient as an anti-inflammatory effect [13].

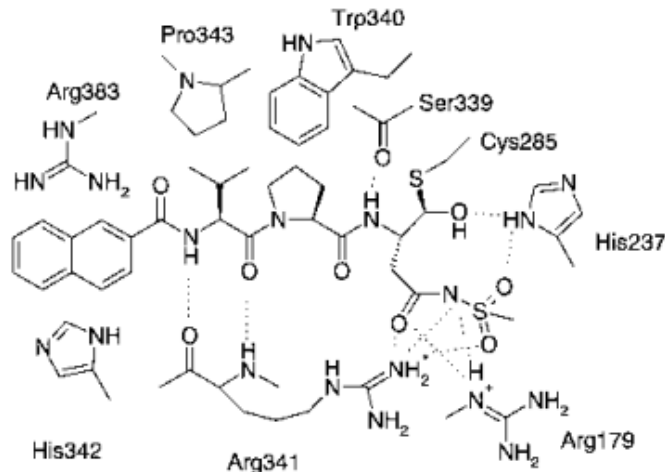


**Figure 2.9** Caspase-3 and its tetrapeptide inhibitor (Ac-DVAD-FMK) [13].

In addition, there are other studies that conducted that tetrapeptide inhibitors were useful. Tetrapeptidic inhibitors for caspase-1 was found to be successful in terms of blocking maturation, release of IL-1 $\beta$  in cultured cells and animals as well as inflammation in animal models such as Ac-YEVD-pNA. This tetrapeptide was found better than Ac-YVAD-pNA due to its high value of kinetic constants [206]. Ac-DEVD-CHO and Ac-YVAD-CHO were other inhibitors for caspase-1 with  $K_i$  of 17 nM and 0.76 nM, respectively [9]. Rano et al. (1997) found that the optimal tetrapeptide recognition sequence for caspase-1 was WEHD. However, many studies showed that YVAD is the best tetrapeptide but Rano et al. (1997) found Ac-WEHD-CHO had  $K_i$  of 56 pM. Despite having an obvious difference between the sequences of human pro-IL-1 $\beta$  (YVHD), pro-IL-1 $\beta$  was thought to have additional endogenous substrates [235].

On the other hand, there are studies related to using less than four residues in a drug due to penetration problems and being nonspecific (tetrapeptides can bind onto other caspases.) L-709049 (Ac-YVAD-H) was used as a model for improving the A-NapCO-VPD-CH<sub>2</sub>Oph compound because it was found to be a potent caspase-1 inhibitor. L-

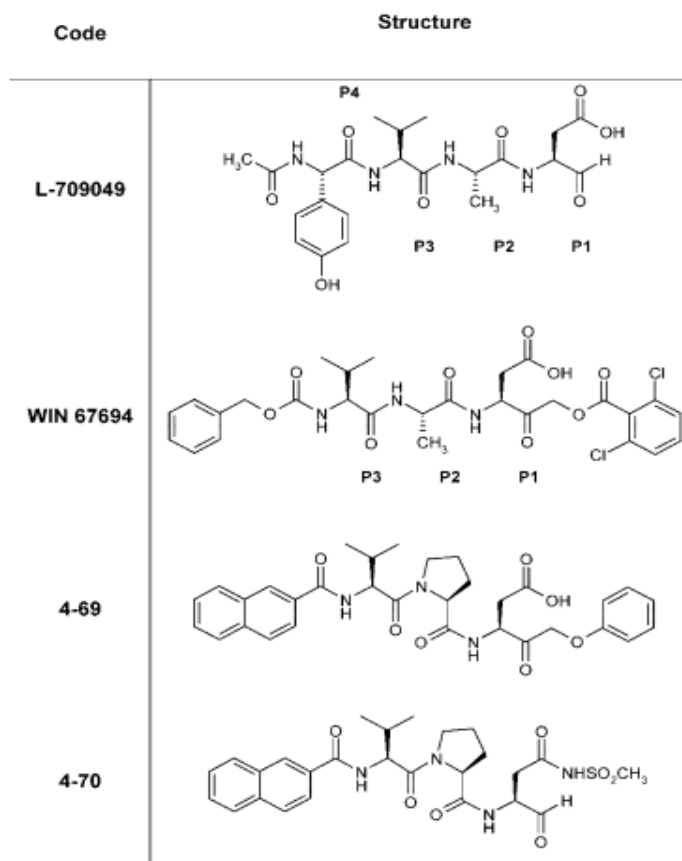
709049 forms a covalent bond with Cys285 of caspase-1 and makes three hydrogen bonds. They found 27c as a potent inhibitor. Thus, 27c is the most hydrophilic drug with high potency and cell permeability whose interaction with caspase-1 is given in Figure 2.10. Additionally, this compound was found successful *in vivo* studies in mice [205]. Karanewsky et al. (1998) studied the caspase inhibitors by considering the four residues inhibitors and their modifications. Firstly, they kept the P1 position as Asp residue due to the specific need of ICE and they found that bicyclic structures were high potent inhibitors for ICE [236].



**Figure 2.10** The interaction between 27c and ICE [205].

The tripeptide irreversible inhibitor Ac-VAD-CH<sub>2</sub>O-DCB and a reversible inhibitor Ac-YVAD-CHO reduce the IL-1 $\beta$  level in mice for rheumatoid arthritis. The treatment for rheumatoid arthritis can be very long owing to the long-term dosing need but the toxicity can be minimized by using a small molecule inhibitor. z-VAD-FMK was also determined to inhibit the activity of caspase-1 which had an effect on the apoptotic pathway. This peptide drug also affected the decrease in tissue death following spinal cord injury because caspase-1 stimulates the expression of IL-1 $\beta$  and IL-18 [13]. Hayashi et al. (2001) found

that the broad caspase family inhibitor (z-VAD-FMK) was good for inhibition of IL-1 $\beta$  [237].



**Figure 2.11** The structures in study of Wagner et al. (2006) [23].

Wagner et al. (2006) determined that L-709049 whose structure is given in Figure 2.11 inhibited caspase-1 with a  $K_i$  of 0.76 nM and found responsible for low levels of IL-1 $\beta$  but had poor penetration into cells on account of weak cellular activity. *N*-Ac-Tyr fragment of L-709049 was replaced with a benzyloxycarbonyl group and called WIN 67694. This tripeptide resulted in a decrease in IL-1 $\beta$  levels in the *in vivo* mouse model. It was an irreversible peptide inhibitor that can covalently bond with Cys285. The 4-69 and 4-70 tripeptides were determined to have weaker activity. Thus, researchers found that there

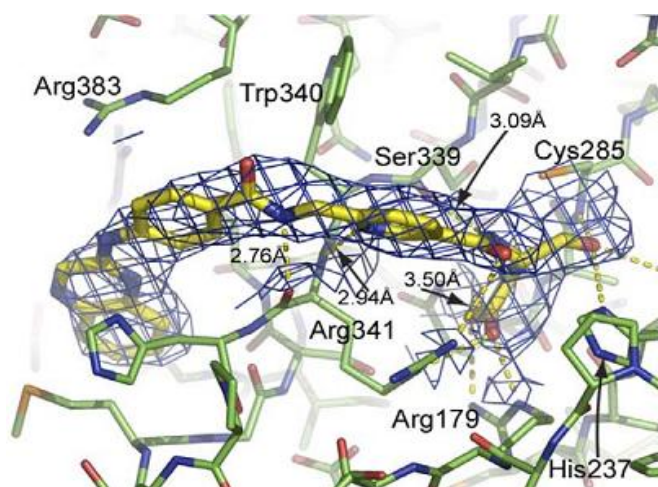


is hydrogen bonding between inhibitor and peptide existed along peptide backbone of inhibitor [23].

Linton et al. (2002) showed that the Ac-DEVD-H inhibitor can be shortened to a potent inhibitor of dipeptide for caspases such as ICE [238]. Nedev et al. (2005) also focused on dipeptides instead of tetrapeptides since the known commercial tetrapeptides is not good in terms of penetration. Additionally, using tetrapeptides is not suitable for therapeutic usage. Moreover, they decided to use phenylglycine which is not a natural amino acid due to potent caspase inhibition. For ICE, the benzyloxycarbonyl-phenylglycyl-aspartyl benzoyloxymethyl ketone (Z-Phg-Asp-CH<sub>2</sub>OCO-Ph) was found as a potent inhibitor [239]. z-Asp-CH<sub>2</sub>OC(O)-2,6-Cl<sub>2</sub>Ph contains only a single amino acid residue that was determined to be sufficient for caspase-1 inhibition [240].

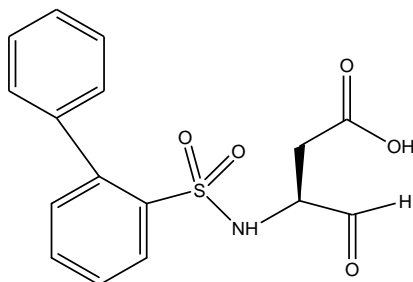
In some cases, there is a need for addition of peptides to drugs. Schlosser et al. (2001) used Ac-AAVALLPAVLLALLAP-YVAD.CHO for caspases-1 inhibition in pancreatic carcinoma cells. It works in terms of inducing cell death in carcinoma cells and also modulates the production levels of bcl-2 family proteins. The C-terminal YVAD.CHO is for specific inhibition of caspases-1. The other peptides (residues 1 to 16) are for cell permeability without cytotoxic effects which corresponds to hydrophobic region of signal peptide of the Kaposi fibroblast [241].

Fahr et al. (2006) studied tethering for identification of caspase potent inhibitors. Their initial structure contained tricyclic fragment attached to a methylene scaffold bearing an Asp residue. By using tethering, they studied small fragments that specifically bind in the S4 region of caspase-1. Firstly, they considered a hydrophobic P2 substitute and then focused on P4. Finally, they found the structure *11* (Figure 2.12) as a potent inhibitor due to its crystal structure with caspase-1 and being energetically preferable [242].



**Figure 2.12** Crystal structure of compound *11* (PDB ID code: 1RWW). Yellow dashed lines represent hydrogen bonds [242].

Many laboratories have also developed non-peptide based caspase-1 inhibitors to overcome the difficulties related to cellular activity and oral bioavailability [23]. There have been many different nonpeptide fragment based drugs studied for ICE inhibition for instance, nitric oxide and serine proteinase inhibitor as intracellular inhibitors inhibit caspase-1 [203]. Fiorucci et al. (2001) also claimed that NO-NSAIDs (Nitric-oxide releasing nonsteroidal anti-inflammatory drugs (NO-NSAIDs) are caspase inhibitors. These drugs are formed by adding a nitroxybutyl moiety to a conventional NSAID and inhibit inflammation by cyclo-oxygenase (COX)-dependent and -independent effects. [243]. Shahripour et al. (2001) studied ICE inhibitor containing a diphenyl ether sulfonamide and held an X-ray crystal structure of a representative member bound to caspase-1. They aimed to increase favorable prime side hydrophobic interactions by using their initial model from their previous study [201]. Shahripour et al. (2002) also studied small nonpeptide inhibitors for ICE by modifying the known potent inhibitor Ac-YVAD-CO-(CH<sub>2</sub>)<sub>5</sub>-Ph (19 nM) with getting the K<sub>i</sub> to better values. In the end, they determined structure *12* whose structure is illustrated at Figure 2.13 was the best among all trials with a K<sub>i</sub> value of 1.6 μM [244].



**Figure 2.13** Structure 12 [244].

Ross et al. (2007) found that VRT-018858 (*Pralnacasan*) was successful for the inhibition of caspase-1 and Ku et al. (2001) determined the  $K_i$  of this nonpeptide vaccine as 1.3 nM. It was defined as neuroprotective in rats [245-246].

The certain weak electrophiles were found to be capable of deactivating cysteine proteases and these inhibitors were also applied on caspase-1. The alternative leaving groups were developed for the inhibition of ICE. In addition, ketone inhibitors were found to be time-dependent inhibitors of cysteine proteases. His237 and Cys285 were found to be important due to taking part in the catalytic reaction and covalent bond [247]. Peptide (Acyloxy) methyl ketones using an appropriate peptide sequence (Ac-YVAD-CH<sub>2</sub>-OC(O)Ar) were also found to be potent, competitive and irreversible inhibitors for ICE by Thornberry et al. (1994). By using Cys285 as active site, this compound was selective for ICE [248]. In addition, Brady et al. (1999) evaluated 619 aspartic ketones with different types of prime-side groups as inhibitors of caspase-1 [249].

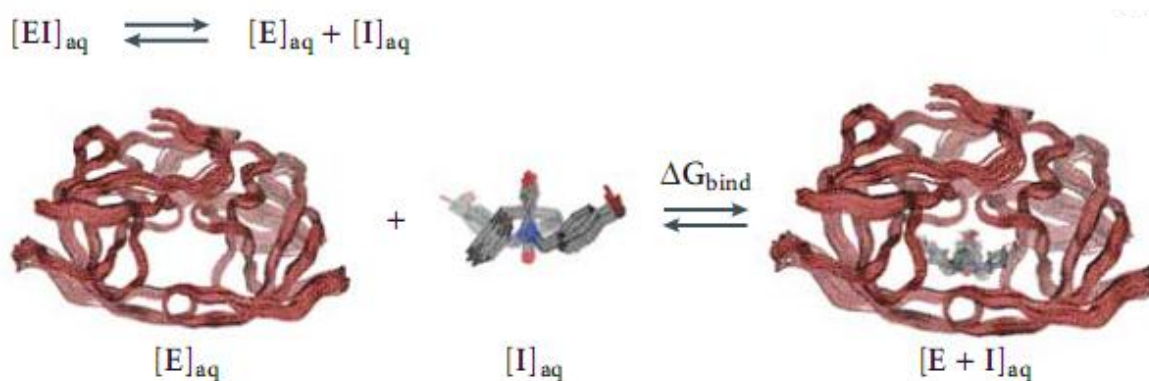
There are also natural inhibitors of caspase-1. For example, cowpox serpin CrmA which is a serpin as an antiapoptotic is a potent inhibitor with  $K_i$  lower than 20 nM and selective inhibitor for caspase-1 but it can affect the other Group I caspases (caspase-4 and caspase-5) [10, 250]. Baculoviral protein p35 also inhibits caspase-1 and serpin PI-9 was found successful in caspase-1-mediated cell death in vitro and in smooth muscle cells [251].

### 2.3 Docking

The binding sites show a selective recognition of ligands as found by many experiment techniques like X-ray crystallography. When these active sites are known, the explanation of binding mode and the determination of suitable ligands are useful using computational methods. There are some problems about usage of computational methods because the methods find different answers to real situations. A powerful search technique and a good understanding the process of binding are needed for the solution of docking problem [252].

Docking methods are the most commonly used tools for predicting of binding modes of compounds on the target proteins. Consequently, many drugs designed by using docking tools in addition to experimental work are found to be suitable for many diseases. That is why drug companies rely on computational works for designing new potential drugs [253]. Prediction of the structure of the intermolecular complex formed between ligand and protein can be done by molecular docking. From rigid complex to full flexibility of ligand, lots of docking strategies have been attempted. Additionally, accepting some parts of the receptor as flexible is also considered since both binding site and binding orientation can be affected [254]. Carlson et al. (2002) stated that designed drugs onto proteins should not be taken as rigid. The historic lock-and-key and induced-fit models are not valid anymore due to theory that states ligands can bind to a protein in any conformation. That is why an ensemble of conformational states happens. Anfinsen proposed “thermodynamic hypothesis”, which states proteins fold to a conformation in which the free energy of the molecule is minimized. This hypothesis is a basis for most of the methods for protein structure prediction [255]. Since three dimensional structure is the basic factor that determines the protein function. Ligand binding is a basic component of the protein function. Binding is the recognition of proteins to their ligands for transport, signal transduction or catalysis. In light of this, drug discovery is done considering the binding of

a ligand on protein. The determination of a functional site location from the literature or other methods is also crucial for computational analysis of binding [256]. Docking predicts the structure of complex between inhibitor and enzyme under equilibrium conditions as seen in Figure 2.14 [257].



**Figure 2.14** The binding of inhibitor Dmp323 to HIV protease (PDB ID=1BVE) [257].

The free energy of binding ( $\Delta G$ ) is depended on binding affinity as follows:

$$\Delta G = -RT \ln K_A \quad (1)$$

$$K_A = K_i^{-1} = \frac{[EI]}{[E][I]} \quad (2)$$

Conformation does not affect  $K_A$  whereas biological activity depends on conformation. Steric, electrostatic, hydrogen bonding, inhibitor strain and enzyme strain are crucial for  $[EI]$  [257].

There are numerous methods for docking the ligands on macromolecules such as molecular dynamics, Monte Carlo methods, genetic algorithms and fragment-based methods [258]. Docking methods are composed of two components which are a search

strategy and a scoring function. The features of estimated search space for a ligand with four rotatable bonds and six rigid-body alignment parameters will need 2,000,000 years of computational time with a rate of 10 configurations per second since the active site will be  $10^3 \text{ \AA}^3$  with  $6 \times 10^{14}$  configurations. In order to overcome this difficulty; constraints, restraints and approximations are used for making the problem more practical to handle.

*Systematic search:* The combinatorial explosion problem is seen in this algorithm. DOCK, FLOG, Hammerhead are some of the example programmes in this class. All docking programmes use different stepwise search algorithm for making the ligand incrementally grown in the active site [257].

*Random search:* The random formation from using a single ligand or a population of ligands is done. Genetic algorithms (GA) and Monte Carlo algorithms are the most known methods which are applied in AutoDock, DOCK and GOLD. AutoDock 3.0, DIVALI and DARWIN are other programs that use GA. However, AutoDock 3.0 and DIVALI consider the AMBER force field whereas DARWIN uses the CHARMM-AA molecular mechanics force field. Tabu search algorithm is also an example which uses the differences of RMSD between new and old molecules such as in the case of PRO\_LEADS. EUDOC and SYSDOC are other computer tools for systematic search [257-258].

*Simulation Methods:* Molecular dynamics and Monte Carlo are examples that consider energy minimization in a system. DOCK and AutoDock include such methods [257]. Molecular dynamics (MD) simulations can be performed with AMBER or CHARMM by Newton's equations of motions. The quality of the results from a standard MD depends on starting configuration and that is why MD can find local minimum. Monte Carlo (MC) simulations contain the combination of atomistic potential energy and stochastic optimization methods. A major benefit of MC compared to MD as gradient based techniques is using a single energy function. Energy barriers can be passed with choice of moves. For instance, Metropolis MC chooses random Cartesian moves based on

Boltzmann probability. In initial versions of AutoDock, Metropolis MC simulated annealing was used with AMBER force field. Prodock also uses MC for binding the flexible ligands to a flexible binding site but performs each move after that it rejects the structure. Glide and DockVision also perform MC based docking [258].

*The fragment-based methods* divide the ligand into separate portions initially. After that, the divided ligands were docked and then modified at the portions with linking of fragments. Due to taking decisions about which functional group of ligand is vital, a base fragment choice is crucial. *de novo* design techniques are used. As examples of *de novo* techniques, most known tools are FlexX, ADAM, Hammerhead, BUILDER, CONCEPTS, CONCERTS, DLD/MCSS, Genstar, Group-Build, Grow, HOOK, DOCK, Legend, LUDI, MCDNLG, SMOG and SPROUT [258]. Docking ligands on receptor can also be done by considering points of *complementary* between protein and ligand. FTDOCK, LIGIN and SANDOCK are the examples for programs that use complementary similarity. This similarity is surface and shape-chemical complementary for LIGIN and SANDOCK, respectively. On the other hand, FTDOCK uses electrostatic interactions [258].

As a second problem, scoring functions are considered. These functions should distinguish the experimental binding modes from all other modes with the minimum computational time and highest accuracy. Force fields like AMBER, OPLS, CHARM; empirical free energy scoring functions and knowledge based functions are some of the scoring techniques [258]. Force-field based scoring is applied that is considering all of the energy contributions (van der Waals, electrostatic) on the system. However, there are some limitations in force-field scores. Such type of scoring does not include solvation and entropic terms. Besides, it needs a cut-off distance definition for defining forces. Empirical scoring functions use experimental data of binding energies and/or conformations. The information is taken from X-ray structural data. Knowledge based scoring functions consider getting the experimental structures rather than binding energies. Additionally,

consensus scoring is also be used that combines different scores. However, it has limited usage on account of calculation errors that result from correlation of different scoring functions [257].

There are more than 60 docking programmes and 30 scoring functions but only some of them are used commonly such as AutoDock. Each docking programme is based on a method for exploring the conformational space of the ligand and/or protein and a scoring function [253]. Taylor et al. (2002) claimed that using one docking programme is not enough. Different methods should be used in order to investigate the binding mode of ligand. However, fast and accurate docking of ligands is still a concern [258]. Dastmalchi et al. (2005) stated that GOLD and AutoDock were the best among all possible docking programmes due to having best algorithms for ligand and protein flexible docking [259-260]. Tiwari et al. (2009) stated that AutoDock results were accordance with experimental data of *closo*- and *nido*-carboranyl antifolates [261].

### 2.3.1. AutoDock

AutoDock applies a grid-based technique for evaluating the binding modes of trial conformations. Lamarckian genetic algorithm is the primary method for docking in which a population of trial conformations is created and then biological evolution methods such as mutation, crossover are used in order to get candidates with lowest binding energy. Lamarckian allows passing the local minima information to the next populations. In addition to Lamarckian, simulated annealing search (for global minima) and a traditional genetic algorithm search methods also exist in AutoDock. Semiempirical free energy force field is used in AutoDock. Incorporation of intramolecular energies into the predicted free energy of binding is applied on force field. The AutoDock incorporates explicit conformational modeling of specified sidechains which is valuable in the analysis of covalent docking. However, adding flexibility can result in some problems not only in



AutoDock but also in the other programmes such as large conformational space and intensive evaluation of the receptor energy.

AutoDock Tools satisfy the need of graphical user interface. Moreover, it formats the input files with defining the charges, the rotatable bonds, active site and volume of space searched in docking simulation. AutoDock Tools is implemented in an object-oriented programming language Python. Python Molecular Viewer (PMV) is needed for this program. AutoDock Tools contains Viewer Framework (for visualization), Deja vu (three dimensional viewing visualization) and MolKit (for representation of molecules) [262].

Autodock 4.0 calculates free binding energy with respect to this formula:

$$\Delta G_{\text{binding}} = \Delta G_{\text{vdW}} + \Delta G_{\text{elec}} + \Delta G_{\text{hbond}} + \Delta G_{\text{desolv}} + \Delta G_{\text{tors}} \quad (3)$$

where

$\Delta G_{\text{vdW}}$ = 12-6 Lennard-Jones potential

$\Delta G_{\text{elec}}$ = Coulombic with Solmajer-dielectric, where Coulombic term ( $\epsilon$ ) is modified with a distance dependent approach,  $\epsilon(r)$ .  $r$  shows the distance [263].

$\Delta G_{\text{hbond}}$ = 12-10 Potential with Goodford Directionality, which considers length, orientation and chemical nature of hydrogen bond. Goodford et al. (1989) used experimental results to formulate hydrogen bond interactions [264].

$\Delta G_{\text{desolv}}$ =Stouten Pairwise Atomic Solvation Parameters which are determined by solvated system of BPTI in water comparing with its crystal structure. Individual atomic contributions were assumed to be used for solvation of a protein. Thus, volumes occupied by atoms in a solvated system were considered [265].

$\Delta G_{\text{tors}}$ =Number of rotatable bonds [262]

AutoDock converts the binding energy into  $K_i$ . As a result;

$$\Delta G = RT \ln K_i \quad (4)[266]$$

There are two steps for force field evaluation. Firstly, the intramolecular energy of change from unbound states of ligand and protein to bound one is considered. After that, intramolecular energy of combination of ligand and protein is evaluated [267].

Cincilla et al. (1999) used AutoDock 4.0 with run number of 150, maximum number of energy evaluation of 2500000 and maximum number of energy generation with 27000. They also modified AutoDock 3.0 and found that both AutoDock 4.0 and modified 3.0 showed a similar approach [268]. Ziemys et al. (2004) used AutoDock with 0.375 Å spacing for energy grid maps of atomic interaction and resulted in 47.25 Å cubic box dimension that covered all protein. LGA was used as docking algorithm with setting the number of individuals in the population, maximum number of energy evaluation, the maximum number as 50, 500000, 27000 [260]. Unal et al. (2010) studied peptide design on HIV-1 protease, scytalidocarboxyl peptidase B, SPG-40, PDF and ConA by AutoDock. In AutoDock, ADT was used for addition of Gasteiger charges and polar hydrogens. Docking parameters were kept small (population size=150, number of runs=10, maximum number of energy evaluation= 250000, number of generations=27000) for fast computation time and the rapid rank of aminoacids/dipeptides was desired since docking part was the prolonged part. They used the bound conformation of protein-peptide complex, the binding energy and  $K_i$  (inhibition constant) value [269-270]. Yoshimori et al. (2004) developed a novel aminoacid positional fitness (APF) score in order to do a screening of caspase inhibitory peptides by AutoDock 3.0 for caspase-3, -7, -8 and -9. Additionally, Ac-DNLD-CHO for caspase-3 inhibition was found by computational screening strategy. They used XXXD structure for this purpose by defining the optimized number of energy evaluation number with considering  $\Delta G_{\text{bind}}$  and  $K_i$  values. Firstly, they formed 8000 peptides library. After that, the calculation of free energies of peptides was done by randomly selection of peptides with AutoDock 3.0 and then APF scores were determined [271].

### 2.3.2. GOLD

GOLD is an automated ligand docking programme. It applies genetic algorithm for analyzing full range of ligand conformational flexibility and partial flexibility of macromolecule. GOLD also considers the ligand displacement with loosely bound water molecule on docking [252]. GOLD consist of three main parts which are:

*1. Scoring function for ranking the different binding modes:* Goldscore function is a molecular function with four terms. The protein-ligand hydrogen bond score, the protein-ligand van der Waals score, the intramolecular hydrogen bonds in the ligand and intramolecular strain in the ligand are the four terms and summation of these will give GOLD Fitness score.

*2. Mechanism for placing the ligand in the binding site:* GOLD adds fitting points to hydrogen bonding groups on protein and ligand. Acceptor points on the ligand on donor points in the protein are also mapped in addition to vice versa of this case. Hydrophobic fitting points are generated in the protein cavity.

*3. A search algorithm:* Genetic algorithm is used for determination of binding modes. Dihedrals of ligand rotatable bonds, ligand ring geometries, dihedrals of protein OH groups and  $\text{NH}_3^+$  groups and the mappings of fitting points (position of ligand in docking place) are modified variables with starting random choices. The key settings that affect docking in GOLD are the number of dockings and the number of GA operators. The required time and determination of global optimum are depended on the GA settings [272].

Hartshorn et al. (2007) stated that GOLD optimizes the protein polar hydrogens during docking process as a beneficial feature. However, sometimes the information related to the polar hydrogens can be known and it will be more useful to fix the positions of hydrogens according the information known. A property of fixing hydrogens is not included in GOLD. Moreover, the addition of X-Ray water molecules within 6 Å was done and the binding modes were estimated with a good estimation in GOLD. The authors also

advised the Goldscore function to be developed since there can be misled scores on account of hydrogen bonds involving charged groups for example [273].

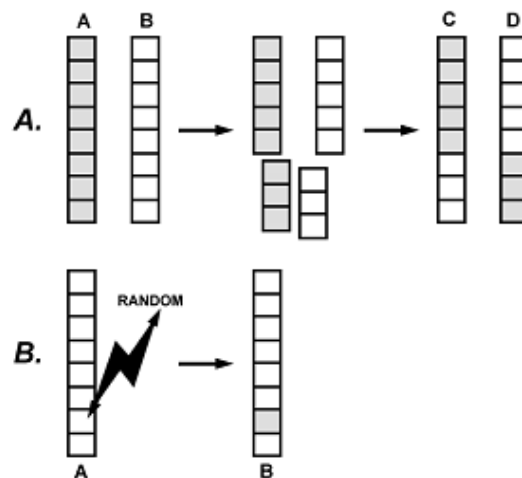
Löser et al. (2010) used GOLD for docking of tripeptidyl ligands on ICE with default docking settings. Mutations, migration and operator weights for crossover were used as 95, 10 and 95, respectively. 30 independent experiments that consist of 100 structures and up to 100000 operations were applied on ligands. Docking was stopped when  $\text{rmsd}=1.5 \text{ \AA}$  between the top three solutions and a maximum distance of  $5 \text{ \AA}$  was used between hydrogen donor and fitting points for allowing the poor nonbonded contacts. Besides, no constraints between ligand and protein were used. Manual identification of best-docked conformation of ligand was performed [274].

## 2.4 Genetic Algorithm

### 2.4.1. The theory

Many hard computational problems were started to be solved with genetic algorithm (GA) in the 1970s and became popular in the late 1980s. In the 1990s, the biological problems were also started to be analyzed such as protein sequence prediction [255]. Genetic algorithm is found on the basis of “survival of fittest” derived from the theory of evolution defined by Charles Darwin in *The Origin of Species*. This algorithm contains a set of potential solutions in which chromosome is defined as encoding of search space. *Chromosome* is a set of string of symbols for instance aminoacid letters in a peptide, natural numbers or nodes of graph. Strings are used as encodings in genetic algorithms and the length of these strings can be changed too [275]. Holland (1992) studied on living organisms which were excellent problem solvers due to obeying the rules of evolution. By the help of these rules, they can survive and pass through the next generation. However, the ones that are weak for handling the problems related to evolution are failed and cannot continue their existence. Holland (1992) worked on the mathematical analysis of adaptation

and survival. He found that recombination of genes and mating were the keys of evolution. Genetic algorithm was written with depending on these keys of evolution by Holland and it gave satisfying results for any type of problems on account of combining the partial solutions. Initially, each member of a population is scored and ranking is done with these scores of members. By a random choice, two members are selected and then they were combined with crossover from randomly selected crossover point. In chromosomes, two gametes meet to produce a zygote and crossover in genetic algorithm resembles the chromosomes. As a third step, mutation with low rate can hinder the uniform next populations that is incapable of further evolution. The formed offsprings (new members) replace with the low-scored members of population. With this approach, genetic algorithm can find the target regions of solution space [276].



**Figure 2.15 a) Crossover and b) Mutation [260].**

It cannot be claimed that GA will always find the optimal solution but it can find a possible combination of features that will survive under the specified conditions. The aim is to have a population of solutions whose size is maintained by placing the bad ones with the ones that have better survival. The survival level of each member of population is

understood in terms of fitness function. Possible solutions in population are demonstrated with strings and *replication*, *mutation* and *crossover* are the three genetic operators for diversifying the population. Fitness value is evaluated for each member after starting GA with  $N$  random solutions. In the *replication* stage, the members that will be replicated are selected according to their fitness values. In *crossover* stage, the  $N$  strings are matched randomly in pairs for having  $N/2$  pairs. A cut point was chosen randomly and the strings are swapped that position. As a result, two new strings are held. Each string can also be subjected to *mutation* at a specified rate. In Figure 2.15, the demonstration of crossover and mutation was done in which each column was a chromosome (an individual) or each square was the gene. There are many types of *mutation* and *crossover*. For example, single-point and n-point are the types of *crossover*. In the same manner, mutation can be done by considering one point or multiple points of chromosome [275]. These steps happen for one generation and many of such calculations continues until eventually some optimal point is held. In the end, population will contain a set of solutions with very good performance. The performance of algorithm is dependent on design and implementation details. However, a well-performed algorithm is based on trial and error procedure [255].

The selection of parents is a stochastic process and depends on the fitness score of the chromosomes in a population. Roulette wheel selection (proportionate selection) is a popular way for selection [275].

#### **2.4.2. Usage for ligand search**

There are  $20^3$  possibilities if we want to design a 3 sequenced peptide structure for caspase-1. It will not be possible to try each possibility with experimental or computational studies in other words the number of possibilities will be too high to try one by one. A logical and time-efficient method of determination of peptide sequence information is

required. For finding the correct structure with minimum time, genetic algorithm was used [269-270].

Unal et al. (2010) studied peptide design NF- $\kappa$ B protein with GA. They determined that the programme determined the hepta-peptide sequence without any prior knowledge and conservation of aminoacids at specific positions was seen. They claimed that the local minima problem was eliminated with high mutation rate [270, 277]. In a similar manner, Cincilla et al. (2009) used massive processing algorithm (MPA) that was based on genetic algorithm. In this algorithm, AutoDock 3.0 was used and estimated binding free energy was used as *fitness* value of each member of the *population*. [268]. Abe et al. (2007) produced the peptides by using a simple genetic algorithm. They used docking score as fitness function in addition to 1.0 crossover rate and single aminoacid mutation rate of 5 to 20 % [278]. Kernytsky et al. (2009) used GA for determination of some protein features which will be useful in enzymatic activity definition [279]. Unger et al. (2004) also stated that dihedral angles and Cartesian coordinates can also be used in GA for three dimensional structure predictions of proteins [255]. Moreover, Yang et al. (2003) studied the prediction of three dimensional peptide chain structures with an improved version of genetic algorithm. They called their method as intergeneration projection micro genetic algorithm (IP- $\mu$ GA) and their fitness function is RMSD [280].

Budin et al. (2001) also used genetic algorithm on caspase-1 with IIBC and removing its water and inhibitor molecules. However, in their study they directly took the most known four sequenced peptide of Ac-WEHD-CHO. After that they kept the D (ASP) residue of inhibitor constant and modified the other three residues with GA. In the end, they found Ac-WDTD-CHO as new peptide inhibitor [281].

## 2.5 Pharmacophore analysis

Pharmacophore modeling of ligand-protein complex is one of the crucial tools in developing novel drugs or analyzing the existing drugs. When pharmacophore analysis is

performed, hydrogen bonding, charge transfer, electrostatic as well as hydrophobic interactions in ligand-protein complex can be visualized. These basic interactions are the keys of determining the binding mode of ligands.

Barreca et al. (2007) used pharmacophore modeling in order to investigate the candidates for chemical scaffolds by LigandScout. Pharmacophore model was used in virtual screening for HIV-1 non-nucleoside reverse transcriptase inhibitors [282]. Thus, LigandScout was also applied on HIV protease, HIV reverse transcriptase, influenza virus neuraminidase, human rhinovirus coat protein and hepatitis C virus RNA polymerase with considering the PDB structures of these receptor-inhibitor complexes by Steindl et al. (2006). They claimed that fast virtual screening of compounds against these systems yielded successful activity profiling in most cases [283]. Moreover, using LigandScout was found to be feasible and intuitiveness by Wolber et al. (2006). They used three dimensional pharmacophore structure for analyzing the macromolecule-ligand interaction with aligning small ligands onto pharmacophore [284].

LigandScout uses interaction between ligand and protein for visualization of three dimensional pharmacophores. When the complex of ligand and macromolecule is loaded, LigandScout automatically shows pharmacophore model. In this model, the vectors show hydrogen acceptors and donors while bunch of vectors are used for positive and negative ionizable regions. Moreover, spheres demonstrate the lipophilic areas and excluded volume spheres. Excluded volume spheres are regions where any potential ligand cannot access. LigandScout uses steric restrictions for formation of excluded volumes automatically [282]. Hydrogen bond interactions are considered when covalently bound hydrogen with a positive partial charge interacts with another atom that has negative partial charge. The distance of hydrogen bonds is 3.8 Å for considering all interactions. However, some geometric constraints were implemented in order to increase selectivity such as using  $sp^2$  and  $sp^3$  donor atoms. Besides, the model assumes an ideal hydrogen bond angle of 180



degrees and if the difference between both directions exceeds 34 degrees, the hydrogen bond will be accepted as broken [285]. Hydrogen bond donors are determined if functional groups of nonacidic hydroxyls (all OHs except sulfonic, sulfinic, carboxylic, phosphonic or phosphinic acids), thiols, acetylenic hydrogens, and NHs (except tetrazoles and trifluoromethyl sulfonamide hydrogens) are formed by hydrogen bond donors. The hydrophobic areas are found by scoring the atoms according to hydrophobicity and then tolerance radius of 1.5 Å is added. Charge transfer interactions are also visualized with considering positive (PI) or negative (NI) ionizable areas. Basic amines, basic secondary amidines, basic primary amidines, basic guanidines, and positive charges not adjacent to a negative charge are examples of PI whereas trifluoromethyl sulfonamide hydrogens, sulfonic acids, phosphonic acids, sulfinic, carboxylic or phosphinic acids, tetrazoles, negative charges which are not neutralized by an adjacency to a positive charge are the examples of NI [285].

## 2.6 GNM (Gaussian Network Model)

GNM has been used for many purposes in protein world such as dynamics of proteins, features of protein fluctuations, conserved structures in proteins by using the same logic with a single parameter for harmonic potential as Tirion did (1996). It can be even applied on supramolecular structures [286-287]. It is a coarse-grained model and contact information of the structure is used. Mode decomposition is applied for identifying the slowest and fastest modes in a given protein system [288]. It is firstly applied on proteins by Bahar et al. (1997) for determination of thermal fluctuations. Crystal structures of proteins were used considering  $C^\alpha$  for determination of temperature factors in GNM. 3D structures of proteins were adapted into an elastic network [289]. In addition, Haliloğlu et al. (1997) studied GNM for folded proteins [290]. HIV-1 protease domain motions as well as folding cores were also determined with GNM by Bahar et al. (1998). Fast and slow modes were found to be associated with stability and function, respectively [291].

Distance vector  $R_{ij}$  demonstrates the distance between  $i^{\text{th}}$  and  $j^{\text{th}}$  residues as seen from Figure 2.16. Fluctuations in distance vector can be calculated by assuming *isotropic* and *Gaussian* as follows [286]:

$$R_j - R_j^0 - R_i - R_i^0 = \Delta R_j - \Delta R_i = \Delta R_{ij} = R_{ij} - R_{ij}^0 \quad (5)$$

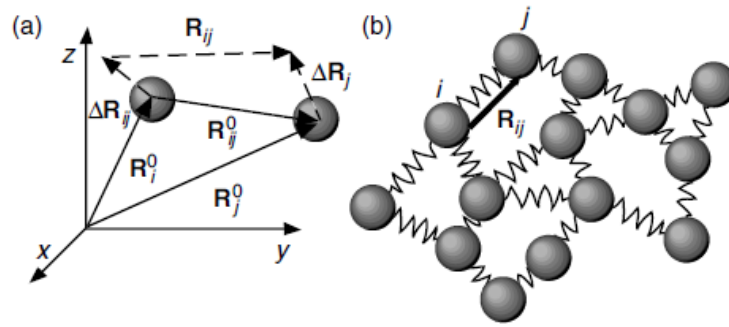
As a result, potential of network of N nodes (residues) will be evaluated as [286]:

$$\begin{aligned} V_{GNM} &= \frac{\gamma}{2} \left[ \sum_{i,j} \Gamma_{ij} \left[ \Delta X_i - \Delta X_j \right]^2 + \left[ \Delta Y_i - \Delta Y_j \right]^2 + \left[ \Delta Z_i - \Delta Z_j \right]^2 \right] \\ &= \frac{\gamma}{2} \left[ \sum_{i,j} \Gamma_{ij} \left[ \Delta R_i - \Delta R_j \right]^2 \right] = \frac{\gamma}{2} \left[ \sum_{i,j} \Gamma_{ij} \left[ \Delta R_{ij} \right]^2 \right] \end{aligned} \quad (6)$$

where  $\Delta X_i$ ,  $\Delta Y_i$  and  $\Delta Z_i$  are components of  $\Delta R_i$ . and  $\gamma$  is force constant (Hookean pairwise potential used by Tirion (1996)) [287, 290, 292], which will be constant for all springs.  $\Gamma_{ij}$  is the  $ij^{\text{th}}$  element of Kirchhoff (connectivity) matrix, which is an N x N symmetric matrix for a protein of N residues [288].  $\Gamma_{ij}$  is defined as:

$$\Gamma_{ij} = \begin{cases} -1 & \text{if } i \neq j \text{ and } R_{ij} \leq r_c \\ 0 & \text{if } i \neq j \text{ and } R_{ij} > r_c \\ -\sum_{j, j \neq i} \Gamma_{ij} & \text{if } i = j \end{cases} \quad (7)$$

where  $r_c$  is cut off distance [286].



**Figure 2.16** Definition of fluctuations [286].

GNM normal modes can be found by eigenvalue decomposition of  $\Gamma$ .

$$\Gamma = U\Lambda U^T \quad (8)$$

in which  $U$  shows eigenvectors ( $u_i$ ) of  $\Gamma$  and  $\Lambda$  represents eigenvalues ( $\lambda_i$ ) of  $\Gamma$ . Transpose of  $U$  equals to its inverse ( $U^T=U^{-1}$ ) and the first eigenvalue of the matrix is zero due to translational invariance of the system. Eigenvalues are usually found in an ascending order ( $0 < \lambda_2 < \dots < \lambda_n$ ) for  $1 \leq i \leq n$ . In addition, eigenvectors and eigenvalues demonstrate shapes and frequencies of modes [286, 293].

Inverse of  $\Gamma$  can be found with ignoring the zero eigenvalue by [290]:

$$\Gamma^{-1} = \sum_{k=2}^n \lambda_k^{-1} u_k u_k^T \quad (9)$$

Demirel et al. (1998) applied GNM for identification of hot residues in folded structures of chymotrypsin inhibitor 2, cytochrome c, and related C2 proteins [293]. In another study by Demirel et al. (2005), connectivity (summation of diagonal elements of  $\Gamma$ ) is found to be proportionally related with residue number. Thus, proteins with same number of residues have same connectivity and total eigenvalue [294]. Furthermore, Bahar et al. (1999) plotted correlation map of HIV-1 RT by using the following definition [288]:

$$\Delta R_i \cdot \Delta R_j = \left(\frac{3kT}{\gamma}\right) \Gamma_{ij} = \left(\frac{3kT}{\gamma}\right) \sum_k (\lambda_k^{-1} u_k u_k^T)_{ij} \quad (10)$$

is the equilibrium correlation between the fluctuations of residues  $i^{\text{th}}$  and  $j^{\text{th}}$ .  $K$  is Boltzmann constant and  $T$  is temperature [288]. The mean-square distance fluctuations between  $i^{\text{th}}$  and  $j^{\text{th}}$  residues, which also demonstrate mobilities of individual residues, can be evaluated by [292, 295]:

$$\Delta R_{ij}^2 = (\Delta R_i - \Delta R_j)^2 = \Delta R_i^2 + \Delta R_j^2 - 2\Delta R_i \cdot \Delta R_j \quad (11)$$

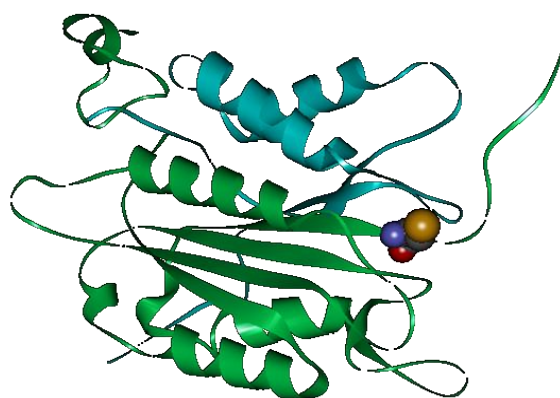
After the start of usage in proteins, GNM studies has been continued and improved for proteins. Haliloglu et al. (2009) studied native proteins and applied GNM model for describing correlations between energy and residue fluctuations by considering HIV protease [296]. In another study by Tuzmen et al. (2011), a more detailed study related to prediction of binding sites and interaction paths in a protein was done. Several proteins were considered with not only unbound (no ligand) but also bound (ligand) by GNM. Comparisons with unbound and bound protein structures demonstrated the success of this method [297].

## Chapter 3

### 3. METHODS

#### 3.1 Caspase-1

Caspase-1 with PDB ID of 2HBQ is used as target protein in this study. 2HBQ belongs to the crystal structure of wild type human caspase-1. Its structure is experimentally determined from X-Ray diffraction with a resolution of 1.8 Å which is demonstrated in Figure 3.1 [8, 17].



**Figure 3.1** Caspase-1 in flat ribbon representation. Green and blue chains are A and B chains, respectively. Orange atom is the sulphur atom of Cys285 (active residue) [17].

## 3.2 Computational methods for ICE drug design

### 3.2.1. Docking

Löser et al. (2010) studied molecular docking on 2HBQ, which is the same caspase-1 structure in this study, without minimization of caspase-1. Moreover, they found potent and selective inhibitors for caspase-1 which are tripeptidyl benzyl- or cyclohexylamines with computational docking methods. The verification with inhibition experiments for found ligands is also performed [274].

#### 3.2.1.1. AutoDock

A python script is written for AutoDock runs which initially convert protein and ligand .pdb structures into .pdbqt file format by using prepare\_receptor4.py and prepare\_ligands4.py python scripts in MGL Tools 1.5.6. By using .pdbqt files, prepare\_gpf4.py and prepare\_dpf4.py python scripts in MGL Tools make ready grid parameter file (.gpf) and docking parameter file (.dpf) for ligand. Genetic algorithm is selected in which the population size, run, maximum number of energy evaluations, number of generations are 150, 50, 2500000 and 27000, respectively for AutoDock runs (AutoDock 4.2). Spacing is taken as 0.4 Å. Crossover, mutation and elitism are 0.80, 0.02 and 1, respectively. AutoDock runs are performed on a Linux system server that has 2 CPU Intel Pentium 2.4 Ghz on 32 nodes with 1 GB RAM memory. Eight tripeptides are analyzed in 45 minutes under these parameters at the same time. Besides, all peptide torsion angles are made flexible with AUTOTORS utility of AutoDock. Adding gasteiger charges and polar hydrogen, and grid map determination are done by MGL Tools 1.5.6. Lamarckian Genetic Algorithm is used as docking search parameter [262].

In python script, autogrid4 and autodock4 are used for getting .glg and .dpf files from .gpf and .dpf files, respectively. Dockings are done on the sulphur atom of Cys285

because it is determined as the most appropriate place for binding on caspase-1 (the detailed information about determination of Cys285 is given in Results part). As a result of AutoDock runs, free binding energy of docked ligand onto protein is found in .dlg file of ligand that is the summation of intermolecular energy (van Der Waals energy, Hydrogen bonds, desolvation energy, electrostatic energy), final total internal energy, torsional free energy of peptide, unbound system energy and estimated inhibition constant [262]. Calculated free binding energies of one ligand are used in Boltzmann Distribution with its degeneracies (degeneracy is the frequency of a free binding energy).

$$\text{Free Energy of Binding} = \frac{\sum g_i E_i e^{\frac{-E_i}{N_A k_B T}}}{\sum g_i e^{\frac{-E_i}{N_A k_B T}}} \quad (12) [298]$$

where  $E_i$  (kcal/mol) is free energy binding of run,  $g_i$  is degeneracy of run,  $k_B$  (kcal/K) is Boltzmann constant,  $N_A$  (1/mol) is the Avogadro's constant and  $T$  (K) is temperature. Final result from Boltzmann calculation gives the free binding energy of inhibitor designed since taking only the lowest score of free binding energy is not appropriate. Consideration of all possible conformations should be done for reaching the correct result.

### 3.2.1.2. GOLD

GOLD 4.1.1 is also used in docking experiments. GOLD runs are performed on 2 CPU and 1.96 GB of RAM with Windows XP Version 2002. Chemscore kinase is used as template and not only Goldscore but also Chemscore  $\Delta G$  values are considered.

Initially in GOLD runs, .pdb files of ligands are converted into .mol2 files by Accelrys Discovery Studio 2.5 and Racoon [299-300]. On the other hand, protein is given with its .pdb file. Dockings are done on the sulphur atom of Cys285. Aminoacids in docking surface are done flexible which are Arg179, His237, Ile239, Gln283 as well as

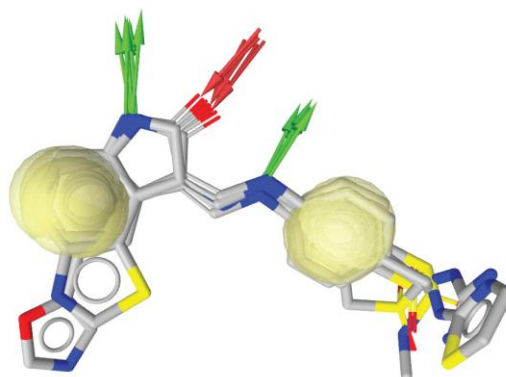
Cys285 for getting more consistent results with reality. Early termination is not allowed due to considering all possible dockings and atoms are selected within 10Å. Hydrogen addition is also done. Detect cavity choice is selected in order to avoid ligands to bind unrelated surfaces. Genetic algorithm search option is selected with population size of 100, number of operations of 100 000, crossover frequency of 95, mutation frequency of 95 and migration frequency of 10 [301]. A tripeptide is docked approximately in 5-10 minutes under these conditions in GOLD. At the end of docking run, all results are given in a gold.conf file. When this file is loaded into GOLD, all results are seen.

### **3.2.2. Computational ligand analysis tools**

#### **3.2.2.1. LigandScout**

LigandScout 2.02 is used for pharmacophore analysis. When .pdb file of a ligand-protein complex is given to LigandScout; hydrophobic interactions (yellow spheres), a negative ionizable area (red bunch), a hydrogen bond donor (green arrow) and two hydrogen bond donors (red arrow) are seen clearly (Figure 3.2). It also gives TPSA (topological polar surface area) and cLogP (an estimate of a compound's overall lipophilicity) for ligand in the complex. .pdb file is loaded into LigandScout and necessary observations are done such as TPSA, cLogP, molecular weight, hydrogen acceptor and hydrogen donor [285].





**Figure 3.2** A sample demonstration of a ligand in LigandScout.[285]

#### 3.2.2.2. iLib Diverse

Java based software iLib Diverse 1.02 is a fast tool for formation of diversified or focused libraries [302]. Lipinski Rule of 5 is used as a filter [217] and many chemical groups are considered, which are acids, alcohols, aliphatics, alkenes and alkynes, amides and imides, amines, benzenes, carbocycles, esters, ethers, Fx groups, halogens, heterocycles, ketones and aldehydes, pharm-groups, phosphor and sulphur.

iLib Diverse is used in peptide based drug search for ICE. First, number of fragments is defined and chemical groups are transported in each fragment. Lipinski Rule of 5 is used as filter and formation of library is performed. Except formation of D-aminoacid part, default values are used. For getting .pdbqt file format for Autodock Vina and .mol2 format for GOLD, a combination of Accelrys Discovery Studio, Marvin Beans [303] and Racoon [299] are used.

#### 3.2.2.3. HyperChem

HyperChem Release 7 is a molecular modeling and simulation programme. It is used for analyzing minimized energy of a ligand on caspase-1 after docking on GOLD. The best docking in GOLD is selected for getting .pdb file of protein-ligand complex in order to

perform HyperChem calculations. After loading complex .pdb structure, molecular mechanics optimization of complex is done and then ligand is deleted. The default values are used such as 0.1 kcal/ (Å x mol) RMS and Polak-Ribiere algorithm (a conjugate gradient method). Following step is to do molecular mechanics optimization of caspase-1. Difference between complex and caspase-1 is evaluated. Minimized energy for a ligand which is close to the optimized energy (local minimum) is determined. [304-305].

#### **3.2.2.4. Accelrys Discovery Studio Visualizer, LigPlot and CS ChemBio3D Pro**

Accelrys Discovery Studio Visualizer 2.5 is used for visualization of proteins as well as ligands. It is also used for analysis of hydrogen bonds (distance is 2.5 Å) and pi interactions between ligand and caspase-1 with default parameters. Accelrys Discovery Studio Visualizer is also used for minimization of procaspase-1. This is performed by ‘Clean Geometry’ property of programme [300]. Thus, ChemBio3D Pro is used for sketching and minimizing ligands with its ‘minimize energy’ feature in 3D [306]. Bonds between ligands and caspase-1 are also studied by LigPlot v.4.5.3. Structure of ligand-caspase-1 complex, which is taken from the best conformation of AutoDock run, is given to LigPlot. Programme gives information related to hydrogen bonds and hydrophobic interactions. Default parameters are used [307].

#### **3.2.3. Genetic Algorithm**

Genetic algorithm is used in order to find a candidate peptide sequence for a tripeptide structure for ICE inhibition. Algorithm code is written in a Python code which is given to a Linux system server that had 2 CPU Intel Pentium 2.4 Ghz on 32 nodes with 1 GB RAM memory. Code is mainly composed of peptide formation, docking, reading the .dlg files, ranking and new population formation whose flowchart is given in Figure 3.3.

*Peptide formation* is done initially with a random choice by a Python script in order to have 64 populated tripeptides sequences (*initial population*= $S_1$ ) then tripeptide structures are given to VMD programme. In VMD, *mutate* command is used for converting a temporary tripeptide .pdb file into the desired tripeptide [277]. As a consequence, three dimensional structural information of tripeptides (.pdb file) is formed. AutoDock is used as docking programme in genetic algorithm code. Since 250,000 is used as maximum number of energy evaluations, eight tripeptides is docked approximately in 26 minutes at the same time. As a result, 64 populated list of tripeptides is finished about 4 hours. Next step is to *read* these .dlg files with the same Python code. This forms *fitness value* of genetic algorithm.

The best five tripeptides are kept constant without no change (*elitism*) after *ranking* according to *fitness values* and the rest of 59 structures are sent to *single mutation*. Using *crossover* in addition to *mutation* would be meaningless for a tripeptide. Besides, using *single crossover* would be same as using *single mutation* since structures are short (3 peptides). Consequently, only *single mutation* (shown by M) is considered with random selection of mutated peptide in the given structure.  $M[P_i, X]$  gives the mutated residue in which  $P_i$  is parent  $i$  and  $X$  is the number of the residue to be mutated in  $P_i$ . In other words, an *offspring* ( $O_j$  where  $j$  is the number of new peptide in *new population*,  $S_{m+1}$  where  $m$  denotes the parent population number) is formed by *parent* ( $P_i$ ). After *mutation*, *new population* (*generation*) is formed and then sent to the beginning of for loop. Until reaching a convergent behavior, code is kept on [270].

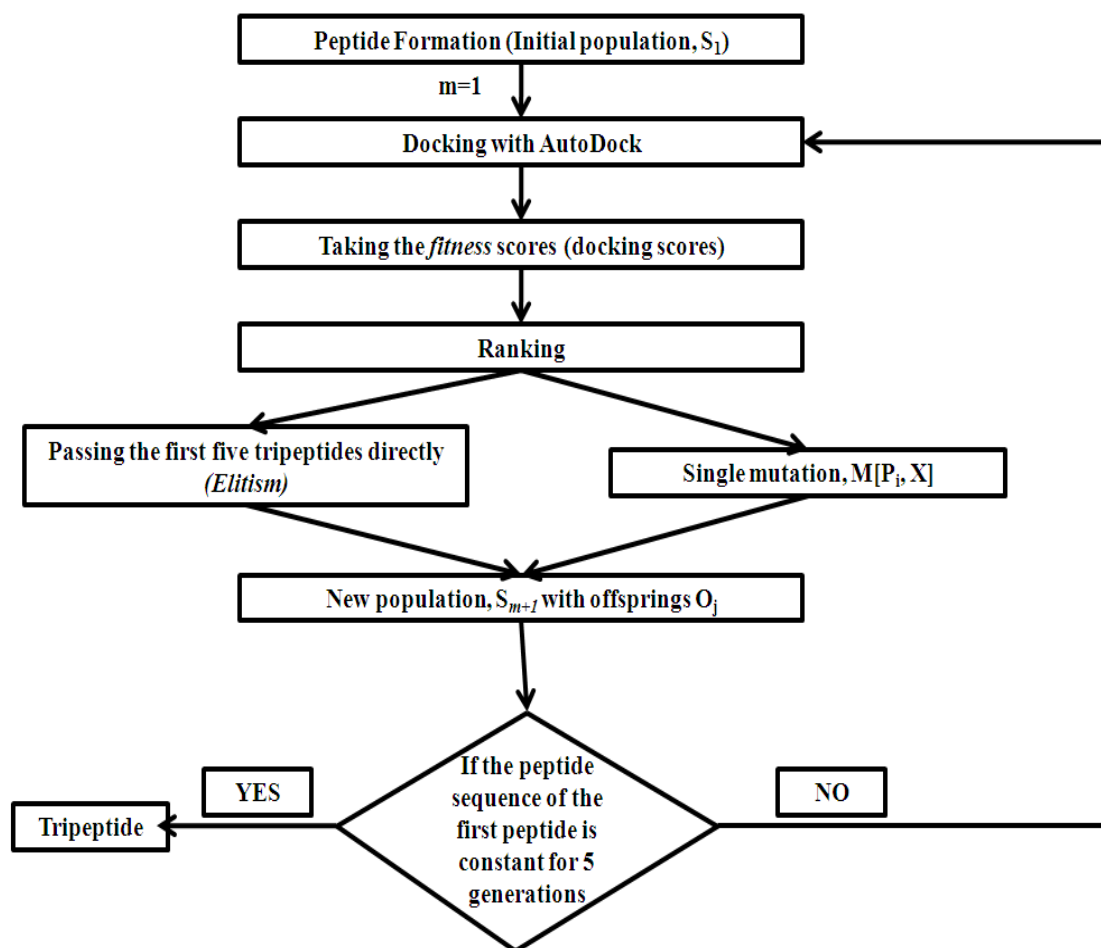


Figure 3.3 Genetic algorithm flowchart.

#### 3.2.4. Complete Enumeration

Complete enumeration of tripeptides is done by a combination of Autodock Vina, GOLD and Autodock. 8,000 tripeptides are all combinations of 20 aminoacids. Tripeptides are formed by HyperChem [304-305]. The first 500 tripeptides are determined by Autodock Vina and then given to GOLD twice. GOLD runs are lasted for one week and the ones which had negative Goldscore and/or positive Chemscore  $\Delta G$  are eliminated. Total of 156 tripeptides is found and a second run is done on GOLD. 87 structures are finally

determined by looking at Goldscore and Chemscore values. Autodock is used for final elimination of 87 structures.

### 3.3 Conformational Factors in the Design of Peptide Drugs

For a given tripeptide, the conformation of each residue is known and fixed in the bound state. In the unbound state, in solution, the tripeptide spends most of its time in its most probable conformations. In order to bind, the peptide has to go from the most probable state to the bound state, and in doing so has to pass over a free energy barrier. There may be several different highly probable states in solution, but we consider only the most probable state as the starting point in this study. We next proceed to calculate the minimum energy state of the tripeptide in the free state, and then calculate the optimum path over the energy landscape through which each of its residues go from the minimum energy conformation of the free state to the conformation in the bound state. We use the term optimum path to mean the path that exhibits the lowest energy barrier. The energies mentioned here should strictly be free energies. However, since free energy calculations require excessive times, we consider energies only. Decision based on energy calculations only safely eliminates the highly improbable paths, but the optimum path obtained in this manner is not necessarily the optimum free energy path. Another assumption of the model is that each residue of the tripeptide changes conformation irrespective of the conformation of its neighboring residues. This assumption is based on the Flory independent residue hypothesis [308], which is shown not to be strictly valid, but is still a good approximation [309].

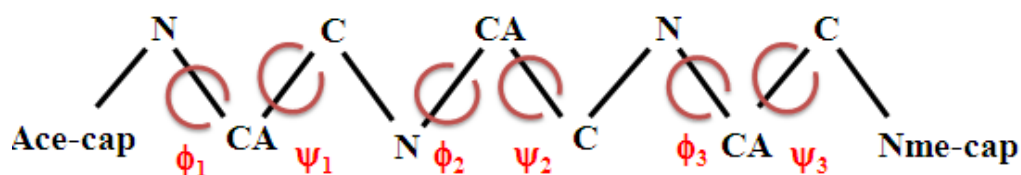
#### 3.3.1. Determination of States

All combinations of  $\alpha$  and  $\beta$  conformations are used for forming tripeptides by HyperChem Release 7 [304-305]. In other words, a tripeptide is formed in 8 ( $2^3$  where 2 is

$\alpha$  and  $\beta$  states, 3 is the number of peptides) conformations. In all tripeptides formed, Ace- and -Nme are used as caps. Besides, each of 8 conformations of a tripeptide is solvated in water box and minimized with geometry optimization tool of HyperChem. Polak-Ribiere algorithm is used with default values except maximum cycles. Number of maximum cycles is 32767 and Bio-Charmm force field is used. After minimization in water which lasts 15-50 minutes (duration depends on conformation of the tripeptide), single-point calculation which lasts about 10-20 seconds is performed for only tripeptide in water (water molecules are not considered). The tripeptide conformation with highest single-point energy is further minimized with geometry optimization in a void space with OPLS force field (lasts 5-15 minutes). Components of OPLS and Bio-Charmm force fields are bond, angle, torsion, non-bonded, electrostatic and hydrogen-bonded interactions. Final structure is used for the determination of states of three aminoacids. These states are the states taken from HyperChem calculations.

Each tripeptide is docked 100 times with AutoDock. At the end of 100 dockings, the best conformations are written in a pdb file by MGL Tools [262]. All formed .pdb files are analyzed according to their aminoacid states. After determining states from 100 AutoDock runs, the most similar states with HyperChem are taken as AutoDock states.

While determining states, there are a total of 6 angles. 3 of them  $\psi$  while other 3 are  $\phi$ . Each couple of  $\phi$  and  $\psi$  determines a state in Ramachandran plot. The definitions of torsion angles are illustrated in Figure 3.4.



**Figure 3.4** The definitions of torsion angles.

We use information from the coil library of the Protein Data Bank to construct the states of the  $\varphi$  and  $\psi$  angle pairs of a residue [310-311]. A given state is obtained when samples from the coil library cluster around a region. In the earlier work of Karplus et al. [312], eleven states were identified [269]. In more recent work, Unal et al identified a total of twenty one states, which are depicted in Table 3.1 and Figure 3.5 [313].

**Table 3.1** Definitions of 21 states [313].

State	Notation	Definition	Angles
1	$\varepsilon'$	Mirror image of the extended region $\varepsilon$ .	$(-180 \leq \phi \leq -60$ and $-180 \leq \psi \leq -150)$ or $(-120 \leq \phi \leq -90$ and $-150 \leq \psi \leq -120)$
2	$\varepsilon$	The extended regions, $\varphi > 0$ , $\psi \sim +180^\circ$ .	$(30 \leq \phi \leq 180$ and $-180 \leq \psi \leq -150)$ or $(30 \leq \phi \leq 150$ and $-150 \leq \psi \leq -120)$ or $(60 \leq \phi \leq 180$ and $150 \leq \psi \leq 180)$ or $(90 \leq \phi \leq 150$ and $120 \leq \psi \leq 150)$
3	$\alpha_R$	Right-handed alpha helix.	$(-180 \leq \phi \leq -30$ and $-90 \leq \psi \leq -30)$
4	$\gamma$	Tight turn region.	$(60 \leq \phi \leq 120$ and $-90 \leq \psi \leq -30)$
5	$\delta_R$	The right handed bridge region between two $\beta$ strands.	$(-120 \leq \phi \leq -30$ and $-30 \leq \psi \leq 0)$ or $(-150 \leq \phi \leq -60$ and $0 \leq \psi \leq 30)$
6	$\delta_L$	Mirror image of the $\delta_R$ region.	$(60 \leq \phi \leq 150$ and $-30 \leq \psi \leq 30)$
7	$\zeta$	Region observed mostly in residues preceding PRO.	$(-180 \leq \phi \leq -120$ and $30 \leq \psi \leq 90)$
8	$\gamma'$	Inverse tight turn region.	$(-120 \leq \phi \leq -60$ and $30 \leq \psi \leq 90)$
9	$\alpha_L$	Mirror image of $\alpha_R$ .	$(30 \leq \phi \leq 90$ and $30 \leq \psi \leq 90)$ or $(30 \leq \phi \leq 60$ and $90 \leq \psi \leq 120)$
10	$\beta_S$	Extended beta sheet forming region.	$(-90 \leq \phi \leq -30$ and $120 \leq \psi \leq 150)$ or $(-120 \leq \phi \leq -60$ and $150 \leq \psi \leq 180)$
11	$\beta_P$	Region with extended polyproline-like helices.	$(-150 \leq \phi \leq -30$ and $90 \leq \psi \leq 120)$ or $(-180 \leq \phi \leq -90$ and $120 \leq \psi \leq 150)$ or $(-180 \leq \phi \leq -120$ and $150 \leq \psi \leq 180)$
12	$\delta_R'$		$(-180 \leq \phi \leq -120$ and $-30 \leq \psi \leq 0)$ or $(-180 \leq \phi \leq -150$ and $0 \leq \psi \leq 30)$
13	$\varepsilon''$		$(-180 \leq \phi \leq -120$ and $-150 \leq \psi \leq -90)$ or $(-120 \leq \phi \leq -90$ and $-120 \leq \psi \leq -90)$
14	$\varepsilon'''$		$(-90 \leq \phi \leq -30$ and $-150 \leq \psi \leq -90)$ or $(-60 \leq \phi \leq -30$ and $-180 \leq \psi \leq -150)$
15	$\gamma''$		$(-60 \leq \phi \leq -30$ and $0 \leq \psi \leq 90)$
16			$(-30 \leq \phi \leq 30$ and $-180 \leq \psi \leq 180)$
17	$\varepsilon\alpha_L$		$(30 \leq \phi \leq 60$ and $120 \leq \psi \leq 180)$ or $(60 \leq \phi \leq 90$ and $90 \leq \psi \leq 150)$
18	$\alpha_L'$		$(90 \leq \phi \leq 150$ and $30 \leq \psi \leq 120)$
19			$(180 \leq \phi \leq 150$ and $-150 \leq \psi \leq 150)$ or $(120 \leq \phi \leq 150$ and $-90 \leq \psi \leq -30)$
20	$\delta_L'$		$(30 \leq \phi \leq 60$ and $-90 \leq \psi \leq 30)$
21	$\varepsilon\gamma$		$(20 \leq \phi \leq 60$ and $-120 \leq \psi \leq -90)$



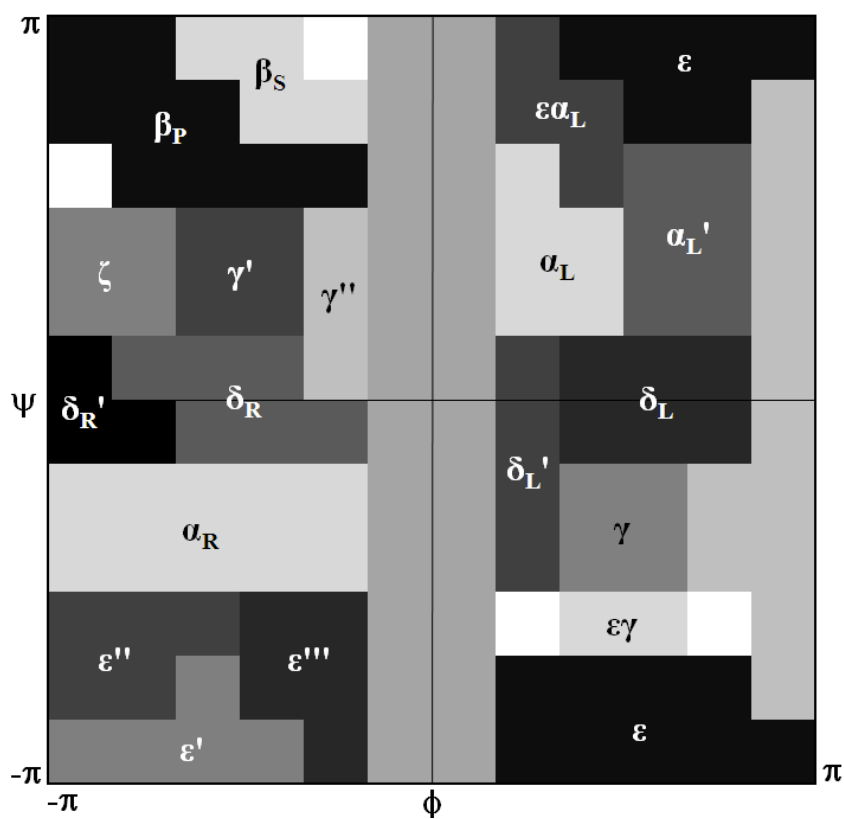


Figure 3.5 The 21 states in Ramachandran map [313].

### 3.3.2. Equilibrium and Kinetic Factors

Analysis of energy changes that a tripeptide meets is done by considering probabilities of bound and minimized states. The probability of a conformation in a given state is calculated according to the expression

$$P_i = \frac{e^{-\beta E_i}}{Z} \quad (13) \quad [314]$$

where  $E_i$  is the energy of the  $i$ 'th state.

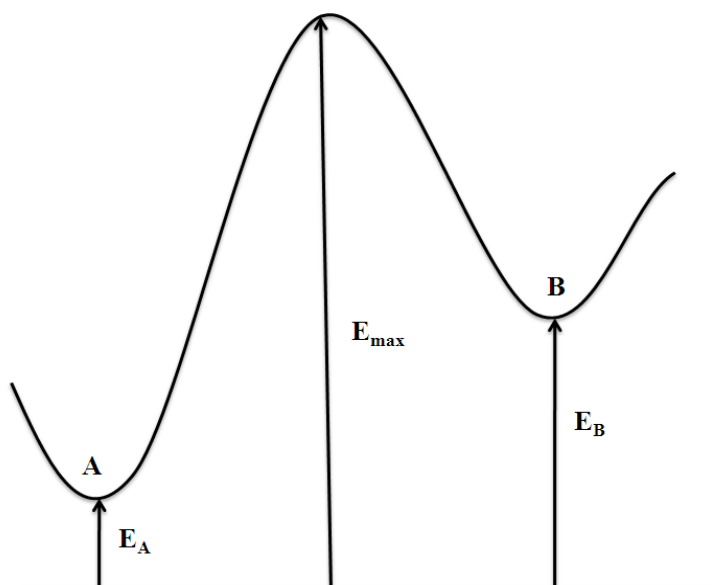
$$Z = \sum_i e^{-\beta E_i} \quad (14) [314]$$

where  $\beta=1/(kT)$  in which  $k$  is the Boltzmann constant and  $T$  is temperature.  $Z$  is the partition function.  $E_A$  and  $E_B$  are the energies of A and B states in Figure 3.6 which are calculated from HyperChem. Rate of reaction between A and B states are given as  $r_{AB}$  and rate of reverse of this reaction is  $r_{BA}$ .

$$r_{AB} = f e^{-\beta E_{act}} = f e^{-\beta(E_{max}-E_A)} \quad (15) [315]$$

$$r_{BA} = f e^{-\beta E_{act}} = f e^{-\beta(E_{max}-E_B)} \quad (16) [315]$$

where  $f$  is a function that is the same in both states.



**Figure 3.6** A schematic representation of energies.

When forward and inverse reactions between A and B states are divided to each other, the resulting ratio equals to ratio between  $P_B$  and  $P_A$ .

$$\frac{r_{AB}}{r_{BA}} = \frac{fe^{-\beta(E_{\max}-E_A)}}{fe^{-\beta(E_{\max}-E_B)}} = \frac{e^{-\beta E_B}}{e^{-\beta E_A}} = \frac{P_B}{P_A} \quad (17)$$

$P_B/P_A$  is used for finding whether a tripeptide goes from its minimized to bound state. For instance if the ratio of  $P_B/P_A$  is small, it means tripeptide has difficulty in passing to its bound state because energy difference between minimized and bound state is large. Energy met by tripeptide is too high to overcome. That is why tripeptide is not a good choice for inhibition of caspase-1. However, tripeptide is found to be a good inhibitor if the  $P_B/P_A$  ratio is high. Since energy difference between minimized and bound conformations is low enough to overcome. All in all, inhibitors for caspase-1 are chosen from tripeptides with high  $P_B/P_A$  ratio.

### 3.3.3. Determination of Path Followed

When a residue in a tripeptide has different state in HyperChem and AutoDock, it is analyzed according to its energy change while passing from HyperChem to AutoDock state by Viterbi decoding. HyperChem (minimized conformation) and AutoDock (bound conformation) states are considered as start and end points, respectively..

A state is affected from its previous neighbor state in 1<sup>st</sup> order Markov process [18]. System is considered as 1<sup>st</sup> order Markov process with unobserved states of  $\varphi$ . The interval of  $\psi$  (between start and end points) is used with many alternatives of  $\varphi$ . Energies of points are evaluated by single-point calculation in HyperChem in OPLS force field [316]. A state is affected from its previous neighbor state [317]. Viterbi decoding is used for determination of the most probable path of unobserved states. In other words, each  $\psi$  has one observed  $\varphi$  which forms path followed by minimized tripeptide conformation to reach its bound state. Notation given in Ewens et al. (2001) and Unal et al. (2010) is used for following definitions [269, 318]:

n: Number of  $\psi$  angles of path that tripeptide follows.

t: Number of grid,  $1 \leq t \leq n$ .

I: Number of  $\varphi$  angle used.

m: Index identifying the  $\varphi$  angle,  $1 \leq m \leq I$ .

S=  $\{S_1, S_2, \dots, S_n\}$ : The torsion angle elements for  $\psi$  angle.

$q_t = S_i$  demonstrates the state of  $t^{\text{th}}$  grid is  $\psi$  angle  $S_i$ .

A= $\{A_1, A_2, \dots, A_I\}$ : The torsion angle elements for  $\varphi$  angle.

$a_t = A_i$  demonstrates the state of  $t^{\text{th}}$  grid is  $\varphi$  angle  $A_i$ .

P =  $(P_{ij}) = \Pr(q_{t+1}=S_j \mid q_t=S_i)$  =The transition probability where energies are taken from HyperChem (kcal/mol). It is calculated as follows:

$$P_{ij} = \frac{\exp\left(-\frac{\Delta E(t+1, j; t, i)}{kT}\right)}{\sum_j^A \exp\left(-\frac{\Delta E(t+1, j; t, i)}{kT}\right)} \quad (18) [314, 318]$$

b =  $(b_{ij}) = \Pr(a_{t+1}=A_j \mid a_t=A_i) = 1$ : The emission probability, which is accepted as 1 since unobserved states of  $\varphi$  are assumed to have equal probability.

$\pi = (\pi_i) = \Pr(q_1=S_i)$ : an initial probability vector.

$\delta_t(j) = \max \Pr(q_1, q_2, \dots, q_t = S_i \text{ and } a_1, a_2, \dots, a_t)$ : The maximum probability of all paths ending at state  $S_i$  at grid  $t$ .

Our aim is to find a path that tripeptide follows while passing from minimized to bound state using minimum energies.  $\operatorname{argmax} \Pr(q_1, q_2, \dots, q_n \mid a_1, a_2, \dots, a_n)$  is used for achieving this. Determination of  $\operatorname{argmax} \Pr(q_1, q_2, \dots, q_n \mid a_1, a_2, \dots, a_n)$  depends on two steps. Initially, forward tracking is used for finding  $\max \Pr(q_1, q_2, \dots, q_n \mid a_1, a_2, \dots, a_n)$ . As a second part of this algorithm, backtracking is applied that finds a  $O_1, O_2, \dots, O_n$  (where  $O_n$  is observed state that has  $A_n$  and  $S_n$  torsion angles) which realizes this maximum.

After forming transition probability matrix, determination of  $\delta_t$  is done.  $\delta_t(i)$  arrays contain the maximum probabilities for each grid.

**Initialization step** only consists of  $\delta_1$  which is  $1 \times I$  vector.

$$\delta_1(i) = \Pr(q_1 = S_i \text{ and } a_1 = p_i | b_i(O_1)) \quad (19) [318]$$

where  $1 \leq i \leq I$ .

**Induction step** contains  $\delta_t$  which is  $1 \times I$  vector

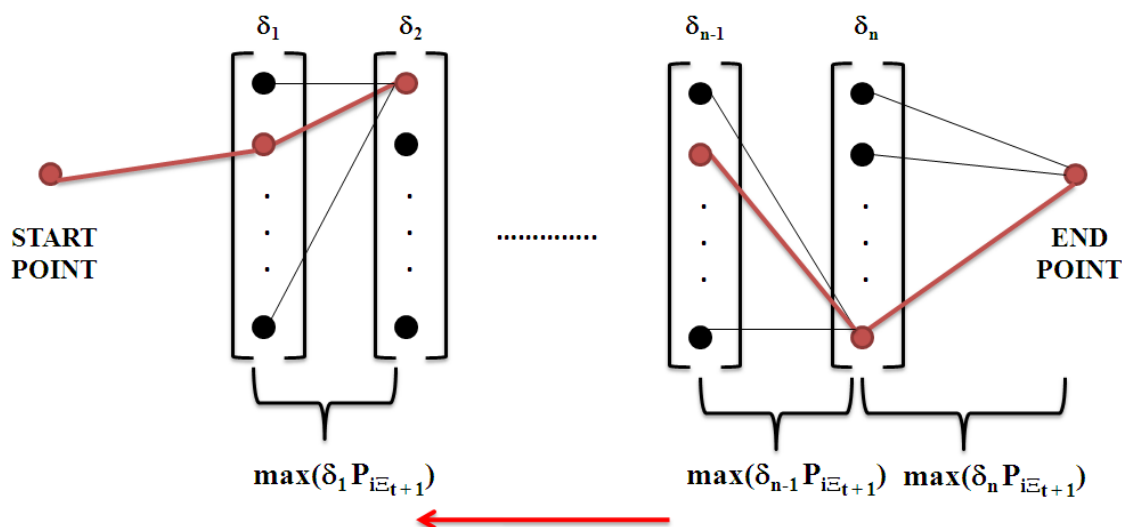
$$\delta_t(j) = \max_i(\delta_{t-1}(i) P_{ij} b_j(a_t | a_{t-1})) \quad (20) [318]$$

where  $1 \leq i, j \leq I$ , and  $2 \leq t \leq n$ .

**Backtracking** shows which path is appropriate for peptide when going from one state to another. The initial and final states are known from minimization and binding. Only middle states are needed to be determined. Consequently, backtracking starts from the last state to previous state by choosing:

$$\Xi_t = \operatorname{argmax}_i \delta_t(i) P_{i\Xi_{t+1}} \quad (21) [269]$$

where  $1 \leq i \leq I$  as shown in Figure 3.7.



**Figure 3.7** Schematic representation of backtracking in Viterbi algorithm. Red arrow shows the direction of backtracking and red dots and lines show  $\max(\delta_t(i) P_{i|E_{t+1}})$  values.

As a result,  $q_t$  which is  $\varphi$  angle is found for  $1 \leq t \leq n$ . Backtracking gives minimum energies from passing one state to another for a residue in tripeptide. Path that is followed by tripeptide is determined by backtracking. Viterbi decoding is written in MATLAB 7.6 [319]. In summary, the algorithm is based on the following steps:

1. The  $\psi$  values between states A and B are divided into  $t$  steps.
2. At each value of  $\psi$ , equally spaced  $m$  points are chosen for  $\varphi$ . This creates a grid of  $txm$  stations between points A and B.
3. The minimum conformational energy of the residue is calculated for each of the  $txm$  points. The minimization is performed by allowing all degrees of freedom of the residue to rearrange for the given value of the  $\psi$  -  $\varphi$  pair.
4. At each of the  $t$  steps, transition probabilities,  $P_{ij}$  ( $1 \leq j \leq m$ ) are calculated
5. Transition probabilities are used for maximum probability,  $\delta$ , determination. Here, emission probabilities,  $b_i$ , are assumed to be 1.
  - a. Initialization step, where  $\delta_1$  is determined by initial transition probability.

- b. Induction step, where  $\delta_t$  is determined ( $2 \leq t \leq n$ ) by using  $\delta_{t-1}$ .
6. Backtracking is done by from the last state to previous state.

### 3.4 The Interaction Matrix for the Caspase-1 Pathway and Correlations

We assume that there are  $n$  proteins in the pathway. We denote the instantaneous states of the  $n$  proteins by an  $n$ -dimensional vector  $\mathbf{R}$  whose  $i$ 'th entry  $R_i$  denotes the state of the  $i$ 'th protein. The state is defined in its most general sense. For example the unbound state of a protein and the bound states are two distinct states. Similarly, a phosphorylated protein is in a different state than an unphosphorylated one.

We also define the function of a protein which results from changes in its state and changes in the states of other proteins that directly or indirectly interacts with it. By function, we mean the activity of the protein that contributes to an experimentally observable event related to that protein. The instantaneous function of the set of  $n$  proteins is given by an  $n$ -dimensional vector  $\mathbf{F}$  whose  $i$ 'th entry  $F_i$  indicates the function of the  $i$ 'th protein.

All protein-protein interactions in the caspase-1 pathway are considered in a connectivity (Kirchhoff) matrix ( $\Gamma$ ), which is related with  $\mathbf{F}$  and has 107 elements. These 107 elements are determined from literature review, and are the main proteins in the caspase-1 pathway.

$\Gamma$  is the connectivity matrix of the pathway which is defined as

$$\Gamma_{ij} = \begin{cases} -1 & \text{if } i \neq j \text{ and } i \& j \text{ interact} \\ 0 & \text{if } i \neq j \text{ and } i \& j \text{ do not interact} \\ -\sum_{j, j \neq i} \Gamma_{ij} & \text{if } i = j \end{cases} \quad (22)$$

The entries of the  $\Gamma$  matrix may be taken in other ways than that of the connectivity matrix. The present definition relates to an undirected graph. Thus, for example, although the phosphorylation of protein  $i$  by protein  $j$  is effectively an unsymmetric interaction, it is ignored in the present notation and only the fact that whether they interact or not is considered.

The only assumption in the model is that the fluctuation of states obeys Gaussian distribution and that the coefficient matrix is given by Eqn.23. We assume that the large fluctuations from the mean being less probable. This leads to the distribution  $W \Delta \mathbf{R}$  of fluctuations

$$W \Delta \mathbf{R} = \frac{\exp\left[-\frac{1}{2} \Delta \mathbf{R}^T \Gamma \Delta \mathbf{R}\right]}{\int \exp\left[-\frac{1}{2} \Delta \mathbf{R}^T \Gamma \Delta \mathbf{R}\right] d \Delta \mathbf{R}} \quad (23) \quad [320]$$

The interaction of protein  $i$  and  $j$  are then obtained as the correlation of fluctuations of the states of these two proteins, i.e.,

$$\langle \Delta R_i \Delta R_j \rangle = \Gamma^{-1}_{ij} \quad (24)$$

This is interaction matrix of correlations of proteins in the Caspase-1 pathway, which are given in Table 3.2. This matrix conveniently allows for the study of knock-out experiments where removal of a given protein from the pathway leads to a new interaction matrix of correlations, and the difference between the two interaction matrices gives the importance of the contributions from the knocked-out protein.



**Table 3.2** The list of used proteins in the ICE pathway.

#	Protein	#	Protein	#	Protein	#	Protein
1	P2X7R	28	TRADD	55	cPLA2	82	Presenilin-1
2	SGT-1	29	TRAF2	56	iPLA2	83	Presenilin-2
3	HSP90	30	RIP2	57	RIG-I	84	Huntingtin
4	SAA	31	Caspase-1	58	IRF3	85	EGFR
5	Amyloid- $\beta$	32	NF- $\kappa$ B	59	IRF7	86	AR
6	NLRP3	33	Pypaf7	60	ERK $\frac{1}{2}$	87	CDK11A
7	CARDINAL	34	Procaspase-11	61	Interferons	88	CDK11B
8	AIM2	35	Caspase-11	62	Rac1	89	Parkin
9	ASC	36	p38 MAPK	63	Pak1	90	MAPT
10	Procaspase1	37	Pro-IL-1 $\beta$	64	MyD88	91	BCAP31
11	MEFV Pyrin	38	Pro-IL-33	65	MD2	92	XIAP
12	PSTPIP-1	39	Pro-IL-18	66	TLR1	93	Pypaf3
13	POP	40	IL-1 $\beta$	67	TLR3	94	Pypaf4
14	BcL-X1	41	IL-33	68	Bid	95	Pypaf5
15	BcL-2	42	IL-18	69	Mal	96	TXNIP
16	NLRP-1	43	Substrates*	70	PYNOD	97	Actin
17	NOD2	44	NOD1	71	ATN1	98	TNFR-II
18	Caspase-5	45	TLR4	72	Pro-caspase-4	99	Stat1
19	Pannexin-1	46	NALP2	73	Pro-caspase-8	100	tp53
20	Flagellin	47	TLR2	74	Caspase-4	101	Sptan-1
21	NLRC4	48	TLR9	75	Pro-IL-1F7	102	TLR5
22	INCA	49	TLR6	76	Caspase-8	103	IRF5
23	COP	50	TLR7	77	SerpinB9	104	Caspase-12
24	Iceberg	51	Biglycan	78	Caspase-14	105	MVK
25	NAIP5	52	TRIF	79	Procaspase-10	106	TNFRSF1A
26	TNF- $\alpha$	53	PI3K	80	PARP1	107	Exogenous inhibitors
27	TNFR-1	54	Phospholipase C	81	Caspase-10		

\*Substrates are the proteins that only interact with caspase-1 in the ICE pathway. These are TIM, GADPH, Aldolase,  $\alpha$ -enolase, TFAP2A, PPAR- $\gamma$ , pyruvate kinase, FGF-2, ATXN-3, NEDD4, Calpastatin, SREPBs and pro-IL-1 $\alpha$ .

## Chapter 4

### 4. RESULTS & DISCUSSIONS

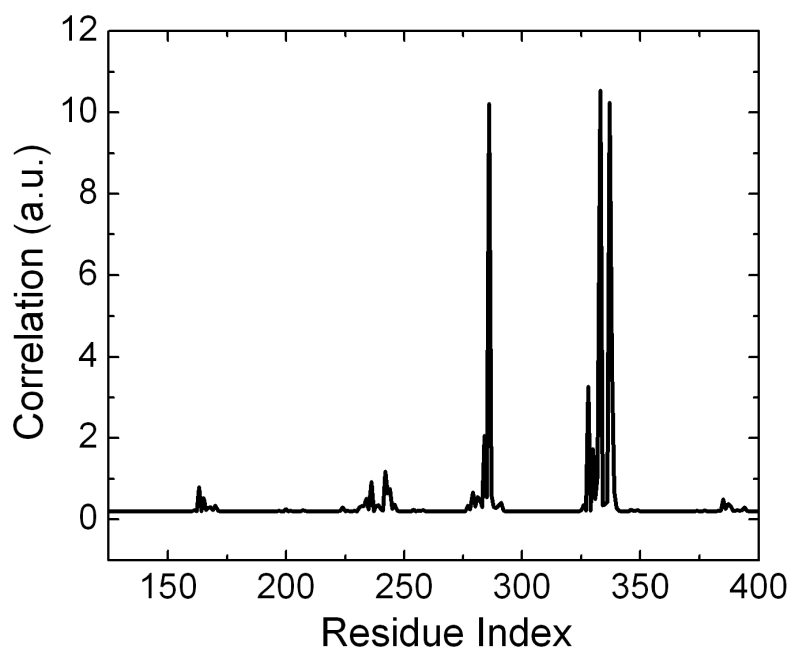
#### 4.1 Peptide-based drug design

##### 4.1.1. Determination of docking place

Suitable and fast methods for finding possible binding sites are generally based on Gaussian approaches [2]. Here, we used the GNM for determining the suitable binding regions [296-297, 316]. GNM assumes that two residues which are within a cutoff distance of each other interact with a harmonic potential. The cutoff distance is chosen as 7.0 Å. These inter-residue interactions lead to an interaction matrix for the protein, whose  $ij$ 'th entry is unity if residues  $i$  and  $j$  are within the cutoff distance. The diagonal elements of the interaction matrix equate to the negative sum of the row elements [290, 297]. The  $ij$ 'th element of the inverse of the interaction matrix indicates the extent of correlations between the  $i$ 'th and  $j$ 'th residues. The correlation,  $C_i$ , of the  $i$ 'th residue with its surroundings, which is a measure of the energy exchange of the residue with its surroundings is given as

$$C_i = \sum_j \Gamma^{-1}_{ij} \quad (25)$$

The surroundings of a residue defined in this formulation include the other residues of the protein as well as any ligands that it binds. Hence, a residue with high  $C_i$  may be regarded as the one that may interact with ligands. The larger eigenvalues of the interaction matrix lead to significant interactions of the residues with their surroundings. The three largest eigenvalues of the interaction matrix are used in obtaining the inverse of the interaction matrix given in Eq. 25 [297]. Results of calculations on this basis are shown in Figure 4.1 where, four regions of binding residues are identified, spanning the ranges 160-171, 230-247, 276-291, and 324-341.



**Figure 4.1** Interactions taken from GNM.

The residues identified in this manner are in general agreement with literature values. The most popular binding sites in the literature include residues of Arg179, His237, Gly238, Gln283, Cys285, Ser339 and Arg341 [22, 244]. There are some inhibitors designed for caspase-1 in PDB data bank that is docked at the sulphur atom of Cys285

(PDB IDs: 1BMQ, 1IBC, 1RWN, 1RWV, 1RWW, 1RWX, 2H4Y, 2H51, 2H54, 2HBR, 2HBY, 3D6F). Taking His237 and Cys285 as the docking residues, or the hotspots, is reasonable since the interaction between ligand and protein is based on histidine-cysteine dyad of enzyme [9, 12]. Moreover, His237 and Cys285 have interaction peaks in Figure 4.1. The allosteric site Cys331 is also found to be a possible docking site by Scheer et al. [17]. In our calculations, we consider all of the hotspots Arg179, His237, Gly238, Gln283, Cys285, Ser 339 and Arg341. The best tripeptides in each hotspot (found by GA) are considered with equilibrium and kinetic factors by Viterbi algorithm. Probabilities for each converged peptide in docking places are given in Table 4.1.

**Table 4.1** Probabilities residues having different states in tripeptides at docking place candidates.

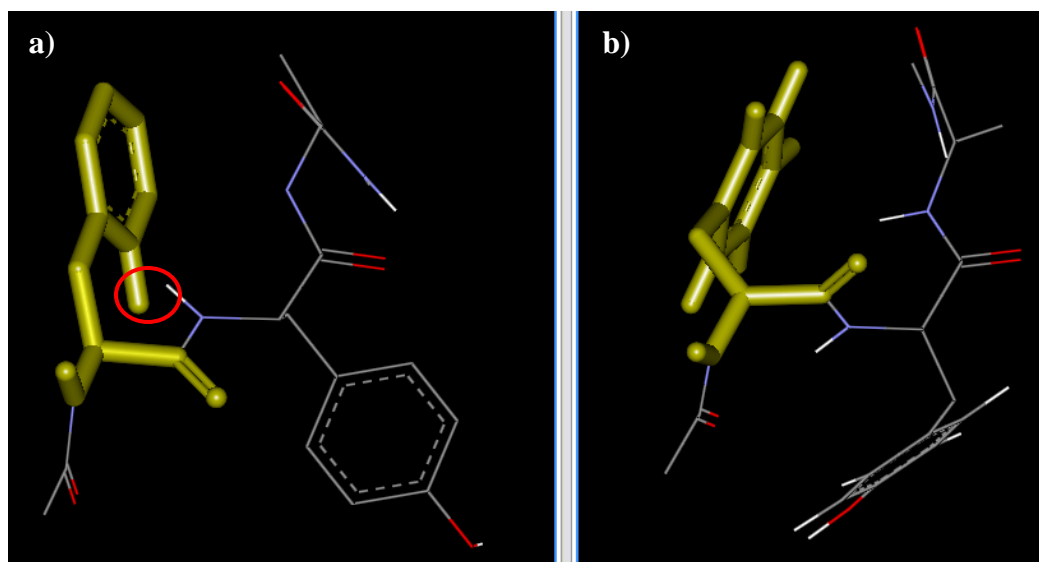
Binding place	Mismatch Residue	State A	State B	Minimized Energy*	Bound Energy*	$\Delta E^*$	$P_B/P_A$
Arg179 (GPS)	P2	16	8	-76.02	-57.94	18.08	$1.59 \times 10^{-13}$
His237 (AWG)	A1	10	8	-39.78	-38.79	0.99	0.20
Gly238 (FYA)	F1	8	10	-95.27	150583.65	150678.92	0
Gln283 (GAP)	G1	18	7	-13.79	-17.04	-3.25	199.84
	P3	16	8	-27.8	-9.14	18.66	$6.17 \times 10^{-14}$
Cys285 (YWG)	G3	8	12	-24.12	-24.24	-0.12	1.22
Cys331 (SCP)	S1	8	10	-19.75	-18.07	1.68	0.06
	P3	16	8	-23.12	-3.52	19.6	$1.33 \times 10^{-14}$
Ser339 (VAP)	V1	8	11	-10.21	-9.94	0.27	0.64
	P3	16	8	-26.87	-7.05	19.82	$9.32 \times 10^{-15}$
Arg341 (WPA)	P2	16	11	-67.19	-42.03	25.16	$1.55 \times 10^{-18}$

\*All energies in (kcal/mol).

The GNM results point to eight candidate residues on the protein surface for binding. These are given in the first column of Table 4.1. The tripeptides obtained by GA are presented in column 1. A tripeptide is given for each binding location on the protein surface. The minimum energy of the free tripeptide is given in column 3. It is obtained by minimizing the energy of the tripeptide. The energy of the bound tripeptide is given in column 4. This energy is obtained by taking the bound conformation, removing the protein, and calculating the energy of the conformation of the tripeptide in the absence of the protein. Column 7 lists the energy difference between the bound conformation and the unbound one. The probabilities of the bound and unbound states are given in column 8 as the ratio  $P_B/P_A$ , where the individual probabilities are obtained by using Eqn.17. Very small values of the ratio points to the impossibility of the bound conformation relative to the free conformation.

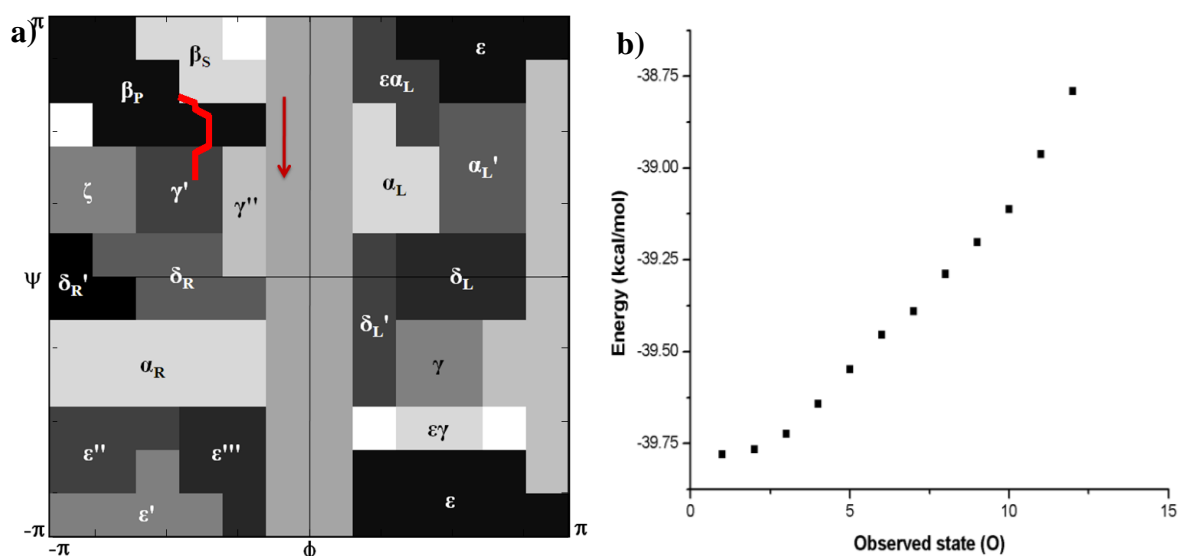
The configuration of the tripeptide upon binding to the protein surface is determined by the torsion states of the three residues. The two possibilities are that the torsion state of a residue in the bound configuration may be the same as the most probable state, or it may differ. The latter is an unfavorable contribution to binding. We refer to this residue as the mismatch residue. In column 2, the mismatch residues of the designed tripeptides are given. In columns 3 and 4, the states of the mismatch residue in the initial and bound configurations are given. In the first case, for example, GPS is converged at Arg179 of the protein and has the minimized state of 10-16-8 and the bound state of 10-8-8. Its 1<sup>st</sup> and 3<sup>rd</sup> residues have the same states but its 2<sup>nd</sup> residue, P2, is a mismatch residue, with state 16 in the most probable free conformation, and with state 8 in the bound conformation. In five of the tripeptides shown in Table 4.1, Proline is the mismatch residue, and changing the torsion angles of proline is expensive due to its ring structure. Thus the binding of the tripeptides GPS, GAP, SCP, VAP and WPA are unfavorable.

The tripeptide FYA binds to Gly238. It has one mismatch residue, F1. When Phe is changing the conformation of Phe from the 8<sup>th</sup> state (most probable) into the 10<sup>th</sup> (bound) results in a large positive energy change, leading to a very small probability ratio. This big energy difference is related to the bound state of FYA since in its bound state, phenyl ring of the FYA residue is started to come closer to the backbone as seen in Figure 4.2. It is about to form a new ring structure but it is impossible to manage a new ring. That is why, high energy is required by the Phe in FYA.



**Figure 4.2** Behavior of Phe residue in FYA in **a)** bound and **b)** minimized conformations.

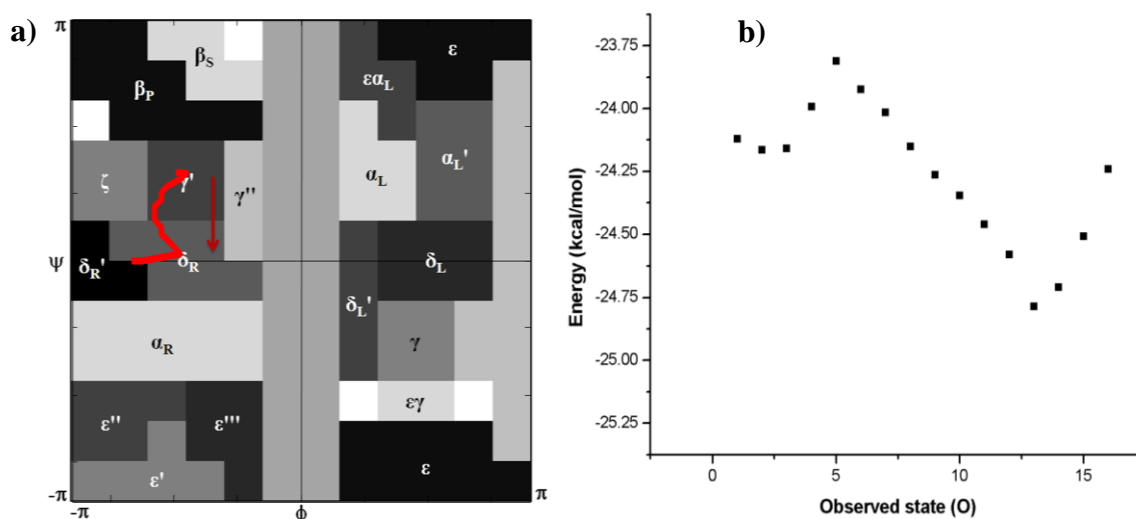
The tripeptide AWG binds to His237 of the protein, which goes from state 10-11-3 to 8-11-3. Its 1<sup>st</sup> residue has different states, 10 and 8, in minimized and bound conformations, respectively. The energy difference between the most probable and the bound state of the tripeptide is about one kcal/mol. The optimum path on the energy landscape is shown in Figure 4.3a, and the corresponding energy profile is shown in Figure 4.3b. These calculations show that AWG may be a good candidate for binding to His237.



**Figure 4.3** VitAL results for AWG.

**a)** Path followed by Ala in AWG in Ramachandran plot. Red arrow shows the direction of path. **b)** Energy calculated by HyperChem versus each observed states by Ala in AWG.

YWG is the converged sequence for Cys285. Its conformation changes from 8-4-8 to 8-4-12, the 3<sup>rd</sup> residue changing its state from 8 to 12. The optimum path on the energy landscape is shown in Figure 4.4a, and the energies are shown in Figure 4.4b. The energy barrier in the transition is not high and the probability ratio for YWG is higher than that of AWG (His237). Consequently, Cys285 is conformationally a more appropriate choice for caspase-1 inhibition.



**Figure 4.4** VitAL results for YWG.

**a)** Path followed by Gly in YWG in Ramachandran plot. Red arrow shows the direction of path. **b)** Energy calculated by HyperChem versus each observed states by Gly in YWG.

#### 4.1.2. Previously designed inhibitors for caspase-1

Some of the inhibitors designed for caspase-1 in literature and their docking results are given in Table 4.2. As seen, all drugs except *minocycline* have high Goldscore values. Besides, HyperChem score of *malonate* is found to be positive. Other HyperChem scores are found to be higher than -31.58 kcal/mol. Additionally, free binding energy of drugs are found between -5.98 kcal/mole and -7.85 kcal/mole from AutoDock. *Minocycline* has small Goldscore value, and *malonate* has positive HyperChem score. As a result, these two drugs have not good docking results. Moreover, *inhibitor 1* and *inhibitor 3* are found to have small docking scores with respect to AutoDock and Chemscore  $\Delta G$ . On the other hand, *inhibitor 3* has a high Goldscore value just like *inhibitor 1*, *inhibitor 2*, *inhibitor 4*, and *z-VAD-FMK*. *Inhibitor 5*, *inhibitor 6* and *Ac-WEHD-CHO* have also high scores of AutoDock but their Goldscore values are at intermediate level.



**Table 4.2** Docking results of previously designed inhibitors.

Inhibitor Name	Goldscore	Chemscore $\Delta G$ (kJ/mole)	HyperChem (kcal/mol)	AutoDock (kcal/mol)	PDB ID	Reference
Minocycline	20.65	-42.42	-30.63	-7.85	-	[321]
Inhibitor 1	70.39	-12.13	-21.06	-6.19	1RWO	[322]
Inhibitor 2	67.66	-30.65	-30.6	-7.78	1RWV	[323]
Inhibitor 3	70.17	-19.02	-23.06	-5.98	1RWW	[242]
Inhibitor 4	64.15	-13.8	-28.58	-7.07	1RWX	[242]
z-VAD-FMK	62.1	-43.22	-28.1	-6.20	2H4Y	[8]
Inhibitor 5	44.74	-11.57	-20.91	-6.24	1RWK	[322]
Inhibitor 6	42.8	-15.61	-22.8	-7.68	1RWM	[322]
Malonate	36.93	-19.81	3.63	-6.97	1SC3	[22]
Ac-WEHD-CHO	50.8	-18.81	-31.58	-6.76	1IBC	[235]

\* **Inhibitor 1:** 4-oxo- 3-(6-[4-(quinoxalin-2-ylamino)-benzoylamino]-2-thiophen-2- yl-hexanoylamino)-pentanoic acid, **Inhibitor 2:** 5-[5- (1-carboxymethyl-2-oxo-propylcarbonyl)-5-phenyl- pentylsulfamoyl]-2-hydroxy-benzoic acid, **Inhibitor 3:** 4-oxo- 3-[(6-([4-(quinoxalin-2-ylamino)-benzoylamino]-methyl)- pyridine-3-carbonyl)-amino]-butyric acid, **Inhibitor 4:** 4-oxo- 3-(6-[4-(quinoxalin-2-yloxy)-benzoylamino]-2-thiophen-2-yl- hexanoylamino)-butyric acid, **Inhibitor 5:** 3-(2- mercapto-acetylamino)-4-oxo-pentanoic acid, **Inhibitor 6:** 4-oxo- 3-[2-(5-([4-(quinoxalin-2-ylamino)-benzoylamino]-methyl)- thiophen-2-yl)-acetylamino]-pentanoic acid

### 4.1.3. Drug determination for caspase-1 inhibition

#### 4.1.3.1. D-amino acid based drug design

There are two mirror-image isomers of aminoacids known as L- and D-enantiomers. L-formed aminoacids are common in proteins and called as natural isomers. D-forms of aminoacids attract attention of many scientists since there is a certain advantage of D-enantiomeric forms. This advantage of D-forms is related to their specificity and hardly cross-reactive with L-forms. Consequently, it becomes hard to degrade them in body. That

is why they can be used for drug design purposes [324]. For instance, Bartzatt (2005) used D-amino acid as a drug carrier for antineoplastic nitrogen mustard groups and determined D-alanine N-mustard is in the level of those clinic antitumor drugs [325]. Moreover, L-amino acids are replaced with D formed ones in order to avoid degradation in body by Hamamoto et al. (2002). It is found that D-amino acid substitutions in human granulysin are beneficial in terms of metabolic degradation [326]. On the other hand, there are no drugs designed for caspase-1 using D-enantiomeric residues. As a result, D-amino acid based drug design is considered with docking programs of GOLD and AutoDock. GOLD can dock many structure faster than AutoDock so firstly dockings are done with GOLD. Moreover, HyperChem is also used for evaluation of minimized energy of ligands docked.

In Table 4.3, scores of best aminoacids in both forms can be seen. There is not so much difference between using either L- or D- form with respect to scores found for Trp, His and Phe.

**Table 4.3** Aminoacids GOLD score results.

Aminoacid Type	Goldscore	Chemscore $\Delta G$ (kJ/mole)	HyperChem (kcal/mol)
L-TRP	41.93	-17.94	-23.33
D-TRP	37.56	-20.13	-27.27
D-PHE	37.08	-23.15	-16.81
D-HIS	36.88	-12.94	-21.11
L-HIS	36.85	-12.55	-16.25
L-PHE	35.82	-23.84	-22.56
D-TYR	35.14	-16.5	-23.5

Consideration of D-amino acid drugs is made for Trp, Phe and Tyr. Chemscore  $\Delta G$  and HyperChem scores of histidine are lower than tyrosine. Besides, scores from GOLD demonstrated using an aminoacid is favorable than *minocycline*, *malonate* and *inhibitor6* despite adding no additional chemical or biological agent. Furthermore, it is observed that using an aminoacid is better than drugs in Table 4.2 on account of Chemscore  $\Delta G$  values.

Besides, some HyperChem scores in Table 4.3 such as D-TRP are in the same with the ones in Table 4.2. In order to improve scores of Table 4.3, addition of chemical groups is done by iLib Diverse with D-Trp, D-Phe and D-Tyr. A random formation of molecules is done by considering all fragments in iLibDiverse.

**Table 4.4** D-aminoacid+ 1 x iLib Diverse results.

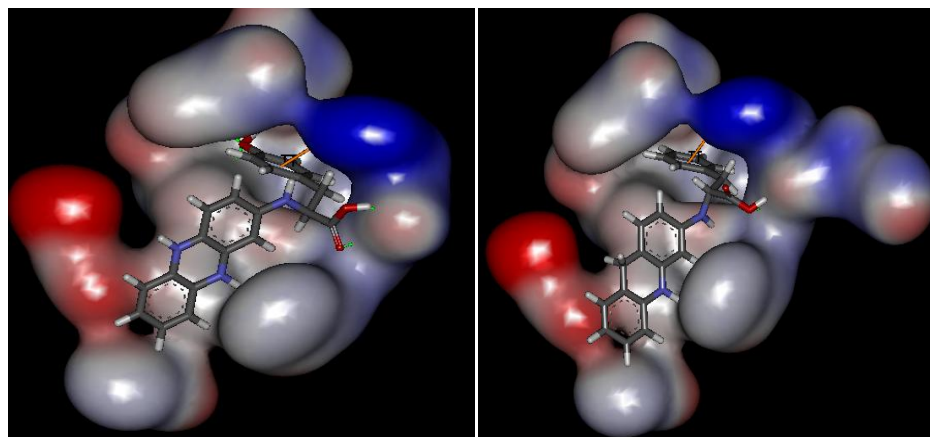
#	Structure	Gold score	Chemscore $\Delta G$ (kJ/mol)	HyperChem (kcal/mol)	AutoDock (kcal/mol)	MW (g/mol)
1	d-tyr+ester(4)	34.07	-20.43	-22.57	-5.60	233
2	d-tyr+ether(17)	44.01	-27.74	-29.7	-5.69	301
3	d-tyr+ester(10)	32.9	-18.76	-32.09	-6.09	315
4	d-tyr+ether(5)	35.22	-25.52	-29.2	-6.34	349
5	d-trp+alcohol(12)	37.46	-22.22	-35.09	-5.55	264
6	d-tyr+monocyclic(24)	42.69	-20.22	-29.64	-5.59	251
7	d-tyr+bicyclic(22)	43.19	-31.17	-37.67	-6.31	359
8	d-tyr+ether(12)	38.4	-26.29	-32.9	-5.62	301
9	d-phe+bicyclic(4)	33.65	-22.09	-29.48	-7.16	342
10	d-phe+bicyclic(7)	37.4	-25.16	-29.48	-6.68	309

Scores in Table 4.4 show that HyperChem values are lower or in the level of inhibitors found previously as seen from Table 4.2. However, Goldscore values in Table 4.4 cannot pass the values found for *inhibitor 1*, *inhibitor 2*, *inhibitor 3*, *inhibitor 4*, *z-VAD-FMK* and *Ac-WEHD-CHO* despite having positive Goldscores that pass values of *minocycline*, *inhibitor 5*, *inhibitor 6*, and *malonate*. Thus, inhibitors in Table 4.4 have better HyperChem scores whereas their Goldscores are not so high like *inhibitor 1*, *inhibitor 2*, *inhibitor 3*, *inhibitor 4*, *z-VAD-FMK* and *Ac-WEHD-CHO*. Chemscore  $\Delta G$ s in Table 4.4 cannot pass only *minocycline* and *z-VAD-FMK*.

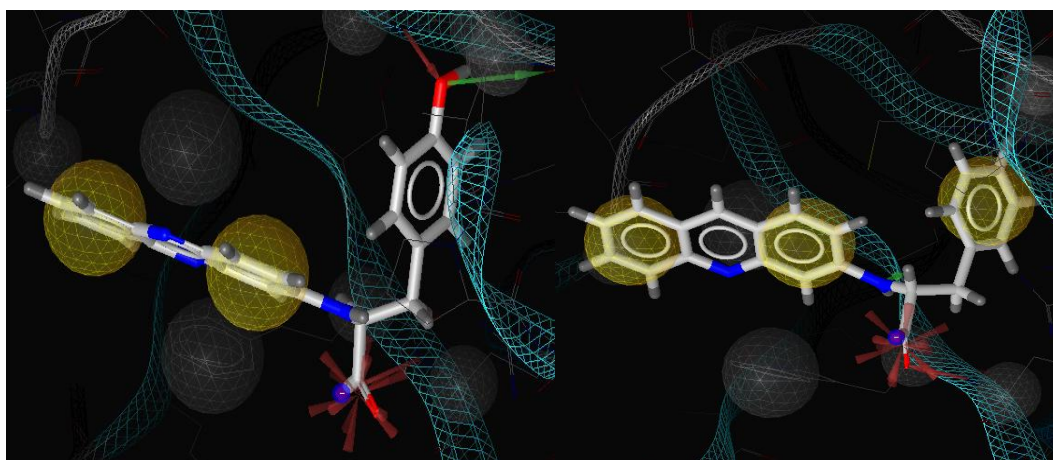
In AutoDock results, the highest value of inhibitors designed previously in Table 4.2 is found to be -7.85 kcal/mol. Despite having no drugs in Table 4.4 that passed this

value, there are some candidates in Table 4.4 that can pass this value by molecule additions. Besides, there are some inhibitors designed in Table 4.2 that passed -5.98 kcal/mol free energy. Consequently, *d-tyr+ether(5)*, *d-tyr+bicyclic(22)*, *d-phe+bicyclic(4)* and *d-phe+bicyclic(7)* are selected according their AutoDock results among ten structures determined as the top 4 drugs found from 1200 structures. Karanewsky et al. (1998) also stated bicyclic structures are potent inhibitors for ICE [236]. Among these 4 structures, *d-tyr+bicyclic(22)* has the best Chemscore  $\Delta G$  and HyperChem values. It has also good Goldscore value. Since best AutoDock results belongs to *d-phe+bicyclic(4)* and *d-tyr+bicyclic(22)*, they are selected two templates for addition of molecules. Other unselected drugs with good AutoDock value, which are *d-tyr+ether(5)* and *d-phe+bicyclic(7)*, had similar AutoDock values with *d-phe+bicyclic(4)* and *d-tyr+bicyclic(22)*. They had also higher values of Chemscore  $\Delta G$  and HyperChem results compared with *d-phe+bicyclic(4)* and *d-tyr+bicyclic(22)*.

Structures of *d-tyr+bicyclic(22)* and *d-phe+bicyclic(4)* are shown in Figure 4.5. Images are held from best AutoDock run with considering the free energy of binding energy and inhibition constant, which are -7.28 kcal/mol and 4.62  $\mu\text{M}$  for *d-tyr+bicyclic(22)* and -8.08 kcal/mol and 1.2  $\mu\text{M}$  for *d-phe+bicyclic(4)*. These values are good since heptapeptides found by Unal et al. (2010), which are larger molecules than *d-tyr+bicyclic(22)* and *d-phe+bicyclic(4)*, had free binding energy and inhibition constant that are lower than -10 kcal/mole and 50 nM [270]. Analysis of interactions between caspase-1 and two structures found showed that *d-tyr+bicyclic(22)* had hydrogen bonds with Cys285, Gly238, Arg341 and Ser 236. Furthermore, it has pi-cation interaction with Arg 179 and Arg 341 as seen in Figure 4.5. For *d-phe+bicyclic(4)*, hydrogen bonds are with Arg341 and Ser339 and pi-cation interactions exist between Arg179 and Arg341, respectively.



**Figure 4.5** Docked structures of *d-tyr+bicyclic(22)* (left) and *d-phe+bicyclic(4)* (right). Ligands are in Stick representation and residues within 4Å are in surface representation. Hydrogen bonds and pi interactions are shown with green and orange lines by Accelrys Discovery Studio 2.5.



**Figure 4.6** Pharmacophores of *d-tyr+bicyclic(22)* (left) and *d-phe+bicyclic(4)* (right).

Pharmacophores of *d-tyr+bicyclic(22)* and *d-phe+bicyclic(4)* are also analyzed by using their docked conformations gained from the best AutoDock run. In Figure 4.6, the *d-tyr+bicyclic(22)* has two hydrophobic interactions, a negative ionizable area, a hydrogen bond donor and two hydrogen bond donors. *d-phe+bicyclic(4)* has three hydrophobic

interactions with one hydrogen acceptor and donor. Moreover, it has a negative ionizable area.

**Table 4.5** D-aminoacid+ 2 x iLib Diverse results.

(*d-tyr+bicyclic(22)* and *d-phe+bicyclic(4)* named as *d-tyr* and *d-phe*)

#	Structure	Goldscore	Chemscore ΔG (kJ/mol)	HyperChem (kcal/mol)	AutoDock (kcal/mol)	MW (g/mol)
11	d-tyr+ester(5)	35.34	-42.34	-33.46	-6.12	417
12	d-tyr+amine(24)	41.31	-31.34	-40.31	-6.53	444
13	d-phe+monocyclic(26)	45.77	-36.75	-32.91	-7.04	409
14	d-phe+ester(12)	30.66	-33.48	-35.03	-7.23	426
15	d-phe+bicyclic(21)	28.23	-35.95	-28.0	-7.38	486
16	d-phe+ester(10)	55.5	-27.23	-32.7	-7.24	440
17	d-phe+monocyclic(10)	38.02	-22.75	-39.05	-7.25	423
18	d-phe+monocyclic(27)	46	-22.51	-29.19	-7.90	439
19	d-phe+bicyclic(18)	40.79	-55.94	-35.49	-7.46	458
20	d-phe+monocyclic(20)	38	-35.08	-30.64	-7.04	426
21	d-phe+monocyclic(37)	39.53	-31.05	-37.15	-6.85	409
22	d-phe+monocyclic(38)	37.71	-35.19	-21.15	-6.46	409
23	d-tyr+pharma(32)	18.69	-45.55	-42.73	-7.99	467
24	d-phe+L-Pro	24.94	-22.37	-33.03	-8.86	455
25	d-phe+L-Gly	31.73	-26.3	-45.19	-8.21	399
26	d-tyr+L-Gly	42.63	-38.98	-25.18	-7.40	430

Using these two determined structures as templates (*d-tyr+bicyclic(22)* and *d-phe+bicyclic(4)* named as *d-tyr* and *d-phe*), search for new drugs are performed with considering the maximum molecular weight that a molecule can have for permeation and adsorption according to Lipinski Rule of 5 [217]. By considering this fact, a second addition of structures is done by iLib Diverse with the most successful compounds. These are esters, amines, monocyclics, bicyclics, pharmaceuticals and L-aminoacid groups. There are some inhibitors in Table 4.5 whose HyperChem, Chemscore ΔG and AutoDock results are generally higher than the values in Table 4.2. However, Goldscore values are

determined to be lower than the Goldscores of *inhibitor 1*, *inhibitor 2*, *inhibitor 3*, *inhibitor 4*, *z-VAD-FMK* and *Ac-WEHD-CHO* but other docking findings in Table 4.5 are determined to be generally higher than these four inhibitors. Consequently, these inhibitors are alternative drugs for caspase-1 inhibition.

*d-tyr-ester(5)* has hydrogen bonds with Gln283, Arg341 and Ser236 and has -7.43 kcal/mol minimum free binding energy (its free energy of binding in Boltzmann Distribution is -6.12 kcal/mol) with 3.56  $\mu$ M inhibition constant. It has a pi-pi interaction with Trp340 and pi-cation interaction with Arg179 and Arg341 as having the top Goldscore.

*d-phe-monocyclic(27)* has a minimum free binding energy of -8.73 kcal/mol with 400.45 nM. It also has hydrogen bonds with Arg179, His237, Gln283, Arg341 and Ser236. Moreover, it has a pi-cation interaction with Arg383 and Arg341.

*d-phe-monocyclic(20)* has minimum free binding energy of -7.62 kcal/mol with 2.58  $\mu$ M. It has hydrogen and pi-cation interaction with Arg341, Ser339 and Arg341, Arg179, respectively.

Furthermore, *d-tyr+pharma(32)* has -9.59 kcal/mol free binding energy with 93.16 nM. It has a pi interaction with Arg341 and hydrogen bonds with Arg179, Gln283, Arg341, Ser339 and Asp288.

*d-phe-L-Pro* which has the minimum binding energy has hydrogen bonds with Arg179, Gln283, Arg341 and Ser236. It has a pi-sigma interaction with His237. Its minimum free binding energy and inhibition constant are -10.24 kcal/mol and 31.36 nM.

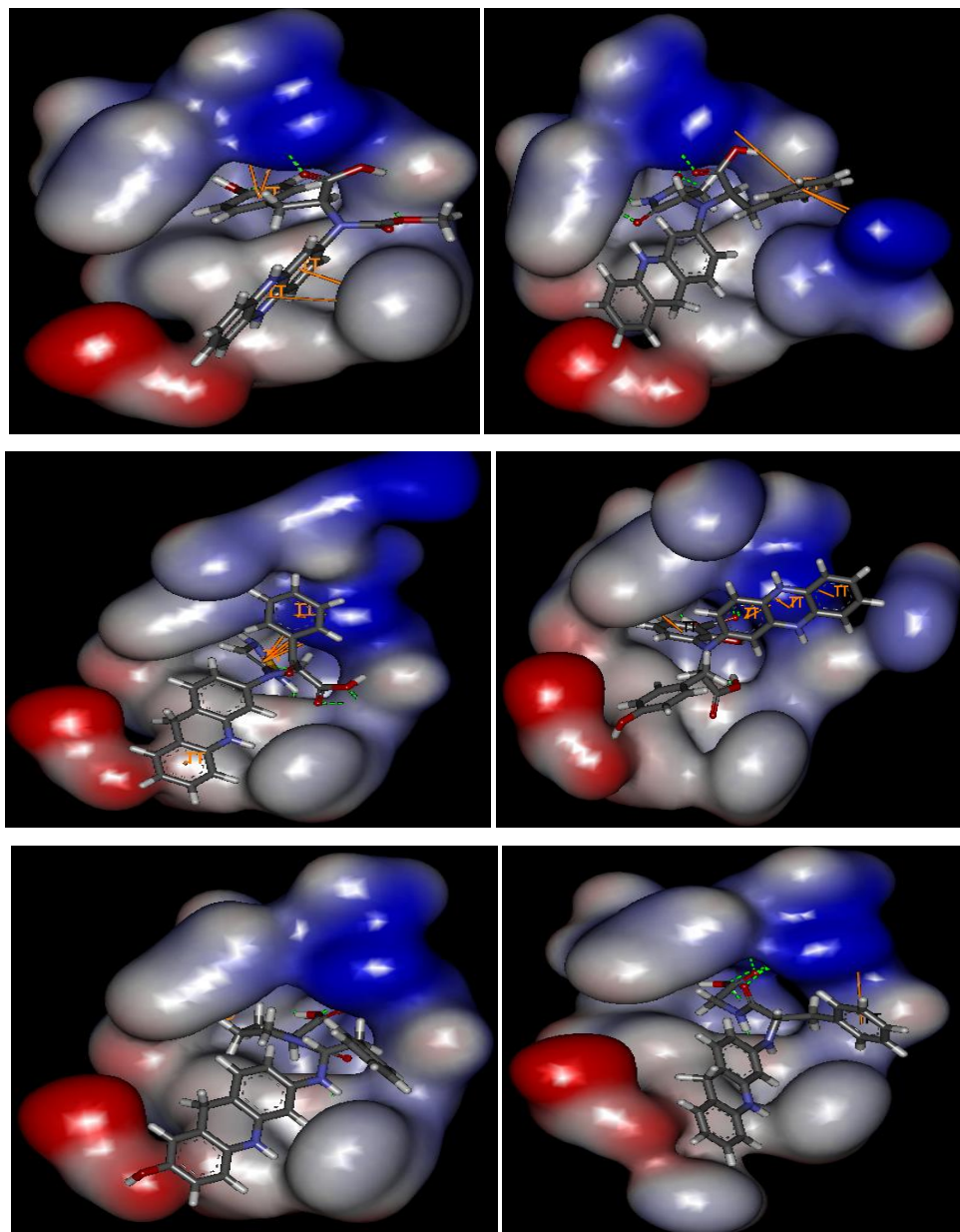
As a final consideration, *d-phe-L-Gly* has hydrogen bonds with Arg179, Gln288, Arg341 and Ser339 with -9.71 kcal/mol minimum binding energy and 76.94 nM inhibition constant. Its only pi interaction is with Arg341. Their docking images are illustrated at Figure 4.7, which are taken from the best AutoDock run by considering free energy of

binding and inhibition constant. Using an additional L-form aminoacid is improved the scores considerably without passing the 500 g/mol level.

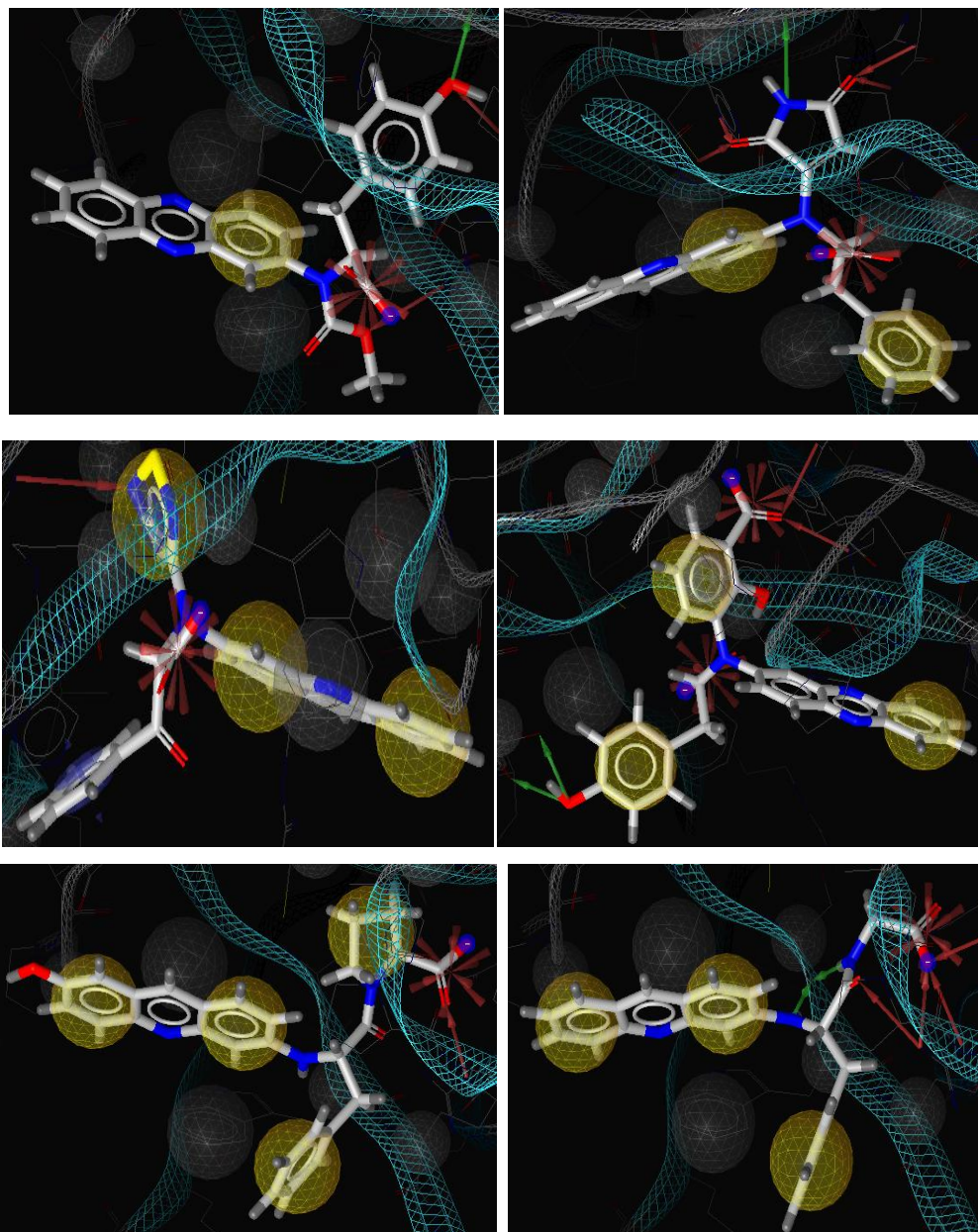
*d-Tyr-ester(5)* has a hydrophobic interaction with two hydrogen bond acceptors, one hydrogen bond donor and one negative ionizable area. *d-Phe-monocyclic(27)* has two hydrophobic interaction with four hydrogen bond acceptors and a donor and also a negative ionizable area. *d-Phe-monocyclic(20)* has three hydrophobic interactions and one hydrogen bond acceptor. It also has a negative ionizable area with an aromatic ring. *d-Tyr+pharma(32)* has two hydrogen bond acceptors and donors with three hydrophobic interactions and two negative ionizable area. *d-Phe-L-Pro* has four hydrophobic interactions with one hydrogen bond acceptor and one ionizable area. *d-Phe-L-Gly* has three hydrophobic interactions with five hydrogen bond acceptors, one hydrogen donor and one negative ionizable area as seen from Figure 4.8.

The molecular structures of inhibitors found in this part are given at Appendix I. There are 27 d-aminoacid based drugs found for caspase-1 inhibition. In Appendix I, representations are given with corresponding structure number.





**Figure 4.7** Docked structures of *d-Tyr-ester(5)* (top left), *d-Phe-monocyclic(27)* (top right), *d-Phe-monocyclic(20)* (middle left), *d-Tyr+pharma(32)* (middle right), *d-Phe-L-Pro* (bottom right) and *d-Phe-L-Gly* (bottom left). Ligands are in Stick representation and residues within 4Å are in Surface representation. Pi (orange lines) and hydrogen (green lines) interactions are shown.



**Figure 4.8** Pharmacophores of *d-Tyr-ester*(5) (top left), *d-Phe-monocyclic*(27) (top middle), *d-Phe-monocyclic*(20) (top right), *d-Tyr+pharma*(32) (bottom right), *d-Phe-L-Pro* (bottom middle) and *d-Phe-L-Gly* (bottom right).

#### 4.1.3.2. Aspartic acid and malonate based drug search

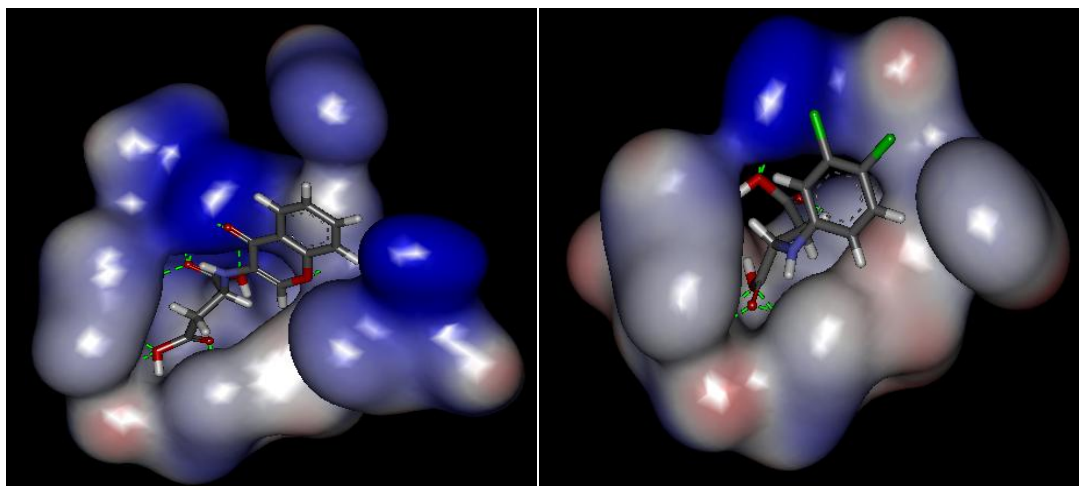
There are many drugs for caspase-1 that contains an aspartic acid residue since caspase-1 cleaves IL-1 $\beta$  from Asp116-Ala117. Moreover, there is a certain need of Asp residue in caspase-1 inhibitors in the P<sub>1</sub> position [1]. Additionally, *malonate* is found to be successful in terms of caspase-1 inhibition as a small inhibitor by Romanowski et al. [22]. Asp and *malonate* are used as templates for adding molecular structures taken from iLib Diverse 1.02. Fragment number is chosen as 2. One of them is Asp or *malonate* and the other fragment is taken from iLib Diverse database. Following step is to use GOLD in order to eliminate the structures formed fastly in an efficient manner. According to positive Goldscore and negative Chemscore  $\Delta G$  values, 9 structures are found good in terms of ICE inhibition. These structures are also docked by AutoDock with 2,500,000 number of energy evaluation. Energies of ligands are also minimized by using HyperChem. Docking results are given at Table 4.6.

**Table 4.6** Docking results of Asp/Malonate + 1 x iLib Diverse.

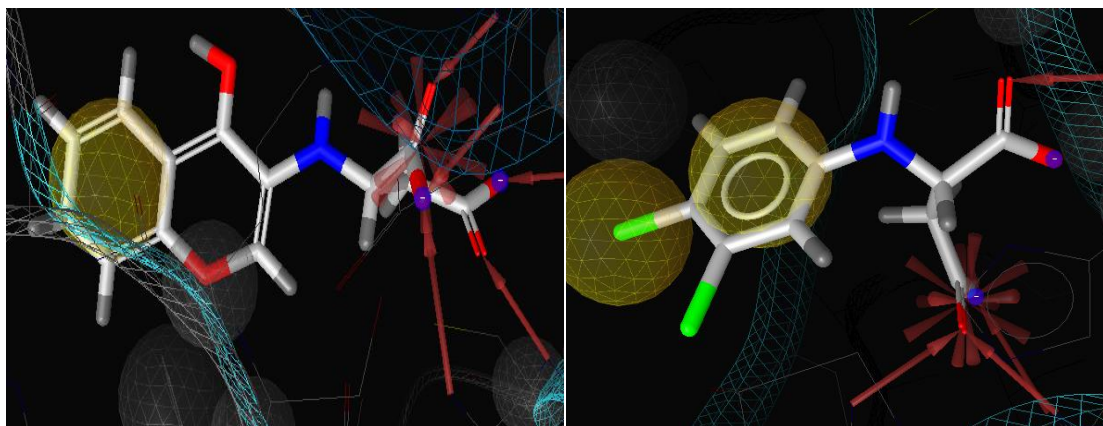
#	Name	Goldscore	Chemscore $\Delta G$ (kJ/mol)	HyperChem (kcal/mol)	AutoDock (kcal/mol)
27	Bicyclic-asp(13)	42.63	-20.52	-29.01	-6.51
28	Monocyclic-asp(34)	41.18	-19.09	-26.81	-6.25
29	Monocyclic-asp(7)	40.53	-15.06	-26.8	-5.98
30	Monocyclic-asp(17)	38.92	-20.06	-23.78	-5.62
31	Benzene-asp(12)	38.63	-16.45	-23.64	-6.81
32	Monocyclic-asp(6)	38.4	-17.68	-25.87	-5.69
33	Ester-asp(7)	38.34	-14.3	-30.99	-5.92
34	Ketone-asp(2)	37.53	-18.21	-20.87	-5.41
35	Bicyclic-asp(4)	37.06	-6.7	-27.89	-5.65

There is no *malonate* containing inhibitors in Table 4.6 due to lower docking scores than Asp. As a result, using Asp is better than *malonate*. *Bicyclic-asp(13)* is found to have the highest scores of Goldscore and Chemscore  $\Delta G$  that are 42.63 and -20.52 kJ/mol, respectively. Its free energy binding energy and HyperChem score are also good with a value of -6.51 kcal/mol and -29.01 kcal/mol. *Benzene-asp(12)* has the highest free binding energy with a value of -6.81 kcal/mol. In order to improve the scores for reaching values in Table 4.2, a fragment addition onto *Bicyclic-asp(13)* and *Benzene-asp(12)* are done.

Before that, analysis of docked structures of these two inhibitors is done. *Bicyclic-asp(13)* has -7.86 kcal/mol minimum free energy binding energy with an inhibition constant of 1.74  $\mu\text{M}$ . It has hydrogen bonds with Arg179, His237, Gly238, Cys285 and Arg341. *Benzene-asp(12)* has hydrogen bonds with Arg179, Gly238, Gln283, Cys285, Arg341 and Ser236 with -8.18 kcal/mol minimum free binding energy and 1.01  $\mu\text{M}$   $K_i$ . Docked structures of these compounds are shown at Figure 4.9.



**Figure 4.9** Docked structures of *Bicyclic-asp(13)* (left) and *Benzene-asp(12)* (right). Ligands are in Stick representation and residues within 4Å are in Surface representation. Pi interactions (orange lines) and hydrogen bonds (green lines) are also demonstrated.



**Figure 4.10** Pharmacophores of *Bicyclic-asp(13)* (left) and *Benzene-asp(12)* (right).

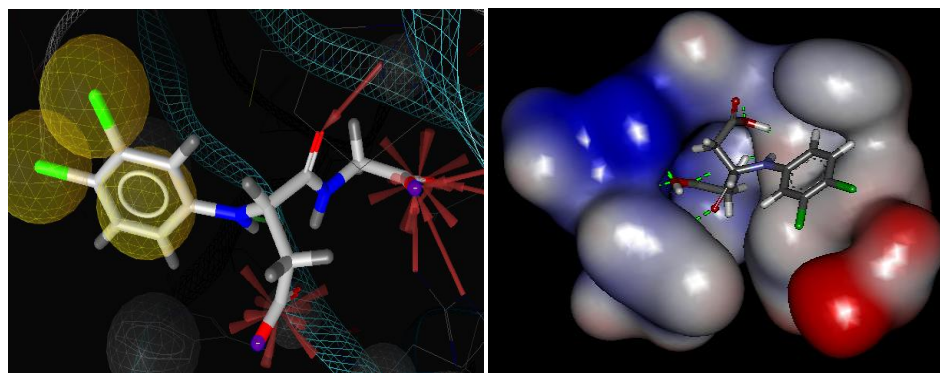
*Bicyclic-asp(13)* has one hydrophobic area with a negative ionizable area. It also has five hydrogen acceptor with cLogP and TPSA values of -4.66 and 125.66, respectively. Its molecular weight is 277.232 g/mol. *Benzene-asp(12)* has two hydrophobic areas with a negative ionizable area. It has 276.075 g/mol molecular weight with cLogP and TPSA values of -2.477 and 92.29, respectively. It also has five hydrogen bond acceptors as given in Figure 4.10.

Two structures are kept constant and additional molecules are used with the best groups for Asp, which are aminoacids, ketones, monocyclic, bicyclic, ester and benzene. *Benzene-asp(12)* has the best results after addition of molecules onto its structure. *Benzene-asp(12)* is called as *a1* in Table 4.7. There are some ketone based structures which are already found suitable candidates for ICE inhibition by Thornberry et al. (1994) and Brady et al. (1999) [247-248].

**Table 4.7** Docking results for 10 structures from 2 x iLib Diverse.

#	Name	Goldscore	Chemscore $\Delta G$ (kJ/mol)	HyperChem (kcal/mol)	AutoDock (kcal/mol)
36	L-Gly-a1(11)	28.62	-17.91	-31.59	-6.86
37	ketone-a1(15)	21.62	-17.16	-28.96	-5.14
38	monocyclic-a1(17)	27.4	-16.05	-32.88	-6.56
39	L-Cys -a1(18)	18.86	-12.4	-25.81	-6.64
40	monocyclic-a1(23)	22.42	-14.59	-29.73	-5.91
41	monocyclic-a1(4)	27.33	-17.08	-17.94	-5
42	L-Ala-a1(6)	18.93	-14.56	-17.57	-6.08
43	ketone-a1(7)	23.93	-14.34	-28.95	-5.6
44	ketone-a1(8)	24.2	-18.25	-25.33	-6.01
45	esters-a1(9)	20.62	-14.2	-32.7	-5.21

*L-Gly-a1(11)* is determined to be the best structure in Table 4.7 due to its good HyperChem and AutoDock results. It has three hydrophobic areas with two negative ionizable areas and six hydrogen acceptors. It also has one hydrogen donor with -3.36 cLogP and 121.39 TPSA. Its molecular weight is 333.127 g/mol and minimum free binding energy is -8.65 kcal/mol with  $K_i$  of 454.05 nM. It also has hydrogen bonds with Arg179, His237, Gln283, Arg341 and Ser339. Docked conformation of this inhibitor is given at Figure 4.11 with its pharmacophore.

**Figure 4.11** Pharmacophore (left) and docked conformation (right) of *L-Gly-a1(11)*.

Since *L-Gly-ai(11)* has low molecular weight with good docking results, an addition to its structure by iLib Diverse is also done but there is no candidate structure that improves docking. As seen from the comparison of docking scores of D-enantiomeric and Asp based inhibitors, using D-enantiomeric form of aminoacid can be an alternative way for using Asp residue.

#### 4.1.3.3. Tripeptide sequence prediction by Genetic Algorithm

*YWG* is found to be the first member of *last population* of genetic algorithm which is the same tripeptide for the last 5 generations. Rest of peptides in Table 4.8 belongs to other members of the same *population* which is in the best peptides in population. *XWG*, where X can be any aminoacid, structure of peptide is converged as the last two peptides of the tripeptide. Without any priori information, GA gives a tripeptide with a 2<sup>nd</sup> residue of Trp and a 3<sup>rd</sup> residue of Gly. Trp is an aromatic polar residue whereas Gly is nonpolar. Both residues are neutral [327]. Thus, one of the tripeptides in is *DWG* whose 1<sup>st</sup> residue is Asp. Aspartic acid was previously used many times as the 1<sup>st</sup> residue in caspase-1 inhibitors [12-13].

**Table 4.8** Docking scores of converged tripeptides.

#	Name	Goldscore	Chemscore $\Delta G$ (kJ/mol)	AutoDock (kcal/mol)	MW (g/mol)
46	AWG	54.69	-21.82	-7.68	368.39
47	DWG	67.21	-24.29	-7.37	412.4
48	GWG	61.48	-21.38	-7.66	354.37
49	YWG	67.96	-27.64	-7.69	460.49

\*The AutoDock values are repeated at 2,500,000 number of energy evaluations.

As *fitness* value in GA, AutoDock scores are used which are given in Table 4.8. Unal et al. (2010) also used AutoDock binding energy as fitness value for peptide based

drug of NF- $\kappa$ B. A heptapeptide was found with a value of -9.37 kcal/mol fitness score [270]. When a comparison is made between our and Unal et al.'s (2010) values, fitness values in Table 4.8 are good for a tripeptide. Additionally, Goldscore and Chemscore  $\Delta G$  values are determined to be good which also validate the tripeptides in Table 4.8.

#### 4.1.3.4. Analysis of drugs in this study for inhibition of ICE

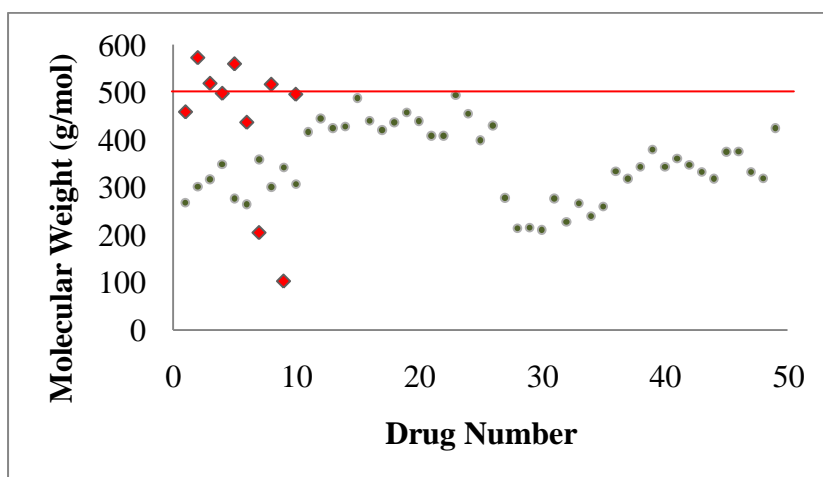
There are 49 structures determined for caspase-1 inhibition. Despite their satisfactory results, an analysis related to their drug-likeness is done.

The most popular set of rules is described by Lipinski et al. (2001) [217]. According to this set of rules, plots in Figure 4.12, Figure 4.13, Figure 4.14, Figure 4.15 and Figure 4.16 are done in which red lines show maximum limit of an orally bio-available drug given by Lipinski et al. [217]. Half of literature drugs pass the 500 Da (g/mole) limit whereas none of found drugs in this study passes. Not only literature drugs but also found drugs in this study do not pass the limit for cLogP value. For hydrogen acceptor case, one of literature drugs pass the limit while there is no found drug passed maximum limit. Additionally, there is no literature and found drugs pass maximum value of hydrogen donor. As a result, we can claim that found drugs are more suitable than literature drugs in terms of Lipinski rule of 5. In addition to Lipinski, found drugs are suitable with respect to Ghose et al. (1999) rules of molecular weight and ClogP (they are approximately same range with Lipinski rule of 5) which considers high drug-likeness of compounds [212]. Besides, Walters & Murcko [MW=200-500, HDO=0-5, HAC=0-10] filter demonstrates found drugs are suitable [216].

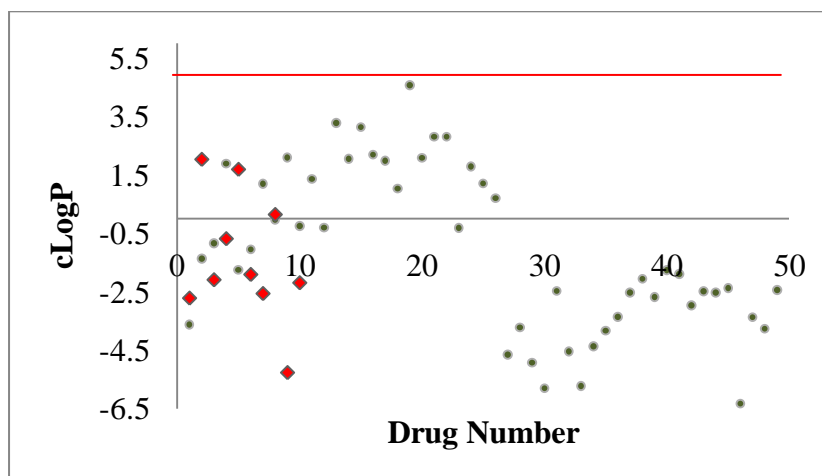
According to Palm filter that only consider the polar surface area (TPSA), some of found drugs with  $TPSA > 139 \text{ \AA}^2$  are absorbed with a 10%. On the other hand, there are completely absorbed drugs due to  $63 \text{ \AA}^2$  TPSA in some of the drugs found as clear from Figure 4.14 [219]. For blood brain permeable drugs analysis, Murcko et al. (1999) found



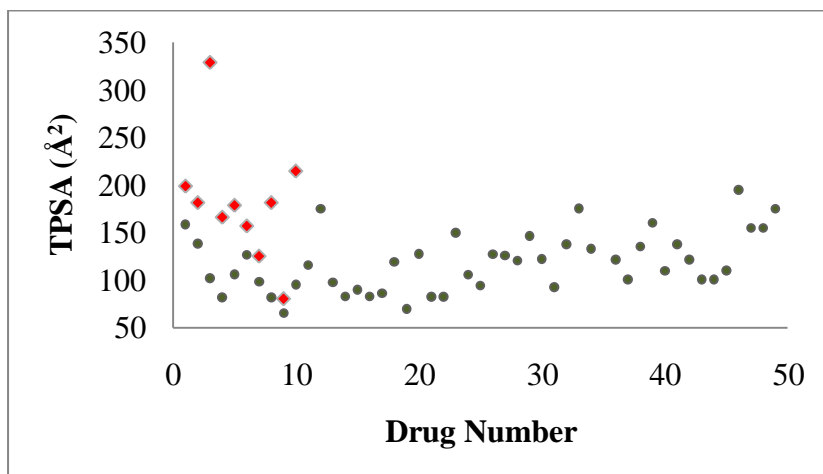
that  $MW=200-450$ ,  $ClogP=0-5.2$ ,  $HAC= \max.4$ ,  $HDO=\max.3$ . There are also some candidates that fall into these rules if plots in Figure 4.12, Figure 4.13, Figure 4.14, Figure 4.15, Figure 4.16 and the limits of Murcko filter are considered [221].



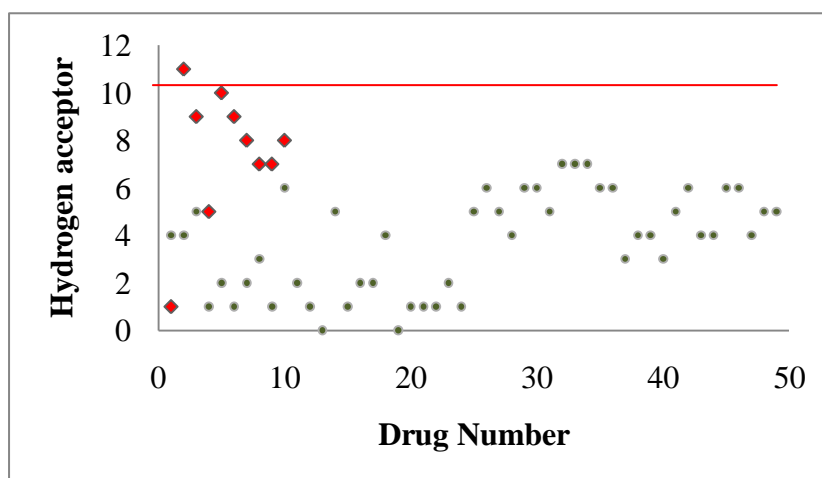
**Figure 4.12** MW of literature (red) and found drugs (green) with Lipinski limits [217].



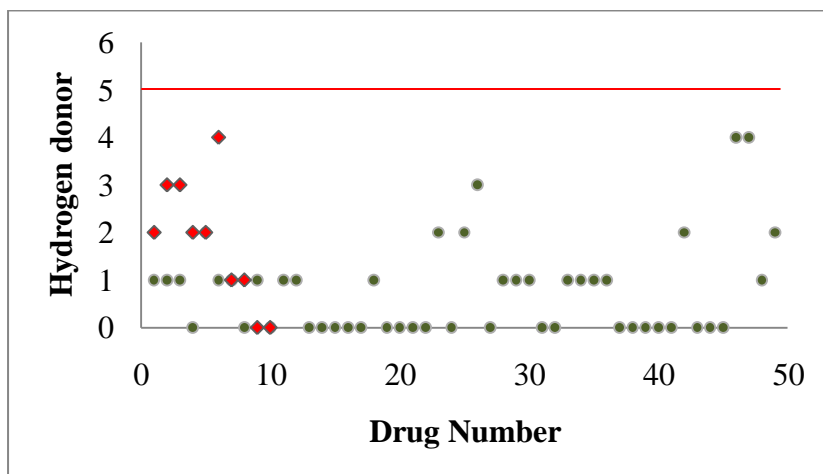
**Figure 4.13** cLogP of literature (red) and found drugs (green) with Lipinski limits [217].



**Figure 4.14** TPSA of literature (red) and found drugs (green).



**Figure 4.15** Hydrogen acceptor of literature (red) and found drugs (green) with Lipinski limits [217].



**Figure 4.16** Hydrogen donor of literature (red) and found drugs (green) with Lipinski limits [217].

#### 4.1.4. Multi-target drug determination considering the ICE pathway

Pathway analysis is done for ICE in order to find drugs that can both affect ICE itself and its pathway proteins. In literature review part, a detailed knowledge related to ICE pathway is given. If found drugs for caspase-1 in this study are also found suitable for other proteins in the activation pathway of caspase-1, it will be easier to deactivate ICE function. However, any protein in this pathway cannot be considered owing to possible affects on unrelated systems.

First, 3D structures of all proteins in the ICE pathway are searched from PDB data bank [311]. Structures of 1F2H for TRADD (Tumor necrosis factor receptor type 1-associated DEATH domain protein), 1F3V for complex between TRADD and TRAF (Tumor necrosis factor receptor-associated factor), 1PN5 for pyrin domain of NLRP1, 1UCP for pyrin domain of ASC, 2KN6 for ASC containing CARD, 2OQ0 and 3B6Y for AIM2, 3KAT for CARD domain of NLRP1 and 3E4C for procaspase-1 are found.

After finding these structures, a search for each protein is done in order to identify any effects on unrelated systems. TRADD and TRAF form complex that can activate

caspase-1 but they can also activate caspase-8 which affects the apoptotic pathways [328-329]. Consequently, inhibition of this complex is not suitable due to affecting apoptotic pathway. Despite being a necessary component for caspase-1 activation, NLRP1 has also interactions with caspase-2 and caspase-9 by forming a complex called apoptosome. Inhibition of NLRP1 will also affect the apoptotic pathways just like TRADD-TRAF2 complex so it is not logical to inhibit this protein [330]. ASC has also a proapoptotic role by activating caspase-9 as a result it will not be useful to inhibit this protein [79]. Moreover, AIM2 controls cell proliferation and deactivating this protein will affect the cell growth [331]. AIM2 is also able to activate a wild-type PYD–caspase-9 [122].

As seen inhibiting one protein can affect other systems, using procaspase-1 as inhibition target can solve this problem since it only goes to active form of caspase-1. It is not related to other proteins or systems. Elliot et al. (2009) studied the structure of procaspase-1 by mutating the Cys285 residue into Ala285 for analyzing the autoactivation of caspase-1 [332]. 3E4C is the PDB ID of the mutated structure of procaspase-1. In order to mutate back into its wild type, Ala285 is mutated into Cys285 with mutate function of HyperChem Release 7 software. After that, final structure is saved and then this structure is minimized with Accelrys Discovery Studio Visualizer 2.5. The reason of using Accelrys Discovery for minimization is related to number of residues in structure. Since there are too much residues for minimization in HyperChem, Accelrys is selected. Moreover, using NAMD will need too much time on account of residue number. However, Accelrys gives minimized structure in a short time. After that, minimized structure is also minimized two more times. Consequently, there are non-minimized, one time minimized and three times minimized structures for procaspase-1.

All found drugs in this study are docked to these three procaspase-1s by AutoDock. Docking place is selected as Asp297 as the main cleavage site for procaspase-1 [332]. In all three structures, drugs with bad scores always found to be bad in others and vice versa for

the ones with good scores. After that, a detailed GOLD runs are done for three procaspase-1s. Docking procedure is the same but docking place is taken as C atom of Asp297. Moreover, Lys296, Asp297 and Ser298 are made flexible without considering water molecules. Good valued inhibitors in three cases are selected whose docking scores are shown in Table 4.9.

**Table 4.9** Drugs for procaspase-1.

Name	Goldscore	Chemscore $\Delta G$ (kJ/mol)	Autodock (kcal/mol)	Name	Goldscore	Chemscore $\Delta G$ (kJ/mol)	Autodock (kcal/mol)
<b>10</b>	55.61	-20.48	-7.11	<b>21</b>	62.39	-28.39	-8.02
<b>11</b>	63.1	-27.53	-6.97	<b>22</b>	65.28	-30.33	-8.39
<b>12</b>	62.01	-33.87	-8.31	<b>23</b>	62.27	-35.5	-6.93
<b>13</b>	64.25	-26.64	-8.09	<b>24</b>	63.34	-26.66	-7.20
<b>14</b>	63.97	-30.23	-7.70	<b>25</b>	69.3	-29.06	-8.37
<b>15</b>	66.57	-29.8	-9.68	<b>26</b>	74.61	-32.62	-7.21
<b>16</b>	73.4	-30.83	-8.65	<b>3</b>	50.65	-21.94	-6.38
<b>17</b>	63.5	-26.75	-7.99	<b>4</b>	52.5	-25.16	-7.50
<b>18</b>	62.54	-26.34	-8.42	<b>7</b>	70.07	-31.68	-7.16
<b>19</b>	65.12	-32.3	-9.64	<b>AWG</b>	64.31	-22.73	-7.73
<b>2</b>	54.13	-20.45	-7.37	<b>DWG</b>	69.79	-23.36	-7.12
<b>20</b>	59.52	-27.27	-8.21	<b>YWG</b>	78.65	-26.13	-8.95

\* The chemical structures of numbered inhibitors are given in Appendix-I.

There are drug candidates for both inhibiting procaspase-1 and caspase-1. By these inhibitors, there will be two targets (procaspase-1 and ICE) in order to deactivate caspase-1 functions.

## 4.2 Conformational Factors in the Design of Peptide Drugs: Application to the Design of Caspase-1 Inhibitors.

### 4.2.1. Equilibrium and Kinetic Factor Analysis of Tripeptides

Minimized (A) and bound (B) states are found for 4 tripeptides in Table 4.10. AWG, DWG, GWG and YWG have only one different state between minimized and bound conformations. Calculation of minimization and binding energies for AWG, DWG, GWG and YWG is done considering energy change in a residue having different states. According to states, path and energy behavior of passing from minimized to bound conformation are determined for 4 tripeptides specified from GA. Knowing the energies at states A and B, probabilities of passing state A to B are evaluated as seen in Table 4.10.

**Table 4.10** Energy change and probabilities of residues having different states.

Peptides	Mismatch Residue	State A	State B	Minimized Energy*	Bound Energy*	$\Delta E^*$	$P_B/P_A$
AWG	A1	8	5	-102.48	-97.01	5.47	$1.34 \times 10^{-4}$
DWG	D1	3	8	-99.01	-98.75	0.26	0.655
GWG	G3	19	8	-107.15	-101.04	6.11	$4.73 \times 10^{-5}$
YWG	G3	8	12	-24.12	-24.24	-0.12	1.22

\*All energies in (kcal/mol).

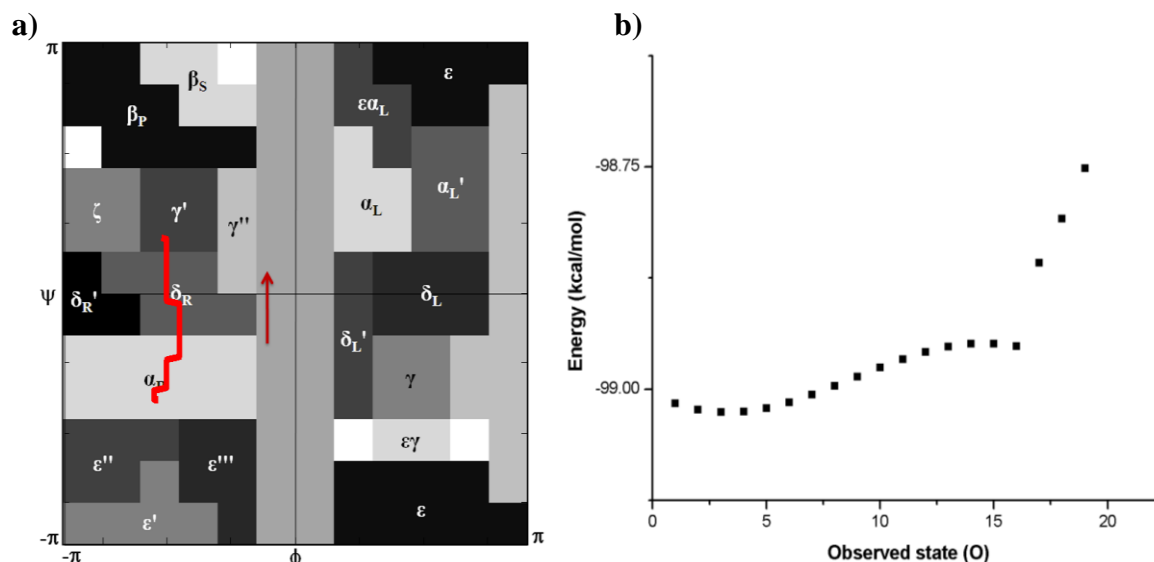
As clear from  $\Delta E$  and  $P_B/P_A$  values in Table 4.10, YWG has the highest probability ( $P_B/P_A$ ) due to low value of energy difference between A and B states. Thus, DWG can be an appropriate candidate because of its high probability. On the other hand, AWG and GWG are not suitable due to their low probabilities for passing from state A to B.

### 4.2.2. Determination of Path Followed by Tripeptides

Paths followed by tripeptides in Table 4.10 are found with considering energy behavior of these tripeptides by Viterbi decoding. In Figure 4.4, the results for YWG are

already given which indicates its suitability for changing its conformation from A to B. YWG is a good candidate because it has the highest probability and energy behavior, which does not make high energy difference.

1<sup>st</sup> residue (Asp) of DWG, which is another good peptide structure, has different states in its minimized (8) and bound (5) conformations. Asp changes its state by following the path given in Figure 4.17a. In addition, DWG makes a slight decrease and starts to rise after 5<sup>th</sup> observed state in its energy behavior given in Figure 4.17b. After making a small decrease at 15<sup>th</sup> and 16<sup>th</sup> observed states, a sharp increase in energies of Asp. However, its energy does not differ while changing state as seen in Figure 4.17b.

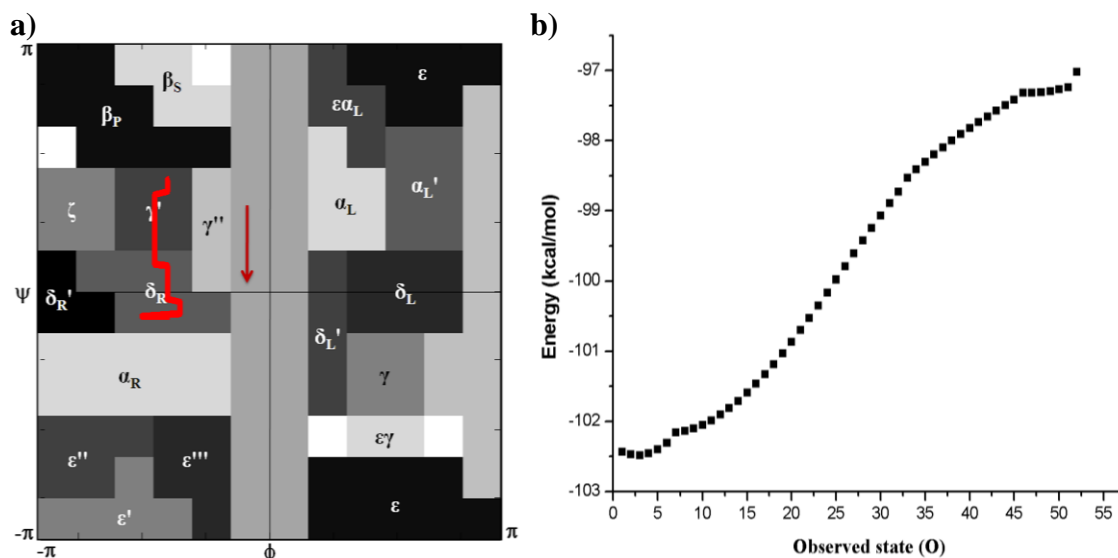


**Figure 4.17** VitAL results for DWG.

**a)** Path followed by Asp in DWG in Ramachandran plot. Red arrow shows the direction of path. **b)** Energy calculated by HyperChem versus each observed states by Asp in DWG.

1<sup>st</sup> residue (Ala) of AWG, which is not a good candidate, goes from 8<sup>th</sup> to 5<sup>th</sup> state as given in Figure 4.18a. Its energy change during this path is also plotted in Figure 4.18b. Energies of Ala start to increase from minimized to bound states but there is a rapid decrease at the beginning. After 3<sup>rd</sup> observed state, energy values of Ala increase steadily

up to last state. Furthermore, there seems to be a plateau after 47<sup>th</sup> observed state. Its energy change is high as changing its state.

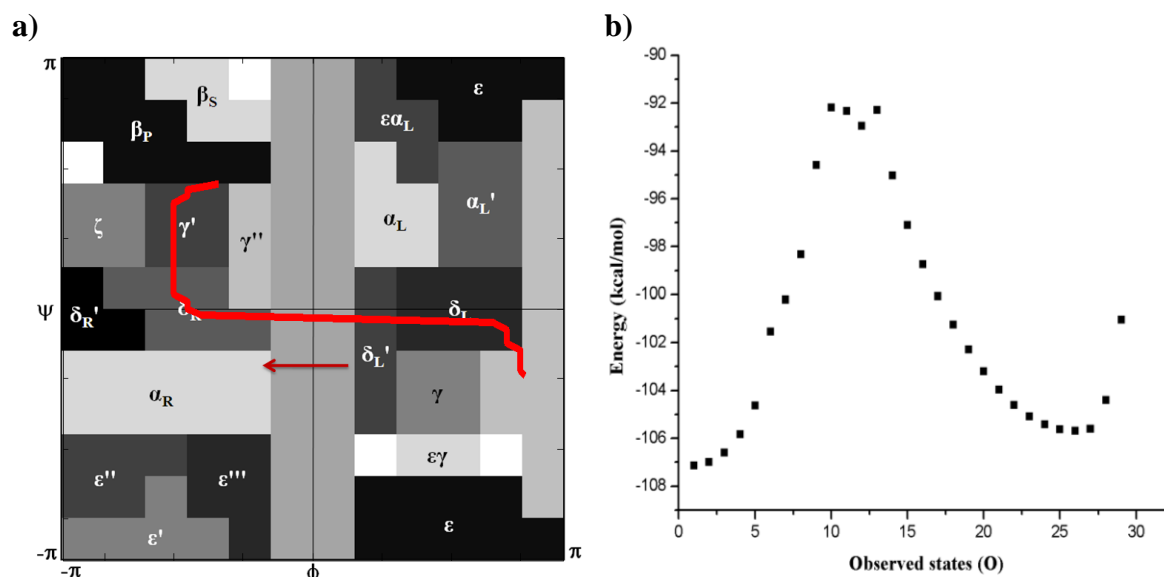


**Figure 4.18** VitAL results for AWG.

**a)** Path followed by Ala in AWG in Ramachandran plot. Red arrow shows the direction of path. **b)** Energy calculated by HyperChem versus each observed states by Ala in AWG.

States of GWG are only different in 3<sup>rd</sup> residue (Gly) because Gly goes from 19<sup>th</sup> to 8<sup>th</sup> states. Its way is very different from other peptides due to passing through different regions which are distinct from each other as demonstrated in Figure 4.19a. It is clear that passing from minimized to bound conformation requires passing high activation energy for Gly residue as clear from Figure 4.19b. Consequently, this peptide is the worst peptide for ICE inhibition.





**Figure 4.19** VitAL results for GWG.

a) Path followed by Gly in GWG in Ramachandran plot. Red arrow shows the direction of path. b) Energy calculated by HyperChem versus each observed states by Gly.

In addition, tripeptides for Cys285 are also determined by complete enumeration. The found tripeptides from complete enumeration are AWG, DWG, DFF, SFF, AWE and DYF. As seen, complete enumeration and GA give AWG and DWG. On the other hand, there are some different tripeptides found from complete enumeration. The conformational factors analysis by VitAL is also performed for DFF, SFF, AWE as well as DYF. The results for these tripeptides given in Appendix II and indicate these tripeptides are not suitable for ICE inhibition. Consequently, using GA is better approach than using complete enumeration

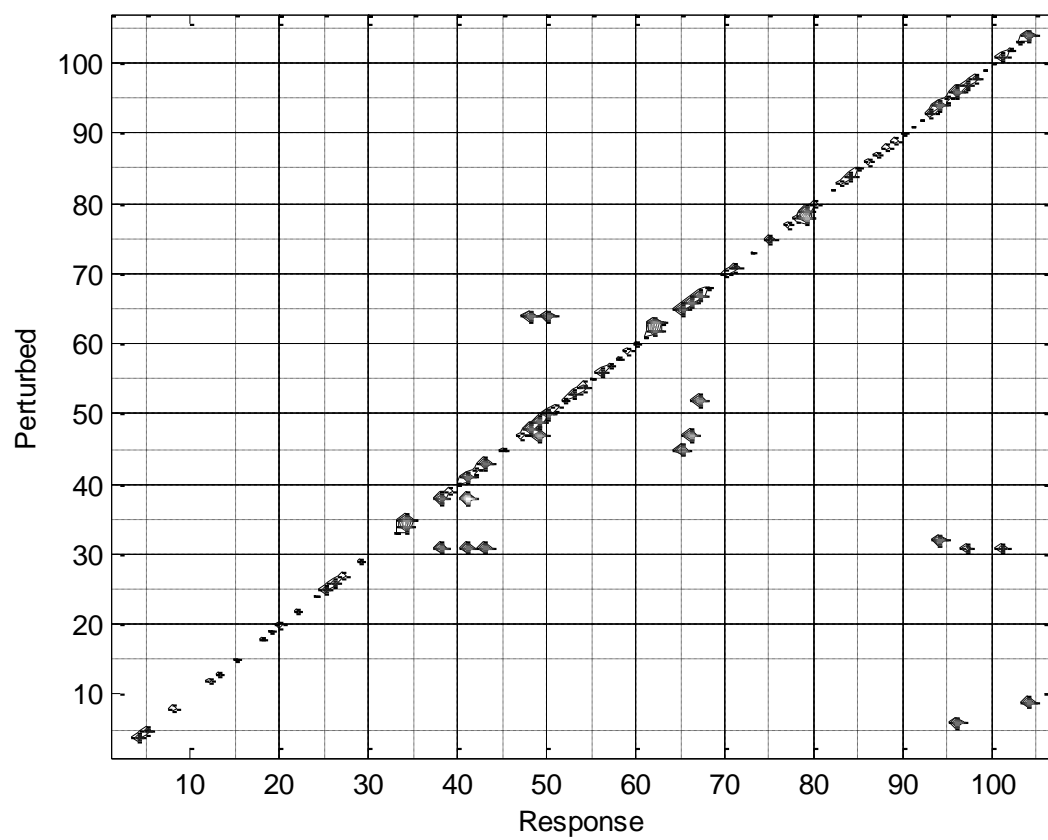
YWG and DWG are candidate peptides for caspase-1 due to their good computational docking scores and probabilities of passing from minimization to binding. In this part, investigation of their binding properties is done. Ace- and -nme are used as caps for these peptides and Ace- is one of the most favorable initial cap for caspase-1 inhibitors [13, 333]. However, there is no specific end cap for caspase-1 inhibitors in literature.

Despite having the same 2<sup>nd</sup> (Trp) and 3<sup>rd</sup> (Gly) residues with YWG and DWG, AWG and GWG are not suitable. Tyr and Asp residues are polar whereas Ala and Gly are not [327]. This may be a reason for AWG and GWG to change their conformation during binding process. In addition, DWG has Asp residue in its structures which was previously determined experimentally as a need for caspase-1 inhibitors [13].

YWG and DWG have hydrogen bonds with Cys285 and Asp288 while AWG and GWG do not. Thus, AWG has no hydrogen bond. The hydrogen bonds with Cys285 and Asp288 can be effective on binding process considerably since YWG and DWG, which are good peptides, have bonds with Cys285 and Asp288.

### 4.3 Knockout Analysis on the ICE pathway

Knockout proteins in the ICE pathway are determined by zero perturbation value as given in Figure 4.20. Perturbation means perturbing a protein in the system with decreasing its force constant. Force constant, which is taken -1 normally when there is an interaction between two proteins in the ICE pathway, is taken as zero for zero perturbation. In other words, interactions of a protein are removed in the ICE pathway when it is perturbed with zero force constant. That is why it can also be called as knockout of a protein. There are 9 knockout proteins in the ICE pathway as given in Table 4.11.



**Figure 4.20** Knockout analysis.

**Table 4.11** Results of Knockout Analysis.

#	Protein name	Defected proteins from knockout	The names of defected proteins from knockout
6	NLRP3	96	TXNIP
9	ASC	104	Caspase-12
31	Caspase-1	38, 41, 43, 97, 101	pro-IL-33, IL-33, substrates, actin, sptan-1
32	NF- $\kappa$ B	94	PYPAF4
38	Pro-IL-33	41	IL-33
45	TLR4	65	MD2
47	TLR2	49, 66	TLR1, TLR6
52	TRIF	67	TLR3
64	MyD88	48, 50	TLR7, TLR9

### ***1. Consequences of NLRP3 knockout:***

The knockout of NLRP3 from the system most strongly influences TXNIP (Thioredoxin-interacting protein) as clear from Figure 4.20. Three signaling pathways have been proposed for the activation of NLRP3 by Jin et al. [334-335]. In one of the signaling pathways, TXNIP serves as a link between ROS (Reactive oxygen species) and inflammasome activation. TXNIP interacts with TRX (Thioredoxin) in the steady state and upon addition of inflammasome activators (MSU (monosodium urate), H<sub>2</sub>O<sub>2</sub> and R-837), ROS is produced, causing TXNIP to dissociate from TRX and binds to NLRP3 as Zhou et al. found [336]. In other signaling pathway, potassium (K<sup>+</sup>) efflux is necessary. As a third pathway, uptake of crystals and particulates (alum, silica, amyloid- $\beta$ ) disrupts the phagolysosome acidic compartment and cathepsin B is released that activates NLRP3.

TXNIP specifically interacts with the LRR (Leucine-rich repeat protein) and NACHT (NTPase domain) domains of NLRP3. On the other hand, there are some debates about TXNIP since it is not clear whether TXNIP is a regulator or an activator as Franchi et al. stated [103]. Our knockout study shows that TXNIP and NLRP3 interaction is important.

### **2. Consequences of ASC knockout:**

Caspase-12 is determined to be affected mostly due to the knockout of ASC as seen in Figure 4.20. Caspase-12 is known to interact with ASC. Besides, ASC is a member of inflammasome structure and interacts with caspase-1. When caspase-12 interacts with ASC, ASC cannot bind caspase-1. As a result of this, caspase-12 inhibits caspase-1 activation as Zhou et al. and Miao et al. said [70, 337].

### **3. Consequences of Caspase-1 knockout:**

Caspase-1 knockout mostly affects substrates [*pro-IL-1 $\alpha$* , *TIM (triosephosphate isomerase)*, *GADPH (glyceraldehyde-3-phosphate dehydrogenase)*, *aldolase*,  *$\alpha$ -enolase*, *TFAP2A (Transcription factor AP-2-alpha)*, *PPAR- $\gamma$  (Peroxisome proliferator-activated receptor gamma)*, *pyruvate kinase*, *FGF-2 (Heparin-binding growth factor 2)*, *ATXN-3 (Ataxin-3)*, *NEDD4 (E3 ubiquitin-protein ligase NEDD4)*, *calpastatin*, *SREPBs (Sterol Regulatory Element-Binding Proteins)*], *pro-IL-33*, *IL-33*, *actin* and *sptan-1 (Spectrin alpha chain, brain)*.

IL-33 is the active form of *pro-IL-33* after caspase-1 cleavage as Dinerello et al. stated [1]. In our analysis, *pro-IL-33* and *IL-33* are determined to be defected in knockout analysis of caspase-1. *Calpastatin*, *NEDD4*, *ATXN-3*, *TIM*, *GADPH*, *pyruvate kinase*, *actin*, *aldolase*,  *$\alpha$ -enolase*, *TFAP2A*, *PPAR- $\gamma$*  and *sptan-1* are substrates of ICE as Wang et al. (1998), Harvey et al. (1998), Wellington et al. (1998), McIntire et al. (2009), Shen et al. (2010) said [175-177, 192-193]. The knockout analysis for ICE also shows *calpastatin*, *NEDD4*, *ATXN-3*, *TIM*, *GADPH*, *pyruvate kinase*, *actin*, *aldolase*,  *$\alpha$ -enolase*, *TFAP2A*, *PPAR- $\gamma$*  and *sptan-1* are affected from caspase-1 knockout.

On the other hand; *pro-IL-1 $\alpha$* , *FGF-2* and *SREPBs* are not substrates of ICE but their functioning depend on caspase-1 as found from knockout analysis of the ICE pathway. Keller et al. (2008) and Yu et al. (2008) also stated *FGF-2*, *pro-IL-1 $\alpha$*  and *SREPBs* are dependent on caspase-1 activity [25, 169].

#### **4. Consequences of NF- $\kappa$ B knockout:**

NF- $\kappa$ B knockout is effective on PYPAF4 (NACHT, LRR and PYD domains-containing protein 4) as clear from Table 4.11. Kinoshita et al. stated (2005) that PYPAF4 deactivates NF- $\kappa$ B activated by cytokines [118]. PYPAF4 directly interacts with IKK $\alpha$  (Inhibitor of nuclear factor kappa-B kinase subunit alpha) for inhibition of NF- $\kappa$ B. In our analysis, NF- $\kappa$ B knockout defects PYPAF4, whose function depends on NF- $\kappa$ B as Fiorentino et al. found [338].

#### **5. Consequences of Pro-IL-33 knockout:**

The knockout of pro-IL-33 affects its active form, IL-33 as found from Table 4.11. When pro-IL-33 is knocked out, IL-33 is the most affected protein in the ICE pathway. IL-33 is simply the active form of pro-IL-33 as Dinerello et al. said [1]. Knockout analysis also shows the relation between pro-IL-33 and IL-33.

#### **6. Consequences of TLR4 knockout:**

MD2 (myeloid differentiation protein) is defected from the knockout of TLR4 as seen in Table 4.11. TLR4 requires MD2 for sensing cellular endotoxin levels, which are needed for inflammatory responses as Prohinar et al. reported [339]. Certain need of MD2 is also observed from the knockout analysis for TLR4.

When LPS is recognized, LPS-binding protein and CD14 (Monocyte differentiation antigen CD14) are triggered for transporting LPS. LPS-binding protein and CD14 are accessory proteins that transport LPS to TLR4 and MD2. TLR4 forms a multimer complex with MD2 for LPS recognition as Park et al. claimed [340]. It cannot directly bind onto triggers, it needs MD2.

#### **7. Consequences of TLR2 knockout:**

Our knockout analysis clearly demonstrates TLR2 knockout mostly affects TLR1 and TLR6. Shenk et al. (2009) stated that TLR2 can recognize diacylated and triacylated protein with either attaching with TLR1 or TLR6. In other words, TLR2 can form complex

with TLR1 or TLR6. In both structures, TLR2-TLR6 and TLR2-TLR1, there are some differences. Fatty acid moieties interact with each TLR structure in the TLR2-TLR1 complex. However; TLR2-TLR6 complex needs hydrogen bonds between glycerol and the peptide backbone of the ligand and the LRR11 loops of TLRs [341]. That is why two different complexes of TLR2 sense different ligands. For instance, human cytomegalovirus and modified vaccinia virus Ankara (MVA) are related with TLR2-TLR1 and TLR2-TLR6 complexes; respectively. On the other hand, hepatitis C is mediated by both complexes [342].

#### **8. Consequences of TRIF knockout:**

As found from Figure 4.20, knockout analysis on the ICE pathway shows TRIF (TIR domain-containing adapter molecule) knockout affects TLR3 most importantly.

Han et al. (2010) said that TLR3-mediated signaling depends on TRIF, which interacts TLR3 by TIR (The Toll/Interleukin-1 receptor) domains [343]. Additionally, Yamamoto et al. (2003) studied TRIF knockout mice. TRIF knockout was determined to have defects in IFN- $\beta$  (Interferon beta) and IRF3 (Interferon regulatory factor 3) activations, which require TLR3 dependent path stimulated by dsRNA and/or poly(I:C) triggers [344]. Consequently, experimental result of Yamamoto et al. (2003) and our knockout analysis for TRIF-TLR3 relation agree with each other.

#### **9. Consequences of MyD88 knockout:**

MyD88 knockout mostly affected TLR7 as well as TLR9. MyD88 is an adaptor protein for TLR7 and TLR9. The relation of TLR7 and TLR9 with MyD88 is also stated in Kawai et al.'s study. Kawai et al. (2010) said TLR7-MyD88 complex is activated when ssRNA is sensed. For TLR9-MyD88 complex, bacterial DNA is a trigger. Both TLR7-MyD88 and TLR9-MyD88 complexes have roles in NF- $\kappa$ B and IRF7 (Interferon regulatory factor 7) activations. Activation of NF- $\kappa$ B and IRF7 will produce inflammatory cytokines and type I IFN; respectively [124].

## Chapter 5

### 5. CONCLUSION

The computational studies are done for caspase-1 related to peptide based drug design, conformational factors in its peptides and knockout analysis on its pathway.

Initially determination of docking place is done for caspase-1 by GNM and VitAL. Cys285 is determined to be suitable for docking studies as was claimed in the literature [9, 12]. Search for peptide based drugs for caspase-1 is also performed with comparison of drugs from the literature. ICE exists in the cytoplasm of cell. Therefore, small molecules (<500 Da) have to be used to reach its target. Moreover, D-enantiomer peptides have not been considered for ICE. Asp residue based drugs are always considered. Bicyclic and ketone structures, based on Asp and D-enantiomeric residues are obtained as good inhibitors in accordance with previous experimental works of Brady et al. and Karanewsky et al. [236, 249]. Besides, D-enantiomeric based drugs have better docking results than Asp residue. Additionally, tripeptide sequence prediction is analyzed by genetic algorithm (GA). The tripeptide X-Trp-Gly, where X can be any amino acid, converges in GA. What is more, there are found common drugs for both caspase-1 and procaspase-1. As a result, multi-target inhibition of caspase-1 functions can be done.

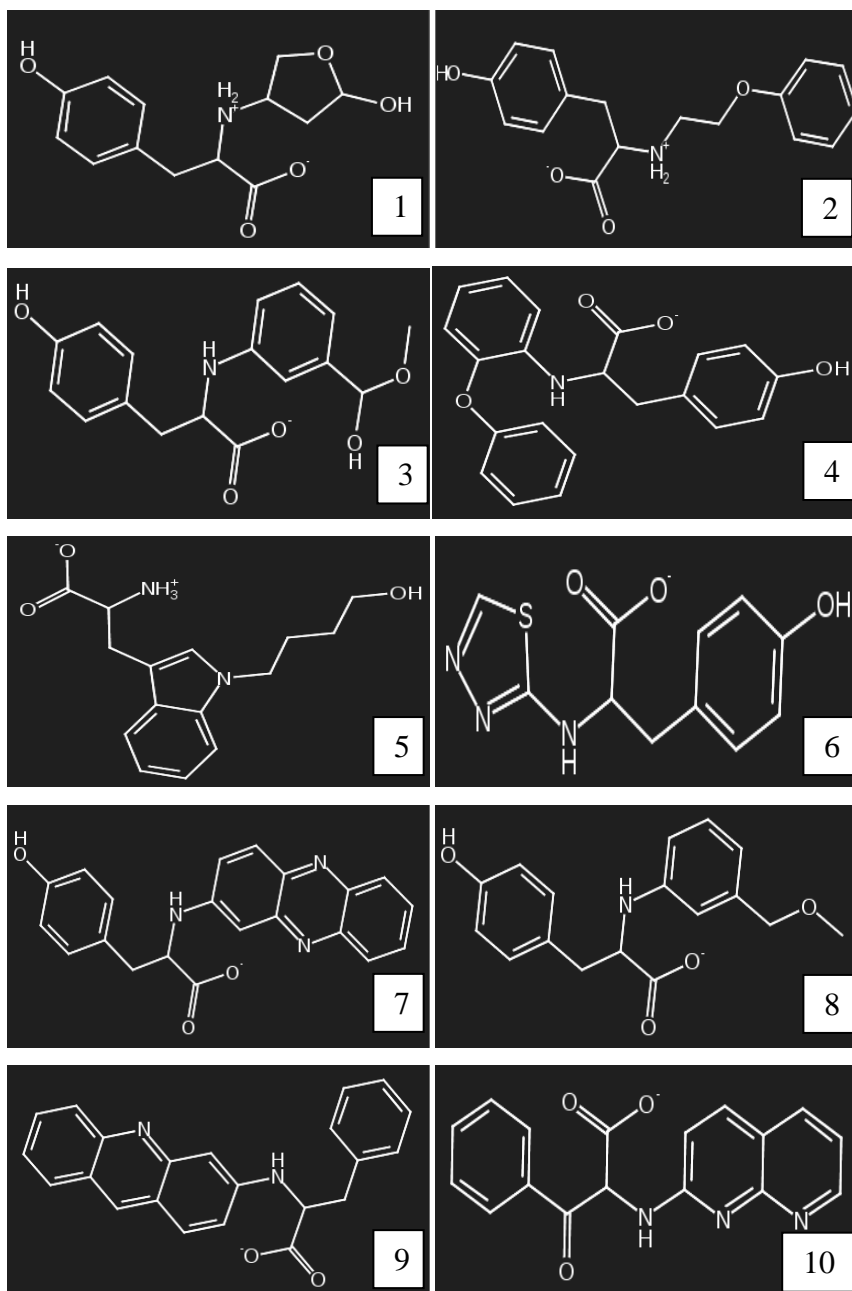
A more detailed analysis of tripeptides found from GA is done by conformational factors with VitAL. The paths followed by tripeptides in the Ramachandran Map are determined with energy behaviors. Consequently, Tyr-Trp-Gly is found to be the most appropriate tripeptide sequence in terms of conformational factors of its minimized and



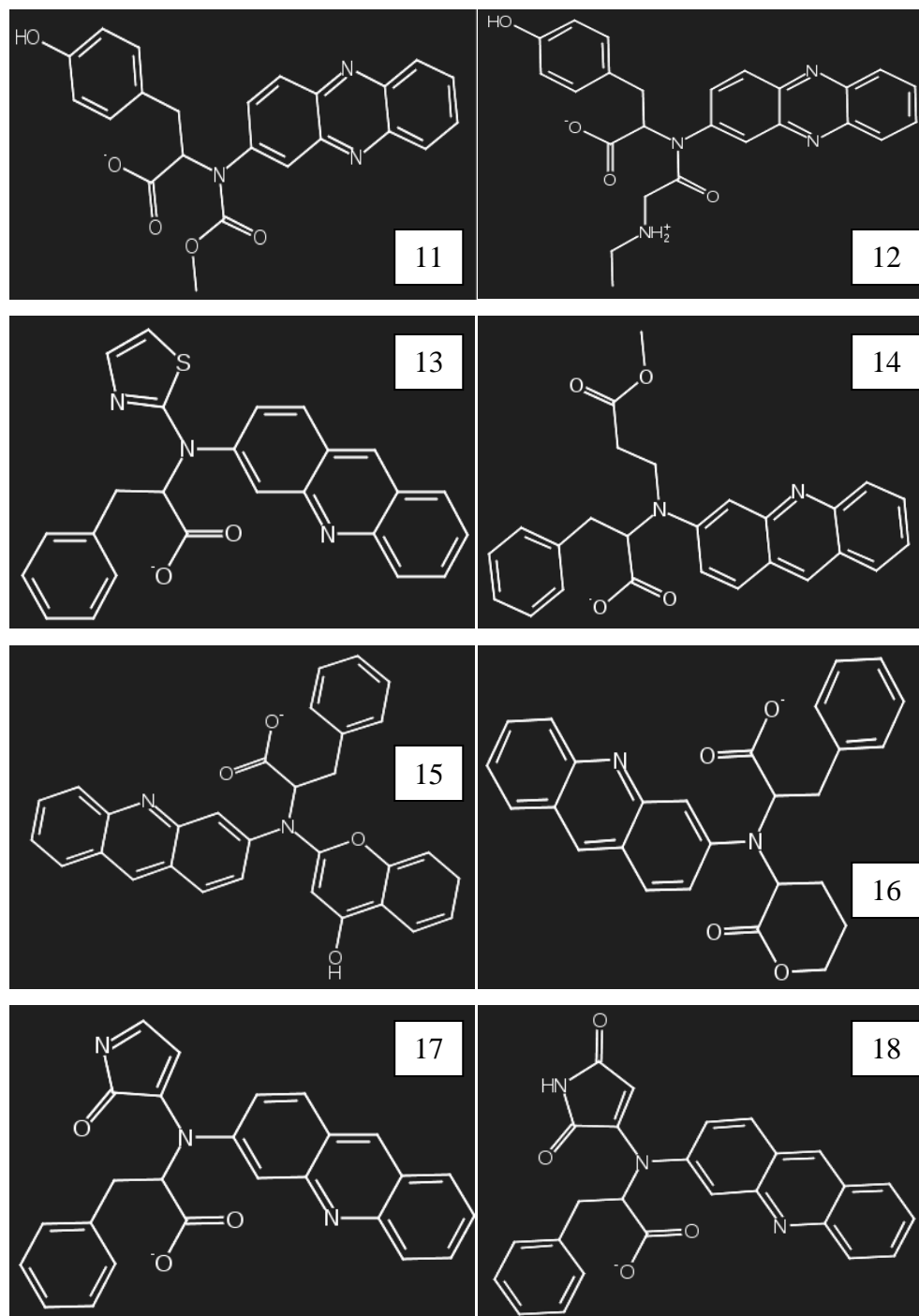
bound structures. Asp-Trp-Gly is also found to be suitable. On the other hand, Ala-Trp-Gly and Gly-Trp-Gly are not suitable due to their high energy changes during changing state. The tripeptides determined from complete enumeration are also analyzed by VitAL. Ala-Trp-Gly and Asp-Trp-Gly are also determined from complete enumeration. However; there are other found peptides which are different from the results of GA. Besides, conformational factors are analyzed by VitAL for these different tripeptides. VitAL demonstrate these different tripeptides are not suitable for ICE inhibition conformationally. This novel method uses VitAL and determines the ability of peptide to change its conformation from minimization to binding.

Knockout of the proteins in the ICE pathway shows affected proteins in the system. Knockout of NLRP3, ASC, caspase-1, NF- $\kappa$ B, pro-IL-33, TLR4, TLR2, TRIF and MyD88 are crucial in the ICE pathway.

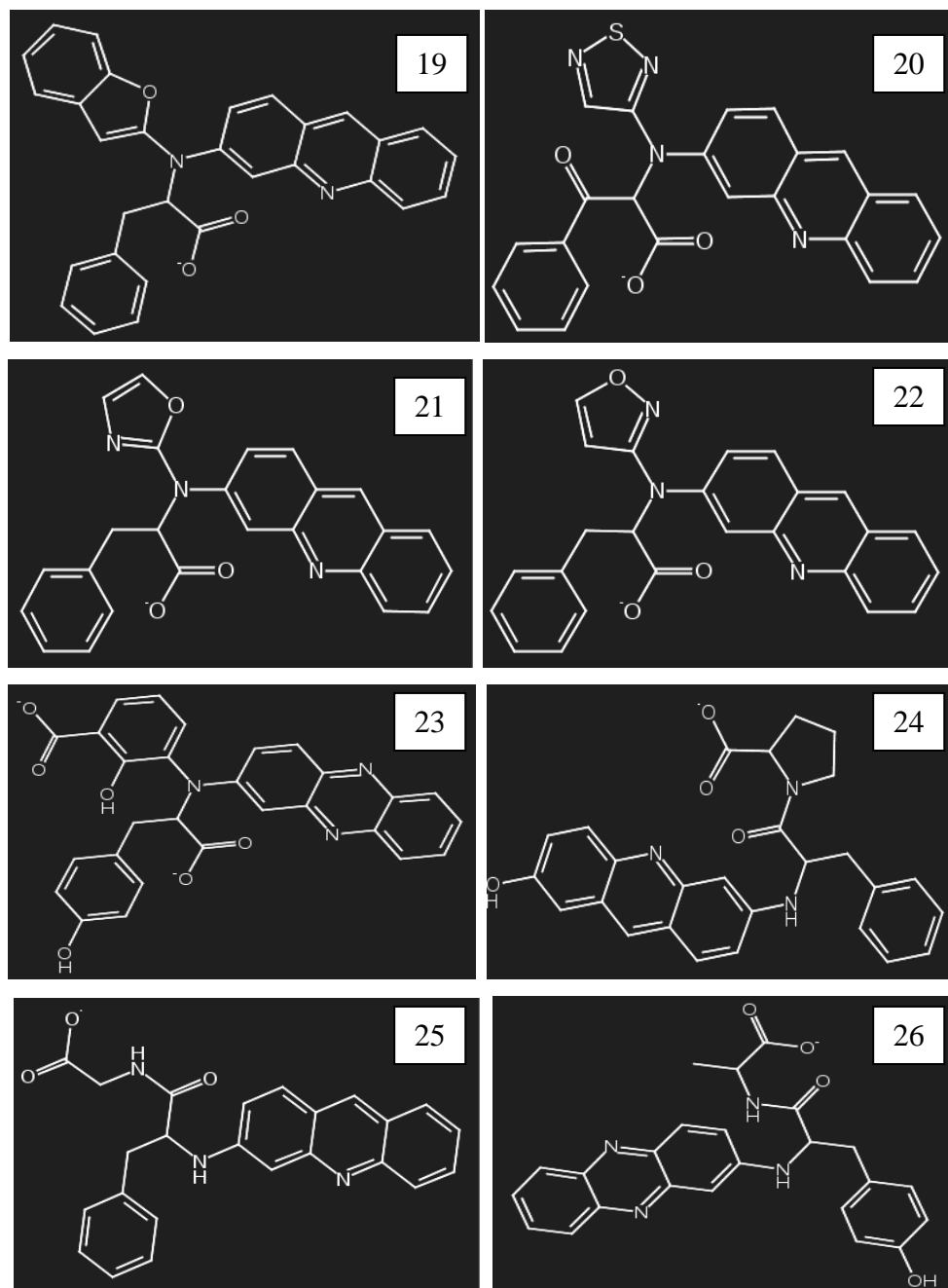
## 6. APPENDIX-I



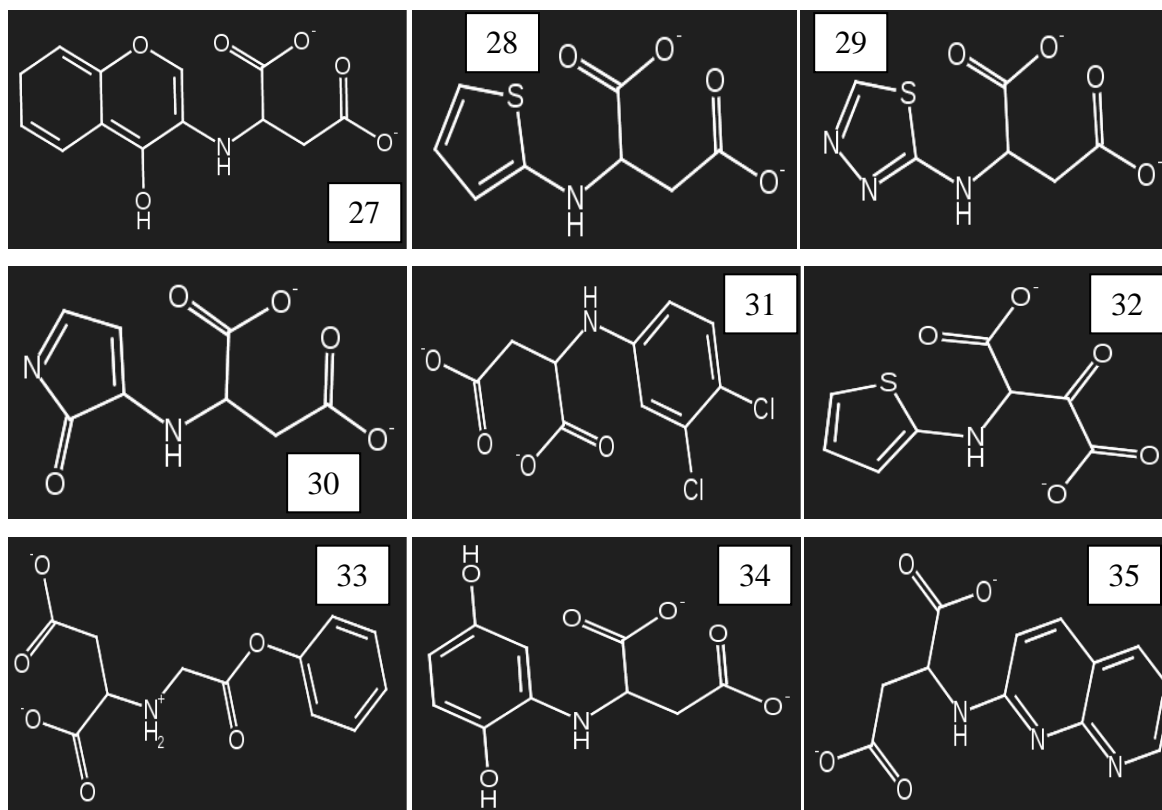
**Figure 6.1** The molecular structures of D-aminoacid+ 1 x iLib Diverse.



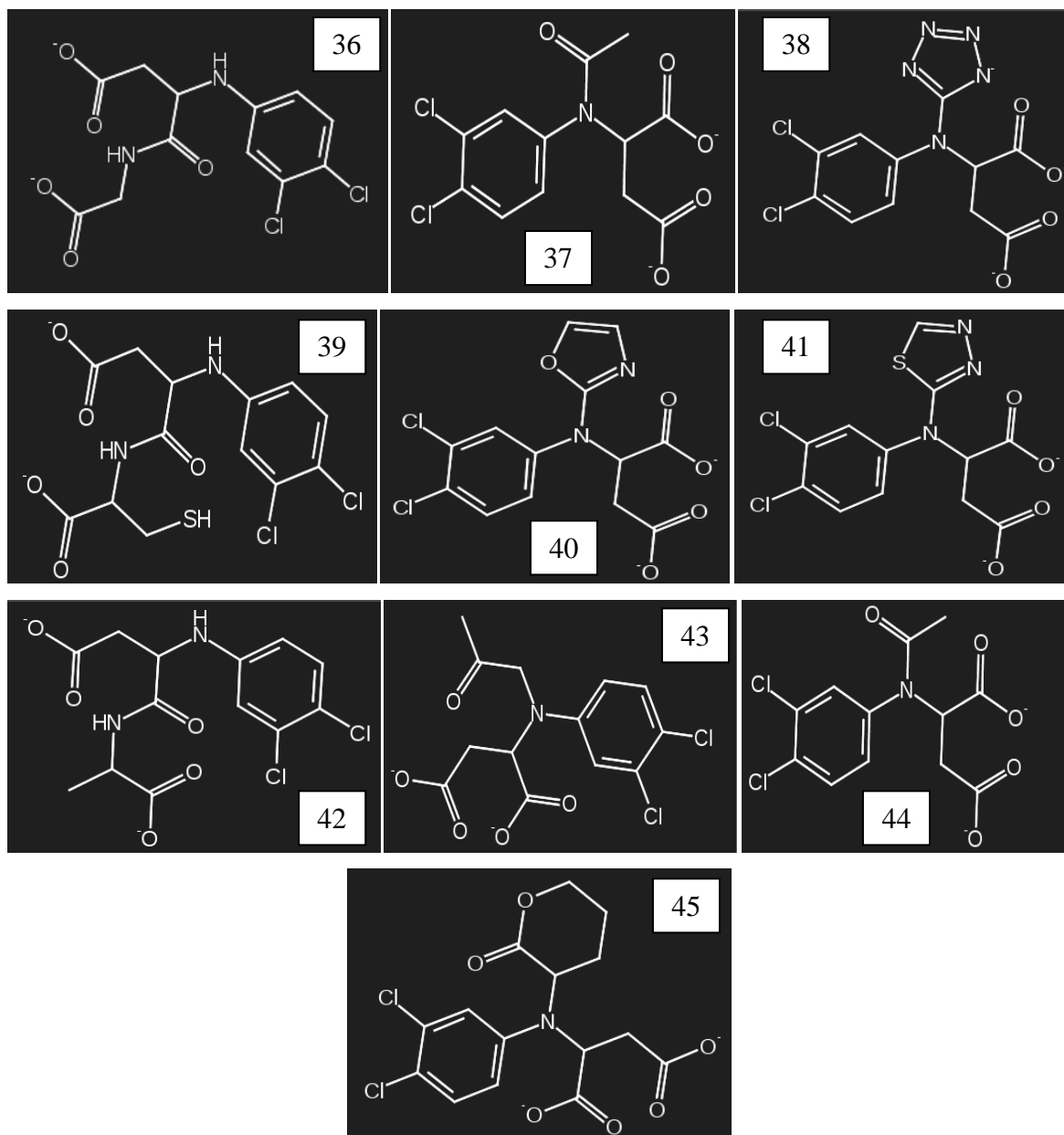
**Figure 6.2** The molecular structures of D-aminoacid+ 2 x iLib Diverse.



**Figure 6.3** The molecular structures of D-aminoacid+ 2 x iLib Diverse.



**Figure 6.4** The molecular structures of 9 molecules from Asp/Malonate + 1 x iLib Diverse.



**Figure 6.5** The molecular structures of 10 molecules from Asp/Malonate + 2 x iLib Diverse.

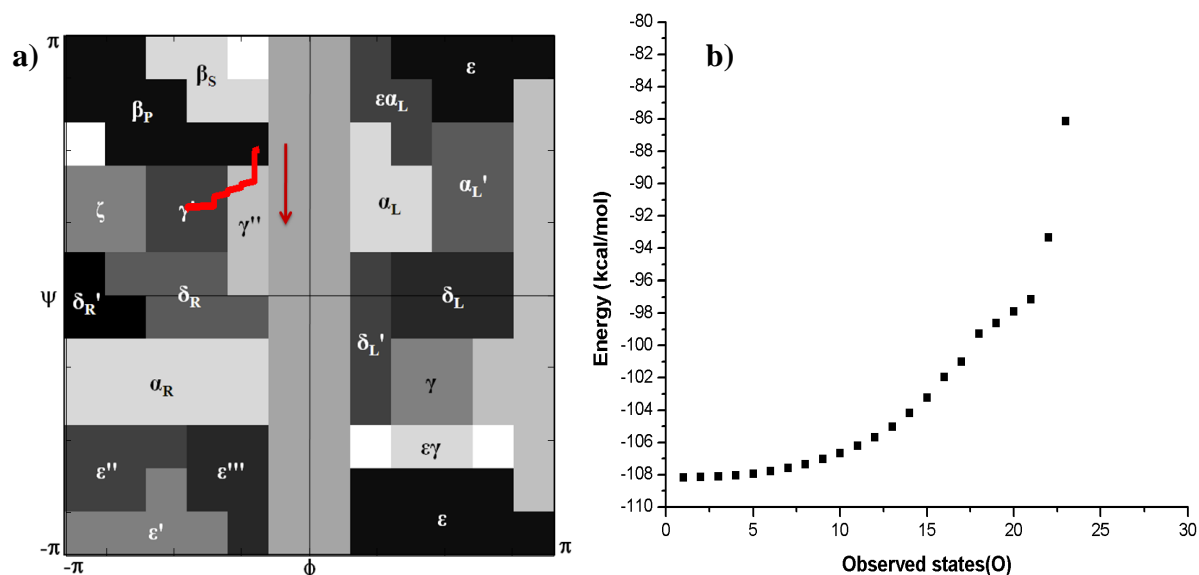
## 7. APPENDIX-II

Determination of minimized and bound energies given in Table 7.1 shows that AWE, DFF, DYF and SFF pass through high energies during state change. Furthermore, there are two residues having different states in DFF, DYF as well as SFF and these residues make large energy difference during state change.

**Table 7.1** Energy change and probabilities of residues having different states.

Peptides	Mismatch Residue	State A	State B	Minimized Energy*	Bound Energy*	$\Delta E^*$	$P_B/P_A$
AWE	E3	11	8	-108.17	-86.11	22.06	$2.42 \times 10^{-6}$
DFF	D1	10	8	-86.89	-59.01	27.88	$1.84 \times 10^{-20}$
	F3	8	12	-97.21	-84.94	12.27	$2.06 \times 10^{-9}$
DYF	D1	11	8	-91.24	-80.18	11.06	$1.48 \times 10^{-8}$
	Y2	8	5	-45.39	-39.055	6.335	$3.28 \times 10^{-5}$
SFF	S1	5	3	-50.3	-43.14	7.16	$8.54 \times 10^{-6}$
	F3	8	3	-47.33	-31.17	16.16	$3.63 \times 10^{-12}$

As one of the unfavorable peptides, AWE goes from 8-8-11 to 8-8-8. Its 3<sup>rd</sup> residue is the only one that changes its state as seen in Figure 7.1a. Although having one aminoacid having different states, it encounters high energy change while going from minimization to binding as given in Figure 7.1b. Especially in the last observed states, a dramatic increase is observed for the Glu residue in AWE in spite of having a smooth exponential rise at initial observed states.

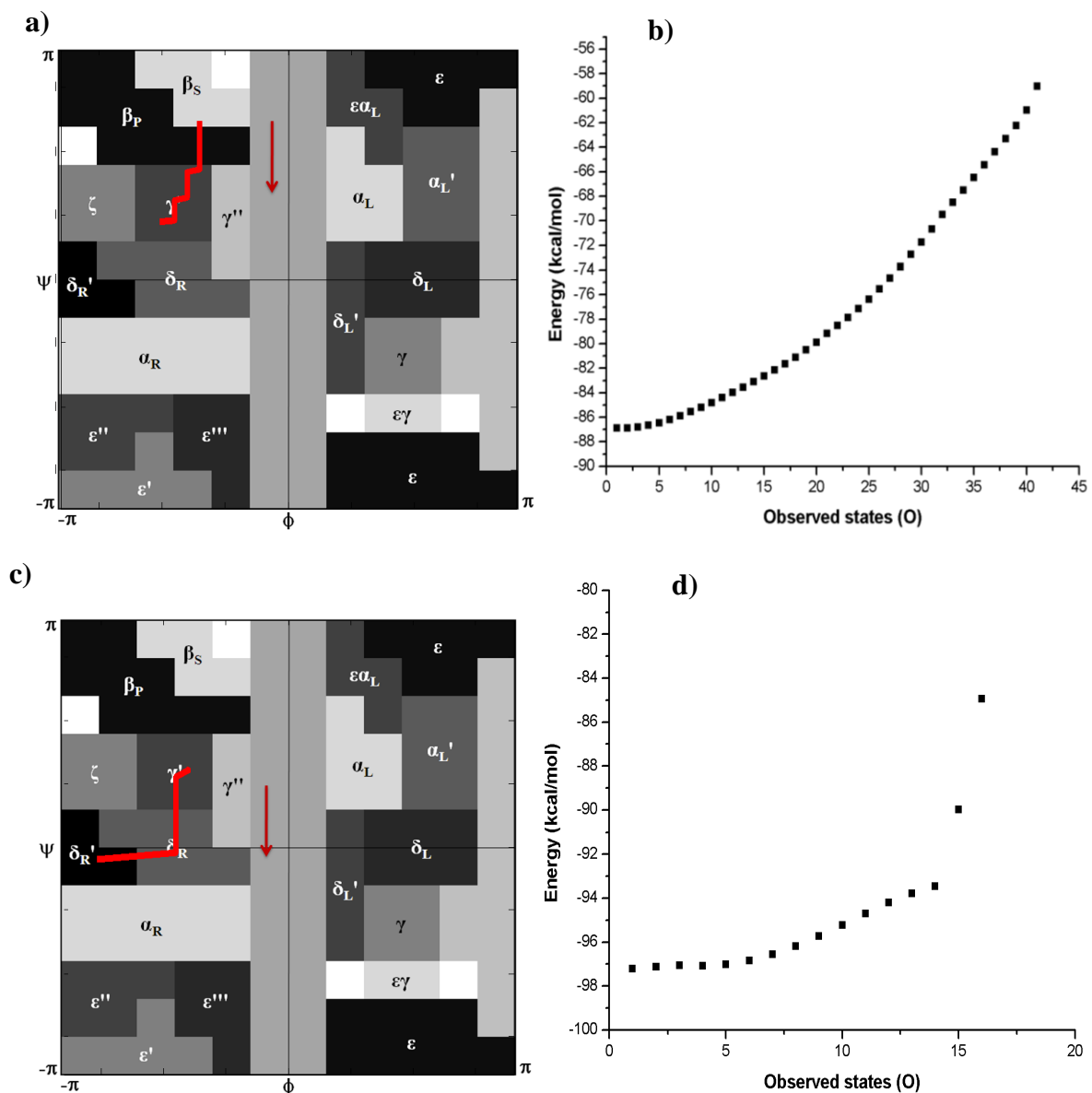


**Figure 7.1** VitAL results for AWE

**a)** Path followed by Glu in AWE in Ramachandran plot. Red arrow shows the direction of path. **b)** Energy calculated by HyperChem versus each observed states by Glu in AWE.

DFF is the worst scored peptide due to the lowest probabilities of its aminoacids. These low probabilities come from high energy changes. For example, its first residue Asp has an increasing energy behavior during state change from 10<sup>th</sup> to 8<sup>th</sup> states as given in Figure 7.2b. Its 3<sup>rd</sup> residue Phe also demonstrates a sharp increase at the last three observed states in Figure 7.2d while going from 8 to 12. The paths followed by residues having different state are shown in Figure 7.2a and Figure 7.2c.

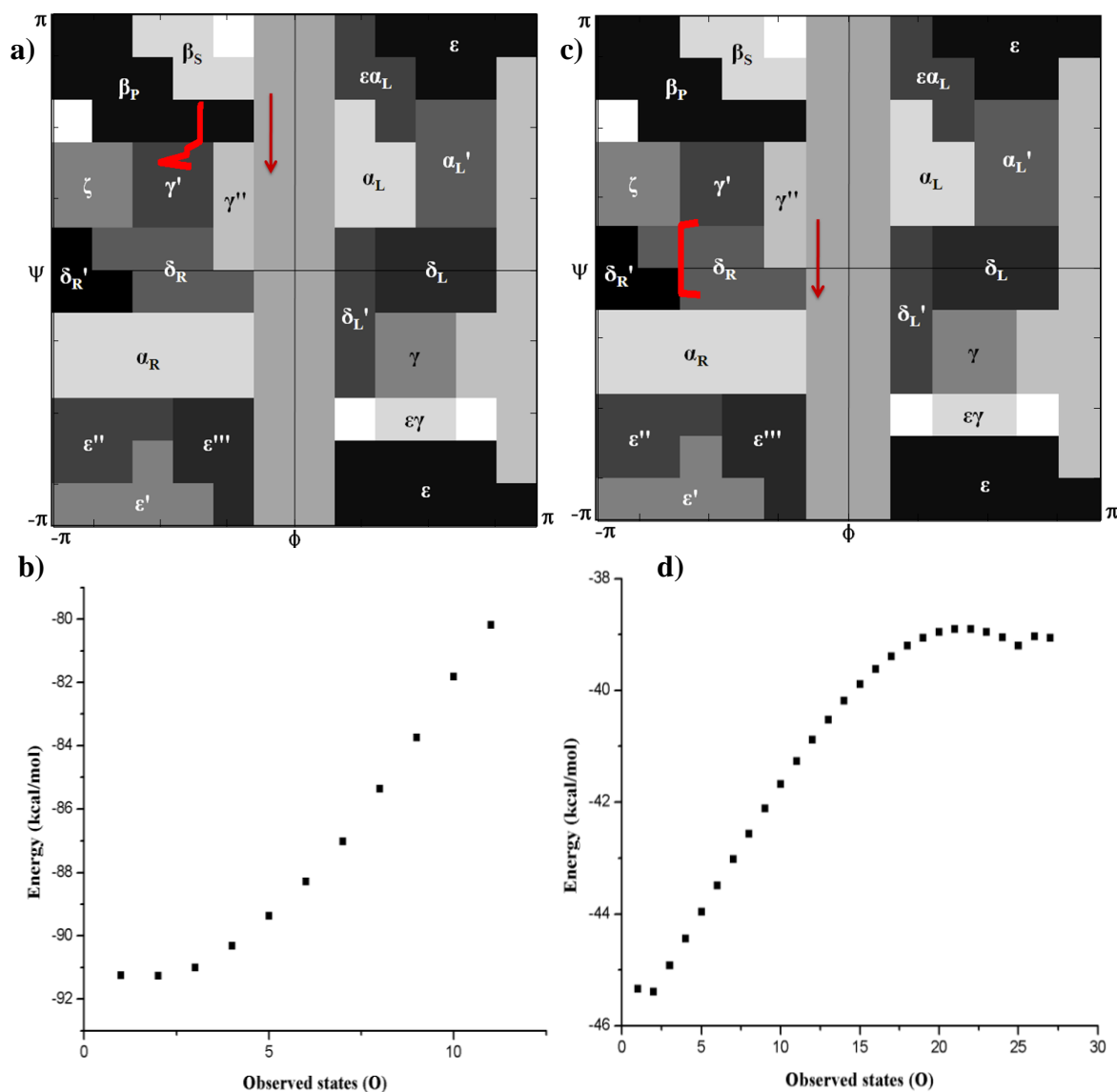




**Figure 7.2** VitAL results for DFF

**a)** Path followed by Asp in DFF in Ramachandran plot. Red arrow shows the direction of path. **b)** Energy calculated by HyperChem versus each observed states by Asp. **c)** Path followed by Phe in DFF in Ramachandran plot. Red arrow shows the direction of path. **d)** Energy calculated by HyperChem versus each observed states by Phe.

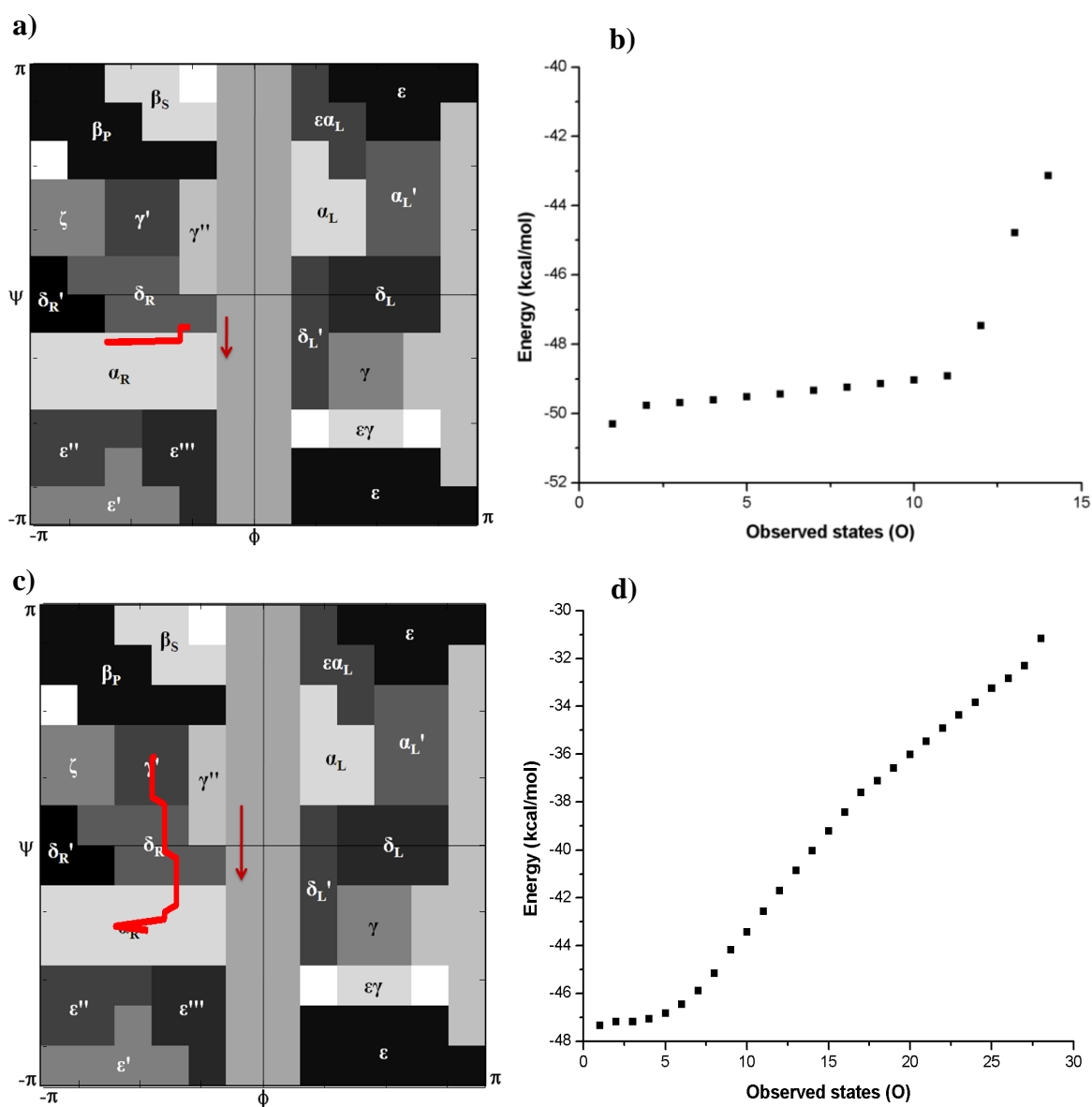
DYF has two residues having different state. 1<sup>st</sup> residue changes its state from 11<sup>th</sup> (minimized) to 8<sup>th</sup> (bound) and 2<sup>nd</sup> aminoacid passes from 8<sup>th</sup> (minimized) to 5<sup>th</sup> (bound). 1<sup>st</sup> aminoacid (Asp) makes a change while passing from 8<sup>th</sup> to 11<sup>th</sup> state as seen in Figure 7.3a and Figure 7.3c. Energy behavior of Asp makes an exponential growth up to last conformation Figure 7.3b. Tyr makes a sudden decrease in energy. It rises up to 22<sup>nd</sup> observed state and forms a peak. There is also a small fluctuation in energy at the last three observed states of Tyr as seen in Figure 7.3d.



**Figure 7.3** VitAL results for DYF.

**a)** Path followed by Asp in DYF in Ramachandran plot. Red arrow shows the direction of path. **b)** Energy calculated by HyperChem versus each observed states by Asp. **c)** Path followed by Tyr in DYF in Ramachandran plot. Red arrow shows the direction of path. **d)** Energy calculated by HyperChem versus each observed states by Tyr.

SFF has two residues having different states (1<sup>st</sup> and 3<sup>rd</sup> aminoacids which go from 5 to 3 and 8 to 3; respectively.) and the paths followed by these residues are given in Figure 7.4a and Figure 7.4c. 1<sup>st</sup> residue has an energy behavior in which a sharp increase happens against smooth rise at initial observed states in Figure 7.4b. What is more, its 3<sup>rd</sup> residue also has high energy difference between minimization and binding as clear from Figure 7.4d which demonstrates SFF is not a good choice for ICE inhibition.



**Figure 7.4** VitAL results for SFF.

**a)** Path followed by Ser in SFF in Ramachandran plot. Red arrow shows the direction of path. **b)** Energy calculated by HyperChem versus each observed states by Ser. **c)** Path followed by Phe in DFF in Ramachandran plot. Red arrow shows the direction of path. **d)** Energy calculated by HyperChem versus each observed states by Phe.

**BIBLIOGRAPHY**

1. Dinerello, C.A., *Immunological and inflammatory functions of the interleukin-1 family*. Annual Review of Immunology, 2009. **27**: p. 519-550.
2. Vanhee, P., et al., *Computational design of peptide ligands*. Trends Biotechnol, 2011. **29**(5): p. 231-9.
3. Simon, A. and J.W. van der Meer, *Pathogenesis of familial periodic fever syndromes or hereditary autoinflammatory syndromes*. Am J Physiol Regul Integr Comp Physiol, 2007. **292**(1): p. R86-98.
4. Park, K.K., et al., *Promoting axon regeneration in the adult CNS by modulation of the PTEN/mTOR pathway*. Science, 2008. **322**(5903): p. 963-6.
5. Hayashi, M. and J.D. Lee, *Role of the BMK1/ERK5 signaling pathway: lessons from knockout mice*. J Mol Med, 2004. **82**(12): p. 800-8.
6. Harding, H.P., et al., *Ppp1r15 gene knockout reveals an essential role for translation initiation factor 2 alpha (eIF2alpha) dephosphorylation in mammalian development*. Proc Natl Acad Sci U S A, 2009. **106**(6): p. 1832-7.
7. Gustin, J.P., et al., *Knockin of mutant PIK3CA activates multiple oncogenic pathways*. Proc Natl Acad Sci U S A, 2009. **106**(8): p. 2835-40.
8. Datta, D., et al., *An allosteric circuit in caspase-1*. J Mol Biol, 2008. **381**(5): p. 1157-67.
9. Nicholson, D.W. and N.A. Thornberry, *Caspases: killer proteases*. Trends Biochem Sci, 1997. **22**(8): p. 299-306.
10. Stennicke, H.R., C.A. Ryan, and G.S. Salvesen, *Reprieval from execution: the molecular basis of caspase inhibition*. Trends Biochem Sci, 2002. **27**(2): p. 94-101.
11. Stennicke, H.R. and G.S. Salvesen, *Catalytic properties of the caspases*. Cell Death Differ, 1999. **6**(11): p. 1054-9.
12. Wei, Y., et al., *The structures of caspases-1, -3, -7 and -8 reveal the basis for substrate and inhibitor selectivity*. Chem Biol, 2000. **7**(6): p. 423-32.
13. O'Brien, T. and D. Lee, *Prospects for caspase inhibitors*. Mini Rev Med Chem, 2004. **4**(2): p. 153-65.
14. Stennicke, H.R. and G.S. Salvesen, *Properties of the caspases*. Biochim Biophys Acta, 1998. **1387**(1-2): p. 17-31.
15. Jarvelainen, H.A., A. Galmiche, and A. Zychlinsky, *Caspase-1 activation by Salmonella*. Trends Cell Biol, 2003. **13**(4): p. 204-9.
16. Shi, Y., *Mechanisms of caspase activation and inhibition during apoptosis*. Mol Cell, 2002. **9**(3): p. 459-70.
17. Scheer, J.M., M.J. Romanowski, and J.A. Wells, *A common allosteric site and mechanism in caspases*. Proc Natl Acad Sci U S A, 2006. **103**(20): p. 7595-600.
18. Fantuzzi, G. and C.A. Dinarello, *Interleukin-18 and interleukin-1 beta: two cytokine substrates for ICE (caspase-1)*. J Clin Immunol, 1999. **19**(1): p. 1-11.
19. Wang, S., et al., *Murine caspase-11, an ICE-interacting protease, is essential for the activation of ICE*. Cell, 1998. **92**(4): p. 501-9.
20. Sekiyama, A., et al., *A stress-induced, superoxide-mediated caspase-1 activation pathway causes plasma IL-18 upregulation*. Immunity, 2005. **22**(6): p. 669-77.
21. Wang, L., et al., *PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates activation of NF-kappa B and caspase-1-dependent cytokine processing*. J Biol Chem, 2002. **277**(33): p. 29874-80.
22. Romanowski, M.J., et al., *Crystal structures of a ligand-free and malonate-bound human caspase-1: implications for the mechanism of substrate binding*. Structure, 2004. **12**(8): p. 1361-71.

23. Wagner, G. and S. Laufer, *Small molecular anti-cytokine agents*. Med Res Rev, 2006. **26**(1): p. 1-62.
24. Swaan, P.W., et al., *Sequential processing of human ProIL-1beta by caspase-1 and subsequent folding determined by a combined in vitro and in silico approach*. Pharm Res, 2001. **18**(8): p. 1083-90.
25. Yu, H.B. and B.B. Finlay, *The caspase-1 inflammasome: a pilot of innate immune responses*. Cell Host Microbe, 2008. **4**(3): p. 198-208.
26. Ting, J.P., et al., *The NLR gene family: a standard nomenclature*. Immunity, 2008. **28**(3): p. 285-7.
27. Bruey, J.M., et al., *PAN1/NALP2/PYPAF2, an inducible inflammatory mediator that regulates NF-kappaB and caspase-1 activation in macrophages*. J Biol Chem, 2004. **279**(50): p. 51897-907.
28. Bryant, C. and K.A. Fitzgerald, *Molecular mechanisms involved in inflammasome activation*. Trends Cell Biol, 2009. **19**(9): p. 455-64.
29. Latz, E., *The inflammasomes: mechanisms of activation and function*. Curr Opin Immunol, 2010. **22**(1): p. 28-33.
30. Bauernfeind, F., et al., *Inflammasomes: current understanding and open questions*. Cell Mol Life Sci, 2010.
31. Parsons, M.J. and P.M. Vertino, *Dual role of TMS1/ASC in death receptor signaling*. Oncogene, 2006. **25**(52): p. 6948-58.
32. *String Database 8.3*.
33. KEGG Pathway NOD-like receptor signaling pathway, H.s.
34. Silverman, W.R., et al., *The pannexin 1 channel activates the inflammasome in neurons and astrocytes*. J Biol Chem, 2009. **284**(27): p. 18143-51.
35. Bruey, J.M., et al., *Bcl-2 and Bcl-XL regulate proinflammatory caspase-1 activation by interaction with NALP1*. Cell, 2007. **129**(1): p. 45-56.
36. Gurcel, L., et al., *Caspase-1 activation of lipid metabolic pathways in response to bacterial pore-forming toxins promotes cell survival*. Cell, 2006. **126**(6): p. 1135-45.
37. Martinon, F., K. Burns, and J. Tschopp, *The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-beta*. Mol Cell, 2002. **10**(2): p. 417-26.
38. Franchi, L., et al., *The inflammasome: a caspase-1-activation platform that regulates immune responses and disease pathogenesis*. Nat Immunol, 2009. **10**(3): p. 241-7.
39. Yu, J.W., et al., *Cryopyrin and pyrin activate caspase-1, but not NF-kappaB, via ASC oligomerization*. Cell Death Differ, 2006. **13**(2): p. 236-49.
40. Damiano, J.S., et al., *Heterotypic interactions among NACHT domains: implications for regulation of innate immune responses*. Biochem J, 2004. **381**(Pt 1): p. 213-9.
41. Mayor, A., et al., *A crucial function of SGT1 and HSP90 in inflammasome activity links mammalian and plant innate immune responses*. Nat Immunol, 2007. **8**(5): p. 497-503.
42. Petrilli, V., et al., *The inflammasome: a danger sensing complex triggering innate immunity*. Curr Opin Immunol, 2007. **19**(6): p. 615-22.
43. Creagh, E.M., H. Conroy, and S.J. Martin, *Caspase-activation pathways in apoptosis and immunity*. Immunol Rev, 2003. **193**: p. 10-21.
44. Inohara, et al., *NOD-LRR proteins: role in host-microbial interactions and inflammatory disease*. Annu Rev Biochem, 2005. **74**: p. 355-83.
45. Lamkanfi, M., et al., *Caspase-1 inflammasomes in infection and inflammation*. J Leukoc Biol, 2007. **82**(2): p. 220-5.
46. Razmara, M., et al., *CARD-8 protein, a new CARD family member that regulates caspase-1 activation and apoptosis*. J Biol Chem, 2002. **277**(16): p. 13952-8.
47. Yamamoto, M., et al., *ASC is essential for LPS-induced activation of procaspase-1 independently of TLR-associated signal adaptor molecules*. Genes Cells, 2004. **9**(11): p. 1055-67.

48. Kang, S.J., et al., *Dual role of caspase-11 in mediating activation of caspase-1 and caspase-3 under pathological conditions*. J Cell Biol, 2000. **149**(3): p. 613-22.
49. Lin, X.Y., M.S. Choi, and A.G. Porter, *Expression analysis of the human caspase-1 subfamily reveals specific regulation of the CASP5 gene by lipopolysaccharide and interferon-gamma*. J Biol Chem, 2000. **275**(51): p. 39920-6.
50. Vancompernelle, K., et al., *Atractyloside-induced release of cathepsin B, a protease with caspase-processing activity*. FEBS Lett, 1998. **438**(3): p. 150-8.
51. Martinon, F. and J. Tschopp, *Inflammatory caspases: linking an intracellular innate immune system to autoinflammatory diseases*. Cell, 2004. **117**(5): p. 561-74.
52. Kahlenberg, J.M., et al., *Potentiation of caspase-1 activation by the P2X7 receptor is dependent on TLR signals and requires NF-kappaB-driven protein synthesis*. J Immunol, 2005. **175**(11): p. 7611-22.
53. Ogura, Y., F.S. Sutterwala, and R.A. Flavell, *The inflammasome: first line of the immune response to cell stress*. Cell, 2006. **126**(4): p. 659-62.
54. Lamkanfi, M., et al., *INCA, a novel human caspase recruitment domain protein that inhibits interleukin-1beta generation*. J Biol Chem, 2004. **279**(50): p. 51729-38.
55. Warren, S.E., et al., *Multiple Nod-like receptors activate caspase 1 during Listeria monocytogenes infection*. J Immunol, 2008. **180**(11): p. 7558-64.
56. Wu, J., T. Fernandes-Alnemri, and E.S. Alnemri, *Involvement of the AIM2, NLRC4, and NLRP3 inflammasomes in caspase-1 activation by Listeria monocytogenes*. J Clin Immunol, 2010. **30**(5): p. 693-702.
57. Hsu, L.C., et al., *A NOD2-NALP1 complex mediates caspase-1-dependent IL-1beta secretion in response to Bacillus anthracis infection and muramyl dipeptide*. Proc Natl Acad Sci U S A, 2008. **105**(22): p. 7803-8.
58. Pedra, J.H., S.L. Cassel, and F.S. Sutterwala, *Sensing pathogens and danger signals by the inflammasome*. Curr Opin Immunol, 2009. **21**(1): p. 10-6.
59. Strober, W., et al., *Signalling pathways and molecular interactions of NOD1 and NOD2*. Nat Rev Immunol, 2006. **6**(1): p. 9-20.
60. Adams, J.M. and S. Cory, *Apoptosomes: engines for caspase activation*. Curr Opin Cell Biol, 2002. **14**(6): p. 715-20.
61. Marina-Garcia, N., et al., *Pannexin-1-mediated intracellular delivery of muramyl dipeptide induces caspase-1 activation via cryopyrin/NLRP3 independently of Nod2*. J Immunol, 2008. **180**(6): p. 4050-7.
62. Sutterwala, F.S., et al., *Critical role for NALP3/CIAS1/Cryopyrin in innate and adaptive immunity through its regulation of caspase-1*. Immunity, 2006. **24**(3): p. 317-27.
63. Pan, Q., et al., *MDP-induced interleukin-1beta processing requires Nod2 and CIAS1/NALP3*. J Leukoc Biol, 2007. **82**(1): p. 177-83.
64. Uematsu, S. and S. Akira, *Toll-like receptors and innate immunity*. J Mol Med, 2006. **84**(9): p. 712-25.
65. Faustin, B., et al., *Mechanism of Bcl-2 and Bcl-X(L) inhibition of NLRP1 inflammasome: loop domain-dependent suppression of ATP binding and oligomerization*. Proc Natl Acad Sci U S A, 2009. **106**(10): p. 3935-40.
66. Faustin, B., et al., *Reconstituted NALP1 inflammasome reveals two-step mechanism of caspase-1 activation*. Mol Cell, 2007. **25**(5): p. 713-24.
67. Ferrari, D., et al., *The P2X7 receptor: a key player in IL-1 processing and release*. J Immunol, 2006. **176**(7): p. 3877-83.



68. Takenouchi, T., et al., *Lysophospholipids and ATP mutually suppress maturation and release of IL-1 beta in mouse microglial cells using a Rho-dependent pathway*. J Immunol, 2008. **180**(12): p. 7827-39.
69. Petrilli, V., et al., *Activation of the NALP3 inflammasome is triggered by low intracellular potassium concentration*. Cell Death Differ, 2007. **14**(9): p. 1583-9.
70. Miao, E.A., et al., *TLR5 and Ipaf: dual sensors of bacterial flagellin in the innate immune system*. Semin Immunopathol, 2007. **29**(3): p. 275-88.
71. Wickliffe, K.E., S.H. Leppla, and M. Moayeri, *Anthrax lethal toxin-induced inflammasome formation and caspase-1 activation are late events dependent on ion fluxes and the proteasome*. Cell Microbiol, 2008. **10**(2): p. 332-43.
72. Pelegrin, P. and A. Surprenant, *The P2X(7) receptor-pannexin connection to dye uptake and IL-1beta release*. Purinergic Signal, 2009. **5**(2): p. 129-37.
73. Qu, Y. and G.R. Dubyak, *P2X7 receptors regulate multiple types of membrane trafficking responses and non-classical secretion pathways*. Purinergic Signal, 2009. **5**(2): p. 163-73.
74. Andrei, C., et al., *Phospholipases C and A2 control lysosome-mediated IL-1 beta secretion: Implications for inflammatory processes*. Proc Natl Acad Sci U S A, 2004. **101**(26): p. 9745-50.
75. Verhoef, P.A., et al., *Maitotoxin induces biphasic interleukin-1beta secretion and membrane blebbing in murine macrophages*. Mol Pharmacol, 2004. **66**(4): p. 909-20.
76. Kersse, K., et al., *A phylogenetic and functional overview of inflammatory caspases and caspase-1-related CARD-only proteins*. Biochem Soc Trans, 2007. **35**(Pt 6): p. 1508-11.
77. Stehlik, C. and A. Dorfleutner, *COPs and POPs: modulators of inflammasome activity*. J Immunol, 2007. **179**(12): p. 7993-8.
78. Papin, S., et al., *The SPRY domain of Pypin, mutated in familial Mediterranean fever patients, interacts with inflammasome components and inhibits proIL-1beta processing*. Cell Death Differ, 2007. **14**(8): p. 1457-66.
79. Richards, N., et al., *Interaction between pypin and the apoptotic speck protein (ASC) modulates ASC-induced apoptosis*. J Biol Chem, 2001. **276**(42): p. 39320-9.
80. Lidar, M. and A. Livneh, *Familial Mediterranean fever: clinical, molecular and management advancements*. Neth J Med, 2007. **65**(9): p. 318-24.
81. Chae, J.J., et al., *The B30.2 domain of pypin, the familial Mediterranean fever protein, interacts directly with caspase-1 to modulate IL-1beta production*. Proc Natl Acad Sci U S A, 2006. **103**(26): p. 9982-7.
82. Balci-Peynircioglu, B., et al., *Pypin, product of the MEFV locus, interacts with the proapoptotic protein, Siva*. J Cell Physiol, 2008. **216**(3): p. 595-602.
83. Chae, J.J., et al., *The familial Mediterranean fever protein, pypin, is cleaved by caspase-1 and activates NF-kappaB through its N-terminal fragment*. Blood, 2008. **112**(5): p. 1794-803.
84. Martinon, F., A. Mayor, and J. Tschopp, *The inflammasomes: guardians of the body*. Annu Rev Immunol, 2009. **27**: p. 229-65.
85. Shoham, N.G., et al., *Pypin binds the PSTPIP1/CD2BP1 protein, defining familial Mediterranean fever and PAPA syndrome as disorders in the same pathway*. Proc Natl Acad Sci U S A, 2003. **100**(23): p. 13501-6.
86. Kuijk, L.M., et al., *HMG-CoA reductase inhibition induces IL-1beta release through Rac1/PI3K/PKB-dependent caspase-1 activation*. Blood, 2008. **112**(9): p. 3563-73.
87. Sutterwala, F.S., et al., *NALP3: a key player in caspase-1 activation*. J Endotoxin Res, 2006. **12**(4): p. 251-6.
88. Meylan, E., J. Tschopp, and M. Karin, *Intracellular pattern recognition receptors in the host response*. Nature, 2006. **442**(7098): p. 39-44.

89. Martinon, F., et al., *Gout-associated uric acid crystals activate the NALP3 inflammasome*. Nature, 2006. **440**(7081): p. 237-41.
90. Martinon, F., et al., *Identification of bacterial muramyl dipeptide as activator of the NALP3/cryopyrin inflammasome*. Curr Biol, 2004. **14**(21): p. 1929-34.
91. Fantuzzi, G., D.A. Reed, and C.A. Dinarello, *IL-12-induced IFN-gamma is dependent on caspase-1 processing of the IL-18 precursor*. J Clin Invest, 1999. **104**(6): p. 761-7.
92. Shirasu, K., *The HSP90-SGT1 chaperone complex for NLR immune sensors*. Annu Rev Plant Biol, 2009. **60**: p. 139-64.
93. Martinon, F., *Detection of immune danger signals by NALP3*. J Leukoc Biol, 2008. **83**(3): p. 507-11.
94. Ye, Z. and J.P. Ting, *NLR, the nucleotide-binding domain leucine-rich repeat containing gene family*. Curr Opin Immunol, 2008. **20**(1): p. 3-9.
95. Babelova, A., et al., *Biglycan, a danger signal that activates the NLRP3 inflammasome via toll-like and P2X receptors*. J Biol Chem, 2009. **284**(36): p. 24035-48.
96. Schroder, K. and J. Tschopp, *The inflammasomes*. Cell, 2010. **140**(6): p. 821-32.
97. Sutterwala, F.S. and R.A. Flavell, *NLR4/IPAF: a CARD carrying member of the NLR family*. Clin Immunol, 2009. **130**(1): p. 2-6.
98. Takeda, K., T. Kaisho, and S. Akira, *Toll-like receptors*. Annu Rev Immunol, 2003. **21**: p. 335-76.
99. Willingham, S.B., et al., *Microbial pathogen-induced necrotic cell death mediated by the inflammasome components CIAS1/cryopyrin/NLRP3 and ASC*. Cell Host Microbe, 2007. **2**(3): p. 147-59.
100. Dostert, C., et al., *Innate immune activation through Nalp3 inflammasome sensing of asbestos and silica*. Science, 2008. **320**(5876): p. 674-7.
101. Wang, X., et al., *Protective role of Cop in Rip2/caspase-1/caspase-4-mediated HeLa cell death*. Biochim Biophys Acta, 2006. **1762**(8): p. 742-54.
102. Franchi, L., T. Eigenbrod, and G. Nunez, *Cutting edge: TNF-alpha mediates sensitization to ATP and silica via the NLRP3 inflammasome in the absence of microbial stimulation*. J Immunol, 2009. **183**(2): p. 792-6.
103. Franchi, L., et al., *Inflammasomes as microbial sensors*. Eur J Immunol, 2010. **40**(3): p. 611-5.
104. Kanneganti, T.D., et al., *Critical role for Cryopyrin/Nalp3 in activation of caspase-1 in response to viral infection and double-stranded RNA*. J Biol Chem, 2006. **281**(48): p. 36560-8.
105. Koo, I.C., et al., *ESX-1-dependent cytolysis in lysosome secretion and inflammasome activation during mycobacterial infection*. Cell Microbiol, 2008. **10**(9): p. 1866-78.
106. Cruz, C.M., et al., *ATP activates a reactive oxygen species-dependent oxidative stress response and secretion of proinflammatory cytokines in macrophages*. J Biol Chem, 2007. **282**(5): p. 2871-9.
107. Lee, S.H., C. Stehlik, and J.C. Reed, *Cop, a caspase recruitment domain-containing protein and inhibitor of caspase-1 activation processing*. J Biol Chem, 2001. **276**(37): p. 34495-500.
108. Chen, G.Y. and G. Nunez, *Sterile inflammation: sensing and reacting to damage*. Nat Rev Immunol, 2010. **10**(12): p. 826-37.
109. Kanneganti, T.D., et al., *Pannexin-1-mediated recognition of bacterial molecules activates the cryopyrin inflammasome independent of Toll-like receptor signaling*. Immunity, 2007. **26**(4): p. 433-43.
110. Eisenbarth, S.C., et al., *Crucial role for the Nalp3 inflammasome in the immunostimulatory properties of aluminium adjuvants*. Nature, 2008. **453**(7198): p. 1122-6.
111. Morishige, T., et al., *Titanium dioxide induces different levels of IL-1beta production dependent on its particle characteristics through caspase-1 activation mediated by reactive oxygen species and cathepsin B*. Biochem Biophys Res Commun, 2010. **392**(2): p. 160-5.

112. Church, L.D., G.P. Cook, and M.F. McDermott, *Primer: inflammasomes and interleukin 1beta in inflammatory disorders*. Nat Clin Pract Rheumatol, 2008. **4**(1): p. 34-42.
113. Schilling, T. and C. Eder, *Importance of lipid rafts for lysophosphatidylcholine-induced caspase-1 activation and reactive oxygen species generation*. Cell Immunol, 2010. **265**(2): p. 87-90.
114. Fubini, B. and A. Hubbard, *Reactive oxygen species (ROS) and reactive nitrogen species (RNS) generation by silica in inflammation and fibrosis*. Free Radic Biol Med, 2003. **34**(12): p. 1507-16.
115. Boucher, M.J., et al., *MEK/ERK signaling pathway regulates the expression of Bcl-2, Bcl-X(L), and Mcl-1 and promotes survival of human pancreatic cancer cells*. J Cell Biochem, 2000. **79**(3): p. 355-69.
116. Brough, D., P. Pelegri, and N.J. Rothwell, *Pannexin-1-dependent caspase-1 activation and secretion of IL-1beta is regulated by zinc*. Eur J Immunol, 2009. **39**(2): p. 352-8.
117. Kahlenberg, J.M. and G.R. Dubyak, *Mechanisms of caspase-1 activation by P2X7 receptor-mediated K<sup>+</sup> release*. Am J Physiol Cell Physiol, 2004. **286**(5): p. C1100-8.
118. Kinoshita, T., et al., *PYPAF3, a PYRIN-containing APAF-1-like protein, is a feedback regulator of caspase-1-dependent interleukin-1beta secretion*. J Biol Chem, 2005. **280**(23): p. 21720-5.
119. Clem, R.J., et al., *Modulation of cell death by Bcl-XL through caspase interaction*. Proc Natl Acad Sci U S A, 1998. **95**(2): p. 554-9.
120. Passer, B.J., et al., *Interaction of Alzheimer's presenilin-1 and presenilin-2 with Bcl-X(L). A potential role in modulating the threshold of cell death*. J Biol Chem, 1999. **274**(34): p. 24007-13.
121. van de Craen, M., et al., *Identification of caspases that cleave presenilin-1 and presenilin-2. Five presenilin-1 (PS1) mutations do not alter the sensitivity of PS1 to caspases*. FEBS Lett, 1999. **445**(1): p. 149-54.
122. Fernandes-Alnemri, T., et al., *AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA*. Nature, 2009. **458**(7237): p. 509-13.
123. Hornung, V., et al., *AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC*. Nature, 2009. **458**(7237): p. 514-8.
124. Kawai, T. and S. Akira, *The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors*. Nat Immunol, 2010. **11**(5): p. 373-84.
125. Poeck, H., et al., *Recognition of RNA virus by RIG-I results in activation of CARD9 and inflammasome signaling for interleukin 1 beta production*. Nat Immunol, 2010. **11**(1): p. 63-9.
126. Bergsbaken, T., S.L. Fink, and B.T. Cookson, *Pyroptosis: host cell death and inflammation*. Nat Rev Microbiol, 2009. **7**(2): p. 99-109.
127. Takeda, K. and S. Akira, *TLR signaling pathways*. Semin Immunol, 2004. **16**(1): p. 3-9.
128. Grenier, J.M., et al., *Functional screening of five PYPAF family members identifies PYPAF5 as a novel regulator of NF-kappaB and caspase-1*. FEBS Lett, 2002. **530**(1-3): p. 73-8.
129. Geddes, B.J., et al., *Human CARD12 is a novel CED4/Apaf-1 family member that induces apoptosis*. Biochem Biophys Res Commun, 2001. **284**(1): p. 77-82.
130. Lu, C., et al., *Nucleotide binding to CARD12 and its role in CARD12-mediated caspase-1 activation*. Biochem Biophys Res Commun, 2005. **331**(4): p. 1114-9.
131. Silveira, T.N. and D.S. Zamboni, *Pore formation triggered by Legionella spp. is an Nlrc4 inflammasome-dependent host cell response that precedes pyroptosis*. Infect Immun, 2010. **78**(3): p. 1403-13.
132. Suzuki, T., et al., *Differential regulation of caspase-1 activation, pyroptosis, and autophagy via Ipaf and ASC in Shigella-infected macrophages*. PLoS Pathog, 2007. **3**(8): p. e111.
133. Miao, E.A., et al., *Innate immune detection of the type III secretion apparatus through the NLRC4 inflammasome*. Proc Natl Acad Sci U S A, 2010. **107**(7): p. 3076-80.

134. Lamkanfi, M., et al., *The Nod-like receptor family member Naip5/Birc1e restricts Legionella pneumophila growth independently of caspase-1 activation*. J Immunol, 2007. **178**(12): p. 8022-7.
135. Case, C.L., S. Shin, and C.R. Roy, *Asc and Ipaf Inflammasomes direct distinct pathways for caspase-1 activation in response to Legionella pneumophila*. Infect Immun, 2009. **77**(5): p. 1981-91.
136. Fortier, A., Chastellier, C., Balor, S., Gros, P., *Birc1e/Naip5 rapidly antagonizes modulation of phagosome maturation by Legionella pneumophila*. Cellular Microbiology, 2007. **9**(4): p. 910–923.
137. Wang, Y., et al., *PYNOD, a novel Apaf-1/CED4-like protein is an inhibitor of ASC and caspase-1*. Int Immunol, 2004. **16**(6): p. 777-86.
138. Dinerallo, C.A., *The interleukin-1 family: 10 years of discovery*. The FASEB Journal, 1994. **8**: p. 1314-1325.
139. O'Neill, L.A., *Targeting signal transduction as a strategy to treat inflammatory diseases*. Nat Rev Drug Discov, 2006. **5**(7): p. 549-63.
140. Shaw, M.H., et al., *NOD-like receptors (NLRs): bona fide intracellular microbial sensors*. Curr Opin Immunol, 2008. **20**(4): p. 377-82.
141. Druilhe, A.S., S.M.; Razmara, M.; Ahmad, M.; Alnemri, E.S., *Regulation of IL-1 $\beta$  generation by Pseudo-ICE and ICEBERG, two dominant negative caspase recruitment domain proteins*. Cell Death and Differentiation, 2001. **8**: p. 649 -657.
142. Ferrero-Miliani, L., Nielsen, O.H., Andersen, P.S., Girardin, S.E., *Chronic inflammation: importance of NOD2 and NALP3 in interleukin-1 $\beta$  generation*. Clinical and Experimental Immunology, 2006. **147**: p. 227–235.
143. *Uniprot Web*
144. Legler, D.F., et al., *Recruitment of TNF receptor 1 to lipid rafts is essential for TNF $\alpha$ -mediated NF- $\kappa$ B activation*. Immunity, 2003. **18**(5): p. 655-64.
145. Mischeau, O. and J. Tschopp, *Induction of TNF receptor 1-mediated apoptosis via two sequential signaling complexes*. Cell, 2003. **114**(2): p. 181-90.
146. Yoo, N.J., et al., *Nod1, a CARD protein, enhances pro-interleukin-1 $\beta$  processing through the interaction with pro-caspase-1*. Biochem Biophys Res Commun, 2002. **299**(4): p. 652-8.
147. Humke, E.W., et al., *ICEBERG: a novel inhibitor of interleukin-1 $\beta$  generation*. Cell, 2000. **103**(1): p. 99-111.
148. Stehlik, C., et al., *Apoptosis-associated speck-like protein containing a caspase recruitment domain is a regulator of procaspase-1 activation*. J Immunol, 2003. **171**(11): p. 6154-63.
149. Inohara, N., et al., *RICK, a novel protein kinase containing a caspase recruitment domain, interacts with CLARP and regulates CD95-mediated apoptosis*. J Biol Chem, 1998. **273**(20): p. 12296-300.
150. Sun, Y., et al., *Lysophosphatidylcholine-induced apoptosis in H19-7 hippocampal progenitor cells is enhanced by the upregulation of Fas Ligand*. Biochim Biophys Acta, 2009. **1791**(1): p. 61-8.
151. Inohara, N., et al., *Nod1, an Apaf-1-like activator of caspase-9 and nuclear factor- $\kappa$ B*. J Biol Chem, 1999. **274**(21): p. 14560-7.
152. Gaggero, A., et al., *A novel isoform of pro-interleukin-18 expressed in ovarian tumors is resistant to caspase-1 and -4 processing*. Oncogene, 2004. **23**(45): p. 7552-60.
153. Goldstein, D.R., et al., *Critical role of the Toll-like receptor signal adaptor protein MyD88 in acute allograft rejection*. J Clin Invest, 2003. **111**(10): p. 1571-8.
154. Mariathasan, S., *ASC, Ipaf and Cryopyrin/Nalp3: bona fide intracellular adapters of the caspase-1 inflammasome*. Microbes Infect, 2007. **9**(5): p. 664-71.
155. Kanneganti, T.D., et al., *Bacterial RNA and small antiviral compounds activate caspase-1 through cryopyrin/Nalp3*. Nature, 2006. **440**(7081): p. 233-6.
156. Carneiro, L.A., L.H. Travassos, and D.J. Philpott, *Innate immune recognition of microbes through Nod1 and Nod2: implications for disease*. Microbes Infect, 2004. **6**(6): p. 609-16.

157. Miggin, S.M., et al., *NF-kappaB activation by the Toll-IL-1 receptor domain protein MyD88 adapter-like is regulated by caspase-1*. Proc Natl Acad Sci U S A, 2007. **104**(9): p. 3372-7.
158. Scully, C., T. Hodgson, and H. Lachmann, *Auto-inflammatory syndromes and oral health*. Oral Dis, 2008. **14**(8): p. 690-9.
159. Walev, I., et al., *Potassium regulates IL-1 beta processing via calcium-independent phospholipase A2*. J Immunol, 2000. **164**(10): p. 5120-4.
160. Eder, C., *Mechanisms of interleukin-1beta release*. Immunobiology, 2009. **214**(7): p. 543-53.
161. Luschen, S., et al., *Cleavage of human cytosolic phospholipase A2 by caspase-1 (ICE) and caspase-8 (FLICE)*. Biochem Biophys Res Commun, 1998. **253**(1): p. 92-8.
162. Srinivasula, S.M., et al., *Molecular ordering of the Fas-apoptotic pathway: the Fas/APO-1 protease Mch5 is a CrmA-inhibitable protease that activates multiple Ced-3/ICE-like cysteine proteases*. Proc Natl Acad Sci U S A, 1996. **93**(25): p. 14486-91.
163. Hu, S., et al., *Caspase-14 is a novel developmentally regulated protease*. J Biol Chem, 1998. **273**(45): p. 29648-53.
164. Davis, B.K., H. Wen, and J.P. Ting, *The Inflammasome NLRs in Immunity, Inflammation, and Associated Diseases*. Annu Rev Immunol, 2010.
165. Annand, R.R., et al., *Caspase-1 (interleukin-1beta-converting enzyme) is inhibited by the human serpin analogue proteinase inhibitor 9*. Biochem J, 1999. **342 Pt 3**: p. 655-65.
166. Mariathasan, S. and D.M. Monack, *Inflammasome adaptors and sensors: intracellular regulators of infection and inflammation*. Nat Rev Immunol, 2007. **7**(1): p. 31-40.
167. Stehlik, C., et al., *The PAAD/PYRIN-only protein POP1/ASC2 is a modulator of ASC-mediated nuclear-factor-kappa B and pro-caspase-1 regulation*. Biochem J, 2003. **373**(Pt 1): p. 101-13.
168. Castell, J.V., et al., *Interleukin-6 is the major regulator of acute phase protein synthesis in adult human hepatocytes*. FEBS Lett, 1989. **242**(2): p. 237-9.
169. Keller, M., et al., *Active caspase-1 is a regulator of unconventional protein secretion*. Cell, 2008. **132**(5): p. 818-31.
170. Kumar, S., et al., *Interleukin-1F7B (IL-1H4/IL-1F7) is processed by caspase-1 and mature IL-1F7B binds to the IL-18 receptor but does not induce IFN-gamma production*. Cytokine, 2002. **18**(2): p. 61-71.
171. Zhang, W.H., et al., *Fundamental role of the Rip2/caspase-1 pathway in hypoxia and ischemia-induced neuronal cell death*. Proc Natl Acad Sci U S A, 2003. **100**(26): p. 16012-7.
172. Sutton, V.R., et al., *Initiation of apoptosis by granzyme B requires direct cleavage of bid, but not direct granzyme B-mediated caspase activation*. J Exp Med, 2000. **192**(10): p. 1403-14.
173. Walensky, L.D., et al., *Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix*. Science, 2004. **305**(5689): p. 1466-70.
174. Milhas, D., et al., *Caspase-10 triggers Bid cleavage and caspase cascade activation in FasL-induced apoptosis*. J Biol Chem, 2005. **280**(20): p. 19836-42.
175. Wellington, C.L., et al., *Caspase cleavage of gene products associated with triplet expansion disorders generates truncated fragments containing the polyglutamine tract*. J Biol Chem, 1998. **273**(15): p. 9158-67.
176. Harvey, K.F., et al., *Caspase-mediated cleavage of the ubiquitin-protein ligase Nedd4 during apoptosis*. J Biol Chem, 1998. **273**(22): p. 13524-30.
177. Wang, K.K., et al., *Caspase-mediated fragmentation of calpain inhibitor protein calpastatin during apoptosis*. Arch Biochem Biophys, 1998. **356**(2): p. 187-96.
178. Gu, Y., et al., *Cleavage of poly(ADP-ribose) polymerase by interleukin-1 beta converting enzyme and its homologs TX and Nedd-2*. J Biol Chem, 1995. **270**(32): p. 18715-8.

179. Bae, S.S., et al., *Proteolytic cleavage of epidermal growth factor receptor by caspases*. FEBS Lett, 2001. **491**(1-2): p. 16-20.
180. Puto, L.A., et al., *p21-activated kinase 1 (PAK1) interacts with the Grb2 adapter protein to couple to growth factor signaling*. J Biol Chem, 2003. **278**(11): p. 9388-93.
181. Liu, Y.F., R.C. Deth, and D. Devys, *SH3 domain-dependent association of huntingtin with epidermal growth factor receptor signaling complexes*. J Biol Chem, 1997. **272**(13): p. 8121-4.
182. Bonaccorsi, L., et al., *Androgen receptor and prostate cancer invasion*. Int J Androl, 2003. **26**(1): p. 21-5.
183. Tang, D., R. Gururajan, and V.J. Kidd, *Phosphorylation of PITSLRE p110 isoforms accompanies their processing by caspases during Fas-mediated cell death*. J Biol Chem, 1998. **273**(26): p. 16601-7.
184. Beyaert, R., et al., *Cleavage of PITSLRE kinases by ICE/CASP-1 and CPP32/CASP-3 during apoptosis induced by tumor necrosis factor*. J Biol Chem, 1997. **272**(18): p. 11694-7.
185. Chen, S., et al., *The C-terminal kinase domain of the p34cdc2-related PITSLRE protein kinase (p110C) associates with p21-activated kinase 1 and inhibits its activity during anoikis*. J Biol Chem, 2003. **278**(22): p. 20029-36.
186. Kahns, S., et al., *Caspase-1 and caspase-8 cleave and inactivate cellular parkin*. J Biol Chem, 2003. **278**(26): p. 23376-80.
187. Gamblin, T.C., et al., *Caspase cleavage of tau: linking amyloid and neurofibrillary tangles in Alzheimer's disease*. Proc Natl Acad Sci U S A, 2003. **100**(17): p. 10032-7.
188. Takashima, A., et al., *Presenilin 1 associates with glycogen synthase kinase-3beta and its substrate tau*. Proc Natl Acad Sci U S A, 1998. **95**(16): p. 9637-41.
189. Moore, D.J., et al., *Parkin mediates the degradation-independent ubiquitination of Hsp70*. J Neurochem, 2008. **105**(5): p. 1806-19.
190. Ng, F.W., et al., *p28 Bap31, a Bcl-2/Bcl-XL- and procaspase-8-associated protein in the endoplasmic reticulum*. J Cell Biol, 1997. **139**(2): p. 327-38.
191. Breckenridge, D.G., et al., *Caspase cleavage product of BAP31 induces mitochondrial fission through endoplasmic reticulum calcium signals, enhancing cytochrome c release to the cytosol*. J Cell Biol, 2003. **160**(7): p. 1115-27.
192. McIntire, C.R., G. Yeretssian, and M. Saleh, *Inflammasomes in infection and inflammation*. Apoptosis, 2009. **14**(4): p. 522-35.
193. Shen, J., et al., *Caspase-1 recognizes extended cleavage sites in its natural substrates*. Atherosclerosis, 2010. **210**(2): p. 422-9.
194. Millar, J.K., S. Christie, and D.J. Porteous, *Yeast two-hybrid screens implicate DISC1 in brain development and function*. Biochem Biophys Res Commun, 2003. **311**(4): p. 1019-25.
195. Gupta, S., et al., *A nuclear protein tyrosine phosphatase activates p53 and induces caspase-1-dependent apoptosis*. FEBS Lett, 2002. **532**(1-2): p. 61-6.
196. Sadasivam, S., et al., *Caspase-1 activator Ipaf is a p53-inducible gene involved in apoptosis*. Oncogene, 2005. **24**(4): p. 627-36.
197. Townsend, P.A., et al., *STAT-1 interacts with p53 to enhance DNA damage-induced apoptosis*. J Biol Chem, 2004. **279**(7): p. 5811-20.
198. Kobierski, L.A., S. Srivastava, and D. Borsook, *Systemic lipopolysaccharide and interleukin-1beta activate the interleukin 6: STAT intracellular signaling pathway in neurons of mouse trigeminal ganglion*. Neurosci Lett, 2000. **281**(1): p. 61-4.
199. Garrote, J.A., et al., *Celiac disease pathogenesis: the proinflammatory cytokine network*. J Pediatr Gastroenterol Nutr, 2008. **47 Suppl 1**: p. S27-32.

200. Vigers, G.P., et al., *X-ray crystal structure of a small antagonist peptide bound to interleukin-1 receptor type I*. J Biol Chem, 2000. **275**(47): p. 36927-33.
201. Shahripour, A.B., et al., *Structure-based design of caspase-1 inhibitor containing a diphenyl ether sulfonamide*. Bioorg Med Chem Lett, 2001. **11**(20): p. 2779-82.
202. Ming, X., et al., *Caspase-1 expression in multiple sclerosis plaques and cultured glial cells*. J Neurol Sci, 2002. **197**(1-2): p. 9-18.
203. Siegmund, B., *Interleukin-1beta converting enzyme (caspase-1) in intestinal inflammation*. Biochem Pharmacol, 2002. **64**(1): p. 1-8.
204. Los, M., et al., *Anticancer drugs of tomorrow: apoptotic pathways as targets for drug design*. Drug Discov Today, 2003. **8**(2): p. 67-77.
205. Okamoto, Y., et al., *Peptide based interleukin-1 beta converting enzyme (ICE) inhibitors: synthesis, structure activity relationships and crystallographic study of the ICE-inhibitor complex*. Chem Pharm Bull (Tokyo), 1999. **47**(1): p. 11-21.
206. Talanian, R.V., et al., *Substrate specificities of caspase family proteases*. J Biol Chem, 1997. **272**(15): p. 9677-82.
207. Sifringer, M., et al., *Activation of caspase-1 dependent interleukins in developmental brain trauma*. Neurobiol Dis, 2007. **25**(3): p. 614-22.
208. Fink, K.B., et al., *Reduction of post-traumatic brain injury and free radical production by inhibition of the caspase-1 cascade*. Neuroscience, 1999. **94**(4): p. 1213-8.
209. Tringali, G., et al., *Depolarization and a NO-donor stimulate interleukin-1beta release from the rat hypothalamus via a mechanism involving caspase 1*. Neurosci Lett, 1999. **276**(2): p. 119-22.
210. Sims, J.E. and D.E. Smith, *The IL-1 family: regulators of immunity*. Nat Rev Immunol, 2010. **10**(2): p. 89-102.
211. Hasegawa, A., et al., *Caspase inhibitors increase the rate of recovery of neural stem/progenitor cells from post-mortem rat brains stored at room temperature*. J Biosci Bioeng, 2009. **107**(6): p. 652-7.
212. Ghose, A.K., V.N. Viswanadhan, and J.J. Wendoloski, *A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases*. J Comb Chem, 1999. **1**(1): p. 55-68.
213. Mozziconacci, J.C., Arnoult, E., Baurin, N., Marot, C., Allory, L.M., *Preparation of a molecular database from a set of 2 million compounds for virtual screening applications: Gathering, structural analysis and filtering*. Institut de Chimie Organique et Analytique, Universite d'Orleans.
214. Lee, M.L. and G. Schneider, *Scaffold architecture and pharmacophoric properties of natural products and trade drugs: application in the design of natural product-based combinatorial libraries*. J Comb Chem, 2001. **3**(3): p. 284-9.
215. Oprea, T.I., *Property distribution of drug-related chemical databases*. J Comput Aided Mol Des, 2000. **14**(3): p. 251-64.
216. Walters, W.B., Murcko, M.A., *Library filtering systems and prediction of drug-like properties*. Methods and Principles in Medicinal Chemistry, 2000. **10**: p. 15-32.
217. Lipinski, C.A., et al., *Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings*. Adv Drug Deliv Rev, 2001. **46**(1-3): p. 3-26.
218. Fichert, T., M. Yazdanian, and J.R. Proudfoot, *A structure-permeability study of small drug-like molecules*. Bioorg Med Chem Lett, 2003. **13**(4): p. 719-22.
219. Palm, K., et al., *Polar molecular surface properties predict the intestinal absorption of drugs in humans*. Pharm Res, 1997. **14**(5): p. 568-71.
220. Veber, D.F., et al., *Molecular properties that influence the oral bioavailability of drug candidates*. J Med Chem, 2002. **45**(12): p. 2615-23.

221. Ajay, G.W. Bemis, and M.A. Murcko, *Designing libraries with CNS activity*. J Med Chem, 1999. **42**(24): p. 4942-51.
222. van de Waterbeemd, H., et al., *Estimation of blood-brain barrier crossing of drugs using molecular size and shape, and H-bonding descriptors*. J Drug Target, 1998. **6**(2): p. 151-65.
223. Tan, M.L., P.F. Choong, and C.R. Dass, *Recent developments in liposomes, microparticles and nanoparticles for protein and peptide drug delivery*. Peptides, 2010. **31**(1): p. 184-93.
224. Purcell, A.W., J. McCluskey, and J. Rossjohn, *More than one reason to rethink the use of peptides in vaccine design*. Nat Rev Drug Discov, 2007. **6**(5): p. 404-14.
225. Stenberg, P., K. Luthman, and P. Artursson, *Prediction of membrane permeability to peptides from calculated dynamic molecular surface properties*. Pharm Res, 1999. **16**(2): p. 205-12.
226. Unal, E.B., A. Gursoy, and B. Erman, *Conformational energies and entropies of peptides, and the peptide-protein binding problem*. Phys Biol, 2009. **6**(3): p. 036014.
227. Vives, E., J. Schmidt, and A. Pelegrin, *Cell-penetrating and cell-targeting peptides in drug delivery*. Biochim Biophys Acta, 2008. **1786**(2): p. 126-38.
228. Stevenson, C.L., *Advances in peptide pharmaceuticals*. Curr Pharm Biotechnol, 2009. **10**(1): p. 122-37.
229. Dao, T. and D.A. Scheinberg, *Peptide vaccines for myeloid leukaemias*. Best Pract Res Clin Haematol, 2008. **21**(3): p. 391-404.
230. Apostolopoulos, V., *Peptide-based vaccines for cancer: are we choosing the right peptides?* Expert Rev Vaccines, 2009. **8**(3): p. 259-60.
231. Gozes, I., *Neuroprotective peptide drug delivery and development: potential new therapeutics*. Trends Neurosci, 2001. **24**(12): p. 700-5.
232. Beck, A., et al., *Peptides as tools and drugs for immunotherapies*. J Pept Sci, 2007. **13**(9): p. 588-602.
233. Iwai, T., et al., *Glucagon-like peptide-1 inhibits LPS-induced IL-1beta production in cultured rat astrocytes*. Neurosci Res, 2006. **55**(4): p. 352-60.
234. Shaji, J. and V. Patole, *Protein and Peptide drug delivery: oral approaches*. Indian J Pharm Sci, 2008. **70**(3): p. 269-77.
235. Rano, T.A., et al., *A combinatorial approach for determining protease specificities: application to interleukin-1beta converting enzyme (ICE)*. Chem Biol, 1997. **4**(2): p. 149-55.
236. Karanewsky, D.S., et al., *Conformationally constrained inhibitors of caspase-1 (interleukin-1 beta converting enzyme) and of the human CED-3 homologue caspase-3 (CPP32, apopain)*. Bioorg Med Chem Lett, 1998. **8**(19): p. 2757-62.
237. Hayashi, Y., et al., *Immunohistochemical investigation of caspase-1 and effect of caspase-1 inhibitor in delayed neuronal death after transient cerebral ischemia*. Brain Res, 2001. **893**(1-2): p. 113-20.
238. Linton, S.D., et al., *Acyl dipeptides as reversible caspase inhibitors. Part 2: further optimization*. Bioorg Med Chem Lett, 2002. **12**(20): p. 2973-5.
239. Nedev, H.N., et al., *Synthesis and evaluation of novel dipeptidyl benzoyloxymethyl ketones as caspase inhibitors*. Biochem Biophys Res Commun, 2005. **336**(2): p. 397-400.
240. Dolle, R.E., et al., *P1 aspartate-based peptide alpha-((2,6-dichlorobenzoyl)oxy)methyl ketones as potent time-dependent inhibitors of interleukin-1 beta-converting enzyme*. J Med Chem, 1994. **37**(5): p. 563-4.
241. Schlosser, S., et al., *Inhibition of caspase-1 induces cell death in pancreatic carcinoma cells and potentially modulates expression levels of bcl-2 family proteins*. FEBS Lett, 2001. **491**(1-2): p. 104-8.
242. Fahr, B.T., et al., *Tethering identifies fragment that yields potent inhibitors of human caspase-1*. Bioorg Med Chem Lett, 2006. **16**(3): p. 559-62.



243. Fiorucci, S., *NO-releasing NSAIDs are caspase inhibitors*. Trends Immunol, 2001. **22**(5): p. 232-5.
244. Shahripour, A.B., et al., *Structure-based design of nonpeptide inhibitors of interleukin-1beta converting enzyme (ICE, caspase-1)*. Bioorg Med Chem, 2002. **10**(1): p. 31-40.
245. Ross, J., et al., *A selective, non-peptide caspase-1 inhibitor, VRT-018858, markedly reduces brain damage induced by transient ischemia in the rat*. Neuropharmacology, 2007. **53**(5): p. 638-42.
246. Ku, G., Ford, P., Raybuck, S.A., Harding, M.W., Randle, J.C., *Selective interleukin-1 $\beta$  converting enzyme (ICE/caspase-1) inhibition with pralnacasan (HMR 3480/VX-740) reduces inflammation and joint destruction in murine type ii collagen-induced arthritis (CIA)*. Arthritis Rheum, 2001. **44**: p. S241.
247. Brady, K.D., *Bimodal inhibition of caspase-1 by aryloxymethyl and acyloxymethyl ketones*. Biochemistry, 1998. **37**(23): p. 8508-15.
248. Thornberry, N.A., et al., *Inactivation of interleukin-1 beta converting enzyme by peptide (acyloxy)methyl ketones*. Biochemistry, 1994. **33**(13): p. 3934-40.
249. Brady, K.D., et al., *A catalytic mechanism for caspase-1 and for bimodal inhibition of caspase-1 by activated aspartic ketones*. Bioorg Med Chem, 1999. **7**(4): p. 621-31.
250. Garcia-Calvo, M., Peterson, E.P., Leiting, B., Ruel, R., Nicholson, D. W., Thornberry, N. A., *Inhibition of human caspases by peptide-based and macromolecular inhibitors*. The Journal of Biological Chemistry, 1998. **49**: p. 32608-32613.
251. LeBlanc, A.C., *Natural cellular inhibitors of caspases*. Prog Neuropsychopharmacol Biol Psychiatry, 2003. **27**(2): p. 215-29.
252. Jones, G., et al., *Development and validation of a genetic algorithm for flexible docking*. J Mol Biol, 1997. **267**(3): p. 727-48.
253. Moitessier, N., et al., *Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go*. Br J Pharmacol, 2008. **153 Suppl 1**: p. S7-26.
254. Sousa, S.F., P.A. Fernandes, and M.J. Ramos, *Protein-ligand docking: current status and future challenges*. Proteins, 2006. **65**(1): p. 15-26.
255. Unger, R., *The genetic algorithm approach to protein structure*. Structure and Bonding, 2004. **110**: p. 153-175.
256. Campbell, S.J., et al., *Ligand binding: functional site location, similarity and docking*. Curr Opin Struct Biol, 2003. **13**(3): p. 389-95.
257. Kitchen, D.B., et al., *Docking and scoring in virtual screening for drug discovery: methods and applications*. Nat Rev Drug Discov, 2004. **3**(11): p. 935-49.
258. Taylor, R.D., P.J. Jewsbury, and J.W. Essex, *A review of protein-small molecule docking methods*. J Comput Aided Mol Des, 2002. **16**(3): p. 151-66.
259. Dastmalchi, S., Hamzeh-Mivehrod, M., *Molecular modeling of human aldehyde oxidase and identification of the key interactions in the enzyme-substrate complex*. DARU, 2005. **13**.
260. Ziemys, A., Rimkute, L., Kulys, J., *Modeling the enantioselective enzymatic reaction with modified genetic docking algorithm*. Nonlinear Analysis: Modeling and Control, 2004. **9**: p. 373-383.
261. Tiwari, R., et al., *Carborane clusters in computational drug design: a comparative docking evaluation using AutoDock, FlexX, Glide, and Surflex*. J Chem Inf Model, 2009. **49**(6): p. 1581-9.
262. Morris, G.M., et al., *AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility*. J Comput Chem, 2009. **30**(16): p. 2785-91.
263. Solmajer, T. and E.L. Mehler, *Electrostatic screening in molecular dynamics simulations*. Protein Eng, 1991. **4**(8): p. 911-7.
264. Boobbyer, D.N., et al., *New hydrogen-bond potentials for use in determining energetically favorable binding sites on molecules of known structure*. J Med Chem, 1989. **32**(5): p. 1083-94.

265. Stouten, P.F.W., Frömmel, C., Nakamura, H., Sander, C., *An effective solvation term based on atomic occupancies for use in protein simulations*. Molecular Simulations, 1993. **10**(2-6): p. 97-120.
266. *Autodock converts binding energy into Ki*. Available from: <http://autodock.scripps.edu/faqs-help/faq/how-autodock-4-converts-binding-energy-kcal-mol-into-ki>.
267. Huey, R., et al., *A semiempirical free energy force field with charge-based desolvation*. J Comput Chem, 2007. **28**(6): p. 1145-52.
268. Cincilla, G., D. Vidal, and M. Pons, *An improved scoring function for suboptimal polar ligand complexes*. J Comput Aided Mol Des, 2009. **23**(3): p. 143-52.
269. Unal, E.B., A. GURSOY, and B. ERMAN, *VitAL: Viterbi algorithm for de novo peptide design*. PLoS One, 2010. **5**(6): p. e10926.
270. Unal, E.B., GURSOY, A., ERMAN, B., *Inhibitor peptide design for NF-KB: Markov Model & Genetic Algorithm*, in *HIBIT*. 2010: Antalya, Turkey.
271. Yoshimori, A., R. Takasawa, and S. Tanuma, *A novel method for evaluation and screening of caspase inhibitory peptides by the amino acid positional fitness score*. BMC Pharmacol, 2004. **4**: p. 7.
272. Verdonk, M.L., et al., *Improved protein-ligand docking using GOLD*. Proteins, 2003. **52**(4): p. 609-23.
273. Hartshorn, M.J., et al., *Diverse, high-quality test set for the validation of protein-ligand docking performance*. J Med Chem, 2007. **50**(4): p. 726-41.
274. Loser, R., et al., *Noncovalent tripeptidyl benzyl- and cyclohexyl-amine inhibitors of the cysteine protease caspase-1*. J Med Chem, 2010. **53**(6): p. 2651-5.
275. Josephus, M., Heijligers, M., *The application of genetic algorithms to high-level synthesis*. 1996, Technische Universiteit Eindhoven.
276. Holland, J.H., *Genetic Algorithm*. 1992: p. 66-72.
277. Humphrey, W., A. Dalke, and K. Schulten, *VMD: visual molecular dynamics*. J Mol Graph, 1996. **14**(1): p. 33-8, 27-8.
278. Abe, K., et al., *Peptide ligand screening of alpha-synuclein aggregation modulators by in silico panning*. BMC Bioinformatics, 2007. **8**: p. 451.
279. Kernysky, A. and B. Rost, *Using genetic algorithms to select most predictive protein features*. Proteins, 2009. **75**(1): p. 75-88.
280. Yang, Z., Liu, G., *3D protein peptide chain search using an improved genetic algorithm*. ICCSA, 2003: p. 322-329.
281. Budin, N., Majeux, N., Tenette-Souaille, C., Caflisch, A., *Structure based ligand design by a Build-up approach and genetic algorithm search in conformational space*. Journal of Computational Chemistry, 2001. **22**: p. 1956-1970.
282. Barreca, M.L., et al., *Structure-based pharmacophore identification of new chemical scaffolds as non-nucleoside reverse transcriptase inhibitors*. J Chem Inf Model, 2007. **47**(2): p. 557-62.
283. Steindl, T.M., et al., *High-throughput structure-based pharmacophore modelling as a basis for successful parallel virtual screening*. J Comput Aided Mol Des, 2006. **20**(12): p. 703-15.
284. Wolber, G., A.A. Dornhofer, and T. Langer, *Efficient overlay of small organic molecules using 3D pharmacophores*. J Comput Aided Mol Des, 2006. **20**(12): p. 773-88.
285. Wolber, G. and T. Langer, *LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters*. J Chem Inf Model, 2005. **45**(1): p. 160-9.
286. Rader, A.J., Chennubhotla, C., Yang, L.W., Bahar, I., *The Gaussian Network Model: Theory and Applications*, in *Normal Mode Analysis. Theory and Applications to Biological and Chemical Systems*, Q. Cui, Bahar, I., Editor. 2006, CRC Press, Taylor & Francis Group. p. 41-64.

287. Tirion, M.M., *Large amplitude elastic motions in proteins from a single-parameter, atomic analysis*. Physical Review Letters, 1996. **77**: p. 1905-1908.
288. Bahar, I., Erman, B., Jernigan, R.L., Atilgan, A.R., Covell, D.G., *Collective motions in HIV-1 reverse transcriptase: Examination of flexibility and enzyme function*. J Mol Biol, 1999. **285**: p. 1023-1037.
289. Bahar, I., A.R. Atilgan, and B. Erman, *Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential*. Fold Des, 1997. **2**(3): p. 173-81.
290. Haliloglu, T., Bahar, I., Erman, B., *Gaussian Dynamics of Folded Proteins*. Physical Review Letters, 1997. **79**: p. 3090-3093.
291. Bahar, I., Atilgan, A.R., Demirel, M.C., Erman, B., *Vibrational dynamics of folded proteins: Significance of slow and fast motions in relation to function and stability*. Physical Review Letters, 1998. **80**: p. 2733-2736.
292. Ozbek, P., et al., *DNABINDPROT: fluctuation-based predictor of DNA-binding residues within a network of interacting residues*. Nucleic Acids Res, 2010. **38**(Web Server issue): p. W417-23.
293. Demirel, M.C., et al., *Identification of kinetically hot residues in proteins*. Protein Sci, 1998. **7**(12): p. 2522-32.
294. Demirel, M.C. and O. Keskin, *Protein interactions and fluctuations in a proteomic network using an elastic network model*. J Biomol Struct Dyn, 2005. **22**(4): p. 381-6.
295. Erman, B., *The gaussian network model: precise prediction of residue fluctuations and application to binding problems*. Biophys J, 2006. **91**(10): p. 3589-99.
296. Haliloglu, T. and B. Erman, *Analysis of Correlations between Energy and Residue Fluctuations in Native Proteins and Determination of Specific Sites for Binding*. Physical Review Letters, 2009. **102**(8): p. 088103.
297. Tuzmen, C., Erman, B., *Identification of Ligand Binding Sites of Proteins Using the Gaussian Network Model*. PLoS ONE, 2011. **6**(1).
298. Widom, B., *Statistical mechanics: A concise introduction for chemists* 2002: Cambridge University Press.
299. Forli, S., *Raccoon/AutoDock VS: an automated tool for preparing AutoDock virtual screenings*. 2010.
300. *Accelrys Discovery Studio Visualizer 2.5* Available from: <http://accelrys.com/products/discovery-studio/>.
301. Kontoyianni, M., McClellan, L.M., Sokol, G.S., *Evaluation of docking performance: Comparative data on docking algorithms*. J. Med. Chem, 2004. **47**: p. 558-565.
302. Schuster, D., et al., *Pharmacophore modeling and in silico screening for new P450 19 (aromatase) inhibitors*. J Chem Inf Model, 2006. **46**(3): p. 1301-11.
303. *Marvin Beans*. Available from: <https://www.chemaxon.com/products/marvin/>.
304. *HyperChem Release 7*. Available from: <http://www.hyper.com/News/PressRelease/Release7Oct2001/tabid/412/Default.aspx>.
305. *HyperChem Release 7 Tools for Molecular Docking (User Manual)*. 2002, Hypercube Inc.
306. *ChemBio3D Pro*. Available from: <http://www.cambridgesoft.com/software/chembio3d/>.
307. Wallace, A., Laskowski, R., *LigPlot*. 2011, UCLB E-LUCID.
308. Flory, P.J., *Statistical Mechanics of Chain Molecules*. 1969, New York: Wiley.
309. Gursoy, A., Keskin, O., Turkay, M., Erman, B., *Relationships between unfolded configurations of proteins and dynamics of folding to the native state*. Journal of Polymer Science, 2006. **44**: p. 3667-3678.
310. *The Protein Coil Library*.
311. *RCSB Protein Data Bank*.

312. Karplus, P.A., *Experimentally observed conformation-dependent geometry and hidden strain in proteins*. Protein Sci, 1996. **5**(7): p. 1406-20.
313. Unal, E.B., Gursoy, A., Erman, B., *Vital Server: Viterbi Algorithm for De Novo Peptide Design*. 2011.
314. Callen, H.B., *Thermodynamics and an introduction to thermostatistics*. Second ed. 1985, New York: John Wiley & Sons.
315. Bahar, I., Erman, B., *Investigation of Local Motions in Polymers by the Dynamic Rotational Isomeric State Model*. Macromolecules, 1987. **20**: p. 1368-1376.
316. Haliloglu, T., A. Gul, and B. Erman, *Predicting important residues and interaction pathways in proteins using Gaussian Network Model: binding and stability of HLA proteins*. PLoS Comput Biol, 2010. **6**(7): p. e1000845.
317. Nilsson, M., *First Order Hidden Markov Model Theory and Implementation Issues*. 2005, Department of Signal Processing Blekinge Institute of Technology. p. 1-93.
318. Ewens, W.J., Grant, G.R., *Statistical methods in bioinformatics: an introduction*. 2001, New York: Springer.
319. *MATLAB 2008*, MathWorks.
320. Klockozkowski, A., Mark, J.E., *Chain Dimensions and Fluctuations in Random Elastomeric Networks. I. Phantom Gaussian Networks in the Undeformed State*. Macromolecules, 1989. **22**: p. 1423-1432.
321. McAllister, J.P., 2nd and J.M. Miller, *Minocycline inhibits glial proliferation in the H-Tx rat model of congenital hydrocephalus*. Cerebrospinal Fluid Res, 2010. **7**: p. 7.
322. O'Brien, T., Fahr, B.T., Sopko, M.M., Lam, J.W., Waal, N.D., Raimundo, B.C., Purkey, H.E., Pham, P., Romanowsk, M.J. , *Structural analysis of caspase-1 inhibitors derived from tethering*. Acta Crystallogr. Sect F Struct. Biol. Cryst. Commun., 2005. **61**: p. 451-458.
323. O'Brien, T., Fahr, B.T., Sopko, M.M., Lam, J.W., Waal, N.D., Raimundo, B.C., Purkey, H.E., Pham, P., Romanowsk, M.J. , *Novel caspase-1 inhibitors discovered using tethering (SM) with extenders*. To be published.
324. Sela, M. and E. Zisman, *Different roles of D-amino acids in immune phenomena*. FASEB J, 1997. **11**(6): p. 449-56.
325. Bartzatt, R.L., *Utilizing a D-amino acid as a drug carrier for antineoplastic nitrogen mustard groups*. Drug Deliv, 2005. **12**(3): p. 141-7.
326. Hamomoto, K., Kida, Y., Zhang, Y., Shimizu, T., Kuwano, K., *Antimicrobial Activity and Stability to Proteolysis of Small Linear Cationic Peptides with D-Amino Acid Substitutions*. Microbiology Immunology, 2002. **46**: p. 741-749.
327. Ophardt, C.E. *Virtual Chembook Characteristics and Properties of Amino Acids*. 2003; Available from: <http://www.elmhurst.edu/~chm/vchembook/561aminostructure.html>.
328. Park, Y.C., et al., *A novel mechanism of TRAF signaling revealed by structural and functional analyses of the TRADD-TRAF2 interaction*. Cell, 2000. **101**(7): p. 777-87.
329. Kim, D.J., et al., *Association of TRAF2 with the short form of cellular FLICE-like inhibitory protein prevents TNFR1-mediated apoptosis*. J Mol Signal, 2008. **3**: p. 2.
330. Zurawek, M., et al., *A coding variant in NLRP1 is associated with autoimmune Addison's disease*. Hum Immunol, 2010. **71**(5): p. 530-4.
331. *AIM2 absent in melanoma 2*. Available from: <http://www.ncbi.nlm.nih.gov/sites/entrez?Db=gene&Cmd=ShowDetailView&TermToSearch=9447>.
332. Elliott, J.M., et al., *Crystal structure of procaspase-1 zymogen domain reveals insight into inflammatory caspase autoactivation*. J Biol Chem, 2009. **284**(10): p. 6546-53.

333. Pereira, N.A. and Z. Song, *Some commonly used caspase substrates and inhibitors lack the specificity required to monitor individual caspase activity*. *Biochem Biophys Res Commun*, 2008. **377**(3): p. 873-7.
334. Jin, C. and R.A. Flavell, *Inflammasome activation. The missing link: how the inflammasome senses oxidative stress*. *Immunol Cell Biol*, 2010. **88**(5): p. 510-2.
335. Jin, C. and R.A. Flavell, *Molecular mechanism of NLRP3 inflammasome activation*. *J Clin Immunol*, 2010. **30**(5): p. 628-31.
336. Zhou, R., et al., *Thioredoxin-interacting protein links oxidative stress to inflammasome activation*. *Nat Immunol*, 2010. **11**(2): p. 136-40.
337. Scott, A.M. and M. Saleh, *The inflammatory caspases: guardians against infections and sepsis*. *Cell Death Differ*, 2007. **14**(1): p. 23-31.
338. Fiorentino, L., et al., *A novel PAAD-containing protein that modulates NF-kappa B induction by cytokines tumor necrosis factor-alpha and interleukin-1beta*. *J Biol Chem*, 2002. **277**(38): p. 35333-40.
339. Prohinar, P., et al., *Expression of functional D299G.T399I polymorphic variant of TLR4 depends more on coexpression of MD-2 than does wild-type TLR4*. *J Immunol*, 2010. **184**(8): p. 4362-7.
340. Park, B.S., et al., *The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex*. *Nature*, 2009. **458**(7242): p. 1191-5.
341. Schenk, M., J.T. Belisle, and R.L. Modlin, *TLR2 looks at lipoproteins*. *Immunity*, 2009. **31**(6): p. 847-9.
342. Delaloye, J., et al., *Innate immune sensing of modified vaccinia virus Ankara (MVA) is mediated by TLR2-TLR6, MDA-5 and the NALP3 inflammasome*. *PLoS Pathog*, 2009. **5**(6): p. e1000480.
343. Han, K.J., et al., *Analysis of a TIR-less splice variant of TRIF reveals an unexpected mechanism of TLR3-mediated signaling*. *J Biol Chem*, 2010. **285**(17): p. 12543-50.
344. Yamamoto, M., et al., *Role of adaptor TRIF in the MyD88-independent toll-like receptor signaling pathway*. *Science*, 2003. **301**(5633): p. 640-3.

**VITA**

Cemre Kocahakimođlu was born in Istanbul, Turkey on November 25, 1985. She graduated from Samsun Tlay Bařaran Anatolian High School with ranking 1<sup>st</sup> in the school in 2004. She received her Bachelor of Science degree from Izmir Institute of Technology (IZTECH) with ranking 1<sup>st</sup> in the chemical engineering department in 2009. She was also a member of IZTECH-Indoor Air Quality Research Group. After 2009, she has been a member of Erman Research Group and teaching-research assistant in the Chemical & Biological Engineering Department in Koc University.