FEBS Letters 580 (2006) 1672–1676

Optimal growth temperature of prokaryotes correlates with class II amino acid composition

Liron Klipcan[a], Ilya Safro[b], Boris Temkin[b], Mark Safro[a,*]

[a] Department of Structural Biology, Weizmann Institute of Science, 76100 Rehovot, Israel
[b] Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, 76100 Rehovot, Israel

Received 17 November 2005; revised 18 January 2006; accepted 7 February 2006

Available online 17 February 2006

Edited by Robert B. Russell

Abstract Partitioning of aminoacyl-tRNA synthetases and their associated amino acids into two classes allows us to distinguish between thermophilic and mesophilic species based only on amino acids composition. The CLASSDB program has been developed for amino acid content analysis in organisms treated individually or pooled together to form a pattern of characteristic properties. A strong correlation has been observed between optimal growth temperature (OGT) of organisms and class II amino acids content. Amino acid composition in organisms closely related phylogenetically but dissimilar in their OGT testifies that thermo-adaptation happens rather rapidly on the time scale of evolution.
© 2006 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Class II aaRSs; Amino acids; Thermophiles; Mesophiles; Evolution; Thermal adaptation

1. Introduction

The aminoacyl-tRNA synthetases (aaRSs) family comprises 20 ancient enzymes that covalently attach amino acids to the corresponding nucleic acid adaptor molecules, tRNA, prior to polypeptide-chain synthesis on ribosome. The utmost importance of aaRSs in the fidelity of the genetic code translation implies that members of the family are probably among the earliest proteins to appear. Surprisingly, the 20 aaRSs are partitioned into two classes of 10 enzymes each [1]. Class I representatives exhibit catalytic domains containing the classical nucleotide-binding Rossmann fold and catalyze aminoacylation reaction of Trp, Tyr, Gln, Glu, Lys-1, Val, Ile, Leu, Met, Arg and Cys substrate amino acids. Class II aaRSs exhibit catalytic domains built around an antiparallel β-sheet flanked by α-helices and are associated with Pro, Thr, Ser, Asp, Asn, Lys-2, His, Ala, Gly and Phe substrate amino acids.

What lies at the heart of the aaRSs separation into two structurally unrelated classes? The remarkable symmetry of the aaRS class division, sequence analyses and assumed dual recognition of opposite sides of the acceptor stems of tRNAs by pairs of class I and class II aaRSs (IleRS and ThrRS; GlnRS and AspRS; TyrRS and PheRS; etc.) led to the hypothesis that distinctions between classes evolves from a gene whose sense and antisense strands code for ancestors of two ancient aaRS classes [2–4]. Class II aaRSs are believed to have roots stretching back into antiquity further than class I enzymes, as they predominantly aminoacylate tRNA with simpler and small amino acids [5,6], which play crucial roles in the formation of protein folds. These amino acids are: Gly, Pro, Ala, Asp, Thr and Ser. Remarkably, these are the same amino acids that were abundant in the classical Miller experiments which attempted to imitate primordial environments [7–9].

As noticed before [10], the biosynthetic pathways of class II aaRSs related amino acids usually involve less intervening stages. Amino acids Ala, Asp, Gly, Ser, Thr and class I Glu, characterized by simpler pathways and play the role of metabolic precursors for other amino acids. Moreover, the GCA rich codons, which apparently are related to ancient codons, are more frequently associated with class II aaRSs [10]. Ancient proteins probably contained a larger proportion of class II aaRSs associated amino acids than the recent ones. Ferredoxin, as it was proposed by Eck and Dayhoff [11], has evolved from repeating sequences patterns of Ala, Asp, Pro, Ser, and Gly, all of which are class II related. However, modern ferredoxin from *Clostridium pastereurianum* contains only 56% of amino acids associated with class II aaRSs. Most likely, ancient signatures of proteins were lost during the long period of evolution. Moreover, analysis of amino acids substitutions recently proposed that currently declining amino acids were among the first to incorporate into the genetic code while the gainer amino acids were added in a later stage [12].

A number of thermophile-specific patterns at the level of nucleotide content, codon usage and amino acid composition have been presented in recent publications [13–16]. Partitioning of aaRSs and their associated amino acids into two natural classes allowed us to distinguish between thermophilic (Bacteria and Archaea) and mesophilic species from different organisms based on amino acids composition. We have also demonstrated that a given clustering of amino acids correlates with already established phylogeny of three primary superkingdoms (Archaea, Bacteria, Eukarya). Further, we revealed the relationship between the optimal growth temperature (OGT) and class II amino acid composition of prokaryotic organisms.

2. Materials and methods

The program CLASSDB is implemented on the Unix platform using Java and Perl with WEB-based Graphic User Interface and is designed for content analysis of different sets of amino acids in protein

*Corresponding author. Fax: +972 8 934 41 36.
E-mail address: mark.safro@weizmann.ac.il (M. Safro).

Abbreviations: aaRSs, aminoacyl-tRNA synthetases; OGT, optimal growth temperature; LUCA, last universal common ancestor

0014-5793/$32.00 © 2006 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.febslet.2006.02.013

sequences (http://safro.weizmann.ac.il/8080). The interface is interactive as it provides an input window, suitable for pasting protein sequences, designed to send amino acid sequences to the program for analysis. Java environment access is required to start the program. The program incorporates two blocks: search engine for the proteins databases and statistical tools. At a preprocessing stage, the application requires a data file from Swiss-Prot or TrEMBL database and converts it into several indexes. The basic operation in this environment is setting up a query that extracts a subset of records from the database. The query is based on the regular logical expression matching database identifiers (naming conventions, source of the organism, sequence features and other options in various logical combinations). The search is performed over database descriptor fields and includes analysis of the logical expressions. Prior to calculations, amino acids subjected to statistical analysis have to be partitioned into two groups. These groups formed by a user: charged and non-charged, or class I and class II associated amino acids, or hydrophilic and hydrophobic, etc. The required sequences can be analyzed in the range of their lengths from MaxLength to MinLength. Each entry includes a checkbox that allows editing or removal of the entry from the pool of the results. The data can be processed using several statistical applications that have built-in textual and graphical representations of the subset's content. All of the above-mentioned functions can be applied to the user's query sequences as well. The application allows storage of both the textual and graphical data on the disk and their printing.

The data on amino acid composition of fully sequenced genomes was obtained from: http://www.ebi.ac.uk/proteome/index.html. The data on OGTs was collected from German Collection of Microorganisms and Cell Cultures (http://www.dsmz.de/).

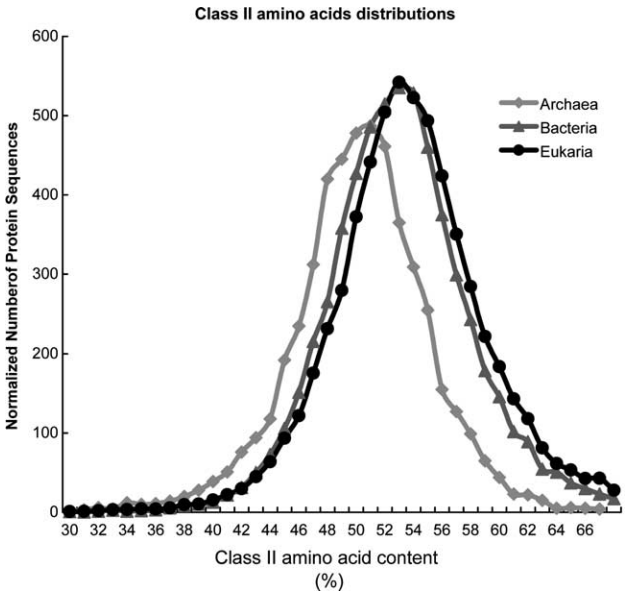

Fig. 1. Class II amino acid content in Archaea, Bacteria and Eukarya. The general distribution of class II amino acids among 5032 Archaeal proteins (marked in rhombuses), 35802 Bacterial proteins (triangles) and 61649 Eukaryotic proteins (circles). The number of proteins in Eukarya and Bacteria are normalized to those (and number of class II associated amino acids accordingly) in Archaea.

3. Results

3.1. Predominance of class II amino acids in mesophiles

The SwissProt [17] database has been tested by CLASSDB for the presence of class II amino acids in various protein sequences. The distributions of class II amino acids among three domains of life Archaea, Bacteria and Eukarya have been calculated. The 61649 protein sequences in Eukarya exhibit the distribution maximum at 54% of class II amino acids content, while 35802 sequences in Bacteria demonstrate maximum at 53%. The 5032 proteins of Archaea exhibits similar distributions of both classes, i.e., ~50% (see Fig. 1). An interesting application is related to the group of ribosomal proteins. Although ribosomal proteins are found to be different among the three primary kingdoms, a considerable number of them are identical in different organisms. Analysis has been carried out on 586 ribosomal proteins of Archaea, 1937 of Bacteria, and 2470 of Eukarya. The analysis exhibits distributions of class II amino acids that are similar to those observed for the entire domains of life and suggests that even individual groups of proteins still preserve characteristic amino acids content and allow distinguishing between different organisms. To eliminate possible artifacts, resulting from unequal number of fundamental folds among kingdoms, the limited set of universally conserved proteins among the three major kingdoms [18,19] has been constructed. Preliminary screening revealed that many thermophilic bacteria display class II amino acids content very similar to Archaea, predominantly represented by thermophilic species in various databases. This gives grounds to separate thermophilic and hyperthermophilic proteins, functioning in Archaea and Bacteria at high temperatures (over 50 °C), into one subgroup. A major portion of bacterial proteins (except for those that are classified as thermophiles) and essentially all proteins in Eukarya are referred to as mesophilic. The subsets have been tested with CLASSDB and the results are presented in Table 1. It has been observed that mesophilic organisms are enriched with class II amino acids, while thermophilic enzymes contain a higher proportion of class I amino acids. Statistical data for pyruvate kinase further illustrates in graphical form predominance of class II amino acids in mesophiles (see Fig. 2).

3.2. Interplay of the amino acid content and OGT

Results described previously [20] appear to propose a different content of individual amino acids in organisms closely related phylogenetically but distinguished by OGT. An additional analysis has been performed on phylogenetically closely related organisms but distinctive in their OGT. The Archaea genus *Methanococcus* are anaerobic methanogens that vary in OGT remaining closely related by virtue of the rRNA phylogenetic tree [21]. *Methanococcus jannaschii* originally isolated from a hydrothermal vent has OGT of 85 °C, while *Methanococcus vannielii* and *Methanococcus voltae* have OGT around 37 °C. It would be reasonable to expect similar distributions of class II content in proteins closely related phylogenetically. Surprisingly, we have found that average content of class II amino acids among 67 proteins from *M. voltae* and 65 from *M. vannielii* is ~53.5%, very similar to Eukarya (~54%) and/or mesophilic Bacteria (~53%), while *M. jannaschii* proteome revealed class II content of ~49%. In addition, moderately mesophilic Archaea proteins demonstrate a Eukaryotic-like distribution (data not shown). Similarly, we analyzed class II content of 36 Bacterial and Archaeal organisms, the OGTs of which are located in different temperature zones (Table 2). As it follows from Fig. 3, a strong correlation has been observed between OGT and class II amino acids composition: a decrease in the proportion of class II amino acids in the organisms is observed, as the OGT rises. To make an estimate of the slope of a given regression, we constructed the similar graphs for ten subsets of amino acids randomly partitioned

Table 1
Class II amino acid content for proteins that *Universally Conserved* between three domains of life

Enzymes	Thermophiles (%)	Bacteria (%)	Eukarya (%)
AlaRS	49.25	54.68	56.40
ThrRS	48.10	50.80	53.57
ArgRS	49.00	52.52	52.00
ValRS	48.80	52.00	53.20
SerRS	47.01	51.41	51.41
S5 ribosomal protein	51.36	57.32	52.38
Thymidylate synthase	49.60	52.47	53.70
Enolase	53.50	56.75	57.55
Triose-phosphate isomerase	53.09	54.72	55.70
Glyceraldehyde 3-phosphate dehydrogenase	53.50	58.40	60.30
Glucosamine-fructose-6-phosphate aminotransferase	51.40	51.50	51.50
Pyruvate-kinase	53.00	58.08	58.98

Thermophiles from Archaea and Bacteria are partitioned into one subgroup. The proteins that belong to Bacteria and Eukarya do not include the thermophilic ones.

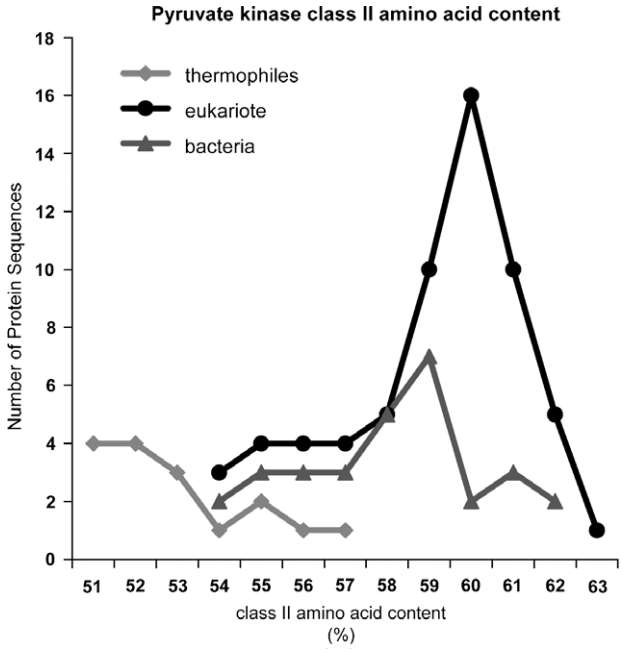

Fig. 2. Class II amino acid distribution in Pyruvate kinase sequences are isolated from different sources. Thermophilic proteins from Archaea and Bacteria were partitioned into one subgroup. The sequences of thermophilic Pyruvate kinase are marked in rhombuses, those from Bacteria (minus thermophiles) marked by triangles and those from Eukarya presented by circles.

into two artificial classes: ''class II-art'' and ''class I-art''. To be consistent with the natural partitioning, we retain ten amino acids for each of the classes. By picking nine arbitrary amino acids of class II coupled with one amino acid of class I, we create a new ''class II-art'' regression presented on graph 1; eight amino acids of class II coupled with those two of class I give rise to another distribution presented in graph 2, etc. Within the subsets of ''class II-art'' we changed only those amino acids that belong to the class I. Thus, graph 1 (see Fig. 4) represents an average over the ten straight lines with different slopes; each line was constructed from the subset that contains nine invariant amino acids of class II and any one amino acid of class I. The graph 2 in Fig. 4 represents an average over the 45 different straight lines (the number of combinations of two amino acids out of a pool of ten that belong to class I) with different slopes. Subsequently for graphs 3, 4, etc., we have limited the number of the averaged subsets by 100 different combinations. Taken together, these graphs comprise 755 random regressions. Broadly speaking, the artificially induced partitioning into two classes can produce OGT-class II content relationship similar to that observed in Fig. 3. However, for the most part, random distributions failed to reveal correlation similar to the natural class I and II partitioning.

Table 2
OGT of selected prokaryotic organisms

Organism	OGT (°C)
Methanopyrus kandleri (A)	98
Pyrococcus abyssi (A)	98
Pyrobaculum aerophilum (A)	98
Pyrococcus furiosus (A)	98
Pyrococcus horikoshi (A)	95
Aeropyrum pernix (A)	90
Archeoglobus fulgidus (A)	85
Methanococcus jan (A)	82
Sulfolobus solfataricus (A)	75
Sulfolobus tokodaii (A)	75
Thermoplasma acidophilum (A)	60
Methanobacterium thermoautotrophicum (A)	60
Thermoplasma volcanium (A)	60
Methanosarchina acetivorans (A)	40
Methanococcus vannielii (A)[a]	37
Methanococcus voltae (A)[a]	37
Haloarcula marismortiui (A)[a]	37
Halobacterium sp. (A)	37
Methanosarchina mazei (A)	37
Aquifex aeolicus (B)	85
Thermotoga maritma (B)	80
Thermoanarebacter tengcongensis (B)	75
Thermus aquaticus (B)[a]	70
Chlorobium tepidum (B)	47
Campylobacter jejuni (B)	37
Clostiridium acetobutylicum (B)	37
Escherichia coli (B)	37
Fusobacterium nucleatum (B)	37
Haemophilus influenzae (B)	37
Helicobacter pylori (B) 26695	37
Bacillus subtilis (B)	30
Deinococcus radiodurans (B)	30
Agrobacterium tumefaciens (B)	26
Xylella fastidiosa (B) 9a5c	26
Xanthomonas campestris (B)	26
Synechocystis sp. (B)	25

[a]Organisms from partially sequenced genomes; A – organisms from Archaea; B – organisms from Bacteria.

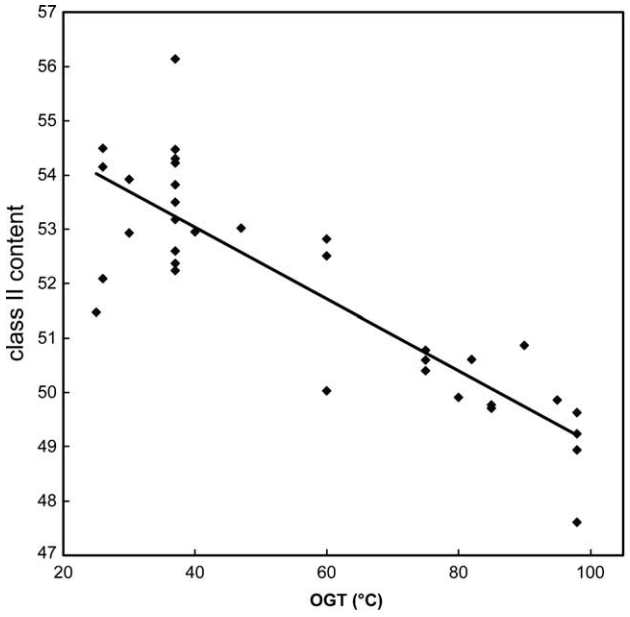

Fig. 3. The graphical representation of the observed correlation between the OGT and class II amino acid content in prokaryotic organisms of Archaea and Bacteria.

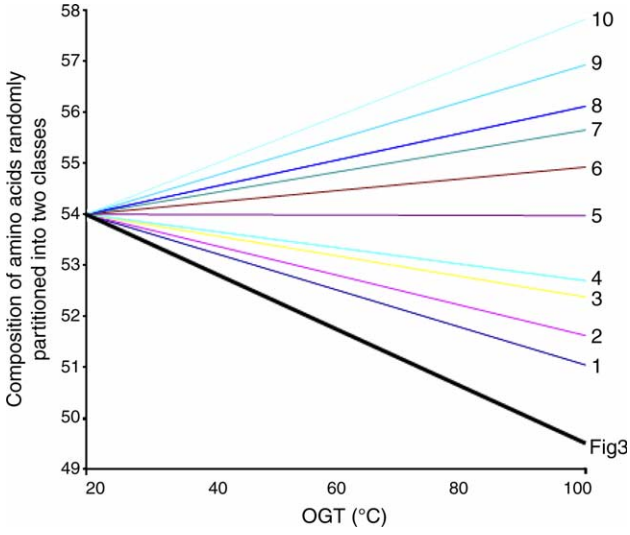

Fig. 4. The graphical representation of the relationships between the OGT and subsets of amino acids randomly distributed into two new classes "class II-art" and "class I-art" in prokaryotic organisms of Archaea and Bacteria. The graphical representation of the regression corresponding to the natural partition into classes I and II is marked as Fig. 3. Straight line numbered 1 is associated with partition into classes when random new "class II-art" contains nine amino acids of natural original class II and one amino acid from class I (9 and 1), line 2 (8 and 2), line 3 (7 and 3), etc. Each line numbered 1, 2, 3, etc., represents an average (up to 100) over regressions where class II amino acids are invariant while class I amino acids vary. All the regressions have been shifted to the common origin to display the differences in their slopes.

4. Discussion

Availability of large number of genomes from organisms that occupy different ecological niches provides new ways for analysis of adaptation processes on molecular level. In particular, worthy of mention are the attempts to trace a tendency of amino acids usage in the organisms that have been adapted to different temperature ranges.

Amino acid composition analysis has been recently used to build phylogenetic tree and to distinguish between thermophilic and mesophilic prokaryotes [14]. Natural classifications of amino acids associated with two classes of aaRSs demonstrate novel and simple approach for study of genome signatures. Suggested classification allows us to distinguish between Archaea, Bacteria and Eukarya based on the amino acids content only. We also show that such analysis is robust even for reduced set of non-homologous proteins.

It is of interest that thermophilic Bacteria and Archaea which in general do not represent one evolutionary group display similar distributions of class II amino acids. As it was pointed by Pe'er et al. [22], amino acid composition analysis also show similar clustering of thermophilic bacteria and archaea. This stems from two probable reasons: (a) massive horizontal gene exchange between the two kingdoms [23] that tangles up the lines in family trees; (b) the proximity of thermophilic bacteria and archaea to the root of the universal phylogenetic tree as it followed from the sequence analysis of 16S rRNA [24]. It is surprising that hyperthermophilic Bacteria and Archaea being located on the deepest branches of rRNA phylogenetic tree, and thus supposed to be ancient, have the predominance of complex amino acids. Yet hyperthermophilic last universal common ancestor (LUCA) hypothesis [25,26] has recently been challenged using a novel method of phylogenetic tree reconstruction [27] and observation that the GC content of rRNA sequences of the LUCA does not support thermophilic origin of life [28]. At present it is generally agreed that thermal stability may be achieved with a number of subtle changes in local weak interactions without significant conformational rearrangements in basic topologies of protein folds [29]. Thus, retaining the set of amino acids that determine the characteristic protein folds, and these are mostly the class II amino acids, proteins can shift their stability profile to thermophilic temperature range. Biasing of statistically significant number of thermophilic proteins (see Fig. 4) by class I amino acids testifies that thermal adaptation is sequenced-based rather than structure-based process. What is more, thermo-adaptation is a rather rapid process on the time scale of evolution since differences in amino acid content have been observed in organisms closely related phylogenetically but distinctive in their OGT (e.g., within genus *Methanococcus* and *Thermus*/*Deinococcus* group).

Amino acid composition analysis testifies that mesophilic organisms are enriched with class II amino acids while predominance of class I associated amino acids in thermophiles has been detected (see Fig. 3). Investigations have been undertaken towards the adaptation processes that exhibit some preferences in amino acid composition of thermophilic proteins [30–32]. Specifically, amino acids that are abundant in thermophiles are Glu, Arg, Lys, Pro, Tyr, Ile and Leu. Most of these amino acids are associated with class I aaRSs.

All Archea, many Bacteria, and the organelles of eukaryotes lack the canonical GlnRS and AsnRS [33]. However within these domains there is a possibility to charge $tRNA^{Asn}$ and $tRNA^{Gln}$ with Asn and Gln, respectively, by using tRNA-dependent amidotransferases [34]. The absence of these aaRSs in Archaea correlates with instability of amino acids Gln and Asn at high temperatures. In addition, genome sequences of some thermophilic Archaea do not contain genes encoding class I CysRS that is consistent with decreasing of cysteine

content in thermophiles. This paradoxical situation was resolved by revealing of *O*-phosphoseryl-tRNA synthetase SepRS, displaying amino acid sequence similarity to α-subunit of PheRS, class II enzyme and capable of charging $tRNA^{Cys}$ with Cys [35]. However, narrow phylogenetic distribution of SepRS makes it unclear whether this enzyme recently diverged from PheRS or, instead, coevolved with PheRS from a common ancestor [35]. Further calculations may then carry out to estimate contribution of specific organisms lacking conventional pathways of tRNA charging in class I and class II amino acid distributions.

Acknowledgements: We are very grateful to E. Trifonov for reading of the early version of the manuscript and valuable discussion and H. Philippe for critical comments and advice. We thank E. Segal for her help in preparation of the manuscript and D. Safro for technical assistance. Supported by the Kimmelman Center for Biomolecular Structure and Assembly. M.S. was aided by a grant from Israel Science Foundation (No. 1034/03-16.1). M.S. holds The Lee and William Abramowitz Professorial Chair of Macromolecular Biophysics.

References

[1] Eriani, G., Delarue, M., Poch, O., Gangloff, J. and Moras, D. (1990) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. Nature 47, 203–206.
[2] Carter Jr., C.W. and Duax, W.L. (2002) Did tRNA synthetase classes arise on opposite strands of the same gene. Mol. Cell. 10, 705–708.
[3] Ribas de Pouplana, L. and Schimmel, P. (2001) Aminoacyl-tRNA synthetases: potential markers of genetic code development. Trends Biochem. Sci. 26, 591–596.
[4] Rodin, S.N. and Ohno, S. (1997) Four primordial modes of tRNA-synthetase recognition, determined by the (G,C) operational code. Proc. Natl. Acad. Sci. USA 94, 5183–5188.
[5] Carter Jr., C.W. (1993) Cognition, mechanism, and evolutionary relationships in aminoacyl-tRNA synthetases. Annu. Rev. Biochem. 62, 715–748.
[6] Hartman, H. (1995) Speculations on the origin of the genetic code. J. Mol. Evol. 40, 541–544.
[7] Miller, S. (1987) Which organic compounds could have occurred on the prebiotic earth. Cold Spring Harb. Symp. Quant. Biol. 52, 17–27.
[8] Trifonov, E.N. and Berezovsky, I. (2002) Molecular evolution from abiotic scratch. FEBS Lett. 527, 1–4.
[9] Trifonov, E. (2004) The triplet code from first principles. J. Biomol. Struct. Dyn. 22, 1–11.
[10] Klipcan, L. and Safro, M. (2004) Amino acid biogenesis, evolution of the genetic code and aminoacyl-tRNA synthetases. J. Theor. Biol. 228, 389–396.
[11] Eck, R. and Dayhoff, M. (1966) Evolution of the structure of ferredoxin based on living relicts of primitive amino acid sequences. Science 152, 363–366.
[12] Jordan, I., Kondrashov, F., Adzhubei, I., Wolf, Y., Koonin, E., Kondrashov, A. and Sunyaev, S. (2005) A universal trend of amino acids gain and loss in protein evolution. Nature 433, 633–637.
[13] Fukuchi, S. and Nishikawa, K. (2001) Protein surface amino acid compositions distinctively differ between thermophilic and mesophilic bacteria. J. Mol. Biol. 309, 835–843.
[14] Kreil, D. and Ouzounis, C. (2001) Identification of thermophilic species by the amino acid compositions deduced from their genomes. Nucleic Acids Res. 29, 1608–1615.
[15] De Farias, S. and Bonato, M. (2002) Preferred codons and amino acid couples in hyperthermophiles. Genome Biol., 3, PREPRINT0006.
[16] Singer, G. and Hickey, D. (2003) Thermophilic prokaryotes have characteristic patterns of codon usage, amino acid composition and nucleotide content. Gene 317, 39–47.
[17] Apweiler, R. (2000) Protein sequence databases (Bork, P., Ed.), Advances in Protein Chemistry. Analysis of Amino Acid Sequences, vol. 54, pp. 31–71, Academic Press, New York.
[18] Peregrin-Alvarez, J., Tsoka, S. and Ouzounis, C. (2003) The phylogenetic extent of metabolic enzymes and pathways. Genome Res. 13, 422–427.
[19] Dandekar, T., Schuster, S., Snel, B., Huynen, M. and Bork, P. (1999) Pathway alignment: application to the comparative analysis of glycolytic enzymes. Biochem. J. 343, 115–124.
[20] McDonald, J., Grasso, A. and Rejto, L. (1999) Patterns of temperature adaptation in proteins from *Methanococcus* and *Bacillus*. Mol. Biol. Evol. 16, 1785–1790.
[21] McCloskey, J.A., Graham, D.E., Zhou, S., Crain, P.F., Ibba, M., Konisky, J., Soll, D. and Olsen, G.J. (2001) Post-transcriptional modification in archaeal tRNAs: identities and phylogenetic relations of nucleotides from mesophilic and hyperthermophilic *Methanococcales*. Nucleic Acids Res. 29, 4699–4706.
[22] Pe'er, I., Felder, C.E., Man, O., Silman, I., Sussman, J.L. and Beckmann, J.S. (2004) Proteomic signatures: amino acid and oligopeptide compositions differentiate among phyla. Proteins 54, 20–40.
[23] Nelson, K., Clayton, R., Gill, S., Gwinn, M., Dodson, R., Haft, D., Hickey, E. and Peterson, J. (1999) Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*. Nature 399, 323–329.
[24] Woese, C.R., Kandler, O. and Wheelis, M. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. USA 87, 4576–4579.
[25] Di Giulio, M. (2003) The universal ancestor and the ancestor of bacteria were hyperthermophiles. J. Mol. Evol. 57, 721–730.
[26] Brooks, D.J., Fresco, J.R. and Singh, M. (2004) A novel method for estimating ancestral amino acid composition and its application to proteins of the Last Universal Ancestor. Bioinformatics 20, 2251–2257.
[27] Brochier, C. and Philippe, H. (2002) Phylogeny: a non-hyperthermophilic ancestor for bacteria. Nature 417, 244.
[28] Galtier, N., Tourasse, N. and Gouy, M. (1999) A non-hyperthermophilic common ancestor to extant life forms. Science 283, 220–221.
[29] Berezovsky, I. and Schakhnovich, E. (2005) Physics and evolution of thermophilic adaptation. Proc. Natl. Acad. Sci. USA 102, 12742–12747.
[30] Chakravarty, S. and Varadarajan, R. (2002) Elucidation of factors responsible for enhanced thermal stability of proteins: a structural genomics based study. Biochemistry 41, 8152–8156.
[31] Tekaia, F., Yeramian, E. and Dujon, B. (2002) Amino acid composition of genomes, lifestyles of organisms, and evolutionary trends: a global picture with correspondence analysis. Gene 297, 51–60.
[32] Vieille, C. and Zeikus, G. (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol. Mol. Biol. Rev. 65, 1–43.
[33] Ibba, M., Becker, H.D., Stathopoulos, C., Tumbula, D.L. and Söll, D. (2000) The adaptor hypothesis revisited. Trends Biochem. Sci. 25, 311–316.
[34] Soll, D. and Ibba, M. (2003) Aminoacyl-tRNA synthetase structure and evolution in: Translation Mechanisms (Lapointe, J. and Brakier-Gingras, L., Eds.), pp. 25–33, Landes Bioscience.
[35] Sauerwald, A., Zhu, W., Major, T., Roy, H., Palioura, S., Jahn, D., Whitman, W., Yates, J., Ibba, M. and Soll, D. (2005) RNA-dependent cysteine biosynthesis in Archaea. Science 307, 1969–1972.