Site-Directed Mutagenesis and ^{1}H NMR Spectroscopy of an Interdomain Segment in the Pyruvate Dehydrogenase Multienzyme Complex of *Escherichia coli*†

Frieda L. Texter,‡,§ Sheena E. Radford,‡ Ernest D. Laue,‡ Richard N. Perham,*,‡ John S. Miles,∥ and John R. Guest∥

Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, England, and Department of Microbiology, University of Sheffield, Sheffield S10 2TN, England

Received June 23, 1987; Revised Manuscript Received September 9, 1987

ABSTRACT: Deletion of two of the three homologous lipoyl domains that form the N-terminal half of each dihydrolipoamide acetyltransferase (E2p) polypeptide chain of the *Escherichia coli* pyruvate dehydrogenase complex can be achieved by in vitro deletion in the structural gene *aceF*. A site-directed mutagenesis of this shortened *aceF* gene was carried out to replace the glutamine residue at position 291 (wild-type numbering) with a histidine residue. Residue 291 is near the middle of a long segment (about 30 amino acid residues) of polypeptide chain, rich in alanine, proline, and charged amino acids, that links the remaining lipoyl domain to the dihydrolipoamide dehydrogenase (E3) binding domain in the E2p chain. A fully active enzyme complex was still assembled, and despite the enormous size of the particle (M_r approximately 4 × 10^6), sharp resonances attributable to the single new histidine residue per E2p chain could be detected in the 400-MHz ^{1}H NMR spectrum of the complex. The sharpness of these resonances, their chemical shifts (7.94 and 7.05 ppm), and the apparent pK_a (6.4) of the side chain were all consistent with this histidine residue being exposed to solvent in a conformationally flexible region of the E2p polypeptide chain. These experiments provide direct proof for the conformational flexibility of this region of polypeptide chain, which is thought to play an important part in the movement of the lipoyl domain required for active site coupling in the enzyme complex. The major sharp resonance (at 1.39 ppm) in the 400-MHz ^{1}H NMR spectrum of the mutated complex, which is derived from the β-CH_3 protons of alanine residues in the long (alanine + proline)-rich region of the E2p chain, was somewhat smaller than expected, suggesting that the insertion of the histidine residue at position 291 had diminished the flexibility of some at least of the alanine residues in this segment of polypeptide chain. Another sharp peak (at 1.52 ppm) was also clearly visible in the spectrum, tentatively attributed to a smaller (alanine + proline)-rich region in the E2p chain at positions 370–377. This could provide for conformational flexibility between the E3-binding domain and the catalytic (and inner core forming) domain of the E2p chain, which may also be contributing to the mechanism of active site coupling in the complex.

The structural core of the pyruvate dehydrogenase multienzyme (PDH) complex of *Escherichia coli* is composed of 24 dihydrolipoyl acetyltransferase (E2p, EC 2.3.1.2) polypeptide chains arranged with octahedral symmetry. To this are bound multiple copies of pyruvate decarboxylase [pyruvate dehydrogenase (lipoamide), E1p, EC 1.2.4.1] and dihydrolipoamide dehydrogenase (E3, EC 1.8.1.4, formerly EC 1.6.4.3) subunits. The lipoyllysine residues, which act as swinging arms to transfer substrate between successive active sites in the mechanism, are housed in lipoyl domains of the E2p chains that protrude between the E1p and E3 components [for reviews, see Reed (1974) and Perham (1983)].

The complex is encoded by three genes, *aceE* (E1p), *aceF* (E2p), and *lpd* (E3), which form the *aceEF-lpd* operon, the complete DNA sequence of which is known (Stephens et al., 1983a–c). The N-terminal half of the E2p polypeptide chain (Stephens et al., 1983b) comprises three highly homologous lipoyl domains, each of about 80 amino acid residues, in tandem array. These can be released by limited proteolysis of the complex (Bleile et al., 1979; Hale & Perham, 1979a) and recovered as three independently folded entities that retain their function as substrates for reductive acetylation by E1p (Packman et al., 1984). The lipoyl domains are separated from each other and linked to a further domain, which consists of about 50 residues and is intimately involved in binding the E3 component (Packman & Perham, 1986), by means of long segments (20–30 amino acids) of polypeptide chain rich in alanine, proline, and charged amino acid residues (Stephens et al., 1983b). The E3-binding domain is in turn linked to the remainder of the E2p chain by a somewhat similar but shorter and less conspicuous sequence (Stephens et al., 1983b; Packman & Perham, 1986, 1987). The C-terminal portion (M_r 28 000) of the E2p chain is currently envisaged as a larger domain which is responsible for binding E1p and other E2p chains as well as providing the acetyltransferase active site (Packman & Perham, 1987; Radford et al., 1987).

Genetic reconstruction of the *aceF* gene has demonstrated that the overall catalytic activity of the PDH complex and the system of active site coupling are not detectably affected by the removal of one or even two lipoyl domains per E2p chain (Guest et al., 1985). The complex encoded by plasmid pGS110, in which each E2p chain contains only one lipoyl domain (Figure 1), is thus a structurally simpler and con-

† This work was supported by grants from the Science and Engineering Research Council (to J.R.G. and R.N.P.) and from Albright College (to F.L.T.). J.S.M. and S.E.R. were supported by research studentships from the Medical Research Council and the Science and Engineering Research Council, respectively. The Cambridge authors are members of the Cambridge Biochemical NMR Group, whose facilities were funded jointly by the SERC and the MRC.

* Author to whom correspondence should be addressed.

‡ University of Cambridge.

§ Permanent address: Albright College, P.O. Box 516, Reading, PA 19603.

∥ University of Sheffield.

0006-2960/88/0427-0289$01.50/0

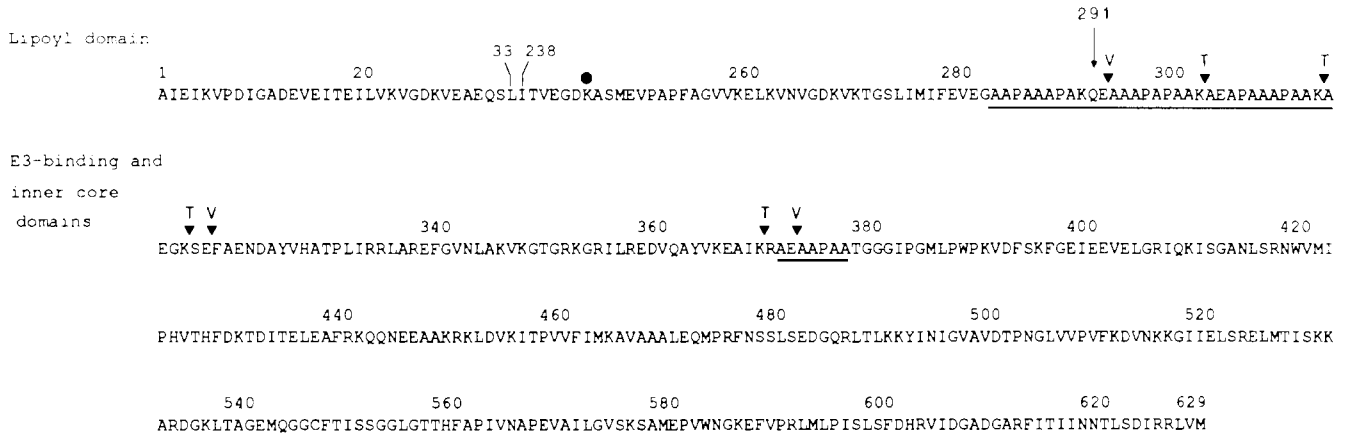

FIGURE 1: Amino acid sequence of the pGS110-encoded E2p chain of the *E. coli* PDH complex inferred from the DNA sequence of the *aceF* structural gene (Stephens et al., 1983b; Guest et al., 1985). The residue numbering is that of the wild-type E2p chain (Stephens et al., 1983b). The covalently bound lipoyl group is indicated (●), and the (alanine + proline)-rich regions are underlined. The cleavage sites expected for limited proteolysis with trypsin (T) and *S. aureus* V8 proteinase (V) are taken from Packman and Perham (1987). The amino acid sequence of the pGS156-encoded complex is identical except that residues 284–295 are deleted. The residue Gln-291 (marked with an arrow) was changed to a histidine residue in the pGS178-encoded complex.

venient system for further study. For example, it has been found that the complex encoded by plasmid pGS156, in which a deletion of 12 residues (positions 284–295) has been engineered in the long (alanine + proline)-rich segment that links the single lipoyl domain to the E3-binding domain (Figure 1), is also without detectable change in catalytic activity or active site coupling (Miles et al., 1987).

Active site coupling in the complex involves movement of the lipoyl domains (Berman et al., 1981; Stepp et al., 1981), and strong evidence in favor of the existence of conformationally flexible regions in the E2p chains has come from ^{1}H NMR spectroscopy (Perham et al., 1981; Roberts et al., 1983). A major source of the unusually sharp signals seen in the ^{1}H NMR spectrum of the complex is thought to be the three long (alanine + proline)-rich sequences which link the lipoyl domains to each other and to the E3-binding domain (Spencer et al., 1984; Packman et al., 1984; Radford et al., 1986, 1987). In addition, there is evidence that other sharp signals in the ^{1}H NMR spectrum of the complex may arise from the shorter (alanine + proline)-rich segment of E2p chain which tethers the E3-binding domain to the catalytic domain (Radford et al., 1987).

In the present paper we describe a site-directed mutagenesis of the pGS110-encoded complex that replaces the glutamine residue at position 291 in the E2p chain (Figure 1) with a histidine residue. Histidine was chosen because it has two characteristic resonances located in the aromatic region of the ^{1}H NMR spectrum (Jardetzky & Roberts, 1981), a region conveniently devoid of sharp signals in the ^{1}H NMR spectra of the wild-type and pGS110-encoded *E. coli* PDH complexes (Perham et al., 1981; Radford et al., 1987). A typical complex (M_r approximately 4×10^6) was found to be assembled and to be fully active. Despite the enormous size of the particle, resonances arising from the single new histidine residue per E2p chain could be detected in the 400-MHz ^{1}H NMR spectrum of the complex. This provides formal proof that the peptide chain surrounding the histidine residue is conformationally flexible and has enabled us to study the long (alanine + proline)-rich region directly. It is likely that such use of protein engineering techniques could find widespread application in studies of conformational flexibility in proteins.

Materials and Methods

Materials. Restriction endonucleases, T_4 DNA ligase, T_4 polynucleotide kinase, and DNA polymerase (Klenow fragment) were purchased from Bethesda Research Laboratories and Boehringer Corp. The 17-mer "universal" primer was from Celltech, and the remaining sequencing primers were either gifts from G. D. Searle Ltd., ICI Pharmaceuticals plc, and New Brunswick Instrument Corp. or were synthesized on an Applied Biosystems 381A machine. α-[^{35}S]Thio-dATP and [2-^{14}C]pyruvate were supplied by New England Nuclear, and [γ-^{32}P]ATP was from Amersham International. Soybean trypsin inhibitor and 3-(trimethylsilyl)-2,2,3,3-tetradeuteriopropionic acid (TSP) were obtained from BDH Biochemicals. Trypsin [N^{α}-tosylphenylalanine chloromethyl ketone (TPCK) treated] was purchased from Worthington, and *Staphylococcus aureus* V8 proteinase was from Miles Laboratories. Deuterium oxide (99.8%) was a product of Norsk Hydro.

E. coli Strains, Plasmids, and Bacteriophages. The PDH complex deletion strain, JRG1342 (*aceEF-lpd* Δ*recA*), used for nutritional complementation tests and to provide a suitable background for mutant PDH complex expression, has been described by Guest et al. (1985). Strains JM101, for propagating M13 phages, and BMH71-18*mutL*, for improving mutagenic frequency, have been described elsewhere (Messing, 1979; Kramer et al., 1984). Plasmid pGS110, which expresses the "one-lipoyl domain" PDH complex, has also been described previously (Guest et al., 1985). Bacteriophages M13mp18 and M13mp19 were from Pharmacia P-L Biochemicals.

Oligonucleotide-Directed Mutagenesis. The 17-mer oligonucleotide GCGAAACACGAAGCGGC (S5) was synthesized by the phosphotriester method (Sproat & Gait, 1984) and purified by denaturing gel electrophoresis (Atkinson & Smith, 1984). The "two-primer extension plus ligation" method (Zoller & Smith, 1984) was used for mutagenesis with primer S5 and the universal (17-mer) sequencing primer. The products of the mutagenesis reaction were used to transfect BMH71-18*mutL*, but the plaques were generated by plating in a lawn of JM101 to avoid unnecessary exposure of the phages to the mutator strain (Carter et al., 1985). Mutant phages were identified by the dot–hybridization procedure using ^{32}P-labeled mutagenic primer as a probe (Zoller & Smith, 1983).

Cassette Replacement. A cassette-replacement method was used, transferring the mutation from M13 to a plasmid for expression of the novel PDH complex. The 1.68 kilobase pair (kbp) *Kpn*I–*Sph*I fragment of pGS110 was replaced by a double-stranded DNA fragment prepared from the mutant M13 clone by primer extension (Hong, 1981) according to the

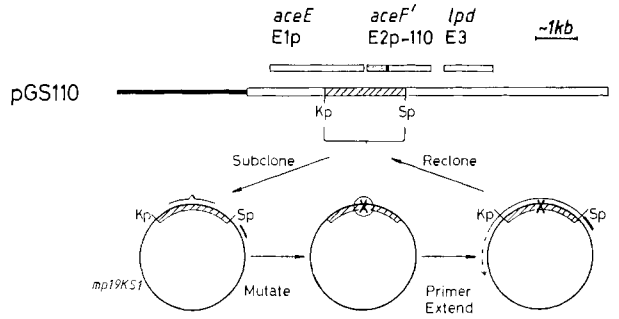

FIGURE 2: Cassette-replacement approach to in vitro mutagenesis of the *aceF'* gene. Plasmid pGS110 contains an *ace-lpd* operon with a mutant *aceF* gene (*aceF'*) encoding the E2p-110 chain with only one lipoyl segment (stippled bar) and (alanine + proline)-rich sequence (solid bar); the E1p and E3 components are wild type. The 1.68-kbp *KpnI–SphI* fragment (Kp–Sp) was subcloned into M13mp19 to give mp19KS1. The remaining 11.5-kbp fragment served as the receptor for recloning the mutant fragment. After primer extension across the mutation, the *KpnI–SphI* fragment was recloned to give a pGS110-like derivative plasmid, pGS178.

procedure of Miles and Guest (1987). The newly synthesized DNA was restricted with *KpnI* plus *SphI* and the relevant fragment cloned into the gel-purified receptor fragment of pGS110 according to standard procedures (Figure 2).

Plasmid Characterization. Small-scale plasmid preparations were made according to Birnboim and Doly (1979). The procedures used for digestion with restriction endonucleases and agarose gel electrophoresis have been described (Guest et al., 1983). Plasmid dot–blots were prepared using the denaturation/neutralization procedure of Maniatis et al. (1982), and the mutant plasmids were identified by using the same hybridization and washing procedure as for the M13 clones.

Nucleotide Sequence Analysis. Single-stranded M13 templates were prepared and sequenced by the dideoxy chain-termination method using α-[^{35}S]thio-dATP and salt-gradient gels (Sanger et al., 1980; Biggin et al., 1983). A series of synthetic oligonucleotides was used to confirm the sequence of the 1.68-kbp *KpnI–SphI* fragment (Figure 3).

Nutritional Complementation Tests. The Ace+/− and lpd+/− nutritional phenotypes of plasmid-containing transformants of JRG1342 were tested by the method of Guest et al. (1985).

Enzyme Purification and Characterization. The pGS110- and pGS178-encoded complexes were purified by the method of Danson et al. (1979), as modified by Guest et al. (1985). PDH complex activity was measured by the NAD^+-reduction assay (Packman et al., 1983). Polypeptide chain ratios (E1p:E2p:E3) were measured by using the radioamidination method of Hale and Perham (1979b), at pH 10.0. To assess the extent of active site coupling, samples were prepared in which the E1p component was partly inhibited by the transition-state analogue thiamin thiothiazolone pyrophosphate (Gutowski & Lienhard, 1976). For each sample the proportion of the E2p component acetylated by incubation with [2-^{14}C]pyruvate, at 37 °C, in the absence of coenzyme A, was then determined to generate an acetylation curve (Packman et al., 1983).

Limited Proteolysis. The pGS110- and pGS178-encoded complexes [5 mg/mL in 20 mM sodium phosphate buffer, pH 7.0, containing 2.7 mM ethylenediaminetetraacetic acid (EDTA) and 0.02% NaN_3] were digested at 0 °C with trypsin (0.5% w/w) or *S. aureus* V8 proteinase (0.5% w/w), as described by Packman et al. (1984). At various times, samples were removed and assayed for PDH activity, and others were taken and examined by polyacrylamide gel electrophoresis in the presence and absence of sodium dodecyl sulfate (SDS) (Packman et al., 1984).

^{1}H NMR Spectroscopy. Samples for ^{1}H NMR spectroscopy were dialyzed exhaustively against ^2H_2O buffer containing 20 mM potassium phosphate, 2.7 mM EDTA, and 0.02% NaN_3, p^{2}H 7.0. ^{1}H NMR spectra were obtained at 400 MHz with a Bruker AM400 spectrometer, using a 10-kHz spectral width and a 90° pulse. The sample temperature was ca. 25 °C. Before Fourier transformation, the free induction decay was multiplied by a negative exponential corresponding to a line broadening of 2 Hz. Chemical shifts are expressed relative to internal TSP. Spin-echo spectra were obtained by using the pulse sequence

$$(90°–t–180°–t–\text{acquire}–T)_n$$

where the variable delay t was 0.5 ms and the relaxation delay T was 10 s (equivalent to $5T_1$ of the methyl protons of TSP). The intensity of the peaks was evaluated by cutting them out from chart paper and comparing the weight of the peaks with that of a peak arising from a known concentration of TSP.

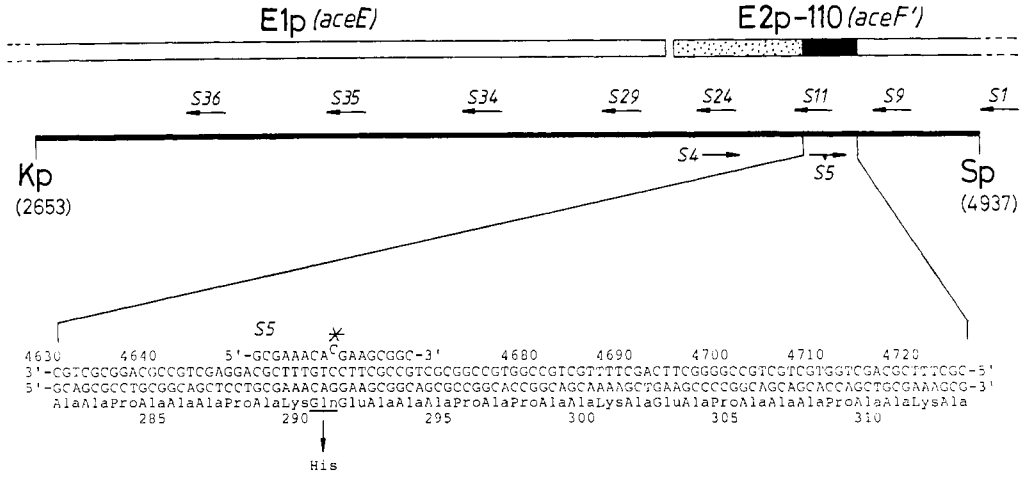

FIGURE 3: Gln-291 → His mutation of the pGS110-encoded E2p chain. The 1.68-kbp *KpnI–SphI* (Kp–Sp) fragment of pGS110 is represented by a solid line. It encodes a C-terminal portion of E1p and an N-terminal portion of E2p-110 containing the lipoyl domain (stippled bar), the (alanine + proline)-rich sequence (solid bar), and the E3-binding domain (open bar). The nucleotide sequence and the 32 residues of the (alanine + proline)-rich sequence are numbered according to Stephens et al. (1983a,b). The 17-mer mutagenic primer S5 was used to direct the mutation of the Gln-291 codon (CAG) to CAC, the preferred codon for histidine. Synthetic primers, used for resequencing the mutant *KpnI–SphI* fragment, are shown as horizontal arrows; a rightward orientation indicates priming on the *KpnI–SphI* fragment when cloned in M13mp19 and leftward when in M13mp18.

To determine the pK_a of the histidine residue (His-291) in the pGS178-encoded complex, the p^2H of the sample (77 mg/mL in the ^2H_2O buffer described above) was adjusted by small additions of either 2.7 M ^2HCl or 0.75 M NaO^2H, and spectra were recorded as already described. The p^2H of the solution was measured directly in the NMR tube by using a 5 mm diameter electrode, and the values recorded were uncorrected meter readings.

Preparation of the Tryptic Core of the pGS178-Encoded Complex for ^1H NMR Studies. The pGS178-encoded complex (20 mg, 26 mg/mL in 20 mM sodium phosphate buffer, pH 7.0, containing 2.7 mM EDTA and 0.02% NaN_3) was incubated with trypsin (0.1% w/w) at 0 °C. When the catalytic activity of the complex had fallen to less than 5% of its original value, the digestion was arrested by adding a tenfold excess (w/w over trypsin) of soyabean trypsin inhibitor. The tryptic core of the complex was then purified by gel filtration on a Superose 12 column in 50 mM sodium phosphate buffer, pH 7.0, containing 2.7 mM EDTA and 0.02% NaN_3. The protein that emerged at the void volume of the column was collected, diluted twofold with 2.7 mM EDTA, and concentrated by ultracentrifugation (122000*g* for 3 h). The protein pellet was resuspended in the 20 mM ^2H_2O buffer described above and dialyzed exhaustively against the same buffer.

RESULTS

Gln-291 → His Mutation in the E2p Chain of the pGS110-Encoded Complex. In order to provide an amino acid residue in the long (alanine + proline)-rich sequence of the pGS110-encoded E2p chain whose side chain would give rise to signals that might be distinguished in the aromatic region of the ^1H NMR spectrum of the intact complex, it was decided to substitute histidine for glutamine at position 291 (Figure 1). Glutamine-291 was selected because it could be changed to an aromatic yet polar amino acid, histidine, by a single point mutation in the *aceF* gene.

The PDH complex is expressed from two transcripts, one encoding the E1p and E2p components (*aceEF*; −4600 nucleotides) and the other encoding all three components (*aceEF-lpd*; −6400 nucleotides) (Spencer & Guest, 1985). This regulatory organization of the PDH complex genes necessitates the manipulation of the entire 6.4-kbp *aceEF-lpd* operon in protein engineering studies. However, because the operon is too large for stable cloning in bacteriophage M13, a cassette-replacement approach was adopted for mutagenesis of the PDH complex (Figure 2). The unique 1.68-kbp *Kpn*I–*Sph*I fragment from pGS110 was cloned into M13mp19 to produce a recombinant phage (mp19KS1) containing the 3′ end of the *aceE* (E1p) gene and a 5′ portion of the *aceF′* (E2p-110) gene encoding the lipoyl domain and long (alanine + proline)-rich segment (Stephens et al., 1983a,b; Figures 2 and 3). This derivative is therefore ideally suited for use as the template in mutagenesis in experiments designed to investigate the (alanine + proline)-rich sequence.

The 17-mer oligonucleotide GCGAAACAC*GAAGCGGC (S5) is complementary to positions 4651–4667 in the *aceF′* gene cloned in mp19KS1 (Figure 3), except for a single C–C mismatch (*) at position 4659. It directs the mutation of the Gln-291 codon (CAG) to CAC, the preferred codon for histidine in *E. coli* (Grosjean & Fiers, 1982). A computer search of the mp19KS1 sequence using SEQFIT (Staden, 1977) revealed no potential hybridization sites except the one of interest.

Following mutagenesis with the S5 and universal primers, some 36 derivatives of mp19KS1 were tested by the dot–hybridization procedure. All gave positive signals at 20 °C, but only three gave signals after washing at their estimated T_m (Wallace et al., 1981) of 58 °C. The putative mutants were reprobed after plaque purification, and all were found to contain the desired mutation when partially sequenced by using the 16-mer primer (S4; Figure 3).

The mutant 1.68-kbp *Kpn*I–*Sph*I fragment was transferred to the gel-purified 11.5-kbp receptor fragment of pGS110 (Figure 2). Because there is no phenotypic selection for the mutation, the restructured plasmids were probed with the mutagenic primer (S5): those containing the desired mutation hybridized at 58 °C whereas pGS110 did not. A representative mutant plasmid, designated pGS178, was characterized by restriction analysis with *Hin*dIII, *Kpn*I, *Nru*I, and *Sph*I and by sequencing the entire 1.68-kbp fragment after recloning in M13mp18 (Figure 3). The restriction patterns were indistinguishable from those of pGS110, and no mutations other than the desired G → C change at position 4659 were found. It was concluded that pGS178 contains an *ace-lpd* operon encoding an E2p-110 chain in which Gln-291 is replaced by a histidine residue.

Functional Characterization of the Mutation. The gross functional consequences of the Gln-291 → His mutation were investigated by transforming an *aceEF-lpd* deletion strain (JRG1342) with pGS178 and testing the transformants on a variety of selective media (Guest et al., 1985). These showed that pGS178 is indistinguishable from pGS110 in its ability to complement the *ace* and *lpd* lesions and indicated that pGS178 expresses a functional PDH complex.

Characterization of the pGS178-Encoded Complex. The pGS178-encoded complex was purified as described under Materials and Methods. When examined by SDS–polyacrylamide gel electrophoresis (Figure 4a), the complex appeared identical with the pGS110-encoded complex as would be expected if the two complexes differ by only a single amino acid residue in E2p. The specific catalytic activity of the pGS178-encoded complex was 29 μmol/(min·mg), which is similar to that [23–30 μmol/(min·mg)] of the pGS110-encoded complex (Guest et al., 1985).

Substitution of His-291 into E2p did not prevent the aggregation of the E1p, E2p, and E3 subunits into a large macromolecular structure, as evidenced by the emergence of the complex at the usual void volume during gel filtration on Sepharose 6B. This is consistent with the fact that the binding sites for E1p, E2p, and E3 are located in the C-terminal half of E2p (between residues 317 and 619) (Packman et al., 1984). The polypeptide chain ratios (E1p:E2p:E3) of the pGS110- and pGS178-encoded complexes were found to be 0.8:1:0.9 and 0.9:1:1.2, respectively.

The intramolecular coupling of active sites in the pGS110- and pGS178-encoded complexes was also assayed (data not shown). The acetylation curves of the two complexes were found to be indistinguishable, and since the departure of this curve from linearity measures the extent of active site coupling in a complex (Packman et al., 1983), it is evident that in both complexes probably all 24 E2p chains can participate in transacetylation reactions. The substitution of His-291 into the E2p chain evidently does not disrupt the assembly, the catalytic activity, or the system of the active site coupling in the enzyme complex.

Limited Proteolysis. The E2p chain in the PDH complex is highly susceptible to proteolysis and is rapidly cleaved with *S. aureus* V8 proteinase and with trypsin, the cleavage sites for which have been mapped (Packman et al., 1984; Packman & Perham, 1987). Since many of these sites (Lys-301, Lys-312, and Lys-316 for trypsin; Glu-292 and Glu-318 for

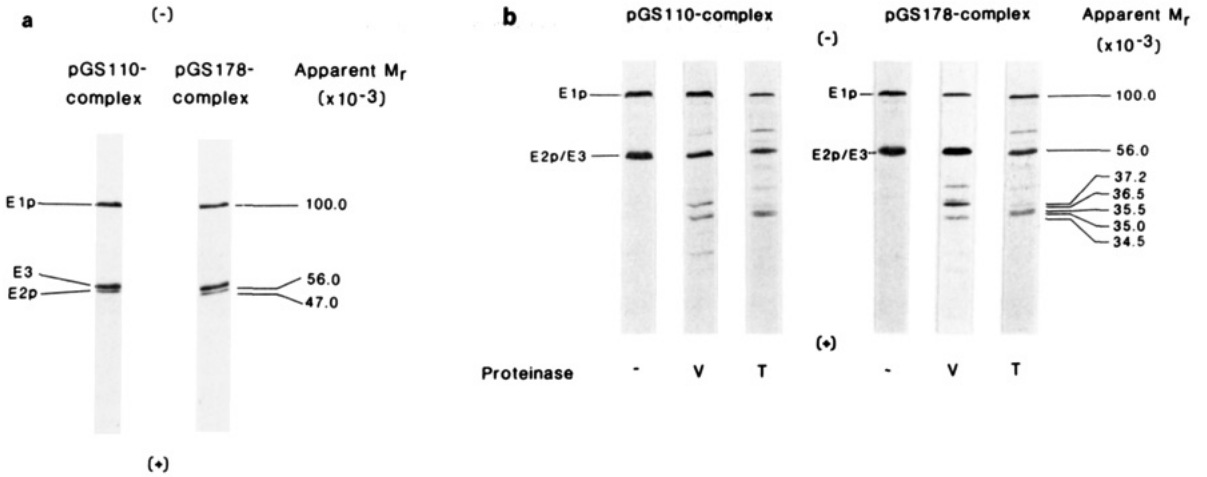

FIGURE 4: SDS–polyacrylamide gel electrophoresis of the pGS110- and pGS178-encoded PDH complexes and the products of their limited proteolysis with trypsin and *S. aureus* V8 proteinase. (a) Examination of the pGS110- and pGS178-encoded PDH complexes by SDS–polyacrylamide gel electrophoresis (20% T, 0.3% C). The gel was stained with Coomassie Blue. (b) Each complex was treated with trypsin (T) and *S. aureus* V8 proteinase (V), as described under Materials and Methods. The products of digestion were examined by SDS–polyacrylamide gel electrophoresis (27% T, 0.3% C). The gel was stained with Coomassie Blue. In this gel system, the E2p and E3 bands are not resolved. The fragment of M_r ca. 45 000 present in the digest of the pGS178-encoded complex with *S. aureus* V8 proteinase probably arises from the E1p subunit.

S. aureus V8 proteinase) are clustered between residues 292 and 316 of E2p (wild-type numbering), the effect of substituting a histidine at residue 291 in E2p was examined.

Tryptic digestion of the pGS110- and pGS178-encoded complexes, as described under Materials and Methods, proceeded at the same rate. Examination of the products of digestion by SDS–polyacrylamide gel electrophoresis (Figure 4b) revealed that cleavage of the pGS178-E2p chain had occurred at all the expected sites, releasing residual fragments of E2p of the appropriate size (36.5, 35.5, and 35.0 kDa for cleavage at Lys-301, Lys-312, and Lys-316, respectively). As expected, the lipoyl domains released from the pGS110- and pGS178-E2p chains did not differ in size (data not shown).

Similarly, cleavage of the two complexes with *S. aureus* V8 proteinase occurred at similar rates (even though the histidine residue lies immediately adjacent to the cleavage site at Glu-292). Cleavage took place at both Glu-292 and Glu-318, as judged by the release of fragments of molecular mass 37.2 and 34.5 kDa (Figure 4b), respectively, and of lipoyl domains of the same apparent size (data not shown).

^{1}H NMR Spectroscopy of the pGS178-Encoded Complex. The aromatic region (ca. 6–9 ppm) of the ^{1}H NMR spectra of the pGS110- and pGS178-encoded complexes are shown in Figure 5 (panel A). Whereas the wild-type complex (Perham et al., 1981) and the pGS110-encoded complex display no sharp signals in this region of the spectrum, the spectrum of the pGS178-encoded complex contains two new sharp peaks (at 7.94 and 7.05 ppm). It is tempting to suppose that these signals arise from the histidine residue that has been engineered into the (alanine + proline)-rich region of the pGS110-encoded E2p chain. The chemical shifts of these two peaks correspond closely to the shifts expected for the C2 and C4 protons (8.12 and 7.14 ppm, respectively) of histidine residues in small peptides (Jardetzky & Roberts, 1981), and attempts to measure the intensity of these peaks (using spin-echo spectra) gave values consistent with ca. 1 mobile histidine residue per E2p chain for each peak.

To confirm that His-291 in the E2p chain is the source of the new sharp signals, the pGS178-encoded complex was submitted to limited proteolysis with trypsin at 0 °C and pH 7.0 (as described under Materials and Methods). The ^{1}H NMR spectrum of this tryptic core complex (Figure 5, panel A) was found to lack both the sharp peaks at 7.94 and 7.05 ppm. This is consistent with the expected tryptic cleavage of the E2p chain at Lys-316, releasing the N-terminal region (including His-291) and leaving a residual tryptic core complex (Packman et al., 1984).

^{1}H NMR studies have often been used to investigate the pK_a values of ionizable residues in proteins (Jardetzky & Roberts, 1981). Histidine residues are particularly amenable to this kind of study since the C2 and C4 protons resonate in the aromatic region of the spectrum and are generally well resolved and more easily distinguished. Spectra of the pGS178-encoded complex were recorded at various p^{2}H values, and the chemical shifts of the supposed C2 and C4 protons of His-291 were measured relative to internal TSP (data not shown). As expected, both peaks shifted upfield when the p^{2}H was raised from 7.0 to ca. 10.0. Similarly, lowering the p^{2}H resulted in a downfield shift of both peaks. A full titration curve and precise pK_a value could not be determined since the enzyme complex precipitated from solution below p^{2}H 5.7. From the curve obtained, however, it is probable that the pK_a value of the histidine residue is ca. 6.4 and lies in the "normal" range, for which values of between 5 and 8 have been reported (Jardetzky & Roberts, 1981).

Aliphatic Region of the ^{1}H NMR Spectrum of the pGS178-Encoded Complex. The aliphatic regions (0–6 ppm) of the ^{1}H NMR spectra of the pGS110- and pGS178-encoded complexes are shown in Figure 5 (panel B). The two spectra are extremely similar, consisting of a broad envelope of overlapping resonances, upon which are superimposed many sharp peaks. For the wild-type and pGS110-encoded complexes, the dominant sharp peak at 1.39 ppm has been assigned to the methyl side chains of alanine residues in the long (alanine + proline)-rich sequence(s) in the E2p chains (Radford et al., 1986, 1987). Closer inspection of the two spectra, however, revealed that the intensity of this peak for the pGS178-encoded complex was significantly smaller than that of the same peak for the pGS110-encoded complex. This was surprising since the number of alanine residues in the long (alanine + proline)-rich regions in the two complexes is identical (Figure 1). By use of spin-echo spectra (which, with a total delay time of 1 ms, reduces the intensity of the broad envelope but leaves the intensity of the sharp peaks virtually undiminished), an attempt was made to compare the intensity of the peaks at 1.39 ppm for the two complexes. For the pGS110- and pGS178-encoded complexes, the intensity of

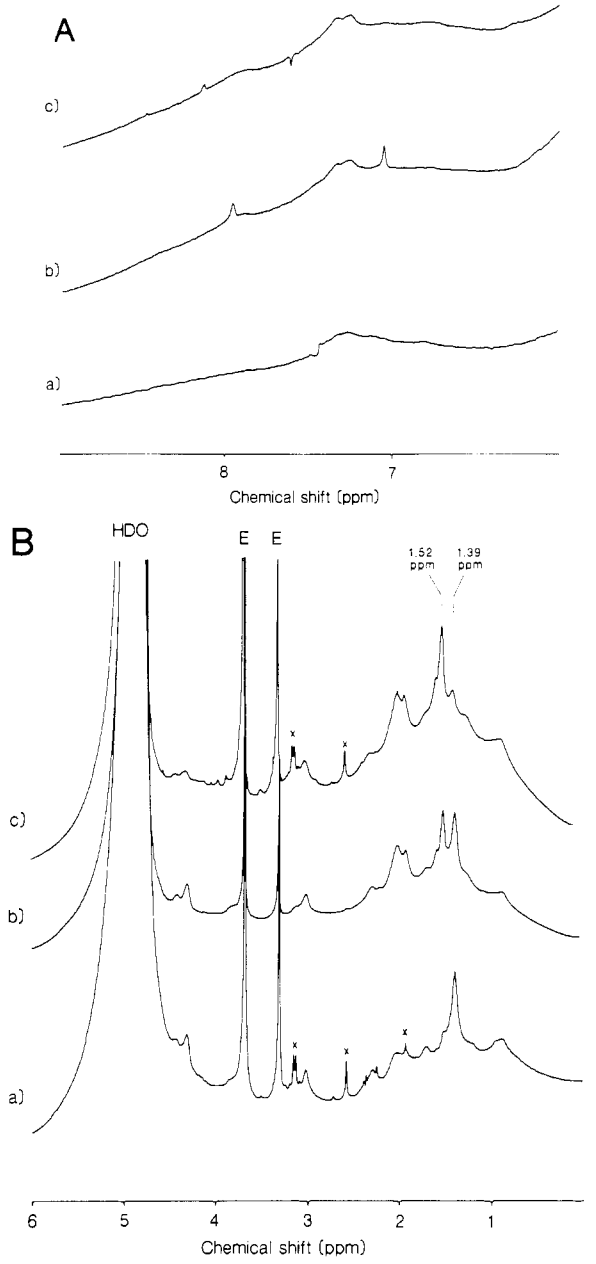

FIGURE 5: 400-MHz ^{1}H NMR spectra of the pGS110- and pGS178-encoded complexes. (Panel A) Aromatic region (ca. 6–9 ppm) of the spectra: (a) pGS110-encoded complex; (b) pGS178-encoded complex; (c) pGS178-encoded tryptic core complex (36-kDa E2p). (Panel B) Aliphatic region (ca. 0–6 ppm) of the spectra: (a) pGS110-encoded complex; (b) pGS178-encoded complex; (c) pGS178-encoded tryptic core (36-kDa E2p). The very sharp resonances marked E and X arise from EDTA and nonprotein contaminants, respectively. The vertical gain of each spectrum has been adjusted according to the protein concentration used, so that the spectra are more directly comparable.

these resonances corresponds to the equivalent of about 16 and 7 mobile alanine residues per E2p chain, respectively. The number of alanine residues in potentially flexible regions of these E2p chains is assumed to be 20, the number of such residues in the long (alanine + proline)-rich segment of E2p, as designated in Figure 1, although the precise length of the flexible region is not yet known (Packman et al., 1984). Given this and the unavoidable errors in calculating the areas of the sharp peaks that derive from the uncertainty in drawing an appropriate base line in the ^{1}H NMR spectra, the correlation between the observed and potential number of flexible alanine residues in the pGS110 complex is good. However, the correlation is significantly poorer for the pGS178-encoded complex. Substitution of a histidine residue at position 291 in the E2p chain of the pGS110-encoded complex appears to impair the conformational flexibility of about half of the alanine residues generating the sharp signal at 1.39 ppm.

The spectrum of the pGS178-encoded complex also clearly displayed another sharp peak (at 1.52 ppm) which was present, but only poorly resolved, in the spectrum of the pGS110-encoded complex (Figure 5, panel B). A peak of similar size resonating at 1.52 ppm is also clearly visible in the 400-MHz ^{1}H NMR spectrum of the pGS156-encoded complex (Radford et al., 1987). One strong possibility is that this peak arises from the short (alanine + proline)-rich region of the E2p chain which separates the E3-binding domain from the C-terminal catalytic domain (Figure 1), about 5 alanine residues per E2p chain being required to account for its size (Radford et al., 1987). To test whether this possibility might also apply to the pGS178-encoded complex, the ^{1}H NMR spectrum of the tryptic core of the complex, generated by cleavage of the E2p chain at Lys-316 (Figure 1), was examined. The peak at 1.39 ppm, assigned to alanine residues in the long (alanine + proline)-rich region of the E2p chain (Radford et al., 1986, 1987), was absent from the spectrum, but that at 1.52 ppm was indeed retained and was undiminished in intensity (Figure 5).

Discussion

The very large size of the *E. coli* PDH complex (M_r ca. 4.5 $\times$ 10^6) and the correspondingly long rotational correlation time of ca. 10^{-5} s (Ambrose & Perham, 1976) suggest that the ^{1}H NMR spectrum of the complex should contain only very broad signals with line widths of the order of 10 kHz. The existence of large sharp peaks in the spectrum is therefore surprising and has been attributed to the presence of conformationally flexible segments within the E2p polypeptide chains (Perham et al., 1981; Roberts et al., 1983). The size and chemical shift (1.39 ppm) of the major sharp peak (Perham et al., 1981) points to the three long sequences rich in alanine, proline, and charged amino acid residues in the E2p chain as being those primarily responsible (Spencer et al., 1984; Packman et al., 1984). This view is strongly supported by ^{1}H NMR spectroscopy of a synthetic peptide with the same sequence as the innermost of the long (alanine + proline)-rich regions of E2p (Radford et al., 1986) and of PDH complexes in which specific deletions of lipoyl domains and the (alanine + proline)-rich sequences had been engineered (Radford et al., 1987).

To test this hypothesis directly, we engineered the replacement of Gln-291 with a histidine residue in the single long (alanine + proline)-rich segment of the E2p chain of the pGS110-encoded PDH complex (Figures 1 and 2). The resultant complex, encoded by plasmid pGS178, was found to be unimpaired in its assembly, as indicated by its content of E1p, E2p, and E3 polypeptide chains (Figure 4), or in its specific catalytic activity and active site coupling. When the complex was examined by ^{1}H NMR spectroscopy (Figure 5), two sharp signals were observed in the aromatic region of the spectrum, with chemical shifts (7.94 and 7.05 ppm) very close to those normally found for the C2 and C4 protons (8.12 and 7.14 ppm, respectively) of histidine residues in small peptides (Jardetzky & Roberts, 1981). The identification of the new histidine residue in E2p as the source of these two peaks was supported by their intensity, which was consistent with one such histidine residue per E2p chain, and by the pH dependence of their chemical shifts. Although a full titration curve could not be obtained because the enzyme complex precipitated at p^{2}H values below 5.7, the peaks appeared to be arising from a side chain with a pK_a of ca. 6.4, which lies in the range

normally expected for histidine residues in peptides (Jardetzky & Roberts, 1981).

There are 23 histidine residues in E1p, 5 in E2p, and 13 in E3 (Stephens et al., 1983a–c), giving a total of about 40 in the protomer of the pGS110-encoded PDH complex (a protomer is defined as a $^1/_{24}$th part of the complex, since the complex is assembled around an octahedral core of 24 E2p chains). None of these produces a signal visible in the aromatic region of the 400-MHz ^{1}H NMR spectrum (Figure 5), presumably because each resides within structured and relatively immobile regions of the various polypeptide chains in this very large enzyme complex. Detection of the signal from the single histidine residue inserted in the long (alanine + proline)-rich segment of each E2p chain of the pGS178-encoded complex provides unequivocal evidence that this residue must be inhabiting a region of polypeptide chain that enjoys significant conformational flexibility with respect to the bulk of the enzyme complex.

Comparison of the aliphatic regions of the ^{1}H NMR spectra of the pGS110- and pGS178-encoded complexes, however, revealed some small but significant differences (Figure 5). The peak at 1.39 ppm, which arises from the methyl side chains of alanine residues in the long (alanine + proline)-rich sequences in E2p (Radford et al., 1986, 1987), was significantly reduced in intensity for the pGS178-encoded complex compared with the pGS110-encoded complex, and the peak at 1.52 ppm was in consequence much better resolved. Since the two complexes differ only by a single amino acid replacement (Gln → His) in the E2p chain, the number of alanine residues in the long potentially flexible region should be the same. The peak arising from these alanine residues (at 1.39 ppm) is unexpectedly small in the pGS178-encoded complex, suggesting that the presence of the new histidine residue has significantly diminished the flexibility of some at least of the alanine residues in this segment of polypeptide chain.

This effect of the histidine residue is not clearly understood. An ionic interaction of the positively charged histidine side chain with a glutamic or aspartic acid residue elsewhere in the enzyme complex is unlikely, since the intensity of the peak at 1.39 ppm did not alter when the p^{2}H of the buffer was changed from ca. 5.7 to 10.0. Nor did the histidine residue affect the susceptibility of the E2p chain to attack by trypsin or *S. aureus* V8 proteinase at nearby sites (Figure 4). An NMR study of a synthetic peptide with the same amino acid sequence as that of the long (alanine + proline)-rich region of the pGS110-encoded E2p polypeptide chain has revealed that this sequence confers some curious structural properties on the peptide in free solution, e.g., the all-trans configuration of its Ala–Pro peptide bonds (Radford et al., 1986). Similar studies of this synthetic peptide with the appropriate Gln → His change may throw light on the effect of the histidine residue, if any, on the conformation of the peptide.

The small sharp peak at 1.52 ppm in the ^{1}H NMR spectra of the wild-type and pGS110- and pGS156-encoded complexes (Radford et al., 1987) and the pGS178-encoded complex (Figure 5) may be arising from the shorter (alanine + proline)-rich sequence in the E2p chain (residues 370–377; Figure 1) that links the E3-binding domain to the large C-terminal catalytic domain (Radford et al., 1987). The better resolution of the peak at 1.52 ppm in the 400-MHz ^{1}H NMR spectra of the pGS156-encoded (Radford et al., 1987) and pGS178-encoded (Figure 5) complexes, compared with the spectra of the wild-type and pGS110-encoded complexes (Radford et al., 1987), is probably due to the smaller size of the adjacent peak at 1.39 ppm in the former two complexes. Conformational flexibility in the region of residues 370–377 of the E2p chain might explain fluorescence data which suggest that the E3 component bound to the PDH complex is quite mobile (Grande et al., 1980). Substitution of the histidine residue at position 291 had no apparent effect on this peak.

The detection of a single introduced histidine residue in the ^{1}H NMR spectrum of an enzyme complex of this large size is a striking result and provides unequivocal evidence for conformational flexibility in the target region of the E2p polypeptide chain. The use of such a marker to probe for conformational flexibility in large proteins is a novel application of site-directed mutagenesis, which can be adapted to study conformational mobility in many other protein systems.

Registry No. S5, 111524-37-9; E1p, 9014-20-4; E2p, 9032-29-5; L-Gln, 56-85-9; L-His, 71-00-1.

REFERENCES

Ambrose, M. C., & Perham, R. N. (1976) *Biochem. J. 155*, 429–432.

Atkinson, T., & Smith, M. (1984) in *Oligonucleotide Synthesis. A Practical Approach* (Gait, M., Ed.) pp 35–81, IRL Press, Oxford and Washington, DC.

Berman, J. N., Chen, G.-X., Hale, G., & Perham, R. N. (1981) *Biochem. J. 199*, 513–520.

Biggin, M. D., Gibson, T. J., & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3963–3965.

Birnboim, H. C., & Doly, J. (1979) *Nucleic Acids Res. 7*, 1513–1524.

Bleile, D. M., Munk, P., Oliver, R. M., & Reed, L. J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4385–4389.

Carter, P., Bedouelle, H., Waye, M. M. Y., & Winter, G. (1985) *Oligonucleotide Site-Directed Mutagenesis in M13*, p 18, Anglian Biotechnology Ltd., Colchester.

Danson, M. J., Hale, G., Johnson, P., Perham, R. N., Smith, J., & Spragg, P. (1979) *J. Mol. Biol. 129*, 603–617.

Grande, H. J., Visser, A. J. W. G., & Veeger, C. (1980) *Eur. J. Biochem. 106*, 361–369.

Grosjean, H., & Fiers, W. (1982) *Gene 18*, 199–209.

Guest, J. R., Roberts, R. E., & Stephens, P. E. (1983) *J. Gen. Microbiol. 129*, 671–680.

Guest, J. R., Darlison, M. G., Spencer, M. E., & Stephens, P. E. (1984) *Biochem. Soc. Trans. 12*, 220–223.

Guest, J. R., Lewis, H. M., Graham, L. D., Packman, L. C., & Perham, R. N. (1985) *J. Mol. Biol. 183*, 743–756.

Gutowski, J. A., & Lienhard, G. E. (1976) *J. Biol. Chem. 251*, 2863–2866.

Hale, G., & Perham, R. N. (1979a) *FEBS Lett. 105*, 263–266.

Hale, G., & Perham, R. N. (1979b) *Biochem. J. 177*, 129–137.

Hong, G. F. (1981) *Biosci. Rep. 1*, 243–252.

Jardetzky, O., & Roberts, G. C. K. (1981) *NMR in Molecular Biology*, Academic, New York.

Kramer, B., Kramer, W., & Fritz, H.-J. (1984) *Cell (Cambridge, Mass.) 38*, 879–887.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Messing, J. (1979) *Recomb. DNA Tech. Bull. 2*, 43–48.

Miles, J. S., & Guest, J. R. (1987) *Biochem. J. 245*, 869–874.

Miles, J. S., Guest, J. R., Radford, S. E., & Perham, R. N. (1987) *Biochim. Biophys. Acta 913*, 117–121.

Norrander, J., Kempe, T., & Messing, J. (1983) *Gene 26*, 101–106.

Packman, L. C., & Perham, R. N. (1986) *FEBS Lett. 139*, 155–158.

Packman, L. C., & Perham, R. N. (1987) *Biochem. J. 242*, 531–538.

Packman, L. C., Stanley, C. J., & Perham, R. N. (1983) *Biochem. J. 213*, 331-338.
Packman, L. C., Hale, G., & Perham, R. N. (1984) *EMBO J. 3*, 1315-1319.
Perham, R. N. (1983) in *Frontiers in Biochemical and Biophysical Studies of Proteins and Membranes* (Liu, T.-Y., Sakakibara, S., Schechter, A. N., Yagi, K., & Yasunobu, K., Eds.) pp 249-258, Elsevier, New York.
Perham, R. N., Duckworth, H. W., & Roberts, G. C. K. (1981) *Nature (London) 292*, 474-477.
Radford, S. E., Laue, E. D., & Perham, R. N. (1986) *Biochem. Soc. Trans. 14*, 1231-1232.
Radford, S. E., Laue, E. D., Perham, R. N., Miles, J. S., & Guest, J. R. (1987) *Biochem. J. 247*, 641-649.
Reed, L. J. (1974) *Acc. Chem. Res. 7*, 40-46.
Roberts, G. C. K., Duckworth, H. W., Packman, L. C., & Perham, R. N. (1983) *Ciba Found. Symp. 93*, 47-62.
Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H., & Roe, B. A. (1980) *J. Mol. Biol. 143*, 161-178.
Spencer, M. E., & Guest, J. R. (1985) *Mol. Gen. Genet. 200*, 145-154.
Spencer, M. E., Darlison, M. G., Stephens, P. E., Duckenfield, I. K., & Guest, J. R. (1984) *Eur. J. Biochem. 141*, 361-374.
Sproat, B. S., & Gait, M. J. (1984) in *Oligonucleotide Synthesis. A Practical Approach* (Gait, M., Ed.) pp 83-115, IRL Press, Oxford and Washington, DC.
Staden, R. (1977) *Nucleic Acids Res. 4*, 4037-4051.
Stephens, P. E., Darlison, M. G., Lewis, H. M., & Guest, J. R. (1983a) *Eur. J. Biochem. 133*, 155-162.
Stephens, P. E., Darlison, M. G., Lewis, H. M., & Guest, J. R. (1983b) *Eur. J. Biochem. 133*, 481-489.
Stephens, P. E., Lewis, H. M., Darlison, M. G., & Guest, J. R. (1983c) *Eur. J. Biochem. 135*, 519-527.
Stepp, L. R., Bleile, D. M., McRorie, D. K., Pettit, F. H., & Reed, L. J. (1981) *Biochemistry 20*, 4555-4560.
Wallace, R. B., Johnson, M. J., Hirose, T., Miyake, T., Kiwashima, E. H., & Itakura, K. (1981) *Nucleic Acids Res. 9*, 879-894.
Zoller, M. J., & Smith, M. (1983) *Methods Enzymol. 100*, 468-500.
Zoller, M. J., & Smith, M. (1984) *DNA 3*, 479-488.

Chemical Modification of the Cu_A Site Affects the Proton Pumping Activity of Cytochrome *c* Oxidase[†,‡]

Thomas Nilsson,[§] Jeff Gelles,[∥] Peter Mark Li,[⊥] and Sunney I. Chan*
Arthur Amos Noyes Laboratory of Chemical Physics, California Institute of Technology, Pasadena, California 91125
Received March 26, 1987; Revised Manuscript Received August 28, 1987

ABSTRACT: Cytochrome *c* oxidase in which the Cu_A site has been perturbed by extensive modification of the enzyme with the thiol reagent *p*-(hydroxymercuri)benzoate has been reconstituted into phospholipid vesicles. The reconstituted vesicles lack respiratory control, and the orientation of the enzyme in the vesicles is similar to that of the native cytochrome *c* oxidase. In the proton translocation assay, the vesicles containing the modified enzyme behave as if they are unusually permeable to protons. When the modified and native proteins were coreconstituted, a substantial portion of the latter became uncoupled as revealed by low respiratory control and low overall proton pumping activity. These results suggest that the modified enzyme catalyzes a passive transport of protons across the membrane. When milder conditions were used for the chemical modification, a majority of the thiols reacted while the Cu_A site remained largely intact. Reconstitution of such a partially modified cytochrome *c* oxidase produced vesicles with respiratory control and proton translocating activity close to those of reconstituted native enzyme. It thus appears that the appearance of a proton leak is related to the perturbation of the Cu_A site. These observations suggest that the structure of Cu_A may be related to the role of this site in the proton pumping machinery of cytochrome *c* oxidase.

Cytochrome *c* oxidase fills a central function in the energy transduction of aerobic organisms. In eukaryotes, the enzyme, which spans the mitochondrial membrane, catalyzes the reduction of molecular oxygen to water. The reductant, cytochrome *c*, donates the electrons from the cytosolic side of the membrane while the protons consumed in the reaction are taken up from the mitochondrial matrix. In addition, one proton is actively transported from the matrix to the cytosol for each electron transferred (Wikström, 1977; Casey et al., 1979a,b). The reaction is thus electrogenic and contributes to the electrochemical potential gradient across the inner mitochondrial membrane. It has been suggested (Wikström et al., 1981) that the proton translocating activity is necessary for the full utilization of the redox span between cytochrome *c* and molecular oxygen. While a great ideal is known about the oxygen reduction in cytochrome *c* oxidase [see Malmström (1982), Naqui et al. (1986), and Wikström et al. (1981) for reviews; Blair et al., 1985], the molecular mechanism of proton translocation remains largely unknown.

The enzyme contains four redox-active metal centers, two irons in cytochromes *a* and a_3 and two copper ions which are

[†] Contribution No. 7574 from the Arthur Amos Noyes Laboratory of Chemical Physics. Supported by Grant GM 22432 from the National Institute of General Medical Sciences, U.S. Public Health Service.
[‡] Dedicated to the memory of Dr. Robert P. Casey.
* Correspondence should be addressed to this author.
[§] Recipient of a postdoctoral fellowship from the Swedish Natural Science Research Council.
[∥] Present address: Department of Cell Biology and Physiology, Washington University Medical School, 660 S. Euclid, St. Louis, MO 63110.
[⊥] National Science Foundation Predoctoral Fellow.

0006-2960/88/0427-0296$01.50/0