# Ancient divergence of long and short isoforms of adenylate kinase: molecular evolution of the nucleoside monophosphate kinase family

Kaoru Fukami-Kobayashi[a], Michiko Nosaka[b], Atsushi Nakazawa[c], Mitiko Gō[a,]*

[a]*Department of Biology, Faculty of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-01, Japan*
[b]*Chemical Engineering, Sasebo College of Technology, Sasebo 857-11, Japan*
[c]*Department of Biochemistry, Yamaguchi University School of Medicine, Ube 755, Japan*

Received 9 March 1996; revised version received 26 March 1996

**Abstract Adenylate kinases (AK) from vertebrates are separated into three isoforms, AK1, AK2 and AK3, based on structure, subcellular localization and substrate specificity. AK1 is the short type with the amino acid sequence being 27 residues shorter than sequences of the long types, AK2 and AK3. A phylogenetic tree prepared for the AK isozymes and other members of the nucleoside monophosphate (NMP) kinase family shows that the divergence of long and short types occurred first and then differentiation in subcellular localization or substrate specificity took place. The first step involved a drastic change in the three-dimensional structure of the LID domain. The second step was caused mainly by smaller changes in amino acid sequences.**

*Key words:* Adenylate kinase; Isozyme; Subcellular localization; Nucleoside monophosphate kinase; Phylogenetic tree; Molecular evolution

## 1. Introduction

Adenylate kinase (AK) is a ubiquitous enzyme which contributes to homeostasis of adenine nucleotides in living cells. Vertebrates have three known isoforms of AKs; AK1 in the cytosol, AK2 in the mitochondrial intermembrane space and AK3 in the mitochondrial matrix. AK1 and AK2 catalyze the reversible transfer of the $\gamma$-phosphate group from MgATP to AMP. The isoform AK3 utilizes GTP more efficiently than ATP (for reviews see [1,2]).

There is size variation among the isoforms; AK1 is the short type consisting of 194 amino acid residues, while AK2 and AK3 are the long type, with 27 additional residues in the central portion of their amino acid sequences. The crystal structures of both types of AKs [3-6], with or without bound substrates have been reported. A comparison of these structures revealed that two lobes undergo large movements with substrate binding [7]. The movements shield the active center of AK from water molecules and prevent ATP hydrolysis. The 3D structure of AK is decomposed into three subdomains, based on the functional roles and induced-fit movements; the NMPbind and LID domains, the moving parts, and the CORE domain that is unaffected by substrate binding [8,9]. The long and short type AKs differ in the LID domain; LID is an 11-residue segment in the short type, whereas that in the long type consists of 38 residues and the difference led to drastic change in the conformation of the LID domain.

AKs isolated from prokaryotes belong to the long type [10,11]. AK localized to the chloroplast is also the long type [12]. If we consider the chloroplast AK and the mitochondrial AKs to have a prokaryotic origin in view of the endosymbiont hypothesis, AKs are apparently subdivided into the eukaryotic short type and the prokaryotic long type. However, there are instances where this classification does not hold. One is the cytosolic AK from *S. cerevisiae* [13]. This enzyme is the long type unlike cytosolic AKs from vertebrates, AK1. The second exception is the long type AK from *Giardia* [14]. *Giardia* belongs to the most primitive group in the Eukaryotae, and does not have mitochondria. It is unlikely, therefore, that the *Giardia* AK was derived from mitochondria. The third is the short type AK from bacteria, *Micrococcus* [15]. These findings led to the postulation that the long and short type of AKs may have diverged before the appearance of eukaryotes, and this would explain why both types of AKs are found in eukaryotes and prokaryotes.

In addition, it was reported that UMP-CMP kinase shows high sequence homology with AK [16,17] to make up nucleoside monophosphate (NMP) kinase family together with AKs. All known UMP-CMP kinases are of the short type. It is thus of interest to determine whether the long and short types of AKs diverged prior to the appearance of eukaryotes. To elucidate the evolutionary relationship of members of the NMP kinase family with different subcellular localizations, different substrate specificities and different sizes, a molecular tree for the family was constructed.

## 2. Materials and methods

34 amino acid sequences of NMP kinases were deduced from nucleotide sequences in DDBJ release 20.0, except that AK from maize chloroplast was obtained from PIR. For an outgroup of the NMP kinase family, three amino acid sequences of guanylate kinase (GK) were collected, in the same manner. Alignment between long and short types was performed by assigning the big gap which reflects size difference to the LID regions. Since the primary structure of the NMPbind domain of GK differs from that of AK [18], the amino acid sequences of GKs were aligned to those of AKs only in regions building up to the CORE and LID domains. Fig. 1 shows an extract from the alignment.

The evolutionary distance among the amino acid sequences was computed from the sequence data in these regions, taking into account the evolutionary similarity [19]. Data on the accepted point mutations and frequencies for the 20 amino acid residues [20] were used for computation. Using the distance matrix, two phylogenetic trees were constructed by the neighbor-joining method [21]; one is for the NMP kinase family only, and the other includes guanylate kinases. When constructing the latter tree, NMPbind domain and its adjacent region (denoted '#' in Fig. 1) was not used for computation of evolutionary distances because of the lack of alignment for the region. To test the reliability of each branch, bootstrap resampling [22] was carried out 1000 times for the respective trees.

---

*Corresponding author. Fax: (81) (52) 789 2977.
E-mail: go@bio.nagoya-u.ac.jp

*PII* S0014-5793(96)00367-5

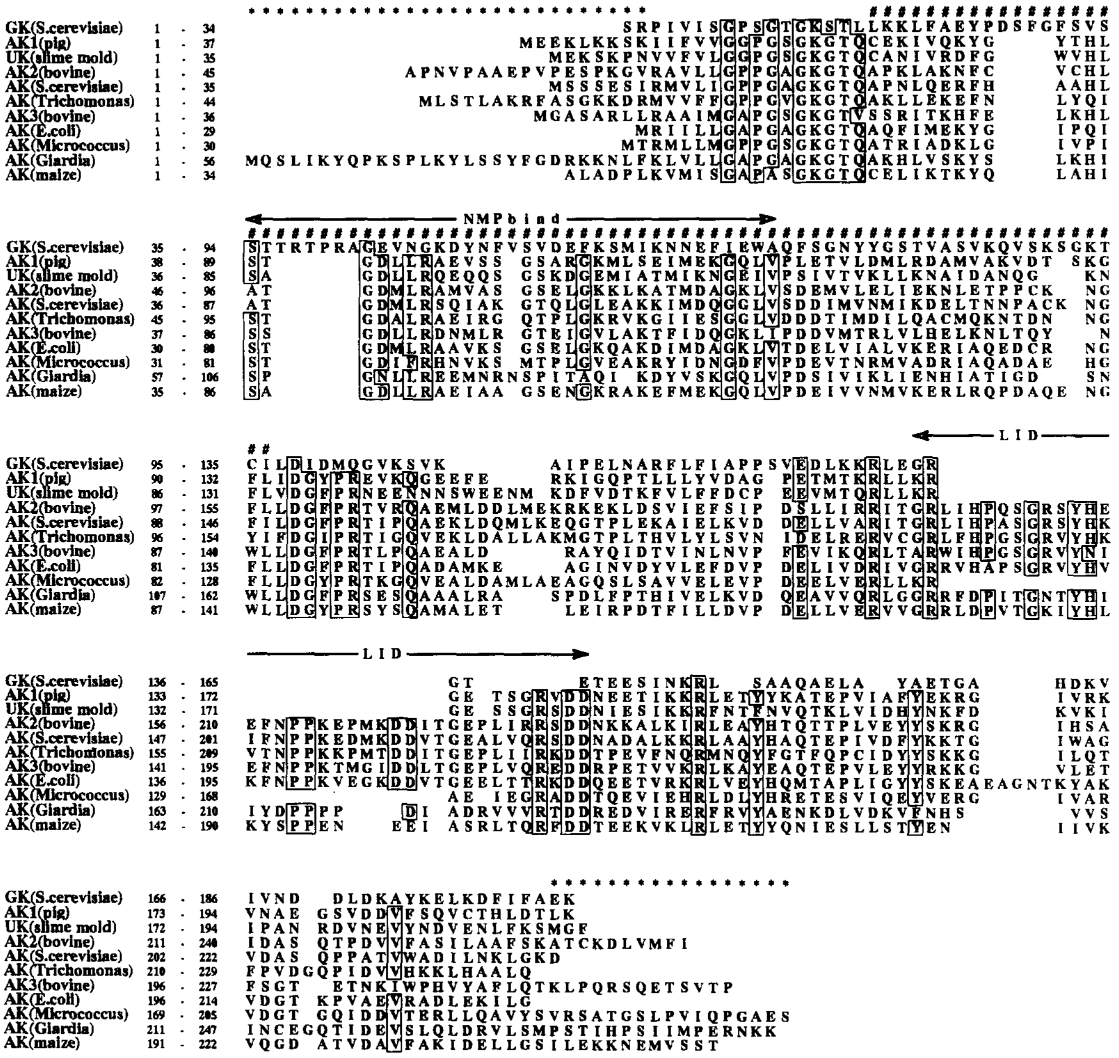


Fig. 1. Alignment of amino acid sequences of the NMP kinase family. Conservative residues are framed. The ranges of NMPbind and LID domains are defined according to the alignment by Vonrhein et al. [9]. Asterisks (*) mark N- and C-terminal regions that were not used for construction of both trees in Fig. 2 because of the irregular lengths among the sequences. Sharps (#) indicate NMPbind domain and its adjacent region that were not used for construction of the tree including GK (Fig. 2B), because GK and AK could not be aligned in the region. For references of sequences in the alignment, see Fig. 2 legend.

## 3. Results

In the tree for the NMP kinase family (Fig. 2A), NMP kinases could be distinctly divided into long and short types. The branch separating the two types appeared in all trials of the bootstrap resampling. The long type AKs from *S. cerevisiae* and *Giardia* were also included in the long type cluster. The only exception is the *Micrococcus* AK. Although this protein belongs to the short type AK (Fig. 1), it clustered with other prokaryotic long type AKs.

The AK from *Giardia* did not cluster with mitochondrial AK2 nor AK3. The result is consistent with the view that the AK from *Giardia* was not derived from mitochondria.

Isoforms with a different subcellular localization, AK1, AK2, AK3 and AK(maize) formed their own subclusters in the long or short type cluster. Bootstrap probabilities indicate that these subclusters are significant. The cytosolic AK from *S. cerevisiae* clustered with the mitochondrial AK2 rather than with the short type AK1. The AK3 subcluster implies that the substrate specificity for GTP was acquired after the AK3 subcluster diverged from the AK2 subcluster. The substrate specificity was changed to accept UMP/CMP in an ancestor of UMP-CMP kinases in the short type cluster, and in an ancestor of the chloroplast AK in the long type cluster, independently of each other.

To locate a root of the NMP kinase family tree, a phylogenetic tree including GKs as an outgroup (Fig. 2B) was constructed. The GK cluster was connected to the branch separating the long and short types (the point designated '*' in the tree), indicating that the divergence of the types of AK is the first event occurred during the evolution of the NMP kinase family. The statistical significance of the location, however, was not obtained; the bootstrap probability of the branch separating the short type AKs and the others was 70%, and that of the branch separating the long type AKs and the others was only 27%. The tree including GKs showed lower probabilities for other branches as compared with the tree for the NMP kinase family only. The poor performance is responsible for exclusion of the NMPbind domain and its adjacent region from computation of evolutionary distance.

## 4. Discussion

Locating the root of the phylogenetic tree on the branch separating the long and short type, as indicated by the GK cluster, the evolution of the NMP kinase family was reconstructed. The long and short type isoforms appeared first, and then isoforms with different subcellular localization or substrate specificities diverged in both types (Fig. 3). The common ancestor of the long and short types had probably either type of LID domain, since the domain is important for the catalytic reaction of the NMP kinase family. It is not certain, however, which type of LID the ancestor had. The phylogenetic relationship of the NMP kinase family implies that in-

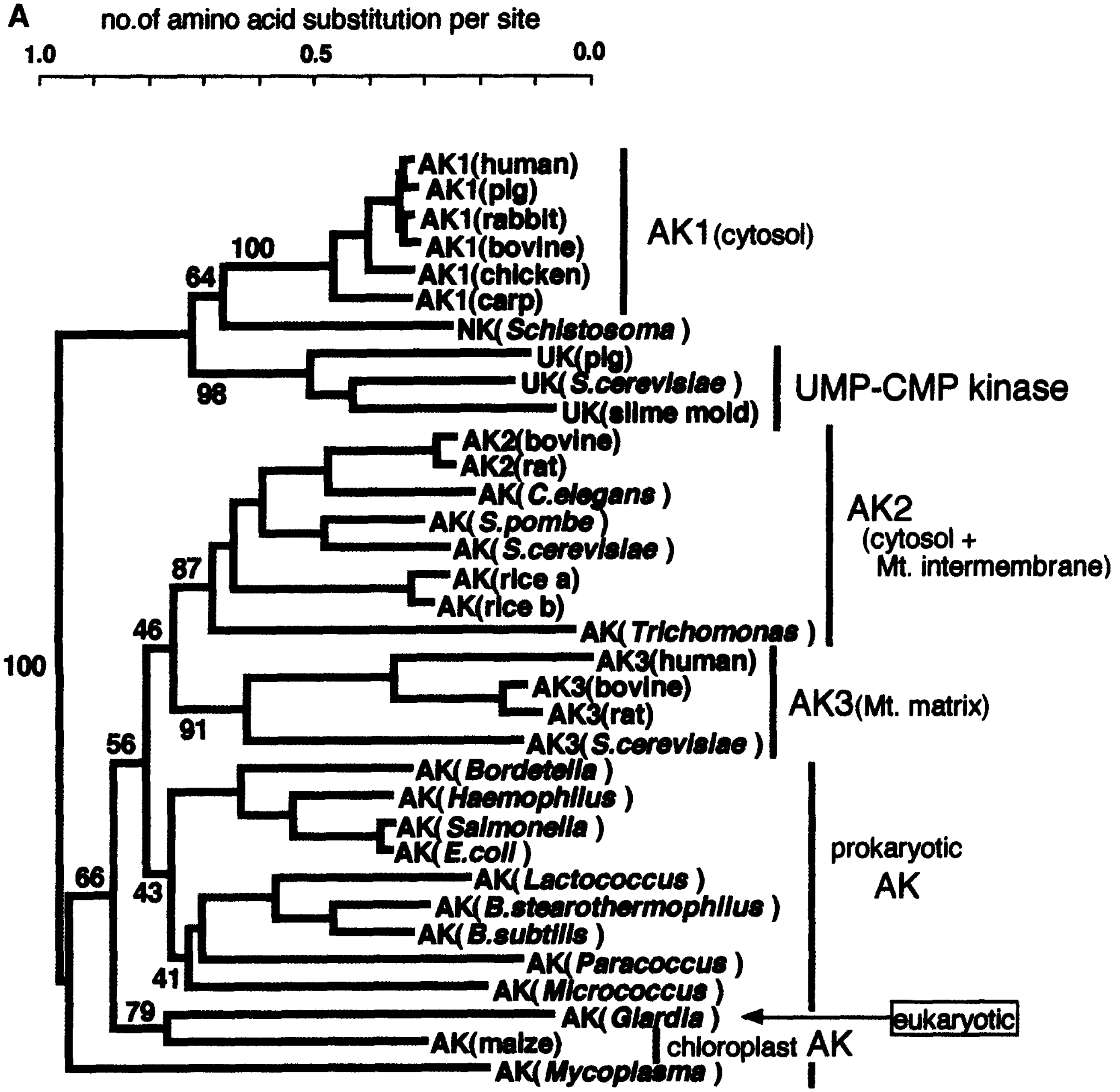
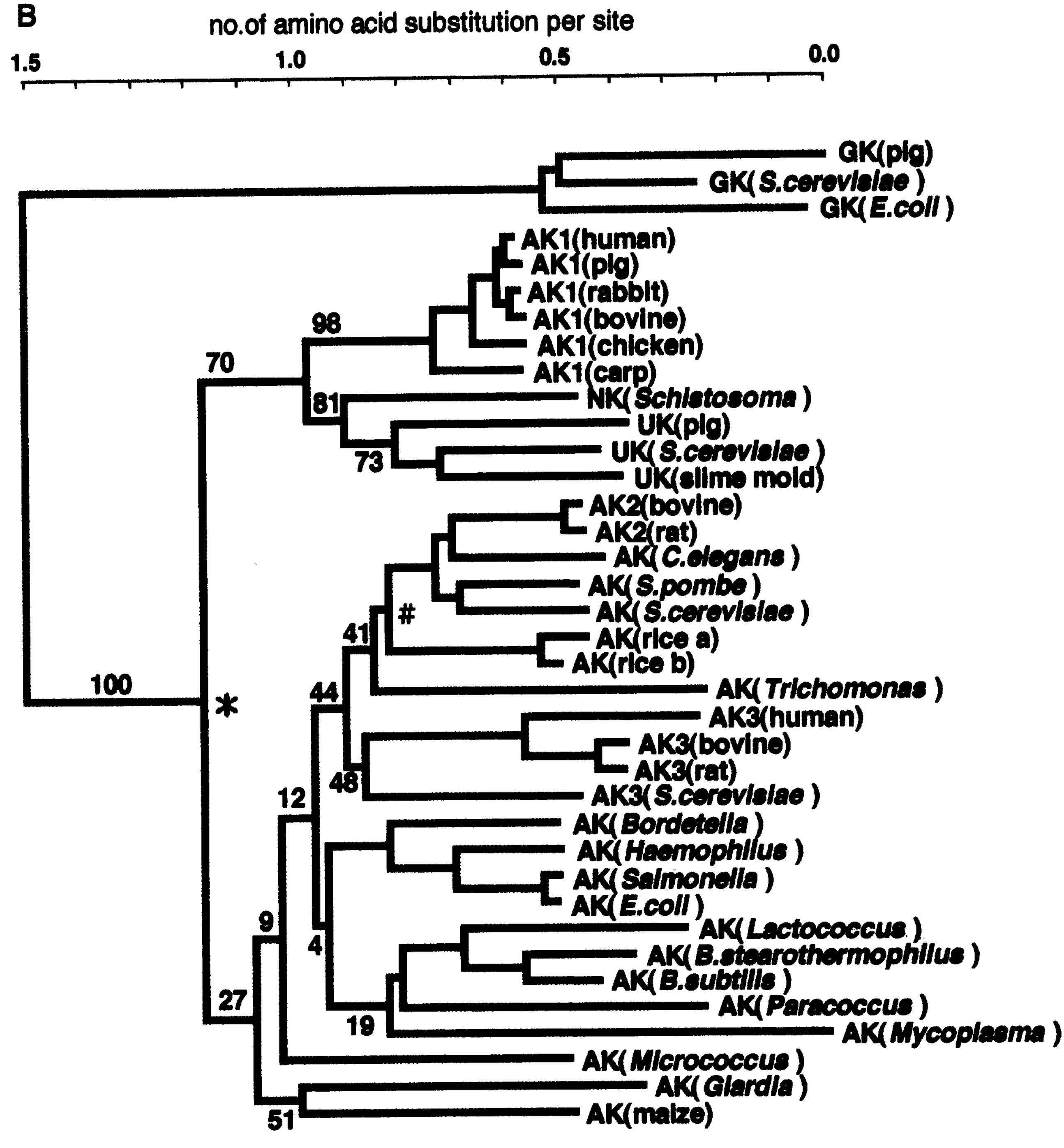


Fig. 2. Phylogenetic trees of the NMP kinase family. (A) The members of the NMP kinase family only. (B) Including guanylate kinases as outgroups. Evolutionary paths of the long and short types of the NMP kinases and those of GKs are shown in green, rose and orange, respectively. Bootstrap probabilities for the relevant branches are indicated. An asterisk indicates the root position of the NMP kinase family tree, which also corresponds to the point of the divergence between the long and short types of this family. A sharp (#) points to the divergence of the eukaryotic three kingdoms in the AK2 subgroup. The sequences used in the tree construction are AK1s from human [33], pig [34], rabbit [35], bovine [35], chicken [36] and carp [37], NK (AK homolog, substrate unknown) from *Schistosoma* [38], UMP-CMP kinases from pig [39] and slime mold [16], uridylate kinase from *S. cerevisiae* [40], AK2s from bovine [41] and rat [42], AKs from *C. elegans* [43], *S. pombe* [27], *S. cerevisiae* [13], rice [44], *Trichomonas* [31], AK3s from human [45], bovine [46], rat [42] and *S. cerevisiae* [47], prokaryotic AKs from *Bordetella* [48], *Haemophilus* [49], *Salmonella* [50], *E. coli* [51], *Lactococcus* [24], *B. stearothermophilus* [52], *B. subtilis* [23], *Paracoccus* [53], *Mycoplasma* [54] and *Micrococcus* [15], protozoan AK from *Giardia* [14], chloroplast AK from maize [12] and GKs from pig [55], *S. cerevisiae* [56] and *E. coli* [57].



sertion or deletion of 20–30 amino acid residues in homologous proteins reflects the more distant evolutionary relationship than do amino acid replacements.

The short type *Micrococcus* AK clustered with the long type prokaryotic AKs in the trees, reflecting a phylogenetic relationship of species from which the prokaryotic AKs derived. The relationship does not support a horizontal transfer of the *Micrococcus* AK from eukaryotes. Similarity on the gene organization of *adk* and its flanking genes of *Micrococcus* to that of other prokaryotes [15,23,24] also rules out the possibility of a horizontal transfer. The *Micrococcus* AK likely evolved independently of the vertebrate AK1; it is considered to have been formed by the secondary loss of the 27 residues in the LID domain from a long type AK at an early stage of divergence of prokaryotic AKs (Fig. 3). Rose et al. [25] reported that removal of 25 residues corresponding to the additional residues in the long type LID domain was not critical for catalysis. The experimental result might be supporting evidence of the secondary loss.

The clustering of AK2 and AK3 with prokaryotic AKs in the constructed tree confirms their mitochondrial origin that is inferred from the subcellular localization of these enzymes. The phylogenetic tree further indicates that a gene duplication had occurred to generate AK2 and AK3 genes in an ancestor of mitochondria before the divergence of the three eukaryotic kingdoms; animals, plants and fungi. It is suggested that the mitochondrial ancestor already had two AK genes when it was endocytosed into an ancestral eukaryotic cell to begin symbiosis.

The *S. cerevisiae* AK belonged to the AK2 cluster in the

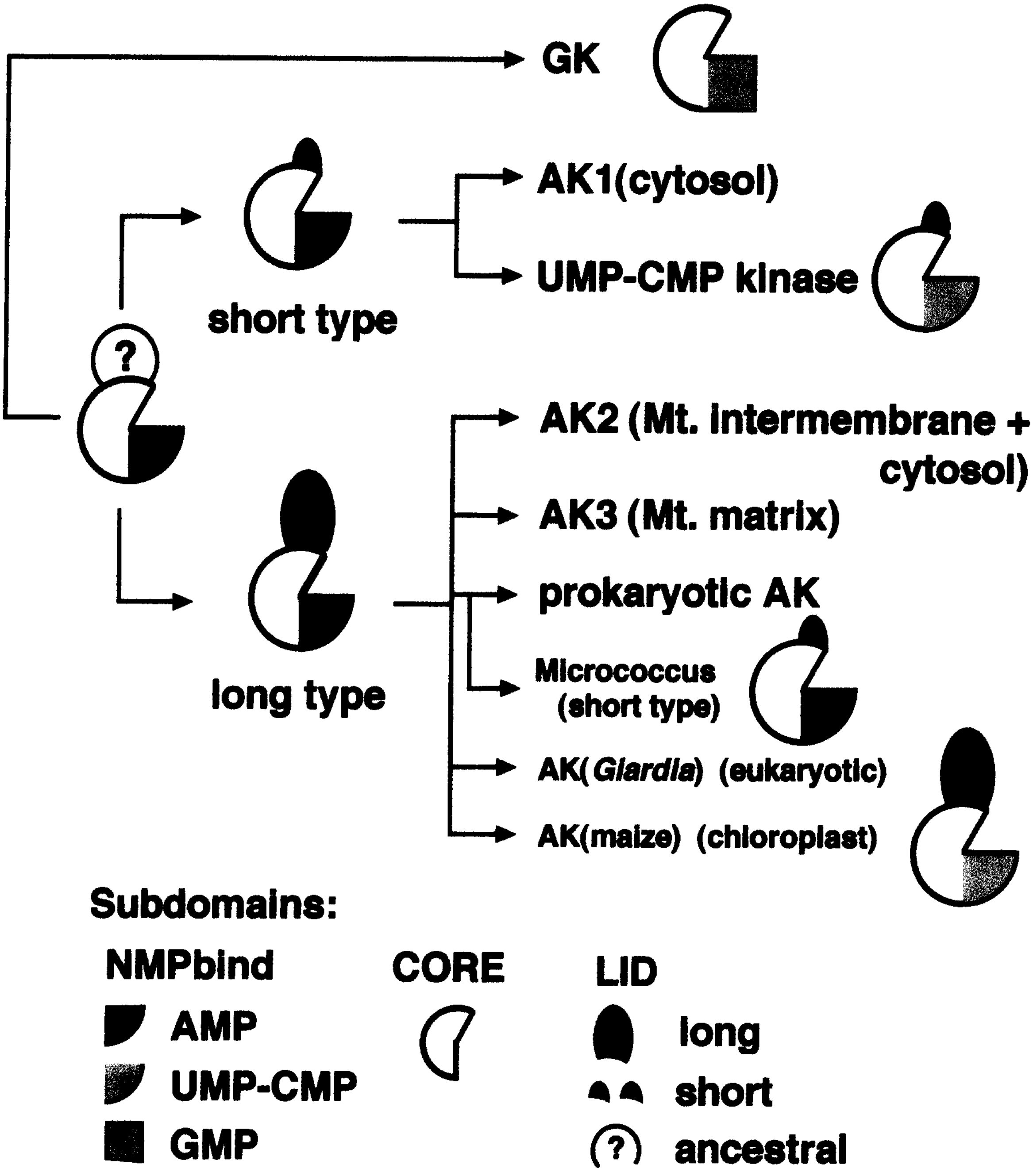


Fig. 3. Putative evolutionary scheme of the NMP kinase family and guanylate kinase.

phylogenetic tree, which seems to be inconsistent with its cytosolic localization. Part of this AK (about 10%), however, has been found to occur in mitochondrial intermembrane [26], where AK2 is localized. Studies on *S. pombe* indicate that the *S. pombe* AK is also present both in the cytosol and in mitochondria [27]. Neither of the AKs from *S. cerevisiae* and *S. pombe* has a cleavable presequence, which would explain their localization. In addition, the AKs from rice have recently been shown to be distributed largely in cytosol [28]. It is thus probable that the subcellular localization of the members in the AK2 cluster is not strictly limited to mitochondrial intermembrane. An experiment shows that the 8 residues from the N-terminal of the *S. cerevisiae* cytosolic AK play a major role in mitochondrial targeting information [29]. This report implies that differentiation of subcellular localization occurred by smaller changes in the sequences than that leading to the divergence of the long and short types.

It has been reported that UMP-CMP kinases from slime mold and pig, *S. cerevisiae* uridylate kinase and maize AK accept UMP/CMP as substrate in addition to AMP. The primary difference in amino acid sequences of these kinases is located at the position corresponding to $Thr^{39}$ in the chicken AK1. All kinases which accept also UMP/CMP have Ala at the position, while the kinases whose substrate is only AMP have Thr or Ser. A site-directed mutagenesis experiment is consistent with the results; it demonstrates that the substrate specificity for UMP/CMP can be acquired by a single replacement of $Thr^{39}$ for Ala in the chicken AK1 [30]. The substrate specificity of *Giardia* AK, which has Pro at the position, has not been documented.

The clustering of hydrogenosomal AK from *Trichomonas* with AK2s is incompatible with the report of Länge et al. [31] that AK from *Trichomonas* has no close relationship to any of the AK subfamilies. The difference is mainly attributed to the method for computation of evolutionary distance; the evolutionary similarity was taken into account in our study, but not so in that of Länge et al. The bootstrap probability of the branch separating the AK2 cluster and the others was 87%. This result indicates a close relationship of the *Trichomonas* AK to mitochondrial AK2 and supports the common evolutionary origin of hydrogenosome and mitochondria.

The estimated number of amino acid substitution per site from the root marked with '*' in Fig. 2B to individual AK2s are roughly twice as much as that from the point of the divergence of the eukaryotic three kingdoms (marked with '#') to individual AK2s. This indicates that the divergence between the long and short type AKs is about twice as old as that of the eukaryotic three kingdoms. It is thus likely that the divergence between the long and short type AKs preceded the divergence between prokaryotes and eukaryotes. Actually, the divergence between the long and short type AKs was shown to have preceded the divergence between prokaryotes and eukaryote *Giardia*. Preliminary results revealed that *Halobacteria* AK belong to the long type cluster (data not shown). The long type subgroup includes AKs from all three urkingdoms. As well as the long type subgruop, it is probable that prokaryotic members will be found in the short type subgroup.

The alternative possibility that AK changed to the short type in an ancestor of eukaryotes, however, cannot be excluded. Since a variety of intracellular bacteria and mycoplasma-like structures are seen in *Giardia* cysts and trophozoites [32], the *Giardia* AK might have been transferred from prokaryotic symbiont other than mitochondria.

*Acknowledgements:* We thank Dr. G. Xu for providing human AK3 sequence, Dr. H. Uchimiya for rice *Adk*-a and *Adk*-b sequence, Dr. M. Müller for *Trichomonas* AK sequence and Dr. T. Okajima for pig UMP-CMP kianse sequence, all of which were unpublished at the time received them. We also thank M. Ohara for critical comments on the manuscript. A part of this work was supported by Grants-in-Aid from the Ministry of Education, Science and Culture, Japan.

## References

[1] Schulz, G.E. (1987) Cold Spring Harb. Symp. Quant. Biol. 52, 429–439.

[2] Nakazawa, A., Yamada, M., Tanaka, H., Shahjahan, M. and Tanabe, T. (1990) in: Isozymes: Structure, Function and Use in Biology and Medicine, pp. 495–514, Wiley-Liss, New York.

[3] Diederichs, K. and Schulz, G.E. (1990) Biochemistry 29, 8138–8144.

[4] Diederichs, K. and Schulz, G.E. (1991) J. Mol. Biol. 217, 541–549.

[5] Dreusicke, D. and Schulz, G.E. (1986) FEBS Lett. 208, 301–304.

[6] Dreusicke, D., Karplus, P.A. and Schulz, G.E. (1988) J. Mol. Biol. 199, 359–371.

[7] Schulz, G.E., Müller, C.W. and Diederichs, K. (1990) J. Mol. Biol. 213, 627–630.

[8] Müller-Dieckmann, H.-J. and Schulz, G.E. (1994) J. Mol. Biol. 236, 361–367.

[9] Vonrhein, C., Schlauderer, G.J. and Schulz, G.E. (1995) Structure 3, 483–490.

[10] Müller, C.W. and Schulz, G.E. (1988) J. Mol. Biol. 202, 909–912.

[11] Müller, C.W. and Schulz, G.E. (1992) J. Mol. Biol. 224, 159–177.

[12] Schiltz, E., Burger, S., Grafmüller, R., Deppert, W.R., Haehnel, W. and Wagner, E. (1994) Eur. J. Biochem. 222, 949–954.

[13] Magdolen, V., Oechsner, U. and Bandlow, W. (1987) Curr. Genet. 12, 405–411.

[14] Rozario, C. and Müller, M. (1995) Mol. Biochem. Parasitol. 71, 279–283.

[15] Ohama, T., Muto, A. and Osawa, S. (1989) J. Mol. Evol. 29, 381–395.

[16] Wiesmüller, L., Noegel, A.A., Bârzu, O., Gerisch, G. and Schleicher, M. (1990) J. Biol. Chem. 265, 6339–6345.

[17] Schricker, R., Magdolen, V., Kaniak, A., Wolf, K. and Bandlow, W. (1992) Gene 122, 111–118.

[18] Stehle, T. and Schulz, G.E. (1992) J. Mol. Biol. 224, 1127–1141.

[19] Fukami-Kobayashi, K. (1994) Mol. Biol. Evol. 11, 99–105.

[20] Jones, D.T., Taylor, W.R. and Thornton, J.M. (1992) Comp. Appl. Biosci. 8, 275–282.

[21] Saitou, N. and Nei, M. (1987) Mol. Biol. Evol. 4, 406–425.

[22] Felsenstein, J. (1985) Evolution 39, 783–791.

[23] Nakamura, K., Nakamura, A., Takamatsu, H., Yoshikawa, H. and Yamane, K. (1990) J. Biochem. Tokyo 107, 603–607.

[24] Koivula, T. and Hemilä, H. (1991) J. Gen. Microbiol. 137, 2595–2600.

[25] Rose, T., Brune, M., Wittinghofer, A., Le Blay, K., Surewicz, W.K., Mantsch, H.H., Bârzu, O. and Gilles, A.M. (1991) J. Biol. Chem. 266, 10781–10786.

[26] Bandlow, W., Strobel, G., Zoglowek, C., Oechsner, U. and Magdolen, V. (1988) Eur. J. Biochem. 178, 451–457.

[27] Konrad, M. (1993) J. Biol. Chem. 268, 11326–11334.

[28] Kawai, M. and Uchimiya, H. (1995) Plant Mol. Biol. 27, 943–951.

[29] Magdolen, V., Schricker, R., Strobel, G., Germaier, H. and Bandlow, W. (1992) FEBS Lett. 299, 267–272.

[30] Okajima, T., Tanizawa, K. and Fukui, T. (1993) FEBS Lett. 334, 86–88.

[31] Länge, S., Rozario, C. and Müller, M. (1994) Mol. Biochem. Parasitol. 66, 297–308.

[32] Adam, R.D. (1991) Microbiol. Rev. 55, 706–732.

[33] Matsuura, S., Igarashi, M., Tanizawa, Y., Yamada, M., Kishi, F., Kajii, T., Fujii, H., Miwa, S., Sakurai, M. and Nakazawa, A. (1989) J. Biol. Chem. 264, 10148–10155.

[34] Heil, A., Müller, G., Noda, L., Pinder, T., Schirmer, H., Schirmer, I. and Von Zobern, I. (1974) Eur. J. Biochem. 43, 131–144.

[35] Kuby, S.A., Palmieri, R.H., Frischat, A., Fischer, A.H., Wu, L.H., Maland, L. and Manship, M. (1984) Biochemistry 23, 2393–2399.

[36] Suminami, Y., Kishi, F., Torigoe, T. and Nakazawa, A. (1988) J. Biochem. Tokyo 103, 611–617.

[37] Reuner, C., Hable, M., Wilmanns, M., Kiefer, E., Schiltz, E. and Schulz, G.E. (1988) Protein Seq. Data. Anal. 1, 335–343.

[38] Cao, M., Akridge, R., Weston, D., Kemp, W.M. and Doughty, B.L. (1992) Exp. Parasitol. 74, 357–359.

[39] Okajima, T., Goto, S., Tanizawa, K., Tagaya, M., Fukui, T., Shimofuruya, H. and Suzuki, J. (1995) J. Biochem. Tokyo 117, 980–986.

[40] Jiang, Z.-R., Abaigar, L.T., Huang, S.-H., Cai, B. and Jong, A.Y. (1991) J. Biol. Chem. 266, 18287–18293.

[41] Kishi, F., Tanizawa, Y. and Nakazawa, A. (1987) J. Biol. Chem. 262, 11785–11789.

[42] Tanabe, T., Yamada, M., Noma, T., Kajii, T. and Nakazawa, A. (1993) J. Biochem. Tokyo 113, 200–207.

[43] Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fraser, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M., Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laisster, N., Latreille, P., Lightning, J., Lloyd, C., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A., Saunders, D., Shownkeen, R., Sims, M., Smaldon, N., Smith, A., Smith, M., Sonnhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Weinstock, L., Wilkinson-Sproat, J. and Wohldman, P. (1994) Nature 368, 32–38.

[44] Kawai, M., Kidou, S., Kato, A. and Uchimiya, H. (1992) Plant J. 2, 845–854.

[45] Xu, G., O'Connell, P., Stevens, J. and White, R. (1992) Genomics 13, 537–542.

[46] Yamada, M., Shahjahan, M., Tanabe, T., Kishi, F. and Nakazawa, A. (1989) J. Biol. Chem. 264, 19192–19199.
[47] Schricker, R., Magdolen, V. and Bandlow, W. (1992) Mol. Gen. Genet. 233, 363–371.
[48] Gilles, A.-M., Sismeiro, O., Munier, H., Fabian, H., Mantsch, H.H., Surewicz, W.K., Craescu, C.C., Bârzu, O. and Danchin, A. (1993) Eur. J. Biochem. 218, 921–927.
[49] Maskell, D.J., Szabo, M.J., Butler, P.D., WIlliams, A.E. and Moxon, E.R. (1991) Mol. Microbiol. 5, 1013–1022.
[50] Gutierrez, J.A. and Csonka, L.N. (1995) J. Bacteriol. 177, 390–400.
[51] Brune, M., Schumann, R. and Wittinghofer, F. (1985) Nucleic Acids Res. 13, 7139–7151.
[52] Glaser, P., Presecan, E., Delepierre, M., Surewicz, W.K., Mantsch, H.H., Bârzu, O. and Gilles, A.-M. (1992) Biochemistry 31, 3038–3043.
[53] Spürgin, P., Tomasselli, A.G. and Schiltz, E. (1989) Eur. J. Biochem. 179, 621–628.
[54] Ohkubo, S., Muta, A., Kawachi, Y., Yamao, F. and Osawa, S. (1987) Mol. Gen. Genet. 210, 314–322.
[55] Zschocke, P.D., Schiltz, E. and Schultz, G.E. (1993) Eur. J. Biochem. 213, 263–269.
[56] Konrad, M. (1992) J. Biol. Chem. 267, 25652–25655.
[57] Gentry, D., Bengra, C., Ikehara, K. and Cashel, M. (1993) J. Biol. Chem. 268, 14316–14321.