

Arginine Kinase from *Nautilus pompilius*, a Living Fossil

SITE-DIRECTED MUTAGENESIS STUDIES ON THE ROLE OF AMINO ACID RESIDUES IN THE GUANIDINO SPECIFICITY REGION*

Received for publication, April 6, 2000, and in revised form May 9, 2000
Published, JBC Papers in Press, May 12, 2000, DOI 10.1074/jbc.M002926200

Tomohiko Suzuki‡, Hideki Fukuta, Hiromi Nagato, and Masahiro Umekawa

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

Arginine kinases were isolated from the cephalopods *Nautilus pompilius*, *Octopus vulgaris*, and *Sepioteuthis lessoniana*, and the cDNA-derived amino acid sequences have been determined. Although the origin and evolution of cephalopods have long been obscure, this work provides the first molecular evidence for the phylogenetic position of Cephalopoda in molluscan evolution. A crystal structure for *Limulus* arginine kinase showed that four amino acid residues (Ser⁶³, Gly⁶⁴, Val⁶⁵, and Tyr⁶⁸) are hydrogen-bonded with the substrate arginine. We introduced three independent mutations, Ser⁶³ → Gly, Ser⁶³ → Thr, and Tyr⁶⁸ → Ser, in *Nautilus* arginine kinase. One of the mutants had a considerably reduced substrate affinity, accompanied by a decreased $V_{\max}$. In other mutants, the activity was lost almost completely. It is known that substantial conformational changes take place upon substrate binding in arginine kinase. We hypothesize that the hydrogen bond between Asp⁶² and Arg¹⁹³ stabilizes the closed, substrate-bound state. Site-directed mutagenesis studies strongly support this hypothesis. The mutant (Asp⁶² → Gly or Arg¹⁹³ → Gly), which destabilizes the maintenance of the closed state and/or perhaps disrupts the unique topology of the catalytic pocket, showed only a very weak activity (0.6–1.5% to the wild-type).

Phosphagen (guanidino) kinases constitute a family of highly conserved enzymes, which catalyzes the reversible transfer of phosphate from phosphagen such as creatine phosphate to ADP. Creatine kinase (CK),¹ the most widely studied member of this family, plays a central role in both temporal and spatial ATP buffering in cells that display high and variable rates of energy turnover (1, 2). The crystal structures have been determined for chicken mitochondrial CK (3) and muscle and brain CKs (4, 5). The remaining major members of this family (arginine kinase (AK), glycocyamine kinase (GK), lombricine kinase (LK), taurocyamine kinase (TK)) have not been investigated as extensively as CK but likely play a similar physiological role (2, 6, 7).

Arginine kinase (AK) is widely distributed in the lower and higher invertebrate groups and is present in many lower chor-

dates but absent in the vertebrates (6, 8). Interestingly, AK evolved at least twice during evolution of phosphagen kinases: first, at an early stage of phosphagen kinase evolution (its descendants are molluscan and arthropod AKs) and second, from CK at a later time in metazoan evolution (9). Conventional wisdom would suggest that AK is the most primitive member of the phosphagen kinase family and that the other members, including CK, arose from tandem gene duplications and subsequent divergence (6, 10, 11).

From an earlier amino acid sequence alignment of CKs, GK, LK, and AKs, we proposed that a region displaying remarkable amino acid deletions (referred to as the guanidino specificity (GS) region) was a possible candidate for the guanidino recognition site (12). This GS region is overlapped partly by the so-called flexible loop in the crystal structures of chicken mitochondrial CK (3) and *Limulus* AK (13). There is a proportional relationship between the size of the deletion in the GS region and the mass of guanidine substrate used (12). Among the amino acid residues in the GS region, we noted previously that the seventh residue, Asp, is conserved in all AK sequences but not in other phosphagen kinases. This residue is not associated with the substrate binding (13), but we assumed that it has a special role in the recognition of the substrate arginine (14).

Mollusks constitute one of the most diverse animal phyla. Living representatives are classically divided into seven classes, of which major ones are Bivalvia, Gastropoda, and Cephalopoda. There are several hypotheses concerning molluscan phylogeny based on morphological features, but the results are still controversial. Sequence comparisons using 18S rRNA, 12S rRNA, or mitochondrial DNA have been used to solve the complicated phylogeny of molluscan classes except Cephalopoda (15–19). However, molecular analyses have not resulted in a clear determination for the phylogeny. Thus, the origin and evolution of cephalopods have long been obscured by the poor fossil record and molecular data.

Nautilus pompilius is the only living representative of the large extinct group of nautiloids, ammonites and belemnites, and is sometimes referred to as a living fossil, like *Latimeria* and *Lingula*. In this report, first, we determined the cDNA-derived amino acid sequences of AKs from three cephalopods *N. pompilius*, *Octopus vulgaris*, and *Sepioteuthis lessoniana* (squid) to elucidate the evolution of molluscan AKs. Next, we examined the functional role of several amino acid residues on the GS region of *Nautilus* AK by site-directed mutagenesis. In particular, we focused on functional significance of Asp⁶², which is not associated with the substrate binding but is conserved in all AK sequences.

Throughout this report, the sequence numbering of *Limulus* AK (20) is used. The actual sequence number in *Nautilus* AK, used in the mutagenesis studies, can be obtained by subtracting 6 from the *Limulus* numbering.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Laboratory of Biochemistry, Faculty of Science, Kochi University, Kochi 780-8520, Japan. Tel.: 88-844-0111; Fax: 88-844-8356; E-mail: suzuki@sc.kochi-u.ac.jp or suzuki@cc.kochi-u.ac.jp.

¹ The abbreviations used are: CK, creatine kinase; AK, arginine kinase; GK, glycocyamine kinase; LK, lombricine kinase; TK, taurocyamine kinase; GS region, guanidino specificity region; TSAC, transition state analog complex; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; bp, base pair(s).

MATERIALS AND METHODS

Isolation of Arginine Kinases—All procedures were carried out at 4–8 °C. The body wall muscle (10 g) of *Nautilus*, *Sepioteuthis*, or *Octopus* was homogenized with 50 ml of 10 mM Tris acetate buffer (pH 8.1) containing 0.1 mM dithiothreitol. The tissue extract was fractionated with 60–80% saturated ammonium sulfate. The precipitate was dissolved in a minimum volume of the same buffer, and applied to a Sephadex G-75 column (3 × 115 cm) equilibrated with the same buffer. The fractions possessing AK activity were pooled and applied to a DEAE-5PW column (7.5 × 75 mm, Tosoh) equilibrated with 20 mM Tris acetate buffer (pH 8.1) containing 0.2 mM dithiothreitol. The column was washed with the same buffer and then eluted with a linear gradient of 0–500 mM NaCl in 20 mM Tris acetate buffer. The purified AK was stored on ice or at –80 °C until use.

Enzyme activity was assayed spectrophotometrically at 25 °C (7). Protein concentration was estimated from the absorbance at 280 nm (the absorbance 0.77 at 280 nm in a 1-cm cuvette corresponds to 1 mg of protein/ml).

SDS-polyacrylamide gel electrophoresis (PAGE) was carried out in 15% acrylamide gels containing 0.087% *N,N'*-methylenebisacrylamide, 0.375 M Tris-HCl (pH 8.8) and 0.1% SDS. The sample was incubated in 0.75% SDS at 100 °C for 5 min in the presence of 2-mercaptoethanol before electrophoresis.

Partial Amino Acid Sequence Determination—The isolated proteins were pyridylethylated and digested with lysyl endopeptidase at an enzyme/substrate ratio of 1:100 in 0.1 M Tris-HCl (pH 8.0) at 37 °C for 6 h. The digested products were isolated by reverse-phase chromatography. The column (Cosmosil 5C₁₈-300, 4.6 × 150 mm) was equilibrated with 0.1% trifluoroacetic acid and eluted with a linear gradient of 0–90% acetonitrile in 0.1% trifluoroacetic acid over 90 min at a flow rate of 1 ml/min. Peptides for sequence analysis were purified further by rechromatography on the same column with a linear gradient of acetonitrile in 10 mM ammonium acetate. The amino acid sequences of the whole protein and peptides were determined using an automated protein sequencer (Applied Biosystems 476A).

cDNA Sequence Determination—mRNA was prepared from the radular muscle of *Nautilus*, *Sepioteuthis*, or *Octopus* with a QuickPrep Micro mRNA purification kit (Amersham Pharmacia Biotech). The single stranded cDNA was synthesized with reaction beads (Ready-To-Go You-Prime First-Strand Beads, Amersham Pharmacia Biotech) using the oligo-dT adapter as a primer. The 3' half of the cDNA was first amplified for 30 cycles, each consisting of 1 min at 94 °C for denaturation, 1 min at 55 °C for annealing, and 1.5 min at 72 °C for primer extension by polymerase chain reaction (PCR). Extra *Taq* DNA polymerase (Takara) was used as the amplifying enzyme. The primers used are the oligo-dT adapter and a 512-fold “universal” redundant oligomer, 5'-GT(A/C/G/T)TGG(A/G/T)(A/C/G/T)AA(T/C)GA(A/G)GA(A/G)GA(T/C)CA, designed for amplification of phosphagen kinases (21). The amplified products were sequenced directly with dye terminator cycle sequencing FS ready reaction kit using a Model 373–18 DNA sequencer (Applied Biosystems). The universal redundant oligomer was used as a sequence primer.

The 5' half of the cDNA was amplified as follows. The single stranded cDNA was newly synthesized with the nonredundant primer (Nauti.R1 for *Nautilus*, Sepio.R1 for *Sepioteuthis*, and Octo.R1 for *Octopus*), and the poly(A)⁺ tail was added to the 3' end with terminal deoxynucleotidyl transferase. Then the 5' half of the cDNA was amplified by the method described above, using the oligo-dT adapter and the nonredundant primer (Nauti.R2 for *Nautilus*, Sepio.R2 for *Sepioteuthis*, and Octo.R2 for *Octopus*). The amplified products were sequenced directly.

Expression of Wild-type *Nautilus* AK and Site-directed Mutagenesis—The open reading frame of 1053 base pairs (bp) of *Nautilus* AK was cloned into the *Bam*HI/*Hind*III site of pMAL-c2 (pMAL/*Nautilus*AK-wild). The maltose binding protein-*Nautilus* AK fusion protein was expressed in *Escherichia coli* TB1 cells by induction with 1 mM isopropyl-1-thio-β-D-galactopyranoside at 27 °C for 18 h. The soluble protein was extracted with B-PER reagent (Pierce), and the fusion protein was purified by affinity chromatography using amylose resin (New England BioLabs). The enzyme assay was done using the purified fusion protein.

The mutations (Asp⁶² (GAC) → Gly (GGC), Asp⁶² (GAC) → Glu (GAG), Ser⁶³ (AGT) → Gly (GGT), Tyr⁶⁸ (TAC) → Ser (TCC), and Arg¹⁹³ (AGA) → Gly (GGA)) were introduced in the template of pMAL/*Nautilus*AK-wild by PCR using mutation-primers (the sequences are shown in Table I). *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 by EasyTrap (version 2, Takara). After blunting

TABLE I
Primers used for site-directed mutagenesis of *Nautilus* AK

Mutation	Primer name	Sequence
Asp ⁶² → Gly	NaAK.mF1	TCTACGCCTGTGACCCAGATGCTT
	NaAK.mR2	TCCCAACACCACTGCCAGGTTTTG
Asp ⁶² → Glu	NaAK.mF1	TCTACGCCTGTGACCCAGATGCTT
	NaAK.mR2	TCCCAACACCACTCTCCAGGTTTTG
Ser ⁶³ → Gly	NaAK.mF1	TCTACGCCTGTGACCCAGATGCTT
	NaAK.mR3	TCCCAACACCACTGCCAGGTTTTG
Ser ⁶³ → Thr	NaAK.mF1	TCTACGCCTGTGACCCAGATGCTT
	NaAK.mR4	TCCCAACACCACTGTCCAGGTTTTG
Tyr ⁶⁸ → Ser	NaAK.mF3	TCTCCGCCTGTGACCCAGATGCTT
	NaAK.mR6	TCCCAACACCACTGTCCAGGTTTTG
Arg ¹⁹³ → Gly	NaAK.mF2	GGATTTCTAAAGCAGCTGGTGG
	NaAK.mR5	GTCTGAATCATTTGAACAAGAAGT

and kination, the DNA was self-ligated. The mutated protein was expressed as described above. The mutated cDNA insert was sequenced completely, and it was confirmed that there is no unexpected mutation.

RESULTS

Arginine kinases (AKs) from the body wall muscles of *Nautilus*, *Octopus*, and *Sepioteuthis* were purified to homogeneity using ammonium sulfate fractionation, Sephadex G-75 gel filtration and DEAE-5PW chromatography. SDS-PAGE, in the presence of a reducing agent, showed that the isolated AKs were highly purified, and the molecular masses were about 40 kDa. The AKs were confirmed to be monomers on the gel filtration column. No N-terminal amino acid of the AKs was detected by protein sequencing, indicating that N termini were blocked. The isolated enzyme was pyridylethylated and digested with lysyl endopeptidase. The peptides were purified by reverse-phase chromatography, and some peptides were sequenced (Fig. 1, A–C).

The enzymatic parameters, K_m for arginine and V_{max} , of the isolated AK enzymes from *Nautilus*, *Octopus*, and *Sepioteuthis* were obtained at 25 °C for the forward reaction (Table II). The K_m values ranged from 1.0 to 2.8 mM, whereas the V_{max} values ranged from 4.3 to 10.4 μmol of P_i/min/mg of protein. These values are comparable to other molluscan AKs (*Solen*, *Meretrix*, and *Aplysia*) determined under the same conditions.

The 3' halves of the cDNAs of *Nautilus*, *Octopus*, and *Sepioteuthis* AKs were amplified by PCR with the phosphagen consensus primer and oligo-dT adapter. Then two new oligomers were designed for amplification of the 5' halves. These 3' and 5' products were sequenced directly, and the nucleotide sequences of 1227 bp for *Nautilus*, 1226 bp for *Octopus*, and 1306 bp for *Sepioteuthis* were determined (Fig. 1, A–C). The open reading frame for *Nautilus* AK is 1053 nucleotides in length and encodes a protein with 350 amino acid residues; those for *Octopus* AK, 1044 nucleotides and 347 residues; those for *Sepioteuthis* AK, 1047 nucleotides and 348 residues. The validity of the cDNA-derived amino acid sequences was supported by chemical sequencing of internal lysyl endopeptidase peptides (81 amino acid residues determined for *Nautilus*, 36 residues for *Octopus*, 43 residues for *Sepioteuthis*). The molecular masses were calculated to be 39,276, 39,176, and 39,040 Da, respectively, consistent with the values estimated by SDS-PAGE. The nucleotide and amino acid sequences has been submitted to the DDBJ (accession numbers AB042330 for *Nautilus*, AB042331 for *Octopus*, and AB042332 for *Sepioteuthis*).

The recombinant maltose binding protein/*Nautilus* AK (wild type) gave enzymatic parameters K_m (0.68 mM) and V_{max} (7.64 μmol of P_i/min/mg of protein) similar to those of the native enzyme. Those for the mutated *Nautilus* AKs are listed in Table II.

A

ATCCCTGCTGCACCATGCTGGACGGGATAAATCTACGAGACCTGAAGAACCCAAAGACTGCAAGTCCCTG 75
 M S D R D K L Y E T L K N A K D C K S L

TTGAAGAAGAAGCTCACCAGTAACTCTCGAGCAGCTGAAGGATAAAGACCAAATTCGGCGGTCCCTGGCC 150
 L K K N L T D E L F E Q L K D K K T K F G G S L A

GACTGCATCCGCTCAGGCTGCAAACTGGAAGTGGTGGTATTACGCTGTGACCCAGATGCTTACACC 225
 D C I R S G C Q N L D S G V G I Y A C D P D A Y T

GTTTTGCACAGTTCTTAATAAGTTATCCAGGATTACCAAGACATCTTCAGTGAGTCATCCAGAACAGAC 300
 V F A P V L N K V I Q D Y H K T S S V S H P E P D

TTTGAGATATTGATAACTTGAATTTTCTGACCTGACCCAGAAGGAAATTAATTGCATCAACAGAGTCCGT 375
 F G D I D N L E F S D L D P E G K L I A S T R V R

GTGGGCAGATCCACGACAGTTTGGCTTCCCTCCAGTGCTGACCAAGAGGACCGTGAATACGGAGAAACT 450
 V G R S H D S F G F P P V L T K E D R V N M E K T

AGTCTGGAGGCTCGAACCTTGAAGGTGATCTTAAGGGTACTTACTTCCCTTGAAGGATGAGCAAGAG 525
 S L E A L N T L E G D L K G T Y F P L Q G M S K E

ATCCAGTCCAAATTAACAGAGGATCACTTCTTGTCAATGATTGACAGATTTCTAAAGCAGCTGGTGTAT 600
 I Q S K L T E D H F L F N D S D R F L K A A G G Y

GACGACTGGCAATTTGGTGGGGAATTTCCATAATAAGACAAGACGTTTGGTGGTCAATGAAGAGGAT 675
 D D W P I G R G I F H N K D K T F L V W V N E E D

→ Nauti.R2 ← Nauti.R1
 CACCTTCGGTGTATTCATGCAGATGGGTGGAGACTTGGTCAAGTGTATGCCAGACTTGTAAAGCTATCAA 750
 H L R L I S M Q M G G D L G Q V Y A R L V K A I K

ACATTGGAGAAGAAGCTATCTTGTCTGCCTGCTAAATGGGGTATCTGACCTTCTGCTACGAACCTGGGG 825
 T L E K K L S F A R L P N M G Y L T F C P S N L G

ACAACCTGAGAGCTAGTGTGCACATCAGAATTCCTCAAGTTTCTGCCCTACCAAACTCAATGATATTTGCT 900
 T T L R A S V H I R I P K V S A L P N F N D I C A

AAGCACAACCTGAGGCCAGAGGTATCCACGGTGAACACGACAGGTAGGTGGTATCTATGATATCTCGAAC 975
 K H N L Q A R G I H G E H T Q S V G G I Y D I S N

AAGCGAAGATGGGACTACAGAAATGAAGCTGTGAGGAAATGAGATAGGTGAGGAAATATCGAGCTG 1050
 K R M G L T E I E A V E Q E M R I G V E K I I E L

GAGAAGGCTCGAGTGATTATATTTTTTCAAAAGTGTACAGGTTCTAGTCTCAAATTTTGTGTTGTT 1125
 E K S L Q

CCTTTTGACATGTTGGGGTTTTTCTACTTGATTGTAATACACCTGAAATAGTAAACCCCACTATGCAAT 1200

GTAAATCTTCTGCTGTCACACGCA 1227

B

AATACTTGAACATTAACTAAGTTATAACTATGGCTGAAGAACTTTCAAGGAGCTTCAGGAAGCAAGGAATGT 75
 M A E E L F K E L Q E A K E C

CATTCTTTATTGAAGAAATCTCACTAAGGAGCGTGTGATAAGTTGAAACTTTAAAGACCAAGTTTGGTGGT 150
 H S L L L K K H L T K E R V D K L K T L K T K F G G

ACTTTAGCAGATGCTCCGTTGAGGTTGCAAAACCTGACAGCGGTGTGGTATCTACGCTTGCATCCTGAT 225
 T L A D C I R S G C K N P D S G V G I Y A C D P D

GCATACCTGTATTTGCTGACGCTTATGCGCTCATCATGGACTACCAATAAATGATAAAGTTCAACATCCA 300
 A Y T Y F A D V L D A V I M D Y H K I D K V Q H P

GTTCCAGACTTGGTGATGCTCAATAACTTGAACATCGGTGATCTTGCATCTGGCAGCTTGATTGTCTTACA 375
 V P D F G D V N N L N I G D L D P S G S L I V S T

CGTGTCCGCTAGGAGGAGCCAGCAGATTTGGTTTCCACCAAGTTCTTAAGAAAGAGGATCGTGTAAATG 450
 R V R V G R S H D S F G F P P V L K K E D R V K M

GAACAAGTATCCGTTGAGGCTTTGAAATCCCTTACAGGTGAATGGCTGAATATACCCCTCTCCACCATG 525
 E Q V S V E A L K S L T G E L A G N Y Y P L S T M

ACCCCTGACGTCAGAAAGCACTGATGATCATTTCTCTTTAAGCACTCTGACAGATTCTTGAAGGCAGCT 600
 T P D V Q K Q L T D D H F L F N D S D R F L K A A

→ AATGGATATGATGACTGGCTACTGGCAGAGGCTACTTCACTGCTCAATAAGACTTTCTGTGTTGGTCAAT 675
 N G Y D D W P T G R G I Y F S A N K T F L C W V N

→ GAGGAAGTCACTTGAAGTTGATTTCCATGCGAAGGGTGGTAACTTGGGTGAAGTCTACAGGAGATTGGTTGAA 750
 E E D H L R L I S M Q K G G N L G E V Y R R L V E

← Oct.R2
 GCTATCCATCAATGGAGAAGAACTGAAGTTTGAAGAAGGACACATGGGTACTTGAATCTGCTCCCAAGC 825
 A I H Q M E K K L K F A K K D N M G Y L T F C P S

← Oct.R1
 AACTTGGTACTACCTAGAGCTAGTGTGATATTAAGATCCCTAAGCTTTCCAGAGAGCAGACTTCAAACT 900
 N L G T T M R A S V H I K I P K L S Q R A D F K T

ATCTGCGCAAAATACCATCTCCAGGCCAGAGGTATCCATGGTGAACATACAGAAAGTGTGGCGGTGTCTATGAT 975
 I C D K Y H L Q A R G I H G E H T E S V G G V Y D

ATCTCTAACAACCGCGTATGGGTCTACAGAGTACGAGGCTGTCACTGAAATGATGAGAGGTGTCAACGAGATC 1050
 I S N K R R M G L T E Y E A V T E M M R G V N E I

ATTCTGGAGAAAACACAGTTAACTTTTTTTTTTAATGTAAATAAAACGATTTGGGATTTGATAACAAA 1125
 I R E E N N S

AGTAAGTCTACTATGTTGAAGTCTGTTTGAATAATCTCAGATACCAACTTTTAGACTGTAAAGGCTTTTGGC 1200

CATGCGAGTGAATAAAACCTTAA

1226

C

CAATCAAGCTATTGTTCTTTCTACTTTAGGATATTAGACTTACAGCTACACCATGCTGACTCAGAAGAGCTCT 75
 M S D S E E L F

TCAAAGTTTTGAGGAGCCAGGACTGCAAACTTTGCTCAAAAAAATCTTACCCAGAGCGTTTAAACAAGC 150
 K V L Q D A K D C K S L L K K N L T P E R F N K L

TGAAGGCTCTCAAGTCAAATTTGGTGGTACTTTAGCTGATTGCATCCGTTCAAGTTGAAGAACCCTGACAGCG 225
 K G L K S K F G G T L A D C I R S G C K N P D S G

GTGTGGTATCTATGCTGTGACCTGATGCATACACCACTTTGCTGAAGTCTTGATGCCGTTATCATGGACT 300
 V G I Y A C D P D A Y T T F A E V L D A V I M D Y

ATCACAAGATTGACAAGTTACACATCCCAACCCAGCTTTGGTGATGTCAATAACTTGGGCTTCGGTGACTTAG 375
 H K I D K V T H P Q P S F G D V N N L G F G D L D

ATCTCTTGGTGATCTGATGCTCAACCCGTGTCGTGTGGCAGAAGCCATGCTAGCTATGCTTCCCTCCAG 450
 P S G D L I V S T R V R V G R S H A S Y G F P P V

TTCTCTTAAGGAGGACCGTATCAACATGGAGAAGTTTCAGTTGAAGCTCTGCAGAGCTTGAGGGTGAAGTGA 525
 L S K E D R I N M E K V S V A E L Q S L E G E L S

GTGGCAGTACTATCTCTCACTGGAATGACCCAGAGACTCAACAAAGTTGACTGCTGACCATTTCTCTTCA 600
 G Q Y Y P L T G M T P E T Q Q K L T A D H F L F N

ATGATTCCGACAGGTTTTTGAAGGCACTGGTGGCTATGATGATTGCCCAACCCGAGAGGAATTTACTTTAAC 675
 D S D R F L K A A G G Y D W P T G R G I Y F N Q

→ AGGACAAGACTCTTGTGCTGGTCAATGAGGAGGATCACCTCAGATTGATCTCTAGCAGAAAGGAGCAACC 750
 D K T F L C W V N E E D H L R L I S M Q K G G N L

← TGGCTGAAGTTTACAAGAGTTGGTCAAGGCTATTGAACAGATGGAGAAGAACTGACCTTTGCAAGAAGGATA 825
 A E V Y K R L V K A I E Q M E K K L T F A K K D N

← Septo.R2 ← Septo.R1
 ACATGGGTTATCTTCTTCTGCCCTAGCAACTTGGGTACCACTGAGAGTAGTGTGACATTAAGATCCCAA 900
 M G Y L T F C P S N L G T T M R A S V H I K I P K

AGGTCTCACAGATGCTGACTTTAAGGAAGTCTGTGATAGATACAATGCAAGCCAGAGGATTCACGGTGAGC 975
 V S Q M P D F K E V C D R Y N L Q A R G I H G E H

ACACAGAAAGCGTTGGTGGAGTCTATGACATCTCAACAAACAGTAGGATGGGTCTTACAGAGTTGAGGCTGTTA 1050
 T E S V G G V Y D I S N K R R M G L T E F E A V T

CAGAGATGATGAGGGGTGTCAACGAGATCATCGTTTGGAGAAAGCTGAACCTGAAGACCATATATGAAAAAGA 1125
 E M M R G V N E I R L E K S

AAAATCCCAAGTAGTTTTTAACAGCAAAATCTGGGATAAACAACCAATAAATAATGATAATTTCAAAGTTC 1200

ACCGTCTCTTCTTCTAGGAATCTTAACCTTCCCAACATAGGCTTATGATTCTGCTAAATGCTTACCAAAA 1275

GATGATAATAAAAGTTGAAACCCCTGAGGG 1306

TABLE II
Enzymatic parameters of molluscan AKs

All parameters were obtained at 25 °C under the same conditions. Values in parentheses are the relative value (%) to *Nautilus* recombinant wild type (WT) AK. The parameters are the average of at least two runs.

Species ^a	Origin	K_m for arginine	$V_{\max}$ or V^b
		mM	$\mu\text{moles } P_i/\text{min/mg of protein}$
<i>Solen</i> ^c	Native	1.48	5.9
<i>Meretrix</i> ^c	Native	0.43	4.5
<i>Aplysia</i>	Native	0.6	ND
<i>Octopus</i>	Native	2.82	4.3
<i>Septoteuthis</i>	Native	1.74	4.9
<i>Nautilus</i>	Native	1.02	10.4
	Recombinant WT	0.68 (100)	7.64 (100)
	Asp ⁶² → Gly	ND	0.045 (0.6) ^b
	Asp ⁶² → Glu	0.67 (99)	0.252 (3.3)
	Ser ⁶³ → Gly	3.45 (516)	0.39 (5.1)
	Ser ⁶³ → Thr	ND	0.022 (0.3) ^b
	Tyr ⁶⁸ → Ser	ND	0.0 (inactive)
	Arg ¹⁹³ → Gly	ND	0.11 (1.5) ^b

^a Data for *Solen*, *Meretrix*, and *Aplysia* are unpublished results by T. Suzuki, T. Taniguchi, Y. Unemi, and T. Higashi.

^b For the very weak enzyme activity, the exact $V_{\max}$ could not be obtained. These values are the observed rates in the presence of 9.5 mM arginine.

^c Two-domain AKs.

DISCUSSION

Evolution of Cephalopod AKs and Phylogenetic Position of *Nautilus*

The cDNA-derived amino acid sequences of *Nautilus*, *Octopus*, and *Septoteuthis* AKs were aligned with that of *Limulus* (horseshoe crab) AK (20) in Fig. 2, which is the only AK whose crystal structure is known (13). The sequence of *Nautilus* AK showed high percentage identity (69–73%) with those of *Octopus* and *Septoteuthis* AKs. The identity between *Octopus* and *Septoteuthis* AKs is 83%. These cephalopod AK sequences still have a strong homology (53–55%) with that of the arthropod *Limulus* AK.

We have constructed a phylogenetic tree with the maximum likelihood method (22) for 36 cDNA-derived amino acid sequences of AK, CK, GK, and LK so far known, including four major molluscan classes Bivalvia, Polyplacophora, Gastropoda, and Cephalopoda. In Fig. 3, a part of the tree with all molluscan sequences is shown. The similar topology is obtained with the Neighbor-Joining method (23). The tree clearly shows that the sequences of the clams *Pseudocardium*, *Solen*, and *Corbicula* (Bivalvia) diverged first in molluscan AKs and was followed by the radiation of Polyplacophora (chiton *Liolophura*), Gastropoda (*Nordotis*, *Battilus*, *Cellana*, and *Aplysia*), and Cephalopoda (*Nautilus*, *Octopus*, and *Septoteuthis*). Although the tree could not make clear the order of divergence of the three classes Polyplacophora, Gastropoda, and Cephalopoda, it suggests that the biological position of the chiton *Liolophura*, which is believed to be one of the primitive molluscs, should be reconsidered. Although the origin and evolution of cephalopods have long been debated, this work provides the first molecular evidence, as far as we know, for the phylogenetic position of Cephalopoda in molluscan evolution. Consistent with the morphological data, the branching pattern in Cephalopoda suggests that *Nautilus* (subclass Tetrabranchia, order Nautiloidea) is the most primitive, and *Octopus* (subclass Dibranchia, order Octapod) and *Septoteuthis* (subclass Dibranchia, order Decapoda) share a common immediate ancestor.

We described before that the amino acid sequences of molluscan AKs could be used as a molecular clock to elucidate the divergence time of molluscs, because the enzymes appear to be derived from a single gene strain and have a moderate evolutionary rate (24). A maximum likelihood clock-like tree constructed from the 36 amino acid sequences allows us to estimate the divergence time of the interested molluscan species. If we assume the divergence between Mollusca and Arthropoda occurred about 550 million years before present (25), the radiation time of Polyplacophora, Gastropoda, and Cephalopoda is estimated to be 290 million years before present (data not shown).

Role of Amino Acid Residues of the GS Region in AK Function

Substrate-binding Residues—Amino acid sequences around the GS region of phosphagen kinases are shown in Fig. 4. This GS region is overlapped partly by the so-called flexible loop in the crystal structure of chicken mitochondrial CK, which has been speculated to move nearer to the active site to exclude water during catalysis (3). Clearly, there is a proportional relationship between the size of the deletion in GS region and the mass of guanidine substrate used. Namely, LK and AK, which use relatively large guanidine substrates, lombricine and arginine, have a five-amino acid deletion in this region, CK has one amino acid deletion, and GK has no deletion. It is possible that the introduction of an amino acid deletion as in the GS region in Fig. 4 shortens the flexible loop and allows the active site to accommodate a larger guanidine substrate.

Recently, a crystal structure for the transition state analog complex (TSAC) of the AK from the horseshoe crab *Limulus* has appeared (13). The structure showed that four amino acid residues (Ser⁶³, Gly⁶⁴, Val⁶⁵, and Tyr⁶⁸ in *Limulus* AK sequence), located in the GS region (see Fig. 4), are hydrogen-bonded with the amino and carboxylate groups of the substrate arginine. These residues are conserved in all three AK sequences from *Nautilus*, *Octopus*, and *Septoteuthis*, suggesting that the substrate-binding mechanism of cephalopod AKs is the same as that of *Limulus* AK.

Among molluscan AKs, the four residues associated with substrate binding are conserved in most AKs. However, residue 64 or 65 is replaced by Ser, Lys, or Cys in the two-domain AKs from the clams *Pseudocardium*, *Corbicula*, and *Solen* and the sea anemone *Anthopleura*. Thus it appears that the two-domain AKs may have a unique substrate-binding system.

We introduced two independent mutations, Ser⁶³ → Gly and Ser⁶³ → Thr, in the 63rd position of *Nautilus* AK (Table II). At least one of the mutant proteins (Ser⁶³ → Gly) had a considerably reduced substrate affinity (K_m = 3.45 mM) compared with the wild type (K_m = 0.68 mM), accompanied also by a decreased $V_{\max}$ (less than 5% to the wild-type). In the case of Ser⁶³ → Thr mutation, the AK activity was lost almost completely, probably due to a steric hindrance.

The Tyr⁶⁸ is conserved in all AKs, including two-domain clam AKs, except in the sea cucumber *Stichopus* AK, which evolved not from the AK gene but from the CK gene (9). *Stichopus* AK has a completely different substrate-binding system (see the sequence of *Stichopus* AK in Fig. 4), which will be described elsewhere. The aromatic side chain of Tyr⁶⁸ forms a hydrogen-bond with the amino group of the substrate arginine in *Limulus* AK, and thus the Tyr⁶⁸ appears to be the most important residue in substrate binding. Consistent with this, the enzyme activity of *Nautilus* Tyr⁶⁸ → Ser mutation was lost

FIG. 1. Nucleotide and derived amino acid sequences of cDNAs of *N. pompilius* (A), *O. vulgaris* (B), and *S. lessoniana* (C) AKs. Arrows indicate primers used for amplification. Amino acid sequences determined chemically are shown by dotted lines.

FIG. 2. Alignment of amino acid sequences of *N. pompilius*, *O. vulgaris*, and *S. lessoniana* AKs with that of *Limulus* AK. Four amino acid residues (Ser⁶³, Gly⁶⁴, Val⁶⁵, and Tyr⁶⁸ in *Limulus* sequence numbering) interacting with the amino and carboxylate groups of the substrate arginine are boxed by gray lines. Asp⁶² and Arg¹⁹³ (in *Limulus* sequence numbering) are indicated by #. The numbering in Fig. 2 is for the "consensus sequence" to the four sequences, and the *Limulus* Arg¹⁹³ corresponds to the position 195 in Fig. 2.

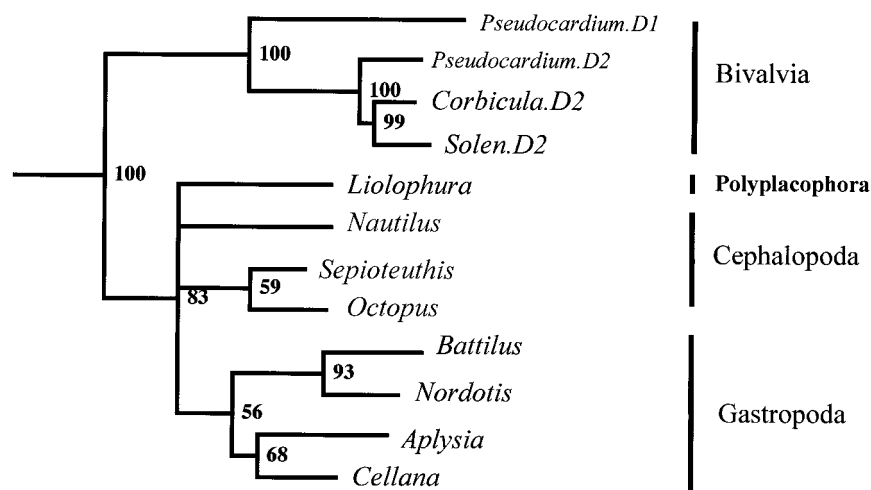
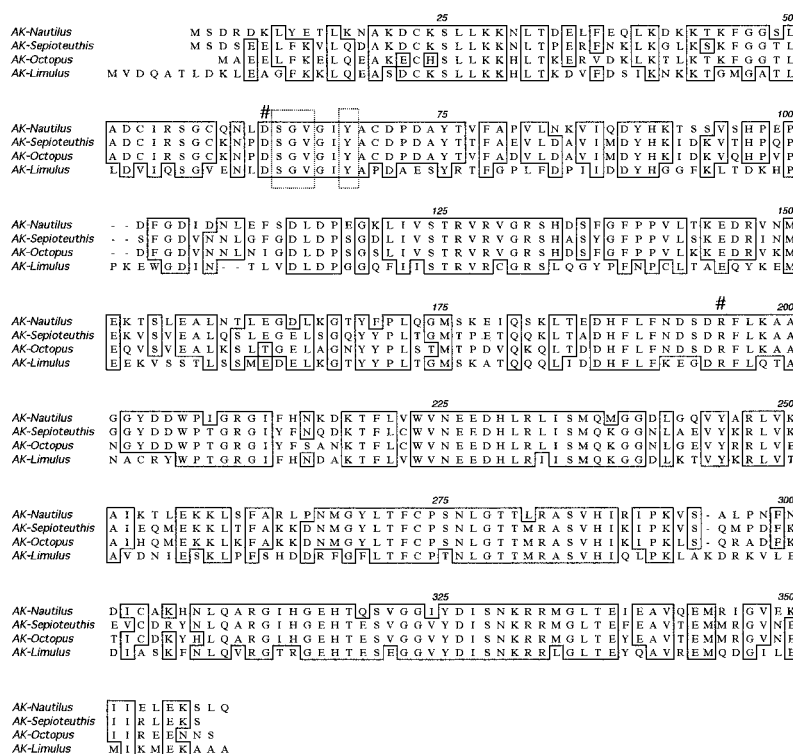


FIG. 3. A phylogenetic tree for amino acid sequences of molluscan AKs. This tree was constructed with the maximum likelihood method (17). The references for sources of sequences: *Pseudocardium* (11); *Liolophura* and *Battillus* (9); *Nordotis* (8); *Corbicula.D2*, *Solen.D2*, *Aplysia*, and *Cellana* (Suzuki *et al.*²).

completely (Table II).

Role of Asp⁶²—We noted previously that the Asp⁶² in the GS region is conserved in most AK sequences, including *Stichopus* AK, which evolved from the CK gene, but not in other phosphagen kinases CK, GK, and LK (9, 14) (see Fig. 4). The TSAC crystal structure of *Limulus* AK indicates that Asp⁶² is not associated with the substrate binding, and the authors have not referred to any functional role (13). On the other hand, we assumed previously that Asp⁶² has a special role in the recognition of the substrate, arginine (14), and the first domain of *Pseudocardium* two-domain AK might not retain a complete enzyme activity, because the Asp⁶² was replaced by a Gly residue (Fig. 4).

The TSAC structure of *Limulus* AK indicates that it consists of two domains; an N-terminal small α -helical domain (N-domain) of residues 1–111 and a C-terminal domain (C-domain) possessing an 8-stranded antiparallel β -sheet flanked by 7 α -helices of residues 112–357 (13). One of the two substrates, ATP or ADP, is accommodated in the C-domain, and another

substrate, arginine or arginine phosphate, contacts mainly with the N-domain. The catalytic center, where reversible transfer of the phosphate occurs, is located in the C-domain. We found that in the TSAC structure of *Limulus* AK, viewed by SwissPdbViewer (PDB Id: 1BG0), the side chain of Asp⁶² in the N-domain is hydrogen-bonded with that of Arg¹⁹³ in the C-domain (Fig. 5), and that Arg¹⁹³ is conserved in all AKs. Here we hypothesize that the hydrogen bond between Asp⁶² and Arg¹⁹³ plays a crucial role in AK function, namely, that the two domains are maintained in a favorable topology by the hydrogen bond, where the two kinds of substrates are accessible enough for reaction.

If the 3-dimensional structure of *Nautilus* AK is homologous with that of *Limulus* AK (13), site-directed mutagenesis studies using *Nautilus* AK strongly support the above hypothesis. Both of the mutations Asp⁶² → Gly and Arg¹⁹³ → Gly, which break the hydrogen bonding completely, markedly reduced enzyme activity (less than 1.5% of the wild-type). The same result was obtained from a mutagenesis study using the Asp⁶² → Gly

	56	62	63	65	68	71
AK- <i>Limulus</i>	SGVENLD		SGVGIYAPD			
AK- <i>Nautilus</i>	SGCQNLD		SGVGIYACD			
AK- <i>Octopus</i>	SGCKNPD		SGVGIYACD			
AK- <i>Sepioteuthis</i>	SGCKNPD		SGVGIYACD			
AK-Lobster	SGVENLD		SGVGIYAPD			
AK-Shrimp	SGVENLD		SGVGIYAPD			
AK-GrassHopper	SGLENHD		SGVGIYAPD			
AK- <i>C.elegance</i>	SGVANLD		SGVGVIYAPD			
AK-SeaAnemone.D1	SGVNLND		SGVGVIYAGD			
AK-SeaAnemone.D2	SGVENLD		SSCGVIYAGD			
AK- <i>Battilus</i>	SGCLNLD		SGVGIYACD			
AK- <i>Nordotis</i>	SGCLNLD		SGVGIYACD			
AK- <i>Liolophura</i>	SGAKNLD		SGVGVIYASD			
AK-Solen.D2	SGCLHMG		SKVGIYACD			
AK- <i>Corbicula</i> .D2	SGCIHLG		SKVGIYACD			
AK- <i>Pseudocardium</i> .D1	SGALHPG		SKVGIYACD			
AK- <i>Pseudocardium</i> .D2	SGSLHLD		SKVGIYACD			
AK-SeaCucumber	NGVDNPD		FHLGLLAGD			
CK-Human.M	TGVDNPGHPF		IMTVGCVAGD			
CK-Rabbit.M	TGVDNPGHPF		IMTVGCVAGD			
CK-Chicken.M	TGVDNPGHPF		IMTVGCVAGD			
CK- <i>Torpedo</i>	TGVDNPGHPF		IMTVGCVAGD			
CK-Human.B	TGVDNPGHPY		IMTVGCVAGD			
CK-Rabbit.B	TGVDNPGHPF		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</i>	TGVDNPGNKFYGKKTGCVFGD					

FIG. 4. Alignment of the amino acid sequence of a possible substrate recognition region (GS region) of phosphagen kinases. The accession numbers of the enzymes in the DNA or protein data bases (from EMBL, GenBank™, DDBJ, or SwissProt) are as follows: AK *Limulus* (U09809); AK, *Nautilus* (AB042330); AK, *Octopus* (AB042331); AK, *Sepioteuthis* (AB042332); AK, lobster (X68703); AK, shrimp (P51545); AK, grasshopper (U77580); AK, *C. elegans* (Z75714); AK, sea anemone (AB008014); AK, *Battilus* (AB008011); AK, *Nordotis* (D26104); AK, *Liolophura* (AB008012); AK, *Pseudocardium* (AB017255); AK, sea cucumber (AB025275); CK, Human.M (KIHUCM); CK, Rabbit.M (KIRBCM); CK, Chicken.M (KICHCM); CK, *Torpedo* (AAA49278); CK, Human.B (KIHUCB); CK, Rabbit.B (KIRBCB); CK, Chicken.B (AAA48687); CK, Chicken.Mi (A27708); CK, Human.Mi (A30789); LK, *Eisenia* (AB008013); GK, *Neanthes* (AB017256). AK, *Solen*.D2 and AK, *Corbicula*.D2, unpublished.

mutant of *Stichopus* AK.² In addition, the results of mutation Asp⁶² → Glu provides strong evidence for the functional role of Asp⁶². This mutation, like the former two mutations, could be expected not to change the substrate affinity for arginine, but it could be expected to have ability to form a hydrogen bond with Arg¹⁹³. Enzymatic parameters are consistent with this idea;

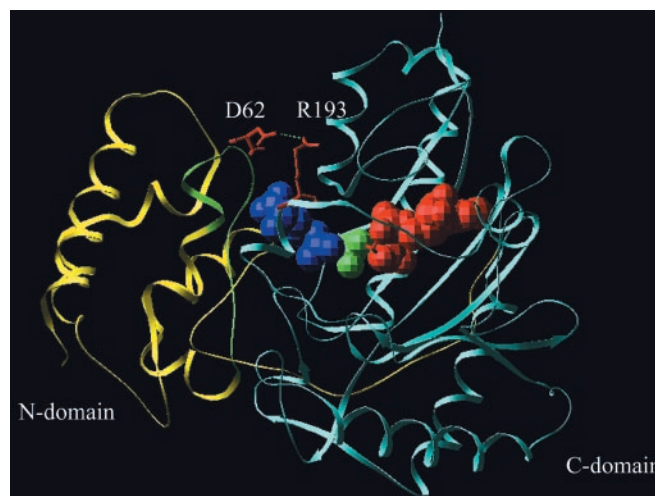


FIG. 5. A hydrogen-bond between Asp⁶² and Arg¹⁹³ in AK. This image was made by the SwissPdbViewer using *Limulus* AK (12) as a template (PDB Id: 1BG0). N-terminal small domain (N-domain) is colored in yellow, and C-terminal large domain (C-domain) in light blue. The GS region in the N-domain, which is overlapped partly by the so-called flexible loop in the crystal structure of chicken mitochondrial CK (21) and *Limulus* AK (12), is colored in green. The substrate arginine, nitrate, and ADP are shown by blue, green, and red balls, respectively.

the K_m (0.67 mM) is the same as that (0.68 mM) of wild-type, and the V_{max} is reduced largely but not so much as that of the Asp⁶² → Gly or Arg¹⁹³ → Gly mutations, probably due to a partial formation of hydrogen bond between Glu⁶² and Arg¹⁹³.

Zhou *et al.* (13) have suggested that it is the precise orientation of the substrates in the catalytic pocket of AK that dominates the forces driving catalysis. Substantial conformational changes take place upon substrate binding in CK (26). The recent crystal structures of CK (3–5) and AK (13) provide very compelling insight into the nature of these structural changes. The extent domain and flexible loop movements during catalysis have recently been modeled by Forstner *et al.* (27) for CK and by Zhou *et al.* (28) for AK.

Our results and the above models show that, upon substrate binding, the flexible loop bearing Asp⁶² moves nearer to the active center to exclude water during catalysis, and the hydrogen bond between Asp⁶² and Arg¹⁹³ is formed to link the N- and C-domain of AK (Fig. 5). This bond stabilizes the closed, substrate-bound state. The mutant (Asp⁶² → Gly or Arg¹⁹³ → Gly) destabilizes the maintenance of the closed state and/or perhaps disrupts the unique topology of the catalytic pocket; therefore, these mutants show only a very weak activity (0.6–1.5% to the wild-type, see Table II). The mutant (Asp⁶² → Glu), while allowing hydrogen bonding between Glu⁶² and Arg¹⁹³, may also disturb the topology of the catalytic pocket.

Our results are highly consistent with prevailing views on the catalytic mechanism of phosphagen kinases. Maintenance of the precise three-dimensional structure of the active site is critical for optimum efficiency of enzyme function. Thus, we conclude that Asp⁶² (or Arg¹⁹³) is one of the key residues in AK function.

Acknowledgments—We thank Dr. Ross Ellington for critical reading of the manuscript and giving us helpful suggestions and Takuya Higashi for performing experiments on amino acid sequencing of *Nautilus* AK.

REFERENCES

1. Meyer, R. A., Sweeney, H. L., and Kushmerick, M. J. (1984) *Am. J. Physiol.* **246**, C365–C377.
2. Wallimann, T., Wyss, M., Brdiczka, D., Nicolay, K., and Eppenberger, H. M. (1992) *Biochem. J.* **281**, 21–40.
3. Fritz-Wolf, K., Schnyder, T., Wallimann, T., and Kabsch, W. (1996) *Nature*

² T. Suzuki, Y. Yamamoto, and M. Umekawa, unpublished.

- 381, 341–345
4. Rao, J. K., Bujacz, G., and Wlodawer, A. (1998) *FEBS Lett.* **439**, 133–137
5. Eder, M., Schlattner, U., Becker, A., Wallimann, T., Kabsch, W., and Fritz-Wolf, K. (1999) *Protein Sci.* **8**, 2258–2269
6. Watts, D. C. (1971) in *Biochemical Evolution and the Origin of Life* (Schoffeniels, E., ed) pp. 150–173, North-Holland, Amsterdam
7. Ellington, W. R. (1989) *J. Exp. Biol.* **143**, 177–194
8. Watts, D. C. (1975) *Symp. Zool. Soc. Lond.* **36**, 105–127
9. Suzuki, T., Kamidochi, M., Inoue, N., Kawamichi, H., Yazawa, Y., Furukohri, T., and Ellington, W. R. (1999) *Biochem. J.* **340**, 671–675
10. van Thoai, N. (1968) in *Homologous Enzymes and Biochemical Evolution* (van Thoai, N., and Roche, J., eds) pp. 199–229, Gordon and Breach, New York
11. Suzuki, T., Kawasaki, Y., and Furukohri, T. (1997) *Biochem. J.* **328**, 301–306
12. Suzuki, T., Kawasaki, Y., Furukohri, T., and Ellington, W. R. (1997) *Biochim. Biophys. Acta* **1348**, 152–159
13. Zhou, G., Somasundaram, T., Blanc, E., Parthasarathy, G., Ellington, W. R., and Chapman, M. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 8449–8454
14. Suzuki, T., Kawasaki, Y., Unemi, Y., Nishimura, Y., Soga, T., Kamidochi, K., Yazawa, Y., and Furukohri, T. (1998) *Biochim. Biophys. Acta* **1388**, 253–259
15. Winneppenninckx, B., Backeljau, T., and De Wachter, R. (1995) *Mol. Biol. Evol.* **12**, 641–649
16. Winneppenninckx, B., Backeljau, T., and De Wachter, R. (1996) *Mol. Biol. Evol.* **13**, 1306–1317
17. Steiner, G., and Muller, M. (1996) *J. Mol. Evol.* **43**, 58–70
18. Adamkewicz, S. L., Harasewych, M. G., Blake, J., Saudek, D., and Bult, C. J. (1997) *Mol. Biol. Evol.* **14**, 619–629
19. Winneppenninckx, B., Backeljau, T., and De Wachter, R. (1994) *Nautilus* **2**, (suppl.) 98–110
20. Strong, S. J., and Ellington, W. R. (1995) *Biochim. Biophys. Acta* **1246**, 197–200
21. Suzuki, T., and Furukohri, T. (1994) *J. Mol. Biol.* **237**, 353–357
22. Strimmer, K., and von Haeseler, A. (1996) *Mol. Biol. Evol.* **13**, 964–969
23. Felsenstein, J. (1993) PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle, U. S. A.
24. Suzuki, T., Ban, T., and Furukohri, T. (1997) *Biochim. Biophys. Acta* **1340**, 1–6
25. Goodman, M., Pedwaydon, J., Czelusniak, J., Suzuki, T., Gotoh, T., Moens, L., Shishikura, F., Walz, D., and Vinogradov, S. (1988) *J. Mol. Evol.* **27**, 236–249
26. Forstner, M., Kriechbaum, M., Laggner, P., and Wallimann, T. (1996) *J. Mol. Struct.* **383**, 217–227
27. Forstner, M., Kriechbaum, M., Laggner, P., and Wallimann, T. (1998) *Biophys. J.* **75**, 1016–1023
28. Zhou, G., Ellington, W. R., and Chapman, M. S. (2000) *Biophys. J.* **78**, 1541–1550