The EMBO Journal Vol.16 No.6 pp.1436–1443, 1997

The structure of the translational initiation factor IF1 from *E.coli* contains an oligomer-binding motif

Marco Sette, Paul van Tilborg, Roberto Spurio[1], Robert Kaptein, Maurizio Paci[2], Claudio O.Gualerzi[1] and Rolf Boelens[3]

Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands, [1]Laboratorio di Genetica, Dipartimento di Biologia MCA, Università degli Studi di Camerino, via Camerini, 62032 Camerino and [2]Dipartimento di Scienze e Tecnologie Chimiche, Università degli Studi di Roma 'Tor Vergata', via della Ricerca Scientifica, 00133 Rome, Italy

[3]Corresponding author

The structure of the translational initiation factor IF1 from *Escherichia coli* has been determined with multidimensional NMR spectroscopy. Using 1041 distance and 78 dihedral constraints, 40 distance geometry structures were calculated, which were refined by restrained molecular dynamics. From this set, 19 structures were selected, having low constraint energy and few constraint violations. The ensemble of 19 structures displays a root-mean-square deviation versus the average of 0.49 Å for the backbone atoms and 1.12 Å for all atoms for residues 6–36 and 46–67. The structure of IF1 is characterized by a five-stranded β-barrel. The loop connecting strands three and four contains a short 3_{10} helix but this region shows considerably higher flexibility than the β-barrel. The fold of IF1 is very similar to that found in the bacterial cold shock proteins CspA and CspB, the N-terminal domain of aspartyl-tRNA synthetase and the staphylococcal nuclease, and can be identified as the oligomer-binding motif. Several proteins of this family are nucleic acid-binding proteins. This suggests that IF1 plays its role in the initiation of protein synthesis by nucleic acid interactions. Specific changes of NMR signals of IF1 upon titration with 30S ribosomal subunit identifies several residues that are involved in the interaction with ribosomes.

Keywords: NMR spectroscopy/protein–RNA interactions/protein structure/protein synthesis/ribosome binding

Introduction

Initiation of protein biosynthesis consists of a number of interrelated steps preceding the formation of the first peptide bond. In *Escherichia coli*, in addition to mRNA, fMet-tRNA and ribosomal subunits, the initiation mechanism requires the presence of three additional proteins, the initiation factors IF1, IF2 and IF3, and at least one GTP molecule. The three initiation factors both influence the kinetics of formation of the ternary complex, increasing the rate of the process by over two orders of magnitude, and enhance the stability of the ternary complex (for reviews, see Hershey, 1987; Gualerzi and Pon, 1990).

Initiation factor IF1, the smallest of the three factors, is a monomeric protein consisting of 71 residues (Pon *et al.*, 1979). IF1 is a highly conserved element of the prokaryotic translational apparatus and has been found also in the chloroplasts of several plants. The gene encoding IF1 in *E.coli*, *infA*, has been identified and cloned (Sands *et al.*, 1987). The expression of *infA* in *E.coli* is under growth-rate control (Cummings *et al.*, 1991) and its inactivation cannot be tolerated by the cell without loss of viability (Cummings and Hershey, 1994), indicating that one or more of the activities of IF1 must be essential for the cell.

Several functions have been reported for IF1. These include: (i) the enhancement of the rate of 70S ribosome dissociation and subunit association (Godefroy-Colburn *et al.*, 1975); (ii) the stimulation of the activity of IF2 and IF3 in the formation of the 30S initiation complex (Wintermeyer and Gualerzi, 1983; Pon and Gualerzi, 1984); and (iii) the modulation of the interaction of IF2 with the ribosome, increasing its affinity for the 30S ribosomal subunit when IF1 is bound and indirectly favouring its release when IF1 is ejected (Stringer *et al.*, 1977; Celano *et al.*, 1988). In addition, by binding to the A-site of the 30S ribosomal subunit, IF1 may contribute to the fidelity of the selection of the initiation site of the mRNA (Moazed *et al.*, 1995).

Equilibrium binding studies have shown that IF1 binds to the 30S ribosomal subunit in a 1:1 ratio (Zucker and Hershey, 1986; Celano *et al.*, 1988). The binding affinity depends strongly upon the ionic strength and upon the presence of IF2 and IF3 which increase its affinity for the ribosome; the K_a ranging from 5×10^5 to 2.5×10^8 M^{-1}. Interaction with 50S ribosomal subunits was also observed, but the affinity is considerably lower than for the 30S ribosomal subunits. Stable interaction with 70S ribosomes has never been observed; in fact, the addition of 50S to 30S subunits bearing IF1 promoted the release of the factor, suggesting that recycling occurs when the two subunits join to form a 70S initiation complex (Celano *et al.*, 1988).

There are only a limited number of structural studies performed on IF1. Secondary structure predictions and circular dichroism (CD) measurements indicated a very low content (5–12%) of α-helix, and a variable value (7–50%) for β-sheet content (Pawlik *et al.*, 1981). A number of amino acids crucial for the function of IF1 were also identified by site-directed mutagenesis (Gualerzi *et al.*, 1989; Spurio *et al.*, 1991).

Here we present the structure determination of IF1, using multidimensional nuclear magnetic resonance (NMR) spectroscopy, as well as the dynamic behaviour of the protein. It is found that the structure of IF1 is very

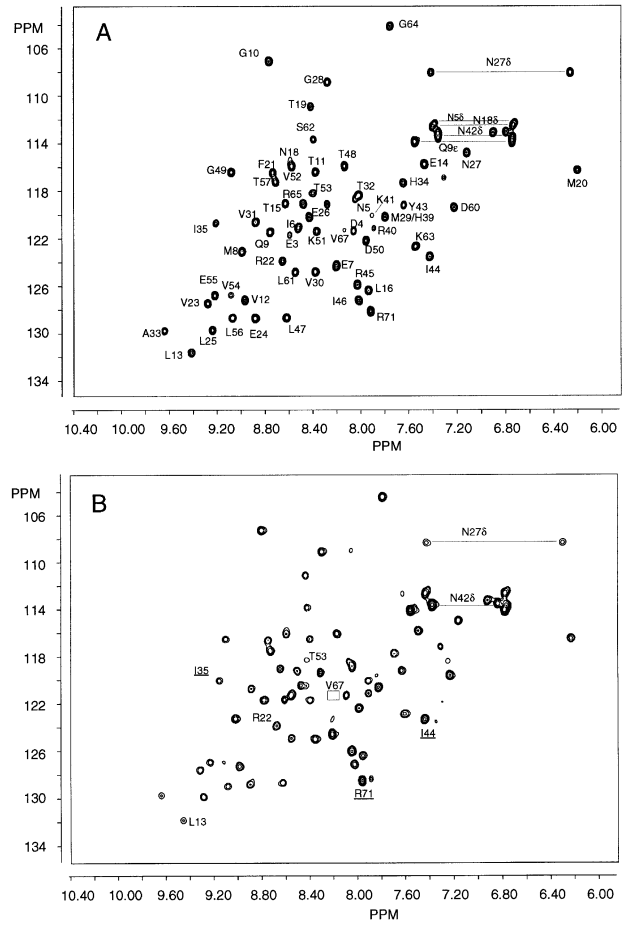

Fig. 1. 2D (^{15}N,^{1}H)-HSQC spectra of initiation factor IF1 from *E.coli* in the absence (**A**) and presence (**B**) of 30S ribosomal subunits. For (A) the sample contained 200 mM KCl. For (B) the sample contained 500 mM KCl and 30S ribosomal subunits in a molar ratio of 30S to IF1 of 7.5×10^{-3}. Since both spectra were recorded at a pH of 6.5 to avoid denaturation of the ribosomes, not all amide protons are visible in (A). Underlined residues have a different chemical shift in the presence of ribosomes; the others are selectively broadened. A box indicates the resonance of Val67, the signal of which disappears in the presence of ribosomes.

similar to nucleic acid-binding proteins of the oligomer-binding (OB) fold family, suggesting that IF1 interacts with the ribosomes through nucleic acid interactions as well. Several residues which are involved in the binding to the 30S ribosomal subunit have been identified.

Results

For the NMR assignments of initiation factor IF1, we made use of ^{15}N-labelled material. A 2D (^{15}N,^{1}H)-HSQC spectrum is shown in Figure 1A. The sequential assignment procedure was performed according to standard methods (Wüthrich, 1986). All the spin systems were obtained from TOCSY spectra and most sequential connectivities using a NOESY spectrum with a short mixing time. A 3D NOESY-(^{15}N,^{1}H)-HSQC spectrum of ^{15}N-labelled IF1 was used for confirmation of the assignment and for resolving ambiguities due to spectral overlap. Finally, a 3D HNHA spectrum was used to check the assignments of the resonances, in particular of threonines. Most of the residues were assigned completely. The resonance assignments of IF1 at pH 5.1 are available as supplementary material.

The structure determination of IF1 is based on NOE data as well as $^{3}J_{HNH\alpha}$ and $^{3}J_{H\alpha H\beta}$ coupling constants. A total of 1041 NOEs was obtained from analysis of the NOESY spectra with different mixing times. These spectra contain many long-range NOEs, characteristic for a β-sheet structure. The existence of several β-strands was confirmed by the large $^{3}J_{HNH\alpha}$ coupling constants as observed in the 3D HNHA spectrum. Figure 2A shows the distribution of NOEs over the sequence of IF1.

Using distance geometry followed by restrained molecular dynamics, 40 structures were calculated, 19 of which were selected on the basis of low constraint energy and few constraint violations. Figure 3B shows the overlay of these structures. The root-mean-square deviation (r.m.s.d.) versus the average for the well-defined region (residues 6–36 and 46–67) is 0.49 Å for the backbone atoms and 1.12 Å for all atoms. A summary of structural statistics is given in Table I and Figure 2B–E.

The solution structure of IF1 consists of a five-stranded β-sheet, arranged as a β-barrel. The five β-strands, as shown in Figure 3A, range from residues 6 to 16, 19 to 26, 29 to 36, 50 to 56 and 62 to 66, respectively, based on the long-range NOE data. The first strand contains a β-bulge at residues 13 and 14. All strands are oriented anti-parallel, except strands III and V. While the turns connecting strands I and II (type IV turn) and strands II and III (type I turn) seem well defined, this is not the case for the region connecting strands III and IV (residues 37–49) and the short loop connecting strands IV and V (residues 57–62). Medium-range NOEs in the region 38–44 indicate the presence of a 3_{10}-helical structure. Heteronuclear (^{15}N,^{1}H)-NOE data show that the loop from residue 36 to residue 49 and both the N- and C-terminal regions have considerable flexibility on the nanosecond time scale (Figure 2F). The core of the protein is remarkably rigid. The coordinates of the structures and the NMR data will be deposited at the protein data bank in Brookhaven.

Using the DALI Server software (Holm and Sander, 1993), the fold of IF1 was identified as the OB fold. Proteins belonging to this family all bind polysaccharides or polynucleotides (Murzin, 1993). Figure 4 shows the similarity in topology between IF1, the cold shock protein A (CspA), the N-terminal domain of aspartyl-tRNA synthetase and staphylococcal nuclease. The major difference between these proteins seems to be the relative orientation of the five β-strands and in the length of the loop connecting strands three and four. Interestingly, except for CspA, a helix is present in this loop, as was also found with IF1. On the other hand, the shear number for the β-barrel of IF1 is S = 8, similar to CspA but different from the other two proteins (S = 10). These proteins homologous to IF1 all bind single-stranded nucleic acids.

The highest degree of structural homology is observed with the *E.coli* CspA. The r.m.s.d. along the Cα atoms of residues 6–36, 46–56, 62–65 and 67–70 of IF1 with the corresponding parts of CspA is 2.66 Å. CspA (also called CS7.4) is the product of the cold shock gene *cspA*, and has a 61% identity to CspB, the equivalent protein found in *Bacillus subtilis*. Cold shock proteins were found to

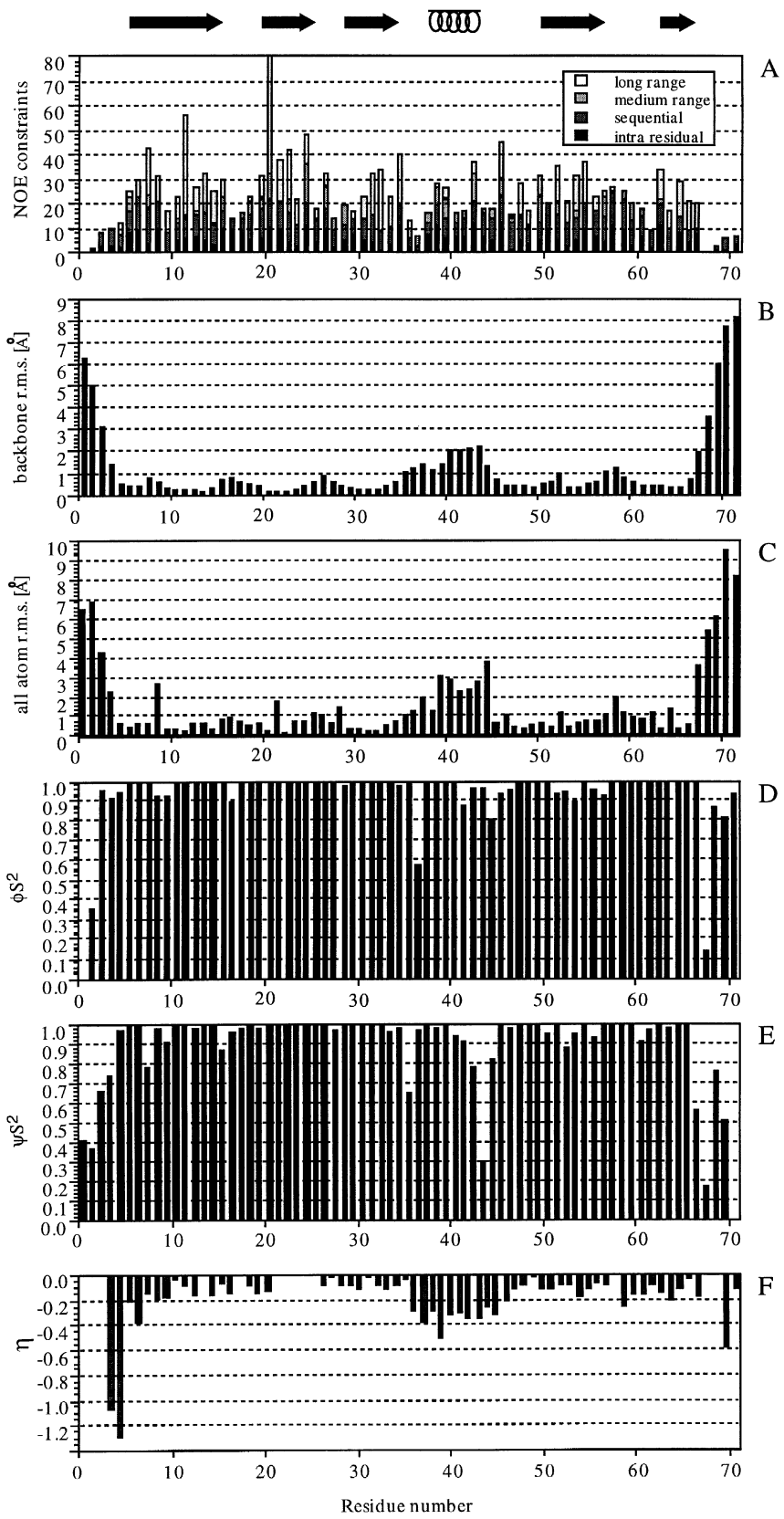

Fig. 2. Overview of structural parameters of IF1. (**A**) Number of distance constraints per residue. The constraints are classified in long-range, medium-range, sequential and intra residual; (**B**) and (**C**) r.m.s.d. for the backbone Cα, N, CO atoms and all atoms versus the residue number, respectively. The structures are superimposed on the backbone atoms of residues 5–35 and 46–67; (**D**) the φ angular order parameter; (**E**) the ψ angular order parameter; (**F**) for the results of the (^{1}H-^{15}N) NOE, values of η (according to the definitions given by Peng and Wagner, 1992) versus the residue number are given. The secondary structure is indicated at the top of (A).

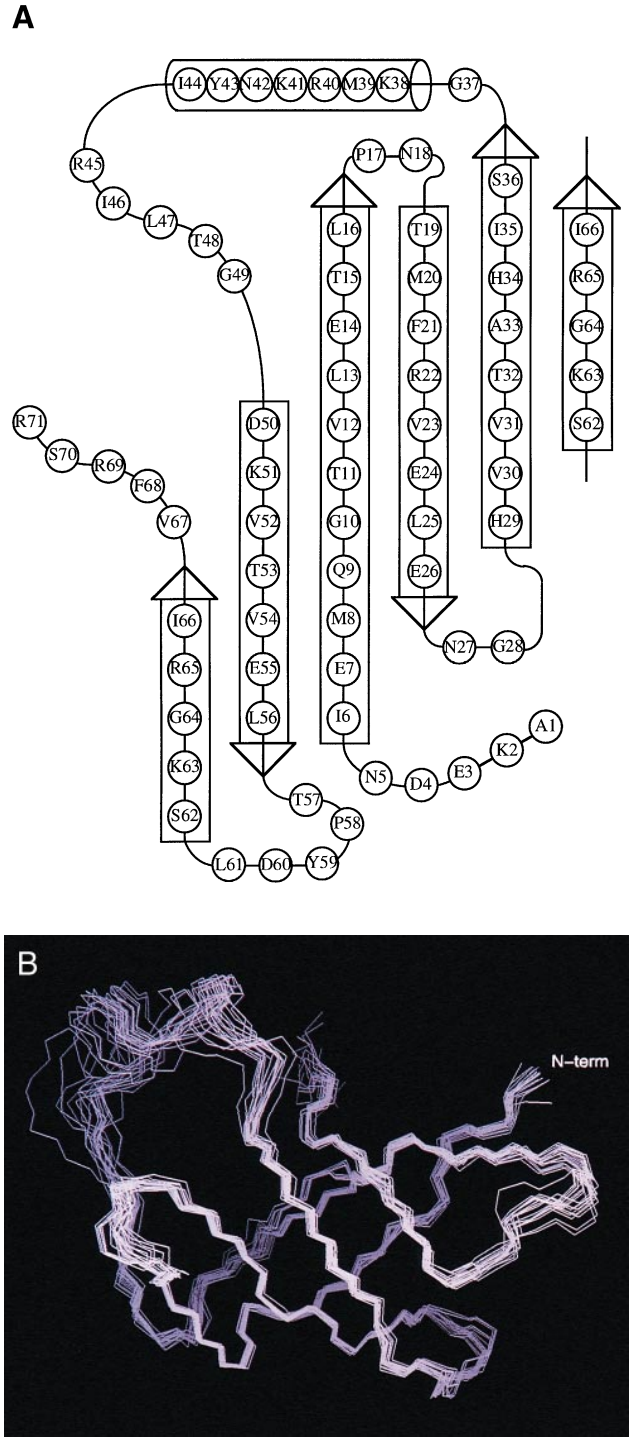

Fig. 3. Folding topology (**A**) and 3D structure (**B**) of initiation factor IF1 from *E.coli*. (A) Arrows and cylinders indicate the β-strands and 3_{10} helix of IF1, respectively. The fifth strand is shown twice to indicate the barrel shape. The set of 19 structures superimposed on the backbone Cα, N, CO atoms of residues 6–35 and 46–67. The first four and the last five residues are omitted from the figure because they are severely disordered.

play a role in improving cell viability after cold shock (Willimsky *et al.*, 1992). These cold shock proteins have nucleic acid-binding properties (both RNA and DNA, both single- and double-stranded) and have sequence homology with eukaryotic Y-box-binding proteins (Wolffe *et al.*,

Table I. Structural statistics for IF1 solution structure

No. of distance constraint violations >0.35 Å	2
Maximum distance constraint violation (Å)	0.64
R.m.s. deviation from average structure (Å)[a]	
Backbone (N,Cα,CO)	0.49 ± 0.16
All heavy atoms	1.12 ± 0.25
Average No. of bad non-covalent contacts[b]	1
Percentage of residues[b,c] with ϕ/ψ in:	
Most favoured regions	60.4
Additional allowed regions	36.1

	Values in IF1	Ideal values[d]
χ1 (g–)	58.2 ± 21.4	64.1 ± 15.7
χ1 (t)	189.7 ± 16.6	183.6 ± 16.8
χ1 (g+)	–68.2 ± 18.6	–66.7 ± 15.0
χ2 (t)	174.7 ± 16.6	177.4 ± 18.5
ω	179.1 ± 9.0	180.0 ± 5.8
Chirality Cα	33.0 ± 3.2	33.9 ± 3.5

[a]Residues 6–36 and 46–67.
[b]Calculated from the final set of structures with PROCHECK; g(–), gauche minus; t, trans; g(+), gauche plus.
[c]Residues excluding glycine and proline.
[d]Ideal values from Morris *et al.* (1992; PROCHECK).

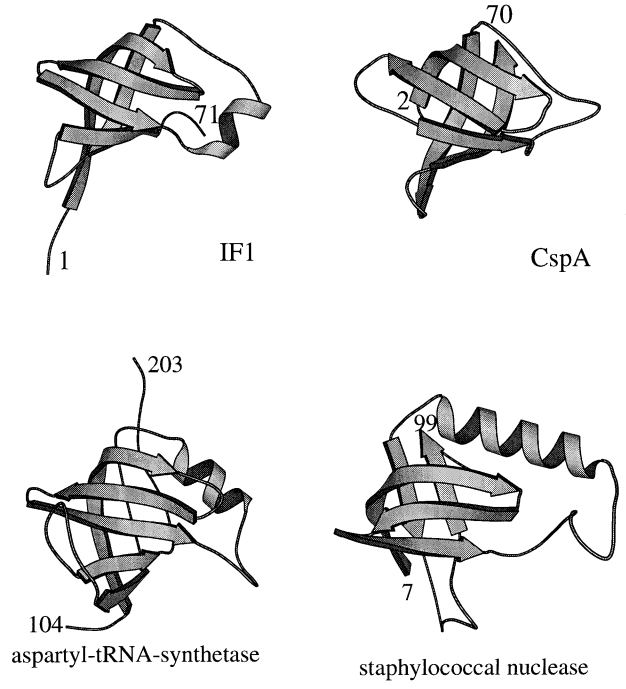

Fig. 4. Comparison of IF1 with CspA and the corresponding parts of the N-terminal domain of aspartyl-tRNA synthetase and staphylococcal nuclease. The coordinates of CspA, the N-terminal domain of aspartyl-tRNA synthetase and staphylococcal nuclease were obtained from the Protein Data Bank in Brookhaven. The structures were generated with the Molscript software (Kraulis, 1991).

1992; Lee *et al.*, 1994). CspA was reported to have an RNP1 and a rudimentary RNP2 motif (Schindelin *et al.*, 1994). For IF1, no RNP1 region could be located, but a rudimentary RNP2 region similar to CspA was found at residues 31–36.

A titration of 30S ribosomal subunits to IF1 was monitored by a series of (^{15}N,^{1}H)-HSQC spectra. An example is given in Figure 1B. Selective broadening or shifting of individual cross-peaks was taken as a measure

of specific interaction. At a molar ratio of 7.5×10^{-3} 30S ribosomal subunits to IF1, all the signals disappeared, probably due to the high overall molecular weight of the complex. It should be pointed out that the molecular weight of the 30S particle alone would be ~900 000 Da. Subsequently, the salt concentration was increased in a number of steps, which allowed most signals to become visible again. Analysis of the relative changes of cross-peak intensities over all spectra of the salt titration revealed that some cross-peaks were broadened more significantly than others by ribosome binding. The resonances affected were those of the backbone amides of Thr53 and Val67 and the side chain amides of Asn42 and Asn27 and, to a lesser extent, those of the backbone amides of Leu13 and Arg22. Furthermore, the resonances of the ^{15}N backbone amides of Ile35, Ile44 and Arg71 are shifted $>$0.2 p.p.m. during the binding process. Also, the side chain resonances of His34 are strongly broadened in the presence of 30S ribosomal subunits while those of His29 are shifted to a large extent. The aforementioned residues are located over a large part of the protein, but all on one side. This indicates that IF1 uses a large surface for interacting with the 30S ribosomal subunits.

Discussion

We have reported here the elucidation of the three-dimensional (3D) structure of *E.coli* translational initiation factor IF1 in solution. The most relevant conclusion derived from this work is that IF1 displays a striking structural resemblance to proteins belonging to the OB family, a class of proteins of which many members interact with single-stranded nucleic acids. Particularly relevant among these are the major cold shock protein CspA (Schindelin *et al.*, 1993, 1994; Schnuchel *et al.*, 1993; Newkirk *et al.*, 1994), the N-terminal domain of aspartyl-tRNA synthetase (Ruff *et al.*, 1991) and the C-terminal domain of methionyl-tRNA^{fMet} transformylase (Schmitt *et al.*, 1996). These similarities may give important insight into the mechanism of action of IF1. Thus, the indication that IF1 binds to the 30S ribosomal subunit through an ionic interaction with 16S rRNA (Zucker and Hershey, 1986; Celano *et al.*, 1988) and the finding that IF1 protects specific bases of the 16S rRNA from chemical modification *in situ* (Moazed *et al.*, 1995) are most relevant in the context of this discussion. These facts, in addition to the data (summarized below) concerning the mechanism by which IF1 binds to the 30S ribosomal subunit, strongly suggest that the nucleic acid-binding surface of IF1 corresponds to an active site of this molecule involved in the interaction with an rRNA moiety of the 30S ribosomal subunit. The structural homology between IF1 and the β-barrel domains of aspartyl-tRNA synthetase and methionyl-tRNA transformylase, implicated in the interaction with the anticodon of the RNA and in the specific recognition of the acceptor stem of the initiator tRNA respectively, suggests that a specific interaction with the initiator tRNA might play a role in the mechanism of action of IF1. It should be noted, however, that, unlike the other structurally homologous proteins which have a well documented nucleic acid-binding activity, evidence for extra-ribosomal interactions of nucleic acids with IF1 is weak. In fact, we are only aware of an early report demonstrating that IF1 can bind to several single- and double-stranded polynucleotides and induce alterations of their secondary structures; thus, IF1 was found to denature the helical hairpin form of poly(U), the single-stranded stacked form of poly(C) and the poly(A)–poly(U) duplex and to alter the structure of double-stranded poly(C) and single- and double-stranded poly(A) (Schleich *et al.*, 1980). Although the significance of these interactions remains as yet unclear, it is possible to imagine that one or more of these activities might ensue upon binding of IF1 to its RNA target, be it a site on the 16S rRNA or a specific portion of the initiator tRNA molecule. Another potential target of the IF1 nucleic acid-binding surface may also be the ribosome-bound mRNA. In this context, it should be recalled that CspA was found to affect not only transcription, acting as a transcriptional enhancer of cold shock genes (La Teana *et al.*, 1991; Jones *et al.*, 1992), but also protein synthesis by stimulating translation of its own mRNA (Brandi *et al.*, 1996).

We have also described experiments aimed at identifying the specific amino acid residues of IF1 involved in the interaction with the 30S ribosomal subunit. Similar attempts had been carried out in the past (Paci *et al.*, 1983). Upon gradual addition of increasing (yet sub-stoichiometric) amounts of deuterated 30S ribosomal subunits to IF1, intermediate exchange dynamics between free and ribosome-bound factor caused selective line broadening and chemical shift changes against the background of its gradually disappearing ^{1}H-NMR spectrum. From the present results, it is suggested that the following residues are involved in or affected by the binding of IF1 to the 30S ribosomal subunit: Leu13, Arg22, Asn27, His29, His34, Ile35, Asn42, Ile44, Thr53, Val67 and Arg71. For His29 and His34 this corresponds to the results previously described by Paci *et al.* (1983).

Subsequently, several amino acid residues were analysed by mutagenesis (Spurio *et al.* 1991). It was found that a short form of IF1 lacking the C-terminal residues Ser70 and Arg71 has the same ability as native IF1 to form a complex with 30S ribosomal subunits and to stimulate the binding of fMet-tRNA to the ribosomes. Both activities are lost in a protein lacking residues 69–71, indicating that Arg69 is important for the functionality of the protein. The two histidines (His29 and His34) were also mutated separately and it was observed that replacement of His34 with Asp (but not with Tyr) severely impaired the binding of IF1 to the 30S ribosomal subunit while His29→Asp mutation allowed substantial binding to the 30S ribosomal subunit but produced complexes which were essentially inactive. Furthermore, His29 mutants proved to be defective in the ejection of IF1 from the ribosome during 30S–50S association (Gualerzi *et al.*, 1989). Finally, it was found that Lys38, although in itself not essential (substitution with Ile or Thr produced no loss of activity), could not be conservatively replaced by the bulkier Arg residue without loss of activity (Spurio *et al.*, 1991). Taken together, the previous mutagenesis data and the results obtained in the present study on the binding of IF1 to the 30S ribosomal subunit are in good agreement. The only discrepancy seems to be R71, for which we observed a change in chemical shift. Although R71 may be involved in ribosome binding, this interaction is not vital for the activity of IF1. The residues affected by ribosome binding

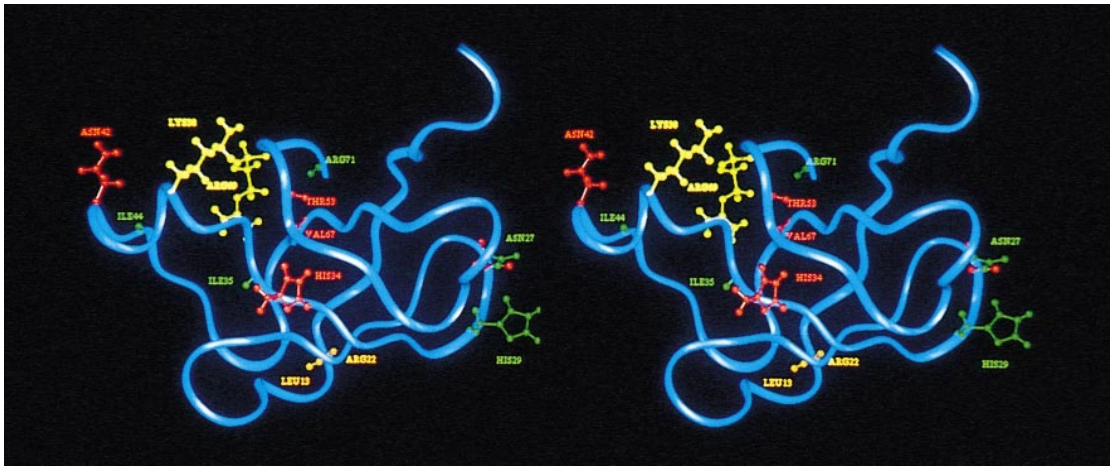

Fig. 5. Stereo drawing of IF1 indicating residues involved in ribosomal interaction. The figure shows those residues for which signals broaden upon binding to 30S subunits in red and those residues for which signals display a significant shift in green. Arg69 and Lys38, which were shown by mutagenesis to affect ribosome binding, are shown in yellow.

are highlighted in the 3D structure presented in Figure 5. As seen from the figure, all the residues implicated so far in this activity are exposed to the solvent and located on the same side of the protein.

In conclusion, although the function of IF1 has been unclear until now, it has been found that the fold of IF1 resembles that of several nucleic acid-binding proteins with an OB fold. NMR titration studies with 30S ribosomal subunits show a number of interactions involving specific residues. Combined with previous biochemical data, this strongly indicates that IF1 performs its function by direct interaction with RNA in the ribosomal complex. Recent work by Moazed *et al.* (1995) shows that this may be 16S rRNA.

Materials and methods

Purification of ^{15}N-labelled IF1

The assembly of the synthetic modular *infA** gene encoding *E.coli* translational initiation factor IF1, the construction of the hyperexpressing vector (pXR201) and the properties of the product have been described previously (Calogero *et al.*, 1987). *Escherichia coli* JM109 transformed with pXR201 carrying *infA** under the control of the λpL promoter (Calogero *et al.*, 1987) and with pcI857 expressing a ts λcI repressor, was grown overnight at 30°C in 500 ml of modified M9 minimal medium containing: (4.4 g/l) KH_2PO_4; (10.4 g/l) $K_2HPO_4{\cdot}3H_2O$; (50 mg/l) $MgSO_4{\cdot}7H_2O$; (7 mg/l) $(NH_4)2SO_4FeSO_4{\cdot}6H_2O$; (5 g/l) glucose; (10 mg/l) thiamine; (60 mg/l) ampicillin; (25 mg/l) kanamycin and 0.5 g/l of $^{15}NH_4Cl$ (ISOTEC Inc.). This culture was used to inoculate 6.5 l of the same medium and growth was continued at the same temperature to an $A_{600\ nm}$ = 1.95 when IF1 hyperproduction was induced by shifting the culture to 42°C by addition of 3 l of the same medium pre-warmed to 70°C. After 20 min at 42°C, the temperature of the culture was shifted to 37°C and the incubation continued for 1 h, at which time the cells were harvested by centrifugation. Sixty grammes of labelled cells were lysed by sonification and treated with 10 g/l of polyethyleneimine. The supernatant containing IF1, obtained after centrifugation at 10 000 r.p.m. for 30 min, was dialysed against buffer A [20 mM Tris–HCl pH 7.1; 10% glycerol; 1 mM EDTA; 0.1 M NH_4Cl; 6 mM β-mercaptoethanol; 0.2 mM phenylmethylsulfonyl fluoride (PMSF)] and loaded onto a phosphocellulose column (3×20 cm) equilibrated with the same buffer. Elution was performed with a 0.1–0.7 M NH_4Cl linear gradient in buffer A, and fractions containing IF1, as monitored by SDS–PAGE, were pooled and concentrated by batch elution with buffer A containing 1 M NH_4Cl from a small phosphocellulose column (2.5×3 cm). The sample was then loaded onto a Superdex^{75} column (Pharmacia; HiLoad 26/60) equilibrated with buffer B (20 mM K phosphate buffer pH 7.6; 0.15 M NaCl; 1 mM dithiothreitol). Following gel filtration, IF1 was >95% pure, as judged by SDS–PAGE. The final yield was ~2.0 mg/l of culture. Uniformly ^{15}N-labelled IF1 was concentrated and resuspended in a buffer containing 200 mM KCl and 20 mM KH_2PO_4 with the pH adjusted to 5.1 using an AMICON stirred cell. The final sample consists of 3.5 mM in protein, as determined by UV spectroscopy (ε_{280} = 2560 M^{-1} cm^{-1}). A 1 mM sample of IF1, in the same buffer plus 5 mM $MgCl_2$, was used for the ribosome binding tests: in this case the final pH was 6.5.

NMR spectroscopy

Two-dimensional NMR spectra were recorded on a Varian Unity Plus 750 MHz spectrometer. The experiments were recorded at 300 K using a spectral width of 16.7 p.p.m. and in 95% H_2O–5% D_2O, pH 5.1. Proton chemical shifts are referred to the water resonance at 4.72 p.p.m.; nitrogen chemical shifts were calculated by indirect referencing (Wishart *et al.*, 1995). Clean-TOCSY spectra (Griesinger *et al.*, 1988) were recorded with mixing times of 30, 50 and 90 ms. NOESY spectra (Jeener *et al.*, 1982) were recorded with mixing times of 50, 100 and 150 ms. A 2D rotating frame Overhauser enhancement spectroscopy (ROESY) spectrum (Bothner-By *et al.*, 1984) with a 50 ms mixing time was also recorded. Six hundred t_1 increments typically were recorded for homonuclear experiments. These experiments were processed using shifted sine-bell or shifted squared sine-bell window functions in both domains. Zero-filling in t_1 was applied to obtain a final matrix of 2048×2048 data points. Processing was carried out on Silicon Graphics INDY and INDIGOII workstations using the TRITON software.

Most heteronuclear 3D experiments were recorded on a Bruker AMXT 600 MHz (equipped with a triple resonance probe). A spectral width of 33 p.p.m. was used for ^{15}N, thus folding the Nε signals of arginines. 3D TOCSY-HSQC and 3D NOESY-HSQC spectra (Marion *et al.*, 1989) were recorded with mixing times of 50 and 120 ms, respectively. After linear prediction in t_2, the final matrix consisted of 1024×128×1024 data points. To determine the $^3J_{HNH\alpha}$ coupling constants, a 3D HNHA experiment (Vuister and Bax, 1993) was performed at 600 MHz. In addition, a 3D HMQC-NOESY-HSQC (Ikura *et al.*, 1990) was recorded at 750 MHz for identifying helical contacts.

Titration studies of 30S ribosomal subunits and IF1

The preparation of *E.coli* 30S ribosomal subunits and the experimental design to study the interaction of ^{15}N-labelled IF1 with the 30S ribosomal subunit were essentially as previously described (Paci *et al.*, 1983). Several 2D (^{15}N,^{1}H)-HSQC spectra (Bodenhausen and Ruben, 1980) were recorded (Δ = 5 ms) to follow the titration of 30S ribosomal subunits to IF1. A 2D (^{15}N,^{1}H)-HSQC spectrum with Δ = 22 ms was performed to observe the connectivities between the HD2 and the HE1 protons with the ND1 and NE2 nitrogen of the ring of the histidines. These spectra were recorded at pH 6.5 in order to prevent denaturation of the 30S ribosomal subunits.

Structure calculations

Distances were extracted from the NOESY with a mixing time of 100 ms. The contacts were classified as strong, medium and weak. The

upper bounds of the distance constraints were set to 2.8, 3.5 and 5.0 Å, respectively. Additional constraints from the 150 ms NOESY were set to 6.0 Å to take spin diffusion into account. Lower bounds were set to the sum of the Van der Waals radii of the involved atoms. Pseudoatom corrections were applied for methylene and aromatic ring protons (Wüthrich *et al.*, 1983). Constraints involving methyl groups were corrected for the three-proton intensity and a pseudoatom correction of 0.3 Å was applied (Koning *et al.*, 1990). In total, 1041 distance constraints were obtained, of which 102 are medium-range NOEs and 266 long-range NOEs. From the ratio of cross-peak to diagonal-peak in the 3D HNHA experiment, 56 $^3J_{HNH\alpha}$ were identified. The corresponding ϕ angles were converted to dihedral constraints with a range of $\pm$ 30° to account for inaccuracies in the measured coupling constants. Also 22 rotameric states relative to the $\chi 1$ angles were determined from the combined analysis of the data. Structures were generated using the program DG-II (Biosym Technologies, Inc.). Forty structures were generated, by embedding in four dimensions and, after projection of the coordinates into three dimensions, simulated annealing with a simple force field. Subsequently, these structures were optimized with the program Discover (Biosym Technologies, Inc.) by molecular dynamics and energy minimization, using the Consistent Valence Force Field. The protocol consisted of a short restrained energy minimization of 600 steps, followed by 10 ps of restrained molecular dynamics *in vacuo* at a constant temperature of 300 K, with a time step of 1 fs. Finally, a restrained energy minimization of 600 steps was carried out. Charges were included during the calculations. A cut-off distance of 12 Å was used for the non-bonded interactions. The structure refinement was done as an iterative procedure. With an initial set of ~200 constraints, a number of structures were calculated. Inspection of these structures helped in reducing several ambiguities in the assignment of additional NOESY cross-peaks. With the additional constraints, a new set of structures was calculated, until a final set of 1041 distance constraints was obtained. From the 40 structures, 18 were discarded on the basis of higher constraint energy and three others on the basis of high constraint violations in loop IV.

Acknowledgements

We would like to thank Dr Marco Tessari for many helpful discussions and suggestions. This work was supported in part by grants from the Italian MURST and National Research Council (P.F. Ingegneria Genetica) to C.O.G. and Progetto Strategico 'Biologia Strutturale' to M.P. and by grants from the Netherlands Foundation for Chemical Research (SON) and the Netherlands Organization for Scientific Research (NWO). The 750 MHz NMR spectra were recorded at the SON NMR Large Scale Facility (Utrecht) which is supported by the Large Scale Facility program of the European Union.

References

Bodenhausen,G. and Ruben,D.J. (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.*, **69**, 185–189.

Bothner-By,A.A., Stephens,R.L. and Lee,Y.M. (1984) Structure determination of a tetrasaccharide: transient nuclear Overhauser effect in the rotating frame. *J. Am. Chem. Soc.*, **106**, 811–813.

Brandi,A., Pietroni,P., Gualerzi,C.O. and Pon,C.L. (1996) Post-transcriptional regulation of *CspA* expression in *Escherichia coli*. *Mol. Microbiol.*, **19**, 231–240.

Calogero,R.A., Pon,C.L. and Gualerzi,C.O. (1987) Chemical synthesis and *in vivo* hyperexpression of a modular gene coding for *Escherichia coli* translational initiation factor IF1. *Mol. Gen. Genet.*, **208**, 63–69.

Celano,B., Pawlik,R.T. and Gualerzi,C.O. (1988) Interaction of *Escherichia coli* translation initiation factor IF1 with ribosomes. *Eur. J. Biochem.*, **178**, 351–355.

Cummings,H.S. and Hershey,J.W.B. (1994) Translation initiation factor IF1 is essential for cell viability in *Escherichia coli*. *J. Bacteriol.*, **176**, 198–205.

Cummings,H.F., Sands,J.F., Goreman,P.C., Fraser,J. and Hershey,J.W.B. (1991) Structure and expression of the *infA* operon encoding translational initiation factor 1. *J. Biol. Chem.*, **266**, 16491–16498.

Godefroy-Colburn,T., Wolfe,A.D., Dondon,J., Grunberg-Manago,M., Dessen,P. and Pantaloni,D. (1975) Light-scattering studies showing the effects of initiation factors on the reversible dissociation of *E.coli* ribosomes. *J. Mol. Biol.*, **94**, 461–478.

Griesinger,C., Otting,G., Wüthrich,K. and Ernst,R.R. (1988) Clean TOCSY for ^{1}H-proton spin systems identification in macromolecules. *J. Am. Chem. Soc.*, **110**, 7870–7872.

Gualerzi,C.O. and Pon,C.L. (1990) Initiation of mRNA translation in prokaryotes. *Biochemistry*, **29**, 5881–5889.

Gualerzi,C.O., Spurio,R., La Teana,A., Calogero,R., Celano,B. and Pon,C.L. (1989) Site-directed mutagenesis of *Escherichia coli* translation initiation factor IF1. Identification of the amino acids involved in its ribosomal binding and recycling. *Protein Eng.*, **3**, 133–138.

Hershey,J.W.B. (1987) Protein synthesis. In Neidhardt,F.C., Ingraham,J.L., Low,K.B., Magasanik,B., Schaechter,M. and Umbarger,H.E. (eds), *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology.* American Society for Microbiology, Washington, DC, pp. 613–647.

Holm,L. and Sander,C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.*, **233**, 123–138.

Ikura,M., Bax,A., Clore,G.M. and Gronenborn,A. (1990) Detection of nuclear Overhauser effects between degenerate amide proton resonances by heteronuclear three-dimensional nuclear magnetic resonance spectroscopy. *J. Am. Chem. Soc.*, **112**, 9020–9022.

Jeener,J., Meier,B.H., Bachman,P. and Ernst,R.R. (1982) Investigations of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.*, **71**, 4546–4553.

Jones,P.G., Krah,R., Tafuri,S.R. and Wolffe,A.P. (1992) DNA gyrase, CS7.4 and the cold shock response in *Escherichia coli*. *J. Bacteriol.*, **174**, 5798–5802.

Koning,M.M.G., Boelens,R. and Kaptein,R. (1990) Calculation of the nuclear Overhauser effect and the determination of proton–proton distances in the presence of internal motions. *J. Magn. Resonance*, **90**, 111–123.

Kraulis,P. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.

La Teana,A., Brandi,A., Falconi,M., Spurio,R., Pon,C.L. and Gualerzi,C.O. (1991) Identification of a cold shock enhancer of the *Escherichia coli* gene encoding nucleoid protein H-NS. *Proc. Natl Acad. Sci. USA*, **88**, 10907–10911.

Lee,S.J., Xie,A., Jiang,W., Etchegaray,J.P., Jones,P. and Inouye,M. (1994) Family of the major cold-shock protein, CspA (CS7.4), of *Escherichia coli*, whose members show a high sequence similarity with the eukaryotic Y-box binding proteins. *Mol. Microbiol.*, **11**, 833–839.

Marion,D., Driscoll,P.C., Kay,L.E., Wingfield,P.T., Bax,A., Gronenborn,A.M. and Clore,G.M. (1989) Overcoming the overlap problem in the assignment of ^{1}H NMR spectra of larger proteins by use of three dimensional heteronuclear ^{1}H-^{15}N Hartmann–Hahn multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1β. *Biochemistry*, **28**, 6150–6156.

Moazed,D., Samaha,R.R., Gualerzi,C. and Noller,H.F. (1995) Specific protection of 16S rRNA by translational initiation factors. *J. Mol. Biol.*, **248**, 207–210.

Morris,A.L., McArthur,M.W., Hutchinson,E.G. and Thornton,J.M. (1992) Stereochemical quality of protein structure coordinates. *Proteins*, **12**, 345–364.

Murzin,A.G. (1993) OB (oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.*, **12**, 861–867.

Newkirk,K., Feng,W., Jiang,W., Tejero,R., Emerson,S.D., Inouye,M. and Montelione,G.T. (1994) Solution NMR structure of the major cold shock protein (CspA) from *Escherichia coli*. Identification of a binding epitope for DNA. *Proc. Natl Acad. Sci. USA*, **91**, 5114–5118.

Paci,M., Pon,C.L. and Gualerzi,C.O. (1983) High resolution ^{1}H-n.m.r. study of the interaction between initiation factor IF1 and 30S ribosomal subunits. *EMBO J.*, **2**, 521–526.

Pawlik,R.T., Littlechild,J., Pon,C.L. and Gualerzi,C. (1981) Purification and properties of *E.coli* initiation factors. *Biochem. Int.*, **2**, 421–428.

Peng,J.W. and Wagner,G. (1992) Mapping of the spectral densities of the N–H bond motions in Eglin c using heteronuclear relaxation experiments. *Biochemistry*, **31**, 8571–8586.

Pon,C.L. and Gualerzi,C.O. (1984) Mechanism of protein biosynthesis in prokaryotic cells. Effect of initiation factor IF1 on the initial rate of 30S initiation complex formation. *FEBS Lett.*, **175**, 203–207.

Pon,C.L., Wittmann-Liebold,B. and Gualerzi,C. (1979) Structure–function relationships in *Escherichia coli* initiation factors. II. Elucidation of the primary structure of initiation factor IF1. *FEBS Lett.*, **101**, 157–160.

Ruff,M., Krishnaswamy,S., Boeglin,M., Poterszman,A., Mitschler,A., Podjarny,A., Rees,B., Thierry,J.C. and Moras,D. (1991) Class II aminoacyl transfer RNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with $tRNA^{Asp}$. *Science*, **252**, 1682–1689.

Sands,J.F., Cummings,H.S., Sacerdot,C., Dondon,L., Grunberg-Manago,M. and Hershey,J.W.B. (1987) Cloning and mapping of *infA*, the gene for protein synthesis initiation factor IF1. *Nucleic Acids Res.*, **15**, 5157–5168.

Schindelin,H., Marahiel,M.A. and Heinemann,U. (1993) Universal nucleic acid-binding domain revealed by crystal structure of the *B.subtilis* major cold-shock protein. *Nature*, **364**, 164–168.

Schindelin,H., Jiang,W., Inouye,M. and Heinemann,U. (1994) Crystal structure of CspA, the major cold shock protein of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **91**, 5119–5123.

Schleich,T., Verwolf,G.L. and Twombly,K. (1980) A circular dichroism study of *Escherichia coli* initiation-factor 1 binding to polynucleotides. *Biochim. Biophys. Acta*, **609**, 313–320.

Schmitt,E., Guillon,J.M., Meinnel,T., MacHulam,Y., Dardel,F. and Blanquet,S. (1996) Molecular recognition governing the initiation of translation in *Escherichia coli*. *Biochimie*, **78**, 543–554.

Schnuchel,A., Wiltscheck,R., Czisch,M., Herrler,M., Willimsky,G., Graumann,P., Marahiel,M.A. and Holak,T.A. (1993) Structure in solution of the major cold-shock protein from *Bacillus subtilis*. *Nature*, **364**, 169–171.

Spurio,R., Paci,M., Pawlik,R.T., La Teana,A., Di Giacco,B.V., Pon,C.L. and Gualerzi,C.O. (1991) Site-directed mutagenesis and NMR spectroscopic approaches to the elucidation of the structure–function relationships in translation initiation factors IF1 and IF3. *Biochimie*, **73**, 1001–1006.

Stringer,E.A., Sarkar,P. and Maitra,U. (1977) Function of initiation factor 1 in the binding and release of initiation factor 2 from ribosomal initiation complexes in *Escherichia coli*. *J. Biol. Chem.*, **252**, 1739–1744.

Vuister,G. and Bax,A. (1994) Quantitative *J* correlation: a new approach for measuring homonuclear three-bond $J(H^NH^\alpha)$ coupling constants in ^{15}N-enriched proteins. *J. Am. Chem. Soc.*, **115**, 7772–7777.

Willimsky,G., Bang,H., Fischer,G. and Marahiel,M.A. (1992) Characterization of *cspB*, a *Bacillus subtilis* inducible cold shock gene affecting cell viability at low temperatures. *J. Bacteriol.*, **174**, 6326–6335.

Wintermeyer,W. and Gualerzi,C.O. (1983) Effect of *Escherichia coli* initiation factors on the kinetics of NAcPhe-$tRNA^{Phe}$ binding to 30S ribosomal subunits. A fluorescence stopped-flow study. *Biochemistry*, **22**, 690–694.

Wishart,D.S., Bigam,C.G., Yao,J., Abildgaard,F., Dyson,H.J., Oldfield,E., Markley,J.L. and Sykes,B.D. (1995) ^1H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR*, **6**, 135–140.

Wolffe,A.P., Tafuri,S., Ranjan,M. and Familari,M. (1992) The Y-box factors: a family of nucleic acid binding proteins conserved from *Escherichia coli* to man. *New Biol.*, **4**, 290–298.

Wüthrich,K. (1986) *NMR of Proteins and Nucleic Acids*. John Wiley & Sons, New York.

Wüthrich,K., Billeter,M. and Braun,W. (1983) Pseudo-structures for the 20 common amino acids for use in studies of proteins conformations by measurements of intramolecular proton–proton distance constraints with nuclear magnetic resonance. *J. Mol. Biol.*, **169**, 949–961.

Zucker,F.H. and Hershey,J.W.B. (1986) Binding of *Escherichia coli* protein synthesis initiation factor IF1 to 30S ribosomal subunits measured by fluorescence polarization. *Biochemistry*, **25**, 3682–3690.

Received on September 20, 1996; revised on November 18, 1996