A computer graphics model of frog γ-crystallin based on the three-dimensional structure of calf γ-II crystallin

L.J. Summers, T.L. Blundell, G.G. Gause+ and S.I. Tomarev+

Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, England and +*Institute of Developmental Biology, USSR Academy of Sciences, Vavilov Str. 26, Moscow, USSR*

Received 11 August 1986

Molecular models for *Rana* γ-1 and γ-2 crystallins have been constructed using computer graphics on the basis of the protein primary structure derived from the complementary DNA sequence and the three-dimensional structure of calf γ-II crystallin that has been defined at high resolution by X-ray analysis. The models show that the cores of the two domains are conserved as hydrophobic, with the polypeptide chain arranged as a four Greek-key motif structure. Although many lysines replace arginines at equivalent positions in mammalian proteins, the *Rana* crystallins also have an extensive series of ion pairs on their surface; these are strongly implicated in their function as stable structural molecules, which are highly conserved in the evolution of the vertebrate eye lens.

γ-Crystallin *Computer graphics model*

1. INTRODUCTION

The γ-crystallins form a class of low molecular mass, lens-specific proteins, which contribute towards the maintenance of the smooth gradient of refractive index essential for lens transparency. They are present in the lens as monomers and function probably both as spacefillers between the larger, oligomeric α- and β-crystallins, and as proteins which control the level of hydration of the lens. The hydration varies between species being relatively high in the soft accommodating human lens and low in the harder lens of the myopic rat; the hydration also varies between the core and the cortex.

The γ-crystallins of the bovine lens were first studied by classical protein sequencing techniques [1,2] but the sequence of γ-II crystallin has been modified by crystallographic studies [3,4] and by sequencing of γ-II crystallin cDNA [5]. The complete sequences of six rat γ-crystallins have been defined from their cDNA and genomic sequences by Schoenmakers et al. [6]; these crystallins are called γ-1-1, γ-1-2, γ-2-1, γ-2-2, γ-3-1 and γ-4-1 in which rat γ-1-2 appears to be equivalent to bovine γ-II. The sequences of murine γ-crystallins have been published [7]. The sequences of two human crystallins, γ-1-2 and γ-2-1 have also been reported [6]. All γ-crystallins are highly homologous (see table 1) but may have different functional roles: bovine γ-IV (and possibly the corresponding human and rat γ-2-2, γ-3-1 and γ-4-1 crystallins) is found in the nucleus of the lens and bovine γ-II (rat and human γ-1-2) in the cortex.

The three-dimensional structure of bovine γ-II crystallin has been shown by X-ray analysis to consist of two globular domains each comprising two similar Greek-key motifs of approx. 40 amino acids in antiparallel β-pleated sheet structure [3,4,8]. A similar structure is found in bovine γ-III [9] and γ-IV [10] crystallins.

Little is known at the molecular level of the structure of γ-crystallins in lower vertebrates. Recently, however, partial sequences (one of them almost complete) of two γ-crystallins, from the frog, *Rana temporaria*, have been defined from their cloned cDNAs [11,12]. Structural studies of more cDNA clones [13] demonstrate that in *R.*

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

temporaria the family of γ-crystallins consists of at least six nonidentical members. The sequences of these γ-crystallins are clearly homologous with other γ-crystallins, although the degree of homology varies from 50 to 80% [13]. Moreover some of these γ-crystallins have profound differences in their amino acid composition. Thus in *Rana* γ-2 crystallin the Lys/Arg ratio is equal to 1:1; in mammalian γ-crystallins this ratio is around 0:1. In this paper we have used interactive computer graphics to investigate the three-dimensional structure of the *Rana* γ-crystallins; on the assumption of a close similarity of *Rana* γ-crystallins with bovine γ-II crystallin defined by X-ray analysis. We discuss the differences of sequences in terms of their structural implications and their consequences for intermolecular interactions.

Table 1

The sequences of γ-crystallins from a bovine [5], human [6], rat [6] and frog [11,12] arranged as four structural equivalent Greek-key motifs

Motif 1

	1	10	20	30
Bovine γII	G K I T F Y E D R G F Q G H C Y E C S S	D C P N L Q P Y F S	R C N S I R V D S
Human γ1-2	G K I T F Y E D R A F Q G R S Y E C T T	D C P N L Q P Y F S	R C N S I R V E S
Human γ2-1	G K I T F Y E E Q G F Q G R S Y E T T T	D C P N L Q P Y F S	R C N S I R V E S
Rat γ1-1	G K I T F Y E D R G F Q G R C Y E C S S	D C P N L Q T Y F S	R C N S I R V D S
Rat γ1-2	G K I T F F E D R G F Q G R C Y E C S S	D C P N L Q T Y F S	R C N S V R V D S
Rat γ2-1	G K I T F Y E D R G F Q G R C Y E C S S	D C P N L Q T Y F S	R C N S V R V D S
Rat γ2-2	G K I T F Y E D R G F Q G R H Y E C S T	D H S N L Q P Y F S	R C N S V R V D S
Rat γ3-1	G K I T F Y E D R G F Q G R H Y E C S T	D H S N L Q P Y F S	R C N S V R V D S
Rat γ4-1	G K I T F Y E D R G F Q G R H Y E C S T	D H S N L Q P Y F S	R C N S V R V D S
Frog γ-1	. . . . .	. . . . .	. . . . .
Frog γ-2	. . . . . Y E D R N F Q G R C Y E C S G	D C A D L H S Y F S	R C N S I K V D S

Motif 3

	90	100	110	120
Bovine γII	F R M R I Y E R D D F R G Q M S E I T D	D C P S L Q D R F H L S	E V H S L N V L E
Human γ1-2	Y R M K I Y D R D E L R G Q M S E L T D	D C L S V Q D R F H L T	E I H S L N V L E
Human γ2-1	H R L R L Y E R E D H K G L M M E L S E	D C P S I Q D R F H L S	E I R S L H V L E
Rat γ1-1	H R I R L Y E R D D Y R G L V S E L T E	D C S C I H D R F R L N	E I Y S M H V L E
Rat γ1-2	Y R M R I Y E R D D F R G Q M S E I T D	D C L S L Q D R F H L S	E I H S L N V M E
Rat γ2-1	H R M R L Y E K E D H K G V M M E L S E	D C S C I Q D R F H L S	E V R S L H V L E
Rat γ2-2	H R I R L Y E R E D Y R G Q M V E F T E	D C P S L Q D R F H F N	E I Y S L N V L E
Rat γ3-1	H R I R I Y E R E D Y R G Q M V E I T D	D C P H L Q D R F H F S	D F H S F H V M E
Rat γ4-1	H R I R I Y E R E D Y R G Q M V E I T D	D C P H L Q D R F H F S	D F H S F H V I E
Frog γ-1	F R L R I Y E R E E F R G Q M M E F T E	D C P Q V H E E F N Y H	D I H S C N V L E
Frog γ-2	H K M K I Y E K E E L K G Q M L E V L E	D C P S V F E H F K N H	D I N S C N V L E

Motif 2

	40	50	60	70	80
Bovine γII	G C W M L Y E R P N Y Q G H Q Y F L R R G D Y P D Y Q Q	W M C F N D S I R S C R L I P Q H T C T
Human γ1-2	G C W M I Y E R P N Y Q G H Q Y F L R R G E Y P D Y Q Q	W M G L S D S I R S C C L I P P H S G A
Human γ2-1	G C W M L Y E R P N Y Q G Q Q Y L L R R G E Y P D Y Q Q	W M G L S D S I R S C C L I P Q . T V S
Rat γ1-1	G C W M L Y E R P N Y Q G Y Q Y F L R R G D Y P D Y Q Q	W M G F S D S I R S C R S I P Y . T S S
Rat γ1-2	G C W M L Y E R P N Y Q G H Q Y F L R R G D Y P D Y Q Q	W M G F S D S I R S C R L I P Q H S G T
Rat γ2-1	G C W M L Y E R P N Y Q G H Q Y F L R R G D Y P D Y Q Q	W M G F S D S I R S C R L I P H . T G S
Rat γ2-2	G C W M L Y E Q P N F T G C Q Y F L R R G D Y P D Y Q Q	W M G F S D S V R S C R L I P H . A G S
Rat γ3-1	G C W M L Y E Q P N F T G C Q Y F L R R G D Y P D Y Q Q	W M G F S D S V R S C R L I P H . S S S
Rat γ4-1	G C W M L Y E Q P N F T G C Q Y F L R R G D Y P D Y Q Q	W M G F S D S V R S C H L I P H . S S S
Frog γ-1	. . W M L Y E H P N Y T G H Q Y F L R R G E Y P D F Q Q	W M G L N D S I R S C R V I P Q H R G S
Frog γ-2	G C W M I Y E R P N F L G H Q Y F L K K G E Y P N Y Q Q	W M G F S D S V R S C K V I P Q Q K G P

Motif 4

	130	140	150	160	170
Bovine γII	G S W V L Y E M P S Y R G R Q Y L L R P G E Y R R Y L D	W G A M N A K V G S L R R V M D F Y
Human γ1-2	G S W I L Y E M P N Y R G R Q Y L L R P G E Y R R F L D	W G A P N A K V G S L R R V M D L Y
Human γ2-1	G C W V L Y E L P N Y R G R Q Y L L R P Q E Y R R C G D	W G A M D A K A G S L R R V V D L Y
Rat γ1-1	G S W V L Y E M P N Y R G R Q Y L L R P G D Y R R Y H D	W G A M D A K V G S L R R V M D L Y
Rat γ1-2	G C W V L Y E M P S Y R G R Q Y L L R P G E Y R R Y L D	W G A A N A K V G S F R R V M D F Y
Rat γ2-1	G C W V L Y E M P N Y R G R Q Y L L R P Q E Y R R Y H D	W G A V D A K A G S L R R V V D L Y
Rat γ2-2	G C W V L Y E M T N Y R G R Q Y L L R P G E Y R R Y H D	W G A M N A R V G S L R R V M D F Y
Rat γ3-1	G Y W V L Y E M P N Y R G R Q Y L L R P G E Y R R Y H D	W G A M N A R V G S L R R I M D F Y
Rat γ4-1	G Y W V L Y E M P N Y R G R Q Y L L R P R E Y R R Y H D	W G A M N A R V G S L R R I M D Y Y
Frog γ-1	G H W I L Y E Q P N Y R G R Q Y Y L R P G E Y R R Y T E	W G A V T P R V G S F R R V Q E M F
Frog γ-2	G H W I F Y E Q P N Y R G R Q Y F L K P G E Y K R F S D	W G S L N A R V S S F R R V L D S

Positions in columns are topologically equivalent. The numbering is that of γ-II crystallin

2. MATERIALS AND METHODS

The sequences of the two *Rana* γ-crystallins were aligned with the sequences of bovine, rat and human γ-crystallins. The coordinates of bovine γ-II crystallin [4] which have been refined at 1.6 Å resolution were used as a basis for modelling the three-dimensional structure of the *Rana* γ-crystallins. The main chain and conserved sidechains of the *Rana* γ-crystallins were placed in identical positions to those of bovine γ-II crystallin. The sidechains of residues of the core were arranged in positions closely approximating those of γ-II crystallin, so that short interatomic contacts were avoided. The program FRODO of Jones [14], extended by I.J. Tickle for an Evans and Sutherland picture system, was used for modelling. Sidechains of surface residues were also placed so that hydrogen bonds, ion pairs and hydrophobic interactions were optimized. The resulting models were tested for an even packing and maintenance of the hydrophobic core. Finally the structure has been optimized using a force-field refinement method which minimizes the total energy using the potential of Weiner et al. [15] and a program written by Dr I. Haneef of our laboratory. The resulting models were visualized by the program MIDAS [16] using the Evans and Sutherland picture system 2.

3. RESULTS AND DISCUSSION

Table 1 shows the sequences of the γ-crystallins aligned together. No deletions or insertions are necessary relative to bovine γ-II crystallin for the *Rana* γ-1 and γ-2, rat γ-1-2 and human γ-1-2 sequences.

Fig.1 shows a stereo view of the complete model of *Rana* γ-2 and fig.2 shows a view of *Rana* γ-1 in which motif 1 has the sequence of bovine γ-II crystallin as this motif has not been sequenced. Fig.3 illustrates the arrangement of charged groups on the molecular surfaces.

We now consider the sequences in terms of the three-dimensional models. To facilitate this we have arranged table 1 so that each line corresponds to a Greek-key motif of the three-dimensional structure [3] and each column corresponds to residues at topologically equivalent positions in the four motifs. This makes it immediately evident that Gly-13 and Ser-34 (bovine γ-II crystallin numbering) are residues which are conserved in *Rana* motifs just as in all other γ-crystallin motifs. The glycine is necessary for the folding of strands

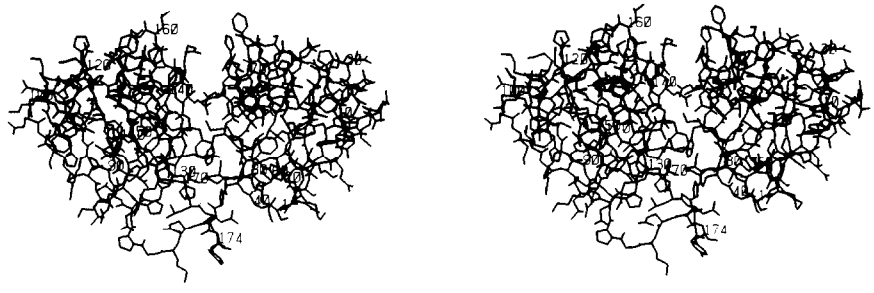

Fig.1. A stereo view of the model of *Rana* γ-2 crystallin.

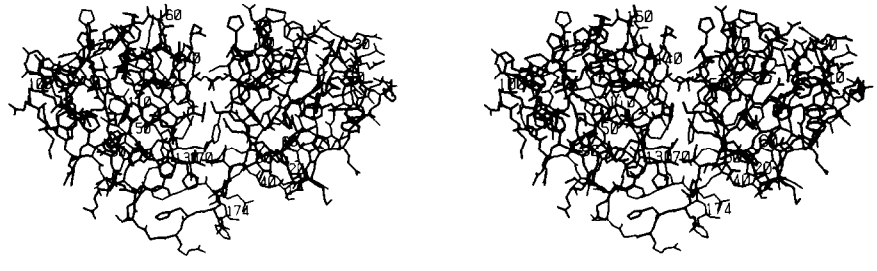

Fig.2. A stereo view of the model of *Rana* γ-1 crystallin in which motif 1 has the sequence of bovine γ-II crystallin.

a and *b*, which form a hairpin, onto the β-pleated sheet contributed to by strands *a*, *b* and *d* of the Greek-key motif and strand *c* of the other motif within the same domain (see [3,4]). The folding of the hairpin is also facilitated by the existence of a glutamate at positions equivalent to Glu-7; this allows a right-handed helical conformation in an otherwise β-sheet stretch of polypeptide. Glutamates are found at these positions in all the *Rana* motif sequences. The fold of the hairpin buries residues on strand *d* topologically equivalent to Ser-34, the γ-oxygen of which hydrogen-bonds to the NH of residue 11.

Residue 11 is at the β-turn at the end of the hairpin and is generally hydrophobic. Topologically equivalent residues in *Rana* γ-crystallin motifs are phenylalanine, tyrosine or leucine, all of which are found in mammalian sequences. The residues equivalent to Tyr-6 interact with this hydrophobic group and are conserved as tyrosines in all γ-crystallins.

The hydrophobic group equivalent to residue 11 is protected from the solvent by an acidic sidechain, an aspartate or a glutamate, and this is conserved as acidic in the *Rana* sequence as in all the other known motifs of γ-crystallins. In motifs 1, 2 and 4, the acidic sidechain is close to that of a basic group which is usually an arginine (equivalent to Arg-36) but in motifs 1 and 2 of *Rana* γ-2 crystallin it is a lysine (see figs 2 and 3). A better ion pair is formed in the solvent facing-motif 1. In motif 3, the topologically equivalent residue is asparagine or histidine; in *Rana* sequences it is asparagine (Asn-125). Residue 152 is an arginine or lysine in all γ-crystallins including those of *Rana* and this makes an alternative ion pair. In motifs 1, 2 and 4 the equivalent residue is a proline (Pro-23) in many γ-crystallin motif sequences. This usually forms a contact with the hydrophobic residue at position 11. In motifs 2 and 3 this is proline in *Rana* sequences but in motif 1 or *Rana* γ-2 it is uniquely alanine. This residue is serine in rat γ-2-2, γ-3-1 and γ-4-1; otherwise it is proline.

These observations indicate that *Rana* γ-crystallin sequences can form the characteristic Greek-key conformations in each of the four motifs.

The cores of each of the domains of all known γ-crystallins are closely packed and hydrophobic [8,10]. The two β-sheets of each domain form a wedge with a narrow region at the bottom which broadens out towards the top. The bottom of the wedge-shaped core is occupied mainly by tryptophans at positions 42 and 131. These are conserved in the *Rana* sequences. In fact they form

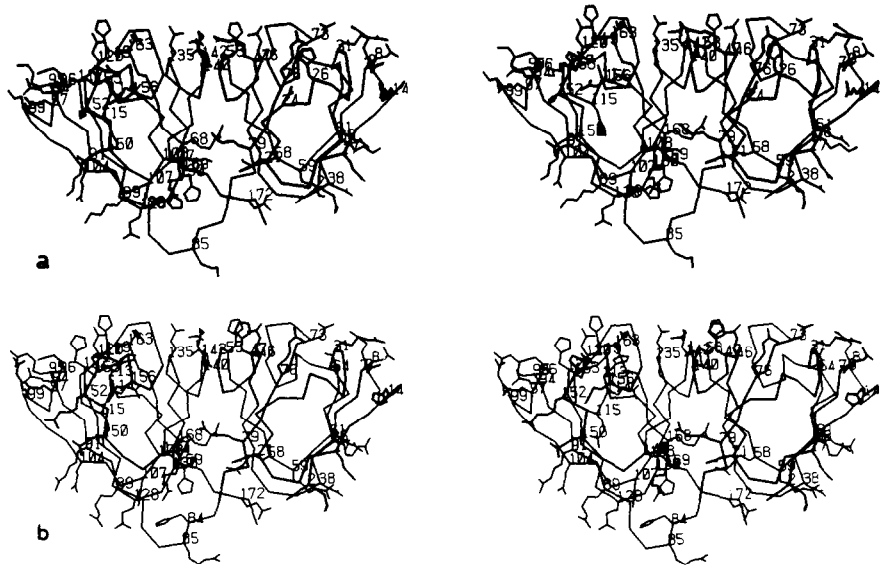

Fig.3. Stereo views of *Rana* γ-2 (a) and γ-1 (b) crystallins in which the virtual bonds between Cα-atoms are shown to indicate the mainchain and the sidechains of the charged groups only are drawn.

part of the hydrophobic surfaces of the two sheets of each domain. The two sheets are slightly displaced from a perfect 2-fold symmetry so that alternating residues interdigitate as can be seen in the schematic fig.4. We have compared the equivalent residues which make up the domain cores of the bovine γ-II and *Rana* γ-1 and γ-2. For domain 1, the contributing residues are remarkably well conserved between bovine and *Rana* crystallins with only one difference and that very conservative (Leu ⟶ 44 Ile). For domain 2, there is more variation corresponding to the greater changes in motif 3. The largest changes are at positions 133 and 167 in the centre of the core. They are retained as hydrophobic in *Rana* sequences, but in *Rana* γ-2 they are both phenylalanines rather than leucines. This appears to force the two opposing sheets slightly further apart. Residues 103 and 105 on the edge of the sheets are generally larger and more hydrophobic especially in *Rana* γ-1 sequence; this emphasises the displacements and leads to an extensive hydrophobic core region.

At the wide part of the wedge there are hydrophobic residues contributed by the intersheet strands. These include tryptophans (68 and 157) which are completely conserved, tyrosines or phenylalanines (65 and 154) and other hydrophobic residues at the topologically equivalent pairs of positions: 71 and 160, 75 and 164, 25 and 112, 29 and 116. The model building shows that these residues in the *Rana* sequences are consistent with the existence of a hydrophobic core

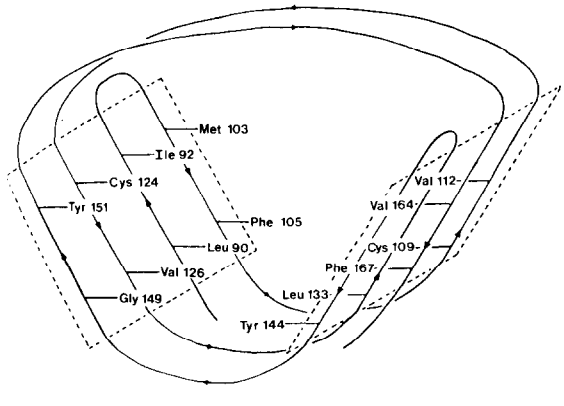

Fig.4. A schematic representation of the hydrophobic core residues of the COOH-terminal domain of *Rana* γ-1 crystallin.

in each of the domains. They also show a general conservation of many of the aromatic-sulphur interactions which have been previously considered to be stabilizing and possibly form delocalized orbitals [4]. These may be of particular importance in the stabilization of free radicals generated in the lens.

The interdomain residue remains hydrophobic in the *Rana* crystallins although the larger phenylalanine and tyrosine at position 145 may slightly displace the relative positions of the two domains. Nevertheless its aromatics can form good interactions with Met-43 on the other domain. From these considerations we deduce that the *Rana* γ-crystallins form the two-domain structures analogous to that of bovine γ-II.

In many proteins the surface residues have varied widely in evolution. In the mammalian γ-crystallins there is an extensive network of surface ion pairs (see fig.3) with a majority of arginines and few lysines contributing the cations. In *Rana* γ-crystallins we have already noted the conservation of the ion pairs which are associated with each folded structure. Most of the other ion pairs are conserved. Many of the arginine positions in motifs 1 and 3 are occupied by lysines in *Rana* γ-2. This may weaken the ion pair interactions as lysines are less likely to be fixed in position than arginines which have fewer degrees of freedom in the sidechains and are less hydrophobic.

Thus the details of the three-dimensional organization of *Rana* γ-crystallins are very similar to those of mammalian crystallins, despite considerable divergence in sequence amounting to 35–40%. It thus appears that the γ-crystallin fold is extremely conservative in evolution and essential for the existence of the stable protein ensemble of the lens.

ACKNOWLEDGEMENTS

We thank Dr Christine Slingsby, Dr Huub Driessen, Dr Peter Lindley, Dr David Moss, Ms Shankari Mylvaganam and other colleagues for many helpful and stimulating discussions.

REFERENCES

[1] Croft, L.R. (1972) Biochem. J. 128, 961–970.
[2] Slingsby, C. and Croft, L.R. (1978) Exp. Eye Res. 26, 291–304.

[3] Wistow, G., Turnell, W.G., Summers, L.J., Slingsby, C., Moss, D.S., Miller, L., Lindley, P.F. and Blundell, T.L. (1983) J. Mol. Biol. 170, 175–202.
[4] Summers, L.J., Wistow, G., Narebor, M., Moss, D.S., Lindley, P.F., Slingsby, C., Blundell, T.L., Bartunik, H. and Bartels, K. (1984) Peptide Protein Rev. 3, 147–168.
[5] Bhat, S.P. and Spector, A. (1984) DNA 3, 287.
[6] Schoenmakers, J.G.G., Den Dunnen, J.T., Moorman, R.J.M., Jongblood, R., Van Leen, R.W. and Lubsen, N.H. (1984) in: Human Cataract Formation, Ciba Found. Symp. 106, pp.208–218, Pitman, London.
[7] 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.
[8] Blundell, T.L., Lindley, P.F., Miller, L., Moss, D.S., Slingsby, C., Tickle, I.J., Turnell, W.G. and Wistow, G. (1981) 289, 771–777.
[9] Chirgadze, Y.N., Sergeev, Y.V., Pomenkova, N.P. and Oreshin, V.D. (1981) FEBS Lett. 131, 81–85.
[10] Driessen, H.P.C. and White, H. (1985) in: Molecular Replacement (Machin, P. ed.) Daresbury Laboratory, SERC.
[11] Tomarev, S.I., Krayev, A.S., Skryabin, K.G., Bayev, A.A. and Gause, G.G. (1982) FEBS Lett. 146, 315–318.
[12] Tomarev, S.I., Zinovieva, R.D., Dolgilevich, S.M., Luchin, S.V., Krayev, A.S., Skryabin, K.G. and Gause, G.G. (1984) FEBS Lett. 171, 297–302.
[13] Gause, G.G., Tomarev, S.I., Zinovieva, R.D., Arutyunyan, K.G. and Dolgilevich, S.M. (1986) Eurage 6, 171–179.
[14] Jones, T.A. (1978) J. Appl. Crystallogr. 11, 268–272.
[15] Weiner, S.J., Kollman, P.A., Case, D.A., Singh, V.C., Ghio, C., Alayon, G., Profota, S. and Weiner, P. (1984) J. Am. Chem. Soc. 106, 765–784.
[16] Tickle, I.J. (1982) MIDAS, A Program for Display of Molecular Structures, Birkbeck College, London University.