Hypothesis

Animal deoxyribonucleoside kinases: 'forward' and 'retrograde' evolution of their substrate specificity[1]

Jure Piškur[a,*], Michael P.B. Sandrini[a], Wolfgang Knecht[a,2], Birgitte Munch-Petersen[b]

[a] *BioCentrum-DTU, Eukaryote Molecular Biology Group, Building 301, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark*
[b] *Department of Life Sciences and Chemistry, Roskilde University, Roskilde, Denmark*

Received 18 November 2003; revised 19 January 2004; accepted 19 January 2004

First published online 2 February 2004

Edited by Takashi Gojobori

Abstract Deoxyribonucleoside kinases, which catalyse the phosphorylation of deoxyribonucleosides, are present in several copies in most multicellular organisms and therefore represent an excellent model to study gene duplication and specialisation of the duplicated copies through partitioning of substrate specificity. Recent studies suggest that in the animal lineage one of the progenitor kinases, the so-called dCK/dGK/TK2-like gene, was duplicated prior to separation of the insect and mammalian lineages. Thereafter, insects lost all but one kinase, dNK (EC 2.7.1.145), which subsequently, through remodelling of a limited number of amino acid residues, gained a broad substrate specificity.

Key words: Deoxyribonucleoside kinase; Nucleic acid precursor; Evolution; Gene duplication; Substrate specificity

1. Introduction

Gene and genome duplications play a major role in biological evolution [1,2]. Genes are continuously duplicated but often one of the copies is subsequently removed from the genome. However, when a set of parallel genes is created, gene copies can be functionally specialised through accumulation of mutations and permanent preservation within the genome. Preservation can be achieved either by origination of a new function by one of the duplicated genes, or by partitioning of the ancestral function(s) between the two duplicated genes [3,4]. Deoxyribonucleoside kinases, usually present in several copies, represent an excellent model to study evolution through gene duplication, preservation and specialisation of the duplicated copies through partitioning of the substrate specificity.

Deoxyribonucleoside kinases catalyse the phosphorylation of deoxyribonucleosides (dN) to the corresponding deoxyribonucleoside monophosphates (dNMP). They are the key enzymes in the salvage of deoxyribonucleosides originating extracellularly from food/medium or from intracellular breakdown of DNA [5]. Subsequently, dNMPs are phosphorylated into diphosphates and triphosphates, which are the direct precursors of DNA. The physiological importance of the deoxyribonucleoside kinases in humans has recently been emphasised by the finding that mutated or deficient mitochondrial kinases result in the mitochondrial DNA depletion syndrome. The affected individuals suffer from devastating myopathy or hepatocerebral failures and have a life span of less than 4 years [6,7]. On the other hand, the salvage pathway is of the utmost importance in treatment of several viral and cancer diseases, through phosphorylation of anti-viral and anti-cancer pro-drugs, such as aciclovir, zidovudine and gemcitabine [5].

2. Insect multisubstrate enzyme

Deoxyribonucleoside kinases exhibit a high degree of diversity among the organisms that have so far been analysed. Detailed characterisation and comparison of the modern enzymes from different organisms can help to reconstruct the likely structures of the ancient kinase(s) and their evolution into the present modern forms. Mammals have four deoxyribonucleoside kinases with overlapping specificities: cytoplasmic thymidine kinase 1 (TK1) phosphorylates only thymidine (dT), mitochondrial thymidine kinase 2 (TK2) phosphorylates deoxycytidine (dC) and dT, cytoplasmic deoxycytidine kinase (dCK) phosphorylates deoxyadenosine (dA), deoxyguanosine (dG) and dC, and mitochondrial deoxyguanosine kinase (dGK) phosphorylates only purine dNs, dA and dG [5]. In the fruit fly, only a single deoxyribonucleoside kinase exists (DmdNK, EC 2.7.1.145), capable of phosphorylating all four natural substrates. This kinase was found to be the fastest known deoxyribonucleoside kinase enzyme [8]. Recently, also silk moth and mosquito have been shown to contain a multisubstrate kinase, [9,10], suggesting that insects have only one deoxyribonucleoside kinase, dNK. This kinase was initially assumed to represent a direct descendant of the original enzyme, a dCK/dGK/TK2-like kinase, a living fossil with a broad substrate specificity [8]. This would position the gene duplication(s) and gene product specialisation after the separation of the insect and mammalian lineages. When the sequences of vertebrate, insect and bacterial deoxyribonucleoside kinases are aligned, the results point out that all modern kinase genes indeed originated from a common progenitor kinase by a series of gene duplications. However, phylogenetic analysis also showed that the insect dNKs group together

*Corresponding author. Fax: (45)-45-932809.
E-mail address: jp@biocentrum.dtu.dk (J. Piškur).

[1] Dedicated to Professor Morten Kielland-Brandt on the occassion of his 60th birthday.
[2] Present address: AstraZeneca R&D, Mölndal, Sweden.

0014-5793/04/$30.00
doi:10.1016/S0014-5793(04)00081-X

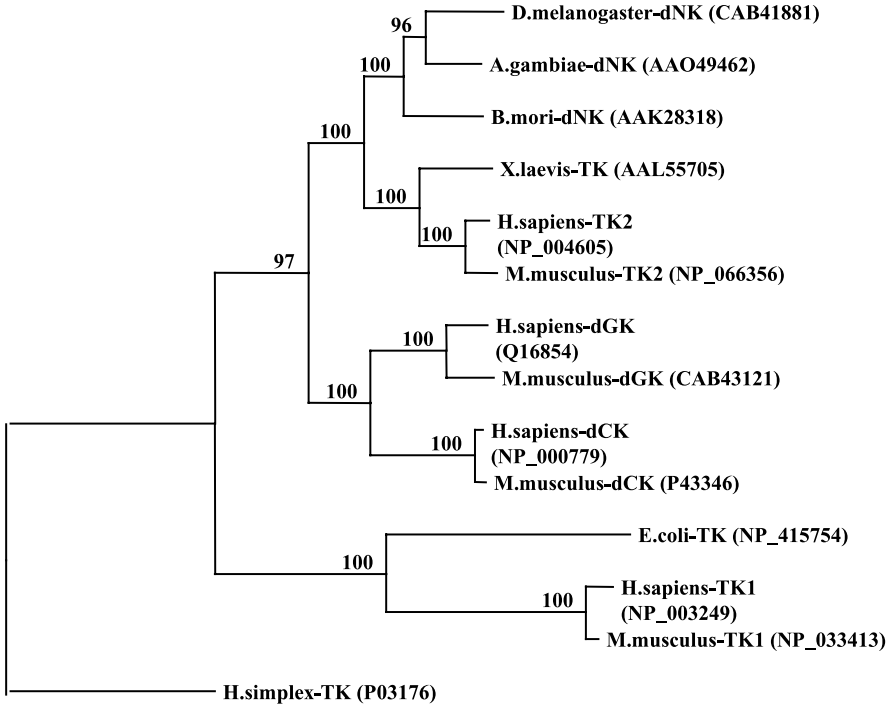

Fig. 1. Phylogenetic relationship among animal deoxyribonucleoside kinases (amino acid sequence accession numbers are in brackets). Insect broadly specific dNK kinases group together with other animal TK2 kinases, which can phosphorylate only pyrimidine substrates. Apparently, also dNK, TK2, dCK and dGK kinases share the common progenitor, the so-called dCK/dGK/TK2-like kinase. The phylogenetic analysis was done using ClustalX 1.81 with standard settings and TREECON 1.3b by the neighbour joining (NJ) method with Poisson correction for distance calculation (details of the phylogenetic analysis can be found in [2]). Human Herpes simplex virus type 1 thymidine kinase is used as outgroup. The numbers given are frequencies (%) at which a given branch appeared in 100 bootstrap replications.

with TK2-like enzymes (Fig. 1), which are highly specialised for pyrimidine substrates [9], and not at the base of the modern dCK, dGK and TK2 enzymes. Therefore, insect dNKs do not necessarily represent living fossils: versions of the original non-specialised enzyme.

3. A limited number of crucial amino acid residues

The DmdNK enzyme has recently also been intensively studied by means of structural and molecular biology methods. DmdNK mutants generated by random hypermutagenesis were isolated based on their ability to increase the sensitivity of a transformed thymidine-deficient *Escherichia coli* strain to various (deoxy)ribonucleoside analogues. When these random DmdNK mutants were analysed and their sequences compared to the 3D structures of different kinases, it was possible to identify the key amino acid residues, which determine the substrate specificity [11,12]. These data are now also useful for reconstruction of the origin of the substrate specificity of the animal kinases. One mutant, MuC, carrying the V84A mutation, increased the sensitivity of *E. coli* only towards dC analogues, like ddC and araC [11]. These results overlapped with the observation that the DmdNK valine at position 84 is present in all TK2-like kinases, but otherwise all eukaryotic dCKs have alanine at this position (Fig. 2). In addition, in DmdNK V84 is one of the substrate binding pocket residues [12]. Thus, this residue may be significant in determining substrate specificity. When the DmdNK and dGK 3D structures were compared, it was apparent that the predominantly pyrimidine-specific DmdNK has a large substrate cleft, while purine-specific dGK has a tighter pocket. DmdNK has an extended cleft at the 5-position of the substrate pyrimidine ring, surrounded by V84, M88 and A110 [12]. These residues are conserved in all TK2-like enzymes but not in dCK- and dGK-like enzymes. The corresponding cavity could therefore be necessary to accommodate the methyl group of thymidine.

The broadly specific enzyme DmdNK has also been modified into different mutant forms with limited substrate specificity. The above-mentioned three amino acid residues, V84/M88/A110, were changed into the corresponding residues found among dCK- and dGK-like enzymes ([13], Fig. 2). As expected, the site-directed mutagenesis of these sites provided new mutant enzymes with changed substrate specificity. For example, the triple DmdNK mutant V84S/M88R/A110D, mimicking the substrate binding site of dGK, lost the pyrimidine specificity but gained the dGK substrate specificity, phosphorylating exclusively purine substrates. These results demonstrated that only one to three changes were necessary to convert the substrate specificities from predominantly dT into dA, dC or dG as the most preferred substrate. However, the modern mammalian TK2 enzymes share the DmdNK V84/M88/A110 pattern (Fig. 2) and exhibit substrate preference only for pyrimidine substrates. Therefore, also amino acid residues other than V84/M88/A110 are involved in fine-tuning the substrate specificity.

Recently, the three-dimensional crystal structure of human dCK was also solved [14]. The last two of the three residues corresponding to the above-mentioned positions in DmdNK, V84, M88 and A110, were mutated back into the DmdNK-specific residues thus reversing the mutations presented above and creating the double dCK mutant: R104M, D133A [14].

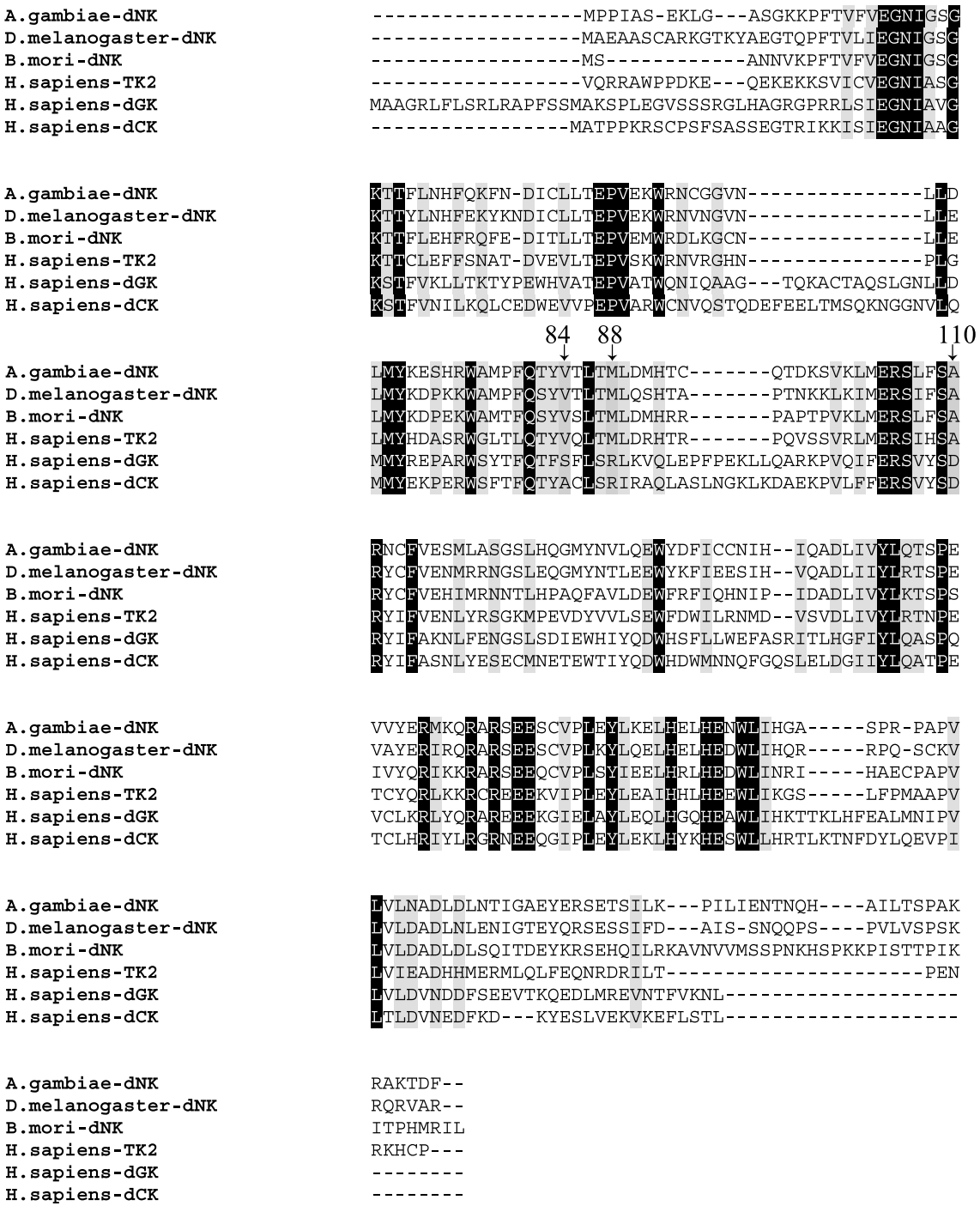

Fig. 2. Alignment of human and insect deoxyribonucleoside kinases. *Bombyx mori* dNK (AAK28318), *Drosophila melanogaster* dNK (CAB41881), *Homo sapiens* TK2 (NP_004605) with a truncated N-terminus, *H. sapiens* dCK (NP_000779), *H. sapiens* dGK (Q16854), *Anopheles gambiae* dNK (AAO49462). The conserved residues are shadowed and the positions employed in the structure–function relationship studies [13,14], DmdNK V84/M88/A110, are highlighted.

Interestingly, this mutant gained the ability to phosphorylate dT. Furthermore, the triple dCK mutant A100V/R104M/D133A, which was created to imitate the DmdNK cavity, was found to have a 30-fold increased k_{cat} for dC phosphorylation, generating a faster enzyme than the wild type dCK [14]. A similar scenario, based on a limited number of single amino acid changes at different positions, could have operated during the evolutionary history of animal deoxyribonucleoside kinases.

4. Retrograde evolution

We assume that the earliest deoxyribonucleoside kinase had a broad substrate specificity and that it has during evolution been duplicated several times and subsequently the resulting copies specialised. While many bacteria have TK1-like enzymes, a phylogenetic equivalent of animal dCK, dGK or TK2 enzymes has so far not been found outside eukaryotes. Thus, one of the first eukaryotes was likely to have a TK1-like

kinase and a dCK/dGK/TK2-like kinase, the latter being the progenitor of the modern mammalian dCK, dGK and TK2 kinases. The situation with only two types of deoxyribonucleoside kinases, TK1 and a dCK/dGK/TK2-like one, is preserved in plants [15]. This 'living fossil' situation provides direct evidence that the common progenitor of plants and animals had a TK1-like kinase and at least one additional kinase from the dCK/dGK/TK2 family. Consequently, the common progenitor of insects and mammals also had at least two kinases.

Upon further rounds of duplication of the progenitor dCK/dGK/TK2-like kinase the resulting copies each specialised for a limited number of substrates. As illustrated above, only a limited number of point mutations has had to accumulate to achieve the final substrate specialisation. The recent results on DmdNK and human dCK, and especially the presence of at least two kinases in less related eukaryotes, suggest that DmdNK does not represent the original, pre-duplication stage, enzyme; rather this enzyme has undergone retrograde evolution from a specialised TK2-like enzyme converting predominantly dT into an enzyme with a broad substrate specificity. The reason could be that the insect progenitor initially had several kinases but at some point lost all but the TK2-like kinase. This enzyme was subsequently under evolutionary pressure to broaden its substrate specificity. Again, only a limited number of point mutations has been necessary to remodel the specialised enzyme into a modern multisubstrate deoxyribonucleoside kinase. While the kinase gene duplications and the initial specialisations are likely to have taken place before the separation of the insect and mammalian lineages, the modern kinases have still preserved the plasticity to easily change their substrate specificity.

Acknowledgements: This work was partially supported by the Danish Research Council and Danish Cancer Society. The authors thank J. Rawls for several useful suggestions.

References

[1] Ohno, S. (1970) Evolution by Gene Duplication, George Allen and Unwin, London.
[2] Langkjær, R.B., Cliften, P.F., Johnston, M. and Piskur, J. (2003) Nature 421, 484–852.
[3] Sidow, A. (1996) Curr. Opin. Genet. Dev. 6, 715–722.
[4] Force, A., Lynch, M., Pyckett, F.B., Amores, A., Yan, Y. and Postlethwait, J. (1999) Genetics 151, 1531–1545.
[5] Arner, E.S.J. and Eriksson, S. (1995) Pharmacol. Ther. 67, 155–186.
[6] Mandel, H., Szargel, R., Labay, V., Elpeleg, O., Saada, A., Shalata, A., Anbinder, Y., Berkowitz, D., Hartman, C., Barak, M., Eriksson, S. and Cohen, N. (2001) Nat. Genet. 29, 337–341.
[7] Saada, A., Shaag, A., Mandel, H., Nevo, Y., Eriksson, S. and Elpeleg, O. (2001) Nat. Genet. 29, 342–344.
[8] Munch-Petersen, B., Piskur, J. and Søndergaard, L. (1998) J. Biol. Chem. 273, 3926–3931.
[9] Knecht, W., Petersen, G.E., Munch-Petersen, B. and Piskur, J. (2002) J. Mol. Biol. 315, 529–540.
[10] Knecht, W., Petersen, G.E., Sandrini, M.P.B., Søndergaard, L., Munch-Petersen, B. and Piskur, J. (2003) Nucleic Acids Res. 31, 1665–1672.
[11] Knecht, W., Munch-Petersen, B. and Piskur, J. (2000) J. Mol. Biol. 301, 827–837.
[12] Johansson, K., Ramaswamy, S., Ljungcrantz, C., Knecht, W., Piskur, J., Munch-Petersen, B., Eriksson, S. and Eklund, H. (2001) Nature Struct. Biol. 8, 616–620.
[13] Knecht, W., Sandrini, M.P.B., Johansson, K., Eklund, H., Munch-Petersen, B. and Piskur, J. (2002) EMBO J. 21, 1873–1880.
[14] Sabini, E., Ort, S., Nerjahn, C., Konrad, M. and Lavie, A. (2003) Nature Struct. Biol. 10, 513–519.
[15] The *Arabidopsis* Genome Initiative (2000) Nature 408, 796–815.