Volume 224, number 1, 97–103 FEB 05227 November 1987

Structure and order of the protein and carbohydrate domains of prothrombin fragment 1

K. Harlos, C.W.G. Boys, S.K. Holland, M.P. Esnouf* and C.C.F. Blake

*Laboratory of Molecular Biophysics, University of Oxford, Rex Richards Building, South Parks Road, Oxford OX1 3QU and *Nuffield Department of Clinical Biochemistry, Radcliffe Infirmary, Oxford OX2 6HE, England*

Received 21 August 1987

The three-dimensional structure of prothrombin fragment 1 has been determined by X-ray crystallography at 3.8 Å resolution. The fragment is composed of a number of structural units, some of which are ordered while others are disordered. The ordered part of the structure includes a compact kringle unit, a helical domain and a carbohydrate chain. The kringle structure is organized around a close pair of buried disulfide bridges. One of its carbohydrate chains, that attached to Asn 101, is fully ordered, but the carbohydrate chain attached to Asn 77 appears to be disordered. The calcium binding unit is composed of a disordered part containing all ten γ-carboxyglutamic acid residues and an ordered part forming the helical domain. The highly conserved residues Phe 41, Trp 42 and Tyr 45, which form a hydrophobic cluster on the first helix, interact around a crystallographic two-fold axis with the equivalent residues in another molecule to form a dimer in the crystal.

Prothrombin fragment 1; Kringle; Carbohydrate structure; Protein crystallography

1. INTRODUCTION

The zymogens of blood coagulation and fibrinolysis have modular structures which show a considerable degree of commonality [1]. All, for example, contain a serine proteinase domain in the C-termini of their polypeptide chains, and between 3 and 6 small disulphide-bridged homology units in their N-termini. The proteinase domain gives rise to the enzymatic activity on activation, while the N-terminal units appear to be associated with binding interactions with other members of the cascade that give specificity and control to the activation process. Prothrombin is a characteristic member of this group of proteins [2]. The serine proteinase at the C-terminus of its polypeptide chain becomes thrombin on activation. At the N-terminus there are three small domains: residues 1–65 correspond to a vitamin K-dependent domain that contains 10 γ-carboxyglutamic acid (Gla) residues which form high-affinity binding sites for calcium ions; residues 66–156 and 157–248 represent a pair of kringle units, each with a characteristic pattern of 3 disulphide bridges, the second of which is involved in binding prothrombin to its protein cofactor, factor V. Homologous vitamin K-dependent regions are also present in coagulation factors VII, IX and X, and in proteins C, S and Z, while kringle units are also found in factor XII and the fibrinolytic proteins, plasminogen, urokinase and tissue plasminogen activator.

X-ray analyses of members of this group of proteins would shed more light on their complex structure/function relationships. Unfortunately the whole proteins of coagulation and fibrinolysis do not crystallise readily, and crystals suitable for X-

Correspondence address: K. Harlos or C.C.F. Blake, Laboratory of Molecular Biophysics, University of Oxford, Rex Richards Building, South Parks Road, Oxford OX1 3QU, England

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50 © 1987 Federation of European Biochemical Societies

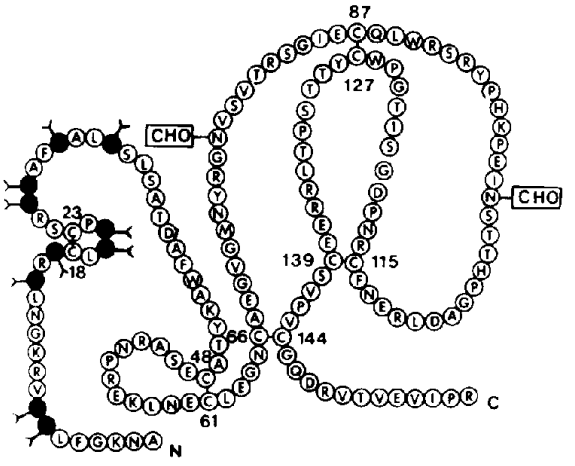

Fig.1. The amino acid sequence of bovine prothrombin fragment 1. Reproduced from Olsson et al. (1982) with kind permission.

ray analysis have been obtained only from thrombin [3], and from fragment 1 of prothrombin [4–6]. Fragment 1 is an autocatalytic product of prothrombin, consisting (see fig.1) of residues 1–156, corresponding to the vitamin K-dependent region and the first kringle unit, which carries two large carbohydrate chains. X-ray analyses of fragment 1 at 4 Å resolution [6], and 3.5 Å resolution [7] yielded preliminary structural information on the calcium-bound and calcium-free forms, respectively. Extension of the resolution to 2.8 Å [8] has provided much basic information on the kringle. We report here an X-ray study of the calcium-free form of fragment 1 that adds information on the carbohydrate structure, the interaction of calcium with the vitamin K region, and the kringle binding site.

2. EXPERIMENTAL

2.1. *Protein crystallisation and derivatisation*

Crystals of fragment 1 were grown by vapour diffusion [9] using the hanging drop method. Large crystals were grown at pH 6.2–6.7 from protein concentration of 25 mg/ml using 26% polyethylene glycol 8000 with 0.1 M ammonium sulphate. The crystals belong to the tetragonal space group $P4_32_12$ with cell dimensions similar to those reported earlier [4,5]. Although large, the crystals grown by this procedure exhibited very variable diffraction patterns which caused much difficulty in subsequent work.

The structure of the crystals was solved using a single isomorphous derivative and the Wang procedure for solvent flattening [10]. The only usable isomorphous derivative was prepared by soaking fragment 1 crystals in solutions of Baker's dimercurial [11] at 10 mM concentrations. A single major site was observed, presumably representing the bound mercury atom, surrounded by a number of minor sites which probably represent alternative locations of the second mercury atom.

2.2. *Data collection and phasing*

Three-dimensional data sets were collected to 3.8 Å resolution for both native and derivative crystals, using a five-circle diffractometer. Because of crystal instability no anomalous scattering data could be collected from the derivative crystals between 6 Å and 3.8 Å resolution. The phases in this region were improved by Wang's procedure [10] for interactive single isomorphous replacement (ISIR). This procedure increased the figure of merit from 0.51, obtained by using the dimercurcial derivative alone, to 0.79 for the 2673 reflections to 3.8 Å resolution, and greatly improved the quality of the map. Table 1 shows the variation in

Table 1

Refinement statistics

Resolution (Å)	Number of reflections	R_K	R_C	$\langle F_H \rangle / \langle E \rangle$
40.0–11.8	112	0.096	0.281	3.35
11.8– 8.5	169	0.087	0.424	2.89
8.5– 6.9	216	0.078	0.382	3.50
6.9– 6.0	208	0.104	0.453	3.24
6.0– 5.1	440	0.113	0.348	1.98
5.1– 4.5	501	0.110	0.518	1.86
4.5– 4.1	546	0.118	0.514	1.66
4.1– 3.75	445	0.134	0.561	1.57

R_K = Kraut R factor = $||F_{PH}(obs)| - |F_{PH}(calc)||/|F_{PH}(obs)|$, R_C, Cullis R factor for centric reflections; $\langle F_H \rangle$, root mean square heavy atom structure factor; $\langle E \rangle$, root mean square lack of closure error

phasing with resolution. Detailed crystallographic analysis will be published elsewhere.

3. RESULTS

3.1. *The crystal structure*

A striking feature of the electron density of this crystal form is the presence of very large liquid volumes, even greater than the anticipated 60%. As can be seen in fig.2 at any one z-level elongated fragment 1 molecules line up end-to-end to form parallel chains separated from one another by about 50 Å of liquid. As a result of the crystallographic four-fold screw axis, at $z \pm 1/4$ similarly spaced chains of molecules run at right angles to those above (and below) to form a three-dimensional lattice-work whose 'holes' are filled with liquid. From each protein molecule a long forked electron density feature projects into the centre of the liquid regions: this represents the carbohydrate moiety linked to Asn 101.

Although calculated at a medium resolution of 3.8 Å the electron density of the protein part of fragment 1 is defined remarkably well. The course of the polypeptide chain is marked by a continuous run of electron density covering residues 36–156. Some segments of the α-helical structure can be readily identified, as can the larger side-chains and disulphide bridges. Since the molecular structure is clearly organized into a number of discrete units or domains, the following description will be in terms of these separate units.

3.2. *The kringle domain*

The kringle homology unit is conventionally defined as the chain between and including the disulphide bridge, 66–144, with the additional residues 145–156 in fragment 1 corresponding to the larger part of the link between kringle 1 and kringle 2. In the crystal, all these residues (66–156) form a compactly folded structural unit which is clearly the dominating element of the structure of fragment 1 (figs 3 and 4). The polypeptide chain of the kringle is organised almost entirely in extended conformations, with a small number of reverse turns or other sharp bends required to fold the chain into a compact form. There is a possibility of one helix-like turn near the carbohydrate link at residue 101, but otherwise what regular structure is present is of the β-type. There is no real indication

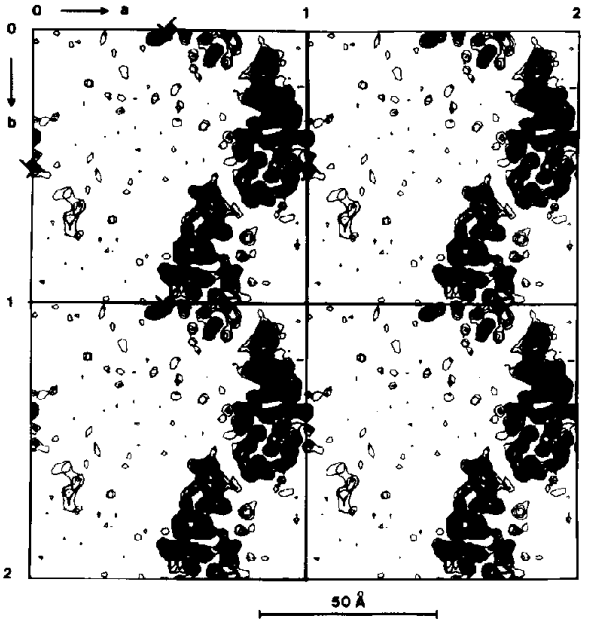

Fig.2. Sections 23/64 to 27/64 perpendicular to the c axis of the electron density map of several unit cells of fragment 1. The crystallographic four-fold screw axis of one unit cell is indicated. Contoured in arbitrary units from 12 to 96 in steps of 12.

of any β-sheet structures, but there do appear to be a number of antiparallel β-hairpin or ribbon structures. At the present resolution it is not possible to define hydrogen-bond interactions with great precision, but our map suggests that residues 64–66 and 74–76; 80–82 and 83–85; 87–90 and 111–116; 140–143 and 148–151 are the most likely candidates for β-interactions.

The overall shape of the kringle is a triaxial ellipsoid of dimensions 39 Å × 30 Å × 22 Å with the polypeptide chain running mostly parallel to the surface. However one segment of chain residues 114–119 runs through the core which is otherwise formed from two disulphide bridges, 87–127 and 115–139, and some large buried hydrophobic side chains. These include Met 72, Tyr 74, Trp 90, Ile 100, Trp 126 and Tyr 128. The two disulphide bridges are located approximately at right angles to one another and with their sulphur atoms in contact with one another. These disulphide bridges appear to be the characteristic and dominant

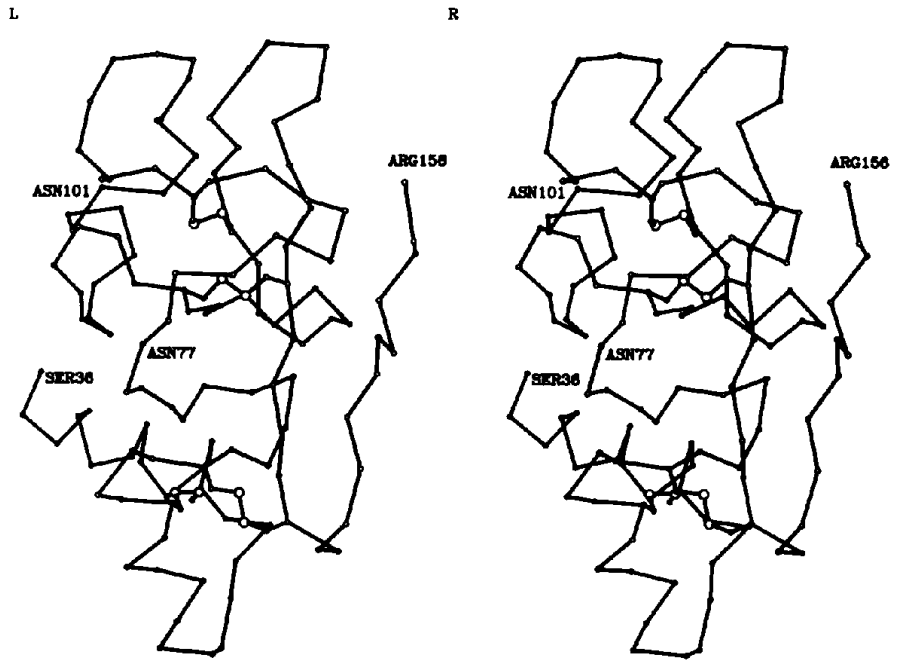

Fig.3. Stereo view of the C_α-backbone of fragment 1 looking down the *c* axis.

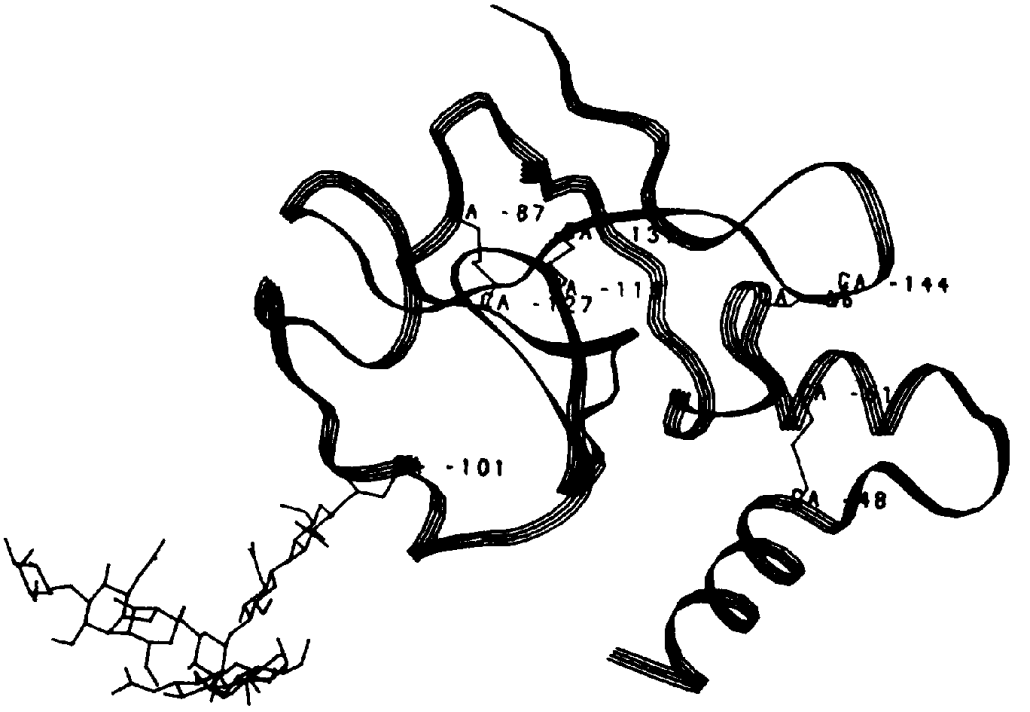

Fig.4. Ribbon drawing of the C_α-backbone of fragment 1 with the carbohydrate attached to Asn 101 included and the disulphide bridges marked. View approximately down the *a* axis.

structural motif of the kringle domain. The pattern of invariant residues in the kringle fold is largely determined by these aspects of the three-dimensional structure, which have been more thoroughly described elsewhere [8].

The C-terminal carboxyl group and possibly the C-terminal arginine side-chain make contact with a neighbouring fragment 1 kringle near a shallow cavity located between residues 121 and 136. It is interesting to note that the corresponding residues have been implicated in the binding site of kringle 4 of plasminogen, which exhibits specificity for arginine or lysine residues [12–14]. It should be noted however that kringle 1 of prothrombin does not exhibit this binding capacity, but it is possible that this function is only weakly expressed on kringle 1, perhaps through an amino acid substitution, but able to reveal itself in the very high concentration conditions represented by the crystal.

3.3. *The carbohydrate chains*

Fragment 1 has two carbohydrate chains, *N*-linked to Asn 77 and Asn 101. Three slightly different polysaccharide chains A, B and C have been reported for prothrombin [15] which are all of the biantennary type and which all possess the usual inner core of Asn-linked carbohydrates. It is at present not known which particular sequence is linked to Asn 77 or Asn 101. In the electron density map an extensive forked chain of density emerges from the position of side-chain 101, but very little density and certainly no more than the equivalent of one sugar residue is associated with Asn 77 and we must assume that this polysaccharide chain is disordered. There is no obvious environmental reason for this distinction in the ordering of the carbohydrate chains. The sugar chain from 101 projects away from the parent kringle structure and does not interact with any other ordered structure in the protein.

The electron density emerging from Asn 101 has been fitted on the graphics system (see fig.5) with sugar sequence C of Mizuochi [15]

Fig.5. Fitting of the carbohydrate chain to density emerging from Asn 101. View approximately down the *a* axis.

since there was no indication of (α2–6) linked neuraminic acid side branches at the appropriate positions. The fitting showed that the forking of the electron density occurred at mannose residue 3 and that both of the branches are well ordered up to the galactose. At the present resolution of 3.8 Å the Man(α1–3)Man link cannot be distinguished from the Man(α1–6)Man link at the forking point and therefore the branches can only be assigned in a tentative way. As can be seen from fig.5 one branch contains a rather bulky piece of density at the end. This density probably represents the terminal sialic acid as well as the galactose, although at the moment only the galactose has been fitted.

3.4. *The calcium binding unit*

In prothrombin the high-affinity calcium binding unit is represented by residues 1–65, with the Gla residues restricted to the first 33 residues. In proteins such as factors VII, IX and X, and protein C [16–19], the calcium binding region shows a close homology over residues 1–48, but then enters

NeuAcα2–6Galβ1–4GlcNAcβ1–2Manα1–3
\
Manβ1–4GlcNAcβ1–4GlcNAcβ1-Asn
/
NeuAcα2–6Galβ1–4GlcNAcβ1–2Manα1–6

into the EGF-like homology units with no equivalent to residues 49–65 of prothrombin. The electron density maps of fragment 1 show essentially no features for residues 1–35, but residues 36–65 are well defined as a compact unit formed from two α-helices, residues 37–48 and residues 54–63, linked by a disulphide bridge between residues 48 and 61. The missing region in factors VII, IX and X and protein C therefore corresponds almost precisely with the second α-helix.

A striking feature of the helical unit is the presence of a cluster of aromatic hydrophobic residues, Phe 41, Trp 42 and Tyr 45 located on the outer edge of the first helix. This cluster of residues is highly conserved and the three residues are invariant in prothrombin and the three blood coagulation proteins listed above. In the crystal these residues are not exposed to the solvent but interact around a crystallographic two-fold axis with the equivalent region in a neighbouring molecule (fig.6). Phe 41 and Tyr 45 are involved in close face-to-face interactions, while Trp 42 makes interactions with both of them and with Leu 58 and Leu 62 from helix 54–63 of the symmetry related molecule. As this hydrophobic interaction represents a major intermolecular interaction in the crystal, the two molecules are packed in the crystal like a dimer. In view of the numerous reported dimerizations of fragment 1 in solution in the presence of calcium [20,21], this hydrophobic interaction could well be one step in the dimer formation of fragment 1 in solution.

The lack of density for residues 1–35, which includes the complete Gla-containing region, must indicate that either this part of the molecule is absent in the crystal or that it is spatially disordered. Although the Leu 35–Ser 36 bond is a well-defined cleavage site for chymotrypsin [22] we have no evidence that the Gla-containing region has been cleaved from the crystalline protein. On the other hand, we do have evidence of the calcium binding ability of crystalline fragment 1 through soaking experiments using either calcium or strontium, which show a strong metal binding site about 10 Å on the N-terminal side of residue 36, i.e. possibly near Gla 30 and 34. The reasonable inference of this experimental result, which will be described in more detail elsewhere, is that the Gla-containing region is present but disordered.

4. DISCUSSION

A major aspect of fragment 1 is that its molecular structure is divided into a number of autonomous units some of which are ordered, while others are disordered. The kringle, which is the largest unit, has a well-ordered structure organised around a close pair of buried disulphide bridges. Of its two similar carbohydrate chains, that attached to Asn 101 is fully ordered, but that attached to Asn 77 is disordered. The calcium binding unit seems to be composed of two parts: residues 1–35 containing all ten invariant γ-carboxyglutamic acid residues appear to be totally

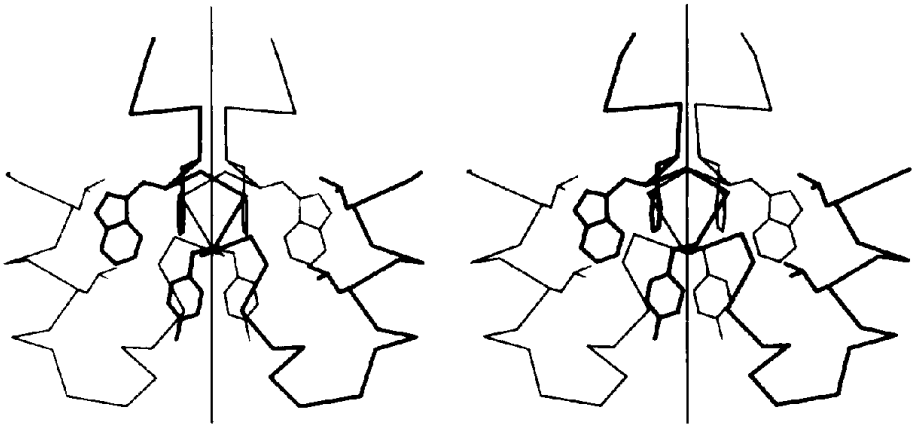

Fig.6. Stereo view of the hydrophobic interaction around the crystallographic two-fold axis with the involved residues marked.

disordered, while residues 36–65 form a well ordered helical structure.

It is possible that this pattern of structural order and disorder has functional significance. For example the indication of an ordered Gla-containing region in the crystals of the calcium-bound form of fragment 1 [6] and the evidence presented here, and in Park and Tulinsky [8], of disorder in the absence of calcium, is strongly indicative of a disorder-order transition in residues 1–35 on the binding of calcium ions. Dimerization of fragment 1 is induced by calcium ions [20,21] and the involvement of the helical part of the calcium binding unit in the dimer-like interaction in the crystal is suggested by an extensive hydrophobic interface. The binding site located in other kringles [12–14] appears to be used by prothrombin kringle 1 as one of the major interaction sites in the crystals.

ACKNOWLEDGEMENTS

We acknowledge the financial support of the Medical Research Council for the project. We would also like to thank the Medical Research Council, the Deutsche Forschungsgemeinschaft (grant Ha 1291/1-2) and the EPA Cephalosporin Fund for financial support to one of us (K.H.). We are grateful to Dr D.I. Stuart for valuable assistance with the data processing. C.C.F.B. and M.P.E. are members of the Oxford Enzyme Group which is supported by the Science and Engineering Research Council.

REFERENCES

[1] Patthy, L. (1985) Cell 41, 657.
[2] Magnusson, S., Sottrup-Jensen, L., Petersen, T.E. and Claeys, H. (1975) in: Proteases and Biological Control (Reich, E. et al. eds) vol.2, p.123, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[3] Tsernoglou, D. and Petsko, G.A. (1977) in: Chemistry and Biology of Thrombin (Lundblad et al. eds) p.123, Ann Arbor Science Publishers, Ann Arbor.
[4] Aschaffenburg, R., Blake, C.C.F., Burridge, J.M. and Esnouf, M.P. (1977) J. Mol. Biol. 114, 575.
[5] Hu-Kung, W., Tulinsky, A. and Nelsestuen, G.L. (1980) J. Biol. Chem. 255, 10523.
[6] Olsson, G., Andersen, L., Lindquist, O., Sjölin, L., Magnusson, S., Peterson, T.E. and Sottrup-Jensen, L. (1982) FEBS Lett. 145, 317.
[7] Tulinsky, A., Park, C.H. and Rydel, T.J. (1985) J. Biol. Chem. 260, 10771.
[8] Park, C.H. and Tulinsky, A. (1986) Biochemistry 25, 3977.
[9] Davies, D.R. and Segal, D.M. (1971) Methods Enzymol. 22, 266.
[10] Wang, B.C. (1985) Methods Enzymol. 115B, 90.
[11] Blundell, T.L. and Johnson, L.N. (1976) Protein Crystallography, p.209, Academic Press, London.
[12] Winn, E.S., Hu, S., Hochschwender, S.M. and Laursen, R.A. (1980) Eur. J. Biochem. 104, 579.
[13] Hochschwender, S.M. and Laursen, R.A. (1981) J. Biol. Chem. 256, 11172.
[14] Trexler, M., Vali, Z. and Patthy, L. (1982) J. Biol. Chem. 257, 7401.
[15] Mizuochi, T., Yamashita, K., Fujikawa, K., Kisiel, W. and Kobata, A. (1979) J. Biol. Chem. 254, 6419.
[16] Hagen, F.S., Gray, C.L., O'Hara, P., Grant, F.J., Saari, G.C., Woodbury, R.G., Hart, C.E., Insley, M., Kisiel, W., Kurachi, K. and Davie, E. (1986) Proc. Natl. Acad. Sci. USA 83, 2412.
[17] Bucher, D., Nebelin, E., Thomsen, J. and Stenflo, J. (1976) FEBS Lett. 68, 293.
[18] Enfield, D.L., Ericsson, L.H., Walsh, K.A., Neurath, J. and Titani, K. (1975) Proc. Natl. Acad. Sci. USA 72, 16.
[19] Fernlund, P., Tufvesson, A. and Stenflo, J. (1978) Proc. Natl. Acad. Sci. USA 75, 5889.
[20] Prendergast, F.G. and Mann, K.G. (1977) J. Biol. Chem. 252, 840.
[21] Jackson, C.M., Peng, C.W., Brenckle, G.M., Jonas, A. and Stenflo, J. (1979) J. Biol. Chem. 254, 5020.
[22] Morris, H.R., Dell, A., Petersen, T.E., Sottrup-Jensen, L. and Magnusson, S. (1976) Biochem. J. 153, 663.