EYE-LENS PROTEINS: THE THREE-DIMENSIONAL STRUCTURE OF β-CRYSTALLIN PREDICTED FROM MONOMERIC γ-CRYSTALLIN

Graeme WISTOW, Christine SLINGSBY and Tom BLUNDELL

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

and

Huub DRIESSEN, Wilfried DE JONG and Hans BLOEMENDAL

Department of Biochemistry, University of Nijmegen, Geert Grooteplein noord, 6500 HB Nijmegen, The Netherlands

Received 14 August 1981

1. Introduction

Recent sequence determinations [1–4] have demonstrated that two of the three classes of mammalian lens-specific proteins, the β- and γ-crystallins, form a single superfamily of proteins, accounting for >50% of the protein in the mammalian lens.

Although the unrelated lens protein, α-crystallin, is composed of two primary gene products which associate to large varying aggregates of M_r-values averaging 800 000 [5–7], β-crystallins are an even more heterogeneous class. It is likely that there are at least 5 polypeptides involved, forming a wide range of aggregates of M_r 50 000–200 000 [8,9]. The predominant β-subunit has been designated βBp [10], and is present in all vertebrate classes [11]. The bovine βBp monomer has an M_r of 23 200 and has been sequenced [1]. When isolated it forms a stable dimer (C. S., G. Berbers, unpublished).

In contrast, γ-crystallins are monomers, proteins of $M_r \sim 20\ 000$. At least 5 closely related gene products have been identified in the bovine lens [2–4]. The sequence of γII and partial sequences of γIIIa,b and γIVa,b have been published. The three-dimensional structure of γII has been determined at 2.6 Å resolution [12]. This structure shows a rather exact 2-fold repeat and a slightly less exact 4-fold repeat in the tertiary fold, which has implications for the role of the molecule as a stable, structural protein. It may also imply evolutionary development by gene multiplication. However, the amino acid sequence does not show such a strong internal homology except for some residues at particular features in the γII fold which do display a 4-fold repeat [12]. There is, however, a general conservation of functional types of residue between the 4 folding motifs of the structure.

Using predictive and computer graphics techniques, we now show that the principal component of the β-crystallins, βBp [10] should have a three-dimensional structure involving 4 structurally homologous motifs similar to that determined for γ-crystallin by X-ray analysis [12]. However, in βBp, there are extensions of the polypeptide chain at both N- and C-termini which we predict form flexible extended arms which may be involved in the self-association to dimers and possibly higher oligomers.

2. Materials and methods

The complete sequence of γII is arranged to show the equivalent residues of the 4 folding motifs in vertical columns (table 1). Residues 16–193 of βBp are then aligned with the γII sequence for maximum homology and residues which are conserved or conservatively varied are boxed.

Four-state secondary structure predictions were made by the methods of Chou and Fasman [13] and Garnier et al. [14]. For the latter, the decisions constants used were those which gave a prediction for γII closest to the structure determined by X-ray analysis. Helix, turn and coil conformations are not shown in table 1. The helical regions of the γII structure are

rather irregular and do not constitute α-helix. The methods used give mixed predictions for these regions and also for the unusual foldover feature between strands a and b (see [12]).

Using the coordinates of γII as a guide, the homologous part of the βBp sequence was modelled using the program FRODO [17] for interactive computer graphics on an Evans and Sutherland picture system II (by T. A. Jones as modified by I. J. Tickle). The main-chain conformation of βBp was constructed with torsion angles identical to those for γII except in regions adjacent to deletions or insertions. The program FITZ, written by Dr G. Taylor, was used to facilitate docking of macromolecules on the Evans and Sutherland picture system II.

3. Results and discussion

Whilst there is no homology between the sequences of α- and γ-crystallin, there is a marked similarity between those of γII and βBp. This relationship is even more marked when considering the modified γII-crystallin sequence suggested by X-ray analysis (see table 1: this is the sequence referred to throughout. All residues are numbered relative to this sequence for both proteins to facilitate comparison.). In particular the 4-fold repeated residues conserved within the γII sequence are present in that of βBp.

When the sequence of the two proteins were analysed for secondary structure tendencies, the γII sequence, which served as a control, gave a secondary

Table 1

Motif 1	βBp	16 Pro	Lys	Ile	20 Ile	Ile	Phe	Glu	Gln	Glu	Asn	Phe	His	**Gly**			30 His	Ser	Gln	Glu	Leu	Asx	Pro	Gly	Asx
Motif 1	γII	1 Gly	2 Lys	3 Ile	4 Thr	5 Phe	6 **Tyr**	7 Glu	8 Asp	9 Arg	10 Gly	11 Phe	12 Gln	13 **Gly**			14 His	15 Cys	16 Tyr	17 Gln	18 Cys	19 Ser	20 Ser		21 Asn
Motif 3	βBp	His	Lys	Ile	Thr	110 Leu	**Tyr**	Glu	Asn	Pro	Asn	Phe	Thr	**Gly**	Lys	120 Lys	Met	Glu	Val	Ile	Asp	Asp		
Motif 3	γII	84 Phe	85 Arg	86 Met	87 Arg	88 Ile	89 **Tyr**	90 Glu	91 Arg	92 Asp	93 Asp	94 Phe	95 Arg	96 **Gly**		97 Gln	98 Met	99 Ser	100 Glu	101 Ile	102 Thr	103 Asp		

Motif 2	βBp	Gly	Pro	Trp	Val	60 Gly	**Tyr**	Glu	Gln	Ala	Asn	Cys	Lys	**Gly**		70 Glu	Gln	Phe	Val	Phe	Glu	Lys	Gly
Motif 2	γII	40 His	41 Ser	42A Trp	43A Phe	44A Val	45A **Tyr**	47 Gln	48 Arg	49 Pro	50 Asp	51 Tyr	51A Arg	52 **Gly**		53 His	54A Gln	54B Tyr	54C Met	54D Leu	55 Gln	56 Arg	57 Gly
Motif 4	βBp	Gly	Thr	150 Trp	Val	Gly	**Tyr**	Gln	Tyr	Pro	Gly	Tyr	Arg	160 **Gly**		Leu	Gln	Tyr	Leu	Leu	Glu	Lys	Gly
Motif 4	γII	123 Gly	124 Ser	125 Trp	125A Val	125B Ile	125C **Tyr**	126 Glu	127 Met	128 Pro	129 Ser	130 Tyr	131 Arg	132 **Gly**		133 Arg	134 Gln	135 Tyr	136 Leu	137 Leu	138 Arg	139 Pro	140 Gly

←——— a ———→　　　　　　　　　　←——— b ———→

The complete sequence of γII is arranged to show the equivalent residues of the four folding motifs in vertical columns. Residues 16–193 of βBp are then aligned with the γII sequence for maximum homology and those residues which are conservatively varied are boxed. Residues which are absolutely conserved in all folding motifs are shown in bold type

structure prediction that corresponded well with the structure determined by X-ray analysis. The equivalent regions of βBp gave similar results (table 1).

Clearly then, it is reasonable to suggest that the tertiary folds of the homologous regions of the two proteins are the same. The only major differences are the N- and C-terminal extensions of βBp. The intervening βBp sequence requires only 4 inserted residues to fit that of γII.

Supporting evidence for this assumption can be found in ORD and CD spectroscopic studies of β- and γ-crystallins where all agree that the proteins contain regions of β-structure with little or no α-helix [15,16].

The relative alignments of the βBp and γII sequences were established not only on the basis of sequence homology and on the secondary structure predictions, but also on the need when model building to form a close-packed hydrophobic core. A stereo view of the model is shown in fig.1.

Insertions at 20a, 67a, 67b and 96a lengthened 3 loops but did not significantly alter the character of the surrounding regions. The changes made were found to be concerted and to make structural 'sense'. For example, Tyr 62 in γII is replaced by tryptophan in βBp, while the nearby Trp 64 is conserved. With no other change it is difficult to accommodate the bulky new side-chain without it protruding into solvent or greatly disturbing the position of the conserved Trp 64. However, the extension of the following loop by Arg 67a and Arg 67b allows the tryptophan side-

Table 1 (continued)

```
                40
Pro Cys Leu Lys Glu Thr Gly     Val Glu            Lys | Ala | Gly | Ser | Val | Leu |     | Val | Gln Ala
                                                                      50

22  23  24  25  26A 27  28      29  30             31  | 32  | 33  | 34  | 35  | 36  | 37  | 38  39
Asn Cys Leu Gln Glu Pro Tyr     Phe Ser            Arg | Cys | Asn | Ser | Ile | Arg | Val | Asp Val

                130
Asp Val Pro Ser Phe His Ala His Gly Tyr Gln Glu    Lys | Val | Ser | Ser | Val | Arg |     | Val | Gln Ser
                                                         140

104A 105 106A 107 108 109 110 111 112 113 114 115  116 | 117 | 118 | 119A | 119B | 119C | 120 | 121 122
Asp  Cys Pro  Ser Leu Gln Asp Arg Phe His Leu Thr  Glu | Val | Asn | Ser  | Val  | Arg  | Val | Leu Glu
```

```
                    80                               90                                                          100
Glu | Tyr | Pro Arg Trp Asp Ser     Trp | Thr Ser Ser Arg Arg Thr Asp Ser | Leu | Ser | Ser | Leu | Arg | Pro | Ile | Lys Val Asp Ser Gln Glu

58  | 59  | 60  61  62  62A 63      64  | 65  66  67      68  69  70 | 71  | 72  | 73  | 74  | 75  | 76  | 77  | 78  79  80  81  82  83
Asn | Tyr | Pro Gln Tyr Gly Gln     Trp | Met Gly Phe     Asp Asp Ser| Ile | Arg | Ser | Cys | Arg | Leu | Ile | Pro Gln His Thr Gly Thr

      170                                            180                                                         190
Asp | Tyr | Lys Asp Ser Gly Asp     Phe | Gly Ala Pro     Gln Pro Gln | Val | Gln | Ser | Val | Arg | Arg Ile | Arg Asp Met Gln

141 | 142 | 143 144 145 146 147     148 | 149 150 151     152 153 154 | 155 | 156 | 157 | 158 | 159 | 160 161 | 162 163 164 165
Glu | Tyr | Arg Arg Tyr Leu Asp     Trp | Gly Ala Met     Asn Ala Lys | Val | Gly | Ser | Leu | Arg | Arg Val | Met Asp Phe Tyr
```

← c →   ←——————— d ———————→

(———) Indicates that both secondary structure analyses predicted extended (β-sheet) conformation; (. . .) indicates that one or other method predicted extended conformation. The four strands of β-sheet experimentally found to be common to all four motifs of γ-II are indicated as a, b, c and d
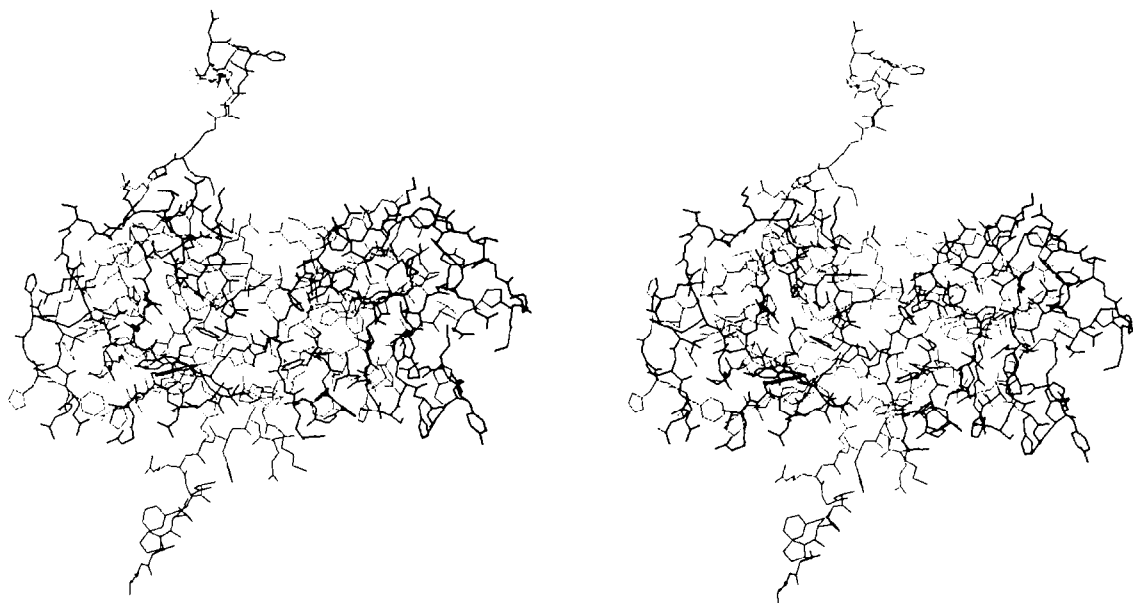
Fig.1. A stereo pair showing a model of the βBp protomer. The 'arms' are shown in an extended form as used in the compact dimer model (fig.3).

chain to be swung round into a more internal environment and to be protected from the solvent to some extent by the lengthened loop. The two inserted arginine residues then lie conveniently close to two glutamate residues, 47 and 53 which, respectively, replace glutamine and histidine residues in the γII sequence.

In spite of this tendency for little change or conservative change, there is one part of the molecular surface where βBp-crystallin appears to be more hydrophobic than γII. Each of the 4 folding motifs of γII is characterised by an unusual foldover feature involving the only residues to be completely conserved within the sequence. In motif I of γII, the chain folds between Glu 7 and Gly 13 (conserved). This causes Phe 11 to pack down on to the internal, conserved Tyr 6. Two other conserved residues, Ser 34 and Arg 36 stabilise the structure. There is a hydrogen bond from Oγ of residue 34 to N of residue 11, and the arginine residue protects the whole assembly from solvent with its guanidinium moiety close to the polar OH group of Tyr 6. In motif I of βBp the residues Glu 7, Gly 13, Phe 11 and Ser 34 are present. However, Tyr 6 is replaced by the more hydrophobic Phe 6 and Arg 36 by Leu 36. In the 8 'equivalent' motifs of γII and βBp this is the only place where the tyrosine and arginine residues are not conserved. Fur-
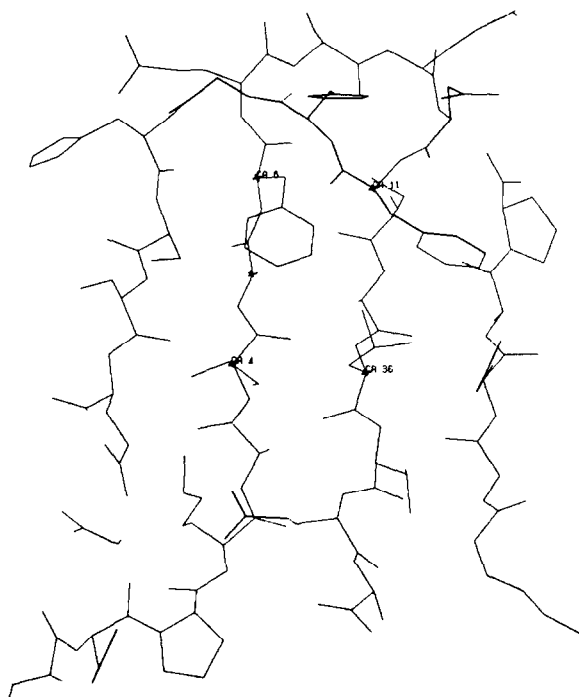
Fig.2. The hydrophobic patch on the N-terminal domain of βBp.

Table 2
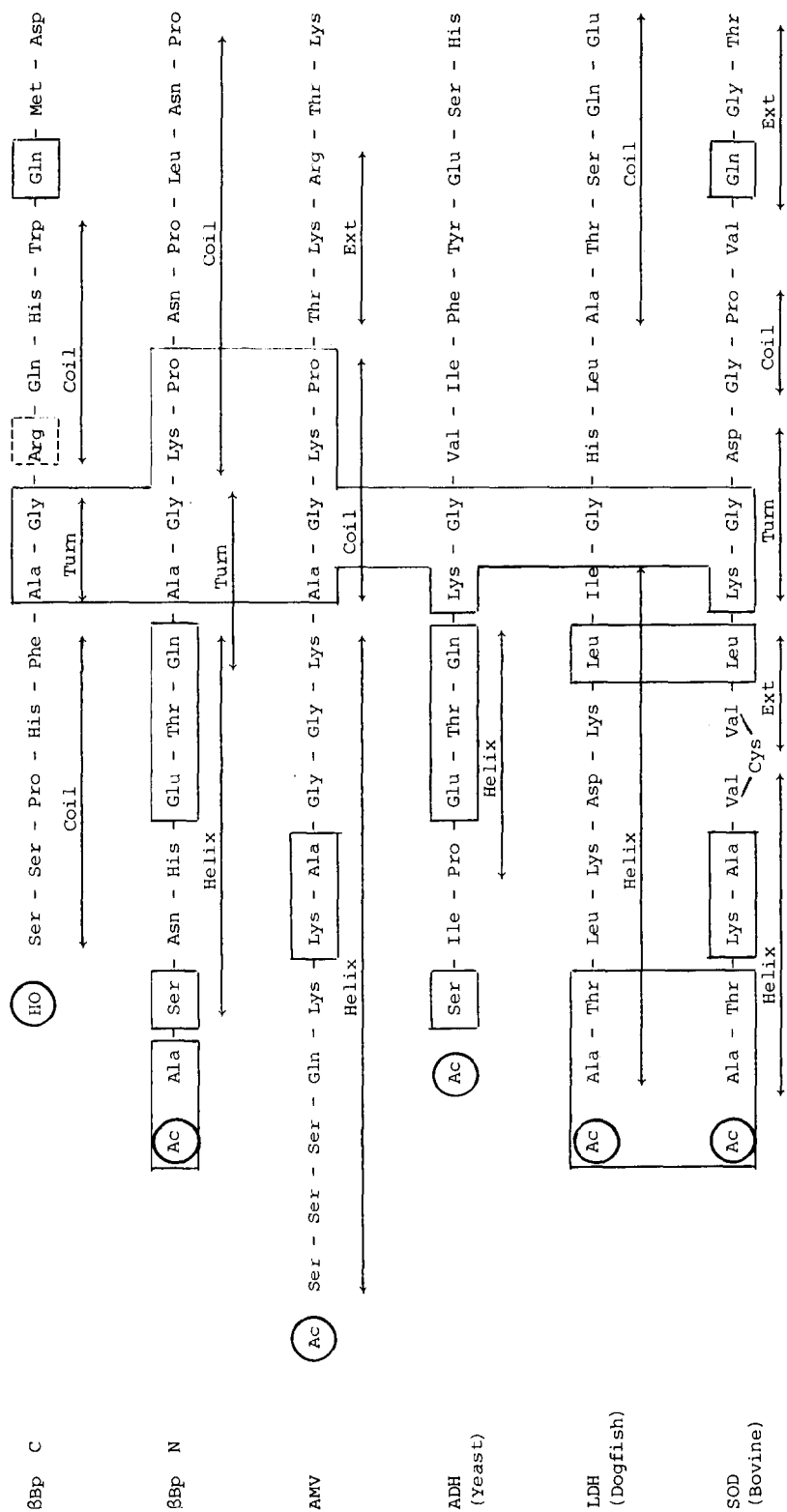
N-terminal sequences of AMV[23,24], ADH[25], LDH[22] and SOD[26] compared with the N- and C-terminal sequences of βBp [1]. The C-terminal sequence is included in the reverse direction. Beneath each sequence is the secondary structure predicted by the methods of Garnier et al. [14]. Areas of sequence homology are indicated by surrounding enclosures. AC = acetyl

thermore, Thr 4 in motif I of γII is replaced by the non-polar Ile 4 in βBp. This is similar to the situation in the interdomain contact region of γII where the equivalent positions in motifs II and IV are occupied by hydrophobic residues.

Overall, this produces a patch of non-polar residues on folding unit 1 of βBp (fig.2) and provides the most obvious site for potential intermolecular contact on the part of the predicted structure homologous

with γII. There is another, smaller patch at the other end of the molecule where Glu 100 is replaced by valine in motif III in the C-terminal domain of βBp, and Arg 87 is replaced by the uncharged Thr 87. However, the major differences between monomeric γII and dimeric βBp are the N- and C-terminal extensions of the latter.

Chou and Fasman [13] and Garnier et al. [14] prediction methods indicate that the 14 C-terminal
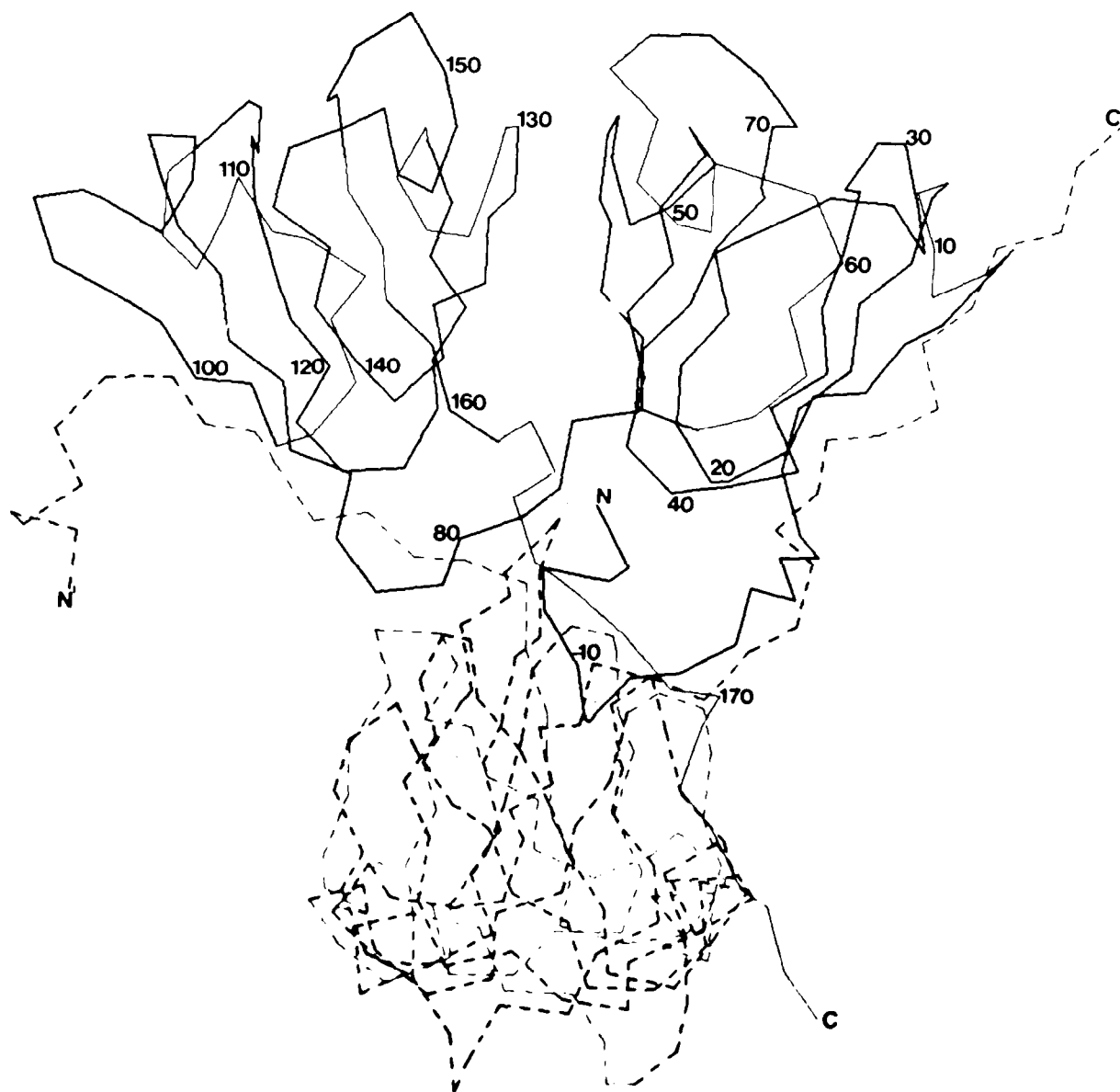
Fig.3. The compact dimer model of βBp. Two protomers pack together as indicated to form a structure with pseudo 222 symmetry. Only Cα positions are shown. One molecule is in continuous, the other in dotted lines.

and 15 N-terminal residues of βBp have little ten-
dency to form regular secondary structures. However,
the predictions suggest that the 6 residues at the
N-terminus could form a short helix, followed by a
reverse turn. This is reminiscent of the N-terminal
arms of spherical virus coat proteins [18,19] and lac-
tate dehydrogenase (LDH) [20,21] which are involved
in the formation of oligomers. Indeed, there are some
similarities in sequence and predicted secondary
structure for the N-terminal regions of the oligomeric
proteins, spiny dogfish LDH [22], alfalfa mosaic virus
coat protein (AMV) [23,24] yeast alcohol dehydro-
genase (ADH) [25] and bovine superoxide dismutase
(SOD) [26] and the N-terminal 'arm' of βBp (table 2).
In all cases a predicted helical region extends from
the acetylated N-terminus to end in turn or coil con-
formation at or around the common glycine residue
(position 9 in βBp).

This may be a general structural feature enhancing
the stability of certain oligomeric proteins and we

suggest that the N- and C-terminal regions of βBp
form such structures. It is interesting that, although it
lacks the predicted terminal helix, the putative
C-terminal arm of βBp has some of these features
arranged in a sense antiparallel to the N-terminal arms
(table 2). The tertiary interactions of N-terminal
arms, where they are known, differ widely and we
have investigated various alternatives for the struc-
ture of the βBp dimer.

Using a combination of the model-building pro-
gram FRODO [17] and FITZ we have explored vari-
ous dimeric structures. As the monomer has pseudo
2-fold symmetry the intuitively most simple arrange-
ment for a dimer would have pseudo 222 symmetry.
A compact model of this type is shown in fig.3. In
this arrangement there are contacts between Phe 172
of the C-terminal arm and the hydrophobic surface
region particularly involving Phe 11 and Leu 36 of
the N-terminal domain of the other molecule. Trp 166
packs against the flat hydrophobic surface formed by
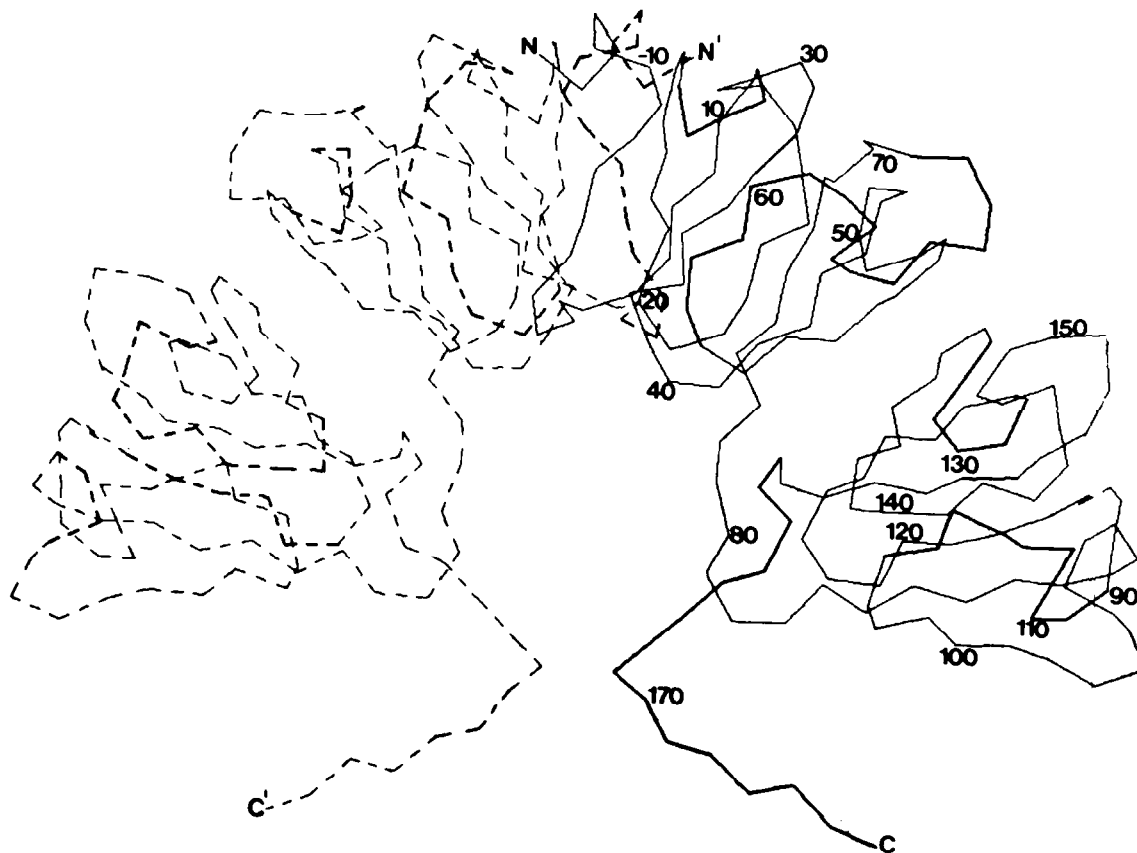
Fig.4. An extended dimer model of βBp viewed perpendicular to the 2-fold axis between monomers. Again, one molecule is
shown in dotted lines.

Gly 40 and Pro 1, Glu 58 and Arg 169 are close enough to form a salt bridge, and the non-bulky residues Gly 170 and Ala 171 wrap across between Glu 58 and Leu 36. In this arrangement, the N-terminal arm of one molecule interacts with the C-terminal domain of the other leading to contacts between Val 100 and Gly −7/Ala −8 and a possible salt bridge between Lys −6 and Asp 102. Finally, the connecting peptides of each molecule lie close to each other so that an antiparallel β-sheet interaction may be formed between residues 78−81 of both molecules.

An alternative, less compact model, involves an isologous interaction between two N-terminal domains (fig.4). This is analogous to the contact between the two globular domains within the monomer. A threonine at residue 4 of γII, equivalent to hydrophobic residues 43A and 125A in the intramolecular contact region, is replaced by isoleucine in βBp. With two N-terminal domains coming together, the two isoleucines would have equivalent roles to residues 43A and 125A. The twist between the two domains would then leave the hydrophobic patch at Phe 11/Leu 36 exposed, forming a possible interaction site for the N-terminal arm of the other molecule.

In the model presented here, Gly −7 and Ala −8 of the N-terminal arm warp across the opposing hydrophobic patch at Leu 36. There are also possibilities for a contact between Pro −5 and Gly 57 as well as an ionic bond between Lys −6 and Glu 58. This model allows for the formation of mixed trimers and higher aggregates through interactions between the free C-terminal arm and other types of β-crystallin subunits.

Further biochemical and spectroscopic analysis may be able to throw more light on the structure of the dimer. However, a detailed model must await a full X-ray study of βBp-crystallin which we hope will eventually become possible using the crystals we have obtained.

Acknowledgements

References

[1] Driessen, H. P. C., Herbrink, P., Bloemendal, H. and De Jong, W. W. (1981) Eur. J. Biochem. in press.

[2] Croft, L. R. (1972) Biochem. J. 128, 961−970.

[3] Slingsby, C. and Croft, L. R. (1978) Exp. Eye Res. 26, 291−304.

[4] Croft, L. R. (1980) Proc. Int. Soc. Eye Res. 1, 118.

[5] Bloemendal, H., Berns, T., Zweers, A., Hoenders, H. and Benedetti, E. L. (1972) Eur. J. Biochem. 24, 401−406.

[6] Van der Ouderaa, F. J., De Jong, W. W. and Bloemendal, H. (1973) Eur. J. Biochem. 39, 207−222.

[7] Van der Ouderaa, F. J., De Jong, W. W., Hilderink, A. and Bloemendal, H. (1974) Eur. J. Biochem. 49, 157−168.

[8] Herbrink, P., Van Westreenen, H. and Bloemendal, H. (1975) Exp. Eye Res. 20, 541−548.

[9] Bindels, J. G., Koppers, A. and Hoenders, H. J. (1981) submitted.

[10] Herbrink, P. and Bloemendal, H. (1974) Biochim. Biophys. Acta 336, 370−382.

[11] Zigler, J. S. and Sidbury, J. B. (1976) Comp. Biochem. Physiol. 55B, 19−24.

[12] Blundell, T. L., Lindley, P. F., Miller, L., Moss, D. S., Slingsby, C., Tickle, I. J., Turnell, W. G. and Wistow, G. J. (1981) Nature 289, 771−777.

[13] Chou, P. Y. and Fasman, G. D. (1974) Biochemistry 13, 222−245.

[14] Garnier, J., Osguthorpe, D. J. and Robson, B. (1978) J. Mol. Biol. 120, 97−120.

[15] Armand, G., Balazs, E. A. and Testa, M. (1970) Exp. Eye Res. 10, 143−150.

[16] Horwitz, J. (1976) Exp. Eye Res. 23, 471−481.

[17] Jones, T. A. (1978) J. Appl. Crystallog. 11, 268−272.

[18] Harrison, S. C., Olsen, A. J., Schutt, C. E., Winkler, F. K. and Bricogne, G. (1978) Nature 276, 368−373.

[19] Abad-Zapatero, C., Abdel-Meguid, S. S., Johnson, J. E., Leslie, A. G. W., Rayment, I., Rossman, M. G., Suck, D. and Tsukihara, T. (1980) Nature 286, 33−39.

[20] Adams, M. J., Ford, G. C., Koekoek, R., Lentz, P. J., McPherson, A., Rossman, M. G., Smiley, I. E., Schevitz, R. W. and Wonacott, A. J. (1970) Nature 227, 1098−1103.

[21] Holbrook, J. J., Liljas, A., Steindal, S. J. and Rossmann, M. G. (1975) in: The Enzymes (Boyer, P. D. ed) 3rd edn, vol. 2, pp. 191−192, Academic Press, New York.

[22] Taylor, S. S. (1977) J. Biol. Chem. 252, 1799−1806.

[23] Van Beynum, G. M. A., De Graaf, J. M., Castel, A., Kraal, B. and Bosch, L. (1977) Eur. J. Biochem. 72, 63−78.

[24] Collot, D., Peter, R., Das, B., Wolff, B. and Duranton, H. (1977) Biochim. Biophys. Acta 492, 267−283.

[25] Jörnvall, H. (1977) Eur. J. Biochem. 72, 425−442.

[26] Steinman, H. M., Naik, V. R., Abernathy, J. L. and Hill, R. L. (1974) J. Biol. Chem. 249, 7326−7338.