# Partial sequence homology of human *myc* oncogene protein to beta and gamma crystallins

M. James C. Crabbe

*Nuffield Laboratory of Ophthalmology, Oxford University, Walton Street, Oxford OX2 6AW, England*

Received 18 December 1984

The human cellular *myc* gene is one of about 20 cellular oncogenes which code for a variety of proteins including protein kinases and growth factors [1]. The human gene is related to the avian myelocytomatosis leukaemia virus MC29 [2] and produces a binding protein which may be involved in regulation of gene expression [3] and cellular differentiation and proliferation [4]. The crystallins are proteins in the eye lens synthesised at different stages of cell differentiation and proliferation, and whose short range order is necessary for lens transparency [5,6]. Computer-based sequence comparisons show that beta Bp and gamma II crystallins, which show partial sequence homology and conservation of 'Greek Key' motives [7,8] are also partially homologous to two regions on the human *myc* protein, though this protein probably does not conserve the 'Greek Key' structural motives.

*Oncogene* c-myc *protein* *Crystallin* *Lens*

## 1. ANALYSIS AND DISCUSSION

The sequence of the human *c-myc* gene protein [2] was compared to those of the alpha, beta and gamma crystallins, using a micro-computer-based sequence comparison program ZPEP. After searching for the best linear homologies, modifications using insertions and deletions were made, similar to those used to fit beta and gamma crystallins [7], and incorporating subsequent sequence revisions. The *c-myc* protein, which showed close homology to the avian myeloblastosis cell oncogene protein, gave two regions of significant homology, with the bovine beta Bp crystallin [9] and bovine gamma II crystallin [8], as shown in fig.1. These homologies are unlikely to have arisen by chance. There was no significant homology with alpha crystallin, with the Drosophila heat shock proteins HS 22, HS 23, HS 26 and HS 27, or with bovine lens leucine aminopeptidase. The regions of maximum homology shown in fig.1 correspond to the four 'Greek Key' motives of the gamma and beta crystallins [8,10]. There are two main regions of homology on the *c-myc* protein, corresponding to residues 12–197 and 152–337. While there is some overlap, it is interesting that the *c-myc* protein is synthesised from two exons of roughly equal length, which are separated by an intron. The homology is 27% for beta and gamma crystallins, 21% for gamma crystallin and the first section of the *c-myc* protein, 19% for beta crystallin and the first section of the *c-myc* protein, 17% for beta crystallin and the second section of the *c-myc* protein, 17% for the first and second sections of the *c-myc* protein, and 13% for gamma crystallin and the second section of the *c-myc* protein. In addition, there are several very close amino acid substitutions, marked by a single line on the figure, and several others, not marked, where substitutions result in amino acids of similar charge or hydrophobicity. It therefore seems that the beta and gamma crystallins, until now thought to be unique proteins from the eye lens, are in fact related to a protein from a very different source. What is remarkable about the beta and gamma crystallins, is that while the sequence homology between them is partial, the structural homology is very high [10], consequent upon conservation of

*Published by Elsevier Science Publishers B.V. (Biomedical Division)*
00145793/85/$3.30 

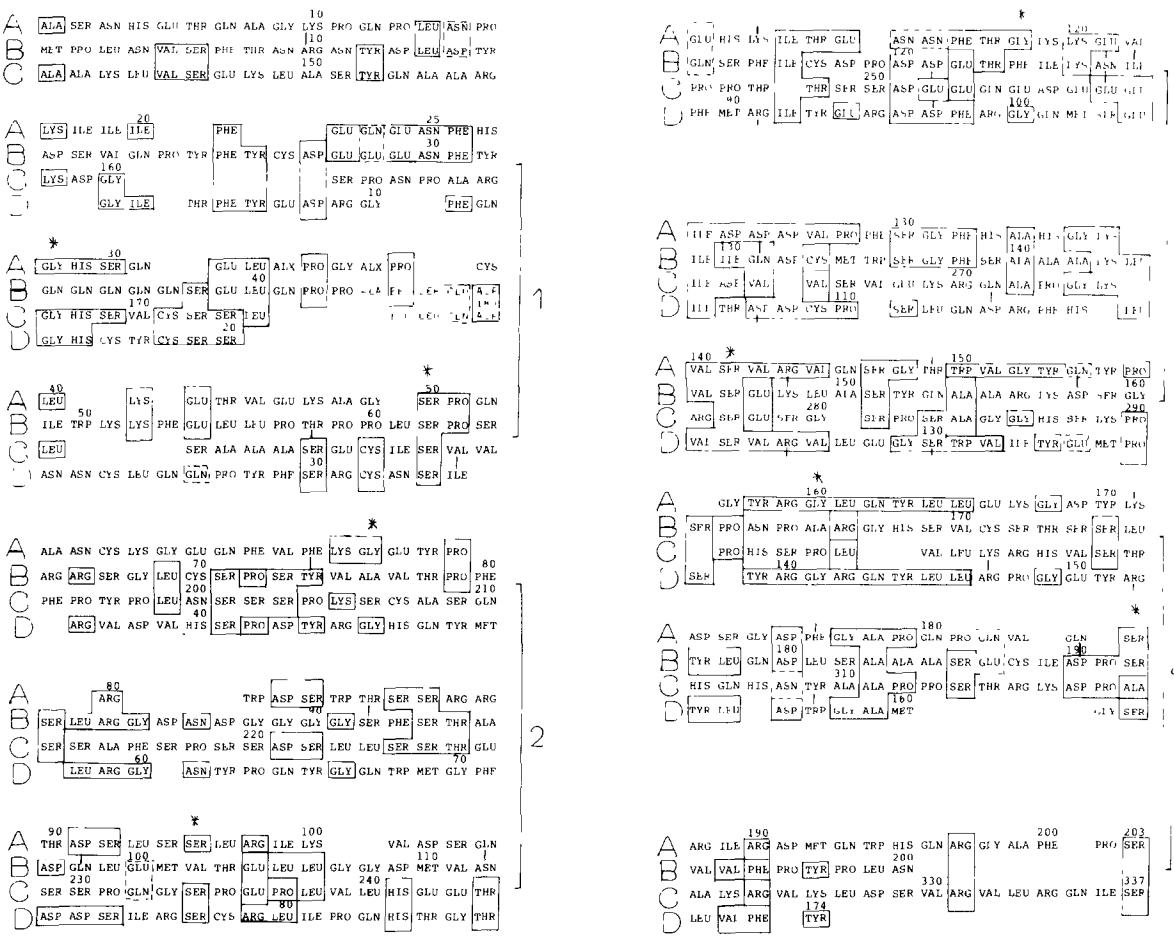

Fig.1. Optimal alignment of (A) bovine beta Bp crystallin; (B) N-terminal section (residues 1–200) of human *c-myc* gene protein; (C) part of the C-terminal section (residues 141–337) of human *c-myc* gene protein; (D) bovine gamma II crystallin. Residues 338–437 of the *c-myc* protein, which show no significant homology are omitted. Chains show numbering of the three-letter amino acid code as in [8,9]. Occasionally, amino acids in the crystallin proteins deleted in the sequence matching are omitted for clarity of presentation, so that the *c-myc* sequence is continuous. Identical interchain residues are shown in solid boxes. Dotted boxes indicate like residues where amidation has occurred: aspartate and asparagine, glutamate and glutamine. Solid vertical lines represent conservative amino acid substitutions. Asterisks represent highly conserved glycine and serine residues for 'Greek Key' motive formation. Square brackets numbered 1–4 indicate the four 'Greek Key' crystallin motives. Sequence comparisons were performed on a BBC/Torch microcomputer using the ZPEP protein sequence comparison program obtained from Dr M. Bishop, Department of Zoology, Cambridge University. Protein sequences were on the NEWAT [17] Doolittle protein sequence database, and used with kind permission.

key glycine, serine and aromatic residues throughout the four motives [5,11] which have been marked on the figure by an asterisk. With the *c-myc* protein, we have two regions of the sequence which are almost of the level of sequence homology with beta and gamma crystallins as the two crystallins themselves, and yet the key residues are poorly conserved. Thus, here we have a case of partial sequence homology with what appears to be little structural homology, and it seems unlikely that the *c-myc* protein is composed of four domains of 8 'Greek Keys'. However, it is possible that the gene coding separately for each 'Greek Key' motive in the beta-gamma gene family [12] was derived from the same ancestral gene as the *c-myc* oncogene. This genetic relationship may

underlie a more subtle functional relationship. We know that cell differentiation in the lens is linked to the appearance of different lens crystallins. In the bovine lens, initial synthesis of gamma IV is gradually replaced by gamma III and gamma II, and finally, as gamma synthesis stops, by beta S [13]. The beta crystallins increase as development proceeds, while two embryonic alpha crystallin subunits are lost [14]. In the rat lens beta and gamma crystallins, detectable by immunofluorescence in differentiated lens fibre cells, are not found in epithelial cells, while alpha crystallins, which are partially homologous to small heat-shock proteins [15], are found in both epithelial and fibre cells [16]. It is interesting that the lens has no apparent malignant potential in vivo, so there is no 'cancer of the lens'. Perhaps the alterations in amino acid sequence and consequent structural modifications from the *c-myc* protein to the crystallins have resulted in a functional change from a protein involved in cell proliferation and transformation to a protein involved in cell differentiation and a highly ordered protein structure within the cell. If this is the case, then the current studies on the structure, genetics and chemical modifications of lens proteins could have hitherto unsuspected rewards for unfolding the mechanisms of carcinogenesis.

## ACKNOWLEDGEMENTS

I thank the MRC for support, Professor R.F. Doolittle for kind permission to use the Doolittle protein sequence database, Dr J.J. Harding for helpful conversations, and Mr A.J. Bron for kind assistance and provision of facilities.

## REFERENCES

[1] Bishop, J.M. (1983) Ann. Rev. Biochem. 52, 301–354.
[2] Colby, W.W., Chen, E.Y., Smith, D.H. and Levinson, A.D. (1983) Nature 301, 722–725.
[3] Kingston, R.E., Baldwin, A.S. and Sharp, P.A. (1984) Nature 312, 280–282.
[4] Pfeifer-Ohlsson, S., Goustin, A.S., Rydnert, J., Wahlstrom, T., Bjersing, L., Stehelin, D. and Ohlsson, R. (1984) Cell 38, 585–596.
[5] Harding, J.J. and Crabbe, M.J.C. (1984) in: The Eye (Davson, H. ed.) 3rd edn, pp.207–492, Academic Press, New York.
[6] Delaye, M. and Tardieu, A. (1983) Nature 302, 415–417.
[7] Driessen, H.C.P., Herbrink, P., Bloemendal, H. and De Jong, W.W. (1980) Exp. Eye Res. 31, 243–246.
[8] Blundell, T., Lindley, P., Miller, L., Moss, D., Slingsby, C., Tickle, I., Turnell, B. and Wistow, G. (1981) Nature 289, 771–777.
[9] Driessen, H.C.P., Herbrink, P., Bloemendal, H. and De Jong, W.W. (1981) Eur. J. Biochem. 121, 83–91.
[10] Slingsby, C., Miller, L.R. and Berbers, G.A.M. (1982) J. Mol. Biol. 157, 191–194.
[11] Summers, L., Slingsby, C., White, H., Narebor, M., Moss, D., Miller, L., Mahadevan, D., Lindley, P., Driessen, H., Blundell, T., Den Dunnen, J., Moormann, R., Van Leen, R. and Schoenmakers, J. (1984) in: Human Cataract Formation (Nugent, J. and Whelan, J. eds) pp.219–230, Ciba Foundation Symposium 106, Books.
[12] Piatigorsky, J. (1984) Cell 38, 620–621.
[13] Slingsby, C. and Croft, L.R. (1973) Exp. Eye Res. 17, 369–376.
[14] Van Kamp, G.J., Van Kleef, F.S.M. and Hoenders, H.J. (1961) Biochim. Biophys. Acta 48, 563–570.
[15] Ingolia, T.D. and Craig, E.A. (1982) Proc. Natl. Acad. Sci. USA 79, 2360–2364.
[16] McAvoy, J.W. (1978) J. Embryol. Exp. Morph. 45, 149–165.
[17] Doolittle, R.F. (1981) Science 214, 149–159.