Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 3097–3101, April 1990
Biochemistry

Cloning of a chicken liver cDNA encoding 5-aminoimidazole ribonucleotide carboxylase and 5-aminoimidazole-4-*N*-succinocarboxamide ribonucleotide synthetase by functional complementation of *Escherichia coli pur* mutants

(*de novo* purine nucleotide synthesis/nucleotide sequence/structure–function relationship/*E. coli purEK purC*/mutagenesis)

ZENGDAO CHEN, JACK E. DIXON, AND HOWARD ZALKIN

Department of Biochemistry, Purdue University, West Lafayette, IN 47907

Communicated by Charles Yanofsky, January 23, 1990 (received for review December 5, 1989)

ABSTRACT **We have used functional complementation of *Escherichia coli pur* mutants to clone avian cDNA encoding 5-aminoimidazole ribonucleotide (AIR) carboxylase–5-aminoimidazole-4-*N*-succinocarboxamide ribonucleotide (SAICAR) synthetase, the bifunctional enzyme catalyzing steps 6 and 7 in the pathway for *de novo* purine nucleotide synthesis. Mutational analyses have been used to establish the structure–function relationship: NH_2–SAICAR synthetase–AIR carboxylase–COOH. The amino acid sequence of the SAICAR synthetase domain is homologous to that of bacterial *purC*-encoded enzymes, and the sequence of the following AIR carboxylase domain is homologous to that of bacterial *purE*-encoded enzymes. In *E. coli*, AIR carboxylase is the product of genes *purEK* with the purK subunit postulated to have a role in CO_2 binding. The avian enzyme lacks sequences corresponding to purK yet functions in *E. coli*. Functional complementation of *E. coli pur* mutants can be used to clone additional avian cDNAs for *de novo* purine nucleotide synthesis.**

Within the past several years many of the genes for *de novo* purine nucleotide synthesis have been cloned from bacteria (1–5). This work has provided insights into regulation of *de novo* purine nucleotide synthesis in *Escherichia coli* (6) and *Bacillus subtilis* (7) as well as providing improved sources for enzyme isolation. Corresponding studies with higher eukaryotes have lagged, although the *Drosophila melanogaster Gart* locus was cloned by complementation of a yeast mutation (8). Given the importance of purine nucleotides for RNA and DNA synthesis and energy metabolism plus the demonstration that defects in purine salvage lead to human disease (9), we have searched for methods to isolate cDNAs for *de novo* purine nucleotide synthesis from animals. Avian cDNA libraries would appear to be advantageous owing to the fact that birds use the *de novo* pathway for both biosynthesis and for excretion of nitrogen as uric acid.

We have found recently (unpublished data) that screening of an avian cDNA library by functional complementation of *E. coli pur* mutants affords one method to obtain several cDNAs for *de novo* purine nucleotide synthesis. Using this approach, we have isolated cDNAs that encode seven of the nine enzymes for synthesis of IMP (Fig. 1). These enzymes include the trifunctional glycinamide ribonucleotide (GAR) synthetase–AIR synthetase–GAR transformylase (corresponding to *E. coli purDMN*; J. Aimi, J. Williams, J.E.D., and H.Z., unpublished data), adenylosuccinate lyase (*E. coli purB*) (26), and AICAR transformylase–IMP cyclohydrolase (*E. coli purH*; L. Ni, J.E.D., and H.Z., unpublished data). Here we report the cloning of chicken liver AIR carboxylase–SAICAR synthetase by functional complementation of *E. coli pur* mutants and a structure–function characterization for this bifunctional enzyme. We show that the chicken liver cDNA encodes a 426-amino acid enzyme that is similar to a bacterial PurC-PurE fusion protein in which the SAICAR synthetase domain (PurC) corresponds to amino acids 1 to ≈258, and residues 260–426 encode a PurE-like AIR carboxylase function. Surprisingly, the avian enzyme lacks completely a PurK-like AIR carboxylase domain or subunit. Thus, the chicken AIR carboxylase is fundamentally different from the bacterial *purEK*-encoded heterodimer, and the role of the bacterial *purK*-encoded subunit is called into question.

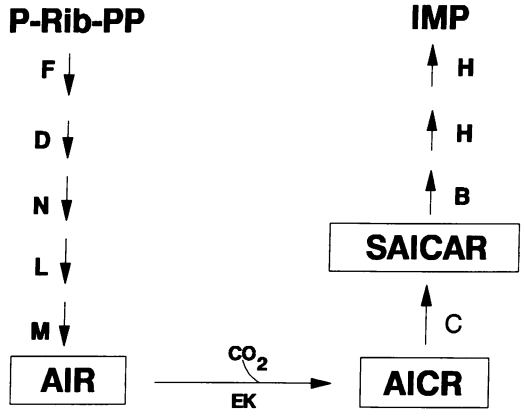

FIG. 1. Pathway for *de novo* purine nucleotide biosynthesis. *E. coli pur* genes are identified by letters. Substrates and products enclosed in boxes are discussed in the text. P-Rib-PP, 5-phosphoribosyl 1-pyrophosphate; AICR, 5-amino-4-imidazolecarboxylate ribonucleotide; AIR, 5-aminoimidazole ribonucleotide; SAICAR, 5-aminoimidazole-4-*N*-succinocarboxamide ribonucleotide.

MATERIALS AND METHODS

Bacteria, Plasmids, and Growth Conditions. *E. coli* strains TX209 (Δ*lac purK*), NK6051 (Δ[*gpt-lac*]*5 purE79*::Tn*10 relA1 spoT1 thi-1*), and NK6056 (Δ[*gpt-lac*]*5 purC*::Tn*10 relA1 spoT1 thi-1*) were used to screen a chicken liver cDNA library. Bacteria were grown in minimal medium (1) containing 0.1% acid-hydrolyzed casein and 0.2 μg of thiamin per ml. Adenine (50 μg/ml) was added as a purine supplement as required. *In vivo* excision and recircularization of recombinant pBluescript plasmids from a λZAP chicken liver cDNA

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

Abbreviations: GAR, glycinamide ribonucleotide; AIR, 5-aminoimidazole ribonucleotide; SAICAR, 5-aminoimidazole-4-*N*-succinocarboxamide ribonucleotide.

library (Stratagene) was carried out as described (10). Strains harboring pBluescript plasmids were grown in medium containing 100 μg of ampicillin per ml. Plasmid pT7-7 was used to determine the molecular mass of the cDNA-encoded enzyme in *E. coli* (11).

RNA Methods. RNA was extracted (12) from frozen chicken liver and subjected to one cycle of oligo(dT)-cellulose chromatography (13). For RNA blot-hybridization (Northern) analysis, RNA was fractionated by electrophoresis on a 1.0% agarose gel containing formaldehyde (14), transferred to nitrocellulose, and probed with ^{32}P-labeled single-stranded DNA. Primer extension mapping (14) with three oligonucleotides that hybridized to mRNA nucleotides 72–90 (probe 1), 145–173 (probe 2), and 269–298 (probe 3) numbered as in Fig. 3 was used to determine the mRNA 5′ end. A sequence corresponding to the 5′ end of the AIR carboxylase–SAICAR synthetase mRNA was amplified by using the polymerase chain reaction (PCR; ref. 15) and cloned (16) as follows. The 173-nucleotide cDNA synthesized by primer extension of probe 2 was isolated by electroelution from a 6% polyacrylamide sequencing gel and tailed with poly(A) using dATP and terminal deoxynucleotidyltransferase. Second-strand synthesis was carried out with a $(dT)_{17}$ primer–adaptor (14, 16). Finally, the double-stranded cDNA product was amplified by the PCR technique using the adaptor of the $(dT)_{17}$ primer–adaptor and an internal primer, probe 4, corresponding to nucleotides 150–166 attached to a 10-nucleotide *Eco*RI adaptor to facilitate cloning. The amplified DNA fragments were digested with *Xho* I and *Eco*RI, and the 185-base-pair (bp) product was purified by polyacrylamide gel electrophoresis and ligated into the *Sal* I and *Eco*RI sites of pUC118 and pUC119 for nucleotide sequencing.

DNA Sequencing. Overlapping deletions were constructed by the unidirectional exonuclease III procedure (17). DNA sequencing was by the dideoxynucleotide chain-termination method (18) using Sequenase (United States Biochemical). The entire sequence was determined on both DNA strands from overlapping clones.

Mutations. Oligonucleotide-directed mutagenesis was by the method of Kunkel *et al.* (19); 3′ deletions were constructed by removal of specific restriction fragments.

RESULTS

Isolation of cDNA Clones for *de novo* Purine Nucleotide Synthesis by Functional Complementation of *E. coli* Mutants. A chicken liver cDNA library constructed in a phage λZAP vector was converted into a plasmid expression library by *in vivo* excision (10). This plasmid expression library was screened for cDNAs encoding enzymes of *de novo* purine nucleotide biosynthesis by complementation of a number of *E. coli pur* mutants. A plasmid encoding chicken AIR carboxylase–SAICAR synthetase was initially isolated by functional complementation of *E. coli purK* mutant TX209. In *E. coli* the two AIR carboxylase subunits are encoded by genes *purEK* (3). Approximately 10^4 transformed cells were spread on each of six Petri plates containing selective medium lacking a source of purines. Two colonies were obtained. Plasmid DNA isolated from these colonies was identical. Each plasmid contained an insert of ≈2.3 kilobases (kb) that was cleaved at an internal *Eco*RI site to yield fragments of 0.87 and 1.4 kb. The plasmid, designated pZD1, retransformed strain TX209, as well as strains NK6056(*purC*) and NK6051(*purE*) to Pur$^+$ at high efficiency (10^4–10^5 transformants per μg of DNA) but did not transform *purB* or *purM* mutants to Pur$^+$. Gene *purC* encodes SAICAR synthetase. Nucleotide sequence analysis, described below, established that plasmid pZD1 contains a 2.3-kb fragment of chicken "purCKE" cDNA.

Analysis of AIR Carboxylase–SAICAR Synthetase RNA and Cloning of the mRNA 5′ End. A blot of poly(A)$^+$ RNA indicated a single message of approximately 2.6 kb (Fig. 2*A*). This suggested that the 2.3-kb insert, although capable of complementing *pur* mutants, was most likely not a full-length clone. The 5′ end of the poly(A)$^+$ RNA was mapped by primer extension with reverse transcriptase and three oligonucleotides designated probes 1–3. Primer extension mapping with probe 2 gave a cDNA of 173 nucleotides (Fig. 2*B*). Probes 1 and 3 gave cDNAs of ≈90 and ≈298 nucleotides, respectively (not shown). The three primer extension reactions define a single 5′ end 69 bp upstream of the first nucleotide in the pZD1 insert.

To clone the 5′ end of mRNA, the cDNA resulting from the primer extension with probe 2 was amplified by the PCR and cloned into pUC118 as described. The inserts from four transformants contained an identical 166-bp 5′ segment of purCKE cDNA flanked by the *Xho* I adaptor–oligo(dT) primer at the 5′ end and *Eco*RI adaptor at the 3′ end. This

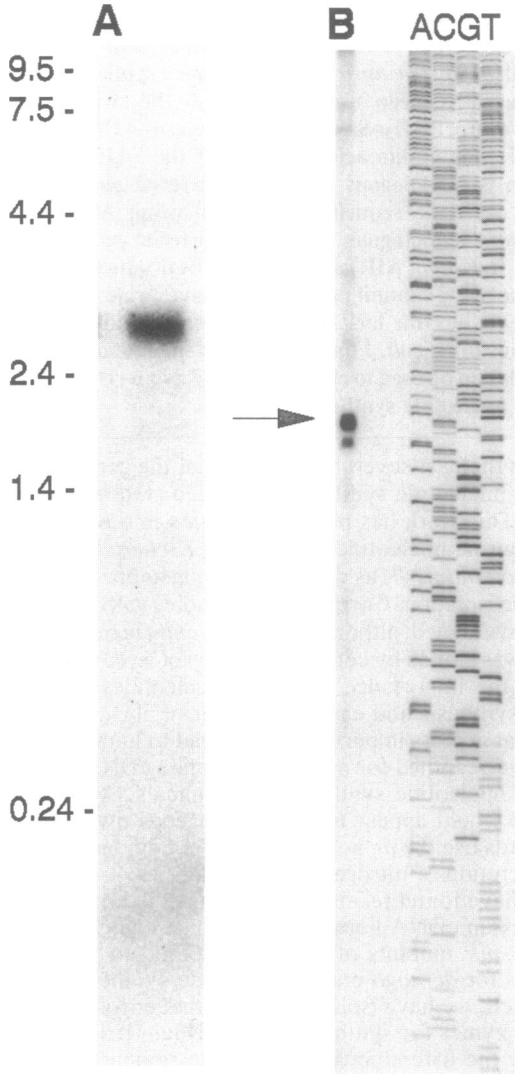

FIG. 2. RNA analysis. (*A*) Northern blot of chicken liver RNA. After electrophoresis a nitrocellulose membrane containing 5 μg of poly(A)$^+$ RNA was hybridized and washed as described (20). Exposure was for 16 hr at −80°C. An RNA ladder (BRL) was used as the size standard (kb). (*B*) Primer extension mapping of the 5′ end of AIR carboxylase–SAICAR synthetase mRNA. The reaction mixture included 5 μg of poly(A)$^+$ RNA, primer 2, and avian myeloblastosis virus reverse transcriptase. A major band corresponding to 173 nucleotides is shown next to a sequencing ladder size standard obtained from pZD1 with primer 2.

sequence completes the AIR carboxylase–SAICAR synthetase coding region.

Nucleotide Sequence of cDNA. The nucleotide sequence of the 2.3-kb insert from pZD1 was obtained from a set of overlapping deletions. The insert in pZD1 corresponds to nucleotides 71–2392, and the PCR-amplified product corresponds to nucleotides 1–166. The nucleotide sequence of a 2392-bp cDNA and derived amino acid sequence are shown in Fig. 3. Translation was deduced to initiate at the first AUG at positions 59–61. There is a single open reading frame that is terminated after Leu-426 by a UAA codon. The UAA codon is followed by a 1053-bp 3′ untranslated region. Since the cDNA insert in pZD1 did not contain a 3′ poly(A)$^+$ tail, the sequence is truncated at the 3′ end. However, based on the size of poly(A)$^+$ RNA and a potential 3′ polyadenylylation signal at nucleotides 2341–2347, the nucleotide sequence shown in Fig. 3 should be nearly full-length. From the functional complementation of *E. coli purC* and *purEK* mutants and sequence homology with microbial genes encoding AIR carboxylase and SAICAR synthetase, discussed below, we conclude that the derived amino acid sequence shown in Fig. 3 is that of the bifunctional chicken AIR carboxylase–SAICAR synthetase.

Synthesis of AIR Carboxylase–SAICAR Synthetase in *E. coli*. The AIR carboxylase–SAICAR synthetase coding region was subcloned into a phage T7-based vector for expression in *E. coli* (11). For this purpose an *Nde* I site was constructed at the 5′ ATG, and an *Nde* I–*Hin*dIII fragment of the cDNA from nucleotide 59 to nucleotide 1738 was ligated into the corresponding sites in pT7-7 to yield plasmid pZD4. Labeling with [^{35}S]methionine detected a protein of ≈46 kDa from plasmid pZD4 but not from the pT7-7 vector. Rifampicin inhibited host-cell protein synthesis but not that of the 46-kDa protein (data not shown). The observed mass of 46 kDa is close to the value of 47,245 calculated from the derived amino acid sequence.

Structure–Function Relationship. Mutations were constructed *in vitro* to identify functional domains in the purCKE coding sequence. The results of an initial series of deletions that extended into the 3′ end of the cDNA are shown in Fig. 4. Deletions with endpoints in the cDNA at positions 1738, 1611, and 1391 retained PurC, PurK, and PurE function in *E. coli*. These results are consistent with the position of the translation termination codon at nucleotides 1337–1339. A further deletion into the coding sequence that removes amino acids downstream of Val-349 abolished PurK and PurE function, while all three functions were lost in a deletion that

```
                                                                    ggcgctcagtttctttcctgtgccttaggtgcctgtcgccgccgccgctgtcgccgcc    58
ATGGCGCCCCGCCGCGCCTCAGAACTGAAACTTGGTAAAAAAGTTAATGAGGGTAAAACGAAAGAAGTGTATGAGCTG  133
MetAlaProAlaAlaSerGluLeuLysLeuGlyLysLysValAsnGluGlyLysThrLysGluValTyrGluLeu      25
CCGGATATCCCGGGATGCGTTCTGATGCAGTCCAAGGACCAAATAACAGCGGGGAATGCCGCTAGGAAGGACCGG  208
ProAspIleProGlyCysValLeuMetGlnSerLysAspGlnIleThrAlaGlyAsnAlaAlaArgLysAspArg      50
ATGGAGGGGAAGGCTGCCATCTCCAACACCACCACCAGCTGCGTGTTCCAGCTGCTGCAGGAAGCAGGAATCAAA  283
MetGluGlyLysAlaAlaIleSerAsnThrThrThrSerCysValPheGlnLeuLeuGlnGluAlaGlyIleLys      75
ACGGCTTTTGTCAGGAAACAGAGTGACACAGCTTTCATAGCAGCTCACTGTGAAATGATCCCAATTGAATGGGTC  358
ThrAlaPheValArgLysGlnSerAspThrAlaPheIleAlaAlaHisCysGluMetIleProIleGluTrpVal     100
TGCAGAAGAATTGCTACTGGCTCTTTCCTCAAAAGAAACCCTGGTGTCAAAGAAGGCTATAAGTTTTACCCACCT  433
CysArgArgIleAlaThrGlySerPheLeuLysArgAsnProGlyValLysGluGlyTyrLysPheTyrProPro     125
AAAATTGAGATGTTTTACAAGGATGATGCCAATAATGATCCACAGTGGTCTGAGGAGCAGCTAATTGAAGCAAAA  508
LysIleGluMetPheTyrLysAspAspAlaAsnAsnAspProGlnTrpSerGluGluGlnLeuIleGluAlaLys     150
TTCTCTTTTGCTGGACTTACTATTGGCAAGACTGAAGTGGATATTATGGCTCGTTCTACTCAAGCTATTTTTGAG  583
PheSerPheAlaGlyLeuThrIleGlyLysThrGluValAspIleMetAlaArgSerThrGlnAlaIlePheGlu     175
ATCCTGGAAAAATCATGGCAGCCCCAAAACTGCACTCTGGTGGACCTGAAGATTGAATTTGGTGTTAATATTTTG  658
IleLeuGluLysSerTrpGlnProGlnAsnCysThrLeuValAspLeuLysIleGluPheGlyValAsnIleLeu     200
ACCAAAGAAATTGTTCTTGCTGATGTTATTGATAATGATTCATGGAGACTGTGGCCATCGGGAGACAGAAGCCAG  733
ThrLysGluIleValLeuAlaAspValIleAspAsnAspSerTrpArgLeuTrpProSerGlyAspArgSerGln     225
CAGAAGGACAAACAGTCCTATCGAGATCTGAAGGAAGTGACTCCTGAAGCATTGCAGATGGTTAAGAGAAACTTT  808
GlnLysAspLysGlnSerTyrArgAspLeuLysGluValThrProGluAlaLeuGlnMetValLysArgAsnPhe     250
GAATGGGTTGCAGAAAGAGTAGAGTTGCTTCTGAAAACAAAGAGCCAAGGTAGAGTTGTGGTATTGATGGGATCT  883
GluTrpValAlaGluArgValGluLeuLeuLeuLysThrLysSerGlnGlyArgValValValLeuMetGlySer     275
ACTTCTGACCTTGGCCACTGTGAGAAAATAAAGAAGGCATGTGCAACCTTTGGAATTCCTTGTGAGTTAAGAGTA  958
ThrSerAspLeuGlyHisCysGluLysIleLysLysAlaCysAlaThrPheGlyIleProCysGluLeuArgVal     300
ACCTCCGCTCACAAAGGGCCAGATGAAACTCTGAGGATCAAAGCAGAATATGAAGGAGATGGAATCCCGACTGTG 1033
ThrSerAlaHisLysGlyProAspGluThrLeuArgIleLysAlaGluTyrGluGlyAspGlyIleProThrVal     325
TTTGTTGCAGTAGCTGGCAGAAGCAATGGTTTAGGGCCAGTAATGTCTGGTAACACTGCTTACCCTGTTGTCAAC 1108
PheValAlaValAlaGlyArgSerAsnGlyLeuGlyProValMetSerGlyAsnThrAlaTyrProValValAsn     350
TGTCCTCCCCTCTCATCTGACTGGGGTGCTCAGGATGTGTGGTCCTCTCTCAGACTGCCCAGTGGTCTTGGCTGT 1183
CysProProLeuSerSerAspTrpGlyAlaGlnAspValTrpSerSerLeuArgLeuProSerGlyLeuGlyCys     375
CCTACTACTCTGTCACCTGAAGGAGCTGCTCAGTTTGCTGCCCAGATTTTTGGTTTAAACAACCACTTGGTATGG 1258
ProThrThrLeuSerProGluGlyAlaAlaGlnPheAlaAlaGlnIlePheGlyLeuAsnAsnHisLeuValTrp     400
GCCAAACTGCGATCAAACATGTTAAATACATGGATCTCTTTGAAGCAGGCTGACAAAAAATTGCGGGAGTGCACC 1333
AlaLysLeuArgSerAsnMetLeuAsnThrTrpIleSerLeuLysGlnAlaAspLysLysLeuArgGluCysThr     425
TTGtaagtccaaccaacaagtaactcctcctcgctacacaaagatagtggcgtgcatatgcatttgtatcaggat 1408
Leu*                                                                            426

ttgcagttggatgatgtgagacatcttttcttacagaaggaagcatcttcagagcagcagctgagtctgctatgt 1483
cctcattccttttcttgtgtatgtttttggttttttagtatagccacttagtaaaagatgagaaggtaatattt 1558
tagttatccttctgttgatccaaaatagtgctgttctgctctttagtgtatactgttaactactgatgatggtta 1633
tgtcaacacaatgattgatttggtgtgatttaaatagaaaccccctttttttttcccccctgtaacagtaacattct 1708
tccatagtaataatctatgtcctgaagcttattaactcaatgggactgctaactaaaactagtgctgatgtagtg 1783
ctgtaataatggtatgctttcttttagcaataaaatggcaacatatgctgtattaataagaaaataggctaagaa 1858
tgagtaaaagcctgaaagaaataactgccttatttaaaagaaaaacatatatttttttcatgtgtatgtgtgtatg 1933
tatatacaccccatcagtaactttaccacatcatttgagtgagttaacagttctcttgtttgtgagaagctctaa 2108
ttttgaatggaaccataactaggcttgttgcctacttgctgtataactcttgactggtccagtcttttgtccttc 2183
aaatagtagacagatctactgtgtatgtatgtgatgctgtcaaattatccttatacacatcaatgtatctatatt 2258
ccctcactctgtttctatgggaattaaactgaactgaaggtgggaataaacagtaaagcaacagaatacacgggc 2333
aagcctaagcatgactgtcttgntctttgcacacaacattctttaaagtaggtactgattccaagaaatatgcaa 2308
ttactcaatgtatgtacaatttatgcttttataattaaataaactatttaatgcaaggaattcgatatcaagctt 2383
atcgatacc                                                                      2392
```

FIG. 3. Nucleotide sequence of chicken liver AIR carboxylase–SAICAR synthetase cDNA and derived amino acid sequence.

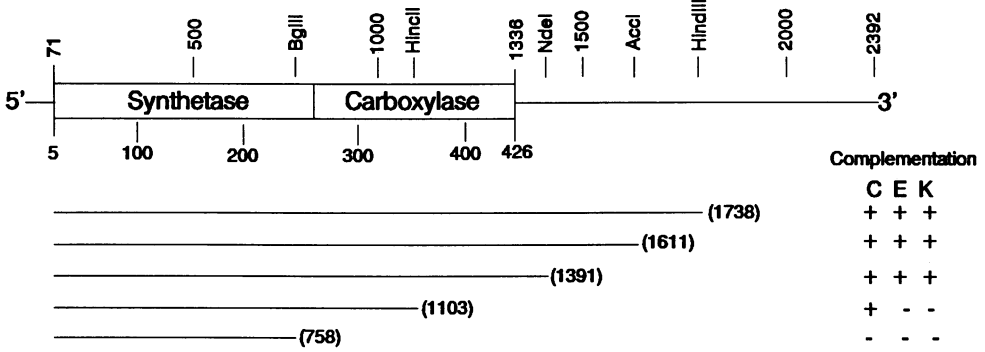

FIG. 4. Structure–function analysis. Deletions were constructed in pZD1 by removal of specific restriction fragments. Enzyme function was determined by complementation of *E. coli pur* mutants. Representation of cDNA: 5′ thin line, *lacZ′* and polylinker; open box, AIR carboxylase–SAICAR synthetase coding region; 3′ thin line, 3′ flanking sequence. Nucleotides are numbered above the line, and amino acids, below the line as in Fig. 3. Complementation of *E. coli pur* mutants is shown at the right. The intact cDNA complemented *purC*, *purE*, and *purK* strains.

extends to nucleotide 758 and removes all amino acids distal to Arg-233. Thus, the overall organization is NH_2–SAICAR synthetase–AIR carboxylase–COOH, with AIR carboxylase function extending beyond Asp-234.

To further define the SAICAR synthetase and AIR carboxylase boundaries, a series of amino acid replacements and small deletions were constructed by oligonucleotide-directed mutagenesis. Mutations were constructed at positions that are conserved in the microbial sequences and near the synthetase–carboxylase boundary (see Fig. 5). The function of mutated enzymes expressed from multicopy plasmids was assessed by measurement of growth rates in *E. coli pur* mutants. In this assay, stimulation of the growth rate by adenine indicates that the function of the chicken enzyme is limiting for purine nucleotide synthesis. The analysis of these mutations is given in Table 1. Basal level function for the unmutagenized plasmid is shown in the first line. With non-mutagenized cDNA, excess SAICAR was synthesized and complemented the *purC* mutant, but synthesis of 5-amino-4-imidazolecarboxylate ribonucleotide (AICR in Fig. 1) was limiting, and the growth rates of purE and purK mutants were stimulated by adenine. Two deletion mutations, $\Delta(Gly^{17}, Lys^{18})$ and $\Delta(Glu^{255}, Arg^{256}, Val^{257}, Glu^{258})$ specifically abolished complementation of *purC* (SAICAR synthetase), whereas the double-replacement mutation Ser-302 → Ala, His-304 → Tyr and the double-deletion mutation $\Delta(Ala^{330}, Gly^{331})$ abolished complementation of *purEK* (AIR carboxylase). These results are consistent with the deletion analysis and indicate that the SAICAR domain extends from the NH_2 terminus to at least Glu-255 and that amino acids between Ser-302 and the COOH terminus are needed for AIR carboxylase activity. In neither of the AIR carboxylase mutations shown in Table 1 nor in the 3′ deletions were *purE* and *purK* functions differentially inactivated. The basis for

FIG. 5. Alignment of chicken (*Gallus gallus*) AIR carboxylase–SAICAR synthetase and microbial sequences: *E. coli purC* (J. M. Smith, personal communication). *B. subtilis purC* and *purE* (5), *E. coli purE* (3), and *Schizosaccharomyces pombe ADE6* (21) were aligned by using FASTA (22).

Table 1. Functional analysis of mutant AIR carboxylase–SAICAR synthetase in *E. coli*

Enzyme*	Relative growth rate†		
	purC	*purE*	*purK*
Wild type	1.0	0.80	0.48
Mutant			
Δ(Gly^{17}, Lys^{18})	<0.05‡	0.75	0.67
Asp-134 → Val, Ala-135 → Val	0.68	0.80	0.81
Trp-218 → Leu	0.81	0.92	0.75
Δ(Glu^{255}, Arg^{256}, Val^{257}, Glu^{258})	0.08	0.92	0.92
Ser-302 → Ala, His-304 → Tyr	0.92	<0.05	<0.05
Δ(Ala^{330}, Gly^{331})	1.0	<0.05	<0.05

*Enzymes are fusions in which the *lacZ'* and polylinker sequence from pBluescript SK$^-$ is ligated to the codon for Ala-5 in pZD1. Deletion mutations are designated by Δ.
†Growth rate in medium without adenine divided by the growth rate in medium containing adenine.
‡No growth in 10 hr.

stimulation of PurK function by mutagenesis of the PurC domain is not understood.

DISCUSSION

We have screened a chicken liver cDNA library using functional complementation of *E. coli pur* mutants to isolate a cDNA that encodes AIR carboxylase–SAICAR synthetase. In addition, clones have been isolated that complement *purDMN* (trifunctional GAR synthetase–AIR synthetase–GAR transformylase), *purB* (adenylosuccinate lyase) and the avian homolog of *purH*. Complementation of *purF* and *purL* were not obtained by this procedure. To our knowledge, of the steps for *de novo* purine nucleotide synthesis in birds or mammals, only the cDNA encoding IMP dehydrogenase (23) from Chinese hamster had been isolated previously, although after completion of this work, we learned that Schild *et al.* (24) had cloned AIR carboxylase–SAICAR synthetase and two other cDNAs for *de novo* purine nucleotide biosynthesis from a human HepG2 cDNA library by functional complementation of yeast mutants.

The nearly full-length cDNA of 2392 bp encodes a bifunctional AIR carboxylase–SAICAR synthetase of calculated M_r 47,245, which is less than the value of 52,000 reported for the enzyme purified from chicken liver (25). A deletion analysis shown in Fig. 4 shows that sequence downstream of an *Nde* I site at position 1391 is not required for enzyme function in *E. coli* and supports the position of translation termination following Leu-426.

The assignment of structural domains in chicken AIR carboxylase–SAICAR synthetase has been made by mutational analyses and sequence alignments with bacterial and yeast genes. The results indicate an organization of SAICAR synthetase at the NH_2 terminus followed by AIR carboxylase. The amino acid sequence alignment in Fig. 5 shows correspondence of residues 1–259 of the chicken enzyme with amino acids throughout the entire length of the SAICAR synthetase from *E. coli* and *B. subtilis*. In this alignment there are 31% amino acid identities of the chicken sequence with the 237-residue *purC*-encoded SAICAR synthetase from *E. coli*. Statistical analysis by the FASTA program (22) gives a significance score of 32 SD above the mean of randomized sequences, a highly significant value for this alignment. Amino acids 255–258 in the chicken enzyme are required for SAICAR synthetase activity (Table 1) and should correspond to the COOH-terminal segment of this domain. According to the alignment in Fig. 5, the NH_2-terminal segment of the AIR carboxylase domain starts near residue 260. There is 27–30% identity of amino acids 260–426 of the chicken enzyme with the *purE*-encoded sequences from *E. coli* and *B. subtilis* and 30% identity with the 176 amino acid COOH-terminal PurE domain in yeast ADE6. This overall identity suggests that chicken SAICAR synthetase–AIR carboxylase residues 1–258 evolved from a *purC*-related ancestral gene, and residues 260–427 evolved from a *purE*-related ancestor. Therefore, it is obvious that the chicken AIR carboxylase is unique and does not contain a domain similar to the *purK*-encoded subunit of 356–379 amino acids as in the *E. coli* or *B. subtilis* enzymes or a domain of similar size in the yeast enzyme. Attempts to obtain a statistically significant alignment of microbial PurK sequences with that of the avian enzyme sequence were unsuccessful and support the conclusion that the AIR carboxylase–SAICAR synthetase does not contain the equivalent of a microbial PurK domain. The microbial AIR carboxylase PurK subunit is thought to function in CO_2 binding (3). Either chicken AIR carboxylase can function without a PurK domain or this function has been incorporated into the PurE domain. In any event, avian PurE-like AIR carboxylase can carry out reaction 6 in the pathway and thus complements both *purE* and *purK* mutations. Expression of cloned microbial genes and the avian cDNA now permits a detailed investigation of the mechanism of the AIR carboxylase reaction.

We thank Junko Aimi, Carolyn Minth, Ronda Rolfes, Honqi Wang, John Williams, Limin Zheng, and Gaochao Zhou for assistance and David Schild for providing a copy of his manuscript prior to publication. This work was supported by Grant GM24658 from the Public Health Service (H.Z.) and Grant 18024 (J.E.D.) from the National Institute of Diabetes and Digestive and Kidney Diseases. Computer facilities were supported by National Institutes of Health Grant AI27713. This is journal paper no. 12361 from the Purdue University Agricultural Experiment Station.

1. Tso, J. Y., Zalkin, H., vanCleemput, M., Yanofsky, C. & Smith, J. M. (1982) *J. Biol. Chem.* **260**, 3525–3531.
2. Smith, J. M. & Daum, H. A., III (1986) *J. Biol. Chem.* **261**, 10632–10636.
3. Tiedeman, A. A., Keyhani, J., Kamholz, J., Daum, H. A., Gots, J. S. & Smith, J. M. (1989) *J. Bacteriol.* **171**, 205–212.
4. Schendel, F. J., Mueller, E., Stubbe, J., Shiau, A. & Smith, J. M. (1989) *Biochemistry* **28**, 2459–2471.
5. Ebbole, D. J. & Zalkin, H. (1987) *J. Biol. Chem.* **262**, 8274–8287.
6. Rolfes, R. J. & Zalkin, H. (1988) *J. Biol. Chem.* **263**, 19653–19661.
7. Ebbole, D. J. & Zalkin, H. (1989) *J. Biol. Chem.* **264**, 3553–3561.
8. Henikoff, S. (1987) *BioEssays* **6**, 8–13.
9. Boss, G. R. & Seegmiller, J. E. (1982) *Annu. Rev. Genet.* **16**, 297–328.
10. Short, J. M., Fernandez, J. M., Sorge, J. A. & Huse, W. D. (1988) *Nucleic Acids Res.* **16**, 7583–7600.
11. Tabor, S. & Richardson, C. C. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1074–1078.
12. Turpen, T. H. & Griffith, O. M. (1986) *Biotechniques* **4**, 13–15.
13. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
14. Zheng, L., Andrews, P. C., Hermodson, M. A., Dixon, J. E. & Zalkin, H. (1990) *J. Biol. Chem.* **265**, 2814–2821.
15. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239**, 487–491.
16. Frohman, M. A., Dush, M. K. & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8998–9002.
17. Henikoff, S. (1984) *Gene* **28**, 351–359.
18. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
19. Kunkel, T. A., Roberts, J. D. & Zakour, R. A. (1987) *Methods Enzymol.* **154**, 367–382.
20. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5201–5205.
21. Szankasi, P., Heyer, W.-D., Schuchert, P. & Kohli, J. (1988) *J. Mol. Biol.* **204**, 917–925.
22. Lipman, D. J. & Pearson, W. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
23. Collart, F. R. & Huberman, E. (1988) *J. Biol. Chem.* **263**, 15769–15772.
24. Schild, D., Brake, A. J., Kiefer, M. C., Young, D. & Barr, P. J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 2916–2920.
25. Patey, C. A. H. & Shaw, G. (1973) *Biochem. J.* **135**, 543–545.
26. Aimi, J., Badylak, J., Williams, J., Chen, Z., Zalkin, H. & Dixon, J. E. (1990) *J. Biol. Chem.*, in press.