

# **Diabetes Mellitus caused by mutations in human insulin: Analysis of impaired receptor binding of insulins *Wakayama, Los Angeles and Chicago* using Pharmacoinformatics**

Md Ataul Islam<sup>1</sup>, Sagar Bhayye<sup>2</sup>, Adebayo A. Adeniyi<sup>3</sup>, Mahmoud E.S. Soliman<sup>3</sup>, Tahir S. Pillay<sup>1, 4\*</sup>

<sup>1</sup>*Department of Chemical Pathology, Faculty of Health Sciences, University of Pretoria and National Health Laboratory Service Tshwane Academic Division, Pretoria, South Africa.*

<sup>2</sup>*Department of Chemical Technology, University of Calcutta, 92, A. P. C. Road, Kolkata – 700 009, India.*

<sup>3</sup>*Discipline of Pharmaceutical Sciences, School of Health Sciences, University of KwaZulu-Natal, Durban, 4000, South Africa*

<sup>4</sup>*Division of Chemical Pathology, University of Cape Town, Cape Town, South Africa*

Email IDs:        MA Islam: ataul.islam80@gmail.com  
                      S Bhayye: bhayye.sagar@gmail.com  
                      TS Pillay: tspillay@gmail.com

This work was supported by the University of Pretoria Vice Chancellor's post-doctoral fellowship and National Research Foundation (NRF), South Africa Innovation post-doctoral fellowship schemes.

Correspondence should be addressed to T.S. Pillay, Department of Chemical Pathology, Faculty of Health Sciences, University of Pretoria, Private Bag X323, Arcadia, Pretoria, 0007  
Email: tspillay@gmail.com  
Phone: +27-123192114  
Fax: +27-123283600

## Abstract

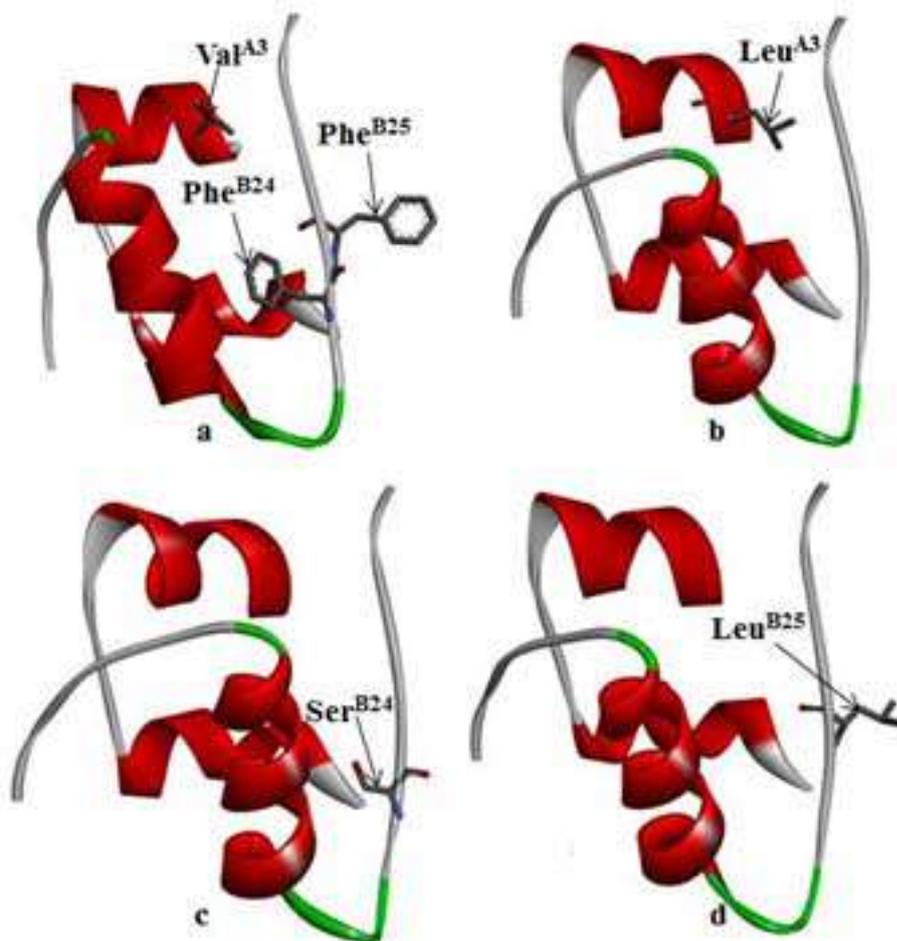
Several naturally occurring mutations in the human insulin gene are associated with diabetes mellitus. The three known mutant molecules, *Wakayama*, *Los Angeles* and *Chicago* were evaluated using molecular docking and molecular dynamics (MD) to analyse mechanisms of deprived binding affinity for insulin receptor (IR). Insulin *Wakayama*, is a variant in which valine at position A3 is substituted by leucine, while in insulin *Los Angeles* and *Chicago*, phenylalanine at position B24 and B25 are replaced by serine and leucine respectively. These mutations show radical changes in binding affinity for insulin receptor. The ZDOCK server was used for molecular docking while AMBER 14 was used for the molecular dynamics study. The published crystal structure of insulin receptor bound to natural insulin was also used for molecular dynamics. The binding interactions and molecular dynamics trajectories clearly explained the critical factors for deprived binding to the insulin receptor. The surface area around position A3 was increased when valine was substituted by leucine, while at position B24 and B25 aromatic amino acid phenylalanine replaced by non-aromatic serine and leucine might be responsible for fewer binding interactions at the binding site of insulin receptor that leads to instability of the complex. In the MD simulation the normal mode analysis (NMA), rmsd trajectories and prediction of fluctuation indicated instability of complexes with mutant insulin in order of insulin native insulin < insulin *Chicago* < insulin *Los Angeles* < insulin *Wakayama* molecules which corresponds to the biological evidence of the differing affinities of the mutant insulins for the IR.

**Key words:** Insulin; insulin receptor; protein-protein docking, ZDOCK, molecular dynamics

## Introduction

Diabetes mellitus (DM) is a heterogeneous group of metabolic diseases characterized by chronic hyperglycaemia resulting from decreased insulin secretion and/or increased insulin resistance. Insulin is a globular protein molecule critical for hormonal regulation of mammalian metabolism and plays a key role in the regulation of blood glucose levels and fat storage. Insulin consists of two chains, A and B with 21 and 30 amino residues respectively constrained by one intra- and two inter-chain disulphide bonds (Adams et al., 1969). The A chain (InsA) contains two  $\alpha$ -helices ranging from amino acid 1 to 8, and 12 to 18, while in case of B chain (InsB) the helix range is amino acids 9 to 19. The cellular activities of insulin are initiated by insulin binding to the cell surface insulin receptor (IR), which is a disulphide-linked ( $\alpha\beta$ )<sub>2</sub> hetero-dimer. The extracellular portion of each  $\alpha\beta$  protomer contains six domains

(L1, CR, L2, FnIII-1, FnIII-2 and FnIII-3) and an insert domain (ID) within FnIII-2. The  $\alpha$ -chain component of the ID is terminated by a segment termed  $\alpha$ CT, spanning residues 704–719. Defective binding of structurally abnormal insulin or its precursor, pro-insulin can result in diabetes mellitus. Mutations in naturally occurring insulin molecules affect insulin activity through alterations in binding to the insulin receptor. There are several naturally occurring mutations in the insulin molecule which have been described in humans and are associated with diabetes mellitus. These include insulin *Wakayama* ( $\text{Val}^{3A} \rightarrow \text{Leu}^{3A}$ ) where valine is mutated to leucine at A3, insulin *Los Angeles* ( $\text{Phe}^{24B} \rightarrow \text{Ser}^{24B}$ ) and insulin *Chicago* ( $\text{Phe}^{25B} \rightarrow \text{Leu}^{25B}$ ) in which phenylalanine is substituted by serine and leucine at B24 and B25 respectively. These substitutions occur within the crevice adjoining the IR binding surface and impair receptor binding by many-fold in affected individuals (Wan et al., 2005; Wollmer et al., 1981). The orientations of amino acids  $\text{Val}^{3A}$ ,  $\text{Leu}^{3A}$ ,  $\text{Phe}^{24B}$ ,  $\text{Ser}^{24B}$ ,  $\text{Phe}^{25B}$  and  $\text{Leu}^{25B}$  in natural and mutant insulin molecules are depicted in Figure 1.



**Figure 1:** Orientation of amino acids at position A3, B24 and B25 in native/natural and mutant insulin molecules. a) Natural insulin, b) Insulin *Wakayama*, c) Insulin *Los Angeles* and d) *Insulin Chicago*.

These mutations on the natural insulin molecule especially at A3, B24 and B25 reduce the binding affinity for IR many-fold(Hayashi et al., 2011). Moreover there are substantial differences in the biological activity between individual mutant molecules. There has been no published pharmacoinformatics analysis of this reduction in binding affinity and the aim of this study was to correlate the biological activity with the structural analysis.

It has been reported that Val<sup>A3</sup>, present within the N-terminal  $\alpha$ -helix of A-chain plays an important role in the insulin fold and is recognised as an essential element(Hua et al., 1996; Nakagawa & Tager, 1992; Weiss et al., 2000). The side chains of Val<sup>A3</sup> pack within the non-polar crevice between InsA and InsB which adjoin the classical receptor binding surface (E. N. Baker et al., 1988; Blundell et al., 1971; Nakagawa & Tager, 1992; Pullen et al., 1976). The crevice is surrounded by sequence-continuous side chains of Ile<sup>A2</sup> and Leu<sup>B11</sup>, the internal disulphide bridge between Cys<sup>A6</sup> and Cys<sup>A11</sup>, Tyr<sup>A19</sup> and Leu<sup>B11</sup> in the hydrophobic core and two amino residues of InsB such as Tyr<sup>B26</sup> and Pro<sup>B28</sup> which project from an overlying B-chain. It was expected that substitution of Val<sup>A3</sup> by Leu<sup>A3</sup> would not change conservation and stability because leucine demonstrates a greater intrinsic  $\alpha$ -helical propensity than valine(Lyu, Liff, Marky, & Kallenbach, 1990; O'Neil & DeGrado, 1990; Padmanabhan, Marqusee, Ridgeway, Laue, & Baldwin, 1990). However, it was observed that such substitutions reduced activity by 500-fold relative to that of normal insulin molecule (Nakagawa & Tager, 1992). Sequences of natural insulin are remarkable for an invariant phenylalanine present at positions B24 and B25 and for the conservation of phenylalanine and tyrosine at positions B25 and B26 respectively (Conlon, 2001). Phe<sup>B24</sup>, Phe<sup>B25</sup> and Tyr<sup>B26</sup> amino residues play crucial role for the stability, self-assembly and activity of the insulin molecule(Mirmira, Nakagawa, & Tager, 1991; Mirmira & Tager, 1989; Zakova et al., 2013). In 1969, it was shown that (Adams et al., 1969; E. N. Baker et al., 1988) on the basis of the crystal structure of the zinc insulin hexamer that residues at the positions B24 to B26 participate in an aromatic rich dimer interface and this interface occurs three times in the hexamers. Owing to substitution of aromatic phenylalanine at positions B24 and B25 by non-aromatic serine and leucine respectively in insulin *Los Angeles* and *Chicago*, binding affinity towards IR is reduced significantly(Weitzel, Bauer, & Eisele, 1978; Wollmer et al., 1981).

Apart from crystallography, molecular docking approaches which takes two or more structures as input and predicts the structure of their complex, are useful for analysing protein-protein interactions. The 3D complex of protein-protein interaction can be used to estimate the effect of mutations(Benedix, Becker, de Groot, Caflisch, & Bockmann, 2009;

Kamisetty, Ramanathan, Bailey-Kellogg, & Langmead, 2011; Moretti et al., 2013), and thus for protein design(Ahmad et al., 2012; D. Baker, 2006; Chen & Keating, 2012; Kortemme & Baker, 2004) and determining the functional consequences of mutations associated with disease (Nisius, Sha, & Gohlke, 2012). Aspects of the conformational variability of a protein molecule can be modelled using the molecular dynamics (MD) simulation method. In MD simulation the detailed time evolution of a system of particles may be evaluated by integrating equation of motion for the particles on their potential energy surface. The MD method may provide effective modelling of protein-protein interactions at a reasonable computing cost. In order to understand the mechanism of deprived binding affinity of the insulin molecules for the IR the present study evaluated protein-protein molecular docking by considering mutant insulin molecules as ligands and recently crystallized IR as receptor. Furthermore, MD simulations were performed to analyse the stability of complexes and exploration to provide insight into the binding interactions between mutant insulins and binding residues of IR.

## **Materials and Methods**

In order to understand impaired binding interactions of mutant insulin molecules to the IR, protein-protein docking and molecular dynamics study were performed using the online ZDOCK server(Pierce et al., 2014) and Amber 14 (<http://ambermd.org/#Amber14>) respectively. The crystal complex of the insulin receptor with insulin molecule (PDB ID: 3W14)(Menting et al., 2013), and natural insulin (PDB: 3IR0) were obtained from RCSB Protein Data Bank (PDB)(Parasuraman, 2012). The DeepView(Guex & Peitsch, 1997)program was used to mutate the residues on natural insulin and the spatial orientation was maintained. The online server contact map analysis (CMA)(Sobolev et al., 2005) was used to analyse the binding interactions between insulin and IR.

### ***Protein-protein docking***

In the current research ZDOCK server(Pierce et al., 2014) was used for the protein-protein docking study. The mutant insulin molecules and insulin receptor were considered as ligand and receptor molecule respectively. The ZDOCK is the rigid-body protein-protein docking program which uses the Fast Fourier Transform algorithm(Chu & Alan, 2000) that enables an efficient global docking search on a 3D grid. This program ranks top hundred probable predictions by a combination of shape complementarily, electrostatic and statistical potential

terms for scoring. It is reported that ZDOCK achieves high predictive accuracy with more than 70% success in the top thousand predictions(Pierce, Hourai, & Weng, 2011). Jobs of the server are run on a dedicated computer server at the University of Massachusetts Medical School(Pierce et al., 2014) which provides interfaces of input two protein structures and email.

### ***Molecular dynamics***

In order to check stability and binding orientation MD simulations were performed for crystal complex of insulin bound to the IR, and best docked complexes between mutant insulins and IR. The MD was carried out using GPU-based PMEMD programme and the protein structures were parameterised and represented with amber force field (leaprc.ff99SB) as implemented in Amber 14(Gotz et al., 2012). The generalized Born (GB) solvation model was used for the implicit solvent with the igb set to 5 which has been shown to agree better with the Poisson-Boltzmann treatment in calculating the electrostatic part of the solvation free energy(Feig et al., 2004). In the production MD of the GB simulations, the Langevin thermostat was used with a collision frequency of 2.0 ps<sup>-1</sup> and used an improved GB-Neck of the GB-OBC model that correspond to mbondi2 radii(Onufriev, Bashford, & Case, 2004) as our starting radii set. It has been shown that GB-OBC worked quite well with mbondi2 radii(Nguyen, Roe, & Simmerling, 2013). It is also reported that slight increases in structural stability of the complexes can be achieved in the performance of GB model (HCT or OBC GB models) by setting the atomic intrinsic radii used to define the dielectric boundary of hydrogens bound to nitrogen, H(N) from their Bondi radii of 1.2 Å to 1.3 Å (forming the mbondi2 radius set)(Mongan, Simmerling, McCammon, Case, & Onufriev, 2007). Initial minimization of all the systems was carried out for 1000 steps of steepest descent followed by 1000 steps of heating up of the system to 300.00 K and equilibration. The MD production was carried out for at least 10ns for each of the system. The frame were recorded at every 200 steps of simulation. The analysis of the conformational changes in the trajectory of the protein was analysed using bio3d(Grant, Rodrigues, ElSawy, McCammon, & Caves, 2006) and the secondary structures were predicted using DSSP(Kabsch & Sander, 1983). The visualization of the results was done using VMD(Humphrey, Dalke, & Schulten, 1996), chimera(Pettersen et al., 2004) and Pymol Molecular Graphics System, Version 1.7.2 Schrödinger, LLC.

## Results and discussion

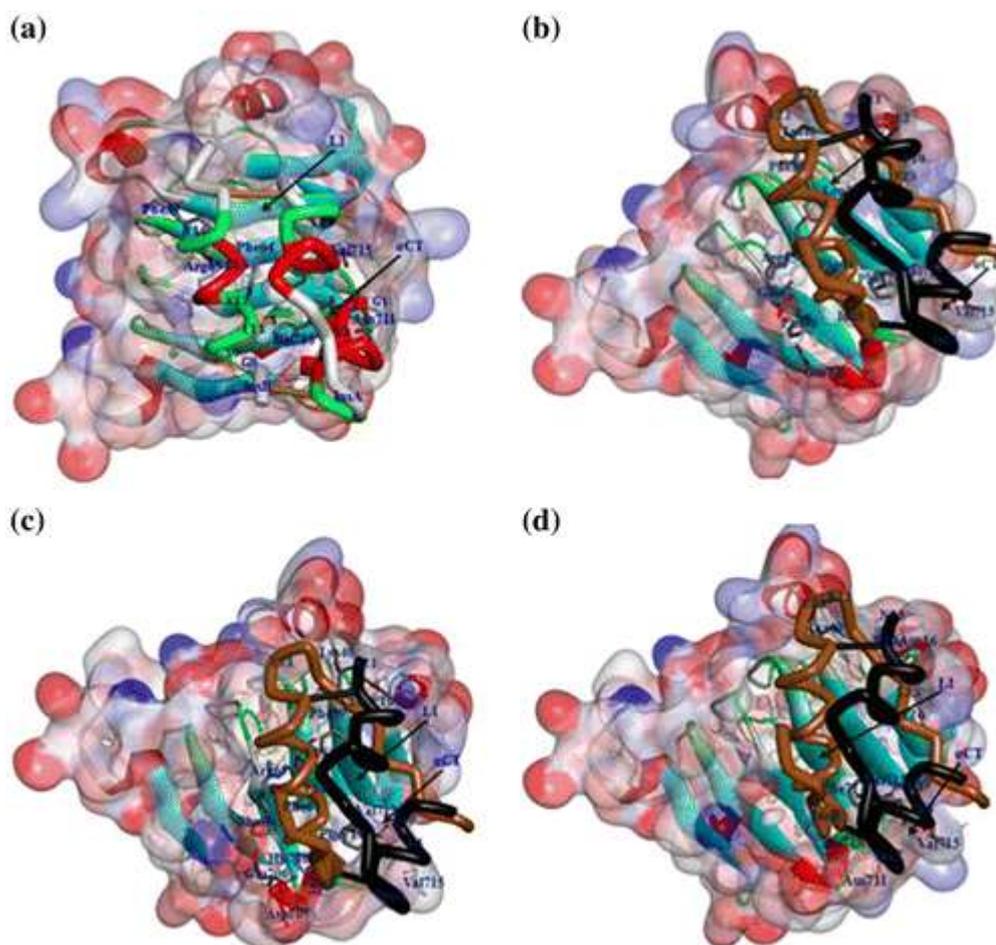
### Results

#### *Crystal complex of IR and native insulin*

The crystal structure (PDB ID: 3W14) of IR bound with native insulin is depicted in Figure 2a, which explains that helix of InsB (B7-B19) engages with the C-terminal end of L1 domain but InsA has no interaction with domain L1. The domain  $\alpha$ CT interacts with both InsA and InsB of the insulin molecule. It was found that two important amino residues, His710 and Phe714 present in the  $\alpha$ CT domain were found to be critical for interaction with both chains of the insulin molecule. His710 inserts into a pocket surrounded by insulin residues Val<sup>A3</sup>, Gly<sup>B8</sup>, Ser<sup>B9</sup> and Val<sup>B12</sup>, while Phe714 occupies a hydrophobic crevice formed by amino residues Gly<sup>A1</sup>, Ile<sup>A2</sup>, Tyr<sup>A19</sup>, Leu<sup>B11</sup>, Val<sup>B12</sup> and Leu<sup>B15</sup> (Ward, et al., 2013). Another amino acid, Asn711 was directed towards Gly<sup>A1</sup>, Val<sup>A3</sup> and Glu<sup>A4</sup>. Two amino residues of InsB such as Val<sup>B12</sup> and Tyr<sup>B16</sup> were found to be critical to interact with L1 domain. Val<sup>B12</sup> was connected to Phe39, Phe64 and Arg65, while Tyr<sup>B16</sup> clashed with Phe39 (Ward et al., 2013). It was observed that  $\alpha$ CT domain is an important domain for interactions with both chains of native insulin molecule.

#### *Molecular docking*

The molecular docking study was performed with the mutant insulin molecules, *Wakayama*, *Los Angeles* and *Chicago* in the receptor cavity of IR. The native insulin molecule bound in the IR in crystal structure (PDB ID: 3W14) was deleted to create the space for the variant insulin molecules. In order to remove any structural strain or conformational changes the mutated insulin molecules were considered for 1ns molecular dynamics simulation analysis. The rmsd vs time along with number of frame vs total, potential and kinetic energies are given in Figure S1 (supplementary file). The simulated conformers were aligned with respective insulin molecule to calculate rmsd values. The rmsd values for insulin *Wakayama*, insulin *Los Angeles* and insulin *Chicago* were found to be 1.001Å, 1.043Å and 1.005Å respectively. The aligned structures are portrayed in Table S2 in supplementary material. The low rmsd values clearly explained that docking using simulated and non-simulated mutant insulin molecules at the IR will give comparable binding interactions. Further both simulated and non-simulated mutant insulin molecules were considered for molecular docking study and were found to have similar binding interactions. The best docked pose of each docking study is given in Figure 2b, c, d.



**Figure 2:** Binding interactions between insulin molecules and IR: a) Crystal structure (PDB ID: 3W14); b) insulin *Wakayama*; c) insulin *Los Angeles*; and d) insulin *Chicago*.

### *Insulin Wakayama:*

The best docked pose insulin *Wakayama* molecule (Figure 2b) explains that Tyr<sup>A19</sup> and Asn<sup>A21</sup> are able to create connections with L1 domain of IR. Tyr<sup>A19</sup> is connected to Arg65, while Asn<sup>A21</sup> is surrounded by Leu37, Phe39, Phe64, Phe96 and Glu97. Both InsA and InsB of the insulin *Wakayama* interact with  $\alpha$ CT domain of IR. Asn<sup>A21</sup>, Glu<sup>B21</sup>, Arg<sup>B22</sup> and Phe<sup>B25</sup> were found to be important for interactions with Glu706, Val713, Phe714 and Val715. InsB of the variant insulin molecule also bound to L1 domain through Phe39, Lys40, Arg65 and Glu97. Interestingly the Leu<sup>A3</sup> in the mutant insulin molecule was not found to create any connection with IR. With the exception of the length of the main chain of valine and leucine the structures of both amino acids are almost similar. However, the orientation and position of the methyl group present in the both amino acids are different and this might be a possible explanation for no interaction with IR via Leu<sup>A3</sup> in the variant insulin molecule.

### *Insulin Los Angeles:*

The binding interactions between insulin *Los Angeles* and IR is depicted in the Figure 2c. The complex suggests that both InsA and InsB of variant insulin *Los Angeles* molecule interact with domains L1 and  $\alpha$ CT. The best docked pose explains that Tyr<sup>A19</sup>, Cys<sup>A20</sup>, Asn<sup>A21</sup>, Ser<sup>B9</sup>, Val<sup>B12</sup>, Leu<sup>B15</sup>, Glu<sup>B21</sup>, Ser<sup>B24</sup> and Phe<sup>B25</sup> are crucial amino residues for connecting to domain L1 of IR through Phe39, Lys40, Phe64, Arg65 and Phe96. Two important amino residues Phe714 and Val715 belonging to  $\alpha$ CT domain successfully interact with Val<sup>A3</sup> and Glu<sup>A4</sup> respectively. Glu706, Asp707, His710 and Val713 along with Phe714 and Val715 of the domain  $\alpha$ CT were found to be critical to establish binding interactions with InsB of the insulin *Los Angeles* through Cys<sup>B7</sup>, Ser<sup>B9</sup>, His<sup>B10</sup>, Glu<sup>B13</sup>, Tyr<sup>B26</sup> and Thr<sup>B30</sup>. It is important to note that Ser<sup>B24</sup> in the variant insulin *Los Angeles* only connects with L1 domain but not with  $\alpha$ CT domain. Owing to reduction of surface volume at positions 24 of InsB (Serine is smaller in size compared to phenylalanine) as well as lack of aromatic ring in Ser<sup>B24</sup>, there was a lack of interaction with IR  $\alpha$ CT domain.

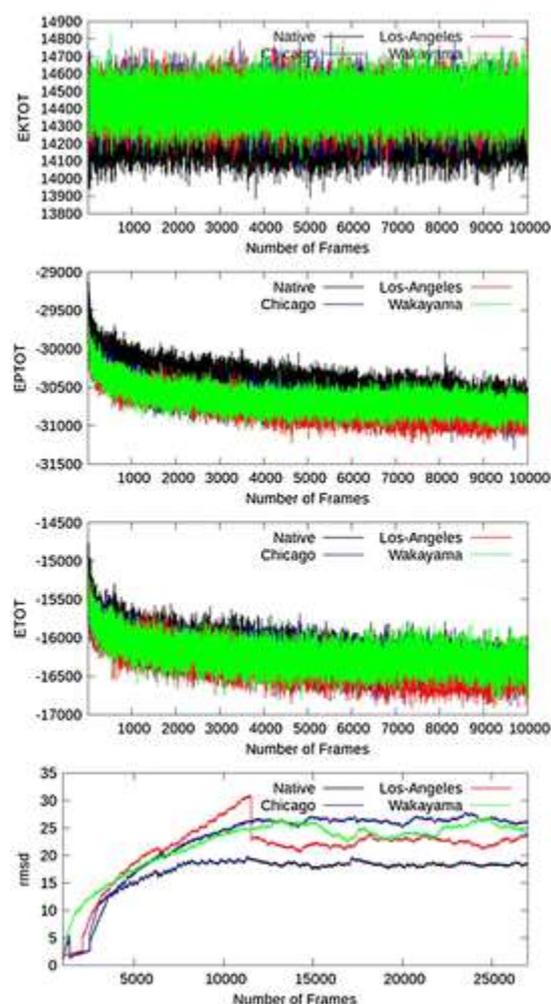
### *Insulin Chicago:*

Insulin *Chicago* was considered as ligand and IR as receptor for molecular docking study. The best docked complex is shown in the Figure 2d. The complex shows that Asn16 and Lys40 are important amino residues of L1 domain of IR for binding interactions with Tyr<sup>A19</sup>, Cys<sup>A20</sup> and Asn<sup>A21</sup>, while Val<sup>A3</sup>, Glu<sup>A4</sup>, Cys<sup>A7</sup> and Thr<sup>A8</sup> are connected to  $\alpha$ CT through His710, Asn711, Val713, Phe714 and Val715. InsB of the variant insulin *Chicago* was unable to establish any binding interaction with L1 domain. However, two amino acids belonging to InsB, Glu<sup>B21</sup> and Arg<sup>B22</sup> clashed with  $\alpha$ CT domain via Val713, Phe714 and Val715. It has been reported that Phe<sup>B24</sup>, Phe<sup>B25</sup> and Tyr<sup>B26</sup> are important amino residues for interaction with IR (Mirmira et al., 1991). However the docking results of insulin *Chicago* shows that none of the three amino acids are involved in binding interactions. The possible reason may be lack of aromatic ring in Leu<sup>B25</sup> and variation on orientation of the functional groups in the insulin *Chicago* compared to native insulin.

### *Molecular dynamics*

Molecular dynamics simulations are likely to play an increasingly important role in the development of novel pharmacological therapeutics. Thus it can be expected that MD simulation would provide more reliable structural information on the mutations in insulin. In order to compare the dynamic stability of the best docked poses of insulin molecules inside

the IR the MD simulations were carried out based on protocols described in the materials and methods section to predict the impact of mutation on insulin molecules. MD gives information on the stability of binding interactions between the ligand and receptor cavity. The published crystal structure of IR, bound with native insulin (PDB ID: 3W14) and best docked complexes of IR with mutant insulin molecules were taken into consideration for molecular dynamics study.

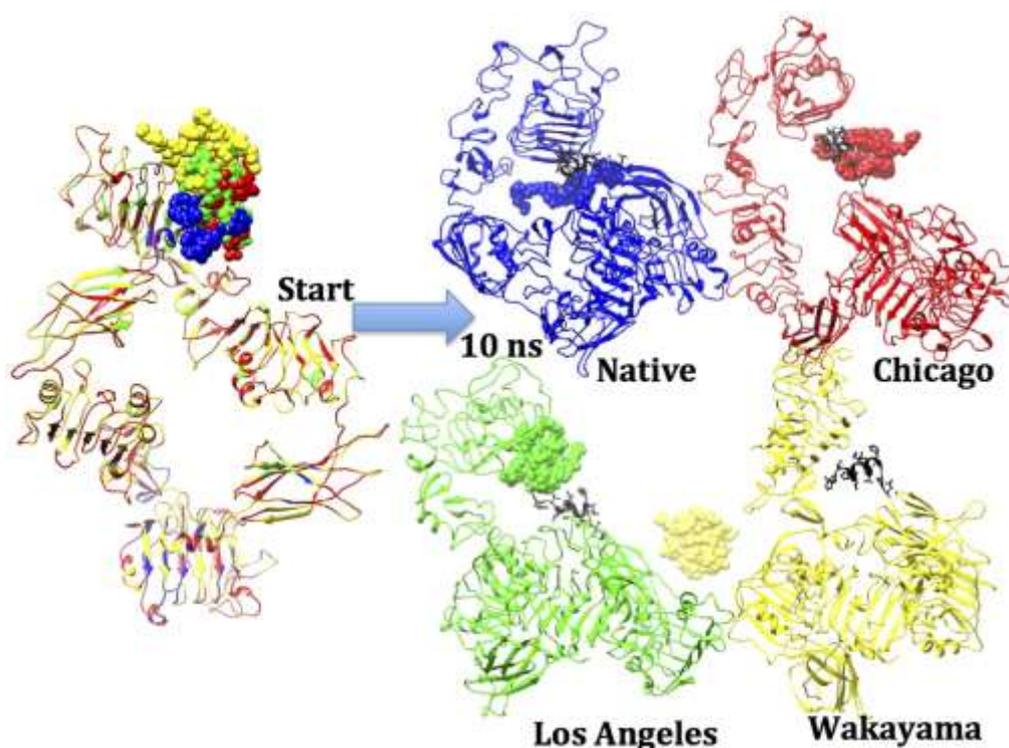


**Figure 3:** The total kinetic (EKTOT), total potential (EPTOT), total energy (ETOT) of the complexes during the simulation and the rmsd of the displacement of the backbone structure of the complexes.

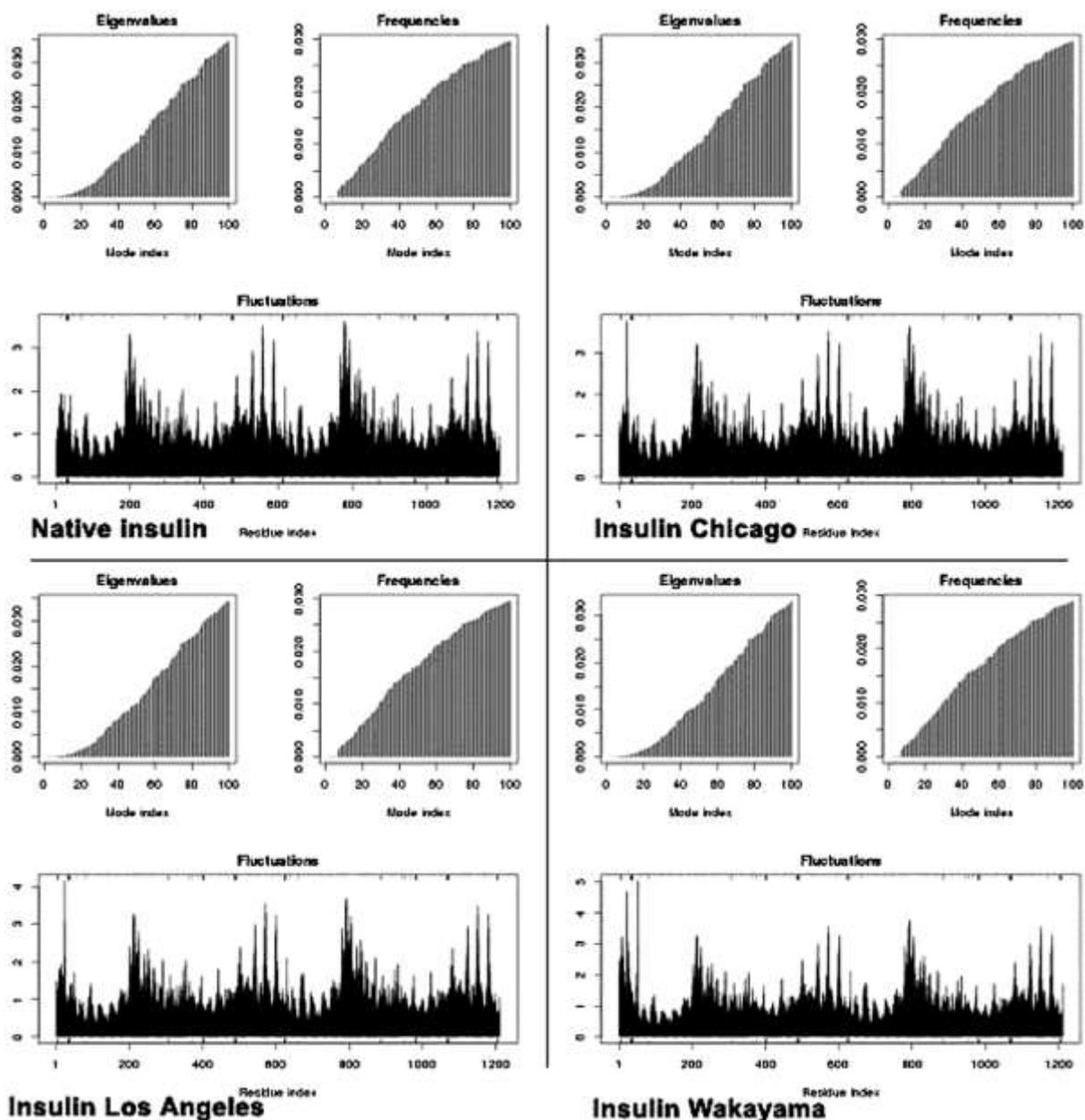
The best docked pose of each insulin *Wakayama*, *Los Angeles* and *Chicago* in the IR along with crystal structure of IR bound with native insulin were considered for MD study. The production of the molecular dynamics of the complexes were at least for 10ns for each of the complexes as shown in the energy plots of Figure 3. The energy plot is shown from 7000ps to the last ps in order to show clearly the fluctuation in the energy from at the points of energy equilibrium. The geometry stability was obtained from 12000 steps of the simulation as

shown by the rmsd plot of Figure 3. The rmsd of the complexes from their respective initial geometries during the trajectory is in the order of complexes of insulin *Chicago*> insulin *Wakayama*>insulin *Los Angeles*> native insulin.

Further trajectory analysis of the MD was done using the frames from 10000 steps (4ns) to the last step of each of the complexes. The sample of the geometries of the insulin complexes of native insulin, insulin *Chicago*, insulin *Los Angeles* and insulin *Wakayama* with the IR at the first step of MD and at the 10 ns of the MD as shown in Figure 4. The correlation of the common IR part of the complexes with that of the native insulin before the MD was observed to be 0.098, 0.070 and 0.062Å while at the 10ns of the MD the correlation changes to 25.241, 21.964 and 20.087Å respectively. Even though the change in the rmsd of the IR of the insulin *Wakayama* is lower than that of insulin *Chicago* and insulin *Los Angeles*, its insulin complex was found to be unstable during the MD simulation. At 10ns of the simulation, the insulin chains of *Wakayama* have completely dissociated from the IR (measured to a distance of 95.036 Å from closest atom of other aligned insulin complexes as shown in Figure 4) but those of native insulin, insulin *Chicago* and insulin *Los Angeles* still maintain a stable interaction with the insulin receptor.



**Figure 4:** The trajectory of the complexes at step one before the md and at time step of 10 ns during the md. The complexes as distinguished with colour: NAT (blue), CHI (red), LOS (green) and WAK (yellow).



**Figure 5:** Summary plot of NMA of the complexes with the schematic representation of the secondary structure added to the top and bottom margins of the fluctuation plot (helices black and strands gray).

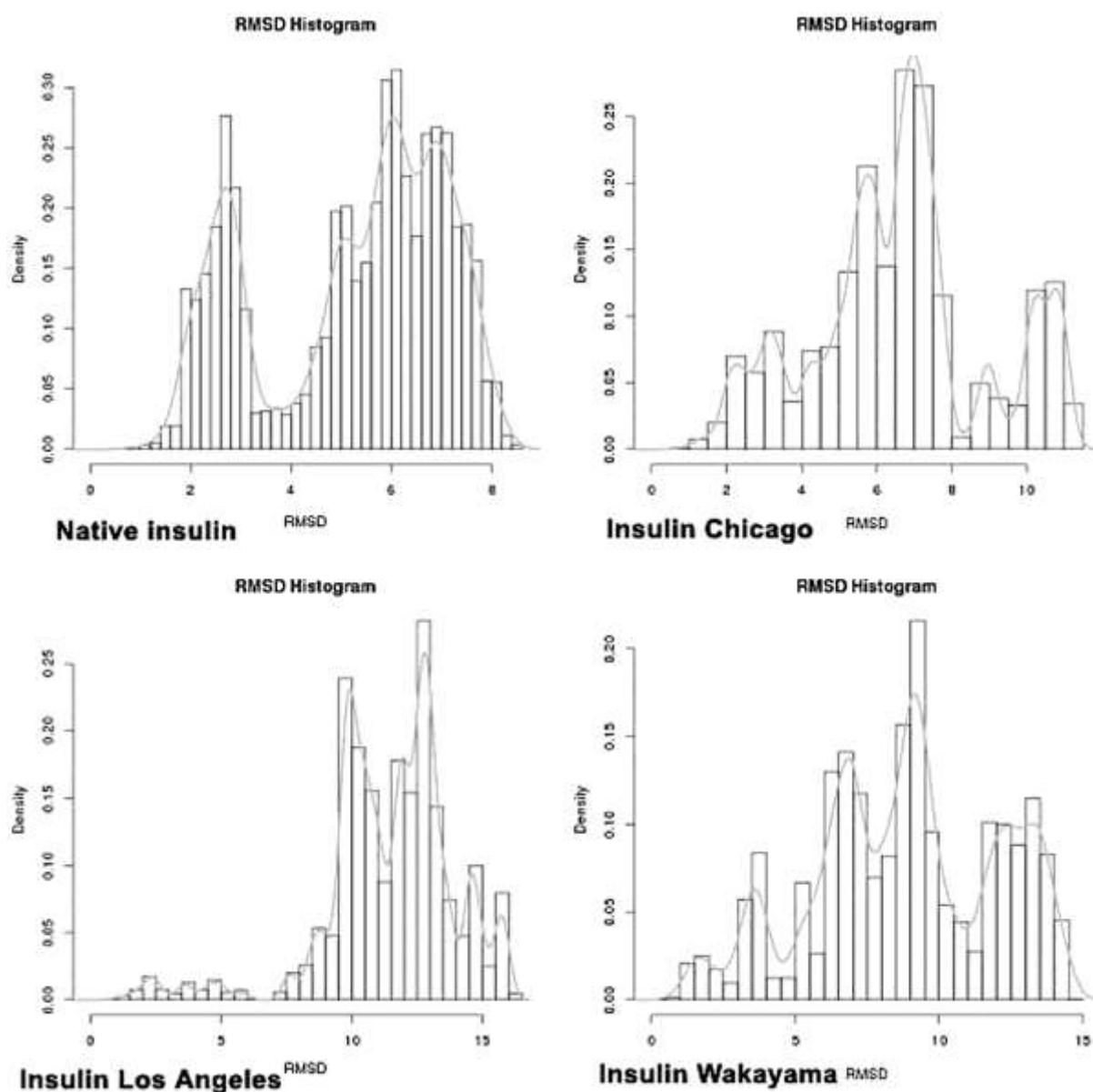
The geometrical analysis of the complexes before the trajectory also show some important distinguishing features in the complexes. The normal mode analysis (NMA) of the complexes shows that insulin *Los Angeles* has similar directions of vector fields with that of complex of native insulin while insulin *Chicago* and insulin *Wakayama* are in opposite directions. Using the NMA simulation techniques to probe large-scale motions in complexes, the highest predicted functional motions in complexes are shown to take place in the mutated region of insulin *Wakayama* (Figure 5). Generally in all the complexes, the larger fluctuations were

predicted to occur at the loop regions in the fluctuation plot of Figure 5). The prediction of fluctuation shows a gradual increase in the fluctuation of the insulin part of the complexes in the order of native insulin < insulin *Chicago* < insulin *Los Angeles* < insulin *Wakayama*. Based on the NMA prediction for insulin *Wakayama*, it would be expected in the course of dynamics to have high fluctuation in the insulin region which we eventually observed to be true.

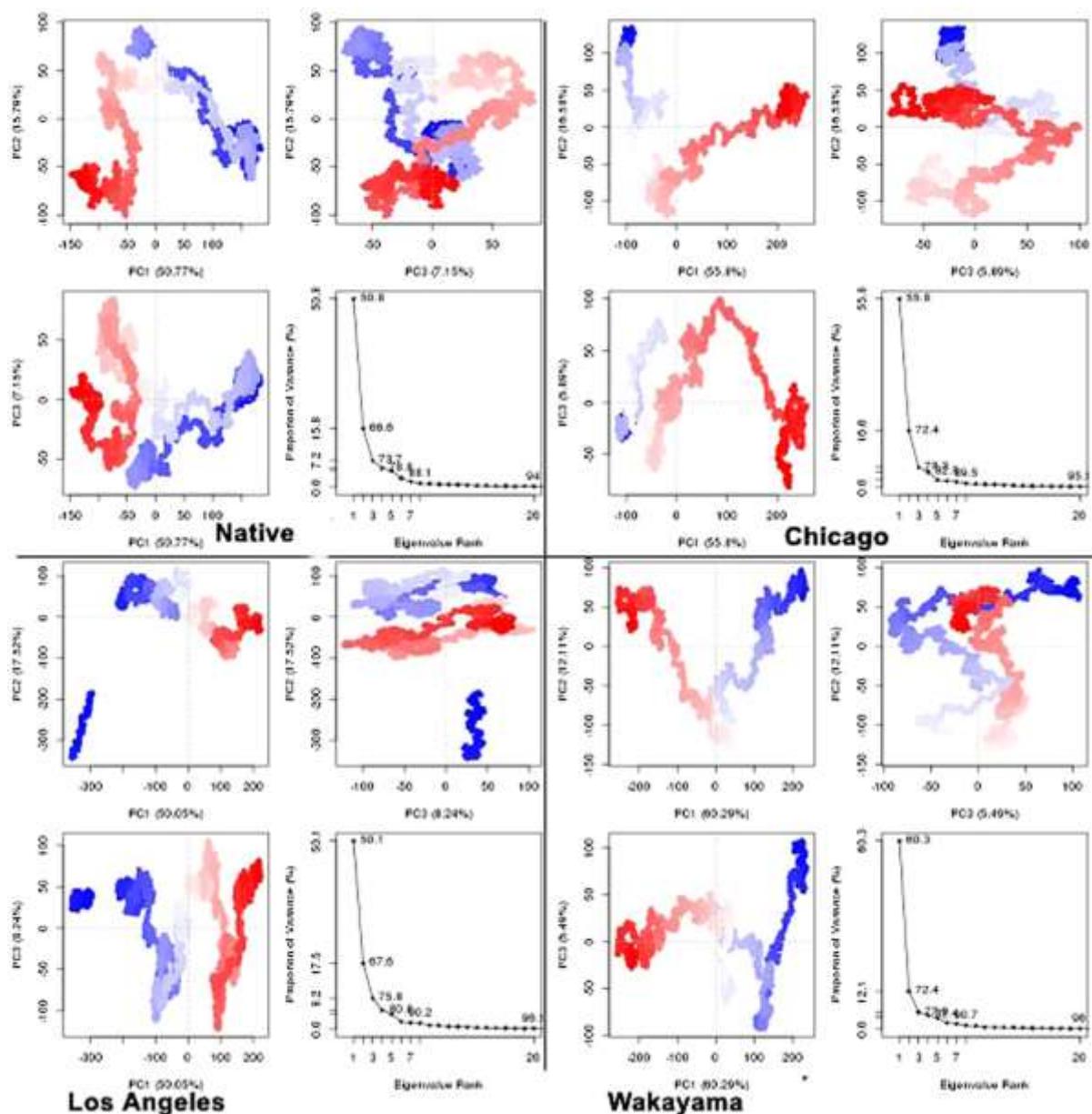
### *Conformational changes in the complexes*

The rmsd of the sampled conformation during the molecular dynamics of complexes with native insulin, insulin *Chicago*, insulin *Los Angeles* and insulin *Wakayama* were found to be predominantly around 6, 7, 13 and 9 respectively from the starting geometries as shown in the histogram of distribution plots in Figure 6. In order to examine the relationship between different conformations sampled during the trajectory, the principal component analysis (PCA) of the trajectory was carried out. The conformational differences are described by the axes of maximal variance of the distribution of structures using principal components (orthogonal eigenvectors). The 3 dimensions (i.e. first three PC) sufficiently capture 73.1, 78.2, 75.8 and 77.0 percent of the total variance in molecular dynamics trajectory of the complexes of native insulin, insulin *Chicago*, insulin *Los Angeles* and insulin *Wakayama* respectively as shown in Figure 7. There are periodic jumps between the different conformers of each complex throughout the trajectory has clearly shown by the continuous colour scale of the PC from blue to white to red as shown in Figure 7. In order to focus on the relationships between individual structures in terms of their major structural displacements, further these distinct conformers were highlighted by clustering structures in PC space as shown in Figure 8. From the results of the clustering, there are two major groups of conformation for each of the complexes. Figure 8 shows the distinct groups of conformations along the PC1 plane for complexes with native insulin (the larger grouping centred around -50 and a second grouping at +150), insulin *Chicago* (the larger grouping centred around -50 and a second grouping at +230), insulin *Los Angeles* (a very small grouping at -330 and a very large grouping centred around +100) and insulin *Wakayama* (one grouping centred around -200 and a longer grouping at +40). The contribution of the residual fluctuation to the conformational changes along the PC1 and PC2 are shown for each complex in Figure 9. The highest contribution of the insulin residues to the observed conformational variances observed along both PC1 and PC2 is found in the complex of insulin *Wakayama* and the next

significant contribution is observed in the complex of native insulin. The figure also shows the groups of the IR residues that significantly contribute to the conformational changes in the complexes of native insulin (200-250, 400-500), insulin *Chicago* (500-600, 1100-1200), LOS (200-300, 1100-1200) and insulin *Wakayama* (300-400, 750-850).



**Figure 6:** The RMSD values showing the sampled conformations during the molecular dynamics of each of the complexes from its starting structure.



**Figure 7:** PCA results of the trajectory of the complexes with instantaneous conformations (i.e. trajectory frames) coloured from blue to red in order of time.

### Discussion

Naturally occurring mutations of the human insulin molecule are associated with diabetes mellitus and this has given insight into structure-function relationships. The present research was focused on three important mutant insulin molecules, *Wakayama*, *Los Angeles* and *Chicago* which have very low binding affinity for IR compared to the native insulin molecule. In the case of insulin *Wakayama*, the molecular docking results suggested that the amino residue at position A3 was unable to connect L1 while valine at the same position in native insulin molecule successfully formed interactions with  $\alpha$ CT domain of IR (Figure 2).

MD study of insulin *Wakayama* and docked complex between insulin *Wakayama* and IR indicated instability of the molecule. RMSD and RMSF of both insulin *Wakayama* and docked complex explained that due to mutation at A3 the stability was disturbed. The branched-chain aliphatic residues changed within the non-polar crevice as valine was replaced by leucine. Such alteration is easy to tolerate in the protein molecule and mutant insulin molecule is almost similar in structure to the native insulin molecule. Observation of several groups (Derewenda et al., 1991; Hua, Shoelson, Kochoyan, & Weiss, 1991; Ludvigsen, Olsen, & Kaarsholm, 1998; Wan, Xu, Chu, Katsoyannis, & Weiss, 2003; Xu et al., 2002) and computational output of the current work explained that crevice related to A3 opens due to binding with IR which enables docking of Val<sup>A3</sup> within non-polar pocket of receptor. Due to steric hindrance between Leu<sup>A3</sup> and IR, the stability of the complex is reduced and this leads to impaired binding affinity of insulin *Wakayama* towards IR. In the MD study of insulin *Los Angeles* and *Chicago* it was observed that the average RMSD and RMSF of docked complexes mutant insulin molecules are higher than crystal complex of native insulin. Moreover the molecular docking results suggested that InsB of insulin *Chicago* failed to connect with L1 domain of the IR. Amino residue at B24 in *insulin Los Angeles* was unable to connect to the domain  $\alpha$ CT. It is reported that the aromatic side chain is favourable for interactions with IR. In both mutant insulin *Los Angeles* and *Chicago*, aromatic amino residue phenylalanine substituted by non-aromatic serine and leucine respectively leads to changes in binding interaction with IR. The charged or polar side chains reduce the hormone binding (Pandeyarajan et al., 2014) and in insulin *Los Angeles*, the number of polar side chains are increased and this may reduce the binding affinity towards IR.

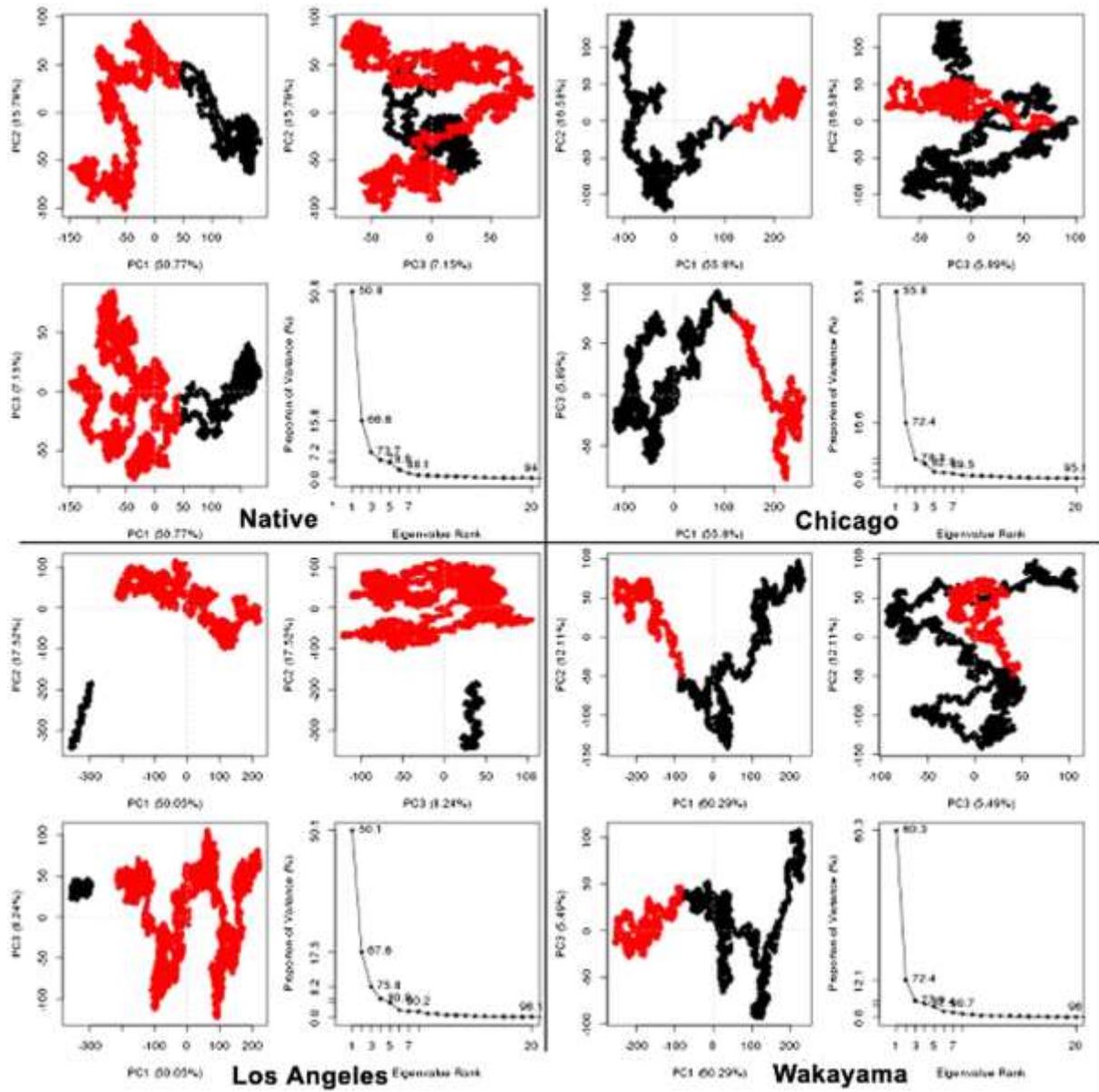
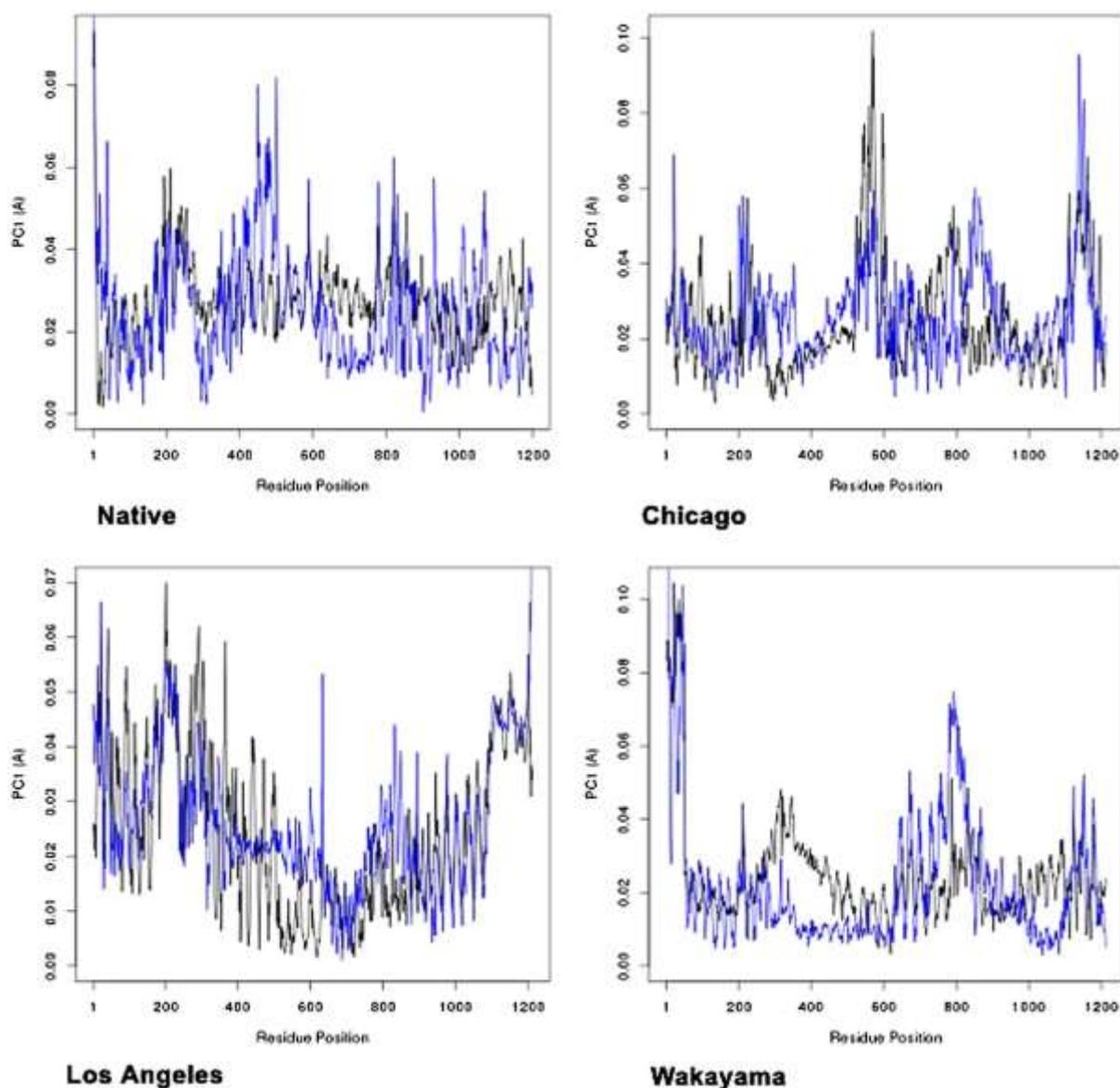


Figure 8: Clustering of the PCA results of the trajectory of the complexes.



**Figure 9:** Residue-wise contributions to the PC1 (black) and PC2 (blue).

## Conclusion

In this study we examined the binding of mutant insulins *Wakayama*, *Los Angeles* and *Chicago* using pharmacoinformatics analysis of the reduced binding affinity to the IR. The molecular docking study of all three mutant insulin was carried out to analyse the binding interaction with the binding amino residues of the IR. The study explained that due to small size and less surface area of the replaced amino acids the binding interactions were reduced and this altered the binding affinity towards IR. In order to confirm the possible reason for this, molecular dynamics study also performed. The rmsd of the protein backbone of the complexes during the trajectory were in the order of insulin *Chicago*>insulin *Wakayama*>insulin *Los Angeles*>native insulin. Aligning the trajectory of complexes with

insulin *Chicago*, insulin *Los Angeles* and insulin *Wakayama* with the native insulin at 10ns of simulation, insulin *Chicago* has the highest rmsd followed by insulin *Los Angeles* while insulin *Wakayama* is least. It was also observed that, the insulin part of native insulin, *Chicago* and *Los Angeles* still maintain the stable interaction with IR but a complete dissociation of the insulin from IR was observed for insulin *Wakayama* during the MD simulation. The PC analysis also shows that the highest contribution of the insulin residues to the two distinct conformational variances observed along both PC1 and PC2 is found in the complex of insulin *Wakayama* and the next significant contribution is observed in the complex with native insulin. Therefore the trajectories of the MD successfully explained that possible reasons may be the smaller size and variation of the size chains of the amino acids which affects the binding interactions as well as the stability of the complex. The outcome of the current research can assist in identifying new possible insulin analogues for the effective treatment of diabetes.

## Acknowledgement

MA Islam and TS Pillay were funded by the University of Pretoria Vice Chancellor's post-doctoral fellowship and National Research Foundation (NRF), South Africa Innovation post-doctoral fellowship schemes.

## References

- Adams, M. J., Blundell, T. L., Dodson, E. J., Dodson, G. G., Vijayan, M., Baker, E. N., . . . Sheat, S. (1969). Structure of Rhombohedral 2 Zinc Insulin Crystals. *Nature*, *224*, 491 - 495
- Ahmad, J. N., Li, J., Biedermannova, L., Kuchar, M., Sipova, H., Semeradtova, A., . . . Maly, P. (2012). Novel high-affinity binders of human interferon gamma derived from albumin-binding domain of protein G. *Proteins*, *80*(3), 774-789. doi: 10.1002/prot.23234
- Baker, D. (2006). Prediction and design of macromolecular structures and interactions. *Philos Trans R Soc Lond B Biol Sci*, *361*(1467), 459-463. doi: 10.1098/rstb.2005.1803
- Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., . . . et al. (1988). The structure of 2Zn pig insulin crystals at 1.5 Å resolution. *Philos Trans R Soc Lond B Biol Sci*, *319*(1195), 369-456
- Benedix, A., Becker, C. M., de Groot, B. L., Cafilisch, A., & Bockmann, R. A. (2009). Predicting free energy changes using structural ensembles. *Nat Methods*, *6*(1), 3-4. doi: 10.1038/nmeth0109-3
- Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. C., . . . Vijayan, M. (1971). Atomic positions in rhombohedral 2-zinc insulin crystals. *Nature*, *231*(5304), 506-511

- Chen, T. S., & Keating, A. E. (2012). Designing specific protein-protein interactions using computation, experimental library screening, or integrated methods. *Protein Sci*, 21(7), 949-963. doi: 10.1002/pro.2096
- Chu, E., & Alan, G. (2000). *Inside the FFT Black Box: Serial and Parallel Fast Fourier Transform Algorithms*: CRC Press.
- Conlon, J. M. (2001). Evolution of the insulin molecule: insights into structure-activity and phylogenetic relationships. *Peptides*, 22(7), 1183-1193
- Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X., & Markussen, J. (1991). X-ray analysis of the single chain B29-A1 peptide-linked insulin molecule. A completely inactive analogue. *J Mol Biol*, 220(2), 425-433
- Feig, M., Onufriev, A., Lee, M. S., Im, W., Case, D. A., & Brooks, C. L., 3rd. (2004). Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures. *J Comput Chem*, 25(2), 265-284. doi: 10.1002/jcc.10378
- Gotz, A. W., Williamson, M. J., Xu, D., Poole, D., Le Grand, S., & Walker, R. C. (2012). Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J Chem Theory Comput*, 8(5), 1542-1555. doi: 10.1021/ct200909j
- Grant, B. J., Rodrigues, A. P., ElSawy, K. M., McCammon, J. A., & Caves, L. S. (2006). Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics*, 22(21), 2695-2696. doi: 10.1093/bioinformatics/btl461
- Guex, N., & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, 18(15), 2714-2723. doi: 10.1002/elps.1150181505
- Hayashi, T., Nozaki, Y., Nishizuka, M., Ikawa, M., Osada, S., & Imagawa, M. (2011). Factor for adipocyte differentiation 158 gene disruption prevents the body weight gain and insulin resistance induced by a high-fat diet. *Biol Pharm Bull*, 34(8), 1257-1263. doi: JST.JSTAGE/bpb/34.1257 [pii]
- Hua, Q. X., Hu, S. Q., Frank, B. H., Jia, W., Chu, Y. C., Wang, S. H., . . . Weiss, M. A. (1996). Mapping the functional surface of insulin by design: structure and function of a novel A-chain analogue. *J Mol Biol*, 264(2), 390-403. doi: 10.1006/jmbi.1996.0648
- Hua, Q. X., Shoelson, S. E., Kochoyan, M., & Weiss, M. A. (1991). Receptor binding redefined by a structural switch in a mutant human insulin. *Nature*, 354(6350), 238-241. doi: 10.1038/354238a0
- Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: visual molecular dynamics. *J Mol Graph*, 14(1), 33-38, 27-38
- Kabsch, W., & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22(12), 2577-2637. doi: 10.1002/bip.360221211
- Kamisetty, H., Ramanathan, A., Bailey-Kellogg, C., & Langmead, C. J. (2011). Accounting for conformational entropy in predicting binding free energies of protein-protein interactions. *Proteins*, 79(2), 444-462. doi: 10.1002/prot.22894
- Kortemme, T., & Baker, D. (2004). Computational design of protein-protein interactions. *Curr Opin Chem Biol*, 8(1), 91-97. doi: 10.1016/j.cbpa.2003.12.008
- Ludvigsen, S., Olsen, H. B., & Kaarsholm, N. C. (1998). A structural switch in a mutant insulin exposes key residues for receptor binding. *J Mol Biol*, 279(1), 1-7. doi: 10.1006/jmbi.1998.1801
- Lyu, P. C., Liff, M. I., Marky, L. A., & Kallenbach, N. R. (1990). Side chain contributions to the stability of alpha-helical structure in peptides. *Science*, 250(4981), 669-673
- Menting, J. G., Whittaker, J., Margetts, M. B., Whittaker, L. J., Kong, G. K., Smith, B. J., . . . Lawrence, M. C. (2013). How insulin engages its primary binding site on the insulin receptor. *Nature*, 493(7431), 241-245. doi: 10.1038/nature11781

- Mirmira, R. G., Nakagawa, S. H., & Tager, H. S. (1991). Importance of the character and configuration of residues B24, B25, and B26 in insulin-receptor interactions. *J Biol Chem*, *266*(3), 1428-1436
- Mirmira, R. G., & Tager, H. S. (1989). Role of the phenylalanine B24 side chain in directing insulin interaction with its receptor. Importance of main chain conformation. *J Biol Chem*, *264*(11), 6349-6354
- Mongan, J., Simmerling, C., McCammon, J. A., Case, D. A., & Onufriev, A. (2007). Generalized Born model with a simple, robust molecular volume correction. *J Chem Theory Comput*, *3*(1), 156-169. doi: 10.1021/ct600085e
- Moretti, R., Fleishman, S. J., Agius, R., Torchala, M., Bates, P. A., Kastiris, P. L., . . . Baker, D. (2013). Community-wide evaluation of methods for predicting the effect of mutations on protein-protein interactions. *Proteins*, *81*(11), 1980-1987. doi: 10.1002/prot.24356
- Nakagawa, S. H., & Tager, H. S. (1992). Importance of aliphatic side-chain structure at positions 2 and 3 of the insulin A chain in insulin-receptor interactions. *Biochemistry*, *31*(12), 3204-3214
- Nguyen, H., Roe, D. R., & Simmerling, C. (2013). Improved Generalized Born Solvent Model Parameters for Protein Simulations. *J Chem Theory Comput*, *9*(4), 2020-2034. doi: 10.1021/ct3010485
- Nisius, B., Sha, F., & Gohlke, H. (2012). Structure-based computational analysis of protein binding sites for function and druggability prediction. *J Biotechnol*, *159*(3), 123-134. doi: 10.1016/j.jbiotec.2011.12.005
- O'Neil, K. T., & DeGrado, W. F. (1990). A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science*, *250*(4981), 646-651
- Onufriev, A., Bashford, D., & Case, D. A. (2004). Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins*, *55*(2), 383-394. doi: 10.1002/prot.20033
- Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T. M., & Baldwin, R. L. (1990). Relative helix-forming tendencies of nonpolar amino acids. *Nature*, *344*(6263), 268-270. doi: 10.1038/344268a0
- Pandeyarajan, V., Smith, B. J., Phillips, N. B., Whittaker, L., Cox, G. P., Wickramasinghe, N., . . . Weiss, M. A. (2014). Aromatic anchor at an invariant hormone-receptor interface: function of insulin residue B24 with application to protein design. *J Biol Chem*, *289*(50), 34709-34727. doi: 10.1074/jbc.M114.608562
- Parasuraman, S. (2012). Protein data bank. *J Pharmacol Pharmacother*, *3*(4), 351-352. doi: 10.4103/0976-500X.103704
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, *25*(13), 1605-1612. doi: 10.1002/jcc.20084
- Pierce, B. G., Hourai, Y., & Weng, Z. (2011). Accelerating protein docking in ZDOCK using an advanced 3D convolution library. *PLoS One*, *6*(9), e24657. doi: 10.1371/journal.pone.0024657
- Pierce, B. G., Wiehe, K., Hwang, H., Kim, B. H., Vreven, T., & Weng, Z. (2014). ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics*, *30*(12), 1771-1773. doi: 10.1093/bioinformatics/btu097
- Pullen, R. A., Lindsay, D. G., Wood, S. P., Tickle, I. J., Blundell, T. L., Wollmer, A., . . . Gammeltoft, S. (1976). Receptor-binding region of insulin. *Nature*, *259*(5542), 369-373
- Sobolev, V., Eyal, E., Gerzon, S., Potapov, V., Babor, M., Prilusky, J., & Edelman, M. (2005). SPACE: a suite of tools for protein structure prediction and analysis based on complementarity and environment. *Nucleic Acids Res*, *33*(Web Server issue), W39-43. doi: 10.1093/nar/gki398
- Wan, Z. L., Huang, K., Xu, B., Hu, S. Q., Wang, S., Chu, Y. C., . . . Weiss, M. A. (2005). Diabetes-associated mutations in human insulin: crystal structure and photo-cross-linking studies of a-chain variant insulin Wakayama. *Biochemistry*, *44*(13), 5000-5016. doi: 10.1021/bi047585k

- Wan, Z. L., Xu, B., Chu, Y. C., Katsoyannis, P. G., & Weiss, M. A. (2003). Crystal structure of allo-Ile(A2)-insulin, an inactive chiral analogue: implications for the mechanism of receptor binding. *Biochemistry*, 42(44), 12770-12783. doi: 10.1021/bi034430o
- Ward, C. W., Menting, J. G., & Lawrence, M. C. (2013). The insulin receptor changes conformation in unforeseen ways on ligand binding: sharpening the picture of insulin receptor activation. *Bioessays*, 35(11), 945-954, doi/10.1002/bies.201370111. doi: 10.1002/bies.201300065
- Weiss, M. A., Hua, Q. X., Jia, W., Chu, Y. C., Wang, R. Y., & Katsoyannis, P. G. (2000). Hierarchical protein "un-design": insulin's intrachain disulfide bridge tethers a recognition alpha-helix. *Biochemistry*, 39(50), 15429-15440
- Weitzel, G., Bauer, F. U., & Eisele, K. (1978). Structure and activity of insulin, XVI. Semisyntheses of desheptapeptide-(B24--30)- up to destriptide-(B28--30)-insulin with lysine or alanine in place of arginine in position B22: influence on the three-step-increase of activity in positions B24--26 (Phe-Phe-Tyr). *Hoppe Seylers Z Physiol Chem*, 359(8), 945-958
- Wollmer, A., Strassburger, W., Glatter, U., Dodson, G. G., McCall, M., Gattner, H. G., . . . Rittel, W. (1981). Two mutant forms of human insulin. Structural consequences of the substitution of invariant B24- or B25-phenylalanine by leucine. *Hoppe Seylers Z Physiol Chem*, 362(6), 581-591
- Xu, B., Hua, Q. X., Nakagawa, S. H., Jia, W., Chu, Y. C., Katsoyannis, P. G., & Weiss, M. A. (2002). Chiral mutagenesis of insulin's hidden receptor-binding surface: structure of an allo-isoleucine(A2) analogue. *J Mol Biol*, 316(3), 435-441. doi: 10.1006/jmbi.2001.5377
- Zakova, L., Kletvikova, E., Veverka, V., Lepsik, M., Watson, C. J., Turkenburg, J. P., . . . Brzozowski, A. M. (2013). Structural integrity of the B24 site in human insulin is important for hormone functionality. *J Biol Chem*, 288(15), 10230-10240. doi: 10.1074/jbc.M112.448050