LIGAND DOCKING TO PROTEINS UNDERGOING LARGE CONFORMATIONAL CHANGES

by

Damla Üçkan

B.S., Chemical Engineering, Istanbul University, 2013

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Chemical Engineering

Boğaziçi University

2019

ACKNOWLEDGEMENTS

I would like to thank my advisor Prof.Pemra Doruker in the very beginning since this work would not have been completed if it were not for her. No words can describe her endless support, always positive attitude and such valuable experience she was more than happy to share with me. Both in the courses and our thesis period, she was always there with her enormous academic knowledge together with her courage that always reminds me if you believe you can achieve. I will always express my gratitude and love to her in every single step of my life.

I would like to state my gratitudes also to Prof.Turkan Haliloglu for accepting to be my supervisor in the last stage and her supportive behavior which gave me more resistance through my researches. I am very thankful to Assoc. Prof. E. Demet Akten for her time by giving her feedbacks with valuable suggestions which helped me to improve this study.

I feel very lucky to have such friendly people in Polymer Research Center, all my friends were always so open to share the information and giving their feedbacks for my progress. Thanks to all of them for being with me.

A very special thanks to my one and only companion during my whole academic life, Doga Findik for being my closest friend and always beside me with his knowledge and computational support. Thanks to this period, we will be friends forever.

I would like to give my biggest thanks to my family for their unconditional love, support and encourage to keep my working and academic life in balance.

Last but not least, to all scientists I was inspired by their academic mind and knowledge during my research, I would like to show all my respect and gratitude for their contribution to the science world.

ABSTRACT

LIGAND DOCKING TO PROTEINS UNDERGOING LARGE CONFORMATIONAL CHANGES

Conformational transitions and ligand binding to the enzyme adenylate kinase were investigated by computational docking. Autodock was used to investigate the dynamics of 92 atomistic conformers, which have been previously obtained by a conformational sampling algorithm- ClustENM. First the internal geometry of the atomistic ClustENM conformers were validated by the MolProbity server. Four different ligands, namely AP5, ATP, ADP and AMP, were docked on each conformer to determine the favorable docking poses in terms of low energies and ligand RMSD values. Docking of ATP to the side of the LID domain gave the best results among all ligands, which is accompanied by LID closure to various extents. Minimum RMSD is obtained where ATP ligand docked to LID domain with 0.63 Å. For multiple conformers, successful poses were found for docking of any ligand to the LID side, whereas dockings to the side of the NMP domain did not yield any successful poses for any of the ligands. Results are mostly in compliance with GOLD procedure especially for the ATP case. To analyze the changes in dynamics for the protein-ATP poses, Dynomics web server was used. Moreover, key protein-ligand interactions were determined by Ligplot, which pointed to the interactions around the phosphate groups of the ligand in the closed crystal structure. Intermediate structures with half closed LID were successful for docking of the ligands, indicating a combination of induced fit and conformational search algorithms.

ÖZET

BÜYÜK YAPISAL DEĞİŞİKLİKLER GEÇİREN PROTEİNLERE LİGAND BAĞLANMA UYGULAMALARI

Adenilat kinaz enzimine ligand bağlanma ve buna bağlı yapısal değişiklikleri incelemek adına bir doking (bağlanma) programı kullanılmıştır. Daha önceki bir çalışmada ClustENM adlı konformer örnekleme algoritması ile elde edilmiş 92 ara yapının dinamikleri, Autodock programıyla incelenmiştir. Öncelikle ClustENM ile elde edilmiş bu yapıların iç yapı geometrilerini doğrulama amaçlı bir çalışma yapılarak, Molprobity server kullanılmıştır. En düşük enerji ve en düşük RMSD değerlerini verecek uygun bağlanma alanlarını inceleyebilmek için, AP5, ATP, ADP ve AMP olarak dört farklı ligand yapısı seçilmiştir. Tüm ligandlar için en iyi sonuçlar LID tarafında bir kapanmaya yönelen ATP ligandı ile elde edilmiştir. En düşük RMSD değeri yine ATP'nin LID tarafına bağlandığı durum için, 0.63 Å olarak belirlenmiştir. Farklı bir çok konformer için alternatif ligandların LID tarafına bağlandığı seçenekler başarılı sonuçlar verirken, aynı ligandların NMP tarafına bağlanmalarında başarı elde edilememiştir. Özellikle ATP ile yapılan çalışmalar, farklı bir bağlanma programı olan GOLD ile tutarlı sonuçlar vermiştir. Protein-ATP pozlarının dinamiğindeki değişiklikleri analiz etmek adına Dynomics server kullanılmıştır. Ayrıca, Ligplot ile kapalı kristal yapıdaki fosfat gruplarının etkileşimlerini işaret eden, önemli protein-ligand etkileşimlerini incelemek adına bir çalışma yapılmıştır. Yarı kapalı LID yönelimi gösteren ara yapıların başarılı sonuçlar vermesi, kilit-anahtar kuramı ve yapı arama algoritmaları arasında bir kombinasyon örneği gösteren bir bağlanma metodunu işaret etmektedir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

A_{ij}	Amber force field parameter
B_{ij}	Amber force field parameter
C_β	Carbon Beta
C_{ij}	Depth of O-H, N-H and S-H bonds
D_{ij}	Depth of O-H, N-H and S-H bonds
$E(t)$	Directionality of hydrogen bond
F	Free Energy Surface
I_L	Intermediate state in the L Basin
I_N	Intermediate state in the N Basin
k_B	Boltzman's constant
Q_{LC}	Fraction of Native Contacts between LID and CORE Domains
Q_{NC}	Fraction of Native Contacts between NMP and CORE Domain
$r_{0_{ij}}$	Residue pairs
S	Solvation parameter
$S_{(hb_{ext})}$	Protein-Ligand Hydrogen Bonding
$S_{(hb_{int})}$	Hydrogen Bond Interactions in the Ligand
$S_{(int)}$	Total intramolecular strain
$S_{(tors)}$	Torsional strain contribution
$S_{(vdw_{ext})}$	Protein-Ligand van der Waals scores
$S_{(vdw_{int})}$	Van der waals intramolecular strain contribution
T	Temperature
V	Volume
V_{bind}	Ligand Binding Energy
V_{bound}^{L-L}	Intramolecular energy of Ligand for Bound State
V_{bound}^{P-P}	Intramolecular energy of Protein for Bound State
V_{bound}^{P-L}	The Change in Intermolecular energies between the Bound State
$V_{unbound}^{L-L}$	Intramolecular energy of Ligand for Unbound State

$V_{unbound}^{P-P}$	Intramolecular energy of Protein for Unbound State
$V_{unbound}^{P-L}$	The Change in Intermolecular energies between the Unbound State
W_{elec}	Electrostatic interaction term
W_{hbound}	Hydrogen bond interaction term
W_{sol}	Desolvation potential
W_{vdw}	Dispersion/repulsion interaction term
Å	Angstrom
ΔG	Energy of Ligand and Protein in the Unbound State
ΔS	Free Energy of Binding State
σ	Interaction range
χ	The Function of Protein Conformational Change

LIST OF ACRONYMS/ABBREVIATIONS

ADK	Adenylate Kinase
ADP	Adenosine-5'-diphosphate
AMP	Adenosine monophosphate
ANM	Anisotropic Network Model
AP5	BIS(Adenosine)-5'-pentaphosphate
ATP	Adenosine-5'-triphosphate
BC	Bound Closed
BO	Bound Open
DIMS	Dynamic Importance Sampling
DNA	Deoxyribo Nucleic Acid
ENM	Elastic Network Models
GA	Genetic Algorithm
GNM	Gaussian Network Model
MD	Molecular Dynamics
Mg	Magnesium Ion
NMA	Normal Mode Analysis
PDB	Protein Data Bank
PMF	Potential of Mean Force
RMSD	Root-Mean-Square Deviation
sRMSD	Standardized Root-Mean-Square Difference
UC	Unbound Closed
UO	Unbound Open

1. INTRODUCTION

Proteins are macromolecules composed of amino acids and synthesized using the information encoded in genes. Proteins are key players in each and every process within the cell and their specific functions are dictated by the conformational dynamics and flexibility.

A protein or polypeptide coded by a gene and, is a linear chain of amino acid residues with a specific sequence called the primary structure. To investigate a protein's structure, mainly four distinct types of structures are considered; primary, secondary, tertiary and quaternary. While the secondary structure represents the local interactions and regularly folded units such as helices and strands, the tertiary structure is the three-dimensional folded structure of a polypeptide stabilized by nonlocal interactions such as salt bridges, disulfide bonds or hydrogen bonds. The tertiary structure, also called the native structure, is uniquely related and thus describes the biological function of a monomeric protein. Quaternary structure refers to a protein composed of more than one polypeptide chain in its functional form.

The relationship between protein structure and function can be described through the so-called conformational dynamics of the folded chain. The three-dimensional structures of conformations of the protein chain can be determined by experimental methods, such as x-ray crystallography and nuclear magnetic resonance techniques. However, these approaches give an idea about only few populated states but not the intermediate states (Grant *et al.*, 2010). In this perspective, computational methods, like molecular dynamic simulations (MD) or coarse-grained techniques, such as elastic network models (ENM), become a necessity in order to provide an explanation to the conformational transitions and dynamics at the molecular level.

As proteins bind to molecules, they may undergo small to large structural changes, which are called as conformational transitions. These transitions need to be elucidated in detail in order to be able to control the function of a protein. The focus

of this thesis will be the enzyme adenylate kinase (ADK), which is considered as an appropriate model to study large scale conformational changes. The objective is to determine which conformations of the highly flexible ADK prefer to bind to its specific substrates/inhibitors. For this aim, ADK conformers previously generated by an unbiased ENM-based conformational search methodology (Kurkcuoglu, 2015), named as ClustENM, will be utilized as they can describe ADK conformational space from the open state to the closed one and vice versa. Docking of ADK's different substrates (AMP, ADP and ATP) and an inhibitor (AP5) will be performed to uncover each ligand's binding preferences in terms of different flexible domains and conformations. ADK presents a challenging case of a highly flexible receptor for assessing the overall performance of the ClustENM method for docking purposes.

2. LITERATURE SURVEY

2.1. Binding Procedure

The ligand-binding process or more specifically substrate-enzyme recognition is currently viewed as a combination of pre-existing equilibrium and post-binding rearrangements, as reviewed by Bahar and coworkers (Bahar *et al.*, 2007; Tobi and Bahar, 2005). As Figure 2.1 shows pre-existing and post binding stages visually, in pre-existing equilibrium, apo-protein originally samples an ensemble of conformations, which in a dynamic equilibrium before any binding processes. As the protein is flexible, it can be observed with many conformational states in apo-form, whereas ligand prefers to bind one of these conformers to have optimal interactions and form a complex structure. The stabilization of the final complex is subject to further arrangements, called induced fit (Bahar *et al.*, 2007; Tobi and Bahar, 2005).

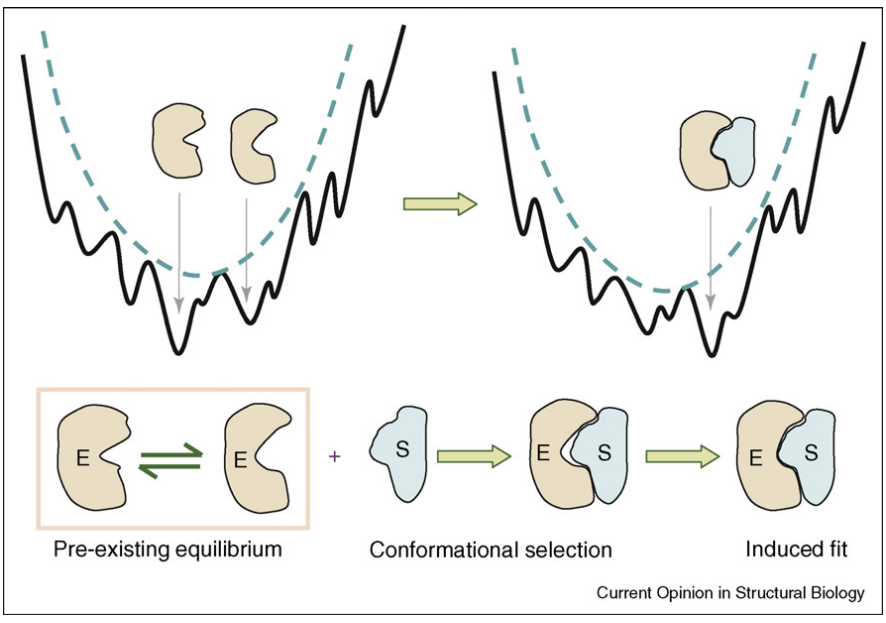


Figure 2.1. Pre-existing and post binding rearrangements (Bahar *et al.*, 2007; Tobi and Bahar, 2005).

By a detailed looking to binding procedures, there are two different and more detailed concepts raised and reformed the subject with new perspectives.

The main concept of biological network studies, which is called allostery, is mainly defined as the coupling between ligand binding and protein conformational changes. There are two representative models to describe these conformational changes, as induced-fit and population-shift models.

In Figure 2.1, induced-fit mechanism is roughly described, however in a different study by Okazaki and Takada (2008), this mechanism is examined in details together with population-shift model. Further investigations are also studied by Whitford *et al.* (2012). The simplest explanation of shifting model is proteins in the unbound states exist in both open and closed conformers, and during the binding process the dominant population shifts from open state to closed one. This mechanism arisen from the Monod-Changeux model of allostery. Recent studies revealed that depending on the ligand structure, the binding mechanism can change between these two models, since the induced-fit model was mainly suggested for protein-protein and protein-DNA binding, whereas population-shift model was offered for antigen-antibody binding and substrate binding to enzyme structures.

To investigate these binding models, protein dynamics needs to be understood in detail, which corresponds to global energy landscape in this context, and it is roughly described with Figure 2.2. below;

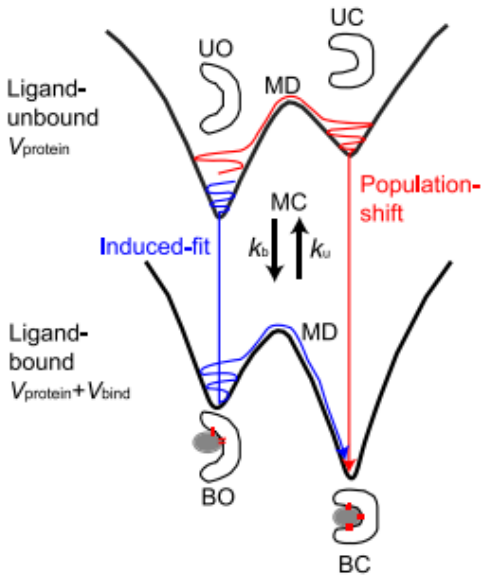


Figure 2.2. Energy landscape model.

In energy landscape funnel model, there can be multiple minima which protein can change its conformation (Frauenfelder *et al.*, 1991; Gershenson *et al.*, 2014; Leach and Leach, 2001; Onuchic *et al.*, 1997). From unbound-open to bound-open state, induced-fit scenario is arisen by considering a protein is at unbound state in the beginning and jump into bound state which is followed by a conformational change to reach the bound-closed state in the end.

In the population-shift scenario, there are two different unbound states as open and closed ones. From unbound closed state, a protein can bind with a ligand and reach bound-closed state with the minimum energy required.

The question here is that what kinds of binding processes prefer to proceed with induced-fit or population-shift mechanism. A series of molecular simulations is performed by using glutamine as a binding model (archetypical hinge-bending motion) by Okazaki and Takada (2008), and two-dimensional free-energy surfaces are investigated by looking the function of protein conformational change (χ) and the ligand-binding energy (V_{bind}) in Figure 2.3 below;

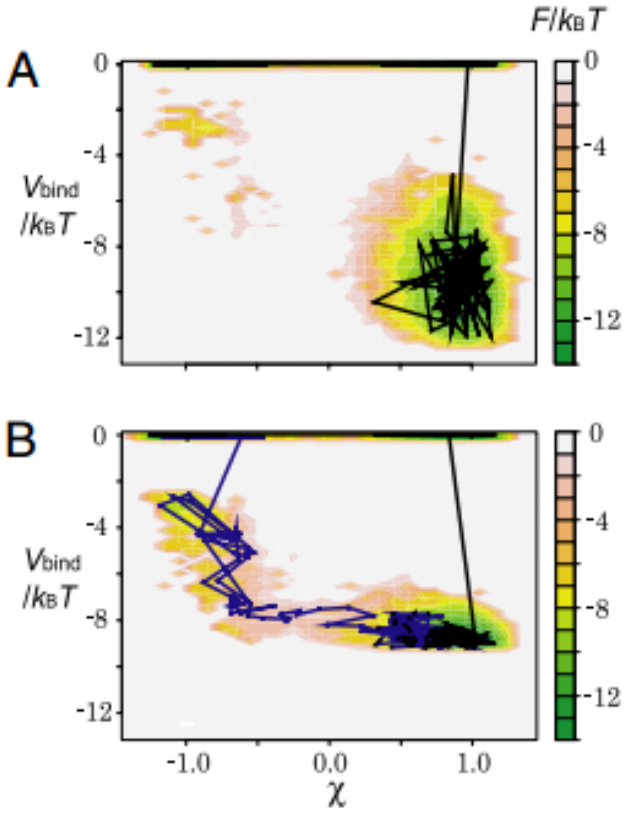


Figure 2.3. Function of protein conformational change and binding energy.

The main point for Figure 2.3 is, in the case (A), short-ranged of interactions are applied with $\sigma = 0.05r_{0_{ij}}$ whereas in case (B) with longer and stronger interactions, $\sigma = 0.15r_{0_{ij}}$.

From the black line (representative trajectory) in case A, it can be seen that there is a direct shift from UC (unbound-closed) to BC (bound-closed) state which lead us to population-shift mechanism. On the contrary, in case B where longer-range interactions are applied, an alternative induced-fit pathway is observed and the binding is followed by the conformational change. It can be definitely said that binding process ended up with a conformational change in this case.

As a result of this study, weak and short-ranged ligand interaction results in the population-shift mechanism, whereas strong and long-ranged ligand interactions encourages the induced-fit mechanism during the simulations.

2.2. Ligand Structure

2.2.1. Adenylate kinase (ADK) conformational transitions

ADK is commonly considered as a model system for studying conformational transitions in terms of ligand binding. ADK is recognized as an essential enzyme that maintains the energy balance of free adenylate nucleotides AMP(Adenosine monophosphate), ADP(Adenosine-5'-diphosphate) and ATP(Adenosine-5'-triphosphate) within the cell by catalyzing the reaction below;

$$Mg^{2+}ATP + AMP \rightleftharpoons Mg^{2+}ADP + ADP \tag{2.1}$$

ADK has three domains that exhibit large motions between the fully open and fully closed states. CORE is the stable domain, while the other two domains close over it during binding. LID and NMP domains bind to substrates ATP and AMP, respectively. Such an open-to-closed transition is depicted in Figure 2.4. The left and right panels represent the respective fully open (4AKE(Müller *et al.*, 1996)) and fully closed (1AKE(Müller and Schulz, 1992)) crystal structures, which have a root mean squared distance (RMSD) of 7 Å in between. Full closure of LID and NMP domains are stabilized in the presence of the inhibitor AP5(Bis(adenosine)-5'-pentaphosphate), which means its two substrates. In fact, recent work has shown that ADK is capable of exhibiting large-scale conformational changes without any ligand binding(Hanson *et al.*, 2007; Kessel and Ben-Tal, 2010).

The conformational degrees of freedom of ADK is commonly described by two variables: LID-CORE and NMP-CORE angles, shown in Figure 2.5. These angles were first introduced by Beckstein *et al.* (2009) to explore the free energy surface of ADK. LID-CORE angle is generated from the centers of geometry of the backbone and C_β atoms based on residues 179-185 (CORE), 115-125 (CORE-LID) and 125-153 (LID).

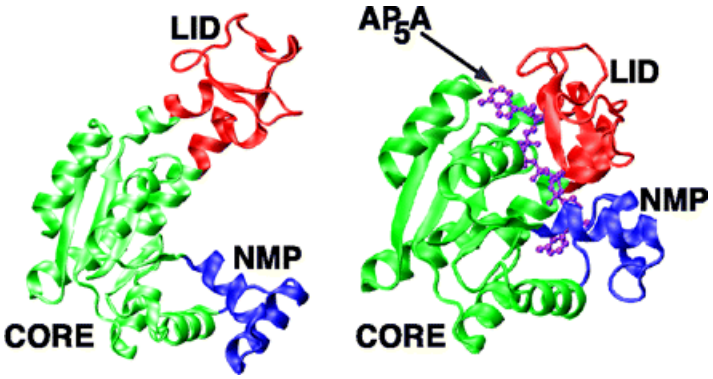


Figure 2.4. ADK in open and closed states (left and right, respectively) as observed in x-ray crystal structures (Arora and Brooks, 2007).

Whereas NMP-CORE angle is formed by using the same principle based on residues 115-125 (CORE-LID), 90-100 (CORE) and 35-55 (NMP).

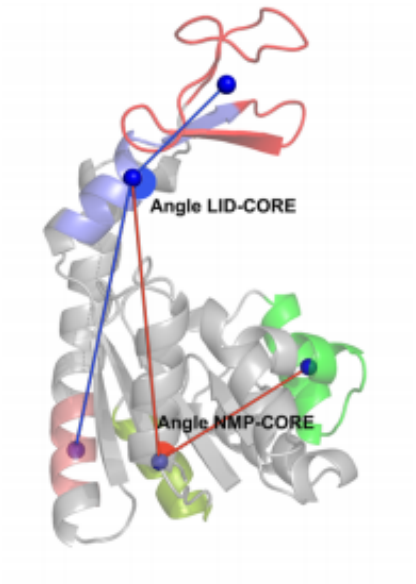
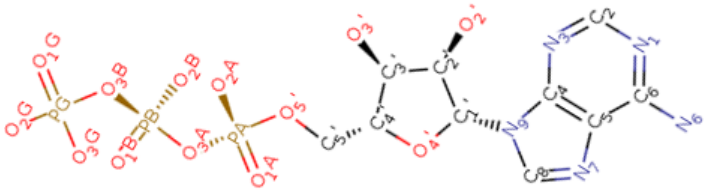
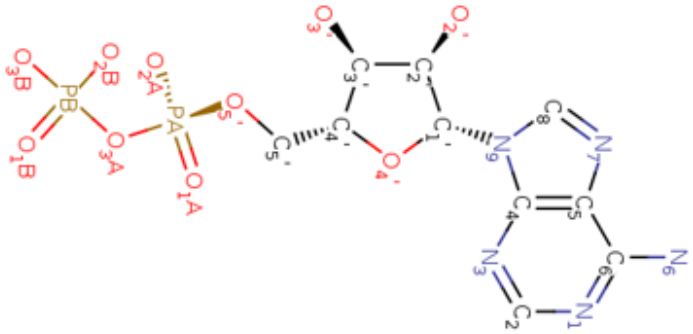


Figure 2.5. LID-CORE and NMP-CORE angles that define the conformational flexibility of ADK.

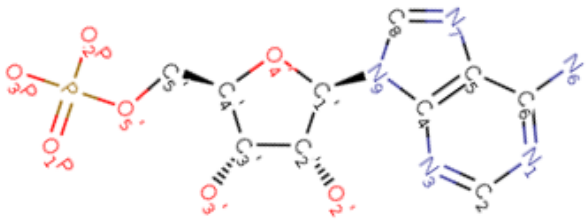
Three substrates of ADK based on the reversible reaction are AMP, ADP and ATP, for which chemical representations are provided in Figures 2.6(a), 2.6(b), 2.6(c).



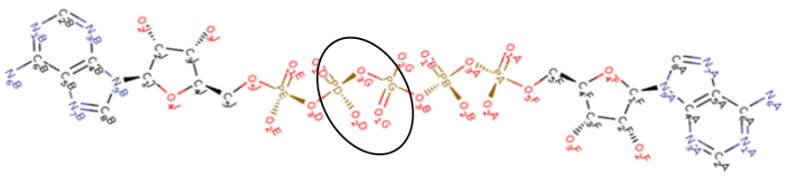
(a) ATP ($C_{10}H_{16}N_5O_{13}P_3$).



(b) ADP ($C_{10}H_{15}N_5O_{10}P_2$).



(c) AMP ($C_{10}H_{14}N_5O_7P_1$).



(d) AP5 ($C_{20}H_{29}N_{10}O_{22}P_5$).

Figure 2.6. Molecule representations of ATP, ADP, AMP and AP5.

The fourth structure in Figure 2.6(d) belongs to ligand AP5. As it can easily be seen AP5 can be used to generate ATP and AMP structures, by breaking a phosphate bond from the circled area. Thus, the position of the ligands ATP and AMP, which are missing in available crystal structures of ADK, can be generated with this principle.

2.3. Functional Dynamic Models

2.3.1. Go Model

In a recent study by Wang *et al.* (2012) explicit modeling of ligands was performed to study the functional dynamic properties of adenylate kinase. A coarse-grained model is used for ADK, where each amino acid is represented by one bead or two beads dependent on their location and properties. Go model (Hills and Brooks, 2009), constitutes a smooth and funnel-like approximation to the folding landscape, is used to investigate the effect of binding affinity of different ligands (Wang *et al.*, 2012). Simulations are performed by Gromacs 4.0.5. (Hess *et al.*, 2008) on four different models, namely Model A (ligand-free), Model M (with AMP), Model T (with ATP) and Model MT (with both ligands). Using explicit-ligand models, it is possible to investigate the effect of ligand-binding strength on the population distribution of ADK. With the increase of AMP and ATP binding strength, the population of the closed state increases, while the population of the open state decreases dramatically.

In Figure 2.7 (Wang *et al.*, 2012), the free energy landscape of apo ADK shows two intermediate states (L and N) and two major pathways. Q_{LC} indicates the fraction of native contacts between LID and CORE domains, whereas Q_{NC} is that between NMP and CORE domain. On the map, letters O and C correspond to the fully open and fully closed states, respectively. I_N is an intermediate state in the N basin, which represents a specific structure with closed LID domain and open NMP. For I_L, LID is open while NMP is closed. The 'I_N pathway' and the 'I_L pathway' are the corresponding pathways passing through I_N and I_L, respectively.

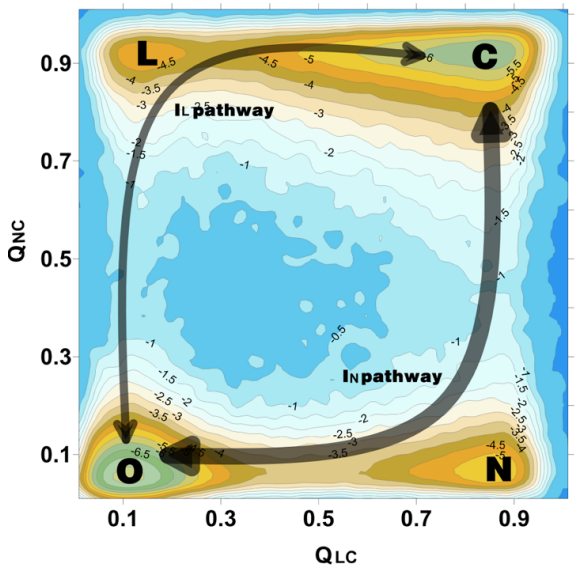


Figure 2.7. Free energy surface of apo ADK as a function of Q_{NC} and Q_{LC} and two major pathways (Wang *et al.*, 2012).

The findings of the study could be summarized as follows (Wang *et al.*, 2012);

The weight of I_N pathway decreased as AMP binding strength increased in Model M. Also in Model M, as NMP-Core interactions increasing, NMP domain tends to the closed state, whereas LID domain to the open state. In Model T, while ATP binding strength increased, the weight of I_L pathway decreased. It can also be observed in this model that during increment of LID-Core interactions, NMP domain towards to open state while LID domain to the closed one. In Model MT, the increase of closed population slows down as the ligand concentration increases, whereas when the binding strength increases, closed population reaches 90% of the population. In summary, this study revealed that the affinity between ligand and its objective enzyme has an important effect on the transition process to achieve the closed structure.

2.3.2. DIMS (Dynamic Importance Sampling)

In a different study by Beckstein *et al.* (2009), DIMS (Dynamic Importance Sampling) molecular dynamic method is developed for simulating conformational changes and applied to ADK transition. This method applies a known bias to the dynamics so that only "important" trajectories such as the ones those represent transitions are sampled. As a result, DIMS method can simulate rare transition events between two known structures, which cannot be easily obtained by conventional MD simulations. The advantage of DIMS algorithm against other conventional methods is the trajectories are indiscriminate and probabilistic, which give more accurate opinion than a single targeted pathway.

DIMS is used to generate a group of 330 trajectories between open and closed structures of apo-ADK. In Figure 2.8, some of these trajectories are shown for ADK, together with two-dimensional free energy surface or potential of mean force (PMF) (Beckstein *et al.*, 2009). Calculation of PMF is carried out by umbrella sampling based on two domain angles, namely NMP-CORE and LID-CORE angles, as the coordinates. These angles represent the main hinges in ADK structure, which are related to the motions of LID and NMP domains relative to slightly stable CORE domain. 45 ADK crystal structures are also presented on the PMF by circles and squares (end structures). DIMS pathways sample biologically relevant motions, which pass through available crystal structures and low energy regions.

Also, in Figure 2.8, PMF indicates two favorable regions which are parallel to LID angle axis. The area below 55^o of NMP-CORE angle is energetically more favorable. The movement of the LID domain is almost barrier-less at a certain level of NMP angle. Thus, the main barrier for the transition pathway in the apo state comprises of the hinge motion of NMP domain. Consequently, it can be seen that open and closed states consist a large range of domain movements especially in LID-CORE angle.

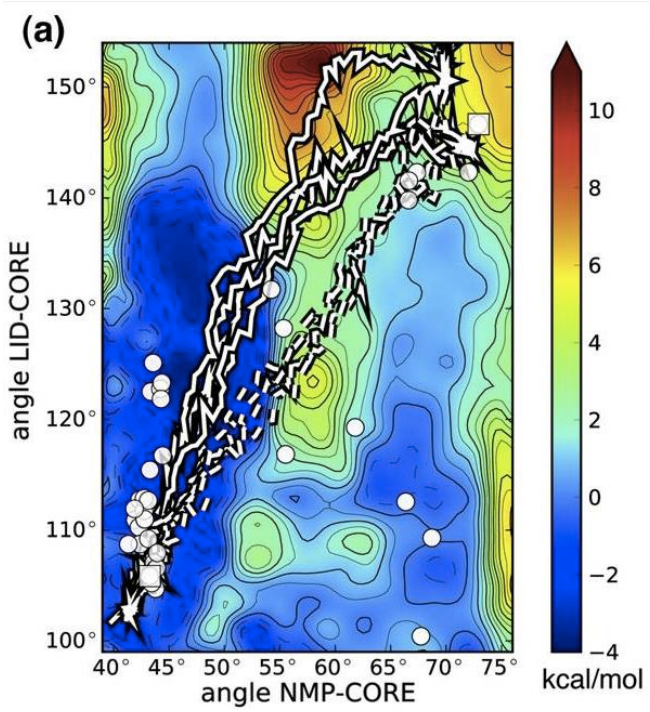


Figure 2.8. Trajectories for ADK, together with two-dimensional free energy surface or potential of mean force (PMF) (Beckstein *et al.*, 2009).

2.3.3. ANMPathway

ANMPathway is another method developed by Das *et al.* (2014) to generate an energetically favorable pathway between two endpoints corresponding to experimentally known structures. Two elastic network model representations related with the two endpoints are combined to construct a two-state potential, which presents a cusp hyper surface in the configuration space. In order to define the transition state, a structure which has equal energies from both surfaces is identified and treated as the transition state. By generating the steepest descent energy minimization trajectories starting from the transition state, conformers are obtained and the related transition pathway is constructed, which passes through low energy regions on the PMF.

2.3.4. ClustENM

In classical coarse-grained approaches as elastic network model (ENM), harmonic interactions are applied to get an idea about the dynamics of supramolecular systems (Atilgan *et al.*, 2001; Bahar *et al.*, 1997; Hinsen *et al.*, 1999)

In recent thesis study of Kurkcuoglu (2015), an ENM-based methodology, the so-called ClustENM, is developed for generating conformers accessible to an initial structure only. The conformers are generated in an unbiased iterative procedure, where the collective modes are combined for deforming the structure(s) with clustering and energy minimization routines. Figure 2.9. represents the probability of generated conformers based on apo ADK structure using LID and NMP angles described before. Highly populated regions are shown in red color, which correspond to low-energy regions on the PMF, indicating the reliability of atomistic conformers generated by ClustENM, because it is always expected to have highly populated regions with a minimum energy required to get an optimum binding.

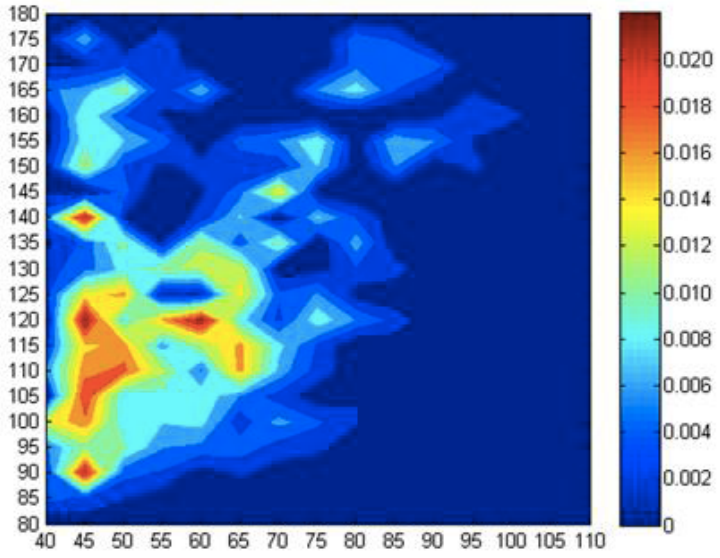


Figure 2.9. Probability of generated conformers for ADK based on LID-CORE (y-axis) and NMP-CORE (x-axis) angles (Kurkcuoglu, 2015).

Figure 2.9 also reveals that beside the closed LID and closed NMP states, there is a half and half open and closed state which is highly occupied within the angles NMP-Core: 55-60°and LID-Core: 115-120°. This state is also obtained by different computational and experimental studies (Seyler and Beckstein, 2014).

Another result that can be observed from Figure 2.9 is that the angles NMP-Core: 74°, LID-Core: 146°which correspond to the fully-open state is not populated. This result is also valid in different studies which claim that fully-open state is not favorable (Bhatt and Zuckerman, 2010; Daily *et al.*, 2010; Feng *et al.*, 2009; Kantarci-Carsibasi *et al.*, 2008; Lou and Cukier, 2006; Lu and Wang, 2008)

It is extremely difficult and generally not successful to dock a ligand to a protein undergoing conformational transitions by using its apo form only. This issue has been addressed by Flores and Gerstein, who developed a new methodology to generate ligand binding conformations for hinge bending proteins (Flores and Gerstein, 2011). Hinge-bending motions generally involve the slowest degree of freedom of a specific protein, however among the proteins those have been worked on that study, results are not satisfactory for ADK. Thus, in this thesis the aim is to focus on ADK's conformational transitions and evaluate the ClustENM conformers and conventional docking algorithms in this respect.

2.4. Conformer generation by ClustENM

ENM is a coarse-grained normal mode analysis, which efficiently calculates the large-scale collective modes of folded proteins, which are known to be correlated with the conformational changes experienced by the apo protein (Tama and Sanejouand, 2001). It was first formulated by Tirion (1996) and later extended to low-resolution versions (Atilgan *et al.*, 2001; Bahar *et al.*, 1997; Doruker *et al.*, 2000). In elastic network model approach, each amino acid is represented by a node located at its alpha carbon and close-neighbouring node pairs (only the ones within the bounds of a certain cut-off) form connected by harmonic springs, as a result forming an elastic network.

ClustENM is an ENM-based conformer generation method developed in Kurkcuoglu (2015)'s Ph.D. thesis and generating procedure of ADK conformers is represented in Figure 2.10 as a schematic version. This method is applied to generate atomistic conformers for proteins with large conformational changes (Kurkcuoglu *et al.*, 2016; Kurkcuoglu and Doruker, 2016). ClustENM methodology starts with the application of ENM on an energetically minimized native structure to extract the global modes (Case *et al.*, 2012; Duan *et al.*, 2003; Hawkins *et al.*, 1995, 1996). New structures are obtained by deformation of the structure along a combination of few collective modes. Next step is to cluster these generated structures and minimize the energy of representative structures. A 'generation' is formed by these conformers and same steps are performed to obtain new generations.

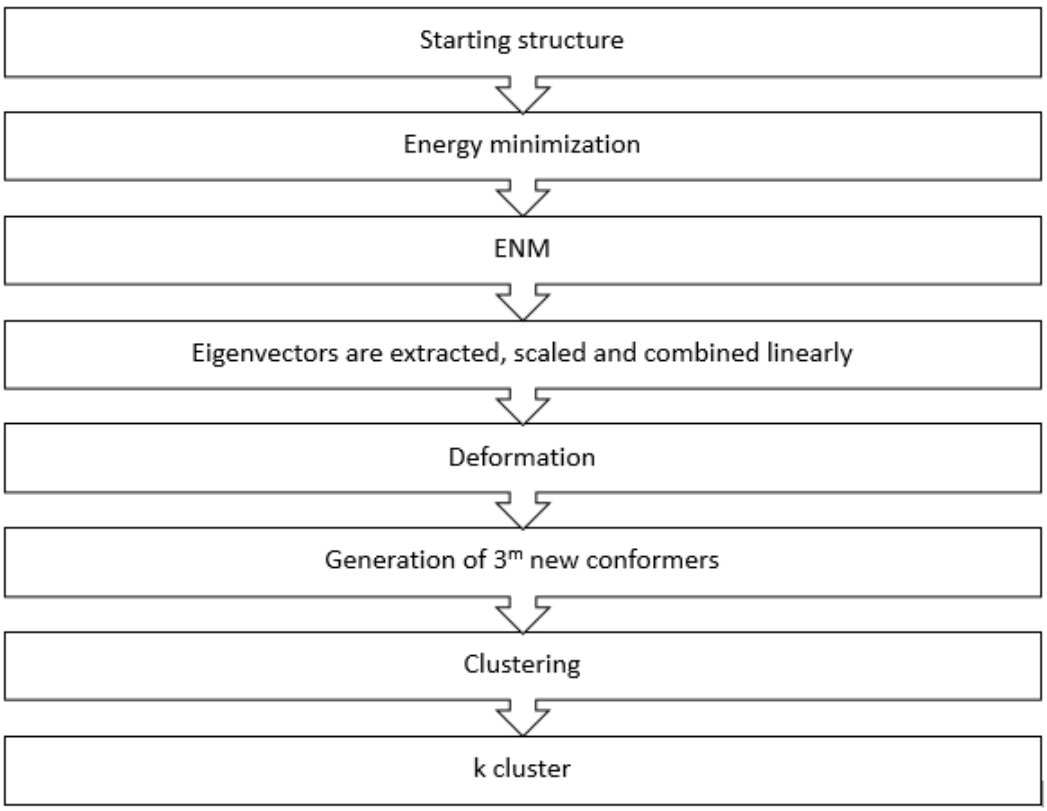


Figure 2.10. Schematic representation of ClustENM procedure.

2.5. Docking Procedure

Docking simulations require two steps; a search method and the force field employment.

2.5.1. Search Method

A search method to explore conformational spaces which are applicable for that specific system, is required for a general docking procedure.

In the basis of the search methods, the aim is to find specific arrangements for ligand and protein by looking at genetic algorithm translation, orientation and conformation of the ligand. These arrangements can be considered as state variables of the ligand and each represents a 'gene'. In the end, the concept leads us to natural genetics subject which includes genotype and phenotype terms by considering also atomic coordinates of the relevant structures.

With the improvements on docking operations, many different methods were suggested to improve the efficiency of genetic algorithm and a local search method is risen to realized that purpose (Morris *et al.*, 1998). Further suggestions came up by developing an adaptive local search method based on previous work by Solis and Wets (1981), a new hybrid approach arisen which is called Lamarckian genetic algorithm and has better performed in comparison with GA alone.

2.5.1.1. Lamarckian Genetic Algorithm. In most of the genetic algorithms, Mendelien genetics are used with the basis of Darwin evolution characteristics. Figure 2.11 indicates a specific mapping between genotype and phenotypes by comparing Darwinian and Lamarckian search algorithms;

In Figure 2.11, the transfer from genotype to phenotype is seen in the right hand side, this part represents Darwinian algorithm. When an inverse mapping can occur

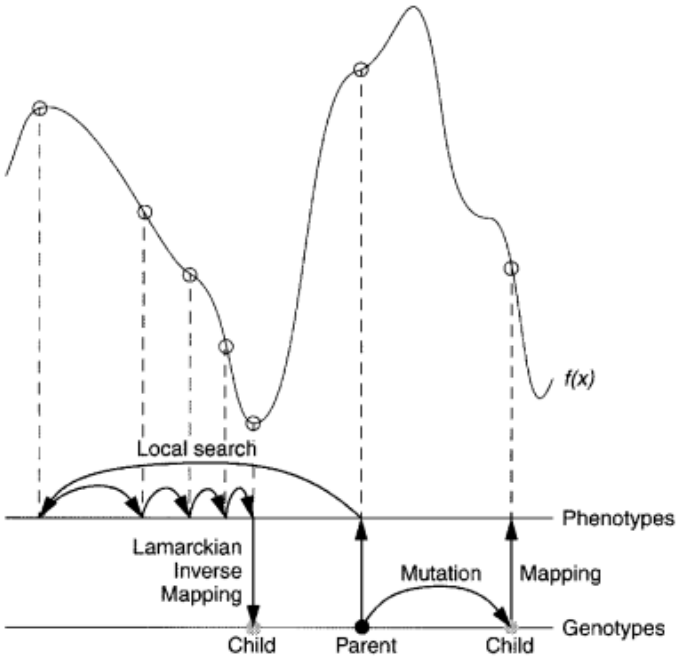


Figure 2.11. A specific mapping between genotype and phenotypes by comparing Darwinian and Lamarckian search algorithms (Morris *et al.*, 1998).

(left-hand side of the figure), a given phenotype leads us to its specific genotype, this is called Lamarckian genetic algorithm which is an implicity to Jean Batiste de Lamarck's.

In the right hand side, two options which lead a genotype to its specific phenotype are shown as a general transfer and secondly the result of a genetic mutation on the parent's genotype with the corresponding phenotype. In local search part, by using an inverse mapping function, it is aimed to reach a specific local minimum and at that time from a given phenotype, corresponding genotype arises. To reach this local minimum, number of iterations are applied, however in molecular docking operations local search is considered as a continuous study, so inverse mapping is not necessary in such studies. In the end of the process, parent's genotype replace with the resultant genotype.

This method, i.e. Lamarckian genetic algorithm is seen as the most efficient search method within many theoretical studies (Huey *et al.*, 2007).

2.5.2. Force Field

After a search method approach to explore available conformational spaces, a force field is employed to investigate the energetics of each individual conformation.

Binding evaluation in terms of force field approach is shown in Figure 2.12 below. In the first step of the binding evaluation by the force field, the intramolecular energetics is investigated from the transition of unbound states of ligand and protein separately. The intermolecular energetics of the combined ligand and protein structure (bound state) is evaluated as the second step and the free energy of the binding process is evaluated as the difference between these two steps.

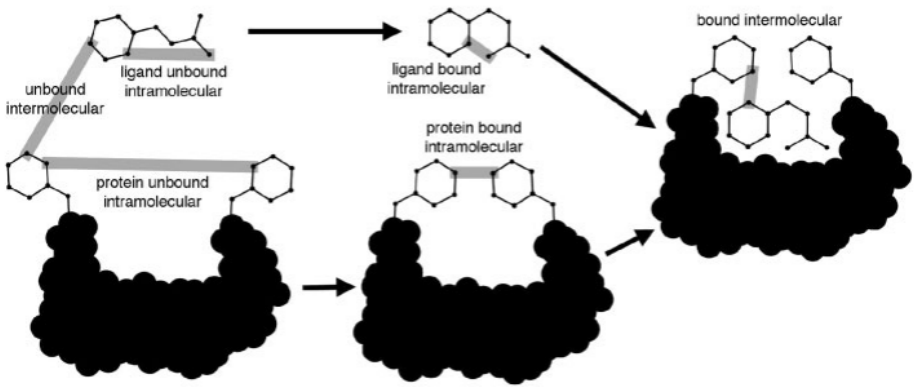


Figure 2.12. Force field approach to binding evaluation(Huey *et al.*, 2007).

$$\Delta G = (V_{bound}^{L-L} - V_{unbound}^{L-L}) + (V_{bound}^{P-P} - V_{unbound}^{P-P}) + (V_{bound}^{P-L} - V_{unbound}^{P-L} + \Delta S) \qquad (2.2)$$

The first four terms are the intramolecular energies for the bound and unbound states for the ligand and the protein, respectively. The third parenthesis includes the change in intermolecular energy between the bound an unbound states.

$$V = W_{vdw} \sum_{ij} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) + W_{hbound} \sum_{ij} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + W_{elec} \sum_{ij} \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} +$$

$$\mathrm{W}_{sol} \sum_{ij} (S_i V_j + S_j V_i) \mathrm{e}^{(} - r_{ij}^2 / 2\sigma^2) \tag{2.3}$$

ΔG is the energy of the ligand and the protein in their unbound states, separately. The difference between ΔG and the energy of the ligand-protein complex is equal to free energy of binding state. The pair-wise terms contain the terms for dispersion/repulsion, electrostatics, desolvation and hydrogen bonding. The first term in Equation 2.3 is coming from dispersion/repulsion interactions, parameters A and B are indicated in the Amber force field by Weiner *et al.* (1984). Hydrogen bond interactions are indicated with the second term and C and D are appointed to the depth of O-H, N-H and S-H bonds. The term E(t) indicates the directionality of the hydrogen bond and it is dependent on the distance from bonding geometry. With a screened Coulomb potential,indicated in the study by Mehler and Solmajer (1991), electrostatic interactions are utilized in the third term, and the final term is the desolvation potential depending the atom volumes(V), together with a solvation parameter(S) and an exponential factor which is driven by the distance. In order to choose the available conformers of any ligand-protein docking system, it is also important to investigate energetics of these conformers, since the best option includes a conformational state with a minimum energy requirement.

3. COMPUTATIONAL METHODS & MATERIALS

In this thesis, atomistic ADK conformers generated by ClustENM and some crystal structures of ADK will be used for ligand docking by Autodock, which will also be explained in the following sections.

3.1. Docking Procedure

Docking is a method used in molecular modeling for investigating the possible poses/orientations, in which a ligand molecule binds to a receptor protein to form a complex structure. Docking applications play a significant role for predicting the affinity of the molecules and very important for drug design to understand the reaction behind any disease and its corresponding treatment.

Docking can simply be applied rigidly as in 'lock and key' model, where 'key' corresponds to 'ligand' and 'lock' to 'protein'. More realistically, the protein and/or ligand can be considered flexible, which means that the protein and/or ligand can arrange their structures to obtain the 'best-fit' conformation. Docking studies are based on molecular recognition process and aspire to achieve an optimized conformation between ligand and protein, and minimize free energy of the overall system. They are mainly based on a search algorithm to determine the orientation of the protein-ligand complex and a force field or scoring function that detects the energy of each pose as mentioned in detailed in previous sections. Autodock and GOLD softwares will be used in this work (Huey *et al.*, 2007; Jones *et al.*, 1997).

3.1.1. Autodock

In Autodock's approach, the difference between separate energies of protein and ligand in unbound state and the energy of ligand-protein complex is taken equal with the free energy of binding state. To evaluate this approach, intramolecular energy of the transition from unbound to bound state is measured separately and intermolecular

energies are measured to form a bound complex (Goodsell *et al.*, 1996; Morris *et al.*, 1998, 2009; Trott and Olson, 2010).

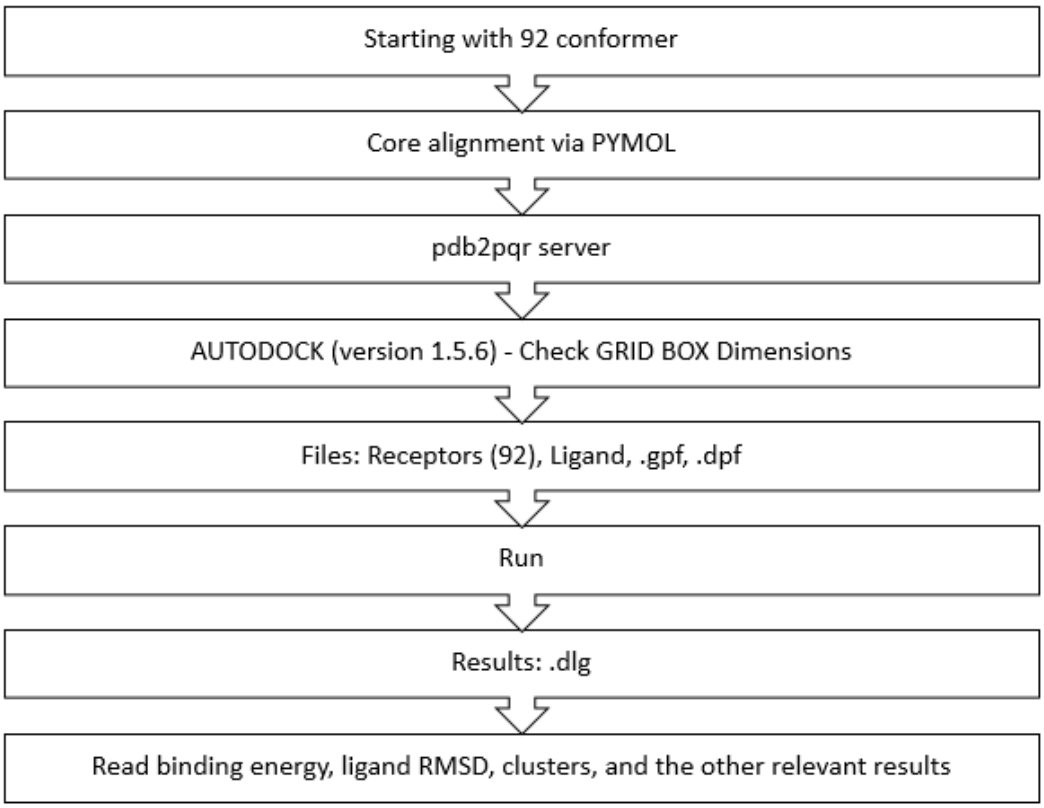


Figure 3.1. Schematic representation of computational steps for Autodock.

3.1.1.1. Core Alignment by Pymol - sRMSD Calculation. In the first step of the docking procedure, core alignment is applied to all 92 intermediate structures in order to get reference ligand RMSD distances in comparison with the closed state, 1AKE. These data will be a column of all our result tables in the next sections.

-Why we chose core alignment instead of overall alignment?

The overall alignment results will also be indicated in the tables, however an alignment to the core domain which is the most stationary among three of them (LID-CORE-NMP),is also necessary to get more accurate results.

By creating sRMSD figures and reading all the data, core alignment ligand RMSD values are taken into consideration as these are more accurate to understand whether we have the similar structure to the closed one with an optimum binding energy or not. If all the tables are investigated carefully, it can be seen that there is always a difference between overall and core aligned RMSD values and it is better to evaluate intermediate states through the stationary (core) domain.

To get even more specified results, it is aimed to see RMSD differences between 1AKE and 92 conformers, for LID and NMP domains separately. In order to read these regional RMSD values Pymol, is used and consequently, sRMSD figures for all the cases mentioned in Section 4 are generated. First, specified regions (residues) are marked both in 1AKE and each conformer one by one, than with the help of command rms_cur, all RMSD values are obtained for each specific region.

3.1.1.2. Setting up Box Dimensions. In order to investigate specific bindings to specific areas, dimensions are selected for each cases considering their targeted parts;

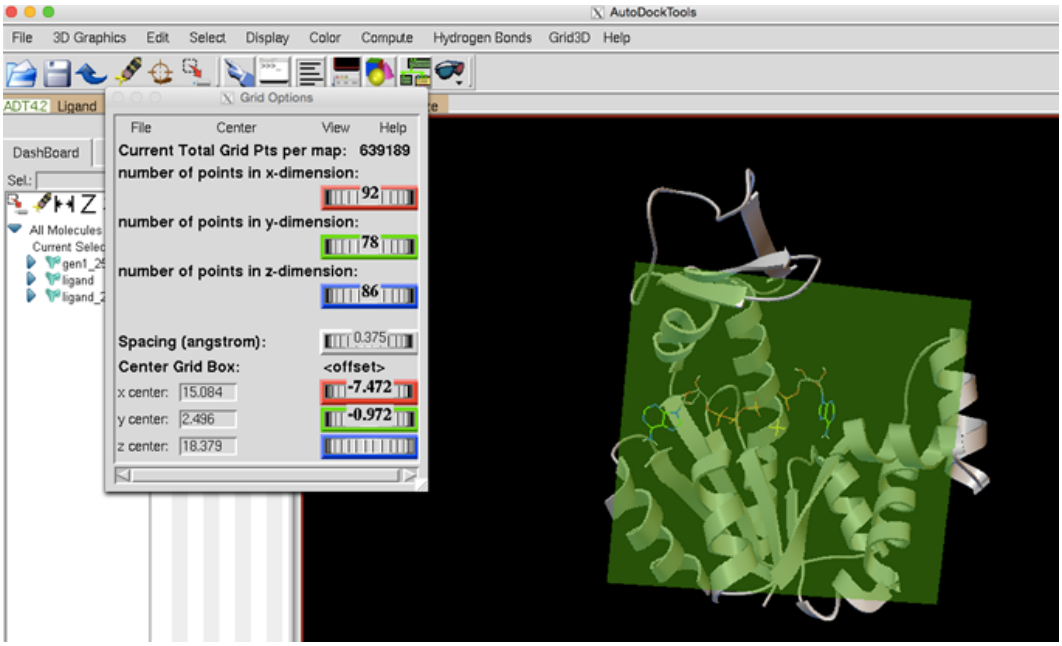


Figure 3.2. Box dimensions of ATP_Blind case (box is large).

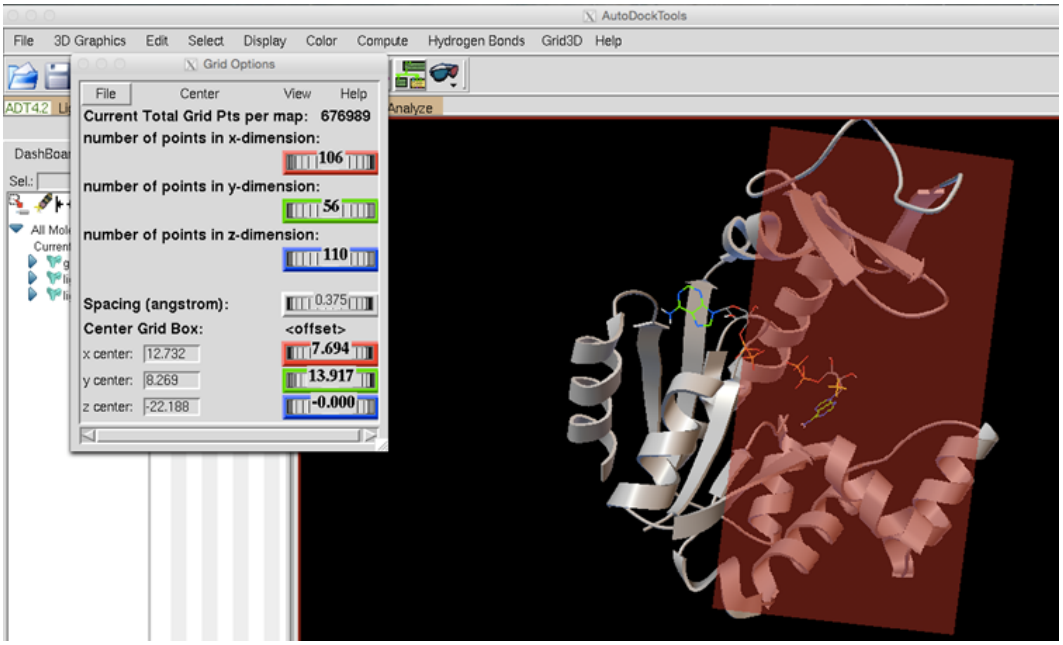


Figure 3.3. Box dimensions of ADP_NMP case (box is quite specific).

3.1.1.3. Preparation of Files (Data Loading). Autodock only requires .pdbqt version as starting files, however after the alignment procedure by Pymol, we still have .pdb files. By loading all the files to the server and to get .pqr files is the second step through the docking procedure. It is important to make this change with pdb2pqr server to determine the protonation states and to add hydrogen atoms to the required residues. Grid section of Autodock program helps to convert the .pqr file to .pdbqt version. Once all necessary files are ready i.e. conformers.pdbqt, ligand.pdbqt, .gpf (grid parameter file)(Figure 3.4(a))and .dpf (docking parameter file)(Figure 3.4(b)), result file (.dlg) can be obtained to read binding energy and ligand RMSD data for each conformer.

```
npts 60 96 118                          # num.grid points in xyz
gridfld receptor.maps.fld               # grid_data_file
spacing 0.375                           # spacing(A)
receptor_types A C HD N NA OA SA        # receptor atom types
ligand_types A C NA OA N P HD           # ligand atom types
receptor receptor.pdbqt                 # macromolecule
gridcenter 22.591 3.479 18.379          # xyz-coordinates or auto
smooth 0.5                              # store minimum energy w/in rad(A)
map receptor.A.map                      # atom-specific affinity map
map receptor.C.map                      # atom-specific affinity map
map receptor.NA.map                     # atom-specific affinity map
map receptor.OA.map                     # atom-specific affinity map
map receptor.N.map                      # atom-specific affinity map
map receptor.P.map                      # atom-specific affinity map
map receptor.HD.map                     # atom-specific affinity map
elecmap receptor.e.map                  # electrostatic potential map
dsolvmap receptor.d.map                 # desolvation potential map
dielectric -0.1465                      # <0, AD4 distance-dep.diel;>0, constant
```

(a) Grid parameter file.

```
autodock_parameter_version 4.2      # used by autodock to validate parameter set
outlev 1                            # diagnostic output level
intelec                             # calculate internal electrostatics
seed pid time                       # seeds for random generator
ligand_types A C NA OA N P HD       # atoms types in ligand
fld receptor.maps.fld               # grid_data_file
map receptor.A.map                  # atom-specific affinity map
map receptor.C.map                  # atom-specific affinity map
map receptor.NA.map                 # atom-specific affinity map
map receptor.OA.map                 # atom-specific affinity map
map receptor.N.map                  # atom-specific affinity map
map receptor.P.map                  # atom-specific affinity map
map receptor.HD.map                 # atom-specific affinity map
elecmap receptor.e.map              # electrostatics map
desolvmap receptor.d.map            # desolvation map
move ligand.pdbqt                   # small molecule
about 24.6854 5.9589 19.5973        # small molecule center
tran0 random                        # initial coordinates/A or random
quaternion0 random                  # initial orientation
dihe0 random                        # initial dihedrals (relative) or random
torsdof 11                          # torsional degrees of freedom
rmstol 2.0                          # cluster_tolerance/A
extnrg 1000.0                       # external grid energy
e0max 0.0 10000                     # max initial energy; max number of retries
ga_pop_size 150                     # number of individuals in population
ga_num_evals 25000000               # maximum number of energy evaluations
ga_num_generations 27000            # maximum number of generations
ga_elitism 1                        # number of top individuals to survive to next generation
ga_mutation_rate 0.02               # rate of gene mutation
ga_crossover_rate 0.8               # rate of crossover
ga_window_size 10                   #
ga_cauchy_alpha 0.0                 # Alpha parameter of Cauchy distribution
ga_cauchy_beta 1.0                  # Beta parameter Cauchy distribution
set_ga                              # set the above parameters for GA or LGA
sw_max_its 300                      # iterations of Solis & Wets local search
sw_max_succ 4                       # consecutive successes before changing rho
sw_max_fail 4                       # consecutive failures before changing rho
sw_rho 1.0                          # size of local search space to sample
sw_lb_rho 0.01                      # lower bound on rho
ls_search_freq 0.06                 # probability of performing local search on individual
set_psw1                            # set the above pseudo-Solis & Wets parameters
unbound_model bound                 # state of unbound ligand
ga_run 100                          # do this many hybrid GA-LS runs
analysis                            # perform a ranked cluster analysis
```

(b) Docking parameter file.

Figure 3.4. Autodock parameter files.

3.1.1.4. Result Reading –Calculation of Ligand RMSD. Five main information are taken from the file .dlg and put into result tables; binding energy, ligand RMSD, cluster, poses and overall-backbone RMSD. Each 13 cases those will be analyzed in Section 4, has 92 .dlg files separately. An example of .dlg file is given Figure 3.5 below;

CLUSTERING HISTOGRAM

Clus-ter Rank	Lowest Binding Energy	Run	Mean Binding Energy	Num in Clus	Histogram 5 10 15 20 25 30 35
1	-3.82	62	-3.79	12	############
2	-3.43	71	-3.42	26	##########################
3	-3.39	79	-3.37	4	####
4	-3.35	28	-3.33	5	#####
5	-3.27	43	-3.25	3	###
6	-3.05	23	-3.01	6	######
7	-3.02	49	-3.02	1	#
8	-3.02	12	-2.99	7	#######
9	-2.91	36	-2.90	4	####
10	-2.89	25	-2.89	1	#
11	-2.89	97	-2.89	1	#
12	-2.84	80	-2.82	15	###############
13	-2.83	18	-2.83	2	##
14	-2.70	82	-2.69	2	##
15	-2.68	77	-2.68	1	#
16	-2.40	68	-2.34	5	#####
17	-2.36	73	-2.35	2	##
18	-2.34	69	-2.34	1	#
19	-2.30	54	-2.30	1	#
20	-1.90	3	-1.90	1	#

Number of multi-member conformational clusters found = 13, out of 100 runs.

RMSD TABLE

Rank	Sub-Rank	Run	Binding Energy	Cluster RMSD	Reference RMSD	Grep Pattern
1	1	62	-3.82	0.00	4.90	RANKING
1	2	26	-3.81	0.04	4.90	RANKING
1	3	63	-3.81	0.01	4.90	RANKING
1	4	11	-3.81	0.09	4.92	RANKING
1	5	35	-3.81	0.07	4.92	RANKING
1	6	94	-3.81	0.05	4.89	RANKING
1	7	20	-3.80	0.04	4.89	RANKING
1	8	4	-3.79	0.05	4.89	RANKING
1	9	16	-3.79	0.05	4.90	RANKING
1	10	1	-3.78	0.10	4.88	RANKING
1	11	61	-3.77	0.08	4.90	RANKING
1	12	78	-3.73	0.14	4.87	RANKING
2	1	71	-3.43	0.00	1.91	RANKING
2	2	74	-3.43	0.03	1.92	RANKING
2	3	42	-3.43	0.04	1.89	RANKING
2	4	24	-3.43	0.06	1.87	RANKING
2	5	81	-3.43	0.07	1.95	RANKING

Figure 3.5. .dlg file of gen1_1 from ADP_Blind docking.

3.1.1.5. Autodock Simulation Protocol. Starting from 92 ADK conformers generated by ClustENM, Lamarckian Genetic Algorithm is applied for all the cases for ligands below:

- AP5,
- ATP with 11 active bonds,
- ATP with 1 active bond,
- ADP,
- AMP with 7 active bonds and
- AMP with 1 active bond.

For each docking procedure; 100 runs have completed with 0.375 Å spacing, and 25,000,000 and 27,000 maximum number of energy evaluations and maximum number of generations, respectively. The ligands ATP, AMP and AP5 are taken from the closed structure of ADK, 1AKE and the ligand ADP from 4CF7. RMSD between these two monomers (1AKE, 4CF7) is 2.21 Å.

3.1.2. GOLD

As a second protocol to verify the results, GOLD docking program is chosen. Among all other docking programs like PRO_LEADS, FlexX and FRED, GOLD gives a bigger success rate in different works. In a recent paper written by Verdonk *et al.* (2003), even 68% success rate is obtained among a validation set of 305 complexes which is the best option among all the other docking protocols.

GOLD program includes two main parts; scoring function and search algorithm. It also considers a mechanism for locating the ligand in the binding site with a unique understanding, called fitting points. These fitting points mainly targeted to hydrogen bonding groups both on protein and ligand structure and they make a reorientation between acceptor and donor groups. This mechanism creates a degree of freedom in the binding site which plays an important role in resulted binding energy values.

Concerning scoring function called Goldscore function, this step is mainly based on a ranking level of different binding modes with below terms (Azam and Abbasi, 2013);

$$Fitness = S_{(hb_{ext})} + 1.3750 \times S_{(vdw_{ext})} + S_{(hb_{int})} + 1.0000 \times S_{(int)} \qquad (3.1)$$

$$S_{(int)} = S_{(vdw_{int})} + S_{(tors)} \tag{3.2}$$

where S_{hb_ext} is the protein-ligand hydrogen bonding and S_{vdw_ext} is the protein-ligand van der Waals scores. S_{hb_int} is the hydrogen bond interactions in the ligand, whereas S_{vdw_int} comes from intramolecular strain in the ligand because of its contribution to the fitness. In the last part of Gold docking procedure, a search algorithm is applied to get possible binding modes and project binding free energies in that specific docking. Genetic algorithm (GA) is the one which is used in Gold protocol and mainly affects the accuracy and the docking time with its parameters, the number of dockings and the number of GA operations.

Different than Autodock, starting files will be .mol2 this time both for receptor (protein) and ligand uploads. To compare the results accurately, binding site is selected exactly the same with grid file of AutoDock for each cases. GA runs is set to 20 and fitness function is chosen as CHEMPLP. Result file is '.rnk' which shows 20 poses with run numbers, together with their score and RMSD values.

The comparison table which includes exact data of this study will be given in Section 4.6.

4. RESULTS AND DISCUSSIONS

In the first section, the ClustENM conformers are assessed in terms of internal geometry. Then, docking results are presented in detail using the crystal structures and the ClustENM conformers. Dockings are mainly performed using the Autodock program and additional verification is provided by GOLD software.

4.1. Evaluation of ClustENM Conformers

All 92 ClustENM conformers were examined in terms of their sRMSD compared to the closed structure (1AKE). After alignment of the core regions in Pymol, the regional sRMSD's were evaluated for the LID domains (residues 122-159) of 1AKE with respect to those of the 92 conformers one by one. Same procedure is applied also for evaluating sRMSD for NMP domains (residues 30-59). The conformers from different ClustENM generations are shown in Figures 4.1(a) for generations 1 to 4 and 4.1(b) for generations 5 to 7. Each generation is distinguished using different markers in the figures and in the rest of the text. The starting open structure is indicated by the black diamond on the figure. The generated conformers are very diverse as the range of sRMSD are between 3.280-23.783 Å and 4.595-20.214 Å for LID and NMP domains, respectively.

In Appendix A, Table A.1, the naming of ClustENM conformers are provided together with LID-CORE and NMP-CORE angles. However, it was mentioned in Kurkcuoglu (2015)'s work that the LID-Core and NMP-Core angles do not uniquely define the domain positioning in ADK. Specifically, two conformers could have similar angle values, however their RMSD may be quite large. For example, the difference may arise from the positioning of the LID with respect to the CORE, as similar LID-CORE angles may be in opposite directions, such as one being more closed and the other more open with respect to the apo forms, leading to large RMSD differences. Therefore, sRMSD values that can clearly define domain closure will be used to determine the approach to the target closed state in this thesis, while some of the angle data will still

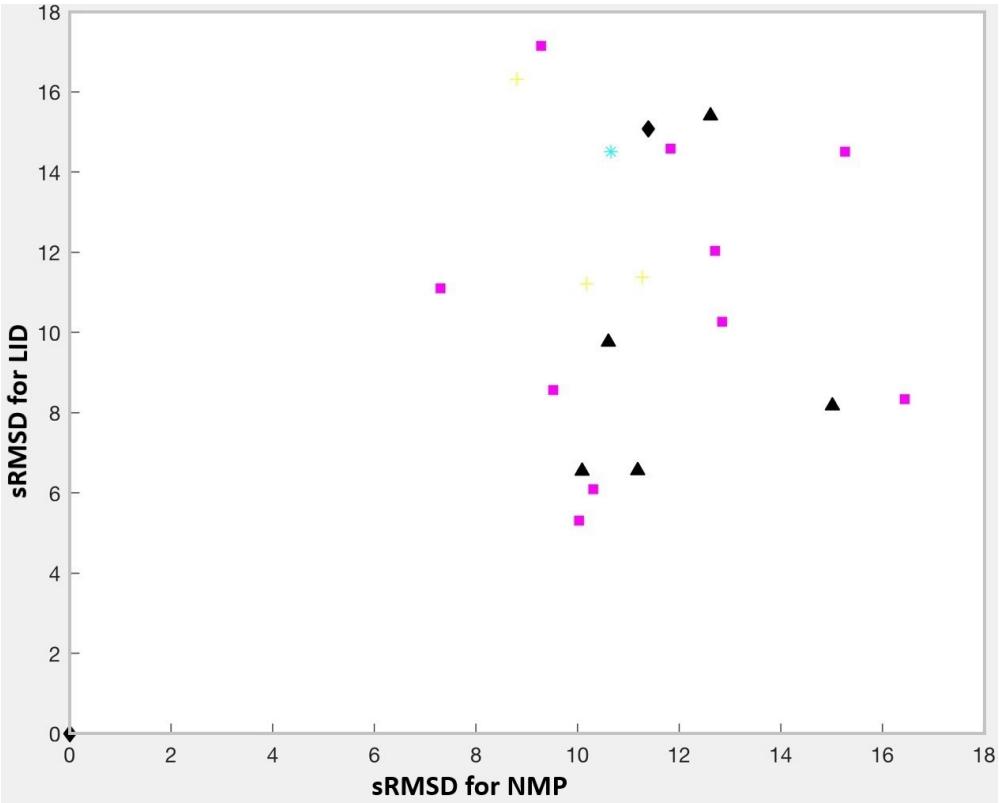
be provided in Appendix A.

Table 4.1. Details of ClustENM generations.

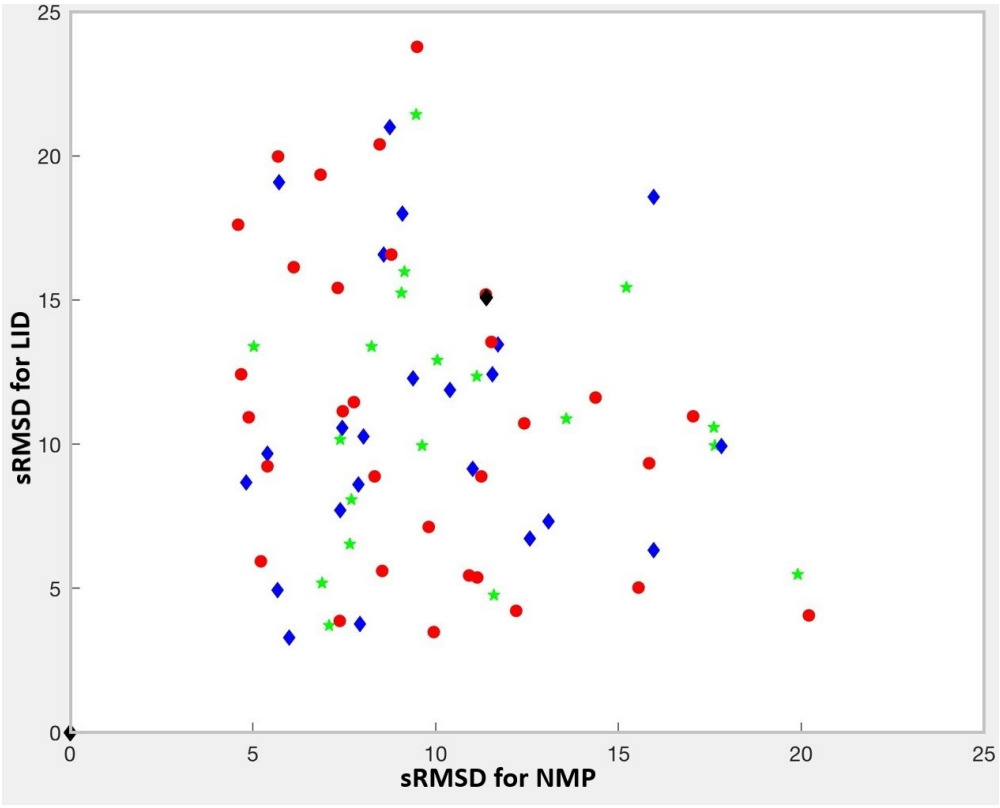
				sRMSD (Å), Min-Max	
Conformers	**No of Conformers**	**Marker**	**Color in Figure 4.1**	**NMP**	**LID**
4AKE	crystal	diamond	black	11.4	15.1
gen1	1	asterisk	cyan	10.6	14.5
gen2	3	plus	yellow	8.8-11.3	11.2-16.3
gen3	5	triangle	black	10.1-15.1	6.5-15.4
gen4	10	square	magenta	7.3-16.4	5.3-17.1
gen5	19	star	green	5.1-17.6	3.7-21.4
gen6	23	diamond	blue	4.8-17.8	3.3-20.9
gen7	31	circle	red	4.6-20.2	3.5-23.8

As the generations increase, more diverse conformers are obtained, as well as conformers that are closer to the closed structure, in terms of sRMSDs of both LID and NMP domains. In order to get a better idea about sRMSD levels in terms of closure, three intermediate structures are shown with different sRMSD values of 3.864, 8.593 and 15.412 Å for LID domain and 4.819, 8.327 and 15.022 Å for NMP. When investigating LID closure as shown in Figure 4.2, NMP sRMSD's are kept nearly same levels and same procedure is followed for NMP evaluation (Figure 4.3).

It is clearly observed that through sRMSD levels from 3.864 to 15.412 Å, there is an opening trend on each domains. It can be considered within the levels 3-4 Å,domains are closed. Within the range 8-10 Å,domains are half closed, and after a certain point as 15 Å, it can be considered as open domains.

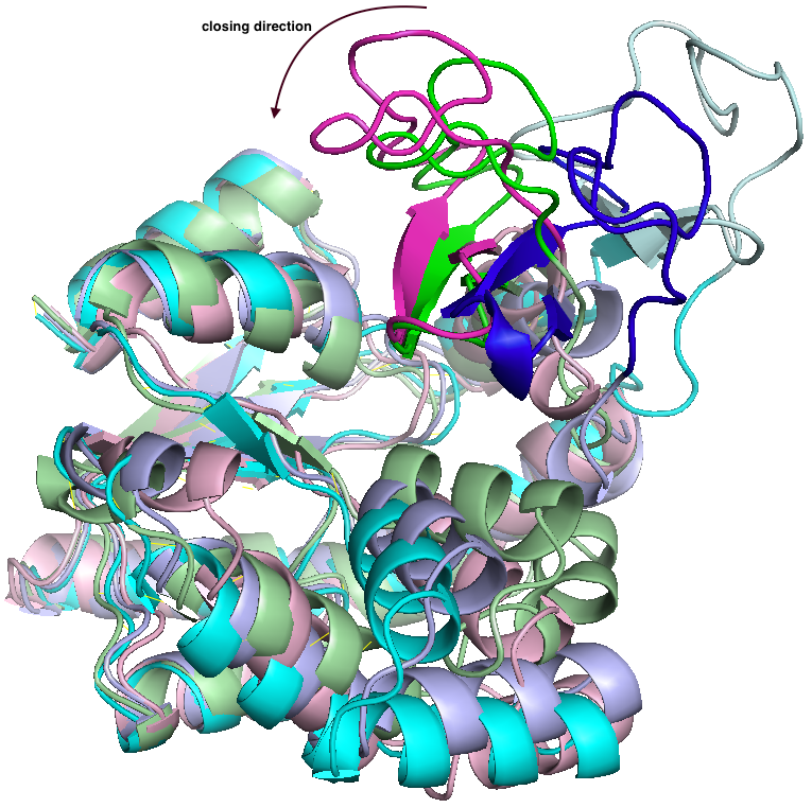


(a) Generations 1, 2, 3 and 4.

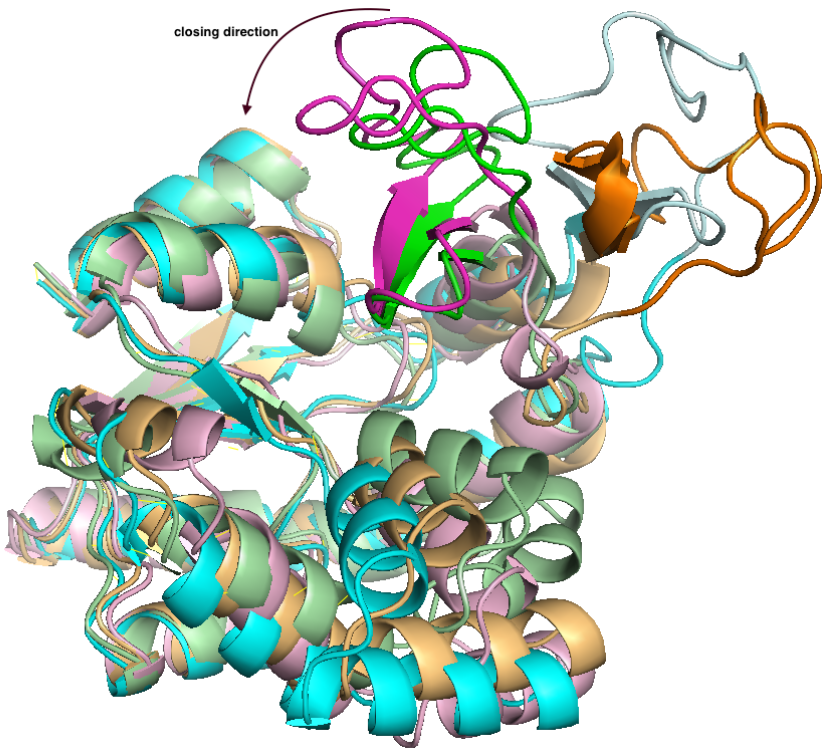


(b) Generations 5, 6 and 7.

Figure 4.1. sRMSD representation of whole generations.

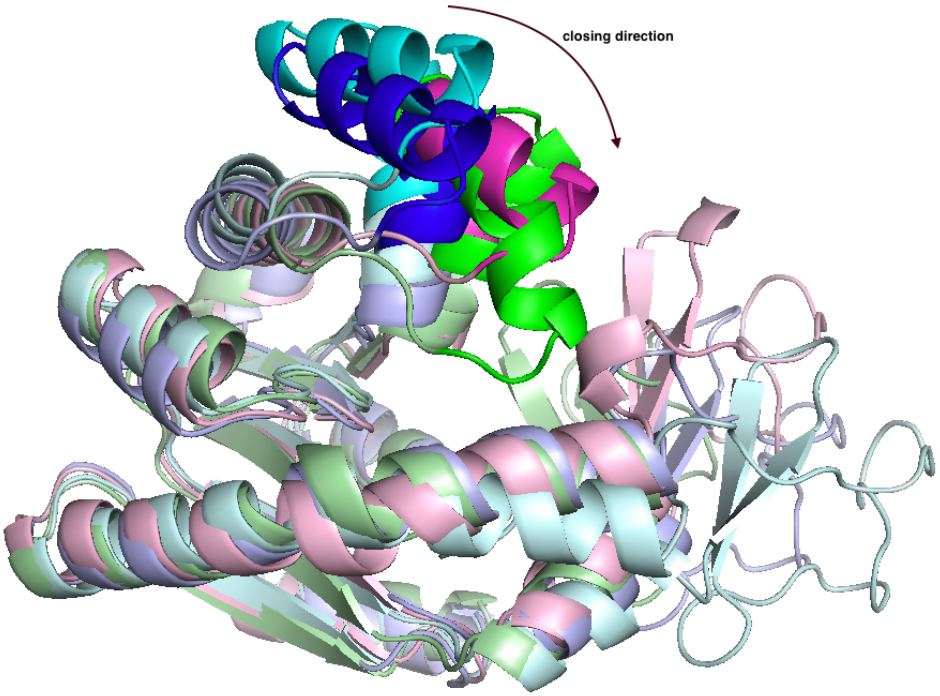


(a) gen7_27 with 3.864 Å(magenta) and gen6_17 with 8.593 Å(blue) sRMSD.

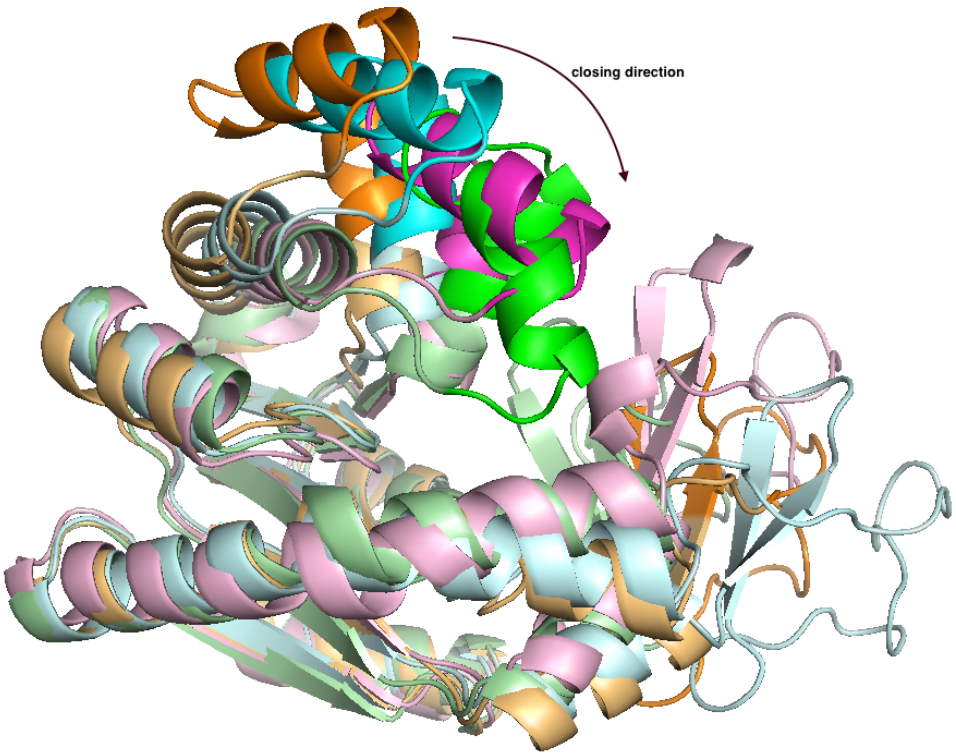


(b) gen7_27 with 3.864 Å(magenta) and gen7_30 with 15.412 Å(orange) sRMSD.

Figure 4.2. LID sRMSD levels upon closure. 1AKE(green), 4AKE(cyan).



(a) gen6_10 with 4.819 Å(magenta) and gen7_11 with 8.327 Å(blue) sRMSD.



(b) gen6_10 with 4.819 Å(magenta) and gen3_3 with 15.022 Å(orange) sRMSD.

Figure 4.3. NMP sRMSD levels upon closure.1AKE(green), 4AKE(cyan).

In this thesis, the aim is to assess all ClustENM conformers, also named as intermediates, in terms of docking/ binding to the substrates of ADK. Before docking, a structure validation web server, Molprobity, was used to check whether the internal geometry of these atomistic conformers are reliable (Chen *et al.*, 2010; Davis *et al.*, 2007), in terms of all atom contacts, Ramachandran plots and side chain rotamers. Molprobity assessment details can be found in Table A.2 for each conformer and the initial and reference crystal structures (4AKE and 1AKE).

When we compare the original crystal structures and their minimized versions from Table A.2, Ramachandran backbone torsion angles of the minimized versions are much better. The percentage of backbone torsions to fall into favored regions is expected to be above 98%. Additionally, the original crystal structures (1AKE and 4AKE) have higher number of poor rotamers than the minimized ones, as a result their overall Molprobity scores and percentiles are the lowest. If we set the minimized crystal structures as the reference, the ClustENM conformers seem to have satisfactory internal geometry. The Molprobity score is a composite one, with the lower values reflecting better geometry.

The number of poor rotamers is mostly 1 or 2, which is also the case in the crystal structures. Higher values of poor rotamers (3 or 4) are observed for only a limited number of conformers. The Ramachandran values falling in the favored region are quite good as well, which are close to the expected percentage of 98. Finally, the Molprobity scores of all conformers and minimized structures fall in the 100th percentile, indicating that the internal geometry of all atomistic conformers are satisfactory.

Table 4.2. Summary of Molprobity results.

	Molprobity Score Percental	**No of Poor Rotamers**	**Ramachandran Favored (%)**
1AKE_min	100th	2	99.01
4AKE_min	100th	1	98.52
gen1	100th	1	97.54
gen2	100th	1	94-98
gen3	100th	1	97-98
gen4	100th	1-4	96-98
gen5	100th	1-2	96-99
gen6	100th	1-3	96-99
gen7	100th	1-2	96-99

4.2. Docking to Crystal Structures

Initially dockings on reference crystal structures for ADK are performed in order to see if satisfactory results can be obtained and to have a reference for ClustENM conformer dockings. Table 4.3 summarizes the Autodock results on two complexes, namely 1AKE and 4CF7. 1AKE is bound AP5, which is an inhibitor mimicking its two substrates. As such, all three substrates, ATP, ADP and AMP, could in fact be generated from AP5. As there are no crystal structures bound to either ATP or AMP, 1AKE is used to obtain the position of ATP and AMP using Pymol in order to use them as reference for calculating ligand RMSDs.

By using Dali server (see Table A.3), structural alignment indicated that 4CF7 is similar to 1AKE and it is bound to two ADPs and Mg ion. We will utilize this structure for ADP dockings and also to assess the effect of Mg ion during the docking procedure.

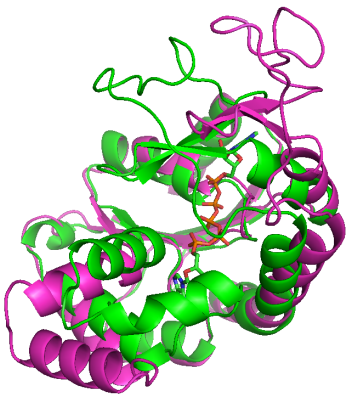
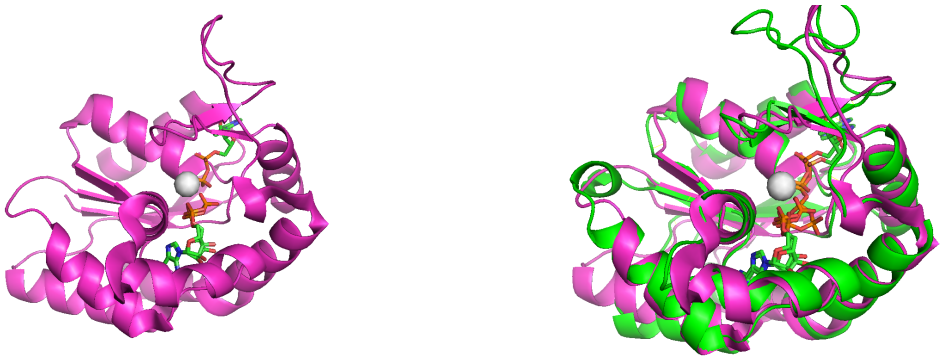


Figure 4.4. Alignment of 4AKE to 1AKE(Müller *et al.*, 1996; Müller and Schulz, 1992).



(a) 4CF7(Kerns *et al.*, 2015). (b) Alignment of 4CF7 to 1AKE.

Figure 4.5. Structural representation of 4CF7 and its alignment to 1AKE.

Table 4.3. AUTODOCK results for crystal structures.

Crystal	**LID**	**NMP**	**Energy (kcal/mol)**	**RMSD (Å)**
1AKE	AP5	-	-16.90	0.47
1AKE	ATP	-	-9.07	0.71
1AKE	AMP	-	-6.64	0.85
1AKE	-	ATP	-10.80	0.15
1AKE	-	AMP	-8.68	0.54
1AKE	-	ADP	-12.83	0.50
4CF7	ADP	-	-9.33	0.53
4CF7	ADP+MG	ADP	-5.91	0.55
4CF7	ADP	ADP	-10.48	0.51

The torsion angle between the adenine and ribose rings is considered flexible in all dockings, amounting to two flexible torsions in AP5 and one for ATP, ADP and AMP. The lowest energy is observed for the largest ligand AP5, which compliments the smallest RMSD value corresponding to the best pose. All ligands are independently docked either to the LID or NMP side. Both dockings give quite satisfactory ligand RMSDs that are less than 1 Å. Scores, i.e. energies, are comparatively lower on the NMP side for each ligand. As expected, the energy for the smallest ligand AMP is lower than the rest.

To investigate the effect of Mg^{+2} on docking results, 4CF7 is chosen as reference with two ADPs and Mg^{+2}. In the presence of ADP and Mg^{+2} bound to the LID, the second ADP is docked to the NMP side. When Mg^{+2} is also present in these dockings, much lower score is obtained indicating that its presence will not improve the dockings to ClustENM conformers. This result is also in line with a study by Zeller and Zacharias (2015) which is saying that with the presence of MG ion, ATP resulted a slightly more open ATP LID ensemble comparing with only ATP_LID case (with the absence of Mg ion).

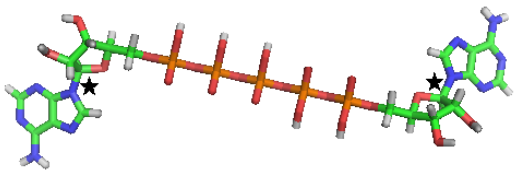
4.3. Docking to Intermediate Conformers

Table 4.4 summarizes the different type of dockings performed on ClustENM conformers, together with the box dimensions. The simulation box covers either the LID or the NMP domain binding site (Figure 3.3). A third case, where both binding sites are covered at the same time by using a larger box (Figure 3.2), corresponds to the blind docking case. Each docking is named according to these simulation box types.

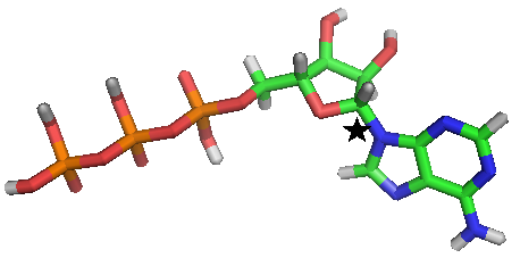
Moreover, the ligand flexibility may be different in each case. The torsion angle between the adenine and ribose rings is considered flexible in all dockings, which means two flexible torsions in AP5 and one for ATP, ADP and AMP. Such minimal number of flexible torsion angles, which are indicated by the stars in Figure 4.6 is preferable for easier sampling of protein-ligand conformations during docking. For comparison, further dockings with fully flexible ATP and AMP are also performed, which include 11 and 7 flexible torsion angles, respectively. Figures 4.7, 4.8 and 4.9 give the torsion tree representations for fully flexible ATP and AMP, and AMP with one flexible bond, respectively. These figures are from Autodock, where green torsion angles are flexible and the purple ones are kept fixed. Fully flexible ligand dockings are explicitly stated in Table 4.4 under docking type (first column). Moreover the reference structure for the complex (either 1AKE or 4CF7), which is used for calculating the ligand RMSD, is also listed in the last column.

Table 4.4. Dockings to ADK intermediate conformers.

Docking Type (Ligand_box)	**No of Flexible Bonds**	**Box Dimensions (Å×Å×Å)**	**Reference Complex**
AP5_Blind	2	80 × 80 × 90	1AKE
ATP_Blind	1	92 × 78 × 86	1AKE
ATP_Blind_Flex	11	6 × 78 × 78	1AKE
ATP_LID	1	60 × 96 × 118	1AKE
ATP_LID_Flex	11	60 × 96 × 118	1AKE
ATP_NMP	1	78 × 88 × 88	1AKE
ADP_Blind	1	110 × 110 × 110	4CF7
ADP_LID	1	96 × 56 × 110	4CF7
ADP_NMP	1	106 × 56 × 110	4CF7
AMP_LID	1	52 × 108 × 116	1AKE
AMP_NMP	1	78 × 96 × 118	1AKE
AMP_NMP_Flex	7	78 × 96 × 118	1AKE
AMP_NMP_wATP_Flex	7	78 × 96 × 104	1AKE



(a) AP5.



(b) ATP.

Figure 4.6. AP5 and ATP ligand structures by Pymol.

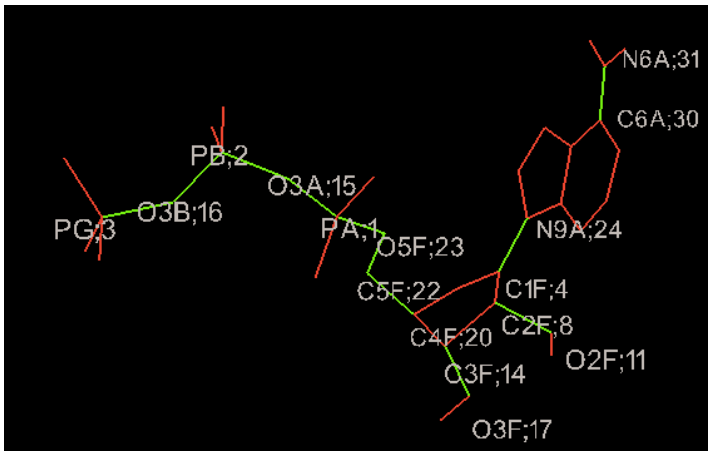


Figure 4.7. Torsion Tree representation of fully flexible ATP by Autodock.

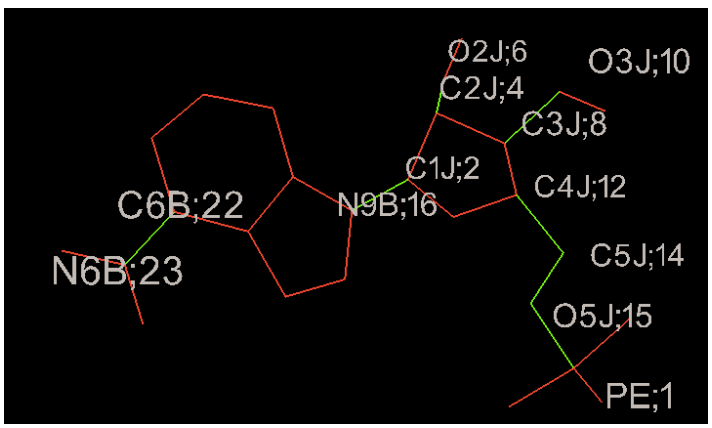


Figure 4.8. Torsion Tree representation of fully flexible AMP.

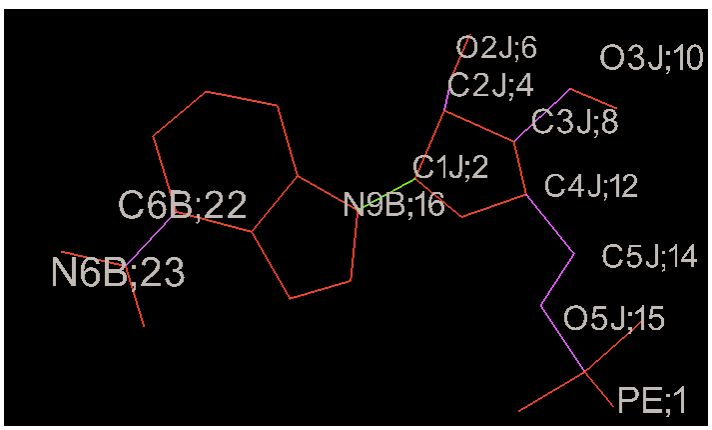


Figure 4.9. Torsion Tree representation of AMP with 1 flexible bond.

The order of presentation in this section will be docking of AP5, ATP, ADP and finally AMP, i.e. in decreasing complexity of the ligand structures. Since AP5 is the biggest molecule among the ligands, it will be hard to get accurate results with fully flexible ligand. That is why the fully flexible case is not applied for AP5.

4.3.1. AP5 Dockings

In the previous docking studies on ClustENM conformers (Kurkcuoglu, 2015), the inhibitor AP5 was docked to seven conformers representing diverse states of ADK from open to closed. The results indicated that AP5 is more favorably docked to intermediate states from generations 4 and 7 with partially closed LID and open NMP domains. This indicated an induced fit mechanism may be in play since the closed states were not accessible to the ligand. In this thesis, dockings to 92 ClustENM conformers are performed to compare the binding preferances of each ligand to ADK.

AP5_Blind docking results are presented in Figure 4.10 based on the sRMSD values of the conformers. In this figure and also in the following ones, the conformers are colored according to the ligand RMSD and docking scores. After each docking perfomed on a specific conformer, 100 protein-ligand poses are obtained, which are further clustered based on the ligand RMSDs with respect to the reference crystal structure (1AKE). If among the clusters, there is one with low ligand RMSD (less than 2.5 Å for AP5 and 2 Å for other ligands), that conformer is colored according to its best score (meaning lowest energy value); red, yellow and green represent scores in descending order. In this case, green ones represents a gap between -5 and -3.18 kcal/mol scoring values. The reported best cluster is not necessarily the first cluster for each conformer. The remaining conformers are shown in gray, which means that they do not have satisfactory ligand RMSD and/or score values.

Among the successful conformers that are colored, half closed-LID states are populated for AP5 docking results. This result is also in agreement with Kurkcuoglu (2015)'s thesis as stated "the conformer generation procedure leads to successful docking poses especially 'in-between' ADK state with a partially closed LID and open NMP

and average success for fully closed domains". Lower ligand RMSD values with best scores (lowest energies in red color) were obtained for gen4_8 and gen7_2, with the corresponding scores of -7.12 and -6.52 kcal/mol.

Tables 4.5 and 4.6 summarize the results of all dockings performed. For AP5, there are 8 conformers with ligand RMSD less than 2 Å, which are almost evenly distributed among generation 3 and above. But the scores of all these conformers with low ligand RMSD are not satisfactory, i.e. have relatively high energy.

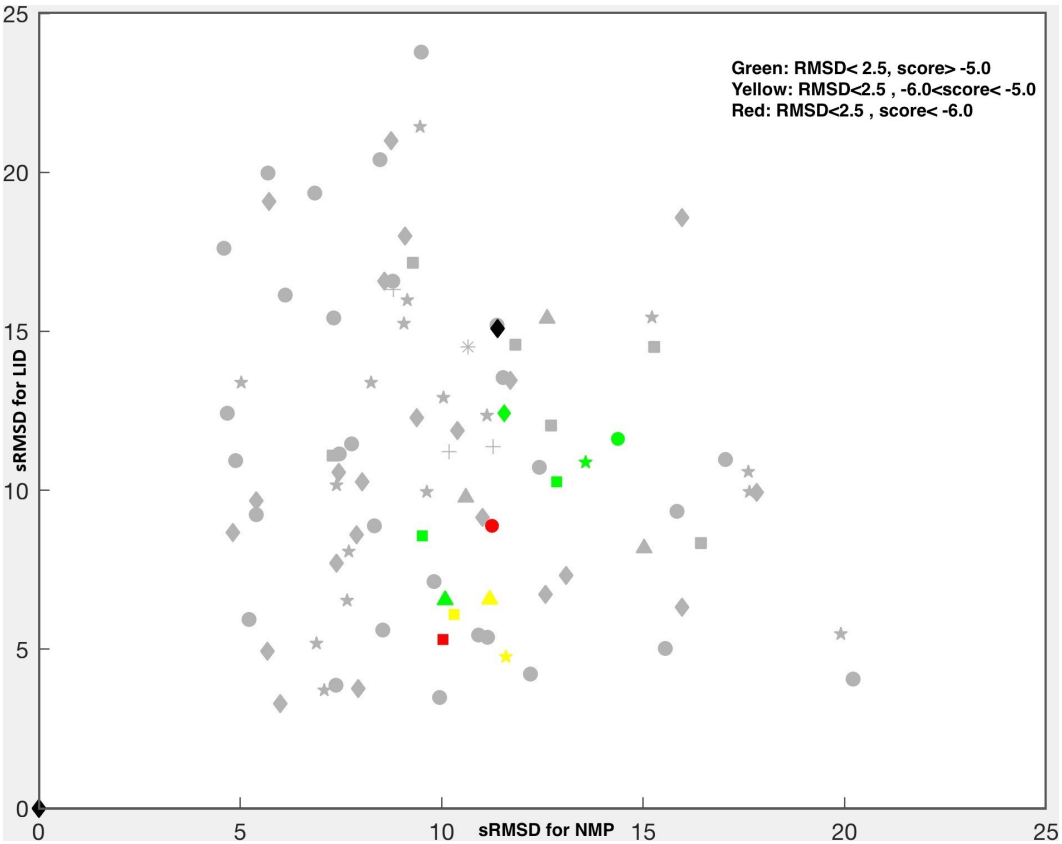


Figure 4.10. sRMSD representation of AP5 dockings.

4.3.2. ATP_Blind Dockings

In the investigation of blind ATP dockings, two different cases of flexibility are considered in order to observe the effect of flexible torsion angles.

With only one flexible bond in the ligand structure, successful conformers populate both intermediate and closed states of the LID domain, whereas NMP domain positioning seems to be not important for ATP binding, as seen in Figure 4.11. The reference ligand for docking is taken on the LID side of ATP blind dockings, and the ligand RMSDs are also reported with respect to this ATP molecule. Therefore it is reasonable to observe that the dockings are insensitive to NMP domain closure. There are 25 conformers with RMSD less than 2 Å with the minimum energy of -7.66 kcal/mol, for which generations 5 and 6 are the most populated.

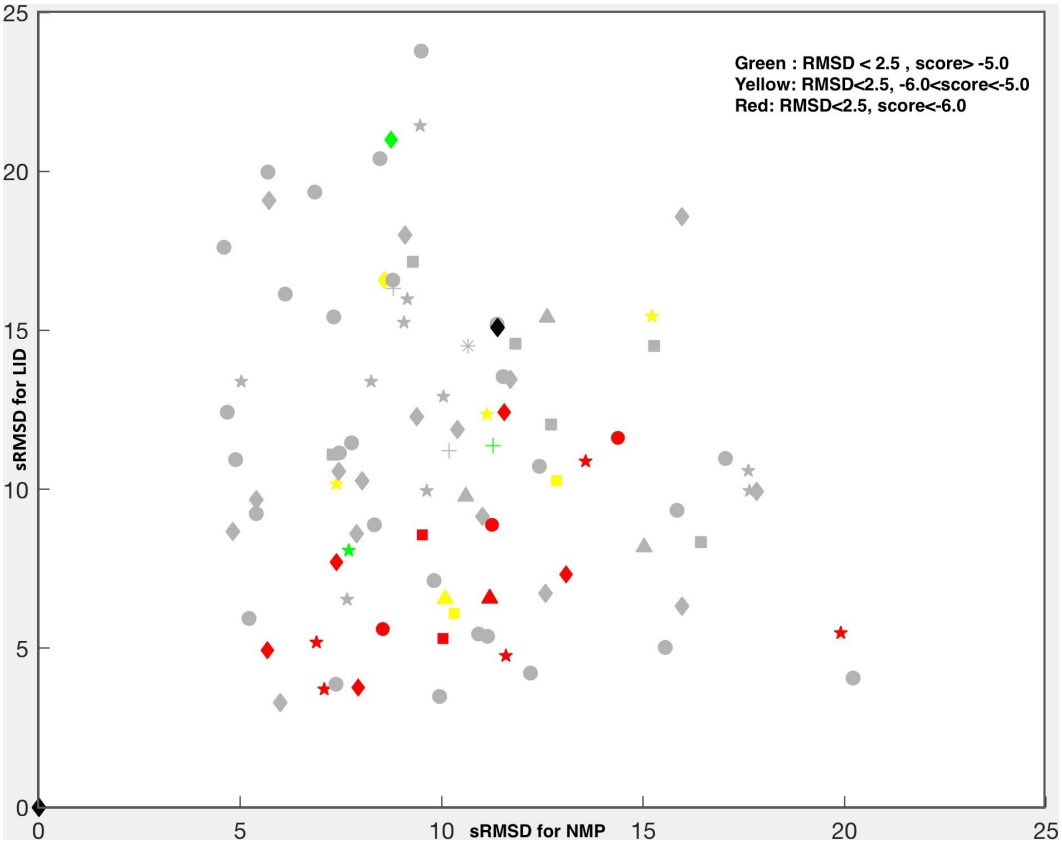


Figure 4.11. sRMSD representation of ATP_Blind dockings.

4.3.3. ATP_Blind_Flex Dockings

When ATP is fully flexible, the distribution in Figure 4.12 is more populated for relatively closed LID and NMP domains. Increasing the flexibility of the ligand leads to lower energies of binding, i.e. higher scores, in comparison to the case with a single

flexible bond. In summary, 14 conformers have satisfactorily low ligand RMSDs and quite high scores. It can be said that with this case, it is obtained more clear picture with only best scores together with minimum RMSD values.

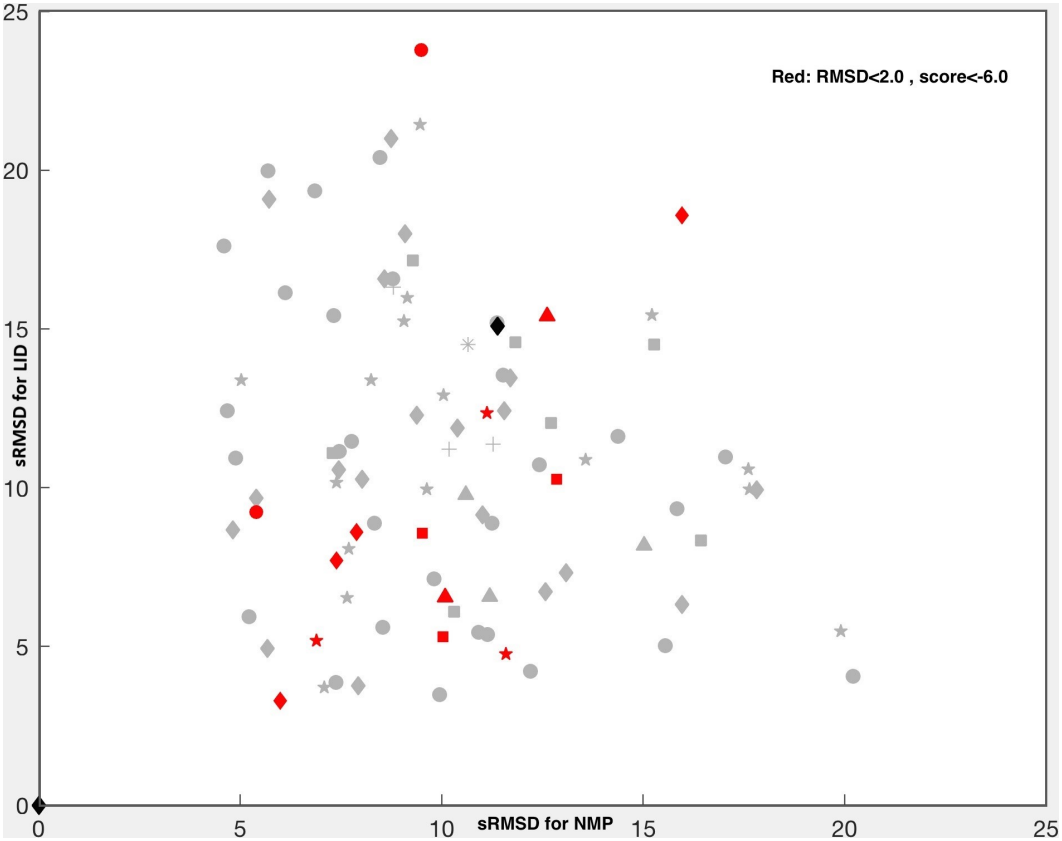


Figure 4.12. sRMSD representation of ATP_Blind_Flex dockings.

4.3.4. ATP_LID Dockings

Next the dockings are performed by confining the box to the LID domain side, which is the correct position for the ligand ATP. Many red poses (which means the conformers have smaller RMSD and lower energy values) are obtained in Figure 4.13. There are 27 conformers with low ligand RMSDs, most of which populate the intermediate to closed states of LID domain. The lowest RMSD is 0.63 Å with a score of -6.0 kcal/mol.

**By looking all 92 conformers' energy and RMSD values, there are 13 conformers which give scoring values lower than -6 kcal/mol, however ligand RMSD values bigger than 4 Å. Since these are not on the correct location but only give better energies, it will be called mislead population in the future sections of this work. It is obtained more accurate conformers as 17 ones (low energy together with low RMSDs).

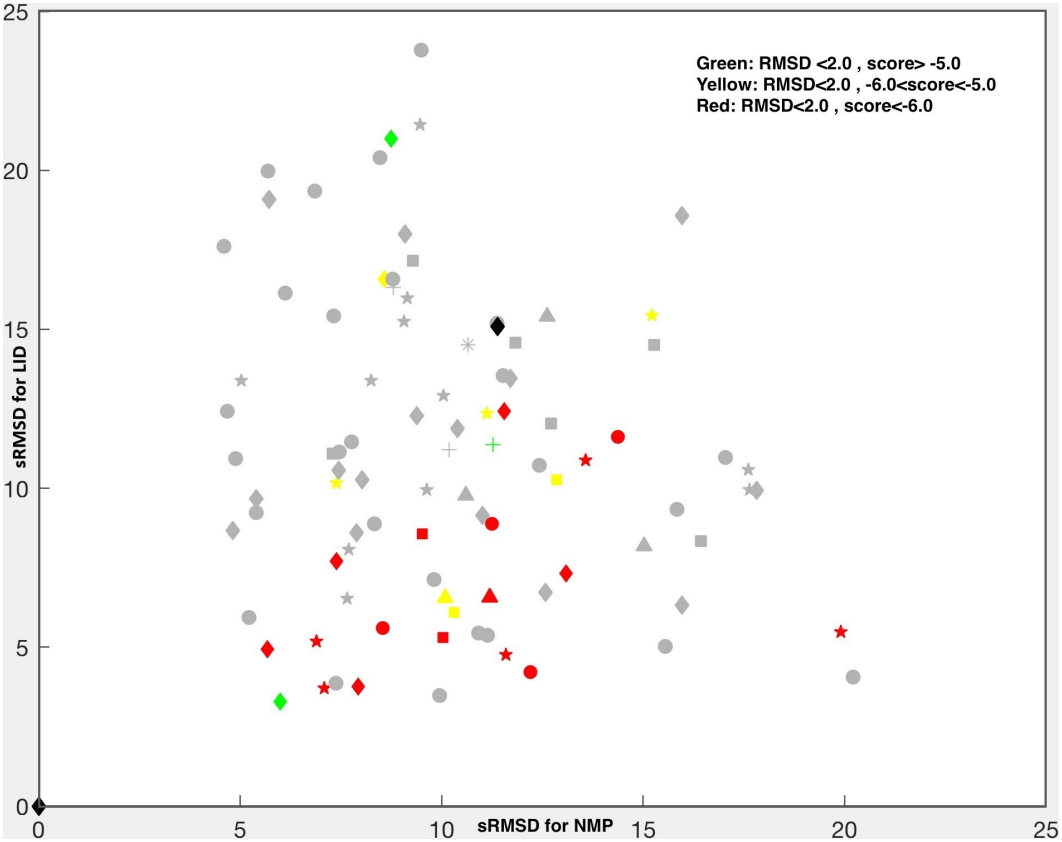


Figure 4.13. sRMSD representation of ATP_LID dockings.

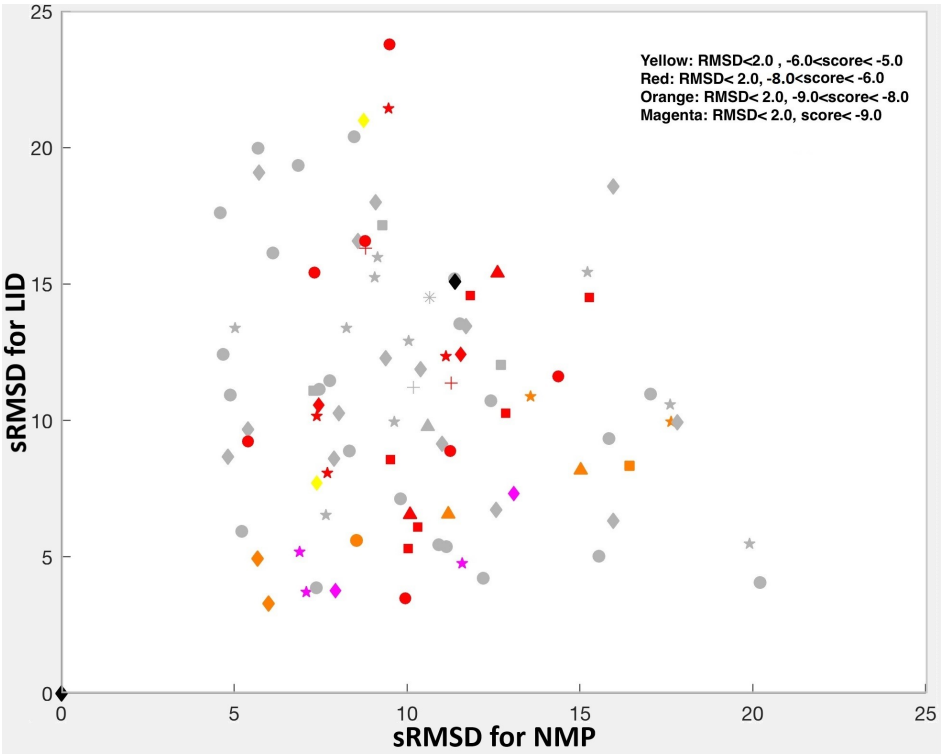
4.3.5. ATP_LID_Flex

To see the effect of fully flexible ligand ATP, the grid box is again set to the LID domain. For this case, the results are much better in terms of scores/energies, which present a minimum value of -9.71. There are 38 conformers with low ligand RMSDs that are populated in generations 4 to 7.

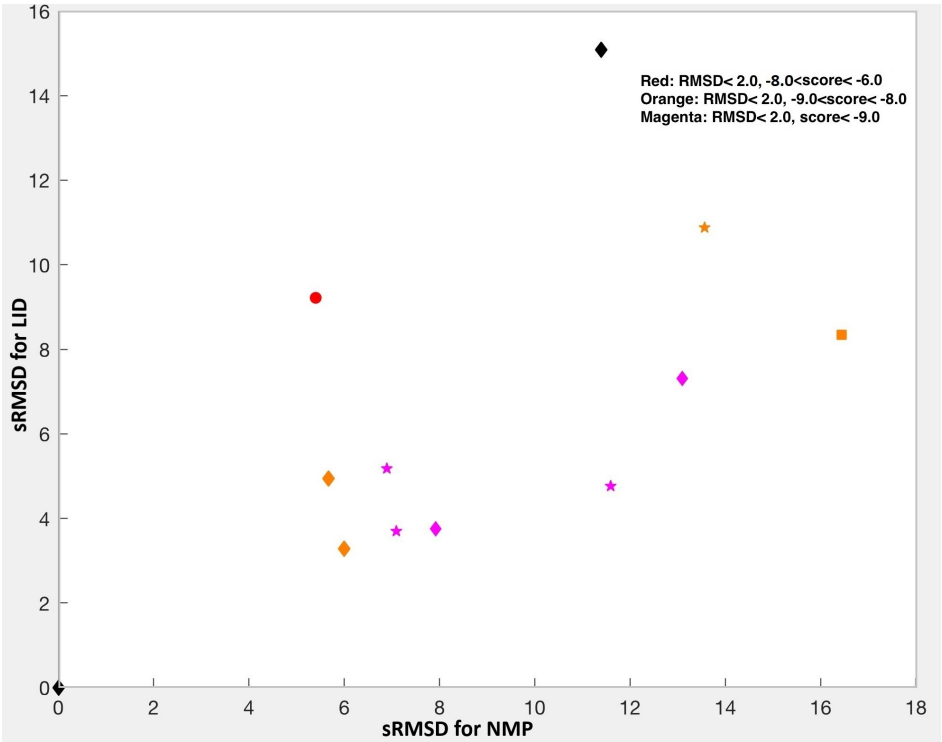
Referring to Figure 4.14(a), there are more diverse conformers with high scores, again the intermediate to closed LID ones being more populated. The fully flexible structure is more likely to have some poses, where two domains are relatively closed with low sRMSDs. Additional colors (orange and magenta) reflect the higher scores, i.e. energies lower than -8.0 kcal/mol on this figure. In summary, the fully flexible ligand more easy accommodates itself on quite diverse conformations of the receptor.

When some conformers are examined in detail which give red values as RMSD lower than 2 Å and scoring between -6 to -8 kcal/mol, it is observed that these conformers have also bigger RMSD's in different clusters and related poses. For instance, gen7_24 has 41 clusters in total and for first and second ones it gives RMSD values 12.92 and 6.09 Å,respectively. The red dote in Figure 4.14(a) corresponds to nineteenth cluster which gives 1.77 Å RMSD together with -6.20 kcal/mol energy level. Same situation is valid for gen5_8 which is also located in the same areas (with bigger LID sRMSD levels) in Figure 4.14(a). In Figure 4.14(b), it is indicated the conformers which they have energy level below -7 kcal/mol together with RMSD level under 2 Å,but only in first two clusters this time. It can be seen that the distribution nearly on the same areas with ATP_LID case,since we have only 1 or 4 clusters in one flexible bond cases which is consistent when we look only to first or second clusters in fully flexible case.

**By looking all 92 conformers' energy and RMSD values, there are 12 conformers which give scoring values lower than -9 kcal/mol, however ligand RMSD values bigger than 4 Å.



(a) ATP_LID_Flex all conformers.



(b) Conformers given best scores in first two clusters.

Figure 4.14. sRMSD representation of ATP_LID_Flex dockings.

4.3.6. ATP_NMP Dockings

To observe if ATP can dock to the NMP domain, the box is placed on NMP side. The lowest ligand RMSD is 2.87 Å with a score of -5.1 kcal/mol. Figure 4.15 displays several conformers provided that the RMSD threshold is kept quite high, indicating unsuccessful dockings. Thus, successful dockings of ATP are only possible to the LID domain.

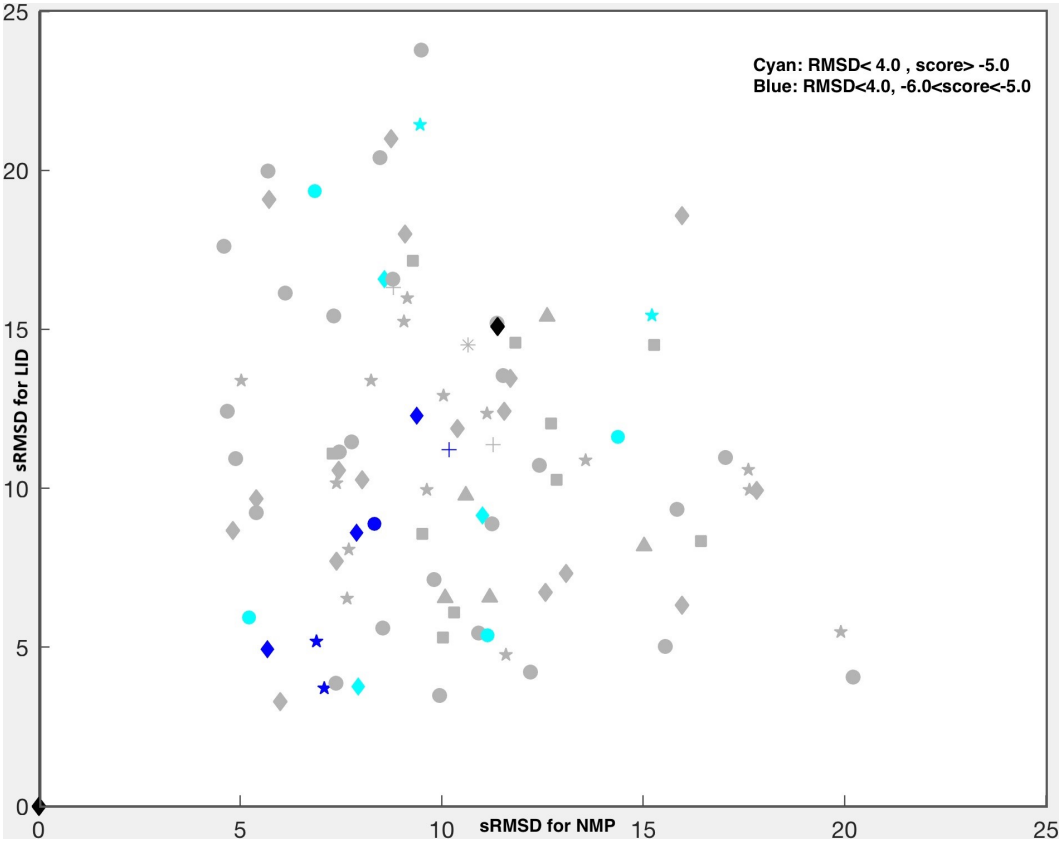


Figure 4.15. sRMSD representation of ATP_NMP dockings.

4.3.7. ADP_Blind Dockings

Next, the dockings of ADP will be considered. ADP_Blind case is presented in Figure 4.16. Ligand prefers to bind intermediate states where LID domain is partially closed. The results are similar to ATP_Blind, but the scores are comparatively lower, as ADP lacks one phosphate group. Another observation is that the higher sRMSD values for NMP are more populated, which was also observed in ATP_LID_Flex.

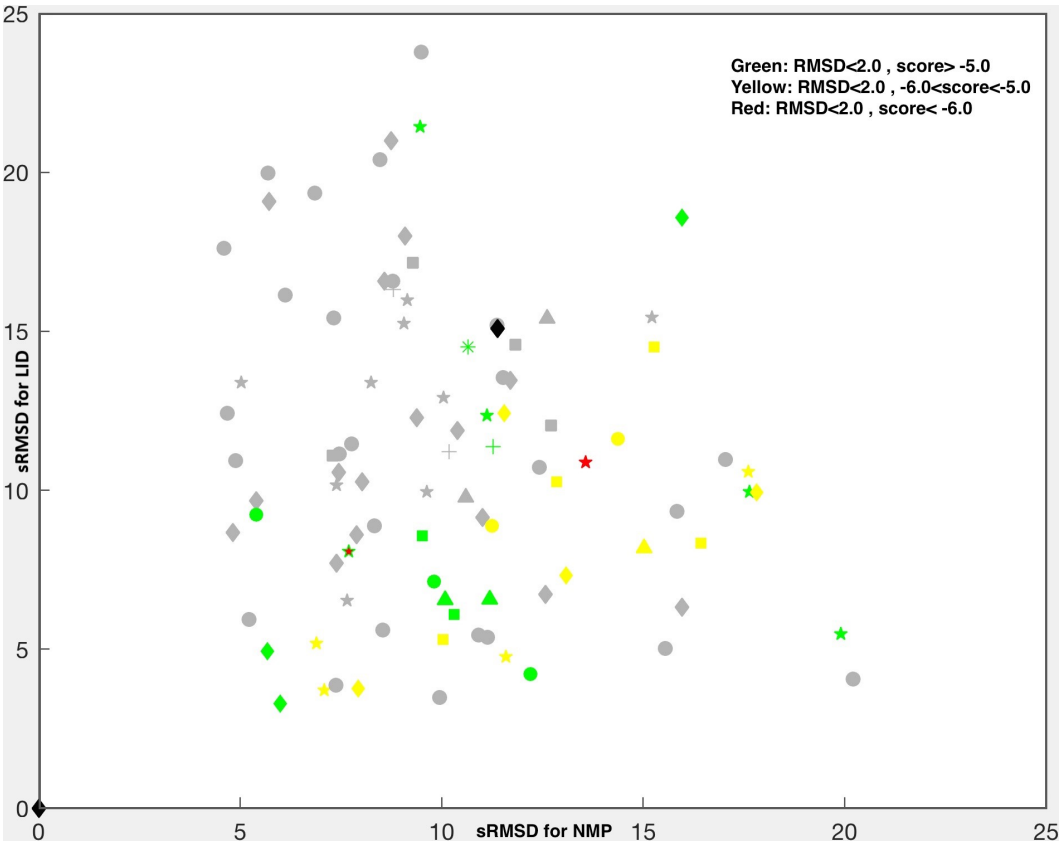


Figure 4.16. sRMSD representation of ADP_Blind dockings.

The one and only best pose in this case is shown in below Figure 4.17, which arisen from generation 5.

With the surface representation, it is clear that gen5_11 has an open NMP and closed LID structure together with a minimum energy necessity. Figure 4.17 helps us to prove this direction by analyzing its .dlg file in Autodock program. This is a clear

picture of the preference of ligand ADP to bind the LID domain and gives a partially open NMP structure.

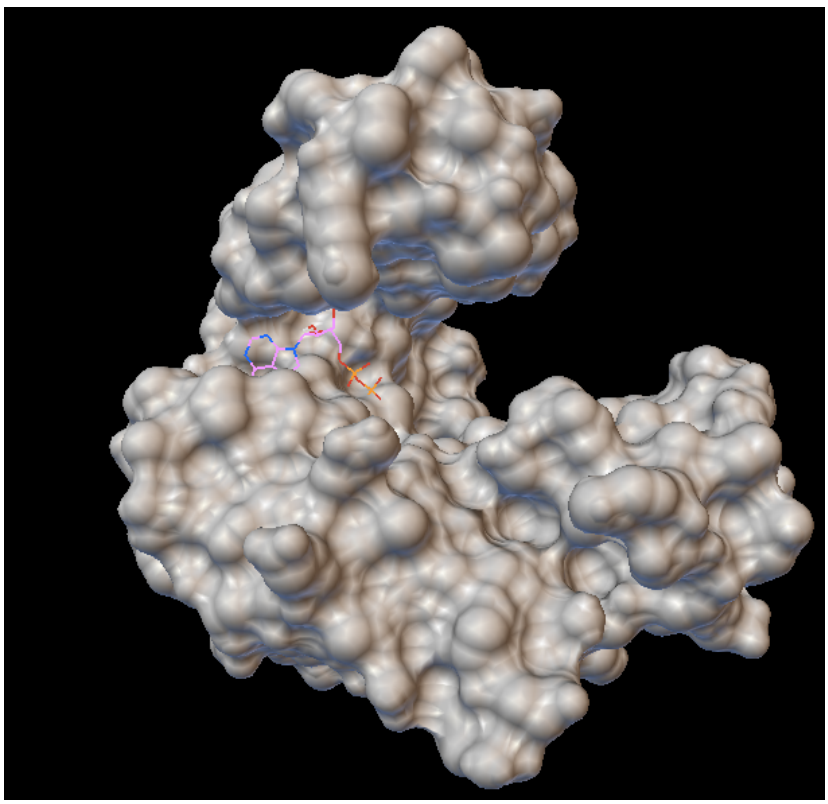
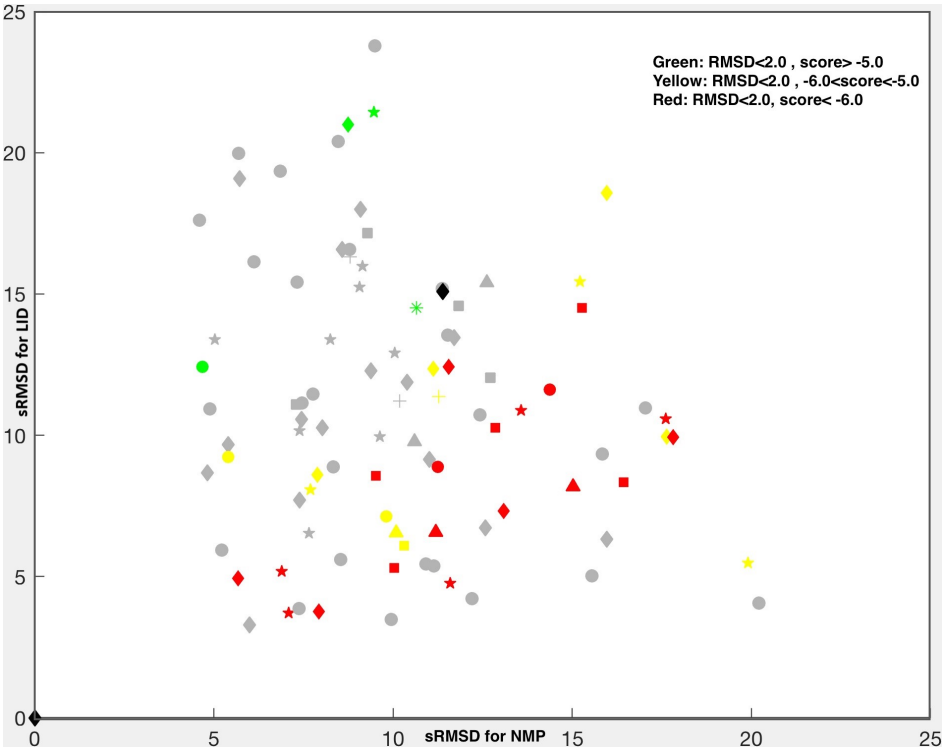


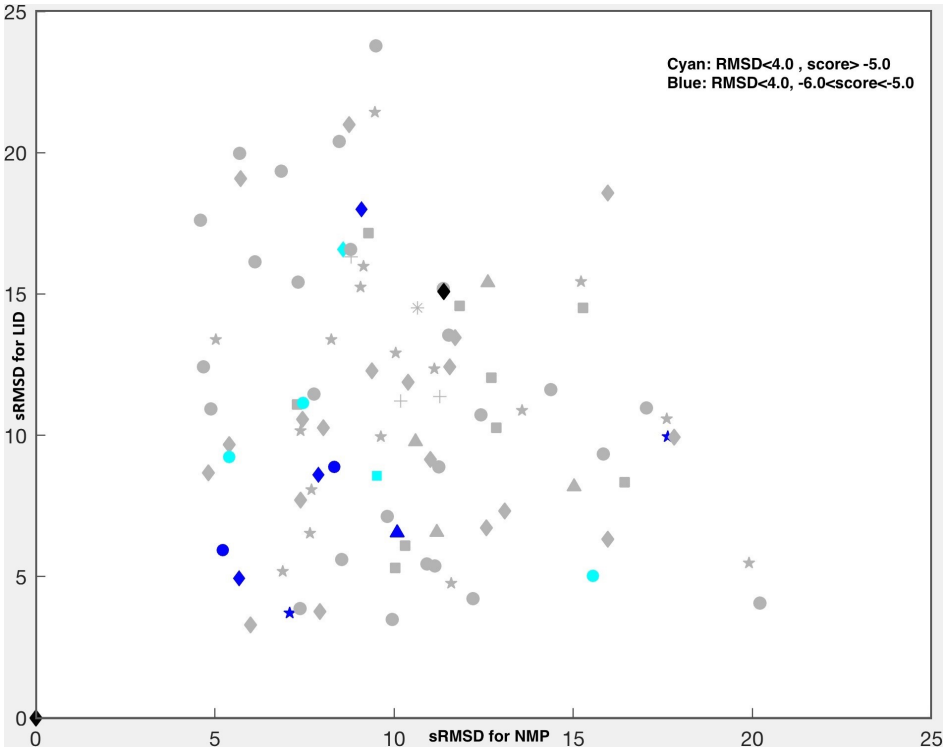
Figure 4.17. Molecular surface representation of gen5_11.

4.3.8. ADP_LID and ADP_NMP Dockings

In this section, grid box dimensions are set to more specific regions, namely the LID or the NMP side, which are presented in Figures 4.18(a) and 4.18(b), respectively. The scores of the best (low RMSD) poses increase especially for the LID case. As a result, ADP also prefers to locate towards LID domain, exhibiting similar behavior as ATP.



(a) ADP_LID.



(b) ADP_NMP.

Figure 4.18. sRMSD representations of ADP domain dockings.

**By looking all 92 conformers' energy and RMSD values, there are 3 mislead conformers for ADP_LID and 8 for ADP_NMP cases, which gives lower energy but high RMSD values. For ADP_LID case, 19 conformers give consistent results in terms of both low energy and low RMSD values.

When ADP ligand is specifically aimed to NMP domain, no best poses (red dots) can be obtained. Some conformers could be shown after revising the RMSD threshold to 4.00, not like in LID case (2.00). The lowest RMSD is 2.81 for the NMP side.

4.3.9. AMP_LID and AMP_NMP Dockings

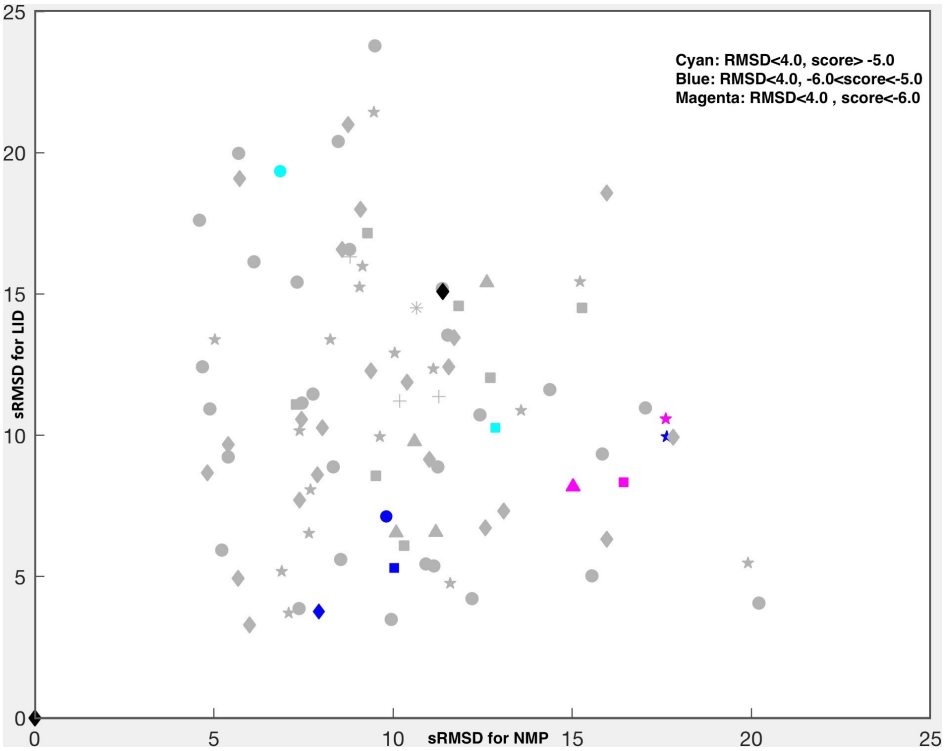
The last ligand is AMP, which is docked to LID and NMP domains independently. Both cases do not yield successful results. Only in the case LID domain, there is a conformer gen6_22 with an RMSD lower than 2.0. If the RMSD threshold is increased to 4.0, some conformers could be shown in Figure 4.19(a). It can be seen from Figure 4.19, AMP has a slight tendency to dock to the LID domain.

4.3.10. AMP_NMP_Flex and AMP_NMP_wATP_Flex Dockings

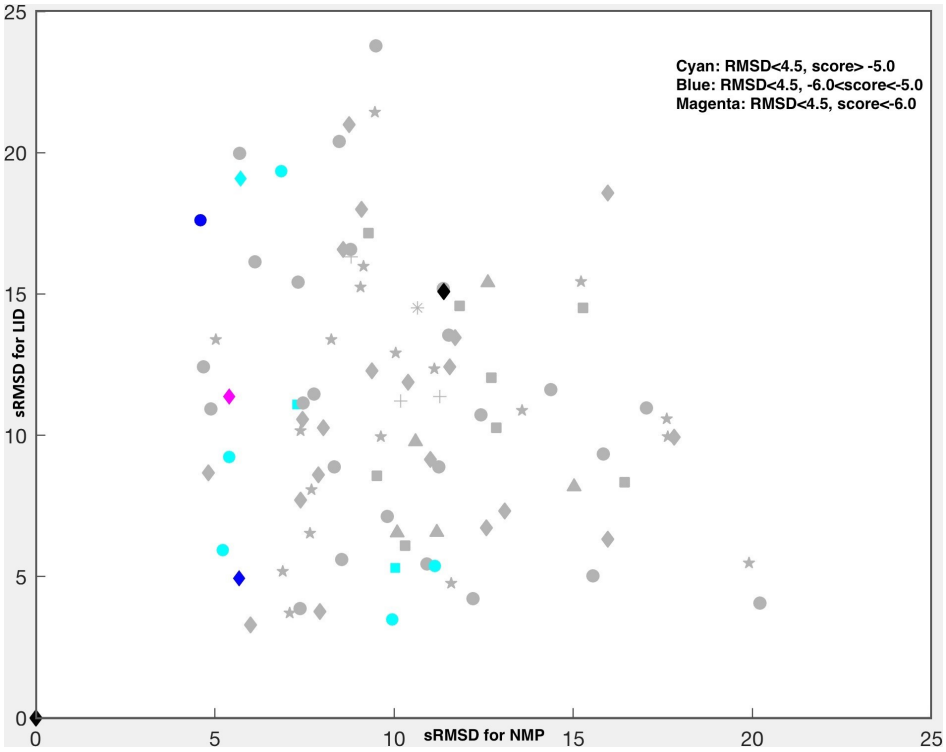
Modeling AMP as fully flexible did not improve the results much in terms of NMP domain binding. Generation 6 also gave the one and only best pose which is shown in Figure 4.20, with an RMSD of 3.0 Å and -5.7 kcal/mol binding energy value. As a result, we do not observe satisfactory docking of AMP to either of the domains.

**In order to understand if the presence of ATP facilitates the docking of AMP to NMP domain, AMP is docked to the NMP side with an ATP molecule already bound to the LID domain.

By using Pymol and Autodock, we created a complex structure with a bound ATP to a specific conformer. The aim of this case is to see if AMP still prefer or have these good results for binding LID domain with the presence of ATP on LID side in the first place.



(a) AMP_LID (RMSD limit 4.00).



(b) AMP_NMP (RMSD limit 4.00).

Figure 4.19. sRMSD representations of AMP domain dockings.

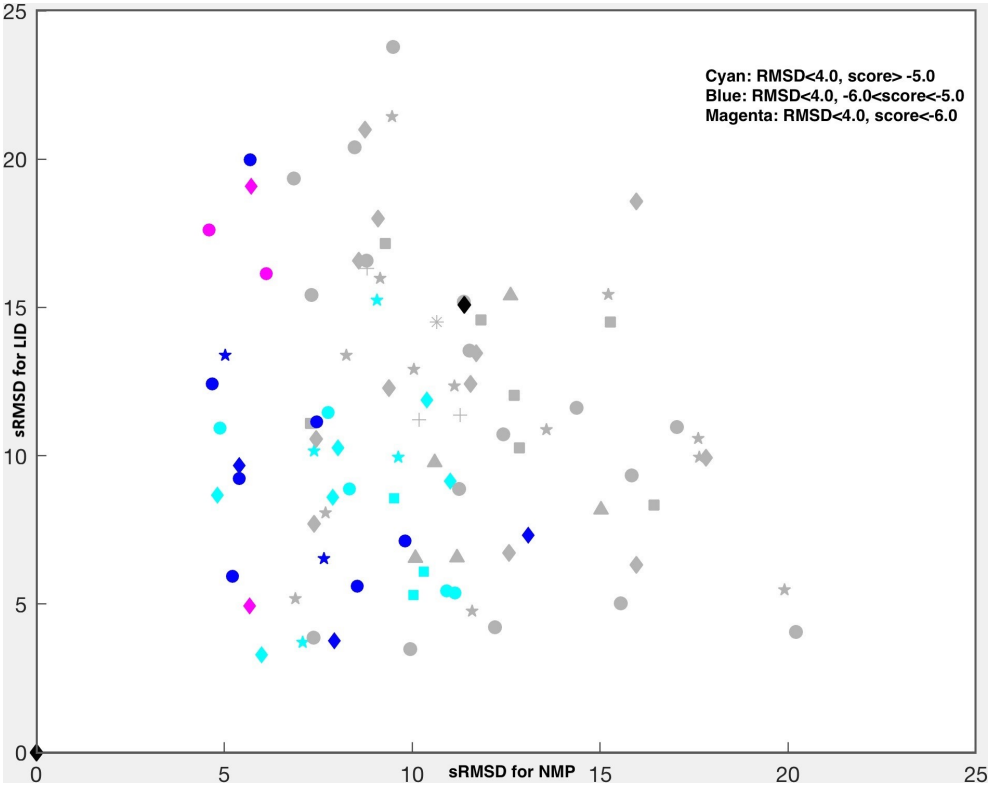


Figure 4.20. sRMSD representation of AMP_NMP_Flex dockings.

We applied this case only 31 conformers since their RMSD values are lower than 4.00 Å (ease the complexity to eliminate unlikely conformers).

All conditions are the same with other AMP cases;

- Ligand: AMP (7 active bonds), inside the LID domain there is ATP molecule (with 11 flexible bonds)- it is preferred to proceed with fully flexible ligands to get maximum movement ability

With the presence of other ligands (ATP in this case) AMP does not prefer to bind as much as possible in the absence of them. It may not be said that there is an order with the binding process however it can be related with the charge balance. With the presence of ATP and AMP together, there is a majority of negative charges, so it may cause a driving force to keep AMP away from the LID domain.

As we investigated in the beginning of this study, AMP seems to prefer to bind NMP domain in crystal structure. But when it comes to the intermediate states, it was observed that conformers most likely preferred to bind LID side just like ATP. As we performed the runs case by case by comparing specifically the binding effects of ATP and AMP separately, it may be related that ATP has more phosphate groups so more tend to interact with different residues. With the presence of both ATP and AMP ligands, it is a reliable result to get the domination of ATP binding.

Table 4.5. Summary table of Autodock results.

Docking	**min.RMSD Conformers**		**RMSD<2Å Conformers**	
	Ligand RMSD (Å)	**Score**	**No of Conformers**	**Min/Max Score (kcal/mol)**
AP5_Blind	1.69	-4.59	8	-6.52/-3.18
ATP_Blind	0.67	-5.95	25	-7.66/-4.90
ATP_Blind_Flex	1.40	-8.11	14	-8.33/-6.30
ATP_LID	0.63	-5.99	27	-7.75/-4.55
ATP_LID_Flex	1.31	-6.62	38	-9.71/-5.64
ATP_NMP	2.87	-5.08	0	-
ADP_Blind	0.70	-5.87	33	-6.14/-3.30
ADP_LID	0.87	-6.19	35	-7.22/-4.57
ADP_NMP	2.81	-5.03	0	-
AMP_LID	1.88	-5.44	1	-5.44
AMP_NMP	3.54	-4.72	0	-
AMP_NMP_Flex	2.94	-5.65	0	-
AMP_NMP_wATP_Flex	4.88	-5.72	0	-

Table 4.6. Generation distribution of dockings.

Docking	No of Conformers with RMSD < 2 Å						
	Gen1	**Gen2**	**Gen3**	**Gen4**	**Gen5**	**Gen6**	**Gen7**
AP5_Blind			1	2	2	1	2
ATP_Blind		1	2	4	8	7	3
ATP_Blind_Flex			2	3	3	4	2
ATP_LID		1	2	4	8	8	4
ATP_LID_Flex		2	4	7	9	8	8
ADP_Blind	1	1	3	6	10	7	5
ADP_LID	1	1	3	6	11	8	5
AMP_LID						1	

4.4. Protein-Ligand Interactions

Among the dockings discussed in the previous section, docking of ATP and ADP to the LID domain produced satisfactory results in terms of ligand RMSDs and scores. Almost one third of ClustENM conformers with intermediate and relative closed positions of the LID domain produced successful binding poses. The best procedure was the ATP_LID_Flex, which chosen to analyze the key protein-ligand interaction.

For this aim, the LigPlot program is used for producing a list of protein-ligand interactions in terms of H-bonds (Laskowski and Swindells, 2011; Wallace *et al.*, 1995). The interactions of diverse conformers/poses, shown in Figures 4.22, 4.23 and 4.24 are compared to the crystal structure, 1AKE. Table 4.7 lists the interactions of the selected poses. The amino acids those are italic in the table are the ones also observed in the crystal structure as H-bonds. Mostly, conformers with low ligand RMSD are chosen, but some unsuccessful ones are also provided for comparison, namely gen1_1, gen4_2 and gen7_19.

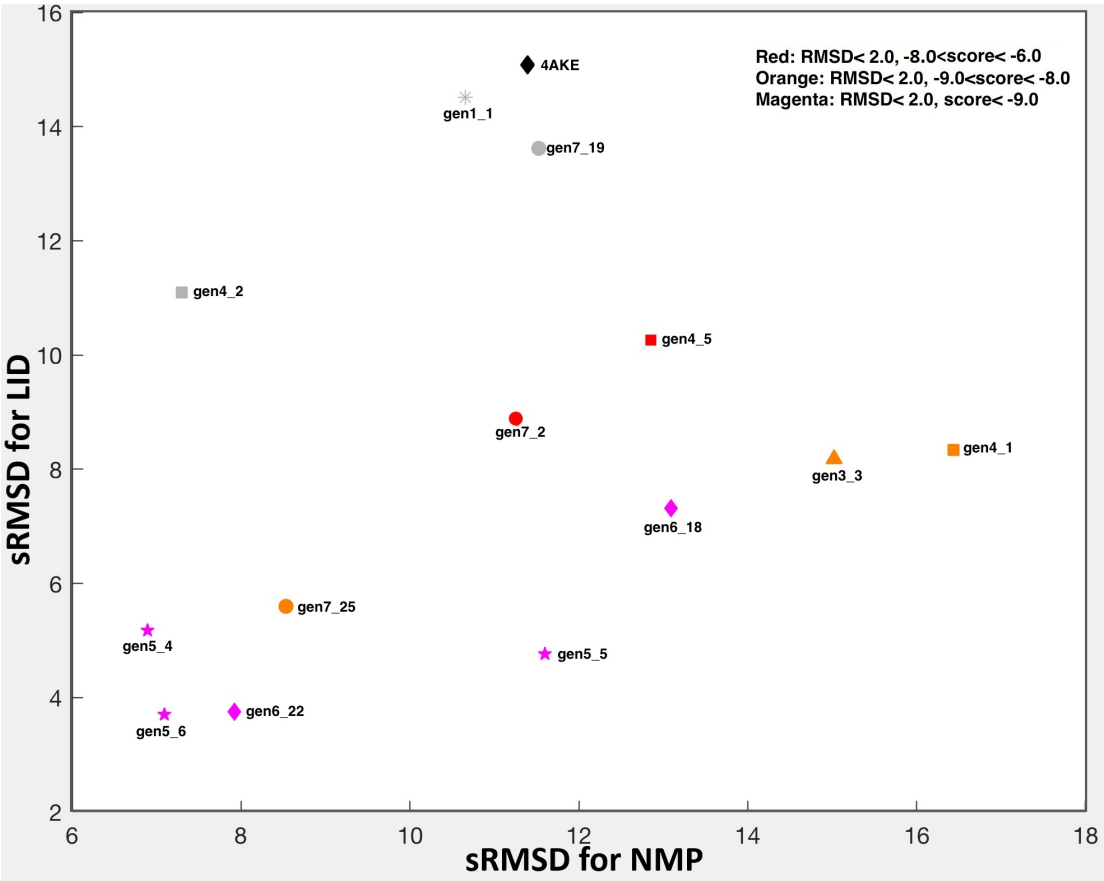


Figure 4.21. Conformers chosen for protein-ligand interaction analysis.

All chosen structures from generation 5 and 6 are the best conformers with the minimum energy and rmsd values and gives the maximum common hydrogen and non bonded amino acid numbers with 1AKE.

Table 4.7. Protein-ligand interaction lists using Ligplot.

Conformers (Ligand RMSD)	**Hydrogen Bond Interactions**
1AKE (LID Interactions)	G10, A11, G12, K13, G14, T15, R123, K200
gen1_1 (11.57 Å)	G7, *G10*, *K13*, *G14*, E170
gen3_3 (2.17 Å)	*G10*, *G12*, *K13*, *R123*, N138, *K200*
gen4_1 (2.46 Å)	*G10*, *G12*, *K13*, *G14*, *R123*, N138
gen4_2 (8.68 Å)	*G10*, *G12*, *K13*, *R36*, D84, *R123*, *R167*
gen4_5 (1.60 Å)	*G10*, *G12*, *K13*, Q28, D84, *R123*, *R167*
gen5_4 (1.95 Å)	*G10*, *G12*, *K13*, *G14*, *R123*, Y133, N138, *K200*
gen5_5 (1.84 Å)	*G10*, *A11*, *G12*, *K13*, *G14*, *R123*, N138, G198, *K200*
gen5_6 (1.73 Å)	*G10*, *A11*, *G14*, N138, *K200*
gen6_18 (1.65 Å)	*G10*, *G12*, *K13*, *G14*, N138, *K200*
gen6_22 (1.81 Å)	*G10*, *A11*, *G12*, *G14*, *T15*, *R123*, N138, *K200*
gen7_2 (3.93 Å)	*G10*, *K13*, *G14*, Y133, N138
gen7_19 (12.17 Å)	R124, K136, D159
gen7_25 (1.68 Å)	*A11*, *K13*, *R123*, A127, T155, *R156*

Structures gen3_3, gen4_1, gen4_5, gen7_2 and gen7_25 also have lower values of required energy and rmsd. It can be seen from Table 4.7, these conformers also have mostly common interactions with the crystal structure. Gen7_19 is the worst case with LID_Core angle 13.50°and NMP_Core angle 11.50°, and it is consistent with Ligplot study that we do not have any common amino acid in terms of hydrogen bond interactions.

Structures gen1_1 and gen4_2 are higher rmsd valued conformers with 5 to 10 Å(energies are still good), however we still have many mutual amino acids. To have a close look to one of these structures gen1_1 with Figure 4.22; it can be seen that there are some touching areas with 1AKE ligand (red), but also a gap between benzene rings.

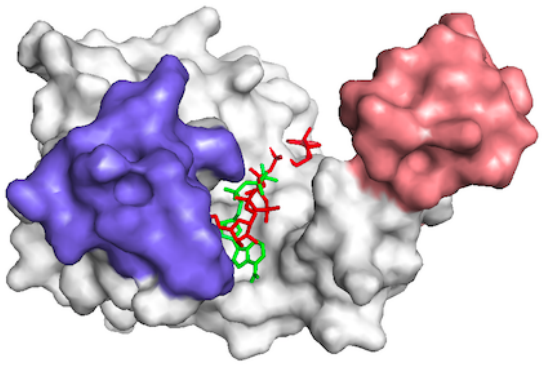


Figure 4.22. Surface representation of gen1_1 with ligand AP5(red).

Just to compare with the worst case gen7_19, there is no contact between the ligands as shown in Figure 4.23 below.

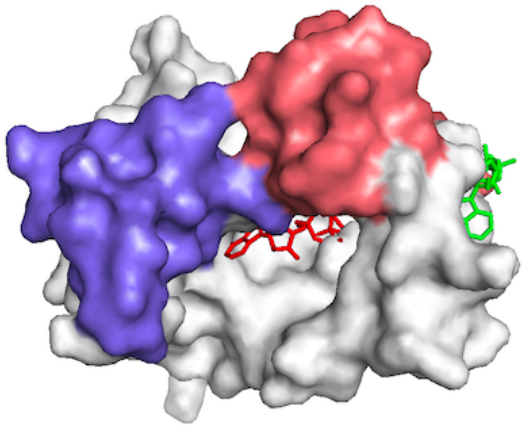


Figure 4.23. Surface representation of gen7_19 with ligand AP5(red).

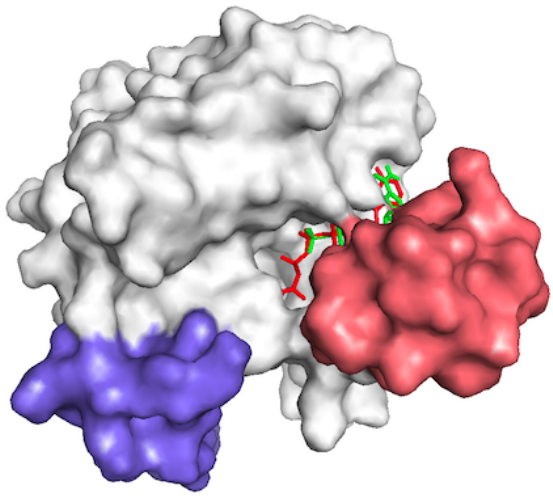


Figure 4.24. Surface representation of gen5_6 with ligand AP5(red).

Most common interactions almost all conformers include are GLY10, GLY12, GLY14, LYS200 and ARG123. These interactions are the ones who guide the ligand and all hydrogen bondings are interacting with phosphate groups. Figure 4.25 is showing all ligand interactions in the crystal structure, 1AKE.

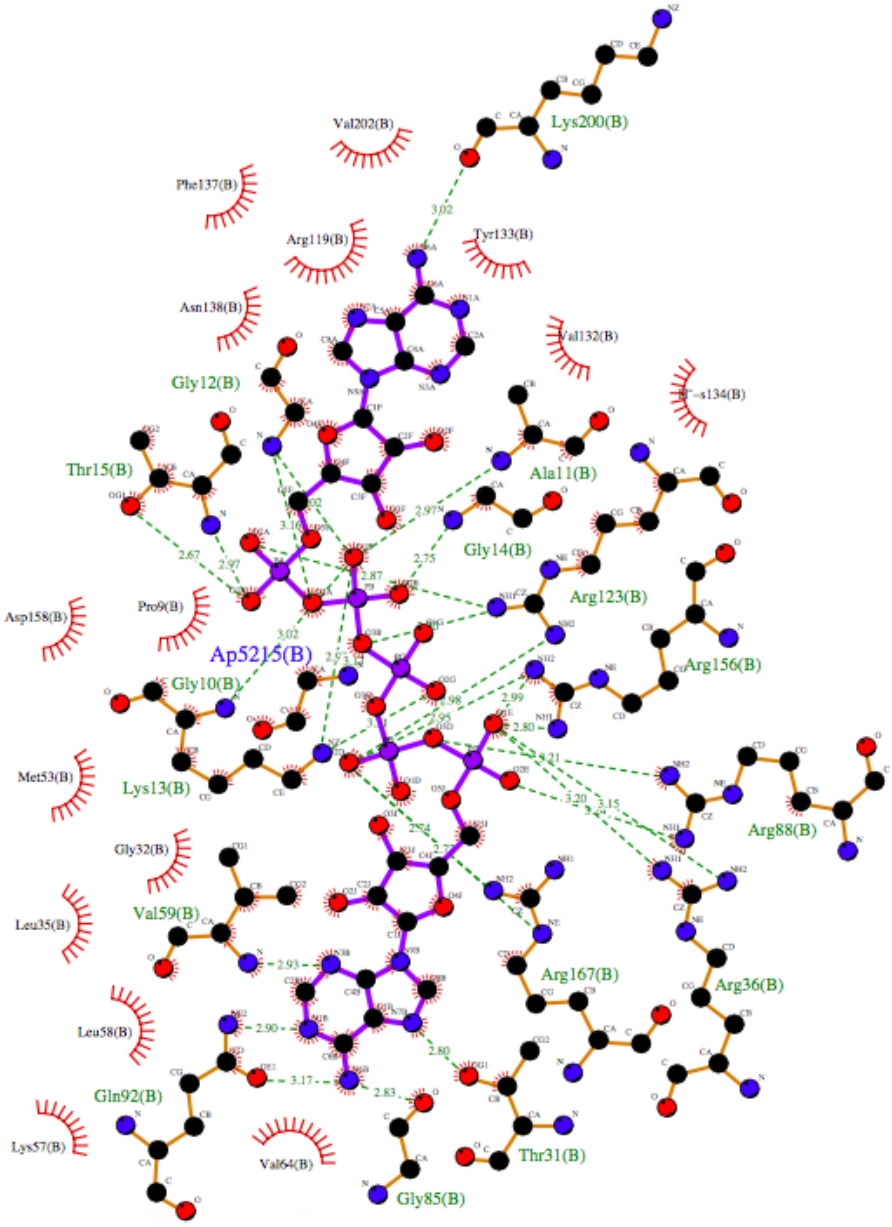
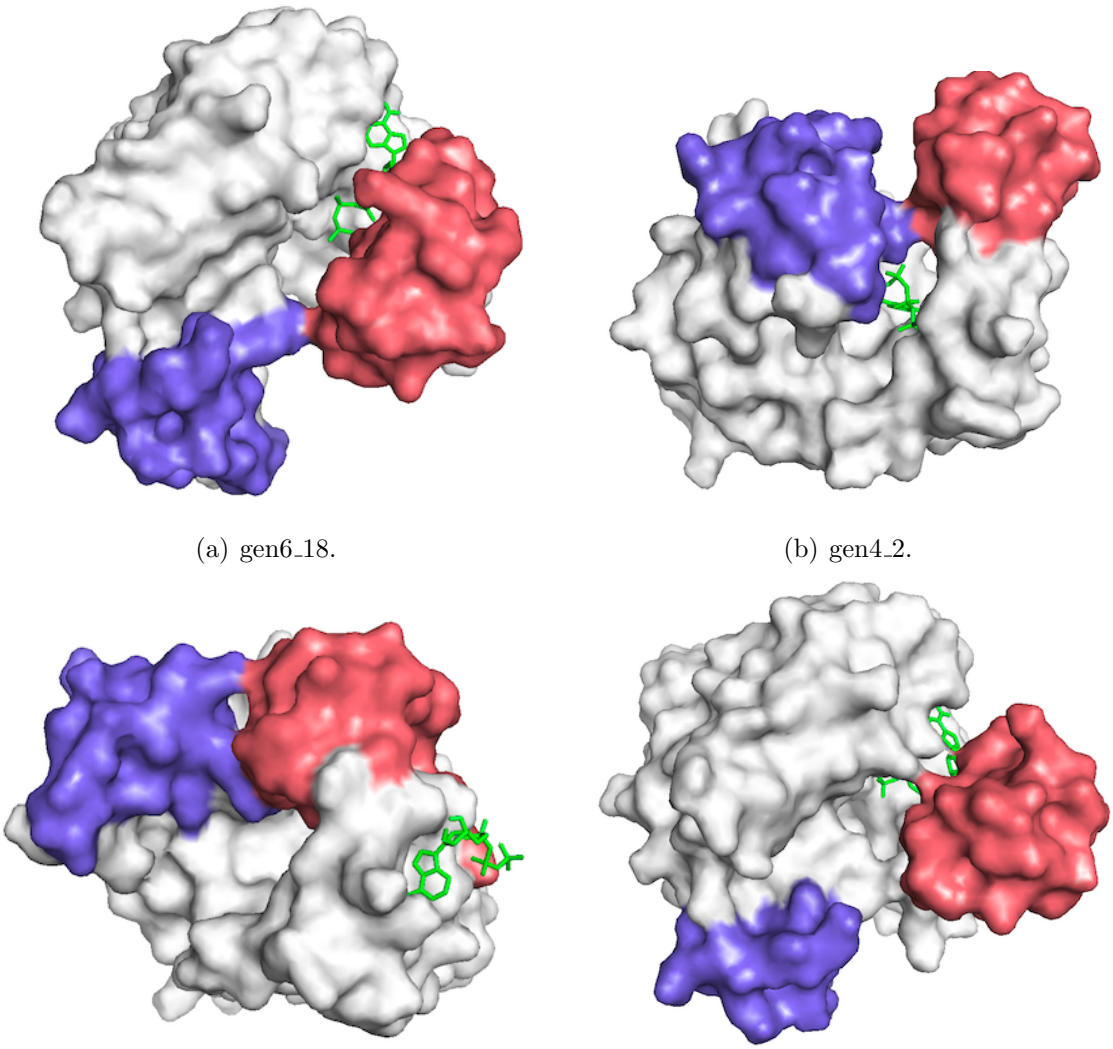


Figure 4.25. Bonding interactions of 1AKE.

4.5. Protein Dynamics Using Elastic Network Models

In this section, the effect of bound ligands on the vibrational dynamics of the protein is analyzed using the Dynomics server(Li *et al.*, 2017). The server applies both Gaussian network model (GNM) and anisotropic network model (ANM)(Atilgan *et al.*, 2001; Bahar *et al.*, 1997; Haliloglu *et al.*, 1997) by forming a network of harmonic springs between close-neighboring residue pairs, represented by alpha-carbon atoms. First the complex structure/pose will be used to obtain the dynamics without the ligand. In order to observe the changes in protein dynamics due to binding, the ligand is introduced using the reduced model approach (Bahar *et al.*, 2009; Ming and Wall, 2005). To have a similar level of coarse-graining with amino acid residues, ATP is represented by four nodes placed at atoms: PA, PG, O2F and N6A. Then, the results of the two calculations will be compared in terms of changes in mode collectivities, and mean squared fluctuations of residues.

In Figures 4.26(a), 4.26(b), 4.26(c) and 4.26(d), four structures are shown from the case ATP_LID_Flex, for which degree of collectivity and mean square fluctuations results will be graphically presented in order to see the effect of ligand binding to different conformers; namely 1AKE (closed crystal structure), gen6_18 (half-closed LID and open NMP domains), gen4_2 (open LID and closed NMP), gen7_19 (open LID and NMP) and gen5_6 (closed LID and NMP).



(a) gen6_18. (b) gen4_2.

(c) gen7_19. (d) gen5_6.

Figure 4.26. Surface representations of chosen structures on Dynomics study.

4.5.1. Degree of Collectivity (GNM)

Degree of collectivity figures are shown for the cases of with and without ligand. In general a significant change does not take place in any of the conformers. Collectivity is a measure of correlated motions of large entities such as domains. Only slight changes are observed, either increase or decrease, in the global mode collectivities. Only for the case of the conformer gen7_19 (LID and NMP open), in which ATP binds to the CORE domain outside the cavity, no change is detected.

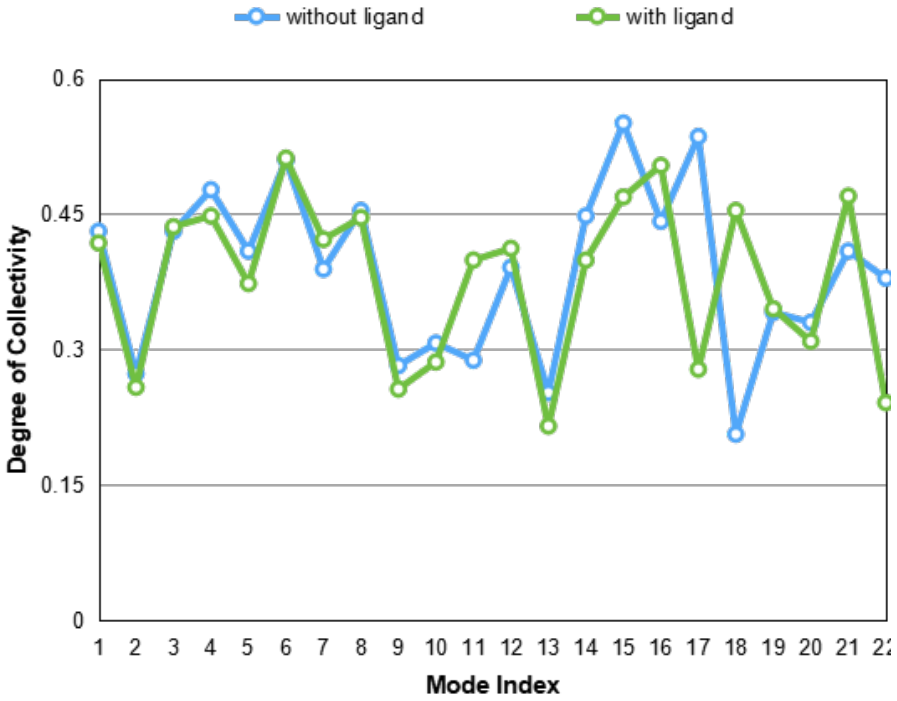


Figure 4.27. Degree of collectivity of 1AKE.

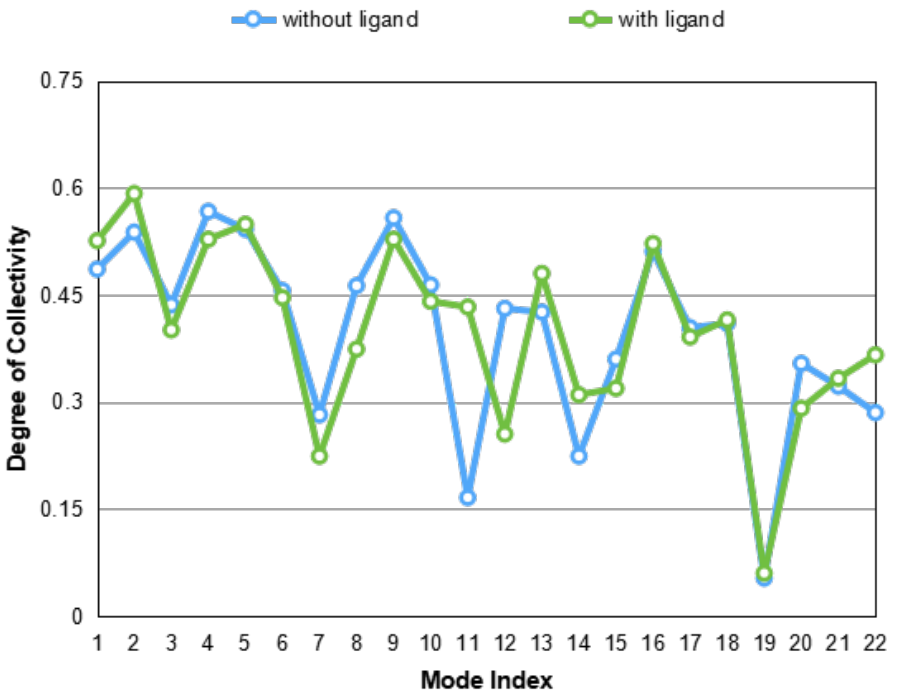


Figure 4.28. Degree of collectivity of gen6_18.

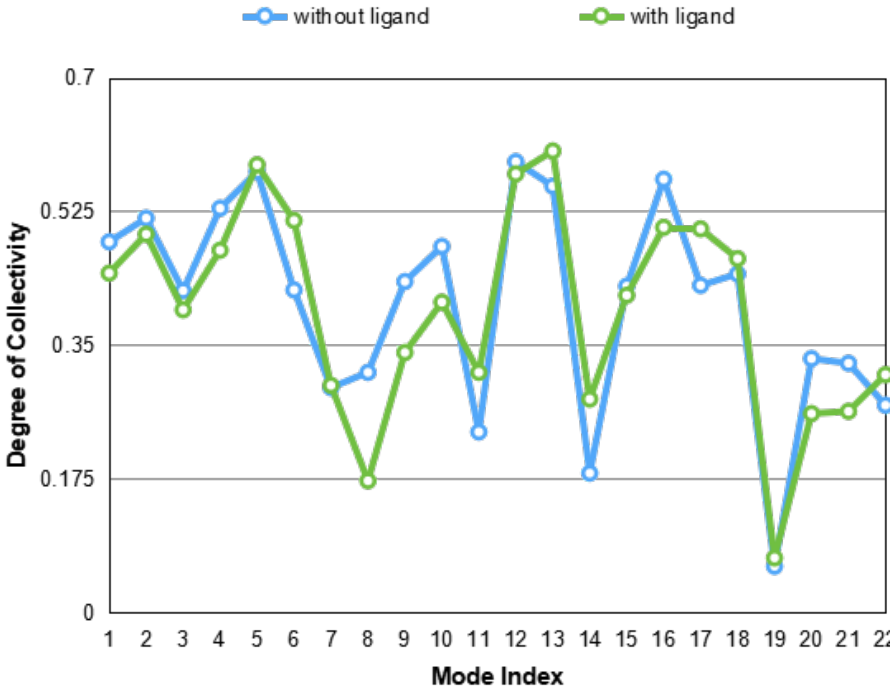


Figure 4.29. Degree of collectivity of gen4_2.

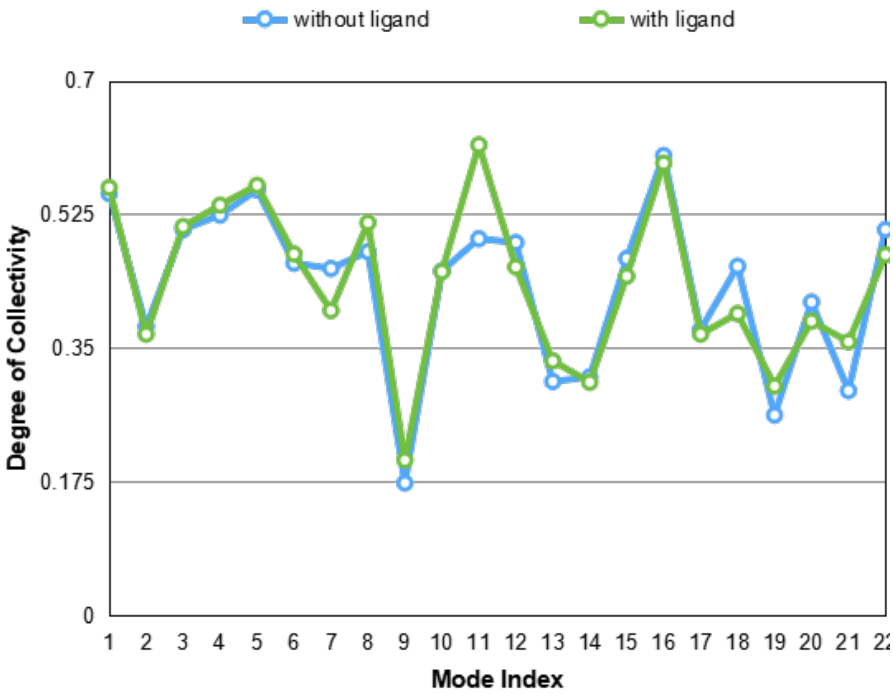


Figure 4.30. Degree of collectivity of gen7_19.

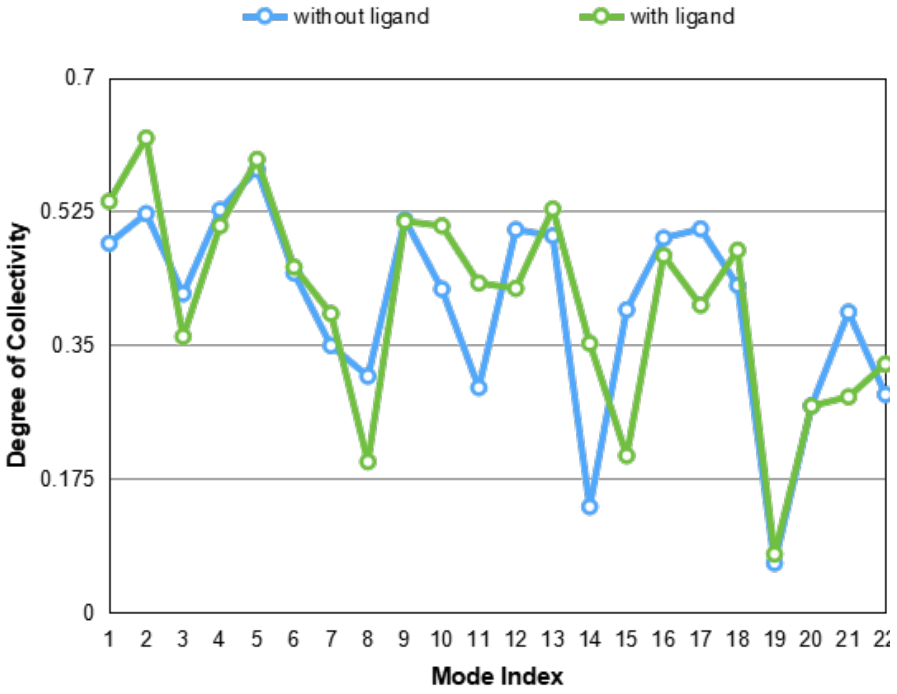
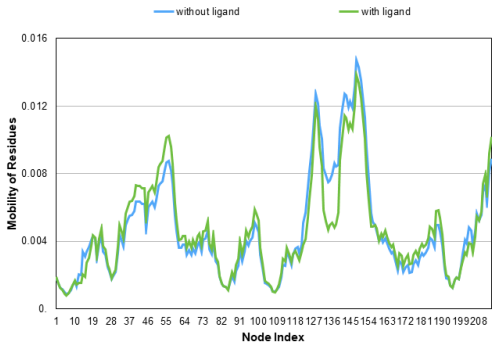


Figure 4.31. Degree of collectivity of gen5_6.

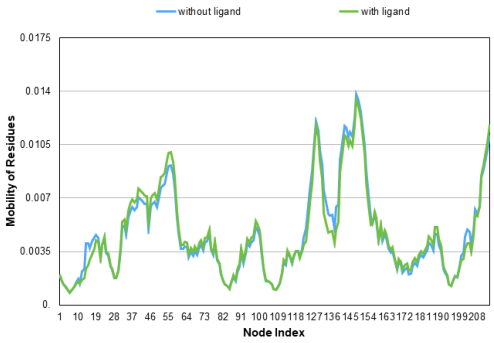
4.5.2. Cumulative contribution of the slowest 10 modes to the mobility of residues (GNM)

Mobility of the residues is also investigated to see the effect of ligand binding on the mean squared fluctuation of residues. For the crystal structure 1AKE, gen6_18 (LID domain half-closed) and gen5_6 (both domains closed), residue mobility is decreasing between 130-140 with the presence of ligand, since ligand binding is happening on these residues (LID domain residues correspond 125-153). The mobilities are nearly the same with and without ligand for the cases with both open LID and NMP and the case where only NMP closed.

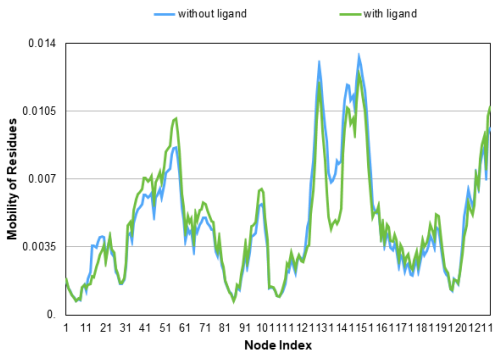
For gen5_6 (LID and NMP closed), the only unique area is obtained between residues 41-60 where the mobility is higher with the presence of ligand, and these residues correspond to the area where NMP prefers to bind (NMP domain residues correspond 35-55). It is relevant to have such different result since we chose a structure that ligand want to bind to LID domain in the first place and it is not a natural tendency towards NMP domain.



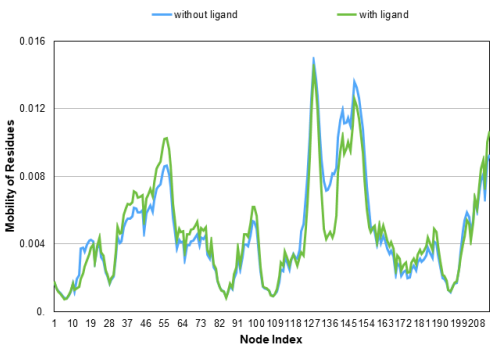
(a) gen3_3.



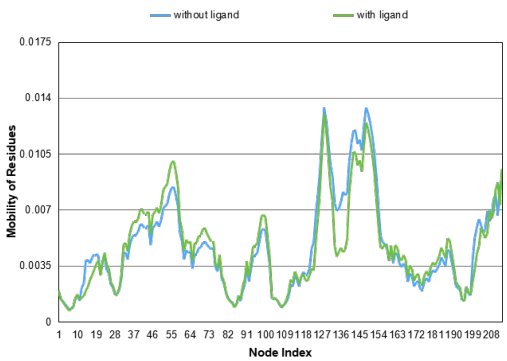
(b) gen4_1.



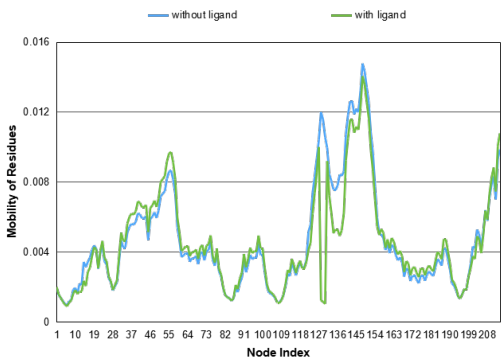
(c) gen5_4.



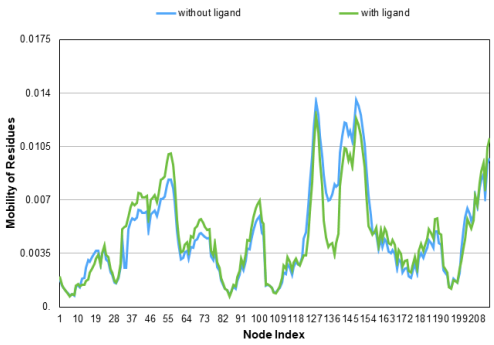
(d) gen5_5.



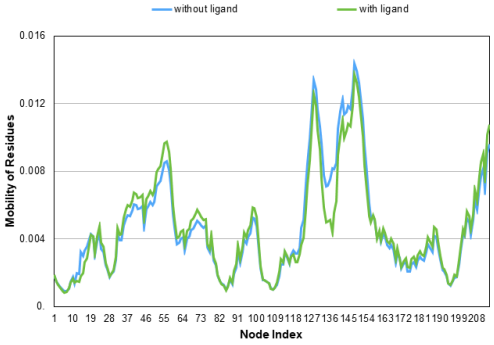
(e) gen5_6.



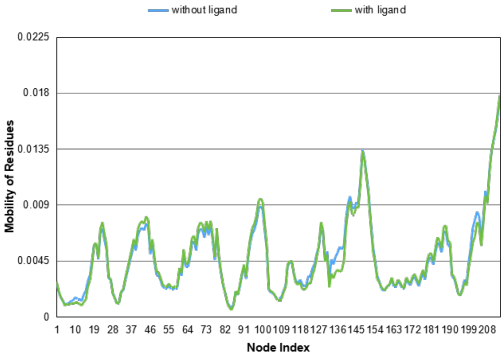
(f) gen6_18.



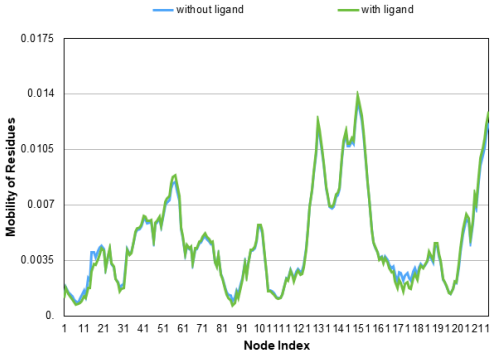
(g) gen6_22.



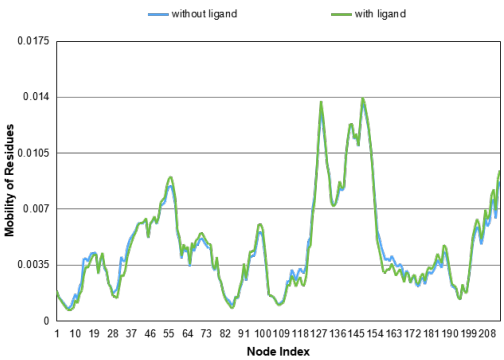
(h) gen7_2.



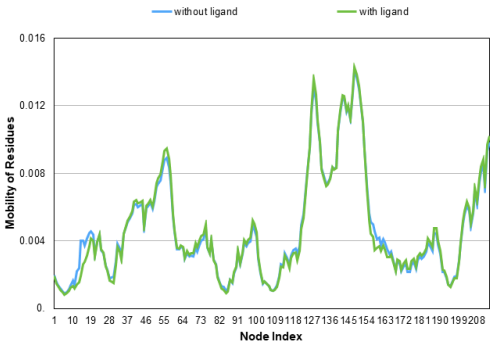
(i) 1AKE.



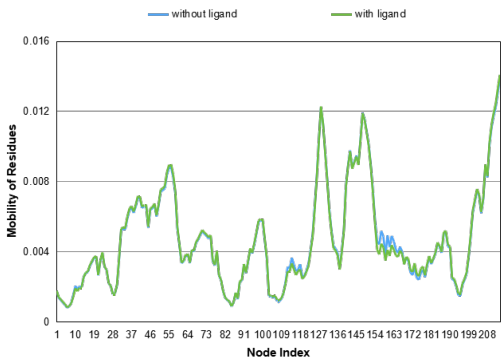
(j) gen1_1.



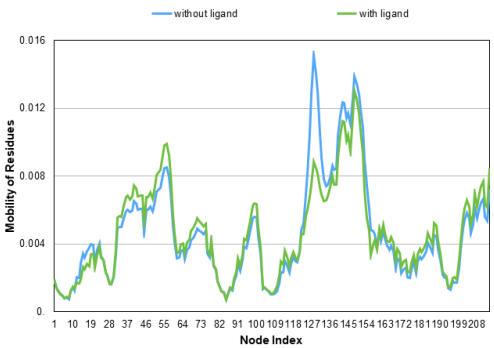
(k) gen4_2.



(l) gen4_5.



(m) gen7_19.



(n) gen7_25.

Figure 4.32. Mean square fluctuations of residues for chosen structures.

4.6. Verification with an Alternative Docking Program

In order to verify the results of ATP_LID_Flex and AMP_NMP_Flex with an alternative docking program, GOLD (Jones *et al.*, 1997; Verdonk *et al.*, 2003) was used to dock fully flexible ligands. Several conformers with different ligand RMSDs (low to high values in Autodock) were chosen for this purpose. Tables 4.8 and 4.9 summarize the results for ATP and AMP dockings,respectively. The first and second rows for GOLD gives the results from the first (best scoring) cluster and the average values for all clusters, respectively. In the Autodock case, first row gives the first (best scoring) cluster results and other rows give alternative populated clusters listed in Appendix A.5

- It needs to be compared with ATP_LID_Flex and AMP_NMP_Flex results in Autodock case, since GOLD considers all the ligands as fully flexible itself. Represented conformers in Tables 4.8 and 4.9 are chosen according to biggest, middle and lowest RMSD values among all of them.
- Result 1: ATP results are more consistent with Autodock.
- Result 2: AMP cases are much better in terms of RMSD but not of energy values in Autodock cases.

It can be said that in general Autodock results are in agreement with GOLD results in terms of the ligand RMSDs. GOLD produces the same level of higher ligand RMSDs for almost all conformers in the tables. As a result of this comparison, it can be said that we can rely on the detailed work carried out with Autodock in the previous sections.

Table 4.8. GOLD results of ATP_LID_Flex.

	GOLD Score	**GOLD RMSD (Å)**	**Autodock Score (kcal/mol)**	**Autodock RMSD - Å (Cluster)**
gen1_1	41.93 36.32	11.8 9.9	-7.93	11.57
gen3_4	65.14 53.94	3.8 7.6	-9.85 -6.95	10.80 1.89 (2)
gen4_5	69.24 61.75	9.1 8.8	-9.13	7.34
gen5_3	71.30 54.26	11.1 9.5	-8.75	8.99
gen6_18	62.44 53.38	2.22 5.4	-9.60	1.65
gen6_23	63.47 54.24	10.6 10.1	-8.53 -8.31 -7.85	6.60 7.35 (3) 4.78 (4)
gen7_2	58.64 52.72	6.1 5.9	-8.67	3.93
gen7_6	47.32 44.86	10.8 11.1	-8.09	8.60
gen7_18	44.79 37.69	10.9 11.1	-8.40	8.10

Table 4.9. GOLD results of AMP_NMP_Flex.

	GOLD Score	**GOLD RMSD (Å)**	**Autodock Score (kcal/mol)**	**Autodock RMSD - Å (Cluster)**
gen2_3	33.41 30.86	5.5 5.6	-6.54	5.57
gen4_5	29.74 28.42	11.9 6.5	-6.56 -6.45	9.83 6.07 (2)
gen5_6	33.10 30.87	5.4 5.6	-6.36	15.90
gen5_14	33.51 37.94	8.3 5.5	-7.16	8.00
gen6_11	46.18 42.75	4.3 4.9	-7.02	4.64
gen6_12	39.34 36.81	4.4 5.4	-6.11 -5.64	3.07 2.95 (2)
gen6_17	42.76 37.94	6.5 5.5	-6.84	4.96
gen7_29	38.94 37.27	4.6 5.3	-6.54	4.53

5. CONCLUSION AND RECOMMENDATIONS

In this thesis, a computational docking study is performed using Autodock to understand binding mechanisms of ligands to the enzyme ADK, undergoing large conformational changes (Huey *et al.*, 2007; Morris *et al.*, 2009). Atomistic conformers of ADK have previously been obtained by an ENM-based comformational search methodology called ClustENM (Kurkcuoglu, 2015). These intermediate structures, as they are called 'conformers' in this thesis, are performed by aligning all of them to the closed structure of ADK, 1AKE before starting the docking procedure.

First, the internal geometry of ClustENM conformers are validated by using Molprobity server (Chen *et al.*, 2010; Davis *et al.*, 2007). All seven generations gave the best Molprobity scores as 100th, together with minimized crystal structures. Dockings of ligands to the reference crystal structures are carried out to have a reference for the other dockings to the intermediate conformers. All the cases give satisfactory results in terms of ligand RMSD values less than 1Å. And the lowest energy is obtained in the docking of biggest ligand, AP5, to crystal structure, 1AKE. All ligands, AP5, ATP, ADP and AMP are successfully docked to LID or NMP domains, with energies being lower on the NMP domain side.

Next step is to performed 13 different rounds of dockings, specifically docking of AP5, ATP, ADP and AMP to different domain locations by setting up the related docking areas (i.e. box parameters). Each case consists of docking to 92 conformers to be evaluated separately in terms of scoring (binding energy) and ligand RMSD. ATP_LID_Flex is the best case within all these docking cases, since it has 38 conformers less than 2Å RMSD level. For all the dockings to NMP domain, there is no any conformer with such RMSD level, so it is confirmed that ligands mainly prefer to bound with LID domain, instead of NMP. Also the minimum ligand RMSD is obtained with the case where ATP is docked to LID domain, as 0.63 Å. If it is aimed to see the success of generations within these docking cases, generation 5 and 6 give the maximum number of conformers with RMSD less than 2Å, including the best case ATP_LID_Flex.

Ligplot and dynomics server are utilized to analyze the protein-ligands interactions and changes in the dynamics upon binding by following GNM and ANM procedures for the ATP docking to LID domain with fully flexible ligand structure (Laskowski and Swindells, 2011; Wallace *et al.*, 1995). Mainly 13 conformers are selected presenting a wide range of ligand RMSD values. Except the worst case, gen7_19, each conformers has some common interactions with the closed structure, 1AKE, and almost all of these conformers has the common residues which guide the ligand and interacting with phosphate groups. Same conformers are further investigated with the Dynomics server to observe changes in degree of mode collectivities and mean square fluctuations upon binding (Li *et al.*, 2017). In general, mobility decreases between the residues 130-140 with the presence of ligand for almost all the cases which LID domain is closed or partially closed. For the cases both LID and NMP open, and only NMP closed, mobilities are nearly the same with and without ligand.

As a last part to verify the docking results with a different procedure GOLD, some conformers are investigated both for ATP_LID_Flex and AMP_NMP_Flex cases (Jones *et al.*, 1997; Verdonk *et al.*, 2003). Results are mostly in accordance with Autodock procedure especially for ATP case, and it is observed that there is no any inconsistent results.

As for all ligands, the best docking results are obtained to LID domain, there may be a sequence of events of binding where LID may be the first one ligands prefer to bind and NMP comes after. There are many intermediate conformers, to which all the ligands prefer to bind with minimum energy and RMSD levels, which points to a combination of induced fit and population shift mechanisms of binding.

There is not such a detailed and successful docking study to ADK in the literature that investigates diverse flexible conformers. Therefore ClustENM proves to be a computationally efficient unbiased conformational search algorithm suitable for docking purposes.

For further studies, it is recommended to perform MD simulations to successful docking poses to observe their stability. Also a second round of ClustENM procedure could be applied to the the successful poses with the ligand to observe whether the closed structure will be approached in a more effective way.

REFERENCES

Arora, K. and Brooks, C. L. (2007). Large-scale allosteric conformational transitions of adenylate kinase appear to involve a population-shift mechanism. *Proceedings of the National Academy of Sciences*, 104(47):18496–18501.

Atilgan, A. R., Durell, S., Jernigan, R. L., Demirel, M., Keskin, O., and Bahar, I. (2001). Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophysical journal*, 80(1):505–515.

Azam, S. S. and Abbasi, S. W. (2013). Molecular docking studies for the identification of novel melatoninergic inhibitors for acetylserotonin-o-methyltransferase using different docking routines. *Theoretical Biology and Medical Modelling*, 10(1):63.

Bahar, I., Atilgan, A. R., and Erman, B. (1997). Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Folding and Design*, 2(3):173–181.

Bahar, I., Chennubhotla, C., and Tobi, D. (2007). Intrinsic dynamics of enzymes in the unbound state and relation to allosteric regulation. *Current opinion in structural biology*, 17(6):633–640.

Bahar, I., Lezon, T. R., Bakan, A., and Shrivastava, I. H. (2009). Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins. *Chemical reviews*, 110(3):1463–1497.

Beckstein, O., Denning, E. J., Perilla, J. R., and Woolf, T. B. (2009). Zipping and unzipping of adenylate kinase: atomistic insights into the ensemble of open↔ closed transitions. *Journal of molecular biology*, 394(1):160–176.

Bhatt, D. and Zuckerman, D. M. (2010). Heterogeneous path ensembles for conformational transitions in semiatomistic models of adenylate kinase. *Journal of chemical theory and computation*, 6(11):3527–3539.

Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., ... Kollman, P. A. (2012). Amber 12. *University of California, San Francisco.*

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., ... Richardson, D. C. (2010). Molprobity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, 66(1):12–21.

Daily, M. D., Phillips Jr, G. N., and Cui, Q. (2010). Many local motions cooperate to produce the adenylate kinase conformational transition. *Journal of molecular biology*, 400(3):618–631.

Das, A., Gur, M., Cheng, M. H., Jo, S., Bahar, I., and Roux, B. (2014). Exploring the conformational transitions of biomolecular systems using a simple two-state anisotropic network model. *PLoS computational biology*, 10(4):e1003521.

Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., ... Richardson, J. S. (2007). Molprobity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic acids research*, 35(suppl_2):W375–W383.

Doruker, P., Atilgan, A. R., and Bahar, I. (2000). Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: Application to α-amylase inhibitor. *Proteins: Structure, Function, and Bioinformatics*, 40(3):512–524.

Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., ... Lee, T. (2003). A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of computational chemistry*, 24(16):1999–2012.

Feng, Y., Yang, L., Kloczkowski, A., and Jernigan, R. L. (2009). The energy profiles of atomic conformational transition intermediates of adenylate kinase. *Proteins: Structure, Function, and Bioinformatics*, 77(3):551–558.

Flores, S. C. and Gerstein, M. B. (2011). Predicting protein ligand binding motions with the conformation explorer. *BMC bioinformatics*, 12(1):417.

Frauenfelder, H., Sligar, S. G., and Wolynes, P. G. (1991). The energy landscapes and motions of proteins. *Science*, 254(5038):1598–1603.

Gershenson, A., Gierasch, L. M., Pastore, A., and Radford, S. E. (2014). Energy landscapes of functional proteins are inherently risky. *Nature chemical biology*, 10(11):884.

Goodsell, D. S., Morris, G. M., and Olson, A. J. (1996). Automated docking of flexible ligands: applications of autodock. *Journal of Molecular Recognition*, 9(1):1–5.

Grant, B. J., Gorfe, A. A., and McCammon, J. A. (2010). Large conformational changes in proteins: signaling and other functions. *Current opinion in structural biology*, 20(2):142–147.

Haliloglu, T., Bahar, I., and Erman, B. (1997). Gaussian dynamics of folded proteins. *Physical review letters*, 79(16):3090.

Hanson, J. A., Duderstadt, K., Watkins, L. P., Bhattacharyya, S., Brokaw, J., Chu, J.-W. and Yang, H. (2007). Illuminating the mechanistic roles of enzyme conformational dynamics. *Proceedings of the National Academy of Sciences*, 104(46):18055–18060.

Hawkins, G. D., Cramer, C. J., and Truhlar, D. G. (1995). Pairwise solute descreening of solute charges from a dielectric medium. *Chemical Physics Letters*, 246(1-2):122–129.

Hawkins, G. D., Cramer, C. J., and Truhlar, D. G. (1996). Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *The Journal of Physical Chemistry*, 100(51):19824–19839.

Hess, B., Kutzner, C., Van Der Spoel, D., and Lindahl, E. (2008). Gromacs 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of chemical theory and computation*, 4(3):435–447.

Hills, R. D. and Brooks, C. L. (2009). Insights from coarse-grained go models for protein folding and dynamics. *International journal of molecular sciences*, 10(3):889—-905.

Hinsen, K., Thomas, A., and Field, M. J. (1999). Analysis of domain motions in large proteins. *Proteins: Structure, Function, and Bioinformatics*, 34(3):369–382.

Huey, R., Morris, G. M., Olson, A. J., and Goodsell, D. S. (2007). A semiempirical free energy force field with charge-based desolvation. *Journal of computational chemistry*, 28(6):1145–1152.

Jones, G., Willett, P., Glen, R. C., Leach, A. R., and Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. *Journal of molecular biology*, 267(3):727–748.

Kantarci-Carsibasi, N., Haliloglu, T., and Doruker, P. (2008). Conformational transition pathways explored by monte carlo simulation integrated with collective modes. *Biophysical journal*, 95(12):5862–5873.

Kerns, S. J., Agafonov, R. V., Cho, Y.-J., Pontiggia, F., Otten, R., Pachov, D. V., ... Thai, V. (2015). The energy landscape of adenylate kinase during catalysis. *Nature structural & molecular biology*, 22(2):124.

Kessel, A. and Ben-Tal, N. (2010). *Introduction to proteins: structure, function, and motion*. CRC Press.

Kurkcuoglu, Z. (2015). *Elastic Network Model Based Approaches for Conformer Generation and Docking Applications*. Ph.D. Thesis, Bogazici University.

Kurkcuoglu, Z., Bahar, I., and Doruker, P. (2016). Clustenm: Enm-based sampling of essential conformational space at full atomic resolution. *Journal of chemical theory and computation*, 12(9):4549–4562.

Kurkcuoglu, Z. and Doruker, P. (2016). Ligand docking to intermediate and close-to-bound conformers generated by an elastic network model based algorithm for highly flexible proteins. *PloS one*, 11(6):e0158063.

Laskowski, R. A. and Swindells, M. B. (2011). Ligplot+: multiple ligand–protein interaction diagrams for drug discovery.

Leach, A. R. and Leach, A. (2001). *Molecular modelling: principles and applications*. Pearson education.

Li, H., Chang, Y.-Y., Lee, J. Y., Bahar, I., and Yang, L.-W. (2017). Dynomics: dynamics of structural proteome and beyond. *Nucleic acids research*, 45(W1):W374–W380.

Lou, H. and Cukier, R. I. (2006). Molecular dynamics of apo-adenylate kinase: a distance replica exchange method for the free energy of conformational fluctuations. *The journal of physical chemistry B*, 110(47):24121–24137.

Lu, Q. and Wang, J. (2008). Single molecule conformational dynamics of adenylate kinase: energy landscape, structural correlations, and transition state ensembles. *Journal of the American Chemical Society*, 130(14):4772–4783.

Mehler, E. L. and Solmajer, T. (1991). Electrostatic effects in proteins: comparison of dielectric and charge models. *Protein Engineering, Design and Selection*, 4(8):903–910.

Ming, D. and Wall, M. E. (2005). Allostery in a coarse-grained model of protein dynamics. *Physical review letters*, 95(19):198103.

Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A. J. (1998). Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function. *Journal of computational chemistry*, 19(14):1639–1662.

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. and Olson, A. J. (2009). Autodock4 and autodocktools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry*, 30(16):2785–2791.

Müller, C., Schlauderer, G., Reinstein, J., and Schulz, G. E. (1996). Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding. *Structure*, 4(2):147–156.

Müller, C. W. and Schulz, G. E. (1992). Structure of the complex between adenylate kinase from escherichia coli and the inhibitor ap5a refined at 1.9 å resolution: A model for a catalytic transition state. *Journal of molecular biology*, 224(1):159–177.

Okazaki, K.-i. and Takada, S. (2008). Dynamic energy landscape view of coupled binding and protein conformational change: induced-fit versus population-shift mechanisms. *Proceedings of the National Academy of Sciences*, 105(32):11182–11187.

Onuchic, J. N., Luthey-Schulten, Z., and Wolynes, P. G. (1997). Theory of protein folding: the energy landscape perspective. *Annual review of physical chemistry*, 48(1):545–600.

Seyler, S. L. and Beckstein, O. (2014). Sampling large conformational transitions: adenylate kinase as a testing ground. *Molecular Simulation*, 40(10-11):855–877.

Solis, F. J. and Wets, R. J.-B. (1981). Minimization by random search techniques. *Mathematics of operations research*, 6(1):19–30.

Tama, F. and Sanejouand, Y.-H. (2001). Conformational change of proteins arising from normal mode calculations. *Protein engineering*, 14(1):1–6.

Tirion, M. M. (1996). Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Physical review letters*, 77(9):1905.

Tobi, D. and Bahar, I. (2005). Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. *Proceedings of the National Academy of Sciences*, 102(52):18908–18913.

Trott, O. and Olson, A. J. (2010). Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2):455–461.

Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W., and Taylor, R. D. (2003). Improved protein–ligand docking using gold. *Proteins: Structure, Function, and Bioinformatics*, 52(4):609–623.

Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995). Ligplot: a program to generate schematic diagrams of protein-ligand interactions. *Protein engineering, design and selection*, 8(2):127–134.

Wang, Y., Gan, L., Wang, E., and Wang, J. (2012). Exploring the dynamic functional landscape of adenylate kinase modulated by substrates. *Journal of chemical theory and computation*, 9(1):84–95.

Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., ... Weiner, P. (1984). A new force field for molecular mechanical simulation of nucleic acids and proteins. *Journal of the American Chemical Society*, 106(3):765–784.

Whitford, P. C., Sanbonmatsu, K. Y., and Onuchic, J. N. (2012). Biomolecular dynamics: order–disorder transitions and energy landscapes. *Reports on Progress in Physics*, 75(7):076601.

Zeller, F. and Zacharias, M. (2015). Substrate binding specifically modulates domain arrangements in adenylate kinase. *Biophysical journal*, 109(9):1978–1985.

APPENDIX A: APPLICATION

Table A.1. Naming of conformers.

Conformers	**Conformers_New**	**theta_LID**	**theta_NMP**
gen1_25	gen1_1	140.4500	66.1629
gen2_25.5	gen2_1	146.5539	60.0893
gen2_25.15	gen2_2	136.3910	69.7786
gen2_25.17	gen2_3	127.0634	62.9480
gen3_5.10	gen3_1	117.1970	82.4394
gen3_5.23	gen3_2	121.4745	70.0799
gen3_10.5	gen3_3	125.5382	74.6383
gen3_17.20	gen3_4	115.6110	64.4432
gen3_17.23	gen3_5	121.3328	65.0502
gen4_5.21	gen4_1	113.7299	54.2704
gen4_10.18	gen4_2	105.6298	71.0551
gen4_10.21	gen4_3	137.8842	62.8551
gen4_15.5	gen4_4	123.4789	78.4431
gen4_20.2	gen4_5	118.3940	64.2951
gen4_20.23	gen4_6	128.2947	66.7596
gen4_23.1	gen4_7	115.6972	85.9548
gen4_23.3	gen4_8	115.1064	70.7786
gen4_23.13	gen4_9	117.0612	67.6379
gen4_23.23	gen4_10	119.0748	85.2033
gen5_1.1	gen5_1	136.8855	70.0882
gen5_1.11	gen5_2	134.9310	60.5797
gen5_1.20	gen5_3	128.5925	58.8109
gen5_3.17	gen5_4	112.5624	76.9471

Table A.1. Naming of conformers. (cont.)

Conformers	**Conformers_New**	**theta_LID**	**theta_NMP**
gen5_3.20	gen5_5	107.5008	82.2528
gen5_3.26	gen5_6	101.4032	61.4685
gen5_5.12	gen5_7	128.2408	46.9494
gen5_5.17	gen5_8	115.6613	56.2676
gen5_5.23	gen5_9	112.8396	57.8434
gen5_13.21	gen5_10	108.8710	91.3227
gen5_13.26	gen5_11	117.8282	92.0375
gen5_18.5	gen5_12	119.1728	88.7287
gen5_18.12	gen5_13	126.3866	67.8602
gen5_18.16	gen5_14	118.5458	58.9627
gen5_18.27	gen5_15	111.8360	68.8180
gen5_21.13	gen5_16	114.2051	58.2765
gen5_21.17	gen5_17	139.1489	58.3331
gen5_21.19	gen5_18	138.1923	63.2909
gen5_23.25	gen5_19	126.8694	59.2855
gen6_11.24	gen6_1	125.9466	57.7901
gen6_11.26	gen6_2	136.0297	57.1974
gen6_12.1.5	gen6_3	136.2893	60.5558
gen6_12.5	gen6_4	128.2331	67.2633
gen6_12.11	gen6_5	102.6788	66.2079
gen6_12.13	gen6_6	97.6537	66.4624
gen6_12.15	gen6_7	144.9892	60.2720
gen6_12.19	gen6_8	150.9046	62.1157

Table A.1. Naming of conformers. (cont.)

Conformers	**Conformers_New**	**theta_LID**	**theta_NMP**
gen6_16.5	gen6_9	126.7218	49.4850
gen6_16.25	gen6_10	127.2570	49.6849
gen6_16.27	gen6_11	134.2865	44.1812
gen6_17_1.20	gen6_12	111.9879	90.5086
gen6_17.20	gen6_13	126.1455	59.6817
gen6_17.23	gen6_14	111.0518	54.1268
gen6_17.25	gen6_15	119.8197	54.4776
gen6_19.4	gen6_16	114.9249	82.1451
gen6_20.5	gen6_17	120.8935	55.8513
gen6_21.20	gen6_18	105.4108	71.8139
gen6_25.2	gen6_19	134.8335	70.7159
gen6_25.24	gen6_20	120.2621	60.3548
gen6_26.20	gen6_21	117.8028	87.1222
gen6_26.23	gen6_22	111.2931	61.2190
gen6_27.25	gen6_23	110.5398	62.1617
gen7_2.13	gen7_1	132.1399	63.0368
gen7_2.23	gen7_2	147.8603	51.5516
gen7_4.2	gen7_3	127.5567	70.5363
gen7_4.16	gen7_4	110.0194	71.4995
gen7_5_1.12	gen7_5	100.0847	65.9500
gen7_5.12	gen7_6	100.2971	71.1184
gen7_5.15	gen7_7	102.0581	57.7218
gen7_5.18	gen7_8	95.9357	68.4086

Table A.1. Naming of conformers. (cont.)

Conformers	**Conformers_New**	**theta_LID**	**theta_NMP**
gen7_5.20	gen7_9	95.4933	64.9624
gen7_5.25	gen7_10	122.9715	81.3281
gen7_11.11	gen7_11	122.3585	67.1199
gen7_11.15	gen7_12	105.5108	87.2172
gen7_13.17	gen7_13	119.1942	53.8249
gen7_15.5	gen7_14	114.8990	66.8928
gen7_15.15	gen7_15	114.1909	61.1167
gen7_15.26	gen7_16	107.2285	64.0771
gen7_19.20	gen7_17	116.9462	60.9942
gen7_19.25	gen7_18	114.9032	62.3086
gen7_19.27	gen7_19	121.3951	47.3724
gen7_20.26	gen7_20	112.2744	51.5374
gen7_23.3	gen7_21	102.3736	58.8942
gen7_23.18	gen7_22	148.2683	55.6706
gen7_23.23	gen7_23	132.7636	51.3865
gen7_23.26	gen7_24	108.0942	89.5974
gen7_24.18	gen7_25	114.4303	71.2517
gen7_25.1	gen7_26	132.1813	43.5482
gen7_25.7	gen7_27	162.7087	56.0417
gen7_25.10	gen7_28	145.9724	45.9045
gen7_25.23	gen7_29	124.9552	45.6707
gen7_26.5	gen7_30	133.7632	57.3453
gen7_27.18	gen7_31	122.6911	54.1358

Table A.2. Molprobity results.

	Molprobity Score	**Poor Rotamers**	**Ramachandran Favored (%)**
1AKE	1.87 / 82nd	A54 - ASP A72-ILE A74-GLN A75-GLU A78-ARG A157-LYS B1-MET B2-ARG B40-LYS B74-GLN B102-ASN B160-GLN B191-THR	96.93
4AKE	2.53 / 48th	A16-GLN A78-ARG A91-PRO A148-VAL A155-THR A160-GLN A162-GLU A166-LYS A170-GLU A195-LYS A206-ARG B18-GLN B45-LEU B75-GLU B125-VAL B142-VAL B148-VAL B174-MET	95.05
1AKE_min	0.55 / 100th	19-PHE 78-ARG5	99.01
4AKE_min	0.50 / 100th	195-LYS	98.52
gen1_1	0.60 / 100th	195-LYS	97.54

Table A.2. Molprobity results. (cont.)

	Molprobity Score	**Poor Rotamers**	**Ramachandran Favored (%)**
gen2_1	0.67 / 100th	195-LYS	94.19
gen2_2	0.60 / 100th	195-LYS	97.54
gen2_3	0.50 / 100th	195-LYS	98.03
gen3_1	0.60 / 100th	195-LYS	97.54
gen3_2	0.60 / 100th	195-LYS	97.54
gen3_3	0.60 / 100th	195-LYS	97.54
gen3_4	0.50 / 100th	195-LYS	98.03
gen3_5	0.60 / 100th	195-LYS	97.54
gen4_1	0.50 / 100th	195-LYS	98.52
gen4_2	0.85 / 100th	163-THR 170-GLU 195-LYS	97.04
gen4_3	1.00 / 100th	109-PHE 131-ARG 163-THR 195-LYS	96.55
gen4_4	0.67 / 100th	195-LYS	97.04
gen4_5	0.50 / 100th	195-LYS	98.03
gen4_6	0.60 / 100th	195-LYS	97.54
gen4_7	0.50 / 100th	195-LYS	98.03
gen4_8	0.55 / 100th	109-PHE 195-LYS	98.03
gen4_9	0.72 / 100th	163-THR 195-LYS	97.04
gen4_10	0.72 / 100th	163-THR 195-LYS	97.04
gen5_1	0.50 / 100th	195-LYS	98.03
gen5_2	0.60 / 100th	195-LYS	97.54

Table A.2. Molprobity results. (cont.)

	Molprobity Score	**Poor Rotamers**	**Ramachandran Favored (%)**
gen5_3	0.50 / 100th	195-LYS	98.03
gen5_4	0.50 / 100th	195-LYS	98.03
gen5_5	0.64 / 100th	109-PHE 195-LYS	97.54
gen5_6	0.50 / 100th	195-LYS	98.52
gen5_7	0.67 / 100th	195-LYS	97.04
gen5_8	0.77 / 100th	195-LYS	96.06
gen5_9	0.67 / 100th	195-LYS	97.04
gen5_10	0.64 / 100th	163-THR 195-LYS	97.54
gen5_11	0.72 / 100th	163-THR 195-LYS	97.04
gen5_12	0.64 / 100th	163-THR 195-LYS	97.54
gen5_13	0.60 / 100th	195-LYS	97.54
gen5_14	0.77 / 100th	195-LYS	96.06
gen5_15	0.60 / 100th	195-LYS	97.54
gen5_16	0.50 / 100th	195-LYS	98.52
gen5_17	0.55 / 100th	47-LYS 195-LYS	99.01
gen5_18	0.50 / 100th	195-LYS	98.03
gen5_19	0.50 / 100th	195-LYS	98.52
gen6_1	0.67 / 100th	195-LYS	97.04
gen6_2	0.60 / 100th	195-LYS	97.54
gen6_3	0.67 / 100th	195-LYS	97.04
gen6_4	0.60 / 100th	195-LYS	97.54

Table A.2. Molprobity results. (cont.)

	Molprobity Score	**Poor Rotamers**	**Ramachandran Favored (%)**
gen6_5	0.60 / 100th	195-LYS	97.54
gen6_6	0.60 / 100th	195-LYS	97.54
gen6_7	0.72 / 100th	195-LYS	96.55
gen6_8	0.50 / 100th	195-LYS	98.03
gen6_9	0.72 / 100th	195-LYS	96.55
gen6_10	0.77 / 100th	109-PHE 195-LYS	96.55
gen6_11	0.67 / 100th	195-LYS	97.04
gen6_12	0.72 / 100th	195-LYS	96.55
gen6_13	0.68 / 100th	47-LYS 163-THR 195-LYS	99.01
gen6_14	0.86 / 100th	163-THR 195-LYS	95.57
gen6_15	0.72 / 100th	195-LYS	96.55
gen6_16	0.55 / 100th	47-LYS 195-LYS	98.52
gen6_17	0.50 / 100th	195-LYS	98.03
gen6_18	0.72 / 100th	163-THR 195-LYS	97.04
gen6_19	0.50 / 100th	195-LYS	98.03
gen6_20	0.50 / 100th	195-LYS	98.52
gen6_21	0.64 / 100th	163-THR 195-LYS	97.54
gen6_22	0.50 / 100th	195-LYS	98.03
gen6_23	0.72 / 100th	195-LYS	96.55
gen7_1	0.50 / 100th	195-LYS	98.52

Table A.2. Molprobity results. (cont.)

	Molprobity Score	**Poor Rotamers**	**Ramachandran Favored (%)**
gen7_2	0.50 / 100th	195-LYS	98.03
gen7_3	0.64 / 100th	47-LYS 195-LYS	97.54
gen7_4	0.55 / 100th	47-LYS 195-LYS	98.52
gen7_5	0.67 / 100th	195-LYS	97.04
gen7_6	0.77 / 100th	195-LYS	96.06
gen7_7	0.67 / 100th	195-LYS	97.04
gen7_8	0.72 / 100th	195-LYS	96.55
gen7_9	0.82 / 100th	131-ARG 195-LYS	96.06
gen7_10	0.60 / 100th	195-LYS	97.54
gen7_11	0.60 / 100th	195-LYS	97.54
gen7_12	0.67 / 100th	195-LYS	97.04
gen7_13	0.67 / 100th	195-LYS	97.04
gen7_14	0.60 / 100th	195-LYS	97.54
gen7_15	0.64 / 100th	47-LYS 195-LYS	97.54
gen7_16	0.72 / 100th	195-LYS	96.55
gen7_17	0.50 / 100th	195-LYS	98.52
gen7_18	0.55 / 100th	47-LYS 195-LYS	98.03
gen7_19	0.55 / 100th	138-ASN 195-LYS	98.03
gen7_20	0.68 / 100th	47-LYS 163-THR 195-LYS	99.01
gen7_21	0.50 / 100th	195-LYS	98.03

Table A.2. Molprobity results. (cont.)

	Molprobity Score	**Poor Rotamers**	**Ramachandran Favored (%)**
gen7_22	0.64 / 100th	163-THR 195-LYS	97.54
gen7_23	0.67 / 100th	195-LYS	97.04
gen7_24	0.77 / 100th	163-THR 195-LYS	96.55
gen7_25	0.82 / 100th	109-PHE 195-LYS	96.06
gen7_26	0.77 / 100th	195-LYS	96.06
gen7_27	0.67 / 100th	195-LYS	97.04
gen7_28	0.72 / 100th	195-LYS	96.55
gen7_29	0.72 / 100th	195-LYS	96.55
gen7_30	0.60 / 100th	195-LYS	97.54
gen7_31	0.77 / 100th	109-PHE 195-LYS	96.55

Table A.3. Dali server structure comparison.

	Chain	**z**	**RMSD**	**lali**	**nres**	**%id**
4IKE	A	28.5	1.7	199	206	47
4IKE	B	24.9	2.0	186	206	48
4JL5	A	29.5	1.5	199	203	47
4JL5	B	28.6	1.8	199	203	47
4CF7	A	28.8	1.7	199	203	47
4CF7	B	29.1	1.7	199	203	47
1AKE	4CF7-A	28.8	1.7	199	203	47
1AKE	4JL5-A	29.5	1.5	199	203	47

Table A.4. Autodock results of ATP_LID.

Confor-mers(no of clusters)	**Binding energy (kcal/mol)**	**Ligand RMSD (Å)**	**Cluster**	**Poses**	**Overall-backbone RMSD (Å)**
gen1_1 (3)	-5.00 -4.91	11.25 7.92	1 2	4 63	6.37
gen2_1 (4)	-5.61 -5.43	3.32 7.89	1 2	49 35	5.04
gen2_2	-5.13	6.75	1	96	6.38
gen2_3 (4)	-5.16 -4.99	11.92 1.38	1 3	1 82	5.81
gen3_1	-5.29	6.10	1	99	4.78
gen3_2	-5.85	0.87	1	100	4.12
gen3_3	-6.38	2.27	1	99	6.35
gen3_4	-6.26	0.81	1	100	4.68
gen3_5	-5.51	2.28	1	90	7.25
gen4_1	-7.38	2.76	1	99	6.84
gen4_2	-6.15	6.60	1	98	4.33
gen4_3	-6.44	6.17	1	99	5.92
gen4_4 (4)	-4.97 -4.81	6.69 8.78	1 2	31 56	6.97
gen4_5	-5.71	1.33	1	100	6.21
gen4_6	-5.99	0.63	1	100	3.92
gen4_7	-6.11	1.61	1	82	4.80
gen4_8	-7.17	1.66	1	100	4.23
gen4_9	-6.23	2.30	1	99	7.88
gen4_10	-5.46	2.71	1	78	6.58
gen5_1	-5.07	2.18	1	100	6.18
gen5_2	-5.35	1.89	1	92	4.83

Table A.4. Autodock results of ATP_LID. (cont.)

Conformers (no of clusters)	**Binding energy (kcal/mol)**	**Ligand RMSD (Å)**	**Cluster**	**Poses**	**Overall-backbone RMSD (Å)**
gen5_3 (4)	-5.03 -5.02	7.10 2.14	1 2	24 62	4.01
gen5_4	-6.46	1.66	1	100	3.54
gen5_5	-7.14	1.53	1	100	4.23
gen5_6	-7.34	1.75	1	100	2.89
gen5_7	-5.28	7.93	1	100	5.73
gen5_8	-5.12	10.53	1	88	8.27
gen5_9	-5.48 -5.06	6.64 7.41	1 2	55 45	6.32
gen5_10	-7.14	1.35	1	100	6.25
gen5_11	-5.83	1.73	1	89	7.82
gen5_12	-5.86	8.33	1	86	5.80
gen5_13	-5.71	7.01	1	97	4.47
gen5_14	-6.33	6.39	1	100	4.75
gen5_15	-6.03	6.49	1	99	3.41
gen5_16	-7.28	1.87	1	61	7.26
gen5_17	-7.05	3.63	1	93	7.85
gen5_18	-5.66	7.03	1	98	7.66
gen5_19	-5.90	1.53	1	72	5.95
gen6_1	-6.35	1.99	1	100	3.95
gen6_2 (4)	-5.66 -5.48	10.94 2.23	1 2	23 61	4.62
gen6_3	-5.54 -5.08	1.50 6.10	1 2	10 90	6.96

Table A.4. Autodock results of ATP_LID. (cont.)

Conformers (no of clusters)	**Binding energy (kcal/mol)**	**Ligand RMSD (Å)**	**Cluster**	**Poses**	**Overall-backbone RMSD (Å)**
gen6_4 (3)	-5.15 -4.86	11.63 6.62	1 2	3 72	7.07
gen6_5	-6.00	5.25	1	100	4.91
gen6_6	-5.16	6.14	1	100	5.93
gen6_7	-5.73	6.70	1	92	4.43
gen6_8	-5.41	7.24	1	82	5.79
gen6_9 (4)	-5.68 -5.20	11.29 6.59	1 4	9 73	6.34
gen6_10	-6.27	6.47	1	100	3.69
gen6_11	-5.46	5.30	1	74	3.72
gen6_12	-6.26	1.78	1	96	2.80
gen6_13	-6.52	4.46	1	88	7.72
gen6_14 (3)	-4.98 -4.87	11.19 1.75	1 2	18 80	7.79
gen6_15	-4.55	1.65	1	90	2.72
gen6_16 (4)	-6.58 -5.69	2.17 8.37	1 2	50 48	6.39
gen6_17	-6.76	5.22	1	100	3.71
gen6_18	-7.58	0.90	1	100	4.89
gen6_19	-6.67 -6.61	1.59 6.61	1 2	38 62	6.31
gen6_20 (4)	-5.19 -5.16	10.04 6.04	1 2	3 81	5.21
gen6_21	-5.74	2.14	1	100	9.08
gen6_22	-6.35	1.65	1	81	3.59

Table A.4. Autodock results of ATP_LID. (cont.)

Conformers (no of clusters)	**Binding energy (kcal/mol)**	**Ligand RMSD (Å)**	**Cluster**	**Poses**	**Overall-backbone RMSD (Å)**
gen6_23	-6.13	7.55	1	99	4.70
gen7_1	-6.15 -6.12	1.89 6.73	1 2	8 92	7.05
gen7_2	-7.75	1.40	1	100	5.04
gen7_3 (6)	-6.05 -5.35	1.83 9.03	1 2	18 53	4.81
gen7_4	-5.80	10.33	1	92	7.14
gen7_5	-5.89	5.70	1	98	4.09
gen7_6 (4)	-3.68 -3.66	7.77 7.10	1 2	38 50	6.69
gen7_7	-4.78	7.70	1	91	7.37
gen7_8	-6.93	6.74	1	100	6.27
gen7_9 (5)	-5.62 -5.41	5.89 9.28	1 2	29 37	5.38
gen7_10	-5.60	7.62	1	93	6.44
gen7_11	-5.82	6.21	1	100	4.82
gen7_12	-7.10 -6.09	11.07 5.27	1 2	24 76	4.91
gen7_13	-5.29	9.17	1	100	7.13
gen7_14 (5)	-5.20 -5.07	9.76 7.14	1 3	39 56	5.78
gen7_15	-5.94	6.14	1	96	5.04
gen7_16	-6.66	6.15	1	89	4.01
gen7_17	-5.68	8.21	1	62	4.50

Table A.4. Autodock results of ATP_LID. (cont.)

Conformers (no of clusters)	**Binding energy (kcal/mol)**	**Ligand RMSD (Å)**	**Cluster**	**Poses**	**Overall-backbone RMSD (Å)**
gen7_18 (7)	-5.25 -3.63	7.39 13.42	1 2	41 36	5.06
gen7_19	-4.61	6.36	1	90	5.55
gen7_20	-6.29	5.85	1	98	7.68
gen7_21	-3.94	7.74	1	84	3.59
gen7_22 (4)	-5.21 -4.60	7.79 5.70	1 3	36 39	7.40
gen7_23 (5)	-5.44 -4.79	2.17 8.53	1 2	9 66	5.06
gen7_24	-4.98	2.27	1	100	8.62
gen7_25	-6.27 -4.98	1.66 10.70	1 2	65 35	3.83
gen7_26	-5.07	10.32	1	99	4.07
gen7_27 (4)	-4.32 -3.57	11.57 14.78	1 3	46 42	3.42
gen7_28	-5.67	6.18	1	84	3.92
gen7_29 (3)	-4.81 -4.47	9.52 4.14	1 2	18 81	2.52
gen7_30 (3)	-4.28 -4.23	2.68 6.88	1 2	45 44	5.94
gen7_31	-5.64	6.48	1	65	3.95

Table A.5. Autodock results of ATP_LID_Flex.

Confor-mers (no of clusters)	**Binding energy (kcal/mol)**	**Ligand RMSD (Å)**	**Cluster**	**Poses**	**Overall-backbone RMSD (Å)**
gen1_1 (43)	-7.93 -7.11 -5.82	11.57 5.20 3.82	1 6 32	4 14 1	6.37
gen2_1 (33)	-8.75 -7.94 -7.25	11.28 3.66 5.11	1 3 10	6 12 13	5.04
gen2_2 (48)	-8.02 -7.16	7.23 1.80	1 6	15 1	6.38
gen2_3 (40)	-8.39 -7.97 -6.95	2.52 10.76 1.98(min.)	1 3 16	7 11 7	5.81
gen3_1 (30)	-9.31	4.54	1	18	4.78
gen3_2 (29)	-10.16 -9.02 -6.95	11.62 4.06 1.89	1 3 5	5 13 11	4.12
gen3_3 (35)	-9.81 -8.86	2.17 1.71(min.)	1 1	10 10	6.35
gen3_4 (33)	-9.85 -8.75 -8.88	10.80 1.89 3.54	1 2 3	6 7 11	4.68
gen3_5 (44)	-8.85 -8.25 -7.26	7.67 5.67 1.62	1 3 13	9 11 2	7.25
gen4_1 (28)	-10.29 -8.31 -9.33	2.46 1.73(min.) 10.03	1 1 3	11 2 14	6.84
gen4_2 (29)	-8.79 -8.66 -6.92	7.33 8.68 5.26	1 2 4	4 16 9	4.33
gen4_3 (36)	-8.36 -8.14 -6.78	6.64 6.11 4.81	1 5 9	7 12 11	5.92

Table A.5. Autodock results of ATP_LID_Flex. (cont.)

Confor-mers (no of clusters)	**Binding energy (kcal/mol)**	**Ligand RMSD (Å)**	**Cluster**	**Poses**	**Overall-backbone RMSD (Å)**
gen4_4 (33)	-8.07 -7.84 -7.43	6.11 7.84 4.62	1 2 7	9 27 3	6.97
gen4_5 (29)	-9.13 -8.62 -7.48	7.34 9.96 1.60	1 3 9	7 16 4	6.21
gen4_6 (31)	-9.06 -7.39	3.94 1.85	1 15	14 3	3.92
gen4_7 (40)	-9.34 -6.78	9.21 1.76	1 5	2 6	4.80
gen4_8 (27)	-9.66 -7.72	2.38 1.80	1 2	8 27	4.23
gen4_9 (36)	-9.33 -7.23	11.56 1.71	1 3	14 11	7.88
gen4_10 (40)	-8.94 -7.26	5.33 1.50	1 16	15 2	6.58
gen5_1 (33)	-9.13 -7.11	2.70 2.02(min.)	1 1	19 19	6.18
gen5_2 (38)	-8.60 -6.62	6.99 1.75	1 2	6 10	4.83
gen5_3 (44)	-8.75 -8.16 -6.62	8.99 2.51 1.31	1 3 16	8 14 4	4.01
gen5_4 (28)	-9.19 -7.70	1.95 1.85(min.)	1 1	56 56	3.54
gen5_5 (27)	-9.23	1.84	1	24	4.23
gen5_6 (26)	-9.71	1.73	1	31	2.89

Table A.5. Autodock results of ATP_LID_Flex. (cont.)

Conformers (no of clusters)	**Binding energy (kcal/mol)**	**Ligand RMSD (Å)**	**Cluster**	**Poses**	**Overall-backbone RMSD (Å)**
gen5_7 (34)	-8.24 -7.83 -6.65	6.15 6.54 3.84	1 2 5	4 21 7	5.73
gen5_8 (33)	-7.65 -7.57 -6.54	5.52 10.29 1.70	1 2 12	9 12 2	8.27
gen5_9 (39)	-8.12 -6.07	7.81 3.16	1 25	17 2	6.32
gen5_10 (42)	-8.94 -8.44	3.38 1.74	1 2	25 11	6.25
gen5_11 (33)	-9.07 -7.92	12.30 2.60	1 2	3 14	7.82
gen5_12 (45)	-8.82 -8.55 -5.81	6.95 8.94 4.29	1 2 28	11 13 1	5.80
gen5_13 (33)	-8.27	6.95	1	18	4.47
gen5_14 (34)	-8.88 -7.44	8.41 7.08	1 12	8 11	4.75
gen5_15 (35)	-9.23 -7.52	5.52 3.54	1 3	17 10	3.41
gen5_16 (34)	-9.53 -8.91	2.08 9.78	1 8	4 13	7.26
gen5_17 (34)	-9.69	2.71	1	16	7.85
gen5_18 (34)	-9.37 -8.69	9.95 1.57	1 5	18 1	7.66
gen5_19 (38)	-8.38 -6.84	4.71 1.71	1 26	16 3	5.95
gen6_1 (34)	-8.32 -8.23 -5.98	6.39 8.01 1.69	1 5 28	3 14 2	3.95

Table A.5. Autodock results of ATP_LID_Flex. (cont.)

Conformers (no of clusters)	**Binding energy (kcal/mol)**	**Ligand RMSD (Å)**	**Cluster**	**Poses**	**Overall-backbone RMSD (Å)**
gen6_2 (40)	-8.16 -7.39	10.95 1.84	1 4	4 10	4.62
gen6_3 (37)	-8.15 -5.39	5.81 3.07	1 14	19 3	6.96
gen6_4 (41)	-8.42 -7.47 -6.24	7.39 6.57 2.27	1 3 18	9 15 1	7.07
gen6_5 (31)	-8.17 -7.36	4.37 6.12	1 7	8 19	4.91
gen6_6 (26)	-8.52 -6.63	6.22 4.17	1 11	20 2	5.93
gen6_7 (28)	-8.68	6.70	1	28	4.43
gen6_8 (43)	-9.68 -8.64 -4.73	7.38 8.77 5.50	1 2 36	13 12 1	5.79
gen6_9 (46)	-8.67 -8.38 -5.46	6.69 8.51 2.62	1 4 40	2 11 1	6.34
gen6_10 (35)	-8.87 -8.24	4.95 8.13	1 6	6 8	3.69
gen6_11 (38)	-8.50 -7.33 -4.68	5.33 5.09 3.58	1 7 38	5 11 1	3.72
gen6_12 (24)	-8.70	1.72	1	30	2.80
gen6_13 (40)	-9.05 -8.73	3.60 2.38	1 3	1 5	7.72
gen6_14 (35)	-8.33 -5.64	5.93 1.67	1 17	12 4	7.79
gen6_15 (28)	-8.24	1.91	1	18	2.72

Table A.5. Autodock results of ATP_LID_Flex. (cont.)

Conformers (no of clusters)	**Binding energy (kcal/mol)**	**Ligand RMSD (Å)**	**Cluster**	**Poses**	**Overall-backbone RMSD (Å)**
gen6_16 (33)	-10.80 -9.96	2.04 10.10	1 2	6 27	6.39
gen6_17 (42)	-8.78 -8.01 -6.85	10.74 5.18 2.20	1 4 20	6 13 1	3.71
gen6_18 (25)	-9.60 -9.49	1.65 2.78	1 2	17 19	4.89
gen6_19 (31)	-8.51 -6.50	11.35 1.70	1 3	3 8	6.31
gen6_20 (34)	-8.65 -6.83	5.44 3.06	1 6	19 7	5.21
gen6_21 (32)	-8.87 -7.85 -6.90	11.36 3.67 2.11	1 6 17	4 18 2	9.08
gen6_22 (28)	-9.60	1.81	1	33	3.59
gen6_23 (29)	-8.53 -8.31 -7.85	6.60 7.35 4.78	1 3 4	9 30 3	4.70
gen7_1 (37)	-9.18 -6.62	2.48 1.68	1 7	20 3	7.05
gen7_2 (37)	-8.67 -7.80	3.93 1.61	1 7	18 8	5.04
gen7_3 (35)	-9.17 -7.47	10.20 3.63	1 9	17 8	4.81
gen7_4 (38)	-9.66 -8.81	10.56 2.81	1 2	14 1	7.14
gen7_5 (38)	-9.46 -6.75	5.41 3.44	1 4	19 6	4.09
gen7_6 (40)	-8.09 -8.02 -6.12	8.60 9.36 2.42	1 2 13	2 14 5	6.69

Table A.5. Autodock results of ATP_LID_Flex. (cont.)

Conformers (no of clusters)	**Binding energy (kcal/mol)**	**Ligand RMSD (Å)**	**Cluster**	**Poses**	**Overall-backbone RMSD (Å)**
gen7_7 (37)	-7.29 -6.46	5.97 2.11	1 7	20 1	7.37
gen7_8 (35)	-9.01 -6.39	6.39 2.15	1 23	15 5	6.27
gen7_9 (32)	-8.51 -6.69	5.53 2.20	1 7	3 1	5.38
gen7_10 (38)	-7.84 -7.38 -6.47	6.46 8.19 1.93	1 5 16	10 14 1	6.44
gen7_11 (46)	-8.48 -7.86 -6.89	5.14 6.50 2.50	1 4 25	5 6 1	4.82
gen7_12 (34)	-9.40 -7.66	9.03 2.83	1 9	11	4.91
gen7_13 (37)	-8.27 -6.77	7.26 4.46	1 10	22 4	7.13
gen7_14 (27)	-8.74 -8.01	7.63 7.18	1 3	5 28	5.78
gen7_15 (30)	-9.10 -8.16	6.83 6.41	1 2	21 23	5.04
gen7_16 (28)	-8.24	5.98	1	26	4.01
gen7_17 (38)	-9.02	7.15	1	21	4.50
gen7_18 (37)	-8.40 -7.26 -4.72	8.10 8.86 6.05	1 4 21	7 13 1	5.06
gen7_19 (45)	-8.50 -5.37	8.83 12.17	1 26	6 9	5.55
gen7_20 (38)	-9.22 -9.01	3.61 7.54	1 3	1 15	7.68

Table A.5. Autodock results of ATP_LID_Flex. (cont.)

Confor-mers (no of clusters)	**Binding energy (kcal/mol)**	**Ligand RMSD (Å)**	**Cluster**	**Poses**	**Overall-backbone RMSD (Å)**
gen7_21 (28)	-8.49 -7.37	7.99 1.86	1 6	28 1	3.59
gen7_22 (24)	-7.98 -6.55 -5.38	9.67 5.74 2.26	1 6 21	5 22 1	7.40
gen7_23 (46)	-9.79	2.50	1	17	5.06
gen7_24 (41)	-9.03 -7.59 -6.20	12.92 6.09 1.77	1 2 19	1 18 1	8.62
gen7_25 (29)	-9.01 -8.12	8.68 1.68	1 8	23 4	3.83
gen7_26 (31)	-8.48	4.22	1	11	4.07
gen7_27 (35)	-8.60	7.13	1	4	3.42
gen7_28 (29)	-8.67	5.82	1	7	3.92
gen7_29 (35)	-8.65 -7.27	9.52 2.24	1 8	4 12	2.52
gen7_30 (41)	-7.82 -6.47	2.36 1.78	1 13	11 3	5.94
gen7_31 (33)	-8.42 -7.02	2.38 1.72(min.)	1 1	22 22	3.95

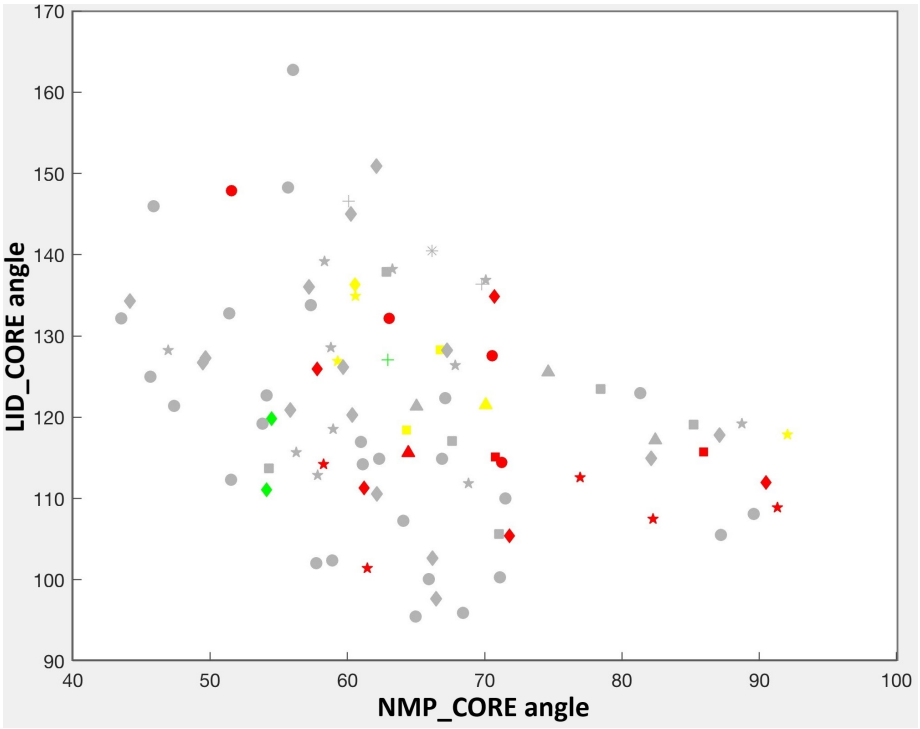


Figure A.1. LID-CORE and NMP-CORE angles of ATP_LID.

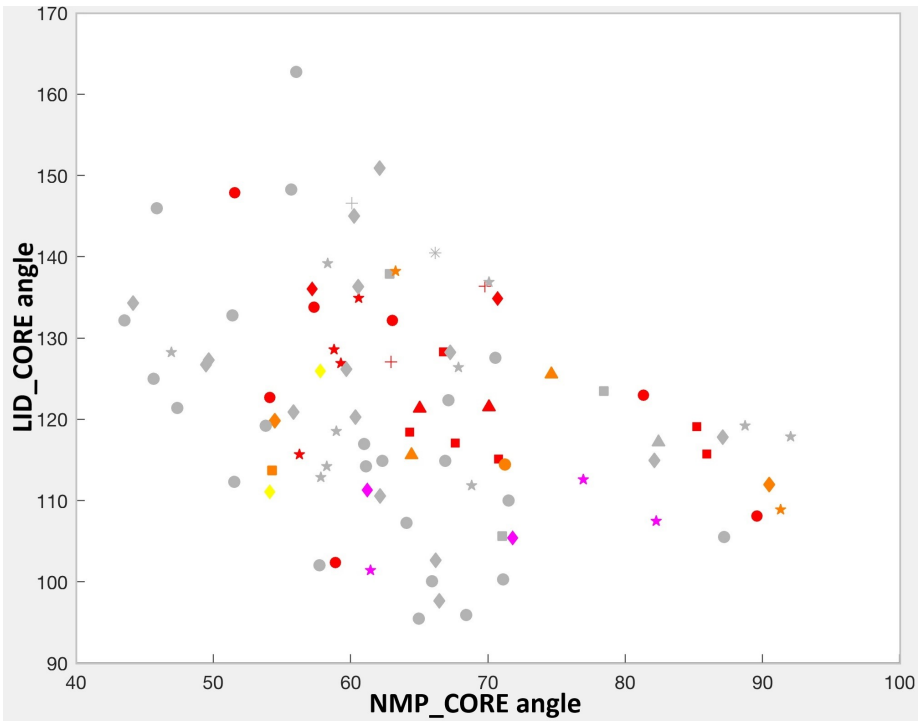


Figure A.2. LID-CORE and NMP-CORE angles of ATP_LID_Flex.