INSIGHT INTO CANCER TARGETS AND LIGAND BINDING LANDSCAPE USING BIOINFORMATICS AND INTEGRATED MOLECULAR MODELING TOOLS

By

UMAR NDAGI

215081668

A thesis submitted to College of Health Sciences, University of KwaZulu-Natal, in fulfilment of the requirements of the degree of

Doctor of Philosophy

In

Medical Science (Pharmaceutical Chemistry)

2017

This is to certify that the contents of this thesis are the original research work of Umar Ndagi

As the candidate supervisor, I have approve this thesis for submission

Supervisor

Sign:----------Name: Prof Mahmoud E Soliman Date ----------

Abstract

The alarming rate of varying types of cancer diseases in human remains a global burden requiring drastic treatment in which, a prominent method of combating it is through enzyme-based drug design. Metastatic castration-resistant prostate cancer (mCRPC) and triple-negative breast cancer (TNBC) are deadly forms of prostate and breast cancers, respectively. The later cancerous growth has been linked to non-receptor tyrosine (Src/p38) kinase as a potential targeted enzyme for possible chemotherapeutic control while, mCRPC have recently been linked to retinoic acid-related orphan-receptor gamma (ROR-γ).

Most studies on ROR-γ usually relate it as an orphan due to low or zero possibility to identify potential inhibitor for this receptor. Amazingly, promising inhibitors of ROR-γ and their therapeutic potential were currently identified and evaluated experimentally, among which inhibitor XY018 has appreciable bioactivity. However, molecular understanding of the conformational features of XY018-ROR-γ complex is still elusive. Herein, we provide the first account of conformation details of XY018-ROR-γ using multiple computational approaches. Comparative molecular dynamics (MD) simulation of XY018-ROR-γ and hydroxycholesterol bound ROR-γ (HC9-ROR-γ) were carried out. This was widened to binding free energy calculation (MM/GBSA), principal component analysis (PCA), root mean square fluctuation (RMSF), radius of gyration (RoG) and ligand-residue interaction network. In addition, the *in silico* study was optimized to predict toxicity and biological activity of the identified ligand.

Findings from this study revealed that: (1) hydrophobic packing contributes significantly to binding free energy, (2) Ile136 and Leu60 exhibited high hydrogen-bond in both systems, (3) XY018-ROR-γ displayed a relatively high loop region residue fluctuation compared to ROR-γ bound to natural ligand HC9-ROR-γ, (4) electrostatic interactions are potential binding force in XY018-ROR-γ complex compared to HC9-ROR-γ, (5) XY018-ROR-γ assumes a rigid conformation which is highlighted by a decrease in residual fluctuation, (6) XY018 could potentially induce pseudoporphyria, nephritis, and interstitial nephritis but potentially safe in renal failure.

In vivo examination of UM-164 as a bioactive moiety against Src/p38 kinase was recently reported in literature. This ligand is a promising lead compound for developing the first targeted therapeutic strategy against triple-negative breast cancer (TNBC). However, the conformational features of UM-164 in complex with Src remained poorly explored towards the rational design of novel Src dual inhibitor. Similar to XY018-ROR-γ investigation, a

comprehensive account on the conformational features of Src-UM-164 and the influence of UM-164 binding to the Src using different computational approaches was also provided. This was carried out through MD simulation, principal component analysis (PCA), thermodynamic calculations, dynamic cross-correlation (DCCM) analysis and ligand-residue interaction network profile, as well as toxicity testing.

Analysis of results from this investigation revealed that: (1) the binding of UM-164 to Src induces a more stable and compact conformation on the protein structure; (2) UM-164 binding to Src induces highly correlated motions in the protein; (3) high fluctuation exhibited by the loops in Src-UM-164 system support the experimental evidence that UM-164 binds the DFG-out inactive conformation of Src; (4) a relatively high binding free energy estimated for the Src-UM-164 system is affirmative of its experimental potency; (5) hydrophobic packing contributes significantly to the drug binding in Src-UM-164; (6) a relatively high H-bond formation in Src-UM-164 indicates enhanced drug-protein interaction; (7) UM-164 is relatively less toxic than Dasatinib, therefore, is potentially safer.

Furthermore, a mutant form of Src was also investigated due to its drug resistivity character. Thr91 mutation was found to induce a complete loss of protein conformation required for drug fitness in c-Src. Computational studies were carried out on this mutant enzyme in complex with UM-164 as described in Src wild-type. A notable observation from binding free energy analysis results is that, a reduction in binding affinity up to -13.416 kcal/mol was estimated for this mutated candidate compared to the wild-type-UM-164.

This entire work provides an invaluable contribution to the understanding of dynamics of the orphan nuclear receptor (ROR-γ) and non-receptor tyrosine kinase (Src) which could largely contribute to the design of novel inhibitors to minimise the chances of drug resistance in castrated resistance prostate cancer and triple negative breast cancer, respectively.

Declaration 1-Plagiarism

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

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

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

This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other resources have been quoted, then their words have been re-written but the general information attributed to them has been referenced..

Where their exact words have been used, then their writing has been placed in italics and inside quotation marks and referenced.

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 and/or include research presented in this thesis is stated.

Signed----------

Declaration 2-Publications

1. Umar Ndagi, Ndumiso N Mhlongo and Mahmoud E Soliman (2016) Re-emergence of an Orphan Therapeutic target for the Treatment of Resistance Prostate Cancer, A Thorough conformational and Binding Analysis for ROR-γ protein. Journal of *Biomolecular Structure & Dynamics* " Manuscript ID TBSD-2016-0664 (**Accepted for publication**)

   Contributions:
   Umar Ndagi·: Conducted all the calculations and manuscript writing.
   Ndumiso N Mhlongo·: Provide editorial support
   Mahmoud E Soliman·: Supervisor.
   Appendix A : Acceptances letter

2. Umnar Ndagi, Ndumiso N Mhlongo and Mahmoud E Soliman (2016). Emergence of a Promising Lead compound in the Treatment of Triple Negative Breast Cancer: An Insight into Conformational Features and Ligand Binding Landscape of c-Src Protein with UM-164. *Journal of Biomolecular Structur e& Dynamics* Manuscript ID TBSD-2016-0690. **(Submitted for publication)**
   Umar Ndagi·: Conducted all the calculations and manuscript writing.
   Ndumiso N Mhlongo·: Provide editorial support
   Mahmoud E Soliman·: Supervisor.
   Appendix B : Acknowledgement letter.

3. Umar Ndagi, Ndumiso N Mhlongo and Mahmoud E Soliman (2016). The impact of Thr91 mutation on c-Src resistance to UM-164: Molecular dynamics study revealed new opportunity for drug design. *Molecular Biosystems* MB-ART-12-2016-000848 **(Submitted for publication)**
   Umar Ndagi·: Conducted all the calculations and manuscript writing.
   Ndumiso N Mhlongo·: Provide editorial support
   Mahmoud E Soliman·: Supervisor.
   Appendix C : Acknowledgement letter.

4. Umar Ndagi, Ndumiso Mhlongo and Mahmoud E Soliman (2016). "Metal Complexes in Cancer Therapy – an update from Drug Design Perspective. *Drug Design, Development and Therapy* **(Accepted for publication)**
   Umar Ndagi·: Conducted all the calculations and manuscript writing.

Ndumiso N Mhlongo·: Provide editorial support.

Mahmoud E Soliman·: Supervisor.

Appendix D: Acceptance letter.

Research output

1 Publications

1. Umar Ndagi, Ndumiso N Mhlongo and Mahmoud E Soliman (2016) Re-emergence of an Orphan Therapeutic target for the Treatment of Resistance Prostate Cancer, A Thorough conformational and Binding Analysis for ROR-γ protein. Journal of *Biomolecular Structure & Dynamics* " Manuscript ID TBSD-2016-0664 (**Accepted for publication**)
2. Umar Ndagi, Ndumiso Mhlongo and Mahmoud E Soliman (2016). "Metal Complexes in Cancer Therapy – an update from Drug Design Perspective. *Drug Design, Development and Therapy* **(Accepted for publication)**

2 Submitted for publication

1. Umar Ndagi, Ndumiso N Mhlongo and Mahmoud E Soliman (2016). Emergence of a Promising Lead compound in the Treatment of Triple Negative Breast Cancer: An Insight into Conformational Features and Ligand Binding Landscape of c-Src Protein with UM-164. *Journal of Biomolecular Structur e& Dynamics* Manuscript ID TBSD-2016-0690. **(Submitted for publication)**
2. Umar Ndagi, Ndumiso N Mhlongo and Mahmoud E Soliman (2016). The impact of Thr91 mutation on c-Src resistance to UM-164: Molecular dynamics study revealed new opportunity for drug design. *Molecular Biosystems* MB-ART-12-2016-000848 **(Submitted for publication)**

3 Conferences

Submitted an abstract titled “The impact of Thr91 mutation on c-Src resistance to UM-164: Molecular dynamics study revealed new opportunity for drug design” for an oral presentation at 6th Pharmaceutical Sciences World Congress scheduled for 21-24 May,2017 in Stock holm Sweden.

Acknowledgement

The grace of Almighty which has guided me throughout the course of my life

I extend my gratitude to the following individuals and partners

- My supervisor Prof. Mahmoud E Soliman for his constant support
- Dr.Ndumiso N Mhlongo for his guidance throughout the course of my studies
- My colleagues for their cooperation and support
- My family and friends for their support
- CHPC for the computational resources
- IBBU, Niger state, Nigeria and UKZN College of Health Sciences for Financial support
- My late father Adamu Usman Ndagi for his endless love
- My mother for her consistent prayers and words of advise
- My Spouse Fatima for her continuous support throughout the course my study
- Rahmat Abubakar is also appreciated for her support

List of figures

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Appendix

All supporting materials are available here

List of abbreviations

γ	Gamma
ADT	Androgen deprivation therapy
C	Carbon
CRPC	Castrated resistance prostate cancer
DNA	Deoxyribonucleic Acid
LBD	Ligand binding domain
MD	Molecular dynamics
MM	Molecular mechanics
MM/GBSA	Molecular Mechanics Poisson Boltzman-Surface Area
PASS	Prediction of activity spectra for biologically active substance
PDB	Protein data bank
SCLC	Small cell lung cancer
TNBC	Triple negative breast cancer
Trx	Thioredoxin reductase

Table of contents

CHAPTER 1

1.1 Background

Cancer constitutes an enormous burden on the society affecting mostly the economically under-developed countries.[1] The population growth and ageing account for the increase in the occurrence of cancer cases.[1] In addition, increased prevalence of established predisposing factors such as smoking, overweight, physical inactivity, dietary habits,[1] industrial revolutions and water pollution are associated with the scourge of cancer. It is a known fact that cancer remain one of the major public health problem worldwide[2] and leading cause of death in the United States of America[3] with 21% and 29% been the estimated new cases of prostate and breast cancer among male and female, respectively.[3]

Similarly, lung and bronchus cancer ranked the highest leading cause of death from cancer among United States citizens in 2016 with estimated 26% and 27% death from female and male, respectively.[3] However, in 2012 there was an upsurge in the number of cancer cases across the world, this according to records from comprehensive surveillance database (GLOBOCAN) about 14.1 million estimated new cases of cancer were recorded out of which 8.2 million deaths occur worldwide.[1,4] This again is complicated by incidences of drug resistance and adverse effects of cytotoxic drugs.

Some of these challenges form the basis of revolutionization of cancer research, hence the need for molecular understanding of proteins involved in pathological condition such as cancer, most of which are the prospective drug target.

In recent years, however, computational approach to the process of drug design and discovery has immensely complemented the experimental discoveries and to certain extents validates its results particularly as it relates to the understanding of complex biological systems. Thus, the computational method is an approach used for rational drug design, it involves the investigation of relationship between chemical structures and biological activity.[5]

In one report, unique strategies was developed to understand the conformational features and ligand binding landscape of experimental dual-kinase inhibitor (UM-164) against a target receptor (c-Src).[6] Similarly, in another report, conformational features and ligand binding landscape of orphan nuclear receptor (ROR-γ) [7] in complex with hybrid of compound (XY018) was examined, detailed of conformational features and ligand binding landscape of these receptors and their known inhibitors were evaluated. The findings of this studies can provide

important insights that will assist in further design of novel inhibitors. The unique strategies employed include molecular docking, molecular dynamics (MD) and predictive toxicity and biological testing.

Although an array of therapeutic targets has been discovered in cancer cells,[8] therapeutic effects of approved drugs (chemotherapeutic agents) on some of these targets is hampered mainly by mutation[9] and metabolic changes in the cells[10] resulting to resistance to chemotherapy. However, cancer chemotherapy is the mainstay of the treatment of various stages of cancer.[11] Therefore, resistance to chemotherapeutic agents results in therapeutic failure and eventually death.[11] In a related development, host related factor such as pharmacokinetic resistance[11] also contribute to chemotherapeutic resistance, in this case, alteration in pharmacokinetic parameters such as absorption, distribution, metabolism and elimination (ADME) decrease the bioavailability of oral drugs consequently reduce both time and amount of cancer tissues exposed to these drugs.

Sequel to this, molecular understanding of target proteins is the current focus of cancer research, it revealed details of protein conformation, therefore form the baseline for rational drug design.[12] Herein, effects of mutation on the dynamics of c-Src and ligand binding were evaluated to provide deeper insight into the effect of the mutation on the dynamics of c-Src-UM-164 complex and adding new dimensions to experimental work that has been previously conducted.

A recent study placed UM-164 at an advantage over the current drugs that are clinically used in the treatment of triple negative breast cancer (TNBC).[6] This is mainly due to its anti-proliferative mechanism, by binding to the DFG-out inactive conformation of its target kinases.[6] Resistance to c-Src has been well investigated, in one of the studies, the resistance mutations were said to be at the alarming rate and often limit the success of inhibitors used in new targeted cancer therapies.[13] Mutation in c-Src commonly occur at gatekeeper position in the hinged region[13] in which small amino acid side chain threonine (Thr) is exchanged for a larger hydrophobic residue isoleucine or methionine (Ile or Met),[13] this single point mutation induces conformational changes in the receptor causing resistance to chemotherapeutic agents.

Therapeutic potential of metal complexes in cancer therapy has attracted a lot of interest mainly because metals exhibit unique characteristics, such as redox activity, variable coordination modes and reactivity towards the organic substrate.[14] These properties become an attractive probe in the design of metal complexes[14] that will selectively bind to the biomolecular target

with resultant alteration in cellular mechanism of proliferation. Several metal-based compounds have been synthesized with promising anticancer properties,[14] some of which are already in use in clinical practice for diagnosis and treatment while some are undergoing clinical trials.[14]

1.2 Rationale for this study

Comprehensive knowledge of the structure and functions of retinoic acid related orphan nuclear receptor alpha (ROR-γ) is crucial to the development of its potential and effective inhibitors. Here, ROR-γ is examined to better understand the conformational features and its ligand binding landscape with known inhibitor (XY018). This receptor has been implicated in castrated resistance prostate cancer (CRPC) and has remained orphan receptor because of unidentifiable inhibitor.

Recently, ROR-γ inhibitors were identified and their therapeutic potentials were evaluated, inhibitor such as XY018 was found to be most successful in the inhibitory profile. Since ROR-γ is experimentally well characterised, molecular understanding of its conformational features and ligand binding landscape of its complex is one aspect that warrants deeper investigation. X-ray crystallography does not give adequate information on the general mobility of residues. Therefore, it is difficult to infer the precise dynamics features of the enzyme structure. In line with this, molecular dynamic simulations provides a robust tool to explore the conformational landscape of a biological system.

Before now, there not been any conformational study on XY018-ROR-γ complex. This is first of its kind. Thus, we seek to simulate XY018-ROR-γ complex using molecular dynamics in order to provide an atomistic insight into experimental work that has already been done.

In this work, we firstly conducted MD simulations to accurately probe the dynamics of XY018-ROR-γ complex. Secondly, we examined the impact of binding of XY018 on the dynamics of ROR-γ by using numerous post-dynamics tools such as principal component analysis (PCA), root mean square fluctuation (RMSF), the radius of gyration (RoG), binding free energy calculation, per-residue decomposition analysis and ligand-residue network profile. Thirdly, we incorporated predictive toxicity and biological testing to the analysis to predict possible toxicity and other biological activities that may be inherent in the XY018. Findings from this study should provide an invaluable contribution to the understanding of dynamics of ROR-γ in complex with XY018 which could form a baseline for the design of new potential ROR-γ inhibitors.

Similarly, the role of c-Src in triple negative breast cancer is experimentally well-studied, thus, understanding of conformational features of c-Src in complex with dual kinase inhibitor (UM-164) would provide an atomistic insight to the dynamics of c-Src-UM-164 complex and adding new dimensions to experimental work that has been previously conducted. At present, no conformational study has been done on this complex. Herein, we examined the conformational features of c-Src-UM-164 to provide deeper insight into the dynamics of c-Src-UM-164 complex and adding new dimensions to experimental work that has been previously conducted. We also studied the impact of the mutation on the dynamics of c-Src and ligand binding. Therefore, MD simulations, principal component analysis (PCA), dynamic cross correlation (DCCR) and residue interaction network (RIN) analysis were used in a report herein to provide the molecular understanding of the impact of the mutation on the binding of UM-164. Finding from this study would provide deeper insight into the effect of the mutation on the dynamics of c-Src-UM-164 complex and adding new dimensions to experimental work that has been previously conducted. Therefore, this study can provide important insights that will assist in the further design novel dual kinase inhibitor to eliminate the chances of drug resistance in triple negative breast cancer.

Metal-based compounds synthesised recently are products of drug design targeted at achieving specific objectives that otherwise could not be achieved by original compounds, due to structural modifications that enhanced their pharmacodynamics and pharmacokinetics properties. To this end, a comprehensive review on "metal complexes in cancer therapy an-update from drug design perspective" is included herein to provide an overview of previous reviews on the cytotoxic effects of metal-based compounds while focusing more on newly designed metal-based compounds and their cytotoxic effect on the cancer cell line, as well as new approaches to metal-based drug design in cancer therapy.

1.3 Aims and objectives

This study has four major aims:

1. To investigate the conformational features and ligand binding landscape of ROR-γ complex in order to accurately probe into the protein dynamics and the impact of ligand (XY018) binding on the protein dynamics. To accomplish this, the following objectives were outlined:

1.1 To accurately determine the best-docked ligand and conformation using molecular docking.

1.2 To validate the docking result using MD simulations in order to guarantee that the docked complex remain stable in the active site of the protein within a specific time scale.

1.3 To perform MD simulations of XY018-ROR-γ as well as HC9-ROR-γ (ROR-γ complex with natural ligand) in order to compare the conformational features and determine how well is XY018 is best fitted to the active site of the receptor

1.4 To employ the (molecular mechanics) MM/GBSA method and per-residue energy decomposition to quantitatively estimating the energy contributions and contributions of each active site residue towards the ligand binding.

1.5 To determine the atomic behaviour of the proteins during simulations using PCA.

1.6 To compute ligand residue interactions on post dynamic structure in order to establish the nature of the reaction and residues involve in the reaction.

1.7 To employ predictive toxicity and biological testing to determine the inherent toxicity and biological activity of the XY018.

2. To investigate the conformational features and ligand binding landscape of c-Src-UM-164 complex in order to accurately probe into the protein dynamics and the impact of ligand (UM-164) binding on the protein dynamics. To accomplish this, the following objectives were outlined:

2.1 To perform MD simulations of Src-UM16, apo and Src-Dasatinib in order to determine the protein dynamics.

2.2 To employ the MM/GBSA method and per-residue energy decomposition to quantitatively estimating the energy contributions and contributions of each active site residue towards the ligand binding.

2.3 To determine the atomic behaviour of the proteins during simulations using PCA.

2.4 To compute ligand residue interactions on post dynamic structure to establish the nature of the reaction and residues involve in the reaction.

2.5 To monitor the hydrogen bond formation during the course of MD simulations.

2.6 To further examine the conformational changes of receptor upon ligand binding.

2.7 To employ predictive toxicity and biological testing to determine the toxicity of UM-164.

3. To provide the molecular understanding of the impact of the mutation on UM-164 binding to c-Src. To accomplish this, the following objectives were outlined:

3.1 To perform MD simulations of wild type, apo and mutant in order to determine the protein dynamics.

3.2 To provide a deeper understanding of drug resistance mechanism of mutant protein to UM-164 using MM/GBSA, PCA, dynamic cross correlation analysis (DCCs) and residue interaction network (RIN).

4. To provide a comprehensive review on the new approach to metal-based drug design.

To accomplish this, the following objectives were outlined:

4.1 We provide an overview of previous reviews of the cytotoxic effects of metal-based compounds while focusing more on newly designed metal-based compounds and their cytotoxic effect on the cancer cell line.

1.4 Overview of this work

This work is divided into 8 chapters, with this one included:

Chapter 2: This chapter gives a general background on cancer, including global statistic on the scourge of cancer. Different types of cancers with a particular emphasis on prostate cancer and breast cancer. We also highlighted on drug resistance mechanism and new approaches to cancer drug design.

Chapter 3: This chapter briefly introduces computation chemistry, various molecular modeling and simulation techniques including their applications. Attempts are made here to explain theoretically some computational methods with highlighted multiple computational tools applied in the study of XY018-ROR-γ and c-Src-UM-164 with main the focus on molecular dynamic simulations and binding free energy calculation.

Chapter 4 (accepted)

This chapter presents results from the study titled “Re-emergence of an Orphan Therapeutic Target for the Treatment of Resistant Prostate Cancer – A thorough Conformational and Binding Analysis for ROR-γ Protein”.

Chapter 5 (submitted for publication)

This chapter presents results from the study titled “Emergence of a Promising Lead Compound in the Treatment of Triple Negative Breast Cancer: An Insight into Conformational Features and Ligand Binding Landscape of c-Src Protein with UM-164”.

Chapter 6 (submitted for publication)

This chapter presents results from the study titled “The impact of Thr91 mutation on c-Src resistance to UM-164: Molecular dynamics study revealed new opportunity for drug design”

Chapter 7 (Published)

This chapter presents a review article titled “Metal Complexes in Cancer Therapy – an update from Drug Design Perspective”.

Chapter 8:

This chapter is overall the conclusion of entire thesis and future work.

References

(1) Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-tieulent, J., and Jemal, A. (2015) Global Cancer Statistics, 2012. *CA a cancer J. Clin. 65*, 87–108.

(2) Ma, X., and Yu, H. (2006) Global burden of cancer. *Yale J. Biol. Med. 79*, 85–94.

(3) Siegel, R. L., Miller, K. D., and Jemal, A. (2016) Cancer statistics, 2016. *CA. Cancer J. Clin. 66*, 7–30.

(4) Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D. M., Forman, D., and Bray, F. (2015) Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer 136*, E359–E386.

(5) Finn, P. W., Finn, P. W., and Kavraki, L. E. (1999) Computational Approaches to Drug Design. *Agorithmic 25*, 347–371.

(6) Gilani, R. A., Phadke, S., Bao, L. W., Lachacz, E. J., Dziubinski, M. L., Brandvold, K. R., Steffey, M. E., Kwarcinski, F. E., Graveel, C. R., Kidwell, K. M., Merajver, S. D., and Soellner, M. B. (2016) UM-164: A Potent c-Src/p38 Kinase Inhibitor with In Vivo Activity against Triple-Negative Breast Cancer. *Clin. Cancer Res. 22*, 5087–5096.

(7) Wang, J., Zou, J. X., Xue, X., Cai, D., Zhang, Y., Duan, Z., Xiang, Q., Yang, J. C., Louie, M. C., Borowsky, A. D., Gao, A. C., Evans, C. P., Lam, K. S., Xu, J., Kung, H.-J., Evans, R. M., Xu, Y., and Chen, H.-W. (2016) ROR-γ drives androgen receptor expression and represents a therapeutic target in castration-resistant prostate cancer. *Nat. Med. 22*, 488–496.

(8) Guo, S., Zou, J., and Wang, G. (2013) Advances in the proteomic discovery of novel therapeutic targets in cancer. *Drug Des. Devel. Ther. 7*, 1259–71.

(9) Housman, G., Byler, S., Heerboth, S., Lapinska, K., Longacre, M., Snyder, N., and Sarkar, S. (2014) Drug resistance in cancer: an overview. *Cancers (Basel). 6*, 1769–92.

(10) Lukey, M. J., Wilson, K. F., and Cerione, R. A. (2013) Therapeutic strategies impacting cancer cell glutamine metabolism. *Future Med. Chem. 5*, 1685–700.

(11) Alfarouk, K. O., Stock, C.-M., Taylor, S., Walsh, M., Muddathir, A. K., Verduzco, D., Bashir, A. H. H., Mohammed, O. Y., Elhassan, G. O., Harguindey, S., Reshkin, S. J., Ibrahim, M. E., and Rauch, C. (2015) Resistance to cancer chemotherapy: failure in drug response from ADME to P-gp. *Cancer Cell Int. 15*, 71.

(12) Hoelder, S., Clarke, P. A., and Workman, P. (2012) Discovery of small molecule cancer drugs: Successes, challenges and opportunities. *Mol. Oncol. 6*, 155–176.

(13) Getlik, M., Grütter, C., Simard, J. R., Klüter, S., Rabiller, M., Rode, H. B., Robubi, A., and Rauh, D. (2009) Hybrid compound design to overcome the gatekeeper T338M mutation in cSrc. *J. Med. Chem.*

(14) Frezza, M., Hindo, S., Chen, D., Davenport, A., Schmitt, S., Tomco, D., and Dou, Q. P. (2010) Novel metals and metal complexes as platforms for cancer therapy. *Curr Pharm Des 16*, 1813–1825.

CHAPTER 2

Cancer

2.1 Introduction

Cancer is as a progressive abnormal proliferation of any of the different kind of cells in the body.[1] However, recent studies have revealed that cancers are associated with abnormal changes in the body metabolic state,[1, 2] this abnormality is linked to changes in the structure and function of mitochondrial,[2] with genetic implication as the primary focus. The propensity to spread and invade (metastasis) related body tissues is the primary cause complications and death in most cancer patients. Usually, un restricted cell growth marks the termination of biological process of the disease.[2] It is a well-known fact that cancers are highly heterogeneous these characteristic is driven by two principal factors namely; the genetic alteration and phenotypic selection resulting from evolution.[3] This are the major reasons for the diverse forms of cancer which at present more than 200 types exist,[4] while the number of validated cancer target remained outside the reach of pharmacological regulation.[5] They are either undruggable or druggable,[5] however, variations in interactive amino acids (molecular network) during cancer development necessitate different drug targeting strategy in the early and late phase of carcinogenesis.[6]

2.2 Current status of global cancer burden

Cancer constitutes an enamours burden on society[7] affecting more the economically less developed countries.[7] The incidences of cancer are relatively high because of growth and ageing of the population,[7] as well as an increase in the prevalence of predisposing factors such as smoking, overweight, physical inactivity, change in reproductive patterns associated with urbanisation and economic development or industrial revolution.[7] In 2012, based on GLOBOCAN (comprehensive surveillance database managed by an international association of cancer registries) estimate, about 14.9 million new cancer cases and 8.2 million cancer associated deaths were recorded worldwide.[7] However, the burden has shifted more to economically less developed countries which currently account for 57% cases and 65% of cancer-associated deaths worldwide.[7,8] In male, five most common sites of cancer diagnosed in 2012 were lungs, prostate, colorectal, stomach and liver, among which lung cancer is the leading cause of cancer death in both developed and economically less developed countries.[7] Breast cancer is the leading cause of cancer death among females in less developed countries[7,

[9] compare to developed countries.[10,11] Other forms of cancer with high mortality rate in developed countries include colorectal cancer among males and female and prostate cancer among males. Similarly, the incidence of cancers combined are nearly twice as high in developed countries than less developed countries in both males and females, however, mortality rates are only 8% to 15% higher in more developed countries.[7]

In United State of America, based on the statistical data from the National Center for Health Statistics, about 1,685,210 new cancer cases and 595,690 cancer-associated deaths were projected to occur in 2016.[11] There is relatively stable trend of incidences of cancer in women but declining by 3.1% per year in men (from 2009-2012),[11] as a result of recent rapid decline in the diagnosis of prostate cancer.[11] Similarly, cancer associated death has dropped by 23% since 1991, translating to about 1.7 million death averted through 2012.[11] However, death rates are relatively increasing for cancers of liver, pancreas and uterine corpus, therefore, cancer is now leading cause of death in 21 state in United states of America.[11]

2.3 The quest for better understanding of cancer

2.3.1 Molecular networks in cancer

The protein-protein interaction network provides general information about cellular functions and biological process.[12] The cellular arrangement of proteins is such that some act as a hub, highly connected to the others, while some have few interactions.[12] The dysfunction of some of these interaction results in many pathological conditions such as cancer. Therefore, a good understanding of network properties of cancer-related proteins will help to explain their role in carcinogenesis.[12] Cancer is generally immortal and is said to be the product of complex cell transformation process.[6] This process usually began with mutation or DNA-rearrangements,[6] which disrupts the original composition of phenotypes in the cell. This results in cell population with variable chromatin organisation, gene expression patterns leading to formation of interatomic composition.[6] These diverse cell population with variable characteristics such as increased level of stochastic processes (noise) phenotypic plasticity[13, 14] and by an increase in network entropy of protein-protein interaction.

The developmental process of cancer is continuously described as system level, network phenomenon,[15, 16] where the networks easily adapt to the changes in parameters (network evolution) such as network connectivity, edge weights, diameter centrality and motifs or modules.[17, 18] The identification of these networks changes have predictive potential in retrospect and about future development of complex system represented by the network.[6] Therefore, assessment of molecular network changes in the progression of cancer requires a

good understanding of time-frames both in clinical sampling and network analysis.[6] The autonomy of cancer cells is based mainly on the changes in its signalling network,[6] formed by interconnected signalling pathway and gene regulatory network.[19] Thus, changes in expression level of signalling proteins in cancer cells may result in activation of cancer-related pathways, this may lead to rewiring of the whole signal network as showed in the case of human epidermal growth factor.[20] **Figure 1** depicts the appropriate drug targeting strategies**.**

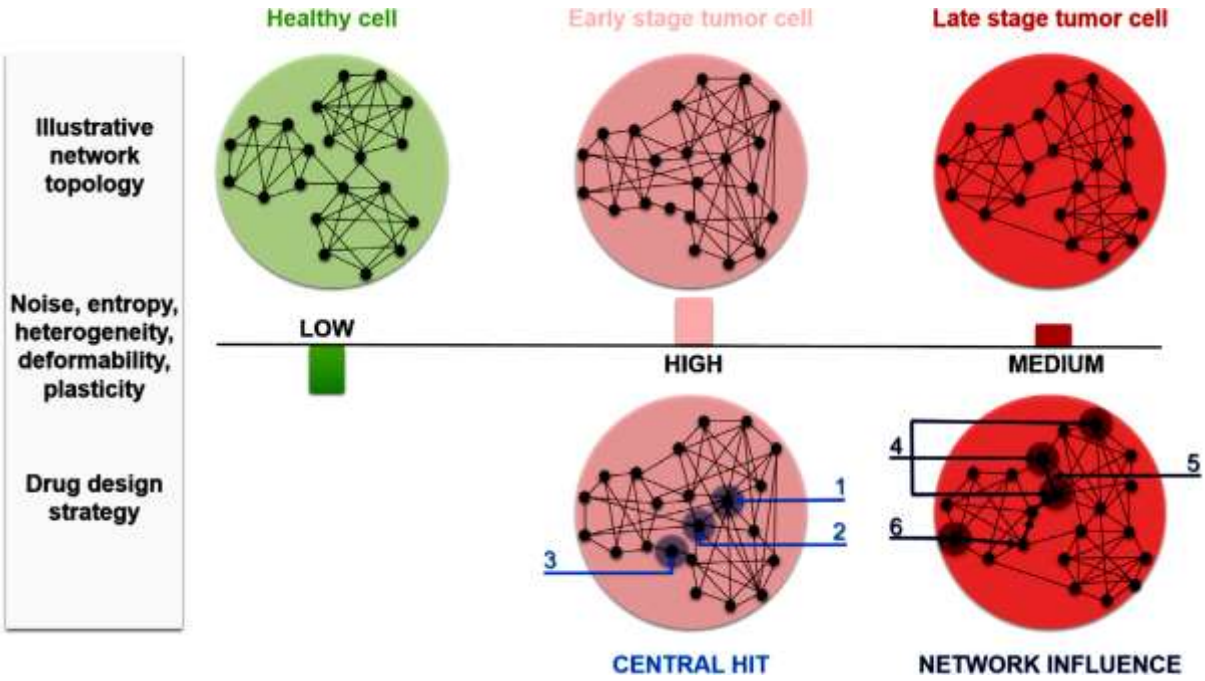


Figure 1. System level development of cancer that determines the most appropriate drug targeting strategy adopted from Dávid *et al.* [6]

It worth noting that diverse cancer targets have been identified, however, developing drugs that will act on these targets have been quite challenging. Changes in residue networks have been identified as the major contributing factor, this leads to the development of dual network strategy, the central hit strategy and network influence strategy to target various disease including cancer.[6] In line with this, molecular understanding of cancer is paramount to the successful design of anticancer drug.

2.3.2 Cancer as a metabolic disease

The initial perception of cancer as a genetic disease is gradually eroding,[21, 22] this may be attributed to Warburg discovery which states that, cancer is actually caused by a defect in the cellular metabolic energy, that is primarily related to the function of the mitochondria.[22] This understanding has changed the view and perception of scientist about the real cause of cancer.

In the light of foregoing, study have shown that cancer is primarily a disease of metabolic disorder associated with energy production from respiration and fermentation.[21, 22] These disturbances is related to abnormalities in structure and function of mitochondria[22], [23] as seen in **Figure 2** It is assumed the growth and progression can be checked adequate monitoring as the body transit from fermentable metabolites to respiratory metabolites[23] considering the evolutional theory of Charles Darwin.[23] This is only true because genetic variation that exist among individuals, therefore individualising metabolic therapy as a baseline in cancer treatment will require individualising the therapy base on individual body physiological state.[23]

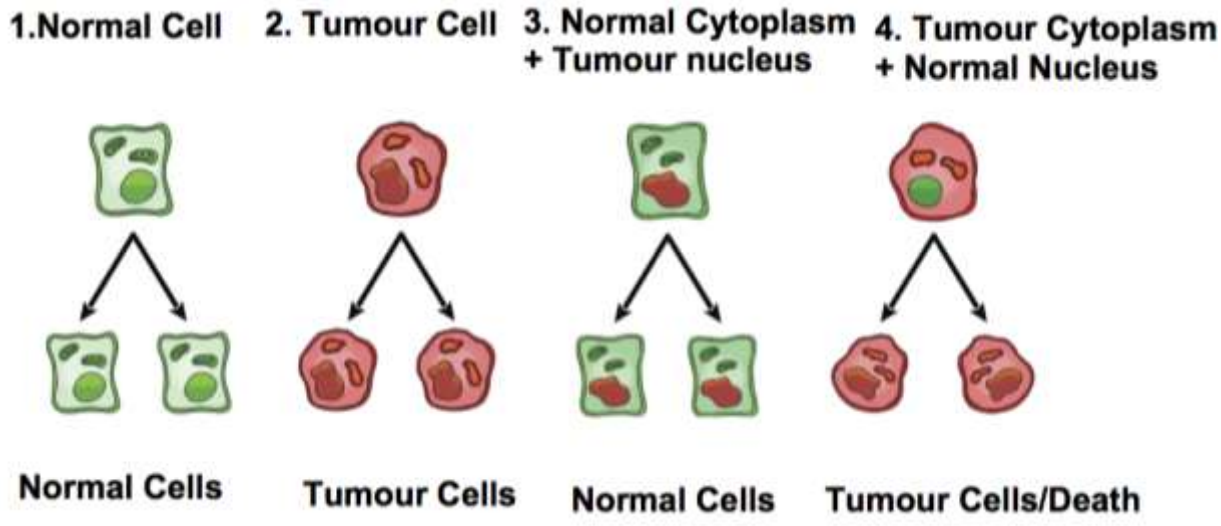


Figure 2. Role of nucleus and mitochondria in the origin of tumour, normal cell depicted in green with mitochondrial and nuclear morphology indicative of normal respiratory and gene expression, adopted from Thomas *et al.* [23]

Placing cancer as single type disease rather than a complex form will enhance the individualisation of therapy.[23, 24, 25] Many scientists have advanced their reasons to consider cancer as a metabolic disorder, this was better supported with variations in the nuclear gene theory in relation to somatic cells.[23] It revealed that tumour growth is suppressed when the cytoplasm from non-tumour cells with functional mitochondria interacts with nuclei from tumour cells [26] as shown in **Figure 2** . Based on this findings, it can be concluded that tumour cell grow depends on the availability of functional mitochondrial rather than specific nuclear mutation. However, there is enhancement in cell growth when nuclei of none-tumorigenic cells are allowed to interacts with with cytoplasm from tumour cells.[27,28]

In Warburg theory, the key points to note are: (1) Decrease in respiration mimics tumorigenesis consequently cancer.[29] (2) There is progressive compensation of energy from respiration by energy from glycolysis due to decrease in respiratory energy.[29, 30] (3) Continue degradation of

lactate in the presence of oxygen.[29, 30] (4) Deficiency in respiration ultimately becomes irreversible[31] as shown in **Figure 3** Therefore applying Warburg theory to cancer management, one will think that the rational strategy for cancer management is the principal focus of the drug that selectively target tumour metabolic pathway.[23] This strategy is applicable to the most cancers irrespective of tissue origin, since they all exhibit similar metabolic features, restricting glucose or targeting glucose is fundamental to the successful management of cancer. [23] This is because glucose is the major source of energy in tumour metabolism resulting from fermentation.[23] Similarly, normal cells in body survive through glucose from glycolytic pathway. It is, therefore, important that normal cells are protected from therapies that disrupt glycolytic pathway.[23]

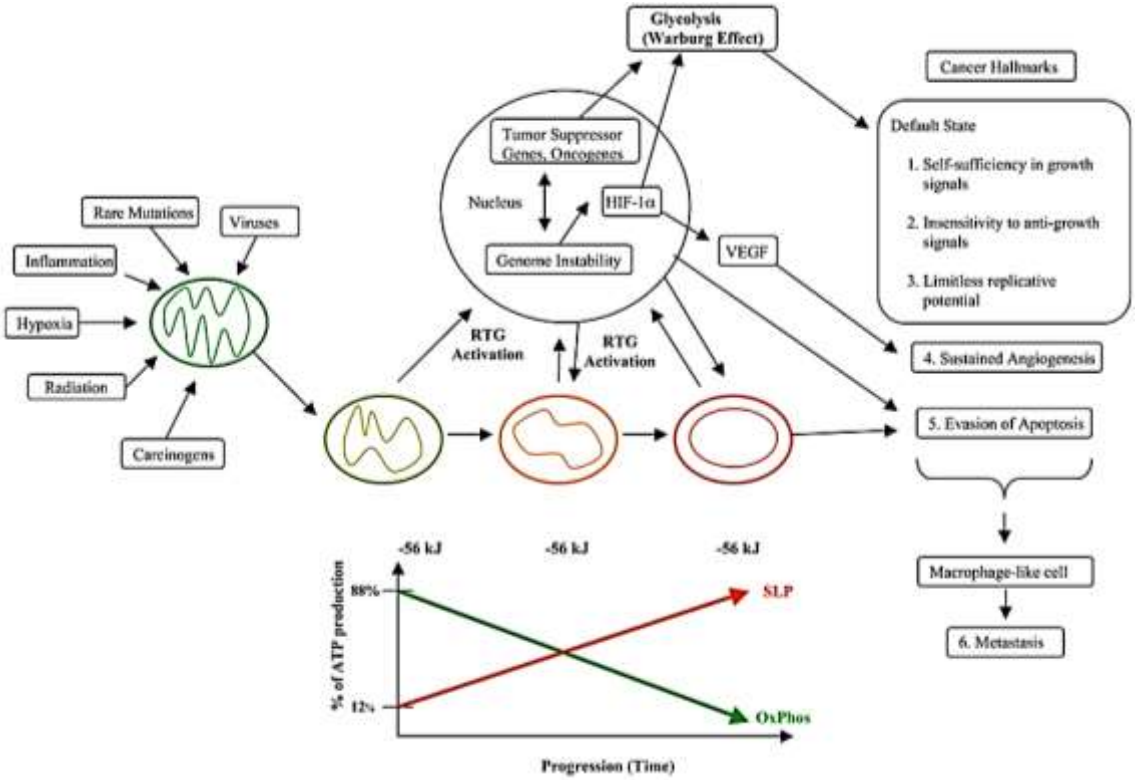


Figure 3. Linking the hall marks of cancer to impaired energy metabolism adopted from Thomas *et al.*[2]

2.3.3 Implication of metabolism for novel therapeutics

In human system, the presence of glucose and ketones within the prescribed normal range under fasting condition that produces anti-angiogenic, anti-inflammatory and pro-apoptotic effects,[23] is termed metabolic management stage.[32] The human system when assumed this stage, low dose of a variety of drugs can be used to target energy metabolism in any surviving tumour cells.[33] Cytotoxic drugs such as Imatinib and trastuzumab used in the management of BCR-ABL leukaemia cells and ErbB2-positive breast cancers derived therapeutic success due to

their selective target properties on the glucose signalling pathway.[34, 35] In a related development, calorie restricted ketogenic diets (KDs) will target any cancer regardless of any form of mutation involved.[36, 37, 38] It is also been speculated that, dietary management another means targeting metabolic pathway with minimal cells toxicity.[38] It is important to note that, KD is therapeutically effective on restrictive intake, it predispose individual to insulin insensitivity when taking in un-restricted form.[39] In related development, targeting glutamine will result in inhibition of systemic metastatic cancer cells,[40, 41] because of presence of multiple cellular macrophages,[42 , 43] glutamine is the major source of energy for macrophage and cells of immune system.[44] However, glutamine said to be possess chemo-preventive effect, therefore further research is required to establish the role of glutamine in this regard. A study has shown that synergistic interactions exist between the KD and hyperbaric oxygen therapy (HBO_2T).[23] The presence of KD lead to decrease glucose level necessary for glycolysis and also reduced NADPH levels for anti-oxidant potential through the pentose-phosphate pathway.[23] The increase in reactive oxygen species (ROS) in tumour cells is associated with the presence of HBO_2T, while ketones have a protective effect on normal cells against ROS damage and potential central nervous system (CNS) oxygen toxicity.[45, 46] The inability to use ketones and the dependency on glucose for energy by the cancer cells make them selectively vulnerable to this therapy.[23] However, this therapy is only effective against those tumour cells with mitochondria, it is unclear at present whether this therapy will be effective on those tumour cells with few or no mitochondria.[47] Therefore, metabolism in cancer cells is a promising research field that required in-depth molecular understanding in our quest to drive cancer from being a terminal illness to a treatable condition.

2.3.4 Mutation in cancer

The majority of the human tumour are highly heterogeneous, this heterogeneity results from a mutator phenotype.[48] This rely on premise that normal mutation rate is insufficient to account for multiple mutations found in human cancer,[48] this mean that cancer cells must have exhibited a mutator phenotype early during their initial phase of growth.[48] The development of cancer is a gradual process particular in human solid tumours, there is usually an interval of 20 years from carcinogenic exposure to clinical detection.[48] During this period, cancer cells actively divide, invade and metastasize.[48] The proposal from mutator phenotype states that, phenotypes result from mutations in genes that can maintain genetic stability in normal cells. Thus, mutations in genetic stability genes can cause mutation in other genes that govern genetic stability,[48] thereby initiating a cascade of mutations throughout the genome. In the midst of

this, some of the mutant will confer a selective advantage(s), allowing the mutated cells to expand and achieve clonal dominance.[48] Mutations are known to be inheritable changes in the nucleotide sequence of DNA which include chromosomal abnormalities. However, new evidence emerged challenging the belief that multiple mutations in human cancer were based mainly on chromosomal aberrations, [48,49] in this case, multiple abnormal chromosomes are found in most solid tumours and are rarely diagnostic of tumour type nor indicative of prognosis.[48] Although there is an exception one of which involve translocations and rearrangement of specific chromosomes that are instrumental in the diagnosis of several hematologic malignancies and sarcomas.[48] In most tumours, there are different regions that exhibits abnormalities in comparative genomic hybridization (CGH),[50] a technique that measure changes in DNA copy number [48, 50] and spectral karyotyping (SKY).[51] Similarly, tumours are said to display amplification of a segment of DNA at high frequencies.[52] They also exhibit loss of heterozygosity as the consequence of deletions in one of the parental chromosomes.[53] In single cancer cells, there are thousands of mutations suggesting that in the human tumour at the time of diagnosis harboured billions of different mutations[48]. Mutation in most gene and regulatory sequence are presence in one or more cell within a tumour.[48] Therefore, when the cells are exposed to chemotherapeutic agents, tumour cells with acquired mutation that render them resistance will preferentially proliferate and eventually increase the population of the resistant strain.[48] This account for the relative ease to which cancer cells acquire resistance to drugs.[48, 54]

2.3.5 Drug resistance in cancer

Chemotherapy remains a mainstay in the clinical management of cancer,[55] this is however faced with a lot of challenges ranging from patient drug intolerance to various forms of drug resistance. Drug resistance in cancer is multifactorial, it could be: (1) Pharmacodynamics drug and (2) Pharmacokinetics drug resistance.

2.3.5.1 Pharmacodynamics drug resistance

In pharmacodynamics resistance, changes in the target sites (receptor) are usually the major challenges. The receptors are modified as result of age[56] or genetic mutations.[57, 58] The re-modification of the receptor or target site affects drug binding. The orientation of drug in the receptor conformational space is absolutely important to bring about the desired pharmacologic effect.[59] Aging is usually accompanied by a decrease in functionality of vital organs of the body[60] resulting in changes in quantity and type of amino acids present in these tissues, this

affects the amount of receptors available for the binding and receptor conformation space required for the drug to bind and elicit it pharmacological effect hence the variation in clinical response to chemotherapeutic agents.

On the other hand, mutational effect premise on the effect of genetic variation on patient drug response.[57, 61] Mutation affects the drug binding in two ways: (1) By altering the molecular interaction between the active site residues and drug atoms.[57, 61] (2) By altering protein conformations. These changes resulting from mutation are sometimes referred to as intrinsic target resistance and are believed to be the cause of exaggerated drug effect. Resistant to chemotherapeutic agents can be acquired during treatment of a tumour that was initially sensitive and suddenly develop resistant due to mutations arising from drug treatment.[62] As the knowledge on the molecular biology of cancer advanced, cancer drug design has shifted toward agents that target specific molecular alteration in cancer.[62]

2.3.5.2 Pharmacokinetics drug resistance

In pharmacokinetic drug resistance, exposure of cancer to the drug molecule is highly limited. This is usually associated with the various mechanism of drug absorption, distribution (transport) elimination (metabolism and excretion).

Absorption of orally administered chemotherapeutic agent goes through normal route like any other oral formulation. However, in the presence of permeability glycol-protein (p-gp) also known as multi-drug resistance protein (MDR) found in the gastrointestinal tract (GIT) [63,64] and small intestines (primary site for epithelial absorption).[65] Co-administration of chemotherapeutic agents leads to over-expression of p-gp thereby decreasing the bioavailability of some of these agents including imatinib.[66] Variation among individual eg genetic polymorphism and change in the body physiological state also play important role in the expression of p-gp in GIT.[67,68] Thus, causing erratic absorption of some chemotherapeutic agents including paclitaxel.[66] The presence of food also reduces the GI viscosity thereby reducing the rate and the extent of disintegration, dissolution and absorption of drugs from GIT. Careful monitoring of drug-food interaction is important in administering chemotherapeutics to avoid unwanted effect and achieve sufficient bioavailability.[69, 70, 71]

***Distribution*:** Drugs are transported via the plasma to the body tissues, the volume of distribution (V_d) which is a hypothetical volume gives a summary of drug distribution into the tissues,[72, 73] increase in volume V_d mean that more of the drug is distributed into the tissue at a relatively lower concentration. Many factors determine drug distribution in the body some which include: Body weight, plasma protein and circadian rhythm.[63]

Body weight is an important factor that determines drug dosing and dosage adjustment, cancer patients often lose weight as the tumour progresses [74,75] therefore, continuous dose adjustment based on current body weight is absolutely important in chemotherapy.

Plasma proteins are mainly responsible for drug distribution in the body. Within the blood, drug exists either as free or are bound to plasma proteins[63] and albumin and alpha 1-glycoprotein are the major binding proteins. Albumins are carriers of acidic drugs, while alpha 1-glycoprotein (AAG) carries the basic or neutral lipophilic drugs.[76, 77, 78] The AAG is an important carrier of an anticancer drug, therefore variation in the plasma concentration of AAG affects the activity of several anticancer drugs including imatinib.[79] Variation in plasma AAG also affects the anticancer activity of Gefitinib[80] hence institution of therapeutic drug monitoring is absolutely important in chemotherapy.

Circadian rhythm, the cardiac timing system is made up peripheral oscillators found in the most tissue of the body and central pacemaker located in the suprachiasmatic nucleus of the hypothalamus.[81] The circadian rhythm has been implicated in pathophysiologies of several diseases,[82] drug action[83,84] and drug disposition (pharmacokinetics).[85] The concentration of plasma protein varies with time, minimum concentration reached around 4:00am and at about 8:00am the concentration begin to increase.[86] Age variation also affects the cardiac rhythm, the rhythm can be masked in the younger ages,[87] and become aggravated with age. Therefore, adequate and proper dosage timing is very important in chemotherapy for optimum plasma concentration around the tumour sites.

Elimination: Drugs are eliminated from the body mainly through metabolism and excretion.

Drug metabolism is a biological strategy of drug detoxification which involves the conversion of drugs from active to inactive form through several processes. Similarly, metabolism can also reflect the conversion of drugs from inactive to the active form, particularly in pro-drugs. Essentially, enzymes are the major players of metabolism. However, some of these enzymes can be induced to increase the metabolic activity of decrease the metabolic activity. The increase in metabolic activity result in reduction of plasma half-life of the drugs thereby exposing the cancer cells to a sub-optimal concentration which eventually lead to drug resistance. While the decrease in metabolism prolongs the plasma half-life of the drugs which expose the patients to the severe adverse effect of anticancer drugs. This is more reason why the therapeutic drug monitoring is recommended for patients with terminal illnesses such cancer.

Excretion: The kidney is the primary organ by which drugs are eliminated from the body, other routes include biliary, skin and the lungs. Variation in renal excretion of drugs could be due to

gender,[88, 89] physiological state of individuals, race and ethnicity.[90] Therefore, changes in the rate of glomeruli filtration affect the elimination half-life of anticancer drugs with a direct consequence of the systemic concentration of the drug. It is important to complete assessment of renal and biliary system of patients on anticancer drugs at an interval within the period of therapy to allow for dosage adjustment when the need arises.

2.4 Potential sites of drug target in cancer

Cancer target sites continue to evolve certainly due to the mutational effect that results in modification and re-modification receptor sites that are either at one time sensitive to the particular anticancer drug or newly discovered receptor site. Currently, there are over 5000 targets in cancer [91] out which only few are druggable, this is despite concerted research effort to develop therapeutics that can bind to these targets and halt cancer progression. The confluence of progress between the cancer research and therapeutic outcome in cancer management result in current research focus in cancer. In line with this, most recent cancer therapies are focusing on the use kinase inhibitors[92] with which deoxyribonucleic acid (DNA) repairs are the possible targets.[93]

2.4.1 DNA repair pathway as a potential target for cancer therapy

Deoxyribonucleic acid (DNA) damage is one of the major impact of chemotherapeutic agents on cancer cells. However, cancer derived a means of survival through DNA repair pathway, therefore inhibitors of specific DNA repair pathway could be very efficacious when used in combination with DNA-damaging chemotherapeutic agents.[94] Similarly, alteration in DNA repair pathway arising during cancer development can make some cancer cells reliant on a reduced set of DNA repair pathway for survival.[94] Similarly, as the cancer cells become more mutagenic, genetic streamlining occur leading to a deficiency in one or more DNA repair pathway accompanying by compensatory activities that increase the levels of proteins responsible for the DNA repairs in the same pathway or in a different one. [94, 95] Therefore, interrupting DNA repair pathway such a way that can shut down the compensatory mechanisms pathway can lead to cell death, this is the goal of the research surrounding the DNA repair pathway.[94] Sequel to this, it is possible to develop an agent that selectively inhibit one of this pathway in such cancers and could prove effective as a single regimen therapy.[94] The DNA damaging agents commonly in use in cancer therapy include:

2.4.1.1 Platinum-based complexes in cancer therapy

Therapeutic potential of metal complexes which include platinum complexes in cancer therapy has attracted a lot of interest mainly because metals exhibit unique characteristics, such as redox activity, variable coordination modes and reactivity towards the organic substrate.[96] These properties become an attractive probe in the design of metal complexes [96] that will selectively bind to the biomolecular target with resultant alteration in cellular mechanism of proliferation. Cytotoxic drugs such as platinum-containing drugs cause DNA damage by binding to N7 reactive centre on purine residue thereby inducing changes in DNA leading to apoptotic cell death.[97] This includes cisplatin, carboplatin, oxaliplatin, satraplatin, ormaplatin and iproplatin to mention but few. The major challenge with the use of these compounds is their dose-limiting toxicities as well as the development of drug resistance.[98] Details on these drugs can be found in published article which forms part of this thesis "Metal complexes in cancer therapy an up-date from drug design perspective. Structure of some platinum complexes used in cancer therapy are represented in **Figure 4**

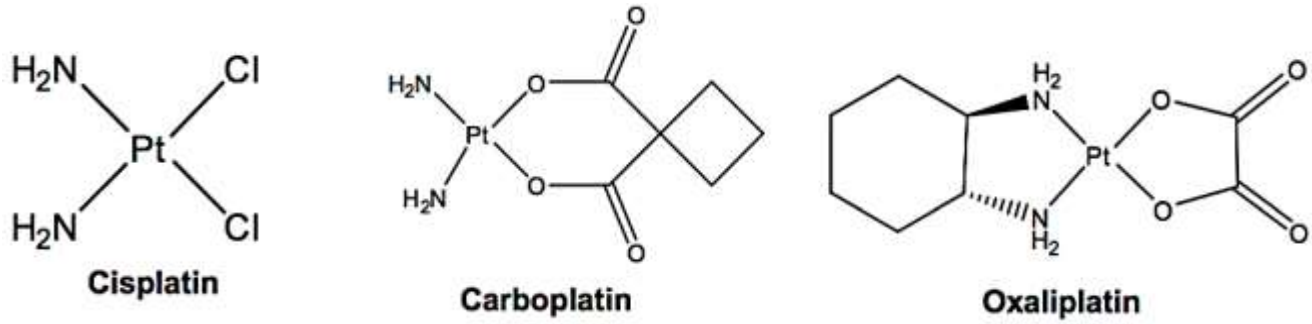


Figure 4. 2D structure of selected Platinum compounds with DNA effect

2.5 Castrated resistance prostate cancer (CRPC)

Prostate cancer (carcinoma of the prostate) is the development of cancer in the prostate gland. It is the fourth most common cause of cancer-associated death [99] and is third most common cause of cancer death in men .[99] Prostate cancer can also develop resistance despite androgen-deprivation therapy (ADT), [100] this is called castrated resistance prostate cancer (CRPC) which is usually associated with continuous rise in the serum prostate specific antigens (PSA).[100] Castrated resistance prostate cancer presents in a different form, ranging from rising PSA level without metastases or any form of symptom despite ADT to metastases with a significant presence of cancer symptoms.[100] Prognosis in CRPC is usually associated with so many factors including performance status, presence of bone pain, extent of disease on bone scan and serum level of alkaline phosphatase.[100] In 90% of men with CRPC bone metastases usually occur with

high chances of producing significant morbidity such as pain, pathological fracture, spinal cord compression and bone marrow failure.[100] Others symptoms include paraneoplastic effects such as anaemia, weight loss, fatigue, hypercoagulability and increased susceptibility to infection.[100]

2.5.1 Treatment of castrated resistance prostate cancer

In castrated resistance prostate cancer (CRPC) majority of patients' diseases condition usually progressed to the stage of resisting anti-androgens therapy.[101] This limit the available treatment options. In the past the prostate cancer was considered unresponsive to the chemotherapy until the advent of palliative treatment mitoxantrone combined with prednisolone in the treatment of CRPC.[101] Docetaxel-based chemotherapy has shown some significant advantage over mitoxantrone in patients with metastatic CRPC.[101] Cabazitaxel has recently shown to improve the survival in patients with metastatic CRPC who progressed after docetaxel-based chemotherapy.[101] It has also been demonstrated that Sipuleucel-T improved overall survival in patients with asymptomatic or mildly symptomatic metastatic CRPC.[101, 102] A lot of research have been carried on different targets and prospective inhibitors in CRPC, some drugs are currently under clinical trial to ascertain their efficacy and safety in the treatment of CRPC. Drugs such as Satraplatin are currently in phase III clinical trial, a combination of Satraplatin and prednisolone results in significant improvement in radiographic progression-free survival (PFS).[103, 104] In a related development, a phase II clinical trial of Sunitinib (tyrosine kinase inhibitor) in patients with chemotherapy-naïve or docetaxel-refractory CRPC shows >50% prostate specific antigen (PSA) reduction in one patient per treatment group.[105] Several drugs are medical domains that are currently used as chemotherapeutic agents in the management of CRPC. However, the progression of the disease condition led to the discovery of retinoic acid-related orphan receptor- γ (ROR-γ) implicated in the CRPC. Receptors such as alpha and beta gamma type constitute family of nuclear receptors that function as ligand-dependent transcription factor.[106]

2.5.2 ROR-γ, a new target in castrated resistance prostate cancer

In recognition of the role of ROR-γ in prostate cancer, recent studies examined the relationship between ROR-γ and prostate cancer.[107] Results from immunoblotting analysis of ROR-γ protein in cancer and noncancerous cell lines revealed that appreciable level of inhibition of growth of androgen-sensitive human prostate adenocarcinoma cells (LNcap) and their CRPC-derivatives C4-2B cells was achieved when ROR-γ was knocked down by different RORC siRNAs.[107] Such a high level of growth inhibition was observed in androgen-sensitive and

CRPC model,[107] as the ROR-γ is knocked down there is resultant induction of apoptosis, which is marked by activation of caspase-3 and caspase-7 and the breakdown of poly (ADP-ribose) polymerase 1(PARP1), resulting in reduced expression of the protein relevant to oncogenesis, proliferation and survival.[107] Most recent studies continued to place ROR-γ as an orphan because of unidentifiable ligand (inhibitor) suitable for its binding site.[108] A most recent study identified three different inhibitors of ROR-γ.[107] These inhibitors were labelled as XY018, SR2211 and XY011 **(Figure 5**) based on their structural differences and effect on the cell line.[107] In this regard, the first two inhibitors demonstrated sufficient level of suppression on the expression of important proliferation and survival proteins.[107] Therefore, the inhibitor XY018 is a promising lead compound in CRPC,[107] **Figure 5** represent the 3D structure of XY018 bound to ROR-γ. Detailed discussion on this topic can be found in one of the chapter of this thesis.

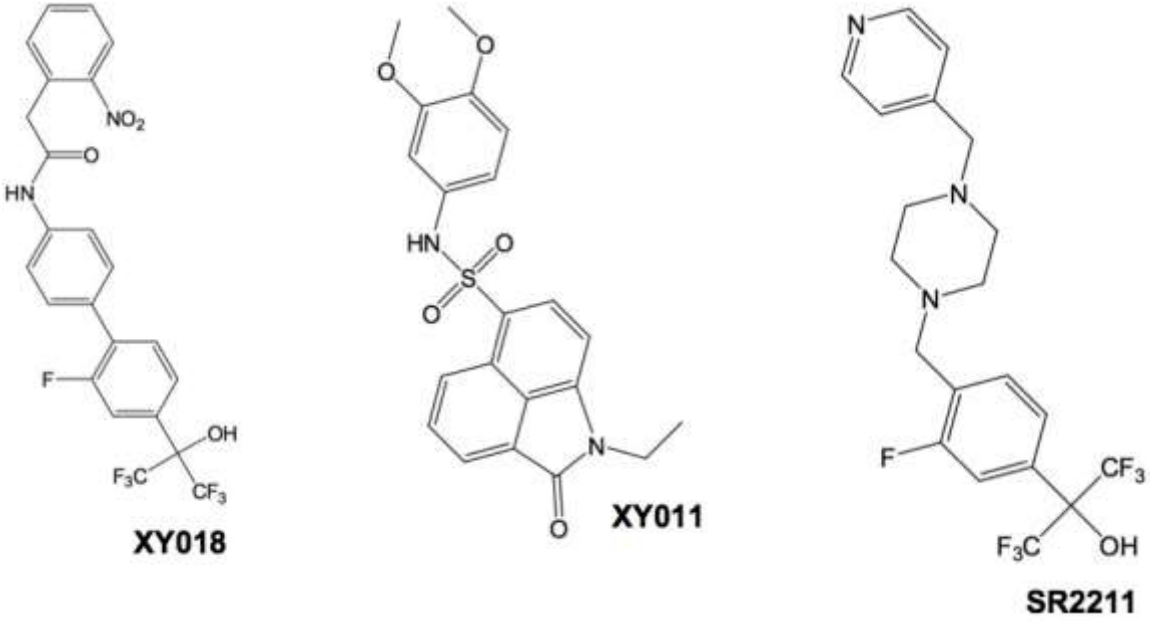


Figure 5. 2D structure of experimental ROR-γ inhibitors

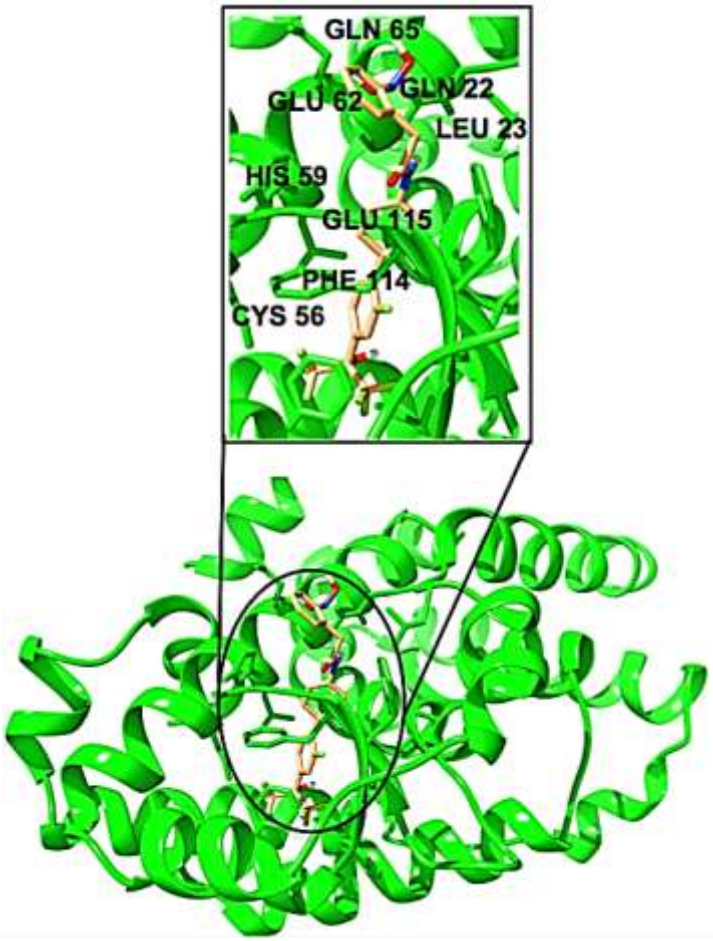


Figure 6. 3D structure of XY018 bound to ROR-γ

2.6 Breast cancer and potential drug targets

Breast cancer remains among the most frequently diagnosed and life-threatening cancer in women,[109] and third leading cause of cancer-associated death among women in the United State of America (US). [11, 110] Genetically, breast cancers (BCs) are heterogeneous,[111] with respect to their gene composition, gene expression, and phenotypes which eventually yield current classifications of 5 subtypes.[111] The triple negative subtypes are more life threatening due to their potential to metastasize and a tendency of local reoccurrence.[112] They are usually associated with the absence of oestrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor-2 (ErbB2/HER-2).[113] They are characterised by classical ductal histology, high grade, high mitotic and cell proliferations rates.[113] The triple negative cancer (TNBC) is firmly associated with poor prognosis, poor disease-free survival (PDFS) and cancer-specific survival (CSS).[113] The local reoccurrence is marked with increasing number of positive lymph nodes,[113] this suggest the reason for high risk of reoccurrence in patients with TNBC in the first 3 to 5 years after diagnosis.[113]

Study have shown that only a few therapeutic options and conventional chemotherapy may probably be the only effective treatment for patients after surgery.[113] A more recent study established a relationship between TNBCs and sensitivity to inhibition of c-Src (non-receptor tyrosine kinase), in an attempt to identify predictive markers response to chemotherapy.[111] In this study, a dual kinase inhibitor known as UM-164 was discovered and had profound activity against c-Src and p38 kinases,[111] as shown in **Figure 6**. This inhibitor is said to be a promising lead compound for developing the first targeted therapeutic strategy against triple-negative breast cancer (TNBC).[111] c-Src is the cellular homolog of the viral oncogene v-Src[114] and an archetype member of a family of non-receptor tyrosine kinases that play important roles in a variety of signalling pathways that involve proliferation, differentiation, survival, motility, and angiogenesis.[114] Overexpression of c-Src plays an important role in oncogenic proliferation, migration, and invasion of TNBC cell lines.[111] This claim is supported by molecular studies that continued to show that c-Src plays a significant role in clinically important pathways in breast cancer,[114] such as steroid and peptide hormone pathway.[114] On the basis of these reports, c-Src represents an attractive target in TNBC.[115, 116] Although previous studies placed more emphasis on the use Dasatinib, Bosutinib and Saracatinib particularly as a combination therapy in the treatment of TNBC.[116] These drugs act by binding the active conformation of the kinase, in addition, resistance to Dasatinib has emerged.[117] Cumulatively, these factors placed other drugs (Dasatinib and Bosutinib) at the disadvantage over dual kinase inhibitor UM-164 (c-Src/p38 inhibitor) which act in a specific inactive conformation (DFG-out) as shown in **Figure 7** while **Figure 8** represent structure of UM-164 and Dasatinib.

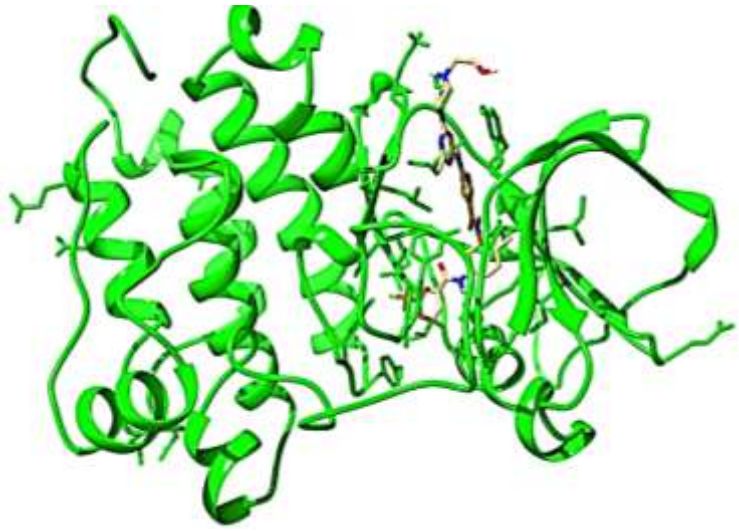


Figure 7. 3D structure of UM-164 bound to a non-receptor tyrosine kinase (c-Src)

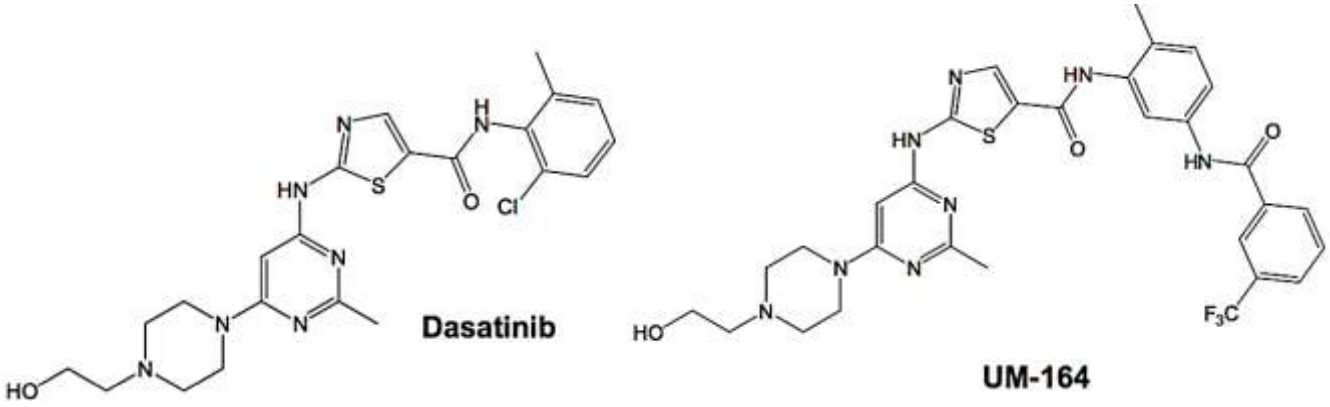


Figure 8. 2D structures of Dasatinib and experimental c-Src inhibitor (UM-164)

References

(1) Fearon, K., Strasser, F., Anker, S. D., Bosaeus, I., Bruera, E., Fainsinger, R. L., Jatoi, A., Loprinzi, C., MacDonald, N., Mantovani, G., Davis, M., Muscaritoli, M., Ottery, F., Radbruch, L., Ravasco, P., Walsh, D., Wilcock, A., Kaasa, S., and Baracos, V. E. (2011) Definition and classification of cancer cachexia: an international consensus. *Lancet Oncol. 12*, 489–495.

(2) Seyfried, T. N., Flores, R. E., Poff, A. M., and D'Agostino, D. P. (2014) Cancer as a metabolic disease: implications for novel therapeutics. *Carcinogenesis 35*, 515–27.

(3) Alizadeh, A. A., Aranda, V., Bardelli, A., Blanpain, C., Bock, C., Borowski, C., Caldas, C., Califano, A., Doherty, M., Elsner, M., Esteller, M., Fitzgerald, R., Korbel, J. O., Lichter, P., Mason, C. E., Navin, N., Pe'er, D., Polyak, K., Roberts, C. W. M., Siu, L., Snyder, A., Stower, H., Swanton, C., Verhaak, R. G. W., Zenklusen, J. C., Zuber, J., and Zucman-Rossi, J. (2015) Toward understanding and exploiting tumor heterogeneity. *Nat. Med. 21*, 846–853.

(4) Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D. M., Forman, D., and Bray, F. (2015) Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer 136*, E359–E386.

(5) Lazo, J. S., and Sharlow, E. R. (2016) Drugging Undruggable Molecular Cancer Targets. *Annu. Rev. Pharmacol. Toxicol. 56*, 23–40.

(6) Gyurkó, D. M., Veres, D. V, Módos, D., Lenti, K., Korcsmáros, T., and Csermely, P. (2013) Adaptation and learning of molecular networks as a description of cancer development at the systems-level: potential use in anti-cancer therapies. *Semin. Cancer Biol. 23*, 262–269.

(7) Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015) Global cancer statistics, 2012. *CA. Cancer J. Clin. 65*, 87–108.

(8) Jemal, A., Center, M. M., DeSantis, C., and Ward, E. M. (2010) Global Patterns of Cancer Incidence and Mortality Rates and Trends. *Cancer Epidemiol. Biomarkers Prev. 19*, 1893–1907.

(9) Anderson, B. O., Cazap, E., El Saghir, N. S., Yip, C.-H., Khaled, H. M., Otero, I. V, Adebamowo, C. A., Badwe, R. A., and Harford, J. B. (2011) Optimisation of breast cancer management in low-resource and middle-resource countries: executive summary of the Breast Health Global Initiative consensus, 2010. *Lancet Oncol. 12*, 387–398.

(10) Autier, P., Boniol, M., Gavin, A., and Vatten, L. J. (2011) Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. *BMJ 343*, d4411–d4411.

(11) Siegel, R. L., Miller, K. D., and Jemal, A. (2016) Cancer statistics, 2016. *CA. Cancer J.*

Clin. 66, 7–30.

(12) Kar, G., Gursoy, A., and Keskin, O. (2009) Human Cancer Protein-Protein Interaction Network: A Structural Perspective. *PLoS Comput Biol 5*.

(13) Kitano, H. (2004) Opinion: Cancer as a robust system: implications for anticancer therapy. *Nat. Rev. Cancer 4*, 227–235.

(14) Huang, S., Ernberg, I., and Kauffman, S. (2009) Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective. *Semin. Cell Dev. Biol. 20*, 869–76.

(15) Dulbecco, R. (1989) Cancer progression: The ultimate challenge. *Int. J. Cancer 44*, 6–9.

(16) Hornberg, J. J., Bruggeman, F. J., Westerhoff, H. V., and Lankelma, J. (2006) Cancer: A Systems Biology disease. *Biosystems 83*, 81–90.

(17) Dorogovtsev, S. N., and Mendes, J. F. F. (2002) Evolution of networks. *Adv. Phys.* 51, 1079

(18) Holme, P., and Saramäki, J. (2012) Temporal Networks. *Phys. Rep*. 519, 97-125

(19) Lin, C.-C., Chen, Y.-J., Chen, C.-Y., Oyang, Y.-J., Juan, H.-F., and Huang, H.-C. (2012) Crosstalk between transcription factors and microRNAs in human protein interaction network. *BMC Syst. Biol. 6*, 18.

(20) Klinke, D. J. (2010) Signal Transduction Networks in Cancer: Quantitative Parameters Influence Network Topology. *Cancer Res. 70*.

(21) Lichtenstein, A. V. (2010) Cancer: evolutionary, genetic and epigenetic aspects. *Clin. Epigenetics 1*, 85–100.

(22) Meidenbauer, J. J., Mukherjee, P., and Seyfried, T. N. (2015) The glucose ketone index calculator: a simple tool to monitor therapeutic efficacy for metabolic management of brain cancer. *Nutr. Metab. (Lond). 12*, 12.

(23) Seyfried, T. N., Flores, R. E., Poff, A. M., and D'Agostino, D. P. (2014) Cancer as a metabolic disease: implications for novel therapeutics. *Carcinogenesis 35*, 515–27.

(24) Fojo, T., and Parkinson, D. R. (2010) Biologically Targeted Cancer Therapy and Marginal Benefits: Are We Making Too Much of Too Little or Are We Achieving Too Little by Giving Too Much? *Clin. Cancer Res. 16*.

(25) Nowell, P. C. (1976) The clonal evolution of tumor cell populations. *Science 194*, 23–8.

(26) Shay, J. W., and Werbin, H. (1987) Are mitochondrial DNA mutations involved in the carcinogenic process? *Mutat. Res. Genet. Toxicol. 186*, 149–160.

(27) Israel, B. A., and Schaeffer, W. I. (1988) Cytoplasmic mediation of malignancy. *In Vitro Cell. Dev. Biol. 24*, 487–90.

(28) Petros, J. A., Baumann, A. K., Ruiz-Pesini, E., Amin, M. B., Sun, C. Q., Hall, J., Lim, S., Issa, M. M., Flanders, W. D., Hosseini, S. H., Marshall, F. F., and Wallace, D. C. (2005) mtDNA mutations increase tumorigenicity in prostate cancer. *Proc. Natl. Acad. Sci. U. S. A. 102*, 719–24.

(29) Warburg, O. (1931) The Metabolism of tumours : investigations from the Kaiser Wilhelm Institute for Biology, Berlin-Dahlem. Richard R. Smith, New York.

(30) Warburg, O. (1956) On the origin of cancer cells. *Science 123*, 309–14.

(31) Weinhouse, S. (1956) On respiratory impairment in cancer cells. *Science 124*, 267–9.

(32) Zuccoli, G., Marcello, N., Pisanello, A., Servadei, F., Vaccaro, S., Mukherjee, P., and Seyfried, T. N. (2010) Metabolic management of glioblastoma multiforme using standard therapy together with a restricted ketogenic diet: Case Report. *Nutr. Metab. (Lond). 7*, 33.

(33) Seyfried, T. N. (2012) Metabolic Management of Cancer, in *Cancer as a Metabolic Disease*, pp 291–354. John Wiley & Sons, Inc., Hoboken, NJ, USA.

(34) Gottschalk, S., Anderson, N., Hainz, C., Eckhardt, S. G., and Serkova, N. J. Imatinib (STI571)-Mediated Changes in Glucose Metabolism in Human Leukemia BCR-ABL–Positive Cells.

(35) Zhao, Y., Liu, H., Liu, Z., Ding, Y., Ledoux, S. P., Wilson, G. L., Voellmy, R., Lin, Y., Lin, W., Nahta, R., Liu, B., Fodstad, O., Chen, J., Wu, Y., Price, J. E., and Tan, M. (2011) Overcoming trastuzumab resistance in breast cancer by targeting dysregulated glucose metabolism. *Cancer Res. 71*, 4585–97.

(36) Mukherjee, P. (2004) Antiangiogenic and Proapoptotic Effects of Dietary Restriction on Experimental Mouse and Human Brain Tumors. *Clin. Cancer Res. 10*, 5622–5629.

(37) Simone, B. A., Champ, C. E., Rosenberg, A. L., Berger, A. C., Monti, D. A., Dicker, A. P., and Simone, N. L. (2013) Selectively starving cancer cells through dietary manipulation: methods and clinical implications. *Futur. Oncol. 9*, 959–976.

(38) Stafford, P., Abdelwahab, M. G., Kim, D. Y., Preul, M. C., Rho, J. M., and Scheck, A. C. (2010) The ketogenic diet reverses gene expression patterns and reduces reactive oxygen species levels when used as an adjuvant therapy for glioma. *Nutr. Metab. (Lond). 7*, 74.

(39) Zhou, W., Mukherjee, P., Kiebish, M. A., Markis, W. T., Mantis, J. G., and Seyfried, T. N. (2007) The calorically restricted ketogenic diet, an effective alternative therapy for malignant brain cancer. *Nutr. Metab. (Lond). 4*, 5.

(40) Laura M. Shelton,Leanne C. Huysentruyt, and Thomas N. Seyfried (2010). Glutamine targeting inhibits systemic metastasis in the VM-M3 murine tumor model. *Int J Cancers.* 2478-2485

(41) Yuneva, M. (2008) Finding an 'Achilles' heel' of cancer: The role of glucose and glutamine metabolism in the survival of transformed cells. *Cell Cycle 7*, 2083–2089.
(42) Lazova, R., Laberge, G. S., Duvall, E., Spoelstra, N., Klump, V., Sznol, M., Cooper, D., Spritz, R. A., Chang, J. T., and Pawelek, J. M. (2013) A Melanoma Brain Metastasis with a Donor-Patient Hybrid Genome following Bone Marrow Transplantation: First Evidence for Fusion in Human Cancer. *PLoS One 8*, e66731.
(43) Seyfried, T. N., and Huysentruyt, L. C. (2013) On the origin of cancer metastasis. *Crit. Rev. Oncog. 18*, 43–73.
(44) Newsholme, P. (2001) Why is L-glutamine metabolism important to cells of the immune system in health, postinjury, surgery or infection? *J. Nutr. 131*, 2515S–22S; discussion 2523S–4S.
(45) Veech, R. L. (2004) The therapeutic implications of ketone bodies: the effects of ketone bodies in pathological conditions: ketosis, ketogenic diet, redox states, insulin resistance, and mitochondrial metabolism. *Prostaglandins. Leukot. Essent. Fatty Acids 70*, 309–19.
(46) Veech, R. L., Chance, B., Kashiwaya, Y., Lardy, H. A., and Cahill, G. F. (2001) Hypothesis Paper Ketone Bodies, Potential Therapeutic Uses. *IUBMB Life 51*, 241–247.
(47) Elliott, R. L., Jiang, X. P., and Head, J. F. (2012) Mitochondria organelle transplantation: introduction of normal epithelial mitochondria into human cancer cells inhibits proliferation and increases drug sensitivity. *Breast Cancer Res. Treat. 136*, 347–354.
(48) Loeb, L. A., Loeb, K. R., and Anderson, J. P. (2003) Multiple mutations and cancer. *Proc. Natl. Acad. Sci. U. S. A. 100*, 776–81.
(49) Jackson, A. L., and Loeb, L. A. (2001) The contribution of endogenous sources of DNA damage to the multiple mutations in cancer. *Mutat. Res. Mol. Mech. Mutagen. 477*, 7–21.
(50) Kallioniemi, A., Kallioniemi, O. P., Sudar, D., Rutovitz, D., Gray, J. W., Waldman, F., and Pinkel, D. (1992) Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science 258*, 818–21.
(51) Stanchescu, R., Betts, D. R., Yekutieli, D., Ambros, P., Cohen, N., Rechavi, G., Amariglio, N., and Trakhtenbrot, L. SKY analysis of childhood neural tumors and cell lines demonstrates a susceptibility of aberrant chromosomes to further rearrangements.*bjh. 275-288*
(52) Tlsty, T. D., Margolin, B. H., and Lum, K. (1989) Differences in the rates of gene amplification in nontumorigenic and tumorigenic cell lines as measured by Luria-Delbrück fluctuation analysis. *Proc. Natl. Acad. Sci. U. S. A. 86*, 9441–5.
(53) Watatani, M., Inui, H., Nagayama, K., Imanishi, Y., Nishimura, K., Hashimoto, Y., Yamauchi, E., Hojo, T., Kotsuma, Y., Yamato, M., Matsunami, N., and Yasutomi, M. (2000)

Identification of high-risk breast cancer patients from genetic changes of their tumors. *Surg. Today 30*, 516–522.

(54) Luqmani, Y. A. (2005) Mechanisms of drug resistance in cancer chemotherapy. *Med. Princ. Pract. 14 Suppl 1*, 35–48.

(55) Twelves, C., Jove, M., Gombos, A., and Awada, A. (2016) Cytotoxic chemotherapy: Still the mainstay of clinical practice for all subtypes metastatic breast cancer. *Crit. Rev. Oncol. Hematol. 100*, 74–87.

(56) Mitchell, S. J., Kane, A. E., Hilmer, S. N., Mitchell, S. J., Kane, A. E., and Hilmer, S. N. (2011) Age-related changes in the hepatic pharmacology and toxicology of paracetamol. *Curr. Gerontol. Geriatr. Res. 2011*, 624156.

(57) Lahti, J. L., Tang, G. W., Capriotti, E., Liu, T., and Altman, R. B. (2012) Bioinformatics and variability in drug response: a protein structural perspective. *J. R. Soc. Interface 9*, 1409–37.

(58) Bunney, T. D., Wan, S., Thiyagarajan, N., Sutto, L., Williams, S. V, Ashford, P., Koss, H., Knowles, M. A., Gervasio, F. L., Coveney, P. V, and Katan, M. (2015) The Effect of Mutations on Drug Sensitivity and Kinase Activity of Fibroblast Growth Factor Receptors: A Combined Experimental and Theoretical Study. *EBioMedicine 2*, 194–204.

(59) Drews, J. (2006) Case histories, magic bullets and the state of drug discovery. *Nat. Rev. Drug Discov. 5*, 635–640.

(60) Tosato, M., Zamboni, V., Ferrini, A., and Cesari, M. (2007) The aging process and potential interventions to extend life expectancy. *Clin. Interv. Aging 2*, 401–12.

(61) Weigelt, J. (2010) Structural genomics—Impact on biomedicine and drug discovery. *Exp. Cell Res. 316*, 1332–1338.

(62) Holohan, C., Van Schaeybroeck, S., Longley, D. B., and Johnston, P. G. (2013) Cancer drug resistance: an evolving paradigm. *Nat. Rev. Cancer 13*, 714–726.

(63) Alfarouk, K. O., Stock, C.-M., Taylor, S., Walsh, M., Muddathir, A. K., Verduzco, D., Bashir, A. H. H., Mohammed, O. Y., Elhassan, G. O., Harguindey, S., Reshkin, S. J., Ibrahim, M. E., and Rauch, C. (2015) Resistance to cancer chemotherapy: failure in drug response from ADME to P-gp. *Cancer Cell Int. 15*, 71.

(64) Thorn, M., Finnstrom, N., Lundgren, S., Rane, A., and Loof, L. (2005) Cytochromes P450 and MDR1 mRNA expression along the human gastrointestinal tract. *Br. J. Clin. Pharmacol. 60*, 54–60.

(65) Wacher, V. J., Salphati, L., and Benet, L. Z. (2001) Active secretion and enterocytic drug metabolism barriers to drug absorption. *Adv. Drug Deliv. Rev. 46*, 89–102.

(66) Sparreboom, A., van Asperen, J., Mayer, U., Schinkel, A. H., Smit, J. W., Meijer, D. K., Borst, P., Nooijen, W. J., Beijnen, J. H., and van Tellingen, O. (1997) Limited oral bioavailability and active epithelial excretion of paclitaxel (Taxol) caused by P-glycoprotein in the intestine. *Proc. Natl. Acad. Sci. U. S. A. 94*, 2031–5.

(67) Cascorbi, I. (2001) Frequency of single nucleotide polymorphisms in the P-glycoprotein drug transporter MDR1 gene in white subjects. *Clin. Pharmacol. Ther. 69*, 169–174.

(68) Dietrich, C. G., Geier, A., and Oude Elferink, R. P. J. (2003) ABC of oral bioavailability: transporters as gatekeepers in the gut. *Gut 52*, 1788–95.

(69) Fujita, K. (2004) Food-drug interactions via human cytochrome P450 3A (CYP3A). *Drug Metabol. Drug Interact. 20*, 195–217.

(70) Singh, B. N., and Malhotra, B. K. (2004) Effects of food on the clinical pharmacokinetics of anticancer agents: underlying mechanisms and implications for oral chemotherapy. *Clin. Pharmacokinet. 43*, 1127–56.

(71) Ruggiero, A., Cefalo, M. G., Coccia, P., Mastrangelo, S., Maurizi, P., and Riccardi, R. (2012) The role of diet on the clinical pharmacology of oral antineoplastic agents. *Eur. J. Clin. Pharmacol. 68*, 115–122.

(72) Wilson, K. (1984) Sex-Related Differences in Drug Disposition in Man. *Clin. Pharmacokinet. 9*, 189–202.

(73) Pleym, H., Spigset, O., Kharasch, E. D., and Dale, O. (2003) Gender differences in drug effects: implications for anesthesiologists. *Acta Anaesthesiol. Scand. 47*, 241–59.

(74) Fearon, K. C., Voss, A. C., Hustead, D. S., and Cancer Cachexia Study Group. (2006) Definition of cancer cachexia: effect of weight loss, reduced food intake, and systemic inflammation on functional status and prognosis. *Am. J. Clin. Nutr. 83*, 1345–50.

(75) Theologides, A. (1977) Nutritional management of the patient with advanced cancer. *Postgrad. Med. 61*, 97–101.

(76) Tillement, J.-P., Zini, R., d'Athis, P., and Vassent, G. (1974) Binding of certain acidic drugs to human albumin: Theoretical and practical estimation of fundamental parameters. *Eur. J. Clin. Pharmacol. 7*, 307–313.

(77) Routledge, P. A. (1986) The plasma protein binding of basic drugs. *Br. J. Clin. Pharmacol. 22*, 499–506.

(78) Ruiz-Cabello, F., and Erill, S. (1984) Abnormal serum protein binding of acidic drugs in diabetes mellitus. *Clin. Pharmacol. Ther. 36*, 691–5.

(79) Arora, B., Gota, V., Menon, H., Sengar, M., Nair, R., Patial, P., and Banavali, S. D. (2013) Therapeutic drug monitoring for imatinib: Current status and Indian experience. *Indian J. Med.*

Paediatr. Oncol. 34, 224–8.

(80) Li, J., Brahmer, J., Messersmith, W., Hidalgo, M., and Baker, S. D. (2006) Binding of gefitinib, an inhibitor of epidermal growth factor receptor-tyrosine kinase, to plasma proteins and blood cells: in vitro and in cancer patients. *Invest. New Drugs 24*, 291–297.

(81) Levi, F., and Schibler, U. (2007) Circadian Rhythms: Mechanisms and Therapeutic Implications. *Annu. Rev. Pharmacol. Toxicol. 47*, 593–628.

(82) Sukumaran, S., Almon, R. R., DuBois, D. C., and Jusko, W. J. (2010) Circadian rhythms in gene expression: Relationship to physiology, disease, drug disposition and drug action. *Adv. Drug Deliv. Rev. 62*, 904–17.

(83) Lévi, F., Okyar, A., Dulong, S., Innominato, P. F., and Clairambault, J. (2010) Circadian Timing in Cancer Treatments. *Annu. Rev. Pharmacol. Toxicol. 50*, 377–421.

(84) Straub, R. H., and Cutolo, M. (2007) Circadian rhythms in rheumatoid arthritis: Implications for pathophysiology and therapeutic management. *Arthritis Rheum. 56*, 399–408.

(85) Prémaud, A., Rousseau, A., Gicquel, M., Ragot, S., Manceau, J., Laurentie, M., and Marquet, P. (2002) An animal model for the study of chronopharmacokinetics of drugs and application to methotrexate and vinorelbine. *Toxicol. Appl. Pharmacol. 183*, 189–97.

(86) Touitou, Y., Carayon, A., Reinberg, A., Bogdan, A., and Beck, H. (1983) Differences in the seasonal rhythmicity of plasma prolactin in elderly human subjects: detection in women but not in men. *J. Endocrinol. 96*, 65–71.

(87) Reinberg, A., Guillet, P., and Relnberg, M. (2016) Differences between young and elderly subjects in seasonal and circadian variations of total plasma proteins and blood ... betweenYoung and ElderlySubjectsin Seasonaland Circadian Variationsof Total Plasma Proteinsand BloodVolumeas Reflectedby.*Clin Chem.* 32, 801-804

(88) Renwick, A. G., and Lazarus, N. R. (1998) Human Variability and Noncancer Risk Assessment— An Analysis of the Default Uncertainty Factor agencies may vary the uncertainty factor in particular circumstances, for example for an incomplete toxicity. *Regul. Toxicol. Pharmacol. 27*, 3–20.

(89) Herak-Kramberger, C. M., Breljak, D., Ljubojević, M., Matokanović, M., Lovrić, M., Rogić, D., Brzica, H., Vrhovac, I., Karaica, D., Micek, V., Dupor, J. I., Brown, D., and Sabolić, I. (2015) Sex-dependent expression of water channel AQP1 along the rat nephron. *Am. J. Physiol. Renal Physiol. 308*, F809–21.

(90) Johnson, J. A. (2000) Predictability of the effects of race or ethnicity on pharmacokinetics of drugs. *Int. J. Clin. Pharmacol. Ther. 38*, 53–60.

(91) Zhao, J., Cheng, F., Wang, Y., Arteaga, C. L., and Zhao, Z. (2016) Systematic

Prioritization of Druggable Mutations in ~5000 Genomes Across 16 Cancer Types Using a Structural Genomics-based Approach. *Mol. Cell. Proteomics 15*, 642–656.

(92) Gross, S., Rahal, R., Stransky, N., Lengauer, C., and Hoeflich, K. P. (2015) Targeting cancer with kinase inhibitors. *J. Clin. Invest. 125*, 1780–1789.

(93) Gavande, N. S., VanderVere-Carozza, P. S., Hinshaw, H. D., Jalal, S. I., Sears, C. R., Pawelczak, K. S., and Turchi, J. J. (2016) DNA repair targeted therapy: The past or future of cancer treatment? *Pharmacol. Ther. 160*, 65–83.

(94) Helleday, T., Petermann, E., Lundin, C., Hodgson, B., and Sharma, R. A. (2008) DNA repair pathways as targets for cancer therapy. *Nat. Rev. Cancer 8*, 193–204.

(95) Kaelin, W. G., and Jr. (2009) Synthetic lethality: a framework for the development of wiser cancer therapeutics. *Genome Med. 1*, 99.

(96) Frezza, M., Hindo, S., Chen, D., Davenport, A., Schmitt, S., Tomco, D., and Dou, Q. P. (2010) Novel metals and metal complexes as platforms for cancer therapy. *Curr. Pharm. Des. 16*, 1813–25.

(97) Dasari, S., and Tchounwou, P. B. (2014) Cisplatin in cancer therapy: molecular mechanisms of action. *Eur. J. Pharmacol. 740*, 364–78.

(98) Cheung-Ong, K., Giaever, G., and Nislow, C. (2013) DNA-Damaging Agents in Cancer Chemotherapy: Serendipity and Chemical Biology. *Chem. Biol. 20*, 648–659.

(99) Jemal, A., Siegel, R., Xu, J., and Ward, E. (2010) Cancer Statistics, 2010. *CA. Cancer J. Clin. 60*, 277–300.

(100) Hotte, S. J., and Saad, F. (2010) Current management of castrate-resistant prostate cancer. *Curr. Oncol. 17 Suppl 2*, S72–9.

(101) Kim, S. J., and Kim, S. Il. (2011) Current treatment strategies for castration-resistant prostate cancer. *Korean J. Urol. 52*, 157–65.

(102) Chrvala, C. A. (2012) The changing landscape of treatment options for metastatic castrate-resistant prostate cancer: challenges and solutions for physicians and patients. *P T 37*, 453–63.

(103) Paola, E. D. Di, Alonso, S., Giuliani, R., Calabrò, F., D'Alessio, A., Regine, G., Cerbone, L., Bianchi, L., Mancuso, A., Sperka, S., Rozencweig, M., and Sternberg, C. N. (2012) An open-label, dose-finding study of the combination of satraplatin and gemcitabine in patients with advanced solid tumors. *Front. Oncol. 2*, 175.

(104) Sternberg, C. N., Whelan, P., Hetherington, J., Paluchowska, B., Slee, P. H. T. J., Vekemans, K., van Erps, P., Theodore, C., Koriakine, O., Oliver, T., Lebwohl, D., Debois, M., Zurlo, A., and Collette, L. (2005) Phase III Trial of Satraplatin, an Oral Platinum plus

Prednisone vs. Prednisone alone in Patients with Hormone-Refractory Prostate Cancer. *Oncology 68*, 2–9.

(105) Dror Michaelson, M., Regan, M. M., Oh, W. K., Kaufman, D. S., Olivier, K., Michaelson, S. Z., Spicer, B., Gurski, C., Kantoff, P. W., and Smith, M. R. (2009) Phase II study of sunitinib in men with advanced prostate cancer. *Ann. Oncol. 20*, 913–920.

(106) Cook, D. N., Kang, H. S., and Jetten, A. M. (2015) Retinoic Acid-Related Orphan Receptors (RORs): Regulatory Functions in Immunity, Development, Circadian Rhythm, and Metabolism. *Nucl. Recept. Res. 2*.

(107) Wang, J., Zou, J. X., Xue, X., Cai, D., Zhang, Y., Duan, Z., Xiang, Q., Yang, J. C., Louie, M. C., Borowsky, A. D., Gao, A. C., Evans, C. P., Lam, K. S., Xu, J., Kung, H.-J., Evans, R. M., Xu, Y., and Chen, H.-W. (2016) Corrigendum: ROR-γ drives androgen receptor expression and represents a therapeutic target in castration-resistant prostate cancer. *Nat. Med. 22*, 692–692.

(108) Roberts, R. (2014) From bench to bedside: the realities of reducing global prostate cancer disparity in black men. *Ecancermedicalscience 8*, 458.

(109) Sharma, G. N., Dave, R., Sanadya, J., Sharma, P., and Sharma, K. K. (2010) Various types and management of breast cancer: an overview. *J. Adv. Pharm. Technol. Res. 1*, 109–26.

(110) Sharma, G. N., Dave, R., Sanadya, J., Sharma, P., and Sharma, K. K. (2010) Various types and management of breast cancer: an overview. *J. Adv. Pharm. Technol. Res. 1*, 109–26.

(111) Gilani, R. A., Phadke, S., Bao, L. W., Lachacz, E. J., Dziubinski, M. L., Brandvold, K. R., Steffey, M. E., Kwarcinski, F. E., Graveel, C. R., Kidwell, K. M., Merajver, S. D., and Soellner, M. B. (2016) UM-164: A Potent c-Src/p38 Kinase Inhibitor with In Vivo Activity against Triple-Negative Breast Cancer. *Clin. Cancer Res. 22*, 5087–5096.

(112) Anders, C. K., and Carey, L. A. (2009) Biology, metastatic patterns, and treatment of patients with triple-negative breast cancer. *Clin. Breast Cancer 9 Suppl 2*, S73–81.

(113) Jiao, Q., Wu, A., Shao, G., Peng, H., Wang, M., Ji, S., Liu, P., and Zhang, J. (2014) The latest progress in research on triple negative breast cancer (TNBC): risk factors, possible therapeutic targets and prognostic markers. *J. Thorac. Dis. 6*, 1329–35.

(114) Finn, R. S., Dering, J., Ginther, C., Wilson, C. A., Glaspy, P., Tchekmedyian, N., and Slamon, D. J. (2007) Dasatinib, an orally active small molecule inhibitor of both the src and abl kinases, selectively inhibits growth of basal-type/'triple-negative' breast cancer cell lines growing in vitro. *Breast Cancer Res. Treat. 105*, 319–326.

(115) Finn, R. S. (2008) Targeting Src in breast cancer. *Ann. Oncol. 19*, 1379–1386.

(116) Mayer, E. L., and Krop, I. E. (2010) Advances in Targeting Src in the Treatment of Breast

Cancer and Other Solid Malignancies. *Clin. Cancer Res. 16*, 3526–3532.
(117) Getlik, M., Grütter, C., Simard, J. R., Klüter, S., Rabiller, M., Rode, H. B., Robubi, A., and Rauh, D. (2009) Hybrid Compound Design To Overcome the Gatekeeper T338M Mutation in cSrc [#]. *J. Med. Chem. 52*, 3915–3926.

CHAPTER 3

Introduction to computation chemistry

3.1 Introduction

Computational chemistry can be define as the use of computer technique to simulate chemical compounds at the atomic or molecular level. It is also referred to as molecular modelling, and principally based on the use of theoretical chemistry incorporated into computer system. The term includes a diverse range of theories and methods which provide a problem solving tools in chemistry.

3.1 Quantum mechanics

In 1990 a scientist called Max Plank reported the absorption and spectral activity of compounds which was labelled as quantum. This was thought to originate from a Latin word *quantu* meaning how much. Quantum is measure of constrained particle quantities [1-4] this particles are said to be influence by gravitational force on the motions, the forces interplay within the object is referred to as Mechanics. The nuclei, electrons and protons are the major components of the molecules, therefore, quantum chemistry involves the study of influence of electromagnetic forces resulting from the activity of nuclear charges. [1-4] The variation in chemical activity is attributed to differences in molecular structure, therefore, adequate knowledge of electron dynamics in the molecules can only be taped through good understanding of quantum chemistry particularly the use of Schrödinger equation.[1-5]

3.2 The Schrödinger equation

The famous Schrödinger equation was proposed and developed by an Australia physicist Erwin Schrodinger in 1926.[4] This equation is presented in its simples form as follow: The

$$H\psi = E\psi \qquad \text{Eq. 3.1}$$

From this equation, H represents molecular Hamilton, ψ a wave function and E is the system energy.[2, 6] From the Eq. 3.1 molecular Hamilton can be derived as:

$$H = T + V \qquad \text{Eq. 3.2}$$

Where T is the kinetic and V potential energy. Considering the kinetic and potential energy of electrons and nuclei of molecules, Hamilton can be represented as:

$$H = -\frac{\hbar^2}{2me}\sum_i \nabla_i^2 - \frac{\hbar^2}{2me}\sum_A \frac{1}{MA}\nabla_A^2 - \sum_A \sum_i \frac{Z_A e^2}{r\,A_i} + \sum_i \sum_j > i \frac{e^2}{r_{ij}} + \sum_A \sum_B > A \frac{Z_A Z}{R_{AB}} \quad \text{Eq. 3.3}$$

In this equation, *A* and *B* are the nuclei and *i* and *j* are the electrons, M_A is the mass of the nuclei *A*, *me* mass of an electron, R_{AB} the distance between the nuclei *A* and *B*, r_{ij} the distance between the electron *i* and *j*, Z_A is the charge of the nucleus *A*, r_{Ai} the distance between the nucleus *A* and electron *i*. The complexity associated with Schrödinger equation pose some challenge in attempt to use it to solve for a molecular system. However, Born-Oppenheimer [1,7] equation can be used to approximation.

3.3 Born-Oppenheimer approximation

The variation between nuclei of an atom and the electron can be optimally utilised in separating motion associated with nuclei and the electron. This mainly because nuclei is much heavier (has slow motion) than electron which move with a faster motion. Therefore, the motions can easily separated[1,7] this is the assumption in Born-Oppenheimer approximation.[8] This equation account for the differences in the mass of electron and the nuclei thus electron can easily replace nuclei. The equation can be presented as follow:

$$H = -\frac{\hbar^2}{2me}\sum_i \nabla_i^2 - \sum_A \sum_i \frac{Z_A e^2}{r\,A_i} + \sum_i \sum_j > i \frac{e^2}{r_{ij}} + \sum_A \sum_B > A \frac{Z_A Z}{R_{AB}} \quad \text{Eq.3.4}$$

The observed differences between the nuclei and electron allow for the Born-Oppenheimer equation be used in reducing the Hamilton and complexity in wave function, given by the following equation:

$$\psi\, r\,(r_{elec}) = \psi\,(r_{elec})\,(r_{nucl}) \quad \text{Eq. 3.5}$$

Eq. 3.1 is re-arranged as:

$$H_{EN}\,\psi\,(r_{el}) = E_{EN}\psi\,(r_{el}) \quad \text{Eq.3.6}$$

Eq. 3.5 and 3.6 are integrated to form equation Eq. 3.7 as follow:

$$(H_{el} + V_{NN})\,\psi\,(rel) = E_{EN}\psi\,(r_{el}) \quad \text{Eq. 3.7}$$

This equation account for energy from variable sources due to differences that exist between the nuclei and the electrons.[3]

3.4 Potential energy surface

The observed relationship in the geometry of the molecular energy and graphical or mathematical equation is termed potential energy surface (PES).[3] This determines the differences in molecular energy and various energy transition states and the coordinates of their nuclei.[3, 9, 10] The PES is better explained by Born-Oppenheimer approximation which put into consideration the observed difference between the nuclei and the electron motions.[8,10] Within the this concept, the position of electron becomes flexible can easily adjust to changes due nuclei movement, this way PES can be related to the motion of the atoms within molecules.[11] **Figure 1** is a graphical representation of two dimensional potential energy surface.

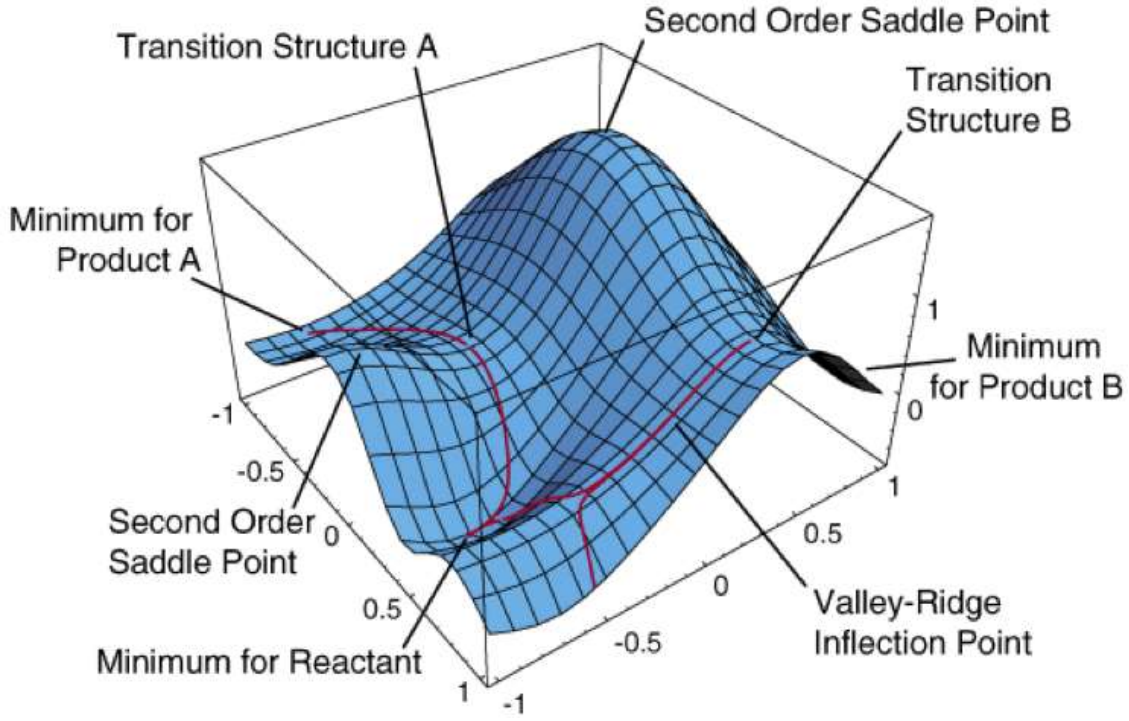


Figure 1. Graphical representation of a two dimensional potential energy surface [11]

The PES exhibits the two main regions; high energy and low energy region respectively. These regions provide a clear understanding of the status of nuclear arrangement in a specific conformational space. [9, 10] In line with this, association between the electrons resulting from prompt response to changes in electron interaction within molecular orbital[12] formed the foundation in which computational chemistry is built upon.[12]

3.5 Molecular mechanics

The energies of molecular structure are conveniently calculate using molecular mechanics, it involves calculating the relative potential energy molecular system.[3, 6, 7] In this case, the electrons and it associated nuclei are treated as a single entity, this is better explained by Born-Oppenheimer approximation as stated earlier

Molecular mechanics (MM) or better referred to as total potential energy of a molecule is described by equation bellow:

$$E_{tot} = E_{str} + E_{bend} + E_{tor} + E_{vdw} + E_{elec} \quad \text{Eq. 3.8}$$

From the above, E_{tot} is the total potential energy, E_{str} bond-stretching energy, E_{bend} bond-angle bending energy, E_{tor} the torsion energy, E_{vdw} the van der Waals forces and E_{elec} electrostatic forces between atoms that are non-chemically bonded. Energy contributions from special treatment of hydrogen bonding and stretch-bend coupling interactions may also be witnessed in MM.

3.5.1 Force field

The energy of a system largely depends on the coordinates of it particles, this relationship can be defined by a set of mathematical equation called force field. Simulation of biomolecules involved the use of force field such as AMBER,[13] CHARM,[14] GROMOS[15] and NAMD[16]. These force fields are product of *ab initio* and are semi-empirically calculated from quantum mechanics or obtained from experimental data.[17] Force fields exhibits variable properties, therefore, they are used for different purpose depending on the complexity of the system involve. Herein, AMBER14 force field[18] was used, whereby the General AMBER Force field (GAFF) parameters were applied for ligand and parameterization and the standard AMBER force field for the protein.

3.6 Molecular dynamics

The most commonly used computer technique in in the simulation of complex system is molecular dynamics (MD). It involved generation of atomic trajectories from of many particles in the system using Newton equation of motion.[19, 20]

$$F_i = m_i \frac{d^2 r_i(t)}{dt^2} \quad \text{Eq. 3.9}$$

Where $ri\ (t) = (x_i\ (t), y_i\ (t), z_i\ (t))$ is the (i) particle position vector, Fi is the force acting on (i) particle at time t and m_i particle mass. Integration of this equation will required specifying the forces that are acting on the particles.[17] MD is associated with time evolution of the system thus position and velocities are transmitted with a precise time interval.[21] The locus particle in the space is given by $ri\ (t)$ and V_i (t) which defines the system's kinetic energy and temperature. With MD dynamics process can be monitored at atomic level.[22]

3.7 Hybrid QM/MM method

Quantum mechanics (QM) and Molecular Mechanics (MM) are associated with certain challenges,[23] this necessitate the harmonisation of the two methods to eliminate set back involve in the use of individual method and enhance the optimum performance of the hybrid type.[23] The QM is associated with good descriptive accuracy therefore when combined with a low computational cost MM result in outstanding algorithm.[23] In hybrid QM/MM the system is sectioned with each section or domain describing a particular function of a molecule as shown in Figure 3.[24] The QM is applied in reactive domain while the MM is applied in non-reactive domain,[25, 26, 27] hence the combination of high precision QM methods with low computation enjoyed good acceptance from the users during molecular simulation.[28] **Figure 2** is a schematic representation of hybrid QM/MM/MD model.

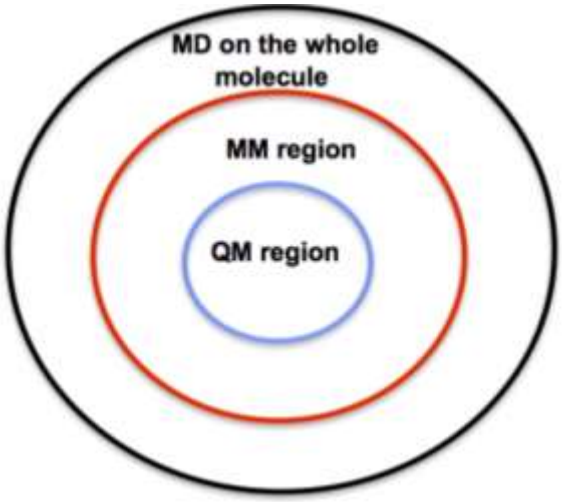


Figure 2. Schematic representation of hybrid QMM/MM/MD model

MM methods is not suitable in addressing the relative interaction between ligand and residue because of its inability to accurately describe changes in the energy in the course of reaction.

In line with this, the hybrid QM/MM play important role in computational drug design and discovery.

3.8 Binding free energy

The binding free energy calculation is an important thermodynamic method that gives detailed information on the interaction between the ligand and protein.[29] It provides good understanding of mechanism of binding, including contributions from both enthalpy and entropy to the molecular recognition.[29] Molecular Mechanics/Generalized-Born Surface Area method (MM/GBSA)[29] are popular methods used to estimate free energy of the binding of small ligands to the biological macromolecule,[29] the calculation gives detailed information on the interaction between the ligand and protein.[29, 30] The following equations were used in binding free energy calculation:

$$\Delta G_{bind} = G_{complex} - G_{receptor} - G_{ligand} \quad \text{Eq.3.10}$$

$$\Delta G_{bind} = E_{gas} + G_{sol} - TS \quad \text{Eq.3.11}$$

$$E_{gas} = E_{int} + E_{vdW} + E_{ele} \quad \text{Eq.3.12}$$

$$G_{sol} = G_{GB} + G_{SA} \quad \text{Eq.3.13}$$

From the equation above, E_{gas} is the energy of the gas phase, E_{int} represents internal energy, E_{ele} represents coulomb while E_{vdW} are the van der Waals energies. E_{gas} is estimated directly from the ff12SB force field. G_{sol} which is the solvation free energy can be broken down to polar and non-polar forms of contribution. The contribution of polar solvation (G_{GB}) is assessed by resolving G_{GB} equation and non-polar solvation (G_{SA}) is determined from the solvent accessible surface area, this can be estimated from water probe radius of 1.4 Å with temperature (T) and total solute entropy (S). The MM/GBSA binding free energy method in Amber 14 was used in this study to calculate the energy of interactive forces to the binding of the ligand. In addition, the energy decomposition analysis per residue was also computed using the same method.

3.9 Principal component analysis

Classical molecular dynamics analysis can only give limited insight into dynamic landscape framework of large and complex biomolecular systems. Therefore, other computational

methods have been developed to handle a large number of explicit degree of freedom.[31, 32, 33, 34] One of such is principal component analysis (PCA), which explore atomic fluctuations experienced within the biological system.

Principal component analysis (PCA) also known as essential dynamics of protein[35] analysis, is a systematic statistical technique applied to reduce the number of dimensions needed to describe the protein dynamics[35] through the decomposition process that screen observed motions from largest to smallest spatial scale.[35] The atomic displacement and conformational changes of protein can be defined[36] using PCA by extracting different modes of the conformation of the protein complex during dynamic simulations such as MD simulation. The direction of motion (eigenvectors) and the extent of motion (eigenvalues) for the biological system can also be determine using PCA.

3.10 Residue interaction network (RIN)

Amino acids are an essential component of biological system with a well-defined network of interaction.[37] The molecular interactions of these amino acids vary among the body tissues and are coded for by the genes, therefore an alteration in the gene sequence affect the network of the interaction of amino acid.[38] This is particularly true when mutation occur, as a result the protein sequence is re-writing and there is an alteration in the network of amino acid. [38] Mutation can affect protein folding and stability, protein function and protein-protein interaction.[39] Certain studies have demonstrated that analysis of interactions from the RIN provides insight on the vital biological interaction of highly complex molecular systems.[40,41]

The RIN analyser software, a java plugin for cystoscape, is used for analysing and visualisation RINs in different proteins and biological systems. It also provides insight into structural and functional properties or connections in different systems. The 3D protein structure network generated by cytoscape can be visualised and explore for a various network of interactions. The basic layout of interaction network display network nodes that represents various amino acid residues, with edges representing the respective molecular interaction between theses amino acids.

3.11 Prediction of activity spectra for biologically active substance (PASS)

Most chemical compounds interact with biological targets different from experimentally recognised targets, such compounds are said to be promiscuous in nature.[42] Compound promiscuity depicts the molecular basis of the pharmacological effects, therefore, assessment of the extent of promiscuity of compounds at different levels of drug research[43] provides a

detailed understanding of other properties of the drug that were probably not anticipated. Toxicity and biological activity of compounds can be predicted by computer tools. The molecular structure in "mol" form is upload into PASS website where the analysis is done and the result of prediction is generated within few minutes. It has been shown that the degree of reliability of these tools varies from one to another.[44,45] However, PASS predicts compound toxicity and biological profile with a mean accuracy of prediction of about 89% to 90%.[44, 45] In this study, PASS was used to predict the toxicity and biological effects of experimental compounds investigated.

References

(1) Schrödinger, E. (1926) An Undulatory Theory of the Mechanics of Atoms and Molecules. *Phys. Rev. 28*, 1049–1070.

(2) Frank Jensen. (1999) Introduction to Computational Chemistry. *Wiley Sons 2nd Ed.*, 1–573.

(3) Atkins, P., de Paula, J. (2006) Atkins physical chemistry 8th edition.

(4) Gray, S. (2008) Introduction to Quantum Mechanics: A Time-Dependent Perspective: David J. Tannor: 9781891389238: Amazon.com: Books. *Sci.* 319, 161–161.

(5) Smith, E. B. (2013) Basic Physical Chemistry. *Imperial college press.*

(6) Lewars, E. (2010). Introduction to the theory and applications of molecular and quantum mechanics. *Computational chemistry*

(7) Rogers, D. W. (2003) Computational Chemistry Using the PC Third Edition. *John Wiley Sons 3rd Ed.*

(8) Born, M., and Oppenheimer, R. (1927) Zur Quantentheorie der Molekeln. *Ann. Phys. 389*, 457–484.

(9) Elsawy, K. M., Hodgson, M. K., and Caves, L. S. D. (2005) The physical determinants of the DNA conformational landscape: an analysis of the potential energy surface of single-strand dinucleotides in the conformational space of duplex DNA. *Nucleic Acids Res. 33*, 5749–62.

(10) Truhlar, D. G. (2001) Potential Energy Surfaces Quantum Mechanical Basis for Adiabatic Potential Energy Surfaces III. Topology of Adiabatic Potential Energy Surfaces IV. Breakdown of the Adiabatic Approximation V. Shapes of Potential Energy Surfaces.*Encyclopedia of physical Science and Technology 3rd edition*

(11) Lewars, E. G. (2011) The Concept of the Potential Energy Surface, in *Computational Chemistry*, pp 9–43. Springer Netherlands, Dordrecht.

(12) Peter Atkins, Ronald Friedman (2005) Molecular quantum mechanics. 4th edition p576

(13) Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem. 26*, 1668–1688.

(14) Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009) CHARMM: The biomolecular simulation program. *J. Comput. Chem. 30*, 1545–1614.

(15) Christen, M., Hünenberger, P. H., Bakowies, D., Baron, R., Bürgi, R., Geerke, D. P., Heinz, T. N., Kastenholz, M. A., Kräutler, V., Oostenbrink, C., Peter, C., Trzesniak, D., and van Gunsteren, W. F. (2005) The GROMOS software for biomolecular simulation: GROMOS05. *J. Comput. Chem. 26*, 1719–1751.

(16) Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem. 26*, 1781–1802.

(17) González, M. A. (2011) Force fields and molecular dynamics simulations. *École thématique la Société Française la Neutron. 12*, 169–200.

(18) Case Tom Darden Thomas E Cheatham III Carlos Simmerling Adrian Roitberg Junmei Wang, D. A., Götz SDSC, A. W., István Kolossváry Budapest, U., and Shaw Francesco Paesani Jian Liu Xiongwu Wu Thomas Steinbrecher, D. *Amber 14 Reference Manual*.

(19) Petrenko, R., and Meller, J. (2010) Molecular Dynamics, in *Encyclopedia of Life Sciences*. John Wiley & Sons, Ltd, Chichester, UK.

(20) Karplus, M. (2003) Molecular dynamics of biological macromolecules: A brief history and perspective. *Biopolymers 68*, 350–358.

(21) Binder, K., Horbach, J., Kob, W., Paul, W., Varnik, F., Allen M P, T. D. J., Binder K, C. G., Frenkel D, S. B., Landau D P, B. K., E, A. B. J. and W. T., E, A. B. J. and W. T., A, R., A, R., C, R. D., S, K. K. and G. G., K, B., Metropolis N, R. A. W. R. M. N. T. A. N. and T. E., K, B., M, C. R. and P., W, K., R, C., D, B. B. J. and T., M, C. D., M, C. D., P, N., H, M. D. and M. M., Tuckerman M E, B. B. J. M. G. J. and K. M. L., Tuckerman M E, H. A., H, M. M., H, S. P. and M. M., Hansen J-P, M. I. R., Evans D J, M. G. P., D, T. B., Hess S, K. M. L. W. P. B. C. S. R. V. H. W. T., J, H. S. and E. D., F, M.-P., J-L, B. L. and B., Binder K, F. P., Chowdhury D, S. L. and S. A., Cetin N, N. K. R. B. and V. A., Amaral L A N, G. A. L. I. P. C. and S. H. E., Amaral L A N, G. A. L. I. P. C. and S. H. E., Mantegna R N, S. H. E., Herrmann

H J, H. J. P. L. S., Brandt A, B. J. B. K., Baschnagel J, B. K. D. P. G. A. A. H. O. K. K. M. W. L. M.-P. F. M. M. P. W. S. S. S. U. W. and T. V, P, B. S. and W., Mainville J, Y. Y. S. E. K. R. S. M. L. K. F. J. and S. G. B., Marro J, B. A. B. K. M. H. and L. J. L., Milchev A, H. D. W. and B. K., K, S. A. and B., Reister E, M. M. and B. K., Paul W, S. G. D. Y. D. Y. F. B. R. S. Z. A. W. L. and R. D., W, P., E, R. P., Doi M, E. S. F., L, V., M, S., C, A. H., Tuckerman M, M. G. J. and B. B. J., Lebowitz J L, P. J. K. and V. L., S, N., G, H. W., Bennemann C, P. W. B. K. and D. B., C, A. H., D, F., C, K. L. P. and M. P., K, B., W, H. J. and K., G, U., L, G. W. and S., van Beest B H W, K. G. J. and van S. R. A., Vollmayr K, K. W. and B. K., M, P. D. L. and C. S., C, M. J., Brébec G, S. R. S. C. B. J. and M. J. C., Scheidler P, K. W. L. A. H. J. and B. K., Horbach J, K. W. and B. K., Polian A, V.-T. D. and R. P., D, T. B., K, V. F. and B., K, H. J. and B., F, M.-P., Baschnagel J, B. C. P. W. and B. K., Binder K, B. J. and P. W., Landau L D, L. E. M., Barnes H A, H. J. H. W. K., J, T. B. D. and E. D., S, G. M., and R, K. (2004) Molecular dynamics simulations. *J. Phys. Condens. Matter 16*, S429–S453.

(22) Wang, W., Donini, O., Reyes, C. M., and Kollman, P. A. (2001) Biomolecular Simulations: Recent Developments in Force Fields, Simulations of Enzyme Catalysis, Protein-Ligand, Protein-Protein, and Protein-Nucleic Acid Noncovalent Interactions. *Annu. Rev. Biophys. Biomol. Struct. 30*, 211–243.

(23) Lin, H., and Truhlar, D. G. (2007) QM/MM: what have we learned, where are we, and where do we go from here? *Theor. Chem. Acc. 117*, 185–199.

(24) Garciavela, and Gerber, R. (1993) Hybrid quantum semiclassical wave packet method for molecular dynamics for photolysis of Ar . . . HCI. *J. Chem. Phys. 98*.

(25) García-Vela, A., and Gerber, R. B. (1993) Hybrid quantum/semiclassical wave packet method for molecular dynamics: Application to photolysis of Ar...HCl. *J. Chem. Phys. 98*, 427.

(26) Sauer, J., and Sierka, M. (2000) Combining quantum mechanics and interatomic potential functions inab initio studies of extended systems. *J. Comput. Chem. 21*, 1470–1493.

(27) Monard, G., Prat-Resina, X., González-Lafont, A., and Lluch, J. M. (2003) Determination of enzymatic reaction pathways using QM/MM methods. *Int. J. Quantum Chem. 93*, 229–244.

(28) Honarparvar, B., Govender, T., Maguire, G. E. M., Soliman, M. E. S., and Kruger, H. G. (2014) Integrated Approach to Structure-Based Enzymatic Drug Design: Molecular Modeling, Spectroscopy, and Experimental Bioactivity. *Chem. Rev. 114*, 493–537.

(29) Genheden, S., and Ryde, U. (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov. 10*, 449–61.

(30) Zheng, G., Xue, W., Wang, P., Yang, F., Li, B., Li, X., Li, Y., Yao, X., Zhu, F., Weintraub, D., Hajos, M., VanderWeide, L. A., Smith, S. M., Trinkley, K. E., Shao, L., Thapar, A., Cooper, M., Hannestad, J., Benson, N., Papakostas, G. I., Vu, A. T., Zhang, P., Millan, M. J., Yan, A., Krishnan, V., Nestler, E. J., Yang, H., Huang, J., Zhu, F., Yamashita, A., Singh, S. K., Krishnamurthy, H., Gouaux, E., Singh, S. K., Yamashita, A., Gouaux, E., Malinauskaite, L., Penmatsa, A., Wang, K. H., Gouaux, E., Penmatsa, A., Wang, K. H., Gouaux, E., Sorensen, L., Andersen, J., Andersen, J., Koldso, H., Grouleff, J., Schiott, B., Andersen, J., Zhou, Z. L., Severinsen, K., Sinning, S., Huang, X., Gu, H. H., Zhan, C. G., Forrest, L. R., Celik, L., Jorgensen, A. M., Gabrielsen, M., Schlessinger, A., Wang, P., Andersen, J., Koldso, H., Ahlqvist, M., Mallipeddi, P. L., Xue, W., Liu, H., Yao, X., Xue, W., Gabrielsen, M., Grouleff, J., Arnold, K., Laskowski, R. A., Jo, S., Jorgensen, W. L., Hornak, V., Dickson, C. J., Joung, I. S., Cheatham, T. E., Wang, J., Bayly, C. I., Wang, J., Hara, Y., Murayama, S., Springborg, M., Kirtman, B., Kollman, P. A., Massova, I., Kollman, P., Weiser, J., Shenkin, P. S., Still, W. C., Tippmann, S., Barer, M. R., Harwood, C. R., Letunic, I., Bork, P., Bymaster, F. P., Tatsumi, M., Kung, H. F., Wang, J., Genheden, S., Ryde, U., Hou, T., Reyes, C. M., Kollman, P. A., Andersen, J., and Wang, H. (2016) Exploring the Inhibitory Mechanism of Approved Selective Norepinephrine Reuptake Inhibitors and Reboxetine Enantiomers by Molecular Dynamics Study. *Sci. Rep. 6*, 26883.

(31) Roman Affentranger, †, Ivano Tavernelli, ‡ and, and Ernesto E. Di Iorio*, †. (2006) A Novel Hamiltonian Replica Exchange MD Protocol to Enhance Protein Conformational Space Sampling. *J. Chem. Theory Comput.2* (2),217–228

(32) Asim Okur, †, Lauren Wickstrom, ‡, Melinda Layten, §., Raphäel Geney, †, Kun Song, †, Viktor Hornak, ⦀ and, and Carlos Simmerling*, †,‡,⦀,. (2006) Improved Efficiency of Replica Exchange Simulations through Use of a Hybrid Explicit/Implicit Solvation Model. *J. Chem. Theory Comput. 420-423*

(33) Xiaolin Cheng, †, Guanglei Cui, †, Viktor Hornak, ‡ and, and Simmerling†‡*, C. (2005) Modified Replica Exchange Simulation Methods for Local Structure Refinement. *J Phys. Chem B.* 8220-8230

(34) Liu, P., Kim, B., Friesner, R. A., and Berne, B. J. (2005) Replica exchange with solute tempering: a method for sampling biological systems in explicit water. *Proc. Natl. Acad. Sci.*

U. S. A. 102, 13749–54.

(35) David, C. C., and Jacobs, D. J. (2014) Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins, in *Methods in molecular biology (Clifton, N.J.)*, pp 193–226.

(36) Wolf, A., and Kirschner, K. N. (2013) Principal component and clustering analysis on molecular dynamics data of the ribosomal L11·23S subdomain. *J. Mol. Model. 19*, 539–549.

(37) Zhu, X., Gerstein, M., and Snyder, M. (2007) Getting connected: analysis and principles of biological networks. *Genes Dev. 21*, 1010–24.

(38) Niroula, A., and Vihinen, M. (2015) Harmful somatic amino acid substitutions affect key pathways in cancers. *BMC Med. Genomics 8*, 53.

(39) Brender, J. R., and Zhang, Y. (2015) Predicting the Effect of Mutations on Protein-Protein Binding Interactions through Structure-Based Interface Profiles. *PLoS Comput. Biol. 11*, e1004494.

(40) Doncheva, N. T., Klein, K., Domingues, F. S., and Albrecht, M. (2011) Analyzing and visualizing residue networks of protein structures. *Trends Biochem. Sci. 36*, 179–182.

(41) Doncheva, N. T., Assenov, Y., Domingues, F. S., and Albrecht, M. (2012) Topological analysis and interactive visualization of biological networks and protein structures. *Nat. Protoc. 7*, 670–685.

(42) Hu, Y., Gupta-Ostermann, D., and Bajorath, J. (2014) Exploring compound promiscuity patterns and multi target activity space. *Comput. Struct. Biotechnol. J. 9*, 1–11.

(43) Hu, Y., Gupta-Ostermann, D., and Bajorath, J. (2014) Exploring Compound Promiscuity Patterns and Multi-Target Activity Spaces. *Comput. Struct. Biotechnol. J. 9*, 1–11.

(44) Lagunin, A., Stepanchikova, A., Filimonov, D., and Poroikov, V. (2000) PASS: prediction of activity spectra for biologically active substances. *Bioinformatics 16*, 747–8.

(45) Poroikov, V., Filimonov, D., Lagunin, A., Gloriozova, T., and Zakharov, A. (2007) PASS: identification of probable targets and mechanisms of toxicity. *Sar Qsar Environ. Res. 18*, 101–110.

CHAPTER 4

Re-emergence of an Orphan Therapeutic Target for the Treatment of Resistant Prostate Cancer – A thorough Conformational and Binding Analysis for ROR-γ Protein

Umar Ndagi[a] Ndumiso N. Mhlongo[a] and Mahmoud E. Soliman[a]*

[a] Molecular Modelling and Drug Design Research Group, School of Health Sciences, University of KwaZulu-Natal, Westville, Durban 4000, South Africa

*Corresponding author: Mahmoud E. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27(0)312608048, Fax: +27 (0)31260 7872

Webpage: http://soliman.ukzn.ac.za/

Abstract

Recent studies have linked a deadly form of prostate cancer known as metastatic castration-resistant prostate cancer (mCRPC) to retinoic acid-related orphan-receptor gamma (ROR-γ). Most of these studies continued to place ROR-γ as orphan because of unidentifiable inhibitor. Recently identified inhibitors of ROR-γ and their therapeutic potential were evaluated, among which inhibitor XY018 was the potent. However, molecular understanding of the conformational features of XY018-ROR-γ complex is still elusive. Herein, molecular dynamics (MD) simulations were conducted on HC9-ROR-γ and XY018-ROR-γ complexes to understand their conformational features at molecular level and the influence of XY018 binding on the dynamics of ROR-γ with the aid of post-dynamic analytical tools. These include; principal component analysis (PCA), radius of gyration (RoG), binding free energy calculation (MM/GBSA), per-residue fluctuation (RMSF) and hydrogen bond occupancy. Findings from this study revealed that (1) hydrophobic packing contributes significantly to binding free energy, (2) Ile136 and Leu60 exhibited high hydrogen-bond occupancy in XY018-ROR-γ and HC9-ROR-γ respectively, (3) XY018-ROR-γ displayed a relatively high loop region residue fluctuation compared to HC9-ROR-γ, (4) electrostatic interactions are a potential binding force in XY018-ROR-γ complex compared to HC9-ROR-γ, (5) XY018-ROR-γ assumes a rigid conformation which is highlighted by a decrease in residual fluctuation, (6) XY018 could potentially induce pseudoporphyria, nephritis, and interstitial nephritis but potentially safe in

renal failure. This study could serve as a base line for the design of new potential ROR-γ inhibitors.

Keywords: ROR-γ, Cancer, PCA, Binding free energy and MD.

1 Introduction

Nuclear receptors (NR) also known as nuclear hormone receptors,[1] are ligand based transcription factors that primarily act as a mediator in hormone transcriptional responses.[2] It directs the activities of distinct target genes by recruiting the host regulatory proteins.[2] Receptors of androgen, oestrogens, glucocorticoid, vitamin D, progesterone, thyroid and retinoic acid related orphan receptor are all forms of nuclear receptors.[1] They are characterised by DNA- binding domains (DBD) and ligand binding domain (LBD).[3] These receptors play a vital role in the array of biological processes such as reproduction, cell differentiation, proliferation and homeostasis, thereby resulting in inflammatory diseases and tumours.[1] Recent studies revealed that a deadly form of prostate cancer known as metastatic castration-resistant prostate cancer (mCRPC) is linked to retinoic-acid receptor-related orphan receptor gamma (ROR-γ).[4] This receptor drives androgen receptor expression by recruiting nuclear receptor coactivator 1 and 3 (SRC-1 and SRC-3) to an ROR response element (RORE) to stimulate androgen receptor (AR) gene transcription.[4] In a related development, a subtype of nuclear receptor, retinoic acid receptor-related gamma-t (ROR-γt) controls the inflammatory activities of T-helper 17 (Th17) cells,[1] and its expression is adequate enough to trigger transcriptional activation of a ROR-γt reporter.[1] The Th17 modulate immune responses, inflammatory and tumours by associating with CD4 cells (CD4+ T helper cells) through the secretion of certain cytokines IL-17A and IL-17F,[5] hence ROR-γt is important for Th17 cell differentiation in response to cytokines.[6]

However, ROR-γ varies from other NR such as ROR-α and ROR-β with a specific mode of expression in the tissues and variable physiological activities. T cells express ROR-β transcript variant by using a unique promoter which yields an isoform referred to as ROR-γt with structural variation from ROR-γ in its N-terminus.[4] Thus, ROR-γ is distinguished from other NR for its unique structural features and physiological functions. Its ability to drive the expression of androgen receptor in castration-resistant prostate cancer (CRPC) represents a milestone in the search for a therapeutic target in CRPC.[4]

In recognition of the role of ROR-γ in prostate cancer, recent studies examined the relationship between ROR-γ and prostate cancer.[4] Results from immunoblotting analysis of ROR-γ protein in cancer and noncancerous cell lines revealed that appreciable level of inhibition of growth of androgen-sensitive human prostate adenocarcinoma cells (LNcap) and their CRPC-derivatives C4-2B cells was achieved when ROR-γ was knocked down by different RORC siRNAs.[4] Such a high level of growth inhibition was observed in androgen-sensitive and CRPC model,[4] as the ROR-γ is knocked down there is resultant induction of apoptosis, which is marked by activation of caspase-3 and caspase-7 and the breakdown of poly (ADP-ribose) polymerase 1(PARP1), resulting into reduced expression of the protein relevant to oncogenesis, proliferation and survival.[4] Most recent studies continued to place ROR-γ as orphan because of unidentifiable ligand (inhibitor) suitable for its binding site.[7] A most recent study identified three different inhibitors of ROR-γ.[4] These inhibitors were labelled as XY018, SR2211 and XY011 based on their structural differences and effect on the cell line.[4] In this regard, the first two inhibitors demonstrated sufficient level of suppression on the expression of important proliferation and survival proteins.[4]

In a related development, ROR-γ was targeted with the same antagonist XY018, SR2211 and XY011 which effectively subdue the messenger RNA (mRNA) and protein expression of complete AR dimension.[4] This results in inhibition of androgen receptor (AR) variants, including AR-V7, C4-2B and VCap cells[4] (cancer cell line). In addition, it was noted that the inhibition of AR and its variant is dose-dependent at the level of mRNA and protein[4]. To assess the effect of ROR-γ on AR, the C4-2B cells were treated with 5 μM of SR2211 for 24hrs,[4] and the results revealed a partial inhibition of AR expression.[4] This is an indication that SR2211 is not very effective in the suppression of AR expression. A study on ROR-γ antagonist in mice revealed that there is sufficient level of inhibition in the growth of AR-positive xenograft tumours including those with AR gene amplification and high-level AR variants,[4] implying that targeting of ROR-γ would form the basis for exploring its clinical importance in prostate cancer.[4] It is also interesting to note that the expression of AR and its variant is strongly subdued when ROR-γ is inhibited, and consequently elimination of AR genome binding.[4]

In all of these studies, the structural properties of retinoic acid receptor-related orphan receptor gamma in complex with XY018 still remains elusive. None of these studies demonstrate the conformational features and ligand binding of ROR-γ with its inhibitor such as XY018. Thus the need for the in-depth conformational analysis of XY018-ROR-γ complex is paramount to

the understanding of the role XY018 as an inhibitor in metastatic castration-resistant prostate cancer (mCRPC).

The mCRPC is a global phenomenon[8] and source of concern because even in the presence of the most effective anti-androgen drugs (AR-pathway targeted therapy)[9] there is evidence of disease progression in some patient.[9] These drugs are becoming completely ineffective in the treatment of mCRPC[10] and may not form primary agent in the future treatment, should a scourge of mCRPC emerge in the absence of known inhibitors to ROR-γ. Previous studies have attributed the absence of good clinical outcome to some anti-androgens to lack of understanding of the role ROR-γ in CRPC.[4] An attempt was made in a study to elucidate the complex structure of ROR-γ with natural ligand (hydroxycholesterol),[7] whereby, the binding domain of the natural ligand and the role of pocket residues in ligand binding[7] was validated. Hydroxycholesterols, the derivatives of cholesterol, are endogenous ligands that have high affinity for RORα.[7] They actively promote coactivator recruitment by ROR-γ and are said to be agonists of the same receptor.[7] They promote ROR-γ – Coactivator interactions via a conserved charge clamp[7] and are known to be potent ROR-γ agonists.[7] Similarly, hydroxycholesterols modulate RORγ-dependent biological processes.[7] This observation has attracted a lot of interest in the quest to identify suitable inhibitors for orphan receptor. Therefore, a good understanding of conformation features and ligand binding landscape of ROR-γ is crucial to the development of potential and effective inhibitors of ROR-γ.

Attempts to elucidate the molecular basis of ROR-γ have been previously carried out.[1,4,7,11] However, findings from such studies suggest that for a good understanding of the dynamics of ROR-γ in the presence of XY018, a long time scale molecular dynamics (MD) simulations of the XY018-ROR-γ complex are required to provide an atomistic insight to conformational features and ligand-binding landscape of this complex.[12] **Figure 1** represent structure of ROR-γ inhibitors investigated experimentally and hydroxycholesterol.

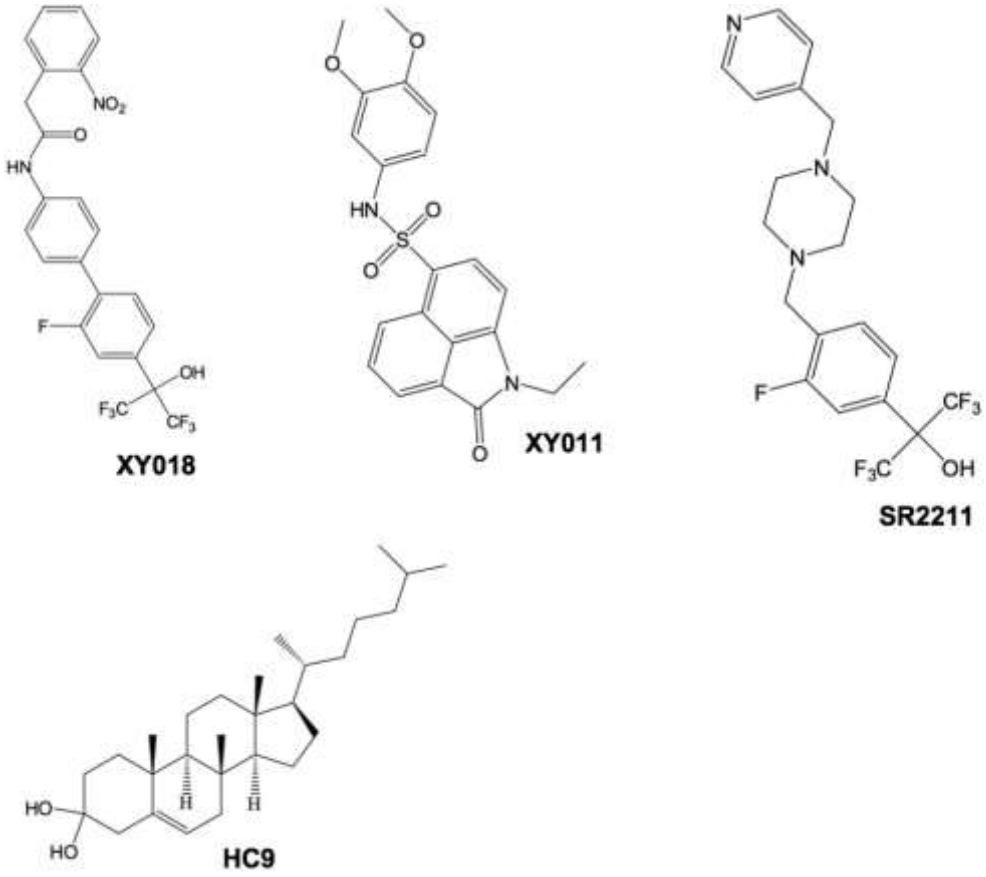


Figure 1. 2D structures of experimental ROR-γ inhibitors and natural ligand (XY018, XY011, SR2211 and HC9).

Numerous post-dynamics techniques have been used to provide molecular understanding of molecular dynamics. The principal component analysis (PCA), also known as essential dynamics analysis, is one of the most popular post-dynamics techniques[13] that is widely used to understand the dynamics of biological systems.[13] Principal component analysis eliminates a wide-range of translational and rotational motions in MD trajectory and correlated motions in atomic simulations of proteins.[13,14] It is an important technique that defines the atomic displacement in a collective manner,[15,16] it can detect major conformational dynamics between the structures[15] and has been used in many studies to determine the difference in motion of protein complexes of different compounds. In this study, we aim to provide a better understanding of conformational features and ligand binding landscape of ROR-γ in complex with XY018 and contributing to the value of the experimental work that has been previously conducted. To achieve this, MD simulations of ROR-γ in complex with a natural ligand (Hydroxycholesterol) and ROR-γ in complex with XY018 were performed to further inspect the effect of XY018 binding on the dynamics of ROR-γ. To prepare for the process of MD simulations, hydroxycholesterol was first docked into ROR-γ. Three other known inhibitors

(XY018, SR2211 and XY011) **(see Figure 1)** were also docked into ROR-γ and the inhibitor with the best docking score was used for MD simulation. A 100ns MD simulations were conducted, followed by binding free energy calculations and PCA to understand the effect of XY018 binding to the dynamic state of the subject protein.[17] Such tools are known to enhance the process of drug discovery.[18] Herein, conformational and structural properties of ROR-γ in complex with XY018 are demonstrated. Such properties may form the basis for which other therapeutics targeting the ROR-γ could be developed. **Figure 2** shows a complex of ROR-γ and hydroxycholesterol (HC9).

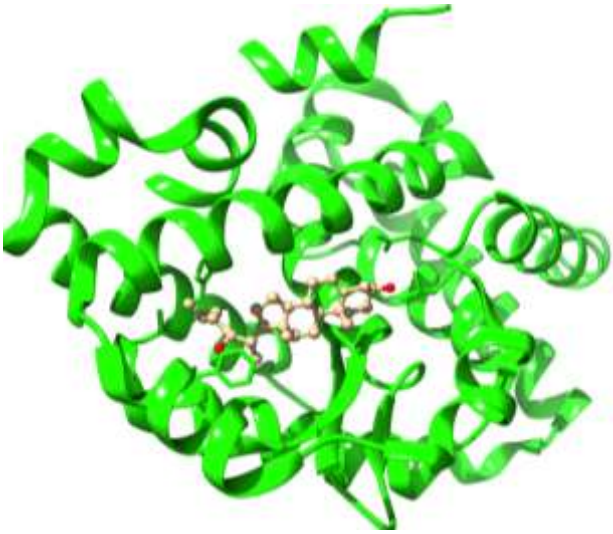


Figure 2. 3D structure of ROR-γ showing the binding position XY018 and hydroxycholesterol studied in this work.

1 Computational methods

1.1 System preparation and molecular docking

The steepest descent method and MMFF94 force field in Avogadro[19] software were used to optimize XY018, SR2211 and XY011 for energy minimization. Hydrogen atoms were deleted using UCSF chimera,[20] in preparation for docking. The X-ray crystal structure of ROR-γ in complex with hydroxycholesterol was obtained from protein data bank (PDB) with code (3L0J),[7] and prepared in UCSF chimera.[20] Water molecules were deleted, hydroxycholesterol was separated from the protein structure, with addition of hydrogen atoms to the protein. **Figure 3** shows the workflow adopted in this study.

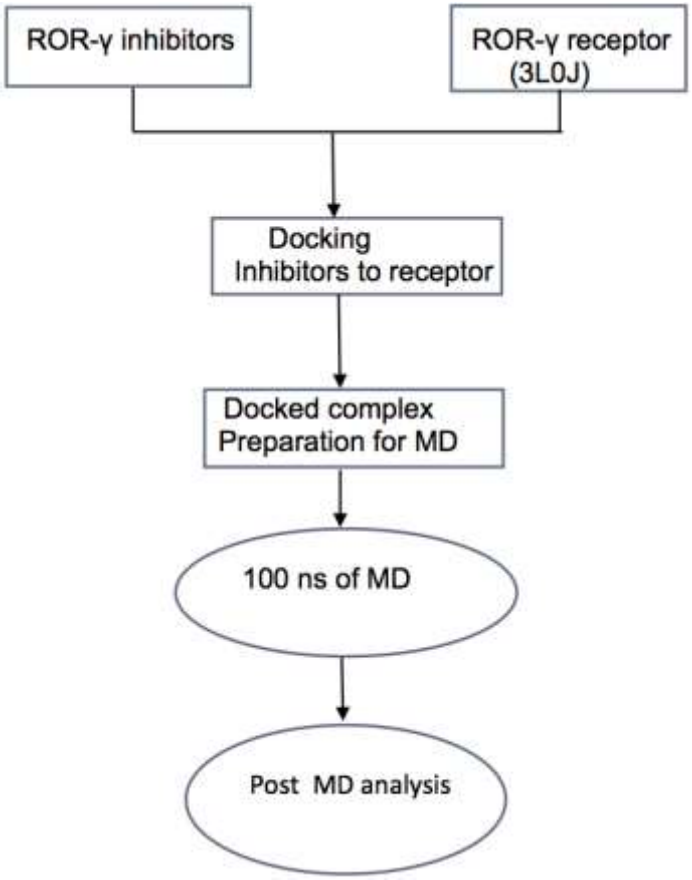


Figure 3. Schematic representation of MD workflow used in the current study.

1.2 Molecular docking

For docking calculations, AutoDock Vina was used.[21] Gasteiger partial charges[21] were allocated during docking. AutoDock Graphical user interface provided by MGL tools was used to outline the AutoDock atom types.[22] The grid box was determined with grid parameters being x = 30 Å, y = 36 Å and z = 36 Å for the dimension while x = -23.546 Å, y = -4.134 Å and z = -24.05 Å for the centre grid and exhaustiveness =8, covering the entire residence of hydroxycholesterol which is at the active site of the receptor[7]. The Lamarckian genetic algorithm[23] was used to generate docked conformations in accordance with their docking score (DS) in a descending order.

1.3 System preparation for MD

The docked complex of XY018-ROR-γ and HC9-ROR-γ were used for simulations in this study. These complexes were visualised using UCSF chimera,[20] ligand and receptor were modified with aid of UCSF chimera[20] and Avogadro[19] software.

1.4 Molecular dynamics simulations

Simulations of XY018-ROR-γ and HC9-ROR-γ system were performed using graphics processor unit (GPU) version of Particle Mesh Ewald Molecular Dynamics (PMEMD) package with Sander module of Amber14.[24,25] The Amber force field FF12SB[26, 27] was applied to describe the protein.[24] The ligand parameters were set using Gasteiger charges in Avogadro,[19] and Antechamber module with the aid of GAFF (generalised Amber force field)[28] The LEaP module implemented in Amber14[24] was used to add hydrogen atoms to the protein and to add counter ions for system neutralization.[24] Each system is enclosed in the TIP3P water box[29] with the protein atoms located 10 Å between the protein surface and the box boundary. The cubic periodic boundary conditions were implemented in all the systems. Long-range electrostatic interactions were treated with particle-mesh Ewald method[30] with a nonbonding cut-off distance of 12 Å. Two minimization steps were adopted, partial minimization and full minimization. The initial energy minimization step of the systems were carried out with a restraint potential of 500 kcal mol^{-1} $Å^{-2}$ applied to the solute, for 1000 steps. Unrestrained conjugate gradient minimization for 1000 steps was conducted for the entire system, with the aid of SANDER module of Amber 14 program. A canonical ensemble (NVT) MD simulations were performed for 50 ps and the system was gradually heated from 0 to 300 K, with harmonic restraints of 5 kcal mol^{-1} $Å^{-2}$ for solute atoms with the aid of Langevin thermostat[31] with a 1ps random collision frequency. The systems were equilibrated at 300 K with a 2 fs time scale in NPT ensemble for 500 ps without any restraint. The Berendsen barostat[32] was used to maintain the pressure at 1 bar. The SHAKE[33] algorithm was used to constrain the bonds of hydrogen atoms in the system. In the absence of restraints, a production run of 100 ns MD was conducted in an isothermal-isobaric (NPT) ensemble using a Berendsen barostat at a pressure of 1 bar and a 2 ps pressure coupling constant. For every 1ps time interval, the coordinates were saved and the trajectories were analysed. Post-MD analysis performed include root mean square fluctuation (RMSF), root mean square deviation (RMSD), the radius of gyration (RoG), hydrogen-bond formation and principal component analysis (PCA) using CPPTRAJ modules in Amber 14, as well as ligand-residue interaction profile. Visualization of trajectories was conducted in Chimera.[20] The results were analysed and plots were generated using Origin[34] and Bio3D[35] software.

2 Post-dynamic analysis

There are several methods applied in analysing protein dynamics, depending on the desired information needed to describe changes that occur within the structure of the protein during simulations.

2.1 Thermodynamic calculations

The binding free energy calculation is an important end point method that gives detailed information on the interaction between the ligand and protein.[36] It provides insight into mechanism of binding, including both enthalpic and enthropic contributions to the molecular recognition.[37] Molecular Mechanics/Generalized-Born Surface Area (MM/GBSA) methods[36] are popular approaches to estimate the free energy of the binding of small ligands to biological macromolecules.[36] They are based on MD simulations of receptor-ligand complex and are therefore intermediate in both accuracy and computational effort between the empirical scoring and alchemical perturbation method.[36] They have been successfully applied to a large number of systems,[36] thus can be used in the present study. For a 100ns trajectory, 1000 snapshots were considered during the calculation of binding free energy, the following set of equations describes the binding free energy calculation:

$$\Delta G_{bind} = G_{complex} - G_{receptor} - G_{ligand} \quad (1)$$

$$\Delta G_{bind} = E_{gas} + G_{sol} - TS \quad (2)$$

$$E_{gas} = E_{int} + E_{vdW} + E_{ele} \quad (3)$$

$$G_{sol} = G_{GB} + G_{SA} \quad (4)$$

From the equation above, E_{gas} is the energy of the gas phase, E_{int} represents internal energy, E_{ele} represents coulomb while E_{vdW} are the van der Waals energies. E_{gas} is estimated directly from the ff12SB[26] force field. G_{sol} which is the salvation free energy can be broken down to polar and non-polar forms contribution. The contribution of polar solvation (G_{GB}) is assessed by resolving G_{GB} equation and non-polar solvation (G_{SA}) is determined from the solvent accessible surface area. This can be estimated from water probe radius of 1.4 Å with temperature (T) and total solute entropy (S). The MM/GBSA binding free energy method in Amber 14 was used to calculate the contribution of each residue to the binding free energy between the XY018, HC9 and the receptor (ROR-γ). In addition, the interaction energy decomposition analysis per residue was also computed using the same method.

2.2 Principal component analysis

The essential dynamics analysis[38] also known as principal component analysis (PCA) is an important method that describes the dynamic nature of proteins. The atomic displacement and conformational changes of a protein can be defined by PCA by extracting different modes of the conformation of the protein complex during dynamic simulations. The eigenvectors (the direction of motion) and the eigenvalues (extent of motion) for the biological system can also be determined using PCA.[39] In this study, the trajectories of the complexes from 100 ns MD were stripped of the solvent molecules and the ions using the CPPTRAJ module in Amber 14.[24] This was conducted prior to MD trajectory processing for PCA. PCA was performed on C-α atoms on 1000 snapshots at 100 ps time interval using in-house script. The first two principal components (PC1 and PC2) were computed and a 2X 2 covariance matrices were generated using Cartesian coordinates of Cα atoms. PC1 and PC2 correspond to first two eigenvectors of a covariant matrices. Origin software[34] was used to construct PCA plots.

3 Result and discussion

3.1 Docking and validation

Docking of a ligand to the protein active site is one of the commonly used methods in molecular modelling to determine the preferred orientation of one molecule (ligand) to another (receptor) when bound together. However, in some instances results from docking studies could be unreliable. This is because even the best-docked conformation may drift away from the active site of the protein within a short time interval of MD simulations. Therefore, to improve the reliability of docking results, we embarked on MD simulations to guarantee that the docked complexes remain stable in the active site of the protein within a specific time scale. In line with this, four ligands (SR2211, XY018, XY011 and hydroxycholesterol) were docked into ROR-γ (PDB 3L0J)[7] the following docking scores were recorded:

Table 1. Docking scores of the compounds on ROR-γ

Serial Number	Compound	Docking scores
1	HC9	-11.1
2	XY018	-10.7
3	SR2211	-10.1
4	XY011	-9.9

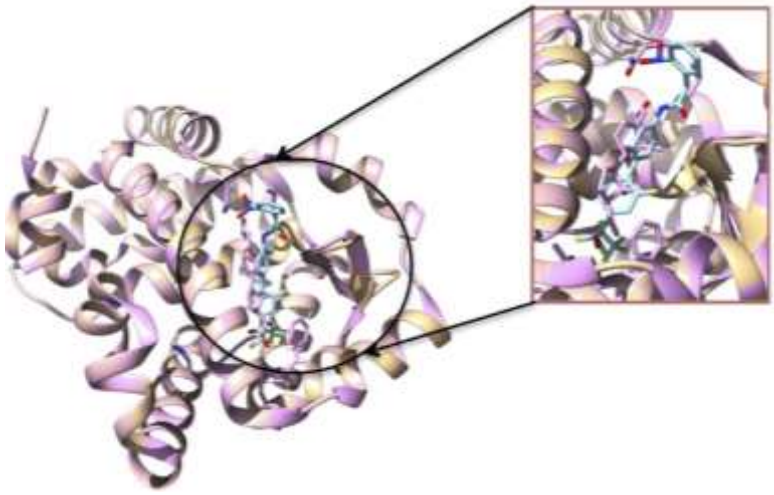


Figure 4. A 3D depiction of docked XY018-RORγ and HC9-RORγ superimposed to validate docking.

A docked complex of ROR-γ with XY018 and ROR-γ with HC9 were superimposed (**Figure 4**) to validate the orientation of the docked ligands on the active site of the receptor. The docked complexes were further subjected to 100 ns MD simulations. It is important to note that due to unavailability of the XY018-ROR-γ complex in the protein data bank (PDB), we used crystal structure of ROR-γ complexed with a natural ligand (hydroxycholesterol) with PDB code 3L0J.[7]

3.2 System stability and MD simulations

Prior to MD trajectory analysis, root mean square deviations (RMSD) and potential energy fluctuations were monitored throughout the 100 ns of MD simulations. This was to guarantee the systems stability through-out a simulation. RMSD was calculated to assess the stability and convergence of the respective systems and the results are presented **Figure 5.** Systems stabilisation and convergence with fluctuations between 0-10000 ps, 15000-20000 ps and 40000-60000 ps (maximum RSMD fluctuation of 1.9 Å) were observed. However, after approximately 60000 ps, the systems converged and fluctuations rested below 1.6 Å for both systems throughout the simulation. The average RMSD of XY018-RORγ system is 1.08 Å

while HC9-RORγ has an average RMSD of 1.00 Å. These account for system stability since a standard parameter defining a stable system is an average RMSD of 2 Angstrom and below.[40] These results show that although slightly, RORγ-XY018 exhibit more flexibility and deviation compared to RORγ-HC9 within the fluctuation regions and claims stability after 60000 ps during MD simulations.

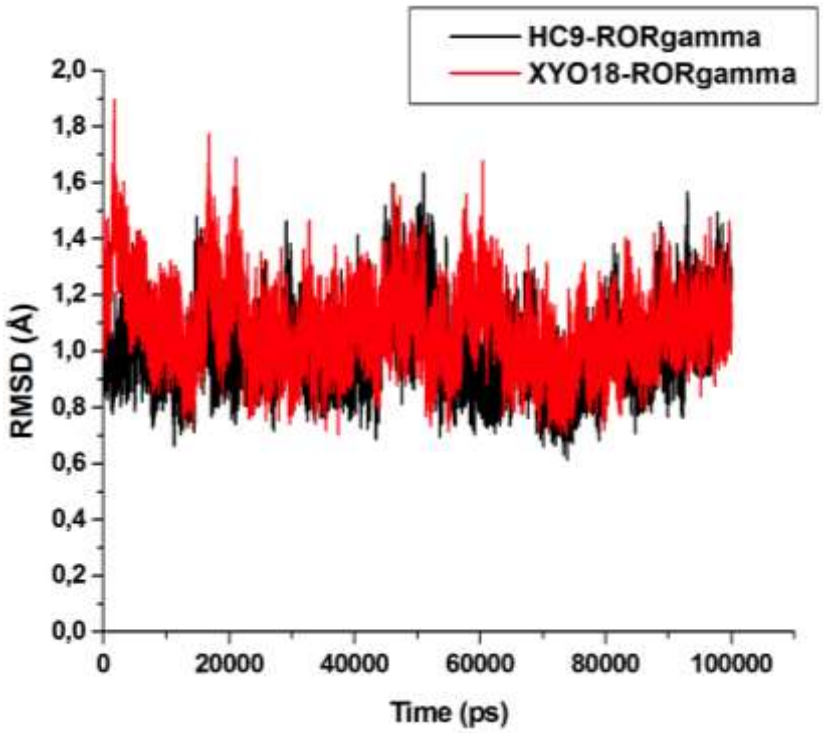


Figure 5. A comparative RMSD plot of XY018-RORγ (red) and HC9-RORγ (black) systems.

3.3 Post MD analysis

3.3.1 Root mean square fluctuation (RMSF)

A protein is made up of a specific sequence of amino acids[41] that allow it to assume a particular conformation.[41] Changes to the protein conformation occur when there is chemical action or mechanical events.[41] Therefore, direct interactions with the protein active site can alter its function. More specifically, the conformational changes that occur as a result of ligand-induced motion during ligand binding.[42] Understanding of ligand-induced conformational changes in the protein is critical to structure-based rational drug design.[43] RMSF is a measure of average atomic mobility of backbone atoms (N, Cα and C) during MD simulation. To understand and explore the structural dynamics (fluctuations) that take place upon the ligand binding, RMSF of both complexes (XY018-RORγ and HC9-RORγ) was calculated.

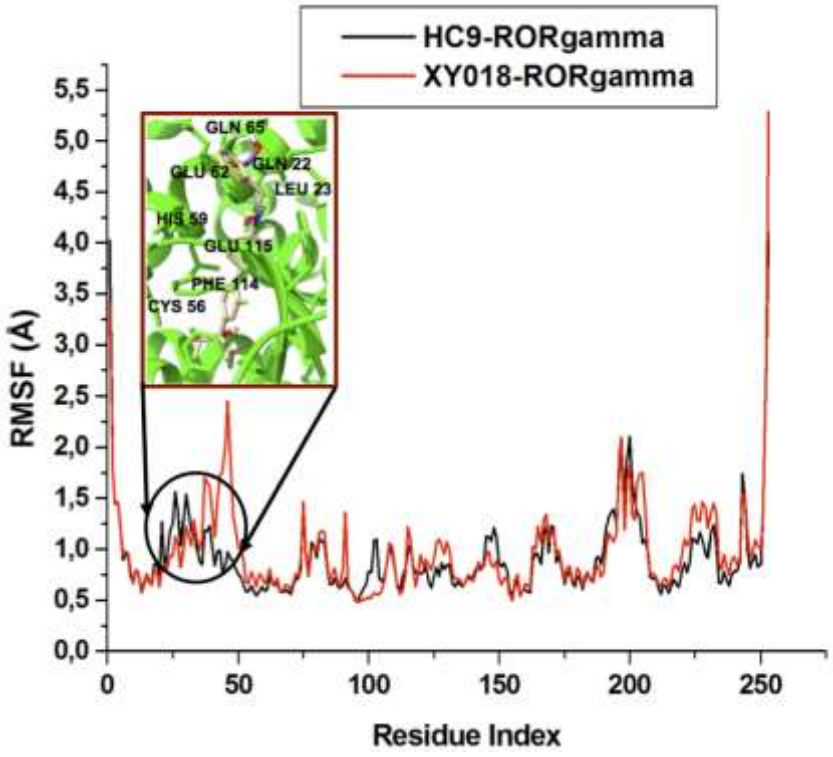


Figure 6. A comparative RMSF plots of RORγ-XY018 (red) and RORγ-HC9 (black) systems

The core of the protein appears to be more rigid compared to the loops (solvent exposed) as shown by RMSF fluctuation in **Figure 6**. Comparatively, the average RMSF value for the two systems does not show significant differences. A lower average RMSF value of 0.92 Å for HC9-RORγ compared to 0.98 Å for XY018-RORγ. This indicate a firmer structure and decreased capacity to fluctuate in HC9-RORγ system and a relatively increased capacity to fluctuate in XY018-RORγ system. The highest residue fluctuation is observed in XY018-RORγ system, particularly in the loop region which comprises residues Asn34, Ile35, Ser37, Lys47 and Ser48 with an RMSF value of 2.5 Å. This is could be due to flexibility of loops that open and accommodate the ligand. The presence of the ligand in the binding region and formation of ligand–residue interaction stabilizes the complex. Subsequent to complex stability, a decrease in flexibility and capacity of hydrophobic active site residues (Gln22, Leu23, His59, Leu60, Glu62, Phe113, Phe114, Glu115, Phe124, Ile136) in a XY018-RORγ system to fluctuate is observed. The binding of XY018 is likely to be responsible for the biochemical changes in the receptor (ROR-γ) [4] owing to the conformational dynamics induced on residues beneath the active site, relative to the binding of the natural ligand (HC9) where the fluctuation is minimal. However, between Gln22, Leu23, Cys56, His59, Glu62 and Gln65, there is notable rigidity. This may be due to the presence of a phenyl group that is interacting

with these residues. This further explains the role of the phenyl group in the inhibitory activity of XY018.

Similarly, residues lying between 99 and 125 (Val99, Met101 and Phe124) with an RMSF value of 0.75 Å do not exhibit noticeable fluctuation, the reason being the presence fluoride groups which contributed to the conformational changes that accompany the binding of XY018 to the ROR-γ.

3.3.2 Radius of gyration (RoG)

The radius of gyration is the distribution of component of an object around the axis,[44] it has been used to gain an insight into molecular stability in the biological system during MD simulations.

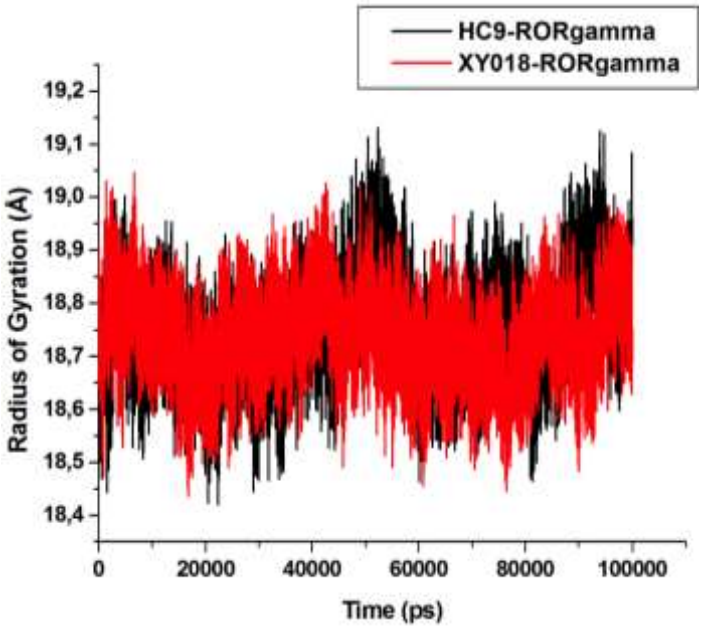


Figure 7. RoG plot of Cα atoms of the XY018-RORγ (red) and HC9-RORγ (black).

In this study, RoG of XY018-RORγ and HC9-RORγ systems were examined. The XY018-RORγ system shows a lower average RoG value of 18.73 Å compared to 18.75 Å of HC9-ROR system **Figure 7**. This observation can be used to draw a correlation between the two systems with relatively no significant difference in their RoG. However, slight increase in RoG is observed between 50000-60000 ps and 90000-95000 ps of MD simulation for HC9-RORγ, which reflects changes associated with the protein expansion within this period, thereby creating windows of opportunity for the solvent molecules to infiltrate the hydrophobic sites. This results in elevated hydrophobicity of the protein surface thus allowing the ease of residue fluctuation around these sites.

3.3.3 Binding free energy and energy decomposition analysis

Molecular Mechanics/Generalized-Born Surface Area (MM/GBSA) methods[36] is a popular approach use to estimate the free energy of the binding of small ligands to biological macromolecules.[36] To assess the various energy contribution to the binding of XY018 and HC9 to RORγ, post-dynamic calculations of binding free energy using MM/GBSA method were performed for the studied systems.

Table 2. MM/GBSA-based binding free energy profile of RORγ-HC9 and RORγ-XY018.

Complex	ΔG_{bind}	ΔE_{ele}	ΔE_{vdW}	ΔG_{gas}	ΔG_{sol}
HC9-RORγ	-57.21±0.32	-12.02±0.32	-59.81±0.26	-71.84±0.36	14.63±0.19
XY018-RORγ	-56.66±0.45	-43.98±1.30	-54.41±0.51	-98.40±1.02	41.73±1.00

Notes: ΔE_{ele}, electrostatic energy; ΔEvdW, van der Waals energy; $G_{binding}$, predicted total binding free energy; sol, solvation energy.

Energy decomposition analysis (**Table 2**) revealed that the estimated binding free energy is higher in HC9-RORγ (-57.2 kcal/mol) compared to XY018-RORγ (-56.66 kcal/mol). The difference in binding affinity (-0.55 kcal/mol) between the two systems is not significant enough to affect drug binding rather shows that HC9 could be better binded to receptor than the drug (XY018). However, the van der Waals contribution to the total binding free energy is higher in HC9-RORγ (-59.81 kcal/mol) with a lower electrostatic energy contribution (-12.02 kcal/mol). The electrostatic energy contribution (-43.98 kcal/mol) from XY018-RORγ to the total binding free energy is more than threefold the total electrostatic energy (-12.02 kcal/mol) contribution from HC9-RORγ to the total binding free energy of the system. Based on these results, electrostatic interactions are the potentially important binding forces between XY018 and ROR-γ. Hydrophobic packing contributes significantly to binding free energy owing to large amount of aromatic and hydrophobic rings within the conformational space, as well as set of hydrophobic residues around the binding pocket. This is evident in the binding free energy contribution, where vdW contribution is relatively higher than the electrostatic contribution as shown in **Table 2**.

3.3.4 Per-residue energy decomposition analysis

In order to assess the energy contribution of individual active site residues to the total binding free energy, and to provide a more detailed understanding of the impact of protein dynamics on the degree of different binding forces. The binding free energy was decomposed into the unit contributions of each active site residue of XY018-RORγ and HC9-RORγ. **Table 3** shows that the major contributors in the XY018-RORγ system were His59, Gln22, Phe114, Phe113,

Leu60 and Leu23 with energy contributions of -2.195, -2.009, -1.959, -1.907, -1.886 and -1.088 kcal/mol respectively. While in HC9-RORγ system the major contributors were Leu60, Met101, Phe124, Ile133, Ala63 and Cys56 with energy contribution of -2.810, -2.244, -1.894, -1.364, -1.353 and -1.264 kcal/mol respectively. There is a variation in per residue energy contribution of the active site residues from the two system. The active site residues in XY018-RORγ system, exhibit relatively lower individual energy contribution to the total binding free energy compared to HC9-RORγ. This may be attributed to the steric effect of XY018 and hydrophobic active site residues. In XY018-RORγ system it was observed that residue Glu62 contributes more to the electrostatic interaction (-5.504 kcal/mol) and has the least van der Waals energy (-0.404 kcal/mol). Similarly, residue Gln22 and Glu115 were also associated with low van der Waal energy (-0.929 and -0.702 kcal/mol respectively) and relatively high electrostatic contribution (-3.008 and -1.721 kcal/mol respectively) to the total binding free energy (**Figure 8 and 9**). The electrostatic contribution is associated with the formation of a hydrogen bond between a nitrogen atom of Glu115 and oxygen atom XY018 (**Figure 12**). However, in the case of residue Ile136, Phe124 and Leu23 electrostatic contributions of -0.135, -0.063 and -0.059 kcal/mol were observed respectively. These residues contribute the least electrostatic energy in the XY018-RORγ system with relatively low van der Waal energy contribution (-0.635, -0.906 and -1.019 kcal/mol).

Table 3. Decomposition of the relative binding free energies on a per-residue basis for XY018-ROR-γ and HC9-ROR-γ

Residues	ΔE_{vdw}	Δ_{ele}	ΔG_{polar}	$\Delta G_{non\text{-}polar}$	$\Delta G_{binding}$
XY018-RORγ					
Gln22	-0.929±0.616	-3.008±1.280	2.090±0.648	-0.163±0.040	-2.009±0.650
Leu23	-1.019±0.392	-0.059±0.101	0.022±0.103	-0.150±0.62	-1.088±0.416
His59	-2.602±0.651	-2.771±1.580	3.487±1.205	-0.308±0.052	-2.195±1.178
Leu60	-1.818±0.348	-0.214±0.134	0.334±0.132	-0.187±0.043	-1.886±0.345
Glu62	-0.404±0.816	-5.504±7.215	5.102±5.980	-0.097±0.051	-0.903±1.084
Phe113	-1.144±0.461	-2.940±1.806	2.319±1.003	-0.142±0.037	-1.907±0.880
Phe114	-1.677±0.408	-0.927±0.910	0.726±0.446	-0.081±0.051	-1.959±0.849
Glu115	-0.702±0.440	-1.721±0.872	1.948±0.905	-0.095±0.073	-0.570±0.647
Phe124	-0.906±0.316	0.063±0.142	0.136±0.128	-0.100±0.046	-0.808±0.325
Ile136	-0.635±0.227	-0.135±0.122	0.145±0.138	-0.070±0.029	-0.695±0.240
HC9-RORγ					
Trp53	-0.866±0.282	-0.083±0.040	0.786±0.187	-0.069±0.024	-0.233±0.226
Cys56	-1.612±0.371	-0.184±0.080	0.652±0.177	-0.120±0.032	-1.264±0.315
His59	-1.357±0.318	-0.090±0.113	0.846±0.492	-0.145±0.033	-0.747±0.407
Leu60	-2.712±0.397	-0.015±0.269	0.234±0.056	-0.316±0.184	-2.810±0.316
Ala63	-1.028±0.114	0.054±0.097	0.235±0.467	-0.143±0.074	-1.353±0.412
Met94	-0.400±0.141	-0.037±0.023	0.128±0.064	-0.047±0.024	-0.357±0.140
Val97	-1.164±0.134	-0.062±0.312	0.057±0.238	-0.097±0.012	-1.266±0.251
Met101	-2.274±0.378	0.067±0.065	0.169±0.195	-0.204±0.031	-2.244±0.391
Val112	-1.216±0.067	0.011±0.109	-0.346±0.231	-0.090±0.074	-1.641±0.431
Phe114	-1.365±0.337	-0.004±0.039	0.645±0.172	-0.127±0.030	-0.842±0.303
Phe124	-1.987±0.295	0.785±0.391	1.005±0.193	-0.127±0.021	-1.894±0.459
Ile133	-1.235±0.292	-0.176±0.093	0.168±0.090	-0.120±0.029	-1.364±0.314
His215	-0.743±0.206	0.029±0.044	0.722±0.346	-0.075±0.027	-0.068±0.373

Notes: ΔE_{ele}, electrostatic energy (kcal/mol); $\Delta E_{vdW,}$ van der Waals energy (kcal/mol); ΔG_{polar}, polar solvation energy (kcal/mol); $\Delta G_{nonpolar}$, nonpolar solvation energy (kcal/mol); $\Delta G_{binding}$ (kcal/mol), total binding free energy

From this analysis, it can deduced that electrostatic energy is pivotal for the binding of residues in XY018-ROR-γ system compared to HC9- ROR-γ. The result from this analysis will provide a guide to the design of potential ROR-γ inhibitors keeping in mind the electrostatic contribution of an individual residue to the total binding free energy.

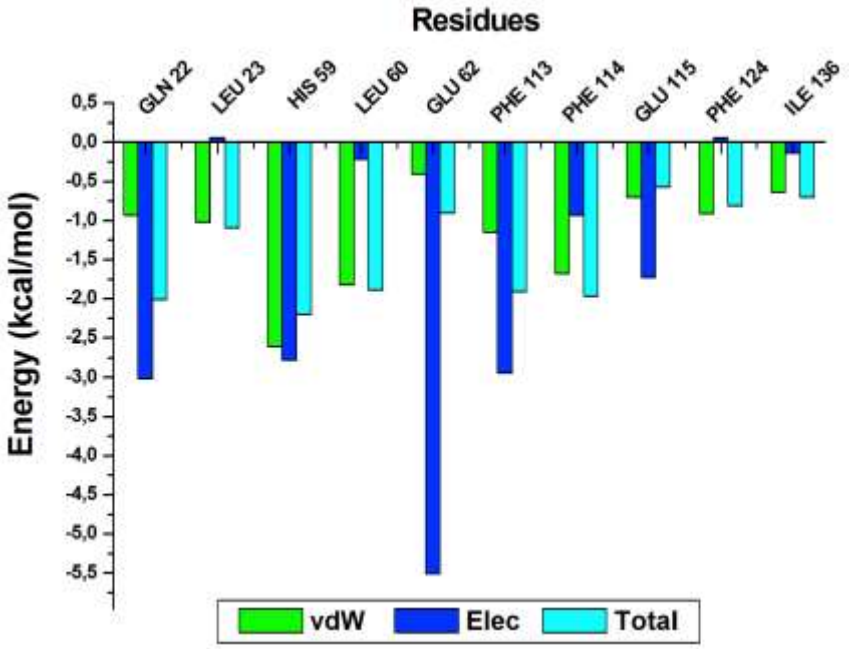


Figure 8. The per-residue decomposition analysis graph of XY018-RORγ.

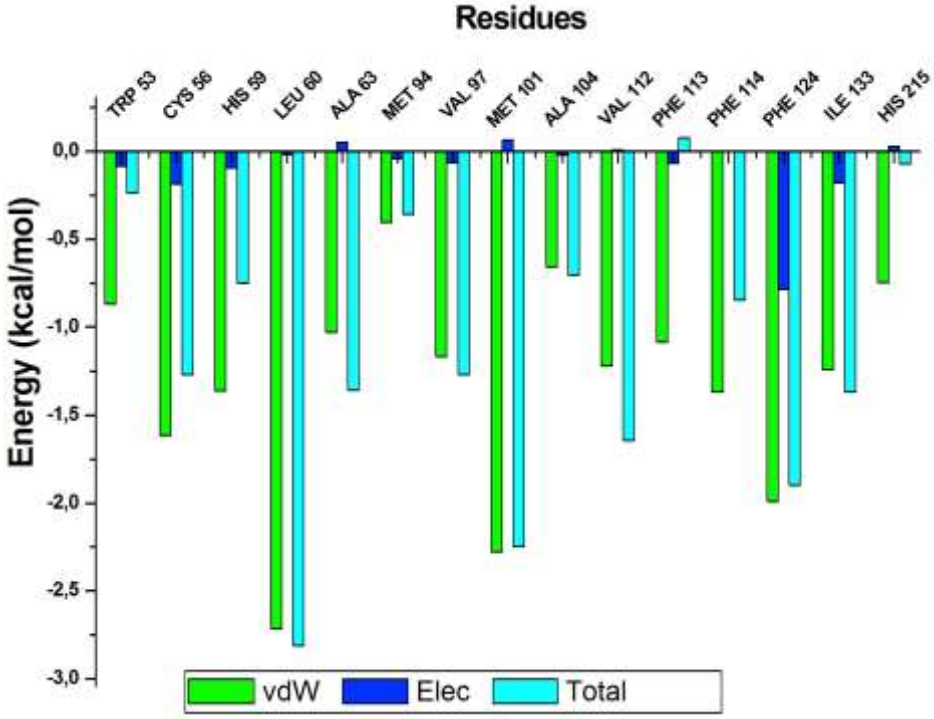


Figure 9. The per-residue decomposition analysis graph of HC9-RORγ.

3.3.5 Principal component analysis (PCA)

A key feature which defines biological function is protein conformation.[45] PCA is one of the principal tools used in determining the behaviour of each atom during a simulation.[46] Here we adopt a clustering method of PC because of its ability to describe different conformational states sampled during a simulation by grouping molecular structure into a subset based on their

conformational similarities.[17] This method of PCA was used to assess the flexibility of ROR-γ during MD simulations. In order to assess conformational motions of XY018-ROR-γ and HC9-ROR-γ. The two systems were projected along the first two principal components or eigenvectors (PC1 vs PC2) direction. The percentage variability or total mean square displacement of atom's positional fluctuation captured in each dimension is characterised by their corresponding eigenvalue.[47]. **Figure 10** shows a plot of XY018-ROR-γ and HC9-ROR-γ waves of conformation in important subspace along the three principal components.

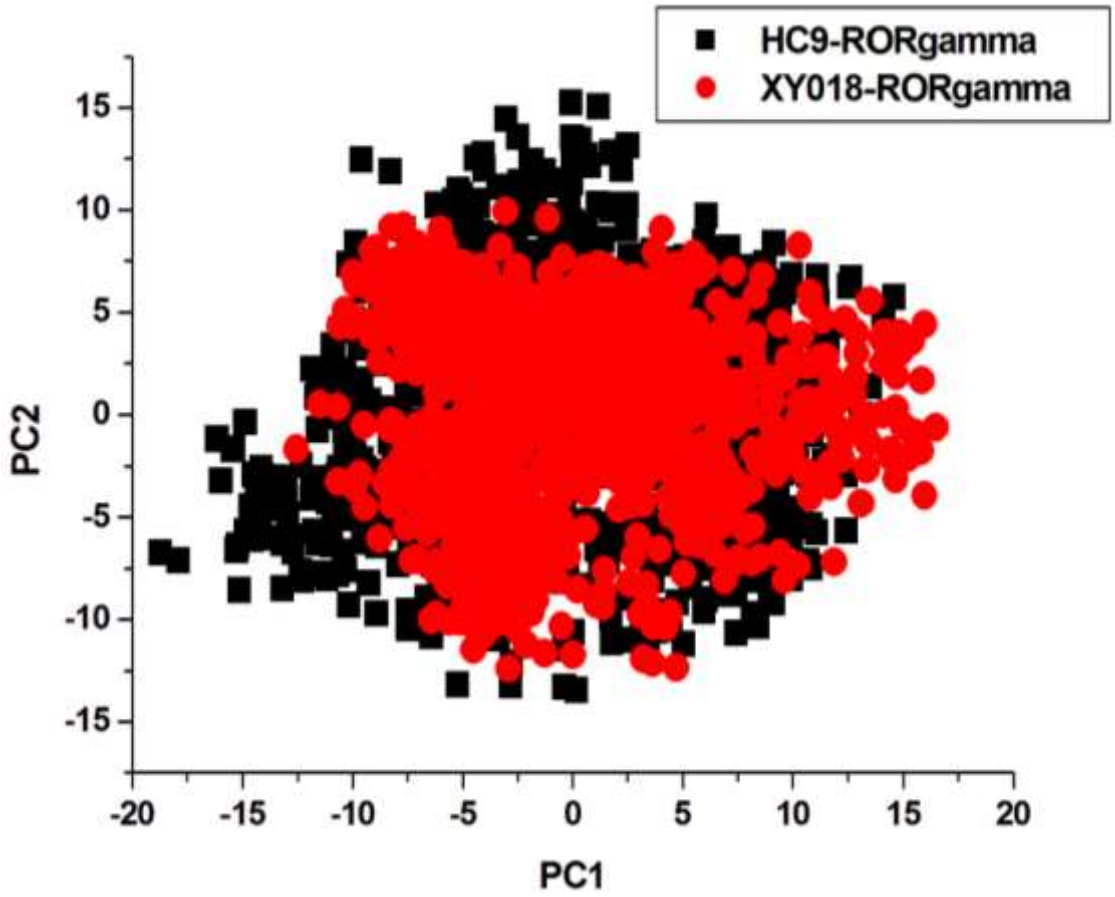


Figure 10. PCA projection of C-α atoms motion constructed by plotting the first two principal components (PC1 and PC2) in conformational space.

A distinct separation of motion was observed with XY018-ROR-γ system displaying a higher correlated motions along PC1 and PC2 compared to HC9-ROR-γ system which displays much more lower correlated motions along PC1 and PC2. It is obvious from the PC plot that XY018-ROR-γ system appears to be more compacted, meaning that the binding of XY018 to the residues in the active site of the receptor induces residue dynamics that results into conformational rigidity, thereby enhancing ligand residue interaction.

3.3.6 Hydrogen bond formation between amino acid residues

Hydrogen (H-bonds) are indispensable in nature and play important role in biological system and maintenance of the protein structural integrity,[48] protein-ligand interaction and catalysis.[48] Therefore, the formation of hydrogen bond between amino acid residues is key to the monitoring of protein conformation. H-bonding is known to be a facilitator of protein-ligand binding,[49] thus we investigate hydrogen bond formation during the course of the simulation of XY018-ROR-γ and HC9-ROR-γ systems. **Figure 11** shows hydrogen bond formation over time during simulation of respective systems. It was observed that more hydrogen bonds are formed in XY018-ROR-γ system compared to HC9-ROR-γ system. Reduction in hydrogen bond formation leads to structural imbalances and conformational dynamics which eventually affects drug binding.[48]

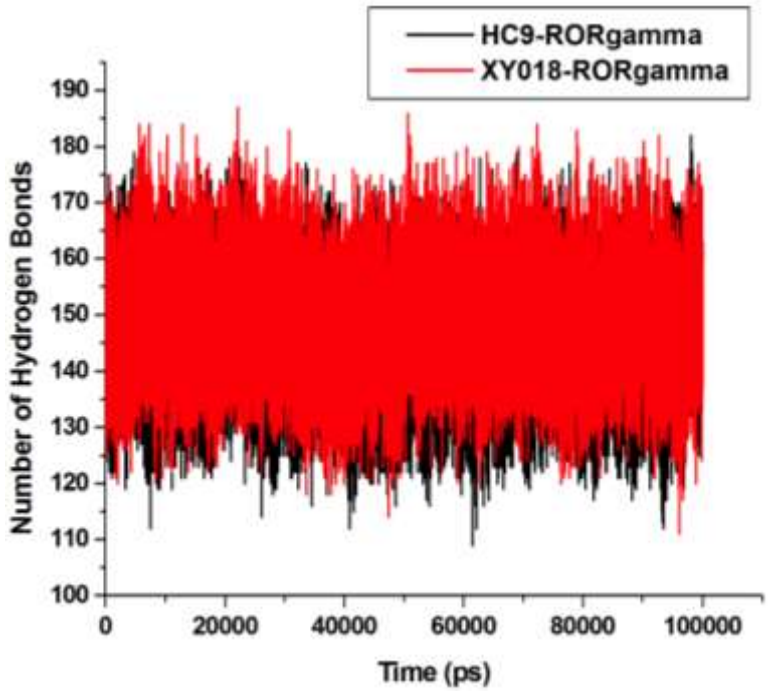


Figure 11. Number of hydrogen bond formation during a simulation over time between XY018-ROR-γ and HC9- ROR-γ.

3.3.7 Ligand-Residue Interaction Network Profile

Interaction of active site residues with the XY018 was examined to gain an insight into ligand-residue interaction. It was observed that the ROR-γ orientation formed a direct hydrogen bond with active site residue Glu115 and Gln22, **Figure 12 (A).** The interactive OH group XY018 that formed a hydrogen bond with Glu115 is essential for the binding of XY018 to ROR-γ. The phenyl group of XY018 interacts with glutamine (Gln22) by accepting a hydrogen through a hydrogen-bond formation. The fluoride atom also interacts with Ile133 as observed in **Figure**

12 (**C**). This explained the significance of these residues to ligand-receptor binding. Essentially, two binding pockets (P) were identified from the fully-minimised XY018-ROR-γ complex, P1 and P2 as observed in **Figure 12** (**B**). The residues Glu62, Gln22, Leu23 and Gln115 accommodates P1, while residues Ile133, Leu127, Met93 and Cys56 accommodates P2. **Figure 13** shows the hydrophobic active site residues and hydrogen bond interaction between ligand and residues.

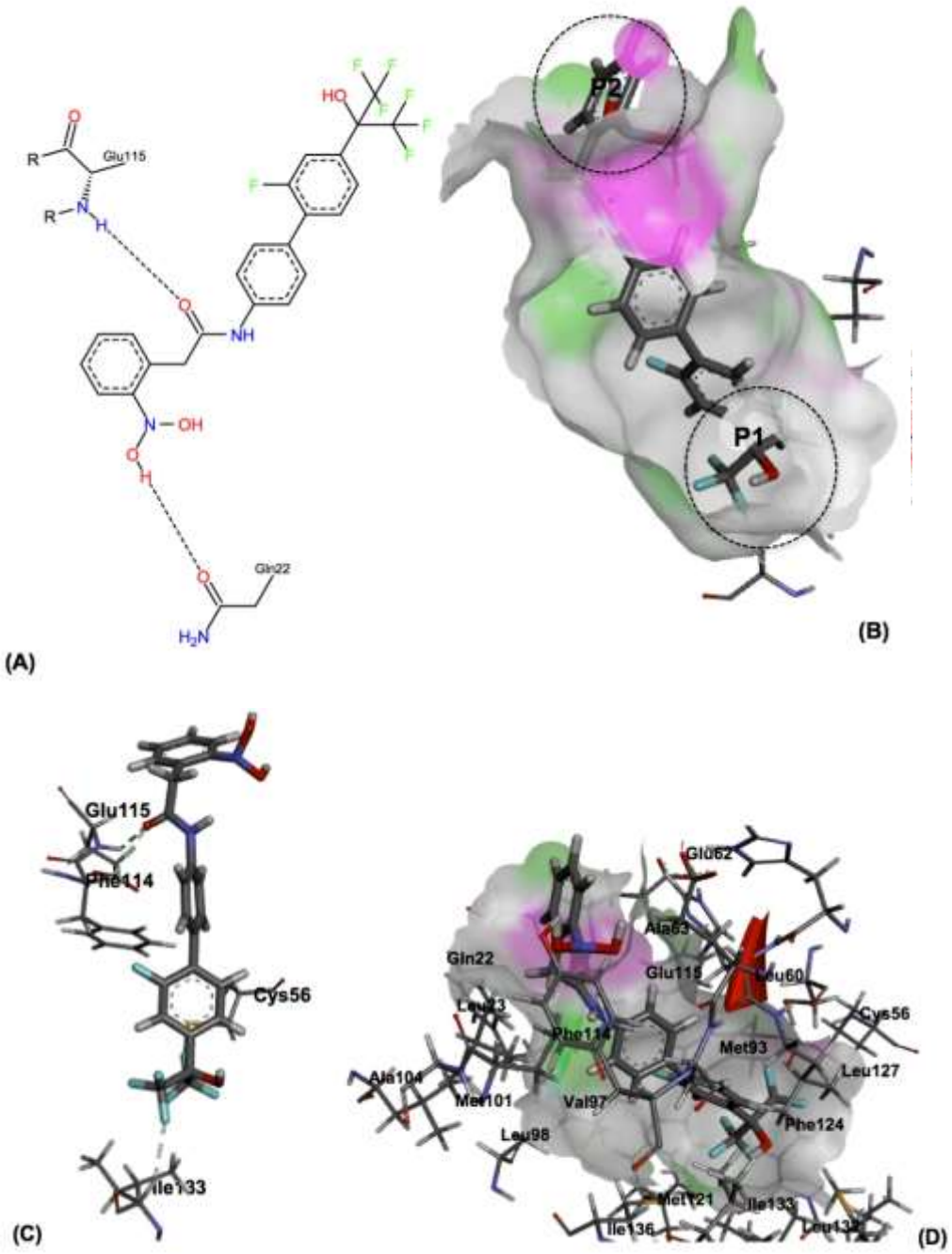


Figure 12. 2D structure of XY018 showing Hydrogen Bond formation with GLN 22 and GLU 115. (**B**) The binding pocket of ROR-γ from the fully-Minimized complex. (**C**) The interaction of Fluoride atom of XY018 with Ile133. (**D**) A network of ligand-residue interaction

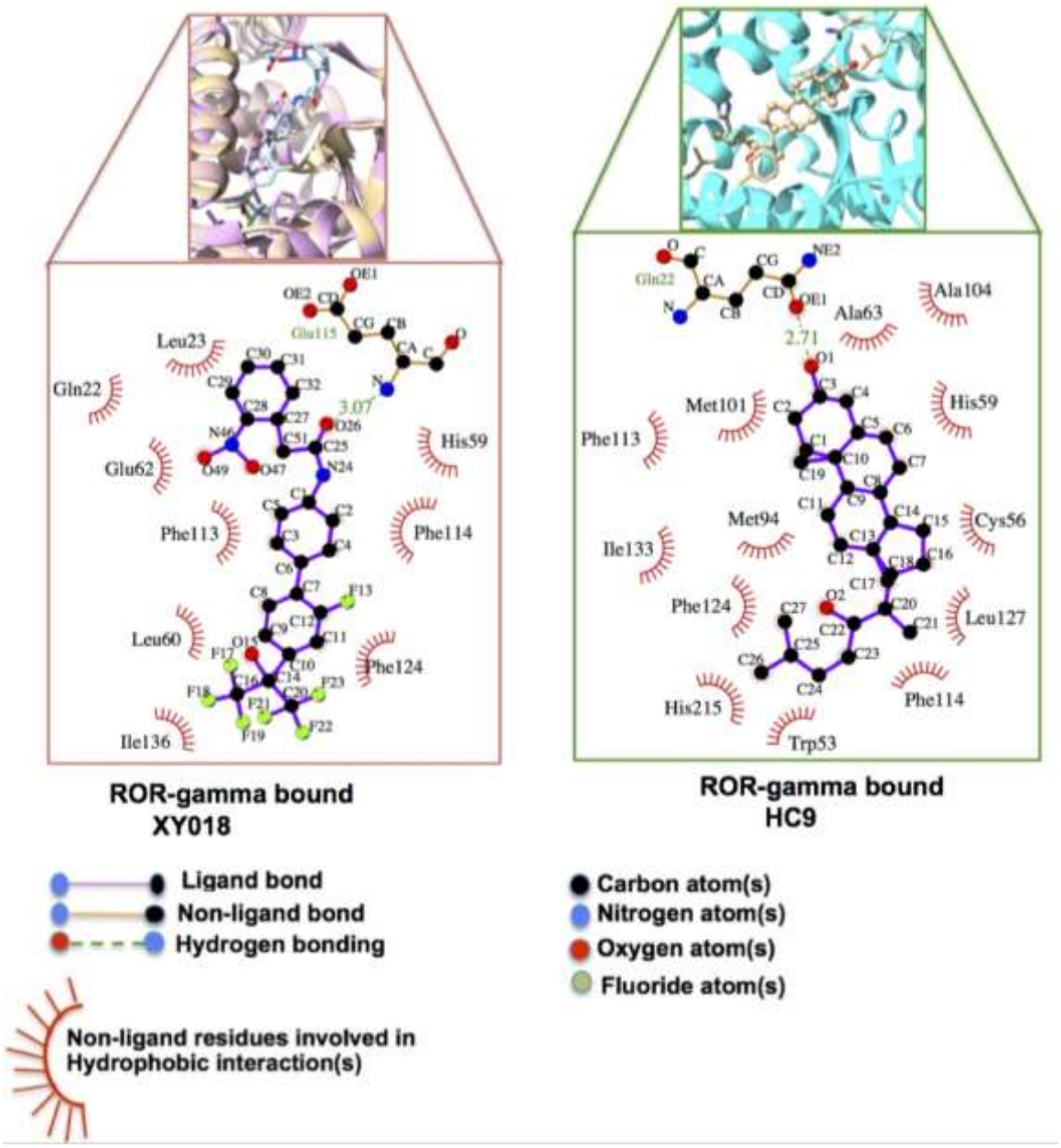


Figure 13. 2D residue-ligand interactions network from fully-minimised complex of XY018-ROR-γ and HC9-ROR-γ using ligplot analysis**.**

3.3.8 Hydrogen-Bond distance and occupancy

Table 4. Hydrogen Bond Occupancy of the interacting residues of XY018-ROR-γ and HC9-ROR-γ

H-bond acceptor	H-bond donor	Number of Frames	Occupancy (%)	Average distance (Å)	Average angle (degree)
RORγ-XY018					
Leu60@O	Ile64@H	866626	86	2.83	160
Leu23@O	Arg103@HH12	79963	79	2.82	155
Ile136@O	Ser140@HG	89628	89	2.73	162
Gln22@OE1	XY018@H50	71927	71	2.75	154
Phe113@O	XY018@H40	49668	49	2.85	156
Phe114@O	His59@HE2	42527	42	2.83	156
His59@O	Ala63@H	29107	29	2.90	158
Glu115@0E1	Arg55@HE	25286	25	2.84	160
Phe124@O	Gly128@H	18875	18	2.85	150
RORγ-HC9					
Leu60@O	Ile64@H	83129	83	2.84	160
His215@NE2	Tyr238@HH	1	12	2.94	135
Ile133@O	Phe137@H	63886	63	2.88	162
Trp53@O	Ala57@H	54311	54	2.87	158
Cys56@O	Leu60@H	69650	69	2.87	161
Gln22@OE1	HC9@H01	64067	64	2.74	162
Met94@O	Leu98@H	52645	52	2.88	157
Leu127@O	Trp53@HE1	36207	36	2.86	155
Phe124@O	Leu127@H	17631	17	2.90	158

Note: No = Number of frames; Å = Angstrom; % = Percentage.

To investigate the stability of XY018-ROR-γ and HC9-ROR-γ systems, hydrogen bond distance and occupancy of the active site residues in the respective systems were monitored throughout the course of simulations (**Table 4**). In XY018-ROR-γ system**,** His59 exhibits maximum H-bond distance (2.90 Å) while Ile136 exhibit the minimum H-bond distance (2.73 Å) with highest H-bond occupancy of 89%. Similarly, Leu60 has the highest H-bond occupancy of 80%, while Gln22 has the least H-bond distance of 2.74 Å in HC9-ROR-γ system. Interestingly, His215 with the highest H-bond distance of 2.94 Å in HC9- ROR-γ system show only 12% H-bond occupancy. Variation in H-bond occupancy and H-bond distance was observed in the respective systems with XY018-ROR-γ having the highest H-bond occupancy compared to HC9-ROR-γ. Therefore, the findings from previous experimental

studies[4] that form basis of our study support the quality of simulations conducted on the two systems.

3.3.9 Predicted toxicity and biological activity

Chemical compounds may interacts with biological targets different from experimentally recognised targets, such compounds are said to be promiscuous in nature. Compound promiscuity depicts the molecular basis of the pharmacological effects,[50] it is important to have a detailed assessment of these compounds to establish the extent of promiscuity among compounds at different level of drug research,[50] thus, predicting toxicity and biological activity of the compounds provides a detailed understanding of other properties of the drug that were probably not anticipated.

Toxicity and biological activity of compounds can be predicted using computer programme. It has been shown that the degree of reliability of these programmes varies from one to another.[51,52] In this study, we use "prediction of activity spectra for biologically active substance" (PASS) [51, 52] to predict possible toxicity and other biological activities that may be inherent on the XY018. The PASS predicts compound toxicity and biological profile with mean accuracy of prediction of about 89% -90%.[51, 52] **Table 5** shows the predicted toxicity and biological activity of XY018.

Table 5. Predicted toxicity and biological activity of XY018 using PASS

Pa	Pi	Activity
Predicted toxicity		
0.794	0.004	Pseudoporphyria
0.649	0.015	Nephritis
0.635	0.009	Interstitial nephritis
0.558	0.045	Coma
0.512	0.056	Allergic dermatitis
0.508	0.053	Keratopathy
0.420	0.004	Papillary necrosis
0.413	0.140	Consciousness alteration
0.348	0.117	Respiratory failure
0.370	0.119	Anaemia
0.358	0.139	Euphoria
0.314	0.154	Allergic contact dermatitis
0.270	0.131	Spermicide
0.303	0.156	Gastrointestinal haemorrhage
0.303	0.156	Gastrointestinal disturbances
0.284	0.192	Hepatotoxicity
0.259	0.221	Drowsiness
0.249	0.214	Nephrotoxic
Predicted biological activity		
0.720	0.005	Atherosclerosis treatment
0.542	0.051	Benzoate-CoA ligase inhibitor

Note: Pa probability of compound being active, Pi probability of compound being inactive. Pa > 0.7 indicates probability of toxicity or biological activity, Pi < 0.5 the compound is unlikely to exhibit toxicity or biological activity, 0.5 < Pa < 0.7 the compound is likely to exhibit toxicity or biological activity but the probability is less and Pa < 0.5 the compound is unlikely to exhibit the activity on experiment.

The result of predicted toxicity test showed that XY018 is likely to cause Pseudoporphyria, nephritis and interstitial nephritis, but less likely to cause conditions such as allergic dermatitis, keratopathy, papillary necrosis, respiratory failure, anaemia, euphoria and allergic contact dermatitis. The drug is potentially safe in hepatic and renal failure. It lacks any adverse effect on male reproductive organ (spermicide) and not likely to cause any gastrointestinal disturbances or haemorrhage, therefore can be formulated for oral intake. It lacks tendency to cause central nervous system effects such as drowsiness and consciousness alteration but may cause a coma if administered in higher doses.

However, the predicted biological activity of XY018 showed that it is highly indicated in the treatment of atherosclerosis and could also serve as benzoate CoA ligase inhibitor. This is one

of the high point discovery of this study, as we suggest that further studies in this regard be expanded.

4 Conclusion

Since the implication of an infant receptor ROR-γ in the castrated resistant prostate cancer, there is a lack of understanding of conformational features of the protein. Herein different computational approaches aimed at providing an in-depth understanding of conformational features of ROR-γ and the influence of XY018 binding to the protein conformation were explored. These approaches include MD simulations, principal component analysis, radius of gyration, binding free energy calculations, hydrogen bond formation and ligand-residue interaction network profile. Findings from this study revealed that; the two systems are relatively stable throughout the period of simulations; hydrophobic packing contributes significantly to binding free energy owing to large amount of aromatic and hydrophobic rings within the active site residues; the energy decomposition analysis revealed that electrostatic interactions are the potentially important binding forces between XY018 and ROR-γ, while van der Waals contributions are more prominent in HC9-ROR-γ system; Ile136 and Leu60 exhibited high hydrogen-bond occupancy in XY018-ROR-γ and HC9-ROR-γ respectively, therefore plays an important role in stabilizing the protein.

Similarly, His59 gives a higher energy contribution to the total binding free energy in XY018-ROR-γ while in HC9-ROR-γ, Leu60 contributed more to the total binding free energy. Analysis of principal components revealed that the binding of XY018 to ROR-γ may be responsible for structural rigidity and decrease in motion observed in the system compared to HC9-ROR-γ. The two system are closely correlated with relatively no significant difference in their RoG. There is a relatively high fluctuation of residues particularly at the loop region of a XY018-ROR-γ system compared to HC9-ROR-γ system as revealed by RMSF. It was also observed that interactive OH group of XY018 that formed a hydrogen bond with Glu115 is essential for the binding of XY018 to ROR-γ. The phenyl group of XY018 interacts with glutamine (Gln22) by accepting a hydrogen through a hydrogen-bond formation. Findings from estimated toxicity and biological testing suggest that XY018 is likely to induce pseudoporphyria, nephritis, and interstitial nephritis but potentially safe in renal failure. However, XY018 is potentially indicated in the treatment of atherosclerosis, one of the most important findings from this study.

Based on the information revealed from our study, the gate to the rational design of potential inhibitors of ROR-γ is opened. We believed that some current methods such as structural and

pharmacophore based virtual screening could be used to explore the properties of known inhibitor (XY018) for the design of new potential inhibitors of ROR-γ.

Acknowledgements

The authors acknowledge the School of Health Science, University of KwaZulu-Natal, Westville campus for financial assistance. The Centre for High Performance Computing (CHPC, www.chpc.ac.za) Cape Town, South Africa, for computational resources

Disclosure

Authors declare no financial and intellectual conflict of interests.

References

(1) Wang, F., Yang, W., Shi, Y., and Le, G. (2014) 3D-QSAR, Molecular docking and Molecular dynamics Studies of a series of RORγt inhibitors. *J. Biomol. Struct. Dyn.* 1–20.

(2) Shi, Y. (2007) Orphan nuclear receptors in drug discovery. *Drug Discov. Today 12*, 440–445.

(3) Aranda, A., and Pascual, A. (2001) Nuclear hormone receptors and gene expression. *Physiol. Rev. 81*, 1269–304.

(4) Wang, J., Zou, J. X., Xue, X., Cai, D., Zhang, Y., Duan, Z., Xiang, Q., Yang, J. C., Louie, M. C., Borowsky, A. D., Gao, A. C., Evans, C. P., Lam, K. S., Xu, J., Kung, H.-J., Evans, R. M., Xu, Y., and Chen, H.-W. (2016) ROR-γ drives androgen receptor expression and represents a therapeutic target in castration-resistant prostate cancer. *Nat. Med.* 1–13.

(5) Luckheeram, R. V., Zhou, R., Verma, A. D., and Xia, B. (2012) Review Article CD4 + T Cells : Differentiation and Functions. *Clinical and Developmental Immunology*. 12

(6) Ivanov, I. I., Mckenzie, B. S., Zhou, L., Tadokoro, C. E., Lepelley, A., Lafaille, J. J., Cua, D. J., and Littman, D. R. (2006) The Orphan Nuclear Receptor ROR g t Directs the Differentiation Program of Proinflammatory IL-17 + T Helper. *Cells.* 1121–1133.

(7) Jin, L., Martynowski, D., Zheng, S., Wada, T., Xie, W., and Li, Y. (2015) Structural Basis for Hydroxycholesterols as Natural ligand. *Mol Endocrinology*. *24*, 923–929.

(8) Roberts, R. (2014) From bench to bedside: the realities of reducing global prostate cancer disparity in black men. *Ecancermedicalscience 8*, 458.

(9) Fitzpatrick, J. M., Bellmunt, J., Fizazi, K., Heidenreich, A., Sternberg, C. N., Tombal, B., Alcaraz, A., Bahl, A., Bracarda, S., Di Lorenzo, G., Efstathiou, E., Finn, S. P., Foss??, S., Gillessen, S., Kellokumpu-Lehtinen, P. L., Lecouvet, F. E., Oudard, S., De Reijke, T. M., Robson, C. N., De Santis, M., Seruga, B., and De Wit, R. (2014) Optimal management of metastatic castration-resistant prostate cancer: Highlights from a European Expert Consensus Panel. *Eur. J. Cancer 50*, 1617–1627.

(10) Lallous, N., Dalal, K., Cherkasov, A., and Rennie, P. S. (2013) Targeting alternative sites on the androgen receptor to treat Castration-Resistant Prostate Cancer. *Int. J. Mol. Sci. 14*, 12496–12519.

(11) Fujita-sato, S., Ito, S., Isobe, T., Ohyama, T., Wakabayashi, K., Morishita, K., Ando, O., and Isono, F. (2011) Structural Basis of Digoxin That Antagonizes ROR □ t Receptor Activity and Suppresses Th17 Cell Differentiation and Interleukin (IL) -17 Production □ *286*, 31409–31417. *J Bio and Chem*. 36, 31409-314017

(12) Klepeis, J. L., Lindorff-Larsen, K., Dror, R. O., and Shaw, D. E. (2009) Long-timescale molecular dynamics simulations of protein structure and function. *Curr. Opin. Struct. Biol. 19*, 120–127.

(13) Sittel, F., Jain, A., and Stock, G. (2014) Principal component analysis of molecular dynamics: On the use of Cartesian vs. internal coordinates. *J. Chem. Phys. 141*.

(14) Maisuradze, G., Liwo, a, and Scheraga, H. (2009) Principal component analysis for protein folding dynamics. *J. Mol. Biol. 385*, 312–329.

(15) Kurylowicz, M., Yu, C. H., and Pom??s, R. (2010) Systematic study of anharmonic

features in a principal component analysis of gramicidin A. *Biophys. J. 98*, 386–395.

(16) Charles C. David and Donald J. Jacobs (2014) Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins. *Methods Mol Biol*. 1084, 193–226

(17) Wolf, A., and Kirschner, K. N. (2013) Principal component and clustering analysis on molecular dynamics data of the ribosomal L11??23S subdomain. *J. Mol. Model. 19*, 539–549.

(18) M Kumalo and M .E Soliman (2015) Per-Residue Energy Footprints-Based Pharmacophore Modeling as an Enhanced In Silico Approach in Drug Discovery : A Case Study on the Identification of Novel b -Secretase1 (BACE1) Inhibitors as Anti- Alzheimer Agents. *Cellular and Molecular Bioengineering.*1-9

(19) Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeerschd, T., Zurek, E., and Hutchison, G. R. (2012) Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform. 4*, 1–17.

(20) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem. 25*, 1605–1612.

(21) Steffen, C., Thomas, K., Huniar, U., Hellweg, A., Rubner, O., and Schroer, A. (2010) TmoleX--a graphical user interface for TURBOMOLE. *J. Comput. Chem. 31*, 2967–2970.

(22) Pacheco, A. B., and Hpc, L. S. U. (2012) Introduction to AutoDock and AutoDock Tools.

(23) Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J. (1998) Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem. 19*, 1639–1662.

(24) Case, D. A., Darden, T., Iii, T. E. C., Simmerling, C., Brook, S., Roitberg, A., Wang, J., Southwestern, U. T., Duke, R. E., Hill, U., Luo, R., Irvine, U. C., Roe, D. R., Walker, R. C., Legrand, S., Swails, J., Cerutti, D., Kaus, J., Betz, R., Wolf, R. M., Merz, K. M., State, M., Seabra, G., Janowski, P., Paesani, F., Liu, J., Wu, X., Steinbrecher, T., Gohlke, H., Homeyer, N., Cai, Q., Smith, W., Mathews, D., Salomon-ferrer, R., Sagui, C., State, N. C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., Berryman, J., and Kollman, P. A. (2014) Amber 14. *Univ. California, San Fr.*

(25) Goetz, A. W., Williamson, M. J., Xu, D., Poole, D., Grand, S. L., and Walker, R. C. (2012) Routine microsecond molecular dynamics simulations with amber - part i: Generalized born. *J. Chem. Theory Comput. 8*, 1542–1555.

(26) Galindo-Murillo, R., Robertson, J. C., Zgarbová, M., Šponer, J., Otyepka, M., Jurečka, P., Cheatham, T. E., and III. (2016) Assessing the Current State of Amber Force Field Modifications for DNA. *J. Chem. Theory Comput. 12*, 4114–27.

(27) Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., and Shaw, D. E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct. Funct. Bioinforma. 78*, 1950–1958.

(28) Wang, J. M., Wolf, R. M., Caldwell, J. W., Kollman, P. a, and Case, D. a. (2004) Development and testing of a general amber force field. *J. Comput. Chem. 25*, 1157–1174.

(29) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926.

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

(31) Johnson, A., Johnson, T., and Khan, A. (2012) Thermostats in Molecular Dynamics Simulations . University of Massachusetts Amherst

(32) Cheng Zhang and Michael W. Deem (2013) Multicanonical molecular dynamics by variable-temperature thermostats and variable pressure barostats. *J.Chem Phys*. 138, 034103

(33) Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys. 23*, 327–341.

(34) E. Seifert (2014) OriginPro 9.1: Scientific Data Analysis and Graphing Software—Software Review. *J. Chem. Inf. Model.* 1552-1552

(35) Skjærven, L., Yao, X.-Q., Scarabelli, G., and Grant, B. J. (2014) Integrating protein structural dynamics and evolutionary analysis with Bio3D. *BMC Bioinformatics 15*, 399.

(36) Genheden, S., and Ryde, U. (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov. 10*, 449–61.

(37) You, W., Huang, Y. M., Kizhake, S., and Natarajan, A. (2016) Characterization of Promiscuous Binding of Phosphor Ligands to Breast-Cancer-Gene 1 (BRCA1) C-Terminal (BRCT): Molecular Dynamics , Free Energy , Entropy and Inhibitor Design . *PloS Comput Biol*. *1*, 1–25.

(38) Amadei, A., Linssen, A. B. M., and Berendsen, H. J. C. (1993) Essential dynamics of proteins. *Proteins Struct. Funct. Genet. 17*, 412–425.

(39) Cocco, S., Monasson, R., and Weigt, M. (2013) From Principal Component to Direct Coupling Analysis of Coevolution in Proteins: Low-Eigenvalue Modes are Needed for Structure Prediction. *PLoS Comput. Biol. 9*.

(40) Carugo, O. (2001) A normalized root-mean-square distance for comparing protein three-dimensional structures. *Protein Sci.* 1470–1473.

(41) Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. *Molecular Biology of cell. 4th edition Geral Sciences New York*

(42) Loeffler, H. H., and Winn, M. D. (2013) Ligand binding and dynamics of the monomeric epidermal growth factor receptor ectodomain. *Proteins Struct. Funct. Bioinforma. 81*, 1931–1943.

(43) Ahmad, E., Rabbani, G., Zaidi, N., Khan, M. A., Qadeer, A., Ishtikhar, M., Singh, S., and Khan, R. H. (2013) Revisiting ligand-induced conformational changes in proteins: essence, advancements, implications and future challenges. *J. Biomol. Struct. Dyn. 31*, 630–648.

(44) Wikipedia, https://en.wikipedia.org/wiki/Radius_of_gyration

(45) Benson, N. C., and Daggett, V. (2008) Dynameomics: large-scale assessment of native protein flexibility. *Protein Sci. 17*, 2038–50.

(46) Desdouits, N., Nilges, M., and Blondel, A. (2015) Principal Component Analysis reveals correlation of cavities evolution and functional motions in proteins. *J. Mol. Graph. Model. 55*, 13–24.

(47) Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A., and Caves, L. S. D. (2006) Bio3d: An R package for the comparative analysis of protein structures. *Bioinformatics 22*, 2695–2696.

(48) Chen, D., Oezguen, N., Urvil, P., Ferguson, C., Dann, S. M., and Savidge, T. C. (2016) Regulation of protein-ligand binding affinity by hydrogen bond pairing. *Sci. Adv. 2*, e1501240.

(49) Chen, D., Oezguen, N., Urvil, P., Ferguson, C., Dann, S. M., and Savidge, T. C. (2016) Regulation of protein-ligand binding affinity by hydrogen bond pairing. *Sci. Adv. 2*, e1501240–e1501240.

(50) Hu, Y., Gupta-Ostermann, D., and Bajorath, J. (2014) Exploring Compound Promiscuity Patterns and Multi-Target Activity Spaces. *Comput. Struct. Biotechnol. J. 9*, 1–11.

(51) Lagunin, A., Stepanchikova, A., Filimonov, D., and Poroikov, V. (2000) PASS: prediction of activity spectra for biologically active substances. *Bioinformatics 16*, 747–8.

(52) Poroikov, V., Filimonov, D., Lagunin, A., Gloriozova, T., and Zakharov, A. (2007) PASS: identification of probable targets and mechanisms of toxicity. *Sar Qsar Environ. Res. 18*, 101–110.

CHAPTER 5

Emergence of a Promising Lead Compound in the Treatment of Triple Negative Breast Cancer: An Insight into Conformational Features and Ligand Binding Landscape of c-Src Protein with UM-164

Umar Ndagi[a] Ndumiso N. Mhlongo[a] and Mahmoud E. Soliman[a]*

[a] Molecular Modelling and Drug Design Research Group, School of Health Sciences, University of KwaZulu-Natal, Westville, Durban 4000, South Africa

*Corresponding author: Mahmoud E. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27(0)312608048, Fax: +27 (0)31260 7872

Webpage: http://soliman.ukzn.ac.za/

Abstract

UM-164, a potent Src/p38 inhibitor, is a promising lead compound for developing the first targeted therapeutic strategy against triple-negative breast cancer (TNBC). However, a lack of understanding of conformational features of UM-164 in complex with Src serves a challenge in the rational design of novel Src dual inhibitors. Herein, we provide first account of in-depth insight into conformational features of Src-UM-164 and the influence of UM-164 binding to the Src using different computational approaches. This involved molecular dynamics (MD) simulation, principal component analysis (PCA), thermodynamic calculations, dynamic cross-correlation (DCCM) analysis and ligand-residue interaction network profile, as well as toxicity testing. Findings from this study revealed that: (1) the binding of UM-164 to Src induces a more stable and compact conformation on the protein structure; (2) UM-164 binding to Src induces highly correlated motions in the protein; (3) high fluctuation exhibited by the loops in Src-UM-164 system support the experimental evidence that UM-164 binds the DFG-out inactive conformation of Src; (4) a relatively high binding free energy estimated for the Src-UM-164 system is affirmative of its experimental potency; (5) hydrophobic packing contributes significantly to the drug binding in Src-UM-164; (6) a relatively high H-bond formation in Src-UM-164 indicates enhanced drug-protein interaction; (7) UM-164 is relatively less toxic than Dasatinib, therefore, is potentially safer. The finding of this study can provide important insights for further design of novel Src inhibitors.

Key words: Src, Molecular dynamics, TNBC, UM-164, Dual kinase inhibitor and DFG-out.

1 Introduction

Breast cancer remains among the most frequently diagnosed and life-threatening cancer in women,[1] and third leading cause of cancer-associated death among women in the United State of America (US).[1,2] Genetically, breast cancers (BCs) are heterogeneous[3] with respect to their gene composition, gene expression, and phenotypes which eventually yield current classifications of 5 subtypes.[3] The triple negative subtypes are more life threatening due to their potential to metastasize and a tendency of local reoccurrence.[4] They are usually associated with the absence of oestrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor-2 (ErbB2/HER-2).[5] They are characterised by classical ductal histology, high grade, high mitotic and cell proliferations rates.[5] The triple negative cancer (TNBC) is firmly associated with poor prognosis, poor disease free survival (PDFS) and cancer specific survival (CSS).[5] The local reoccurrence is marked with increasing number of positive lymph nodes,[5] this suggest the reason for high risk of reoccurrence in patients with TNBC in the first 3 to 5 years after diagnosis.[5] Study have shown that only a few therapeutic options and conventional chemotherapy may probably be the only effective treatment for patients after surgery.[5]

However, other studies have shown that chemotherapeutic agents (neoadjuvant) such as 5-fluorouracil (5-FU), anthracycline, cyclophosphamide, taxanes and platinum compounds used in the treatment of TNBC have a low success rate.[6] This may partly be due to lack of acceptable predictive biomarkers,[3] safety concerns and resistance to some of the compounds.[6] Similarly, a study conducted to define the molecular basis of TNBC classified TNBC as ER, PR and HER-2. These molecular markers provide understanding of the potential targets in the course of therapy,[7] while the prognostic markers such as HER_1, $ALDH_1$, $LOXL_2$, Ki-67, SNCG and LDHB are important in providing prognostic information.[7] This is, however, contrary to the initial belief that the major challenge in treatment of TNBC is a lack of validated molecular target in the tumour.[8]

A study set to determine the influence of tissue inhibitor of metalloproteinases-1 (TIMP-1) on TNBC[9] revealed that TIMP-1 is a biomarker indicative of poor prognosis in TNBC diseased individuals.[9] TIMP-1 may provide attractive therapeutic intervention specifically for TNBC.[9] The relative disposition to TNBC biomarkers such as ER, PR and ErbB2/HER-2 have provided tremendous advances in the treatment of breast cancer in the last 3 decades,[3] this is especially

true in patients whose tumour overexpressed ErbB2 and or hormone receptor.[3] Most of the therapies targeted at TNBC are not effective[3] because of lack of ER, PR and HER-2, however, efforts have been made to surmount these challenges in the last 3 decades, some of which have resulted in a drastic improvement in patient survival[3]. The relative heterogeneity nature of TNBC among breast cancers made it challenging for FDA-approved targeted therapies and cytotoxic agents remain the backbone of chemotherapy in TNBC.[10]

The understanding of clonal evolution and heterogeneity of breast cancer[3] revealed by previous clonal evolution study provide only limited information on the genomic diversity of tumours.[11] Burdened by these challenges, a recent study on clonal evolution used a whole genome and exome single cell sequencing approach[11] to determine heterogeneity and evolution of breast cancer. The result from this study shows that triple negative breast cancer has an increased mutation rate compared to ER^+ tumour cell which did not show any form of mutation.[11] Findings from this study challenged the belief of single pathway target in the treatment of breast cancer.[3,11] To this extent, there is increasing interest in the inhibition of multiple pathways given that TNBC is heterogeneous in nature.[3]

A more recent study established a relationship between TNBCs and sensitivity to inhibition of c-Src (Src), in an attempt to identify predictive markers response to chemotherapy.[3] In this study, a dual kinase inhibitor known as UM-164 was discovered and had profound activity against Src and p38 kinases.[3] This inhibitor is said to be a promising lead compound for developing the first targeted therapeutic strategy against triple negative breast cancer (TNBC).[3] Src is the cellular homolog of the viral oncogene v-Src[12] and an archetype member of a family of non-receptor tyrosine kinases that play important roles in a variety of signalling pathways that involve proliferation, differentiation, survival, motility, and angiogenesis.[12] Overexpression of Src plays an important role in oncogenic proliferation, migration, and invasion of TNBC cell lines.[3] This claim is supported by molecular studies that continued to show that Src plays a significant role in clinically important pathways in breast cancer,[12] such as steroid and peptide hormone pathway.[12] On the basis of these reports, Src represents an attractive target in TNBC.[13, 14] Although previous studies placed more emphasis on the use Dasatinib, Bosutinib and Saracatinib particularly as a combination therapy in the treatment of TNBC.[14] These drugs act by binding the active conformation of the kinase, in addition, resistance to Dasatinib has emerged.[15] Cumulatively, these factors placed other drugs (Dasatinib and Bosutinib) at the disadvantage over UM-164 (Src/p38 inhibitor) which act in a

specific inactive conformation (DFG-out). Structure of UM-164 and Dasatinib are presented in **Figure 1.**

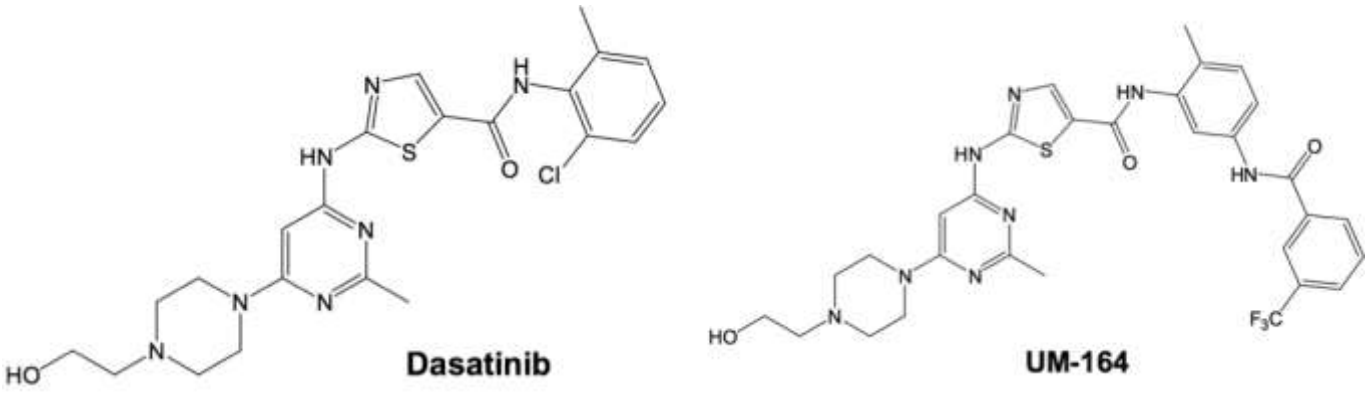


Figure 1. 2D structures of Dasatinib and experimental Src inhibitor (UM-164).

Dasatinib is a potent Src kinase inhibitor,[8] available as oral medication and can be used as a single agent in the treatment of TNBC [8] and other forms of cancer including leukaemia.[16] Phase II clinical trial assessment of Dasatinib as a single agent in TNBC revealed a significant efficacy and safety in patients with advanced TNBC. Structurally, the addition of fluorinated benzene ring to UM-164 marked the major differences between the two, such differences have translated to improve efficacy and safety of the analogue (UM-164) against the original drug Dasatinib.

Approximately 15% of invasive breast cancer is triple negative.[17,18] This huge population is a global phenomenon[19] and a source of concern because women with TNBC experience the peak risk of recurrence within 3 years of diagnosis,[17] and increased mortality rate 5 years post diagnosis.[17] Therefore, the discovery of UM-164 (an analogue of Dasatinib) could be termed as a popular breakthrough in absence of superior therapies in the management of TNBC.

It is known that all clinically used Src inhibitors act by binding the active conformation of the kinase.[3] However, it was hypothesised that inhibiting Src in a specific inactive conformation (DFG-out) would have improved efficacy against TNBCs.[3] Therefore, inhibiting a kinase in the DFG-out inactive conformation can have dramatic effects on the non-catalytic functions of the kinase.[3] Fortunately, experimental results revealed that UM-164 act by binding on the DFG-out inactive conformation of Src.[3] **Figure 2** shows DFG-out inactive conformations of Src complex with UM-164.

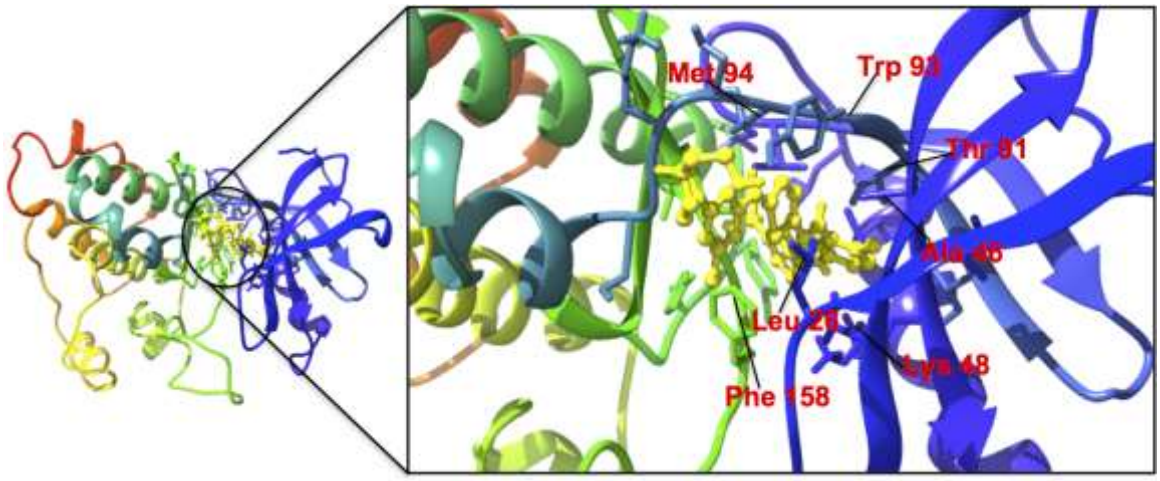


Figure 2. 3D structure of Src DFG-out inactive conformation showing the binding site of UM-164.

In the wake of the emergence of drug resistance in TNBC,[15] and the burden of TNBC among people living with breast cancer[17] across the globe. The need for an in-depth molecular understanding of conformational features and ligand binding landscape of UM-164 complex with Src is key to successful management of TNBC cases. This is especially more important because of growing population of TNBC among the invasive breast cancer. Therefore, a good understanding of conformational features and ligand binding landscape of Src is crucial to the development of new potential and effective inhibitors of Src.

Currently, no conformational studies have been carried out on Src complex with UM-164, thus a detailed analysis of conformational features of Src complex with UM-164 will require long time-scale molecular dynamic simulations that will provide adequate information on the dynamics of the protein. Therefore, to have an atomistic insight to experimental work[20] that have already been conducted on UM-164-complexed Src, we conduct MD simulations to broaden the knowledge on the subject matter.

Various post dynamic techniques have been used to provide a molecular understanding of molecular dynamics. The principal component analysis (PCA), also known as essential dynamics analysis,[21] is one of the most popular post-dynamics techniques [21] that is widely applied to understand the changes in biological systems.[21] PCA eliminates translational and rotational motions in molecular dynamic (MD) trajectory and correlated motions in atomic simulations of proteins.[22] It defines the atomic displacement in a collective manner,[21] and can detect major conformational changes between the structures[23] and has been used in many studies to determine the difference in motion of protein complex of two different compounds.[21] In this study, our objective is to provide insight into the conformational features and ligand

binding landscape of Src in complex with UM-164 and contributing to the understanding of experimental work that has been previously conducted. To achieve this, we performed molecular dynamic simulations of Src complex with UM-164 (Src-UM-164), Src complex with Dasatinib (Src-Dasatinib) and apo to further inspect the effects UM-164 binding on the dynamic state of Src. To facilitate the process of MD, Dasatinib was first docked into Src and the docked conformation was used for MD simulations respectively (see **Figure S1** and **Table S1** of the supplimentary Material). Herein, we perform 150 ns of MD simulations, PCA, dynamic cross-correlation and binding free energy calculations were also conducted to understand the effect of UM-164 binding to the dynamic state of the protein.[23] These tools are known to enhance the process of drug discovery,[24] therefore, provide a platform for the discovery of novel therapeutics.[24] Finding from this study would demonstrate the conformational and structural properties of Src in complex with UM-164, such properties may form the baseline for which other potential therapeutics targeting Src can be developed.

2 Computational methods

2.1 System preparation

The X-ray crystal structure of Src in complex with UM-164, PDB code 4YBJ[25] was obtained from Protein Data Bank (PDB). This is a static (inactive) conformation of Src structure exist as a homodimer with two chains (A and B), however, only chain A was used for simulation in this study to reduce the computation cost. Chimera[26] and Avogadro software package[27] were used to modify and visualise ligand and receptor respectively.

2.2 Molecular dynamic simulations

Simulations of bound Src-UM-164, Src-Dasatinib complexes as well as unbound apo were performed using graphic processor unit (GPU) version of Particle Mesh Ewald Molecular Dynamics (PMEMD) package with Sander module of Amber14.[28, 29] The AMBER force field ff12SB[30, 31] was applied to describe the protein.[31] The ligands parameters were set using Gasteiger charges in Avogadro,[27] and Antechamber module with the aid of GAFF (generalised Amber force field).[32] The LEAP module implemented in Amber14[33] was used to add hydrogen atoms to the protein and to add counter ions for the system neutralization.[33] Each system is enclosed in the TIP3P water box[31] with the protein atoms located 10 Å between the protein surface and the box boundary within the period of simulations. The cubic periodic boundary conditions were implemented in all the systems, long-range electrostatics interaction was treated with particle-mesh Ewald method[31] implemented in Amber14 with a nonbonding cut-

off distance of 12 Å. Two minimization steps were employed, partial minimization and full minimization. The initial energy minimization step of the systems was carried out with a restraint potential of 500 kcal mol^{-1} $Å^{-2}$ apply to the solute, for 1000 steps. Unrestrained conjugated gradient minimization for 1000 steps was conducted for the entire system with aid of SANDER module of Amber 14 program. A canonical ensemble (NVT) MD simulations were performed for 50 ps and the system was gradually heated from 0 to 300 K, with harmonic restraints of 5 kcal mol^{-1} $Å^{-2}$ for solute atoms with the aid of Langevin thermostat[34] with a 1ps random collision frequency. The systems were equilibrated at 300 K with a 2fs time step in NPT ensemble for 500 ps without any restraint and Berendsen barostat[34] was used to maintain the pressure at 1bar. The SHAKE[35] algorithm was used to constrain the bonds of hydrogen atoms in the system. The 2fs time scale and SPFP precision model were used for MD runs. To achieve sufficient sampling and ensure systems convergence, we performed a 150 ns of MD simulations in the absence of restraints, using NPT ensemble at a target pressure of 1 bar and a 2 ps pressure coupling constant. For every 1ps time interval, the coordinates were saved and the trajectories were analysed every 1ps.

Post MD analysis performed include root mean square deviations (RMSD), root mean square fluctuations (RMSF), the radius of gyration, hydrogen-bond occupancy, dynamic cross-correlation and principal component analysis (PCA) using CPPTRAJ[36] modules in Amber 14, as well as ligand-residue interaction profile and predicted toxicity test. Visualisation of trajectories was conducted in the chimera.[26] The results were analysed and plots were generated with aid of Origin[37] and Bio3D[38] software respectively.

2.3 Thermodynamic calculations

The binding free energy calculation is an important thermodynamic method that gives detailed information on the interaction between the ligand and protein.[39] It provides good understanding of mechanism of binding, including contributions from both enthalpy and entropy to the molecular recognition.[39] Molecular Mechanics/Generalized-Born Surface Area method (MM/GBSA)[39] is a popular method used to estimate free energy of binding of small ligands to the biological macromolecule,[39] the calculation gives detailed information on the interaction between the ligand and protein.[39] The binding free energy of Src-UM-164 and Src-Dasatinib systems were calculated using MM/GBSA.[39] For a 150 ns trajectory, 1000 snapshots were considered during the calculation of binding free energy. The following equations described binding free energy calculation:

$$\Delta G_{bind} = G_{complex} - G_{receptor} - G_{ligand} \quad (1)$$

$$\Delta G_{bind} = E_{gas} + G_{sol} - TS \quad (2)$$

$$E_{gas} = E_{int} + E_{vdW} + E_{ele} \quad (3)$$

$$G_{sol} = G_{GB} + G_{SA} \quad (4)$$

From the equation above, E_{gas} is the energy of the gas phase, E_{int} represents internal energy, E_{ele} represents Coulomb while E_{vdW} is the van der Waals energies. E_{gas} is estimated directly from the ff12SB[30] force field. G_{sol} which is the solvation free energy can be broken down to polar and non-polar forms of contribution. The contribution of polar solvation (G_{GB}) is assessed by resolving G_{GB} equation and non-polar solvation (G_{SA}) is determined from the solvent accessible surface area, which can be estimated from water probe radius of 1.4 Å with temperature (T) and total solute entropy (S). The MM/GBSA binding free energy method in Amber 14 was used to calculate the contribution of each residue to the total binding free energy.

2.4 Principal component analysis

Principal component analysis (PCA) also known as essential dynamics of protein[21] analysis is a multivariate statistical technique applied to systematically reduce the number of dimensions needed to describe the protein dynamics[21] through the decomposition process that screen observed motions from largest to smallest spatial scale.[21] The atomic displacement and conformational changes of proteins can be defined[23] using PCA by extracting different modes of the conformation of the protein complex during dynamic simulations. The direction of motion (eigenvectors) and the extent of motion (eigenvalues) of the biological system can also be determined using PCA. Herein, 150 ns of MD trajectories were stripped of the solvent molecules and the ions using the CPPTRAJ module in Amber 14.[36] This was done prior to MD trajectory processing for PCA. Principal component analysis was performed on C-α atoms on 1000 snapshots at 100 ps time interval each. The first two principal components (PC1 and PC2) were computed and 2 X 2 covariance matrices were generated using Cartesian coordinates of Cα atoms. PC1 and PC2 correspond to first two eigenvectors of a covariant matrix. Origin software was used to construct the PC plot. [37]

2.5 Dynamic cross-correlation matrices (DCCM)

The cross-correlation is a 3D matrix representation that graphically displays time correlated information among the residues of the proteins.[40] Residue-based time correlated data can be analysed using visual pattern recognition.[41] To better understand the dynamics of apo, Src-UM-164 and Src-Dasatinib systems, a DCCM was generated to determine cross-correlated displacements of backbone Cα atoms in the trajectories, using the following equation

$$C_{ij} = \langle \Delta r_i * \Delta r_j \rangle / (\langle \Delta r^2_i \rangle \langle \Delta r^2_j \rangle)^{1/2} \quad (5)$$

Where, *i* and *j* represent *i*th and *j*th residues and Δr_i and Δr_j corresponds to the displacement of *i*th and *j*th atom from the mean respectively. The coefficient of cross-correlation $C_{ij,}$ varies between the range -1 to +1, where the upper and lower limits correspond to strong correlated (+) and anti-correlated (-) motions within the period of simulations. The DCCM analysis was conducted using CPPTRAJ module in Amber 14.[36] Matrices were generated and analysed using Origin software.[37, 38]

3 Results and discussion

3.1 System stability MD simulations

In preparation for MD trajectory analysis, RMSD and potential energy fluctuations were monitored throughout the MD simulations. RMSD was calculated to assess the stability and convergence of the respective systems and the results are presented **Figure 3**. Systems stabilisation and convergence with maximum fluctuation of 4.25 Å between 0-12500 ps in apo system was observed, whereas in the Src-Dasatinib system, the highest fluctuation observed was 3.25 Å, at 12500 ps, after which the fluctuation rested below 2.00 Å (**Figure 3)**. However, after approximately 50000 ps, the RMSD trajectories converged and the fluctuation rested below 2.00 Å for all the systems. Similarly, in Src-UM-164 system the peak RMSD of 3.5 Å was reached at about 9000 ps. The average RMSD of 1.58 Å and 1.52 Å was observed in Src-UM-164 and Src-Dasatinib systems respectively, while the apo system has an average RMSD of 1.46 Å. These account for system stability since a standard parameter defining a stable system is an RMSD of 2 Angstroms and below.[42] These results show that Src-UM-164 and Src-Dasatinib systems, exhibit more flexibility and deviation compared to the apo system which appears to be most stable sequel to average RMSD of 1.46 Å.

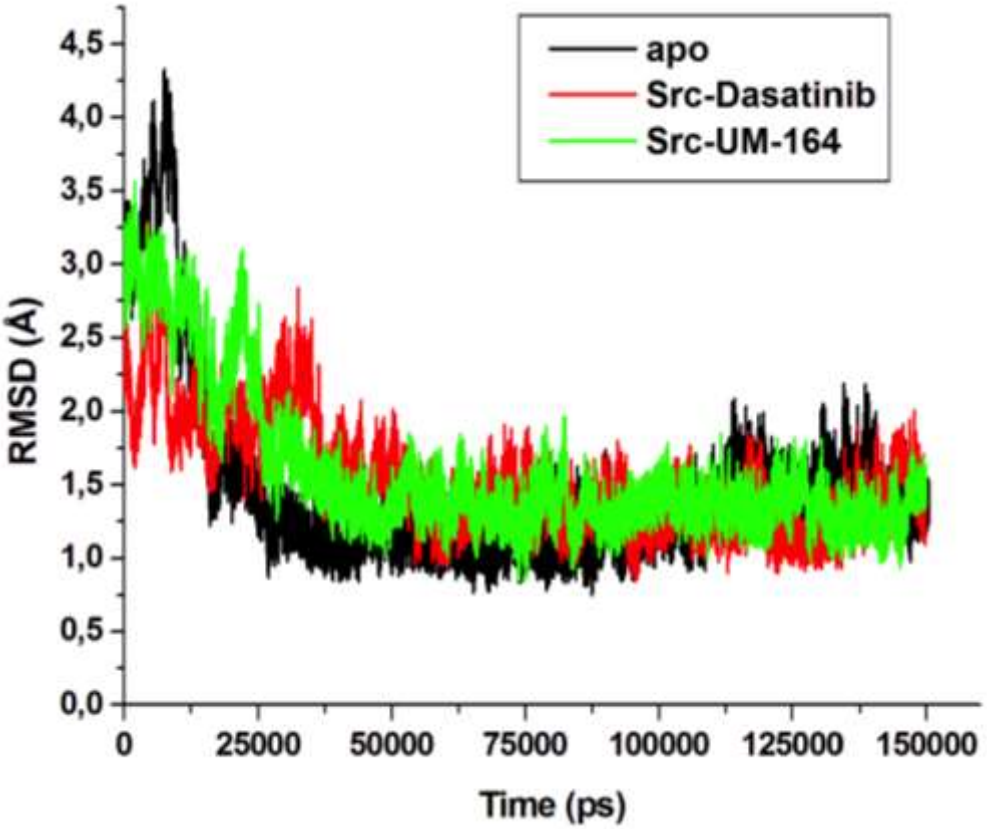


Figure 3. The RMSD plot of apo (black), Src-Dasatinib (red) and Src-UM-164 (green) respectively.

3.2 Root mean square fluctuation (RMSF)

An aggregate of a specific sequence of amino acids make up a proten[43] and play a vital role in conformational features of the protein.[43] Changes to the protein conformation occur when there is a chemical reaction or mechanical events.[43] Therefore, direct interactions of the protein active site residues with a ligand may induce conformational changes in protein structure and alter its function. More specifically, the conformational changes that occur as a result of ligand-induced motion during ligand binding.[44] Understanding ligand-induced conformational changes in the protein structure are critical to structure-based rational drug design.[45] RMSF is a measure of average atomic mobility of backbone atoms (N, Cα and C) during MD simulation.[46] To understand and explore the structural dynamics that take place upon the ligand binding, RMSF of the subject systems was calculated from MD trajectories and the plot is presented in **Figure 4**. The core of the protein appears to be more rigid compared to the loops (solvent exposed) as shown by RMSF plot.

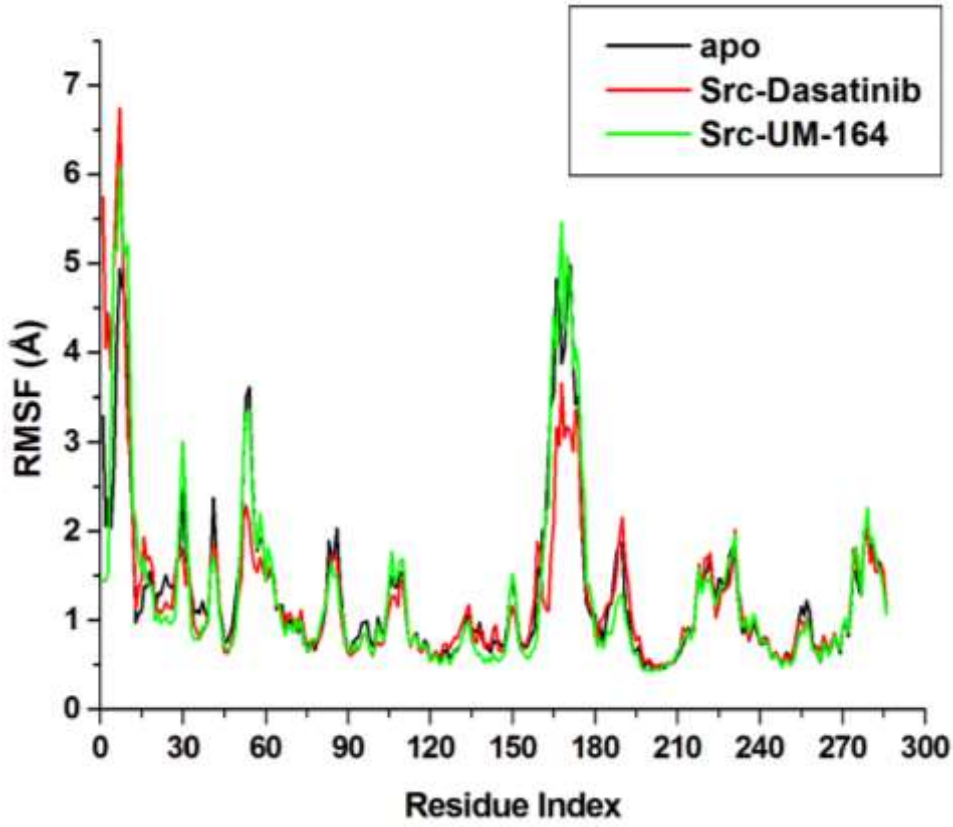


Figure 4. The Cα RMSF plots of apo (black), Src-Dasatinib (red) and Src-UM-164 (green) respectively.

The highest fluctuations were observed in the loop region involving residues Ala171, Arg172, Gln173, Lys176, Pro178 and Lys180 in the respective systems. Similarly, the loop region with His2, Met3, Gln4, Gln6, Lys10, Trp13 and Pro16 also exhibits maximum residue fluctuations with the Src-UM-164 system being the highest amongst the respective systems. In a related development, the loop region involving residues Lys51, Pro52, Met55 and Pro57 exhibit maximum fluctuations in the Src-UM-164 system compared to Src-Dasatinib and apo systems. These loops are called activation loops, [47] when phosphorylated becomes rigid and contribute to the switch from inactive to an active conformation.[47] Therefore, high fluctuation exhibited by the loops in Src-UM-164 system support the experimental evidence that UM-164 binds the DFG-out inactive conformation of Src.[3] Ligand–residue interaction relatively stabilises Src-UM-164 and Src-Dasatinib systems, thus resulting in a slight decrease in flexibility and capacity of atomic fluctuation of the active site residues in the respective systems. The active site residues exhibit a certain level of rigidity, particularly residues Glu18, Leu20, Arg21 and Leu22 in the Src-UM-164 system. While a more prominent rigidity was observed in residue Gly126, Ala128, Ala143, Leu146 and Ile145. This could be attributed to the presence of multiple phenyl rings of UM-164 resulting in steric hindrance of residues in the vicinity.

Similarly, a slight decrease in fluctuation was also observed in residues Leu70, Leu75 and Val76 of Src-UM-164 system. This decrease may be associated with the presence of fluorinated phenyl group that limit the capacity of residues in the vicinity to fluctuate, consequently resulting in their rigidity. The observed residual fluctuations in Src-UM-164 relative to the Src-Dasatinib system may probably account for the experimental potency of UM-164 on the xenograft model.[3] Generally, apo system appears to be more flexible than Src-Dasatinib system, while the Src-UM-164 exhibits less fluctuation (except at the loops) compared to the two systems, therefore, appear to be more stable.

3.3 Radius of gyration (RoG)

The radius of gyration (RoG) is defined as a moment of inertia of Cα atoms from their centre of mass.[48] It has been applied to gain insight into molecular stability in the biological system during molecular dynamic simulations. The RoG of Src-UM-164, Src-Dasatinib and apo systems were evaluated and presented in **Figure 5**.

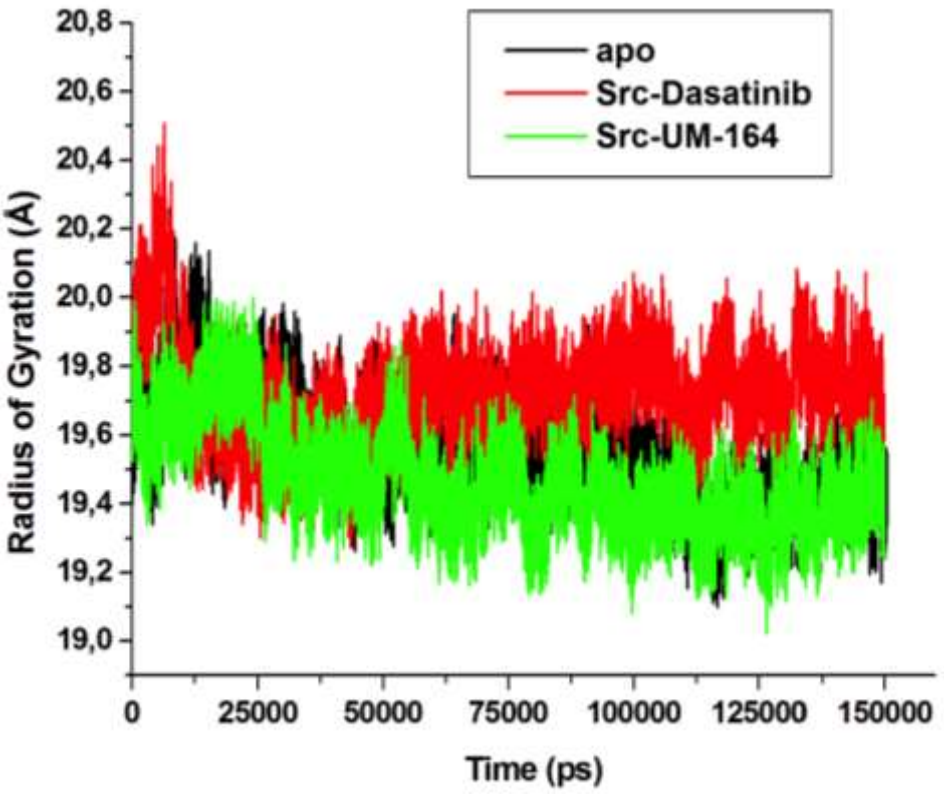


Figure 5. Radius of gyration plot of Cα of the apo (black), Src-Dasatinib (red) and Src-UM-164 (green) respectively.

In this study, we evaluate RoG of apo, Src-Dasatinib and Src-UM-164 systems. There is no remarkable difference in the average RoG of the three systems, as observed in Figure 5. The Src-UM-164 system exhibits a lower average RoG of 19.47 Å, whereas Src-Dasatinib and apo systems had an average RoG of 19.72 Å and 19.56 Å respectively. The asymmetric flexibility of residues in Src-Dasatinib and apo, and the tendency of residues in these systems to remain relatively flexible during the period of simulation, destabilise the mass centre of Src-Dasatinib and apo which results in an increase in average RoG and fluctuation.

3.4 Principal component analysis (PCA)

The protein conformation has been recognised as key features in the determination of biological function,[49] and PCA is one of the principal tools used in determining the flexibility of each atom during a simulation.[50] Here, clustering method of principal components (PC) was adopted because of its ability to describe different conformational states sampled during a simulation by grouping molecular structure into a subset based on their conformational similarities.[23] This method of PCA was used to assess the flexibility of apo, Src-Dasatinib and Src-UM-164 systems during 150 ns MD simulations. In order to gain insight into motions associated with the conformational behaviour of the subject systems, the systems were projected along the first

two principal components (PC1 vs PC2) or eigenvectors direction. The percentage variability or total mean square displacement of atom's positional fluctuation captured in each dimension is characterised by their corresponding eigenvalue.[38] **Figure 6** represents PCA plots of the subject systems (apo, Src-UM-164 and Src-Dasatinib systems) in this study.

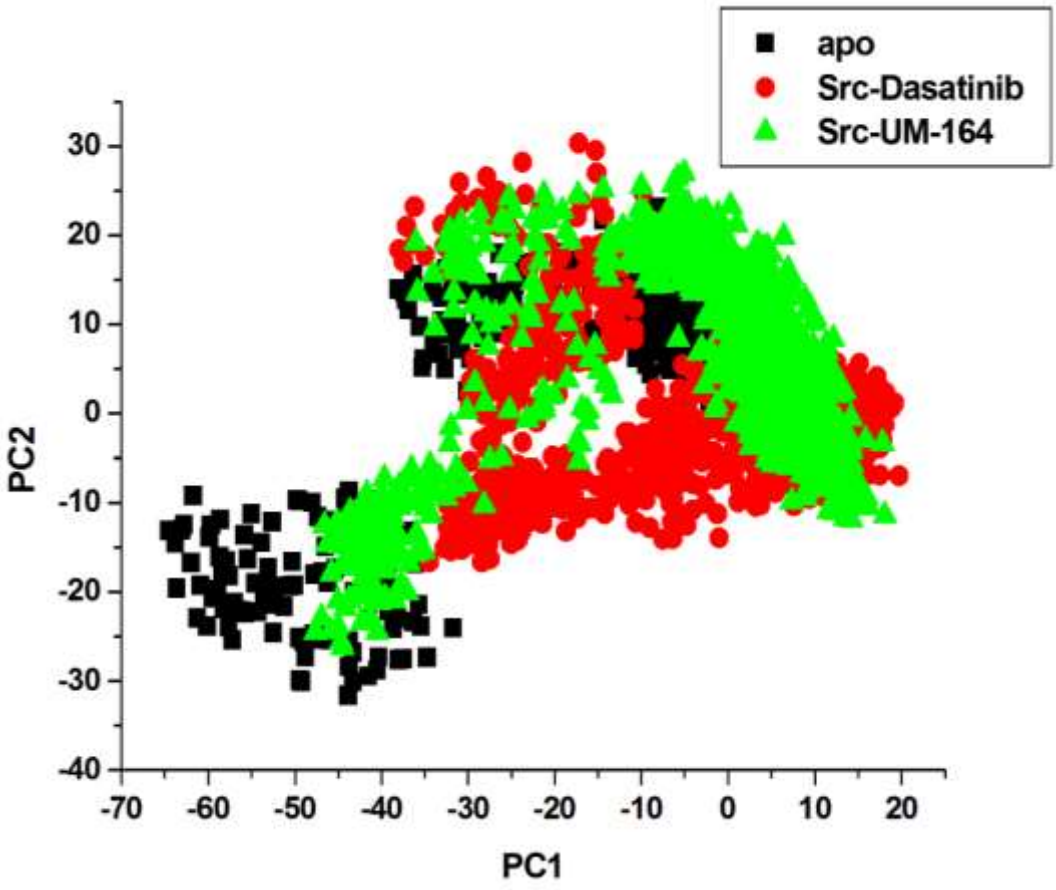


Figure 6. PCA projection of Cα atoms motion constructed by plotting the first two principal components (PC1 and PC2) in conformational space, apo (black) Src-Dasatinib (red) Src-UM-164 (green) respectively.

The PCA plot shows different and detailed waves of conformation in important subspace along the two principal components as evident in **Figure 6**. A distinct separation of motion was observed with apo, Src-UM-164 and Src-Dasatinib systems. However, a more correlated motion was observed in the Src-UM-164 system **(see Figure 6 A**) along the two principal components PC1 and PC2 compared to Src-Dasatinib and apo systems with relatively less correlated motion along PC1 and PC2 respectively. The apo system appears to be more flexible than Src-Dasatinib and Src-UM-164 systems, suggesting that the binding of the respective drugs, Dasatinib and UM-164, to the active site of the protein induces conformational dynamics which is reflected by the PCs as a wave of motion. Similarly, Src-UM-164 is more compacted

than Src-Dasatinib, meaning that the binding of UM-164 to the protein induced a more correlated motion compared to Src-Dasatinib and apo.

3.5 Dynamic cross-correlation matrices (DCCM) analysis

To further examine the conformational changes of Src protein upon the binding of UM-164 and Dasatinib, DCCM analysis was conducted on the positions of the Cα atoms throughout the simulations to determine the presence of correlated motions and the results are presented in **Figure 7**.

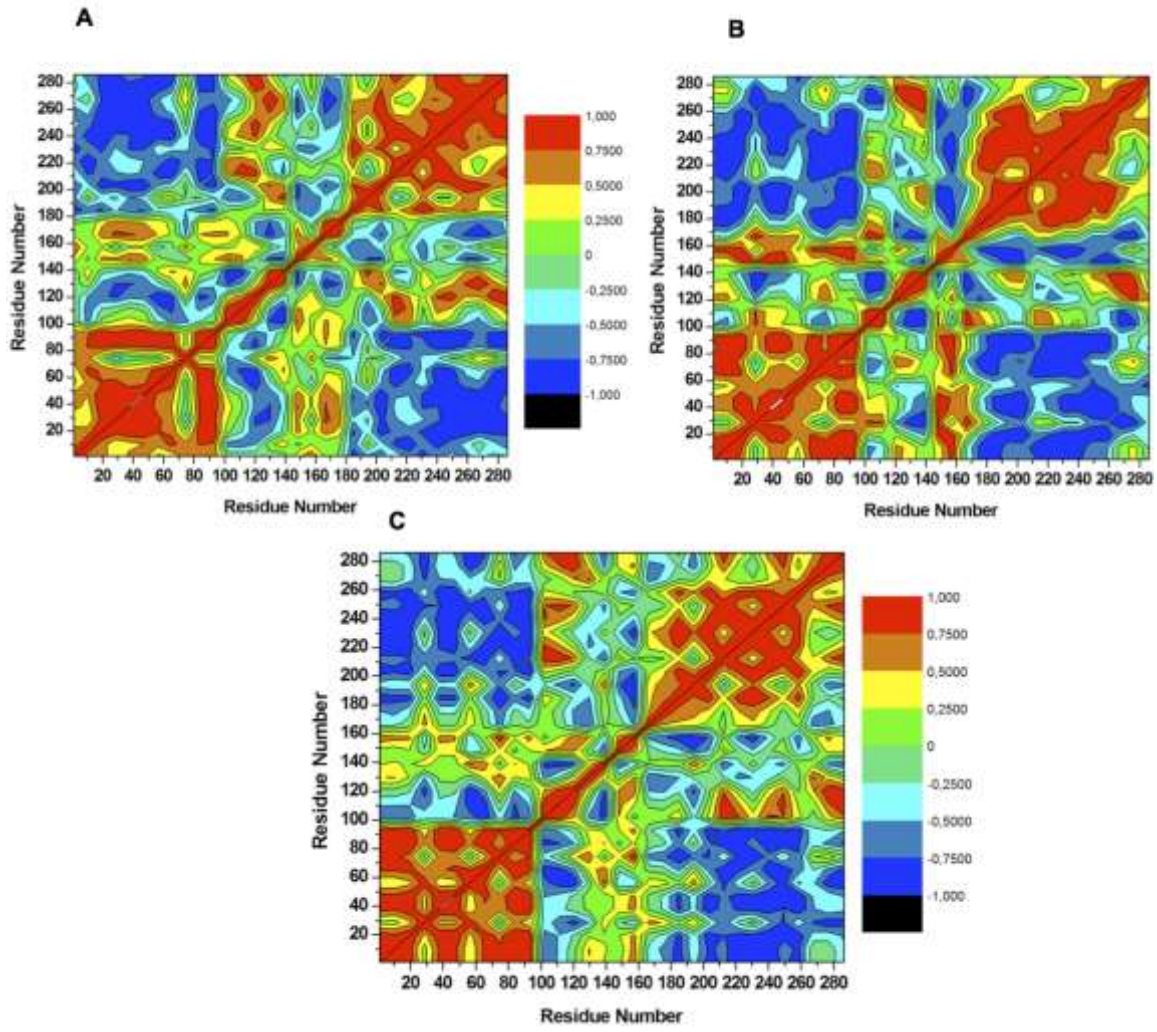


Figure 7. Cross-correlation matrices of Cα atoms fluctuations in apo (**A**), Src-Dasatinib (**B**) and Src-UM-164 (**C**).

The correlated motions (highly-positive) of specific residues are represented as a yellow-red (colour) region, whereas, anti-correlated (highly-negative) movement of specific residues are represented as blue-black (colour) regions. The three systems herein exhibited overall correlated residual motions relative to anticorrelated motions. DCCM analysis revealed that

binding of UM-164 and Dasatinib alter the structure conformation of Src as reflected by changes in the correlated motions and dynamics. In apo conformation (**Figure 7 A**), Met94 correlates with Ser95, while Ala156 slightly correlates with Val155 and Asp157. Anti-correlated residual motions in apo conformation occur between residues 200 - 280 compared to the Src-UM-164 system and Src-Dasatinib system with variable correlated and anticorrelated motions. The Src-UM-164 system exhibits two prominent correlated regions (**Figure 7 B**), namely; residue 1-95 strong correlated region, whereas 122-160 slight correlated region. These regions are among the most dynamic regions in the protein and the majority of hydrophobic active site residues reside within these regions. Upon Dasatinib binding, correlation strongly increases relative to the apo conformation suggesting that ligand binding induced residue dynamics that may have resulted in conformational changes in the protein. Anti-correlated motions are more prominent between residues 180-280 of the Src-Dasatinib system and such residues are located distal from the active site of the protein. Similarly, the Src-UM-164 system (**Figure 7 C**) also exhibit two prominent correlated regions namely; residues in region 1-100 representing strongly correlated residues and residue in region 120-160 represents slightly correlated residues. Higher correlated motions was observed among the residues of Src-UM-164 than Src-Dasatinib, this may probably be due to the conformational changes induced by UM-164 binding to Src relative to those induced by Dasatinib. Therefore, UM-164 impact a more significant conformational changes on the protein compared to Dasatinib, hence experimentally more potent than Dasatinib.

3.6 Hydrogen bond formation between amino acid residues

Hydrogens (H-bonds) are unique and universal in nature, they play a central role in biological systems and maintenance of the protein structural integrity,[51] protein-ligand interaction and catalysis.[51] H-bonds are reported to promote ligand binding affinity by displacing protein-bound water molecules into bulk solvents.[51] They are facilitators of protein-ligand binding,[51, 52] therefore, the formation of hydrogen bond between amino acid residues is key to the monitoring of protein conformation. In line with this, we investigate hydrogen bond formation during the course of the simulations. **Figure 8** shows a hydrogen bond formation over time during simulations of the respective systems.

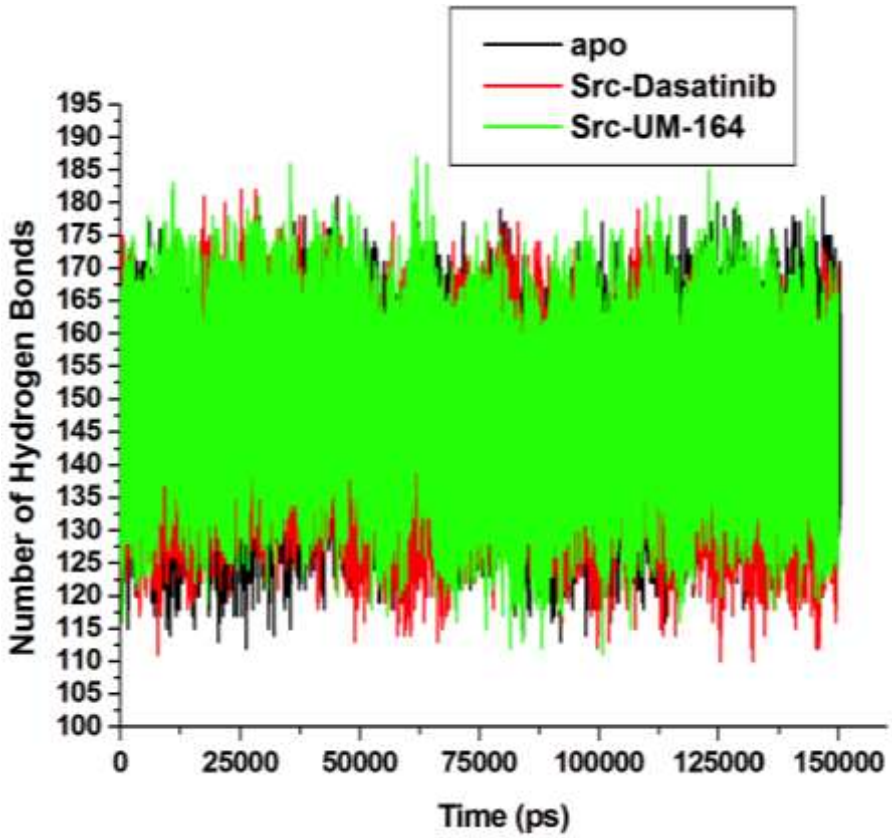


Figure 8. Number of hydrogen bond formation during simulation over time between Apo (black), Src-Dasatinib (red) and Src-UM-164 (green).

The apo system exhibited relatively lower average hydrogen bond formation during a simulation (144.978). However, the average number of H-bond formed in Src-Dasatinib (146.312) system is relatively higher than the value obtained in apo system. In a related development, average H-bond in Src-UM-164 system (148.584) is higher than the values in apo and Src-Dasatinib systems, this can be correlated to the system stability. Meaning that Src-UM-164 with higher average H-bond is more stable than the Src-Dasatinib and apo, it is evident that this result corroborate the result in RoG. Higher H-bond formation enhance receptor-ligand interaction[53] while a decrease in hydrogen bond formation leads to structural imbalances and conformational dynamics which eventually affect drug binding.[51]

To investigate the stability of Src-UM-164, Src-Dasatinib and apo systems, hydrogen bond distance and occupancy of the active site residues in the three systems were monitored throughout the course of simulations and the results are presented in **Table 1**.

Table 1. Hydrogen Bond Occupancy of interactive active site residues of UM-164, Src-Dasatinib and apo.

H-bond acceptor	H-bond donor	Frames (No)	Occupancy (%)	Average distance (Å)	Average angle (degree)
Src-UM-164					
Ala46@O	Thr91@H	129860	88	2.83	161
Val34@O	Leu26@H	128546	85	2.81	155
Met94@O	UM-164@H56	102854	68	2.84	153
Leu146@O	Lys154@H	94939	63	2.87	162
Glu63@O	Met67@H	79596	53	2.87	162
Lys48@O	Ile89@H	56321	37	2.90	161
Val155@O	Val76@H	40569	27	2.90	156
Leu26@O	Leu163@H	37529	25	2.87	161
Src-Dasatinib					
Met94@O	UM-164@H23	132566	88	2.82	159
Leu146@O	Lys154@H	119204	79	2.85	162
Ala46@O	Thr91@H	118906	79	2.86	162
Val34@O	Leu26@H	116780	77	2.81	154
Lys48@O	Ile89@H	59240	39	2.90	161
Val155@O	Val76@H	52682	35	2.89	156
Ser95@O	UM-164@H32	26359	17	2.83	159
Glu63@OE2	Lys48@HZ2	8463	5	2.77	157
Apo					
Val34@O	Leu26@H	106510	70	2.84	153
Thr91@O	Ala46@H	83764	55	2.88	161
Lys48@O	Ile89@H	77496	51	2.88	162
Glu63@O	Met67@H	72287	48	2.87	162
Val155@O	Val76@H	61045	40	2.88	158
Asp157@O	Asn144@H	54932	36	2.84	165
Leu26@O	Arg172@HH12	34006	22	2.82	154
Met94@O	Gly97@H	8882	5	2.91	151

Note: No = Number of frames; A = Angstrom; % = Percentage

The Src-UM-164 system brings into play active site residues with the H-bond that can be monitored within the course of system simulation. These residues include Lys48 and Val155 which exhibit maximum average H-bond distance of 2.90 Å each and Val34 with a minimum average H-bond distance of 2.81 Å. However, Ala46 and Val34 shows the highest occupancy of 88% and 85% respectively, demonstrating the importance of this residue in the Src-UM-164 system. Similarly, Met94 form H-bond with UM-164 and exhibit H-bond occupancy of 68% with a short average distance of 2.84 Å. The decrease in the distance marked the strength of the bond between Met94 and UM-164. In a related development, Met94 in Src-Dasatinib system exhibit highest occupancy of 88% with a lower average bond distance of 2.82 Å

compare to the Src-UM-164 system. This implies that Met94 formed a stronger H-bond with Dasatinib than with UM-164. Leu146 and Leu46 in Src-Dasatinib system exhibit the same H-bond occupancy of 79% and distances of 2.85 Å and 2.86 Å respectively, while Val34 shows the minimum average distance of 2.81 Å in the Src-Dasatinib system. In general, the H-bond occupancy in the Src-UM-164 system is slightly higher than the Src-Dasatinib system. The apo system has Met94 with maximum average H-bond of 2.91 Å and the least H-bond occupancy of 5% while Val34 had a minimum average H-bond distance of 2.84 Å and highest occupancy of 70%. In all three systems, the apo system shows the least H-bond distance and occupancy.

3.7 Residue Interaction Network Profile

Structural modification to improve bioavailability, enhance pharmacokinetic and dynamics (reduce adverse effects) of a drug molecule is one of the objectives of drug design. The mechanism of drug action involves the interaction of receptor specific active site residue with the specific groups in the drug molecule, this result in signal transduction and consequently induction of specific reaction interpreted as drug action. Therefore, it serves a purpose to examine the drug-receptor interaction to gain insight into the role and the type of interaction that is common to residues. Here, it was observed that most of the active site residues form the hydrophobic interaction with the ligand in Src-Dasatinib and Src-UM-164 systems (**Figure 9 A and B**). The UM-164 orientation formed the hydrogen bond with the active site residue Met94 and Thr91 (**Figure 9 B**). Similarly, Dasatinib oriented position formed the hydrogen bond with the active site residues Met94 and Lys96 (**Figure 9 A**). The interactive OH group of UM-164 and Dasatinib formed a hydrogen bond with Met94 and Lys96, therefore, these residues are essential for the binding of the UM-164 and Dasatinib to the active site of Src. The fluorinated phenyl group of UM-164 interacts with hydrophobic residues Asp157 and Met67 of Src. This interaction is completely missing in the Src-Dasatinib system. This may probably explain a high experimental potency exhibited by UM-164. Generally, UM-164 interact with more hydrophobic active site residues compare to Dasatinib this may probably be due to the presence of additional benzene ring in the structure of UM-164. However, hydrophobic interaction is the major interaction between the drug and active site residues in the two systems.

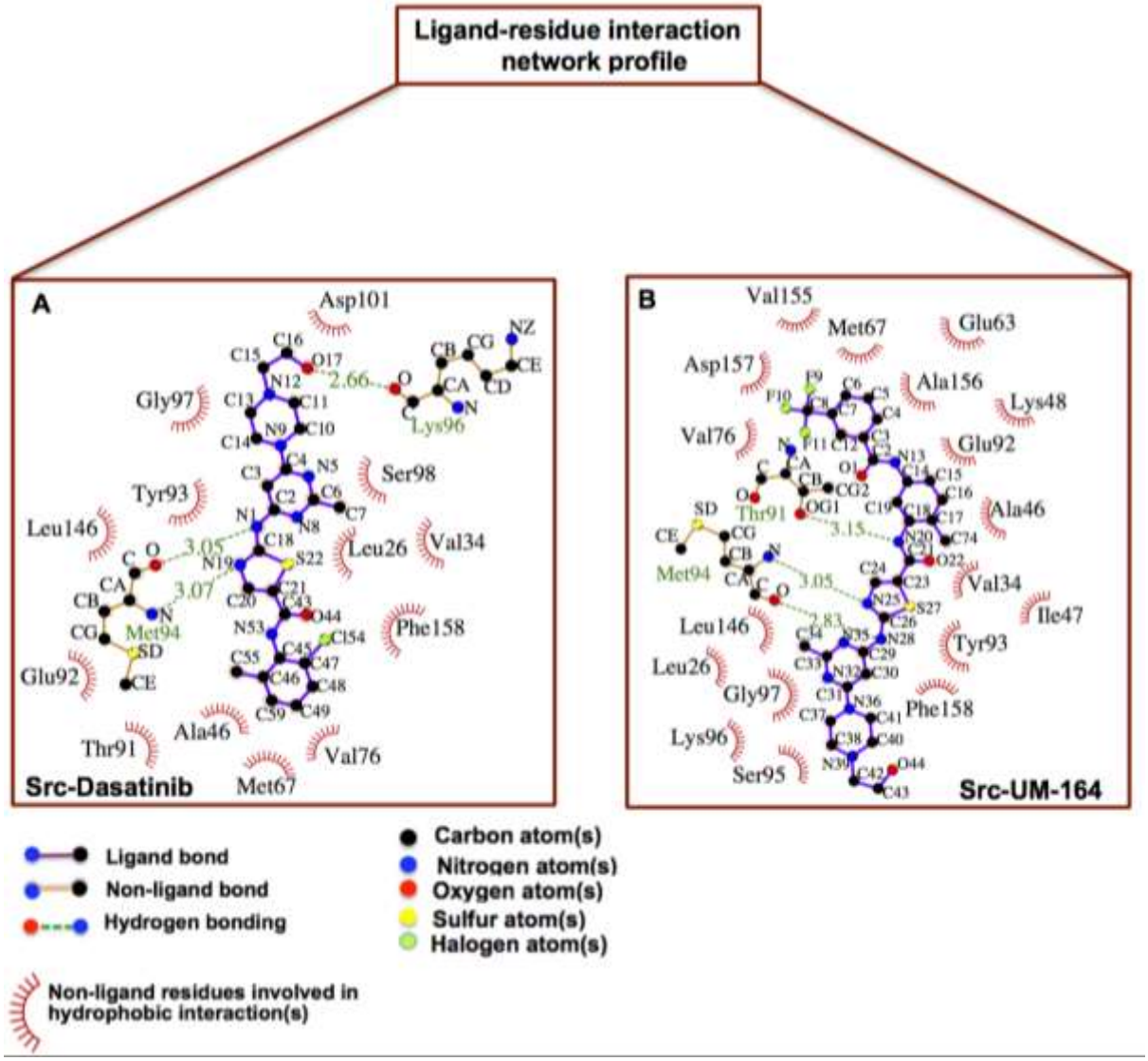


Figure 9. 2D Ligand-residue interactions network from fully-minimised complex of Src-Dasatinib (**A**) and Src-UM-164 (**B**).

3.8 Binding free energy and energy decomposition analyses

Molecular Mechanics/Generalized-Born Surface Area (MM/GBSA) method[39] is a popular approach used to estimate the binding free energy of small ligands to biological macromolecules.[39] This method was used to estimate the total binding energy of UM-164 and Dasatinib to Src and the results were presented in **Table 2.**

Table 2. MM/GBSA based binding free energies profile of Src-UM-164 and Src-Dasatinib.

ΔG_{bind}	ΔE_{ele}	ΔE_{vdW}	ΔG_{gas}	ΔG_{sol}
Src-UM-164				
-82.293±4.320	-121.874±13.653	-83.311±3.332	-205.184±13.836	122.890±12.879
Src-Dasatinib				
-52.610±4.282	-86.520±13.911	-56.780±3.099	-143.301±14.909	90.691±13.276

Notes: ΔE_{ele} = electrostatic energy; ΔE_{vdw} = van der Waals energy; G_{bind} = calculated total binding free energy; G_{sol} = solvation free energy.

It was revealed that the estimated binding free energy is higher in Src-UM-164 (-82.293 kcal/mol) compared to Src-Dasatinib (-52.610 kcal/mol). The difference in binding free energy (-29.683 kcal/mol) between Src-UM-164 and Src-Dasatinib systems is quite significant, meaning that the force of interactions contributes higher energy in the binding of UM-164 to c-Src compare Dasatinib. The Src-UM-164 system also exhibited a relatively higher electrostatic energy (-121.874) contribution to the total binding free energy compared to the Src-Dasatinib system with maximum electrostatic energy contribution of -86.520 kcal/mol. However, van der Waal contribution to the total binding free energy was higher in Src-UM-164 (-83.311 kcal/mol) compare to Src-Dasatinib with van der Waal contribution of -56.780 kcal/mol. Hydrophobic packing contributes significantly to binding free energy in Src-UM-164 owing to a large amount of aromatic and hydrophobic rings within the conformational space, as well as a set of hydrophobic residues around the binding pocket. This is evident in the binding free energy contribution of the Src-UM-164 system, where vdW contribution is relatively higher than the total binding free energy as shown in **Table 2**.

3.8.1 Per-residue energy decomposition analysis

Binding of the ligand to the active site of the receptor is associated with a certain amount of energy contributed by individual residues involve in the interaction. To assess the energy contribution of individual active site residues to the total binding free energy, and to provide a molecular understanding of the impact of protein dynamics on the degree of different binding forces, per-residue energy decomposition analysis was conducted and the result presented in **Table 3**

Table 3. Decomposition of the relative binding free energies on a per-residue basis for Src-Dasatinib and Src-UM-164 systems.

Residues	ΔE_{vdw}	Δ_{ele}	ΔG_{polar}	$\Delta G_{non\text{-}polar}$	$\Delta G_{binding}$
Src-UM-164					
Met94	-1.244±0.612	-6.662±0.838	4.548±0.616	-0.070±0.020	-3.718±0.575
Thr91	-1.670±0.471	-2.174±0.702	1.106±0.303	-0.147±0.022	-2.885±0.684
Lys48	-2.438±0.450	8.176±1.417	-8.449±1.847	-0.159±0.018	-2.871±1.269
Phe158	-2.330±0.398	-0.294±0.398	0.277±0.293	-0.120±0.029	-2.467±0.488
Asp157	-2.425±0.509	-9.667±0.923	9.924±1.031	-0.207±0.023	-2.374±0.548
Leu146	-2.077±0.258	0.724±0.132	-0.825±0.115	-0.174±0.023	-2.353±0.279
Met67	-2.137±0.340	-0.388±0.416	0.400±0.260	-0.176±0.031	-2.301±0.402
Tyr93	-2.418±0.381	-1.580±0.950	1.872±0.760	-0.169±0.041	-2.296±0.556
Leu26	-2.169±0.414	-0.371±0.258	0.644±0.316	-0.242±0.067	-2.139±0.406
Gly97	-1.456±0.343	-0.042±0.527	-0.089±0.454	-0.137±0.026	-1.724±0.555
Src-Dasatinib					
Met94	-1.043±0.643	-6.386±0.968	3.534±0.567	-0.066±0.019	-3.962±0.569
Thr91	-1.609±0.452	-1.495±1.025	1.052±0.313	-0.150±0.024	-2.201±0.968
Leu26	-2.217±0.451	-0.385±0.301	0.724±0.385	-0.318±0.051	-2.195±0.435
Tyr93	-2.459±0.393	-1.268±0.820	1.736±0.712	-0.183±0.049	-2.174±0.488
Leu146	-1.675±0.349	0.725±0.139	-0.751±0.134	-0.114±0.036	-1.813±0.383
Gly97	-1.471±0.227	0.134±0.501	0.065±0.414	-0.137±0.029	-1.409±0.338
Val34	-0.946±0.234	-0.306±0.161	0.183±0.145	-0.069±0.024	-1.138±0.247
Ala46	-1.417±0.259	0.263±0.239	0.138±0.217	-0.111±0.020	-1.127±0.295
Val76	-0.866±0.275	0.388±0.097	-0.460±0.162	-0.046±0.019	-0.983±0.230
Glu92	-0.576±0.126	-13.331±0.628	13.042±0.510	-0.010±0.011	-0.874±0.335

Notes: ΔE_{ele} = electrostatic energy (kcal/mol); ΔE_{vdW} = van der Waals energy (kcal/mol); $\Delta G_{polar\ (kcal/mol)}$ = polar solvation energy (kcal/mol); $\Delta G_{nonpolar\ (kcal/mol)}$ = nonpolar solvation energy (kcal/mol); $\Delta G_{binding\ (kcal/mol)}$ = total binding free energy $_{(kcal/mol)}$.

The binding free energy was decomposed into the unit contributions of each active site residue of Src-UM-164 and Src-Dasatinib systems. The major energy contributors in Src-UM-164 are Met94, Thr91 and Lys48 with -3.718, -2.885 and -2.871 kcal/mol respectively. While in Src-Dasatinib system, the major energy contributors are Met94, Thr91 and Leu26 contributing -3.962, -2.201 and -2.195 kcal/mol respectively. Met94 exhibits exemplary contributions to the total binding free energy in the two system by contributing highest binding free energy to the system making it an important residue the binding of UM-164 and Dasatinib their respective proteins. This result corroborate the result of H-bond formation where met94 exhibit high H-bond occupancy and strength. However, other active site residues in Src-UM-164 systems (**Figure 10**) show significantly higher binding free energy contribution relative to their contribution in Src-Dasatinib systems, a probable reason for a high potency of UM-164 over

Dasatinib in c-Src inhibition. Lys48 in UM-164 is associated with low electrostatic contribution to the total binding free energy (8.176 kcal/mol) and relatively higher van der Waals energy (-2.438 kcal/mol). Similarly, Asp157 had the highest electrostatic contribution of -9.667 kcal/mol and high van der Waals energy contribution of -2.425 kcal/mol to the total binding free energy in Src-UM-164 compared to other active site residues in the same system. This may be due to its interaction with a fluorinated phenyl group in UM-164. Therefore, Met94 and Asp157 are the key residues in Src-UM-164 system contributing major electrostatic forces to the total binding free energy. In contrast to Src-UM-164 system, the per-residue energy contributions from active site residues in Src-Dasatinib exhibit variable electrostatic energy contribution to the total binding free energy. Met94 though contributes the highest electrostatic energy (-6.386 kcal/mol) in Src-Dasatinib system, its contribution is relatively higher in Src-UM-164 system compared to Src-Dasatinib system. The high electrostatic energy contribution from Met94 and Asp157 may probably place UM-164 at advantage over Dasatinib in Src inhibitory profile.

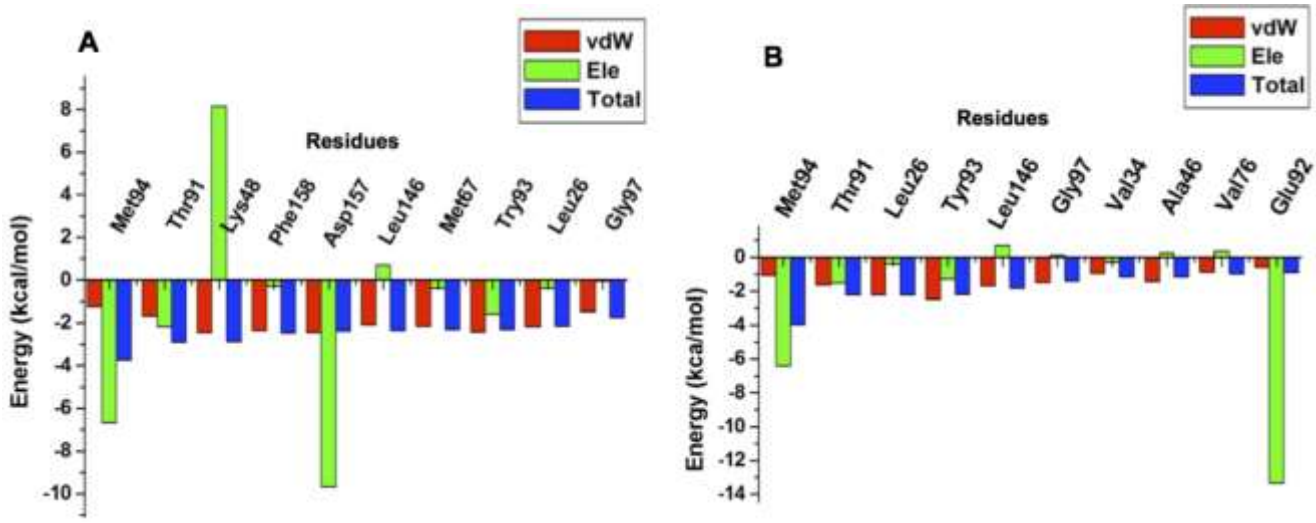


Figure 10. The per-residue energy decomposition analysis graph of Src-UM-164 (A) and Src-Dasatinib (B) system respectively.

3.9 Comparative toxicity test of UM-164 and Dasatinib

Most chemical compounds interact with biological targets different from experimentally recognised targets, such compounds are said to be promiscuous in nature.[54] Compound promiscuity depicts the molecular basis of the pharmacological effects, therefore, assessment of the extent of promiscuity among compounds at different levels of drug research[55] provides a detailed understanding of other properties of the drug that were probably not anticipated.

Toxicity and biological activity of compounds can be predicted by computer tools. It has been shown that the degree of reliability of such tools varies from one to another.[56,57] In this study, the "prediction of activity spectra for biologically active substance" (PASS) [56, 57] programme was employed to compare the possible toxicity of UM-164 and Dasatinib. The PASS predicts compound toxicity and biological profile with a mean accuracy of prediction of about 89% to 90%.[56, 57] **Table 4** shows the predicted toxicity and biological activity of UM-164 and Dasatinib.

Table 4. Comparative predicted toxicity and biological activity of UM-164 and Dasatinib using PASS.

Pa	Pi	Activity	Pa	Pi	Activity
Predicted toxicity for UM-164			**Predicted toxicity for Dasatinib**		
0.781	0.006	Optic neuritis	0.861	0.009	Asthma
0.765	0.009	Optic neuropathy	0.834	0.004	Optic neuropathy
0.730	0.005	Myocarditis	0.828	0.003	Optic neuritis
0.645	0.037	Stomatitis	0.831	0.016	Stomatitis
0.597	0.044	Thrombophlebitis	0.805	0.004	Myocarditis
0.582	0.050	Asthma	0.758	0.013	Thrombophlebitis
0.464	0.052	Cholestasis	0.756	0.019	Hepatitis
0.485	0.106	Hepatitis	0.667	0.038	Tachycardia
0.479	0.123	Conjunctivitis	0.638	0.015	Cholestasis
0.313	0.223	Tachycardia	0.645	0.027	Broncho constrictor
0.315	0.211	Thrombocytopenia	0.626	0.031	Agranulocytosis
0.330	0.212	Consciousness alteration	0.626	0.053	Conjunctivitis
0.346	0.164	Broncho constrictor	0.600	0.052	Thrombocytopenia
0.350	0.177	Inflammation	0.542	0.048	Tremor
0.360	0.121	Agranulocytosis	0.391	0.143	Ocular toxicity

Note: Pa probability of compound being active, Pi probability of compound being inactive. Pa > 0.7 indicates probability of toxicity or biological activity, Pi < 0.5 the compound is unlikely to exhibit toxicity or biological activity, 0.5 < Pa < 0.7 the compound is likely to exhibit toxicity or biological activity but the probability is less and < 0.5 the compound is unlikely to exhibits the activity on experiment.

The predicted toxicity value revealed that UM-164 can induce optic neuritis, optic neuropathy and myocarditis, and is likely to cause conditions such as stomatitis, thrombophlebitis and asthma. These effects are likely to occur at a relatively high dose. The drug is unlikely to cause cholestasis, hepatitis, conjunctivitis, tachycardia, thrombocytopenia, consciousness alteration and bronchoconstriction, therefore potentially safe. Similarly, Dasatinib can induce asthma, optic neuropathy, optic neuritis, stomatitis, myocarditis, thrombophlebitis and hepatitis. Comparatively, UM-164 is likely to exhibit lower side effects compared Dasatinib. PASS revealed that Dasatinib is likely to predispose patients to hepatitis, bronchoconstriction, myocarditis, cholestasis, thrombocytopenia and consciousness alteration these effects are

unlikely to be induced by UM-164 thus, UM-164 may be superior to Dasatinib in terms of safety.

4 Conclusion

The discovery of a novel Src inhibitor, UM-164, has given hope to the discovery and designing of potent Src/p38 inhibitors for the treatment of TNBC. However, understanding of conformational features of Src induced by UM-164, which could open alternative avenues in the treatment of TNBC, are untapped. Herein, different computational approaches aimed at providing an in-depth understanding of the influence of UM-164 binding to Src and the resultant conformational features were explored. Herein, comparative MD simulations of Src-UM-164 and Src-Dasatinib with post-MD analytical approaches including PCA, RoG, thermodynamic calculations, DCCM, ligand-residue interaction network profile and predictive toxicity assessment, were conducted. Findings from this study revealed that the subject systems were relatively stable throughout the simulations with no remarkable difference in their average RoG. The binding of UM-164 to Src induces a more stable and compact conformation of the protein structure, compared to Dasatinib. UM-164 binding induces a more correlated motion in Src relatively to Dasatinib suggesting that ligand binding may have induced residue dynamics that results in conformational changes in the protein. High fluctuation exhibited by the loops in Src-UM-164 system support the experimental evidence that UM-164 binds the DFG-out inactive conformation of Src. The estimated binding free energy is higher in Src-UM-164 compared to Src-Dasatinib, this reflects the relative higher binding capacity of UM-164 to the Src. Hydrophobic packing contributes significantly to binding free energy in Src-UM-164 owing to a large amount of aromatic and hydrophobic rings within the conformational space, as well as a set of hydrophobic residues around the binding pocket of the system. Within the context of this study, Met94 was found to exhibits exemplary contributions to the total binding free energy in the Src-UM-164 and Src-Dasatinib systems by contributing highest binding free energy making it an important residue for the binding of UM-164 and Dasatinib. Src-UM-164 was also found to exhibits higher average H-bond formation making it a more stable system compared to Src-Dasatinib and apo systems. Met94 formed H-bond with UM-164 and Dasatinib with relatively higher occupancy in Src-Dasatinib compare to Src-UM-164. The orientation of UM-164 in the active site of Src allows for hydrophobic interaction with a fluorinated phenyl group, thus may contribute to high potency reported experimentally. The interactive OH group of UM-164 and Dasatinib forms a hydrogen bond with Met94 and Lys96,

implying that these residues are essential for the binding of the UM-164 and Dasatinib to the receptor active site.

UM-164 is potentially safer than Dasatinib regarding the toxicity, thus superior to Dasatinib in both its clinical efficacy and safety. Considering the overall findings from this study, the conformational features of Src-UM-164 system provided in this study can serve as a base-line in the design of novel Src inhibitors with dual inhibitory properties.

Acknowledgements

The authors acknowledge the School of Health Science, University of KwaZulu-Natal, Westville campus for financial assistance. The Centre for High Performance Computing (CHPC, www.chpc.ac.za) Cape Town, South Africa, for computational resources

Disclosure

Authors declare no financial and intellectual conflict of interests.

References

(1) Sharma, G. N., Dave, R., Sanadya, J., Sharma, P., and Sharma, K. K. (2010) Various types and management of breast cancer: an overview. *J. Adv. Pharm. Technol. Res. 1*, 109–26.

(2) Siegel, R. L., Miller, K. D., and Jemal, A. (2016) Cancer statistics *66*, 7–30.

(3) Gilani, R. A., Phadke, S., Bao, L. W., Lachacz, E. J., Dziubinski, M. L., Brandvold, K. R., Steffey, M. E., Kwarcinski, F. E., Graveel, C. R., Kidwell, M., Merajver, S. D., and Soellner, M. B. (2016) UM-164: a potent c-Src/p38 kinase inhibitor with *invivo* activity against triple negative breast cancer. *Clin Cancer Res. 5087–96*.

(4) Anders, C. K., and Carey, L. a. (2010) Biology, Metastatic Patterns and Treatment of Patients with Triple-Negtive Breast Cancer. *Breast 9*, S73–S81.

(5) Jiao, Q., Wu, A., Shao, G., Peng, H., Wang, M., Ji, S., Liu, P., and Zhang, J. (2014) The latest progress in research on triple negative breast cancer (TNBC): Risk factors, possible therapeutic targets and prognostic markers. *J. Thorac. Dis. 6*, 1329–1335.

(6) Gluz, O., Liedtke, C., Gottschalk, N., Pusztai, L., Nitz, U., and Harbeck, N. (2009) Triple-negative breast cancer--current status and future directions. *Ann. Oncol. 20*, 1913–1927.

(7) Jafarzadeh, N., Ashraf, H., Khoshroo, F., Sepehri Shamloo, A., Bidouei, F., and Ghaffarzadehgan, K. (2015) Triple Negative Breast Cancer: Molecular Classification, Prognostic Markers and Targeted Therapies. *Razavi Int. J. Med. 3*.

(8) Finn, R. S., Bengala, C., Ibrahim, N., Roche, H., Sparano, J., Strauss, L. C., Fairchild, J., Sy, O., and Goldstein, L. J. (2011) Dasatinib as a Single Agent in Triple-Negative Breast Cancer: Results of an Open-Label Phase 2 Study. *Clin. Cancer Res. 17*, 6905–6913.

(9) Cheng, G., Fan, X., Hao, M., Wang, J., Zhou, X., and Sun, X. (2016) Higher levels of TIMP-1 expression are associated with a poor prognosis in triple-negative breast cancer. *Mol. Cancer 15*, 30.

(10) Bayraktar, S., and Gl??ck, S. (2013) Molecularly targeted therapies for metastatic triple-negative breast cancer. *Breast Cancer Res. Treat. 138*, 21–35.

(11) Wang, Y., Waters, J., Leung, M. L., Unruh, A., Roh, W., Shi, X., Chen, K., Scheet, P., Vattathil, S., Liang, H., Multani, A., Zhang, H., Zhao, R., Michor, F., Meric-Bernstam, F., and Navin, N. E. (2014) Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature 512*, 1–15.

(12) Finn, R. S., Dering, J., Ginther, C., Wilson, C. a, Glaspy, P., Tchekmedyian, N., and Slamon, D. J. (2007) Dasatinib, an orally active small molecule inhibitor of both the src and abl kinases, selectively inhibits growth of basal-type/'triple-negative' breast cancer cell lines growing in vitro. *Breast Cancer Res. Treat. 105*, 319–326.

(13) Finn, R. S. (2008) Targeting Src in breast cancer. *Ann. Oncol. 19*, 1379–1386.

(14) Mayer, E. L., and Krop, I. E. (2010) Advances in targeting Src in the treatment of breast cancer and other solid malignancies. *Clin. Cancer Res. 16*, 3526–3532.

(15) Getlik, M., Gr??tter, C., Simard, J. R., Kl??ter, S., Rabiller, M., Rode, H. B., Robubi, A., and Rauh, D. (2009) Hybrid compound design to overcome the gatekeeper T338M mutation in cSrc. *J. Med. Chem. 52*, 3915–3926.

(16) Foà, R., Vitale, A., Vignetti, M., Meloni, G., Guarini, A., De Propris, M. S., Elia, L., Paoloni, F., Fazi, P., Cimino, G., Nobile, F., Ferrara, F., Castagnola, C., Sica, S., and Leoni, P. Dasatinib as first-line treatment for adult patients with Philadelphia chromosome–positive acute lymphoblastic leukemia. *Blood.* 118, 6521-6528.

(17) Boyle, P. (2012) Triple-negative breast cancer: Epidemiological considerations and recommendations. *Ann. Oncol. 23*, 8–13.

(18) Tryfonopoulos, D., Walsh, S., Collins, D. M., Flanagan, L., Quinn, C., Corkery, B., McDermott, E. W., Evoy, D., Pierce, A., O'Donovan, N., Crown, J., and Duffy, M. J. (2011) Src: A potential target for the treatment of triple-negative breast cancer. *Ann. Oncol. 22*, 2234–2240.

(19) Edward L. T and Kathy J. H (2016) Strengthening Global Partnership in Breast Cancer Research. *JGO*.1–9.

(20) Klepeis, J. L., Lindorff-Larsen, K., Dror, R. O., and Shaw, D. E. (2009) Long-timescale molecular dynamics simulations of protein structure and function. *Curr. Opin. Struct. Biol. 19*, 120–127.

(21) Charles C. David and Donald J. Jacobs (2014) Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins. *Methods Mol Biol*. 1084, 193–226

(22) Maisuradze, G., Liwo, a, and Scheraga, H. (2009) Principal component analysis for protein folding dynamics. *J. Mol. Biol. 385*, 312–329.

(23) Wolf, A., and Kirschner, K. N. (2013) Principal component and clustering analysis on molecular dynamics data of the ribosomal L11??23S subdomain. *J. Mol. Model. 19*, 539–549.

(24)M Kumalo and M .E Soliman (2015) Per-Residue Energy Footprints-Based Pharmacophore Modeling as an Enhanced In Silico Approach in Drug Discovery : A Case Study on the Identification of Novel b -Secretase1 (BACE1) Inhibitors as Anti- Alzheimer Agents. *Cellular and Molecular Bioengineering.*1-9

(25) Kwarcinski, F. E., Brandvold, K. R., Phadke, S., Beleh, O. M., Johnson, T. K., Meagher, J. L., Seeliger, M. A., Stuckey, J. A., and Soellner, M. B. (2016) Conformation-Selective Analogues of Dasatinib Reveal Insight into Kinase Inhibitor Binding and Selectivity. *ACS Chem. Biol. 11*, 1296–1304.

(26) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem. 25*, 1605–1612.

(27) Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeerschd, T., Zurek, E., and Hutchison, G. R. (2012) Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform. 4*, 1–17.

(28) Walker, R., Supercomputer, S. D., and Roitberg, A. AMBER 14 and GPUs : Creating the World ' s Fastest MD Package.

(29) Goetz, A. W., Williamson, M. J., Xu, D., Poole, D., Grand, S. L., and Walker, R. C. (2012) Routine microsecond molecular dynamics simulations with amber - part i: Generalized born. *J. Chem. Theory Comput. 8*, 1542–1555.

(30) Galindo-Murillo, R., Robertson, J. C., Zgarbova Jir, M., Otyepka, M., Jurec, P., and Cheatham, T. E. (2016) Assessing the Current State of Amber Force Field Modifications for

DNA. *J. Chem. Theory Comput 12*, 4114–4127.

(31) Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., and Shaw, D. E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct. Funct. Bioinforma. 78*, 1950–1958.

(32) Wang, J. M., Wolf, R. M., Caldwell, J. W., Kollman, P. a, and Case, D. a. (2004) Development and testing of a general amber force field. *J. Comput. Chem. 25*, 1157–1174.

(33) Case, D. A., Darden, T., Iii, T. E. C., Simmerling, C., Brook, S., Roitberg, A., Wang, J., Southwestern, U. T., Duke, R. E., Hill, U., Luo, R., Irvine, U. C., Roe, D. R., Walker, R. C., Legrand, S., Swails, J., Cerutti, D., Kaus, J., Betz, R., Wolf, R. M., Merz, K. M., State, M., Seabra, G., Janowski, P., Paesani, F., Liu, J., Wu, X., Steinbrecher, T., Gohlke, H., Homeyer, N., Cai, Q., Smith, W., Mathews, D., Salomon-ferrer, R., Sagui, C., State, N. C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., Berryman, J., and Kollman, P. A. (2014) Amber 14. *Univ. California, San Fr.*

(34) Finnerty, J. (2011) Molecular dynamics meets the physical world : Thermostats and barostats.

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

(36) Roe, D. R., and Cheatham III, T. E. (2013) PTRAJ and CPPTRAJ: software for processing and analysis of molecular synamics trajectory data. *J Chem Theory Com 9*, 3084–3095.

(37) E. Seifert (2014) OriginPro 9.1: Scientific Data Analysis and Graphing Software—Software Review. *J. Chem. Inf. Model.* 1552-1552.

(38) Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A., and Caves, L. S. D. (2006) Bio3d: An R package for the comparative analysis of protein structures. *Bioinformatics 22*, 2695–2696.

(39) Genheden, S., and Ryde, U. (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov. 10*, 449–61.

(40) Arnold, G. E., and Ornstein, R. L. (1997) Molecular dynamics study of time-correlated protein domain motions and molecular flexibility: cytochrome P450BM-3. *Biophys. J. 73*, 1147–1159.

(41) Swaminathan, S. (1991) Investigation of domain structure in proteins via molecular dynamics simulation: application to HIV-1 protease dimer. *J. Am. ...* 2717–2721.

(42) Carugo, O. (2001) A normalized root-mean-square distance for comparing protein three-dimensional structures/. *Protein Sci.* 1470–1473.

(43) Alberts B, Johnson A, Lewis J, et al (2002) Molecular Biology of the Cell 4th edition. *New York: Garland Science.*

(44) Loeffler, H. H., and Winn, M. D. (2013) Ligand binding and dynamics of the monomeric epidermal growth factor receptor ectodomain. *Proteins Struct. Funct. Bioinforma. 81*, 1931–1943.

(45) Ejaz Ahmad , Gulam Rabbani , Nida Zaidi , Mohammad Azam Khan , Atiyatul Qadeer , Mohd Ishtikhar , Saurabh Singh & Rizwan Hasan Khan. (2013) Revisiting ligand-induced conformational changes in proteins: essence, advancements, implications and future

challenges. *Journal of Biomolecular Structure and Dynamics*. 630-648

(46) Vendome, J., Posy, S., Jin, X., Bahna, F., Ahlsen, G., Shapiro, L., and Honig, B. (2011) Molecular design principles underlying β-strand swapping in the adhesive dimerization of cadherins. *Nat. Struct. Mol. Biol. 18*, 693–700.

(47) Pucheta-Martínez, E., Saladino, G., Morando, M. A., Martinez-Torrecuadrada, J., Lelli, M., Sutto, L., D'Amelio, N., and Gervasio, F. L. (2016) An Allosteric Cross-Talk Between the Activation Loop and the ATP Binding Site Regulates the Activation of Src Kinase. *Sci. Rep. 6*, 24235.

(48) Mhlongo, N. N., Ebrahim, M., Skelton, A. A., Kruger, H. G., Williams, I. H., and Soliman, M. E. S. (2015) Dynamics of the thumb-finger regions in a GH11 xylanase Bacillus circulans: comparison between the Michaelis and covalent intermediate. *Rsc Adv. 5*, 82381–82394.

(49) Benson, N. C., and Daggett, V. (2008) Dynameomics: large-scale assessment of native protein flexibility. *Protein Sci. 17*, 2038–50.

(50) Desdouits, N., Nilges, M., and Blondel, A. (2015) Principal Component Analysis reveals correlation of cavities evolution and functional motions in proteins. *J. Mol. Graph. Model. 55*, 13–24.

(51) Chen, D., Oezguen, N., Urvil, P., Ferguson, C., Dann, S. M., and Savidge, T. C. (2016) Regulation of protein-ligand binding affinity by hydrogen bond pairing. *Sci. Adv. 2*, e1501240.

(52) Patil, R., Das, S., Stanley, A., Yadav, L., and Sudhakar, A. (2010) Optimized Hydrophobic Interactions and Hydrogen Bonding at the Target-Ligand Interface Leads the Pathways of Drug-Designing *PLoS One* 16;5(8):e12029

(53) Chen, D., Oezguen, N., Urvil, P., Ferguson, C., Dann, S. M., and Savidge, T. C. (2016) Regulation of protein-ligand binding affinity by hydrogen bond pairing. *Sci. Adv. 2*, e1501240.

(54) Hu, Y., Gupta-Ostermann, D., and Bajorath, J. (2014) Exploring compound promiscuity patterns and multi-target activity spaces. *Comput. Struct. Biotechnol. J. 9*, e201401003.

(55) Hu, Y., Gupta-Ostermann, D., and Bajorath, J. (2014) Exploring Compound Promiscuity Patterns and Multi-Target Activity Spaces. *Comput. Struct. Biotechnol. J. 9*, 1–11.

(56) Lagunin, A., Stepanchikova, A., Filimonov, D., and Poroikov, V. (2000) PASS: prediction of activity spectra for biologically active substances. *Bioinformatics 16*, 747–8.

(57) Poroikov, V., Filimonov, D., Lagunin, A., Gloriozova, T., and Zakharov, A. (2007) PASS: identification of probable targets and mechanisms of toxicity. *Sar Qsar Environ. Res. 18*, 101–110.

CHAPTER 6

The impact of Thr91 mutation on c-Src resistance to UM-164: Molecular dynamics study revealed new opportunity for drug design

Umar Ndagi[a] Ndumiso N. Mhlongo[a] and Mahmoud E. Soliman[a]*

[a] Molecular Modelling and Drug Design Research Group, School of Health Sciences, University of KwaZulu-Natal, Westville, Durban 4000, South Africa

*Corresponding author: Mahmoud E. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27(0)312608048, Fax: +27 (0)31260 7872

Webpage: http://soliman.ukzn.ac.za/

Abstract

The emergence of drug resistance non-receptor tyrosine kinase (c-Src) in triple-negative breast cancer (TNBC) remains a prime interest in relation to the burden of TNBC among people living with breast cancer and drug development. Thr91 mutation was found to induce a complete loss of protein conformation required for drug fitness. Herein, we provide the first account of the molecular impact of the Thr91 mutation on c-Src resistance to experimental drug UM-164 using various computational approaches, including molecular dynamics simulation, principal component analysis (PCA), dynamic cross-correlation matrices (DCCM) analysis, hydrogen bond occupancy, thermodynamics calculation and residue interaction network (RIN). Findings from this study revealed that Thr91 mutation leads to steric conflict between the UM-164 and side chain of methionine (Met91); the mutation distorts the UM-164 optimum orientation on the conformational space of mutant c-Src compared to wild-type; decreases the overall hydrogen bond formation between the residues in the mutant protein; decreases UM-164 binding energy in mutant by -13.416 kcal/mol compared to wild-type; decrease the residue correlation mutant; induces a change in the overall protein structure conformation from inactive to active conformation; distorts ligand atomic interaction network and distorts the residue interaction network. This report provides important insights that will assist in the further design of novel dual kinase inhibitors to minimise the chances of drug resistance in triple negative breast cancer.

Key words: Molecular dynamics, mutation, TNBC, Threonine 91, c-Src and drug resistance.

1 Introduction

Breast cancer is ranked the most frequently diagnosed and life-threatening cancer in women[1] and leading cause of cancer death among women.[1] In the United States of America however, breast cancer is the third leading cause of cancer death.[2] Breast cancers (BCs) are genetically heterogeneous[3] with respect to their gene composition, gene expression, and phenotypes thus, are currently classified into 5 subtypes.[3] The triple negative subtypes are more life threatening because of their potential to metastasize and tendency of local reoccurrence.[4] They are usually associated with the absence of oestrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor-2 (ErbB2/HER-2).[5] They are characterised by classical ductal histology, high grade, high mitotic and cell proliferations rates.[5] The triple negative cancer (TNBC) is highly associated with poor prognosis, poor disease-free survival (PDFS) and cancer-specific survival (CSS).[5] Positive lymph nodes are usually observed in the local reoccurrence,[5] this perhaps is the reason for high risk of reoccurrence in a patient with TNBC in the first 3 to 5 years after diagnosis.[5] The current studies suggests that chemotherapy remains the most effective treatment after surgery.[5] However, other studies have shown that neoadjuvant chemotherapy such as 5- fluorouracil (5-FU), anthracycline, cyclophosphamide, taxanes and platinum compounds used in the treatment of TNBC have low success rate[6] mainly due to lack of acceptable predictive biomarkers,[3] safety concerns and resistance to some of the compounds.[6] On molecular basis, TNBC has been classified as ER, PR and HER-2 and these molecular markers provide an understanding of the potential targets during the course of therapy.[7] The prognostic markers such as HER_1, $ALDH_1$, $LOXL_2$, Ki-67, SNCG and LDHB are important in providing prognostic information.[7]

In a related development, the influence of tissue inhibitor metalloproteinases-1 (TIMP-1) on TNBC[8] revealed that TIMP-1 is a biomarker indicative of poor prognosis in TNBC diseased individuals.[8] Hence TIMP-1 may be an indication of the need for therapeutic intervention specifically for TNBC.[8] Most of the therapies targeted at TNBC are not effective[3] because of a lack of ER, PR and HER-2. However, efforts have been made to overcome these challenges in the last 3 decades, some of which have resulted in a drastic improvement in patient survival[3]. In line with this, cytotoxic drugs remain the backbone of chemotherapy in TNBC[9] even as FDA approved targeted therapies are challenged by the relative heterogeneity of TNBC.

Recently, the clonal evolution study used the whole genome and exome single cell sequencing approach[10] to determine heterogeneity and evolution of breast cancer. Findings from this study

show that triple negative breast cancer has an increased mutation rate compared to ER^+ tumour cell which did not show any form of mutation.[10] These findings are very important in the process of diagnosis, treatment and evolution of chemoresistance in breast cancer.[10] A more recent study established a relationship between TNBCs and c-Src (Src), in an attempt to identify predictive markers response to chemotherapy.[3] In this study, a dual inhibitor (UM-164) was found to have a profound activity against Src and p38 kinases.[3] Src is the cellular homolog of the viral oncogene v-Src[11] and a prototypical member of the family of non-receptor tyrosine kinases that play important role in a variety of signalling pathways that involve proliferation, differentiation, survival, motility, and angiogenesis.[11] Overexpression of Src is known to play an important role in oncogenic proliferation, migration, and invasion of TNBC cell lines.[3] However, the emergence of a resistance mutation in Src is limiting the success of inhibitors used in new targeted cancer therapies[11,12] such as dasatinib, an inhibitor of both c-Src and abl protein.[11] The most common mutations occur at the gatekeeper position in the hinge region[12] in which small amino acid chain (Threonine) is exchanged for larger hydrophobic residue (Isoleucine or Methionine),[12] as shown in **Figure 1.**

Mutation in cancer is a global phenomenal and source of an obstacle to the success in drug design and the treatment of patients with any form of cancer. This is more so because continuous residual network changes in protein impacted by mutation have challenged the potency and efficacy of most effective experimental drug such as UM-164. Therefore, UM-164 could become completely ineffective as primary TNBC treatment in the future, should resistance to UM-164 emerge in the absence of other chemotherapeutic agents with more superior pharmacological effects.

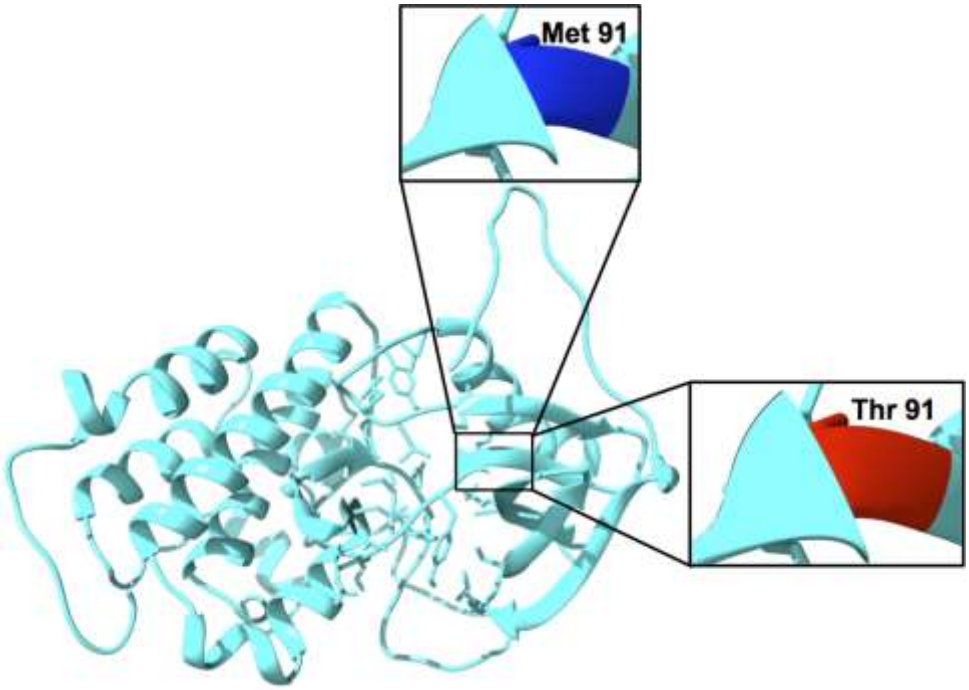


Figure 1. The 3D structure of Src protein showing the position of mutation (red) Thr91, studied in this work

Mutational resistance in Src have been previously studied[12], however, such studies were developed using X-ray crystal structure of Src with variable ligands.[12] Similarly, attempt was made to obtain molecular understanding on how mutation confers resistance to kinase inhibitors.[12, 13, 14] However, findings from these studies revealed molecular changes in the protein (changes from inactive conformation to active conformation resulting from mutation) that require prolonged timescale molecular dynamic (MD) simulations of Src structure.[13, 14] Therefore thorough understanding of conformational features of mutant Src in complex with UM-164 is crucial to the design of potent, effective and relatively less toxic dual kinase inhibitors (Src/p38). Although, detailed conformational features of the Src-UM-164 complex have been conducted and reported in our previous study, however, conformational features of Src mutant and its potential resistance to UM-164 have not been examined. The lead compound against Src inhibition (UM-164) is represented in **Figure 2**.

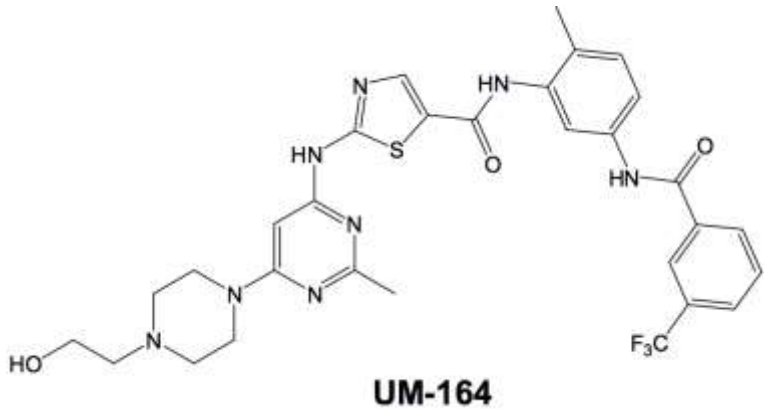


Figure 2. The 2D structure of UM-164 the lead compound in Src in inhibition.

The molecular dynamics calculation and analysis have not been conducted on the Src-UM-164 mutant. Herein, we aim to provide deeper insight into the effect of the mutation on the dynamics of Src and adding new dimensions to experimental work that has been previously conducted. To achieve this, we perform MD simulations of Src-UM-164 (wild type), Src-UM-164-mutant (mutant) and apo. Principal component analysis (PCA), dynamic cross-correlation (DCCA), residue interaction network (RIN) and binding free energy calculations were also conducted to understand the effect of mutation on the binding of UM-164.[15] Findings from this study can provide important insight into the conformational dynamics of mutant complex that will assist in the design of novel Src inhibitors with good pharmacological properties on mutant protein.

Various post dynamic techniques have been used to provide a molecular understanding of molecular dynamics. The principal component analysis (PCA), also known as essential dynamics analysis,[16] is one of the most popular post-dynamics techniques [16] that is widely used to understand the changes in the biological system.[16] PCA eliminates spectrum of translational and rotational motion in molecular dynamic (MD) trajectory and correlated motions in atomic simulations of proteins.[17] It defines the atomic displacement in a collective manner,[16] and can detect major conformational changes between the structures[15] and has been used in many studies to determine the difference in motion of protein complex of two different compounds.[16]

We believe that computational and molecular modelling tools as adopted in this study would provide the necessary information on the subject matter. These tools are known to enhanced the process of drug discovery,[18]therefore, provide a platform for the discovery of novel therapeutics.[18]

1 Computational methods

1.1 System preparation

The X-ray crystal structure of Src in complex with UM-164, PDB code 4YBJ[19] was obtained from Protein Data Bank (PDB). This is an inactive conformation of c-Src structure deposited in PDB as a homodimer with two chains (A and B), in this study, only chain A was used for simulation to reduce the computation cost. *In silico* mutation was carried out to mutate threonine (Thr) at position 91 to methionine (Met) as shown in **Figure 1**. Similarly, Chimera[20] and Avogadro software package[21] were used to modify and visualise ligand and receptor respectively.

1.2 Molecular dynamic simulations

Simulations of Src in complex UM-164 (wild), Src-mutant (mutant) in complex with UM-164 and Src (apo) were performed using graphic processor unit (GPU) version of Particle Mesh Ewald Molecular Dynamics (PMEMD) package with Sander module of Amber14.[22, 23] The AMBER force field ff12SB[24] was applied to describe the protein.[25] The ligands parameters were set using Gasteiger charges in Avogadro,[21] and Antechamber module with the aid of GAFF (generalised Amber force field).[26] The LEAP module implemented in Amber14[27] was used to add hydrogen atoms to the protein and to add counter ions for the system neutralization.[27] Each system is enclosed in the TIP3P water box[25] with the protein atoms located 10 Å between the protein surface and the box boundary within the period of simulations. The cubic periodic boundary conditions were implemented in all the systems, long-range electrostatics interaction was treated with particle-mesh Ewald method[25] implemented in Amber14 with a nonbonding cut-off distance of 12 Å. Two minimization steps were employed, partial minimization and full minimization. The initial energy minimization step of the systems was carried out with a restraint potential of 500 kcal mol^{-1} $Å^{-2}$ apply to the solute, for 1000 steps. Unrestrained conjugated gradient minimization for 1000 steps was conducted for the entire system with aid of SANDER module of Amber 14 program. A canonical ensemble (NVT) MD simulations were performed for 50ps and the system was gradually heated from 0 to 300 K, with harmonic restraints of 5 kcal mol^{-1} $Å^{-2}$ for solute atoms with the aid of Langevin thermostat[28] with a 1ps random collision frequency. The systems were equilibrated at 300 K with a 2fs time step in NPT ensemble for 500 ps without any restraint and Berendsen barostat[28] was used to maintain the pressure at 1bar. The SHAKE[29] algorithm was used to constrain the bonds of hydrogen atoms in the system. The 2fs time scale and SPFP precision model were used for MD runs. In the absence of restraints, a production run of 150

ns MD simulations were conducted in an isothermal-isobaric (NPT) ensemble using a Berendsen barostat[28] at a pressure of 1 bar and a pressure-coupling constant of 2 ps. For every 1ps time interval, the coordinates were saved and the trajectories were analysed every 1ps. Post MD analyses performed include RMSD, RMSF, the radius of gyration, hydrogen-bond occupancy, dynamic cross-correlation and principal component analysis (PCA) using PTRAJ and CPPTRAJ[30] modules in amber 14, as well as residue interaction network (RIN). Visualisation of trajectories were done using the molecular modelling tool in chimera.[20] The results were analysed and plots were generated with aid of origin software[31] and Bio3D[32] software respectively.

1.3 Thermodynamic calculations

The binding free energy calculation is an important thermodynamic method that gives detailed information on the interaction between the ligand and protein.[33] It provides understanding of the mechanism of binding, including contributions from both enthalpy and entropy to the molecular recognition.[33] Molecular Mechanics/Generalized-Born Surface Area method (MM/GBSA)[33] are popular methods used to estimate free energy of the binding of small ligands to the biological macromolecule,[33] the calculation gives detailed information on the interaction between the ligand and protein.[33] The binding free energy of wild type and mutant were both calculated using MM/GBSA.[33] For 150 ns trajectory, 1000 snapshots were considered during the calculation of binding free energy. The binding free energy computed by this method can be represented by following equations:

$$\Delta G_{bind} = G_{complex} - G_{receptor} - G_{ligand} \quad (1)$$

$$\Delta G_{bind} = E_{gas} + G_{sol} - TS \quad (2)$$

$$E_{gas} = E_{int} + E_{vdW} + E_{ele} \quad (3)$$

$$G_{sol} = G_{GB} + G_{SA} \quad (4)$$

From the equation above, E_{gas} is the energy of the gas phase, E_{int} represents internal energy, E_{ele} represents coulomb while E_{vdW} are the van der Waals energies. E_{gas} is estimated directly from the ffSB[24] force field. G_{sol} which is the solvation free energy can be broken down to polar and non-polar forms of contribution. The contribution of polar solvation (G_{GB}) is assessed by resolving G_{GB} equation and non-polar solvation (G_{SA}) is determined from the solvent accessible surface area, this can be estimated from water probe radius of 1.4 Å with temperature (T) and

total solute entropy (S). The MM/GBSA binding free energy method in Amber 14 was used to calculate the contribution of each residue to the binding free energy between the inhibitor (UM-164) and receptors both in wild type and mutant. In addition, the interaction energy decomposition analysis per residue was also computed using the same method.

1.4 Principal component analysis

Principal component analysis (PCA) also known as essential dynamics of protein[16] analysis, is a systematic statistical technique applied to reduce the number of dimensions needed to describe the protein dynamics[16] through the decomposition process that screen observed motions from largest to smallest spatial scale.[16] The atomic displacement and conformational changes of protein can be defined[15] using PCA by extracting different modes of the conformation of the protein complex during dynamic simulations such as MD simulation. The direction of motion (eigenvectors) and the extent of motion (eigenvalues) for the biological system can also be determine using PCA. In this study, the trajectories of the complexes from 150 ns MD simulations were stripped of the solvent molecules and the ions using the CPPTRAJ module in Amber 14,[30] this was done prior to MD trajectory processing for PCA. Principal component analysis was performed on Cα atoms on 1000 snapshots at 100 ps time interval each. Using in-house script, the first two principal components (PC1 and PC2) were computed and a 2 X 2 covariance matrix were generated using Cartesian coordinates of Cα atoms. PC1 and PC2 correspond to first two eigenvectors of a covariant matrix. An origin software[31] was used construct PC plot.

1.5 Dynamic cross-correlation matrices (DCCM) analysis

The cross-correlation is a 3D matrix representation that graphically displays time correlated information among the residues of the proteins.[34] Residue-based time correlated data can be analysed using visual pattern recognition. [34, 35] To better understand the dynamics of apo, wild type and mutant, a DCCM was generated to determine cross-correlated displacements of backbone Cα atoms in the trajectories, using the following equation:

$$C_{ij} = < \Delta r_i * \Delta r_j > / (< \Delta r^2_i > < \Delta r^2_j >)^{1/2} \quad (5)$$

Where, i and j represent ith and jth residues and Δr_i and Δr_j corresponds to the displacement of ith and jth atom from the mean respectively. The coefficient of cross-correlation $C_{ij,}$ varies between the range -1 to +1, where the upper and lower limits correspond to strong correlation (+) and anti-correlated (-) motions within the period of simulations. The DCCM analysis was

carried out using CPPTRAJ module in Amber 14,[30] matrices were generated and analysed using origin software.[31]

2 Results and discussion

2.1 System stability MD simulations

In preparation for MD trajectory analysis, RMSD and potential energy fluctuations were monitored throughout the MD simulations. RMSD was calculated to assess the stability and convergence of the respective systems and the results are presented **Figure 3**. Systems stabilisation and convergence with the maximum fluctuation of 4.25 Å between 0-12500 ps in apo system were observed. However, after approximately 25000 ps, the RMSD trajectories converged and the fluctuation rested below 2.00 Å. In mutant system, the highest fluctuation observed was 2.75 Å, at 15000 ps, after which the fluctuation rested below 2.0 Å. Similarly, in wild system the peak RMSD of 3.50 Å was reached at about 5000 ps. The average RMSD of 1.46 Å and 1.52 Å was observed in apo and mutant systems respectively, while in wild, an average RMSD of 1.58 Å was also observed. These account for system stability since a standard parameter defining a stable system is an RMSD of 2 Angstroms and below.[36]

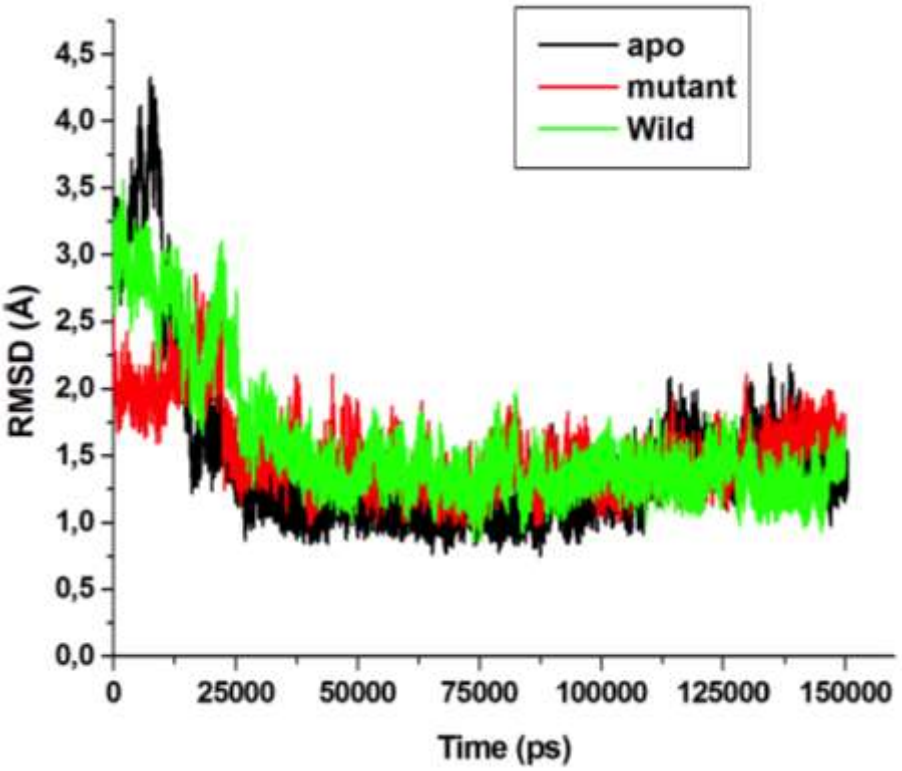
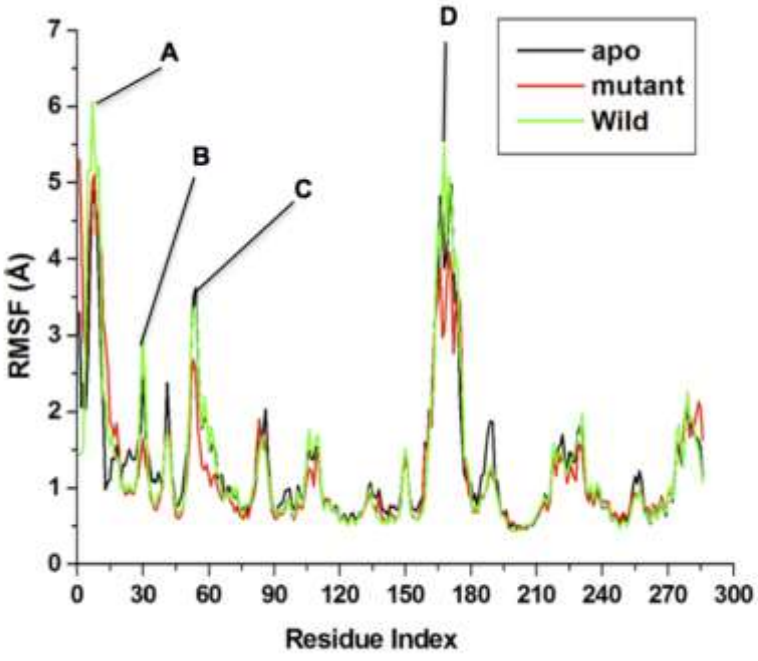
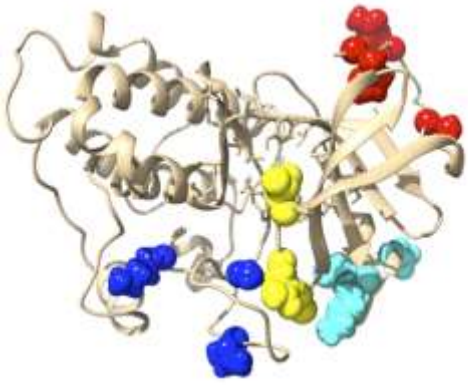


Figure 3. The RMSD plot of apo (black), mutant (red) and wild (green) respectively.

2.2 Root mean square fluctuation (RMSF)

Amino acids are building blocks of proteins[37] and play a vital role in conformational features of the proten.[38] Changes to the protein conformation occurs when there is a chemical reaction or mechanical events.[38] Therefore, direct interactions of the protein active site residues with a ligand may induce conformational changes in protein structure and alter its function. More specifically, the conformational changes that occur as a result of ligand-induced motion during ligand binding.[39] Understanding ligand-induced conformational changes in the protein structure are critical to structure-based rational drug design.[40] RMSF is a measure of average atomic mobility of backbone atoms (N, Cα and C) during MD simulation.[41] To understand and explore the structural dynamics that take place upon the ligand binding, RMSF of the subject systems was calculated from MD trajectories, represented in **Figure 4**. The core of the protein appears to be more rigid compared to the loops (solvent exposed) as shown by RMSF plot.





Colour code	RMSF Fluctuation	Residues
Red	A	His2, Gln6 and Ala12
Yellow	B	Leu26, Leu28 and Cys30
Cyan	C	Lys51, Met55 and Glu58
Blue	D	Ala161, Gln173 and Cys180

Figure 4. RMSF plots of apo (apo), mutant (red) and wild-type (green) systems respectively.

The highest fluctuations was observed in the loop region "**A"** involving residues His2, Gln6 and Ala8, with RMSF values of 4.8 Å, 5.0 Å and 6.0 Å in apo, mutant and wild-type respectively. The loop region "**B**" involving residues Leu26, Leu28 and Cys30 exhibit fluctuations with RMSF values of 2.5 Å, 1.5 Å and 3.0 Å in apo, mutant and wild-type systems respectively. In loop region "**C**" involving residues Lys51, Met55 and Glu58 fluctuations was observed at RMSF values of 3.5 Å, 2.8 Å and 3.48 Å in apo, mutant and wild-type respectively. In a related development, maximum fluctuations were observed in the loop region "**D**" involving residues with Ala161, Gln173 and Cys180 with RMSF values of 5.0 Å, 5.5 Å and 4.0 Å in apo, mutant and wild-type system respectively. These loops are called activation loops[42] and are important in the conversion of inactive conformation of Src to the active conformation[42] after phosphorylation. Therefore, fluctuations exhibited by these loops in apo and wild-type reflects on the status of Src, meaning that inactive form of the protein favours the binding of UM-164. It is known that UM-164 binds to an inactive conformation of Src[3] and a switch from inactive conformation to active conformation rigidifies the protein thereby reducing the tendency of loop fluctuation.[42] Fluctuations in the loops of apo and wild-type systems support the experimental evidence which states that UM-164 binds to DFG-out inactive conformation of Src.[3] However, a decrease in the loops fluctuations observed in a mutant structure may be a result of adoption of an active conformation due to a synergy between the Thr91 pro-mutation which destabilises the inactive conformation required for UM-164 binding.

2.3 Radius of Gyration

The radius of gyration (RoG) is a moment of inertia of Cα atoms from their centre of mass.[43] It has been applied to gain insight into molecular stability in biological systems during molecular dynamic simulations. The RoG of apo, mutant and wild-type systems was evaluated and presented in **Figure 5**. There is no remarkable difference in the average RoG of the three systems. The wild-type system exhibits a lower average RoG of 19.47 Å, whereas apo and mutant systems had an average RoG of 19.56 Å and 19.71 Å respectively.

However, the mutant system exhibits a relatively higher average RoG as evident in **Figure 5**. This increase may be a reflection of highly unstable nature of a mutant complex compared to wild-type. This agreed with the assumption that a mutation decreases the interaction between the amino acids resulting in an unstable moment of inertia.[43]

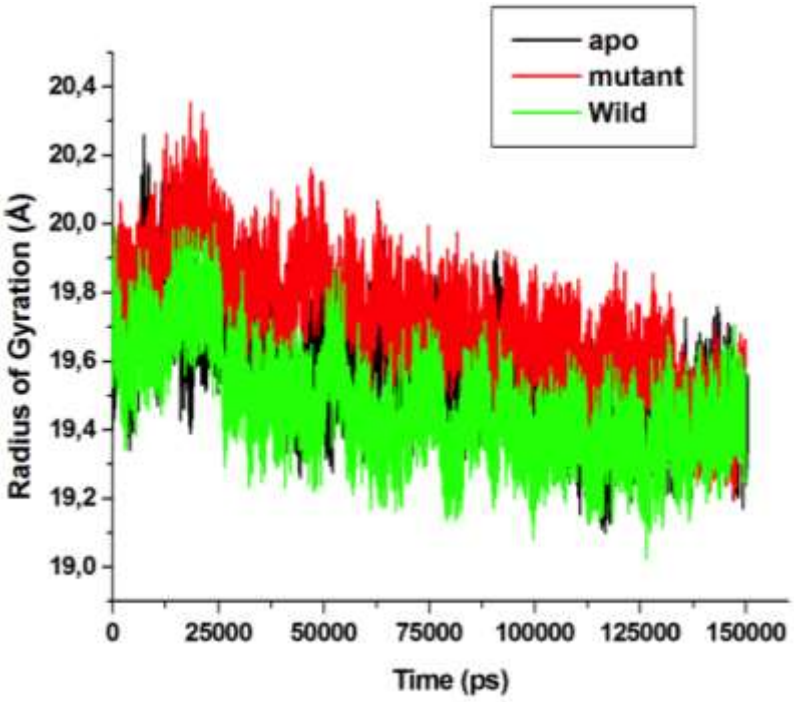


Figure 5. Radius of gyration plot of Cα of the apo (black), mutant (red) and wild-type (green) protein structures respectively.

2.4 Principal component analysis (PCA)

Protein conformation has been recognised as one of the important features in the determination of biological function,[44] and PCA is one of the principal tools used in determining the flexibility of each atom during a simulation.[45] **Figure 6** represents PCA plots of the three systems in this study. Herein, a clustering method of principal components (PC) was adopted because of its ability to describe different conformational states sampled during a simulation by grouping molecular structures into a subset based on their conformational similarities.[46] The percentage variability or total mean square displacement of atom's positional fluctuation captured in each dimension is characterised by their corresponding eigenvalue.[47] This method of PCA was used to assess major conformational changes in apo, mutant and wild-type systems during 150 ns MD simulations. In order to gain insight into motions associated with conformational behaviour of the subject systems, the systems were projected along the first two principal components (PC1 vs PC2) or eigenvectors direction. The PCA plot in **Figure 6** shows different and detailed waves of conformation in important subspace along the two principal components.

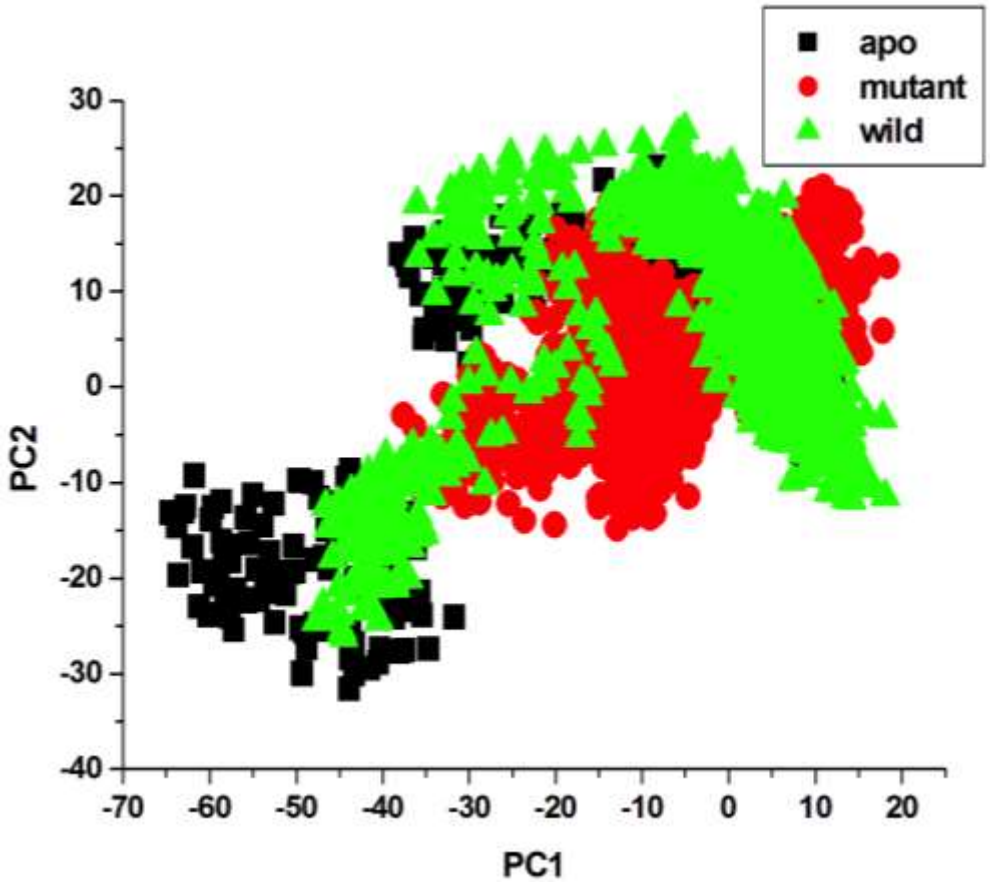


Figure 6. PCA projection of C-α atoms motion constructed by plotting the first two principal components (PC1 and PC2) in conformational space with apo (black), mutant (red) and wild-type (green) colours respectively.

In the PCA plots of the three systems, a distinct separation of motion was observed in the three respective systems. However, a more correlated motion is observed in the wild type system along the two principal components PC1 and PC2 compared to mutant and apo with relatively less correlated motion along PC1 and PC2 respectively. The apo system appears to be more flexible than mutant and wild-type, suggesting that the binding of UM-164 to the protein induces conformational dynamics reflected in the PCs as a wave of motion. Similarly, the binding of UM-164 to the mutant protein impact less dynamic effect on the protein compared to a wild-type. A decrease in fluctuation observed in a wild-type system is a mark of UM-164 impact on the protein. In the mutant system, however, less rigidity was observed compared to a wild-type system this could be due to mutation-induced conformational changes.

2.5 Dynamic cross-correlation matrices (DCCM) analysis

To further examine the conformational changes of receptor upon mutation, DCCM analysis was conducted on the positions of the Cα atoms throughout the simulations to determine the presence of correlated motions (**Figure 7**). The correlated motions (highly-positive) of specific

residues are represented as a yellow-red (colour) region, whereas, anti-correlated (highly-negative) movement of specific residues are represented as blue-black (colour) regions. The three systems herein exhibited overall correlated residual motions relative to anti-correlated motions. DCCM analysis shows that binding of UM-164 alters the structure conformation of c-Src as reflected by changes in the correlated motions and dynamics. In apo conformation (**Figure 7 A**) Met94 correlates with Ser95, while Ala156 slightly correlates with Val155 and Asp157. Anti-correlated residual motions in apo occur between residues 200 - 280 compared to wild-type and mutant with variable correlated and anti-correlated motions.

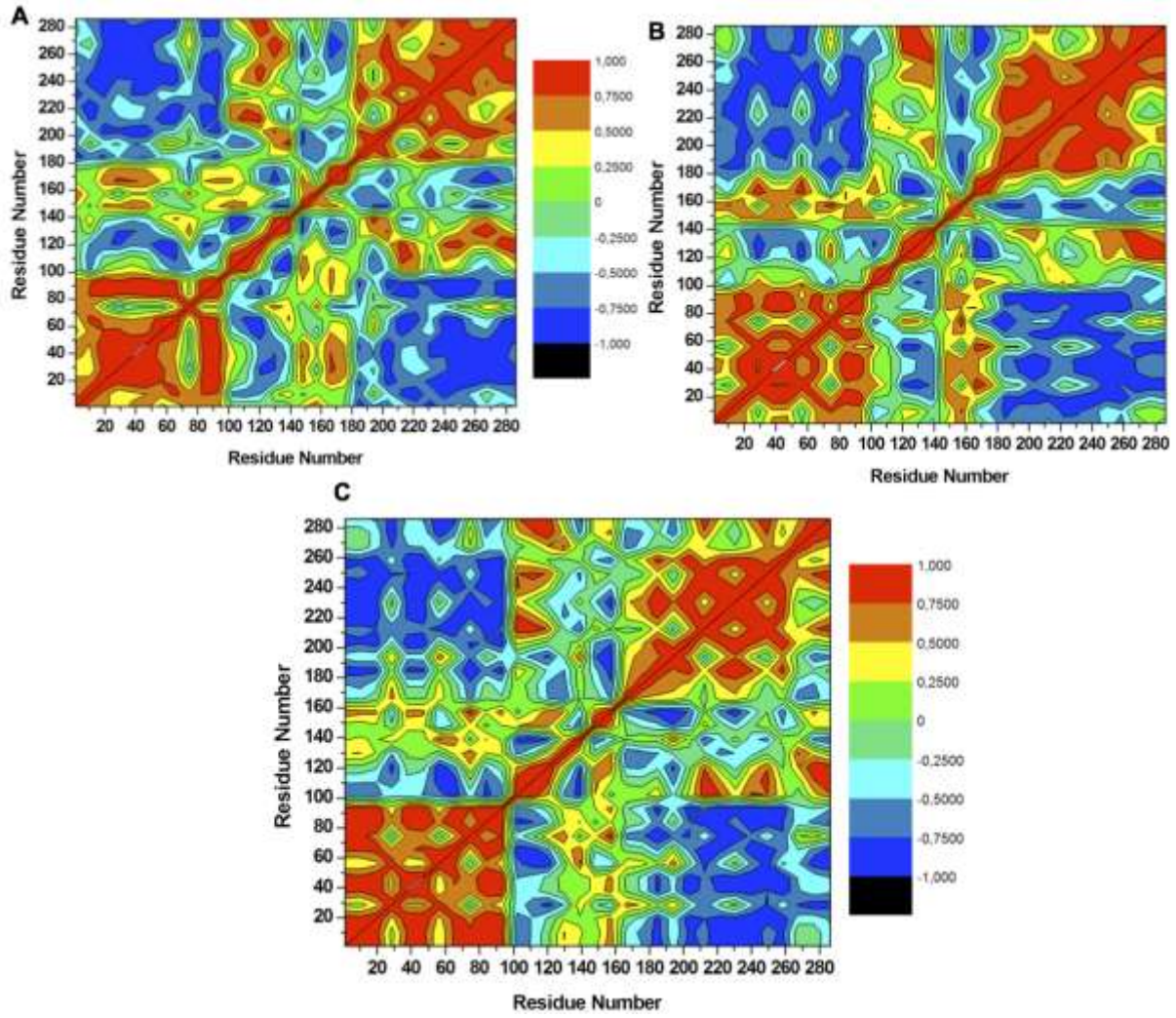


Figure 7. Cross-correlation matrices of the C-α atoms fluctuations in apo (**A**), mutant (**B**) and wild-type (**C**)

The wild-type exhibits two prominent correlated regions (**Figure 7 C**), namely; residue 1-95 strong correlated region, whereas 122-160 slight correlated region. These regions are among the most dynamic regions in the receptor and most hydrophobic active site residues lie within

these regions. Similarly, in the mutant system relative to the wild-system, correlated motions slightly decreases between residues 1-95 where mutation occurred whereas anti-correlated motion increases between residues 180-280. This phenomenon suggests the occurrence of mutation induced residue dynamics that may have resulted in conformational changes in the protein.

2.6 Hydrogen bond formation between amino acid residues

Hydrogens (H-bonds) are universal in nature, they play a central role in biological systems and maintenance of the protein structural integrity,[48] protein-ligand interaction and catalysis.[48] H-bonds are facilitators of protein-ligand binding,[48] therefore, the formation of hydrogen bond between amino acid residues is key to the monitoring of protein conformation. Sequel to this, we investigate hydrogen bond formation during the course of the simulations. **Figure 8** shows a hydrogen bond formation over time during simulations of the respective systems.

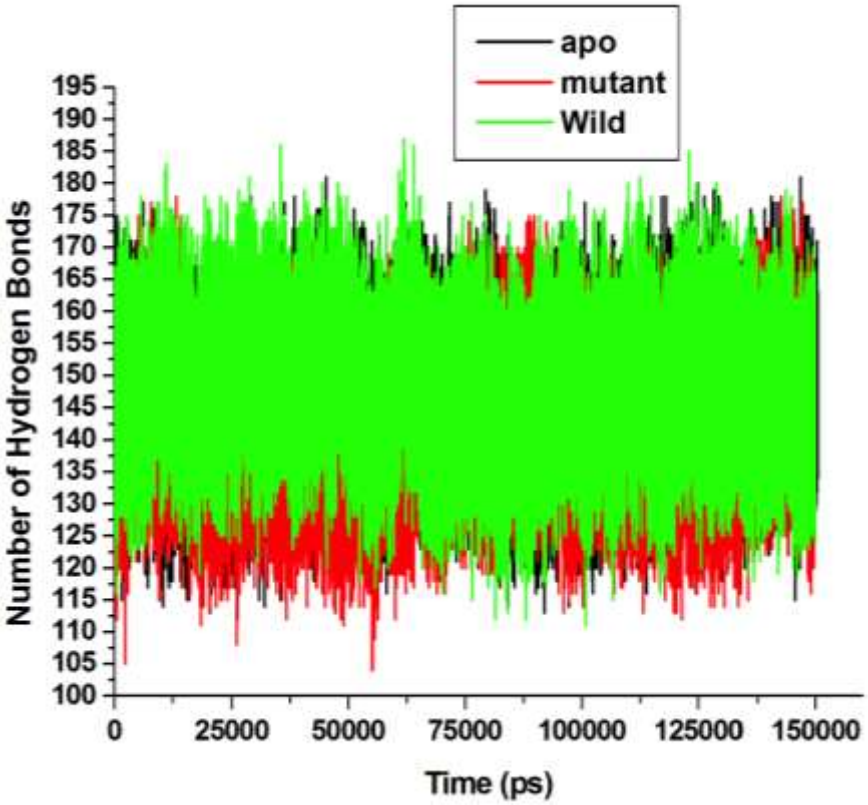


Figure 8. Number of hydrogen bond formation during simulation over time between apo, mutant and wild-type systems.

The mutant system exhibited a lower average H-bond (142.15) formation during a simulation compared to the wild-type and apo systems with average H-bond formations of 148.58 and

146.31 respectively. The decrease in H-bond formation observed in the mutant system could be a result of H-bond destabilisation caused by the mutation in the mutant system. Threonine is a small hydrophilic amino acid which is highly conserved in the core of the protein, it forms a H-bond with UM-164. Mutation of this residue to a large hydrophobic amino acid (methionine) induces local changes in the conformation of the protein leading to loss of H-bond between oxygen atom of thrionine91 and nitrogen (N21) of UM-164 as shown in **Figure 10**. The reduction in H-bond formation leads to structural imbalances and conformational dynamics which eventually affect drug binding.[48]

To further investigate the relative stability of wild-type, mutant and apo systems, H-bond distance and occupancy of the active site residues in the respective systems were monitored throughout the course of simulations and the results are presented in **Table 1**.

Table 1. Hydrogen Bond Occupancy of interactive active site residues of wild-type, mutant and apo systems.

H-bond acceptor	H-bond donor	Frame number	Occupancy (%)	Average distance (Å)	Average angle (degree)
Wild type					
Ala46@O	Thr91@H	129860	88	2.83	161
Met94@O	UM-164@H56	131304	85	2.84	159
Val34@O	Leu26@H	128546	85	2.81	155
Leu146@O	Lys154@H	94939	63	2.87	162
Glu63@O	Met67@H	79596	53	2.87	162
Lys48@O	Ile89@H	56321	37	2.90	161
Val155@O	Val76@H	40569	27	2.90	156
Leu26@O	Leu163@H	37529	25	2.87	161
Mutant					
Met94@O	UM-164@H56	131304	87	2.82	159
Val34@O	Leu26@H	124614	83	2.80	154
Ala46@O	Met91@H	132292	72	2.87	161
Leu146@O	Lys154@H	82337	54	2.87	162
Lys48@O	Ile89@H	77791	51	2.89	161
Ser95@O	UM-164@H45	64287	42	2.83	159
Glu63@O	Met67@H	40877	27	2.89	162
Val155@O	Val76@H	24961	16	2.91	155
Apo					
Val34@O	Leu26@H	106510	70	2.84	153
Thr91@O	Ala46@H	83764	55	2.88	161
Lys48@O	Ile89@H	77496	51	2.88	162
Glu63@O	Met67@H	72287	48	2.87	162
Val155@O	Val76@H	61045	40	2.88	158
Asp157@O	Asn144@H	54932	36	2.84	165
Leu26@O	Arg172@HH12	34006	22	2.82	154
Met94@O	Gly97@H	8882	5	2.91	151

Note: A = Angstrom; % = Percentage

In a wild-type system, Thr91 exhibit highest H-bond occupancy of 88% with an average H-bond distance of 2.83 Å. This remarkable high H-bond occupancy signifies the importance of Thr91 to the conformational state of the protein. Within the same system, maximum H-bond occupancy of 85% was observed between oxygen atom Met94 and H56 atom of UM-164. Similarly Val34 with an average H-bond distance of 2.84 Å exhibit H-bond occupancy of 85%. Lys48 and VaL155 exhibit the highest average H-bond distance of 2.90 Å each with relatively low H-bond occupancy of 37% and 27% respectively. The H-bond occupancy and distance account for the relative contribution of these residues to the stability of the wild system. However, in mutant system Met91 exhibit H-bond occupancy of 72% with an average H-bond

distance of 2.87 Å, this is contrary to the position of threonine in a wild-type system. A decrease in occupancy and increase H-bond distance in the mutant system may be attributed to the hydrophobic nature and steric effect of threonine which may greatly affect drug binding. The apo system has Met94 with maximum average H-bond distance of 2.91 Å and the least H-bond occupancy of 5%, while Val34 had a minimum average H-bond distance of 2.84 Å and a highest H-bond occupancy of 70%. Generally, the wild-type system exhibits highest H-bond occupancy compared to mutant and apo demonstrating relative stability of the system.

2.7 Binding free energy and energy decomposition analysis

Molecular Mechanics/Generalized-Born Surface Area (MM/GBSA) method[49] is a popular approach used to estimate the binding free energy of small ligands to biological macromolecules.[49] This method was used to estimate the total binding energy of UM-164 to wild-type and mutant proteins as presented in **Table 2.**

Table 2. MM/GBSA-based binding free energy profile of UM-164 bound to wild-type and UM-164 bound to mutant protein.

ΔG_{bind}	ΔE_{ele}	ΔE_{vdW}	ΔG_{gas}	ΔG_{sol}
Wild-type				
-82.293±4.320	-121.874±13.653	-83.311±3.332	-205.184±13.836	122.890±12.879
Mutant				
-68.877±4.258	-98.718±13.652	-75.741±3.543	-174.455±14.017	105.577±12.924

Notes: ΔE_{ele} = electrostatic energy; ΔE_{vdW} = van der Waals energy; G_{bind} = calculated total binding free energy; G_{sol} = solvation free energy.

The result of energy decomposition analysis revealed that the estimated binding free energy is higher in wild-type (-82.293 kcal/mol) compared to mutant (-68.877 kcal/mol). The difference in binding energy (-13.416 kcal/mol) between wild-type and mutant is quite significant, meaning that the force of interactions contributes higher energy in the binding of UM-164 to Src in wild-type compared to the mutant. Generally, hydrophobic packing contributes significantly to binding free energy in wild-type and mutant systems owing to multiple number of aromatic and hydrophobic rings (UM-164) within the conformational space, as well as a set of hydrophobic residues around the binding pocket. However, a decrease in energy contribution of interacting forces in a mutant system compared to a wild-type system signifies a remarkable decrease in the potency of UM-164 in the mutant system. Therefore, thr91mutation in Src is likely to cause resistance to UM-164.

2.7.1 Per-residue energy decomposition analysis

To assess the energy contribution of individual active site residues to the total binding free energy, and to provide a molecular understanding of the impact of protein dynamics on the degree of different binding forces, per-residue energy decomposition analysis was conducted and the energy values are represented in **Table 3.**

Table 3. Decomposition of the relative binding free energies on a per-residue basis for Wild-type and mutant systems.

Residues	ΔE_{vdw}	Δ_{ele}	ΔG_{polar}	$\Delta G_{non\text{-}polar}$	$\Delta G_{binding}$
Wild-type					
Met94	-1.244±0.612	-6.662±0.838	4.548±0.616	-0.070±0.020	-3.718±0.575
Thr91	-1.670±0.471	-2.174±0.702	1.106±0.303	-0.147±0.022	-2.885±0.684
Lys48	-2.438±0.450	8.176±1.417	-8.449±1.847	-0.159±0.018	-2.871±1.269
Phe158	-2.330±0.398	-0.294±0.398	0.277±0.293	-0.120±0.029	-2.467±0.488
Asp157	-2.425±0.509	-9.667±0.923	9.924±1.031	-0.207±0.023	-2.374±0.548
Leu146	-2.077±0.258	0.724±0.132	-0.825±0.115	-0.174±0.023	-2.353±0.279
Met67	-2.137±0.340	-0.388±0.416	0.400±0.260	-0.176±0.031	-2.301±0.402
Tyr93	-2.418±0.381	-1.580±0.950	1.872±0.760	-0.169±0.041	-2.296±0.556
Leu26	-2.169±0.414	-0.371±0.258	0.644±0.316	-0.242±0.067	-2.139±0.406
Gly97	-1.456±0.343	-0.042±0.527	-0.089±0.454	-0.137±0.026	-1.724±0.555
Mutant					
Met94	-0.890±0.174	0.595±0.166	-0.606±0.143	-0.005±0.007	-0.906±0.165
Met91	-2.097±0.417	0.423±0.369	-0.146±0.186	-0.198±0.032	-2.018±0.426
Lys48	-1.092±0.296	7.093±0.843	-6.902±0.993	-0.133±0.033	-1.034±0.507
Phe158	-1.804±0.344	-0.345±0.252	0.553±0.161	-0.226±0.049	-1.822±0.363
Asp157	-1.264±0.205	-7.132±0.781	7.641±0.861	-0.154±0.037	-0.909±0.264
Leu146	-1.740±0.271	0.221±0.077	-0.269±0.060	-0.162±0.027	-1.950±0.200
Met67	-1.637±0.346	0.168±0.516	0.095±0.202	-0.180±0.045	-1.553±0.493
Tyr93	-1.981±0.377	-0.966±0.772	1.646±0.659	-0.143±0.035	-1.444±0.466
Leu26	-2.297±0.351	0.351±0.119	-0.243±0.095	-0.317±0.036	-2.507±0.365
Gly97	-0.513±0.087	-0.630±0.064	0.609±0.051	-0.015±0.008	-0.549±0.092

Notes: ΔE_{ele} = electrostatic energy (kcal/mol); ΔE_{vdW} = van der Waals energy (kcal/mol); $\Delta G_{polar\ (kcal/mol)}$ = polar solvation energy (kcal/mol); $\Delta G_{nonpolar\ (kcal/mol)}$ = nonpolar solvation energy (kcal/mol); $\Delta G_{binding\ (kcal/mol)}$ = total binding free energy $_{(kcal/mol)}$.

The binding free energy was decomposed into the unit contributions of each active site residue of wild-type and mutant protein structure, represented graphically in **Figure 9**. In a wild-type structure, all the active site residues listed in Table 4 contribute relatively high binding free energy with an exception of Gly97 (-1.724 kcal/mol) which has a less significant contribution compared to other active site residues of the wild-type. However, Met94 in a wild-type system has the highest binding free energy contribution relative to other active site residues in mutant

and wild-type systems. Therefore, Met94 can be regarded as a very important residue in the binding of UM-164 to Src. The presence of multiple benzene rings in UM-164 contributes to the observed hydrophobic interaction with active site residue in both wild-type and mutant systems as represented in **Figure 10**. In the mutant system, energy contributions from the active site residues were drastically reduced except for Leu26 with a total binding of energy of -2.507 kcal/mol compared to -2.139 kcal/mol in a wild-type system. Threonine (Thr91) in the wild-type system exhibits high electrostatic energy of -2.174 kcal/mol and total binding energy of -2.885 kcal/mol relative to Methionine (Met91) in a mutant system with a relatively low electrostatic energy contribution of 0.423 kca/mol and a total binding energy of -2.018 kcal/mol. However, the van der Waals energy contribution from Met91 (-2.097 kcal/mol) in the mutant is relatively more than the contribution from Thr91 (-1.670 kcal/mol) in a wild-type system. This may be due to a shift from hydrogen bond interaction of Thr91 to the hydrophobic interaction of Met91 in wild-type and mutant systems respectively.

On a close assessment of per-residue energy contribution, the observed significant decrease in total binding free energy from wild-type to mutant is a signal to eminent resistance to UM-164 in an event that Src undergoes mutation.

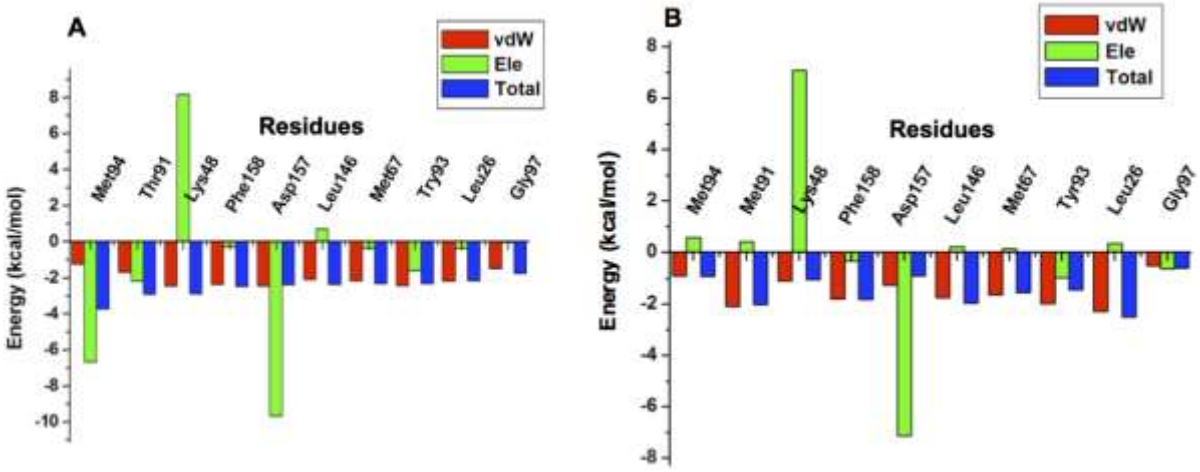


Figure 9. The per-residue energy decomposition analysis graph of wild-type (A) and mutant (B) Src.

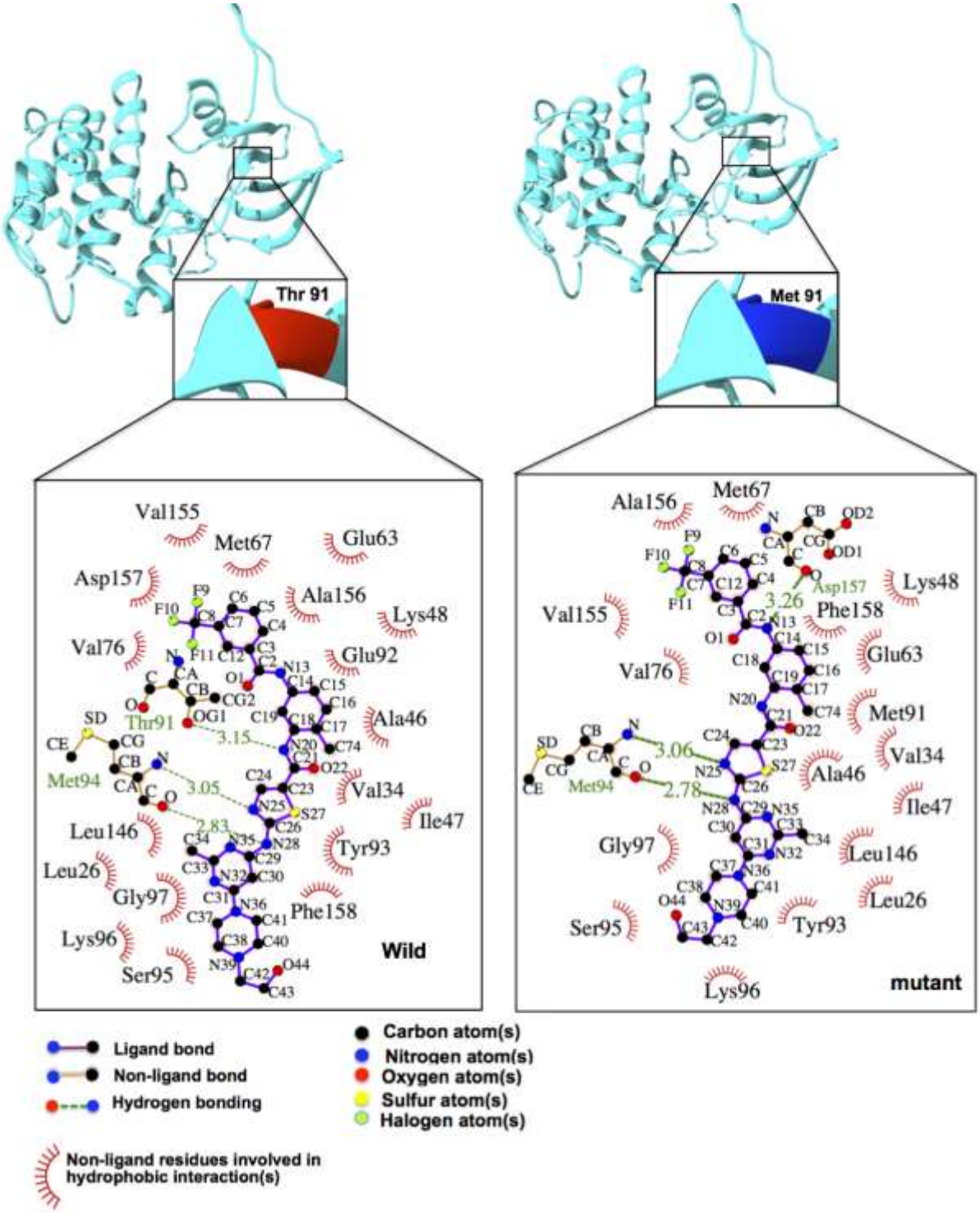


Figure 10. The interactive residues of wild-type (A) and mutant (B) Src complexed with UM-164.

2.8 Residue interaction networks (RINs)

Amino acids are an essential components of the biological system with a well-defined network of interaction. The molecular interactions of these amino acids vary among the body tissues

and are coded for by the genes, therefore an alteration in the gene sequence affect the network of the interaction of amino acids. This is particularly true when a mutation occurs, as a result the protein sequence is re-written and there is an alteration in the network of amino acids. Mutation can affect protein folding and stability,[50] protein function and protein-protein interaction.[50] Therefore an investigation of residue interaction network is imperative to the understanding of the impact of the mutation on the protein. Herein, residue interaction network analysis of protein backbone was used to explore differences in the network of proteins including the wild-type and mutant. This way, we investigate the relationship between key residues of the wild-type and mutant Src structures by generating residue interaction networks using fully-minimised structures from 150 ns MD simulations. RINs plot is presented in **Figure 11.**

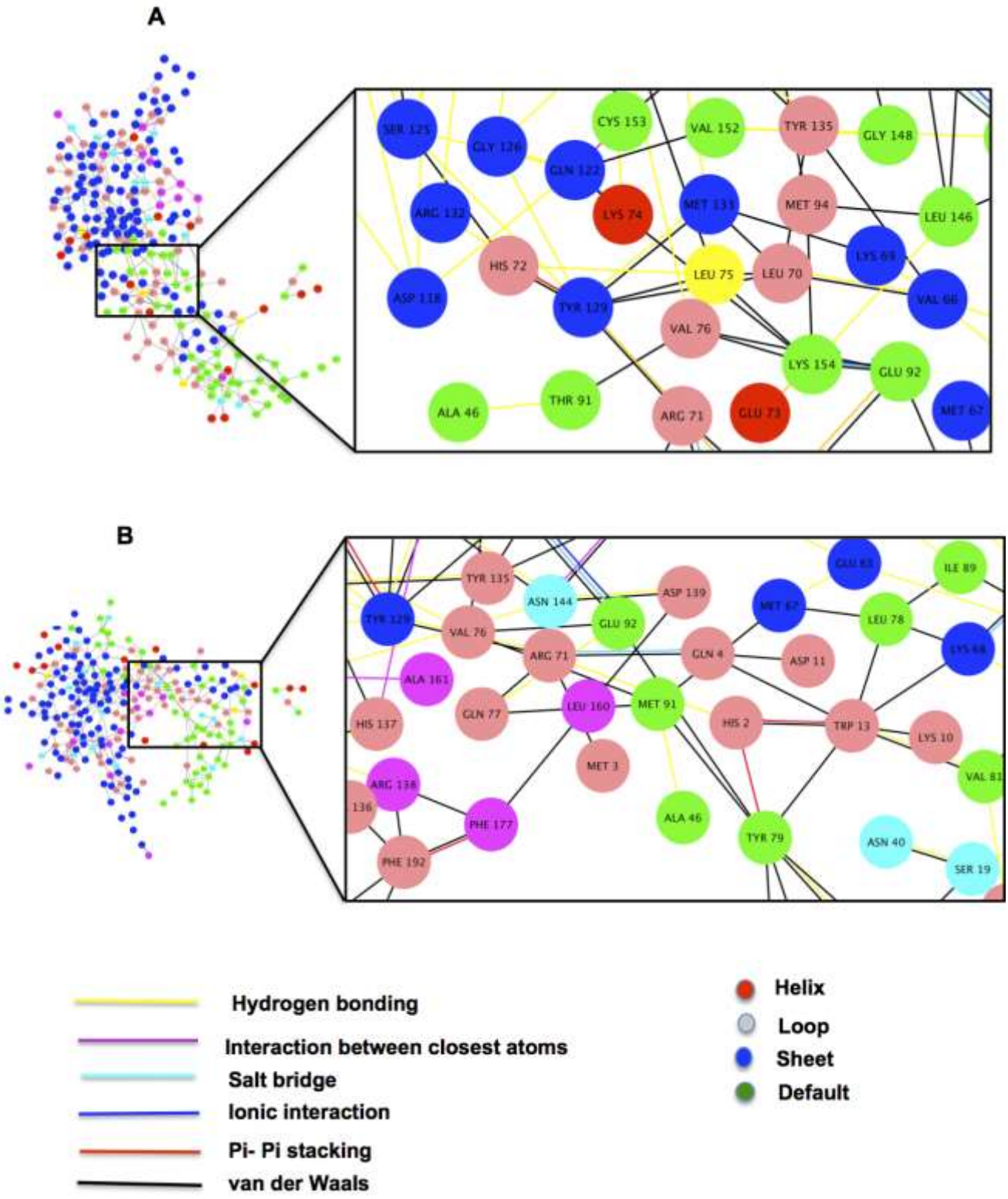


Figure 11. Comparative residue interaction network of wild-type (A) and mutant (B) protein Src structures complexed with UM-164 highlighting the changes in residue network interaction at point 91.

On a close examination of RIN, it is evident that the mutation distorted the overall residue network in a mutant compared to a wild-type structure. Residues interacting with Thr91 in a wild-type system (**Figure 11 A**) within the network space are different from the residues interacting with Met91 in a mutant system (**Figure 11 B**). In a wild-type, Thr91 interacts with Ala46 and Val76, while in a mutant, Met91 interacts with Ala46, His2, Leu160 and Gln4. In a related development, it was observed that Thr91 is involved mainly in H-bond interaction, while Met91 is involved in multiple van der Waal interactions. These changes in the network of interaction affect the drug binding landscape and consequently decrease the potency of UM-164 in mutant Src protein.

1 Conclusion

Currently, molecular understanding of Src resistance to the experimental drug UM-164 is lacking in the literature. Herein, we used different computational approaches aimed at providing multidimensional insight into Src resistance to UM-164. Molecular dynamic simulation, principal component analysis (PCA), dynamic cross-correlation matrices (DCCM) analysis, hydrogen occupancy, thermodynamics calculation and residue interaction network (RIN) led us to the findings that provide adequate information on the impact of the mutation on drug resistance.

These findings revealed that thr91 mutation decrease the capacity of loops to fluctuate; distort the ligand optimum orientation on the conformational space of Src; decrease the hydrogen overall hydrogen bond formation in the mutant protein; decrease the drug binding energy of UM-164 by -13.416 kcal/mol; decrease the residue correlation in mutant system; change the overall protein structure conformation from inactive to active conformation; distort ligand atomic interaction network and residue interaction network. These findings agreed with reported effect of the Src mutation on the binding of some kinase inhibitors such Imatinib and provide atomistic insight into the impact of the mutation on drug binding. These findings can provide important insights that will assist in the further design of novel dual kinase inhibitors to eliminate the chances of drug resistance in triple negative breast cancer.

Acknowledgements

The authors acknowledge the School of Health Science, University of KwaZulu-Natal, Westville campus for financial assistance. The Centre for High Performance Computing (CHPC, www.chpc.ac.za) Cape Town, South Africa, for computational resources

Disclosure

Authors declare no financial and intellectual conflict of interests.

References

(1) Sharma, G. N., Dave, R., Sanadya, J., Sharma, P., and Sharma, K. K. (2010) Various types and management of breast cancer: an overview. *J. Adv. Pharm. Technol. Res. 1*, 109–26.

(2) Siegel, R. L., Miller, K. D., and Jemal, A. (2016) Cancer statistics *66*, 7–30.

(3) Gilani, R. A., Phadke, S., Bao, L. W., Lachacz, E. J., Dziubinski, M. L., Brandvold, K. R., Steffey, M. E., Kwarcinski, F. E., Graveel, C. R., Kidwell, M., Merajver, S. D., and Soellner, M. B. (2016) UM-164: a potent c-Src/p38 kinase inhibitor with *invivo* activity against triple negative breast cancer. *Clin Cancer Res. 5087–96*.

(4) Anders, C. K., and Carey, L. a. (2010) Biology, Metastatic Patterns and Treatment of Patients with Triple-Negtive Breast Cancer. *Breast 9*, S73–S81.

(5) Jiao, Q., Wu, A., Shao, G., Peng, H., Wang, M., Ji, S., Liu, P., and Zhang, J. (2014) The latest progress in research on triple negative breast cancer (TNBC): Risk factors, possible therapeutic targets and prognostic markers. *J. Thorac. Dis. 6*, 1329–1335.

(6) Gluz, O., Liedtke, C., Gottschalk, N., Pusztai, L., Nitz, U., and Harbeck, N. (2009) Triple-negative breast cancer--current status and future directions. *Ann. Oncol. 20*, 1913–1927.

(7) Jafarzadeh, N., Ashraf, H., Khoshroo, F., Sepehri Shamloo, A., Bidouei, F., and Ghaffarzadehgan, K. (2015) Triple Negative Breast Cancer: Molecular Classification, Prognostic Markers and Targeted Therapies. *Razavi Int. J. Med. 3*.

(8) Cheng, G., Fan, X., Hao, M., Wang, J., Zhou, X., and Sun, X. (2016) Higher levels of TIMP-1 expression are associated with a poor prognosis in triple-negative breast cancer. *Mol. Cancer 15*, 30.

(9) Bayraktar, S., and Gl??ck, S. (2013) Molecularly targeted therapies for metastatic triple-negative breast cancer. *Breast Cancer Res. Treat. 138*, 21–35.

(10) Wang, Y., Waters, J., Leung, M. L., Unruh, A., Roh, W., Shi, X., Chen, K., Scheet, P., Vattathil, S., Liang, H., Multani, A., Zhang, H., Zhao, R., Michor, F., Meric-Bernstam, F., and Navin, N. E. (2014) Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature 512*, 1–15.

(11) Finn, R. S., Dering, J., Ginther, C., Wilson, C. a, Glaspy, P., Tchekmedyian, N., and Slamon, D. J. (2007) Dasatinib, an orally active small molecule inhibitor of both the src and abl kinases, selectively inhibits growth of basal-type/'triple-negative' breast cancer cell lines growing in vitro. *Breast Cancer Res. Treat. 105*, 319–326.

(12) Getlik, M., Grütter, C., Simard, J. R., Klüter, S., Rabiller, M., Rode, H. B., Robubi, A., and Rauh, D. (2009) Hybrid compound design to overcome the gatekeeper T338M mutation in cSrc. *J. Med. Chem.*

(13) Young, M. A., Shah, N. P., Chao, L. H., Seeliger, M., Milanov, Z. V., Biggs, W. H., Treiber, D. K., Patel, H. K., Zarrinkar, P. P., Lockhart, D. J., Sawyers, C. L., and Kuriyan, J. (2006) Structure of the kinase domain of an imatinib-resistant Abl mutant in complex with the

aurora kinase inhibitor VX-680. *Cancer Res. 66*, 1007–1014.

(14) Liu, Y., Bishop, A., Witucki, L., Kraybill, B., Shimizu, E., Tsien, J., Ubersax, J., Blethrow, J., Morgan, D. O., and Shokat, K. M. (1999) Structural basis for selective inhibition of Src family kinases by PP1. *Chem. Biol. 6*, 671–8.

(15) Wolf, A., and Kirschner, K. N. (2013) Principal component and clustering analysis on molecular dynamics data of the ribosomal L11. 23S subdomain. *J. Mol. Model. 19*, 539–549.

(16) Charles C. David and Donald J. Jacobs (2014) Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins. *Methods Mol Biol.* 1084: 193–226.

(17) Maisuradze, G., Liwo, a, and Scheraga, H. (2009) Principal component analysis for protein folding dynamics. *J. Mol. Biol. 385*, 312–329.

(18) M Kumalo and M .E Soliman (2015) Per-Residue Energy Footprints-Based Pharmacophore Modeling as an Enhanced In Silico Approach in Drug Discovery : A Case Study on the Identification of Novel b -Secretase1 (BACE1) Inhibitors as Anti- Alzheimer Agents. *Cellular and Molecular Bioengineering.*1-9

(19) Kwarcinski, F. E., Brandvold, K. R., Phadke, S., Beleh, O. M., Johnson, T. K., Meagher, J. L., Seeliger, M. A., Stuckey, J. A., and Soellner, M. B. (2016) Conformation-Selective Analogues of Dasatinib Reveal Insight into Kinase Inhibitor Binding and Selectivity. *ACS Chem. Biol. 11*, 1296–1304.

(20) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem. 25*, 1605–1612.

(21) Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeerschd, T., Zurek, E., and Hutchison, G. R. (2012) Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform. 4*, 1–17.

(22) Walker, R., Supercomputer, S. D., and Roitberg, A. AMBER 14 and GPUs : Creating the World ' s Fastest MD Package.

(23) Goetz, A. W., Williamson, M. J., Xu, D., Poole, D., Grand, S. L., and Walker, R. C. (2012) Routine microsecond molecular dynamics simulations with amber - part i: Generalized born. *J. Chem. Theory Comput. 8*, 1542–1555.

(24) Galindo-Murillo, R., Robertson, J. C., Zgarbova Jir, M., Otyepka, M., Jurec, P., and Cheatham, T. E. (2016) Assessing the Current State of Amber Force Field Modifications for DNA. *J. Chem. Theory Comput 12*, 4114–4127.

(25) Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., and Shaw, D. E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct. Funct. Bioinforma. 78*, 1950–1958.

(26) Wang, J. M., Wolf, R. M., Caldwell, J. W., Kollman, P. a, and Case, D. a. (2004) Development and testing of a general amber force field. *J. Comput. Chem. 25*, 1157–1174.

(27) Case, D. A., Darden, T., Iii, T. E. C., Simmerling, C., Brook, S., Roitberg, A., Wang, J., Southwestern, U. T., Duke, R. E., Hill, U., Luo, R., Irvine, U. C., Roe, D. R., Walker, R. C., Legrand, S., Swails, J., Cerutti, D., Kaus, J., Betz, R., Wolf, R. M., Merz, K. M., State, M., Seabra, G., Janowski, P., Paesani, F., Liu, J., Wu, X., Steinbrecher, T., Gohlke, H., Homeyer, N., Cai, Q., Smith, W., Mathews, D., Salomon-ferrer, R., Sagui, C., State, N. C., Babin, V.,

Luchko, T., Gusarov, S., Kovalenko, A., Berryman, J., and Kollman, P. A. (2014) Amber 14. *Univ. California, San Fr.*

(28) Finnerty, J. (2011) Molecular dynamics meets the physical world : Thermostats and barostats.

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

(30) Roe, D. R., and Cheatham III, T. E. (2013) PTRAJ and CPPTRAJ: software for processing and analysis of molecular synamics trajectory data. *J Chem Theory Com 9*, 3084–3095.

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

(32) Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A., and Caves, L. S. D. (2006) Bio3d: An R package for the comparative analysis of protein structures. *Bioinformatics 22*, 2695–2696.

(33) Genheden, S., and Ryde, U. (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov. 10*, 449–61.

(34) Arnold, G. E., and Ornstein, R. L. (1997) Molecular dynamics study of time-correlated protein domain motions and molecular flexibility: cytochrome P450BM-3. *Biophys. J. 73*, 1147–1159.

(35) Swaminathan, S. (1991) Investigation of domain structure in proteins via molecular dynamics simulation: application to HIV-1 protease dimer. *J. Am. ...* 2717–2721.

(36) Carugo, O., and Pongor, S. (2001) A normalized root-mean-square distance for comparing protein three-dimensional structures. *Protein Sci. 10*, 1470–3.

(37) Spassov, V. Z., Yan, L., and Flook, P. K. (2007) The dominant role of side-chain backbone interactions in structural realization of amino acid code. ChiRotor: a side-chain prediction algorithm based on side-chain backbone interactions. *Protein Sci. 16*, 494–506.

(38) Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002) Protein Function.

(39) Henchman, R. H., Wang, H.-L., Sine, S. M., Taylor, P., and McCammon, J. A. (2005) Ligand-induced conformational change in the alpha7 nicotinic receptor ligand binding domain. *Biophys. J. 88*, 2564–76.

(40) Ahmad, E., Rabbani, G., Zaidi, N., Khan, M. A., Qadeer, A., Ishtikhar, M., Singh, S., and Khan, R. H. (2013) Revisiting ligand-induced conformational changes in proteins: essence, advancements, implications and future challenges. *J. Biomol. Struct. Dyn. 31*, 630–648.

(41) Vendome, J., Posy, S., Jin, X., Bahna, F., Ahlsen, G., Shapiro, L., and Honig, B. (2011) Molecular design principles underlying β-strand swapping in the adhesive dimerization of cadherins. *Nat. Struct. Mol. Biol. 18*, 693–700.

(42) Pucheta-Martínez, E., Saladino, G., Morando, M. A., Martinez-Torrecuadrada, J., Lelli, M., Sutto, L., D'Amelio, N., and Gervasio, F. L. (2016) An Allosteric Cross-Talk Between the Activation Loop and the ATP Binding Site Regulates the Activation of Src Kinase. *Sci. Rep. 6*, 24235.

(43) Bhakat, S., Martin Bc, A. J. M., and Soliman, M. E. S. (2014) An integrated molecular

dynamics, principal component analysis and residue interaction network approach reveals the impact of M184V mutation on HIV reverse transcriptase resistance to lamivudine. *Mol. BioSyst. Mol. BioSyst 10*, 2215–2228.

(44) Assadi-Porter, F. M., Maillet, E. L., Radek, J. T., Quijada, J., Markley, J. L., and Max, M. (2010) Key Amino Acid Residues Involved in Multi-Point Binding Interactions between Brazzein, a Sweet Protein, and the T1R2–T1R3 Human Sweet Receptor. *J. Mol. Biol. 398*, 584–599.

(45) David, C. C., and Jacobs, D. J. (2014) Principal component analysis: a method for determining the essential dynamics of proteins. *Methods Mol. Biol. 1084*, 193–226.

(46) Wolf, A., and Kirschner, K. N. (2013) Principal component and clustering analysis on molecular dynamics data of the ribosomal L11·23S subdomain. *J. Mol. Model. 19*, 539–549.

(47) Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A., and Caves, L. S. D. (2006) Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics 22*, 2695–6.

(48) Chen, D., Oezguen, N., Urvil, P., Ferguson, C., Dann, S. M., and Savidge, T. C. (2016) Regulation of protein-ligand binding affinity by hydrogen bond pairing. *Sci. Adv. 2*, e1501240.

(49) Genheden, S., and Ryde, U. (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov. 10*, 449–61.

(50) Brender, J. R., and Zhang, Y. (2015) Predicting the Effect of Mutations on Protein-Protein Binding Interactions through Structure-Based Interface Profiles. *PLoS Comput. Biol. 11*, e1004494.

CHAPTER 7

Metal Complexes in Cancer Therapy – an update from Drug Design Perspective

Umar Ndagi[a] Ndumiso Mhlongo[a] and Mahmoud E. Soliman[a]*

[a] Molecular Modelling and Drug Design Research Group, School of Health Sciences, University of KwaZulu-Natal, Westville, Durban 4000, South Africa

*Corresponding author: Mahmoud E. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27 0312607413, Fax: +27 031260 779

Webpage: http://soliman.ukzn.ac.za/

Abstract

In the past, metal-based compounds were widely used in the treatment of disease conditions, but the lack of clear distinction between the therapeutic and toxic doses was the major challenge. With the discovery of cisplatin by Barnett Rosenberg in 1960, a milestone in the history of metal-based compounds used in the treatment of cancers was witnessed. This forms the foundation for the modern era of the metal-based anticancer drugs. Platinum drugs, such as cisplatin, carboplatin and oxaliplatin are the mainstay of the metal-based compounds in the treatment of cancer, but the delay in the therapeutic accomplishment of other metal-based compounds hampered the progress of research in this field. Recently, however, there has been an upsurge of activities relying on the structural information, aimed at improving and developing other forms of metal-based compounds, and non-classical platinum complexes whose mechanism of action is distinct from known drugs such as cisplatin. In line with this, many more metal-based compounds have been synthesized by redesigning the existing chemical structure through ligand substitution or building the entire new compound with enhanced safety and cytotoxic profile. However, because of increased emphasis on the clinical relevance of metal-based complexes, few of these drugs are currently on clinical trial and many more are awaiting ethical approval to join the trial. In this study, we seek to give an overview of previous reviews on the cytotoxic effect of metal-based complexes while focusing more on newly designed metal-based complexes and their cytotoxic effect on the cancer cell lines, as

well as new approach to metal-based drug design and molecular target in cancer therapy. We are optimistic that the concept of selective targeting remains the hope of the future in developing therapeutics that would selectively target cancer cells and leave healthy cells unharmed.

Keywords: Cancer, DNA, platinum, metal complexes, apoptosis, selective target.

1 Introduction

Therapeutic potentials of metal-based compounds date back to ancient time.[1] During this period, the ancient Assyrians, Egyptians and Chinese knew about the importance of using metal-based compounds in the treatment of diseases[1], such as the use of cinnabar (mercury sulphide) in the treatment of ailments.[1] The advent of "theoretical science," by Greek philosophers (Empedocles and Aristotle) in 5th and 4th century BC[1], boosted the knowledge of metal-based compounds as a therapeutic agents. This was supported by the information handed down by Pliny and Aulus Cornelius Celsus (Roman physician) on the use of cinnabar in the treatment of trachoma and venereal diseases.[1] In 9th and 11th century, the contributions of ancient scientists such as Rhazes (Al-Razi) and Avicenna (Ibn Sina) was applauded[1]. This is a sequel to the discovery of toxicological effects of mercury in the animals and the use of mercury (quicksilver ointment) for skin diseases respectively.

Arsenic trioxide (ATO) was used as an antiseptic[2] and in the treatment of rheumatoid diseases, syphilis and psoriasis by traditional Chinese medical practitioners[2]. Certainly, arsenic trioxide was among the first compounds suggested for use in the treatment of leukaemia[3] during 18th and 19th century, until in the early 20th century when its use was replaced by radiation and cytotoxic chemotherapy.[3] Therapeutic use of gold and copper can be traced to the history of civilization,[4] where the Egyptians and Chinese were famous users in the treatment of certain disease conditions, such as syphilis.[4] The discovery of platinum compound (cisplatin) by Barnett Rosenberg in 1960s[5] was a milestone in the history of metal-based compounds used in the treatment of cancer.[6] This forms the foundation for the modern era of the metal-based anticancer drug.[5] Despite the wide use of the metal-based compounds, the lack of clear distinction between the therapeutic and toxic dose was a major challenge. This was so because practitioners of ancient time lack adequate knowledge of dose-related biological response.[7] The advent of molecular biology and combinatorial chemistry pave the way for the rational design of chemical compounds to target specific molecules.[7]

Generally, metals are essential components of cells chosen by nature.[8] They are frequently found in the enzyme catalytic domain[9] and are involved in multiple biological processes, from the exchange of electrons to catalysis and structural roles.[9] They are extensively used in cellular activities[9]. Such metals include gallium, zinc, cobalt, silver, vanadium, strontium, manganese, and copper, which are required in trace amounts to trigger catalytic processes.[10] To this end, a balance between cellular need and the amount available in the body is important for the normal physiological state. Comparatively, metals including nickel, cadmium, chromium and arsenic can induce carcinogenesis, hence less beneficial to the body[10]. These limitations have triggered a search for platinum-based compounds that shows lower toxicity, higher selectivity and a broader spectrum of activity.[8,11,12] Platinum(II) complexes such as carboplatin and oxaliplatin as well as other platinum analogues are the products of this search (**Figure 6**). Other metal complexes containing ions such as zinc(II), gold, and copper chelating agents have received considerable interest as anticancer agents.[8,13] Recently, the chemistry of ruthenium and gold-based compounds have received intensive scrutiny, due to renewed interest in providing an alternative to cisplatin, because of their promising cytotoxic and potential anticancer properties.[4,14,15]

Nevertheless, metal-based compounds, especially transition metals exhibit definite properties including their potential to undergo a redox reaction.[5] Therefore, metals and their redox activities are tightly regulated to maintain normal wellbeing.[5-21]

Recently, there is a growing demand for metal-based compounds in the treatment of cancer. This may be due to the scourge of cancer and to the greater extent the level of *in vitro* cytotoxic effect exhibited by metal-based compounds particularly those that are synthesized recently. In addition, ligand substitution and modification of existing chemical structures led to the synthesis of a wide range of metal-based compounds, some of which have demonstrated an enhanced cytotoxic and pharmacokinetic profile. Again, a different approach of cytotoxic drug design is recently been adopted. This involves conjugating metallic compounds with bile acid, steroid, peptide or sugar to allow direct drugs delivery to the cancer cell thereby circumventing some pharmacokinetic challenges. The objective of this review is to provide an overview of previous reviews on the cytotoxic effects of metal-based compounds while focusing more on newly designed metal-based compounds and their cytotoxic effect on the cancer cell line, as well as new approach to metal-based drug design in cancer therapy

2 Properties of metal complexes and metal-based compounds

Transition metals form member elements of the "d" block and are included in groups III-XII of the periodic table.[22] They form unique properties which include:

2.1 Charge variation.

In aqueous solution, metal ions exist as positively charged species. Depending on the existing coordination environment, the charge can be modified to generate species that can be cationic, anionic, or neutral.[23] Most importantly, they form positively charged ions in aqueous solution that can bind to negatively charged biological molecules.[8]

2.2 Structure and bonding.

Relative to organic molecules, metal complexes can aggregate to a wide range of coordination geometries that gives them unique shapes. The bond length, bond angle, and coordination site vary depending on the metal and its oxidation state.[23] In addition to this, metal-based complexes can be structurally modified to a variety of distinct molecular species that confer a wide spectrum of coordination numbers and geometries, as well as kinetic properties that cannot be realised by conventional carbon-based compounds.[8,24,25]

2.3 Metal – ligand interaction.

Different forms of metal-ligand interaction exist, however, these interactions usually lead to the formation of complexes that are unique from those of individual ligands or metals. The thermodynamic and kinetic properties of metal – ligand interactions influence ligand exchange reactions.[23] The ability of metals to undergo this reaction offers a wide range of advantage for the metals to interact and coordinate with biological molecules.[8]

2.4 Lewis acid properties.

Characterised by high electron affinity, most metal ions can easily polarize groups that are coordinated to them, thus facilitating their hydrolysis.[8,23]

2.5 Partially filled d-shell.

For transition metals, the variable number of electrons in the d-shell or f-shell (for lanthanides) influence the electronic and magnetic properties of transition metal complexes.[23]

2.6 Redox activity.

Many transition metals have a tendency to undergo oxidation and reduction reaction.[23] The oxidation state of these metals is an important consideration in the design of the coordination compound. In biochemical redox catalysis, metal ions often serve to activate coordinated substrates and to participate in redox-active sites for charge accumulation.

Metal complexes and metal-based compounds possess the ability to coordinate ligands in a three-dimensional configuration, thereby allowing functionalization of groups that can be shaped to defined molecular targets.[8]

3 The scope of metal complexes in the treatment of cancer

Therapeutic potential of metal complexes in cancer therapy has attracted a lot of interest mainly because metals exhibit unique characteristics, such as redox activity, variable coordination modes and reactivity towards the organic substrate.[8] These properties become an attractive probe in the design of metal complexes[8] that will selectively bind to the biomolecular target with resultant alteration in cellular mechanism of proliferation. **Table 1** below provides a summary of *in vitro* cytotoxic effect of various metal-based compounds within the period of six years with particular reference to their proposed mechanism of action and target.

Table 1. An update on the anticancer activities of metal-based complexes (2010-2016)

Metal complexes	**Molecular formula**	**Proposed Mechanism of Action**	**Target enzymes/cell lines/Therapeutic indications**	***IC50* Range (µM)**
Carbene- metal complexes and related ligands				
Novel Gold(I) and Gold(III) –N Heterocyclic carbenes complexes[26]	$C_{52}H_{44}Au_2N_{12}P_2F_{12}$ $C_{26}H_{24}AuCl_2OF_6N_6P$`	Induction of apoptosis Inhibition of Thioredoxin reductase (TrxR)[26] Induction of reactive oxygen species (ROS)[26]	Thioredoxin reductase. A549, HCT 116, HepG2, MCF7 Chemotherapy of solid tumours.[26]	**$C_{52}H_{44}Au_2N_{12}P_2F_{12}$** 5.2 ± 1.5 (A549) 3.6 ± 4.1 (HCT 116) 3.7 ± 2.3 (HepG2) 4.7 ± 0.8 (MCF7)[26] **$C_{26}H_{24}AuCl_2OF_6N_6P$`** 5.2 ± 3.0 (A549) 5.9 ± 3.6 (HCT 116) 5.1 ± 3.8 (HepG2) 6.2 ± 1.4 (MCF7)[26]
Caffeine-Based Gold(1) N-Heterocyclic carbenes[27]	[Au(Caffeine-2-yielding)$_2$][BF_4][27]	Inhibition of protein poly-(adenosine diphosphate (ADP)-ribose) polymerase 1 (PARP-1)[27]	DNA A2780, A2780R, SKOV3, A549 and HEK-293T	0.54 – 28.4 (A2780) 17.1– 49 (A2780/R) 0.75 – 62.7 (SKOV3) 5.9 – 90.0 (A549) 0.20 – 84 (HEK-293T)
Ester- and Amide-Functionalized Imidazole of N-Heterocyclic Carbene Complexes[28]	{[Im^A]AgCl} {[Im^A]AuCl} {[Im^B]$_2$AgCl {[Im^B]AuCl HIm^ACl = [1,3-bis(2-ethoxy-2-oxoethyl)-1H-imidazol-3-ium chloride]	Inhibition of tyrosine by gold(I) NHC ligands thereby targeting TrxR.[28] CuNHC arrest cell cycle progression in GI phase.[28] Anticancer activity of Ag[1]NHC is based on	TrxR.[28] A375, A549, HCT-15 MCF-7 Human colon adenocarcinoma,[28] Leukaemia and breast cancer.[28]	**{[Im^A]AgCl}** 24.65 (A375) 22.14 (A549) 20.32 (HCT-15) 21.14 (MCF-7) **{[Im^A]AuCl}** 44.64 (A375) 42.37 (A549) 41.33 (HCT-15)

	$HIm^{B}Cl$ = {1,3-bis[2(diethylamino)-2-oxoethyl]-1H-imidazol-3-ium chloride}	highly lipophilic aromatic substituted carbenes.[28]		38.53 (MCF-7) **{$[Im^{B}]_2AgCl$}** 24.46 (A375) 16.23 (A549) 14.11 (HCT-15) 15.31 (MCF-7)[28]
Novel Ruthenium(II) N-Heterocyclic carbenes[29]	$[(\eta^6$ –p-cymene$)_2$ $Ru_2(Cl_2)_2]$ NHC	Mimic iron.[30] Interact with plasmidic DNA[30]	DNA-as target Caki-1, MCF-7 Chemotherapy of solid tumour[31]	13-500 (Caki-1) 2.4-500 (MCF-7)[29]
Caffeine derived Rhodium(I) N-Heterocyclic carbene complexes[32]	[Rh(I)Cl(COD)(NHC)] complexes	Inhibition of TrxR.[32] Increase in reactive oxygen species (ROS) formation.[32] DNA damage.[32] Cell cycle arrest.[32] Decrease mitochondria membrane potential.[32]	TrXR[32] MCF-7, HepG2 MDA-MB-231, HCT-116, LNCaP, Panc-1 and JoPaca-1 Chemotherapy of solid tumour[32]	84 (HepG2) 20 (HCF-7) 23 (MDA-MB-231) 35 (JoPaca-1) 49 (Panc-1) 80 (LNCaP) 9.0 (HCT116)[32]
N- Heterocyclic carbene- Amine Pt(II) complexes[33]	NHC (PtX2)-amine complexes[33]	Nuclear DNA platination.[33]	Target DNA[33] KB3-1, SK-OV3, OVCAR-8, MV-4-11, A2780, A2780/DPP Chemotherapy of solid and non-solid tumours[33]	2.5 (KB3-1) 4.33 (SK-OV3) 1.84 (OVCAR-8) 0.60 (MV-4-11) 4.00 (A2780) 8.5 (A2780/DPP)[33]

2-Hydroxy-3-(hydroxyamino)-4-oxopentan-2 ylidene)benzohydrazide derivatives[34]	[(HL)Cu(OAc)(H_2O)$_2$]·H$_2$O $C_{14}H_{21}N_3O_9Cu$	Bind to DNA[34]	Target DNA[34] HepG2 Chemotherapy of solid tumours[34]	2.24-6.49 (HepG2)[34]
Molybdenum(II) allyl dicarbonate complexes[35]	[Mo(allyl)(CO)$_2$(N-N)(py)]PF_6	DNA fragmentation[35] Induction of apoptosis[35]	Target DNA[35] NALM-6 MCF-7 HT-29 Chemotherapy of solid and non-solid tumours[35]	1.8-13 (NALM-6) 2.1-32 (MCF-7) 1.8-32 (HT-29)[35]
Metal-Arene complexes and other ligands				
Ruthenium(II)-Arene complex[36]	[(η^6- arene)RuII (en)Cl]$^+$	DNA damage[36] Cell cycle arrest[36] Induction of apoptosis[36]	Target DNA AH54 and AH63 Chemotherapy of colorectal cancer[36]	**$C_{15}H_{18}C_lF_6N_2PRu$** 16.6 (AH54)[36] **$C_{16}H_2OC_lF_6N_2PRu$** 10.9 (AH63)[36]
Novel Ruthenium-arene pyridinyl methylene complexes[37]	[(η^6-p-cymene)RuCl (pyridinylmethylene)	DNA binding[37]	Target DNA[37] MCF-7 and HeLa Chemotherapy of solid tumour[37]	07.76-25.42 (MCF-7) 07.10-29.22 (HeLa)[37]

Multi-targeted organometallic Ruthenium(II)-arene[38]	[(η^6 –p-cymene) $RuCl_2$]2-poly ADP-ribose polymerase (PARP) and PARP-1 inhibitors[38]	DNA binding[38] PARP-1 inhibition[38] Transcription inhibition[38]	Target DNA[38] A549, A2780, HCT116, HCC1937 and MRC-5 Chemotherapy of solid tumours[38]	85.1-500 (A549) 38.8-500 (A2780) 46.0-500 (HCT116) 93.3-500 (HCC1937) 143-500 (MRC-5)[38]
Ruthenium(II)-arene complexes with 2-Aryldiazole ligands[39]	[(η^6- arene)RuX(k^2 -N,N-L)]Y	DNA binding[39] Inhibition of CDK1	Target DNA[39] A2780, A2780cis, MCF-7 and MRC-5 Chemotherapy of solid tumours	11-300 (A2780) 11-34 (A2780cis) 26-300 (MCF-7) 25-224 (MRC-5)[39]
Osmiun(II)-arene carbohydrate base anticancer compound[40]	Osmium(II)-bis [dichloride(η^6-p-cymene)	DNA binding[40]	Target DNA[40] CH1, SW480 and A549	50-746 (CH1) 215-640 (SW480) 640 (A549)[40]
Ruthenium(II)-arene complexes with carbosilane methallodendrimers[41]	G_n-[$NH_2Ru(\eta^6$ –p-cymene)Cl_2]m	Interaction with DNA[41] Interaction with Human Serum Albumin (HSA)[41] Inhibition of cathepsin-B[41]	Target DNA[41] HeLa, MCF-7, HT-29 MDA-MB-231 and HEK-239T Chemotherapy of solid and non-solid tumours[41]	6.3-89 (HeLa) 2.5-56.0 (MCF-7) 3.3-41.7 (HT-29) 4-74 (MDA-MB -231) 5.0-51.9(HEK-239T)[41]
Ruthenium(II) complexes with aroylhydrazone ligand[42]	[Ru(η^6-C_6H_6)Cl(L)]	Induction of apoptosis[42] Fragmentation of DNA[42]	Target DNA[42] MCF-7, HeLa, NIH-3T3 Chemotherapy of solid tumour[42]	10.9-15.8 (MCF-7)[42] 34.3-48.7 (HeLa) 152.6-192 (NIH-3T3)
Cyclopentadienyl complexes and other ligands				

Iridium(III) complexes with 2-phenylpyridine ligand[43]	[(η^5 -Cp*)Ir(2- (R′-phenyl)-R-pyridine)Cl	Interaction with DNA nucleobases[43] Catalysis of NADH oxidation[43]	Target DNA[43] A2780, HCT-116, MCF-7 and A549 Chemotherapy of solid tumour[43]	1.18-60 (A2780) 3.7-57.3 (HCT-116) 4.8-28.6 (MCF-7) 2.1-56.67 (A549)[43]
New iron(II) cyclopentadienyl derivative complexes[44]	[Fe(η^5-C_5H_5)(dppe) L][X]	Interaction with DNA[44] Induction of apoptosis[44]	Target DNA HL-60 Chemotherapy of non-solid tumours[44]	0.67-5.89 (HL-60)
Ruthenium(II) cyclopentadienyl complexes with carbohydrate ligand[45]	[Ru(η^5-C_5H_5)(PP) (L)][X]	Induction of apoptosis[45] Activation of caspase-3 and -7 activity[45]	HCT116CC, HeLa Chemotherapy of solid tumours[45]	0.45 (HCT116CC) 3.58 (HeLa)[45]
Ruthenium(II) cyclopentadienyl complexes with phosphane co-ligand[45]	[Ru(η^5-C_5H_5)(PP) (L)][X]	Induction of apoptosis[45]	HeLa Chemotherapy of solid tumour[45]	2.63 (HeLa)[45]
OrganoIridium cyclopentadienyl complexes[46]	[(η^5-Cpx)Ir (L^L′)Z]	Intercalation of DNA[46] Coordinate with DNA guanine[46]	HeLa Chemotherapy of solid tumour[46]	0.23 (HeLa)[46]

Several metal-based compounds have been synthesised with promising anticancer properties, some of which are already in use in clinical practice for diagnosis and treatment while some are undergoing clinical trials. Metal-based compounds synthesised recently are products of drug design targeted at achieving a specific objectives that the original compound could not achieve and such compounds exhibit a different spectrum of cytotoxicity. Compounds in this group include:

4 Platinum complexes and associated ligands

Platinum compounds, particular cisplatin, are the heartbeat of the metal-based compounds in cancer therapy. Clinical use of platinum complexes as an adjuvant in cancer therapy is based on the desire to achieve tumour cell death[47] and the spectrum of activity of the candidate drug,[47] Such complexes are mostly indicated for the treatment of cervical, ovarian, testicular, head and neck, breast, bladder, stomach, prostate and lung cancer. Their anticancer activities are also extended to Hodgkin's and non-Hodgkin's lymphoma, neuroblastoma, sarcomas, melanoma and multiple myeloma.[47] Although resistance to cisplatin emerged, it was the fundamental basis that triggered the search for alternative metallic compounds with improved anticancer and pharmacokinetic properties. On this basis, alternative platinum compounds were derived. Carboplatin, oxaliplatin, satraplatin, omaplatin, aroplatin, enloplatin, zeniplatin, sebriplatin, meboplatin, picoplatin, satraplatin, omaplatin and iproplatin are all products of extensive research of platinum complexes (**Figure 1)**

The united states Food and Drug Administration (FDA) in 1978 approved Platinol,(R) a brand of cis-platin as combination therapy in the management of metastatic testicular, ovarian and bladder cancer.[48] FDA also approved Paraplatin,® a brand of carboplatin as combination therapy in the management of ovarian cancer,[48] numerous other platinum derivatives have been synthesized with established clinical success including oxaliplatin (Eloxatin®) , nedaplatin (Aqupla®) approved for use in Japan and lobaplatin approved for use in China.

Oxaliplatin branded as Eloxatan® was initially launched in France in 1996 and formally available in the countries of Europe in 1999 and launched in United states in 2002.[48] This is a platinum based drug with oxalate and diaminocyclohexane ligand (DACH). The DACH plays major a role in cytotoxicity and protect it against cross-resistance with cis-platin and oxaliplatin. It is licenced to be used as a combination therapy with other chemotherapeutic

agents in the management of colon cancer and non-small-cell-lung cancer.[49] This drug has better safety profile than cis-platin as such is used in patients that cannot tolerate cis-platin.[48]

Nedaplatin branded as Aqupla®, is a platinum derivative of cis-diamine-glycolate which was formally approved in Japan in 1995. The drug is said to have a better safety profile than cis-platin (Less nephrotoxic)[50] and used as combination therapy in the management of urological tumours.[50]

Lobaplatin is a derivative of the platinum compound, represented as 1,2-diammino-l-methyl-cyclobutane-platinum(II)-lactate. The antitumour activities of this compound span through the human lung, ovarian and gastric cancer xenograft.[48] It has non-cross resistance to cis-platin particularly human sensitive cancer cells. Lobaplatin was originally approved for use in the management of patients with chronic myelogenous leukaemia, small cell lung cancer and metastatic cancer.[51] Recently, a phase I clinical trial of dose escalation of lobaplatin in combination with fixed-dose docetaxel in the treatment of human solid tumour was established.[52] In this study, the maximum tolerable dose of lobaplatin when combined with docetaxel for the treatment of solid tumour known to have progressed after chemotherapy was established.[52] Positive results from phase I trials prompt researcher to recommend the same dosage for the phase II clinical trials.[52]

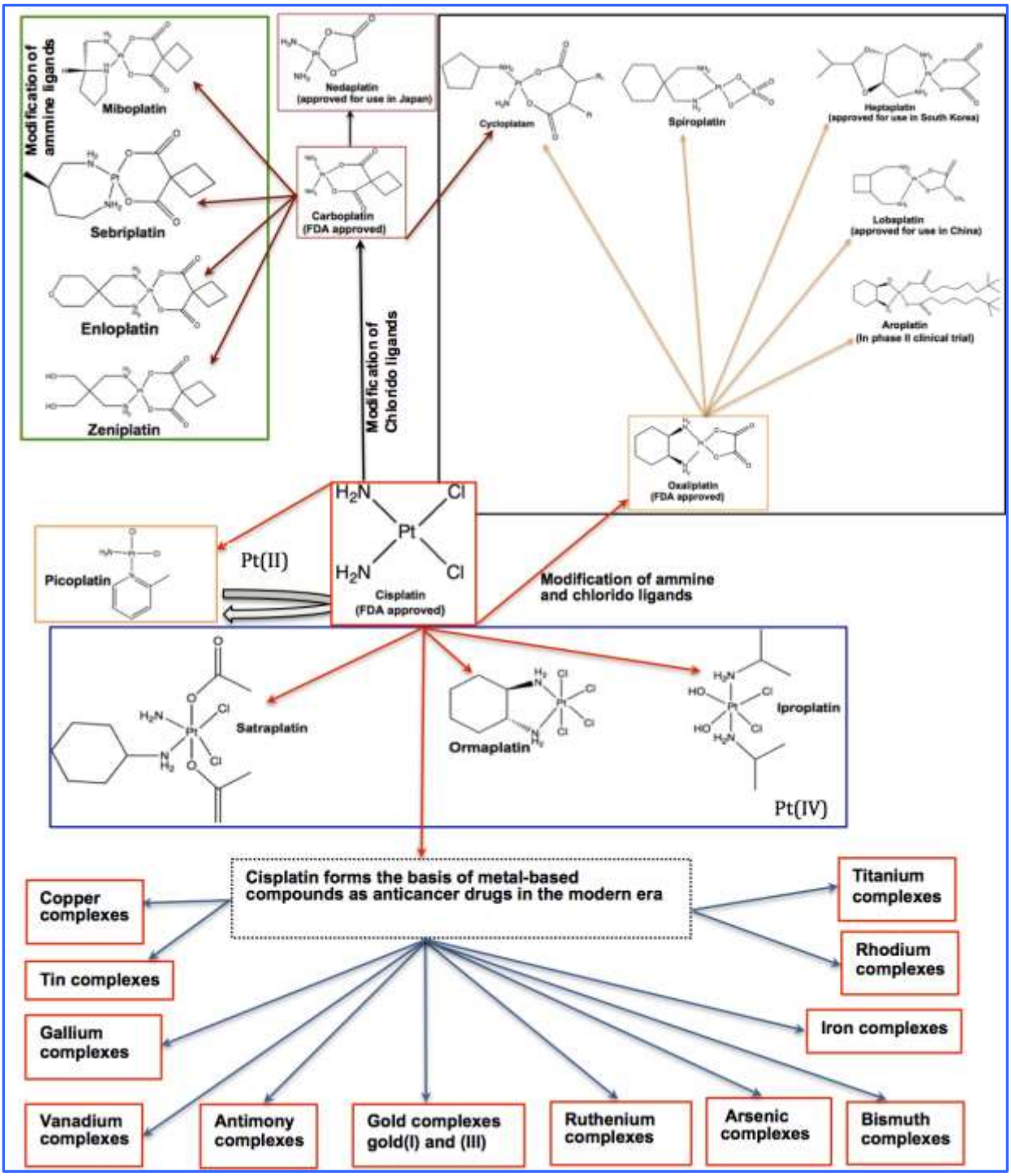


Figure 1. Evolution of organometallic complexes in cancer therapy

Picoplatin is a 2-methylpyridine analogue of cis-platin (formerly known as ZD0473) originally developed to provide steric cover around the platinum centre thereby providing a steric hindrance to the drug and prevents the attack from nucleophiles. It shields it against DNA-repair pathway that enhanced resistance.[53] Preclinical studies[54] revealed a promising anticancer activities in the resistant cell line to cis-platin.[49] However, after the conducted phase II clinical

trial, it was noted that picoplatin offers no superior advantage on the targeted cell line except a significant decrease in neurotoxicities.[49] In a related development, picoplatin is still undergoing phase I and II clinical trials as a treatment for colorectal cancer in combination with 5-flurouracil and leucovorin, also in combination with docetaxel for prostate cancer[55]

Satraplatin is the first orally bioavailable platinum drug. This drug is [bis-amino-dichloro-(cyclohexyl)amine]platinum(IV), exhibits varying pharmacodynamics and pharmacokinetic properties relative to other platinum compounds, hence may possess a different spectrum of anticancer activities.[48] The anticancer activities of satraplatin span through platinum-sensitive and resistant cell lines including cervical, prostate, ovaries and lung cancer.[55] Non-linear pharmacokinetics was one major challenge encountered during the initial studies of satraplatin that led to the study being abandoned.[55] Satraplatin has undergone several phases of clinical trials, phase III clinical trials examined satraplatin and prednisolone combination against refractory cancer,[55] satraplatin is currently targeted in phase I, II and III trial in combination with other drugs such as docetaxel in the treatment of prostate cancer.

Lipoplatin is a liposomal form of cis-platin designed to enhance the pharmacokinetic safety profile and allow dosage manipulation while targeting cancer cells.[55] The liposomes are made of dipalmitoyl phosphotidyl glycerol, soyphosphatidyl choline cholesterol and methoxypolyethylene glycol distreatoyl phosphatidyl ethanolamine.[55] The presence of liposomes offered a circulatory advantage to the drug. Lipoplatin has undergone phase I, II and III clinical trial with main focus on its anticancer activity in small cell lung cancer.[55] It is also being investigated for breast, pancreatic, head and neck anticancers.

ProLindac is an oxaliplatin with hydrophilic (hydroxypropylmethacrylamide) polymer bonded to the active moiety. This drug is designed to target the solid tumour with enhanced retention within the tumour cells.[56] Anticancer activities of proLindac have been evaluated in mice with ovarian carcinoma and B16 melanoma.[55] In this study, an optimum growth inhibitory effect was achieved in addition to a decrease in toxicity towards a normal cell and sustained plasma concentration. This drug has undergone both phase I and phase II clinical trials with profound anticancer activities. It is used as a mono-therapy in advance ovarian cancer resistant to platinum therapy.[57] The drug is currently undergoing phase III clinical trials in the treatment of head and neck cancer.[57]

Platinum drugs discontinued

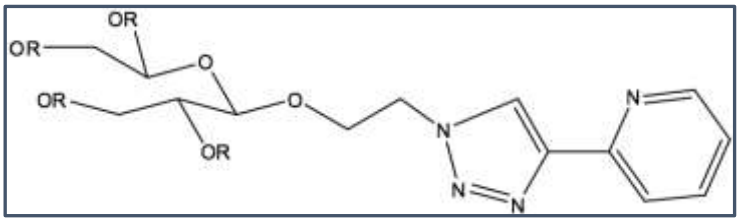
In the midst of several challenges on the use of platinum drugs in the management of a spectrum of cancers, many platinum compounds have been synthesised and demonstrate good *in vitro* cytotoxic activity. However, the use of these compounds for the treatment of cancers is subject to scientific evidence from clinical trials. Such result must reflect; a pharmacokinetic safety profile, pharmacological properties of the drugs and at times social-economic factors of patients is also considered. In this case, studies of platinum compounds are discontinued either due to toxicity or lack of profound anticancer activity on the patient. Many platinum-based drugs have been screened through clinical trials in an attempt to find an alternative to cis-platin, mainly due to dose-related adverse effects or resistance to cis-platin. The lack of superior advantage of these drugs over cis-platin led to the suspension from the study. Platinum drugs discontinued from clinical trials include sebriplatin, spiroplatin, cycloplatam, miboplatin, SPI-077, aroplatin, BBR3464, TRK-710, iproplatin, zeniplatin enloplatin ormaplatin JM 11 and NSC 170898.

5 New platinum complexes as a product of drug design

Recently, more platinum complexes are being synthesized and their anticancer of activities against tumour cell lines are being evaluated. This involves remodelling of the parent compound (platinum) by conjugating it with the different ligand to achieve the desired outcome. In most cases, pharmacokinetic parameters, spectrum of activity and toxicity profile are improved to circumvent those challenges inherent from the parent compound. This includes:

Platinum complexes conjugated with sugar

Carbohydrate (sugar) conjugation may bring about biological and physicochemical changes to the platinum compound.[58] The changes may be in the form of improved solubility, a decrease in adverse effects and improvement in cellular uptake of the drug.[58] Most of the newly synthesized platinum compounds have been evaluated for their anticancer activities. Recently, compounds such as [$PtIICl_2$(AcGlc-pyta)] have been synthesized and their cytotoxicity profiles evaluated *in vitro* against human cervical tumour (HeLa) and compared with the effect of cis-platin and the compound shows less cytotoxicity.[58] **Figure 2** is a typical example of sugar conjugated triazole ligand



AcGlc-pyta (a) R= Ac
Glc-pyta (b) R= H
Figure 2. Sugar conjugated triazole ligands

In order to deliver the platinum drugs directly into the cells, 2-deoxyglucose conjugated (2-DG) platinum(II) (conjugated platinum (II) complex for glucose transporter-1) also known as GLUTI-1 was designed to transport the drug to the cancer cells.[59] This was possibly due to the fact that to maintain cellular homeostasis, growth, and proliferation, malignant cells exhibit a higher rate of glycolysis than the normal cells.[59,60] GLUTI-1 was evaluated for its potentials to transport the drug complex with the aid of comparative molecular docking analysis using the latest GLUTI crystal structure. Molecular dynamics (MD) was used in identifying the key binding site of 2-DG as a substrate for GLUTI.[59] Findings from docking result revealed that the 2-DG conjugated platinum complex can bind to the same binding site as GLUTI substrate.[59] This led to the synthesis of the conjugate and was evaluated for its cytotoxicity. This study provides critical information about the potential of 2-DG conjugated platinum(II) compound in pharmaceutical research and drug development.[59]

In a related development, a recent study on potent glucose platinum conjugate[60] revealed the potential of novel Glc-pts in cancer therapy. The crystal structure of bacterial xylose transporter XyIE recently published was used for the docking,[60] and C6-glucose-platinum derivative was initially optimized using DFT before docking it to the protein. A good hydrogen bond interaction with key glucose binding residues Gln168, Gln288, Tyr298, and Gln175[60] was revealed. The cytotoxicity studies conducted against ovarian cancer cell line (A2780) shows that A2780 cells were the most sensitive to Glc-pts compounds with IC_{50} value of 0.15-0.22 μM.

Chiral platinum complexes

The biological activities of the chiral compound have attracted a great deal of research. These compounds display a high level of selectivity and specificity and are involved in many biological events.[61] To overcome many challenges with cis-platin and other platinum drugs already in clinical use, numerous drug design approaches have been adopted. One of such is

metal-drug-synergism,[61] which can be achieved by combining a known active organic compound with a metal-based complex.[61] The structure of DNA and other biomolecules can be probed with chiral metallic complexes,[61] therefore, biological activities are greatly influenced by nature of chiral molecules present in the system.[61] In a recent study of chiral platinum complexes, the DNA binding characteristic of platinum is combined with chemical properties of phosphines and chloroquine.[61] The cytotoxicity profile of these compounds against cancer cell lines (IC_{50}) is as follows: MDA-MB-231 (200 μM for chloroquine and 2.44 μM for cisplatin phosphine complex), MCF-7 (82.0 μM for CQ and 13.98 μM for cisplatin phosphine complex), A459 (56.53 μM for CQ and 14.42 μM for CDDP), DU-145 (79.50 μM for CQ and 2.33 μM for CDDP), V79-4 (29.85 μM for CQ and 21.60 μM for CDDP) and L929 (25.94 μM for CQ and 16.53 μM for CDDP). In correlation with the biological activities of phosphine chloroquine platinum(II) complexes, it was revealed that the compounds display a lower biological activities even at relatively higher DNA interaction and higher affinity for blood plasma protein.[61] Due to high protein affinity exhibited by this compound, the effect on DNA is greatly reduced and consequently, the cytotoxic effect is also affected. A desired effort is needed to devise means of displacing or reducing plasma binding of these compounds to enhance its interaction with DNA.

Monofunctional platinum(II) complexes

Some platinum complexes have the ability to bind to DNA only through one coordination site, particularly chloride ligand.[62] Many of these compounds have been synthesis with only a few exhibiting *in vitro* cytotoxic effects against a range of cancer cell lines.[62] Phenanthriplatin a monofunctional platinum complex (**Figure 3**) exhibited a relatively better cytotoxic effect than cis-platin. It was thought that the biological activity of these complexes arises from monofunctional adduct. This is a sequel to the failure of antibodies of DNA containing bifunctional platinum adduct to recognise the lesion formed by cis-$[Pt(NH_3)_2(Am)Cl]+$.

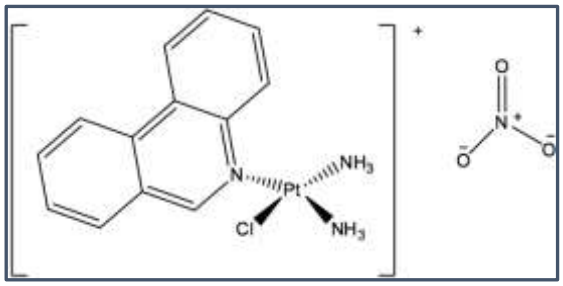


Figure 3. Structure of Phenanthriplatin

This mean that the complexes bind to DNA with a favourable biological effect through a monofunctional adduct formation.[62]

Platinum(IV) complexes and anticancer activity

Since the discovery of cis-platin and its clinical use, there has been an effort for the search of improved platinum drugs based on classical platinum(II)-diamine pharmacophores. This has only yielded few new drugs with a clinical advantage over the parent compound. Therefore new methods are being adopted to explore other available platinum-related compounds. This is centred around platinum(IV) complexes, whose anticancer properties have been recognized since last the decade.[62] The stability and expanded coordination sphere serves as an advantage of overcoming the challenges inherent from platinum (II) compounds. The platinum(IV) complex has six saturated coordination sphere of low spin d^6 with octahedral geometries.[62] This property gives them the kinetic stability over platinum(II) complexes.[62] In a study conducted to determine the cytotoxicity profile of newly synthesised platinum(IV) complexes on ovarian cancer cell line (TOV21G) and colon cancer cell line (HCT-116). Both complexes show dose and time dependent cytotoxicity towards the tested cell line[63] with highest cytotoxic effects on TOV21G. Platinum(IV) complex had a more pronounce cytotoxicity on TOV21G at a lower concentration.[63] However, platinum(IV) complex and cis-platin have a similar cytotoxic effect on HCT-116 cell line.[63]

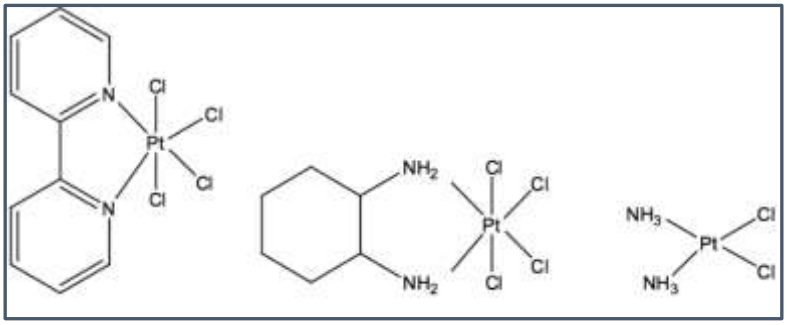


Figure 4. Structure of platinum(IV) complexes under investigation from left to right $[PtCl_4(bipy)]$, $[PtCl_4(dach)]$, *cis*-$[PtCl_2(NH_3)_2]$

This study provides an opportunity for exploring cytotoxic properties of platinum(IV) complexes in the treatment of colon cancer as an alternative to platinum(II) derived anticancer drugs. A good example of compounds under investigation are shown in **Figure 4**.

Table 2. Summary of metal-based compounds undergoing clinical trials in human

Drug name	Developers	Phase of clinical trial	Indications
Picoplatin[64]**(**JM473**)**	Pionard	Phase II	Treatment of colorectal cancer[64] in combination with 5-FU and leucovorin
Lipoplatin™ (Nanoplatin™, Oncoplatin)	Regulon	Phase II[64] & Phase III clinical trial in cancer cells.	Treatment of locally advanced gastric cancer / squamous cell in different cancer cells carcinoma of head and neck[64]
ProLindac™ (AP5046)	Access Pharm	Phase I,II & III trial	Advanced ovarian cancer,[64] head and neck cancer.
Satraplatin (JM216)	Spectrum Pharm and Agennix AG	Phase I, II and III trials	Treatment of colorectal cancer in combination with 5-FU & Leucovorin,[64] in combination with docetaxel in treatment of prostate cancer and treatment of a patient with progressive or relapse NSCLC[64]
NAMIA-A		Phase I	Metastatic tumour (lung, colorecta, melanoma, ovarian and pancreatic)[65]
KP1019		Phase II	Advanced colorecta cancer[65]
^{64}Cu-ATMS		Phase II	PET/CT monitoring therapeutic progress in patient with cervica[66]

Note: PET positron emission tomography, ^{64}Cu-ATSM ^{64}Cu-diacetyl-bis(N^4–methyl thiosemicarbazone), NSCL non-small lung cancer, CT computed tomography scan.

6 Ruthenium complexes in cancer therapy and drug design

For the past few decades, the use of metal-bases complexes in the treatment of cancer has been dominated by a spectrum of challenges emanating from the use of platinum complexes and platinum-derived drugs in clinical practice. Whilst it is virtually impossible to deny the success of platinum drugs in chemotherapy,[67] it is also important to note that it has its own limitations such as drug resistance, limited spectrum of activity and worsening side effects.[67] To surmount these challenges, efforts have been made to critically consider other metal-based complexes with cytotoxic properties, such as ruthenium, gold and iron complexes. Ruthenium compounds developed for this purpose are known to cause fewer and less severe side effects.[67] Ruthenium can form octahedral complexes that give opportunity for exploring more ligands compared to platinum(II) complexes that only form square planar complexes.[67] The most recent approach to ruthenium drug design is the development of a ruthenium organic directing molecule (RODM). In this case, the organic molecule binds to the active site of an enzyme and the attached ruthenium ion binds to nearby residues of the same enzyme.[67] The advantage of this approach is that there is a known biological target of a compound from which enzymological studies can be performed, such as studies of rate of enzyme inhibition.[67]

Other approaches include directed therapy and multinuclear approach. Ruthenium's ability to form multinuclear and supramolecular architecture has also been explored in drug design. This includes ruthenium cluster complexes, ruthenium DNA intercalators, and ruthenium-platinum-mixed metal compounds.[67] In directed therapy, ruthenium is chemically attached to an organic compound with known biological target that directs the drug to the cells thereby increasing the potency of the compound. The peculiarity of ruthenium as one of the least toxic metallic complexes is attributed, in part, to the ability of Ru(III) complex to mimic iron binding serum protein, thereby reducing the concentration of free plasma ruthenium and increase the concentration that reaches the cancer cells compared to healthy cells.[68] Plasma bound ruthenium complexes possess high affinity to cancer cells with transferrin receptors, this brings about diverse pharmacodynamics differences that exist between cancerous and healthy cells,[68] and form the basis of higher cytotoxicity experienced with KP1019 (A) compared to NAMI-A[68](B) see **Figure 5**.

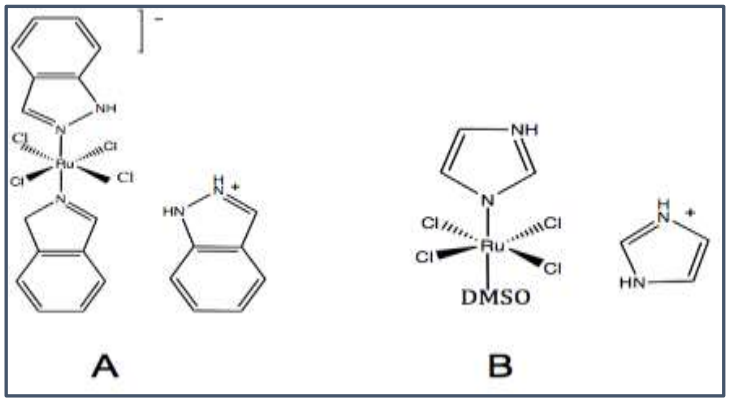


Figure 5. The structure of KP1019 (A) and NAMI-A (B)

The proposed transport mechanism of ruthenium(III) involves both passive diffusion, and transferrin-dependent mechanism.[69] In addition, the reduction of Ru(III) complex to Ru(II) following the cellular uptake of the former could also play an important role.[69] However, research evidence has linked the cytotoxic effect of ruthenium to Ru(II) arene complexes.[68]

Ru(II) arene complexes incorporated into amphiphilic 1,3,5-triaza-7-phosphaadamantane (pta) ligand *i.e* Ru(η^6 – toluene)-(pta)Cl_2, RAPTA-T, and Ru(η^6 –*p*-cymene)(pta)Cl_2, RAPTA-C, (**Figure 6**) exhibits low toxicity *in vivo*[68] and are not cytotoxic, but demonstrate relevant antitumour properties.[68] RAPTA-C selectively bind to histone protein core in chromatin,[70] resulting into aquation of chloride ligands[68] and translates to mild growth inhibition on primary tumours *in vivo.*[71] The combination of RAPT-A compounds gives unique effects, and when applied in combination with other drugs, RAPTA-C (**Figure 6**) produces efficient inhibition of tumour growth at very low doses and devoid of toxic side effects[68]. Anticancer activities of ruthenium complexes are summarised in **Table 1**.

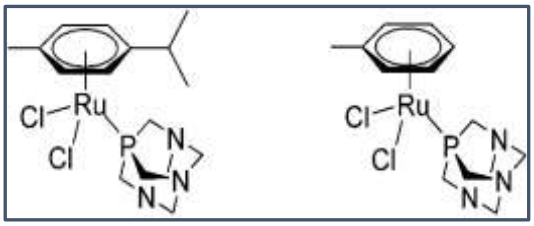


Figure 6. RAPTA-C (on the left) and RAPTA-T (on the right)

Other ruthenium complexes that are recently synthesize and shows the cytotoxic effect on the cancer cell lines include:

Cyclometallated ruthenium compounds

In an attempt to provide the best alternative to the platinum compounds and other metallo-tumour drugs, many approaches to the design of anticancer drugs have been adopted. One of such approaches is he design of cyclometallated ruthenium compound with cytotoxic properties.[72] In a recent study, the cytotoxic effect of four Ru(II) dyes incorporated to cyclometallated ligand phpy-(deprotonated 2-phenylpyridine) was evaluated against HeLa cells.[72] In this study, the cytotoxic activity of all the compound was similar to that of cisplan.[72] The activity of compounds were compared, compound [Ru(phpy)(bpy)- (dppn)]+ (4; bpy = 2,2′-bipyridine, dppn = benzo[i]dipyrido- [3,2-a:2′,3′-c]phenazine) being the most active among all, with an activity of about 6-fold higher than the platinum drug,[72] and possess the ability to disrupt the mitochondria membrane potential.[72] Compounds such as [Ru(phpy)(biq)2]+ (3; biq = 2,2′-biqinoline), have absorption of 640 nm, which shows a profound activity upon irradiation with 633 nm light.[72] Based on this finding, it was concluded that coordinatively saturated cyclometallated Ru dyes are potential class or family of the new metal-based anticancer compounds.[72]

Half-Sandwiched Ruthenium(II) Compounds Containing 5-Fluorouracil Derivatives

Combination of two or more multifunctionality groups bring into play different properties of compounds,[73] this is a popular strategy adopted in the design of new therapeutics.[73] The use of 5 –Fluorouracil in the treatment of cancer has been associated with strong toxicities to gastric system, intestinal mucosa and bone marrow.[73] Thus, attempt are made to improve anticancer property and minimized its side effects by exploiting other prodrugs.[73] In line with this, respective advantages of ruthenium(II) compound and 5-fluorouracil[74] can be explored and the combination results in synergistic action.[74] Half sandwich ruthenium arene complexes allow

the introduction of numerous biologically active groups,[73, 75] this way many ligands can be introduced into the drug either to improve the spectrum of activity against the cancer cells or to modify the pharmacokinetic profile of the drugs. In a recent study, the two novel coordination compounds of half sandwiched ruthenium(II) containing 2-(5-fluorouracil)-yl-N-(pyridyl)-acetamide were synthesised.[73] The result from DNA intercalation binding reveals that the compounds exhibit activity on DNA[73] which may be interpreted as being the cytotoxic effect on cancer cells.

7 Titanocenes

Titanocenes are a class of metal-based cytotoxic agents with a distinct mechanism of action and spectrum of activity from platinum complexes.[76] Studies have shown that titanocenes C [bis-(N,N-dimethylamino-2(N-methylpyrrolyl)-methyl-cyclopentadienyl)titanium(IV) dichloride] inhibit the proliferation of human tumour cell lines with mean IC_{50} of 48.3±32.5 μM. The activity against small cell lung cancer (SCLC) was more profound with a profile different from cis-platin.[76] Both titanocene C and Y has little or no cross-resistance to cis-platin and oxaliplatin invariant HL-60 cell lines. Titanocene C is particularly favoured in the induction of cell cycle arrest at G 1/0-S interphase.[76] In a related development, the cytotoxic effect, and mechanism of action of titanocene difluoride in ovarian cancer was evaluated.[77] In this study, three titanocene difluoride compounds were used, two of which bears carbohydrate moiety (α-D-ribofuranos-5-yl) and one has no substitution.[77] In contrast to the mechanism of action of cisplatin that involve DNA damage, activation of p53 protein and induction of apoptosis,[68] the result from this study shows that the mechanism of action of titanocene difluoride derivatives is mediated *via* the endoplasmic reticulum stress pathway and autophagy.[77] In conclusion, the author stated that the cytotoxic effects of titanocene difluoride are comparable to that of cis-platin and are more effective in cis-platin resistance cell line and therefore recommends that these compounds be considered as potential agents in the management of cis-platin resistance cases.[77]

8 Anticancer properties of copper complexes

Copper complexes exhibit cytotoxic properties with the mechanism of action different from that of the clinically used platinum compound, cis-platin.[78] A spectrum of activities varies among these compounds, depending on the type of ligand attached to the simple copper complex. Cytotoxic effects of copper-based complexes have been investigated based on the assumption that endogenous copper may be less toxic for normal cell relative to cancer cells.[79]

The situation is entirely different, copper can undergo redox activity and competitively bind to the site that could otherwise be occupied by other metals.[79] Copper is an essential cellular element necessary for many biological pathways. It is also a cofactor in enzyme catalytic processes.[79] The role of copper in angiogenesis has been a subject of controversy, generally, the role of metals in this process is still under scrutiny.[79] However, copper complexes are known to mimics superoxide dismutase (SOD),[80] an important antioxidant enzymes which protect cells from harmful radical superoxide through its dismutation to non-toxic molecules[80] they are present almost in all living organism.[80] The native SOD are natural scavengers of free radicals that work very efficiently in the body,[80] therefore, in the presence of an imbalance between the generation of free radicals and the concentration of dismutation enzymes,[80] a low molecular weight mimics of antioxidant enzymes especially SOD with good scavenging activity is required.[80] Copper complexes particularly a low molecular weight are indicated, since they are good mimics of SOD.

It has been reported that a mixture of copper salt with dithiocarbamates (DTCs) and clioquinol (CQ) bind spontaneously to cellular copper to form a proteasome and an apoptosis inducer.[79] Recently, copper(II) complex $[Cu(C_20H_{22}NO_3)_2]\cdot H_2O$ was synthesised and cytotoxic activity evaluated.[81] The complex was investigated for its interaction with calf thymus DNA and BSA using spectroscopic methods, the results revealed that the binding mechanism is a static quenching process.[81] The *in vitro* cytotoxic evaluation study was conducted using MTT assay and the result revealed that a copper complex exhibit enhanced cytotoxicity, high selectivity and dose-dependent cytotoxicity.[81]

In a related development, four novel mononuclear Schiff base copper(II) complexes were synthesised and characterised by X-ray crystallography.[82] This include $[Cu(L)(OAc)]\cdot H_2O$, $[Cu(HL)(C_2O_4)(EtOH)]\cdot EtOH$,[Cu(L)(Bza)]and[Cu(L)(Sal)](HL=1-(2-(2-hydroxypropyl) (aminoethyl)(imino)(methyl)naphthalene-2-ol), Bza = benzoic acid, Sal = salicylic acid). Antiproliferative effects of these complexes were evaluated, the result showed that all the four complexes demonstrated good cytotoxicity against cancer cell lines.[82] Complexes that had salicylate attached to it as an auxiliary ligand showed better anticancer activity relative to the others.[82] It was suggested that the enhanced activity of complex with the highest activity may be due to the presence of a Schiff base complex and a nonsteroidal anti-inflammatory drug that might have contributed to the process of cell death.[82]

9 Gold complexes in cancer therapy

The cytotoxic property of gold complexes has attracted attention recently. This may not be unconnected to the various level of challenges witnessed with clinical use of platinum compounds. Gold(III) complexes are an emerging class of metal complexes with potential antitumor properties alternative to cis-platin. This is mainly due to their outstanding cytotoxic properties exhibited through non-cisplatin antitumor mechanism. The main objective of designing these drugs is to have a product that is very effective, less toxic and selectively binds to the active site of enzymes.[83] The potential of selectivity of gold(III) complexes to thiol-containing enzymes such as thioredoxin reductase (TrxR), make it an attractive probe in designing compounds that can selectively bind to residues in the active site of the enzyme.

Many forms of gold(III) complexes have been synthesized and the anticancer activity evaluated against cancer cell lines. In most cases, the ligands are either Cl, Br, S or P. Other forms of gold(III) complexes have also been synthesized with proven cytotoxicity.[84] Most of the reported cytotoxic gold(III) complexes have a profound effect on cis-platin resistant cell lines.[83] I n a recent study, the cytotoxic effect of an organometallic titanocenes-gold compound [(η-$C_5H_5)_2Ti\{OC(O)CH_2PPh_2AuCl\}_2$] was evaluated *in vitro* against prostate and renal cell line as potential chemotherapeutics in renal cancer.[85] The result showed that the compound acts synergistically because the resulting cytotoxic effect is more pronounced, when compared to monometallic tetanocene dichloride and gold(I) [$\{HOC(O)RPPh_2\}AuCl$] (R = $-CH_2-$ 6, $-4-C_6H_4-$ 7) in renal cancer cell lines.[85]

Thioredoxin reductase: a target for gold(III) complex. The binding of inhibitors (drugs) to the active site of the enzyme, depends on the type of residues (Amino acids) on the catalytic site. The catalytic residues are highly specific for certain compounds. These characteristics provide a platform for designing a compound that can specifically target certain residues resulting in ligand-residue interaction. Since some enzymes possess more than one binding site (binding pocket), such binding sites can be filtered to get the best out of many. The relationship termed as "intimacy" exist between the best binding site and the compounds.

Metal-based compounds bind to specific residues, most of which are involved in enzyme catalysis. The binding of metal-based complexes to these residues results in alteration of cellular processes and consequently cell death (apoptosis).[26] For example, gold(III) complexes bind to thioredoxin reductase (TrxR).[26] The catalytic residues on this enzyme are located between two chains, with each chain contributing to the binding of gold(III) complexes. Other

binding pockets have been identified in the same enzyme with less binding affinity. **Figure 7** depicts the catalytic residues in TrxR.

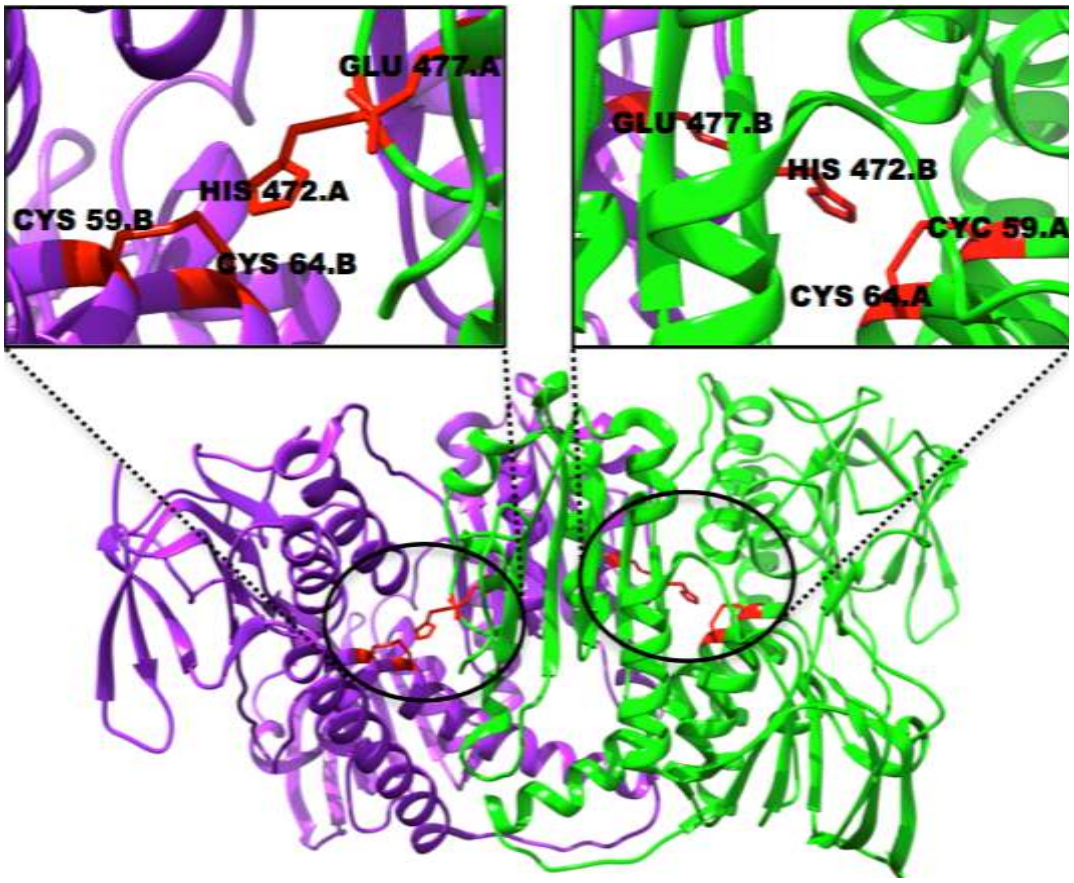


Figure 7. A 3D structure of the thioredoxin reductase homodimer (PDB entry 2J3N), with two chains in green and purple, respectively. The active site residues Cys 59(B), Cys 64(B), His 472(A), and Glu 477(A), represent the possible binding site for the gold(III) compounds.

10 Silver complexes in cancer therapy

Silver complexes have since been known for its antimicrobial activities[86] and are widely used in the treatment of infected wounds and burn cases.[86] In the past, silver complexes did not receive much attention compared to other metals,[86] although silver complexes also demonstrated good cytotoxic activity against many cancer cell lines.[86] Recently, cytotoxic properties of silver(I) complexes has attracted a great deal of interest, this is because most silver(I) complexes have been found to exhibit a greater cytotoxic activity than cisplatin[86] with relatively low toxicity and greater selectivity towards cancer cells.[86] An *in vitro* study conducted to assessed the cytotoxic properties of silver(I) complexes against cancerous B16 (murine melanoma) and non-cancerous 10T1/2 (murine fibroblast) cells line,[86] silver complexes containing hydroxymethylene group exhibited greater cytotoxic activity against

B16 (murine melanoma) than $AgNO_3$, AgSD and cisplatin.[86] These complexes were found to exhibit relatively low toxicity against non-cancerous 10T1/2 (murine fibroblast).[86] Similarly, a study set to determine the anticancer properties of gold(I) and silver(I) N-heterocyclic carbene complexes[87] revealed that these compounds along with cisplatin exhibited similar anticancer activity upon testing on H460 lung cancer cell line.[87]

In a related development, silver complexes of 2,6-disubstituted pyridine ligands were synthesized,[88] **Figure 8** the ligands and the complexes were evaluated *in vitro* with doxorubicin (reference compound) in hepatocellular carcinoma (HepG2), lung adenocarcinoma (A549), colon carcinoma (HT29) and breast adenocarcinoma (MCF7) using MMT method.[88] All the synthesized complex exhibited significant activity more than the corresponding ligands,[88] and most of the prepared silver complexes exhibited outstanding cytotoxic activity against tested cancer cell line compared to the doxorubicin.[88] All these properties placed silver complexes as a promising metal complex to be targeted in future for chemotherapy.

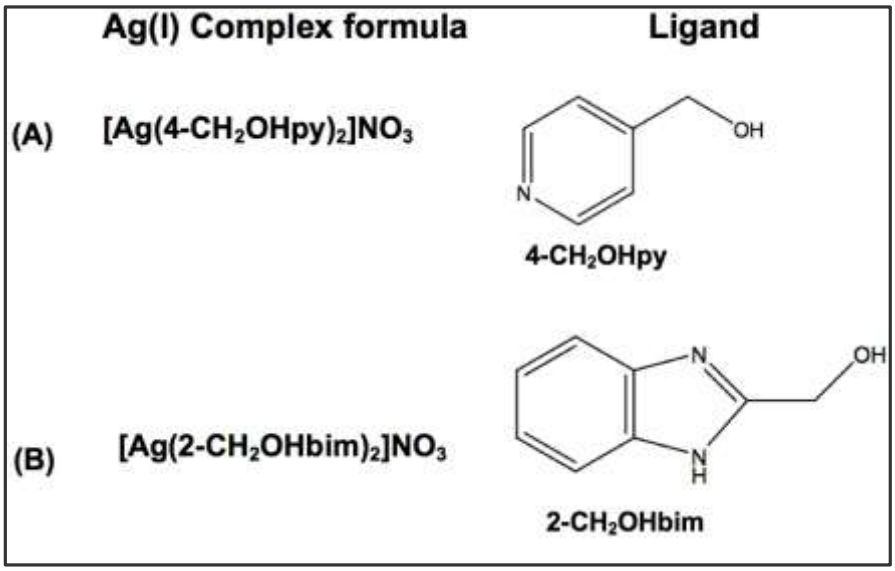


Figure 8. Silver complexes with 2,6-disubstituted pyridine ligands

11 Metallocenes with less attention in cancer therapy

Despite renewed attention on metal-based compounds in cancer therapy, some other compounds have received less attention in the context of cytotoxic effects on cancer cell lines. This includes zirconocene, vanadocene, niobocene and molybdocene.[31]Although, zirconocenes demonstrated a good antiproliferative activities against several cancer cell lines,[31] such as lung

adenocarcinoma, head and neck tumour, anaplastic thyroid cancer, ovarian and colon cancer[31] they need further optimization to be used as anticancer chemotherapy.[31]

12 Relative safety issues associated with metal complexes

The role of metal-based complexes in cancer therapy cannot be over emphasis, because of potential chemotherapeutic and diagnostic properties exhibited by metal-based complexes. However, since the advent of cisplatin,[89] the main goal for drug design and development have been modification of toxicity profile[89] that reflect relative safety of the drug, circumvention of resistant and improvement on the spectrum of activity of the metal complexes.[89] Despite the major breakthrough with cisplatin in the treatment of cancer, the major challenge still remained severe side effects associated with the drug and this include dose-limiting nephrotoxicity, neurotoxicity, ototoxicity and emetogenesis.[89] This lead to development of carboplatin and other platinum-based cytotoxic drugs in response to necessity to reduce the toxic effect of cisplatin.[89] Unfortunately, some platinum-based drugs developed as a result of short-comings emanated from the clinical use of cisplatin are also associated with severe side effects that have prevented regulatory authorities from granting their marketing approval.[64] Drugs in this class include; JM-11 developed by Johnson Matthew company has not been granted marketing approval because it lacks better blood and renal clearance compared to cisplatin,[64] ormaplatin developed by NCI (USA)/ Upjohn has not been granted marketing approval because of its severe and unpredictable cumulative neurotoxicity,[64] zeniplatin developed by American Cyanamid has not been granted marketing approval because of its serious nephrotoxicity,[64] and spiroplatin developed by Bristol Myers has not been granted marketing approval because of its unpredictable renal failure.[64]

Similarly, gold(III) complexes have been found to exhibit toxic effects, the most adverse cases of gold complex toxicity are restricted to skin and mucous membrane as reported in case of blind clinical trial.[90] In a related development, increased ceruloplasmin and copper levels in various tissues have been associated to cancer progression.[91] Most of these adverse effects are dose related and can be circumvented by structural modification of the meta-based complexes to enhance selectivity and reduce unwanted effects on normal cell.

13 Nanoparticles in cancer therapy

Nanotechnology has greatly enhanced drug delivery system[92] and to a large extent provide a means of direct drug delivery to the active site thereby reducing the unwanted effects by limiting the drug effect to specific site leaving other tissues untouched. In cancer medicine,

nanoparticles provides an advanced bioavailability, *in vivo* stability, intestinal absorption, solubility, sustained and targeted delivery and therapeutic effectiveness of several anticancer drugs.[92] Most potent chemotherapeutic agents used in the treatment of cancer have narrow therapeutic index and have been used against several tumour types,[92] however, their cytotoxic effects affect both normal and cancerous cells.[92] This remained a big challenge in the dosing of metal-based complexes in cancer therapy. Therefore, the opportunity provided by nanoparticles to selectively target cancer cells and leave behind healthy cells untouched has gained interest in the design of metal-based cytotoxic drugs.

Metal-Based Nanoparticles

Metal-based nanoparticles are of different shapes and sizes and have been investigated for their role in diagnosis and drug delivery system,[92] most commonly available metal-based nanoparticles include nickel, gold, silver, iron oxide, gadolinium, and titanium dioxide particles.[92] Metal-based nanoparticles provide a large surface area that allowed incorporation of large drug dose.[92] To improve the specificity in the diagnosis of cancers, various types of highly specific and highly sensitive nanoparticle-based (NP) optical imagine platforms are being investigated.[93] NP-based diagnostic platforms offers a major advantage compared with other agents, it can be functionalized to specifically target tumour cells allowing the imaging and therapeutic agent to specifically delivered to those cells.[93] NP can be multifunctional and exhibits optical, magnetic and structural properties that is lacking in single molecule.[93] Since the tumour specific targeting is achieved by conjugating the surface of NPs with a molecule or biomarker that attached to tumour cell receptor, the knowledge of tumour specific receptor, biomarkers, homing proteins, and enzymes that permits selective cellular uptake of therapeutic and diagnostic agents[93] is absolutely important. In tumour targeting and conjugation, some molecules and biomarkers are precisely used, this include; peptides, proteins, nucleic acid and small molecule ligands.[93] It is possible to achieved a synergistic effects by conjugating multifunctional NP with different peptides and loading it with multidrug regimens,[93] thereby by reducing the fraction of each drug in the combination.

14 Selected targets in anticancer drug design

Platinum complexes are widely used, anticancer agents. Their use is primarily based on the pharmacological properties of cisplatin,[94] which act as a model for the design of other metal-based compounds for use in cancer therapy.[94] Since all clinically utilized platinum compounds share the same mechanism of action,[94] many researchers are now seeking to make increasingly

drastic measures to the general molecular framework shared by these compounds to achieve a novel mechanism of cell death.[68] Most of the research in this field is tailored towards ligand substitution method. This way, platinum-based compounds with different rates of substitution reactions at the central atom in the biological system can be developed.[94] These drugs are designed to have superior anticancer activity even in cisplatin-resistant cases and improved patient compliance. The current area of interest in cytotoxic drug design include:

14.1 Sugar targeting.

One of the major characteristics of the cancer cells is indiscriminate cell division. This occurs only in cancer cells relative to normal cells as a result of the continued supply of nutrients necessary for its metabolic process, particularly glucose, for survival.[95] The need for glucose is further aggravated by the altered metabolic states in which many cancer cells exist.[96] In line with this, a facet of bio-sugar can be exploited for drug targeting because of enhanced uptake of glucose by cancer cells.[96] For example, 2αα,3-diaminosugars complexes analogous to oxaliplatin were investigated and found to have promising activity.[51] Several other prospective platinum-based compounds complexed with glucose were also investigated with seemingly promising results.[51]

14.2 Steroidal targeting.

Sex hormones such as testosterone and estrogen play a vital role in drug targeting. In this case, incorporation of steroidal units into non-living group ligand is important, as a result, platinum complex is directed by its targeting unit (steroidal unit) to the tissues expressing the similar steroidal receptor.[51] For instance, estrogen receptor (ER) is a known drug target because of high expression of this protein on the surfaces of some cancer cells, particularly in breast cancer.[51] With the progression of research in this field, another ER has been discovered in addition to a prominent ER, designated as ERα. This estrogen receptor is termed ERβ.[97] It has been documented that the ERβ may even play a more important role by exhibiting antiangiogenic and antiproliferative properties.[97] The linkage of steroidal unit capable of interacting with the ER to a platinum centre can influence the anticancer activity by interfering with biological functions of the receptor or by permitting enhanced uptake of platinum complexes.[51] This way the DNA platination is increased with resultant higher chance of apoptotic cell death.[51] Just as ER targets platinum to the cancer cells expressing ER receptors, testosterone can target platinum to cancer cells expressing androgen receptor (AR).[51,86]

14.3 Bile acid target.

Bile acids are steroidal in nature and have been conjugated to platinum complexes[51] in an effort to deliver compounds directly to the liver cells since a number of transport proteins that take up bile salt from blood are expressed on the hepatic epithelial cells.[51] Several efforts have been made to conjugate platinum complexes to the bile acid in different manners and the resultant complex seems to have promising anticancer activity.[51] For example, a bile acid chelated to dicarboxylate motif bound to a *cis*-diammineplatinum(II) fragment was explored as an orally administered anticancer agent.[51] Preliminary *in vitro* assay revealed activity in cultured murine hepatoma cells,[51] and further research on a syngeneic orthotopic rat model of hepatocellular carcinoma confirmed that the complex had antitumor activity.[51]

Related steroid targeting. The translocator protein (TSPO) commonly known as peripheral benzodiazepine is known to regulate the transport of cholesterol and synthesis of steroids.[98] This protein has been suggested to be an important target in cancer treatment since it is overexpressed in numerous tumour cells.[98] Chelated platinum(II) complexes with bidentate thiazolylimidazopyridine are reported to interact strongly with this receptor.[51]

14.4 Folate targeting.

Folate is an important carbon source for many cellular pathways including DNA, RNA, protein methylation as well as DNA synthesis.[99] Cancer cell growth is rapid with enhanced folate uptake. The role folate in all these processes could conceivably be used as a baseline for drug targeting.[99] However, there is a limitation to the use of folate as a targeting agent of platinum complex. An early study of the interaction of cisplatin with cellular folates suggests that it would not be able to operate as a cytosolic agent in manner analogues to cisplatin.[100] However, researchers are not relenting in their effort to understand the potentials of folate in the selective drug targeting.

14.5 Peptide targeting.

Conjugation of platinum(II) complexes to the peptide resulting in the platination of complexes with anticancer activity.[51] The cyclic peptide c(CNRGC) with Asn-Gly-Arg sequence that targets CD13 receptor is over-expressed on the surface of certain cancer cells.[101] The target complex was more toxic to prostate cancer cells expressing CD13 than non-targeted carboplatin,[51] and the competition assays confirmed that the complex is taken up via the interaction with CD13.[51] Many forms of platinum complexes conjugated with peptides have been screened against cancer cell lines and a reasonable number of them exhibit a promising

anticancer activity.

15 Conclusion

Since the discovery of cisplatin, a great deal of research has been conducted in the therapeutic application of metal-based complexes. These compounds exhibit ambivalent dispositions some are associated with induction and progression of cancers, others have demonstrated efficacy in cancer treatment, whilst some displaying both properties. Despite the challenges emanated from the clinical use of platinum compounds, there is a growing demand for metal-based compounds in cancer therapy. This may be due to the scourge of cancer and to the greater extent the level of *in vitro* cytotoxic effects exhibited by metal-base compounds particularly those that have been synthesized recently. There is excitement among some researchers in this field, about that cancer cells can directly be targeted using a different approach of drug design as highlighted in this study. This development put to rest the fear of toxicity associated with many organometallics since the drugs are directly delivered to the cancer cells leaving behind healthy cells unharmed. However, the pharmacokinetic profile of most these drugs are yet to be ascertained in the human system, but keeping our hope alive that designing an metal-based compound to selectively target cancer cells is a major breakthrough in this field of research. Another approach in the field of anticancer drug design is the use of nanoparticles to target the biomolecules. Such method ensures that drugs are delivered to specific cancer cells. To this extent, the concept of selective targeting remains the hope of the future in developing therapeutics that would selectively target cancer cells, leaving behind healthy cells unharmed.

Acknowledgements

The authors acknowledge the School of Health Science, University of KwaZulu-Natal, Westville campus for financial assistance.

Disclosure

Authors declare no financial and intellectual conflict of interests.

References

1. Norn, S., Permin, H., Kruse, E., and Kruse, P. R. (200) Mercury a major agent in the history of medicine and alchemy. *Dans Medicinhist Årboq*. 36, 21-40.

2. Agnew, J. (2010) Medicine in the Old West: A History, 1850-1900. *McFarland*. 978-0-7864-5603-1.

3. Waxman, S., and Anderson, K. C (2001) History of the development of arsenic derivatives in cancer therapy. *Oncologist 6 Suppl*. 6(2), 3–10.

4. Fricker, S. P. (1996) Medical Uses of Gold Compounds.*Gold Bulleting*. 29 (2), 53–60.

5. Jungwirth, U., Kowol, C. R., Keppler, B. K., Hartinger, C. G., Berger, W., and Heffeter, P (2011) Anticancer activity of metal complexes: involvement of redox processes. *Antioxid. Redox Signal*.15(4), 1085–127.

6. Nicolini, M. (1997) Platinum and Other Metal Coordination Compounds in Cancer Chemotherapy. *Springer*. 54,978-1-4613-1717-3.

7. Ji, H.-F., Li, X.-J., and Zhang, H.-Y. (2009) Natural products and drug discovery. Can thousands of years of ancient medical knowledge lead us to new and powerful drug combinations in the fight against cancer and dementia? *EMBO reports*. 10(3), 194–200.

8. Frezza, M., Hindo, S., Chen, D., Davenport, A., Schmitt, S., Tomco, D., and Dou, Q. P (2010) Novel metals and metal complexes as platforms for cancer therapy. *Curr Pharm Des*.16(16), 1813–1825.

9. Bruijnincx, P., and Sadler, P. (2008) New trends for metal complexes with anticancer activity. Curr. Opin. *Chem. Biol* 12(12), 197–206.

10. Mourino, V., Cattalini, J. P., and Boccaccini, a. R. (2012) Metallic ions as therapeutic agents in tissue engineering scaffolds: an overview of their biological applications and strategies for new developments. *J. R. Soc. Interface*. 9(68), 401–419.

11. Brad T. Benedetti, Erica J. Peterson, Peyman Kabolizadeh, Alberto Martínez, Ralph Kipping, and Nicholas P. Farrell (2011) Effects of Noncovalent Platinum Drug–Protein Interactions on Drug Efficacy: Use of Fluorescent Conjugates as Probes for Drug Metabolism. *Mol pharm*. 8(3),940–948.

12. Bhargava, A., and Vaishampayan, U. N. (2009) Satraplatin: leading the new generation of

oral platinum agents. *Expert Opin. Investig. Drugs*.18(11), 1787–97.

13. Milacic, V., Fregona, D., and Dou, Q. P. (2008) Gold complexes as prospective metal-based anticancer drugs. *Histo Histopathol*. 23(1), 101–108.

14. Bindoli, A., Pia, M., Scutari, G., Gabbiani, C., Casini, A., and Messori, L.(2009) Thioredoxin reductase : A target for gold compounds acting as potential anticancer drugs, Coordination. *Chemistry Review*. 253, 1692–1707.

15. Kostova, I. (20-06) Ruthenium complexes as anticancer agents. *Curr. Med. Chem*. 13(9), 1085–107.

16. Karlenius, T. C., and Tonissen, K. F. (2010) Thioredoxin and cancer: A role for thioredoxin in all states of tumor oxygenation. Cancers Basel. 2(2), 209–232.

17. Arredondo, M., and Núñez, M. T. (2005) Iron and copper metabolism. *Mol. Aspects Med* 26,313–327.

18. Balamurugan, K., and Schaffner, W. (2006) Copper homeostasis in eukaryotes: Teetering on a tightrope. *Biochimica et Biophysica Acta.* 1763, 737–746.

19. Go, Y. M., and Jones, D. P. (2008) Redox compartmentalization in eukaryotic cells. *Biochim. Biophys. Acta - Gen. Subj*. 1780(11), 1273–1290.

20. Gupte, A., and Mumper, R. J. (2009). Elevated copper and oxidative stress in cancer cells as a target for cancer treatment. *Cancer Treat. Rev.* 35(1): 32–46.

21. Hentze, M. W., Muckenthaler, M. U., and Andrews, N. C. (2004). Balancing Acts.Cell, Molecular control of mamalian iron metabolism. 117(3), 285–297.

22. Pattan, S. R., Pawar, S. B., Vetal, S. S., Gharate, U. D., and Bhawar, S. B. (2012).The scope of metal complexes in drug design - A review. *Indian Drugs*. 49(11), 5–12.

23. Haas, K. L., and Franz, K. J. (2010). Application of metal coordination chemistry to explore and manipulate cell biology. *Chem Rev*. 109(10), 4921–4960.

24. Yan, Y. K., Melchart, M., Habtemariam, A., and Sadler, P. J. (2005). Organometallic chemistry, biology and medicine: ruthenium arene anticancer complexes. *Chem. Commun.* 14(38), 4764–4776.

25. Salga, M. S., Ali, H. M., Abdulla, M. A., and Abdelwahab, S. I. (2012). Acute oral toxicity evaluations of some Zinc(II) complexes derived from 1-(2-Salicylaldiminoethyl)piperazine

schiff bases in rats. *Int. J. Mol. Sci*. 13,1393–1404.

26 Saha, K. Das, and Dinda, J. (2014). Novel Gold(I) − and Gold(III) − N-Heterocyclic Carbene Complexes: Synthesis and Evaluation of Their Anticancer Properties. *Organometallics*.33(10), 2544–2548.

27. Stefan, L., Pirrotta, M., Monchaud, D., Bodio, E., Richard, P., Gendre, P. Le, Warmerdam, E., Jager, M. H. De, Groothuis, G. M. M., Picquet, M., and Casini, A. (2014).Caff eine-Based Gold(I) N - Heterocyclic Carbenes as Possible Anticancer Agents: Synthesis and Biological Properties.*Inorg.Chem*.53(4), 2296-2303.

28. Pellei, M., Gandin, V., Marinelli, M., Marzano, C., Yousufuddin, M., Dias, H. V. R., and Santini, C. (2012). Synthesis and Biological Activity of Ester- and Amide-Functionalized Imidazolium Salts and Related Water-Soluble Coinage Metal N - Heterocyclic Carbene Complexes.*Inorg.Chem*. 51(18), 9873-9882.

29. Hackenberg, F., Mu, H., Smith, R., Streciwilk, W., Zhu, X., and Tacke, M. (2013). Novel Ruthenium(II) and Gold(I) NHC Complexes: Synthesis, Characterization, and Evaluation of Their Anticancer Properties. *Organometallics*. 32(19), 5551-5560.

30. Dragutan, I., Dragutan, V., and Demonceau, A. (2015). Editorial of Special Issue Ruthenium Complex: The Expanding Chemistry of the Ruthenium Complexes. *Molecules. 20(9), 17244-17274.*

31. Pedro, Martins., Mara, Marques., Lidia, Coito., Armando, J L.Ponbeiro., Pedro, Viana Baptista., and Alexandra R. Fernandes. (2014).Organometallic Compounds in Cancer Therapy: Past Lessons and Future Directions. *Anti- Cancer Agent in Medicinal Chemistry*. 14(19), 1199-1212.

32. Jing-Jing, Zhang., Julienne, K. Muenzner., Mohamed, A. Abu el Maaty., Bianka, Karge., Rainer, Schobert., Stefan, Wölfl., and Ingo, Ott. (2016). A Multi-target caffeine derived rhodium(I) N-heterocyclic carbene complex: evaluation of the mechanism of action. *Dalton Trans.* 45, 13161.

33.Eloy, L., Saker, L., Poupon, J., Bombard, S., Cresteil, T., Retailleau, P., and Marinetti, A.(2013). Antitumor trans -N-Heterocyclic Carbene − Amine − Pt(II) Complexes: Synthesis of Dinuclear Species and Exploratory Investigations of DNA Binding and Cytotoxicity Mechanisms. *J.Med.Chem*. 56(5), 2074-2086.

34. El-tabl, A. S., El-waheed, M. M. A., Wahba, M. A., and El-fadl, N. A. E. A. (2015). Synthesis , Characterization , and Anticancer Activity of New Metal Complexes Derived from 2-Hydroxy-3- (hydroxyimino) -4-oxopentan-2-ylidene) benzohydrazide. *Bioinorganic Chemistry and Application*. 1, 2-14.

35. Pfeiffer, H. (2012). Synthesis and biological activity of molybdenum carbonyl complexes and their peptide conjugates. *Disertation*. 6-137.

36. Carter, R., Westhorpe, A., Romero, M. J., Habtemariam, A., Gallevo, C. R., Bark, Y., and Menezes, N. (2016). Radiosensitisation of human colorectal cancer cells by ruthenium (II) arene anticancer complexes. *Nat. Publ. Gr*. 6, 20569.

37. Raosaheb, G., Sinha, S., Chhabra, M., and Paira, P. (2016). Bioorganic & Medicinal Chemistry Letters Synthesis of novel anticancer ruthenium – arene pyridinylmethylene scaffolds via three-component reaction. *Bioorg. Med. Chem. Lett.* 26, 2695–2700.

38. Wang, Z., Qian, H., Yiu, S., Sun, J., and Zhu, G. (2014). Multi-targeted organometallic ruthenium (II) – arene anticancer complexes bearing inhibitors of poly (ADP-ribose) polymerase-1 : A strategy to improve cytotoxicity. *J. Inorg. Biochem*. 131: 47–55.

39. Rodr, A. M., and Espino, G. (2014). Derivation of structure − activity relationships from the anticancer properties of ruthenium(II) arene complexes with 2 - aryldiazole ligands. *Inorg.Chem.*53(20), 11274-11288.

40. Hanif, M., Nazarov, A. A., Hartinger, C. G., Kandioller, W., Jakupec, M. A., Arion, V. B., Dyson, J., and Keppler, B. K. (2010). Osmium (II)– versus ruthenium (II)– arene carbohydrate-based anticancer compounds : similarities and differences. *Dalto Trans.* 39(31),7345–7352.

41. Marta, Maroto-Díaz., Benelita, T. Elie., Pilar, Gómez-Sal., Jorge, Pérez-Serrano., Rafael, Gómez., María, Contel., and F. Javier de la Mata. Manuscript, A. (2016). Synthesis and anticancer activity of carbosilane metallodendrimers based on arene ruthenium (II) complexes *Dalton Trans*. 45, 7049-7066.

42. Mohan, N., Muthumari, S., and Ramesh, R. (2016). Ruthenium (II) complexes containing aroylhydrazone ligands. *J. Organomet. Chem.* 807, 45–51.

43. Millett, A. J., Habtemariam, A., Romero-canelo, I., Clarkson, G. J., and Sadler, P. J.(2015). Contrasting Anticancer Activity of Half-Sandwich Iridium(III) Complexes Bearing

Functionally Diverse 2 - Phenylpyridine Ligands. *Organometallics*. 34(11), 2683-2694.

44. Valente, A., Santos, A. M., Côrte-real, L., Robalo, M. P., Moreno, V., Font-bardia, M., Calvet, T., Lorenzo, J., and Garcia, M. H. SC. (2014). DNA interaction and cytotoxicity studies of new ruthenium(II) cyclopentadienyl derivative complexes containing heteroaromatic ligands. *J. Organomet. Chem.* 105(2), 241-249.

45. Fernandes, A. C., Florindo, P., Pereira, D. M., Pedro, M., Rodrigues, C. M. P., and Fátima, M. (2015). Cyclopentadienyl-Ruthenium (II) and Iron (II) Organometallic Compounds with Carbohydrate Derivative Ligands as Good Colorectal Anticancer Agents. *Organometallics.* 58(10), 4339-4334.

46. Qamar, B., Liu, Z., and Hands-portman, I. (2013). Organometallic Iridium(III) Anticancer Complexes with New Mechanisms of Action: NCI-60 Screening, Mitochondrial Targeting, and Apoptosis. *Chem Biol*. 8(6), 1335-1345.

47. Florea, A.-M., and Büsselberg, D. (2011). Cisplatin as an anti-tumor drug: cellular mechanisms of activity, drug resistance and induced side effects. *Cancers (Basel)*. 3(1): 1351–71.

48. Monneret, C. (2011). Platinum anticancer drugs . From serendipity to rational design. *Ann. Pharm. Fr*. 69(6): 286–295.

49. Chan, B. A., and Coward, J. I. G. (2007). Review Article Chemotherapy advances in small-cell lung cancer. *Journal of Thoraxic Disease*. 5(5), S565-S578

50. Uehara, T., Yamate, J., Torii, M., and Maruyama, T. (2011). Comparative Nephrotoxicity of Cisplatin and Nedaplatin : Mechanisms and Histopathological Characteristics. *J. Toxicol Pathol*. 2011; 24(2): 87–94.

51. Timothy, C. Johnstone., Kogularamanan, Suntharalingam., and Stephen J. Lippard. (2016). The Next Generation of Platinum Drugs: Targeted Pt(II) Agents, Nanoparticle Delivery, and Pt(IV) Prodrugs. *Chem Review*.116(5), 3436-3486.

52. Yu Peng., Yue-E Liu., Xiao-Can Ren., Xue-Ji Chen., Hui-Ling Su., Jie Zong., Zeng-Li Feng., Dong-Ying Wang., Qiang Lin., Xian-Shu Gao. (2015). A phase I clinical trial of dose escalation of lobaplatin in combination with fixed - dose docetaxel for the treatment of human solid tumours that had progressed following chemotherapy. *Oncology letters*. 9(1), 67–74.

53. Shah, N., and Dizon, D. S. (2009). New-generation platinum agents for solid tumors. *Future Oncology*. 5(1), 33–42.

54. David A. Karlin, Hazel B. Breitz and Paul I Weiden. (2012). Use of Picoplatin to Treat Colorecta Cancer. *United States Patent Application Publication*. 1, 1-19.

55. Nial, J. Wheate., Shonagh, Walker., Gemma, E. Craig., and Rabbab Oun. (2010). The status of platinum anticancer drugs in the clinic and in clinical trials. *Dalton Trans*. 39(35), 8113-8127.

56. Liu, D., Poon, C., Lu, K., He, C., and Lin, W. (2014). Self-assembled nanoscale coordination polymers with trigger release properties for effective anticancer therapy. *Nat. Commun.* 25(5),4182.

57. Michael G. Apps, Eugene H. Y. Choi and Nial J. Wheate. (2015). The state-of-play and future of platinum drugs. 22(4), R219-233.

58. Ohi, H., Ashizaki, M., Obata, M., Mikata, Y., Tanaka, R., Nishioka, T., Kinoshita, I., Sugai, Y., Okura, I., Ogura, S., Czaplewska, J. A., Gottschaldt, M., Schubert, U. S., Funabiki, T., Morimoto, K., and Nakai, M. (2012). Syntheses , characterization , and antitumor activities of platinum (II) and palladium (II) complexes with sugar-conjugated triazole ligands. Chemistry and Biodiversity. 9(9), 1903–1915.

59. Mi, Q., Ma, Y., Gao, X., Liu, R., Liu, P., Mi, Y., and Fu, X. (2015). 2-Deoxyglucose conjugated platinum (II) complexes for targeted therapy : design , synthesis , and antitumor activity. *Journal of Biomolecular Structure and Dynamics*. 0, 1-12.

60. Dr. Malay, Patra., Dr. Timothy, C. Johnstone., Dr.Kogularamanan, Suntharalingam., and Prof. Stephen, J. Lippard. (2016). A potent glucose-platinum conjugate exploits glucose transporters and preferentially accumulates in cancer cells. *Angew Chem Int Ed Engl*. 55(7), 2550-2554.

61. Ellena, J., Gozzo, C., Cominetti, M. R., Ferreira, A. G., Antonio, M., Ferreira, B., Navarro, M., and Batista, A. A. (2015). Chiral Platinum(II) Complexes Featuring Phosphine and Chloroquine Ligands as Cytotoxic and Monofunctional DNA-Binding Agents. *Inorg.Chem.* 54(24), 11709-11720.

62. Timothy, C., Wilson, J. J., Lippard, S. J., and Link, C. (2013). Monofunctional and Higher-Valent Platinum Anticancer Agents. *Inorg Chem*. 52(21), 12234-12249.

63. Milovanovic, M., and Volarevic, V. (2012). Cytotoxic properties of platinum (IV) and dinuclear platinum (II) complexes and their ligand substitution reactions with guanosine.

Transition Metal Chem.37(5), 481–488.

64. Wheate, N. J., Walker, S., Craig, G. E., and Oun, R. (2012). The status of platinum anticancer drugs in the clinic and in clinical trials. *Dalton Trans.* 39, 8113–8127.

65. Antonarakis, E. S., and Emadi, A. (2010). Ruthenium-based chemotherapeutics: Are they ready for prime time? Cancer Chemother. *Pharmacol.* 66, 1–9.

66. Szymański, P., Fraczek, T., Markowicz, M., and Mikiciuk-Olasik, E. (2012). Development of copper based drugs, radiopharmaceuticals and medical materials. *BioMetals. 25*, 1089–1112.

67. Page, S. (2012). Ruthenium compounds as anticancer agents. *EDUCATION IN CHEMISTRY*. 26–29.

68. Allardyce, C. S., and Dyson, P. J. (2016). Metal-based drugs that break the rules. *Dalton trans*. 3201–3209.

69. Webb, M. I., and Walsby, C. J. (2015). Albumin binding and ligand-exchange processes of the Ru(III) anticancer agent NAMI-A and its bis-DMSO analogue determined by ENDOR spectroscopy. *Dalton Trans*. *44*,17482–93.

70. Babak, M. V., Meier, S. M., Huber, K. V. M., Reynisson, J., Legin, A. A., Jakupec, M. A., Roller, A., Stukalov, A., Gridling, M., Bennett, K. L., Colinge, J., Berger, W., Dyson, P. J., Superti-Furga, G., Keppler, B. K., and Hartinger, C. G. (2015). Target profiling of an antimetastatic RAPTA agent by chemical proteomics: relevance to the mode of action. *Chem. Sci.* 6, 2449–2456.

71. Weiss, A., Ding, X., van Beijnum, J. R., Wong, I., Wong, T. J., Berndsen, R. H., Dormond, O., Dallinga, M., Shen, L., Schlingemann, R. O., Pili, R., Ho, C. M., Dyson, P. J., van den Bergh, H., Griffioen, A. W., and Nowak-Sliwinska, P. (2015). Rapid optimization of drug combinations for the optimal angiostatic treatment of cancer. Angiogenesis. 18, 233–244.

72. Bruno Peña, Amanda David, Christiane Pavani, Mauricio S. Baptista, Jean-Philippe Pellois, Claudia Turro, and Kim R. Dunbar. (2014). Cytotoxicity Studies of Cyclometallated Ruthenium(II) Compounds: New Applications for Ruthenium Dyes. *Organometllics*. 33(5): 1100-1103.

73. Li, Z., Hou, Y., Qin, D., Jin, Z., and Hu, M. (2015). Two Half-Sandwiched Ruthenium (II) Compounds Containing 5-Fluorouracil Derivatives : Synthesis and Study of DNA

Intercalation. *PLOS ONE*. 1–13.

74. Liu, K. G., Cai, X. Q., Li, X. C., Qin, D. A., and Hu, M. L. (2012). Arene-ruthenium(II) complexes containing 5-fluorouracil-1-methyl isonicotinate: Synthesis and characterization of their anticancer activity. *Inorganica Chim. Acta.* 388: 78–83.

75. Motswainyana, W. M., and Ajibade, P. A. (2015). Anticancer Activities of Mononuclear Ruthenium (II) Coordination Complexes. *Advances in Chemistry*. 1-21.

76. Olszewski, U., Claffey, J., Hogan, M., Tacke, M., Zeillinger, R., Bednarski, P. J., and Hamilton, G. (2011). Anticancer activity and mode of action of titanocene C. Invest. *New Drugs*. 29, 607–614.

77. Koubkova, L., Vyzula, R., Karban, J., Pinkas, J., Ondrouskova, E., Vojtesek, B., and Hrstka, R. (2015). Evaluation of cytotoxic activity of titanocene difluorides and determination of their mechanism of action in ovarian cancer cells. Invest. *New Drugs*. 33,1123–1132.

78. Massimiliano DAiuto, G. P. (2014). Organo-Metallic Compounds: Novel Molecules in Cancer Therapy. *Biochem. Pharmacol*. 13(13), 1603-1615.

79. Santini, C., Pellei, M., Gandin, V., Porchia, M., Tisato, F., and Marzano, C. (2014). Advances in copper complexes as anticancer agents. *Chem. Rev*.114, 815–862.

80. Khalid, H., Hanif, M., Hashmi, M. A., Mahmood, T., Ayub, K., and Monim-Ul-Mehboob, M. (2013). Copper complexes of bioactive ligands with superoxide dismutase activity. Mini Rev. *Med. Chem*. 13, 1944–1956.

81. Shokohi-pour, Z., Chiniforoshan, H., Momtazi-borojeni, A. A., and Notash, B. (2015). A novel Schiff base derived from the gabapentin drug and copper (II) complex: Synthesis, characterization, interaction with DNA/protein and cytotoxic activity. *J. Photochem. Photobiol. B Biol.* 162, 34–44.

82. Lian, W.-J., Wang, X.-T., Xie, C.-Z., Tian, H., Song, X.-Q., Pan, H.-T., Qiao, X., and Xu, J.-Y. (2016). Mixed-ligand copper Schiff base complexes: the role of the co-ligand in DNA binding, DNA cleavage, protein binding and cytotoxicity. *Dalt. Trans*. 45(22), 9073–9087.

83. Zou, T., Ching, A., Lum, T., Lok, C.-N., Zhang, J.-J., and Che, C.-M. (2015). Chemical biology of anticancer gold(III) and gold(I) complexes. *Chem. Soc. Rev*. 44, 8786 – 8801.

84. da Silva Maia, P. I., Deflon, V. M., and Abram, U. (2014). Gold(III) complexes in medicinal

chemistry. *Future Med. Chem*. 6, 1515–36.

85. Ferna, J., Elie, B. T., Sulzmaier, F. J., Sanau, M., Ramos, J. W., and Contel, M.(2014). Organometallic Titanocene−Gold Compounds as Potential Chemotherapeutics in Renal Cancer. Study of their Protein Kinase Inhibitory Properties.. 33(42), 6669-6681. *Organometallics*

86. Kalinowska-Lis, U., Felczak, A., Checinska, L., Szablowska-Gadomska, I., Patyna, E., Malecki, M., Lisowska, K., and Ochocki, J. (2015). Antibacterial Activity and Cytotoxicity of Silver(I) Complexes of Pyridine and (Benz)Imidazole Derivatives. X-ray Crystal Structure of [Ag(2,6-di(CH(2)OH)py)(2)]NO(3). *Molecules* . 21(2): 87.

87. Siciliano, T. J., Deblock, M. C., Hindi, K. M., Durmus, S., Panzner, M. J., Tessier, C. A., and Youngs, W. J. (2011). Synthesis and anticancer properties of gold(I) and silver(I) N-heterocyclic carbene complexes. *J. Organomet. Chem*. 696, 1066–1071.

88. Ali, K. A., Abd-Elzaher, M. M., and Mahmoud, K. (2013). Synthesis and Anticancer Properties of Silver(I) Complexes Containing 2,6-Bis(substituted)pyridine Derivatives. *Int. J. Med. Chem.* 2013: 256836.

89. Farrell, N. (1989). Transition Metal Complexes as Drugs and Chemotherapeutic Agents. Met. Complexes as Drugs Chemother. *Agents.* 11: 809–840.

90. Uversky, V. N., Kretsinger, R. H., and Permyakov, Ei. E. a. (2013). Encyclopedia of Metalloproteins. *Springer*. 1, 1-89.

91. Iakovidis, I., Delimaris, I., and Piperakis, S. M. (2011). Copper and its complexes in medicine: a biochemical approach. *Mol. Biol. Int*. 594529.

92. Díaz, M. R., and Vivas-Mejia, P. E. (2013). Nanoparticles as drug delivery systems in cancer medicine: Emphasis on RNAi-containing nanoliposomes. *Pharmaceuticals.* 6,1361–1380.

93. Ventola, C. L. (2012). The nanomedicine revolution: part 2: current and future clinical applications. *P T* . 37, 582–91.

94. Bergamo, A., and Sava, G. (2015). Linking the future of anticancer metal-complexes to the therapy of tumour metastases. *Chem. Soc. Rev*. 44, 8818–8835.

95. Valentino, Laquintana., Adriana, Trapani., Nunzio, Denora., Fan Wang., James M. Gallo.,

and Giuseppe Trapani. (2009). New strategies to deliver anticancer drugs to brain tumors. *Expert Opin Drug Deliv*. 6(10), 1017-1032.

96. Coller, H. A. (2014). Is cancer a metabolic disease? 9582978472. *Am. J. Pathol.* 184, 4–17.

97. Capper, C. P., Rae, J. M., and Auchus, R. J. (2016). The Metabolism, Analysis, and Targeting of Steroid Hormones in Breast and Prostate Cancer. *Horm. Cancer.* 1–16.

98. Joo, H. K., Lee, Y. R., Kang, G., Choi, S., Kim, C., and Ryoo, S. (2015). The 18-kDa Translocator Protein Inhibits Vascular Cell Adhesion Molecule-1 Expression via Inhibition of Mitochondrial Reactive Oxygen Species. *Mol Cells*. *38*, 1064–1070.

99. Crider, K. S., Yang, T. P., Berry, R. J., and Bailey, L. B. (2012). Folate and DNA Methylation : A Review of Molecular Mechanisms and the Evidence for Folate ' s Role. *Am. Soc. Nutr*. 3, 21–38.

100. Zwicke, G. L., Mansoori, G. A., and Jeffery, C. J. (2012). Targeting of Cancer Nanotherapeutics. *Nano Rev*. 1,1–11.

101. Xiao, Y. F., Jie, M. M., Li, B. S., Hu, C. J., Xie, R., Tang, B., and Yang, S. M. (2015). Peptide-Based Treatment: A Promising Cancer Therapy. *J. Immunol. Res.* 761820.

CHAPTER 8

General conclusions and future study recommendations

1 General conclusions

The major aims of this study were to investigate conformational features and ligand binding landscape of orphan nuclear receptor (ROR-γ) complex with the experimental drug (XY018), non-receptor tyrosine kinase (c-Src) complex with the experimental drug (UM-164) and to provide molecular understanding on how the c-Src mutation confers resistance to UM-164. Results from this work confirmed the following conclusions:

1. Hydrophobic packing contributes significantly to binding free energy owing to a large amount of aromatic and hydrophobic rings within the active site residues of ROR-γ.

2. The energy decomposition analysis revealed that electrostatic interactions are the potentially important binding forces between XY018 and ROR-γ, while van der Waals contributions are more prominent in HC9-ROR-γ system

3. Ile136 and Leu60 exhibited high hydrogen-bond occupancy in XY018-ROR-γ and HC9-ROR-γ respectively, therefore plays an important role in stabilising the protein.

4. Analysis of principal components revealed that the binding of XY018 to ROR-γ may be responsible for structural rigidity and decrease in motion observed in the XY018-ROR-γ system. Therefore, XY018 exhibits very high binding affinity to ROR-γ.

5. Ligand-residue interaction revealed that interactive OH group of XY018 that formed a hydrogen bond with Glu115 is essential for the binding of XY018 to ROR-γ, while phenyl group of XY018 interacts with glutamine (Gln22) by accepting a hydrogen through a hydrogen-bond formation.

6. Findings from estimated toxicity and biological testing suggest that XY018 is likely to induce pseudoporphyria, nephritis, and interstitial nephritis but potentially safe in renal failure. However, XY018 is potentially indicated in the treatment of atherosclerosis, one of the most important findings from this study.

7. The binding of the lead drug (UM-164) to Src induces a more stable and compact conformation of the protein structure, compared to Dasatinib.

8. The UM-164 binding induces a more correlated motion in Src relatively to Dasatinib suggesting that ligand binding may have induced residue dynamics that results in conformational changes in the protein.

9. High fluctuation exhibited by the loops in Src-UM-164 complex support the experimental evidence that UM-164 binds the DFG-out inactive conformation of Src.

10. The estimated binding free energy is higher in Src-UM-164 compared to Src-Dasatinib, this reflects the relative higher binding capacity of UM-164 to the Src.

11. The orientation of UM-164 in the active site of Src allows for hydrophobic interaction with a fluorinated phenyl group, thus may contribute to high potency reported experimentally.

12. The interactive OH group of UM-164 forms a hydrogen bond with Met94 and Lys96, implying that these residues are essential for the binding of the UM-164 to the protein.

13. UM-164 is potentially safer than Dasatinib regarding the toxicity, thus may be superior to Dasatinib in both its clinical efficacy and safety.

14. Thr91 mutation in Src decrease the capacity of the loops to fluctuate thereby causing the loop rigidity.

15. Ligand-residue interaction revealed that thr91 mutation distorts UM-164 optimum orientation in the conformational space of Src, thus affecting the binding of UM-164.

16. The mutation also decreases the binding energy of UM-164 by -13.416 kcal/mol. Therefore, the mutant Src may be developing resistance to UM-164.

17. There is a distortion of residue interaction network and an overall decrease in hydrogen bond formation between the residues resulting from thr91 mutation.

2 Future study and recommendations

The lead compounds which were used in this study demonstrated a good safety profile with regards to their toxicity. However, the clinical trial of these compounds is required for the rationalisation of their use in the treatment of cancer. A strategic computational technique presented in this work could serve as an important tool to enhance novel drug discovery and development process. It could provide important insights that will assist in the further design of novel inhibitors to minimise the chance of drug resistance in cancer. It could also provide an invaluable contributions to the understanding dynamics of orphan nuclear (ROR-γ) receptor

and non-receptor tyrosine kinase (c-Src) which would largely contribute to the design of potent inhibitor targeting ROR-γ and c-Src respectively.

Appendix

Supplementary materials chapter 5

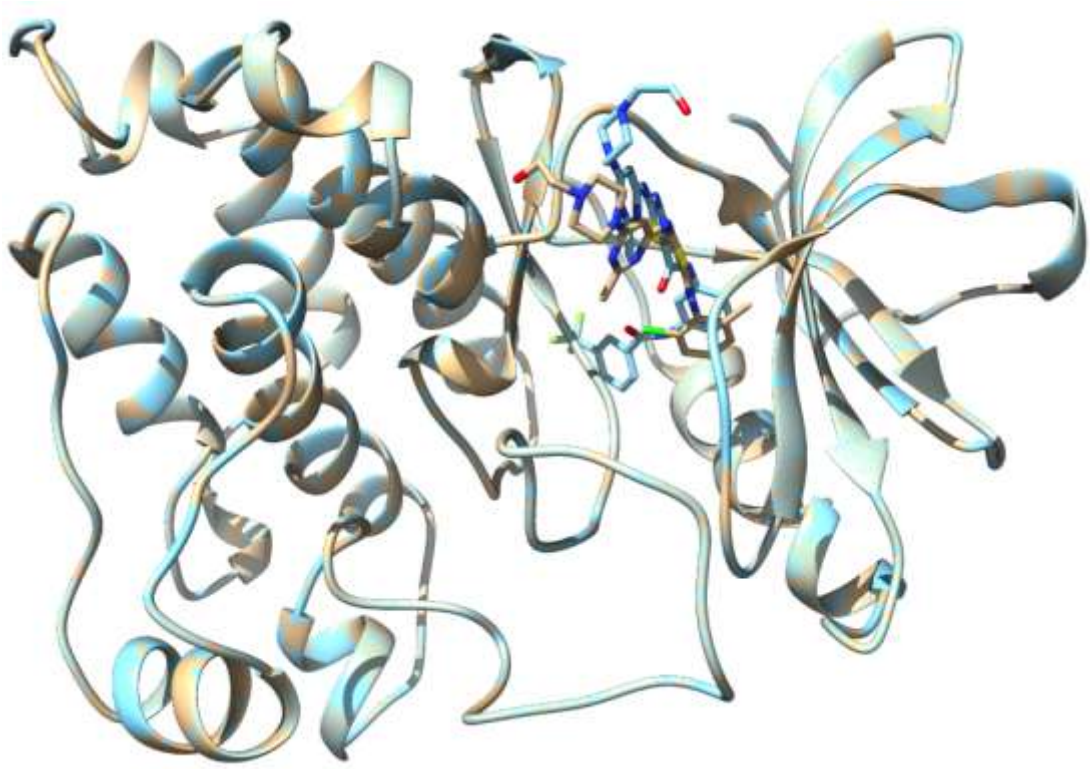


Figure S1. 3D structure of docked Dasatinib superimposed on UM-164-Src complex.

Table S1. Docking scores of Dasatinib and UM-164

Compounds	Docking scores
UM-164	-11.6
Dasatinib	-9.3

Letter of manuscript acceptance for chapter 4

Acknowledgement letter of manuscript submission for chapter 5

Acknowledgement letter of manuscript submission for chapter 6

Letter of manuscript acceptance for chapter 7

Plagiarism report