Published in final edited form as:

Proteins. 2015 April ; 83(4): 662–669. doi:10.1002/prot.24759.

Evaluation of conformational changes in diabetes-associated mutation in insulin A chain: A molecular dynamics study

Yong Hwan Kim^a, Kevin Kastner^b, Badi Abdul-Wahid^b, and Dr. Jesús A. Izaguirre^{a,b}

^a Bioengineering Graduate Program, University of Notre Dame, Notre Dame, IN

^b Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, IN

Abstract

Insulin plays a central role in the regulation of metabolism in humans. Mutations in the insulin gene can impair the folding of its precursor protein, proinsulin, and cause permanent neonatal-onset diabetes mellitus known as Mutant INS-gene induced Diabetes of Youth (MIDY) with insulin deficiency. To gain insights into the molecular basis of this diabetes-associated mutation, we perform molecular dynamics simulations in wild-type and mutant (Cys^{A7} to Tyr or C(A7)Y) insulin A chain in aqueous solutions. The C(A7)Y mutation is one of the identified mutations that impairs the protein folding by substituting the cysteine residue which is required for the disulfide bond formation. A comparative analysis reveals structural differences between the wild-type and the mutant conformations. The analyzed mutant insulin A chain forms a metastable state with major effects on its N-terminal region. This suggests that MIDY mutant involves formation of a partially folded intermediate with conformational change in N-terminal region in A chain that generates flexible N-terminal domain. This may lead to the abnormal interactions with other proinsulins in the aggregation process.

Keywords

Computer simulation; conformation; dynamic; insulin; mutant; metastable state

INTRODUCTION

Insulin is a peptide hormone that regulates blood glucose level and mutations in the insulin gene can lead to the development of diabetes mellitus.^{1,2} Insulin consists of two chains, an A chain (21 residues) and B chain (30 residues), connected by disulfide bonds.³ The native structure of insulin has three α -helices, two in the A chain and one in the B chain. The chains are connected by two interchain disulfide bonds (A7-B7 and A20-B19) and one intrachain disulfide bond in A chain (A6-A11). The hormone is generated by proteolytic conversion of a single-chain precursor, proinsulin (86 residues).⁴ A C chain that connects the C terminus of the B-Chain to the N terminus of the A Chain in proinsulin gets cleaved out in post-Golgi vesicles. The mature hormone is stored within secretory granules of

pancreatic β -cells as hexamers stabilized by zinc and released into the bloodstream in a monomeric form.⁵

There have been many reports of insulin-deficient diabetes caused by heterozygous mutations in the insulin gene.⁶⁻⁸ Although the mutation is heterozygous, it negatively affects the rest of the wild-type insulin production and causes a significant number of cases of permanent neonatal-onset diabetes mellitus known as Mutant INS-gene induced Diabetes of Youth (MIDY)^{9,10} which involves protein misfolding. MIDY mutations impair proinsulin folding and block the formation of wild-type insulin by recruiting the wild-type proinsulins to MIDY mutant proinsulins with exposure of thiol group of cysteine in mice^{11,12} and human.¹³ MIDY mutant triggers increased amounts of improper disulfide-linked protein complexes exhibiting a dominant negative effect with the presence of an unpaired cysteine within proinsulin. This disrupts trafficking from endoplasmic reticulum (ER), which promotes the ER stress and leads to the apoptosis of β -cells eventually.¹⁴

The finding of misfolded proinsulin-mediated diabetes has drawn attention to the aggregation process of proinsulin and insulin. Protein aggregation often involves partially folded conformation with conformational changes. Several studies have shown that structural changes in the native state introduces instability which is required for aggregation.¹⁵ In one paper, it was hypothesized that cysteinyl mutation would generate a partially folded intermediate with differences in secondary structure, such as loss of α -helix in A chain, which is usually shown during protein aggregation due to the decrease in thermostability. To test this hypothesis in molecular detail, we performed MD simulations in aqueous solutions. Molecular dynamics (MD) studies have been conducted in numerous studies to investigate the dynamics of proteins.¹⁶⁻²⁴ Due to the computational cost of simulating full length proinsulin in explicit solvent, we simulated fragments of proinsulin.

We studied the relation between mutant insulin A chain and the changes in protein conformation. The mutant of interest in this MD study, the C(A7Y)Y mutant, is an identified MIDY mutation that impairs the proper protein folding by disturbing the disulfide bond formation with cysteine residue substitution and triggers abnormal interaction with other proteins. Thus, structural analysis of conformations of this mutation from MD simulations would provide hints to its aggregation mechanism. Our comparative study of wild-type and mutant conformations produced useful information regarding the structural changes that may affect the abnormal interactions.

METHODS

MD Simulations

All simulations of insulin A chain were performed and analyzed with the GROMACS software package (version 4.5.3).²⁵ The starting conformation was extracted from the folded proinsulin structure (PDB ID:2KQP)²⁶ and unfolded structure was obtained by running the unfolded structure generator developed by the Freed and Sonicks Groups at University of Chicago.²⁷ UCSF Chimera software²⁸ was used to generate a homology model of the mutant conformation with folded wild-type proinsulin as the template. The AMBER99SB-ILDN-NMR forcefield²⁹ was used. (Data from simulations using other forcefields and

solvent models can be found in the Figure S1 in Supporting Information.) The systems were solvated in a cubic box (box vector 5.5 nm) with TIP3P water molecules. The systems were neutralized by adding Na^+ and Cl^- ions (salt concentration 0.15 mol L^{-1}) and energy-minimized until the maximal force on individual atoms was smaller than $1 \text{ kJ mol}^{-1} \text{ nm}^{-1}$. After minimization, restrained MD runs were performed for 10 ps to fix heavy atoms to their initial positions with a force constant of $1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$. After restrained MD runs, simulations were equilibrated for 2 ns in NVT (constant number of particles, volume, and temperature) followed by 200 ns production runs with a constant number of particles, pressure (P) and temperature. The equation of motion was integrated using a stochastic dynamics integrator with a 2 fs time step. A constant pressure of 1 bar was applied with a coupling constant of 1.0 ps with Parrinello-Rahman coupling scheme.³⁰ Protein, water molecules and ions were coupled separately to a bath at 310 K with a coupling constant of 0.2 ps. The LINCS³¹ and SETTLE³² algorithms were used to constrain all bond lengths and water molecules respectively. The cutoff radius for the Lennard-Jones interactions was set to 1.4 nm. Long-range electrostatic forces were treated with reaction-field³³ with cutoff value of 1.4 nm and epsilon-rf (reaction field dielectric constant) of 78. A frame was extracted every 10 ns of each simulation.

There were a total of 12 independent simulations of the wild-type conformation and 13 independent simulations of the mutant conformation. Simulations used five different starting conformations for wild-type and four different starting conformations for mutant, each with their own initial velocities and ran for 200 ns. There was one simulation of wild-type and two simulations of mutant, each with different initial velocities, using the unfolded conformation obtained by using the unfolded structure generator. For the wild-type, seven additional simulations were run using three different conformations obtained from this trajectory. There were four additional simulations using another conformation of wild-type obtained from one of these seven wild-type trajectories. For the mutant, eleven additional simulations were run using three conformations obtained from the initial trajectory (see Table S1 in Supporting Information). Each conformation, with the exception of the conformations generated by unfolded structure generator, was extracted based on a specific root mean square deviation (RMSD) of interest (3.5 Å, 4.5 Å, 6 Å or 7 Å) compared to the folded structure to explore the protein dynamics in detail by using various states. These conformations were visualized in 3D to examine the validity of structures. New simulations were started with different initial velocities.

Simulation Analysis

We used the RMSD of α -carbon ($\text{C}\alpha$) coordinates of each conformation as a measure of structural similarity between simulated A chain and the native. For RMSD calculation of a mutant A chain, $\text{C}\alpha$ coordinates of the conformation were compared to those of the folded mutant A chain structure. Since $\text{C}\alpha$ RMSD calculation excludes side chain position, α -helix formation has a small effect on $\text{C}\alpha$ RMSD.

The mean-first-passage time (MFPT) was calculated to provide a qualitative measure of the time to fold (see Supporting Information). The MFPT was calculated using a Bayesian

modification of maximum likelihood estimate.³⁴ With n folding simulations out of N total simulations and total required simulation time Θ , the estimators for the mean rate $\langle k \rangle$ is:

$$\langle k \rangle = \frac{n+1}{\Theta} \quad (1)$$

Θ is the sum of the first-passage times for simulations generating folded structures plus the total simulation time of simulations that did not form folded structure.³⁵ The MFPT is returned as $\langle k \rangle^{-1}$.

For clustering of conformations from MD to identify stable states and producing free energy landscapes of systems, `g_cluster` and `g_sham` modules provided by the GROMACS software suite were used. The energy values that are given in the results are the Gibbs-Free energies at 310 K. Structures are considered “folded” at an RMSD cutoff of 0.3 nm for clustering and MFPT calculation. `VMD`³⁶ and `Chimera` applications were used for molecular visualizations. To investigate the structural segments in the insulin A chain, the secondary structures for each conformation were assigned using the Dictionary of Protein Secondary Structure (DSSP) program.³⁷ Stabilization of the α -helix structure was examined by calculating the percentage of conformations with α -helices in each residue.

RESULTS AND DISCUSSION

RMSD Distributions of Wild-Type and Mutant

Here we perform MD simulations to investigate the behavior of the wild-type and mutant (C(A7)Y) insulin A chain. The structural changes of wild-type and mutant insulin A chain during the simulation were investigated by analyzing the C α RMSD distributions shown in Figure 1. Their distribution patterns indicate that formation of the folded conformation is favored and stabilized in a nearly folded state in the mutant. It shows that there are many frames in the mutant trajectories with low RMSDs compared to wild-type. 35% of the mutant conformations have RMSD values between 3.0 Å and 3.9 Å whereas 16% of the wild-type conformations have corresponding RMSD values. For RMSDs lower than 3 Å which are considered to be in the folded state, there are 1.1% of the mutant conformations but these are a lower fraction of conformations compared to those of wild-type (2.3%). Having more compacted conformations with RMSDs between 3 and 5 Å, it may imply that mutant quickly compacts compared to the wild-type although it takes a longer time to fully fold.

Comparison of MFPT in Wild-Type and Mutant

In order to compare the folding time, we calculated MFPTs for wild-type and mutant based on the RMSDs of conformations. As shown in Table 1, MFPTs for wild-type and mutant are 365 ns and 586 ns respectively when 0.3 nm is used as the threshold value for counting conformations as folded. Even though there are many mutant conformations with RMSDs near 0.3 nm, wild-type has a faster MFPT because there are more fractions of conformations below 0.3 nm RMSD in the wild-type. When 0.4 nm is used as the threshold value, MFPT for the mutant is faster (97 ns) than wild-type (106 ns). This suggests that the mutant compacts faster than the wild-type but it stays in the partially folded state longer than the

wild-type. By staying in the problematic state longer, the chance for the mutant to form an improper disulfide bond with another proinsulin would be increased. However, this comparison of MFPT should be interpreted with caution since the starting conformations were manually selected.

Cluster Analysis of the Wild-Type and Mutant Trajectories

To identify metastable states in the trajectories in detail, the conformations sampled during the MD were classified. 78 and 111 clusters were found from the wild-type and the mutant respectively. The most populated representative conformation of insulin A chain for wild-type and mutant according to the cluster analysis are displayed in Figure 2. The most populated conformation of unfolded wild-type (22% of structures) almost reproduces the native state whereas the mutant (25% of structures) generates conformation close to the native state but without the N-terminal helix. From Figure 2, it can be seen that thiol groups in cysteine residues of the wild-type and mutant were positioned to about the same directions compared to those of native state. In the native state, thiol groups of A6 and A10 which form an intrachain disulfide bond were pointing inward and the thiol group of A7 and A20 were exposed. Orientations of thiol groups of the wild-type were about the same to those of the native state. In the mutant, A7 tyrosine residue is exposed to the surface. With its flexible N-terminus and exposed tyrosine residue, the mutant A chain could be responsible for the abnormal interactions which may involve the recruitment of another proinsulin and form the improper A7-B7 disulfide bond to the lone pair of B7 cysteine residue of the mutant.

Free Energy Landscape of the Wild-Type and Mutant

In order to further elucidate the states, free energy landscapes for wild-type and mutant are constructed as a function of the radius of gyration, RMSD and Gibbs free energy in Figure 3 (Free energy landscape in three dimensions are shown in Figure S2 in Supporting Information). It shows a relatively linear pattern for the wild-type insulin A chain whereas a featureless pattern with wider distribution of conformations is shown for the mutant. The lowest free energy conformations were extracted from the trajectory for analysis. There are eight conformations of the lowest energy (0.64 kJ mol^{-1}) for the wild-type and seven conformations below 1 kJ mol^{-1} with the lowest energy of 0.29 kJ mol^{-1} for the mutant. Three of the lowest free energy structure of wild-type have the correct orientations of N-terminus and C-terminus and helix location with two of these conformations reproducing the native state. No mutant conformations with the lowest energy generate the native state. Four of the lowest energy conformations of the mutant form extended C-terminal α -helices without an N-terminal α -helix and two conformations form partial N-terminal α -helices with extended C-terminal α -helices.

Partially folded intermediates are formed in intrinsically disordered proteins (IDP) which can progress to aggregation.³⁸ In addition, simulated conformations generated various compacted conformations differing in RMSDs indicating structural heterogeneity, which is another feature of intrinsically disordered protein.³⁹ Forming a partially folded intermediate with a flexible N-terminus and heterogeneous structures are characteristics of an intrinsically

disordered protein and the suggested intermediate can be developed into the problematic state.

Comparison of Secondary Structure Formation in Wild-Type and Mutant

To investigate the structural changes in detail, we examined the secondary structure dynamics. In Figure 4, we report the population of each residue in insulin A chain in extracted frames adopting α -helix of the DSSP program's secondary structure to see the overall helical propensities. Mutation altered the α -helix propensity quite significantly in the N-terminus. It was previously reported that the substitution of the A7 cysteine residue with serine resulted in the loss of the N-terminal α -helix in A2-A8 residues in an *in vitro* studies.⁴⁰ It is possible that the effect of mutation on the structures is due to the properties of either the absent cysteine or the substituted tyrosine, which is a polar and aromatic residue. In addition to the mutant generating a relatively low helical population in the N-terminus region, the mutant A chain forms an extended α -helix structure. High helical propensities in A9-A12 (a non-helical segment in wild-type) are present extended from the canonical C-terminal α -helix (residue numbers A13-A19).

To provide more insight into the changes of α -helix propensity due to the mutation, we plotted the percentage of simulations which formed an α -helix segment for each residue at least once in Figure 5. In the mutant simulations, about 50% of the simulations do not form an α -helix structure throughout their entire trajectories in the α -helical region in the N-terminus. It suggests that the mutation in the A7 residue interferes with α -helix formation in these residues.

We also report the population of each residue in the insulin A chain over the simulation time to examine the time dependence of residue-dependent α -helix propensity (Supporting Information Figure S3). For the first 100 ns, the difference in populations between wild-type and mutant is not noticeable. From 100-200 ns, there is a distinct difference in populations between wild-type and mutant and it continues to the end of the simulation (400 ns). In addition to the formation of α -helix segments in A2-A8 residues in less than 20% of the populations, high helical propensities are observed in A9-A12 (a non-helical segment in wild-type) extended from the C-terminal α -helix region in the mutant throughout the simulations. The C-terminal α -helix is largely conserved in all the simulations. This suggests that the C-terminus α -helix is a very stable region and possibly has an important role in initial folding of A chain. Also, our results suggests that the partially folded intermediate of C(A7)Y mutant forms after the first disulfide bond formation between residue number A20 and B19 during an early intermediate state⁴¹ since it is hard to form a disulfide bond after formation of α -helix in the C-terminus region.

The large number of simulated conformations of the mutant have similar structures to the native folded state but many of them do not form the N-terminal α -helix region where the mutation was made. Substantial data indicate that a relatively unfolded structure is crucial in the fibrillation and aggregation process which often progresses to the disease states.³⁸ There were several studies showing that protein aggregation is accompanied by changes in secondary and tertiary structures of the protein.^{42,43} Querol et al showed that reduction in helicity decreased thermostability.⁴² It was shown in an *in vitro* study that the aggregation-

induced increase in β -sheet content is often accompanied by a drop in α -helix structure.⁴³ We speculate that this conformation missing the N-terminal α -helix structure is one of the partially folded intermediates that can trigger aggregation by interacting and recruiting another protein with its flexible N-terminus.

CONCLUSIONS

In the present work, we report MD simulation results for MIDY mutant insulin A chain with comparison of conformational changes to that of wild-type. These simulation results enabled us to characterize the possible initiation conformation of MIDY whose mechanism and conformation are not well understood. Our study revealed the effect of substitution of a cysteine residue in the stability of protein conformation. MIDY mutation (C(A7)Y mutation) destabilized α -helix structure in the N-terminus region (residues 2 to 8) where an *in vitro* study previously reported that C(A7)S mutation resulted in the loss of N-terminal α -helix structure. In contrast, wild-type had a higher α -helix propensities in N-terminus region following the native structure. In addition to it, the mutant formed an extended α -helix region in the latter half of the chain. Instead of canonical α -helix region from residue 13 to 19, it had higher α -helix propensities in residues from 9 to 19. In addition to the differences in secondary structure formation, several characteristics of intrinsically disordered protein such as formation of a partially folded intermediate and flexible N-terminus from the loss of α -helix structure and structural heterogeneity were observed. IDPs are commonly involved in the human diseases. Also, the mutant is observed to stay in the partially folded state longer than the wild-type which would increase the chance of undesirable interactions as shown in the MFPT and cluster analysis.

Protein aggregation accompanies a partially folded intermediate with changes in secondary structure such as the loss of α -helix secondary structure. Our simulation results suggest the conformation of mutant partially folded intermediate with a flexible N-terminus region and extended C-terminal α -helix as one of the potential candidates for a partially folded intermediate of A chain that could trigger protein aggregation. Its flexible N-terminus can be responsible for abnormal interactions that could result in recruitment of a proinsulin to the lone pair of cysteine in B7.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The authors are grateful to O. Wiest (University of Notre Dame) for very helpful discussions. We also thank the Center for Research Computing for computational resources. These results were partially funded by NIH grant R01 GM101935 to JI.

REFERENCES

1. Vinik A, Bell G. Mutant Insulin Syndromes. *Horm Metab Res.* 1988; 20:1–10. [PubMed: 3286444]
2. Steiner DF, Tager HS, Chan SJ, Nanjo K, Sanke T, Rubenstein AH. Lessons learned from molecular biology of insulin-gene mutations. *Diabetes Care.* 1990; 13:600–609. [PubMed: 2192846]

3. Steiner DF, Chan SJ. An overview of insulin evolution. *Horm Metab Res.* 1988; 20:443–444. [PubMed: 3049292]
4. Steiner DF. Evidence for a precursor in the biosynthesis of insulin. *NY Acad Sci.* 1967; 30:60–68.
5. Dodson G, Steiner DF. The role of assembly in insulin's biosynthesis. *Curr Opin Struct Biol.* 1998; 8:189–194. [PubMed: 9631292]
6. Stoy J, Edghill EL, Flanagan SE, Ye H, Paz VP, Pluzhnikov A, Below JE, Hayes MG, Cox NJ, Lipkind GM, Lipton RB, Greeley SAW, Patch AM, Ellard S, Steiner DF, Hattersley AT, Philipson LH, Bell GI. Neonatal Diabetes International Collaborative Group. Insulin gene mutations as a cause of permanent neonatal diabetes. *Proc Natl Acad Sci USA.* 2007; 104:15040–15044. [PubMed: 17855560]
7. Colombo C, Porzio O, Liu M, Massa O, Vasta M, Salardi S, Beccaria L, Monciotti C, Toni S, Pedersen O, Hansen T, Federici L, Pesavento R, Cadario F, Federici G, Ghirri P, Arvan P, Iafusco D, Barbetti F. Early Onset Diabetes Study Group of the Italian Society of Pediatric Endocrinology and Diabetes. Seven mutations in the human insulin gene linked to permanent neonatal/infancy-onset diabetes mellitus. *J Clin Invest.* 2008; 118:2148–2156. [PubMed: 18451997]
8. Polak M, Dechaume A, Cave H, Nimri R, Crosnier H, Sulmont V, de Kerdanet M, Scharfmann R, Lebenthal Y, Froguel P, Vaxillaire M, French ND. (Neonatal Diabetes) Study Group. Heterozygous missense mutations in the insulin gene are linked to permanent diabetes appearing in the neonatal period or in early infancy: a report from the French ND (Neonatal Diabetes) Study Group. *Diabetes.* 2008; 57:1115–1119. [PubMed: 18171712]
9. Weiss MA. Proinsulin and the Genetics of Diabetes Mellitus. *J Biol Chem.* 2009; 284:19159–19163. [PubMed: 19395706]
10. Liu M, Hodish I, Haataja L, Lara-Lemus AR, Rajpal G, Wright J, Arvan P. Proinsulin misfolding and diabetes: mutant INS gene-induced diabetes of youth. *Trends Endocrinol Metabolism.* 2010; 21:652–659.
11. Yoshioka M, Kayo T, Ikeda T, Koizumi A. A novel locus, Mody4, distal to D7Mit189 on chromosome 7 determines early-onset NIDDM in nonobese C57BL/6 (Akita) mutant mice. *Diabetes.* 1997; 46:887–894. [PubMed: 9133560]
12. Wang J, Takeuchi T, Tanaka S, Kubo S, Kayo T, Lu D, Takata K, Koizumi A, Izumi T. A mutation in the insulin 2 gene induces diabetes with severe pancreatic beta-cell dysfunction in the Mody mouse. *J Clin Invest.* 1999; 103:27–37. [PubMed: 9884331]
13. Liu M, Haataja L, Wright J, Wickramasinghe N, Hua Q, Phillips N, Barbetti F, Weiss M, Arvan P. Mutant INS-Gene Induced Diabetes of Youth: Proinsulin Cysteine Residues Impose Dominant Negative Inhibition on Wild-Type Proinsulin Transport. *PLoS ONE.* 2010; 5:e13333. [PubMed: 20948967]
14. Liu M, Hodish I, Rhodes CJ, Arvan P. Proinsulin maturation, misfolding, proteotoxicity. *Proc Natl Acad Sci USA.* 2007; 104:15841–15846. [PubMed: 17898179]
15. Ahmad A, Millett I, Doniach S, Uversky V, Fink A. Partially folded intermediates in insulin fibrillation. *Biochemistry.* 2003; 42:11404–11416. [PubMed: 14516191]
16. Karanicolas J, Brooks CL III. Integrating folding kinetics and protein function: Biphasic kinetics and dual binding specificity in a WW domain. *Proc Natl Acad Sci.* 2004; 101:3432–3437. [PubMed: 14981252]
17. Lin YS, Bowman GR, Beauchamp KA, Pande VS. Investigating how peptide length and a pathogenic mutation modify the structural ensemble of amyloid beta monomer. *Biophys J.* 2012; 102:315–324. [PubMed: 22339868]
18. Lindorhff-Larsen K, Pianna S, Dror RO, Shaw DE. How fast-folding proteins fold. *Science.* 2011; 334:517–520. [PubMed: 22034434]
19. Lu Q, Tan YH, Luo R. Molecular dynamics simulations of p53 DNA-binding domain. *J Phys Chem.* 2007; 111:11538–11545.
20. Zagrovic B, Snow CD, Shirts MR, Pande VS. Simulation of folding of a small alpha-helical protein in atomistic detail using worldwide-distributed computing. *J Mol Biol.* 2002; 323:927–937. [PubMed: 12417204]

21. Zhu X, Bora RP, Barman A, Singh R, Prabhakar R. Dimerization of the full-length Alzheimer amyloid beta-peptide (A β 42) in explicit aqueous solution: a molecular dynamics study. *J Phys Chem.* 2012; 116:4405–4416.
22. Chong LT, Snow CD, Rhee YM, Pande VS. Dimerization of the p53 Oligomerization Domain: Identification of a Folding Nucleus by Molecular Dynamics Simulations. *J Mol Biol.* 2005; 345:869–878. [PubMed: 15588832]
23. Dibeneditto D, Rossetti G, Caliendo R, Carloni P. A molecular dynamics simulation-based interpretation of nuclear magnetic resonance multidimensional heteronuclear spectra of α -synuclein dopamine adducts. *Biochemistry.* 2013; 52:6672–6683. [PubMed: 23964651]
24. Wu C, Scott J, Shea JE. Binding of Congo red to amyloid protofibrils of the Alzheimer A β _{9–40} peptide probed by molecular dynamics simulations. *Biophys J.* 2012; 103:550–557. [PubMed: 22947871]
25. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D, Hess B, Lindahl E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics.* 2013; 29:845–854. [PubMed: 23407358]
26. Yang Y, Hua QX, Liu J, Shimizu EH, Choquette MH, Mackin RB, Weiss MA. Solution structure of proinsulin: connecting domain flexibility and prohormone processing. *J Biol Chem.* 2010; 285:7847–7851. [PubMed: 20106974]
27. Jha AK, Colubri A, Freed K, Sosnick TR. Statistical coil model of the unfolded state: Resolving the reconciliation problem. *Proc Natl Acad Sci USA.* 2005; 102:13099–13104. [PubMed: 16131545]
28. Petterson E, Goddard T, Huang C, Couch G, Greenblatt D, Meng E, Ferrin T. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J Comput Chem.* 2004; 25:1605–1612. [PubMed: 15264254]
29. Li D, Bruschweiler R. NMR-Based Protein Potentials. *Angew Chem.* 2010; 122:6930–6932.
30. Parinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics methods. *J Appl Phys.* 1982; 52:7182.
31. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: a linear constraint solver for molecular simulations. *J Comp Chem.* 1997; 18:1463–1472.
32. Miyamoto S, Kollman PA. SETTLE: an analytical version of the SHAKE and RATTLE algorithms for rigid water models. *J Comp Chem.* 1992; 13:952–962.
33. Onsager L. Electric Moments of Molecules in Liquids. *J Am Chem Soc.* 1936; 58:1486–1493.
34. Zagrovic B, Pande V. Solvent viscosity dependence of the folding rate of a small protein: distributed computing study. *J Comp Chem.* 2003; 12:1432–6. [PubMed: 12868108]
35. Izaguirre J, Sweet C, Pande V. Multiscale dynamics of macromolecules using normal mode Langevin. *Pac Symp Biocomput.* 2010:240–251. [PubMed: 19908376]
36. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Molecular Graphics.* 1996; 14:33–38.
37. Kabsch W, Sander C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 1983; 12:2577–2637. [PubMed: 6667333]
38. Uversky V, Fink AL. Conformational constraints of amyloid fibrillation: the importance of being unfolded. *BBA-Proteins and Proteom.* 2004; 1698:131–153.
39. Uversky V. Unusual biophysics of intrinsic disordered proteins. *BBA-Proteins and Proteom.* 2013; 1834:932–951.
40. Hua Q, Nakagawa S, Jia W, Hu S, Chu Y, Katsoyannis P, Weiss M. Hierarchical Protein Folding: Asymmetric Unfolding of an Insulin Analogue Lacking the A7-B7 Interchain Disulfide Bridge. *Biochemistry.* 2001; 40:11299–12311.
41. Jia XY, Guo Z, Wang Y, Xu Y, Duan S, Feng Y. Peptide models of four possible insulin folding intermediates with two disulfides. *Protein Science.* 2003; 12:2412–2419. [PubMed: 14573855]
42. Querol E, Perez-Pons J, Mozo-Villarias A. Analysis of protein conformational characteristics related to thermostability. *Protein Engineering.* 1996; 9:265–271. [PubMed: 8736493]
43. Nielsen L, Frokjaer S, Brange J, Uversky V, Fink AL. Probing the mechanism of insulin fibril formation with insulin mutants. *Biochemistry.* 2001; 40:8397–8409. [PubMed: 11444987]

Figure 1.
RMSD distributions of the wild-type (left) and mutant (right).

Figure 2.

Conformation of the native A chain folded structure is shown on the left. The most populated conformations of wild-type (middle) and mutant (right) are shown with cysteine residues (A6, A7, A10 and A20) obtained from the unfolded structure simulations. A tyrosine residue for mutant in A7 is shown. Chimera software was used to visualize molecular structures.

Figure 3. Free energy landscapes of A chain wild-type (left) and mutant (right) as a function of radius of gyration and RMSD among conformations.

Figure 4. Percent of frames generating α -helix structures based on the DSSP analysis for insulin A chain wild-type (left) and mutant (right).

Figure 5.
Percent of simulations generating α -helix structures based on the DSSP analysis for insulin A chain wild-type (left) and mutant (right).

Table 1

The mean-first-passage time for the wild-type and mutant

Threshold Value (nm)	MFPT _{wild-type} (ns)	MFPT _{mutant} (ns)
0.3	365	586
0.4	106	97