Predicted Structural Alterations in Proinsulin during Its Interactions with Prohormone Convertases[†]

Gregory Lipkind[‡] and Donald F. Steiner*[,‡,§]

Department of Biochemistry and Molecular Biology and The Howard Hughes Medical Institute, The University of Chicago, Chicago, Illinois 60637

Received June 30, 1998; Revised Manuscript Received November 3, 1998

ABSTRACT: The intracellular conversion of proinsulin to insulin occurs via cleavage at the two dibasic sites: Arg^{31}-Arg^{32}, B chain−C-peptide (BC) junction; and Lys^{64}-Arg^{65}, A chain−C-peptide (CA) junction, catalyzed by the subtilisin-like prohormone convertases SPC3 (PC1/PC3) and SPC2 (PC2), respectively. In this report we propose a possible conformational variant of proinsulin that would facilitate the formation of enzyme−substrate complexes at the BC and AC junctions of proinsulin with the substrate binding groove of the two closely related convertases. Productive convertase interaction requires extended peptide conformations in both the CA junction (residues 62−67, LQKRGI) and the BC junction (residues 29−34, KTRREA) and leads to significant perturbations in the normally α-helical N-terminal region of the A chain and the extended C-terminal region of the B chain of the insulin moiety of proinsulin. In this model of the reactive conformation of human proinsulin, both processing sites assume positions that are relatively far apart. The C-peptide was then modeled in an unobtrusive conformation relative to the convertases and the remainder of the substrate, forming an extended loop of length ∼40 Å with a short α-helical segment rather than a random coil. A model of the stereochemical transformations that occur during the processing of proinsulin by SPC2 is presented.

In neuroendocrine cells, many peptide hormones and neuropeptides are synthesized as parts of larger, less active or inactive precursors. An essential step in the formation of bioactive peptides is the selective endoproteolytic cleavage of their precursors, as was demonstrated in the processing of proinsulin to insulin (*1*−*3*). Relative to insulin, proinsulin includes the additional connecting peptide segment (C-peptide) which links the C-terminus of the insulin B chain to the N-terminus of the insulin A chain (Figure 1). The connecting segment ranges in size from 30 (cow) to 35 (human, rat) amino acid residues, while the A and B chains are constant in length (*4*). All the known mammalian proinsulins have pairs of basic residues at either end of the C-peptide: Arg^{31}-Arg^{32} at the B chain−C-peptide junction (site 1) and Lys^{64}-Arg^{65} at the C-peptide−A chain junction (site 2). Both sites are processed in the conversion of proinsulin to insulin (*4*).

Despite its considerably larger molecular size, proinsulin is remarkably similar to insulin in many properties, including solubility, isoelectric point, and self-associative properties (*4*). In solution and in the crystalline state, proinsulin forms dimers and zinc-stabilized hexamers similar to those of insulin (*5*−*7*). Moreover, proinsulin has been cocrystallized in a 1:1 complex with insulin (*8*, *9*). These observations

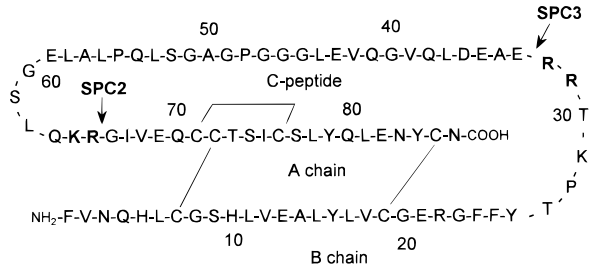

FIGURE 1: Primary structure of human proinsulin. Preferred sites of cleavage by SPC3 and SPC2 are indicated by arrows. Following endoproteolytic cleavage at these sites, the C-terminal basic residues are removed by carboxypeptidase E to generate native insulin and C-peptide (*4*). Processing at only the B chain−C-peptide junction generates des-31,32-proinsulin intermediate while processing at only the C-peptide−A chain junction generates des-64,65-proinsulin intermediate.

strongly suggest that the conformation of the insulin moiety in proinsulin is nearly identical to that of insulin. Recently, Weiss et al. (*10*) also showed by comparative ^{1}H NMR and photochemical dynamic nuclear polarization (photo-CIDNP) studies of human proinsulin and insulin that the A and B chains of proinsulin are similar in structure to those in insulin. The A chain contains 21 amino acid residues and forms 2 α-helices at its N- and C-terminals: A1−A8 and A14−A20. The B chain (30 amino acid residues) includes a central α-helix (B9−B19), which is framed by 2 extended regions on both ends (*11*, *12*). The surface of insulin that is free from contacts in dimers and hexamers is formed primarily by residues located between the two α-helices of the A chain (for example, A2 Ile, A3 Val, A4 Glu, A8 Thr, A19 Tyr).

[†] This work was supported by USPHS Grants DK 13914 and DK 20595 and in part by the Howard Hughes Medical Institute.

* Correspondence should be addressed to this author at The Howard Hughes Medical Institute, 5841 S. Maryland Ave., MC 1028, Room N-216, Chicago, IL 60637. Phone: (773) 702-1334. FAX: (773) 702-4292. E-mail: dfsteine@midway.uchicago.edu.

[‡] Department of Biochemistry and Molecular Biology.

[§] The Howard Hughes Medical Institute.

10.1021/bi981556q CCC: $18.00
Published on Web 12/29/1998

Despite much effort devoted to crystallization of proinsulin, it has not yet been possible to determine the three-dimensional structure of this protein by X-ray diffraction. Usually, the C-peptide is considered as a structureless segment in proinsulin (*4*, *13*, *14*). However, the C-peptide most likely does not exist as a statistical coiled coil. Weiss et al. (*10*) observed significant changes in the chemical shifts of protons of aromatic groups in the insulin moiety of proinsulin (A19 Tyr, B24 Phe) as compared with insulin. These results suggested the possibility that a stable local structure is formed at the CA junction of proinsulin, associated with alterations in insulin-specific core interactions (*10*). This structural deviation was designated as the "CA knuckle," but its precise nature has not been established.

Taking into account that the C-peptide does not obscure those surfaces of insulin which interact to form dimers and hexamers, Steiner (*8*) and Snell and Smyth (*15*) proposed hypothetical arrangements of the connecting peptide. In both models, the C-peptide is considered as a long extended loop overlying the A chain core residues of insulin and with connections to the insulin monomer at both the CA and BC junctions, which are near each other. In this report we consider the structure of proinsulin in relation to its role as a substrate in the processing of proinsulin to insulin.

The intracellular conversion of proinsulin to insulin is realized through proteolytic processing by specialized neuroendocrine endopeptidases. Davidson et al. (*16*) demonstrated the presence in insulinoma secretory granules of two calcium-dependent activities (designated types I and II), which were directed specifically against the BC and CA junctions, respectively. Later it was found that the type II activity corresponded to the subtilisin-like prohormone convertase SPC2 (PC2), while the type I activity corresponded to SPC3 (PC1/PC3) (*17*−*19*).

Due to the recent discovery that these enzymes participate in the processing of proinsulin and many other neuroendocrine precursors, we can consider the optimal conformations of the BC and CA junctions, and of the connecting segment, that would facilitate the binding and cleavage of proinsulin by the prohormone convertases.

Although X-ray investigations of the SPCs have not yet been successful, members of this family of proteases all possess similar catalytic domains that are based on the motifs of the bacterial subtilisins (*20*). Taking advantage of this feature, Siezen et al. (*21*, *22*) have proposed a spatial structure of the catalytic domain of furin (SPC1). We have taken a similar approach in modeling the active sites of PC2 and PC3 (*23*) and have identified the corresponding substrate binding pockets of these convertases, as exemplified by PC3. As in the subtilisins (*24*), the active site grooves of the convertases can bind at least six amino acids residues (P4−P2′)[1] of substrates in extended conformation.

Surveys of the cleavage sites of a large variety of precursors that are processed within the secretory pathway have led to the conclusion that most of them possess multibasic cleavage sites of the general type Arg-X-Lys/Arg-Arg (*26*, *27*). A basic residue at position −4 relative to the scissible bond (P4) may also contribute to substrate selectivity of the prohormone convertases under some conditions, although dibasic cleavage sites are the rule in most prohormone substrates.

In the proposed model of the active sites of SPC2 and SPC3 (*23*), the S1, S2, and S4 sites each contain residues of aspartic or glutamic acid which can interact with the guanidinium or amino groups of arginine or lysine side chains and thereby contribute to substrate selectivity. Moreover, the S1 subsite readily accommodates the relatively long Arg side chain which extends downward such that the interaction of its guanidinium group and the aspartic acid residue (#320 in SPC3)[2] located at the bottom of the cavity is favorable. Both SPC2 and SPC3 have closely similar S4 pockets, which correspond well to the size and shape of the side chain of Arg and provide significant electrostatic stabilization of binding for this residue (when present) through interactions with the side chains of residues Glu^{250} and Asp^{273} in SPC3 and with Glu^{250} in SPC2. It is possible that the presence of Asn in position 273 of SPC2 instead of Asp (as in SPC3) explains the reduced sensitivity of SPC2 to the nature of the residue in the P4 position (*19*). Residues of substrates after the scissible peptide bond (P1′ and P2′) are also of importance for binding to the prohormone convertases. Nakayama et al. (*29*) have found that hydrophobic amino acid residues predominate at position P2′, while hydrophobic aliphatic residues are not suitable at position P1′. Lipkind et al. (*23*) also have shown that despite the high structural homology of the subtilisins and the SPCs, the structure of the loop which lies immediately below the active sites differs due to the presence of a *cis*-peptide bond (Tyr^{167}-Pro^{168}) in this loop in the subtilisins and its absence in the SPCs (*23*).

In our reconstruction of the spatial structure of human proinsulin, we have assumed that residues P4−P2′ of both the CA junction (62−67, LQKRGI) and the BC junction (29−34, KTRREA) adopt extended conformations, suitable for binding with the active sites of prohormone convertases. This approach allowed us to consider the arrangement of the CA and BC junctions relative to the insulin moiety of proinsulin, taking into account possible interactions of the connecting peptide with the surfaces of the active site regions of the enzymes as well as the insulin moiety of proinsulin.

METHODS

Molecular modeling was carried out using the Insight and Discover graphical environments (Biosym Technologies, Inc., San Diego). The molecular mechanics energetic calculations utilized the force field cvff. Molecular models of the active sites of SPC3 and SPC2, proposed by Lipkind et al. (*23*), and the X-ray structure of insulin (*11*) as an initial approximation for the insulin moiety of proinsulin have been used. In determination and description of the angles of rotation around the $C^{\alpha}-N(\varphi)$ and $C^{\alpha}-C'(\psi)$ bonds in the peptide chain, we follow the standard nomenclature (*30*).

RESULTS AND DISCUSSION

The structure of the CA junction was considered using a model consisting of insulin enlarged by inclusion of residues

[1] According to the nomenclature of Schechter and Berger (*25*), the scissible peptide bond is located between the P1 and P1′ substrate residues. Amino acid residues of the substrate are numbered P2, P3, etc. toward the N-terminus. The complementary subsites of the binding region of the enzyme then are numbered S1′, S1, S2, S3, ...

[2] The numbering of amino and residues used here corresponds to that of SPC3 (*28*).

FIGURE 2: Complex of a portion of a representative SPC (SPC2) (yellow labels) with the CA junction (residues 62−65) and the insulin moiety (residues 66−69) of mouse or rat proinsulin I (*45*), shown by space-filling images, and marked by pink labels. The A and B chains are shown as pink and yellow ribbons.

62−65 of proinsulin (LQKR) at the N-terminus of the A chain, while residues at the C-terminus of the B chain (B27−B30) were excluded because normally they are highly flexible (*31*). We noticed that when the P4−P1 residues (62−65, LQKR) of proinsulin were arranged in the active sites of the modeled SPC, in a manner similar to that predicted for the complex with the tetrapeptide model substrate ArgAlaArgArg (*23*), a prohibited overlapping of the surfaces of the convertase active site with the volume of the insulin moiety occurred. It was thus apparent that productive binding of the CA junction of proinsulin by the convertases would require significant perturbations of several residues in the N-terminal A chain α-helix. Accordingly, after fixation of residues 62−65 of proinsulin, we calculated a suitable arrangement of residues Gly^{66}/insulin A1 and Ile^{67}/insulin A2 (the P1′ and P2′ residues, respectively) in the active sites of SPC2, appropriate for attack by the side chain of Ser^{384} on the peptide bond between Arg^{65} (P1) and Gly^{66} and for the formation of the tetrahedral and acylenzyme intermediates. In addition, the side chain of Ile^{67} was located in the surface hydrophobic cavity of the convertase, as formed by Tyr^{343} of SPC2 (loci S2′).

However, to avoid nonbonded repulsions with residues of the surface groove near the active site of the enzyme, it was necessary also to change the conformations of at least two additional residues—Val^{68}/A3 and Glu^{69}/A4—by rotation around their main chain $C^{\alpha}-N$ and $C^{\alpha}-C'$ bonds. This maneuver introduced an approximately perpendicular orientation of the insulin moiety relative to the enzyme surface and resulted in a partial uncoiling of the A chain N-terminal helix. The modeled binding of the CA junction in a complex with the convertase thus has the extended structure, as shown in Figure 2 with angles of rotation $\varphi(C^{\alpha}-N)$ and $\psi(C^{\alpha}-C')$ for residues P4−P4′ (62−69, LQKRGIVE) as follows: −77, 114; −143, 125; −41, 128; −77, 52; −114, 81; −63, 118; −91, −128; −85, 132; respectively. From this it is evident

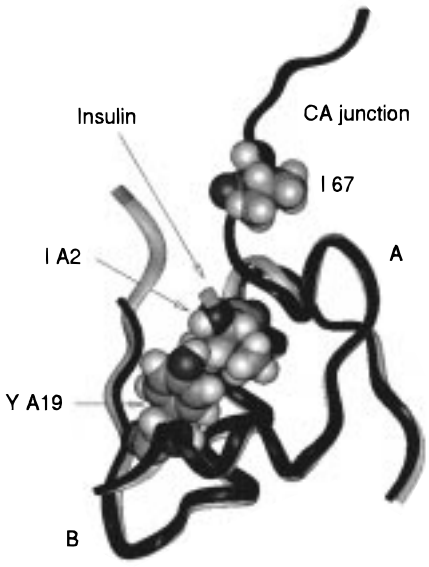

FIGURE 3: Superposition of the CA junction and the partly unfolded insulin moiety of SPC-bound proinsulin (dark ribbon) with the native insulin structure (light ribbon).

that residues A1−A4 of the A domain of proinsulin (proinsulin residues 66−69) must also adopt an extended conformation as part of this β-strand and as illustrated in the superposition of insulin and the CA junction of proinsulin in Figure 3. It is important to note that the strong hydrophobic contact between Tyr^{84}/A19 and Ile^{67}/A2 that normally occurs within the hydrophobic core of insulin is disrupted in this model of proinsulin.

Docking of the CA junction of proinsulin in the active site of the convertases also influences the structure of the BC junction. If we consider the C-terminal residues of the B chain (B26−B30) as they are oriented in the X-ray structure of insulin (*11*), then these residues form prohibited contacts with some of the surface residues of the SPCs. To avoid such repulsions, we have introduced a half β-turn at a likely position—in the middle residues Thr-Pro (B27−B28). This turn allows the C-terminal residues of the B chain to be arranged on the protein surface in an extended conformation that does not perturb the CA interaction (Figure 4). Thus, interaction of the CA junction of proinsulin with a convertase requires that the BC junction also adopt an extended conformation.

In considering the BC junction/convertase interactions, we have used a similar modeling approach. Calculations were made for the P4−P2′ residues of the BC junction (29−34, KTRREA), using residues Thr B27 and Pro B28 as a hinge between the BC junction and the insulin moiety. The angles of rotation $\varphi(C^{\alpha}-N)$ and $\psi(C^{\alpha}-C')$ for P4−P2′ are close to those found for the CA junction residues. For Thr B27 and Pro B28, the angles φ and ψ are as follows: −129, 68; −91, 161; respectively.

Finally, we have combined these predicted conformations of the CA and BC junctions in a model des-(35−61)-fragment of proinsulin by superimposing the conserved regions of the insulin moiety in the two model structures discussed above (Figure 5). This figure shows that for optimal convertase attack the two junctions are widely separated; the distance between the C^{α} atoms of the P1 residues at the two junctions is ∼30 Å. If the CA junction is directed upward relative to the A and B chains of the

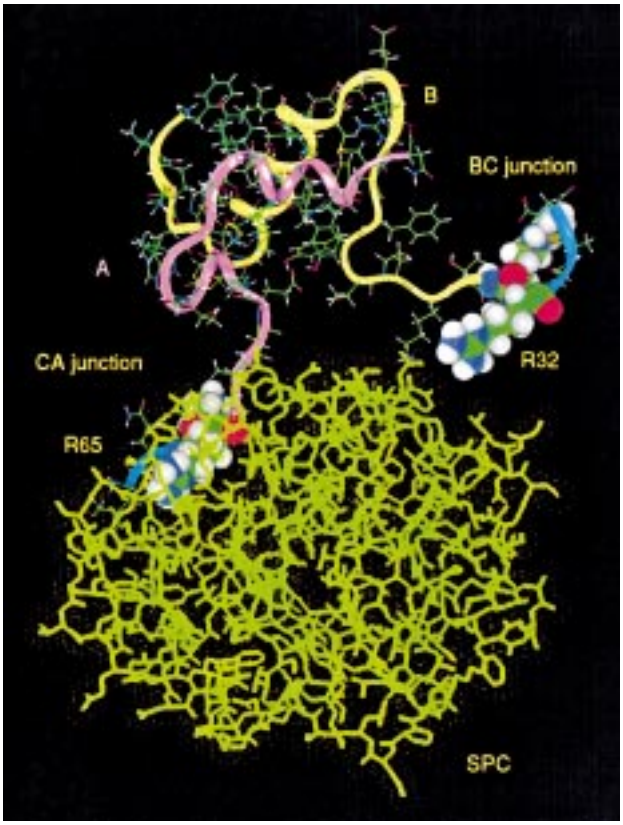

FIGURE 4: Arrangement of the BC junction of proinsulin on the surface of an SPC upon binding of the CA junction in the active site groove.

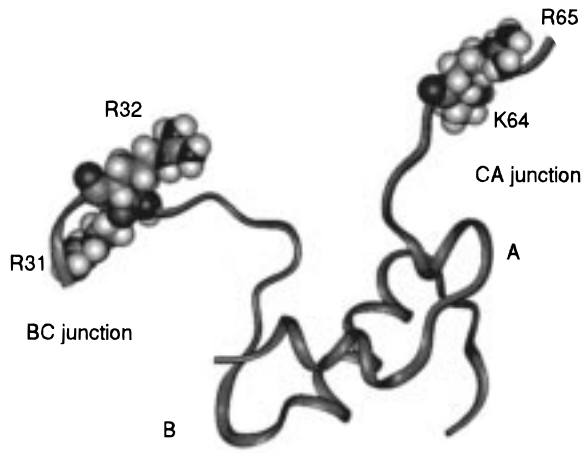

FIGURE 5: Arrangement of the CA and BC junctions of proinsulin relative to its insulin moiety in ribbon representation.

insulin moiety, then the BC junction is located toward the viewer. This arrangement of the CA and BC junctions would thus permit optimal binding and cleavage at either of the dibasic sites; however, it must be borne in mind that simultaneous cleavage at both sites would be prevented by steric interactions between the two attacking convertases.

The above partial proinsulin model requires the presence of at least 15−20 additional connecting residues to span the open distance between the A and B chains. Experimental data on the processing of miniproinsulins, i.e., proinsulins with shortened connecting peptide segments, support this model. For example, des-(38−62)-proinsulin was not processed by either enzyme, SPC2 or SPC3 (*32*), though the connecting peptide (RREAEDLQKR), and the A and B chains contain the same important residues for binding (P4−P2′) at both junctions as does wild-type proinsulin. One reasonable explanation for the lack of processing of this mutant form is that the CA and BC junctions are too close in space, due to the shortened connecting peptide, and therefore lack the flexibility to assume the more extended conformations required for productive convertase binding. In some instances, even small deletions in the C-peptide can be crucial for processing, as in the case of deletion of just four residues (residues 33−36) after the dibasic site at the BC junction of rat proinsulin; this form was not cleaved by SPC3 in AtT20 cells (*33*). In this case, however, cleavage failure may be partially due to unfavorable changes in the P1′ residue (Glu→Pro) resulting from the deletion. The successful cleavage in yeast of a miniproinsulin containing the short connecting sequence RRLQKR may reflect significant differences in the requirements for productive binding to kexin, the *S. cerevisiae* SPC family convertase (*19*).

Another interesting example is the insulin-like growth factor I. Its spatial structure in solution has been resolved by NMR methods (*34*). Despite the presence of two neighboring Arg residues (Arg^{36}-Arg^{37}) near the middle of the C domain of IGF I, convertases do not normally cleave this site. This is easily understood from a conformational point of view, since the rather fixed structure of the P4−P2′ residues in this molecule does not at all coincide with the required substrate conformation for convertase binding. Moreover, the peptide bond between P1 and P1′ residues (Arg^{37}-Ala^{38}) is screened by Pro^{39} and Pro^{28} such that attack by Ser^{382} of the convertase is impossible.

It is important to note that the proposed structure of the des-(35−61)-fragment of proinsulin (Figure 5) allows the formation of dimers similar to those found in crystalline insulin (*11*) and likely does not prevent the formation of hexameric proinsulin. This is an important point in considering whether the substrate is processed normally as a dimer or hexamer (*35*).

We will now proceed to consider the possible structure of the central portion of the connecting peptide segment, as influenced by interactions of the CA and BC junctions of proinsulin with the active sites of the convertases, as modeled in the des-(35−61)-fragment (Figure 5). In the case where the BC junction of the des-(35−61)-fragment is bound, the closure of the gap between the free ends and the arrangement of those regions of the C-peptide outside the active sites of the convertase specify its location close to the external surface of the A chain. Similarly, at the CA junction the only way to avoid prohibited contacts of the connecting peptide with the enzyme volumes is by locating it on the same surface of the A chain (see Figure 4). The foregoing considerations lead to the proposal of an approximate arrangement of the C-peptide relative to the insulin moiety, and also specify the configuration of the C-peptide as an extended loop, joining the remote ends of the BC and CA junctions (∼35−40 Å apart). Moreover, from this model it follows that the C-peptide screens and blocks some residues of the A chain that are believed to participate in the receptor binding of insulin (Val A3, Thr A8, Tyr A19), possibly explaining the significantly reduced binding of proinsulin (about 30-fold) relative to insulin (*4*, *13*). Epitope analysis of the binding of monoclonal antibodies by proinsulin also

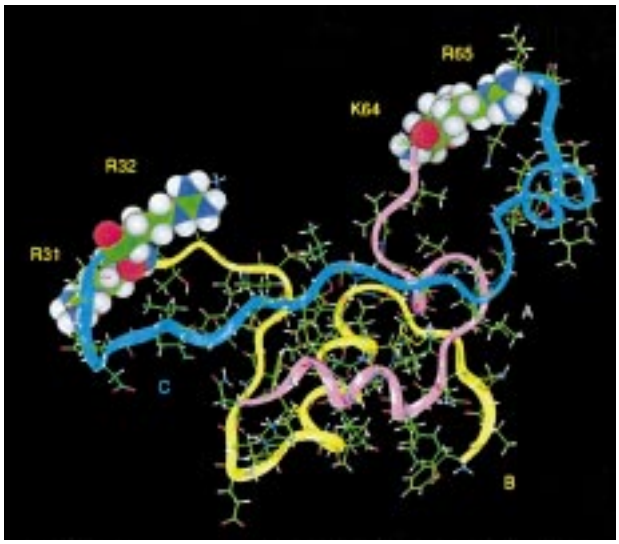

FIGURE 6: Proposed spatial model of human proinsulin in a conformation suitable for its proteolytic conversion.

suggests that the C-peptide may be interacting with the insulin moiety (*36*).

The determination of the probable orientation of the C-peptide in this model allows us to consider a possible variant of its conformation that may occur upon binding of proinsulin with prohormone convertases (Figure 6). Most likely, the sequence 33−36 (EAED) forms a β-turn between the BC junction and the more extended region of the connecting peptide. Then residues 37−44 (LQVGQVEL) form a part of this irregular, extended strand in the direction of the A chain, and their location is determined by the formation of favorable van der Waals contacts with the insulin moiety and the active site region of the convertase bound to the BC junction. The presence of Gly and Pro residues in the region between residues 37 and 44 in the proinsulin of other species does not promote the formation of α-helix by this segment. Continuing onward, the glycine-rich segment in the mid region of the C-peptide, residues 45−48 (GGGP), was packed on the exposed surface of the A chain formed by residues Tyr A19, Gln A5, Gln A15, and Ser A9. Pro[48] at the end of this segment contacts Ser A9 and, through a half-β-turn, changes the course of the connecting peptide in the direction of the CA junction. Here the GGGP motif adopts an extended conformation. Indeed, multiple Gly residues rarely are found in turnlike structures; e.g., a GGG sequence in fibrinopeptide A forms an extended structure (*37*). It is likely that these Gly residues in the middle of the C-peptide facilitate optimal docking with the convertases.

Residues 53−58 (LQPLAL) in the middle of the segment between Pro[48] and the CA junction have a high probability of forming an α-helix. According to Chou and Fasman (*38*), an α-helix is predicted if four out of six adjacent residues are helix-favoring. The sequence 53−58 contains 5 helix-favoring residues (and, moreover, proline [Pro[55]] tends to occur on the N-terminal side of α-helices). It is of interest that the residues of this helix have been deleted in the C-peptides of cow and sheep proinsulins (*4*). The absence of this α-helical region in these proinsulins would not be expected to lead to significant perturbations in the structure of the C-peptide, due to the compactness of residues 54−58. After this α-helix, three relatively flexible residues, E-G-S (59−61), form the second β-turn and join with residue Leu[62] of the CA junction. Again the appropriate conformations for interaction with the convertases of residues in the proximity of the CA junction were taken into account.

This proposed structure emphasizes the ability of the CA and BC junctions of proinsulin to adopt conformations appropriate for the formation of enzyme−substrate complexes with prohormone convertases, a characteristic property of prohormone structure. It is possible that both the CA and BC junctions form stable local structures, as suggested for the CA junction by Weiss et al. (*10*), i.e., the so-called "CA knuckle." Moreover, the connecting C-peptide is unlikely to possess the full flexibility of a random coil and thus may exert a definite conformational strain on the structure of the CA junction, as is indicated by the observation that des-(31,32)-proinsulin is processed by the type II (SPC2) endopeptidase more rapidly than is intact proinsulin (*39*, *40*). These observations support the proposed motif of proinsulin shown in Figure 6 in which the connecting peptide is represented neither as a rigid structure nor as a random coil, but rather assumes the form of an extended loop between the remote CA and BC junctions, running with a distribution of conformations approximately as indicated. The proinsulin C-peptide domain thus resembles a fluctuating string with fixed ends representing the CA and BC junctions appropriately poised for proteolysis.

It is very likely that significant structural perturbations in the A chain of the insulin moiety of proinsulin (Figure 6) take place during proinsulin processing when SPC2 acts first at the CA junction to produce a des-(64,65)-proinsulin intermediate. Possible stereochemical transformations of proinsulin on this processing pathway, which includes successive cleavages at the CA junction and then at the BC junction to generate insulin, are shown in Figure 7.

Traditionally, the insulin processing pathway with initial cleavage at the BC junction to produce des-(31,32)-proinsulin is considered as more dominant (*41*, *42*) because des-(31,32)-proinsulin is a preferred substrate for SPC2 in comparison with proinsulin (*39*). More recent observations in living cells have shown that either SPC3 or SPC2, acting alone, can cleave proinsulin at both junctions to fully convert proinsulin to insulin. Cell lines that express high levels of SPC3, but barely detectable SPC2 (AtT20 cells; *43*) or, contrary, high levels of SPC2 without any detectable presence of SPC3 (αTC1−6 cells;[3] *42*), efficiently process proinsulin to insulin. Moreover, either transfected SPC2 or SPC3 can convert human proinsulin to insulin in the same cellular setting (GH3 cells; *44*). Nonetheless, each prohormone convertase displays a preference for a single site of cleavage: SPC3 at the B chain−C-peptide junction, while SPC2 prefers the C-peptide−A chain junction (*16*, *17*, *39*, *44*). These results suggest that convertase site selectivity in proinsulin may be influenced by the level of convertase expression (*41*, *42*) as well as the nature of the P4 residue at that site (*45*); it is therefore likely that both pathways of proinsulin processing coexist in vivo. In Figure 7 we illustrate a plausible structural basis for the transformation of proinsulin to insulin in which SPC2 first processes the CA junction.

[3] Q. Gong and D. F. Steiner, unpublished results.

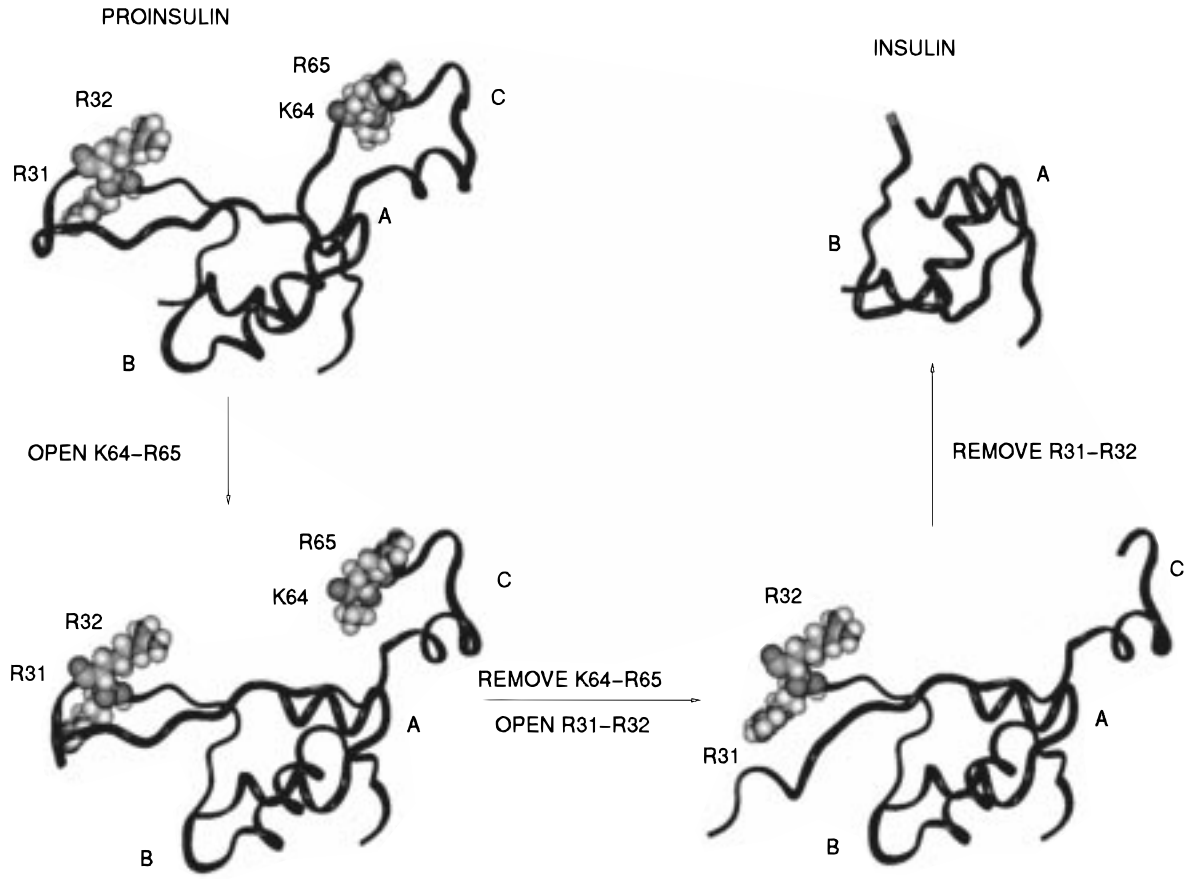

FIGURE 7: Stereochemical transformations of proinsulin to insulin in a processing pathway with a predominance of SPC2.

DISCUSSION

One of the key questions in the field of proteolytic processing of precursors is whether these protein substrates have well-defined native structures which facilitate their intracellular transport and enzymatic cleavage. The lack of a crystal structure for any of the many such precursors, including proinsulin, has left this question unanswered. Since the structure of insulin is known from crystallographic analysis, it has always been tempting to fill in the connecting segment. The available data, including NMR analysis (*10*), suggest that this region is in large part not well-defined, probably existing as an equilibrium among a multitude of possible conformations. What we have done here is to predict a conformation at sites of cleavage in proinsulin such that these fit appropriately into a generic active site of the catalytic domain of the prohormone convertases. Once the optimum arrangements of the cleavage regions relative to the insulin moiety were determined, the remainder of the C-peptide was then modeled according to its predicted secondary structural features in an unobtrusive conformation relative to the convertases and the remainder of the substrate. The resultant structure is highly plausible because it presents the cleavage sites in precisely correct orientations for enzymatic processing.

This model points out that highly significant structural alterations must occur in the N-terminal helix of the A chain domain of proinsulin in order for productive binding and cleavage by the convertase. Unwinding of the beginning of the N-terminal helix in the A domain exteriorizes Ile A2 from its normal juxtaposition with the side chain of Tyr A19 in the hydrophobic core of insulin, allowing its side chain to occupy the P2′ subsite on the convertase. An appropriate question to raise here would be as to whether proinsulin is actually so configured or whether interaction with SPC2 somehow induces this change, perhaps from a more intermediate transitional structure or set of structures. A related question would be whether des-(31,32)-proinsulin more readily adapts such a configuration. Because of the uniqueness of isoleucine in the structures of several animal proinsulins (present only in this position), it may be possible to answer such questions through ^{13}C NMR spectroscopic studies. Some other features of the proposed proinsulin structure also could be tested by mutational analysis (for example, by formation of disulfide bonds between residues predicted to lie in spacial proximity, such as Ser^{74}/A9 and Pro^{48}).

ACKNOWLEDGMENT

We thank Michael A. Weiss for helpful discussions and Rosie Ricks for expert assistance in the preparation of the manuscript.

REFERENCES

1. Steiner, D. F. (1967) *Trans. N.Y. Acad. Sci. 30*, 60−68.
2. Steiner, D. F., and Oyer, P. E. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 473−480.
3. Steiner, D. F., Clark, J. L., Nolan, C., et al. (1969) *Recent Prog. Horm. Res. 25*, 207−282.
4. Steiner, D. F., Bell, G. I., Tager, H. S., and Rubenstein, A. H. (1995) in *Endocrinology* (DeGroot, L., Ed.) pp 1296−1328, W. B. Saunders, Philadelphia, PA.

5. Pekar, A. H., and Frank, B. H. (1972) *Biochemistry 11*, 4013−4016.
6. Rosen, L. S., Fullerton, W. W., and Low, B. W. (1972) *Arch. Biochem. Biophys. 152*, 569−573.
7. Blundell, T. L., Cleasky, A., Murray-Rust, J., et al., unpublished results.
8. Steiner, D. F. (1973) *Nature 243*, 528−530.
9. Low, B. W., Fullerton, W. W., and Rosen, L. S. (1974) *Nature 248*, 339−340.
10. Weiss, M. A., Frank, B. H., Khait, I., Pekar, A., Heiney, R., Shoelson, S. E., and Neuringer, L. J. (1990) *Biochemistry 29*, 8389−8401.
11. Baker, E. N., Blundell, T. L., Cutfield, J. F., et al. (1988) *Philos. Trans. R. Soc.* B319, 389−456.
12. Derewenda, U., and Dodson, G. G. (1993) in *Molecular Structures in Biology* (Diamond, R., et al., Ed.) pp 260−277, Oxford University Press, Oxford, U.K.
13. Murray-Rust, J., McLeod, A. N., Blundell, T. L., and Wood, S. P. (1992) *BioEssays 14*, 325−331.
14. Xie, L., and Tsou, C.-L. (1993) *J. Protein Chem. 12*, 483−487.
15. Snell, C. R., and Smyth, D. G. (1975) *J. Biol. Chem. 250*, 6291−6295.
16. Davidson, H. W., Rhodes, C. J., and Hutton, J. C. (1988) *Nature 333*, 93−96.
17. Smeekens, S. P., Montag, A. G., Thomas, G., Albiges-Rizo, C., Carroll, R., Benig, M., Phillips, L. A., Martin, S., Ohagi, S., Gardner, P., et al. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8822−8826.
18. Bailyes, E. M., Shennan, K. I. J., Seal, A. J., Smeekens, S. P., Steiner, D. F., Hutton, J. C., and Docherty, K. (1992) *Biochem. J. 285*, 391−394.
19. Steiner, D. F., Smeekens, S. P., Ohagi, S., and Chan, S. J. (1992) *J. Biol. Chem. 267*, 23435−23438.
20. Mizuno, K., Nakamura, T., Ohshima, T., Tanaka, S., and Matsuo, H. (1988) *Biochem. Biophys. Res. Commun. 156*, 246−254.
21. Siezen, R. J., de Vos, W. M., Leunissen, J. A. M., and Dijkstra, B. W. (1991) *Protein Eng. 4*, 719−737.
22. Siezen, R. J., Creemers, J. W. M., and van de Ven, W. J. M. (1994) *Eur. J. Biochem. 222*, 255−266.
23. Lipkind, G., Gong, Q., and Steiner, D. F. (1995) *J. Biol. Chem. 270*, 13277−13284.
24. McPhalen, C. A., and James, M. N. G. (1988) *Biochemistry 27*, 6582−6598.
25. Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−182.
26. Bresnahan, P. A., Leduc, R., Thomas, L., Thorner, J., Gibson, H. L., Brake, A. J., Barr, P. J., and Thomas, G. (1990) *J. Cell Biol. 111*, 2851−2859.
27. Hosaka, M., Nagahama, M., Kim, W. S., et al. (1991) *J. Biol. Chem. 266*, 12127−12130.
28. Smeekens, S. P., Chan, S. J., and Steiner, D. F. (1992) *Prog. Brain Res. 92*, 235−246.
29. Nakayama, K., Watanabe, T., Nakagawa, T., et al. (1992) *J. Biol. Chem. 267*, 16335−16340.
30. Creighton, T. E. (1993) *Proteins: Structures and Molecular Properties*. 2nd ed., W. H. Freeman and Co., New York.
31. Kline, A. D., and Justice, R. M., Jr. (1990) *Biochemistry 29*, 2906−2913.
32. Docherty, K., Rhodes, C. J., Taylor, N. A., Shennan, K. I. J., and Hutton, J. C. (1989) *J. Biol. Chem. 264*, 18335−18339.
33. Gross, D. J., Villa-Komaroff, L., Kahn, C. R., Weir, G. C., and Halban, P. A. (1989) *J. Biol. Chem. 264*, 21486−21490.
34. Cooke, R. M., Harvey, T. S., and Campbell, I. D. (1991) *Biochemistry 30*, 5484−5491.
35. Dodson, G., and Steiner, D. (1998) *Curr. Biol. 8*, 189−194.
36. Crowther, N. J., Xiao, B., Jørgensen, P. N., Dodson, G. G., and Hales, C. N. (1994) *Protein Eng. 7*, 137−144.
37. Martin, P. D., Malkowski, M. G., DiMajo, J., Konishi, Y., Ni, F., and Edwards, B. F. P. (1996) *Biochemistry 35*, 13030−13039.
38. Chou, P. Y., and Fasman, G. D. (1978) *Annu. Rev. Biochem. 47*, 251−276.
39. Rhodes, C. J., Lincoln, B., and Shoelson, S. E. (1992) *J. Biol. Chem. 267*, 22719−22727.
40. Rouillé, Y., and Steiner, D. F., unpublished results.
41. Rouillé, Y., Duguay, S. J., Lund, K., et al. (1995) *Front. Neuroendocrinol. 16*, 322−361.
42. Steiner, D. F., Rouillé, Y., Gong, Q., Martin, S., Carroll, R., and Chan, S. J. (1996) *Diabetes Metab. (Paris) 22*, 94−104.
43. Irminger, J.-C., Vollenweider, F. M., Neerman-Arbez, M., and Halban, P. A. (1994) *J. Biol. Chem. 269*, 1756−1762.
44. Kaufmann, J. E., Irminger, J.-C., Mungall, J., and Halban, P. A. (1997) *Diabetes 46*, 978−982.
45. Furuta, M., Carroll, R., Martin, S., Swift, H. H., Ravazzola, M., Orci, L., and Steiner, D. F. (1998) *J. Biol. Chem. 273*, 3431−3437.

BI981556Q