Nucleic Acids Research, Vol. 20, No. 17 4423–4428

Sequence comparison of new prokaryotic and mitochondrial members of the polypeptide chain release factor family predicts a five-domain model for release factor structure

Herman J.Pel, Martijn Rep and Leslie A.Grivell*
Section for Molecular Biology, Department of Molecular Cell Biology, University of Amsterdam, Kruislaan 318, 1098 SM Amsterdam, The Netherlands

Received June 25, 1992; Revised and Accepted August 13, 1992 Accession nos+

ABSTRACT

We have recently reported the cloning and sequencing of the gene for the mitochondrial release factor mRF-1. mRF-1 displays high sequence similarity to the bacterial release factors RF-1 and RF-2. A database search for proteins resembling these three factors revealed high similarities to two amino acid sequences deduced from unassigned genomic reading frames in *Escherichia coli* and *Bacillus subtilis*. The amino acid sequence derived from the *Bacillus* reading frame is 47% identical to *E.coli* and *Salmonella typhimurium* RF-2, strongly suggesting that it represents *B.subtilis* RF-2. Our comparison suggests that the expression of the *B.subtilis* gene is, like that of the *E.coli* and *S. typhimurium* RF-2 genes, autoregulated by a stop codon dependent +1 frameshift. A comparison of prokaryotic and mitochondrial release factor sequences, including the putative *B.subtilis* RF-2, leads us to propose a five-domain model for release factor structure. Possible functions of the various domains are discussed.

INTRODUCTION

Peptide chain release factors (RFs) are required for proper termination of polypeptide chain synthesis. RFs participate in this process by selective binding to translating ribosomes that have encountered a stop codon at their decoding site (1). This binding results in the release of the newly synthesized protein, presumably via an alteration of the catalytic activity of the peptidyl transferase centre from peptide bond formation to peptidyl-tRNA hydrolysis (2). The RFs best characterized so far are the prokaryotic release factors 1 and 2 (RF-1 and RF-2) (for review see 3). These RFs act in a codon-specific manner: RF-1 is required for UAA and UAG-dependent termination, while RF-2 mediates UAA and UGA-dependent termination. The genes encoding RF-1 and RF-2 (*prfA* resp. *prfB*) in the gram-negative bacteria *Escherichia coli* and *Salmonella typhimurium* have been cloned and characterized. The similar functions of RF-1 and RF-2 are reflected by similarity in primary structure: *E.coli* RF-1 and RF-2 share 37.2% identical amino acids. RF-3, a third, poorly characterized prokaryotic release factor, stimulates the activity of RF-1 and RF-2 two- to three-fold but lacks codon-specificity (4). A gene encoding this protein has not been identified so far. There has been a report on the cloning of a rabbit (5) gene that codes for a cytoplasmic release factor. Surprisingly, the deduced amino acid sequence shows no structural similarity to prokaryotic or mitochondrial counterparts, but appears to be near identical to human γ2 interferon-induced protein γ2 (6, 7) and mammalian tryptophanyl-tRNA synthetases (8, 9). For the human protein $tRNA^{Trp}$ acylating activity has been established and experiments are currently under way to test the hypothesis that tRNA charging and peptide chain release are catalyzed by the same protein (10, 11).

Recently, we cloned and characterized the first gene encoding an organellar release factor. The *MRF1* gene of the yeast *Saccharomyces cerevisiae* encodes the mitochondrial release factor mRF-1 (12). mRF-1 exhibits a high sequence similarity to prokaryotic release factors RF-1 and RF-2. Two lines of evidence suggest that mRF-1 is in fact an RF-1 type release factor. First, mRF-1 is more similar to RF-1 than to RF-2 (38.4% resp. 29.8% identity). Second, the stop codon UGA encodes tryptophan in yeast mitochondria. Therefore, a single release factor of the RF-1 type (recognizing both UAA and UAG codons) should be sufficient for termination of yeast mitochondrial polypeptide synthesis. Indeed, the activity of only one, RF-1 type release factor could be detected in rat mitochondria (13), organelles using the same stop codon assignment as yeast mitochondria.

In this study we report the identification of two as yet unassigned prokaryotic reading frames whose predicted products display high sequence similarity to RF-1, RF-2 and mRF-1. A genomic reading frame of the gram-positive bacterium *Bacillus subtilis* is a good candidate for the gene encoding *Bacillus* RF-2. Further analysis of this gene suggests that its expression, like

* To whom correspondence should be addressed

+ GenBank accession no. M34034; DDBJ accession no. D90218

that of the *E.coli* and *S. typhimurium prfB* genes, is autoregulated by a +1 frameshift. A second amino acid sequence with remarkable sequence similarity to RFs was deduced from an unassigned reading frame in the *E.coli pepD* region.

The identification of yeast mitochondrial mRF-1 and the putative *Bacillus* RF-2 enables us to compare these proteins with previously described prokaryotic RFs. Based on this comparison we present a five-domain model for release factor structure. Possible functions of the various domains are discussed.

RESULTS AND DISCUSSION

Identification of the *Bacillus subtilis prfB* Gene

The amino-acid sequence of yeast mRF-1 was used to screen the SwissProt protein database (release 21.0) and the EMBL nucleic acid database (release 30.0) with the FastA respectively the TFastA computer programs developed by Pearson and Lipman (14). The TFastA program translates DNA sequences in all six registers and analyzes the resulting amino-acid sequences with the FastA algorithm. As expected both methods yielded significant similarities between mRF-1 and the four known prokaryotic release factors: RF-1 and RF-2 from both *E.coli* (15, 16, 17) and *S. typhimurium* (18, 19). However, the TFastA search revealed additional high similarities of mRF-1 to the amino acid sequence deduced from two as yet unassigned genomic reading frames of *Bacillus subtilis* and *E.coli*. The same protein sequences were identified when the databases were screened with *E.coli* RF-1 and RF-2.

The *B.subtilis* reading frame is located immediately downstream of and probably in the same operon as the *secA* (*div-341*$^+$) gene (EMBL accession no. D90218; 20). The reading frame starts with an AUG codon and consists of 299 codons. The C-terminal part of the putative protein product is lacking, since no in frame stop codon is present in the sequenced region. Analysis of the amino acid sequence revealed that it is 31.4% identical to mRF-1 and 35.9% identical to *E.coli* RF-1. More important, the sequence appeared to be 47.7% identical to *E.coli* RF-2. Similar values were obtained for *S. typhimurium* RF-1 and RF-2, since these RFs are over 95% identical to their *E.coli* counterparts. These similarities strongly suggest that the identified reading frame represents the *B.subtilis prfB* gene encoding RF-2.

Further compelling evidence comes from a closer inspection of the region upstream of the reading frame. Craigen and Caskey (21) established that a high efficiency +1 frameshift is required for autoregulated expression of the *E.coli prfB* gene. Weiss and co-workers (22, 23) identified the nucleotides involved in this event. Our sequence comparison identifies a stretch of 18 nucleotides located 43 nucleotides upstream of the start codon of the previously identified reading frame that matches perfectly the minimal frameshifting window in *E.coli*, whereas the surrounding sequences show no significant similarity (Fig. 1). Such a frameshift extends the N-terminus of the putative release factor by 27 amino acids and further increases the similarity with the *E.coli* and *S. typhimurium* RF-2 (Fig. 2). These findings thus lend strong support to the idea that we have identified the *B.subtilis prfB* gene and make it also very likely that a similar frameshift is required for a regulated expression of this gene. It is striking that the site of frameshifting is located at exactly the same position when the putative *B.subtilis* RF-2 is aligned to *E.coli* RF-2 (Fig. 2). The same alignment suggests that only a small number of codons are lacking at the 3′ end of the

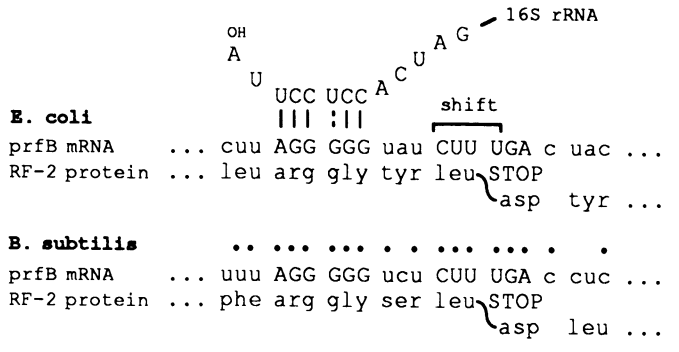

Figure 1. A DNA sequence involved in high efficiency +1 frameshifting in the *E.coli prfB* gene is also present in the putative *prfB* gene of *B.subtilis*. Weiss *et al.* (23) determined that both a 'shifty-stop' (three uracil residues directly preceding a UGA stop codon) and an appropriately spaced Shine-Dalgarno sequence are essential for an efficient reading frame switch in the *E.coli prfB* gene (nucleotides shown in capitals). Nucleotide sequences derived from the *E.coli* (position 452−473 in ref. 15) and the *B.subtilis prfB* gene (position 3191−3212 in ref. 20) are compared and the deduced amino acid sequences are shown. Dots denote positions of identical nucleotides.

identified reading frame. The near 50% identity between the putative *B.subtilis* RF-2 and *E.coli* as well as *S. typhimurium* RF-2s permits a first structural comparison of RF-2 type RFs (see below).

Structural similarities between *B.subtilis* and *E.coli* RF-2 could be anticipated from *in vitro* termination experiments. Lee and co-workers have shown that *B.subtilis* RF-2 is able to direct translational termination in response to UGA on *E.coli* ribosomes (13). Furthermore, antibodies raised against a region of similarity between *E.coli* RF-2 and RF-1 (RF-2 residues 232−295, Fig. 2) cross-reacted with *B.subtilis* RF-2, predicting the presence of a structurally related domain in the latter RF (1). Indeed, the alignment of both RF-2s reveals many positions of identical or related amino acids in this region (Fig. 2). Both in *E.coli* and *S. typhimurium* several mutations influencing translational fidelity turned out to reside in release factor genes (reviewed in 24). Unfortunately, few mutations of this kind have been described for *B.subtilis*, and no mutations have been mapped to the *secA* region (25).

Identification of *Escherichia coli* RF-H

A second amino acid sequence showing significant sequence similarity to mRF-1 and the prokaryotic release factors was deduced from a reading frame in the *E.coli pepD* region (EMBL accession no. M34034; 26). This reading frame is located downstream and on the opposite strand of the *pepD* gene. It starts with an AUG codon (nucleotides 1940 to 1938 in ref. 26), consists of 141 codons, and ends with an UAA codon (nucleotides 1518 to 1516). Based on its similarity to peptide chain release factors we have designated the putative product RF-H (Release Factor Homologue). Alignment of the homologous region of RF-H (residues 46 to 132) with other RFs revealed a similar percentage of identities for both RF-1 and RF-2 type RFs: 46.0% for both mRF-1 and *E.coli* RF-1, 40.2% and 42.5% for *E.coli* and the putative *B.subtilis* RF-2 respectively (see Fig. 2). RF-H may correspond to RF-3. However, a size of approximately 46 kDa has been reported for RF-3 (27). Alternatively, RF-H may represent a new protein factor involved in translational termination. A more detailed characterization of RF-H is necessary in order to assess its relevance.

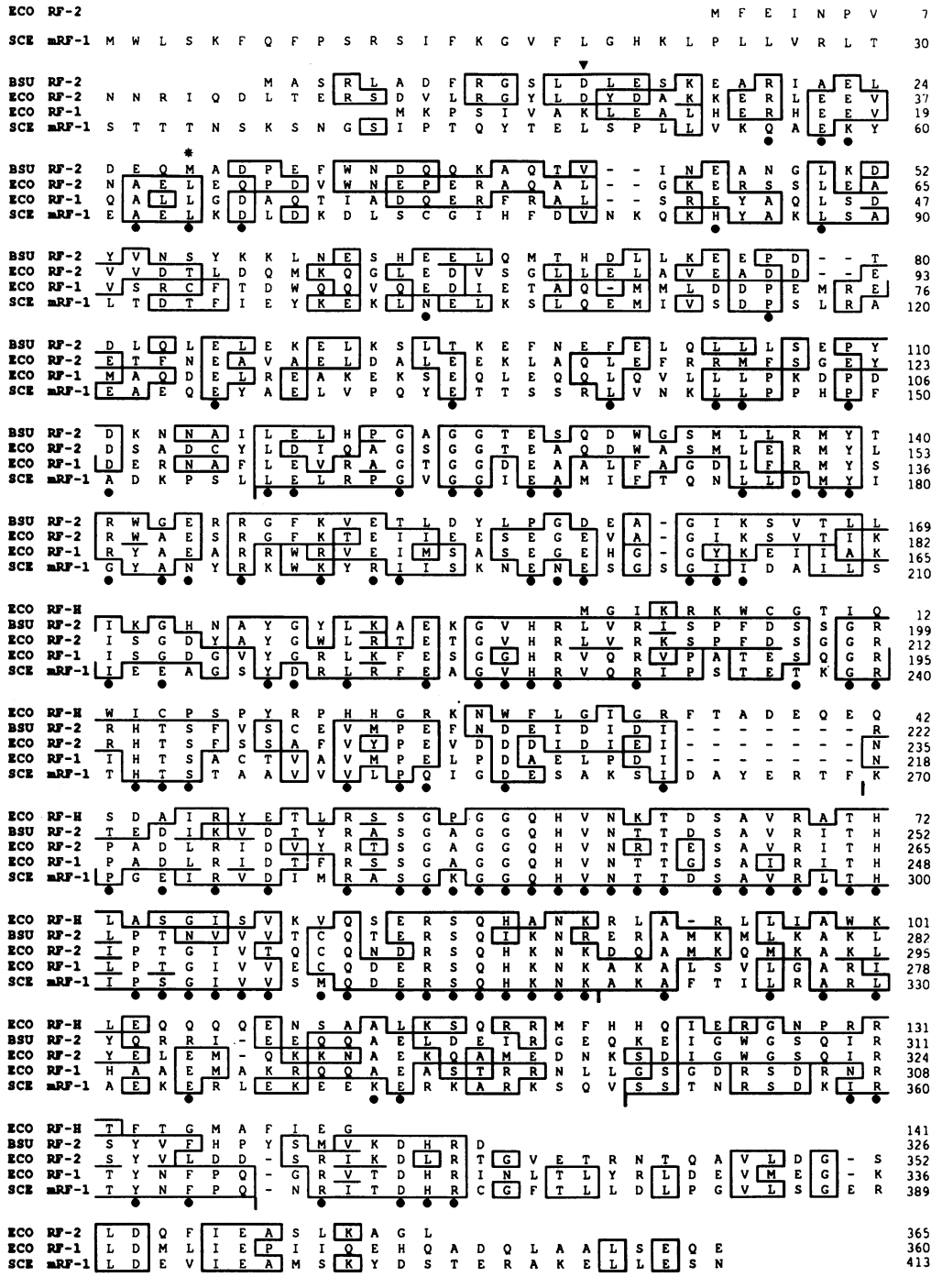

Figure 2. Multiple alignment of the amino acid sequence of *S.cerevisiae* mRF-1 (SCE mRF-1), the *E.coli* release factors 1 (ECO RF-1) and 2 (ECO RF-2), the putative *B.subtilis* RF-2 (C-terminus not yet determined, BSU RF-2) and *E.coli* RF-H (ECO RF-H). Positions of identity between the various release factors are boxed. Black dots indicate positions where three or four RFs (except RF-H) contain identical amino acids. The sequences were aligned by the FastA program of Pearson and Lipman (14). Dashes indicate gaps introduced to optimize the alignment. The arrowhead indicates the position of +1 frameshifting in the RF-2s. The asterisk denotes the previously identified N-terminus of the putative *B.subtilis* RF-2 (20). Vertical lines below the alignment indicate domain borders (see text for details).

Structural comparison of bacterial and mitochondrial RFs

The primary structures of only *E.coli* and *S. typhimurium* RF-1 and RF-2 have been known for several years. The identification of the yeast mitochondrial mRF-1 and the putative *B.subtilis* RF-2 has now allowed us to screen a family of prokaryotic and organellar RFs for the presence of both common and type-specific domains. An alignment of the various members of this family is shown in Figure 2. A detailed inspection of this, combined with an analysis of predicted secondary structures (Fig. 3), prompt us to propose a five-domain model for release factor structure. The sections below describe the structural characteristics that allow a distinction between various domains. Little attention will be paid to RF-H in the current evaluation, since its relevance is at present unknown. Fixed residue numbers

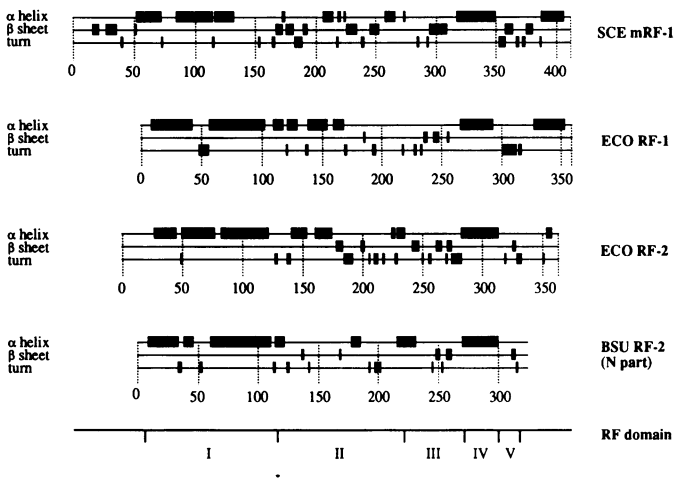

Figure 3. Predicted secondary structure of release factors. The amino acid sequence of four release factors (abbreviations as in Fig. 2) was analyzed for the presence of α helices, β sheets and turns as described by Chou and Fasman (34) and Garnier *et al.* (35). Only secondary structure elements predicted by both algorithms are shown. The plots were generated by the computer program MacVector (version 3.5, Int. Biotech. Inc.). The C-termini of the amino acid sequences are adjusted according to the alignment shown in Figure 2. This implicates that the N-terminus of mRF-1 is shifted in N-terminal direction by about nine amino acids due to the presence of several gaps in the alignment. Roman numbers identify release factor domains I to V.

referring to mRF-1 are used to mark the joining-regions of neighbouring domains.

The multiple alignment demonstrates that mRF-1 has an N-terminal extension when compared to the other RFs. This extension has the characteristics of a cleavable pre-sequence, required for targeting this protein into mitochondria (28 and H.J. Pel, unpublished results). The alignment also reveals that primary structure similarity is not evenly distributed throughout the aligned proteins. Relatively few common residues (positions where three or all four RFs contain identical amino acids indicated, see Fig. 2) are found in the N-terminal part of the alignment (residues 41−156, domain I). However, as shown by the predicted secondary structure of the various RFs (Fig. 3), all RFs likely share extensive a helical structures in this region.

Domain II, the region between residues 157 and 269, as well as domain III, bordered by residues 270 and 319, have characteristics that are quite opposite to those of domain I: no common secondary structure elements are predicted and the number of shared amino acids is high (40.7% and 78.0% for domains II and III respectively). Several RF type-specific residues are present in domain II (Fig. 4). Both the different levels of primary structure similarity and the presence of a relatively large gap in the multiple alignment compel us to consider domains II and III as separate entities. Moffat *et al.* (1) have shown that limited papain treatment of *E.coli* RF-1 as well as RF-2 yields two stable fragments, suggesting that RFs are comprised of two

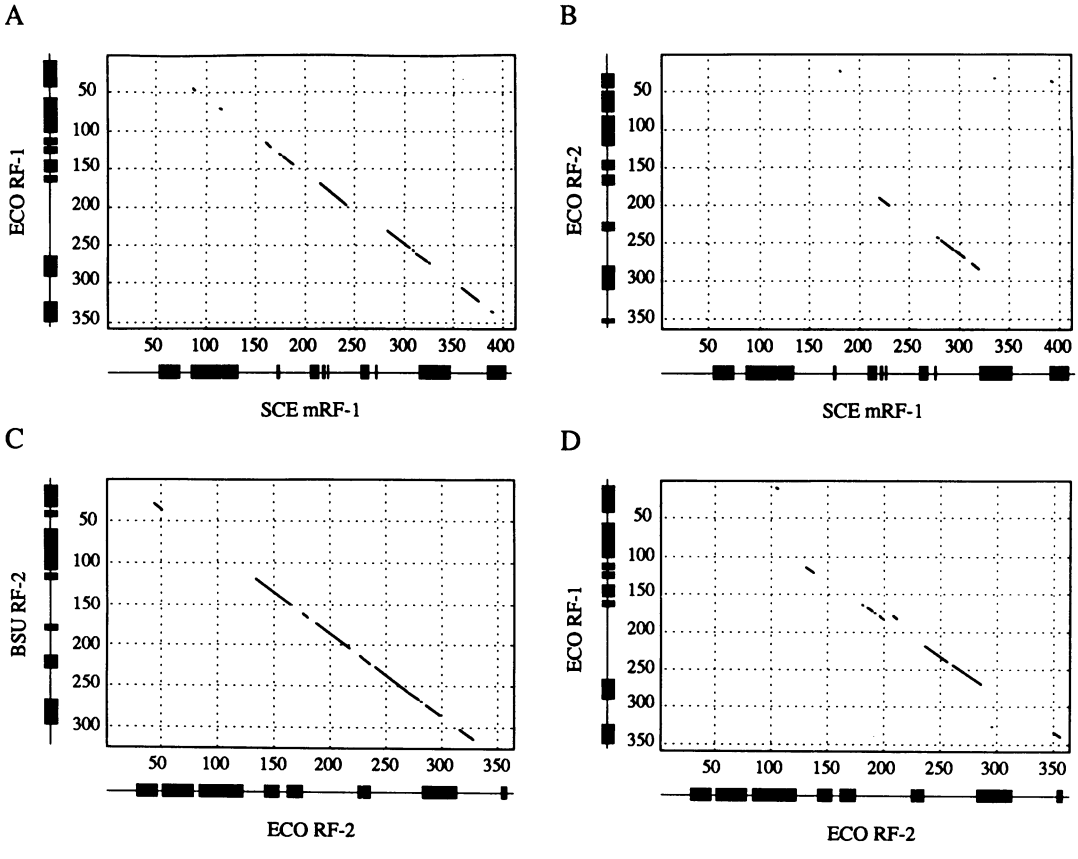

Figure 4. Combined comparison of primary and secondary structure of prokaryotic and mitochondrial release factors. Dot matrix plots of *S.cerevisiae* mRF-1 versus *E.coli* RF-1 and RF-2 and *E.coli* RF-2 versus *E.coli* RF-1 and *B.subtilis* RF-2 were generated by the computer program MacVector (version 3.5, Int. Biotech. Inc.) using the PAM250 scoring matrix and the following parameters: window 9; similarity 60%; hash value 1. The a helical structures predicted in Figure 3 are plotted along the x- and y-axis in order to visualize the alternating pattern of similarities in primary and secondary structure.

separate domains linked by a relatively exposed bridge. The cleavage site was proposed to be located between RF-1 residues 205–215, close to the gap in our alignment. Studies on the ribosomal binding sites of RFs have suggested that these factors have an extended structure (1 and see below). It is tempting to speculate that mRF-1 requires a seven amino acid insertion to accommodate the slightly larger size of yeast mitochondrial ribosomes compared to bacterial ribosomes.

The features of domain IV (residues 320–350) are quite similar to those of domain I: the amino acid similarity is relatively low (22.6%), and an a helical secondary structure is predicted for all RFs. Both length and position of the predicted helix are remarkably similar for the various RFs (Fig. 3). Figure 4 illustrates the alternating pattern of primary and secondary structure similarities, as well as the abruptness of the boundary between domains III and IV.

Domain V (residues 351–366) is characterized by an absence of common sequence motifs at both the primary and secondary structure level. However, we consider it as one of the most interesting domains, since it contains a number of amino acids that are only shared by RFs of the same type. This region, sixteen amino acids in length, contains eight RF-1 type and eight RF-2 type-specific residues but only four common residues (Fig. 2). The dot matrix plots presented in Figure 4 illustrate the presence of type-specific residues (compare 4A and 4C to 4B and 4D). This feature makes domain V a very good candidate for a region that interacts with the decoding centre and that confers stop codon specificity to RFs. An α helical structure predicted for the C-terminal region of both mRF-1 and RF-1 is much smaller in *E.coli* RF-2 (Fig. 3). The fact that the C-terminal part of the putative *B.subtilis* RF-2 has not yet been characterized makes it difficult to predict whether this indicates that domain V has to be extended further in C-terminal direction or whether a sixth domain is present in this region. Table 1 summarizes the characteristic features of the various domains.

Towards an identification of functional release factor domains

Release factors function in peptide chain termination by binding to ribosomes that have encountered a stop codon at their decoding site (1). The binding results in release of the nascent protein, presumably via an alteration of the catalytic activity of the peptidyl transferase centre from peptide bond formation to peptidyl-tRNA hydrolysis (2). In agreement with this function, studies on the ribosomal binding sites of RFs (reviewed in refs. 1 and 3) revealed the decoding site of the small subunit and the peptidyl transferase centre of the large subunit as sites of close contact between RF and ribosome. A third region of interaction encompasses the base of the L7/L12 stalk on the large subunit and the neck region of the concave curved side of the small subunit. Current models for ribosome structure predict the latter region to be quite distant from the first two sites, suggesting that RFs form an extended structure that penetrates deeply into the cleft between the two ribosomal subunits (29, 30).

Assuming that domain V is involved in stop codon recognition, we propose a release factor model in which domains I and II are required for binding at the entrance to the ribosomal interface near the base of the L7/L12 stalk (see Fig. 5). A second function for domain II may be to penetrate into the cleft between the ribosomal subunits, in order to bridge the distance between the stalk region and the other sites of interaction with the ribosome. A detailed study of L11, a ribosomal protein located at the base of the L7/L12 stalk, disclosed that *E.coli* RF-1 and RF-2 interact differently with this part of the ribosome (31, 32). The RF type-specific residues in domain II may reflect the subtle differences in ribosome binding of RF-1 and RF-2 type RFs. Ribosomal protein L11 has also been implicated in RF-3 function (4). *In vitro* termination experiments have led to the assumption that RF-3 stimulates the activity of RF-1 and RF-2 by lowering their K_m for binding to a terminating ribosome (33). RF-3 may therefore modulate RF binding via an interaction with domain I and/or domain II.

We propose that the release factor domains III, IV and V encompass sites of interaction with the functional centres of the ribosome, the peptidyl transferase centre and the decoding site. Domain III is a promising candidate for a site of interaction with protein and RNA components of the peptidyl transferase centre at the base of the central protuberance of the large subunit. Especially the common residues present in this domain may play an important role in peptidyl-tRNA hydrolysis. As explained before, we presume that domain V interacts with the decoding centre on the small subunit and confers stop codon specificity to RFs. Domain IV may assist either domain III or domain V in their proposed functions. Alternatively, this domain may contribute to a correct positioning of domain III and V on respectively the large and small ribosomal subunit.

Table 1. Features of release factor domains

domain	residues	% identity[a]	other features
all	325	34.4	
I	115	14.8	predicted α helix
II	113	40.7	
III	50	78.0	
IV	31	22.6	predicted α helix
V	16	25.0	RF type-specific residues

[a] amino acids identical in three or four release factors as defined in legend figure 2.

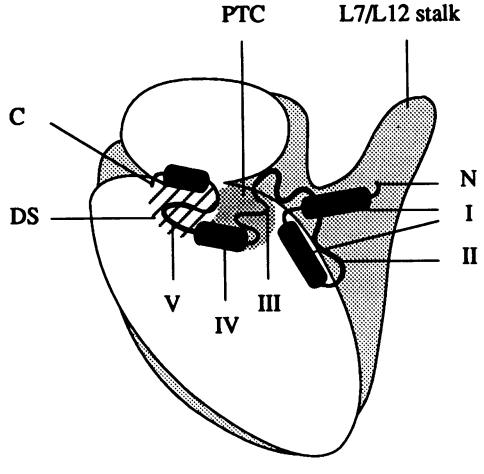

Figure 5. Schematic representation of ribosome binding of *E.coli* RF-1 to a 70S *E.coli* ribosome. Sites of interaction with the large (grey) and small (white) subunit are the decoding site (DS), peptidyl transferase centre (PTC), and the region encompassing the base of the L7/L12 stalk of the large subunit and the concave curved side of the head-neck region of the small subunit. The N- (N) and C-terminus (C) of RF-1 are indicated. Roman numbers identify release factor domains I to V, black boxes refer to α-helical structures. See text for further details.

It is obvious that the presented model for release factor structure needs further refinement. A functional analysis of mutants carrying lesions in specific release factor domains could be very helpful in this respect. However, for this moment we feel that our model could serve as a useful framework for future investigations of release factor mediated stop signal recognition and peptidyl-tRNA hydrolysis.

ACKNOWLEDGEMENTS

H.J.Pel acknowledges the European Molecular Biology Organisation (EMBO) for a Short Term Fellowship (ASTF 5809). We thank Han de Winde for helpful discussions and suggestions. This work was in part supported by a grant to L.A.Grivell from the Netherlands Foundation for Chemical Research (SON), with financial aid from the Netherlands Organization for the Advancement of Research (NWO).

REFERENCES

1. Moffat,J.G., Timms,K.M., Trotman,C.N.A. and Tate,W.P. (1991) *Biochimie*, **73**, 1113–1120.
2. Caskey,C.T. (1980) *Trends Biochem. Sci.*, **5**, 234–237.
3. Craigen,W.J., Lee,C.C. and Caskey,C.T. (1990) *Mol. Microbiol.*, **4**, 861–865.
4. McCaughan,K.K., Ward,C.D., Trotman,C.N.A. and Tate,W.P. (1984) *FEBS Lett.*, **175**, 90–94.
5. Lee,C.C., Craigen,W.J., Muzny,D.M., Harlow,E. and Caskey,C.T. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 3508–3512.
6. Fleckner,J., Rasmussen,H.H. and Justesen,J. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 11520–11524.
7. Buwitt,U., Flohr,T. and Böttger,E.C. (1992) *EMBO J.*, **11**, 489–496.
8. Frolova,L.Y., Sudomoina,M.A., Grigorieva,A.Y., Zinovieva,O.L. and Kisselev,L.L. (1991) *Gene*, **109**, 291–296.
9. Garret,M., Pajot,B., Trezeguet,V., Labouesse,J., Merle,M., Gandar,J.-C., Benedetto,J.-P., Sallafranque,M.-L., Alterio,J., Gueguen,M., Sarger,C., Labouesse,B. and Bonnet,J. (1991) *Biochemistry*, **30**, 7809–7817.
10. Rubin,B.Y., Anderson,S.L., Xing,L., Powell,R.J. and Tate,W.P. (1991) *J. Biol. Chem.*, **266**, 24245–24248.
11. Bange,F.-C., Flohr,T., Buwitt,U. and Böttger,E.C. (1992) *FEBS Lett.*, **300**, 162–166.
12. Pel,H.J., Maat,M.J., Rep.M. and Grivell,L.A. (1992) submitted for publication.
13. Lee,C.C., Timms,K.M., Trotman,C.N.A. and Tate,W.P. (1987) *J. Biol. Chem.*, **262**, 3548–3552.
14. Pearson,W.R. and Lipman,D.J. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2444–2448.
15. Craigen,W.J., Cook,R.G., Tate,W.P. and Caskey,C.T. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3616–3620.
16. Lee,C.C., Kohara,Y., Akiyama,K., Smith,C.L., Craigen,W.J. and Caskey,C.T. (1988) *J. Bacteriol.*, **170**, 4537–4541.
17. Mikuni,O., Kawakami,K. and Nakamura,Y. (1991) *Biochimie*, **73**, 1509–1516.
18. Elliott,T. (1989) *J. Bacteriol.*, **171**, 3948–3960.
19. Kawakami,K. and Nakamura,Y. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 8432–8436.
20. Sadaie,Y., Takamatsu,H., Nakamura,K. and Yamane,K. (1991) *Gene*, **98**, 101–105.
21. Craigen,W.J. and Caskey,C.T. (1986) *Nature*, **322**, 273–275.
22. Weiss,R.B., Dunn,D.M., Atkins,J.F. and Gesteland,R.F. (1987) *Cold Spring Harbor Symp. Quant.Biol.*, **52**, 687–693.
23. Weiss,R.B., Dunn,D.M., Dahlberg,A.E., Atkins,J.F. and Gesteland,R.F. (1988) *EMBO J.*, **7**, 1503–1507.
24. Parker,J. (1989) *Microbiol. Rev.*, **53**, 273–298.
25. Lovett,P.S., Ambulos Jr,N.P., Mulbry,W., Noguchi,N. and Rogers,E.J. (1991) *J. Bacteriol.*, **173**, 1810–1812.
26. Henrich,B., Monnerjahn,U. and Plapp,R. (1990) *J. Bacteriol.* **172**, 4641–4651.
27. Caskey,C.T., Forrester,W.C. and Tate,W.P. (1984) In Clark, B.F.C., Petersen, H.U., (eds.), Gene expression. Munksgaard, Copenhagen, pp. 149–158.
28. Hartl,F.U., Pfanner,N., Nicholson,D.W. and Neupert,W. (1989) *Biochim. Biophys. Acta*, **988**, 1–45.
29. Walleczek,J., Schüler,D., Stöffler-Meilicke,M., Brimacombe,R. and Stöffler,G. (1988) *EMBO J.*, **7**, 3571–3576.
30. Capel,M.S., Kjelgaard,M., Engelman,D.M. and Moore,P.B. (1988) *J. Mol. Biol.*, **200**, 65–87.
31. Tate,W.P., Schulze,H. and Nierhaus,K.H. (1983) *J. Biol. Chem.*, **258**, 12816–12820.
32. Tate,W.P., McCaughan,K.K., Ward,C.D., Sumpter,V.G., Trotman,C.N.A., Stöffler-Meilicke,M., Maly,P. and Brimacombe,R. (1986) *J. Biol. Chem.*, **261**, 2289–2293.
33. Goldstein,J.L. and Caskey,C.T. (1970) *Proc. Natl. Acad. Sci. USA*, **67**, 537–543.
34. Chou,P.Y. and Fasman,G.D. (1978) *Adv. Enzymol.*, **47**, 45–148.
35. Garnier,J., Osguthorpe,D.J. and Robson,B. (1978) *J. Mol. Biol.*, **120**, 97–120.