Biochimie (1991) 73, 1001–1006

Site-directed mutagenesis and NMR spectroscopic approaches to the elucidation of the structure–function relationships in translation initiation factors IF1 and IF3

R Spurio[1,2], M Paci[3], RT Pawlik[1], A La Teana[1,2], BV DiGiacco[1], CL Pon[1,2], CO Gualerzi[1,2]

[1]*Max-Planck-Institut für Molekulare Genetik, Abt Wittmann, Berlin-Dahlem, Germany;*
[2]*Laboratory of Genetics, Department of Biology, University of Camerino, Camerino (MC);*
[3]*Department of Chemistry, University of Rome, Tor Vergata, Rome, Italy*

(Received 3 December 1990; accepted 4 March 1991)

Summary — The recent developments in the knowledge of the structure and structure–function relationships of prokaryotic initiation factors IF1 and IF3 obtained in our laboratory by site-directed mutagenesis, biochemical and NMR-spectroscopic approaches are discussed.

protein synthesis / initiation factors / site-directed mutagenesis / 2-dimensional NMR spectroscopy

Introduction

Our laboratory has been engaged for many years in the investigation of the structural properties of prokaryotic translation initiation factors and the relationships existing between their structure and function. We were always encouraged to pursue this work by the genuine interest shown by the late Dr HG Wittmann and by our conviction that by doing so we were not only contributing to a better understanding of the translational initiation process but also offering a modest contribution to the more general problem of the elucidation of the structure and structure–function relationships in prokaryotic ribosomes.

The initiation factors are three proteins (IF1, IF2 and IF3) required for initiation of mRNA translation with prokaryotic ribosomes. The functional properties of these factors have been reviewed recently [1–3]. In this paper, we summarize the recent, mostly unpublished progress made in our laboratory by use of a combination of approaches towards a better understanding of the structure and the nature of the active sites of the two smaller factors (*ie* IF1 and IF3).

Results and discussion

Initiation factor IF1

Following the assembly of a synthetic modular gene (*infA**) encoding *E coli* initiation factor IF1 and the subsequent massive overproduction of this protein [4], the study of the secondary and tertiary structure of IF1 in solution and the production of large quantities of IF1 molecules altered in their primary structure by cassette mutagenesis became feasible. Due to the persistent failure to obtain workable crystals of IF1, the elucidation of the structure was undertaken by means of high resolution NMR spectroscopy. Preliminary studies using 1-dimensional ^{1}H-NMR spectroscopy in combination with a variety of standard techniques applied to spectroscopy such as pH variation, spin lattice relaxation time measurements and chemical modifications led to the assignment of a number of resonances, particularly those due to protons of the aromatic residues [5]. Subsequent 2-dimensional NMR studies (Paci, Gualerzi, Boelens and Kaptein, unpublished results) led to the complete assignment of the ^{1}H-NMR spectrum in both deuterated and protonated water. The complete sequential assignment of the proton resonances of the polypeptide backbone was obtained from COSY and NOESY spectra. Furthermore, cross-peaks between protons allowed the establishment of proximities between different parts of the backbone of the protein due to β-sheet or α-helix secondary structures. Thus, it was established that the IF1 molecule contains rather extended β-sheet structures. After a run of α-helix (residues 1–13), residues 15–36 as well as 59–65 were found to form stretches of β-sheet interrupted by some turns and by a short α-helical structure between residues 39 and 46.

```
                                  CSY                  ITR
                                  |F     Y  D Y   D   |ITK
                                  ||      V    V      ||
Eco     1  AKEDNIEMQGTVLETLPNTMFRVELENGHVVTAHISGKMRKNYIRILTGD  50
           |||| ||||||||||||| |||||||||| | ||||||| |||||| ||
Bsu     1  AKDDVIEVEGTIVETLPNAMFKVELENGHTVLAHVSGKIRMHFIRILPGD  50

                          ΔΔΔ
                           ΔΔ
                            Δ
Eco    51  KVTVELTPYDLSKGRIVFRSR  71
           |||||| |||| |||| || |
Bsu    51  KVTVELSPYDLTRGRITYRYK  71
```

Fig 1. Sequence comparison of *E coli* and *B subtilis* translation initiation factor IF1 showing the amino acid replacements or deletions introduced by genetic engineering. Eco, *Escherichia coli* IF1 sequence [11]; Bsu, *Bacillus subtilis* IF1 sequence [12]. Δ represents the deletions introduced. For further details see [7].

The analysis of the proximities in space of the protons of the amino acid side-chains from which the 3-dimensional fold of the protein can be obtained is now in progress so that the complete elucidation of the structure of this factor in solution, in spite of the absence of an X-ray crystal structure should soon be available.

Meanwhile, the 2-dimensional NMR spectra of several mutated IF1 molecules obtained by site-directed mutagenesis (*ie* IF1Asp29, IF1Tyr29, IF1Asn34, IF1Tyr34) have been analyzed and compared to the spectra obtained with wild-type (wt) IF1. These experiments, on the one hand, clearly confirmed that the desired amino acid replacement occurred in the correct position, while showing on the other hand, that the wt secondary structure is substantially conserved in all these mutants. Thus, the regions of the COSY and NOESY spectra which are more sensitive to conformational changes of the protein (*ie* the fingerprint region containing the cross-peaks between the resonances of the amide groups and those of the α- and β-protons of the amino acid residues which are particularly sensi-

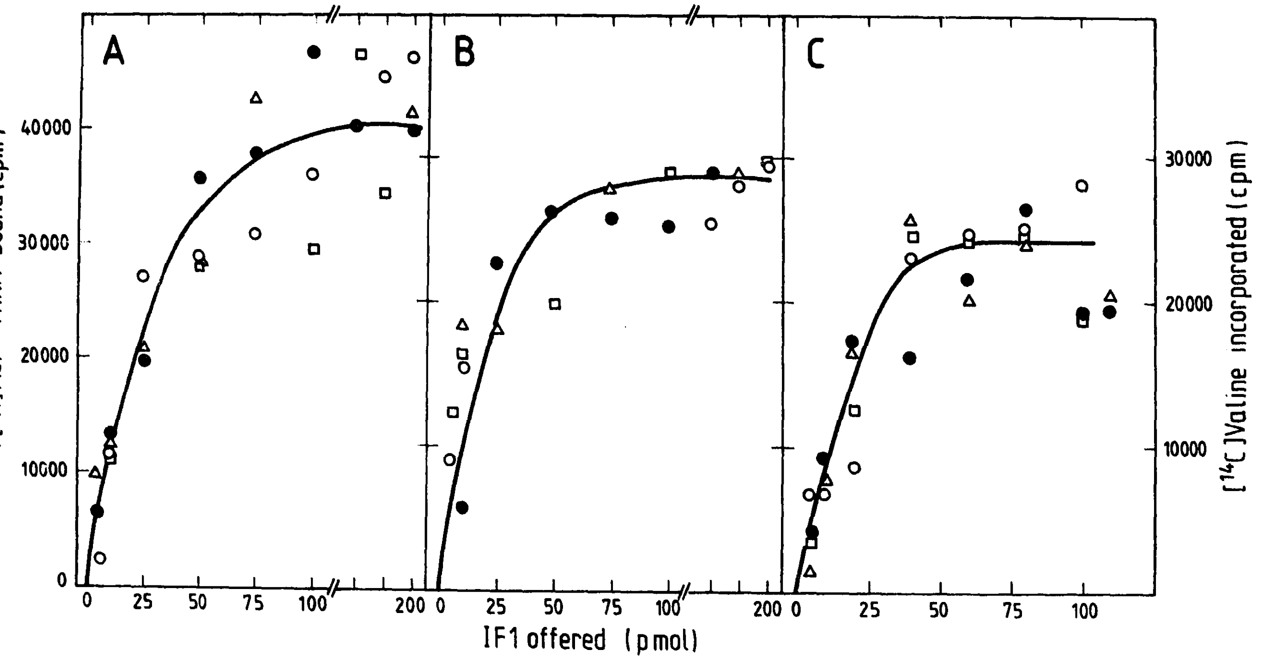

Fig 2. *In vitro* activity of IF1Phe21 mutants. Activity in poly(AUG)-dependent fMet-tRNA binding to (A) 30S subunits, (B) 70S ribosomes and in (C) MS2 RNA-dependent protein synthesis. (•) wt IF1; (○) IF1Cys21; (Δ) IF1Ser 21; (□) IF1Tyr21. The strategy for cassette mutagenesis and the experimental assay conditions are essentially as described in [7].

tive to chemical shifts caused by even slight changes of the chemical and magnetic environment) remained substantially unaffected indicating that no major structural alteration has taken place in the IF1 samples mutated by site-directed mutagenesis. This type of information is of course very important in the evaluation of the phenotypes of the mutated protein obtained from various types of *in vitro* studies. Specific spectral changes involving Arg, Ser, His, Phe and possibly Lys resonances have been observed by ^{1}H-NMR spectroscopy upon titrating the factor with increasing amounts of deuterated subunits indicating that at least one of each of these residues is either involved in or affected by the interaction of the factor with the 30S ribosomal subunit [6]. In figure 1, we present a comparison of the primary structures of the two known bacterial IF1 proteins and indicate the amino acid replacements introduced in the *E coli* molecule and whose effect has been tested sofar. Sequential deletion of the individual carboxy-terminal amino acids and replacement of the two histidine residues (His29 and His34) with either Tyr or Asp residues allowed us to conclude that Arg69 is part of the 30S ribosomal subunit binding site of IF1, that neither one of the two histidines is essential for the binding of the IF1 to the 30S ribosomal subunit nor for any functional activity of the factor with the exception of the 50S subunit-induced ejection of IF1 from the 30S subunit which depends, at least in part, on the presence of His29 [7].

The search for additional amino acid residues involved in the IF1-30S interactions prompted us to substitute Phe21, Arg22, Lys38 and Arg40 with other amino acids. The results presented in figure 2 show that substitution of Phe21 with either Cys, Ser or Tyr does not affect the capacity of IF1 to stimulate in the presence of the other two factors the binding of fMet-tRNA to 30S subunits (fig 2A) and 70S ribosomes (fig 2B) and to stimulate the MS2 RNA-dependent incorporation of valine (fig 2C). Thus, these results indicate that Phe21 is not essential for IF1 activity and strongly suggest that the Phe ring resonances affected by the IF1-30S interaction did not belong to Phe21 but rather to Phe68 which is adjacent to Arg69 which has been, as mentioned above, implicated in this interaction. Substitution of Lys38 with Ile or Thr does not impair the functional activity of IF1 in either the stimulation of the fMet-tRNA binding to 70S ribosomes or MS2 RNA-dependent protein synthesis. Surprisingly, however, the conservative Lys → Arg change in this position of the molecule resulted in a clear-cut reduction of both activities (fig 3A, B). These results indicate that the presence of a positively charged, or hydrophilic amino acid at position 38 is not essential for IF1 function; the likely explanation for the reduced activity of IF1Arg38, on the other hand, could be a steric hindrance due to the bulkier size of this residue. Two additional residues (*ie* Arg22 and Arg40) have been changed (fig 1) and the corresponding proteins purified. The analysis of their properties is still in progress.

Initiation factor IF3

It has been known for some time that IF3 binds to nucleic acids and that its physiological binding to 30S ribosomal subunits depends upon its interaction with the 16S RNA [1]. Several amino acid residues have been implicated in this interaction; thus, early work showed that chemical modification of Tyr109, Lys112 and an arginine residue, presumably Arg114, inhibited the IF3-30S interaction [8]. Two additional residues (*ie* Tyr71 and His139) were also shown by chemical modification to be somehow involved in the interaction. In fact, iodination of Tyr71 was found to prevent the functional interaction of the factor with the ribosomal subunit without impairing the binding to the 30S subunit. Photo-oxidation of the single His residue of *E coli* IF3 (His139) was also found to abolish the IF3 activity. In this case, the results indicated that His139 is not part of the primary nucleic acid binding site of the protein but that it can influence the binding of the factor to the 30S subunit either as part of a secondary binding site or as a constituent of a molecular hinge [8]. In agreement with these data, later ^{1}H-NMR spectroscopic investigations of the interaction of IF3 with either 30S ribosomal subunit or rRNA indicated that these interactions involve one Tyr, one Phe and some Arg and Lys residues of the protein, while a second aromatic residue (a Tyr or a Phe) and possibly the His residue are implicated in the interaction of IF3 with the 30S ribosomal subunit, but not with 16S rRNA. The position of the functionally relevant residues in the primary structure of IF3 and other types of considerations led us to suggest that this factor might contain a dual binding site for the 30S ribosomal subunit [8].

The availability of cloned *infC* genes from both *E coli* and *B stearothermophilus* and the advancement of DNA manipulation and NMR spectroscopic techniques (in particular, site-directed mutagenesis and 2-dimensional NMR spectroscopy) have now allowed us to test directly some of the earlier hypotheses. Several amino acid replacements have been introduced by the gapped duplex site-directed mutagenesis technique in *Bacillus stearothermophilus* and, when appropriate, in *E coli infC*. The amino acid replacements introduced in the IF3 molecule are summarized in figure 4.

As can be seen here, some of the old conclusions are fully supported by the newer results, while others, in particular the hypothesis of a dual IF3 active site, must be rejected in the light of recent structural and functional information.

The primary structures of *E coli* and *B stearothermophilus* IF3 are rather similar with 84 identical resi-

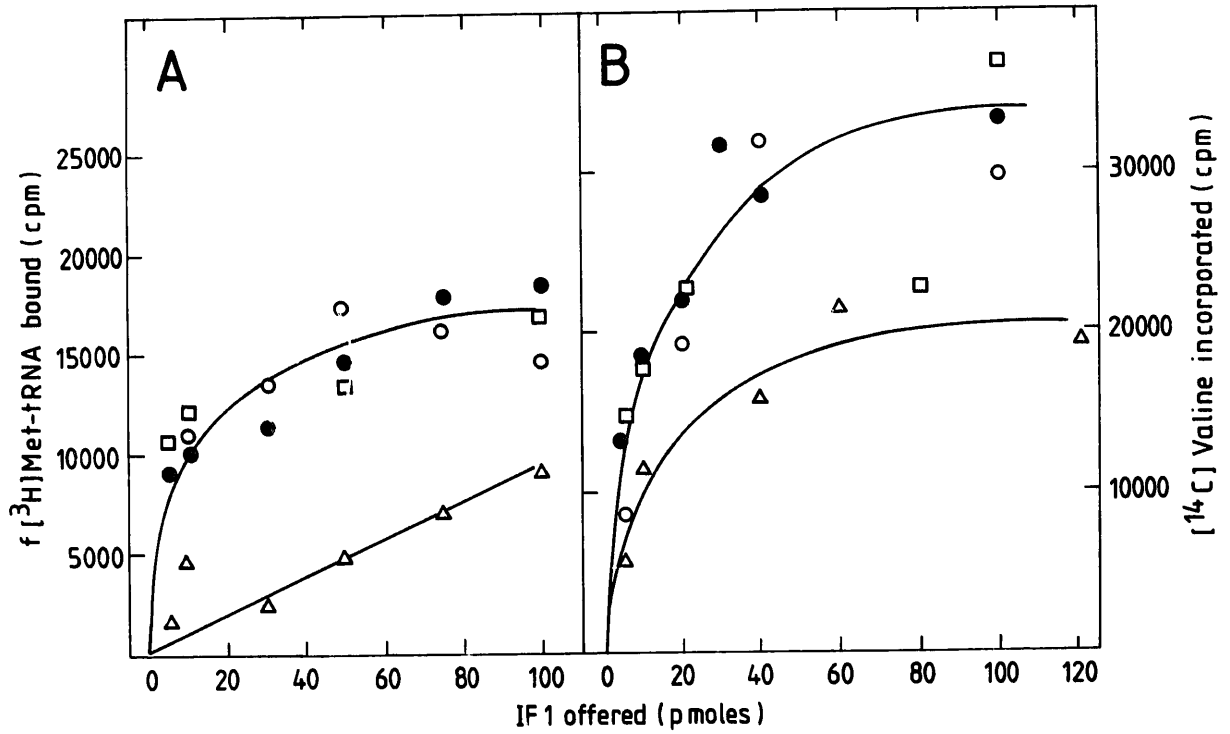

Fig 3. *In vitro* activity of IF1Lys38 mutants. Activity in (A) poly(AUG)-dependent fMet-tRNA binding to 70S ribosomes and in (B) MS2 RNA-dependent protein synthesis. (•) wt IF1; (○) IF1Ile38; (Δ) IF1Arg38; (□) IF1Thr38. The strategy for cassette mutagenesis and the experimental assay conditions are essentially as described in [7].

dues and 21 conservative replacements (fig 4). The thermophilic IF3 is 10 amino acids shorter, however, primarily due to the absence of the first nine N-terminal amino acids. The secondary and tertiary structures of the two proteins also appear to be very similar, at least as far as one can judge from a preliminary comparison of their 1- and 2-dimensional-NMR spectra.

A comparative investigation of the functional properties of *E coli* and *B stearothermophilus* IF3 has shown that the thermophilic factor binds to *E coli* 30S ribosomal subunits with a somewhat higher affinity than the homologous factor. This higher affinity does not impair, however, the normal ejection process when 50S subunits bind to 30S-IF3 complexes. In fact, release of IF3 occurs equally well with the mesophilic and thermophilic factors; furthermore, the two factors were found to be fully interchangeable in an *in vitro* translational system directed by natural mRNA. When binding to other types of 30S ribosomal subunits is studied, however, a strong difference between *B stearothermophilus* and *E coli* IF3 is observed. Thus, interaction of *E coli* IF3 with the small subunit of the moderate thermophile *B stearothermophilus* is strongly reduced in comparison to the homologous factor. This difference is even more pronounced when the binding of *B stearothermophilus* and *E coli* IF3 to the 30S subunit of the extreme thermophile *Thermus thermophilus* is compared. These findings seem to indicate that the mesophilic IF3 becomes progressively less 'fit' for the interaction with increasingly thermophilic ribosomes, while *B stearothermophilus* IF3 shows 'universal' binding properties.

The 2-dimensional NMR (COSY and NOESY) spectra of *E coli* and *B stearothermophilus* IF3 as well as of several IF3 molecules obtained by site-directed mutagenesis have been recorded and analyzed (Paci, Gualerzi, Pon, Boelens and Kaptein, unpublished results). The correlation spectra allowed us to obtain a scalar correlation pattern of many protons belonging to the same aromatic ring (*eg* ortho- and para-protons of the same tyrosine residues) and to make (or confirm) a number of resonance assignments. From the NOESY spectra, on the other hand, which, in addition to the scalar correlations, also show the dipolar correlations (NOE) through space, we were able to establish the existence of dipolar interactions between protons and thus to demonstrate a proximity between some protons belonging to different aromatic residues.

In particular, Tyr109 and Tyr76 showed a clear-cut cross-peak indicating the proximity between their aromatic ring protons, while a meta-proton of Tyr76 gave rise to a cross-peak with one or more of a group of several protons found between 7.1 and 7.5 ppm, possibly belonging to either the ring of Tyr71 or a Phe residue.

As to the exposure of these residues, photo-CIDNP experiments showed that, of all the aromatic residues present in IF3 which are potentially sensitive to this technique (*ie* the three Tyr and the single His), only Tyr76 and His139 are exposed to the solvent while the other two Tyr residues are probably located in a more hydrophobic environment (Paci, Gualerzi, Pon, Boelens and Kaptein, unpublished results). This result agrees with the finding that Tyr76 is by far more accessible to iodination than Tyr71 and 109 and that its modification increases the reactivity of the other two while the transition from monoiodo to di-iodo-tyrosine, which is faster in a more hydrophobic environment, occurs more readily for Tyr109 than for Tyr71 and Tyr76 [9]. Taken together these results clearly indicate that all three tyrosine residues of IF3 are probably clustered together in the tertiary structure and that Tyr109 and Tyr71, previously thought to belong to two separate IF3 active sites, are in reality very close to each other.

Chemical modifications implicated Tyr109 of *E coli* IF3 in the interaction with the 30S subunit. In *B stearothermophilus* IF3, however, this residue is replaced by a Phe residue (Phe98). This conservative change might have been favored by the need for a more stable (stacking?) interaction with rRNA at the higher temperature optimal for the growth of this bacterium; thus, we decided to substitute Tyr109 of *E coli* IF3 with a Phe residue by gapped duplex site-directed mutagenesis and to investigate the structural and functional properties of the resulting protein. A preliminary comparison of the 1- and 2-dimensional NMR spectra of wild-type IF3 (Tyr109) and mutated IF3 (IF3 Phe109) show that the two molecules have only marginal differences in the aromatic region of the spectrum, apart from the expected lack of the ortho- and meta-proton resonances of the tyrosine ring. In addition to confirming the previously published Tyr resonance assignments, the observation of the spectra of the mutated protein indicated that the Tyr → Phe change does not substantially alter the chemical environment of the other aromatic residues.

When the ribosomal binding capacity of the mutated molecule was tested, however, the IF3 Phe109 displayed a slight but distinct reduction of affinity for both *E coli* and *B stearothermophilus* 30S ribosomal subunits. This finding shows that the increased hydrophobicity of the aromatic ring is not in itself a stabilizing element for the IF3-ribosome interaction. The reduction in binding affinity in spite of the conservative nature of the amino acid replacement, on the other hand, seems to lend additional support to the notion that in *E coli* Tyr109 is somehow implicated in the interaction.

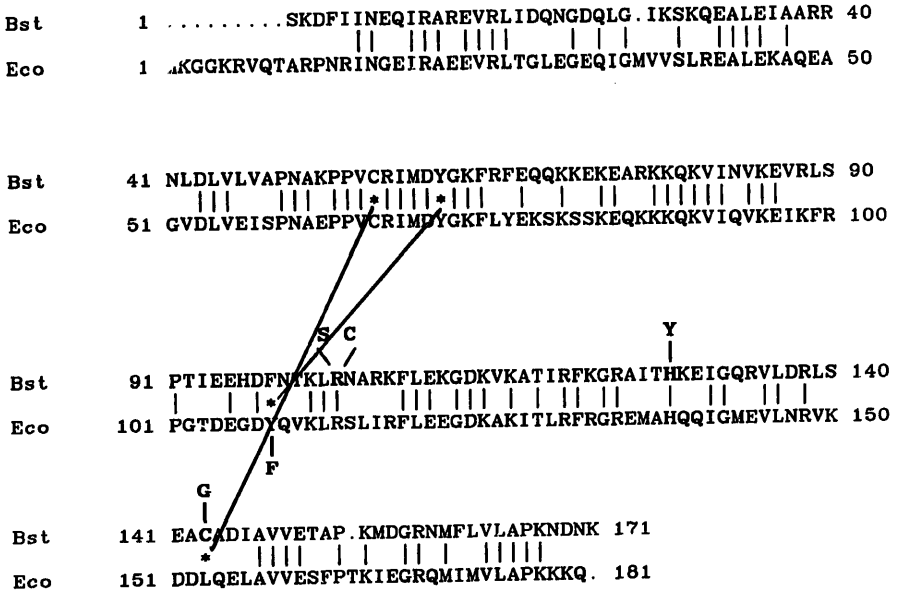

Fig 4. Sequence comparison of *E coli* and *B stearothermophilus* translation initiation factor IF3. The figure shows the amino acid replacements introduced by genetic engineering and the two established proximities between amino acids in the tertiary structure. Eco, *Escherichia coli* and Bst, *Bacillus stearothermophilus* IF3 sequences [13].

Another amino acid changed by site-directed mutagenesis was His128 of *B stearothermophilus* IF3 which was substituted by Tyr giving rise to Bst-IF3Tyr128. The substituted His corresponds to His139 of the *E coli* factor, which, as mentioned earlier, had been implicated in the interactions with 30S subunits [8]. As expected, compared to wt Bst-IF3, this mutant displayed a decreased binding activity for the 30S ribosomal subunit as well as a reduced activity in a typical functional assay for IF3 such as, the dissociation of 30S-poly(U)-NAcPhe-tRNA ternary complexes.Work is in progress to determine whether the structural properties of Bst-IF3Tyr128 and the release of this factor from the 30S subunit are different from those of wt IF3.

Reaction of *E coli* IF3 with 2,3 butanedione (2, 3BD) led to the reversible modification of several Arg residues accompanied by a reversible inactivation of the factor. From the time- and 2,3BD concentration-dependence of the inactivation, it was concluded that the modification of a single Arg residue (out of a possible 15) is responsible for the inactivation of the factor [10]. Peptide analysis of the modified protein led to inconclusive results, as far as the precise identification of the essential arginine was concerned. Nevertheless, the data were consistent with the possibility of Arg114 being the target of the inactivation and its proximity to Tyr109 and Lys112, both implicated in the binding, further strengthened the suspicion that modification of Arg114 might be responsible for the inactivation. Thus, to confirm this premise, Arg103 of Bst-IF3 (which is equivalent to Arg114 of *E coli* IF3) has been replaced by either Cys or Ser residues yielding Bst-IF3Cys109 and Bst-IF3Ser109. The finding of a near normal binding and biological activity, at least with the former mutated protein, strongly argues against the hypothesis of a direct implication of Arg114 (*E coli*) or Arg109 (*B stearothermophilus*) in the binding site of the factor and suggests that the functionally essential Arg residue of IF3 has yet to be identified.

Finally, in addition to the above-mentioned proximity of Tyr71 and 109 (fig 4), at least another tertiary structure constraint appears to be well established for the IF3 molecule. This is the proximity between Cys66 and Leu154 in light of the finding that a disulfide bridge joins the corresponding residues (*ie* Cys56 and Cys143) in *B stearothermophilus* IF3. This proximity is expected to bring into mutual proximity the most hydrophobic regions of the molecule. Consistent with this idea is the finding that the single Cys residue of *E coli* IF3 (*ie* Cys66). which is not essential for either the ribosomal binding or the activity of the factor is located very close to an hydrophobic patch or pocket of the molecule.

Since the thermal stability of many proteins is increased by the presence of -S-S- bridges, the difference between *E coli* and *B stearothermophilus* IF3 in this respect could stem from the necessity to increase the thermal stability of the latter molecule by introducing a covalent bond between two distant parts of the molecule. To test this hypothesis, Cys143 of BstIF3 was changed into a Gly residue and the corresponding protein was tested for both its biological activity and thermal stability in comparison with wt BstIF3. The results obtained indicate that while the binding capacity and the biological activity of BstIF3-Gly143 appear somewhat reduced with respect to the wt molecule, the mutated protein does not show any reduction in thermal stability, but rather, it displayed an increased heat resistance compared to the wt BstIF3.

Acknowledgments

This work was supported in part by grants from the Italian Ministry of Public Education and National Research Council (CNR) to COG Target projects on 'Biotechnology and Bioinstrumentation' and 'Genetic Engineering'.

References

1 Gualerzi CO, Pon CL (1990) *Biochemistry* 29, 5881–5889
2 McCarthy JEG, Gualerzi CO (1990) *Trends Genet* 6, 78–85
3 Gualerzi CO, La Teana A, Spurio R, Canonaco MA, Severini M, Pon CL (1990) *In: The Ribosome, Structure, Function and Evolution* (Hill WE, Dahlberg A, Garrett RA, Moore PB, Schlessinger D, Warner JR, eds) Am Soc Microbiol Washington, DC, 281–291
4 Calogero RA, Pon CL, Gualerzi CO (1987) *Mol Gen Genet* 208, 63–69
5 Paci M, Pon CL, Gualerzi CO (1988) *FEBS Lett* 236, 303–308
6 Paci M, Pon CL, Gualerzi CO (1983) *EMBO J* 2, 521–526
7 Gualerzi CO, Spurio R, La Teana A, Calogero RA, Celano B, Pon CL (1989) *Prot Eng* 3, 133–138
8 Gualerzi CO, Pon CL, Pawlik RT, Canonaco MA, Paci M, Wintermeyer W (1986) *In: Structure, Function and Genetics of Ribosomes* (Hardesty B, Kramer G, eds) Springer-Verlag, NY, 621–641
9 Bruhns J, Gualerzi C (1980) *Biochemistry* 18, 1670–1676
10 Lammi M (1982) Doctoral thesis, University of Calabria
11 Pon CL, Wittmann-Liebold B, Gualerzi C (1979) *FEBS Lett* 101, 157–160
12 Boylan SA, Suh JW, Thomas SM, Price CW (1989) *J Bacteriol* 171, 2553–2562
13 Kimura M, Ernst H, Appelt K (1983) *FEBS Lett* 160, 78–81