Volume 219, number 2, 419-425 FEB 04963 July 1987

Prothoracicotropic hormone has an insulin-like tertiary structure

H. Jhoti, A.N. McLeod, T.L. Blundell, H. Ishizaki*, H. Nagasawa+ and A. Suzuki+

*Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, London University, Malet Street, London WC1E 7HX, England, *Faculty of Science, Nagoyal University, Chikusa, Nagoya 464* and +*Department of Agricultural Chemistry, Faculty of Agriculture, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan*

Received 27 May 1987

A three-dimensional model of PTTH-II has been constructed using interactive computer graphics and energy minimisation techniques, assuming homology with porcine insulin, the structure of which has been determined by X-ray analysis. The model shows that PTTH-II can assume an insulin-like tertiary structure, which is compact with the exception of the sequence variable NH_2-terminal amino acids of the B chain. Most of the hydrophobic core residues including A2 Ile, A6 Cys, A11 Cys, A16 Leu, A20 Cys, B11 Leu, B15 Leu and B19 Cys are identical in PTTH-II and insulins. The glycines at A1, B8 and B23 allow the chain to assume the characteristic tertiary interactions of the insulin fold and although polypeptide chains are shorter at the COOH-termini of the A and B chains and extended at the NH_2-terminus of the B chain, the insulin-like tertiary structure can still be assumed. It is unlikely that PTTH-II forms either dimers or hexamers, characteristic of porcine and human insulin, and the model is consistent with the inability of PTTH-II to bind anti-insulin antibodies or insulin receptors. A hydrophobic surface region of PTTH-II may be involved in intermolecular actions of physiological relevance. We discuss the implications of our model for evolution of this family of hormones and growth factors.

Prothoracicotrophic hormone; Tertiary structure; Sequence homology; Structure modeling; Evolution

1. INTRODUCTION

Prothoracicotrophic hormone (PTTH) is secreted by the brain of the silkworm, *Bombyx mori*, and acts on the prothoracic glands to stimulate the release of ecdysone [1,2]. *B. mori* contains two kinds of PTTH, a low molecular mass (4400 kDa) and a high molecular mass (22 kDa) species. Recently the 4 kDa-PTTH has been shown to be heterogeneous and to contain several species which differ in sequence [3]. The sequence of one of these, PTTH-II, has now been published [4] and has been shown to have two chains with 40% homology with the A and B chains of human insulin. More recently, several more sequences have become available which also have a high degree of homology with insulin. The homology suggests that PTTH is a member of the insulin family of hormones and that it is reasonable to assume an insulin-like tertiary structure. Most strikingly the hydrophobic core residues - A2 Ile, A6 Cys, A11 Cys, A16 Leu, A20 Cys, B11 Leu, B15 Leu and B19 Cys - are identical in PTTH-II and insulins. The shorter COOH-terminus of the B chain that is more reminiscent of relaxin, does not prevent the assumption of an insulin-like tertiary structure [7-10], but it is unlikely that PTTH-II forms either dimers or hexamers, characteristic of porcine and human insulins [5,6]. The model is consistent with

Correspondence address: H. Jhoti, Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, London University, Malet Street, London WC1E 7HX, England

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

the inability of insulin to evoke PTTH activity. A hydrophobic surface region of PTTH-II may be involved in intermolecular interactions of physiological relevance.

2. MATERIALS AND METHODS

The sequence of PTTH-II was aligned with the known sequences of insulins, insulin-like growth factors and relaxins (see [5] for a review) and the identical and conservatively varied residues identified. The aligned sequence of PTTH-II was modelled using the program FRODO [11] on to the three-dimensional structure of insulin [6]. Identical residues were given the same coordinates. The side chains of conservatively varied residues were placed so that side chain torsion angles were similar to those of the residues replaced and occupied the same region of space where this was compatible with good intramolecular interactions. Where there were more radical sequence differences and at chain extensions, main chain torsion angles were made consistent with secondary structure prediction, the maintenance of good tertiary interactions or similar supersecondary structures [12,13]. The structure was energy-minimised and analysed using programs developed by Dr I. Haneef and A. Hemmings at Birkbeck (Haneef, unpublished results; [14]) which use the functions of Weiner et al. [15].

3. RESULTS AND DISCUSSION

Table 1 shows the alignment of the PTTH sequences with the insulin, insulin-like growth factor and relaxin sequences (see [16] for review). PTTH-II has the arrangement of half-cystines that is characteristic of the insulin family. PTTH-II and PTTH-V also have glycines that are found in all insulins at B8 and B23, although in PTTH-IV B23 is serine. In insulins, these residues have torsion angles in which ϕ is positive [5,6], a conformation which is only easily available to glycine that has no side chain. B8 Gly enables the main-chain to turn sharply after the cystine at B7 to form the α-helix between B9 and B19. A further chain reversal between B20 and B23 involves sharp turns at B20 and B23. In most insulins B20 is glycine but this is not the case in many hystricomorph insulins such as those of guinea pig, coypu, casiragua and cuis [17–21]. It has been suggested that these isulins may not have the type II turn [22] between B19 and B21 with a hydrogen bond from B19 CO to B22 NH, but rather a type I or III turn where B20 has a torsion angle allowed for a residue with a side chain. It is thus possible to model the main chain of PTTH-II with a conformation close to that of human insulin, and in an identical way to that of the modelled structures of hystricomorph insulins with such a type III turn between B19 and B22 (fig.1a,b). Apart from the substitution of valine for isoleucine at B2 in PTTH-IV, all the changes in sequence variant PTTHs affect surface residues. These side chain substitutions can be easily accommodated in the modelled structure for PTTH-II except for the substitution of serine for glycine at B23 in PTTH-IV which creates some difficulty in modelling the C^- terminal portion of the B chain. It is possible that this portion of the molecule deviates from the insulin structure and that the C^- terminus of the B chain of the PTTHs is more flexible than the insulin B chain.

A1 Gly appears to be conserved to enable the helix A2–A8 to pack into the tertiary structure, and this glycine is also found in the PTTHs. The other residues in the A chain of PTTH-II are consistent with an insulin-like conformation, as shown in fig.1c and d. In insulin, the interchain disulphides of cystines A7–B7 and A20–B19 lead to a globular structure with a hydrophobic core. Central to this core are hydrophobic residues A2 Ile, A6 Cys, A11 Cys, A16 Leu, A20 Cys, B11 Leu, B15 Leu and B19 Cys. These are all conserved in the PTTH sequences except for the conservative substitution of Val for Ile in PTTH-IV. Other residues such as A3 Val (Val in PTTH-II), A19 Tyr (Tyr), B6 Leu (Tyr), B6 Leu (Tyr), B12 Val (Ala), B18 Val (Leu), B24 Phe (Val), contribute to the hydrophobic core although they are on the surface of the insulin protomer; they are all conserved as hydrophobic in PTTH and modelling shows that they are compatible with an insulin-like tertiary structure. Thus PTTH could form a tertiary structure with a hydrophobic core similar to all known insulins including those of the hagfish [23], the pig and the human [5,24] which have been defined by X-ray analysis. The proposed tertiary structure of PTTH-II is shown in fig.1e.

The absence of a residue equivalent to A21 of insulin and insulin-like growth factors would not

Table 1

The alignment of PTTH-II with representative insulin, insulin-like growth factor and relaxin sequences

	−7	−6	−5	−4	−3	−2	−1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
A-chain																																					
Porcine								G	I	V	E	Q	C	C	T	S	I	C	S	L	Y	Q	L	E	N	Y	C	N									
Human																						E															
Bovine															A		V																				
Rat 1											D																										
Guinea pig											D					G	T		T	R	H			Q	S												
Chinchilla											D								T																		
Casiragua											D					N				R	N			L	T												
Coypu											D					N				R	N			M	S			D									
Porcupine											D					G	V							Q													
Cuis											D					R			T	S				R													
IGF-1											D	E			F	R	S		D		R	R			M		A	P	L	K	P	A	K	S	A	−
IGF-2												E			F	R	S		D		A	L			T		A	T	−	−	P	A	K	S	E	−
Porcine relaxin						R	M	T	L	S		K			E	V	G		I	R	K	D	I	A	R	L	−									
Rat relaxin		PCA	S			G	A	L	L	S					H	I	G		T	R	R	S	I	A	K	L	−									
Shark relaxin		A		T		S	P	A	M	S	I	K			I	Y	G		T	K	K	D	I	S	V	L	−									
PTTH-II											D	E			L	R	P			V	D	V		L	S		−									
PTTH-IV									V		D	E			I	Q	P		T	L	D	V		A	T		−									
B-chain																																					
Porcine								F	V	N	Q	H	L	C	G	S	H	L	V	E	A	L	Y	L	V	C	G	E	R	G	F	F	Y	T	P	K	A
Human																																					T
Bovine																																					
Rat 1										K						P																				S
Guinea pig										S	R						N				T			S		Q	D	D				I			D
Chincilla											K									D							D						M		
Casiragua								Y		G		R					Q			D	T			S		K	H				Y	R	P	S	E	−
Coypu								Y		S		R					Q			D	T			S		R	H				Y	R	P	N	D	−
Porcupine									F	N	R						N			D				V		K	D	K				S	R		D	−
IGF-1									−	G	P	E	T			A	E			D			Q	F			D				Y	F	N	K	P	T
IGF-2						A	T	R	P	S	E	T				G	E			D	T		Q	F			D				Y	F	S	R	P	A
Porcine relaxin				PCA	S	S	T	N	D	F	I	K	A			R	E			R	L	W	V	E	I		V	W	S	−	−	−	−	−	−	−
Rat relaxinR	V	S	E		E	W	M	D	Q	V	I	Q	V			R	G	Y	A	R		W	I	E	V		A	S	V	G	R	L	A	L	−	−
Shark relaxin					PCA	−	−	−	−	−	−	−	−			R	G	F	I	R		I	I	F	A		G	S	R	−	−	−	−	−	−	−
PTTH-II					PCA	Q	P	Q	A	V	H	T	Y			R			A	R	T		A	D	L	W		A		V	D	−	−	−	−	−
PTTH-IV					PCA	E	A	N	V	A	H		Y			R			A	N	T		A	D	L	W	D	T	S	V	E	−	−	−	−	−
PTTH-V							PCA	Q	V	H	T	Y				R			A	R	T		A	N	L	W		A		V	D	−	−	−	−	−

prevent the formation of an insulin-like fold although it might destabilise it. In mammalian insulins, removal of the A21 Asn and B30 Ala leads to changes in conformation as measured by circular dichroism [25]. These are similar to changes observed on dilution of native insulin which have been interpreted as being due to a reduction in the dimer content of an insulin solution [18]. There is strong evidence that dissociation of dimers causes changes in circular dichroism consistent with those observed for desAla desAsn insulin; there is no clear evidence for a difference in three-dimensional structure between this analogue and native insulin. They may also arise partly from rearrangement of B22 Arg which forms an ion pair with A21 α-carboxylate and may move to form a similar interaction with the new carboxylate at A20. In PTTH-II the B22 Arg is replaced by alanine and so this would not occur. A similar lack of a residue equivalent to A21 occurs in relaxins (see Table 1) which also lack the arginine at B22 [7-10,26].

The B chain COOH-terminus of PTTH-II terminates at B25 and thus is shorter than that of insulins. Evidence that only small conformational changes might occur is found in the extensive chemical, biological, spectroscopic and crystallographic studies [27,28] of despentapeptide (B25-B30)-insulin, which show that this modified insulin has an insulin-like conformation and biological activity.

The B chain NH_2-terminus is extended by three residues compared to human insulin in a similar way to that of relaxin. However, with the exception of B3 pyro-glutamic acid in both PTTH and relaxin, there is little homology between any of the sequences. In insulin this region has an extended conformation in 2-Zn crystals which contain a hexamer [5]. The sequences of the region B1-B5 are very variable between insulins of different species. In a similar way, the sequences of four PTTH B chain NH_2-termini show variation in both sequence and length in this region. We have modelled the B chain NH_2-terminus of PTTH-II with a generally extended conformation. However, secondary structure prediction [29] of the B chain NH_2-terminus shows that an α-helical conformation might occur as suggested for rat relaxin B chain NH_2-terminus [30].

In summary it appears that the PTTH-II molecule can have a compact globular structure with the exception of the variable and flexible B chain NH_2-terminus. If this is so, we can now ask questions concerning the arrangement of the surface residues. Do they give clues to formation of oligomers, antigenicity and receptor binding?

Let us first consider dimers, which in most mammalian, avian, fish and cyclostome insulins are stabilised both by hydrogen bonds and hydrophobic interactions [16]. The absence of B26 in PTTH-II makes this impossible as is shown by despentapeptide (B25-B30)-insulin which forms only monomers [32,33]. The four negatively-charged carboxylates from B25 Asp and the COOH-termini of the two molecules clustered around the dyad would also destabilise a dimer similar to that of insulin.

It is fascinating that B10 is histidine in PTTH-II. In most insulins this binds zinc in 2 Zn-hexamers. However, zinc binding is unlikely to be strong in the absence of hexamers and its existence in PTTH is probably an example of convergent evolution as both hagfish insulin, the most primitive vertebrate insulin sequenced, and insulin-like growth factors have no histidine at B10.

The remaining surface of insulin is characterised by mainly hydrophilic residues. The region involving A8-A10 is particularly variable in insulins and together with B2, B3 and B4 forms the main non-contiguous epitope for antibodies to human or porcine insulin raised in rabbits or guinea pigs. A8-A10 of PTTH is also variable but is different from human or bovine insulins. As it is probably less flexible due to the proline at A10, it is not surprising that PTTH-II does not bind antibodies raised in guinea pigs to porcine insulin [3].

There has been some debate about the exact extent of the insulin receptor binding region although it is generally recognised to involve non-continuous regions of the B chain (B23-B26, B12-B16) and A chain (A1, A5, A19-A21) in a surface region centred on B25 Phe. The arrangement of this region is dependent on maintenance of the tertiary structure of insulin. It involves hydrophobic residues and main-chain hydrogen bonding functions surrounded by more polar and charged residues. The nature of almost the whole of this surface region differs in PTTH-II. In particular, the absence of the central, hydrophobic phenylalanine at B25 would also be expected to reduce binding to insulin receptors.

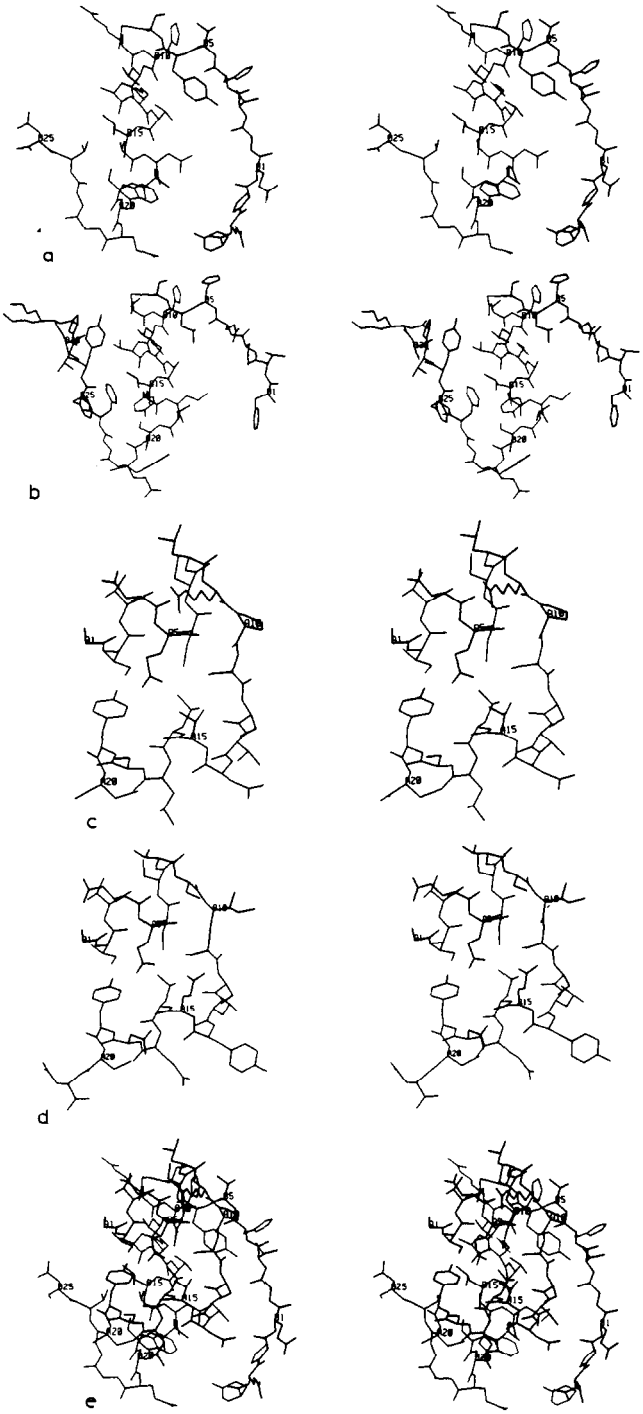

Fig.1. The proposed conformation (a) of the PTTH-II B chain based on the conformation of the porcine insulin B chain (b) defined by X-ray analysis [5,6]. Note the amphipatic nature of the α-helic (B9–B19) and the conservation of glycines at B8 and B23 to allow turns in the main-chain. (c) of the PTTH-II A chain based on the conformation of the porcine insulin A chain (d) defined by X-ray analysis and (e) a stereo view of the proposed model of PTTH-II.

One striking feature of the PTTH-II molecular surface demonstrated by our model is an extensive hydrophobic surface involving A13 Val (Leu), A15 Val (Gln), A17 Leu (Glu) and B18 Leu (Val). It has been suggested that a similar patch in insulin-like growth factors is involved in binding a carrier protein. In a similar way the surface hydrophobic region of PTTH-II may be important in binding other proteins, in the formation of oligomers or in receptor binding, but confirmation of this must await structure function studies by chemical modification, peptide synthesis or site specific mutagenesis.

The fact that PTTH-II almost certainly has an insulin-like tertiary structure is strong evidence for divergent evolution. An ancestral protein of insulin, insulin-like growth factor, relaxin and PTTH appears to have had at least an insulin-like fold of residues equivalent to A1-A20 and B4-B23, three disulphide bridges (A7–B7, A6–A11 and B19–A20) and a hydrophobic core of a volume similar to that of insulin. In all members of the family A2, A16, A19, B6, B11, B12, B15 and B18 are hydrophobic and together with the disulphides at A6–A11 and B19–A20 make up the core. Apart from the cystines only B8 Gly is invariant, probably for maintenance of conformation as outlined above. An insulin-like tertiary structure, is strong evidence for the existence of a single-chain precursor for PTTH. The single-chain precursors proinsulin and prorelaxin are well established and the tertiary structure of PTTH-II would not be expected to form in the absence of such a precursor. The existence of a two-chain insulin-like molecule in insects strongly suggests that the processing of the precursor evolved at an early stage, before the divergence of protostomes and deuterostomes, and that the loss of processing in the insulin-like growth factors I and II could have occurred at a fairly recent stage in prevertebrate deuterostomian evolution.

In many ways PTTH-II is intermediate between relaxins and insulins/insulin-like growth factors. PTTH-II resembles relaxins in the shorter COOH-termini of the A and B chains and the extension at the B chain NH_2-terminus. It resembles insulin in the terminus of the A chain both in length and sequence; there are 18 sequence identities between human insulin and PTTH-II, compared with only 11 between porcine relaxin and PTTH-II. Apart from residues with a structural role, A12 Ser and B21 Glu are conserved on the surface of PTTH-II and most insulins, although A12 appears to be threonine in PTTH-I and -III [3]. In a similar way B9 Arg and B13 Arg are conserved in PTTH-II and relaxins for reasons which are not presently understood.

The ancestral molecule of the insulin family appears to have been monomeric. This is not surprising as insulins are active at their receptors as monomers (see discussion in [5,16]). The most primitive insulin sequenced, that of the hagfish [23] is a dimer and modelling indicates that insulin-like growth factors may also form dimers [34]. Thus the ability to form dimers must have evolved at a prevertebrate stage. The formation of hexamers is a vertebrate phenomenon which appeared first in the elasmobranch insulins (and is lost in some mammalian insulins [35]).

ACKNOWLEDGEMENTS

This work was supported by Grants-in-Aids for Scientific Research (Nos 59360012, 60560133 and 61470134) from the Ministry of Education, Science and Culture of Japan, and by the UK Science and Engineering Research Council.

REFERENCES

[1] Williams, C.M. (1947) Biol. Bull. 93, 89–93.
[2] Ishizaki, H. and Suzuki, A. (1980) in: Neurohormonal Techniques in Insects, (Miller, T.A. ed.) pp. 244–276, Springer Verlag, New York.
[3] Nagasawa, H., Kataoka, H., Isogai, A., Tamura, S., Suzuki, A., Ishizaki, H., Mizoguchi, A., Fujiwara, Y. and Suzuki, A. (1984) Science 226, 1344–1345.
[4] Nagasawa, H., Kataoka, H., Isogai, A., Tamura, S., Suzuki, A., Mizoguchi, A., Fujiwara, Y., Suzuki, A., Takahashi, S. and Ishizaki, H. (1986) Proc. Natl. Acad. Sci. USA 83, 5840–5843.
[5] Blundell, T.L., Dodson, G.G., Mercola, D. and Hodgkin, D.C. (1972) Adv. Protein Chem. 26, 279–402.
[6] Dodson, E.J., Dodson, G.G., Hodgkin, D.C. and Reynolds, C.D. (1979) Can. J. Biochem. 57, 469–479.
[7] Schwabe, C., McDonald, J.K. and Steinetz, B.G. (1976) Biochem. Biophys. Res. Commun. 70, 397–401.

[8] James, R., Niall, H., Kwok, S. and Bryant-Greenwood, G. (1977) Nature 276, 544–546.
[9] Bedarkar, S., Turnell, W.G., Blundell, T.L. and Schwabe, C. (1977) Nature 270, 449–451,
[10] Isaacs, N., James, R. and Niall, H. (1978) Nature 271, 278–281.
[11] Jones, T.A. (1978) J. Appl. Cryst. 11, 268–272.
[12] Sternberg, M.J.E. and Blundell, T.L. (1985) Trends Biotech. 3, 228–235,
[13] Blundell, T.L. (1986) Phil. Trans. R. Soc.
[14] Hemmings, A.M., Foundling, S.I., Sibanda, L., Wood, S.P., Pearl, L.H. and Blundell, T. (1985) Biochem. Soc. Trans. 613, 1036–1041.
[15] Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S. and Weiner, P.J. (1984) Am. Chem. Soc. 106, 765–784.
[16] Blundell, T.L., Humbel, R.E. (1980) Nature 287, 781–787.
[17] Smith, L.F. (1972) Diabetes Suppl. 2, 457–460.
[18] Wood, S.P., Blundell, T.L., Wollmer, A., Lazarus, R. and Neville, R.W.J. (1975) Eur. J. Biochem. 55, 531–542.
[19] Bajaj, M. (1986) J. Biochem. Soc. in press.
[20] Bajaj, M., Wood, S.P., Blundell, T.L. and Waterfield, M. (1986) unpublished results.
[21] Blundell, T.L. and Horuk, R. (1981) Hoppe-Seyler's Z. Physiol. Chem. 362, 727–733.
[22] Venkatachalam, C.M. (1968) Biopolymers 6, 1425–1436.
[23] Cutfield, J.F., Cutfield, S.M., Dodson, E.J., Dodson, G.G., Emdin, S.F. and Reynolds, C.D. (1979) J. Mol. Biol. 132, 85–100.
[24] Chowdbury, S.A., Dodson, G.G., Cleasby, A. and Blundell, T.L. (1982) Proc. FDA-USP Workshop on Drug and Reference Standards for Insulin, 106–115.
[25] Goldman, J. and Carpenter, F.H. (1974) Biochemistry 22, 4566–4574.
[26] Niall, H., Hudson, P., Haley, J., Cronk, M. and Shine, J. (1981) Ann. NY Acad. Sci. 380, 13–21.
[27] Lu, Z. and Yu, R. (1980) Sci. Sin. 13, 1592–1599.
[28] Bi, R.C., Cutfield, S., Dodson, E.J. and Dodson, G.G. (1981) Acta Cryst. 01.3-01 (Suppl.).
[29] Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) J. Mol. Biol. 120, 97–120.
[30] Dodson, G.G. (1982) Int. J. Biol. Macromol. 4, 399–404.
[31] Bentley, A.G., Dodson, E.J., Dodson, G.G., Hodgkin, D.C. and Mercola, D.A. (1969) Nature 261, 166–168.
[32] Arquilla, E.R., Bromer, V.W. and Mercola, D.A. (1969) Diabetes 18, 193.
[33] Brandenburg, D. et al. (1973) in: Peptides 1962 (Hanson, H. and Jakubke, H.D. eds) pp. 270–283, North-Holland, Amsterdam.
[34] Blundell, T.L., Bedarkar, S. and Humbel, R.E. (1983) Fed. Proc. 42, no. 9.
[35] Blundell, T.L. and Wood, S.P. (1975) Nature 197–203.