Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 14385–14390, December 1996
Biochemistry

Deoxycytidine kinase and deoxyguanosine kinase of *Lactobacillus acidophilus* R-26 are colinear products of a single gene

NING MA*, SEIICHIRO IKEDA, SHENYUAN GUO, ANGELA FIENO†, INSHIK PARK‡, STEPHEN GRIMME, TAKEHISA IKEDA, AND DAVID H. IVES

Department of Biochemistry, The Ohio State University, Columbus, OH 43210-1292

Communicated by Gertrude B. Elion, GlaxoWellcome, Research Triangle Park, NC, September 26, 1996 (received for review June 7, 1996)

ABSTRACT **Three of the four deoxynucleoside kinases required for growth of *Lactobacillus acidophilus* R-26 exist as heterodimeric pairs specific for deoxyadenosine (dAK) and deoxycytidine (dCK) or dAK and deoxyguanosine (dGK). However, only two tandem genes, *dak*/*dgk*, are found, and are expressed only as dAK/dGK in transformed *Escherichia coli*. Sequencing peptides spanning 63% of the native dCK subunit revealed a sequence identical to that deduced from *dgk* (beginning MTVIVL···), except that dCK lacks residues 2 and 3 (dCK is M··IVL; dGK is ·TVIVL). Also, mass spectrometry indicates that native dCK and dGK subunits are identical in mass adjusted for the first three residues. Furthermore, the native enzymes have identical isoelectric pH values, indicating an equal number of charged residues. To enable *E. coli* to express peptide having the native dCK sequence, codons 2 and 3 were deleted from the *dgk* portion of the tandem genes, resulting in expression of protein having the specificities and regulatory properties of native dAK/dCK, including heterotropic stimulation of dAK activity by deoxycytidine or dCTP (not deoxyguanosine or dGTP) and end-product inhibition of the respective activities by dATP and dCTP. Subcloning normal and mutant *dgk* yielded homodimeric dGK and dCK, respectively. The dCK homodimer strongly resembles human dCK, with a low K_m for deoxycytidine, the ability to phosphorylate deoxyadenosine and deoxyguanosine at much higher K_m values, and end-product inhibition by dCTP. Thus two distinct and specific enzymes evidently are derived from a single *Lactobacillus* gene. The mechanism by which this occurs *in vivo* has yet to be elucidated.**

The deoxynucleoside kinases provide a salvage pathway to DNA distinct from *de novo* synthesis of the triphosphate precursors. Because they provide a portal to intracellular incorporation into DNA for a variety of important chemotherapeutic agents, they have been the subject of intensive study in many laboratories (for review see ref. 1). In the cytosol of most growing mammalian cells are thymidine kinase (TK) and deoxycytidine kinase (dCK), the latter being capable of phosphorylating deoxyadenosine (dAdo) and deoxyguanosine (dGuo) as well. Despite having been cloned from human and murine cDNAs (refs. 2 and 3; GenBank accession no. U01881), establishing the specificities and structure–function relationships of human dCK has been problematic—due in some measure, we think, to persistent proteolytic modifications (4). Accordingly, a stable prokaryotic model for dCK would be very useful.

While most bacteria have TK, only two genera have so far been shown clearly to express all four deoxynucleoside kinase activities, namely *Lactobacillus* (5) and *Bacillus* (6). In the case of *Lactobacillus acidophilus* R26, a strain from which ribonucleotide reductase is lacking, all four of these salvage enzymes are essential for DNA synthesis. While TK is readily separable, the other three deoxynucleoside kinase activities of *Lactobacillus* are located on separate subunits of two physically similar heterodimeric proteins, dAdo kinase (dAK)/dCK and dAK/dGuo kinase (dGK; ref. 7). This organization of activities contrasts with the *Bacillus subtilis* enzyme, in which dAdo and deoxycytidine (dCyd) compete for a common site, whereas dGK is genetically distinct (6). In the *Lactobacillus* system, each of these substrates is phosphorylated at an independent site, all having comparable K_m values and strong feedback inhibition by their homologous deoxynucleoside triphosphates. In another sense, however, the subunits interact, inasmuch as saturation of either the dCK or dGK subunit by its respective deoxynucleoside or dNTP end product produces a several-fold heterotropic stimulation of turnover at the associated dAK subunit.

Tandem *Lactobacillus* genes for dAK and dGK (*dak*/*dgk*) have been cloned and expressed in *Escherichia coli*, using their common upstream promoter and independent Shine/Dalgarno sequences (8). The DNA sequences encoding these purine-specific subunits have a 65% identity, so a larger difference might be expected in the gene for the pyrimidine-specific dCK. However, no distinct gene for dCK could be found with a variety of probes, but the search was prolonged out of respect for the one-gene/one-polypeptide paradigm. Eventually, the accumulating circumstantial evidence pointed overwhelmingly to dCK arising from the same gene as dGK. We have now been able to show by means of site-directed mutagenesis that both separable enzymes can be expressed from one gene and that the first three N-terminal residues uniquely contribute to their different specificities.

EXPERIMENTAL PROCEDURES

Materials. Radioactive nucleotides [α-^{32}P]dATP, [γ-^{32}P]ATP, [γ-^{33}P]ATP, and [α-thio-^{35}S]dATP were from Amersham or ICN, and radioactive nucleosides [^{3}H]dAdo, [^{3}H]dGuo, and [^{3}H]dCyd were from Moravek Biochemicals (Brea, CA). ATP was from Boehringer Mannheim and other enzyme assay components were from Sigma. The Photogene Nucleic Acid Detection kit and biotin-7-dATP were purchased from Bethesda Research Laboratories (BRL). DNA sequencing kits were from BRL, Epicentre Technologies (Madison, WI), and United States Biochemical. Kits were purchased from the following: mutagenesis kits (Mutagene II) from Bio-Rad, the reverse transcription (RT) kit from BRL, and the TA Cloning kit from Stratagene. *Taq* polymerase for PCR or RT-PCR was from BRL or United States Biochemical. Restriction endonucleases, T4 ligase, and polynucleotide kinase were from Boehringer Mannheim or BRL, and photogene membranes for hybridization were also from BRL. *L.*

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

Abbreviations: TK, thymidine kinase; dCyd, deoxycytidine; dAdo, deoxyadenosine; dGuo, deoxyguanosine; dCK, dCyd kinase; dAK, dAdo kinase; dGK, dGuo kinase; RT, reverse transcription; MALDI, matrix-assisted laser-desorption/ionization.

*Present address: Genetics and Biochemistry Branch, National Institutes of Health, Bethesda, MD 20892.
†Present address: Proctor and Gamble, Cincinnati, OH 45253-8707.
‡Present address: Dong-A University, Pusan 604-714, Korea.

acidophilus R-26 (ATCC 11506; ssp. *johnsonii*) stock was from American Type Culture Collection. *Lactobacillus* culture media, *Lactobacillus* broth, and MRS medium were from Difco. Plasmids pBluescript KS(+) and pUC-19 and *E. coli* XL1-Blue host cells were from Stratagene. All the primers used for sequencing, PCR, and mutagenesis were synthesized at the Biotechnology Center at the Ohio State University. Qiagen columns were purchased from Qiagen (Chatsworth, CA).

Reagents for gel electrophoresis were from Bio-Rad Laboratories and Sigma. Sephacryl S-200 HR was from Pharmacia. PCR was performed on a Perkin–Elmer/Cetus DNA Thermal Cycler or MJ Research PTC-100 Programmable Thermal Controller. Electroporation was carried out on the Gene Pulse from Bio-Rad.

Enzyme Purification and Assays. *E. coli* cells were cultured in *Lactobacilli* MRS Broth (Difco) containing ampicillin, and *L. acidophilus* cells were grown in either *Lactobacilli* MRS Broth (Difco) or modified Hoff–Jorgensen medium (5). The cells were broken with a mini-Bead Beater (Biospec Products, Bartlesville, OK) containing 0.2-mm glass beads for small-scale preparations or by sonication for larger-scale preparations, followed by centrifugation. The extraction buffer was 100 mM Tris (pH 8.0) containing 3 mM EDTA. Extracts were purified by ammonium sulfate precipitation (67% saturation), followed by gel permeation chromatography, yielding products judged by electrophoresis to be at least 80% pure. The native and engineered dAK/dCK and native dAK/dGK were further purified by dATP- and/or dCTP-Sepharose affinity chromatography to homogeneity as described for the purification of native *Lactobacillus* dAK/dCK and dAK/dGK (9, 10). Routine deoxynucleoside kinase assays were as described (11). Substrate concentrations were varied for steady-state kinetics assays, and specific activity is defined as 1 nmol of product per min per mg of protein.

A new spectrophotometric assay procedure better suited for assays requiring millimolar substrate concentrations has been developed; it is particularly useful with alternative substrates having large K_m values. ATP consumed by the deoxynucleoside kinase reaction is regenerated by phosphoenolpyruvate and pyruvate kinase; the resulting pyruvate is converted to lactate by lactic dehydrogenase while monitoring the consumption of NADH at 340 nm. The reaction mixture contains (final concentrations) 0.1 M Tris·HCl (pH 8.0), 12 mM $MgCl_2$, 3 mM phosphoenolpyruvate (monocyclohexylammonium salt), 0.125 mM NADH, 0.02–2.0 mM deoxynucleoside, 5 mM ATP, 12 units of pyruvate kinase (rabbit muscle), and 9 units of D-lactic dehydrogenase (*Leukonostoc mesenteroides*). Because nearly all ATP preparations contain traces of ADP, the complete system (minus deoxynucleoside kinase) is preincubated for 5 min to phosphorylate any ADP, followed by addition of the appropriate deoxynucleoside kinase to start the assay reaction. (Based upon the absorbance change in a blank assay, the total amount of NADH added to the mixture must be adjusted to compensate for any loss during ADP regeneration, thereby ensuring an initial concentration of 0.125 mM NADH for the kinase assay.)

Subunit Separation and Peptide Sequencing. Subunit separation was accomplished by reversed-phase HPLC on a Waters 625 system equipped with a Vydac 214TP54 (C-4) column (Vydac, Hesperia, CA), employing a linear gradient [buffer A consisted of water containing 0.06% trifluoroacetic acid; buffer B was acetonitrile/water, 80:20 (vol/vol), containing 0.054% trifluoroacetic acid]. The two peaks on HPLC were identified by SDS/PAGE run at pH 6.6 (12), according to the assignment of subunits established previously (7). For internal sequences, the dCK subunit was digested with 2% trypsin (wt/wt) in the presence of 2 M urea at 37°C for 24 hr, followed by separation of peptides by reversed-phase HPLC on a Vydac 218TP54 (C-18) column. Sequencing was performed by automated gas-phase Edman degradation chemistry at the Ohio State University Biochemical Instrument Center. N-terminal sequences of intact subunits were obtained as described previously (7), except for the cloned dGK subunit, which was prepared by HPLC (above) and sequenced directly.

Mass Spectrometry. Measurement of the molecular mass of each enzyme subunit was performed at Procter & Gamble with a Vestec VT 2000 matrix-assisted laser desorption/ionization (MALDI) mass spectrometer and a Perkin–Elmer/Sciexx API-III LC/MS/MS triple quadrapole mass spectrometer equipped with nanospray.

Isoelectric Focusing. Horizontal gel slabs, 0.5 mm thick, containing 2% Ampholine 4–6 (Pharmacia) were placed on a Hoefer thermoelectric cooling cell. Enzyme samples were applied at positions remote from the final equilibrium pH and developed at constant power until a field strength of 2000 V was reached. Sample lanes were divided into 2-mm slices, which were immersed in radiochemical assay mixtures. The pH was determined in sections of the adjacent lane.

Site-Directed Mutagenesis. The materials and protocols of the Mutagene kit (version II) from Bio-Rad were used. After phosphorylation by polynucleotide kinase, the mutagenic oligonucleotide was annealed to the single stranded template, derived from pBluescript KS(+) containing the *Lactobacillus Kpn*I fragment, which includes the tandem *dak/dgk* genes with their homologous promoter (8). After polymerization and ligation, the closed circular double-stranded DNA was used to transform *E. coli* strains MV1190 or XL1-Blue. The mutations were confirmed by DNA sequencing, using the dideoxynucleotide termination method and materials and protocol from Sequenase Version II kits (United States Biochemical).

RT-PCR. Amplification and sequencing of mRNA transcripts were by means of reverse transcriptase and PCR (RT-PCR), followed by conventional dideoxy sequencing. The SuperScript kit from BRL was used for the reverse transcriptase procedure, with a few modifications. RT primers were isolated by denaturing polyacrylamide electrophoresis and purified to ensure homogeneity. Purified total *Lactobacillus* RNA was heated at 75°C for 10 min and quenched on ice immediately. Annealing, followed by the RT reaction with RT-primer, occurred at 42°C for 50 min. After stopping the RT reaction with TE buffer, standard PCR was carried out using primers selected to amplify the desired portions of the coding sequence; the 5′ end of one of the primers was labeled with ^{33}P. As controls against extraneous contaminants, PCR was also run with no template and with total RNA. The PCR bands were resolved by electrophoresis on denaturing DNA sequencing gels, subcloned, and sequenced.

RESULTS

Approaches to dCK Gene Cloning. In view of the many physical and kinetic similarities between the native dAK/dGK and dAK/dCK enzyme pairs, we expected that the latter would be encoded by homologous tandem genes, or, perhaps by a gene for dCK alone, with sharing of dAK subunits from a common source. Screening by PCR (8) or colony hybridization, with PCR-generated probes based on known N-terminal sequences (7), which are identical for residues 3–28 of dGK and dCK, repeatedly yielded dAK/dGK clones but none with dCK specificity. Therefore, in an effort to construct a dCK-specific probe, we sought an internal sequence unique to the putative dCK gene, *dck*. Native dAK/dCK, free of dAK/dGK, was resolved into its subunits by reversed-phase HPLC on a C-4 column. The dCK subunit was subjected to digestion by 2% trypsin (wt/wt) for 24 hr (at 37°C, with 2 M urea). Following rechromatography on a C-18 column, 10 peptide fractions were sequenced by Edman gas-phase methodology, yielding 143 clearly identified residues. To our amazement, each of these residues was identical to the corresponding residue deduced from the gene for dGK, except for the first three residues of one peptide (Fig. 1). Thus, with 63% of the dCK residues accounted for, in peptides well-distributed throughout the subunit, including fragments corresponding to the N- and

dCK M--IVLSGPIGAGKSSLTGILSKYLGTNPF
dgk MTVIVLSGPIGAGKSSLTGILSKYLGTNPFYESVDDNPVLPLFYENPKKY
dak MTVIVLSGPIGAGKSSLTSLLAEHLGTQAFYEGVDNNPILPLYYKDMAHY

dCK DDNNVLDRSIYEXA QMNAXIGRATPE
dgk AFLLQVYFLNTRFRSIKSALTDDNNVLDRSIYEDALFFQMNADIGRATPE
dak TFLLNTYLLNHRLAQINQAIRDHNSVSDRSIYEDALFFKMNVDSGIADPT

dCK EVDTYYELLHNMMSELDR NYEQ
dgk EVDTYYELLHNMMSELDRMPKKNPDLLVHIDVSYDTMLKRIQKRGRNYEQ
dak EFKIYDSLLENMMEQAPGNPSKKPDLLIYIHVSLDTMLHRIQKRGRKFEQ

dCK LSYDPTLEDYYKR YYKPXYAKYDYSPKMTIDGDKLDFMASEEDRQEV
dgk LSYDPTLEDYYKRLLRYYKPWYAKYDYSPKMTIDGDKLDFMASEEDRQEV
dak LSTDPSLKDYYARLLSYYEPWYEKYNASPKMMIDGDKYDFVANEDARRKV

dCK LNQI LKEMGKLEDDWKPNLVK
dgk LNQIVAKLKEMGKLEDDWKPNLVK
dak INAIDQKLIDIGNLNX

FIG. 1. Edman gas-phase sequencing of tryptic peptides from native dCK subunits reveals only sequences identical to dGK. Sequences obtained from tryptic peptides of the dCK subunit are shown in boldface type. They are aligned with the complete *deduced* amino acid sequences from the *dgk* and *dak* DNA sequences (8), which are shown in italics. (See second column of Table 1 for actual processed N-terminal sequences of dGK and dAK peptides.) Single residues that could not be assigned without ambiguity are denoted by X, but in each case one of the two possible assignments is the residue found in dGK.

C-terminal peptides of dGK, the only discernible difference between the two subunits is at their N termini, with TVIVL··· for dGK and MIVL··· for dCK. Of particular interest, we note that the sequences surrounding the DRS motif—which in the dGK subunit (13) and in viral TKs (14, 15) has been implicated as contributing to a deoxynucleoside site—are identical. In five instances, we could not identify a particular residue of a peptide unambiguously, owing to the small quantities available and to diminishing repetitive yields; these appear as an "X" on the map. Not all of the Arg- and Lys-containing linkages were fully hydrolyzed, so some of the sequences obtained were from larger peptides containing internal basic amino acids.

For practical reasons, we decided not to attempt recovery and sequencing of the remaining peptides, but opted instead to use the common N- and C-terminal amino acid sequences available to construct a *dgk(dck)* probe by means of PCR applied to genomic DNA. It was expected that the DNA sequence(s) of such probes should reveal any alternative amino acid residues and/or order attributable to dCK among the remaining unsequenced peptides. The probe, spanning almost the entire length of the *dgk* gene, was generated with 5′-ATGACAGTTATTGTATTGAGCG-3′ (the first 22 nt of *dgk* or *dak*) and 5′-AACTAAATTAGGTTTCCAGTC(G/A)TC(T/C)TC-3′. Thus, PCR can reasonably be expected to have amplified both *dgk* and the putative *dck*, presumably producing two very similar products having occasional variations in sequence and/or codon selection. However, when the smaller of the two resulting PCR products (the larger being *dak*/*dgk*) was isolated and sequenced, we found that both its size and DNA sequence were those of *dgk* alone; no doubling of certain base pairs was observed, contrary to what would be expected from two slightly divergent genes. The putative *dck* sequence should have been copied by PCR, except in the unlikely event of too many mismatches occurring in the primer regions due to alternative codon usage. PCR was also employed to construct a hybridization probe from the *dak* gene from base pairs 345–699 (a region with only 40% identity with *dgk*), in case a putative *dck* gene occurs in tandem with a second homologous *dak* gene.

Using the *dgk(dck)*- and *dak*-specific probes, extensive Southern hybridization experiments were conducted with genomic DNA digested with a variety of restriction endonucleases singly and in combination (10 enzymes for *dak*, nine for *dgk*). No unexpected bands were observed; only DNA fragments expected from the tandem *dak*/*dgk* genes were detected in each instance.

Mass Spectrometric and Isoelectric Analysis of dAK, dGK, and dCK Subunits and Heterodimers. There remains the possibility that our various attempts to uncover the basis for the different specificities of the dCK and dGK subunits by comparing their linear structural or coding sequences may have overlooked a critical codon or amino acid substitution at a specificity site. Or, possibly, posttranslational modification(s) might have given rise to two activities from a common polypeptide. Two techniques with sufficient sensitivity to detect the substitution of a single amino acid residue are mass spectrometry and narrow-range isoelectric focusing. Mass spectrometry now offers a level of precision approaching ±1 mass unit under ideal conditions. We have employed both MALDI-MS and nanoelectrospray-MS for comparing the molecular masses of native and cloned dCK, dGK, and dAK subunits. The results, shown in Table 1, provide several important pieces of information. First, they clearly indicate masses for all three subunits that are within experimental error of those predicted from the *dgk* and *dak* DNA sequences alone. Such predictions were based upon the inferred amino acid sequences for dAK and dGK, adjusted in the first three residues according to the N-terminal sequences actually observed for each native or recombinant subunit (shown in the second column of Table 1). The results strongly support the original sequence assignments for dAK and dGK, which, except for deletions within

Table 1. Molecular mass measurements of deoxynucleoside kinase subunits by MALDI and nanoelectrospray mass spectrometry

Subunit	Sequenced N terminus	Calculated MH^{+}* (no. of residues)	MALDI MH^{+}	Δ	Nanospray $MH^{+} \pm \sigma$	Δ
Native dCK†	MIVLSG···	26,116‡ (222)	26,138	+22	26,122 ± 4.8	+6
Native dAK (I)†	MIVLSG···	24,499§ (213)	24,504	+5	24,500 ± 2.3	+1
Native dGK¶	TVIVLSG···	26,185 (223)	26,196	+11	ND	—
Native dAK (II)¶	MIVLSG···	24,499 (213)	24,499	0	24,501 ± 1.7	+2
Recombinant dGK‖	TVIVLSG···	26,185 (223)	26,194	+9	26,187 ± 2.8	+2
Recombinant dAK (II)‖	TVIVLSG···	24,568 (214)	24,561	−7	24,568 ± 3.8	0
Mutant dGK**	MIVLSG···	26,116 (222)	ND	—	26,125 ± 7.5	+9

ND, not determined.
*Assumes sequence deduced from gene is modified at N terminus as determined by Edman sequencing (see previous column).
†From dAK (I)/dCK heterodimer of *L. acidophilus* R-26.
‡Assumes sequence for dCK is same as for dGK, except at the N terminus (second column).
§Assumes sequence for dAK (I) is the same as dAK (II).
¶From dAK (II)/dGK.
‖From recombinant dAK/dGK.
**From mutant dGK/dGK homodimer.

the first three residues, appear to be unmodified by any other posttranslational modification or adducts.

But of greatest significance is the fact that mass comparisons are consistent with dCK and dGK being identical in amino acid content, except for the N-terminal residues noted, again strongly reinforcing the notion that both arise from a common gene in *Lactobacillus*. Also, for the same reason, the different specificities cannot be attributed to side-chain derivatization (e.g., glycosylation, acylation, etc.). Further, a deletion mutant of the recombinant dGK subunit, designed to be expressed with the same N-terminal sequence as native dCK, is found to be indistinguishable from native dCK polypeptide in mass. (Other properties of the mutant are presented below.) The fact that the mass determined for each peptide by nanospray-MS is several units higher than theoretical is believed to be attributable to the presence of some salt in the sample. The dGK subunit prepared by the same procedures was not readable by nanospray-MS, presumably for the same reason.

We also note that dAK(I) and dAK(II), the subunits associated in *Lactobacillus* heterodimers with dCK and dGK, respectively, are identical in size. Moreover, except for processing differences by *E. coli* at the three N-terminal residues (see second column of Table 1), they are equal in mass to the dAK from cloned dAK/dGK. It is likely that both dAK(I) and dAK(II) are products of a common *dak* gene and associate with either dGK or dCK subunits after translation and processing. Because the same N-terminal difference appears between native and recombinant dAK(II), we must assume that (*i*) *Lactobacillus* cells have a novel mechanism for editing mRNA, (*ii*) that they process the protein differently from *E. coli*, or (*iii*) a translation initiating codon other than AUG is used. The *dak* gene encodes MTVIVL, but is expressed as M··IVL in *Lactobacillus* and ·TVIVL in *E. coli*. It is understood that, in *E. coli*, a threonine in the second position results in the excision of the initiating methionine (16), while in the parental *Lactobacillus*, the second and third residues, Thr and Val, are excised somehow, while retaining the initial Met. It is very likely that dCK, which has an N terminus identical to dAK in *Lactobacilli,* is processed by the same unknown mechanism(s).

If the intact proteins, dAK/dGK and dAK/dCK, should differ by as little as one charged amino acid residue, that difference should produce a discernible shift in the isoelectric pH values for the two heterodimers. Therefore, both proteins were subjected to electrofocusing in polyacrylamide gels with ampholites producing a pH 4–6 gradient. The two proteins yielded identical pI values, 5.40 ± 0.03, which again is consistent with these proteins having identical amino acid contents and surface exposure, except for three uncharged residues at their N termini. Moreover, this experimental value closely approximates the calculated pI of 5.43, obtained by applying the DNASTAR program to the amino acids deduced from the DNA sequence.

Switching Specificity by Site-Directed Mutagenesis. Thus far, all the evidence available indicates that a single locus encodes both dGK and dCK. The native *Lactobacillus* peptides appear to differ only in the first three residues, suggesting that these residues might be the key to their specificities. Accordingly, to enable *E. coli* to express a peptide with a sequence identical to the *Lactobacillus* dCK, we constructed two deletion mutants lacking codons 2 and 3 (Thr and Val) of the *dgk* gene. In mutant I (*dak*/*dgk*Δ4-9), which was derived from the original dAK/dGK clone in pBluescript (8), the mutant *dgk* remains in tandem with unaltered *dak*, while in mutant II (*dgk*Δ4-9), the mutant *dgk* was subcloned by itself into pUC19. For each mutant, the translated peptide now has Ile in the penultimate position, preventing *E. coli* methionyl aminopeptidase from excising the terminal Met (16), and resulting, for mutant II at least, in an expressed peptide identical in mass to native dCK (Table 1).

As may be seen in Table 2, we compare the specificities and kinetic behavior of the mutant I gene product with unaltered recombinant dAK/dGK and native dAK/dCK. The mutant protein was quantitatively retained on a dCTP-Sepharose column and purified to homogeneity, exactly as with native dAK/dCK (9), whereas dAK/dGK heterodimer passes through that column (11). Most remarkably, the K_m(dCyd), maximum specific activity, and catalytic efficiency of the mutant I protein toward dCyd turnover are indistinguishable from those of pure native dAK/dCK, within experimental error. The activation of dAK by dCyd, rather than dGuo, also indicates a switch in nucleoside specificity of the mutated peptide. Moreover, end-product regulation of the engineered dCK is by dCTP, not dGTP, as in the unmutated protein (data not shown; see Table 3 for comparable results with mutant II). To simplify the table, results with native dAK/dGK are not shown either, but its K_m and V_{max} with dGuo (11 μM and 2150 nmol/min per mg) are closely reflected in the recombinant protein. Earlier estimates with less pure recombinant dAK/dGK had yielded a K_m for dGuo which was one order higher (13); we now believe the highly active purine nucleoside phosphorylase of *E. coli*, which is not present in the parental *Lactobacilli* (17), was competing for substrate, raising the apparent K_m.

By all of our criteria, the specificity of one subunit of mutant I has been switched from dGK to dCK. But newly discovered with both of these very pure proteins is the fact that dGK or dCK each possesses a secondary activity toward dCyd or dGuo, respectively, but with K_m values that are one to two orders of magnitude higher than for the primary substrates. However, it is very clear from the V_{max}/K_m efficiency ratios that dCyd is the preferred substrate of the mutant enzyme and dGuo is preferred by dGK. The dAK subunit, of course, has its own low-K_m

Table 2. Kinetic properties of native, recombinant, and mutant heterodimers

Kinase	Substrate	K_m(app),* μM	V_{max}, nmol/min per mg	V_{max}/K_m	dAK activation factor: dGuo, 50 μM	dAK activation factor: dCyd, 50 μM
Recombinant	dGuo	7.8	1660	210		
dAK/dGK	dAdo	8.4	120	14	5.1	1.1
	dCyd	1680	600	0.4		
Recombinant mutant	dGuo	250	1000	4.0		
dAK/dGK	dAdo	8.2	140	17	1.2	8.6
(dAK/dCK activity)	dCyd	4.5	2190	490		
Native dAK/dCK	dGuo	170	1070	6.3		
	dAdo	6.9	320	46	1.1	6.1
	dCyd	2.2	2540	1150		

*[MgATP] = 10 mM.

Table 3. Kinetic properties of homodimeric recombinant and mutant kinases

Kinase	Substrate	K_m(app),* μM	V_{max}, nmol/min per mg	V_{max}/K_m	Inhibition, % activity remaining† dGTP	dATP	dCTP
dGK	dGuo	5.8	1280	220	1.3	99	71
homodimer	dAdo	240	200	0.8	0.5	94	69
	dCyd	1800	1190	0.7	0.1	80	72
Mutated dGK	dGuo	110	290	2.6	52	81	4.8
homodimer	dAdo	1800	4200	2.3	58	78	3.0
(dCK activity)	dCyd	2.8	640	230	75	91	11

*[MgATP] = 10 mM.
†For inhibition experiment, [MgATP] = 2 mM, [dN] = 0.02 mM, and [dNTP] = 0.1 mM.

binding site (7), and it is interesting that its K_m is unchanged by this particular mutation on the opposing subunit and that its K_m and specific activity are about the same whether associated with dGK or dCK. (Since dAK activities were measured without activating dGuo or dCyd, the V_{max} values reflect only about one-fifth to one-ninth of its maximum potential.)

We might well ask whether the secondary activities of the two heterodimers are a function of the dGK(dCK) subunit, or of dAK. Subcloning the dGK- or dCK-specific genes permits us to examine the specificities of these subunits in the absence of dAK; they apparently are expressed as homodimers which are more weakly-associated than the heterodimers (S. Guo, unpublished work). Table 3 reveals again the dramatic reversal in substrate affinities produced upon mutagenic deletion from *dgk* of encoded residues 2 and 3. As with the mutant I heterodimer, the mutant homodimer (mutant II) has a K_m(app) for dCyd that is virtually identical to that of native dAK/dCK. In the control dGK homodimer, the K_m for dGuo is also essentially the same as in the native dAK/dGK. However, the specific activities of both homodimers, corrected for minor impurities, are only about one-third of the value with native heterodimers, suggesting that the dAK subunit also affects dCK or dGK turnover somewhat [but not substrate binding (18, 19)]. Both homodimeric enzymes exhibit the secondary, high-K_m substrate specificities seen with the heterodimers. In addition, both now can be seen to have secondary dAK activities as well. Each of the homodimeric enzymes is most strongly inhibited by its primary homologous triphosphate end product, controlling both the primary and secondary activities of a common active site. Thus the switch in nucleoside specificity generated by the mutation also switches the dNTP regulatory specificity, further evidence that the feedback inhibition is mediated through the nucleoside site as has been proposed (7, 20). It is gratifying that both polypeptides behave very much the same as their counterparts associated with dAK as heterodimers, which will allow for independent study of certain properties. [The same cannot be said for dAK, which is nearly inactive until associated with either dGK or dCK subunits (S. Guo, unpublished results).]

We have begun preliminary explorations seeking a possible mechanism by which two N-terminal products might be expressed from a single transcript. To determine if editing might be occurring at the level of mRNA, partial sequencing of the appropriate mRNAs was undertaken, using reverse transcriptase and PCR (RT-PCR). (To avoid possible contamination by existing dGK clones and mutants, this work was carried out in another laboratory, starting with new reagents and *Lactobacillus* cultures.) Copying of the total *Lactobacillus* RNA by reverse transcriptase, using dGK(dCK)-specific primers, was followed by amplification of the region surrounding the 5′ terminus of the transcripts by PCR. The product(s) would be 124 bp if no editing occurs, and 118 bp if the 6 bp corresponding to Thr and Val are excised. No DNA band corresponding to 118 bp was obtained. Three amplification products were obtained, ≈120-140 bp. These were each cloned, using a TA Cloning Kit (Stratagene), and sequenced. The largest band proved to be a portion of 16S ribosomal RNA, whereas the other two exhibited the 5′ terminus of genomic *dgk* only. (The smallest amplified DNA band appeared to have a spurious deletion of about half of the 3′ end of the region enclosed by the primers and is not believed to be significant.) But, from the intermediate band, it seems clear that the editing of codons 2 and 3 is not occurring in the *Lactobacillus* at the level of mRNA.

DISCUSSION

Various lines of evidence, taken together, indicate that only one set of tandem genes in the *Lactobacillus* genome encode dAK, dCK, and dGK and that dCK and dGK share a common coding sequence. "Back-to-back" primers hybridizing with both strands near the beginning of the *dak* gene should have enabled PCR to amplify the entire sequence between two similar *dak* genes lying within 10 kb or so of one another on the *Lactobacillus* genome. No such product could be obtained (unpublished results), suggesting that only one copy of *dak* is used. Further evidence for this is that the native (i.e., wild-type) dAK polypeptides associated with dCK and dGK apparently are identical, both in kinetic behavior (20) and in molecular mass (Table 1). The results of Southern hybridization with genomic DNA digested by a large variety of restriction endonucleases also were consistent with only one copy of the tandem *dak* and *dgk* genes, whether probed with *dak* or *dgk(dck)* DNA. There remains the question as to whether the 5′-primer used in PCR generation of the *dck* probe annealed effectively to the putative *dck* DNA; it is possible that even though the amino acid sequences of dCK and dGK are identical in this region, alternative codon usage conceivably might produce a slightly different DNA base composition. But, considering that the 5′ primer was degenerate at its critical 3′ terminus, and that a nonstringent annealing temperature (14° below melting point) was used, it seems very likely that PCR should have copied any existing *dck* sequence as well as that of *dgk*. We note also that dGK and dAK, which are otherwise quite different peptides, have identical DNA sequences for their amino termini, so it seems even more likely that the same codons would appear at the beginning of dCK's gene. Finally, the fact that the amino acid sequence of dCK is identical to dGK over the 63% of the molecule sequenced, from one end of the subunit polypeptide to the other, and that the two subunits have identical molecular masses when adjusted for the differing three N-terminal residues, makes it most unlikely that some rearrangement within the remaining unsequenced regions could account for differing specificities. Therefore, we conclude that dCK and dGK are products of a common *Lactobacillus* gene, but that *E. coli* appears to be unable to provide the alternative processing necessary to yield both forms.

In support of this view, genetic data obtained by mutagenesis of *dgk* indicate that the initial three codons or residues of a common precursor molecule hold the key to the alternative

specificities of dCK and dGK. Using normal N-terminal processing by the *E. coli* host, mutagenesis of the *dgk* gene by removing codons 2 and 3 (Thr and Val) leaves Ile next to the initiating Met, yielding both the sequence and specificity of wild-type dCK. It is not yet clear whether the two N-terminal sequences, TVIVL··· and MIVL···, contribute to alternative folding conformers, or if they participate directly in binding deoxynucleoside at the active center. It must be noted that within only five residues of the common Ile residue is the glycine-rich motif characteristic of the P-loop of an ATP-binding site (9), which may be presumed to be folded into the active site. Surprisingly, the K_m(dAdo) of the dAK subunit does not seem to be affected significantly by the composition of its first three residues, as may be inferred from Table 2, comparing the results with native dAK (dAK is MIV··) and recombinant dAK which has the sequence TVIV··. Of course, other elements of the dGK (or dCK) and dAK polypeptides also participate in nucleoside binding and/or turnover, and different tertiary structures may be predicted on the basis of the differing 40% of the two molecules. Site-directed replacement of the common Arg-78 by Lys has opposite effects on dGK and dAK, with an increased K_m for dGuo but a decreased K_m for dAdo (13). Moreover, the R78K mutation of dGK alone produced a dramatic increase in the turnover of unmutated dAK.

Given its identical mass, N-terminal sequences and kinetic behavior in comparison with native *Lactobacillus* dCK, we now feel justified in calling the products of mutants I and II engineered "dAK/dCK" or "dCK," respectively. Also, we considered whether this dCK derived from a bacterium resembles human dCK (4, 21–23). Although human and *Lactobacillus* dCK polypeptides are identical in only ≈50 residues (23%), these tend to be grouped at the ATP-binding and putative nucleoside-binding motifs, but are otherwise scattered throughout the entire length, except for their C and N termini, which are staggered and do not overlap (8). However, in terms of its quaternary structure, kinetics, ordered nucleoside specificities, and inhibition by dCTP, the bacterial dCK strongly mimics the human enzyme. dCyd is the primary substrate, as reflected in its relatively low K_m, while dGuo and dAdo have K_m values that are two to three orders larger, just as with the human enzyme.

Examples of altered substrate specificity by the mutagenesis of only one or two residues are accumulating—e.g., the conversion of thymidylate synthase to deoxycytidine methylase (24, 25), adenyl to guanyl cyclase (26), a shift toward pyrimidine specificity of adenylate kinase (27), and even the mechanism and functionality of an enzyme have been changed by an appropriate mutation (28). However, all of these examples involve internal residues, and we know of no comparable case involving the extreme N terminus.

There is clearly much to be learned about the location and conformation of the deoxynucleoside-specific sites and to what extent the N-terminal residues of the dAK, dCK, and dGK subunits differentially contribute to them. The ·DRS· motif of the *Lactobacillus* kinases is thought to be analogous with the conserved ·DRH· motif located within or near the putative active site of herpesviral thymidine kinase (TK); as with viral TK, for which the aspartate is essential (14, 15), Asp-78 is required for both dAK and dGK activities (13). However, it is not yet known what other domains might participate in comprising the active sites of these enzymes. Moreover, the mechanism by which the *Lactobacilli* can, in effect, delete the second and third residues while retaining the initiating methionine has yet to be discovered. The apparent lack of mRNA processing suggests that we must look for a mechanism occurring at the co- or posttranslational level. In *E. coli*, it is known that Thr in the second position stimulates the excision of the initial Met (16), as we have also observed with cloned dAK/dGK. If the subsequent hydrolytic removal of Thr and Val were to occur in *Lactobacillus*, resplicing Met to the Ile of the modified product might be catalyzed by an aminoacyl-tRNA-protein transferase analogous to those which mark proteins for degradation (29–31). Alternatively, the same structure could be created by an initiation complex capable of allowing the codon for valine (GUU) rather than AUG to serve as the initiating codon part of the time, although this seems less likely than if the valine codon were GUG. The latter mechanism has some precedent in the suppressor tRNAs which are capable of partially overriding stop codons; this could explain how both dGK and dCK N-terminal sequences can arise simultaneously. On the other hand, all of the dAK peptide is so modified, despite the fact that its 5′-terminal coding sequence is identical to that of *dgk(dck)*.

We thank Dr. J. Lowbridge for sequencing peptides, M. Lacey and T. Keough for mass spectrometric measurements, Dr. Qin Lu for synthesis of oligonucleotides, Dr. D. Camerini-Otero for generous provision of laboratory facilities, and Dr. C. Breitenberger for helpful comments. This work was supported in part by National Institute of General Medical Sciences Grant R01 GM49635.

1. Arnér, E. S. J. & Eriksson, S. (1995) *Pharmacol. Ther.* **67,** 155–186.
2. Chottiner, E. G., Shewach, D. S., Datta, N. S., Ashcraft, E., Gribbin, D., Ginsburg, D., Fox, I. H. & Mitchell, B. S. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 1531–1535.
3. Karlsson, A., Johansson, M. & Eriksson, S. (1994) *J. Biol. Chem.* **269,** 24374–24378.
4. Kim, M.-Y., Ikeda, S. & Ives, D. H. (1988) *Biochem. Biophys. Res. Commun.* **156,** 92–98.
5. Durham, J. P. & Ives, D. H. (1971) *Biochim. Biophys. Acta* **228,** 9–25.
6. Möllgaard, H. (1980) *J. Biol. Chem.* **255,** 8216–8220.
7. Ikeda, S., Ma, G. T. & Ives, D. H. (1994) *Biochemistry* **33,** 5328–5334.
8. Ma, G. T., Hong, Y. S. & Ives, D. H. (1995) *J. Biol. Chem.* **270,** 6595–6601.
9. Ikeda, S., Swenson, R. P. & Ives, D. H. (1988) *Biochemistry* **27,** 8648–8652.
10. Ikeda, S. & Ives, D. H. (1994) *Biochemistry* **33,** 13373–13381.
11. Ives, D. H. (1984) *Anal. Biochem.* **136,** 416–420.
12. Moos, M., Jr., Nguyen, M. Y. & Liu, T.-Y. (1988) *J. Biol. Chem.* **263,** 6005–6008.
13. Hong, Y. S., Ma, G. T. & Ives, D. H. (1995) *J. Biol. Chem.* **270,** 6602–6606.
14. Black, M. E. & Loeb, L. A. (1993) *Biochemistry* **32,** 11618–11626.
15. Fetzer, J., Folkers, G., Müller, I. & Keil, G. M. (1993) *Virus Genes* **7,** 205–209.
16. Hirel, P.-H., Schmitter, J.-M., Dessen, P., Gayat, G. & Blanquet, S. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 8247–8251.
17. Roush, A. H. & Betz, R. F. (1958) *J. Biol. Chem.* **233,** 261–266.
18. Deibel, M. R., Jr., & Ives, D. H. (1977) *J. Biol. Chem.* **252,** 8235–8239.
19. Chakravarty, R., Ikeda, S. & Ives, D. H. (1984) *Biochemistry* **23,** 6235–6240.
20. Ikeda, S., Chakravarty, R. & Ives, D. H. (1986) *J. Biol. Chem.* **261,** 15836–15843.
21. Kim, M.-Y. & Ives, D. H. (1989) *Biochemistry* **28,** 9043–9047.
22. Bohman, C. & Eriksson, S. (1988) *Biochemistry* **27,** 4258–4265.
23. Datta, N. S., Shewach, D. S., Mitchell, B. S. & Fox, I. H. (1989) *J. Biol. Chem.* **264,** 9359–9364.
24. Liu, L. & Santi, D. V. (1992) *Biochemistry* **31,** 5100–5104.
25. Hardy, L. W. & Nalivaika, E. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 9725–9729.
26. Beuve, A. & Danchin, A. (1992) *J. Mol. Biol.* **225,** 933–938.
27. Okajima, T., Tanizawa, K. & Fukui, T. (1993) *FEBS Lett.* **334,** 86–88.
28. Kim, J., Ruzicka, F. & Frey, P. A. (1990) *Biochemistry* **29,** 10590–10593.
29. Sohl, J. L. (1980) in *Transfer RNA: Biological Aspects*, eds. Söll, D., Abelson, J. N. & Schimmel, P. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 493–505.
30. Varshavsky, A. (1992) *Cell* **69,** 725–735.
31. Abramochkin, G. & Schrader, T. E. (1995) *J. Biol. Chem.* **270,** 20621–20628.