Cloning, Sequencing, Expression, and Site-Directed Mutagenesis of the Gene from *Clostridium perfringens* Encoding Pyruvoyl-Dependent Histidine Decarboxylase[†,‡]

Paul D. van Poelje and Esmond E. Snell*

Departments of Chemistry and Microbiology, The University of Texas at Austin, Austin, Texas 78712

Received July 10, 1989; Revised Manuscript Received August 22, 1989

ABSTRACT: The DNA encoding pyruvoyl-dependent histidine decarboxylase (HisDCase) of *Clostridium perfringens* was cloned, sequenced, and overexpressed in *Escherichia coli*. The gene encodes a single polypeptide of 320 amino acids, M_r 35 526, demonstrating that clostridial HisDCase, which has an $(\alpha\beta)_6$ structure, is synthesized as a precursor (proHisDCase, π_6). No π subunits of proHisDCase were observed in crude or purified preparations of the cloned HisDCase; they appear to undergo rapid cleavage in vivo to the α (M_r 24 887) and β (M_r 10 526) subunits characteristic of this HisDCase. This cleavage occurs between Ser-96 and Ser-97; Ser-97 gives rise to the catalytically essential pyruvoyl group blocking the N-termini of the α subunits of the active enzyme. When Ser-97 was converted to an alanyl residue by site-specific mutagenesis, the expressed, inactive protein (π'_6) contained a single peptide species (π', M_r 35 510) that was not cleaved either in vivo or in vitro. These results support previous conclusions that activation of the wild-type clostridial proenzyme occurs via nonhydrolytic serinolysis. Although clostridial HisDCase has only a 47% sequence similarity to HisDCase from *Lactobacillus* 30a, all of the residues known to be important for substrate binding and catalytic action of the *Lactobacillus* HisDCase are conserved in the *C. perfringens* enzyme. While the encoded N-terminal Met of clostridial HisDCase is removed by *E. coli*, the cloned enzyme retains a 10-residue presequence (NKNLEANRNR) not present in the mature enzyme isolated from *C. perfringens*.

Pyruvoyl-dependent histidine decarboxylases have been purified to homogeneity from four (Gram-positive) bacterial sources: *Lactobacillus* 30a (Rosenthaler et al., 1965), *Lactobacillus buchneri* (Huynh & Snell, 1985a), *Clostridium perfringens* (Recsei et al., 1983b), and *Micrococcus* sp.n. (Semina & Mardashev, 1965). All four enzymes contain pairs of dissimilar subunits and show considerable amino acid sequence homology, particularly at the C-termini of their smaller "β" subunits and the N-termini of their "α" subunits (Huynh & Snell, 1985a). Both the amino acid sequence (Vaaler et al., 1982; Huynh et al., 1984) and the gene sequence (Vanderslice et al., 1986) of the best characterized enzyme of this group, the histidine decarboxylase (HisDCase)[1] [$(\alpha\beta)_6$] from *Lactobacillus* 30a, are known. It is synthesized as an inactive proenzyme (π_6) which undergoes a self-catalyzed serinolysis reaction in which each constituent π subunit cleaves nonhydrolytically between serine residues 81 and 82, with loss of ammonia, to yield a β subunit with Ser-81 at its C-terminus and an α subunit with a catalytically essential pyruvoyl residue (derived from Ser-82) at its N-terminus (Recsei et al., 1983a). A similar proenzyme and activation mechanism have been demonstrated for the HisDCase from *L. buchneri* (Recsei & Snell, 1985) but not for the HisDCase from other organisms.

The X-ray structure of the *Lactobacillus* 30a HisDCase has recently been refined to a resolution of 2.5 Å (Gallagher et al., 1989) and permitted formulation of a tentative molecular mechanism for the decarboxylation reaction. Comparative studies of related, substantially different enzymes would greatly enhance the significance of such mechanistic studies and also clarify features essential for the unique self-activation reaction of the precursors of these enzymes. To this end, we describe here the cloning, sequencing, and expression of the *hdc* gene from *C. perfringens*, and the purification and properties of the recombinant HisDCase. By conversion of the serine residue that gives rise to the pyruvoyl residue at the active site of HisDCase to an alanyl residue by site-specific mutagenesis, we demonstrate in vivo accumulation of a mutant proenzyme (π'_6).

Table I: Bacterial Strains

strain	genotype	reference
C. perfringens (ATCC 13124)	wild type	
E. coli K12		
JM105	Δ(*lacpro*) *thi strA endA sbcBa5 hsdR4*/F'*traD36 proAB lacI*^{q}ZM15	Felton (1983)
GM1	*ara*Δ(*lacpro*)*thi* F'/(*lacpro*) *lacI*^{q}L8	Miller et al. (1977)
BW313	HfrKL16PO/45[*lysA*(61-62) *dut1 ung1 thi1 relA1*	Kunkel et al. (1987)
JM103Y	Δ(*lacpro*) *thi strA supE endA sbcB15 hsdR4*/F'*traD36 proAB lacI*^{q}ZΔM15	Messing et al. (1981)

Materials and Methods

Materials. Materials and their suppliers were as follows: restriction endonucleases, T4 DNA ligase, T4 polynucleotide kinase, DNA molecular weight standards, and phage M13mp18 and M13mp19, New England Biolabs; calf intestinal alkaline phosphatase, Boehringer Mannheim; T4 DNA polymerase, Bio-Rad; Sequenase DNA sequencing kit, U.S.

[†] This work was supported by Grant F-714 from the Robert A. Welch Foundation.

[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J02880.

* Author to whom correspondence should be addressed at the Department of Microbiology, The University of Texas at Austin.

[1] Abbreviations: HisDCase, histidine decarboxylase (EC 4.1.1.22); proHisDCase, prohistidine decarboxylase; DEAE, diethylaminoethyl; Ap^r, ampicillin resistance; IPTG, isopropyl β-D-galactopyranoside; X-gal, 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside; kb, kilobase pair(s); M_r, molecular weight; SDS, sodium dodecyl sulfate.

0006-2960/90/0429-132$02.50/0

Biochemicals; nitrocellulose filters (BA 85) and NA-45 DEAE membranes, Schleicher & Schuell; HA-type nitrocellulose filters, Millipore; [γ-^{32}P]ATP (4500 Ci/mmol), [α-^{35}S]dATP (1350 Ci/mmol), and L-[*carboxyl*-^{14}C]histidine (56 mCi/mmol), New England Nuclear; Sephacryl S-200 and Sephadex G-50, Pharmacia; DE52, Whatman. The mixed-sequence oligonucleotide probe corresponding to amino acid residues 40–45 of HisDCase from *C. perfringens* and the oligonucleotide primers used for sequencing and site-directed mutagenesis (Figure 1) were synthesized locally on an Applied Biosystems Model 381 or 381A DNA synthesizer. All other chemicals were of reagent grade from standard sources.

Bacterial Strains and Growth Media. The bacterial strains used are shown in Table I. *C. perfringens* was grown in 3% Difco tryptic soy broth, 1% glucose, 0.5% yeast extract, and 0.05% sodium thioglycolate (stationary culture, 37 °C). *E. coli* strains JM105, GM1, and BW313 were grown in TYE medium (1% tryptone, 0.5% yeast extract, and 0.5% sodium chloride, pH 7.0, supplemented with 200 μg/mL carbenicillin when required), and *E. coli* JM103Y was grown overnight in minimal medium (Miller, 1972) supplemented with 1% glucose before dilution into TYE medium.

Cloning of the hdc Gene. Standard DNA techniques were used (Maniatis et al., 1982) unless otherwise indicated. Genomic DNA was isolated from *C. perfringens* by the procedure of Nakamura et al. (1979). Restriction enzyme digests of this DNA preparation were subjected to electrophoresis on 1% agarose gels containing 0.5 μg/mL ethidium bromide, then transferred onto nitrocellulose sheets, and hybridized to the ^{32}P-labeled oligonucleotide probe (10^6 cpm/mL) at 46 °C for 16 h. Following five washes in 0.9 M NaCl/0.9 M sodium citrate, pH 7.0 (100 mL each, room temperature), the blots were exposed to Kodak XAR-5 film with an intensifying screen (2–12 h, −70 °C). Appropriate restriction fragments were purified from agarose gels using either NA-45 DEAE membranes (according to the manufacturer's instructions) or Spectrapor dialysis tubing. Purified fragments were ligated into appropriately restricted and phosphatase-treated pUC8 [Ap^r (Vieira & Messing, 1982)] plasmid DNA and transformed into competent JM105 cells which were then plated on TYE plates [containing carbenicillin (200 μg/mL), IPTG (0.5 mM), and X-gal (40 μg/mL)] and grown at 37 °C overnight. Recombinant colonies were transformed to duplicate TYE/carbenicillin plates, grown for 12 h at 37 °C, and then transferred to nitrocellulose filters. The filter-bound colonies were lysed and screened with the labeled oligonucleotide probe as described by Woods (1984). Plasmid DNA was prepared from colonies that bound the probe by the boiling method (Holmes & Quigley, 1981). Restriction analyses of the recombinant plasmids were performed as described for genomic DNA above.

DNA Sequencing. pUC8 plasmid inserts were cloned into phage M13mp18 and/or M13mp19 (Yannisch-Perron et al., 1981). RF DNA was purified from phage-infected cells by the alkaline lysis method (Birnboim & Doly, 1979), and single-stranded template DNA was purified from their culture medium (Carlson & Messing, 1984). Templates were sequenced by using a Sequenase DNA sequencing kit according to the instructions of the manufacturer. Acrylamide gels [4–8%; acrylamide:bis(acrylamide) ratio of 19:1] were run on a Hoefer poker face sequencer. Nucleotide and amino acid sequences were compiled and analyzed by using DNAstar (DNAstar Inc., Madison, WI) software on an IBM XT personal computer. The algorithms of Needleman and Wunsch (1970) and Garnier et al. (1978) were used for the alignment of amino acid sequences and structural predictions, respectively.

Expression of the hdc Gene. The *hdc* gene was expressed in *E. coli* under control of the *lac* promoter of pUC8 as described under Results. All DNA manipulations were performed by using standard procedures. Transformants were selected on TYE plates (200 μg/mL carbenicillin), and their plasmid DNA was prepared for restriction analysis by using the boiling method (Holmes & Quigley, 1981). For the purpose of enzyme purification, recombinant cells were grown in TYE medium supplemented with 200 μg/mL carbenicillin and 1 mM IPTG.

Site-Directed Mutagenesis. The site-specific mutant was constructed by the method of Kunkel et al. (1987). Uracil templates were produced in *E. coli* BW313 (*dut*$^-$, *ung*$^-$). Selection of the mutant DNA strand was performed in *E. coli* JM103Y (*dut*$^+$, *ung*$^+$).

Purification of Cloned HisDCase. All purification steps, excluding the heat treatment and column chromatography, were performed at 4 °C. Unless otherwise indicated, centrifugations were at 43000*g*. Cells from a 2-L culture were harvested in early stationary phase by centrifugation (4000*g*, 5 min), resuspended in 0.2 M potassium phosphate, pH 7.0, and forced through a French pressure cell (1000 psi, 3×). The extract was clarified by centrifugation (30 min). Ammonium sulfate was then added to 50% saturation, and after being stirred for 15 min, the solution was centrifuged (15 min). Additional ammonium sulfate was subsequently added to the supernatant to 95% saturation. After the mixture was stirred for 15 min, the precipitated protein was collected by centrifugation (15 min), redissolved in H_2O, and heated to 70 °C for 2 min with continuous swirling in a conical flask immersed in a water bath maintained at 75 °C. Precipitated protein was removed by centrifugation (15 min) and the supernatant dialyzed against 0.02 M potassium phosphate, pH 7.0 (buffer A), for 4 h. The dialyzed protein was then passed through a DE-52 column (10 × 2.5 cm) equilibrated and eluted with buffer A. HisDCase-containing fractions were pooled, concentrated to 5 mL, and loaded onto a Sephacryl S-200 column (1.5 × 150 cm) which was subsequently eluted with 0.2 M ammonium acetate, pH 4.5. The appropriate fractions were pooled, concentrated to a protein concentration of 10 mg/mL using an Amicon ultrafiltration cell, and stored at 4 °C.

Purification of the β Subunit of Cloned HisDCase. The β subunit was separated from the α subunit, following denaturation of cloned HDC [6 M guanidine hydrochloride/0.05 M potassium phosphate, pH 7.0 (buffer B), 1 h, 37 °C], by gel filtration on a Sephadex G-50 column (1.5 × 160 cm) equilibrated and eluted with buffer B. The β subunit, which eluted last from the column, was then dialyzed exhaustively against deionized water and lyophilized.

SDS Gel Electrophoresis. Electrophoresis was performed in the presence of 1% SDS on slab gels (Laemmli, 1970). Protein bands were visualized with Coomassie blue (Fairbanks et al., 1979).

Protein Sequencing. Protein sequencing was performed either manually by the 4-(*N*,*N*-dimethylamino)benzene 4-isothiocyanate/phenyl isothiocyanate double-coupling method (Chang et al., 1978) or automatically on an Applied Biosystems 477A sequencer.

Results

Cloning and Sequencing of the hdc Gene. Restriction fragments (1.9–2.5 kb) resulting from a *Hin*dIII digest of genomic DNA from *C. perfringens* were purified, cloned into pUC8, and introduced into *E. coli* JM105 cells. One of the

```
                40                  45
A.           Asp Gly Tyr Gly Met Pro
      DNA 3'-CTA CCG ATG CCT TAC GG -5'

     Probe 5'-GA(T/C) GGN TA(T/C) GGN ATG CC -3'


           241                 257
B.  (i) 3'-CG ACT TTA TTT ACT ACG-5'

           496                 512
   (ii) 3'-CG GGT ACA TTA TAC ACG-5'

           740                 756
  (iii) 3'-CG CCA TCT CTA AGA GTA-5'

           988                 1004
   (iv) 3'-CC TCC TCA ATA ATA CTT-5'


          93                          100
          Leu Thr Ala Ser Ala Phe Cys Gly
C.    3'-C GAA TGT CGA AGT CGA AAG ACA CCC-5'
```

FIGURE 1: (A) Nucleotide and corresponding amino acid sequences of the probe used to locate the *hdc* gene of *C. perfringens*. (B) Sequencing primers designed to anneal to the template strand of the *hdc* gene at approximately 250 base pair intervals. (C) Nucleotide and corresponding amino acid sequence of the primer used to convert the codon for Ser-97 to an alanyl codon by site-specific mutagenesis. The bases that differ from the wild-type sequence are underlined.

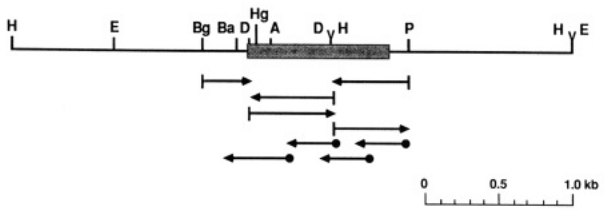

FIGURE 2: Partial restriction map and sequencing strategy of the *hdc* gene and surrounding regions. The coding region of the *hdc* gene is shadowed. Only the restriction sites used in the subcloning of fragments are shown. The arrows show the direction and extent of nucleotide sequencing from the universal primer (|→) and from specific primers (●→). Abbreviations: A, *Afl*II; Ba, *Bal*I; Bg, *Bgl*II; D, *Dra*I; E, *Eco*RI; H, *Hin*dIII; Hg, *Hgi*AI; P, *Pst*I.

500 transformants screened contained a plasmid (pVP-1) with an insert (2.2 kb) that hybridized to an oligonucleotide probe (Figure 1A) corresponding to a pentameric amino acid sequence of *C. perfringens* HisDCase. A restriction map of the insert was constructed (Figure 2) and the coding region of the *hdc* gene located by nucleotide sequencing of appropriate subclones and comparison of their encoded sequences to known amino acid sequences of the HisDCases from *C. perfringens* (Huynh & Snell, 1985a) and *Lactobacillus* 30a (Vaaler et al., 1982; Huynh et al., 1984). As only 57% of the *hdc* gene was found on the *Hin*dIII fragment, we attempted to clone an *Eco*RI fragment (3.1 kb) which restriction and hybridization data showed should contain the complete gene. For unknown reasons, none of 2500 transformants screened contained the desired insert; Vanderslice et al. (1986) had similar difficulties in attempts to clone the complete *hdc* gene and flanking regions from *Lactobacillus* 30a. An *Hgi*AI–*Pst*I fragment (1.1 kb) of the clostridial genome which lacked the *hdc* promoter region and the first 74 bases of the coding region but contained the remainder of the gene was subsequently cloned successfully into pUC8 between the *Sma*I and *Pst*I restriction sites. The insert of this plasmid (pVP-2) was obtained by digesting the genomic DNA from *C. perfringens* with *Hgi*AI, converting the cohesive ends of the resulting fragments to blunt ends with Sequenase, digesting these fragments with *Pst*I, and then purifying fragments of the appropriate size.

The complete sequence of the coding region of the *hdc* gene was determined unambiguously on both strands by sequencing of *Dra*I (0.6 kb) and *Hin*dII–*Pst*I (0.5 kb) fragments cloned into M13mp18 and M13mp19 (see Figure 2). Four oligonucleotide primers (Figure 1B) designed to anneal to one strand of the gene at approximately 250 base pair intervals were used to sequence the insert in pVP-2 and thus confirm the overlap between the aforementioned fragments. The sequence of the *hdc* promotor was determined on both strands as shown in Figure 2.

Nucleotide Sequence of the hdc Gene and the Deduced Amino Acid Sequence of HisDCase. The coding region of the *hdc* gene (Figure 3) comprises 963 nucleotides, initiates with ATG (Met), and terminates in TAA. The G + C content of the gene is 45.5%; that of the overall genome is 24–27% (Smith & Hobbs, 1974). The gene codes for a single polypeptide consisting of 320 amino acids (π subunit, M_r 35 528), and, since HisDCase from *C. perfringens* has an $(\alpha\beta)_6$ structure (Recsei et al., 1983b), the enzyme must be synthesized as a precursor molecule (proHisDCase, π_6). Comparison of this gene-encoded amino acid sequence with the previously determined sequence of the β subunit and the partial sequence at the N-terminus of the α subunit of HisDCase (Huynh & Snell, 1985a) revealed that cleavage of the π subunits of proHisDCase occurs between Ser-96 and Ser-97, that Ser-97 gives rise to the catalytically essential pyruvoyl group blocking the N-terminus of the α subunit, and that Ser-96 forms the carboxyl terminus of the β subunit.

Examination of the nucleotide sequence also revealed a decapeptide "presequence" (NKNLEANRNR) which was not found when the N-terminus of the intact β subunit of the mature enzyme isolated from *C. perfringens* was sequenced (Huynh & Snell, 1985a). It is not known whether removal of this peptide, which presumably occurs by trypsin-like cleavage at the terminal arginine residue in its sequence, takes place intracellularly in *C. perfringens* or during purification of the enzyme following extraction of the cells. If three-dimensional structures of HisDCase from *Lactobacillus* 30a (Parks et al., 1985) and from *C. perfringens* are similar, this presequence, the predicted structure of which is an α-helix, is probably located on the outer surface of the protein. Although the N-terminal peptide is shorter and lacks the hydrophobic region typical of signal peptides (Kreil, 1981), it may be involved in the localization of HisDCase in *C. perfringens*.

The molecular weights of the α and β subunits (24 887 and 10 526, respectively) calculated from the deduced amino acid sequence compare well with those determined earlier for the HisDCase from *C. perfringens* [25 000 and 9000 (Recsei et al., 1983b)]. There are several discrepancies between the sequence of HisDCase reported earlier (Huynh & Snell, 1985a) and the sequence deduced here for the gene product: His-108, His-113, Lys-122, Lys-128, and Thr-161 were previously reported as Val, Ser, Phe, Met, and Leu, respectively. Most of these sequencing errors were made during a single, extended automatic sequence analysis of the α subunit of HisDCase. In addition, the deduced sequence shows the presence of six Cys residues; only three Cys-peptides were isolated from partial digests of this enzyme (Huynh & Snell, 1985b). The number of Cys residues in the cloned enzyme determined by amino acid analysis [following its reduction and carboxymethylation by the procedure of Crestfield et al. 1963)] was in agreement with that indicated by the deduced sequence: 5.4 mol of (carboxymethyl)cysteine and 0.3 mol of cysteic acid were found per mole of $(\alpha\beta)$.

```
                      -35                     -10
5'-AAATACTATAAATGTTAAATTTGTAACAAATATTTAAAGTAGCTATAATTTTACATAATGAAAGGTAAATTATTATTTGGCCAAATAATATATTTTATTCTAAA
                                            1                                   10
                                        Met Asn Lys Asn Leu Glu Ala Asn Arg Asn Arg Thr Leu Ser Glu Gly Ile
ATTTAATAGTTAAATTTATATTGTTTAGGAGTGATTTAA ATG AAT AAA AAT CTT GAA GCA AAC AGA AAC AGG ACT CTT TCT GAA GGA ATA
            20                                      30                                      40
His Lys Asn Ile Lys Val Arg Ala Pro Lys Ile Asp Lys Thr Ala Ile Ser Pro Tyr Asp Arg Tyr Cys Asp Gly Tyr Gly
CAT AAA AAT ATA AAG GTA AGA GCA CCT AAA ATT GAT AAA ACA GCT ATA AGT CCT TAT GAT CGT TAT TGT GAT GGC TAC GGA
                        50                                      60                                      70
Met Pro Gly Ala Tyr Gly Asn Gly Tyr Val Ser Val Leu Lys Val Ser Val Gly Thr Val Lys Lys Thr Asp Asp Ile Leu
ATG CCA GGG GCT TAT GGT AAT GGA TAT GTA TCA GTT CTT AAG GTT TCC GTT GGT ACA GTT AAA AAG ACT GAC GAT ATA TTA
                                    80                                      90
Leu Asp Gly Ile Val Ser Tyr Asp Arg Ala Glu Ile Asn Asp Ala Tyr Val Gly Gln Ile Asn Met Leu Thr Ala Ser*Ser
CTA GAT GGT ATA GTT TCT TAT GAT AGA GCT GAA ATA AAT GAT GCT TAT GTT GGA CAA ATT AAT ATG CTT ACA GCT TCA TCA
        100                                     110                                     120
Phe Cys Gly Val Ala Gly Gln Val Trp Gly His Asp Leu Ala Thr His Asp Ser Ile Ala Asn Asp Glu Ile Lys Pro Leu
TTC TGT GGG GTT GCA GGT CAA GTT TGG GGA CAT GAT TTA GCA ACA CAT GAT TCA ATT GCT AAT GAT GAA ATC AAA CCT TTA
                    130                                     140                                     150
Tyr Glu Leu Lys Gln Phe Asp Gly Thr Pro Leu Lys Val Tyr Asp Ala Lys Pro Leu Leu Glu Ala Gly Ile Glu Leu Phe
TAT GAA TTA AAA CAA TTT GAT GGA ACT CCA CTT AAA GTT TAT GAT GCT AAA CCT TTA TTA GAA GCT GGT ATA GAA CTT TTC
                                160                                     170
Gly Thr Glu Lys Asn Arg Arg Phe Thr Thr Ala Pro Gly Ala His Val Ile Cys Ala Asn Lys Ser Ala Thr Ala Tyr Arg
GGT ACA GAA AAA AAT AGA CGT TTT ACA ACT GCT CCA GGT GCC CAT GTA ATA TGT GCA AAT AAG AGT GCT ACA GCT TAT AGA
    180                                     190                                     200
Pro Lys Glu Asn Arg Pro Leu Lys Glu Gly Glu Ala Tyr Gly Val Trp Ser Phe Ile Ala Leu Ser Leu Ser Asn Asp Arg
CCA AAG GAA AAT AGA CCT TTA AAA GAA GGA GAA GCT TAT GGT GTA TGG TCA TTT ATA GCC TTA TCA CTA TCA AAT GAT AGA
                210                                     220                                     230
Asp His Cys Ala Asp Leu Phe Ile Glu Asp Ala Gly Leu Trp Thr Lys Asn Asp Asn Pro Glu Asp Leu Lys Lys Phe Leu
GAT CAC TGT GCT GAT TTA TTC ATA GAA GAT GCA GGA CTT TGG ACA AAA AAT GAT AAT CCA GAA GAC TTA AAG AAA TTC CTA
                            240                                     250
Glu Asp His Arg Lys Ala Val Thr Trp Ser Val Val Glu Cys Gly Arg Asp Ser His Val Val Phe Glu Arg Thr Tyr Ile
GAA GAT CAT AGA AAA GCT GTA ACT TGG TCA GTA GTT GAA TGC GGT AGA GAT TCT CAT GTT GTT TTT GAA AGA ACT TAC ATA
260                                     270                                     280
Gly Phe Ala Tyr Val Ile Met Lys Pro Gly Glu Ile Gly Asn Ala Leu Thr Cys Ala Pro Tyr Val Thr Leu Ala Arg Asp
GGT TTT GCT TAT GTA ATA ATG AAA CCT GGT GAA ATT GGA AAC GCT TTA ACT TGC GCT CCT TAT GTA ACC TTA GCA AGA GAT
            290                                     300                                     310
Ala Val Pro Ser Glu Gly Phe Pro Ser Leu Asn Arg Ile Ser Leu Ser Gln Trp Leu Asp Asp Met Asn Phe Asp Ser Leu
GCT GTT CCT TCT GAA GGT TTC CCA AGT TTA AAT AGA ATT TCT CTT TCT CAA TGG TTA GAT GAT ATG AAT TTT GAT TCT TTA
                30                                                                           1
Val Asn Pro Ser Lys Lys End                                                              Met Asn Asp Asn Lys Lys Leu Gly Val
GTT AAT CCT AGT AAA AAA TAA TCAATAATAATTTCTAAAACAAAATTATTCTAGGAGGAGTTATT ATG AAT GAT AAT AAA AAA TTA GGT GTA
    10                                      20
Ile Ala Leu Ala Gly Val Val Ile Ser Ala Met Leu Gly Gly Gly Val Tyr Asn Leu Pro Gln Asn Met Ala Gln Ser
ATT GCT CTT GCT GGG GTA GTA ATT AGT GCT ATG CTG GGT GGA GGT GTT TAT AAC CTT CCC CAA AAC ATG GCC CAA TCC-3'
```

FIGURE 3: Sequences of the *hdc* gene of *C. perfringens* and its flanking regions and the deduced amino acid sequence of HisDCase. The putative promoter (−10 and −35 regions) and ribosome binding sites are underlined. The Ser–Ser bond that is cleaved upon activation of proHisDCase is indicated by an asterisk.

Comparison of the Amino Acid Sequences of the ProHisDCases and Derived HisDCases from C. perfringens and Lactobacillus 30a. The alignment of the amino acid sequence of the proHisDCase from *C. perfringens* with the only other proHisDCase for which the entire sequence is known, that of *Lactobacillus* 30a (Figure 4), shows an overall homology of 47.4%. The homology between the nucleotide sequences of their genes (not shown) is higher: 60.5%. The most striking similarities in sequence occur around the cleavage sites of the proenzymes: there is a 71% homology between residues 85–108 of the *C. perfringens* proHisDCase and the corresponding region (residues 71–92) of the *Lactobacillus* proenzyme. The lowest similarities occur near the N- and C-termini of the proteins.

All of the residues implicated in catalysis and substrate binding in the HisDCase from *Lactobacillus* 30a (Gallagher et al., 1989) are conserved in the *C. perfringens* enzyme, indicating that the catalytic mechanisms of these enzymes are closely similar. The conserved residues include those forming the predominantly hydrophobic pocket surrounding the carboxyl group of L-histidine in the *Lactobacillus* enzyme[2] [Ala-80(95), Ala-153(170), Ile-180(197), Phe-195(212), and Ala-260(278)] and the potential proton donors/acceptors [Glu-197(214) and Lys-155(172)], the residues that participate in hydrogen bonding to both imidazole nitrogen atoms of the substrate [Asp-63(78) and Ser-81(96)] and in hydrophobic interactions with the imidazole ring [Ile-59(74) and Phe-83-(98)], and the residues forming a chain of hydrogen bonds near the active site [Ser-81(96), Tyr-62(77), and Glu-66(81)].

There is no homology between the amino acid sequence of pyridoxal phosphate dependent HisDCase from *Morganella morganii* (Vaaler et al., 1986) and that of the pyruvoyl-dependent HisDCase from either *C. perfringens* or *Lactobacillus* 30a.

Flanking Nucleotide Sequences. A typical promoter sequence (Rosenburg & Court, 1979) precedes the clostridial *hdc* gene: a distinct −35 (TTG) region is evident as well as a −10 (TATAAT) region (Figure 3). However, the A cluster at positions −45 to −41, the TG at −16 and −15, and the AT-rich region between −3 and −7, which were postulated to form an "extended" promoter element in other Gram-positive genes (Graves & Rabinowitz, 1986), are not present. A traditional ribosome binding site [(Shine & Delgarno, 1975) AGGA] is present 94 bases downstream from the Pribnow–Schaller box and is separated from the translation initiation codon (ATG) by 9 bases.

A second open reading frame is located 44 bases downstream from the termination codon of the *hdc* gene (Figure 3). This reading frame, which codes for at least 34 amino acids

[2] Numbers in parentheses indicate the corresponding residues in the *C. perfringens* HisDCase.

```
HisDCase          10         20         30         40         50
   CP     NKNLEANRNRTLSEGIHKNIKVRAPKIDKTAISPYDRYCDGYGMPGAYGNGYVSV
                                 K        D  AISPY      GY  PG  GNGYV
  L30a                          SELDAKLNKLGVDRIAISPYKQWTRGYMEPGNIGNGYVTG
                                          10         20         30

          60         70         80         90        100
   Cp     LKVSVGTVKKTDDILLDGIVSYDRAEINDAYVGQINMLTASSFCGVAGQVWGHDL
          LKV  G   K DD  LDGIVSYDRAE   AY GQINM TASSF GV G V G D
  L30a    LKVDAGVRDKSDDDVLDGIVSYDRAETKNAYIGQINMTTASSFTGVQGRVIGYDI
          40         50         60         70         80         90

                   120        130        140        150        160
   Cp     ATHDSIANDEIKPLYELKQFDGTPLKVYDAKPLLEAGIELFGTEKNRRFTTAPGA
            S   D  KPL    Q DG  L  YDAKPL  A  E FGTE  RR   APG
  L30a    --LRSPEVDKAKPLFTETQWDGSELPIYDAKPLQDALVEYFGTEQDRRHYPAPGS
                100        110        120        130        140

          170        180        190        200        210
   Cp     HVICANKSATAYRPKENRPLKEGEAYGVWSFIALSLSNDRDHCADLFIEDAGLWT
            CANK  TA RPK     K G  YGVWS  IA S   D          F EDAG W
  L30a    FIVCANKGVTAERPKNDADMKPGQGYGVWSAIAISFAKDPTKDSSMFVEDAGVWE
           150        160        170        180        190        200

                   230        240        250        260        270
   Cp     KNDNPEDLKKFLEDHRKAVTWSVVECGRDSHVVFERTYIGFAYVIMKPGEIGNAL
                L   LE  RKA    S  ECG D H   FE   IGFAY   M PG IGNA
  L30a    TPNE-DELLEYLEGRRKAMAKSIAECGQDAHASFESSWIGFAYTMMEPGQIGNAI
                210        220        230        240        250

          280                  290        300        310
   Cp     TCAPYVTLARDAVP-------SEGFPSLNRISLSQWLDDMNFDSLVNPSKK
          T APYV L  D  P                          WL  M   SL
  L30a    TVAPYVSLPIDSIPGGSILTPDKDMEIMENLTMPEWLEKMGYKSLSANNALKY
            260        270        280        290        300
```

FIGURE 4: Alignment of the amino acid sequences of the proHisDCases from *C. perfringens* (Cp) and *Lactobacillus* 30a (L30a). Identical amino acids are printed between the two sequences. The boxed residues show the amino acids implicated in catalysis or substrate binding in the *Lactobacillus* HisDCase (Gallagher et al., 1989). All are conserved in the clostridial enzyme. Dashes indicate gaps in the sequence introduced to optimize the alignment.

(only part of it was cloned), is preceded by a ribosome binding site (AGGA) but not by a Pribnow–Schaller box. This suggests that the mRNA encoding this protein could be cotranscribed with that of HisDCase in *C. perfringens*. The N-terminal region of this protein has a 17% homology to the corresponding region of a 173-residue protein encoded by a similarly positioned coding sequence downstream from the *hdc* gene of *Lactobacillus* 30a which is cotranscribed with this gene 15% of the time (Copeland, 1987). Possible roles of these downstream proteins include the regulation of the *hdc* gene, the transport of histidine and/or histamine, or perhaps the metabolism of histamine.

Expression of the hdc Gene in E. coli. A computer-generated restriction analysis of the nucleotide sequences of the *Hin*dIII and *Hgi*AI–*Pst*I fragments, which together contain the complete *hdc* gene (Figure 2), indicated the presence of a unique *Afl*II site in the coding region of HisDCase that was common to both fragments, and a unique *Bal*I site between the Pribnow–Schaller box and the ribosome binding site in the upstream sequence of the gene. These restriction sites made it possible to reassemble the complete coding region of the gene at the *Afl*II site and to remove the Pribnow–Schaller box from the *hdc* promoter, thus allowing the expression of the *hdc* gene in *E. coli* GM1 under control of the *lac* promoter of pUC8 (Figure 5). The correct construction of the expression vector (pVP-4) was verified by sequencing one strand of the insert containing the *hdc* gene using the four specific oligonucleotide primers described earlier. The presence of stop codons in all three reading frames upstream from the ribosome binding site, preceding the coding region of the *hdc* gene, ensured that the gene product was not fused to the β-galactosidase encoded by pUC8. Cells transformed with pVP-4 in the absence of IPTG exhibited <2% of the HisDCase activity of cells induced with IPTG (1 mM); expression of the *hdc* gene was clearly under control of the *lac* promoter region in pVP-4.

Purification and Properties of Cloned HisDCase. HisDCase was purified to homogeneity from transformed *E. coli* GM1 as described under Materials and Methods (see Table

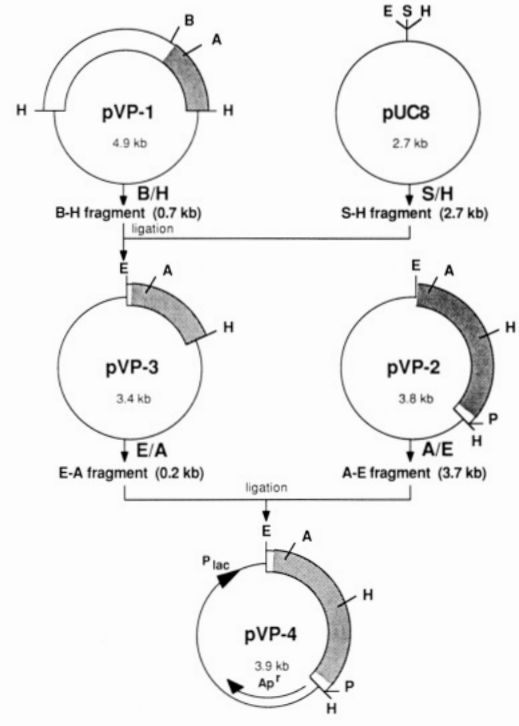

FIGURE 5: Reconstruction and expression of the gene encoding histidine decarboxylase from *C. perfringens* under control of the *lac* promotor of pUC8. The coding region of the *hdc* gene is shaded; the open box represents flanking regions. Abbreviations: A, *Afl*II; B, *Bal*I; E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; S, *Sma*I. See text for details.

II). Purity of the preparation was verified by SDS–polyacrylamide gel electrophoresis (Figure 6, lane 6). About 11% of the protein solubilized by French press extraction at pH 7.0 was HisDCase. In *C. perfringens*, the enzyme makes up approximately 1% of the total protein (Recsei et al., 1983b).

The cloned HisDCase contains dissimilar α and β subunits (Figure 6, lane 6) with approximate molecular weights of 25 000 and 10 000, respectively (Figure 6, lane 7). These

Table II: Purification of Cloned Histidine Decarboxylase

step	vol (mL)	protein[a] (mg)	act.[b] [μmol/(min·mg)]	yield (%)	purification (x-fold)
crude extract	75	1007	2.7	100	
$(NH_4)_2SO_4$ fractionation	24	298	8.4	93	3
heat treatment	22	112	20.4	85	7.7
DEAE-cellulose	55	58	24	52	9.0
Sephacryl S-200	25	34	24.6	31	9.3

[a] Protein concentrations were determined by the method of Lowry et al. (1951) with bovine serum albumin as a standard. [b] Enzyme activity was measured by the release of $^{14}CO_2$ from carboxyl-labeled L-histidine (Recsei & Snell, 1970; Ichiyama, 1970) in 0.2 M ammonium acetate/0.4 M magnesium sulfate, pH 4.5.

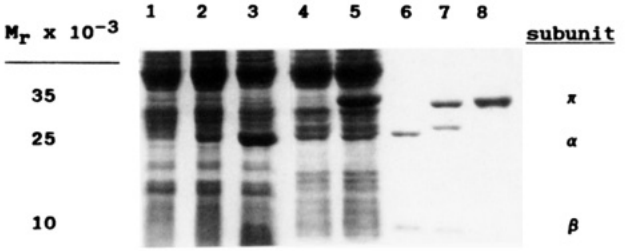

FIGURE 6: Electrophoretic characterization of crude and purified preparations of cloned HisDCase and S97A proHisDCase. Lanes 1 and 4, crude extract of uninduced cells; lanes 2 and 3, crude extract of cells expressing HisDCase induced with IPTG for 1 or 16 h, respectively; lane 5, crude extract of cells expressing S97A pro-HisDCase induced with IPTG for 16 h; lane 6, homogeneous cloned HisDCase; lane 7, partially activated proHisDCase from *Lactobacillus* 30a; lane 8, homogeneous S97A proHisDCase. Crude extracts were prepared by passage of harvested cells, suspended in 0.2 M ammonium acetate, pH 4.5, through a French pressure cell. The proHisDCase from *Lactobacillus* 30a was purified as described previously (van Poelje, 1988). The approximate molecular weights of the subunits of this proHisDCase and the derived HisDCase are as follows: π subunit, 34 000; α subunit, 25 000; β subunit, 10 000 (Huynh et al., 1984; Vaaler et al., 1982).

values agree well with those calculated from the deduced amino acid sequence of the *hdc* gene product (see above). Since cloned HisDCase from *C. perfringens* elutes from a Sephacryl S-200 column at the same point as the HisDCase from *Lactobacillus* 30a [$(\alpha\beta)_6$, M_r 208 000 (Hackert et al., 1981)], the cloned enzyme, like that isolated directly from *C. perfringens*, has an $(\alpha\beta)_6$ structure.

The π subunits of proHisDCase were not observed either in purified preparations of the cloned enzyme (Figure 6, lane 6) or in crude protein preparations from uninduced or induced cells of *E. coli* (Figure 6, lanes 1–3). The recombinant pro-HisDCase from *C. perfringens*, unlike that from *Lactobacillus* 30a (Copeland et al., 1987), thus appears to undergo rapid, complete conversion to HisDCase in *E. coli* cells.

To determine whether the 10 amino acid extension at the N-terminus encoded by the nucleotide sequence of the clostridial *hdc* gene was present in the enzyme expressed in *E. coli*, the β subunit of the cloned HisDCase was isolated (see Materials and Methods). Both manual and automated sequencing failed, suggesting that the N-terminus was blocked. However, the amino acid composition of the β subunit (Table III) clearly shows that the N-terminal decapeptide is present and that N-terminal formylmethionine is completely removed after protein synthesis.

Both the cloned enzyme and the HisDCase isolated from *C. perfringens* exhibit maximal activity at pH 4.5 and an ionic strength of 1.4 M. Under these conditions, the K_m values for L-histidine of cloned HisDCase (0.2 mM) and the enzyme from *C. perfringens* [0.3 mM (Recsei et al., 1983)] are similar, but the cloned enzyme is much more active [V_{max} = 25 vs 7 μmol/(min·mg)]. Whether this increased activity is due to differences in preparative procedures or to presence of the presequence on the β subunit of the cloned enzyme is not yet known. Neither dithiothreitol nor bovine serum albumin, routinely added during assay of the HisDCase from *C. perfringens* in earlier studies (Recsei et al., 1983b), was necessary for optimal activity of the cloned enzyme under the assay conditions used. The approximate p*I* of (cloned) HisDCase, determined by computer analysis of its amino acid sequence, is 6.05.

Table III: Amino Acid Composition of the β Subunit of Cloned HisDCase

amino acid	no. of residues (mol/10 523 g) theoretical[a]	found[b]
Ala	7 (6)	7
Arg	5 (3)	5
Asp (+Asn)	16 (12)	16
Cys	1	1
Glu (+Gln)	4 (3)	4
Gly	9	10
His	1	1
Ile	8	8
Leu	6 (5)	6
Lys	8 (7)	7
Met	2[c]	2
Pro	3	3
Ser	6	6
Thr	5	5
Tyr	7	6
Val	8	9

[a] Encoded by the *hdc* gene. Parentheses enclose the number of residues encoded if the presequence were missing. [b] Average of values obtained from three hydrolyses of the β subunit, rounded to the nearest integer. Samples were hydrolyzed in 6 N HCl at 110 °C in vacuo for 24 h and analyzed on a Beckman 121MB amino acid analyzer. [c] This figure does not include the N-terminal methionine residue.

Construction and Expression of the Gene for S97A Pro-HisDCase. The *Eco*RI–*Pst*I fragment (1.1 kb) of pVP-4 (see Figure 5), which contains the complete *hdc* gene, was subcloned into M13mp18, and the codon corresponding to Ser-97 (TCA) converted to an alanyl codon (GCT) by site-directed mutagenesis as described under Materials and Methods. The oligonucleotide primer employed in this procedure is shown in Figure 1C. Six of the resulting templates were screened by double-lane DNA sequencing; all contained the desired base changes. One of these templates was sequenced in its entirety by procedures described earlier. It was thus confirmed that a single codon had been altered. The RF DNA corresponding to this template was isolated, and the mutant *hdc* gene, excised by digestion with *Eco*RI and *Pst*I, was then cloned into appropriately restricted pUC8, and expressed under control of the *lac* promoter as described for the wild-type gene.

Purification and Properties of S97A ProHisDCase. Induction of the mutant gene in *E. coli* resulted in the appearance of a new protein band corresponding in molecular weight to the π subunit of proHisDCase from *Lactobacillus* 30a (Figure 6, lanes 4 and 5). Purification of this mutant protein to homogeneity was achieved (Figure 6, lane 8) by the procedure described for the cloned enzyme (see Materials and

Methods) but with two exceptions: (a) since the S97A protein lacks all HisDCase activity, purification was monitored by SDS gel electrophoresis, and (b) the heat treatment was conducted at 60 °C because the mutant protein precipitated at higher temperatures. The cloned wild-type enzyme, in contrast, remained soluble and active up to approximately 80 °C (4 min).

The S97A protein eluted from the final Sephacryl S-200 column in the same volume as HisDCase from *Lactobacillus* 30a; its molecular weight thus corresponds to that of a hexamer of π' subunits. The α and β subunits characteristic or HisDCase were not present in the pure preparation; the π' subunits of the S97A proenzyme are clearly not cleaved in vivo. Cleavage also did not occur in vitro within 48 h at 37 °C on exposure to a variety of K^+, Na^+, and NH_4^+ concentrations at pH values between 5 and 8.

Both proHisDCase and HisDCase from *Lactobacillus* 30a bind L-histidine with approximately equal affinity (van Poelje & Snell, 1988). Under the same conditions, no binding of histidine by the S97A protein was observed.

Discussion

We have determined the nucleotide sequence of the *hdc* gene of *C. perfringens* and from it have deduced the amino acid sequence of a proenzyme form of this HisDCase. By expression of the gene in *E. coli* and subsequent characterization of its product, we have established that the cloned proHisDCase undergoes a processing reaction closely similar to that of the proenzyme in *C. perfringens*: a specific peptide bond (between Ser-96 and Ser-97) is cleaved to yield the $(\alpha\beta)_6$ structure of the active enzyme. The cloned HisDCase resembles the enzyme isolated from *C. perfringens* in its subunit composition, its K_m for L-histidine, its pH optimum of catalysis, and the variation of its activity with ionic strength. However, the cloned enzyme has over 3 times the specific activity of the enzyme from *C. perfringens* and retains a blocked decapeptide presequence at the N-terminus of its β subunit (NKNLEANRNR) that was not found in *C. perfringens* enzyme. The function of this presequence and the nature of its blocking group are unknown. The abundance of Asn in the presequence and the presence of a eucaryotic consensus sequence for glycosylation [N-X-T/S (Lennarz & Struck, 1980)] at residues 9–11 of the intact β subunit are of interest.

The rapid and complete conversion of cloned clostridial proHisDCase to HisDCase in *E. coli* is in contrast to the slow rate of cleavage observed (Copeland et al., 1987) for the cloned proHisDCase of *Lactobacillus* 30a in the same organism. Since the latter proenzyme is converted to HisDCase rapidly in *Lactobacillus* 30a, its slow activation in *E. coli* has led to the speculation (Copeland et al., 1987) that its activation is not entirely autocatalytic as originally reported (Recsei & Snell, 1973) but that an "activation factor" absent from *E. coli* cells is required for its rapid conversion to HisDCase. Our observations that the cloned clostridial proHisDCase, which bears a 47% homology to the *Lactobacillus* proenzyme, undergoes complete cleavage to HisDCase in *E. coli* cells in the absence of a clostridial "activation factor" do not support such speculation. Limited evidence indicates that the reduced activation rate of the cloned *Lactobacillus* 30a proenzyme in the recombinant host may result from aberrant folding (van Poelje, 1988).

In view of the high degree of homology at and near their cleavage sites, it is likely that the clostridial proHisDCase is processed at the Ser^{96}–Ser^{97} bond via the same serinolysis reaction observed in the proenzyme from *Lactobacillus* 30a (Recsei et al., 1983a) and *L. buchneri* (Recsei & Snell, 1985). The key role of the hydroxyl group of Ser-97 in this mechanism is supported by the insusceptibility of the S97A proenzyme to cleavage. This evidence, however, is not conclusive since the substitution of Ser-97 by Ala appears to affect proenzyme structure (see below). The preceding serine residue (Ser-96) plays no essential role in the activation reaction as formulated: it is therefore of interest that dissimilar peptide bonds are cleaved during biogenesis of other pyruvoyl enzymes: a Lys–Ser bond in *pro-S*-adenosylmethionine decarboxylase of *E. coli* (Tabor & Tabor, 1987) and a Gly–Ser bond in *pro*-phosphatidylserine decarboxylase, also from *E. coli* (Li & Dowhan, 1988). In each case, however, the catalytically essential pyruvoyl group is derived from the serine residue at the processing site.

Judging from its substantially decreased temperature stability and lack of affinity for histidine, the conformation of the S97A proenzyme differs from that of the clostridial HisDCase and, since the conformations of the *Lactobacillus* 30a proHisDCase and HisDCase are virtually indistinguishable (Hackert et al., 1981), probably from the wild-type clostridial proHisDCase as well. In accordance with this view, a comparison showed that instead of the extended conformation at the activation site predicted for the wild-type proenzyme, the mutant proenzyme has a high probability of forming a helical structure in this region. The perturbation of structure and decreased heat stability as a result of the conservative substitution of a single amino acid residue (and perhaps the consequent loss of a critical hydrogen bond) are not without precedent: substitution of Thr-157 of phage T4 lysozyme by Ile resulted in an 11 °C decrease in T_m and was attributed primarily to the inability of the substituted residue to form a hydrogen bond with the amide nitrogen of Asp-159 (Alber et al., 1987).

Acknowledgments

We thank Gloria Vaaler for helpful discussions regarding cloning procedures, Sandra Smith for amino acid analyses and automated protein sequencing, and Karen Browning for synthesis of oligonucleotides.

References

Alber, T., Sun, D., Wilson, K., Wozniak, J. A., Cook, S. P., & Matthews, B. W. (1987) *Nature 330*, 41–46.

Birnboim, H. C., & Doly, J. (1979) *Nucleic Acids Res. 7*, 1513–1523.

Carlson, J., & Messing, J. (1984) *Bio/Technology 1*, 253–264.

Chang, J. Y., Brauer, D., & Wittmann-Liebold, B. (1978) *FEBS Lett. 93*, 205–214.

Copeland, W. C. (1987) Doctoral Thesis, University of Texas at Austin, Austin, TX.

Copeland, W. C., Vanderslice, P., & Robertus, J. D. (1987) *Protein Eng. 1*, 419–423.

Crestfield, A. M., Stein, W. H., & Moore, S. (1963) *J. Biol. Chem. 238*, 2413–2420.

Fairbanks, G., Steck, T. L., & Wallach, D. F. H. (1971) *Biochemistry 10*, 2606–2617.

Felton, J. (1983) *BioTechniques 1*, 42.

Gallagher, T., Snell, E. E., & Hackert, M. L. (1989) *J. Biol. Chem. 264*, 12737–12743.

Garnier, J., Osguthorpe, D. J., & Robson, B. (1978) *J. Mol. Biol. 120*, 97–120.

Graves, M. C., & Rabinowitz, J. C. (1986) *J. Biol. Chem. 261*, 11409–11415.

Hackert, M. L., Clinger, K., Ernst, S. R., Parks, E. M., & Snell, E. E. (1987) in *Crystallography in Molecular Biology*

(Wilson, K., Ed.) pp 403–411, Plenum, New York.
Hackert, M. L., Meador, W. E., Oliver, R. M., Salmon, J. B., Recsei, P. A., & Snell, E. E. (1981) *J. Biol. Chem. 256*, 687–690.
Holmes, D. S., & Quigley, M. (1981) *Anal. Biochem. 114*, 193–197.
Huynh, Q. K., & Snell, E. E. (1985a) *J. Biol. Chem. 260*, 2789–2803.
Huynh, Q. K., & Snell, E. E. (1985b) *J. Biol. Chem. 260*, 2794–2797.
Huynh, Q. K., Vaaler, G. L., Recsei, P. A., & Snell, E. E. (1984) *J. Biol. Chem. 259*, 2826–2839.
Ichiyama, A. (1970) *J. Biol. Chem. 245*, 1699–1709.
Kreil, G. (1981) *Annu. Rev. Biochem. 50*, 317–348.
Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488–492.
Kunkel, T. A., Roberts, J. D., & Zabour, R. A. (1987) *Methods Enzymol. 154*, 367–382.
Laemmli, U. K. (1970) *Nature (London) 227*, 680–685.
Li, Q.-X., & Dowhan, W. (1988) *J. Biol. Chem. 263*, 11516–11522.
Lowry, G. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. (1951) *J. Biol. Chem. 193*, 265–275.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1983) *Molecular Cloning: A laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Messing, J., Crea, R., & Seeburg, P. H. (1981) *Nucleic Acids Res. 9*, 309–321.
Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Miller, J. H., Ganem, D., Lu, P., & Schmitz, A. (1977) *J. Mol. Biol. 109*, 275–301.
Nakamura, K., Pirtle, R. M., & Inouye, M. (1979) *J. Bacteriol. 137*, 595–604.
Needleman, S. J., & Wunsch, C. D. (1970) *J. Mol. Biol. 48*, 443–453.
Parks, E. H., Ernst, S. R., Hamlin, R., Xuong, Ng. H., & Hackert, M. L. (1985) *J. Mol. Biol. 182*, 455–465.
Recsei, P. A., & Snell, E. E. (1970) *Biochemistry 9*, 1492–1497.
Recsei, P. A., & Snell, E. E. (1973) *Biochemistry 12*, 365–371.
Recsei, P. A., & Snell, E. E. (1985) *J. Biol. Chem. 260*, 2804–2806.
Recsei, P. A., Huynh, Q. K., & Snell, E. E. (1983a) *Proc. Natl. Acad. Sci. U.S.A. 80*, 973–977.
Recsei, P. A., Moore, W. M., & Snell, E. E. (1983b) *J. Biol. Chem. 258*, 439–444.
Rosenberg, M., & Court, D. (1979) *Annu. Rev. Genet. 13*, 319–353.
Rosenthaler, J., Guirard, B. M., Chang, G. W., & Snell, E. E. (1965) *Proc. Natl. Acad. Sci. U.S.A. 54*, 152–158.
Semina, L. A., & Mardashev, S. R. (1965) *Biokhimiya 30*, 100–106.
Shine, J., & Delgarno, C. (1975) *Nature 254*, 34–38.
Smith, L. DS., & Hobbs, G. (1974) in *Bergeys Manual of Determinative Bacteriology*, 8th ed., Waverly Press Inc., Baltimore, MD.
Struck, D. K., & Lennarz, W. J. (1980) in *The Biochemistry of Glycoproteins and Proteoglycans* (Lennarz, W., Ed.) pp 85–160, Plenum, New York.
Tabor, C. W., & Tabor, H. (1987) *J. Biol. Chem. 262*, 16037–16040.
Vaaler, G. L., Recsei, P. A., Fox, J. W., & Snell, E. E. (1982) *J. Biol. Chem. 257*, 12770–12774.
Vaaler, G. L., Brasch, M. A., & Snell, E. E. (1986) *J. Biol. Chem. 261*, 11010–11014.
Vanderslice, P., Copeland, W. C., & Robertus, J. D. (1986) *J. Biol. Chem. 261*, 15186–15191.
Van Poelje, P. D. (1988) Doctoral Thesis, University of Texas at Austin, Austin, TX.
Van Poelje, P. D., & Snell, E. E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8449–8453.
Vieira, J., & Messing, J. (1982) *Gene 19*, 259–268.
Woods, D. (1984) *Focus 6*, 1–2.
Yannisch-Perron, C., Vieira, J., & Messing, J. (1985) *Gene 33*, 103–119.