



UNIVERSITY OF <sup>TM</sup>  
KWAZULU-NATAL

---

INYUVESI  
YAKWAZULU-NATALI

**Combinatorial *in-silico* Modeling and Bioinformatics Analysis of Immune Proteins and  
Small-Molecular Weight Inhibitors: A Potential for Cancer Chemotherapy**

**Opeyemi Seun Soremekun**

**2020**

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

**Combinatorial *in-silico* Modeling and Bioinformatics Analysis of Immune Proteins and  
Small-Molecular Weight Inhibitors: A Potential for Cancer Chemotherapy**

**Opeyemi Seun Soremekun**

**218068269**

**2020**

A thesis submitted to the School of Health Sciences, University of KwaZulu-Natal, Westville, for the degree of Doctor of Philosophy.

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

This is to certify that the content of this thesis are the original research work of Mr Opeyemi Seun Soremekun.

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

Supervisor:

Signed: Mahmoud El-Sayeed Soliman

Name: Prof. Mahmoud El-Sayeed Soliman

## PREFACE

This thesis is divided into eight chapters:

### **Chapter 1:**

This is an introductory chapter that addresses the background, rationale and relevance of the study as well as the proposed aim and objectives. The general outline and structure of the thesis concludes this chapter.

### **Chapter 2:**

This chapter gives a comprehensive background on the immune system, the component and cells of the immune system. This chapter also describes immune system disorders and the regulation of the immune system. A brief description of cancer immunotherapy is also discussed in this chapter. Furthermore, this chapter highlights three different immune proteins that can and have been targeted in cancer immunotherapy by both humanized antibodies and small molecule inhibitors.

### **Chapter 3:**

This chapter discusses *in-silico* techniques that have aided and facilitated the drug design and development pipeline. Molecular modeling, molecular dynamic simulation, virtual screening, identification of non-synonymous Single Nucleotide Polymorphism (nsSNPs) are some of the techniques discussed in this chapter.

### **Chapter 4: (Published work- this chapter is presented in the required format of the journal and is the final version of the accepted manuscript):**

This chapter presents result from the study entitled “ Recruiting monomer for dimer formation: Resolving the antagonistic mechanisms of novel immune check point inhibitors against

Programmed Death Ligand-1 in cancer immunotherapy”. This research has been published in *Molecular Simulation*.

**Chapter 5: (Published work- this chapter is presented in the required format of the journal and is the final version of the accepted manuscript):**

This chapter presents result from the study entitled “Drug Promiscuity: Exploring the polypharmacology potential of 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones as an Inhibitor of the ‘god father’ of Immune Checkpoint”. This research was published *Computational Biology and Chemistry* in 2019.

**Chapter 6: (Published work- this chapter is presented in the required format of the journal and is the final version of the accepted manuscript):**

This chapter presents result from the study entitled “From Genomic Variation to Protein Aberration: Mutational Analysis of Single Nucleotide Polymorphism present in *ULBP6* gene and Implication in Immune Response”. This research was published in *Computers in Biology and Medicine* in 2019.

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

This chapter is a review on application of bioinformatics in cancer immunotherapy, it is entitled “Integrating Bioinformatics strategies in cancer immunotherapy: Current and future perspectives”. This research was published in *Combinatorial Chemistry & High Throughput Screening* in 2020.

**Chapter 8:**

This is the final chapter that proposes future work and concluding remarks.

## ABSTRACT

The immune system carries out pivotal functions in the protection of the body from damaging substances, microbes, and cellular alteration that could affect the health of an individual. The component of the immune system comprises of proteins, cells, and diverse organs. Individuals remain in a good state of health and wholeness if the immune system is working optimally, but if it becomes incapacitated toward fighting off germs or other harmful foreign substances, a diseased state set in. The innate and adaptive systems are the two sub-categories of the immune system. They work synergistically in the defence of the body and fighting off germs that triggered an immune response.

Several proteins have been discovered to play pivotal roles in immune evasion and have therefore become attractive targets. Three of these proteins form the core of this thesis.

Programmed death-ligand 1 (PD-L1), is an immune checkpoint protein which upon binding with another inhibitory checkpoint protein programmed cell death protein 1 (PD-1), elicit a cascade of reaction that leads to the reduction of proliferating antigen-specific T cells. The upregulation of PD-L1 can therefore, lead to evasion of the immune system by cancer cells.

Cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) is made up of three parts (a transmembrane part, extracellular part, and a cytoplasmic part). CTLA-4 has been implicated in the downregulation of immune response and blocking CTLA-4 activity results in a surge in immune functions. It has also been found that CTLA-4 negatively controls the T-cells.

Natural killer group 2, member D (NKG2D) is located on the surface of immune cells where it acts as an activating receptor and regulator of the adaptive and innate immune system upon binding to its constitutive ligands such as UL16-binding protein (ULBP6). ULBP6 is a highly

polymorphic protein, hence, the interaction between NKGD2 and ULBP6 is often altered. This effect has a great impact on the function of NKGD2 as a regulator of the immune system.

Various humanized antibodies have been developed to specifically target PD-L1 and CTLA4. Some have been approved while others are at different phases of clinical trial. Ipilimumab and tremelimumab are designed as CTLA-4 inhibitors. Durvalumab and atezolimumab are designed as anti-PD-L1 inhibitor. However, due to the limitations that have characterized the use of humanized antibodies inhibitors such as production cost, instability, and low tumour penetration, etc, small molecule inhibitors have been considered as a better alternative to humanized antibodies inhibitors.

This thesis explored the mechanism of inhibition of newly synthesised PD-L1 inhibitors (BMS-1166 and BMS-1001). Also, per-residue based virtual screening was employed to predict potential CTLA-4 inhibitors. Computational methods such as molecular docking, molecular dynamic simulation, virtual screening, and SNPinformatics were employed. These computational techniques revealed that BMS-1166 and BMS-1001 caused a motional movement in the monomers of PD-L1 to form a dimer, thereby preventing PD-L1-PD-1 interaction. Although the PD-L1 monomers have the same residues, their affinity for the BMS compounds differ. Two compounds ZINC04515726 and ZINC08985213 were identified as possible targets of CTLA-4. These two compounds elicited favourable interaction with CTLA-4 facilitated by some crucial residues. Furthermore, the non-synonymous Single Nucleotide Polymorphism (nsSNPs) associated with ULBP6 were identified, and the effect of these nsSNPs on the interaction between NKGD2 and ULBP6 was also investigated.

The first study (Chapter 4) investigates the structural dynamics and also provides insights into the mechanism of inhibition of BMS-1166 and BMS-1001 on PD-L1.

The second study (Chapter 5) determines the binding site landscape of CTLA-4 and also employs binding site similarities between unrelated proteins to repurpose an inhibitor to target CTLA-4.

The third study (Chapter 6) identifies deleterious polymorphisms associated with ULBP6. The effect of these polymorphisms on NKGD2-ULBP6 binding as a consequent on immune response is also explored in this chapter.

Chapter 7 gives a detailed report on how the use of bioinformatics tools and strategies have aided and advanced the field of cancer immunotherapy.

This study provides a thorough insight into the *in-silico* design, development and mechanism of action of small molecule inhibitors of PD-L1 and CTLA-4. Furthermore, this study gives insight into the polymorphic nature of ULBP6. Thence, the work presented in this study would serve as a platform towards the design of small molecule inhibitors of CTLA-4 and PD-L1 with high therapeutic and less toxicity.

## DECLARATRION 1 -PLAGIARISM

I, Opeyemi Seun Soremekun, declare that

1. The research reported in this thesis, except where otherwise stated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other person's data, pictures, graphs or other information. Unless specifically acknowledged as being source from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
  - i. Their words have been-rewritten but the general information attributed to them has been referenced.
  - ii. Where the exact words have been used, then their writing has been placed in italics and inside quotation marks and referenced.
5. This thesis does not contain text, graphics, or tables copied and pasted from the internet, unless specifically acknowledged, and the source being detailed in the thesis and in the references sections.

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

Signed: O S. Soremekun



## DECLARATRION 2: LIST OF PUBLICATIONS

1. **Soremekun, O.**, Olotu, F., Agoni, C., Soliman, M. (2019). Recruiting monomer for dimer formation: Resolving the antagonistic mechanisms of novel immune check point inhibitors against Programmed Death Ligand-1 in cancer immunotherapy. *Molecular Simulation*, DOI:10.1080/08927022.2019.1593977. (Published)

### **Contribution:**

Soremekun, O: contributed to the project by performing all the experimental work.

Olotu, F: Assisted in manuscript writing.

Agoni, C: Help in the manuscript preparation

Soliman, M: Supervisor

*Appendix A: pdf version of the publication*

2. **Soremekun, O.**, Olotu, F., Agoni, C., Soliman, M. (2019). Drug Promiscuity: Exploring the polypharmacology potential of 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones as an Inhibitor of the ‘god father’ of Immune Checkpoint. *Computational Biology and Chemistry*, 80:433-440. (Published).

### **Contribution:**

Soremekun, O: contributed to the project by performing all the experimental work.

Olotu, F: Assisted in manuscript writing.

Agoni, C: Help in the manuscript preparation

Soliman, M: Supervisor

*Appendix B: pdf version of the publication*

3. **Soremekun, O.**, Soliman, M. (2019). From Genomic Variation to Protein Aberration: Mutational Analysis of Single Nucleotide Polymorphism present in *ULBP6* gene and Implication in Immune Response. *Computers in Biology and Medicine*, 111:1033354. (Published).

**Contribution:**

Soremekun, O: contributed to the project by performing all the experimental work, manuscript writing and preparation

Soliman, M: Supervisor

*Appendix C: pdf version of the publication*

4. Houda N. Washah, Elliasu Y. Salifu, **Opeyemi Soremekun**, Ahmed A.Elrashedy, Geraldene Munsamy, Fisayo A. Olotu and Mahmoud E.S. Soliman. (2020). Integrating Bioinformatics Strategies in Cancer Immunotherapy: Current And Future Perspectives. *Combinatorial Chemistry & High Throughput Screening*, 2020, 23, 0-0. (Published).

**Contribution:**

Soremekun, O: contributed to the project by manuscript preparation and writing part of the manuscript.

Houda N: Contributed by writing part of the manuscript.

Salifu Y: contributed to the project by manuscript preparation and writing part of the manuscript.

Elrashedy A: Assisted in data collection

Munsamy G: Assisted in data collection

Olotu, F: Assisted in manuscript writing

Soliman, M: Supervisor

*Appendix D: pdf version of the publication*

## RESEARCH OUTPUT

### A- LIST OF PUBLICATIONS

1. **Soremekun, O.**, Olotu, F., Agoni, C., Soliman, M. (2019). Recruiting monomer for dimer formation: Resolving the antagonistic mechanisms of novel immune check point inhibitors against Programmed Death Ligand-1 in cancer immunotherapy. *Molecular Simulation*, DOI:10.1080/08927022.2019.1593977
2. **Soremekun, O.**, Olotu, F., Agoni, C., Soliman, M. (2019). Drug Promiscuity: Exploring the polypharmacology potential of 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones as an Inhibitor of the ‘god father’ of Immune Checkpoint. *Computational Biology and Chemistry*, 80:433-440.
3. **Soremekun, O.**, Soliman, M. (2019). From Genomic Variation to Protein Aberration: Mutational Analysis of Single Nucleotide Polymorphism present in *ULBP6* gene and Implication in Immune Response. *Computers in Biology and Medicine*, 111:1033354.
4. Houda N. Washah, Elliasu Y. Salifu, **Opeyemi Soremekun**, Ahmed A.Elrashedy, Geraldene Munsamy, Fisayo A. Olotu and Mahmoud E.S. Soliman. (2020). Integrating Bioinformatics Strategies in Cancer Immunotherapy: Current And Future Perspectives. *Combinatorial Chemistry & High Throughput Screening*, 2020, 23, 0-0.
5. **Soremekun, O.**, F. Omolabi., Soliman, M. (2020). Identification and Classification of Differentially Expressed Genes Reveals Potential Molecular Signatures Associated with SARS-CoV-2 Infection in Lung Adenocarcinoma Cells. *Informatics in Medicine Unlocked* DOI: 10.1016/j.imu.2020.100384
6. Olotu, F., **Soremekun, O.**, Agoni, C., Soliman, M. (2020) The recent application of 3D-QSAR and docking studies to novel HIV-protease inhibitor drug discovery. *Expert Opinion on Drug Discovery*. 10.1080/17460441.2020.1773428
7. Adewunmi, T., **Soremekun, O.**, Ajadi M., Soliman, M. (2020) Thompson Loop: Opportunities for Antitubercular drug design by Targeting the Weak Spot in Demethylmenaquinone Methyltransferase. *RCS Advances*. DOI: 10.1039/D0RA03206A
8. Adewunmi, T., Ramanchandra, P, **Soremekun, O.**, Soliman, M. (2020) Delving into the Characteristic Features of “Menace” Mycobacterium tuberculosis Homologs: A Structural dynamics and Proteomics Perspectives. *The Protein Journal*. DOI:10.1007/s10930-020-09890-4
9. Akinsiku O., **Soremekun, O.**, Olotu, F., Soliman, M. (2020). Tapping on the Crucial Role of Asp1116 in Selective Drug Targeting of CREB-cAMP- responsive element-

binding Implicated in Prostate Cancer. *Combinatorial Chemistry & High Throughput Screening*. DOI:10.2174/1386207323666200219122057

10. Kubi, P., Oduro, **Soremekun, O** Kariwani, R., Agoni, C., Olotu, F., Soliman M. (2020) Exploring the Ring Potential of 2,4-diaminopyrimidine derivatives towards the Identification of Novel Caspase-1 Inhibitors in Alzheimer's Disease Therapy. *Journal of Molecular Modelling*. DOI: 10.1007/s00894-020-4319-6
11. Olotu, F., **Soremekun, O.**, Agoni, C., Soliman, M. (2020) An update on the pharmacological usage of curcumin: has it failed in the drug discovery pipeline? *Cell Biochemistry and Biophysics*. 10.1007/s12013-020-00922-5
12. **Soremekun, O.**, Agoni, C., Olotu, F. (2019). Learning the Footprints and Fingerprints: Pharmacophore Modelling in the Discovery of Potential Drug Candidates. *Journal of Organic and Biomolecular Simulations*, 1(1), 1-9
13. Kariwani, R., Agoni, C., **Soremekun, O.**, Kubi, P., Oduro, R., Olotu, F., Soliman M. (2019) Same Target, Different Therapeutic Outcomes: The Case of CAY10471 and Fevipiprant on CRTh2 Receptor in Treatment of Allergic Rhinitis and Asthma. *Combinatorial Chemistry & High Throughput Screening*. doi: 10.2174/1386207322666190919113006.
14. Oduro, R., Kariwani, R., **Soremekun, O.**, Agoni, C., Olotu, F., Soliman, M. (2019). From the Explored to the Unexplored; How Computer Aided Drug Design techniques have and can help in Caspase research. *Combinatorial Chemistry & High Throughput Screening*. doi:10.2174/1386207322666190927143026.
15. Ajadi, B., **Soremekun, O.**, Adeniyi, T., Kumalo, H., Soliman, M. (2020). Leveraging on Active Site Similarities; Identification of Potential Inhibitors of Zinc-Finger and UFSP domain Protein (ZUFSP). *Current Pharmaceutical Biotechnology*. 10.2174/1389201021666200730151218

## **B-List of Publications Under Review**

1. Ajadi, B., Elrashedy A, **Soremekun, O.**, Olotu, F., Agoni, C., Soliman. Probing protein-protein interactions and druggable site identification: Mechanistic binding events between Ubiquitin and Zinc finger with UFM1-specific peptidase domain protein (ZUFSP). *Journal of Molecular Modelling*
2. **Opeyemi S. Soremekun**, Chisom Ezenwa, Omotuyi Idowu, Oyekanmi Nashiru, Mahmoud Soliman, and Segun Fatumo. Computational Analysis of Functional Single Nucleotide Polymorphisms Associated with Haemoglobin Subunit Beta (*HBB*) gene. *Computers in Biology and Medicine*.

3. **Soremekun, O.**, F. Omolabi., Adewunmi, T, Soliman, M. Exploring the Effect of Ritonavir and TMC-310911 on SARS-CoV-2 and SARS-CoV Main Proteases – A Potential From Molecular Perspective. *Future Science OA*
4. Subair, T, **Soremekun, O.**, F. Olotu., Soliman, M. Therapeutic Path to Double Knockout: Investigating the Selective Dual-inhibitory Mechanisms of Adenosine Receptors A1 and A2 by a novel methoxy-substituted benzofuran derivative in the Treatment of Parkinson's Disease. *Cell Biochemistry and Biophysics*
5. Okunlola, F., **Soremekun, O.**, F. Olotu., Soliman, M. East to West not Northwest: Structure-based Mechanistic Resolution of 8-hydroxyl replacement and resulting effects on the activities of Imidazole-Based Heme Oxygenase-1 Inhibitors. *Anticancer agents in Medicinal Chemistry*.
6. Akawa, O., **Soremekun, O.**, F. Olotu., Soliman, M. Piercing the fragments together: Dynamical insights into the enhancement of BRD4-BD1 (BET protein) druggability in cancer chemotherapy using a novel 8-Methyl-Pyrrolo[1,2-a]pyrazin-1(2H)-one derivatives. *The Protein Journal*.
7. Adewunmi, T., Elrashedy, A., **Soremekun O.**, Ajadi, M. Soliman, M. Irreversible inhibition dynamics of Mycobacterium tuberculosis antigen 85C SER 124- $\beta$ -isomer-enolphosphorus Cycliphostin: implications for a potential novel TB drug. *Journal of Biomolecular Structure and Dynamics*.

## C-Conferences

- 24<sup>TH</sup> International Conference on Research in Computational Molecular Biology. Virtual Conference. 22<sup>nd</sup> -25<sup>th</sup> June 2020
- 2<sup>nd</sup> Symposium of ACE-B on “Global Health Bioinformatics” Organised by African Centres of Excellence in Bioinformatics, Bamako, Mali. 16<sup>TH</sup> -17<sup>TH</sup> March 2020.
- Expanding the Druggable Proteome with Chemical Biology Organised by EMBL Heidelberg, Germany. 5<sup>th</sup> - 7<sup>th</sup> February 2020.
- Centre for High Performance Computing (CHPC) Conference 2019. Birchood Hotels, Johannesburg, South Africa. 1<sup>st</sup>-5<sup>th</sup> December 2019.
- Introduction to Biostatistics and R organized by SANTHEAfrica. Durban, South Africa. 12<sup>th</sup> -15<sup>th</sup> November 2019.
- H3ABioNet Introduction to Bioinformatics (IBT) Course. UKZN class. Durban, South Africa. August-November 2019.
- Antibiotic Stewardship and Conservation in Africa, Durban, South Africa. 20<sup>th</sup> -23<sup>rd</sup> October 2019.
- Christian-Albrechts-Universitat zu Kiel. Summer School on “Degradomics”. Kiel, Germany. 16<sup>th</sup> - 18<sup>th</sup> September 2019.
- Centre for High Performance Computing (CHPC) Winter School Course. Pretoria, South Africa. 1<sup>st</sup> July-6<sup>th</sup> July 2019

## **ACKNOWLEDGEMENT**

God Almighty the giver of all knowledge

My Supervisor Prof. Mahmoud Soliman for his constant support

My Mum, Siblings, and Fiancée for their support

My colleagues for their support

University of KwaZulu-Natal College of Health Sciences for financial support

Centre for High Performance Computer for the computational resources.

## LIST OF AMINO ACIDS

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

## LIST OF ABBREVIATIONS

CTLA-4	Cytotoxic T-lymphocyte-associated protein 4
nsSNPs	Non-synonymous Single Nucleotide Polymorphism
ULBP6	UL16-binding protein
TNF	Tumour Necrosis Factor
IL	Interleukin
NK	Natural Killer Cells
DC	Dendritic Cells
MC	Mast Cells
CD	Cluster of Differentiation
TCR	T Cell Receptor
BCR	B Cell Receptor
Ig	Immunoglobulin
CI	Cancer Immunotherapy
MHC	Major Histocompatibility Complex
PD-L1	Programmed death-ligand 1
PD-1	Programmed cell death protein 1
ICI	Immune Checkpoint Inhibitors
NKG2D	NH group 2, member D
MDS	Molecular Dynamic Simulation
VS	Virtual Screening
ns	Nanoseconds
PDB	Protein Data Bank
QM	Quantum Mechanics



## LIST OF FIGURES

Figure 2.1 A schematic depiction of the immune system .....	9
Figure 2.2 The cancer immune cycle highlighting molecules involved in each step [30]......	13
Figure 2.3 3D structure of human apo-CTLA-4 monomer showing the A'GFCC'C'' sheet and BED making up the extracellular domain (PDB ID: 3OSK ). .....	14
Figure 2.4: Symmetric apo-CTLA-4 homodimer (B), showing the polyproline motif (C) which is an attractive target site for ligands and the stabilizing loop, showing an interaction between the Cys122 residues of both monomers (A). .....	15
Figure 2.5: 3-dimensional structure of PD-L1 protein with monomer A highlighted in aquamarine, monomer B in cyan, and a small compound bound to the binding interface between monomer A and monomer B. PDB: 5NIX [19] .....	17
Figure 2.6: 3-dimensional structure of ligand ULBP6 in complex with two chains of NKGD2D (PDB ID: 4S0U) [78]. .....	22
Figure 3.1: Graphical description of two-dimensional potential energy surface [7]. .....	35
Figure 3.2: Molecular dynamic simulation basic algorithms and motion calculation [16].....	38
Figure 4.1: 2D structure of soluble immune checkpoint inhibitors; [A] BMS-1001 and [B] BMS-1166. ....	52
Figure 4.2: 3D structure of PD-L1 monomer A (Blue) and monomer B (Gray), in complex with BMS-1001 (Red) and BMS-1166 (Green) at the tunnel-like binding cleft formed upon dimerization. PDB: 5NIX [19]. BMS-1001 was superimposed on the crystal structure. ....	53
Figure 4.3: Structural dynamics of BMS-1001/BMS1166 bound PD-L1s and unbound PD-L1 over the 150ns MD simulation period. Red boxes depict the monomeric motions of BMS-1001-bound PD-L1 while green boxes depict the dynamics of the BMS-1166-bound PD-L1 system. However, blue boxes depict the unbound PD-L1 system. [A] Starting structure averaged over 0-20ns [B] averaged intermediate structure 60-90ns and [C] averaged end structure 90-150ns.....	59
Figure 4.4: Monomeric distances, tunnel formation and mechanistic transition of BMS-1001 and BMS-1166 across both A and B PD-L1 monomers. [A] Tunnel formation and monomeric distance of BMS-1001 bound PD-L1 [B] Tunnel formation and monomeric distance of BMS-1166 bound PD-L1. ....	60
Figure 4.5: Visual representation of the monomeric active site residues interaction of BMS-1001 and BMS-1166 at 30-55ns (A, C) and 60-90ns (B, D). .....	62
Figure 4.6: Conformational analysis plot showing stability and atomistic motions among unbound (blue), BMS-1001 (red) and BMS-1166 (green) bound systems. [A] Whole-structure comparative C- $\alpha$ RMSD plot of the three systems. Blue highlight depicts post-equilibration periods (100-150ns) where distinct separation in motions occurred [B] Comparative C- $\alpha$ RMSD plot (post-equilibration $\rightarrow$ 100-150ns) showing distinct variations in structural stability among the three systems [C] C- $\alpha$ RMSD plot showing active site stability among the unbound (blue), BMS-1001- (red) and BMS-1166- (green) bound systems .....	64
Figure 4.7: Structural stability, positioning and motion of BMS-1001 and BMS-1166 at the tunnel-like binding cleft of PDL-1. [A] Superposition of BMS-1001 (red) and BMS-1166 (green) as they transverse the binding tunnel [B] Comparative C- $\alpha$ RMSD plot depicting the stability and motion of both BMS-1001 (red) and BMS-1166 (green) during the post-equilibration MD-simulation period (100ns-150ns).....	66

Figure 4.8: Structural flexibility induced simultaneously at both PDL-1 monomers (A and B) by the binding of BMS-1001 ( <b>red</b> ) and BMS-1166 ( <b>green</b> ). [A] Superposed monomers A and corresponding C- $\alpha$ RMSF plot showing the degree of flexibility among the unbound ( <b>blue</b> ), BMS-1001- ( <b>red</b> ) and BMS-1166- ( <b>green</b> ) bound PDL-1 systems [B] Superposed monomers B of unbound ( <b>blue</b> ), BMS-1001- ( <b>red</b> ) and BMS-1166- ( <b>green</b> ) bound PDL-1 systems coupled with corresponding C- $\alpha$ RMSF plot showing the degree of flexibility. ....	68
Figure 5.1: 3D crystal structure of CTLA-4. The CC'-Loop is highlighted in green, G-strand in red and the FG-loop in orange. (PDB:5TRU) <sup>12</sup> .....	87
Figure 5.2: Graphical representation of kallikrein and CTLA-4 binding site similarity between Kallikrein 7 and CTLA-4. Similar binding site residues are highlighted in red .....	93
Figure 5.3: 2D structures of ZINC04515726 and ZINC08985213 .....	94
Figure 5.4: Individual energy contributions of crucial binding site residues to high-affinity binding and stabilization of TDSO. [A] Per-residue decomposition plot showing energy contributions of interacting residues [B] Molecular interactions between essential residues and reactive substituent on TDSO. ....	99
Figure 5.5: Individual energy contributions of crucial binding site residues to high-affinity binding and stabilization of ZINC04515726. [A] Per-residue decomposition plot showing energy contributions of interacting residues [B] Molecular interactions between essential residues and reactive substituent on ZINC04515726.....	99
Figure 5.6: Individual energy contributions of crucial binding site residues to high-affinity binding and stabilization of ZINC04515726. [A] Per-residue decomposition plot showing energy contributions of interacting residues [B] Molecular interactions between essential residues and reactive substituent on ZINC04515726.....	100
Figure 5.7: Conformational analysis plot showing stability and atomistic motions among unbound ( <b>black</b> ), TDSO ( <b>red</b> ), ZINC04515726 ( <b>green</b> ) and ZINC08985213 ( <b>blue</b> ) bound systems. [A] Whole-structure comparative C- $\alpha$ RMSD plot of the four systems.....	101
Figure 5.8: [A] C- $\alpha$ RMSD plot showing active site stability among unbound ( <b>black</b> ), TDSO ( <b>red</b> ), ZINC04515726 ( <b>green</b> ) and ZINC08985213 ( <b>blue</b> ) bound systems. [B] Comparative C- $\alpha$ RMSD plot depicting the stability and motion of TDSO ( <b>red</b> ), ZINC04515726 ( <b>green</b> ) and ZINC08985213 ( <b>blue</b> ). [C]Superimposition of TDSO ( <b>red</b> ), ZINC04515726 ( <b>green</b> ) and ZINC08985213 ( <b>blue</b> ) ...	102
Figure 5.9: Structural flexibility induced by the binding of TDSO ( <b>red</b> ), ZINC04515726 ( <b>green</b> ) and ZINC08985213 ( <b>blue</b> ) to unbound ( <b>black</b> ). ....	103
FigS1: 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones showing the interacting pharmacophoric moieties after a 20ns simulation with CTLA-4 .....	108
FigS2: PRED contribution of different residues at the active site (A). 2D ligand interaction plot (B), and (C) Pharmacophore features responsible for FBE contributions .....	109
Figure 6.1: 3-D crystallography structure of ULBP6 protein in complex with NKG2D (PDB ID: 4S0U (14) .....	114
Figure 6.2: A schematic computational workflow .....	116
Figure 6.3: Profile representation of LIMBO stretches in the mutant (V52F) (A), Wild (B), difference in LIMBO chaperone propensity between the mutant and wild protein (C) and molecular visualization of LIMBO chaperone-binding sites colored in pink (D).....	122

Figure 6.4: Profile representation of <b>WALTZ</b> and <b>TANGO</b> stretches present in the mutant ( <b>A</b> ), WALTZ and TANGO stretches present in the wild type ( <b>B</b> ). Difference in WALTZ amyloid propensity and difference in TANGO aggregation between WT and variant ( <b>C</b> ).....	123
Figure 6.5: Functional interaction between ULBP6 and its related genes.....	124
Figure 6.6: Residue interaction network analysis of the wild protein ( <b>green</b> ) and mutant protein ( <b>red</b> ). .....	126
Figure 6.7: DSSP analysis of the mutant and wild protein at 10ns, 50ns and 100ns.....	126
Figure 6.8: Backbone RMSDs are depicted as a function of time for the wild and mutant ULBP6, the mutant is represented with a red colour while the wild as green ( <b>A</b> ). RMSF of the C- $\alpha$ of the mutant (red) and wild (green) ( <b>B</b> ). Radius of gyration of C- $\alpha$ atoms of native and mutant ULBP6 versus time at 300k ( <b>C</b> ). Projection of the motion of the mutant and wild proteins in phase space along the first two principal eigenvectors at 300k ( <b>D</b> ).....	127
Figure 6.9: Schematic diagram depicting change in distance of the interacting domain residue in the course of the simulation run. ....	129
Figure 7.1: Classification of cancer immunotherapy .....	141
Figure 7.2: 3D structures of <b>A</b> ) Cytotoxic antigen – 4 receptor CTLA-4(Light green) in complex with Ipilimumab (brown) PDB code: <b>5TRU</b> <b>B</b> ) Programmed Death 1(PD-1) in complex with Programmed Death Ligand 1 (PD-L1) PDB code: <b>3BIK</b> <b>C</b> ) PD-L1 in complex with Durvalumab inhibitor (PDB code: <b>5X8M</b> ). ....	143
Figure 7.3: Adoptive T cell therapy .....	147
Figure 7.4: Bioinformatics tools for cancer immunology and immunotherapy .....	149

## LIST OF TABLES

Table 2.1 : Activity of CTLA-4 antibody in metastatic melanoma patients .....	16
Table 2.2: Small molecule inhibitors designed to target PD-L1 .....	20
Table 4.1: Per-residue fluctuation values of crucial binding cleft residues in the unbound and bound PD-L1 systems. ....	69
Table 4.2: MM/PBSA binding free energy profiles of BMS-1001 and BMS-1166 to PD-L1 .....	73
Table 5.1: Constituent residues of predicted CTLA-4 binding site .....	92
Table 5.2: Comparative analyses of the pharmacokinetic profile of TDSO, ZINC04515726, and ZINC08985213.....	96
Table 5.3: MM/PBSA binding free energy profiles of TDSO, ZINC04515726, and ZINC08985213 to CTLA-4 .....	97
Table S1: Binding energy scores of TDSO, ZINC04515726, ZINC08985213 with TDSO.....	109
Table 6.1: Damaging and Deleterious SNPs predicted by SIFT, PolyPhen, SNP&GO, Pmut and Phd-SNP with their corresponding scores. ....	120
Table 6.2: LIMBO regions in variant and wild type. For each LIMBO region, the start, end, sequence and score are given.....	122
Table 6.3: <b>TANGO</b> regions in variant and wild type. For each TANGO region, the start, end, sequence and score are given. ....	123
Table 6.4: <b>WALTZ</b> regions in variant and wild type. For each WALTZ region, the start, end, sequence and score are given. ....	123
Table 7.1: Checkpoint inhibitors .....	143
Table 7.2: A summary of some monoclonal antibodies used in immunotherapy of cancer. ....	145

## TABLE OF CONTENT

PREFACE .....	iii
ABSTRACT .....	v
DECLARATRION 1 -PLAGIARISM.....	viii
DECLARATRION 2: LIST OF PUBLICATIONS.....	ix
RESEARCH OUTPUT.....	xi
A- LIST OF PUBLICATIONS .....	xi
B-List of Publications Under Review .....	xii
C-Conferences.....	xiii
ACKNOWLEDGEMENT.....	xiv
LIST OF AMINO ACIDS .....	xv
LIST OF ABBREVIATIONS .....	xvi
LIST OF FIGURES.....	xvii
LIST OF TABLES.....	xx
TABLE OF CONTENT .....	xxi
CHAPTER 1 .....	1
1.0 Introduction .....	1
1.1 Background and Rational .....	1
1.2 Aims and Objectives.....	2
1.3 Novelty and Significance of Research .....	3
CHAPTER 2 .....	5
2. Background on Immune System .....	5
2.1 Introduction .....	5
2.2 The innate immunity.....	6
2.3 Adaptive Immunity .....	7
2.4 Physiological control of the immune System.....	9
2.5 Disorders of human immunity .....	9
2.6 Cancer and the Immune System .....	10
2.7 Cancer and Cancer Immunotherapy (CI).....	11
2.8 Cancer immune cycle .....	12
2.9 Cytotoxic T-lymphocyte-associated antigen .....	13
2.10 Programmed Death Ligand -1 .....	16
2.11 UL16-Binding Protein .....	20
CHAPTER 3 .....	32
3.0 Computational Strategies Used in Drug Design and SNPinformatics approach for nsSNPs Identification .	32
3.1 Introduction .....	32
3.2 Quantum Mechanics (QM) .....	33
3.2.1 Schrödinger’s wave function theory .....	34
3.2.2 The Born-Oppenheimer approximation.....	34
3.2.3 Potential Energy Surface (PES).....	35
3.3 Molecular Mechanics.....	36
3.3.1 Force Field.....	37
3.4 The Concept of Molecular Dynamic Simulation.....	37
3.4.1 Molecular Dynamics Trajectory Analysis .....	39
3.5 Binding Free Energy .....	41
3.6 Molecular Docking .....	42
3.7 Virtual Screening.....	42
3.8 Identification of Deleterious nsSNPs.....	43
CHAPTER 4 .....	49
Abstract.....	50
4.1 Introduction.....	51
4.2 Computational Methods.....	54
4.2.1. System Preparation.....	54
4.2.2. Binding free energy estimation.....	56
4.2.3. Dynamic cross-correlation matrix (DCCM).....	56
4.2.4. Principal component analysis (PCA).....	57
4.3. Results and Discussion.....	58
4.3.1 Monomeric binding of BMS compounds (BMS-1001 and BMS-1166) induced systematic PD-L1 dimerization .....	58

4.3.2	<i>BMS compounds induced high conformational flexibility favorable for monomeric motion and dimerization</i> .....	62
4.3.2.1	Conformational stability and residual fluctuation .....	62
3.2.2.	Principal component analysis of structural dynamics and motion .....	71
4.3.3.	BMS-1001 and BMS-1166 exhibited ‘monomeric interchange’ and systematic high-affinity binding prior to PD-L1 dimerization.....	72
4.4	Conclusion .....	76
CHAPTER 5	.....	84
	Abstract.....	85
5.1	Introduction .....	86
5.2	Computational Method .....	87
5.2.2	Binding site similarity identification.....	87
5.2.3	Pharmacophore Model Creation using Per Residue Energy Decomposition (PRED) Based Approach .....	88
5.2.4	Drug Likelihood Assessment .....	89
5.2.6	System preparation and Molecular dynamic (MD) simulations.....	89
5.2.6	Post-dynamic analysis.....	91
5.2.7	Thermodynamic Calculation .....	91
5.3	Results and Discussion .....	92
5.3.1	Binding Site Prediction and Validation.....	92
5.3.1.2	Pharmacophore Model Creation .....	93
5.3.1.3	Generation of Ligands.....	94
5.3.2	Drug Likelihood Assessment of TDSO, ZINC04515726 and ZINC08985213.....	95
5.3.2.1	ADME Result .....	95
5.3.3	Post-Molecular Dynamics Simulation Analysis .....	96
5.3.3.1	Differential binding of TDSO, ZINC04515726 and ZINC08985213 to CTLA-4.....	96
5.3.3.2	Structural elucidation in the binding of TDSO, ZINC04515726, and ZINC08985213 to CTLA-4 .....	100
5.4	Conclusion.....	103
CHAPTER 6	.....	111
	Abstract.....	112
6.1	Introduction .....	113
6.2	Material and Methods .....	116
6.2.1	SNPs Retrieval .....	116
6.2.3	Impact of mutation on protein-protein interaction and Molecular Dynamics Simulation .....	118
6.2.4	Principal Component Analysis (PCA) .....	118
6.3	Results.....	119
6.3.1	Identification of deleterious and damaging SNPs .....	119
6.3.2	Phenotype Analysis .....	121
	LIMBO prediction.....	121
	Waltz and Tango prediction.....	122
6.3.3	Investigation of ULBP6 Gene’s interactions and Appearance in Networks in STRING Database.....	124
6.3.4	Impact of V52F polymorphism on structural integrity of ULBP6 .....	125
6.3.5	Investigation of Structural conformation in ULBP6 .....	127
6.4	Discussion .....	129
6.5	Conclusion.....	131
CHAPTER 7	.....	137
	Abstract.....	138
7.0	Introduction .....	139
7.1	The immune system’s mechanism of action against cancer.....	140
7.2	Techniques used in immunotherapy of cancer.....	141
7.3	Immune checkpoint inhibitors and Monoclonal Antibodies.....	142
7.3.1	Immune checkpoint Inhibitors. ....	142
7.3.2	Monoclonal Antibodies (mAbs) .....	144
7.4	Cancer Vaccine.....	145
7.5	Adoptive T-cell therapy (ACT) .....	146
7.6	Combinatorial cancer immunotherapy.....	147
7.7	Application of Bioinformatics strategies in Cancer immunotherapy .....	148
7.8	Next Generation Sequencing .....	149
7.9	Integrated data analysis and network modelling technique .....	150
7.10	HLA Typing .....	151
7.11	Epitope Prediction .....	152
7.12	Sample heterogeneity.....	152

7.13 Modeling of tumor–immune cell interactions .....	153
7.14 Future Perspectives .....	153
7.16 Conclusion.....	154
APPENDIX.....	169
Appendix A.....	169
Appendix B.....	170
Appendix C.....	170
Appendix D.....	170

# CHAPTER 1

## 1.0 Introduction

### *1.1 Background and Rational*

The immune system is an advanced system comprising of specialized cells, tissues, and proteins that fights off malignancies, toxins, and pathogens [1]. The immune system is grouped into innate and adaptive immune responses. The main function of the immune system is in defence of the host against foreign assaults [2]. Once the functions of the immune system are compromised, it leads to the onset of diseases. However, on the flip side, hypersensitive diseases could result if the immune system produces excessive and unwanted reactions.

Immune proteins produced by different immune cells perform various immunological functions in the body. Ranging from immunosuppression, activation of other immune proteins, and the pathogenesis of diseases. Up-regulation/downregulation or aberration of some of these proteins could lead to the onset of diseases. Examples of these proteins include major histocompatibility complex, Interleukin 6, T-cell receptor, Fc receptor, Interleukin 4, C1Q complex, Interleukin 18, IL1B, PD-L1, CTLA-4, PD-1, and Perforin, etc. Immune checkpoint proteins like PD-L1, CTLA-4, PD-1, LAG-1, etc. have been implicated in the onset of cancer, hence they have become an attractive target in cancer immunotherapy [3,4].

Although great success has been recorded in the target of some immune proteins using humanized antibodies, the production cost, elimination of immunogenicity, improved tumour penetration, stability, and amenability for oral administration, small-molecular weight compounds has been



looked to as a therapeutic option for the otherwise costly humanized antibodies. However, the exploration of small molecular weight compounds is still at its infancy, therefore, the use of computational strategies could facilitate this process and provide a template for the tailored design of immune protein inhibitors. This process could also provide insight into the mechanism of actions of these drugs which otherwise may have not been possible.

### *1.2 Aims and Objectives*

The aims and objectives of this research are listed as follow:

1. To unravel the mechanism of inhibition and structural perturbation of PD-L1 upon binding to BMS-1166 and BMS-1001. This will provide a cognizant understanding into how BMS-1166 and BMS-1001 disrupt PD-1-PD-L1 interactions. This will be achieved by:
  - 1.1 Performing molecular docking and extended molecular dynamic simulation run on the BMS-1001/BMS-1166 bound PD-L1 and the apoenzyme PD-L1.
  - 1.2 Using Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PB-SA) thermodynamic calculation to compute the differential binding of BMS-1166 and BMS-1001 on PD-L1.
2. To use a Per-residue decomposition-based virtual screening to propose potential small molecular weight inhibitors of CTLA-4 by:
  - 2.1 Using active site similarity drug repurposing approach to identify proteins possessing similar binding sites as CTLA-4 and to identify a potential lead compound.
  - 2.2 Performing an initial molecular dynamic simulation to determine structural ensembles and pharmacophoric moieties on CTLA-4 and a lead compound.
  - 2.3 Performing a library search of a compound database (ZINC database), to identify potential hits based on the pharmacophoric moieties identified.

- 2.4 Assessing the drug likeness of the hits compounds by carrying out chemoinformatic analysis.
- 2.5 Validating the hit compounds by estimating their binding energy and structural stability.
3. To identify and assess the impact of non-synonymous Single Nucleotide Polymorphism (nsSNPs) present in ULBP6 by:
  - 3.1 Retrieving all nsSNPs associated with ULBP6 present in NCBI-dbSNPs database
  - 3.2 Investigating whether the SNPs retrieved are damaging, disease-causing, deleterious, benign or neutral through the use of some computational tools.
  - 3.3 Using molecular dynamics simulation to investigate the structural impact of the identified SNPs.

### **1.3 Novelty and Significance of Research**

Despite the successes that have been recorded in the use of humanized antibodies as inhibitors for CTLA-4 and PD-L1, the limitations that have characterized humanized antibodies drugs has necessitated the introduction of small molecule inhibitors. Nevertheless, the design and development of small-molecular weight inhibitors still lie behind humanized antibodies drugs. Identification of potential small molecule inhibitors and unraveling their mechanism of inhibition using *in-silico* strategies could facilitate the drug development pipeline of some forms of cancer.

### Reference

- [1] Nicholson LB. The immune system. *Essays Biochem.* 2016;60:275–301.

- [2] Marshall JS, Warrington R, Watson W, et al. An introduction to immunology and immunopathology. *Allergy, Asthma Clin. Immunol.* 2018;14:1–10.
- [3] Immunity I. CHAPTER 14 Innate Immunity. *Immunology* [Internet]. 2006;125:1–4.  
Available from: <http://immuneweb.xxmu.edu.cn/wenzhai/pdf/003305.pdf>.
- [4] Bonilla FA, Oettgen HC. Adaptive immunity. *J. Allergy Clin. Immunol.* 2010;125:S33–S40.

## CHAPTER 2

### 2. Background on Immune System

#### 2.1 Introduction

The normal homeostasis of humans is constantly threatened by pathogenic and non-pathogenic microbes which in themselves contain an array of allergenic and toxic substances [1]. These microbes comprise of obligate and beneficial microorganism which are necessary for the normal functioning of organs and tissue, however, their activities are always regulated [2]. Pathogenic microbes possess diverse means with which they replicate, spread, and elicit a detrimental effect on the host [2]. As the host immune system eliminates these pathogenic microbes, some protective mechanisms are also put in place to avoid damage to the host tissue and beneficial microbes [3,4]. The host immune system leverage on the distinct structural difference that exists on the toxin or pathogen to distinguish it from the host system [3,4]. This discriminatory host-pathogen/toxin mechanism is quite important to enable the host to eliminate external assaults without causing damage to its tissues or organs.

In general, the immune system comprises of an aggregate of cells, chemicals, and process that protects intestinal tracts, respiratory passages, skin and other parts of the body from foreign bodies such as viruses, cancer cells, and microbes. Aside from the chemical and structural barriers employed by the body for defence, the immune system is broadly divided into adaptive immunity and innate immunity. The innate immune system is always the first line of defence against foreign assaults, the innate immune response is usually activated immediately there is contact with a

foreign body. The innate immune system does not have immunological memory, hence unable to recognise the antigen in the event of future exposure [3,4]. The adaptive immune system can keep memory of antigen on the first contact, hence facilitating effective and rapid response during future infection [2].

## **2.2 The innate immunity**

The innate immune system comprises of defensive barriers, anatomical, physiological, endocytic/phagocytic, and inflammatory defence. The anatomical barrier includes mucous membrane, skin, physiologic barriers mediated by temperature, low pH, and chemical mediators. Specialized cells like neutrophils, blood monocytes, and macrophages make up the phagocytic/endocytic barrier system [3]. Innate immunity helps in recruiting immune cells to points of inflammation and infection by producing chemokines and cytokines. Cytokines released early at the infection stage are tumour necrosis factor (TNF), interleukin 1 (IL-1), and interleukin 6 (IL-6). Dysregulated production of IL-6, TNF, and IL-1 has been implicated in autoimmune diseases or inflammation, making them an attractive therapeutic target in the treatment of autoimmune diseases. As stated above, several cells are required for innate immunity, such as basophils, dendritic cells, innate lymphoid cells, phagocytes, mast cells, eosinophils, and natural killer (NK) cells [2,3]. Macrophages and neutrophils are two sub-division of phagocytes that act in engulfing and killing of pathogens [5,6]. Neutrophils exterminate pathogens via enzymatic degradation and phagocytosis. They are the highest-circulating leukocytes found in the body [7]. Macrophages have been found to play the most central and important activity in the innate immune response. Mature macrophages differentiate from monocytes and localize at site that is prone to

pathogenic assaults, once a pathogen is encountered, macrophages activate microbial effector pathways [8].

Dendritic cells functions unify the adaptive and innate immunity [9]. Mast cells (MC) are a derivative of CD34+ hematopoietic progenitors, they are typically not found in the circulation [10]. MCs perform important function in acute inflammatory response [10]. Basophils are granulocytes emanating from progenitor cells in bone marrow, they are not always found in tissues, they are usually recruited to inflammatory sites [11]. Eosinophils, are eosinophilic or “acid-loving” as a result of their large acidophilic cytoplasmic granules. They destroy parasites that are too large for phagocytosis [12]. Like any other granulocytes, development and differentiation occur in the bone marrow [12]. Natural Killer (NK) cells perform crucial function in the destruction of virus-infected cells, this is carried out via releasing granzyme and perforins into the blood circulation. NK cells exhibit wide tissue distribution and phenotypic viability [13].

### **2.3 Adaptive Immunity**

The adaptive immune response is activated when the innate immune response fails to eliminate pathogenic assault. The major role of the adaptive immune system is to discriminate self-antigen and non-self-antigen, production of effector pathways that annihilate definite pathogens, and establishment of memory cells that can eliminate pathogens should the cells get infected again [2]. B lymphocytes, antibody-producing cells, and effectors of cellular response are the major cells of adaptive immunity. T-lymphocytes emanate and mature at the hematopoietic cells and the thymus. T-cells receptors (TCR) are distinctive antigen-binding receptors expressed on the membrane of T-cells which are responsible for recognising antigens from tumours, pathogens, and the environment [14]. Peripheral T-Cells encompasses naïve T cells which have the potential of

responding to fresh antigens [15]. T-cells do not perform the same function throughout life stages, they are found in practically every organ and tissue in the body including mucosal and barrier sites, lymphoid tissue, exocrine organs, fat, and the CNS [14,15]. Most of the T-cells are found within the mucosal sites, lymphoid tissues, and the skin. B cells originate from hematopoietic stem cells, unlike other cells, B cell expresses B cell receptors (BCR) on their membrane, these receptors enable B cell to recognise and bind to a specific antigen, against which it elicits antibody response [16,17]. At the developmental stage in the bone marrow, B-cells undergoes positive and negative selection. The positive selection takes place through a signalling that is independent of antigen using pre-BCR and the BCR [17]. B-cell development and signalling are halted when BCR and ligand interaction are not exhibited [16,17]. Negative selection is facilitated when self-antigen binds with BCR, upon this binding, B cells undergo the following, receptor editing, anergy, clonal deletion, or it disregards signal and carry on with development [16]. B cells also produce antibodies that have also help in facilitating effective immune response [2,18]. IgA, IgD, IgE, IgG, and IgM are the major antibodies produced by B cells. These antibodies perform a different biological function and their recognition patterns are different [18,19].

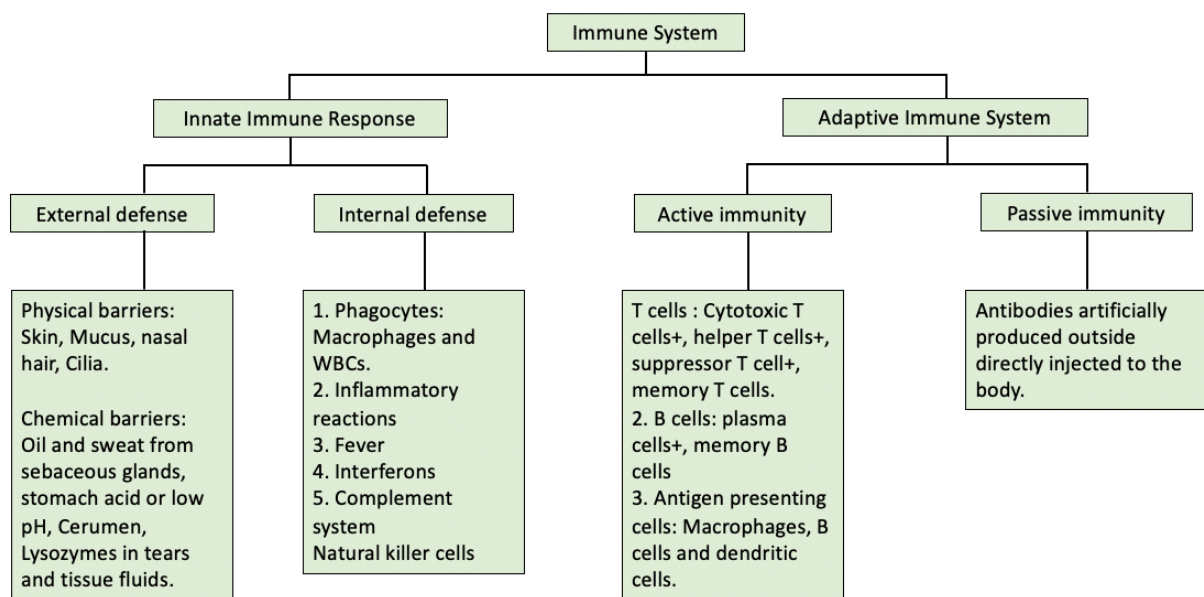


Figure 2.1 A schematic depiction of the immune system

## **2.4 Physiological control of the immune System**

In conjunction with other systems, the immune system performs a vital task in the physiological control of the host system. This pathological regulation is facilitated by nutrition and diet, hormones, vitamin D, repair and regeneration, and sleep and rest. Hormones function as immunomodulators by disrupting the responsiveness of the immune system [20]. The female sex hormones act as immunostimulators while the male hormones act as immunosuppressants [21]. Nutrition has also been implicated in immune response, there is a correlation between diseases such as obesity and diabetes which has been reported to impact immune response. Mineral deficiency can also affect the immune response [22].

## **2.5 Disorders of human immunity**

As much as the immune system acts in body defence, the immune system can also be overwhelmed or overcome by foreign bodies. Some mutations also have been reported to disrupt the immune response [23]. These disorders/diseases are categorized into autoimmunity, immunodeficiencies, and hypersensitivities [23]. Immunodeficiencies are characterized by the inability of the immune system to elicit immune response due to the inactivation of some immune system components [23]. Autoimmunity is described as the inability of the immune system to differentiate between self-antigen and foreign-antigen thus attacking itself. The set of disorders that arise due to this immune malfunction is called autoimmune disorders. Hypersensitivity is an immune phenomenon where immune response destroys the body's own tissue cells. Four classes of hypersensitivity (I-IV) have been identified, this identification is based on the mechanism elicited and the duration of the



hypersensitivity [24,25]. Type 1 hypersensitivity is characterized by a quick response and is facilitated by IgE. Type II is facilitated by IgG and IgM, this hypersensitivity is brought about by antibody-antigen binding which marks the host's cell for annihilation. [24,25]. Type III hypersensitivity is triggered by complement proteins, IgG, antigens, and IgM. Unlike Type 1 hypersensitivity which is immediate, type IV is triggered between 2 to 3 days. For this reason, it is called delayed-type or cell-mediated hypersensitivity [24,25]. Inflammation can arise without any known cause, although it is the first immune response to infections [24,25].

## **2.6 Cancer and the Immune System**

Cancer is a deadly disease resulting from unregulated cell division, about a century and half ago, the connection between cancer and the immune system was first detailed by Rudolph Virchow [26]. The immune system respond to cancerous cell by employing three basic mechanisms which include: 1) detection of foreign substances “non-self” antigens or aberrant cells, 2) activation of effector activity of cells such as macrophages, natural killer cells, etc to target and annihilate the corrupted/aberrated cells, and 3) establishment of memory cells through the adaptive immune responses for future defence of the body against the same assault. However, over the years, cancer cells have developed mechanisms with which they use in evading immune responses and subsequent destruction by immune cells. This birthed a new treatment option in the field of cancer therapy called cancer immunotherapy [26].

## **2.7 Cancer and Cancer Immunotherapy (CI)**

Immune surveillance which is the ability of the host's immune system to annihilate malignant cells during initial transformation have been leveraged on in treating cancer cells [26]. Genetic and epigenetic changes have been implicated in the onset of tumorigenesis and subsequently the hallmarks of cancer [26]. These epigenetic and genetic changes also elicit the production of neo-antigen which makes the neoplastic cells noticeable by the immune system and targeted for destruction [26,27]. Despite all the mechanisms launched by the immune system to obliterate aberrated cells, tumour cells still bypass identification and consequent destruction by the immune system [26,27]. This immune escape is possible through the development of multiple resistance mechanisms such as induction of tolerance, immune invasion, and coordinated destruction of T cell signalling [26,27]. Over the years, scientists have come to understand that cancerous cells are efficient in repressing the host's immune response, this has necessitated treatment options such as surgery, chemotherapy, and radiation therapy to be explored in cancer treatment [26]. The approval of sipuleucel-T and ipilimumab has revitalized the field of CI. CI was adjudged the "breakthrough of the year" by Science in 2013 due to the success rate recorded from different treatments in the clinical phase [26]. William Coley in 1893 conceptualized the idea that the immune system can recognize and control tumour growth using live bacteria as an immune stimulant to treat cancer. Tumour cells evade immune recognition and elimination using diverse mechanisms, this has limited the therapeutic efficiency of CI [28]. Mechanisms employed by tumour cells to bypass immune surveillance include augmented differentiation of immune effector cells, production of immunosuppressive cytokines, an intrusion of myeloid-derived suppressor cells, and tumour-related macrophages and decreased expression of MHC antigens [29]. The major objective of CI is to revive the repressed immune system to reactivate dormant immune response mechanisms to fight tumour cells thereby eliminating cancer [26].

## **2.8 Cancer immune cycle**

For a cancer cell to be effectively annihilated by an immune response, a sequential process which is interlinked must be initiated and permitted to proceed. Chen and Mellman described this process and called it the cancer-immunity cycle [30]. The cancer immune cycle is described stepwise below [31]:

1. Neoantigens produced via oncogenesis are seized by DCs for processing.
2. The captured antigens arrested by DCs present on MHC1 and MHCII molecules are presented to T cells.
3. This presentation to T cells leads to priming and activation of effector T cell responses.
4. The T cells are trafficked to the tumour cells
5. Infiltration of tumour cells by T cells
6. Identification of tumour cells by T cells
7. Tumour cells are killed (Figure 2.2).

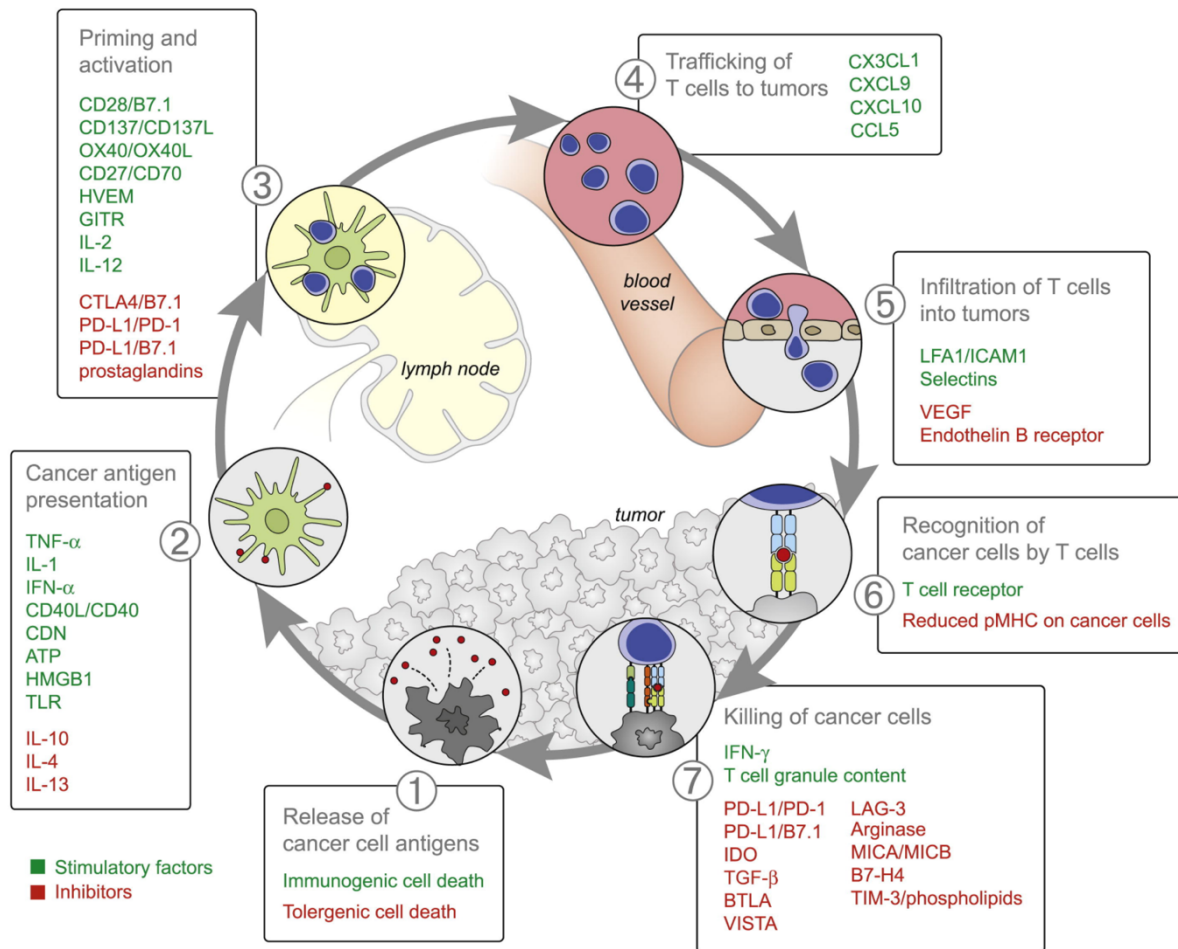


Figure 2.2 The cancer immune cycle highlighting molecules involved in each step [30].

## 2.9 Cytotoxic T-lymphocyte-associated antigen

Cytotoxic T-lymphocyte-associated antigen (CTLA-4) was first discovered by Pierre Goisten and his team [32], shortly after that, a group published their findings on the ability of CTLA-4 to act as a negative controller of T cell activation through gene knockout in mice [33,34]. CTLA-4 also referred to CD152, is a glycoprotein type membrane protein produced by activated effector T cells (Teffs) [35], whose major function is the negative control of T-cell activity [36]. CTLA-4 comprises three important regions, ectodomain, a transmembrane domain, and cytoplasmic tail

[37].

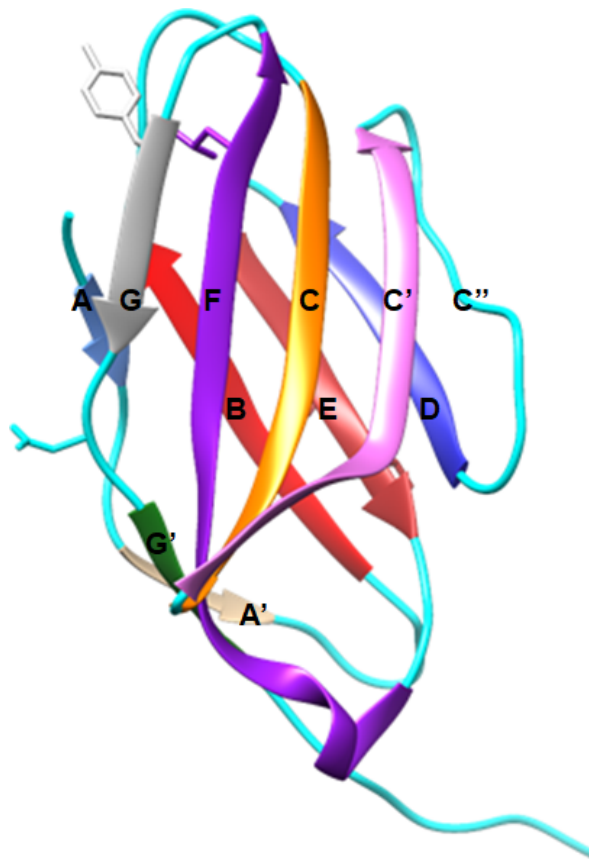


Figure 2.3 3D structure of human apo-CTLA-4 monomer showing the A'GFCC'C'' sheet and BED making up the extracellular domain (PDB ID: 3OSK ).

The homodimeric form connected by disulphide bond is always found in the membrane while the monomeric form is soluble [36]. The extracellular domain has two beta-sheets comprising the A'GFCC'C'' loop and the ABED strands (Figure 2.3). The dimeric nature of CTLA-4 is due to the covalent disulphide bond formed by the cysteine residues (Cys122) found within the stabilizing loop connecting the ectodomain and the transmembrane domain. The polyproline part (MYPPPY) found in the loops connecting the F and G strands, serves as the ligand-binding sites, B7-1 and B7-2 [38] (Figure 2.4). CTLA-4 is located in the intracellular compartment during the resting phase of T-cells. Upon activation by binding to CD28, CTLA-4 moves to the surface of T-cells where it

is expressed [39–41]. The amount of transported and expressed CTLA-4 on the surface of T cell is dependent on how strong the stimulatory signal is [42]. Upon expression on T cell surface, inhibition by CTLA-4 is facilitated when B7-1 and B7-2 present on triggered monocytes, and B cells binds. CTLA-4 activity goes beyond costimulation blockade, it has been reported that CTLA-4 activation impedes IL-2 transcription inside active T cells, consequently altering their advancement in the cell cycle [43].

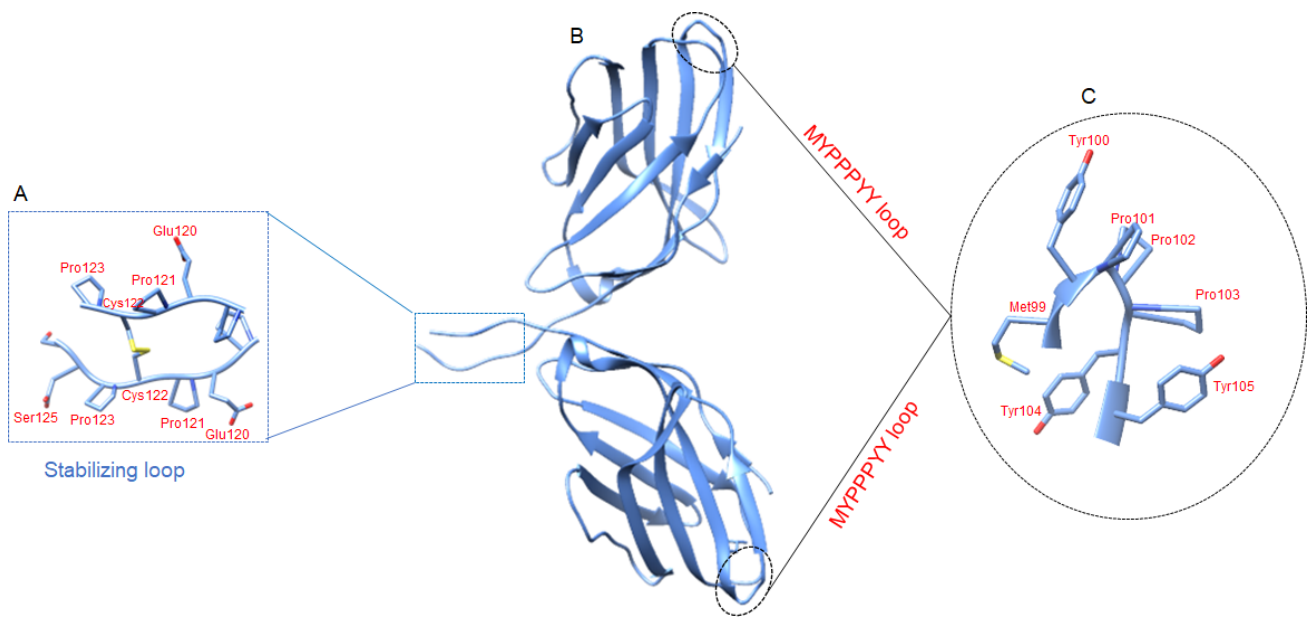


Figure 2.4: Symmetric apo-CTLA-4 homodimer (B), showing the polyproline motif (C) which is an attractive target site for ligands and the stabilizing loop, showing an interaction between the Cys122 residues of both monomers (A).

In addition, CTLA-4 is also involved in the dephosphorylation of important TCR signalling kinases and RAS pathway proteins such as ZAP-70, LCK, FYN, SHP-1, PP2A, and SHP-2 [41]. Traces of CTLA-4 has also been found in expressed Tregs [41]. Tumour cells can alter the activation of immune costimulatory pathways hence, represses antitumour immune response. Kwon et al., used tumour cells gotten from TRAMP (transgenic adenocarcinoma of the mouse prostate) mouse to reveal that *in-vivo* blocking of CTLA-4 can strengthen T-cell assisted reversion

of subcutaneous prostate cancer in C57BL/6 mice [44]. Allison et al. demonstrated that the blocking of CTLA-4 by antibodies incubated with tumour lysate-enriched dendritic cells lowered Tregs, impeded metastatic growth, and increased CD8+ lymphocytes [41,45]. Anti-CTLA-4 humanized antibodies strengthened anti-prostate cancer immune response in tumour infected cells and consequently slowing down tumour growth [41]. IgG1 $\kappa$  (ipilimumab) and IgG2 (tremelimumab) antibodies have been developed to target CTLA-4 in patients with cancer [46].

Table 2.1 : Activity of CTLA-4 antibody in metastatic melanoma patients

<b>Antibody</b>	<b>Combination</b>	<b>Dose</b>	<b>Reference</b>
Ipilimumab	None	3 mg/kg	[47]
Ipilimumab	Gp100 peptides	3 mg/kg	[46,48]
Ipilimumab	Gp100, tyrosinase, MART-1 peptides	0.3-3 mg/kg	[49]
Ipilimumab	Il-2	0.1-3 mg/kg	[50]
Ipilimumab	None	3-9 mg/kg	[51]
Tremelimumab	None	0.01-15 mg/kg	[52]
Tremelimumab	None	10-15 mg/kg	[53]

## 2.10 Programmed Death Ligand -1

Programmed death-ligand 1 (PD-L1) is an immune checkpoint protein translated from the human gene CD274 [54]. It is made up of 290 amino acid residues translated from the Cd274 gene, with 7 exons, and it is found on chromosome 9 [54]. PD-L1 is a 40kDa type 1 transmembrane protein

also referred to as cluster of differentiation 279 that is critical in repressing the immune system and self-tolerance promotion via T cell inflammatory activity suppression [55,56]. This T cell suppression helps to avert autoimmune diseases however, the capability of the immune system to annihilate cancerous cells is altered [55]. Structurally, it has two extracellular regions, cytoplasmic domain, and a transmembrane domain (Figure 2.5).

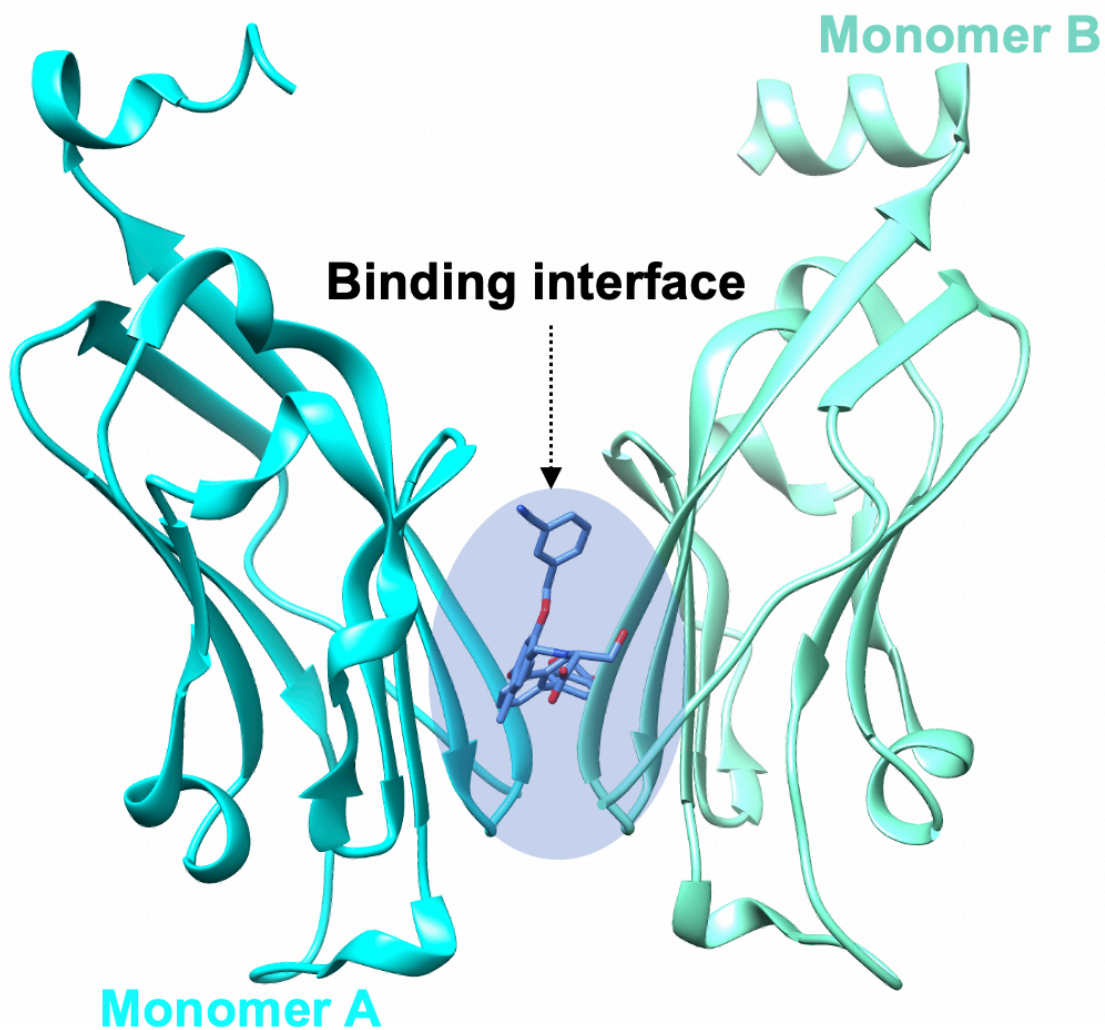


Figure 2.5: 3-dimensional structure of PD-L1 protein with monomer A highlighted in aquamarine, monomer B in cyan, and a small compound bound to the binding interface between monomer A and monomer B. PDB: 5NIX [19]



PD-L1 was initially referred to as a regulatory protein B7-H1 but was renamed PD-L1 after it was discovered to bind PD-1 as a ligand [58]. The binding of PD-L1 to its receptor PD-1 modulates inhibition or activation of myeloid cells, B cells, and T cells. This binding activates a cascade of reaction that involves the obstruction of TCR-mediated T cell proliferation and IL-2 synthesis. It has been found that PD-L1 binds to other receptors such as B7 (CD80) [59].

The expression of PD-L1 is upregulated in different forms of cancer, it often contributes to poor prognosis and a great determinant of anti-PD-1/PD-L1 antibodies [59]. Cells that do not express PD-L1 on their surfaces have also been found to react to anti-PD-1 antibodies. Regulation of PD-L1 expression is mediated by interferons, macrophages, monocytes, microRNAs, and epigenetic regulation [60].

Immune checkpoint inhibitors (ICI), principally those targeting PD-L1 and PD-1 showed favourable therapeutic activity against various hematologic diseases and malignancies [54]. The binding of PD-1 and PD-L1 represses T cell movement, proliferation, and production of cytotoxic agents. Anti-PD-1/PD-L1 agents interrupt the PD-L1/PD-1 interactions consequently causing the reversal of T cell repression, and intensify the constitutive antitumour immune response to release long-term anti-tumour response [55]. There are three PD-L1 inhibitors ratified by the FDA for the treatment of different forms of cancers. There are however, other different PD-L1 inhibitors still undergoing clinical trials.

Atezolizumab is an anti-PD-L1 IgG monoclonal Ab that is wholly humanized. The Fc domain is engineered with an alteration that abolishes antibody-dependent cell toxicity to avert exhaustion of T cells exhibiting PD-L1. Atezolizumab occludes the binding of PD-L1 with PD-1 on tumour-

infiltrating immune cells. Induction of change in the level of cellular cytokine and temporary increase in CXCL11, IL-18, and IFN $\gamma$  through an increased level of circulating CD8<sup>+</sup> T cells has been detailed in preclinical investigations [61,62]. Via the blockade of PD-L1, atezolizumab minimizes immunosuppressive communications and concurrently elevates T cell-assisted immune response [61]. Durvalumab is an anti-PD-L1 IgG monoclonal Ab that is wholly humanized possessing strong binding and selectivity for PD-L1. It disrupts the binding of CD80 and PD-1 [63]. Durvalumab is modified to intercept antibody-dependent cell toxicity on T cells exhibiting PD-L1. The progression of human tumour cells in a xenograft model accommodating co-implanted human T cells was significantly halted by durvalumab [63]. Avelumab is yet another anti-PD-L1 IgG monoclonal Ab that is wholly humanized. It averts the binding of PD-L1 with T cell receptors, B7.1, and PD-1 thereby facilitating T cell resuscitation and cytokine production [63,64]. Avelumab possesses a wild-type IgG1 Fc region with which it interacts with Fc- $\gamma$  receptor on NK cells and generates tumour-guided antibody-dependent cell toxicity in preclinical studies [64].

PD-L1 comprises of three parts, IgC-like extracellular domain, a transmembrane domain, and cytoplasmic region. In the interaction of PD-L1 with PD-1, some crucial residues have been implicated. For instance, Q75 of PD-1 forms 3 pairs of H-bond with D26 and R125 of PD-L1. T76 of PD-1 forms a bond with Y123 of PD-L1. These bonds serve as a molecular platform for the design of small molecular weight inhibitors. Small molecular weight inhibitors that have been designed to target PD-L1 are listed in (Table 2.2).

Table 2.2: Small molecule inhibitors designed to target PD-L1

<b>Drug</b>	<b>Stage</b>	<b>Treatment</b>	<b>Reference</b>
AUNP-12	Preclinical Phase	Melanoma	[65]
DPPA-1	Preclinical Phase	Colon cancer	[66]
TPP-1	Preclinical Phase	LCLC	[67]
BMS-202	Preclinical Phase		[65,68]
CA-170	Clinical Phase	NSCLC	[69]
JQ1	Clinical Phase	Lymphoma	[70]
eFT508	Clinical Phase	Liver cancer	[71]
Osimertinib	Preclinical Phase	NSCLC	[72]
PlatycodinD	Preclinical Phase	NSCLC	[73]
PD-LYLSO	Preclinical Phase		[74]
Curcumin	Preclinical Phase	Breast cancer	[75]
Metformin	Preclinical Phase	Breast cancer	[76]

### 2.11 UL16-Binding Protein

NH group 2, member D (NKG2D) is a crucial immune protein that is involved in anti-cancer and anti-pathogen immune responses [77]. It is usually expressed in NKC,  $\alpha\beta$ T, and  $\gamma\delta$ T of human cells [77]. A characteristic feature of NKG2D is its binding with an array of various NKG2D ligands found of the cell surface [78].

The interacting partner of NKG2D in humans are the MICA/B and ULBP proteins. ULBP was discovered recently to be one of the highly polymorphic NKG2D ligands. The two haplotypes

ULBP0601 and ULBP0602 make up about 70% of the human population [77,79]. ULBP/RAET gene is located on chromosome 6q24.2-q25.3. The polymorphic forms of ULBP6 have been implicated in alopecia areata [80] and nephropathy [81]. ULB6 has ten loci [77], four are pseudogenes while five are translated into functional transcripts [77]. ULBP0601 and ULBP0602 have different binding ability with NKG2D, this disparity enhances different capacity to trigger effector cells [78]. The proposed mechanism of this differential binding is via a highly polymorphic amino acid change at R106L [78]. As revealed by the structure of NKG2D-ULBP0602, ULBP0602 inserts R106L into a non-polar cavity of NKGD2 to form a unique interface with hydrophobic hotspot [78].

The capability of NKGD2 to bind with different constitutive ligands (ULBP1 to ULBP6, MICA, and MICAB) has been suggested to be via rigid body adaptation or induced fit [78], however, ULBP3 and ULBP6 bind distinctly with NKGD2 [78]. Despite the difference in interaction profiles between NKGD2 and its ligands, E96 and D189 of ULBP6 interacts with K of patch A and B of NKGD2 respectively. The involvement of ULBP1 in cancer immune response has been established, for instance, Kamei et al., pointed out that the expression level of ULBP1 and NKG2D influence the overall survival of patients with gastric cancer [82]. Also, Rolle et al., discovered that the defence mechanism of the host to Human Cytomegalovirus infection is via the expression of ULBPs [83].

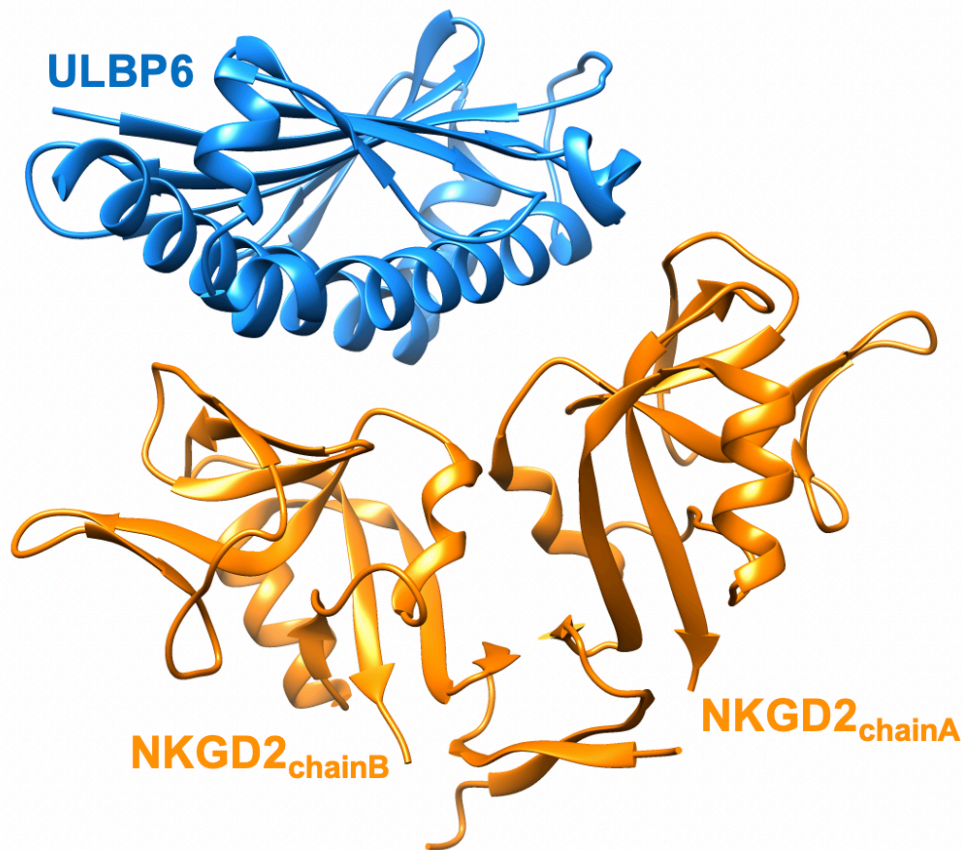


Figure 2.6: 3-dimensional structure of ligand ULBP6 in complex with two chains of NKGD2D (PDB ID: 4S0U) [78].

### Reference

- [1] Nicholson LB. The immune system. *Essays Biochem.* 2016;60:275–301.
- [2] Marshall JS, Warrington R, Watson W, et al. An introduction to immunology and immunopathology. *Allergy, Asthma Clin. Immunol.* 2018;14:1–10.
- [3] Immunity I. CHAPTER 14 Innate Immunity. *Immunology* [Internet]. 2006;125:1–4. Available from: <http://immuneweb.xmu.edu.cn/wenzhai/pdf/003305.pdf>.
- [4] Bonilla FA, Oettgen HC. Adaptive immunity. *J. Allergy Clin. Immunol.* 2010;125:S33–S40.
- [5] Prame Kumar K, Nicholls AJ, Wong CHY. Partners in crime: neutrophils and

- monocytes/macrophages in inflammation and disease. *Cell Tissue Res.* 2018;371:551–565.
- [6] Butterfield T a, Best TM, Merrick M a. The Dual Roles of Neutrophils and. *J. Athl. Train.* 2006;41:457–465.
- [7] Rosales C. Neutrophil: A cell with many roles in inflammation or several cell types? *Front. Physiol.* 2018;9:1–17.
- [8] Nenasheva T, Gerasimova T, Serdyuk Y, et al. Macrophages Derived From Human Induced Pluripotent Stem Cells Are Low-Activated “Naïve-Like” Cells Capable of Restricting Mycobacteria Growth. *Front. Immunol.* 2020;11:1–15.
- [9] Patente TA, Pinho MP, Oliveira AA, et al. Human dendritic cells: Their heterogeneity and clinical application potential in cancer immunotherapy. *Front. Immunol.* 2019;10:1–18.
- [10] Cruvinel WM, Júnior DM, Araújo JAP, et al. Immune system - part I fundamentals of innate immunity with emphasis on molecular and cellular mechanisms of inflammatory response. *Rev. Bras. Reumatol.* 2010;50:443–461.
- [11] Min B, Brown MA, Legros G. Understanding the roles of basophils: Breaking dawn. *Immunology.* 2012;135:192–197.
- [12] Wen T, Rothenberg ME. The Regulatory Function of Eosinophils. *Myeloid Cells Heal. Dis.* 2017;4:257–269.
- [13] Paul S, Lal G. The molecular mechanism of natural killer cells function and its importance in cancer immunotherapy. *Front. Immunol.* 2017;8.
- [14] Pennock ND, White JT, Cross EW, et al. T cell responses: Naïve to memory and everything in between. *Am. J. Physiol. - Adv. Physiol. Educ.* 2013;37:273–283.
- [15] Kumar B V, Connors T, Farber DL. *life.* 2019;48:202–213.
- [16] Mårtensson IL, Almqvist N, Grimsholm O, et al. The pre-B cell receptor checkpoint. *FEBS*

- Lett. 2010;584:2572–2579.
- [17] Lebien TW, Tedder TF. B lymphocytes: How they develop and function. *Blood*. 2008;112:1570–1580.
- [18] Chiu ML, Goulet DR, Teplyakov A, et al. Antibody Structure and Function: The Basis for Engineering Therapeutics. *Antibodies*. 2019;8:55.
- [19] Schroeder HWJ, Cavacini L. Structure and Function of Immunoglobulins (author manuscript). *J. Allergy Clin. Immunol.* 2010;125:S41–S52.
- [20] Moriyama A, Shimoya K, Ogata I, et al. Secretory leukocyte protease inhibitor (SLPI) concentrations in cervical mucus of women with normal menstrual cycle. *Mol. Hum. Reprod.* 1999;5:656–661.
- [21] Fimmel S, Zouboulis CC. Influence of physiological androgen levels on wound healing and immune status in men. *Aging Male*. 2005;8:166–174.
- [22] Langley-Evans SC, Carrington LJ. Diet and the developing immune system. *Lupus*. 2006;15:746–752.
- [23] Aw D, Silva AB, Palmer DB. Immunosenescence: Emerging challenges for an ageing population. *Immunology*. 2007;120:435–446.
- [24] Baldo BA, Pham NH. Drug allergy: Clinical aspects, diagnosis, mechanisms, structure-activity relationships. *Drug Allergy Clin. Asp. Diagnosis, Mech. Struct. Relationships*. 2013.
- [25] Dispenza MC. Classification of hypersensitivity reactions. *Allergy Asthma Proc*. 2019;40:470–473.
- [26] Sambhi M, Bagheri L, Szewczuk MR. Current challenges in cancer immunotherapy: Multimodal approaches to improve efficacy and patient response rates. *J. Oncol.* 2019;2019.

- [27] Farkona S, Diamandis EP, Blasutig IM. Cancer immunotherapy: The beginning of the end of cancer? *BMC Med.* 2016.
- [28] Yang Y. Cancer immunotherapy: Harnessing the immune system to battle cancer. *J. Clin. Invest.* 2015;125:3335–3337.
- [29] Prodeus A, Abdul-Wahid A, Fischer NW, et al. Targeting the PD-1/PD-L1 immune evasion axis with DNA aptamers as a novel therapeutic strategy for the treatment of disseminated cancers. *Mol. Ther. - Nucleic Acids.* 2015;4:e237.
- [30] Chen DS, Mellman I. Oncology meets immunology: The cancer-immunity cycle. *Immunity.* 2013;39:1–10.
- [31] Motz GT, Coukos G. Deciphering and Reversing Tumor Immune Suppression. 2014;39:61–73.
- [32] Golstein P., Brunet, JF, Luciani MF, Dosseto-Roux M, Suzan M, Mattei P. A new member of the immunoglobulin superfamily - CTLA-4. *Nature.* 1987;328(6127):267-270.
- [33] Waterhouse P, Penninger JM, Timms E, et al. Lymphoproliferative Disorders with Early Lethality in Mice Deficient in Ctla-4. *Science.* 1995;270(5238):985-988.
- [34] Tivol EA, Borriello F, Schweitzer AN, et al. Loss of CTLA-4 Leads to Massive Lymphoproliferation and Fatal Multiorgan Tissue Destruction , Revealing a Critical Negative Regulatory Role of CTLA-4. 1995;3:541–547.
- [35] Li B, Zhao Y, Yang W, Huang Y, Cui R, Li X. Evolving Roles for Targeting CTLA-4 in Cancer Immunotherapy. *Cell Physiol Biochem.* 2018;47721–734.
- [36] Ramagopal UA, Liu W, Garrett-Thomson SC, et al. Structural basis for cancer immunotherapy by the first-in-class checkpoint inhibitor ipilimumab. *Proc. Natl. Acad. Sci.* 2017;114:E4223–E4232.



- [37] Kumar SP, Jha PC. Multi-level structure-based pharmacophore modelling of caspase-3-non-peptide complexes: Extracting essential pharmacophore features and its application to virtual screening. *Chem. Biol. Interact.* 2016;254:207–220.
- [38] Scheipers P. Role of the CTLA-4 Receptor in T Cell Activation and Immunity Physiologic Function of the CTLA-4 receptor. 1998;103–115.
- [39] Linsley PS, Bradshaw J, Greene JA, et al. Intracellular trafficking of CTLA-4 and focal localization towards sites of TCR engagement. *Immunity.* 1996;4:535–543.
- [40] Chang J-F, Phillips J, Kufe D, et al. MUC1 can function as a potent negative regulator of T cell activation. *FASEB J.* 1998;12:405–413.
- [41] Buchbinder E, Hodi FS, Buchbinder E, et al. Cytotoxic T lymphocyte antigen-4 and immune checkpoint blockade Find the latest version : Cytotoxic T lymphocyte antigen-4 and immune checkpoint blockade. *J. Clin. Invest.* 2015;125:3377–3383.
- [42] Egen JG, Allison JP. Cytotoxic T lymphocyte antigen-4 accumulation in the immunological synapse is regulated by TCR signal strength. *Immunity.* 2002;16:23–35.
- [43] Brunner MC, Chambers CA, Chan FK, et al. CTLA-4-Mediated inhibition of early events of T cell proliferation. *J. Immunol.* 1999;162:5813–5820.
- [44] Kwon ED, Hurwitz AA, Foster BA, et al. Manipulation of T cell costimulatory and inhibitory signals for immunotherapy of prostate cancer. *Proc. Natl. Acad. Sci. U. S. A.* 1997;94:8099–8103.
- [45] Leach DR, Krummel MF, Allison JP. Enhancement of antitumour immunity by CTLA-4 blockade. *Science (80-. ).* 1996;271:1734–1736.
- [46] Keler T, Halk E, Vitale L, et al. Activity and Safety of CTLA-4 Blockade Combined with Vaccines in Cynomolgus Macaques. *J. Immunol.* 2003;171:6251–6259.

- [47] Hodi FS, Mihm MC, Soiffer RJ, et al. Biologic activity of cytotoxic T lymphocyte-associated antigen 4 antibody blockade in previously vaccinated metastatic melanoma and ovarian carcinoma patients. *Proc. Natl. Acad. Sci. U. S. A.* 2003;100:4712–4717.
- [48] Phan GQ, Yang JC, Sherry RM, et al. Cancer regression and autoimmunity induced by cytotoxic T lymphocyte-associated antigen 4 blockade in patients with metastatic melanoma. *Proc. Natl. Acad. Sci. U. S. A.* 2003;100:8372–8377.
- [49] Sanderson K, Scotland R, Lee P, et al. Autoimmunity in a phase I trial of a fully human anti-cytotoxic T-lymphocyte antigen-4 monoclonal antibody with multiple melanoma peptides and montanide ISA 51 for patients with resected stages III and IV melanoma. *J. Clin. Oncol.* 2005;23:741–750.
- [50] Maker A V., Phan GQ, Attia P, et al. Tumour regression and autoimmunity in patients treated with cytotoxic T lymphocyte-associated antigen 4 blockade and interleukin 2: A phase I/II study. *Ann. Surg. Oncol.* 2005;12:1005–1016.
- [51] Maker A V., Yang JC, Sherry RM, et al. Inpatient dose escalation of anti-CTLA-4 antibody in patients with metastatic melanoma. *J. Immunother.* 2006;29:455–463.
- [52] Reuben JM, Lee BN, Li C, et al. Biologic and immunomodulatory events after CTLA-4 blockade with ticitimumab in patients with advanced malignant melanoma. *Cancer.* 2006;106:2437–2444.
- [53] Ribas A, Camacho LH, Lopez-Berestein G, et al. Antitumour activity in melanoma and anti-self responses in a phase I trial with the anti-cytotoxic T lymphocyte-associated antigen 4 monoclonal antibody CP-675,206. *J. Clin. Oncol.* 2005;23:8968–8977.
- [54] Akinleye A, Rasool Z. Immune checkpoint inhibitors of PD-L1 as cancer therapeutics. *J. Hematol. Oncol.* 2019;12:1–13.

- [55] Shen X, Zhang L, Li J, et al. Recent findings in the regulation of programmed death ligand 1 expression. *Front. Immunol.* 2019;10:1–25.
- [56] Keir ME, Butte MJ, Freeman GJ, et al. PD-1 and its ligands in tolerance and immunity. *Annu. Rev. Immunol.* 2008;26:677–704.
- [57] Skalniak L, Zak KM, Guzik K, et al. Small-molecule inhibitors of PD-1/PD-L1 immune checkpoint alleviate the PD-L1-induced exhaustion of T-cells. *Oncotarget.* 2017;
- [58] Phoenix CH, Chambers KC. Sexual performance of old and young male rhesus macaques following treatment with GnRH. *Physiol. Behav.* 1990;47:513–517.
- [59] Paterson AM, Brown KE, Keir ME, et al. The PD-L1:B7-1 pathway restrains diabetogenic effector T cells in vivo. *J. Immunol.* 2011;187:1097–1105.
- [60] Ciardiello F., Normanno N, Martinelli E, Troiani T, Pisconti S, Cardone C, Nappi A, Bordonaro AR, Rachiglio M, Lambiase M, Latiano TP, Modoni G, Cordio S, Giuliani F, Biglietto M, Montesarchio V, Barone C, Tonini G, Cinieri S, Febbraro A, Rizzi D, De Vita MEC-G investigators. *Annals of Oncology Advance Access.* *Ann Oncol.* 2016;Jun:1055–1061.
- [61] Deng R, Bumbaca D, Pastuskovas C V., et al. Preclinical pharmacokinetics, pharmacodynamics, tissue distribution, and tumour penetration of anti-PD-L1 monoclonal antibody, an immune checkpoint inhibitor. *MAbs [Internet].* 2016;8:593–603. Available from: <http://dx.doi.org/10.1080/19420862.2015.1136043>.
- [62] Inman BA, Longo TA, Ramalingam S, et al. Atezolizumab: A PD-L1-blocking antibody for bladder cancer. *Clin. Cancer Res.* 2017;23:1886–1890.
- [63] Diggs LP, Hsueh EC. Utility of PD-L1 immunohistochemistry assays for predicting PD-1/PD-L1 inhibitor response. *Biomark. Res.* 2017;5:1–6.

- [64] Grenga I, Donahue RN, Lepone LM, et al. A fully human IgG1 anti-PD-L1 MAb in an in vitro assay enhances antigen-specific T-cell responses. *Clin. Transl. Immunol.* 2016;5:e83.
- [65] Wu Q, Jiang L, Li S cheng, et al. Small molecule inhibitors targeting the PD-1/PD-L1 signaling pathway. *Acta Pharmacol. Sin.* [Internet]. 2020; Available from: <http://dx.doi.org/10.1038/s41401-020-0366-x>.
- [66] Li C, Zhang N, Zhou J, et al. Peptide Blocking of PD-1/PD-L1 interaction for cancer immunotherapy. *Cancer Immunol. Res.* 2018;6:178–188.
- [67] Macrocyclic Peptides Useful as Immunomodulators. c/o WO2016077518A1 Company S. PCT. 2016;
- [68] CHUPAK, Louis S.; ZHENG X c/o. WO 2015/034820 A1. 2015;198.
- [69] Sasikumar PG, Ramachandra M. Small-Molecule Immune Checkpoint Inhibitors Targeting PD-1/PD-L1 and Other Emerging Checkpoint Pathways. *BioDrugs* [Internet]. 2018;32:481–497. Available from: <https://doi.org/10.1007/s40259-018-0303-4>.
- [70] Fang W, Zhang J, Hong S, et al. EBV-driven LMP1 and IFN- $\gamma$  up-regulate PD-L1 in nasopharyngeal carcinoma: Implications for oncotargeted therapy. *Oncotarget.* 2014;5:12189–12202.
- [71] Azuma K, Ota K, Kawahara A, et al. Association of PD-L1 overexpression with activating EGFR mutations in surgically resected nonsmall-cell lung cancer. *Ann. Oncol.* [Internet]. 2014;25:1935–1940. Available from: <https://doi.org/10.1093/annonc/mdu242>.
- [72] Wang CW, Klionsky DJ. The molecular mechanism of autophagy. *Mol. Med.* 2003;9:65–76.
- [73] Huang MY, Jiang XM, Xu YL, et al. Platycodin D triggers the extracellular release of prograded death Ligand-1 in lung cancer cells. *Food Chem. Toxicol.* [Internet].

- 2019;131:110537.
- [74] Nicholas Dias, Yung Peng RK. deubiquitination and Stabilization of PD-L1 by CSN5. *Physiol. Behav.* 2017;176:139–148.
- [75] Viollet B, Guigas B, Sanz Garcia N, et al. Cellular and molecular mechanisms of metformin: An overview. *Clin. Sci.* 2012;122:253–270.
- [76] Eikawa S, Nishida M, Mizukami S, et al. Immune-mediated antitumour effect by type 2 diabetes drug, metformin. *Proc. Natl. Acad. Sci. U. S. A.* 2015;112:1809–1814.
- [77] Ñ RAE, Ñ JAT, Hair JR, et al. ULBP6 / RAET1L is an additional human NKG2D ligand. 2009;3207–3216.
- [78] Zuo J, Willcox CR, Mohammed F, et al. A disease-linked ULBP6 polymorphism inhibits NKG2D-mediated target cell killing by enhancing the stability of NKG2D ligand binding. 2017;8904:1–15.
- [79] Antoun A, Jobson S, Cook M, et al. Single nucleotide polymorphism analysis of the NKG2D ligand cluster on the long arm of chromosome 6: Extensive polymorphisms and evidence of diversity between human populations. *Hum. Immunol.* [Internet]. 2010;71:610–620. Available from: <http://dx.doi.org/10.1016/j.humimm.2010.02.018>.
- [80] Petukhova L, Duvic M, Hordinsky M, et al. Genome-wide association study in alopecia areata implicates both innate and adaptive immunity. *Nature.* 2010;466:113–117.
- [81] Jayne A, Diane M, Chris C, et al. Targeted genome-wide investigation identifies novel SNPs associated with diabetic nephropathy. 2010;77–82.
- [82] Kamei R, Yoshimura K, Yoshino S, et al. Expression levels of UL16 binding protein 1 and natural killer group 2 member D affect overall survival in patients with gastric cancer following gastrectomy. *Oncol. Lett.* 2018;15:747–754.

- [83]. Rölle A, Mousavi-Jazi M, Eriksson M, et al. Effects of Human Cytomegalovirus Infection on Ligands for the Activating NKG2D Receptor of NK Cells: Up-Regulation of UL16-Binding Protein (ULBP)1 and ULBP2 Is Counteracted by the Viral UL16 Protein. *J. Immunol.* 2003;171:902–908.

## CHAPTER 3

### 3.0 Computational Strategies Used in Drug Design and SNPinformatics approach for nsSNPs Identification

#### 3.1 Introduction

Over the past decades, the process of drug discovery, design, and development has gained admirable attention due to the cost implication, time, and manpower required of the traditional drug discovery pipeline [1]. It has been estimated that the traditional route of drug development cost around 1 billion dollars and it takes about 10-15 years [1,2]. These limitations necessitated the introduction of computer-aided drug design (CAAD). With this technique, the production cost of a new drug could be reduced by 50% [1,2].

Molecular modelling is one of the CADD methods used in drug design and discovery process. Other computational methods employed in the drug design and pipeline include molecular docking, virtual screening, homology modelling, molecular dynamic simulation, quantitative structure-activity relationship models (QSAR), etc. Molecular modelling details the generation, description, and engineering of the three-dimensional architecture of a compound and its corresponding physicochemical properties [3]. Molecular Modelling incorporates a wide range of computer-aided techniques and experimental data to speculate biological and molecular properties [3]. The concept of molecular modelling emanated from the postulation of molecular orbitals and classical mechanical programs [3]. Molecular mechanics (MM) and quantum mechanics (QM) are the two forms of molecular modelling theory used in elucidating the structural change and thermodynamics of a system.

### **3.2 Quantum Mechanics (QM)**

Classical mechanics originate from examination of matter at the macroscopic level. In classical mechanics, an atom has a designated location in space that can be identified. The description of the atom is only limited by the reliability of the instruments used for measurement [4]. The classical mechanical system can be solved using different techniques such as Newtonian and Hamiltonian mechanics. Due to the identification of particles that cannot be effectively described by classical mechanics, quantum mechanics was introduced. Quantum mechanics describe the study and characteristics of matter and its relationship with energy at the atomic and subatomic levels [4].

Quantum mechanics can be used to explain some biological processes, most especially during bond formation and breakage, atom transfer and electron excitation. QM can also be used to describe the behaviour of particles at the sub-atomic level in the 3D-space of a molecule. Electrons are portrayed with the aid of continuous electron density techniques and the thermodynamic characteristics of the system are computed using Schrödinger's wave function theory. For a system that cannot be described by the Schrödinger's wave function theory, such as large systems, the energetics and electron density can be computed with the aid of Born-Oppenheimer approximation theory [4].



### 3.2.1 Schrödinger's wave function theory

Erwin Schrödinger hypothesized the quantum mechanical model of an atom. Building upon the Bohr atomic model, which pontificate that electrons revolve around the orbits of a nucleus. Using a mathematical model, Schrödinger expresses the possibility of finding an electron on a particular path. QM describes particles as a wave function without a definite position or momentum. The possibility of each observation may be detected by wave function [4,5]. The Schrödinger equation is described below:

$$\boxed{H\Psi = E\Psi} \quad (\text{Eq 3.2.1})$$

Where H = Hamiltonian operator.

E = Energy eigenvalues of the system

$\Psi$  = wave function.

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

$$\boxed{\mathbf{H = \left[ -\frac{\hbar^2}{8\pi^2} \sum_i \frac{1}{mj} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \right] + \sum_i \sum_{<j} \left( \frac{e_i e_j}{r_{ij}} \right)}} \quad (\text{Eq 3.2.3})$$

### 3.2.2 The Born-Oppenheimer approximation

It possesses some advantages over Schrödinger's equation. In the description of characteristics of atoms in computational chemistry, it is paramount to have a clear cut distinction between the electronic and nuclear motions. The nuclei revolve slowly around the orbit, unlike the electrons which travel at the speed of light. Born-Oppenheimer approximation permits the Schrödinger equation to be distinctly expressed as an electronic and nuclear equation. The Hamiltonian operator of Born-Oppenheimer approximation is expressed as:

$$\mathbf{T}^{elect} = \left[ -\frac{\hbar^2}{8\pi^2 m} \sum_i^{electrons} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \right] \quad (\text{Eq 3.2.4})$$

### 3.2.3 Potential Energy Surface (PES)

PES is a mathematical description of the interconnection between the vibrational energy, geometry, and nuclear probability distribution of a molecule computed by deciphering the time-dependent Schrödinger equation. PES is borne out of the Born-Oppenheimer approximation. PES graphically describes the ratio of the potential energy to the geometry of a molecule. Mathematically, it can be said that the potential energy is directly proportional to the geometry[6,7].

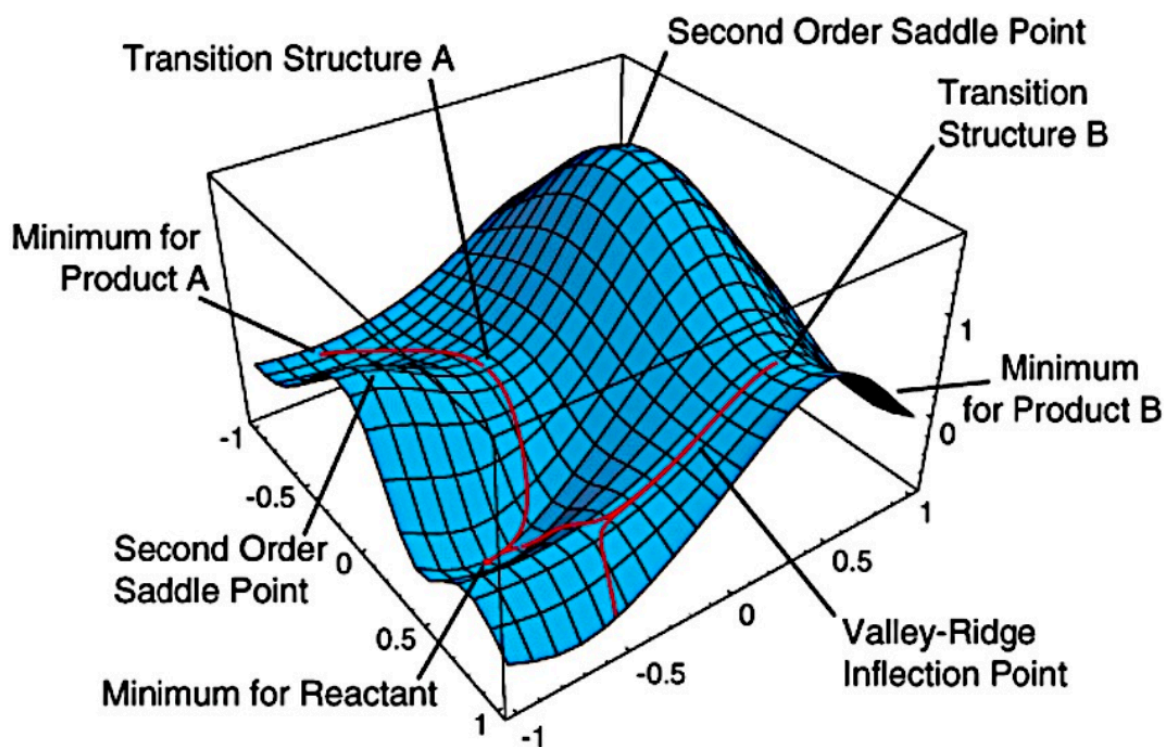


Figure 3.1: Descriptive illustration of 2D PES [7].

### 3.3 Molecular Mechanics (MM)

MM is an experimental method used in the computation of properties of molecules such as heat of formation, vibrational frequencies, dipole moment, strain energy, and molecular geometry [8].

MM or force field methods employ traditional models to project the kinetics of a molecule as a function of its structural configuration. This gives room for prediction of (1) relative energies between conformers (2). Equilibrium geometries and transition states [8]. MM has been successfully used in the description of complex structures especially those in the biological system. It engages the traditional Newtonian mechanics to detail large biological systems such as hydrocarbons, large molecular weight proteins, and membrane fragments [8].

In the studied system, the atomic nuclei are described as an isolated system while the electrons are omitted. Elimination of electrons and the treatment of nuclear and electronic motion as a different system is important as suggested by Born-Oppenheimer approximation. The differential energy of the systems is more critical than potential energy values. The total potential energy can be estimated by the equation below:

$$\mathbf{E}_{\text{tot}} = \mathbf{E}_{\text{str}} + \mathbf{E}_{\text{bend}} + \mathbf{E}_{\text{tor}} + \mathbf{E}_{\text{vdw}} + \mathbf{E}_{\text{elec}}. \quad (\text{Eq 3.2.4.1})$$

Where:

$E_{\text{tot}}$  = Total potential energy

$E_{\text{str}}$  = Bond-stretching energy

$E_{\text{bend}}$  = Bond angle-bending energy

$E_{\text{tor}}$  = Torsional energy

$E_{\text{vdw}}$  = van der Waals interactions

$E_{\text{elec}}$  = Electrostatic interactions

### *3.3.1 Force Field*

The force field is a mathematical expression that is used in the description of the relationship that exists between the coordinates of particles in a molecule and the energetics of the system. The most commonly used forcefields in molecular modelling simulation include AMBER [9], GROMOS [10], NAMD, and CHARMM [11]. However, they differ in the number of cross-terms, functional form and information for parameter fittings. Force fields used in simulating systems such as protein or DNA, use basic functional forms without cross-terms and use Lennard-Jones potential as van der Waals energy called harmonic/diagonal force fields. For the simulation of small or medium-size molecules, a high degree of accuracy is required, hence, they possess several cross-terms and exponential-type potential for van der Waals. They are known as class II force fields.

### **3.4 The Concept of Molecular Dynamic Simulation**

Since the introduction of Molecular Dynamic Simulation (MDS) in the 70s [12,13], it has moved beyond the mere simulation of several hundreds of atoms to biological systems such as proteins in solution, membrane-embedded proteins, nucleosomes, DNA, ribosomes [14,15]. Simulating a system having 50,000-100,000 atoms has become conventional, also, simulating systems with a larger amount of atoms like 500,000 has become possible with the right computer resources. This advancement is due to the simplification of MD calculations (Figure 3.2) and the introduction of high-performance computing (HPC).

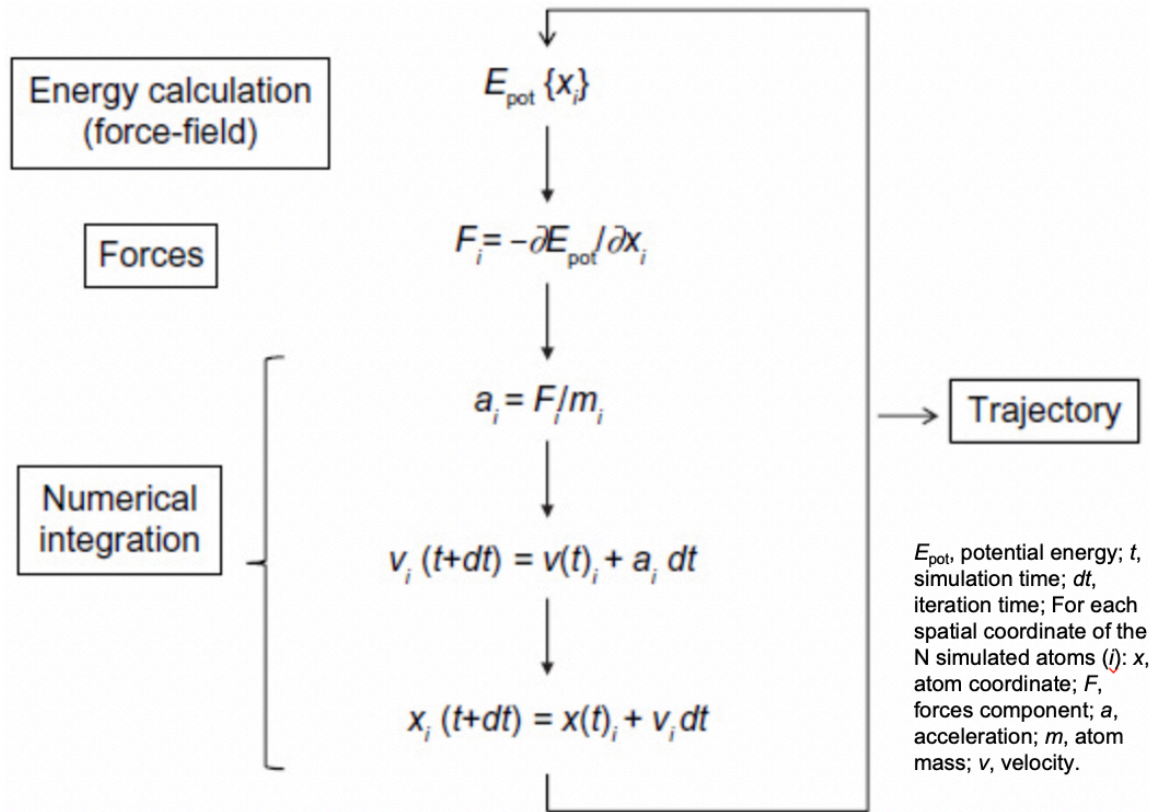


Figure 3.2: Molecular dynamic simulation basic algorithms and motion calculation [16].

The model used for simulation is gotten from structural modelling data or experimental structures such as those from X-ray crystallography and Nuclear Magnetic Resonance [17]. The simulated system can either be represented with atomistic representation or coarse-grained representation. The best production is obtained using atomistic representation while coarse-grained representation is employed when huge systems and extended simulations are necessary. Furthermore, Newton's equation of motion is used to generate the atomic trajectories of the different particles present in the system. It is mathematically expressed as:

$$\mathbf{F}_i = m_i \frac{d^2 \mathbf{r}_i(t)}{dt^2} \quad (\text{Eq 3.3.1})$$

Where  $r_i(t)$  is defined as the particle position vector,

$t$  = time-evolution

$m$  = mass of the particle

$F_i$  = Interacting force

### 3.4.1 Molecular Dynamics Trajectory Analysis

MD trajectories are generated during the simulation run. They are defined as the successive snapshots having velocity vectors and positional coordinates and time evolution of the system. The nature of the study determines the kind of trajectory analysis to be carried out. For this study, post-analysis carried out are tailored to the study scope.

The convergence of a system is used to describe the dynamics of a protein using bond angles gyration and bond types in the course of protein unfolding. The level of deviation of a system termed Root Mean Square Deviation (RMSD) can be determined by estimating the distance of the backbone atoms of superimposed systems. RMSD can also be used in estimating the distance of atomic backbone of other non-protein systems such as small organic compounds. RMSD is mathematically represented as:

$$\mathbf{RMSD} = \left( \frac{\sum_N (R_i - R_i^0)^2}{N} \right)^{\frac{1}{2}} \quad (\text{Eq 3.4.1})$$

Where:  $N$  = Number of atoms in the system

$R_i$  = Vector position of particle  $i$

$O$  = Initial conformation

The Radius of Gyration (RoG) is defined as the root mean square distance of the atoms relative to their centre of gravity. It can be used to determine the level of compactness of a protein or in the estimation of protein folding/unfolding dynamics. It is represented as :

$$\mathbf{r^2_{gyr}} = \frac{(\sum_{i=1}^n w_i (r_i - r^-)^2)}{\sum_{i=1}^n w_i} \quad (\text{Eq 3.4.2})$$

Where :

$r_i$  = position of  $i$ th atom  
 $r$  = centre of gravity of atom  $i$ .

Conformational characteristics of a system can be estimated with the aid of Root Mean Square Fluctuations (RMSF). RMSF is the estimation of the divergence of a particle in relation to its reference position over a period of time. It can also be used to determine the level of fluctuations of protein residues. Mathematically, it is expressed with the equation below:

$$\mathbf{sRMSF} = \frac{(\overline{RMSF_i - RMSF})}{\sigma(RMSF)} \quad (\text{Eq 3.4.3})$$

Where :  $RMSF_i$  = RMSF of the  $i$ th residue

$\overline{RMSF}$  = Average RMSD

Principal Component Analysis (PCA) is a standard mathematical parameter used in the evaluation of the correlations between a large data set. It can be used to determine the displacement of a loop or any region of interest in a protein. PCA application in MDS is referred to as essential dynamics because only important motional movement of a data is retrieved from the millions of structural snapshots. This structural motions are subsequently filtered from the highest to lowest fluctuations and represented graphically with the aid of the covariance matrix [18]. The newly generated data coordinates are described as principal components and are ordered along at least two principal components [18].

### 3.5 Binding Free Energy

The computation of the binding free energy between a protein and a ligand is crucial in drug design and development [19] as it can be used to unravel the mechanism of binding between the protein and a compound. The most widely used free energy calculation method is the Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PB-SA) and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) [20]. Other free energy calculation methods include thermodynamic integration [21], free energy perturbation [22], and replica exchange [23]. The estimation of binding free energy is dependent on the molecular dynamic simulation and the force field employed. Using an explicit force field produces precise binding affinities. The binding free energy between a ligand and receptor can be expressed with the equation below:

$$G_{bind} = G_{complex} - G_{receptor} - G_{ligand} \quad (\text{Eq 3.4.1.1})$$

$$\Delta G_{bind} = \Delta E_{gas} + \Delta G_{sol} - TS \quad (\text{Eq 3.4.1.2})$$

$$E_{gas} = E_{int} + E_{vdw} + E_{ele} \quad (\text{Eq 3.4.1.3})$$

$$G_{sol} = \Delta G_{GB/PB} + \Delta G_{SA} \quad (\text{Eq 3.4.1.4})$$

$$G_{SA} = \gamma SASA \quad (\text{Eq 3.4.1.5})$$

Where:  $E_{gas}$  = Gas phase energy

$E_{int}$  = Internal Energy

$E_{ele}$  = Coulomb energy

$E_{vdw}$  = van der Waals energies

$G_{sol}$  = Solvation free energy

SASA = Solvent accessible surface area

$G_{SA}$  = non-polar solvation energy

TS = Total entropy S of the solute at temperature T



### 3.6 Molecular Docking

Molecular docking is a method that explores the interactivity between a compound and a protein, it helps to unravel the interaction pattern of a compound in the active site of a protein [24]. There are two basic steps necessary to carry out molecular docking: assessment of the ligand configuration and orientation within the binding site (sampling methods) and the estimation of the binding score (scoring schemes). The success of any docking process is dependent on the availableness of a 3-dimensional structure either obtained from homology modelling, X-Ray crystallography or NMR. The location and knowledge of the binding site are also important. The binding site of a protein is usually known before docking, the binding cavity of a protein can as well be predicted by comparison with protein with similar structure and functions. In the absence of these, binding site determining tools such as POCKET [25], GRID [26], PASS [27], MMC [28], SurfNet [29], MetaPocket [30], SiteMap [31], etc can be engaged in characterizing the binding cavity. Blind docking is a docking process without the full knowledge of the binding site. When the protein and the compound are both handled as inflexible entities, it is referred to as rigid ligand and rigid receptor docking process. The search space only considers the three rotational and translational degrees of freedom. Software which adopt this method include DOCK [18], FLOG [32], and FTDOCK [33]. FlexX ([34] and AutoDock [35], use a docking technique where the compound is flexible and the receptor remains inflexible. The receptor and the ligand can be both flexible.

### 3.7 Virtual Screening

Virtual screening (VS) is the prediction of potentially bioactive compounds from a library of small molecules using computational tools [36]. VS has become a desired computational method in the field of drug design and development as various *in-silico* based strategies are constantly being developed, improved, and publicly available for use [36]. Being an easy-to-use technique, public

and private organisations employ VS methodologies to save resources and time [37]. VS is always carried out in a sequential workflow integrating various methodologies, which filters and discards unwanted compounds. Compounds that go through the filtering process are referred to as hit compounds and tested experimentally to evaluate their biological activities. VS methods are classified as ligand-based (leverages on similarities between the compound of interest with active compounds) and receptor based (uses complementarity between the compound of interest with the target protein binding site).

Pharmacophore based virtual screening is a new form of VS that employs pharmacophoric features of a compound's functional groups such as hydrogen bond acceptors, hydrogen bond donors, cations, hydrophobic areas, aromatics. These pharmacophoric moieties are then used to such a library of compounds to retrieve compounds with similar pharmacophoric properties.

### **3.8 Identification of Deleterious nsSNPs**

Structural and functional analysis is the end point of investigating nsSNPs. Knowledge of the architectural change within a protein can help provide insight into the pathogenesis of a disease and further identification of therapeutics [38]. The structure-function balance of a protein has been found to be disrupted as a result of nsSNPs [39].

Sorting Intolerant from Tolerant (SIFT) [40], Protein Variation Effect Analyzer (PROVEAN) [41], Polymorphism Phenotyping v2 (Polyphen2) [42], and PhD-SNP [43], can be used to determine deleterious nsSNPs.

## Reference

- [1] Kore PP, Mutha MM, Antre R V., et al. Computer-Aided Drug Design: An Innovative Tool for Modeling. *Open J. Med. Chem.* 2012;02:139–148.
- [2] Daina A, Blatter MC, Baillie Gerritsen V, et al. Drug Design Workshop: A Web-Based Educational Tool to Introduce Computer-Aided Drug Design to the General Public. *J. Chem. Educ.* 2017;94:335–344.
- [3] Nadendla RR. Molecular modeling: A powerful tool for drug design and molecular docking. *Resonance.* 2004;9:51–60.
- [4] Lin S-K. *Fundamentals of Quantum Chemistry: Molecular Spectroscopy and Modern Electronic Structure Computations.* Int. J. Mol. Sci. 2001.
- [5] Schrödinger E. An Undulatory Theory of the Mechanics of Atoms and Molecules. *Phys. Rev.* 28, 1049.
- [6] Jain M. Fundamental forces and laws: a brief review. *Textbook of Engineering Physics, Part 1.* PHI Learning Pvt. Ltd. p. 10. ISBN 978-81-203-3862-3.
- [7] Lewars EG. *Computational chemistry: Introduction to the theory and applications of molecular and quantum mechanics.* *Comput. Chem. Introd. to Theory Appl. Mol. Quantum Mech.* 2011.
- [8] de Visser S, Hofer T. Quantum Mechanical/Molecular Mechanical Approaches for the Investigation of Chemical Systems-Recent Developments and advanced Applications. *Front. Chem.* [Editorial]. 2018 13 september.
- [9] Case R. AMBER. Amber 2019\_Reference manual. *J. Chem. Inf. Model.* 2013;53:1689–1699. Available from: <http://ambermd.org/contributors.html>.
- [10] Horta BAC, Merz PT, Fuchs PFJ, et al. A GROMOS-Compatible Force Field for Small

- Organic Molecules in the Condensed Phase: The 2016H66 Parameter Set. *J. Chem. Theory Comput.* 2016;12:3825–3850.
- [11] Allouche A. Software News and Updates Gabedit — A Graphical User Interface for Computational Chemistry Softwares. *J. Comput. Chem.* 2012;32:174–182.
- [12] Zhang H, Duan L, Guo X, Cong, Y, Feng G, Yuchen L, Zhang J. Accelerated Molecular Dynamics Simulation for Helical Proteins Folding in Explicit Water. *front. Chem.*, 2019;7:540
- [13] Warshel A, Levitt M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* 1976;103:227–249.
- [14] Kobori T, Iwamoto S, Takeyasu K, et al. Biopolymers Volume 85 / Number 4 295. *Biopolymers.* 2007;85:392–406.
- [15] Brandman R, Brandman Y, Pande VS. A-site residues move independently from P-site residues in all-atom molecular dynamics simulations of the 70S bacterial ribosome. *PLoS One.* 2012;7:1–8.
- [16] Hospital A, Goñi JR, Orozco M, et al. Molecular dynamics simulations: Advances and applications. *Adv. Appl. Bioinforma. Chem.* 2015;8:37–47.
- [17] Berendsen HJC, Postma JPM, Van Gunsteren WF, et al. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 1984;81:3684–3690.
- [18] Tan Y, Hanson JA, Chu J, et al. VIP MSM UBER WIP PCA ESSENTIAL Protein Dynamics. *Proteins.* 2014. Available from: <http://link.springer.com/10.1007/978-1-62703-658-0>.
- [19] Kalra P, Reddy T V., Jayaram B. Free energy component analysis for drug design: A case

- study of HIV-1 protease-inhibitor binding. *J. Med. Chem.* 2001;44:4325–4338.
- [20] Hou T, Wang J, Li Y, et al. Assessing the Performance of the MM\_PBSA and MM\_GBSA Methods. 1. The Accuracy.pdf. *J. Chem. Inf. Model.* 2011;51:69–82.
- [21] Beale PD. Acoustic crystal thermodynamic integration method. *Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.* 2002;66:1–7.
- [22] Park H, Jeon YH. Free energy perturbation approach for the rational engineering of the antibody for human hepatitis B virus. *J. Mol. Graph. Model.* 2011;29:643–649.
- [23] Woods CJ, King MA, Essex JW. Replica-Exchange-Based Free-Energy Methods. *Lect. Notes Comput. Sci. Eng.* 2006;49:252–259.
- [24] Cui M, Xuan-Yu M, Hong-Xing Z, Mezei M. Molecular Docking: A powerful approach for structure-based drug discovery. *Curr comput Aided Drug Des. Bone.* 2011;23:1–7(2).
- [25] DG. L, LJ. B. POCKET: a computer graphics method for identifying and displaying protein cavities and their surrounding amino acids. *J. Mol. Graph.* 1992;10:229–234.
- [26] Kastenholtz MA, Pastor M, Cruciani G, et al. GRID/CPCA: A new computational tool to design selective ligands. *J. Med. Chem.* 2000;43:3033–3044.
- [27] Brady GP, Stouten PFW. Fast prediction and visualization of protein binding pockets with PASS. *J. Comput. Aided. Mol. Des.* 2000;14:383–401.
- [28] Mezei M. A new method for mapping macromolecular topography. *J. Mol. Graph. Model.* 2003;21:463–472.
- [29] Laskowski RA. SURFNET: A program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J. Mol. Graph.* 1995;13:323–330.
- [30] Huang B. Metapocket: A meta approach to improve protein ligand binding site prediction. *Omi. A J. Integr. Biol.* 2009;13:325–330.

- [31] Halgren TA. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* 2009;
- [32] Miller MD, Kearsley SK, Underwood DJ, et al. FLOG: A system to select “quasi-flexible” ligands complementary to a receptor of known three-dimensional structure. *J. Comput. Aided. Mol. Des.* 1994;8:153–174.
- [33] Gabb HA, Jackson RM, Sternberg MJE. Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J. Mol. Biol.* 1997;272:106–120.
- [34] Rarey M, Kramer B, Lengauer T, et al. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* 1996;261:470–489.
- [35] Morris GM, Goodsell DS, Pique ME, et al. AutoDock Version 4.2. User Guid. 2009;1–49.
- [36] Gimeno A, Ojeda-Montes MJ, Tomás-Hernández S, et al. The light and dark sides of virtual screening: What is there to know? *Int. J. Mol. Sci.* 2019;20.
- [37] Shoichet BK. Virtual screening of chemical libraries Problems with virtual screening. *Nature.* 2006;432:862–865.
- [38] Soremekun OS, Soliman MES. From genomic variation to protein aberration: Mutational analysis of single nucleotide polymorphism present in ULBP6 gene and implication in immune response. *Comput. Biol. Med.* 2019;111:103354.
- [39] Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet.* 2013.
- [40] Vaser R, Adusumalli S, Leng SN, et al. protocol UPDATE SIFT missense predictions for genomes. *Nat. Protoc.* 2015;11:1–9.
- [41] Choi Y, Chan AP. PROVEAN web server: A tool to predict the functional effect of amino

- acid substitutions and indels. *Bioinformatics*. 2015;31:2745–2747.
- [42] Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7:248–249.
- [43] Capriotti E, Fariselli P. PhD-SNPg: A webserver and lightweight tool for scoring single nucleotide variants. *Nucleic Acids Res*. 2017;45:W247–W252.

## CHAPTER 4

### **Recruiting monomer for dimer formation: Resolving the antagonistic mechanisms of novel immune check point inhibitors against Programmed Death Ligand-1 in cancer immunotherapy**

Opeyemi S. Soremekun<sup>a</sup>, Fisayo A. Olotu<sup>a</sup>, Clement Agoni<sup>a</sup>, Mahmoud E. S. Soliman<sup>a\*</sup>

<sup>a</sup>Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

\*Corresponding Author: Mahmoud E.S. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872



## **Abstract**

The design of small molecule antagonists against Programmed Death Ligand-1 (PD-L1) has been the recent highlight of the immune checkpoint blockade therapy. This interventive approach has been potentiated by the development of BMS compounds; BMS-1001 and BMS-1166, which exert their therapeutic activity by inducing dimerization of PD-L1; a molecular mechanism that has remained unclear. For the first time, we resolve the dynamical events that underlie the antagonistic mechanisms of BMS-1001 and BMS-1166 when bound to PD-L1. Our findings revealed that upon binding a PD-L1 monomer, the BMS-compounds gradually facilitated the ‘inbound’ motion of another PD-L1 monomer in the same conformational phase space up till dimer formation. Moreover, the non-liganded PD-L1 monomer exhibited the highest structural flexibility and atomistic motions relative to the BMS-liganded monomer. Interestingly, the BMS compounds exhibited mechanistic transitions from the monomeric binding site (monomer A) where they were initially bound, to the second monomeric site (monomer B) where they were strongly bound, followed by eventual high-affinity interactions at the tunnel-like binding cleft formed upon the dimerization of both PD-L1 monomers. These findings present a model that describes the mechanism by which the BMS compounds induce PD-L1 dimerization and could further enhance the design of highly selective and novel monomeric recruiters of PD-L1 in cancer immunotherapy.

**Keywords:** Cancer Immunotherapy, Immune Checkpoint, Programmed Death Ligand-1, Binding cleft, Molecular Dynamic Simulation

## 4.1 Introduction

Over the years, traditional anticancer therapeutic strategies have been categorized under three pillars namely; surgery, radiation and chemotherapy. Although these therapeutic approaches have offered substantial interventions in the eradication primary tumor the problem of relapse still persist, which result from residual malignant cells and/or tumor metastases [1][2]. This necessitated a new approach regarded as cancer immunotherapy, which provides potent and alternative treatment options [3] that involve the utilization of naturally derived or synthetically generated compounds to stimulate or enhance the immune system to fight cancer [4]. Hence, cancer immunotherapy was described by science magazine as “2013’s breakthrough of the year” due to the recorded success in cancer treatment [5]. The basic role of the immune system is to protect the body against foreign pathogens and infections, and more importantly, the humoral immunity employs antibodies produced by the B cells to neutralize and eradicate foreign pathogens and toxins [6] while cellular immunity responds through recognition of antigens, activation of antigen presenting cells (APCs), activation and proliferation of T cells [7]. Cancer immunotherapy has shown its potency among numerous therapeutic options [1], with the incorporation of cancer vaccinations, chimeric antigen receptor (CAR) T cell therapy and immune checkpoint blockade (ICB) therapy [8]. Immune checkpoint inhibitors are a class of drugs that have been designed to increase immune response against cancer cells [9]. Moreover, in the immune system, there are different checkpoints pathways that focus on the activation of T cells, [1] and are crucially regulated by biomolecules such as CTLA-4, PD-L1, lymphocyte activation gene-3 (LAG-3) and T cell immunoglobulin and mucin domain containing protein 3 (Tim-3) [10]. These, over the years, have become attractive targets for driving the cytotoxic T cells attack of malignant cells for destruction due to the hypo-responsive and exhaustive influence they have on the immune system via regulation [11], [12]. According to the classic checkpoint hypothesis, immune checkpoint

inhibitors activate the immune system by blocking the expression of co-inhibitory molecules such as CTLA4, PD-1 and PD-L1 which in turn re-activate the pre-existing tumor response [9], [13]. As previously reported, antagonizing the interaction between PD-1 and PD-L1 reverts T cell phenotypic exhaustion, which in turn leads to the efficient killing of cancer cells [14]. In recent years, antibodies were widely employed to antagonize the interactions between PD-1 and PD-L1, a biological association that promotes immunosuppressiveness, which is employed by cancer cells to evade T-cell induced death [15], [16]. However, due to lower production costs, higher stability, improved tumor penetration, amenability for oral administration and elimination of immunogenicity, small-molecular weight inhibitors present a more viable option in place of antibodies as immune check point inhibitors [17]. Most recently, two small molecule antagonists (BMS-1001 and BMS-1166) that directly and selectively disrupt the association between PD-1 and PD-L1 have been identified and patented by Bristol-Myers Squibb (BMS) [18].

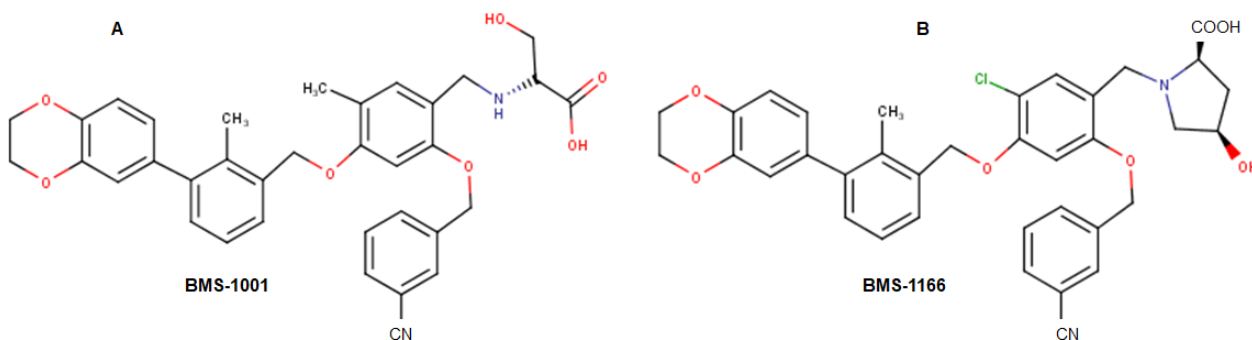


Figure 4.1: 2D structure of soluble immune checkpoint inhibitors; [A] BMS-1001 and [B] BMS-1166.

According to previous reports, these BMS compounds bind to and induce PD-L1 dimerization thereby preventing canonical interactions with PD-1. Mechanistically, these compounds bind to a PD-L1 monomer (<sub>A</sub>PD-L1) and facilitate the recruitment of another PD-L1 monomer (<sub>B</sub>PD-L1) to form a <sub>AB</sub>PD-L1 thereby blocking interactions with the naturally occurring PD1. In other words, the ability of these compounds to induce the formation of PD-L1 dimers underlies its antagonistic

functions since PD-L1 dimerization presents a hurdle for the canonical interactions between PD-L1 monomers and its intrinsic ligand, PD-1, resulting in the activation of T cells [19], [20].

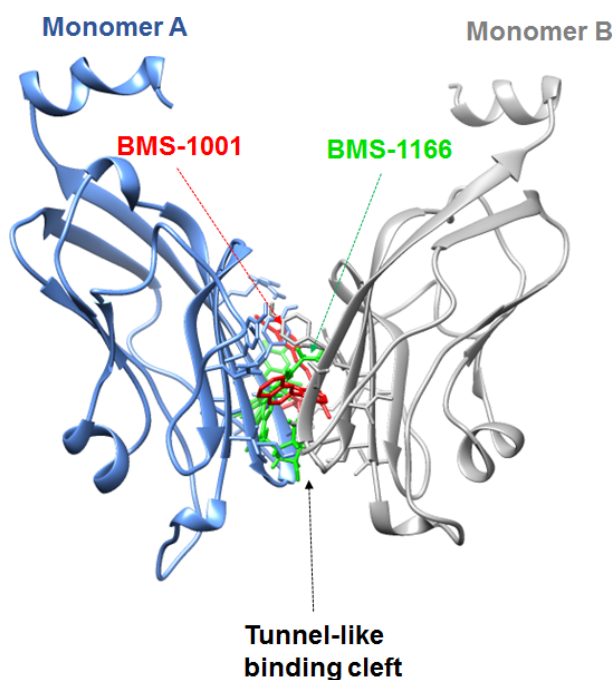


Figure 4.2: 3D structure of PD-L1 monomer A (Blue) and monomer B (Gray), in complex with BMS-1001 (Red) and BMS-1166 (Green) at the tunnel-like binding cleft formed upon dimerization. PDB: 5NIX [19]. BMS-1001 was superimposed on the crystal structure.

Therefore, the development and use of these novel small molecule inhibitors in ICB therapy present an attractive prospect to understand the underlying structural and molecular events involved in the blockage of PD-1 and PD-L1. Molecular elucidation of the transient mechanisms by which BMS compounds (BMS-1001 and BMS-1166) induce PD-L1 dimerization following an initial interaction with a PD-L1 monomer would provide a model that would be helpful to achieve the structure-based design of monomeric binders. Therefore, in this study, we employed a computational paradigm to investigate the structural and molecular events associated with PD-L1 dimerization as induced by antagonistic BMS compounds; BMS-1001 and BMS-1166. We believe this model would provide essential insights and interesting possibilities towards the future design

of novel high-affinity small-molecule antagonists and monomeric recruiters of PD-L1 in ICB therapy.

## 4.2 Computational Methods

### 4.2.1. System Preparation

X-ray crystallographic structures of dimerized PD-L1s in complex with BMS-1001 and BMS-1166 were retrieved from Protein Data Bank with entries 5NIU and 5NIX respectively [19]. These were prepared by the removal of co-crystallized molecules such as 1,2-ethanediol (EDO) and crystallographic water using UCSF Chimera [21]. Moreover, missing residues which majorly constituted terminal amino acids were added with the aid of MODELLER, an *in-silico* tool used for structural remodeling [22]. In order to elucidate the mechanisms of BMS-1001/BMS-1166-induced dimerization, which was in line with the aim of this study, it was important to obtain a starting structure that could possibly reflect the recruitment of a PD-L1 monomer in the presence of another BMS-bound PD-L1 monomer as experimentally depicted in a previous study [19]. To this effect, the ligand-dimer complexes were first separated into non-complexed monomers using the Graphical User Interface (GUI) of UCSF Chimera. The <sub>B</sub>PD-L1 was first removed from the dimerized complex (5NIU) leaving a BMS-1001-bound <sub>A</sub>PD-L1. This was followed by the molecular protein-protein repositioning of <sub>B</sub>PD-L1 using UCSF-Chimera integrated Autodock Vina interface on manually defined co-ordinates which differs from the crystallized ‘dimer’ structure. The same paradigm was employed for preparing BMS-1166 bound systems (5NIX) to generate the desirable starting structures which entail ligand-bound <sub>A</sub>PD-L1s with a distant <sub>B</sub>PD-L1 monomer in the same conformational space phase as shown in Figure 4.3. In the same vein, a replica system devoid of the small molecule antagonists (BMS-1001/BMS-1166) was prepared,

representative of the unbound form. This was necessary to clearly define the dimerizing effects of both BMS compounds, which underlies their antagonistic functions. Altogether, these systems were set up for an MD simulation of 150ns using in-house protocols previously reported [23]–[25]. The Graphical Processor Unit (GPU) version (Particle Mesh Ewald Molecular Dynamics - PMEMD) of the AMBER14 software in addition to its integrated modules were used to perform MD simulations [26]. The ANTECHAMBER module was used to parameterize the ligands and generate atomic partial charges (Gasteiger - gaff) using the bcc charge scheme while the FF14SB force-field was used to define the PD-L1 parameters [27]. Protein modification, renaming and protonation (histidine) were done using the in-house *pdb4amber* script. This was followed by the generation of topology and parameter files for the respective PD-L1-BMS1001/BMS1166 complexes using the LEAP module. This was also used for system neutralization via the addition of counter ions at a constant pH (cpH) and solvation within a TIP3P water box of 10Å. Using a restraint potential of 500kcal/mol Å, partial minimization of 2500 steps was carried out followed by 5000 full minimization steps without energy restraints. Gradual system thermalization from 0-300k was also carried out in a canonical ensemble (NVT) for 50ps with the aid of a Langevin thermostat and a 5kcal/mol Å<sup>2</sup> harmonic restraint. System equilibration was then carried out for 1000ps at 300k without restraints while atmospheric pressure was maintained at 1bar using the Berendsen barostat [28]. This was followed by an MD production run of 150ns where all atomic hydrogen bonds were constricted using the SHAKE algorithm [29]. Moreover, at every 1ps, resulting coordinates and trajectories were saved, which were analyzed with the integrated CPPTRAJ and PTRAJ modules [30]. Furthermore, data plots were made with Origin data analytical tool [31] while the Graphical User Interface (GUI) of UCSF Chimera was used for visual analysis.

#### 4.2.2. Binding free energy estimation

The Molecular Mechanics/ Poisson-Boltzmann Surface Area (MM/PBSA) method was used to estimate the differential binding of BMS-1001 and BMS-1166 to PD-L1 [32]. This is an end-point energy calculation that can be used to predict binding strength and affinities of small molecules towards their respective target proteins. MM/PBSA is mathematically represented as follows;

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

$$(2) \quad E_{\text{gas}} = E_{\text{int}} + E_{\text{vdw}} + E_{\text{ele}}$$

$$(3) \quad G_{\text{sol}} = G_{\text{GB}} + G_{\text{SA}}$$

$$(4) \quad G_{\text{SA}} = \gamma \text{SASA}$$

Where  $E_{\text{gas}}$ , the gas-phase energy,  $E_{\text{int}}$  is the internal energy;  $E_{\text{ele}}$  and  $E_{\text{vdw}}$  are the Coulomb and van der Waals energies, respectively. Also,  $G_{\text{sol}}$  is the solvation free energy and  $G_{\text{GB}}$  is the polar solvation contribution.  $G_{\text{SA}}$  is the non-polar solvation contribution and was estimated by the solvent accessible surface area (SASA) determined using a water probe radius of 1.4 Å. The surface tension constant  $\gamma$  was set to 0.0072 kcal/(mol·Å<sup>2</sup>).

#### 4.2.3. Dynamic cross-correlation matrix (DCCM)

This is graphically employed to represent dynamical motions of constituent protein residues in correlation with time [33], [34]. Therefore, in this study, we investigated the inter-residue dynamics and motions of PD-L1 when bound by BMS-1001 and BMS-1166 with respect to their antagonistic activities. We generated the DCCM to estimate the cross-correlated displacement of protein backbone atoms from the resulting trajectories across the MD simulation time based on the equation below:

$$C_{ij} = \langle \Delta r_i * \Delta r_j \rangle / (\langle \Delta r_i^2 \rangle \langle \Delta r_j^2 \rangle)^{1/2}$$

From the equation,  $\Delta r_i$  and  $\Delta r_j$  depict the mean displacement the  $i$ th and  $j$ th atoms which are represented by  $i$  and  $j$  respectively.  $C_{ij}$  is the coefficient of cross-correlation which varies from -1 to +1, which corresponds to strongly correlated (+) and anti-correlated (-) motions of constituent residues during the period of MD simulation. This analysis was performed using the integrated CPPTRAJ module [30] of AMBER14 and resulting matrices were plotted with the aid of the Origin data analysis software [31].

#### 4.2.4. *Principal component analysis (PCA)*

This was carried out to obtain necessary insights into the 3D conformational and dynamical changes that occurred in PD-L1 when distinctively bound by both BMS-1001 and BMS-1166. This method is used to describe the magnitude and direction of protein motions which are depicted by eigenvalues and eigenvectors respectively [25], [35]. Hence, with the aid of the integrated CPPTRAJ module in AMBER14, we computed the first two principal components (PC1 and PC2) from the dynamics of the protein C- $\alpha$  atoms. Moreover, the conformational behaviors of the unbound and bound PD-L1 systems were projected along the first two eigenvectors (ev1/PC1 vs ev2/PC2) using the Cartesian coordinates of C- $\alpha$  atoms.



### 4.3. Results and Discussion

#### *4.3.1 Monomeric binding of BMS compounds (BMS-1001 and BMS-1166) induced systematic PD-L1 dimerization*

The mechanisms by which the selective ‘monomeric’ targeting of PD-L1 by BMS compounds cause structural dimerization has remained unresolved till date even though it underlies the basis of ICB therapy. According to previous reports, PD-L1 exist in a single monomeric state where it is able to exhibit canonical interactions with PD-1 resulting in cancer immunosuppression and evasion of T-cell induced death [15], [16]. Moreover, BMS compounds reportedly binds initially to a PD-L1 monomer ( ${}_A$ PD-L1) and elicit monomeric recruitment ( ${}_B$ PD-L1) to form a dimer with a tunnel-like binding cleft in between [19]. Firstly, we visually examined the structural events that occurred along the MD trajectories to monitor the dynamics of monomeric recruitment with regards to the binding of the BMS compounds.

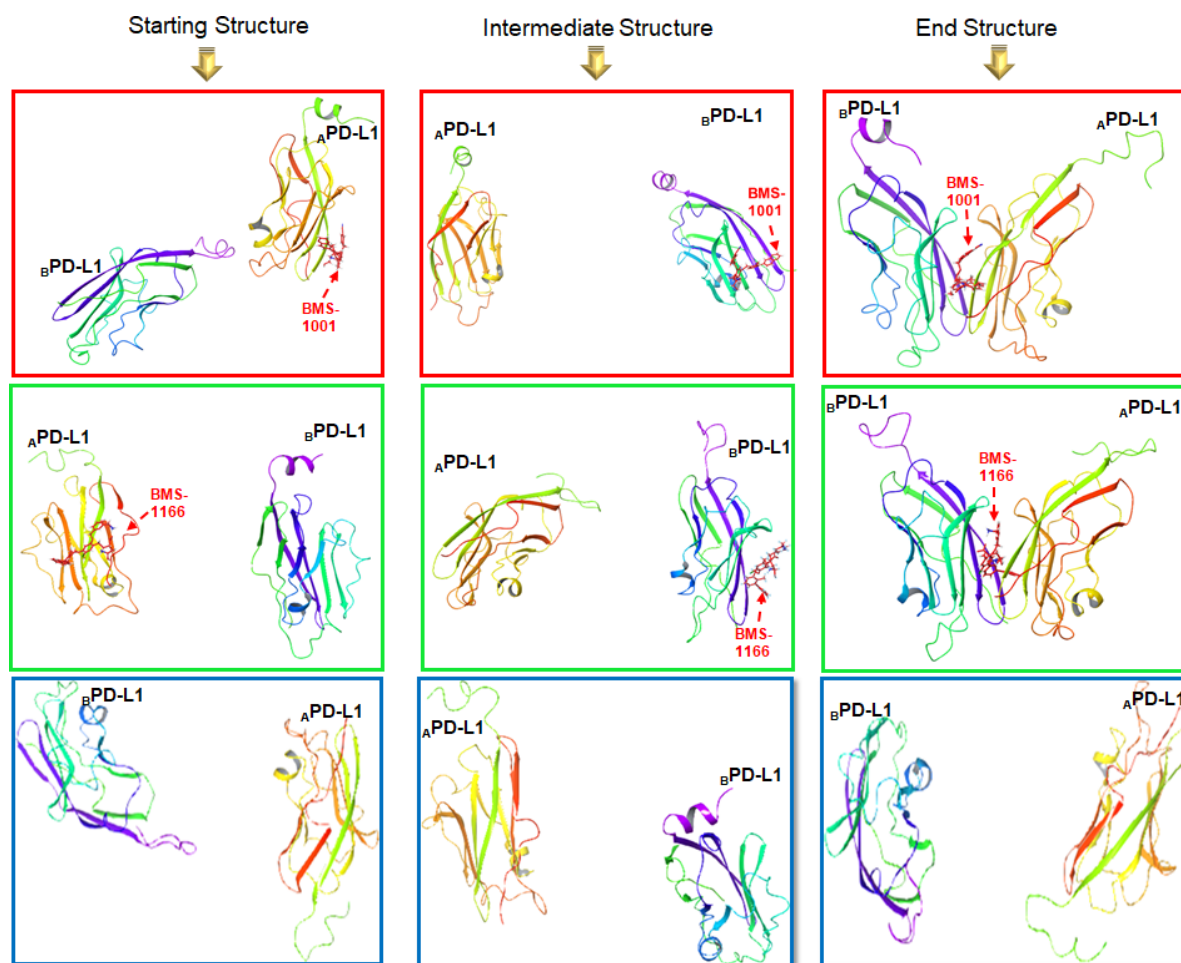


Figure 4.3: Structural dynamics of BMS-1001/BMS1166 bound PD-L1s and unbound PD-L1 over the 150ns MD simulation period. Red boxes depict the monomeric motions of BMS-1001-bound PD-L1 while green boxes depict the dynamics of the BMS-1166-bound PD-L1 system. However, blue boxes depict the unbound PD-L1 system. [A] Starting structure averaged over 0-20ns [B] averaged intermediate structure 60-90ns and [C] averaged end structure 90-150ns.

Interestingly, in the BMS-bound systems, we observed that the arrangements of the starting structures were maintained until about 20ns where monomeric distances gradually reduced, and the monomers existed in close proximities. These dynamical monomeric motions in the conformational phase space then proceeded until 90ns where we observed that both monomers associated to form a PD-L1 dimer and a tunnel-like cleft at the region where the BMS compounds were bound (Figure 4.3). This was evidenced by the gradual decrease in monomeric distances as

shown in Figure 4.3, which is derived from measuring distances between  $^A$ Tyr39 and  $^B$ Val185. As estimated, in the starting structure, positional distances between the two monomers ( $^A$ PDL1 and  $^B$ PDL1) were 76.79Å and 72.83Å for the BMS-1001- and BMS-1166-bound systems respectively, which gradually decreased in the average intermediate structure to 65.14Å (BMS-1001) and 69.23Å (BMS-1166) prior to eventual dimerization. On the contrary, while these events gradually occurred in the BMS-bound systems, the unbound PD-L1 system maintained its monomeric arrangement over the simulation period indicating that the PD-L1 dimerization is solely based on the presence of the BMS compounds (Figure 4.3).

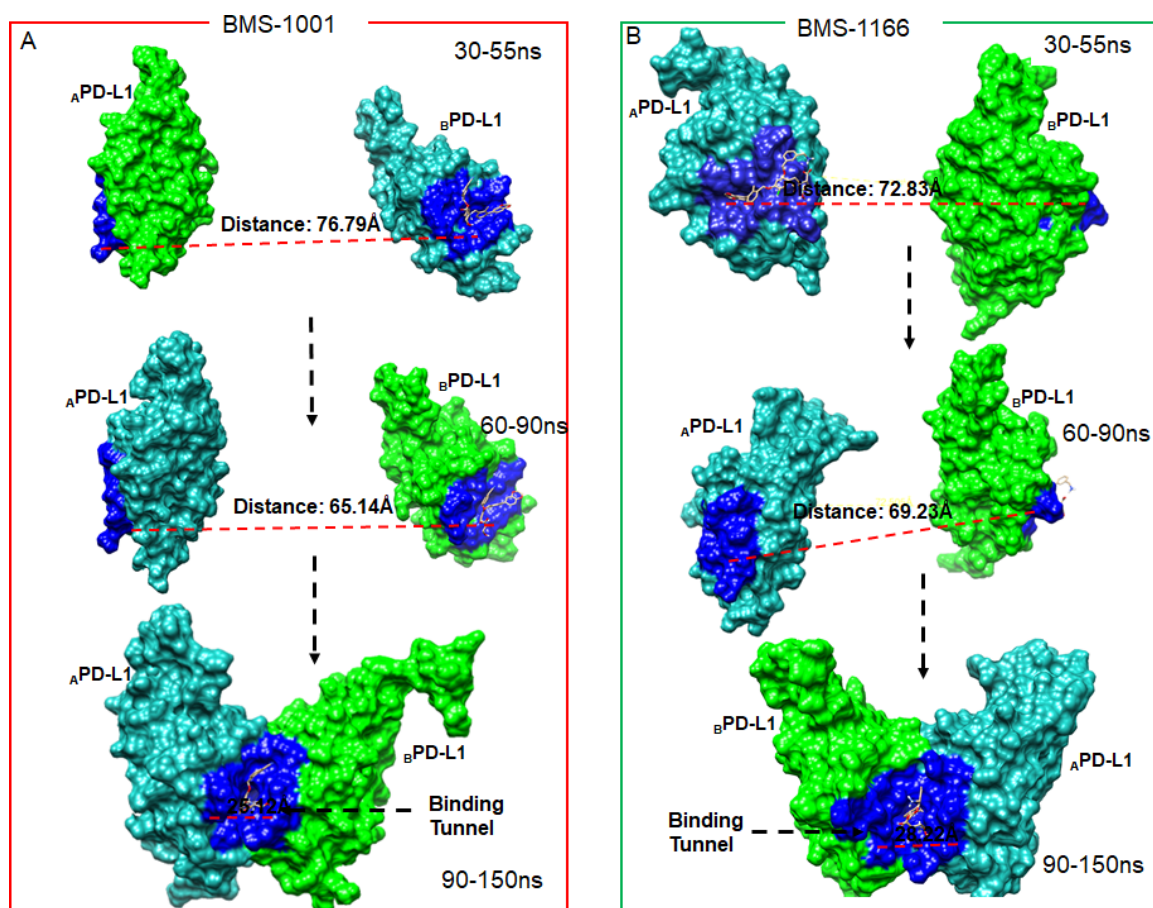


Figure 4.4: Monomeric distances, tunnel formation and mechanistic transition of BMS-1001 and BMS-1166 across both A and B PD-L1 monomers. [A] Tunnel formation and monomeric distance of BMS-1001 bound PD-L1 [B] Tunnel formation and monomeric distance of BMS-1166 bound PD-L1.

Another interesting finding was the concurrent dynamical ‘switch’ or exchange of both BMS compounds at each monomeric binding sites. At trajectory 30-55ns, averaged structure revealed that BMS-1001 was bound to selected site residues such as Met224, Val185, Ile163 and Val177 on monomer A via weak interactions (Figure 4.5A). Presumably, the nature of these interactions permitted a possible transition to monomer B which occurred at about 60-90ns and characterized by the occurrence of high-affinity interactions, which include the strong hydrogen bonds and attractive salt bridge with Arg108 and Asp105 respectively (Figure 4.5B). Relative to BMS-1166, a similar ligand transition pattern between both <sub>A</sub>PD-L1 and <sub>B</sub>PD-L1 monomers was observed along the trajectories as revealed by the averaged structures; where Thr3, Phe2, Lys107 contributed to the binding of BMS-1166 via high-affinity and strong hydrogen interactions while Met98 formed attractive salt-bridge as shown in (Figure 4.5D). Moreover, the mechanistic transitions of both BMS compounds occurred prior to the eventual formation of the tunnel-like binding cleft upon PD-L1 dimerization. We could propose that the dynamical motions of the PD-L1 monomers was accompanied by a concurrent binding site interchange of the BMS compounds prior to dimerization.

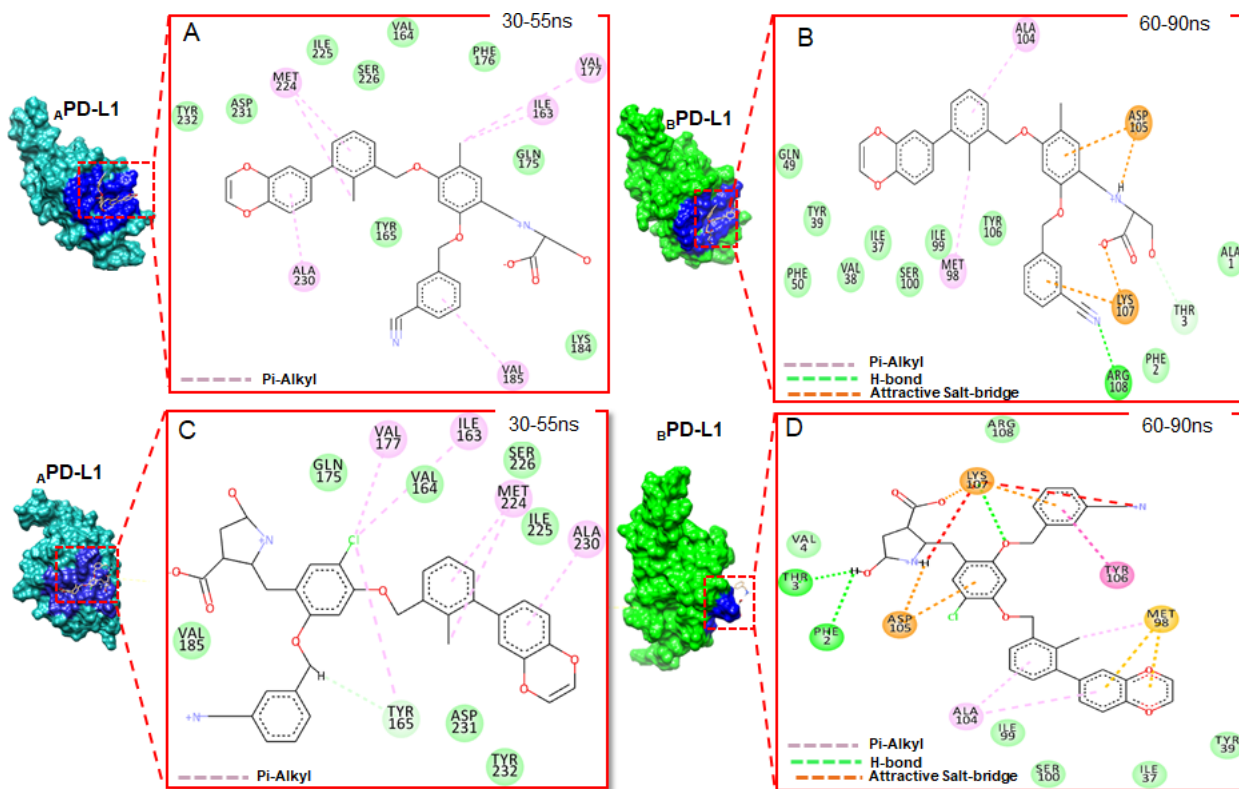


Figure 4.5: Visual representation of the monomeric active site residues interaction of BMS-1001 and BMS-1166 at 30-55ns (A, C) and 60-90ns (B, D).

### 4.3.2 BMS compounds induced high conformational flexibility favorable for monomeric motion and dimerization

#### 4.3.2.1. Conformational stability and residual fluctuation

We proceeded to measure the dynamical events that accompany the formation of  $_{AB}$ PD-L1 dimer as induced by the monomeric-binding of BMS-1001 and BMS-1166 respectively. We believe these insights would provide structural details that characterize the monomeric recruitment of PD-L1 in the presence of the BMS compounds and mechanistic details into their motions in the conformational phase space. To understand these distinct dynamical motions in relation to the binding of these antagonists, we employed metrics such as the root mean square deviation (RMSD),  $C\alpha$  root mean square fluctuation (RMSF) and radius of gyration (RoG), which are

appropriate to determine C $\alpha$  atomistic deviations and structural stability with the target protein over the MD simulation time.

Structural stability was first investigated among the unbound and BMS-bound systems using the C- $\alpha$  RMSD parameter in order to ensure that systems were appropriately equilibrated. Our findings showed that the systems converged and attained stability early in the production run (~10ns) until about 100ns where there was a distinct separation in atomistic motions among the three systems until the end of the simulation period (Figure 4.6). Interestingly, in agreement with the results earlier presented (Section 3.1), this dynamic event correlates with the dynamic of the molecules. In other words, the association of the PD-L1 monomers in both BMS-1001 and BMS-1166-bound systems involved considerable conformational alterations and instability, a structural attribute that characterize biological protein-protein interactions. These could further indicate that the binding of the BMS compounds played a crucial role in facilitating the interactions between both monomers with respect to PD-1 blockade.

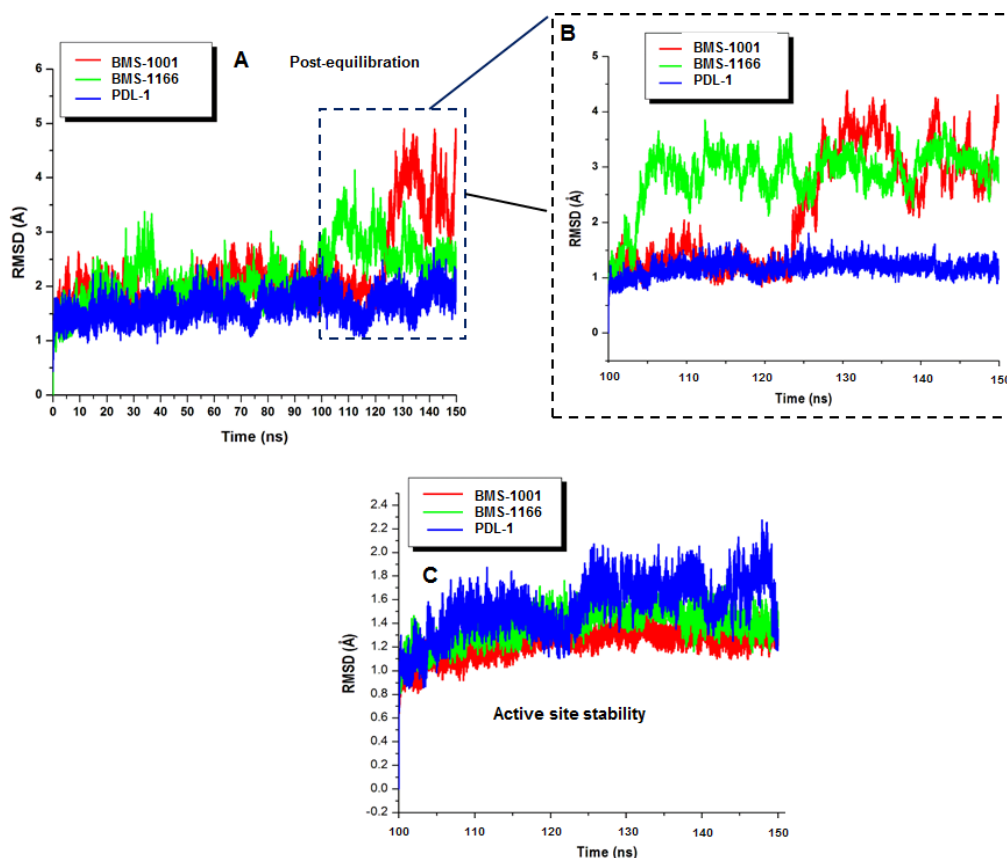


Figure 4.6: Conformational analysis plot showing stability and atomistic motions among unbound (**blue**), BMS-1001 (**red**) and BMS-1166 (**green**) bound systems. [A] Whole-structure comparative C- $\alpha$  RMSD plot of the three systems. **Blue** highlight depicts post-equilibration periods (100-150ns) where distinct separation in motions occurred [B] Comparative C- $\alpha$  RMSD plot (post-equilibration  $\rightarrow$  100-150ns) showing distinct variations in structural stability among the three systems [C] C- $\alpha$  RMSD plot showing active site stability among the unbound (blue), BMS-1001- (**red**) and BMS-1166- (**green**) bound systems

According to the plot, the unbound system maintained a stable structure throughout the simulation time (RMSD < 2Å) while the binding of BMS-1001 caused a notable deviation in PDL-1 structure (RMSD  $\sim$ 5.5Å). Likewise, atomistic deviation (RMSD) of  $\sim$ 4Å characterized the overall structure of PD-L1 when bound by BMS-1166. Taken together, we can suggest that the binding of these compounds induced notable instability in the structure of PD-L1 relative to inhibition. On the average, unbound PDL-1 system had RMSD value of 1.66Å while BMS-1001- and BMS-1166-bound systems had mean RMSD values of 2.28Å and 2.21Å.

In order to accurately measure the degree of structural instability induced by both PD-L1 antagonists, we carried out time-specific investigations into the occurrences from ~100ns where distinct separation in motion occurred among the three systems. This is clearly presented in Figure 4.6B, and accordingly, structural instability was highly induced in PD-L1 when bound by both BMS-1001 and BMS-1166, which is highly favourable for structural dimerization. In addition, we evaluated the distinct motions of the BMS compounds over the simulation period in relation to their roles in dimerization from the separate monomeric PDLs, most importantly from the period at which we presumed dimer formation (~100ns → ~150ns). As estimated, relative to their respective positioning and motions at the tunnel-like binding cleft upon PD-L1 dimerization, BMS-1001 has a RMSD value of 2.28Å while BMS-1166 exhibited a deviation of 2.21Å. However, the disparate motions exhibited by both compounds lasted until ~140ns where they both attained stability in dynamical motion (Figure 4.6A). Therefore, from this plot we can deduce that the deviations in motions early in the simulation period could be related to their ‘monomeric interchange’ as earlier mentioned earlier, which was later stabilized by the formation of the binding tunnel and corresponding interactions with constituent residues. Moreover, the dynamics of the PD-L1 binding sites were evaluated from ~100ns in relation to the non-active sites region. Our findings showed that while the non-active site regions exhibited high structural instability favorable for dimerization, the BMS-bound active sites showed lower motions indicative of the stabilizing effects of BMS-1001 and BMS-1166 respectively (Figure 4.6C).



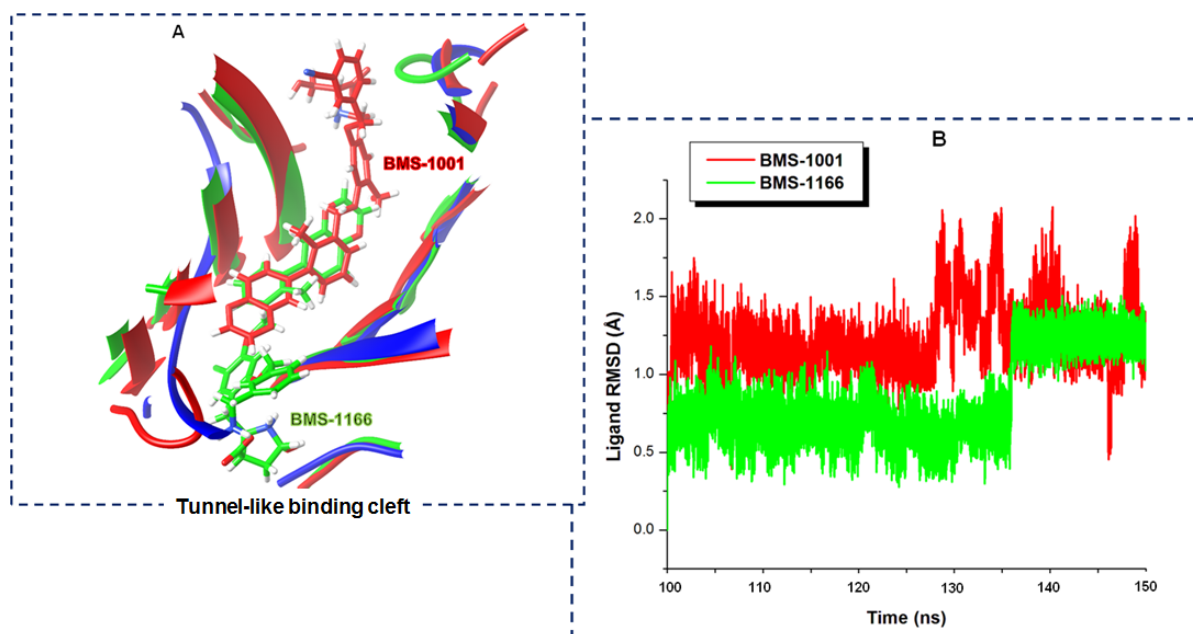


Figure 4.7: Structural stability, positioning and motion of BMS-1001 and BMS-1166 at the tunnel-like binding cleft of PDL-1. [A] Superposition of BMS-1001 (**red**) and BMS-1166 (**green**) as they transverse the binding tunnel [B] Comparative C- $\alpha$  RMSD plot depicting the stability and motion of both BMS-1001 (**red**) and BMS-1166 (**green**) during the post-equilibration MD-simulation period (100ns-150ns).

Accordingly, while the active site (C- $\alpha$ ) of unbound PD-L1 had an RMSD of  $\sim 2.4\text{\AA}$ , the BMS-bound systems had lower deviations at their active site regions. On the average, BMS-1001- and BMS-1166-bound systems had active site RMSDs of  $1.4\text{\AA}$  and  $1.7\text{\AA}$  respectively. In addition, a reduction in atomistic motions among active site residues could be due to the inward pulling effects elicited by complementary interactions with the bound BMS compounds. Further structural insights were provided by measuring the radius of gyration (RoG) with regards to the time-specific structural events ( $\sim 100\text{ns}$ - $150\text{ns}$  post convergence) in the active and non-active site regions with reference to the RMSD. According to previous studies, high RoG value indicates an increase in motion of constituent residues around their C- $\alpha$  atoms, which also corresponds to a notable loss in structural compactness [23], [36]–[38]. On the contrary, a low RoG value suggests a structurally compact protein with minimal motions in constituent atoms. Our findings revealed that the binding of BMS-1001 and BMS-1166 induced considerably high atomistic motions across the secondary

structure of PD-L1 as compared to the unbound system with low RoG (Figure S1). Also, these could imply that while unbound PD-L1 was structurally compact (due to limited motions), the binding of these compounds increased structural motions and activity which could in turn favour dynamical motions of both monomers with respect to dimer formation which corroborates earlier results and previous experimental reports [19], [20]. As estimated, unbound PD-L1 had mean RoG values of 18.7Å while BMS-1001- and BMS-1166-bound PD-L1 had mean RoG values of 20.2Å and 20.0Å respectively. Estimations of the root mean square fluctuation (RMSF) and DCCM for all simulated models were also conducted to gain additional insights into PD-L1 structural occurrences when bound by BMS-1001 and BMS-1166. With regards to RMSF analyses, both monomers (A and B) were evaluated differently in order to define the distinct monomeric motions relative to dimerization as induced by the binding of the BMS compounds. As observed in Figure 4.8, high residual fluctuation occurred concurrently in both monomers of PD-L1 when bound by BMS-1001 and BMS-1166, depictive of a structural characteristic that is important for the dynamical motions of proteins with regards to biological interactions with other biomolecules. As estimated, both monomers (<sub>A</sub>PD-L1 and <sub>B</sub>PD-L1s) for the BMS-1001 and BMS-1166 systems was highly flexible in structure over the simulation period in contrast to the unbound system which demonstrated low structural flexibility and motion. This is in agreement to the results presented above and further indicates that the binding of these compounds induced structural flexibility in PD-L1 monomers which is suitable for dimerization. On the other hand, low flexibility in the unbound system could be due to minimal structural and monomeric motions in the absence of the BMS compounds.

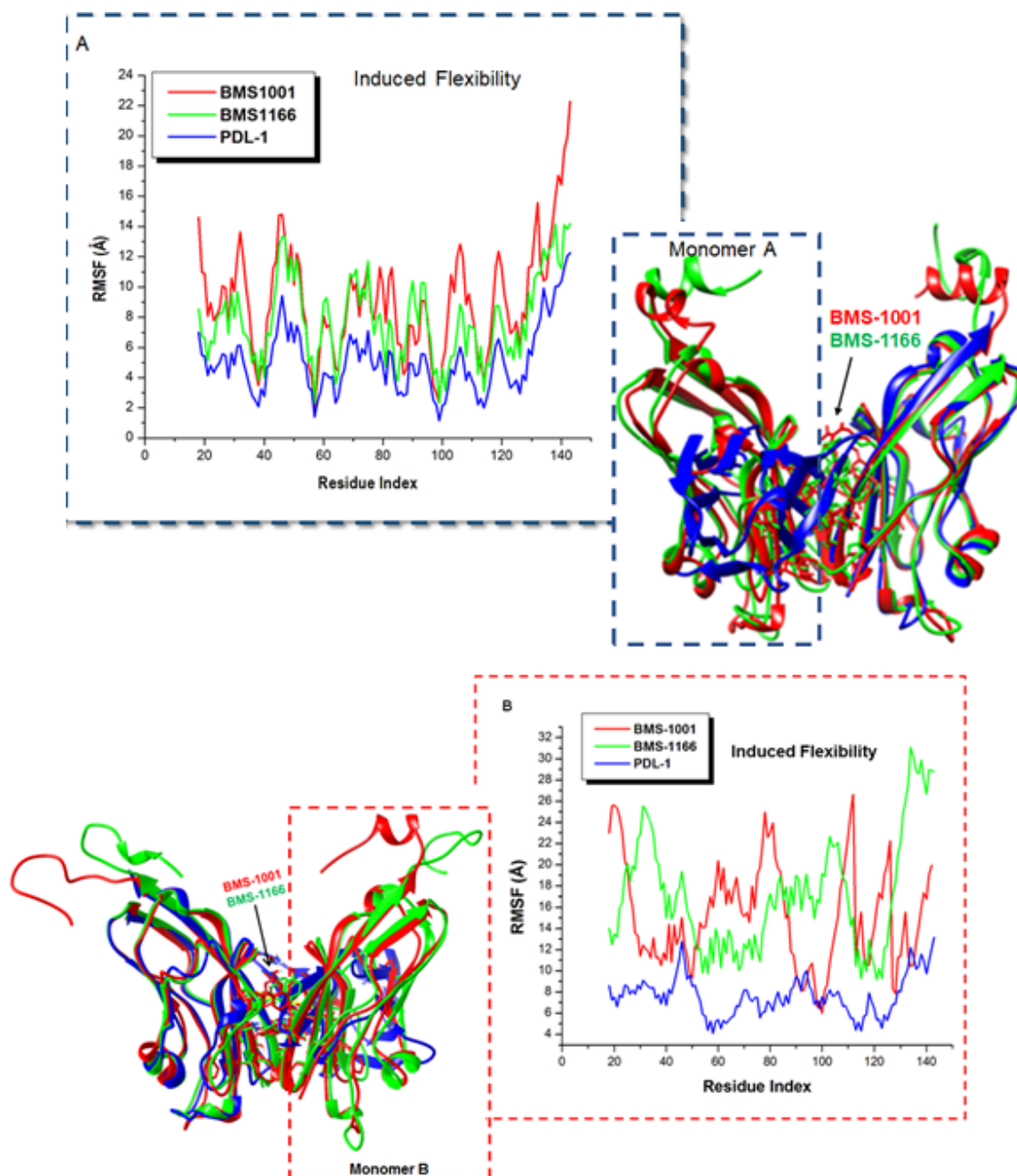


Figure 4.8: Structural flexibility induced simultaneously at both PDL-1 monomers (A and B) by the binding of BMS-1001 (**red**) and BMS-1166 (**green**). [A] Superposed monomers A and corresponding C- $\alpha$  RMSF plot showing the degree of flexibility among the unbound (**blue**), BMS-1001- (**red**) and BMS-1166- (**green**) bound PDL-1 systems [B] Superposed monomers B of unbound (**blue**), BMS-1001- (**red**) and BMS-1166- (**green**) bound PDL-1 systems coupled with corresponding C- $\alpha$  RMSF plot showing the degree of flexibility.

Upon the presumed dimer formation, we further estimated the per-residue fluctuation of specific binding cleft residues that played key interactive roles with BMS-1001 and BMS-1166 (Table 4.1).

Our findings revealed that these residues exhibited high fluctuation, which, suggestively, could be due to the traversing motions and optimal positioning of both compounds at the tunnel-like binding cleft. Taken together, high structural flexibility induced in both PD-L1 monomers with regards to the respective binding of BMS-1001 and BMS-1166 could indicate high structural activity, an occurrence that could in turn enhance PD-L1 protein-protein interactions (dimerization). This reflected the characteristic antagonistic activity elicited by both compounds as previously reported [11], [19], [20].

Table 4.1: Per-residue fluctuation values of crucial binding cleft residues in the unbound and bound PD-L1 systems.

Core active loop residues	Unbound PD-L1	BMS-1001-bound PD-L1	BMS-1166-bound PD-L1
	<b>Per-residue fluctuation (Å)</b>		
<b><sub>A</sub>Tyr56</b>	3.13	4.01	5.45
<b><sub>A</sub>Met115</b>	2.63	4.93	5.13
<b><sub>A</sub>Ala121</b>	4.83	9.47	6.70
<b><sub>A</sub>Asp122</b>	4.08	8.35	5.48
<b><sub>A</sub>Tyr123</b>	3.32	6.94	6.08
<b><sub>A</sub>Lys124</b>	3.55	7.14	5.21
<b><sub>A</sub>Arg125</b>	3.85	7.73	7.05
<b><sub>B</sub>Ile54</b>	5.91	14.09	9.81
<b><sub>B</sub>Tyr56</b>	4.38	15.70	9.72
<b><sub>B</sub>Asn63</b>	5.05	16.28	11.65
<b><sub>B</sub>Val68</b>	6.61	16.95	10.04
<b><sub>B</sub>Val76</b>	5.58	19.75	10.87
<b><sub>B</sub>Met115</b>	4.33	15.57	9.26
<b><sub>B</sub>Ala121</b>	5.36	15.96	9.17

On the average, BMS-1001-bound <sub>A</sub>PD-L1 had a mean RMSF values of 9.1Å while the corresponding <sub>B</sub>PD-L1 had 15.2Å while it was 7.8Å and 17.0Å in the BMS-1166-bound <sub>A</sub>PD-L1 and <sub>B</sub>PD-L1 respectively. Also, it is important to mention that in comparison to the BMS-bound

<sub>A</sub>PD-L1s, the non-liganded <sub>B</sub>PD-L1s exhibited higher flexibility prior to dimerization which could indicate that structural motions were higher in relation to monomeric recruitment. In other words, <sub>B</sub>PD-L1s exhibited higher motions than the liganded-<sub>A</sub>PD-L1s in the conformational phase space, which could correlate with the mechanisms of BMS-induced dimerization as previously reported [19]. On the contrary, unbound <sub>A</sub>PD-L1 and <sub>B</sub>PD-L1 had lower RMSF values of 5.1Å and 7.6Å respectively, predictive of a relatively low structural flexibility. Taken together, the binding of both compounds concurrently induced characteristic flexibility in PD-L1 structure across both monomers, favourable for dimerization.

To further investigate the inhibitory impact of BMS-1001 and BMS-1066 on the inter-residual dynamics and motions of PD-L1, dynamic cross correlation matrix (DCCM) analysis was performed. This method was used to obtain necessary insights into the correlated and anti-correlated motions of constituent residues with respect to their C- $\alpha$  atoms over the simulation period. Positive correlated motions were depicted by a yellow to deep red coloration whereas anti-correlated (negative) motions were represented by a cyan to black coloration. Plots generated for the simulated protein models (across both monomers) are displayed in Figure S2. As shown, correlated motions were highly prominent in the PD-L1 systems bound by BMS-1001 and BMS-1166 as compared to the unbound system which had lesser correlated motions but higher dominance of anti-correlated residual motions (cyan/blue colorations). In agreement with reports earlier presented, these findings depict that the binding of these antagonists induce high motions among the constituent atoms of PD-L1 which correspond to high structural activity.

### 3.2.2. Principal component analysis of structural dynamics and motion

PCA also known as essential dynamics analysis is one of the most important post-dynamic analysis used to study the trajectory of a system over a simulation period [39]. The most significant fluctuation modes of a protein together with the motion of the system in terms of planarity of motion (eigenvectors) and its magnitude (eigenvalues) can be determined using PCA. PCA analysis was performed on the protein C- $\alpha$  atoms using the CPPTRAJ module in AMBER14 to compute the first two components (PC1 and PC2). Figure 4.8 shows the overall motional shifts across two principle components for the free proteins and ligand-protein complex.

To gain insights into the significance of such motions with regards to our study, the conformational attributes of unbound and bound PD-L1 systems were projected directionally along the first two principal components or eigenvectors (ev1/PC1 vs ev2/PC2). Accordingly, distinct conformations and motions were observed in the essential subspace along both principal components. The plot showed a clear and distinct separation of motion among the three systems, with the BMS-bound PD-L1 exhibiting a highly dispersed motion as compared to the unbound system, which showed a more compact motion. Induced dispersion among the protein systems bound by BMS-1001 and BMS-1166 could only imply that the binding of these compounds enhanced structural activity unlike in the unbound system which maintained structural compactness across the simulation period. These findings further corroborates above-presented results which altogether reflects that the binding of these compounds stabilized a high activity conformation favorable for dimerization as previously reported [18]–[20]. Suggestively, high activity conformation relative to PD-L1 dimerization as induced by these BMS antagonistic could disallow its association with PD-1 relative to immune checkpoint activities in T cell exhaustion.

### **4.3.3. BMS-1001 and BMS-1166 exhibited ‘monomeric interchange’ and systematic high-affinity binding prior to PD-L1 dimerization**

The antagonistic mechanism of the BMS-compounds towards PD-L1 is centered on their ability to induce dimerization thereby preventing an association with PD-1 which promotes tumor progression and survival by escaping tumor neutralizing immune surveillance. Association of PD-L1 with PD-1 of T-cells lead to T Cell dysfunction, exhaustion, neutralization, and interleukin-10 (IL-10) production in tumor mass [13].

According to previous reports, the BMS-compounds bind to a PD-L1 monomer and facilitate monomeric recruitment leading to the formation of a dimer molecule [19]. Interestingly, from our previous findings, as presented in Section 3.1 above, we proposed a model that possibly depicted the dynamical motions of the respective monomers in the conformational phase with respect to the starting structure employed in this study. Moreover, we observed that the BMS-compounds systematically transited from the initial monomer ( $_A$ PD-L1), to which they were first bound, onto the second monomer prior to dimerization (Figure 4.4 and 5). As shown in the initial trajectory (0-30ns), the BMS-compounds maintained their starting position at the first monomeric binding site ( $_A$ PD-L1) until 30ns where we observed that it was bound at a similar binding site on monomer  $_B$ PD-L1, and eventually at the tunnel-like binding cleft upon the formation of PD-L1 dimer. This could represent an important insight into the mechanistic occurrences that accompany BMS-induced dimerization of PD-L1 in cancer immunotherapy. The MM/PBSA analysis was further employed to deduce the binding strength and affinity of the BMS compounds with regards to the dimerized protein form. This estimated the binding free energies involved in complex formation in relation to their antagonistic activities. Using trajectories at which the PD-L1 dimer was formed in both BMS-1001 and BMS-1166-bound systems, our estimations revealed that both compounds exhibited considerably high and favorable negative energies indicative of a strong and high affinity

binding. Accordingly, BMS-1001 had a  $\Delta G$  value of -20.5kcal/mol while BMS-1166 had a  $\Delta G$  value of -18.1kcal/mol. Relatively, we estimated the monomeric  $\Delta G$  values of the BMS compounds to each monomer at the respective simulation periods. We observed that both BMS-1001 and BMS-1166 exhibited similar  $\Delta G$  towards the PD-L1 monomers even when interchanged at their respective ‘monomeric’ binding sites as earlier mentioned. BMS-1001 had an estimated  $\Delta G$  of 23.64kcal/mol and 23.75kcal/mol at the initial (0-30ns) and intermediate (60-90ns) trajectories while BMS-1166 had 23.43kcal/mol and 23.66kcal/mol respectively. At the final trajectory, BMS-1001 and BMS-1166 had  $\Delta G$  values of 23.78 kcal/mol and 25.96kcal/mol respectively. MM/PBSA estimations for the distinct MD simulation periods are presented in Table 4.2. Therefore, these considerably high negative free energy values could correlate with a high-affinity binding that is essential for the systematic dimerization of PD-L1 in the presence of the respective BMS-compounds [19].

Table 4.2: MM/PBSA binding free energy profiles of BMS-1001 and BMS-1166 to PD-L1

<b>Complexes</b>	<b><math>\Delta E_{vdw}</math> (kcal/mol)</b>	<b><math>\Delta E_{ele}</math> (kcal/mol)</b>	<b><math>\Delta G_{gas}</math> (kcal/mol)</b>	<b><math>\Delta G_{sol}</math> (kcal/mol)</b>	<b><math>\Delta G_{bind}</math> (kcal/mol)</b>
<b>MD1 BMS-1001</b>	-35.9	39.12	36.0	28.1	23.64
<b>MD2 BMS-1001</b>	-28.12	45.12	34.0	27.1	23.75
<b>MD3 BMS-1001</b>	-34.8	-27.9	-42.7	30.8	23.78
<b>MD1 BMS-1166</b>	-26.78	40.15	37.92	26.87	23.43
<b>MD2 BMS-1166</b>	-27.42	44.91	38.00	27.32	23.66
<b>MD3 BMS-1166</b>	-27.9	-48.0	-39.0	29.7	25.96

$\Delta E_{ele}$  = electrostatic energy;  $\Delta E_{vdw}$  = van der Waals energy;  $\Delta G_{bind}$  = total binding free energy;  $\Delta G_{sol}$  = solvation free energy  $\Delta G$  = gas phase free energy.



Upon dimerization, additional insights into the interactions and positioning of BMS-1001 and BMS-1166 at the tunnel-like binding cleft were gained by molecular visualizations using the Graphical User Interface of UCSF Chimera and Discovery Studio 2016 Client [40]. These revealed intermolecular interactions that occurred at the ‘dimeric’ binding cleft of PD-L1, which is systematically induced after the binding of a BMS compound (BMS-1001/BMS-1166) to a PD-L1 monomer, as elucidated in the model presented in this study. As observed at each of the ‘dimeric’ binding sites, both compounds transversed the tunnel-like binding cleft while high-affinity interactions occurred between constituent ligand groups and key residues lying on either side of the monomers within and outside the tunnel to achieve optimal binding modes and stability. Outside the tunnel, key residues such as <sub>A</sub>Asp122, <sub>A</sub>Tyr123, <sub>A</sub>Lys124 and <sub>A</sub>Arg125 formed strong high-affinity interactions with the (2R)-2-amino-3-hydroxypropanoic acid and 3-cyanobenzyl substituents; which majorly constitute hydrogen bonds and attractive charge (ionic) interactions that account for the stability of BMS-1001 at this region. Moreover, key interactions that accounted for high-affinity binding and stability of BMS-1001 at the other side of the tunnel (binding cleft) were elicited by deep hydrophobic pocket residues; <sub>B</sub>Ala121 ( $\pi$ -alkyl) and <sub>B</sub>Asp122 (CH--O), with the buried 2,3-dihydro-1,4-benzodioxine moiety.

Also, the orientations of <sub>A</sub>Tyr56 and <sub>B</sub>Tyr56 played important roles in the stability of BMS-1001 at the binding cleft of PD-L1 possibly via ring  $\pi$ - $\pi$  stacking. Likewise, BMS-1166 had a similar orientation at the binding cleft of PD-L1 wherein it transverses the tunnel-like binding cleft and elicit high-affinity interactions at both sides of the tunnel. At the outer region of the tunnel, residues such as <sub>B</sub>Asp122, <sub>B</sub>Tyr123 and <sub>B</sub>Lys124, <sub>B</sub>Arg125 existed in high-affinity (hydrogen and ionic) interactions with the (2R, 4R)-4-hydroxypyrrolidine-2-carboxylic acid moiety while  $\pi$ - $\pi$  stacking

of the 3-cyanobenzyl ring by <sub>B</sub>Tyr123 complements its role in BMS-1166 stability around this region.

The role of <sub>A</sub>Tyr56 was also crucial to the stability of BMS-1166 while it transverses the tunnel possibly due to its involvement in ring  $\pi$ - $\pi$  stacking interactions with the intermediate 5-chlorophenyl ring. Likewise, the contributions of <sub>A</sub>Tyr123 (NH--O) and <sub>A</sub>Asp122 (CH--O) towards the stability of BMS-1166 encompass their interactions with the 2,3-dihydro-1,4-benzodioxine moiety at the deep hydrophobic pocket. Furthermore, per-residue energy decomposition analyses were used to estimate the individual energy contributions of these crucial residues to the stability of BMS-1001 and BMS-1166 respectively at the binding cleft. Energy contributions towards BMS-1001 stability at the binding cleft were considerably high among key tunnel residues (inside and outside) as earlier mentioned in comparison with other interacting residues.

Outside the tunnel, <sub>A</sub>Asp122, <sub>A</sub>Tyr123, <sub>A</sub>Lys124, and <sub>A</sub>Arg125 had high electrostatic energy contributions of -10.5kcal/mol, -5.4kcal/mol, -10.1kcal/mol and -6.9kcal/mol respectively, which could be due to the occurrence of strong hydrogen and ionic interactions. Inside the tunnel, key residues such as <sub>B</sub>Ala121 and <sub>B</sub>Asp122 that interacted with the buried 2,3-dihydro-1,4-benzodioxine moiety of BMS-1001 had van der Waals and electrostatic energies of -3.6kcal/mol and -3.4kcal/mol respectively. In addition, <sub>A</sub>Tyr56 (inside: 2,3-dihydro-1,4-benzodioxine) and <sub>B</sub>Tyr56 (outside: 4-methylphenoxy ring) which elicited ring  $\pi$ - $\pi$  stacking of BMS-1001 at the binding tunnel had van der Waals energy contributions of -2.6kcal/mol and -3.4kcal/mol respectively. Similarly, residues such as <sub>A</sub>Asp122 and <sub>A</sub>Tyr123 had high electrostatic energies of -5.3kcal/mol and -6.2kcal/mol respectively. Likewise, crucial residues at the outer tunnel region that interacted with the (2R, 4R)-4-hydroxypyrrolidine-2-carboxylic acid moiety and the 3-cyanobenzyl ring had considerably high energy contributions, which further substantiate high-

affinity interactions earlier reported (Figure 4.6A). Accordingly, <sub>B</sub>Asp122, <sub>B</sub>Lys124 and <sub>B</sub>Arg125 had electrostatic energy contributions of -9.5kcal/mol, -10.7kcal/mol, and -12.4kcal/mol respectively while <sub>B</sub>Tyr123 and <sub>A</sub>Tyr56 had van der Waals energy contributions of -4.2kcal/mol and -3.2kcal/mol respectively. Taken together, considerably high energy contributions among these crucial residues at the target site of PD-L1 account for high-affinity binding and stability of both BMS-1001 and BMS-1166 while they transverse the tunnel-like binding cleft to elicit their antagonistic functions.

#### **4.4 Conclusion**

The use of small molecule compounds as immune checkpoint inhibitors represent a paradigm shift in the cancer immunotherapy (immune checkpoint blockade - ICB). This therapeutic strategy is being initiated with the identification of novel BMS compounds; BMS-1001 and BMS-1166 which specifically target PD-L1 and prevents its association with PD-1 resulting in the reactivation of T cells and effective killing of cancer cells. Hence it was necessary to investigate the activities of these antagonistic compounds coupled with the mechanisms by which they induce monomeric recruitment of the distinct PD-L1 monomers, so as to obtain necessary insights which could aid the future design of highly effective antagonists and monomeric recruiters in ICB therapy. Appropriate investigations were carried out using MD simulation and MM/PBSA techniques, which are sufficient to provide structural and molecular occurrences associated with the binding of these compounds to PD-L1. Using a structurally separated starting structures for the unbound and BMS-bound systems, our result revealed that the BMS-1001 and BMS-1166 bound monomers elicited the recruitment of the other monomer, as the simulation progresses, there was ligand exchange from the monomer A to monomer B, this is due to the weak binding interaction that existed between the ligands and monomer A active site residues, however, monomer B active site

residues provided a much stronger interaction. MM/PBSA estimations revealed that both compounds were bound to PD-L1 with considerably high-affinities as evidenced by relatively high  $\Delta G$  values. Moreover, molecular visualization revealed that both compounds were positioned across the tunnel-like binding cleft, and while they traversed this region, they maintained high-affinity interactions (strong hydrogen and ionic bonds) with key active site residues both within and outside the tunnel via their substituent groups. More importantly were the high affinity interactions elicited by <sub>A</sub>Asp122, <sub>A</sub>Tyr123, <sub>A</sub>Lys124 and <sub>A</sub>Arg125 with the (2R)-2-amino-3-hydroxypropanoic acid and 3-cyanobenzyl substituents of BMS-1001 outside the tunnel and <sub>B</sub>Ala121 and <sub>B</sub>Asp122, inside the tunnel (deep hydrophobic pocket). Also, with regards to BMS-1166, high-affinity interactions occurred outside the tunnel between <sub>B</sub>Asp122, <sub>B</sub>Tyr123, <sub>B</sub>Lys124 and <sub>B</sub>Arg125 with the 3-cyanobenzyl and (2R, 4R)-4-hydroxypyrrolidine-2-carboxylic acid substituents while at the other side of the tunnel, <sub>A</sub>Tyr123 and <sub>A</sub>Asp122 played crucial role in stabilizing BMS-1166 via interactions with its 2,3-dihydro-1,4-benzodioxine moiety. The important roles of these residues were further revealed by per-residue decomposition analyses which estimated high individual energy contributions and could underlie the considerable binding strength and stability demonstrated by these compounds. Comparative insights into the structural dynamics and motions of PD-L1 in its unbound and bound states were provided by post-MD conformational analysis using C- $\alpha$  RMSD, C- $\alpha$  RMSF, C- $\alpha$  RoG, DCCM and PCA. An interesting and consistent pattern observed was that the binding of both compounds induced distinctively high structural motions and activity in contrast to the unbound model which demonstrated structural compactness and rigidity across the MD simulation period. Binding of the ligand to a monomer elicited the recruitment of the other monomer inducing dimerization of the monomers, there was also a ligand exchange from monomer A unto monomer B. This transition in the structural

dynamics and motion of PD-L1 when bound by BMS-1001 and BMS-1166 could be seen as the mechanism of PD-L1/PD-1 association. High per-residue fluctuation of key residues that interacted with the substituents of both compounds along the tunnel-like binding cleft could as well imply their high activity and contributions towards the favorable positioning and binding of BMS-1001 and BMS-1166. Taken together, BMS-1001 and BMS-1166 bind strongly to PD-L1 due to high-affinity interactions with crucial residues at the tunnel-like binding cleft and induce high structural activity favorable for dimerization, which could in turn prevent association with PD-1.

### **Conflict of interest**

The authors declare none.

### **Acknowledgement**

The authors acknowledge the College of Health Sciences, UKZN for their financial and infrastructural support and at the same time thank the Centre for High Performance Computing (CHPC, [www.chpc.ac.za](http://www.chpc.ac.za)), Cape Town, for computational resources.

### **Reference**

- [1] H. Zhang and J. Chen, “Current status and future directions of cancer immunotherapy,” *J. Cancer*. **2018**, *9*, 1773–1781.
- [2] E. J. Jacobs, C. C. Newton, V. L. Stevens, P. T. Campbell, S. J. Freedland, and S. M. Gapstur, “Daily aspirin use and prostate cancer-specific mortality in a large cohort of men with nonmetastatic prostate cancer,” *J. Clin. Oncol.* **2014**, *32*, 3716–3722.
- [3] M. J. Smyth, “Multiple approaches to immunotherapy - The new pillar of cancer treatment,” *Immunology and Cell Biology*. **2017**.
- [4] J. Dine, R. Gordon, Y. Shames, M. K. Kasler, and M. Barton-burke, “Immune Checkpoint

Inhibitors: An Innovation in Immunotherapy for the Treatment and Management of Patients with Cancer of immunotherapies reflects a promising new approach to cancer treatment involving activation of the immune system against cancer. Immune checkpoint inhibitors: An innovation in immunotherapy for the treatment and management of patients with cancer,” *Asia Pac J Oncol Nurs.* **2017**, *4*, 127–35.

- [5] K. Chatamra, “Cancer and immunology.,” *Asian Pacific J. Allergy Immunol.* **1991**, *9*, 71–73.
- [6] C. Berek, “Eosinophils: Important players in humoral immunity,” *Clin. Exp. Immunol.* **2016**, *183*, 57–64.
- [7] J. F. Ponte *et al.*, “Enhancement of humoral and cellular immunity with an anti-glucocorticoid-induced tumour necrosis factor receptor monoclonal antibody,” *Immunology.* **2010**, *130*, 231–242.
- [8] I. Mellman, G. Coukos, and G. Dranoff, “Cancer immunotherapy comes of age,” *Nature.* **2014**, *480*, 480–489.
- [9] S. Aspeslagh, A. Marabelle, J.-C. Soria, and J.-P. Armand, “Upcoming innovations in lung cancer immunotherapy: focus on immune checkpoint inhibitors.,” *Chinese Clin. Oncol.* **2015**, *4*, 48.
- [10] A. Makkouk, G. Weiner, "Cancer immunotherapy and breaking immune tolerance-new approaches to an old challenge," *NIH Public Access*,” **2016**, *75*, 5–10.
- [11] M. S. Barbee, A. Ogunniyi, T. Z. Horvat, and T. O. Dang, “Current status and future directions of the immune checkpoint inhibitors ipilimumab, pembrolizumab, and nivolumab in oncology,” *Ann. Pharmacother.* **2015**, *49*, 907–937.
- [12] N. D. Shore, “Advances in the understanding of cancer immunotherapy,” *BJU Int.* **2015**,

116, 321–329.

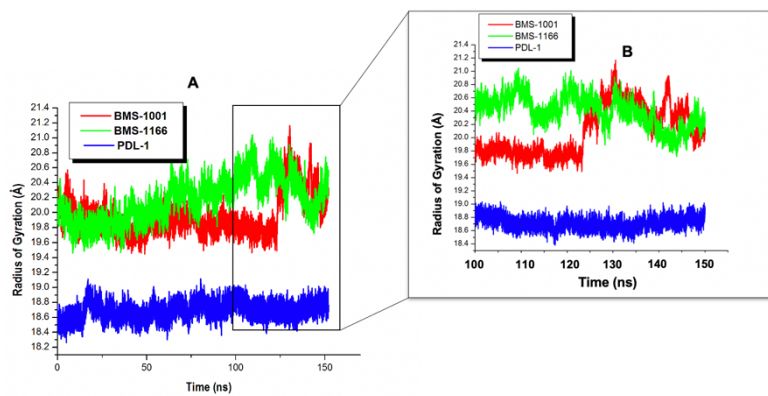
- [13] H. O. Alsaab *et al.*, “PD-1 and PD-L1 checkpoint signaling inhibition for cancer immunotherapy: mechanism, combinations, and clinical outcome,” *Front. Pharmacol.* **2017**, 8, 1–15.
- [14] L. M. Francisco, P. T. Sage, and A. H. Sharpe, “The PD-1 pathway in tolerance and autoimmunity,” *Immunological Reviews.* **2010**, 236, 219–242.
- [15] X. Wang, F. Teng, L. Kong, and J. Yu, “PD-L1 expression in human cancers and its association with clinical outcomes,” *OncoTargets and Therapy.* **2016**, 9, 5023–5039.
- [16] S. Ostrand-Rosenberg, L. A. Horn, and S. T. Haile, “The Programmed Death-1 Immune-Suppressive Pathway: Barrier to Antitumor Immunity,” *J. Immunol.* 2014, 193, 3835–3841.
- [17] M. M. Zhan, X. Q. Hu, X. X. Liu, B. F. Ruan, J. Xu, and C. Liao, “From monoclonal antibodies to small molecules: The development of inhibitors targeting the PD-1/PD-L1 pathway,” *Drug Discov. Today.* **2016**, 21, 1027–1036.
- [18] A. F. Abdel-Magid, “Inhibitors of the PD-1/PD-L1 Pathway Can Mobilize the Immune System: An Innovative Potential Therapy for Cancer and Chronic Infections,” *ACS Med. Chem. Lett.* **2015**, 6, 489–490.
- [19] L. Skalniak *et al.*, “Small-molecule inhibitors of PD-1/PD-L1 immune checkpoint alleviate the PD-L1-induced exhaustion of T-cells,” *Oncotarget*, **2017**.
- [20] S. A. Andrei *et al.*, “Stabilization of protein-protein interactions in drug discovery,” *Expert Opinion on Drug Discovery.* **2017**, 12, 925–940.
- [21] E. F. Pettersen *et al.*, “UCSF Chimera - A visualization system for exploratory research and analysis,” *J. Comput. Chem.* **2004**, 25, 1605–1612.
- [22] N. Eswar *et al.*, “Comparative protein structure modeling using MODELLER.,” *Curr.*

- Protoc. Protein Sci.* **2007**, 2, 2.9.
- [23] F. A. Olotu and M. E. S. Soliman, “From mutational inactivation to aberrant gain-of-function: Unraveling the structural basis of mutant p53 oncogenic transition,” *J. Cell. Biochem.* **2018**, 119, 2646–2652.
- [24] M. Abdullahi, F. A. Olotu, and M. E. Soliman, “Allosteric inhibition abrogates dysregulated LFA-1 activation: Structural insight into mechanisms of diminished immunologic disease,” *Comput. Biol. Chem.* **2018**, 73, 49–56.
- [25] M. Lawal, F. A. Olotu, and M. E. S. Soliman, “Across the blood-brain barrier: Neurotherapeutic screening and characterization of naringenin as a novel CRMP-2 inhibitor in the treatment of Alzheimer’s disease using bioinformatics and computational tools,” *Comput. Biol.* **2018**, 98, 168–177.
- [26] D. A. Case *et al.*, “The Amber biomolecular simulation programs,” *J. Comput. Chem.* **2005**, 16, 1668–1688.
- [27] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, “Development and Testing of a General Amber Force Field,” *J. Comput. Chem.* **2004**, 25, 1157–1174.
- [28] H. J. C. Berendsen *et al.*, “Molecular dynamics with coupling to an external bath Molecular dynamics with coupling to an external bath,” *J. Chem. Phys.* **2012**, 3684, 926–935.
- [29] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, “Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes,” *J. Comput. Phys.* **1977**, 23, 327–341.
- [30] D. R. Roe and T. E. Cheatham III, “PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data,” *J Chem Theory Com.* **2013**, 9, 3084–3095.
- [31] E. Seifert, “OriginPro 9.1: Scientific data analysis and graphing software - Software

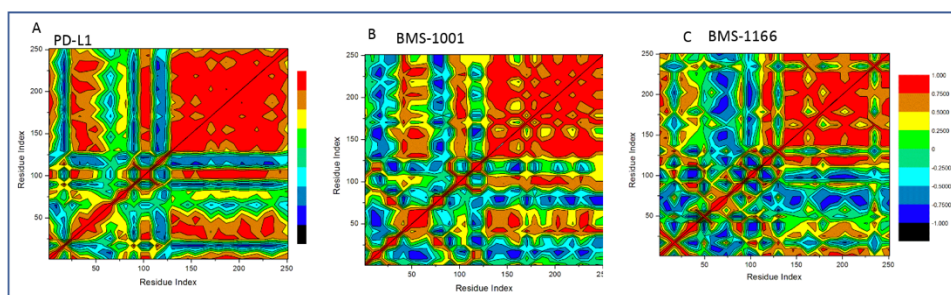


- review,” *J. Chem. Inf. Model.* **2014**, *54*, 1552.
- [32] P. A. Kollman *et al.*, “Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models,” *Acc. Chem. Res.* **2000**, *33*, 889–897.
- [33] E. A. Adeniji, F. A. Olotu, and M. Soliman, “Exploring the lapse in druggability: Sequence Analysis, Structural Dynamics and binding site characterization of K-RasG12C variant, a feasible oncotherapeutics target.,” *Anticancer. Agents Med. Chem.* **2018**, *18*.
- [34] G. E. Arnold and R. L. Ornstein, “Molecular dynamics study of time-correlated protein domain motions and molecular flexibility: Cytochrome P450BM-3,” *Biophys. J.* **1997**, *73*, 1147–1159.
- [35] C. C. David and D. J. Jacobs, “Principal component analysis: a method for determining the essential dynamics of proteins.,” *Methods Mol. Biol.* **2014**, *1084*, 193–226.
- [36] M. Salleh, A.B., Rahim, A.S., Rahman, R.N., Leow, T.C., and Basri, “The role of Arg157Ser in improving the compactness and stability of ARM Lipase,” *J Comput. Sci. Syst. Biol.* **2012**, *5*, 38–46.
- [37] M. I. Lobanov, N. S. Bogatyreva, and O. V Galzitskaia, “Radius of gyration is indicator of compactness of protein structure,” *Mol. Biol. (Mosk)*. **2008**, *42*, 701–706.
- [38] A. A. El Rashedy, F. A. Olotu, and M. E. S. Soliman, “Dual drug targeting of mutant Bcr-Abl induces inactive conformation: New strategy for the treatment of chronic myeloid leukemia and overcoming monotherapy resistance,” *Chem. Biodivers.* **2018**.
- [39] F. Sittel, A. Jain, and G. Stock, “Principal component analysis of molecular dynamics: On the use of Cartesian vs. internal coordinates,” *J. Chem. Phys.* **2014**, *141*.
- [40] D. S. BIOVIA, “Discovery Studio 2016 Client,” *San Diego Dassault Systèmes*, **2016**.

## SUPPLEMENTARY



FigureS1: Radius of gyration of the whole system (A) and post equilibration (B) RoG.



FigS2: Correlated motion plot of simulated protein models (across both monomers) over the simulation period

## CHAPTER 5

### **Drug Promiscuity: Exploring the polypharmacology potential of 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones as an Inhibitor of the ‘god father’ of Immune Checkpoint**

Opeyemi S. Soremekun<sup>a</sup>, Fisayo A. Olotu<sup>a</sup>, Clement Agoni<sup>a</sup>, Mahmoud E. S. Soliman<sup>a\*</sup>

<sup>a</sup>Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

\*Corresponding Author: Mahmoud E.S. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872

## **Abstract**

High production cost, instability, low tumor penetration are some of the shortcomings that have characterized and undermined the use of antibodies as a target for Cytotoxic T-lymphocytes associated protein 4 (CTLA-4). Design and discovery of small molecule inhibitors have therefore become a *sine qua non* in targeting immune proteins implicated in immune disorders. In this study, we utilized a drug repositioning approach to explore the characteristic feature of unrelated proteins to have similar binding sites and the promiscuity of drugs to repurpose an existing drug to target CTLA-4. CTLA-4 and Kallikrein-7 were found to have similar binding sites, we therefore used 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones (TDSO) which is an inhibitor of Kallikrein-7 as our lead compound. High throughput screening using TDSO as a lead compound resulted in 9 hits with ZINC04515726 and ZINC08985213 having the highest binding score. We went ahead to investigate the interaction of these compounds with CTLA-4 by conducting a molecular dynamic simulation. Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) estimations revealed that TDSO had the highest binding energy value of -28.51Kcal/mol, with ZINC04515726 and ZINC08985213 having -23.76Kcal/mol and -21.03Kcal/mol respectively. The per-residue decomposition highlighted Tyr24, Ala25, Gly28, Ala30, Tyr53 and Asn72 as having significantly high electrostatic energy contributions and the main contributing residues to the binding of TDSO, ZINC04515726 and ZINC08985213 to Cytotoxic T lymphocytes CTLA-4. Summarily, from the results gathered, we proposed that TDSO can be an effective immune check point small molecule inhibitor against the suppression of T-cell activation, proliferation, and tumor cell eradication.

**Keywords:** Cytotoxic-lymphocytes associated protein 4, Drug repurposing, Polypharmacology,

## 5.1 Introduction

Cancer Immunotherapy has become a standard treatment for a wide range of cancer malignancies and is considered to be the “fifth pillar” of cancer therapy, after surgery, chemotherapy, radiation and targeted therapy <sup>1</sup>. Immune system does not just prevent the body against infectious organisms but also help in eradicating malignancies when stimulated <sup>2,3</sup> through strengthening the host immune responses against tumors, annihilating signals produced by cancer cells that repress immune responses, or supplying modified immune system components <sup>4</sup>. With the understanding of the functionality of the immune system, small molecules, peptides, recombinant antibodies, vaccines and cellular therapeutic modalities are being employed in manipulating the immune system for cancer treatment <sup>5</sup>. The use of immune checkpoint inhibitor in cancer immunotherapy include generating antibodies against the cytotoxic T-lymphocytes associated protein 4 (CTLA-4), the programmed death receptor 1 (PD-1) or its ligand (PD-L1)<sup>6</sup>. CTLA-4 is mainly expressed on activated CD8<sup>+</sup> effector T Cells and negates the early-stage T-cell activation and cell cycle progression. CTLA-4 is homologous to CD28, and it possesses the same ligands, including B7-1 (CD80) and B7-2 (CD86), thus blocking the activation and proliferation of antigen-activated T cells secondary to CD28/B7 interaction <sup>7</sup> (Figure 5.1). The inhibition of CTLA-4 signaling is extremely promising to recover the suppression of T-cell activation, proliferation, and infiltration into tumors. This process can result in enhanced anti-tumor immunity and tumor cell eradication <sup>8</sup>. FDA in 2011 approved anti-CTLA-4 antibodies ipilimumab and tremelimumab for the treatment of metastatic melanoma; this marked the beginning of a new era for cancer immunotherapy <sup>9</sup>. Ipilimumab and tremelimumab work by blocking the binding of CTLA-4 to its ligand (CD80 AND CD86) <sup>10</sup>. Due to the lower production costs, higher stability, improved tumor penetration, amenability for oral administration and elimination of immunogenicity, small-molecular weight inhibitors are expected to be a viable option in place of antibodies as immune check point inhibitors

<sup>11</sup>. Unfortunately, there are no available small molecule inhibitors that targets CTLA-4. Hence, we used drug repositioning and pharmacophore approach to propose drugs that can effectively serve as an inhibitor of CTLA-4.

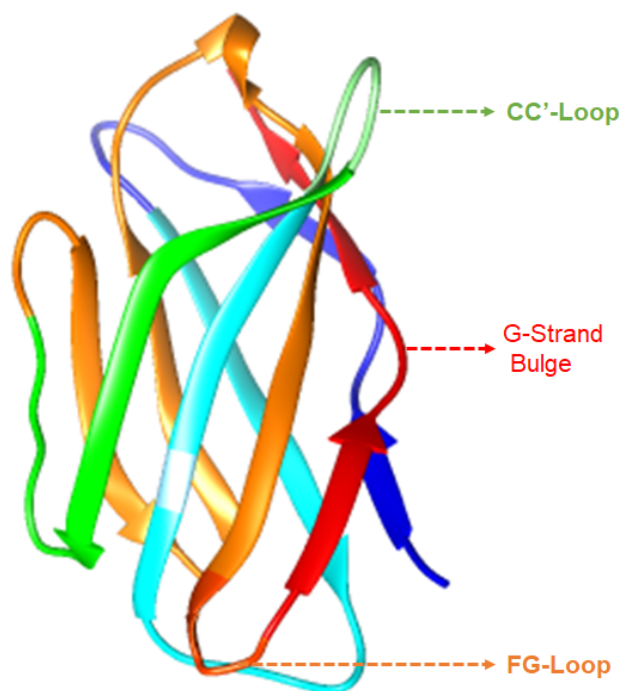


Figure 5.1: 3D crystal structure of CTLA-4. The CC'-Loop is highlighted in green, G-strand in red and the FG-loop in orange. (PDB:5TRU) <sup>12</sup>

## 5.2 Computational Method

### 5.2.2 Binding site similarity identification

The complexity of protein and the dynamics of their active site have made it possible for drugs to modulate multiple targets implicated in diseases. This binding site similarity has been explored in repurposing drugs for effective therapeutic potentials. We used Binding DB database ([bindingdb.org/bind/index.jsp](http://bindingdb.org/bind/index.jsp)), a publicly accessible database which contains approximately 20000 experimentally determined binding affinities of protein-ligand complexes, chemical

structures, substructure and similarity<sup>13</sup>. Therefore this drug repositioning methodology which identifies novel target for existing drugs using bioinformatics databases and algorithms as described by Heinrich *et al.*, 2010 was used <sup>14</sup> . This structural approach explores binding site similarities between the target protein and other related/unrelated proteins to predict novel binders. Therefore, we employ a similar principle to identify small molecule compounds/antagonists that are likely to block the immune checkpoint activity of CTLA-4.

### *5.2.3 Pharmacophore Model Creation using Per Residue Energy Decomposition (PRED) Based Approach*

Identification of small molecule CTLA-4 antagonists was followed up by the modeling of novel compounds that could also act in a similar inhibitory capacity towards the target protein.

Contrary to other conventional pharmacophore modelling methods, we used an in house PRED approach to define the pharmacophoric architecture of the ligand- receptor in order to retrieve a more tailored potential hit <sup>15,16</sup>. This pharmacophore model explore both the structural features of the proteins as well as the chemical features of ligands<sup>15</sup>. To generate a PRED-based pharmacophore model, PRED decomposition was computed from MM/PBSA calculations after 20ns MD simulations of the (TDSO-CTLA-4) complex. Pharmacophoric features based on the receptor-ligand interaction obtained from this short run MD simulation was selected. Residues Ala25, Ala30, Thr31, Val33, Met55, Gly71 and Asn72 were found to be the highest contributing residues that interact with the ligand (FigS5A). These ligand features were set as a query to generate a PRED-based pharmacophore model in ZINCPharma <sup>17</sup>. Furthermore, the PRED-based pharmacophore model (Figure S1) was used to screen the ZINC database <sup>18</sup> for compounds with similar features to obtain the novel hits. The filter was set to query Zinc purchasable compounds

of molecular weight  $\leq 500$  for hit screening while rotatable bonds was set at  $\leq 10$ . Lipinski's "rule of five" was also set as cut-off. Hits were sorted by increasing RMSD Scores.

#### *5.2.4 Drug Likelihood Assessment*

An Online software SwissADME <sup>19</sup> was used to determine the physiochemical descriptors, pharmacokinetic properties and drug-like nature of the screened compounds. SwissADME uses the "Brain or Intestinal Estimated permeation, (BOILED-Egg)" method which computes the lipophilicity and polarity of small molecules.

#### *5.2.6 System preparation and Molecular dynamic (MD) simulations*

The x-ray crystal structure of CTLA-4 in complex with an antibody check point inhibitor Ipilimumab with Protein Data Bank code (5TRU) <sup>12</sup> was obtained from RSCB Protein Data Bank <sup>20</sup>. The Ipilimumab was deleted to obtain an unbound CTLA-4 protein. CTLA-4 preparation as well as visualization was conducted in Chimera <sup>21</sup> and Avogadro Visualization <sup>22</sup> software. 2D-structure of the ligands were setup for auto-optimization using UFF forcefield and steepest descent algorithm on Avogadro 1.2.0 <sup>22</sup>. This was carried out to adjust inter-connective bonds and eliminate clashes. Further preparations of the ligands and CTLA-4 were carried out using the Molegro Molecular Viewer (MMV) software <sup>23</sup>. The ligands were docked into the protein using automated Autodock Tools docking procedure which entails grid box 'active-site' definition <sup>24</sup>. The docking score was validated by re-docking using Chimera <sup>25</sup> and HADDOCK <sup>26</sup>. The results are presented in TableS1. The MD simulation was performed using the GPU version of the PMEMD engine provided with AMBER package, FF14SB variant of the AMBER force field was used to describe the protein. ANTECHAMBER was used to generate partial charges for the ligand by utilizing the



restrained electrostatic potential (RESP) and the General Amber Force Field (GAFF) procedures. The leap module of AMBER 14 allowed the addition of hydrogen atoms, as well as Na<sup>+</sup> and Cl<sup>-</sup> counter ions for neutralization to both the Apo and bound system. An orthorhombic box of TIP3P water molecules at a distance of 9Å from all the protein atoms was used to solvate all the systems prior to simulation <sup>27</sup>. An initial minimization of 2000 steps were carried out within an applied restraint potential of 500 kcal mol<sup>-1</sup> Å<sup>-2</sup> for both solutes were performed for 1000 steps using a steepest decent method followed by a 1000 step of conjugate gradients. An additional full minimization of 10000 steps was further carried out by conjugate gradient algorithm without restrain.

A gradual heating MD simulation from 0 k to 300 K was executed for 50ps, such that the system maintained a fixed number of atoms and fixed volume, *i.e.*, a canonical ensemble (NVT). The solutes within the system are imposed with a potential harmonic restraint of 10 kcal mol<sup>-1</sup> Å<sup>-2</sup> and collision frequency of 1.0 ps<sup>-1</sup>. Following heating, an equilibration estimating 500ps of each system was conducted; the operating temperature was kept constant at 300 k. Additional features such as number of atoms and pressure were also kept constant mimicking an isobaric-isothermal ensemble (NPT). The system pressure was maintained at 1 bar using Berendsen barostat.

The total time for the MD simulation was 150 ns. In each simulation the SHAKE algorithm was employed to constrict the bonds of hydrogen atoms. The step size of each simulation was 2 fs and an SPFP precision model was used. The simulations coincided with isobaric-isothermal ensemble (NPT), with randomized seeding, constant pressure of bar maintained by the Berendsen barostat, a pressure-coupling constant of 2ps, a temperature of 300 K and Langevin thermostat with collision frequency of 1.0ps<sup>-2</sup>.

### 5.2.6 Post-dynamic analysis

The trajectories were saved and further analysis done using PTRAJ<sup>28</sup> module, CPPTRAJ<sup>28</sup> module was used to analyze Radius of Gyration, root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF).

### 5.2.7 Thermodynamic Calculation

The free binding energy of the CTLA-4 active site was analyzed by the Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA)<sup>29</sup>. Binding free energy calculation is an endpoint energy calculation that provides valuable information about ligand-receptor association. Calculation of binding free energies considered 1500 snapshots from each 10ns trajectory with the average values of the trajectories being computed. From each snapshot, binding free energy ( $\Delta G_{\text{bind}}$ ) was computed from the following:

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

$$(2) \quad E_{\text{gas}} = E_{\text{int}} + E_{\text{vdw}} + E_{\text{ele}}$$

$$(3) \quad G_{\text{sol}} = G_{\text{GB}} + G_{\text{SA}}$$

$$(4) \quad G_{\text{SA}} = \gamma \text{SASA}$$

Where  $E_{\text{gas}}$ , the gas-phase energy, was calculated using the FF99SB force field;  $E_{\text{int}}$  is the internal energy;  $E_{\text{ele}}$  and  $E_{\text{vdw}}$  are the Coulomb and van der Waals energies, respectively.  $G_{\text{sol}}$  is the solvation free energy.  $G_{\text{GB}}$  is the polar solvation contribution.  $G_{\text{SA}}$  is the non-polar solvation contribution and was estimated by the solvent accessible surface area (SASA) determined using a water probe radius of 1.4 Å. The surface tension constant  $\gamma$  was set to 0.0072 kcal/(mol·Å<sup>2</sup>).

## 5.3 Results and Discussion

### 5.3.1 Binding Site Prediction and Validation

The binding site of CTLA-4 was predicted using site map <sup>30</sup>, and cross-validated with the aid of MetaPocket site identification tool <sup>31</sup>. A binding pocket was predicted, and constitutive residues are shown in Table 5.1. This binding pocket was chosen based on the drugability (Dscore) score, which measures the ability of a protein to tightly bind small molecules.

Table 5.1: Constituent residues of predicted CTLA-4 binding site

Biological target	Binding site identification (residues)	
	Site Map	MetaPocket
CTLA-4	Tyr24, Ala25, Ser26, Gly28, Lys29, Ala30, Thr31, Val33, Tyr53, Met55, Gly56, Ser69, Gly71, Asn72, Leu95	Tyr24, Val33, Tyr53, Met55, Gly56, Gly73, Asn74, Ala30, Ser71, Ser72, Thr31, Gln75, Lys29, Gly28, Leu97, Ser26, Ala25, Met54, Asn57

According to the methods of Heinrich 2010, CTLA-4 was queried on a binding database ([www.bindingdb.org](http://www.bindingdb.org)) to obtain proteins similar in binding site to CTLA-4 <sup>13</sup>. In this regard, a protein with an identical binding site to CTLA-4 was identified, which had a small molecule; 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones (*TDSO*) bound to it. Moreover, in relation to CTLA-4, binding site residues in Kallikrein 7 include Asn189, Gly193, Ser146, Val149, Ala183, and Thr96 (Figure 5.2) indicative of a unique similarity.

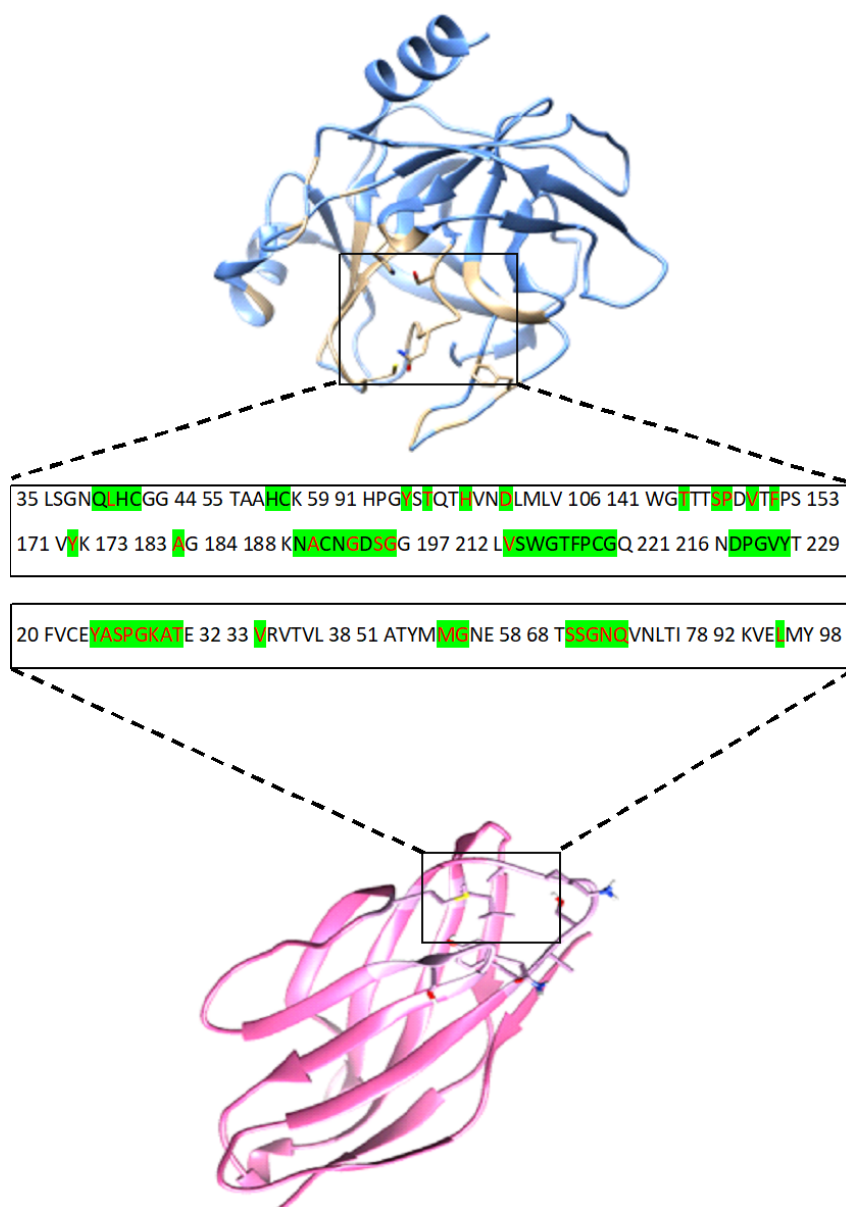


Figure 5.2: Graphical representation of kallikrein and CTLA-4 binding site similarity between Kallikrein 7 and CTLA-4. Similar binding site residues are highlighted in red

### 5.3.1.2 Pharmacophore Model Creation

The similarity between CTLA-4 and TDSO-bound Kallikrein 7 (PDB: 5Y9L) binding sites informed its (TDSO) selection as the lead compound used in modeling novel antagonists of CTLA-4. To obtain complementary structural architecture of the steric and electronic features that were

important to ensure supramolecular interaction of TDSO with CTLA-4, an intermediate 20ns MD simulation was carried out on the TDSO-CTLA-4 complex. Afterwards, a pharmacophore model or theory was obtained as shown in FigS1, which highlights the fundamental pharmacophoric features that served as a starting point for virtual screening to identify potential CTLA-4 antagonists.

### 5.3.1.3 Generation of Ligands

The pharmacophoric moieties that exhibited the most significant interaction with complementary residues of the binding pocket were employed to screen for novel compounds with similar structural and chemical features in Zinc database (contains 4.6 million small molecules). In addition, the Lipinski's Rule of Five<sup>32</sup> was used as a search criteria to streamline the resulting compounds to drug-like substances, which were then screened virtually using Autodock Vina's with an exhaustiveness set at 8.0. The grid box coordinates *X*, *Y*, and *Z* are centered on 16.62, 18.77, and 17.57 respectively while generated results were saved in the pdbqt format. This pharmacophore-based query led to the generation of 9 structurally similar compounds that belong to the ZINC Purchasable Subset. Further selection was carried out based on comparative binding energies, which led to the identification and retrieval of two compounds with the highest negative binding energies (*ZINC04515726* and *ZINC08985213*) in relation to others (Fig.5.3).

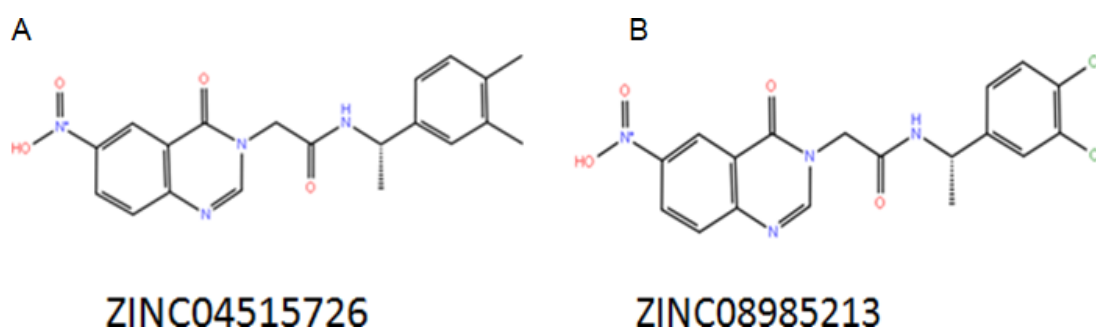


Figure 5.3: 2D structures of ZINC04515726 and ZINC08985213

### **5.3.2 Drug Likelihood Assessment of TDSO, ZINC04515726 and ZINC08985213**

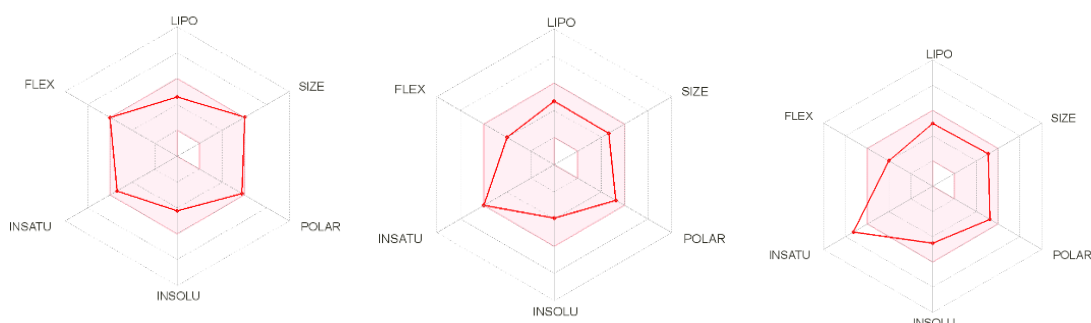
#### *5.3.2.1 ADME Result*

To access the physiochemical descriptors, pharmacokinetic properties and drug likelihood of *TDSO*, *ZINC04515726* and *ZINC08985213*, SwissADME <sup>19</sup> was used. In the design of therapeutics, the pharmacokinetic properties, safety, potency and selectivity are always considered. Lipophilicity (LogP) is a very essential feature to be considered in the interaction between a chemical compound and its biological target. The permeability of a compound across the lipid bilayer is a function of the value of its lipophilicity, hepatic clearance or solubility <sup>33</sup>. A chemical compound having a LogP value between 2 and 3 shows a highly favorable potential of achieving permeability and first pass clearance <sup>111</sup>. From Table 2 below, it is seen that *ZINC04515726*, *ZINC08985213* and *TDSO* have LogP values of 2.43, 2.38 and 2.53 respectively, indicating that these three compounds have the potential to achieve membrane permeability and first pass clearance. The gastrointestinal absorption of the compounds is high indicating that they can easily be absorbed by the GIT, which is favorable for their respective biological transport towards CTLA-4. Likewise, their inability to pass through the blood brain barrier (BBB) could indicate their suitability for targeting cancers cells not found in the central nervous system (CNS).

Table 5.2: Comparative analyses of the pharmacokinetic profile of TDSO, ZINC04515726, and ZINC08985213

	<b>TDSO</b>	<b>ZINC04515726</b>	<b>ZINC08985213</b>
<b>Molecular Formula</b>	$C_{24}H_{28}OCIN_5O_5$	$C_2OH_2ON_4O_4$	$C_{18}H_{14}C_{12}N_4O_4$
<b>Molecular Weight (g/mol)</b>	501.96	380.40	421.23
<b>Lipophilicity (LogP)</b>	2.53	2.43	2.38
<b>Water Soluble</b>	Soluble	Moderately Soluble	Poorly Soluble
<b>GIT absorption</b>	High	High	High
<b>BBB Permeability</b>	Not Permeable	Not Permeable	Not Permeable
<b>Bioavailability Score</b>	0.55	0.55	0.55
<b>Synthetic accessibility</b>	4.24	3.30	3.18
<b>Druglikeness (Lipinski)</b>	Yes	Yes	Yes

**“Bioled egg” structure**



### 5.3.3 Post-Molecular Dynamics Simulation Analysis

#### 3.3.1 Differential binding of TDSO, ZINC04515726 and ZINC08985213 to CTLA-4

The MM/PBSA is used to calculate the free energy difference between two different solvated conformations of the same molecule <sup>34</sup>. To determine the effect of inhibitor binding to CTLA-4, post-dynamic calculation of MM/PBSA was carried out. These binding affinities provide us with

the information of each amino acid residue toward the total binding affinity. As estimated, the three compounds exhibited favorable binding as evidenced by their high negative energies. From Table 3, TDSO had a  $\Delta G$  value of -28.51 kcal/mol while ZINC04515726 and ZINC08985213 have  $\Delta G$  values of -23.76 kcal/mol and -21.03 kcal/mol respectively, which could suggest strong and high affinity binding.

Table 5.3: MM/PBSA binding free energy profiles of TDSO, ZINC04515726, and ZINC08985213 to CTLA-4

	$\Delta E_{\text{ele}}$ (kcal/mol)	$\Delta E_{\text{vdw}}$ (kcal/mol)	$\Delta E_{\text{gas}}$ (kcal/mol)	$\Delta G_{\text{sol}}$ (kcal/mol)	$\Delta G_{\text{bind}}$ (kcal/mol)
TDSO	-6.35 ± 3.41	-36.01 ± 3.41	-42.36 ± 7.88	13.86 ± 6.80	-28.51 ± 3.57
ZINC04515726	-5.48 ± 5.51	-33.28 ± 4.70	-38.76 ± 6.02	15.00 ± 5.12	-23.76 ± 4.66
ZINC08985213	-8.41 ± 7.93	-26.18 ± 5.46	-34.59 ± 10.04	13.56 ± 7.73	-21.03 ± 5.43

$\Delta E_{\text{ele}}$  = electrostatic energy;  $\Delta E_{\text{vdw}}$  = van der Waals energy;  $\Delta G_{\text{bind}}$  = total binding free energy;  $\Delta G_{\text{sol}}$  = solvation free energy  $\Delta E_{\text{gas}}$  = gas phase free energy.

Furthermore, we investigated key binding site residues and their respective roles in the differential binding of TDSO, ZINC04515726 and ZINC08985213 towards CTLA-4. This was carried out by decomposing the  $\Delta G$  binding free energy into electrostatic, van der Waals and total energies using an MMPB/SA-integrated per-residue energy decomposition method. As shown in Figure 5A, 6A and 7A, Tyr24, Ala25, Gly28, Ala30, Tyr53 and Asn72 play crucial roles in the binding of the respective compounds as evidenced by considerably high electrostatic energy contributions. Tyr24, Ala25, Gly28, Ala30, Tyr53 and Asn72 had electrostatic contributions of -7.63, -8.23, -7.76, -8.76, -7.99 and -7.92 kcal/mol respectively towards TDSO, while in ZINC04515726, energy contributions of these residues were -7.4, -8.18, -7.27, -8.67, -7.6, and -80.4 kcal/mol respectively. These residues also had -7.63, -8.13, -7.75, -8.85, -7.58 and -7.68 kcal/mol electrostatic energy contributions towards ZINC08985213 binding to CTLA-4. Moreover, in the three systems, there were appreciable van der Waals energy contributed by Tyr24, Tyr53, and Leu95, TDSO had -1.34,



-1.41 and -1.1 kcal/mol contributed respectively. ZINC08985213 had -1.43, -1.41 and -1.06 kcal/mol contribution, while Tyr53, and Leu95 contributed mostly to the van der Waals energy in the binding of ZINC04515726 to CTLA-4 with energy contributions of 1.41 and 1.11 kcal/mol respectively.

The nature and types of the occurring interactions were further visualized with the aid of Discovery studio 2016 Client to gain additional insights into the mechanistic interactions of the respective ligands with key active site residues with respect to high-affinity binding and stability. As shown in Figures 5.4B, 5.5B and 5.6B, strong hydrogen bonds (NH--O, CH--O and OH--O) as well as weak pi-pi stacked, pi-alkyl and pi-sulfur interactions were observed, which altogether contributed to the binding of the respective ligands to CTLA-4. These high-affinity interactions could also account for the high electrostatic energies observed on the per residue decomposition plot which altogether could possibly underlie the high-affinity inhibitory activity of these compounds towards CTLA-4. Although the mechanisms of charge transfer or complementary bond formation were not investigated in this study, we propose that the use of other quantitative techniques such as QM or QM/MM in future studies would provide more insights.

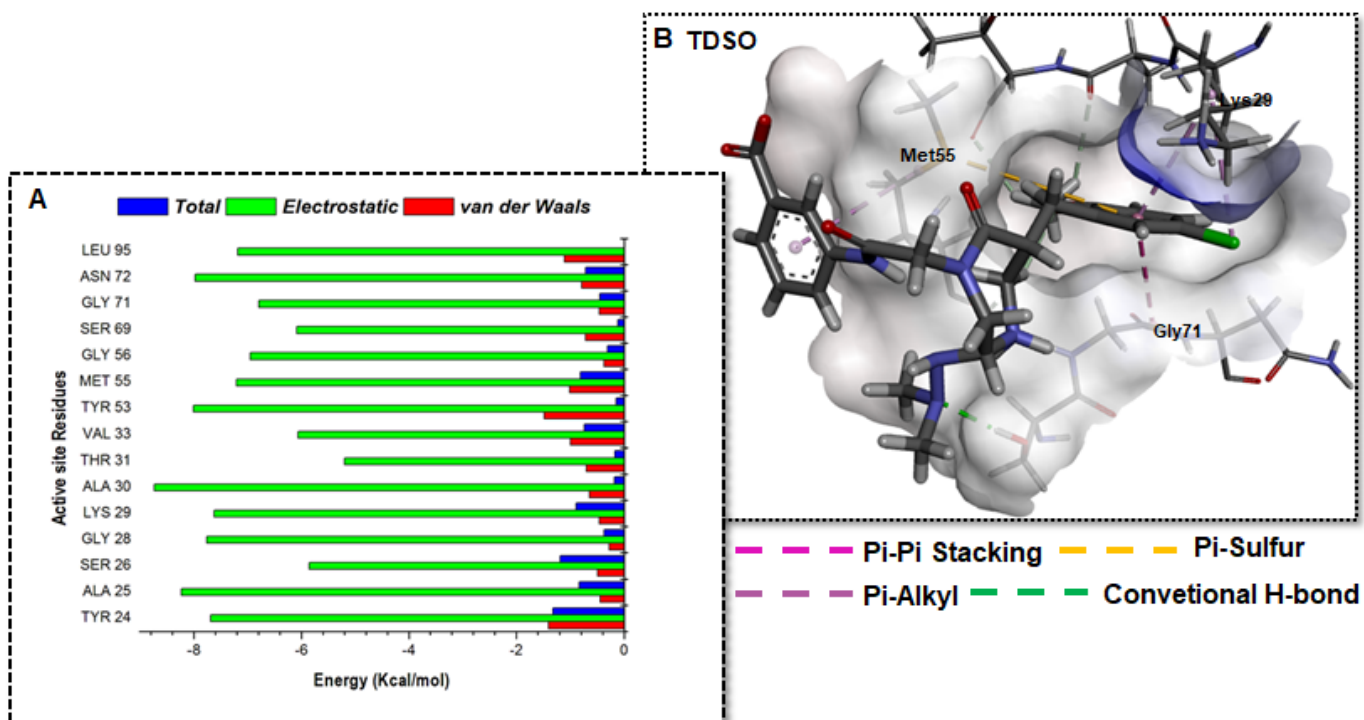


Figure 5.4: Individual energy contributions of crucial binding site residues to high-affinity binding and stabilization of TDSO. [A] Per-residue decomposition plot showing energy contributions of interacting residues [B] Molecular interactions between essential residues and reactive substituent on TDSO.

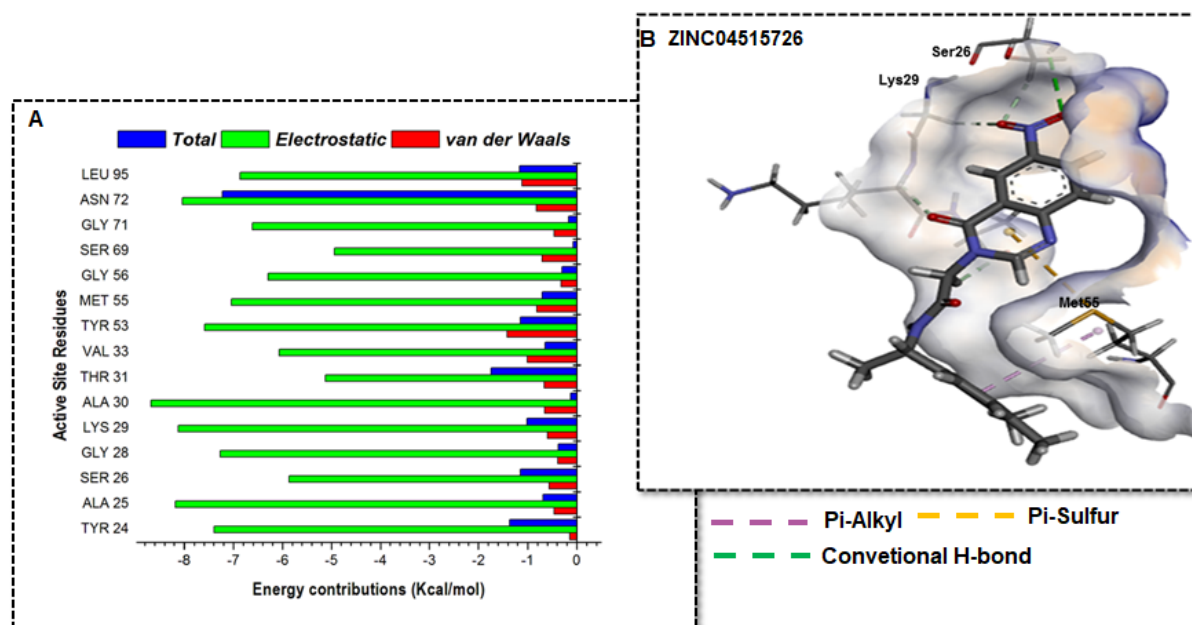


Figure 5.5: Individual energy contributions of crucial binding site residues to high-affinity binding and stabilization of ZINC04515726. [A] Per-residue decomposition plot showing energy contributions of

interacting residues [B] Molecular interactions between essential residues and reactive substituent on ZINC04515726.

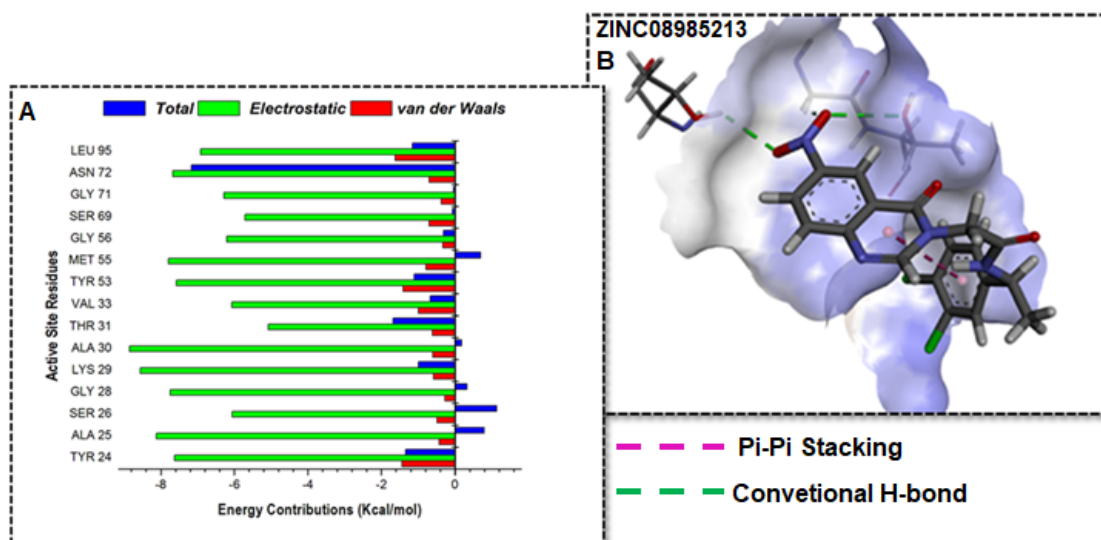


Figure 5.6: Individual energy contributions of crucial binding site residues to high-affinity binding and stabilization of ZINC04515726. [A] Per-residue decomposition plot showing energy contributions of interacting residues [B] Molecular interactions between essential residues and reactive substituent on ZINC04515726.

### 5.3.3.2 Structural elucidation in the binding of TDSO, ZINC04515726, and ZINC08985213 to CTLA-4

To understand the structural events associated with the inhibitory activity of TDSO, ZINC04515726, and ZINC08985213 towards CTLA-4, it was necessary to analyze the structural stability of the simulated systems. As such, the root of mean square deviation (RMSD) of C- $\alpha$  atom across the 150ns simulation period was monitored. As shown in Figure 5.7A, the systems converged early in the simulation period until about 10ns where the systems showed distinct separations in motions. Whole structural RMSD analyses revealed that unbound CTLA-4 exhibited the highest C- $\alpha$  deviation (RMSD  $\geq 2\text{\AA}$ ) compared to the bound protein-forms which had lower motions (Figure 5.7A). Moreover, the binding of the respective ligands elicited notable structural perturbations on the proteins correlative of their conformational instabilities. This could

also have possible effects on protein motions thereby impeding their involvements in pathologic interactions. Estimated mean values for unbound CTLA-4 was 1.48Å while values of 1.05Å, 1.16Å and 1.13Å were estimated for TDSO-, ZINC08985213- and ZINC04515726-bound systems respectively.

We also investigated the impacts of the respective compounds on the architecture of CTLA-4 binding pocket with respect to their antagonistic roles. As presented in Figure 5.7B, atomistic motions were higher at the active site of unbound CTLA-4 while the presence of the ligands lowered C- $\alpha$  motions of constituent active site residues. This could be due to possible ‘pulling-effects’ elicited by complementary ligand-residual interactions that are crucial for high-affinity binding (Figure 5.7B).

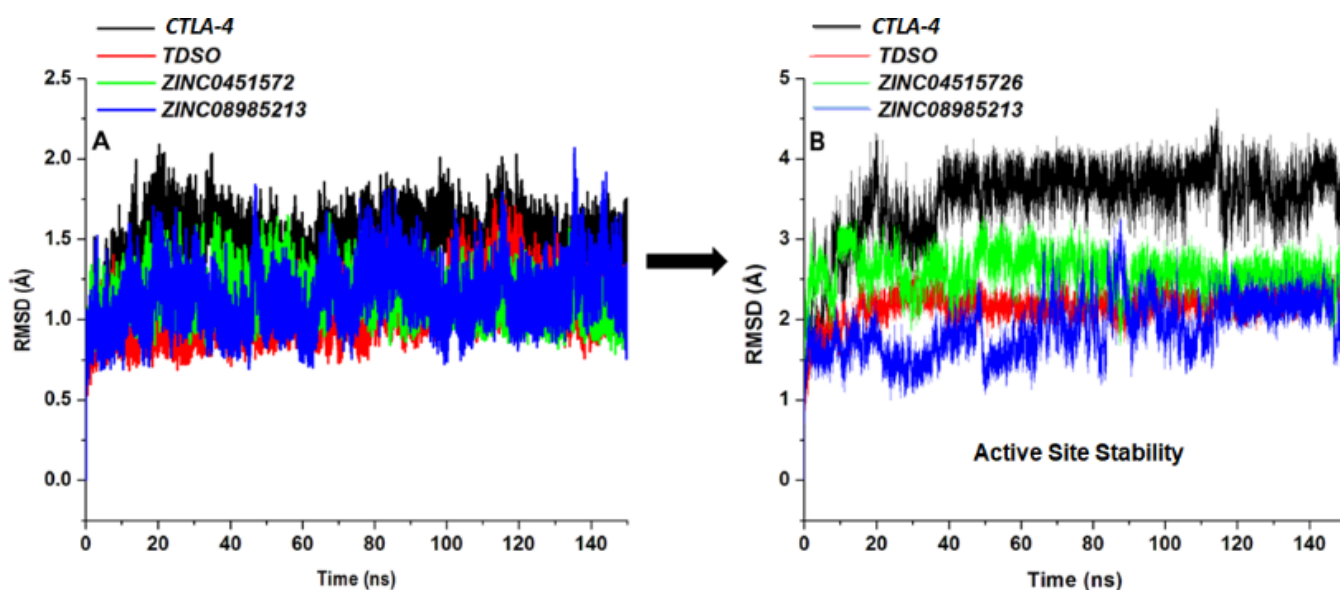


Figure 5.7: Conformational analysis plot showing stability and atomistic motions among unbound (**black**), TDSO (**red**), ZINC04515726 (**green**) and ZINC08985213 (**blue**) bound systems. [A] Whole-structure comparative C- $\alpha$  RMSD plot of the four systems.

The motions of the respective ligands at the binding pockets of CTLA-4 was also estimated as presented in Figure 5.8. Findings revealed that the compounds showed uniform and stable motions

until about 90ns where the patterns of motions appeared distinct. While TDSO showed a more stable motion at the binding pocket, the identified hits; ZINC04515726 and ZINC08985213, demonstrated similar alterations in motions. Consequently, the smaller sizes of both hits, relative to TDSO could favor their optimal motions at the binding cavity.

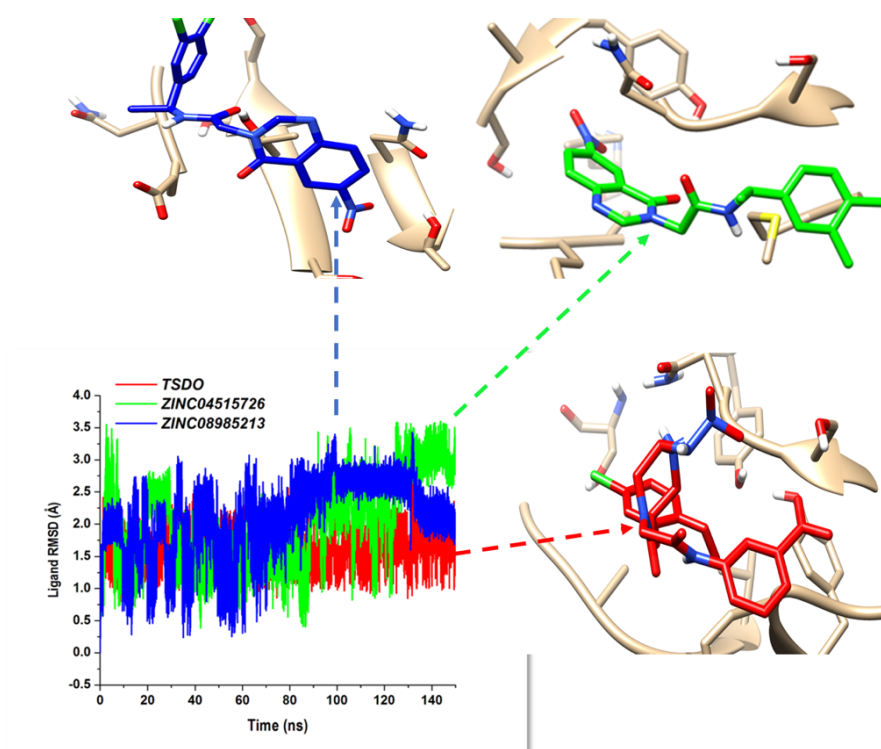


Figure 5.8: [A] C- $\alpha$  RMSD plot showing active site stability among unbound (**black**), TDSO (**red**), ZINC04515726 (**green**) and ZINC08985213 (**blue**) bound systems. [B] Comparative C- $\alpha$  RMSD plot depicting the stability and motion of TDSO (**red**), ZINC04515726 (**green**) and ZINC08985213 (**blue**). [C] Superimposition of TDSO (**red**), ZINC04515726 (**green**) and ZINC08985213 (**blue**)

Under physiological conditions, proteins like any other molecules undergo fluctuations in their structures<sup>35</sup>. Therefore, the dynamics of a protein is equally as important as its structure. The interaction with a protein active site can alter the protein's functions. RMSF is calculated to determine the mobility of individual residues within a protein<sup>22</sup>. To understand and explore the structural dynamics (fluctuations) that occur upon ligand binding, RMSF was carried out on the systems. Figure 5.9 clearly shows high residual fluctuations of the ligand when compared to

CTLA-4. However, TDSO had a lower flexibility from residues 80 – 114. This residual fluctuation is possibly due to the impact of the ligands upon binding.

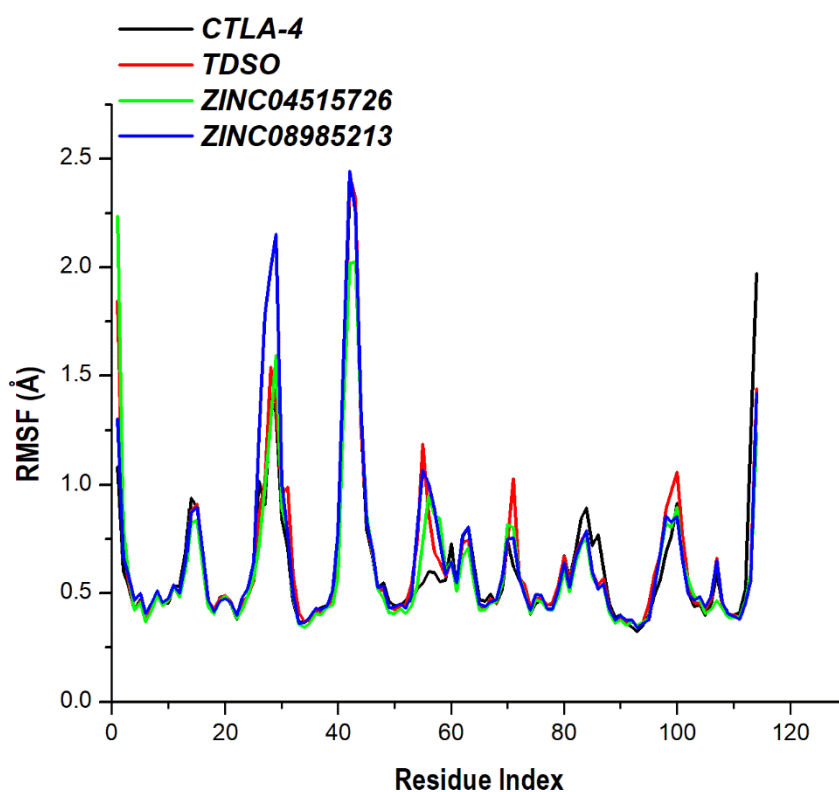


Figure 5.9: Structural flexibility induced by the binding of TDSO (red), ZINC04515726 (green) and ZINC08985213 (blue) to unbound (black).

#### 5.4 Conclusion

Antibodies have always been used in restoring T-Cell exhaustion, but due to the financial cost incurred in the production of these antibodies, small molecule inhibitors are thought to be a viable option to monoclonal antibodies. A drug repositioning and pharmacophore modeling approach was used to discover 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones which is an inhibitor of Kallikrien-7. The chemoinformatic analysis of ZINC04515726 and ZINC08985213 showed they are drug-like having obeyed the Lipinski rule of five. MMGBSA revealed that 1, 3, 6-trisubstituted

1, 4-diazepane-7-ones has the highest binding energy value compared to ZINC04515726 and ZINC08985213. The aforementioned results suggest that 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones can be a good anti-CTLA-4.

### **Conflict of interest**

The authors declare none.

### **Acknowledgement**

The authors acknowledge the College of Health Sciences, UKZN for their financial and infrastructural support and at the same time thank the Centre for High Performance Computing (CHPC, [www.chpc.ac.za](http://www.chpc.ac.za)), Cape Town, for computational resources

### **Reference**

1. Chatamra, K. Cancer and immunology. *Asian Pacific J. Allergy Immunol.* **9**, 71–73 (1991).
2. Mc Granahan, N. *et al.* Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351**, 1463–1469 (2016).
3. Soremekun, O. S., Olotu, F. A., Agoni, C. & Mahmoud, E. S. Recruiting monomer for dimer formation : resolving the antagonistic mechanisms of novel immune check point inhibitors against Programmed Death Ligand-1 in cancer immunotherapy. *Mol. Simul.* **0**, 1–13 (2019).
4. Chen, G., Bodogai, M., Tamehiro, N., Shen, C. & Dou, J. Cancer Immunotherapy: Theory and Application. *J. Immunol. Res.* **2018**, 1–2 (2018).
5. Marin-Acevedo, J. A. *et al.* Next generation of immune checkpoint therapy in cancer: new developments and challenges. *J. Hematol. Oncol.* **11**, 39 (2018).
6. Thallinger, C. *et al.* Review of cancer treatment with immune checkpoint inhibitors Current concepts, expectations, limitations and pitfalls. *Wien Klin Wochenschr* **130**, 85–91 (2018).

7. Shen, M. & Ren, X. Highlights on immune checkpoint inhibitors in non-small cell lung cancer. *Tumor Biol.* **39**, (2017).
8. Bielinska, A. U. *et al.* Distinct Pathways of Humoral and Cellular Immunity Induced with the Mucosal Administration of a Nanoemulsion Adjuvant. *J. Immunol.* **192**, 2722–2733 (2014).
9. Barbee, M. S., Ogunniyi, A., Horvat, T. Z. & Dang, T. O. Current status and future directions of the immune checkpoint inhibitors ipilimumab, pembrolizumab, and nivolumab in oncology. *Ann. Pharmacother.* **49**, 907–937 (2015).
10. Aspeslagh, S., Marabelle, A., Soria, J.-C. & Armand, J.-P. Upcoming innovations in lung cancer immunotherapy: focus on immune checkpoint inhibitors. *Chinese Clin. Oncol.* **4**, 48 (2015).
11. Zhan, M. M. *et al.* From monoclonal antibodies to small molecules: The development of inhibitors targeting the PD-1/PD-L1 pathway. *Drug Discov. Today* **21**, 1027–1036 (2016).
12. Ramagopal, U. A. *et al.* Structural basis for cancer immunotherapy by the first-in-class checkpoint inhibitor ipilimumab. *Proc. Natl. Acad. Sci.* **114**, E4223–E4232 (2017).
13. Liu, T., Lin, Y., Wen, X., Jorissen, R. N. & Gilson, M. K. BindingDB: A web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res.* **35**, 198–201 (2007).
14. Henrich, S. *et al.* Computational approaches to identifying and characterizing protein binding sites for ligand design. *J. Mol. Recognit.* **23**, 209–219 (2010).
15. Cele, F., Ramesh, M. & Soliman, M. Per-residue energy decomposition pharmacophore model to enhance virtual screening in drug discovery: a study for identification of reverse transcriptase inhibitors as potential anti-HIV agents. *Drug Des. Dev. Ther.* **10**, 1365–1377



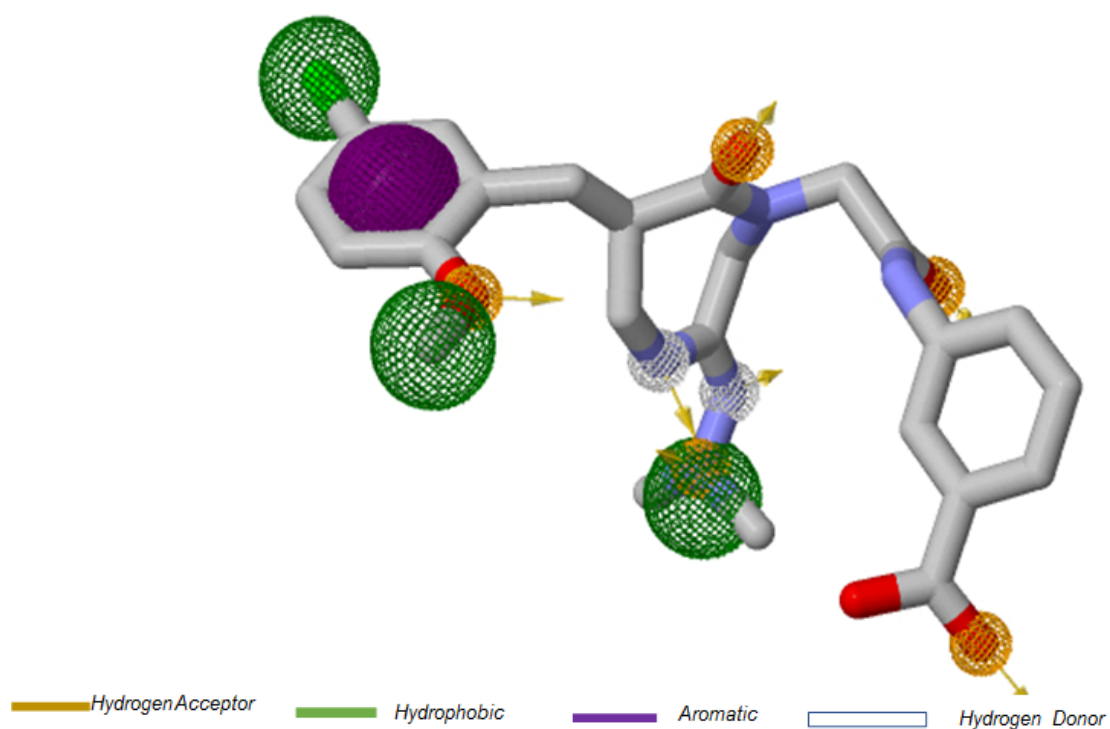
- (2016).
16. Kumalo, H. M. & Soliman, M. E. Per-Residue Energy Footprints-Based Pharmacophore Modeling as an Enhanced In Silico Approach in Drug Discovery: A Case Study on the Identification of Novel  $\beta$ -Secretase1 (BACE1) Inhibitors as Anti-Alzheimer Agents. *Cell. Mol. Bioeng.* **9**, 175–189 (2016).
  17. Koes, D. R. & Camacho, C. J. ZINCPharmer : pharmacophore search of the ZINC database. **40**, 409–414 (2012).
  18. Irwin, J. J. & Shoichet, B. K. for Virtual Screening. **45**, 177–182 (2006).
  19. Daina, A., Michielin, O. & Zoete, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **7**, 1–13 (2017).
  20. Skalniak, L. *et al.* Small-molecule inhibitors of PD-1/PD-L1 immune checkpoint alleviate the PD-L1-induced exhaustion of T-cells. *Oncotarget* (2017). doi:10.18632/oncotarget.20050
  21. Yang, Z. *et al.* UCSF Chimera, MODELLER, and IMP: An integrated modeling system. *J. Struct. Biol.* **179**, 269–278 (2012).
  22. Hanwell, M. D. *et al.* Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* **4**, 1–17 (2012).
  23. Windows, M. M. V & X, M. O. S. molegro molecular viewer user manual. (2010).
  24. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–61 (2010).
  25. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and

- analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
26. Dominguez, C., Boelens, R. & Bonvin, A. M. J. J. HADDOCK : A Protein - Protein Docking Approach Based on Biochemical or Biophysical Information. 1731–1737 (2003).
  27. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *Joufile///C/Users/HP/Desktop/CTLF/Calculating Struct. Free. Chem. Phys.* **79**, 926–935 (1983).
  28. David A. Case. AmberTools12 Reference Manual. *Russell J. Bertrand Russell Arch.* 535 (2012).
  29. Kollman, P. A. *et al.* Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* **33**, 889–897 (2000).
  30. Halgren, T. A. Identifying and Characterizing Binding Sites and Assessing Druggability. 377–389 (2009).
  31. MetaPocket : A Meta Approach to Improve Protein Ligand Binding Site Prediction. **13**, 325–330 (2009).
  32. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Deliv. Rev.* **23**, 3–25 (1997).
  33. Sciences, H. DETERMINATION OF THE PERMEABILITY OF BIOLOGICAL MEMBRANES TO VARIOUS CHEMICAL MARKERS , INCLUDING ANTI-HIV DRUGS Erina Pretorius ( Basson ) December 2009. (2009).
  34. Genheden, S. & Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-

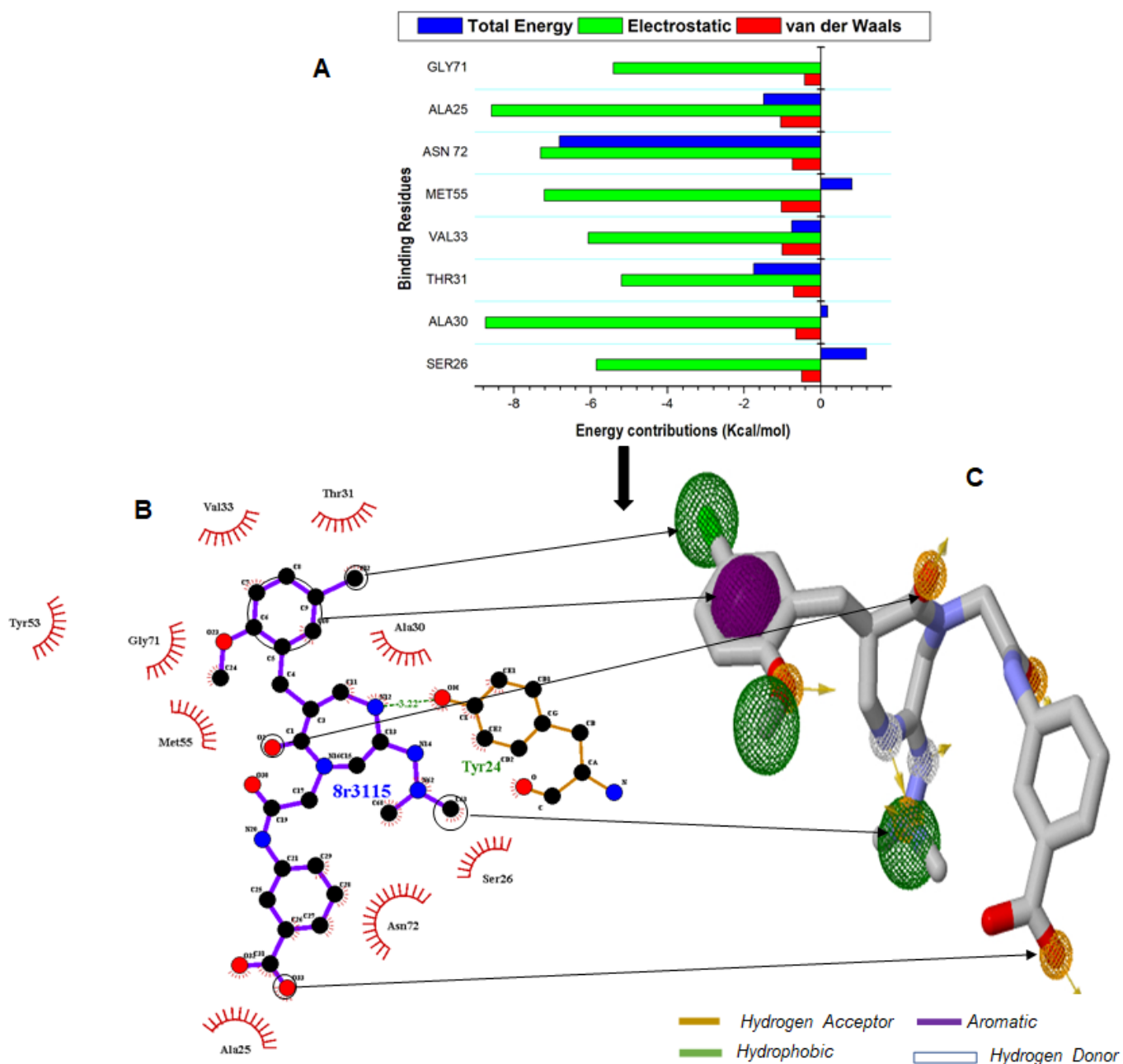
binding affinities. *Expert Opin. Drug Discov.* **10**, 449–461 (2015).

35. Fuglebakk, E., Echave, J. & Reuter, N. Measuring and comparing structural fluctuation patterns in large protein datasets. *Bioinformatics* **28**, 2431–2440 (2012).

### Supplementary



FigS1: 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones showing the interacting pharmacophoric moieties after a 20ns simulation with CTLA-4



FigS2: PRED contribution of different residues at the active site (A). 2D ligand interaction plot (B), and (C) Pharmacophore features responsible for FBE contributions

Table S1: Binding energy scores of TDSO, ZINC04515726, ZINC08985213 with TDSO

---

	<b>Chimera</b>	<b>HADDOCK</b>
<b>TDSO</b>	-5.3	-5.4
<b>ZINC04515726</b>	-5.5	-5.4
<b>ZINC08985213</b>	-5.1	-5.3

---

## CHAPTER 6

### **From Genomic Variation to Protein Aberration: Mutational Analysis of Single Nucleotide Polymorphism present in *ULBP6* gene and Implication in Immune Response**

Opeyemi S.Soremekun<sup>a</sup> and Mahmoud E. S. Soliman<sup>a\*</sup>

<sup>a</sup>Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

\*Corresponding Author: Mahmoud E.S. Soliman

Email: [soliman@ukzn.ac.za](mailto:soliman@ukzn.ac.za)

<https://soliman@ukzn.ac.za>

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872

## **Abstract**

**Background:** Genetic polymorphisms have been identified as one of the underlying factors in disease pathogenesis and drug resistance since they account for protein dysfunctionality, or in some cases, aberrancy. This explains the high degree of inactivity that characterize the polymorphic variants of ULBP6 binding protein, which in turn disrupts its primary interaction with human Natural Killer Group 2-member D (NKG2D) and accounts for an impediment to immuno-surveillance. The possible identification of deleterious non-synonymous Single Nucleotide Polymorphisms (nsSNPs) present in the *ULBP6* gene is essential for the development of novel gene therapies to prevent the translation of dysfunctional protein variants.

**Methods/Results:** In this study, for the first time, we employed an SNP-informatics approach (SNPs retrieval, pathogenic/mutational analysis, phenotypic analysis and structural analysis) and molecular dynamics techniques to identify and characterize undesirable SNPs coupled with their impact on ULBP6 structural activities relative to dysfunctionality. V52F was predictively pathogenic amongst SNPs studied. Conformational and dynamic studies revealed that in comparison to wildtype ULBP6 (ULBP6<sup>wt</sup>), pathogenic ULBP6<sup>V52F</sup> demonstrated considerable structural inactivity, which could in turn impede biological protein-protein interactions. Moreover, ULBP6<sup>V52F</sup> showed relatively limited motions in the conformational space as deduced from estimations of structural stability, fluctuations and principal components. **Conclusion:** This study provides a workable paradigm for investigating pathological nsSNPS using computational platforms which findings present ULBP6<sup>V52F</sup> as a novel and attractive immunotherapeutic target in combatting immune associated disorders. This study also provides an insight into future investigations of pathological nsSNPS using computational platforms.

**Keywords:** Single Nucleotide Polymorphism, ULBP6, Bioinformatics, Molecular Dynamic Simulation, Mutation

## 6.1 Introduction

The human immune defense is a well-developed system that has evolved over time in the protection of the body against invading pathogenic assaults and toxins (1). Basically, there are four mechanisms with which this protection is elicited. Firstly, through physical barrier such as skin and mucosal tissues that protects internal tissues from assault (2). The second line of defense is mediated by antimicrobial enzymes such as lysozyme, anti-microbial peptides such as defensins (3), and the action of the complement system, this line of defense helps to initiate the third line of defense, the innate immune system. Complement binding to a pathogen marks it for destruction by phagocytes, which initiate an inflammatory response (4). Inflammation recruits monocytes and dendritic cells, which through activation of their Toll-like receptors (TLRs) recruit a range of cytotoxic effector cells such as neutrophils, lymphocytes and natural killer (NK) cells (5). These cells elicit the last defense mechanism, which is recruitment of cells that specifically target pathogen and form immune memory in case of repeat infection (6). NK cells possess surface receptors that recognize ligands that are up-regulated on the surface of infected or transformed cells (5).

Natural Killer group 2, member D (NKG2D), is one of the receptor that is expressed on the surface of NK cells, (7), NKG2D is activated upon binding to its complementary ligands, MHC class-I chain related proteins (MIC) A, MICB and UL-16-binding proteins (ULBP1-6) (8). The ULBP6 gene cluster is found on chromosome 6q24.2-q25.3 and comprises of ten loci (9). ULBP1-6 is one of the recently discovered human NKG2D ligands and is highly polymorphic, with two variants (*ULBP0601* and *ULBP0602*)(8). This high polymorphic tendencies is probably due to evolutionary attempts to maintain viral resistance, and cancer immunoediting (10) Eagle et al., reported the



absence of ULBP6 in primary human fibroblasts, however, upon infection with two different strains of HCMV, ULBP6 transcript was expressed (11). ULBP6 polymorphisms have also been implicated in diabetic nephropathy (12), alopecia areata (AA) (13) and endometriosis (8).

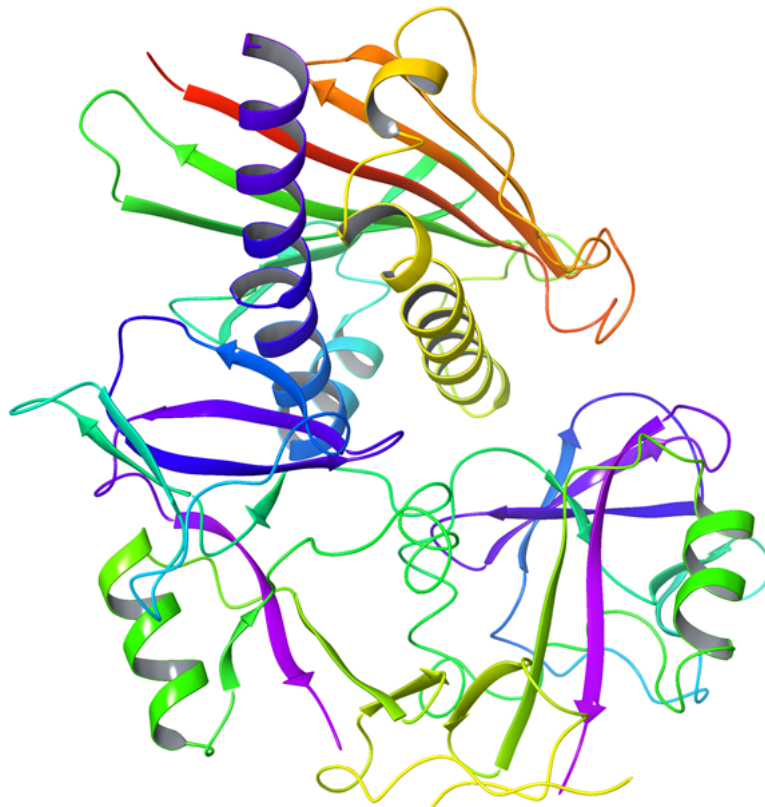


Figure 6.1: 3-D crystallography structure of ULBP6 protein in complex with NKG2D (PDB ID: 4S0U (14)

In the binding of ULBP6 and NKG2D, some amino acids play crucial role in enhancing the binding potential of the two proteins. Zuo et al., 2017, describes this binding and highlighted K186, E201, K197 of NKG2DA as forming hydrogen bond interactions with residues E154, Y184 and D189 of ULBP6. Likewise, in the NKG2B monomer, residues S151, N207 and K197 formed hydrogen interaction with D99, E103 and R44 of ULBP6 residues (14). Furthermore, binding analysis revealed that the binding of NKG2D to ULBP6 is quite similar to ULBP3 (15).

Single nucleotide polymorphisms (SNPs) account for most genetic variation observed in humans (16). This genetic disparity caused by SNPs occurring at genomic exons alters the translation result leading to structural and functional changes in the mutated protein. However, not all SNPs affect the structure and function of a protein, some are pathogenic while others nonpathogenic (16). Taking cognizant of the role ULBP6 play in immune response, a study of the impact of its polymorphism will contribute immensely to scientific research. Therefore, in this study we used computational and bioinformatics analysis (Fig 6.2) to map out the most deleterious and disease-associated SNP in the *ULBP6* gene, putting into account their structural consequences at the molecular level. We used SIFT (17), PolyPhen (18), PhD-SNP (19), PMut (20) and SNP&GO (21) tools to prioritize the deleterious disease-associated SNPs from the available SNP datasets obtained from dsSNPs database. To examine the molecular and structural basis of predicted disease-associated SNP, we then carried out molecular dynamics simulation of the native and mutant ULBP6.

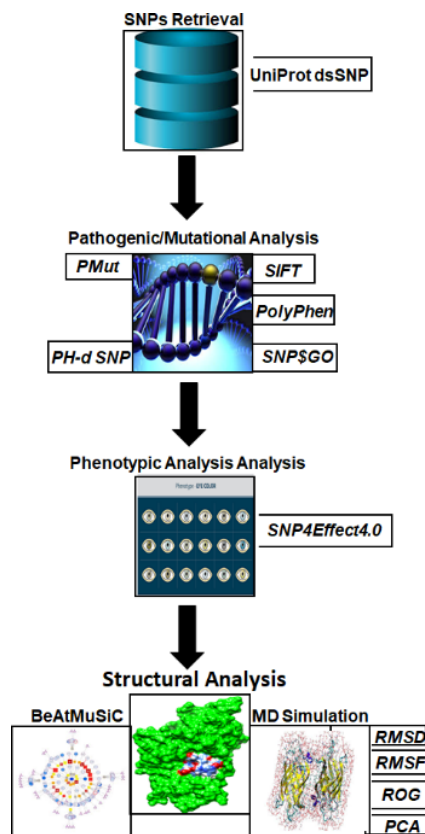


Figure 6.2: A schematic computational workflow

## 6.2 Material and Methods

### 6.2.1 SNPs Retrieval

Human ULBP6 (accession ID: Q5VY80) protein sequence was retrieved from Uniprot database (22). SNP data used for our computational analysis was obtained from dsSNPs (<http://www.ncbi.nlm.nih.gov/snp/>) (23). 3-D crystallography structure of ULBP6 was obtained from Protein Data Bank (PDB ID: 4S0U) (24). In order to build the mutant ULBP6, a point mutation was induced at position 52 of the wild ULBP6 protein using UCSF chimera (25).

### 6.2.2 Analysis of the functional impact of coding SNP

Single nucleotide polymorphism occurring in a protein coding region may impact some structural defect on the 3D structure and consequently leading to some pathological effects. We used SIFT (26) and Polyphen2 (27) to study the functional impact of the nsSNPs occurring in ULBP6. SIFT

is a sequence-based homology program that evaluates amino acids substitution (26). The input SNPs' rs-ID retrieved from nsSNPs database were submitted to SIFT server for analysis. SIFT uses the query sequence to execute multiple alignment in order to predict amino acids substitution present in the sequence and further classify them as either tolerated or deleterious. Prediction is given as tolerance index (TI) score ranging from 0.0 to 1. SNPs with TI score under 0.05 were regarded to be deleterious; those more pronounced than or equal to 0.05 were considered to be tolerated. SIFT has a sensitivity and specificity of 83% and 52% respectively (17). PolyPhen determines if the amino acid substitution in the sequence occurs in a conserved region and predicts if the substitution has any deleterious impact on the structure of the protein (27). PolyPhen employs the position-specific independent count (PSIC) score to ascertain the functional and structural impact of the substitution. A PSIC score of 1.0 is regarded as damaging. The SNPs are classified as benign, possibly damaging and probably damaging (27). To get an accurate prediction, we used nsSNPS that were jointly predicted by the two servers to be deleterious and damaging. Predicted damaging SNPs were further analyzed using SNP&GO (28), PhD-SNP (29). SNPs&GO retrieves data from protein sequence, evolutionary information, and functions as encoded in the gene ontology terms (28). It predicts if the SNP is neutral or pathological with the reliability index score (29). PhD-SNP is a support vector machine-based classifier that employs a supervised learning strategy to categorize the disease-causing mutation that are present in a query dataset. PhD-SNP is a neural network-based program which is trained on large database of neutral and pathological mutation. It employs three functions including, mutation parameters, solvent accessibility, and residue and sequence properties to calculate the pathogenicity index of a SNP ranging from 0 to 1. Mutation with index greater than 0.5 are regarded to be pathologically significant (19).

### 6.2.3 Impact of mutation on protein-protein interaction and Molecular Dynamics Simulation

Particle Mesh Ewald Molecular Dynamics - PME MD of the AMBER18 software in addition to its integrated modules were used to perform MD simulations (30). FF14SB force-field was used to define the proteins (31). ULBP6 modification, renaming and protonation (histidine) were done using an in-house *pdb4amber* script. Parameter and topology files were generated for ULBP6 using the LEAP module. This was also used for system neutralization. Restraint potential of 500kcal/mol Å, partial minimization of 2500 steps was carried out and a full minimization of 5000 steps. Gradual system thermalization from 0-300k was also carried out. System equilibration was then carried out for 1000ps at 300k without restraints while atmospheric pressure was kept constant at 1bar using the Berendsen barostat (32). Afterwards, an MD run of 100ns was carried out (33). At every 1ps, coordinates and trajectories generated were saved. They were further analysed using CPPTRAJ and PTRAJ (34).

### 6.2.4 Principal Component Analysis (PCA)

We used PCA to obtain necessary insights into the 3D conformational and dynamical changes that occurred in the wild and mutant proteins. This method is used to describe the magnitude and direction of protein motions which are depicted by eigenvalues and eigenvectors respectively (35,36). Hence, with the aid of the integrated CPPTRAJ module in AMBER18, we computed the first two principal components from the dynamics of the protein C- $\alpha$  atoms.

## 6.3 Results

### 6.3.1 Identification of deleterious and damaging SNPs

Identification of pathological genetic variants in proteins are useful in the foundation of genome level study especially in cancer and some neurodegenerative diseases, this provide a better insight for personalized and targeted based drug administration. We retrieved a total of 23 SNPs from the dbSNP database to identify deleterious point mutation in ULBP6 gene. We simultaneously run the 23 predicted SNPs on SIFT, PolyPhen, SNP&GO, Pmut and Phd-SNP. The SIFT server was used to determine the tolerance index of all 23 retrieved SNPs by evolutionary conservation analysis. A SIFT score value of  $<0.05$  was considered to be deleterious. Out of the input polymorphic data from the 23 SNPs, 8 were predicted to be deleterious with a tolerance index score less than or equal to 0.05. To further analyze the structural impact of SNPs, the 23 SNPs retrieved from dbSNP were submitted to the PolyPhen server. A total of 12 were predicted to be damaging having a PSIC score of  $>1.5$  while 11 were considered as benign. 2 SNPs were predicted to be disease causing with SNP&GO, 3 SNPs were predicted as diseased with Phd-SNP (Table 6.1). Taken together, only rs145336297 with V52F was jointly predicted by 4 out of the 5 tools as disease causing.

Table 6.1: Damaging and Deleterious SNPs predicted by SIFT, PolyPhen, SNP&GO, Pmut and Phd-SNP with their corresponding scores.

SNP ID	Amino Acid Change	SIFT	Score	PolyPhen	Score	SNP&GO	Score	PMut	Score	Phd-SNP	Score
rs1543547	R26G	Tolerated	1	Probably Damaging	0.976	Neutral	9	Neutral	0.02	Neutral	8
rs61730071	T147I	Tolerated	0.862	Benign	0.001	Neutral	9	Neutral	0.15	Neutral	7
rs61748301	M157V	Tolerated	0.112	Probably Damaging	0.636	Neutral	8	Neutral	0.22	Neutral	7
rs7989700	V74I	Tolerated	0.588	Probably Damaging	0.021	Neutral	9	Neutral	0.06	Neutral	5
rs137965307	T38I	Tolerated	0.106	Probably Damaging	0.984	Neutral	7	Neutral	0.15	Neutral	3
rs138504598	L62P	Deleterious	0.001	Probably Damaging	1.000	Disease	7	Neutral	0.34	Disease	0
rs139656748	F151L	Deleterious	0.001	Probably Damaging	0.998	Neutral	2	Neutral	0.08	Neutral	9
rs139983884	P76S	Tolerated	0.351	Benign	0.167	Neutral	8	Neutral	0.08	Neutral	1
rs140013602	I143I	Deleterious	0.037	Benign	0.020	Neutral	8	Neutral	0.10	Neutral	7
rs140239332	E58Ter	Tolerated	0.322	Probably Damaging	1.000	Neutral	8	Neutral	0.08	Neutral	9
rs141330153	D99N	Tolerated	0.062	Possibly Damaging	0.877	Neutral	7	Neutral	0.15	Neutral	6
rs141345751	G192E	Tolerated	1	Benign	0.005	Neutral	6	Neutral	0.01	Neutral	8
rs141551402	M1V	Deleterious	0	Benign	0.000	Neutral	9	Neutral		Neutral	8
rs141680152	L231F	Tolerated	0.741	Benign	0.001	Neutral	10	Neutral	0.12	Neutral	5
rs145336297	V52F	Deleterious	0.042	Probably Damaging	1.000	Disease	5	Neutral	0.08	Disease	0.73
rs145690445	A123E	Deleterious	0.003	Probably Damaging	0.999	Neutral	5	Neutral	0.23	Neutral	2
rs147008347	I191T	Tolerated	0.595	Benign	0.000	Neutral	8	Neutral	0.07	Neutral	7
rs147287797	M180T	Tolerated	0.348	Benign	0.027	Neutral	7	Neutral	0.02	Neutral	8
rs147068392	S208G	Deleterious	0.008	Probably Damaging	0.945	Neutral	10	Neutral	0.31	Neutral	4
rs147812292	P207T	Tolerated	0.645	Benign	0.086	Neutral	9	Neutral	0.09	Neutral	7
rs148198102	E173K	Deleterious	0.004	Probably Damaging	0.983	Neutral	7	Neutral	0.20	Disease	4
rs151006207	P76H	Tolerated	0.318	Benign	0.045	Neutral	8	Neutral	0.08	Neutral	5
rs151268095	C232Y	Tolerated	1.000	Benign	0.006	Neutral	7	Neutral	0.13	Neutral	6

### 6.3.2 Phenotype Analysis

SNPEffect4.0 tool (37) was used to detect the phenotypic changes induced by the V52F mutation. The results are predicted in the form of TANGO, LIMBO, WALTZ and ddG scores for the mutant protein.

#### LIMBO prediction

LIMBO is a chaperone binding site predictor of Hsp70 chaperones, designed from peptide binding data and structural modeling. The total LIMBO score for ULBP6 is 1409.53 and mutation can increase ( $dLIMBO > 50$ ), decrease ( $dLIMBO < -50$ ) or not affect chaperone binding ( $dLIMBO$  between  $-50$  and  $50$ ). In this case,  $dLIMBO$  equals  $0.00$  which means that the mutation does not affect the chaperone binding tendency of ULBP6. In figure 6.3.9 and 6.3.10, the position of LIMBO stretches in the wild type and variant proteins can be seen, represented by a bar or profile representation. In table 2, the short stretches are listed for both wild type and mutant. To compare the effect of the mutation to the WT, we also show a difference profile (Figure 6.3C), that plots the difference between WT protein and the variant.

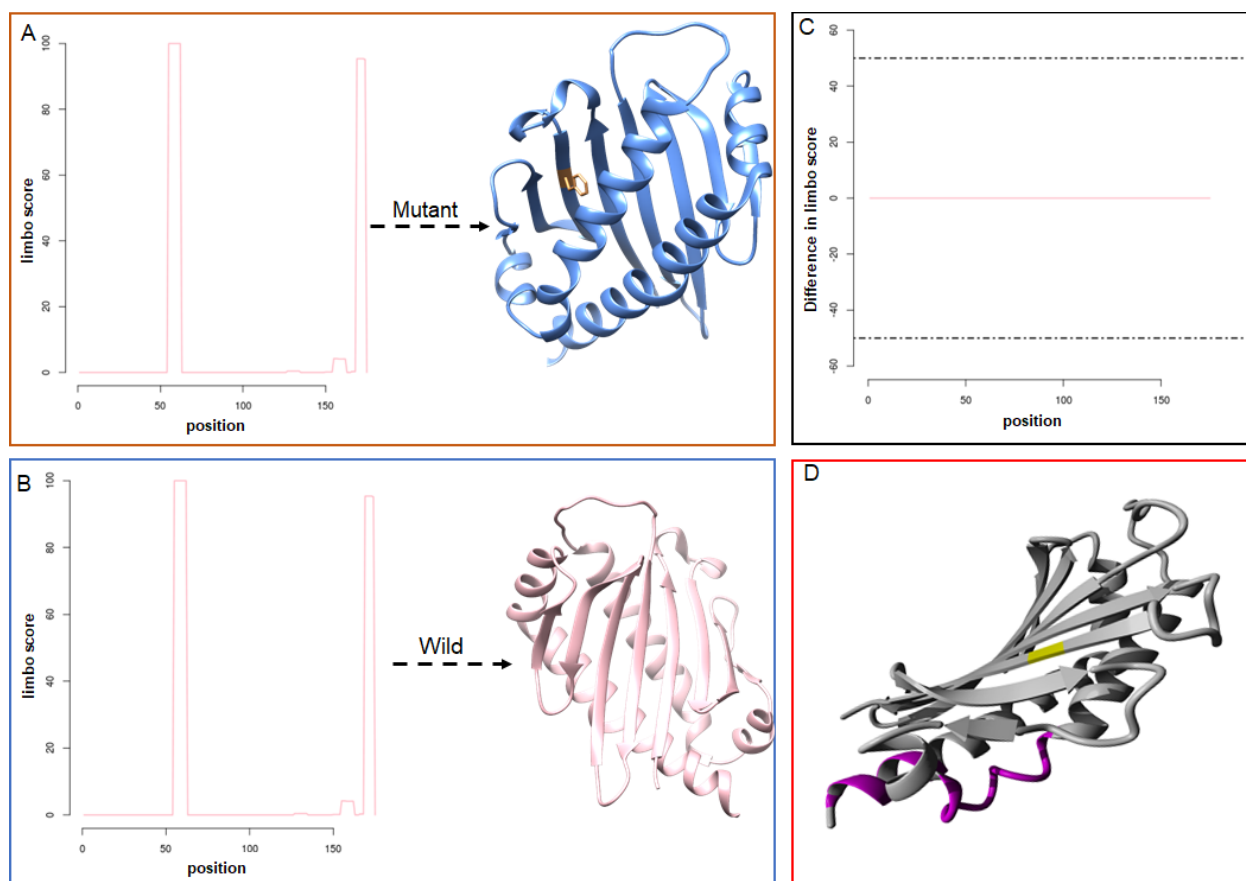




Figure 6.3: Profile representation of LIMBO stretches in the mutant (V52F) (A), Wild (B), difference in LIMBO chaperone propensity between the mutant and wild protein (C) and molecular visualization of LIMBO chaperone-binding sites colored in pink (D).

Table 6.2: LIMBO regions in variant and wild type. For each LIMBO region, the start, end, sequence and score are given.

Number	Start	End	Stretch	Score
<b>Wild Type</b>				
1	54	62	VTMAWKAQ	99.99
2	168	174	FLMGMD	95.34
<b>Mutant</b>				
1	54	62	VTMAWKAQ	99.99
2	168	174	FLMGMD	95.34

#### *Waltz and Tango prediction*

WALTZ is an algorithm that accurately and specifically predicts amyloid-forming regions in protein sequences. The total WALTZ score predicted for ULBP6 is 618.25 and the impact of mutation can increase ( $D_{waltz} > 50$ ), decrease ( $D_{waltz} < -50$ ) or not affect amyloid propensity ( $D_{waltz}$  between -50 and 50). ULBP6 has a  $D_{waltz}$  score of 0.01, this suggests that the mutation has no effect on the amyloid propensity of ULBP6. We further showed the difference in WALTZ score between the mutant and native protein (Figure 6.4C). TANGO predicts the aggregation prone region in a protein sequence. The total TANGO score for ULBP6 is 234.67. Mutation can increase ( $dTANGO > 50$ ), decrease ( $dTANGO < -50$ ) or not affect aggregation propensity ( $dTANGO$  between -50 and 50). For V52F mutation,  $dTANGO$  equals 0.00, this signify that the mutation has no effect on the aggregation tendency of the protein. To compare the effect of the mutation to the native protein, we presented the difference in TANGO score in Figure 6.4C.

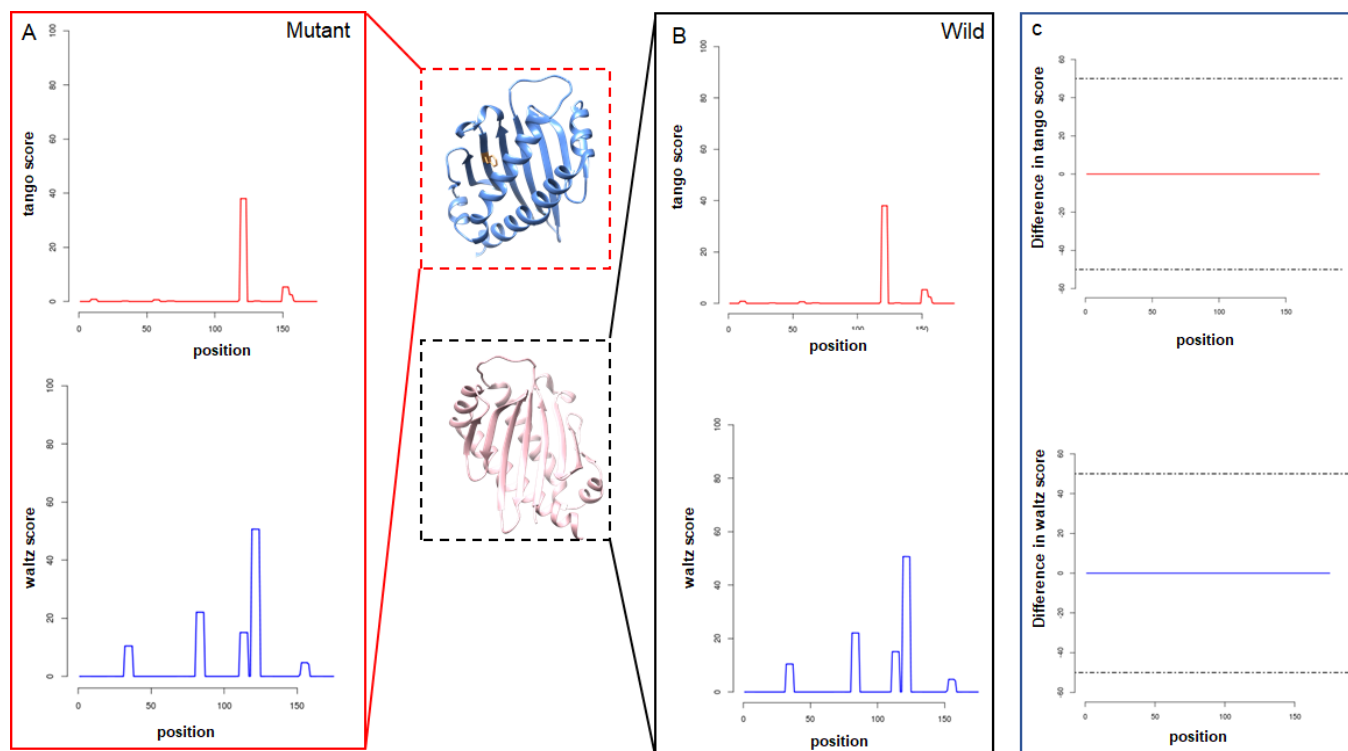


Figure 6.4: Profile representation of **WALTZ** and **TANGO** stretches present in the mutant (A), **WALTZ** and **TANGO** stretches present in the wild type (B). Difference in **WALTZ** amyloid propensity and difference in **TANGO** aggregation between WT and variant (C).

Table 6.3: **TANGO** regions in variant and wild type. For each **TANGO** region, the start, end, sequence and score are given.

Number	Start	End	Stretch	Score
Wild Type	No <b>TANGO</b> stretches present			
Mutant	No <b>TANGO</b> stretches present			

Table 6.4: **WALTZ** regions in variant and wild type. For each **WALTZ** region, the start, end, sequence and score are given.

Number	Start	End	Stretch	Score
<b>Wild Type</b>				
1	31	37	TFLHYD	10.45
2	80	86	QLENYT	2207
3	110	116	WQFSID	15.09
	118	124	TFLLFD	50.67
<b>Mutant</b>				
1	31	37	TFLHYD	10.45

2	80	86	QLENYT	22.07
3	110	116	WQFSID	15.09
4	118	124	TFLLFD	50.67

### 6.3.3 Investigation of ULBP6 Gene's interactions and Appearance in Networks in STRING Database

For better comprehension of a gene's function, knowledge of its specific interacting partners is important. We used STRING database (<http://string-db.org>) to study the interacting genes with ULBP6. STRING database provides a critical assessment and integration of protein-protein interactions, including direct (physical) as well as indirect (functional) associations. From the interaction network, ULBP6 has direct interaction with killer cell lectin-like receptor subfamily K member 1 (KLRK1) (Fig. 6.5).

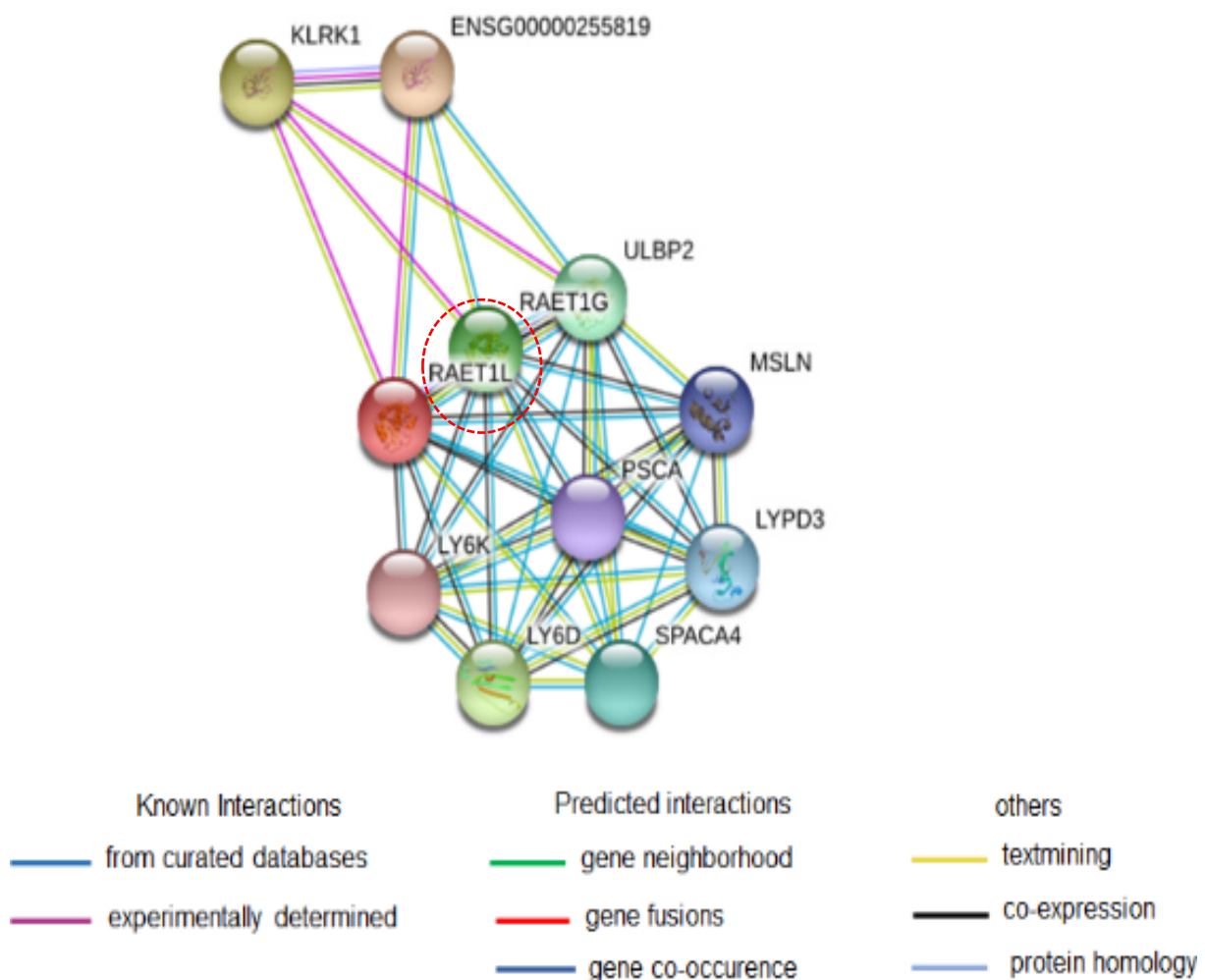


Figure 6.5: Functional interaction between ULBP6 and its related genes.

#### 6.3.4 *Impact of V52F polymorphism on structural integrity of ULBP6*

The formation of protein complex plays an essential role in the regulation of numerous biological processes (38). Availability of computational tools to evaluate the impact of mutations on protein-protein binding can therefore be a valuable tool (38). Here we used BeAtMuSiC to predict the effect of V52F mutations on NKG2D-ULBP6 binding. BeAtMuSiC is a coarse-grained predictor of changes in binding free energy induced by point mutation. BeAtMuSiC (38) predicted the binding affinity to be 1.79kcal/mol. The empirical protein design forcefield FoldX was used to calculate the difference in free energy of the mutation  $\Delta\Delta G$  (delta delta G). The mutation from Val to Phe at position 52 resulted in a  $\Delta\Delta G$  of 8.54kcal/mol. This implies that the mutation severely reduces the protein stability and alters the NKG2D-ULBP6 interaction. Residue interaction network was investigated using Cytoscape (39). The interaction network showed that residue V52 in the wild protein formed hydrogen bonds with Ser151 and Val 97, it also elicited van der Waals interaction with Met125 and Asp107, however in the mutant, phenylalanine formed hydrogen bond with Tyr64, Val98, Val39 (Fig 6.6). In order to investigate the changes in secondary structure of the proteins during the simulation period, we used STRIDE (40) to predict the secondary structure change at 10ns, 50ns and 100ns. As shown in Fig. 6.7, the major differences observed in the mutant in relative to the wild protein is observed in Pro176, Asn177 and Leu178. These residues were seen to change from a helix configuration in the wild protein to a coil in the mutant protein.

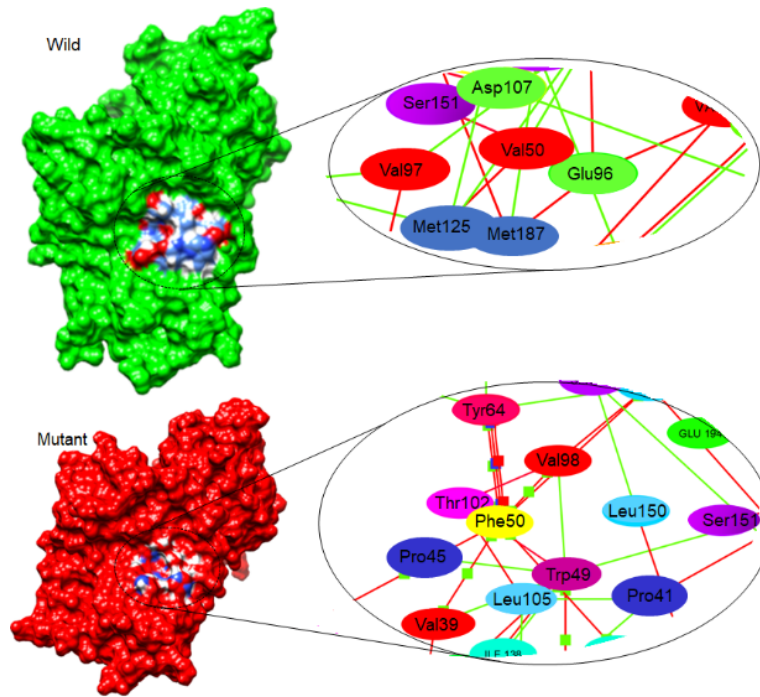


Figure 6.6: Residue interaction network analysis of the wild protein (**green**) and mutant protein (**red**).

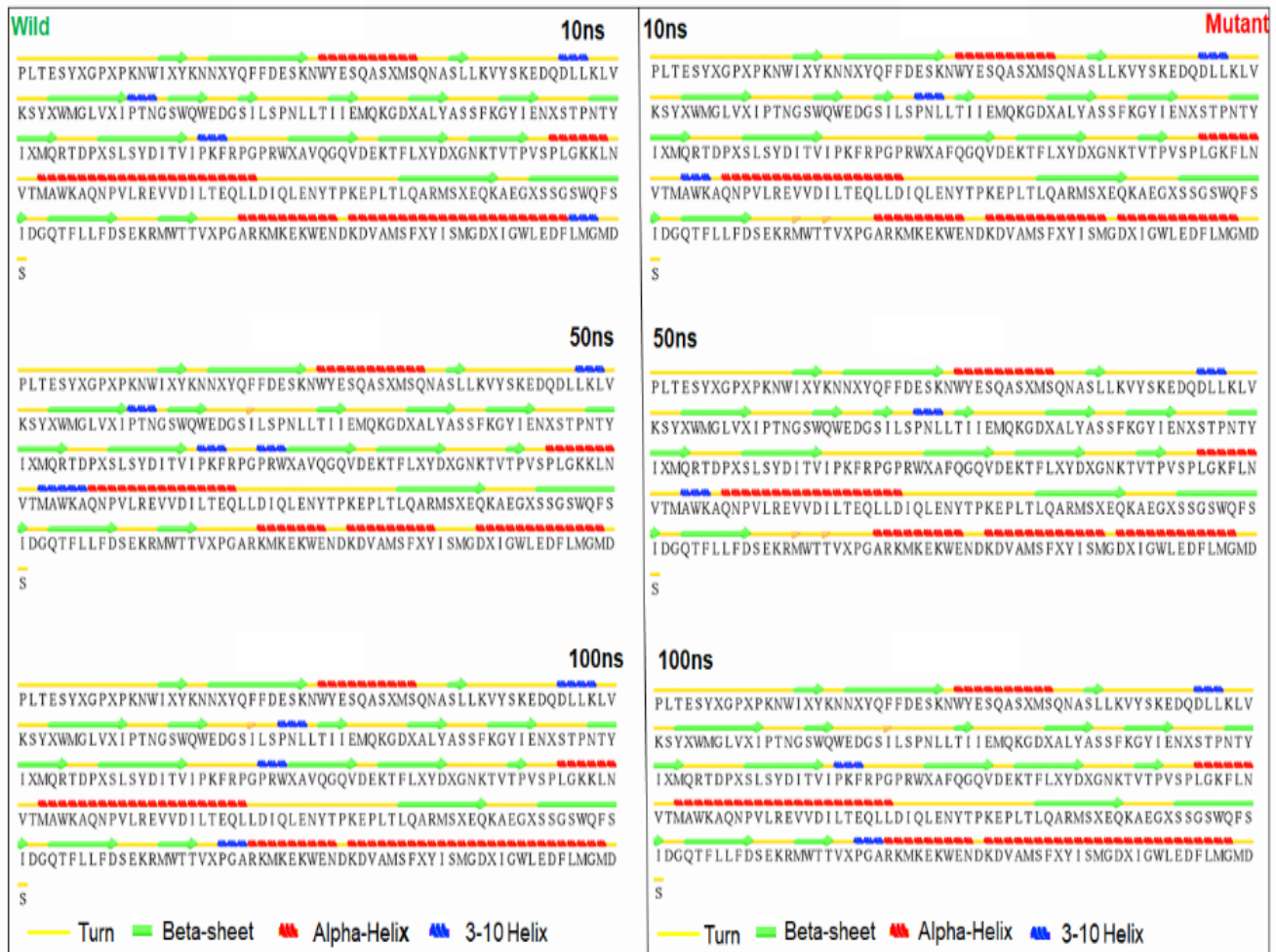


Figure 6.7: DSSP analysis of the mutant and wild protein at 10ns, 50ns and 100ns.

### 6.3.5 Investigation of Structural conformation in ULBP6

To understand the structural impact of V52F mutation, we carried out protein-protein docking using FRODOCK (41) and PRODIGY (44), afterwards, we carried out a molecular dynamic simulation of the wild and mutant protein. We calculated RMSD for all the C $\alpha$  atoms from the initial structure using the equation:

$$RMSD = \left( \frac{\sum_N (R_1 - R_1^0)^2}{N} \right)^{\frac{1}{2}}$$

N represents the total number of atoms in the complex. Ri denotes the vector position of the C $\alpha$  atom in the reference conformation of particle i. This is calculated after aligning the structure to the initial conformation (O) using the least square fitting.

RMSF was estimated using the equation:

$$sRMSF_i = \frac{(RMSF_i - \overline{RMSF})}{\sigma(RMSF)}$$

RMSFi represents the RMSF of the i<sup>th</sup> residue, from which the average RMSF is subtracted. This is then divided by the RMSF's standard deviation to yield the resultant standardized RMSF.

The backbone RMSD values of the wild and Mutant protein during the production phase relative to the starting structures were plotted (Figure 6.8A) to obtain an estimate of the MD trajectory quality and convergence.

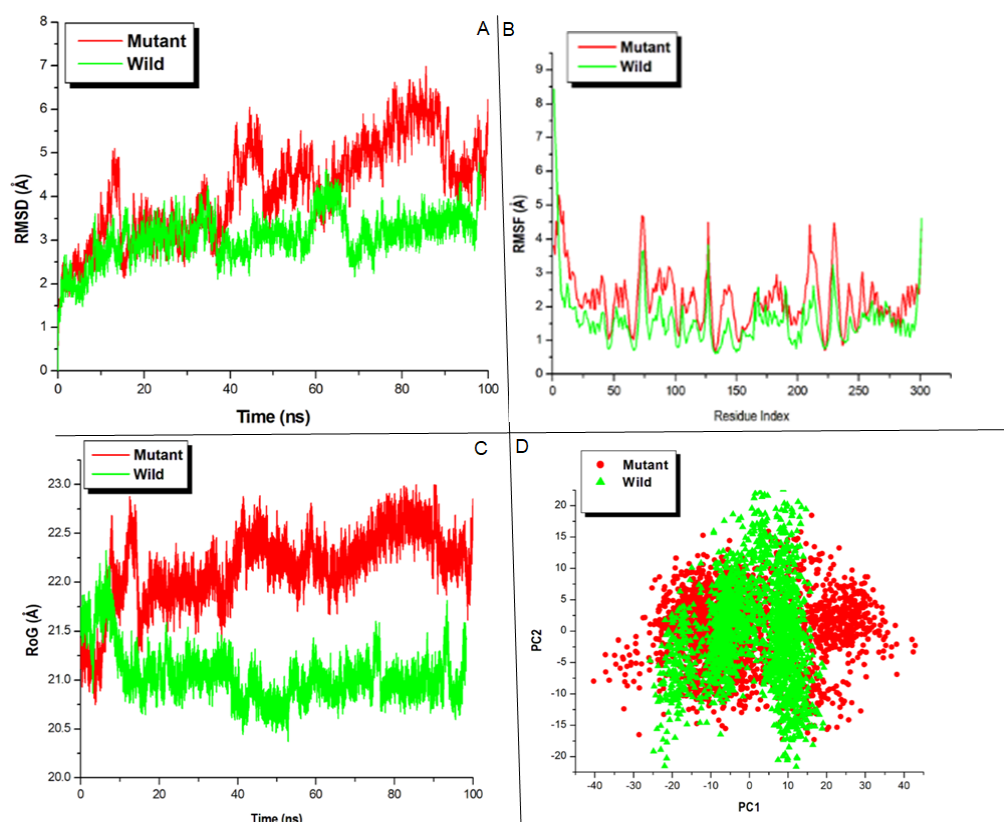


Figure 6.8: Backbone RMSDs are depicted as a function of time for the wild and mutant ULBP6, the mutant is represented with a red colour while the wild as green (A). RMSF of the C- $\alpha$  of the mutant (red) and wild (green) (B). Radius of gyration of C- $\alpha$  atoms of native and mutant ULBP6 versus time at 300k (C). Projection

of the motion of the mutant and wild proteins in phase space along the first two principal eigenvectors at 300k (D).

The RMSD result showed that the mutant protein had an average RMSD value of 4.46Å while the wild protein had an average value of 3.07Å. This showed that the wild protein had lesser motional movement when compared to the mutant protein. We used RMSF to explore the residual movement of the systems in the course of the simulation period. As seen in the plot, the mutant exhibited higher residual fluctuation with an average RMSF value of 2.33Å while the wild protein showed a lesser residual fluctuation with an average RMSF of 1.62Å (Figure 8A). The PCA showed that the mutant protein occupies more area in the conformational space with higher trace value and changes in direction of the clusters when compared to the wild protein (Figure 6.8D). Furthermore, the ROG plot of the two systems showed dissimilar compactness trend between the mutant and wild ULBP6 (Figure 6.8C). The mutant protein had an average ROG value of 22.2Å while the wild protein exhibited a lower ROG average value of 21.1Å. In order to have an insight into the structural impact of V52F mutation on protein-protein interaction, we monitored the distance between the hotspot residues present in the mutant and wild protein in the course of the simulation. In the mutant protein, the distance between S151 and D99 increased from 1.67Å at 20ns to 8.0Å at 100ns. N207-E103 increased from 11.2Å at 20ns to 14.2Å at 100ns and K197-R44 increased from 13.3Å at 20ns to 14.7Å at 100ns (Fig. 6.9). In the wild protein, these distances were seen to remain almost constant from the beginning of the simulation to the end. This suggests that the mutation caused high motional movement of the mutant proteins, this corroborate the high residual fluctuations seen in the RMSF plot.

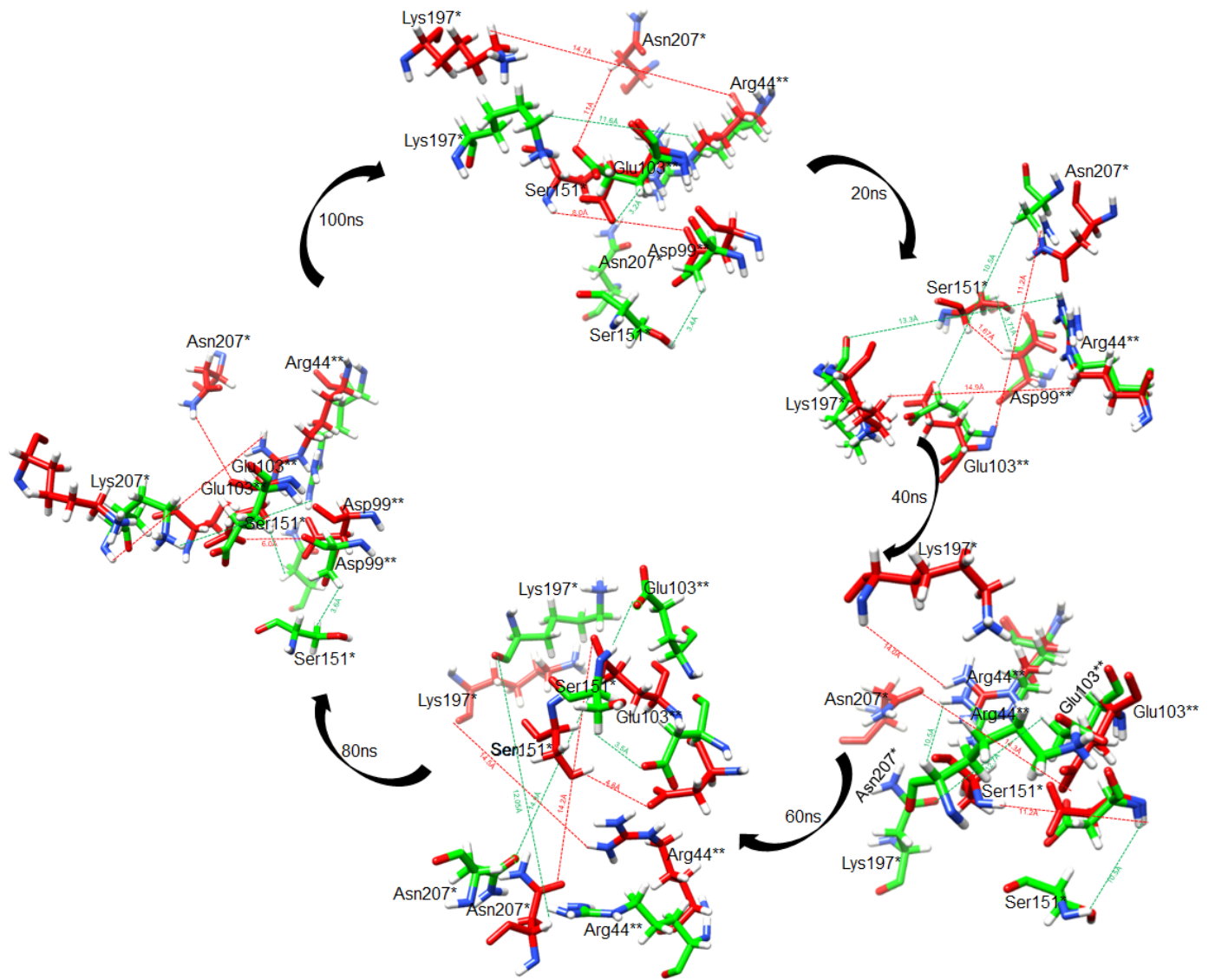


Figure 6.9: Schematic diagram depicting change in distance of the interacting domain residue in the course of the simulation run.

## 6.4 Discussion

Alteration in genetic sequences can be spontaneous, evolutionary adaptation, or chemical induction. This alteration has been classified under two broad categories; mutation and polymorphism (42). While mutation in most cases lead to the onset of diseases, impact of polymorphic variant varies (42). The structure-function of a protein is greatly affected when the SNP occur in the protein coding region, consequently leading to a pathogenic state (43). As a result of the roles they play, attention has been focused on them. The effect of these SNPs can be determined using bioinformatics tools. In this present study, we used PolyPhen, SIFT, SNP&GO, Phd-SNP and PMut to determine the effect of SNPs found in ULBP6. The SNPs data was retrieved from Uniprot dbSNPS. V52F with SNP ID rs145336297 was predicted by the tools to be deleterious and damaging. We performed a protein-protein docking and further ran a 100ns MD simulation to reveal the effect of V52F mutation on NKGD-ULBP6 binding.



Protein-Protein docking using PRODIGY (44), showed the wild protein had a binding affinity of  $-8.6\text{kcalmol}^{-1}$  while the mutant had  $6.3\text{kcalmol}^{-1}$ . This result corroborated the prediction of BeAtMuSiC (38). Protein-protein interaction as revealed by STRING indicated that ULBP6 has direct interaction with killer cell lectin-like receptor subfamily K member 1 (KLRK1), which functions as an activating and costimulatory receptor involved in immune surveillance upon binding to various cellular stress-inducible ligands. The interaction network showed that residue V52 in the wild protein formed hydrogen bonds with Ser151 and Val 97, it also elicited van der Waals interaction with Met125 and Asp107, however in the mutant, phenylalanine formed hydrogen bond with Tyr64, Val98, and Val39. These hydrogen bond elicited between Phe52 and surrounding residues could provide an explanation why the instability at the protein-protein hot spot was high in the mutant as compared to the wild. Molecular dynamic simulation is a powerful computational tool that has been used in the study of diseases and macromolecule behavior (45–47). It is cheaper and fast when compared to experimental studies, it also provide a framework upon which experimental studies can be designed (48–50). Of notice in the MD simulation studies is the change in distance between S151-D99, N207-E103 and K197-R44 (Fig 6.9). Zou et al., 2017, has also highlighted these residues to be paramount to the mechanistic binding between NKG2D and ULBP6 (14). As the simulation progresses, the distance between these residues increased in the mutant, however in the wild protein, the distance remained slightly unchanged from the beginning of the simulation to the end. Interaction between NKG2D and ULBP6 facilitate the cytotoxicity of NK cells and acts as a co-stimulatory component of T cells (51), if this interaction is distorted, NKG2D-ULBP6 will be unable to perform its role as immune system regulator. Taken together, our results showed that the instability of NKG2D-ULBP6 as a result of V52F mutation could lead to the dislodgement of ULBP6, resulting to inability of NKG2D to perform its immune-regulatory role.

## **6.5 Conclusion**

Non-synonymous SNPs (nsSNPS) is responsible for most human pathological conditions and they are mostly found in the exons of protein. In this study, ULBP6 gene was investigated to evaluate the influence of functional SNPs through computation methods. 23 nsSNPS were retrieved from the dbSNP database to identify deleterious point mutation in ULBPS gene. Taken together, with the arrays of bioinformatics tools employed and computational analysis carried out, we predicted V52F to have a possible impact on NKG2D-ULBP6 binding, although it is not located in the interacting region.

## **Conflict of interest**

The authors declare none.

## **Funding**

This research was not funded

## **Author's Contributions**

Opeyemi Soremekun contributed to the study design, technical methods and writing of manuscripts. Mahmoud Soliman contributed to the manuscript proof-reading and attested to the study design and methods.

## **Acknowledgement**

The authors acknowledge the College of Health Sciences, UKZN for their financial and infrastructural support and at the same time thank the Centre for High Performance Computing (CHPC, [www.chpc.ac.za](http://www.chpc.ac.za)), Cape Town, for computational resources.

## **Reference**

1. Yang Y. Cancer immunotherapy: Harnessing the immune system to battle cancer. *J Clin Invest.* 2015;125(9):3335–7.
2. Eyerich S, Eyerich K, Traidl-hoffmann C, Biedermann T. Cutaneous Barriers and Skin Immunity : Differentiating A Connected Network. *Trends Immunol.* 2018;39(4):315–27.
3. Lisowski P, Strzałkowska N, Józ A, Jarczak J, Kos EM, Krzy J, et al. Defensins : Natural component of human innate immunity q. 2013;74:1069–79.

4. Smyth MJ. Multiple approaches to immunotherapy - The new pillar of cancer treatment. *Immunology and Cell Biology*. 2017;
5. Purdy AK, Campbell KS. Natural killer cells and cancer: Regulation by the killer cell ig-like receptors (KIR). *Cancer Biol Ther*. 2009;8(23):2211–20.
6. Berek C. Eosinophils: Important players in humoral immunity. *Clin Exp Immunol*. 2016;183(1):57–64.
7. Verheyden S, Bernier M, Demanet C. Identification of natural killer cell receptor phenotypes associated with leukemia. *Leukemia* [Internet]. 2004;18(12):2002–7. Available from: <http://www.nature.com/articles/2403525>
8. González-foruria I, Santulli P, Chouzenoux S. Soluble Ligands for the NKG2D Receptor Are Released during Endometriosis and Correlate with Disease Severity. 2015;1–15.
9. Chromosome H, Radosavljevic M, Cuillerier B, Wilson MJ, Clément O, Wicker S, et al. A Cluster of Ten Novel MHC Class I Related Genes on. 2002;79(1):114–23.
10. Hunter S. ( 1 ) Understanding the molecular role of NKG2D ligands in lymphoid stress recognition in cancer ( 2 ) Investigating neuroblastoma mediated immunosuppression A thesis submitted to the University of Birmingham for the degree of School of Biosciences. (1).
11. Ñ RAE, Ñ JAT, Hair JR, Jafferji I, Trowsdale J. ULBP6 / RAET1L is an additional human NKG2D ligand. 2009;3207–16.
12. Jayne A, Diane M, Chris C, Maxwell AP, Fogarty DG, Warren T, et al. Targeted genome-wide investigation identifies novel SNPs associated with diabetic nephropathy. 2010;(2009):77–82.
13. Petukhova L, Duvic M, Hordinsky M, Norris D, Price V, Shimomura Y, et al. Genome-wide association study in alopecia areata implicates both innate and adaptive immunity. *Nature*. 2010;466(7302):113–7.
14. Zuo J, Willcox CR, Mohammed F, Davey M, Hunter S, Khan K, et al. A disease-linked ULBP6 polymorphism inhibits NKG2D-mediated target cell killing by enhancing the stability of NKG2D ligand binding. 2017;8904(May):1–15.
15. Radaev S, Rostro B, Brooks AG, Colonna M, Sun PD. Conformational plasticity revealed by the cocrystal structure of NKG2D and its class i MHC-like ligand ULBP3. *Immunity*.

- 2001;15(6):1039–49.
16. Dokholyan N V. Predicting the functional consequences of non-synonymous single nucleotide polymorphisms in. *Sci Rep*. 2017;(June):1–18.
  17. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res* [Internet]. 2003;31(13):3812–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12824425><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC168916>
  18. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Vol. 2, *Current Protocols in Human Genetics*. 2013.
  19. Capriotti E, Fariselli P. PhD-SNPg: A webserver and lightweight tool for scoring single nucleotide variants. *Nucleic Acids Res*. 2017;45(W1):W247–52.
  20. López-Ferrando V, Gazzo A, De La Cruz X, Orozco M, Gelpí JL. PMut: A web-based tool for the annotation of pathological variants on proteins, 2017 update. *Nucleic Acids Res*. 2017;45(W1):W222–8.
  21. Capriotti E, Calabrese R, Fariselli P, Martelli P, Altman RB, Casadio R. WS-SNPs&GO: a web server for predicting the deleterious effect of human protein variants using functional annotation. *BMC Genomics* [Internet]. 2013;14(Suppl 3):S6. Available from: <http://www.biomedcentral.com/1471-2164/14/S3/S6>
  22. Bateman A, Martin MJ, O'Donovan C, Magrane M, Alpi E, Antunes R, et al. UniProt: The universal protein knowledgebase. *Nucleic Acids Res*. 2017;45(D1):D158–69.
  23. Sherry ST, Ward M, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. 2001;29(1):308–11.
  24. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. *Nucleic Acids Res*. 2000;28(1):235–42.
  25. Yang Z, Lasker K, Schneidman-Duhovny D, Webb B, Huang CC, Pettersen EF, et al. UCSF Chimera, MODELLER, and IMP: An integrated modeling system. *J Struct Biol*. 2012;179(3):269–78.
  26. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. protocol UPDATE SIFT missense

- predictions for genomes. *Nat Protoc.* 2015;11(1):1–9.
27. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Bork P, Kondrashov AS, et al. HHS Public Access. 2010;7(4):248–9.
  28. Calabrese R, Capriotti E, Fariselli P, Martelli PL, Ñ RC. Functional Annotations Improve the Predictive Score of Human Disease-Related Mutations in Proteins. 2009;
  29. Capriotti E, Calabrese R, Casadio R. Sequence analysis Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. 2006;22(22):2729–34.
  30. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, et al. The Amber biomolecular simulation programs. *J Comput Chem.* 2005 Dec;26(16):1668–88.
  31. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and Testing of a General Amber Force Field. *J Comput Chem.* 2004;25:1157–74.
  32. Berendsen HJC, Postma JPM, Gunsteren WF Van, Dinola A, Haak JR, Berendsen HJC, et al. Molecular dynamics with coupling to an external bath Molecular dynamics with coupling to an external bath. *J Chem Phys.* 2012;3684(May 2016):926–35.
  33. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys.* 1977;23(3):327–41.
  34. Roe DR, Cheatham III TE. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Com.* 2013;9(7):3084–95.
  35. David CC, Jacobs DJ. Principal component analysis: a method for determining the essential dynamics of proteins. *Methods Mol Biol.* 2014;1084:193–226.
  36. Lawal M, Olotu FA, Soliman MES. Across the blood-brain barrier: Neurotherapeutic screening and characterization of naringenin as a novel CRMP-2 inhibitor in the treatment of Alzheimer’s disease using bioinformatics and computational tools. *Comput Biol Med.* 2018 Jul;98:168–77.
  37. Baets G De, Durme J Van, Reumers J, Maurer-stroh S. SNPeffect 4 . 0 : on-line prediction of molecular and structural effects of protein-coding variants. 2012;40(November 2011):935–9.
  38. Dehouck Y, Kwasigroch JM, Rooman M, Gilis D. BeAtMuSiC : prediction of changes in protein

- protein binding affinity on mutations. 2013;41(May):333–9.
39. Paul Shannon 1, Andrew Markiel 1, Owen Ozier, 2 Nitin S. Baliga, 1 Jonathan T. Wang, 2 Daniel Ramage 2, Nada Amin 2, Benno Schwikowski, 1, 5 and Trey Ideker<sup>2, 3, 4 5</sup>, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res* [Internet]. 2003;13(22):6. Available from: <http://ci.nii.ac.jp/naid/110001910481/>
  40. Heinig M, Frishman D. STRIDE: A web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res*. 2004;32(WEB SERVER ISS.):500–2.
  41. Garzon JI, López-Blanco JR, Pons C, Kovacs J, Abagyan R, Fernandez-Recio J, et al. FRODOCK: A new approach for fast rotational protein-protein docking. *Bioinformatics*. 2009;25(19):2544–51.
  42. Alibrandi S. ( FORMERLY : CAPSULA BURNEA ) Review roles of single-nucleotide polymorphisms in healthy subjects and disease. 2018;13(27):118–9.
  43. Kumar A, Purohit R. Computational screening and molecular dynamics simulation of disease associated nsSNPs in CENP-E. *Mutat Res - Fundam Mol Mech Mutagen* [Internet]. 2012;738–739(1):28–37. Available from: <http://dx.doi.org/10.1016/j.mrfmmm.2012.08.005>
  44. Xue LC, Rodrigues JP, Kastritis PL, Bonvin AM, Vangone A. PRODIGY: A web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics*. 2016;32(23):3676–8.
  45. Soremekun OS, Olotu FA, Agoni C, Soliman MES. Recruiting monomer for dimer formation: resolving the antagonistic mechanisms of novel immune check point inhibitors against Programmed Death Ligand-1 in cancer immunotherapy. *Mol Simul* [Internet]. 2019;45(10):777–89. Available from: <https://doi.org/10.1080/08927022.2019.1593977>
  46. Soremekun OS, Olotu FA, Agoni C, Soliman MES. Drug promiscuity: Exploring the polypharmacology potential of 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones as an inhibitor of the ‘god father’ of immune checkpoint. *Comput Biol Chem*. 2019;80.
  47. Khan S, Biji I, Betz RM, Soliman MES. Reversible versus irreversible inhibition modes of ERK2: A comparative analysis for ERK2 protein kinase in cancer therapy. *Future Med Chem*.

2018;10(9):1003–15.

48. Lazarova M. Virtual Screening – Models , Methods and Software Systems. Int Sci Conf Comput Sci. 2008;(January 2008):55–60.
49. Yang Y, Liu H, Yao X. Understanding the molecular basis of MK2-p38 $\alpha$  signaling complex assembly: insights into protein-protein interaction by molecular dynamics and free energy studies. Mol Biosyst. 2012 Aug;8(8):2106–18.
50. Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Discov [Internet]. 2015;10(5):449–61. Available from: <http://www.tandfonline.com/doi/full/10.1517/17460441.2015.1032936>
51. Zuo J, Willcox BE, Moss P. ULBPs : regulators of human lymphocyte stress recognition. 2017;8(63):106157–8.

## CHAPTER 7

### **Integrating Bioinformatics strategies in cancer immunotherapy: Current and future perspectives**

Houda N. Washah<sup>1</sup>, Elliasu Y. Salifu<sup>1</sup>, Opeyemi Soremekun<sup>1</sup>, Ahmed A. Elrashedy<sup>1</sup>, Geraldene Munsamy<sup>1</sup>, Fisayo A. Olotu<sup>1</sup>, Mahmoud E. S. Soliman<sup>1\*</sup>

<sup>1</sup>Molecular Bio-computation and Drug Design Lab, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

\*Corresponding Author: Mahmoud E.S. Soliman

Email: [soliman@ukzn.ac.za](mailto:soliman@ukzn.ac.za)

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



## **Abstract**

For the past few decades, the mechanism of immune response to cancer has been exploited extensively and a significant attention has been given into tackling the therapeutic potential of the immune system. Cancer immunotherapy has been established as a promising innovative treatment for many forms of cancer. Immunotherapy has gained its prominence through various strategies including; cancer vaccines, monoclonal antibodies (mAbs), adoptive T cell cancer therapy and the immune checkpoint therapy. However, the full potential of cancer immunotherapy is yet to be attained. Recent studies have identified the use of bioinformatics tools as a viable option to help transform the treatment paradigm of several tumors by providing a therapeutically efficient method of cataloging, predicting and selecting immunotherapeutic targets which are known bottlenecks in the application of immunotherapy. Herein, we gave insightful overview of the types of immunotherapy techniques used currently, their mechanisms of action and discussed some bioinformatics tools and databases applied in the immunotherapy of cancer. This review also provides some future perspective in the use of bioinformatics tools for immunotherapy.

**Keywords:** Cancer Immunotherapy, immune system, Monoclonal antibody, Bioinformatics.

## 7.0 Introduction

Cancer is still a major global health concern due to its high incidence and rapid mortality rate for the past decades [1]. Cancer is a broad term for a collection of related diseases that involves abnormal cell growth with the tendency of spreading to other parts of the body [1, 2]. The global statistics of cancer reported for 2018 identifies cancer as the second leading cause of mortality ranking behind cardiovascular disease worldwide [3]. The report also suggests cancer has the potential of becoming the leading cause of death worldwide with a projected increase of 18.1million new cases by the 21<sup>st</sup> century [3]. Cancer is known to develop in various parts of the body including the lung, breast, prostate area, liver, cervix and many more. Among the common cancer types, lung cancer is identified to be frequently diagnosed among populations and is reported to be the leading cause of mortality with an estimated 18.4% of all cancer deaths [3]. female breast cancer closely followed this with an estimated 11.6% and then prostate cancer with 7.1% of the total cancer deaths [3].

Current available therapeutic strategies employed in treating cancer include surgery, radioactivity, targeted therapy, immunotherapy and chemotherapy [4, 5]. These treatment options has provided notable advances towards eradicating primary tumors, however the prevalence of disease relapse remain on the rise as a result of residual malignant cells [6, 7]. As such the surge for viable therapeutic options that will eliminate resistant malignant tumor cells is warranted. Cancer results in the formation of many abnormal cells comprising of old and ruptured cells as well as the formation of unwanted new cells which collectively grow and become malignant cells known as tumors [8]. Cancer cells develop when the gene that controls growth and differentiation in a normal cell undergoes alterations [9]. The implicated genes are categorized as oncogenes where they aid cell growth and reproduction and Tumor suppressor genes where they impede cell division and survival. Targeting these tumor cells has been exploited to be important towards developing therapeutic agents to eliminate the cancer epidemic[10]. The immune system has proven to be an attractive route for defeating cancer in recent times due to its ability to induce anti-tumor response [11–14].

## 7.1 The immune system's mechanism of action against cancer

Clinical studies suggest the existence of a close relationship between the immune systems and the treatment of cancer [7, 9]. The body's Immune response comprises the humoral and cellular immunities, both of which are mediated by the B and T lymphocytes also known as B cells and T cells respectively [7, 15]. The B cells produce antibodies that neutralize extracellular microbes and toxins while T cells are responsible for eliminating intracellular microbes by recognizing antigens and also activate antigen presenting cells (APCs) [16, 17]. The immune system plays a very important role towards eliminating cancer through innate immunity and adaptive immunity [18]. It has been established that cancer as a genetic entity triggers both the innate and adaptive immune response during its evolution [16, 19]. The innate immune system can stimulate the response from B cells and T cells by releasing signals [19]. While the adaptive immune system is known to consist of B cells, CD8+ cytotoxic T cells as well as CD4+ helper T cells [16]. Antigen presenting cells (APCs) play a crucial role in the immune process as it bridges the innate and adaptive immune systems by identifying unfamiliar antigens and presenting them to T lymphocytes during an immune response. The natural killer (NK) cells of the T lymphocyte produce a pleiotropic cytokine known as interferon gamma (IFN- $\gamma$ ) [20]. This IFN- $\gamma$  plays an important role on the interface of innate and adaptive immune systems by signaling an increase in expression of Major histocompatibility complex (MHC I) and also induce the expression of MHC II molecules on target cells, and therefore increases their ability to display antigenic peptides to the cytotoxic T lymphocytes to trigger an immediate immune response [21]. In case these peptides are derived from a tumor associated antigen (TAA), IFN- $\gamma$  could result in increased TAA-specific cytolytic CD8+ T lymphocytes (CTLs) activation and T cell mediated tumor killing [20, 21]. The CTLs form the basis of immune response towards combating cancer [22]. Cancer immunotherapy encompasses a wide scope of techniques that aim to improve the immune response against tumors. As such we focus on the techniques employed in cancer immunotherapy and give an overview of bioinformatics strategies applied in cancer immunotherapy as well as some future perspectives in cancer immunotherapy.

## 7.2 Techniques used in immunotherapy of cancer

The last few decades have witnessed ground breaking research geared towards immuno-oncology studies. This has provided an intriguing advances that suggests tumors are recognized by the immune system and the growth of these tumors can be restrained for a longer period through the process of immune-surveillance [23]. The early efforts to harness the immune system in cancer control was pioneered by Dr William B. Coley in the 1890's. Coley worked on the first immunotherapy to treat cancer by using toxins extracted from *Streptococcus erysipelatis* and *Bacillus prodigious* [24]. However, there was a limited clinical efficacy as a result of the tumor cells evading the immune system unrecognised. As such there has been tireless efforts by researchers to provide insights on how cancer evades the immune system with the aim of developing novel pathways to eliminate the disease. Cancer immunotherapy has been categorized into active and passive immunotherapies based on their ability to engage the immune system of a host against cancer. As shown in Figure 7.1. Passive and active immunotherapy has further been divided in various strategies. The cancer immunotherapy strategies discussed in this review include; checkpoint therapy, cancer vaccines, adoptive T cell therapies and monoclonal antibodies.

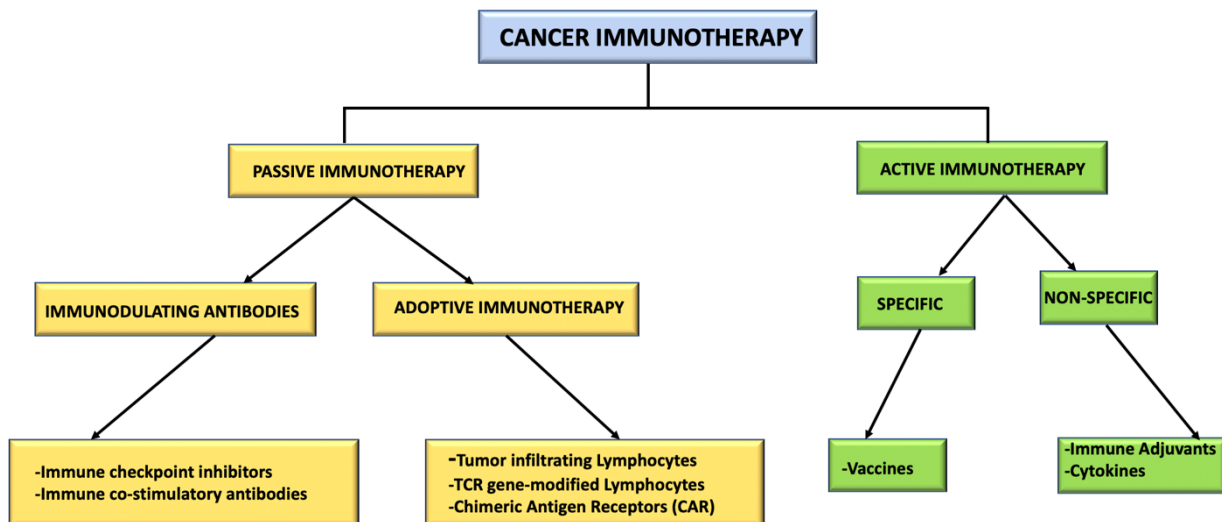


Figure 7.1: Classification of cancer immunotherapy

## 7.3 Immune checkpoint inhibitors and Monoclonal Antibodies

### 7.3.1 Immune checkpoint Inhibitors.

The application of Immune checkpoint inhibitors has gained recognition as the most advanced approach to therapeutically explore the antitumor activity in cancer cells[25, 26]. The immune checkpoint strategy has a lot of recent clinical success especially towards patients with several malignant cancer types. This type of immunotherapeutic strategy focuses on target identification and is mediated by ligand-receptor interactions. The strategy derives its idea from the ability of the immune system to clearly distinguish between normal cells of the body and foreign ones in order to trigger an immune response[27, 28]. As such the immune system has checkpoints which enables easy identification of foreign cells to initiate an attack. Some of the suitable immune checkpoint targets include; Programmed death 1 (PD-1), T-cell immunoglobulin and mucin domain-3 (TIM-3), Lymphocyte-activation gene 3 (LAG-3) and Cytotoxic antigen-4 (CTLA-4)[29, 30]. However, Programmed death 1 (PD-1) and Cytotoxic antigen-4 (CTLA-4) are the most widely studied[27, 31]. However, cancer cells usually develop ways to use these checkpoints to evade attacks from the immune system. As such drugs that target these checkpoints are readily developed and hold a lot of promise as cancer treatment. The checkpoint immune technique has been identified as one of the most remarkable immunomodulating therapies of present-day [32].

PD-1 is a checkpoint protein found on T cells of the immune system. It helps in keeping the immune response on check by acting as an off control to prevent T cell from eliciting an attack on other body cells [33]. This is possible when it binds to another protein known as Programmed Death-Ligand 1 (PD-L1), which helps to prevent the T cells from killing other cells. Some cancer cells are known to possess large amounts of PD-L1, an adaptation that helps in evading attacks from the immune system[33–35]. However this binding of PD-1 to PD-L1 can be blocked by some specific antibodies known as monoclonal antibodies [19]. Several drugs have been developed that target PD-1 and are very promising towards treating cancer. Examples of drugs that target PD-1 include; Pembrolizumab (Keytruda)[36], Nivolumab (Opdivo) [37], Cemiplimab (Libtayo)[38]. Also, examples of drugs that target PD-L1 include; Atezolizumab (Tecentriq) [39], Avelumab (Bavencio)[40], Durvalumab (Imfinzi) 3D structure shown in figure 7.2C [41]. The CTLA-4 checkpoint protein is expressed in the

regulatory T cells but is only upregulated in the conventional T cell activation. It also acts an “off switch” when bound to CD08 and CD86 on the surface of antigen-presenting cells (APCs)[42]. Ipilimumab is known inhibitor of CTLA-4 protein (shown in Figure 7.2A) and acts by keeping this protein active to enable an attack on cancer cells [12][43]. A summary of checkpoint inhibitors and the cancer type involved are shown in table 7.1. Nonetheless, some notable setbacks in the use of immune checkpoint therapies is that, some patients fail to respond to this technique and also the drugs used in this therapy can cause failure of the immune system to identify cancer cells and may rather attack body cells causing undesirable outcomes [32].

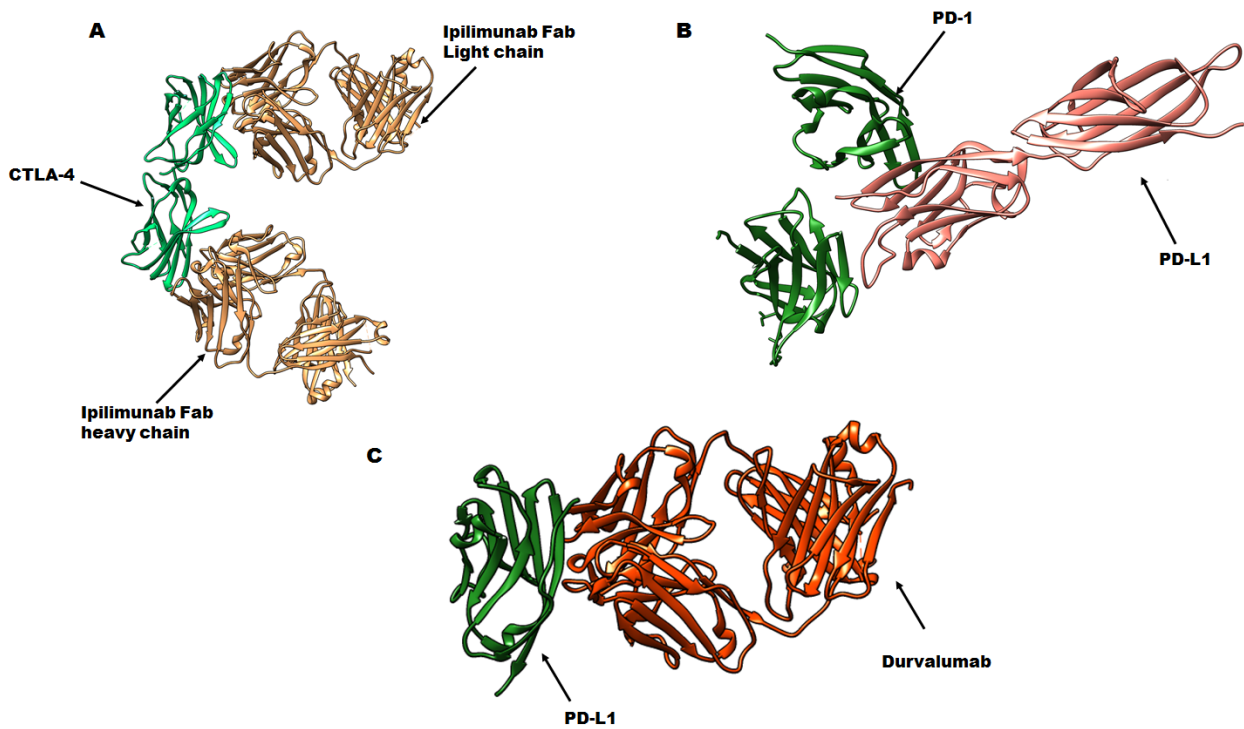


Figure 7.2: 3D structures of A) Cytotoxic antigen – 4 receptor CTLA-4(Light green) in complex with Ipilimumab (brown) PDB code: **5TRU** B) Programmed Death 1(PD-1) in complex with Programmed Death Ligand 1 (PD-L1) PDB code: **3BIK** C) PD-L1 in complex with Durvalumab inhibitor (PDB code: **5X8M**).

Table 7.1: Checkpoint inhibitors

<i>Cancer Type</i>	<i>Drug name</i>	<i>Status</i>	<i>Reference</i>
<b>Melanoma</b>	Ipilimumab	FDA approved	[44][45][46]
	Pembrolizumab	FDA approved	[47]
<b>Melanoma, prostate</b>	Enoblituzumab	Phase I-III	[48][49]
	Ipilimumab	Phase I-III	[50]

<b>Multiple cancers</b>	Tremelimumab	Phase I-III	[51]
	Nivolumab	Phase I-III	[52]
	Pembrolizumab	Phase I-III	[36]
	AMP-224	Phase I	[53]
	Pidilizumab	Phase I-III	[54]
	Atezolizumab	Phase I-III	[55]
	Avelumab	Phase I-III	[56]
	BMS-936559	Phase I	[57]
	IMP321	Phase I	[58]
	BMS-986016	Phase I	[59]
<b>Renal Cell Carcinoma</b>	Ipilimumab	FDA approved	[50]

### 7.3.2 Monoclonal Antibodies (mAbs)

This cancer treatment strategy has been acknowledged as one of the most successful therapeutic strategies for both hematologic malignancies and solid tumors for the past decade [7, 60, 61]. Monoclonal antibodies are identical antibodies produced in the laboratory from a single clone of immune cells. During the process, a mouse is vaccinated with the target antigen. This stimulates the B cells in the spleen to produce antibodies against the target antigen. The spleen of the mouse is removed and the B cells isolated and fused with a tumor cell to form hybridoma cells. This is reportedly necessary because tumor cells divide more easily and rapidly than B cells. These hybridoma cells then reproduce rapidly to make cloned cells which all constitute the same antibody. The resulting antibodies are purified and directed to target a specific part of deregulated signals transduction pathways in cancer or hinder with immunological processes [61, 62]. MAbs have a high specificity as such they easily recognize and bind to a single antigen binding site making them useful in diagnosing and treating disease. They can combine with an anticancer drug to accurately locate and target only the cancer cells whilst avoiding the healthy ones. In addition to this, mAbs carry special markers which enable easy detection of where cancerous cells are starting to build. Monoclonal antibodies can be used to trigger the body's own immune system to recognize and eliminate cancer cells. Recently, FDA approved quite a number of mAbs for the treatment of both solid tumors and hematological malignancies [61, 63]. Available reports suggests that using monoclonal antibodies against PD1 and CTLA-4 has been

regarded as an important breakthrough in cancer immunotherapy[64]. This strategy proved to be efficacious towards patients with metastatic melanoma evidenced by the antitumor response and an increased in survival rate overall of patients treated with ipilimumab, a known monoclonal antibody that targets CTLA-4 in humans [64]. Clinical trials of monoclonal antibodies are ongoing for several types of cancers.

Table 7.2: A summary of some monoclonal antibodies used in immunotherapy of cancer.

<b>Antibody</b>	<b>Mechanism</b>	<b>Uses</b>	<b>Target</b>	<b>Current Status</b>	<b>References</b>
<b>Cetuximab</b>	Binds to and inhibits EGFR	Metastatic colorectal ,Head & Neckcancer	EGFR	FDA approved	[65]
<b>Bevacizumab</b>	Blocks angiogenesis by inhibiting VEGF	Glioblastoma, Lung, Colon cancer	VEGF	FDA approved	[66]
<b>Trastuzumab</b>	Targets HER2 to induce an immune response	Metastatic Breast cancer	HER2	FDA approved	[67]
<b>Panitumomab</b>	Binds to EGFR to prevent its activation	Metastatic colorectal cancer	EGFR	FDA approved	[68]

**EGFR:** Epidermal Growth Factor Receptor, **VEGF:** Vascular Endothelial Growth Factor, **HER2:** Human Epidermal Growth Factor Receptor 2

#### 7.4 Cancer Vaccine

Another immunotherapeutic technique used in treating cancer is the use of vaccines. Most vaccines work by introducing a non-infectious version of a disease-causing microbe into an individual, thus providing a better stimulus to activate disease-specific T cells and to develop immunological memory. Immune memory cells can destroy microbes rapidly, and prevent infection. This form of procedure resulted in eradicating smallpox disease[69]. Cancer vaccine therapy differs from immunotherapies in that it initiates the immune system’s activation in other to re-gain equilibrium between tumor cells and normal cells of the body [11, 70]. Like viral-targeted vaccines, cancer vaccines do not prevent infection but rather activate the immune system to combat a disease that already exists [69]. Most cancer vaccines are made of cancer cells, cell parts or pure antigens. The immune cells of a patient are often isolated and exposed to antigens of cancer, and once activated, these immune cells are reintroduced into the patient's body and are better able to suppress cancer cells [69]. Cancer cells develop from normal body



cells, these cancer cells can be recognized as antigens (foreign substance) by immune cells known as dendritic cells [71]. These dendritic cells act as a commander to the immune system by initiating phagocytosis of cancer cells. Signals are sent to the lymphocytes to fight against these foreign substances [72]. The first anticancer vaccine was approved in 2010 by the Food and Drugs Authority (FDA). The vaccine is known as sipuleucel-T (Provenge) and is actively used for patients with metastatic, castration-resistant prostate cancer (CRPC)[73]. Sipuleucel-T is developed to stimulate T-cell immune responses against prostatic acid phosphatase (PAP), an antigen that is expressed on most prostate cancer cells [74]. Recently, the FDA gave approval for the use of hepatitis B virus vaccines (HBV) and human papilloma virus vaccines (HPV) , these two vaccines work by stimulating the immune system with tumor peptides and antigens [9]. Cancer vaccines have varying side effects from patient to patient however the regularly reported side effect include the inflammation at the site of the administered injection [72].

### **7.5 Adoptive T-cell therapy (ACT)**

In this immunotherapy strategy, T-cells are infused into a cancer patient with the aim of identifying, targeting and eliminating tumor cells from the body [75]. The mechanism of action in this technique involve, extracting specific tumor cytotoxic T cells from a patient's blood and developing at the laboratory and then infused back to the patient to attack cancer cells as shown in figure 7.3 [76]. This mechanism helps the immune system in fighting cancer cells. The T cells can sometimes be subjected to a genetic modification and cultured in the laboratory to be more specific in targeting cancer cells[77, 78]. Adoptive T cell therapy consists of chimeric antigen receptor T-cell (CAR T-cell) and tumor-infiltrating lymphocyte (TIL)[76, 77]. The main difference between these two types relies on the laboratory modifications. In the TILs method, tumor specific cytotoxic T cells are extracted from the patients tumor and modified in the laboratory with substances that easily activates lymphocytes and then given back to the patient whiles in the CAR T-cell method T cells extracted from a patient's blood is modified by adding a gene that targets a special receptor known as chimeric antigen receptor which is known to bind to a protein on cancer cells [33, 79]. Large quantities of these activated lymphocytes as well CAR T cells are then infused back to the patient to help the immune system fight cancer [43,

79]. ACT has gained prominence as a substantial strategy and shows a great potential of being a curative technique for several cancers [76]. Nonetheless, the widespread to solid tumors of this therapy is one of ACT's main future goals due to the difficulties in finding suitable target antigens and also for tumor immunosuppression and complex tumor microenvironment. Furthermore, ACT requires optimization to reduce toxicity and increase the effectiveness of anti-tumors. Recent reports suggests several ACT-based therapies are currently in the late-phase clinical testing, and some T cell therapies are already attaining regulatory approval for the treatment of patients with B cell malignancies.

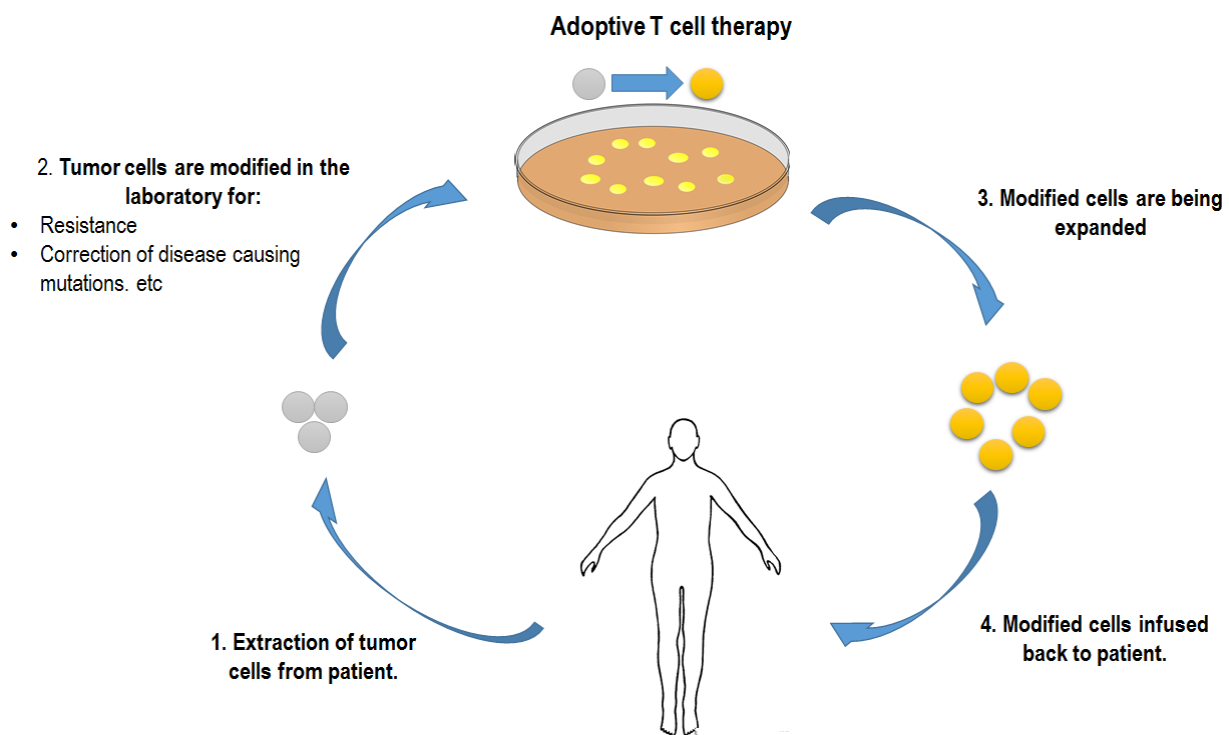


Figure 7.3: Adoptive T cell therapy

## 7.6 Combinatorial cancer immunotherapy

Regardless of the notable clinical success of immunotherapy over other cancer treatment options, the approach is only favorable to some patients as such patients whose immune system fail to elicit an immune response or faced by immune evasion are left untreated[80]. This has led to the development of a combinatorial cancer treatment approach in recent times. This approach involves combining cancer immunotherapy along with other existing anticancer treatment strategies, an approach which has gained prominence significantly in preclinical or clinical studies[32]. A combinatorial strategy involving the

checkpoint inhibitors with drugs that increase tumor immunogenicity, reduce tumor burden and reverse tumor mediated immune suppression has recently been shown to be an effective and long-lasting antitumor response technique[15, 80]. This approach has shown remarkable success in several cancers and is beneficial to patients who fail to respond to checkpoint monotherapy.

Anticancer treatment strategy such as radiotherapy is mostly used in combination with surgeries or chemotherapy, however there are emerging regimens involving a combination of radiotherapy and immunotherapy treatments for cancer[81]. Ongoing studies suggest a combinatorial therapy of radiotherapy and immunotherapies has a high potential of boosting abscopal response rates and will possibly expand the use of radiotherapy in the treatment for both local and metastatic disease [82, 83]. Nonetheless, various preclinical studies are continuously being carried out to explore the role of radiation in combination with immune checkpoint inhibitors. Also, a combination of epigenetic therapy and immunotherapy such as checkpoint immunotherapy has gained a significant recognition in recent studies [84]. Epigenetic modulators in combination with checkpoint inhibitors are known to increase T cell infiltration in tumor microenvironments (TME), reduce Myeloid-derived suppressor cells (MDSC) in TME and augment surface expression of immune checkpoints[84].

### **7.7 Application of Bioinformatics strategies in Cancer immunotherapy**

Recent studies have shown the increase in the development of tumor tolerance to immunotherapy after a limited period of successful treatment. Cataloging, predicting, and selecting immunotherapy targets can be extensively addressed using existing bioinformatics tools and biological databases. Identification and selection of antigens has many different features that depends both on the type and on the application of antigens. Identification of potential antigens de novo from genomic sequence using bioinformatics tools is very difficult, since the expression of proteins is regulated by an array of complex regulatory mechanisms, many of which are poorly understood[85]. As part of a long-established practice, tumor antigens are identified in vitro from serum by screening cDNA phage libraries using immunoassays [86] or proteomics-based screening [87], but bioinformatics tools are perfectly suited to aid this process, either by actively recognizing novel tumor antigens or by organizing and accessing information about known tumor antigens in accessible databases. The techniques

discussed in this review include Next generation sequencing, Epitope prediction, integrated data analysis and network modelling, HLA typing, Sampling Heterogeneity and Modeling of tumor-immune cell interactions (Figure 7.4 ). All these techniques employ various tools in executing their function.

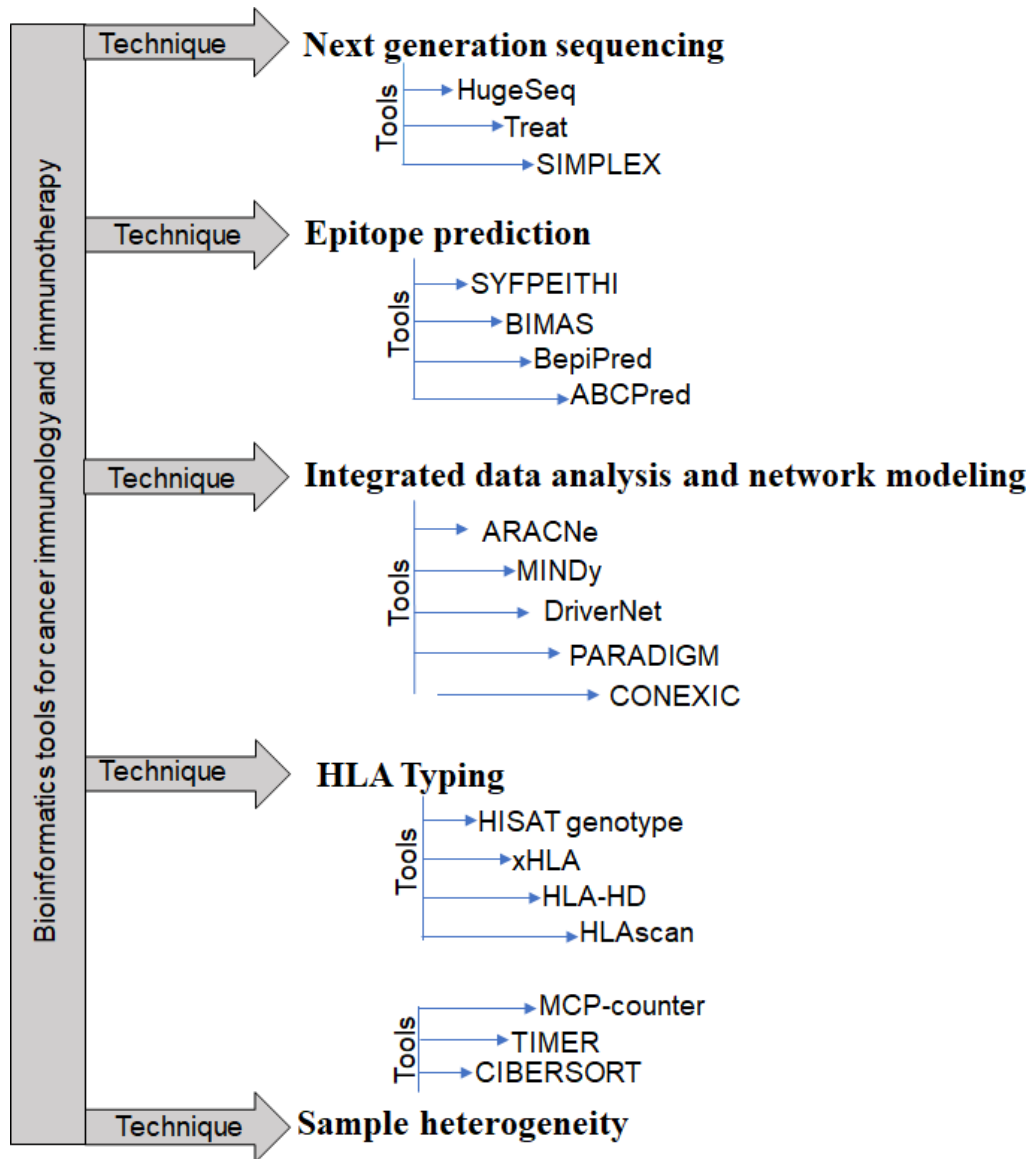


Figure 7.4: Bioinformatics tools for cancer immunology and immunotherapy

### 7.8 Next Generation Sequencing

Next Generation Sequencing (NGS), a largely parallel sequencing technology that has revolutionized the biological sciences [88] [85]. NGS Provides quantitative insights into the tumor cell's molecular machinery, Improves transcript and gene expression profiling as well as detection of alternative splicing and enables the discovery of single nucleotide (SNV) variants, insertions, amplifications, deletions and inter chromosome rearrangements throughout the genome and transcriptome [85]. NGS uses tools such

as HugeSeq, a fully integrated pipeline for NGS analysis from aligning reads to identification and annotation of all types of variants (SNPs, Indels, CNVs, SVs)[89]. Sailana and colleagues applied HugeSeq to investigate the association of HSG with alopecia and mental retardation syndrome [90]. From this research they were able to map a novel gene (APMR2) to chromosome 3q26.2-q26.31 [90]. Furthermore, using the same techniques, APMR3 gene was mapped on chromosome 18q11.2-q12.2 [90]. Another tool used in NGS is the SIMPLEX, an autonomous analysis pipeline for the analysis of NGS exome data, covering the workflow from sequence alignment to SNP/DIP identification and variant annotation[85]. It supports input from various sequencing platforms and exposes all available parameters for customized usage[85].

### **7.9 Integrated data analysis and network modelling technique**

The integrated data analysis and network modelling technique enables one to comprehensively study molecular mechanisms of cancer cells and their interactions within the immune system [91]. The tools used in implementing this technique include: ARACNe[92] and MINDy[93] aids in the reconstruction of gene co-expression networks, genes (nodes) with comparable global expression profiles over samples (tumor/patients). These tools may also be integrated to identify key transcriptional regulators [91]. ARACNe was used to obtain database of transcriptional interactions by using *Arabidopsis thaliana* root samples. Using ARACNe, Montes et al discovered that the transcriptional regulatory networks derived from this newly constructed database successfully recover previously identified root transcriptional modules and to propose new transcription factors for the short root/scarecrow and plethora pathways [94]. DriverNet[95], PARADIGM and CONEXIC[96] are additional tools for implementing the integrated data analysis and network modelling technique which are used to address the functional importance of specific genome alterations, identify which pathways are affected. As well as identify mutations likely to be drivers in tumor progression [91]. Figure 7.4 summarizes examples of bioinformatics techniques and tools for immunology and immunotherapy. DriverNet was used to identify driver genes through the integration of multi-omics data such as somatic mutation, gene expression, and copy number alterations [97].

## 7.10 HLA Typing

HLAs possess very high polymorphic region and has been touted as the most polymorphic segment present in the human genome [98]. This high polymorphism has made it difficult in unraveling HLA-mediated immunogenic correlation of immuno-pathologies [98]. HLA typing is basically employed in organ transplants to get a corresponding potential donor that match with a patient [99]. As technology advances, HLA typing as concurrently advanced; HLA typing was originally carried out using hybridization strategy, serological typing, cellular typing, DNA-based typing and restriction fragment length polymorphism (RFLP) techniques. Polymerase Chain Reaction approach and direct sequencing capillary electrophoretic analyses were later employed [100], with current development in computation, second and third generation sequencing platforms are now employed in HLA typing. This current approach has shown great potential and high predictability [100, 101]. Several HLA typing platforms have been reported such as HISAT genotype [102], xHLA [99], HLA-HD [103], HLAScan [104]. There has been rise in the amount of HLA sequences and new alleles deposited in HLA databases such as the IMG/HLA database [105], CWD HLA allele [106]. HLA typing software which have been developed for easy and rapid HLA typing include ATHLATES [107], POLYSOLVER [108], OpiType [109]. HLA typing of genes have been applied to unreported haplotypes with the aim of unraveling the complexity of HLA-related pathological conditions and associated clinical subtypes. A typical example in the application of HLA typing in psoriasis, novel associations of HLA-DPB1 and BTNL2 genes and five loci among HLA-C, -B, -DPB1, and BTNL2 were discovered with the aid of NGS [98]. Furthermore, specific HLA associations were also reported for the overall risk of psoriasis vulgaris and risk of its specific subphenotypes, such as psoriatic arthritis and cutaneous psoriasis [98].

### **7.11 Epitope Prediction**

B-cells are very important component of the adaptive system, this is due to the long-lasting protection they provide to combat foreign assault or pathogenic onslaught [110]. Some cancer treatment and prevention strategies involve the use of vaccines. Bioinformatics have provided easy and fast means of vaccine development, however, it is expedient to first distinguish between epitopes that are immuno-protective and those that are not [111]. B-cell identification is paramount in many medical, biological and immunological applications such as development of vaccines, control of disease and diagnosis [112]. Experimental techniques used in the identification of epitope include protein crystallography, ELISA and peptide-chip however, these methods are time consuming, low-accuracy, expensive and low throughput generation [110]. Due to these experimental shortcomings, computational strategies have been developed to corroborate or as stand-alone tool for epitope prediction. BepiPred [113, 114], DiscoTope [114, 115], CBtope [116], and ABCpred [117]. Epitope prediction has found useful application in the design of vaccines against some diseases. Qamar et al designed used some in silico bioinformatic approaches to design epitope-based vaccines against middle east respiratory syndrome coronavirus [118]. They identified conserved B and -T cell for the MERS-COV spike (S) protein that may perform a significant role in eliciting the resistance response to MERS-COV infection [118].

### **7.12 Sample heterogeneity**

Deconvolution is the estimation of the relative immune cells that is present in a sample after DNA/RNA sequencing [119]. This is made possible through the recognition of biomarkers and employing their level of expression to discriminate cell types [120]. Tools used in deconvolution evaluation include CIBERSORT [121], TIMER [122], and MCP-counter [123]. They make use of expression data gotten from cell types generated by Immunological Genome Project (ImmGen) [124]. TIMER has the capacity to evaluate 6 immune subtypes, CIBERSORT evaluate 22 immune cell subtypes and MCP-counter evaluates 8 immune cell subtypes and 2 stromal cell subtypes. MCP-counter evaluates absolute abundances while CIBERSORT and TIMER evaluates relative frequencies of immune cell subtypes, while [125].

### **7.13 Modeling of tumor–immune cell interactions**

There is long history of theoretical studies and simulation techniques involving mathematical and computational approaches to study tumor progression and tumor–immune cell interaction. Computational modeling has provided essential insight into studying intra-tumor heterogeneity and the vital interplay between the tumor and the surrounding microenvironment [126]. Computational models allow the investigation the tumor-enhancing effects of the immune system providing indispensable insight aiding in the management of tumor-immune interactions. Modeling has consecutively provided a quantitative time- and cost-effective avenue to study the physical and chemical interactions in tumor initiation growth[127]. The integration of molecular modeling techniques further complements experimental platforms by providing in-depth knowledge into clonal dynamics and microenvironmental cues over time. There are various number of techniques which include deterministic models, stochastic models, Petri nets, cellular automata, agent-based model, and hybrid approaches [128]. Integration of one or more of these models will help improve immunotherapies for cancer treatment.

### **7.14 Future Perspectives**

The advent of computational methods in immunotherapy of cancer has helped transform the treatment paradigm of several tumors and is expected to influence the therapeutic efficiency of immunotherapeutic strategies in future. The application of computational methods in cancer immunotherapy however has some limitations. Firstly, the identification of T cell epitopes using computational techniques are not yet as accurate as peptide binding prediction algorithms. Moreover, the availability of tumor sequences represents a hold-up in conservation and variability, nonetheless a remedy for this limitation is plausible in the near future as high-throughput screening is becoming readily affordable and more efficient. Additionally, these current methods may not adequately capture the issue of intra-tumor genetic diversity as such may affect the efficacy of immunotherapy and other cancer therapies alike. The use of proteomics analyses in immunotherapy has shown to be more prolific in the past years in contrast to genomic analyses. Some promising advances are currently being made



in the wet laboratory to address these limitations through diverse ways. This will progressively boost the need for bioinformatics tools in cancer immunotherapy.

### **7.16 Conclusion**

Cancer immunotherapy has gained a lot of recognition since its incision in the past decades by introducing very efficient strategies using the body's own immune response. The active cancer vaccine, immune checkpoint inhibitors, T cell adoptive therapies and the like are among these developments that have contributed to increased survival rate among patients. Despite these advances emerging research in the interdisciplinary fields of bioinformatics, immunotherapy and immunology is significant to further enhance therapeutic advantage and reduce side effects. Studies presented in this review provide a general overview of cancer immunotherapy giving some insights in the relationship of cancer and the immune system. Further to this, some immunotherapeutic strategies were discussed including cancer vaccine, immune checkpoint inhibitors, T cell adoptive therapies, monoclonal antibodies and combinatorial cancer immunotherapy. Finally, we discussed some of the application of bioinformatics tools in cancer immunotherapy and some future perspectives. As tumor cells' response to immunotherapies is gradually being unravelled and the body of biological tumor data grows, so will the need for bioinformatics to organize, store, and analyse these data.

### **Conflict of interest**

Authors declare no conflicts of interest.

## References

1. Registry, P. C. (2019). Global Cancer Observatory. *Malaysia Cancer Statistics*.
2. Facts, G. C. (2007). Global Cancer Facts & Figures. *Cancer*.
3. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 68(6), 394–424. doi:10.3322/caac.21492
4. Sharma, P., Hu-Lieskovan, S., Wargo, J. A., & Ribas, A. (2017). Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. *Cell*. doi:10.1016/j.cell.2017.01.017
5. Rius, M., & Lyko, F. (2012). Epigenetic cancer therapy: Rationales, targets and drugs. *Oncogene*. doi:10.1038/onc.2011.601
6. Barton, M. K. (2015). Daily aspirin may reduce mortality from prostate cancer with risk of high recurrence. *CA: A Cancer Journal for Clinicians*. doi:10.3322/caac.21263
7. Borghaei, H., Smith, M. R., & Campbell, K. S. (2009). Immunotherapy of cancer. *European Journal of Pharmacology*. doi:10.1016/j.ejphar.2009.09.067
8. Cavallo, F., De Giovanni, C., Nanni, P., Forni, G., & Lollini, P. L. (2011). 2011: The immune hallmarks of cancer. In *Cancer Immunology, Immunotherapy*. doi:10.1007/s00262-010-0968-0
9. Makkouk, A., & Weiner, G. (2015). Cancer Immunotherapy and Breaking Immune Tolerance- New Approaches to an Old Challenge. *Cancer research*, 75(1), 5. doi:10.1158/0008-5472.CAN-14-2538
10. Thommen, D. S. (2019). The First Shall (Be) Last: Understanding Durable T Cell Responses in Immunotherapy. *Immunity*. doi:10.1016/j.immuni.2018.12.029
11. Speiser, D. E., & Flatz, L. (2014). Cancer immunotherapy drives implementation science in oncology. *Human vaccines & immunotherapeutics*, 10(11), 3107–10. doi:10.4161/21645515.2014.983000
12. Yang, Y. (2015). Cancer immunotherapy: Harnessing the immune system to battle cancer. *Journal of Clinical Investigation*. doi:10.1172/JCI83871
13. Tovoli, F., Casadei-Gardini, A., Benevento, F., & Piscaglia, F. (2019). Immunotherapy for

- hepatocellular carcinoma: A review of potential new drugs based on ongoing clinical studies as of 2019. *Digestive and Liver Disease*. doi:10.1016/j.dld.2019.05.006
14. Kruger, S., Ilmer, M., Kobold, S., Cadilha, B. L., Endres, S., Ormanns, S., ... Von Bergwelt-Baildon, M. (2019). Advances in cancer immunotherapy 2019 - Latest trends. *Journal of Experimental and Clinical Cancer Research*. doi:10.1186/s13046-019-1266-0
  15. Nisbet, I. (2016). Cancer immunotherapy comes of age (Finally!). *Australasian Biotechnology*.
  16. Binder, R. J. (2014). Functions of heat shock proteins in pathways of the innate and adaptive immune system. *Journal of immunology (Baltimore, Md.: 1950)*, 193(12), 5765–71. doi:10.4049/jimmunol.1401417
  17. Voena, C., & Chiarle, R. (2016). Advances in cancer immunology and cancer immunotherapy. *Discovery medicine*.
  18. Miller, J. F. A. P., & Sadelain, M. (2015). The journey from discoveries in fundamental immunology to cancer immunotherapy. *Cancer Cell*. doi:10.1016/j.ccell.2015.03.007
  19. Woo, S.-R., Corrales, L., & Gajewski, T. F. (2015). Innate Immune Recognition of Cancer. *Annual Review of Immunology*, 33(1), 445–474. doi:10.1146/annurev-immunol-032414-112043
  20. Trinchieri, G., & Perussia, B. (1985). Immune interferon: a pleiotropic lymphokine with multiple effects. *Immunology Today*. doi:10.1016/0167-5699(85)90080-5
  21. Farrar, M. A., & Schreiber, R. D. (2003). The Molecular Cell Biology of Interferon-gamma and its Receptor. *Annual Review of Immunology*. doi:10.1146/annurev.iy.11.040193.003035
  22. Bevan, M. J. (2004). Helping the CD8+ T-cell response. *Nature Reviews Immunology*. doi:10.1038/nri1413
  23. Ribatti, D. (2016). The concept of immune surveillance against tumors: The first theories. *Oncotarget*, 8(4), 7175–7180. doi:10.18632/oncotarget.12739
  24. Coley, W. B. (1910). The Treatment of Inoperable Sarcoma by Bacterial Toxins (the Mixed Toxins of the Streptococcus erysipelas and the Bacillus prodigiosus). *Proceedings of the Royal Society of Medicine*, 3(Surg Sect), 1–48.
  25. Johnson, D. B., Sullivan, R. J., & Menzies, A. M. (2017). Immune checkpoint inhibitors in challenging populations. *Cancer*. doi:10.1002/cncr.30642

26. Ito, A., Kondo, S., Tada, K., & Kitano, S. (2015). Clinical Development of Immune Checkpoint Inhibitors. *BioMed Research International*. doi:10.1155/2015/605478
27. Mahoney, K. M., Freeman, G. J., & McDermott, D. F. (2015). The next immune-checkpoint inhibitors: Pd-1/pd-11 blockade in melanoma. *Clinical Therapeutics*. doi:10.1016/j.clinthera.2015.02.018
28. Spain, L., Diem, S., & Larkin, J. (2016). Management of toxicities of immune checkpoint inhibitors. *Cancer Treatment Reviews*. doi:10.1016/j.ctrv.2016.02.001
29. Ceeraz, S., Nowak, E. C., Burns, C. M., & Noelle, R. J. (2014). Immune checkpoint receptors in regulating immune reactivity in rheumatic disease. *Arthritis Research and Therapy*. doi:10.1186/s13075-014-0469-1
30. Pardoll, D. M. (2012). The blockade of immune checkpoints in cancer immunotherapy. *Nature Reviews Cancer*. doi:10.1038/nrc3239
31. Yuan, J., Hegde, P. S., Clynes, R., Foukas, P. G., Harari, A., Kleen, T. O., ... Fox, B. A. (2016). Novel technologies and emerging biomarkers for personalized cancer immunotherapy. *Journal for immunotherapy of cancer*, 4, 3. doi:10.1186/s40425-016-0107-3
32. Sathyanarayanan, V., & Neelapu, S. S. (2015). Cancer immunotherapy: Strategies for personalization and combinatorial approaches. *Molecular Oncology*. doi:10.1016/j.molonc.2015.10.009
33. Hamanishi, J., Mandai, M., Iwasaki, M., Okazaki, T., Tanaka, Y., Yamaguchi, K., ... Fujii, S. (2007). Programmed cell death 1 ligand 1 and tumor-infiltrating CD8+ T lymphocytes are prognostic factors of human ovarian cancer. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.0611533104
34. Okazaki, T., & Honjo, T. (2007). PD-1 and PD-1 ligands: From discovery to clinical application. *International Immunology*. doi:10.1093/intimm/dxm057
35. Taube, J. M., Klein, A., Brahmer, J. R., Xu, H., Pan, X., Kim, J. H., ... Anders, R. A. (2014). Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clinical Cancer Research*. doi:10.1158/1078-0432.CCR-13-3271

36. Robert, C., Schachter, J., Long, G. V., Arance, A., Grob, J. J., Mortier, L., ... Ribas, A. (2015). Pembrolizumab versus Ipilimumab in Advanced Melanoma. *New England Journal of Medicine*. doi:10.1056/nejmoa1503093
37. Larkin, J., Chiarion-Sileni, V., Gonzalez, R., Grob, J. J., Cowey, C. L., Lao, C. D., ... Wolchok, J. D. (2015). Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. *New England Journal of Medicine*. doi:10.1056/nejmoa1504030
38. Migden, M. R., Rischin, D., Schmults, C. D., Guminski, A., Hauschild, A., Lewis, K. D., ... Fury, M. G. (2018). PD-1 Blockade with Cemiplimab in Advanced Cutaneous Squamous-Cell Carcinoma. *New England Journal of Medicine*. doi:10.1056/NEJMoa1805131
39. Fehrenbacher, L., Spira, A., Ballinger, M., Kowanetz, M., Vansteenkiste, J., Mazieres, J., ... Rittmeyer, A. (2016). Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): A multicentre, open-label, phase 2 randomised controlled trial. *The Lancet*. doi:10.1016/S0140-6736(16)00587-0
40. Tsang, K.-Y., Boyerinas, B., Jochems, C., Fantini, M., Heery, C. R., Madan, R. A., ... Schlom, J. (2019). Antibody dependent cellular cytotoxicity activity of a novel anti-PD-L1 antibody, avelumab (MSB0010718C), on human tumor cells. *Journal of Clinical Oncology*. doi:10.1200/jco.2015.33.15\_suppl.3038
41. Approved. (2010). *Chemical & Engineering News*. doi:10.1021/cen-v040n029.obc
42. Farkona, S., Diamandis, E. P., & Blasutig, I. M. (2016). Cancer immunotherapy: The beginning of the end of cancer? *BMC Medicine*. doi:10.1186/s12916-016-0623-5
43. Voena, C., & Chiarle, R. (2016). Advances in cancer immunology and cancer immunotherapy. *Discovery medicine*, 21(114), 125–33.
44. Syn, N. L., Teng, M. W. L., Mok, T. S. K., & Soo, R. A. (2017). De-novo and acquired resistance to immune checkpoint targeting. *The Lancet. Oncology*, 18(12), e731–e741. doi:10.1016/S1470-2045(17)30607-1
45. *FDA approves new treatment for a type of late-stage skin cancer*. (2011). U.S. Food and Drug Administration (FDA).
46. Pollack, A. (2011). Approval for Drug That Treats Melanoma. *The New York Times*.

47. Redman, J. M., Gibney, G. T., & Atkins, M. B. (2016). Advances in immunotherapy for melanoma. *BMC medicine*, *14*, 20. doi:10.1186/s12916-016-0571-0
48. Rizvi, N. A., Loo, D., Baughman, J. E., Yun, S., Chen, F., Moore, P. A., ... Tolcher, A. W. (2016). A phase 1 study of enoblituzumab in combination with pembrolizumab in patients with advanced B7-H3-expressing cancers. *Journal of Clinical Oncology*. doi:10.1200/jco.2016.34.15\_suppl.tps3104
49. Shenderov, E., Demarzo, A., Boudadi, K., Allaf, M., Wang, H., Chapman, C., ... Antonarakis, E. S. (2018). Phase II neoadjuvant and immunologic study of B7-H3 targeting with enoblituzumab in localized intermediate- and high-risk prostate cancer. *Journal of Clinical Oncology*. doi:10.1200/jco.2018.36.15\_suppl.tps5099
50. Hodi, F. S., O'Day, S. J., McDermott, D. F., Weber, R. W., Sosman, J. A., Haanen, J. B., ... Urba, W. J. (2010). Improved survival with ipilimumab in patients with metastatic melanoma. *New England Journal of Medicine*. doi:10.1056/NEJMoa1003466
51. Ribas, A., Kefford, R., Marshall, M. A., Punt, C. J. A., Haanen, J. B., Marmol, M., ... Hauschild, A. (2013). Phase III randomized clinical trial comparing tremelimumab with standard-of-care chemotherapy in patients with advanced melanoma. *Journal of Clinical Oncology*. doi:10.1200/JCO.2012.44.6112
52. Larkin, J., Chiarion-Sileni, V., Gonzalez, R., Grob, J. J., Cowey, C. L., Lao, C. D., ... Wolchok, J. D. (2015). Combined nivolumab and ipilimumab or monotherapy in untreated Melanoma. *New England Journal of Medicine*. doi:10.1056/NEJMoa1504030
53. Duffy, A. G., Makarova-Rusher, O. V., Pratt, D., Kleiner, D. E., Fioravanti, S., Walker, M., ... Greten, T. F. (2016). A pilot study of AMP-224, a PD-L2 Fc fusion protein, in combination with stereotactic body radiation therapy (SBRT) in patients with metastatic colorectal cancer. *Journal of Clinical Oncology*. doi:10.1200/jco.2016.34.4\_suppl.560
54. Armand, P., Nagler, A., Weller, E. A., Devine, S. M., Avigan, D. E., Chen, Y. Bin, ... Gordon, L. I. (2013). Disabling immune tolerance by programmed death-1 blockade with pidilizumab after autologous hematopoietic stem-cell transplantation for diffuse large b-cell lymphoma: Results of an international phase II trial. *Journal of Clinical Oncology*.

doi:10.1200/JCO.2012.48.3685

55. Santini, F. C., & Rudin, C. M. (2017). Atezolizumab for the treatment of non-small cell lung cancer. *Expert Review of Clinical Pharmacology*. doi:10.1080/17512433.2017.1356717
56. Boyerinas, B., Jochems, C., Fantini, M., Heery, C. R., Gulley, J. L., Tsang, K. Y., & Schlom, J. (2015). Antibody-dependent cellular cytotoxicity activity of a Novel Anti-PD-L1 antibody avelumab (MSB0010718C) on human tumor cells. *Cancer Immunology Research*. doi:10.1158/2326-6066.CIR-15-0059
57. Gay, C. L., Bosch, R. J., Ritz, J., Hataye, J. M., Aga, E., Tressler, R. L., ... Eron, J. J. (2017). Clinical trial of the anti-PD-L1 antibody BMS-936559 in HIV-1 infected participants on suppressive antiretroviral therapy. *Journal of Infectious Diseases*. doi:10.1093/infdis/jix191
58. Brignone, C., Escudier, B., Grygar, C., Marcu, M., & Triebel, F. (2009). A phase I pharmacokinetic and biological correlative study of IMP321, a novel MHC class II agonist, in patients with advanced renal cell carcinoma. *Clinical Cancer Research*. doi:10.1158/1078-0432.CCR-09-0068
59. Lipson, E., Gopal, A., Neelapu, S. S., Armand, P., Spurgeon, S., Leonard, J. P., ... al., et. (2016). Initial experience administering BMS-986016, a monoclonal antibody that targets lymphocyte activation gene (LAG)-3, alone and in combination with nivolumab to patients with hematologic and solid malignancies. *Journal for immunotherapy of cancer. Conference: 31st annual meeting and associated programs of the society for immunotherapy of cancer, SITC 2016. United states. Conference start: 20161109. Conference end: 20161113*. doi:10.1186/s40425-016-0173-6
60. Riethmüller, G., Schneider-Gädicke, E., & Johnson, J. P. (1993). Monoclonal antibodies in cancer therapy. *Current Opinion in Immunology*. doi:10.1016/0952-7915(93)90129-G
61. Henricks, L. M., Schellens, J. H. M., Huitema, A. D. R., & Beijnen, J. H. (2015). The use of combinations of monoclonal antibodies in clinical oncology. *Cancer Treatment Reviews*, 41(10), 859–867. doi:10.1016/J.CTRV.2015.10.008
62. Shore, N. D. (2015). Advances in the understanding of cancer immunotherapy. *BJU International*, 116(3), 321–329. doi:10.1111/bju.12692
63. Maleki, L. A., Baradaran, B., Majidi, J., Mohammadian, M., & Shahneh, F. Z. (2013). Future

- prospects of monoclonal antibodies as magic bullets in Immunotherapy. *Human Antibodies*, 22(1–2), 9–13. doi:10.3233/HAB-130266
64. Hodi, F. S., O’Day, S. J., McDermott, D. F., Weber, R. W., Sosman, J. A., Haanen, J. B., ... Urba, W. J. (2010). Improved Survival with Ipilimumab in Patients with Metastatic Melanoma. *New England Journal of Medicine*. doi:10.1056/nejmoa1003466
  65. Beljanski, V. (2007). Cetuximab. In *xPharm: The Comprehensive Pharmacology Reference*. doi:10.1016/B978-008055232-3.63727-4
  66. Beljanski, V. (2007). Bevacizumab. In *xPharm: The Comprehensive Pharmacology Reference*. doi:10.1016/B978-008055232-3.63725-0
  67. Findlay, V. J., & Scholar, E. (2007). Trastuzumab. In *xPharm: The Comprehensive Pharmacology Reference*. doi:10.1016/B978-008055232-3.63738-9
  68. List of Cleared or Approved Companion Diagnostic Devices (In Vitro and Imaging Tools) | FDA. (n.d.).
  69. Naran, K., Nundalall, T., Chetty, S., & Barth, S. (2018). Principles of Immunotherapy: Implications for Treatment Strategies in Cancer and Infectious Diseases. *Frontiers in Microbiology*, 9(December), 1–23. doi:10.3389/fmicb.2018.03158
  70. Ye, Z., Qian, Q., Jin, H., & Qian, Q. (2018). Cancer vaccine: learning lessons from immune checkpoint inhibitors. *Journal of Cancer*, 9(2), 263–268. doi:10.7150/jca.20059
  71. Hurley, L. P., Bridges, C. B., Harpaz, R., Allison, M. A., O’ Leary, S. T., Crane, L. A., ... Kempe, A. (2016). Physician Attitudes Toward Adult Vaccines and Other Preventive Practices, United States, 2012. *Public health reports (Washington, D.C.: 1974)*, 131(2), 320–30. doi:10.1177/003335491613100216
  72. Wong, K. K., Li, W. A., Mooney, D. J., & Dranoff, G. (2016). Advances in Therapeutic Cancer Vaccines. In *Adv Immunol* (Vol. 130, pp. 191–249). doi:10.1016/bs.ai.2015.12.001
  73. Higano, C. S., Small, E. J., Schellhammer, P., Yasothan, U., Gubernick, S., Kirkpatrick, P., & Kantoff, P. W. (2010). Sipuleucel-T. *Nature Reviews Drug Discovery*. doi:10.1038/nrd3220
  74. Kantoff, P. W., Higano, C. S., Shore, N. D., Berger, E. R., Small, E. J., Penson, D. F., ... Schellhammer, P. F. (2010). Sipuleucel-T Immunotherapy for Castration-Resistant Prostate



- Cancer. *New England Journal of Medicine*. doi:10.1056/nejmoa1001294
75. Perica, K., Varela, J. C., Oelke, M., & Schneck, J. (2015). Adoptive T Cell Immunotherapy For Cancer. *Rambam Maimonides Medical Journal*. doi:10.5041/rmmj.10179
76. June, C. H. (2007). Adoptive T cell therapy for cancer in the clinic. *Journal of Clinical Investigation*. doi:10.1172/JCI32446
77. June, C. H. (2007). Principles of adoptive T cell cancer therapy. *Journal of Clinical Investigation*. doi:10.1172/JCI31446
78. Restifo, N. P., Dudley, M. E., & Rosenberg, S. A. (2012). Adoptive immunotherapy for cancer: Harnessing the T cell response. *Nature Reviews Immunology*. doi:10.1038/nri3191
79. Stanton, S. E., & Disis, M. L. (2016). Clinical significance of tumor-infiltrating lymphocytes in breast cancer. *Journal for ImmunoTherapy of Cancer*. doi:10.1186/s40425-016-0165-6
80. Vilgelm, A. E., Johnson, D. B., & Richmond, A. (2016). Combinatorial approach to cancer immunotherapy: strength in numbers. *Journal of Leukocyte Biology*. doi:10.1189/jlb.5ri0116-013rr
81. Pilonis, K. A., Vanpouille-Box, C., & Demaria, S. (2015). Combination of Radiotherapy and Immune Checkpoint Inhibitors. *Seminars in Radiation Oncology*. doi:10.1016/j.semradonc.2014.07.004
82. Weichselbaum, R. R., Liang, H., Deng, L., & Fu, Y. X. (2017). Radiotherapy and immunotherapy: A beneficial liaison? *Nature Reviews Clinical Oncology*. doi:10.1038/nrclinonc.2016.211
83. Formenti, S. C., & Demaria, S. (2013). Combining radiotherapy and cancer immunotherapy: A paradigm shift. *Journal of the National Cancer Institute*. doi:10.1093/jnci/djs629
84. Dunn, J., & Rao, S. (2017). Epigenetics and immunotherapy: The current state of play. *Molecular Immunology*. doi:10.1016/j.molimm.2017.04.012
85. Charoentong, P., & Angelova, M. (2012). Bioinformatics for cancer immunology and immunotherapy. *Cancer Immunology and Immunotherapy*, 61, 1885–1903. doi:10.1007/s00262-012-1354-x
86. Sioud, M., Hansen, M., & Dybwad, A. (2000). Profiling the immune responses in patient sera

with peptide and cDNA display libraries. *International journal of molecular medicine*.

87. Hanash, S. (2003). Disease proteomics. *Nature*. doi:10.1038/nature01514
88. Rønn, L., Benito, O., Stein, M., Ole, B., & Vladimir, W. (2014). Bioinformatics for cancer immunotherapy target discovery. *Cancer Immunology and Immunotherapy*. doi:10.1007/s00262-014-1627-7
89. Lam, H. Y. K., Pan, C., Clark, M. J., Lacroute, P., Chen, R., Haraksingh, R., ... Snyder, M. (2012). Detecting and annotating genetic variations using the HugeSeq pipeline. *Nature Biotechnology*. doi:10.1038/nbt.2134
90. Sailani, M. R. (n.d.). M. Reza Sailani, PhD.
91. Narang, V., Decraene, J., & Gouaillard, A. (2012). Systems immunology : a survey of modeling formalisms , applications and simulation tools. *Immunology Research*, 251–265. doi:10.1007/s12026-012-8305-7
92. Margolin, A. A., Nemenman, I., Basso, K., Wiggins, C., Stolovitzky, G., Favera, R. D., & Califano, A. (2006). ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics*. doi:10.1186/1471-2105-7-S1-S7
93. Wang, K., Saito, M., Bisikirska, B. C., Alvarez, M. J., Lim, W. K., Rajbhandari, P., ... Califano, A. (2009). Genome-wide identification of post-translational modulators of transcription factor activity in human B cells. *Nature Biotechnology*. doi:10.1038/nbt.1563
94. Montes, R. A. C., Coello, G., González-aguilera, K. L., Marsch-martínez, N., Folter, S. De, & Alvarez-buylla, E. R. (2014). Art:10.1186/1471-2229-14-97.
95. Bashashati, A., Haffari, G., Ding, J., Ha, G., Lui, K., Rosner, J., ... Shah, S. P. (2012). DriverNet: uncovering the impact of somatic driver mutations on transcriptional networks in cancer. *Genome biology*. doi:10.1186/gb-2012-13-12-r124
96. Akavia, U., Litvin, O., Kim, J., Mozes, E., Kotliar, D., Tzur, Y., ... Pe'er, D. (2009). Abstract B70: Conexic: A Bayesian framework to detect drivers and their function uncovers an endosomal signature in melanoma. doi:10.1158/0008-5472.fbc09-b70
97. Zhang, T., & Zhang, D. (2017). Integrating omics data and protein interaction networks to prioritize driver genes in cancer. *Oncotarget*, 8(35), 58050–58060.

doi:10.18632/oncotarget.19481

98. Kishore, A., & Petrek, M. (2018). Next-Generation Sequencing Based HLA Typing : Deciphering Immunogenetic Aspects of Sarcoidosis, 9(October), 1–8. doi:10.3389/fgene.2018.00503
99. Xie, C., Xuan, Z., Wong, M., Piper, J., Long, T., Kirkness, E. F., ... Bloom, K. (2017). Fast and accurate HLA typing from short-read next-generation sequence data with xHLA, 114(30), 8059–8064. doi:10.1073/pnas.1707945114
100. Gandhi, M. J., Ferriola, D., Huang, Y., Duke, J. L., & Monos, D. (2017). Targeted Next-Generation Sequencing for Human Leukocyte Antigen Typing in a Clinical Laboratory, 141(June). doi:10.5858/arpa.2016-0537-RA
101. Access, O. (n.d.). We are IntechOpen , the world ’ s leading publisher of Open Access books Built by scientists , for scientists TOP 1 %.
102. Kim, D., Paggi, J., & Salzberg, S. L. (2018). HISAT-genotype : Next Generation Genomic Analysis Platform on a Personal Computer.
103. Kawaguchi, S., Matsuda, F., Higasa, K., Shimizu, M., & Yamada, R. (2017). HLA-HD : An accurate HLA typing algorithm for next-generation sequencing data, (March), 788–797. doi:10.1002/humu.23230
104. Ka, S., Lee, S., Hong, J., Cho, Y., Sung, J., Kim, H., & Kim, H. (2017). HLAScan : genotyping of the HLA region using next-generation sequencing data, 1–11. doi:10.1186/s12859-017-1671-3
105. Robinson, J., Halliwell, J. A., Hayhurst, J. D., Flicek, P., Parham, P., & Marsh, S. G. E. (2015). The IPD and IMGT / HLA database : allele variant databases, 43(November 2014), 423–431. doi:10.1093/nar/gku1161
106. Mack, S. J., Cano, P., Hollenbach, J. A., He, J., Hurley, C. K., Middleton, D., ... Fischer, G. F. (2014). NIH Public Access, 81(4), 194–203. doi:10.1111/tan.12093.Common
107. Liu, C., Yang, X., Duffy, B., Mohanakumar, T., Mitra, R. D., Zody, M. C., & Pfeifer, J. D. (2013). ATHLATES : accurate typing of human leukocyte antigen through exome sequencing, 41(14), 1–8. doi:10.1093/nar/gkt481

108. Shukla, S. A., Rooney, M. S., Rajasagi, M., Tiao, G., Dixon, P. M., Lawrence, M. S., ... Getz, G. (2015). A N A L Y S I S Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes, *33*(11). doi:10.1038/nbt.3344
109. Schubert, B., Mohr, C., Sturm, M., Feldhahn, M., & Kohlbacher, O. (2014). Sequence analysis OptiType : precision HLA typing from next-generation sequencing data, *30*(23), 3310–3316. doi:10.1093/bioinformatics/btu548
110. Jespersen, M. C., Mahajan, S., Peters, B., & Nielsen, M. (2019). Antibody Specific B-Cell Epitope Predictions : Leveraging Information From Antibody-Antigen Protein Complexes, *10*(February), 1–10. doi:10.3389/fimmu.2019.00298
111. Delisi, C., & Berzofsky, J. A. Y. A. (1985). T-cell antigenic sites tend to be amphipathic structures *j \_ k*, *82*(October), 7048–7052.
112. Shirai, H., Prades, C., Vita, R., Marcatili, P., Popovic, B., Xu, J., ... Ikeda, K. (2015). B[1] Shirai H, Prades C, Vita R, et al. Biochimica et Biophysica Acta Antibody informatics for drug discovery ☆. *BBA - Proteins Proteomics* [Internet]. 2015;1844:2002–2015. Available from: <http://dx.doi.org/10.1016/j.bbapap.2014.07.006>. *BBA - Proteins and Proteomics*, *1844*(11), 2002–2015. doi:10.1016/j.bbapap.2014.07.006
113. Jespersen, M. C., Peters, B., Nielsen, M., & Marcatili, P. (2017). epitope prediction using conformational epitopes, *45*(May), 24–29. doi:10.1093/nar/gkx346
114. Guedes, R. L. M., Rodrigues, C. M. F., Coatnoan, N., Cosson, A., Cadioli, F. A., Garcia, H. A., ... de Vasconcelos, A. T. R. (2019). A comparative in silico linear B-cell epitope prediction and characterization for South American and African Trypanosoma vivax strains. *Genomics*, *111*(3), 407–417. doi:10.1016/j.ygeno.2018.02.017
115. Andersen, P. H., Nielsen, M., & Lund, O. L. E. (2006). Prediction of residues in discontinuous B-cell epitopes using protein 3D structures, 2558–2567. doi:10.1110/ps.062405906.2558
116. Ansari, H. R., & Raghava, G. P. S. (2010). Identification of conformational B-cell Epitopes in an antigen from its primary sequence. *Immunome Research*, *6*(1), 6. doi:10.1186/1745-7580-6-6
117. Saha, S., & Raghava, G. P. S. (2006). SHORT COMMUNICATION Prediction of Continuous

- B-Cell Epitopes in an Antigen Using Recurrent Neural Network, *48*(August), 40–48.  
doi:10.1002/prot
118. Shey, R. A., Ghogomu, S. M., Esoh, K. K., Nebangwa, N. D., Shintouo, C. M., Nongley, N. F., ... Souopgui, J. (2019). In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. *Scientific Reports*, *9*(1), 1–18. doi:10.1038/s41598-019-40833-x
  119. Abbas, A. R., Wolslegel, K., Seshasayee, D., Modrusan, Z., & Clark, H. F. (2009). Cellular Activation Patterns in Systemic Lupus Erythematosus, *4*(7). doi:10.1371/journal.pone.0006098
  120. Abbas, A. R., Baldwin, D., Ma, Y., Ouyang, W., Gurney, A., Martin, F., ... Lookeren, M. Van. (2005). Immune response in silico ( IRIS ): immune-specific genes identified from a compendium of microarray expression data, 319–331. doi:10.1038/sj.gene.6364173
  121. Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., ... Ash, A. (2016). HHS Public Access, *12*(5), 453–457. doi:10.1038/nmeth.3337.Robust
  122. Li, B., Severson, E., Pignon, J., Zhao, H., Li, T., Novak, J., ... Liu, X. S. (2016). Comprehensive analyses of tumor immunity: implications for cancer immunotherapy, 1–16. doi:10.1186/s13059-016-1028-7
  123. Becht, E., Giraldo, N. A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., ... Reyniès, A. De. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biology*, 1–20. doi:10.1186/s13059-016-1070-5
  124. Heng, T. S. P., Painter, M. W., Immunological, T., & Project, G. (2008). The Immunological Genome Project: networks of gene expression in immune cells, *9*(10), 1091–1094.
  125. Hammerbacher, J., & Snyder, A. (2017). Informatics for cancer immunotherapy, *28*(Supplement 12), 56–73. doi:10.1093/annonc/mdx682
  126. Schoenfeld, J. D. (2013). We Are All Connected: Modeling the Tumor- Immune Ecosystem. *TRENDS in CANCER*, *16*, 7–9. doi:10.1016/j.trecan.2018.08.006
  127. Onofrio, A. (2008). Metamodeling tumor – immune system interaction , tumor evasion and immunotherapy. *Mathematical and Computer Modelling*, *47*, 614–637.

doi:10.1016/j.mcm.2007.02.032

128. Alex, T., Id, A., Taiwo, R., Id, A., Khanyile, S., Masamba, P., ... Kappo, A. P. (2017). Structural Analysis and Epitope Prediction of MHC Vaccine Development. doi:10.3390/vaccines6010001

## **CHAPTER 8**

### **8.1 Conclusion**

We used a computer algorithm and extended simulation strategy to explore the mechanism of inhibition of newly synthesized small molecule inhibitors (BMS-1166 and BMS-1001) on immune checkpoint PD-L1 protein. Furthermore, at the time of this research, there is no small molecule inhibitor designed to target the immune checkpoint CTLA-4 protein. We carried out a per-residue based pharmacophore base virtual screening leveraging on active site similarity to identify potential small molecule CTLA-4 inhibitors. Natural Killer group 2, member D (NKG2D), is one of the receptors that is expressed on the surface of NK cell and it plays a crucial role in the immune response. NKG2D is activated upon binding to its complementary ligands, UL-16-binding proteins (ULBP6). In the binding of ULBP6 and NKG2D, some amino acids play a crucial role in enhancing the binding potential of the two proteins. Mutation in any of these mutations alters NKG2D-ULBP6 interaction. The results presented in this thesis could be used as a platform for the design of small molecular weight inhibitors for immune proteins.

Our findings revealed that :

1. The mechanism of action of BMS-1001 and BMS-1166 could be via eliciting a motional movement between the PD-L1 monomers which led to the formation of a dimer.
2. This PD-L1 dimer formation disrupted the PD-L1-PD-1 interaction, hence preventing immunosuppression activity of PD-L1-PD-1.
3. The unbound PD-L1 monomer had the highest flexibility and atomic motion when compared to the bound PD-L1 monomer
4. Even though the monomers have the same residues. BMS-1001 and BMS-1166 showed a mechanistic transition from monomer A to monomer B, this suggests that monomer B has a stronger affinity for BMS-1001 and BMS-1166 than monomer A.
5. Binding sites similarities between CTLA-4 and Kallikrein-7 proteins could be leveraged on to repurpose a Kallikrein-7 inhibitor as an anti-CTLA-4.
6. 9 hits were identified when 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones (TDSO) was used in high throughput screening of a large compound database. Two compounds, ZINC04515726 and ZINC08985213 had the highest binding affinity for CTLA-4.

7. V52F mutation in ULBP6, identified by SNPinformatics strategy altered NKGD2-ULBP6 interaction.
8. The mutant ULBP6 elicited high structural inactivity when compared to the wildtype ULBP6.

## **8.2 Future Perspective**

The utilization of small molecular weight molecules as an alternative to humanized antibodies for targeting immune proteins especially immune checkpoint proteins constitutes a paradigm shift in immunotherapy. Although BMS-1001 and BMS-1166 have been experimentally validated. ZINC04515726 and ZINC08985213 which were identified by screening a library need to be validated experimentally both *in-vivo* and *in-vitro* to explore their therapeutic effect. A structure based virtual screening can be carried out on ULBP6 to identify inhibitors with the potential of reversing the effect of the mutation V52F. Furthermore, a mutagenesis study may be carried out to have an in-depth understanding of the effect of V52F on the activity of ULBP6.

## **APPENDIX**

### **Appendix A**



**Soremekun, O.**, Olotu, F., Agoni, C., Soliman, M. (2019). Recruiting monomer for dimer formation: Resolving the antagonistic mechanisms of novel immune check point inhibitors against Programmed Death Ligand-1 in cancer immunotherapy. *Molecular Simulation*, DOI:10.1080/08927022.2019.1593977. (Published).

## **Appendix B**

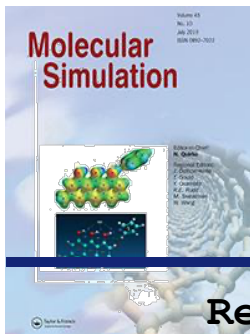
**Soremekun, O.**, Olotu, F., Agoni, C., Soliman, M. (2019). Drug Promiscuity: Exploring the polypharmacology potential of 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones as an Inhibitor of the ‘god father’ of Immune Checkpoint. *Computational Biology and Chemistry*, 80:433-440. (Published).

## **Appendix C**

**Soremekun, O.**, Soliman, M. (2019). From Genomic Variation to Protein Aberration: Mutational Analysis of Single Nucleotide Polymorphism present in *ULBP6* gene and Implication in Immune Response. *Computers in Biology and Medicine*, 111:1033354. (Published).

## **Appendix D**

Houda N. Washah, Elliasu Y. Salifu, **Opeyemi Soremekun**, Ahmed A.Elrashedy, Geraldene Munsamy, Fisayo A. Olotu and Mahmoud E.S. Soliman. (2020). Integrating Bioinformatics Strategies in Cancer Immunotherapy: Current And Future Perspectives. *Combinatorial Chemistry & High Throughput Screening*, 2020, 23, 0-0. (Published).



## Recruiting monomer for dimer formation: resolving the antagonistic mechanisms of novel immune check point inhibitors against Programmed Death Ligand-1 in cancer immunotherapy

Opeyemi S. Soremekun, Fisayo A. Olotu, Clement Agoni & Mahmoud E. S. Soliman

To cite this article: Opeyemi S. Soremekun, Fisayo A. Olotu, Clement Agoni & Mahmoud E. S. Soliman (2019) Recruiting monomer for dimer formation: resolving the antagonistic mechanisms of novel immune check point inhibitors against Programmed Death Ligand-1 in cancer immunotherapy, Molecular Simulation, 45:10, 777-789, DOI: [10.1080/08927022.2019.1593977](https://doi.org/10.1080/08927022.2019.1593977)

To link to this article: <https://doi.org/10.1080/08927022.2019.1593977>



Published online: 21 Mar 2019.



Submit your article to this journal



Article views: 125



View related articles



View Crossmark data



Citing articles: 3 View citing articles



# Recruiting monomer for dimer formation: resolving the antagonistic mechanisms of novel immune check point inhibitors against Programmed Death Ligand-1 in cancer immunotherapy

Opeyemi S. Soremekun, Fisayo A. Olotu, Clement Agoni and Mahmoud E. S. Soliman

Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Durban, South Africa

## ABSTRACT

The design of small molecule antagonists against Programmed Death Ligand-1 (PD-L1) has been the recent highlight of the immune checkpoint blockade therapy. This interventive approach has been potentiated by the development of BMS compounds; BMS-1001 and BMS-1166, which exert their therapeutic activities by inducing dimerisation of PD-L1; a molecular mechanism that has remained unclear. For the first time, we resolve the dynamical events that underlie the antagonistic mechanisms of BMS-1001 and BMS-1166 when bound to PD-L1 using an all-atom molecular dynamics (MD) simulations approach and free binding energy Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) calculations. Time-scale dynamical findings revealed that upon binding a PD-L1 monomer, the BMS-compounds gradually facilitated the 'inbound' motion of another PD-L1 monomer in the same conformational phase space up till dimer formation. Moreover, the non-liganded PD-L1 monomer exhibited the highest structural flexibility and atomistic motions relative to the BMS-liganded monomer as revealed by post-MD trajectory analyses using root mean square deviation (RMSD) and root mean square fluctuations (RMSF) parameters. Trajectory investigations into ligand motions also revealed that the BMS compounds exhibited mechanistic transitions from the monomeric binding site (monomer A) where they were initially bound, to the second monomeric site (monomer B) where they were strongly bound, followed by eventual high-affinity interactions at the tunnel-like binding cleft formed upon the dimerisation of both PD-L1 monomers. These findings present a model that describes the mechanism by which the BMS compounds induce PD-L1 dimerisation and could further enhance the design of highly selective and novel monomeric recruiters of PD-L1 in cancer immunotherapy.

## ARTICLE HISTORY

Received 20 December 2018  
Accepted 4 March 2019

## KEYWORDS


Cancer immunotherapy;  
immune checkpoint;  
Programmed Death Ligand-1; binding cleft; molecular dynamic simulation

## 1. Introduction

Over the years, traditional anticancer therapeutic strategies have been categorised under three pillars namely; surgery, radiation and chemotherapy. Although these therapeutic approaches have offered substantial interventions in the eradication primary tumour the problem of relapse still persist, which result from residual malignant cells and/or tumour metastases [1,2]. This necessitated a new approach regarded as cancer immunotherapy, which provides potent and alternative treatment options [3] that involve the utilisation of naturally derived or synthetically generated compounds to stimulate or enhance the immune system to fight cancer [4]. Hence, cancer immunotherapy was described by science magazine as '2013s breakthrough of the year' due to the recorded success in cancer treatment [5]. The basic role of the immune system is to protect the body against foreign pathogens and infections, and more importantly, the humoral immunity employs antibodies produced by the B cells to neutralise and eradicate foreign pathogens and toxins [6] while cellular immunity responds through recognition of antigens, activation of antigen presenting cells (APCs), activation and proliferation of T cells [7]. Cancer immunotherapy has shown its potency

among numerous therapeutic options [1], with the incorporation of cancer vaccinations, chimeric antigen receptor (CAR) T cell therapy and immune checkpoint blockade (ICB) therapy [8]. Immune checkpoint inhibitors are a class of drugs that have been designed to increase immune response against cancer cells [9]. Moreover, in the immune system, there are different checkpoints pathways that focus on the activation of T cells, [1] and are crucially regulated by biomolecules such as CTLA-4, PD-L1, lymphocyte activation gene-3 (LAG-3) and T cell immunoglobulin and mucin domain containing protein 3 (Tim-3) [10]. These, over the years, have become attractive targets for driving the cytotoxic T cells attack of malignant cells for destruction due to the hypo-responsive and exhaustive influence they have on the immune system via regulation [11,12]. According to the classic checkpoint hypothesis, immune checkpoint inhibitors activate the immune system by blocking the expression of co-inhibitory molecules such as CTLA4, PD-1 and PD-L1 which in turn re-activate the pre-existing tumour response [9,13]. As previously reported, antagonising the interaction between PD-1 and PD-L1 reverts T cell phenotypic exhaustion, which in turn leads to the efficient killing of cancer cells [14]. In recent years, antibodies were widely employed to

CONTACT Mahmoud E.S. Soliman [soliman@ukzn.ac.za](mailto:soliman@ukzn.ac.za) Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences, University of Kwa-Zulu-Natal, Westville Campus, Durban 4001, South Africa

 Supplemental data for this article can be accessed at <https://doi.org/10.1080/08927022.2019.1593977>

© 2019 Informa UK Limited, trading as Taylor & Francis Group

antagonise the interactions between PD-1 and PD-L1, a biological association that promotes immunosuppressiveness, which is employed by cancer cells to evade T-cell induced death [15,16]. However, due to lower production costs, higher stability, improved tumour penetration, amenability for oral administration and elimination of immunogenicity, small-molecular weight inhibitors present a more viable option in place of antibodies as immune check point inhibitors [17]. Most recently, two small molecule antagonists (BMS-1001 and BMS-1166) that directly and selectively disrupt the association between PD-1 and PD-L1 have been identified and patented by Bristol-Myers Squibb (BMS) [18] (Figure 1).

According to previous reports, these BMS compounds bind to and induce PD-L1 dimerisation thereby preventing canonical interactions with PD-1. Mechanistically, these compounds bind to a PD-L1 monomer (APD-L1) and facilitate the recruitment of another PD-L1 monomer (BPD-L1) to form a ABPD-L1 thereby blocking interactions with the naturally occurring PD1. In other words, the ability of these compounds to induce the formation of PD-L1 dimers underlies its antagonistic functions since PD-L1 dimerisation presents a hurdle for the canonical interactions between PD-L1 monomers and its intrinsic ligand, PD-1, resulting in the activation of T cells [19,20] (Figure 2).

Therefore, the development and use of these novel small molecule inhibitors in ICB therapy present an attractive prospect to understand the underlying structural and molecular events involved in the blockage of PD-1 and PD-L1. Molecular elucidation of the transient mechanisms by which BMS compounds (BMS-1001 and BMS-1166) induce PD-L1 dimerisation following an initial interaction with a PD-L1 monomer would provide a model that would be helpful to achieve the structure-based design of monomeric binders. Monomeric recruitment has been previously reported as the mechanism by which vascular endothelial growth factors (VEGFs) dimerises in the presence of  $Zn^{2+}$ ,  $Co^{2+}$  or  $Cu^{2+}$  to form interlocked hydrophobic surfaces [21]. Maschio et al., also used MD simulation to explore the dimerisation mechanism of  $\beta_2$ -microglobulin, they proposed initiation of salt bridges between monomers leading to a 'zipping-like' mechanism [22]. Therefore, in this study, we employed a computational paradigm to model the structural and molecular events associated with PD-L1 dimerisation as induced by antagonistic BMS compounds; BMS-1001 and BMS-1166. Many of the immune checkpoint proteins are very difficult

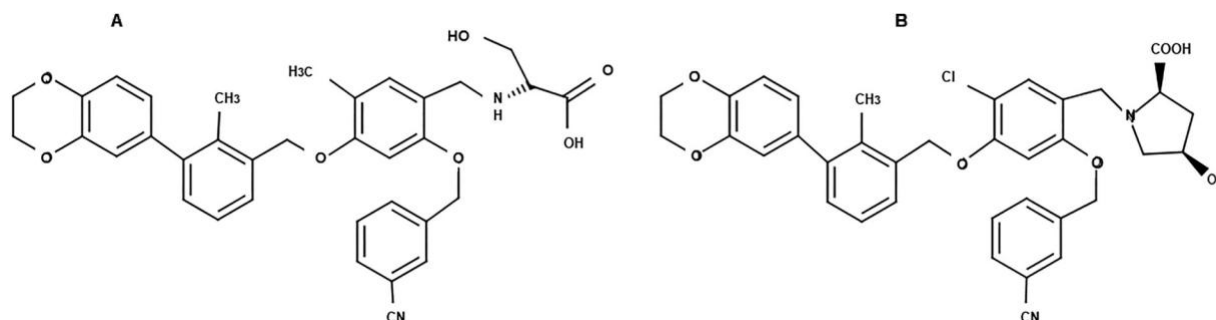


Figure 1. 2D structure of soluble immune checkpoint inhibitors; [A] BMS-1001 and [B] BMS-1166.

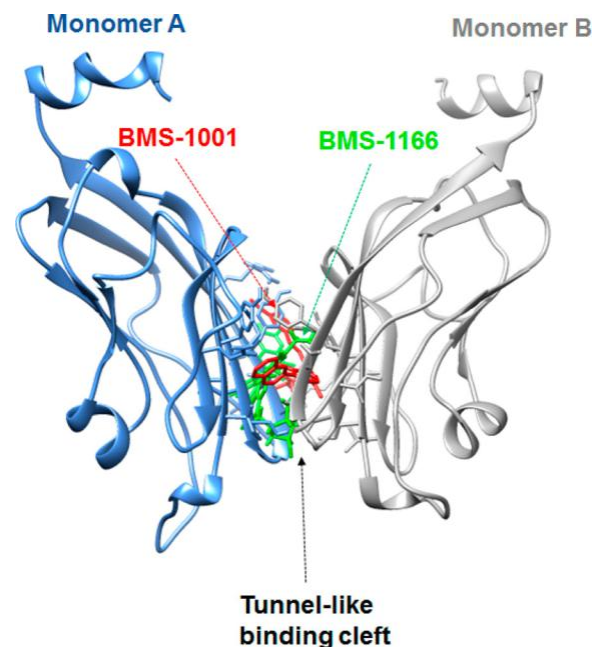


Figure 2. (Colour online) 3D structure of PD-L1 monomer A (Blue) and monomer B (Gray), in complex with BMS-1001 (Red) and BMS-1166 (Green) at the tunnel-like binding cleft formed upon dimerisation. PDB: 5NIX [20]. BMS-1001 was superimposed on the crystal structure.

to target with small molecule inhibitors due to unresolved mechanistic mode of action. We believe this model would provide essential insights into this mechanism and interesting possibilities towards the future design of novel high-affinity small-molecule antagonists and monomeric recruiters of PD-L1 in ICB therapy.

Molecular dynamics (MD) simulation in recent years has metamorphosed from simulating several hundreds of atoms to systems with biological activity, including entire proteins in solution with explicit solvent representations, membrane embedded proteins, or large macromolecular complexes [23]. Other recent applications of MD simulation has been reported in the study of disease associated nsSNPs [24–27].

## 2. Computational methods

### 2.1. System preparation

X-ray crystallographic structures of dimerised PD-L1s in complex with BMS-1001 and BMS-1166 were retrieved from

Protein Data Bank with entries 5NIU and 5NIX respectively [20]. These were prepared by the removal of co-crystallized molecules such as 1,2-ethanediol (EDO) and crystallographic water using UCSF Chimera [28]. Moreover, missing residues which majorly constituted terminal amino acids were added with the aid of MODELLER, an in-silico tool used for structural remodelling [29]. In order to elucidate the mechanisms of BMS-1001/BMS-1166-induced dimerisation, which was in line with the aim of this study, it was important to obtain a starting structure that could possibly reflect the recruitment of a PD-L1 monomer in the presence of another BMS-bound PD-L1 monomer as experimentally depicted in a previous study [20]. To this effect, the ligand-dimer complexes were first separated into non-complexed monomers using the Graphical User Interface (GUI) of UCSF Chimera. The  $\beta$ PD-L1 was first removed from the dimerised complex (5NIU) leaving a BMS-1001-bound  $\Delta$ PD-L1. This was followed by the molecular protein-protein repositioning of  $\beta$ PD-L1 using UCSF-Chimera integrated Autodock Vina interface on manually defined co-ordinates which differs from the crystallized 'dimer' structure. The same paradigm was employed for preparing BMS-1166 bound systems (5NIX) to generate the desirable starting structures which entail ligand-bound  $\Delta$ PD-L1s with a distant  $\beta$ PD-L1 monomer in the same conformational space phase as shown in Figure 3. In the same vein, a replica system devoid of the small molecule antagonists (BMS-1001/ BMS-1166) was prepared, representative of the unbound form. This was necessary to clearly define the dimerising effects of both BMS compounds, which underlies their antagonistic functions. Altogether, these systems were set up for an MD simulation of 150 ns using in-house protocols previously reported [30–32]. The Graphical Processor Unit (GPU) version (Particle Mesh Ewald Molecular Dynamics - PMEMD) of the AMBER14 software in addition to its integrated modules were used to perform MD simulations [33]. The ANTECHAMBER module was used to parameterise the ligands and generate atomic partial charges (Gasteiger - gaff) using the bcc charge scheme while the FF14SB force-field was used to define the PD-L1 parameters [34]. Protein modification, renaming and protonation (histidine) were done using the in-house pdb4amber script. This was followed by the generation of topology and parameter files for the respective PD-L1-BMS1001/BMS1166 complexes using the LEAP module. This was also used for system neutralisation via the addition of counter ions at a constant pH (cpH) and solvation within a TIP3P water box of 10 Å. Using a restraint potential of 500kcal/mol Å, partial minimisation of 2500 steps was carried out followed by 5000 full minimisation steps without energy restraints. Gradual system thermalisation from 0 to 300 k was also carried out in a canonical ensemble (NVT) for 50ps with the aid of a Langevin thermostat and a 5kcal/mol Å<sup>2</sup> harmonic restraint. System equilibration was then carried out for 1000ps at 300k without restraints while atmospheric pressure was maintained at 1bar using the Berendsen barostat [35]. This was followed by an MD production run of 150 ns where all atomic hydrogen bonds were constricted using the SHAKE algorithm [36]. Moreover, at every 1ps, resulting coordinates and trajectories were saved, which were analyzed with the integrated CPPTRAJ and PTRAJ modules [37]. Furthermore, data plots were made

with Origin data analytical tool [38] while the Graphical User Interface (GUI) of UCSF Chimera was used for visual analysis.

## 2.2. Binding free energy estimation

The Molecular Mechanics/ Poisson-Boltzmann Surface Area (MM/PBSA) method was used to estimate the differential binding of BMS-1001 and BMS-1166 to PD-L1 [39]. This is an end-point energy calculation that can be used to predict binding strength and affinities of small molecules towards their respective target proteins. MM/PBSA is mathematically represented as follows:

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

$$E_{\text{gas}} = E_{\text{int}} + E_{\text{vdw}} + E_{\text{ele}} \quad (2)$$

$$G_{\text{sol}} = G_{\text{GB}} + G_{\text{SA}} \quad (3)$$

$$G_{\text{SA}} = g\text{SASA} \quad (4)$$

where  $E_{\text{gas}}$ , the gas-phase energy,  $E_{\text{int}}$  is the internal energy;  $E_{\text{ele}}$  and  $E_{\text{vdw}}$  are the Coulomb and van der Waals energies, respectively. Also,  $G_{\text{sol}}$  is the solvation free energy and  $G_{\text{GB}}$  is the polar solvation contribution.  $G_{\text{SA}}$  is the non-polar solvation contribution and was estimated by the solvent accessible surface area (SASA) determined using a water probe radius of 1.4 Å. The surface tension constant  $\gamma$  was set to 0.0072 kcal/(mol·Å<sup>2</sup>).

## 2.3. Dynamic cross-correlation matrix (DCCM)

This is graphically employed to represent dynamical motions of constituent protein residues in correlation with time [40,41]. Therefore, in this study, we investigated the inter-residue dynamics and motions of PD-L1 when bound by BMS-1001 and BMS-1166 with respect to their antagonistic activities. We generated the DCCM to estimate the cross-correlated displacement of protein backbone atoms from the resulting trajectories across the MD simulation time based on the equation below:

$$C_{ij} = \frac{1}{2} \left( \frac{r_i \cdot r_j}{|r_i| |r_j|} \right)$$

From the equation,  $r_i$  and  $r_j$  depict the mean displacement the  $i$ th and  $j$ th atoms which are represented by  $i$  and  $j$  respectively.  $C_{ij}$  is the coefficient of cross-correlation which varies from -1 to +1, which corresponds to strongly correlated (+) and anti-correlated (-) motions of constituent residues during the period of MD simulation. This analysis was performed using the integrated CPPTRAJ module [37] of AMBER14 and resulting matrices were plotted with the aid of the Origin data analysis software [38].

## 2.4. Principal component analysis (PCA)

This was carried out to obtain necessary insights into the 3D conformational and dynamical changes that occurred in PD-L1 when distinctively bound by both BMS-1001 and BMS-1166. This method is used to describe the magnitude and direction of protein motions which are depicted by eigenvalues and eigenvectors respectively [31,42]. Hence, with the aid of the

integrated CPPTRAJ module in AMBER14, we computed the first two principal components (PC1 and PC2) from the dynamics of the protein C- $\alpha$  atoms. Moreover, the conformational behaviours of the unbound and bound PD-L1 systems were projected along the first two eigenvectors (ev1/PC1vs ev2/ PC2) using the Cartesian coordinates of C- $\alpha$  atoms.

### 3. Results and discussion

#### 3.1. Monomeric binding of BMS compounds (BMS-1001 and BMS-1166) induced systematic PD-L1 dimerisation

The mechanisms by which the selective ‘monomeric’ targeting of PD-L1 by BMS compounds cause structural dimerisation has remained unresolved till date even though it underlies the basis of ICB therapy. According to previous reports, PD-L1 exist in a single monomeric state where it is able to exhibit canonical interactions with PD-1 resulting in cancer immunosuppression and evasion of T-cell induced death [15,16]. Moreover, BMS compounds reportedly binds initially to a PD-L1 monomer ( $A_{PD-L1}$ ) and elicit monomeric recruitment ( $B_{PD-L1}$ ) to form a dimer with a tunnel-like binding cleft in between [20]. Firstly, we visually examined the structural events that occurred

along the MD trajectories to monitor the dynamics of monomeric recruitment with regards to the binding of the BMS compounds. (Figure 3)

Interestingly, in the BMS-bound systems, we observed that the arrangements of the starting structures were maintained until about 20 ns where monomeric distances gradually reduced, and the monomers existed in close proximities. These dynamical monomeric motions in the conformational phase space then proceeded until 90 ns where we observed that both monomers associated to form a PD-L1 dimer and a tunnel-like cleft at the region where the BMS compounds were bound (Figure 3). This was evidenced by the gradual decrease in monomeric distances as shown in Figure 4, which is derived from measuring distances between  $A_{Tyr39}$  and  $B_{Val185}$ . As estimated, in the starting structure, positional distances between the two monomers ( $A_{PDL1}$  and  $B_{PDL1}$ ) were 76.79 Å and 72.83 Å for the BMS-1001- and BMS-1166-bound systems respectively, which gradually decreased in the average intermediate structure to 65.14 Å (BMS-1001) and 69.23 Å (BMS-1166) prior to eventual dimerisation. On the contrary, while these events gradually occurred in the BMS-bound systems, the unbound PD-L1 system maintained its monomeric arrangement over the simulation period indicating

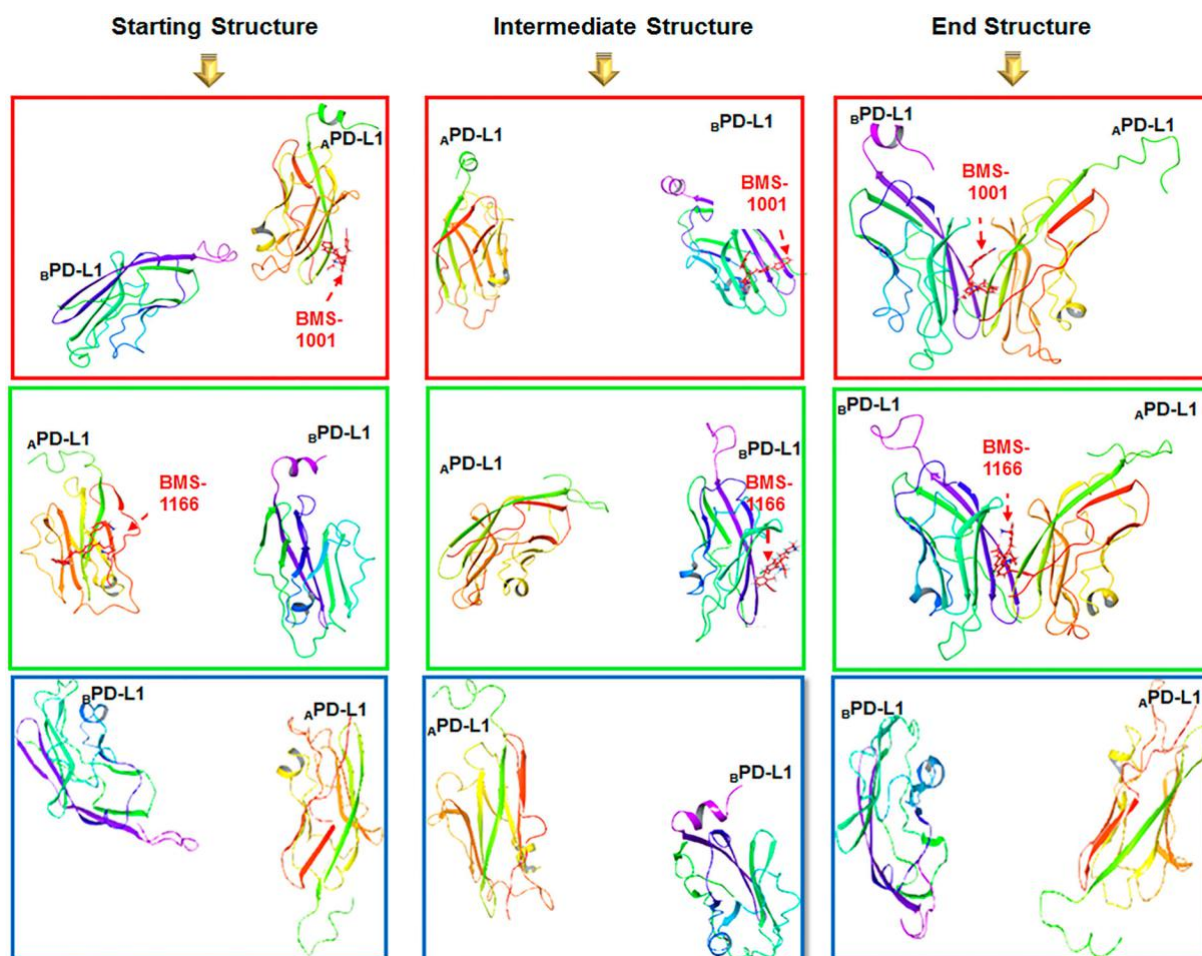


Figure 3. (Colour online) Structural dynamics of BMS-1001/BMS1166 bound PD-L1s and unbound PD-L1 over the 150 ns MD simulation period. Red boxes depict the monomeric motions of BMS-1001-bound PD-L1 while green boxes depict the dynamics of the BMS-1166-bound PD-L1 system. However, blue boxes depict the unbound PD-L1 system. [A] Starting structure averaged over 0–20 ns [B] averaged intermediate structure 60–90 ns and [C] averaged end structure 90–150 ns.

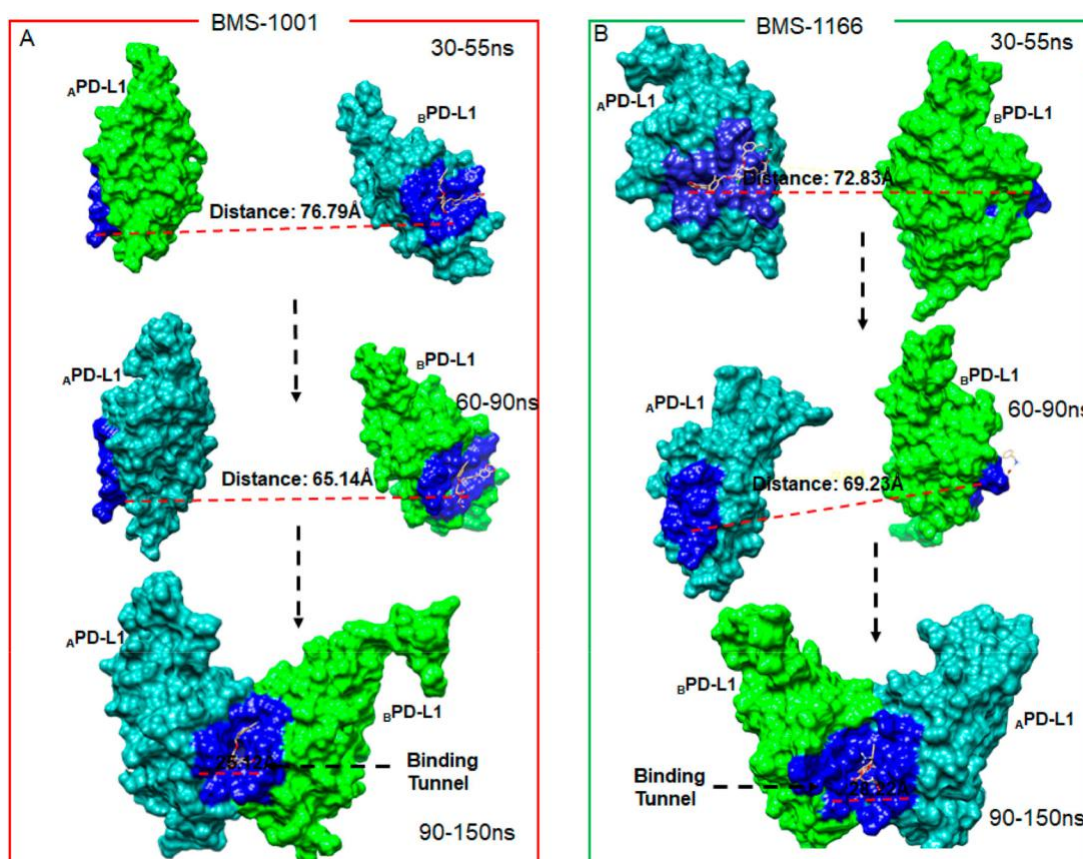


Figure 4. (Colour online) Monomeric distances, tunnel formation and mechanistic transition of BMS-1001 and BMS-1166 across both A and B PD-L1 monomers. [A] Tunnel formation and monomeric distance of BMS-1001 bound PD-L1 [B] Tunnel formation and monomeric distance of BMS-1166 bound PD-L1.

that the PD-L1 dimerisation is solely based on the presence of the BMS compounds (Figure 3).

Another interesting finding was the concurrent dynamical ‘switch’ or exchange of both BMS compounds at each monomeric binding sites. At trajectory 30–55 ns, averaged structure revealed that BMS-1001 was bound to selected site residues such as Met224, Val185, Ile163 and Val177 on monomer A via weak interactions (Figure 5A). Presumably, the nature of these interactions permitted a possible transition to monomer B which occurred at about 60–90 ns and characterised by the occurrence of high-affinity interactions, which include the strong hydrogen bonds and attractive salt bridge with Arg108 and Asp105 respectively (Figure 5B). Relative to BMS-1166, a similar ligand transition pattern between both A<sub>PD-L1</sub> and B<sub>PD-L1</sub> monomers was observed along the trajectories as revealed by the averaged structures; where Thr3, Phe2, Lys107 contributed to the binding of BMS-1166 via high-affinity and strong hydrogen interactions while Met98 formed attractive salt-bridge as shown in (Figure 5D). Moreover, the mechanistic transitions of both BMS compounds occurred prior to the eventual formation of the tunnel-like binding cleft upon PD-L1 dimerisation. We could propose that the dynamical motions of the PD-L1 monomers was accompanied by a concurrent binding site interchange of the BMS compounds prior to dimerisation.

3.2. BMS compounds induced high conformational flexibility favourable for monomeric motion and dimerisation

### 3.2.1. Conformational stability and residual fluctuation

We proceeded to measure the dynamical events that accompany the formation of A<sub>B</sub>PD-L1 dimer as induced by the monomeric-binding of BMS-1001 and BMS-1166 respectively. We believe these insights would provide structural details that characterise the monomeric recruitment of PD-L1 in the presence of the BMS compounds and mechanistic details into their motions in the conformational phase space. To understand these distinct dynamical motions in relation to the binding of these antagonists, we employed metrics such as the root mean square deviation (RMSD), C $\alpha$  root mean square fluctuation (RMSF) and radius of gyration (RoG), which are appropriate to determine C $\alpha$  atomistic deviations and structural stability with the target protein over the MD simulation time.

Structural stability was first investigated among the unbound and BMS-bound systems using the C- $\alpha$  RMSD parameter in order to ensure that systems were appropriately equilibrated. Our findings showed that the systems converged and attained stability early in the production run (~10 ns) until about 100 ns where there was a distinct separation in atomistic motions among the three systems until the end of the simulation period (Figure 6). Interestingly, in agreement with the

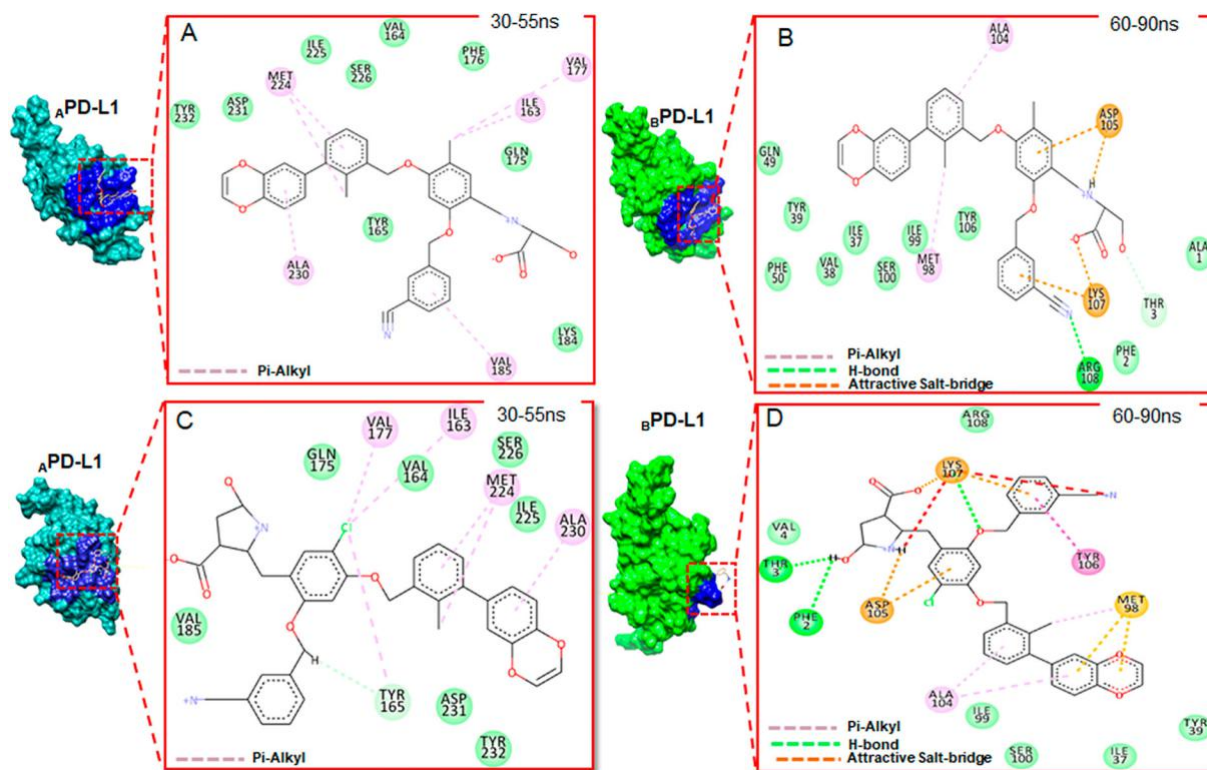


Figure 5. (Colour online) Visual representation of the monomeric active site residues interaction of BMS-1001 and BMS-1166 at 30–55 ns (A, C) and 60–90 ns (B, D).

results earlier presented (Section 3.1), this dynamic event correlates with the duration at

molecules. In other words, the association of the PD-L1 monomers in both BMS-1001 and BMS-1166-bound systems involved considerable conformational alterations and instability, a structural attribute that characterise biological protein-protein interactions. These could further indicate that the binding of the BMS compounds played a crucial role in facilitating the interactions between both monomers with respect to PD-1 blockade.

According to the plot, the unbound system maintained a stable structure throughout the simulation time (RMSD < \* Å) while the binding of BMS-1001 caused a notable deviation in PDL-1 structure (RMSD ~5.5 Å). Likewise, atomistic deviation (RMSD) of ~4 Å characterised the overall structure of PD-L1 when bound by BMS-1166. Taken together, we can suggest that the binding of these compounds induced notable instability in the structure of PD-L1 relative to inhibition. On the average, unbound PDL-1 system had RMSD value of 1.66 Å while BMS-1001- and BMS-1166-bound systems had mean RMSD values of 2.28 Å and 2.21 Å.

In order to accurately measure the degree of structural instability induced by both PD-L1 antagonists, we carried out time-specific investigations into the occurrences from ~100 ns where distinct separation in motion occurred among the three systems. This is clearly presented in Figure 6B, and accordingly, structural instability was highly induced in PD-L1 when bound by both BMS-1001 and BMS-1166, which is highly favourable for structural dimerisation. In addition, we evaluated the distinct motions of the BMS compounds over the simulation period in relation to their roles in

dimerisation from the separate monomeric PDLs, most importantly from the period at which we presumed dimer formation (~100 ns → ~150 ns). As estimated, relative to their respective positioning and motions at the tunnel-like binding cleft upon PD-L1 dimerisation, BMS-1001 has a RMSD value of 2.28 Å while BMS-1166 exhibited a deviation of 2.21 Å. However, the disparate motions exhibited by both compounds lasted until ~140 ns where they both attained stability in dynamical motion (Figure 6A). Therefore, from this plot we can deduce that the deviations in motions early in the simulation period could be related to their ‘monomeric inter-change’ as earlier mentioned earlier, which was later stabilised by the formation of the binding tunnel and corresponding interactions with constituent residues. Moreover, the dynamics of the PD-L1 binding sites were evaluated from ~100 ns in relation to the non-active sites region. Our findings showed that while the non-active site regions exhibited high structural instability favourable for dimerisation, the BMS-bound active sites showed lower motions indicative of the stabilising effects of BMS-1001 and BMS-1166 respectively (Figure 6C). (Figure 7)

Accordingly, while the active site (C-α) of unbound PD-L1 had an RMSD of ~2.4 Å, the BMS-bound systems had lower deviations at their active site regions. On the average, BMS-1001- and BMS-1166-bound systems had active site RMSDs of 1.4 Å and 1.7 Å respectively. In addition, a reduction in atomistic motions among active site residues could be due to the inward pulling effects elicited by complementary interactions with the bound BMS compounds. Further structural insights were provided by measuring the radius of gyration (RoG) with regards to the time-specific structural events (~100ns-



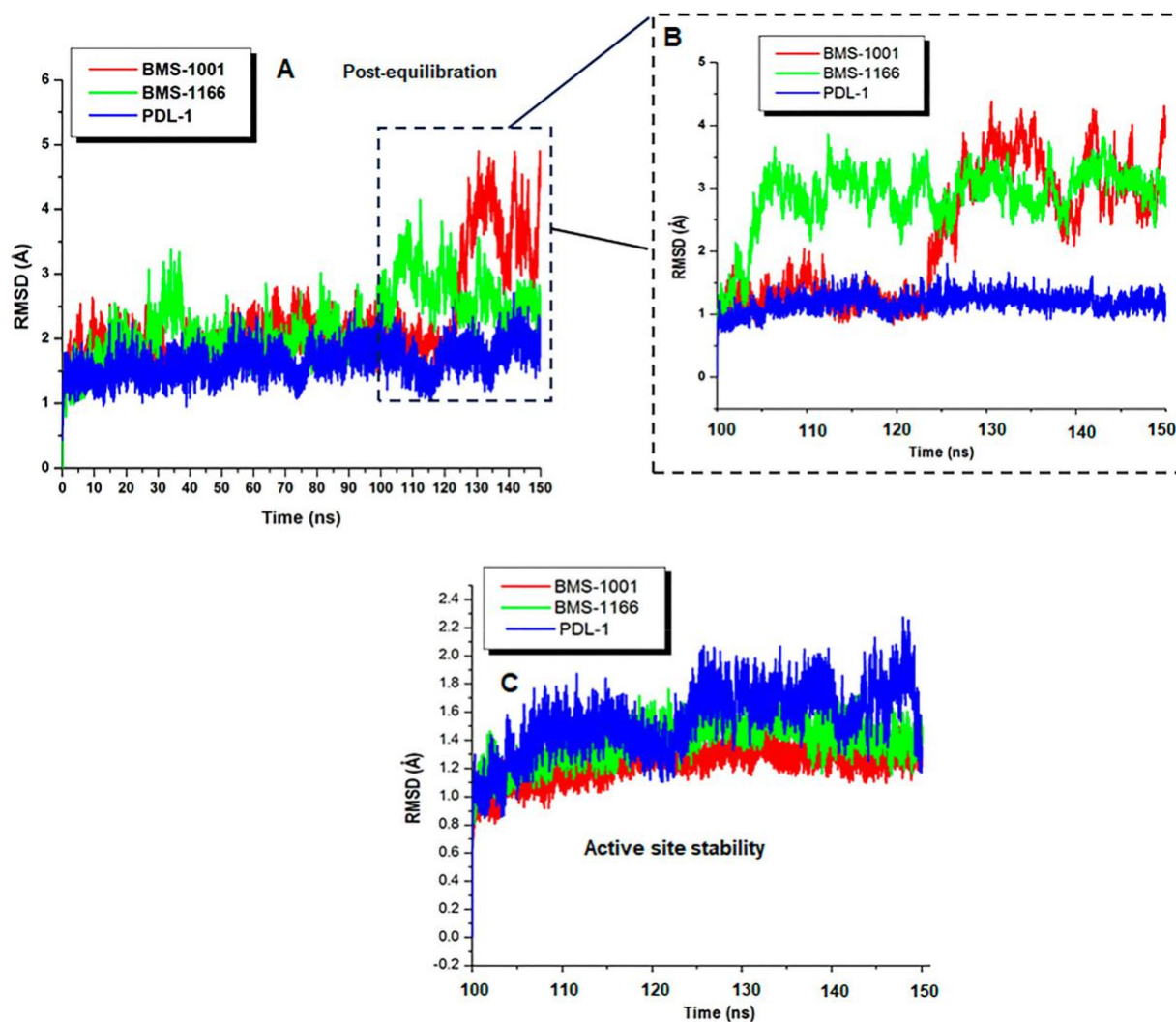


Figure 6. (Colour online) Conformational analysis plot showing stability and atomic motions among unbound (blue), BMS-1001 (red) and BMS-1166 (green) bound systems. [A] Whole-structure comparative C- $\alpha$  RMSD plot of the three systems. Blue highlight depicts post-equilibration periods (100–150 ns) where distinct separation in motions occurred [B] Comparative C- $\alpha$  RMSD plot (post-equilibration  $\rightarrow$  100–150 ns) showing distinct variations in structural stability among the three systems [C] C- $\alpha$  RMSD plot showing active site stability among the unbound (blue), BMS-1001- (red) and BMS-1166- (green) bound systems.

150 ns post convergence) in the active and non-active site regions with reference to the RMSD. According to previous studies, high RoG value indicates an increase in motion of constituent residues around their C- $\alpha$  atoms, which also corresponds to a notable loss in structural compactness [32,43–45]. On the contrary, a low RoG value suggests a structurally compact protein with minimal motions in constituent atoms. Our findings revealed that the binding of BMS-1001 and BMS-1166 induced considerably high atomistic motions across the secondary structure of PD-L1 as compared to the unbound system with low RoG. Also, these could imply that while unbound PD-L1 was structurally compact (due to limited motions), the binding of these compounds increased structural motions and activity which could in turn favour dynamical motions of both monomers with respect to dimer formation which corroborates earlier results and previous experimental reports [19,20]. As estimated, unbound PD-L1 had mean RoG values of 18.7 Å while BMS-1001- and BMS-1166-bound PD-L1 had mean RoG values of 20.2 Å and 20.0 Å respectively. Estimations of the root mean square fluctuation

(RMSF) and DCCM for all simulated models were also conducted to gain additional insights into PD-L1 structural occurrences when bound by BMS-1001 and BMS-1166. With regards to RMSF analyses, both monomers (A and B) were evaluated differently in order to define the distinct monomeric motions relative to dimerisation as induced by the binding of the BMS compounds. As observed in Figure 8, high residual fluctuation occurred concurrently in both monomers of PD-L1 when bound by BMS-1001 and BMS-1166, depictive of a structural characteristic that is important for the dynamical motions of proteins with regards to biological interactions with other biomolecules. As estimated, both monomers (APD-L1 and BPD-L1s) for the BMS-1001 and BMS-1166 systems was highly flexible in structure over the simulation period in contrast to the unbound system which demonstrated low structural flexibility and motion. This is in agreement to the results presented above and further indicates that the binding of these compounds induced structural flexibility in PD-L1 monomers which is suitable for dimerisation. On the other hand, low flexibility in the unbound system could be due to minimal

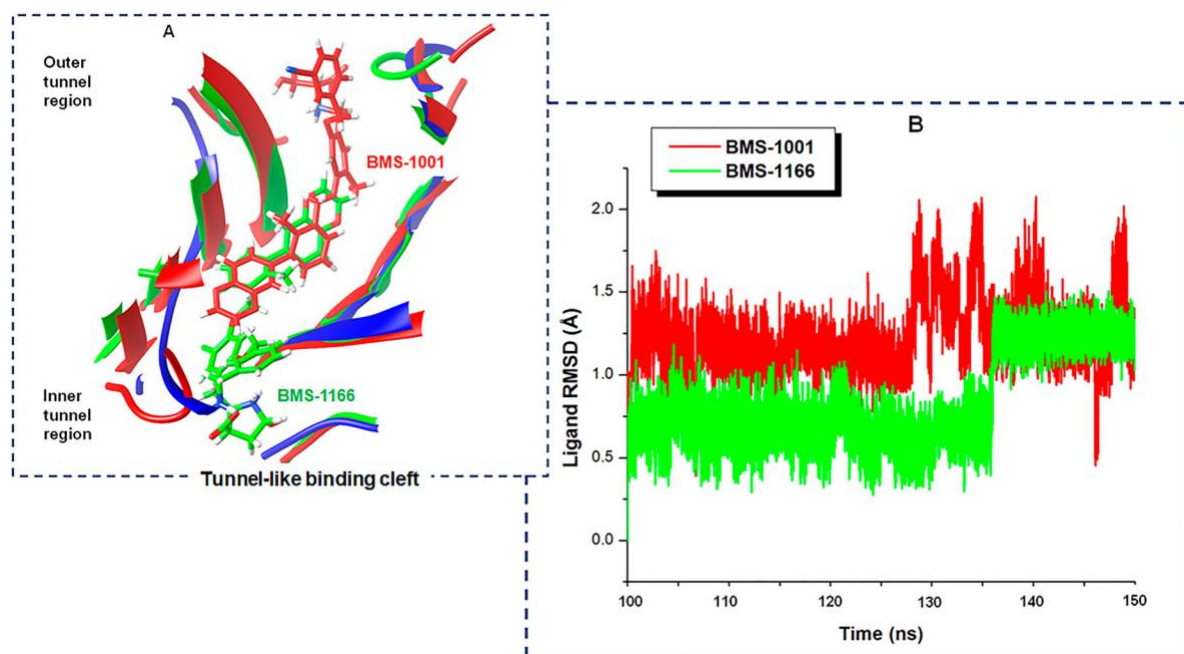


Figure 7. (Colour online) Structural stability, positioning and motion of BMS-1001 and BMS-1166 at the tunnel-like binding cleft of PDL-1. [A] Superposition of BMS-1001 (red) and BMS-1166 (green) as they transverse the binding tunnel [B] Comparative C- $\alpha$  RMSD plot depicting the stability and motion of both BMS-1001 (red) and BMS-1166 (green) during the post-equilibration MD-simulation period (100ns–150 ns).

Table 1. Per-residue fluctuation values of crucial binding cleft residues in the unbound and bound PD-L1 systems.

Core active loop residues	Unbound PD-L1	BMS-1001-bound PD-L1	BMS-1166-bound PD-L1
		Per-residue fluctuation (Å)	
ATyr56	3.13	4.01	5.45
AMet115	2.63	4.93	5.13
AAla121	4.83	9.47	6.70
AAsp122	4.08	8.35	5.48
ATyr123	3.32	6.94	6.08
ALys124	3.55	7.14	5.21
AArg125	3.85	7.73	7.05
Ble54	5.91	14.09	9.81
BTyr56	4.38	15.70	9.72
BAsn63	5.05	16.28	11.65
BVal68	6.61	16.95	10.04
BVal76	5.58	19.75	10.87
BMet115	4.33	15.57	9.26
BAla121	5.36	15.96	9.17

structural and monomeric motions in the absence of the BMS compounds (Figure 8).

Upon the presumed dimer formation, we further estimated the per-residue fluctuation of specific binding cleft residues that played key interactive roles with BMS-1001 and BMS-1166 (Table 1). Our findings revealed that these residues exhibited high fluctuation, which, suggestively, could be due to the tra-versing motions and optimal positioning of both compounds at the tunnel-like binding cleft. Taken together, high structural flexibility induced in both PD-L1 monomers with regards to the respective binding of BMS-1001 and BMS-1166 could indi-cate high structural activity, an occurrence that could in turn enhance PD-L1 protein-protein interactions (dimerisation). This reflected the characteristic antagonistic activity elicited by both compounds as previously reported [11,19,20].

On the average, BMS-1001-bound APD-L1 had a mean RMSF values of 9.1 Å while the corresponding BPD-L1 had 15.2 Å while it was 7.8 Å and 17.0 Å in the BMS-1166-bound APD-L1 and BPD-L1 respectively. Also, it is important to mention that in comparison to the BMS-bound APD-L1s, the non-liganded BPD-L1s exhibited higher flexibility prior to dimerisation which could indicate that structural motions were higher in relation to monomeric recruitment. In other words, BPD-L1s exhibited higher motions than the liganded-APD-L1s in the conformational phase space, which could corre-late with the mechanisms of BMS-induced dimerisation as previously reported [20]. On the contrary, unbound APD-L1 and BPD-L1 had lower RMSF values of 5.1 Å and 7.6 Å respectively, predictive of a relatively low structural flexibility. Taken together, the binding of both compounds concurrently induced characteristic flexibility in PD-L1 structure across both monomers, favourable for dimerisation.

To further investigate the inhibitory impact of BMS-1001 and BMS-1066 on the inter-residual dynamics and motions of PD-L1, dynamic cross correlation matrix (DCCM) analysis was performed. This method was used to obtain necessary insights into the correlated and anti-correlated motions of con-stituent residues with respect to their C- $\alpha$  atoms over the simu-lation period. Positive correlated motions were depicted by a yellow to deep red colouration whereas anti-correlated (nega-tive) motions where represented by a cyan to black colouration. Plots generated for the simulated protein models (across both monomers) are displayed in Figure S6. As shown, correlated motions were highly prominent in the PD-L1 systems bound by BMS-1001 and BMS-1166 as compared to the unbound sys-tem which had lesser correlated motions but higher dominance of anti-correlated residual motions (cyan/blue colorations). In agreement with reports earlier presented, these findings depict

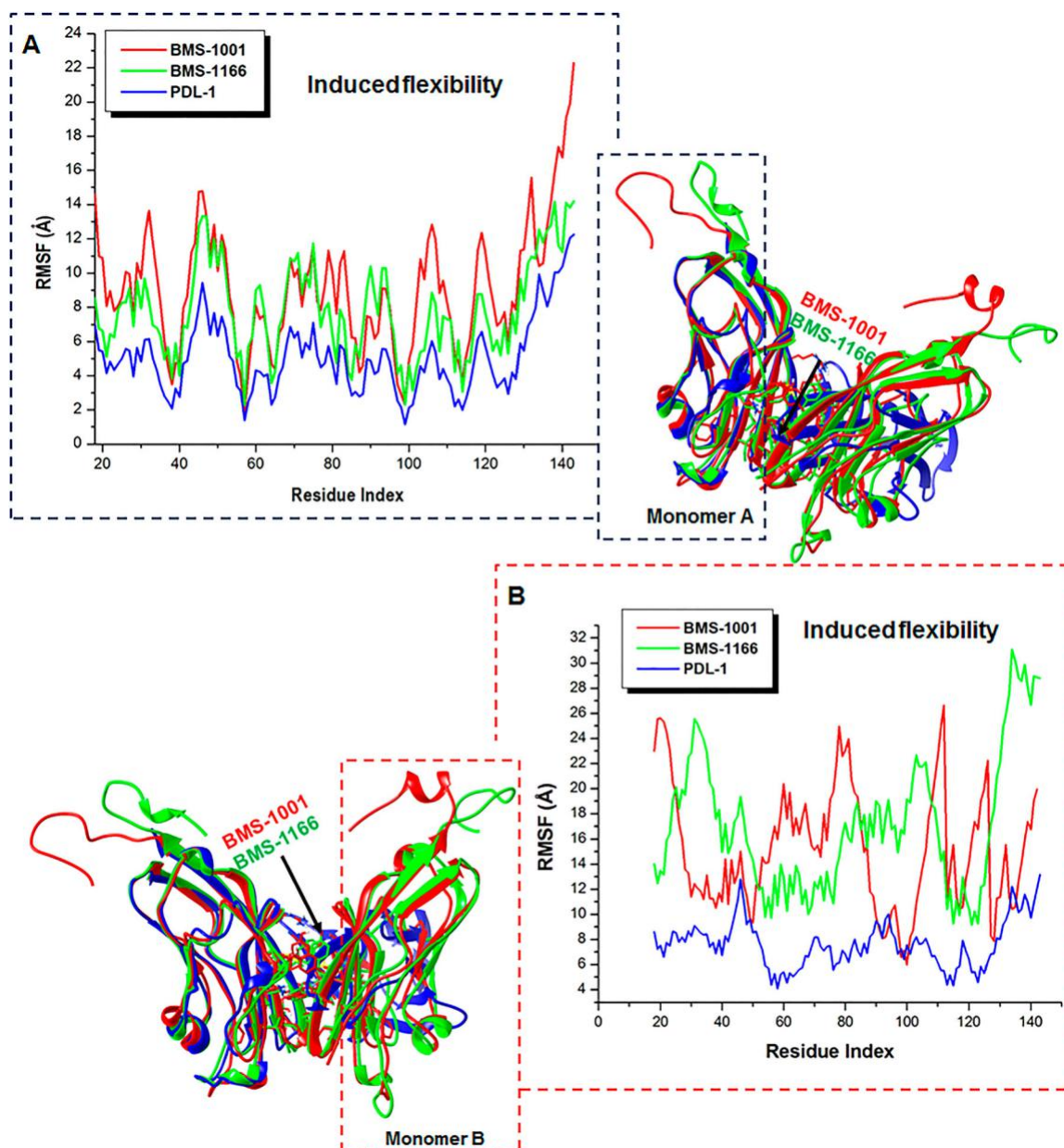


Figure 8. (Colour online) Structural flexibility induced simultaneously at both PDL-1 monomers (A and B) by the binding of BMS-1001 (red) and BMS-1166 (green). [A] Superposed monomers A and corresponding C- $\alpha$  RMSF plot showing the degree of flexibility among the unbound (blue), BMS-1001- (red) and BMS-1166- (green) bound PDL-1 systems [B] Superposed monomers B of unbound (blue), BMS-1001- (red) and BMS-1166- (green) bound PDL-1 systems coupled with corresponding C- $\alpha$  RMSF plot showing the degree of flexibility.

that the binding of these antagonists induce high motions among the constituent atoms of PD-L1 which correspond to high structural activity.

### 3.2.2. Principal component analysis of structural dynamics and motion

PCA also known as essential dynamics analysis is one of the most important post-dynamic analysis used to study the trajectory of a system over a simulation period [46]. The most significant fluctuation modes of a protein together with the motion of the system in terms of planarity of motion (eigenvectors) and its magnitude (eigenvalues) can be determined using PCA.

PCA analysis was performed on the protein C- $\alpha$  atoms using the CPPTRAJ module in AMBER14 to compute the first two

components (PC1 and PC2). Figure 8 shows the overall motional shifts across two principle components for the free proteins and ligand-protein complex.

To gain insights into the significance of such motions with regards to our study, the conformational attributes of unbound and bound PD-L1 systems were projected directionally along the first two principal components or eigenvectors (ev1/PC1 vs ev2/PC2). Accordingly, distinct conformations and motions were observed in the essential subspace along both principal components. The plot showed a clear and distinct separation of motion among the three systems, with the BMS-bound PD-L1 exhibiting a highly dispersed motion as compared to the unbound system, which showed a more compact motion. Induced dispersion among the protein systems bound by

BMS-1001 and BMS-1166 could only imply that the binding of these compounds enhanced structural activity unlike in the unbound system which maintained structural compactness across the simulation period. These findings further corroborates above-presented results which altogether reflects that the binding of these compounds stabilised a high activity conformation favourable for dimerisation as previously reported [18–20]. Suggestively, high activity conformation relative to PD-L1 dimerisation as induced by these BMS antagonistic could disallow its association with PD-1 relative to immune checkpoint activities in T cell exhaustion.

### 3.3. BMS-1001 and BMS-1166 exhibited 'monomeric interchange' and systematic high-affinity binding prior to PD-L1 dimerisation

The antagonistic mechanism of the BMS-compounds towards PD-L1 is centred on their ability to induce dimerisation thereby preventing an association with PD-1 which promotes tumour progression and survival by escaping tumour neutralising immune surveillance. Association of PD-L1 with PD-1 of T-cells lead to T Cell dysfunction, exhaustion, neutralisation, and interleukin-10 (IL-10) production in tumour mass [13].

According to previous reports, the BMS-compounds bind to a PD-L1 monomer and facilitate monomeric recruitment leading to the formation of a dimer molecule [20]. Interestingly, from our previous findings, as presented in Section 3.1 above, we proposed a model that possibly depicted the dynamical motions of the respective monomers in the conformational phase with respect to the starting structure employed in this study. Moreover, we observed that the BMS-compounds systematically transited from the initial monomer ( $\text{APD-L1}$ ), to which they were first bound, onto the second monomer prior to dimerisation (Figures 4 and 5). As shown in the initial trajectory (0–30 ns), the BMS-compounds maintained their starting position at the first monomeric binding site ( $\text{APD-L1}$ ) until

ns where we observed that it was bound at a similar binding site on monomer  $\text{BPD-L1}$ , and eventually at the tunnel-like binding cleft upon the formation of PD-L1 dimer. This could represent an important insight into the mechanistic occurrences that accompany BMS-induced dimerisation of PD-L1 in cancer immunotherapy. The MM/PBSA analysis was further employed to deduce the binding strength and affinity of the BMS compounds with regards to the dimerised protein form. This estimated the binding free energies involved in complex formation in relation to their antagonistic activities. Using trajectories at which the PD-L1 dimer was formed in both BMS-1001 and BMS-1166-bound systems, our estimations revealed that both compounds exhibited considerably high and favourable negative energies indicative of a strong and high affinity

binding. Accordingly, BMS-1001 had a  $G$  value of  $-20.5\text{kcal/mol}$  while BMS-1166 had a  $G$  value of  $-18.1\text{kcal/mol}$ . Relatively, we estimated the monomeric  $G$  values of the BMS compounds to each monomer at the respective simulation periods. We observed that both BMS-1001 and BMS-1166 exhibited similar  $G$  towards the PD-L1 monomers even when interchanged at their respective 'monomeric' binding sites as earlier mentioned. BMS-1001 had an estimated  $G$  of  $23.64\text{kcal/mol}$  and  $23.75\text{kcal/mol}$  at the initial (0–30 ns) and

Table 2. MM/PBSA binding free energy profiles of BMS-1001 and BMS-1166 to PD-L1.

Complexes	Evdw (kcal/mol)	Eele (kcal/mol)	Ggas (kcal/mol)	Gsol (kcal/mol)	Gbind (kcal/mol)
MD1 BMS-1001	-35.9	39.12	36.0	28.1	23.64
MD2 BMS-1001	-28.12	45.12	34.0	27.1	23.75
MD3 BMS-1001	-34.8	-27.9	-42.7	30.8	23.78
MD1 BMS-1166	-26.78	40.15	37.92	26.87	23.43
MD2 BMS-1166	-27.42	44.91	38.00	27.32	23.66
MD3 BMS-1166	-27.9	-48.0	-39.0	29.7	25.96

$E_{\text{ele}}$  = electrostatic energy;  $E_{\text{vdW}}$  = van der Waals energy;  $G_{\text{bind}}$  = total binding free energy;  $G_{\text{sol}}$  = solvation free energy  $G$  = gas phase free energy.

intermediate (60–90 ns) trajectories while BMS-1166 had  $23.43\text{kcal/mol}$  and  $23.66\text{kcal/mol}$  respectively. At the final trajectory, BMS-1001 and BMS-1166 had  $G$  values of  $23.78\text{kcal/mol}$  and  $25.96\text{kcal/mol}$  respectively. MM/PBSA estimations for the distinct MD simulation periods are presented in Table 2. Therefore, these considerably high negative free energy values could correlate with a high-affinity binding that is essential for the systematic dimerisation of PD-L1 in the presence of the respective BMS-compounds [20].

Upon dimerisation, additional insights into the inter-actions and positioning of BMS-1001 and BMS-1166 at the tunnel-like binding cleft were gained by molecular visualisations using the Graphical User Interface of UCSF Chimera and Discovery Studio 2016 Client [47]. These revealed inter-molecular interactions that occurred at the 'dimeric' binding cleft of PD-L1, which is systematically induced after the binding of a BMS compound (BMS-1001/BMS-1166) to a PD-L1 monomer, as elucidated in the model presented in this study. As observed at each of the 'dimeric' binding sites, both compounds transversed the tunnel-like binding cleft while high-affinity interactions occurred between constituent ligand groups and key residues lying on either side of the monomers within and outside the tunnel to achieve optimal binding modes and stability (Figure S2). Outside the tunnel, key residues such as  $\text{AAsp122}$ ,  $\text{ATyr123}$ ,  $\text{ALys124}$  and  $\text{AArg125}$  formed strong high-affinity interactions with the (2R)-2-amino-3-hydroxypropanoic acid and 3-cyanobenzyl substituents; which majorly constitute hydrogen bonds and attractive charge (ionic) interactions that account for the stability of BMS-1001 at this region (Figure S2B). Moreover, key interactions that accounted for high-affinity binding and stability of BMS-1001 at the other side of the tunnel (binding cleft) were elicited by deep hydrophobic pocket residues;  $\text{BAla121}$  ( $\pi$ -alkyl) and  $\text{BAsp122}$  (CH–O), with the buried 2,3-dihydro-1,4-benzodioxine moiety.

Also, the orientations of  $\text{ATyr56}$  and  $\text{BTyr56}$  played important roles in the stability of BMS-1001 at the binding cleft of PD-L1 possibly via ring  $\pi$ - $\pi$  stacking (Figure S2B). Likewise, BMS-1166 had a similar orientation at the binding cleft of PD-L1 wherein it transverses the tunnel-like binding cleft and elicit high-affinity interactions at both sides of the tunnel (Figure S3). At the outer region of the tunnel, residues such as  $\text{BAsp122}$ ,  $\text{BTyr123}$  and  $\text{BLys124}$ ,  $\text{BArg125}$  existed in high-affinity (hydrogen and ionic) interactions with the (2R, 4R)-4-hydroxypyrrolidine-2-carboxylic acid moiety while  $\pi$ - $\pi$  stacking of the 3-cyanobenzyl ring by  $\text{BTyr123}$  complements its role in BMS-1166 stability around this region.

The role of  $\text{ATyr56}$  was also crucial to the stability of BMS-1166 while it transverses the tunnel possibly due to its involvement in ring  $\pi$ - $\pi$  stacking interactions with the intermediate 5-chlorophenyl ring. Likewise, the contributions of  $\text{ATyr123}$  (NH-O) and  $\text{AAsp122}$  (CH-O) towards the stability of BMS-1166 encompass their interactions with the 2,3-dihydro-1,4-benzodioxine moiety at the deep hydrophobic pocket. Furthermore, per-residue energy decomposition analyses were used to estimate the individual energy contributions of these crucial residues to the stability of BMS-1001 and BMS-1166 respectively at the binding cleft. As shown in Figure S4, energy contributions towards BMS-1001 stability at the binding cleft were considerably high among key tunnel residues (inside and outside) as earlier mentioned in comparison with other interacting residues.

Outside the tunnel,  $\text{AAsp122}$ ,  $\text{ATyr123}$ ,  $\text{ALys124}$ , and  $\text{AArg125}$  had high electrostatic energy contributions of  $-10.5\text{kcal/mol}$ ,  $-5.4\text{kcal/mol}$ ,  $-10.1\text{kcal/mol}$  and  $-6.9\text{kcal/mol}$  respectively, which could be due to the occurrence of strong hydrogen and ionic interactions as shown in Figure S4B. Inside the tunnel, key residues such as  $\text{BAla121}$  and  $\text{BAsp122}$  that interacted with the buried 2,3-dihydro-1,4-benzodioxine moiety of BMS-1001 had van der Waals and electrostatic energies of  $-3.6\text{kcal/mol}$  and  $-3.4\text{kcal/mol}$  respectively. In addition,  $\text{ATyr56}$  (inside: 2,3-dihydro-1,4-benzodioxine) and  $\text{BTyr56}$  (outside: 4-methylphenoxy ring) which elicited ring  $\pi$ - $\pi$  stacking of BMS-1001 at the binding tunnel had van der Waals energy contributions of  $-2.6\text{kcal/mol}$  and  $-3.4\text{kcal/mol}$  respectively. Similarly, residues such as  $\text{AAsp122}$  and  $\text{ATyr123}$  had high electrostatic energies of  $-5.3\text{kcal/mol}$  and  $-6.2\text{kcal/mol}$  respectively (Figure S5).

Likewise, crucial residues at the outer tunnel region that interacted with the (2R, 4R)-4-hydroxypyrrolidine-2-carboxylic acid moiety and the 3-cyanobenzyl ring had considerably high energy contributions, which further substantiate high-affinity interactions earlier reported (Figure 6A). Accordingly,  $\text{BAsp122}$ ,  $\text{BLys124}$  and  $\text{BArg125}$  had electrostatic energy contributions of  $-9.5\text{kcal/mol}$ ,  $-10.7\text{kcal/mol}$ , and  $-12.4\text{kcal/mol}$  respectively while  $\text{BTyr123}$  and  $\text{ATyr56}$  had van der Waals energy contributions of  $-4.2\text{kcal/mol}$  and  $-3.2\text{kcal/mol}$  respectively. Taken together, considerably high energy contributions among these crucial residues at the target site of PD-L1 account for high-affinity binding and stability of both BMS-1001 and BMS-1166 while they transverse the tunnel-like binding cleft to elicit their antagonistic functions.

#### 4. Conclusion

The use of small molecule compounds as immune checkpoint inhibitors represent a paradigm shift in the cancer immunotherapy (immune checkpoint blockade - ICB). This therapeutic strategy is being initiated with the identification of novel BMS compounds; BMS-1001 and BMS-1166 which specifically target PD-L1 and prevents its association with PD-1 resulting in the reactivation of T cells and effective killing of cancer cells. Hence it was necessary to investigate the activities of these antagonistic compounds coupled with the mechanisms by which they induce monomeric recruitment of the distinct PD-L1 monomers, so as to obtain necessary insights

which could aid the future design of highly effective antagonists and monomeric recruiters in ICB therapy. Appropriate investigations were carried out using MD simulation and MM/PBSA techniques, which are sufficient to provide structural and molecular occurrences associated with the binding of these compounds to PD-L1. Using a structurally separated starting structures for the unbound and BMS-bound systems, our result revealed that the BMS-1001 and BMS-1166 bound monomers elicited the recruitment of the other monomer. As the simulation progresses, there was ligand exchange from the monomer A to monomer B, this is due to the weak binding interaction that existed between the ligands and monomer A active site residues, however, monomer B active site residues provided a much stronger interaction. MM/PBSA estimations revealed that both compounds were bound to PD-L1 with considerably high-affinities as evidenced by relatively high G values. Moreover, molecular visualisation revealed that both compounds were positioned across the tunnel-like binding cleft, and while they traversed this region, they maintained high-affinity interactions (strong hydrogen and ionic bonds) with key active site residues both within and outside the tunnel via their substituent groups. More importantly were the high affinity interactions elicited by  $\text{AAsp122}$ ,  $\text{ATyr123}$ ,  $\text{ALys124}$  and  $\text{AArg125}$  with the (2R)-2-amino-3-hydroxypropanoic acid and 3-cyanobenzyl substituents of BMS-1001 outside the tunnel and  $\text{BAla121}$  and  $\text{BAsp122}$ , inside the tunnel (deep hydrophobic pocket). Also, with regards to BMS-1166, high-affinity interactions occurred outside the tunnel between  $\text{BAsp122}$ ,  $\text{BTyr123}$ ,  $\text{BLys124}$  and  $\text{BArg125}$  with the 3-cyanobenzyl and (2R, 4R)-4-hydroxypyrrolidine-2-carboxylic acid substituents while at the other side of the tunnel,  $\text{ATyr123}$  and  $\text{AAsp122}$  played crucial role in stabilising BMS-1166 via interactions with its 2,3-dihydro-1,4-benzodioxine moiety. The important roles of these residues were further revealed by per-residue decomposition analyses which estimated high individual energy contributions and could underlie the considerable binding strength and stability demonstrated by these compounds. Comparative insights into the structural dynamics and motions of PD-L1 in its unbound and bound states were provided by post-MD conformational analysis using C- $\alpha$  RMSD, C- $\alpha$  RMSF, C- $\alpha$  RoG, DCCM and PCA. An interesting and consistent pattern observed was that the binding of both compounds induced distinctively high structural motions and activity in contrast to the unbound model which demonstrated structural compactness and rigidity across the MD simulation period. Binding of the ligand to a monomer elicited the recruitment of the other monomer inducing dimerisation of the monomers, there was also a ligand exchange from monomer A unto monomer B. This transition in the structural dynamics and motion of PD-L1 when bound by BMS-1001 and BMS-1166 could be seen as the mechanism of PD-L1/PD-1 association. High per-residue fluctuation of key residues that interacted with the substituents of both compounds along the tunnel-like binding cleft could as well imply their high activity and contributions towards the favourable positioning and binding of BMS-1001 and BMS-1166. Taken together, BMS-1001 and BMS-1166 bind strongly to PD-L1 due to high-affinity interactions with crucial residues at the tunnel-like binding cleft and induce high structural activity favourable

for dimerisation, which could in turn prevent association with PD-1. Our results describe the model with which BMS -1001 and BMS-1166 recruit other monomer for dimer formation. This mechanism can further provide an explanation into the mechanism of dimer formation of protein and can be explored in the design of small molecule inhibitors of PD-L1.

## Acknowledgement

The authors acknowledge the College of Health Sciences, UKZN for their financial and infrastructural support and at the same time thank the Centre for High Performance Computing (CHPC, www.chpc.ac.za), Cape Town, for computational resources.

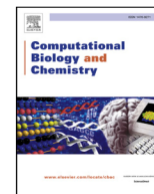
## Disclosure statement

No potential conflict of interest was reported by the authors.

## Reference

- [1] Zhang H, Chen J. Current status and future directions of cancer immunotherapy. *J Cancer*. 2018;9(10):1773–1781. doi:10.7150/jca.24577.
- [2] Jacobs EJ, Newton CC, Stevens VL, et al. Daily aspirin use and prostate cancer-specific mortality in a large cohort of men with nonmetastatic prostate cancer. *J Clin Oncol*. 2014;32(33):3716–3722. doi:10.1200/JCO.2013.54.8875.
- [3] Smyth MJ. Multiple approaches to immunotherapy - the new pillar of cancer treatment. *Immunol Cell Biol*. 2017. doi:10.1038/icc.2017.
- [4] Dine J, Gordon R, Shames Y, et al. Immune checkpoint inhibitors: an innovation in immunotherapy for the treatment and management of patients with cancer of immunotherapies reflects a promising new approach to cancer treatment involving activation of the immune system against cancer. Immune checkpoint inhibitors: an innovation in immunotherapy for the treatment and management of patients with cancer. *Asia Pac J Oncol Nurs*. 2017;4:127–135. doi:10.4103/apjon.apjon\_4\_17.
- [5] Chatamra K. Cancer and immunology. *Asian Pac J Allergy Immunol*. 1991;9(2):71–73. doi:10.21873/anticanres.11144.
- [6] Berek C. Eosinophils: important players in humoral immunity. *Clin Exp Immunol*. 2016;183(1):57–64. doi:10.1111/cei.12695.
- [7] Ponte JF, Ponath P, Gulati R, et al. Enhancement of humoral and cellular immunity with an anti-glucocorticoid-induced tumour necrosis factor receptor monoclonal antibody. *Immunology*. 2010;130(2):231–242. doi:10.1111/j.1365-2567.2009.03228.x.
- [8] Mellman I, Coukos G, Dranoff G. Cancer immunotherapy comes of age. *Nature*. 2014;480(7378):480–489. doi:10.1038/nature10673.
- [9] Aspeslagh S, Marabelle A, Soria J-C, et al. Upcoming innovations in lung cancer immunotherapy: focus on immune checkpoint inhibitors. *Chin Clin Oncol*. 2015;4(4):48. doi:10.3978/j.issn.2304-3865.2015.12.06.
- [10] Makkouk A, Weiner GJ. Cancer immunotherapy and breaking immune tolerance: new approaches to an old challenge. *Cancer Res*. 2016;75(1):5–10. doi:10.1158/0008-5472.CAN-14-2538.
- [11] Barbee MS, Ogunniyi A, Horvat TZ, et al. Current status and future directions of the immune checkpoint inhibitors ipilimumab, pembrolizumab, and nivolumab in oncology. *Ann Pharmacother*. 2015;49(8):907–937. doi:10.1177/1060028015586218.
- [12] Shore ND. Advances in the understanding of cancer immunotherapy. *BJU Int*. 2015;116(3):321–329. doi:10.1111/bju.12692.
- [13] Alsaab HO, Sau S, Alzhrani R, et al. PD-1 and PD-L1 checkpoint signaling inhibition for cancer immunotherapy: mechanism, combinations, and clinical outcome. *Front Pharmacol*. 2017;8(AUG):1–10. doi:10.3389/fphar.2017.00561.
- [22] Francisco LM, Sage PT, Sharpe AH. The PD-1 pathway in tolerance and autoimmunity. *Immunol Rev*. 2010. doi:10.1111/j.1600-065X.2010.00923.x.
- [23] Ostrand-Rosenberg S, Horn LA, Haile ST. The programmed death-1 immune-suppressive pathway: barrier to antitumor immunity. *J Immunol*. 2014;193(8):3835–3841. doi:10.4049/jimmunol.1401572.
- [24] Wang X, Teng F, Kong L, et al. PD-L1 expression in human cancers and its association with clinical outcomes. *Oncotargets Ther*. 2016. doi:10.2147/OTT.S105862.
- [25] Zhan MM, Hu XQ, Liu XX, et al. From monoclonal antibodies to small molecules: the development of inhibitors targeting the PD-1/PD-L1 pathway. *Drug Discov Today*. 2016;21(6):1027–1036. doi:10.1016/j.drudis.2016.04.011.
- [26] Abdel-Magid AF. Inhibitors of the PD-1/PD-L1 pathway can mobilize the immune system: an innovative potential therapy for cancer and chronic infections. *ACS Med Chem Lett*. 2015;6(5):489–490. doi:10.1021/acsmchemlett.5b00148.
- [27] Andrei SA, Sijbesma E, Hann M, et al. Stabilization of protein-protein interactions in drug discovery. *Expert Opin Drug Discov*. 2017. doi:10.1080/17460441.2017.1346608.
- [28] Skalniak L, Zak KM, Guzik K, et al. Small-molecule inhibitors of PD-1/PD-L1 immune checkpoint alleviate the PD-L1-induced exhaustion of T-cells. *Oncotarget*. 2017. doi:10.18632/oncotarget.20050.
- [29] Reille-seroussi M, Gagey-eilstein N, Broussy S, et al. Biophysical studies of the induced dimerization of human VEGF receptor 1 binding domain by Divalent metals competing with VEGF-A. *PLOS ONE*. 2016;11–23. doi:10.1371/journal.pone.0167755.
- [30] Maschio MC, Brancolini G, Corni S. Salt bridges regulate in silico dimers formation for  $\beta$  2 -microglobulin amyloidogenic variants. *J Self-Assembly Mol Electron*. 2019;6:35–60. doi:10.13052/jsame2245-4551.6.003.
- [31] Brandman R, Brandman Y, Pande VS. A-site residues move independently from P-site residues in all-atom molecular dynamics simulations of the 70S bacterial ribosome. *PLoS ONE*. 2012;7(1):1–10. doi:10.1371/journal.pone.0029377.
- [32] Kumar A, Purohit R. Computational investigation of pathogenic nsSNPs in CEP63 protein. *Gene*. 2012a;503(1):75–82. doi:10.1016/j.gene.2012.04.032.
- [33] Kumar A, Purohit R. Mutation research / fundamental and molecular mechanisms of mutagenesis computational screening and molecular dynamics simulation of disease associated nsSNPs in CENP-E. *Mutat Res - Fundam Mol Mech Mutagen*. 2012b;738–739:28–37. doi:10.1016/j.mrfmmm.2012.08.005.
- [34] Kumar A, Purohit R. Use of Long Term molecular dynamics simulation in predicting cancer associated SNPs. *PLoS Comput Biol*. 2014;10(4):e1003318. doi:10.1371/journal.pcbi.1003318.
- [35] Kumar A, Rajendran V, Sethumadhavan R, et al. In silico prediction of a disease-associated STIL mutant and its affect on the recruitment of centromere protein J (CENPJ). *FEBS Open Bio*. 2012;2:285–293. doi:10.1016/j.fob.2012.09.003.
- [36] Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera - a visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25(13):1605–1612. doi:10.1002/jcc.20084.
- [37] Eswar N, Webb B, Marti-Renom MA, et al. Comparative protein structure modeling using MODELLER. *Curr Protoc Protein Sci*. 2007;2(November):Unit 2.9. doi:10.1002/0471140864.ps0209s50.
- [38] Abdullahi M, Olotu FA, Soliman ME. Allosteric inhibition abrogates dysregulated LFA-1 activation: structural insight into mechanisms of diminished immunologic disease. *Comput Biol Chem*. 2018;73:49–60. doi:10.1016/j.compbiolchem.2018.02.002.
- [31] Lawal M, Olotu FA, Soliman MES. Across the blood-brain barrier: neurotherapeutic screening and characterization of narin-genin as a novel CRMP-2 inhibitor in the treatment of Alzheimer's disease using bioinformatics and computational tools. *Comput Biol Med*. 2018;98:168–177. doi:10.1016/J.COMPBIOMED.2018.05.012.
- [32] Olotu FA, Soliman MES. From mutational inactivation to aberrant gain-of-function: Unraveling the structural basis of mutant p53 oncogenic transition. *J Cell Biochem*. 2018;119(3):2646–2652. doi:10.1002/jcb.26430.

- [33] Case DA, Cheatham TE, Darden T, et al. The amber biomolecular simulation programs. *J Comput Chem*. 2005;26(16):1668–1688. doi:10.1002/jcc.20290.
- [34] Wang J, Wolf RM, Caldwell JW, et al. Development and testing of a general amber force field. *J Comput Chem*. 2004;25:1157–1174.
- [35] Berendsen HJC, Postma JPM, Van Gunsteren WF, et al. Molecular dynamics with coupling to an external bath. *J Chem Phys*. 2012;3684(May 2016):926–935. doi:10.1063/1.448118.
- [36] Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys*. 1977;23(3):327–341. doi:10.1016/0021-9991(77)90098-5.
- [37] Roe DR, Cheatham III TE. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Com*. 2013;9(7):3084–3095. doi:10.1021/ct400341p.
- [38] Seifert E. Originpro 9.1: scientific data analysis and graphing software - software review. *J Chem Inf Model*. 2014;54(5):1552. doi:10.1021/ci500161d.
- [39] Kollman PA, Massova I, Reyes C, et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res*. 2000;33(12):889–897. doi:10.1021/ar000033j.
- [40] Adeniji EA, Olotu FA, Soliman M. ). Exploring the lapse in drug-gability: sequence analysis, structural dynamics and binding site characterization of K-RasG12C variant, a feasible oncotherapeutic target. *Anticancer Agents Med Chem*. 2018;18. doi:10.2174/1871520618666180718110231.
- [41] Arnold GE, Ornstein RL. Molecular dynamics study of time-correlated protein domain motions and molecular flexibility: cytochrome P450BM-3. *Biophys J*. 1997;73(3):1147–1159. doi:10.1016/S0006-3495(97)78147-5.
- [42] David CC, Jacobs DJ. Principal component analysis: a method for determining the essential dynamics of proteins. *Methods Mol Biol (Clifton, N.J.)*. 2014;1084:193–226. doi:10.1007/978-1-62703-658-0\_11.
- [43] El Rashedy AA, Olotu FA, Soliman MES. Dual drug targeting of mutant Bcr-Abl induces inactive conformation: new strategy for the treatment of chronic myeloid leukemia and overcoming mono-therapy resistance. *Chem Biodiver*. 2018. doi:10.1002/cbdv.201700533.
- [44] Lobanov MI, Bogatyreva NS, Galzitskaia OV. Radius of gyration is indicator of compactness of protein structure. *Molekuliarnaia Biologiia*. 2008;42(4):701–706. doi:10.1134/S0026893308040195.
- [45] Salleh AB, Rahim AS, Rahman RN, et al. The role of Arg157Ser in improving the compactness and stability of ARM Lipase. *J Comput Sci Syst Biol*. 2012;5:38–46.
- [46] Sittel F, Jain A, Stock G. Principal component analysis of molecular dynamics: on the use of Cartesian vs. internal coordinates. *J Chem Phys*. 2014;141(1). doi:10.1063/1.4885338.
- [47] Biovia DS. Discovery Studio 2016 Client. San Diego: Dassault Systèmes; 2016.



## Research Article

# Drug promiscuity: Exploring the polypharmacology potential of 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones as an inhibitor of the ‘god father’ of immune checkpoint

Opeyemi S. Soremekun, Fisayo A. Olotu, Clement Agoni, Mahmoud E.S. Soliman

Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

## ARTICLE INFO

## Keywords:

Cytotoxic-lymphocytes associated protein 4  
Drug repurposing  
Polypharmacology

## ABSTRACT

High production cost, instability, low tumor penetration are some of the shortcomings that have characterized and undermined the use of antibodies as a target for Cytotoxic T-lymphocytes associated protein 4 (CTLA-4). Design and discovery of small molecule inhibitors have therefore become a sine qua non in targeting immune proteins implicated in immune disorders. In this study, we utilized a drug repositioning approach to explore the characteristic feature of unrelated proteins to have similar binding sites and the promiscuity of drugs to re-purpose an existing drug to target CTLA-4. CTLA-4 and Kallikrein-7 were found to have similar binding sites, we therefore used 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones (TDSO) which is an inhibitor of Kallikrein-7 as our lead compound. High throughput screening using TDSO as a lead compound resulted in 9 hits with ZINC04515726 and ZINC08985213 having the highest binding score. We went ahead to investigate the inter-action of these compounds with CTLA-4 by conducting a molecular dynamic simulation. Molecular Mechanics/ Poisson-Boltzmann Surface Area (MM/PBSA) estimations revealed that TDSO had the highest binding energy value of -28.51Kcal/mol, with ZINC04515726 and ZINC08985213 having -23.76Kcal/mol and -21.03Kcal/mol respectively. The per-residue decomposition highlighted Tyr24, Ala25, Gly28, Ala30, Tyr53 and Asn72 as having significantly high electrostatic energy contributions and the main contributing residues to the binding of TDSO, ZINC04515726 and ZINC08985213 to Cytotoxic T lymphocytes CTLA-4. Summarily, from the results gathered, we proposed that TDSO can be an effective immune check point small molecule inhibitor against the suppression of T-cell activation, proliferation, and tumor cell eradication.

## 1. Introduction

Cancer Immunotherapy has become a standard treatment for a wide range of cancer malignancies and is considered to be the “fifth pillar” of cancer therapy, after surgery, chemotherapy, radiation and targeted therapy (Chatamra, 1991). Immune system does not just prevent the body against infectious organisms but also help in eradicating malignancies when stimulated (Mc Granahan et al., 2016; Soremekun et al., 2019) through strengthening the host immune responses against tumors, annihilating signals produced by cancer cells that repress immune responses, or supplying modified immune system components (Chen et al., 2018). With the understanding of the functionality of the immune system, small molecules, peptides, recombinant antibodies, vaccines and cellular therapeutic modalities are being employed in manipulating the immune system for cancer treatment (Marin-Acevedo et al., 2018). The use of immune checkpoint inhibitor in cancer

Corresponding author.

E-mail address: [soliman@ukzn.ac.za](mailto:soliman@ukzn.ac.za) (M.E.S. Soliman).

immunotherapy include generating antibodies against the cytotoxic T-lymphocytes associated protein 4 (CTLA-4) (Fig. 1), the programmed death receptor 1 (PD-1) or its ligand (PD-L1) (Thallinger et al., 2018). CTLA-4 is mainly expressed on activated CD8<sup>+</sup> effector T Cells and negates the early-stage T-cell activation and cell cycle progression. CTLA-4 is homologous to CD28, and it possesses the same ligands, including B7-1 (CD80) and B7-2 (CD86), thus blocking the activation and proliferation of antigen-activated T cells secondary to CD28/B7 inter-action (Shen and Ren, 2017). The inhibition of CTLA-4 signaling is extremely promising to recover the suppression of T-cell activation, proliferation, and infiltration into tumors. This process can result in enhanced anti-tumor immunity and tumor cell eradication (Bielinska et al., 2014). FDA in 2011 approved anti-CTLA-4 antibodies ipilimumab for the treatment of metastatic melanoma; this marked the beginning of a new era for cancer immunotherapy (Barbee et al., 2015). Ipilimumab and tremelimumab work by blocking the binding of CTLA-4 to its ligand

<https://doi.org/10.1016/j.compbiolchem.2019.05.009>

Received 19 March 2019; Received in revised form 30 April 2019; Accepted 21 May 2019  
Available online 22 May 2019

1476-9271/© 2019 Published by Elsevier Ltd.



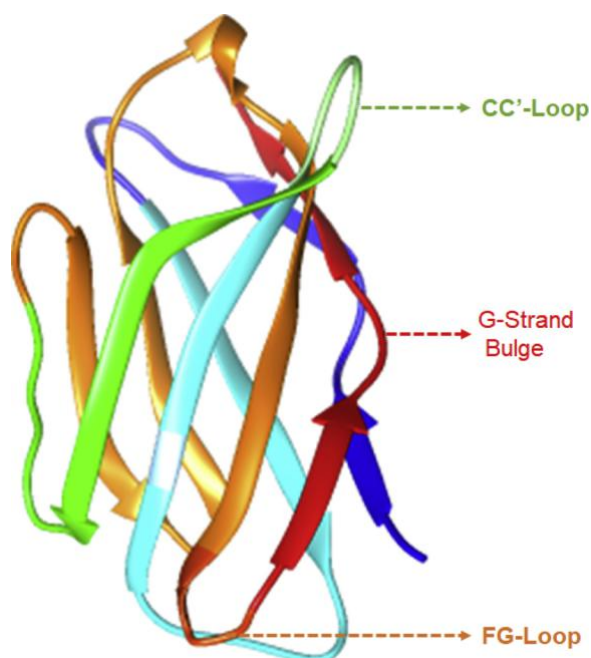


Fig. 1. 3D crystal structure of CTLA-4. The CC'-Loop is highlighted in green, G-strand in red and the FG-loop in orange. (PDB:5TRU) (Ramagopal et al., 2017).

(CD80 AND CD86) (Aspesslagh et al., 2015). Due to the lower production costs, higher stability, improved tumor penetration, amenability for oral administration and elimination of immunogenicity, small-molecular weight inhibitors are expected to be a viable option in place of antibodies as immune check point inhibitors (Zhan et al., 2016). Unfortunately, there are no available small molecule inhibitors that target CTLA-4. Hence, we used drug repositioning and pharmacophore approach to propose drugs that can effectively serve as an inhibitor of CTLA-4.

## 2. Computational method

### 2.1. Binding site similarity identification

The complexity of protein and the dynamics of their active site have made it possible for drugs to modulate multiple targets implicated in diseases. This binding site similarity has been explored in repurposing drugs for effective therapeutic potentials. We used bindingdb database, a publicly accessible database which contains approximately 20,000 experimentally determined binding affinities of protein-ligand complexes, chemical structures, substructure and similarity (Liu et al., 2007). Therefore this drug repositioning methodology which identifies novel target for existing drugs using bioinformatics databases and algorithms as described by Heinrich et al., 2010 was used (Heinrich et al., 2010). This structural approach explores binding site similarities between the target protein and other related/unrelated proteins to predict novel binders. Therefore, we employ a similar principle to identify small molecule compounds/antagonists that are likely to block the immune checkpoint activity of CTLA-4.

### 2.2. Pharmacophore Model Creation using Per Residue Energy Decomposition (PRED) Based Approach

Identification of small molecule CTLA-4 antagonists was followed up by the modeling of novel compounds that could also act in a similar inhibitory capacity towards the target protein.

Contrary to other conventional pharmacophore modelling methods, we used an in house PRED approach to define the pharmacophoric architecture of the ligand-receptor in order to retrieve a more tailored

potential hit (Cele et al., 2016; Kumalo and Soliman, 2016). This pharmacophore model explore both the structural features of the proteins as well as the chemical features of ligands (Cele et al., 2016). To generate a PRED-based pharmacophore model, PRED decomposition was computed from MM/PBSA calculations after 20 ns MD simulations of the (TDSO-CTLA-4) complex. Pharmacophoric features based on the receptor-ligand interaction obtained from this short run MD simulation was selected. Residues Ala25, Ala30, Thr31, Val33, Met55, Gly71 and Asn72 were found to be the highest contributing residues that interact with the ligand (Fig. S2A). These ligand features were set as a query to generate a PRED-based pharmacophore model in ZINCPharma (Koes and Camacho, 2012). Furthermore, the PRED-based pharmacophore model (Fig. S1) was used to screen the ZINC database (Irwin and Shoichet, 2006) for compounds with similar features to obtain the novel hits. The filter was set to query Zinc purchasable compounds of molecular weight  $\leq 500$  for hit screening while rotatable bonds was set at  $\leq 10$ . Lipinski's "rule of five" was also set as cut-off. Hits were sorted by increasing RMSD Scores.

### 2.3. Drug likeliness assessment

An Online software SwissADME (Daina et al., 2017) was used to determine the physicochemical descriptors, pharmacokinetic properties and drug-like nature of the screened compounds. SwissADME uses the "Brain or Intestinal Estimated permeation, (BOILED-Egg)" method which computes the lipophilicity and polarity of small molecules.

### 2.4. System preparation and Molecular dynamic (MD) simulations

The x-ray crystal structure of CTLA-4 in complex with an antibody check point inhibitor Ipilimumab with Protein Data Bank code (5TRU) (Ramagopal et al., 2017) was obtained from RSCB Protein Data Bank (Skalniak et al., 2017). The Ipilimumab was deleted to obtain an un-bound CTLA-4 protein. CTLA-4 preparation as well as visualization was conducted in Chimera (Yang et al., 2012) and Avogadro Visualization (Hanwell et al., 2012) software. 2D-structure of the ligands were setup for auto-optimization using UFF forcefield and steepest descent algorithm on Avogadro 1.2.0 (Hanwell et al., 2012). This was carried out to adjust inter-connective bonds and eliminate clashes. Further preparations of the ligands and CTLA-4 were carried out using the Molegro Molecular Viewer (MMV) software (Windows, M. M. V and X, 2010). The ligands were docked into the protein using automated Autodock Tools docking procedure which entails grid box 'active-site' definition (Trott and Olson, 2010). The docking score was validated by re-docking using Chimera (Pettersen et al., 2004) and HADDOCK (Dominguez et al., 2003). The results are presented in Table S1. The MD simulation was performed using the GPU version of the PMEMD engine provided with AMBER package, FF14SB variant of the AMBER force field was used to describe the protein. ANTECHAMBER was used to generate partial charges for the ligand by utilizing the restrained electrostatic potential (RESP) and the General Amber Force Field (GAFF) procedures. The leap module of AMBER 14 allowed the addition of hydrogen atoms, as well as  $\text{Na}^+$  and  $\text{Cl}^-$  counter ions for neutralization to both the Apo and bound system. An orthorhombic box of TIP3P water molecules at a distance of 9 Å from all the protein atoms was used to solvate all the systems prior to simulation (Jorgensen et al., 1983). An initial minimization of 2000 steps were carried out within an applied restraint potential of 500 kcal mol<sup>-1</sup> Å<sup>-2</sup> for both solutes were performed for 1000 steps using a steepest descent method followed by a 1000 step of conjugate gradients. An additional full minimization of 10,000 steps was further carried out by conjugate gradient algorithm without restraint.

A gradual heating MD simulation from 0 K to 300 K was executed for 50 ps, such that the system maintained a fixed number of atoms and fixed volume, i.e., a canonical ensemble (NVT). The solutes within the system are imposed with a potential harmonic restraint of 10 kcal mol<sup>-1</sup>

1 Å<sup>-2</sup> and collision frequency of 1.0 ps<sup>-1</sup>. Following heating, an equilibration estimating 500 ps of each system was conducted; the operating temperature was kept constant at 300 K. Additional features such as number of atoms and pressure were also kept constant mimicking an isobaric-isothermal ensemble (NPT). The system pressure was maintained at 1 bar using Berendsen barostat.

The total time for the MD simulation was 150 ns. In each simulation the SHAKE algorithm was employed to constrict the bonds of hydrogen atoms. The step size of each simulation was 2 fs and an SPFP precision model was used. The simulations coincided with isobaric-isothermal ensemble (NPT), with randomized seeding, constant pressure of bar maintained by the Berendsen barostat, a pressure-coupling constant of 2 ps, a temperature of 300 K and Langevin thermostat with collision frequency of 1.0ps<sup>-2</sup>.

## 2.5. Post-dynamic analysis

The trajectories were saved and further analysis done using PTRAJ (David, 2012) module, CPPTRAJ (David, 2012) module was used to analyze Radius of Gyration, root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF).

## 2.6. Thermodynamic calculation

The free binding energy of the CTLA-4 active site was analyzed by the Molecular Mechanics/ Poisson-Boltzmann Surface Area (MM/PBSA) (Kollman et al., 2000). Binding free energy calculation is an endpoint energy calculation that provides valuable information about ligand-receptor association. Calculation of binding free energies considered 1500 snapshots from each 10 ns trajectory with the average values of the trajectories being computed. From each snapshot, binding free energy ( $G_{\text{bind}}$ ) was computed from the following:

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

$$E_{\text{gas}} = E_{\text{int}} + E_{\text{vdw}} + E_{\text{ele}} \quad (2)$$

$$G_{\text{sol}} = G_{\text{GB}} + G_{\text{SA}} \quad (3)$$

$$G_{\text{SA}} = \gamma \text{SASA} \quad (4)$$

Where  $E_{\text{gas}}$ , the gas-phase energy, was calculated using the FF99SB force field;  $E_{\text{int}}$  is the internal energy;  $E_{\text{ele}}$  and  $E_{\text{vdw}}$  are the Coulomb and van der Waals energies, respectively.  $G_{\text{sol}}$  is the solvation free energy.  $G_{\text{GB}}$  is the polar solvation contribution.  $G_{\text{SA}}$  is the non-polar solvation contribution and was estimated by the solvent accessible surface area (SASA) determined using a water probe radius of 1.4 Å. The surface tension constant  $\gamma$  was set to 0.0072 kcal/(mol Å<sup>2</sup>).

## 3. Results and discussion

### 3.1. Binding site prediction and validation

The binding site of CTLA-4 was predicted using site map (Halgren, 2009), and cross-validated with the aid of MetaPocket site identification tool (Huang, 2009). A binding pocket was predicted, and constitutive residues are shown in Table 1. This binding pocket was chosen based on the drugability (Dscore) score, which measures the ability of a protein to tightly bind small molecules.

Table 1

Constituent residues of predicted CTLA-4 binding site.

Binding site identification (residues)		
Biological target	Site Map	MetaPocket
CTLA-4	Tyr24, Ala25, Ser26, Gly28, Lys29, Ala30, Thr31, Val33, Tyr53, Met55, Gly56, Ser69, Gly71, Asn72, Leu 95	Tyr24, Val33, Tyr53, Met55, Gly56, Gly73, Asn74, Ala30, Ser71, Ser72, Thr31, Gln75, Lys29, Gly28, Leu97, Ser26, Ala25, Met54 Asn57

According to the methods of Heinrich 2010, CTLA-4 was queried on a binding database (www.bindingdb.org) to obtain proteins similar in binding site to CTLA-4 (Liu et al., 2007). In this regard, a protein with an identical binding site to CTLA-4 was identified, which had a small molecule; 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones (TDSO) bound to it. Moreover, in relation to CTLA-4, binding site residues in Kallikrein 7 include Asn189, Gly193, Ser146, Val149, Ala183, and Thr96 (Fig. 2) indicative of a unique similarity.

### 3.1.1. Pharmacophore model creation

The similarity between CTLA-4 and TDSO-bound Kallikrein 7 (PDB: 5Y9L) binding sites informed its (TDSO) selection as the lead compound used in modeling novel antagonists of CTLA-4. To obtain complementary structural architecture of the steric and electronic features that were important to ensure supramolecular interaction of TDSO with CTLA-4, an intermediate 20 ns MD simulation was carried out on the TDSO-CTLA-4 complex. Afterwards, a pharmacophore model or theory was obtained as shown in Fig. S1, which highlights the fundamental pharmacophoric features that served as a starting point for virtual screening to identify potential CTLA-4 antagonists.

### 3.1.2. Generation of ligands

The pharmacophoric moieties that exhibited the most significant interaction with complementary residues of the binding pocket were employed to screen for novel compounds with similar structural and chemical features in Zinc database (contains 4.6 million small molecules). In addition, the Lipinski's Rule of Five (Lipinski et al., 1997) was used as a search criteria to streamline the resulting compounds to drug-like substances, which were then screened virtually using Autodock Vina's with an exhaustiveness set at 8.0. The grid box coordinates X, Y, and Z are centered on 16.62, 18.77, and 17.57 respectively while generated results were saved in the pdbqt format. This pharmacophore-based query led to the generation of 9 structurally similar compounds that belong to the ZINC Purchasable Subset. Further selection was carried out based on comparative binding energies, which led to the identification and retrieval of two compounds with the highest negative binding energies (ZINC04515726 and ZINC08985213) in relation to others (Fig. 3).

### 3.2. Drug likeliness assessment of TDSO, ZINC04515726 and ZINC08985213

#### 3.2.1. ADME result

To access the physicochemical descriptors, pharmacokinetic properties and drug likeliness of TDSO, ZINC04515726 and ZINC08985213, SwissADME (Daina et al., 2017) was used. In the design of therapeutics, the pharmacokinetic properties, safety, potency and selectivity are always considered. Lipophilicity (LogP) is a very essential feature to be considered in the interaction between a chemical compound and its biological target. The permeability of a compound across the lipid bilayer is a function of the value of its lipophilicity, hepatic clearance or solubility (Sciences, H., 2009). A chemical compound having a LogP value between 2 and 3 shows a highly favorable potential of achieving permeability and first pass clearance (Muheem et al., 2016). From Table 2 below, it is seen that ZINC04515726, ZINC08985213 and TDSO have LogP values of 2.43, 2.38 and 2.53 respectively, indicating that these three compounds have the potential to achieve membrane

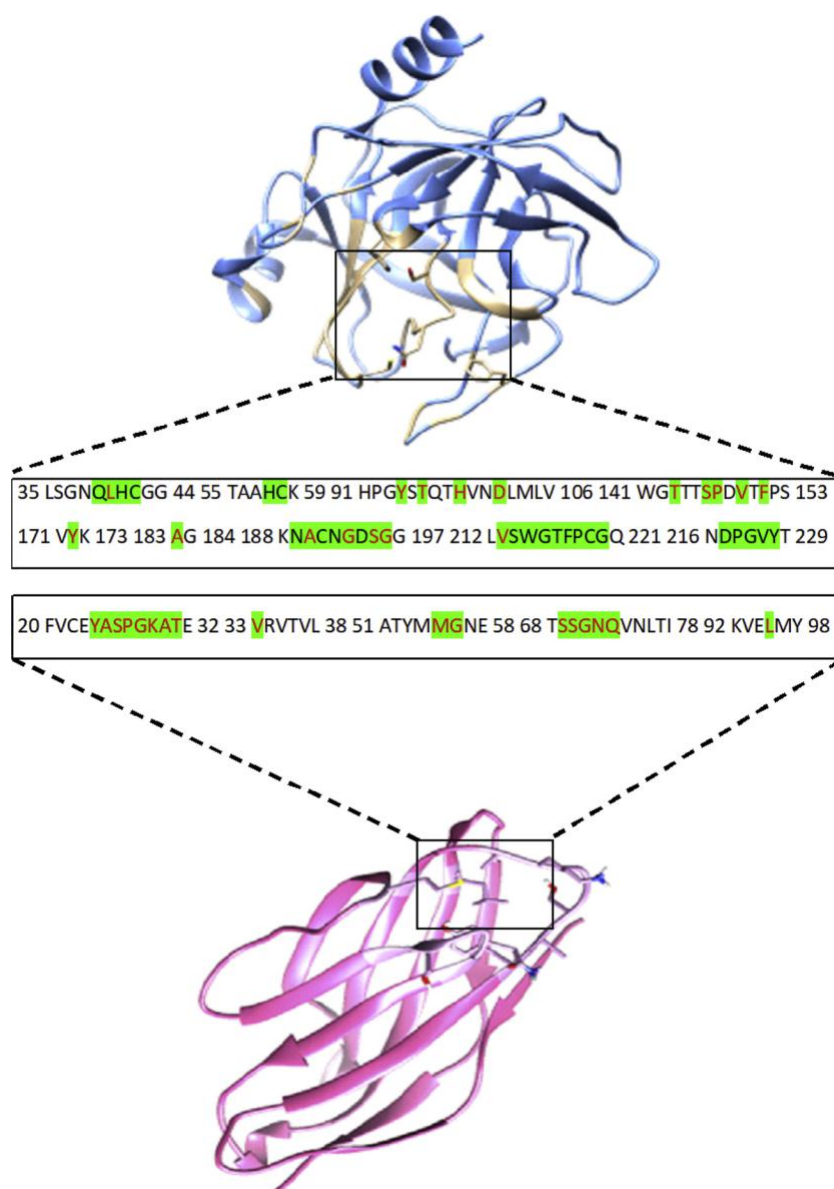


Fig. 2. Graphical representation of kallikrein and CTLA-4 binding site similarity between Kallikrein 7 and CTLA-4. Similar binding site residues are highlighted in red.

permeability and first pass clearance. The gastrointestinal absorption of the compounds is high indicating that they can easily be absorbed by the GIT, which is favorable for their respective biological transport towards CTLA-4. Likewise, their inability to pass through the blood brain barrier (BBB) could indicate their suitability for targeting cancers

cells not found in the central nervous system (CNS).

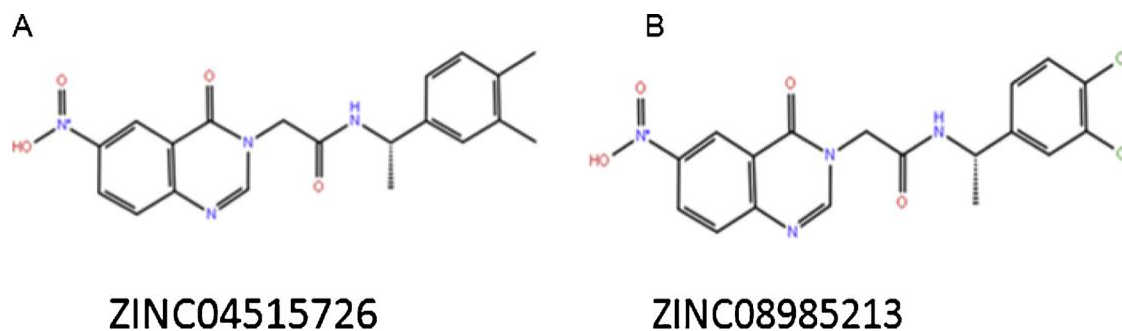


Fig. 3. 2D structures of ZINC04515726 and ZINC08985213.

Table 2

Comparative analyses of the pharmacokinetic profile of TDSO, ZINC04515726, and ZINC08985213.


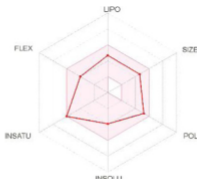

	TDSO	ZINC04515726	ZINC08985213
Molecular Formula	C <sub>24</sub> H <sub>28</sub> OCIN <sub>5</sub> O <sub>5</sub>	C <sub>2</sub> OH <sub>2</sub> ON <sub>4</sub> O <sub>4</sub>	C <sub>18</sub> H <sub>14</sub> N <sub>12</sub> O <sub>4</sub>
Molecular Weight (g/mol)	501.96	380.40	421.23
Lipophilicity (LogP)	2.53	2.43	2.38
Water Soluble	Soluble	Moderately Soluble	Poorly Soluble
GI <sub>T</sub> absorption	High	High	High
BBB Permeability	Not Permeable	Not Permeable	Not Permeable
Bioavailability Score	0.55	0.55	0.55
Synthetic accessibility	4.24	3.30	3.18
Druglikeness (Lipinski)	Yes	Yes	Yes
"Bioled egg" structure			

Table 3

MM/PBSA binding free energy profiles of TDSO, ZINC04515726, and ZINC08985213 to CTLA-4.

	E <sub>ele</sub> (kcal/mol)	E <sub>vdw</sub> (kcal/mol)	E <sub>gas</sub> (kcal/mol)	G <sub>sol</sub> (kcal/mol)	G <sub>bind</sub> (kcal/mol)
TDSO	-6.35 ± 3.41	-36.01 ± 3.41	-42.36 ± 7.88	13.86 ± 6.80	-28.51 ± 3.57
ZINC04515726	-5.48 ± 5.51	-33.28 ± 4.70	-38.76 ± 6.02	15.00 ± 5.12	-23.76 ± 4.66
ZINC08985213	-8.41 ± 7.93	-26.18 ± 5.46	-34.59 ± 10.04	13.56 ± 7.73	-21.03 ± 5.43

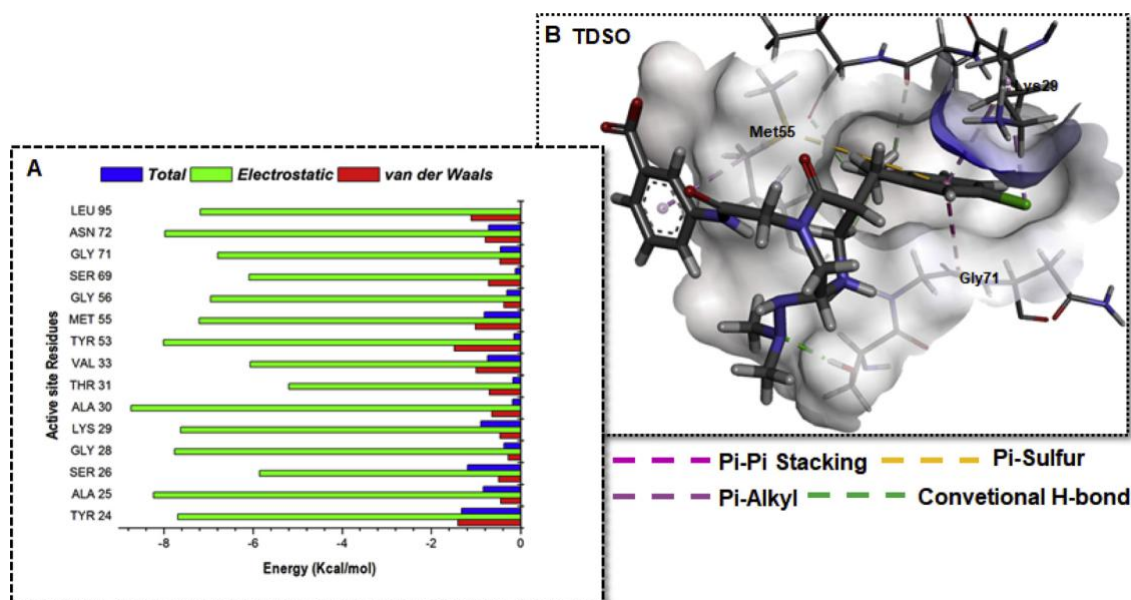
E<sub>ele</sub> = electrostatic energy; E<sub>vdw</sub> = van der Waals energy; G<sub>bind</sub> = total binding free energy;G<sub>sol</sub> = solvation free energy E<sub>gas</sub> = gas phase free energy.

Fig. 4. Individual energy contributions of crucial binding site residues to high-affinity binding and stabilization of TDSO. [A] Per-residue decomposition plot showing energy contributions of interacting residues [B] Molecular interactions between essential residues and reactive substituent on TDSO.

### 3.3. Post-molecular dynamics simulation analysis

#### 3.3.1. Differential binding of TDSO, ZINC04515726 and ZINC08985213 to CTLA-4

The MM/PBSA is used to calculate the free energy difference between two different solvated conformations of the same molecule (Genheden and Ryde, 2015). To determine the effect of inhibitor binding to CTLA-4, post-dynamic calculation of MM/PBSA was carried out. These binding affinities provide us with the information of each amino acid residue toward the total binding affinity. As estimated, the three compounds exhibited favorable binding as evidenced by their

high negative energies. From Table 3, TDSO had a G value of -28.51 kcal/mol while ZINC04515726 and ZINC08985213 have G values of -23.76 kcal/mol and -21.03 kcal/mol respectively, which could suggest strong and high affinity binding.

Furthermore, we investigated key binding site residues and their respective roles in the differential binding of TDSO, ZINC04515726 and ZINC08985213 towards CTLA-4. This was carried out by decomposing the G binding free energy into electrostatic, van der Waals and total energies using an MMPB/SA-integrated per-residue energy decomposition method. As shown in Figs. 5A–7 A, Tyr24, Ala25, Gly28, Ala30, Tyr53 and Asn72 play crucial roles in the binding of the respective

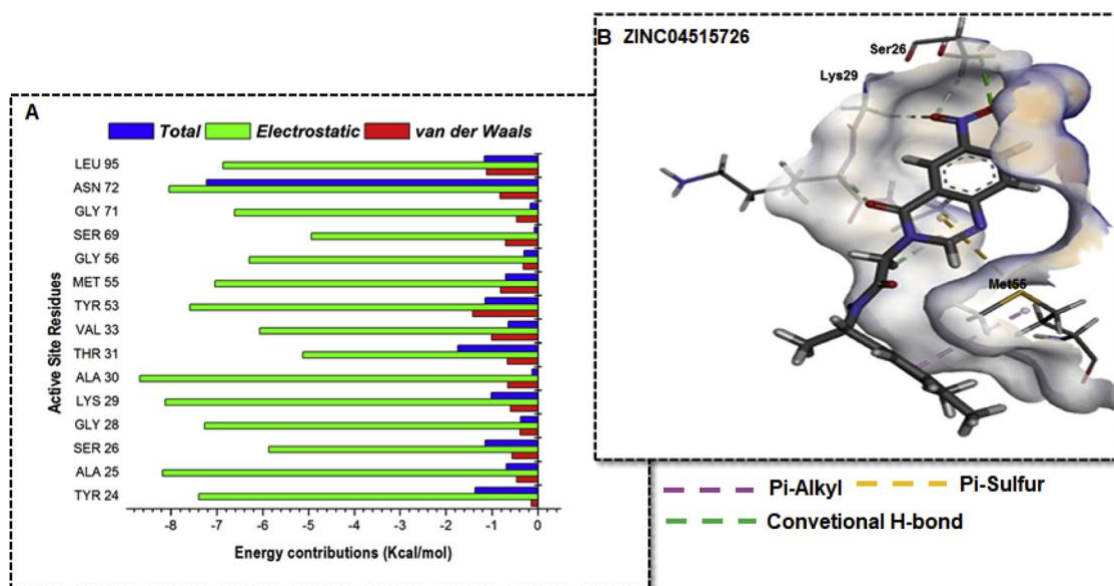


Fig. 5. Individual energy contributions of crucial binding site residues to high-affinity binding and stabilization of ZINC04515726. [A] Per-residue decomposition plot showing energy contributions of interacting residues [B] Molecular interactions between essential residues and reactive substituent on ZINC04515726.

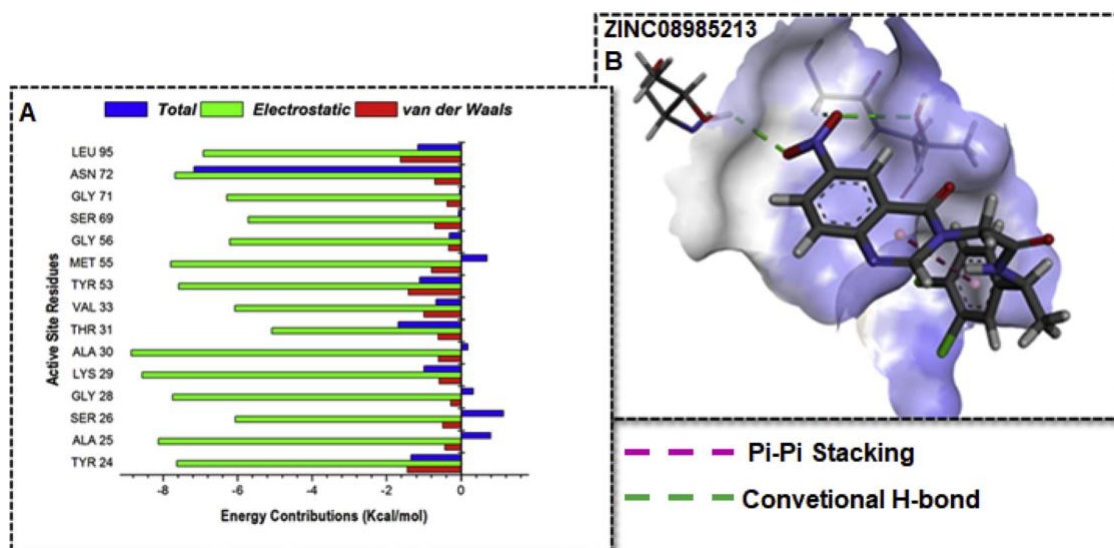


Fig. 6. Individual energy contributions of crucial binding site residues to high-affinity binding and stabilization of ZINC08985213. [A] Per-residue decomposition plot showing energy contributions of interacting residues [B] Molecular interactions between essential residues and reactive substituent on ZINC08985213.

compounds as evidenced by considerably high electrostatic energy contributions. Tyr24, Ala25, Gly28, Ala30, Tyr53 and Asn72 had electrostatic contributions of  $-7.63$ ,  $-8.23$ ,  $-7.76$ ,  $-8.76$ ,  $-7.99$  and  $-7.92$  kcal/mol respectively towards TDSO, while in ZINC04515726, energy contributions of these residues were  $-7.4$ ,  $-8.18$ ,  $-7.27$ ,  $-8.67$ ,  $-7.6$ , and  $-80.4$  kcal/mol respectively. These residues also had  $-7.63$ ,  $-8.13$ ,  $-7.75$ ,  $-8.85$ ,  $-7.58$  and  $-7.68$  kcal/mol electrostatic energy contributions towards ZINC08985213 binding to CTLA-4. Moreover, in the three systems, there were appreciable van der Waals energy contributed by Tyr24, Tyr53, and Leu95, TDSO had  $-1.34$ ,  $-1.41$  and  $-1.1$  kcal/mol contributed respectively. ZINC08985213 had  $-1.43$ ,  $-1.41$  and  $-1.06$  kcal/mol contribution, while Tyr53, and Leu95 contributed mostly to the van der Waals energy in the binding of ZINC04515726 to CTLA-4 with energy contributions of  $1.41$  and  $1.11$  kcal/mol respectively.

The nature and types of the occurring interactions were further visualized with the aid of Discovery studio 2016 Client to gain additional

insights into the mechanistic interactions of the respective ligands with key active site residues with respect to high-affinity binding and stability. As shown in Figs. 4B–6 B, strong hydrogen bonds (NHeO, CHeO and OHeO) as well as weak pi-pi stacked, pi-alkyl and pi-sulfur interactions were observed, which altogether contributed to the binding of the respective ligands to CTLA-4. These high-affinity interactions could also account for the high electrostatic energies observed on the per residue decomposition plot which altogether could possibly underlie the high-affinity inhibitory activity of these compounds towards CTLA-

Although the mechanisms of charge transfer or complementary bond formation were not investigated in this study, we propose that the use of other quantitative techniques such as QM or QM/MM in future studies would provide more insights.

### 3.3.2. Structural elucidation in the binding of TDSO, ZINC04515726, and ZINC08985213 to CTLA-4

To understand the structural events associated with the inhibitory activity of TDSO, ZINC04515726, and ZINC08985213 towards CTLA-4,

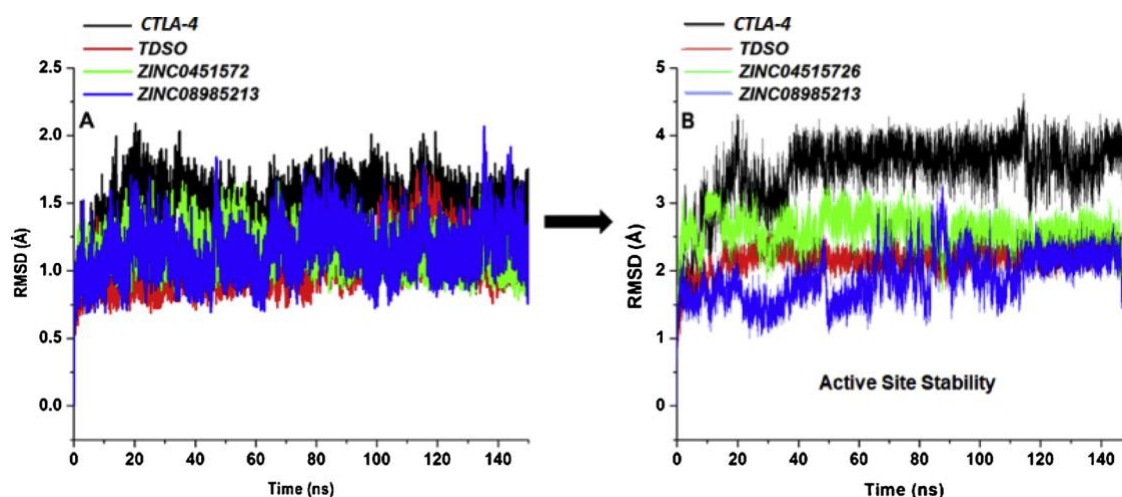


Fig. 7. Conformational analysis plot showing stability and atomistic motions among unbound (black), TDSO (red), ZINC0451572 (green) and ZINC08985213 (blue) bound systems. [A] Whole-structure comparative C- $\alpha$  RMSD plot of the four systems.

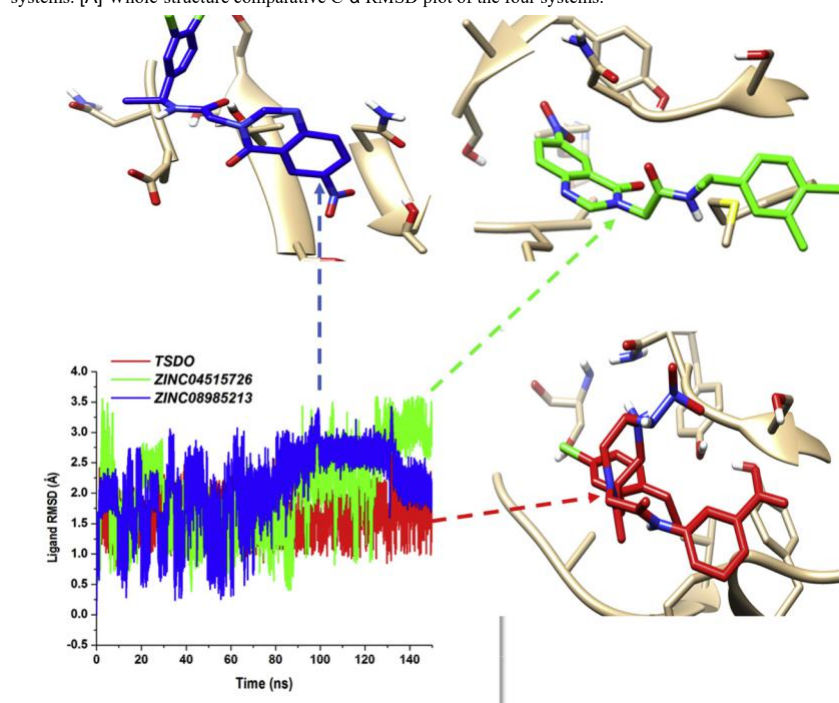


Fig. 8. [A] C- $\alpha$  RMSD plot showing active site stability among unbound (black), TDSO (red), ZINC0451572 (green) and ZINC08985213 (blue) bound systems. [B] Comparative C- $\alpha$  RMSD plot depicting the stability and motion of TDSO (red), ZINC0451572 (green) and ZINC08985213 (blue). [C] Superimposition of TDSO (red), ZINC0451572 (green) and ZINC08985213 (blue).

it was necessary to analyze the structural stability of the simulated systems. As such, the root of mean square deviation (RMSD) of C- $\alpha$  atom across the 150 ns simulation period was monitored. As shown in Fig. 7A, the systems converged early in the simulation period until about 10 ns where the systems showed distinct separations in motions. Whole structural RMSD analyses revealed that unbound CTLA-4 exhibited the highest C- $\alpha$  deviation (RMSD  $\geq 2$  Å) compared to the bound protein-forms which had lower motions (Fig. 7A). Moreover, the binding of the respective ligands elicited notable structural perturbations on the proteins correlative of their conformational instabilities. This could also have possible effects on protein motions thereby impeding their involvements in pathologic interactions. Estimated mean values for unbound CTLA-4 was 1.48 Å while values of 1.05 Å, 1.16 Å and 1.13 Å were estimated for TDSO-, ZINC08985213- and ZINC0451572-bound systems respectively.

We also investigated the impacts of the respective compounds on the architecture of CTLA-4 binding pocket with respect to their antagonistic roles. As presented in Fig. 7B, atomistic motions were higher at

the active site of unbound CTLA-4 while the presence of the ligands lowered C- $\alpha$  motions of constituent active site residues. This could be due to possible ‘pulling-effects’ elicited by complementary ligand-residual interactions that are crucial for high-affinity binding (Fig. 7B).

The motions of the respective ligands at the binding pockets of CTLA-4 was also estimated as presented in Fig. 8. Findings revealed that the compounds showed uniform and stable motions until about 90 ns where the patterns of motions appeared distinct. While TDSO showed a more stable motion at the binding pocket, the identified hits; ZINC0451572 and ZINC08985213, demonstrated similar alterations in motions. Consequently, the smaller sizes of both hits, relative to TDSO could favor their optimal motions at the binding cavity.

Under physiological conditions, proteins like any other molecules undergo fluctuations in their structures (Fuglebakk et al., 2012). Therefore, the dynamics of a protein is equally as important as its structure. The interaction with a protein active site can alter the protein's functions. RMSF is calculated to determine the mobility of individual residues within a protein (Hanwell et al., 2012). To understand

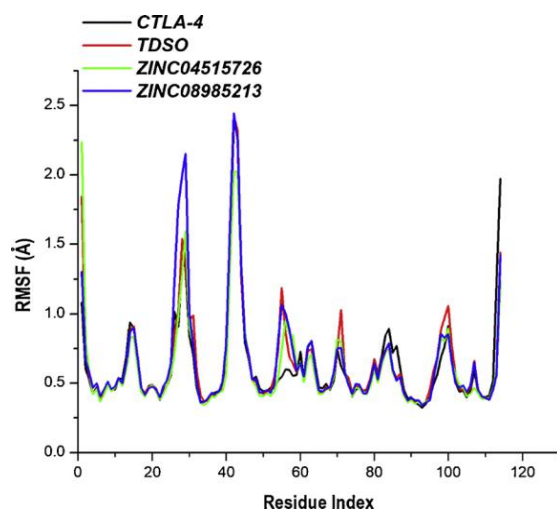


Fig. 9. Structural flexibility induced by the binding of TDSO (red), ZINC04515726 (green) and ZINC08985213 (blue) to unbound (black).

and explore the structural dynamics (fluctuations) that occur upon ligand binding, RMSF was carried out on the systems. Fig. 9 clearly shows high residual fluctuations of the ligand when compared to CTLA-

[14] However, TDSO had a lower flexibility from residues 80–114. This residual fluctuation is possibly due to the impact of the ligands upon binding.

#### 4. Conclusion

Antibodies have always been used in restoring T-Cell exhaustion, but due to the financial cost incurred in the production of these anti-bodies, small molecule inhibitors are thought to be a viable option to monoclonal antibodies. A drug repositioning and pharmacophore modeling approach was used to discover 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones which is an inhibitor of Kallikrein-7. The chemoinformatic analysis of ZINC04515726 and ZINC08985213 showed they are drug-like having obeyed the Lipinski rule of five. MMGBSA revealed that 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones has the highest binding energy value compared to ZINC04515726 and ZINC08985213. The aforementioned results suggest that 1, 3, 6-tri-substituted 1, 4-diazepane-7-ones can be a good anti-CTLA-4.

#### Conflict of interest

The authors declare none.

#### Acknowledgement

The authors acknowledge the College of Health Sciences, UKZN for their financial and infrastructural support and at the same time thank the Centre for High Performance Computing (CHPC, www.chpc.ac.za), Cape Town, for computational resources.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.combiolchem.2019.05.009>

#### References

Chatamra, K., 1991. Cancer and immunology. *Asian Pacific J. Allergy Immunol.* 9, 71–73.

- Mc Granahan, N., et al., 2016. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* 351, 1463–1469.
- Soremekun, O.S., Olotu, F.A., Agoni, C., Mahmoud, E.S., 2019. Recruiting monomer for dimer formation: resolving the antagonistic mechanisms of novel immune checkpoint inhibitors against programmed Death Ligand-1 in cancer immunotherapy. *Mol. Simul.* 0, 1–13.
- Chen, G., Bodogai, M., Tamehiro, N., Shen, C., Dou, J., 2018. Cancer Immunotherapy: theory and application. *J. Immunol. Res.* 2018, 1–2.
- Marin-Acevedo, J.A., et al., 2018. Next generation of immune checkpoint therapy in cancer: new developments and challenges. *J. Hematol. Oncol.* 11, 39.
- Thallinger, C., et al., 2018. Review of cancer treatment with immune checkpoint inhibitors: Current concepts, expectations, limitations and pitfalls. *Wien Klin Wochenschr* 130, 85–91.
- Shen, M., Ren, X., 2017. Highlights on immune checkpoint inhibitors in non-small cell lung cancer. *Tumor Biol.* 39.
- Bielinska, A.U., et al., 2014. Distinct pathways of humoral and cellular immunity induced with the mucosal administration of a nanoemulsion adjuvant. *J. Immunol.* 192, 2722–2733.
- Barbee, M.S., Ogunniyi, A., Horvat, T.Z., Dang, T.O., 2015. Current status and future directions of the immune checkpoint inhibitors ipilimumab, pembrolizumab, and nivolumab in oncology. *Ann. Pharmacother.* 49, 907–937.
- Aspeshlagh, S., Marabelle, A., Soria, J.-C., Armand, J.-P., 2015. Upcoming innovations in lung cancer immunotherapy: focus on immune checkpoint inhibitors. *Chinese Clin. Oncol.* 4, 48.
- Zhan, M.M., et al., 2016. From monoclonal antibodies to small molecules: the development of inhibitors targeting the PD-1/PD-L1 pathway. *Drug Discov. Today* 21, 1027–1036.
- Ramagopal, U.A., et al., 2017. Structural basis for cancer immunotherapy by the first-in-class checkpoint inhibitor ipilimumab. *Proc. Natl. Acad. Sci. U. S. A* 114, E4223–E4232.
- Liu, T., Lin, Y., Wen, X., Jorissen, R.N., Gilson, M.K., 2007. BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res.* 35, 198–201.
- Henrich, S., et al., 2010. Computational approaches to identifying and characterizing protein binding sites for ligand design. *J. Mol. Recognit.* 23, 209–219.
- Cele, F., Ramesh, M., Soliman, M., 2016. Per-residue energy decomposition pharmacophore model to enhance virtual screening in drug discovery: a study for identification of reverse transcriptase inhibitors as potential anti-HIV agents. *Drug Des. Dev. Ther.* 10, 1365–1377.
- Kumalo, H.M., Soliman, M.E., 2016. Per-residue energy footprints-based pharmacophore modeling as an enhanced in silico approach in drug discovery: a case study on the identification of novel  $\beta$ -Secretase1 (BACE1) inhibitors as anti-alzheimer agents. *Cell. Mol. Bioeng.* 9, 175–189.
- Koes, D.R., Camacho, C.J., 2012. ZINCPharmer: pharmacophore search of the ZINC database. *Nucleic Acids Res.* 40, 409–414.
- Irwin, J.J., Shoichet, B., 2006. K. for Virtual Screening 45, 177–182.
- Daina, A., Michielin, O., Zoete, V., 2017. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* 7, 1–13.
- Skalniak, L., et al., 2017. Small-molecule inhibitors of PD-1/PD-L1 immune checkpoint alleviate the PD-L1-induced exhaustion of T-cells. *Oncotarget* 8 (42), 72167–72181. <https://doi.org/10.18632/oncotarget.20050>.
- Yang, Z., et al., 2012. UCSF Chimera, MODELLER, and IMP: an integrated modeling system. *J. Struct. Biol.* 179, 269–278.
- Hanwell, M.D., et al., 2012. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* 4, 1–17.
- Windows, M. V & X, 2010. M. O. S. Molegro Molecular Viewer User Manual.
- Trott, O., Olson, A.J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461.
- Petersen, E.F., et al., 2004. UCSF Chimera - A visualization system for exploratory re-search and analysis. *J. Comput. Chem.* 25, 1605–1612.
- Dominguez, C., Boelens, R., Bonvin, A.M.J.J., 2003. HADDOCK: A Protein-Protein Docking Approach Based on Biochemical or Biophysical Information. pp. 1731–1737.
- Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L., 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935.
- David, A., 2012. Case. AmberTools12 reference manual. Russell J. Bertrand Russell Arch. 535.
- Kollman, P.A., et al., 2000. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* 33, 889–897.
- Halgren, T.A., 2009. Identifying and characterizing binding sites and assessing drug-ability. *J. Chem. Inf. Model.* 377–389.
- Huang, B., 2009. MetaPocket: A Meta Approach to Improve Protein Ligand Binding Site Prediction. *OMICS* 13, 325–330.
- Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J., 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 23, 3–25.
- Sciences, H., 2009. Determination of The Permeability of Biological Membranes to Various Chemical Markers, Including Anti-Hiv Drugs. *Erina Pretorius (Basson) December 2009.*
- Genheden, S., Ryde, U., 2015. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* 10, 449–461.
- Fuglebak, E., Echave, J., Reuter, N., 2012. Measuring and comparing structural fluctuation patterns in large protein datasets. *Bioinformatics* 28, 2431–2440.
- Muheim, A., 2016. A review on the strategies for oral delivery of proteins and peptides and their clinical perspectives. *Saudi Pharm J.* 24 (4), 413–428.



## From genomic variation to protein aberration: Mutational analysis of single nucleotide polymorphism present in *ULBP6* gene and implication in immune response



Opeyemi S. Soremekun, Mahmoud E.S. Soliman\*

Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban, 4001, South Africa

### ARTICLE INFO

#### Keywords:

Single nucleotide polymorphism  
ULBP6  
Bioinformatics  
Molecular dynamic simulation  
Mutation

### ABSTRACT

**Background:** Genetic polymorphisms have been identified as one of the underlying factors in disease pathogenesis and drug resistance since they account for protein dysfunctionality, or in some cases, aberrancy. This explains the high degree of inactivity that characterizes the polymorphic variants of ULBP6 binding protein, which in turn disrupts its primary interaction with human Natural Killer Group 2-member D (NKG2D) and accounts for an impediment to immuno-surveillance. The possible identification of deleterious non-synonymous Single Nucleotide Polymorphisms (nsSNPs) present in the *ULBP6* gene is essential for the development of novel gene therapies to prevent the translation of dysfunctional protein variants.

**Methods/results:** In this study, for the first time, we employed an SNP-informatics approach (SNPs retrieval, pathogenic/mutational analysis, phenotypic analysis, and structural analysis) and molecular dynamics techniques to identify and characterize undesirable SNPs coupled with their impact on ULBP6 structural activities relative to dysfunctionality. V52F was predictively pathogenic amongst SNPs studied. Conformational and dynamic studies revealed that in comparison to wildtype ULBP6 (ULBP6<sup>WT</sup>), pathogenic ULBP6<sup>V52F</sup> demonstrated considerable structural inactivity, which could, in turn, impede biological protein-protein interactions. Moreover, ULBP6<sup>V52F</sup> showed relatively limited motions in the conformational space as deduced from estimations of structural stability, fluctuations, and principal components.

**Conclusion:** This study provides a workable paradigm for investigating pathological nsSNPs using computational platforms which findings present ULBP6<sup>V52F</sup> as a novel and attractive immunotherapeutic target in combatting immune-associated disorders.

### 1. Introduction

The human immune defense is a well-developed system that has evolved in the protection of the body against invading pathogenic assaults and toxins [1]. There are four mechanisms with which this protection is elicited. Firstly, through physical barriers such as skin and mucosal tissues that protect internal tissues from assault [2]. The second line of defense is mediated by antimicrobial enzymes such as lysozyme, anti-microbial peptides such as defensins [3], and the action of the complement system, this line of defense helps to initiate the third line of defense, the innate immune system. Complement binding to a pathogen marks it for destruction by phagocytes, which initiate an inflammatory response [4]. Inflammation recruits monocytes and dendritic cells, which through activation of their Toll-like receptors (TLRs) recruit a range of cytotoxic effector cells such as neutrophils,

lymphocytes and natural killer (NK) cells [5]. These cells elicit the last defense mechanism, which is the recruitment of cells that specifically target pathogens and form immune memory in case of repeat infection [6]. NK cells possess surface receptors that recognize ligands that are up-regulated on the surface of infected or transformed cells [5].

Natural Killer group 2, member D (NKG2D), is one of the receptors that is expressed on the surface of NK cells [7], NKG2D is activated upon binding to its complementary ligands, MHC class-I chain-related proteins (MIC) A, MICB and UL-16-binding proteins (ULBP1-6) [8] (Fig 1). The ULBP6 gene cluster is found on chromosome 6q24.2-q25.3 and comprises of ten loci [9]. ULBP1-6 is one of the recently discovered human NKG2D ligands and is highly polymorphic, with two variants (*ULBP0601* and *ULBP0602*) [8]. This high polymorphic tendency is probably due to evolutionary attempts to maintain viral resistance, and cancer immunoeediting [10] Eagle et al., reported the absence of ULBP6

\* Corresponding author.

E-mail address: [soliman@ukzn.ac.za](mailto:soliman@ukzn.ac.za) (M.E.S. Soliman).

<https://doi.org/10.1016/j.complbiomed.2019.103354>

Received 18 March 2019; Received in revised form 8 July 2019; Accepted 8 July 2019  
0010-4825/ © 2019 Published by Elsevier Ltd.



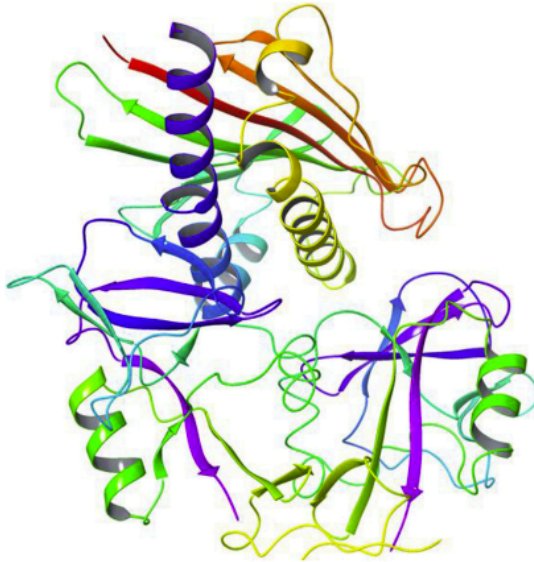


Fig. 1. 3-D crystallography structure of ULBP6 protein in complex with NKG2D (PDB ID: 4S0U [14]).

in primary human fibroblasts, however, upon infection with two different strains of HCMV, ULBP6 transcript was expressed [11]. ULBP6 polymorphisms have also been implicated in diabetic nephropathy [12], alopecia areata (AA) [13] and endometriosis [8].

In the binding of ULBP6 and NKG2D, some amino acids play a crucial role in enhancing the binding potential of the two proteins. Zuo et al., 2017, describes this binding and highlighted K186, E201, K197 of NKG2DA as forming hydrogen bond interactions with residues E154, Y184, and D189 of ULBP6. Likewise, in the NKG2B monomer, residues S151, N207 and K197 formed hydrogen interaction with D99, E103, and R44 of ULBP6 residues [14]. Furthermore, the binding analysis revealed that the binding of NKG2D to ULBP6 is quite similar to ULBP3 [15].

Single nucleotide polymorphisms (SNPs) account for most genetic variation observed in humans [16]. This genetic disparity caused by SNPs occurring at genomic exons alters the translation result leading to structural and functional changes in the mutated protein. However, not all SNPs affect the structure and function of a protein, some are pathogenic while others nonpathogenic [16]. Taking cognizant of the role ULBP6 play in immune response, a study of the impact of its polymorphism will contribute immensely to scientific research. Therefore, in this study we used computational and bioinformatics analysis (Fig. 2) to map out the most deleterious and disease-associated SNP in the *ULBP6* gene, putting into account their structural consequences at the molecular level. We used SIFT [17], PolyPhen [18], PhD-SNP [19], PMut [20] and SNP&GO [21] tools to prioritize the deleterious disease-associated SNPs from the available SNP datasets obtained from the dsSNPs database. To examine the molecular and structural basis of predicted disease-associated SNP, we then carried out molecular dynamics simulation of the native and mutant ULBP6.

## 2. Material and methods

### 2.1. SNPs retrieval

Human ULBP6 (accession ID: Q5VY80) protein sequence was retrieved from the Uniprot database [22]. SNP data used for our computational analysis was obtained from dsSNPs (<http://www.ncbi.nlm.nih.gov/snp/>) [23]. 3-D crystallography structure of ULBP6 was

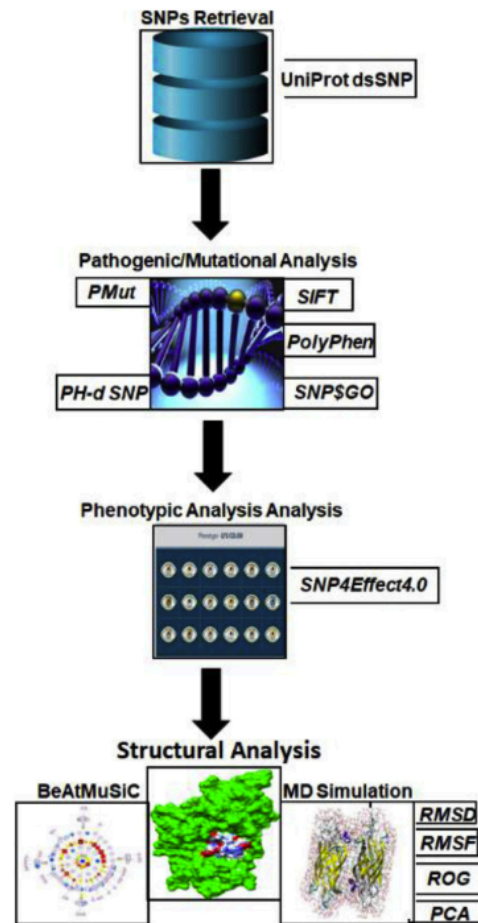


Fig. 2. A schematic computational workflow.

obtained from Protein Data Bank (PDB ID: 4S0U) [24]. To build the mutant ULBP6, a point mutation was induced at position 52 of the wild ULBP6 protein using UCSF chimera [25].

### 2.2. Analysis of the functional impact of coding SNP

Single nucleotide polymorphism occurring in a protein-coding region may impact some structural defect on the 3D structure and consequently leading to some pathological effects. We used SIFT [26] and Polyphen2 [27] to study the functional impact of the nsSNPs occurring in ULBP6. SIFT is a sequence-based homology program that evaluates amino acid substitution [26]. The input SNPs' rs-ID retrieved from the nsSNPs database were submitted to the SIFT server for analysis. SIFT uses the query sequence to execute multiple alignments to predict amino acid substitution present in the sequence and further classify them as either tolerated or deleterious. Prediction is given as tolerance index (TI) score ranging from 0.0 to 1. SNPs with TI score under 0.05 were regarded to be deleterious; those more pronounced than or equal to 0.05 were considered to be tolerated. SIFT has a sensitivity and specificity of 83% and 52% respectively [17]. PolyPhen determines if the amino acid substitution in the sequence occurs in a conserved region and predicts if the substitution has any deleterious impact on the structure of the protein [27]. PolyPhen employs the position-specific independent count (PSIC) score to ascertain the functional and structural impact of the substitution. A PSIC score of 1.0 is regarded as

**Table 1**  
Damaging and Deleterious SNPs predicted by SIFT, PolyPhen, SNP&GO, Pmut and Phd-SNP with their corresponding scores.

SNP ID	Amino Acid Change	SIFT	Score	PolyPhen	Score	SNP&GO	Score	PMut	Score	Phd-SNP	Score
rs1543547	R26G	Tolerated	1	Probably Damaging	0.976	Neutral	9	Neutral	0.02	Neutral	8
rs61730071	T147I	Tolerated	0.862	Benign	0.001	Neutral	9	Neutral	0.15	Neutral	7
rs61748301	M157V	Tolerated	0.112	Probably Damaging	0.636	Neutral	8	Neutral	0.22	Neutral	7
rs7989700	V74I	Tolerated	0.588	Probably Damaging	0.021	Neutral	9	Neutral	0.06	Neutral	5
rs137965307	T38I	Tolerated	0.106	Probably Damaging	0.984	Neutral	7	Neutral	0.15	Neutral	3
rs138504598	L62P	Deleterious	0.001	Probably Damaging	1.000	Disease	7	Neutral	0.34	Disease	0
rs139656748	F151L	Deleterious	0.001	Probably Damaging	0.998	Neutral	2	Neutral	0.08	Neutral	9
rs139983884	P76S	Tolerated	0.351	Benign	0.167	Neutral	8	Neutral	0.08	Neutral	1
rs140013602	I143I	Deleterious	0.037	Benign	0.020	Neutral	8	Neutral	0.10	Neutral	7
rs140239332	E58Ter	Tolerated	0.322	Probably Damaging	1.000	Neutral	8	Neutral	0.08	Neutral	9
rs141330153	D99N	Tolerated	0.062	Possibly Damaging	0.877	Neutral	7	Neutral	0.15	Neutral	6
rs141345751	G192E	Tolerated	1	Benign	0.005	Neutral	6	Neutral	0.01	Neutral	8
rs141551402	M1V	Deleterious	0	Benign	0.000	Neutral	9	Neutral	0.02	Neutral	8
rs141680152	L231F	Tolerated	0.741	Benign	0.001	Neutral	10	Neutral	0.12	Neutral	5
rs145336297	V52F	Deleterious	0.042	Probably Damaging	1.000	Disease	5	Neutral	0.08	Disease	0.73
rs145690445	A123E	Deleterious	0.003	Probably Damaging	0.999	Neutral	5	Neutral	0.23	Neutral	2
rs147008347	I191T	Tolerated	0.595	Benign	0.000	Neutral	8	Neutral	0.07	Neutral	7
rs147287797	M180T	Tolerated	0.348	Benign	0.027	Neutral	7	Neutral	0.02	Neutral	8
rs147068392	S208G	Deleterious	0.008	Probably Damaging	0.945	Neutral	10	Neutral	0.31	Neutral	4
rs147812292	P207T	Tolerated	0.645	Benign	0.086	Neutral	9	Neutral	0.09	Neutral	7
rs148198102	E173K	Deleterious	0.004	Probably Damaging	0.983	Neutral	7	Neutral	0.20	Disease	4
rs151006207	P76H	Tolerated	0.318	Benign	0.045	Neutral	8	Neutral	0.08	Neutral	5
rs151268095	C232Y	Tolerated	1.000	Benign	0.006	Neutral	7	Neutral	0.13	Neutral	6

damaging. The SNPs are classified as benign, possibly damaging and probably damaging [27]. To get an accurate prediction, we used nsSNPs that were jointly predicted by the two servers to be deleterious and damaging. Predicted damaging SNPs were further analyzed using SNP&GO [28], Phd-SNP [29]. SNP&GO retrieves data from protein sequence, evolutionary information, and functions as encoded in the gene ontology terms [28]. It predicts if the SNP is neutral or pathological with the reliability index score [29]. Phd-SNP is a support vector machine-based classifier that employs a supervised learning strategy to categorize the disease-causing mutation that is present in a query dataset. Phd-SNP is a neural network-based program which is trained on a large database of neutral and pathological mutation. It employs three functions including, mutation parameters, solvent accessibility, and residue and sequence properties to calculate the pathogenicity index of an SNP ranging from 0 to 1. Mutation with an index greater than 0.5 is regarded to be pathologically significant [19].

### 2.3. Impact of mutation on protein-protein interaction and molecular dynamics simulation

Particle Mesh Ewald Molecular Dynamics - PMEMD of the AMBER18 software in addition to its integrated modules were used to perform MD simulations [30]. FF14SB force-field was used to define the proteins [31]. ULBP6 modification, renaming, and protonation (histidine) were done using an in-house *pdb4amber* script. Parameter and topology files were generated for ULBP6 using the LEAP module. This was also used for system neutralization. Restraint potential of 500 kcal/mol Å, partial minimization of 2500 steps was carried out and a full minimization of 5000 steps. Gradual system thermalization from 0 to 300 k was also carried out. System equilibration was then carried out for 1000 ps at 300 k without restraints while atmospheric pressure was kept constant at 1 bar using the Berendsen barostat [32]. Afterward, an MD run of 100ns was carried out [33]. At every 1ps, coordinates and trajectories generated were saved. They were further analyzed using CPPTRAJ and PTRAJ [34].

### 2.4. Principal component analysis (PCA)

We used PCA to obtain necessary insights into the 3D conformational and dynamical changes that occurred in the wild and mutant proteins. This method is used to describe the magnitude and direction

of protein motions which are depicted by eigenvalues and eigenvectors respectively [35,36]. Hence, with the aid of the integrated CPPTRAJ module in AMBER18, we computed the first two principal components from the dynamics of the protein C-α atoms.

## 3. Results

### 3.1. Identification of deleterious and damaging SNPs

Identification of pathological genetic variants in proteins is useful in the foundation of genome-level study especially in cancer and some neurodegenerative diseases, this provides better insight for personalized and targeted based drug administration. We retrieved a total of 23 SNPs from the dbSNP database to identify deleterious point mutations in the ULBP6 gene. We simultaneously run the 23 predicted SNPs on SIFT, PolyPhen, SNP&GO, Pmut, and Phd-SNP. The SIFT server was used to determine the tolerance index of all 23 retrieved SNPs by evolutionary conservation analysis. A SIFT score value of < 0.05 was considered to be deleterious. Out of the input polymorphic data from the 23 SNPs, 8 were predicted to be deleterious with a tolerance index score less than or equal to 0.05. To further analyze the structural impact of SNPs, the 23 SNPs retrieved from dbSNP were submitted to the PolyPhen server. A total of 12 were predicted to be damaging having a PSIC score of > 1.5 while 11 were considered as benign. 2 SNPs were predicted to be disease-causing with SNP&GO, 3 SNPs were predicted as diseased with Phd-SNP (Table 1). Taken together, only rs145336297 with V52F was jointly predicted by 4 out of the 5 tools as disease-causing.

### 3.2. Phenotype analysis

SNPeffect4.0 tool [37] was used to detect the phenotypic changes induced by the V52F mutation. The results are predicted in the form of TANGO, LIMBO, WALTZ and ddG scores for the mutant protein.

#### 3.2.1. LIMBO prediction

LIMBO is a chaperone binding site predictor of Hsp70 chaperones, designed from peptidebinding data and structural modeling. The total LIMBO score for ULBP6 is 1409.53 and mutation can increase (dLIMBO > 50), decrease (dLIMBO < -50) or not affect chaperonebinding (dLIMBO between -50 and 50). In this case, dLIMBO

**Table 2**

LIMBO regions in variant and wild type. For each LIMBO region, the start, end, sequence, and score are given.

Number	Start	End	Stretch	Score
Wild Type				
1	54	62	VTMAWKAQ	99.99
2	168	174	FLMGMD	95.34
Mutant				
1	54	62	VTMAWKAQ	99.99
2	168	174	FLMGMD	95.34

equals 0.00 which means that the mutation does not affect the chaperone binding tendency of ULBP6. In Fig. 3A and B, the position of LIMBO stretches in the wild type and variant proteins can be seen, represented by a bar or profile representation. In Table 2 the short stretches are listed for both wild type and mutant. To compare the effect of the mutation to the WT, we also show a difference profile (Fig. 3C), that plots the difference between WT protein and the variant.

### 3.2.2. Waltz and Tango prediction

WALTZ is an algorithm that accurately and specifically predicts amyloid-forming regions in protein sequences (Table 4). The total WALTZ score predicted for ULBP6 is 618.25 and the impact of mutation can increase ( $D_{\text{waltz}} > 50$ ), decrease ( $D_{\text{waltz}} < -50$ ) or not affect amyloid propensity ( $D_{\text{waltz}}$  between  $-50$  and  $50$ ). ULBP6 has a  $D_{\text{waltz}}$  score of 0.01, this suggests that the mutation does not affect the amyloid propensity of ULBP6. We further showed the difference in the WALTZ score between the mutant and native protein (Fig. 4C). TANGO predicts the aggregation-prone region in a protein sequence. The total TANGO score for ULBP6 is 234.67. Mutation can increase ( $d\text{TANGO} > 50$ ), decrease ( $d\text{TANGO} < -50$ ) or not affect aggregation

**Table 3**

TANGO regions in variant and wild type. For each TANGO region, the start, end, sequence, and score are given.

Number	Start	End	Stretch	Score
Wild Type				
No TANGO stretches present				
Mutant				
No TANGO stretches present				

**Table 4**

WALTZ regions in variant and wild type. For each WALTZ region, the start, end, sequence, and score are given.

Number	Start	End	Stretch	Score
Wild Type				
1	31	37	TFLHYD	10.45
2	80	86	QLENYT	22.07
3	110	116	WQFSID	15.09
	118	124	TFLFLD	50.67
Mutant				
1	31	37	TFLHYD	10.45
2	80	86	QLENYT	22.07
3	110	116	WQFSID	15.09
4	118	124	TFLFLD	50.67

propensity ( $d\text{TANGO}$  between  $-50$  and  $50$ ). For V52F mutation,  $d\text{TANGO}$  equals 0.00, this signifies that the mutation does not affect the aggregation tendency of the protein (Table 3). To compare the effect of the mutation to the native protein, we presented the difference in TANGO score in Fig. 4C.

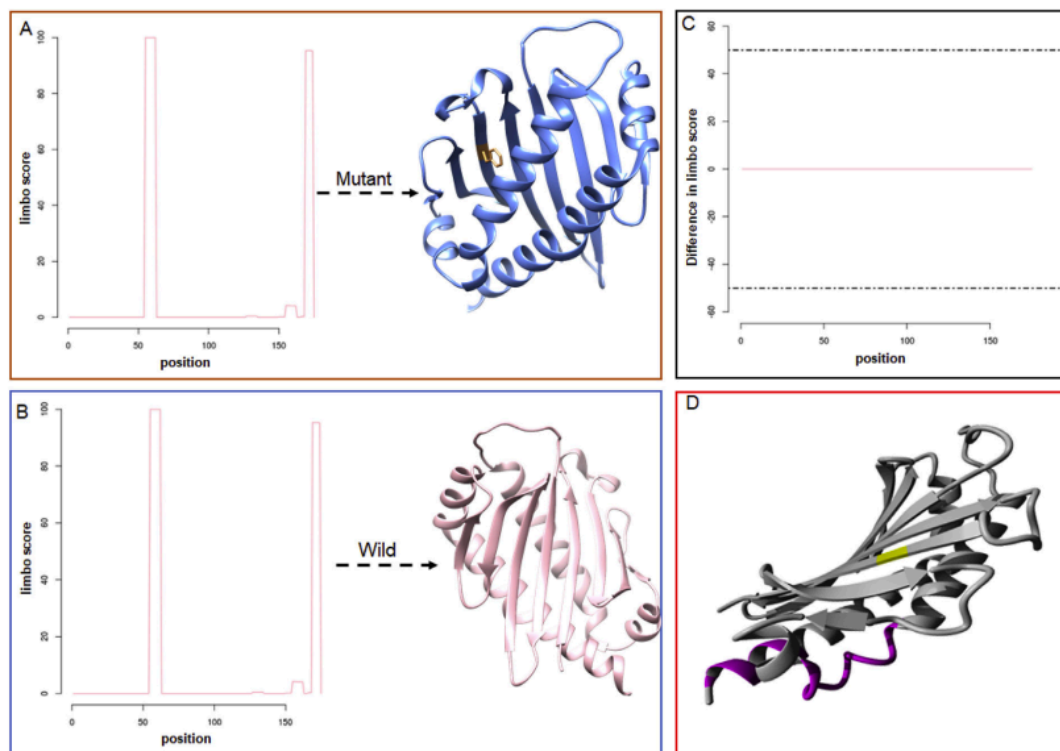


Fig. 3. Profile representation of LIMBO stretches in the mutant (V52F) (A), Wild (B), difference in LIMBO chaperone propensity between the mutant and wild protein (C) and molecular visualization of LIMBO chaperone-binding sites colored in pink (D).

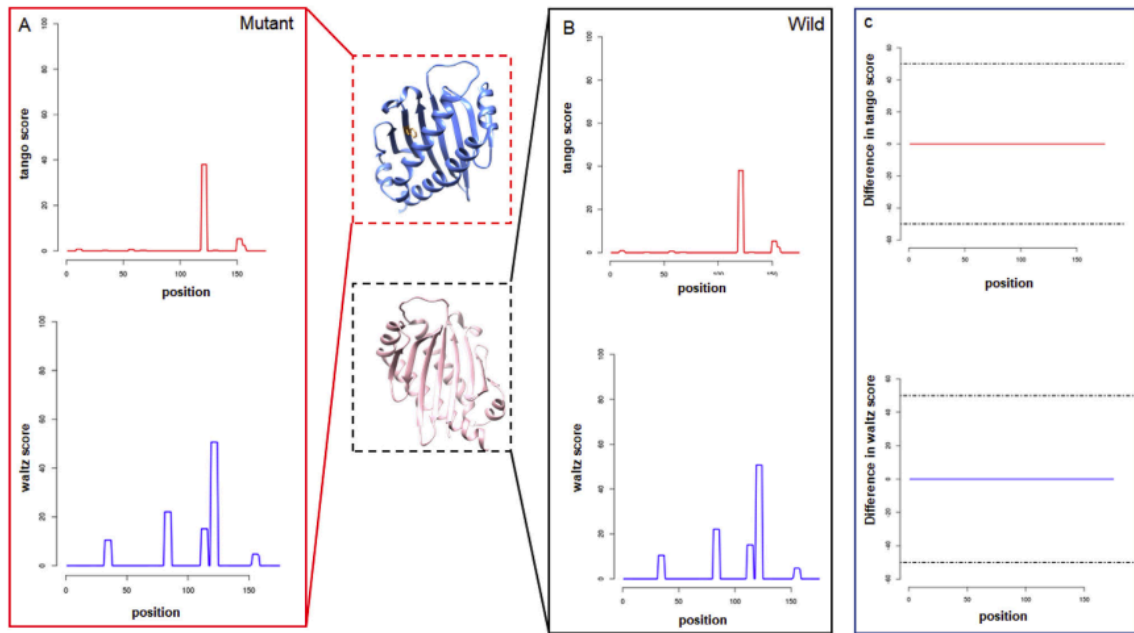


Fig. 4. Profile representation of WALTZ and TANGO stretches present in the mutant (A), WALTZ and TANGO stretch present in the wild type (B). Difference in WALTZ amyloid propensity and difference in TANGO aggregation between WT and variant (C).

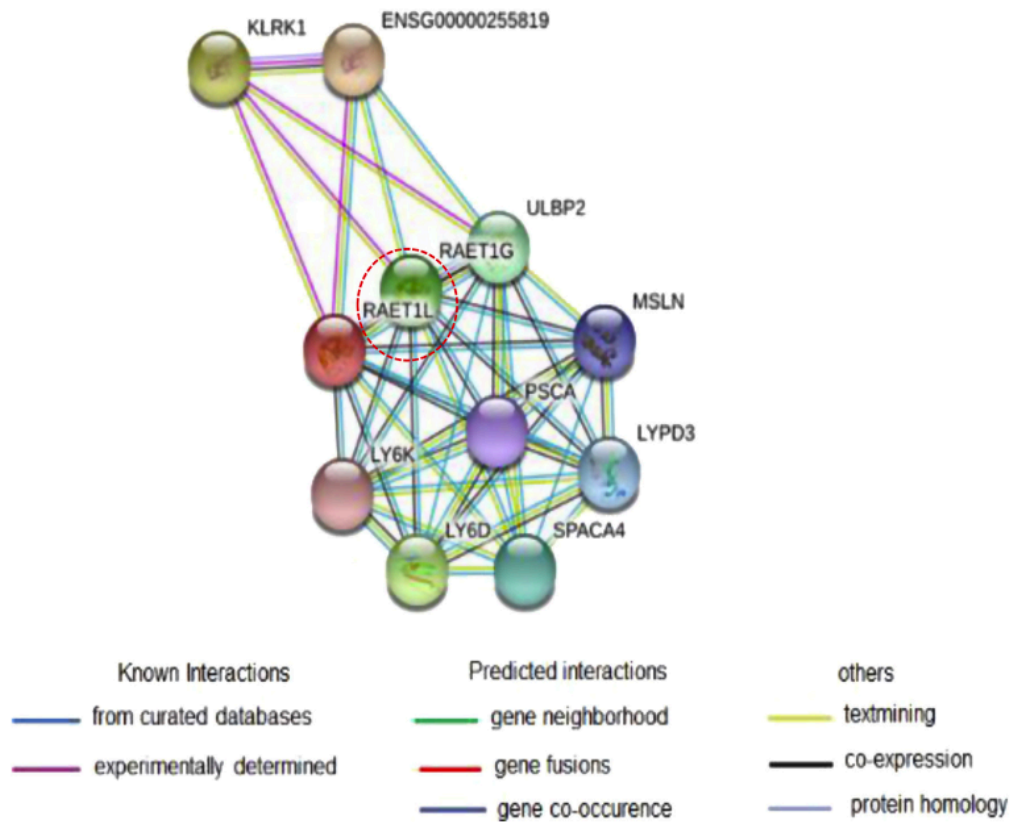


Fig. 5. Functional interaction between ULBP6 and its related genes.

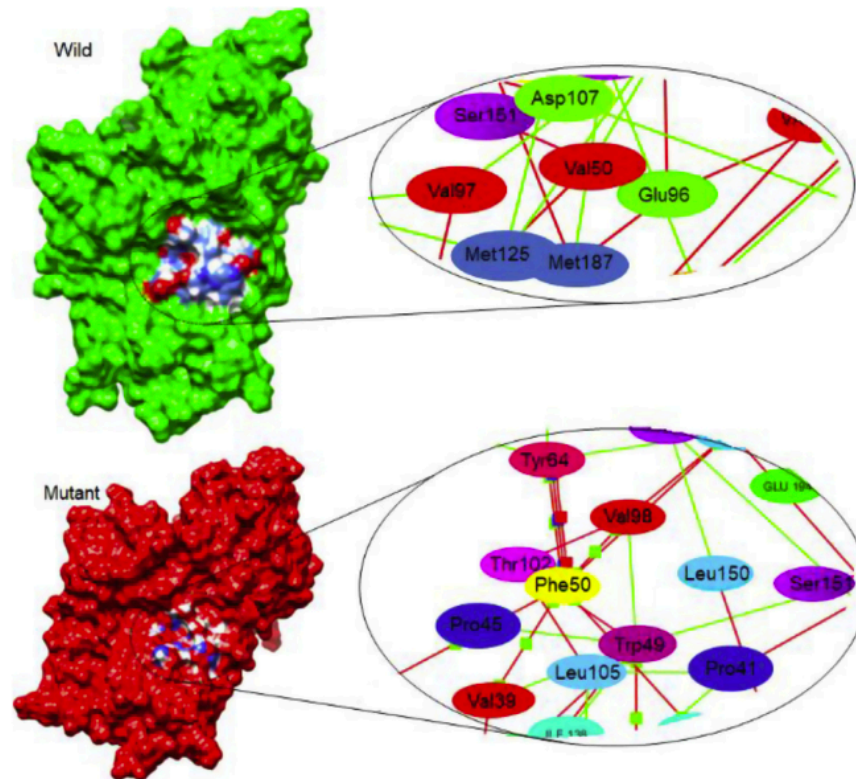


Fig. 6. Residue interaction network analysis of the wild protein (green) and mutant protein (red).

### 3.3. Investigation of ULBP6 Gene's interactions and appearance in networks in STRING database

For better comprehension of a gene's function, knowledge of its specific interacting partners is important. We used STRING database (<http://string-db.org>) to study the interacting genes with ULBP6. STRING database provides a critical assessment and integration of protein-protein interactions, including direct (physical) as well as indirect (functional) associations. From the interaction network, ULBP6 has direct interaction with killer cell lectin-like receptor subfamily K member 1 (KLRK1) (Fig. 5).

### 3.4. Impact of V52F polymorphism on the structural integrity of ULBP6

The formation of a protein complex plays an essential role in the regulation of numerous biological processes [38]. The availability of computational tools to evaluate the impact of mutations on protein-protein binding can therefore, be a valuable tool [38]. Here we used BeAtMuSiC to predict the effect of V52F mutations on NKG2D-ULBP6 binding. BeAtMuSiC is a coarse-grained predictor of changes in binding free energy induced by point mutation. BeAtMuSiC [38] predicted the binding affinity to be 1.79 kcal/mol. The empirical protein design forcefield FoldX was used to calculate the difference in the free energy of the mutation  $\Delta\Delta G$  (delta delta G). The mutation from Val to Phe at position 52 resulted in a  $\Delta\Delta G$  of 8.54 kcal/mol. This implies that the mutation severely reduces protein stability and alters NKG2D-ULBP6 interaction. Residue interaction network was investigated using Cytoscape [39]. The interaction network showed that residue V52 in the wild protein formed hydrogen bonds with Ser151 and Val 97, it also elicited van der Waals interaction with Met125 and Asp107, however in the mutant, phenylalanine formed a hydrogen bond with Tyr64, Val98,

Val39 (Fig. 6). To investigate changes in the secondary structure of the proteins during the simulation period, we used STRIDE [40] to predict the secondary structure change at 10 ns, 50 ns, and 100 ns. As shown in Fig. 7, the major differences observed in the mutant in relative to the wild protein is observed in Pro176, Asn177, and Leu178. These residues were seen to change from a helix configuration in the wild protein to a coil in the mutant protein.

### 3.5. Investigation of structural conformation in ULBP6

To understand the structural impact of V52F mutation, we carried out protein-protein docking using FRODOCK [41] and PRODIGY [44], afterward, we carried out a molecular dynamics simulation of the wild and mutant protein. We calculated RMSD for all the C $\alpha$  atoms from the initial structure using the equation:

$$RMSD = \left( \frac{\sum_N (R_i - R_i^0)^2}{N} \right)^{\frac{1}{2}}$$

N represents the total number of atoms in the complex.  $R_i$  denotes the vector position of the C $\alpha$  atom in the reference conformation of particle  $i$ . This is calculated after aligning the structure to the initial conformation (O) using the least square fitting.

RMSF was estimated using the equation:

$$sRMSF_i = \frac{(RMSF_i - \overline{RMSF})}{\sigma(RMSF)}$$

RMSF $_i$  represents the RMSF of the  $i$ th residue, from which the average RMSF is subtracted. This is then divided by the RMSF's standard deviation to yield the resultant standardized RMSF.

The backbone RMSD values of the wild and Mutant protein during

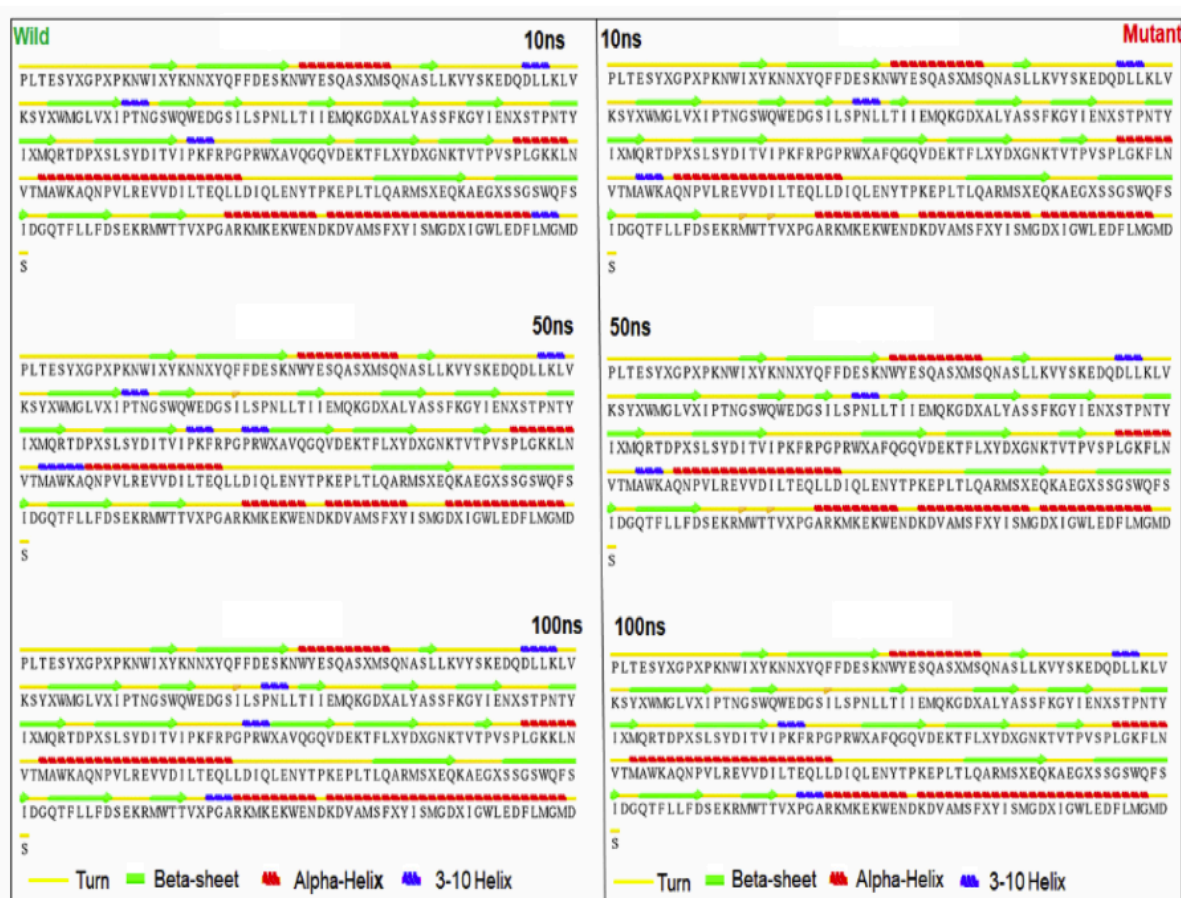


Fig. 7. DSSP analysis of the mutant and wild protein at 10 ns, 50 ns, and 100 ns

the production phase relative to the starting structures were plotted (Fig. 8A) to obtain an estimate of the MD trajectory quality and convergence.

The RMSD result showed that the mutant protein had an average RMSD value of 4.46 Å while the wild protein had an average value of 3.07 Å. This showed that the wild protein had lesser motional movement when compared to the mutant protein. We used RMSF to explore the residual movement of the systems in the course of the simulation period. As seen in the plot, the mutant exhibited higher residual fluctuation with an average RMSF value of 2.33 Å while the wild protein showed a lesser residual fluctuation with an average RMSF of 1.62 Å (Fig. 8A). The PCA showed that the mutant protein occupies more area in the conformational space with higher trace value and changes in direction of the clusters when compared to the wild protein (Fig. 8D). Furthermore, the ROG plot of the two systems showed a dissimilar compactness trend between the mutant and wild ULBP6 (Fig. 8C). The mutant protein had an average ROG value of 22.2 Å while the wild protein exhibited a lower ROG average value of 21.1 Å. To have an insight into the structural impact of V52F mutation on protein-protein interaction, we monitored the distance between the hotspot residues present in the mutant and wild protein in the course of the simulation. In the mutant protein, the distance between S151 and D99 increased from 1.67 Å at 20ns to 8.0 Å at 100 ns N207-E103 increased from 11.2 Å at 20ns to 14.2 Å at 100ns and K197-R44 increased from 13.3 Å at 20ns to 14.7 Å at 100 ns (Fig. 9). In the wild protein, these distances were seen to remain almost constant from the beginning of

the simulation to the end. This suggests that the mutation caused high motional movement of the mutant proteins, this corroborates the high residual fluctuations seen in the RMSF plot.

#### 4. Discussion

Alteration in genetic sequences can be spontaneous, evolutionary adaptation, or chemical induction. This alteration has been classified under two broad categories; mutation and polymorphism [42]. While mutation in most cases leads to the onset of diseases, the impact of polymorphic variant varies [42]. The structure-function of a protein is greatly affected when the SNP occurs in the protein-coding region, consequently leading to a pathogenic state [43]. As a result of the roles they play, attention has been focused on them. The effect of these SNPs can be determined using bioinformatics tools. In this present study, we used PolyPhen, SIFT, SNP&GO, Phd-SNP and PMut to determine the effect of SNPs found in ULBP6. The SNPs data was retrieved from Uniprot dbSNPs. V52F with SNP ID rs145336297 was predicted by the tools to be deleterious and damaging. We performed a protein-protein docking and further ran a 100ns MD simulation to reveal the effect of V52F mutation on NKGD-ULBP6 binding. Protein-Protein docking using PRODIGY [44], showed the wild protein had a binding affinity of  $-8.6\text{kcalmol}^{-1}$  while the mutant had  $6.3\text{kcalmol}^{-1}$ . This result corroborated the prediction of BeAtMuSiC [38]. Protein-protein interaction as revealed by STRING indicated that ULBP6 has direct interaction with killer cell lectin-like receptor subfamily K member 1 (KLRK1), which

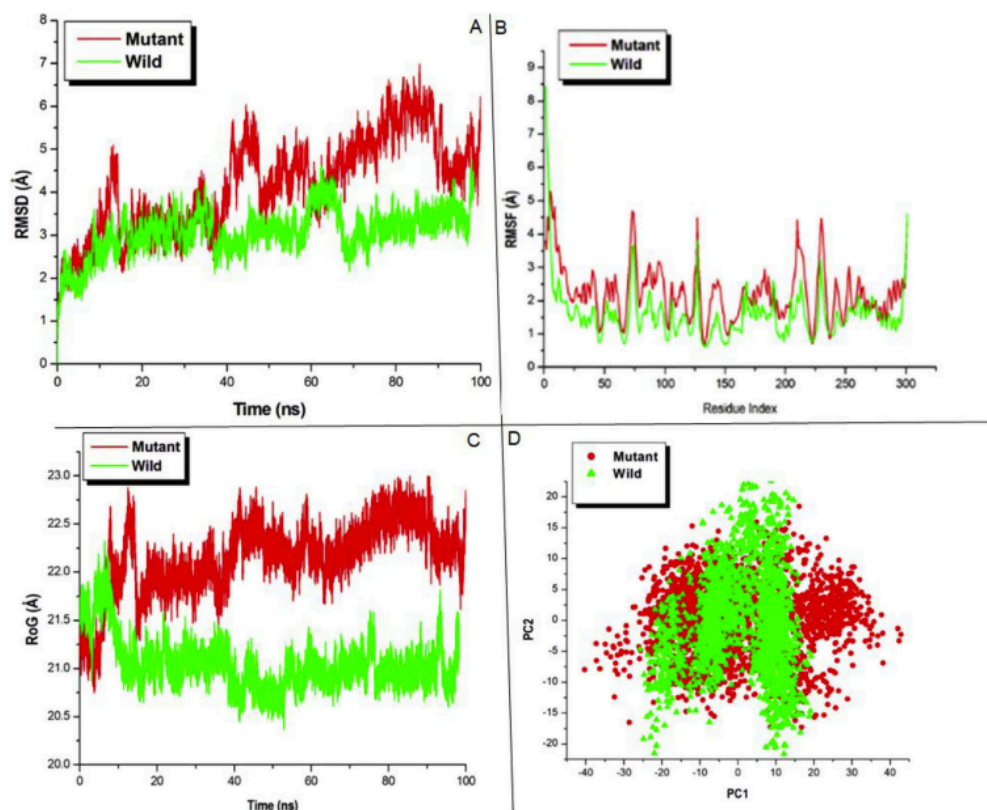


Fig. 8. Backbone RMSDs are depicted as a function of time for the wild and mutant ULBP6, the mutant is represented with a red colour while the wild as green (A). RMSF of the C- $\alpha$  of the mutant (red) and wild (green) (B). Radius of gyration of C- $\alpha$  atoms of native and mutant ULBP6 versus time at 300 k (C). Projection of the motion of the mutant and wild proteins in phase space along the first two principal eigenvectors at 300 k (D).

functions as an activating and costimulatory receptor involved in immune surveillance upon binding to various cellular stress-inducible ligands. The interaction network showed that residue V52 in the wild protein formed hydrogen bonds with Ser151 and Val 97, it also elicited van der Waals interaction with Met125 and Asp107, however in the mutant, phenylalanine formed hydrogen bond with Tyr64, Val98, and Val39. These hydrogen bonds elicited between Phe52 and surrounding residues could provide an explanation of the instability at the protein-hot spot was high in the mutant as compared to the wild. Molecular dynamics simulation is a powerful computational tool that has been used in the study of diseases and macromolecule behavior [45–47]. It is cheaper and fast when compared to experimental studies, it also provides a framework upon which experimental studies can be designed [48–50]. Of notice in the MD simulation studies is the change in distance between S151-D99, N207-E103, and K197-R44 (Fig. 9). Zou et al., 2017, has also highlighted these residues to be paramount to the mechanistic binding between NKG2D and ULBP6 [14]. As the simulation progresses, the distance between these residues increased in the mutant; however, in the wild protein, the distance remained slightly unchanged from the beginning of the simulation to the end. Interaction between NKG2D and ULBP6 facilitates the cytotoxicity of NK cells and acts as a co-stimulatory component of T cells [51], if this interaction is distorted, NKG2D-ULBP6 will be unable to perform its role as immune system regulator. Taken together, our results showed that the instability of NKG2D-ULBP6 as a result of V52F mutation could lead to the dislodgement of ULBP6, resulting in the inability of NKG2D to perform its immune-regulatory role.

## 5. Conclusion

Non-synonymous SNPs (nsSNPs) are responsible for most human pathological conditions and they are mostly found in the exons of protein. In this study, ULBP6 gene was investigated to evaluate the influence of functional SNPs through computation methods. 23 nsSNPs were retrieved from the dbSNP database to identify deleterious point mutation in ULBP6 gene. Taken together, with the arrays of bioinformatics tools employed and computational analysis carried out, we predicted V52F to have a possible impact on NKG2D-ULBP6 binding, although it is not located in the interacting region.

## 6. Conflicts of interest

The authors declare none.

## 7. Funding

This research was not funded.

## 8. Author's contributions

Opeyemi Soremekun contributed to the study design, technical methods and writing of manuscripts. Mahmoud Soliman contributed to the manuscript proof-reading and attested to the study design and methods.

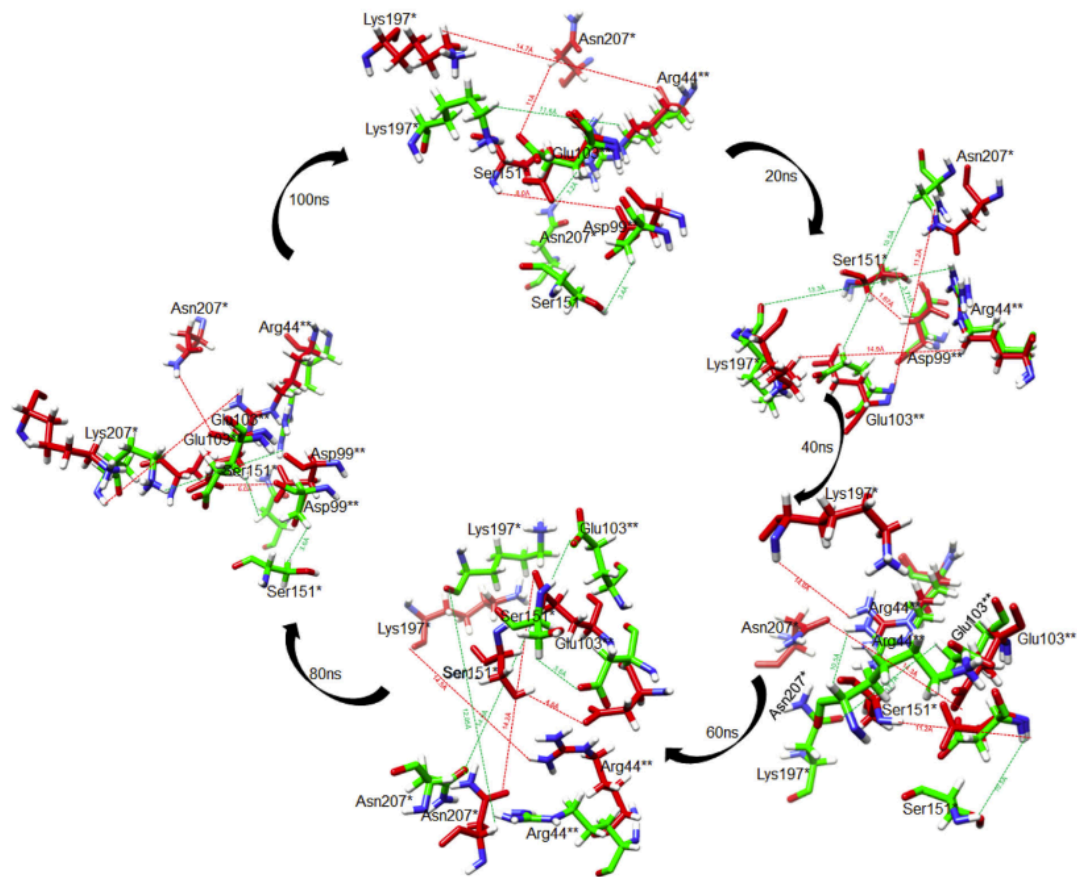


Fig. 9. Schematic diagram depicting the change in distance of the interacting domain residue in the course of the simulation run.

#### Acknowledgement

The authors acknowledge the College of Health Sciences, UKZN for their financial and infrastructural support and at the same time thank the Centre for High Performance Computing (CHPC, [www.chpc.ac.za](http://www.chpc.ac.za)), Cape Town, for computational resources.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.combiomed.2019.103354>.

#### References

- [1] Y. Yang, Cancer immunotherapy: harnessing the immune system to battle cancer, *J. Clin. Invest.* 125 (9) (2015) 3335–3337.
- [2] S. Eyerich, K. Eyerich, C. Traidl-hoffmann, T. Biedermann, Cutaneous barriers and skin immunity: differentiating a connected network, *Trends Immunol.* 39 (4) (2018) 315–327.
- [3] P. Lisowski, N. Strzałkowska, A. Józ, J. Jarczak, E.M. Kos, J. Krzy, et al., Defensins, Natural component of human innate immunity *q* 74 (2013) 1069–1079.
- [4] M.J. Smyth, Multiple approaches to immunotherapy - the new pillar of cancer treatment, *Immunol. Cell Biol.* 95 (4) (2017) 323–324.
- [5] A.K. Purdy, K.S. Campbell, Natural killer cells and cancer: regulation by the killer cell ig-like receptors (KIR), *Cancer Biol. Ther.* 8 (23) (2009) 2211–2220.
- [6] C. Berek, Eosinophils: important players in humoral immunity, *Clin. Exp. Immunol.* 183 (1) (2016) 57–64.
- [7] S. Verheyden, M. Bernier, C. Demanet, Identification of natural killer cell receptor phenotypes associated with leukemia, *Leukemia* 18 (12) (2004) 2002–2007.
- [8] I. González-foruria, P. Santulli, S. Chouzenoux, Soluble ligands for the NKG2D receptor are released during endometriosis and correlate with disease severity, *PLoS One* (2015) 1–15.
- [9] M. Radosavljevic, B. Cuillerier, M.J. Wilson, O. Clément, S. Wicker, et al., A cluster of ten novel MHC class I related genes on human chromosome 6q24.2-q25.3, *Genomics* 79 (1) (2002) 114–123.
- [10] Hunter S. report Understanding the Molecular Role of NKG2D Ligands in Lymphoid Stress Recognition in Cancer ( 2 ) Investigating Neuroblastoma Mediated Immunosuppression A Thesis Submitted to the University of Birmingham for the degree of School of Biosciences. (1).
- [11] R.A. Eagle, J.A. Traherne, J.R. Hair, I. Jafferji, J.Á. Trowsdale, ULBP6/RAET1L is an additional human NKG2D ligand, *Eur. J. Immunol.* (2009) 3207–3216.
- [12] A. Jayne, M. Diane, C. Chris, A.P. Maxwell, D.G. Fogarty, T. Warren, et al., Targeted genome-wide investigation identifies novel SNPs associated with diabetic nephropathy, *HUGO J.* (2010) 77–82 2009.
- [13] L. Petukhova, M. Duvic, M. Hordinsky, D. Norris, V. Price, Y. Shimomura, et al., Genome-wide association study in alopecia areata implicates both innate and adaptive immunity, *Nature* 466 (7302) (2010) 113–117.
- [14] J. Zuo, C.R. Wilcox, F. Mohammed, M. Davey, S. Hunter, K. Khan, et al., A disease-linked ULBP6 polymorphism inhibits NKG2D-mediated target cell killing by enhancing the stability of NKG2D ligand binding, *Sci. Signal.* 890 (2017) 1–15.
- [15] S. Radaev, B. Rostro, A.G. Brooks, M. Colonna, P.D. Sun, Conformational plasticity revealed by the cocrystal structure of NKG2D and its class I MHC-like ligand ULBP3, *Immunity* 15 (6) (2001) 1039–1049.
- [16] N.V. Dokholyan, Predicting the functional consequences of non-synonymous single nucleotide polymorphisms in IL8 gene, *Sci. Rep.* (2017) 1–18 June.
- [17] P.C. Ng, S. Henikoff, SIFT: predicting amino acid changes that affect protein function, *Nucleic Acids Res.* 31 (13) (2003) 3812–3814.
- [18] I. Adzhubei, D.M. Jordan, S.R. Sunyaev, Predicting functional effect of human missense mutations using PolyPhen-2, *Current Protocols in Human Genetics* 2 (2013).
- [19] E. Capriotti, P. Fariselli, PhD-SNPg: A webserver and lightweight tool for scoring single nucleotide variants, *Nucleic Acids Res.* 45 (W1) (2017) W247–W252.
- [20] V. López-Ferrando, A. Gazzo, X. De La Cruz, M. Orozco, J.L. Gelpi, PMut: a web-based tool for the annotation of pathological variants on proteins, 2017 update, *Nucleic Acids Res.* 45 (W1) (2017) W222–W228.
- [21] E. Capriotti, R. Calabrese, P. Fariselli, P. Martelli, R.B. Altman, R. Casadio, WS-SNPs



- &GO: A web server for predicting the deleterious effect of human protein variants using functional annotation, *BMC Genomics* 14 (Suppl 3) (2013) S6.
- [22] A. Bateman, M.J. Martin, C. O'Donovan, M. Magrane, E. Alpi, R. Antunes, et al., UniProt: the universal protein knowledgebase, *Nucleic Acids Res.* 45 (D1) (2017) D158–D169.
- [23] S.T. Sherry, M. Ward, M. Kholodov, J. Baker, L. Phan, E.M. Smigielski, et al., dbSNP: the NCBI database of genetic variation, *Nucleic Acids Res.* 29 (1) (2001) 308–311.
- [24] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, et al., The protein data bank, *Nucleic Acids Res.* 28 (1) (2000) 235–242.
- [25] Z. Yang, K. Lasker, D. Schneidman-Duhovny, B. Webb, C.C. Huang, E.F. Pettersen, et al., UCSF Chimera, modeller, and IMP: an integrated modeling system, *J. Struct. Biol.* 179 (3) (2012) 269–278.
- [26] R. Vaser, S. Adusumalli, S.N. Leng, M. Sikic, P.C. Ng, Protocol update SIFT missense predictions for genomes, *Nat. Protoc.* 11 (1) (2015) 1–9.
- [27] I.A. Adzhubei, S. Schmidt, L. Peshkin, V.E. Ramensky, P. Bork, A.S. Kondrashov, et al., A method and server for predicting damaging missense mutations, *Nat. Methods* 7 (4) (2010) 248–249.
- [28] R. Calabrese, E. Capriotti, P. Fariselli, P.L. Martelli, R.C. À, Functional annotations improve the predictive score of human disease-related mutations in proteins, *Hum. Mutat.* (2009) 1237–1244.
- [29] E. Capriotti, R. Calabrese, R. Casadio, Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information, *Bioinformatics* 22 (22) (2006) 2729–2734.
- [30] D.A. Case, T.E. Cheatham, T. Darden, H. Gohlke, R. Luo, K.M. Merz, et al., The Amber biomolecular simulation programs, *J. Comput. Chem.* 26 (16) (2005) 1668–1688.
- [31] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, *J. Comput. Chem.* 25 (2004) 1157–1174.
- [32] H.J.C. Berendsen, J.P.M. Postma, W.F. Van Gunsteren, A. Dinola, J.R. Haak, H.J.C. Berendsen, et al., Molecular dynamics with coupling to an external bath Molecular dynamics with coupling to an external bath, *J. Chem. Phys.* 3684 (2012) 926–935.
- [33] J.P. Ryckaert, G. Cicotti, H.J.C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J. Comput. Phys.* 23 (3) (1977) 327–341.
- [34] D.R. Roe, T.E. Cheatham III, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, *J. Chem. Theory Comput.* 9 (7) (2013) 3084–3095.
- [35] C.C. David, D.J. Jacobs, Principal component analysis: a method for determining the essential dynamics of proteins, *Methods Mol. Biol.* 1084 (2014) 193–226.
- [36] M. Lawal, F.A. Olotu, M.E.S. Soliman, Across the blood-brain barrier: neurotherapeutic screening and characterization of naringenin as a novel CRMP-2 inhibitor in the treatment of Alzheimer's disease using bioinformatics and computational tools, *Comput. Biol. Med.* 98 (2018) 168–177.
- [37] G. De Baets, J. Van Durme, J. Reumers, S. Maurer-stroh, SNPeff 4.0: on-line prediction of molecular and structural effects of protein-coding variants, 40 (2012) 935–939 November 2011.
- [38] Y. Dehouck, J.M. Kwasigroch, M. Rooman, D. Gilis, BeAtMuSic: prediction of changes in protein – protein binding affinity on mutations, *Nucleic Acids Res.* 41 (2013) 333–339.
- [39] Paul Shannon, Andrew Markiel, Ozier Owen, S. Nitin, Baliga, 1 Jonathan T. Wang, 2 Daniel Ramage 2, Nada Amin 2, Benno Schwikowski, 1, 5 and Trey Ideker2, 3, 4 5, et al. Cytoscape: a Software Environment for Integrated Models of Biomolecular Interaction Networks, *Genome Res.* 13 (22) (2003) 6.
- [40] M. Heinig, D. Frishman, STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins, *Nucleic Acids Res.* 32 (2004) 500–502 WEB SERVER ISS..
- [41] J.I. Garzon, J.R. López-Blanco, C. Pons, J. Kovacs, R. Abagyan, J. Fernandez-Recio, et al., FRODOCK: a new approach for fast rotational protein-protein docking, *Bioinformatics* 25 (19) (2009) 2544–2551.
- [42] S. Alibrandi, Review roles of single-nucleotide polymorphisms in healthy subjects and disease, *Embj* 13 (27) (2018) 118–119.
- [43] A. Kumar, R. Purohit, Computational screening and molecular dynamics simulation of disease associated nsSNPs in CENP-E, *Mutat. Res. Fundam. Mol. Mech. Mutagen.* 738–739 (1) (2012) 28–37.
- [44] L.C. Xue, J.P. Rodrigues, P.L. Kastiris, A.M. Bonvin, A. Vangone, PRODIGY: a web server for predicting the binding affinity of protein-protein complexes, *Bioinformatics* 32 (23) (2016) 3676–3678.
- [45] O.S. Soremekun, F.A. Olotu, C. Agoni, M.E.S. Soliman, Recruiting monomer for dimer formation: resolving the antagonistic mechanisms of novel immune check point inhibitors against Programmed Death Ligand-1 in cancer immunotherapy, *Mol. Simul.* 45 (10) (2019) 777–789.
- [46] O.S. Soremekun, F.A. Olotu, C. Agoni, M.E.S. Soliman, Drug promiscuity: exploring the polypharmacology potential of 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones as an inhibitor of the 'god father' of immune checkpoint, *Comput. Biol. Chem.* 80 (2019).
- [47] S. Khan, I. Bji, R.M. Betz, M.E.S. Soliman, Reversible versus irreversible inhibition modes of ERK2: a comparative analysis for ERK2 protein kinase in cancer therapy, *Future Med. Chem.* 10 (9) (2018) 1003–1015.
- [48] M. Lazarova, Virtual screening – models , methods and software systems, *Int Sci Conf Comput Sci* (2008) 55–60.
- [49] Y. Yang, H. Liu, X. Yao, Understanding the molecular basis of MK2-p38 $\alpha$  signaling complex assembly: insights into protein-protein interaction by molecular dynamics and free energy studies, *Mol. Biosyst.* 8 (8) (2012 Aug) 2106–2118.
- [50] S. Genheden, U. Ryde, The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities, *Expert Opin. Drug Discov.* 10 (5) (2015) 449–461.
- [51] J. Zuo, B.E. Willcox, P. Moss, ULBPs: regulators of human lymphocyte stress recognition, *Oncotargets* 8 (63) (2017) 106157–106158.

REVIEW ARTICLE

# Integrating Bioinformatics Strategies in Cancer Immunotherapy: Current And Future Perspectives

Houda N. Washah, Elliasu Y. Salifu, Opeyemi Soremekun, Ahmed A. Elrashedy, Geraldene Munsamy, Fisayo A. Olotu and Mahmoud E.S. Soliman

<sup>1</sup>*Molecular Bio-computation and Drug Design Lab, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa*

---

ARTICLE HISTORY

---

Received: November 21, 2019  
 Revised: December 21, 2019  
 Accepted: February 26, 2020

DOI:  
 10.2174/1386207323666200427113734

**Abstract:** For the past few decades, the mechanisms of immune responses to cancer have been exploited extensively and significant attention has been given into utilizing the therapeutic potential of the immune system. Cancer immunotherapy has been established as a promising innovative treatment for many forms of cancer. Immunotherapy has gained its prominence through various strategies, including cancer vaccines, monoclonal antibodies (mAbs), adoptive T cell cancer therapy, and immune checkpoint therapy. However, the full potential of cancer immunotherapy is yet to be attained. Recent studies have identified the use of bioinformatics tools as a viable option to help transform the treatment paradigm of several tumors by providing a therapeutically efficient method of cataloging, predicting and selecting immunotherapeutic targets, which are known bottlenecks in the application of immunotherapy. Herein, we gave an insightful overview of the types of immunotherapy techniques used currently, their mechanisms of action, and discussed some bioinformatics tools and databases applied in the immunotherapy of cancer. This review also provides some future perspectives in the use of bioinformatics tools for immunotherapy.

**Keywords:** Cancer immunotherapy, immune system, monoclonal antibody, bioinformatics.

## 1. INTRODUCTION

Cancer is still a major global health concern due to its high incidence and rapid mortality rate for the past decades

\* Cancer is a broad term for a collection of related diseases that involves abnormal cell growth with the tendency of spreading to other parts of the body [1, 2]. The global statistics, reported for the year 2018, identified cancer as the second leading cause of mortality ranking behind cardiovascular diseases, worldwide [3]. The report also suggested that cancer has the potential of becoming the leading cause of death worldwide with a projected increase of 18.1million new cases by the 21<sup>st</sup> century [3]. Cancer is known to develop in various parts of the body, including the lung, breast, prostate area, liver, cervix, and many more. Among the common cancer types, lung cancer is identified to be frequently diagnosed among populations and is reported to be the leading cause of mortality with an estimated 18.4% of all cancer deaths [3]. Female breast

cancer closely followed this with an estimated 11.6% and then prostate cancer, with 7.1% of the total cancer deaths [3].

Current available therapeutic strategies employed in treating cancer include surgery, radioactivity, targeted therapy, immunotherapy, and chemotherapy [4, 5]. These treatment options have provided notable advances towards eradicating primary tumors, however, the prevalence of disease relapse remains on the rise as a result of residual malignant cells [6, 7]. A search for viable therapeutic options that will eliminate resistant malignant tumor cells is warranted. Cancer results is a result of of many abnormal cells, and includes old and ruptured cells as well as the formation of unwanted new cells, which collectively grow and become malignant cells known as tumors [8]. Cancer cells develop when the gene that controls growth and differentiation in normal cells undergoes alterations [9]. The implicated genes are categorized as oncogenes, where they aid cell growth and reproduction and tumor suppressor genes that impede cell division and survival. Targeting these tumor cells has been exploited to be important in developing therapeutic agents to eliminate the cancer epidemic[10]. The immune system has proven to be an attractive route for defeating cancer in recent times due to its ability to induce anti-tumor response [11-14].

---

\*Address correspondence to this author at the Molecular Bio-computation and Drug Design Lab, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa;  
 E-mail: soliman@ukzn.ac.za

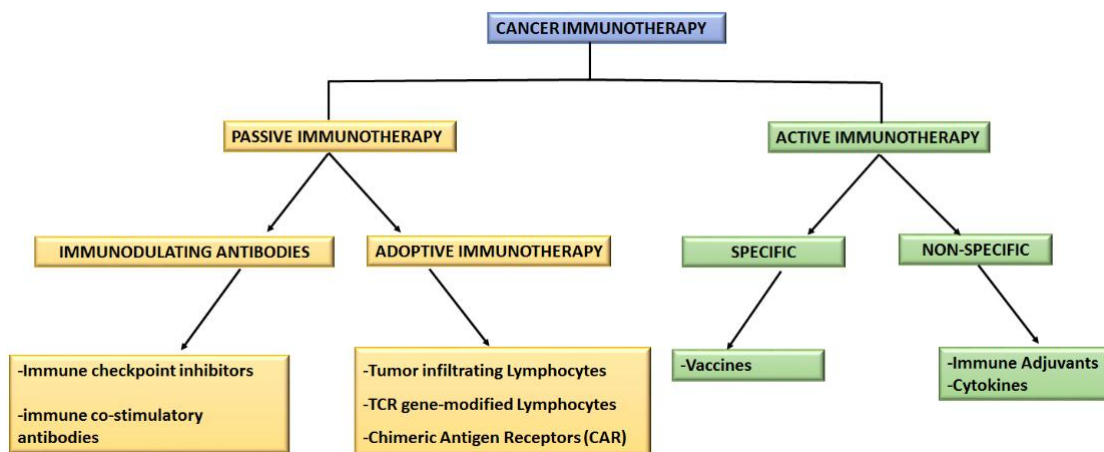


Fig. (1). Classification of cancer immunotherapy.

### 1.1. The Immune System's Mechanism of Action Against Cancer

Clinical studies suggest the existence of a close relationship between the immune systems and the treatment of cancer [7, 9]. The body's immune response comprises the humoral and cellular immunities, both of which are mediated by the B and T lymphocytes, also known as B cells and T cells, respectively [7, 15]. The B cells produce antibodies that neutralize extracellular microbes and toxins, while T cells are responsible for eliminating intracellular microbes by recognizing antigens and activate antigen-presenting cells (APCs) [16, 17]. The immune system plays a very important role in eliminating cancer through innate and adaptive immunity [18]. It has been established that cancer, as a genetic entity, triggers both the innate and adaptive immune response during its evolution [16, 19]. The innate immune system can stimulate the response from B cells and T cells by releasing signals [19]. While the adaptive immune system is known to consist of B cells, CD8<sup>+</sup> cytotoxic T cells, as well as CD4<sup>+</sup> helper T cells [16]. Antigen-presenting cells (APCs) play a crucial role in the immune process as they bridge the innate and adaptive immune systems by identifying unfamiliar antigens and presenting them to T lymphocytes during an immune response. The natural killer (NK) cells of the T lymphocyte produce a pleiotropic cytokine known as interferon-gamma (IFN- $\gamma$ ) [20]. This IFN-

plays an important role in the interface of innate and adaptive immune systems by signaling an increase in expression of Major histocompatibility complex (MHC I) and induce the expression of MHC II molecules on target cells and therefore, increasing their ability to display antigenic peptides to the cytotoxic T lymphocytes to trigger an immediate immune response [21]. In case these peptides are derived from a tumor-associated antigen (TAA), IFN- $\gamma$  could result in increased TAA-specific cytolytic CD8<sup>+</sup> T lymphocytes (CTLs) activation and T cell-mediated tumor killing [20, 21]. The CTLs form the basis of the immune response towards combating cancer [22]. Cancer immunotherapy encompasses a wide scope of techniques that aim to improve the immune response against tumors. This review focuses on the techniques employed in cancer immunotherapy and gives an overview of bioinformatics strategies applied in cancer immunotherapy as well as some future perspectives in cancer immunotherapy.

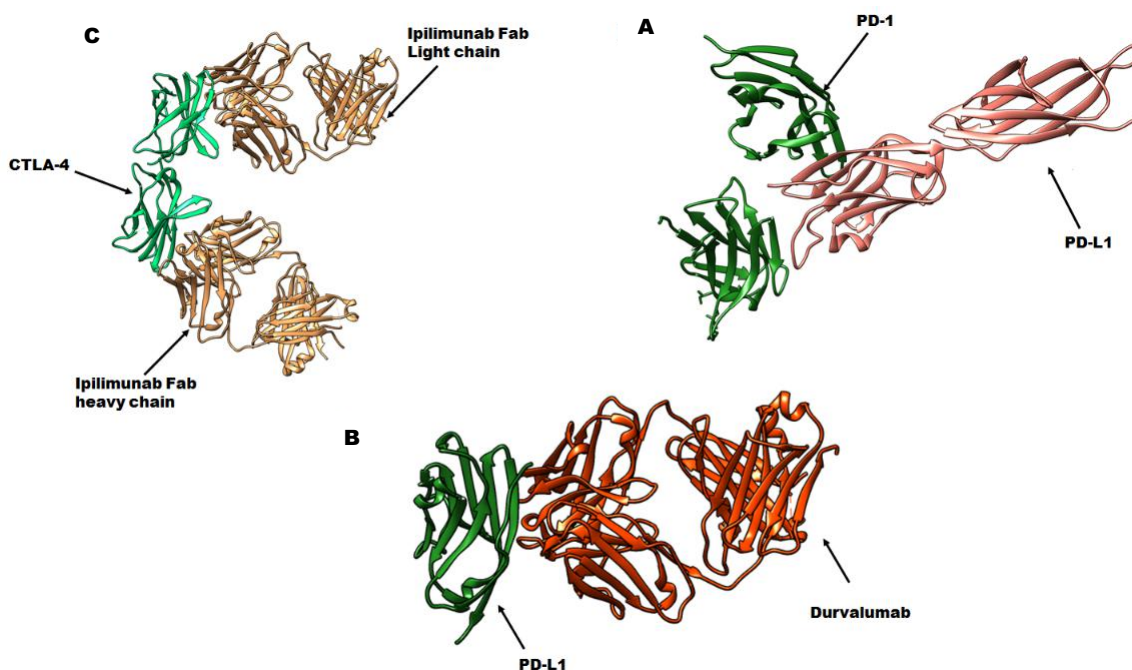
### 1.2. Techniques used in Immunotherapy of Cancer

The last few decades have witnessed groundbreaking research geared towards immuno-oncology studies. This has provided intriguing advances that suggest tumors are recognized by the immune system and the growth of these tumors can be restrained for a longer period through the process of immune surveillance [23]. The early efforts to harness the immune system in cancer control were pioneered by Dr. William B. Coley in the 1890s. Coley worked on the first immunotherapy to treat cancer by using toxins extracted from *Streptococcus erysipelatis* and *Bacillus prodigious* [15]. However, there was a limited clinical efficacy as a result of unrecognized tumor cells evading the immune system. Since then, there has been tireless efforts by researchers to provide insights on how cancer evades the immune system with the aim of developing novel pathways to eliminate the disease. Cancer immunotherapy has been categorized into active and passive immunotherapies, based on their ability to engage the immune system of a host against cancer. As shown in Fig. (1), passive and active immunotherapy has further been divided in various strategies. The cancer immunotherapy strategies discussed in this review include; checkpoint therapy, cancer vaccines, adoptive T cell therapies and monoclonal antibodies.

### 1.3. Immune Checkpoint Inhibitors and Monoclonal Antibodies

#### 1.3.1. Immune Checkpoint Inhibitors

The application of immune checkpoint inhibitors has gained recognition as the most advanced approach to therapeutically explore the antitumor activity in cancer cells [25, 26]. The immune checkpoint strategy has a lot of recent clinical success, especially towards patients with several cancer types. This type of immunotherapeutic strategy focuses on target identification and is mediated by ligand-receptor interactions. The strategy derives its idea from the ability of the immune system to clearly distinguish between normal cells of the body and foreign ones to trigger an immune response [27, 28]. The immune system has checkpoints that enable easy identification of foreign cells to initiate an attack. Some of the suitable immune checkpoint targets include; Programmed death 1 (PD-1), T-cell



**Fig. (2).** 3D structures of **A)** Programmed Death 1(PD-1) in complex with Programmed Death Ligand 1 (PD-L1) PDB code: **3BIK** **B)** PD-L1 in complex with Durvalumab inhibitor (PDB code: **5X8M**) **C)** Cytotoxic antigen – 4 receptor CTLA-4(Light green) in complex with Ipilimumab (brown) PDB code: **5TRU**

immunoglobulin and mucin domain-3 (TIM-3), Lymphocyte-activation gene 3 (LAG-3), and Cytotoxic antigen-4 (CTLA-4) [29, 30]. However, Programmed death [10] (PD-1) and Cytotoxic antigen-4 (CTLA-4) are the most widely studied [27, 31]. Cancer cells usually develop ways to use these checkpoints to evade attacks from the immune system. Therefore the drugs that target these checkpoints are readily developed and hold a lot of promise as a cancer treatment. The checkpoint immune technique has been identified as one of the most remarkable immunomodulating therapies of present-day [32].

PD-1 is a checkpoint protein found in T cells of the immune system. It helps in keeping the immune response in check by acting as an off control to prevent T cells from eliciting an attack on other body cells [33]. This is possible when it binds to another protein known as Programmed Death-Ligand 1 (PD-L1) as shown in Fig 2A, which helps to prevent the T cells from killing other cells. Some cancer cells are known to possess large amounts of PD-L1, an adaptation that helps in evading attacks from the immune system [33-35]. However, this binding of PD-1 to PD-L1 can be blocked by some specific antibodies known as monoclonal antibodies [19]. Several drugs have been developed that target PD-1 and are very promising towards treating cancer. Examples of drugs that target PD-1 include Pembrolizumab (Keytruda) [36], Nivolumab (Opdivo) [37], cemiplimab (Libtayo)[38]. Examples of drugs that target PD-L1 include Atezolizumab (Tecentriq) [39], Avelumab (Bavencio)[40], Durvalumab (Imfinzi), their 3D structure shown in Fig. (2B) [41]. The CTLA-4 checkpoint protein is expressed in the regulatory T cells but is only upregulated in the conventional T cell activation. It also acts as an “off switch” when binding to CD80 and CD86 on the surface of antigen-presenting cells (APCs)[42]. Ipilimumab is a known inhibitor of CTLA-4 protein (shown in Fig. 2C) and acts by keeping this protein

active to enable an attack on cancer cells [12, 43]. A summary of checkpoint inhibitors and the cancer type involved are shown in Table 1. Some notable setbacks in the use of immune checkpoint therapies are: some patients fail to respond to this technique and also the drugs used in this therapy can cause the failure of the immune system to identify cancer cells and may rather attack body cells causing undesirable outcomes [32].

#### 1.4. Monoclonal Antibodies (mAbs)

This cancer treatment strategy has been acknowledged as one of the most successful therapeutic strategies for both hematologic malignancies and solid tumors for the past decade [7, 60, 61]. Monoclonal antibodies are identical antibodies produced in the laboratory from a single clone of immune cells. During the process, a mouse was vaccinated with the target antigen. This stimulated the B cells in the spleen to produce antibodies against the target antigen. The spleen of the mouse was removed and the B cells were isolated and fused with a tumor cell to form hybridoma cells. This was necessary because tumor cells divide more easily and rapidly than B cells. These hybridoma cells then reproduced rapidly to make cloned cells, which all constituted the same antibody. The resulting antibodies were purified and directed to target a specific part of deregulated signals transduction pathways in cancer or hinder immunological processes [61, 62]. MAbs have a high specificity as they easily recognize and bind to a single antigen- binding site making them useful in diagnosing and treating a disease. They can combine with an anticancer drug to accurately locate and target only the cancer cells whilst avoiding the healthy ones. In addition to this, mAbs carry special markers that enable easy detection of where cancerous cells are starting to build. Monoclonal antibodies can be used to trigger the body’s immune system to

**Table 1. Checkpoint inhibitors.**

Cancer Type	Drug name	Status	Reference
Melanoma	Ipilimumab	FDA approved	[44][45][46]
	Pembrolizumab	FDA approved	[47]
Melanoma, Prostate	Enoblituzumab	Phase I-III	[48][49]
Multiple cancers	Ipilimumab	Phase I-III	[50]
	Tremelimumab	Phase I-III	[51]
	Nivolumab	Phase I-III	[52]
	Pembrolizumab	Phase I-III	[36]
	AMP-224	Phase I	[53]
	Pidilizumab	Phase I-III	[54]
	Atezolizumab	Phase I-III	[55]
	Avelumab	Phase I-III	[56]
	BMS-936559	Phase I	[57]
	IMP321	Phase I	[58]
BMS-986016	Phase I	[59]	
Renal Cell Carcinoma	Ipilimumab	FDA approved	[50]

**Table 2. A summary of some monoclonal antibodies used in immunotherapy of cancer.**

Antibody	Mechanism	Uses	Target	Current Status	References
Cetuximab	Binds to and inhibits EGFR	Metastatic colorectal ,Head & Neck cancer	EGFR	FDA approved	[65]
Bevacizumab	Blocks angiogenesis by inhibiting VEGF	Glioblastoma, Lung, Colon cancer	VEGF	FDA approved	[66]
Trastuzumab	Targets HER2 to induce an immune response	Metastatic Breast cancer	HER2	FDA approved	[67]
Panitumumab	Binds to EGFR to prevent its activation	Metastatic colorectal cancer	EGFR	FDA approved	[68]

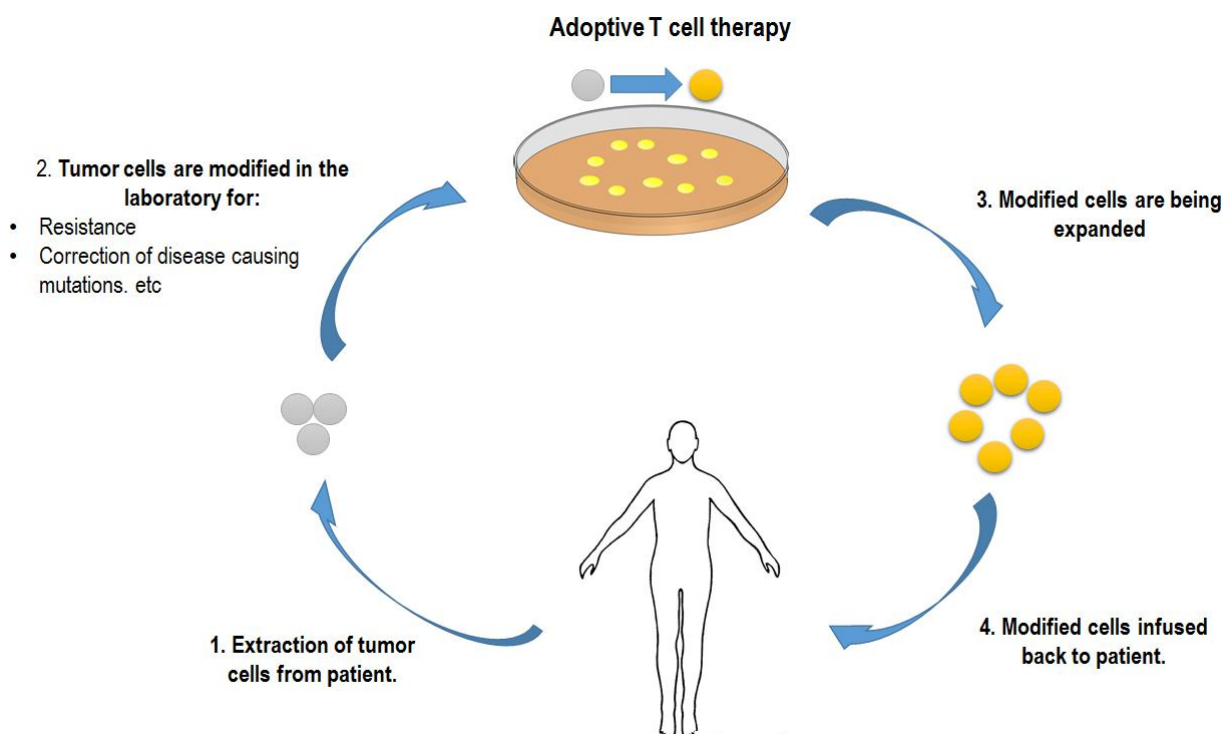
**EGFR:** Epidermal Growth Factor Receptor, **VEGF:** Vascular Endothelial Growth Factor, **HER2:** Human Epidermal Growth Factor Receptor 2.

recognize and eliminate cancer cells. Recently, the FDA approved several mAbs for the treatment of both solid tumors and hematological malignancies [61, 63]. Available reports suggest that using monoclonal antibodies against PD1 and CTLA-4 has been regarded as an important breakthrough in cancer immunotherapy [64]. This strategy proved to be efficacious towards patients with metastatic melanoma evidenced by the antitumor response and an increase in overall survival rate of patients treated with ipilimumab, a known monoclonal antibody that targets CTLA-4 in humans [64]. Clinical trials of monoclonal antibodies are ongoing for several types of cancers.

### 1.5. Cancer Vaccine

Another immunotherapeutic technique used in treating cancer is the use of vaccines. Most vaccines work by introducing a non-infectious version of a disease-causing microbe into an individual, thus providing a better stimulus to activate disease-specific T cells and to develop immunological memory. Immune memory cells can destroy microbes rapidly and prevent infection. This form of procedure has resulted in eradicating smallpox disease [69].

Cancer vaccine therapy differs from immunotherapies in that it initiates the immune system's activation in order to regain equilibrium between tumor cells and normal cells of the body [11, 70]. Like viral-targeted vaccines, cancer vaccines do not prevent infection but rather activate the immune system to combat a disease that already exists [69]. Most cancer vaccines are made of cancer cells, cell parts, or pure antigens. The immune cells of a patient are often isolated and exposed to antigens of cancer, and once activated, these immune cells are reintroduced into the patient's body and so that they are better able to suppress cancer cells [69]. Cancer cells develop from normal body cells, these cancer cells can be recognized as antigens (foreign substance) by immune cells known as dendritic cells [71]. These dendritic cells act as a commander to the immune system by initiating phagocytosis of cancer cells. Signals are sent to the lymphocytes to fight against these foreign substances [72]. The first anti-cancer vaccine was approved in 2010 by the Food and Drugs Authority (FDA). The vaccine is known as sipuleucel-T (Provenge) and is actively used for patients with metastatic, castration-resistant prostate cancer (CRPC)[73]. Sipuleucel-T is developed to stimulate T-cell immune responses against prostatic acid phosphatase (PAP),



**Fig. (3).** Adoptive T cell therapy.

an antigen that is expressed on most prostate cancer cells [39]. Recently, the FDA approved the use of hepatitis B virus vaccines (HBV) and human papillomavirus vaccines (HPV). These two vaccines work by stimulating the immune system with tumor peptides and antigens [9]. Cancer vaccines have varying side effects from patient to patient, however, the regularly reported side effect includes the inflammation at the site of the administered injection [72].

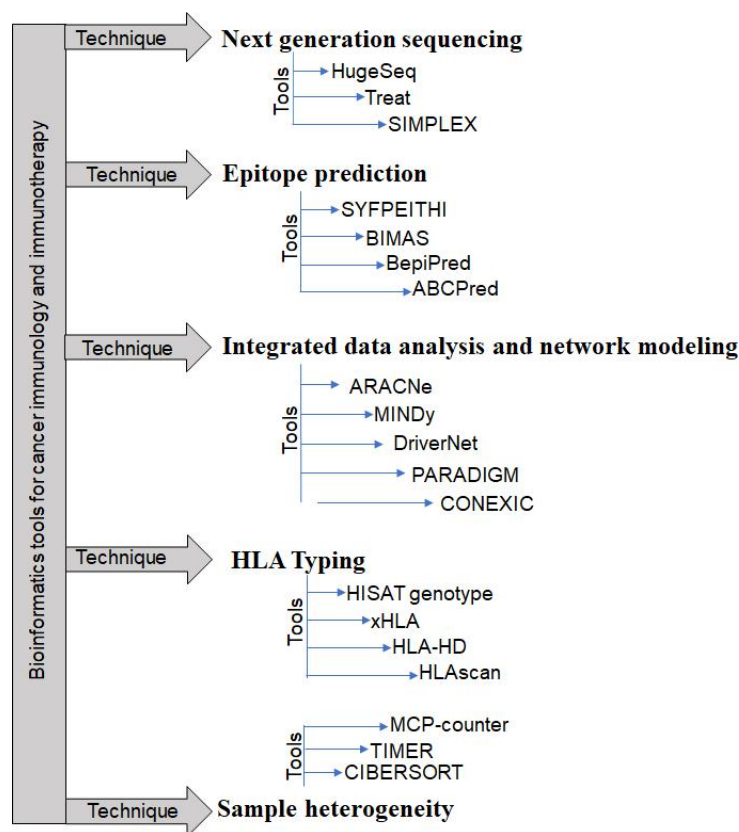
### 1.6. Adoptive T-cell Therapy (ACT)

In this immunotherapy strategy, T- cells are infused into a cancer patient to identify, target, and eliminate tumor cells from the body [75]. The mechanism of action in this technique involves extracting specific tumor cytotoxic T cells from a patient's blood and developing in the laboratory and then infused back into the patient to attack cancer cells, as shown in Fig. (3) [76]. This mechanism helps the immune system in fighting cancer cells. The T cells can sometimes be subjected to genetic modification and cultured in the laboratory to be more specific in targeting cancer cells [77, 78]. Adoptive T cell therapy consists of chimeric antigen receptor T -cell (CAR T-cell) and tumor-infiltrating lymphocyte (TIL)[76, 77]. The main difference between these two types relies on laboratory modifications. In the TILs method, tumor-specific cytotoxic T cells are extracted from the patients' tumor and modified in the laboratory with substances that easily activate lymphocytes and then given back to the patient, while in the CAR T-cell method, T cells extracted from a patient's blood are modified by adding a gene that targets a special receptor known as chimeric antigen receptor, known to bind to a protein on cancer cells [33, 79]. Large quantities of these activated lymphocytes, as well as CAR T cells, are then infused back to the patient to help the immune system fight cancer [43, 79]. The ACT has

gained prominence as a substantial strategy and shows great potential of being a curative technique for several cancers [38]. Nonetheless, the widespread of this therapy in solid tumors is one of ACT's main future goals due to the difficulties in finding suitable target antigens and also for tumor immunosuppression and complex tumor microenvironment. Furthermore, ACT requires optimization to reduce toxicity and increase the effectiveness of anti-tumors. Recent reports suggest that several ACT-based therapies are currently in the late-phase clinical testing, and some T cell therapies are already attaining regulatory approval for the treatment of patients with B cell malignancies.

### 1.7. Combinatorial Cancer Immunotherapy

Regardless of the notable clinical success of immunotherapy over other cancer treatment options, the approach is only favorable to some patients, as patients whose immune system fails to elicit an immune response, or faced by immune evasion, are left untreated[80]. This has led to the development of a combinatorial cancer treatment approach in recent times. This approach involves combining cancer immunotherapy along with other existing anticancer treatment strategies, an approach that has gained prominence significantly in preclinical or clinical studies[32]. A combinatorial strategy involving the checkpoint inhibitors with drugs that increase tumor immunogenicity, reduce tumor burden, and reverse tumor mediated immune suppression has recently been shown to be an effective and long-lasting antitumor response technique[15, 80]. This approach has shown remarkable success in several cancers and is beneficial to patients who fail to respond to checkpoint monotherapy.



**Fig. (4).** Bioinformatics tools for cancer immunology and immunotherapy.

Anticancer treatment strategy such as radiotherapy is mostly used in combination with surgeries or chemotherapy, however, there are emerging regimens involving a combination of radiotherapy and immunotherapy treatments for cancer [81]. Ongoing studies suggest that a combination of radiotherapy and immunotherapies has a high potential of boosting abscopal response rates and will possibly expand the use of radiotherapy in the treatment for both local and metastatic diseases [82, 83]. Nonetheless, various preclinical studies are continuously being carried out to explore the role of radiation in combination with immune checkpoint inhibitors. A combination of epigenetic therapy and immunotherapy, such as checkpoint immunotherapy, has gained significant recognition in recent studies [84]. Epigenetic modulators, in combination with checkpoint inhibitors, are known to increase T cell infiltration in tumor microenvironments (TME), reduce Myeloid -derived suppressor cells (MDSC) in TME and augment surface expression of immune checkpoints [84].

### 1.8. Application of Bioinformatics Strategies in Cancer Immunotherapy

Recent studies have shown an increase in the development of tumor tolerance to immunotherapy after a limited period of successful treatment. Cataloging, predicting, and selecting immunotherapy targets can be extensively addressed using the existing bioinformatics tools and biological databases. Identification and selection of antigens have many different features that depend both on the type and the application of antigens. Identification of

potential antigens de novo from genomic sequence using bioinformatics tools is very difficult since the expression of proteins is regulated by an array of complex regulatory mechanisms, many of which are poorly understood [85]. As part of a long-established practice, tumor antigens are identified in vitro from serum by screening cDNA phage libraries, using the immunoassays [86] or proteomics -based screening [87], but bioinformatics tools are perfectly suited to aid this process, either by actively recognizing novel tumor antigens or by organizing and accessing information about the known tumor antigens in databases. The techniques discussed in this review include next-generation sequencing, epitope prediction, integrated data analysis and network modeling, HLA typing, sampling heterogeneity and modeling of tumor-immune cell interactions (Fig. 4). All these techniques employ various tools in executing their function.

### 1.9. Next-Generation Sequencing

Next-Generation Sequencing (NGS) is a large parallel sequencing technology that has revolutionized the biological sciences [88] [85]. NGS provides quantitative insights into the tumor cell's molecular machinery, Improves transcript and gene expression profiling as well as detection of alternative splicing and enables the discovery of single nucleotide (SNV) variants, insertions, amplifications, deletions and inter chromosome rearrangements throughout the genome and transcriptome [85]. NGS uses tools such as HugeSeq, a fully integrated pipeline for NGS analysis from aligning reads to identification and annotation of all types of

variants (SNPs, Indels, CNVs, SVs)[89]. Sailana and colleagues applied HugeSeq to investigate the association of HSG with alopecia and mental retardation syndrome [90]. From this research, they were able to map a novel gene (APMR2) to chromosome 3q26.2-q26.31 [90]. Furthermore, using the same techniques, the APMR3 gene was mapped on chromosome 18q11.2 -q12.2 [90]. Another tool used in NGS is the SIMPLEX, an autonomous analysis pipeline for the analysis of NGS exome data, covering the workflow from sequence alignment to SNP/DIP identification and variant annotation[85]. It supports input from various sequencing platforms and exposes all available parameters for customized usage [85].

### 1.10. Integrated Data Analysis and Network Modeling Technique

The integrated data analysis and network modeling technique enables one to comprehensively study molecular mechanisms of cancer cells and their interactions within the immune system [91]. The tools used in implementing this technique include ARACNe[92] and MINDy[93], which aid in the reconstruction of gene co-expression networks, genes (nodes) with comparable global expression profiles over samples (tumor/patients). These tools may also be integrated to identify key transcriptional regulators [91]. ARACNe was used to obtain a database of transcriptional interactions by using *Arabidopsis thaliana* root samples. Using ARACNe, Montes et al. discovered that the transcriptional regulatory networks, derived from this newly constructed database, successfully recover previously identified root transcriptional modules and propose new transcription factors for the short root/scarecrow and plethora pathways [33]. DriverNet[95], PARADIGM, and CONEXIC[96] are additional tools for implementing the integrated data analysis and network modeling technique that are used to address the functional importance of specific genome alterations, identify which pathways are affected. As well as identify mutations likely to be drivers in tumor progression [41]. Fig. (4) summarizes examples of bioinformatics techniques and tools for immunology and immunotherapy. DriverNet was used to identify driver genes through the integration of multi-omics data, such as somatic mutation, gene expression, and copy number alterations [97].

#### 1.11. HLA Typing

HLAs possess very high polymorphic region and have been touted as the most polymorphic segments present in the human genome [98]. This high polymorphism has made it difficult to unravel the HLA-mediated immunogenic correlation of immuno-pathologies [98]. HLA typing is employed in organ transplants to get a corresponding potential donor that matches with a patient [99]. As technology advances, HLA typing has concurrently advanced. It was originally carried out by using hybridization strategy, serological typing, cellular typing, DNA-based typing and restriction fragment length polymorphism (RFLP) techniques. Polymerase Chain Reaction approach and direct sequencing capillary electrophoretic analyses were later employed [100]. With current development in computation, second and third-generation sequencing platforms are now employed in HLA typing. This current approach has shown great potential and

high predictability [100, 101]. Several HLA typing platforms have been reported, such as HISAT genotype [102], xHLA [99], HLA-HD [103], HLAScan [104]. There has been a rise in the amount of HLA sequences and new alleles deposited in HLA databases, such as the IMG/HLA database [105], CWD HLA allele [106]. HLA typing software, which has been developed for easy and rapid HLA typing, includes ATHLATES [107], POLYSOLVER [108], and OpiType [48]. HLA typing of genes has been applied to unreported haplotypes to unravel the complexity of HLA-related pathological conditions and associated clinical subtypes. A typical example is the application of HLA typing in psoriasis, novel associations of HLA-DPB1 and BTNL2 genes and five loci among HLA-C, -B, -DPB1, and BTNL2 were discovered with the aid of NGS [98]. Furthermore, specific HLA associations were also reported for the overall risk of psoriasis vulgaris and the risk of its specific subphenotypes, such as psoriatic arthritis and cutaneous psoriasis [98].

### 1.12. Epitope Prediction

B-cells are a very important component of the adaptive system, this is due to the long-lasting protection they provide to combat foreign assault or pathogenic onslaught [110]. Some cancer treatment and prevention strategies involve the use of vaccines. Bioinformatics has provided easy and fast means of vaccine development, however, it is expedient to first distinguish between epitopes that are immuno-protective and those that are not [111]. B-cell identification is paramount in many medical, biological, and immunological applications, such as the development of vaccines, control of disease and diagnosis [112]. Experimental techniques used in the identification of epitope include protein crystallography, ELISA, and peptide-chip, however, these methods are time-consuming, expensive, have a low-accuracy and low throughput generation [110]. Due to these experimental shortcomings, computational strategies have been developed to corroborate or as a stand-alone tool for epitope prediction. BepiPred [113, 114], DiscoTope [114, 115], CBtope [116], and ABCpred [117]. Epitope prediction has found useful applications in the design of vaccines against some diseases. Qamar et al. used some in silico bioinformatic approaches to design epitope-based vaccines against middle east respiratory syndrome coronavirus [118]. They identified conserved B and -T cells for the MERS-COV spike (S) protein that may perform a significant role in eliciting the resistance response to MERS-COV infection [118].

### 1.13. Sample Heterogeneity

Deconvolution is the estimation of the relative immune cells that are present in a sample after DNA/RNA sequencing [119]. This is made possible through the recognition of biomarkers and employing their level of expression to discriminate cell types [120]. Tools used in deconvolution evaluation include CIBERSORT [121], TIMER [122], and MCP-counter [123]. They make use of expression data gotten from cell types generated by the Immunological Genome Project (ImmGen) [124]. TIMER has the capacity to evaluate 6 immune subtypes, CIBERSORT evaluates 22 immune cell subtypes, and MCP-



counter evaluates 8 immune cell subtypes and 2 stromal cell subtypes. MCP-counter evaluates absolute abundances, while CIBERSORT and TIMER evaluate relative frequencies of immune cell subtypes [125].

#### 1.14. Modeling of Tumor-immune Cell Interactions

There is a long history of theoretical studies and simulation techniques involving mathematical and computational approaches to study tumor progression and tumor-immune cell interaction. Computational modeling has provided essential insight into studying intra-tumor heterogeneity and the vital interplay between the tumor and the surrounding microenvironment [126]. Computational models allow the investigation of the tumor-enhancing effects of the immune system providing indispensable insight aiding in the management of tumor-immune interactions. Modeling has consecutively provided a quantitative time- and cost-effective avenue to study the physical and chemical interactions in tumor initiation growth [127]. The integration of molecular modeling techniques further complements experimental platforms by providing in-depth knowledge into clonal dynamics and microenvironmental cues over time. There are a number of techniques which include deterministic models, stochastic models, Petri nets, cellular automata, agent-based model, and hybrid approaches [128]. The integration of one or more of these models will help improve immunotherapies for cancer treatment.

#### 1.15. Future Perspectives

The advent of computational methods in immunotherapy of cancer has helped transform the treatment paradigm of several tumors and is expected to influence the therapeutic efficiency of immunotherapeutic strategies in the future. The application of computational methods in cancer immunotherapy, however, has some limitations. First, the identification of T cell epitopes, using computational techniques, is not yet as accurate as peptide binding prediction algorithms. Moreover, the availability of tumor sequences represents a hold-up in conservation and variability, nonetheless, a remedy for this limitation is plausible in the near future as high-throughput screening is becoming readily affordable and more efficient. Additionally, these current methods may not adequately capture the issue of intra-tumor genetic diversity as such may affect the efficacy of immunotherapy and other cancer therapies alike. The use of proteomics analyses in immunotherapy has shown to be more prolific in the past years in contrast to genomic analyses. Some promising advances are currently being made in the wet laboratory to address these limitations in diverse ways. This will progressively boost the need for bioinformatics tools in cancer immunotherapy.

#### CONCLUSION

Cancer immunotherapy has gained a lot of recognition since its incision in the past decades by introducing very efficient strategies of using the body's own immune response. The active cancer vaccine, immune checkpoint inhibitors, T cell adoptive therapies are among these

developments that have contributed to increased survival rate among patients. Despite these advances, emerging research in the interdisciplinary fields of bioinformatics, immunotherapy, and immunology is significant to further enhance therapeutic advantage and reduce side effects. Studies presented in this review provide a general overview of cancer immunotherapy giving some insights into the relationship of cancer and the immune system. Furthermore, some immunotherapeutic strategies were discussed, including cancer vaccine, immune checkpoint inhibitors, T cell adoptive therapies, monoclonal antibodies, and combinatorial cancer immunotherapy. Finally, we discussed some of the applications of bioinformatics tools in cancer immunotherapy and some future perspectives. As tumor cells' response to immunotherapies is gradually being unraveled and the body of biological tumor data grows, so will the need for bioinformatics to organize, store, and analyze these data.

#### CONSENT FOR PUBLICATION

Not applicable.

#### FUNDING

None.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

#### ACKNOWLEDGEMENTS

Declared none.

#### REFERENCES

- [1] Registry, P.C. Global Cancer Observatory. *Malaysia Cancer Statistics*; , **2019**.
- [2] Bray, G.C. Global Cancer Facts & Figures. *Cancer*, **2007**.
- [3] Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.*, **2018**, *68*(6), 394-424. <http://dx.doi.org/10.3322/caac.21492> PMID: 30207593
- [4] Sharma, P.; Hu-Lieskovan, S.; Wargo, J.A.; Ribas, A. Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. *Cell*, **2017**, *168*(4), 707-723. <http://dx.doi.org/10.1016/j.cell.2017.01.017> PMID: 28187290
- [5] Rius, M.; Lyko, F. Epigenetic cancer therapy: rationales, targets and drugs. *Oncogene*, **2012**, *31*(39), 4257-4265. <http://dx.doi.org/10.1038/onc.2011.601> PMID: 22179827
- [6] Barton, M.K. Daily aspirin may reduce mortality from prostate cancer with risk of high recurrence. *CA Cancer J. Clin.*, **2015**, *65*(2), 83-84. <http://dx.doi.org/10.3322/caac.21263> PMID: 25640813
- [7] Borghaei, H.; Smith, M.R.; Campbell, K.S. Immunotherapy of cancer. *Eur. J. Pharmacol.*, **2009**, *625*(1-3), 41-54. <http://dx.doi.org/10.1016/j.ejphar.2009.09.067> PMID: 19837059
- [8] Cavallo, F.; De Giovanni, C.; Nanni, P.; Forni, G.; Lollini, P.L. The immune hallmarks of cancer. *Cancer Immunology; Immunotherapy*, **2011**.
- [9] Makkouk, A.; Weiner, G.J. Cancer immunotherapy and breaking immune tolerance: new approaches to an old challenge. *Cancer Res.*, **2015**, *75*(1), 5-10. <http://dx.doi.org/10.1158/0008-5472.CAN-14-2538> PMID: 25524899

- [10] Thommen, D.S. The First Shall (Be) Last: Understanding Durable T Cell Responses in Immunotherapy. *Immunity*, **2019**, *50*(1), 6-8. <http://dx.doi.org/10.1016/j.immuni.2018.12.029> PMID: 30650381
- [11] Speiser, D.E.; Flatz, L. Cancer immunotherapy drives implementation science in oncology. *Hum. Vaccin. Immunother.*, **2014**, *10*(11), 3107-3110. <http://dx.doi.org/10.4161/21645515.2014.983000> PMID: 25625923
- [12] Yang, Y. Cancer immunotherapy: harnessing the immune system to battle cancer. *J. Clin. Invest.*, **2015**, *125*(9), 3335-3337. <http://dx.doi.org/10.1172/JCI83871> PMID: 26325031
- [13] Tovoli, F.; Casadei-Gardini, A.; Benevento, F.; Piscaglia, F. Immunotherapy for hepatocellular carcinoma: A review of potential new drugs based on ongoing clinical studies as of 2019. *Dig. Liver Dis.*, **2019**, *51*(8), 1067-1073. <http://dx.doi.org/10.1016/j.dld.2019.05.006> PMID: 31208929
- [14] Kruger, S.; Ilmer, M.; Kobold, S.; Cadilha, B.L.; Endres, S.; Ormanns, S.; Schuebbe, G.; Renz, B.W.; D'Haese, J.G.; Schloesser, H.; Heinemann, V.; Subklewe, M.; Boeck, S.; Werner, J.; von Bergwelt-Baildon, M. Advances in cancer immunotherapy 2019 - latest trends. *J. Exp. Clin. Cancer Res.*, **2019**, *38*(1), 268. <http://dx.doi.org/10.1186/s13046-019-1266-0> PMID: 31217020
- [15] Nisbet, I. Cancer immunotherapy comes of age (Finally!). *Australas. Biotechnol.*, **2016**. PMID: 27011048
- [16] Binder, R. J. Functions of heat shock proteins in pathways of the innate and adaptive immune system. *Journal of immunology (Baltimore, Md. : 1950)*, **2014**, *193*(12), 5765-5771. <http://dx.doi.org/10.4049/jimmunol.1401417>
- [17] Voena, C.; Chiarle, R. Advances in cancer immunology and cancer immunotherapy. *Discov. Med.*, **2016**, *21*(114), 125-133. PMID: 27011048
- [18] Miller, J.F.A.P.; Sadelain, M. The journey from discoveries in fundamental immunology to cancer immunotherapy. *Cancer Cell*, **2015**, *27*(4), 439-449. <http://dx.doi.org/10.1016/j.ccell.2015.03.007> PMID: 25858803
- [19] Woo, S.-R.; Corrales, L.; Gajewski, T.F. Innate immune recognition of cancer. *Annu. Rev. Immunol.*, **2015**, *33*(1), 445-474. <http://dx.doi.org/10.1146/annurev-immunol-032414-112043> PMID: 25622193
- [20] Trinchieri, G.; Perussia, B. Immune interferon: a pleiotropic lymphokine with multiple effects. *Immunol. Today*, **1985**, *6*(4), 131-136. [http://dx.doi.org/10.1016/0167-5699\(85\)90080-5](http://dx.doi.org/10.1016/0167-5699(85)90080-5) PMID: 25289500
- [21] Farrar, M.A.; Schreiber, R.D. The Molecular Cell Biology of Interferon-gamma and its Receptor. *Annu. Rev. Immunol.*, **2003**. <http://dx.doi.org/10.1146/annurev.iy.11.040193.003035> PMID: 8476573
- [22] Bevan, M.J. Helping the CD8(+) T-cell response. *Nat. Rev. Immunol.*, **2004**, *4*(8), 595-602. <http://dx.doi.org/10.1038/nri1413> PMID: 15286726
- [23] Ribatti, D. The concept of immune surveillance against tumors. The first theories. *Oncotarget*, **2017**, *8*(4), 7175-7180. <http://dx.doi.org/10.18632/oncotarget.12739> PMID: 27764780
- [24] Coley, W.B. The Treatment of Inoperable Sarcoma by Bacterial Toxins (the Mixed Toxins of the Streptococcus erysipelas and the Bacillus prodigiosus). *Proceedings of the Royal Society of Medicine*, **1910**, *3*(Surg Sect), pp. 1-48.
- [25] Johnson, D.B.; Sullivan, R.J.; Menzies, A.M. Immune checkpoint inhibitors in challenging populations. *Cancer*, **2017**, *123*(11), 1904-1911. <http://dx.doi.org/10.1002/cncr.30642> PMID: 28241095
- [26] Ito, A.; Kondo, S.; Tada, K.; Kitano, S. Clinical Development of Immune Checkpoint Inhibitors. *BioMed Res. Int.*, **2015**, *2015*605478. <http://dx.doi.org/10.1155/2015/605478> PMID: 26161407
- [27] Mahoney, K.M.; Freeman, G.J.; McDermott, D.F. The next immune-checkpoint inhibitors: Pd-1/pd-l1 blockade in melanoma. *Clin. Ther.*, **2015**, *37*(4), 764-782. <http://dx.doi.org/10.1016/j.clinthera.2015.02.018> PMID: 25823918
- [28] Spain, L.; Diem, S.; Larkin, J. Management of toxicities of immune checkpoint inhibitors. *Cancer Treat. Rev.*, **2016**, *44*, 51-60. <http://dx.doi.org/10.1016/j.ctrv.2016.02.001> PMID: 26874776
- [29] Ceeraz, S.; Nowak, E.C.; Burns, C.M.; Noelle, R.J. Immune checkpoint receptors in regulating immune reactivity in rheumatic disease. *Arthritis Res. Ther.*, **2014**, *16*(5), 469. <http://dx.doi.org/10.1186/s13075-014-0469-1> PMID: 25606596
- [30] Pardoll, D.M. The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer*, **2012**, *12*(4), 252-264. <http://dx.doi.org/10.1038/nrc3239> PMID: 22437870
- [31] Yuan, J.; Hegde, P.S.; Clynes, R.; Foukas, P.G.; Harari, A.; Kleen, T.O.; Kvistborg, P.; Maccalli, C.; Maecker, H.T.; Page, D.B.; Robins, H.; Song, W.; Stack, E.C.; Wang, E.; Whiteside, T.L.; Zhao, Y.; Zwierzina, H.; Butterfield, L.H.; Fox, B.A. Novel technologies and emerging biomarkers for personalized cancer immunotherapy. *J. Immunother. Cancer*, **2016**, *4*, 3. <http://dx.doi.org/10.1186/s40425-016-0107-3> PMID: 26788324
- [32] Sathyanarayanan, V.; Neelapu, S.S. Cancer immunotherapy: Strategies for personalization and combinatorial approaches. *Mol. Oncol.*, **2015**, *9*(10), 2043-2053. <http://dx.doi.org/10.1016/j.molonc.2015.10.009> PMID: 26548534
- [33] Hamanishi, J.; Mandai, M.; Iwasaki, M.; Okazaki, T.; Tanaka, Y.; Yamaguchi, K. Programmed cell death 1 ligand 1 and tumor-infiltrating CD8+ T lymphocytes are prognostic factors of human ovarian cancer. *Proceedings of the National Academy of Sciences*, **2007**. <http://dx.doi.org/10.1073/pnas.0611533104>
- [34] Okazaki, T.; Honjo, T. PD-1 and PD-1 ligands: from discovery to clinical application. *Int. Immunol.*, **2007**, *19*(7), 813-824. <http://dx.doi.org/10.1093/intimm/dxm057> PMID: 17606980
- [35] Taube, J.M.; Klein, A.; Brahmer, J.R.; Xu, H.; Pan, X.; Kim, J.H.; Chen, L.; Pardoll, D.M.; Topalian, S.L.; Anders, R.A. Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin. Cancer Res.*, **2014**, *20*(19), 5064-5074. <http://dx.doi.org/10.1158/1078-0432.CCR-13-3271> PMID: 24714771
- [36] Robert, C.; Schachter, J.; Long, G.V.; Arance, A.; Grob, J.J.; Mortier, L.; Daud, A.; Carlino, M.S.; McNeil, C.; Lotem, M.; Larkin, J.; Lorigan, P.; Neyns, B.; Blank, C.U.; Hamid, O.; Mateus, C.; Shapira-Frommer, R.; Kosh, M.; Zhou, H.; Ibrahim, N.; Ebbinghaus, S.; Ribas, A. KEYNOTE-006 investigators. Pembrolizumab versus Ipilimumab in Advanced Melanoma. *N. Engl. J. Med.*, **2015**, *372*(26), 2521-2532. <http://dx.doi.org/10.1056/NEJMoa1503093> PMID: 25891173
- [37] Larkin, J.; Chiarion-Sileni, V.; Gonzalez, R.; Grob, J.J.; Cowey, C.L.; Lao, C.D.; Schadendorf, D.; Dummer, R.; Smylie, M.; Rutkowski, P.; Ferrucci, P.F.; Hill, A.; Wagstaff, J.; Carlino, M.S.; Haanen, J.B.; Maio, M.; Marquez-Rodas, I.; McArthur, G.A.; Ascierto, P.A.; Long, G.V.; Callahan, M.K.; Postow, M.A.; Grossmann, K.; Szno, M.; Dreno, B.; Bastholt, L.; Yang, A.; Rollin, L.M.; Horak, C.; Hodi, F.S.; Wolchok, J.D. Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. *N. Engl. J. Med.*, **2015**, *373*(1), 23-34. <http://dx.doi.org/10.1056/NEJMoa1504030> PMID: 26027431
- [38] Migden, M.R.; Rischin, D.; Schmultz, C.D.; Guminski, A.; Hauschild, A.; Lewis, K.D.; Chung, C.H.; Hernandez-Aya, L.; Lim, A.M.; Chang, A.L.S.; Rabinowits, G.; Thai, A.A.; Dunn, L.A.; Hughes, B.G.M.; Khushalani, N.I.; Modi, B.; Schadendorf, D.; Gao, B.; Seebach, F.; Li, S.; Li, J.; Mathias, M.; Booth, J.; Mohan, K.; Stankevich, E.; Babiker, H.M.; Brana, I.; Gil-Martin, M.; Homs, J.; Johnson, M.L.; Moreno, V.; Niu, J.; Owonikoko, T.K.; Papadopoulos, K.P.; Yancopoulos, G.D.; Lowy, I.; Fury, M.G. PD-1 Blockade with Cemiplimab in Advanced Cutaneous Squamous-Cell Carcinoma. *N. Engl. J. Med.*, **2018**, *379*(4), 341-351. <http://dx.doi.org/10.1056/NEJMoa1805131> PMID: 29863979
- [39] Fehrenbacher, L.; Spira, A.; Ballinger, M.; Kowanzet, M.; Vansteenkiste, J.; Mazieres, J.; Park, K.; Smith, D.; Artal-Cortes, A.; Lewanski, C.; Braiteh, F.; Waterkamp, D.; He, P.; Zou, W.; Chen, D.S.; Yi, J.; Sandler, A.; Rittmeyer, A. POPLAR Study Group. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. *Lancet*, **2016**, *387*(10030), 1837-1846. [http://dx.doi.org/10.1016/S0140-6736\(16\)00587-0](http://dx.doi.org/10.1016/S0140-6736(16)00587-0) PMID: 26970723

- [40] Tsang, K.-Y.; Boyerinas, B.; Jochems, C.; Fantini, M.; Heery, C.R.; Madan, R.A. ... Schlom, J. (2019). Antibody dependent cellular cytotoxicity activity of a novel anti-PD-L1 antibody, avelumab (MSB0010718C), on human tumor cells. *J. Clin. Oncol.*, [http://dx.doi.org/10.1200/jco.2015.33.15\\_suppl.3038](http://dx.doi.org/10.1200/jco.2015.33.15_suppl.3038)
- [41] Approved. (2010). *Chem. Eng. News*, <http://dx.doi.org/10.1021/cen-v040n029.obc>
- [42] Farkona, S.; Diamandis, E.P.; Blasutig, I.M. Cancer immunotherapy: the beginning of the end of cancer? *BMC Med.*, **2016**, *14*, 73. <http://dx.doi.org/10.1186/s12916-016-0623-5> PMID: 27151159
- [43] Voena, C.; Di Giacomo, F.; Panizza, E.; D'Amico, L.; Boccalatte, F.E.; Pellegrino, E.; Todaro, M.; Recupero, D.; Tabbò, F.; Ambrogio, C.; Martino, C.; Bonello, L.; Pulito, R.; Hamm, J.; Chiarle, R.; Cheng, M.; Ruggeri, B.; Medico, E.; Inghirami, G. The EGFR family members sustain the neoplastic phenotype of ALK+ lung adenocarcinoma via EGR1. *Oncogenesis*, **2013**, *2*(4)e43 <http://dx.doi.org/10.1038/oncsis.2013.7> PMID: 23567620
- [44] Syn, N.L.; Teng, M.W.L.; Mok, T.S.K.; Soo, R.A. De-novo and acquired resistance to immune checkpoint targeting. *Lancet Oncol.*, **2017**, *18*(12), e731-e741. [http://dx.doi.org/10.1016/S1470-2045\(17\)30607-1](http://dx.doi.org/10.1016/S1470-2045(17)30607-1) PMID: 29208439
- [45] FDA approves new treatment for a type of late-stage skin cancer. *U.S. Food and Drug Administration (FDA)*, **2011**.
- [46] Pollack, A. Approval for Drug That Treats Melanoma. *The New York Times*, **2011**.
- [47] Redman, J.M.; Gibney, G.T.; Atkins, M.B. Advances in immunotherapy for melanoma. *BMC Med.*, **2016**, *14*, 20. <http://dx.doi.org/10.1186/s12916-016-0571-0> PMID: 26850630
- [48] Rizvi, N.A.; Loo, D.; Baughman, J.E.; Yun, S.; Chen, F.; Moore, P.A. ... Tolcher, A. W. (2016). A phase 1 study of enoblituzumab in combination with pembrolizumab in patients with advanced B7-H3-expressing cancers. *J. Clin. Oncol.*, [http://dx.doi.org/10.1200/jco.2016.34.15\\_suppl.tps3104](http://dx.doi.org/10.1200/jco.2016.34.15_suppl.tps3104)
3. Shenderov, E.; Demarzo, A.; Boudadi, K.; Allaf, M.; Wang, H.; Chapman, C. ... Antonarakis, E. S. (2018) . Phase II neoadjuvant and immunologic study of B7-H3 targeting with enoblituzumab in localized intermediate- and high-risk prostate cancer. *J. Clin. Oncol.*, [http://dx.doi.org/10.1200/jco.2018.36.15\\_suppl.tps5099](http://dx.doi.org/10.1200/jco.2018.36.15_suppl.tps5099)
- [50] Hodi, F.S.; O'Day, S.J.; McDermott, D.F.; Weber, R.W.; Sosman, J.A.; Haanen, J.B.; Gonzalez, R.; Robert, C.; Schadendorf, D.; Hassel, J.C.; Akerley, W.; van den Eertwegh, A.J.; Lutzky, J.; Lorigan, P.; Vaubel, J.M.; Linette, G.P.; Hogg, D.; Ottensmeier, C.H.; Lebbé, C.; Peschel, C.; Quirt, I.; Clark, J.L.; Wolchok, J.D.; Weber, J.S.; Tian, J.; Yellin, M.J.; Nichol, G.M.; Hoos, A.; Urba, W.J. Improved survival with ipilimumab in patients with metastatic melanoma. *N. Engl. J. Med.*, **2010**, *363*(8), 711-723. <http://dx.doi.org/10.1056/NEJMoa1003466> PMID: 20525992
- [51] Ribas, A.; Kefford, R.; Marshall, M.A.; Punt, C.J.A.; Haanen, J.B.; Marmol, M.; Garbe, C.; Gogas, H.; Schachter, J.; Linette, G.; Lorigan, P.; Kendra, K.L.; Maio, M.; Trefzer, U.; Smylie, M.; McArthur, G.A.; Dreno, B.; Nathan, P.D.; Mackiewicz, J.; Kirkwood, J.M.; Gomez-Navarro, J.; Huang, B.; Pavlov, D.; Hauschild, A. Phase III randomized clinical trial comparing tremelimumab with standard-of-care chemotherapy in patients with advanced melanoma. *J. Clin. Oncol.*, **2013**, *31*(5), 616-622. <http://dx.doi.org/10.1200/JCO.2012.44.6112> PMID: 23295794
- [52] Larkin, J.; Chiarion-Sileni, V.; Gonzalez, R.; Grob, J.J.; Cowey, C.L.; Lao, C.D.; Schadendorf, D.; Dummer, R.; Smylie, M.; Rutkowski, P.; Ferrucci, P.F.; Hill, A.; Wagstaff, J.; Carlino, M.S.; Haanen, J.B.; Maio, M.; Marquez-Rodas, I.; McArthur, G.A.; [64] Ascierto, P.A.; Long, G.V.; Callahan, M.K.; Postow, M.A.; Grossmann, K.; Sznol, M.; Dreno, B.; Bastholt, L.; Yang, A.; Rollin, L.M.; Horak, C.; Hodi, F.S.; Wolchok, J.D. ... Wolchok, J. D. (2015). Combined nivolumab and ipilimumab or monotherapy in untreated Melanoma. *N. Engl. J. Med.*, **2015**, *373*(1), 23-34. <http://dx.doi.org/10.1056/NEJMoa1504030> PMID: 26027431
- [53] Duffy, A.G.; Makarova-Rusher, O.V.; Pratt, D.; Kleiner, D.E.; Fioravanti, S.; Walker, M. ... Greten, T. F. (2016). A pilot study of AMP-224, a PD-L2 Fc fusion protein, in combination with stereotactic body radiation therapy (SBRT) in patients with metastatic colorectal cancer. *J. Clin. Oncol.*, [http://dx.doi.org/10.1200/jco.2016.34.4\\_suppl.560](http://dx.doi.org/10.1200/jco.2016.34.4_suppl.560)
- [54] Armand, P.; Nagler, A.; Weller, E.A.; Devine, S.M.; Avigan, D.E.; Chen, Y.B.; Kaminski, M.S.; Holland, H.K.; Winter, J.N.; Mason, J.R.; Fay, J.W.; Rizzieri, D.A.; Hosing, C.M.; Ball, E.D.; Uberti, J.P.; Lazarus, H.M.; Mapara, M.Y.; Gregory, S.A.; Timmerman, J.M.; Andorsky, D.; Or, R.; Waller, E.K.; Rotem- Yehudar, R.; Gordon, L.I. Disabling immune tolerance by programmed death-1 blockade with pidilizumab after autologous hematopoietic stem-cell transplantation for diffuse large B-cell lymphoma: results of an international phase II trial. *J. Clin. Oncol.*, **2013**, *31*(33), 4199-4206. <http://dx.doi.org/10.1200/JCO.2012.48.3685> PMID: 24127452
- [55] Santini, F.C.; Rudin, C.M. Atezolizumab for the treatment of non-small cell lung cancer. *Expert Rev. Clin. Pharmacol.*, **2017**, *10*(9), 935-945. <http://dx.doi.org/10.1080/17512433.2017.1356717> PMID: 28714780
- [56] Boyerinas, B.; Jochems, C.; Fantini, M.; Heery, C.R.; Gulley, J.L.; Tsang, K.Y.; Schlom, J. Antibody-dependent cellular cytotoxicity activity of a Novel Anti-PD-L1 antibody avelumab (MSB0010718C) on human tumor cells. *Cancer Immunol. Res.*, **2015**, *3*(10), 1148-1157. <http://dx.doi.org/10.1158/2326-6066.CIR-15-0059> PMID: 26014098
- [57] Gay, C.L.; Bosch, R.J.; Ritz, J.; Hataye, J.M.; Aga, E.; Tressler, R.L.; Mason, S.W.; Hwang, C.K.; Grasele, D.M.; Ray, N.; Cyktor, J.C.; Coffin, J.M.; Acosta, E.P.; Koup, R.A.; Mellors, J.W.; Eron, J.J. AIDS Clinical Trials 5326 Study Team. Clinical trial of the anti-PD-L1 antibody BMS-936559 in HIV-1 infected participants on suppressive antiretroviral therapy. *J. Infect. Dis.*, **2017**, *215*(11), 1725-1733. <http://dx.doi.org/10.1093/infdis/jix191> PMID: 28431010
- [58] Brignone, C.; Escudier, B.; Grygar, C.; Marcu, M.; Triebel, F. A phase I pharmacokinetic and biological correlative study of IMP321, a novel MHC class II agonist, in patients with advanced renal cell carcinoma. *Clin. Cancer Res.*, **2009**, *15*(19), 6225-6231. <http://dx.doi.org/10.1158/1078-0432.CCR-09-0068> PMID: 19755389
- Lipson, E.; Gopal, A.; Neelapu, S.S.; Armand, P.; Spurgeon, S.; Leonard, J.P. Initial experience administering BMS-986016, a monoclonal antibody that targets lymphocyte activation gene (LAG)-3, alone and in combination with nivolumab to patients with hematologic and solid malignancies. *Journal for immunotherapy of cancer. Conference: 31st annual meeting and associated programs of the society for immunotherapy of cancer, SITC 2016.*, United states. Conference start: 20161109. Conference end: 20161113. **2016**.
- Riethmüller, G.; Schneider-Gädick, E.; Johnson, J.P. Monoclonal antibodies in cancer therapy. *Curr. Opin. Immunol.*, **1993**, *5*(5), 732-739. [http://dx.doi.org/10.1016/0952-7915\(93\)90129-G](http://dx.doi.org/10.1016/0952-7915(93)90129-G) PMID: 8240735
- [61] Henricks, L.M.; Schellens, J.H.M.; Huitema, A.D.R.; Beijnen, J.H. The use of combinations of monoclonal antibodies in clinical oncology. *Cancer Treat. Rev.*, **2015**, *41*(10), 859-867. <http://dx.doi.org/10.1016/j.ctrv.2015.10.008> PMID: 26547132
- [62] Shore, N.D. Advances in the understanding of cancer immunotherapy. *BJU Int.*, **2015**, *116*(3), 321-329. <http://dx.doi.org/10.1111/bju.12692> PMID: 24612369
- [63] Maleki, L.A.; Baradaran, B.; Majidi, J.; Mohammadian, M.; Shahneh, F.Z. Future prospects of monoclonal antibodies as magic bullets in immunotherapy. *Hum. Antibodies*, **2013**, *22*(1-2), 9-13. <http://dx.doi.org/10.3233/HAB-130266> PMID: 24284304
- Hodi, F.S.; Chiarion-Sileni, V.; Gonzalez, R.; Grob, J.-J.; Rutkowski, P.; Cowey, C.L.; Lao, C.D.; Schadendorf, D.; Wagstaff, J.; Dummer, R.; Ferrucci, P.F.; Smylie, M.; Hill, A.; Hogg, D.; Marquez-Rodas, I.; Jiang, J.; Rizzo, J.; Larkin, J.; Wolchok, J.D. Nivolumab plus ipilimumab or nivolumab alone versus ipilimumab alone in advanced melanoma (CheckMate 067): 4-year outcomes of a multicentre, randomised, phase 3 trial. *Lancet Oncol.*, **2018**, *19*(11), 1480-1492. [http://dx.doi.org/10.1016/S1470-2045\(18\)30700-9](http://dx.doi.org/10.1016/S1470-2045(18)30700-9) PMID: 30361170
- [65] Beljanski, V. Cetuximab. *xPharm: The Comprehensive Pharmacology Reference.*, **2007**.

- <http://dx.doi.org/10.1016/B978-008055232-3.63727-4>
- [66] Beljanski, V. Bevacizumab. *xPharm: The Comprehensive Pharmacology Reference.*, **2007**.  
<http://dx.doi.org/10.1016/B978-008055232-3.63725-0>
- [67] Findlay, V. J.; Scholar, E. Trastuzumab. *xPharm: The Comprehensive Pharmacology Reference.*, **2007**.  
<http://dx.doi.org/10.1016/B978-008055232-3.63738-9>
- [68] *List of Cleared or Approved Companion Diagnostic Devices (In Vitro and Imaging Tools) | FDA.*, n.d.
- [69] Naran, K.; Nundalall, T.; Chetty, S.; Barth, S. Principles of Immunotherapy: Implications for Treatment Strategies in Cancer and Infectious Diseases. *Front. Microbiol.*, **2018**, *9*(December), 3158.  
<http://dx.doi.org/10.3389/fmicb.2018.03158> PMID: 30622524
- [70] Ye, Z.; Qian, Q.; Jin, H.; Qian, Q. Cancer vaccine: learning lessons from immune checkpoint inhibitors. *J. Cancer*, **2018**, *9*(2), 263-268.  
<http://dx.doi.org/10.7150/jca.20059> PMID: 29344272
- [71] Hurley, L. P.; Bridges, C. B.; Harpaz, R.; Allison, M. A.; O' Leary, S. T.; Crane, L. A. Physician Attitudes Toward Adult Vaccines and Other Preventive Practices, United States, 2012. *Public health reports (Washington, D.C. : 1974)*, **2016**, *131*(2), 320-330.  
<http://dx.doi.org/10.1177/003335491613100216>
- [72] Wong, K.K.; Li, W.A.; Mooney, D.J.; Dranoff, G. Advances in Therapeutic Cancer Vaccines. *In Adv Immunol.*, **2016**, Vol. 130, pp. 191-249.
- [73] Higano, C. S.; Small, E. J.; Schellhammer, P.; Yasothan, U.; Gubernick, S.; Kirkpatrick, P.; Kantoff, P. W. Sipuleucel-T. *Nature Reviews Drug Discovery.*, **2010**.  
<http://dx.doi.org/10.1038/nrd3220>
- [74] Kantoff, P.W.; Higano, C.S.; Shore, N.D.; Berger, E.R.; Small, E.J.; Penson, D.F.; Redfern, C.H.; Ferrari, A.C.; Dreicer, R.; Sims, R.B.; Xu, Y.; Frohlich, M.W.; Schellhammer, P.F. IMPACT Study Investigators. Sipuleucel-T immunotherapy for castration-resistant prostate cancer. *N. Engl. J. Med.*, **2010**, *363*(5), 411-422.  
<http://dx.doi.org/10.1056/NEJMoa1001294> PMID: 20818862
- [75] Perica, K.; Varela, J.C.; Oelke, M.; Schneck, J. Adoptive T cell immunotherapy for cancer. *Rambam Maimonides Med. J.*, **2015**, *6*(1)e0004  
<http://dx.doi.org/10.5041/RMMJ.10179> PMID: 25717386
- [76] June, C.H. Adoptive T cell therapy for cancer in the clinic. *J. Clin. Invest.*, **2007**, *117*(6), 1466-1476.  
<http://dx.doi.org/10.1172/JCI32446> PMID: 17549249
- [77] June, C.H. Principles of adoptive T cell cancer therapy. *J. Clin. Invest.*, **2007**, *117*(5), 1204-1212.  
<http://dx.doi.org/10.1172/JCI31446> PMID: 17476350
- [78] Restifo, N.P.; Dudley, M.E.; Rosenberg, S.A. Adoptive immunotherapy for cancer: harnessing the T cell response. *Nat. Rev. Immunol.*, **2012**, *12*(4), 269-281.  
<http://dx.doi.org/10.1038/nri3191> PMID: 22437939
- [79] Stanton, S.E.; Disis, M.L. Clinical significance of tumor-infiltrating lymphocytes in breast cancer. *J. Immunother. Cancer*, **2016**, *4*, 59. <http://dx.doi.org/10.1186/s40425-016-0165-6> PMID: 27777769
- [80] Vilgelm, A.E.; Johnson, D.B.; Richmond, A. Combinatorial approach to cancer immunotherapy: strength in numbers. *J. Leukoc. Biol.*, **2016**, *100*(2), 275-290.  
<http://dx.doi.org/10.1189/jlb.5R10116-013RR> PMID: 27256570
- [81] Pilonis, K.A.; Vanpouille-Box, C.; Demaria, S. Combination of radiotherapy and immune checkpoint inhibitors. *Semin. Radiat. Oncol.*, **2015**, *25*(1), 28-33.  
<http://dx.doi.org/10.1016/j.semradonc.2014.07.004> PMID: 25481263
- [82] Weichselbaum, R.R.; Liang, H.; Deng, L.; Fu, Y.X. Radiotherapy and immunotherapy: a beneficial liaison? *Nat. Rev. Clin. Oncol.*, **2017**, *14*(6), 365-379.  
<http://dx.doi.org/10.1038/nrclinonc.2016.211> PMID: 28094262
- [83] Formenti, S.C.; Demaria, S. Combining radiotherapy and cancer immunotherapy: a paradigm shift. *J. Natl. Cancer Inst.*, **2013**, *105*(4), 256-265.  
<http://dx.doi.org/10.1093/jnci/djs629> PMID: 23291374
- [84] Dunn, J.; Rao, S. Epigenetics and immunotherapy: The current state of play. *Mol. Immunol.*, **2017**, *87*, 227-239.  
<http://dx.doi.org/10.1016/j.molimm.2017.04.012> PMID: 28511092
- [85] Charoentong, P.; Angelova, M.; Efremova, M.; Gallasch, R.; Hackl, H.; Galon, J.; Trajanoski, Z. Bioinformatics for cancer immunology and immunotherapy. *Cancer Immunol. Immunother.*, **2012**, *61*(11), 1885-1903.  
<http://dx.doi.org/10.1007/s00262-012-1354-x> PMID: 22986455
- [86] Sioud, M.; Hansen, M.; Dybwad, A. Profiling the immune responses in patient sera with peptide and cDNA display libraries. *Int. J. Mol. Med.*, **2000**, *6*(2), 123-128.  
<http://dx.doi.org/10.3892/ijmm.6.2.123> PMID: 10891554
- [87] Hanash, S. Disease proteomics. *Nature*, **2003**, *422*(6928), 226-232. <http://dx.doi.org/10.1038/nature01514> PMID: 12634796
- [88] Olsen, L.R.; Campos, B.; Barnkob, M.S.; Winther, O.; Brusica, V.; Andersen, M.H. Bioinformatics for cancer immunotherapy target discovery. *Cancer Immunol. Immunother.*, **2014**, *63*(12), 1235-1249.  
<http://dx.doi.org/10.1007/s00262-014-1627-7> PMID: 25344903
- [89] Lam, H.Y.K.; Pan, C.; Clark, M.J.; Lacroute, P.; Chen, R.; Harakasingh, R.; O'Huallachain, M.; Gerstein, M.B.; Kidd, J.M.; Bustamante, C.D.; Snyder, M. Detecting and annotating genetic variations using the HUGO pipeline. *Nat. Biotechnol.*, **2012**, *30*(3), 226-229.  
<http://dx.doi.org/10.1038/nbt.2134> PMID: 22398614
- [90] Sailani, M. R.; M. Reza Sailani, PhD. n.d.
- [91] Narang, V.; Decraene, J.; Wong, S.Y.; Aiswarya, B.S.; Wasem, A.R.; Leong, S.R.; Gouaillard, A. Systems immunology: a survey of modeling formalisms, applications and simulation tools. *Immunol. Res.*, **2012**, *53*(1-3), 251-265.  
<http://dx.doi.org/10.1007/s12026-012-8305-7> PMID: 22528121
- [92] Margolin, A.A.; Nemenman, I.; Basso, K.; Wiggins, C.; Stolovitzky, G.; Dalla Favera, R.; Califano, A. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics*, **2006**, *7*(Suppl. 1), S7.  
<http://dx.doi.org/10.1186/1471-2105-7-S1-S7> PMID: 16723010
- [93] Wang, K.; Saito, M.; Bisikirska, B.C.; Alvarez, M.J.; Lim, W.K.; Rajbhandari, P.; Shen, Q.; Nemenman, I.; Basso, K.; Margolin, A.A.; Klein, U.; Dalla-Favera, R.; Califano, A. Genome-wide identification of post-translational modulators of transcription factor activity in human B cells. *Nat. Biotechnol.*, **2009**, *27*(9), 829-839.  
<http://dx.doi.org/10.1038/nbt.1563> PMID: 19741643
- [94] Montes, R. A. C.; Coello, G.; González-aguilera, K. L.; Marschmarte, N.; Folter, S.; De, J.; Alvarez-buylla, E. R. **2014**.  
<http://dx.doi.org/10.1186/1471-2229-14-97>
- [95] Bashashati, A.; Haffari, G.; Ding, J.; Ha, G.; Lui, K.; Rosner, J.; Huntsman, D.G.; Caldas, C.; Aparicio, S.A.; Shah, S.P. DriverNet: uncovering the impact of somatic driver mutations on transcriptional networks in cancer. *Genome Biol.*, **2012**, *13*(12), R124.  
<http://dx.doi.org/10.1186/gb-2012-13-12-r124> PMID: 23383675
- [96] Akavia, U.; Litvin, O.; Kim, J.; Mozes, E.; Kotliar, D.; Tzur, Y. *Abstract B70: Conexic: A Bayesian framework to detect drivers and their function uncovers an endosomal signature in melanoma.*, **2009**.  
<http://dx.doi.org/10.1158/0008-5472.fbc09-b70>
- [97] Zhang, T.; Zhang, D. Integrating omics data and protein interaction networks to prioritize driver genes in cancer. *Oncotarget*, **2017**, *8*(35), 58050-58060. <http://dx.doi.org/10.18632/oncotarget.19481> PMID: 28938536
- [98] Kishore, A.; Petrek, M. Next-Generation Sequencing Based HLA Typing: Deciphering Immunogenetic Aspects of Sarcoidosis. *Front. Genet.*, **2018**, *9*(October), 503.  
<http://dx.doi.org/10.3389/fgene.2018.00503> PMID: 30410504
- [99] Xie, C.; Xuan, Z.; Wong, M.; Piper, J.; Long, T.; Kirkness, E. F. *Fast and accurate HLA typing from short-read next-generation sequence data with xHLA*, **2017**, *114*(30), 8059-8064. <http://dx.doi.org/10.1073/pnas.1707945114>
- [100] Gandhi, M.J.; Ferriola, D.; Huang, Y.; Duke, J.L.; Monos, D. Targeted Next-Generation Sequencing for Human Leukocyte Antigen Typing in a Clinical Laboratory: Metrics of Relevance and Considerations for Its Successful Implementation. *Arch. Pathol. Lab. Med.*, **2017**, *141*(6), 806-812.  
<http://dx.doi.org/10.5858/arpa.2016-0537-RA> PMID: 282340



