The EMBO Journal vol.4 no.9 pp.2385-2388, 1985

Structural details of the binding of guanosine diphosphate to elongation factor Tu from *E. coli* as studied by X-ray crystallography

T.F.M.la Cour, J.Nyborg, S.Thirup[1] and B.F.C.Clark

Division of Biostructural Chemistry, Department of Chemistry, Aarhus University, 8000 Aarhus C, Denmark

[1]Present address: Biomedical Center, Uppsala University, Uppsala, Sweden

Communicated by B.F.C.Clark

Structural details of the guanosine diphosphate binding to a modified form of elongation factor Tu from *Escherichia coli*, resulting from X-ray crystallographic studies, are reported. The protein elements that take part in the nucleotide binding are located in four loops connecting β-strands with α-helices. These loops correspond to regions in primary sequences which show a high degree of homology when compared with other prokaryotic and eukaryotic elongation factors and initiation factor 2.

Key words: elongation factor Tu/GDP binding/sequence homology

Introduction

More than 50% of known proteins recognize or bind phosphates as part of their function, for example in the form of nucleic acids or nucleotides. In protein biosynthesis, the purine nucleotides play a crucial role as regulators of molecular functions and as sources of energy. At the elongation step in particular, guanosine di- and triphosphates are important as co-factors both for elongation factor Tu (EF-1α in eukaryotes) and for elongation factor G (EF-2 in eukaryotes). Other processes in which guanine nucleotides are important are initiation of protein biosynthesis, hormone action and tubulin assembly. Recently, human bladder p21 protein, which also binds GTP (Shih *et al.*, 1980), has attracted attention because of its transformation into an oncogene by a single point mutation. Attempts have been made to predict the three-dimensional structure and to explain the oncogenic properties of the transformed product by comparison of its sequence with those of proteins of known primary structure. Comparisons have thus been made with both nucleotide-binding (Gay and Walker, 1983) and dinucleotide-binding (Wierenga and Hol, 1983) proteins. Furthermore, sequence comparisons have been made among GTP-binding proteins (Halliday, 1984; Leberman and Egner, 1984; Möller and Amons, 1985), but the lack of a detailed spatial description of the binding mode of guanine nucleotide to a protein has until now made such analysis rather speculative.

To facilitate such comparisons, we describe here the structure of the region involved in the GDP/GTP binding in prokaryotic elongation factor Tu as determined by single-crystal X-ray diffraction of a modified form of EF-Tu:GDP from *Escherichia coli*. The folding pattern of the protein around the diphosphate moiety is found to be very similar to that observed in other nucleotide-binding proteins, and the residues involved in the binding of GDP are highly conserved, not only among EF-Tus from different species but also between EF-Tu and both EF-G and IF-2 from *E. coli*. The sequence homology between EF-Tu and *ras* proteins and its implications are the subject of another article (McCormick *et al.*, 1985).

Results

Here we shall limit ourselves to a description of the binding site of the co-factor guanosine diphosphate (GDP). Not all of the interactions described between side chains and GDP are actually seen in the electron-density map, but several are postulated from model building studies.

The GDP molecule binds to the surface of domain I (Clark *et al.*, 1984), which has a typical α/β type structure (Levitt and Chothia, 1976). In the case of EF-Tu, the central core consists of a twisted β-sheet made up of five parallel β-strands and one anti-parallel β-strand. The connections between the parallel strands contain six α-helices, as shown in Figure 1. The nucleotide-binding site is situated at the carboxy end of the β-sheet, such as is found in other nucleotide-binding proteins (Brändén, 1980), and the points of interaction with the protein, i.e., the amino acids involved, are situated in four loops connecting β-strands with α-helices.

The guanine ring is partly buried in a cavity defined by the residues Leu(175), Thr(26) and the side chain of Lys(136) (for the primary structure of EF-Tu from *E. coli*, see Jones *et al.*, 1980). As reported earlier (Rubin *et al.*, 1981), this is in agreement with spectroscopic studies indicating no involvement of π-π interactions with aromatic residues in the protein. Leu(175) is located at the N-terminal end of helix E, and Thr(26) is located in the loop connecting strand a and helix A. At the bottom of this cavity, Asn(135) can form hydrogen bonds to O(6) of guanine, while Asp(138) can form hydrogen bonds to N(1) and/or N(2) of guanine; this may explain the high specificity of EF-Tu towards guanine nucleotides as compared with other purine and pyridimine nucleotides. The replacement of O(6) with sulphur lowers the affinity of EF-Tu for GTP by two orders of magnitude, and its replacement by hydrogen completely abolishes binding (Wittinghofer *et al.*, 1977; Eccleston, 1981). The residues Asn(135) and Asp(138) are both located in the loop connecting strand e with helix D (see Figure 2). Also located in this loop is Lys(136), which by virtue of its long flexible sidechain is able to form hydrogen bonds to the peptide carbonyl groups of Asp(21) and of His(22). This explains the apparent lack of reactivity of Lys(136) towards, for example, modification with ethyl acetamide (Antonsson and Leberman, 1984).

The ribose ring is positioned with its 'endo side' against the protein, leaving the 2′- and 3′-hydroxyl groups exposed to the solvent, in agreement with observations on modification or substitution of these sites (Miller and Weissbach, 1977; Suck and Kabsch, 1981). This arrangement, which places the guanine ring in the *anti*-conformation with respect to the ribose (Morikawa *et al.*, 1978), is necessary in order to position the pyrophosphate in close contact with the protein.

The binding site for the diphosphate residue of the GDP is made up of the loop connecting strand a with helix A. It is commonly

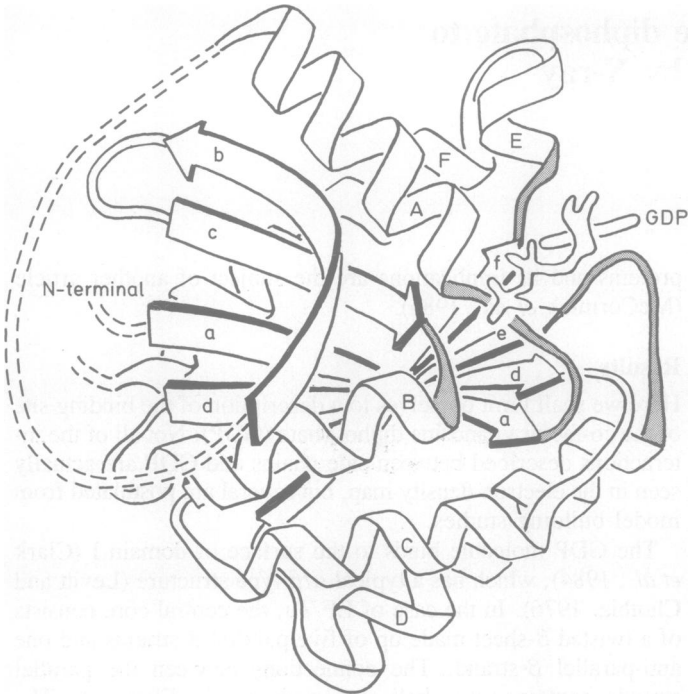

Fig. 1. A structural cartoon showing domain I of EF-Tu from *E. coli*. The picture corresponds to Figure 2 of Rubin *et al.* (1981), but note the change in the order of the secondary structural elements and their connections. Arrows represent β-strands and curled ribbons represent α-helices. The loop connecting helix A with strand b is poorly defined or missing at this stage of the refinement, and the same applies to the first 10 N-terminal residues. The four shadowed loops to the right in the figure are the loops involved in the nucleotide binding described in the text. Also shown is the location of the GDP molecule with the guanine base seen edge on.

found amongst nucleotide-binding proteins that the (pyro)phosphate-binding site is associated with a loop connecting an α-β unit in which the α-helix is approximately anti-parallel to the β-strand, as seen in the Rossmann fold. The advantages of such an arrangement have been summarized by Schulz and Schirmer (1979). In the case of EF-Tu, this loop is rather long and flexible and contains the residues Gly(18)-His-Val-Asp-His-Gly(23). It appears that the interaction with the α-phosphate is limited to a close contact between the pro-R oxygen of the phosphate and Gly(23). This is consistent with biochemical and spectroscopic studies on the binding of GDP to bacterial EF-Tu, which have shown that replacement of the pro-R oxygen with sulphur on the α-phosphate reduces the affinity of this GDP analogue for EF-Tu by a factor of 25 as compared with the replacement at the pro-S oxygen (Goody and Leberman, 1979; Wittinghofer *et al.*, 1982). Furthermore, n.m.r. and e.p.r. studies have shown that the binding of the β-phosphate to the protein is much stronger than the corresponding α-phosphate interaction, and that this interaction involves a Mg(II) ion (Eccleston *et al.*, 1981; Wittinghofer *et al.*, 1982; Leupold *et al.*, 1983; Kalbitzer *et al.*, 1984). It was concluded that two β-oxygens interact with the protein, whereas the third β-oxygen is a ligand of the metal ion. The other ligands in the first coordination sphere of the metal ion are two groups from the protein and three water molecules. A similar arrangement has been found in the Mn:ADP complex at the ATPase site of myosin subfragment 1 (Webb *et al.*, 1982).

In the search for a Mg ion in the electron density map, a peak was located in a pocket formed by the β-phosphate and the loop connecting strand c and helix B. This peak has about half the height of the β-phosphate peak and is located 0.38 nm from the centre of this phosphate group. The loop between strand c and

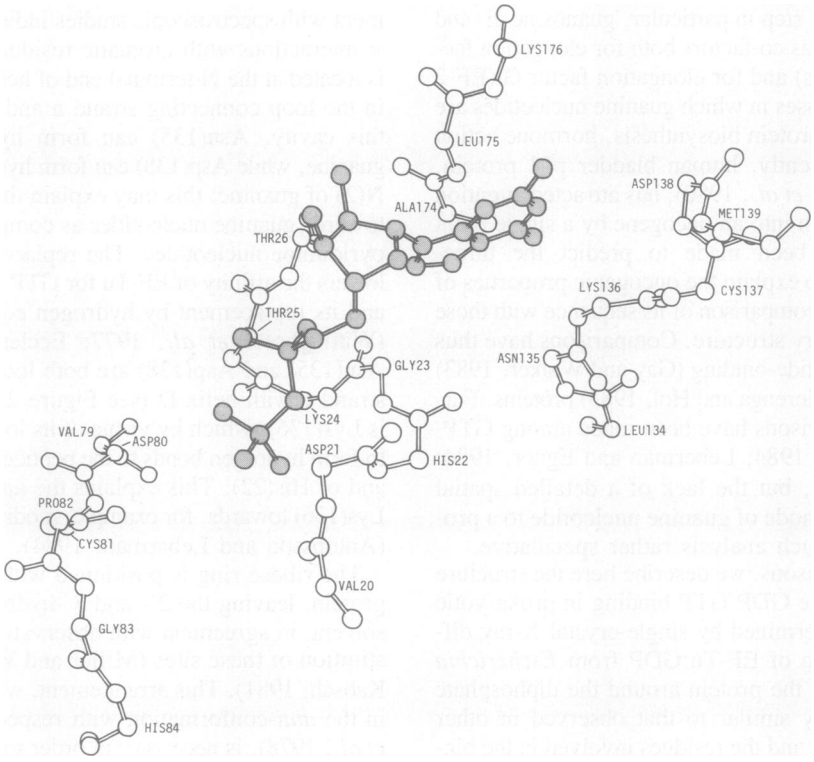

Fig. 2. Close-up picture of the guanosine diphosphate molecule and its environment. The picture is seen from a direction corresponding to a 90° rotation about a vertical axis in Figure 1. The GDP molecule is shown in a darker shade, and for clarity only main-chain atoms of the protein are shown.

helix B is situated diametrically opposite the β-phosphate and consists of residues Asp(80)-Cys-Pro-Gly-His(84). This pocket does not seem large enough to accommodate a γ-phosphate without moving the loop to widen the pocket. It is conceivable that movement of this loop will influence the position of helix B, which – because it forms the interface between domain I and domain II – in turn could influence the relative position of these two domains.

Discussion

For the first time, the detailed structural binding pattern of a guanine nucleotide to a guanine nucleotide-dependent protein is reported. In prokaryotic EF-Tu, the binding takes place on the surface of a rigid part of the structure composed of a central β-sheet flanked by α-helices. The residues involved in the binding are situated in loops connecting the sheet with helices. This is a pattern commonly found in nucleotide-binding proteins, and although folding patterns (i.e., the orders of strands and the helical connections) may differ from protein to protein, some general features are observed, and these are indeed found in EF-Tu. A constant feature is a loop connecting a β-strand at the carboxy edge of a β-sheet with an approximately anti-parallel helix, giving a favourable interaction between the dipole moment of the helix and negatively-charged groups (Hol *et al.*, 1978). Owing to the right-handedness of the connections and the reversed strand order in the sheet, the α-helical connections alternate from one side of the sheet to the other (Brändén, 1980), requiring a somewhat hydrophobic sheet. The length of the loop varies in size, depending on the type of nucleotide-co-factor/substrate but, in order to facilitate close contact between phosphate groups and the helix, glycine residues are always found in the loop close to and/or at the beginning of the helix.

In EF-Tu, this loop is relatively large, consisting of the sequence Gly(18)-His-Val-Asp-His-Gly(23), continuing in the helix with Lys(24)-Thr(25). It is clear from the model that the position of Gly(23) is dictated by the close approach of the α-phosphate to the helix. A comparison of bacterial EF-Tu (Jones *et al.*, 1980) and mitochondrial EF-Tu (Nagata *et al.*, 1983), chloroplast EF-Tu (Montandon and Stutz, 1983) and eukaryotic EF-1α (van Hemert *et al.*, 1984; Nagata *et al.*, 1984), EF-G (Ovchinnikov *et al.*, 1982) and IF-2 (Sacerdot *et al.*, 1984) shows that there is a high degree of conservation of residues in this loop. The largest deviations are His(22), which is Ala in EF-G and Ser in EF-1α and Gly(18), which is also Ala in EF-G.

The sequence of this loop Gly-X-X-X-X-Gly-Lys is a common motif amongst purine-nucleotide binding or processing proteins. In addition to the already mentioned initiation and elongation factors, it is found in P21 proteins, transducing G-proteins, adenylate kinases, ATPases, Fe-nitrogenases, Rec A proteins, GTP:AMP phosphotransferase and myosin.

The loop responsible for the guanine base interaction in *E. coli* EF-Tu, Asn(135)-Lys-Cys-Asp(138), also seems to be strongly conserved. A comparison again between bacterial, mitochondrial and chloroplast EF-Tu, EF-1α, EF-G and IF-2 shows that the only variable residue is the Cys which is replaced by Met in EF-G and EF-1α, Ile in EF-2, Glu in chloroplast and Val in mitochondrial EF-Tu. The observed protection of Cys(137) against sulphydryl reagents in the presence of nucleotide (Miller *et al.*, 1971; Wittinghofer and Leberman, 1976) fits very neatly with the location of this residue in the EF-Tu:GDP structure. In the presence of GDP, this loop containing the cysteine is held in a conformation that maintains the Cys side chain in a half-buried, half-exposed position. On removal of the nucleotide, the loop is no longer held in its conformationally rigid state by interaction with the base, and this leads to full exposure of the Cys side chain to the solvent.

The loop Asp(80)-Cys-Pro-Gly-His(84) contains the other cysteine that has been implicated in the functioning of EF-Tu. Alkylation of Cys(81) abolishes the binding of aminoacyl-tRNA (Miller *et al.*, 1971; Arai *et al.*, 1974) and, conversely, the formation of the ternary complex protects this cysteine from modification with, for example, TPCK (Jonák *et al.*, 1980). The environment of this cysteine in the three-dimensional structure is somewhat similar to that of Cys(137), not fully exposed but half buried. As indicated in the Results, the loop containing Cys(81) may undergo conformational changes, depending on whether GDP or GTP is bound, so it is not possible to conclude from the model whether or not Cys(81) is actively involved in the aminoacyl-tRNA binding. However, this loop is also highly conserved amongst the elongation factors. The only variable residue appears again to be the cysteine, which is some other, at present unknown, residue in EF-Tu from *Thermus thermophilus* (Y.Kaziro, personal communciation), is an Ala in the known EF-1αs and is a Thr in EF-G and IF-2. In view of these observations, it is probable that the results of modification and protection merely reflect an indirect steric effect on the conformation of this loop and hence on the functional state of the molecule.

A knowledge of the interactions between specific groups in macromolecular reactions is a prerequisite for the description and understanding of these complex chemical processes. By means of the newly-developed techniques in protein engineering, subtle changes in very restricted regions of a molecule can be made. We are currently attempting to use the results described above as a basis for further investigation of the mechanism of aminoacyl-tRNA recognition and GTP hydrolysis by applying the technique of site-directed mutagenesis to the EF-Tu molecule.

Materials and methods

In our earlier reports on the structure of elongation factor Tu from *E. coli* (Morikawa *et al.*, 1978; Rubin *et al.*, 1981; Clark *et al.*, 1984), we described the molecule as consisting of three domains. Domain I consists of residues 60–241, arranged in a rather compact structure with one central six-stranded β-sheet surrounded by six α-helices. This domain contains the GDP-binding site. Domain II consists of residues 300–393, and the only secondary structural elements found in this domain are β-strands. Domain III consists of residues 1–36 and 256–299 and contains less well-defined secondary structure. The remaining residues 37–59 and 242–255 constitute connections between domains.

At that time, all model building had been carried out by fitting molecular skeletal model parts to an electron density map based on MIR phases from three derivatives, as reported earlier. The model included 56% of the sequence, and an R-factor calculation showed an agreement of 47% between observed and calculated amplitudes to a resolution of 0.29 nm. In order to improve and extend our model, the following procedure was adopted. Structure factors were calculated based on 56% of the total scattering matter with the use of temperature factors of 0.3 nm^2 obtained from Wilson statistics. The calculated phases were combined with the observed MIR phases and assigned a figure of merit based on Sim weighting (Bricogne, 1976). By using these combined phases and as amplitudes 2 x F(obs) – F(calc), a new electron density map was calculated. On the basis of this map, it was possible to include most of the density corresponding to domain III in the model. This was done, together with minor adjustments to the rest of the model, by utilizing the graphics display system at the Wallenberg Laboratory with the program system FRODO, written by T.A.Jones (Jones, 1978). A new R-factor calculation, including only atoms which were positioned in density in the map, gave a value of 45% and included 62% of the theoretical protein:co-factor scattering matter.

Up to this point, data from the orthorhombic crystal form, which contains two molecules per asymmetric unit, had been used. At this point, it was possible to analyse the data from the tetragonal crystal form as well. Tetragonal crystals had been prepared as described by Morikawa *et al.* (1978), but, instead of proceeding from intact protein, the starting material was subjected to a mild trypsin treat-

ment. Before being placed in the crystallization dish, the protein was treated with a 1% trypsin solution (Worthington, NJ) at 0°C for 30 min, and the reaction was then stopped with soya bean trypsin inhibitor (Sigma, St. Louis) at 3% concentration. The crystallization procedure was identical to that described for the orthorhombic crystals. Three-dimensional diffraction data were collected on film using the oscillation method, and the films were digitized using an Optronics film scanner. Two crystals were used to collect a whole set of native data up to 0.26 nm resolution. The merging and symmetry R-factor for this data set was 4%. The orthorhombic model was then transformed into the tetragonal cell by applying a 45° rotation around the c-axis to all coordinates in the orthorhombic cell. This model was subjected to least-squares refinement with constraints applied to the molecular geometry using the program CORELS (Sussman *et al.*, 1977). The refinement included ~350 residues and GDP, and the R-factor (including data to 0.26 nm resolution) is 33%.

It then became clear than an alternative tracing of the main-chain was possible, placing the N-terminal residues in domain I. This revised interpretation found support in several observations. Sequence homologies amongst several GTP binding proteins clearly showed four very homologous regions amongst the first 200 residues and a better fit of side chain densities was obtained building the N-terminal sequence in strand a and helix A. Since this change is virtually only a replacement of a few side chains, we have not yet attempted to refine the revised model. Our present model described in the text hence consists of the GDP-binding domain I containing residues 1 – 200, domain III containing residues 201 – 300 and domain II containing the C-terminal residues 301 – 393. We have worked out a similar model using our orthorhombic data, and the differences in the geometry of GDP binding in the overall structure of domain I are insignificant. More details on the refinements of the tetragonal and the orthorhombic structures and their results are the subject of an article in preparation.

Acknowledgements

We are indebted to Dr. D.Miller who has provided the purified EF-Tu material. We wish to thank Drs. M.Kjeldgaard and L.Nørskov-Lauritsen for helpful discussions and our colleagues in the Molecular Structure Group at the Biomedical Center, Uppsala University, for their help in using their graphics display system. Our recent model building has also benefitted significantly by discussions of unpublished work with Dr. F.McCormick. We are grateful to the Danish Natural Science Research Council for its continuing support.

References

Antonsson,B. and Leberman,R. (1984) *Eur. J. Biochem.*, **141**, 483-487.
Arai,K.-I., Kawakita,M., Nakamura,S., Ishikawa,I. and Kaziro,Y. (1974) *J. Biochem.*, **76**, 523-534.
Brändén,C.-I. (1980) *Qt. Rev. Biophys.*, **13**, 317-388.
Bricogne,G. (1976) *Acta Crystallogr.*, **A32**, 832-847.
Clark,B.F.C., la Cour,T.F.M., Nielsen,K.M., Nyborg,J., Petersen,H.U., Siboska,G.E. and Wikman,F.P. (1984) in Clark,B.F.C. and Petersen,H.U. (eds.), *Gene Expression. The Translational Step and its Control*, Munksgaard, Copenhagen, pp. 127-148.
Eccleston,J.F. (1981) *Biochemistry (Wash.)*, **20**, 6265-6272.
Eccleston,J.F., Webb,M.R., Ash,D.E. and Reed,G.H. (1981) *J. Biol. Chem.*, **256**, 10774-10777.
Gay,N.J. and Walker,J.E. (1983) *Nature*, **301**, 262-264.
Goody,R.S. and Leberman,R. (1979) *FEBS Lett.*, **102**, 269-272.
Halliday,K.R. (1984) *J. Cyclic Nucleotide Protein Phosphorylation Res.*, **9**, 435-448.
Hol,W.G.J., van Duijnen,P.T. and Berendsen,H.J.C. (1978) *Nature*, **273**, 443-446.
Jonák,J., Smrt,J., Holy,A. and Rychlik,I. (1980) *Eur. J. Biochem.*, **105**, 315-320.
Jones,M.D. Petersen,T.E., Nielsen,K.M., Magnusson,S., Sottrup-Jensen,L., Gausing,K. and Clark,B.F.C. (1980) *Eur. J. Biochem.*, **108**, 507-526.
Jones,T.A. (1978) *J. Appl. Crystallogr.*, **11**, 268-272.
Kalbitzer,H.R., Goody,R.S. and Wittinghofer,A. (1984) *Eur. J. Biochem.*, **141**, 591-597.
Leberman,R. and Egner,U. (1984) *EMBO J.*, **3**, 339-341.
Leupold,C.M., Goody,R.S. and Wittinghofer,A. (1983) *Eur. J. Biochem.*, **135**, 237-241.
Levitt,M. and Chothia,C. (1976) *Nature*, **261**, 552-557.
McCormick,F., Clark,B.F.C., la Cour,T.F.M., Kjeldgaard,M., Nørskov-Lauritsen,L. and Nyborg,J. (1985) *Science (Wash.)*, in press.
Miller,D.L., Hachmann,J. and Weissbach,H. (1971) *Arch. Biochem. Biophys.*, **144**, 115-121.
Miller,D.L. and Weissbach,H. (1977) in Weissbach,H. and Pestka,S. (eds.), *Molecular Mechanisms of Protein Biosynthesis*, Academic Press, NY, pp. 323-373.
Möller,W. and Amons,R. (1985) *FEBS Lett.*, **186**, 1-7.
Montadon,P.-E. and Stutz,E. (1983) *Nucleic Acids Res.*, **11**, 5877-5892.
Morikawa,K., la Cour,T.F.M., Nyborg,J., Rasmussen,K.M., Miller,D.L. and Clark,B.F.C. (1978) *J. Mol. Biol.*, **125**, 325-338.
Nagata,S., Tsunetsugu-Yokota,Y., Naito,A. and Kaziro,Y. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6192-6196.
Nagata,S., Nagashima,K., Tsunetsuga-Yokota,Y., Fujimura,K., Miayazaki,M. and Kaziro,Y. (1984) *EMBO J.*, **3**, 1825-1830.
Ovchinnikov,Y.A., Alakhov,Y.B., Bundulis,Y.P., Dovgas,N.V., Kozlov,V.P., Motuz,L.P. and Vinokurov,L.M. (1982) *FEBS Lett.*, **139**, 130-135.
Rubin,J.R., Morikawa,K., Nyborg,J., la Cour,T.F.M., Clark,B.F.C. and Miller,D.L. (1981) *FEBS Lett.*, **129**, 177-179.
Sacerdot,C., Dessen,P., Hershey,J.W.B., Plumbridge,J.A., Grunberg-Manago,M. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7787-7791.
Schulz,G.E. and Shirmer,R.H. (1979) *Principles of Protein Structure*, published by Springer-Verlag, NY.
Shih,T.Y., Papageorge,A.G., Stokes,P.E., Weeks,M.O. and Scolnick,E.M. (1980) *Nature*, **287**, 686-691.
Suck,D. and Kabsch,W. (1981) *FEBS Lett.*, **126**, 120-122.
Sussman,J.L., Holbrook,S.R., Church,G.M. and Kim,S.-H. (1977) *Acta Crystallogr.*, **A33**, 800-804.
van Hemert,F.J., Amons,R., Pluijms,W.J.M., van Ormondt,H. and Möller,W. (1984) *EMBO J.*, **3**, 1109-1113.
Webb,M.R., Ash,D.E., Leyh,T.S., Trentham,D.R. and Reed,G.H. (1982) *J. Biol. Chem.*, **257**, 3068-3072.
Wierenga,R.K. and Hol,W.G.M. (1983) *Nature*, **302**, 842-845.
Wittinghofer,A. and Leberman,R. (1976) *Eur. J. Biochem.*, **62**, 373-382.
Wittinghofer,A., Warren,W.F. and Leberman,R. (1977) *FEBS Lett.*, **75**, 241-243.
Wittinghofer,A., Goody,R.S., Roesch,P. and Kalbitzer,H.R. (1982) *Eur. J. Biochem.*, **124**, 109-115.

Received on 17 December 1984; revised on 18 June 1985