

Letter to the Editor

The Crystal Structure of eEF1A Refines the Functional Predictions of an Evolutionary Analysis of Rate Changes Among Elongation Factors

Eric A. Gaucher,* Ujjwal K. Das,* Michael M. Miyamoto,† and Steven A. Benner*

*NASA Astrobiology Institute, Department of Chemistry and †Department of Zoology, University of Florida

The complex interplay between sequence, folded structure, and function in a divergently evolving protein family can be reflected in the rate at which individual sites accumulate amino acid replacements. Thus, detecting shifts in the evolutionary rates of individual sites offers a way to identify potential instances of functional divergence among proteins (Benner 1989; Naylor and Gerstein 2000; Gaucher, Miyamoto, and Benner 2001; Marin et al. 2001; Wang and Gu 2001). The covarion model of Fitch and Markowitz (1970) was among the first to consider protein structure and function in this way. The ability to detect rate-shift sites relies on molecular evolutionary approaches (Gu 1999, 2001; Benner and Gaucher 2001; Galtier 2001; Landgraf, Xenarios, and Eisenberg 2001), whereas functional explanations for these changes can be derived from experiments in structural biology (Golding and Dean 1998). This combination defines a powerful approach for studies within the field of functional genomics in general.

Gaucher, Miyamoto, and Benner (2001) illustrated how such a combined approach can provide experimentally testable hypotheses that relate individual amino acids to specific functional differences among proteins. They analyzed elongation factor (EF) sequences and identified those sites with shifting rates between the EF-Tu and eEF1A orthologs from bacteria and eukaryotes, respectively. Their comparisons relied on maximum likelihood (ML) methods to estimate the evolutionary rates of sites in one group, then in the other. Those sites with the greatest rate changes between EF-Tu and eEF1A were identified from a frequency histogram of their individual differences.

EF-Tu and eEF1A (formerly known as EF-1 α) are essential for translation in bacteria and eukaryotes (Krab and Parmeggiani 1998; Negrutskii and El'skaya 1998). Despite their similar overall roles in translation and very slow rates of evolution, these two orthologous proteins differ in several of their more specific functions. For example, bacterial EF-Tu binds GDP \sim 100-fold tighter than GTP, in contrast to the near-identical affinities of eEF1A for both. EF-Tu regenerates its active form via the single-subunit nucleotide exchange factor EF-Ts. In contrast, eEF1A is regenerated by the multisubunit nucleotide exchange factor eEF1B (formally known as EF-1 $\beta\gamma\delta$). eEF1B is composed of the subunits α , β , and γ ,

with α responsible for eEF1A binding and for the catalytic activity of nucleotide exchange. eEF1A interacts with the actin of the eukaryotic cytoskeleton and may thereby play a role in cellular transformation and apoptosis (Duttaroy et al. 1998). EF-Tu can have no such role in bacteria.

Our previous analysis of rate changes between EFs identified 17 positions that were evolving more rapidly in bacteria than eukaryotes versus 19 sites that were evolving faster in eukaryotes than bacteria (Gaucher, Miyamoto, and Benner 2001). These 36 positions with the greatest rate differences were evaluated for their potential roles in the functional divergence between EFs by mapping them onto the available tertiary structures of bacterial EF-Tu (Nissen et al. 1995; Song et al. 1999). The resultant conclusions were presented as a set of testable hypotheses that awaited further determination of the tertiary structure for eukaryotic eEF1A.

While our paper was in press, Andersen et al. (2000) published the crystal structure of the eEF1B α -bound-state of eEF1A (fig. 1). This new tertiary structure confirms that eEF1A and EF-Tu use the same homologous regions and majority of conserved sites to bind aminoacyl tRNAs (aa-tRNAs). However, the two EFs are now seen to differ markedly in their binding to nucleotide exchange factors (fig. 1). These differences exist even though the mechanism of nucleotide exchange between eEF1A and EF-Tu appears the same (Andersen et al. 2000). EF-Tu binds to EF-Ts through α -helix D and coils at the surfaces of domains 1 and 3, respectively. In contrast, although present in eEF1A, these secondary structural elements are not involved in its binding to eEF1B α . Instead, eEF1A binds to its nucleotide exchange factor through a distinct hydrophobic pocket and other unique contacts on the surfaces of domains 2 and 1, respectively.

The original 36 sites with shifting rates corresponded to those with rate differences of ≥ 2 standard deviations (SD, 1.52 replacements/site/unit evolutionary time) from the mean (-0.03 replacements/site/unit evolutionary time) for all 380 aligned positions of EF-Tu versus eEF1A (Gaucher, Miyamoto, and Benner 2001). Although used in this way, these cutoffs were not viewed as rigorous thresholds of statistical significance but were rather treated as conservative approximations with heuristic value. One obvious reason for this conservative interpretation was that these rate differences were not standardized for their absolute values (Sokal and Rohlf 1981, pp. 417–421). This potential bias can be corrected by log transforming the individual rates before subtracting them between the two groups.

This concern (and others) was addressed in the current study by two different methods. First, the original rate estimates of Gaucher, Miyamoto, and Benner (2001)

Key words: tertiary structures, protein function, evolutionary rates, rate changes, covarions, elongation factors.

Address for correspondence and reprints: Eric A. Gaucher, NASA Astrobiology Institute, Department of Chemistry, 440 Leigh Hall, University of Florida, Gainesville, Florida 32611-7200. E-mail: gaucher@ufl.edu.

Mol. Biol. Evol. 19(4):569–573. 2002

© 2002 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

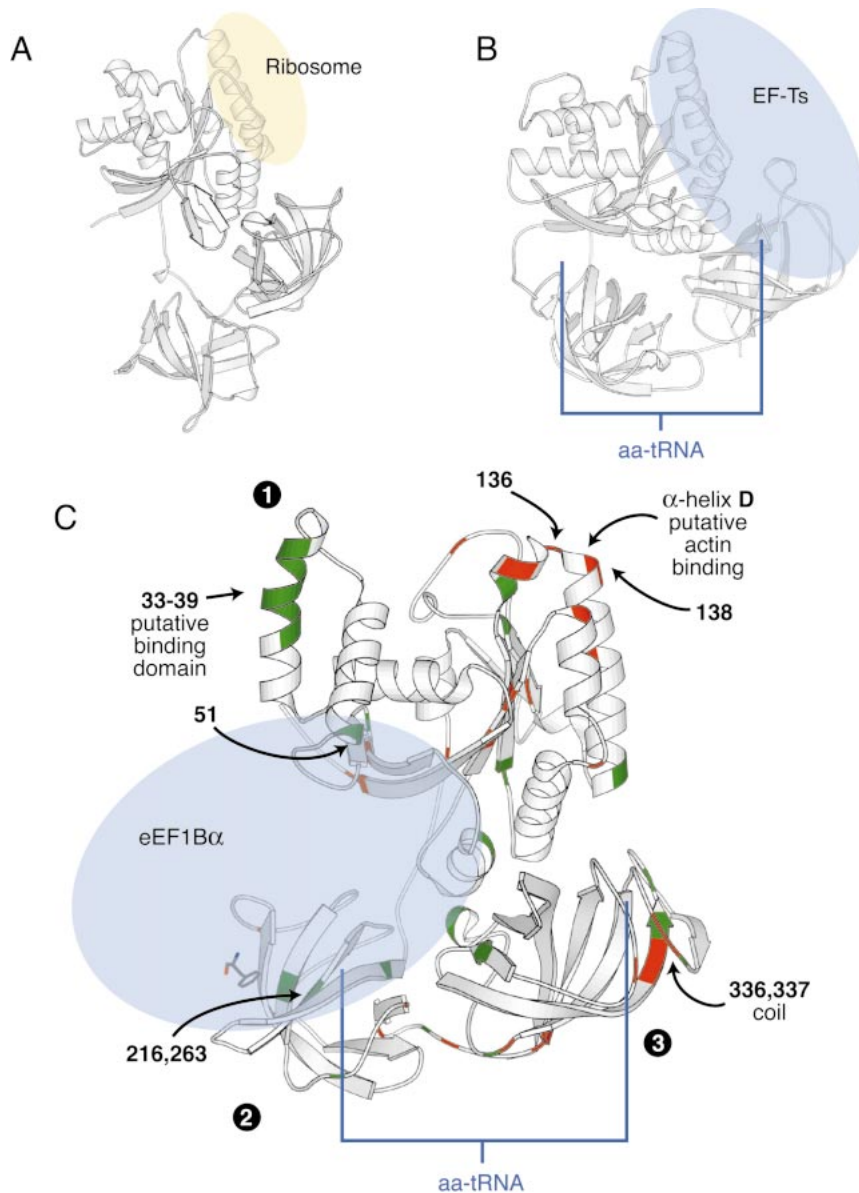


FIG. 1.—Tertiary structures of: (A) EF-Tu bound to GDP:EF-Ts from *Escherichia coli* (Stark et al. 1997; Song et al. 1999); (B) EF-Tu bound to GTP:aa-tRNA from *Thermus aquaticus* (Nissen et al. 1995); and (C) eEF1A bound to eEF1B α from *Saccharomyces cerevisiae* (Andersen et al. 2000). These three orthologues are 393, 405, and 458 residues long, respectively. *Saccharomyces cerevisiae* displays 67% and 68% sequence similarity to *E. coli* and *T. aquaticus*, respectively. Structures are drawn with Molscrip (Kraulis 1991). Green and red identify those sites from the posterior probability analysis, which are evolving slower in eukaryotes than bacteria, and vice versa, respectively (table 1). Those binding regions, secondary structures, and sites that are highlighted in the text are labeled. In (C), the three domains of EFs are numbered and phenylalanine 163 from eEF1B α is shown in stick representation. This residue is highlighted, as it is necessary for interactions with the unique hydrophobic pocket on the surface of eEF1A (Andersen et al. 2000).

were log transformed before reanalysis by their quantile-based method. In this way, 24 sites (rather than 36) were identified with significant rate differences between bacteria and eukaryotes (mean and SD of 0.11 and 1.47 replacements/site/unit evolutionary time for all 380 positions, respectively, table 1). By emphasizing only those sites with the greatest (corrected) rate differences, this approach offered a conservative alternative to the other adopted method (see below).

Next, the EF sequences were analyzed by the posterior probability method of Gu (1999, 2001), using the

same gene phylogeny and ML conditions as in Gaucher, Miyamoto, and Benner (2001). This approach relied on the Jones, Taylor, and Thornton model, with site-to-site heterogeneity in rates according to the gamma distribution, to calculate the posterior probability of a rate-shift site from the replacement differences of bacteria versus eukaryotes. This posterior probability method is based on a solid statistical foundation (Gu 1999, 2001), has been used successfully by its author and others in studies of protein functional divergence (Naylor and Orstein 2000; Wang and Gu 2001), and is now available

Table 1
The 49 Positions of EFs with Posterior Probabilities of Functional Divergence of $\geq 95\%$

Sites	Properties, Function
Evolving Slower in Bacteria	
<u>4</u> , 69 , <u>160</u> , <u>172</u> , <u>288</u> , <u>290</u> , <u>350</u>	Surface, no known function
<u>106</u> , <u>136</u> , <u>144</u> , <u>337</u>	Surface, in proximity to EF-Ts, may be involved in ribosome binding
103 , <u>133</u> , 138 , 336	Surface, all residues bind EF-Ts, may be involved in ribosome binding
189	On loop connecting domains 1 and 2
<u>282</u> , 325–326	Surface, in proximity to aa-tRNA binding
<u>253</u> , <u>277</u> , 305	Surface, all residues bind aa-tRNA
<u>96</u> , <u>98</u>	Interior, structural
Evolving Slower in Eukaryotes	
<u>33</u> , 34 , 35 , 36 , <u>37</u> , <u>38</u> , <u>39</u>	Surface, possible localization sites
<u>131</u>	Surface, helix, binds GTP/GDP
153 , 163	Interior, no known function
<u>51</u> , <u>82</u> , <u>203</u> , <u>263</u> , <u>271</u> , 327 , 329	Surface, in proximity to aa-tRNA binding
<u>67</u> , 123 , <u>216</u> , <u>286</u> , 311 , <u>335</u> , <u>351</u> , <u>357</u>	Surface, possible localization sites

NOTE.—Positions are numbered according to the multiple sequence alignment of Gaucher, Miyamoto, and Benner (2001). The properties and functions of these sites are based on biochemical, cellular, and structural data for bacterial EFs (Krab and Parmeggiani 1998; Negrutskii and El'skaya 1998). Those sites that are also significant according to the log-transformed versus original rate differences are underlined and italicized, respectively. In all, 21 and 28 sites overlap between those from the Bayesian analysis versus log-transformed and original rate differences, respectively. In turn, 17 positions overlap among all three approaches (bold). In addition to the 21 sites, 3 other positions (31, 40, and 42) are significant according to their log-transformed rate differences. In addition to the 28 sites, 8 additional positions (32, 102, 117, 135, 141, 176, 269, and 322) are significant according to their original rate differences (Gaucher, Miyamoto, and Benner 2001).

in a computer package for the analysis of rate-shift sites and structural biology (<http://xgu1.zool.iastate.edu>).

The overall estimate of functional divergence (θ) between EF-Tu and eEF1A was 0.71 (SD = 0.04). As θ ranges from 0 to 1 (with 0 indicating functional conservation), this estimate was indicative of significant functional divergence between the two EFs. A total of 49 sites was then identified with posterior probabilities of $\geq 95\%$ (table 1). Of these 49 sites, 21 overlapped with the 24 from the log-transformed rate differences. This implies that the second set of sites is a subset of the first and that the analysis based on log-transformed rate differences is more conservative than the posterior probability method.

The new tertiary structure of eEF1A:eEF1B α is characterized by a closed compact configuration that is more similar to EF-Tu:GTP:aa-tRNA than to its EF-Tu:GDP:EF-Ts counterpart (fig. 1). This similarity and difference support the hypothesis that eEF1A does not undergo a major conformational change between its active (GTP) and inactive (GDP) states, as does EF-Tu (Negrutskii and El'skaya 1998). This conclusion also identifies a large central pocket of EF-Tu:GDP as a potential site for an antibiotic or drug that is specifically targeted at EF-Tu in bacteria.

The availability of tertiary structures for both eEF1A and EF-Tu now provides an opportunity to more rigorously test the potential functional importance of those sites with significant rate changes. Eight positions with slower rates in bacteria than eukaryotes were identified by Gaucher, Miyamoto, and Benner (2001) from in and around the binding regions of EF-Tu to EF-Ts. These sites were hypothesized as important to the interactions of EF-Tu with EF-Ts but not of eEF1A with eEF1B α . The new posterior probabilities and crystal structure of eEF1A now strengthen this hypothesis. The posterior probability analysis identifies eight sites from the same binding region as above that are evolving

slower in bacteria than eukaryotes (103, 106, 133, 136, 138, 144, 336, and 337) (table 1). In turn, the new tertiary structure of eEF1A maps these eight positions to α -helix **D** and a coil on the surfaces of domains 1 and 3, respectively (fig. 1). However, in contrast to EF-Tu, these secondary structures are not involved in the binding of eEF1A to its nucleotide exchange factor. Thus, these eight positions may be evolving faster in eukaryotes than bacteria because they are under less functional constraint for the binding of eEF1A to eEF1B α .

In addition, at least some of the four sites from α -helix **D** (136, 138, 144, and 153) may be evolving slower in bacteria than eukaryotes because of differences in ribosome binding (table 1 and fig. 1). In EF-Tu, α -helix **D** binds to the L7/L12 stalk of the ribosome, as well as to EF-Ts (Wieden, Wintermeyer, and Rodnina 2001). In contrast, ribosome binding in eEF1A remains only poorly understood. Thus, the four sites from α -helix **D** that are more conserved in bacteria may be under stronger functional constraints because of a unique interaction between EF-Tu and its ribosome. This hypothesis and the previous one for the distinct binding of EF-Tu to EF-Ts are complementary.

Five positions with slower rates in eukaryotes than bacteria (32–36) were predicted to form a unique α -helix at the surface of eEF1A, in combination with or separate from an adjacent insertion (Gaucher, Miyamoto, and Benner 2001). Given its charged and hydrophobic residues, this unique α -helix was assigned a putative binding function in eEF1A. In EF-Tu, no such binding interactions were assigned to these sites, as they were neither conserved nor part of a rigid secondary structure. The posterior probabilities and tertiary structure for eukaryotes reconfirms the status of these five sites (table 1) and document the existence of this α -helix between positions 32–39 in eEF1A (fig. 1). These results support the hypothesis that this unique secondary structure confers distinct binding properties onto eEF1A. Biochemical and structural studies of eEF1A, bound to its ribo-

some and to additional ligands, are now needed to identify which factor(s) interacts with this unique secondary element. Interestingly, the Gag polyprotein of the human immunodeficiency virus interacts specifically with an unknown binding domain of eEF1A somewhere between its first 74 residues (Cimarelli and Luban 1999).

The posterior probability analysis and tertiary structure of eEF1A corroborate the above hypotheses of functional divergence between EFs. In turn, they also support the refinement of other hypotheses (Gaucher, Miyamoto, and Benner 2001). Three sites with slower rates in eukaryotes than bacteria (51, 216, and 263) were identified from around the binding regions of EF-Tu to aa-tRNA. Thus, these three positions were hypothesized as important in the interactions of eEF1A (but not EF-Tu) with this ligand. Although the posterior probability analysis reconfirms the status of these sites (table 1), the recent tertiary structure of eEF1A shows that these three positions contribute to a salt bridge and reside within a unique hydrophobic pocket at the surfaces of domains 1 and 2 for eEF1B α binding, respectively (fig. 1). Thus, the new crystal structure of eEF1A supports a modified version of the original hypothesis that these three sites are evolving slower in eukaryotes because of functional constraints from aa-tRNA binding. Rather, it documents that these slower rates are a consequence of the unique interactions of eEF1A with its nucleotide exchange factor.

α -helix **D** is conserved in both EF-Tu and eEF1A, even though it is not involved in the binding of the latter to eEF1B α . Instead, this conserved structure on the surface of both EFs may be responsible for the binding of eEF1A to the actin of the eukaryotic cytoskeleton (Gaucher, Miyamoto, and Benner 2001). This possibility is supported by the sequence similarity between α -helix **D** of eEF1A and the actin-binding region of depatin (Yang et al. 1990). eEF1A occurs at both the nucleus and ribosomes and binds both charged and uncharged tRNAs (Negrutskii and El'skaya 1998). Taken together, these arguments raise the intriguing corollary that these sites may be responsible in eukaryotes for the ability of eEF1A to channel tRNAs between the nucleus and ribosomes (Grosshans, Simos, and Hurt 2000).

The new crystal structure of eEF1A allows for stronger hypotheses about the importance of particular sites in the functional divergence of EFs. In turn, these refined hypotheses can now be directly tested with standard molecular biology techniques such as site-directed mutagenesis and yeast 2-hybrids. These laboratory experiments complete an integrated research program that holds considerable promise for the resolution of difficult problems in functional genomics.

Acknowledgments

We thank R. A. Kiltie, D. L. Swofford, M. R. Tennant, and X. Gu for their suggestions about our research. This work was supported by a post-doctoral fellowship from the National Research Council and NASA Astrobiology to E.A.G. by the Department of Zoology, Uni-

versity of Florida, and by a grant from NASA Exobiology to S.A.B.

LITERATURE CITED

- ANDERSEN, G. R., L. PEDERSEN, L. VALENTE, I. CHATTERJEE, T. G. KINZY, M. KJELDGAARD, and J. NYBORG. 2000. Structural basis for nucleotide exchange and competition with tRNA in the yeast elongation factor complex eEF1A, eEF1B α . *Mol. Cell* **6**:1261–1266.
- BENNER, S. A. 1989. Patterns of divergence in homologous proteins as indicators of tertiary and quaternary structure. *Adv. Enzyme Regul.* **28**:219–236.
- BENNER, S. A., and E. A. GAUCHER. 2001. Evolution, language and analogy in functional genomics. *Trends Genet.* **17**:414–418.
- CIMARELLI, A., and J. LUBAN. 1999. Translation elongation factor 1-alpha interacts specifically with the human immunodeficiency virus type 1 Gag polyprotein. *J. Virol.* **73**:5388–5401.
- DUTTARROY, A., D. BOURBEAU, X. L. WANG, and E. WANG. 1998. Apoptosis rate can be accelerated or decelerated by overexpression or reduction of the level of elongation factor-1 alpha. *Exp. Cell Res.* **238**:168–176.
- FITCH, W. M., and E. MARKOWITZ. 1970. An improved method for determining codon variability in a gene and its application to the rate of fixation of mutations in evolution. *Biochem. Genet.* **4**:579–593.
- GALTIER, N. 2001. Maximum-likelihood phylogenetic analysis under a covarion-like model. *Mol. Biol. Evol.* **18**:866–873.
- GAUCHER, E. A., M. M. MIYAMOTO, and S. A. BENNER. 2001. Function-structure analysis of proteins using covarion-based evolutionary approaches: elongation factors. *Proc. Natl. Acad. Sci. USA* **98**:548–552.
- GOLDING, G. B., and A. M. DEAN. 1998. The structural basis of molecular adaptation. *Mol. Biol. Evol.* **15**:355–369.
- GROSSHANS, H., G. SIMOS, and E. HURT. 2000. Transport of tRNA out of the nucleus—direct channeling to the ribosome? *J. Struct. Biol.* **129**:288–294.
- GU, X. 1999. Statistical methods for testing functional divergence after gene duplication. *Mol. Biol. Evol.* **16**:1664–1674.
- . 2001. Maximum-likelihood approach for gene family evolution under functional divergence. *Mol. Biol. Evol.* **18**:453–464.
- KRAB, I. M., and A. PARMEGGIANI. 1998. EF-Tu, a GTPase odyssey. *Biochim. Biophys. Acta* **1443**:1–22.
- KRAULIS, P. J. 1991. Molscript—a program to produce both detailed and schematic plots of proteins structures. *J. Appl. Crystallogr.* **24**:946–950.
- LANDGRAF, R., I. XENARIOS, and D. EISENBERG. 2001. Three-dimensional cluster analysis identifies interfaces and functional residue clusters in proteins. *J. Mol. Biol.* **307**:1487–1502.
- MARIN, I., M. A. FARES, F. GONZALEZ-CANDELAS, E. BARRIO, and A. MOYA. 2001. Detecting changes in the functional constraints of paralogous genes. *J. Mol. Evol.* **52**:17–28.
- NAYLOR, G. J., and M. GERSTEIN. 2000. Measuring shifts in function and evolutionary opportunity using variability profiles: a case study of the globins. *J. Mol. Evol.* **51**:223–233.
- NEGRUTSKII, B. S., and A. V. EL'SKAYA. 1998. Eukaryotic translation elongation factor 1-alpha, structure, expression, functions, and possible role in aminoacyl-tRNA channeling. *Prog. Nucleic Acid Res. Mol. Biol.* **60**:47–78.
- NISSEN, P., M. KJELDGAARD, S. THIRUP, G. POLEKHINA, L. RESHETNIKOVA, B. F. CLARK, and J. NYBORG. 1995. Crystal

- structure of the ternary complex of Phe-tRNA(Phe), EF-Tu, and a GTP analog. *Science* **270**:1464–1472.
- SOKAL, R. R., and F. J. ROHLF. 1981. *Biometry*. 2nd edition. Freeman, New York.
- SONG, H., M. R. PARSONS, S. ROWSELL, G. LEONARD, and S. E. V. PHILIPS. 1999. Crystal structure of intact elongation factor EF-Tu from *Escherichia coli* in GDP conformation at 2.05 angstrom resolution. *J. Mol. Biol.* **285**:1245–1256.
- STARK, H., M. V. RODNINA, J. RINKE-APPEL, R. BRIMACOMBE, W. WINTERMEYER, and M. VAN HEEL. 1997. Visualization of elongation factor Tu on the *Escherichia coli* ribosome. *Nature* **389**:403–406.
- WANG, Y. F., and X. GU. 2001. Functional divergence in the caspase gene family and altered functional constraints: statistical analysis and prediction. *Genetics* **158**:1311–1320.
- WIEDEN, H.-J., W. WINTERMEYER, and M. V. RODNINA. 2001. A common structural motif in elongation factor Ts and ribosomal protein L7/L12 may be involved in the interaction with elongation factor Tu. *J. Mol. Evol.* **52**:129–136.
- YANG, F., M. DEMMA, V. WARREN, S. DHARMAWARDHANE, and J. CONDEELIS. 1990. Identification of an actin-binding protein from *Dictyostelium* as elongation factor 1a. *Nature* **347**:494–496.

RODNEY HONEYCUTT, reviewing editor

Accepted October 19, 2001