Pyruvate Carboxylase from *Rhizobium etli*: Mutant Characterization, Nucleotide Sequence, and Physiological Role

MICHAEL F. DUNN,* SERGIO ENCARNACIÓN, GISELA ARAÍZA, M. CARMEN VARGAS, ARACELI DÁVALOS, HUMBERTO PERALTA, YOLANDA MORA, AND JAIME MORA

Departamento de Ecología Molecular, Centro de Investigación sobre Fijación de Nitrógeno, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico

Received 20 May 1996/Accepted 13 August 1996

Pyruvate carboxylase (PYC), a biotin-dependent enzyme which catalyzes the conversion of pyruvate to oxaloacetate, was hypothesized to play an important anaplerotic role in the growth of *Rhizobium etli* during serial subcultivation in minimal media containing succinate (S. Encarnación, M. Dunn, K. Willms, and J. Mora, J. Bacteriol. 177:3058–3066, 1995). *R. etli* and *R. tropici pyc*::Tn5-mob mutants were selected for their inability to grow in minimal medium with pyruvate as a sole carbon source. During serial subcultivation in minimal medium containing 30 mM succinate, the *R. etli* parent and *pyc* mutant strains exhibited similar decreases in growth rate with each subculture. Supplementation of the medium with biotin prevented the growth decrease of the parent but not the mutant strain, indicating that PYC was necessary for the growth of *R. etli* under these conditions. The *R. tropici pyc* mutant grew normally in subcultures regardless of biotin supplementation. The symbiotic phenotypes of the *pyc* mutants from both species were similar to those of the parent strains. The *R. etli pyc* was cloned, sequenced, and found to encode a 126-kDa protein of 1,154 amino acids. The deduced amino acid sequence is highly homologous to other PYC sequences, and the catalytic domains involved in carboxylation, pyruvate binding, and biotinylation are conserved. The sequence and biochemical data show that the *R. etli* PYC is a member of the α_4 homotetrameric, acetyl coenzyme A-activated class of PYCs.

The biotin-dependent enzyme pyruvate carboxylase (PYC) (pyruvate: CO_2 ligase [ADP forming] [EC 6.4.1.1]) plays an important anaplerotic role in many organisms by catalyzing the carboxylation of pyruvate to form oxaloacetate (OAA), which is used to replenish tricarboxylic acid cycle intermediates used for energy generation and biosynthesis (for reviews, see references 4 and 54). In bacteria, anaplerotic CO_2 fixation is critical for the maintenance of the OAA pool during growth on compounds that are utilized via pyruvate (1, 28). Although PYC fulfills this function in many prokaryotes (1, 13, 46, 50, 68), enteric bacteria as well as some other groups convert phosphoenolpyruvate (PEP) to OAA in the reaction catalyzed by PEP carboxylase (PPC) (10, 37). These two pathways for OAA synthesis are not mutually exclusive, since some bacteria produce both PYC and PPC (18, 28, 38, 45, 55).

In eukaryotes and some prokaryotes (e.g., a thermophilic *Bacillus* strain [33] and *Rhodobacter capsulatus* [39]), the PYC holoenzyme is an α_4 homotetramer composed of 113- to 130-kDa subunits (4). Each subunit polypeptide contains three domains that work in concert to catalyze the overall reaction, which occurs as the sum of the two half-reactions shown below, where ENZ represents the enzyme.

$$MgATP + HCO_{3}^{-} + ENZ-biotin \underbrace{Mg^{2^{+}}, acetyl-CoA}_{MgADP + P_{i}} + ENZ-biotin-CO_{2}^{-}$$
(1)

$$ENZ-biotin-CO_2^- + pyruvate$$
(2)
$$ENZ-biotin + OAA$$

In reaction 1, the ATP-dependent biotin carboxylase domain carboxylates a biotin prosthetic group linked to a specific lysine residue in the biotin carboxyl carrier protein (BCCP) domain. Acetyl coenzyme A (acetyl-CoA) activates reaction 1 by increasing the rate of bicarbonate-dependent ATP cleavage. In reaction 2, the BCCP domain donates the CO₂ to pyruvate in a reaction catalyzed by the transcarboxylase domain (4).

The PYC holoenzymes from *Pseudomonas citronellolis* and *Azotobacter vinelandii* are exceptional in having an $\alpha_4\beta_4$ structure in which the α subunit (65 kDa) contains the catalytic domains (12, 26, 55). In contrast to the α_4 PYCs, all of which require or are activated by acetyl-CoA (33, 39, 42), the $\alpha_4\beta_4$ PYCs are fully active in the absence of this effector (12, 56). Although bacterial PYCs from both structural classes have been studied at the biochemical and physiological levels (12, 13, 26, 38, 39, 49, 55, 60), little genetic characterization has been done and only one nucleotide sequence for a putative prokaryotic PYC has been obtained (57).

Previous results from this laboratory led us to hypothesize that PYC plays an important role in determining if strains of *Rhizobium etli* and *R. tropici* display an aerobic or fermentative metabolism in vitro (17, 18). We showed that *R. etli* CE3 develops a fermentative metabolism, with a concomitant decrease in growth rate, during serial subcultivation in minimal medium (MM) containing succinate as a sole carbon source. The fermentative metabolism was prevented when the medium was supplemented with biotin or thiamine. The presence of either of these vitamins increased the activities of several carbon-metabolic enzymes, most notably PYC and pyruvate and oxoglutarate dehydrogenases, during subcultivation. Although the activities of all three enzymes were markedly increased with either supplement, biotin stimulated PYC to a greater

^{*} Corresponding author. Mailing address: Departamento de Ecología Molecular, Centro de Investigación sobre Fijación de Nitrógeno, Universidad Nacional Autónoma de México, A.P. 565-A, Cuernavaca, Morelos, Mexico. Phone: (73) 13-9944. Fax: (73) 17-5094. Electronic mail address: mike@cifn.unam.mx.

Strain or plasmid	Relevant characteristics ^a	Reference or source
R. etli strains		
CE3	Derivative of wild-type strain CFN42, Sm ^r Nal ^r Suc ⁺ Glc ⁺ Pyr ⁺	7
12-53	CE3 pyc::Tn5-mob, Sm^r Nal ^r Km^r Suc ⁺ Glc [±] Pyr ⁻	This study
R. tropici strains		
CFN299	Wild-type strain, Sp ^r Nal ^r Suc ⁺ Glc ⁺ Pyr ⁺	36
2-176	CFN299 pyc::Tn5-mob, Sp ^r Nal ^r Km ^r Suc ⁺ Glc ^{\pm} Pyr ⁻	This study
E. coli strains		-
S17-1	294 recA; mobilizing strain for Tn5	56
DH5a	$recA1 \ lacU169 \ \phi 80 \overline{d} lacZ\Delta M15$	Bethesda Research Laboratories
HB101	F^- recA13; host for R. etli pLAFR1 genomic library	Laboratory stock
Plasmids		-
pLAFR1	Broad-host-range cloning vector, Tc ^r	21
pSUP5011	pBR233 Tn5-mob, Ap ^r Cm ^r Km ^r	56
pRK2013	Helper plasmid, Km ^r	20
pBluescript II SK (+)	Cloning vector, Ap ^r	Stratagene
pHPD5g	pBluescript containing the Tn5-mob insertion and flanking regions from strain 12-53	This study
pDG1	pBluescript with an <i>Eco</i> RI fragment containing the Tn5-mob insertion from strain 12-53	This study
pDG1-2	<i>Eco</i> RI- <i>Hin</i> dIII subclone of pDG1	This study
p2-176	pBluescript with an <i>Eco</i> RI fragment containing the Tn5-mob insertion from strain 2-176	This study
p2-1	EcoRI-HindIII subclone of p2-176	This study
pPC	pLAFR1 harboring 15.2 kb of R. etli genomic DNA containing pyc; Tc ^r	This study
pPC1	EcoRI deletion subclone of pPC containing R. etli pyc	This study

TABLE 1. Bacterial strains and plasmids used in this study

^{*a*} Suc, succinate: Glc, glucose; Pyr, pyruvate; +, normal growth; \pm , reduced growth; -, no growth.

degree than it did the dehydrogenases, whereas thiamine had the opposite effect. In contrast to R. etli, R. tropici CFN299 maintained high levels of these activities during subcultivation in unsupplemented MM and did not develop a fermentative metabolism (18). We chose to further study the effect of biotin supplementation on the growth of R. etli and to test our hypothesis that PYC was the major site at which biotin acts in promoting an aerobic metabolism in this species (15, 18). A corollary to this hypothesis is that R. tropici might maintain its growth in unsupplemented medium because it does not lose PYC activity under these conditions. Our approach, as reported here, was to inactivate the pyc gene in R. etli and R. tropici and characterize the free-living and symbiotic phenotypes of the mutants. To begin a comparative genetic and biochemical analysis of the PYCs produced by these species, we cloned and sequenced the R. etli pyc and also report some characteristics of the enzyme.

(Preliminary reports of part of this work were presented previously [15, 17].)

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. Bacterial strains and plasmids used in this study are listed in Table 1. Rhizobia were grown either in PY rich medium (8) or in MM (5) containing a carbon source and filter-sterilized biotin as indicated for each experiment. To prepare MM plates, carbon sources were added to a final concentration of 20 mM and media were solidified with 1.5% purified agar (Merck, Darmstadt, Germany). For the growth of *Rhizobium* strains in batch cultures, cells were prepared and media were inoculated and incubated as described previously (18). To culture cells in the presence of radiolabeled biotin, MM-succinate (10 mM) containing 1 ng of unlabeled biotin plus 0.05 μ Ci (0.28 ng) d-[8,9-³H]biotin per ml was inoculated and incubated for 18 h as described previously (18) except that 1 liter of medium in a 2.8-liter flask was used. Strains of *Escherichia coli* were grown in LB medium (51). Cultures of *Rhizobium* spp. or *E. coli* were supplemented with antibiotics at the following concentrations (in micrograms per milliliter): spectinomycin, 200; streptomycin, 200; nalidixic acid, 20; tetracycline, 10; kanamycin, 30; and carbenicilin, 100.

Mutagenesis of *Rhizobium* **spp. and selection of** *pyc* **mutants.** Cells were mutagenized by a standard biparental mating with *E. coli* S17-1/pSUP5011 as the donor strain. Clones containing Tn5-mob insertions were selected by spreading dilutions of 24-h matings onto PY medium containing spectinomycin or strep-

tomycin (for strains CFN299 or CE3, respectively), nalidixic acid, and kanamycin. After incubation for 3 days at 30°C, colonies were patched onto MM containing succinate or pyruvate as a sole carbon source. Clones able to grow on succinate but unable to grow on pyruvate were selected and further characterized.

Isolation of the *R. etli pyc* **gene.** A genomic DNA bank (Table 1) prepared from *R. etli* CE3 was mobilized to the *R. tropici pyc* mutant 2-176 (Table 1), using pRK2013 as a helper strain. Mating mixtures were incubated on PY medium for 24 h at 30°C, resuspended in 0.85% NaCl, and spread onto MM with glucose or pyruvate as a sole carbon source and containing spectinomycin, nalidixic acid, and tetracycline. One colony appearing on a MM-glucose plate after 3 days of incubation was purified by restreaking and used for further studies. The complementing plasmid isolated from this clone was designated pPC (Table 1).

DNA manipulations. Routine DNA manipulations were performed by standard techniques (51). Southern blot hybridizations to detect transposon insertions were performed with ³²P-labeled Tn5-mob prepared by using a Megaprime kit (Amersham International) and were performed as described previously (51). Indigenous plasmids from the pyc mutants were isolated and electrophoretically separated by the Eckhardt procedure (16). The DNA fragment containing the Tn5-mob insertion in R. tropici 2-176 (Table 1) was isolated by completely digesting genomic DNA with EcoRI, ligating the fragments into pBluescript, and selecting carbenicillin- and kanamycin-resistant transformants of E. coli DH5a. The plasmid obtained was designated p2-176. The Tn5-mob insertion from R. etli 12-53 was cloned by a similar strategy except that partial EcoRI digests of the genomic DNA were used so that the insertion and several contiguous EcoRI fragments were cloned. This plasmid was designated pHPD5g. The EcoRI fragment containing the insertion was subcloned from pHPD5g and designated pDG1 (Table 1). DNA flanking either end of the Tn5-mob insertions in p2-176 and pDG1 were subcloned as HindIII-EcoRI fragments (where the HindIII site is within the transposon and the EcoRI site is within the genomic DNA) to obtain plasmids pDG1-2 and p2-1, respectively (Table 1). The nucleotide sequence of 100 bp near the Tn5-mob insertion subcloned from R. tropici 2-176 (plasmid p2-1) was determined by using an EcoRI primer and was 91% identical to a segment of the cloned R. etli pyc sequence. By correlating this information with the distance of the sequenced portion from the Tn5-mob insertion in p2-1, the probable site of the insertion was determined. This approximation is based on the assumption that the R. tropici pyc is similar in size (see Discussion) and sequence to the R. etli pyc. The Tn5-mob insertion site in R. etli 12-53 was localized by sequencing the DNA from pDG1-2 with a primer specific for the *XhoI* site located near the border of the transposon. The DNA sequencing described above was performed by the Sanger dideoxy method (53). Doublestranded sequencing of the R. etli pyc was performed by Medigene GmbH, Munich, Germany, using an Applied Biosystems (Foster City, Calif.) 373A DNA sequencing system. Nucleotide and protein sequence homology searches were made by using the BLAST program (2) via the National Institute for Biotech-

TABLE 2. Effect of culture carbon source and biotin supplementation on PYC activity in *R. etli* CE3^{*a*}

Carbon source	Culture biotin concn (µg/ml)	PYC activity (nmol/min/mg of protein)
Succinate	0	$6.3 \pm 0.5 (1.0)$
	0.00005	$5.2 \pm 1.5(0.83)$
	0.00025	$24.2 \pm 6.5 (3.8)$
	0.05	$28.2 \pm 4.7 (4.5)$
	0.5	27.3 ± 2.8 (4.3)
	1.0	$26.5 \pm 2.5 (4.2)$
Glucose	0	$10.5 \pm 1.6(1.0)$
	1.0	28.9 ± 2.3 (2.8)
Pyruvate	0	$16.9 \pm 6.3 (1.0)$
-	1.0	$42.4 \pm 2.5(2.5)$

^{*a*} Cells were grown for 16 h in MM containing the indicated carbon source at 10 mM with or without biotin supplementation as indicated. Activities are the means \pm standard errors for values obtained in two to three separate experiments. For each carbon source, values in parentheses are normalized to the activity of cells cultured without added biotin, set at unity.

nology Information server. Multiple protein sequence alignments were made with Clustal W version 1.5 (63) at the Baylor College of Medicine-Human Genome Center server.

Enzyme and protein assays. For enzyme activity screening, cells grown for 3 days on PY plates containing the appropriate antibiotics were washed twice in sterile PY and used to inoculate 100 ml of PY medium to an initial A_{540} of 0.05. Cultures were incubated at 30°C for 16 h with shaking at 200 rpm. The late-logphase cells were harvested by centrifugation (9,800 \times g, 10 min) and washed twice in 20 mM potassium phosphate buffer (pH 7.2) containing 0.85% NaCl. The cells were resuspended in 5 ml of cell lysis buffer (20 mM potassium phosphate [pH 7.2], 5 mM MgCl₂, 1 mM EDTA, 1 mM dithiothreitol, 0.75% glycerol) and lysed by sonication as described previously (18). To prepare protein extracts from cells grown in the presence of [³H]biotin or for the partial purification of PYC, cells were harvested and washed as described above, resuspended in 10 ml of cell lysis buffer per g (fresh weight), and sonicated in 10 1-min bursts using a MSE sonicator (Dalon Scientific Ltd., London, United Kingdom) with a 1-cm probe at 75% of maximum power, with cooling of the sample to 5°C between bursts. Cell-free supernatants were obtained by centrifugation at 90,000 $\times g$ for 1.5 h at 4°C. Enzyme assays were performed as described previously (18) except that PYC was assayed in the presence of 20 µM acetyl-CoA. Protein contents of culture biomass and cell lysates were determined as described previously (18).

Partial purification of PYC from R. etli. Batch cultures of strain CE3/pPC1 (Table 1) were grown for 16 h in a total of 3 liters of MM-pyruvate (30 mM) containing biotin (1 µg/ml) and tetracycline. Cell extracts were prepared as described above, and proteins were precipitated at 4°C by the addition of solid (NH₄)₂SO₄. Following equilibration for 20 min at each final salt concentration, precipitates were collected by centrifugation at $10,000 \times g$ for 10 min. The precipitate from the 40 to 60% of saturation fraction contained the majority of the PYC activity and was stored overnight at -80°C. The protein was redissolved in a small volume of equilibration buffer (EQB; 20 mM potassium phosphate [pH 8.0], 5 mM MgSO₄, 20 mM KCl, 5% glycerol, 1 mM dithiothreitol, 1 mM EDTA) and desalted on a Sephadex G-25 column equilibrated in EQB. The desalted sample was chromatographed on a DEAE-cellulose column (1 by 14 cm) equilibrated in EQB and eluted at 1.5 ml/min. Following the elution of nonbinding proteins, the column was eluted with a linear 20 to 250 mM gradient of KCl in EQB over 134 min. Fractions containing PYC activity were pooled and precipitated with (NH₄)₂SO₄ to 60% saturation. The preparation was centrifuged and the pellet was stored as described above. The protein pellet was redissolved in a small volume of 1 mM NaCl (unbuffered) and desalted on a G-25 column equilibrated in the same buffer. The desalted preparation was loaded onto a column (0.7 by 8 cm) of Bio-Gel HTP hydroxylapatite (Bio-Rad Corporation, Hercules, Calif.) equilibrated in the same buffer. The column was washed with about 2 column volumes of 1 mM NaCl, and bound proteins were eluted with a 50-ml linear gradient of 0 to 0.3 M potassium phosphate at 0.5 ml/min. Active fractions were pooled, precipitated, and stored as described above. Before use in enzyme assays, the precipitate was redissolved in a minimal amount of cell lysis buffer and dialyzed against the same buffer. All purification steps were carried out at 4°C.

Detection of biotin-containing proteins by Western blotting (immunoblotting) and [³H]biotin labeling. Proteins in cell extracts were separated in sodium dodecyl sulfate (SDS)-10% polyacrylamide gels (23) and electroblotted onto nitrocellulose for 1 h at 300 mA as described by Towbin et al. (65). Biotincontaining proteins on the blots were detected by using a Sigma-Blot protein detection kit (Sigma Chemical Corp., St. Louis, Mo.) as instructed by the manufacturer but omitting the protein biotinylation step so that only proteins naturally containing biotin would bind the streptavidin-horseradish peroxidase conjugate. Blots were imaged with a pdi Desktop scanner and Diversity One software (pdi Inc., Huntington Station, N.Y.). Protein bands were manually contoured prior to quantitation. Radioactivity in biotin-containing proteins prepared from cultures grown in the presence of [³H]biotin was quantitated by excising the bands, dissolving them in 5 ml of scintillation fluid (9), and counting in a scintillation counter. Portions of membrane excised from areas in the lanes not containing biotinylated proteins were included as background checks. Specific radioactivity was calculated by dividing the counts per minute obtained for a band by its contour quantity (units of optical density [OD] multiplied by area [square millimeters]) obtained by scanning densitometry.

Nodulation and nitrogen fixation assays. Seeds of *Phaseolus vulgaris* cv. Negro jamapa were surface disinfected and germinated as described previously (8). Cultures of the *pyc* mutants and their respective parent strains were grown to saturation overnight in PY media containing the appropriate antibiotics (Table 1), washed once in 0.85% NaCl, and resuspended to an OD at 540 nm of 0.05. One milliliter of this suspension was used to inoculate 3-day-old seedlings, and plants were maintained aseptically under greenhouse conditions. At 12, 15, 18, 25, 32, 40, and 55 days postinoculation, plants were harvested and analyzed for nodule number and dry weight, plant dry weight, and acetylene reduction activity as described previously (11). Seed yields were determined at 116 days postinoculation. Nodule microsymbiont occupancy was determined at all time points by plating nodule homogenates on PY medium with or without the antibiotics (Table 1) required to differentiate each strain.

Nucleotide sequence accession number. The nucleotide sequence data presented in this report are available under GenBank accession number U51439.

RESULTS

Physiological and biochemical characterization of the PYC from *R. etli* CE3. Because PYC activity in strain CE3 is significantly affected by growth conditions (17, 18), we examined the effects of culture carbon source and biotin supplementation on the production of PYC in this strain (Table 2). In cells cultured in MM without added biotin, PYC activities were 2.7-fold higher in pyruvate than in succinate, while glucose-grown cells had an intermediate level of activity. Supplementation of the media with 1 μ g of biotin per ml increased PYC activities, relative to unsupplemented media, 4.2-fold in succinate, 2.8-fold in glucose, and 2.5-fold in pyruvate. In cells grown in MM-succinate, a biotin concentration of 0.25 ng/ml was sufficient to induce PYC activity to near-maximal levels (Table 2).

For use as a standard on Western blots, PYC was partially purified from *R. etli* CE3 by ion-exchange chromatography (14). On Western blots developed to detect biotinylated proteins, this preparation gave an intense band which migrated with an apparent molecular weight (MW) of 120,000 (Fig. 1). In blots of extracts prepared from cells grown in MM-succinate



FIG. 1. Effect of culture medium biotin supplementation on the production of biotin-containing proteins in *R. etli* CE3. Biotin-containing proteins were detected in the Western blot as described in Materials and Methods. Lanes: 1, biotinylated protein standards with molecular masses as indicated; 2, partially purified *R. etli* PYC; 3, extract from cells cultured in MM-succinate (10 mM) containing 1 μ g of biotin per ml; 4, extract from cells cultured in MM-succinate (10 mM) without added biotin. Lanes 3 and 4 were each loaded with 110 μ g of total protein.

TABLE	3.	Effect	of assa	y reaction	n mixture	composition
		on the	activity	of PYC	from R. e	tli

Reaction mixture ^a	Activity (mU) ^b	Relative rate
Standard	43.22	1.00
Standard – pyruvate	0.48	0.01
Standard – ATP	1.05	0.02
Standard $- \text{HCO}_3^-$	1.84	0.04
Standard $-$ MgCl ₂	NA^{c}	0.00
Standard – acetyl-CoA	1.72	0.04
Standard + 2 μ M acetyl-CoA	2.75	0.06
Standard + 5 μ M acetyl-CoA	14.59	0.34
Standard + 20 μ M acetyl-CoA ^d	44.49	1.03
Standard + 50 μ M acetyl-CoA	43.33	1.00
Standard + 200 µM acetyl-CoA	44.25	1.02
Standard + 1 mM L-Asp	34.73	0.80
Standard + 5 mM L-Asp	4.22	0.10
Standard + 10 mM L-Asp	2.16	0.05
Standard + 10 mM 2-OG	45.00	1.04
Standard + 10 mM L-Glu	45.72	1.06

^{*a*} The standard reaction mixture contained (per milliliter) *N*-2-hydroxyethylpiperazine-*N*'-2-ethanesulfonic acid (HEPES) buffer (pH 8.0), 100 mM; MgCl₂, 10 mM; ATP, 2.5 mM; NaHCO₃, 20 mM; NADH, 0.2 mM; acetyl-CoA, 20 μ M; malate dehydrogenase; 0.5 U; and pyruvate, 10 mM. Substrates and amino or organic acid stocks used in the reactions were prepared in 50 mM HEPES buffer (pH 7.0).

^b One unit of activity catalyzes the reduction of 1 nmol of NADH per min. The specific activity of the enzyme preparation used in all assays was 2,470 nmol/ min/mg of protein.

^c NA, no activity detected.

^d Note that this reaction mixture is identical to that used in the standard assay.

without added biotin, PYC accounted for greater than 99% of the total biotin-containing proteins detected; a 51-kDa protein (see below) comprised the remainder of the biotinylated protein in the sample (Fig. 1). In extracts obtained from cells cultured in the presence of biotin, the quantity of PYC protein increased 5.3-fold (Fig. 1), in fairly close agreement with the 4.2-fold increase in enzymatic activity observed under the same conditions (Table 2). Two additional biotin-containing proteins with MWs of 51,000 and 14,000 were induced to easily detectable levels when cells were grown in biotin-supplemented medium (Fig. 1). By scanning densitometry, PYC and the 51- and 14-kDa proteins accounted for 79, 19.4, and 1.6%, respectively, of the total biotinylated protein in the sample.

When strain CE3 was grown in MM-succinate supplemented with [³H] biotin, the quantity of biotin in the media was reduced 93% after 18 h of growth. On Western blots prepared from these cultures, specific radioactivities for PYC and the 51and 14-kDa bands were in the range of 800 to 6,000 cpm/OD \times mm², while background radioactivity in sample blanks, prepared as described in Materials and Methods, were less than 1% of these values.

Holo-PYC was overproduced when strain CE3 containing plasmid-borne copies of the *R. etli pyc* (pPC1; Table 1) was grown in biotin-containing media. The enzyme was purified 12-fold to a specific activity of 2,470 nmol/min/mg of protein. The activity of the enzyme was substantially dependent on the presence of substrates (pyruvate, MgATP, and bicarbonate) in the reaction mixture. In the absence of acetyl-CoA, the enzyme retained 4% of its activity relative to that measured in the standard assay (Table 3). From these data, we calculate that the acetyl-CoA concentration necessary for half-maximal activation of the enzyme is 8.8 μ M. The enzyme was inhibited 50% by 2.8 mM L-aspartate but was not inhibited by 2-oxoglutarate or L-glutamate (Table 3). When CE3/pPC1 was grown in MMsuccinate lacking biotin, virtually all of the PYC produced was present in the apo form (14). Genetic and phenotypic characterization of the *Rhizobium pyc* mutants. Putative *pyc* mutants were selected from populations of Tn5-mob-containing clones on the basis of their inability to grow on solid MM-pyruvate. Enzyme assays showed that two clones (of 607 screened) with this phenotype isolated from *R. tropici* CFN299 were both PYC negative, and one, strain 2-176 (Table 1), was used in further studies. For *R. etli* CE3, 6,500 clones were screened and 3 having a pyruvate-negative phenotype were isolated. Of these three clones, only strain 12-53 (Table 1) lacked PYC activity.

Hybridization of a Tn5-mob probe against *Eco*RI fragments of genomic DNA prepared from the pyc mutants and the parent strains showed that each mutant contained a single Tn5-mob insertion, and no hybridization was observed against DNA isolated from the parent strains. The hybridizing band for mutant 12-53 was approximately 8.5 kb, while that for mutant 2-176 was 8.0 kb. By correlating the size of the hybridizing fragment from R. etli 12-53 with the restriction enzyme map of pHPD5g (Table 1) (47), we found that the DNA surrounding the insertion had the same restriction sites as the cloned R. etli pyc (47) and that the insertion site was in a fragment corresponding to the 0.86-kb EcoRI fragment of pPC (see Fig. 6a). When indigenous plasmids from the pyc mutants were separated on Eckhardt gels and hybridized with Tn5mob, no hybridization was observed. As a positive control, plasmids isolated from an R. etli mutant with a Tn5-mob insertion in the b plasmid (61) were run in the same gel and gave a strong hybridization signal (19).

The Tn5-mob-containing DNA fragments cloned from both mutants were analyzed by partial sequencing. The insertion site in *R. tropici* 2-176 was approximated as described in Materials and Methods and is estimated to occur approximately midway in the *pyc* open reading frame (15). The insertion site in *R. etli* 12-53 is very near the 5' end of the gene (following nucleotide 117) and is in the same *Eco*RI fragment predicted by hybridization and restriction enzyme mapping.

Growth curves for the mutant and parent strains in MMpyruvate (10 mM) show that the mutants of either species were unable to grow on this carbon source except when it was supplemented with 1 mM L-aspartate (Fig. 2A). In MM-glucose, the *R. etli* and *R. tropici pyc* mutants grew slowly and gave



FIG. 2. Growth of *Rhizobium* wild-type and *pyc* mutant strains in MM-pyruvate (10 mM) with and without 1 mM L-aspartate (A) and in MM-glucose (10 mM) (B). Strains tested were *R. etli* CE3 (closed circles), *R. tropici* CFN299 (open circles), *R. etli pyc* mutant 12-53 (open boxes), and *R. tropici pyc* mutant 2.176 (closed boxes). Symbols in panel A for growth in pyruvate-containing aspartate: closed triangles, 12-53; open triangles, 2-176.



FIG. 3. Growth of *R. etli* parent strain CE3 (closed symbols) and *pyc* mutant 12-53 (open symbols) during subcultivation in MM containing 30 mM succinate with or without 1 μ g of biotin per ml. (A) Strain CE3 without biotin; (B) strain 12-53 without biotin; (C) strain CE3 with biotin; (D) strain 12-53 with biotin. Symbols: circles, first subculture; squares, second subculture; triangles, third subculture.

final cell yields of 41 and 34% in comparison with the respective parent strains (Fig. 2B). Supplementation of the pyruvate or glucose minimal medium with biotin (1 μ g/ml) or thiamine (10 μ g/ml) did not allow growth of the mutants (19).

During serial subcultivation in MM containing 10 or 30 mM succinate without added biotin, the *R. tropici* parent strain



FIG. 4. Western blot analysis of PYC produced by the *Rhizobium pyc* mutants and complemented strains. Cells were grown in biotin-supplemented media as indicated in Materials and Methods. The positions of PYC and the 116-kDa protein standard are indicated. Sample lanes were loaded with 12.4 μ g of total protein.

and pyc mutant showed no differences in growth rate or final cell yield (19). In MM containing 10 (19) or 30 mM (Fig. 3A and B) succinate without added biotin, R. etli pyc mutant 12-53 decreased its growth rate in a manner similar to that of the parent strain. In biotin-supplemented MM containing 10 mM succinate, both the mutant and the parent strain grew well, with the mutant having only a slightly lower final cell yield in the third subculture (19). Under these conditions, thiamine supplementation of the MM also allowed good growth of strain 12-53 in subcultures (19). A physiological effect of the pyc mutation in R. etli was much more apparent during subculturing in MM containing 30 mM succinate, in which case addition of biotin to the medium allowed continuous growth of the parent strain (Fig. 3C) but not of the mutant (Fig. 3D). The effect of thiamine supplementation on the growth of the mutant has not been determined under these conditions.

Enzyme activities in the pyc mutants. Preliminary PYC activity assays of cell extracts detected a low to intermediate level of apparent PYC activity in preparations from both pyc mutants. We therefore assayed PYC in assay reaction mixtures with and without ATP in order to differentiate true, ATPdependent PYC activity from nonspecific background activity (Table 4). The apparent PYC activity detected in the mutant strains, in contrast to that of the parent strains, was not significantly decreased when ATP was excluded from the assays of cell lysates prepared from cells grown in PY or MM-succinate (Table 4), indicating that this apparent PYC activity is actually due to an ATP-independent enzyme and not to PYC. The lack of detectable, biotinylated PYC protein on Western blots of proteins prepared from the mutants (Fig. 4) supports this conclusion. The activities of several additional carbon-metabolic enzymes were assayed and did not differ significantly in the mutants and their respective parent strains (Table 5).

TABLE 4. PYC specific activities in Rhizobium wild-type and pyc mutant strains

			Mean (nmol/min/mg	g of protein) \pm SE ^{<i>a</i>}		
Strain		PY (rich) medium		MM-succinate (10 mM) + biotin (1 µg/ml)		
	Assay + ATP	Assay – ATP	ATP-dependent activity	Assay + ATP	Assay – ATP	ATP-dependent activity
R. etli						
CE3	43.5 ± 1.7	13.2 ± 0.8	30.3	28.4 ± 2.0	4.5 ± 0.5	23.9
12-53	9.9 ± 1.6	13.0 ± 1.6	NA^b	0.95 ± 0.95	3.1 ± 3.4	NA
R. tropici						
CFN299	57.1 ± 1.3	22.0 ± 1.6	35.1	55.2 ± 1.3	15.6 ± 2.0	39.6
2-176	9.8 ± 0.7	7.3 ± 0.9	2.5	11.8 ± 2.6	13.0 ± 0.5	NA

^a Assay + ATP is the standard PYC assay (see Table 3); assay - ATP lacked only ATP. ATP-dependent activity is the difference between these two values. ^b NA, no activity detected.

TABLE 5. Activities of selected carbon-metabolic enzymes in the parent and pyc mutant strains^{*a*}

	A	Activity (nmol/m	in/mg of protein	1)
Enzyme	R. etli R. tropici		opici	
	CE3	12-53	CFN299	2-176
PDH	28 ± 8	24 ± 2	21 ± 5	20 ± 3
ODH	25 ± 6	20 ± 1	54 ± 13	58 ± 9
MDH	$2,142 \pm 645$	$2,421 \pm 861$	$3,518 \pm 631$	$4,058 \pm 845$
РҮК	190 ± 14	170 ± 22	130 ± 5	164 ± 20
PCK	130 ± 11	122 ± 13	161 ± 16	168 ± 28
PPC	52 ± 15	44 ± 20	54 ± 20	30 ± 2
$ME(NAD^+)$	152 ± 21	129 ± 24	103 ± 5	75 ± 8
ME (NADP ⁺)	88 ± 43	77 ± 24	97 ± 7	86 ± 3

^{*a*} Duplicate PY cultures were harvested at 16 h, and lysates prepared from each were assayed in duplicate. Results are the means \pm standard errors from two separate experiments. Abbreviations: PDH, pyruvate dehydrogenase (EC 1.2.2.2); ODH, oxoglutarate dehydrogenase (EC 1.2.4.2); MDH, malate dehydrogenase (EC 1.1.1.37); PYK, pyruvate kinase (EC 2.7.1.40); PCK, phosphoenol pyruvate carboxykinase (EC 4.1.1.49); PPC, phosphoenolpyruvate carboxylase (EC 4.1.1.31); ME, malic enzyme (assayed with NAD⁺ [EC 1.1.1.39] or NADP⁺ [EC 1.1.1.40] as indicated).

Cloning and characterization of the PYC gene from R. etli. Plasmid pPC (Table 1) was isolated from an R. etli genomic DNA bank on the basis of its ability to restore the growth phenotype of R. tropici 2-176. As expected, the presence of pPC in either pvc mutant 2-176 or pvc mutant 12-53 restored the normal growth phenotype in MM-pyruvate (Fig. 5). In MM-succinate (10 mM) supplemented with biotin, the level of ATP-dependent PYC activity produced by the complemented R. tropici mutant was 84 U/mg of protein, or about twofold higher than that of the parent strain (Table 4). When pPC was introduced into the R. etli mutant, the ATP-dependent PYC activity was 182 U/mg of protein, a 7.6-fold increase over that produced by the parent strain (Table 4). In Western blots of cell extracts prepared from the complemented mutants, the PYC band, absent in the mutants, was restored (Fig. 4). Densitometric quantitation of the PYC protein produced by the complemented mutants relative to the parent strains (Fig. 4) revealed 3.2- and 6.6-fold increases for strains 2-176/pPC and 12-53/pPC, respectively. These values are in fairly close agreement with the increase in PYC activity determined in the enzyme assays.

The growth phenotypes and PYC activities of the *R. tropici* mutant containing subclones of pPC are shown in Fig. 6a. Partial nucleotide sequencing of *Eco*RI fragments derived from subclone pPC1 indicated that the putative PYC-coding region was located in the 3' half of the fragment, and subsequent sequencing of this region localized the *pyc* open reading frame as shown in Fig. 6b. The 3,462-bp sequence of the *R. etli pyc* is predicted to encode a protein of 1,154 amino acids with an MW of 126,008, which is similar to the subunit MW of 120,000 estimated by SDS-polyacrylamide gel electrophoresis (PAGE) (Fig. 1). The putative initiation codon (TTG) is preceded by a ribosomal binding site sequence (GGAGG [25]) 9 nucleotides upstream.

Nucleotide sequence similarity comparisons showed that the *R. etli pyc* had significant homology only to *pyc* or other biotin carboxylase sequences in the databases (14). The deduced amino acid sequence of the *R. etli pyc* is 44% identical to that of the putative *Mycobacterium tuberculosis* PYC and 47% identical to sequences of both yeast and human PYCs (Fig. 7). The N terminus is 77% identical to the first 26 amino acid residues of the *R. capsulatus* PYC recently determined by amino

acid sequencing (39) (Fig. 7). The *R. etli* PYC sequence (1,154 amino acids) is intermediate in length between those of yeast and human PYCs (both containing 1,178 amino acids including the leader sequence) and the putative *M. tuberculosis* (1,124 residues) PYC (Fig. 7 and 8).

A comparison of substrate-binding-site motifs in the R. etli sequence (Fig. 7) with those found in related proteins is shown schematically in Fig. 8. In the amino-terminal segment of the R. etli PYC, the hexapeptide GGGGRG matches the GGGG (R/K)G sequence found in all biotin-containing enzymes (67) and is believed to function in ATP binding (22, 48) (Fig. 7 and 8). A variant of this motif also occurs in carbamylphosphate synthetase, which shares some mechanistic properties with PYC (4) but does not require a biotin cofactor (29) (Fig. 8). A Cys residue (Cys-237 [Fig. 7]) is conserved 62 residues downstream of the ultimate Gly residue in this site and is believed to be involved in the CO₂ fixation reaction catalyzed by biotindependent carboxylases (32) (Fig. 7 and 8). A second region proposed to be involved in ATP binding is present in carbamylphosphate synthetase and biotin-dependent carboxylases (34) and is also conserved in the R. etli sequence (Fig. 7). In the intermediate region of the protein, a putative pyruvate-binding motif, FLTEDPWER, has significant homology with those found in the transcarboxylase domains of Mycobacterium, yeast, and human PYCs (Fig. 7) as well as in the Propionibacterium shermanii 5s subunit of transcarboxylase (Fig. 8). Tryptophan fluorescence studies with transcarboxylase indicate that the Trp residue present in this motif is involved in pyruvate binding (30). In the carboxy-terminal segment of the enzyme, a putative biotin-binding site (AMKM) is identical to those found in the other PYCs (Fig. 7) as well as the BCCP domains of other biotin-dependent enzymes (Fig. 8). Beginning 29 residues upstream of the biotinyl-lysine in all of the PYC sequences (Fig. 7) and in the α subunit of propionyl-CoA carboxylase (Fig. 8), a PX(P/A) sequence occurs; this sequence is proposed to be involved in the recognition and biotinylation of the biotin carboxylases (32).

Symbiotic phenotypes of the *pyc* **mutants in bean plants.** At all time points assayed, the symbiotic phenotypes of the *R. etli* and *R. tropici pyc* strains were statistically indistinguishable from those of the parent strains by the criteria of nitrogen



FIG. 5. Phenotypic complementation of the *Rhizobium pyc* mutants for growth in MM-pyruvate (10 mM) when complemented with the cloned *pyc* gene (pPC) from *R. etli*. Strains tested were *R. etli pyc* mutant 12-53 (open boxes), 12-53/pPC (open triangles), *R. tropici pyc* mutant 2-176 (closed boxes), and 2-176/pPC (closed triangles). Growth curves for the mutant strains are reproduced from Fig. 2A.



FIG. 6. (a) Subcloning of the *R. etli pyc* gene. Subclones were generated by deleting *Eco*RI fragments from plasmid pPC (Table 1). Complementation refers to the restoration of the wild-type growth phenotype in MM-pyruvate. PYC specific activities (nanomoles per minute per milligram of protein) were determined in 12-h PY (for pPC) or MM-succinate plus biotin (all others) cultures with cell extracts from strain 2-176 containing the indicated plasmid. E, *Eco*RI; B, *Bam*HI; H, *Hind*III; nd, not detected. (b) Sequencing strategy and the location of the *R. etli* pyc open reading frame (ORF) in pPC1.

fixation activity, number of nodules formed, and nodule and plant dry weights. Final seed yields of plants inoculated with the mutant or parent strains were indistinguishable (40). Microsymbionts isolated from nodules at all time points retained the expected antibiotic resistances.

DISCUSSION

Biotin stimulates the growth of a variety of fast-growing Rhizobium species under laboratory conditions (19, 59, 66), and the enzyme activities affected by this vitamin have been investigated in R. etli and R. tropici (18). We now extend these observations by showing that in biotin-supplemented MM, R. etli CE3 produces three biotinylated polypeptides, the 120kDa PYC subunit and the 51- and 14-kDa proteins, neither of which we have identified (Fig. 1). The unidentified biotin proteins were also produced by pyc mutant strain 12-53 (14), indicating that they are not degradation products of PYC. Extracts of R. tropici contained biotinylated proteins with sizes very similar to those produced by R. etli (14), and because the pyc mutant 2-176 lacked only a 120-kDa protein produced by the wild-type strain (Fig. 4), this protein probably represents the PYC subunit in this species. Thus, it appears that the PYCs of R. etli and R. tropici are similar in size. The results of our hybridization studies indicate that the pyc genes in these species are not borne on any of their indigenous plasmids but are instead chromosomal.

When strain CE3 was grown in MM-succinate without added biotin, PYC was easily detected in Western blots but at a level of less than 20% of that produced in biotin-containing

cultures (Fig. 1). Our data show that strain CE3 can effectively take up exogenously supplied biotin from the medium and incorporate it into PYC and the other biotin-containing proteins. However, under some conditions in which biotin is limiting, this strain appears to produce insufficient biotin to fully biotinylate PYC (Fig. 1) (19). It is important to note several lines of evidence suggesting that R. etli CE3 is not a biotin auxotroph. (i) The growth arrest observed when strain CE3 is subcultured in MM without added biotin occurs only after two successive subcultures (18, 19), whereas biotin auxotrophs show no or very little growth when inoculated into medium lacking the vitamin (24, 59). (ii) Rhizobium meliloti 1021, like R. etli CE3, decreases its growth rate in successive MM subcultures in the absence of biotin (18, 19). Nevertheless, strain 1021 is a biotin prototroph from which genetically defined biotin auxotrophs have been isolated. These auxotrophs do not grow in minimal media lacking biotin (59). (iii) The activity of acetyl-CoA carboxylase, a biotin-requiring enzyme, is maintained in R. etli during subculturing in MM lacking biotin (18), suggesting that sufficient biotin is being produced to maintain the enzyme in its active (biotinylated) form. (iv) Pimelic acid (10 µg/ml), the earliest known precursor in the biotin biosynthetic pathway (24), can substitute for biotin in allowing strain CE3 to grow in subcultures (14). This finding suggests that R. etli is capable of synthesizing biotin de novo. Collectively, these data indicate that R. etli CE3 is a biotin prototroph which regulates biotin synthesis (or the efficiency of its ligation to biotin-requiring carboxylases) during growth on succinate. We note parenthetically that thiamine, which can substitute for

Re	EARLENE CONTRACTOR CONTRAC	59
Rc	MFE-KILMANRGBIAIRVLRAANBLVI	26
Mt		55
SC	MSQRKFAGLRUNFNLLGEKNGLIVANRCEIPIRLFRTAHELSMQTVAISSHEDRGSTHAKADEAVVIGEV	71
пъ		69
Re	R-DLGPIESYLSIDEVIRVAKLSGADAIHPGYGLLSESPEFVDACNKAGIIFIGPKADTMRQLGNKVAARNLAISVGVPVVPATEPLPDDMAEVA	153
Mt	HPVHAYLSVDEIVATARRAGADAIYPGYGFLSENPDLAAACAAAGISFVGPSAEVLELAGNKSRAIAAAREAGLPVLMSS-APSASVDBLL	145
Sc	-GQYTP VGAY LAID EII SIAQKHQVDFIH PGYGFL SENSE FAD KVVKAGITWIGP PAEVID SVGDK VSARNLAAKANV PTV PGT PGPIETVEBAL	165
HS	LAP VQAY LHI PDII KVA KENNVD AVH PGYGFLS ERADFAQA CQDA GVRFIGPS PEVVRKMGDK VER RAIA I AAGV PVV PGT DAPITSLHEAH	181
	ATT hinding submotif	
Re	KMAAAIGYPYMLKASWGGGGGMRVIISEADLAKEVTBAKREAMAAFGKDEVYLEKLVERARHVESOILGDTHGNVVHLFERDČSVORRNOKVVE	248
Mt	SVA AGMPFPLFVK AVAGGGGGGGRGNRVGDIA AL PEAIEAASR BAESA FGDPTVYLEQAV INPRHIBVQ I LADNLGDV IH LYERDC SVQRRHQKVIE	240
Sc	DFVNEYGYPVIIK AAFGGGGRGMRVVREGDDVADAFQRATSE ARTA FGNGTCFVERFLDKPKHIEVQLLADNHGNVVHLFERDCSVQRRHQKVVE	260
Hs	EFSNTYGFPIIFKAAYGGGGGGGGWVVHSYEELEENYTRAYSEAWPAFGNGALFVEKFIEKPRHIEVQILGDQYGNILHLYERDCSIQRHQKVVE	276
	+P+ KA GGGGRGMR + + A EA AFG ++E+ + +H+E Q+L D G+++HL+ERDCS+QRR+QKV+E	
Re	RAFA DIGHT SALMULA SALM	343
Mt	LAPAPHLDAELRYKMCVDAVAFARHIGYSCAGTVEFLLDER-GEYVFIEMNPRVQVEHTVTEEITDVDLVASQLRIAAGETLEQLGLR-QEDI	331
Sc	VAPAKTLPREVRDAILTDAVKLAKECGYRNAGTABFLVDNQ-NRHYFIBINPRIQVBHTITBEITGIDIVAAQIQIAAGASLPQLGLF-QDKI	351
Hs	IA PAAHL DPQLRTRLTSDS VKLAKQVGYENAGTVBFLVDRH-GKHYFIEVNSRLQVBHTVTBEIT DVDLVHAQIHVAEGRSLPDLGLR-QENI	367
	APA L R + ++ A Y AGT E+L+D + FIE+N R+QVEHT+TE +T +D+V +Q+ + G G+ Q+ I	
Re	RLNGHALOCRVTTEDPEHNFIPDYGRITAYRSASGFGIRLDGGTSYSGATTTERVYDPLLVKUTAWADNDLEATSENDEALPEETEGUAUNLUEL	420
Mt	APHGAALQCRITTEDPANGFRPTRACSARCDPPAVPVSAWTAAPTWRRNOPVLRLHAGOADLSGRDLPTAVS RARRAIABFRIRGVSTNIPFL	424
Sc	TTRGFÅ I QCRITTED PAKN FOP DTGRIEVYRSAGGNGVRLDGGNAY AGTI ISPHY DSMLVKCSCSGSTYEIVRRKMIRALIEFRIRGVKT NIPFL	446
Hs	R IN GCALQC RVTTED PAPTFQP DTGRIEVFR SGEGMGIRLDNASAFQG AVISPHYDSLLVKVIAHGKDHPTAATKMSRALABFRVRGVKT NIAFL	462
	G A+QCR+TTEDP F P + + + RA+ EFR+RGV TN+ FL	
Re	BAIIGHPKFRDNSYTTRFIDTTPELF00VKRODRATKLLTYLADVTVNGHPBAKDRPKPLENAARPVVPYANGNGVKDGTKOLLD	523
Mt	QAVLDDPDFRAGRVTTSFIDERPQLLTARASADRGTKILNFLADVTVNNPYG-SRPSTIYPDDKLPDLDLRAAPPAGSKQRLV	506
Sc	LTLLTNP VFIEGTYWGTFIDDT PQLFQMVSSQNRAQKLLHYLADVADNG-SSIKGQIGLPKLKSNPSVPHLHDAQGNVINVTKSAPPSGWRQVLL	540
Hs	QNVLNNQQFLAGTVDTQFIDENPDVFQLRPAQNRAQKLLHYLGHVMVNG-PTTPIPVKASPSPTDPVVPAVPIGPPPAGFRDILL	546
	++ F FID P++ +K K+L +L V N +P G + L Durante binding methic	
Re	TLGPKKFGRWWRNEKEVILTDTTMEDGHOSLLATEWETYDIAEIAGTYSHALDNILSLEUWGGATEDUSMENTAEDTWERELALDEGADNILLOW	618
Mt	KLGPEGFARWLRESAAVGVTDTTFRDAHQSLLATRVRTSGLSRVAPYLARTMPOLLSVECWGGATYDVALRFLKEDPWERLATLRAAMPNILCIOM	601
Sc	EKG PAEFAROVROFNOTT, IMD TTWR DAHOSI, LATRUR THDIATTAPTAHAILAGAFAL ROWGATTOWARF, HED DWORLDKI, SUVD NIDEO M	COF
~~		635
Hs	REGPEGFARAVR NHPGLL LMDTTFR DAHQSLLATRVR THDLKKI APYVAHNFSKLFSMENWGGATFDVAM RFLYECPWRR LQELR ELIP NIRFQ M	641
Hs	REGPEGFARAVRNHPGLLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFSKLFSMENWGGATFDVAMRFLYECPWRRLQELRELIPNIRFQM GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM	641
Hs Re	REGPEGFARAVRNHPGLLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFSKLFSMENWGGATFDVAMRFLYECFWRRLQELRELIPNIRFQM GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVKYFVROAAKGGIDLFRVFDCLNWVENMRVSMDATAREN-KLCRAATCYTGDTLNSARPKYDLKYYTNLAVELRKAGA	635 641 712
Hs Re Mt	REGPEGFARAVR NHPGLLLMD TTFR DAHQSLLATRVR THDLKKI APYVAHNFSKLFSMENWGGATFDVAMRFLYECP WRR LQELR ELIP NIRFQ M GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICYTGDILNSARPKYDLKYYTNLAVELEKAGA LLRGRNTVGYTPY PEIVTSAFVQEATATGIDIF RIFDAL NNIESM RPAID AVRETGSAIAEVAMCYTGDLTDPGEQLYTLDY YLKLAEQIVDAGA	635 641 712 696
Hs Re Mt Sc	REGPEGFARAVR NHPGLL LMD TTFR DAHQSLLATRVR THDLKKI APYVAHNFSKLFSMENWGGATFDVAMR FLYECP WRR LQELR ELIP NIRFQ M GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIABEN-KLCEAAICYTGDILNSARPKYDLKYYTNLAVELEKAGA LLRGRNTVGYTPY PEIVTSAF VQEATATGIDIF RIFDAL NNIE SM RPAID AVRETGSAIAEVAMCYTGDLTDPGEQLYTLDY YLKLAEQIVDAGA LLRGANGVAY SSLPDNAIDHFVKQAKDNSVDIFRVFDAL NDLEQLKVGVDAVKKAG-GVVEATVCFSG DMLQPG-KKYNLDY YLEIAEKIVQMGT	641 712 696 728
Hs Re Mt Sc Hs	REG PEGFARAVR NHPGLL LMD TTFR DAH QSLLATRVR THELIKT APVVAHNFSKLFSMENN GGATFDVAM FLYECP WER LQELR ELIP NIRFQ M GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICYTGDLINSARPKYDLKYYTNLAVELEKAGA LLRGANGVGYTYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICYTGDLTDPGEQLYTLDY YLKL AEQIVDAGA LLRGANGVAY SSLPDNAIDHFVKQAKDNSVDIFRVFDAL NDLEQLKVGVDAVKKAG-GVVEATVCFSGDMLQPG-KKYNLDY YLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFRVFDSLNYLPNMLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT	635 641 712 696 728 735
Re Mt Sc Hs	REG PEGFARAVR NHPGLL LMD TTFR DAH QSLLATRVR THDLKKI APVVAHNFSKL FSMENN GGATFDVAM FLYECP WRR LQELR EL IP NIRFQ M GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVXYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICYTGDLINSARPKYDLKYYTNLAVELEKAGA LLRGANGVGYTYPPDNVVXYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICYTGDLTDPGEQLYTLDY YLKL AEQIVDAGA LLRGANGVAY SSLPDNAIDHFVKQ AKDNSVD IFRVFDAL NDLEQLKVGVD AVKRAG-GVVE ATVCFSG DMLQPG-KKYNLDY YLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFFCEVAKENGMDVFRVFDSL NYLPNMLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT LLRG N VGY P+ F A +D+FR+FD LN + + ++A + E + + GD+ Y L YY +A + G	635 641 712 696 728 735
Hs Re Mt Sc Hs Re	REG PEGFARAVR NHPGLL LMD TTFR DAH QSLLATRVR THDLKKI APVVAHNFSKL FSMENN GGATFDVAM FLYEC W RK LQELK BLI PARFQ M GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVXYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICYTGDLTDPGEQLYTLDY YLKL AEQIVDAGA LLRGANGVGYTYPP PEIVTSAF VQEATATGIDIF RIFDAL NNIE SM RPAID AVRETGSAIAEVAMCYTGDLTDPGEQLYTLDY YLKL AEQIVDAGA LLRGANGVAY SSLPDNAIDHFVKQ AKDNSVD IFRVFDAL NDLEQLKVGVD AVKKAG-GVVE ATVCFSG DMLQPG-KKYNLDY YLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFFCEVAKENGMDVFRVFDSL NYLPN MLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQY YMGL AEALVRAGT LLRG N VGY P+ F A +D+FR+FD LN + + ++A + E + + GD+ Y L YY +A + G HIIAVKDMAGLLKPAAAKVLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW	635 641 712 696 728 735 806
Hs Re Mt Sc Hs Re Mt	REG PEGFARAVR NHPGLL LMD TTFR DAH QSLLATRVR THDLKKI APVVAHNFSKLFSMENNGGATF DVAM FLYEC, WRR LQELE BLIP NIRFQ M GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICYTGDLTDPGEQLYTLDY YLKLAEQIVDAGA LLRGANGVGYTPY PEIVTSAF VQEATATGIDIF RIF DAL NNIE SM RPAID AVRETGSAIAEVAMCYTGDLTDPGEQLYTLDY YLKLAEQIVDAGA LLRGANGVAY SSLP DNAIDHF VKQ AKDNSVD IFRVFDAL NDLEQLKVGVD AVKKAG-GVVE ATVCFSG DMLQPG-KKYNLDY YLEIAEKIVQMGT LLRGANAVGYTNYP DNVVFKFCEVAKENGMDVFRVFDSL NYLPN MLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQY YMGLAEALVRAGT LLRG N VGY P+ F A +D+FR+FD LN + + ++A + E + + GD+ Y L YY +A + G HIIAVKDMAGLLKPAAAKVLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDMAGLLKPAAAKVLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALAGSERDPGLDPAW	635 641 712 696 728 735 806 790
Hs Re Mt Sc Hs Re Mt Sc	REGPEGFARAVR NHPGLL LMD TTFR DAH QSLLATRVR THDLKKI APVVAHNFSKLFSMENNGGATF DVAM FLYEC, WR LQELE ELIP NIRFQ M GP F +R + DTT RD HQSLLATR+RT + +A + +E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENMRVSMDAIAEEN-KLCEAAICYTGDLINSARPKYDLKYYTNLAVELEKAGA LLRGANGVGYTPYPEIVTSAF VQEATATGIDIF RIF DAL NNIE SM RPAID AVRETGSAIAEVAMCYTGDLTDPGEQLYTLDY YLKLAEQIVDAGA LLRGANGVAY SSLPDNAIDHF VKQ AKDNSVD IFRVFDAL MDLEQLKVGVD AVKRAG-GVVE ATVCFSG DMLQPG-KKYNLDY YLKLAEQIVDAGA LLRGANAVGYTNYPDNVVFFCEVAKENGMDVFRVFDSL NYLPNMLLGMEAAGSAG-GVVEATVCFSG DMLQPG-KKYNLDY YLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFFCEVAKENGMDVFRVFDSL NYLPNMLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQY YMGLAEALVRAGT LLRG N VGY P+ F A +D+FR+FD LN + + ++A + E + + GD+ Y L YY +A + G HIIAVKDMAGLLKPAAAKVLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDMAGLLKPAAAKLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALAGSERDPGLDPAW HVLAIKDMAGTMKPAAAKLLIGSL RAKYPDLPIHVHTHDSAGTRVSMTACALAGADV DVAINSMSGLTSQPSINALLASLEGN-IDTGINVEH	535 641 712 696 728 735 806 790 822
Hs Re Mt Sc Hs Re Mt Sc Hs	REG PEGFARAVR NHPGLL LMD TTFR DAH QSLLATRVR THDLKKI APVVAHNFSKL FSMENNGGATF DVAM FLVEC KUR LQELF BLIP NIRFQ M GP F +R + DTT RD HQSLLATRVR THDLKKI APVVAHNFSKL FSMENNGGATF DVAM FLVEC ELIP NIRFQ M GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICYTGDLTDPGEQLYTLDY LKLAEQIVDAGA LLRGANGVAY SSLPDNAIDHF VKQAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICYTGDLTDPGEQLYTLDY LKLAEQIVDAGA LLRGANGVAY SSLPDNAIDHF VKQAKDNSVD IFRVFDAL MDLEQLKVGVDAVKRAG-GVVEATVCFSGDMLQPG-KKYNLDY YLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFFCEVAKENGMDVFRVFDSL NYLPNMLLGMEAAGSAG-GVVEATVCFSGDMLQPG-KKYNLDY YLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFFCEVAKENGMDVFRVFDSL NYLPNMLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQY YMGLAEALVRAGT LLRG N VGY P+ F A +D+FR+FD LN + + ++A + E + + GD+ Y L YY +A + G HIIAVKDMAGLLKPAAAKVLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDMAGLLKPAAAKLJFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALASLEGN-IDTG INVEH HLGIKDMAGTMKPAAAKLLIGSL RAKYPDLPIHVHTHDSAGTRVASMTACALAGADVV DVAINSMSGLTSQPSIMALLASLEGN-IDTG INVEH HLLCIKDMAGLLKPTACTMLVSSL RDRFPDLPLH HTHH G A A A A A A D A GD VV DVAADSMGATSQPSMGALVACTRGTPLDTEVPMER H+++ +KDMG4 ++P A L ++R	835 641 712 696 728 735 806 790 822 830
Hs Re Mt Sc Hs Re Mt Sc Hs	REGPEGFARAUR NHPGLL LMD TTFR DAH QSLLATRUR THDLKKI APVUAHNFSKL FSMENNGGATF DUAM FLYEC, WUR LQELER ELIP NIRFQ M GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENMRVSMDAIAEEN-KLCEAAICYTGDLTDPGEQLYTLDY LKLAEQIVDAGA LLRGANGVAY SSLPDNAIDHF VKQAKDSVD IFRVFDAL MDLEQLKVGVDAVKRAG-GVVEATVCFSGDMLQPG-KKYNLDY YLKLAEQIVDAGA LLRGANGVAY SSLPDNAIDHF VKQAKDNSVD IFRVFDAL MDLEQLKVGVDAVKRAG-GVVEATVCFSGDMLQPG-KKYNLDY YLKLAEQIVDAGA LLRGANAVGYTNYPDNVVFFCEVAKENGMDVFRVFDSL NYLPNMLLGMEAAGSAG-GVVEATVCFSGDMLQPG-KKYNLDY YLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFFCEVAKENGMDVFRVFDSL NYLPNMLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQY YMGLAEALVRAGT LLRG N VGY P+ F A +D+FR+FD LN + + ++A + E + + GD+ Y L YY +A + G HIIAVKDMAGLLKPAAAKVLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDMAGLLKPAAAKLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALASLEGN-IDTG INVEH HLCIKDMAGLLKPTACTMLVSSL RDRFPDL PLH IHTHAPSGAGVAAMLACAQAGADVV DVAADSMSGMTSQPSMGALVACTRGTPLDTEVPMER H++ +KDMAG ++P A L +LR LP+H HTH G A A AG D VD A ++G TSQP + +++ D +	835 641 712 696 728 735 806 790 822 830
Hs Re Mt Sc Hs Re Mt Sc Hs Re	REGPEGFARAVR NHPGLL LMD TTFRDAHQSLLATRVR THDLKKI APVVAHNFSKLFSMENNGGATFDVAM FLYECP KUR LQELF BLIP NIRFQ M GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENMRVSMDAIAEEN-KLCEAAICYTGDLINSARPKYDLKYTNLAVELEKAGA LLRGANGVAYSSLPDNAIDHFVKQAKGGIDLFRVFDCLNWVENMRVSMDAIAEEN-KLCEAAICYTGDLTDPGEQLYTLDY YLKLAEQIVDAGA LLRGANGVAYSSLPDNAIDHFVKQAKGGIDLFRVFDCLNWVENMRVSMDAIAEEN-KLCEAAICYTGDLTDPGEQLYTLDY YLKLAEQIVDAGA LLRGANGVAYSSLPDNAIDHFVKQAKGDNSVDIFRVFDALMDLEQLKVGVDAVKKAG-GVVEATVCFSGDMLQPG-KKYNLDYYLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFRVFDSLNYLPNMLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT LLRG N VGY P+ F A +D+FR+FD LN + + ++A + E + + GD+ Y L YY +A + G HIIAVKDMAGLLKPAAAKVLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDMAGLLKPAAAKLLIGSL RAKYPDLPIHVHTHDSAGTRVASMTACALAGADVVDVAANDALSGNTSQPSMGALVACTRGTPLDTEVPMER HILCIKDMAGLLKPTACTMLVSSL RDFFDL PLH IHTHAPSGAGVAAMLACQQ GADVV DVAADSMS GMT SQPSMGALVACTRGTPLDTEVPMER H++ +KDMAG ++P A L +LR LP+H HTH G A A AG D VD A ++G TSQP + +++ D + IRRISFYWEAVRNQYAAFESDLKGPASEVYLHEMPGGQFTNLKEQARSLGLETRWHQVAQAYADANQMFGDIVKVTPSSKVVGDMALMNVSQD	635 641 712 696 728 735 806 790 822 830 899
Hs Re Mt Sc Hs Re Mt Sc Hs Re Mt	REG PEGFARAVR NHPGLL LMD TTFR DAH QSLLATRVR THDLKKI APVVAHNFSKLFSMENNGGATF DVAM FLVEC. BLIP NLRFQ M GP F +R + DTT RD HQSLLATRYR T + +A + +E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEN-KLCEAAICYTGDLINSARPKYDLKYTLAVELEKAGA LLRGRNTVGYTPY PEIVTSAF VQEATATGIDIF RIF DAL NNIE SM RPAID AVRETGSAIAEVAMCYTGDLTDPGEQLYTLDY YLKLAEQIVDA GA LLRGANGVAYSSLP DNAIDHF VKQ AKDNSVD IFRVFDAL MDLEQLKVGVD AVKKAG-GVVE ATVCFSO DMLQPG-KKYNLDYYLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFRVFDSL NYLPN MLLGMEAAGSAG-GVVEATVCFSO DMLQPG-KKYNLDYYLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFRVFDSL NYLPN MLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT LLRG N VGY P+ F A +D+FR+FD LN + + ++A + E + + GD+ Y L YY +A + G HIIAVKDMAGLLKPAAAKVLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDMAGLLKPAAAKLLIGSL RAKYPDLPIHVHTHDSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDFGLDPAW HVLAIKDMAGLLKPAAAKLLIGSL RAKYPDLPIHVHTHDSGGAATRVASMTACALAGADV DVAINSMSGLTSQPSINALLASLEGN-IDTGINVEH H LLCIKDMAGLKPTACTMLVSSL RDRFPDL PLH IHTHAPSGAGVAAML ACAQAGADVV DVAINSMSGTSQPSMGALVACTRGTPLDTEVPMER H++ +KDMAG ++P A L +LR LP+H HTH G A A AG D VD A ++G TSQP + +++ D + IRRISFYWEAVRNQYAAFESDLKGPASEVYLHEMPGGQFTNLKEQARSLGLETRWHQVAQAYADANQMFGDIVKVTPSSKVVGDMALMMVSQD VCALEPYWEALRKVYAPFESGLFGPTGRV YHHEIPGGQLSN LRQQAIALGCGRFEEIEEAYAGADRVLGRLVKVTPTSKVVGDLALALVGAG	635 641 712 696 728 735 806 790 822 830 899 883
Re Mt Sc Hs Re Mt Sc Hs Re Mt Sc	REG PEGFARAUR NHPGLL LMD TTFR DAH QSLLATRUR THDLKKI APVUAHNFSKLFSMENNGGATF DUAMR FLYECP WRR LQELR ELIP NIRFQ M GP F +R + DTT RD HQSLLATR+RT + +A + +E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICYTGDLTDPGEQLYTLDY LKLAEQIVDAGA LLRGRNTVGYTPY PEIVTSAF VQEATATGIDIF RIF DAL NNIE SM RPAID AVRETGSAIAE VAMCYTGDLTDPGEQLYTLDY YLKLAEQIVDAGA LLRGANGVAYSSLP DNAIDHF VKQ AKDNSVD IFRVFDAL NDLEQLKVGVD AVKKAG-GVVE ATVCFSC DMLQPGKKYNLDY YLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFRVFDSL NYLPNMLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQY YMGLAEALVRAGT LLRGN VGY P+ F A +D+FR+FD LN + + ++A + E + + GD+ Y L YY +A + G HIIAVKDMAGLLKPAAAKVLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDMAGLLKPAAAKLLIGSL RAKYPDLPIHVHTHDSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDMAGLKPTACTMLVSSL RDRFPDL PLH HTHDFGGQLASYVA AWHA GADAVDCAAAPLAGTTSQPALSSIVAAAHTEYDTG LSLSA HILGIKDMAGLKPTACTMLVSSL RDRFPDL PLH HTHDFGGQLASYVA AWHA GADAVDCAAAPLAGTTSQPSINALLASLEGN-IDTG INVEH HILCIKDMAGLKFTACTMLVSSL RDRFPDL PLH HTHAPSGAGVAAML ACAQA GADVV DVAADSMS GMTSQPSMGALVACTRGTPLDTEVPMER H++ +KDMAG ++P A L +LR LP+H HTH G A A AG D VD A ++G TSQP + +++ D + IRRISFYWEAVRQYAAFESDLKGPASEVYLHEMPGGQFTNLKEQARSLGLETRWHQVAQAYADANQMFGDIVKVTPSSKVVGDMALMMVSQD VCALEPYWE ALRKVY APFE SGLEGP TGRV YHH EIPGGQLSN LQQ AIALGLGDR FEEIEEAYAGADRVLGRLVKVTPTSKVVGDLALALVGAG VRELDAYWE ALRKVY APFE SGLEGP DEVY OH EIPGGQLSN LQQ AIALGLGDR FEEIEEAYAGADRVLGRLVKVTPTSKVVGDLA LALVGAG VRELDAYWE ALRKVY APFE SGLEGP DEVY DH EIPGGQLSN LQQ AIALGLGDR FEEIEEAYAGADRVLGRLVKVTPTSKVVGDLA KEMVSND	835 641 712 696 728 735 806 790 822 830 822 830 899 883 9155
Re Mt Sc Hs Re Mt Sc Hs Re Mt Sc Hs	REO PEGFARAUR NHPGLL LMD TTFR DAH QSLLATRUR THDLKKI APVUAHNFSKLFSMENNGGATF DUAMR FLIVEC WER LQELR ELIP NIRFQ M GP F +R + DTT RD HQSLLATRYR THDLKKI APVUAHNFSKLFSMENNGGATF DUAMR FLIVEC WER LQELR ELIP NIRFQ M GP F +R + DTT RD HQSLLATRYR T + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYP DNVVKYFVRQAAKGGID LFRVFDCLNWVENNRVSMDAIAEN-KLCEAAICYTGDLINSARPKYD LKYTNLAVELEKAGA LLRGRNTVGYTPY PEIVTSAF VQEATATGID IF RIF DAL NNIE SM RPAID AVRETGSAIAE VAMCYTGDLIDPGEQLYTL DY YLKLAEQIVDA GA LLRGANGVAYSSLP DNAIDHF VKQ AKDNSVD IF RVF DAL MDLE QLKVGVD AVKKAG-GVVE ATVCFSO DMLQPGKKYNL DY YLEIAEKIVQMGT LLRGANAVGYTNYP DNVVFKFCEVAKENGMD VFRVFDSL NYLPN MLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQY YMGLAEALVRA GT LLRG N VGY P+ F A +D+FR+FD LN + + ++A + E + + GD+ Y L YY +A + G HIIAVKDMAGLLKPAAAKVLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDMAGLLKPAAAKLLIGSL RAKYPDLPIHVHTHDSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDMAGLLKPAAAKLLIGSL RAKYPDLPIHVHTHDSGGAFRVASMTACALAGADV DVAINSMSGLTSQPSINALLASLEGN-IDTGINVEH H LCIKDMAGLKPTACTMLVSSL RDRFPDL PLH IHTHAPS GAGVAAML ACAQAGADVV DVAINSMS GLTSQPSMGALVACTRGTPLDTEVPMER H++ +KDMAG ++P A L +LR LP+H HTH G A A AG D VD A ++G TSQP + +++ D + IRRISFYWEAVRNQYAAFESDLKGPASEVYLHEMPGGQFTNLKEQARSLGLETRWHQVAQAYADANQMFGDIVKVTPSSKVVGDMALKMVSQD VCALEPYWE ALRKVY APFESGLEGP ASEVYLHEMPGGQFTNLKEQARSLGLETRWHQVAQAYADANQMFGDIVKVTPSSKVVGDMALKMVSQD VCALEPYWE ALRKVY APFESGLEGP DPEVYOH EIPGGQLTNLFQ AQUG GLGGWAQTKRAYREANVLLGDIVKVTPTSKVVGDLALALVGAG VRELDAYWEGARGLYAAFDCTATMKSGNSDV YH EIPGGQLTNLHFQ AQUG GLGSKFKEVKKAYLEANVLLGDIVKVTPSSKVVGDLALALVGAG VRELDAYWEGARGLYAAFDCTATMKSGNSDV YENE IPGGQLTNLHFQ AQUG GLGSKFKEVKKAYLEANVLLGDIVKVTPSSKVVGDLAKFMVSNK VFDYSEYWEGARGLYAAFDCTATMKSGNSDV YENE IPGGQLTNLHFQ AQUG GLGSKFKEVKKAYLEANVLLADIVKTPSSKVVGDLAKFMVSNK VFDYSEYWEGARGLYAAFDCTATMKSGNSDV YENE IPGGQUTNLHFQ AQUG GLGSKFKEVKKAYLEANVLLADIVKVTPSSKVVGDLAQFM VNK	835 641 712 696 728 735 806 790 822 830 822 830 899 883 915 925
HS ReMtSC HS RMtSC HS RMtC HS RMtC SHS	REGPEGFARAUR NHPGLL LMD TTFR DAH QSLLATRUR THDLKKI APVUAHNFSKLFSMENNGGATFDVAM FLYECH WER LJELRE LIP NIRFQ M GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEN-KLCEAAICYTGDLINSARPKYDLKYTNLAVELEKAGA LLRGRNTVGYTPY PEIVTSAFVQEATATGIDIFRIFDAL NDLE SM RPAID AVRETGSAIAEVAMCYTGDLIDPGEQLYTLDY YLKLAEQIVDAGA LLRGANGVAYSSLPDNAIDHFVKQAKGGIDLFRVFDCLNWVENNRVSMDAIAEN-KLCEAAICYTGDLIDPGEQLYTLDY YLKLAEQIVDAGA LLRGANGVAYSSLPDNAIDHFVKQAKGNSVDIFRVFDAL NDLEQLKVGVDAVKKAG-GVVEATVCFSGDMLQPGKKYNLDYYLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFRVFDSLNYLPNNLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT LLRGN VGY P+ F A +D+FR+FD LN + + ++A + E + + GD+ Y L YY +A + G HIIAVKDMAGLLKPAAAKVLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDMAGLLKPAAAKKLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDMAGLLKPAAAKLLIGSL RAKYPDLPIHVHTHDSGGATRVASMTACALAGADVVDVAAINSM GLTSQPSINALLASLEGN-IDTGINVEH HILGIKDMAGTMKPAAAKLLIGSL RAKYPDLPIHVHTHDSGGFTNLKEQARSLGLETRWHQVAQAYADANQMFGDIVKVTPSSKVVGDMALMMVSQD VCALEPYWEALRKVYAPFESGLKGPASEVYLHEMPGGQFTNLKEQARSLGLETRWHQVAQAYADANQMFGDIVKVTPSSKVVGDMALMMVSQD VCALEPYWEALRKVYAPFESGLKGPDPEVYHEIPGGQLSNLRQQ AIALGLGDR FEEIEEAYAGADRVLGDIVKVTPSSKVVGDMALMMVSQD VCALEPYWEALRKVYAPFESGLKGPDPEVYHEIPGGQLSNLRQQ AIALGLGDR FEEIEEAYAGADRVLGDIVKVTPSSKVVGDMALMMVSQD VCALEPYWEALRKVYAPFESGLKGPDPEVYHEIPGGQLSNLRQQ AIALGLGSKFKEVKKAYVEANQMLGDLIKVTPSSKVVGDLAKFMVSNK VFDYSEYWEGARGLYAAFDCTATMKSGNSDVYENEIPGGQNTNLHFQAHSMGLGSKFKEVKKAYVEANQMLGDLIKVTPSSKVVGDLAKFMVSNK + YW R Y+ F+ VY +E+PGGQ NL QA +GL ++ + AY A+ + G ++KVTP SK+VGD+A +V	835 641 712 696 728 735 806 790 822 830 822 830 899 883 915 925
HS Re Mtc HS KS Re Mtc Scs Re Mtc Scs Re	REG PEGFARAVR NHPCLL LMD TTFR DAHQ S LLATAVR THDLKKI APVVAHNFSKLFSMENWGGATT DVAM R LYBC P WRR LQELR ELIP NIRFQ M GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DV+HRFL E PW RL +R PN+QM LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENMRVSMDAIAEEN-KLCEAAICYTGDILNSARPKYDLKYYTNLAVELEKAGA LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENMRVSMDAIAEEN-KLCEAAICYTGDILNSARPKYDLKYYTNLAVELEKAGA LLRGANGVGYTYP PELVTSAF VQEATATG I DI F RIF DAL NNIE SM RPAID AVRE TGSATAE VAMCYTG DLTDPGEQLITTL DY YLKL AEQIVDA GA LLRGANGVGYTSSLPDNAIDHF VKQAKDNSVD I FRVFDAL NDLEQLKVGVD AVKKAG-GVVEATVCFSG DMLQPG-KKYNLDY YLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFKFCEVAKENOMDVFRVFDSLNVLPN MLLGM EAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGL AEALVRA GT LLRG N VGY P+ F A +D+FR+FD LN + +++A + E + + GD+ Y L YY +A + G HIIAVKDMAGLLKPTAAAKVLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDMAGLLKPTAAKKLFKALREAT-GLPIHFHTHDTFGQLASYVA AWHA GADAVDGAAAPLAGTTSQPALSSIVAAAAHTEYDTG LSLSA HILGIKDMAGLLKPTACTMLVSSL RSRF-DL PVHLHTHDTPGQLASYVA AWHA GADAVDA NDALSGNTSQPSINALLASLEGN-IDTG INVEH HLCIKDMAGLLKPTACTMLVSSL RSRF-DL PVHLHTHDFSGAGVAAMLACAQA GADVV DVA INSMS GLTSQPSINALLASLEGN-IDTG INVEH HLCIKDMAGLKPTACTMLVSSL RAKYPDLPIH HTHDFSGAGVAAMLACAQA GADVV DVA NASMS GMT SQPSMGALVACTRGTPLDTEVPMER H++ +KDMAG ++P A L +LR LP+H HTH G A A AG D VD A ++G TSQP + +++ D + IRRISFYWEAVRNQYAAFESDLKGPASEVYLHEMPGGQFTNLKEQARSLGLETRWHQVAQAYADANQMFGDIVKVTPTSKVVGDLA LALVGAG VCALEPYW BALRKVY APFESGLKGPASEVYLHEMPGGQFTNLKEQARSLGLETRWHQVAQAYADANQMFGDIVKVTPTSKVVGDLA LALVGAG VRELDAY WAEMRLLYSCF EADLKGPDPE VQH BIPGGQLSN LRQQ AIALG LGSKFKEVKKA YVEANQMLG DLIKVTPTSKVVGDLA KFMVSNK VFDYSEYW BGARGLXAAFDCTATMKSGNSDV YENE IPGGQUNL QA +GL ++ + AY A++ G ++KVTP SK+VGD+A +V LTVADVVSPDREVSFPESVVSMLKGDLGQPPSGWPEALQKKALKGEKP-YTVRPGSLLKEADLBAERKVIEKKLEREVSDFEFASYLMYPKVFTD	835 641 712 696 728 735 806 790 822 830 822 830 829 833 915 925 925
HS ReMtcs HS ReMtcs HS ReMtcs HS Ref KS Ref KS Ref KS	REG PEGFARAVR NHPGLL LMD TTFR DANQS LLATRVR THDLKKI A PYVAHNFSKLFSMENW GGAT FDVAM RF LVBC P WRL LQELR BLIP NIFFQ M GP F +R + DTT RD HQSLLATRVR THDLKKI A PYVAHNFSKLFSMENW GGAT FDVAM RF LVBC P WRL LQELR BLIP NIFFQ M GP F +R + DTT RD HQSLLATRVR THDLKKI A PYVAHNFSKLFSMENW GGAT FDVAM RF LVBC P WRL LQELR BLIP NIFFQ M GP F +R + DTT RD HQSLLATRVR THDLKKI A PYVAHNFSKLFSMENW GGAT FDVAM RF LVBC P WRL QELR BLIP NIFFQ M GP F +R + DTT RD HQSLLATRVR THDLKKI A PYVAHNFSKLFSMENW GGAT FDVAM RF LVBC P WRL QELR BLIP NIFFQ M GP F +R + DTT RD HQSLLATRVR THDLKKI A PYVAHNFSKLFSMENW GGAT FDVAM RF LVBC P WRL QELR BLIP NIFFQ M GP F +R + DTT RD HQSLLATRVR THDLKKI A PYVAHNFSKLFSMENW GGAT FDVAM RF LVBC P WRL QELR BLIP NIFFQ M GP F +R + DTT RD HQSLLATRVR THDLKKI A PYVAHNFSKLFSMENW GGAT FDVAM RF LVBC P WRL QELR BLIP NIFFQ M GP F +R + DTT RD HQSLLATRVR THDLKKI A PYVAHNFSKLFSMENG GAT FDVAM RF LVBC P WRL QELR BLIP NIFFQ M LLRGANAVGYTNYPDNVVFKJEVGAKNEGGID FRVFDCLNWVENNRVSMDAIABEN-KLCBAAICYTGDLTDPGEQLYTDY LAVELEKAGA LLRGANAVGYTNYPDNVVFKJEVGAKNENG D FRVFDALNIES SM RPAID AVRE TGSAIAEVAMCYTGDLVE A TVCFSG DML QPG-KK YNLDY YLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFKJEVGAKNENG DVFVFDSLNYLPNMLLGMEAAGSAG-GVVBAAISYTGDVADPSRTK YSLQY YMGLAEALVRA GT LLRG N VGY P+ F A +D+FR+FD LN + + ++A + E + + GD+ Y L YY +A + G HILAKDMAGLLKPAAKVLJKALRBAT-GLPIH FHTHDTSGIAAATVLAAVEAGVAAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDMAGLLKPAAKVLJKALRBAT-GLPIH FHTHDTSGIAAATVLAAVEAGVAAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDMAGLLKPTACTMLVSSL RDRFDLP H HTHDSGGAVAAMLAACAQA GADV DAADNSS GMTSQPSMGALVACTRGTPLDTEVPMER HILCIKDMAGLLKPTACTMLVSSL RDRFPDL PLH HTHDSGGAVAAMLACAQA GADV DVAADSNS GMTSQPSMGALVACTRGTPLDTEVPMER H++ +KDMAG ++P A L +LR LP+H HTH G A A AG D VD A ++G TSQP + +++ D + IRRISFYWEAVRNQVAAFESDLKGPSEVYLHEMPGGQFTNLKEQASLGLEETWHQVQAQAYADANQMFGDIVKVTPTSKVVGDLAKFWYSNK VFDYSEYMEARRLLYSCF BADLKGPDEPVYQH EIPGGQLTNLLFQ AQUGLGEGWAQ TKRA YREANVLLGDIVKVTPTSKVVGDLAKFWYSNK VFDYSEYMEGARGLYAAFDCTATMKSGNSDV YENE IPGGQLTNLLFQ AQUGLGEGWAQ TKRA YEAN NULLGDLIKVTPSKVVGDLAKFWYSNK V	635 641 712 696 728 735 806 790 822 830 822 830 829 833 915 925 925 993 993
HS Retcs Retcs Retcs Retcs Retcs Retcs	REG PEGFARAVR.NHPGLL LMD TFFR DARQS LLATRVR THDLKKI APVVAHNFSKLFSMENWGGATFDVAMR FLYECFWRR LOELR ELIP NIFFQ M GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICYTGDILNSARPKYDLKYYTNLAVELEKAGA LLRGRNTVGYTPY DNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICYTGDILNSARPKYDLKYYTNLAVELEKAGA LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICYTGDLTDPGEQLYTLDYYLKLAEQIVDAGA LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICYTGDLTDPGEQLYTLDYYLKLAEQIVDAGA LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICYTGDLTDPGEQLYTLDYYLKLAEQIVDAGA LLRGANGVGYTNYPDNVVFKFCEVAKENGMDVFVFDSLNYLPNMLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT LLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFVFDSLNYLPNMLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT LLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFVFDSLNYLPNMLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT LLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFVFDSLNYLPNMLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT LLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFVFDSLNYLPNMLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT LLRGANAGYNFPAACRLVSLTKALREAT-GLPIHFHTHDTSGIAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW HVLAIKDNAGLKPPAAORLVSALRSRF-DLFVHLHFHDTFGGQLASYVAAWHA GADAVDGAAAPLAGTTSQPALSSIVAAAAHTEYDTGLSLSA HILGIKDNAGLKPPAAORLVSSL RDRFPDL PLH.HTHDTPSGAGVAAMLACAQA GADVVDVAINSMSGLTSQPSMGALVACTRGTPLDTEVPMER H++ +KDMAG ++P A L +LR LP+H HTH G A A AG D VD A ++G TSQP + +++ D + IRRISFYWEAVRNQYAAFESDLKGPASEVYLHENPGGQFTNLKEQARSLGLETRWHQVAQAYADANQNFGDIVKVTPTSKVVGDLAALALWVSQD VCALEPYW EALAKVY AFFE SGLFGPGPV YQH EIPGGQLSN LRQQALALGLGGR FEEIEEAYAGADRVLGRLVKVTPTSKVVGDLAALALWVSQD VCALEDAYWAEMRLLSCF BADLKGPDPEVYQH EIPGGQLSN LRQQALGLGEWAQTKRA YREANYLLGDIVKVTPTSKVVGDLAALALWVSQD VARLEDAYAEMRLLSCF BADLKGPDPEVYQH EIPGGQLSN LRQQALGLGEWAQTKRA YREANYLLGDIVKVTPTSKVVGDLAALALWVSQD VARLEDAYAEMRLLSCF BADLKGPDPEVYQH EIPGGQLSN LRQAALALGLGBR FEEIEENAAGADRUG DLIKVTPTSKVVGDLAALALWONG + YW R Y+ F+ VY +E+PGGQ NL QA +GL +	635 641 712 696 728 735 806 790 822 830 822 830 829 883 915 925 925 993 968 1009
HS RetSchs Rtchs Rtchs Rtchs Rtchs Rtchs Rtchs Rtchs Rtchs Rtchs Rtchs	REG PEGFARAUR NHEGLL LADT TFR DANGS LLATRUR THD LKNI APVUAHNESKLESMENNGGATF DUAMR FLUEC WRR LQELE ELIP NIFFO M GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DV+RRFLUEC WRR LQELE ELIP NIFFO M LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIABEN-KLEBAAICYTGDILNSARPKYDLKYYTNLAVELEKAGA LLRGRNTVGYPP PELVTSAFVQEATATGIDIF RIF DAL NNIESM RPAID AVRETGSAIABVAMCYTGDLTDPGEQLYTLDY YLKLAEQIVDAGA LLRGANGVGYTNYPDNVVKYFVRQAKGOIDLFRVFDCLNWVENNRVSMDAIABEN-KLEBAAICYTGDILNSARPKYDLKYYTNLAVELEKAGA LLRGANGVGYTNYPDNVVKYFVRQAKGOIDLFRVFDAL NDLEQLKVGVDAVKKAG-GVVEATVCFSGDMLQPG-KKYNLDYYLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFKICEVAKENGMDVFRVFDAL NDLEQLKVGVDAVKKAG-GVVEATVCFSGDMLQPG-KKYNLDYYLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFKICEVAKENGMDVFRVFDSLNYLPMLLGMEAAGSAG-GVVEAAISYTGDVDDPSRTKYSLQYYMGLAEALVRAGT LLRGNNAGLLKPAAKVLFKALRBAT-GLPIHFHTHDTSGIAAATVLAAVBAGVDAVDAANDALSGNTSQFCLGSIVBALSGSERDFGLDPAW HVLAIKDMAGLLKPAAKVLFKALRBAT-GLPIHFHTHDTSGIAAATVLAAVBAGVDAVDAANDALSGNTSQPCLGSIVBALSGSERDFGLDPAW HVLAIKDMAGLLKPAAKKLJIGSL RAKYPDLPIWHTHDSAGIAAATVLAAVBAGVDAVDAANDALSGNTSQPCLGSIVBALSGSERDFGLDPAW HVLAIKDMAGLLKPAAAKLIGSL RDRFPDLPIHFHTHDTSGIAAATVLAAVBAGVDAVDAANDALSGNTSQPCLGSIVBALSGSERDFGLDPAW HVLAIKDMAGLLKPAAAKULJGSL RDRFPDLPIHHTHDFGGQLASYVAAWHA GADVVDVAINSMSGMTSQPSMGALVACTRGTPLDTEVPMER HLICIKDMAGLKFTACTMLVSSL RDRFPDLPIHHTHTBSGGQTNLKEQAANLACAQAGADVVDVAINSMSGMTSQPSMGALVACTRGTPLDTEVPMER HL++KDMAG ++P A L +LR LP+H HTH G A A AG D VDA ++G TSQP + ++++ D + IKRISFYWBAVRNQYAAFBSDLKGPASEVYLHEMPGGQFTNLKEQASLGLEFRWHQVAQAYADANQMFGDIVKVTPTSKVVGDLALMVSQD VCALEPYWBALRKVY APFESGLKGPDEVYQB BIPGGQLSN LRQQ AIALGLGDRFEEIEEAYAGADRVLGRLVKVTPTSKVVGDLALMVSQD VCALEPYWBALRKLYSCF BADLKGPDEVYQB BIPGGQLSN LKQQ AIALGLGBCKKEKKKAYDENAQMLGDLIKVTPTSKVVGDLALMVSQD VCALEPYWBALRKLYSCF BADLKGPDEVYQB BIPGGQLSN LKQA AIGLGSKFKEVKKAYDEANQULGDLIKVTPTSKVVGDLALMVSQD VCALEPYWBALRKLSSCF BADLKGP-OPEVYGB BIPGGQLSN LKQA AIGLGSKFKEVKKAYDEANQULGDLIKVTPTSKVVGDLALFVVSNK VFDYSEYBBASKSFDSVVSHLKGDLGOPPSGNPEALQKKALKGERP-YTVRPGSLLKBADLDABEKKVEKKYDESAVGDALANVSQD VCALEPYWBALRKLYAFFFELIGD PGGAP EPFRSVLKL	635 641 712 696 728 735 806 790 822 830 822 830 822 830 829 823 915 925 925 993 968 1009 1019
HS RetSchs Rtchs Rtchs Rtchs Rtchs Rtchs Rtchs Rtchs Rtchs Rtchs Rtchs	REG PEGFARAVR NIPGGL LMD TFFR DAHQS LLATRVR THD LKNI APVVARMFSK FSMENWGGATF DVAMRF LYECPWR LQELR ELIP NIR QM GP F +R + DTT RD HQS LLATRVR THD LKNI APVVARMFSK FSMENWGGATF DVAMRF LYECPWR LQELR ELIP NIR QM LLRGANGVGYTNYPDNVVKYFVRQAAKGGID LFRVFDCLNWVENNRVSMDAIABEN-KLCEAAICYTGD LNSARPKYDLKYYTNLAVELEKAGA LLRGRMTVG TPPY PELVTSAF VQEATAGGID JFRIFDALMNIESM RPAID AVRETGSAIAE VAMCYTGD LTDPGEQLYTL DYYLK LAEQIVDAGA LLRGANGVGTNYPDNVVKYFVRQAAKGGID LFRVFDCLNWVENNRVSMDAIABEN-KLCEAAICYTGD LTDPGEQLYTL DYYLK LAEQIVDAGA LLRGANGVGTNYPDNVVKYFVRQAKDNSVD JFRVFDALNDLEQLKVGVDAVKKAG-GVVE ATVC FSG DMLQPG-KKNNLDYYLKIAVELEKAGA LLRGANGVGTNYPDNVVFKGEVAKENGMDVFFVD DSL NYLPMLLGM EAAGSAG-GVVE ATVC FSG DMLQPG-KKNNLDYYLKIAVELEKAGA LLRGANGVGTNYPDNVVFKGEVAKENGMDVFFVD DSL NYLPMLLGM EAAGSAG-GVVE ATVC FSG DMLQPG-KKNNLDYYLKIAVELEKAGA LLRGANAVGTNYPDNVVFKGEVAKENGMDVFFVD DSL NYLPMLLGM EAAGSAG-GVVE ATVC FSG DMLQPG-KKNNLDY YLEIAEKIVQMGT LLRGANAVGTNYPDNVVFKGEVAKENGMDVFFVD DSL NYLPMLLGM EAAGSAG-GVVE ATSG FOLDGSIVEALSGERDPGLDPAK HVLAIKDMAGLLFPAAQRLVSALREAT-GLPIHPHTHDTSGIAAATYLAAVBAGVDAVDAANDALSGNTSQPCLGSIVEALSGEERDPGLDPAW HVLAIKDMAGLLFPAAQRLVSALRSRF-DLPVHLHTHDTSGIAAATYLAAVBAGVDAVDAANDASS GMT SQPSIGALVACTRGTPLDTEVPMER HILCIKDMAGLKFTACTMLVSSL RDRFPDLPIH HHTHDSAGTRVASMTACALAGAD GADVV DVAINSMS GLT SQPSIGALVACTRGTPLDTEVPMER HLCIKDMAGLKFTACTMLVSSL RDRFPDLPIH HHTH G A A AG D VD A ++G TSQP + +++ D + IRRISPYWEAVRNQYAAFBESDLKGPASEVYLHEMPGQQFTNLKEQARSLGLETRWHQVAQAVDANDANGMFGDIVKVTPSSKVVGDMALMNVSQD VCALEFYWBALRKVY APFE SGLFGPTGRV YHH BIPGQQLSM LRQQ AIALG LGDR FEETEEN YAGADRVLGRUVKVTPTSKVVGDLA LALVGAG VEDJSEYW BAARGLLXSCF FADLKGPDEFVYQH BIPGQQLSM LRQQ AIALG LGSKFKEVKKA YVEANQLG RLVKVTPTSKVVGDLA LALVGAG VFDYSEYW BAARGLLXAFDCTATMKSGNSDV YENS IPGQQYTNLHFQ AHSMG LGSKFKEVKKA YDEANQLGDLWKVTPSSKVVGDLA LALVGAG VFDYSEYW BAARGLYAAFDCTATMKSGNSDV YENS IPGQQYNLHFQ AHSMG LGSKFKEVKKA VEANQULG RLVKVTPTSKVVGDLA LALVGAG VSADEFASDPARFGIPE SVLGFLRGEL GDP PGWPFBLQKKALKGEKP-YTVRPGSLLKEADLDAERKVIEKKLEREVSDFEFASYLMYFKVFTD VSADEFASDPARFGIPE SVLGFLRGEL GDP PGWPFBLYKKKLTCPP G	635 641 712 696 728 735 806 790 822 830 822 830 822 830 829 823 915 925 925 993 968 1009 1019
Hs Ret	REG PEGFARAVR.NHPGLL LMDTFR DARQSLLATRVR.THDLKKI APVVARHFSKLFSMENNGGATFDVARRFLYCC WRR LQELR &LIP NIRFQ M GP F +R + DTT RD HQSLLATRVR.T + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSNDAIAEEN-KLCEAAICYTGDLINSARPKYDLKYYTNLAVELEKAGA LLRGRNTVGYTPY PEIVTSAFVQEATATGIDIF RIFDALNNIESM RPAID AVRETGSATAEVAMCYTGDLTDPGEQLTLDY YLKLAEQIVDAGA LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDCLNWVENNRVSNDAIAEEN-KLCEAAICYTGDLINSARPKYDLKYYTNLAVELEKAGA LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDALNNIESM RPAID AVRETGSATAEVAMCYTGDLTDPGEQLTLDY YLKLAEQIVDAGG LLRGANAVGYTNYPDNVVFFGCVAKENGMDVFRVFDALNDLEQLKVGVDAVKKAG-GVVEATVCFGGDMLQPG-KKYNLDYYLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFKJCVAKENGMDVFRVFDALNDLEQLKVGVDAVKKAG-GVVEATVCFSGDMLQPG-KKYNLDYYLEIAEKIVQMGT LLRGANAVGYTNYPDNVVFKJCVAKENGMDVFRVFDSLWJLPNMLLGMEAAGSAG-GVVEAAISTGDADPSRTKYSLOYYMGLAEALVVRGT LLRGN VGY P+ F A +D+FR+FD LN + + ++A + E + + GD+ Y L YY +A + G HIIAVKDMAGLLKPAAAKVLFKALREAT-GLPIHFHTHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAW EVLAIKDMAGLKPPAAQRUVSALSSRF-DLFVHLHTHDTPGGLASVVA WHA AGDAVDGAANPLAGTTS QPALSSIVAAAAHTEYDTG LSLSA HIIGIKDMAGTMKPAAKLLIGSLAKYYDPL HWYTHDSGGTAVASWTACALA GADVV DVAINSMS GLTS QPSINALLASLEGS-HDTGTINVEH HILCIKDMAGLKPFACTMLVSSL RDRFPDL PLWIH HTHAPSGAGVAAMLACAQA GADVV DVAINSMS GMTSQPSMGALVACTRGTPLDTEVPMER H++ +KDMAG ++P A L +LR LP+H HTH G A A AG D VD A ++G TSQP + +++ D + IRRISFYWEAVRNQYAAFBSDLKGPASEVYLHENPGGQFTNLKEQAISLGLETRWHQVAQAYADANQMFGDIVKVTPSSVVGDLALMVSQD VCALEPYWSALRKVY AFF SGLP PTGRV YHH EIPGGLSN LRQQAIALGCEGEFEELEEXAYAGARVLGRLVKVTPTSKVVGDLALALVGAG Y ELDAYWAARMLLYSCF RADLKGPDSEVYLHENPGGQTTNLKEQAILGGLSEGEFEELEEXAYAGARVLGRLVKVTPTSKVVGDLALALVGAG Y WR Y+ F+ VY +E+FGGQ NL QA +GL ++ + +AY A+ G ++KVTP SKVVGDLALALVGAG Y WR YFY SELUGFLGGYGFPEFEGULGYRGFPEFENJUKNKKKYCRDALAADELSSVGAKRQATLNNLLFFSPYKE LTVADVVSPDREVSFPESVSNLKGDLGQPSGWPEALQKKALKGEKP-YTVRPGSLLKEALDLABERKVIEKKLEREVSDFFEASYLMPKVFTD VSADEFASDPARGIPSSVLGFLGGLGQPGSGPEFELGAVLKKALKGEKP-YTVRPGSLLKEALLDABERKVIEKKLEREVSDFFEASYLMPKVFTD LSSDARRGLYAAFD	635 641 712 696 728 735 806 790 822 830 822 830 829 833 915 925 993 968 1009 1019
HS Retts Retts Retts Retts Retts Retts Retts	REG PEGFARAVR NHPGLLLMD TFR DAHQSLLATRVR THDLKKI APVVARNFSKLFSNEDNGGATF DVANR FLYECP WRR LQELR ELIP NIRFQ M GP F +R + DTT RD HQSLLATRVR TH LKKI APVVARNFSKLFSNEDNGGATF DVANR FLYECP WRR LQELR ELIP NIRFQ M GP F +R + DTT RD HQSLLATRVR TH + A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LLRGANGVGY TNYPDNVVKY FVRQAAKGGIDLFRVFDCLNWVENNRVSMDAIAEEN-KLCEAAICY TGDLIDSARP KYDLKYYTNLAVELEKAGA LLRGANGVGY TNYPDNVVKY FVRQAAKGGIDLFRVFDAL NDLE OLKVGVD AVKKAG-GVVE ATVCFSG DMLQPG-KK YNLDY YLEIAEKIVQMGT LLRGANGVAY SSLPDNAIDHF VKQ AKDNSVD IF RVFDAL NDLE OLKVGVD AVKKAG-GVVE ATVCFSG DMLQPG-KK YNLDY YLEIAEKIVQMGT LLRGANGVAY SSLPDNAIDHF VKQ AKDNSVD IF RVFDAL NDLE OLKVGVD AVKKAG-GVVE ATVCFSG DMLQPG-KK YNLDY YLEIAEKIVQMGT LLRGANGVYTNYPDNVVFKFCEVAKENGMD VFKVPSL NYLPM NLLGMEAAGSAG-GVVE ATSY GDVADPSSTK YSLQY YMGL AEALVRAGT LLRGN VGY P+ F A +D+FR+FD LN + ++A + E + + GD+ Y L YY +A + G HIIAVKDNAGLLKPAAAKVLFKALREAT-GLP IHFHTHDTSGIAAATVLAAVEAGVDAVDAAMDALSGNTSQPCLGS IVEALSGSERDPGLDPAW HVLAIKDMAGLKPP PA QRLVSALRSRF-DL PVH HTHDFGGLASVVA MHA GADAV DCAAAPLAGTT SQ PALSS IVAAAAHTE YD TO LSLSA HILGIKDMAGTMKPAAKLLIGSL RAKYPDLP IHVTH DSAGTRVASMTACALAGADAV DVAAANS GUT SQ PSINALLASLEGN- IDTG INVEH H+L KLMAG LKPP PA QRLVSALRSRF-DL PVH HTHDFGGQLSVX MAHA GADAV DVAADSMS GWT SQ PSINALLASLEGN- IDTG INVEH H++ +KDMAG ++P A L +LR LP+H HTH G A A AG D VD A ++G TSQP + +++ D + IRRISFYWEAVRNQYAAFESDLKGPSEVYLHENPGGQFTNLKEQARSLGLETRWHQVAQAYAADANQMFGDIVKVTPTSKVVGDLAKHVSQD VCALEPYWEALKKVY APFESDLKGPSEVYLHENPGGQTTNLKEQARSLGLETRWHQVAQAYAADANQMFGDIVKVTPTSKVVGDLAKFMV SNK VFDYSEYWEALKKVY APFESDLKGPDEVYH EIFGGQLSN ERQQ AIAG GGGAF WTAQLAGD EILSVGARQATINRLLFSSVVGDLAKFW SNK VFDYSEYWEALKAVYDGAAFESDLKGPDEVYLHENPGGQTTNLHFQ HABNG LSSKFKEVKXA YVERNUML DLINVTPTSKVVGDLAKFW SNK VFDYSEYWEALKAVDCFRADLKGPLGYNGVFPERSUVLKALKGEKP-YTVRPGSLLKEADLDAERKVIEKKLEREVSDFEFASYLMFVFTD VSADBFABDPARGIPESVSMLKGDLGOP PSGWPEFLRAALGGRGAAF PTAQLAADD EILSS GARQATINRLLFSSTVWGHA +V LTVADVVSPDREVSFPESVVSLLKGDLGOP GGP PSFPRSVLKNKLKGRGAF PTAQLAADD ELSKYDCHAGATGAAVRKAEPGNAAHV VNENKEN	635 641 712 696 728 735 806 790 822 830 822 830 822 830 829 883 915 925 993 968 1009 1019
HS Rettchs	REG PEGFARAVR NHFGLLLMD TFR DAHQSLLATRVR THDLKKI APVVARNFSKLFSNEDNGGATF DVANR FLYECP WER LQELR ELIP MIR OM GP F +R + DTT RD HQSLLATRVR TH LKKI APVVARNFSKLFSNEDNGGATF DVANR FLYECP WER LQELR ELIP MIR OM LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDLWVENNRVSMDAIAEEN-KLCEAAICYTGDILNSARPKYDLKYYTNLAVELEKAGA LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDLINNESM RPAID AVRETGSAIAEVAMCYTGDLTDPGECLYTLDY VI.KLAEQIVDAGA LLRGANGVGYTNYPDNVVKYFVRQAAKGGIDLFRVFDLINNESM RPAID AVRETGSAIAEVAMCYTGDLTDPGECLYTLDY VI.KLAEQIVDAGA LLRGANGVGYSSLPDAMIDHFYKQAKGNOV TRVFDSLNVLPNLEGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT LLRGANAVGYTNYPDNVVFKFCEVAKENGMDVTRVFDSLNVLPNLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT LLRGANAVGYTNYPDNVVFKFCEVAKENGMDVTRVFDSLNVLPNLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT LLRGANAVGYTNYPDNVFKFCEVAKENGMDVTRVFDSLNVLPNLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT LLRGANAVGYTNYPDNVFKFCEVAKENGMDVTRVFDSLNVLPNLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT LLRGANAVGYTNYPDNVFKFCEVAKENGMDVTRVFDSLNVLPNLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEALVRAGT LLRGANAVGYLPPA F A +D+FR+FD LN + + ++A + E + + GD+ Y LYY +A + G HILGIKDMAGLKPPAAKLLIKSLSKER-GLFVHHHDTSGIAAATVLAAVEAGVDAVDAANDAADALSGNTSOPCLGSIVEALSGSERDPGLDPAM HVLAIKDMAGLKPPAAKLLIKSLEKSF-DLFVHHHDTSGIAAATVLAAVEAGVDAVDAANDALSGNTSOPCLGSIVEALSGSERDFOLDFLDTEVPMER HILGIKDMAGTMKPAAKLLIKSLRSF-DLFVHHHDTSGGQTNNLKACALAGADVVDVANDSNSGMTSOPSIMALLASLEGON-IDTGINVEH HILGIKDMAGTMKPAAKLLIKSLRSFPDLFLHIHTAFSGAGVAAMLACAAGAADVVDVANDSNSGMTSOPSIMALLASLEGON-IDTGINVEH HILGIKDMAGTMKPAAKLLISSL RDRFPDLFLHIHTAFSGAGVANMLACALAGADAVVDAADANQMFGDIVKVTPSSKVUGDLAKINSQD VCALEPYWEILKKVAPFSSGLEGP-ASEVYLHEMPGGQTNLKEQARSLGLETKWQVAQAYADANQMFGDIVKVTPSSKVUGDLAKINSQD VCALEPYWEILKKVAPFSSGLEGP-ASEVYLHEMPGGQTNLKEQARSLGLEEKWUQAQAYADANQMFGDIVKVTPSSKVUGDLAKFMVSNK VFDVSEYMEARARLLYSCFEADLKGPDEVYCH BIPGGQLNLLFQAQLGLGEGVAQTKRA YREANVLLGDIVKVTPSSKVUGDLAKFMVSNK VFDUSEYMEARARVLYPFSSUUGPGAPSGNPEALQKKALKGERP-YTVRPGSLLKEAUCAGAAVKARYARSUVGLAKFMVSNK VCALEPYWEILKKVYAPFSSGLEGP-DEVYHENGATHIKSQD VCALEPYWEILKKVYAPFSS	635 641 712 696 728 735 806 790 822 830 822 830 822 830 829 883 915 925 993 968 1009 1019 1087 1061 1103
HS Retts Retts HS HS Retts Retts HS Sts Retts Retts HS Sts Retts HS Retts HS Retts HS Retts HS HS HS HS HS HS HS HS HS HS HS HS HS	REG PEGFARAVR NHPGLL LMD TFR DARGS LLATRVR THDLKKI APYVARNESKLESMENNGGAT FOVARKE LUSCURRELQELRE LIP NIREQM GP F +R + DTT RD HQSLLATRVR THDLKKI APYVARNESKLESMENNGGAT FOVARKE LUSCURRELQELRE LIP NIREQM GP F +R + DTT RD HQSLLATRVR TH LKKI APYVARNESKLESMENNGGAT FOVARKE LUSCURRELQELRE LIP NIREQM GR F +R + DTT RD HQSLLATRVR THOLKKI APYVARNESKLESMENNGGAT FOVARKE LUSCURRELQELRE LIP NIREQM GR F +R + DTT RD HQSLLATRVR THOLKKI APYVARNESKLESMENNGGAT FOVARKE LUSCURRELQELRE LIP NIREQM GR F +R + DTT RD HQSLLATRVR THOL NURSMESSAGE AVACYTOD LINSARPKYDLKYTNLAVELSKAGA LLRGANGVGYTNYD PONVYKYVRQAAKGGID LFRVFDALNDLEQLKVCVD AVKKAG-GVVE AVACYTOD LKNGAPKYDLKYTNLAVELSKAGA LLRGANAVGYTNYD NUVYKYCQAKONSVD IFRVFDALNDLEQLKVCVD AVKKAG-GVVE AVISTOD VADPSRTKYSLQYTNGLAEALVRAGT LLRGANAVGYTNYD NUVYKYCQAKONSVD IFRVFDALNDLEQLKVCVD AVKKAG-GVVE AVISTOD VADPSRTKYSLQYTNGLAEALVRAGT LLRGANAVGYNYD PNUVYKYCQAKONSVD IFRVFDALNDLEQLKVCVD AVKKAG-GVVE AVISTOD VADPSRTKYSLQYTNGLAEALVRAGT LLRGANAVGYNYD NUVYKYCQAKONSVD IFRVFDALNDLEQLKVCVD AVKKAG-GVVE AVISTOD VADPSRTKYSLQYTNGLAEALVRAGT LLRGANAGYNYD PNUVYKYCQAKONSVD IFRVFDALNULEYLEN LANGAAVALASAGGAGVAVDANDALSGNTSOPCLGS IVEALSSERDPGDLPAM HVLAIKDMAGLKPAAKVULFKALRART-GLPIHFHTHDTSGIAATVLAAVEAQVDAVDAANDALSGNTSOPCLGS IVAALAAHTE VD TG LSLSA HILGIKDMAGLKPTACTMLVSSL RAKYPDLP IHVIHTDASGAGAVAAMIA CAQA GADVV DVA INSMS GLT SOPSMGALVACTRGTPLD TEVPMER H1 LKVNAGLKPTACTMLVSSL RAKYPDLP IHVIHTDASGAGAVAAMIA CAQA GADVV DVA INSMS GLT SOPSMGALVACTRGTPLD TEVPMER H1 + KONAG + P A L +LR LP+H HTH G A A G D VD A + 4G TSOP + +++ D + IRRISFYWEAVRQYAAFESDLKGPASEVYLHENPGGQTNLKEQAILGLGCRFEEIEEAYAGADRVLGRLVKVTPSSKVUGDALHNVSQD VCALEPYWEALRKVY APF SOLKGPASEVYLHENPGGQTNLKEQAILGLGCRFEEIEEAYAGADRVLGRLVKVTPSSKVUGDALAHNVSQD VCALEPYWEALRKVY APF SOLKGP-OPEYYQH BIPGGQLTNLFQAQAIGG LSKKEWKKA VVEANQUG LLKKVTPSSKVUGDALAKNVS PN VFDVSETWEARGKUJAAFDCTATMKSONSDV YENE IPGGQUTNLFQAQAGGAACAR TAQLAADDCIAKWAT VEANQUG LLKKVTPSSKVUGDLAKEWYSNK VFDVSETWEARGRUGAAAFDCTATMKSONSDVYENE IPGGQUTNLFQAQAAR TAQLAADDCIAKKAVEANGLAGARVIEKKVOGDEAVANYSNK VFDVSET	635 641 712 696 728 735 806 790 822 830 822 830 822 830 829 883 915 925 993 968 1009 1019 1087 1061 1103 1112
HS Ret Sch Sch Ret Sch Sch Ret Sch Sch Sch Sch Sch Sch Sch Sch Sch Sch	REGPEGFARLVR NHPOLL LND TFFR DAHQS LLATRVR THDLKKI APYVANNESKLFSMENNGGATFDVAMFLIVECPWER LQELE BLIP NIRFON GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DVAMFLIVECPWER LQELE BLIP NIRFON GP F +R + DTT RD HQSLLATR+RT + +A + ++E WGGAT+DVAMFLIVECPWER LQELE BLIP NIRFON LLRGANGYGTNYD DNVVKYURQAARGGIDJFRVPCLNWVENNKVSMDAIAED-KLCEAAICYTGDILNSARPKYDLKYTNLAVELEKAGA LLRGANGYGTSDID DNIDHFVKQ AKDNSVDIFRVPDLNNLESM RPAID AVKTGSAILGUTPGEGOLTVLDY YLKIAGOVUDAG LLRGANAYGTNYD DNVVFKCEVAKENGMUVFRVPDLNNLEGMEAAGSAG-GVVEAAIST GDVADPSRTKSLQY YMGLAEALVRAGT LLRGANAYGTNYD DNVVFKCEVAKENGMUVFRVPDLNNLEGMEAAGSAG-GVVEAAIST GDVADPSRTKSLQY YMGLAEALVRAGT LLRGANAYGTNYD DNVVFKCEVAKENGMUVFRVPDLNNLEGMEAAGSAG-GVVEAAIST GDVADPSRTKSLQY YMGLAEALVRAGT LLRGANAYGTNYD DNVVFKCEVAKENGMUVFRVPDLNNLLGMEAAGSAG-GVVEAAIST GDVADPSRTKSLQY YMGLAEALVRAGT LLRGANAYGTNYD DNVVFKCEVAKENGMUVFRVPDLNNLLGMEAAGSAG-GVVEAAIST GDVADPSRTKSLQY YMGLAEALVRAGT LLRGANAYGTNYD DNVVFKCEVAKENGMUVFRVPDLNNLLGMEAAGSAG-GVVEAAIST GDVADPSRTKSLQY YMGLAEALVRAGT LLRGANAYGTNYD DNVVFKCEVAKENGMUVFRVPDLNNLLGMEAAGSAG-GVVEAAIST GDVADPSRTKSLQOF YMGLAEALVRAGT LLRGANAYGTNYD DNVVFKLEKALREAT-GLPIHFHTHDTSGIAAATVLAAVENGVDAVDAAMDALSGNTSQPCLGSIVEALSGSERDPGLDPAM HVLAIKDNAGLKPPAAAKLLIGSLRAKYPDLPIHHTHDSGOLSJUSVALAVCRGVAVVDAADSNSGJUSGPSNGALVACTRGTPLDTEVPMER HLGIKDNAGTMRPAAAKLLISSLRDRFPDLPIH HTHDSGGQTNNLKAQAQAQADVDVAADSNSGJUSGPSNGALVACTRGTPLDTEVPMER HILGIKDNAGTMRPAAAKLLAGSLRAKYPDLPIHHTHDSGOLSNIKQAGAGAVVDVADASNSGJUSGPSNGALVACTRGTPLDTEVPMER HTH+ *KDMGLKPTACTMLVSSLRDRFPDLPIHHTHDSGOLSNIRQQAIALGLGDAFEEIEEAYAGADRUGRLVKVTPTSKVVGDLAKHVSNG VCALEPYWEALRKVYAFFSGLKGPASEVYHENPGQUTNLKFQARSLGLEFRWHQVAQAYADANQNGFDIVKVTPSSKVVGDNALMMSQD VCALEPYWEALRKVYAFFSGLKGSDVYNH BIPGGUTNLKFQARSLGLEFRWHQVAQAYADANQHGDIVKVTPTSKVVGDLAKEWSNG VREDLAWEMERLLYSCF RADLKGP-DEVYH HEIPGGULSNIRQQAIALGLGDAFKEVKXYVENGNULGRLVKVTPTSKVVGDLAKEWSNG VREDLAWEMERGRUTAALAFDCTATMKSGNSDVFRNEIFPGQUTNLKFQARGACHKRAKTREANLLGDIVKVTPTSKVVGDLAKEWSNG VY YE W R Y+ F+ VY +E+PGQQ NL QA +GL +++ +AY A+ G ++KVTP SKVVGDALKENVSNG VY YE WEAGRUTAALAFDCTATMKSGNSDVFRNEIFFGULG	635 641 712 696 728 735 806 790 822 830 822 830 829 883 915 925 993 968 1009 1019 1087 1061 1103 1112
HS RetS RetS RS RE RE RE RE RE RE RE RE RE RE	REG PEGFARVTNNHPGLLEND TFR DANGSLEATRVR THDLKKI APYVANNGSKLPSMENNGGATFDVAMRELING LING LING VELIDITAL OF LEEL PARE OF LANGE ARAVTNNHPGLLEND TFR DANGSLEATRVR THDLKKI APYVANNGSATFDVAMRELING OF LANGE ARAVTNNHPGLLEND THE AND AND ALL ALL PARE OF LANGE ARAVTNNHPGLLEND THE AND ALL ALL PARE OF LANGE AND ALL ALL ALL ALL ALL ALL ALL ALL ALL AL	635 641 712 696 728 735 806 790 822 830 899 883 915 925 925 993 968 1009 1019 1087 1061 1103 1112
HS Retcs	REG PEGFARAVR NHEGL LIND TFE DARQSLLATRAR THDLKKT APYVANNESKLESMENNOGAFT DVANKFLEC WAR LOGLABLIP NITHEN GP F +R + DIT RD HQSLLATRART + +A + ++E WGGAT+DV++RFL E PW RL +R PN+ QM LRGANGVGYTNYPDNVKYFVQAARGGID LFRVFDCLNWVENMRVSMDAIAEEN-KLCGAAICYTGDLINSARFKJLKYTNLAVSLERAGA LLRGANGVGYTNYPDNVKSTYVQAARGGID LFRVFDCLNWVENMRVSMDAIAEEN-KLCGAAICYTGDLINSARFKJLKYTNLAVSLERAGA LLRGANGVAT SSLPDMAIDH VKQAKNOSVD IFRVFDALNDLEOLKVGVD AVKRAG-GVVE ATVCFSG DMLQFG-KKYNLDYYLEIAEKIVQMGT LLRGANAVGTNYPDNVVFKYCCUAKENGMDVRVFDSLNYLPNHLLGHEAGSGAG-GVVEAAVCYFGDLTDGEOLYTLDYYLEIAEKIVQMGT LLRGANAVGTNYPDNVVFKYCCUAKENGMDVRVFDSLNYLPNHLLGHEAGSGAG-GVVEAAISYTGDVADDSRTKYSLQYTMGLAEALVRAGT LLRGANAVGTNYPDNVVFKYCCUAKENGMDVRVFDSLNYLPNHLLGHEAGSGAG-GVVEAAISYTGDVADDSRTKYSLQYTMGLAEALVRAGT LLRGANAVGTNYPDNVVFKYCCUAKENGMDVRVFDSLNYLPNHLGHEAGSGAG-GVVEAAISYTGDVADDSRTKYSLQYTMGLAEALVRAGT LLRGANAVGTNYPDNVVFKYCCUAKENGMDVRVFDSLNYLPNHLGHEAGSGAG-GVVEAAISSTGDVLGSIVEALSGEEDFGLDFAH HYLAIKDNAGLKPFAAAKLJFKALREAT-GLPIHPHTHDTSGIAAATVLAAVBGVDAVDAANDALSGNTSQPCLGSIVEALSGEEDFGLDPAM HYLAIKDNAGLKPFAAAKLJFKALREAT-GLPIHPHTHDTSGIAAATVLAAVBAQDAVDGAAANDALSGNTSQPCLGSIVEALSGEEDFGLDPAM HYLAIKDNAGLKPFAAAKLJFKALREAT-GLPIHPHTHDTSGGLASVYAAWHAGADAVDGAAVDAANDALSGNTSQPCLGSIVEALSGEEDFGLDPEN HILGIKDNAGTMFFAAAKLJIGSIRAKYPDLPIHVITHBSAGTRVASMTACALAGADVDVAANDANSGLTSGPSINALLASLEGN-IDTGINVEH HILCIKDNAGLKPFACTMUSSSK DERPPLDIHITHEAPSGAGVAAMLACAGAGDVVDVAADSMSGMTGGPSMLSSENVGDMALHMVSQD VCALEPYWBALRKVIAFFSGLKGPASEVYLHEMPGGQFTNLKEQARSLGLETRWHQVAQAYADANQMPGDIVKVTPSSKVVGDMALHMVSQD VCALEPYWBALRKVIAFFSGLKGPDEFVYUHEIFGGQLFNLKEQARSLGLETRWHQVAQAYADANQMPGDIVKVTPSSKVVGDLAALAUVGG + YW R YFF SCLPGPDEFVYUHEIFGGQLFNLKEQARSLGLETRWHQVAQAYADANQMFGDIKKVTPTSKVVGDLAALAUVGG + YW R YFF VU Y E+PEGQLIN LLFQ AQUGLGCMAGTKRYKWYTTSKVVGDLAKEWVSNK VFDYSEYWBARGLIYAAFDCTATMKSGNSDV YENE IPGGQUTNLHFQ AHSM GLGSKFKEVKKA YVEANQHLGGLIKVTPSSKUVGDLAALAUVGG + YW R YFF + VY +E+PEGQUIN LQA +GL +++ +AY A++G ++KVTP SKVUGDLAFWVSNK VFDYSEYBBESTSPSUSKLKGDLGOPPSGWFEAQKALLGGREFSTVVNGSLLAKKUTGCHAKVVTYSSKUVGDLAALAUVGNG + YW R YFF + VY +E+PEGQU	635 641 712 696 728 735 806 790 822 830 899 883 915 925 925 993 968 1009 1019 1087 1061 1103 1112
HS Retcs	REG PEGFARAVR NHPGLL IMD TFE BAR QSLLATRVR THDLKKI APYVARNESKI PSMENNGGAF DVARKFIZEC WRR LOGLAR LIP NIRFON GP F +R + DIT RD HQSLLATRVR THDLKKI APYVARNESKI PSMENNGGAF DVARKFIZEC WRR LOGLAR LIP NIRFON GP F +R + DIT RD HQSLLATRVR THDLKKI APYVARNESKI PSMENNGGAF DVARKFIZEC WRR LOGLAR LIP NIRFON LLRGANVGYTNY PDNVVKY FYQARKGGI DFRVFDCLMWYENNRVSMDAIA BEN-KLCKAAI (YTGDLINSAR FKYLKYTNLAVELEKAGA LLRGANVGYTEY PENVYESAF VGRTATAG DIF RIFDALNNIESK BRADIA AKREGA LAKAI (YTGDLINSAR FKYLKYTNLAVELEKAGA LLRGANVGYTEY PENVYESAF VGRTATAG DIF RIFDALNNIESK BRADIA AKREGA LAVCAG GUVE ATVCFSG DMLQPG-KKYNLDY YLEIAEKIVONGT LLRGANVGYTEY PDNVVFKCEVA KENGND VRVY PDALNDLEQLKVGVD AVKKAG-GVVE ATVCFSG DMLQPG-KKYNLDY YLEIAEKIVONGT LLRGN NVGT P+ F A + D+FRFFD LN + + +A + E + GD+ Y L YY +A + G HIIAVKDNAGLLKPAAKVLFKALREAT-GLFIHFITHDTSGIAAATVLAVERGVDAVDAANDALSGNTSQPCLGSIVEALSGSERDPGLDPAM HVLAIKDNAGLLKPAA, QELVSALESRF-DLPVHLETHDTPGGQLASVXA WHH GDAVDGAAPLAAGTTSQPALSSI VAALASGSERDPGLDPAM HVLAIKDNAGLLKPAA, QELVSALESRF-DLPVHLETHDTPGGQLASVXA WHH GDAVDGAAPLAAGTSQPALGSI VAAAHTEYDTGG SLSA HILGIKDNAGLLKPTA.CTMLVSSL BESRF-DLPVHLETHDTPGGQLASVA AWHA GADAVDVAN INSMS GHTSQPSMGALVACTRGTEDDTEVPHMER H++ +KDMAG ++P A L +LR LP+H HTH G A A AG D V D A ++G TSQP + +++ D + IRRISTYWEAVRMQYAAFESDLKGPASEVILHEMPGGQTSNLKEQARSIGLETRWHQVAQAYADANGVFGDIVKVTPTSKVVGDLALKHWSQD VGALEEYW HALRKVATPTSSLVGPDFRVYHHEIPGGQLSN LRQQ AIALGAGA GADV DVAANSMS GHTSQPSMGALVACTRGTEDDTEVPHMER H++ +KDMAG ++P A L +LR LP+H HTH G A A AG D V D A ++G TSQP + +++ D + IRRISTYWEAVRMQYAAFESDLKGPDFRVYLHEMPGGQTSN LRGQAIALLGAGA GADV DAANSMSGRUVGDLKVTPTSKVVGDLALKWYSN VFDVSE WWEARGLAXATDCTATMKSGNSDV VENE IPGGQLSN LRQQ AIALGAGA GADV DAANSMSGRUVGLAKVTTSSKVGDNALKHWSQD VCALEPYW HALRKVLAPC -DFRV VLHEMPGGQTSN LHFQAQAG ALGGENAATKRA KEN NYLLGD LIXVTPTSKVVGDLAKWYSNK VFDVSE WWEARGLAXATDCTATMKSGNSDV VENE IPGGQLSN LRQQ AIALGAGA AKRA KEN NYLLGD LIXVTPTSKVVGDLAKHWSNK VFDVSE WWEARGLAXATDCTATMKSGNSDV VENE IPGGQLSN LLFQAQCH FEXVY RAGAGUAKRA KEN NYLDGD LSRAEAFAGALVANAGUAVSKYLFY DE SLEGPGDFRVVLHENGVLGAN AGAGACHAC	635 641 712 696 728 735 806 790 822 830 899 883 915 925 925 993 968 1009 1019 1087 1061 1103 1112

FIG. 7. Comparison of the deduced amino acid sequence of the *R. etli* (Re) PYC with the N-terminal 26 amino acid residues of the *R. capsulatus* (Rc) PYC and the entire deduced amino acid sequences of the *M. tuberculosis* (Mt), *Saccharomyces cerevisiae* (Sc), and *Homo sapiens* (Hs) PYCs. Amino acids identical to those in the Re PYC are in boldface, and residues conserved in all sequences are listed below the alignments, with plus signs indicating conservative amino acid substitutions. Sequence motifs discussed in the text are overbarred. Dashes indicate gaps introduced into the sequences to maximize alignment. Sequences were taken from the references cited in Fig. 8 except for the *R. capsulatus* (39) and *S. cerevisiae* (34) PYCs.



FIG. 8. Schematic representation of the deduced domain structure of the *R. etli* PYC based on sequence homologies to other proteins containing defined sequence motifs. The functional domains are indicated in the key, and some of their associated amino acid (aa) sequence motifs are shown in the diagrams. The extent of the shading representing each of the various motifs was determined by homology comparisons of the following proteins with the *R. etli* PYC: transcarboxylase (TC) domain, *P. shermanii* TC 5s subunit; biotin carboxylase (BC) and BCCP domains, *E. coli* acetyl-CoA carboxylase (ACC) BC and BCCP subunits, respectively. Amino acid residues corresponding to the numbering in the sequence motifs are in boldface. Sequences were taken from the references indicated: human PYC (67); *M. tuberculosis* PYC (57); *P. shermanii* TC 5s subunit (64); human propionyl-CoA carboxylase (PCC) α subunit (31); rat carbamylphosphate synthase (CPS) (44); *E. coli* ACC BC subunit (32); *E. coli* ACC BCCP subunit (41); and *P. shermanii* TC 1.3s subunit (35).

biotin in maintaining an aerobic metabolism in *R. etli*, is a required cofactor for the biotin synthase of *E. coli* (6).

The biochemical data presented show that the R. etli PYC is composed of subunits of approximately 120 kDa, and analysis of the enzyme by gel chromatography indicate that the native enzyme is approximately 400 kDa (15). These data are consistent with the enzyme having an α_4 structure. This conclusion is supported by the finding that the partially purified PYC (Table 3) has regulatory properties similar to those of many α_4 PYCs (4, 54). The relatively small variation in R. etli PYC activity during growth on different carbon sources (Table 2) is similar to that observed with the allosterically regulated α_4 PYCs produced by Bacillus subtilis (13) and B. licheniformis (49). In contrast, the $\alpha_4\beta_4$ PYCs from A. vinelandii (55) and P. citronellolis (62) are highly inducible and are not subject to regulation by effectors. The simultaneous presence of the α_4 PYC and PPC (Table 5) in R. etli is worthy of note. Although an allosterically regulated PPC occurs in many bacterial species which produce the $\alpha_4\beta_4$ form of PYC (12, 28, 45, 55), PPC has been shown to be absent in species producing the α_4 subclass of PYC (1, 13, 46, 50, 68). This finding raises the question of how

PYC and PPC are regulated in *R. etli*, since many bacterial PPCs are allosterically regulated in a manner very similar to that of the α_4 PYCs (42, 55).

Under what growth conditions do R. etli and R. tropici utilize PYC? When grown on glucose or other carbon sources utilized via the C3 compounds pyruvate or PEP, bacteria anaplerotically produce \bar{C}_4 compounds by using CO_2 -fixing enzymes (PYC or PPC) or the glyoxylate shunt (1). For example, the growth of Bacillus megaterium on C6 or C3 compounds requires PYC unless a source of C_4 compounds (e.g., aspartate) is available or unless the glyoxylate cycle is functioning (1). The R. etli and R. tropici pyc and parent strains grew equally well in rich medium (14, 19) and to similar final cell densities during the first subculture in MM-succinate (Fig. 3) (15, 19). This finding is consistent with the belief that PYC activity is not required for growth on a C₄ carbon source (MM-succinate) or in medium in which a precursor thereof (in PY) is present (1, 13). The mutant strains of both species grew to some extent in glucose (Fig. 2B), a substrate which can be catabolized to PEP in Rhizobium spp. (50). The PEP produced from glucose could be converted to OAA in the reaction catalyzed by PPC, thereby allowing some growth of the mutants. In contrast, pyc mutants of B. subtilis do not grow on glucose probably because PPC is not produced by this species (13). The Rhizobium mutants failed to grow on pyruvate (Fig. 2A) as demonstrated for pyc mutants isolated from other species (39, 46, 68). Our failure to detect PEP synthase, or pyruvate or orthophosphate dikinase (either of which can convert pyruvate to PEP [14]), in R. etli and R. tropici indicates that OAA synthesis from pyruvate (via PEP) cannot occur and prevents the growth of the mutants on this carbon source. We believe that the Rhizobium mutants are able to grow in pyruvate medium supplemented with aspartate (Fig. 2A) because OAA is produced by the action of aspartate aminotransferase (54), for which there is considerable activity in R. etli and R. tropici (14). The production of OAA catalyzed by malate dehydrogenase appears to be possible in the pyc mutants (Table 5), although it has been proposed that during growth on C_3 or C_6 substrates, this reaction produces insufficient OAA for both energy generation (via the tricarboxylic acid cycle) and biosynthesis. Consequently, the anaplerotic activity of PYC is necessary for growth under these conditions (13). Finally, the production of C_4 metabolites via the glyoxylate shunt is unlikely to occur in the pyc mutants, since the enzymes required for this pathway in R. etli CE3 (14) and R. tropici CFN299 (27) are present at significant levels only during growth in acetate-containing medias, as previously documented for other bacterial systems (43).

Our hypothesis that PYC is the enzyme upon which biotin acts in preventing the fermentative metabolism in R. etli was based on results of experiments using MM containing 10 mM succinate (18). Under these conditions, biotin promoted almost equally good growth of the pyc mutant and parent strains in subcultures, suggesting that a biotin-dependent system, distinct from PYC, is the site for biotin action in the mutant. We do not know, however, if this alternative biotin-dependent system is also the site at which biotin acts in the parent strain, or whether it merely compensates for a lack of PYC in the mutant. PYC does appear to have a clear role in allowing R. etli (but not R. tropici) to grow continuously in serial subcultures in media containing 30 mM succinate (Fig. 3). This conclusion is based on the fact that biotin supplementation was not able to restore the growth of the R. etli mutant during subculturing, indicating that PYC is responsible for the growth-maintaining effect of biotin under these conditions. Although a requirement for PYC during growth on succinate appears to be inconsistent with the generally accepted metabolic role for this enzyme (see above), we speculate that growth in higher concentrations of succinate might cause a greater flux, via malic enzyme, of succinate-derived carbon to pyruvate rather than to OAA (via malate dehydrogenase). If this in fact does occur, PYC could be essential in generating sufficient OAA for biosynthetic and energy-generating purposes. This hypothesis remains to be tested.

The deduced amino acid sequence of the *R. etli pyc* has significant similarity to the PYC sequences from a diverse group of organisms (Fig. 7) and contains a biotin carboxylase domain in its N-terminal region, a BCCP domain in its C-terminal region, and a transcarboxylase domain, with a binding site specific for pyruvate, in its central region (Fig. 8). The subunit MW of the protein as determined by SDS-PAGE or as predicted from the deduced amino acid sequence of the cloned gene is similar to those of eukaryotic α_4 PYCs (4). This finding extends the observation that there is a strong evolutionary conservation among biotin-containing carboxylases (29, 52) and a prokaryotic PYC. Our sequence comparison results also indicate that the putative *M. tuberculosis pyc* (57) does indeed encode an α_4 PYC.

Putative *pyc* mutants of *Rhizobium leguminosarum* by. trifolii (50) and *R. leguminosarum* by. viciae (3) were reported to be symbiotically proficient on clover and pea plants, respectively. The findings from our in planta experiments are consistent with these data and support the generally accepted proposal that the major carbon sources used by bacteroids are dicarboxylic acids such as malate or succinate (58). Interestingly, by using very low inoculum densities, the importance of biotin in the infective capability of *R. meliloti* was recently demonstrated, although the biochemical basis for this effect was not determined (59). Because the activity of PYC in *R. etil* is greatly affected by biotin, it would be of interest to determine if our mutant strain is altered in competitiveness or growth in the bean rhizosphere.

ACKNOWLEDGMENTS

We are grateful to Patricia Bustos for technical assistance in DNA sequencing, Armando Rivera and Antonia Jaimes for performing the plant experiments, and Gabriela Guerrero for assistance with digitized images. For suggestions on sequence analysis, we acknowledge Miguel Cevallos; for helpful comments on the manuscript, we thank Alberto Mendoza and the two anonymous reviewers. We gratefully acknowledge the superior work done by Brigitte Obermaier of Medigene in the final nucleotide sequence determinations.

This work was supported by a grant from Consejo Nacional de Ciencia y Tecnología, México CONACyT no. N9111-0954, and by DGAPA-UNAM no. IN202393 and IN213095.

REFERENCES

- Al-ssum, R. M., and P. J. White. 1977. Activities of anaplerotic enzymes and acetyl coenzyme A carboxylase in biotin-deficient *Bacillus megaterium*. J. Gen. Microbiol. 100:203–206.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403–410.
- Arwas, R., A. Ř. Glenn, I. A. McKay, and M. J. Dilworth. 1986. Properties of double mutants of *Rhizobium leguminosarum* which are defective in the utilization of dicarboxylic acids and sugars. J. Gen. Microbiol. 132:2743– 2747.
- Attwood, P. V. 1995. The structure and the mechanism of action of pyruvate carboxylase. Int. J. Biochem. Cell Biol. 27:231–249.
- Beringer, J. E. 1974. R factor transfer in *Rhizobium leguminosarum*. J. Gen. Microbiol. 84:188–198.
- Birch, O. M., M. Fuhrmann, and N. M. Shaw. 1995. Biotin synthase from Escherichia coli, an investigation of the low molecular weight and protein components required for activity in vitro. J. Biol. Chem. 270:19158–19165.
- Bravo, A., B. Becerril, and J. Mora. 1988. Introduction of the *Escherichia coli* gdhA gene into *Rhizobium phaseoli*: effect on nitrogen fixation. J. Bacteriol. 170:985–988.
- Bravo, A., and J. Mora. 1988. Ammonium assimilation in *Rhizobium phaseoli* by the glutamine synthetase-glutamate synthase pathway. J. Bacteriol. 170: 980–984.
- Bray, G. A. 1960. A simple efficient liquid scintillator for counting aqueous solutions in a liquid scintillation counter. Anal. Biochem. 1:279–285.
- Bridger, G. P., and T. K. Sundaram. 1976. Occurrence of phosphoenolpyruvate carboxylase in the extremely thermophilic bacterium *Thermus aquaticus*. J. Bacteriol. 125:1211–1213.
- Cevallos, M. A., S. Encarnación, A. Leija, Y. Mora, and J. Mora. 1996. Genetic and physiological characterization of a *Rhizobium etli* mutant strain unable to synthesize poly-β-hydroxybutyrate. J. Bacteriol. 178:1646–1654.
- Cohen, N. D., J. A. Duc, H. Beegan, and M. F. Utter. 1979. Quaternary structure of pyruvate carboxylase from *Pseudomonas citronellolis*. J. Biol. Chem. 254:9262–9269.
- Diesterhaft, M. D., and E. Freese. 1973. Role of pyruvate carboxylase, phosphenolpyruvate carboxykinase and malic enzyme during growth and sporulation of *Bacillus subtilis*. J. Biol. Chem. 248:6062–6070.
- 14. Dunn, M. F., and G. Araíza. Unpublished results.
- Dunn, M. F., and J. Mora. 1995. Biochemical and genetic characterization of pyruvate carboxylase in *Rhizobium etli* and *R. tropici*, p. 576. *In* I. A. Tikhonovich, N. A. Provorov, V. I. Romanov, and W. E. Newton (ed.), Nitrogen fixation: fundamentals and applications. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Eckhardt, T. 1978. A rapid method for the identification of plasmid deoxyribonucleic acid in bacteria. Plasmid 1:584–588.
- Encarnación, S., M. Dunn, A. Leija, M. C. Vargas, H. Peralta, and J. Mora. 1995. Alternative metabolic programs in *Rhizobium*, p. 577. *In* I. A. Tikhonovich, N. A. Provorov, V. I. Romanov, and W. E. Newton (ed.), Nitrogen

fixation: fundamentals and applications. Kluwer Academic Publishers, Dordrecht, The Netherlands.

- Encarnación, S., M. Dunn, K. Willms, and J. Mora. 1995. Fermentative and aerobic metabolism in *Rhizobium etli*. J. Bacteriol. 177:3058–3066.
- 19. Encarnación, S., and M. C. Vargas. Unpublished results.
- Figurski, D. H., and D. R. Helinski. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans.* Proc. Natl. Acad. Sci. USA 76:1648–1652.
- Friedman, A. M., S. R. Long, S. E. Brown, W. J. Buikema, and F. M. Ausubel. 1982. Construction of a broad host range cosmid cloning vector and its use in the genetic analysis of *Rhizobium* mutants. Gene 18:289–296.
- Fry, D. C., S. A. Kuby, and A. S. Mildvan. 1986. ATP-binding site of adenylate kinase: mechanistic implications of its homology with ras-encoded p21, F1-ATPase, and other nucleotide-binding proteins. Proc. Natl. Acad. Sci. USA 83:907–911.
- Garfin, D. E. 1990. One-dimensional gel electrophoresis, p. 425–441. *In* M. P. Deutcher (ed.), Guide to protein purification. Academic Press Inc., New York.
- Gloeckler, R., I. Ohsawa, D. Speck, C. Ledoux, S. Bernard, M. Zinsius, D. Villeval, T. Kisou, K. Kamogawa, and Y. Lemoine. 1990. Cloning and characterization of the *Bacillus sphaericus* genes controlling the bioconversion of pimelate into dethiobiotin. Gene 87:63–70.
- Gold, L. 1988. Posttranscriptional regulatory mechanisms in *Escherichia coli*. Annu. Rev. Biochem. 57:199–233.
- Goss, J. A., N. D. Cohen, and M. F. Utter. 1981. Characterization of the subunit structure of pyruvate carboxylase from *Pseudomonas citronellolis*. J. Biol. Chem. 256:11819–11825.
- 27. Hernández-Lucas, I. Personal communication.
- Higa, A. I., S. R. Milrad de Forchetti, and J. J. Cazzulo. 1976. CO₂-fixing enzymes in *Pseudomonas fluorescens*. J. Gen. Microbiol. 93:69–74.
- Knowles, J. R. 1989. The mechanism of biotin-dependent enzymes. Annu. Rev. Biochem. 58:195–221.
- Kumer, G. K., F. C. Haase, N. F. B. Phillips, and H. G. Wood. 1988. Involvement and identification of a tryptophanyl residue at the pyruvate binding site of transcarboxylase. Biochemistry 27:5978–5983.
- Lamhonwah, A., F. Quan, and R. A. Gravel. 1987. Sequence homology around the biotin-binding site of human propionyl-CoA and pyruvate carboxylase. Arch. Biochem. Biophys. 254:631–636.
- Li, S., and J. E. Cronan. 1992. The gene encoding the biotin carboxylase subunit of *Escherichia coli* acetyl-CoA carboxylase. J. Biol. Chem. 267:855– 863.
- Libor, S., T. K. Sundaram, R. Warwick, J. A. Chapman, and S. M. W. Grundy. 1979. Pyruvate carboxylase from a thermophilic *Bacillus*: some molecular characteristics. Biochemistry 18:3647–3653.
- Lim, F., C. P. Morris, F. Occhiodoro, and J. C. Wallace. 1988. Sequence and domain structure of yeast pyruvate carboxylase. J. Biol. Chem. 263:11493– 11497.
- Maloy, W. L., B. U. Bowien, G. K. Zwolinski, G. Kumar, H. G. Wood, L. H. Ericsson, and K. A. Walsh. 1979. Amino acid sequence of the biotinyl subunit from transcarboxylase. J. Biol. Chem. 254:11615–11622.
- Martínez, E., R. Palacios, and F. Sánchez. 1987. Nitrogen-fixing nodules induced by *Agrobacterium tumefaciens* harboring *Rhizobium phaseoli* plasmids. J. Bacteriol. 169:2828–2834.
- McAlister, L. E., E. L. Evans and T. E. Smith. 1981. Properties of a mutant Escherichia coli phosphoenolpyruvate carboxylase deficient in coregulation by intermediary metabolites. J. Bacteriol. 146:200–208.
- Milrad de Forchetti, S. R., and J. J. Cazzulo. 1976. Some properties of the pyruvate carboxylase from *Pseudomonas fluorescens*. J. Gen. Microbiol. 93: 75–81.
- Modak, H. V., and D. J. Kelly. 1995. Acetyl-CoA-dependent pyruvate carboxylase from the photosynthetic bacterium *Rhodobacter capsulatus*: rapid and efficient purification using dye-ligand affinity chromatography. Microbiology 141:2619–2628.
- 40. Mora, Y., S. Encarnación, and M. Dunn. Unpublished results.
- Muramatsu, S., and T. Mizuno. 1989. Nucleotide sequence of the *fabE* gene and flanking regions containing a bent DNA sequence of *Escherichia coli*. Nucleic Acids Res. 17:3982.
- Myers, D. E., B. Tolbert, and M. F. Utter. 1983. Activation of yeast pyruvate carboxylase: interactions between acyl coenzyme A compounds, aspartate, and substrates of the reaction. Biochemistry 22:5090–5096.

- 43. Nunn, W. D. 1987. Two-carbon compounds and fatty acids as carbon sources, p. 285–301. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 2. American Society for Microbiology, Washington, D.C.
- Nyunoya, H., K. E. Broglie, E. E. Widgren, and C. J. Lusty. 1985. Characterization and derivation of the gene coding for mitochondrial carbamyl phosphate synthetase I of rat. J. Biol. Chem. 260:9346–9356.
- O'Brien, R. W., D. T. Chuang, B. L. Taylor, and M. F. Utter. 1977. Novel enzymic machinery for the metabolism of oxalacetate, phosphoenolpyruvate, and pvruvate in *Pseudomonas citronellolis*, J. Biol. Chem. 252:1257–1263.
- Payne, J., and J. G. Morris. 1969. Pyruvate carboxylase in *Rhodopseudomo-nas spheroides*. J. Gen. Microbiol. 59:97–101.
- 47. Peralta, H., and M. Dunn. Unpublished results.
- Post, L. E., D. J. Post, and F. M. Raushel. 1990. Dissection of the functional domains of *Escherichia coli* carbamoyl phosphate synthetase by site-directed mutagenesis. J. Biol. Chem. 265:7742–7747.
- Renner, E. D., and R. W. Bernlohr. 1972. Characterization and regulation of pyruvate carboxylase of *Bacillus licheniformis*. J. Bacteriol. 109:764–772.
- Konson, C. W., and S. B. Primrose. 1979. Carbohydrate metabolism in *Rhizobium trifolii*: identification and symbiotic properties of mutants. J. Gen. Microbiol. 112:77–88.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Samols, D., C. G. Thornton, V. L. Murtif, G. K. Kumar, F. C. Haase, and H. G. Wood. 1988. Evolutionary conservation among biotin enzymes. J. Biol. Chem. 263:6461–6464.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463–5467.
- Scrutton, M. C. 1978. Fine control of the conversion of pyruvate (phosphoenolpyruvate) to oxaloacetate in various species. FEBS Lett. 89:1–9.
- Scrutton, M. C., and B. L. Taylor. 1974. Isolation and characterization of pyruvate carboxylase from *Azotobacter vinelandii* OP. Arch. Biochem. Biophys. 164:641–654.
- Simon, R. 1984. High frequency mobilization of gram-negative bacterial replicons by the *in vitro* constructed Tn5-mob transposon. Mol. Gen. Genet. 196:413–420.
- 57. Smith, D. R. 1994. GenBank accession number U00024.
- Streeter, J. G. 1991. Transport and metabolism of carbon and nitrogen in legume nodules. Adv. Bot. Res. 18:129–187.
- Streit, W. R., C. M. Joseph, and D. A. Phillips. 1996. Biotin and other water-soluble vitamins are key growth factors for alfalfa root colonization by *Rhizobium meliloti* 1021. Mol. Plant-Microbe Interact. 9:330–338.
- Sundaram, T. K. 1973. Physiological role of pyruvate carboxylase in a thermophilic *Bacillus*. J. Bacteriol. 113:307–315.
- 61. Taboada, H. Personal communication.
- Taylor, B. L., S. Routman, and M. F. Utter. 1975. The control of the synthesis of pyruvate carboxylase in *Pseudomonas citronellolis*: evidence from double labeling studies. J. Biol. Chem. 250:2376–2382.
- 63. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673–4680.
- 64. Thornton, C. G., G. K. Kumar, B. C. Shenoy, F. C. Haase, N. F. B. Phillips, V. M. Park, W. J. Magner, D. P. Hejlik, H. G. Wood, and D. Samols. 1993. Primary structure of the 5S subunit of transcarboxylase as deduced from the genomic DNA sequence. FEBS Lett. 330:191–196.
- Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA 76:4350–4354.
- West, P. M., and P. W. Wilson. 1940. Biotin as a growth stimulant for the root nodule bacteria. Enzymologia 8:152–162.
- Wexler, I. D., Y. Du, M. V. Lisgaris, S. K. Mandal, S. O. Freytag, B. Yang, T. Liu, M. Kwon, M. S. Patel, and D. S. Kerr. 1994. Primary amino acid sequence and structure of human pyruvate carboxylase. Biochim. Biophys. Acta 1227:46–52.
- Willison, J. C. 1988. Pyruvate and acetate metabolism in the photosynthetic bacterium *Rhodobacter capsulatus*. J. Gen. Microbiol. 134:2429–2439.