



Computational study of the activity, dynamics, energetics and conformations of insulin analogues using molecular dynamics simulations: Application to hyperinsulinemia and the critical residue B26



Anastasios Papaioannou^{a,b,*}, Serdar Kuyucak^b, Zdenka Kuncic^{a,b,*}

^a Charles Perkins Centre, University of Sydney, Sydney, NSW, Australia

^b School of Physics, University of Sydney, NSW, Australia

ARTICLE INFO

Keywords:

Molecular dynamics
Insulin
Mutations
Diabetes

ABSTRACT

Due to the increasing prevalence of diabetes, finding therapeutic analogues for insulin has become an urgent issue. While many experimental studies have been performed towards this end, they have limited scope to examine all aspects of the effect of a mutation. Computational studies can help to overcome these limitations, however, relatively few studies that focus on insulin analogues have been performed to date. Here, we present a comprehensive computational study of insulin analogues—three mutant insulins that have been identified with hyperinsulinemia and three mutations on the critical B26 residue that exhibit similar binding affinity to the insulin receptor—using molecular dynamics simulations with the aim of predicting how mutations of insulin affect its activity, dynamics, energetics and conformations. The time evolution of the conformers is studied in long simulations. The probability density function and potential of mean force calculations are performed on each insulin analogue to unravel the effect of mutations on the dynamics and energetics of insulin activation. Our conformational study can decrypt the key features and molecular mechanisms that are responsible for an enhanced or reduced activity of an insulin analogue. We find two key results: 1) hyperinsulinemia may be due to the drastically reduced activity (and binding affinity) of the mutant insulins. 2) Y26_BS and Y26_BE are promising therapeutic candidates for insulin as they are more active than WT-insulin. The analysis in this work can be readily applied to any set of mutations on insulin to guide development of more effective therapeutic analogues.

1. Introduction

Insulin is a hormone produced by the pancreas that is mainly responsible for the maintenance of metabolic homeostasis [1]. Proinsulin, the precursor of insulin [2,3], is converted to the final form of insulin within the secretory granules by enzymatic removal of the C-peptide at the sites of dibasic amino acids [4,5]. The final structure of insulin comprises two polypeptide chains named A and B. A-chain consists of 21 amino acids and contains two α -helices, while B-chain consists of 30 amino acids and contains a central α -helix. The two chains are stabilized by two inter-chain disulphide bridges between A7-B7 and A20-B19, while A-chain is further stabilized by an intra-chain disulphide bridge between A6-A11. The first crystal structure of insulin was determined by Hodgkin et al. in 1969 [6], followed by other structures of insulin and its analogues in the storage form (hexamers and dimers) [7–12].

Insulin binds to the extracellular domain (ectodomain) of the insulin receptor (IR) to initiate the cellular activities of insulin [13]. IR belongs

to the family of tyrosine kinase receptors and is a disulphide-linked $(\alpha\beta)_2$ homodimer. Two binding sites that interact with the IR have been identified in insulin [14]. The first one, called the primary binding site, consists of the hormone-dimerizing residues and interacts with the α CT segment of one IR α -chain and the central β -sheet of the L1 domain of the other α -chain [14–19]. The second binding site consists of the hormone-hexamerizing residues and is suggested to interact with the IR site at the junction of FnIII-1 and 2 of the α -chain [14–16,18–21]. The first crystal structure of the primary binding site of insulin in complex with IR was recently determined [22], identifying the conformational changes necessary for insulin to bind to IR. Many other studies have since confirmed these results and also shown that conformational changes occur in the IR as well [22–30]. All these studies have revealed a series of conformational changes during insulin activation. Initially, insulin B-chain C-terminal (BC-CT) moves away from the protein, exposing its hydrophobic core [19,23–31]. The detachment of BC-CT enables its insertion between the α CT segment and L1 domain of IR, while the B-chain α -helix of insulin is oriented in parallel with the α CT

* Corresponding authors at: Charles Perkins Centre, University of Sydney, Sydney, NSW, Australia.

E-mail addresses: anastasios.papaioannou@sydney.edu.au (A. Papaioannou), zdenka.kuncic@sydney.edu.au (Z. Kuncic).

<http://dx.doi.org/10.1016/j.bbrep.2017.04.006>

Received 10 March 2017; Accepted 13 April 2017

Available online 18 April 2017

2405-5808/ © 2017 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

segment, creating strong interactions between the key residues of insulin and the α CT segment and L1 domain [19,22,24,29,32,33].

Inability of insulin to bind to the IR can result in several diseases, the main one being diabetes mellitus. Diabetes mellitus is a group of metabolic diseases, in which there is a dramatic decrease of insulin secretion from the pancreas and/or insulin resistance is increased, that is, cells fail to respond to the produced insulin. Although the intrinsic causes of diabetes have not yet been identified, some mutations in the insulin gene can lead to the production of structurally abnormal insulin or proinsulin with reduced biological activity and binding affinity to the IR, and hence can cause diabetes. To date, several insulin and proinsulin mutations have been associated with diabetes. Three mutant insulins have been identified to cause hyperinsulinemia, a condition in which the level of insulin circulating in the blood is higher than normal in non-diabetic people. Tager et al. discovered the first mutant human insulin, designated insulin Chicago (F25_BL^{*}) [34]. Subsequently, a second mutant insulin was reported, designated insulin Los Angeles (F25_BS) [35–37], and the third mutant insulin discovered in Japan, designated insulin Wakayama (V3_AL) [38–40]. All three mutant insulins have been shown to have significantly reduced biological activity and binding affinity to the IR [41–45], with the insulin Wakayama being the one with the lowest activity and affinity. We note that these mutations are relatively rare and do not always lead to diabetes.

Due to the rapid rise in the number of diabetic people, there is great deal of interest in the development of insulin analogues with enhanced activity and affinity, which can be used for therapeutic purposes. Extensive mutagenesis and structural studies revealed the significant role of the BC-CT—and in particular the triplet of the B-chain aromatic residues F24_B, F25_B and Y26_B—in the activation process of insulin and the binding interface [16,22–26,29–31,46–49]. The F24_B residue has been shown to act as a hinge as the BC-CT opens during the activation of insulin (BC-CT opening) [23,25,26,30,46], while the Y26_B residue has been shown to be a critical residue for the activation of insulin [26,30,31,47–49]. These two residues are also part of the hydrophobic core of insulin, which contains the hydrophobic residues L11_B, V12_B, L15_B, F24_B, Y26_B, I2_A and V3_A.

Although a significant number of experimental studies focus on insulin and its complex structure with the IR [16,17,20–29,31–33,46–50], as well as on mutations on the critical residues of insulin for predicting potential therapeutic candidates [23,25–27,31,46–49], there are still a lot of key features that cannot be solved or are hard to explain using only experimental approaches. There have been relatively few computational studies investigating the conformational changes and dynamics of insulin [25,30,31,51,52], and even less focusing on insulin mutations [25,31,53,54]. Therefore, we present a comprehensive computational study of mutant insulins, focusing on changes in the activity, dynamics, energetics and conformations of insulin due to mutations using molecular dynamics (MD) simulations. We consider two groups of insulin analogues. The first group involves the mutant insulins that cause hyperinsulinemia, namely, F24_BS, F25_BL and V3_AL, which will be called “hyperinsulinemia analogues”. The second group involves three recently identified mutations on the critical residue B26 that exhibit a similar or enhanced binding affinity to IR, namely, Y26_BA, Y26_BE and Y26_BS [49], which will be called “critical residue analogues”. The activity of the insulin analogues is examined by calculating the probabilities of the conformations of insulin, namely closed/inactive, open/active and wide-open/active (the conformation that fits the receptor binding interface) [30]. The effect of the mutations on the dynamics and energetics of insulin is studied by the probability density function (PDF) and the potential of mean force (PMF) calculations, respectively. Furthermore, conformational analysis is performed to determine how these mutations affect the structure of insulin using several methods, such as the calculation of the hinge interaction and the strongest interactions in the hydrophobic core, principal component analysis (PCA), correlation maps and Root

Mean Square Deviation (RMSD) distributions. The mutant insulin results are compared with those of WT insulin, which was extensively characterized in our previous studies [30,55].

2. Methods

2.1. Modeling and preparation of insulin analogues

The crystal structure of insulin determined at 1.8 Å resolution is available from the Protein Data Bank (PDB ID: 2G4M) [56]. The hyperinsulinemia analogues, F24_BS, F25_BL and V3_AL, and the critical residue analogues, Y26_BA, Y26_BE and Y26_BS, were generated from the crystal structure of insulin using the Mutator plugin in VMD [57]. The VMD software was also used to build and prepare all the simulations.

The following protocol was used for minimization and equilibration in all insulin analogues. Each analogue was embedded in a water box with periodic boundary conditions and solvated with approximately 3500 water molecules. Then, the simulation system was neutralized and ionized by setting the NaCl concentration to 0.1 M by randomly placing Na⁺ and Cl⁻ ions in the water box. The x, y, z dimensions of the simulation box were 55, 53 and 46 Å, respectively. After this protocol, the system was subjected to energy minimization. The simulation systems were equilibrated in two stages after energy minimization. In the first stage, the correct water densities were obtained by using 1 atm pressure coupling and restraining the protein atoms. During the second stage, the protein was gradually relaxed by reducing the restraints, applied on the side chain and backbone atoms of the protein, from $k=10\text{--}0.1\text{ kcal/mol/\AA}^2$. It is important to mention that a typical time scale for equilibration of a side chain after a mutation is a few ns. However, this assumes a water environment where the side chain can easily rotate, which is not the case for insulin, because all the mutations are applied on its hydrophobic pocket. In the hydrophobic pocket, the network of interactions in combination with the fact that other residues are also involved in the rearrangement process restrict any movement. Thus, the equilibration process in most of the mutations took a few hundreds of ns, which make this step very important for avoiding misleading results. The equilibrated system was then ready for production run without any restraints on the protein atoms. The RMSD of the protein backbone atoms was used to monitor the relaxation and equilibration of the protein.

2.2. Molecular Dynamics (MD) simulations

MD simulations were carried out using the NAMD software [58] and the CHARMM36 force field [59,60]. Periodic boundary conditions were employed in combination with the Particle-Mesh Ewald (PME) algorithm to compute the long-range electrostatic interactions. An NpT ensemble was used with the temperature and pressure maintained at 300 K and 1 atm, respectively, via the Langevin coupling with a damping coefficient of 5 ps⁻¹. The Lennard-Jones (LJ) interactions were switched off within a distance of 10.5–13.5 Å using a switching function. A time step of 2 fs was used and the trajectory data were recorded at every 15 ps in all simulations. The MD simulations were performed for a total time of 600 ns after equilibration for each insulin analogue except for Y26_BA, which was run for 1.2 μs to ensure adequate sampling (see the PMF calculations for more details). The analyses and molecular graphics were performed using the VMD [57], UCSF Chimera [61] and Matlab [62] software. The statistical analyses in the “Results and discussion” section were based on the trajectory data collected from the long MD simulations of the insulin analogues and our previous study of WT insulin [30].

2.3. Probability Density Function (PDF) calculations

The PDF of a variable describes the relative likelihood for this variable to occur at a particular value. In our MD simulations, this

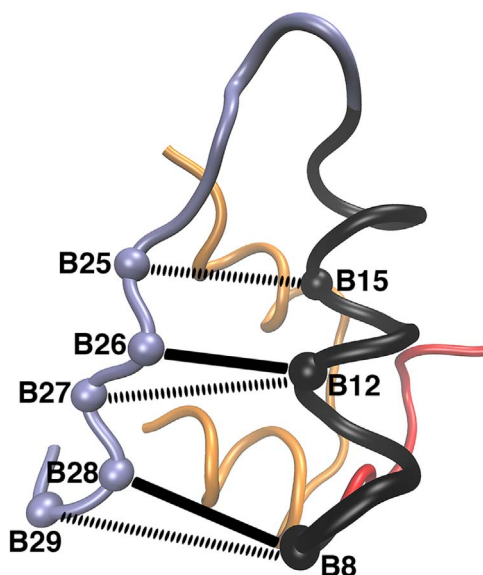


Fig. 1. Illustration of the distances between the C_{α} atoms of the BC-CT residues B25–B29, and their closest partner in the B-chain α -helix. The two critical distances, namely B26(C_{α})-B12(C_{α}) and B28(C_{α})-B8(C_{α}), that provide the best criterion for the BC-CT opening and are used in the PDF and PMF calculations are shown in solid lines.

variable was chosen as the distances between the C_{α} atoms of the critical residues of the BC-CT from their closest partner in the B-chain α -helix, namely, B26(C_{α})-B12(C_{α}) and B28(C_{α})-B8(C_{α}) (Fig. 1). These two particular distances were chosen because they provide the best measure for the activation process of insulin, i.e. the opening of the BC-CT [30]. The PDF frequency histograms represent the probability of a distance occurring during the MD simulations. The histograms are characterized by three different colors, which define the three distinct conformations of BC-CT, closed, open and wide-open, as observed in our previous MD simulations of insulin [30]. The inactive state of insulin corresponds to its closed conformation. The transition from the inactive to the active state occurs when the hydrophobic core breaks due to the entry of water molecules and BC-CT opens. The active conformation that fits to the IR binding interface corresponds to the wide-open conformation of insulin. The threshold for the transition from the closed to the open conformation was determined by the number of water molecules within the hydrophobic core, i.e., a zero number of water molecules implies a closed conformation, whereas a non-zero number, an open conformation. The threshold for the transition from the open to the wide-open conformation was determined by superimposing the trajectories of insulin obtained from the MD simulations on the complex structure of insulin bound to the IR (PDB ID: 4OGA) [23]. In the case of the F24_BS analogue, the thresholds for the transition from the closed to the open and from the open to the

wide-open conformation were readjusted because the BC-CT residues assumed a new configuration after the mutation. This new configuration is driven by the creation of a network of hydrogen bonds, which causes the B24–B26 residues to move slightly further away from the core, and the B27–B30 residues to move in parallel to the B-chain α -helix but not in the direction of the wide-open conformation. We discuss the configuration of the BC-CT of this particular analogue in more detail in the “Results and discussion” section.

2.4. Potential of Mean Force (PMF) calculations

The PMF is commonly used in MD simulations to examine the free energy changes in a system with respect to a specific reaction coordinate. The energetics of the BC-CT opening of the simulated insulin analogues were quantitatively characterized by the PMF calculations using the distances in Fig. 1 as reaction coordinates. The trajectory data from the long MD simulation provided sufficient sampling to calculate the PMF directly from the PDF data using the Boltzmann equation

$$W = -k_B T \ln(\rho/\rho_0) \quad (1)$$

where W is the free energy, k_B is the Boltzmann constant, T is the temperature, ρ is the density obtained from the distance distributions, and ρ_0 is chosen as the maximum density. The PMF calculation provides an estimate of the free energy change required for activation of insulin.

Convergence of the PMF calculations was checked to verify that the simulation time was adequate for sampling. Fig. 2 shows PMF curves at increasing simulation times for WT-insulin and the V3_AL and Y26_BA analogues (all the other analogues studied were found to have similar PMF behavior with that of V3_AL and are not shown here). In WT-insulin and V3_AL analogue, the PMF curves begin to acquire their final form after ≈ 300 ns and converge after 500 ns of simulation time. This confirms that the simulation time of 600 ns is sufficient for the PMF calculations. However, in Y26_BA, the PMF curves acquire a new form after 500 ns and start to converge after 1000 ns. Thus, a simulation time of 1200 ns appears to be adequate for a converged PMF calculations for this analogue.

2.5. Principal Component Analysis (PCA)

PCA, also known as Essential Dynamics Analysis, is a useful tool for analyzing the relationship between different conformers in proteins. PCA has been extensively applied to both distributions of experimental structures and MD trajectories to provide insights into the nature of conformational differences in a range of proteins and biomolecules [63–69]. Using MD trajectories, PCA can reveal variances and correlations between motions in different parts of a protein. PCA transforms the data to a new coordinate system using an orthogonal linear transformation. The resulting coordinates, named principal components

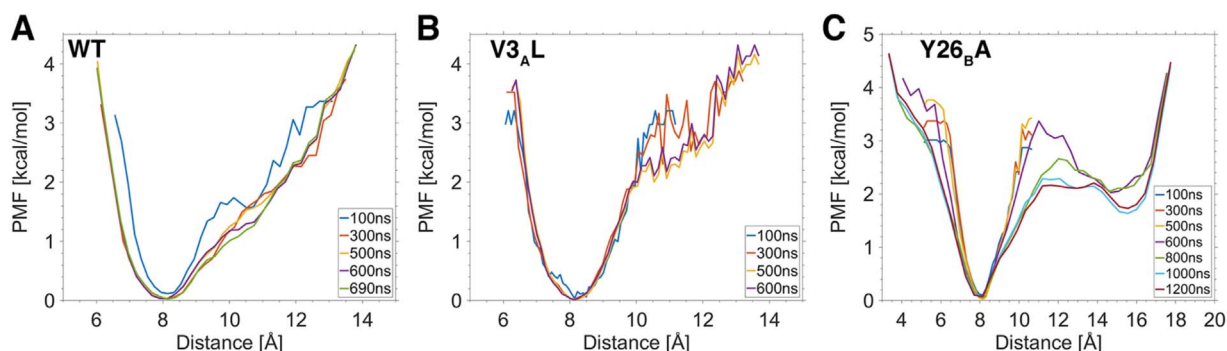


Fig. 2. Convergence analysis of the PMF calculation in (A) WT-insulin (B) the V3_AL analogue and (C) the Y26_BA analogue. In WT-insulin and V3_AL analogue, the PMF calculation converges after 500 ns, whereas in Y26_BA, it converges after 1000 ns, suggesting that a total simulation time of 600 ns and 1200 ns is adequate for sampling in WT-insulin/V3_AL analogue and Y26_BA analogue, respectively. All the other analogues display similar behavior to V3_AL.

(orthogonal eigenvectors), project the variance of the conformers of the protein, with the greatest variance lying on the first principal component, the second greatest variance on the second principal component and so on. The number of principal components that can sufficiently capture over 70% of the total variance of the system can provide a useful description retaining most of the total variance of the system.

In the insulin analogues and WT-insulin, only the C_{α} atoms of the BC-CT residues were chosen for PCA because only those residues are involved in the activation process. All the MD trajectories were superposed according to the most invariant region, which is the B-chain α -helix in insulin. This technique was used to avoid underestimating the true atomic displacement and to clearly highlight the domain rearrangement [70]. Here we use it to highlight the detachment of BC-CT from its hydrophobic core during activation of insulin.

2.6. Correlation map

A correlation matrix is useful for measuring the dependence between multiple variables. In our case, the linear correlation matrix was calculated using the RMSDs of the C_{α} atoms of the BC-CT residues throughout each MD simulation by employing the Pearson correlation coefficient [71]. The correlation matrix was used to determine whether there is collective (correlated) movement between the BC-CT residues.

3. Results and discussion

3.1. Quantifying activity

Activity of the insulin analogues was initially investigated by inspecting the time series of the B26(C_{α})-B12(C_{α}) and B28(C_{α})-B8(C_{α}) distances for the hyperinsulinemia and the critical residue analogues shown in Fig. 3. The time series provide a very useful qualitative description of the activity of the analogues compared to WT-insulin, and indicate a decreased activity for all the hyperinsulinemia analogues, while all the critical residue analogues display an increased activity. To examine the relative activities of each analogue more quantitatively, the average probabilities of the conformations were calculated from the individual probabilities for the distances between the C_{α} atoms of the BC-CT residues B25-B29 (B30 is not included because it is the C-terminal residue), and their closest partners in the B-chain α -helix (Fig. 1). The average probabilities of the conformations of the hyperinsulinemia and the critical residue analogues are presented in Table 1. It is evident from Table 1 that only the critical residue analogues show a higher probability of the active conformations (both open and wide-open conformations) than WT-insulin, while the hyperinsulinemia mutations make insulin more inactive. The rank order of activity, from most to least active, is: Y26_BS > Y26_BE > Y26_BA > WT > V3_AL > F25_BL > F24_BS. Interestingly, the hyperinsulinemia analogues show even lower probability of the wide-open conformation than WT-insulin, which rarely attains this conformation [30,55]. On the other hand, Y26_BA, and to a greater extent Y26_BE and Y26_BS, show substantially higher probabilities of the wide-open conformation compared to WT-insulin. Thus, the critical residue analogues are not only more active than WT-insulin, but they can also engage the receptor binding site more rapidly. The weaker activity of F24_BS compared to WT-insulin is a surprising result. One would expect that substitution of a hydrophobic side chain with a polar one in the hinge residue would weaken the hinge and expose the hydrophobic core, leading to a more active insulin. In fact, the BC-CT residues in the F24_BS analogue rearrange to form a hydrogen-bond network, which leads to a more stable hinge and hydrophobic core. We discuss this rearrangement in more detail in the “Elucidating the conformational changes due to a mutation” section below.

3.2. Unravelling the dynamics and energetics

As the BC-CT does not behave like a rigid rod, its opening cannot be adequately described by a single distance. As mentioned in Fig. 1, the BC-CT dynamics is best described by both the B26(C_{α})-B12(C_{α}) and B28(C_{α})-B8(C_{α}) distances. The data from the time series of these distances (Fig. 3) were converted into probability densities. The PDFs of these two distances for the hyperinsulinemia and the critical residue analogues are shown in Fig. S1 and S2, respectively. The effect of the mutations on the energetics of the BC-CT opening were further quantified by calculating the PMF profiles from the PDFs. The two classes of analogues behave quite differently, so we discuss them separately.

Inspection of the PDFs for the hyperinsulinemia analogues and WT insulin (Fig. S1) shows that the B26(C_{α})-B12(C_{α}) and B28(C_{α})-B8(C_{α}) distance distributions are qualitatively similar with all peaking around the closed conformation. Thus, the PDFs from the two distances exhibit similar dynamic behavior. Comparison of the hyperinsulinemia analogues with WT-insulin shows that there are substantial differences between the open and wide-open probabilities. The wide-open conformation is a relatively rare event in WT-insulin (probability \approx 2%). In contrast, the wide-open conformation is almost completely suppressed in all the hyperinsulinemia analogues, implying that these analogues are even less effective than WT-insulin. The PMFs in Fig. 4A-B help to interpret the result in terms of free energy changes during the conformational transitions. The PMF in WT-insulin rises steadily similar to a harmonic potential, thus a relatively large amount of energy is required to attain its wide-open conformation. The PMFs in all three hyperinsulinemia analogues exhibit a much steeper rise around the open and wide-open conformations than that of WT-insulin. Thus, energetically it is much more difficult to activate these analogues compared to WT-insulin.

We next consider the critical residue analogues. The PDFs for the B26(C_{α})-B12(C_{α}) distance (Fig. S2A) are seen to exhibit a different dynamic behavior, namely, they all peak in the open conformation while the PDFs for the B28(C_{α})-B8(C_{α}) distance (Fig. S2B) peak in the closed conformation, consistent with the behavior of PDFs in WT-insulin and hyperinsulinemia analogues. This happens because the mutations on Y26 weakens the hydrophobic core, allowing entry of water molecules which increases the B26(C_{α})-B12(C_{α}) distance but not the B28(C_{α})-B8(C_{α}) distance. This effect can also be seen in the corresponding PMFs in Fig. 4C-D, where the minima in the B26(C_{α})-B12(C_{α}) PMFs are shifted from the closed conformation in WT-insulin to the open conformation in the critical residue analogues but no such shift occurs in the B28(C_{α})-B8(C_{α}) PMFs. The PDFs for the critical residue analogues are much broader compared to WT-insulin and the hyperinsulinemia analogues, indicating more readily accessible open and wide-open conformations. From the corresponding PMFs (Fig. 4C-D), the free energy required for making such conformational changes in the critical residue analogues is estimated around 1–2 kcal/mol, which is much lower than WT-insulin, especially for the wide-open conformations.

3.3. Elucidating the conformational changes due to a mutation

The residue B24 has been shown to play a hinge role during opening of the BC-CT [23,25,26,30,46,55]. The strength of the interaction of this hinge residue with its closest partner in the B-chain α -helix (B15) was shown to play a significant role in the activity and stability of insulin [30,55]. Fig. 5 shows the time series for the B24(C_{α})-B15(C_{α}) distance for WT-insulin and its analogues. Only in the F24_BS analogue does the hinge distance appear to exhibit a different pattern of fluctuations compared to the rest of the analogues and WT-insulin. This is explained by the mutation of the large hydrophobic F residue to the smaller polar S residue at the hinge position. Inspection of the MD trajectories for the F24_BS analogue shows that the B24 side chain

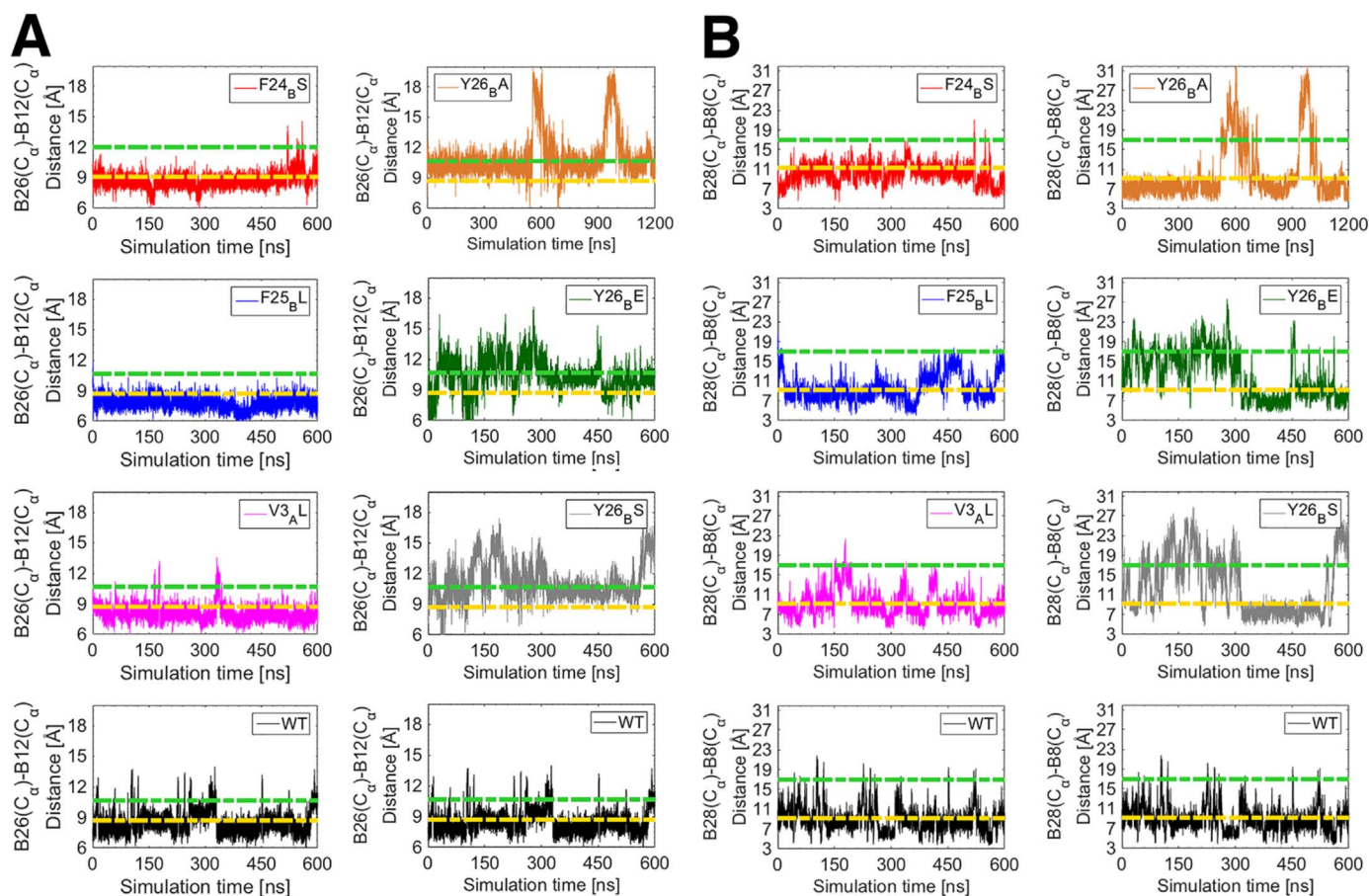


Fig. 3. Time series for the hyperinsulinemia and critical residue analogues using (A) the $B26(C_{\alpha})-B12(C_{\alpha})$ distance and (B) the $B28(C_{\alpha})-B8(C_{\alpha})$ distance. The corresponding plots for WT-insulin are also shown for comparison in the bottom row. The threshold from the closed to the open conformation and from the open to the wide-open conformation are represented by yellow and green dashed lines in the plots, respectively.

relocates after equilibration and creates a network of hydrogen bonds with its neighboring residues in the A-chain and B-chain (Fig. S3). This results in an even tighter hinge than in WT-insulin, which is consistent with the PMFs in Fig. 4A, i.e., the $F24_B S$ PMF rises more sharply than that of WT-insulin.

Apart from the hinge interaction, the strength of the interactions in the hydrophobic core of insulin has also been shown to significantly affect its activity and stability [55]. In particular, a stable hydrophobic core in combination with a strong and stable hinge interaction reduce insulin activity, and vice-versa. Table 2 presents the strongest interactions in the hydrophobic core of WT-insulin and insulin analogues. It is evident from Table 2 that the hinge interactions are stable and strong in WT-insulin and all its analogues during the simulations, including the $F24_B S$ analogue. All the hyperinsulinemia analogues show more stable interactions in the hydrophobic core compared to WT-insulin, which make them less active. The reason why $F24_B S$ is less active than WT-insulin becomes apparent upon consideration of the changes in the core. As described above, the $F24_B S$ mutation creates a network of hydrogen bonds between $S24_B(OG)$, $Y19_A(O)$ and $Y26_B(N)$ (Fig. S3),

which stabilizes the position of the hinge residue. In addition, the relocation of the B26 residue (e.g., the $B26-A2$ and $B26-A3$ distances in Table 2 become shorter) causes the hydrophobic side chain of the $Y26_B$ residue to be buried further in the hydrophobic core, strengthening the interactions with the other residues in the core. In contrast, the B26 distances are uniformly larger for the critical residue analogues compared to WT-insulin (Table 2). They display a much weaker and unstable hydrophobic core, leading to more active insulin analogues compared to WT-insulin, which corroborates our calculations of the probabilities of the conformations (Table 1). Consequently, the strength and the stability of the interactions within the core, as well as the hinge interaction, are closely correlated with the activity of insulin.

The average distances and the standard deviations were calculated using the trajectory data from the MD simulations.

To analyze the relationship between the conformations occurring during the MD simulations, PCA was performed on the simulation trajectories for the insulin analogues and WT-insulin. Fig. S4 and S5 show the PCA results for the hyperinsulinemia and critical residue analogues compared to WT-insulin, respectively. In all cases, the first 3

Table 1

Average probabilities of the conformations of the hyperinsulinemia and critical residue analogues compared to WT insulin.

Conformation	WT	Hyperinsulinemia analogues			Critical residue analogues		
		$F24_B S$	$F25_B L$	$V3_A L$	$Y26_B A$	$Y26_B E$	$Y26_B S$
Closed/Inactive	0.73	0.83	0.80	0.77	0.57	0.41	0.41
Open/Active	0.25	0.17	0.19	0.22	0.27	0.39	0.32
Wide-open/Active	0.02	0.00	0.01	0.01	0.16	0.20	0.27

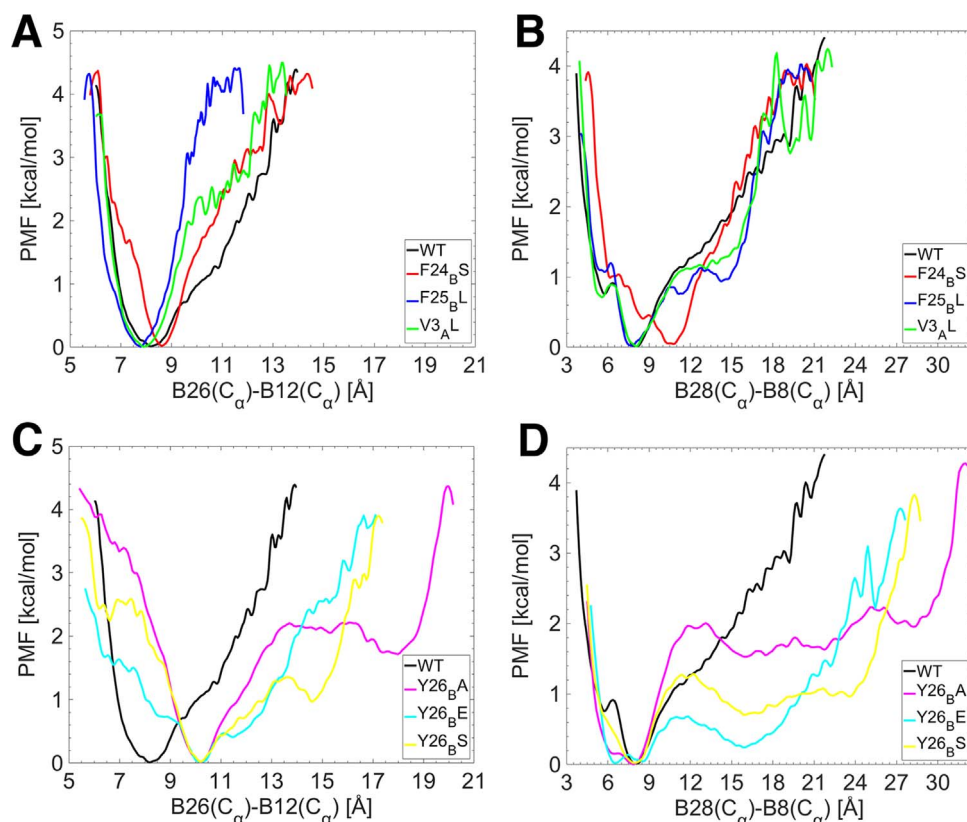


Fig. 4. Comparison of PMFs for the hyperinsulinemia and critical residue analogues against those of WT-insulin using (A, C) $B26(C_{\alpha})$ - $B12(C_{\alpha})$ distance and (B, D) $B28(C_{\alpha})$ - $B8(C_{\alpha})$ distance, respectively.

principal components (eigenvalues) can adequately describe more than 80% of the total variance of the BC-CT in MD trajectories (Fig. S4D and S5D). The color scale in the PCA plots refers to the time evolution of the BC-CT conformers during the MD simulations. The blue color indicates the initial simulation time, which progressively changes to white, denoting the midpoint of the simulation time, and then progressively changes to red, denoting the final simulation time. The PCA results clearly show that the transitions between different conformers throughout the MD simulations are stochastic in nature in all the analogues and WT-insulin. However, the different conformers vary among the insulin analogues and WT-insulin. A key difference between the conformers of WT-insulin and the insulin analogues can be seen in the PCA plots using the first two principal components (Fig. S4A and S5A). While in insulin the same set of conformers are involved throughout the evolution of its MD simulation, in the insulin analogues, different sets of conformers are introduced at different times during the evolution of the MD simula-

tions.

Fig. 6 shows the correlation maps constructed for the insulin analogues and WT-insulin to further examine the conformational changes in the BC-CT, especially the correlation between the movement of the BC-CT residues. The correlation matrices are presented as heat maps with values from -1 (blue, negative correlation), to 0 (white, no correlation), to 1 (red, positive correlation). Surprisingly, in all cases, no clear linear correlation between the movement of all the BC-CT residues is evident, with only the Y26BA analogue exhibiting a strong positive correlation among almost all the BC-CT residues except for the last one. WT-insulin exhibits three distinct groups of residues with strongly positively correlated movement: B24-B25, B25-B26 and B27-B30. Among all the analogues, only V3AL shows similar behavior to WT-insulin, with the only difference being that the three groups of residues are completely uncorrelated to each other for V3AL. The F24BS analogue seems to have similar behavior to WT-insulin regarding the

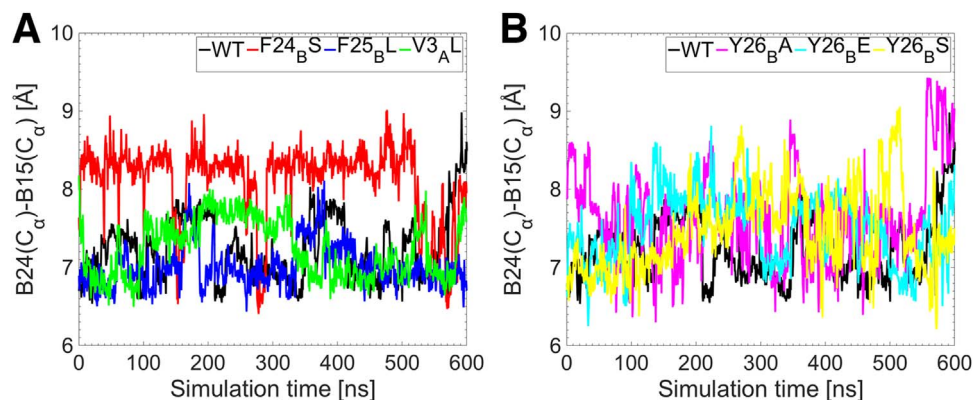


Fig. 5. Time series of the $B24(C_{\alpha})$ - $B15(C_{\alpha})$ distance which forms a hinge between BC-CT and B-chain α -helix (A) in the hyperinsulinemia analogues and (B) in the critical residue analogues, compared to WT-insulin in both cases.

Table 2
Strongest interactions in the hydrophobic core of WT-insulin and insulin analogues.

Residues	Average Distance [\AA]						
	WT	F24 _B S	F25 _B L	V3 _A L	Y26 _B A	Y26 _B E	Y26 _B S
B24-B15 ^a	4.6 ± 0.4	4.7 ± 0.5	4.6 ± 0.4	4.5 ± 0.3	4.9 ± 0.9	4.7 ± 0.6	4.8 ± 0.6
B26-A2 ^b	8.1 ± 3.4	5.9 ± 2.0	5.0 ± 0.7	6.9 ± 3.0	9.4 ± 2.3	10.8 ± 2.1	10.6 ± 2.5
B26-A3 ^b	7.1 ± 2.5	6.2 ± 1.8	5.0 ± 0.8	5.8 ± 1.2	10.7 ± 2.8	12.0 ± 2.2	10.7 ± 2.3
B26-B11 ^c	7.0 ± 2.8	7.9 ± 1.9	5.5 ± 0.8	6.0 ± 1.8	12.2 ± 2.5	13.5 ± 2.1	12.4 ± 2.1
B26-B12 ^c	5.6 ± 2.4	5.5 ± 1.8	4.4 ± 0.8	4.7 ± 1.3	10.8 ± 2.6	10.8 ± 2.2	10.8 ± 2.2
B26-B15 ^c	6.5 ± 2.1	6.1 ± 2.3	5.6 ± 0.7	5.9 ± 1.2	9.6 ± 1.8	10.9 ± 1.4	9.9 ± 1.6

^a The hinge interaction in WT-insulin and the insulin analogues.

^b The interactions between the critical residue (B26) and the hydrophobic residues of the first α -helix of the A-chain that belong to the hydrophobic core.

^c The interactions between the critical residue (B26) and the hydrophobic residues of the B-chain α -helix that belong to the hydrophobic core.

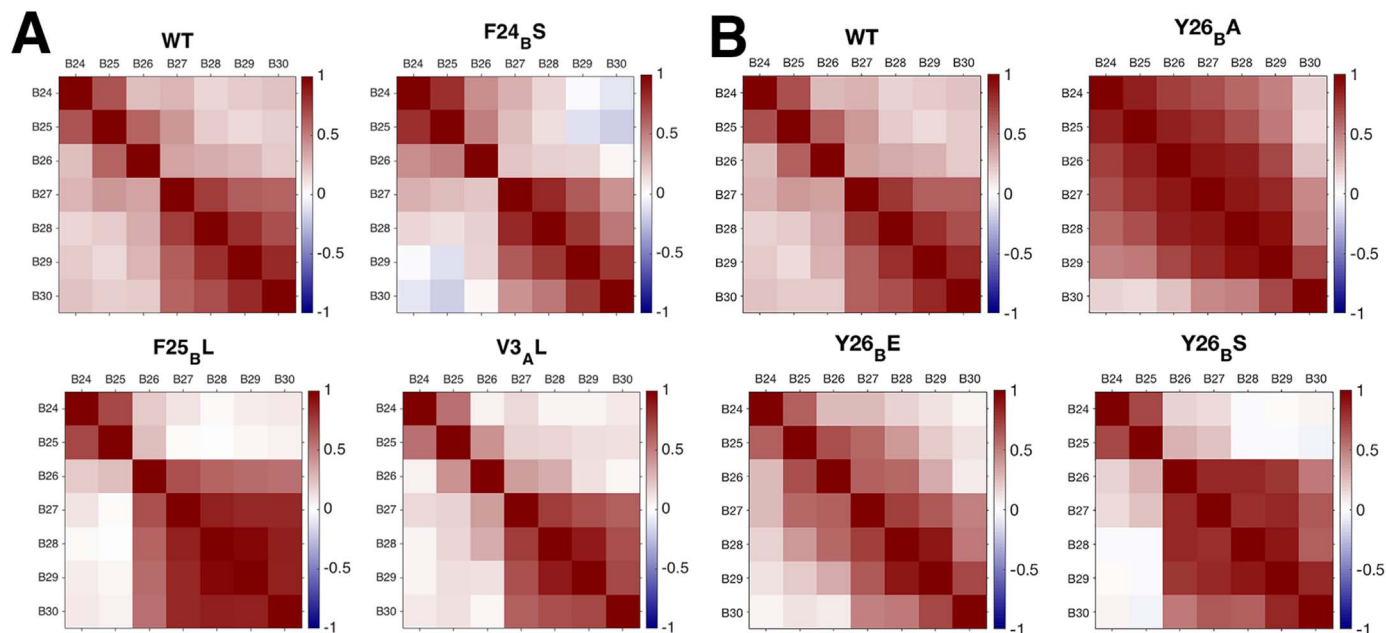


Fig. 6. Correlation maps of the (A) hyperinsulinemia and (B) critical residue analogues compared to WT insulin. WT-insulin is comprised of 3 distinct groups of positively correlated residues, while similar correlation pattern can only be seen in V3_AL and to a lesser extent in F24_BS. The Y26_BA analogue shows the strongest positively correlated behavior, while F25_BL and Y26_BS exhibit strong positively correlated behavior but to a lesser extent and Y26_BE does not show any pattern in its correlation map.

correlated movement of the B27-B30 residues, however its B24-B26 residues are correlated with each other in a different manner, composing two distinct group of correlated residues, namely B24-B25 and B26 alone. Additionally, the B25-B29 and B25-B30 residues of F24_BS are negatively correlated, which is not the case for WT-insulin. Another interesting case is the F25_BL and Y26_BS analogues, which display a strong positively correlated behavior, consisting of only two distinct groups of correlated residues, with the second one including most of the BC-CT residues, i.e. B26-B30. Lastly, in the Y26_BE analogue, the only clear correlations between the movements of the residues are those that involve the closest neighboring BC-CT residues and there is no consistent pattern in its correlation map.

The RMSD throughout the MD simulations was calculated using the initial conformation of each analogue and WT-insulin as a reference, which represents the closed/inactive conformation in all cases. The RMSD was measured using two groups of atoms, the C_α atoms of the residues of the whole insulin and the C_α atoms of only the BC-CT residues. Fig. S6 shows the histogram distributions of the RMSDs using these two groups of atoms. Compared to WT-insulin, the RMSD distributions for all the hyperinsulinemia analogues show similar behavior to that of WT-insulin. It is also evident from Fig. S6 that the RMSD distributions for F24_BS are shifted to higher values due to the relocation that occurs in its B24-B26 residues. In the case of the critical residue analogues, a shift to higher values is seen in addition to more

spread RMSD distributions. This means that the critical residue analogues display distributions with more active conformations due to their more active nature. These results are consistent with the probabilities of the conformations of WT-insulin and its analogues in Table 1.

4. Conclusions

Here, we have presented a comprehensive computational study of insulin analogues, where the activity, dynamics, energetics and conformations of the analogues are systematically analyzed using MD simulations. In particular, we have determined whether a mutation has a positive or negative effect on the activity of insulin using the activity of WT-insulin as a reference. The cause of the difference in the activity was identified by examining the dynamics and energetics of the insulin analogues. These calculations reveal insights into the effect of the mutation on the dynamics of the BC-CT opening (i.e., insulin activation), and the energy needed for the transitions between the conformations of insulin. We have also examined the effect of each mutation on the hydrophobic core and the hinge residue, which have been shown to be correlated with the activity and stability of insulin [55]. Conformational analysis includes the investigation of the time evolution of the conformations using PCA and the examination of the correlation between the movement of the BC-CT residues via correlation maps.

Through such conformational analysis, we have decrypted the key molecular features and mechanisms that are responsible for an enhanced or decreased activity of insulin due to a mutation.

In this study, we focused on two different groups of insulin analogues. The first consists of the three insulin analogues that have been identified with hyperinsulinemia [34–40], while the second group consists of three insulin analogues, which involve mutations on the critical residue B26 and have been shown to have similar or enhanced binding affinity to the IR as WT-insulin [49]. Interestingly, all the hyperinsulinemia analogues show significantly reduced activity compared to WT-insulin. Also, energetically it is more difficult for these analogues to attain the wide-open active conformations, which becomes apparent in the PMF profiles (Fig. 4). This suggests that these three mutants are less active than WT-insulin, and in combination with their significantly reduced binding affinity to the IR [41–45], this may force the pancreas to produce more insulin to cope with the extensive need for more effective insulin. This may explain the excess levels of insulin in the blood due to hyperinsulinemia. In contrast, all the critical residue analogues show much more active behavior than WT-insulin—conformational transitions to more active states are energetically much easier for these analogues. The significantly enhanced activity for Y26_BE and Y26_BS, in particular, in combination with their enhanced binding affinity [49], suggests that these two analogues may be promising potential therapeutic candidates.

Experimental studies deal with many challenges when examining the effect of mutations, which limits their application. For example, in most mutagenesis studies of insulin, only a few critical residues are considered (e.g., B24 [23,25,46] and B26 [26,31,47–49]), and the differences in the activity and binding affinity compared to WT-insulin are determined for several mutations. Exploration of all the possible mutations on a particular residue is rarely attempted as it can be very time-consuming, with no guarantee of success. Furthermore, it is impractical to crystallize all the possible mutant insulins individually and/or in complex with the IR. Hence, it is difficult to explore and identify the key molecular features of the insulin analogues that have an enhanced binding affinity and activity or vice versa.

Computational studies have their own challenges, especially with regard to the validation of results. Long MD simulations are necessary for sufficient sampling to validate the calculations as shown in this study, where the analogues were run for 600–1200 ns. In a recent computational study, binding of the hyperinsulinemia analogues to the insulin receptor was analyzed [53]. The analogues were docked to the IR using computational docking tools. All the mutations were applied on the wide-open conformation of WT-insulin, and docking was attempted using this configuration. However, using the wide-open conformation of the F24_BS analogue for docking is not accurate. Binding of this analogue to the IR like WT-insulin may not be feasible, as it cannot attain its wide-open conformation according to our results. Thus, it is essential to identify and analyze the conformational changes that occur in the structure of insulin due to a mutation prior to the investigation of its binding to the IR. Also, it is important to note that there is no unique active state for WT-insulin or its analogues, but instead a series of active states, as shown in the PMF profiles.

These experimental and computational challenges can be partially or fully overcome using the computational techniques presented here. While we have considered two particular groups of insulin analogues, namely hyperinsulinemia and critical residue analogues, the analysis and the methods used can be readily applied to any set of mutations for designing more effective insulin analogues. Summarizing, the computational study presented here shows that the effect of mutations on insulin properties can be effectively investigated using simulation methods. By applying these methods, one can investigate the effect of a number of mutations to reveal and identify the crucial molecular features of these mutations on insulin activation and binding. The most significant and interesting mutations can then be selected for further investigation using experimental methods to validate the results and gain new

insights. Thus, a combination of computational studies with experimental methods can facilitate the ultimate goal of finding insulin analogues with enhanced affinity, activity, and stability.

Acknowledgments

All the MD simulations were performed using the two High Performance Computing (HPC) facilities in Australia, namely, the Avoca cluster at the Victorian Life Sciences Computation Initiative (VLSI, Melbourne, Victoria, Australia) and the Raijin cluster at the National Computational Infrastructure (NCI, Canberra, Australian Capital Territory, Australia). This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Transparency document

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.bbrep.2017.04.006>.

References

- [1] G. Dodson, D. Steiner, The role of assembly in insulin's biosynthesis, *Curr. Opin. Struct. Biol.* 8 (1998) 189–194.
- [2] D.F. Steiner, P.E. Oyer, The biosynthesis of insulin probable precursor of insulin by a human islet cell adenoma, *Proc. Natl. Acad. Sci. USA* 57 (1967) 473–480.
- [3] D.F. Steiner, D. Cunningham, L. Spigelman, B. Altan, Insulin biosynthesis: evidence for a precursor, *Science* 157 (1967) 697–700.
- [4] D.F. Steiner, S.Y. Park, J. Stoy, L.H. Phillipson, G.I. Bell, A brief perspective on insulin production, *Diabetes Obes. Metab.* 11 (2009) 189–196.
- [5] M.A. Weiss, Proinsulin and the genetics of diabetes mellitus, *J. Biol. Chem.* 284 (2009) 19159–19163.
- [6] M.J. Adams, T.L. Blundell, E.J. Dodson, G.G. Dodson, M. Vijayan, E.N. Baker, M.M. Harding, D.C. Hodgkin, B. Rimmer, S. Sheat, Structure of rhombohedral 2 zinc insulin crystals, *Nature* 224 (1969) 491–495.
- [7] E.N. Baker, T.L. Blundell, J.F. Cutfield, S.M. Cutfield, E.J. Dodson, G.G. Dodson, D.M. Hodgkin, R.E. Hubbard, N.W. Isaacs, C.D. Reynolds, K. Sakabe, N. Sakabe, N.M. Vijayan, The structure of 2Zn pig insulin crystals at 1.5 Å resolution, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 319 (1988) 369–456.
- [8] G. Bentley, E. Dodson, G. Dodson, D. Hodgkin, D. Mercola, Structure of insulin in 4-zinc insulin, *Nature* 261 (1976) 166–168.
- [9] G.D. Smith, D.C. Swenson, E.J. Dodson, G.G. Dodson, C.D. Reynolds, Structural stability in the 4-zinc human insulin hexamer, *Proc. Natl. Acad. Sci. USA* 81 (1984) 7093–7097.
- [10] U. Derewenda, Z. Derewenda, E.J. Dodson, G.G. Dodson, C.D. Reynolds, G.D. Smith, C. Sparks, D. Swenson, Phenol stabilizes more helix in a new symmetrical zinc insulin hexamer, *Nature* 338 (1989) 594–596.
- [11] M.L. Brader, M.F. Dunn, Insulin hexamers: new conformations and applications, *Trends Biochem. Sci.* 16 (1991) 341–345.
- [12] E. Ciszak, J.M. Beals, B.H. Frank, J.C. Baker, N.D. Carter, G.D. Smith, Role of C-terminal B-chain residues in insulin assembly: the structure of hexameric LysB28ProB29-human insulin, *Structure* 3 (1995) 615–622.
- [13] P. De Meyts, J. Whittaker, Structural biology of insulin and IGF1 receptors: implications for drug design, *Nat. Rev. Drug Discov.* 1 (2002) 769–783.
- [14] P. De Meyts, Insulin and its receptor: structure, function and evolution, *Bioessays* 26 (2004) 1351–1362.
- [15] P. De Meyts, W. Sajid, J. Palsgaard, A., M. Theede, L. Gauguin, H. Aladdin, J. Whittaker, Insulin and IGF-I receptor structure and binding mechanism, in: A.R. Saltiel, J.E. Pessin (Eds.), *Mechanisms of Insulin Action*. Springer, New York, 2007, pp. 1–32.
- [16] C.W. Ward, M.C. Lawrence, Ligand-induced activation of the insulin receptor: a multi-step process involving structural changes in both the ligand and the receptor, *Bioessays* 31 (2009) 422–434.
- [17] B.J. Smith, K. Huang, G. Kong, S.J. Chan, S. Nakagawa, J.G. Menting, S.Q. Hu, J. Whittaker, D.F. Steiner, P.G. Katsoyannis, C.W. Ward, M.A. Weiss, M.C. Lawrence, Structural resolution of a tandem hormone-binding element in the insulin receptor and its implications for design of peptide agonists, *Proc. Natl. Acad. Sci. USA* 107 (2010) 6771–6776.
- [18] C.W. Ward, M.C. Lawrence, Similar but different: ligand-induced activation of the insulin and epidermal growth factor receptor families, *Curr. Opin. Struct. Biol.* 22 (2012) 360–366.
- [19] P. De Meyts, Insulin/receptor binding: the last piece of the puzzle? *Bioessays* 37 (2015) 389–397.
- [20] N.M. McKern, M.C. Lawrence, V.A. Streltsov, M.Z. Lou, T.E. Adams, G.O. Lovrecz, T.C. Elleman, K.M. Richards, J.D. Bentley, P.A. Pilling, P.A. Hoyne, K.A. Cartledge, T.M. Pham, J.L. Lewis, S.E. Sankovich, V. Stoichevska, E. Da Silva, C.P. Robinson, M.J. Frenkel, L.G. Sparrow, R.T. Fernley, V.C. Epa, C.W. Ward, Structure of the insulin receptor ectodomain reveals a folded-over conformation, *Nature* 443 (2006) 218–221.

- [21] L. Whittaker, C. Hao, W. Fu, J. Whittaker, High-affinity insulin binding: insulin interacts with two receptor ligand binding sites, *Biochemistry* 47 (2008) 12900–12909.
- [22] J.G. Menting, J. Whittaker, M.B. Margetts, L.J. Whittaker, G.K. Kong, B.J. Smith, C.J. Watson, L. Zakova, E. Kletvikova, J. Jiracek, S.J. Chan, D.F. Steiner, G.G. Dodson, A.M. Brzozowski, M.A. Weiss, C.W. Ward, M.C. Lawrence, How insulin engages its primary binding site on the insulin receptor, *Nature* 493 (2013) 241–245.
- [23] J.G. Menting, Y. Yang, S.J. Chan, N.B. Phillips, B.J. Smith, J. Whittaker, N.P. Wickramasinghe, L.J. Whittaker, V. Pandeyarajan, Z. Wan, S.P. Yadav, J.M. Carroll, N. Strokes, C.T. Roberts Jr, F. Ismail-Beigi, W. Milewski, D.F. Steiner, V.S. Chauhan, C.W. Ward, M.A. Weiss, M.C. Lawrence, Protective hinge in insulin opens to enable its receptor engagement, *Proc. Natl. Acad. Sci. USA* 111 (2014) E3395–E3404.
- [24] C.W. Ward, J.G. Menting, M.C. Lawrence, The insulin receptor changes conformation in unforeseen ways on ligand binding: sharpening the picture of insulin receptor activation, *Bioessays* 35 (2013) 945–954.
- [25] L. Zakova, E. Kletvikova, V. Veverka, M. Lepsik, C.J. Watson, J.P. Turkenburg, J. Jiracek, A.M. Brzozowski, Structural integrity of the B24 site in human insulin is important for hormone functionality, *J. Biol. Chem.* 288 (2013) 10230–10240.
- [26] J. Jiracek, L. Zakova, E. Antolikova, C.J. Watson, J.P. Turkenburg, G.G. Dodson, A.M. Brzozowski, Implications for the active form of human insulin based on the structural convergence of highly active hormone analogues, *Proc. Natl. Acad. Sci. USA* 107 (2010) 1966–1970.
- [27] T. Glendorf, A.R. Sorensen, E. Nishimura, I. Pettersen, T. Kjeldsen, Importance of the solvent-exposed residues of the insulin B-chain alpha-helix for receptor binding, *Biochemistry* 47 (2008) 4743–4751.
- [28] J. Whittaker, L.J. Whittaker, C.T. Roberts, N.B. Phillips, F. Ismail-Beigi, M.C. Lawrence, M.A. Weiss, α -Helical element at the hormone-binding surface of the insulin receptor functions as a signaling element to activate its tyrosine kinase, *Proc. Natl. Acad. Sci. USA* 109 (2012) 11166–11171.
- [29] J.G. Menting, C.F. Lawrence, G.K. Kong, M.B. Margetts, Ward, C.W., M.C. Lawrence, Structural congruency of ligand binding to the insulin and Insulin/Type 1 insulin-like growth factor hybrid receptors, *Structure* 23 (2015) 1–12.
- [30] A. Papaioannou, S. Kuyucak, Z. Kuncic, Molecular dynamics simulations of insulin: elucidating the conformational changes that enable its binding, *PLoS ONE* 10 (12) (2015) e0144058.
- [31] L. Zakova, E. Kletvikova, M. Lepsik, M. Collinsov, C.J. Watson, J.P. Turkenburg, J. Jiracek, A.M. Brzozowski, Human insulin analogues modified at the B26 site reveal a hormone conformation that is undetected in the receptor complex, *Acta Crystallogr. D Biol. Crystallogr.* 70 (2014) 2765–2774.
- [32] T. Croll, B.J. Smith, J. Whittaker, M.B. Margetts, M.A. Weiss, C.W. Ward, M.C. Lawrence, Higher-resolution structure of the human insulin receptor ectodomain: multi-modal inclusion of the insert domain, *Structure* 24 (2016) 469–476.
- [33] B. Xu, K. Huang, Y.C. Chu, S.Q. Hu, S. Nakagawa, S. Wang, R.Y. Wang, J. Whittaker, P.G. Katsoyannis, M.A. Weiss, Decoding the cryptic active conformation of a protein by synthetic photocrosslinking: Insulin inserts a detachable arm between receptor domains, *J. Biol. Chem.* 284 (2009) 14597–14608.
- [34] H. Tager, B. Give, B. Baldwin, M. Mako, J. Markese, A. Rubenstein, J. Olefsky, M. Kobayashi, O. Kolterman, R. Poucher, A structurally abnormal insulin causing human diabetes, *Nature* 281 (1979) 122–125.
- [35] M. Haneda, S.J. Chan, S.C.M. Kwok, A.H. Rubenstein, D.F. Steiner, Studies on mutant human insulin gene: identification and sequence analysis of a gene encoding Ser B24 insulin, *Proc. Natl. Acad. Sci. USA* 80 (1980) 6366–6370.
- [36] S. Shoelson, M. Fickova, M. Haneda, A. Nahum, G. Musso, E.T. Kaiser, A.H. Rubenstein, H. Tager, Identification of a mutant human insulin predicted to contain a serine-for-phenylalanine substitution, *Proc. Natl. Acad. Sci. USA* 80 (1983) 7390–7394.
- [37] S. Shoelson, M. Haneda, P. Blix, A. Nanjo, T. Sanke, K. Inouye, D. Steiner, A. Rubenstein, H. Tager, Three mutant insulin in man, *Nature* 302 (1983) 540–543.
- [38] K. Nanjo, T. Sanke, M. Miyano, K. Okai, R. Sowa, M. Kondo, S. Nishimura, K. Iwo, K. Miyamura, B.D. Given, Diabetes due to secretion of a structurally abnormal insulin (insulin Wakayama), *J. Clin. Invest.* 77 (1986) 514–519.
- [39] H. Sakura, Y. Iwamoto, Y. Sakamoto, T. Kuzuya, H. Hirata, Structurally abnormal insulin in a diabetic patient: patient characterization of the mutant insulin A3(Val-to-Leu) isolated from the pancreas, *J. Clin. Invest.* 78 (1986) 1666–1672.
- [40] K. Nanjo, M. Miyano, M. Kondo, T. Sanke, S. Nishimura, K. Miyamura, K. Inouye, B.D. Given, S.J. Chan, K.S. Polonsky, Insulin Wakayama: familial mutant insulin syndrome in Japan, *Diabetologia* 30 (1987) 87–92.
- [41] T. Hayashi, Y. Nozaki, M. Nishizuka, M. Ikawa, S. Osada, M. Imagawa, 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 (2011) 1257–1263.
- [42] S.H. Nakagawa, H.S. Tager, Importance of aliphatic side-chain structure at positions 2 and 3 of the insulin A chain in insulin-receptor interactions, *Biochemistry* 31 (1992) 3204–3214.
- [43] G. Weitzel, F.U. Bauer, K. Eisele, 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 (1978) 945–958.
- [44] A. Wollmer, W. Strassburger, U. Glatter, G.G. Dodson, M. McCall, H.G. Gattner, W. Danho, D. Brandenburg, W. Rittel, 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 (1981) 581–592.
- [45] Z.L. Wan, K. Huang, B. Xu, S.Q. Hu, S. Wang, Y.C. Chu, P.G. Katsoyannis, M.A. Weiss, Diabetes-associated mutations in human insulin: crystal structure and photo-cross-linking studies of a-chain variant insulin Wakayama, *Biochemistry* 44 (2005) 5000–5016.
- [46] V. Pandeyarajan, B.J. Smith, N.B. Phillips, L. Whittaker, G.P. Cox, N. Wickramasinghe, J.G. Menting, Z.L. Wan, J. Whittaker, F. Ismail-Beigi, M.C. Lawrence, M.A. Weiss, Aromatic anchor at an invariant hormone-receptor interface: function of insulin residue B24 with application to protein design, *J. Biol. Chem.* 289 (2014) 34709–34727.
- [47] L. Zakova, T. Barth, J. Jiracek, J. Barthova, S. Zorad, Shortened insulin analogues: marked changes in biological activity resulting from replacement of TyrB26 and N-methylation of peptide bonds in the C-terminus of the B-chain, *Biochemistry* 43 (2004) 2323–2331.
- [48] L. Zakova, L. Kazdova, I. Hanclova, E. Protivinska, M. Sanda, M. Budesinsky, J. Jiracek, Insulin analogues with modifications at position B26. Divergence of binding affinity and biological activity, *Biochemistry* 47 (2008) 5858–5868.
- [49] V. Pandeyarajan, N.B. Phillips, N. Rege, M.C. Lawrence, J. Whittaker, M.A. Weiss, Contribution of TyrB26 to the function and stability of insulin: Structure-activity relationships at a conserved hormone-receptor interface, *J. Biol. Chem.* 291 (2016) 12978–12990.
- [50] L. Kosinova, V. Veverka, P. Novotna, M. Collinsova, M. Urbanova, N.R. Moody, J.P. Turkenburg, J. Jiracek, A.M. Brzozowski, L. Zakova, Insight into the structural and biological relevance of the T/R transition of the N-terminus of the B-chain in human insulin, *Biochemistry* 53 (2014) 3392–3402.
- [51] V. Zoete, M. Meuwly, M. Karplus, A comparison of the dynamic behavior of monomeric and dimeric insulin shows structural rearrangements in the active monomer, *J. Mol. Biol.* 342 (2004) 913–929.
- [52] F.S. Legge, A. Budi, H. Treutlein, I. Yarovsky, Protein flexibility: multiple molecular dynamics simulations of insulin chain B, *Biophys. Chem.* 119 (2006) 146–157.
- [53] M.A. Islam, S. Bhayye, A.A. Adeniyi, M.E. Soliman, T.S. Pillay, Diabetes mellitus caused by mutations in human insulin: analysis of impaired receptor binding of insulins Wakayama, Los Angeles and Chicago using pharmacoinformatics, *J. Biomol. Struct. Dyn.* 14 (2016) 1–14.
- [54] V. Zoete, M. Meuwly, M. Karplus, Study of the insulin dimerization: binding free energy calculations and per-residue free energy decomposition, *Proteins* 61 (2005) 79–93.
- [55] A. Papaioannou, S. Kuyucak, Z. Kuncic, Elucidating the activation mechanism of the insulin-family proteins with molecular dynamics simulations, *PLoS ONE* 11 (8) (2016) e0161459.
- [56] C. Mueller-Dieckmann, S. Panjikar, A. Schmidt, S. Mueller, J. Kuper, A. Geerloff, M. Willmanns, R.K. Singh, P.A. Tucker, M.S. Weiss, On the routine use of soft X-rays in macromolecular crystallography. Part IV. Efficient determination of anomalous substructures in biomacromolecules using longer X-ray wavelengths, *Acta Crystallogr. D Biol. Crystallogr.* 63 (2007) 366–380.
- [57] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, *J. Mol. Graph.* 14 (1996) 33–38.
- [58] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, Scalable molecular dynamics with NAMD, *J. Comput. Chem.* 26 (2005) 1781–1802.
- [59] A.D. MacKerell, B. Brooks, C.L. Brooks, L. Nilsson, B. Roux, Y. Won, M. Karplus, CHARMM: The energy function and its parameterization with an overview of the program, in *The Encyclopedia of Computational Chemistry*, Schleyer, P. v. R., et al., editors. John Wiley and Sons, Chichester. pp. 271–277, 1998.
- [60] B.R. Brooks, C.L. Brooks III, A.D. Jr Mackerell, L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caffisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, M. Karplus, CHARMM: the biomolecular simulation program, *J. Comput. Chem.* 30 (2009) 1545–614.
- [61] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera—a visualization system for exploratory research and analysis, *J. Comput. Chem.* 25 (2004) 1605–12.
- [62] *Matlab and Statistics Toolbox Release R, The MathWorks, Inc., Natick, Massachusetts, United States, 2015b.*
- [63] S. Hayward, B.L. de Groot, Normal modes and essential dynamics, *Methods Mol. Biol.* 443 (2008) 89–106.
- [64] A.E. Garcia, Large-amplitude nonlinear motions in proteins, *Phys. Rev. Lett.* 68 (1992) 2696–2699.
- [65] A. Amadei, A.B.M. Linssen, H.J.C. Berendsen, Essential dynamics of proteins, *Protein: Struct. Funct. Genet.* 17 (1993) 412–425.
- [66] R. Abseher, L. Horstink, C.W. Hilbers, M. Nilges, Essential spaces defined by NMR structure ensembles and molecular dynamics simulation show significant overlap, *Proteins* 31 (1998) 370–382.
- [67] L.S. Caves, J.D. Evanseck, M. Karplus, Locally accessible conformations of proteins: multiple molecular dynamics simulations of crambin, *Protein Sci.* 7 (1998) 649–666.
- [68] K.M. ElSawy, M.K. Hodgson, L.S. Caves, The physical determinants of the DNA conformational landscape, *Nucleic Acids Res.* 33 (2005) 5749–5762.
- [69] D.M. van Aalten, D.A. Conn, B.L. de Groot, H.J. Berendsen, J.B. Findlay, A. Amadei, Protein dynamics derived from clusters of crystal structures, *Biophys. J.* 73 (1997) 2891–2896.
- [70] B.J. Grant, A.P. Rodrigues, K.M. ElSawy, J.A. McCammon, L.S. Caves, Bio3d: an R package for the comparative analysis of protein structure, *Bioinformatics* 22 (2006) 2695–2696.
- [71] K. Pearson, Notes on regression and inheritance in the case of two parents, *Proc. R. Soc. Lond.* 58 (1895) 240–242.