Agric. Chem. Biotechnol. **41**(8), 568-571 (1998)

Article

A Comparison of Three Dimensional Structures of Insulin, Proinsulin and Preproinsulin Using Computer Aided Molecular Modeling

Mina Oh, K. Hun Mok[1] and Yoongho Lim*

Department of Applied Biology and Chemistry, Konkuk University, Seoul 133-701, Korea
[1]Biomolecular Structure Research Unit, Korea Research Institute of Bioscience and Biotechnology

Received November 26, 1998

The conformations of human insulin precursors, proinsulin and preproinsulin, are described in terms of molecular dynamics simulations. Despite the presence of the C-peptide and/or the signal peptide, molecular dynamics calculations utilizing the hydration shell model over a period of 500 ps indicate that the native conformations of the A and B chains are well conserved in both cases. These results further support the NMR spectroscopy results that the C-peptide is relatively disordered and does not influence the overall conformation of the native structure. The robustness of the native structure as demonstrated by experiment and simulation will permit future protein engineering applications, whereby the expression or purification yields can be improved upon sequence modification of the C-peptide and/or the signal peptide.

Keywords : *insulin, preproinsulin, proinsulin, computer aided molecular modeling.*

Insulin is a small, globular protein of molecular weight 5700 Da with two polypeptide chains, the A chain (21 residues) and B chain (30 residues). These two are joined by two disulfide bonds at A7-B7 and A20-B19, with a third disulfide bond existing within the A chain, A6-A11.[1)] Insulin is synthesized in the pancreatic β-cells as an inactive 103-residue single-chain precursor, preproinsulin, with has a 17-residue N-nerminal signal peptide that directs its passage into secretory vesicles. proteolytic removal of the signal peptide produces the 86-residue proinsulin, which is constituted of the A and B chains and the C peptide, a connecting peptide of 35 residues positioned between N-terminal of the A chain and C-terminal of the B chain. When the elelevation of blood glucose triggers insulin secretion, proinsulin is converted into active insulin.

Insulin exists as a monomer at low concentrations in neutral conditions.[2)] At higher concentrations it dimerizes[3-5)] and, in the presence of zinc, assembles further to a hexamer. The crystallographic structures of monomeric, dimeric, and hexameric insulin have all been determined.[6)]

Hua and Weiss suggested in 1991 that each structure of the A and B chains is conserved when preproinsulin is converted into insulin.[7)] This means that the existence of the flexible C chain positioned between the N-terminal of the A chain and the C-terminal of the B chain does not influence the structures of the A and B chains of preproinsulin. If this is true, the biosynthesis of the full C-peptide region of preproinsulin is not required per se to obtain functional human insulin. Thus, in order to commercially produce insulin time- and cost-efficiently, a shortened C-peptide may offer a better solution.

To enlighten these possibilities, the structural studies of native human insulin, proinsulin, and preproinsulin using Computer Aided Molecular Modeling (CAMM) were carried out in this study. Using the previously determined three dimensional crystallographic structure of des-B1 Phe bovine insulin as a template structure, the structure of human insulin was modeled through CAMM. A comparison of the partial structures of preproinsulin with the three dimensional structure of insulin based on CAMM is reported here.

Materials and Methods

The initial structure of native human insulin used in CAMM was taken from the X-ray structure of des-B1 Phe bovine insulin (Protein Data Bank accession number: 2INS),[8)] a dimer composed of 2 asymmetric units. The two molecules particularly differed in the spatial arrangement of the aromatic ring of B25 Phe. Molecule 1 was chosen for this study. Human and bovine insulins differed by three amino acid residues as shown in Fig. 1. To obtain

*Corresponding author
Phone: 82-2-450-3760; Fax: 82-2-3436-5776
E-mail: yoongho@kkucc.konkuk.ac.kr

Abbreviations: CAMM, computer aided molecular modeling; CVFF, consistent-valence forcefield; MD, molecular dynamics; PDB, protein data bank.

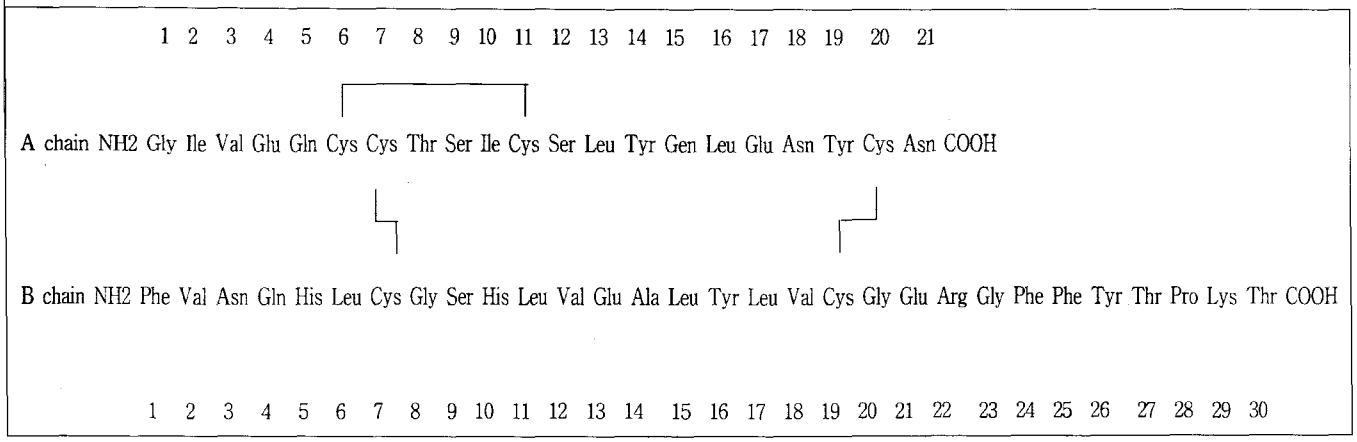

Fig. 1. The primary structure of human insulin. In bovine insulin, A8, A10 and B30 are replaced with Ala, Val and Ala, respectively.

the structure of human insulin, the residues A8 Ala, A10 Val, and B30 Ala of bovine insulin were replaced with A8 Thr, A10 Ile, and B30 Thr, respectively. Additionally, B1 Phe was attached to the N-terminal of the B chain. An MD simulation of human insulin in aqueous solution (5 Å layer of H_2O molecules) was subsequently conducted using the Discover v2.9 module of InsightII 95.0 software (Biosym/MSI). The potentials and partial charges were arranged using the consistent-valence forcefield (CVFF), and simulation was carried out for 500 ps. No other constraints except for the three disulfide bonds were given during the calculation. The three dimensional structures of preproinsulin and proinsulin were subsequently modeled using insulin. For the case of proinsulin, the primary structure of the C chain was initially built separately, then constrained within 8 Å prior to positioning the C-peptide between the N-terminal of the A chain and the C-terminal of the B chain, MD of proinsulin was peformed in a similar fashion as with insulin. MD of preproinsulin was accomplished by further attaching the signal peptide of 17 residues to the N-terminal of proinsulin.

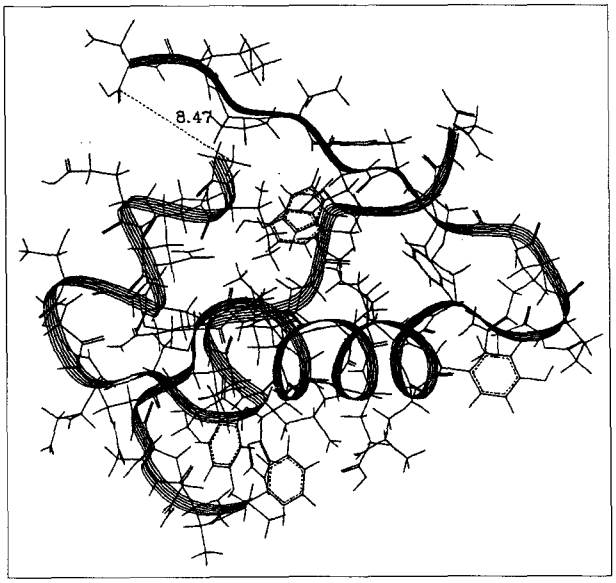

Fig. 2. The ribbon structure of native human insulin calculated from MD.

Results and Discussion

The MD calculations of native human insulin, proinsulin, and preproinsulin were conducted one by one as described above. The three dimensional structure of native human insulin obtained from these calculations (Fig. 2) was compared with the structure obtained from the X-ray crystallographic structure. The result shows a Root Mean Squares (RMS) value of 1.4 Å. Fig. 3 shows good agreement between the three dimensional structure of the X-ray determined molecule and that of the post-MD for native insulin. All residues were spatially superimposable in the backbones as well as the side chains except for the side chains of A21 Asn, B29 Lys, and B30 Thr. Only the side chains of the C-terminal regions of the A and B chains were placed in the opposite direction. The secondary structure of insulin obtained from the X-ray determined structure contained three α-helices at B9-B19, A1-A9, and A13-A20, and a β-turn and β-sheet at B20-B23 and B24-B28, respectively. Likewise, the secondary structure by MD calculation of native human insulin contained α-helices, β-turns, and β-strand at the same regions. These results also coincide with the secondary structure obtained from NMR in α-helix regions.

The structure of proinsulin was built by insertion of the C chain between the ends of the A and B chains. The constraint of 8 Å between the N-terminal of the A chain and the C-terminal of the B chain was applied during dynamics calculations in the presence of the C chain. The folding process of the protein could be observed during the dynamics calculation. The 3D structure of proinsulin obtained in this manner was superimposed with that of native insulin (Fig. 4), and the RMS value was determined to be 1.8 Å. The secondary structure by MD of proinsulin agreed well with that of the MD structure of native insulin. For example, the beginning

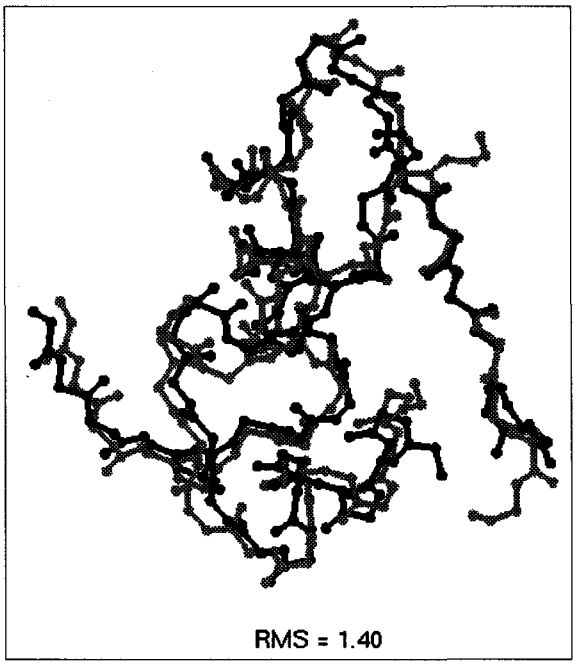

Fig. 3. A comparison of X-ray and MD structures of native human insulin.

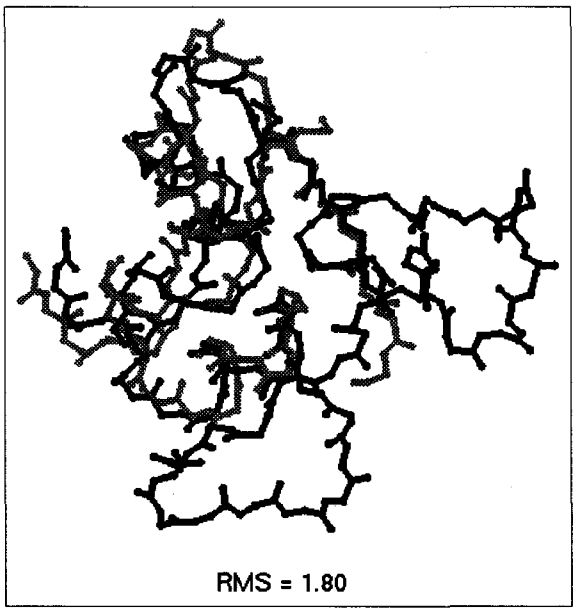

Fig. 4. A comparison of X-ray structure of native human insulin and MD structure of proinsulin.

and the end of helices in proinsulin were identical with those of native insulin. Therefore, it may be regarded that the existence of the C chain of proinsulin did not affect the structure of active insulin.

Preproinsulin was built by linking the signal peptide of 17 amino acids to the N-terminal of the B chain of proinsulin. The energy profile and structures obtained from the MD calculations are shown in Fig. 5. The signal peptide of preproinsulin folded together with the A and B chains in a gradual manner. The three dimensional structure of preproinsulin was superimposed with that of

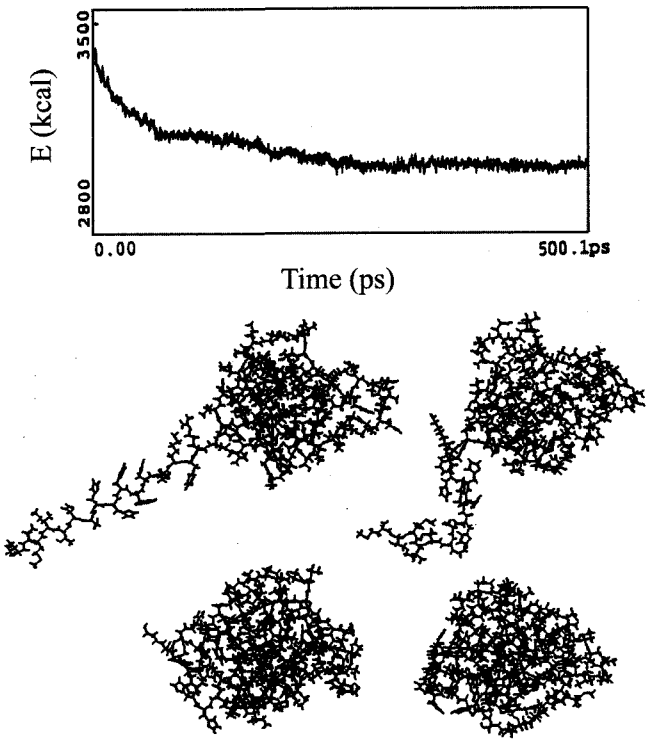

Fig. 5. Structures of preproinsulin obtained from MD and its energy profile.

native insulin (Fig. 6), and the RMS value was 2.91 Å. The secondary structure of the A and B chains by MD of preproinsulin and that of native insulin were likewise well superimposable. These results coincide once again with the secondary structure obtained from NMR in α-helix regions. The distances between these backbones were nearly maintained to a value of 5 Å. For the case of proinsulin. the distance between the N-terminal of the A chain and the C-terminal of the B chain was about 8 Å. However, the result of preproinsulin showed that the

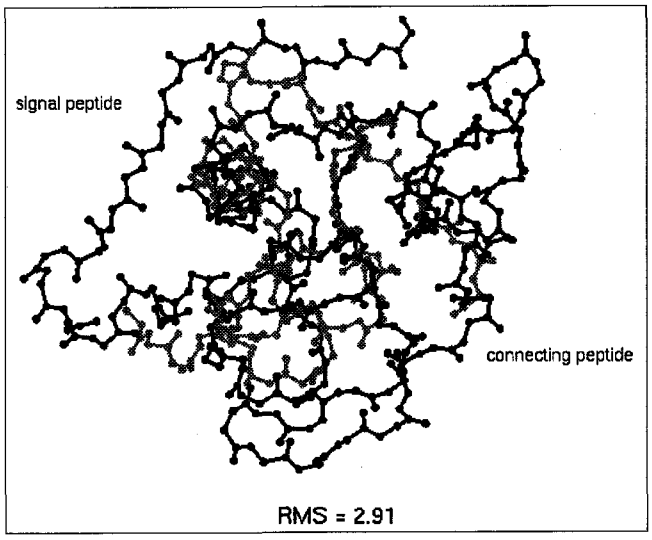

Fig. 6. A comparison of MD structure of native human insulin and that of preproinsulin.

distance was 5.28 Å, and the distance of both ends of the C chain of preproinsulin was 7.33 Å. As shown in Fig. 6, it can be regarded that the three dimensional structures of the A and B chains of preproinsulin are well-conserved like those of native insulin.

Acknowledgments. The authors wish to acknowledge the financial support of the Korea Research Foundation made in the program year of 1997. Software for CAMM was supported by MSI (San Diego, CA).

References

1 Brown, H., Sanger, F. and Kitai, R. (1955) The structure of sheep and pig insulins. *Biochem. J.*, **60**, 556-565.

2. Frank, B. H., Pekar, A. H. and Veros, A. J. (1972). Insulin and proinsuliin conformation in solution. *Diabetes* [suppl. 2] **21**, 486-491.

3. Blundell T. L., Dodson, G. G., Hodgkin, D. C. and Mercola, D. A. (1972) Insulin: The structure in the crystal and its reflection in chemistry and biology. *Adv. Protein Chem.* **26**, 379-402.

4. Milthorpe, B. K., Nicol, L. W. and Jeffrey, P. D. (1977) The polymerization pattern of zinc(II) insulin at pH 7. *Biochem. Biophys. Acta* **495**, 195-202.

5. Goldman, J. and Carpenter, F. (1974) Zinc binding insulin dichroism and equilibrium sedimentation studies on insulin (bovine) and several of its derivatives. *Biochemistry* **13**, 4566-4574.

6. Smith, G. D., Duax, W. L., Dodson, E. J., Dodson, G. G. and Reynolds, C. G. (1982) The structure of des-phe B1 bovine insulin. *Acta Crystalog.* **38**, 3028.

7. Hua, Q. and Weiss, M. A. (1991) Comparative 2D NMR studies of human insulin and des-pentapeptide insulin: Sequential resonance assignment and implications for protein dynamics and receptor recognition. *Biochemistry* **30**, 5505-5515.

8. Kruger, P., Burger, W. S., Wollmer, A., Gunstern, W. F. V. and Dodson, G. G. (1987) The simulated dynamics of the insulin monomer and their relationship to the molecule's structure. *Eur. Biophys. J.* **14**, 449-459.