X-ray analysis (1. 4-Å resolution) of avian pancreatic polypeptide: Small globular protein hormone

(satiety factor/graphics/zinc/insulin/glucagon)

T. L. BLUNDELL, J. E. PITTS, I. J. TICKLE, S. P. WOOD, AND C.-W. WU

Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX England

Communicated by Jerome Karle, April 1, 1981

The crystal structure of avian pancreatic poly-ABSTRACT peptide (aPP), a 36-residue polypeptide with some hormonal properties, has been determined by using single isomorphous replacement and anomalous scattering to 2.1-Å resolution. The phases were extended to 1.4-Å resolution by using a modified tangent formula. The molecule contains two regions of secondary structurean extended polyproline-like helix (residues 1-8) and an α -helix (residues 14-31) -that run roughly antiparallel. The packing together of nonpolar groups from these regions gives the molecule a hydrophobic core in spite of its small size. The aPP molecules form a symmetrical dimer in the crystal stabilized principally by interlocking of nonpolar groups from the α -helices. The aPP dimers are crosslinked by coordination of Zn²⁺; three aPP molecules contribute ligands to each zinc. The coordination geometry is a distorted trigonal bipyramid. The properties of the aPP molecule in solution are consistent with expectations based on the crystal structure. The aPP molecule has several general features in common with the pancreatic hormones insulin and glucagon. All three hormones have complex mechanisms for self-association. Like insulin, aPP seems to have a stable monomeric structure but its biological activity seems to depend on the more flexible COOH-terminal region analogous to the flexible NH₂-terminal region of glucagon.

Pancreatic polypeptide (PP) is a 36-amino acid peptide found in the endocrine pancreas (1, 2). There is close sequence homology among the mammalian peptides although mammalian and avian PPs differ at about 20 positions (Fig. 1). All sequences have an amidated COOH terminus, a feature known to occur in other polypeptide hormones such as gastrin, secretin, and oxytocin (3). There is now evidence that PP is synthesized as a larger precursor, in the same way as other pancreatic polypeptide hormones (4).

The ultrastructural appearance of the PP-producing cells also strongly argues for an endocrine function for this peptide. PP is located in membrane-enclosed granules (5) like those of the alpha and beta pancreatic cells in which glucagon and insulin are synthesized. The PP cells are located on the periphery of the islets at the head of the pancreas, to the exclusion of glucagon cells which occur at the periphery of the islets of the tail end (6). In certain fishes such as the teleost *Cottus scorpius*, which has two principal islets, the pyloric region has mainly PP cells and the splenic contains no PP cells (7).

PP is released into the circulation partly as a result of vagal cholinergic stimulation after feeding or in response to hypoglycemia. Although the half-life is of the order of 5 min the levels remain increased for several hours, indicating a continuous release (8). Injections of bovine PP (bPP) decrease food intake and body weight in the hyperglycemic ob/ob mouse (9), and injections of either the bovine or avian PP (aPP) cause New Zealand obese mice to revert to normal (10). Although these observations indicate that PP may act as a satiety factor, the peptide appears to have other physiological effects in addition, including increased gut motility and "gastrin-like" actions (11, 12). Recent results indicate that PP is a member of a larger family of homologous peptides found in the gut and in the brain (13).

We have purified aPP from turkey pancreas by using described techniques (14). The peptide has been crystallized, and preliminary x-ray analysis showed that the crystals contained dimers with approximately 50% α -helical content (14, 15). In this paper we discuss the extension of these studies to 1.4-Å resolution. We show that the molecule has a compact globular structure involving α -helical and polyproline-like helical secondary structures; we compare and contrast it with the known structures of two other pancreatic hormones, insulin and glucagon; and we discuss the implications of the conformation to its molecular biology.

MATERIALS AND METHODS

Purification and crystallization were carried out as described (14) although Zn^{2+} were found to be necessary for crystallization (15). The crystals produced in the presence of zinc were identical to those reported earlier (14, 16) with space group C2 and cell dimensions a = 34.18 Å, b = 32.92 Å, c = 28.44 Å, and $\beta = 105.30^{\circ}$. Data were collected by using a four-circle diffractometer and the peak scan methods of Tickle (17, 18), from the native and a derivative crystal prepared by soaking with 2 mM HgCl₂ for 24 hr. Until 2.1-Å resolution, the whole reciprocal sphere was measured for both the native and derivative, but between 2.1 and 1.4 Å only one equivalent was measured for the native. The data were corrected for Lorentz, polarization, absorption (19), and radiation damage; equivalent reflections were merged to give an agreement index,

$$R = \frac{\sum_{h} \sum_{i} |I_{(h)i} - \bar{I}_{(h)}|}{\sum_{h} \sum_{i} I_{(h)i}} \approx 3.6\%,$$

 $\bar{I}_{(h)}$ being the mean of measurements and $I_{(h)i}$ being the *i*th measurement of reflection h for both native and derivative.

Phases to 2.1-Å resolution calculated by using the method of single isomorphous replacement with anomalous scattering had an overall mean figure of merit of 0.79. Phase extension to 1.4 Å was achieved by using the modified tangent formula of Hull and Irwin (20).

The electron density map calculated by using phases to 2.1and 1.4-Å resolution was displayed on an Evans and Sutherland Computer Graphics System II and interpreted by using the program FRODO written by Jones (21) for a Vector General

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: PP, pancreatic polypeptide; bPP, bovine PP; aPP, avian PP.



FIG. 1. Amino acid sequences of turkey (aPP) and bovine (bPP) pancreatic polypeptides.

graphics system and modified for the Evans and Sutherland system by T. A. Jones and I.J.T. The atomic positions of 35 amino acids and the Zn^{2+} were used to calculate structure factors giving an agreement value of 27.9% for use in a difference Fourier analysis. The latter confirmed the position of the remaining amino acid (tyrosine-36) and the position of 37 welldefined water molecules. The coordinates have been deposited with the Brookhaven Data Bank.

RESULTS AND DISCUSSION

Electron Density Map. The electron density map at 3.0-Å resolution allowed the chain to be followed and indicated that the Zn^{2+} crosslink three molecules to stabilize the crystal lattice. Parts of the electron density map at 1.4-Å resolution are shown in Fig. 2, and these demonstrate high quality. The 1.4-Å electron density map gives a clear indication of the atomic positions. Most rings (phenyl, proline, etc.) have clear holes in the center, carbonyl oxygens of peptides are well-defined, and densities for leucines, valines, glutamate, etc. are clearly bifurcated. There is no evidence of a strong "squared structure" effect in the 1.4-Å electron density map although the contrasts between oxygen and carbon atoms, and between well-ordered

atoms and those with high thermal parameters, were heightened. This near-atomic-resolution electron density map allowed an unequivocal interpretation of the molecular conformation, although the relatively high thermal vibration of the terminal residue 36 necessitated confirmation of the atomic positions from the difference Fourier calculation. Many water molecules were shown by clearly resolved electron density.

Conformation of the Protomer. The conformation of the aPP protomer illustrated in Fig. 3 comprises two well-defined sections of secondary structure folded onto each other. The first stretch, consisting of residues 1–8, includes proline residues at positions 2, 5, and 8 and appears to have a poly-proline or collagen-like conformation. The mean value of the torsion angles ϕ and ψ for residues 2–7 is -70° and 138°, close to that reported for models of collagen. The conformation has an approximate 3-fold screw axis which brings the three proline side chains onto the same side of the polyproline-like helix so that they can all contribute to the hydrophobic core of the molecule.

The second well-defined region of secondary structure is an α -helix. Residues 14–31 have good α -helical conformation with mean values of ϕ and ψ as -63.6° and -41.0°. In this region, the hydrogen bond distances between carbonyl oxygen (O) of



FIG. 2. Electron density for the 1.4-Å resolution map of aPP showing the fit for residues phenylalanine-20 (Left) and tyrosine-7 to aspartate-11 (Right).

Biochemistry: Blundell et al.

residue *i* and amide nitrogen (N) of residue i + 4 range between 2.73 and 3.23 Å with a mean of 3.03 Å. The angles C—Ô—N vary between 136° and 166° with a mean of 149°. Residue 13 does not have an α -helical conformation but its carbonyl oxygen (O) makes distances to N(16) of 3.01 Å and to N(17), of 3.64 Å. The conformation at 32 is relaxed from a strict α -helix although a good hydrogen bond (3.1 Å) is made between O(28) and N(32). At 33 the conformation is that of a left-handed α -helix although the residue (arginine) has a side-chain and at residue 34 the conformation becomes more extended (see Fig. 3) although hydrogen bond contacts O(29)–N(33), O(29)–N(34), and O(30)–N(33) are made at the end of the α -helix. There is also a hydrogen bond O γ (32)–O(32) which may contribute to destabilization of the helix at this point.

The turn at residues 33 and 34 allows the COOH terminus to be oriented away from the α -helix axis. Residues 9–13 form a turn so that the polyproline-like helix and the α -helix lie approximately antiparallel with an angle of $\approx 152^{\circ}$ between the helix axes. This is achieved by glycine-9 assuming a conformational angle with ϕ positive leading into a type I β -bend between residues 9 and 12 with a hydrogen bond O(9)–N(12). Residues 12 and 13 have an extended chain conformation accommodating proline-13 and allowing space between the two helices.

The contacts between the polyproline-like helix and the α helix are hydrophobic and are established by an interdigitation of the three proline side groups between the mainly nonpolar side groups of one face on the α -helix (see Fig. 3), giving van der Waals' contacts of proline-2 with valine-30 and tyrosine-27, proline-5 with tyrosine-27, asparagine-23, leucine-24, and phenylalanine-20, and proline-8 with phenylalanine-20 and leucine-17. A further hydrophobic contact occurs between the side chains of tyrosine-7 and phenylalanine-20. This arrangement of groups allows an extensive region to be buried and undoubtedly accounts for the surprising stability of the conformation of this peptide in the monomeric form (22, 23). Furthermore, this interaction of a polyproline-like structure with an α -helix may form a general mode of interaction between collagens (the prolines of which would be available on the surface) and α -helical proteins, equivalent to the well-established and commonly occurring interactions between α -helices (24) and between an α helix and a β -sheet (25). There are also a number of specific interactions involving charged amino acid side chains which may stabilize the structure. These include a hydrogen bond or ionic interaction between: aspartate-10 carboxylate and tyrosine-7 hydroxyl, glutamate-15 carboxylate and arginine-19, aspartate-16 carboxylate and arginine-19, and aspartate-22 carboxylate and O(18) and O(8). These charged groups help to shield the hydrophobic core from water.

These features are characteristic of globular proteins in general. In fact, the requirement for a hydrophobic core implies a lower limit to the length of a chain that can adopt such a conformation. It is interesting that glucagon which has 29 amino acid residues (only 7 less than PP) has few tertiary interactions (26) and is flexible in dilute aqueous solutions where it is a monomer. It adopts a well-defined conformation only on selfassociation to trimers and higher oligomers (26) or in certain nonpolar solvents (27, 28). Previously defined globular proteins having no cofactor are somewhat larger and, like insulin (51 amino acids), often are folded in the form of a larger precursor and contain disulfide bonds to stabilize the conformer. PP is the smallest stable globular protein yet described and achieves economical and simple tertiary interaction through hydrophobic packing of two secondary structures. As in the case of insulin, there is an extensive area of the surface that is hydrophobic in the protomer; in aqueous solution, this leads to self-association.

Structure of the Dimer. The protomers are related by 2-fold crystallographic axes to form dimers with extensive hydrophobic interactions. The atoms of the dimer are shown along the 2-fold axis in Fig. 4 *Left*; the side chains involved in the interactions are shown from a direction at right angles to this in Fig. 4 *Right*. The two molecules have contacts both between the α -helices and the polyproline-like helices.

The axes of the α -helices are separated by ≈ 10 Å and make an angle of $\pm 30^{\circ}$. C. Chothia pointed out to us that the packing of helices is best described by the fitting of ridges into grooves (personal communication). Fig. 4 *Right* shows that these are formed by side chains of $i \pm 4$ in one helix and $i \pm 3$ in the second helix, giving a class III interaction (24). Because the helices are related by a 2-fold axis, the ridges of importance change on each helix in the vicinity of the axis. This arrangement also implies that the rows of side chains $i \pm 1$ are approximately parallel.

Around the 2-fold axis there is a series of aromatic groups tyrosine-7, phenylalanine-20, and tyrosine-21—which cluster together and form the center of a very hydrophobic core of the dimer. The aromatic rings of tyrosine-7 of the two molecules are stacked almost on top of each other, although the hydroxyl function is in a more hydrophilic region to hydrogen-bond to aspartate-10. Three leucines, two valines, and one isoleucine contribute to the hydrophobic interactions, as well as the C β , $C\gamma$, and C δ of glutamine-25 which shields the edge of this region from the solvent. The extensive contacts account for the



FIG. 3. (Left) Main chain atoms of aPP molecule, showing the relationship between the collagen-like helix (NH₂-terminal region) and the α -helix. (Right) Stereo diagram of the complete aPP molecule with an orientation similar to Left.



FIG. 4. (Left) Stereo diagram of the aPP dimer with side chains viewed down the 2-fold axis. (Right) Sections through van der Waals' spheres of side-chain atoms from the two PP molecules of a dimer, showing the interlocking of side chains from the α -helices at the interface. (Diagram was kindly provided by C. Chothia.)

stability of the dimer of aPP with $K_d \approx 5 \times 10^{-8}$ M and the hydrophobic nature of the interactions is consistent with the observation that the dimerization is driven entropically (29).

Coordination of Zinc in the Crystal Structure. Fig. 5 Left shows the arrangement of the dimers viewed along the 2-fold axis. Zn^{2+} , one per protomer, link the dimers which are symmetry-related by 2-fold axes. The linear arrays so formed are further linked through the Zn^{2+} to give a face-centered twodimensional lattice. In this way, each Zn^{2+} is coordinated to three protomers, and this phenomenon evidently is the reason for the decrease of solubility of the aPP observed in the crystallization experiments.

The coordination of the Zn^{2+} is shown in Fig. 5 *Right*. Each Zn^{2+} has five ligands: the NH_2 -terminal nitrogen and carbonyl oxygen of glycine-1 from one protomer, the amide of asparagine-23 from a second protomer, N^e of the imidazole of histidine-34 of a third, and a water molecule. The groups are arranged in a distorted trigonal bipyramidal geometry with the water molecule and carbonyl oxygen in the axial positions.

There are other intermolecular interactions that stabilize the crystal structure. One of these involved hydrogen bonding between arginine-33 guanidinium of one molecule and the main chain of the polyproline-like helix of a second. Hydrogen bonds are formed between carbonyl O of arginine-33 and peptide NH of glutamine-4 N^e of arginine-33 and carbonyl O of glutamine-4, and Nⁿ of arginine-33 and carbonyl O of proline-5. The arginine appears to be well-ordered, implying a strong interaction. Arginine and aspartate residues (for example, arginine-35 and aspartate-11) also give extended regions of ion pairs at the ends of the helices between molecules related by the translation of the unit cell.

There also are many well-ordered water molecules in the crystal structure. All free oxygens have at least one water molecule bound. At present we are reserving discussion of the water structure until we have made further progress with the refinement at 0.98-Å resolution for which we now have x-ray data.

Conformation of Mammalian PPs. Mammalian PP molecules are highly homologous one with another, although only 16 residues of aPP are conserved in bPP (Fig. 1). Can bPP adopt the same conformation and form dimers?

We have approached this problem by using computer graphics to construct a model of bPP and by using predictive methods to generate the secondary structure. These show that bPP



FIG. 5. (Left) Stereo diagram of the backbone atoms of two aPP dimers, illustrating the crosslinking by Zn^{2+} . (Right) Disposition of ligands (H₂O and asparagine-23, glycine-1, and histidine-34) from three aPP molecules to the Zn^{2+} , forming a distorted trigonal bipyramid.

Biochemistry: Blundell et al.

would be expected to have a helical structure between residues 14 and 31 in the same way as that observed in the aPP molecule. Furthermore, the side chains of the α -helix which pack against proline-2, -5, and -8 are conserved hydrophobic and allow equivalent tertiary interactions to be formed. The very conservative variation in the residues 9–12 allows formation of the β bend, and the proline can easily be accommodated at position 14 at the NH₂ terminus of the α -helix instead of position 13 as in aPP. More difficulty was experienced in the replacement of histidine-34 by proline, and this change may be expected to alter the flexibility and conformation of the COOH-terminal residues. Nevertheless, our model is consistent with a conserved general conformation, especially with respect to the helical content as has been demonstrated by circular dichroism (23, 30). The bovine molecule would also be expected to form dimers, and this has been confirmed (22, 23, 29), although Znlinked oligomers found in crystals of aPP could not occur because histidine-34 and asparagine-23 are not present in bPP.

Structure and Biology. Comparison of the structure and biological properties of three pancreatic hormones-insulin, glucagon, and PP-is instructive. All three peptides are stored in membrane-enclosed granules in specific islet cells, and for insulin and glucagon there is evidence that the storage form is aggregated and sometimes crystalline (26, 31-34). Indeed, in several cases, electron microscopy of sections of storage granules reveals crystal shapes and lattice dimensions similar to those of crystals grown in vitro. Although as yet there have been no reports of crystalline inclusions in PP-producing cells, dimers are likely to be present in granules of all species, and the Zn²⁺-linked extended structure described here for aPP crystals may, like the zinc hexamers of insulin granules, play a role in the stabilization of the PP storage granules (30).

When secreted into the circulation, oligomeric forms of insulin and glucagon dissociate to monomers, which appear to be important in receptor binding. This is also likely to be the case for mammalian PPs although for aPP in chickens, circulating levels are high and the association is tight, and so dimers may be relevant to receptor binding (22, 31, 35).

For insulin and glucagon, it has been suggested that the surface regions involved in dimerization are involved in receptor binding (26, 31, 35). In view of the many similarities between these hormones, it is tempting to speculate that the hydrophobic surface of the PP monomer, which is involved in dimerization, is also important in receptor binding. For glucagon, the nonhelical NH2 terminus is rather flexible and there is evidence that histidine-1 is more involved with target cell activation than with stabilizing the hormone-receptor complex (33, 34). PP seems to be inactivated on removal of tyrosine-36 and this region is relatively flexible as it extends away from the bulk of the molecule (36). Thus, it is possible that the COOH-terminal region involving tyrosine-36 of PP may have an activating role similar to that of the NH2 terminus of glucagon. However, no convenient bioassay or receptor assay has yet been reported for PP, and developments in this area are required to test the ideas presented here concerning the PP receptor interactions.

We thank Prof. A. Wollmer, Prof. J. Kimmel, Dr. N. Lazarus, Mr. D. Barlow, Dr. R. Chance, Prof. S. Falkmer, Dr. T. Schwartz, and Dr. C. Chothia for helpful discussions and for making available unpublished data. We are grateful to the Science Research Council (U.K.) for financial assistance.

- Kimmel, J. R., Pollock, H. G. & Hazelwood, R. L. (1968) En-1. docrinology 83, 1323-1330.
- 2 Floyd, J. C., Fayans, S. S., Pek, S. & Chance, R. E. (1977) Recent Prog. Horm. Res. 33, 519-570.
- 3. Dayhoff, M. O. (1972) Atlas of Protein Sequence and Structure (McGregor & Werner), Vol. 5.
- Schwartz, T. W., Gingerich, R. L. & Tager, H. S. (1980) J. Biol. 4. Chem. 255, 11494–11498.
- Larsson, L.-I., Sundler, F., Häkanson, R., Pollock, H. G. & 5. Kimmel, J. R. (1974) Histochemistry 42, 377-382.
- Orci, L., Baeteus, D. & Ravazzola, M. (1976) Life Sci. 19, 1811-6. 1816
- Van Noorden, S. & Falkmer, S. (1980) Invest. Cell Pathol. 3, 21-7. 35.
- 8. Schwartz, T. W., Stenguist, B., Olbe, L. & Stadil, F. (1979) Gastroenterology 76, 14-19.
- 9. Malaisse-Lagae, F., Carpentier, J.-L., Patel, Y. C., Malaisse, W. J. & Orci, L. (1977) Experientia 33, 915-917.
- 10.
- Gates, R. J. & Lazarus, N. R. (1977) Hormone Res. 8, 189–202. Taylor, I. L., Solomon, T. E., Walsh, J. H. & Grossman, M. I. 11. (1979) Gastroenterology 76, 524-528.
- Björnsson, O. G., Adrian, T. E., Dawson, J., McCloy, R. E., Greenburg, G. R., Bloom, S. R. & Chadwick, V. S. (1979) Eur. 12. J. Clin. Invest. 9, 293-301.
- 13.
- Tatemoto, K. & Mutt, V. (1980) Nature (London) 285, 417-419. Wood, S. P., Pitts, J. E., Blundell, T. L., Tickle, I. J. & Jenkins, 14. J. A. (1977) Eur. J. Biochem. 78, 119-126.
- Pitts, J. E., Blundell, T. L., Tickle, I. J. & Wood, S. P. (1979) in 15. Proceedings of the Sixth American Peptide Symposium, eds. Gross, E. & Meienhofer, J. (Pierce Chemical, Rockford, IL) pp. 1011-1016.
- 16. Kimmel, J. R., Hayden, L. J. & Pollock, H. G. (1975) J. Biol. Chem. 250, 9369-9376.
- 17. Blundell, T. L. & Johnson, L. N. (1976) Protein Crystallography (Academic, London).
- 18. Tickle, I. J. (1975) Acta Crystallogr. B31, 329-333.
- 19. North, A. C. T., Phillips, D. C. & Mathews, F. S. (1968) Acta Crystallogr. A24, 351-359.
- Hull, E. & Irwin, M. J. (1978) Acta Crystallogr. A34, 863-868. 20.
- 21. Jones, T. A. (1978) J. Appl. Crystallogr. 11, 268-272
- Noelkin, M. E., Chang, P. J. & Kimmel, J. R. (1980) Biochem-22 istry 19, 1838-1844.
- 23. Blundell, T. L., Dockerill, S., Pitts, J. E., Wood, S. P. & Tickle, I. J. (1977) FEBS Symp. 45, 249-258.
- 24. Chothia, C., Levitt, M. & Richardson, D. (1977) Proc. Natl. Acad. Sci. USA 74, 4130-4134.
- 25. Sternberg, M. J. E. & Thornton, J. M. (1977) J. Mol. Biol. 110, 285-296.
- 26. Sasaki, K., Dockerill, S., Adamiak, D. A., Tickle, I. J. & Blundell, T. L. (1975) Nature (London) 257, 751-757.
- 27. Contaxis, C. C. & Epand, R. M. (1974) Can. J. Biochem. 52, 456-468.
- Schneider, A. B. & Edelhoch, H. (1972) J. Biol. Chem. 247, 28. 4987-4991.
- 29. Chang, P. J., Noelken, M. E. & Kimmel, J. R. (1980) Biochemistry 19, 1844–1849.
- 30. Pitts, J. E. (1980) Dissertation (Sussex Univ., Brighton, England).
- Blundell, T. L., Dodson, G. G., Hodgkin, D. C. & Mercola, D. A. (1972) Adv. Protein Chem. 26, 279-402. 31.
- 32. Blundell, T. L., Dockerill, S., Sasaki, K., Tickle, I. J. & Wood, S. P. (1976) Metabolism 25, 1331-1336.
- 33. Blundell, T. L. (1979) Trends Biochem. Sci. 4, 51-54.
- Blundell, T. L. (1979) Trends Biochem. Sci. 4, 80-83. 34.
- Pullen, R. A., Lindsay, D. G., Wood, S. P., Tickle, I. J., Blun-dell, T. L., Wollmer, A., Krail, G., Brandenburg, D., Zahn, H., 35. Gliemann, J. & Gammeltoft, S. (1976) Nature (London) 259, 369-373.
- 36. Lin, T.-M., Evans, D. C., Chance, R. E. & Spray, G. F. (1977) Am. J. Physiol. 232, E311-E315.