Comparative Study of Covalent & Non-covalent Drug Inhibitory Investigation: Targeting HSP72 Protein in Cancer Therapy Using Molecular Modelling Techniques

Mr Aimen Khalefa Aljoundi

217075071

2021

A thesis submitted to the School of Health Sciences, University of Kwa-Zulu Natal, Westville, in fulfilment of the degree of Doctor of Philosophy

Comparative Study of Covalent & Non-covalent Drug Inhibitory Mechanism Investigation: Targeting HSP72 Protein in Cancer Therapy Using Molecular Modelling Techniques

Mr Aimen Khalefa Aljoundi

217075071

2021

This is the thesis in which the chapters are written as a set of discrete research publications, with an overall introduction and final summary. Typically, these chapters will have been published in internationally recognized, peer-reviewed journals.

This is to certify that the contents of this thesis are the original research work of Mr Aimen Khalefa Aljoundi.

As the candidate's supervisor, I have approved this thesis for submission.

Supervisor:

Signed: [redacted] Name: Prof Mahmoud Soliman

Date: July 2, 2021

PREFACE

This thesis is divided into six chapters, including the following:

Chapter 1:

This chapter introduces the background of the study, its rationale, relevance and the aims and objectives. This chapter also includes the summarized outline of this thesis.

Chapter 2:

This chapter provides a comprehensive review on the molecular protein target that was explored in this thesis and its involvement in the protein degradation process. Also, currently existing strategic therapeutic interventions in relation to this target were discussed coupled with their respective advantages and limitations. The historical background, function and structural characteristics of these approaches were appropriately elucidated.

Chapter 3:

This chapter explains computer aided drug design by discussing a various molecular dynamic approaches and applications. Moreover, the relevant application of these diverse techniques was appropriately discussed, and most especially with regards to the combinatorial *in silico* methods succinctly applied in the studies that constitute this thesis.

Chapter 4: **(Published work – this chapter is presented in accordance to the required format of the journal in which it was accepted for publication, hence it is the final version of the accepted manuscript).**

In this chapter, we highlight and provide a justification from the computational point of view and computational techniques suitable for covalent and non-covalent bad and good advantage. We also discuss the workability of each method, lapses and shortcomings.

This review article has already been published in the journal of *Protein Journal* (IF = 1.317) in 2020.

Chapter 5: (Published work – this chapter is presented in accordance to the required format of the journal in which it was accepted for publication, hence it is the final version of the accepted manuscript).

This chapter covers the second objective of this thesis, including a comparative study of the covalent versus covalent inhibition of the nucleotide binding domain of HSP72 protein. Upon the binding of these two covalent bond inhibitions via lysine56 and cysteine17, the optimal underlying mechanism was evaluated based on the structural and conformational molecular dynamic variations of the enzyme.

This research article has also been published in the journal of *Molecules* (IF = 3.267) in 2020.

Chapter 6: (Published work – this chapter is presented in accordance to the required format of the journal in which it was accepted for publication, hence it is the final version of the accepted manuscript).

This chapter demonstrates the two different binding modes of the HSP72 protein. evaluates the third objective of the thesis: to compare the inhibition modes of HSP72 protein when bound covalently and non-covalently at the ATP binding domain. In addition, a detailed molecular understanding of ATP nucleotide binding domain structural mechanisms of inhibition was provided.

This research article has been published in the *computer in biology and medicine* (IF = 3.434) in 2021.

Chapter 7: (Published work- this chapter is presented in the required format of the journal and is the final version of the submitted manuscript).

This chapter entitled “Comparison of irreversible inhibition targeting HSP72 protein - the resurgence of covalent drug developments”, describes the preferential binding of the third-generation covalent inhibitor to the ATP binding domain.

This research article has been published in the *Molecular Simulation* Journal (IF = 1.73).

ABSTRACT

Cancer is the most complicated and diverse disease that has been menacing human beings worldwide. Up to date, important advancement has been done to improve the existing therapeutic interventions in the treatment and management of cancer. However, the side effect of these drugs that are mostly associated with the "off-target" effects is a perpetual failure in cancer drug development. Therefore, efficient regimen with minimal toxicities and high drug target selectivity should be achieved. Covalent inhibition is an emerging field in drug discovery and a very distinct category of therapeutics that reduces adverse side effects and possible interactions that lead to drug resistance due to its attainable reactivity and high selectivity.

The Heat shock proteins (HSPs) play a crucial role in the clearance of damaged proteins by encouraging proteotoxicity and proteins acclamation. This process occurs by avoiding unsuitable stress-induced protein aggregation, ensure suitable refolding of denatured proteins, and promoting their degradation; thus, the involvement of this enzyme in many human diseases, including cancer. In this study, we delve into the structural features of one of the most crucial enzymatic targets of the stress proteins, the Heat shock proteins72. In drug development, the integration of computational techniques including molecular dynamic simulations, docking and molecular modelling has allowed drug developers to screen and syntheses millions of compounds and thus screen out possible lead drugs. Computer-Aided Drug Design has been validated as a cost-effective strategy to fast trace the drug discovery process due to these in silico methods.

One of the characteristics of the HSP72 is its ability to be targeted either covalently or non-covalently through small drug molecules. Therefore, the above-mentioned methods, amongst several other computational tools were employed out in this study to provide insights into conformational changes that explain potential covalent and non-covalent inhibitory mechanisms, binding sites assessment features leading to promising small molecule inhibitor candidates.

These combinatorial computational studies offer an inclusive *in silico* perspective to fill the gap in drug design studies about targeting protein degradation, thus providing insights toward the structural characteristics of the pivotal target and describing promising drug developments.

OKUFINQIWE

Umdlavuza isifo esiyinkimbinkimbi futhi esinhlobonhlobo kunazo zonke izifo ezihlupha abantu emhlabeni jikelele. Ukuthuthuka kwezobuchwepheshe okusesikhathini futhi okubalulekile ikhona osekudlale iqhaza elinzulu ekuphucukiseni ukungenelela kokwelapha [izindlela zokwelapha] okusetshenziselwa ekwalapheni kanye nokulawula umdlavuza. Kepha, imiphumela eqhamuka ngokungehloso uma udla imishanguzo ehlobene kakhulukazi nemiphumela "engashayi emhlolweni [engaqondiwe]" ikhomba ukwehluleka okuqhubekayo mayelana nokuthuthukiswa kwemishanguzo yomdlavuza. Ngako-ke, izinhlelo eziphumelelayo ezinobuthi okuyingcosana kanye nemishanguzo ekwazi ukukhetha okuphokopholelwe izona izinto okufanele zizuzwe. Ukuvinjelwa okuvamile (i-*covalent inhibition*) kungumkhakha osafufusa ocwaningweni lwemishanguzo futhi kunguhlobo lokwelapha oluqavile olukwazi ukunciphisa imiphumela emibi [eqhamuka uma udla imishanguzo] kanye nokwenzelana [ngenxa kwe-*attainable reactivity* kanye nokubandlulula okuphezulu (*high selectivity*)] okungenzeka okuholela ekungasebenzeni kwemishanguzo.

Ama-*heat shock proteins* (HSPs) adlala indima ebalulekile ekususeni amaprotheni alimeleyo ngoku gqugquzela amaprotheni ukuthi ajwayele ushintsho oselukhona emzimbeni kanye nemiphumela yokulimala ngokweqile noma ukungasongeki ngendlela efanele kwamaprotheni. Le nqubo yenzakala ngokunqanda ukuhlangana kwamaprotheni angafanelekile adalwa ukucindezela kwengcindezi, iphinde futhi iqinisekise ukusongana ngendlela efanekileyo yamaprotheni abesengasongekile ngokwendalo, futhi iyaku khuthaza ukuncipha kwamaprotheni: kanjalo, nokubandakanyeka kwale-*enzyme* ezifweni eziningi eziphatha abantu, okubalwa kuzo umdlavuza. Kulolu cwaningo, sihlaziya izici zezakhiwo zama-*enzyme* ahlosiwe amaprotheni engcindezi abaluleke kunawo wonke, lawa ama-*heat shock proteins72* (HSP72). Ekuthuthukisweni kwemishanguzo, ukusetshenziswa kobuchwepheshe okubalwa kukho i-*molecular dynamic simulation*, i-*docking*, kanye ne-*moleculur modelling* sekuvumele osokhemisi ukuthi bakhone ukuhlanganisa baphinde bahlole izigidi zezinhlanganisela ukuze bekwazi ukukhipha imishanguzo okungenzeka ibe usizo. Lolu cwaningo olwenziwa ngamakhompyutha (ama-*in silico methods*) seluqinisekisile i-*Computer-Aided Drug Design* njengendlela [yokucwaninga] ebiza kahle ezo phuthumisa inqubo yokutholakala kwemishanguzo.

Esinye sezici zama-HSP72 ikhono lawo lokuthi kwaziwe ukufinyelelwa kuwo *covalently* okanye *non-covalently* ngokusebenzisa amagqamuzana (*molecules*) amancane emishanguzo.

Ngaleyo ndlela, izindlela zocwaningo ezishiwo ngenhla, kanye nobunye ubucwepheshe obusebenzisa amakhompyutha, ikhona okusetshenzisiwe kulolu cwaningo ukuze umcwaningi ekwazi ukunikeza ukuqonda ngoshintsho olwenzeka esakhiweni sama-*macromolecule* (*conformational change*) okuyilona okungenzeka likwazi ukucacisa ngezindlela zokuvimbela ezi-*covalent* okanye *non-covalent* [izici zokuvivinya zezindawo zokuhlanganisa ezingaholela ezithikamizweni ezithembisayo ezincanyana zamagqamuzana (*molecule*) aqhudelanayo].

Lezi zinhlobo zocwaningo ezisebenzisa izindlela zokubala ezibizwa ngokuthi i-*combinatorial* zikhona ukunikeza umbono ophelele ngocwaningo olwenziwa ngamakhompyutha (ama-*in silico methods*) ukuze kwazeke ukuvala isikhala esikhona ngolwazi oluphathelene nocwaningo lokuklama imishanguzo emelene nokunqoba ukukhahlazeka kwamaprotheni, ngaleyo ndlela kwazeka ukuthi kutholakale imibono ngesakhiwo sezici eziqondiwe ezinohlonze kanye nokuchaza ngezindlela ezithembisayo zokuthuthukiswa kwemishanguzo.

DECLARATION I – PLAGIARISM

I, Aimen K Aljoundi, declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.

2. This thesis has not been submitted for any degree or examination at any other university.

3. This thesis does not contain other person's data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This thesis does not contain other person's writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

> a. their words have been re-written, but the general information attributed to them has been referenced.
>
> b. where their exact words have been used, then their writing has been placed in italics and inside quotation marks and referenced.

5. This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being detailed in the thesis and in the references sections.

A detailed contribution to publications that form part or/and include research presented in this thesis is stated (include publications submitted, accepted, in press and published).

Signed: **A.Aljoundi**

DECLARATION II – LIST OF PUBLICATIONS

1 **Aimen Aljoundi**, Imane Bjij, Ahmed El Rashedy, Mahmoud E.S. Soliman (2020). Covalent Versus Non-covalent Enzyme Inhibition: Which Route Should We Take? A Justification of the Good and Bad from Molecular Modelling Perspective. *The Protein Journal*. DOI:10.1007/s10930-020-09884-2

Contribution:

Aimen Aljoundi: carried out the conceptualization, sourcing of relevant literatures/data, writing and editing of the manuscript.
Imane Bjij: contributed to the project by assisting in data collation and manuscript writing.
Ahmed El Rashedy: contributed to the project by assisting in data collation.
Mahmoud E.S. Soliman: Supervisor

2 **Aimen Aljoundi**, Ahmed El Rashedy, Patrick Appiah-Kubi, Mahmoud E.S. Soliman (2020). Coupling of HSP72 α-helix Subdomains by the Unexpected Irreversible Targeting of Lysine-56 over Cysteine-17; Coevolution of Covalent Bonding. *Molecules Journal* DOI: 10.3390/molecules25184239

Contribution:

Aimen Aljoundi: carried out the conceptualization, sourcing of relevant literatures/data, writing and editing of the manuscript.
Ahmed El Rashedy: contributed to the project by assisting in data collation.
Patrick Appiah-Kubi: manuscript editing.
Mahmoud E.S. Soliman: Supervisor.

3. **Aimen Aljoundi**, Ahmed El Rashedy, Mahmoud E.S. Soliman (2021) Distinguishing the optimal binding mechanism through reversible and irreversible inhibition analysis of HSP72 protein in cancer therapy, *Computer in biology and medicine*. DOI: 10.1016/j.compbiomed.2021.104301

Contribution:

Aimen Aljoundi: carried out the conceptualization, sourcing of relevant literatures/data, writing and editing of the manuscript.

Ahmed El Rashedy: contributed to the project by assisting in data collation and manuscript editing.
Mahmoud E.S. Soliman: Supervisor.

4. **Aimen Aljoundi**, Ahmed El Rashedy, Mahmoud E.S. Soliman (2021) Comparison of irreversible inhibition targeting HSP72 protein - the resurgence of covalent drug developments, *Journal of Molecular Simulation*. DOI: 10.1080/08927022.2021.1949457

Contribution:

Aimen Aljoundi: carried out the conceptualization, sourcing of relevant literatures/data, writing and editing of the manuscript.
Ahmed El Rashedy: contributed to the project by assisting in data collation and manuscript editing.
Mahmoud E.S. Soliman: Supervisor.

RESEARCH OUTPUT

A- List of publications

1. **Aimen Aljoundi**, Imane Bjij, Ahmed El Rashedy, Mahmoud E.S. Soliman (2020). Covalent Versus Non-covalent Enzyme Inhibition: Which Route Should We Take? A Justification of the Good and Bad from Molecular Modelling Perspective. *The Protein Journal*. DOI:10.1007/s10930-020-09884-2

2. **Aimen Aljoundi**, Ahmed El Rashedy, Patrick Appiah-Kubi, Mahmoud E.S. Soliman (2020). Coupling of HSP72 α-helix Subdomains by the Unexpected Irreversible Targeting of Lysine-56 over Cysteine-17; Coevolution of Covalent Bonding. *Molecules Journal* DOI: 10.3390/molecules25184239

3 **Aimen Aljoundi**, Ahmed El Rashedy, Mahmoud E.S. Soliman (2021) Distinguishing the optimal binding mechanism through reversible and irreversible inhibition analysis of HSP72 protein in cancer therapy, *Computer in biology and medicine*. DOI: 10.1016/j.compbiomed.2021.104301

4. **Aimen Aljoundi**, Ahmed El Rashedy, Mahmoud E.S. Soliman (2021) Comparison of irreversible inhibition targeting HSP72 protein - the resurgence of covalent drug developments. *Journal of Molecular Simulation*. DOI: 10.1080/08927022.2021.1949457

DEDICATION

In the name of Allah, the Most Gracious and the Most Merciful Alhamdulillah, all praises to Allah for the strengths and His blessings in completing this thesis.

This study is wholeheartedly dedicated to my beloved parents, Dr Khalefa Aljoundi and Mrs Salima Mansour, who have been my source of inspiration and gave me strength when I thought of giving up, who continually provide their moral, spiritual and emotional support.

To my beloved siblings especially my sister Entisar Aljoundi and her husband my brother-in-law Dr. Omar Moumin for their advice and encouragement to finish this study and believing in my wild dreams, aspirations and ambitions, thank you.

ACKNOWLEDGEMENTS

In the name of Allah, the Most Gracious and the Most Merciful Alhamdulillah, all praises to Allah for the strengths and His blessings in completing this thesis.

Foremost, I would like to express my sincere gratitude to my supervisor Prof. Mahmoud Soliman for the continuous support and advice of my PhD study and research, for his motivation, patience, immense knowledge and enthusiasm. His guidance helped me a lot in all the hard times of research and writing of this thesis. I could not have imagined a better supervisor, advisor and mentor for my PhD study.

A big appreciation goes to all my colleagues at the Molecular Bio-computation and Drug Design Laboratory for their all-round support.

To all my friends, thank you for your understanding and encouragement in my many, many moments of crisis. Your friendship makes my life a wonderful experience. I cannot list all the names here, but you are always on my mind.

Finally, I would like to acknowledge University of Kwa-Zulu Natal, Westville Campus, Durban, South Africa, for the financial support throughout the course of my study and CHPC for their resources and technical support.

LIST OF ABBREVIATIONS

Å	Angstrom
ATP	Adenosine Triphosphate
ADP	Adenosine di-phosphate
CADD	Computer-Aided Drug Design
CACDD	Computer-Aided Covalent Drug Design
CMDs	Covalent Molecular Dynamics Simulation
DCCM	Dynamic Cross Correlation Matrix
DSG	Deoxyspergualin
FDA	Food and Drug Administration
FP	Fluorescence Polarization
GAFF	General amber Force Field
HSPs	Heat Shock Proteins
HSPA1A	Heat Shock Protein 70kDa protien1A
HSPA1B	Heat shock Protein 70kDa protein 1B
HSP70	Heat shock Protein 70
HSP72	Heat shock Protein 72
kDA	kilo Dalton
MMV	Molegro Molecular Viewer
MC	Monte Carlo
MD	Molecular Dynamics

MM	Molecular Mechanics
MM/GBSA	Molecular Mechanics/Generalized Born Surface Area
MM/PBSA	Molecular Mechanics/Poisson-Boltzmann Surface Area
NS	Nano Second
NBD	Nucleotide-binding domain
NMR	Nuclear Magnetic Resonance
PME	Particle mesh Ewald
PES	Potential energy surface
PCA	Principal component analysis
PDB	Protein data bank
PME	Particle-mesh Ewald method
QM	Quantum Mechanics
QM/MM	Quantum Mechanics/Molecular Mechanics
RDD	Rational Drug Design
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
RoG	Radius of Gyration
SBD	Substrate-binding domain
SASA	Surface Accessible Surface Area
vDW	van der Waals
VS	Virtual screening

WHO	World Health Organization
α	Alpha
β	Beta
ΔG	Free Binding energy
3D	Three-Dimensional

LIST OF AMINO ACIDS

Three Letter Code	Amino Acid
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic Acid
Cys	Cysteine
Gln	Glutamine
Glu	Glutamic Acid
Gly	Glycine
His	Histidine
Ile	Isoleucine
Leu	Leucine
Lys	Lysine
Met	Methionine
Phe	Phenylalanine
Pro	Proline
Ser	Serine
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

CHAPTER 1

1. Introduction

1.1 Background and rationale of the study

Cancer has been threatening people worldwide since ancient ages, being one of the non-communicable diseases that are responsible for the larger part of deaths across the world (Yin *et al.*, 2019). According to the World Health Organization (WHO), cancer is the ultimate cause of death before the age of 70 years in most countries globally (WHO, 2018). Cancer is the main cause of death and the major obstacle for improving life expectancy around the world in the 21st century (Bray *et al.*, 2018; Siegel, Miller and Jemal, 2019).

The growth and division of cells that result from the uncontrolled cellular changes in the human body lead to the development of cancer. Different forms of cancer cause cells to grow and divide hastily, while others cause cells to grow and divide a slower rate (Baskar *et al.*, 2014). There are several types of cancer; however, all the cancerous cells grow continuously, divide and leads to the formation of new abnormal cells (Boothman *et al.*, 2019). Few classes of cancerous cells move from one part of the body to others through the circulation of blood, where they start growing. Tumors are cancerous and non-cancerous (benign). Benign tumors are not life-threatening and does not grow (Sudhakar, 2009).

Besides lifestyle choices, such as smoking, excess body weight or poor nutrition, genetic factors can contribute to the disease's development. Several elements affect the ways that DNA communicates with cells and directs their division and death (Bouvard *et al.*, 2009; Cogliano *et al.*, 2011; Hecht, 1999). Conventional approaches for cancer treatment compromise a combination of chemotherapy, radiation therapy, and surgery. Some patients benefit from personalized cancer treatment that is usually based on the tumour genetic profile (Block *et al.*, 2015). However, the drug cannot always characterize the proteins involved in cancer, and sometimes it affects healthy tumour cells; thus, finding selective inhibitors is a pivotal objective in drug discovery (Visscher, Michelle R Arkin and Dansen, 2016).

Drug selectivity is a key concept to achieve the proper drug effect and overcome interactions with undesirable targets that can be essential for biological processes in the human body

(Huggins, Sherman and Tidor, 2012). In order to increase the potency and selectivity of therapeutics, scientists have extended their hunt for different strategies that can contribute to the path through drug development. The covalent inhibition of pathogenic proteins is one of the reemerging strategies holding the attention of chemical biologist, because of their proven reactivity and selectivity to the biological receptors (Johnson, Weerapana and Cravatt, 2010; Liu *et al.*, 2013). Covalent drugs have contributed to the curing of different diseases and pharmacological problems that could not be solved by the non-covalent inhibitors, due to their peculiar mechanism of action, allowing a prolonged linkage with the target (Zhao and Bourne, 2018). Altogether, the desirable features of the covalent inhibitors explain the importance of covalent drugs in the pharmaceutical industry, as this class of therapeutics occupies, approximately, 30% of all drugs on the market, with anti-cancer drugs accounting for 21%. (Singh *et al.*, 2011).

Hsp70 is the main stress-inducing member of the chaperone family and is coded by two genes, HSPA1A and HSPA1B, producing structure isoforms of HSP72 (Kampinga *et al.*, 2009). HSP72 is symmetrical to the 78 kDa glucose-regulated protein of 377 residues of amino acid that form in the nucleotide-binding domain (NBD), which play a senior role in coordinating unfolding protein response, HSP72 is identified as high level effective protein in malignant tumors of many origins and improve cancer cell survival (Nylandsted *et al.*, 2000, 2002). Thus, HSP72 is demonstrated as a promising anti-cancer target due to its key role in expressing at high levels in malignant tumors of numerous origins (Jaattela, 1999). Hsp70 forms a transient chaperone interaction with substrate proteins through its C-terminal substrate-binding domain(SBD), according to their functional mechanism and molecular structure (Mayer and Bukau, 2005). Controlling the transient chaperone interaction when the nucleotide binds allosterically to the N-terminal nucleotide-binding domain (NBD). The affinity of the substrate-binding domain (SBD) for substrates decreases by 10 to 400 folds when ATP is binding to the N-terminal of the nucleotide-binding domain (NBD). The inhibition of the nucleotide-binding domain (NBD) is identified as one of the most promising strategies for HSP72 function inhibition (Zhuravleva, Clerico and Gierasch, 2012), which can be covalently or non-covalently targeted by small molecules inhibitors, will undergo various studies reported herein.

The development of drug discovery is expensive and challenging, starting from basic scientific research to FDA approval passing through preclinical and clinical trials (Mohs and Greig,

2017). Computational methods and resources may be implemented in most stages of drug discovery from identifying targets, to drug optimization and preclinical testing. With such a blighted disease, like cancer, computational techniques including molecular modeling and molecular dynamic simulations, docking, identification of pharmacophoric hot spots and virtual screening allow drug developers to synthesis millions of compounds to funnel out possible lead drugs which may then be validated experimentally. This strategy overcomes the concept of "walking in the dark" with experimental screening, thus reducing the drug discovery timeline (Sliwoski *et al.*, 2014; Ou-Yang *et al.*, 2012).

The necessity of computational methods in drug discovery and development activity is hastily expanding in terms of growth, application and popularity. Computer-aided covalent drug design is one of the most growing field in the area of drug discovery (Baillie, 2016; Lonsdale and Ward, 2018; Kapetanovic, 2008). Molecular modeling techniques are widely used in drug design and development and can play a key role in selecting compounds for synthesis and also offering atomic level justifications which can be used to direct further design efforts. Various modeling techniques used in drug design were only evolved for the application on non-covalent inhibitors. Most of the techniques can also be functional to covalent drug design; however, certain concerns or improved software packages or tool are mandatory sometimes, to preserve the covalent bond between the inhibitor and target (De Cesco *et al.*, 2017; Lonsdale and Ward, 2018). The detailed molecular techniques useful for covalent drug design have been introduced in chapter three of this thesis.

In this study, due to the lack of fundamental research about the rational drug designing of covalent inhibitors of the HSP72, we have utilized key computational techniques to fill the gap in the literature regarding the inhibition of this biological target, thus giving insights toward oncogenic drug targets and designing potential inhibitors against cancer.

1.2 Aims and objective of this study

The main purpose of the current study is to provide structural and molecular insights on crucial target in protein degradation process, heat shock protein 72 (HSP72), as well as fill the gap in literature regarding covalent inhibition of biological targets in drug discovery by implementing computational techniques. Coupled *in silico* applications of strategic drug design methods were utilized to uncover the differences between the covalent inhibition and the non-covalent

inhibition of HSP72 and optimize the therapeutic activities of compounds and improving their pharmacokinetics in relation to their use against this target protein.

To accomplish this, specific sub-aims were outlined including the following:

1. To explore the molecular insights regarding the development and modifications of computational techniques suitable for covalent systems, their lapses, shortcomings and recent advancements. To achieve this, we;
    - Providing insights into the covalent drugs and the advantages and limitations that can be related to this class of therapeutics in drug discovery.
    - Listing the different tools and techniques that can be used by a computational scientist to study the covalently targeted systems.
    - Bringing to light an expert opinion about the development of the covalent inhibitors and the progress associated with these drugs.

2. To establish the most optimal binding mechanism for the inhibition of the HSP72 and highlight the structural dynamics upon binding of two different inhibitors (covalent and non-covalent) in the catalytic site by:
    - Performing molecular dynamic simulations to compare the structural adjustment of covalently and non-covalently inhibited HSP72, with respect to its APO form.
    - Utilizing a wide variety of post-molecular dynamic analysis techniques to characterize the binding landscape of the HSP72 enzyme and assess its stability and to demonstrate any structural alterations associated with the binding inhibitor.

3. To investigate the preferential binding mode of a covalent inhibitor towards the HSP72 and determine the implicated structural and molecular dynamics. In other words, we highlight the binding mechanism of one covalent inhibitor and one non-covalent inhibitor into two different binding sites and we establish the impact of the binding mechanism on the selectivity of the inhibitors. This objective was achieved by:

- Applying the covalent molecular dynamic approach to study two covalent systems: when ligand is bound to the Lysine56 amino acid and when the same ligand is bound to Cystine17 amino acid.
- Utilizing different post-molecular dynamic analysis techniques to characterize the binding landscape of the enzyme in the two bound systems and to clarify previous experimental results on the lysine56 covalent bond as the favored binding of the covalent inhibitor.

4. To combine various *in silico* techniques to virtually screen and propose potential covalent inhibitors of the HSP72. This may be achieved by:

    - Appropriate ligand and protein target are selected, docked covalently or non-covalently using respective docking tools.
    - This was followed by parameterization of the ligand, protein and covalent complex respectively.
    - Combining molecular docking with covalent molecular dynamic simulations of 250 ns for unbound enzyme, enzyme-inhibitor complex.
    - Validating the most favourable ligands by assessing the stability and conformation structure changing of each system following exposure to molecular dynamic simulations.

1.3 Novelty and significance of the work

Extraordinary efforts have been giving in the past decades toward a novel class of drugs known as 'covalent drugs', where structure-based strategies are involved in producing small molecule inhibitors that inhibit their targeted protein through a covalent bond with a particular residue. This strategy has been proved promising especially in cancer therapy to overcome the challenges associated with selectivity, potency and efficacy of drugs (Singh *et al.*, 2011; Visscher, Michelle R. Arkin and Dansen, 2016; De Cesco *et al.*, 2017). As opposed to the rapidly growing areas of drug design, computational studies on covalent inhibition are very limited, for instance, there have been insufficient reported computational studies which comparatively account the different binding inhibition modes of drug action related to HSP72. Therefore, we chose to provide insights into structural dynamics and binding landscapes upon covalent and non-covalent drug inhibition HSP72. With the use of computer aided drug

development approach, a comprehensive *in silico* perspective is offered to shed light on potential structural features that allow for the inhibition of this protein as well as amino acids implicated in enzyme activity. Defining the binding landscape will offer prospective design of selective and unique inhibitors with critical pharmacophoric features that will aid in developing targeted and effective small molecule inhibitors.

To this end, the work conferred in this thesis is considered to be a major platform in the development of research to design more highly selective potent covalent inhibitors for Heat shock protein (HSP72) with the potential to overcome challenges related to selectivity in cancer therapy.

References

Baillie, T. A. (2016) ‘Targeted Covalent Inhibitors for Drug Design’, *Angewandte Chemie - International Edition*, pp. 13408–13421. doi: 10.1002/anie.201601091.

Baskar, R. *et al.* (2014) ‘Biological response of cancer cells to radiation treatment.’, *Frontiers in molecular biosciences*. Frontiers Media SA, 1, p. 24. doi: 10.3389/fmolb.2014.00024.

Block, K. I. *et al.* (2015) ‘Designing a broad-spectrum integrative approach for cancer prevention and treatment’, *Seminars in Cancer Biology*, 35, pp. S276–S304. doi: 10.1016/j.semcancer.2015.09.007.

Boothman, A.-M. *et al.* (2019) ‘Impact of Patient Characteristics, Prior Therapy, and Sample Type on Tumor Cell PD-L1 Expression in Patients with Advanced NCSLC Screened for the ATLANTIC Study.’, *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. doi: 10.1016/j.jtho.2019.04.025.

Bouvard, V. *et al.* (2009) ‘A review of human carcinogens--Part B: biological agents.’, *The Lancet. Oncology*, 10(4), pp. 321–2.

Bray, F. *et al.* (2018) ‘Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.’, *CA: a cancer journal for clinicians*, 68(6), pp. 394–424. doi: 10.3322/caac.21492.

De Cesco, S. *et al.* (2017) ‘Covalent inhibitors design and discovery’, *European Journal of Medicinal Chemistry*. Elsevier Masson SAS, 138, pp. 96–114. doi: 10.1016/j.ejmech.2017.06.019.

Cogliano, V. J. *et al.* (2011) ‘Preventable Exposures Associated With Human Cancers’, *JNCI Journal of the National Cancer Institute*, 103(24), pp. 1827–1839. doi: 10.1093/jnci/djr483.

Hecht, S. S. (1999) ‘Tobacco smoke carcinogens and lung cancer.’, *Journal of the National Cancer Institute*, 91(14), pp. 1194–210.

Huggins, D. J., Sherman, W. and Tidor, B. (2012) ‘Rational Approaches to Improving Selectivity in Drug Design’, *Journal of Medicinal Chemistry*, 55(4), pp. 1424–1444. doi: 10.1021/jm2010332.

Jaattela, M. (1999) ‘Heat shock proteins as cellular lifeguards’, *Annals of Medicine*.

Johnson, D. S., Weerapana, E. and Cravatt, B. F. (2010) ‘Strategies for discovering and derisking covalent, irreversible enzyme inhibitors’, *Future Medicinal Chemistry*, 2(6), pp. 949–964. doi: 10.4155/fmc.10.21.

Kampinga, H. H. *et al.* (2009) ‘Guidelines for the nomenclature of the human heat shock proteins.’, *Cell stress & chaperones*, 14(1), pp. 105–11. doi: 10.1007/s12192-008-0068-7.

Kapetanovic, I. M. (2008) ‘Computer-aided drug discovery and development (CADDD): In silico-chemico-biological approach’, *Chemico-Biological Interactions*, 171(2), pp. 165–176. doi: 10.1016/j.cbi.2006.12.006.

Liu, Q. *et al.* (2013) ‘Developing irreversible inhibitors of the protein kinase cysteinome’, *Chemistry and Biology*. doi: 10.1016/j.chembiol.2012.12.006.

Lonsdale, R. and Ward, R. A. (2018) ‘Structure-based design of targeted covalent inhibitors’, *Chemical Society Reviews*. doi: 10.1039/C7CS00220C.

Mayer, M. P. and Bukau, B. (2005) 'Hsp70 chaperones: Cellular functions and molecular mechanism', *Cellular and Molecular Life Sciences*. Birkhäuser-Verlag, 62(6), pp. 670–684. doi: 10.1007/s00018-004-4464-6.

Mohs, R. C. and Greig, N. H. (2017) 'Drug discovery and development: Role of basic biological research', *Alzheimer's and Dementia: Translational Research and Clinical Interventions*. Elsevier Inc., 3(4), pp. 651–657. doi: 10.1016/j.trci.2017.10.005.

Nylandsted, J. *et al.* (2000) 'Selective depletion of heat shock protein 70 (Hsp70) activates a tumor-specific death program that is independent of caspases and bypasses Bcl-2', *Proceedings of the National Academy of Sciences*, 97(14), pp. 7871–7876. doi: 10.1073/pnas.97.14.7871.

Nylandsted, J. *et al.* (2002) 'Eradication of glioblastoma, and breast and colon carcinoma xenografts by Hsp70 depletion.', *Cancer research*, 62(24), pp. 7139–42.

Ou-Yang, S.-S. *et al.* (2012) 'Computational drug discovery.', *Acta pharmacologica Sinica*. Nature Publishing Group, 33(9), pp. 1131–40. doi: 10.1038/aps.2012.109.

Siegel, R. L., Miller, K. D. and Jemal, A. (2019) 'Cancer statistics, 2019.', *CA: a cancer journal for clinicians*, 69(1), pp. 7–34. doi: 10.3322/caac.21551.

Singh, J. *et al.* (2011) 'The resurgence of covalent drugs', *Nature Reviews Drug Discovery*. Nature Publishing Group, 10(4), pp. 307–317. doi: 10.1038/nrd3410.

Sliwoski, G. *et al.* (2014) 'Computational methods in drug discovery.', *Pharmacological reviews*. American Society for Pharmacology and Experimental Therapeutics, 66(1), pp. 334–95. doi: 10.1124/pr.112.007336.

Sudhakar, A. (2009) 'History of Cancer, Ancient and Modern Treatment Methods', *Journal of Cancer Science & Therapy*, 01(02), pp. i–iv. doi: 10.4172/1948-5956.100000e2.

Visscher, M., Arkin, Michelle R and Dansen, T. B. (2016) 'Covalent targeting of acquired cysteines in cancer', *Current Opinion in Chemical Biology*, 30, pp. 61–67. doi: 10.1016/j.cbpa.2015.11.004.

Visscher, M., Arkin, Michelle R. and Dansen, T. B. (2016) 'Covalent targeting of acquired cysteines in cancer', *Current Opinion in Chemical Biology*, 30, pp. 61–67. doi: 10.1016/j.cbpa.2015.11.004.

WHO (2018) *Global Health Observatory, Geneva: World Health Organization*.

Yin, X. *et al.* (2019) 'Direct costs of both inpatient and outpatient care for all type cancers: The evidence from Beijing, China', *Cancer Medicine*, p. cam4.2184. doi: 10.1002/cam4.2184.

Zhao, Z. and Bourne, P. E. (2018) 'Progress with covalent small-molecule kinase inhibitors', *Drug Discovery Today*, pp. 727–735. doi: 10.1016/j.drudis.2018.01.035.

Zhuravleva, A., Clerico, E. M. and Gierasch, L. M. (2012) 'An interdomain energetic tug-of-war creates the allosterically active state in Hsp70 molecular chaperones.', *Cell*. Elsevier, 151(6), pp. 1296–307. doi: 10.1016/j.cell.2012.11.002.

CHAPTER 2

2. Cancer, implicated heat shock proteins in therapeutic interventions and the impact of covalent vs non-covalent inhibition

2.1 Cancer overview

Cancer is a complicated and diverse disease that has been threateningly evading people worldwide. The scientific and medical society is expecting about 20 million of new cases yearly and proportional increase by 2025 (Ferlay *et al.*, 2012). Cancer implies the aberrance development of biochemical and cellular factors (Zhang *et al.*, 2008). In fact, normal cells undertake the continual cellular cycle of growth, division and death, cancer cells violate this normal order and rather than die, they exhibit continuous growth and division and multiplication which results in tumor development (Hanahan and Weinberg, 2011). Videlicet, disrupting the natural process of the proliferation and cell death generates tumors that eventually turn into cancer. This can affect all types of tissues and cells in the body, hence the various existing types of cancer. The disease is specifically unique since the growth of cancer can be distinctive in individual patient.

Several factors have been identified to cause cancer and they basically include genetic predisposition and environmental exposure to carcinogens (Bouvard *et al.*, 2009; Cogliano *et al.*, 2011; Hecht, 1999). Genetic predisposition entails the inheritance of defective genes from members of the same family. This, according to Knudson two hit hypothesis explains that a person with genetic mutations is highly predisposed to cancer development as life progresses (Knudson, 2001). The role of age in cancer predisposition has also been considered due to the propensities of cancer-inducing DNA mutations as aging occurs, hence age has been considered as a risk factor in cancer incidences. Likewise, pathogenicity of virulent organisms such human papillomavirus (HPV), hepatitis B and C, and Epstein-Barr virus have been related to carcinogenesis while immunosuppressive factors or diseases such as Human immunodeficiency virus (HIV) can also increase the tendency of cancer development (Samaras *et al.*, 2010; Plummer *et al.*, 2016; Haque *et al.*, 2018; WHO, 2018) (Figure 1.2).

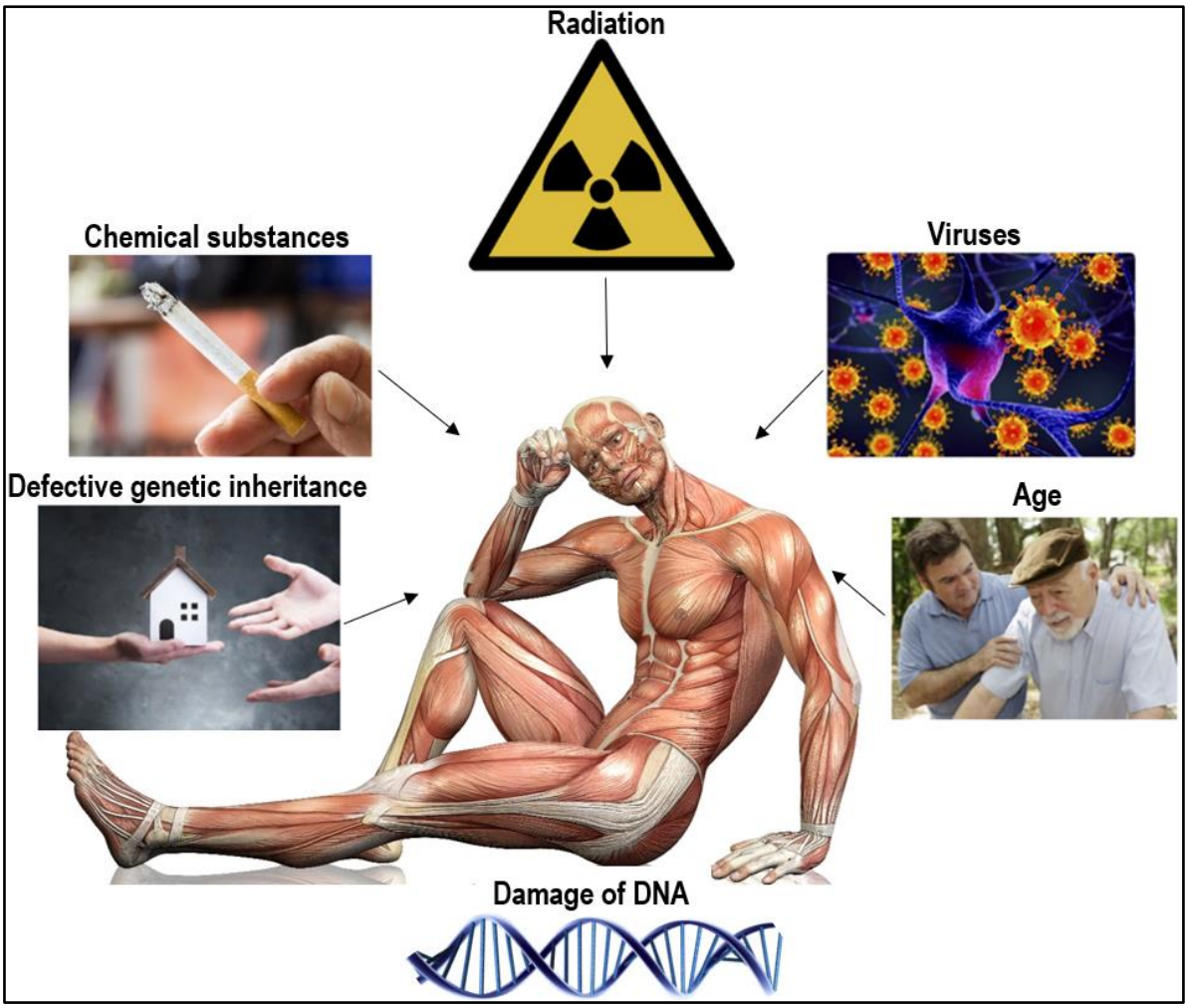


Figure 2. 1 Risk factors that contribute to cancer incidence.

Other than these above listed factors, the role of environmental or lifestyle factors have been greatly reported with regards to carcinogenesis (Knudson, 2001). These factors entail the intake of toxic substances, smoking or diet, which altogether are basically constitute carcinogenic; chemical substances that possess DNA-damaging abilities and they include arsenic, X-ray or gamma radiation, asbestos, tobacco, arsenic and carbon compounds from exhaust fumes (mainly poly aromatic hydrocarbons - PAHs). Free radicals are generated in the presence of carcinogens, which damage cellular biomolecules such as DNA and proteins and prevent their biological functions.

Current treatment strategies for cancer include radiotherapy, surgery and systemic treatment options which consist of cytotoxic chemotherapy, immunotherapy, hormonal therapy, and targeted or personalized therapies (Block *et al.*, 2015). The adoption of precision medicine, targeted therapies and immunotherapies have recorded wide success in the treatment and

management of cancer diseases. These approaches have resulted in the development of targeted drugs such as imatinib, gefatinib, erlotinib, sunitinib, nivolumab among several others, which are either used singly or in combination (Palumbo *et al.*, 2013; Bode and Dong, 2017; Shin, Bode and Dong, 2017). In various cases, this method selects drugs based on the upregulation of a pathway in the melanoma, although the drug does not inevitably categorize among the expression of the protein in healthy vs unhealthy or tumor tissue. Conventional forms of treatments are less effective and harmful in different types of cancers. Scientific developments have enabled the identification of signaling pathways and molecular targets specific and related to cancerous cells which results in treatments with better efficacy, selectivity and less toxicity, indeed, verification of compounds has depended on target validation first (Block *et al.*, 2015). Heat shock proteins (HSPs) are vastly expressed in numerous types of carcinomas. The circulating HSPs levels, along with the antibodies, are excellent biomarkers for analysing the aggressiveness and stage of certain cancer types. HSPs are involved in tumor cell proliferation, differentiation, invasion and metastasis (Wang *et al.*, 2014; Das *et al.*, 2019). Heat shock induces proteotoxicity and aggregation of proteins with heat shock proteins (HSPs), encouraging heat tolerance by preventing excessive protein aggregation caused by stress, supporting in refolding denatured proteins properly and, promoting their degradation if necessary (Mathew and Morimoto, 1998; D. Whitley, Goldberg and Jordan, 1999; Kleinjung *et al.*, 2003). In many cancers, heat shock protein (HSP) molecular chaperones are elevated, and its overexpression signals a poor prognosis in various cancer types in terms of response to therapy and survival (Cornford *et al.*, 2000; Ciocca *et al.*, 2003). HSP elevated expression in malignant cells plays a major role in protection against malignancy-related spontaneous apoptosis as well as, therapy-generated apoptosis, mechanisms that may underlie HSP's role in tumor treatment resistance and progression (Volloch and Sherman, 1999).

Hsp70 protein members are among the highly preserved proteins and play a critical part in these processes (Wu, Hunt and Morimoto, 1985). HSP72 is main stress-inducing member of the Hsp70 chaperone family and is encoded by two genes, HSPA1A and HSPA1B, which generate isoforms structure of HSP72 (Kampinga *et al.*, 2009). HSP72 is demonstrated as a promising anti-cancer target due to its key role in expressing at high level in malignant tumors of many origins (Jaattela, 1999). More aspects of this auspicious protein target, as well as the emplacement of the HSP70 family in the cancer proliferation and treatment, will be highlighted in the current section of this study. In addition, we discuss the advantages of therapeutic measures that have been previously used to target it besides their shortcomings.

2.2 Background on the Heat shock protein (HSPs) and cancer

The heat-shock response in all organisms, induced by a wide range of the variety of stimuli increases the expression of a protein family called Heat-shock proteins (HSPs) (Kroeger, 1960; Tissiéres, Mitchell and Tracy, 1974) that act as chaperones of molecules and they display cytoprotective properties for cells under stress (Li and Werb, 1982; Nollen *et al.*, 1999). Stress beyond a certain outset induces aggregation and misfolding of proteins and disruption of regulatory complexes (Yahara, 1996). Five principal heat shock proteins conserved classes which have chaperone activity: Hsp100, Hsp90, Hsp70, Hsp60, Hsp27 and the small heat-shock protein class (sHSPs) (Calderwood, 2018). Members of the HSP family are transported to different cellular compartments, either expressed regulated inductively or constitutive. ATP-dependent molecular chaperones are HSPs of large molecular weight; small HSPs act in an ATP-independent manner (Kumar *et al.*, 2016). Molecular chaperone genes may either be constitutively synthesized or, in the case of inducible HSP genes, are either silent or weakly expressed in non-stressed cells but activate after stress at dramatic rates to cope with the ensuing protein folding crisis (Feder and Hofmann, 1999). Inducible HSP genes each contain heat shock transcription factor 1 (HSF1) binding sites within their promoters, a protein that detects stress exposure almost spontaneously, changes from a monomer to a trimer type, and powerfully induces HSP genes (Calderwood and Murshid, 2017). HSPs are involved not only in the growth of tumors, but also in evaluating their response to treatment; therefore, disrupting HSPs as protein target involved in this process is a strong approach in anti-cancer therapy (Kumar *et al.*, 2016; Calderwood, 2018) (Figure 2.2).

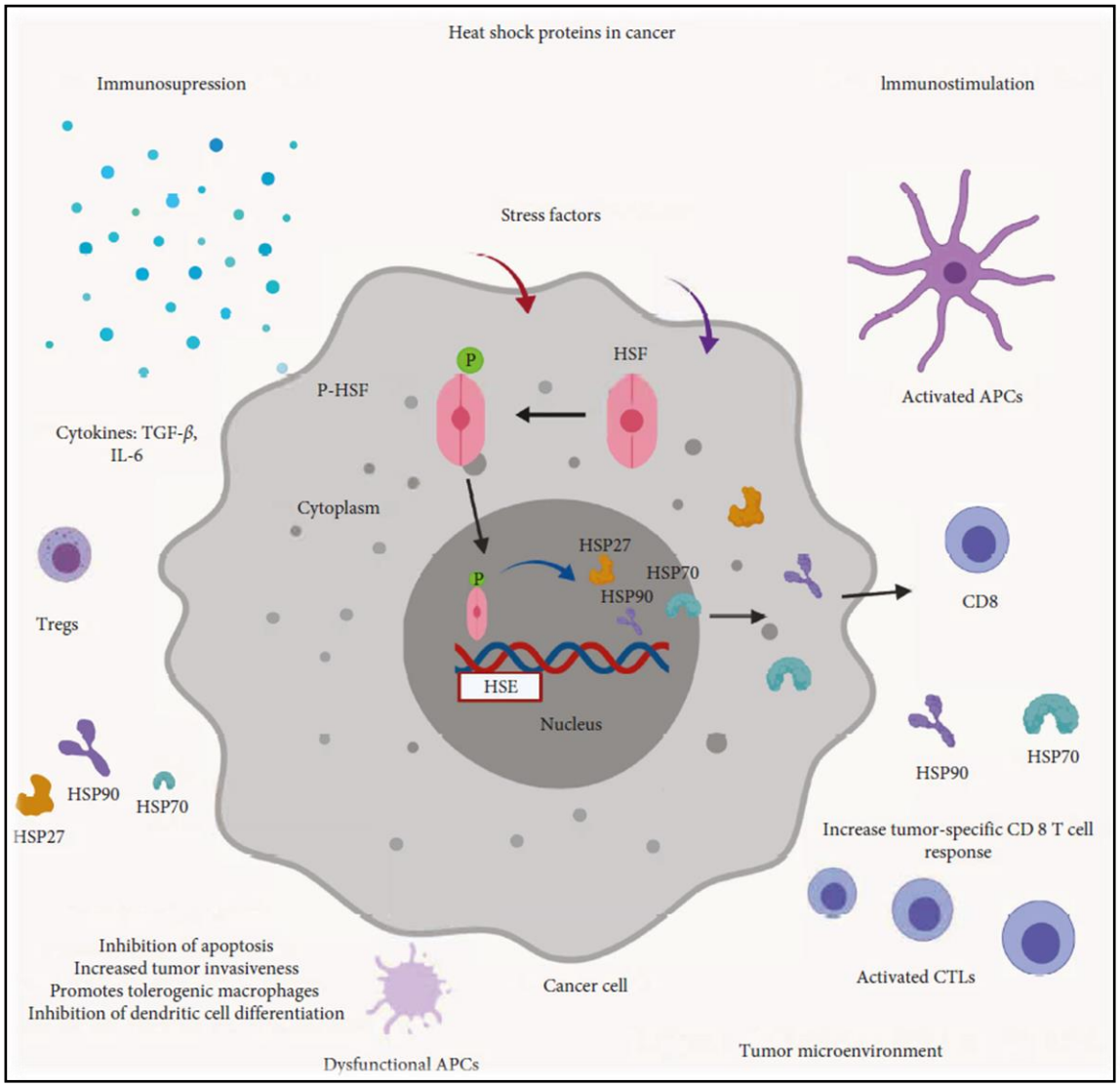


***Figure 2.2 Heat shock proteins and cancer. Cancer cells are incurred to several stress factors from the milieu extracellular of the tumor microenvironment. The stress factors activate heat shock transcription factors (HSFs) by facilitating their separation from heat shock proteins and phosphorylating them. The heat shock transcription factors bind with heat shock elements (HSE) where they translocated into the nucleus and initiated the transcription of heat shock proteins like HSP27, HSP70 (HSP72), and HSP90. The HSPs enhance the survival and proliferation of cancer cells by activating their cellular protection machinery, which is revealed in immunosuppressive conditions. The HSPs may also be maintaining a fine balance between cell death and survival by stimulating the anticancer immune response under optimal conditions. (Adapted from* (Das *et al.*, 2019)).**

HSPs most notably HSP27, HSP72 and HSP90 have been revealed as a standard inducers in many types of cancer (Ciocca and Calderwood, 2005; Arrigo and Gibert, 2014). Although the

increases caused by these mechanisms are not entirely understood, it has been proposed that HSPs proliferation in many cancer cells by responding to folding stresses that arise through tumour progression and malignant transformation (Workman, 2004). The accumulation of oncogenic proteins, many of which are mutated, and the friable, flexible structures that often accompany heavy enzymatic activity push for an increase in protein folding influence. (Workman, 2004; Gray, Stevenson and Calderwood, 2007). Since molecular chaperones work in a stoichiometric scheme, often associating in long-lived protein complexes with unhinged clients, instead of enzymes that operate in an attached and release catalytic manner, massive intracellular concentrations of HSPs are required for effective folding. (Ellis, 2007). It has been envisaging that this folding increasing in cancer, activates the heat shock response, leading to HSF1-dependent increases in transcription of HSP genes (Volloch and Sherman, 1999). Tumour cells have thus come to be observed as 'addicted to chaperones', becoming dependent on a continual supply of large concentrations of HSPs, and responding with some adversity to their withdrawal (Workman, 2004; WORKMAN *et al.*, 2007; Trepel *et al.*, 2010). However, the activation in terms of chaperone accumulation would be like the same, whether induced through an oncogenic signalling cascade or increased folding demand (Calderwood, 2018). Many of the main mechanisms that drive malignant transformation and tumor progression, such as escape from apoptosis stimulus-independent growth and senescence, de novo angiogenesis, metastasis, and invasion of surrounding tissues, are all induced by HSPs. (Workman, 2004; Trepel *et al.*, 2010; Hanahan and Robert A. Weinberg, 2011; Gibert *et al.*, 2012; Colvin *et al.*, 2014). Despite the fact that a narrow number of studies have focused on the critical function of HSPs in tumor models, the available evidence indicates different modes of action depending on the driver oncogene that intercedes transformation (Figure 2.3.)

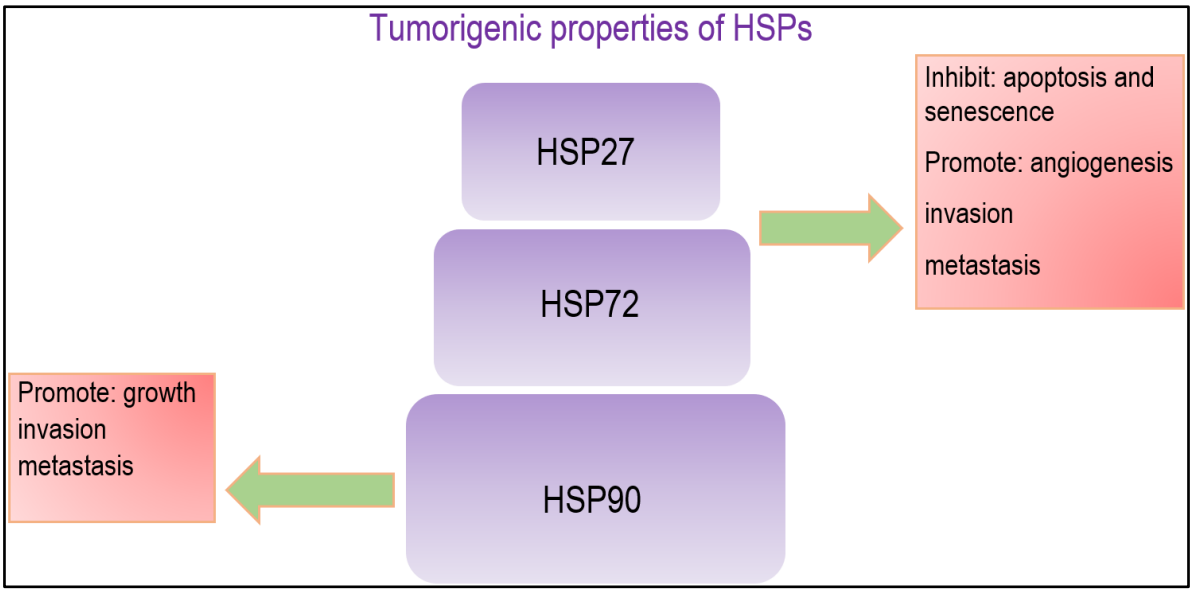


Figure 2. 3 Tumorigenic effects of intracellular HSPs. Elevated levels of Hsp70 and Hsp27 are particularly effective in inhibiting oncogene-induced apoptosis and senescence, also important in granting de novo angiogenesis, invasion and metastasis. Heightened levels of Hsp90 affect most areas of tumorigenesis, including stimulus-independent tumour cell growth, escape from senescence, invasion of surrounding tissues and metastasis.

2.3 HSP72 and implications in cancer therapy

Heat shock proteins play a crucial role in the damaged proteins clearance by inducing proteotoxicity and proteins acclamation. This process occurs by preventing unsuitable stress-induced protein accumulation, ensure proper refolding of denatured proteins, and, the promotion of their degradation if necessary (Mathew and Morimoto, 1998; D Whitley, Goldberg and Jordan, 1999; Nollen and Morimoto, 2002). Some cancers display elevated levels of Hsp70, and their expression has been associated with disease prognosis, cell proliferation and resistance to chemotherapy. As a consequence of this increase; thereby, cancer cells seem to be susceptible to agents that inhibit the elimination of aggregated or misfolded proteins generated by protein synthesis. (Hideshima *et al.*, 2005; Wu *et al.*, 2010; Zhu, Dunner and McConkey, 2010).

Hsp70 proteins are vastly preserved proteins and play a critical role in these processes (Wu, Hunt and Morimoto, 1985). The most important chaperon, a stress-inducing conserved member of the Hsp70 family, is known as HSP72, is coded by two genes, HSPA1A and HSPA1B

(Kampinga *et al.*, 2009). HSP72 is homologous to the 78 kDa glucose-regulated protein of 377 residues of amino acid that form in the nucleotide-binding domain (NBD), which plays a senior role in establishing the unfolding protein response. HSP72 is recognized at high levels in malignant tumors of many origins (Jäättelä, 1999) and enhances cancer cell survival (Nylandsted *et al.*, 2000, 2002). Therefore, the inhibition of HSP72 is identified to be a successful pathway in anti-tumor therapy (Powers *et al.*, 2010). (Figure 2.4)

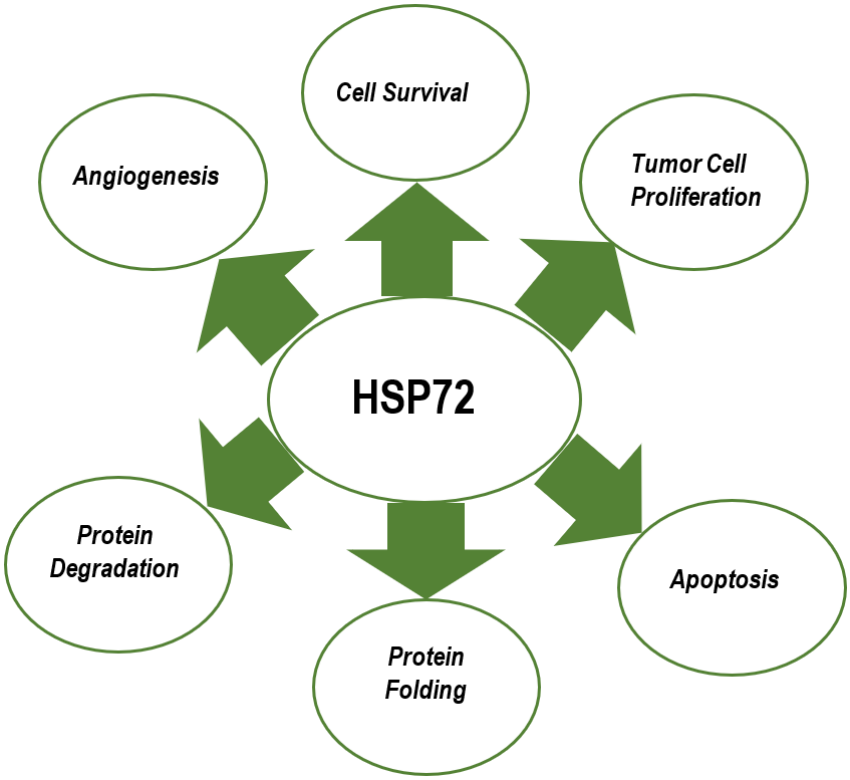


Figure 2. 4 Several roles of HSP72 in cancer related cellular processes

Hsp70 different functions are proficient through a transient chaperone interaction with substrate proteins through its C-terminal substrate-binding domain (SBD) (Mayer and Bukau, 2005). Controlling the transient chaperone interaction when the nucleotide binds allosterically to the N-terminal nucleotide-binding domain (NBD). The affinity of the substrate-binding domain (SBD) for substrates decreases from 10 to 400 folds when ATP is binding to the N-terminal of the nucleotide-binding domain (NBD). Hence, the inhibition of the nucleotide-binding domain (NBD) is considered one of the most promising strategies for HSP72 function inhibition (Zhuravleva, Clerico and Gierasch, 2012). The NBD is composed of two adjacent lobes (lobe I and lobe II), which connect to the base via a deep nucleotide groove. Each lobe consists of two subdomains (IB, IIB, IA and IIA) (Sharma and Masison, 2009; Penkler *et al.*,

2017). IA and IIA Domains are linked to IB and IIB, respectively, by flexible hinges and control access to the nucleotide-binding sites (Chiappori *et al.*, 2012).

HSP72 helps to maintain the probity of protein structures and their transfer by facilitating precise protein folding and degradation of misfolded or old proteins across several cell organelles (Sharma and Verma, 2018). Stress-induced HSP72 expression allows cells to deal with great quantities of unfolded or denatured proteins. Moreover, the functions of HSP72 involve transportation of precursor proteins into cellular comportments; degradation of unstable proteins and protein complexes; protein folding (cytosol, mitochondria and endoplasmic reticulum); regulatory proteins control and protein refolding (Voos, 2003; Young *et al.*, 2004; Hennessy *et al.*, 2005). (Figure 2.5)

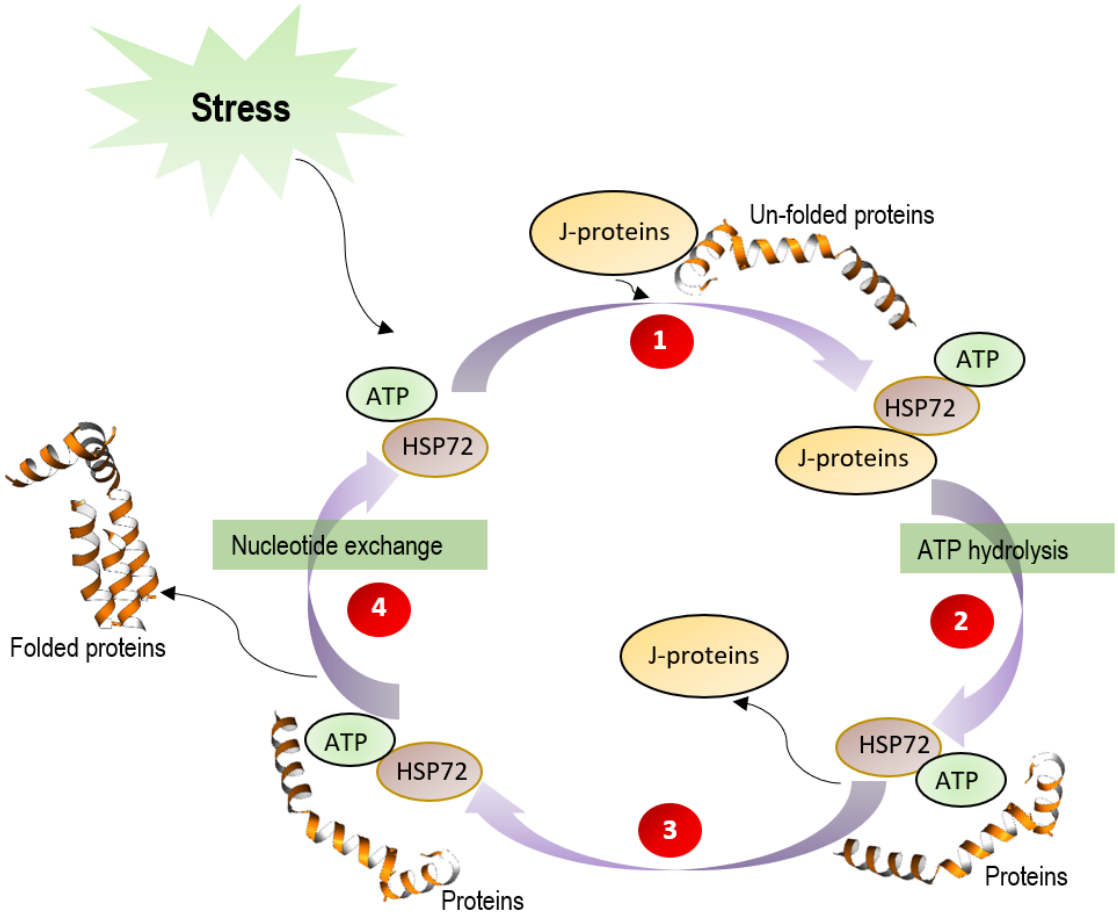


Figure 2. 5 Diagrammatic presentation of the chaperone functions of HSP72* (Young *et al.*, 2004; Hennessy *et al.*, 2005). *HSP72 binds to unfolded protein in the presence of J-proteins using ATP. ATP hydrolysis stimulates J-protein release that further led to release of correctly folded protein(s), and HSP72 become available for the next cycle

2.4 Important strategies in inhibiting HSP72

2.4.1 Non-covalent inhibition

Non-covalent interactions are essential in the development of a new substance with specific properties (Hagel *et al.*, 2011; Halgren, 2009; Swinney, 2004). These interactions have been identified in different fields correlated to chemical pharmacological, physical, materials and biological sciences (Smith *et al.*, 2009; Bauer, 2015; Potashman and Duggan, 2009). The study of non-covalent interactions such as hydrogen bonding, van der Waals interactions, electrostatic interactions, hydrophobic interactions, and salt bridges has been a major focus in drug development (Zhanting Li *et al.*, 2015). Understanding these interactions and the physical mechanisms that underpin them is critical for improving the current drug design strategy. As a consequence, many efforts have been dedicated to explaining the structural, geometrical, energetic, and thermodynamic properties of these interactions. The strong hydrogen bonds (Figure 2.6:A), with their high directional characteristics is very effective in the supramolecular assembly of molecules (Desiraju, 2011). In sight of their importance in many science fields, the debate on the nature, properties, and importance of hydrogen bonds are increasing. However, the van der Waals interaction, outcome from a momentary random fluctuation in the circulation movement of the electrons of any atom (Figure 2.6:C), which give upturn to a transient differing circulation of electrons, a transient electric dipole (Zhang, 2013). van der Waals interactions are responsible for the constancy between molecules of solids and nonpolar liquids, heptane as example, cannot form ionic interactions or hydrogen bonds with other molecules (Hunter, 2013). The present of these strong interactions, overrule most of the effect of van der Waals interactions, and the strength efficiency of these interactions decreases rapidly with increasing distance; thus, these non-covalent bonds can form only when atoms are quite near to one another (Finkelstein, 2007). Nonetheless, Ionic bonds (Figure 2.6:B), generate from the electrostatic attraction between the positive and negative charges of ions (Mahmudov *et al.*, 2017). The Hydrophobic bonds occur between nonpolar molecules, such as hydrocarbons, in an aqueous environment (Privalov, 1988). In most cases, compared to their natural substrate analogues, non-covalent reversible inhibitors bind to the target active site with a higher affinity. Since the bound inhibitor is non-covalently bonded at the active site and is reversibly released from the active site, enzyme deactivation is not irreversible in these cases (Puhan *et al.*, 2017).

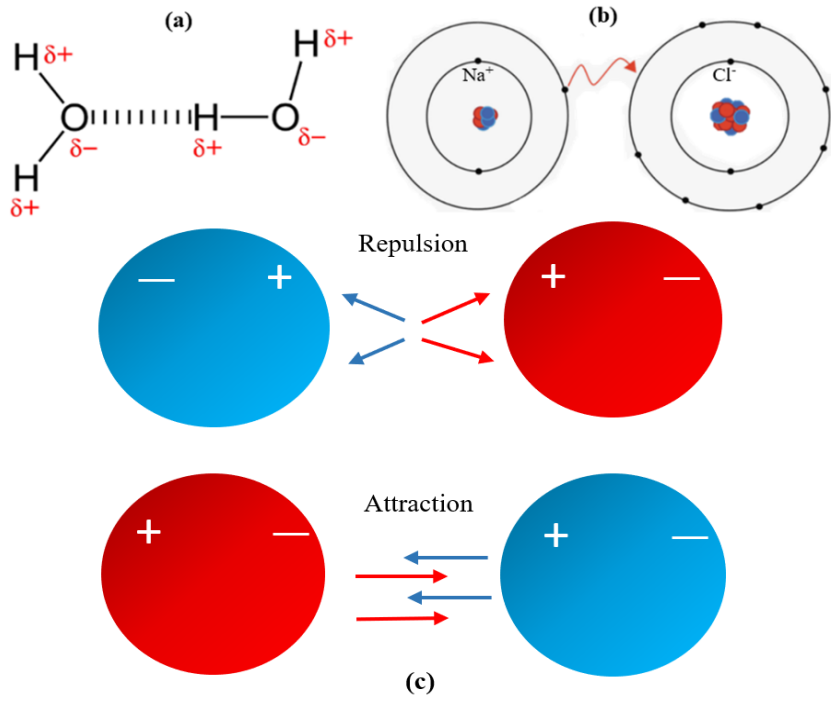


Figure 2. 6 Non-covalent interactions: (A) Strong hydrogen non-covalent bond. (B) Ionic bond between two atoms Na^+ and Cl^- by electrostatic force. (C) Van der Waal interaction force.

2.4.2 Historical framework of covalent inhibitors in drug discovery

As a validated approach to avoid acquired resistance in malignancy, Covalent inhibition is a rapidly emerging paradigm in the process of drug discovery. (Schwartz *et al.*, 2014) (Pettinger, Jones and Cheeseman, 2017). The process of developing small molecule inhibitors to inhibit the target protein covalently dates back to the establishment of drug discovery field. Aspirin (acetylsalicylic acid) was manufactured by Bayer in the late 1800s century to treat pain, and inflammation. Until the 1970s, it was unknown that aspirin works by the covalent mechanism of inhibition of cyclooxygenase-1 and -2 (COX-1 and 2) enzymes (Bauer, 2015). Numerous covalent inhibitors in the past were invented accidentally. Recent development has seen a huge modification in this outlook, as covalent inhibition expressed potential for targets where earlier attempts to investigate non-covalent small molecule inhibitors have flopped (Lonsdale and Ward, 2018).

Covalent drugs have resurged in recent years, and various reports have acclaimed the general advantages of designing covalent inhibitors (John, Long and Aye, 2017). These drugs occupy a very distinguish class in the history of therapeutics. There are several drugs in the market,

however very less have been intentionally designed to interact covalently with the biological target (De Cesco *et al.*, 2017).

In drug discovery area, the commercial and clinical progress of covalent drugs has urged an improved and more thoughtful detection of covalent mechanism (Strelow, 2017). Drugs that form a covalent bond to their target have typically been considered as different from conventional non-covalent drugs. There has been a hesitancy to apply a covalent mode of action in drug discovery, despite of having many examples of efficacious drugs through a covalent mechanism (Figure 2.7) (Singh *et al.*, 2011).

The small molecule activators and inhibitors that link to a target non-covalently and reversibly are always the focus of conventional structure-based drug design. Enhancements in potency and selectivity are usually attained by improving shape and non-covalent interfaces (van der Waals interactions, hydrogen bonds and salt bridges) among inhibitor and a target binding site. (Lonsdale and Ward, 2018).

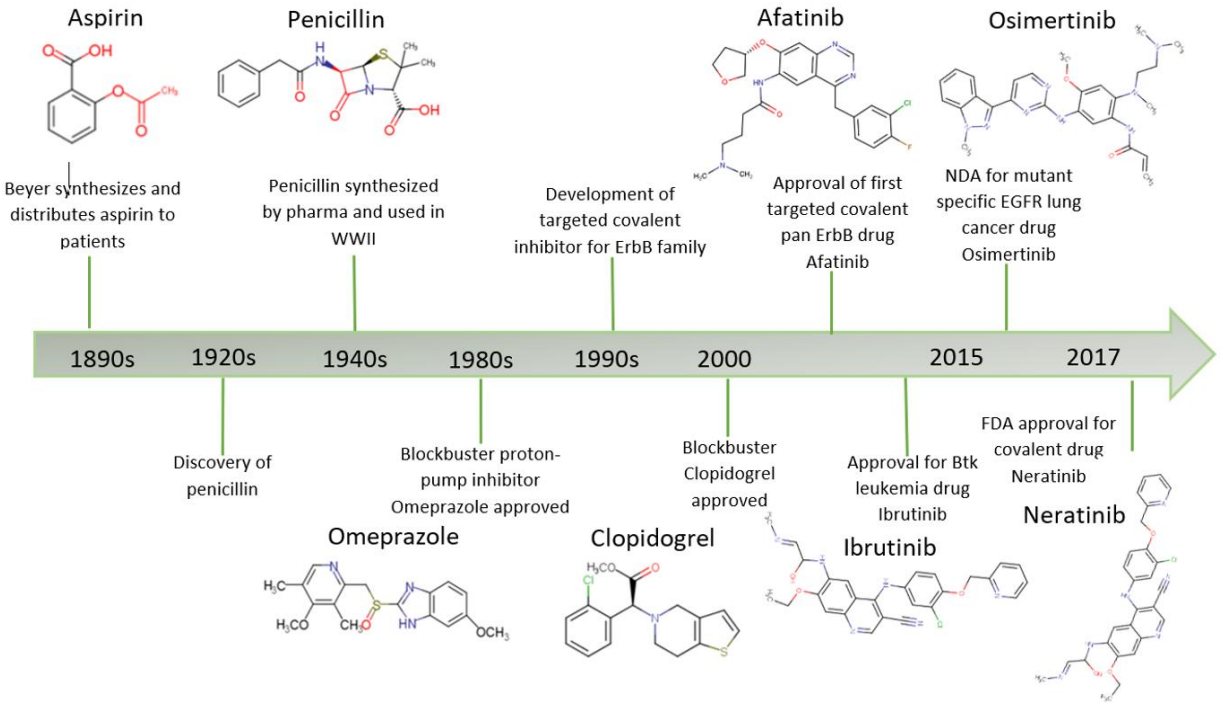


Figure 2. 7 Timeline representation of approved covalent inhibitors in history.

In drug discovery enterprises, a non-conventional approach coined as 'covalent inhibition' has invaded the cognizance of an increasing number of drug hunting companies. In the covalent mechanism of inhibition, a small molecule is designed not only to bind to a target protein through the conventional non-covalent mechanism but also to endure a bond-forming action

that results in a durable drug-protein association. The resulting covalent bond is irreversible and sufficiently long-lived within the half-life of the target protein, resulting in a drug-protein complex that is not subject to classical equilibrium (Bauer, 2015). In the first step (A) a covalent drug [D] links to its target enzyme [E] through a non-covalent mechanism of inhibition to generate an enzyme-drug complex [ED]. In another step (B), second chemical reaction results among the drug [D] and enzyme [E] to produce covalent irreversible complex [E-D] by leading conformational alterations in the final complex (Figure 2.8) (Khan *et al.*, 2018). A most appropriate example of covalent inhibition is afatinib, a proven drug for metastatic non-small lung cancer.

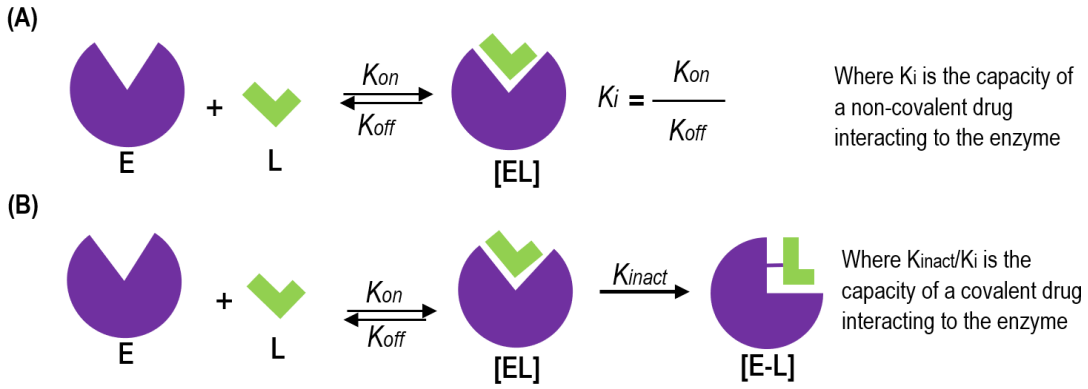


Figure 2. 8 Drug-enzyme interactions. Association of a Ligand (L) interacting with Enzyme (E) in (A) conventional non-covalent and (B) covalent mode of action. Where is L: Ligand; E: Enzyme.

2.4.2.1 *Warhead moiety*

The warhead moiety plays a significant role in covalent inhibition by regulating the covalent-interaction-induced toxicity of targeting. Researchers integrate various warhead moieties, intending the balance of risk versus treatment (Zhao and Bourne, 2018). There is resurging attention in small molecule inhibitors that involve their target through covalent inhibition. Cysteine's thiol (-SH) is enriched with improved reactivity, making it the nucleophile of interest for covalent inhibition with an inhibitor allying an electrophilic trap with a cysteine amino acid residue in a target of interest (Lagoutte, Patouret and Winssinger, 2017). For instance, the deprotonated thiolate anion is a highly selective nucleophilic moiety which is used as a crucial catalytic residue in cysteine proteases and phosphatases. Cysteines also function as

an essential non-catalytic residue in stabilizing tertiary structure of a protein by forming a disulphide-cross link and direction of enzyme cofactors (metals).

To date, most covalent inhibition approaches have been discovered by targeting the highly nucleophilic cysteine thiolate. Different types of electrophilic warhead moieties that can counter with nucleophilic amino acid residues like cysteine, tyrosine or lysine have been investigated in designing covalent inhibitors as showing below. The most extensive chemical reaction used to achieve irreversible binding is Michael addition reaction (Figure 2.9).

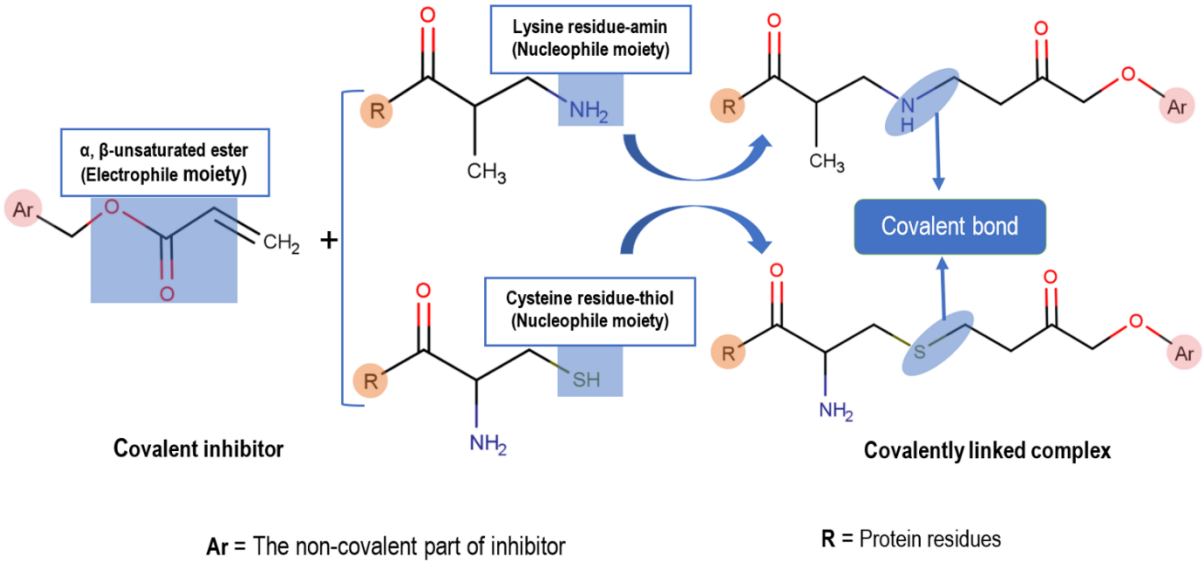


Figure 2. 9 A schematic representing Michael addition as a popular example of covalent reaction between, in this case, α, β-unsaturated ester of the ligand and both cysteine and lysine residues of the protein (these two residues can be any electrophile and nucleophile moieties).

The commonly used functional groups to sustain this addition reaction involve, vinyl sulfonates, quinones, acrylamides, alkynyl amides and propargylic acid derivatives (Figure 2.10). The efficiency to target distinctly altered cysteine amino acid residue may provide further means of attaining selectivity (Liu *et al.*, 2013).

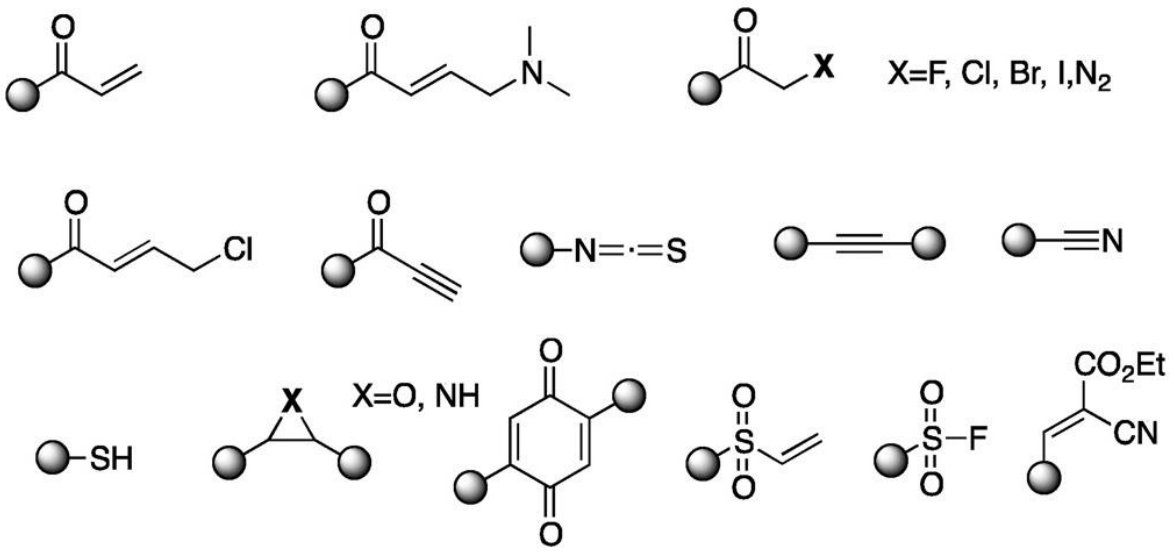


Figure 2. 10 Warhead moieties used in covalent drugs (Adapted from (Liu *et al.*, 2013)***).***

2.4.2.2 Benefits and risks of covalent drugs in cancer therapy

Approximately, 30% of drugs available in the market work via a covalent mechanism, however, most of them were not manufactured as covalent inhibitors and were discovered by serendipity (Mah, Thomas and Shafer, 2014). Drugs that form a covalent bond to their target protein have been treated as a separate class of drugs from conventional non-covalent drugs. This class of drugs has verified to be a highly profitable class in pharmaceutical production. In the United States, the three top-most selling drugs that achieved blockbusters (lansoprazole, esomeprazole and clopidogrel) are covalent (Singh *et al.*, 2011). There are many examples of successful covalent drugs, and the illustrative examples are shown in (Figure 2.11).

Covalent drugs are used to cure various diseases like cancer, often link with a specific amino acid residue in target protein in a highly selective manner and are irreversible by generating prolonged effects (Serafimova *et al.*, 2012). According to Barf and Kaptein, while perceiving covalent inhibition as a method, there can be a significant balance between the benefits and potential risks of covalent inhibitors (Barf and Kaptein, 2012).

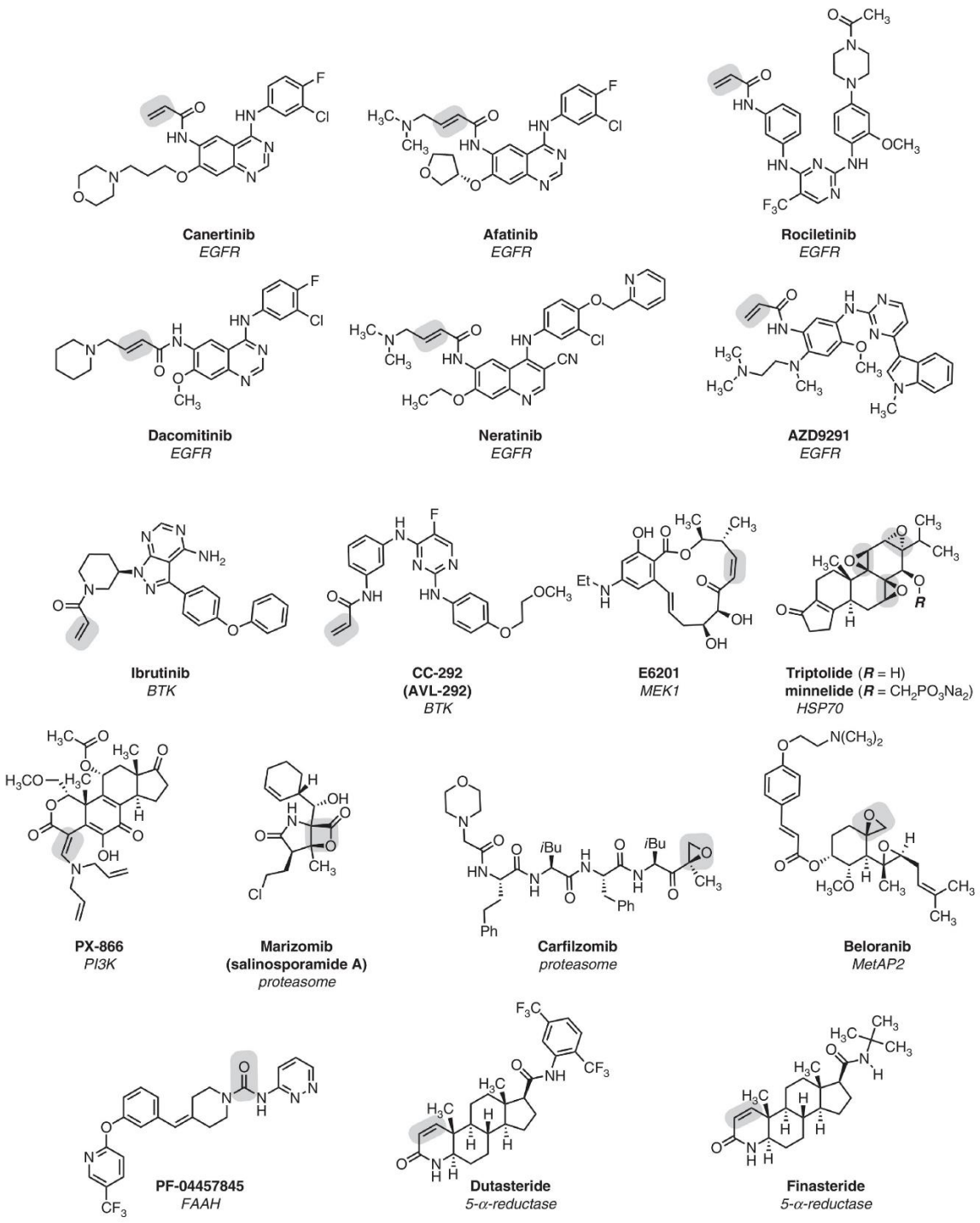


Figure 2. 11 Few examples of covalent drugs. Electrophilic warhead moiety is highlighted in grey color (Adapted from (Bauer, 2015)***).***

Covalent inhibitors are the class of the small molecule compounds that inactivate the enzyme through the formation of a covalent attachment between the nucleophilic region of the enzyme and electrophilic region of the inhibitor. The formation of the covalent bond results to the loss

of enzymes function may be either because it alters the active site amino acid residue important for the chemical reaction or it does not allow access to enzymes catalytic or allosteric site (Rempel and Withers, 2008).

Although the risks of covalent drugs are well known, yet the long-lasting duration of covalent bond creates a number of potential benefits including: (i) enhanced biochemical efficacy, (ii) less frequent dosing decrease a patient's drug burden without compromising overall drug efficiency, (iii) most efficient approach when a complete loss of enzymes activity is needed, (iv) potential to circumvent drug resistance mechanism and (v) a dissociation of pharmacokinetics (PK) from pharmacodynamics (PD) because PD relies on the synthesis of a new enzyme (Figure 2.12) (Mah, Thomas and Shafer, 2014).

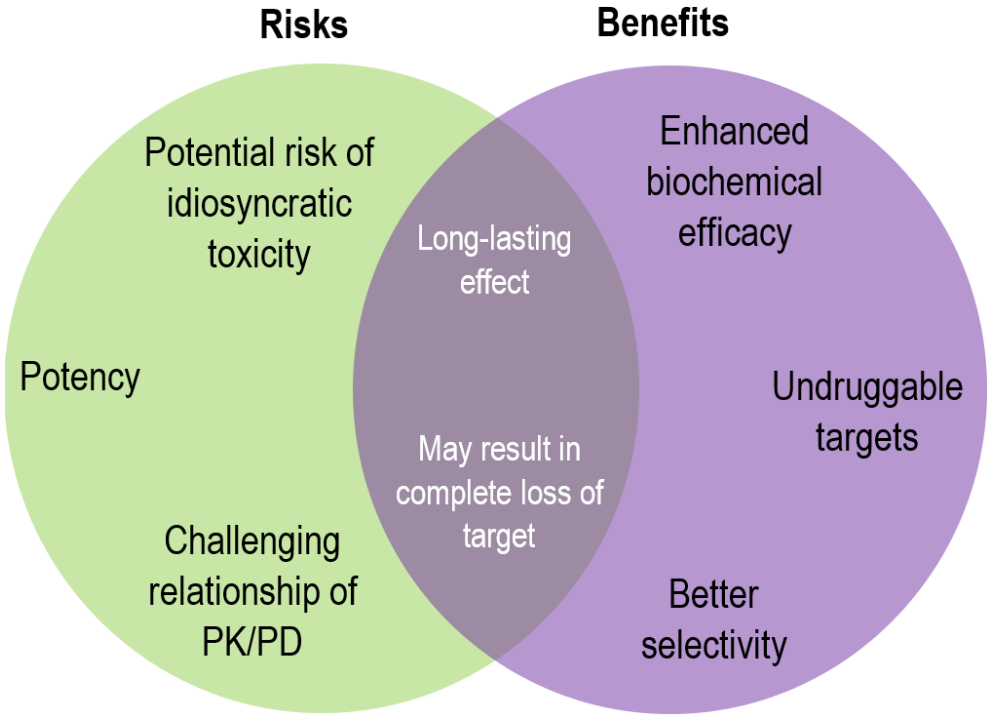


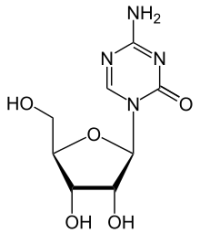
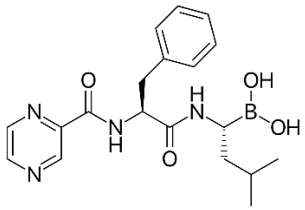
Figure 2. 12 Benefits and risks of covalent drugs (Image prepared by author).

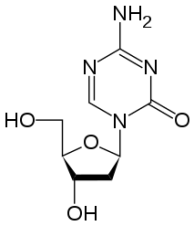
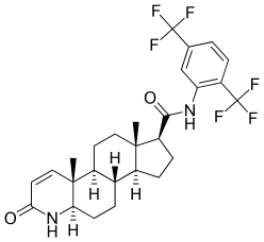
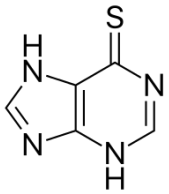
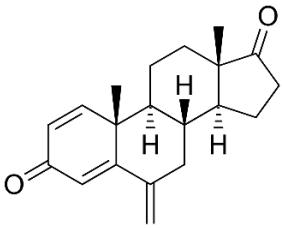
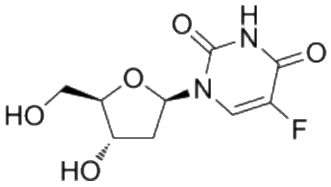
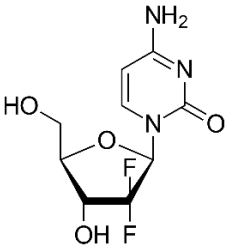
Although concerns are associated with the toxicity, covalent drugs have had a major influence on human health. Remarkably, the FDA approved 27 drugs in 2013, out of which three were covalent drugs. Finally, covalent mode of inhibition may be a misjudged approach for focusing on critical targets and 'undruggable' modalities in human diseases (Bauer, 2015).

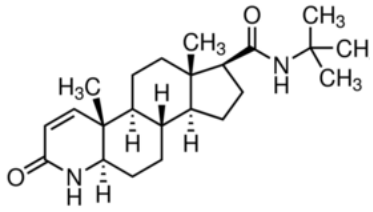
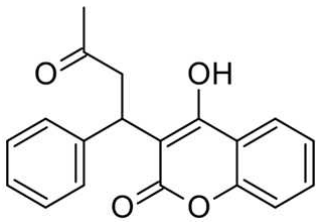
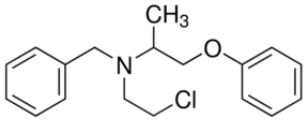
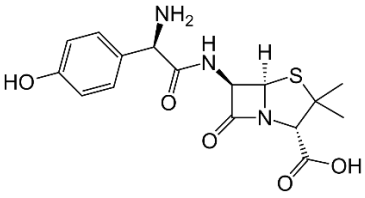
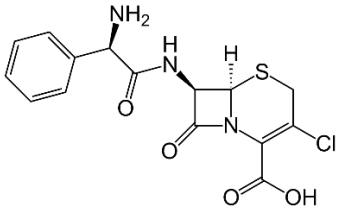
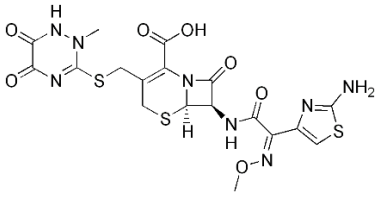
2.4.2.3 Clinical developments of covalent drugs

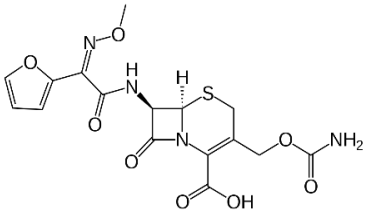
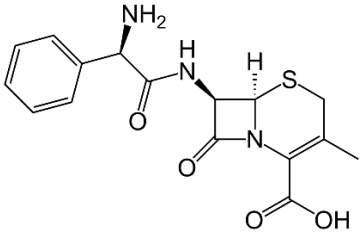
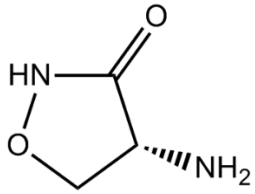
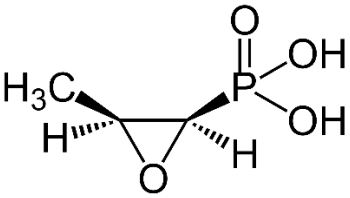
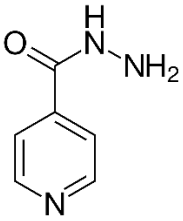
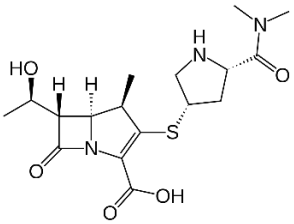
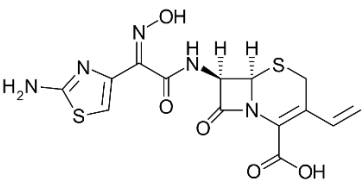
Covalent drug design is a rapidly growing and promising field in drug discovery, however principle for the rational design of these small molecule compounds have emerged recently by enabling the development of this therapeutic class (Singh *et al.*, 2011). These drugs have many benefits over conventional non-covalent drugs; low dosage is required to indicate the needed pharmaceutical effect, produce balance and long-lived links and are resistant to mutations (Ruddraraju *et al.*, 2017). There is a broad scope of development in this field, more specifically by targeting non-catalytic amino acid residues other than catalytic cysteines. The range and number of covalent drugs, and the interest of how to develop them, is constantly increasing. There are many schemes for covalent drug discovery projects; the preference will be affected by the quantity of preceding information and/or screening library databases (Lonsdale and Ward, 2018). Computational chemistry methods can have a key effect on covalent drug design and will sustain to be an important feature of the expansion of this interesting field (Lonsdale and Ward, 2018). Although most pharmaceutical industries hesitate to design covalent inhibitors due to toxicity, there are many examples of FDA approved covalent drugs. Table 2.1 summarizes the covalent drugs based on their biological target and therapeutic domain.

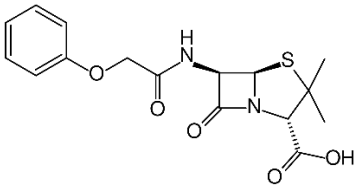
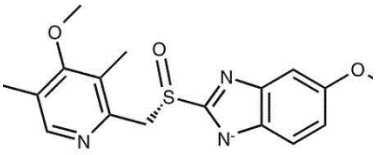
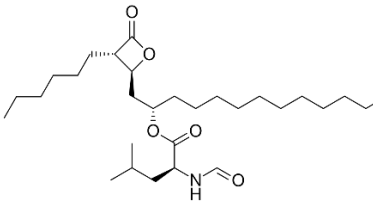
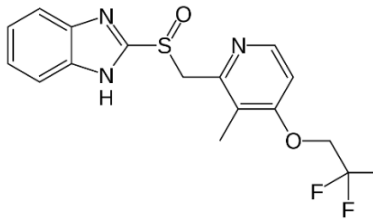
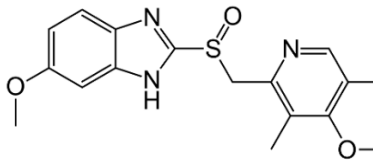
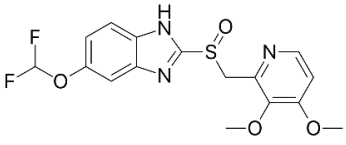
Table 2. 1. List of FDA approved covalent drugs.

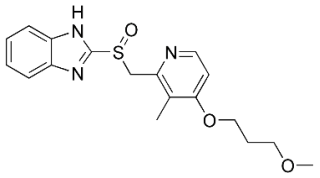
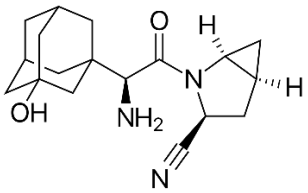
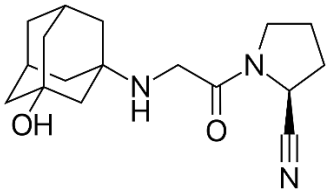
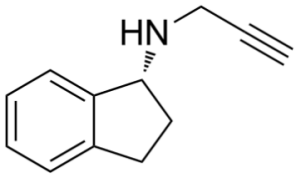
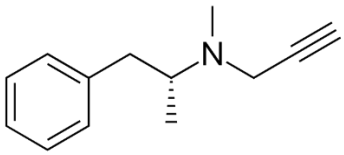
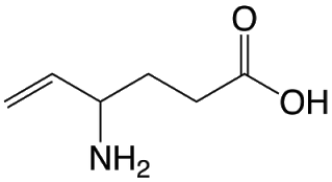
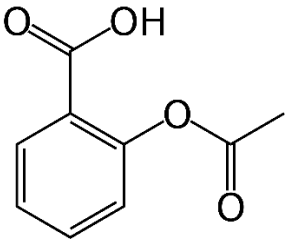
Drug	Structure	Target	Domain
Azacytidine		Methyltransferase	Cancer
Bortezomib		Proteasome	Cancer

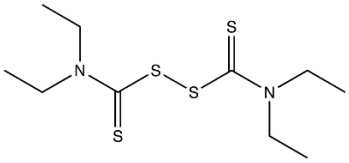
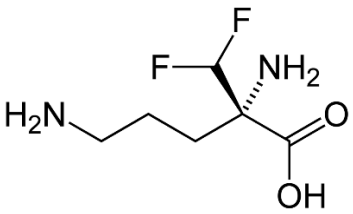
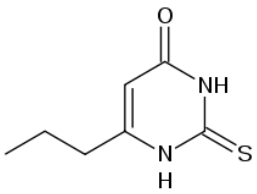
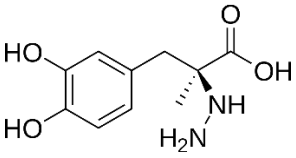
Decitabine/ **AzadC**		Methyltransferase	Cancer
Dutasteride/ **Avodart**		5-α-Reductase	Cancer
Mercaptopurine/Purinthol		Purine-nucleotide synthesis	Cancer
Exemestane/ **Aromasin**		Aromatase	Cancer
Floxuridine		Thymidylate synthase	Cancer
Gemcitabine/ **Gemzar**		Ribonucleoside reducatse	Cancer

Proscar/ Finasteride		5-α-Reductase		Cancer
Warfarin		Vitamin K reductase		Cardio-vascular
Phenoxy-benzamine hydrochloride		α-Adrenoceptor		Cardio-vascular
Amoxicillin		PBP		Anti-infective
Cefaclor/ Ceclor		PBP		Anti-infective
Ceftriaxone/ Rocephin		PBP		Anti-infective

Cefuroxime axetil		PBP	Anti-infective
Cephalexin/ Keflex		PBP	Anti-infective
D-cycloserine/ Seromycin		Alanine racemase	Anti-infective
Fosfomycin/ Monurol		UDP-N-acetylglucosamine -3-enolpyruvyl-transferase	Anti-infective
Isoniazid		Enol-acyl carrier protein reductase	Anti-infective
Meropenem		PBP	Anti-infective
Omnicef		PBP	Anti-infective

Penicillin V		PBP	Anti-infective
Nexium/ Esomeprazole		H+/K+ ATPase	Gastro-intestinal
Orlistat		Lipase	Gastro-intestinal
Prevacid/ Iansoprazole		H+/K+ ATPase	Gastro-intestinal
Prilosec/ Omeprazole		H+/K+ ATPase	Gastro-intestinal
Protonix/ Pantoprazole		H+/K+ ATPase	Gastro-intestinal

Aciphex/ **Rabeprazole**		H+/K+ ATPase	Gastro-intestinal
Saxagliptin/ **Onglyza**		DPP-IV	Anti-diabetic
Vildagliptin/ **Eugreas**		DPP-IV	Anti-diabetic
Rasagiline		MAO-B	Parkinson's disease
Selegiline		MAO-B	Anti-epileptic
Vigabatrin/ **Sabril**		GABA-Aminotransferase	Inflammation
Aspirin		Cyclooxygenase	Chronic alcoholism

Disulfiram/ Antabuse		Aldehyde dehydrogenase	Hirsutism
Eflornithine		Ornithine decarboxylase	Hyperthyroidism
Propylthiouracil/Procasil		Thyroxine-5-deiodinase	CNS
Carbidopa/ Lodosyn		DOPA decarboxylase	

References

Arrigo, A.-P. and Gibert, B. (2014) 'HspB1, HspB5 and HspB4 in Human Cancers: Potent Oncogenic Role of Some of Their Client Proteins', *Cancers*, 6(1), pp. 333–365. doi: 10.3390/cancers6010333.

Barf, T. and Kaptein, A. (2012) 'Irreversible protein kinase inhibitors: Balancing the benefits and risks', *Journal of Medicinal Chemistry*, pp. 6243–6262. doi: 10.1021/jm3003203.

Bauer, R. A. (2015) 'Covalent inhibitors in drug discovery: From accidental discoveries to avoided liabilities and designed therapies', *Drug Discovery Today*, pp. 1061–1073. doi: 10.1016/j.drudis.2015.05.005.

Block, K. I. *et al.* (2015) 'Designing a broad-spectrum integrative approach for cancer prevention and treatment', *Seminars in Cancer Biology*, 35, pp. S276–S304. doi: 10.1016/j.semcancer.2015.09.007.

Bode, A. M. and Dong, Z. (2017) 'Precision oncology- the future of personalized cancer medicine?', *npj Precision Oncology*, 1(1), p. 2. doi: 10.1038/s41698-017-0010-5.

Bouvard, V. *et al.* (2009) 'A review of human carcinogens--Part B: biological agents.', *The Lancet. Oncology*, 10(4), pp. 321–2.

Calderwood, S. K. (2018) 'Heat shock proteins and cancer: intracellular chaperones or extracellular signalling ligands?', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1738), p. 20160524. doi: 10.1098/rstb.2016.0524.

Calderwood, S. K. and Murshid, A. (2017) 'Molecular Chaperone Accumulation in Cancer and Decrease in Alzheimer's Disease: The Potential Roles of HSF1', *Frontiers in Neuroscience*, 11(APR). doi: 10.3389/fnins.2017.00192.

De Cesco, S. *et al.* (2017) 'Covalent inhibitors design and discovery', *European Journal of Medicinal Chemistry*. Elsevier Masson, 138, pp. 96–114. doi: 10.1016/J.EJMECH.2017.06.019.

Chiappori, F. *et al.* (2012) 'Molecular Mechanism of Allosteric Communication in Hsp70 Revealed by Molecular Dynamics Simulations', *PLoS Computational Biology*. Public Library of Science, 8(12), p. e1002844. doi: 10.1371/journal.pcbi.1002844.

Ciocca, D. R. *et al.* (2003) 'Hsp25 and Hsp70 in rodent tumors treated with doxorubicin and lovastatin', *Cell Stress and Chaperones*, 8(1), pp. 26–36. doi: 10.1379/1466-1268(2003)8<26:HAHIRT>2.0.CO;2.

Ciocca, D. R. and Calderwood, S. K. (2005) 'Heat shock proteins in cancer: Diagnostic, prognostic, predictive, and treatment implications', *Cell Stress and Chaperones*, 10(2), pp. 86–103. doi: 10.1379/CSC-99r.1.

Cogliano, V. J. *et al.* (2011) 'Preventable Exposures Associated With Human Cancers', *JNCI Journal of the National Cancer Institute*, 103(24), pp. 1827–1839. doi: 10.1093/jnci/djr483.

Colvin, T. A. *et al.* (2014) 'Hsp70–Bag3 Interactions Regulate Cancer-Related Signaling Networks', *Cancer Research*, 74(17), pp. 4731–4740. doi: 10.1158/0008-5472.CAN-14-0747.

Cornford, P. A. *et al.* (2000) 'Heat shock protein expression independently predicts clinical outcome in prostate cancer', *Cancer Research*, 60(24), pp. 7099–7105.

Das, J. K. *et al.* (2019) 'Heat shock proteins in cancer immunotherapy', *Journal of Oncology*, 2019, pp. 1–9. doi: 10.1155/2019/3267207.

Desiraju, G. R. (2011) 'A bond by any other name', *Angewandte Chemie - International Edition*, 50(1), pp. 52–59. doi: 10.1002/anie.201002960.

Ellis, R. J. (2007) 'Protein misassembly: Macromolecular crowding and molecular chaperones', in *Advances in Experimental Medicine and Biology*. New York, NY: Springer New York, pp. 1–13. doi: 10.1007/978-0-387-39975-1_1.

Feder, M. E. and Hofmann, G. E. (1999) 'HEAT-SHOCK PROTEINS, MOLECULAR CHAPERONES, AND THE STRESS RESPONSE: Evolutionary and Ecological Physiology', *Annual Review of Physiology*, 61(1), pp. 243–282. doi: 10.1146/annurev.physiol.61.1.243.

Ferlay, J. *et al.* (2012) *GLOBOCAN 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012 v1.0*. International Agency for Research on Cancer.

Finkelstein, A. V. (2007) 'Average and extreme multi-atom Van der Waals interactions: Strong coupling of multi-atom Van der Waals interactions with covalent bonding', *Chemistry Central Journal*, 1(1), pp. 1–9. doi: 10.1186/1752-153X-1-21.

Gibert, B. *et al.* (2012) 'Targeting heat shock protein 27 (HspB1) interferes with bone metastasis and tumour formation in vivo', *British Journal of Cancer*, 107(1), pp. 63–70. doi: 10.1038/bjc.2012.188.

Gray, P. J., Stevenson, M. A. and Calderwood, S. K. (2007) 'Targeting Cdc37 Inhibits Multiple Signaling Pathways and Induces Growth Arrest in Prostate Cancer Cells', *Cancer Research*, 67(24), pp. 11942–11950. doi: 10.1158/0008-5472.CAN-07-3162.

Hagel, M. *et al.* (2011) 'Selective irreversible inhibition of a protease by targeting a noncatalytic cysteine', *Nature Chemical Biology*, 7(1), pp. 22–24. doi: 10.1038/nchembio.492.

Halgren, T. A. (2009) 'Identifying and Characterizing Binding Sites and Assessing Druggability', *Journal of Chemical Information and Modeling*. American Chemical Society, 49(2), pp. 377–389. doi: 10.1021/ci800324m.

Hanahan, D. and Weinberg, Robert A. (2011) 'Hallmarks of Cancer: The Next Generation', *Cell*, 144(5), pp. 646–674. doi: 10.1016/j.cell.2011.02.013.

Hanahan, D. and Weinberg, Robert A. (2011) 'Hallmarks of Cancer: The Next Generation', *Cell*, 144(5), pp. 646–674. doi: 10.1016/j.cell.2011.02.013.

Haque, M. *et al.* (2018) 'Health care-associated infections - an overview.', *Infection and drug resistance*. Dove Press, 11, pp. 2321–2333. doi: 10.2147/IDR.S177247.

Hecht, S. S. (1999) 'Tobacco smoke carcinogens and lung cancer.', *Journal of the National*

Cancer Institute, 91(14), pp. 1194–210.

Hennessy, F. *et al.* (2005) 'Not all J domains are created equal: Implications for the specificity of Hsp40-Hsp70 interactions', *Protein Science*, 14(7), pp. 1697–1709. doi: 10.1110/ps.051406805.

Hideshima, T. *et al.* (2005) 'Small-molecule inhibition of proteasome and aggresome function induces synergistic antitumor activity in multiple myeloma.', *Proceedings of the National Academy of Sciences of the United States of America*, 102(24), pp. 8567–72. doi: 10.1073/pnas.0503221102.

Hunter, C. A. (2013) 'Van der Waals interactions in non-polar liquids', *Chemical Science*, 4(2), pp. 834–848. doi: 10.1039/c2sc21666c.

Jaattela, M. (1999) 'Heat shock proteins as cellular lifeguards', *Annals of Medicine*.

Jäättelä, M. (1999) 'Heat shock proteins as cellular lifeguards', *Annals of Medicine*, 31(4), pp. 261–271. doi: 10.3109/07853899908995889.

John, M., Long, C. and Aye, Y. (2017) 'Cell Chemical Biology Perspective Privileged Electrophile Sensors: A Resource for Covalent Drug Development', *Cell Chemical Biology*, 24, pp. 787–800. doi: 10.1016/j.chembiol.2017.05.023.

Kampinga, H. H. *et al.* (2009) 'Guidelines for the nomenclature of the human heat shock proteins.', *Cell stress & chaperones*, 14(1), pp. 105–11. doi: 10.1007/s12192-008-0068-7.

Khan, S. *et al.* (2018) 'Reversible versus irreversible inhibition modes of ERK2: A comparative analysis for ERK2 protein kinase in cancer therapy', *Future Medicinal Chemistry*, 10(9), pp. 1003–1015. doi: 10.4155/fmc-2017-0275.

Kleinjung, T. *et al.* (2003) 'Heat shock protein 70 (Hsp70) membrane expression on head-and-neck cancer biopsy - A target for natural killer (NK) cells', *International Journal of Radiation Oncology Biology Physics*, 57(3), pp. 820–826. doi: 10.1016/S0360-3016(03)00629-1.

Knudson, A. G. (2001) 'Two genetic hits (more or less) to cancer', *Nature Reviews Cancer*, 1(2), pp. 157–162. doi: 10.1038/35101031.

Kroeger, H. (1960) 'The induction of new puffing patterns by transplantation of salivary gland nuclei into egg cytoplasma of Drosophila', *Chromosoma*, 11(1), pp. 129–145. doi: 10.1007/BF00328649.

Kumar, S. J. *et al.* (2016) 'Targeting Hsp70: A possible therapy for cancer', *Cancer Letters*. Elsevier Ireland Ltd, 374(1), pp. 156–166. doi: 10.1016/j.canlet.2016.01.056.

Lagoutte, R., Patouret, R. and Winssinger, N. (2017) 'Covalent inhibitors: an opportunity for rational target selectivity', *Current Opinion in Chemical Biology*, pp. 54–63. doi: 10.1016/j.cbpa.2017.05.008.

Li, G. C. and Werb, Z. (1982) 'Correlation between synthesis of heat shock proteins and development of thermotolerance in Chinese hamster fibroblasts.', *Proceedings of the National Academy of Sciences*, 79(10), pp. 3218–3222. doi: 10.1073/pnas.79.10.3218.

Liu, Q. *et al.* (2013) 'Developing irreversible inhibitors of the protein kinase cysteinome', *Chemistry and Biology*. doi: 10.1016/j.chembiol.2012.12.006.

Lonsdale, R. and Ward, R. A. (2018) 'Structure-based design of targeted covalent inhibitors', *Chemical Society Reviews*. doi: 10.1039/C7CS00220C.

Mah, R., Thomas, J. R. and Shafer, C. M. (2014) 'Drug discovery considerations in the development of covalent inhibitors', *Bioorganic and Medicinal Chemistry Letters*, pp. 33–39. doi: 10.1016/j.bmcl.2013.10.003.

Mahmudov, K. T. *et al.* (2017) 'Non-covalent interactions in the synthesis of coordination compounds: Recent advances', *Coordination Chemistry Reviews*, 345, pp. 54–72. doi: 10.1016/j.ccr.2016.09.002.

Mathew, A. and Morimoto, R. I. (1998) 'Role of the heat-shock response in the life and death of proteins', in *Annals of the New York Academy of Sciences*, pp. 99–111. doi: 10.1111/j.1749-6632.1998.tb08982.x.

Mathew, A. and MORIMOTO, R. I. (1998) 'Role of the Heat-Shock Response in the Life and Death of Proteins', *Annals of the New York Academy of Sciences*, 851(1 STRESS OF LIF), pp. 99–111. doi: 10.1111/j.1749-6632.1998.tb08982.x.

Mayer, M. P. and Bukau, B. (2005) 'Hsp70 chaperones: Cellular functions and molecular mechanism', *Cellular and Molecular Life Sciences*. Birkhäuser-Verlag, 62(6), pp. 670–684. doi: 10.1007/s00018-004-4464-6.

Nollen, E. A. A. *et al.* (1999) 'In Vivo Chaperone Activity of Heat Shock Protein 70 and Thermotolerance', *Molecular and Cellular Biology*, 19(3), pp. 2069–2079. doi: 10.1128/MCB.19.3.2069.

Nollen, E. A. A. and Morimoto, R. I. (2002) 'Chaperoning signaling pathways: molecular chaperones as stress-sensing "heat shock" proteins.', *Journal of cell science*, 115(Pt 14), pp. 2809–16.

Nylandsted, J. *et al.* (2000) 'Selective depletion of heat shock protein 70 (Hsp70) activates a tumor-specific death program that is independent of caspases and bypasses Bcl-2', *Proceedings of the National Academy of Sciences*, 97(14), pp. 7871–7876. doi: 10.1073/pnas.97.14.7871.

Nylandsted, J. *et al.* (2002) 'Eradication of glioblastoma, and breast and colon carcinoma xenografts by Hsp70 depletion.', *Cancer research*, 62(24), pp. 7139–42.

Palumbo, M. O. *et al.* (2013) 'Systemic cancer therapy: achievements and challenges that lie ahead', *Frontiers in Pharmacology*, 4, p. 57. doi: 10.3389/fphar.2013.00057.

Penkler, D. *et al.* (2017) 'Perturbation–Response Scanning Reveals Key Residues for Allosteric Control in Hsp70', *Journal of Chemical Information and Modeling*. American Chemical Society, 57(6), pp. 1359–1374. doi: 10.1021/acs.jcim.6b00775.

Pettinger, J., Jones, K. and Cheeseman, M. D. (2017) 'Lysine-Targeting Covalent Inhibitors', *Angewandte Chemie - International Edition*, 56(48), pp. 15200–15209. doi: 10.1002/anie.201707630.

Plummer, M. *et al.* (2016) 'Global burden of cancers attributable to infections in 2012: a synthetic analysis', *The Lancet Global Health*, 4(9), pp. e609–e616. doi: 10.1016/S2214-109X(16)30143-7.

Potashman, M. H. and Duggan, M. E. (2009) 'Covalent Modifiers: An Orthogonal Approach to Drug Design', *Journal of Medicinal Chemistry*. doi: 10.1021/jm8008597.

Powers, M. V. *et al.* (2010) 'Targeting HSP70: The second potentially druggable heat shock protein and molecular chaperone?', *Cell Cycle*, 9(8), pp. 1542–1550. doi: 10.4161/cc.9.8.11204.

Privalov, P. L. (1988) 'Hydrophobic Interactions in Proteins', *Protein Structure and Protein Engineering*, pp. 6–15. doi: 10.1007/978-3-642-74173-9_2.

Puhan, M. A. *et al.* (2017) 'Fisetin Acts on Multiple Pathways to Reduce the Impact of Age and Disease on CNS Function', 37(4), pp. 784–790. doi: 10.1183/09031936.00063810.The.

Rempel, B. P. and Withers, S. G. (2008) 'Covalent inhibitors of glycosidases and their applications in biochemistry and biology', *Glycobiology*, pp. 570–586. doi: 10.1093/glycob/cwn041.

Ruddraraju, K. V. *et al.* (2017) 'Covalent inhibition of protein tyrosine phosphatases', *Mol. BioSyst.*, 13(7), pp. 1257–1279. doi: 10.1039/C7MB00151G.

Samaras, V. *et al.* (2010) 'Chronic bacterial and parasitic infections and cancer: a review.', *Journal of infection in developing countries*, 4(5), pp. 267–81.

Schwartz, P. A. *et al.* (2014) 'Covalent EGFR inhibitor analysis reveals importance of reversible interactions to potency and mechanisms of drug resistance', *Proceedings of the National Academy of Sciences*, 111(1), pp. 173–178. doi: 10.1073/pnas.1313733111.

Serafimova, I. M. *et al.* (2012) 'Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles', *Nature Chemical Biology*, 8(5), pp. 471–476. doi: 10.1038/nchembio.925.

Sharma, D. and Masison, D. C. (2009) 'Hsp70 structure, function, regulation and influence on yeast prions.', *Protein and peptide letters*. NIH Public Access, 16(6), pp. 571–81. doi: 10.2174/092986609788490230.

Sharma, P. C. and Verma, R. (2018) 'Implication of HSP70 in the Pathogenesis of Gastric Cancer', in, pp. 113–130. doi: 10.1007/978-3-319-89551-2_6.

Shin, S. H., Bode, A. M. and Dong, Z. (2017) 'Addressing the challenges of applying precision oncology', *npj Precision Oncology*. Nature Publishing Group, 1(1), p. 28. doi: 10.1038/s41698-017-0032-z.

Singh, J. *et al.* (2011) 'The resurgence of covalent drugs', *Nature Reviews Drug Discovery*. Nature Publishing Group, 10(4), pp. 307–317. doi: 10.1038/nrd3410.

Smith, A. J. T. *et al.* (2009) 'Beyond picomolar affinities: quantitative aspects of non-covalent and covalent binding of drugs to proteins.', *Journal of medicinal chemistry*. NIH Public Access, 52(2), pp. 225–33. doi: 10.1021/jm800498e.

Strelow, J. M. (2017) ‘A Perspective on the Kinetics of Covalent and Irreversible Inhibition’, *Journal of Biomolecular Screening*, pp. 3–20. doi: 10.1177/1087057116671509.

Swinney, D. C. (2004) ‘Biochemical mechanisms of drug action: what does it take for success?’, *Nature Reviews Drug Discovery*. Nature Publishing Group, 3(9), pp. 801–808. doi: 10.1038/nrd1500.

Tissiéres, A., Mitchell, H. K. and Tracy, U. M. (1974) ‘Protein synthesis in salivary glands of Drosophila melanogaster: Relation to chromosome puffs’, *Journal of Molecular Biology*, 84(3), pp. 389–398. doi: 10.1016/0022-2836(74)90447-1.

Trepel, J. *et al.* (2010) ‘Targeting the dynamic HSP90 complex in cancer’, *Nature Reviews Cancer*, 10(8), pp. 537–549. doi: 10.1038/nrc2887.

Volloch, V. Z. and Sherman, M. Y. (1999) ‘Oncogenic potential of Hsp72’, *Oncogene*, 18(24), pp. 3648–3651. doi: 10.1038/sj.onc.1202525.

Voos, W. (2003) ‘A new connection: Chaperones meet a mitochondrial receptor’, *Molecular Cell*, 11(1), pp. 1–3. doi: 10.1016/S1097-2765(03)00002-9.

Wang, H. *et al.* (2014) ‘Heat Shock Proteins at the Crossroads between Cancer and Alzheimer’s Disease’, *BioMed Research International*, 2014, pp. 1–9. doi: 10.1155/2014/239164.

Whitley, D, Goldberg, S. P. and Jordan, W. D. (1999) ‘Heat shock proteins: a review of the molecular chaperones.’, *Journal of vascular surgery*, 29(4), pp. 748–51. doi: 10.1016/s0741-5214(99)70329-0.

Whitley, D., Goldberg, S. P. and Jordan, W. D. (1999) ‘Heat shock proteins: A review of the molecular chaperones’, *Journal of Vascular Surgery*, 29(4), pp. 748–751. doi: 10.1016/S0741-5214(99)70329-0.

WHO (2018) ‘WHO | Key facts about cancer’, *WHO*. World Health Organization.

Workman, P. (2004) ‘Altered states: selectively drugging the Hsp90 cancer chaperone’, *Trends in Molecular Medicine*, 10(2), pp. 47–51. doi: 10.1016/j.molmed.2003.12.005.

WORKMAN, P. *et al.* (2007) ‘Drugging the Cancer Chaperone HSP90: Combinatorial Therapeutic Exploitation of Oncogene Addiction and Tumor Stress’, *Annals of the New York Academy of Sciences*, 1113(1), pp. 202–216. doi: 10.1196/annals.1391.012.

Wu, B., Hunt, C. and Morimoto, R. (1985) ‘Structure and expression of the human gene encoding major heat shock protein HSP70.’, *Molecular and Cellular Biology*, 5(2), pp. 330–341. doi: 10.1128/MCB.5.2.330.

Wu, W. K. K. *et al.* (2010) ‘Macroautophagy modulates cellular response to proteasome inhibitors in cancer therapy.’, *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy*, 13(3), pp. 87–92. doi: 10.1016/j.drup.2010.04.003.

Yahara, I. (1996) ‘Stress-inducible cellular responses. Introduction.’, *EXS*, 77, pp. XI–XII. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8856965.

Young, J. C. *et al.* (2004) ‘Pathways of chaperone-mediated protein folding in the cytosol’, *Nature Reviews Molecular Cell Biology*, 5(10), pp. 781–791. doi: 10.1038/nrm1492.

Zhang, L. (2013) ‘The van der Waals force and gravitational force in matter’, pp. 1–14. Available at: http://arxiv.org/abs/1303.3579.

Zhang, W.-K. *et al.* (2008) ‘Occurrence of cancer at multiple sites: Towards distinguishing multigenesis from metastasis’, *Biology Direct*. BioMed Central, 3, p. 14. doi: 10.1186/1745-6150-3-14.

Zhanting Li, L.-Z. W. *et al.* (2015) *Hydrogen Bonded Supramolecular Structures, Lecture Notes in Chemistry*. doi: 10.1007/978-3-662-45756-6.

Zhao, Z. and Bourne, P. E. (2018) ‘Progress with covalent small-molecule kinase inhibitors’, *Drug Discovery Today*, pp. 727–735. doi: 10.1016/j.drudis.2018.01.035.

Zhu, K., Dunner, K. and McConkey, D. J. (2010) ‘Proteasome inhibitors activate autophagy as a cytoprotective response in human prostate cancer cells.’, *Oncogene*, 29(3), pp. 451–62. doi: 10.1038/onc.2009.343.

Zhuravleva, A., Clerico, E. M. and Gierasch, L. M. (2012) ‘An interdomain energetic tug-of-war creates the allosterically active state in Hsp70 molecular chaperones.’, *Cell*. Elsevier, 151(6), pp. 1296–307. doi: 10.1016/j.cell.2012.11.002.

CHAPTER 3

3. Computer-aided drug design and molecular modeling approaches for biomolecular structural analyses and selective therapeutic targeting

3.1 Introduction to molecular modeling

Molecular modeling approaches have made tremendous progress in past years and have become a necessary element of various biological and physio-chemical studies. There are two main objectives for this expansion: first, there has been an uncontrollable increase in the structural data available for proteins, from X-ray crystallography, electron microscopy and Nuclear Magnetic Resonance (NMR) studies. Second, convoying this development in the field of structural biology has approached remarkable improvements in computational methodology and hardware. As computers revive to enhance in speed and proficiency, we are able to handle more complex systems (Saxena *et al.*, 2009). Molecular modeling is one of the most promptly growing scientific fields by comprising a broad range of computational and theoretical tools, used to simulate small biologically active systems with the objective to understand their behavior at an atomistic level (Kore *et al.*, 2012).

Drug design and discovery is an important topic in pharmaceutical research as it is highly time consuming and cost-effective method to manufacture new drug compounds. There are several computational approaches that having critical impact in the area of computer aided drug design (CADD) over many years. These approaches can be divided into three basic fields: conformational modeling (complexes of macromolecules and small molecules), property modeling (biological, chemical and physical properties) and molecular design (to optimize these properties). Thus, computational approaches have given a significant hope to pharmaceutical industries to discover new potential drug targets which in turn imitate the time and success of implementing clinical trials for identifying new drugs (Baldi, 2010). Starting from 1960s, these computational approaches have served a strong platform for drug discovery and biomolecular structure analysis (Leach, 2001). Since the early 1980s, developments in the field of molecular biology, computational chemistry and protein crystallography have significantly aided Rational Drug Design (RDD) examples and the efficiency of their binding affinity estimations (Roy, Sengupta and De, 2001).

Most of the approaches utilized in drug designing were advanced for the reversible non-covalent inhibitors and have widely used everywhere. However, these techniques may also be functional for covalent inhibitors, special attentions or software with modified packages are necessary, to form the covalent bond between the substrate and ligand. Here, we proposed some molecular modeling approaches which are useful for covalent drug design and discovery (Lonsdale and Ward, 2018).

Some of the major techniques includes: Covalent docking, Covalent Molecular Dynamics simulation (CMDs), Molecular Mechanics (MM), Quantum Mechanics (QM) and Quantum Mechanics/Molecular Mechanics (QM/MM). These techniques are broadly used for rational drug discovery and design developments, to understand the molecular binding mechanism of the drug (Ramírez, 2016).

There are two fundamental molecular modeling principles that may be used to define the conformational and energetic modifications to drug-target complex:

- Quantum Mechanics

- Molecular Mechanics

By combining these two principles of molecular modeling with molecular dynamic simulations, flexibility of the target and landscape of the inhibitor binding can be studied (Lewars, 2010).

In this chapter, quantum mechanics, molecular mechanics and molecular dynamic simulations will be explained on, by providing insight into the rationale behind the chosen energy descriptors for this study. The principle behind each of the computational tools employed in the analysis will also be further described.

3.2 The principle of Quantum Mechanics

Quantum mechanics (QM) is the most successful field in physics. In the early 20s century, principle of QM has been developed and divided into two types. German scientists: Planck, Born, Jordan, Heisenberg and English-borne Dirac initiated the development of Matrix mechanics. Later, Erwin Schrödinger in 1926 developed Wave mechanics, that play a key role in the understanding of quantum phenomena (Trabesinger, 2009).

The theory of Quantum elucidates the behavioral features of sub-atomic particles, like nuclei and electrons (Atkins and Friedman, 2011). QM plays a significant role in different approaches in molecular biology, like formation/breaking of the bond, electron excitation and atomic transfer. Any property of a particular system in a 3D-space can be predicted by QM. Electrons are mapped by employing the continuous electron density approach and system energetic is calculated by using Schrödinger's theory of wave function. For complex systems, density of electrons may be estimated using the Born-Oppenheimer approximation theory (Cui, 2016; Shen, Wu and Yang, 2016)

3.2.1 The Schrödinger Wave Function Theory

In the early 1900s, Austrian physicist Erwin Schrödinger projected the quantum mechanical model of the atom. Elaborating on the Bohr's model of atoms which suggests that electrons are organized in concentric circular orbits nearby the nucleus, Schrödinger developed the mathematical equations to explain the probability of positioning an electron on the precise path. The model is represented as a nucleus which is bounded by a low- and high-density electron cloud. According to theory of quantum mechanics, all particles are explained as wave function with no described momentum or position unless they are perceived. The wave function determined the probability of every possible observation (Atkins and Friedman, 2011)(Leach, 2001).

The Schrödinger equation creates the basic core of QM, as Schrödinger himself observed that by accumulating the properties of an atom, being the mass and charge, to the equation, he was capable to predict the sequence of shapes presenting the wave pattern of electrons in an atom (Bahrami *et al.*, 2014).

The Schrödinger wave equation:

$$\mathbf{H\Psi = E\Psi} \qquad \text{(Eq 3.2.1)}$$

H is the Hamiltonian operator which contains derivatives with respect to the location of atom, E is the eigenvalues of energies in the system and Ψ is the wave function. To imitate an appropriate physical model of Schrödinger's equation, the wave function must be normalized,

single valued, continuous and anti-symmetric. H is the sum of atom's total potential energy (V) and kinetic energy (T):

$$\mathbf{H = T + V} \qquad \text{(Eq 3.2.2)}$$

Where H is defined as follows:

$$\mathbf{H} = \left[-\frac{\mathbf{h^2}}{\mathbf{8\pi^2}}\sum_{\mathbf{i}}\frac{\mathbf{1}}{\mathbf{m_j}}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\right]+\sum_{\mathbf{i}}\sum_{<\mathbf{j}}\left(\frac{\mathbf{e_i e_j}}{\mathbf{r_{ij}}}\right) \qquad \text{(Eq 3.2.3)}$$

The Schrödinger equation is extremely complicated when solving molecular systems as it may comprise thousands of atoms, thus proved to be non-executable (Barde *et al.*, 2015)(Nakatsuji, 2004). Although, another QM theory, The Born-Oppenheimer Approximation, recompenses for molecular structure rather than atomic structure.

3.2.2 The Born-Oppenheimer Approximation Theory

Born-Oppenheimer approximation theory has been proposed by the two physicists Max Born and J. Robert Oppenheimer in the year 1927. This theory explains about the uncoupling of the nuclei wave function to that of the electrons (Born and Oppenheimer, 1927). Electrons are taken to be of less weight than nuclei by having improved velocity and can move rapidly to nuclei movement. Distribution of electron inside a molecule is described by the position of the nuclei (Belkic *et al.*, 1986)(Liehr, 1957). This enables for the Schrödinger equation to be resolved for the kinetic energy of the electrons as this energy will remain constant for the nuclei. The variation in electron and nuclei's velocities allow Born-Oppenheimer to be applied by minimizing the complexity of the wave function of the Hamiltonian equation (Huang and Yi, 2009). Now the simplified wave function:

$$\mathbf{\Psi(r_{elec}) = \Psi(r_{elec})\ (\Psi(r_{ncul})} \qquad \text{(Eq 3.2.3)}$$

Equation 3.2.1 is converted into:

$$\mathbf{H_{EN}\Psi(r_{elec}) = E_{EN}\Psi(r_{elec})} \qquad \text{(Eq 3.2.4)}$$

H_{EN} symbolizes a modification of terms based activity to stabilize the nuclear position (V_{NN}). Equation 3.2.5 represents E_{EN}, initiated from two sources being the unstable electron co-ordinates and fixed co-ordinates of nuclear.

$$\mathbf{(H_{el} + V_{NN})\ \Psi\ (r_{el}) = E_{EN}\Psi(r_{el})} \qquad \text{(Eq 3.2.5)}$$

We use the Schrödinger equation to illustrate electronic motion within a molecule. The Approximation is believed to be more precise when applied to ground electronic states. After the equation has been resolved, stable positions of interest of the equilibrated conformation may be evaluated and the potential energy surface and curve can be formed (Jecko, 2014)(Matsika, 2010).

3.2.3 Potential Energy Surface as an Application of Quantum Mechanics

The potential energy surface is an efficient graphical/mathematical illustration among the molecular vibrational motions of a molecule, its geometry and its nuclear probability distribution by explaining the time-dependent Schrödinger's equation. The model of potential energy surface is based on the Born- Oppenheimer approximation theory, where the electrons fluctuate according to the positional states of the nuclei. Potential energy surface is further taken as the potential of an atoms movement to collide into each other within a molecule (Atkins and Friedman, 2011). A potential energy surface presents excessive regions of potential energy, denoting high-energy nuclear activities and low regions of energy denoting low nuclear energy conformations (Figure 3.1). This energy can be employed in computational chemistry to determine the lowest state of energy and the positional geometry of a molecule at this state (Jensen, 2007).

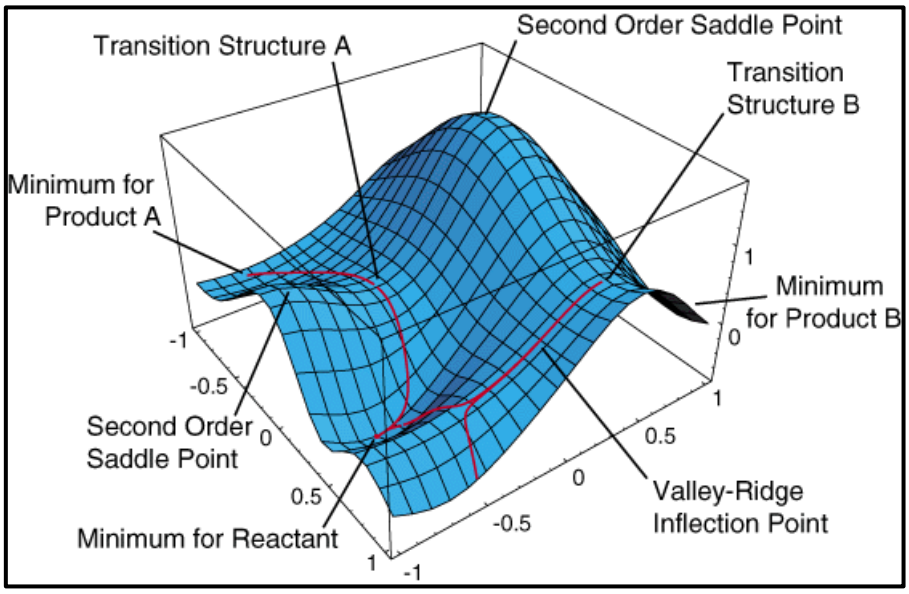


Figure 3.1: Graphical representation of a two-dimensional potential energy surface (PES) (Adapted from The California State University 2017).

3.3 The principle of Molecular Mechanics

One of the major difficulties in the field of chemistry is to be able to comprehend the chemical features of a molecule, for example, its solubility, stability and reactivity. The basic concept behind the molecular mechanics (MM) is that a molecular complex can be observed as a microscopic mechanical system. These molecular systems are connected by mechanical springs that regulate their angles, covalent bonds among successive bonds, rotations around the bonds and so on. Atoms repel or attract with each other due to the non-bonded classical potentials that identify the non-bonded interatomic energies. In mathematical expressions, this explanation needs the composition of a potential function integrating entirely conventional terms. This potential energy function is then utilized to calculate all the appropriate forces for the Newton's equation of motions, that finally explains the dynamics of the molecular system (Vanommeslaeghe, Guvench and MacKerell, 2014)(Lewars and Lewars, 2011).

In MM, simple algebraic terms are used to define the compounds total energy without requiring to calculate electron density or wave function as with QM (Tsai, 2002)(Davidson, 1993). Various methods are used in rational drug design which determines potentially suitable molecules before experimental analysis. MM simulations also enables the formation of atomistic models based on promising energy estimations (Poltev, 2015).

3.3.1 Potential Energy Function

In force filed methods, all atoms are categorized as the building blocks and electrons are not treated to be separate particles. This means that instead of solving the Schrödinger's equation, precise information about the bonds should be specified. In force field techniques, molecules are specified by a ball and spring model with distinctive atoms size and different length of bonds. It was seen that diverse molecules might have similarity in their structures because of the atoms they are made up of. The model was termed as "atom types" and is dependent on the number of atoms and chemical bonds holding it in place (Jensen, 2007; Tsai, 2002).

The sum of potential energy includes the extended sum of every single potential inter/intra molecular constituents, comprising:

1. Bond stretching (between directly bonded atoms)

$$\mathbf{E_r} = \sum \mathbf{K_r}(\mathbf{r} - \mathbf{r_0})^2 \qquad \text{(Eq 3.3.1.1)}$$

2. Angle bending (atoms surrounded to same central atoms)

$$\mathbf{E_\theta} = \sum \mathbf{K_\theta}(\boldsymbol{\theta} - \boldsymbol{\theta_0})^2 \qquad \text{(Eq 3.3.1.2)}$$

3. Bond torsion

$$\mathbf{E_\phi} = \sum \mathbf{K_\phi}[\mathbf{1} + \mathbf{cos}(\mathbf{n}\boldsymbol{\phi} - \boldsymbol{\phi_0})] \qquad \text{(Eq 3.3.1.3)}$$

4. Non-bonded interactions (electrostatic and van der Waals)

$$\mathbf{E_{nb}} = \left[\Sigma\sum\left(\frac{\mathbf{A_{ij}}}{\mathbf{r_{ij}}^{12}} - \frac{\mathbf{B_{ij}}}{\mathbf{r_{ij}}^{6}}\right)\right] + \left[\Sigma\sum\left(\frac{\mathbf{q_i q_j}}{\mathbf{D r_{ij}}}\right)\right] \qquad \text{(Eq 3.3.1.4)}$$

Where: K_r, K_θ, K_ϕ are force constants for bond, angle and dihedral angles and r_o, θ_o, ϕ_o are the equilibrium distance, angle and phase angle. r_{ij} is the parameter for distance and A_{ij} and B_{ij} are the parameters for van der Waals. D is the molecular dielectric constant; q_i and q_j are charge points. Atoms are usually treated as spheres and bonds as springs in MM. The characteristics

stated above are the simplest to explain mathematically where atoms are treated as spheres with characteristics radii.

The final functional equation for potential energy function is:

$$\mathbf{E_{total}} = \mathbf{E_r} + \mathbf{E_\theta} + \mathbf{E_\phi} + \mathbf{E_{nb}} \qquad (\text{Eq. } 3.3.2)$$

Currently, there are various types of force fields in use; although, they vary by the function of individual term of energy, the type of information used for parameters fitting and the number of cross terms. The two common types may be considered while designing the force fields:

1. Force fields implemented on big systems like protein or DNA, have a comparatively understandable function with no cross terms and use the Lennard-Jones potential as van der Waals energy. These are called diagonal/harmonic force fields.

2. Force fields implemented on medium to small size of systems, have to sustain more accuracy. These systems are having number of cross terms and use an exponential-type potential for van der Waals energy. These are called "class II" force fields.

The most extensively used force fields in biomolecular simulations are AMBER (Wang *et al.*, 2004), GROMOS (Hermans *et al.*, 1984), OPLS-AA (Damm *et al.*, 1997), CHARMM (Brooks *et al.*, 2009), ENCAD (Levitt *et al.*, 1995) and NAMD (Phillips *et al.*, 2005).

In this study, the AMBER force field (Wang *et al.*, 2004) was used for the defining the molecular systems.

3.4 The Hybrid Quantum Mechanics/Molecular Mechanics (QM/MM)

For modeling the non-covalent interactions in large complexes, MM based approaches are useful like protein-ligand systems. Although, this approach is not fully implemented on the covalent system as it does not allow the reactivity effects. QM techniques are necessary in order to suitably model chemical reactions; however, these techniques are usually restricted to small model systems.

A well-known approach for large systems is to use the hybrid quantum mechanics/molecular mechanics approach (QM/MM). In this method, small number of atoms such as the ligand and target residue are employed with the QM method, enabling the chemical reaction to be formed.

The other region of the substrate and solvent are used with the MM method. The benefit of the QM/MM method over the common QM method is that the force of the surrounding solvent and protein is contained their geometry and effects based on the local electrostatic environment. QM/MM methods may be utilized to reveal mechanistic aspects of enzymatic reactions that can be useful in designing covalent inhibitors (Lonsdale and Ward, 2018). However, QM/MM is widely used and proven to be very efficient in simulating systems where energy profiles, reaction pathways, electronic contributions, chemical reactivity in biomolecular systems are being studied. Nevertheless, it is also highly cost effective and still the high-level description of QM region remains out of reach. Also, the free energy calculations of QM/MM approach require extensive configurational sampling. Therefore, for biomolecular simulations where long MD is required, such as conformational analysis, protein folding etc., QM/MM would not serve as a feasible choice. For this type of study, where a long MD simulation is required, all-atom MM covalent approach is the most feasible and appropriate. It has been a challenge in literature to perform all-atom covalent simulations and that is truly one of the biggest strengths in our current report besides. Therefore, we present a less costly approach for describing irreversible covalent interactions (Khan, Bjij, Fisayo, *et al.*, 2018).

3.5 The Principle of Molecular Dynamic Simulations

Molecular dynamics (MD) first started in the late 1950s, where Rahman, Alder and Wainwright established the techniques for the dynamics of liquids. The domain of computational chemistry has advanced since then. From 1970s, the method has developed from simulating hundreds to thousands atoms to systems containing macromolecules like DNA or protein (Hospital *et al.*, 2015). There are mainly two types of simulation methods:

- Classical MD
- Monte Carlo (MC)

In past years, many hybrid approaches have also been launched. One of the leading benefit of MD over MC is that, it enables the dynamical characteristics of a system like time-dependent reactions and rheological properties (Liu *et al.*, 2018). MD is specifically useful in molecular biology and biochemistry as on atom scale, it allows to determine and classify the dynamic event which can affect the biological properties of a system (Jarosaw Meller, 2001).

Conventional/classical MD includes Newton's equations of motions into computational algorithm. On atomistic level, MD simulations offers high probability real-time mechanistic

and conformational interpretations of different chemical reactions which are based on the highly advanced physical and mathematical algorithm (Figure 3.2) (Vitalini *et al.*, 2015)(Monticelli and Tieleman, 2013).

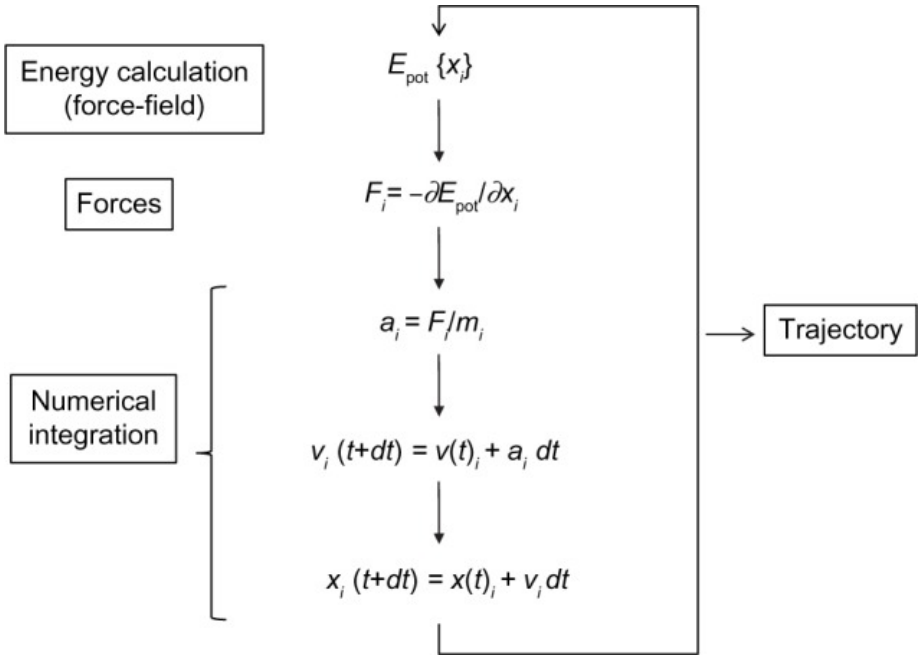


Figure 3.2 Basic algorithm of Molecular Dynamics. Where E_{pot} = potential energy; t = simulation time; dt = iteration time; x = tom co-ordinates; F = force component; a = acceleration; m = atom mass and v = velocity (Image adapted from (*Hospital et al.*, 2015)***).***

The MD simulation allows analyzing the particles interacting in the system throughout a required time-period, by generating a dynamical trajectory which can be further explored. The whole objective of this computational approach is to use the Newton's equations to study and understand the structural dynamics and energies of a biological system. The below mentioned conditions for initial particles are needed:

1. A suitable force field to describe the forces between atoms such as AMBER or CHARMM.

2. Velocities and positions of individual particles.

3. Conditions for the boundary that need to be employed.

The classical equation of motion can be solved then:

$$\mathbf{F_i} = \mathbf{m_i} \frac{\mathbf{d^2 r_i(t)}}{\mathbf{dt^2}} \qquad \text{(Eq 3.4)}$$

Where $r_i(t)$ is the particle position vector, t is time-evolution, m is the mass of the particle and F_i represents the interacting force on the particle.

MD can be divided into four continuous technical steps which may be repetitive millions of types to produce a trajectory (Haile *et al.*, 1993). The steps are as follows:

1. The basic condition of the biomolecular system are explained as:

    - The co-ordinates of each atom

    - Bond characteristics between each atom

    - Acceleration of atoms

2. Every atom's potential energy is calculated.

3. The potential energies from step 2 are then used to study the equations of motions.

4. The new "state" of the system must be saved, and the co-ordinates of the atoms should be changed and step forward in the simulation in taken. The process then starts back from step 1.

Once the trajectory is completely produced, quantitative evaluation of the system's time-evolution can progress.

3.5.1 Covalent molecular dynamic simulations

In silico techniques can explain the relationship between covalent inhibition and produced conformational changes, offers a more accurate approach for investigating covalent systems contrary to classical simulation techniques. Covalent Molecular Dynamic simulations (CMDs) is a technical modification of the MD method which allows to overcome the disputes accompanied with the parameterization of a covalent bond. Therefore, we proposed the first standard, accurate and reliable protocol for the simulations of covalent systems (Khan, Bjij, Fisayo, *et al.*, 2018). This enables the simultaneous rationalization of covalent bond formation and corresponding changes in the conformation of the system. This method specifically illustrates the interface between target protein and covalent inhibitors. This method has been

positively validated by few of our recent published studies (Khan, Bjij, Betz, *et al.*, 2018; Khan, Bjij, Fisayo, *et al.*, 2018; Bjij *et al.*, 2018). It includes the primary retrieval of the suitable ligand and proteins, system preparation and the final step of MD simulation using all atom MM approach. Detailed steps for studying covalent systems are mentioned below:

3.5.2 Specific criteria for the selection of ligands and proteins

Although there are numerous criteria to select the ligand and their respective target proteins but not all are appropriate for studying the covalent systems, therefore it is essential to focus on the characteristics and basis of selection. The choice of ligand and protein is depending on the existence of reactive moieties that can be electrophilic and sometimes nucleophilic in nature. Reactive moieties are described according to their compatibility to either donate (nucleophiles) or receive (electrophiles) electrons, which results in the covalent bond formation. The process of covalent bond formation is known as Michael addition that happens between a Michael donor and a Michael acceptor. The common residues involved in the formation of covalent bond are cysteine (mostly), lysine, tyrosine, threonine, arginine, histidine, aspartate and glutamate. These residues are selective for covalent inhibition in their un-protonated states. The most appropriate residues are the ones that positioned in the catalytic site of the proteins, however the covalent inhibition of adenosine triphosphate (ATP) binding site and allosteric site residues are also gaining interest. These criteria have been utilized in the covalent inhibition of numerous proteins involved in disease pathogenesis (Liu and Rost, 2003) (Punthasee *et al.*, 2017; Ruddraraju *et al.*, 2017).

3.5.3 Molecular Dynamics Post-Analysis

Trajectories of MD are generated from the production run of the simulation. These trajectories can be described as snapshots of the sequences which are considered by velocity vectors and positional co-ordinates and report the time evolution of the system in phase space (Likhachev, Balabaev and Galzitskaya, 2016)(Jarosaw Meller, 2001).

While selecting the analytical software, the below specifications are necessary:

1. Software for qualitative visualization that will generate high quality images/snapshots and also show the video clips for trajectory analysis.

2. The software must have fast processors that can entertain data with large volume.

3. The software should have variety of options for analysis.

The preferred post-dynamic approaches and calculations should be related to the type of the MD study; although, specific quantitative calculation is required to support any visual systemization.

For the purpose of this type of investigation, the post-dynamic analysis of the trajectories is important to observe the:

- Conformational stability and energetics of the biomolecular system,

- Dynamics or variability of the biomolecular system,

- Appearances of the system's small molecule binding landscape and the thermodynamic fluctuations of energies with the clustered trajectory.

3.5.3.1 Stability of the System

Convergence:

This term can be utilized to explain dynamics of the protein based on types of bonds and bond angle vibrations at the time of protein unfolding. This integration for equilibrium and the illustration of a conformational plateau and final energetic is necessary for a MD trajectory to be more precise and reproducible (Amadei, Ceruso and Di Nola, 1999). At this plateau the ligand-protein system is presented to show energetic constant conformations.

Root Mean Square Deviation (RMSD):

The divergence of a system can be evaluated by the spatial modification between two static structures of the uniform trajectory. The RMSD of a trajectory is described as:

$$\mathbf{RMSD} = \left(\frac{\sum_{\mathbf{N}}(\mathbf{R_i} - \mathbf{R_1^0})^2}{\mathbf{N}}\right)^{\frac{1}{2}} \qquad \text{(Eq 3.5)}$$

Here, N is the total number of atoms in the system, R_i is the vector position of the Cα atom of particle i in the reference conformation that is calculated after aligning the structure to the primary conformation (O) using the least square fitting.

The average RMSD can be computed by taking the average of the frame numbers in each trajectory and can be calculated for the ligand, receptor and complex of a system (Orry and Abagyan, 2012; Forster, 2002).

Radius of Gyration (RoG):

In a protein, the radius of gyration can be described as the root mean square distance of the atoms from their center of gravity. This enables for the calculation of a compactness of a protein system with the trajectory. The RoG of a system can be based on the below mentioned equation:

$$\mathbf{r^2gyr} = \frac{(\sum_{i-1}^{n} \mathbf{w_i} (\mathbf{r_i} - \mathbf{r^-})^2)}{\sum_{i-1}^{n} \mathbf{w_i}} \qquad \text{(Eq 3.6)}$$

Where, r_i is the position of the ith atom and r is the center weight of atom i. The average RoG can be computed by taking the average frame number in a trajectory (Lobanov, Bogatyreva and Galzitskaya, 2008).

3.5.3.2 Conformational fluctuations of system

Root Mean Square Fluctuation (RMSF):

RMSF of a protein evaluates the residual Cα atom fluctuations based on the average structure of the protein with system's trajectory. This further extends to hypothesize the flexible regions of a protein based on the calculated RMSF (Bornot, Etchebest and De Brevern, 2011).

The below equation is applied to calculate the RMSF:

$$\mathbf{sRMSF} = \frac{(\mathbf{RMSF_i} - \overline{\mathbf{RMSF}})}{\boldsymbol{\sigma}\,(\mathbf{RMSF})} \qquad \text{(Eq 3.8)}$$

Where, $RMSF_i$ is the RMSF of the i_{th} residue, from which average RMSF is subtracted. This is further divided by the RMSF's standard deviation to achieve the resultant standardized RMSF.

This technique is different from RMSD and RoG as it is calculated as the total residual fluctuation with the trajectory and is not evaluated at each frame in the trajectory.

3.5.3.3 Free energy binding landscape

Free binding energy calculations is an essential technique that may explain the mechanism of ligand-protein binding involving the contribution of entropy and enthalpy (Ylilauri and Pentikäinen, 2013). Binding of free energy estimation leads to advancement of different algorithms and methods comprising thermodynamic integration, molecular energy calculations, free energy perturbation and linear integration energy etc. From all the free energy calculations, Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) and Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) approaches are the most efficient and accurate in computing binding free energies for biological macromolecules. Opposing to molecular docking, MM/GBSA and MM/PBSA do not account on a big data set to explain diverse parameters in each energy term. The above mentioned approaches utilize the combination of MM in terms of solvent model to calculate the absolute binding energy which is a average number of frames in the trajectory (Rastelli *et al.*, 2010). The free binding energy (ΔG) estimated by these techniques for a protein system like ligand, receptor and complex can be illustrated as:

$$\Delta G_{bind} = G_{complex} - G_{receptor} - G_{ligand} \quad \text{(Eq 3.7.1)}$$

$$\Delta G_{bind} = E_{gas} + G_{sol} - TS \quad \text{(Eq 3.7.2)}$$

$$E_{gas} = E_{int} + E_{vdw} + E_{ele} \quad \text{(Eq 3.7.3)}$$

$$G_{sol} = G_{GB/PB} + G_{SA} \quad \text{(Eq 3.7.4)}$$

$$G_{SA} = \gamma SASA \quad \text{(Eq 3.7.5)}$$

Here, E_{gas} symbolizes gas-phase energy consisting the internal energy E_{int}; Coulomb energy E_{ele} and the van der Waals energy E_{vdw}. The E_{gas} was calculated from the FF14SB. G_{sol} is the solvation free energy that was computed from the polar states, non-polar states and $G_{GB/PB}$ energy contribution. G_{SA} is the non-polar solvation energy. G_{SA} was calculated from the solvent accessible surface area (SASA) by utilizing a water probe radius of 1.4 Å, while the polar solvation $G_{GB/PB}$ involvement was calculated by solving the equation of GB/PB. T and S signify the total entropy of the temperature and solute respectively. The MM/GBS and MM/PBSA algorithms propose quantifiable analysis of the binding affinity of the inhibitor to the protein and thus enables to explain molecular docked structures (Suenaga *et al.*, 2012).

3.5.3.4 Dynamic cross correlation matrix of a system

Dynamic Cross Correlation Matrix (DCCM) is used to plot the fluctuations between the residues either in or out of phase throughout the simulation. The cross-correlation constant differs from +1 (completely correlated movements) to -1 (completely anti-correlated movements).

The equation used to explain the DCCM is:

$$\mathbf{C_{ij}} = \frac{< \mathbf{\Delta ri}.\mathbf{\Delta rj} >}{\left(< \mathbf{\Delta r_i^2} >< \mathbf{\Delta r_j^2} >\right)^{\frac{1}{2}}} \qquad (Eq\ 3.9)$$

The cross-correlation coefficient (C_{ij}) differs within a value of -1 to +1 and the higher and lower limit resembles to a fully correlated and anti-correlated movements throughout the simulation process.

Where, i and j denotes ith and jth residues respectively and Δr_i or Δr_j symbolizes displacement vectors relate to ith and jth residues. Dynamic cross correlation maps have become very effective in calculating residual movements that arise from ligand binding or in the occurrence of protein mutations (Tiberti, Invernizzi and Papaleo, 2015)(Kasahara, Fukuda and Nakamura, 2014).

3.6 Additional Computer-Aided Drug Design techniques utilized in the study

3.6.1 Covalent docking

Covalent docking has become a valuable computational tool applied in structure-based design of covalent inhibitors. Its application gained prominence upon the recent resurgence of covalent inhibitors. The traditional docking techniques compute the binding modes of non-covalent inhibitors. These techniques are restricted because they are not able to deal with the formation of covalent bond. Steric repulsion among the reactive residues and covalent moiety may inhibit the appropriate docking poses from being recognized. Currently, there are numerous docking tools and servers available for covalent docking. The accurate functional group has been implicit into the docking software and then docking poses are traced in which the covalent bond between ligand and receptor has been formed. This unique process of covalent bond formation makes the covalent docking technique different from the conventional docking

methods. Some of the user friendly docking tools and packages that successfully execute covalent docking involve: AutoDock (Morris *et al.*, 2009), CovalentDock (Ouyang *et al.*, 2013), CovDock (Zhu *et al.*, 2014), FlexX (Rarey *et al.*, 1996) and Dockovalent (London *et al.*, 2014). Although, few of these software can hardly apply in the virtual screening of large libraries of molecules, as they need a manual definition of the covalent bond. Covalent docking has been efficiently implemented in various studies as a validation of its importance as an essential tool in computer aided covalent drug design (Khan, Bjij, Fisayo, *et al.*, 2018) (Lonsdale and Ward, 2018).

The utilized computer-aided drug design techniques have been performed in this Study as the following chapters show where the Molecular modeling approaches have been carried out to exhibit the crucial result of this Study.

References

Amadei, A., Ceruso, M. A. and Di Nola, A. (1999) 'On the convergence of the conformational coordinates basis set obtained by the Essential Dynamics analysis of proteins' molecular dynamics simulations', *Proteins: Structure, Function and Genetics*, 36(4), pp. 419–424. doi: 10.1002/(SICI)1097-0134(19990901)36:4<419::AID-PROT5>3.0.CO;2-U.

Atkins, P. W. and Friedman, R. (2011) 'Molecular Quantum Mechanics', *Quantum*, 134, p. 588.

Bahrami, M. *et al.* (2014) 'The Schrödinger-Newton equation and its foundations', *New Journal of Physics*, 16. doi: 10.1088/1367-2630/16/11/115007.

Baldi, A. (2010) 'Computational approaches for drug design and discovery: An overview', *Systematic Reviews in Pharmacy*, 1(1), p. 99. doi: 10.4103/0975-8453.59519.

Barde, N. P. *et al.* (2015) 'Deriving time dependent Schrödinger equation from Wave-Mechanics, Schrödinger time independent equation, classical and Hamilton-Jacobi equations', *Leonardo Electronic Journal of Practices and Technologies*, 14(26), pp. 31–48.

Belkic, D. *et al.* (1986) 'The first Born approximation for charge transfer collisions', *Journal of Physics B: Atomic and Molecular Physics*, 19(18), pp. 2945–2953. doi: 10.1088/0022-3700/19/18/023.

Bjij, I. *et al.* (2018) 'Exploring the Structural Mechanism of Covalently Bound E3 Ubiquitin Ligase: Catalytic or Allosteric Inhibition?', *Protein Journal*. Springer US, 37(6), pp. 500–509. doi: 10.1007/s10930-018-9795-5.

Born, M. and Oppenheimer, J. R. (1927) 'Born-Oppenheimer approximation', *Ann. Phys. (Leipzig)*, 84(may), p. 457. doi: 10.1007/978-3-662-44593-8.

Bornot, A., Etchebest, C. and De Brevern, A. G. (2011) 'Predicting protein flexibility through the prediction of local structures', *Proteins: Structure, Function and Bioinformatics*, 79(3), pp. 839–852. doi: 10.1002/prot.22922.

Brooks, B. R. *et al.* (2009) 'CHARMM: The biomolecular simulation program', *Journal of Computational Chemistry*. Wiley Subscription Services, Inc., A Wiley Company, 30(10), pp. 1545–1614. doi: 10.1002/jcc.21287.

Cui, Q. (2016) 'Quantum mechanical methods in biochemistry and biophysics', *The Journal of Chemical Physics*, 145(14), p. 140901. doi: 10.1063/1.4964410.

Damm, W. *et al.* (1997) 'OPLS all-atom force field for carbohydrates', *Journal of Computational Chemistry*, 18(16), pp. 1955–1970. doi: 10.1002/(SICI)1096-987X(199712)18:16<1955::AID-JCC1>3.0.CO;2-L.

Davidson, E. R. (1993) 'Molecular Mechanics and Modeling: Overview', *Chemical Reviews*, p. 2337. doi: 10.1021/cr00023a600.

Forster, M. J. (2002) 'Molecular modelling in structural biology.', *Micron (Oxford, England : 1993)*, 33(4), pp. 365–384. doi: Pii S096804328(01)00035-X\rDoi 10.1016/S0968-

4328(01)00035-X.

Haile, J. M. *et al.* (1993) 'Molecular Dynamics Simulation: Elementary Methods', *Computers in Physics*, 7(6), p. 625. doi: 10.1063/1.4823234.

Hermans, J. *et al.* (1984) 'A consistent empirical potential for water–protein interactions', *Biopolymers*, 23(8), pp. 1513–1518. doi: 10.1002/bip.360230807.

Hospital, A. *et al.* (2015) 'Molecular dynamics simulations: Advances and applications', *Advances and Applications in Bioinformatics and Chemistry*, 8, pp. 37–47. doi: 10.2147/AABC.S70333.

Huang, X. L. and Yi, X. X. (2009) 'Born-Oppenheimer approximation in open systems', *Physical Review A*, 80(3), p. 7. doi: 10.1103/PhysRevA.80.032108.

Jarosaw Meller (2001) 'Molecular Dynamics', *Encyclopedia of Life Sciences*, pp. 1–8. doi: 10.1021/jp909004y.

Jecko, T. (2014) 'On the mathematical treatment of the born-Oppenheimer approximation', *Journal of Mathematical Physics*, 55(5). doi: 10.1063/1.4870855.

Jensen, F. (2007) *Introduction to Computational Chemistry*, *Angewandte Chemie International Edition*. doi: 10.1007/s00214-013-1372-6.

Kasahara, K., Fukuda, I. and Nakamura, H. (2014) 'A novel approach of dynamic cross correlation analysis on molecular dynamics simulations and its application to Ets1 dimer-DNA complex', *PLoS ONE*, 9(11). doi: 10.1371/journal.pone.0112419.

Khan, S., Bjij, I., Fisayo, O., *et al.* (2018) 'Covalent Simulations of Covalent/Irreversible Enzyme Inhibition in Drug Discovery – A Reliable Technical Protocol', *Future medicinal chemistry*, p. Accepted.

Khan, S., Bjij, I., Betz, R. M., *et al.* (2018) 'Reversible versus irreversible inhibition modes of ERK2: A comparative analysis for ERK2 protein kinase in cancer therapy', *Future Medicinal Chemistry*, 10(9), pp. 1003–1015. doi: 10.4155/fmc-2017-0275.

Kore, P. P. *et al.* (2012) 'Computer-Aided Drug Design: An Innovative Tool for Modeling', *Open Journal of Medicinal Chemistry*, 02(04), pp. 139–148. doi: 10.4236/ojmc.2012.24017.

Leach, A. R. (2001) 'Molecular Modelling: Principles and applications', in *Molecular Modelling: Principles and applications*, pp. 26–161. doi: 10.1016/S0097-8485(96)00029-0.

Levitt, M. *et al.* (1995) 'Potential energy function and parameters for simulations of the molecular dynamics of proteins and nucleic acids in solution', *Computer Physics Communications*, 91(1–3), pp. 215–231. doi: 10.1016/0010-4655(95)00049-L.

Lewars, E. G. (2010) *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics*, *Computational Chemistry*. doi: 10.1017/CBO9781107415324.004.

Lewars, E. G. and Lewars, E. G. (2011) 'Molecular Mechanics', *Computational Chemistry*, pp. 45–83. doi: 10.1007/978-90-481-3862-3.

Liehr, A. D. (1957) ‘On the use of the Born-Oppenheimer approximation in molecular problems’, *Annals of Physics*, 1(3), pp. 221–232. doi: 10.1016/0003-4916(57)90009-X.

Likhachev, I. V., Balabaev, N. K. and Galzitskaya, O. V. (2016) ‘Available Instruments for Analyzing Molecular Dynamics Trajectories’, *The Open Biochemistry Journal*, 10(1), pp. 1–11. doi: 10.2174/1874091X01610010001.

Liu, J. and Rost, B. (2003) ‘Domains, motifs and clusters in the protein universe’, *Current Opinion in Chemical Biology*, 7(1), pp. 5–11. doi: 10.1016/S1367-5931(02)00003-0.

Liu, X. *et al.* (2018) ‘Expert Opinion on Drug Discovery Molecular dynamics simulations and novel drug discovery Molecular dynamics simulations and novel drug discovery’, *Expert Opinion on Drug Discovery*, 13(1), pp. 23–37. doi: 10.1080/17460441.2018.1403419.

Lobanov, M. Y., Bogatyreva, N. S. and Galzitskaya, O. V. (2008) ‘Radius of gyration as an indicator of protein structure compactness’, *Molecular Biology*, 42(4), pp. 623–628. doi: 10.1134/S0026893308040195.

London, N. *et al.* (2014) ‘Covalent docking of large libraries for the discovery of chemical probes’, *Nature Chemical Biology*, 10(12), pp. 1066–1072. doi: 10.1038/nchembio.1666.

Lonsdale, R. and Ward, R. A. (2018) ‘Structure-based design of targeted covalent inhibitors’, *Chemical Society Reviews*. doi: 10.1039/C7CS00220C.

Matsika, S. (2010) ‘The Born-Oppenheimer approximation’, *The Journal of chemical physics*, 133(22), p. 224103. doi: 10.1063/1.3511638.

Monticelli, L. and Tieleman, D. P. (2013) ‘Force fields for classical molecular dynamics’, *Biomolecular Simulations: Methods and Protocols*, 924(January 2013), p. 617. doi: 10.1007/978-1-62703-017-5.

Morris, G. M. *et al.* (2009) ‘AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility.’, *Journal of computational chemistry*. NIH Public Access, 30(16), pp. 2785–91. doi: 10.1002/jcc.21256.

Nakatsuji, H. (2004) ‘Scaled Schrodinger equation and the exact wave function’, *Physical Review Letters*, 93(3), pp. 030403–1. doi: 10.1103/PhysRevLett.93.030403.

Orry, A. J. W. and Abagyan, R. (2012) *Methods in Molecular Biology: Homology Modelling*, *Humana Press*. doi: 10.1007/978-1-62703-239-1_1.

Ouyang, X. *et al.* (2013) ‘CovalentDock: Automated covalent docking with parameterized covalent linkage energy estimation and molecular geometry constraints’, *Journal of Computational Chemistry*, 34(4), pp. 326–336. doi: 10.1002/jcc.23136.

Phillips, J. C. *et al.* (2005) ‘Scalable molecular dynamics with NAMD’, *Journal of Computational Chemistry*, 26(16), pp. 1781–1802. doi: 10.1002/jcc.20289.

Poltev, V. (2015) *Handbook of Computational Statistics, American Chemical Society*. doi: 10.1198/tech.2005.s314.

Punthasee, P. *et al.* (2017) ‘Covalent Allosteric Inactivation of Protein Tyrosine Phosphatase

1B (PTP1B) by an Inhibitor-Electrophile Conjugate', *Biochemistry*, 56(14), pp. 2051–2060. doi: 10.1021/acs.biochem.7b00151.

Ramírez, D. (2016) 'Computational Methods Applied to Rational Drug Design', *The Open Medicinal Chemistry Journal*, 10(1), pp. 7–20. doi: 10.2174/1874104501610010007.

Rarey, M. *et al.* (1996) 'A fast flexible docking method using an incremental construction algorithm', *Journal of Molecular Biology*, 261(3), pp. 470–489. doi: 10.1006/jmbi.1996.0477.

Rastelli, G. *et al.* (2010) 'Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA', *Journal of Computational Chemistry*, 31(4), pp. 797–810. doi: 10.1002/jcc.21372.

Roy, K., Sengupta, C. and De, A. U. (2001) 'Theoretical aspects of rational drug design - An overview', *Journal of Scientific and Industrial Research*, 60(9), pp. 699–716.

Ruddraraju, K. V. *et al.* (2017) 'Covalent inhibition of protein tyrosine phosphatases', *Mol. BioSyst.*, 13(7), pp. 1257–1279. doi: 10.1039/C7MB00151G.

Saxena, A. *et al.* (2009) 'The Basic Concepts of Molecular Modeling', *Methods in Enzymology*, pp. 307–334. doi: 10.1016/S0076-6879(09)67012-9.

Shen, L., Wu, J. and Yang, W. (2016) 'Multiscale Quantum Mechanics/Molecular Mechanics Simulations with Neural Networks', *Journal of Chemical Theory and Computation*, 12(10), pp. 4934–4946. doi: 10.1021/acs.jctc.6b00663.

Suenaga, A. *et al.* (2012) 'An efficient computational method for calculating ligand binding affinities', *PLoS ONE*, 7(8). doi: 10.1371/journal.pone.0042846.

Tiberti, M., Invernizzi, G. and Papaleo, E. (2015) '(Dis)similarity Index to Compare Correlated Motions in Molecular Simulations', *Journal of Chemical Theory and Computation*, 11(9), pp. 4404–4414. doi: 10.1021/acs.jctc.5b00512.

Trabesinger, A. (2009) 'History of quantum theory: The short version', *Nature Physics*, p. 383. doi: 10.1038/nphys1279.

Tsai, C. S. (2002) 'Molecular modeling: molecular mechanics', in *An introduction to Computational Biochemistry*, pp. 285–314.

Vanommeslaeghe, K., Guvench, O. and MacKerell, A. D. (2014) 'Molecular Mechanics', *Current Pharmaceutical Design*, 20(20), pp. 3281–3292. doi: 10.2174/13816128113199990600.

Vitalini, F. *et al.* (2015) 'Dynamic properties of force fields Dynamic properties of force fields', *Journal of Chemical Physics*, 142, pp. 1–12. doi: 10.1063/1.4909549.

Wang, J. *et al.* (2004) 'Development and Testing of a General Amber Force Field', *Journal of Computational Chemistry*, 25, pp. 1157–1174.

Ylilauri, M. and Pentikäinen, O. T. (2013) 'MMGBSA as a tool to understand the binding affinities of filamin-peptide interactions', *Journal of Chemical Information and Modeling*, 53(10), pp. 2626–2633. doi: 10.1021/ci4002475.

Zhu, K. *et al.* (2014) 'Docking covalent inhibitors: A parameter free approach to pose prediction and scoring', *Journal of Chemical Information and Modeling*, 54(7), pp. 1932–1940. doi: 10.1021/ci500118s.

CHAPTER 4

Covalent versus Non-covalent Enzyme Inhibition: Which Route should we take? A Justification of the Good and Bad from Molecular Modelling perspective

Aimen Aljoundi[1], Imane Bjij[1,2], Ahmed El Rashedy[1], and Mahmoud E. S. Soliman[1*]

[1] Molecular Bio-Computation & Drug Design Lab, School of Health Sciences, University of KwaZulu-Natal, Westville, Durban 4000, South Africa.

[2] Université Cadi Ayyad, Faculté des Sciences Semlalia, Département de Chimie, Av. My Abdellah, BP2390 Marrakech, Morocco.

[*] Corresponding author: Mahmoud E. Soliman, email: soliman@ukzn.ac.za

Telephone: +27 31 260 8048, Fax: +27 31 260 7872

Abstract

The pace and efficiency of drug target strategies have been emanating debates among researchers in the field of drug development. Covalent inhibitors possess significant advantages over non-covalent inhibitors, such that covalent warheads can target rare residues of a particular target protein, thus leading to the development of highly selective inhibitors. However, toxicity can be a real challenge related to this class of therapeutics. From the challenges of irreversible drug toxicity to the declining reactivity of reversible drugs, herein we provide justifications from the computational point of view. It was evident that both classes had its merits, however, with the increase in drug resistance, covalent inhibition seemed more suitable. There also seems to be enhanced selectivity of the covalent systems, proving its use as a therapeutic regimen worldwide. We believe that this study will assist researchers in making informed decisions on which drug class to choose as lead compounds in the drug discovery pipeline.

Keywords: Covalent inhibition; Non-covalent inhibition; Merits; Pitfalls

3.1 Introduction

The attenuation of biological effects of protein through the binding of small molecules defines the general process of its inhibition. The common mechanism of this process depends on the equilibrium of reaction between the functional groups of the drug and the active site residues of the protein. The created equilibrium between the bound and unbound states of the complex is a result of some or all the interactions such as, hydrogen bonds, dipole-dipole interactions, van der Waals interactions, ionic bonds and London dispersion interactions. In addition to these non-covalent interactions, the proteins can be covalently modified in the presence of certain chemical moieties that form a covalent bond with the appropriate amino acid in the protein. Creating this covalent bond in the ligand-receptor complex, establishes a re-emerging class of drugs in drug discovery called covalent drugs [1].Covalent inhibition is a rapidly evolving paradigm in the area of drug discovery as a validated approach to avoid acquired resistance in malignancy. The process of evolving small molecules to inhibit the target receptor covalently has become an important task in the drug discovery field [2]. In the late 1800s, Aspirin (acetylsalicylic acid) was manufactured by the Bayer company to treat inflammation and pain by covalently inhibiting of cyclooxygenase (COX)-1 and -2 enzymes. This further paved the way for the development of other covalent drugs, which have evolved over time [3]. Figure 4.1 illustrates a timeline of the covalent drug evolution from the last two centuries. Covalent bonds are usually formed by the interaction between a nucleophilic cysteine, serine, threonine or rarely, lysine, and a reactive functional group of the ligand such as hydroxyl, epoxy or carbonyl, thus leading to the formation of covalent adducts [4].

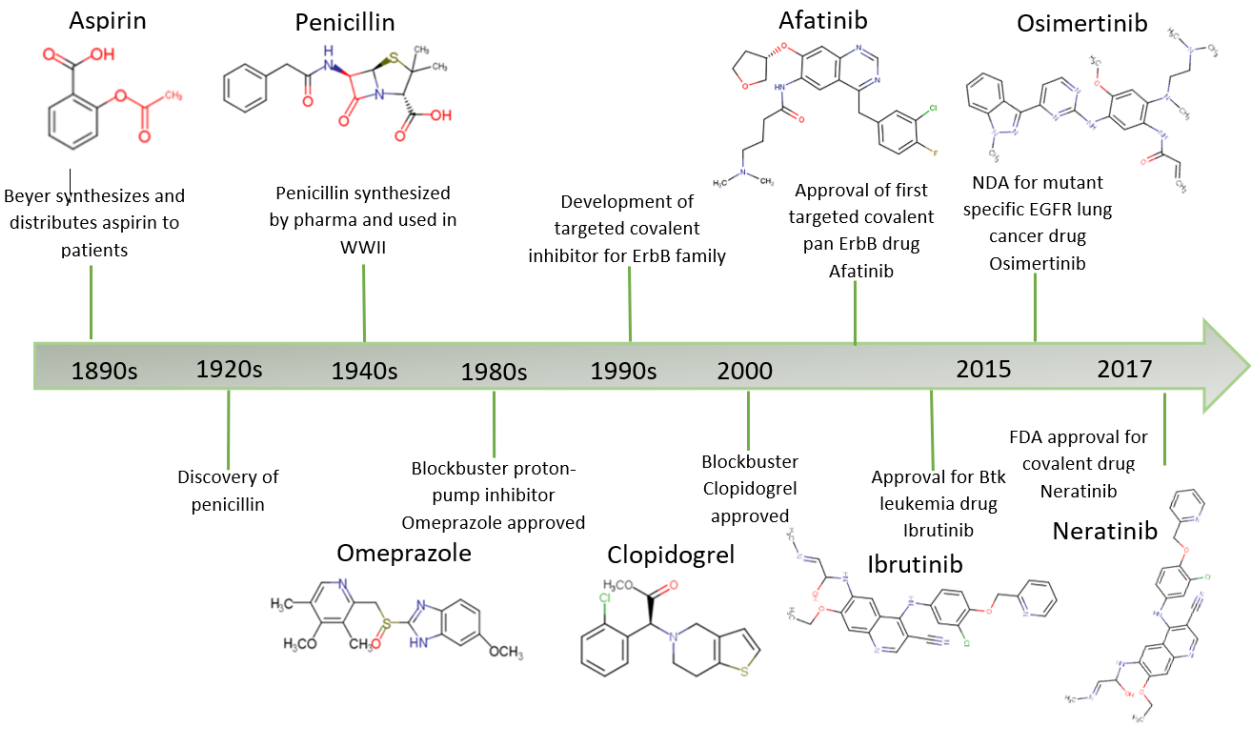


Figure 4.1*: *Timeline representation of approved covalent inhibitors in history.

This bond is irreversible within the half-life of the target protein and sufficiently long lived, resulting in a drug-protein complex, which is not subject to classical equilibrium (Figure 4.2) [5]. On the other hand, the conventional mechanism for non-covalent inhibitors is usually dependent on the development of small inhibitors that bind to a target non-covalently and the subsequent enhancement of potency and selectivity. This is usually attained by enhancing the shape, and non-covalent interfaces (van der Waals interactions, hydrogen bonds, salt bridges, etc.) among the inhibitor and target binding site residues. The binding mechanism of covalent and non-covalent drugs are shown in figure 2 [6].

Although there are numerous examples of efficacious drugs through covalent mechanisms, the clinical and commercial progress of covalent drugs has urged an improved and more thoughtful pursuit of irreversible and covalent mechanisms in the drug discovery process [7].

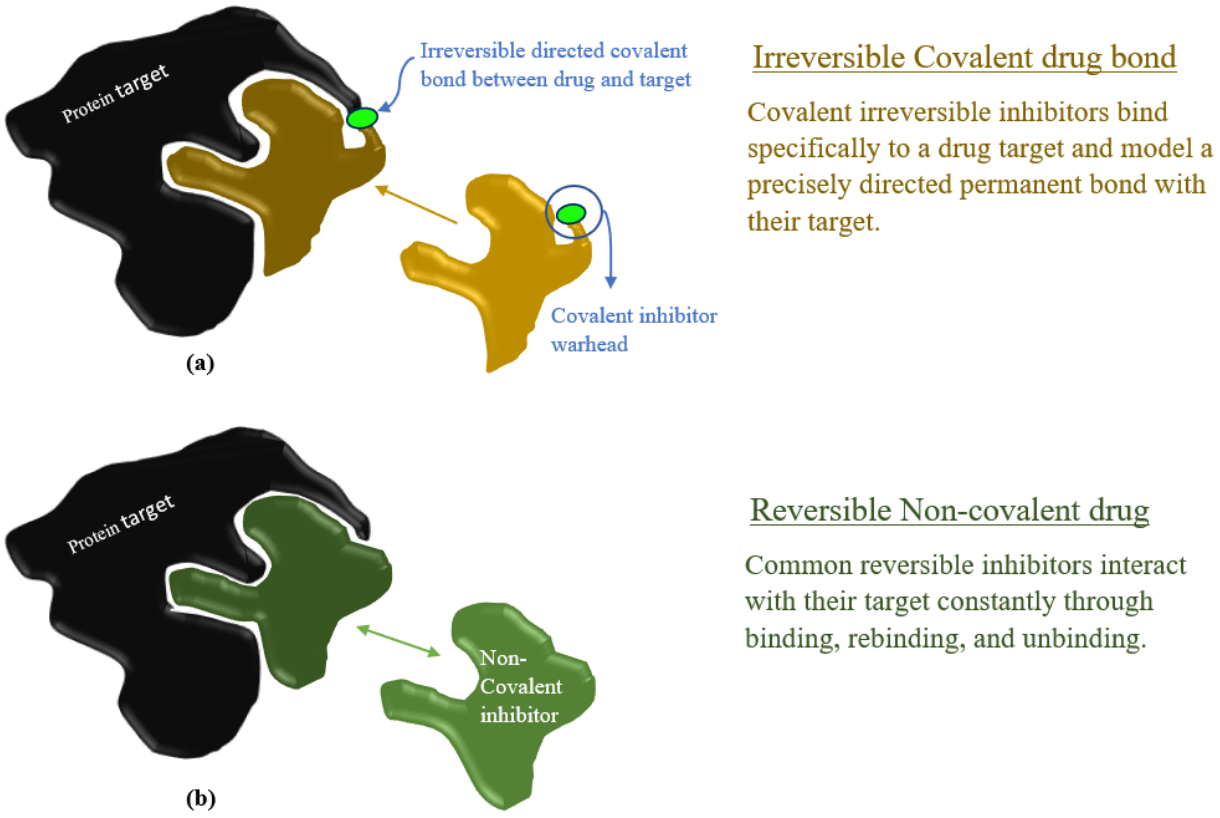


***Figure 4.2**: The mechanism of (a) covalent binding through a functional group in the ligand and (b) non-covalent binding to the drug target protein.*

With the prominent rise in interest toward the rational design of inhibitors and considering the ambiguities surrounding the advantages and disadvantages in drug treatments using covalent and non-covalent inhibition, we opted to provide a detailed understanding of covalent and non-covalent inhibition. We also elucidate on each mechanism's merits and pitfalls, as well as introduce molecular modelling features related to each of these inhibitory modes.

3.2 Non-covalent approach in drug discovery- Merits and hurdles

Non-covalent interactions play an essential role in the structure-based design of new substituent with specific properties [8,9]. They have been recognized in different fields related to chemical pharmacological, biological, physical, and materials sciences [10–12]. Non-covalent interactions like hydrogen bonding, hydrophobic interactions, van Waals interactions, electrostatic interaction and salt bridges, has been the main focus in designing and improving drugs [13,14]. Understanding these interactions and their physical basis is of significant interest

towards improving the currently available drug design strategies. Hence, many efforts have been devoted to explaining the structural, energetic, geometric, and thermodynamic properties of these interactions. The strong hydrogen bonds (Figure 4.3a), with their high directional characteristics, are very efficient in the supramolecular assembly of molecules [15]. In spite of their importance in many scientific domains, controversy on the properties, nature, and importance of hydrogen bonds are increasing. The van der Waals interactions, on the other hand, result from a temporary random fluctuation in circulation of the electrons of atom (Figure 4.3c), which give an upsurge to a transient differing circulation of electrons, a transient electric dipole [16]. van der Waals interactions are responsible for the consistency between molecules of solids and nonpolar liquids. An example of this is heptane, which cannot form hydrogen bonds or ionic interactions with other molecules [9]. When these strong interactions are present, they overrule most of the effects of van der Waals interactions. The strength efficacy of the interactions subsequently decrease with increasing distance; thus, these noncovalent bonds can form only when atoms are quite close to one another [17]. Nevertheless, Ionic bonds (Figure 4.3:b) are an outcome of electrostatic attraction between the negative and positive charges of ions [18]. The Hydrophobic bonds occur between nonpolar molecules, such as hydrocarbons, in an aqueous environment [19,20]. The non-covalent reversible inhibitors, in most cases, bind to the target's active site with relatively greater affinity as compared to that of their natural substrate analogues. The enzyme inactivation is not irreversible in these cases, as the bound inhibitor is non-covalently bound at the active site and is reversibly released from the active site [21].

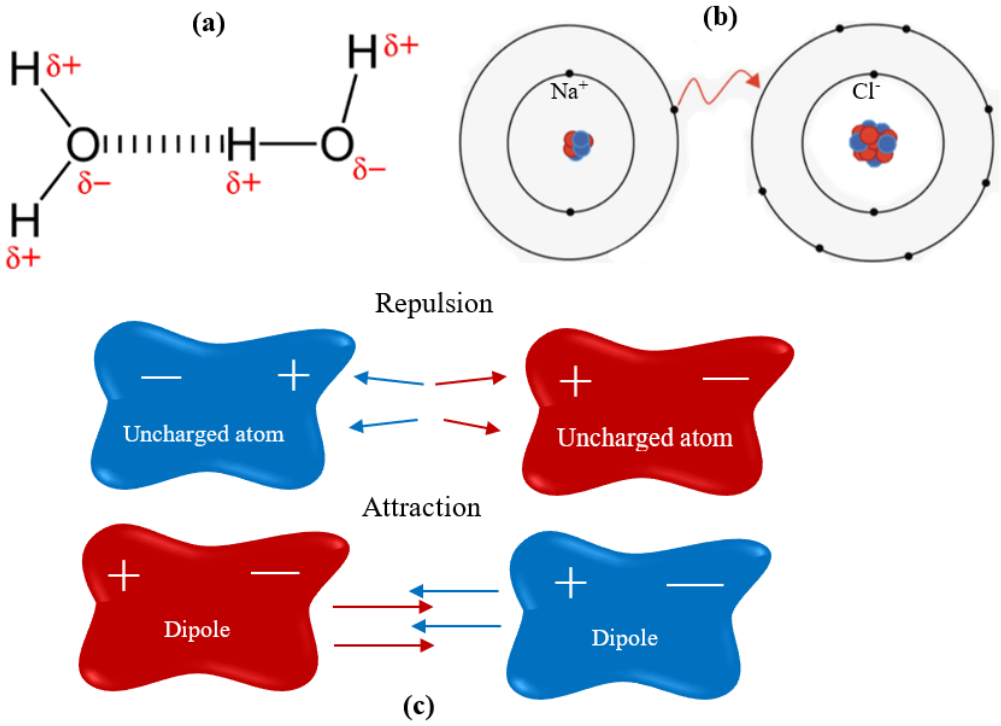
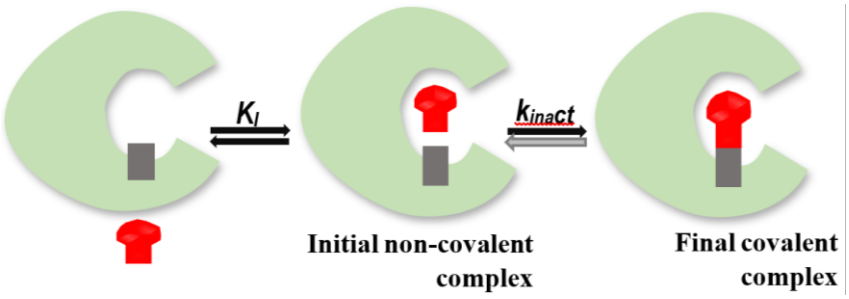


Figure 4.3: Non-covalent interactions: (A) Strong hydrogen non-covalent bond. (B) Ionic bond between two atoms Na^+ and Cl^- by electrostatic force. (C) Van der Waal interaction force.

3.3 Covalent approach in drug Discovery-Merits and hurdles

Covalent drugs have made an extensive positive effect on human health and also have been permitted in curing different disease symptoms. In the pharmaceutical industry, these therapeutic drugs have already proved to be a very beneficial class. Approximately, 30% of the overall drugs in the market that target enzymes are covalent drugs [22]. It is estimated worldwide that around 26 covalent drugs account for more than US$33 billion sales [8]. Inhibitory interactions are defined by the creation of covalent bonds between the electrophilic moiety warhead of the inhibitor and the efficient side chain moieties of definite residues at the active site of the enzyme target [23]. Since the serendipitous discovery of the COX-1 covalent drugs like Aspirin and few residues like cysteine, threonine, histidine, tyrosine, arginine, glutamate and aspartate and lysine. These residue are selective for targeting due to their reactive side chains in their unprotonated states [3,23]. The advancement of selective irreversible covalent inhibitors has been the focus of intensive comprehensive exploration. Protein targets offer the amino acid residue for an appropriate covalent drug to directly affect with the reaction mechanism. Covalent inhibitors are regularly designed to mimic the natural substrate but have

been adjusted in such a way that the reaction cannot continue to achievement, and the drug remains covalently bound to the receptor, irreversibly rendered unable to perform its activity [6]. The covalent bond interaction is completed in two basic steps, the first of which involves the reversible association of the inhibitor with its biological target in an orientation. The faintly electrophilic functionality "warhead" of the ligand is then carried into proximity to a correctly placed nucleophilic residue on the protein. In the second step, an impulsive reaction occurs between the contributing functional groups on the ligand and protein, thus producing the covalently-adapted, inactivated protein. The action of a targeted covalent inhibitor can be defined by the generic mechanism shown in Equation 1, where *Ki* refers to the noncovalent binding constant (affinity of the ligand for the target) and *kinact* to the rate constant for enzyme inhibition [7].



Equation 1. ***The generic mechanism of action of a target-specific covalent inhibitor***

The ligand warhead moiety is vital for the covalent reaction to occur and to control the eventually induced toxicity of enzyme [24]. In consideration of the equity of treatment versus peril, researchers consolidate different kinds of warheads [25], the reactivity of a covalent warhead will not be suitable reactivity variety. In this case it might be acceptable to tune the reactivity by making structural adjustments around the warhead or modify it or replace it entirely, For instance, in the event of cysteine warhead targeting, making the warhead more electrophilic will probably increase the reactivity, however, proper optimization of electrophilicity is needed to control toxicity [6]. Figure 4.4 represents the popular Michael reaction, which occurs between the cysteine residue of proteins and electrophilic warheads (in this example we used the α, β-unsaturated ester due its wide use in the process of drug conception).

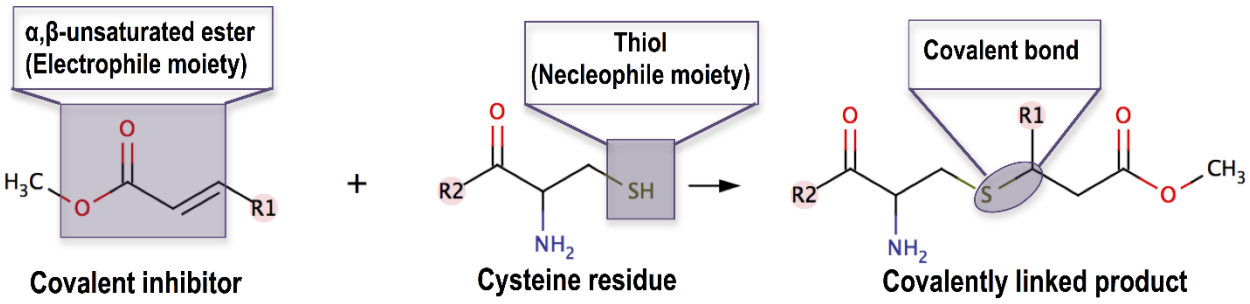


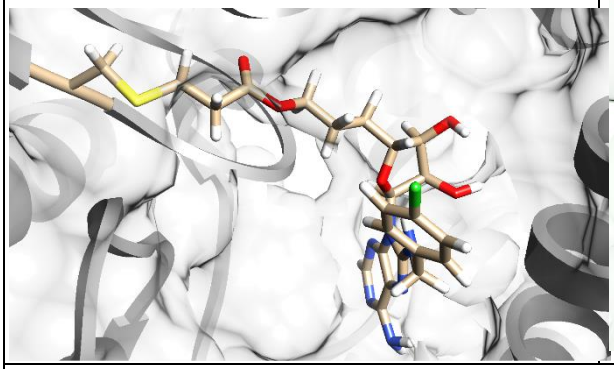
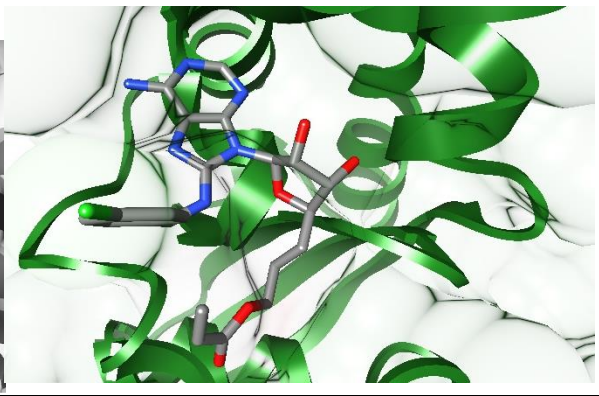
Figure 4.4 A schematic representing Michael addition as a popular example of covalent reaction between, in this case, α, β-unsaturated ester of the ligand and cysteine residue of the protein (these two residues can be any electrophile and nucleophile moieties).

A main challenge for the treatment of infectious diseases and cancer is the development of drug resistance due to mutations at the binding site of a drug target. However, it seems that covalent inhibitors may preserve activity against drug resistant mutations that are gained after treatment with non-covalent inhibitors [26]. Indeed, covalent drug design is an increasing and promising area, the amount and variety of covalent inhibitors and knowledge reaches in this field of drug discovery is continuously increasing. There are plenty strategies for covalent drug design development projects and computational chemistry techniques can have vital impact on covalent inhibitor design and will remain to be a significant aspect to the development of this class of therapeutics.

Computational and molecular modeling tools have become a close counterpart to experiments in the understanding of molecular aspects of biological systems. The computational simulations of non-covalent biological systems have been widely explored by scientists in the field of biological and pharmaceutical chemistry [26]. However, the computational studies of covalent inhibition of target proteins have been limited due to the shortage of technical protocols for *in silico* simulations of this kind of irreversible systems. These techniques aid the simulation of biological interactions thereby providing molecular and structural insights into the direct binding of lead compounds to biological targets coupled with the dynamics and induced alterations in the structure–function relationship of target proteins. Our recent reports have covered this gap by providing our in-house technical protocol on the *in silico* modeling and tackling the challenges related to the simulation of covalent systems [27–29]. A summary of the protocol is provided in table 4.1. The computational preparation for the non-covalent and

covalent bonds has a similarity process with a different step, as it focused on the covalent bond preparation is more favorable than the non-covalent bond, especially after docking preparation to antechamber steps alongside the Tleap preparation table 1.

Table 4.1: Computational preparation for Covalent bond and Non-covalent bond:

<table>
<tr><th>Covalent preparation</th><th>Non-covalent preparation</th></tr>
<tr><td colspan="2">Docking</td></tr>
<tr><td>Using Auto-dock vina for covalent docking after a protein ligand charge and hydrogen preparation, by posing the ligand near to the residue, which we create the covalent bond with, making it close to each other approximately 2.5Å or 4 Å.
  </td><td>Protein and ligand preparation by add charges and hydrogen or removal in certain cases, using Auto-dock vina for docking the ligand with the receptor at the right pose in the active site, with no bond attached with the ligand.
  </td></tr>
<tr><td colspan="2">Molecular dynamic (MD) preparations</td></tr>
<tr><td>Creating covalent bond by using some softwares for instance Maestro software, make the ligand and residue as one complex under one name and residue number covalently attached.
Antechamber
Tleap with residue charge inserted and changed
Dabble for the complex
Minimization with certain restrains from high restrains to low restrains in several steps
Heating,
Equilibration in several steps
MD simulation.
checking the covalent Bond and making sure after every step the Bond still save and not broken with the whole complex inside the Water box.</td><td>Unlike the Covalent preparations, non-covalent complex is not requiring using any softwares for the ligand and receptor after docking, using the MD preparation steps straight after docking.
Antechamber with ligand charge
Tleap
Minimization (Partial Minimization, Full Minimization)
Heating
Equilibration
MD Simulation
Checking the complex and water box before going to MD simulation step.</td></tr>
<tr><td colspan="2">Molecular dynamic simulation (MDs)</td></tr>
</table>

Molecular dynamic simulation going on short several steps of Nano second for the irreversible covalent bond, to make sure the binding of the covalent bond is not affected during the simulation periods.	Non- covalent molecular dynamic simulation, going on long several steps of Nano second, with complex and Water box checking after every MD period.

Moreover, some aspects are important for these types of interactions such as interactions strength, environment through which it acts distance parameters. Examples for small range interactions for these forces are shown in Table 4.2.

Table 4.2: Example for small range interactions forces:

Nature of bond interaction	Type of Force	Distance	Energy (Kcal/mol)
Ionic bond interaction	Coulombic force	2.8 **Å** 2.0 **Å**	180 NaCl 240 LiF
Covalent bond interaction	Electrostatic force (sharing to pairs of electrons, wave function overlap)	N/A	170 Dimond 283 SiC
Hydrogen bond interaction	Strong type of directional dipole-dipole interaction	N/A	7 HF
Van der Waals	Dispersion forces, dipole-induced dipole forces, dipole-dipole forces	Important in the range from a few **Å** to hundred **Å**	2.4 CH_4

3.4 The essence docking protocol of molecular modeling and drug design perspective for both covalent and non-covalent inhibitors

Molecular docking is a computational process implemented on structure-based rational drug design to identify right conformations of small molecule ligands and also to estimate the strength of the ligand-protein interaction, commonly one receptor and one ligand [30–32]. The most common docking tools are AutodockVina, Autodock [33], and maestro *Schrödinger* [34] and more other softwares. Nevertheless, many other methods similar to them mainly focus on the docking between the two molecules through non-covalent interactions (van der Waals interaction, the electrostatics interaction and hydrogen bonding) Figure 4.3, using other knowledge or pragmatic to characterize these non-covalent interactions [35]. However, not all

drugs bind non-covalently to the protein active site, there are other classifications of drugs, called the covalent drugs [7]. Nowadays docking of ligands that are bound to a receptor through non-covalent interactions is relatively imitative. The majority of docking methods development research has been focused on the effective prediction of the binding modes of non-covalent inhibitors [36,37]. Because of the reaction between the drug and the protein that needs to be taken into consideration, docking ligands that bind covalently to the receptor it's been complicated [7]. Most covalent docking software's are only successful in predicting the binding energy between a nucleophilic receptor and electrophilic ligand; However, Different routines have been developed to perform covalent docking of the inhibitors to the target proteins Figure 5.

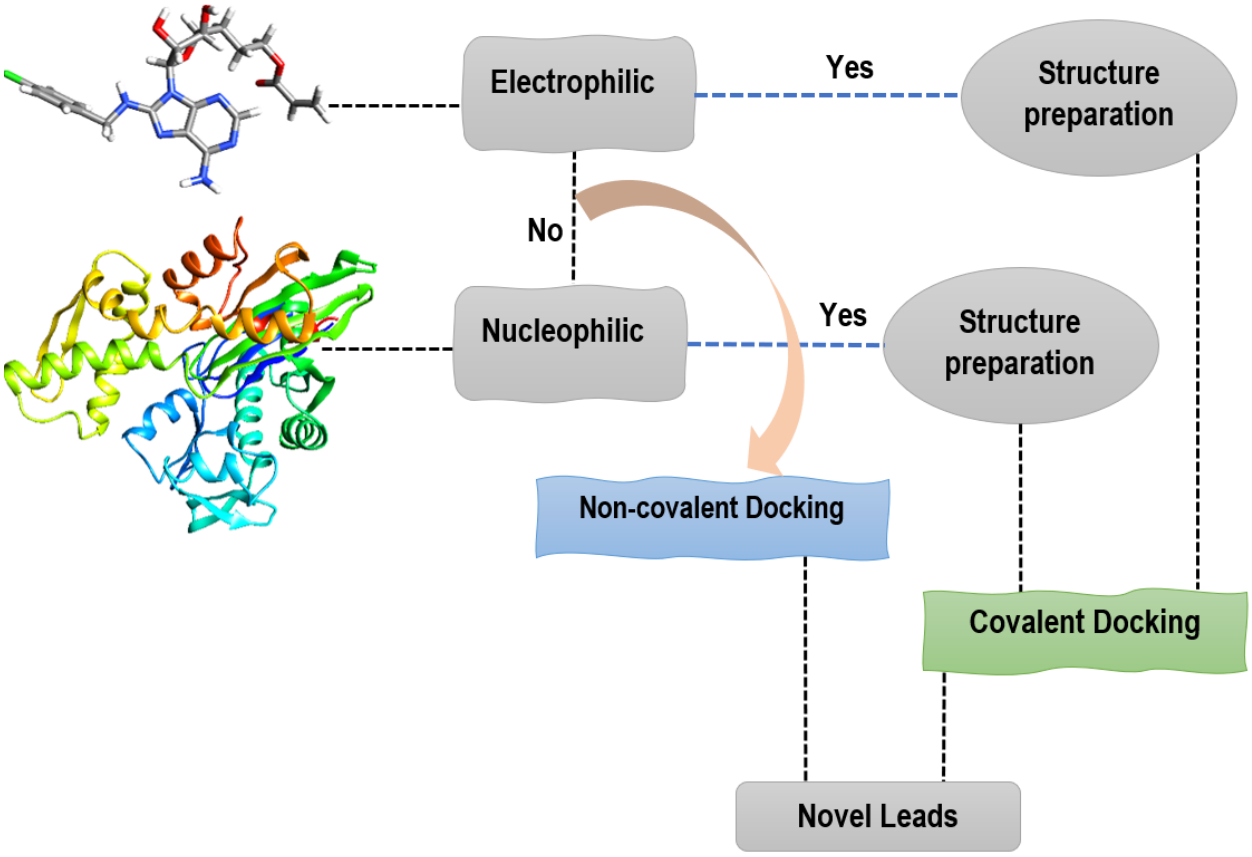


Figure 4.5: Scheme describing the workflow of covalent docking in drug discovery.

Docking methods for the covalent bond is deferent from the non-covalent bond, although of using same software's [38]. Essentially the docking of covalent bond is between the inhibitor and the residue "link atom", creating the covalent bond for the inhibitor with a certain residue it's the main reason for the covalent docking. Using a certain software's like Auto-dock vina and Maestro *Schrödinger* for instance [33,38], helping to create the perfect covalent bond and

create a ligand residue complex. The perfect docking for covalent bond is almost so important that the molecular dynamic simulation results depend on it. Making the distance between the inhibitor and the enzyme specific residue close to each other approximately 2.5Å or 4 Å is essential for covalent docking as it showing in table 4.1, after create the covalent bond between the ligand and the inhibitor in such docking distance it's more guaranteed in MD simulation period the covalent bond won't break, and that's so important to carry on the simulation process to get the desired results. However, for the non-covalent docking there's bond between the ligand and the residue, it depends on the right pose for the inhibitor inside the enzyme pocket site and which residues are more attracted to it by using analysis such as (MM/GB-SA-MM/PBSA) [39]. Choosing the right pose for the inhibitor during the docking process with the highest negativity position, is significant for molecular modeling simulation and better analysis results. Docking materials can affect the reversible and irreversible inhibition for the way ligand has been docked (posed, created bond). Docking is importantly essential to express the relationship outcome results between the inhibitor and the enzyme. In covalent binding the ligand binds first through non-covalently interaction with the protein in a pose that benefits the reaction, and then followed by the reaction between the elements to form a covalent bond. Therefore, in Covalent-Dock the interactions between the inhibitor and its protein are modelled the same as in conventional molecular dockings through non-covalent interactions with additional energy contribution from covalent bond formation (Equation (1)).

3.5 Paving the route: guided answer from molecular modelling perspective

The bond's strength interactions describe how strongly each atom is bound to another atom and how much energy is required to break the interaction bond between these atoms. The interactions strength of the covalent and non-covalent is significant in various field of protein folding, drug design and origins of life. Bonds interactions such as ionic bond, covalent bond, hydrogen bond and van der Waal interaction are important to define the interaction strength, form the strongest bond to the weakest starting with the higher strength interaction the ionic bonds, due to the massive electrostatic attraction between two atoms, the ionic bonds are the strongest which including the higher strength interaction. Nevertheless, covalent bond comes next, because it is formed by sharing and overlap of two atoms, and in some cases, it considers stronger than the ionic bond. And then it comes the hydrogen bond and it's followed by Van der Waals interactions.

In terms of deciding the route taken in illness therapy of covalent or non-covalent inhibition, it is always riddled with confusion, despite the advantages and disadvantages of both. From a molecular modeling and drug design perspective, we see that both are equally and not equally from the therapeutic side. In the case of drug mutation, such as those implicated in cancer disease, covalent inhibitors prove to be more effective because of its prolonged stay at the active site [32], resulting high selectivity and activity as compared to the reversible effect of the non-covalent drugs [34,39]. However, non-covalent drugs can be better therapeutics when it comes to evading toxicity that can be associated with the irreversible binding of covalent drugs. Covalent drugs possess some advantages compared to the non-covalent one such as a high binding affinity, higher selectivity and specificity leading to complete target inhibition under non-equilbrium kinetics [40]. Their prolonged activity also helps in decreasing dose frequency and enhancing the drug regimen [41]. In addition, they have a great potency in overcoming the challenge of drug resistance which has caused a significant draw back in the disease treatment using a non-covalent drugs [7]. Popular drugs, afatinib, neratinib and dacomitinib, which have been shown to overcome drug resistance [25,42], occurs when non-covalent drugs such as lapatinib and gefitinib are used as EGFR inhibitors [27,43], also hepatitis C mutant patients have been effectively treated with covalent drugs [44]. Covalent drugs have shown an important strategy for addressing supposedly undruggable and challenging targets in human diseases. Odanacatib, which effectively inhibits Cathepsin K, which was regarded as an undruggable cysteine protease [45]. However, in some cases, the long-lasting effect and the permanent nature of bond between the drug and the target can be associated with drug toxicity. Hence, researchers from academia and industry tend to combine strengths of covalent and non-covalent modes of action by carefully designing reactive compounds that can overcome the drug resistance and still not toxic, especially when the administrated doses are lower. Lipinski has provided guidelines for computational as well as experimental scientists to rationally assess the inhibitors during the conception phase of drugs [46]. Moreover, in the light of the prodigious evolution of computational tools, it is possible to direct the efforts of researchers in every phase of drug development. This, has highly contributed to incline the researches to the reversible covalent drugs that has proved to overcome the aforementioned challenges and also reduce off-target toxicity and drug-drug interaction [4,47].

3.6 Conclusion

In this review, the pros and cons of covalent and non-covalent inhibitory mechanisms were elucidated on. The characteristic features of each class were discussed, including the duration of the drug at the active site region as well as the extrapolated effects on the drug potency and selectivity. In this study we also provided a short summary of our previous in-house protocol, which will be used to identify covalent inhibitors based on pharmacophore-based virtual screening. A final opinion was offered on the suitability of each drug class. It was evident that both classes had its merits, however, with the increase in drug resistance, covalent inhibition seemed more suitable. For example, Michael addition in the case of cysteine residues, thiolate is formed based on the deprotonated cysteine thiol. It then reacts with the functional group of the inhibitor. The rate of covalent modification depends on the pKa of the cysteine residue because of the deprotonation step. More reactive cysteines have lower pKas because of the increased probability that the cysteine will be in the nucleophilic thiolate state. Accordingly, cysteines with low pKas are more susceptible to covalent modification. Thus, calculating the pKas of cysteines could help identify druggable cysteine targets prone to covalent modification, leading to design of more potent and selective inhibitors. Hence, enhanced selectivity of the covalent systems is manageable proving its use as a therapeutic regimen worldwide. We believe that this study will assist researchers in making informed decisions when identifying whether or not to further assess a compound.

References

1. Schwartz PA, Kuzmic P, Solowiej J, *et al.* Covalent EGFR inhibitor analysis reveals importance of reversible interactions to potency and mechanisms of drug resistance. *Proc. Natl. Acad. Sci. U. S. A.* 111(1), 173–178 (2014).

2. Pettinger J, Jones K, Cheeseman MD. Lysine-Targeting Covalent Inhibitors. *Angew. Chemie - Int. Ed.* 56(48), 15200–15209 (2017).

3. Bauer RA. Covalent inhibitors in drug discovery: From accidental discoveries to avoided liabilities and designed therapies. Drug Discov. Today. 20(9), 1061–1073 (2015).

4. Adeniyi AA, Muthusamy R, Soliman MES. New drug design with covalent modifiers. Expert Opin. Drug Discov. 11(1), 79–90 (2016).

5. Powers JC, Asgian JL, Ekici ÖD, James KE. Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem. Rev.* 102(12), 4639–4750 (2002).

6. Lonsdale R, Ward RA. Structure-based design of targeted covalent inhibitors. *Chem. Soc. Rev.* 47(11), 3816–3830 (2018).

7. Singh J, Petter RC, Baillie TA, Whitty A. The resurgence of covalent drugs. *Nat. Rev. Drug Discov.* 10(4), 307–317 (2011).

8. Strelow JM. A Perspective on the Kinetics of Covalent and Irreversible Inhibition. J. Biomol. Screen. 22(1), 3–20 (2017).

9. Hunter CA. Van der Waals interactions in non-polar liquids. *Chem. Sci.* 4(2), 834–848 (2013).

10. Irimia-Vladu M, Głowacki ED, Schwabegger G, *et al.* Natural resin shellac as a substrate and a dielectric layer for organic field-effect transistors. *Green Chem.* 15(6), 1473–1476 (2013).

11. Schneider HJ. Binding mechanisms in supramolecular complexes. *Angew. Chemie - Int. Ed.* 48(22), 3924–3977 (2009).

12. Bonin J, Costentin C, Robert M, Tard D. ProtonÀElectron Transfers. *Structure* [Internet]. , 372–81 (2011). Available from: http://www.ncbi.nlm.nih.gov/pubmed/22029773.

13. Zhanting Li L-ZW, Zhang D-W, Wong H, Ting-Li Z. Hydrogen Bonded Supramolecular Structures [Internet]. Available from: https://books.google.com/books?id=hJ8oBgAAQBAJ&pgis=1.

14. Zhao GJ, Han KL. Hydrogen bonding in the electronic excited state. *Acc. Chem. Res.* 45(3), 404–413 (2012).

15. Desiraju GR. A bond by any other name. *Angew. Chemie - Int. Ed.* 50(1), 52–59 (2011).

16. Zhang L. The van der Waals force and gravitational force in matter. , 1–14 (2013). Available from: http://arxiv.org/abs/1303.3579.

17. Finkelstein A V. Average and extreme multi-atom Van der Waals interactions: Strong coupling of multi-atom Van der Waals interactions with covalent bonding. *Chem. Cent. J.* 1(1), 1–9 (2007).

18. Mahmudov KT, Kopylovich MN, Guedes da Silva MFC, Pombeiro AJL. Non-covalent interactions in the synthesis of coordination compounds: Recent advances. *Coord. Chem. Rev.* [Internet]. 345, 54–72 (2017). Available from: http://dx.doi.org/10.1016/j.ccr.2016.09.002.

19. Hadidi M, Zydney AL. Fouling behavior of zwitterionic membranes: Impact of electrostatic and hydrophobic interactions. *J. Memb. Sci.* [Internet]. 452, 97–103 (2014). Available from: http://dx.doi.org/10.1016/j.memsci.2013.09.062.

20. Privalov PL. Hydrophobic Interactions in Proteins. *Protein Struct. Protein Eng.* , 6–15 (1988).

21. Puhan MA, Chandra D, Mosenifar Z, *et al.* Fisetin Acts on Multiple Pathways to Reduce the Impact of Age and Disease on CNS Function. 37(4), 784–790 (2017).

22. De Cesco S, Kurian J, Dufresne C, Mittermaier AK, Moitessier N. Covalent inhibitors design and discovery. *Eur. J. Med. Chem.* [Internet]. 138, 96–114 (2017). Available from: http://dx.doi.org/10.1016/j.ejmech.2017.06.019.

23. Baillie TA. Targeted Covalent Inhibitors for Drug Design. *Angew. Chemie - Int. Ed.* 55(43), 13408–13421 (2016).

24. Miller RM, Paavilainen VO, Krishnan S, Serafimova IM, Taunton J. Electrophilic fragment-based design of reversible covalent kinase inhibitors. *J. Am. Chem. Soc.* 135(14), 5298–5301 (2013).

25. Liu Q, Sabnis Y, Zhao Z, *et al.* Developing irreversible inhibitors of the protein kinase cysteinome. Chem. Biol. 20(2), 146–159 (2013).

26. Bjij I, Olotu FA, Agoni C, *et al.* Covalent Inhibition in Drug Discovery: Filling the Void in Literature. *Curr. Top. Med. Chem.* 18(13), 1135–1145 (2018).

27. Akher FB, Farrokhzadeh A, Soliman MES. Covalent vs. Non-Covalent Inhibition: Tackling Drug Resistance in EGFR – A Thorough Dynamic Perspective. *Chem. Biodivers.* 16(3) (2019).

28. Khan S, Bjij I, Olotu FA, Agoni C, Adeniji E, Soliman MES. Covalent simulations of covalent/irreversible enzyme inhibition in drug discovery: A reliable technical protocol. *Future Med. Chem.* 10(19), 2265–2275 (2018).

29. Badichi Akher F, Farrokhzadeh A, Olotu FA, Agoni C, Soliman MES. The irony of chirality-unveiling the distinct mechanistic binding and activities of 1-(3-(4-amino-5-(7-methoxy-5-methylbenzo[: B] thiophen-2-yl)-7 H -pyrrolo[2,3- d] pyrimidin-7-yl)pyrrolidin-1-yl)prop-2-en-1-one enantiomers as irreversible covalent FGFR4. *Org. Biomol. Chem.* 17(5), 1176–1190 (2019).

30. Yuriev E, Agostino M, Ramsland PA. Challenges and advances in computational docking: 2009 in review. *J. Mol. Recognit.* 24(2), 149–164 (2011).

31. Ekmekci B, McAnany CE, Mura C. An Introduction to Programming for Bioscientists: A Python-Based Primer. *PLoS Comput. Biol.* 12(6), 1–43 (2016).

32. Kumalo HM, Bhakat S, Soliman MES. Theory and applications of covalent docking in drug discovery: Merits and pitfalls. *Molecules*. 20(2), 1984–2000 (2015).

33. Trott O, Olson A. Autodock vina: improving the speed and accuracy of docking. *J.*

Comput. Chem. 31(2), 455–461 (2010).

34. Madhavi Sastry G, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided. Mol. Des.* 27(3), 221–234 (2013).

35. Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* 261(3), 470–489 (1996).

36. Yuriev E, Ramsland PA. Latest developments in molecular docking: 2010-2011 in review. *J. Mol. Recognit.* 26(5), 215–239 (2013).

37. Jacob RB, Andersen T, McDougal OM. Accessible high-throughput virtual screening molecular docking software for students and educators. *PLoS Comput. Biol.* 8(5) (2012).

38. Kusumaningrum S, Budianto E, Kosela S, Sumaryono W, Juniarti F. The molecular docking of 1,4-naphthoquinone derivatives as inhibitors of Polo-like kinase 1 using Molegro Virtual Docker. *J. Appl. Pharm. Sci.* 4(11), 47–53 (2014).

39. Hou T, Wang J, Li Y, Wang W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.* 51(1), 69–82 (2011).

40. Johnson DS, Weerapana E, Cravatt BF. Strategies for discovering and derisking covalent, irreversible enzyme inhibitors. *Future Med. Chem.* 2(6), 949–964 (2010).

41. Claxton AJ, Cramer J, Pierce C. A systematic review of the associations between dose regimens and medication compliance. *Clin. Ther.* 23(8), 1296–1310 (2001).

42. Yun CH, Mengwasser KE, Toms A V., *et al.* The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc. Natl. Acad. Sci. U. S. A.* 105(6), 2070–2075 (2008).

43. Bjij I, Khan S, Ramharak P, Cherqaoui D, Soliman MES. Distinguishing the optimal binding mechanism of an E3 ubiquitin ligase: Covalent versus noncovalent inhibition. *J. Cell. Biochem.* 120(8), 12859–12869 (2019).

44. Hagel M, Niu D, St Martin T, *et al.* Selective irreversible inhibition of a protease by targeting a noncatalytic cysteine. *Nat. Chem. Biol.* 7(1), 22–24 (2011).

45. Halgren TA. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* 49(2), 377–389 (2009).

46. Swinney DC. Biochemical mechanisms of drug action: What does it take for success? *Nat. Rev. Drug Discov.* 3(9), 801–808 (2004).

47. Smith AJT, Zhang X, Leach AG, Houk KN. Beyond picomolar affinities: Quantitative aspects of noncovalent and covalent binding of drugs to proteins. *J. Med. Chem.* 52(2), 225–233 (2009).

CHAPTER 5

Coupling of HSP72 α-helix Subdomains by the Unexpected Irreversible Targeting of Lysine-56 over Cysteine-17; Coevolution of Covalent Bonding

Aimen Aljoundi, Ahmed El Rashedy, Patrick Appiah-Kubi, and Mahmoud E.S Soliman*

Molecular Bio-computation and Drug Design Laboratory

School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 78

Abstract

Covalent inhibition has recently gained a resurgence of interest in several drug discovery areas. The expansion of this approach is based on evidence elucidating the selectivity and potency of covalent inhibitors when bound to particular amino acids of a biological target. The unexpected covalent inhibition of heat shock protein 72 (HSP72) by covalently targeting Lys-56 instead of Cys-17 was an interesting observation. However, the structural basis and conformational changes associated with this preferential coupling to Lys-56 over Cys-17 remain unclear. To resolve this mystery, we employed structural and dynamic analyses to investigate the structural basis and conformational dynamics associated with the unexpected covalent inhibition. Our analyses reveal that the coupling of the irreversible inhibitor to Lys-56 is intrinsically less dynamic than Cys-17. Conformational dynamics analyses further reveal that the coupling of the inhibitor to Lys-56 induced a closed conformation of the nucleotide-binding subdomain (NBD) α-helices, in contrast, an open conformation was observed in the case of Cys-17. The closed conformation maintained the crucial salt-bridge between Glu-268 and Lys-56 residues, which strengthens the interaction affinity of the inhibitor nearly identical to adenosine triphosphate (ADP/Pi) bound to the HSP72-NBD. The outcome of this report provides a substantial shift in the conventional direction for the design of more potent covalent inhibitors.

Keywords: Covalent MD simulation; HSP72; 8-N-benzyladenosine; Coupling; Principal component analysis.

5.1 Introduction:

Heat shock proteins (HSPs) play a central role in the clearance of damaged proteins by inducing protein aggregation and proteotoxicity. This process occurs by preventing inappropriate stress-induced protein aggregation, ensure proper refolding of denatured proteins, and, if necessary, the promotion of their degradation [1–3]. Recent studies have proven that increased protein synthesis (translation) is vital to the conversion of neoplasms. As a result of this increase, cancer cells appear to be particularly susceptible to agents that inhibit the removal of aggregated or misfolded proteins generated by protein synthesis as a product [4–6]. Hsp70 protein member families are among the highly conserved proteins and play a critical role in these processes [7]. The primary stress-inducing member of the Hsp70 chaperone family is known as HSP72 and is encoded by two genes, HSPA1A and HSPA1B, which generate isoforms of Hsp72 [8]. HSP72 is extremely homologous to the 78 kDa glucose-regulated protein, which plays a significant role in organizing the unfolding protein response[9]. HSP72 is expressed at high levels in malignant tumors of various origins [10] and enhances cancer cell survival [11,12]. Thus, the inhibition of HSP72 is considered to be a successful pathway in anti-tumor therapy [13]. All the different functions of Hsp70s are accomplished through a transient chaperone interaction with substrate proteins through its C-terminal substrate-binding domain (SBD) [14]. The nucleotide binds allosterically to the *N*-terminal nucleotide-binding domain (NBD) to control the transient chaperone interaction. The affinity of the SBD for substrates decreases by 10- to 400-fold when ATP is binding to the NBD. Hence, the inhibition of NBD is considered one of the most promising strategies for HSP72 function inhibition [15]. The NBD consists of two adjacent lobes (lobe I and lobe II), which form a deep nucleotide groove connected to the base. Each lobe consists of two subdomains (IA, IIA, IB, and IIB) [16,17]. Domains IB and IIB are linked to IA and IIA, respectively, by flexible hinges and control access to the nucleotide-binding sites [18] (Figure 5.1).

Several studies have designed potential HSP72 inhibitors, including 2-phenylethynesulfonamide (PES) [19], 15-deoxyspergualin (DSG) [20], natural products Oridonin [21] and Novolactone [22], but upregulation is one of the most challenges associated with drug resistance and poor clinical outcomes [23]. The challenging hurdle to cellular activity for competitive nucleotide inhibitors of HSP72 is due to the highly conserved domain. This conserved domain is mostly occupied by ADP and ATP (ADP, KD ~ 110 nm) in addition to hydrophilic and electrostatic interactions with the nucleotide ribose and phosphate amino acid residues, hence difficult drug binders [24]..

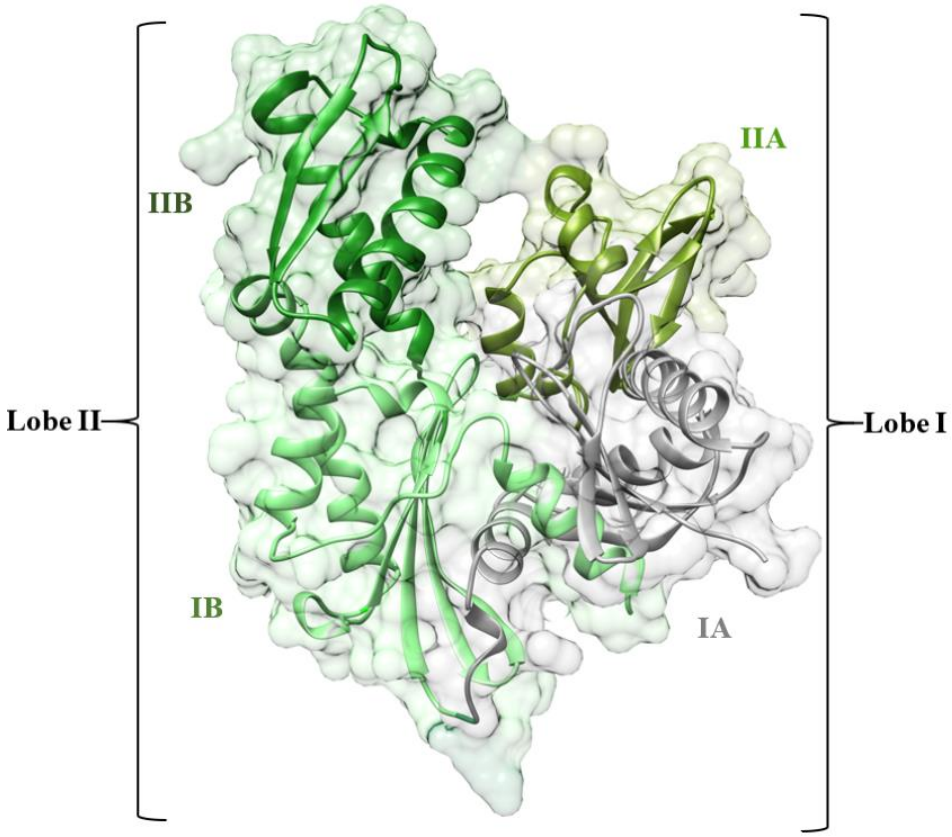


Figure 5.1 The 3-D crystal structure of the HSP72-NBD protein (PDB code: 5MKS). The IA, IIA, IB, and IIB subdomains are shown in green, light-green, oily green, and grey, respectively.

Covalent inhibition is a key approach for high-affinity proteins [25] and has recently sparked interest among the community of pharmaceutical research [26]. Covalent inhibition occurs when the electrophilic moiety of a covalent ligand connects with a nucleophilic residue of a biological target, resulting in an irreversible link between the protein and the drug. For example, it can inhibit the same biological target at a lower concentration than a noncovalent drug due to the long-lasting effects of a covalent drug. [27,28]. An example of a covalent reaction between a ligand and its protein target is shown in Figure 5.2.

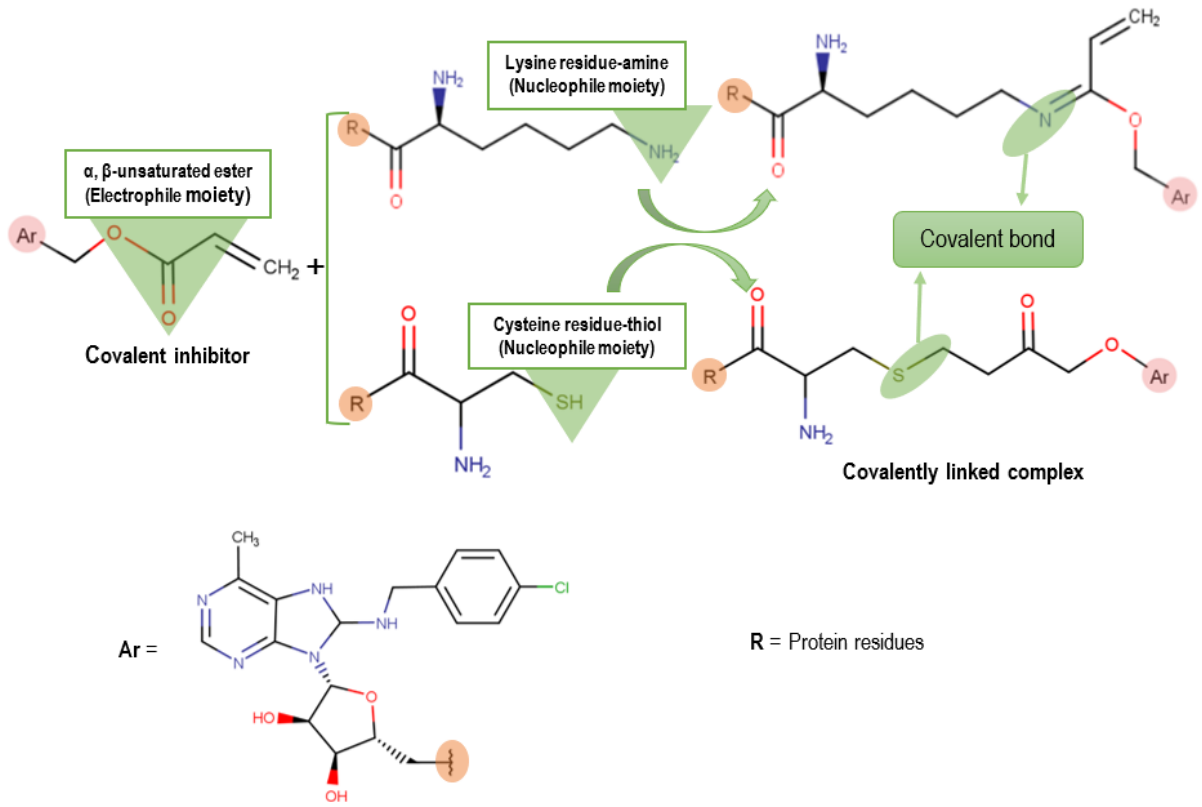


Figure 5.2 A schematic summarizing the covalent reaction mechanism between a covalent inhibitor and the protein residues lysine and cysteine.

In a recent study by Pettinger et al. (2017) using fluorescence polarization (FP) assay and crystallography, the authors observed an unexpected covalent bond interaction between 8-N-benzyladenosine and lysine-56 of the NBD of HSP72 (HSP72-NBD domain). This unexpected covalent bond interaction resulted in the arrest of the NBD via hydrogen-bonding array of the ribose and adenine moieties with the lipophilic para-chlorobenzylamine moiety, parallel with the two α-helices of the binding cavity [29] (Figure 5.3). It is worth highlighting that the observed covalent inhibition of HSP72 via lysine-56 by 8-*N*-benzyladenosine was opposed to the anticipated 8-*N*-benzyladenosine covalent inhibition of HSP72 via cysteine-17.

The unexpected preferential covalent bond formation of 8-*N*-benzyladenosine with lysine-56 over cysteine-17 prompted the need to investigate the conformational plasticity and structural dynamics associated with this unexpected covalent interaction. To accomplish this, we utilized in silico approaches such as covalent molecular dynamics simulation, clustering and principal component analyses to define and compare the structural dynamics of 8-*N*-benzyladenosine-Lys-56 modeled covalent complex with 8-*N*-benzyladenosine-cys-17 covalent complex on HSP72-NBD domains.

Extensive analyses reveal that the coupling of the inhibitor to cysteine-17 is intrinsically more dynamic than to lysine-56, mainly in the IIA and IA α-helices region. Conformational dynamics analysis further reveals that the coupling of 8-*N*-benzyladenosine to lysine-56 induces a closed conformation of the IIB and IIA α-helices of the NBD. In contrast, an open conformation was observed when coupled to the cysteine-17 residue.

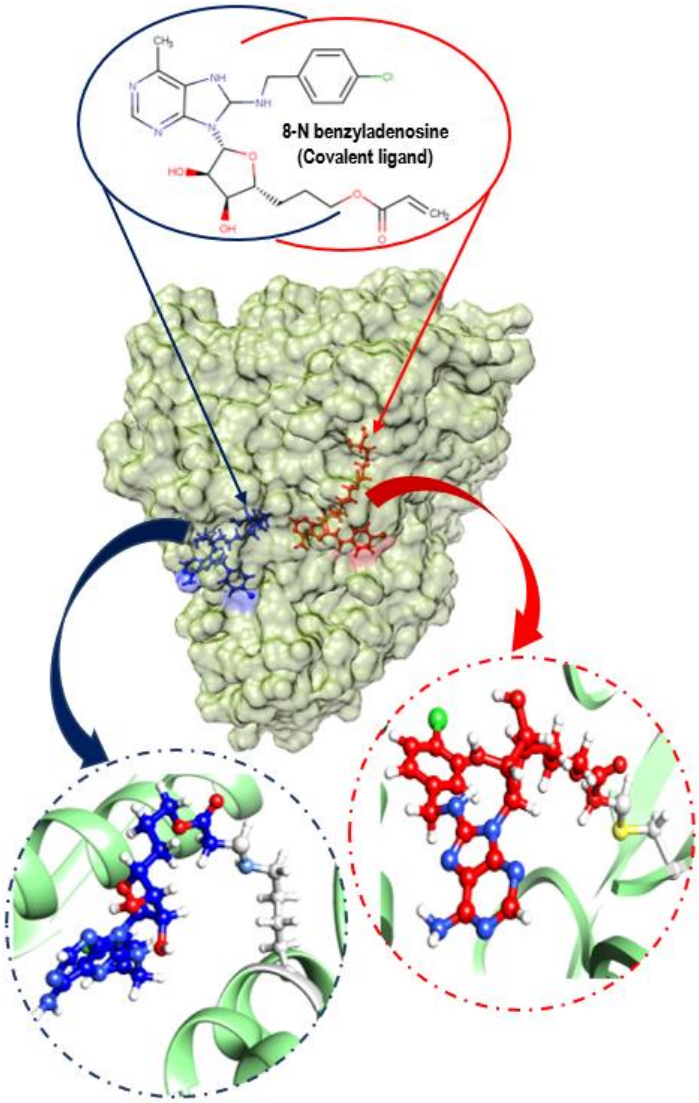


Figure 5.3 Surface view of HSP72 covalently bonded with 8-N-benzyladenosine inhibitor via a cysteine residue (red) and a lysine residue (navy-blue) with different binding pocket. A close view of two covalent bonds (yellow color cysteine residue bonding and light blue for lysine residue bonding)

5.2 Computational methodology

5.2.1 System preparation

The studied models were prepared based on the reported human HSP72-NBD domain crystal structure retrieved from the Protein Data Bank [30] (PDB code: 5MKS) [31] and set up using

the UCSF Chimera software package [32]. The ligand was prepared using MarvinSketch 6.2.1, 2014, Molegro Molecular Viewer (MMV), and Chem-Axon (http://www.chemaxon.com) to ensure that the ligand hybridization state and proper angels were displayed [33,34].

5.2.2 Covalent Docking

Covalent bonds were formed between the inhibitor and lysine-56, cysteine-17 residues of HSP72-NBD. Before the creation of the covalent bond, a non-covalent docking was performed for the inhibitor to ensure an appropriate stable binding mode at HSP72-NBD. This initial non-covalent docking was performed using AutoDock Tools GUI, and AutoDock Vina [35] integrated with Chimera GUI. The AutoDock Tools GUI was used to define the grid box (center: −12.75, −8.433, and 5.19; Size: 18.65, 18.02, and 21.39) at the lysine and cysteine binding site of the protein. The initial non-covalent docking was to appropriately position the inhibitor in the active site to allow the covalent bond to be created subsequently. The binding poses were checked to see if the two atoms that will eventually form the covalent bond between the target residues (lysine-56 and cysteine-17) and the inhibitor was within 3 Å of each other. The inhibitor binding mode that did not meet this condition to allow the bond to be formed was rejected. The Schrödinger Maestro [36] was used to create the covalent bond between the inhibitor and the residues (lysine-56 and cysteine-17). The best covalent complex pose was then selected.

The protein preparation wizard in Maestro Schrödinger was used to optimize the protonation state of the complex, adjust hydrogen atoms, cap acetyl, and methylamide neutral residues. An initial vacuum minimization was performed to resolve any steric clashes and restore normal bond lengths. The prepared receptor and covalent ligand were saved separately and taken for parameterization using the preparatory program. The APO system was run based on our previously reported protocol for non-covalent simulations [37,38].

5.2.3 Covalent molecular dynamic simulation

The two covalent systems were exposed to an all-atom classical covalent molecular dynamics simulation (MD) using the PMEMD package in Amber 14 [39]. The Antechamber module in the Amber 14 was used to provide atom types and atomic ligand partial charges using the FF14SB forcefield [40]. The LEaP program was used to generate a library defining the ligand residue topology. The final system was built, neutralized, and solvated with two Na^+ counter-

ions using the Dabble program [41]. Before starting the simulation process, the studied covalent complexes were placed within a box of TIP3P water molecules with 10 Å distance from the protein [42]. Particle mesh Ewald (PME) method was implemented within Amber14, with direct space and Van der Waals cut-off of 12 Å, to obtain long-range electrostatic interaction. To further relax the complex and remove potential steric clashes, each system was energy minimized for a total of 7500 steps (2500 steps of steepest descent and 5000 conjugate gradient steps) with a 10 kcal/mol/$Å^2$ restraint conditions applied. The systems were heated for 30 ps from 0 to 300 K with an additional 7-ns equilibration performed at a 4-fs integration time step. MD simulation production runs of 250 ns were performed for each system during which the SHAKE algorithm was used to constrict all atomic hydrogen bonds at a 4-fs integration time step [43]. The computational methodology concerning the covalent systems was based on our previously reported [44]. The CPPTRAJ and PTRAJ modules [45] of AMBER14 package were used to analyze resulting trajectories for root mean square deviation (RMSD), root mean square fluctuation (RMSF), solvent accessible surface area (SASA), the radius of gyration (ROG), and secondary structure analysis. The data were expressed in mean ± standard deviation. The obtained data were plotted using Microcal Origin tools [46] and Maestro Schrödinger software [36].

5.2.4 Clustering and Principal component analysis

A principal component analysis (PCA) was performed to describe the internal motion of the complexes using the Bio3D package in R. The process involves the initial construction of the covariance matrix (C) from (x,y,z) coordinate positions of the C-atoms as representatives of residues (N), generating a large matrix of dimension 3N_3N. The covariance matrix was further diagonalized to obtain eigenvectors based on related eigenvalues. This was then projected on the first three eigenvectors (PC1, PC2, and PC3).

5.3 Result and Discussion

To understand the molecular behavior associated with the preferential binding mechanism of 8-*N*-benzyladenosine to Lys-56 over Cys-17 at HPS72-NBD, covalent molecular dynamics simulation was employed to study the structural and dynamical changes of the above two covalent binding models. Post molecular dynamics analyses were carried out covering different aspects, including dynamic conformational stability (RMSD, RMSF), dynamic system variations (SASA, secondary structure analysis, PCA).

5.3.1 Overall Structural Stability and Dynamics of the Simulated Systems:

To assess the structural stability of the studied systems, the root mean square deviation (RMSD) was calculated based on C-α atoms for the Apo and covalent complexes over the 250ns simulation. The systems achieved stable equilibration after 50ns. The recorded average RMSD values for the entire frames of the systems were 2.1 ± 0.27 Å, 3.1 ± 0.42 Å, and 4.8 ± 0.54 Å for Apo-NBD, Lys-NBD, and Cys-NBD, respectively (Figure S 4.1). The Apo system showed a lower RMSD average value, whereas the Lys-NBD induces a relatively more stable protein conformation compared with the Cys-NBD complex.

The covalent binding effects of 8-*N*-benzyladenosine towards amino acid residues of the HSP72-NBD when covalently bonded to Lys-56 or Cys-17 were analyzed by the root mean square fluctuation (RMSF). The mobility of HSP72-NBD C-α RMSF was computed and averaged over 250ns to observe inhibitor binding effects towards HSP72-NBD protein structural flexibility. The computed average atomic fluctuations for Apo-NBD, Lys-NBD, and Cys-NBD were 1.2 ± 0.52 Å, 1.6 ± 0.77 Å, and 2.1 ± 0.94 Å, respectively. The plotted results in Figure 5.4 indicate overall relative lower residue fluctuation in the Lys-NBD complex system compared with the Cys-NBD complex system. This observation suggests that the covalent binding of 8-*N*-benzyladenosine to HSP72-NBD via lysine relatively decreases the overall protein flexibility in contrast to 8-N-benzyladenosine binding to HSP72-NBD via cysteine residue.

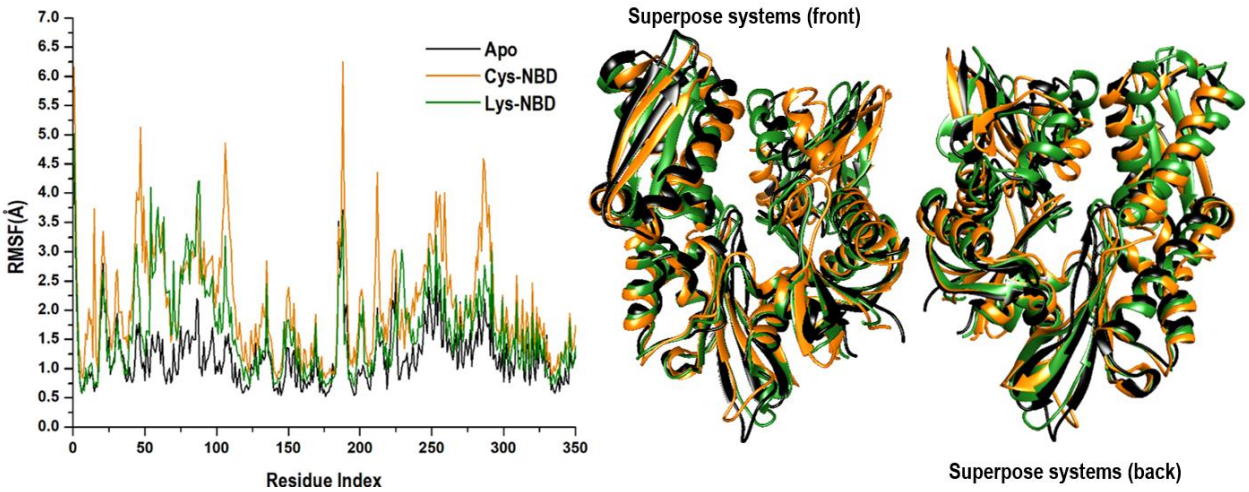


Figure 5.4 The time evolution RMSF of each residue of the protein C-α atoms over 250 ns for Apo (black color), and Cys-NBD (orange color), and Lys-NBD (green color). Superposed crystal structures of the studied systems are also illustrated to show differences in fluctuations.

Furthermore, to obtain insight into how the protein surface interacts with solvent molecules and how it relates to the compactness of the hydrophobic protein core, the solvent-accessible surface area (SASA) of the protein upon ligand binding was calculated (Figure 5.5). This was accomplished by calculating the surface area of the protein visible to solvent across the 250 ns MD simulation, which is vital for biomolecular stability [47]. The overall SASA indicates that the Lys-NBD protein surface is relatively less exposed to solvent molecules compared with Cys-NBD inhibition. The computed average SASA values for Apo-NBD, Lys-NBD, and Cys-NBD inhibition systems were 17374.3 ± 327.7 Å, 17975.5 ± 518.2 Å, and 18289.5 ± 680.2 Å, respectively.

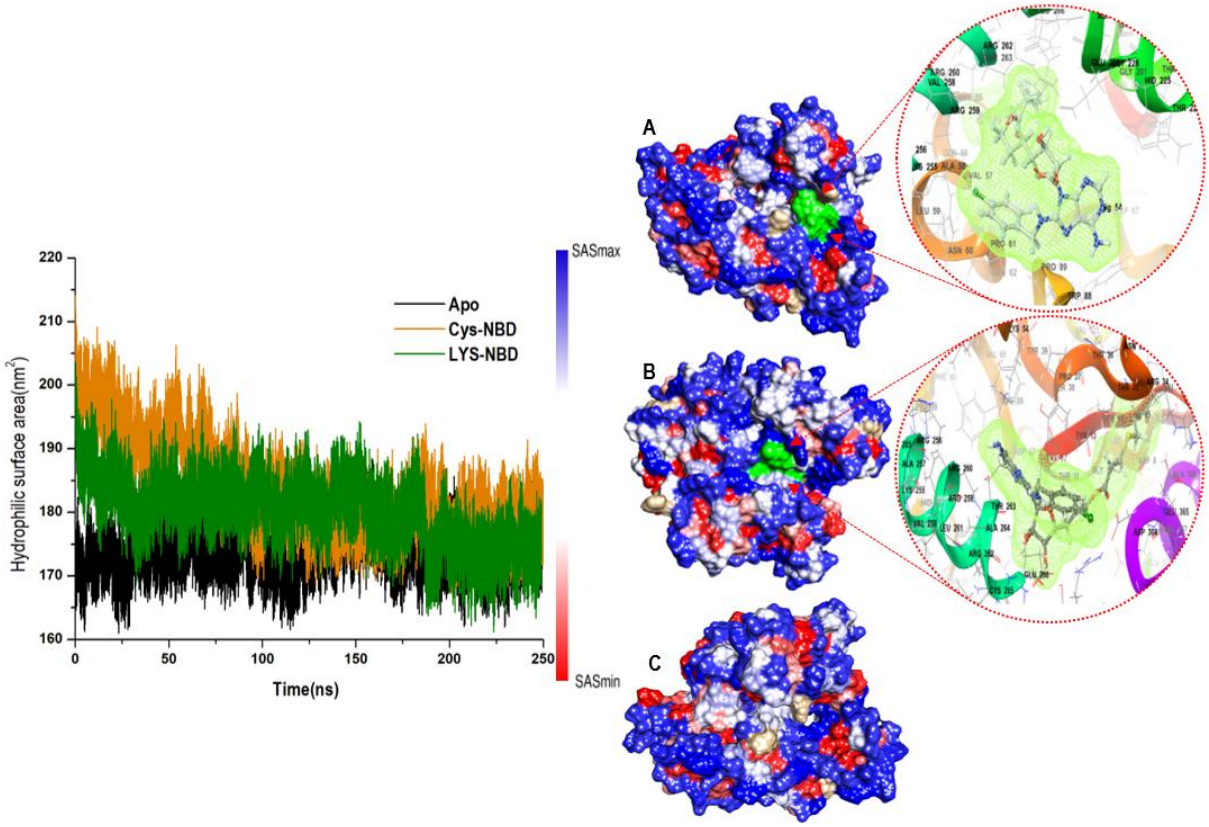


Figure 5.5 Solvent accessible surface area (SASA) backbone atoms relative to the starting minimized structure over 250 ns for Apo, Cys-NBD, and Lys-NBD. Areas with higher SASA values and lower SASA values are shown in blue and red, respectively, for Cys-NBD (**A**), and Lys-NBD (**B**) and Apo (**C**).

5.3.2 Analysis of Secondary Structure Variation

To further gain an additional structural understanding of the cysteine-NBD targeting, and the lysine-NBD targeting of HSP72, the DSSP classification for each amino acid was calculated (Figure 5.6). The segment of residues 50–70, which was suggested in the stabilization of the closed formation [48–50], remains as α-helix-para-turn (green-blue-brown) conformation

throughout the simulation when bound to cysteine. In contrast, it becomes 3_{10}helices-bend (dark green-red) when bound to lysine is binding. The α-helix in this segment disappears almost entirely and thus potentially destroys the NBD as it is no longer in close proximity. The solvent-exposed salt-bridge between Glu-268 and Lys-56 is absent [51].

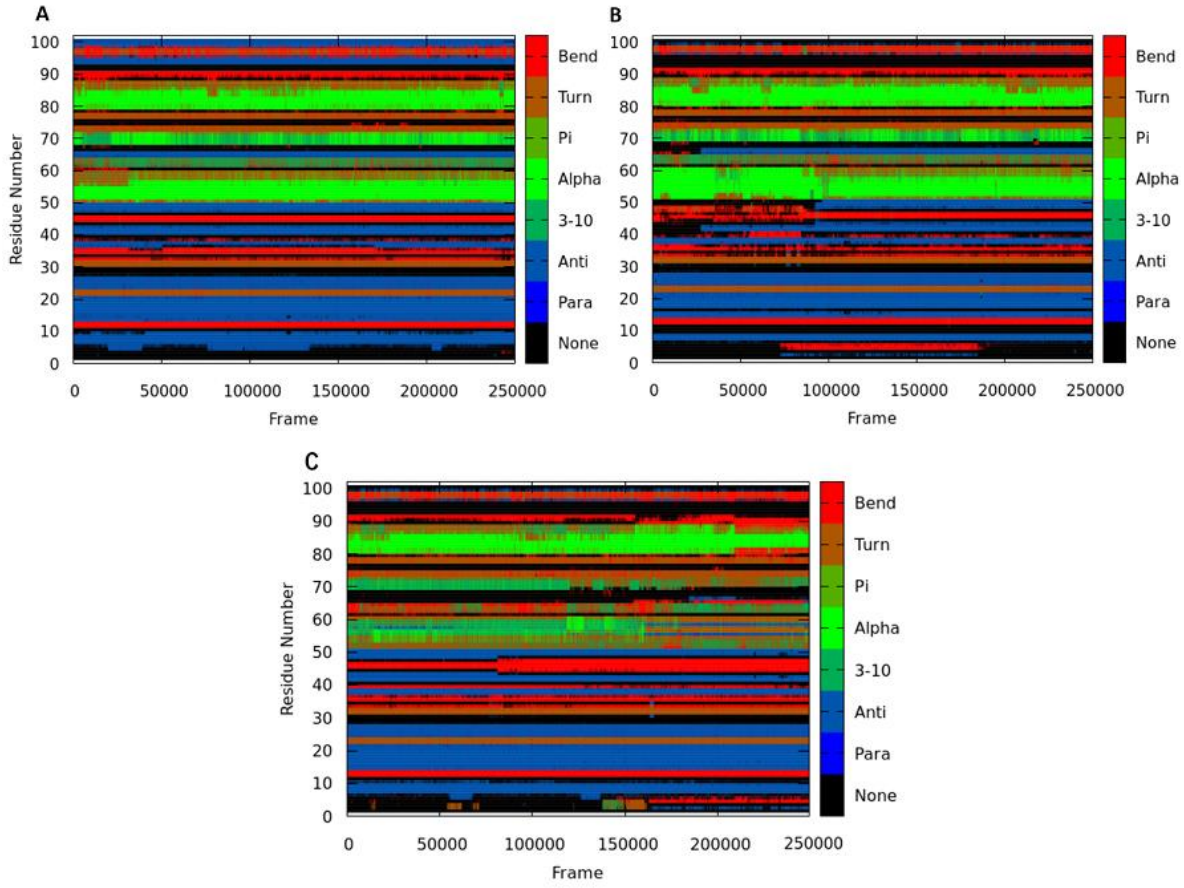


Figure 5.6 DSSP classification for the time evolution of the secondary structural elements for Apo (**A**), Cys-NBD (**B**), and Lys-NBD (**C**).

5.3.3 Conformational clustering and principal component analysis

Principal component analysis and clustering were performed for Cys-17-HSP72 and Lys-56-HSP72 8-*N*-benzyladenosine covalent complexes to observe the overall concerted motion of HSP72-NBD protein. The PCA provides insight into the conformational changes of macromolecules, such as proteins [52]. The structural distribution of conformational changes

and the proportion of variance of the captured eigenvectors are shown in Figure 5.7A, C. The first three principal components accounted for 67.8% and 53.2% of the total variance observed in the MD trajectories for Cys-17-HSP72 and Lys-56-HSP72 covalent complexes, respectively. The magnitude of principal component 1 (PC1) was observed to be the highest for the Cys-17-HSP72 complex (51.2%); however, a relatively lower PC1 of 30.8% was observed for Lys-56-HSP72 covalent complex. The observed PC1 variance suggests that the inhibitor-induced radical conformational changes in the protein structure when couple to Cys-17, resulting in a dynamic rearrangement of the IIA and IA helices of HSP72-NBD (Figure 5.7B). The clustering analysis also shows conformational distribution variance along the first, second, and third principal components with each dot representing a single complex conformation.

The clustering and principal component analyses suggest that HSP72 undergoes a large conformational change in the NBD when the inhibitor is coupled to Cys-17 compared with Lys-56 (Figure 5.7). The coupling of the inhibitor to cysteine-17, therefore, shows intrinsically more dynamic residue mobility largely in the IIA α-helix region compared with the lysine-56 inhibitor complex (Figure 5.7B, D).

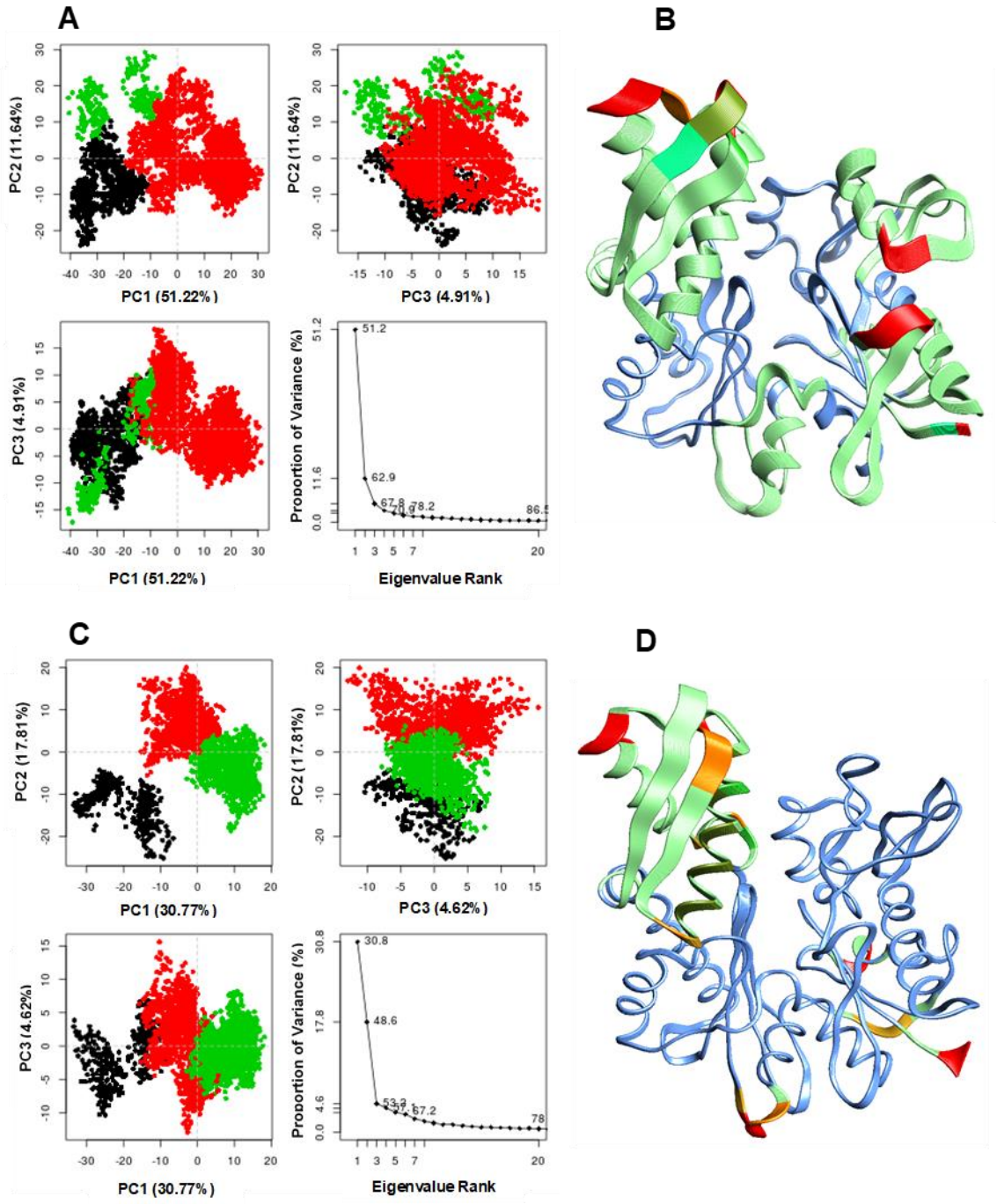


Figure 5.7 Clustering and principal component analysis on 250ns of equidistance conformations using the Bio3D package in R. The plots show the first three eigenvectors for Cys-17 HSP72 complex (**A**) and Lys-56 HSP72 complex (**C**) Conformers are colored according to the k-means clustering: cluster 1, black; 2, red; 3, green. dominant motions and captured eigenvector variance; and residue mobility for Cys-17 HSP72 complex (**B**) and Lys-56 sHSP72 complex (**D**). The color scale from blue, green, to red depicts low to high atomic displacements.

5.3.4 Understanding structural dynamics upon 8-N-benzyladenosine coupling to Lysine and Cysteine of HSP72-NBD α-helices.

To further elucidate the structural basis of the preferential covalent coupling of 8-N-benzyladenosine with Lys-56 over Cys-17, we compared the subdomain helix dynamic motions throughout the simulation. The observed structural changes in the helix motion show a different mechanism of the NBD characterized by an opened conformation and a closed conformation of the IIB and IIA subdomains for Cys-17 HSP72 complex and Lys-56 HSP72 complex, respectively (Figure 5.8). The Lys-56-HSP72 8-*N*-benzyladenosine complex binding results in the closure of the two helices for the subdomain IIB and IIA, thus, inducing a closed conformation. On the other hand, an open conformation of the two helices for the subdomain IIB and IIA was observed for the Cys-17-HSP72 8-*N*-benzyladenosine complex.

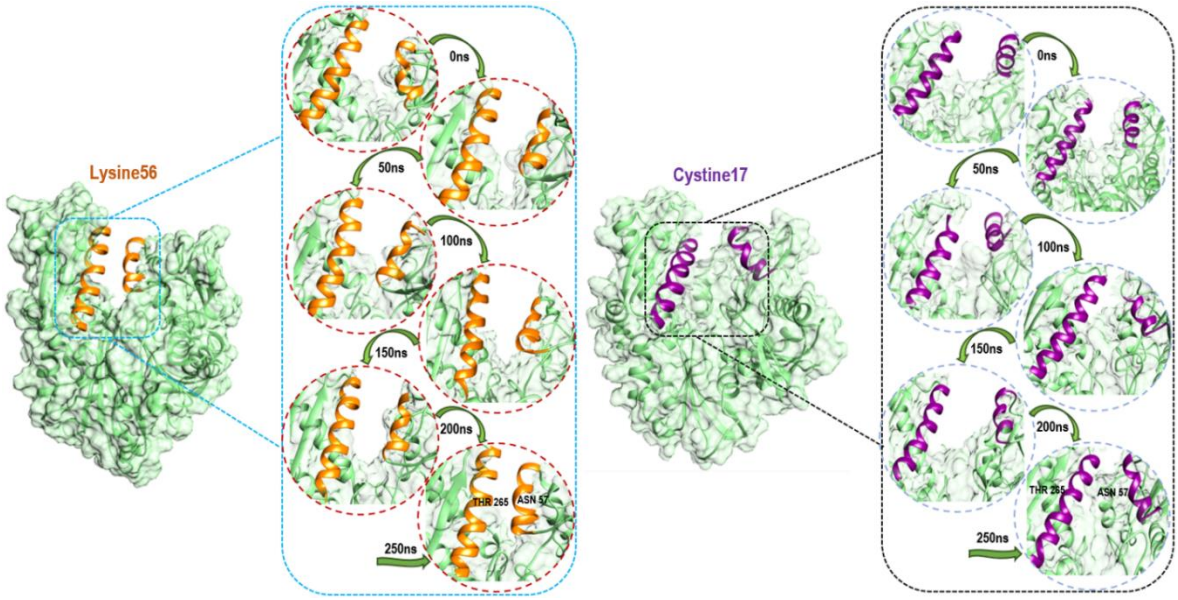


Figure 5.8 Conformational dynamic comparison of α-helix coupling upon 8-*N*-benzyladenosine binding to Lys56 **(**Orang color**)** and Cys17 **(**violet color**)** showing the closed conformation of HSP71-NBD.

The inhibition of HSP72-NBD via lysine-56 decreased the inter-residue distance between two opposite IIB and IIA subdomain helix residues Thr-265 and Asn-57 with an average distance of 9.24 Å. However, a distance of 26.14 Å was observed between residues Thr-265 and Asn-57 via the cysteine-17 inhibition model (Figure 5.9). The coupling of 8-*N*-benzyladenosine to Lys-56 over Cys-17 residue strengthens the interaction affinity of 8-*N*-benzyladenosine and induced a closed conformation of the IIB and IIA nucleotide-binding subdomain.

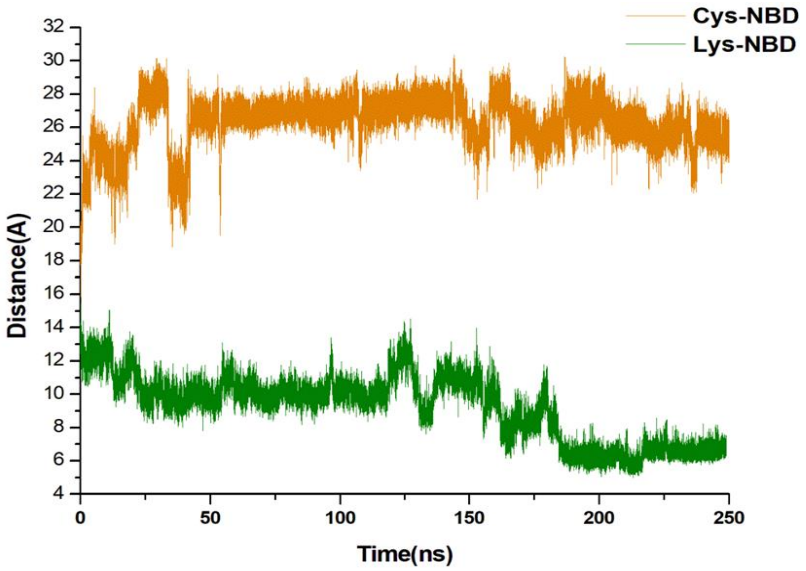


Figure 5.9 Distance between C- α residues involving THR265, and ASN57 α-helices of the NBD. The average distances were found to be 9.24 Å and 26.14 Å, respectively, for Lys-NBD and Cys-NBD conformations.

Further analysis of the 8-*N*-benzyladenosine interaction mechanism with Lys-56 and Cys-17 reveals a salt-bridge interaction between Lys-56 and Glu-268 in the Lys-56-HSP72-NBD complex (Figure 5.10A), which is absent in the Cys-17-HSP72-NBD complex (Figure 5.10B). The observed Lys-56-8-*N*-benzyladenosine interaction and the closed conformation is nearly identical to the closed conformation of ADP/Pi bound to HSP72-NDB [51]. In Cys-17 inhibitor coupling, the bulk of the inhibitor is sandwiched between Glu-268 and Lys-56 of the two helices, which prevent the possible formation of the Glu-268 and Lys-56 salt bridge inducing an open conformation (Figure 5.10B).

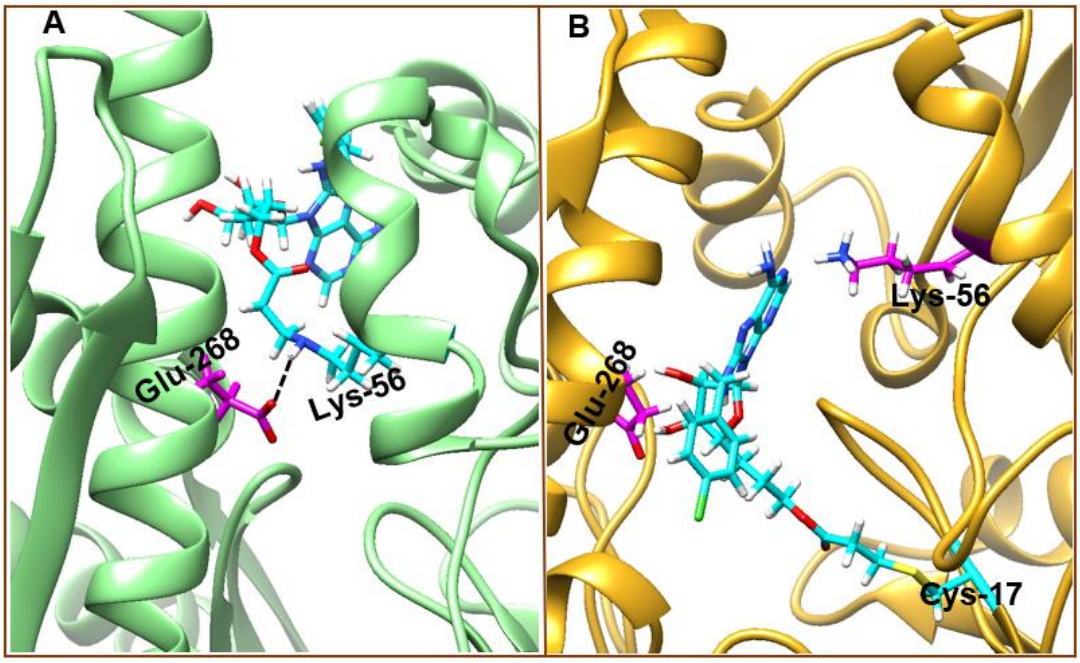


Figure 5.10 Lys-56-8-N-benzyladenosine (**A**) and Cys-17-8-N-benzyladenosine (**B**) coupled HSP72-NBD irreversible interaction mechanism. Crucial salt-bridge between Glu-268 and Lys-56 residues were missing in the Cys-17 coupled conformation.

5.4 Conclusion

The human heat shock protein 72 (HSP72) represents a vital therapeutic target during the critical stages of oncogenesis and progression of human cancers. In this study, we performed covalent molecular dynamic simulation followed by extensive analyses to decipher the structural basis and conformational dynamics associated with the unexpected preferential coupling of 8-*N*-benzyladenosine to lysine-56 over cysteine-17 in HSP72-NBD models. The results reveal that the irreversible binding of 8-*N*-benzyladenosine to lysine-56 over cysteine-17 in HSP72-NBD represents the most stable conformation with minimal intrinsic dynamics. Clustering and PCA showed that the first three principal components accounted for 67.8% and 53.2% of the total variance for Cys-17-HSP72 and Lys-56-HSP72 covalent complexes, respectively. The magnitude of principal component 1 (PC1) was observed to be the highest for the Cys-17-HSP72 complex (51.2%); however, a relatively lower PC1 of 30.8% was observed for Lys-56-HSP72 covalent complex.

The conformational dynamics analysis further reveals what the experimental study could not capture and explain, that the coupling of 8-*N*-benzyladenosine to Lysine-56 induces a closed conformation of the IIB and IIA α-helices of the nucleotide-binding subdomain. In contrast, an open conformation was observed in coupling to Cysteine-17 residue. Interestingly, the close confirmation maintained the crucial salt-bridge between Glu-268 and Lys-56 residues, which strengthens the interaction affinity of 8-*N*-benzyladenosine nearly identical to ADP/P_i bound to the HSP72-NBD. It is rare for non-catalytic lysine residues to form a covalent bond with an inhibitor. The recent unexpected covalent formation between lysine-56 and 8-N-benzyladenosiness would assist with the design of more potent and highly selective covalent inhibitors for HSP72 with the potential to overcome drug resistance challenges and represent a novel therapeutic approach for inhibiting HSP72 oncoprotein.

Acknowledgments

The authors acknowledge the College of Health Science of University of KwaZulu-Natal for funding and the Centre for High-Performance Computing (CHPC), Cape Town, South Africa, for computational resources (www.chpc.ac.za).

Conflict of interest

The authors declare no conflict of interest.

References

1. Mathew A, MORIMOTO RI. Role of the Heat-Shock Response in the Life and Death of Proteins. *Ann. N. Y. Acad. Sci.* 851(1 STRESS OF LIF), 99–111 (1998).

2. Nollen EAA, Morimoto RI. Chaperoning signaling pathways: molecular chaperones as stress-sensing "heat shock" proteins. *J. Cell Sci.* 115(Pt 14), 2809–16 (2002).

3. Whitley D, Goldberg SP, Jordan WD. Heat shock proteins: a review of the molecular chaperones. *J. Vasc. Surg.* 29(4), 748–51 (1999).

4. Hideshima T, Bradner JE, Wong J, *et al.* Small-molecule inhibition of proteasome and aggresome function induces synergistic antitumor activity in multiple myeloma. *Proc. Natl. Acad. Sci. U. S. A.* 102(24), 8567–72 (2005).

5. Wu WKK, Sakamoto KM, Milani M, *et al.* Macroautophagy modulates cellular response to proteasome inhibitors in cancer therapy. *Drug Resist. Updat.* 13(3), 87–92 (2010).

6. Zhu K, Dunner K, McConkey DJ. Proteasome inhibitors activate autophagy as a cytoprotective response in human prostate cancer cells. *Oncogene*. 29(3), 451–62 (2010).

7. Wu B, Hunt C, Morimoto R. Structure and expression of the human gene encoding major heat shock protein HSP70. *Mol. Cell. Biol.* 5(2), 330–341 (1985).

8. Kampinga HH, Hageman J, Vos MJ, *et al.* Guidelines for the nomenclature of the human heat shock proteins. *Cell Stress Chaperones*. 14(1), 105–11 (2009).

9. Nawrocki ST, Carew JS, Dunner K, *et al.* Bortezomib inhibits PKR-like endoplasmic reticulum (ER) kinase and induces apoptosis via ER stress in human pancreatic cancer cells. *Cancer Res.* 65(24), 11510–9 (2005).

10. Jäättelä M. Heat shock proteins as cellular lifeguards. *Ann. Med.* 31(4), 261–271 (1999).

11. Nylandsted J, Rohde M, Brand K, Bastholm L, Elling F, Jaattela M. Selective depletion of heat shock protein 70 (Hsp70) activates a tumor-specific death program that is independent of caspases and bypasses Bcl-2. *Proc. Natl. Acad. Sci.* 97(14), 7871–7876 (2000).

12. Nylandsted J, Wick W, Hirt UA, *et al.* Eradication of glioblastoma, and breast and colon carcinoma xenografts by Hsp70 depletion. *Cancer Res.* 62(24), 7139–42 (2002).

13. Powers M V., Jones K, Barillari C, Westwood I, Montfort RLM van, Workman P. Targeting HSP70: The second potentially druggable heat shock protein and molecular chaperone? *Cell Cycle*. 9(8), 1542–1550 (2010).

14. Mayer MP, Bukau B. Hsp70 chaperones: Cellular functions and molecular mechanism. *Cell. Mol. Life Sci.* 62(6), 670–684 (2005).

15. Zhuravleva A, Clerico EM, Gierasch LM. An interdomain energetic tug-of-war creates the allosterically active state in Hsp70 molecular chaperones. *Cell*. 151(6), 1296–307 (2012).

16. Penkler D, Sensoy Ö, Atilgan C, Tastan Bishop Ö. Perturbation–Response Scanning Reveals Key Residues for Allosteric Control in Hsp70. *J. Chem. Inf. Model.* 57(6),

1359–1374 (2017).

17. Sharma D, Masison DC. Hsp70 structure, function, regulation and influence on yeast prions. *Protein Pept. Lett.* 16(6), 571–81 (2009).

18. Chiappori F, Merelli I, Colombo G, Milanesi L, Morra G. Molecular mechanism of allosteric communication in Hsp70 revealed by molecular dynamics simulations. *PLoS Comput. Biol.* 8(12), e1002844 (2012).

19. Leu JI-J, Pimkina J, Frank A, Murphy ME, George DL. A small molecule inhibitor of inducible heat shock protein 70. *Mol. Cell*. 36(1), 15–27 (2009).

20. Brodsky JL. Selectivity of the molecular chaperone-specific immunosuppressive agent 15-deoxyspergualin. *Biochem. Pharmacol.* 57(8), 877–880 (1999).

21. Hassan AQ, Kirby CA, Zhou W, *et al.* The Novolactone Natural Product Disrupts the Allosteric Regulation of Hsp70. *Chem. Biol.* 22(1), 87–97 (2015).

22. O'Regan L, Sampson J, Richards MW, *et al.* Hsp72 is targeted to the mitotic spindle by Nek6 to promote K-fiber assembly and mitotic progression. *J. Cell Biol.* 209(3), 349–358 (2015).

23. Halgren TA. Identifying and Characterizing Binding Sites and Assessing Druggability. *J. Chem. Inf. Model.* 49(2), 377–389 (2009).

24. Yun C-H, Mengwasser KE, Toms A V, *et al.* The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc. Natl. Acad. Sci. U. S. A.* 105(6), 2070–5 (2008).

25. Singh J, Petter RC, Baillie TA, Whitty A. The resurgence of covalent drugs. *Nat. Rev. Drug Discov.* 10(4), 307–317 (2011).

26. D.S. J, E. W, B.F. C. Strategies for discovering and derisking covalent, irreversible enzyme inhibitors. *Future Med. Chem.* (2010).

27. Liu Q, Sabnis Y, Zhao Z, *et al.* Developing Irreversible Inhibitors of the Protein Kinase Cysteinome. *Chem. Biol.* 20(2), 146–159 (2013).

28. Pettinger J, Le Bihan YV, Widya M, van Montfort RLM, Jones K, Cheeseman MD. An Irreversible Inhibitor of HSP72 that Unexpectedly Targets Lysine-56. *Angew. Chemie - Int. Ed.* 56(13), 3536–3540 (2017).

29. Berman HM, Battistuz T, Bhat TN, *et al.* The Protein Data Bank. *Biol. Crystallogr.* 58, 899–907 (2002).

30. Schlecht R, Scholz SR, Dahmen H, *et al.* Functional analysis of Hsp70 inhibitors. *PLoS One*. 8(11), 1–12 (2013).

31. Pettersen Ef Fau - Goddard TD, Goddard Td Fau - Huang CC, Huang Cc Fau - Couch GS, *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* (2004).

32. Windows MM V, X MOS. Molegro Molecular Viewer User Manual. , 145 (2011).

33. Kusumaningrum S, Budianto E, Kosela S, Sumaryono W, Juniarti F. The molecular docking of 1,4-naphthoquinone derivatives as inhibitors of Polo-like kinase 1 using Molegro Virtual Docker. *J. Appl. Pharm. Sci.* 4(11), 47–53 (2014).

34. Science M. First Principles Studies on the Interaction of O 2 with X @ Al 12 (X 5 Al 2 , P 1 , C , Si) Clusters. 12 (2010).

35. Trott O, Olson A. Autodock vina: improving the speed and accuracy of docking. *J. Comput. Chem.* 31(2), 455–461 (2010).

36. Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided. Mol. Des.* 27(3), 221–234 (2013).

37. Mhlongo NN, Ebrahim M, Skelton AA, Kruger HG, Williams IH, Soliman MES. Dynamics of the thumb-finger regions in a GH11 xylanase Bacillus circulans: comparison between the Michaelis and covalent intermediate. *RSC Adv.* 5(100), 82381–82394 (2015).

38. Ramharack P, Oguntade S, Soliman MES. Delving into Zika virus structural dynamics-a closer look at NS3 helicase loop flexibility and its role in drug discovery. *RSC Adv.* 7(36), 22133–22144 (2017).

39. Götz AW, Williamson MJ, Xu D, Poole D, Le Grand S, Walker RC. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* 8(5), 1542–1555 (2012).

40. Lindorff-Larsen K, Piana S, Palmo K, *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins*. 78(8), 1950–8 (2010).

41. Betz R. Dabble Zendo. (2017).

42. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79(2), 926–935 (1983).

43. Gonnet P. P-SHAKE: A quadratically convergent SHAKE in O (n2). *J. Comput. Phys.* (2007).

44. Khan S, Bjij I, Betz RM, Soliman MES. Reversible versus irreversible inhibition modes of ERK2: A comparative analysis for ERK2 protein kinase in cancer therapy. *Future Med. Chem.* 10(9), 1003–1015 (2018).

45. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* 9(7), 3084–3095 (2013).

46. Seifert E. OriginPro 9.1: Scientific data analysis and graphing software - Software review. *J. Chem. Inf. Model.* 54(5), 1552–1552 (2014).

47. Pan L, Patterson JC, Deshpande A, Cole G, Frautschy S. Molecular Dynamics Study of Zn(Aβ) and Zn(Aβ)2. *PLoS One*. 8(9), 70681–70688 (2013).

48. Wijffels G, Dalrymple B, Kongsuwan K, Dixon N. Conservation of Eubacterial Replicases. *IUBMB Life*. 57(6), 413–419 (2005).

49. Richmond TJ. Solvent accessible surface area and excluded volume in proteins. Analytical equations for overlapping spheres and implications for the hydrophobic effect. *J. Mol. Biol.* 178(1), 63–89 (1984).

50. Wisniewska M, Karlberg T, Lehtiö L, *et al.* Crystal Structures of the ATPase Domains

of Four Human Hsp70 Isoforms: HSPA1L/Hsp70-hom, HSPA2/Hsp70-2, HSPA6/Hsp70B', and HSPA5/BiP/GRP78. *PLoS One*. 5(1), e8625 (2010).

51. Sriram M, Osipiuk J, Freeman B, Morimoto R, Joachimiak A. Human Hsp70 molecular chaperone binds two calcium ions within the ATPase domain. *Structure*. 5(3), 403–14 (1997).

52. Shida M, Arakawa A, Ishii R, *et al.* Direct inter-subdomain interactions switch between the closed and open forms of the Hsp70 nucleotide-binding domain in the nucleotide-free state. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 66(3), 223–232 (2010).

53. Cheeseman MD, Westwood IM, Barbeau O, *et al.* Exploiting Protein Conformational Change to Optimize Adenosine-Derived Inhibitors of HSP70. *J. Med. Chem.* 59(10), 4625–4636 (2016).

CHAPTER 6

Distinguishing the optimal binding mechanism through reversible and irreversible inhibition analysis of HSP72 protein in cancer therapy

Aimen Aljoundi[1], Ahmed El Rashedy[1], and Mahmoud E.S Soliman[1*]

[1]Molecular Bio-computation and Drug Design Laboratory

School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 78

Abstract

Over the past two decades, covalent inhibitors have gained much interest and are living up to their reputation as a powerful tool in drug discovery. Covalent inhibitors possess several significant advantages, including increased biochemical efficiency, prolonged duration and the ability to target shallow, solvent-exposed substrate-binding domains. One of the enzymes that have been both covalently and non-covalently targeted is the heat shock protein 72 (HSP72). This elevated enzyme expression in cancer cells may be responsible for tumorigenesis and tumor progression by providing chemotherapy resistance. A critical gap remains in the molecular understanding of the structural mechanism's covalent and non-covalent binding to HSP72. In this study, we explore the most optimal binding mechanism in the inhibition of the HSP72. Based on the molecular dynamic analyses, it was evident that the non-covalent complex showed more stability than the covalent complex. The covalent ligand, however, was more able to induce and stabilize the sealed conformation of the HSP72-NBD ATP binding domain throughout the. Also, the non-covalent ligand does not induce any significant conformational change as it remained close to the shape of the unbound complex; and the affinity is only dependent on the multiple hydrogen bonds in contrast to the covalent ligand. This is supported by the secondary structure elements and principal component analysis that was more dominant in the covalently inhibited complex. Covalent bond induced the α-helices sealed conformation of the HSP72-NBD; based on our findings, this will prevent other small molecules from interacting at the ATP binding site domain. Moreover, inhibition of the ATP binding domain can directly affect the ATPs protein folding mechanism of the HSP72 enzyme. The essential dynamic analysis presented in this report compliments the binding mechanism of HSP72, establishing covalent inhibition as the preferred method of inhibiting. The findings from this study may assist in the design of more target-specific HSP72 covalent inhibitors exploring the surface-exposed lysine residues.

Keywords: Covalent MD simulation; HSP72; covalent; Conjunction; Principal component analysis.

6.1 Introduction

Most small-molecule drugs are made to interact with their biological targets under balanced binding conditions, where the desired drug-protein interaction is a hasty and reversible process [1]. Depending on the concentration and the exposure length, the drug can utilize irreversible and reversible inhibition [2]. Non-covalent interactions can be divided into various groups, such as electrostatic, π-effects, van der Waals forces, and hydrophobic effects [3]. Upon the drug binding, these non-bonding interactions are leading forces for the affinity, specificity, and selectivity of the drug, leading to increasing the stability of the Drug-protein complex [4]. Even though these interactions optimization has mostly dominated the advance in the drug discovery process, including covalent modification of protein targets, represent an emerging filed in the drug development process [5–7].

Covalent bonds vary from non-covalent interaction in that electrons are shared between two contributing atoms; therefore, the covalent bonds are much stronger than the non-covalent forces [8]. Targeted covalent inhibitors typically react via a two-step mechanism of action (Figure 6.1). Firstly, the target protein enzyme (E) interacts reversibly with covalent Drug/ligand (L) to generate a non-covalent enzyme-ligand complex (EL). The tendency of the reversible complex recognizes by the free concentration of the drugs and the equilibrium constant K_i. The reversible complex can then undergo covalent bond formation as identified by first-order kinetic, to generate the covalent complex (E-L) by inducing conformational changes in the complex.

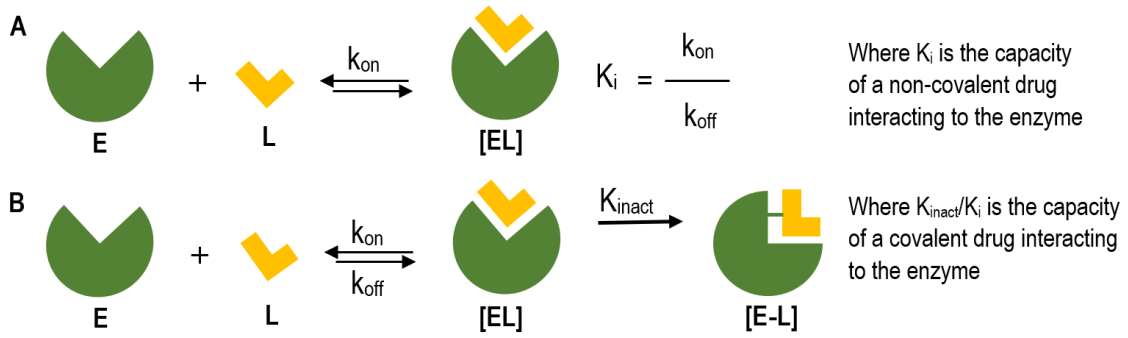


Figure 6.1 Drug-enzyme interaction. Association of a Ligand (L) interacting with Enzyme (E) in **(A)** conventional non-covalent and **(B)** covalent mode of action. Where is L: Ligand; E: Enzyme.

HSP72 is induced in an HSF1-dependent manner when the cell is undergoing stress and is overexpressed in several cancer cell types [9]. This overexpression is correlated with metastasis, poor prognosis, and resistance to chemotherapy in patients. Thus, inhibition of HSP72 is considered to be a successful pathway for anti-tumor therapy [10]. All the different functions of Hsp70s are accomplished through a transient chaperone interaction with substrate proteins through its C-terminal substrate-binding domain (SBD) [11]. This interaction is allosterically controlled by the nucleotide bound to the N-terminal nucleotide binding domain (NBD). In the nucleotide-free and ADP bound state, the affinity for substrates is high, but substrate association and dissociation rates are low. ATP binding to the NBD increases association and dissociation rates by orders of magnitude, thereby decreasing the substrates affinity by 10 to 400-fold [12–14]. Hence, the nucleotide-binding domain's inhibition is considered one of the most promising strategies for HSP72 function inhibition [15]. The NBD consists of two adjacent lobes (lobe I and lobe II), which form a deep nucleotide groove connected to the base (Figure 6.2). Each lobe consists of two subdomains (IA, IIA, IB, and IIB) [16,17]. Domains IB and IIB are linked to IA and IIA, respectively, by flexible hinges and control access to the nucleotide-binding sites [18].

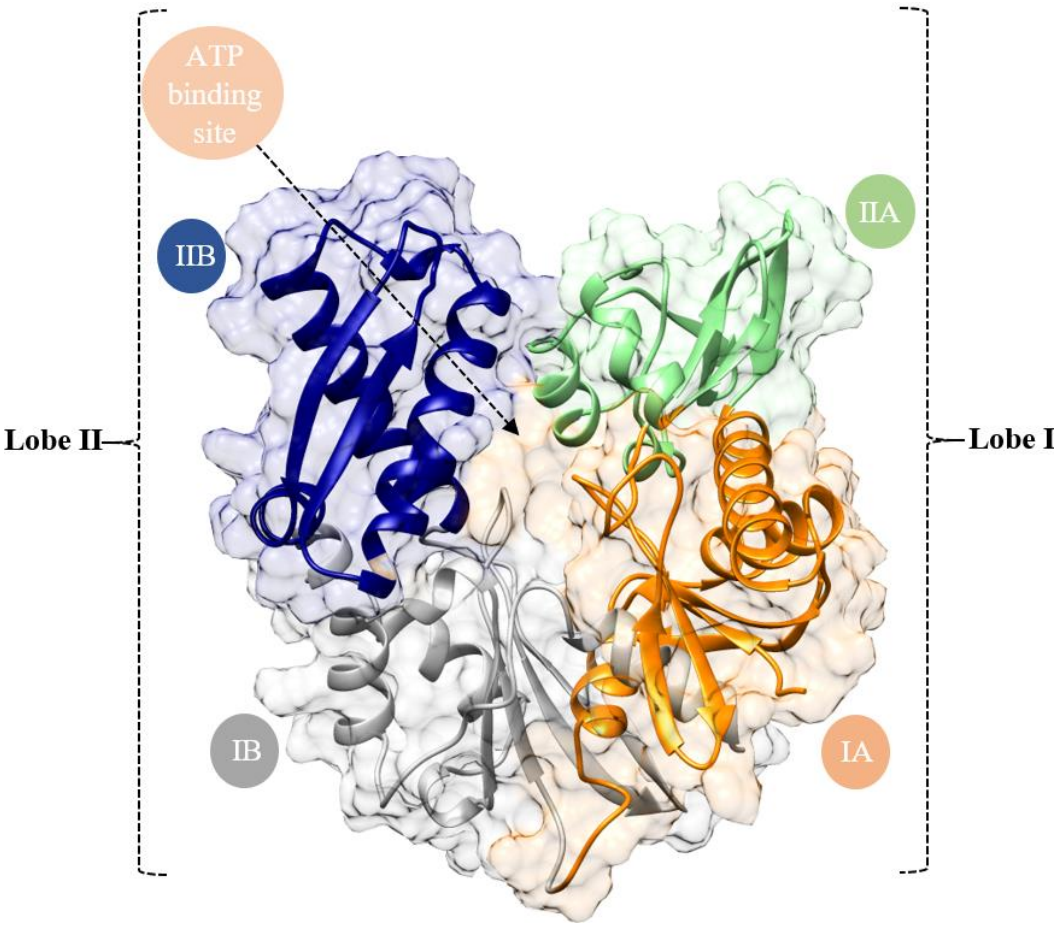


Figure 6.2 The 3-D crystal structure of the HSP72-NBD protein (PDB code: 5MKS). The IIB, IB, IIA, and IA subdomains are shown in blue, grey, orange, and green.

One of the enzymes that have been both covalently and non-covalently targeted is the Heat shock proteins (HSP72) chaperone, which play a significant role in the unfolding protein response organisation by clearance of damaged cells induced by protein aggregation and proteotoxicity [19]. Recently, Pettinger et al. (2017), fluorescence polarization (FP) assay, and crystallography had confirmed that the 8-N-benzyladenosine could irreversibly target the non-catalytic nucleophilic lysine 56 residue on HSP72 [20]. On the other hand, the adenosine-derived inhibitors can reversibly target the ATPase binding domain of the HSP72 [21]. The crucial difference between the two ligands is their ability to induce and stabilize the closed conformation of the HSP72-NBD [22]. Based on the discrepancies involving covalent versus non-covalent forces and the ambiguity surrounding the mechanistic inhibition of HSP 72

enzyme, we were prompted to study the difference in the binding mechanism between covalent ligand 3-[(2~{R},3~{S},4~{R},5~{R})-5-[6-azanyl-8-[(4-chlorophenyl)methylamino]purin-9-yl]-3,4-bis(oxidanyl)oxolan-2-yl]propyl prop-2-enoate **(TI8)** [20], and non-covalent ligand 4-{(2R,3S,4R)-5-[(R)-6-Amino-8-(3,4-dichloro-benzylamino)-purin-9-yl]-3,4-dihydroxy-tetrahydro-furan-2-ylmethoxymethyl}-benzonitrile **(3FD)** [21] (Figure 6.3).

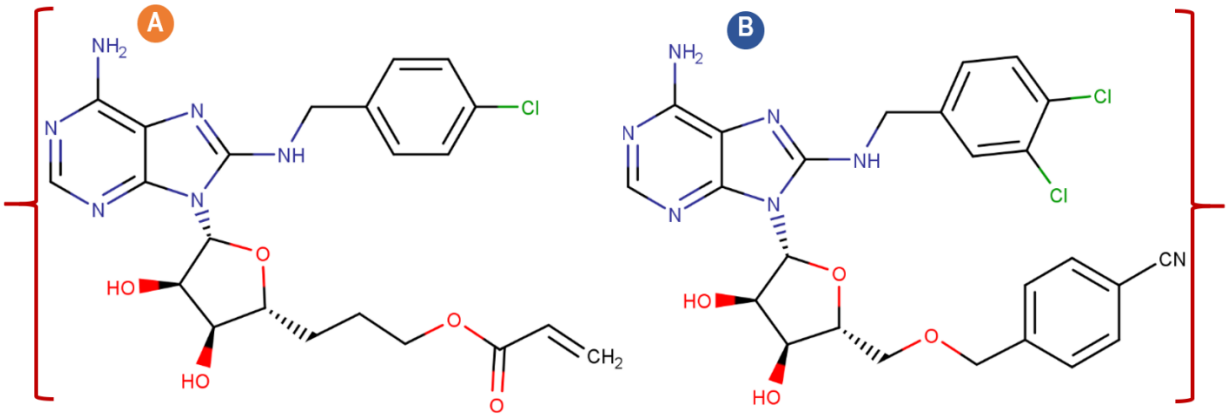


Figure 6.3 2D structure of [A] TI8 and [B] 3FD

In this study, we aim to establish the most optimal binding mechanism amongst covalent and non-covalent inhibition of HSP72. This was accomplished by molecular dynamic simulation, which compared the structural modification of covalent and non-covalent inhibition of HSP72 by comparing two modes of inhibition of HSP72, covalent (irreversible) and non-covalent (reversible) with respect to its APO (apoenzyme correspond to the free form of the enzyme) form in order to gain a better understanding of underlying structural dynamics and critical implications in each binding mechanism (Figure 6.4).

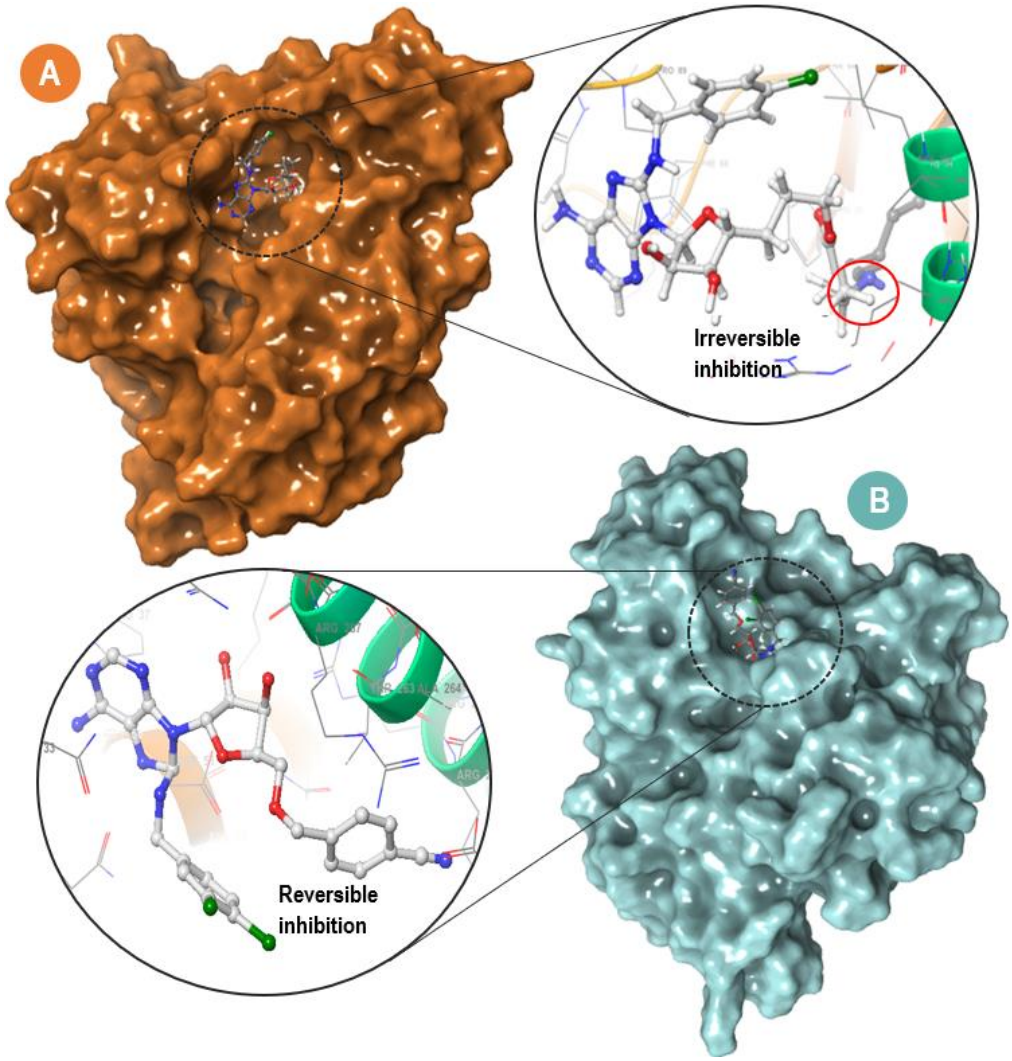


Figure 6.4 The surface view of HSP72 [**A**] covalently bonded ligand **[B]** and non-covalent bonded ligand. A close view of two inhibition conditions (red color circle shows the covalent bond).

6.2 Computational methodology

6.2.1 System preparation

The studied models were prepared based on the reported human HSP72-NBD domain crystal structure retrieved from the Protein Data Bank [23] (PDB code: 5MKS) [24] and set up using the UCSF Chimera software package [25]. The ligand was prepared using MarvinSketch 6.2.1, 2014, Molegro Molecular Viewer (MMV), and Chem-Axon (http://www.chemaxon.com) to ensure that the ligand hybridization state and proper angels were displayed [26,27].

6.2.2 Covalent docking

Covalent bonds were formed between the inhibitor and lysine-56 residues of HSP72-NBD. Before the creation of the covalent bond, a non-covalent docking was performed for the inhibitor to ensure an appropriate stable binding mode at HSP72-NBD. This initial non-covalent docking was performed using AutoDock Tools GUI, and AutoDock Vina [28] integrated with Chimera GUI. The AutoDock Tools GUI was used to define the grid box (center: -12.75, -8.433, and 5.19; Size: 18.65, 18.02, and 21.39) at the lysine binding site of the protein.The initial non-covalent docking was to appropriately position the inhibitor in the active site to allow the covalent bond to be created subsequently. The binding poses were checked to see if the two atoms that will eventually form the covalent bond between the target lysine residue and the inhibitor was within 3 Å of each other. The inhibitor binding mode that did not meet this condition to allow the bond to be formed was rejected. The Schrödinger Maestro 11.5 [29,30] was used to create the covalent bond between the inhibitor and the lysine residues . The best covalent complex pose was then selected.

The protein preparation wizard in Maestro Schrödinger was used to optimize the protonation state of the complex, adjust hydrogen atoms, cap acetyl, and methylamide neutral residues. An initial vacuum minimization was performed to resolve any steric clashes and restore normal bond lengths. The prepared receptor and covalent ligand were saved separately and taken for parameterization using the preparatory program. Furthermore, the non-covalent ligand docking was performed using AutoDock Tools GUI, and AutoDock Vina [27] integrated with Chimera GUI. The AutoDock Tools GUI was used to define the grid box (center: -15.57, -10.21, and 10.21; Size: 12.57, 12.46, and 11.95) at the ATP binding site. The APO system was run based on our previously reported protocol for non-covalent simulations [31,32].

6.2.3 Molecular dynamics simulation with Covalent inhibitor

The covalent system was exposed to an all-atom covalent molecular dynamics simulation (cMD) using the PEMED package in Amber14 [33]. Using the Antechamber module in the Amber14 to provide atom types and atomic ligand partial charges using the FF14SB forcefield [34], while the LEaP module was used to create a library of covalent ligand residue topology, whereas the Dabble program [35] was used to build, neutralize and solvate the system with two Na+ counter-ions. The covalent system was inserted within a box of TIP3P water molecules with 10Å distance from the protein [36] and minimized and heated at 300 K and to stabilize the covalent during simulation gradually. Moreover, the production cMD run lasted for 350ns with a 2-fs time step using graphics processing unit version of particle mesh Ewald (PME) molecular dynamics in AMBER14 [37]. The computational methodology concerning the covalent systems was based on our in-house protocol reported previously [38].

6.2.4 MD simulation with non-Covalent inhibitor

In a non-covalent system (ncMD), to parametrize the ligand, we run antechamber, and simulation was done as explained in the section on cMD simulation. AMBER force field FF14SB was applied to define a protein [39]. We were using the LEaP module of AMBER-14 to add the missing hydrogens and counter ions to neutralize the system for the protein. We use a box of TIP3P water model [40] to include the corresponding complexes with a distance of 10 °A among the protein solute atom and surface. The particle mesh Ewald (PME) method was carried out within Amber14, with direct space and Van der Waals cut-off of 12 Å, to obtain long-range electrostatic interaction [41]. Ahead of the non-covalent MD simulations, system minimization was conducted with the restraint potential of 10 Å to consider the solute molecule using steepest descent 500 steps, followed by steps of 1000 conjugate gradient minimization process. The simulated systems were minimized and heated at 300 K and steadily equilibrated

to ensure system stability during simulations. This was followed by MD simulation production runs of 350ns for each system to analyze the motions in the trajectory, during which the SHAKE algorithm was used to constrict all atomic hydrogen bonds [42].

For all simulation systems, The CPPTRAJ and PTRAJ modules [43] of the AMBER14 package were used to calculate resulting coordinate, and trajectories for Root mean square deviation (RMSD), dynamics cross-correlation matrix (DCCM), Root mean square fluctuation (RMSF), Solvent accessible surface area (SASA), the radius of gyration (ROG) and secondary structure analysis. The obtained data were plotted using Microcal Origin tools (www.originlab.com) [44].

6.2.5 DCCM analysis

We used dynamic cross-correlation analysis to investigate the fluctuations and movements in the backbone of the α carbon atoms [45]. The cross-correlation elements Cij between Cα atoms of residues i and j of proteins can be computed based on structural extracted from MD trajectories using the following equations [46–49] :

$$\mathbf{C_{ij}} = \frac{< \mathbf{\Delta ri}.\mathbf{\Delta rj} >}{\left(< \mathbf{\Delta r_i^2} >< \mathbf{\Delta r_j^2} >\right)^{\frac{1}{2}}} \qquad \textbf{(Equation 1)}$$

Where Δ ri is the displacement of the ith Cα aom relative to its averaged position. The Δri is the displacement of the ith Cα atom relative to its averaged position. Significantly correlated movements are symbolized by C_{ij} = 1, while C_{ij} = -1 symbolized highly anticorrelated movements in the trajectory. The divergence of motion from 1 and -1 indicates that i and j movements are anticorrelated.

The DCCM matrix was carried out using the CPPTRAJ package in Amber 14, and the matrices were plotted and evaluated using Origin software (www.originlab.com) [44].

6.2.6 Clustering and Principal component analysis

Principal Component Analysis (PCA) is a multivariate statistical approach used to systematically minimize the number of dimensions needed to characterize protein dynamics through a decomposition process that filters movements from the largest to the smallest spatial scales [50–54]. Principal component analysis (PCA) was performed to the complexes using the Bio3D package in R. The process involves the initial construction of the covariance matrix (C) from (x,y,z) coordinate positions of the C-atoms as representatives of residues (N), generating a large matrix of dimension 3N_3N. The covariance matrix was diagonalized to obtain eigenvectors based on related eigenvalues and plotted on the first three eigenvectors (PC1, PC2, and PC3).

6.3 Result and Discussion

6.3.1 Dynamic Conformational Stability and Fluctuation:

It was significant to ensure that all the systems were converged and structurally stable during the 350ns MD production run to avoid disrupted motions and simulation artifacts. So that, Cα -RMSD for the three systems was estimated (Figure S 6.1). The systems achieved stable equilibration after 75ns. The recorded average RMSD values for the entire frames of the systems were 2.15 Å, 3.15 Å, and 1.68 Å for Apo, Cov-NBD, and Non-NBD, respectively. The RMSD plot indicates that the Non-NBD complex was more stable compared to the Apo and Cov-NBD. The Cov-NBD complex did not achieve satisfactory equilibrium throughout the entire simulation because it is binding to non -catalytic nucleophile lysine residue. [20] Besides, the non-covalent binding depends on the multiple hydrogen bonds between the compound and the target protein[17].

Afterwards, we calculated the Root mean square fluctuation (RMSF) values to invested the inhibitory effect of the covalent and non-covalent inhibitor towards the amino acid residues of the HSP 72-NBD. The computed average atomic fluctuation of Apo, Non-NBD, and Cov-NBD were 1.2 Å, 1.60 Å, and 1.16 Å, respectively. The plotted results in (Figure 6.5) indicate overall relative a higher residue fluctuation at the α-helix region on the Cov-NBD compared to the Non-NBD as the Cov-NBD has induced and stabilized the closed conformation of the HSP72-NBD throughout the simulation in comparison to the Non-NBD.

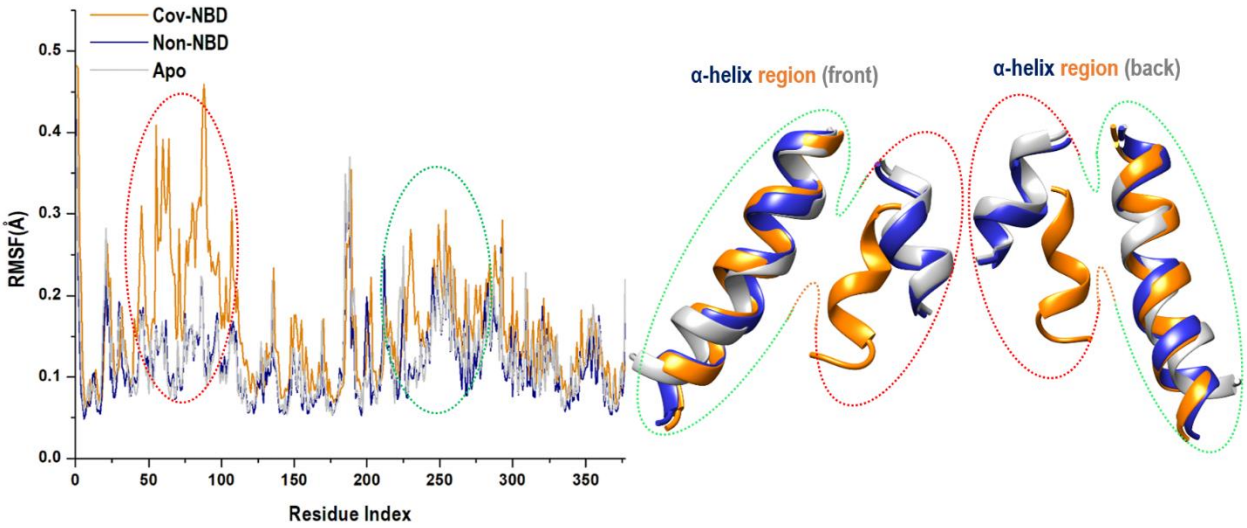


Figure 6.5 The time evolution RMSF of each residue of the protein Cα atom over 350 ns for Apo **(**Gray color**)**, and Non-NBD **(**blue color**)**, and Cov-NBD **(**orang color**)**. We are superposed of the α-helix region crystal structures of the studied systems to show differences in fluctuations in the α helix region.

The radius of gyration (RoG) is an indicator of protein compactness after ligand binding. The RoG was determined by calculating the mass-weighted Root mean square distance of the collection of atoms from their common center of mass during the molecular dynamic simulation [55]. The average ROG values were 21.52 Å, 21.61 Å, and 21.41 Å for Apo, Cov-NBD and Non-NBD

inhibition, respectively. This pattern suggested that the non-covalent interaction residues with HSP 72 showed relatively rigid, stable protein structure compared to the covalent interaction

residues (Figure S 6.2). The observed lower ROG value for Non-NBD than Cov-NBD reflects a similar pattern, as shown in RMSF and RMSD values.

Moreover, to know how the protein surface interrelates with solvent atoms and how it relays to the compactness of the hydrophobic protein core, the solvent-accessible surface area (SASA) of the protein upon ligand binding was calculated (Figure 6.6). This was accomplished by computing the surface area of the protein observable to solvent across the 350 ns MD simulation, which is vital for biomolecular stability [56]. The computed average SASA values for Apo-NBD, Cov-NBD, and Non-NBD inhibition systems were 17399.3 Å, 17677.6 Å, and 16905.8 Å, respectively. The overall SASA indicates that the Non-NBD protein surface is relatively less exposed to solvent molecules compared to Cov -NBD inhibition.

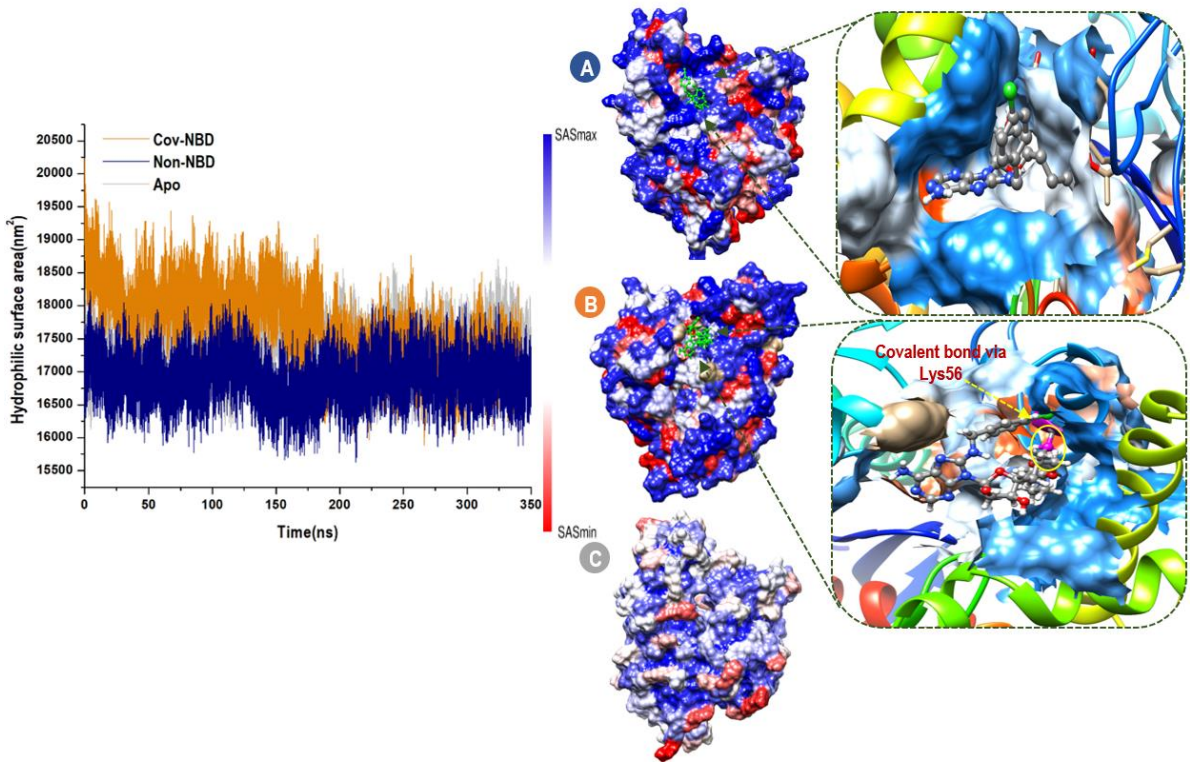


Figure 6.6 Solvent accessible surface area **(**SASA**)** backbone atoms relative to the starting minimized structure over 350ns for Apo, Non-NBD, and Cov-NBD. Areas with higher SASA values and lower SASA values are shown in blue and red, respectively, for Non-NBD **[A]**, Cov-NBD **[B],** and Apo **[C]**.

Different enzyme dynamics between the three conditions were examined using a dynamic cross-correlation matrix and plotted by observing the correlated movements of all residues.

DCM plots are represented by different colors, highly positive regions range from green to olive-green (strongly correlated), and highly negative range from light blue to black (strong anticorrelated) movement for specific residues, respectively.

As evident in Figure 6.7, The Cov-NBD system enzyme displayed less correlated motions compared to the Non-NBD complex. Moreover, the α-helices region showed a highly correlated motion in a Cov-NBD complex system compared to a Non-NBD system. These results are correlated with RMSF, which justified that the α-helices region exhibits a highly correlated motion during the simulation per time. Taken together, these results suggested that the conformation flexibility of the α-helices region was improved by covalent inhibitor binding.

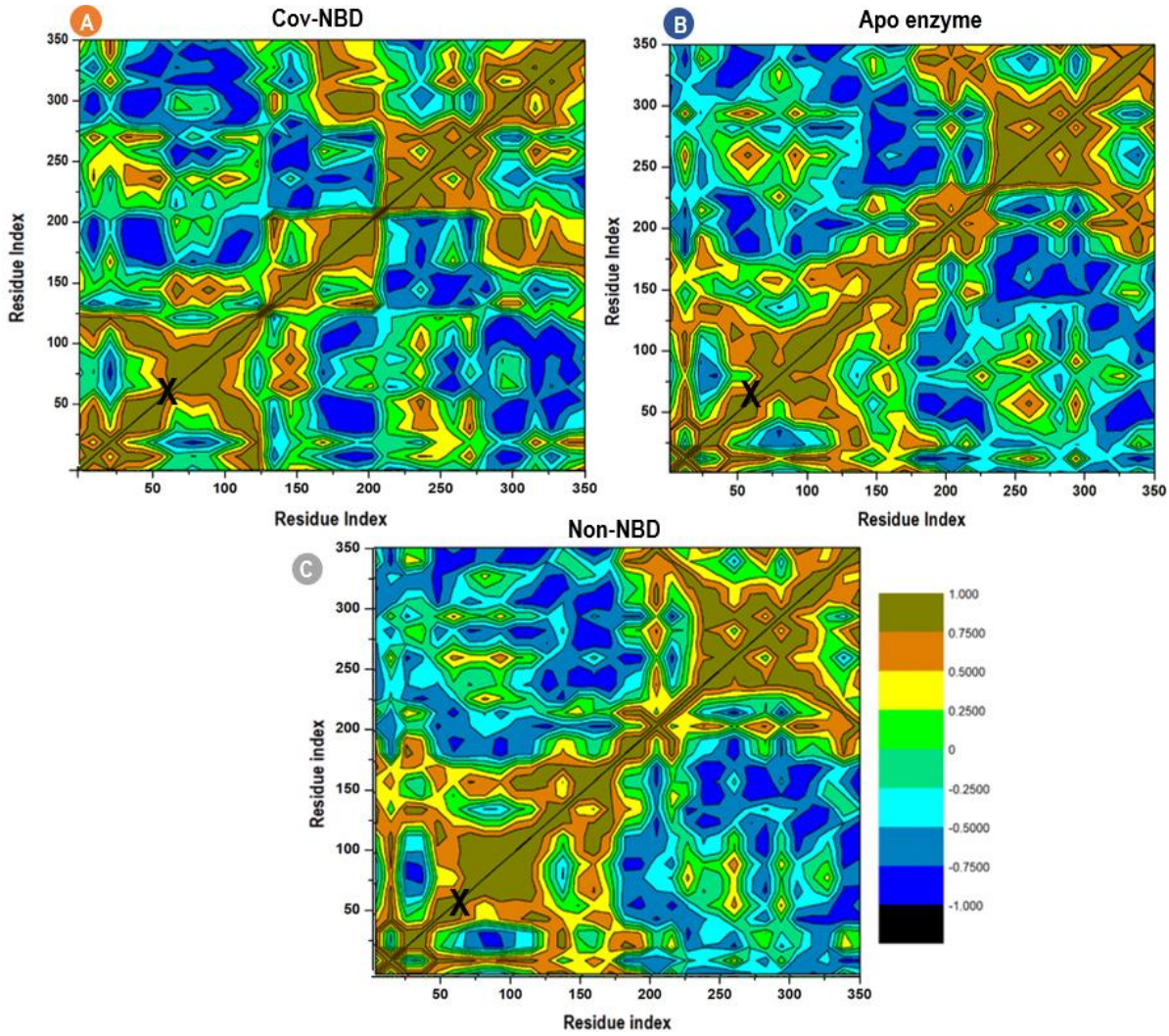


Figure 6.7 Dynamic cross-correlation matrix analyses for HSP72 covalent **(A)** and non -covalent **(C)** with unbound Apo **(B)**. Numbers closer to 1 indicate high correlation, while those closer to -1 indicate anticorrelation between pairs of residues. X denoted the α-helices site of HSP72.

6.3.2 Analysis of Secondary Structure Variation

To additionally increase the structural understanding of the Cov-NBD targeting, and the Non-NBD targeting of HSP72, the DSSP classification for each amino acid was calculated and showed its time course in (Figure 6.8). The segment of residues 50–70, which was suggested in the stabilization of the closed formation [57–59], remains as α-helix-para-turn (green-blue-brown) conformation throughout the simulation when bound to covalent binding, whereas it becomes 3_{10}helices-bend (dark green-red) when bound to non-covalent is binding. The α-helix in this segment disappears almost entirely and potentially destroys the nucleotide-binding domain as it is no longer close, which allows a less residues interaction compared to covalent inhibition. The solvent-exposed salt-bridge between glutamic acid 268 and lysine 56 is absent [22].

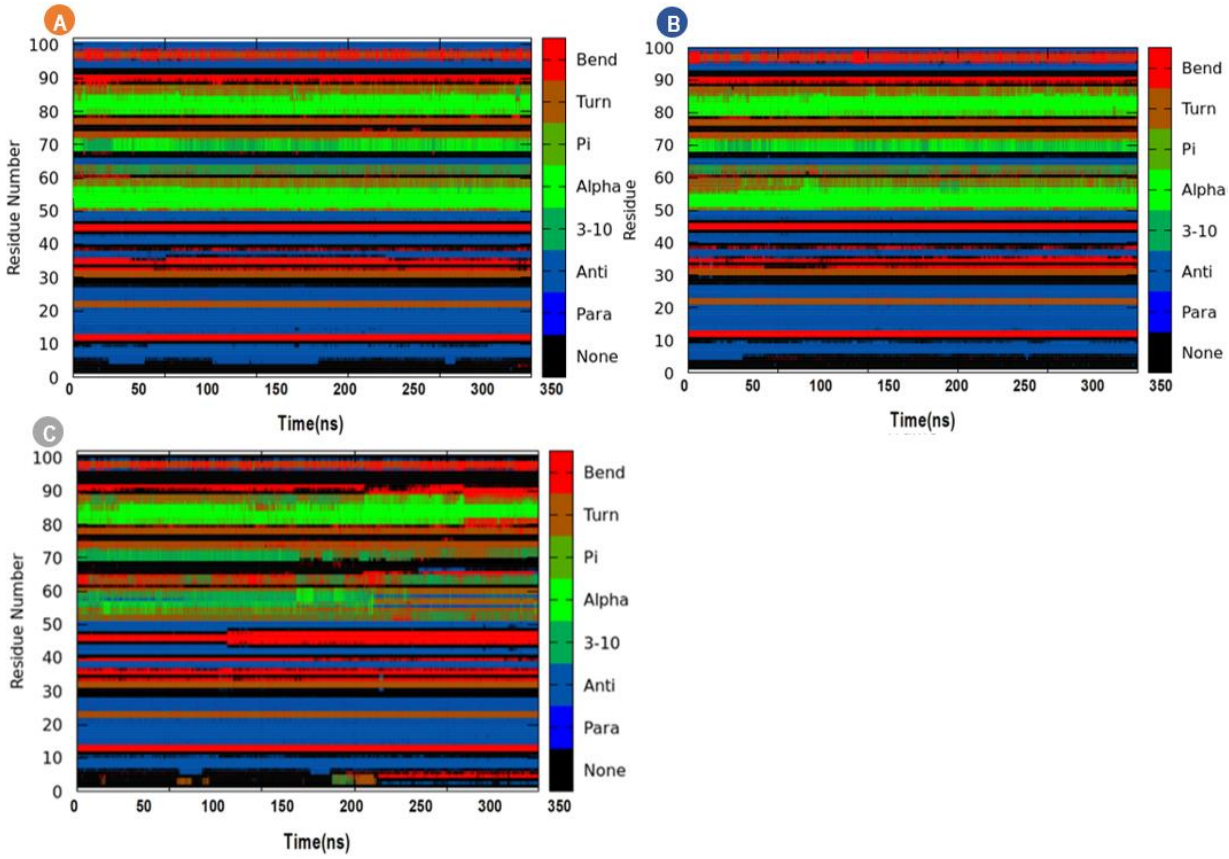


Figure 6.8 DSSP classification for the time evolution of the secondary structural elements for Apo **[A]**, Non-NBD **[B]**, and Cov-NBD **[C].**

6.3.3 Conformational clustering and principal component analysis

Principal component analysis are largly used to understand the conformational changes due to eithier the mutation or drug binding ,and this help for producing potentional inhibitors [60–62]. In this study, Principal component analysis and clustering have been carried out for Cov-NBD HSP72 and Non-NBD HSP 72 complexes to realize the overall concerted motion of HSP 72-NBD protein. The structural distribution of conformational changes and the proportion of variance of the captured eigenvectors are shown in (Figure 6.8B & C). The first three principal components accounted for 53.2% and 47.4 % of the total variance observed in the MD trajectories for Cov-HSP72 and Non -HSP72 complexes. The clustering analysis also shows conformational distribution variance along with the first, second, and third principal components with each dot representing a single complex conformation. For the Cov-HSP72 complex, the PC 1/3 and PC 1/2 conformational subspace highlight distinct thermodynamic periodic jumps with a significant energy barrier. However, the Non-NBD complex displayed a uniform and overlapping conformational PC 1/3 and PC 1/2 subspace with a minimal energy barrier. The clustering and principal component analyses suggest the HSP72 undergoes a massive conformational change of the nucleotide-binding domain with reversible ligand compared to the irreversible ligand. The sealed conformation conjunction of the irreversible inhibitor shows intrinsically more dynamic residue mobility, mainly in the IIA α-helix region compared to the reversible inhibitor complex (Figure 6.9A & D).

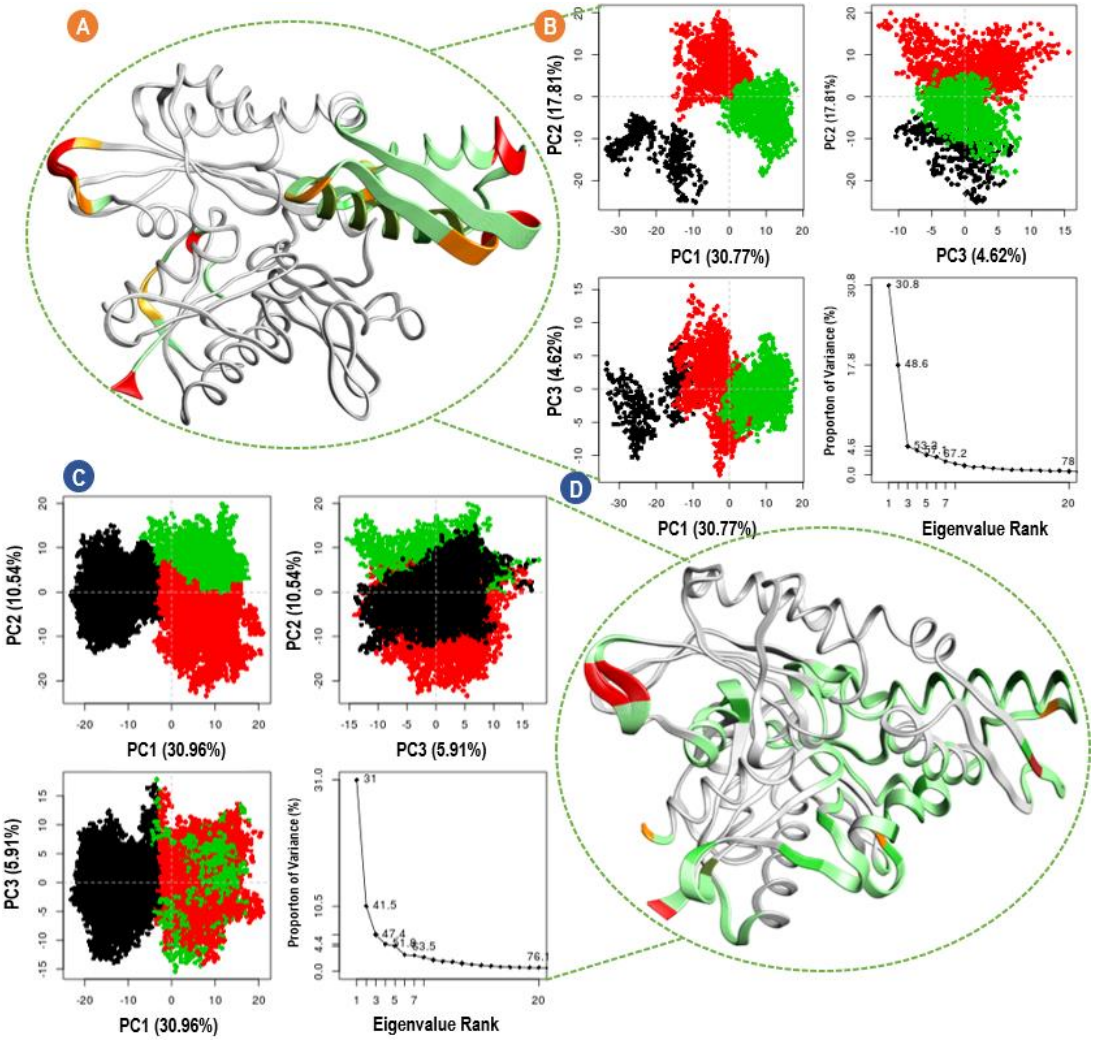


Figure 6.9 Clustering and principal component analysis on 350ns of equidistance conformations using the Bio3D package in R. The plots show the first three eigenvectors for Cov-NBD HSP72 complex **[B]** and Non-NBD HSP72 complex **[C]**, Conformers are colored according to the k-means clustering: cluster 1, black; 2, red; 3, green. Dominant motions and captured eigenvector variance; and residue mobility for Cov-NBD complex **[A]** and Non-NBD complex **[D].** The color scale from grey, green, to red, depicts low to high atomic displacements.

6.3.4 Understand the structural dynamics of the HSP72-NBD α-helices upon covalent and non-covalent inhibition.

To further gain an additional structural understanding of the preferential covalent inhibition conjunction over the non-covalent inhibition, we essentially compared the subdomain α-helix dynamic motions throughout the simulation. The observed structural changes in the α-helix motion show a different mechanism of the NBD characterized by a wide conformation and a sealed confirmation of the IIB and IIA subdomains for Non-NBD complex and Cov-NBD complex, respectively (Figure 6.10 and Figure 6.11). The Cov-NBD complex binding results

in the closure of the two helices for the subdomain IIB and IIA, thus, inducing a sealed confirmation as displaying in (Figure 6.10). On the other hand, a wide conformation of the two α-helices for the subdomain IIB and IIA was observed for the Non-NBD complex (Figure 6.11).

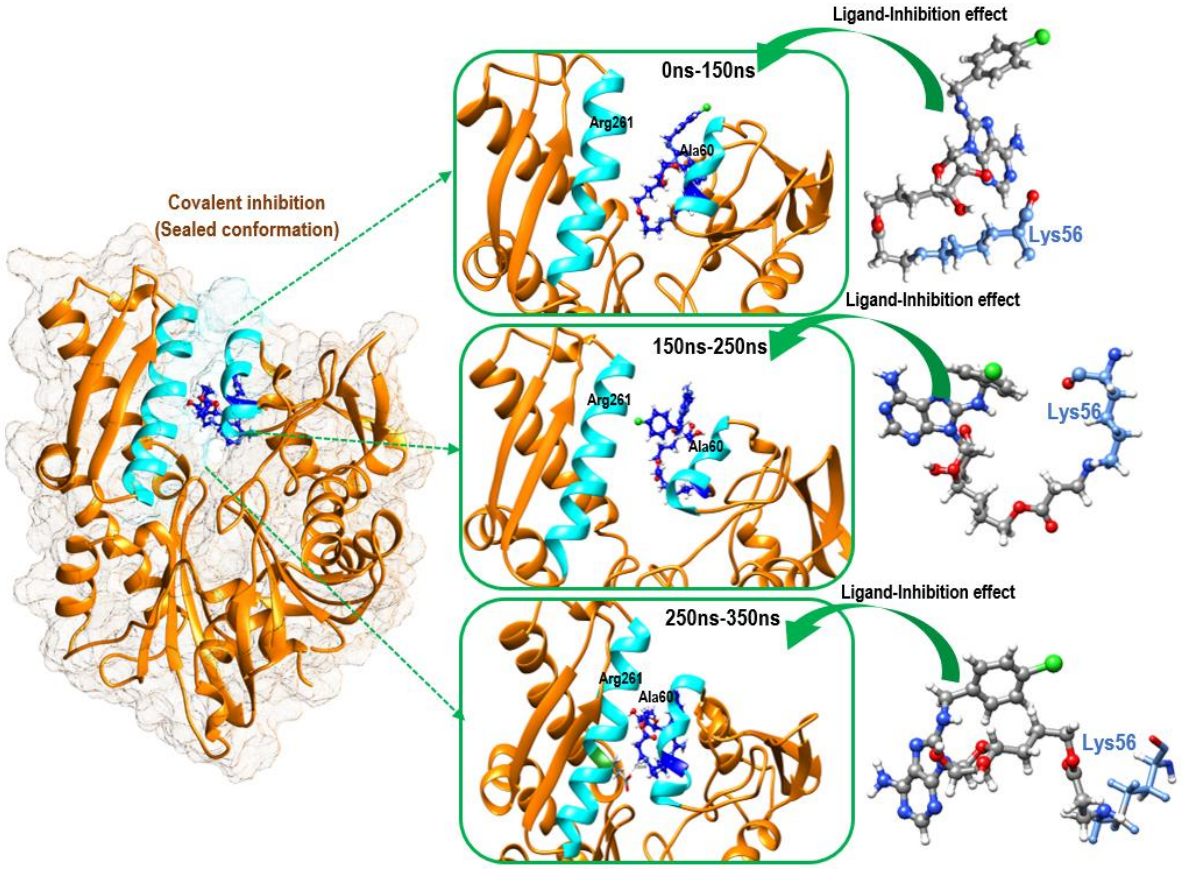


Figure 6.10 The conformational dynamic of α-helix conjunction upon 8-N-benzyladenosine binding to Lys56 via covalent bond showing the sealed confirmation of Cov-NBD of HSP72.

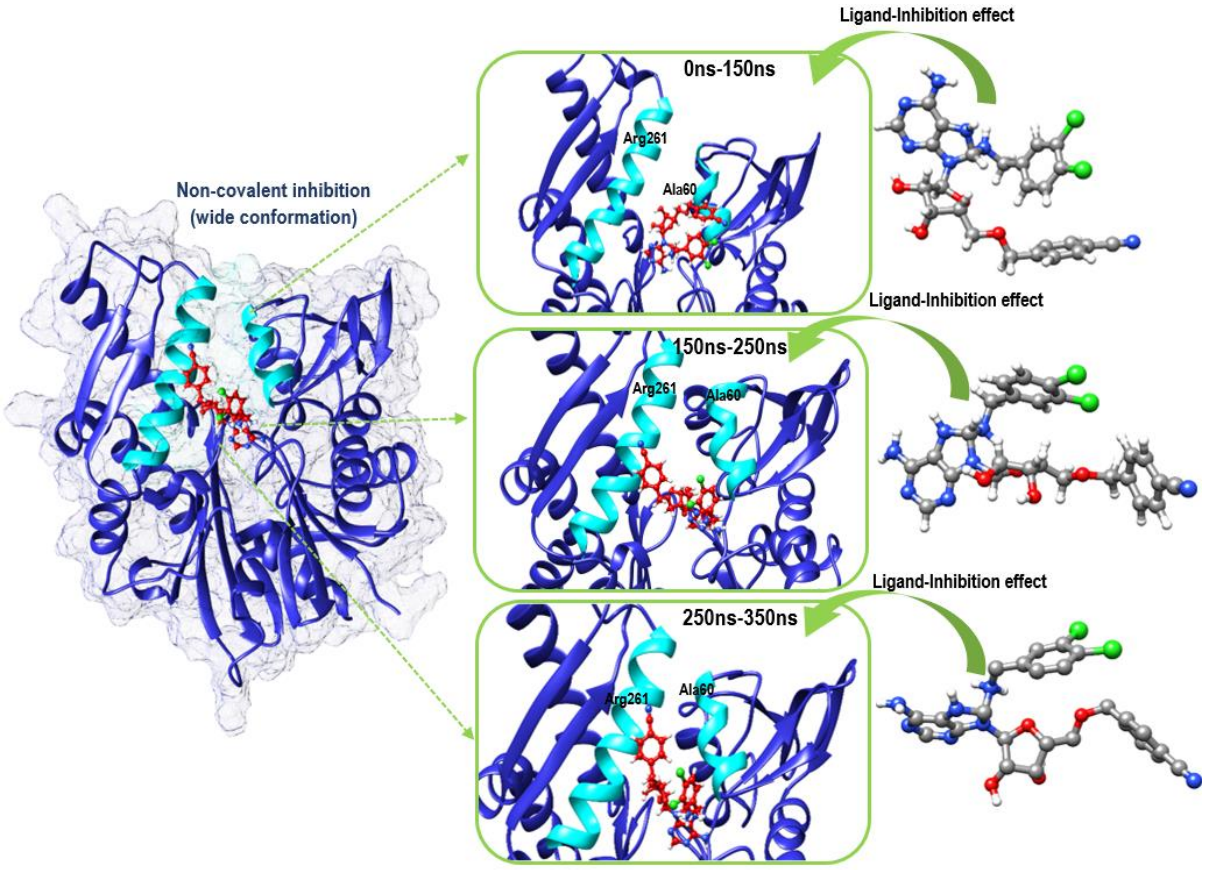


Figure 6.11 The conformational dynamic of α-helix conjunction upon non-covalent binding at ATP binding site via non-covalent bond showing the wide conformation of Non-NBD of HSP72.

The inhibition of HSP72-NBD as a result of the conjunction of irreversible inhibitor was captured by observing the decline in the inter-residue distance between two opposite IIB and IIA subdomain helix residues Arg-261 and Ala-60 with an average distance of 6.25 Å. However, a distance of 14.63 Å was found between residues Arg-261 and Ala-60 in the conjunction of reversible inhibitor (Figure 6.12). The irreversible inhibitor prompted a sealed confirmation of the IIB and IIA nucleotide-binding subdomain.

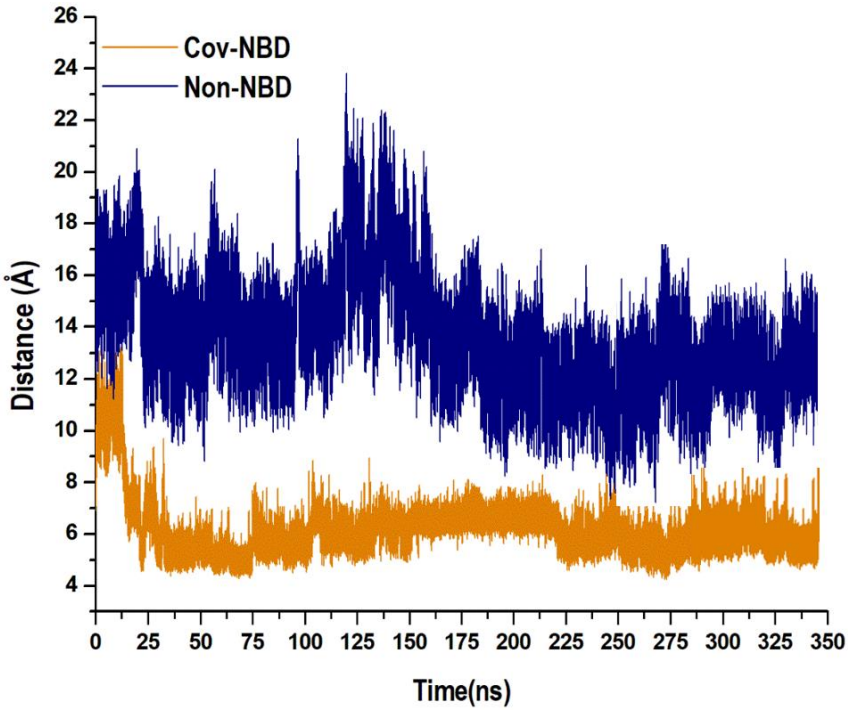


Figure 6.12 Distance between C- α residues involving ALA60 and ARG261 α-helices of the NBD. The average distances were found to be 6.25 Å and 14.63 Å, respectively, for Cov-NBD HSP72 and Non-NBD HSP72 conformations.

Further evaluation of the mechanism of the interaction as a reversible and irreversible inhibitor reveals a salt-bridge interaction between Lys-56 and Glu-268 in the irreversible HSP72-NBD complex (Figure 6.13B), which is absent in the reversible HSP72-NBD complex (Figure 6.13A). The detected covalent bond of the irreversible complex and the sealed confirmation is nearly identical to the sealed confirmation of ADP/Pi bound to HSP72-NDB [22]. Non-covalent inhibition conjunction, the distance wide open between Glu-268 and Lys-56 of the two α-helices, prevents the possible formation of the Glu-268 and Lys-56 salt bridge and generate a wide conformation (Figure 6.13A). Furthermore, other residues interact for a close distance, which helps in inducing the sealed conjunction of the α-helix subdomain as showing (Figure 6.13A & B).

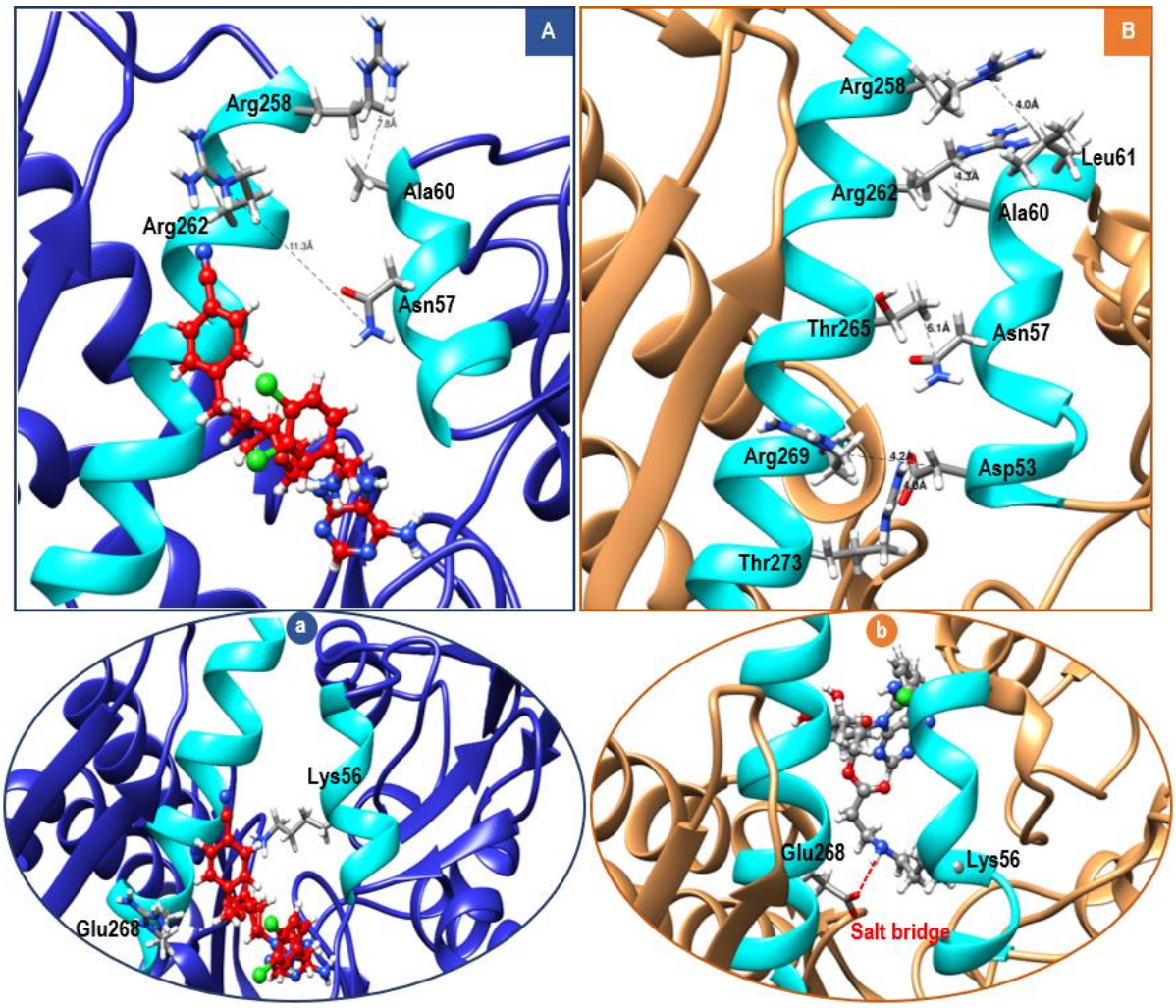


Figure 6.13 Non-NBD (**A**) and Cov-NBD(**B**) conjunction HSP72-NBD covalent and non-covalent interaction mechanisms respect the bond distance between the conjunction residues. Crucial salt-bridge between Glu-268 and Lys-56 residues were missing in the Non-NBD wide conformation.

6.4 Conclusion

The human heat shock protein 72 (HSP72) represents a vital therapeutic target during the critical stages of oncogenesis and progression of human cancers. This study was undertaken to compare two inhibition modes of HSP72-NBD: covalent and non-covalent in order to gain a better understanding of the underlying structural dynamics of each binding mechanism. The performed MD analyses provided in this study established the most optimal ligand binding mechanism at the HSP72-NBD enzyme. Molecular mechanism revealed that the covalent ligand was more able to induce and stabilize the closed conformation of the HSP72-NBD throughout the simulation than the non-covalent ligand. This was further confirmed by extensive analyses of the structural basis and conformational dynamics associated with the preferential conjunction of the covalent ligand over the non-covalent ligand in HSP72-NBD models. Covalent bonds induced the α-helices sealed conformation of the nucleotide-binding subdomain; based on our results, this will prevent other small molecules from interacting at the ATP binding site domain. Furthermore, inhibition of the ATP binding domain can directly affect the ATPs protein folding mechanism of the HSP72 enzyme. Interestingly, the sealed confirmation maintained the crucial salt-bridge between glutamic acid 268 and lysine 56 residues, strengthening the interaction affinity of covalent ligand bound to the HSP72-NBD. Disturbance in HSP72-NBD sealed confirmation upon covalent inhibition was also revealed by the DCCM, which showed a highly correlative residual movement in the α-helices of the NBD compared with the non-covalent inhibition. This study will assist with the design of more potent and highly selective covalent inhibitors for HSP72 with the potential to overcome drug resistance challenges and represent a novel therapeutic approach for inhibiting HSP72 oncoprotein.

Acknowledgements

The authors acknowledge the College of Health Science of the University of KwaZulu-Natal for funding and the Centre for High-Performance Computing (CHPC), Cape Town, South Africa, for computational resources (www.chpc.ac.za).

Conflict of interest

The authors declare no conflict of interest.

References

1. Bauer RA. Covalent inhibitors in drug discovery: From accidental discoveries to avoided liabilities and designed therapies. Drug Discov. Today20(9), 1061–1073 (2015).

2. Johnson DS, Weerapana E, Cravatt BF. Strategies for discovering and derisking covalent, irreversible enzyme inhibitors. *Future Med. Chem.* 2(6), 949–964 (2010).

3. Rahman A, Hoque MM, Khan MAK, Sarwar MG, Halim MA. Non-covalent interactions involving halogenated derivatives of capecitabine and thymidylate synthase: a computational approach. *Springerplus*. 5(1), 1–18 (2016).

4. Zhou P, Huang J, Tian F. Specific Noncovalent Interactions at Protein-Ligand Interface: Implications for Rational Drug Design. *Curr. Med. Chem.* 19(2), 226–238 (2012).

5. Ke X, Ng VWL, Ono RJ, *et al.* Role of non-covalent and covalent interactions in cargo loading capacity and stability of polymeric micelles. *J. Control. Release*. 193, 9–26 (2014).

6. Varma AK, Patil R, Das S, Stanley A, Yadav L, Sudhakar A. Optimized hydrophobic interactions and hydrogen bonding at the target-ligand interface leads the pathways of Drug-Designing. *PLoS One*. 5(8) (2010).

7. Wei Y, Ma L, Zhang L, Xu X. Noncovalent interaction-assisted drug delivery system with highly efficient uptake and release of paclitaxel for anticancer therapy. *Int. J. Nanomedicine*. 12, 7039–7051 (2017).

8. Gomila RM, Frontera A. Covalent and Non-covalent Noble Gas Bonding Interactions in XeFn Derivatives (n = 2–6): A Combined Theoretical and ICSD Analysis. *Front. Chem.* [Internet]. 8 (2020). Available from: https://www.frontiersin.org/article/10.3389/fchem.2020.00395/full.

9. Kumar SJ, Stokes J, Singh UP, *et al.* Targeting Hsp70: A possible therapy for cancer. *Cancer Lett.* [Internet]. 374(1), 156–166 (2016). Available from: http://dx.doi.org/10.1016/j.canlet.2016.01.056.

10. Powers M V., Jones K, Barillari C, Westwood I, Montfort RLM van, Workman P. Targeting HSP70: The second potentially druggable heat shock protein and molecular chaperone? *Cell Cycle*. 9(8), 1542–1550 (2010).

11. Mayer MP, Bukau B. Hsp70 chaperones: Cellular functions and molecular mechanism. *Cell. Mol. Life Sci.* 62(6), 670–684 (2005).

12. Mayer MP, Schröder H, Rüdiger S, Paal K, Laufen T, Bukau B. Multistep mechanism of substrate binding determines chaperone activity of Hsp70. *Nat. Struct. Biol.* (2000).

13. Schmid D, Baici A, Gehring H, Christen P. Kinetics of molecular chaperone action. *Science (80-.).* [Internet]. 263(5149), 971–973 (1994). Available from: https://www.sciencemag.org/lookup/doi/10.1126/science.8310296.

14. Pierpaoli E V., Gisler SM, Christen P. Sequence-Specific Rates of Interaction of Target Peptides with the Molecular Chaperones DnaK and DnaJ †. *Biochemistry* [Internet]. 37(47), 16741–16748 (1998). Available from: https://pubs.acs.org/doi/10.1021/bi981762y.

15. Zhuravleva A, Clerico EM, Gierasch LM. An interdomain energetic tug-of-war creates the allosterically active state in Hsp70 molecular chaperones. *Cell*. 151(6), 1296–307 (2012).

16. Penkler D, Sensoy Ö, Atilgan C, Tastan Bishop Ö. Perturbation–Response Scanning Reveals Key Residues for Allosteric Control in Hsp70. *J. Chem. Inf. Model.* 57(6), 1359–1374 (2017).

17. Sharma D, Masison DC. Hsp70 structure, function, regulation and influence on yeast prions. *Protein Pept. Lett.* 16(6), 571–81 (2009).

18. Chiappori F, Merelli I, Colombo G, Milanesi L, Morra G. Molecular Mechanism of Allosteric Communication in Hsp70 Revealed by Molecular Dynamics Simulations. *PLoS Comput. Biol.* 8(12), e1002844 (2012).

19. Nawrocki ST, Carew JS, Dunner K, *et al.* Bortezomib inhibits PKR-like endoplasmic reticulum (ER) kinase and induces apoptosis via ER stress in human pancreatic cancer cells. *Cancer Res.* 65(24), 11510–9 (2005).

20. Pettinger J, Le Bihan YV, Widya M, van Montfort RLM, Jones K, Cheeseman MD. An Irreversible Inhibitor of HSP72 that Unexpectedly Targets Lysine-56. *Angew. Chemie - Int. Ed.* 56(13), 3536–3540 (2017).

21. Williamson DS, Borgognoni J, Clay A, *et al.* Novel adenosine-derived inhibitors of 70 kDa heat shock protein, discovered through structure-based design. *J. Med. Chem.* 52(6), 1510–1513 (2009).

22. Cheeseman MD, Westwood IM, Barbeau O, *et al.* Exploiting Protein Conformational Change to Optimize Adenosine-Derived Inhibitors of HSP70. *J. Med. Chem.* 59(10), 4625–4636 (2016).

23. Berman HM, Battistuz T, Bhat TN, *et al.* The Protein Data Bank. *Biol. Crystallogr.* 58, 899–907 (2002).

24. Schlecht R, Scholz SR, Dahmen H, *et al.* Functional analysis of Hsp70 inhibitors. *PLoS One*. 8(11), 1–12 (2013).

25. Pettersen Ef Fau - Goddard TD, Goddard Td Fau - Huang CC, Huang Cc Fau - Couch GS, *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* (2004).

26. Windows MM V, X MOS. Molegro Molecular Viewer User Manual. , 145 (2011).

27. Kusumaningrum S, Budianto E, Kosela S, Sumaryono W, Juniarti F. The molecular docking of 1,4-naphthoquinone derivatives as inhibitors of Polo-like kinase 1 using Molegro Virtual Docker. *J. Appl. Pharm. Sci.* 4(11), 47–53 (2014).

28. Trott O, Olson A. Autodock vina: improving the speed and accuracy of docking. *J. Comput. Chem.* 31(2), 455–461 (2010).

29. Madhavi Sastry G, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided. Mol. Des.* 27(3), 221–234 (2013).

30. Friesner RA, Murphy RB, Repasky MP, *et al.* Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes.

J. Med. Chem. [Internet]. 49(21), 6177–6196 (2006). Available from: https://pubmed.ncbi.nlm.nih.gov/17034125/.

31. Mhlongo NN, Ebrahim M, Skelton AA, Kruger HG, Williams IH, Soliman MES. Dynamics of the thumb-finger regions in a GH11 xylanase Bacillus circulans: comparison between the Michaelis and covalent intermediate. *RSC Adv.* 5(100), 82381–82394 (2015).

32. Ramharack P, Oguntade S, Soliman MES. Delving into Zika virus structural dynamics-a closer look at NS3 helicase loop flexibility and its role in drug discovery. *RSC Adv.* 7(36), 22133–22144 (2017).

33. Götz AW, Williamson MJ, Xu D, Poole D, Le Grand S, Walker RC. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* 8(5), 1542–1555 (2012).

34. Lindorff-Larsen K, Piana S, Palmo K, *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins*. 78(8), 1950–8 (2010).

35. Betz R. Dabble Zendo [Internet]. (2017). Available from: https://doi.org/10.5281/ZENOD%0AO.836914.

36. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79(2), 926–935 (1983).

37. Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, Walker RC. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh ewald. *J. Chem. Theory Comput.* 9(9), 3878–3888 (2013).

38. Khan S, Bjij I, Betz RM, Soliman MES. Reversible versus irreversible inhibition modes of ERK2: A comparative analysis for ERK2 protein kinase in cancer therapy. *Future Med. Chem.* 10(9), 1003–1015 (2018).

39. Perez A, MacCallum JL, Brini E, Simmerling C, Dill KA. Grid-Based Backbone Correction to the ff12SB Protein Force Field for Implicit-Solvent Simulations. *J. Chem. Theory Comput.* 11(10), 4770–4779 (2015).

40. Van Der Spoel D, Van Maaren PJ. The origin of layer structure artifacts in simulations of liquid water. *J. Chem. Theory Comput.* 2(1), 1–11 (2006).

41. Harvey MJ, De Fabritiis G. An implementation of the smooth particle mesh Ewald method on GPU hardware. *J. Chem. Theory Comput.* 5(9), 2371–2377 (2009).

42. Gonnet P. P-SHAKE: A quadratically convergent SHAKE in O (n2). *J. Comput. Phys.* 220(2), 740–750 (2007).

43. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* 9(7), 3084–3095 (2013).

44. Seifert E. OriginPro 9.1: Scientific data analysis and graphing software - Software review. *J. Chem. Inf. Model.* 54(5), 1552 (2014).

45. Kasahara K, Fukuda I, Nakamura H. A novel approach of dynamic cross correlation analysis on molecular dynamics simulations and its application to Ets1 dimer-DNA complex. *PLoS One*. 9(11) (2014).

46. Ichiye T, Karplus M. Collective motions in proteins: A covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins Struct. Funct. Bioinforma.* [Internet]. 11(3), 205–217 (1991). Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/prot.340110305.

47. Levy RM, Srinivasan AR, Olson WK, McCammon JA. Quasi-harmonic method for studying very low frequency modes in proteins. *Biopolymers* [Internet]. 23(6), 1099–1112 (1984). Available from: https://pubmed.ncbi.nlm.nih.gov/6733249/.

48. Yan F, Liu X, Zhang S, Su J, Zhang Q, Chen J. Effect of double mutations T790M/L858R on conformation and drug-resistant mechanism of epidermal growth factor receptor explored by molecular dynamics simulations. *RSC Adv.* [Internet]. 8(70), 39797–39810 (2018). Available from: https://pubs.rsc.org/en/content/articlehtml/2018/ra/c8ra06844e.

49. Laberge M, Yonetani T. Molecular dynamics simulations of hemoglobin a in different states and bound to DPG: Effector-linked perturbation of tertiary conformations and HbA concerted dynamics. *Biophys. J.* [Internet]. 94(7), 2737–2751 (2008). Available from: /pmc/articles/PMC2267116/?report=abstract.

50. London KP-T, Edinburgh undefined, Philosophical and D, 1901 undefined. LIII. On lines and planes of closest fit to systems of points in space. *Taylor Fr.* [Internet]. . Available from: https://www.tandfonline.com/doi/pdf/10.1080/14786440109462720.

51. Hotelling H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* [Internet]. 24(6), 417–441 (1933). Available from: /record/1934-00645-001.

52. Multivariate Statistical Methods: A Primer, Fourth Edition - Bryan F.J. Manly, Jorge A. Navarro Alberto - Google Books [Internet]. . Available from: https://books.google.co.za/books?hl=en&lr=&id=bCMNDgAAQBAJ&oi=fnd&pg=PP1&dq=Manly+B.+Multivariate+statistics—a+primer.+Chapman+%26+Hall/CRC%3B+Boca+Raton,+FL:+1986&ots=FGfjmay1aR&sig=s6i0YROUrYcbyENcg-KQPN45_HE&redir_esc=y#v=onepage&q&f=false.

53. Abdi H, Williams LJ. Principal component analysis. Wiley Interdiscip. Rev. Comput. Stat.2(4), 433–459 (2010).

54. Principal Component Analysis [Internet]. Springer-Verlag, New York. Available from: http://link.springer.com/10.1007/b98835.

55. Pan L, Patterson JC. Molecular Dynamics Study of Zn(Aβ) and Zn(Aβ)2. *PLoS One*. 8(9), 70681–70688 (2013).

56. Richmond TJ. Solvent accessible surface area and excluded volume in proteins. Analytical equations for overlapping spheres and implications for the hydrophobic effect. *J. Mol. Biol.* 178(1), 63–89 (1984).

57. Wisniewska M, Karlberg T, Lehtiö L, *et al.* Crystal Structures of the ATPase Domains of Four Human Hsp70 Isoforms: HSPA1L/Hsp70-hom, HSPA2/Hsp70-2, HSPA6/Hsp70B', and HSPA5/BiP/GRP78. *PLoS One*. 5(1), e8625 (2010).

58. Sriram M, Osipiuk J, Freeman B, Morimoto R, Joachimiak A. Human Hsp70 molecular chaperone binds two calcium ions within the ATPase domain. *Structure*. 5(3), 403–14 (1997).

59. Shida M, Arakawa A, Ishii R, *et al.* Direct inter-subdomain interactions switch between the closed and open forms of the Hsp70 nucleotide-binding domain in the nucleotide-free state. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 66(3), 223–232 (2010).

60. Chen J, Wang X, Pang L, Zhang JZH, Zhu T. Effect of mutations on binding of ligands to guanine riboswitch probed by free energy perturbation and molecular dynamics simulations. *Nucleic Acids Res.* 47(13), 6618–6631 (2019).

61. Li Y, Deng L, Ai S-M, *et al.* Insights into the molecular mechanism underlying CD4-dependency and neutralization sensitivity of HIV-1: a comparative molecular dynamics study on gp120s from isolates with different phenotypes. *RSC Adv.* [Internet]. 8(26), 14355–14368 (2018). Available from: http://xlink.rsc.org/?DOI=C8RA00425K.

62. Ahrari S, Khosravi F, Osouli A, *et al.* MARK4 protein can explore the active-like conformations in its non-phosphorylated state. *Sci. Rep.* [Internet]. 9(1), 12967 (2019). Available from: http://www.nature.com/articles/s41598-019-49337-0.

CHAPTER 7

Comparison of irreversible inhibition targeting HSP72 protein - the resurgence of covalent drug developments

Aimen Aljoundi[1], Ahmed El Rashedy[1], and Mahmoud E.S Soliman[1*]

[1]Molecular Bio-computation and Drug Design Laboratory

School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 78

Abstract

The covalent inhibition mechanism of action, which overcomes competition with high-affinity, high-abundance substrates of challenging protein targets, can deliver effective chemical probes and drugs. Heat shock protein 72 (HSP72) expressed in cancer cells may be responsible for tumorigenesis and tumor progression by providing resistance to chemotherapy. In this study, we explore the most optimal binding mechanism in the inhibition of the HSP72 by two different covalent generation, TCI2 and TCI8. However, the structural basis and conformational changes associated with this preferential covalent inhibition to HSP72-TCI2 over HSP72-TCI8 remain unclear. Our results revealed that HSP72-TCI2 covalent complex was more able to stabilize and induce better interaction with high correlated dynamic motion of HSP72-NBD at ATP binding site throughout the simulation in comparison to the HSP72-TCI8 covalent complex. This is supported by the dynamic cross-correlation and principal component analysis that was more dominant in the HSP72-TCI2 inhibited complex. The insights demonstrating the above binding mechanism of HSP72 establish TCI2 covalent inhibition as the preferred method of inhibiting the HSP72 protein. This investigation aids in the understanding of the structural mechanism of HSP72 inhibition and would assist in the design of more potent covalent inhibitors of HSP72.

Keywords: MD simulations; HSP72; Covalent bond; Principal component analysis; ATP binding domain.

7.1 Introduction

Irreversible covalent drug inhibition is a developing model; however, critical gaps in unravelling molecular determinants' efficacy still persist. The development of covalent drugs, particularly in cancer therapeutics, has recently lightened interest among the pharmaceutical research community. Despite many of our most important drugs using an enzyme's irreversible covalent inhibition [1], concerns about idiosyncratic toxicity led to the near exclusion of this mechanism of action (MOA) from drug discovery programs [2]. The recent renaissance in covalent inhibitors is primarily due to their inherent advantage over reversible counterparts with high-affinity, high-abundance natural substrates for antagonizing proteins [3]. While the previous generation of irreversible inhibitor drugs were discovered by serendipity or were natural products, the rational design strategy for modern targeted covalent inhibitors (TCIs) focuses on exploiting high resolution small-molecule/protein X-ray crystal structures of high-affinity reversible ligands to target active site, solvent exposed cysteine residues with sparingly reactive electrophiles [4]. Unfortunately, the rarity of cysteine in the proteome has limited its application [5], leading to an increased interest in targeting other potentially nucleophilic residues [6–8].

TCIs use both reversible and irreversible occupancy of a protein depending on exposure duration and concentration [9]. For rational design, identifying a TCI only through an IC_{50} value may be restricting, as with increasing preincubation time, the inhibitor (I) will eventually become more potent. TCIs typically react via a two-step (MOA), the target protein enzyme (E) interacts reversibly with TCI (L) to make a non-covalent enzyme-ligand complex (EL). The tendency of the reversible complex recognizes by the free concentration of the drugs and the equilibrium constant K_i. The reversible complex can then bear covalent bond formation as identified by first-order kinetic, to give the covalent complex (E-L) and inducing conformational changes in the complex. Covalent inhibitors can be divided into two

chemotypes: a warhead acting as a reactive group function to which the entire structure is attached (Figure 7.1).

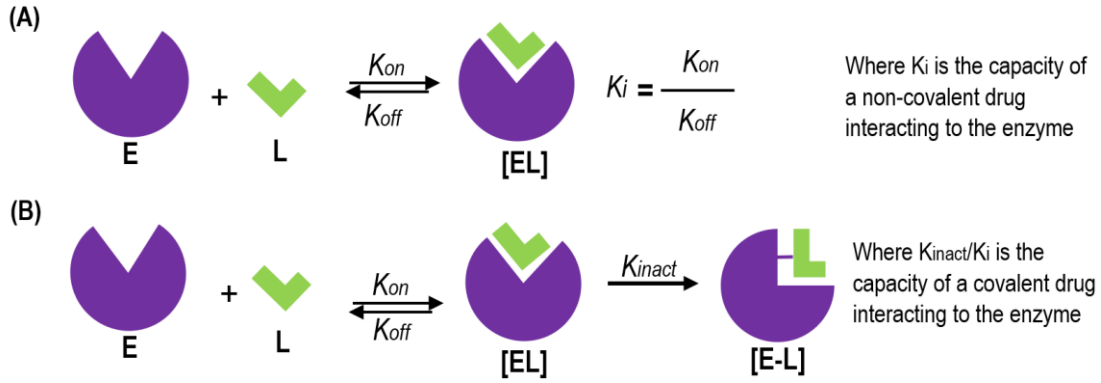


Figure 7.1 Drug-enzyme interaction. Association of a L interacting with E in **[A]** conventional non-covalent and **[B]** covalent mode of action. Where is L: Ligand; E: Enzyme.

Heat shock 72 kDa protein (HSP72) is a member of the HSP70 family of molecular chaperones. It is an ATPase that binds misfolded proteins, stabilizing the cellular environment and allowing the cell to return to homeostasis [10]. HSP72 is induced in an HSF1-dependent manner when the cell is undergoing stress and is overexpressed in several cancer cell types [11]. This overexpression is correlated with metastasis, poor prognosis, and resistance to chemotherapy in patients. Thus, inhibition of HSP72 is considered to be a successful pathway for anti-tumour therapy [12].

The different functions of Hsp70s are achieving through a transient chaperone interaction with substrate proteins through its C-terminal substrate-binding domain (SBD) [13]. The nucleotide binds allosterically to the N-terminal nucleotide-binding domain (NBD) to control the transient chaperone interaction. Substrate-binding domain affinity for substrates decreases by 10 to 400 folds when ATP is binding to the N-terminal of the nucleotide-binding domain. Hence, the nucleotide-binding domain inhibition is considered one of the most promising approaches for HSP72 function inhibition [14]. The NBD consists of two adjacent lobes (lobe I and lobe II),

which form a deep nucleotide groove connected to the base [15]. Each lobe composes of two subdomains (IA, IIA, IB, and IIB) [16,17], domains IB and IIB are linked to IA and IIA, respectively, by flexible hinges and control access to the nucleotide-binding sites.

Targeting lysine residues with TCIs presents a number of unique challenges and is still in its infancy [18], but the greater prevalence of lysine in the proteome could result in more wide-ranging applications of the irreversible inhibitor paradigm than has so far been possible through the rational targeting of cysteine [19]. Recently , Pettinger et al. reported the discovery that lysine-56 can be selectivity targeted with an acrylate-derived **TCI 8** ,and **TCI 2** (Figure 7.2) [20]. TCI2 was demonstrate a >100-fold enhancement in covalent efficiency compared to TCI8. Yet, to date, little is known about the peculiar binding mechanism of this inhibitors and the details behind the detected high affinity of TCI2 over TCI8 when targeting the covalent binding sites of HSP72 (Figure 7.3).

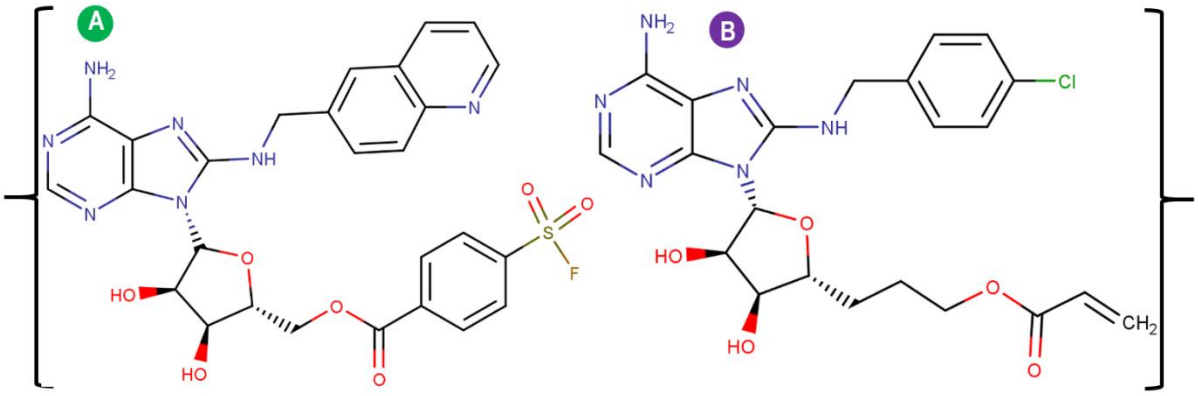


Figure 7.2: 2D structure of **[A]** TCI2 and **[B]** TCI8.

In this study, we aim to compare two covalent binding modes and stablish the most optimal binding mechanism amongst first generation and third generation covalent inhibition of HSP72. This was accomplished by molecular dynamic simulation, which compared the structural modification of covalent inhibition of HSP72 by comparing one mode of inhibition of HSP72, covalent (irreversible) with respect to its APO (apoenzyme correspond to the free

form of the enzyme) form in order to gain a better understanding of underlying structural dynamics and critical implications in each binding mechanism (Figure 3).

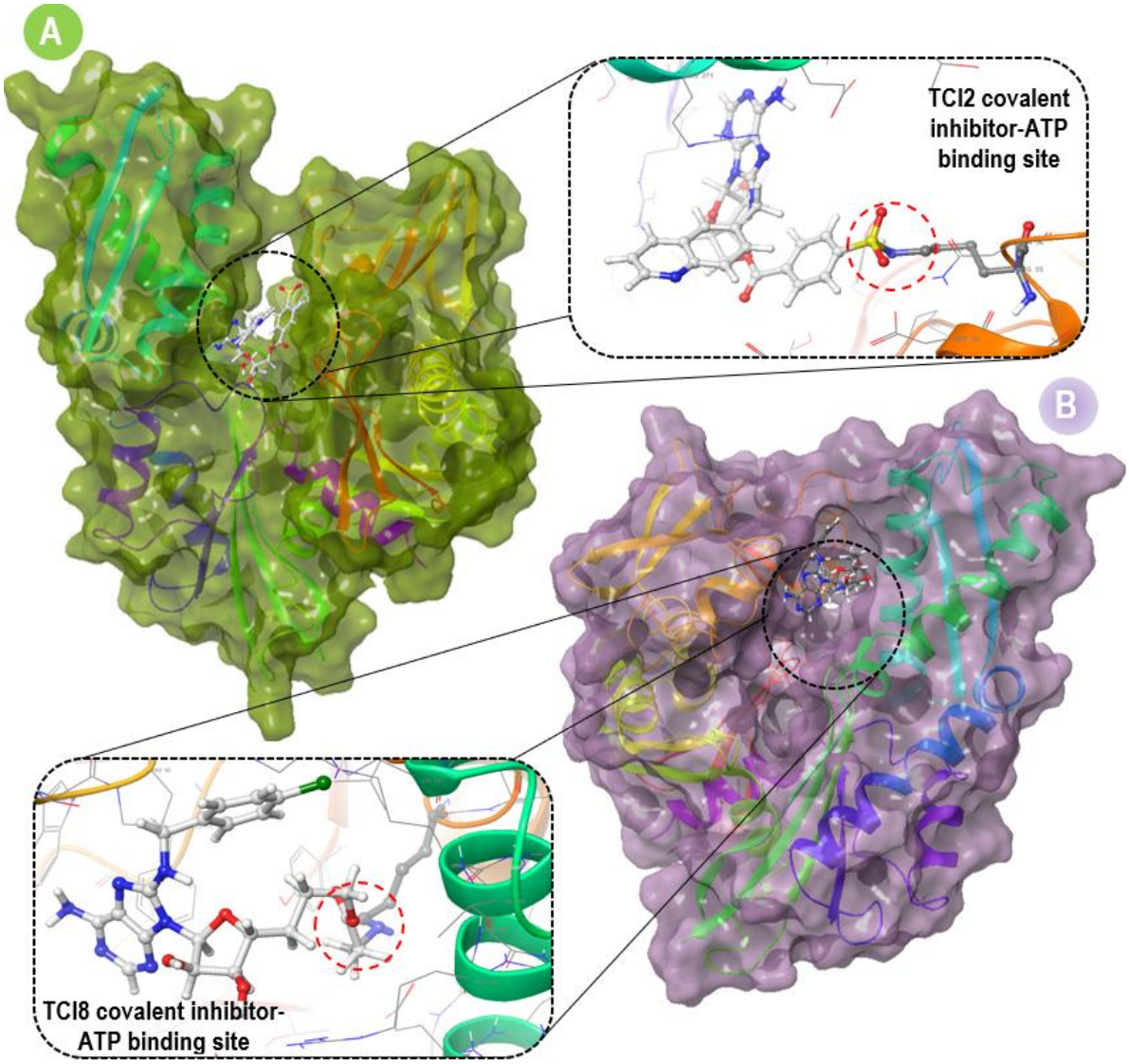


***Figure 7.3*:** The surface view of HSP72 **[A]** covalently bonded with TCI2 **[B]** and covalently bonded with TCI8. A close view of two covalent inhibition conditions (red color circle shows the covalent bond via Lys56).

7.2 Computational methodology

7.2.1 System preparation and MD simulation

The X-ray crystal structures of heat shock protein; HSP72 retrieved from the Protein Data Bank [21] (PDB code: 5MKS) [22]. These structures were then prepared for molecular dynamic (MD) simulation using the UCSF Chimera software package [23]. MarvinSketch 6.2.1, 2014,

and Molegro Molecular Viewer (MMV), was used for the ligands preparation and ensured that the ligands proper angels and hybridization state were displayed [24,25]. AutoDock Tools GUI was used to describe the grid box at lysine-56 and ATP binding site of the protein [26]. The Lamarckian Genetic algorithm was used to perform docking calculations [27]. The prepared systems protonation states were optimized using Maestro Schrödinger [28], as for the covalent systems, the covalent bond was built up, necessary hydrogen atoms were corrected, and capping neutral residues to ensure protein stability during the simulation. The Apo system was run based on our previous in-house protocol for non-covalent simulations [29,30].

7.2.2 Simulation protocols

7.2.2.1 Molecular dynamics simulation with Covalent inhibitors

The two covalent systems were exposed to an all-atom classical covalent molecular dynamic simulation (MD) using the PMEMD package in Amber 14 [38]. The Antechamber module in the Amber14 was used to provide atom types and atomic ligand partial charges using the FF14SB forcefield [39]. The LEaP program was used to generate a library defining the ligand residue topology. The final system was built, neutralized, and solvated with two Na+ counter-ions using the Dabble program [40]. Before starting the simulation process, the studied covalent complexes were placed within a box of TIP3P water molecules with 10 Å distance from the protein [41]. Particle mesh Ewald (PME) method was implemented within Amber14, with direct space and Van der Waals cut-off of 12 Å, to obtain long-range electrostatic interaction. To further relax the complex and remove potential steric clashes, each system was energy minimized for a total of 7500 with a 10 Å restraint conditions applied. The systems were heated for 30 ps from 0 to 300 K with an additional 7-ns equilibration performed at a 2-fs integration time step. MD simulation production runs of 250ns were performed for each system during which the SHAKE algorithm was used to constrict all atomic hydrogen bonds [42]. The computational methodology regarding the covalent systems was based on our previously

reported [43]. The CPPTRAJ and PTRAJ modules [44] of AMBER14 package were used to analyze resulting trajectories for root mean square deviation (RMSD), root mean square fluctuation (RMSF), solvent accessible surface area (SASA), the radius of gyration (ROG), and secondary structure analysis. The data were expressed in mean ± standard deviation. The obtained data were plotted using Microcal Origin tools [45] and Maestro Schrödinger software [35].

7.2.3 Clustering and Principal component analysis

Principal component analysis (PCA) was performed to describe the internal motion of the complexes using the Bio3D package in R. The process involves the initial construction of the covariance matrix (C) from (x,y,z) coordinate positions of the C-atoms as representatives of residues (N), generating a large matrix of dimension 3N_3N. The covariance matrix was diagonalized to obtain eigenvectors based on related eigenvalues and plotted on the first three eigenvectors (PC1, PC2, and PC3).

7.2.4 DCCM analysis

We used dynamic cross-correlation analysis to investigate the fluctuations and movements in the backbone of the α carbon atoms [31]. Cross-correlation components for i and j Cα atoms are shown in equation 1:

$$\mathbf{C_{ij}} = \frac{< \mathbf{\Delta ri}.\mathbf{\Delta rj} >}{\left(< \mathbf{\Delta r_i^2} >< \mathbf{\Delta r_j^2} >\right)^{\frac{1}{2}}} \qquad \textbf{(Equation 1)}$$

Where $r_i = C\alpha^i$ = standard time throughout the MD trajectory. Significantly correlated movements are symbolized by $C_{ij} = 1$, while $C_{ij} = -1$ symbolized highly anticorrelated movements in the trajectory. The divergence of motion from 1 and -1 indicates that i and j movements are anticorrelated. The DCCM matrix was carried out using the CPPTRAJ package

in Amber 14, and the matrices were plotted and evaluated using Origin software (www.originlab.com) [32].

7.3 Result and Discussion

7.3.1 Dynamic Conformational Stability and Fluctuations

MD simulations were carried out to investigate the inhibition performance and interactions of the potential ligands with HSP 72 targets. The validation of system stability is essential to trace disrupted motions and avoid artifacts that may arise during the course of the simulation. In this study, Root-Mean-Square Deviation (RMSD) was calculated to measure the systems' stability during the 250 ns simulations (Figure 7.4). The recorded average RMSD values for all frames of systems HSP72-Apo, HSP72-TCI8, and HSP72-TCI2 were 2.11 Å, 3.10 Å, and 2.31Å, respectively. These results revealed that HSP72-TCI2 complex acquired a relatively more stable conformation compared to HSP72-TCI8 system.

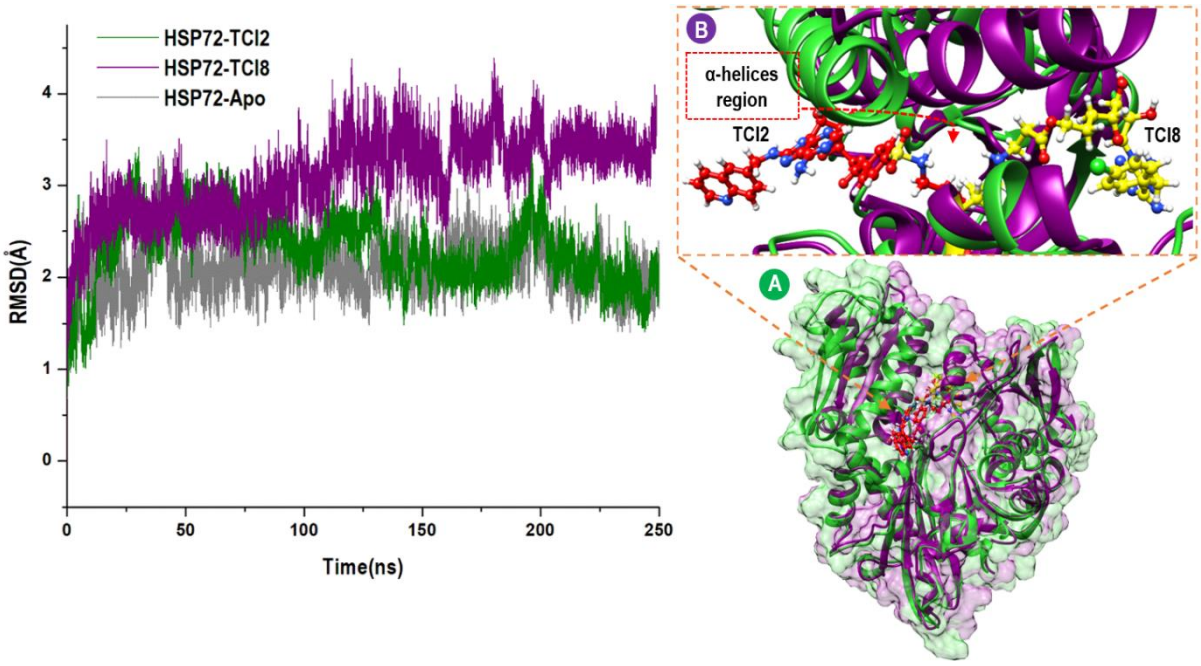


Figure 7.4: Comparative C-α RMSD plots showing the degree of stability and convergence of the studied systems over the 250ns MD simulation time. Crystal structure superimpose **[A]** and all ligand atoms alignment **[B],** were is the TCI2 (**red**) and TCI8 (**yellow**), with deferent binding pose direction at the ATP binding site α-helices region.

The evolution of protein structure flexibility upon ligand binding is essential in probing residue performance and their association with the ligand during MD simulation. HSP72 residual fluctuations were evaluated using Root-Mean-Square Fluctuation (RMSF) algorithm to assess the effect of inhibitor binding towards the respective targets over 250 ns simulations. The computed average atomic fluctuation of HSP72-Apo, HSP72-TCI8, and HSP72-TCI2 were 0.97Å, 1.61Å, and 1.60 Å, respectively. The plotted results in Figure 7.5 indicate overall relative lower residue fluctuation in the HSP72-TCI2 complex system compared with the HSP72-TCI8 complex system. This observation suggests that the covalent binding of TCI2 to HSP72 via Lys56 relatively decreases the overall protein flexibility in contrast to TCI8 covalent bond.

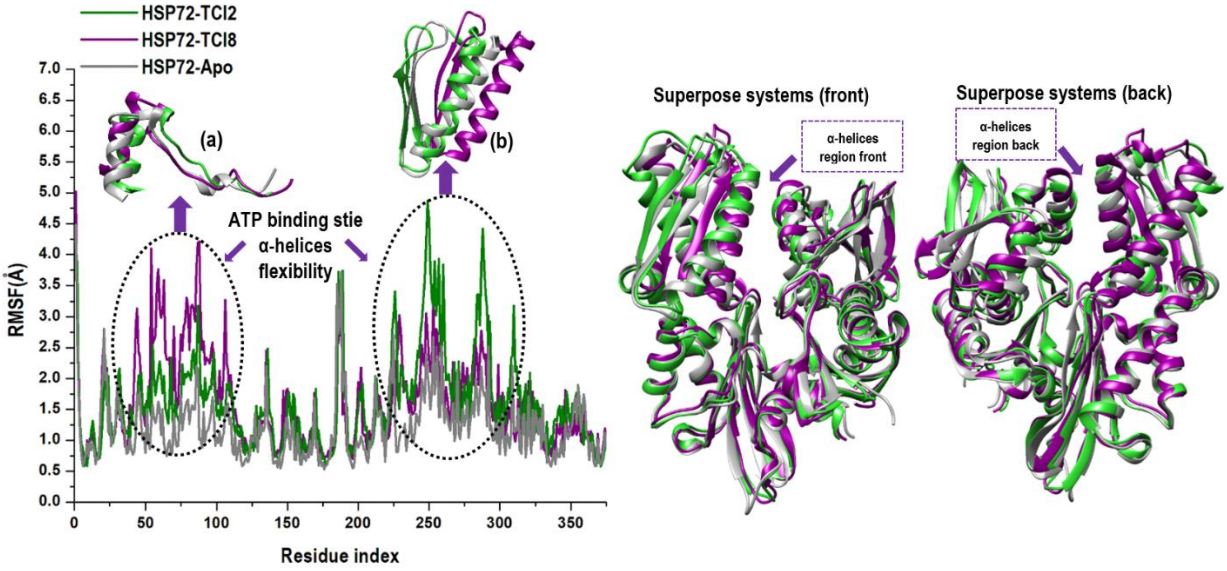


***Figure 7.5*:** The time evolution RMSF of each residue of the protein Cα atom over 250 ns for HSP72-Apo **(Gray)**, and HSP72-TCI2 **(green)**, and HSP72-TCI8 **(magenta)**. Superposed of the α-helix region crystal structures of the studied systems to show differences in fluctuations in the α helix region. Comparative Cα RMSF plot showing the degree of major flexibility of certain N terminal NBD. ATP binding site α-helices flexibility **[a]** and **[b]** (Inset).

The radius of gyration (RoG) is an indicator of protein compactness after ligand binding. The RoG was calculated by measuring the mass-weighted root mean square distance of the set of atoms from their shared center of mass [33]. The average Rg values are 21.52 Å, 21.70 Å, 21.63 Å, for HSP72-Apo, HSP72-TCI8, and HSP72-TCI2, respectively. This pattern suggested that

the HSP72-TCI2 interaction residues with HSP 72 showed relatively rigid, stable protein structure compared to the HSP72-TCI8 interaction residues (Figure S 7.1). The observed lower ROG value for HSP72-TCI2 than HSP72-TCI8 reflects a similar pattern, as shown in RMSF and RMSD values. Moreover, to indicate how the protein surface interrelates with solvent atoms and how it relays to the compactness of the hydrophobic protein core, the solvent-accessible surface area (SASA) of the protein upon ligand binding was calculated (Figure 7.6). This was accomplished by computing the surface area of the protein observable to solvent across the 250ns MD simulation, which is vital for biomolecular stability [34]. The computed average SASA values for HSP72-Apo, HSP72-TCI8, and HSP72-TCI2 inhibition systems were 17374.3 Å, 18006.26 Å, and 18086.26Å, respectively. The overall SASA indicates that the HSP72-TCI8 protein surface is relatively less exposed to solvent molecules compared to HSP72-TCI2 inhibition. Nevertheless, HSP72-TCI8 is 18.0% of the molecules exposed to the ligand, or just under 82% is buried in the receptor, while the HSP72-TCI2 is 18.6% exposed to the ligand. Or more importantly for these reasons, 18% is exposed to the solvent.

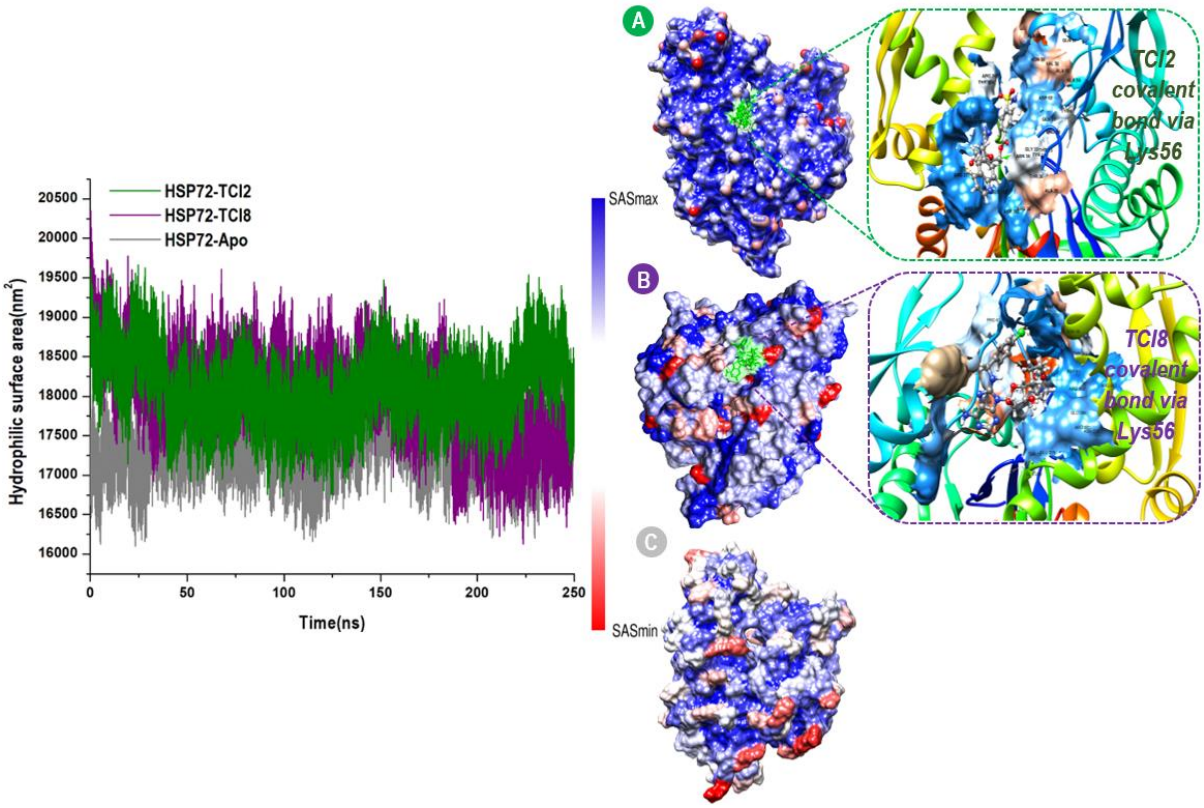


***Figure 7.6*:** Solvent accessible surface area **(**SASA**)** backbone atoms relative to the starting minimized structure over 250 ns for HSP72-Apo, HSP72-TCI2, and HSP72-TCI8. Areas with higher SASA values and lower SASA values are shown in blue and red, respectively, for HSP72-TCI2 **[A]**, HSP72-TCI8 **[B],** and HSP72-Apo **[C]**. A close view for both ligands SASA values pocket-site.

Different enzyme dynamics between the three conditions were examined using a dynamic cross-correlation matrix and plotted by observing the correlated movements of all residues. DCM plots are represented by different colors, highly positive regions range from green to dark-red (strongly correlated), and highly negative range from light blue to black (strong anticorrelated) movement for specific residues, respectively.

As evident in (Figure 7.7), The HSP72-TCI2 system enzyme displayed higher correlated motions compared to the HSP72-TCI8 complex. Moreover, the α-helices region where the covalent bond made showed a highly correlated motion in a HSP72-TCI2 complex system compared to a HSP72-TCI8 system. These results are correlated with RMSF and RMSD which justified that the α-helices region binding site exhibits a highly correlated motion during the

simulation per time. Taken together, these results suggested that the conformation flexibility and stability of the entire enzyme was improved by HSP72-TCI2 inhibitor binding. The HSP72-TCI2 covalent binding site shows a remarkable positively high correlating atoms motion of the α-helices binding region, and mostly for the complete enzyme compare to HSP72-TCI8 complex. These findings agree with the conformation changes of the HSP72 enzyme, which it agrees on the other hand of the strong covalent bond inhibition of the TCI2 toward HSP72 protein.

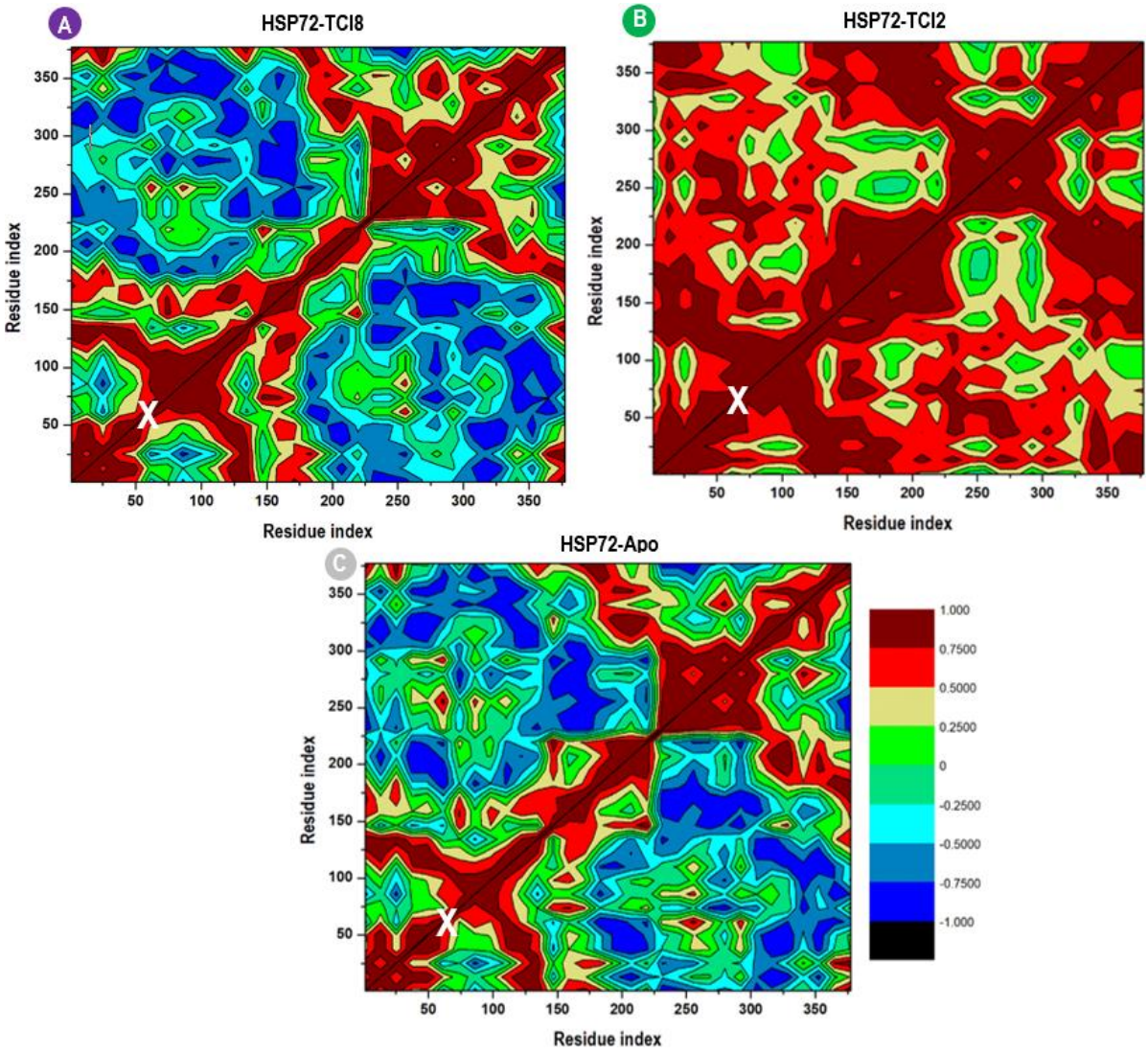


***Figure 7.7*:** Dynamic cross-correlation matrix analyses for HSP72-TCI8 **[A]** and HSP72-TCI2 **[B]** with unbound HSP72-Apo **[C]**. Numbers closer to 1 indicate high correlation, while those closer to -1 indicate anticorrelation between pairs of residues. X denoted the α-helices binding site of HSP72 via Lysine56.

7.3.2 Residual overall motions clustering and principal component analysis

Principle component analysis can convert the high-dimensional data of protein dynamics into the low-dimensional space to obtain a series of eigenvectors and eigenvalues that replicate overall motions in the protein [35,36]. The PCA analysis can be applied to any system and permits to study the effect of any varying parameters, by reducing the density of the collective motion [37,38], which is associated with the phase space behavior related to protein functions and stability. Consequently, it is often used to distinguish different conformational variances, which are involved in protein open-close mechanism and folding of ion channels, and conformational dynamics [39–41].

Principle component analysis and clustering have been carried out for HSP72-TCI2 and HSP72-TCI8 complexes to realize the overall concerted motion of HSP72 protein. The structural distribution of conformational changes and the proportion of variance of the captured eigenvectors are shown in (Figure 7.8A & B). The first three principal components accounted for 47.3 % and 53.2% of the total variance observed in the MD trajectories for HSP72-TCI2 and HSP72-TCI8 complexes. The clustering analysis also shows conformational distribution variance along with the first, second, and third principal components with each dot representing a single complex conformation. For the HSP72-TCI2 complex, the PC 1/3 and PC 1/2 conformational subspace highlight distinct thermodynamic periodic jumps with a significant energy barrier. However, the HSP72-TCI8 complex displayed a uniform and overlapping conformational PC 1/3 and PC 1/2 subspace with a higher motion energy barrier. The clustering and principal component analyses suggest the HSP72 protein undergoes minimal fluctuated conformational change and more stable structure of the nucleotide-binding domain (NBD) with HSP72-TCI compared to HSP72-TCI8.

The first PC1 pdbs were rendered (Figure 7.8 a & b) to getting insights into the regions of protein being fluctuated during the simulation and Figure 8 a & b depicts conformational variations on timestep. It can be seen from Figure 8 a & b that, residues mobility of HSP72-

TCI8 (mostly α-helix region), of ATP binding site of HSP72 have demonstrated highest flexibility than HSP72-TCI2. Notably, other high fluctuations were observed for HSP72-TCI8, within deferent regions. Thus, this compressed analysis supports earlier RMSF calculation.

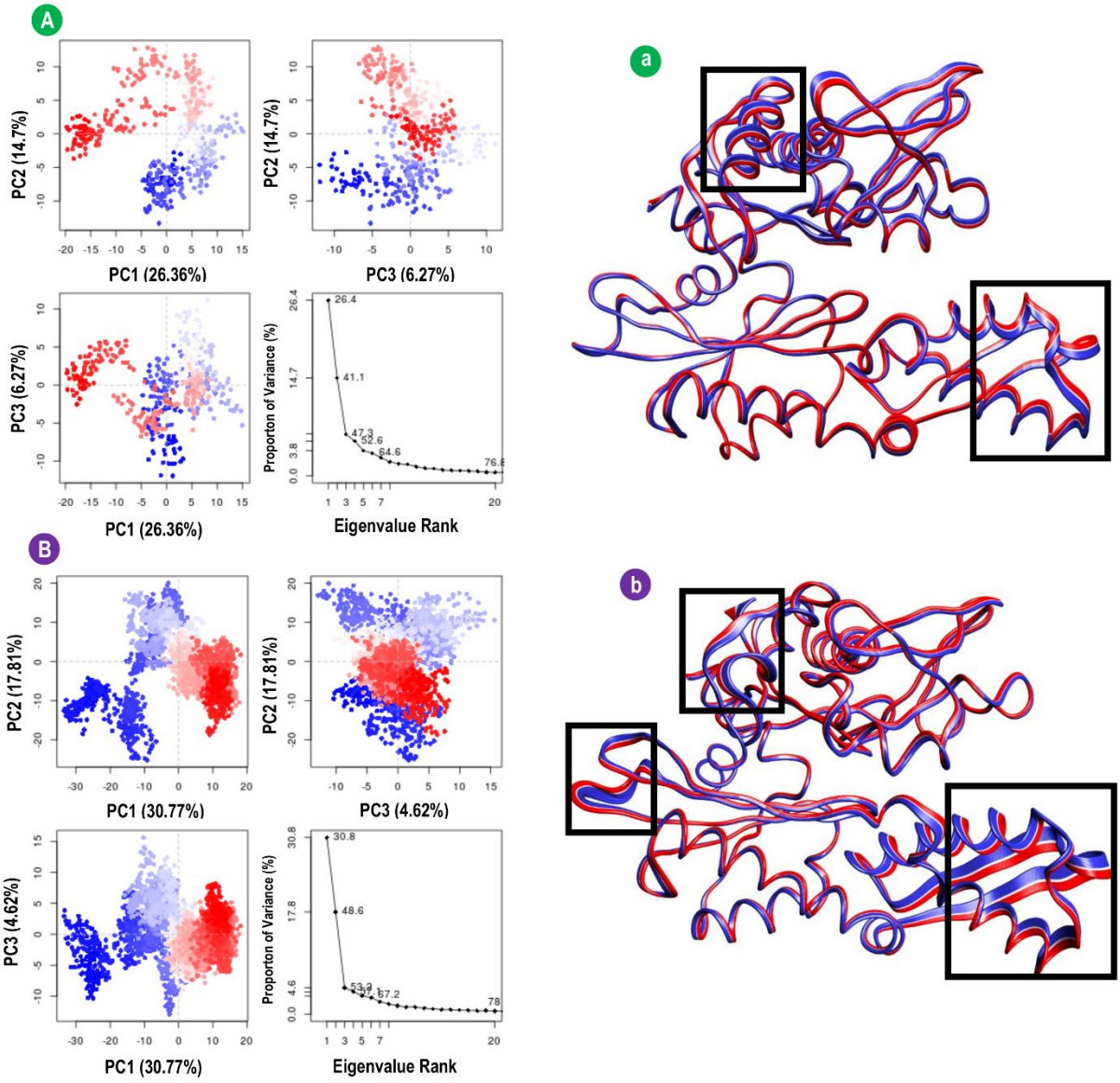


***Figure 7.8*:** Clustering and principal component analysis on 250ns of equidistance conformations using the Bio3D package in R. The plots show the first three eigenvectors for HSP72-TCI2 complex **[A]** and HSP72-TCI8 complex **[B]**, Conformers are colored according to the k-means clustering: cluster 1, blue; 2, red; 3, white. Dominant motions and captured eigenvector variance; and residue mobility for HSP72-TCI2 complex **[a]** and HSP72-TCI8 complex **[b].** The rectangular black shaped box marks the most flexible region.

7.3.3 Understand the structural dynamics motion of the covalent inhibitors at the HSP72 ATP binding site domain.

Binding to N-terminal of the nucleotide-binding domain ATP site (NBD) can stop the folding circle process. Hence, nucleotide-binding domain (NBD) inhibition is considered one of the most promising strategies for HSP72 function inhibition. Accordingly, the ATP binding domain is the most favourable binding site for the ligands binding and usually comes with a high inhibition effect on HSP72 [42]. To further gain an additional structural understanding of the preferential covalent inhibition of HSP72-TCI2 over HSP72-TCI8 inhibition, we essentially compared both covalent bonds dynamic motions at ATP binding subdomain α-helix site throughout the simulation. Inhibiting the HSP72 by TCI2 covalently bonded via Lys56, showing more stability and high correlated motion for the protein, TCI8 showed less as compare. Moreover, our findings revealed the favourable covalent bond for HSP72-TCI2 complex comprehensive and exact according to the previous analyses data. Interestingly, HSP72-TCI2 covalent bond shows an attractive, dynamic motion towards the ATP binding domain, as a result, during the 250ns MD time (Figure 7.9). It also makes it one of the main reasons it is more favourable, stable, less fluctuate and more correlated as a covalent bond compares to the HSP72-TCI8 complex.

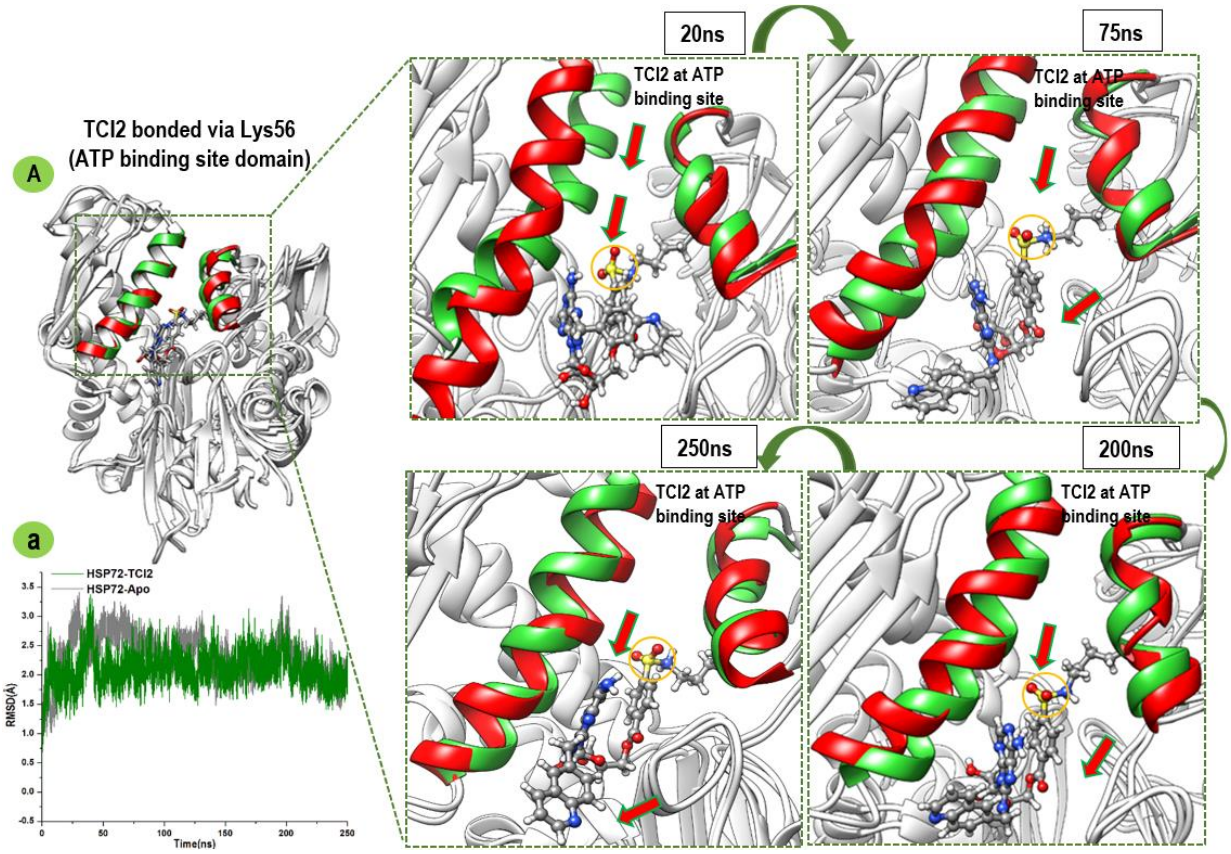


Figure 7.9: The conformational dynamic motion of HSP72-TCI2 covalent bond toward ATP binding domain are indicated in red arrows **[A].** Comparative C-α RMSD plots showing the degree of stability induced by TCI2 at the ATP site **[a]**. Also shown is the ATP binding α-helix site superposition of unbound (**red**) and HSP72-TCI2 (**green**) form of the protein. (orange circle shows the covalent bond via Lys56).

As evident in (Figure 7.9 &7.10), both covalent bonds' orientation can be explained based on the previous findings. The HSP72-TCI2 system enzyme displayed high dynamic motions directed towards the ATP binding site compared to the HSP72-TCI8 complex as it moved behind and above ATP binding subdomain α-helix site. The observed structural changes in the NBD α-helix motion show a different inhibition mechanism caused by both covalent bonds; hence, controlling the ATP binding site can be define as the highest inhibition of HSP72 (Figure 7.9 &7.10). Noticeably, the HSP72-TCI2 controlling the binding site by heading to the ATP binding groove during the 250ns, and that can be enough reason to explain the high correlation dynamic motion for HSP72-TCI2.

On the other hand, HSP72-TCI8 can be observed in a different direction of the covalent bond, heading above and behind the ATP binding site, displaying less correlation dynamic motion and less stability during the simulation. (Figure 7.10) The α-helices region where the covalent bond made showed a highly conformation motion and flexibility in a HSP72-TCI8 system compared to a HSP72-TCI2 system due to the high fluctuation.

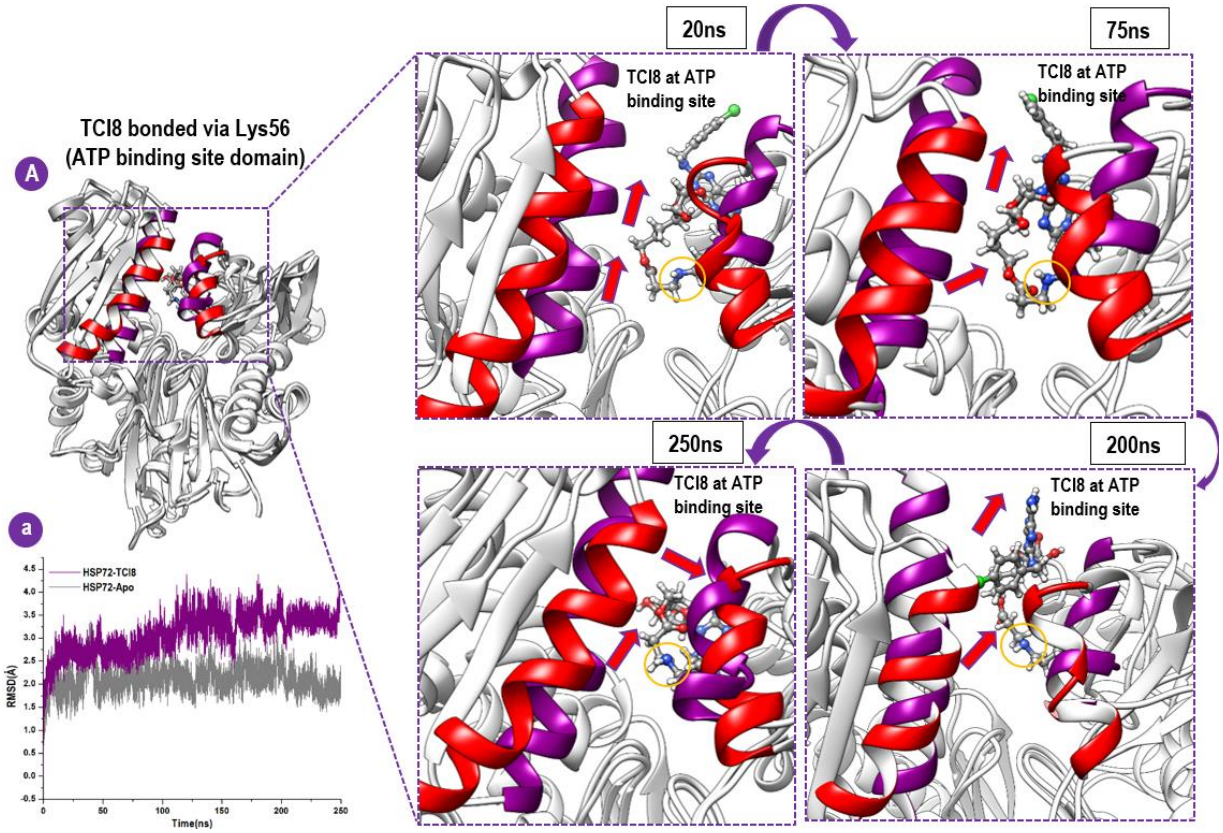


***Figure 7.10*:** The conformational dynamic motion of HSP72-TCI8 covalent bond toward ATP binding domain are indicated in red arrows **[A].** Comparative C-α RMSD plots showing the degree of instability and disruption induced by TCI8 at the ATP site **[a]**. Also shown is the ATP binding α-helix site superposition of unbound (**red**) and HSP72-TCI2 (**magnet**) form of the protein. (orange circle shows the covalent bond via Lys56).

Furthermore, ligand-residue interaction is fundamental to the identification of key binding site residues and their corresponding interactions with the irreversible drugs. Nevertheless, they don't consider as a strong evidence for the covalent bond interactions, since the drug is directly bonded with a certain amino acid, beside the remarkable immediate effect on the enzyme as compare to the non-covalent drug, regardless, of their benefits and side effects. For these

reasons, molecular visualization was used to obtain insights into the activity of both irreversible inhibitors covalently bonded via Lys56 at the ATP binding domain site of HSP72 with respect to their affinities, stabilizations and covalent-inhibitions. As shown in (Figure 7.9), certain residues of the hydrophobic covalent binding site elicited strong interactions with TCI2 inhibitor, which accounted for its stability and high affinity. These interactions vary from strong ionic interactions, salt bridges, conventional and non-conventional hydrogen bond types to weak aromatic π interactions.

To understand the interaction dynamics of HSP72-TCI2 and HSP72-TCI8 at the binding site of HSP72 covalently bonded via lysine56, we noted certain residues that interacted directly, steadily and intermittently across the 250ns MD simulation time. These modes of interactions could relatively provide essential details and insights into the covalent molecular basis and mechanisms of HSP72-TCI2 selectivity and high-affinity binding towards HSP72.

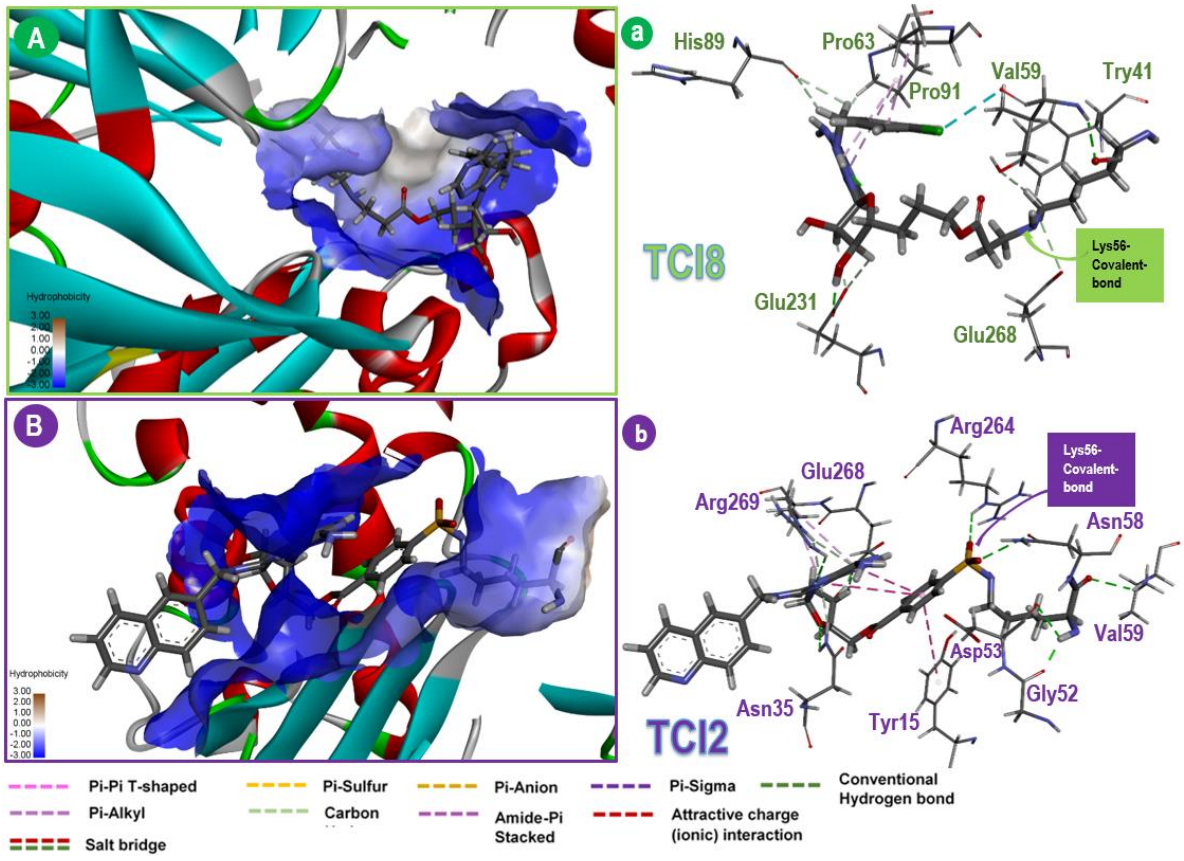


***Figure 7.11*:** Molecular visualization of TIC8 and TCI2 at the ATP binding sites (hydrophobic pockets of covalent bonds) of **[A]** HSP72-TCI8 **[B]** HSP72-TCI2. Inter-molecular interactions

between two covalent inhibitors as mentioned and binding site residues in HSP72 complexes are shown in **[a]** and **[b]** respectively. Hydrophobic surface representations of TCI8 and TCI2 covalent-bounds binding sites. The degree of hydrophobicity ranges from least (-3 → blue) to highest (+3 → deep brown).

As shown in (Figure 7.11) above, among all interactions caused by HSP72-TCI2 and HSP72-TCI8 covalent binding, only two amino acids in common for both inhibitor interactions Val59 and Glu268, the rest of them are different. Moreover, the numbers of the interacted residues for both inhibitors are nine remarkable interacted residues (Arg269, Glu268, Arg264, Asn58, Asn35, Tyr15, Asp53, Gly52 and Val59) for HSP72-TCI2 and seven significant residues for the HSP72-TCI8 (His89, Pro63, Pro91, Val59, Try41, Glu268 and Glu231), as shown in figure 11. Consequently, perhaps one of the reasons why is HSP72-TCI2 is better stability and less fluctuate more than HSP72-TCI8 is the high number of the interacted residue which forms direct hydrogen bonds alongside other bonds. Thus, as we mentioned, the key interaction residue is not strong evidence for the covalent bond outcome, but the bonded amino acid itself, and how it interacts covalently with the irreversible drug. Therefore, the formed covalent bond with TCI2 via lysine 56 shows a better affinity, acceptance and favourable as a strong bond for the inhibition of HSP72 irreversibly. Hence, based on these findings, the covalent bond of HSP72-TCI8 via lysine56 will eventually be disconnected as an effect of irreversible drug way faster compared to HSP72-TCI2, and that is the essential extraordinary meaning of the enhancement and development for the covalent drugs generation after another.

7.4 Conclusion

The human heat shock protein 72 (HSP72) represents a vital therapeutic target during the critical stages of oncogenesis and progression of human cancers. This study was undertaken to compare two covalent inhibition modes of HSP72 at the nucleotide binding domain: HSP72-TCI2 and HSP72-TCI8 in order to gain a better understanding of the underlying structural dynamics of each binding mechanism. The performed MD analyses provided in this study established the most optimal ligand binding mechanism at the HSP72-NBD enzyme. Molecular mechanism revealed that the HSP72-TCI2 was more able to inhibit and induce stabilization of the HSP72-NBD ATP binding site throughout the simulation than the HSP72-TCI8 complex. This was further confirmed by extensive analyses of the structural basis and conformational dynamics associated with the preferential inhibition of the covalent TCI2 over the TCI8 covalent ligand in HSP72-NBD models.

Our findings revealed that the dynamic motion of the irreversible inhibitor TCI2 toward ATP binding site domain significantly destabilizes the protein, coupled with a decrease in surface-exposed residues. These observed perturbations, which occur as a result of TCI2 irreversible inhibition, could in turn alter substrate binding and catalysis, indicative of enzyme inhibition. Disruption in HSP72-NBD conformation upon HSP72-TCI2 irreversible inhibition was also revealed by the DCCM which showed a highly correlative residual movement in the protein structure compared with the HSP72-TCI8 and unbound systems. This study will assist with the design different generations of more potent and highly selective covalent inhibitors for HSP72 with the potential to overcome drug resistance challenges and represent a novel therapeutic approach for inhibiting HSP72 oncoprotein.

Acknowledgements

The authors acknowledge the College of Health Science of the University of KwaZulu-Natal for funding and the Centre for High-Performance Computing (CHPC), Cape Town, South Africa, for computational resources (www.chpc.ac.za).

Conflict of interest

The authors declare no conflict of interest.

References

1. Baillie TA. Targeted Covalent Inhibitors for Drug Design. *Angew. Chemie Int. Ed.* [Internet]. 55(43), 13408–13421 (2016). Available from: http://doi.wiley.com/10.1002/anie.201601091.

2. Johnson DS, Weerapana E, Cravatt BF. Strategies for discovering and derisking covalent, irreversible enzyme inhibitors. *Future Med. Chem.* [Internet]. 2(6), 949–964 (2010). Available from: http://www.future-science.com/doi/10.4155/fmc.10.21.

3. Bauer RA. Covalent inhibitors in drug discovery: from accidental discoveries to avoided liabilities and designed therapies. *Drug Discov. Today* [Internet]. 20(9), 1061–1073 (2015). Available from: https://linkinghub.elsevier.com/retrieve/pii/S135964461500183X.

4. Engel J, Richters A, Getlik M, *et al.* Targeting Drug Resistance in EGFR with Covalent Inhibitors: A Structure-Based Design Approach. *J. Med. Chem.* [Internet]. 58(17), 6844–6863 (2015). Available from: https://pubs.acs.org/doi/10.1021/acs.jmedchem.5b01082.

5. Marino SM, Gladyshev VN. Cysteine Function Governs Its Conservation and Degeneration and Restricts Its Utilization on Protein Surfaces. *J. Mol. Biol.* [Internet]. 404(5), 902–916 (2010). Available from: https://linkinghub.elsevier.com/retrieve/pii/S0022283610010156.

6. Jakob CG, Upadhyay AK, Donner PL, *et al.* Novel Modes of Inhibition of Wild-Type Isocitrate Dehydrogenase 1 (IDH1): Direct Covalent Modification of His315. *J. Med. Chem.* [Internet]. 61(15), 6647–6657 (2018). Available from: https://pubs.acs.org/doi/10.1021/acs.jmedchem.8b00305.

7. Kharenko OA, Patel RG, Brown SD, *et al.* Design and Characterization of Novel Covalent Bromodomain and Extra-Terminal Domain (BET) Inhibitors Targeting a Methionine. *J. Med. Chem.* [Internet]. 61(18), 8202–8211 (2018). Available from: https://pubs.acs.org/doi/10.1021/acs.jmedchem.8b00666.

8. Mukherjee H, Grimster NP. Beyond cysteine: recent developments in the area of targeted covalent inhibition. *Curr. Opin. Chem. Biol.* [Internet]. 44, 30–38 (2018). Available from: https://linkinghub.elsevier.com/retrieve/pii/S1367593118300048.

9. Strelow JM. A Perspective on the Kinetics of Covalent and Irreversible Inhibition. *SLAS Discov. Adv. Sci. Drug Discov.* [Internet]. 22(1), 3–20 (2017). Available from: http://journals.sagepub.com/doi/10.1177/1087057116671509.

10. Fernández-Fernández MR, Valpuesta JM. Hsp70 chaperone: A master player in protein homeostasis [version 1; peer review: 3 approved]. *F1000Research* [Internet]. 7, 1497 (2018). Available from: https://f1000research.com/articles/7-1497/v1.

11. Kumar SJ, Stokes J, Singh UP, *et al.* Targeting Hsp70: A possible therapy for cancer. *Cancer Lett.* [Internet]. 374(1), 156–166 (2016). Available from: http://dx.doi.org/10.1016/j.canlet.2016.01.056.

12. Powers M V., Jones K, Barillari C, Westwood I, Montfort RLM van, Workman P. Targeting HSP70: The second potentially druggable heat shock protein and molecular chaperone? *Cell Cycle*. 9(8), 1542–1550 (2010).

13. Mayer MP, Bukau B. Hsp70 chaperones: Cellular functions and molecular mechanism. *Cell. Mol. Life Sci.* 62(6), 670–684 (2005).

14. Zhuravleva A, Clerico EM, Gierasch LM. An interdomain energetic tug-of-war creates the allosterically active state in Hsp70 molecular chaperones. *Cell*. 151(6), 1296–307 (2012).

15. Chiappori F, Merelli I, Colombo G, Milanesi L, Morra G. Molecular Mechanism of Allosteric Communication in Hsp70 Revealed by Molecular Dynamics Simulations. *PLoS Comput. Biol.* 8(12), e1002844 (2012).

16. Penkler D, Sensoy Ö, Atilgan C, Tastan Bishop Ö. Perturbation–Response Scanning Reveals Key Residues for Allosteric Control in Hsp70. *J. Chem. Inf. Model.* 57(6), 1359–1374 (2017).

17. Sharma D, Masison DC. Hsp70 structure, function, regulation and influence on yeast prions. *Protein Pept. Lett.* 16(6), 571–81 (2009).

18. Pettinger J, Jones K, Cheeseman MD. Lysine-Targeting Covalent Inhibitors. *Angew. Chemie Int. Ed.* [Internet]. 56(48), 15200–15209 (2017). Available from: http://doi.wiley.com/10.1002/anie.201707630.

19. Hacker SM, Backus KM, Lazear MR, Forli S, Correia BE, Cravatt BF. Global profiling of lysine reactivity and ligandability in the human proteome. *Nat. Chem.* [Internet]. 9(12), 1181–1190 (2017). Available from: http://www.nature.com/articles/nchem.2826.

20. Pettinger J, Carter M, Jones K, Cheeseman MD. Kinetic Optimization of Lysine-Targeting Covalent Inhibitors of HSP72. *J. Med. Chem.* 62(24), 11383–11398 (2019).

21. Berman HM, Battistuz T, Bhat TN, *et al.* The Protein Data Bank. *Biol. Crystallogr.* 58, 899–907 (2002).

22. Schlecht R, Scholz SR, Dahmen H, *et al.* Functional analysis of Hsp70 inhibitors. *PLoS One*. 8(11), 1–12 (2013).

23. Pettersen EF, Goddard TD, Huang CC, *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* 25(13), 1605–1612 (2004).

24. Kusumaningrum S, Budianto E, Kosela S, Sumaryono W, Juniarti F. The molecular docking of 1,4-naphthoquinone derivatives as inhibitors of Polo-like kinase 1 using Molegro Virtual Docker. *J. Appl. Pharm. Sci.* 4(11), 47–53 (2014).

25. Thomsen R, Christensen MH. MolDock: A new technique for high-accuracy molecular docking. *J. Med. Chem.* 49(11), 3315–3321 (2006).

26. Allouche A. Software News and Updates Gabedit — A Graphical User Interface for Computational Chemistry Softwares. *J. Comput. Chem.* 32, 174–182 (2012).

27. Trott O, Olson A. Autodock vina: improving the speed and accuracy of docking. *J. Comput. Chem.* 31(2), 455–461 (2010).

28. Madhavi Sastry G, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided. Mol. Des.* 27(3), 221–234 (2013).

29. Mhlongo NN, Ebrahim M, Skelton AA, Kruger HG, Williams IH, Soliman MES. Dynamics of the thumb-finger regions in a GH11 xylanase Bacillus circulans:

comparison between the Michaelis and covalent intermediate. *RSC Adv.* 5(100), 82381–82394 (2015).

30. Ramharack P, Oguntade S, Soliman MES. Delving into Zika virus structural dynamics-a closer look at NS3 helicase loop flexibility and its role in drug discovery. *RSC Adv.* 7(36), 22133–22144 (2017).

31. Kasahara K, Fukuda I, Nakamura H. A novel approach of dynamic cross correlation analysis on molecular dynamics simulations and its application to Ets1 dimer-DNA complex. *PLoS One*. 9(11) (2014).

32. Seifert E. OriginPro 9.1: Scientific data analysis and graphing software - Software review. *J. Chem. Inf. Model.* 54(5), 1552 (2014).

33. Pan L, Patterson JC. Molecular Dynamics Study of Zn(Aβ) and Zn(Aβ)2. *PLoS One*. 8(9), 70681–70688 (2013).

34. Richmond TJ. Solvent accessible surface area and excluded volume in proteins. Analytical equations for overlapping spheres and implications for the hydrophobic effect. *J. Mol. Biol.* 178(1), 63–89 (1984).

35. Yan F, Liu X, Zhang S, Su J, Zhang Q, Chen J. Electrostatic interaction-mediated conformational changes of adipocyte fatty acid binding protein probed by molecular dynamics simulation. *J. Biomol. Struct. Dyn.* [Internet]. 37(14), 3583–3595 (2019). Available from: https://www.tandfonline.com/doi/full/10.1080/07391102.2018.1520648.

36. Shukla R, Singh TR. Virtual screening, pharmacokinetics, molecular dynamics and binding free energy analysis for small natural molecules against cyclin-dependent kinase 5 for Alzheimer's disease. *J. Biomol. Struct. Dyn.* [Internet]. 38(1), 248–262 (2020). Available from: https://www.tandfonline.com/doi/full/10.1080/07391102.2019.1571947.

37. Kitao A, Hirata F, Gō N. The effects of solvent on the conformation and the collective motions of protein: Normal mode analysis and molecular dynamics simulations of melittin in water and in vacuum. *Chem. Phys.* [Internet]. 158(2–3), 447–472 (1991). Available from: https://linkinghub.elsevier.com/retrieve/pii/0301010491870827.

38. Kitao A, Go N. Investigating protein dynamics in collective coordinate space. *Curr. Opin. Struct. Biol.* [Internet]. 9(2), 164–169 (1999). Available from: https://linkinghub.elsevier.com/retrieve/pii/S0959440X99800232.

39. Maisuradze GG, Liwo A, Scheraga HA. Principal Component Analysis for Protein Folding Dynamics. *J. Mol. Biol.* [Internet]. 385(1), 312–329 (2009). Available from: https://linkinghub.elsevier.com/retrieve/pii/S0022283608012886.

40. Maisuradze GG, Liwo A, Scheraga HA. Relation between Free Energy Landscapes of Proteins and Dynamics. *J. Chem. Theory Comput.* [Internet]. 6(2), 583–595 (2010). Available from: https://pubs.acs.org/doi/10.1021/ct9005745.

41. Spellmon N, Sun X, Sirinupong N, Edwards B, Li C, Yang Z. Molecular Dynamics Simulation Reveals Correlated Inter-Lobe Motion in Protein Lysine Methyltransferase SMYD2. *PLoS One* [Internet]. 10(12), e0145758 (2015). Available from: https://dx.plos.org/10.1371/journal.pone.0145758.

42. Bauer D, Meinhold S, Jakob RP, *et al.* A folding nucleus and minimal ATP binding

domain of Hsp70 identified by single-molecule force spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* [Internet]. 115(18), 4666–4671 (2018). Available from: http://www.pnas.org/lookup/doi/10.1073/pnas.1716899115.

CONCLUSION AND FUTURE PERSPECTIVES

1. Conclusion

Cancer has continued to render existing therapies ineffective due to incessant incidences of resistance, apoptotic evasion and disease relapse. Hence, the search for novel treatment strategies have been continuous, particularly those that incorporate the identification of crucial oncogenic targets with the design of highly selective inhibitory molecules; an approach that reduces unwanted interactions with biological non-targets, which basically engenders toxic effects.

Covalent drug design is a rapidly growing and promising field in the drug discovery. Despite of the various examples of effective covalent drugs, the fundamentals for the rational design of these small compounds have only developed recently. The number and knowledge of covalent inhibitors is continuously increasing every day. There is a further possibility of major advancement in this area, more specifically around targeting the non-catalytic amino acid residues. Covalent inhibition has emerged as a therapeutic approach for targeting 'undruggable' pathogenic proteins. Likewise, computer aided covalent drug design have contributed greatly to the drug discovery and development of covalent inhibitors in cancer therapy.

Heat shock proteins (HSPs) play a crucial role in the clearance of damaged proteins by inducing proteins acclamation and proteotoxicity. HSP72 is a crucial enzyme in the protein degradation process that has been proved to involve in various human disease including cancer. This enzyme encloses the ATP binding domain that can be targeted either covalently or non-covalently. In this project, we sought to expand the structural and molecular characteristics of the HSP72 as an important target in cancer therapy. Then, create a CADD route map to identify the potential inhibitors of this enzyme and finally, implement the route map to investigate preferred inhibitors for the enzyme.

The major objectives of this project were:

- Covering the existing gap in drug discovery literature regarding covalent inhibition and highlighting its advancements and drawbacks and assemble the computational tools available for analyzing these systems.

- The covalent inhibition at the N-terminal of the nucleotide-binding domain (NBD) site of the HSP72 was determined as the optimal binding mechanism mode over the non-covalent inhibition using molecular dynamics simulations.
- The same covalent inhibitor was used to study the structural variations between the Lysine56 covalent bond and cysteine17 covalent bond of the HSP72. These comparisons have aided in the understanding of the structural and conformational dynamics and the underlying inhibition mechanisms of this peculiar enzyme. Various analyses have been performed and were validated by calculating the binding free energies.
- Comparison of two novel and potent covalent inhibitors as different generation and implement it to identify the optimal binding inhibition profile at the N-terminal of the nucleotide-binding domain of the HSP72 enzyme.

Upon commencement of the study, there were limitations of *in silico* methods to simulate covalent complexes using all atom MM approach. There were difficulties associated with the use of currently available methods that are inaccurate, unreliable and irreproducible. These methods were failed to define and to parameterize the formation of covalent bond and structural dynamics of protein associated to covalent inhibition. Hence, we established the first accurate, well-documented and more reliable protocol for simulating the covalent inhibitor-enzyme complex. This protocol is reproducible and has also been validated in our recent studies. The protocol is based on the all-atom MM approach. The ability to explore covalent interactions within the protein and covalent inhibitor using this technical protocol will assist significantly towards the development of covalent inhibition in the therapy of incurable diseases. Overall, this study has provided valuable insights into the design and development of covalent inhibitors through molecular modelling and CADD.

2. Future perspectives

Approximately 30% of the overall drugs available in the market are inactivating their enzymes by covalent mode of inhibition. Covalent ligands have many potentials containing better biochemical efficacy of target interruption, prolonged duration of action that sustain molecule's pharmacokinetics and lower drug dosing. The covalent inhibition approach is gaining acceptance as a favorable tool in pharmaceutical industries and will also make major effect in

future to design more potent covalent inhibitors.

The work reported here, reveals an appropriate and reliable protocol to simulate irreversible covalent inhibitors of target proteins and can be used as a beneficial reference for computational chemists towards the design and development of covalent inhibitors. Furthermore, the study described herein could serve as a scaffold for the synthetic design of highly effective and selective of novel lead compounds against the activity of HSP72 in malignant cells. This could present a ‘low cost, less toxic’ ideal starting material for drug design in cancer therapeutics.

APPENDIX

APPENDIX A

Supplementary Documents for published Article: Aimen Aljoundi, Ahmed El Rashedy, Patrick Appiah-Kubi, and Mahmoud E.S Soliman (2020) “Coupling of HSP72 α-helix subdomains by the unexpected irreversible targeting of Lysine-56 over Cysteine-17; coevolution of covalent bonding”, *Molecules Journal.*

APPENDIX B

Supplementary Documents for Accepted for publication Article: Aimen Aljoundi, Ahmed El Rashedy, Mahmoud E.S Soliman (2021) “Distinguishing the optimal binding mechanism through reversible and irreversible inhibition analysis of HSP72 protein in cancer therapy” *Journal of Computer in Biology and Medicine.*

APPENDIX B

Supplementary Documents for Submitted Article: Aimen Aljoundi, Ahmed El Rashedy, Mahmoud E.S Soliman (2021) “Comparison of irreversible inhibition targeting HSP72 protein - the resurgence of covalent drug developments” *Journal of Molecular simulation.*

Coupling of HSP72 α-helix subdomains by the unexpected irreversible targeting of Lysine-56 over Cysteine-17; coevolution of covalent bonding

Supplementary Material:

Aimen Aljoundi, Ahmed El Rashedy, Patrick Appiah-Kubi, and Mahmoud E.S Soliman

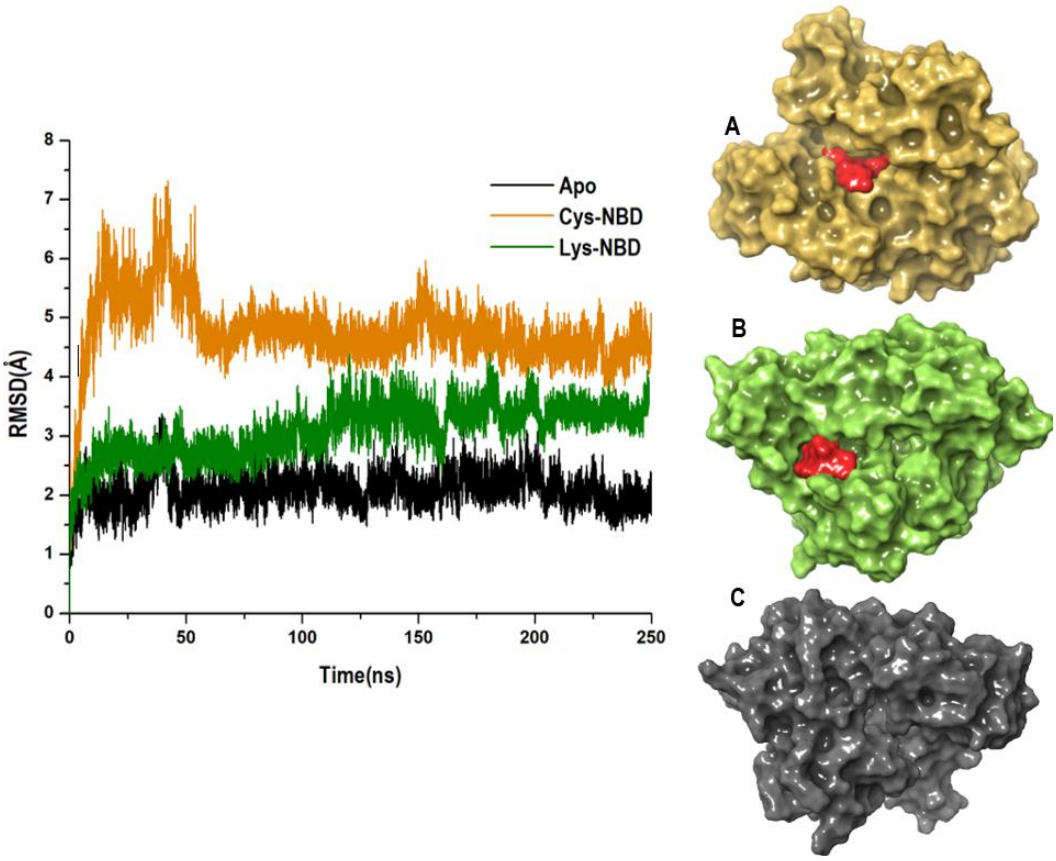


Figure S5.1. Comparative C-α RMSD plots showing the degree of stability and convergence of the studied systems over the 250ns MD simulation time; **[A]** Cysteine-NBD**[B]** Lysine-NBD **[C]** Apo.

Distinguishing the optimal binding mechanism through reversible and irreversible inhibition analysis of HSP72 protein in cancer therapy

Supplementary Material:

Aimen Aljoundi, Ahmed El Rashedy, Mahmoud E.S Soliman

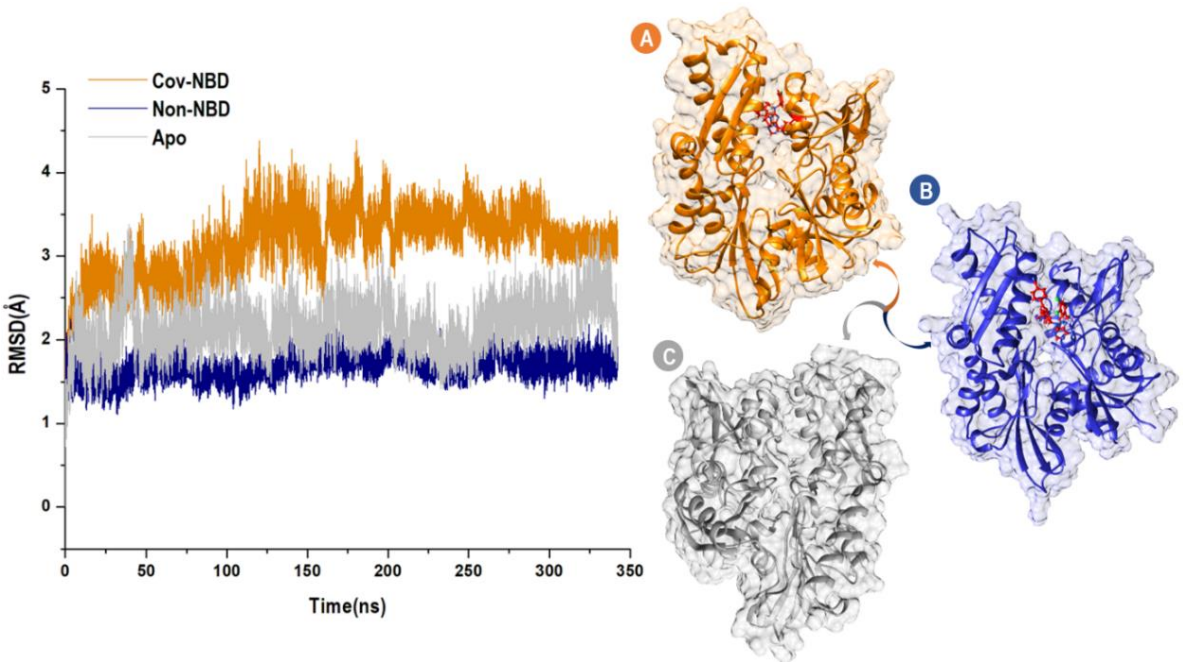


Figure S6.1. Comparative C-α RMSD plots showing the degree of stability and convergence of the studied systems over the 350ns MD simulation time; **[A]** Cov-NBD **[B]** Non-NBD **[C]** Apo.

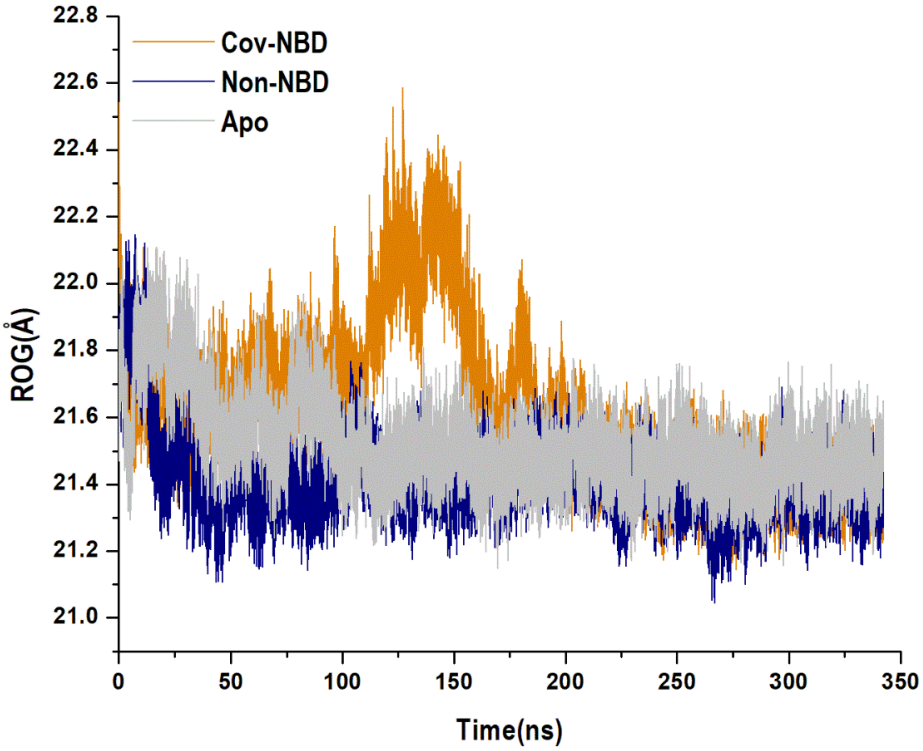


Figure S6.2. The radius of Gyration **(**ROG**)** of Cα atoms of NBD protein residues Apo-NBD **(**grey color**)**, and Non-NBD **(**Blue color**)**, and Cov-NBD **(**orange color**)**.

Comparison of irreversible inhibition targeting HSP72 protein - the resurgence of covalent drug developments

Supplementary Material:

Aimen Aljoundi, Ahmed El Rashedy, Mahmoud E.S Soliman

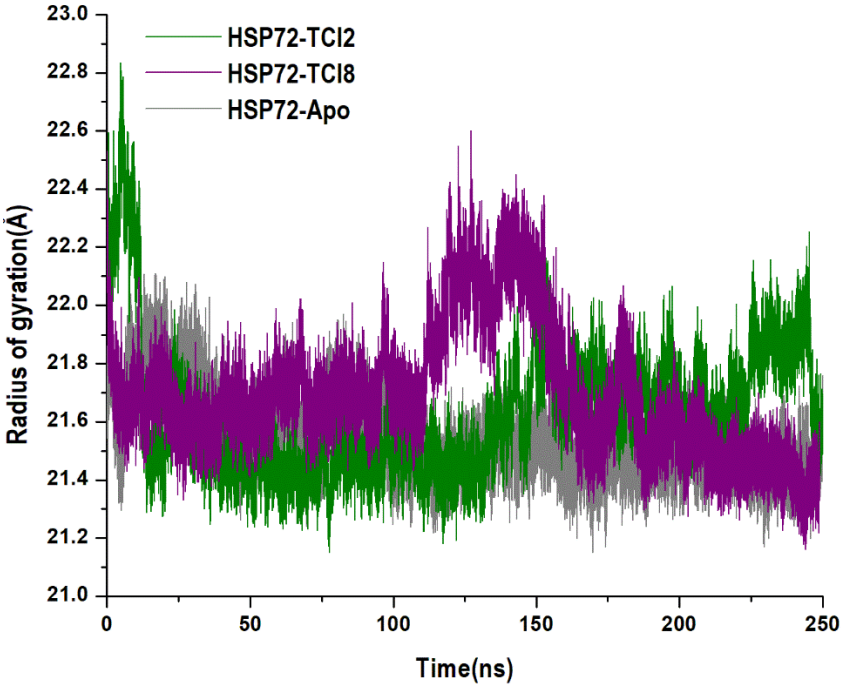


Figure S7.1: The radius of Gyration **(**ROG**)** of Cα atoms of HSP72 protein residues HSP72-Apo **(grey)**, HSP72-TCI2 **(green**), and HSP72-TCI8 **(magnet)**.