The EMBO Journal vol.4 no.10 pp.2597–2602, 1985

βs-Crystallin: structure and evolution of a distinct member of the βγ-superfamily

Yvonne Quax-Jeuken, Huub Driessen[1], Jack Leunissen, Wim Quax, Wilfried de Jong and Hans Bloemendal

Department of Biochemistry, University of Nijmegen, Geert Grooteplein Noord 21, 6525 EZ Nijmegen, The Netherlands, and [1]Department of Crystallography, Laboratory of Molecular Biology, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK

Communicated by H.Bloemendal

The nucleotide sequence of the cDNA of bovine lens βs-crystallin has been determined, and the derived amino acid sequence has been confirmed by amino acid compositions and partial sequences of the tryptic peptides of this monomeric protein. βs-Crystallin has a length of 177 residues, corresponding to a mol. wt. of 20 773, and a blocked N-terminal serine. Comparison of βs with the known sequences of other β- and γ-crystallins, and computer construction of a phylogenetic tree of these sequences, shows βs to be more closely related to the monomeric γ-crystallins than to the oligomeric β-crystallins. Also the tertiary structure of βs modelled by interactive computer graphics on the coordinates of γII-crystallin, revealed similarities with the γ-crystallins which might explain its monomeric behavior: the presence of a very short N-terminal 'arm' as compared with the β-crystallins; a distribution of charged residues on the surface as in the γ-crystallins; and finally the nature of certain residues of its inter-domain contacts. βs-Crystallin seems to be an old and isolated offshoot of the γ-family, and, considering its ancient origin, might well be present in other, non-mammalian, vertebrate classes.

Key words: crystallin/lens/molecular evolution/three-dimensional structure

Introduction

Much information has become available on the molecular evolution of genes and proteins from comparative studies of protein structure and gene organization. The crystallins are suitable objects for such a study (de Jong, 1981). These structural proteins from the vertebrate eye lens account for 80–90% of the soluble protein fraction (for reviews, see Bloemendal, 1982; Harding and Crabbe, 1984). There are three classes of crystallins in the mammalian lens: α, β and γ. The β- and γ-crystallins form a superfamily of related polypeptides, which most likely are derived from a common ancestral gene (Driessen *et al.*, 1981; Wistow *et al.*, 1981). These crystallins are built up of two similar globular domains (Blundell *et al.*, 1981), which show sequence homology, suggesting an intragenic duplication in their ancestral gene. The structure of each domain again is made up of two 'Greek key' motifs, reflecting an even earlier duplication event. This tertiary structure is directly related to the gene organization of the β- and γ-crystallins in that each of the four predicted structural motifs of the β-crystallin polypeptides is encoded by a separate exon (Inana *et al.*, 1983), while in the γ-crystallin genes an intron separates the two domain sequences (Schoenmakers *et al.*, 1984).

The main difference between the β- and γ-crystallin polypeptides is that β-chains are organized in dimers and trimers (β_L) or higher aggregates (β_H), while the γ-crystallins are monomeric. Structural analysis revealed that β-crystallins have an N-terminal extension, which may be involved in dimer formation or interaction with other polypeptides (Berbers *et al.*, 1983a, 1983b). The absence of such an extension at the N terminus may explain the monomeric nature of the γ-crystallins. In this context a remarkable member of the β-crystallin family is the βslow polypeptide, which, however, is monomeric. It was designated as a β-crystallin, because it is N-terminally blocked, like the other β-crystallins, and has an isoelectric point in the β-crystallin range (van Dam, 1966). Also its amino acid composition resembles that of β-crystallins more than one of the γ-crystallins. Furthermore its mol. wt., determined by SDS-gel electrophoresis, is higher than that of γ-crystallins and in the range of the β-crystallin polypeptides (Kabasawa *et al.*, 1977).

To determine, whether βs-crystallin belongs either to the β- or to the γ-family, or possibly represents a distinct branch in the evolution of β- and γ-crystallins, we performed sequence analysis on a βs clone isolated from a bovine lens complementary DNA bank (Quax-Jeuken *et al.*, 1984). Information about the tertiary structure of βs-crystallin was obtained by three-dimensional model building, using interactive computer graphics, based on the structure of bovine γII-crystallin. Furthermore an evolutionary tree for the βγ-crystallins was constructed to place βs in the branching arrangement of the other β- and γ-crystallins.

Results

Nucleotide sequence of pBLβs and primary structure of βs-crystallin

The nucleotide sequence of the cDNA clone is 592 bases long excluding the poly(A) and G.C tails introduced during the cloning procedure. This βs clone contains the coding sequence for 172 amino acids, and the total 3'-non-coding part of the mRNA (76 bases) including the poly(A) addition signal AATAAA (Proudfoot and Brownlee, 1976). Since our cDNA clone lacked the coding sequence of the N-terminal residues, we performed a 'primer extension' experiment with the synthetic primer (CTTGAAAGTTTTTGTC), which is homologous with the nucleotides 47–62 in Figure 1, and total bovine lens mRNA. After hybridization and extension the coding sequence for six N-terminal amino acids was determined, the start codon AUG and 10 bases of the 5'-non-coding region. The total length of the 5'-non-coding part of the mRNA could be estimated at ~50 bases. The nucleotide sequence of the sense strand (mRNA sequence) and its translated amino acid sequence (consisting of 177 residues, mol. wt. 20 773) are shown in Figure 1.

To support the correctness of the derived amino acid sequence and to prove its correspondence with the original βs-crystallin described by van Dam (1966), we isolated bovine lens βs for determination of amino acid composition and partial sequences. Electrophoretically pure βs was obtained by gel filtration of total

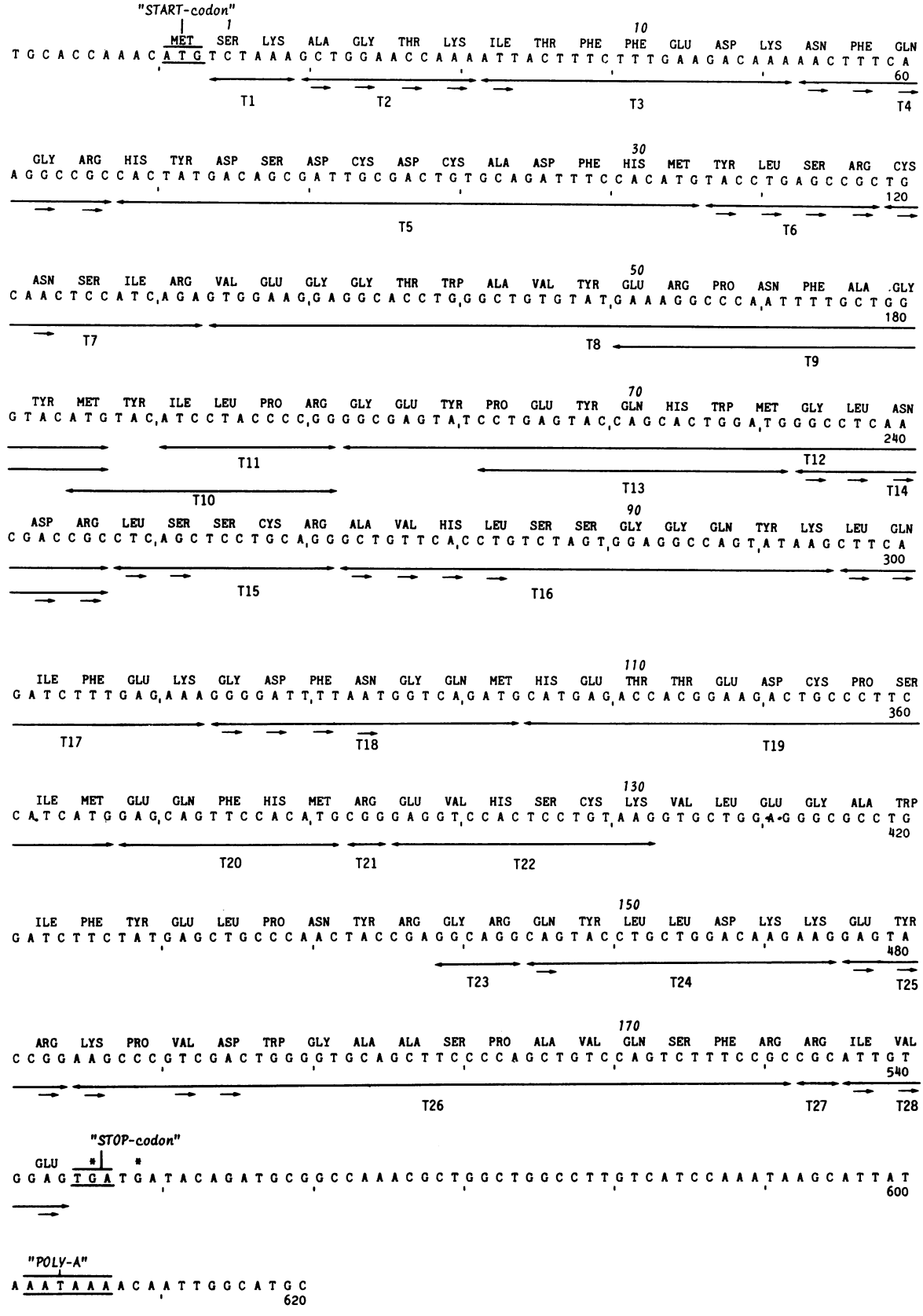

Fig. 1. Nucleotide sequence of βs mRNA. Sequence of the βs-crystallin mRNA (combined sequences of the cDNA in pBLβs and the RNA sequence derived by the 'primer extension' experiment) and the corresponding amino acid sequence are shown. G- and C-tails are not included. The tryptic peptides (T1 – T28) isolated from S-β-aminoethylated βs-crystallin are indicated. Some peptides were subjected to dansyl-Edman degradation, which confirmed the sequence of some residues (short arrows).

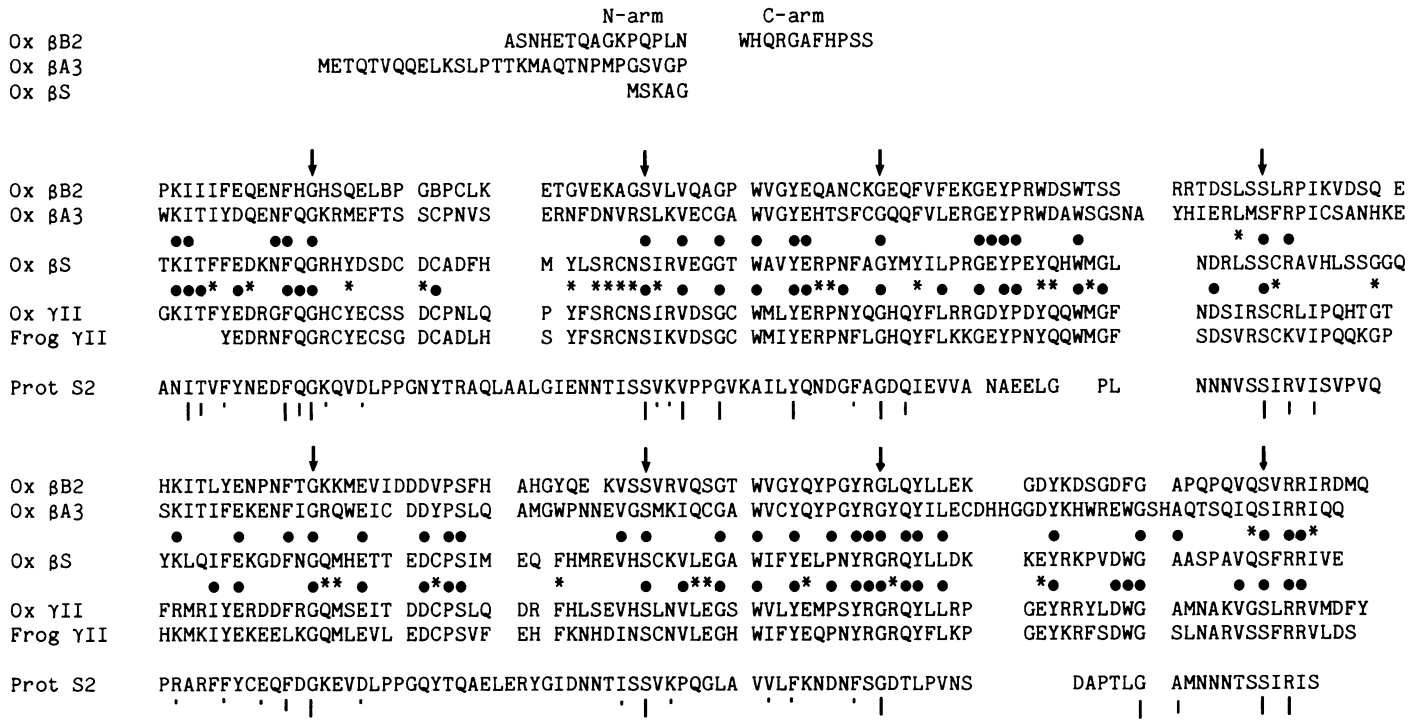

Fig. 2. Alignment of βs-crystallin with representative sequences of the βγ-family and with *M. xanthus* protein S. Bovine βB2- and βA3-crystallin, and bovine γII- and frog γII-crystallin are included as the most divergent representatives from the eight β- and γ-crystallin sequences used in the tree construction (Figure 3). The sequences corresponding to the two structural domains are aligned separately. The remaining N- and C-terminal extensions of the β-sequences are given at the top of the figure. Residues shared by βs and the two β-sequences are indicated by ● above the βs sequence; those in common with the two γ-sequences by ● below the βs-sequence. The symbol * above and below the βs-sequence indicates, where βs shares residues uniquely with βA3 and βB2, or with the γ's, respectively. Arrows indicate the glycine and serine residues, which are essential for the 'Greek key' motifs, and are absolutely conserved in all βγ-crystallins as well as in protein S. The symbols under the protein S sequence emphasize its sequence similarity with the βγ-crystallins: (|) residues in common with all five shown βγ-sequences; (') in common with four of the five βγ-sequences; (·) in common with, or chemically very similar (F-Y; D-E; K-R; I-V) to, at least three of the five βγ-sequences.

water-soluble lens protein over Sephadex G-75, followed by CM-52 ion exchange chromatography. After two-dimensional gel electrophoresis it appeared as a single spot of apparent mol. wt. 22 000 and pI 7.0. For unknown reasons this same material always appeared as a doublet on one-dimensional SDS-PAGE.

The amino acid composition of this βs (data not shown) fully agreed with the compositions described by van Dam (1966) and Croft (1973), and with the composition derived from the nucleotide sequence. A large number of peptides were purified from tryptic digests of cyanogen bromide fragments of βs, and partial sequence determinations were performed. All isolated peptides could be fitted into the derived amino acid sequence, as shown in Figure 1. This proves the authenticity of the βs-crystallin. As can be expected, the large and hydrophobic peptide from residues 131 – 145 could not be isolated by the peptide mapping procedure. The isolation of the blocked and uncharged peptide Ser-Lys from a tryptic digest of total βs agreed with the proposed N-terminal sequence of βs. The blocking group is most likely an N^{α}-acetyl group, as in the βBp chain and in the α-crystallins (Driessen *et al.*, 1981; Bloemendal, 1982). The N-terminal sequence Ser-Lys is indeed suitable to become acetylated (Driessen *et al.*, 1985). Our results thus contradict the N-terminal acetyl-Trp reported earlier to be present in βs (Croft, 1973).

Evolutionary relationship of βs-crystallin

The amino acid sequence of βs was compared with those of other known β- and γ-crystallins, as well as with the sequence of protein S of the bacterium *Myxococcus xanthus*, which recently has been shown to have considerable similarity with the βγ superfamily (Wistow *et al.*, 1985). The alignment of these sequences (Figure 2) clearly reveals the greater homology of βs with the γ-chains (27 uniquely shared residues) than with the β-chains (only three uniquely shared residues). The N-terminal extension, characteristic for the β-crystallins, is very short in βs, in which respect it also resembles the γ-crystallins more.

The percentage sequence homology of βs with the γ-crystallins (49 – 55%) is clearly higher than with the β-crystallins (34 – 37%) (Table I), but is in both instances considerably lower than amongst the γ- and β-crystallins themselves. This already reflects the isolated position of βs in relation to the β- and γ-crystallin families. A more detailed analysis of the evolutionary relationship of βs has been performed by constructing a most parsimonious phylogenetic tree of the βγ-crystallin superfamily. Inclusion of protein S as an outgroup in the tree construction procedure enabled the assignment of a root to the βγ-tree, which in turn enabled us to conclude that βs is indeed more related to the γ- than to the β-crystallins (Figure 3). It can also be concluded that the separation of the βs-gene from the γ-crystallin family occurred well before the divergence of the ancestors of frog and mammals. This means that βs-crystallin might, in principle, be present in all tetrapod vertebrates, and perhaps in fishes as well.

Predicted tertiary structure of βs-crystallin

Bovine βs shows the same 4-fold internally repeated structure of four motifs which is present in all βγ-crystallin polypeptides for which the sequence has been determined (e.g., Berbers *et*

Table I. Sequence similarities between βs and other βγ-crystallins

	Ox βB2	Ox βA3	Ox βS	Frog γII	Ox γII	Prot S2
Ox βB2	100.0	44.6	33.7	33.3	33.9	19.0
Ox βB1	56.7	50.0	36.0	36.2	36.1	20.7
Rat βB1	55.1	50.5	34.3	35.1	34.4	19.6
Ox βB3	55.9	47.8	37.1	35.6	38.3	21.7
Rat βB3	53.4	44.6	36.0	34.5	37.2	21.2
Ox βA3	44.6	100.0	34.6	31.7	35.1	20.3
Mouse β23	43.0	96.2	34.1	31.1	34.1	19.8
Frog β23	42.5	89.7	36.2	33.3	35.7	19.8
Ox βS	33.7	34.6	100.0	53.6	53.2	17.1
Frog γII	33.3	31.7	53.6	100.0	63.1	21.9
Frog γI	34.1	32.9	54.5	65.2	69.2	17.3
Rat γ1-1	33.3	32.4	49.1	59.5	81.6	20.2
Rat γ2-1	32.2	31.9	50.9	61.9	81.0	17.9
Human γ2-1	32.8	33.0	52.6	58.3	75.3	19.7
Human γ1-2	32.2	31.9	53.8	66.1	83.9	19.5
Rat γ1-2	33.9	33.5	54.3	62.5	91.4	19.5
Ox γII	33.9	35.1	53.2	63.1	100.0	22.4
Rat γ2-2	33.3	32.4	50.9	63.7	81.0	20.8
Rat γ3-1	34.4	33.5	50.9	61.3	79.9	19.7
Rat γ4-1	33.3	32.4	50.3	60.7	78.2	19.1
Prot S2	19.0	20.3	17.1	21.9	22.4	100.0

The percentages of sequence similarity between the optimally aligned sequences are given. Also the distantly related sequence of protein S of *M. xanthus* is included for comparison.

al., 1984; Schoenmakers *et al.*, 1984). This is reflected in the three-dimensional structure of bovine γII-crystallin, which consists of two similar globular domains, each comprising two similar 'Greek key' motifs (Blundell *et al.*, 1981; Wistow *et al.*, 1983; Summers *et al.*, 1984). The two domains pack together with a single connecting peptide, and are related by a pseudo-2-fold axis. In each domain the two 'Greek key' motifs form a pair of four-stranded anti-parallel β-pleated sheets, each of which is composed of three strands from one motif, and one from the other. In the sequence of βs those residues which are crucial for a γII-like structure are completely conserved: Gly 17, 56, 105 and 146 have torsion angles which cannot be formed with other residues, while Ser 38, 81, 128 and 171 provide stabilization of a 'Greek key' motif by hydrogen bonding to a main chain NH. In the core of βs the four tryptophans occupy the same central positions as in the γII structure. The sequence of βs could therefore easily be modelled by computer graphics methods into the four-motif structure of γII-crystallin.

Discussion

The structural analysis of bovine βs-crystallin completes the picture of the presently known members of the βγ superfamily. The primordial βγ gene must have originated in the ancestral line of the vertebrates, from some pre-existing gene. The protein S of *M. xanthus* may give us a vague inkling of the nature of such an ancestral gene (Wistow *et al.*, 1985). An even more distant similarity may exist between βγ-crystallins and the human c-*myc*

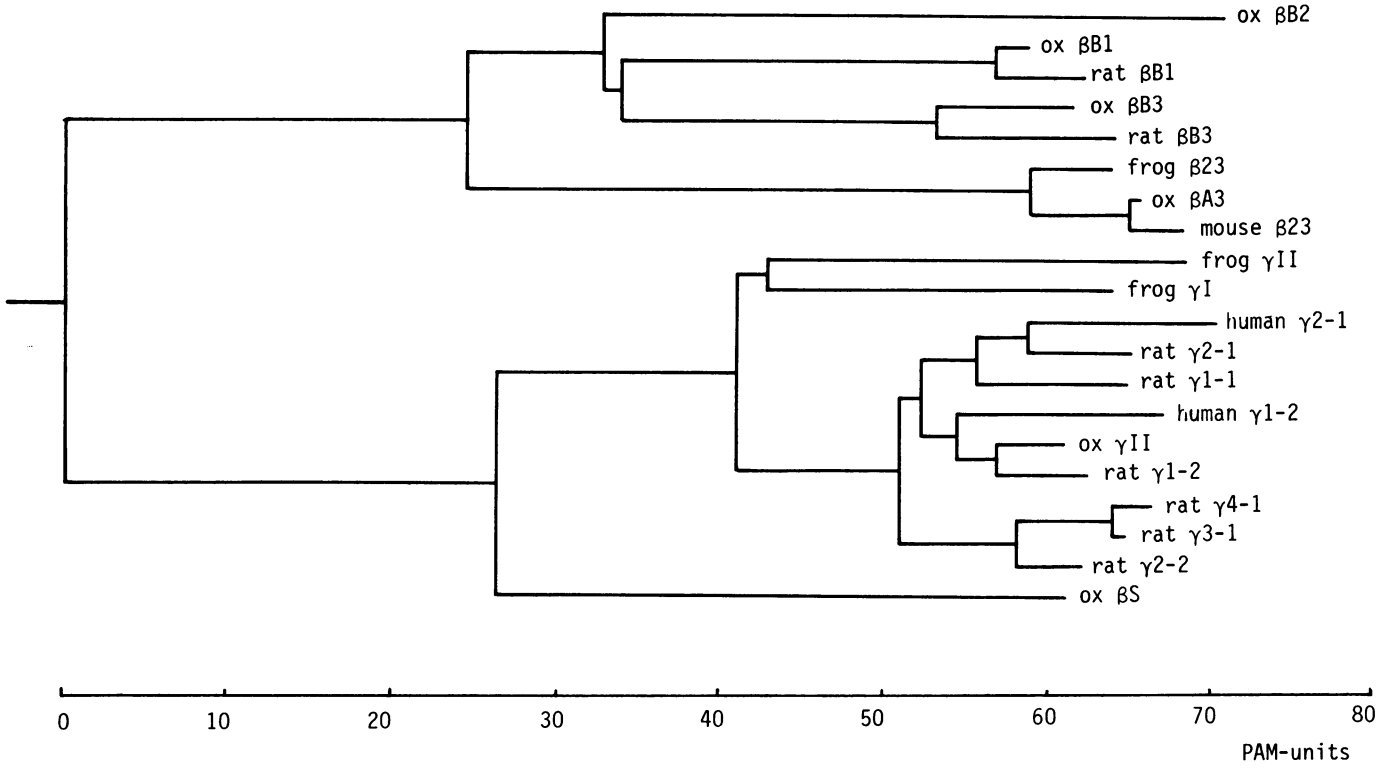

Fig. 3. Phylogenetic tree of the βγ-superfamily. This tree was constructed from the augmented difference matrix of the β- and γ-sequences as mentioned in Materials and methods, using protein S as outgroup, which provided the root of the tree. The scale is given in augmented percentage sequence divergence (PAM-units; Dayhoff *et al.*, 1978). The four recently published mouse γ-sequences (Breitman *et al.*, 1984), which closely resemble the rat sequences, have not been included, but would not have altered the branching arrangements. Also the use of the unaugmented difference matrix provided the same topology, although the branch lengths obviously become progressively shorter towards the root of the tree. When the motif order 2-1-4-3 for protein S was used (as proposed by Wistow *et al.*, 1985), instead of 1-2-3-4, the only difference in the topology was that the frog γI-sequence joined the branch to the mammalian γ-crystallins. Calculation of the total number of required amino acid replacements for the tree shown in this figure (285 replacements), showed that it is indeed only marginally less parsimonious, by a single replacement, to have a monophyletic rather than paraphyletic origin of the two frog γ-chains.

oncogene product (Crabbe, 1985).

Intragenic duplications, leading to the four-motif structure of the βγ-crystallins, followed by complete gene duplications, have given rise to the ancestors of the present-day families of β-, βs- and γ-crystallins. The γ-crystallin genes further duplicated, giving rise, in the rat, to six genes, of which five are closely linked (Schoenmakers *et al.*, 1984). Also the ancestral β-crystallin gene has repeatedly duplicated early in evolution, resulting in the presence of multiple β-crystallin subunits in all investigated vertebrates. At least seven different β-chains are present in the bovine lens (Berbers *et al.*, 1984), and orthologous gene products are found in mouse, man and chicken (Piatigorsky *et al.*, 1984). In the course of evolution the β-crystallins have acquired an N-terminal sequence, which probably extends from the compact two-domain structure of the protein, and possibly favors the interaction with other proteins. These N-terminal arms are, at least in rat βB1, encoded by a separate exon, and show little or no sequence similarity amongst the different orthologous β-chains (Berbers *et al.*, 1984; Piatigorsky *et al.*, 1984).

Only a single type of βs polypeptide, without any charge or size heterogeneity, seems to be present in the bovine lens, which suggests it to be the product of a single gene. Considering its ancient divergence from the γ-crystallin lineage, the βs gene might well be present in all vertebrate classes. Apart from bovine, βs has been reported in rat and human lenses (Croft, 1973; Kabasawa *et al.*, 1977; Bindels *et al.*, 1981; Zigler *et al.*, 1981; Harding and Crabbe, 1984), although without solid identification. Also the low mol. wt. protein from avian lenses has been suggested to be βs-like, rather than a γ-crystallin (McDevitt and Croft, 1977). Treton *et al.* (1984) could indeed not detect any γ-crystallin mRNA or γ-crystallin genes or pseudogenes in the chicken.

The present results establish a molecular mass of 20 773 for βs, which should settle the controversy about its size, variously reported as 22 000 and 28 000 (van Dam, 1966; Bindels *et al.*, 1981, 1982; Zigler *et al.*, 1981), depending on the method of analysis. Whether, apart from βs and γ-crystallin, yet another low mol. wt. monomeric protein, designated as γ_H-crystallin, is present in mammalian lenses is still a matter of debate (Harding and Crabbe, 1984).

In the course of evolution, βs and γ-crystallins not only have diverged in their structure, but also in their expression during development. While the synthesis of γ-crystallins sharply decreases after fetal life, the synthesis of βs-crystallin increases with age, both in bovine and in human lenses (van Dam, 1966; Croft, 1973; Slingsby and Croft, 1973).

From comparison of the sequences of βs and the βγ-crystallin polypeptides, it is apparent that the overall length of βs is more similar to that of the γ's. However, it does not have the very small C-terminal tail (Phe-Tyr) of the γ's. In this respect it resembles the acidic β-chains (βA3, β23). At the N terminus there is an extension, which is normally characteristic for the β-crystallin polypeptides (Berbers *et al.*, 1984). There are indications that these N-terminal arms play an important role in the aggregation behavior of the β-subunits (Wistow *et al.*, 1981; Berbers *et al.*, 1983b). It seems very likely that the N-terminal arm of βs with only four residues is too short, compared with a normal length of 12 – 58 residues, to have a significant effect on aggregation. One of the important conditions for β-type behavior would therefore not be fulfilled.

Moreover, when comparing the charged residues of βs with those of the β- and the γ-families, it is found that most of the conserved ones in the γ-family are also conserved in βs, while this polypeptide has not a single conserved charged residue which is specific for the β-family. The distribution of these residues on the surface of βs is fairly even and therefore very similar to those of the γ's. This is in contrast to the charge distribution in the β-chains, which show distinct hydrophilic patches on top of the N-terminal and at the bottom of the C-terminal domain. This may be important for their aggregation interactions (Wistow *et al.*, 1981; H.Driessen, D.Mahadevan, C.Slingsby and T.Blundell, unpublished data). This means that a second important condition for β-type behavior is not met.

The β-crystallin polypeptides appear to secure the compactness of their structure, as compared with γII, by the possible presence of two ion pairs in their interdomain region at both sides of the central pseudo-2-fold axis (H.Driessen, D.Mahadevan, C.Slingsby and T.Blundell, unpublished data). The residues which make up these ion pairs have been replaced by basic residues only in the case of the γ-crystallin polypeptides (Schoenmakers *et al.*, 1984). In this respect βs appears to take an intermediate position between the β- and γ-crystallins. Arg 83 most likely forms an ion pair with Asp 152, while Arg 173 cannot have an ionic interaction with Pro 62.

The features of the βs structure described above make it clear that although this polypeptide has some β-like structural characteristics the overall structure is more like that of the γ-crystallins. One therefore wonders if the name γF (γFast) might be a better description for this polypeptide.

Materials and methods

Sequence analysis of the bovine βs cDNA clone

The cloning of the bovine βs cDNA in the pBR322 plasmid has been described elsewhere (Quax-Jeuken *et al.*, 1984). The insert of the cDNA clone pBLβs (two *Pst*I fragments) and the DNA fragments resulting from digestion of the insert with *Hae*III, *Pvu*I or *Sma*I were subcloned in M13 Mp8 and Mp9 vectors and grown in JM103 rec(A^-) (Messing *et al.*, 1981). Single-stranded DNAs from recombinant bacteriophages were sequenced by the dideoxy method (Sanger *et al.*, 1980). Both DNA strands were sequenced completely. To determine the nucleotide sequence of the 5' part of the βs mRNA, dideoxy sequence determination with bovine mRNA (50 μg/ml) as template and a synthetic primer of 16 bases (kindly supplied by Professor Van Boom, Leiden) was carried out in 10 mM $MgCl_2$, 50 mM KCl, 50 mM Tris-HCl (pH 8.3), 1.5 mM dithiothreitol. The concentrations of deoxyribonucleoside triphosphates and dideoxyribonucleoside triphosphates were 25 μM and 5 μM, respectively. The reactions were carried out at 37°C for 20 min in a volume of 5 μl with avian myeloblastosis virus reverse transcriptase (2 units per reaction).

Protein-chemical analysis of βs-crystallin

The water-soluble protein (800 mg) from adult bovine lens cortices was fractionated over a Sephadex G-75 column (150 x 4 cm) in 0.05 M Tris-HCl buffer, pH 7.5, containing 0.05 M NaCl and 0.001 M Na_2-EDTA. The βs-crystallin was incompletely separated from γ-crystallin, and further purified by ion-exchange chromatography over carboxymethylcellulose (Whatman CM-52). Up to 100 mg of protein was applied onto a 20 × 1 cm bed of CM-52, and eluted at room temperature, at 19 ml/h, with a gradient ranging from 200 ml 0.1 M Na-acetate, pH 5.0, to 200 ml 0.3 M Na-acetate, pH 5.0. The purity and identity of pooled fractions were checked by one- and two-dimensional gel electrophoresis (Laemmli, 1970; O'Farrell, 1975).

Purified βs-crystallin was reduced and aminoethylated, and either subjected to tryptic digestion or to cyanogen bromide cleavage, using procedures described earlier (Driessen *et al.*, 1981). The CNBr fragments were separated by gel filtration over Sephadex G-50 sf in 5% acetic acid, and digested with trypsin (2% for 3 h at 37°C in 0.1 M NH_4HCO_3, pH 8.9). Peptides were isolated by high-voltage paper electrophoresis, pH 6.5, followed by descending chromatography. Neutral peptides were, if necessary, further purified by re-electrophoresis at pH 3.5. Amino acid compositions were determined on a Rank-Hilger Chromaspek analyzer, after hydrolysis for 22 h in 6 N HCl at 110°C. Partial sequences were determined by dansyl-Edman degradation (Driessen *et al.*, 1981).

Evolutionary tree construction of the βγ-crystallins

The amino acid sequence of βs was aligned for maximum homology with the sequences of bovine γII (Bhat and Spector, 1984), βB1, βB2, βB3 and βA3

(Berbers *et al.*, 1984; Quax-Jeuken *et al.*, 1984; Gorin and Horwitz, 1984), human γ1-2 and γ2-1 (Schoenmakers *et al.*, 1984; Summers *et al.*, 1984), rat βB1 and βB3 (den Dunnen *et al.*, 1985), the six rat γ-chains (Schoenmakers *et al.*,1984), mouse β23 (Inana *et al.*, 1983), frog β23 (Gause *et al.*, 1984), γI and γII (Tomarev *et al.*, 1982, 1984). Also the sequence of protein S of the bacterium *M. xanthus* (Inouye *et al.*, 1983), which recently has been reported to show structural and sequence similarity with the βγ-crystallins (Wistow *et al.*, 1985), has been included in the alignment, taking special care to align the residues which are of key importance for determining the 'Greek key' motifs.

From the aligned sequences a distance matrix was constructed, counting only the similarities within the aligned domain sequences (thus omitting N- and C-terminal extensions). Gaps within the domain sequences were counted as mismatches, unless the gap was present in both chains under consideration. Before tree construction the values in the distance matrix were augmented, using the PAM-scale of Dayhoff *et al.* (1978), to correct for superimposed replacements, and hence, to obtain more realistic values for the evolutionary distances between the different sequences. Because the PAM-scale is discontinuous, rounding off towards the next higher PAM-value was often necessary in the process of augmentation of the distance values. As a result the present branch lengths in the phylogenetic tree are slightly larger than, or equal to the real PAM-distances. Phylogenetic trees were constructed from the augmented distance matrix, using the computer program FITCH, version 2.5, as supplied by J.Felsenstein in his PHYLIP package. This program is based on the algorithm of Fitch and Margoliash (1967), with the improvement that negative branch lengths are not allowed (Prager and Wilson, 1978). Several different input orders were used but provided the same topology.

Molecular model building of βs-crystallin

The protein sequences of βs- and γII-crystallin were aligned in such a way as to obtain maximum homology. The structure of βs was modelled on γII using the program FRODO (Jones, 1978) for interactive computer color graphics on an Evans and Sutherland Picture System II (by T.A.Jones as modified by I.J.Tickle) (Wistow *et al.*, 1981; Inana *et al.*, 1983). The main chain conformation was built with the torsion angles of the γII structure, except in the regions adjacent to the insertion. The region of the insertion was constructed using the predictions of programs analyzing the sequence for secondary structure elements. Then for all residues the unchanged side chains were placed in the same position while non-identical side chains were put into similar positions, and adjusted interactively to minimize unfavorable contacts and to optimize van der Waals and ionic interactions keeping in mind the need to form a closely packed hydrophobic core. Finally the N-terminal extension was built in a random configuration since prediction techniques failed to induce a specific secondary structure.

Acknowledgements

We are grateful to Marlies Versteeg for performing the amino acid analyses and to Gerrit Groenewoud for technical assistance. We thank Professor Van Boom for the synthesis of the 'βs-primer', Miss L.Summers for kindly providing us with the 1.6 Å coordinates of γII and Dr. C.Slingsby and Professor T.L.Blundell for valuable discussions. This investigation has been carried out in part under the auspices of the Netherlands Foundation for Chemical Research (SON) and with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO). Huub Driessen is the recipient of an EMBO fellowship.

References

Berbers,G.A.M., Hoekman,W.A., Bloemendal,H., de Jong,W.W., Kleinschmidt,T. and Braunitzer,G. (1983a) *FEBS Lett.*, **161**, 225-229.

Berbers,G.A.M., Brans,A.M.M., Hoekman,W.A., Slingsby,C., Bloemendal,H. and de Jong,W.W. (1983b) *Biochim. Biophys. Acta*, **748**, 213-219.

Berbers,G.A.M., Hoekman,W.A., Bloemendal,H., de Jong,W.W. Kleinschmidt,T. and Braunitzer,G. (1984) *Eur. J. Biochem.*, **139**, 467-479.

Bhat,S.P. and Spector,A. (1984) *DNA*, **3**, 287-295.

Bindels,J.G., Koppers,A. and Hoenders,H.J. (1981) *Exp. Eye Res.*, **33**, 333-343.

Bindels,J.G., de Man,B.M. and Hoenders,H.J. (1982) *J. Chromatogr.*, **252**, 255-267.

Bloemendal,H. (1982) *CRC Crit. Rev. Biochem.*, **12**, 1-39.

Blundell,T., Lindley,P., Miller,L., Moss,D., Slingsby,C., Tickle,E., Turnell,B. and Wistow,G. (1981) Nature, **289**, 771-777.

Breitman,M.L., Lok,S., Wistow,G., Piatigorsky,J., Treton,J.A., Gold,R.J.M. and Tsui,L.-C. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7762-7766.

Crabbe,M.J.C. (1985) *FEBS Lett.*, **181**, 157-159.

Croft,L.R. (1973) *CIBA Found. Symp.*, **19**, 207-224.

Dayhoff,M.O., Schwartz,R.M., Orcutt,B.C. (1978) in Dayhoff,M.O. (ed.), *Atlas of Protein Sequence and Structure*, Vol. **5**, Suppl. 3, 343-352.

de Jong,W.W. (1981) in Bloemendal,H. (ed.), *Molecular and Cellular Biology of the Eye Lens*, John Wiley & Sons, pp. 221-278.

den Dunnen,J.T., Moormann,R.J.M. and Schoenmakers, J.G.G. (1985) *Biochim. Biophys. Acta*, **824**, 295-303.

Driessen,H.P.C., Herbrink,P., Bloemendal,H. and de Jong,W.W. (1981) *Eur. J. Biochem.*, **121**, 83-91.

Driessen,H.P.C., De Jong,W.W., Tesser,G.I. and Bloemendal,H. (1985) *CRC Crit. Rev. Biochem.*, **18**, in press.

Fitch,W.M. and Margoliash,E. (1967) *Science (Wash.)*, **155**, 279-284.

Gause,G.G.,Jr., Tomarev,S.I., Zinovieva,R.D. and Dolgilevich,S.M. (1984) in Belmonte,C. (ed.), *Proceedings of the International Society for Eye Research*, Vol. **3**, p. 137.

Gorin,M.B. and Horwitz,J. (1984) *Curr. Eye Res.*, **3**, 939-948.

Harding,J.J. and Crabbe,M.J.C. (1984) in Davson,H. (ed.), *The Eye*, 3rd edition, Academic Press, pp. 207-492.

Inana,G., Piatigorsky,J., Norman,B., Slingsby,C. and Blundell,T. (1983) *Nature*, **302**, 310-315.

Inouye,S., Franceschini,T. and Inouye,M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6829-6833.

Jones,T.A. (1978) *J. Appl. Crystallogr.*, **11**, 268-272.

Kabasawa,I., Tsunematsu,Y., Barber,G.W. and Kinoshita,J.H. (1977) *Exp. Eye Res.*, **24**, 437-448.

Laemmli,U.K. (1970) *Nature*, **227**, 680-685.

McDevitt,D.S. and Croft,L.R. (1977) *Exp. Eye Res.*, **25**, 473-481.

Messing,J., Crea,R. and Seeburg,P. (1981) *Nucleic Acids Res.*, **9**, 309-321.

O'Farrell,P.H. (1975) *J. Biol. Chem.*, **250**, 4007-4021.

Piatigorsky,J., Nickerson,J.M., King,C.R., Inana,G., Hejtmancik,J.F., Hawkins,J.W., Borras,T., Shinohara,T., Wistow,G. and Norman,B. (1984) *CIBA Found. Symp.*, **106**, 191-207.

Prager,E.M. and Wilson,A.C. (1978) *J. Mol. Evol.*, **11**, 129-142.

Proudfoot,N.J. and Brownlee,G.G. (1976) *Nature*, **263**, 211-214.

Quax-Jeuken,Y., Janssen,C., Quax,W., van den Heuvel,R. and Bloemendal,H. (1984) *J. Mol. Biol.*, **180**, 457-472.

Sanger,F., Coulson,A.R., Barrel,B.G., Smith,A.J.H. and Roe,B.A. (1980) *J. Mol. Biol.*, **143**, 161-178.

Schoenmakers,J.G.G., den Dunnen,J., Moormann,R.J.M., Jongbloed,R., van Leen,R.W. and Lubsen,N.H. (1984) *CIBA Found. Symp.*, **106**, 208-218.

Slingsby,C. and Croft,L.R. (1973) *Exp. Eye Res.*, **17**, 369-376.

Summers,L., Wistow,G., Narebor,M., Moss,D., Lindley,P., Slingsby,C., Blundell,T., Bartunik,H. and Bartels,K. (1984) in Hearn,M. (ed.), *Peptide and Protein Reviews*, Vol. **3**, Marcel Dekker, NY and Basel, pp. 147-168.

Tomarev,S.I., Krayev,A.S., Skryabin,K.G., Bayev,A.A. and Gause,G.G.,Jr. (1982) *FEBS Lett.*, **146**, 315-318.

Tomarev,S.I., Zinovieva,R.D., Chalovka,P., Krayev,A.S., Skryabin,K.G. and Gause,G.G.,Jr. (1984) *Gene*, **27**, 301-308.

Treton,J.A., Jones,R.E., King,C.R. and Piatigorsky,J. (1984) *Exp. Eye Res.*, **39**, 513-522.

van Dam,A.F.M. (1966) *Exp. Eye Res.*, **5**, 255-266.

Wistow,G., Slingsby,C., Blundell,T., Driessen,H., de Jong,W. and Bloemendal,H. (1981) *FEBS Lett.*, **133**, 9-16.

Wistow,G., Turnell,B., Summers,L., Slingsby,C., Moss,D., Miller,L., Lindley,P. and Blundell,T. (1983) *J. Mol. Biol.*, **170**, 175-202.

Wistow,G., Summers,L. and Blundell,T. (1985) *Nature*, **315**, 771-773.

Zigler,J.S., Horwitz,J. and Kinoshita,J.H. (1981) *Exp. Eye Res.*, **32**, 21-30.

Received on 29 May 1985; revised on 11 July 1985