Slow Rate of Evolution of α A Chain of α -Crystallin

THE ever-increasing amount of comparative sequence data confirms that, in general, each protein evolves at a specific, more or less constant rate^{1–3}. The rate of acceptance of mutations in a particular gene is apparently influenced by the biological role and the size of the corresponding protein². Knowledge of the rate and pattern of amino acid substitutions in more and functionally diverse proteins is, however, required to understand better the principles governing the evolution of proteins. It is moreover important to determine which residues in a given polypeptide chain are constant throughout evolution, because they can be considered to be of vital importance for the appropriate functioning of that protein.

Whereas most evolutionary studies have been performed on enzymes or other proteins with well known biological activities, we have started a comparative study of α -crystallin, a structural protein from the vertebrate eye lens which has no relationship to any other group of proteins for which sequence data are available. α -Crystallin offers some attractive properties for evolutionary studies: it is present in all vertebrates and can easily be isolated in considerable quantities.

In the bovine lens α -crystallin makes up 30% of the soluble protein. It is isolated as an aggregate with an average molecular weight of 800,000 (ref. 4) and is composed of two types of chains, each with a molecular weight of about 20,000, the acidic α A and basic α B chains, occurring in a ratio of 2:1. Both α A and α B chains are usually found in two electrophoretically distinguishable forms, designated α A₁ and α A₂, and α B₁ and α B₂ respectively, which are otherwise very similar⁵. α A₁ has the same amino acid composition as α A₂ and is thought to be derived from α A₂ by *in vivo* deamidation⁶. Interest in α -crystallin has recently increased, since it seemed possible to isolate and translate the mRNA coding for the α A₂ chain⁷. The complete amino acid sequence of the bovine α A₂ chain has been established⁸. This provided the basis for a comparative study of α A chains from different mammals. The soluble lens proteins from ten mammals (cow, pig, horse, dog, rabbit, rat, mouse, guinea pig, rhesus monkey and human)

Fig. 1 Acrylamide gel electrophoresis in Tris-glycine buffer, pH 8.5, containing 6 M urea²⁵. Samples are: bovine α -crystallin (a), and water-soluble lens proteins from cow (b), pig (c), guinea pig (d), mouse (e), rabbit (f), dog (g) and rhesus monkey (h). In rat, horse and human (not shown) the α A chains have the same mobility as in the other species.

were compared by acrylamide gel electrophoresis (Fig. 1). No difference in mobilities of the α A chains was observed among these species. This indicates that no substitutions involving

charge differences exist between these chains, apart from the rather unlikely possibility that in a particular chain a substitution from a neutral to an acidic residue occurs together with a substitution from a neutral to a basic residue (or *vice versa*). More detailed studies were performed on the α A chains from cow, pig, dog and rabbit. α -Crystallin was prepared from aqueous lens extracts by isoelectric precipitation, followed by gel filtration on Sephadex G-200⁹. The α A chain was isolated by column chromatography on SE-Sephadex at pH 3.2 in 7 M urea⁵. The α A chain was reduced and aminoethylated¹⁰—and digested with 1% (by weight) trypsin in 0.1 M NH_4HCO_3 , pH 8.9, for 4 h at 37° C. The pH was lowered to 6.5 with 1 M HCl and the insoluble peptides removed by centrifugation. The soluble peptides were subjected to peptide mapping (Fig. 2). No significant differences were observed among the peptide maps of cow, pig and rabbit aminoethylated α A chains. The peptide map of canine AE- α A differed from the others in that peptide T11 had a considerably higher R_F .

Peptides from fingerprints stained with 0.02% ninhydrin were eluted with 6 M HCl and hydrolysed for amino acid analysis. The results of these analyses, which will be published in full elsewhere, revealed only very few differences among the four species. Peptides which were identical in composition to the corresponding bovine ones were assumed to have the same sequence, which is reasonable considering the unlikelihood of reciprocal substitutions within a single short peptide. Differences in composition from the corresponding bovine peptides were observed in pig T3, T6, T18 and T20, in dog T3, T11 and T18, and in rabbit T1 and T18. These peptides were isolated by preparative paper electrophoresis and chromatography, using the systems described in the legend to Fig. 2. The unstained peptides were eluted from the paper with 10% acetic acid and sequenced according to the dansyl-Edman technique¹¹. Dansyl amino acids were identified by polyamide thin-layer chromatography¹². Assignment of amide groups was based on electrophoretic mobilities of peptides¹³.

The results are summarised in Table 1 and compared to the known bovine sequence⁸. All amino acid replacements can be accounted for by single base substitutions in the DNA. They are moreover of a chemically conservative character¹⁴, with the exception of the glutaminyl-leucyl interchange in dog T11.

Both dog and pig have alanine in position 13 where the other species have threonine. Since cow and pig are more closely related than pig and dog, at least two identical substitutions in different evolutionary lineages are needed to account for the occurrence of alanine or threonine in peptide T3.

Although this might seem to be a rare coincidence, it is known that in cytochrome *c* from 20% to 27% of all substitutions are in fact caused by parallel mutations¹⁵. The non-randomness of the evolutionary process is shown also by the presence of several substitutions in T1 and T18, whereas long stretches of the chain are unchanged. Comparative studies of other proteins have likewise demonstrated that although mutagenic events probably occur randomly over the entire gene, the distribution of accepted substitutions is non-random^{16–18}. Such departure from randomness reflects the constraints on the structure of a protein that are imposed by its function.

The analysed tryptic peptides of the four species account together for 137 residues of the total number of 173 residues in the α A chain. The remaining thirty-six residues are present in the insoluble 'core' peptides T4 and T9 which have not yet been analysed in pig, dog and rabbit. On the basis of our analyses we can make a first estimate of the rate of evolution of the mammalian α A chain. Among the investigated 137 residues we have observed the following number of substitutions (in parentheses the approximate time of divergence of the respective species^{19–21}: between cow and pig four (50 m.y.);

Table 1 Amino Acid Substitutions in α A Chain of Mammalian α -Crystallins

		T1	10	T2	T3	20	30	T4
Bovine	Ac-Met-Asp-Ile-Ala-Ile-Gln-His-Pro-Trp-Phe-Lys-Arg-Thr-Leu-Gly-Pro-Phe-Tyr-Pro-Ser-Arg-Leu-Phe-Asp-Gln-Phe-Phe-Gly-Glu-Gly-Leu-Phe-Glu-Tyr-							
Pig	← Ile-Ala → ↔ Ala → (not determined)							
Dog	← Ile-Ala → ↔ Ala →							
Rabbit	← Val-Thr → ↔ Thr →							
	40		50	T5	T6	60		T7-8
Bovine	-Asp-Leu-Leu-Pro-Phe-Leu-Ser-Ser-Thr-Ile-Ser-Pro-Tyr-Tyr-Arg-Gln-Ser-Leu-Phe-Arg-Thr-Val-Leu-Asp-Ser-Gly-Ile-Ser-Glu-Val-Arg-Ser-Asp-Arg-Asp-							
Pig	(not determined) ← Val →							
Dog	← Ile →							
Rabbit	← Ile →							
	70	T9	80	T10	90	T11	100	T12
Bovine	-Lys-Phe-Val-Ile-Phe-Leu-Asp-Val-Lys-His-Phe-Ser-Pro-Glu-Asp-Leu-Thr-Val-Lys-Val-Gln-Glu-Asp-Phe-Val-Glu-Ile-His-Gly-Lys-His-Asn-Glu-Arg-Gln-							
Pig	→ (not determined) ← Gln →							
Dog	→ ← Leu →							
Rabbit	→ ← Gln →							
	T13	110	T14	T15	T16	120	T17a	130
Bovine	-Asp-Asp-His-Gly-Tyr-Ile-Ser-Arg-Glu-Phe-His-Arg-Arg-Tyr-Arg-Leu-Pro-Ser-Asn-Val-Asp-Gln-Ser-Ala-Leu-Ser-Cys-Ser-Leu-Ser-Ala-Asp-Gly-Met-Leu-							
Pig	→ ← → ← → ← → ←							
Dog	→ ← → ← → ← → ←							
Rabbit	→ ← → ← → ← → ←							
	140	150	T18	160	T19	T20	170	
Bovine	-Thr-Phe-Ser-Gly-Pro-Lys-Ile-Pro-Ser-Gly-Val-Asp-Ala-Gly-His-Ser-Glu-Arg-Ala-Ile-Pro-Val-Ser-Arg-Glu-Glu-Lys-Pro-Ser-Ser-Ala-Pro-Ser-Ser-COOH							
Pig	→ Val-Pro → Val → ← Thr →							
Dog	→ Val-Pro → Val → ← Ser →							
Rabbit	→ Val-Gln → Leu → ← Ser →							

The sequences which are identical in all four species are indicated by drawn lines. Except where the amino-acid compositions of peptides differed from the corresponding bovine ones, the sequences for the other species were deduced from the bovine sequence. Peptides which differed in composition were

sequentially degraded by the dansyl-Edman method (→). Peptides T4 and T9, which are insoluble at pH 6.5 and therefore not present on the peptide map, have not yet been investigated. Sites of cleavage with trypsin are indicated (↓). Because rabbit T1, like bovine²⁷, is blocked at the N-terminus, the

relevant sequence was obtained by four degradation steps on the N-terminal cyanogen bromide fragment of the α A chain. Porcine T20 was split with thermolysin into two fragments, as indicated. Both thermolytic peptides were sequenced.

Fig. 2 Peptide map of the soluble tryptic peptides from 2 mg bovine aminoethylated αA_2 chain. Electrophoresis was carried out in pyridine-acetic acid-water (25 : 1 : 225, by volume; pH 6.5) for 2 h at 50 V cm^{-1} , and followed by descending chromatography in the second dimension, using *n*-butanol-pyridine-acetic acid-water (15 : 10 : 3 : 12, by volume)²⁶. Staining was performed with 0.02% ninhydrin in acetone. Peptides are numbered in the order in which they occur in the αA chain. The methionine-containing peptides T1 and T17b are partially oxidised. Peptide T18 hardly stained with ninhydrin, due to its N-terminal sequence Ile-Pro. The arrow indicates the position of T11 in aog. Peptides were eluted from the paper with 0.5 ml of 6 M HCl for amino acid analysis. Hydrolysis under reduced pressure was performed for 20 h at 110° C. Cysteine was detected as aminoethylcysteine and tryptophan by staining the peptide map with Ehrlich's reagent.

cow-dog, three (70 m.y.); cow-rabbit, five (80 m.y.); pig-dog, three (70 m.y.); pig-rabbit, seven (80 m.y.); dog-rabbit, six (80 m.y.). Provided that these 137 residues are representative of the whole αA chain of 173 residues, we can calculate from these data that in the αA chain the average time needed for fixation of one amino acid substitution per 100 residues is between 40 and 50 m.y. This corresponds to about two accepted point mutations (PAMs) per 100 m.y.¹⁴ and characterises the αA chain as a slowly evolving protein (see Table 6-1 in ref. 3).

The slow rate of evolution of the αA chain compared with other proteins is also evident from Table 2. Although the slow evolutionary rate of the αA chain makes it less suitable for phylogenetic studies among the mammalian orders, it will probably be a convenient protein to study the lines of evolution between more distantly related taxa within the vertebrate subphylum.

Table 2 Sequence Comparison of Homologous Proteins and Rates of Evolution^{3,18,22}

Protein	% Difference from cow			PAMs ¹⁴ per 100 m.y.
	Pig	Dog	Rabbit	
Histone IV	0			0.06
α -Crystallin, αA chain	3		4	2
Cytochrome c	0	3	4	3
Insulin A and B chain	4	4	6	4
Parathyroid hormone				
Haemoglobin α chain	12	20	18	11
Haemoglobin β chain	16	19	21	13
Cytochrome b5	16		19	
Pancreatic RNase	22			33
Proinsulin C-peptide	42	58		
Fibrinopeptides A and B	51	68	70	90

It seems that the rate of evolution of proteins slows down with increasing molecular weight, other factors being equal, because interior amino acid side chains are subjected to more stringent stereochemical requirements than exterior side chains². Although the αA chain itself has a rather low molecular weight (19,830), its slow rate of evolution apparently correlates better with the high molecular weight of the α -crystallin aggregate (average 800,000).

The slow rate of evolution of the αA chain agrees with the high degree of immunological cross reactivity between α -crystallins from different mammalian species^{23,24}, although in fact no data are available on isolated αA chains.

If indeed the rate of evolution of proteins is limited by selective constraints on them, we must conclude that the αA chain is subjected to strong selective forces, hardly tolerating any change in the primary structure. Since α -crystallin is probably devoid of substitution-sensitive centres, like the substrate- or coenzyme-binding sites in enzymes, the selection must work solely at the level of the interactions within and between the polypeptide chains composing the molecule. Apparently most amino acid substitutions in the αA chain would impair the integrity of the α -crystallin aggregate to such an extent that the appropriate functioning of the lens becomes seriously affected.

We thank M. Versteeg and G. Groenewoud for performing the amino acid analyses. Kiam Liem-The kindly provided rabbit α -crystallin. Lenses were obtained through the cooperation of the Central Animal Laboratory of the University of Nijmegen. This work was supported in part by the Netherlands Foundation for Chemical Research and by financial aid from the Netherlands Organization for the Advancement of Pure Research.

W. W. DE JONG
J. M. VAN AMELSVOORT
F. J. VAN DER OUDERAA
H. BLOEMENDAL

Department of Biochemistry,
University of Nijmegen,
Nijmegen

Received August 20; revised October 9, 1973.

- ¹ Ohta, T., and Kimura, M., *J. molec. Evol.*, **1**, 18 (1971).
- ² Dickerson, R. E., *J. molec. Evol.*, **1**, 26 (1971).
- ³ McLaughlin, P. J., and Dayhoff, M. O., in *Atlas of Protein Sequence and Structure* (edit. by Dayhoff, M. O.), **5**, 47 (National Biomedical Research Foundation, Silver Spring, Maryland, 1972).
- ⁴ Bloemendal, H., Berns, A., Zweers, A., Hoenders, H., and Benedetti, E. L., *Eur. J. Biochem.*, **24**, 401 (1972).
- ⁵ Schoenmakers, J. G. G., Gerding, J. J. T., and Bloemendal, H., *Eur. J. Biochem.*, **11**, 472 (1969).
- ⁶ Bloemendal, H., Berns, A. J. M., van der Ouderaa, F., and De Jong, W. W., *Expl Eye Res.*, **14**, 80 (1972).
- ⁷ Berns, A. J. M., Schreurs, V. V. A. M., van Kraaikamp, M. W. G., and Bloemendal, H., *Eur. J. Biochem.*, **33**, 551 (1973).
- ⁸ Van der Ouderaa, F. J., De Jong, W. W., and Bloemendal, H., *Eur. J. Biochem.*, **39**, 207 (1973).
- ⁹ Schoenmakers, J. G. G., Hoenders, H. J., and Bloemendal, H., *Expl Eye Res.*, **7**, 172 (1968).
- ¹⁰ Raftery, M. A., and Cole, R. D., *J. biol. Chem.*, **241**, 3457 (1966).
- ¹¹ Gray, W. R., and Smith, J. F., *Analyt. Biochem.*, **33**, 36 (1970).
- ¹² Woods, K. R., and Wang, K. T., *Biochim. biophys. Acta*, **133**, 369 (1967).
- ¹³ Offord, R. E., *Nature*, **211**, 591 (1966).
- ¹⁴ Dayhoff, M. O., Eck, R. V., and Park, C. M., in *Atlas of Protein Sequence and Structure* (edit. by Dayhoff, M. O.), **5**, 88 (National Biomedical Research Foundation, Silver Spring, Maryland, 1972).
- ¹⁵ Boulter, D., Ramshaw, J. A. M., Thompson, E. W., Richardson, M., and Brown, R. H., *Proc. R. Soc. Lond.*, **B181**, 441 (1972).
- ¹⁶ Jukes, T. H., and Holmquist, R., *J. molec. Biol.*, **64**, 163 (1972).

- ¹⁷ Boyer, S. H., Crosby, E. F., Noyes, A. N., Fuller, G. F., Leslie, S. E., Donaldson, L. J., Vrablik, G. R., Schaefer, E. W., and Thurmon, T. F., *Biochem. Genet.*, **5**, 405 (1971).
- ¹⁸ Barnard, E. A., Cohen, M. S., Gold, M. H., and Kim, J. K., *Nature*, **240**, 395 (1972).
- ¹⁹ Young, J. Z., *The Life of Vertebrates*, second ed. (Oxford University Press, Oxford, 1962).
- ²⁰ Romer, A. S., *Vertebrate Paleontology*, third ed. (University of Chicago Press, Chicago, 1966).
- ²¹ McKenna, M. G., *Ann. N.Y. Acad. Sci.*, **167**, 217 (1969).
- ²² Peterson, J. D., and Steiner, D. F., *J. biol. Chem.*, **247**, 4866 (1972).
- ²³ Björk, I., *Expl Eye Res.*, **7**, 129 (1968).
- ²⁴ Mehta, P. D., and Lerman, S., *Comp. Biochem. Physiol.*, **38A**, 637 (1971).
- ²⁵ Bloemendal, H., in *Electrophoresis. Theory, Methods and Applications* (edit. by Bier, M.), **2**, 379 (Academic Press, New York and London, 1967).
- ²⁶ Waley, S. G., and Watson, J., *Biochem. J.*, **55**, 328 (1953).
- ²⁷ Hoenders, H. J., and Bloemendal, H., *Biochim. biophys. Acta*, **147**, 183 (1967).