Deduced Amino Acid Sequence of *Escherichia coli* Adenosine Deaminase Reveals Evolutionarily Conserved Amino Acid Residues: Implications for Catalytic Function[†,‡]

Zengyi Chang,[§] Per Nygaard,[∥] A. Craig Chinault,[§,⊥] and Rodney E. Kellems*[,§,⊥]

Verna and Marrs McLean Department of Biochemistry and Institute for Molecular Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, and Enzyme Division, University Institute of Biological Chemistry B, Copenhagen, Denmark

Received July 30, 1990; Revised Manuscript Received November 6, 1990

ABSTRACT: The goal of the research reported here is to identify evolutionarily conserved amino acid residues associated with enzymatic deamination of adenosine. To do this, we isolated molecular clones of the *Escherichia coli* adenosine deaminase gene by functional complementation of adenosine deaminase deficient bacteria and deduced the amino acid sequence of the enzyme from the nucleotide sequence of the gene. Nucleotide sequence analysis revealed the presence of a 996-nucleotide open reading frame encoding a protein of 332 amino acids having a molecular weight of 36 345. The deduced amino acid sequence of the *E. coli* enzyme has approximately 33% identity with those of the mammalian adenosine deaminases. With conservative amino acid substitutions the overall sequence homology approaches 50%, suggesting that the structures and functions of the mammalian and bacterial enzymes are similar. Additional amino acid sequence analysis revealed specific residues that are conserved among all three adenosine deaminases and four AMP deaminases for which sequence information is currently available. In view of previously published enzymological data and the conserved amino acid residues identified in this study, we propose a model to account for the enzyme-catalyzed hydrolytic deamination of adenosine. Potential catalytic roles are assigned to the conserved His 214, Cys 262, Asp 295, and Asp 296 residues of mammalian adenosine deaminases and the corresponding conserved amino acid residues in bacterial adenosine deaminase and the eukaryotic AMP deaminases.

The catalytic power of an enzyme is measured by its ability to enhance the rate of a specific chemical reaction. One of the highest rate enhancements known is that achieved by adenosine deaminase (EC 3.5.4.4), an enzyme that catalyzes the hydrolytic deamination of adenosine with a rate enhancement of approximately 2×10^{12} (Frick et al., 1987). The catalytic power of adenosine deaminase is encompassed within a single polypeptide chain and is achieved without the aid of cofactors. Thus, all of the functional groups required to catalyze the hydrolytic deamination of adenosine reside within the properly configured amino acid residues of a single polypeptide chain. The amino acid sequences of the murine (Yeung et al., 1985) and the human (Wiginton et al., 1984) adenosine deaminases have been deduced from the nucleotide sequence of the cloned cDNAs. Functional studies have indicated that adenosine deaminases from different sources are similar with regard to substrate specificity and kinetic constants (Wolfenden, 1969; Wolfenden et al., 1969; Agarwal et al., 1975; Spector et al., 1983). Because of the apparent simplicity and catalytic efficiency of adenosine deaminase, the enzyme has been the focus of extensive investigations in an effort to understand the mechanism of catalysis. A variety of studies have implicated an enzyme sulfhydryl (Ronca et al., 1967; Wolfenden et al., 1967; Weiss et al., 1987), one or more carboxyl groups (Kurz & Frieden, 1983; Kati & Wolfenden, 1989a), and one or more imidazole groups (Orsi et al., 1972; Kurz & Frieden, 1983) in the catalytic process. The availability of a variety of substrates and inhibitors (Chassy & Suhadolnik, 1967; Cha et al., 1975; Wolfenden et al., 1977; Agarwal et al., 1977) has played an important role in efforts to understand substrate specificity and catalytic mechanism. The transition state of the enzyme-catalyzed reaction appears to involve a tetrahedral intermediate (Wolfenden, 1969; Evans & Wolfenden, 1970, 1973; Frick et al., 1986; Kurz & Frieden, 1987; Jones et al., 1989) which is formed during an addition–elimination reaction at the sixth position of the purine ring. Recent studies illustrate the importance of the hydroxyl group at the tetrahedral carbon of transition-state analogues for binding to the enzyme (Schramm & Baker, 1985; Jones et al., 1989; Kati & Wolfenden, 1989a).

Recombinant DNA technology has contributed to the structural and functional analysis of adenosine deaminase in two ways. First, molecular cloning and sequencing of murine (Yeung et al., 1985) and human (Wiginton et al., 1984) adenosine deaminase cDNAs allowed the amino acid sequence of each enzyme to be deduced. Second, the construction of bacteria capable of overproducing murine adenosine deaminase has made it possible to purify large quantities of the enzyme for crystallization and X-ray diffraction analysis (Wilson et al., 1988). Genetic engineering also allows the introduction of specific amino acid substitutions in proteins of interest by site-specific mutagenesis (Leatherbarrow & Fersht, 1987). However to utilize this last approach, it is important to have some rational basis for choosing the specific amino acid changes to be made. For example, such choices can be made on the basis of detailed structural information that reveals amino acids present at the active site that are likely to provide functional groups that participate in the catalytic pathway.

[†] This research was supported by Grant AI25255 from the National Institutes of Health.

[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J05320.

[§] Verna and Marrs McLean Department of Biochemistry, Baylor College of Medicine.

[∥] Enzyme Division, University Institute of Biological Chemistry B.

[⊥] Institute for Molecular Genetics, Baylor College of Medicine.

0006-2960/91/0430-2273$02.50/0

However, such high-resolution structural information is not yet available for adenosine deaminase. An alternative and completely independent approach which can often reveal the identity of functionally important amino acid residues or sequence motifs is the identification of evolutionarily conserved amino acid sequences among functionally related proteins (Reichardt & Berg, 1988). For this reason we have cloned the *Escherichia coli* adenosine deaminase gene, deduced the amino acid sequence of the enzyme, and identified four amino acid sequences, of at least four amino acid residues in length, which have remained unchanged since the evolutionary divergence of bacteria and mammals. Two of these amino acid sequences are also conserved in four AMP deaminases, enzymes which also catalyze a hydrolytic deamination from the sixth position of an adenine ring. Additional sequence analysis revealed that a histidine, a cysteine, and a pair of aspartic acid residues are conserved among the three adenosine deaminases (bacterial, mouse, and human) and the four AMP deaminases (yeast, rat, and two human isoforms) for which sequence information is currently available. On the basis of a consideration of previously published enzymolgoical data and the evolutionarily conserved amino acids identified in this study, we propose a mechanism for the enzyme-catalyzed deamination of adenosine. Potential catalytic roles are assigned to four specific amino acid residues; His 214, Cys 262, Asp 295, and Asp 296 of the mammalian adenosine deaminases and the corresponding conserved amino acid residues in *E. coli* adenosine deaminase and four eukaryotic AMP deaminases.

Materials and Methods

Bacterial Strains and Growth Conditions. Sϕ3834 (*rpsL*, Δ*add-uid-man*, *metB*, *guaA*, *uraA*::Tn10) is an *E. coli* strain which has two features important to this work: one is the deletion of the *add* gene (Jochimsen et al., 1975), and the other is a mutation of the *guaA* gene (Hove-Jensen & Nygaard, 1989). Sϕ3834 was constructed as follows: *E. coli* strain Sϕ3822 was derived from Sϕ198 (*rpsL*, *metB*, *guaA*) by P1 transduction with phage grown on Sϕ1608 (HfrR4, *pyrA*35, *purM*48, *metB*1, *uraA*::Tn10). Colonies were selected on L-broth with 5 μg/mL tetracycline. The resulting strain is designated Sϕ3822 (*rpsL*, *metB*, *guaA*, *uraA*::Tn10). Strain Sϕ3834 was derived from strain Sϕ230 (*rpsL*, Δ*add-uid-man*, *metB*) by P1 transduction with phage grown on strain Sϕ3822. Again, colonies were selected on L-broth plus 5 μg/mL tetracycline. Note that *uraA* encodes a protein for uracil transport at low concentrations and has been disrupted by a Tn10 insertion in strain Sϕ1608. It is not relevant to the present study but was used to move the closely linked *guaA* gene into Sϕ3822 by P1 transduction.

Minimal medium contained (per liter of salt buffer) 50 mg of methionine, 2 mL of glycerol, 0.2 g of $MgSO_4$, 0.01 g of $CaCl_2$, and 1 mg of thiamin. The salt buffer contained (per liter) 13.6 g of KH_2PO_4, 2 g of $(NH_4)_2SO_4$, and 0.0005 g of $FeSO_4 \cdot 7H_2O$ with the pH adjusted to 7.2 by KOH. When required, guanine, hypoxanthine, and uracil were added at 15 μg/mL. When DAP[1]-supplemented minimal medium was made, 20 mg of 2,6-diaminopurine was added to the above solution. Antibiotics, when appropriate, were also added. Solutions were sterilized by either autoclaving or filtration. All the chemicals were from Sigma Chemical Co.

Construction and Screening of Plasmid Libraries. A plasmid library of *E. coli* genomic DNA was made by insertion of fractionated 8–15-kb DNA fragments from *Sau*3A partially digested *E. coli* genomic DNA into *Bam*HI-digested and phosphatase-treated pUC9 plasmid vector (Jayakumar et al., 1989). The plasmid library was transformed into the Sϕ3834 cells and selected on L-broth/ampicillin plates. The colonies thus obtained were pooled and reselected on DAP minimal medium plates. The secondary plasmid library was made similarly. Plasmid DNA from the positive clone (padd7.7; see Results) was partially digested with *Sau*3A, and 1–3-kb DNA fragments were purified with GeneClean (Bio 101) and inserted into *Bam*HI-digested and phosphatase-treated Bluescript II KS+ plasmid (Stratagene, La Jolla, CA). The library was transformed into competent Sϕ3834 cells and selected on DAP minimal medium.

Detection of Adenosine Deaminase Activity. After growing in 10 mL of L-broth-rich medium overnight at 37 °C, cells were pelleted and resuspended in 200 μL of homogenization buffer (Nygaard, 1978) with 1 mM phenylmethanesulfonyl fluoride. The cells were lysed by sonication (Branson sonicator–cell disruptor 185). Cell debris was eliminated by ultracentrifugation (Beckman, TL-100) at 45 000 rpm for 30 min. The resulting supernatant was assayed for adenosine deaminase activity. All operations were performed at 4 °C.

The adenosine deaminase activity was analyzed by both spectrophotometry (Nygaard, 1978) and zymogram analysis (Yeung et al., 1983). The spectrophotometric assay is based on the decrease of absorption at 265 nm resulting from the deamination of adenosine to inosine. The zymogram analysis is based on a specific histochemical reaction that is dependent on adenosine deaminase activity and was used to identify adenosine deaminase activity after electrophoretic separation of supernatant fractions under nondenaturing conditions on an Authentifilm thin agarose gel (Innovative Chemistry) in a chilled electrophoresis chamber (100 V, 45 min).

DNA Sequencing. *Exo*III/mung bean nuclease deletions were made according to the supplier's recommendations [Stratagene and Sambrook et al. (1989)] except for the following modifications. The plasmid DNA was completely digested with the 3′-overhang restriction enzyme first and then cut with the 5′-overhang enzyme [instead of vice versa as described by Sambrook et al. (1989)]. After mung bean nuclease treatment, the DNA was run on a low-melting agarose gel, and the DNA fragments were purified with GeneClean for better ligation. Plasmid DNA from minilysate preparations was used for double-strand dideoxy termination sequencing. The boiling method of preparing minilysates was used (Ausubel et al., 1987), with two phenol/chloroform/isoamylalcohol (25:24:1) extractions before and after the 2-propanol precipitations. The double-strand sequencing was performed by the Sequenase method according to the supplier's recommendation (U.S. Biochemicals). The reverse primer and universal primer of Bluescript II KS+ plasmid were used for sequencing the two complementary DNA strands. Sequence analysis was performed with the MicroGenie sequence analysis program (Version 5.5, Beckman Instruments, Inc.).

Results

Cloning Strategy: Functional Complementation of Adenosine Deaminase Deficient Bacteria. The phenotypic selection strategy we used stems directly from the approach used previously to clone full-length murine adenosine deaminase cDNA by functional complementation of an adenosine deaminase deficient *E. coli* strain (Yeung et al., 1985). However, the particular *E. coli* strain used in the earlier work, Sϕ200, was not entirely satisfactory due to a relatively high reversion frequency. To alleviate this problem, we constructed a new

[1] Abbreviations: EHNA, *erythro*-9-(2-hydroxy-3-nonyl)adenine; DAP, 2,6-diaminopurine; GR, guanosine; G, guanine.

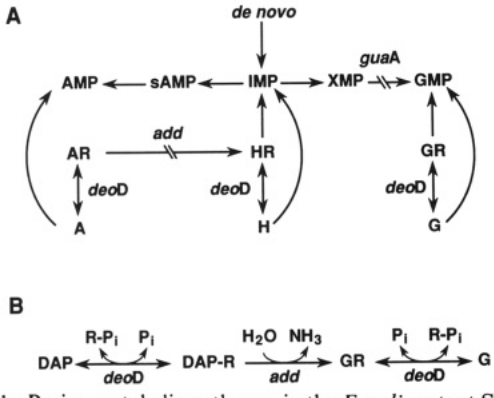


FIGURE 1: Purine metabolic pathways in the *E. coli* mutant Sϕ3834. Panel A depicts an abbreviated version of purine metabolism in *E. coli*. Sϕ3834 (*rpsL*, Δ*add-uid-man*, *metB*, *guaA*, *uraA*::Tn10) is a guanine-requiring nutritional auxotroph which has a deletion in the adenosine deaminase gene (*add*). The mutation in the *guaA* gene results in the loss of the guanylate synthetase activity and provides the basis for the guanine (G) or guanosine (GR) reqiurement. Additional abbreviations: AR, adenosine; A, adenine; HR, inosine; H, hypoxanthine. Panel B shows the pathway leading to the synthesis of guanine and guanosine from 2,6-diaminopurine (DAP) and ribose 1-phosphate (RPi) by adenosine deaminase (*add*) and purine nucleoside phosphorylase (*deoD*). Sϕ3834 can grow on DAP if adenosine deaminase is provided because GMP can be synthesized from guanosine or guanine as indicated.

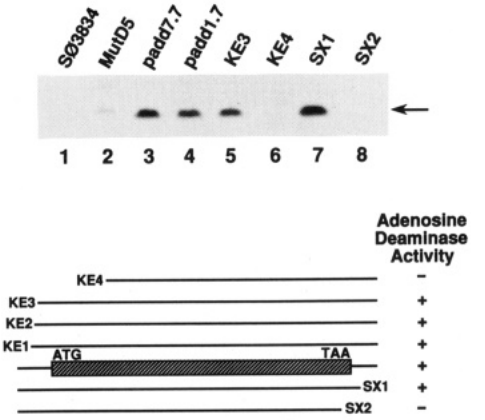


FIGURE 2: Zymogram analysis of bacterial transformants to test for the presence of adenosine deaminase activity. Zymogram analysis of adenosine deaminase activity in cell lines and transformants is shown in the upper panel, and the deletion mutants used to define the boundaries of the *add* gene are shown in the lower panel. Lane 1 is from the adenosine deaminase deficient Sϕ3834 cells and serves as a negative control for bacterial adenosine deaminase activity. Lane 2 is from an adenosine deaminase positive strain, MutD5 (Schaaper, 1988), and serves as a positive control for adenosine deaminase activity. Lanes 3 and 4 are from the two positive clones selected for growth on DAP minimal medium. Lanes 5–8 are from the deletion mutants shown at the bottom. The position of the open reading frame between the ATG and TAA sequences is indicated by the hatched box in the lower panel.

mutant, Sϕ3834 (*rpsL*, Δ*add-uid-man*, *metB*, *guaA*, *uraA*::Tn10). The Δ*add-uid-man* refers to a deletion of the adenosine deaminase gene (*add*) and two neighboring genes (Jochimsen et al., 1975). The *guaA* mutation renders this strain a guanine-requiring nutritional auxotroph by preventing the conversion of XMP to GMP. The reversion frequency for the *guaA* mutation is less than 10^{-8}. The relevant *E. coli* purine metabolic pathways and the specific mutations affecting these pathways in Sϕ3834 are shown in Figure 1A (Nygaard, 1983).

The metabolic basis for the phenotypic selection protocol is as follows. In the presence of purine nucleoside phosphorylase (encoded by *deoD*) and adenosine deaminase (encoded by *add*), guanosine (GR) and guanine (G) can be produced from 2,6-diaminopurine (DAP) as shown in Figure 1B. The resulting guanosine and guanine can be converted directly to GMP as shown in Figure 1A, thus circumventing the genetic block to GMP synthesis resulting from the *guaA* mutation. In strain Sϕ3834, a *guaA* mutant lacking an *add* gene, DAP cannot satisfy the guanine nucleotide needs of the cell unless a functional *add* gene is provided. Thus, Sϕ3834 cells grow on minimal medium when supplemented with guanosine or guanine, but not with DAP. However when transformed with a recombinant plasmid encoding mouse or human adenosine deaminase, Sϕ3834 cells grow on minimal medium supplemented with DAP (data not shown). Under these conditions such transformants are killed by the addition of inhibitors of adenosine deaminase such as EHNA [*erythro*-9-(2-hydroxy-3-nonyl)adenine], purine riboside, or deoxycoformycin (data not shown), indicating that adenosine deaminase is essential for growth. Thus, the DAP selection medium appeared to be suitable for cloning the *E. coli add* gene by functional complementation of Sϕ3834.

Complementation Cloning of the Bacterial add Gene. A plasmid library of *E. coli* genomic DNA (Jayakumar et al., 1989) was constructed and used to transform the adenosine deaminase deficient mutant, Sϕ3834. Transformants were initially selected on L-broth plus ampicillin. Transformants were then pooled and replated on minimal medium plus DAP, and one positive clone was obtained. This clone, padd7.7, has an insert of approximately 7.7 kb of genomic DNA. The following two observations indicate that the padd7.7 contains the *add* gene rather than the *guaA* gene (which may be recovered by our selection scheme). First, Sϕ3834 cells transformed by padd7.7 did not grow on minimal medium without supplementation with DAP; if the *guaA* gene had been cloned, the transformants would have been able to grow on minimal medium alone (see Figure 1A). Second, a high level of adenosine deaminase activity was detected in cell extracts from the transformed cells (Figure 2, lane 3). The much higher levels of adenosine deaminase activity in cells transformed with padd7.7, as compared to a control bacterial extract (MutD5, Figure 2, lane 2), presumably results from gene dosage effects reflecting the presence of multiple copies of padd7.7 present per transformed cell.

The *add* gene was further delineated by preparing a secondary plasmid library from padd7.7. This was accomplished by *Sau*3A partial digestion of padd7.7, gel purification of DNA fragments in the size range of 1–3 kb, and insertion into *Bam*H1-digested and phosphatase-treated Bluescript II KS+ vector DNA. The resulting recombinant plasmids were introduced into Sϕ3834, and transformants were selected on minimal medium plus DAP. Plasmid minilysates were prepared from a number of transformants, and one transformant, padd1.7, was identified as having the smallest insert of approximately 1.7 kb. The presence of the *add* gene was indicated by the presence of adenosine deaminase activity in cell extracts (Figure 2, lane 4). The location of the *add* gene within padd1.7 was determined by preparing a nested set of 5′ and 3′ deletion mutants (a subset of which is shown in Figure 2, lower panel) and testing extracts of transformed Sϕ3834 for adenosine deaminase activity. From the results shown in Figure 2 we conclude that one end of the *add* gene lies between the end points of subclones KE3 and KE4, whereas the other end of the *add* gene is situated between the end points of subclones SX1 and SX2.

Nucleotide Sequence of the E. coli add Gene. Unambiguous sequence from both strands of padd1.7 was obtained for the region believed to encode adenosine deaminase. The sequence

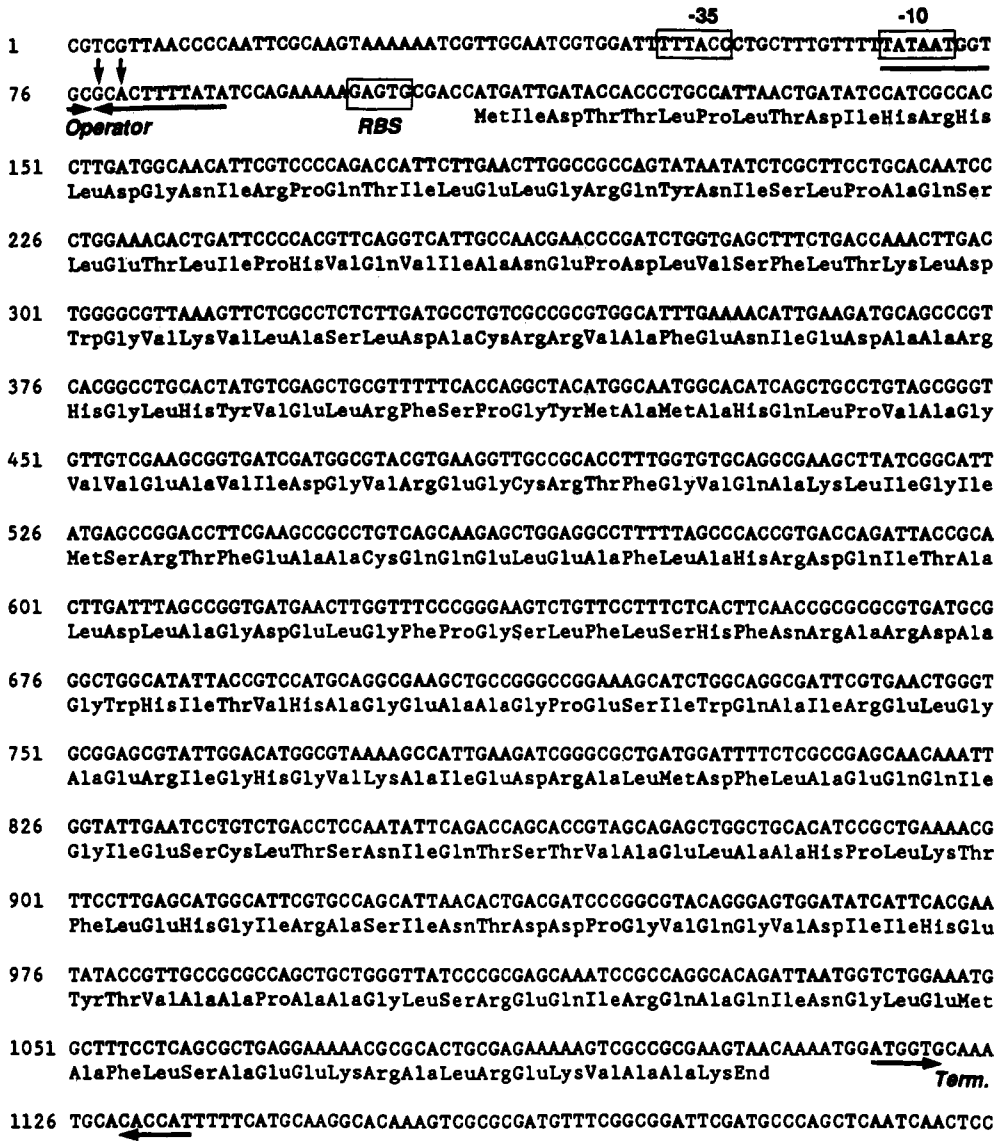
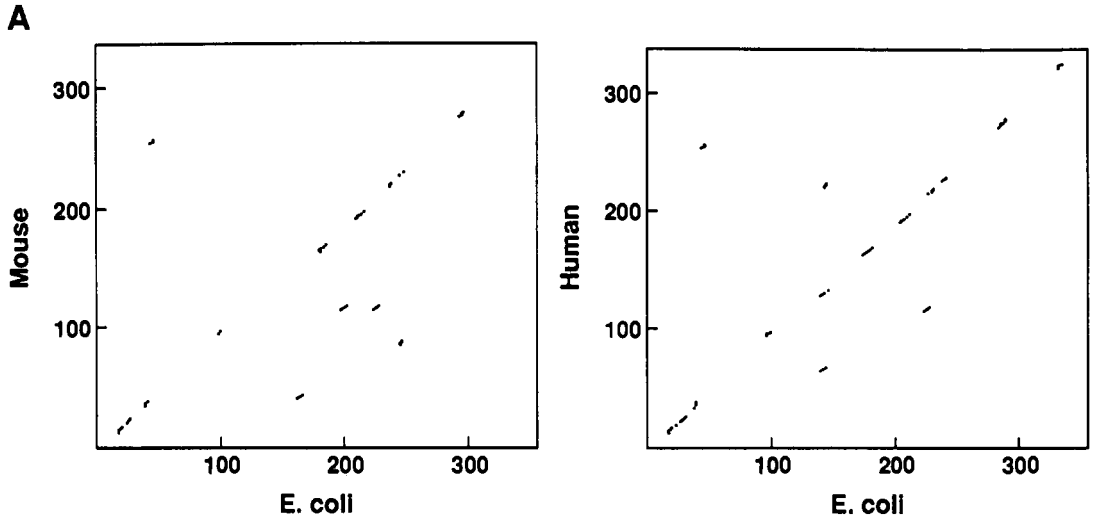


FIGURE 3: Nucleotide sequence of the *E. coli* adenosine deaminase (*add*) gene and flanking region. Shown below the nucleotide sequence is the amino acid sequence encoded by the open reading frame. Putative regulatory elements are indicated. These include promoter elements at −10 and −35, an operator which consists of a region of dyad symmetry (inverted arrows) flanking potential transcriptional initiation sites (downward arrows), a ribosome binding site (RBS), and ρ-independent terminator (Term.) possessing elements of dyad symmetry (indicated by the inverted arrows).

(Figure 3) included a 996-nucleotide open reading frame, bounded on one end by a translation start codon (ATG) and on the other end by a translation stop codon (TAA). This open reading frame encodes a protein of 332 amino acids having a molecular weight of 36 345 and lies within the boundaries delineated in Figure 2 for the functional *add* gene. The deduced amino acid sequence shows considerable similarity to those of mouse and human adenosine deaminase (Figure 4). The *E. coli* sequence shows 33% identity and 49% similarity with that of the mouse adenosine deaminase sequence and 32% identity and 47% similarity with that of the human adenosine deaminase sequence. On the basis of the functional evidence presented in Figure 2 and the sequence similarities illustrated in Figure 4, we conclude that the 996-nucleotide open reading frame in padd1.7 encodes *E. coli* adenosine deaminase.

Sequence motifs which may serve as promoter, operator, ribosome binding site, and ρ-independent transcription terminator for the *add* gene are depicted in Figure 3. Procaryotic promoters usually contain two consensus hexamers (McClure, 1985; Hawley & McClure, 1983): one is the −10 TATAAT sequence; the other is the −35 TTGACA. The putative −10 sequence for the *E. coli add* gene is completely conserved, whereas the putative −35 sequence (TTTACC) is less well conserved. In *E. coli*, transcription usually startes at a G or an A nucleotide which is four to eight bases downstream from the −10 sequence (McClur, 1985; Hawley & McClure, 1983). Thus, the G and A nucleotides, which are five and seven bases downstream of the −10 sequence, are potential transcription start sites for the *add* gene (shown as downward arrows in Figure 3). An operator usually consists of a DNA sequence having dyad symmetry which encompasses part of the promoter sequence and part of the transcription start site (McClure, 1985). On the basis of these criteria, the sequence TATAATGGTGC GCAGTTTTATA, which has partial dyad symmetry and covers the putative promoter and transcription start site, may serve as an operator that could be involved in the metabolic regulation of *add* gene expression (Nygaard, 1978). The consensus sequence for ribosome binding, GAGG or AGGA, is usually centered 8–13 bases upstream of the translation initiation codon (Gold et al., 1981). The GAGTG sequence, which is centered eight bases upstream from the ATG codon, is a likely candidate for a ribosome binding site.



B

```
E. coli     MIDT T L PLTDIHRHLDGNIRPQTILELGRQYNISLPAQSLETLIPHVQVIANEPDLVSFLTKLDWGVKVLA SL
Mouse       MAQTPAFNKPKVELHVHLDGAIKPETILYFGKKRGIALPADTVEEL RNIIGMDKPLSLPGFLAKFDYYMPVIAGCR
Human       MAQTPAFDKPKVELHVHLDGSIKPETILYYGRRRGIALPANTAEGLL NVIGMDKPLTLPDFLAKFDYYMPAIAGCR
Conserved   M--T-+---P--++H-HLDG-I+P-TIL--G+---I+LPA-+-E-L---+--+-----L--FL+K-D+-+--+A---

            DACRRVAFENIEDAARHGLHYVELRFSPGYMAMAHQLP    VA G      VVEAVIDGVREGCRTFGVQAKLIGI
            EAIKRIAYEFVEMKAKEGVVYVEVRYSPHLLANSKVDPMPWNQTEGDVTPDDVVDLVNQGLQEGEQAFGIKVRSI L
            EAIKRIAYEFVEMKAKEGVVYVEVRYSPHLLANSKVEPIPWNQAEGDLTPDEVVALVGQGLQEGERDFGVKARSI L
            +A-+R+A+E-+E--A+-G+-YVE+R+SP--+A-++--P-----+-G------VV--V--G+-EG---FG+--+-I-+

            MSRTFEAACQQELEAFLAHRDQIT ALDLAGDELGFPGSLFLSH FN RARDAGWHITVHAGEAAGPESIWQAIR
            CCMRHQPSWSLEVLELCKKYNQKTVVAMDLAGDETIEGSSLFPGHVEAYEGAVKNGIHRTVHAGEVGSPEVVREAVD
            CCMRHQPNWSPKVVELCKKYQQQTVVAIDLAGDETIPGSSLLPGHVQAYQEAVKSGIHRTVHAGEVGSAEVVKEAVD
            -----------+-----+--Q-T--A+DLAGDE-----SL---H------A---G-H-TVHAGE----E-+--A+-

            ELGAERIGHGVKAIEDRALMDFLAEQQIGIESCLTSNIQTSTVAELAAHPLKTFLEHGIRASINTDDPGVQGVDII
            ILKTERVGHGYHTIEDEALYNRLLKENMHFEVCPWSSYLTGAWDPKTTHAVVRFKNDKANYSLNTDDPLIFKSTLD
            ILKTERLGHGYHTLEDQALYNRLRQENMHFEICPWSSYLTGAWKPDTEHAVIRLKNDQANYSLNTDDPLIFKSTLD
            -L-+ER+GHG-+++ED-AL---L---++--E-C--S---T--+---+-H-+----------S+NTDDP-+----+-

            HEYTVAAPAAGLSREQIRQAQINGLEMAFLSAEEKRALREKVAAK
            TDYQMTKKDMGFTEEEFKRLNINAAKSSFLPEEEKKELLERLYREYQ
            TDYQMTKRDMGFTEEEFKRLNINAAKSSFLPEDEKRELLDLLYKAYGMPPSASAGQNL
            -+Y-++----G-+-E--+--+IN----+FL--+EK+-L-+-+----------------
```

FIGURE 4: Amino acid sequence comparisons of *E. coli*, murine, and human adenosine deaminases. Panel A shows the dot matrix comparisons. To score positive in the comparison, four amino acids out of five must be conserved. Panel B shows the sequence alignment comparisons with identical residues listed on the bottom line where (+) indicates a conservative amino acid substitution and (−) indicates a nonconservative amino acid substitution.

Table I

Enzymes	Conserved Sequences[a]	References[b]
Adenosine Deaminases		
E.coli	^{194}TVHAGE-----(44)-----C^{244}-----(31)------SINTDDP^{279}	(1)
Mouse	^{212}TVHAGE-----(44)-----C^{262}-----(31)------SLNTDDP^{297}	(2)
Human	^{212}TVHAGE-----(44)-----C^{262}-----(31)------SLNTDDP^{297}	(3)
AMP Deaminases		
Yeast	^{421}MHAH-----(206)-----C^{631}-----(74)------SLSTDDP^{709}	(4)
Rat	^{362}VHAG-----(207)-----C^{573}-----(74)------SLSTDDP^{651}	(5)
Human(M)	^{362}VHAG-----(207)-----C^{573}-----(74)------SLSTDDP^{651}	(6)
Human(L)	^{362}VHAD-----(207)-----C^{573}-----(74)------SLSTDDP^{651}	(7)

[a] The number of amino acid residues between the conserved sequences is indicated in parentheses. The amino acids depicted in the model presented in Figure 5 are underlined. [b] References are as follows: (1) this manuscript; (2) Yeung et al., 1985; (3) Wiginton et al., 1984; (4) Meyer et al., 1989; (5) Sabina et al., 1987; (6) Sabina et al., 1990; (7) Sabina et al., personal communication.

The sequence ATGGTGCAAATGCACCACCATTTTT, which is eight bases downstream from the translation stop codon, encodes an RNA with the potential to form a stem–loop followed by a run of uridylic acid residues and very likely serves as a ρ-independent transcription terminator (Platt, 1986).

Conserved Amino Acid Sequences Present in Adenosine Deaminases and AMP Deaminases. Although there is a great deal of overall amino acid sequence conservation between the *E. coli* and mammalian adenosine deaminases, there are only four stretches of four or more contiguous amino acid residues that have been completely conserved with regard to sequence and highly conserved with regard to position (Figure 4). The distance between two of the conserved elements, TVHAGE and SL(I)NTDDP, has been exactly conserved at 76 amino acids for all three adenosine deaminases. In addition, the only conserved cysteine among the three enzymes occurs between these two conserved elements with identical spacing between each element and the cysteine as illustrated in Table I. The amino acid sequence of the three adenosine deaminases were also compared with those of four AMP deaminases; yeast, rat,

and the human muscle (M) and liver (L) forms. AMP deaminase is a larger protein than adenosine deaminase with monomeric molecular weight of approximately 96 000 (Meyer et al., 1989) and is a tetramer in its native form. Nevertheless, it has been suggested that adenosine deaminases and AMP deaminases have similar catalytic mechanisms due to the similarity of the reactions catalyzed and the fact that deoxycoformycin derivatives are potent inhibitors of both enzymes (Frieden et al., 1980). Two of the four conserved sequence elements shared among the three adenosine deaminases are also present in the four AMP deaminases (Table I). The SLSTDDP sequence differs only in the conservative substitution of a serine for an asparagine. In the mammalian AMP deaminases the VHA sequence is completely conserved, and the HA portion is conserved in the yeast enzyme (Table I). A conserved cysteine with conserved spacing between these two regions is also found. Thus, sequence analysis of three adenosine deaminases and four AMP deaminases revealed the presence of a highly conserved sequence motif, SLN(S)TDDP, a conserved dipeptide, HA, and a conserved cysteine residue. Such highly conserved amino acids are likely to provide functional groups that play important roles in the catalytic process of these enzymes (Reichardt & Berg, 1988).

Discussion

The molecular cloning and sequence analysis of the *E. coli* adenosine deaminase gene has enabled us to identify amino acid residues that are evolutionarily conserved among adenosine deaminases and AMP deaminases, enzymes which catalyze the hydrolytic deamination from the C-6 position of adenine nucleosides and nucleotides. These highly conserved amino acid residues (Table I) are likely to play functional roles in the catalytic process of these enzymes. Participation of certain of these conserved amino acids in the catalytic mechanism of the adenosine deaminase reaction has been suggested by a number of previous studies (Kati & Wolfenden, 1989a,b; Kurz & Frieden, 1983; Jones et al., 1989; Orsi et al., 1972; Weiss et al., 1987; Wolfenden et al., 1967; Ronca et al., 1967). These studies indicated that two or more ionizable groups appear to participate in the enzyme reaction. Orsi et al. (1972) indicated that the rate-limiting step of the adenosine deaminase catalyzed reaction appears to be governed by a single ionizing group with a pK_a of 4.8, compatible with a carboxyl group. Similarly, Kurz and Frieden (1983) found that an ionizable group with a pK_a of 4.97 is rate limiting and suggested that it may be a carboxyl group. The involvement of a sulfhydryl group for the adenosine deaminase catalyzed reaction was implied by chemical modification studies (Wolfenden et al., 1967; Ronca et al., 1967) and pH profile studies (Orsi et al., 1972; Wolfenden et al., 1967; Weiss et al., 1987). The involvement of a sulfhydryl group was also confirmed by the analysis of nitrogen-15 and solvent deuterium isotope effects (Weiss et al., 1987). Kati and Wolfendon (1989a,b) found that a single hydroxyl group at the C-6 position of 6-hydroxyl-1,6-dihydropurine ribonucleoside, a transition-state analogue inhibitor of adenosine deaminase, increased the affinity for the enzyme by approximately 10^8-fold. They suggested that strong hydrogen bonding may occur between the 6-hydroxyl and a charged carbonyl group at the active site. In addition, an imidzole group, probably in the protonated form, has been implicated in enzyme function (Orsi et al., 1972; Kurz & Frieden, 1983). Thus, available data suggest that the adenosine deaminase catalyzed hydrolytic deamination process includes the following events: (1) the reversible donation of a proton to the N-1 position of the purine ring by an enzyme sulfhydryl, (2) the removal of a proton from the water molecule that attacks the C-6 position of the purine ring, (3) the stabilization of the resulting C-6 tetrahedral intermediate via strong hydrogen bonding to the 6-hydroxyl group, and (4) the donation of a proton to the 6-amino group to form the leaving group.

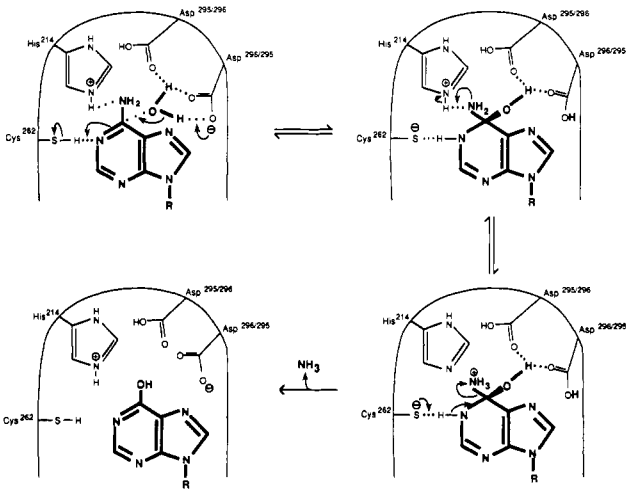


FIGURE 5: Enzymatic deamination of adenosine: a model. In view of previously published enzymological data and the conserved amino acid residues identified in this study, we propose a model to account for the enzyme-catalyzed deamination of adenosine. Potential catalytic roles are assigned to the conserved Cys 262, His 214, Asp 295, and Asp 296 residues of mammalian adenosine deaminases and the corresponding amino acid residues in bacterial adenosine deaminase and the eukaryotic AMP deaminases (see Table I). The two aspartates (Asp 295 and Asp 296) may function in a manner similar to that proposed for aspartic acid proteinases [see Discussion and Davies (1990), Foundling et al. (1987), and Rich et al. (1985)].

On the basis of highly conserved amino acid sequences we have identified in the present study and on the mechanistic data described in the preceding paragraph, we propose a model for the enzyme-catalyzed hydrolytic deamination of adenosine (Figure 5). We propose that the four catalytic functions referred to above are carried out by Cys 262, Asp 295, Asp 296, and His 214 of the mammalian adenosine deaminases and the corresponding conserved amino acids in the bacterial enzyme. According to this model (Figure 5) Cys 262 is the enzyme sulfhydryl believed to share a proton with the N-1 position of the purine ring. One of the aspartic acid residues (Asp 295 or Asp 296) serves as a general base which is believed to remove a proton from the attacking water molecule, and the two aspartates together provide the strong hydrogen bonding to the 6-hydroxyl which is postulated to stabilize the tetrahedral intermediate (Kati & Wolfenden, 1989a, 1989b). The two aspartate residues (Asp 295 and Asp 296) are likely to function in a manner similar to that postulated for the aspartic acid proteinases. For the latter enzymes the reaction is believed to proceed by general-base catalysis, in which the carboxyl groups provided by the two highly conserved aspartates deprotonate the attacking water molecule and stabilize the tetrahedral intermediate (Davies, 1990; Foundling et al., 1987; Rich et al., 1985). His 214 of the mammalian adenosine deaminases is postulated to serve as a general acid by donating a proton to the amino group which is eliminated as ammonia from the C-6 position. Other highly conserved amino acid residues deserve mention. The serine, asparagine, and threonine, which are present along with the two aspartates in the conserved seven amino acid sequence, are likely candidates to participate in hydrogen bonding to substrates or intermediates in the enzyme-catalyzed reaction. The asparagine is conservatively replaced with serine in the AMP deaminases. Because these specific amino acids are so highly conserved in

all adenosine deaminases and AMP deaminases for which sequence information is available, we postulate that these amino acids carry out similar functions in both classes of enzyme.

The model depicted in Figure 5 makes testable predictions. The model predicts that the side chains of the four highly conserved amino acids are present at the active site of the enzyme. A high-resolution map of murine adenosine deaminase should be available in the near future and will provide a direct structural test of this prediction. The model also predicts that amino acid substitutions at the four amino acids depicted in Figure 5, as well as changes within or near the highly conserved sequence motifs shown in Table I, will affect catalytic activity. The model predicts that the highly conserved amino acids illustrated in Table I will likely be found in other enzymes that catalyze a hydrolytic deamination from an adenine ring. Likely candidates include the enzyme that converts adenosine to inosine in the anticodon of many tRNAs (Elliott & Trewyn, 1984; Grosjean et al., 1987) and the RNA unwinding activity that deaminates up to 40% of adenosine residues in double-stranded RNA (Bass & Weintraub, 1988; Wagner et al., 1989).

As a functional test of the model proposed in Figure 5, site-specific amino acid substitutions are being genetically engineered into the recombinant murine adenosine deaminase. Because each mutant must be subjected to a thorough structural and functional analysis in order to accurately understand the basis for any changes in enzyme activity, this work will require considerable time and effort to complete. However, as a preliminary test of the model, we surveyed the naturally occurring amino acid substitutions associated with complete loss or drastic reduction of adenosine deaminase activity in humans suffering from immunodeficiency syndromes (Akeson et al., 1989; Hirschhorn et al., 1989). One of the amino acid substitutions falls within the conserved SLNTDDP sequence, a proline to glutamine mutation at position 297 (Hirschhorn et al., 1989). Another occurs immediately in front of the conserved TVHAGE sequence, a mutation causing an arginine to histidine change at amino acid position 211 (Akeson et al., 1988). Thus, naturally occurring missense mutations within or immediately adjacent to the highly conserved amino acid sequence motifs are associated with a loss of enzyme function. Genetically engineered adenosine deaminses and AMP deaminases with amino acid substitutions at the conserved cysteine, histidine, or aspartic acid residues are being constructed. The detailed structural and functional analysis of such site-directed mutants will allow a direct test of the model presented in Figure 5.

Acknowledgments

We appreciate the critical comments and helpful discussion provided by Hiram F. Gilbert, Florante A. Quiocho, V. Ramamurthy, Frederick B. Rudolph, and David K. Wilson. We are especially grateful to Richard L. Sabina for sharing with us unpublished data regarding the sequence of the human L isoform of AMP deaminase. Bacterial strain Sϕ1608 was kindly provided by Bente Mygind.

References

Agarwal, R. P., Sagar, S. M., & Parks, R. E., Jr. (1975) *Biochem. Pharmacol. 24*, 693–701.

Agarwal, R. P., Spector, T., & Parks, R. E., Jr. (1977) *Biochem. Pharmacol. 26*, 359–367.

Akeson, A. L., Wiginton, D. A., Dusing, M. R., States, J. C., & Hutton, J. J. (1988) *J. Biol. Chem. 263*, 16291–16296.

Akeson, A. L., Wiginton, D. A., & Hutton, J. J. (1989) *J. Cell. Biochem. 39*, 217–228.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Smith, J. A., Seidman, J. G., & Struhl, K. (1987) *Current Protocols in Molecular Biology*, Greene Publishing Associates and Wiley Interscience, New York.

Bass, B. L., & Weintraub, H. (1988) *Cell 55*, 1089–1098.

Cha, S., Agarwal, R. P., & Parks, R. E., Jr. (1975) *Biochem. Pharmacol. 24*, 2187–2197.

Chassy, B. M., & Suhadolnik, R. J. (1967) *J. Biol. Chem. 242*, 3655–3658.

Davies, D. R. (1990) *Annu. Rev. Biophys. Biophys. Chem. 19*, 189–215.

Elliott, M. S., & Trewyn, R. W. (1984) *J. Biol. Chem. 259*, 2407–2410.

Evans, B. E., & Wolfenden, R. (1970) *J. Am. Chem. Soc. 92*, 4751–4752.

Evans, B. E., & Wolfenden, R. V. (1973) *Biochemistry 12*, 392–398.

Foundling, S. I., Cooper, J., Watson, F. E., Cleasby, A., Pearl, L. H., Sibanda, B. L., Hemmings, A., Wood, S. P., Blundell, T. L., Valler, M. J., Norey, C. G., Kay, J., Boger, J., Dunn, B. M., Leckie, B. J., Jones, D. M., Atrash, B., Hallett, A., & Szelke, M. (1987) *Nature 327*, 349–352.

Frick, L. F., Smal, E., Baker, D. C., & Wolfenden, R. (1986) *Biochemistry 25*, 1616–1621.

Frick, L. F., Neela, J. P. M., & Wolfenden, R. (1987) *Bioorg. Chem. 15*, 100–108.

Frieden, C., Kurz, L. C., & Gilbert, H. R. (1980) *Biochemistry 19*, 5303–5309.

Gold, L., Pribnow, D., Schmeider, T., Shinedling, S., Singer, B. S., & Stormo, G. (1981) *Annu. Rev. Microbiol. 35*, 365–403.

Grosjean, H., Haumont, E., Droogmans, L., Carbon, P., Fournier, M., de Henau, S., Doi, T., Keith, G., Gangloff, J., Kretz, K., & Trewyn, R. (1987) in *Biophosphates and Their Analogues-Synthesis, Structure, Metabolism and Activity* (Bruzik, K. S., & Stec, W. J., Eds.) pp 355–378, Elsevier Science Publishers B.V., Amsterdam.

Hawley, D. K., & McClure, W. R. (1983) *Nucleic Acids Res. 11*, 2237–2255.

Hirschhorn, R., Tzall, S., Ellenbogen, A., & Orkin, S. H. (1989) *J. Clin. Invest. 83*, 497–501.

Hove-Jensen, B., & Nygaard, P. (1989) *J. Gen. Microbiol. 135*, 1263–1273.

Jayakumar, A., Hwang, S. J., Fabiny, J. M., Chinault, A. C., & Barnes, E. M., Jr. (1989) *J. Bacteriol. 171*, 996–1001.

Jochimsen, B., Nygaard, P., & Vestergaard, T. (1975) *Mol. Gen. Genet. 143*, 85–91.

Jones, W., Kurz, L. C., & Wolfenden, R. (1989) *Biochemistry 28*, 1242–1247.

Kati, W. M., & Wolfenden, R. (1989a) *Science 243*, 1591–1593.

Kati, W. M., & Wolfenden, R. (1989b) *Biochemistry 28*, 7919–7927.

Kurz, L. C., & Frieden, C. (1983) *Biochemistry 22*, 382–389.

Kurz, L. C., & Frieden, C. (1987) *Biochemistry 26*, 8450–8457.

Leatherbarrow, R. J., & Fersht, A. R. (1987) in *Enzyme Mechanisms* (Page, M. I., & Williams, A., Eds.) pp 78–96, The Royal Society of Chemistry, London.

McClure, W. (1985) *Annu. Rev. Biochem. 54*, 171–204.

Meyer, S. L., Kvalnes-Krick, K. L., & Schramm, V. L. (1989) *Biochemistry 28*, 8734–8743.

Nygaard, P. (1978) *Methods Enzymol. 51*, 508–512.
Nygaard, P. (1983) in *Metabolism of Nucleotides, Nucleosides and Nucleobases in Microorganism* (Munch-Petersen, A., Ed.) pp 27–93, Academic Press, London.
Orsi, B. A., McFerran, N., Hill, A., & Bingham, A. (1972) *Biochemistry 11*, 3386–3392.
Platt, T. (1986) *Annu. Rev. Biochem. 55*, 339–372.
Rich, D. H., Bernatowicz, M. S., Agarwal, N. S., Kawai, M., Salituro, F. G., & Schmidt, P. G. (1985) *Biochemistry 24*, 3165–3173.
Reichardt, J. K., & Berg, P. (1988) *Nucleic Acids Res. 16*, 9017–9026.
Ronca, G., Bauer, C., & Rossi, C. A. (1967) *Eur. J. Biochem. 1*, 434–438.
Sabina, R. L., Marquetant, R., Desai, N. M., Kaletha, K., & Holmes, E. W. (1987) *J. Biol. Chem. 262*, 12397–12400.
Sabina, R. L., Movisak, T., Clark, P., Eddy, R., Shows, T. B., Morton, C. C., & Holmes, E. W. (1990) *J. Biol. Chem. 265*, 9423–9433.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning, a Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Schaaper, R. M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8126–8130.
Schramm, V. L., & Baker, D. C. (1985) *Biochemistry 24*, 641–646.
Spector, T., Jones, T. E., & Beacham, L. M., III (1983) *Biochem. Pharmacol. 32*, 2505–2509.
Suguna, K., Padlan, E. A., Smith, C. W., Carlson, W. D., & Davies, D. R. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 7009–7013.
Wagner, R. W., Smith, J. E., Cooperman, B. S., & Nishikura, K. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2647–2651.
Weiss, P. M., Cook, P. F., Hermes, J. D., & Cleland, W. W. (1987) *Biochemistry 26*, 7378–7384.
Wiginton, D. A., Adrian, G. S., & Hutton, J. J. (1984) *Nucleic Acid Res. 12*, 2439–2446.
Wilson, D. K., Rudolph, F. B., Harrison, M. L., Kellems, R. E., & Quiocho, F. A. (1988) *J. Mol. Biol. 200*, 613–614.
Wolfenden, R. (1969) *Biochemistry 8*, 2409–2412.
Wolfenden, R., Sharpless, T. K., & Allan, R. (1967) *J. Biol. Chem. 242*, 977–983.
Wolfenden, R., Kaufman, J., & Macon, J. B. (1969) *Biochemistry 8*, 2412–2415.
Wolfenden, R., Wentworth, D. F., & Mitchell, G. N. (1977) *Biochemistry 16*, 5071–5077.
Yeung, C. Y., Riser, M. E., Kellems, R. E., & Siciliano, M. J. (1983) *J. Biol. Chem. 258*, 8330–8337.
Yeung, C. Y., Ingolia, D. E., Roth, D. B., Shoemaker, C., Al-Ubaidi, M. R., Yen, J. Y., Ching, C., Bobonis, C., Kaufman, R. J., & Kellems, R. E. (1985) *J. Biol. Chem. 260*, 10299–10307.