J Mol Evol (1998) 47:649–655

JOURNAL OF MOLECULAR EVOLUTION

Evolutionary Relationship Between Translation Initiation Factor eIF-2γ and Selenocysteine-Specific Elongation Factor SELB: Change of Function in Translation Factors

Patrick J. Keeling,[1,*] Naomi M. Fast,[2] Geoff I. McFadden[1]

[1] Plant Cell Biology Research Centre, Botany School, University of Melbourne, Parkville VIC 3052, Australia
[2] Department of Biochemistry, Dalhousie University, Halifax, N.S., Canada B3H 4H7

Received: 2 January 1998 / Accepted: 1 June 1998

Abstract. Eubacterial and eukaryotic translation initiation systems have very little in common, and therefore the evolutionary events that gave rise to these two disparate systems are difficult to ascertain. One common feature is the presence of initiation, elongation, and release factors belonging to a large GTPase superfamily. One of these initiation factors, the γ subunit of initiation factor 2 (eIF-2γ), is found only in eukaryotes and archaebacteria. We have sequenced eIF-2γ gene fragments from representative diplomonads, parabasalia, and microsporidia and used these new sequences together with new archaebacterial homologues to examine the phylogenetic position of eIF-2γ within the GTPase superfamily. The archaebacterial and eukaryotic eIF-2γ proteins are found to be very closely related, and are in turn related to SELB, the selenocysteine-specific elongation factor from eubacteria. The overall topology of the GTPase tree further suggests that the eIF-2γ/SELB group may represent an ancient subfamily of GTPases that diverged prior to the last common ancestor of extant life.

Key words: Translation initiation — eIF-2 — SELB — EF-Tu — Selenocysteine

Correspondence to: P.J. Keeling; *e-mail:* pkeeling@bio.indiana.edu
* *Present address:* Department of Biology, Indiana University, Bloomington, IN 47405, USA

Introduction

Translation initiation entails the assembly of charged initiator tRNA, the ribosome, and the message itself, in such a way that translation will commence at the appropriate codon. This is a very complex process and many of the interactions remain controversial (Kozak 1992; McCarthy and Brimacombe 1994). Moreover, in those eukaryotes and eubacteria that have been examined, translation initiation appears to take place by analogous, but not homologous mechanisms: different factors are employed, and many key events take place in a different order (for contrast, see Kozak 1984).

Particularly interesting in this regard is the path of import of the initiator tRNA to the small subunit of the ribosome. In both eukaryotes and eubacteria this tRNA is imported as part of a ternary complex with GTP and an initiation factor (IF-2 in eubacteria and eIF-2 in eukaryotes). In the canonical view of eukaryotic initiation, this ternary complex binds the small subunit first, then the message is recruited through the recognition of its cap and the cap-binding proteins (for review see Merrick 1992; Pain 1996). In eubacteria, on the other hand, the messenger RNA and small subunit may bind one another by Shine–Dalgarno base-pairing and other interactions. These interactions help determine the site of initiation and can form before the ternary complex is engaged (for review see Kozak 1984; McCarthy and Brimacombe 1994).

Furthermore, the translation factor components of the

ternary complex, IF-2 in eubacteria and eIF-2 in eukaryotes, are not the same protein (despite their unfortunately similar names). Both factors are GTPases, however, eubacterial IF-2 is a single polypeptide, while eIF-2 is composed of three heterologous subunits, α, β, and γ. The α and β subunits of eIF-2 are not detectably related to IF-2 and have no recognizable eubacterial equivalents. The γ subunit of eIF-2 is a member of the same GTPase superfamily as eubacterial IF-2, but the two proteins are only distantly related. This superfamily includes translation initiation, elongation, and release factors, which can be divided by the functional property of either requiring an external guanine nucleotide exchange factor or having a guanine nucleotide exchange domain (Keeling and Doolittle 1995a). eIF-2γ belongs to the class requiring a guanine nucleotide exchange factor (in the case of eIF-2, this accessory factor is eIF-2B), while eubacterial IF-2 homologues have a nucleotide exchange domain (Keeling and Doolittle 1995a). That these factors are not closely related to one another shows that they independently acquired their present role in translation initiation, so where did they come from originally? One way to tell is to identify their closest relatives within the GTPase superfamily.

Eukaryotic and archaebacterial homologues of eubacterial IF-2 have now been identified and these proteins were all found to form a subfamily that diverged prior to the divergence of the three domains (Keeling et al. 1996). Unfortunately the cellular activity of the archaebacterial and eukaryotic proteins is not known, although the eukaryotic protein is known to have an essential function in yeast (Sutrave et al. 1994) and has never been observed in the translation initiation complex.

As for eIF-2, the sequences of the *Methanococcus jannaschii* and *Archaeoglobus fulgidus* genomes revealed the presence of genes encoding proteins highly similar to all three subunits (Bult et al. 1996; Klenk et al. 1997). Apparently eIF-2 predates the archaebacterial–eukaryote divergence (we continue to use ''eIF-2'' to avoid confusion with ''IF-2''), but no eubacterial orthologue of eIF-2γ has been identified, and the position of the archaebacterial and eukaryotic eIF-2γ proteins within the GTPase superfamily has not been adequately addressed. Here we have analyzed this question, seeking to find a eubacterial orthologue of eIF-2γ and, also, to clarify the position of eIF-2γ within the whole of the GTPase superfamily. We have found eIF-2γ to be strongly and specifically related to the eubacterial elongation factor specific for selenocysteine, known as SELB (Forchammer et al. 1989). This SELB/eIF-2γ group appears, moreover, to be yet another ancient subfamily that arose prior to the divergence of the three domains.

Materials and Methods

Amplification and Sequencing of eIF-2γ Genes. Polymerase chain reaction was used to amplify a fragment of eIF-2γ from the diplomonad *Giardia lamblia* strain WB (ATCC 30957), the parabasalia *Trichomonas vaginalis* strain NIH-C1 (ATCC 30001) and *Monocercomonas* sp. strain Ns-1PRR (ATCC 50210), and the microsporidian *Nosema locustae* (ATCC 30860). Genomic DNA from the parabasalia were gifts from M. Müller, and that of *G. lamblia* and *N. locustae* from A.J. Roger. Amplification reactions consisted of 30 cycles with an annealing temperature of 57°C and contained both *Taq* and *Pfu* polymerases. The primers CGCCAGGCCACSATHAAYATHGGNAC and CCGCCTGGCTTGTTNACRTCRAA were used to amplify a fragment of approximately 650 bp that contains the four regions of eIF-2γ that can be reliably aligned to other related translation factors. Amplification products were isolated from agarose gels and cloned using the TA vector pCR2.1. In each case at least two individual clones were sequenced on both strands by either LiCor or ABI automated sequencers, and in no case were discrepancies observed between any clones from any one organism.

Phylogenetic Analysis. New eIF-2γ genes were aligned with all other reported eIF-2γ sequences and with representatives of related translation factors, EF-Tu, SELB, EF-1α, EF-G, EF-2, and IF-2. The *C. elegans* eIF-2γ sequence used here is derived from a composite of expressed sequence tags from the *C. elegans* EST database. The *A. fulgidus* eIF-2γ gene was found using TBLASTP (Gish and States 1993) to search the finished sequence, and the complete sequence retrieved from the TIGR FTP site. This sequence was subsequently confirmed in the published *A. fulgidus* genome (Klenk et al. 1997).

Trees of eIF-2γ alone were inferred from 215 amino acids (encompassing the length of the amplification products reported here) by distance, parsimony, and maximum likelihood. Corrected distances were calculated according to the Dayhoff PAM 250 substitution matrix, and trees constructed by neighbor-joining, using the PROTDIST and NEIGHBOR programs from the PHYLIP 3.5c package (Felsenstein 1993). Parsimony trees were inferred using 50 random addition heuristic searches with tree bisection and reconnection by PAUP 3.1.1 (Swofford 1993). In both cases 100 bootstrap resampling replicates were performed. Protein maximum-likelihood trees were inferred using ProtML from the MOLPHY 2.2 package (Adachi and Hasegawa 1994) by exhaustively searching all topologies of the 12 taxa analyzed using the JTT transition probability matrix (Jones et al. 1992) corrected for the observed frequency of amino acids in the data set (jf option). Estimated bootstrap confidence was calculated by the resampling estimated log likelihood (RELL) method (Hasegawa and Kishino 1994), and the support at each node was collated using mol2con (perl script provided by Arlin Stoltzfus). The relative-likelihood support was also calculated using TreeCons (Jermiin et al. 1997) using a class V weighting scheme and α values ranging from 0.5 to 0.01. The significance cutoff was not observed to affect the RLS, and therefore only the results from a cutoff of 0.1 are shown.

Trees of the larger data set including other related factors were inferred from 116 positions deemed to be clearly homologous between the disparate factors. These positions are found in four blocks of sequences used by other authors analyzing these proteins (Iwabe et al. 1989; Bourne et al. 1991; Keeling et al. 1996; Baldauf et al. 1996). The alignment is available upon request. The 12 eIF-2γ genes in Fig. 1 were included, as well as 27 EF-Tu genes, 4 SELB genes, and 22 EF-1α genes from eukaryotes and archaebacteria. This tree was left unrooted or rooted using one of three selections of outgroup sequences, consisting of either eight EF-2 and EF-G and six IF-2 homologues (as shown in Fig. 3) or groups of EF-2, EF-G, or IF-2 separately. Two GTPases from the *M. jannaschii* and *A. fulgidus* genomes were also analyzed, but since these were found to be very divergent, all analyses were also carried out both including and excluding them, and they were eventually excluded, as they appear to branch more or less at random within the tree.

Corrected distance and parsimony trees of this data set were inferred as described for eIF-2γ alone, but in the case of parsimony, the large number of equally parsimonious topologies within the EF-Tu and EF-1α subfamilies confounded the analysis. Protein maximum-

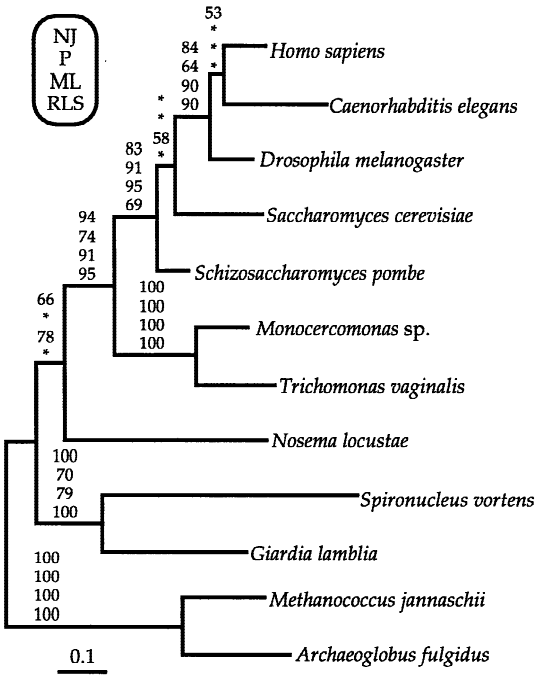

Fig. 1. Protein maximum-likelihood tree of eIF-2γ sequences, also showing bootstrap support for each node from parsimony and neighbor-joining, and RLS. Where support is less than 50%, an *asterisk* is shown.

likelihood trees were also inferred, but the large number of taxa could be dealt with only by partially constraining the tree. Subfamilies composed of clearly robust groups were constrained and left unresolved. Altogether there were six of these groups: EF-Tu, EF-1α (archaebacteria and eukaryotes together), eubacterial SELB, eIF-2γ (archaebacterial and eukaryotes together), and an outgroup consisting of one of the combinations of EF-G, EF-2, and IF-2 described for distance analyses. The branching order between the subfamilies in this partially constrained, partly unresolved tree was then exhaustively searched as described for the maximum-likelihood tree of eIF-2γ alone.

Results

Identification of eIF-2γ Genes from Diplomonads, Parabasalia, and Microsporidia. A fragment of the eIF-2γ consisting of approximately 650 bp was amplified from genomic DNA of the diplomonad *Giardia lamblia,* the parabasalia *Trichomonas vaginalis* and *Monocercomonas* sp., and the microsporidian *Nosema locustae.* This fragment covers the amino-terminal half of the gene, which contains four blocks of amino acid residues that are highly conserved between all members of the GTPase translation factors (Bourne et al. 1991). Since the regions outside these blocks are not readily aligned between the disparate GTPase factors, their absence does not affect the position of eIF-2γ in these analyses. In each case a single PCR product was obtained with a high degree of sequence similarity to other eIF-2γ genes, and all contain an insertion that is unique to eIF-2γ (Gaspar et al. 1994). This insertion is between 33 and 37 amino acids in eukaryotes (but only 20 and 23 in the archaebacteria) and is rich in proline, cysteine, and serine residues.

Phylogeny of eIF-2γ. Trees of eIF-2γ alone were inferred since the amount of data that can be used is far more than can be aligned between more distantly related members of the superfamily. Prior to the addition of these new sequences, eIF-2γ genes had been characterized only from animals, fungi, and expressed sequence tags from plants (which are too short to be included), and while this work was in progress eIF-2γ was also identified in the archaebacteria, *M. jannaschii* (Bult et al. 1996) and *A. fulgidus* (Klenk et al. 1997).

A maximum-likelihood tree of eIF-2γ with *M. jannaschii* and *A. fulgidus* as the outgroup is shown in Fig. 1. Neighbor-joining yielded a topology identical to that shown, and parsimony resulted in five equally parsimonious trees where the branching order of *Giardia, Spironucleus,* and *Nosema* were unresolved. Overall, this tree resembles those inferred from many other molecules: diplomonads branch very early, followed by microsporidia, parabasalia, and animals and fungi. Unfortunately, however, there are still relatively few eIF-2γ sequences known, so the tree cannot be adequately compared to other eukaryotic phylogenies. It should also be noted that evidence is accumulating to suggest that, contrary to this tree, the microsporidia may be specifically related to fungi (for reviews see Müller 1997; Keeling and McFadden 1998). Microsporidia generally branch early among eukaryotes in molecular trees based on components of the translation machinery, such as small subunit ribosomal RNA or elongation factors (Vossbrink et al. 1987; Kamaishi et al. 1996) but not in trees based on tubulin or chaperonin proteins (Keeling and Doolittle 1996; Edlind et al. 1996; Germot et al. 1997; Hirt et al. 1997). In these trees microsporidia are found to be related to fungi, suggesting that they are not primitive or ancient at all. The eIF-2γ tree is consistent with other translation proteins genes in placing the microsporidian deep in the eukaryotes, however, like microsporidian rRNAs and elongation factors, the *N. locustae* eIF-2γ is a highly divergent protein, and its position in the tree should likely be considered with caution.

Relationship of eIF-2γ to Eubacterial Selenocysteine-Specific Elongation Factor, SELB. The phylogenetic position of the eIF-2γ subfamily within the GTPase superfamily was resolved using a larger data set with representative sequences from the most closely related members of the family, EF-Tu, EF-1α, and SELB. Trees were inferred either unrooted or rooted with a selection of the more distantly related factors. Figure 2 shows the neighbor-joining tree of 79 factors with EF-2, EF-G, and IF-2 as an outgroup. In this tree eIF-2γ is most closely related to SELB, the eubacterial selenocysteine-specific elongation factor, and together these two form a group

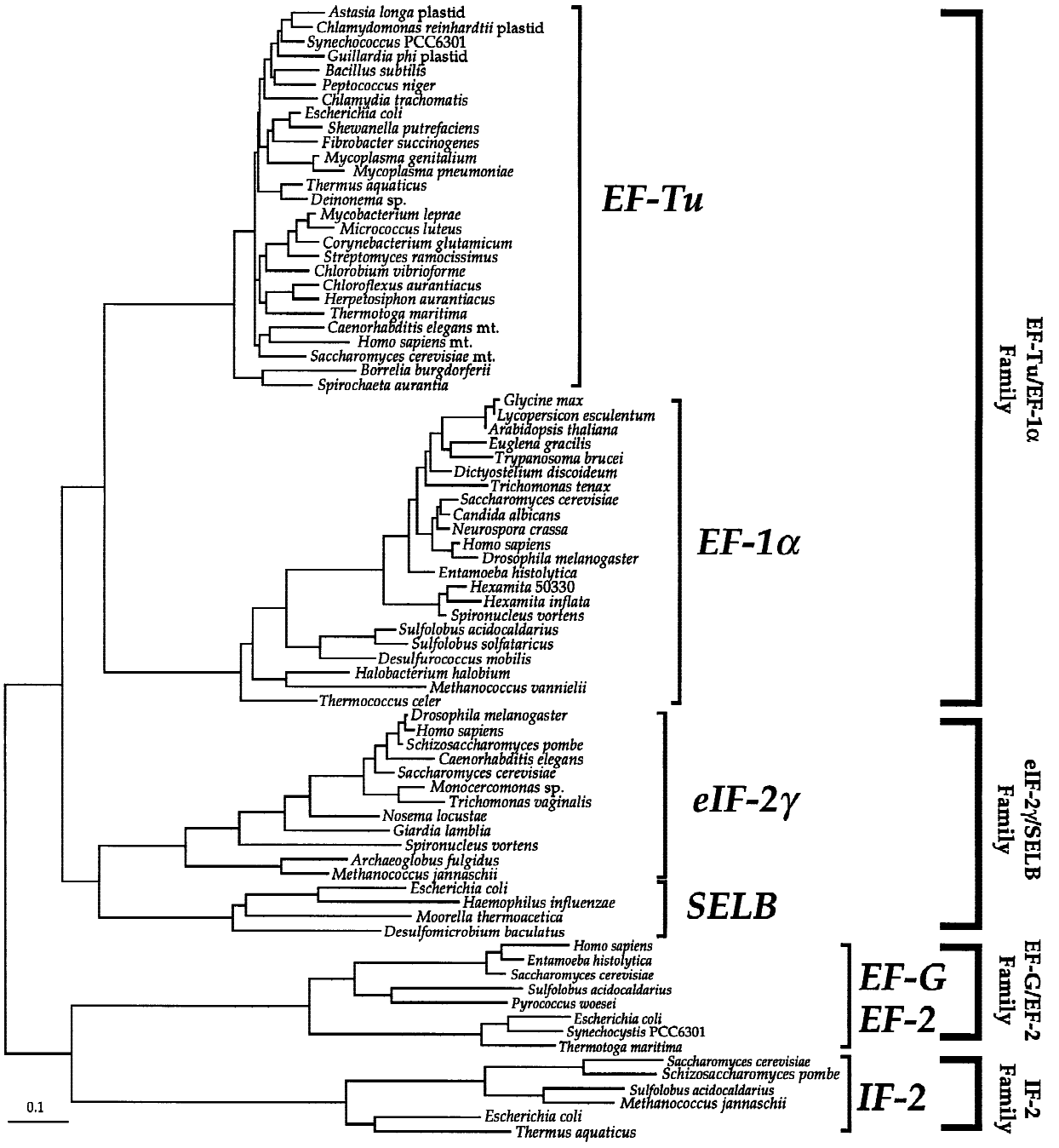

Fig. 2. Neighbor-joining tree consisting of 79 sequences using EF-2, EF-G, and IF-2 sequences as an outgroup. The groups indicated by *brackets to the right* are fairly robust in most cases, but the root of the nonrecycling factors is very uncertain (see Fig. 3).

that diverged prior to the divergence of the three domains. In all, there are four such groups, EF-2 and EF-G, EF-1α and EF-Tu, the IF-2-like group, and now this eIF-2γ and SELB group. All four groups mirror one another in that archaebacteria and eukaryotes are sister groups to the exclusion of eubacteria, as originally found for elongation factors (Iwabe et al. 1989).

However, a great evolutionary distance separates the outgroup and ingroup factors, and the composition of the outgroup was accordingly found to have a marked affect on the position of the root. In Fig. 3, four trees are shown, three with different combinations of EF-2, EF-G, and IF-2 as an outgroup and one unrooted. Each of these trees is congruent with the topology of the unrooted tree, but in each case the root falls at a different branch in both distance and maximum-likelihood analyses. To explain topologies B, C, and D, lateral transfers or the loss of a great number of paralogous genes must be evoked. Topology A, however, needs no special explanation, as it might simply represent the duplication and divergence of proteins prior to the divergence of eubacteria.

Notwithstanding the lack of certainty in the root of the

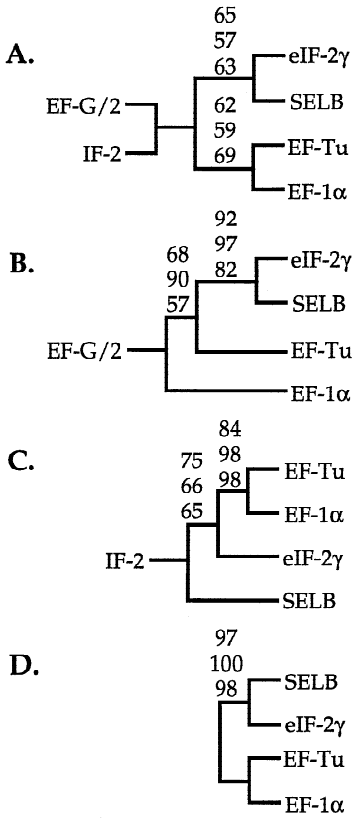

Fig. 3 Comparison of maximum-likelihood and neighbor-joining topologies resulting from eight data sets differing in the choice of outgroup. In each data set the maximum-likelihood and neighbor-joining trees were in agreement. *Numbers* at each nonconstrained node correspond to (from *top to bottom*) RELL bootstrap from protein maximum-likelihood, neighbor-joining bootstrap, and relative-likelihood support. Each of the three outgroups results in a topology compatible with the unrooted tree (tree D) but with different root.

tree, there is a relatively consistent and strongly supported relationship between eIF-2γ and eubacterial SELB (the single exception being topology C, which is the most poorly supported). Moreover, the relationship between eIF-2γ and SELB garners additional support from length variation in the region corresponding to the effector loop of these factors. This variable loop has been noted to be unusually short in eIF-2γ (Gaspar et al. 1995), and as Fig. 4 shows, the effector loop of SELB is exactly the same length (28 amino acids shorter than EF-1α).

Discussion

The family of GTPases involved in translation is growing ever more complex as new members are found and their relationships examined. Related factors have now been identified in all three domains and in all three main steps in translation: initiation, elongation, and termination. In some cases the role of orthologues is common in the three domains, as in EF-1α and EF-Tu or EF-2 and EF-G. However, this is not always found to be the case. For instance, archaebacterial and eukaryotic orthologues of eubacterial IF-2 have now been identified (Keeling et al. 1996), but these proteins likely do not have the same, if any, role in translation initiation. This can be said with some confidence for eukaryotes since the IF-2 orthologue has never been identified in many well-studied translation initiation complexes, and perhaps more importantly, the task performed by eubacterial IF-2 is known to be performed by eIF-2 in eukaryotes. In archaebacteria there is no direct information on translation initiation, but the system appears to be much like that of eukaryotes. Archaebacteria have a small suite of eukaryote-like initiation factors homologous to eIF-1A, eIF-2, eIF-2B, eIF-5A, and the eIF-4 family (Bartig et al. 1994; Keeling and Doolittle 1995b; Built et al. 1996; Klenk et al. 1997). The activity of these proteins has not been studied, but their collective presence and the concomitant absence of the eubacterial factors IF-1 and IF-3 suggest that the archaebacterial translation initiation complex resembles that of eukaryotes.

The presence of orthologous proteins with different jobs in different domains tells us that at least some of these proteins have switched function during evolution. In the case of IF-2 this is almost certainly the case, but until the present function of the archaebacterial and eukaryotic proteins is known, we cannot be certain what this switch might have entailed.

There is now also a good case for a change of function with eIF-2γ and SELB, and in this case both eubacterial and eukaryotic orthologues are of defined function. Moreover, both eukaryotes and eubacteria have nonhomologous but functionally analogous proteins to SELB and eIF-2γ, respectively. As mentioned earlier, eubacteria use IF-2 to the same end as eukaryotes use eIF-2. Similarly, eukaryotes do not have SELB but use a mechanism to incorporate selenocysteine cotranslationally that relies instead on other factors that bind the mRNA in the downstream untranslated region (Shen et al. 1995a, b). It would appear, therefore, that in the lineage leading to either eubacteria or eukaryotes, this protein switched function as in the IF-2 subfamily. Once again, however, it is impossible to say what the ancestral function might have been.

The phylogenetic place of eIF-2γ has not been well studied previously, having in the past been observed only to be related to EF-Tu (Gaspar et al. 1994; Keeling and Doolittle 1995a). With the relationship between SELB and eIF-2γ shown here, the evolutionary history of both factors is much clearer. Having said so much, however, it is also clear that as more such factors are identified and included in phylogenetic analyses, the relationships between the various subgroups of GTPase translation factors are becoming increasingly difficult to resolve.

Acknowledgments. This work was supported by grants from the Medical Research Council of Canada (to W.F. Doolittle) and the Aus-

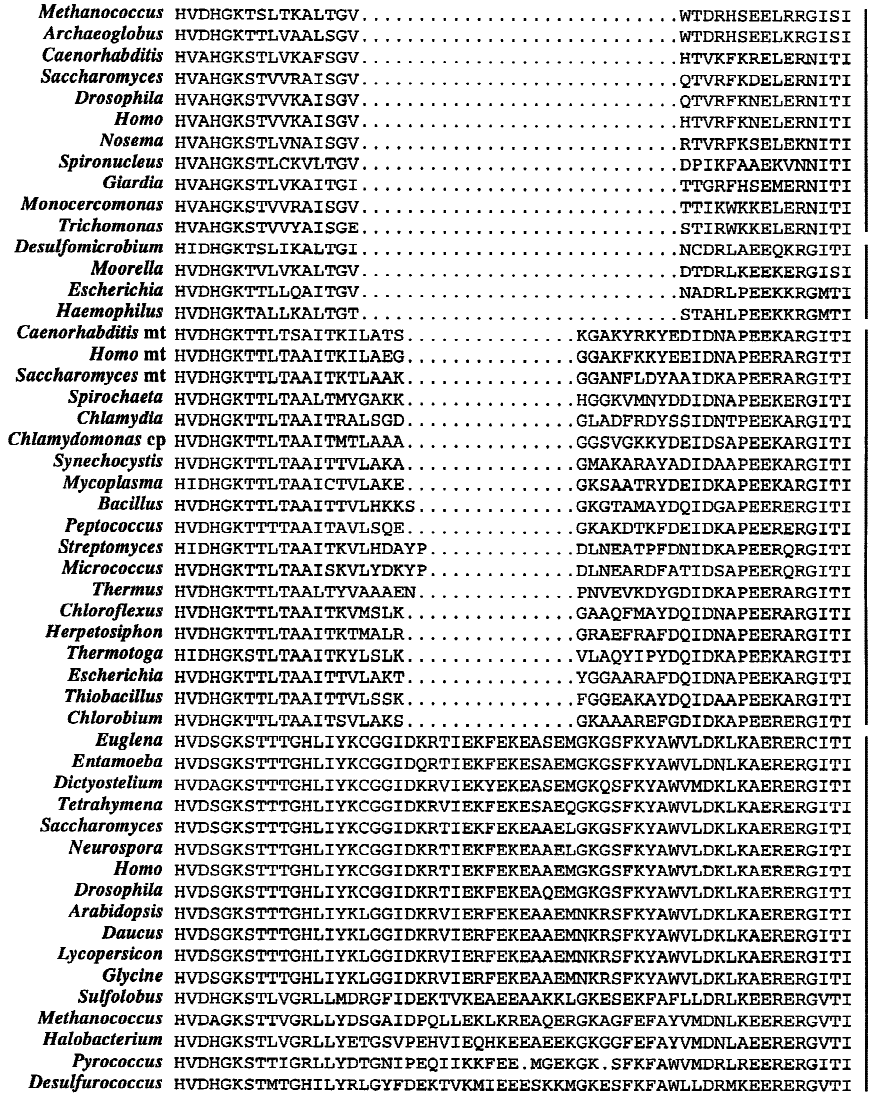

Fig. 4. Partial alignment of the inferred amino acid sequence surrounding the effector loop, showing size variation between translation factor subfamilies. The sizes of the eIF-2γ and SELB effector loops are identical and different from the sizes of all other subfamilies, supporting their relationship to one another. These characters cannot be polarized since the effector loop is variable in EF-G, EF-2, and IF-2 as well.

tralian Research Council (to G.I.M.). P.J.K. was supported by a fellowship from the Natural Sciences and Engineering Research Council of Canada (NSERC). N.M.F. was supported by grants from NSERC and the Killam Trust. We would like to thank A.J. Roger and M. Müller for providing genomic DNA samples, S.L. Baldauf and H.-P. Klenk for sharing unpublished observations, and A. Stoltzfus for the use of his mol2con program. We would also like to thank W.F. Doolittle for supporting this work and for helpful discussions and critical comments on the manuscript. New sequences have been deposited in GenBank under the accession numbers AF005075–AF005078.

References

Adachi J, Hasegawa M (1992) Computer science monographs, No. 27. MOLPHY: Programs for molecular phylogenetics. I. PROTML: Maximum likelihood inference of protein phylogeny. Institute of Stastical Mathematics, Tokyo

Baldauf SL, Palmer JD, Doolittle WF (1996) The root of the universal tree and the origin of eukaryotes based on elongation factor phylogeny. Proc Natl Acad Sci USA 93:7749–7754

Bartig D, Lmekemeier K, Frank J, Lottspeich F, Klink F (1992) The archaebacterial hypusine-containing protein. Eur J Biochem 204: 751–758

Bourne HR, Sanders DA, McCormick F (1991) The GTPase superfamily: Conserved structure and molecular mechanism. Nature 349: 117–127

Bult CJ, White O, Olsen GJ, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM, Clayton RA, Gocacyne JD, Kerlavage AR, Dougherty BA, Tomb JF, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL, Geoghagen NSM, Weidman JF, Fuhrmann JL, Presley EA, Nguyen TR, Utterback JM, Kelley JD, Peterson PW, Sadow MC, Hanna MD, Cotton D, Hurst MA, Roberts KM, Kaine BP, Borodovsky M, Klenk H-P, Fraser CM, Smith HO, Woese CR, Venter JC (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii.* Science 273:1058–73

Edlind TD, Li J, Visvesvara GS, Vodkin MH, McLaughlin GL, Katiyar SK (1996) Phylogenetic analysis of β-tubulin sequences from amitochondrial protozoa. Mol Phylogenet Evol 5:359–367

Felsenstein J (1993) PHYLIP (phylogeny inference package), v. 3.5. University of Washington, Seattle

Forchhammer K, Leinfelder W, Bock A (1989) Identification of a novel translation factor necessary for the incorporation of selenocysteine into protein. Nature 342:453–456

Gaspar NJ, Kinzy TG, Scherer BJ, Humbelin M, Hershey JW, Merrick WC (1994) Translation initiation factor eIF-2. Cloning and expres-

sion of the human cDNA encoding the gamma-subunit. J Biol Chem 269:3415–3422

Germot A, Phillippe H, Le Guyader H (1997) Evidence for loss of mitochondria in microsporidia from a mitochondrial-type HSP70 in *Nosema locustae.* Mol Biochem Parasitol 87:159–168

Gish W, States DJ (1993) Identification of protein coding regions by database similarity search. Nat Genet 3:266–272

Hasegawa M, Kishino H (1994) Accuracies of the simple methods for estimating the bootstrap probability of a maximum likelihood tree. Mol Biol Evol 11:142–145

Hirt RP, Healy B, Vossbrinck CR, Canning EU, Embley TM (1997) A mitochondrial Hsp70 orthologue in *Vairimorpha necatrix:* Molecular evidence that microsporidia once contained mitochondria. Curr Biol 7:995–998

Iwabe N, Kuma K-I, Hasegawa M, Osawa S, Miyata T (1989) Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc Natl Acad Sci USA 86:9355–9359

Jermiin LS, Olsen GJ, Mengersen KL, Easteal S (1997) Majority-rule consensus of phylogenetic trees obtained by maximum-likelihood analysis. Mol Biol Evol 14:1296–1302

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matricies from protein sequences. Comput Appl Biosci 8:275–282

Kamaishi T, Hashimoto T, Nakamura Y, Nakamura F, Murata S, Okada N, Okamoto K-I, Shimzu M, Hasegawa M (1996) Protein phylogeny of translation elongation factor EF-1α suggests microsporidians are extremely ancient eukaryotes. J Mol Evol 42:257–263

Keeling PJ, Doolittle WF (1995a) Archaea: Narrowing the gap between prokaryotes and eukaryotes. Proc Natl Acad Sci USA 92:5761–5764

Keeling PJ, Doolittle WF (1995b) An archaebacterial e1F-1A: new grist for the mill. Mol Microbiol 17:399–400

Keeling PJ, Doolittle WF (1996) Alpha-tubulin from early-diverging eukaryotic lineages and the evolution of the tubulin family. Mol Biol Evol 13:318–326

Keeling PJ, McFadden GI (1998) Origins of microsporidia. Trends Microbiol 6:19–23

Keeling PJ, Baldauf SL, Zillig W, Doolittle WF, Klenk H-P (1996) An *infB*-homologue in *Sulfolobus acidocaldarius.* Syst Appl Microbiol 19:312–321

Klenk HP, Clayton RA, Tomb JF, White O, Nelson KE, Ketchum KA, Dodson RJ, Gwinn M, Hickey EK, Peterson JD, Richardson DL, Kerlavage AR, Graham DE, Kyrpides NC, Fleischmann RD, Quackenbush J, Lee NH, Sutton GG, Gill S, Kirkness EF, Dougherty BA, McKenney K, Adams MD, Loftus B, Peterson S, Reich CI, McNeil LK, Badger JH, Glodek A, Zhou L, Overbeek R, Gocayne JD, Weidman JF, McDonald L, Utterback T, Cotton MD, Spriggs T, Artiach P, Kaine BP, Sykes SM, Sadow PW, D'Andrea KP, Bowman C, Fujii C, Garlan SA, Mason TM, Olsen GJ, Fraser CM, Smith HO, Woese CR, Venter JC (1997) The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus.* Nature 390:364–370

Kozak M (1983) Comparison of initiation of protein synthesis in procaryotes, eukaryotes, and organelles. Microbiol Rev 47:1–45

Kozak M (1992) A consideration of alternative models for the initiation of translation in eukaryotes. Crit Rev Biochem Mol Biol 27:385–402

McCarthy JEG, Brimacombe R (1994) Prokaryotic translation: The interactive pathway leading to initiation. Trends Genet 10:402–407

Merrick WC (1992) Mechanism and regulation of eukaryotic protein synthesis. Microbiol Rev 56:291–315

Pain VM (1996) Initiation of protein synthesis in eukaryotic cells. Eur J Biochem 236:747–771

Shen Q, Leonard JL, Newburger PE (1995a) Structure and function of the selenium translation element in the 3′-untranslated region of human cellular glutathione peroxidase mRNA. RNA 1:519–525.

Shen Q, Leonard JL, Newburger PE (1995b) RNA-binding proteins that specifically recognize the selenocysteine insertion sequence of human cellular glutathione peroxidase mRNA. J Biol Chem 270: 30448–30452

Sutrave P, Shafer BK, Strathern JN, Hughes SH (1994) Isolation, identification and characterization of the *FUN12* gene of *Saccharomyces cerevisiae.* Gene 146:209–213

Swafford DL (1993) PAUP: Phylogenetic analysis using parsimony, v. 3.1.1. Illinois Natural History Survey, Champaign

Vossbrinck CR, Maddox JV, Friedman S, Debrunner-Vossbrinck BA, Woese CR (1987) Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature 326:411–414