Stichopus japonicus arginine kinase: gene structure and unique substrate recognition system

Tomohiko SUZUKI¹, Yasufumi YAMAMOTO and Masahiro UMEKAWA

Laboratory of Biochemistry, Faculty of Science, Kochi University, Kochi 780-8520, Japan

Arginine kinase from the sea cucumber *Stichopus japonicus* underwent a unique molecular evolution. Unlike the monomeric 40 kDa arginine kinases from molluscs and arthropods, *Stichopus* arginine kinase is dimeric, the same as cytoplasmic isoenzymes of the vertebrate creatine kinases. Its entire amino acid sequence is more similar to creatine kinases than to other arginine kinases, but the guanidino specificity region (GS region) is of the arginine kinase type. To elucidate its unusual evolution, the structure of the *Stichopus* arginine kinase gene was determined. It consisted of seven exons and six introns, and a part of the exon 2 of the *Stichopus* gene corresponds to the GS region. Compared with the structure of the human muscle creatine kinase gene (seven exons, six introns), the splice junctions of five introns were conserved exactly between the two genes, suggesting that these introns had been conserved for at least 500 million years. The entire sequence of *Stichopus* arginine kinase is distinctly included in the creatine kinase cluster in all tree construction methods examined. On the other hand, if the tree is constructed only from sequences corresponding to *Stichopus* exon 2, it is placed in the arginine kinase cluster. Thus we conclude that *Stichopus* arginine kinase evolved not from the arginine kinase gene but from the

creatine kinase gene, and suggest that its GS region, determining substrate specificity, has been replaced by an arginine kinase type via exon shuffling. In typical arginine kinases four residues, Ser⁶³, Gly⁶⁴, Val⁶⁵ and Tyr⁶⁸ (numbering from the *Limulus polyphemus* sequence), in the GS region are highly conserved and are associated with substrate binding. Among them, Tyr⁶⁸ appears to play a crucial role by forming a hydrogen bond with the substrate, and is conserved exactly in all arginine kinases. However, in *Stichopus* arginine kinase, none of these four conserved residues were present. Nevertheless, the enzyme displays an affinity for the substrate arginine ($K_m = 0.8$ mM) comparable with other arginine kinases. This implies that a completely different substrate-binding system has been developed in *Stichopus* arginine kinase. We propose that the His⁶⁴ in *Stichopus* arginine kinase acts as a substitute for the Tyr⁶⁸ in other arginine kinases, and that the imidazole ring of His⁶⁴ is hydrogen bonded with the substrate arginine, thus stabilizing it.

Key words: convergent evolution, creatine kinase, exon/intron organization, exon shuffling, site-directed mutagenesis.

INTRODUCTION

Phosphagen kinases are enzymes that catalyse the reversible transfer of the high-energy phosphoryl group of ATP to naturally occurring guanidino compounds such as creatine, glycocyamine, taurocyamine, lombricine and arginine. Members of this enzyme family play a key role in the interconnection of energy production and utilization in animals [1–4]. The phosphorylated high-energy guanidine is referred to as a phosphagen.

In vertebrates, phosphocreatine is the only phosphagen, and the corresponding phosphagen kinase is creatine kinase. In invertebrates, at least six phosphagens, phosphoarginine, phosphoglycocyamine, phosphotaurocyamine, phospholombricine, phosphohypotaurocyamine and phosphoopheline, are present in addition to phosphocreatine, and the corresponding kinases for the former four, arginine kinase, glycocyamine kinase, taurocyamine kinase and lombricine kinase, have been identified and characterized [1–5]. The homologous amino acid sequences of creatine, arginine, glycocyamine and lombricine kinases suggest that they have evolved from a common ancestor [6–8]. From an earlier amino-acid-sequence alignment of phosphagen kinases, we proposed that the guanidino specificity (GS) region, displaying

remarkable amino acid deletions, is a possible candidate for the guanidino-recognition site [9]. There is a proportional relationship between the size of the deletion in the GS region and the mass of the guanidine substrate used [9]. The GS region is partly overlapped by the so-called flexible loop in the crystal structures of chicken mitochondrial creatine kinase [10] and *Limulus polyphemus* arginine kinase [11].

Arginine kinase is widely distributed in the lower and higher invertebrate groups [12,13]. Conventional wisdom would suggest that arginine kinase is the most primitive member of the phosphagen kinase family and that the other members, including creatine kinase, arose from tandem gene duplications and subsequent divergence [6,8,12]. The crystal structure of *Limulus* arginine kinase has been determined recently [11]. It consists of two domains, an N-terminal small α -helical domain of residues 1 to 111 (N-domain) and a C-terminal domain possessing an eight-stranded antiparallel β -sheet flanked by seven α -helices of residues 112 to 357 (C-domain) [11]. ATP or ADP is accommodated in the C-domain, and another substrate, arginine or arginine phosphate, contacts mainly with the N-domain. The catalytic centre is located in the C-domain, where reversible transfer of the phosphate is achieved.

Unlike the monomeric 40 kDa arginine kinases from molluscs and arthropods, the *Stichopus japonicus* arginine kinase is dimeric, the same as the vertebrate creatine kinases. Moreover,

Abbreviations used: GS, guanidino specificity; UPG, unweighted pair-group clustering.

¹ To whom correspondence should be addressed (e-mail suzuki@sc.kochi-u.ac.jp or suzuki@cc.kochi-u.ac.jp).

the amino acid sequence of *Stichopus* arginine kinase [14] is completely different from the other arginine kinases in that (1) its overall sequence is more like that of a creatine kinase, and (2) its active site is more like that of an arginine kinase than that of a creatine kinase, yet it is also significantly different from other arginine kinases. These results suggest that arginine kinase evolved at least twice during the evolution of phosphagen kinases: first at an early stage of phosphagen kinase evolution (its descendants are molluscan and arthropod arginine kinases) and second, from creatine kinase at a later time in metazoan evolution [14].

In this paper, we describe the exon/intron organization of the *Stichopus* arginine kinase gene to make clear its evolutionary origin, and to show that it is apparently homologous with the mammalian creatine kinase gene. Furthermore, we propose a unique substrate-recognition system for *Stichopus* arginine kinase which is quite different from that of typical arginine kinases.

MATERIALS AND METHODS

PCR amplification of the *Stichopus* arginine kinase gene

The sea cucumber *Stichopus japonicus* was purchased from the seafood market of Kochi, Japan. The genomic DNA was purified from the longitudinal muscle using a conventional phenol/chloroform method. A 1 µg aliquot of genomic DNA and a pair of non-redundant primers (1 pmol each) were used to amplify a DNA fragment containing the intron. The genomic DNA was amplified for 30 cycles, each consisting of 1 min at 94 °C (denaturation), 1.5 min at 55 °C (annealing) and 5–10 min at 72 °C (primer extension). Extra or LA *Taq* DNA polymerase® (Takara, Tokyo, Japan) was used as the amplifying enzyme. The PCR products were isolated on a 0.7% agarose gel and sequenced directly with a dye-terminator-cycle-sequencing FS ready reaction kit® using a Model 373-18 DNA sequencer (Applied Biosystems, Warrington, U.K.).

Expression of wild-type *Stichopus* arginine kinase and site-directed mutagenesis

The open reading frame of 1113 bp of *Stichopus* arginine kinase was cloned into the *Bam*HI/*Sal*I site of pMAL/-c2 (pMAL-*Stichopus* AK). The maltose-binding protein-*Stichopus* arginine kinase fusion protein was expressed in *Escherichia coli* TB1 cells by induction with 1 mM isopropyl β-D thiogalactoside at 27 °C for 18 h. The soluble protein was extracted with B-PER reagent® (Pierce, Rockford, IL, U.S.A.) and the fusion protein was purified by affinity chromatography using amylose resin (New England Biolabs, Beverly, MA, U.S.A.). The enzyme assay was done using the purified fusion protein.

The mutation Asp⁶² (GAT) to Gly (GGG) was introduced into the template of pMAL/*Stichopus* AK by PCR using the primers: mF1, TTCCATTTGGGTCTTCTTGCCGGG and mR1, CCCGGGATTATCGACACCGTTCTG. The mutation Asp⁶²Phc⁶³ (GATTC) to AspGlyPhe (GATGGGTTC) was introduced into the template of pMAL/*Stichopus* AK by PCR using the primers: mF2, TTCCATTTGGGTCTTCTTGCCGGG and mR2, CCCATCGGGATTATCGACACCGTTCTG. *Pfu* DNA polymerase was used as the amplifying enzyme. The PCR products were digested with *Dpn*I and the target 7000 bp DNA was recovered with EasyTrap Ver.2 (Takara). After blunting and treating with polynucleotide kinase, the DNA was self-ligated. The cDNA insert of the plasmid was sequenced completely and it was confirmed that there was no unexpected mutation. The mutated protein was expressed as described above.

The enzyme activity was assayed spectrophotometrically at 25 °C [5]. The protein concentration was estimated from the absorbance at 280 nm (the absorbance 0.77 at 280 nm in a 1 cm cuvette corresponded to 1 mg of protein/ml).

RESULTS AND DISCUSSION

Evolution of *Stichopus* arginine kinase

The structure of the *Stichopus* arginine kinase gene was constructed from six overlapping fragments amplified separately by PCR. The gene is characterized by six introns and seven exons. All of the introns begin with gt and end with ag. Intron 2 is the longest (approx. 3500 bp) and intron 3 is the shortest (approx. 500 bp). Intron size and the sequence of the exon/intron boundaries of the *Stichopus* arginine kinase gene are listed in Table 1. The gene is estimated to be at least 9 kb in length.

The positions of the introns in the *Stichopus* arginine kinase gene were compared with the human muscle creatine kinase gene (six intron/seven exon structure) [15] and the *Caenorhabditis elegans* arginine kinase gene (two intron/three exon structure) [16] in the alignment of their amino acid sequences (Figure 1). The *C. elegans* gene is the only arginine kinase gene whose structure is known. The splice junctions of *Stichopus* arginine kinase introns 1, 3, 4, 5 and 6 were conserved in precisely the same positions as they were in the gene for human creatine kinase, suggesting that these introns may have been conserved for at least 500 million years [17]. In addition, the position of *Stichopus* arginine kinase intron 2 is close to the position of intron 2 of the human creatine kinase gene. In contrast, there is no conserved intron between the *Stichopus* and *C. elegans* arginine kinase genes. Thus, homology of the exon/intron organization of the genes for *Stichopus* arginine kinase and human creatine kinase clearly indicates that the *Stichopus* arginine kinase gene evolved from a gene for creatine kinase.

When the intron positions of *Stichopus* arginine kinase and human muscle creatine kinase were compared with those of human and chicken mitochondrial creatine kinase genes (eight exon/seven intron structure) [18,19], splice junctions of the two introns, intron 4 and 5 in the *Stichopus* arginine kinase gene, were conserved. Since the mitochondrial creatine kinase separated much earlier (at least before the divergence of deuterostomes and protostomes) from muscle creatine kinase isoenzymes [20], the intron positions of creatine kinase genes appears to be relatively stable.

In spite of the fact that the entire amino acid sequence and gene structure of the *Stichopus* enzyme are more similar to creatine kinases than other arginine kinases, this enzyme displays true arginine kinase activity. In this context, it is of great interest to ascertain how this catalytic activity towards arginine/arginine phosphate is accomplished. We previously showed that the amino acid sequence around the GS region of the *Stichopus* enzyme was of the arginine kinase type [14]; five amino acid deletions in the GS region which may help to accommodate larger guanidine substrates in the active site (the creatine kinase-type has one amino acid deletion).

There are two hypotheses to explain how the GS region of the *Stichopus* enzyme became the arginine kinase-type: (i) the creatine kinase-type sequence of the GS region in the *Stichopus* enzyme has been replaced by arginine kinase-type sequence via exon shuffling [14], similar to the multi-domain serine proteases [21], or (ii) point mutations accumulated in the GS region of the *Stichopus* enzyme and consequently caused the change of enzyme activity from creatine kinase to arginine kinase. Namely, the generation of *Stichopus* arginine kinase is

Intron sequences are shown by lower case letters. The primers indicated were used for amplification.

Intron	Primers	Sequence (estimated intron length)
Intron 1	ScAK.F1/ScAK.R3	GATAATCCCGGctgacgtact... (~ 1000 bp) ...ttctctcgaATTTCATT
Intron 2	ScAK.gF2/ScAK.gR1	AAGAAAAACAggtatttgta... (~ 3500 bp) ...atttctcatgATAACATAA
Intron 3	ScAK.gF3/ScAK.gR4	GTCTCCGAAGgtaagcgatg... (~ 500 bp) ...tcttcaccagCTCTGAATTC
Intron 4	ScAK.gF4/ScAK.gR3	GAGGTATCTGgaagttlac... (~ 1150 bp) ...tatttcacagGCATAATGAC
Intron 5	ScAK.gF5/ScAK.R1	CGTTACAGAGgtacctcttc... (~ 950 bp) ...ttttttcagGTCGAGAAGAC
Intron 6	ScAK.gF6/ScAK.gR5	AGAGGAACAAGtaagaacat... (~ 800 bp) ...tcttcacagGGGGAATT

The amino acid (codon) including the intron is boxed. The number for the intron is for the *Stichopus* arginine kinase gene. Phase 1, 2 and 3 indicate that the intron is inserted after the first, second and third nucleotide of the codon respectively.

Figure 2 For legend please see opposite page

Figure 2 Phylogenetic trees for amino acid sequences of arginine, creatine, glycoylamine and lombricine kinases

The *Stichopus* arginine kinase sequence is boxed. The number at the branching point shows the bootstrap value. The tree is presented as a rectangular cladogram. AK, arginine kinase; CK, creatine kinase; GK, glycoylamine kinase; LK, lombricine kinase. Left panel: (A) UPG, (B) maximum likelihood and (C) Neighbor-Joining trees obtained from comparison of the entire amino acid sequences. Right panel: (A) UPG, (B) maximum likelihood and (C) Neighbor-Joining trees constructed from sequences corresponding to *Stichopus* arginine kinase exon 2 (45 amino acid residues).

typical of adaptive convergent evolution, similar to the case of lysozymes from foregut fermenters [22].

To test the first idea, we constructed a phylogenetic tree using the unweighted pair-group clustering (UPG) method [23], which can give a root for the tree, from 36 amino acid sequences of arginine kinase, creatine kinase, glycoylamine kinase and lombricine kinase belonging to the guanidino kinase family. The tree obtained from the entire amino acid sequence showed that the *Stichopus* enzyme was distinctly placed in the creatine kinase cluster (Figure 2A, left panel), in agreement with the trees constructed using the maximum likelihood method [24] (Figure 2B, left panel), Neighbor-Joining method [23] (Figure 2C, left panel) and Fitch method [23] (results not shown). On the other hand, if the tree was constructed only from sequences corresponding to *Stichopus* arginine kinase exon 2 (residues 64 to 108 in Figure 1), which includes the flexible loop or a part of the GS region, it was placed in the arginine kinase cluster (Figure 2A, right panel). The topology of unrooted trees with the methods of maximum likelihood (Figure 2B, right panel), Neighbor-Joining (Figure 2C, right panel) and Fitch [23] (results not shown) and their bootstrap values also support the assertion that exon 2 of the *Stichopus* enzyme is of the arginine kinase type. Thus we conclude that the *Stichopus* enzyme evolved from the creatine kinase gene and suggest that its GS region, determining substrate specificity, has been replaced by an arginine kinase type via exon shuffling. We assume that the gene for the original

arginine kinase of *Stichopus* was altered and a modified creatine kinase gene evolved as a substitute.

However, shuffling of exon 2 does not alone explain the fact that *Stichopus* arginine kinase has a unique substrate-recognition system (described below). Therefore we can not completely exclude the second possibility that the *Stichopus* enzyme resulted from convergent evolution. To generate the arginine kinase activity, for example, from human muscle creatine kinase, in a manner similar to the *Stichopus* enzyme, at least four amino acid deletions and several independent point mutations must be introduced into the creatine kinase sequence (compare the sequences of *Stichopus* arginine kinase with human creatine kinase in Figure 3). Since such mutations would have to accumulate only in exon 2, this scenario appears to be an unlikely event.

The unique substrate-binding system in *Stichopus* arginine kinase

To probe the nature of substrate binding, *Stichopus* arginine kinase was successfully expressed as a fusion protein with maltose-binding protein in the soluble fraction of *E. coli*. The purified enzyme had a K_m of 0.83 ± 0.05 mM for arginine and a V_{max} of 34.4 ± 4.8 $\mu\text{mol P}_i \cdot \text{min}^{-1} \cdot \text{mg of protein}^{-1}$ for the forward reaction. These values are close to the those for the native enzyme ($K_m = 0.84 \pm 0.04$, $V_{max} = 50.0 \pm 4.5$).

	56	62	63	65	68	71
AK- <i>Limulus</i>	SGVENLD	SGVGIYAPD				
AK-Lobster	SGVENLD	SGVGIYAPD				
AK-Shrimp	SGVENLD	SGVGIYAPD				
AK-Chiton	SGAKNLD	SGVGVIYASD				
AK- <i>Nautilus</i>	SGCQNLD	SGVGIYACD				
AK- <i>Abalone</i>	SGCLNLD	SGVGIYACD				
AK- <i>SeaAnemone.D1</i>	SGVVNLD	SGVGVIYAGD				
AK- <i>Stichopus</i>	NGVDNPD	FHLGLLAGD				
CK-Human.M	TGVDNPGHPF	IMTVGCVAGD				
CK-Chicken.M	TGVDNPGHPF	IMTVGCVAGD				
CK-Torpedo	TGVDNPGHPF	IMTVGCVAGD				
CK-Human.B	TGVDNPGHPY	IMTVGCVAGD				
CK-Chicken.B	TGVDNPGHPF	IMTVGCVAGD				
CK-Chicken.Mi	TGVDNPGHPF	IKTVGMVAGD				
CK-Human.Mi	TGVDNPGHPF	IKTVGMVAGD				
LK- <i>Eisenia</i>	PSVDNT	GRIIGLVAGD				
GK- <i>Neanthes.α</i>	TGVDNPGNKFYGKKTGCVFGD					

Figure 3 Alignment of the amino acid sequence of the GS region of arginine, creatine, glycocyamine and lombricine kinases

Asp⁶², conserved in arginine kinases, is boxed by the dotted line. Ser⁶³, Gly⁶⁴, Val⁶⁴ and Tyr⁶⁸, interacting with the substrate arginine, are boxed by a solid line. Abbreviations as in Figure 2.

We noted previously that Asp⁶² (numbering based on the *Limulus* arginine kinase sequence [25]) in the GS region was conserved in most arginine kinase sequences (Figure 3) [26]. In the crystal structures of chicken mitochondrial creatine kinase and *Limulus* arginine kinase, the GS region is partly overlapped by the so-called flexible loop that has been speculated to move nearer to the active site in order to exclude water during catalysis [10,11]. It should be pointed out that Asp⁶² is not conserved in

any creatine kinases (residue 62 is a conserved Gly in creatine kinases; see Figure 3). Even though *Stichopus* arginine kinase evolved from creatine kinase, Asp⁶² is present, as in other arginine kinases.

The transition-state analogue-complex structure of *Limulus* arginine kinase shows that Asp⁶² is not involved in substrate binding, and the authors have not referred to any functional role [11]. The three-dimensional structure of *Limulus* arginine kinase, viewed by SwissPdbViewer (ProteinDataBank id. 1BG0), shows that the side chain of Asp⁶² in the N-domain is hydrogen-bonded with the side chain of Arg¹⁹³ in the C-domain (Figure 4, left panel). The Arg¹⁹³ is also conserved in all arginine kinases including *Stichopus* arginine kinase (Figure 4, right panel). Thus it is very likely that the N- and C-domains are maintained by the hydrogen bond between Asp⁶² and Arg¹⁹³ in a favourable topology for the enzyme reaction [27].

Site-directed mutagenesis of Asp⁶² to Gly in *Stichopus* arginine kinase, which breaks the hydrogen bond completely, showed a markedly reduced activity of $0.022 \pm 0.002 \mu\text{mol P}_i \cdot \text{min}^{-1} \cdot \text{mg}$ of protein⁻¹, which is less than 1 % of the recombinant wild-type. This result was reinforced by the mutant Asp⁶²Phe⁶³ to Asp-GlyPhe (insertion of Gly). This mutant could also be expected to break the hydrogen-bond between Asp⁶² and Arg¹⁹³, similar to the Asp⁶² to Gly mutation. As expected, its arginine kinase activity ($0.035 \pm 0.004 \mu\text{mol P}_i \cdot \text{min}^{-1} \cdot \text{mg}$ of protein⁻¹) was decreased by nearly as much as it was in the case of the Asp⁶² to Gly mutation.

A crystal structure of *Limulus* arginine kinase showed that the backbone of the three amino acid residues (Ser⁶³, Gly⁶⁴ and Val⁶⁵), located in the GS region, are hydrogen-bonded with the carboxyl group of the substrate arginine [11] (Figure 4, left panel). In addition, the aromatic side chain of Tyr⁶⁸ forms a hydrogen bond with the amino group of the substrate arginine. These four residues are conserved in almost all of the arginine kinases known thus far. However, in *Stichopus* arginine kinase, none of the four residues are present (see the sequence in Figure 3). Nevertheless, this enzyme displays an affinity for the substrate

Figure 4 A proposed structure for binding of the substrate arginine in *Stichopus* arginine kinase

This figure was made using the SwissProteinDataBaseViewer program (left panel) *Limulus* arginine kinase figure constructed using the three-dimensional data published in [11], (right panel) *Stichopus* arginine kinase. The amino acids interacting with the substrate arginine are coloured in yellow.

arginine ($K_m = 0.8$ mM) comparable with other arginine kinases. This suggests that a rather different substrate-binding system has evolved in *Stichopus* arginine kinase. We propose that the His⁶⁴ in *Stichopus* arginine kinase acts as a substitute for the Tyr⁶⁸ in other arginine kinases, namely the imidazole ring of His⁶⁴ is hydrogen-bonded with the substrate arginine, stabilizing it (Figure 4, right panel).

The unique substrate-binding system in *Stichopus* arginine kinase, shown in Figure 4, appears to be in favour of convergent evolution by point mutations that occurred in the *Stichopus* substrate specificity region (exon 2). However we cannot find a reasonable explanation for why the mutations have to accumulate only in exon 2. Therefore, at present, we assume that the shuffling of exon 2 is the most probable event in the unique evolution of *Stichopus* arginine kinase.

We thank Dr. Ross Ellington (Florida State University, FL, U.S.A.) for giving us helpful suggestions.

REFERENCES

- Watts, D. C. (1968) The origin and evolution of phosphagen phosphotransferases. In *Homologous Enzymes and Biochemical Evolution* (Van Thoai, N. and Roche, J., eds.), pp. 279–296, Gordon and Breach, New York
- Van Thoai, N. (1968) Homologous phosphagen phosphokinases. In *Homologous Enzymes and Biochemical Evolution* (Van Thoai, N. and Roche, J., eds.), pp. 199–229, Gordon and Breach, New York
- Morrison, J. F. (1973) Arginine kinase and other invertebrate guanidino kinases. In *The Enzymes* (Boyer, P. C., ed.), pp. 457–486, Academic Press, New York
- Kenyon, G. L. and Reed, G. H. (1986) Creatine kinase: structure–activity relationships. *Adv. Enzymol.* **54**, 367–426
- Ellington, W. R. (1989) Phosphocreatine represents a thermodynamic and functional improvement over other muscle phosphagens. *J. Exp. Biol.* **143**, 177–194
- Muhlebach, S. M., Gross, M., Wirz, T., Walliman, T., Perriard, J. -C. and Wyss, M. (1994) Sequence homology and structure predictions of the creatine kinase isoenzymes. *Mol. Cell. Biochem.* **133/134**, 245–262
- Suzuki, T. and Furukohri, T. (1994) Evolution of phosphagen kinase. Primary structure of glycocyamine kinase and arginine kinase from invertebrates. *J. Mol. Biol.* **237**, 353–357
- Suzuki, T., Kawasaki, Y. and Furukohri, T. (1997) Evolution of phosphagen kinase. Isolation, characterization and cDNA-derived amino acid sequence of two-domain arginine kinase from the sea anemone *Anthopleura japonicus*. *Biochem. J.* **328**, 301–306
- Suzuki, T., Kawasaki, Y., Furukohri, T. and Ellington, W. R. (1997) Evolution of phosphagen kinase. VI. Isolation, characterization and cDNA-derived amino acid sequence of lombricine kinase from the earthworm *Eisenia foetida*, and identification of a possible candidate for the guanidine substrate recognition site. *Biochim. Biophys. Acta* **1348**, 152–159
- Fritz-Wolf, K., Schnyder, T., Wallimann, T. and Kabsch, W. (1996) Structure of mitochondrial creatine kinase. *Nature (London)* **381**, 341–345
- Zhou, G., Somasundaram, T., Blanc, E., Parthasarathy, G., Ellington, W. R. and Chapman, M. (1998) Transition state structure of arginine kinase: implications for catalysis of bimolecular reactions. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8449–8454
- Watts, D. C. (1971) Evolution of phosphagen kinases. In *Biochemical Evolution and the Origin of Life* (Schoffeniels, E., ed.), pp. 150–173, North-Holland, Amsterdam
- Watts, D. C. (1975) Evolution of phosphagens along the chordate line. *Symp. Zool. Soc. Lond.* **36**, 105–127
- Suzuki, T., Kamidochi, M., Inoue, N., Kawamichi, H., Yazawa, Y., Furukohri, T. and Ellington, W. R. (1999) Arginine kinase evolved twice: evidence that echinoderm arginine kinase originated from creatine kinase. *Biochem. J.* **340**, 671–675
- Trask, R. V., Strauss, A. W. and Billadello, J. J. (1988) Developmental regulation and tissue-specific expression of the human muscle creatine kinase gene. *J. Biol. Chem.* **263**, 17142–17149
- The Washington University Genome Sequencing Center (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. The *C. elegans* Sequencing Consortium. *Science* **282**, 2012–2018
- Goodman, M., Pedwaydon, J., Czelusniak, J., Suzuki, T., Gotoh, T., Moens, L., Shishikura, F., Walz, D. and Vinogradov, S. (1988) An evolutionary tree for invertebrate globin sequences. *J. Mol. Evol.* **27**, 236–249
- Haas, R. C., Korenfeld, C., Zhang, Z. F., Perryman, B., Roman, D. and Strauss, A. W. (1989) Isolation and characterization of the gene and cDNA encoding human mitochondrial creatine kinase. *J. Biol. Chem.* **264**, 2890–2897
- Muhlebach, S. M., Wirz, T., Brandle, U. and Perriard, J. C. (1996) Evolution of the creatine kinases. The chicken acidic type mitochondrial creatine kinase gene as the first nonmammalian gene. *J. Biol. Chem.* **271**, 11920–11929
- Pineda, Jr, A. O. and Ellington, W. R. (1999) Structural and functional implications of the amino acid sequences of dimeric, cytoplasmic and octameric mitochondrial creatine kinases from a protostome invertebrate. *Eur. J. Biochem.* **264**, 67–73
- Rogers, J. (1985) Exon shuffling and intron insertion in serine protease genes. *Nature (London)* **315**, 458–459
- Stewart, C. B., Schilling, J. W. and Wilson, A. C. (1987) Adaptive evolution in the stomach lysozymes of foregut fermenters. *Nature (London)* **330**, 401–404
- Felsenstein, J. (1993) PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle, WA, U.S.A.
- Strimmer, K. and von Haeseler, A. (1996) Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13**, 964–969
- Strong, S. J. and Ellington, W. R. (1995) Isolation and sequence analysis of the gene for arginine kinase from the chelicerate arthropod, *Limulus polyphemus*: insights into catalytically important residues. *Biochim. Biophys. Acta* **1246**, 197–200
- Suzuki, T., Kawasaki, Y., Unemi, Y., Nishimura, Y., Soga, T., Kamidochi, K., Yazawa, Y. and Furukohri, T. (1998) Gene duplication and fusion have occurred frequently in the evolution of phosphagen kinases—a two-domain arginine kinase from the clam *Pseudocardium sachalinensis*. *Biochim. Biophys. Acta* **1388**, 253–259
- Suzuki, T., Fukuta, H., Nagato, H. and Umekawa, M. (2000) Arginine kinase from *Nautilus pompilius*, a living fossil. Site-directed mutagenesis studies on the role of amino acid residues in GS (guanidino specificity) region. *J. Biol. Chem.* **275**, 23884–23890

Received 3 April 2000/13 July 2000; accepted 14 August 2000