Shanblatt, S., & Revzin, A. (1984) *Nucleic Acids Res. 13*, 5287–5306.
Shimamoto, N., & Wu, C. W. (1980) *Biochemistry 19*, 842–848.
Stefano, J. E., & Gralla, J. D. (1979) *Biochemistry 18*, 1063–1067.
Stefano, J. E., & Gralla, J. D. (1982) *J. Mol. Biol. 257*, 13924–13929.
Straney, D. C., & Crothers, D. M. (1985) *Cell (Cambridge, Mass.) 43*, 449–459.
Straney, D. C., & Crothers, D. M. (1986) *J. Mol. Biol.* (in press).
Walter, G., Zillig, W., Palm, P., & Fuchs, E: (1967) *Eur. J. Biochem. 3*, 194–201.
Wang, J. C., Jacobson, J. H., & Saucier, J. M. (1977) *Nucleic Acids Res. 4*, 1225–1241.

The Protein Synthesis Initiation Factor 2 G-Domain. Study of a Functionally Active C-Terminal 65-Kilodalton Fragment of IF2 from *Escherichia coli*†

Yves Cenatiempo,‡ Francoise Deville,‡ Jacques Dondon,§ Marianne Grunberg-Manago,§ Christine Sacerdot,§ John W. B. Hershey,∥ Hans Frede Hansen,⊥ Hans Uffe Petersen,*,# Brian F. C. Clark,# Morten Kjeldgaard,# Troels F. M. la Cour,# Kim Kusk Mortensen,# and Jens Nyborg#

Laboratory of Molecular Biology, University of Lyon, 69622 Villeurbanne, France, Institut de Biologie Physico-Chimique, 75005 Paris, France, Department of Biological Chemistry, School of Medicine, University of California, Davis, California 95616, University Department of Clinical Chemistry, State University Hospital, DK-2100 Copenhagen, Denmark, and Division of Biostructural Chemistry, Department of Chemistry, Aarhus University, DK-8000 Aarhus C, Denmark

Received October 24, 1986; Revised Manuscript Received February 20, 1987

ABSTRACT: Protein synthesis initiation factor 2 (IF2) is present in *Escherichia coli* cells as two forms which are expressed from the same gene: IF2α [97.3 kilodaltons (kDa)] and IF2β (79.7 kDa). During isolation, a smaller form, IF2γ, is generated, presumably by partial proteolysis. It has been purified to homogeneity and has an apparent mass of 70 kDa, as determined by sodium dodecyl sulfate–polyacrylamide gel electrophoresis. Immunoelectrophoresis of IF2α and IF2γ shows that IF2γ is immunologically partially identical with IF2α. The sequence of the 15 N-terminal amino acid residues of IF2γ was determined and compared with that of IF2α. The N-terminal amino acid of IF2γ corresponds to Arg-290 of IF2α, suggesting that IF2γ is generated by proteolytic cleavage of the Lys-289–Arg-290 bond of IF2. Assuming a C terminus identical with IF2α, we calculate that IF2γ comprises 601 amino acid residues and has a mass of 64.8 kDa. The truncated protein was tested for activities characteristic of IF2 in three in vitro assays: fMet-tRNA$_f^{Met}$ binding to 70S ribosomes, N-terminal dipeptide synthesis in a DNA-dependent transcription/translation system, and ribosome-dependent GTP hydroly97–7 The specific activities of IF2γ were comparable with, or only slightly less than, thoşe for IF2α, indicating that IF2γ contains the active centers for interaction with fMet-tRNA$_f^{Met}$, ribosomes, and GTP. A central region in the primary structure of IF2 shows extensive sequence homology with a number of GDP-binding proteins and especially with the G-domain of elongation factor Tu (EF-Tu). A predicted secondary structure of this region of IF2 (present in IF2α, IF2β, and IF2γ) is shown. On the basis of our detailed knowledge about the structure of EF-Tu [la Cour, T. F. M., Nyborg, J., Thirups, S., & Clark, B. F. C. (1985) *EMBO J. 4*, 2385–2388; Jurnak, F. (1985) *Science (Washington, D.C.) 230*, 32–36], and the regions of extensive sequence homology between EF-Tu and IF2, a model for the tertiary structure of the IF2 G-domain was built with the aid of a graphic display system. The model obtained is a 17-kDa globular domain approximately 35 Å in diameter.

Initiation of protein synthesis in *Escherichia coli* is promoted by three proteins, initiation factors 1, 2, and 3 (IF1, IF2, and IF3) [for extensive reviews, see Grunberg-Manago (1980) and Maitra et al. (1982)]. IF2 is the largest of the initiation factors and is present in bacterial cells in two size classes, IF2α [97.3 kilodaltons (kDa)] and IF2β (79.7 kDa) (Howe & Hershey, 1982). It interacts with at least three components of the initiation pathway: GTP, fMet-tRNA$_f^{Met}$, and ribosomes. Through these interactions it promotes the binding of fMet-tRNA$_f^{Met}$ to the 30S ribosomal subunit and catalyzes the hydrolysis of GTP after 70S initiation complex formation (Petersen, 1985). The gene for IF2, *infB*, has been cloned and mapped (Plumbridge et al., 1982), and the primary sequence of IF2α has been deduced from the sequence of the gene (Sacerdot et al., 1984). A second translational initiation site in the mRNA accounts for the expression of IF2β from the same gene (Plumbridge et al., 1985). Examination of the amino acid sequence of IF2 revealed an unusual, highly charged repeating structure in the N-terminal half of the molecule and regions in the C-terminal half with homology

† This work was supported by a grant from the CNRS (UA 1176) to Y.C. and F.D.; grants from CNRS (UA 585), MRI (82V1289), INSERM (831.003), and E. I. du Pont de Nemours and Co. to M.G.-M.; grants from the American Cancer Society (NP-70) and from the National Science Foundation (INT83-12982) to J.H.; grants from the Danish Natural Science Research Council (5-17-4.1.04 to B.F.C.C. and H.U.P., 11-5992 to H.U.P., 11-4008 to T.l.C. and J.N., and 5.10.17.05 to J.N.); the Thomas B. Thriges Fund and Direktør Ib Henriksens Fund to T.l.C. and J.N.; and the Niels Bohr Fellowship Committee and Aarhus University Research Fund to H.U.P.

‡ University of Lyon.
§ Institut de Biologie Physico-Chimique.
∥ University of California.
⊥ State University Hospital.
Aarhus University.

0006-2960/87/0426-5070$01.50/0 © 1987 American Chemical Society

to elongation factor Tu (EF-Tu), elongation factor G (EF-G), and a variety of other GTP-binding proteins (Sacerdot et al., 1984).

An approach to studying the function of a factor such as IF2 is to characterize truncated forms of the protein. IF2β is precisely homologous to IF2α except that it lacks 157 amino acid residues of the N terminus (Plumbridge et al., 1985) and it is active in numerous assays of initiation of protein synthesis in vitro. An even shorter form of IF2, called IF2-70K, was purified from cells that overproduce IF2 proteins owing to the presence of a multicopy plasmid carrying *infB* (Dondon et al., 1985). This truncated IF2, which we shall call IF2γ, appears to result from proteolysis of IF2 during purification, since it was not detected in fresh lysates by immunoblotting with anti-IF2α (Howe & Hershey, 1982).

In this paper, we characterize IF2γ functionally and demonstrate that it retains the ability to stimulate assays of initiation in vitro as well as GTPase activity. IF2 contains regions of extensive amino acid sequence homology with several GTP-binding proteins. Of special interest is the homology with EF-Tu, for which three-dimensional structural data are available. We have examined in detail the common GDP-binding domain of IF2α, IF2β, and IF2γ and developed a tertiary structure model of the region based on the X-ray crystallographic model for EF-Tu (Clark et al., 1984; la Cour et al., 1985; Jurnak, 1985).

Materials and Methods

Materials. IF2α and IF2γ from *E. coli* were purified as described by Dondon et al. (1985). Elongation factor Tu was isolated from *E. coli*, as described by Morikawa et al. (1978). Rabbit antisera against IF2α and EF-Tu were raised by subcutaneous injection of 100 μg of each antigen in Freund's complete adjuvant followed by intravenous injection of 100 μg of antigen in 0.09% NaCl 30 days later. Fifty milliliters of blood was recovered 10 days after the boosts. After coagulation, the serum was recovered and was used in the immunoelectrophoresis experiments. Tests with crude protein extracts from *E. coli* showed that the antisera were monospecific against IF2 and EF-Tu, respectively.

Immunoelectrophoresis. Crossed immunoelectrophoresis was performed essentially as described by Høiby and Axelsen (1983). Electrophoresis in the first dimension was at 250 V for 40 min and in the second dimension was at 70 V for 18 h. The amounts of antisera and antigens are indicated in the legend to Figure 1. Crossed-line immunoelectrophoresis (Krøll, 1983) was run as described above, with intermediate agarose gels in the second dimension.

N-Terminal Amino Acid Sequence Determination. One milligram of the protein was dissolved in 0.06 mL of 0.1% (v/v) acetic acid and subjected to automated sequence analysis by repeated Edman degradation on an Applied Biosystems 470A gas-phase sequenator with the MHTFA1 program of M. Hunkapiller [a program modified from Hunkapiller et al. (1983), available at Applied Biosystems, Foster City, CA]. Heptane was used as a supplementary solvent (all chemicals were purchased from Applied Biosystems). The phenylthiohydantoin derivatives of amino acids were characterized by high-performance liquid chromatography (HPLC) on a Hewlett-Packard 1084 liquid chromatograph with a 0.45 cm $\times$ 25 cm column of CN (5-μm particles) and with a sodium acetate/acetonitrile gradient elution system as described by Hunkapiller and Hood (1983). The samples from the sequenator were methylated before HPLC by treating the dried derivatives with acidified methanol (1 M HCl in methanol from Applied Biosystems) for 10 min at 50 °C. Aminobutyric acid was used as internal standard during HPLC for correction of elution time and for quantifying the amino acid derivatives.

N-Terminal Dipeptide Synthesis. The incubation conditions for in vitro dipeptide synthesis in a DNA-directed coupled system were as previously described (Robakis et al., 1983; Plumbridge et al., 1985), except that 4 units of RNasin was added to the incubation mixture. The assay for the dipeptide product was based on the separation of the ^{3}H amino acid on a minicolumn of Dowex 50 WX-8 (H^+ form) as described earlier (Weissbach et al., 1984). In some experiments, translation was uncoupled from transcription. For these, a preliminary 30-min incubation at 37 °C was performed in the absence of aminoacyl-tRNAs, 70S ribosomes, and initiation factors. After the first incubation, rifampicin (29 μg/mL) was added to stop RNA synthesis, and translation was initiated by the addition of the components.

GTPase Activity. IF2-mediated GTP hydrolysis was monitored according to the procedure of Kolakofsky et al. (1968) by measuring the amount of [^{32}P]phosphate liberated during the reaction. The incubation mixture for measuring uncoupled GTPase activity contained in 25 μL 50 mM ammonium chloride, 7 mM β-mercaptoethanol, 800 pmol of [γ-^{32}P]GTP (specific radioactivity 400–650 cpm/pmol), 17 pmol of 70S ribosomes, and IF2α or IF2γ as indicated. Coupled GTPase activity was measured under similar conditions except that the ammonium chloride concentration was increased to 75 mM and the incubation medium was supplemented with 0.145 A_{260} unit of poly(A,U,G) and 13.5 pmol of fMet-tRNA$_f^{Met}$. Unless otherwise indicated, the reactions were carried out for 20 min at 37 °C and then stopped by the addition of 0.1 mL of 1 M perchloric acid and 1 mL of KH_2PO_4. The mixtures were treated as described by Beaudry et al. (1979), and liberated [^{32}P]phosphate was counted in 10 mL of Hydrocount (Baker Chemical, Deventer, The Netherlands). Values were corrected for background GTP hydrolysis detected in the absence of ribosomes.

Molecular Model Building. The computer model of IF2 was built by using an Evans & Sutherland PS300 picture system on a VAX 11/780 and the program Frodo from Alwyn Jones (Biomolecular Center, Uppsala University). The crystallographic model for *E. coli* EF-Tu (la Cour et al., 1985) was used as a framework. Deletions and insertions of amino acid residues were then modeled manually by the distance matrix method of Jones and Thirup (1986). The resulting model was subjected to several cycles of energy minimization by the program of Jack and Levitt (1978).

Results

Immunological Comparison of IF2α and IF2γ. IF2α and IF2γ were isolated from an overproducing strain of *E. coli* that contains a plasmid carring *infB* under the control of the inducible IPL promoter (Dondon et al., 1985). Rabbit antibodies raised against IF2α were used to compare IF2α and IF2γ by crossed immunoelectrophoresis (Figure 1A–C) and crossed-line immunoelectrophoresis (Figure 1D,E) as described under Materials and Methods. The appearance of two partly fused immunoprecipitation peaks (Figures 1C,D) when mixtures of IF2α and IF2γ are tested indicates partial immunological identity between the two antigens and additional antigenic determinants in IF2α compared to IF2γ. The observed serological relatedness of IF2γ and IF2α under nondenaturing conditions agrees with the finding that both proteins react with anti-IF2α following blotting from denaturing sodium dodecyl sulfate (SDS)–polyacrylamide gels (Dondon et al., 1985).

N-Terminal Amino Acid Sequence of IF2γ. IF2γ migrates as a 70-kDa protein on SDS–polyacrylamide gels and is

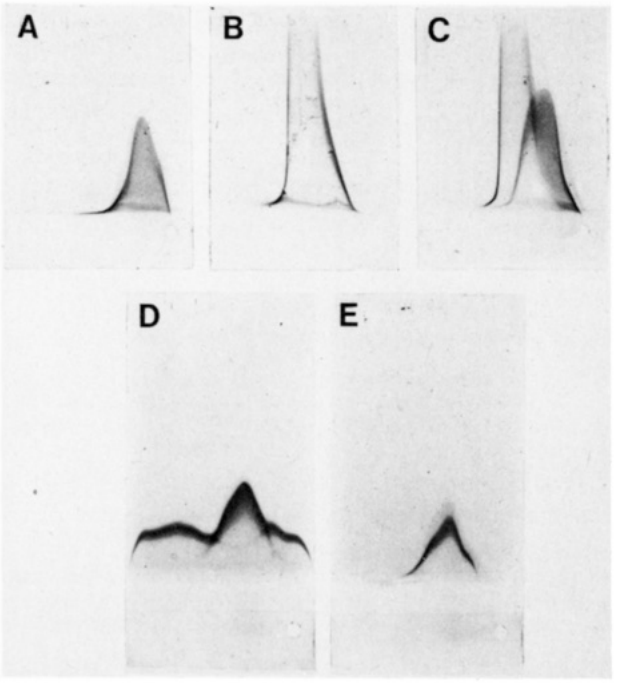

FIGURE 1: Comparison of IF2α and IF2γ by two-dimensional immunoelectrophoresis. Crossed (panels A–C) and crossed-line (panels D and E) immunoelectrophoreses were performed as described under Materials and Methods. Anti-IF2α antiserum was added to the two-dimensional gels at concentrations of 7 μL/cm^2 (panels A–C) and 4 μL/cm^2 (panels D and E). Panel A, 10 μg of IF2α; panel B, 10 μg of IF2γ; panel C, 10 μg each of IF2α and IF2γ; panel D, 5 μg of IF2α and in the intermediate gel 10 μg of IF2γ; panel E, 5 μg of IF2α. The figure shows a photograph of the Coomassie-stained precipitin lines.

therefore considerably smaller than IF2α or IF2β. Since its presence in fresh lysates is not detected, it is likely that IF2γ is generated by partial proteolysis of the native IF2 proteins during their purification. In order to investigate further the structural similarities between IF2α and IF2γ, the sequence of the 15 N-terminal amino acids was determined by Edman degradation and HPLC as described under Materials and Methods and found to be

Arg-Lys-Gly-Ser-Ser-Leu-Gln-Gln-Gly-Phe-Gln-Lys-Pro-Ala-Gln-

Comparison with the IF2α sequence shows an exact correspondence of residues 290–304, indicating that IF2γ may have arisen by cleavage of the Lys-289–Arg-290 bond of IF2α or IF2β. The secondary structure prediction for IF2α (which is described in more detail below) places the two residues in a loop connecting an α-helix and a β-strand (Figure 2); the loop may therefore be exposed and accessible to proteolytic enzymes. A protein beginning at Arg-290 and extending to the C terminus of IF2α comprises 601 amino acid residues and a mass of 64.8 kDa. Since the masses of IF2 proteins, when measured by SDS–polyacrylamide gel electrophoresis (PAGE), appear consistently greater than those calculated from the gene sequence (116 vs. 97.3 kDa for IF2α; 90 vs. 79.7 kDa for IF2β), the sequence value of 64.8 kDa for IF2γ is in good agreement with the value of 70 kDa measured by SDS–PAGE. These results suggest that IF2γ may contain the entire C terminus of IF2α, although this issue has not been addressed experimentally.

Functional Comparison of IF2α and IF2γ. Since IF2γ represents a truncated form of IF2α lacking the N-terminal third of the protein, it was of interest to determine whether it was functionally active. To this end, biological activities of IF2α and IF2γ were compared in a variety of assays in vitro: fMet-tRNA$_f^{Met}$ binding to 70S ribosomes, dipeptide synthesis in a coupled, DNA-directed transcription–translation system, and ribosome-dependent GTP hydrolysis. A comparison of the ability to promote fMet-tRNA$_f^{Met}$ binding to 70S ribosomes in the presence of poly(A,U,G) is shown in Figure 3. We have observed that IF2γ loses activity on repeated freezing and thawing, whereas IF2α is less affected. Therefore, IF2γ appears to be able to interact with both fMet-tRNA$_f^{Met}$ and ribosomes. Under the conditions of the assay, IF2 does not

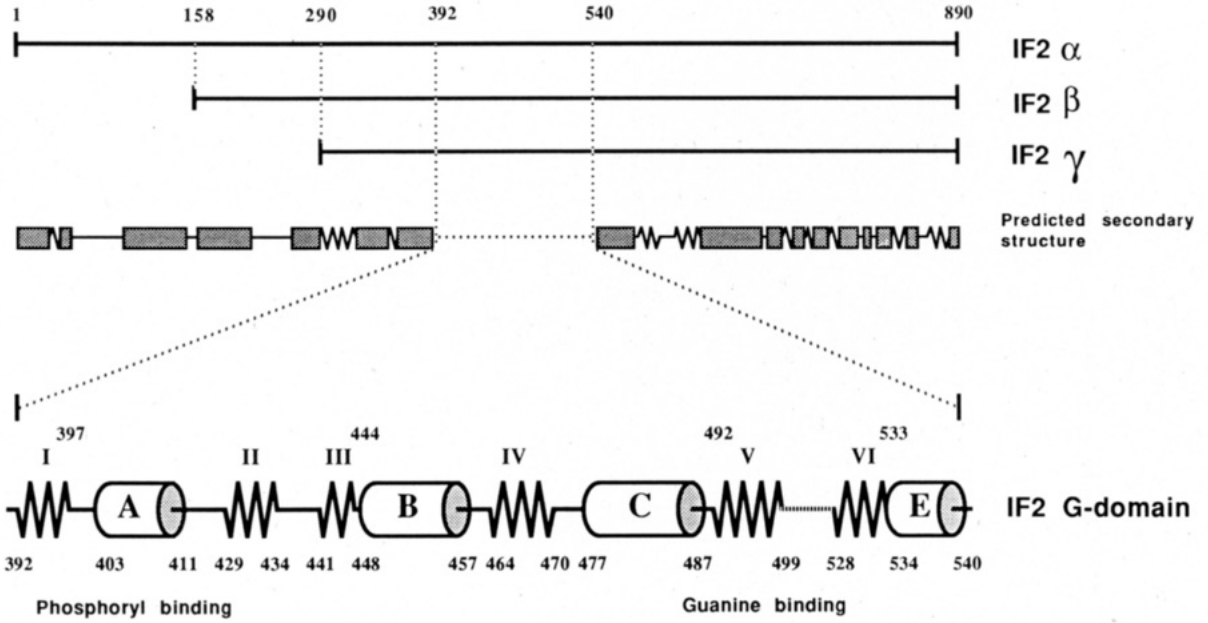

FIGURE 2: Primary and secondary structures of IF2. A schematic representation of the primary structures of IF2α, IF2β, and IF2γ is shown, based on information cited in the text. The secondary structure of IF2α was computed by the methods of Chou and Fasman (1978) and Garnier et al. (1978), and structures in agreement by both methods are shown. Numbers indicate amino acid residue numbers from the N terminus of IF2α (Sacerdot et al., 1984). The IF2 G-domain region (residues 392–540) is shown in an expanded form. Predicted α-helices (A, B, C, and E) are symbolized by cylinders; β-sheets (I–VI) are symbolized by zigzag lines. The dotted line between amino acid residues 499 and 528 indicates the region containing the α-helix D in the computer model (Figure 7) based on sequence homology with EF-Tu (Figure 6). The helix does not appear by secondary structure prediction.

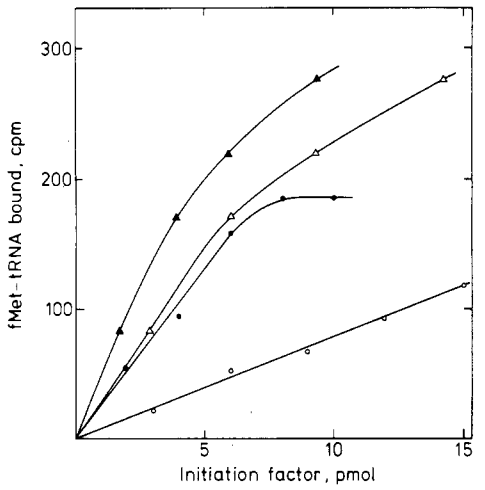

FIGURE 3: fMet-tRNA$_f^{Met}$ binding to 70S ribosomes. Incubation mixtures (100 μL) contained 50 mM tris(hydroxymethyl)aminomethane hydrochloride, pH 7.4, 80 mM ammonium chloride, 5 mM magnesium acetate, 1 mM GTP, 0.15 A_{260} unit of poly(A,U,G), 1.16 A_{260} units of 70S ribosomes, 11.4 pmol of f[^{35}S]Met-tRNA$_f^{Met}$ (specific radioactivity 337 cpm/pmol), and the indicated concentrations of IF2α (solid symbols) and IF2γ (open symbols). The experiment was carried out with freshly prepared factors (triangular symbols) and with factor preparations that had been thawed out 3 times (circular symbols). Mixtures were incubated for 15 min at 30 °C, and nitrocellulose-filtered samples were counted in a liquid scintillation counter. The data were corrected for a background of 189 cpm.

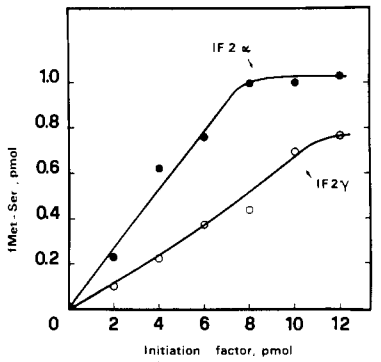

FIGURE 4: Dipeptide synthesis assay. fMet-Ser synthesis from the β-lactamase gene on plasmid pBR322 was determined as described under Materials and Methods. The concentrations of IF2α (solid symbols) and IF2γ (open symbols) are indicated in the figure.

act catalytically, so conclusions about the GTPase function of IF2 cannot be drawn.

A more complex assay of IF2 function during initiation is the dipeptide synthesis assay. The formation of amino-terminal dipeptides in a DNA-dependent protein-synthesizing system reflects accurate initiation of translation, as already demonstrated (Robakis et al., 1983). In such a system, complete dependence on IF2 has been reported previously (Robakis et al., 1981; Peacock et al., 1982). Generally, an excess of IF2 is added as compared to the amount of N-terminal dipeptide synthesized. This can be due to the activity of IF2 itself or alternatively to a binding effect of other components in the system such as RNA polymerase, ribosomes, or other protein factors. The assay was used to test the activity of the two forms of IF2. The synthesis of fMet-Ser, the N-terminal dipeptide of β-lactamase expressed from the plasmid pBR322, was tested in the presence of IF2α and IF2γ (Figure 4). In the linear section of the concentration curve, IF2γ is only half as active as IF2α, but upon addition of saturating amounts of factors, a difference of only 20–25% is detected in the amount of dipeptide synthesized. The time course of fMet-Ser

Table I: Activity of IF2γ in the Dipeptide Synthesis Assay[a]

product of expressed gene	dipeptide product	rel act.
chloramphenicol acetyltransferase	fMet-Glu	85
IF2α	fMet-Thr	79
IF2β	fMet-Ser	77
p15B	fMet-Ala	85
β-lactamase	fMet-Ser	81

[a] The dipeptide assay was carried out with the plasmid pFY16-1 (Plumbridge et al., 1985), which carried four of the genes involved, and pBR322, which carried the β-lactamase gene. Relative activity is defined by 100(dipeptide synthesis catalyzed by IF2γ)/(dipeptide synthesis catalyzed by IF2α).

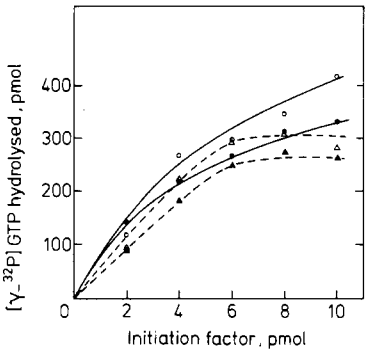

FIGURE 5: GTP hydrolysis assay. Coupled (open symbols) and uncoupled (solid symbols) GTPase reactions were carried out as described under Materials and Methods. The concentrations of IF2α (solid curves and circular symbols) and IF2γ (dashed curves and triangular symbols) are shown in the figure.

synthesis was also examined in the assay where translation is uncoupled from transcription (see Materials and Methods). Both forms of IF2 catalyzed a linear synthesis of dipeptide for at least 90 min, with the slope of the IF2γ line at about 85% of that for IF2α (results not shown). The fact that dipeptide formation can continue in a linear fashion for a long period of time has been observed previously (Cenatiempo et al., 1982) and was explained by the lability of the dipeptidyl-tRNA–ribosome–mRNA complex. As a result, the 70S ribosomes and mRNAs become available for further initiation of protein synthesis. As far as this property is concerned, the proteolyzed form of IF2 reacts as well as the native form.

By use of different plasmids as template, the synthesis of other dipeptide gene products has been tested, namely, chloramphenicol acetyltransferase (fMet-Glu) and all products of the plasmid carrying the *nusA-infB* operon (Table I). Upon addition of a saturating amount (10 pmol) of either IF2α or IF2γ, it appears that the latter is slightly less active (77–85%) than the former.

In the third type of in vitro assay, the GTPase function of both IF2α and IF2γ was tested directly. IF2 catalyzes GTP hydrolysis in the presence of 70S ribosomes (uncoupled) or in the presence of ribosomes, fMet-tRNA$_f^{Met}$, and mRNA (coupled). Both the coupled and uncoupled GTPase functions of IF2α and IF2γ were compared (Figure 5). The activities of the two IF2 forms are indistinguishable in these assays. Coupled synthesis is slightly higher than uncoupled synthesis for the two factors. Under both conditions, the IF2s function catalytically, hydrolyzing about 50 molecules of GTP per molecule of IF2 added. The kinetics of the uncoupled GTPase activities were measured also; both IF2 forms performed comparably (results not shown).

Three-Dimensional Model of the IF2 G-Domain. The relation between the primary structures of IF2α, IF2β, and IF2γ is shown in Figure 2. The N termini of IF2β (Val-158) and

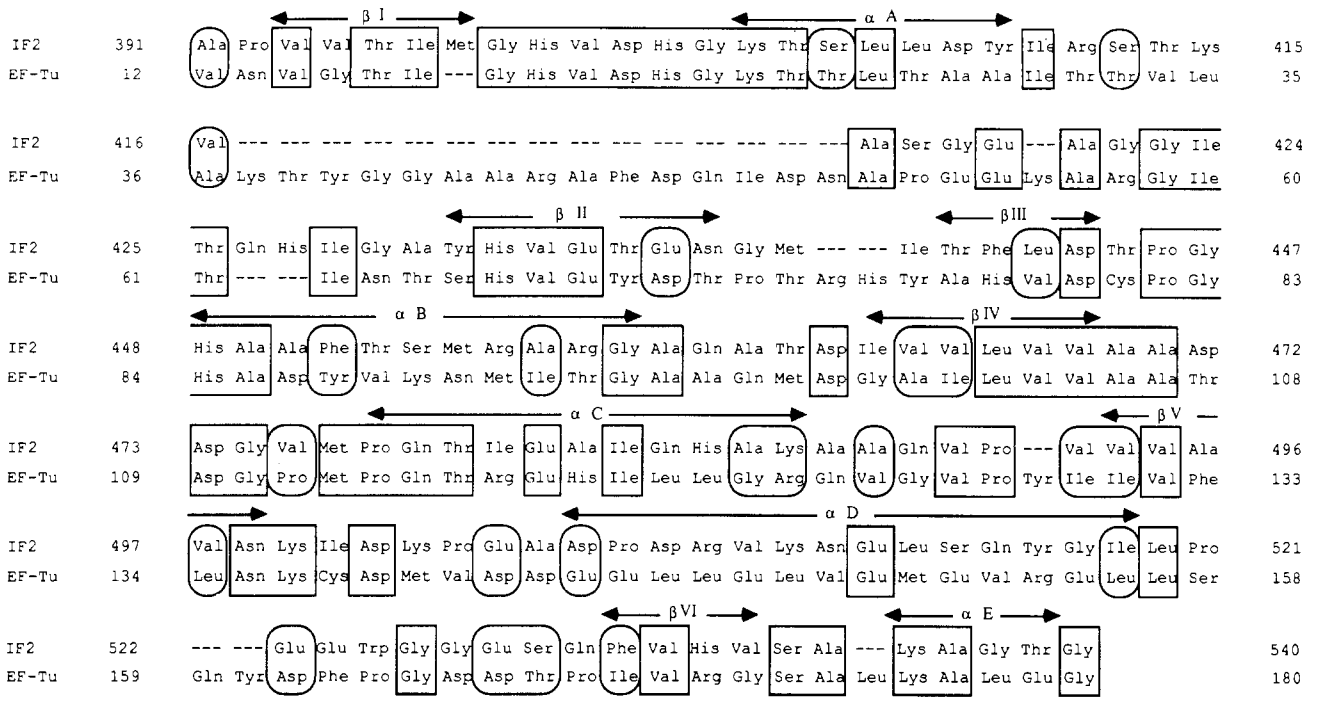

FIGURE 6: Sequence homologies between IF2 and EF-Tu. Homology between the G-domains of IF2 (residues 391–540) and EF-Tu (residues 12–180) was determined by maximizing identical or near-identical residues. Identical residues are in squared boxes and functionally conserved residues are in rounded boxes. Residues absent in the strands compared are shown by dashes. The sequences forming β-strands or α-helices in the three-dimensional model (Figure 7) are indicated.

IF2γ (Arg-290) are indicated, and the figure is drawn on the assumption that all three IF2 forms contain the same C terminus. Two methods (Chou & Fasman, 1978; Garnier et al., 1978) were employed to predict the secondary protein structure on the basis of the amino acid sequence of IF2α (which includes IF2β and IF2γ). The secondary structure obtained by both methods is illustrated in Figure 2. The prediction is shown in more detail for the region (residues 392–540) known to be partially homologous to the GTP-binding region (G-domain) of EF-Tu (Sacerdot et al., 1984). This region is part of all three forms of IF2 and presumably represents its G-domain.

The predicted secondary structure in the homologous G-domain of IF2 contains six β-strands numbered I–VI and four α-helices numbered A, B, C, and E. The α-helix D, which emerges from sequence comparison with EF-Tu (see Figure 6), is not observed in this prediction. The two regions in IF2 that are homologous to the phosphoryl- and guanine-binding regions in EF-Tu are labeled in Figure 2 and are discussed in greater detail below. A striking feature of the arrangement of the α-helices and β-strands is that the segments are found in the same order in IF2 and EF-Tu. This feature strongly suggests that the higher order structures of the two factors are similar. By aligning the elements of secondary structure in the same order, a modified and more extensive homologous arrangement between the primary sequences of IF2 and EF-Tu is possible (Figure 6). Of the 150 residues of IF2 compared to those of EF-Tu, 59 are identical and another 24 are closely related.

A computer model of the G-domain of IF2 was built on the basis of this sequence homology and the crystallographic model of E. coli EF-Tu (Figure 7). The amino acid residues (numbered from the N terminus of IF2α) that are included in this model are 389–411, 424–434, and 443–540. Residues 416–423 align with a region of EF-Tu for which coordinates have not been determined, so they were not included in the IF2 model. The overall folding of the IF2 peptide chain in

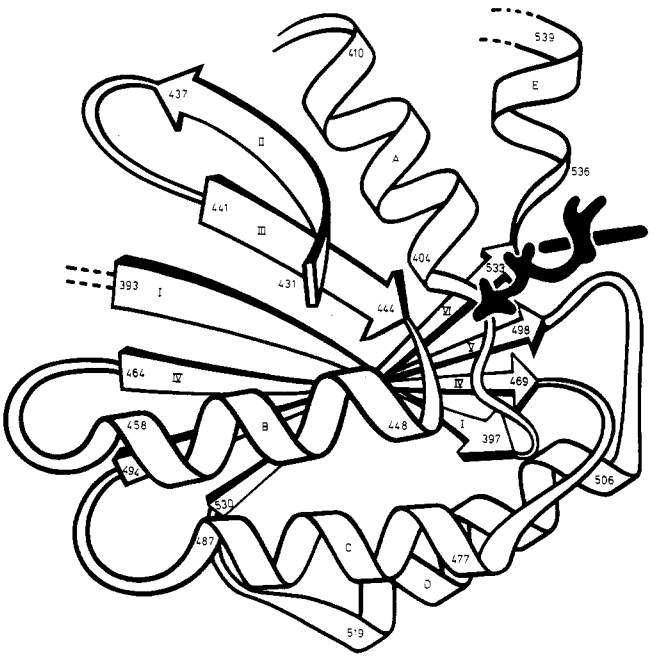

FIGURE 7: Structural cartoon for the tertiary structure model of the IF2 G-domain. The arrows represent β-strands (I–VI), and the curled ribbons represent α-helices (A–E). Arabic numbers indicate IF2α amino acid residues at the ends of secondary structure elements. A GDP molecule is shown in black.

the model is similar to that of EF-Tu (la Cour et al., 1985; McCormick et al., 1985; Jurnak, 1985). One feature is a double Rossmann fold formed by the secondary structural sequence $\beta_{IV}\alpha_C\beta_V\alpha_D\beta_{VI}$. The guanine-binding site is situated at the carboxy end of the sheet, with the guanine moiety wedged in a hydrophobic pocket between β-strands V and VI. The phosphates are situated at the N terminus of the connecting α-helix A, giving a favorable interaction between the dipole moment of the α-helix and the negative charge of the diphosphate group (Wierenga & Hol, 1983). Amino acid

residues 398–403 form a loop connecting β-strand I and α-helix A, and residues 404–405 form the beginning of α-helix A. This loop is responsible for the binding of the pyrophosphate group and exhibits the characteristic sequence Gly-X-X-X-X-Gly-Lys-Thr, where the lysine is directly involved in the binding of the β-phosphate (Möller & Amons, 1985). The segment 498–501, with the sequence -Asn-Lys-Ile-Asp-, is responsible for guanine recognition by hydrogen bonding from Asn-498 to the keto group of the base and from Asp-501 to the amino group.

The deletions and insertions of amino acid residues of IF2 compared with EF-Tu (see Figure 6) were generally found at the termini of secondary structural elements, close to positions that join these with loop connections. The modeling consisted mainly in the shortening or lengthening of these elements. For instance, the absence of IF2 residues corresponding to residues 159–160 in EF-Tu simply meant shortening the α-helix D by two residues, thus making a "shortcut" from residues 158 to 161. The subsequent positions of these two residues were hardly affected by this maneuver.

Discussion

The work reported here establishes that the major active centers for IF2 reside in the C-terminal two-thirds of the protein. The similar specific activities for IF2α and IF2γ in the assays for fMet-tRNA$_f^{Met}$ binding to 70S ribosomes, for dipeptide synthesis, and for GTP hydrolysis indicate that this C-terminal protein is responsible for the interaction of IF2 with fMet-tRNA$_f^{Met}$, ribosomes, and GTP. The results also suggest that the N-terminal third of the molecule does not contribute critically to these interactions but does confer stability to IF2α. The functional role of the highly charged and unusual repeating structures (residues 104–155 and 167–214) that are found in IF2α and some of which are found in IF2β (but that are entirely absent in IF2γ) is not known, but they may take part in pathways other than the initiation of protein synthesis. IF2 has been implicated in transcription (Travers et al., 1980) and secretion (Shiba et al., 1986), although these roles for IF2 are putative and have not been established definitively.

The three-dimensional model of the GTP-binding region of IF2, for which we propose the name IF2 G-domain, is closely related to the structure of domain 1 of elongation factor Tu. It is a globular domain containing 150 amino acids from the central part of the sequence of IF2α. Its mass is approximately 17 kDa, corresponding to one-fourth the size of IF2γ. It is a rigid structure of diameter 35 Å, having a central six-strand parallel β-pleated sheet with right-handed helical connections. The IF2 G-domain has a deep cleft, at the edge of which the γ-phosphate of a GTP molecule would be positioned. The model has been developed on the arguments of sequence homologies with elongation factor Tu. Even if this homology is rather weak in certain regions of the domain, the important requirement for globularity in proteins is fulfilled, namely, that the charged residues be on the surface of the protein and the hydrophobic residues mainly in the interior.

In this central region of IF2α, homology also exists with other guanine nucleotide binding proteins, such as prokaryotic elongation factor G, (Sacerdot et al., 1984; March & Inouye, 1985), eukaryotic elongation factor 2 (Kohno et al., 1986), LepA (March & Inouye, 1985), and the *ras* gene products (Halliday, 1984; Leberman & Enger, 1984; Möller & Amons, 1985; Kohno et al., 1986; March & Inouye, 1985). It thus seems plausible that a common basic folding pattern exists. However, more detailed crystallographic and other physicochemical studies are essential to elucidate the exact structure of the domain.

In addition to the sequence homologies between EF-Tu and IF2, functional similarities reinforce the idea that the two factors share structural features. Both factors bind GTP or GDP and interact with tRNAs. When they are in the GTP-binding conformation, both interact with ribosomes and hydrolyze GTP to GDP and inorganic phosphate. The subsequent factor–GDP complexes bind to ribosomes less tightly, resulting in release of the factors. In addition, ribosome binding and GTPase activities of IF2 and EF-Tu require the presence of ribosomal proteins L7/L12 and are inhibited by thiostreptone, suggesting similar sites of action on the ribosomes.

We have compared IF2α and IF2γ with EF-Tu by different immunoelectrophoretic techniques using rabbit antibodies against IF2α and EF-Tu. No cross-reaction between the initiation factors and the elongation factor was observed (results not shown). However, it is well-known that regions of high amino acid sequence homology are normally found buried in the tertiary structure of the protein molecules and thus would not be expected to exert strong immunogenic activity. Thus, although we suggest a tertiary structure model for the G-binding site of IF2 very similar to that of EF-Tu, it is not surprising that the two translation factors do not cross-react immunologically.

Attempts are now in progress to isolate small fragments of IF2 containing the G-domain. It is hoped that such fragments can be obtained in a form conserving biological activity and possessing sufficient stability for crystallization experiments. Furthermore, our three-dimensional model of the IF2 G-domain can serve as an important clue in future protein engineering experiments on IF2.

Acknowledgments

We thank H. Worm-Leonhard for technical assistance, H. O. Sperling-Petersen, A. Lindahl, and O. Jensen for preparing the drawings and photographs, and L. Heilesen for help in the preparation of the manuscript.

References

Beaudry, P., Sander, G., Grunberg-Manago, M., & Douzou, P. (1979) *Biochemistry 18*, 202–207.

Cenatiempo, Y., Robakis, N., Reid, B. R., Weissbach, H., & Brot, N. (1982) *Arch. Biochem. Biophys. 218*, 572–578.

Chou, P. Y., & Fasman, G. D. (1978) *Annu. Rev. Biochem. 47*, 251–276.

Clark, B. F. C., la Cour, T. F. M., Nielsen, K. M., Nyborg, J., Petersen, H. U., Siboska, G. E., & Wikman, F. P. (1984) in *Gene Expression. The Translational Step and its Control* (Clark, B. F. C., & Petersen, H. U., Eds.) pp 127–145, Munksgaard, Copenhagen.

Dondon, J., Plumbridge, J. A., Hershey, J. W. B., & Grunberg-Manago, M. (1985) *Biochimie 67*, 643–649.

Garnier, J., Osguthorpe, G. J., & Robson, B. (1978) *J. Mol. Biol. 120*, 97–120.

Grunberg-Manago, M. (1980) in *Ribosomes: Structure, Function and Genetics* (Chambliss, G., Craven, G. R., Davies, J., Davis, K., Kahan, L., & Nomura, M., Eds.) pp 445–477, University Park Press, Baltimore.

Halliday, K. (1984) *J. Cyclic Nucleotide Protein Phosphorylation Res. 9*, 435–448.

Høiby, N., & Axelsen, N. H. (1983) *Scand. J. Immunol., Suppl. No. 10*, 125–134.

Howe, J. G., & Hershey, J. W. B. (1982) *Arch. Biochem. Biophys. 214*, 446–451.

Hunkapiller, M. W., & Hood, L. E. (1983) *Methods Enzymol. 91*, 486–493.

Hunkapiller, M. W., Hewick, R. M., Dreyer, W. J., & Hood, L. E. (1983) *Methods Enzymol. 91*, 399–413.
Jack, A., & Levitt, M. (1978) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A34*, 931–935.
Jones, T. A., & Thirup, S. (1986) *EMBO J. 5*, 819–822.
Jurnak, F. (1985) *Science (Washington, D.C.) 230*, 32–36.
Kohno, K., Uchida, T., Ohkubo, H., Nakanishi, S., Nakanishi, T., Fukui, T., Ohtsuka, E., Ikehara, M., & Okada, Y. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 4978–4982.
Kolakofsky, D., Dewey, K. F., Herskey, J. W. B., & Thach, R. E. (1968) *Proc. Natl. Acad. Sci. U.S.A. 61*, 1066–1071.
Krøll, J. (1983) *Scand. J. Immunol., Suppl. No. 10*, 171–173.
la Cour, T. F. M., Nyborg, J., Thirup, S., & Clark, B. F. C. (1985) *EMBO J. 4*, 2385–2388.
Leberman, R., & Egner, U. (1984) *EMBO J. 3*, 339–341.
Maitra, U., Stringer, E. A., & Chaudhuri, A. (1982) *Annu. Rev. Biochem. 51*, 869–900.
March, P. E., & Inouye, M. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7500–7504.
McCormick, F., Clark, B. F. C., la Cour, T. F. M., Kjeldgaard, M., Nørskov-Lauritsen, L., & Nyborg, J. (1985) *Science (Washington, D.C.) 230*, 78–82.
Möller, W., & Amons, R. (1985) *FEBS Lett. 186*, 1–7.
Morikawa, K., la Cour, T. F. M., Nyborg, J., Rasmussen, K. M., Miller, D. L., & Clark, B. F. C. (1978) *J. Mol. Biol. 125*, 325–338.
Peacock, S., Cenatiempo, Y., Robakis, N., Brot, N., & Weissbach, H. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4609–4612.
Petersen, H. U. (1985) *Mat. Fys. Medd. K. Dan. Vidensk. Selsk. 41*, 291–335.
Plumbridge, J. A., Howe, J. G., Springer, M., Touati-Schwartz, D., Hershey, J. W. B., & Grunberg-Manago, M. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 5033–5037.
Plumbridge, J. A., Deville, F., Sacerdot, C., Petersen, H. U., Cenatiempo, Y., Cozzone, A., Grunberg-Manago, M., & Hershey, J. W. B. (1985) *EMBO J. 4*, 223–229.
Robakis, N., Meza-Basso, L., Brot, N., & Weissbach, H. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4261–4264.
Robakis, N., Cenatiempo, Y., Meza-Basso, L., Brot, N., & Weissbach, H. (1983) *Methods Enzymol. 101*, 690–706.
Sacerdot, C., Dessen, P., Hershey, J. W. B., Plumbridge, J. A., & Grunberg-Manago, M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7787–7791.
Shiba, K., Ito, K., Nakamura, Y., Dondon, J., & Grunberg-Manago, M. (1986) *EMBO J. 5*, 3001–3006.
Travers, A. A., Debenham, P. G., & Pongs, O. (1980) *Biochemistry 19*, 1651–1656.
Weissbach, H., Robakis, N., Cenatiempo, Y., & Brot, N. (1984) *Biotechniques 2*, 16–22.
Wierenga, R. K., & Hol, W. G. J. (1983) *Nature (London) 302*, 842–844.

Solution Structure of the Trp Operator of *Escherichia coli* Determined by NMR†

Jean-Francois Lefèvre,‡ Andrew N. Lane,§ and Oleg Jardetzky*
Stanford Magnetic Resonance Laboratory, Stanford University, Stanford, California 94305-5055
Received November 18, 1986; Revised Manuscript Received March 9, 1987

ABSTRACT: We have assigned the majority of the nonexchangeable protons in the NMR spectrum of the 20 base-pair fragment of DNA corresponding to the Trp operator of *Escherichia coli*. The sequence (CGTACTAGTTAACTAGTACG) also contains a Pribnow box (underlined). Variation of the intrinsic spin–lattice relaxation rate constants of the H8's along the sequence indicates that the structure of the oligonucleotide is not regular. Splitting patterns of the H1′ resonances in the deoxyriboses, obtained from a two-dimensional *J*-resolved experiment, allowed the dominant pucker mode of each nucleotide to be determined. Intranucleotide NOEs from the sugar protons H1′, H2′, and H3′ to the base protons were used to determine the conformation of each nucleotide (puckers and glycosidic torsion angles). The relative orientations of nucleotide units (roll, propeller twist, helical twist angle, and pitch) were calculated by using internucleotide NOEs between protons of neighboring nucleotides in the sequence. All these parameters were determined for each step along the 20-mer. The structure belongs to the B family of conformations, but variations of the local geometry are observed from step to step. Some of the variations, such as the roll and the twist angles, can be predicted by the rules of Calladine and Dickerson [Calladine, C. R., & Dickerson, R. E. (1983) *J. Mol. Biol. 166*, 419–441]. The puckers of the deoxyriboses of purines are found mainly in conformations near C2′ endo, while those of the pyrimidines prefer C3′ endo and related conformations. Glycosidic torsion angles obtained for purines are larger than those of pyrimidines. Except for this last observation, the general properties of the operator DNA structure are comparable with those of crystal structures of B DNA of other sequences.

The control of gene expression by regulatory proteins that bind to specific sequences of DNA with affinities an order of magnitude greater than their affinity for bulk DNA must involve highly stereospecific interactions. Knowledge of the detailed structure of these specific sequences, such as the operator sequences binding repressors, is therefore important for the understanding of these interactions and the structural changes involved in the action of corepressors and inducers.

As part of an NMR study (Lane & Jardetzky, 1985a–c; Lefèvre et al., 1985a,b; Lane et al., 1986a,b; Lane, 1986) of the operator–repressor interaction in the Trp operator system

† This work was supported by NIH Grant RR00711 and NIGMS Grant GM33385. A.N.L. is grateful to EMBO for a long-term fellowship and J.F.L. to NATO and the Philippe Foundation for fellowships.
* Author to whom correspondence should be addressed.
‡ Present address: Institut de Biologie Moleculaire et Cellulaire, 67000 Strasbourg, France.
§ Present address: National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, U.K.

0006-2960/87/0426-5076$01.50/0 © 1987 American Chemical Society