Relationship between Mitochondrial NADH–Ubiquinone Reductase and a Bacterial NAD-Reducing Hydrogenase[†,‡]

Stephanie J. Pilkington, J. Mark Skehel, Robert B. Gennis,[§] and John E. Walker*
The Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K.
Received August 15, 1990; Revised Manuscript Received October 31, 1990

ABSTRACT: Bovine mitochondrial NADH–ubiquinone reductase (complex I), the first enzyme in the electron-transport chain, is a membrane-bound assembly of more than 30 different proteins, and the flavoprotein (FP) fraction, a water-soluble assembly of the 51-, 24-, and 10-kDa subunits, retains some of the catalytic properties of the enzyme. The 51-kDa subunit binds the substrate NAD(H) and probably contains both the cofactor, FMN, and also a tetranuclear iron–sulfur center, while a binuclear iron–sulfur center is located in the 24- or 10-kDa proteins. The 75-kDa subunit is the largest of the six proteins in the iron–sulfur protein (IP) fraction, and its sequence indicates that it too contains iron–sulfur clusters. Partial protein sequences have been determined at the N-terminus and at internal sites in the 51-kDa subunit, and the corresponding cDNA encoding a precursor of the protein has been isolated by using a novel strategy based on the polymerase chain reaction. The mature protein is 444 amino acids long. Its sequence, and those of the 24- and 75-kDa subunits, shows that mitochondrial complex I is related to a soluble NAD-reducing hydrogenase from the facultative chemolithotroph *Alcaligenes eutrophus* H16. This enzyme has four subunits, α, β, γ, and δ, and the $\alpha\gamma$ dimer is an NADH oxidoreductase that contains FMN. The γ-subunit is related to residues 1–240 of the 75-kDa subunit of complex I, and the α-subunit sequence is a fusion of homologues of the 24- and 51-kDa subunits, in the order N- to C-terminal. The most highly conserved regions are in the 51-kDa subunit and probably form parts of nucleotide binding sites for NAD(H) and FMN. Another conserved region surrounds the sequence motif CysXXCysXXCys, which is likely to provide three of the four ligands of a 4Fe–4S center, possibly that known as N-3. Characteristic ligands for a second 4Fe–4S center are conserved in the 75-kDa and γ-subunits. This relationship with the bacterial enzyme implies that the 24- and 51-kDa subunits, together with part of the 75-kDa subunit, constitute a structural unit in mitochondrial complex I that is concerned with the first steps of electron transport.

NADH–ubiquinone reductase is the first enzyme in the mitochondrial electron-transport chain and catalyzes the transfer of two electrons from NADH to ubiquinone-10 with concomitant translocation of four protons across the membrane (Wikström, 1984). Of the 30 or more subunits that can be distinguished in the bovine heart enzyme (J.M.S. and J.E.W., unpublished results), only 7 are encoded in the mitochondrial genome (Chomyn et al., 1985, 1986). The complex also contains a variety of prosthetic groups including at least seven iron–sulfur (Fe–S) centers and FMN [see Ragan (1987) for a review]. An understanding of the mechanism of this complicated enzyme would be helped by a precise knowledge of the arrangement of subunits and the disposition of prosthetic groups. Information about the organization of subunits and the sites of some prosthetic groups has been obtained by splitting the complex with chaotropes into the FP (flavoprotein), IP (iron–sulfur protein), and HP (hydrophobic protein) fractions (Hatefi & Stempel, 1969; Galante & Hatefi, 1979. The FP subcomplex is composed of the 51-, 24-, and 10-kDa subunits, and a tetranuclear and a binuclear Fe–S center, respectively, are in the 51-kDa subunit and in either the 24- or the 10-kDa subunit (Ohnishi et al., 1985). It can oxidize NADH by using a variety of electron acceptors, with similar kinetic properties to the intact enzyme complex (Dooijewaard & Slater, 1976a,b). The NADH binding site is in the 51-kDa subunit (Chen & Guillory, 1981; Deng et al., 1990), and, since FMN is likely to accept electrons directly from NADH, its binding site is probably in the same protein (Ingledew & Ohnishi, 1980; Krishnamoorthy & Hinkle, 1988). The IP fraction contains the 75-, 49-, 30-, 18-, 15-, and 13-kDa subunits and at least three Fe–S clusters, and sequence studies have provided evidence for a 4Fe–4S cluster in the 75-kDa subunit (Runswick et al., 1989). The HP fraction is the residuum after removal of FP and IP fractions and contains at least two further Fe–S centers.

In order to gain further knowledge of the binding sites for these various ligands, we have determined the sequences of a number of the subunits of the enzyme, including those of the bovine 75- and 24-kDa subunits (Runswick et al., 1989; Pilkington & Walker, 1989), and that of the 51-kDa subunit is described below. The sequences show that these three subunits are closely related to components of a soluble NAD-dependent hydrogenase from the bacterium *Alcaligenes eutrophus*, a facultative chemolithotroph that can assimilate CO_2 and use H_2 as an energy source. The soluble hydrogenase contains four different subunits, α, β, γ, and δ, encoded in the *hoxS* locus in the order *hoxF* (α), *hoxU* (γ), *hoxY* (δ), and *hoxH* (β) (Tran-Betcke et al., 1990). The $\alpha\gamma$ dimer is an NADH oxidoreductase and contains one molecule of FMN; the other two subunits are concerned with the hydrogenase activity. The sequence of amino acids 1–188 of HoxF is homologous to the 24-kDa subunit of complex I, and the remainder is closely related to the 51-kDa subunit. HoxU is related to amino acids 1–240 of the 75-kDa subunit. These relationships imply that the 51- and 24-kDa subunits, together with the N-terminal part of the 75-kDa subunit of complex

[†] S.J.P. was supported by an MRC research studentship. R.B.G. was a Fullbright Scholar during this work and was supported, in part, as a Fellow of the John Simon Guggenheim Foundation.
[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J05316.
* To whom correspondence should be addressed.
[§] Present address: School of Chemical Sciences, University of Illinois at Urbana-Champaign, Urbana, IL 61801.

0006-2960/91/0430-2166$02.50/0

I, form a structural and functional unit similar to the $\alpha\gamma$ dimer of the hydrogenase. The most highly conserved sequences are in the 51-kDa subunit. They probably form parts of the binding sites for NADH and FMN, and provide ligands for a 4Fe–4S center. At least one other Fe–S center is associated with the N-terminal part of the 75-kDa subunit.

Materials and Methods

Isolation of the 51-kDa Subunit. Complex I was isolated from bovine heart mitochondria as described before (Runswick et al., 1989). Its subunits were fractionated by electrophoresis in the presence of sodium dodecyl sulfate on a 10–25% polyacrylamide gel (Laemmli, 1970). Proteins were detected with Page blue 83 dye. The stained 51-kDa subunit was excised with a razor blade and eluted in a buffer containing 0.1 M ammonium bicarbonate and 0.1% sodium dodecyl sulfate (final volume ca. 500 μL). The extract was dried down; the residue was redissolved in water (50 μL) and precipitated at −20 °C overnight with 95% ethanol (0.95 mL).

Protein Chemical Methods. Partial and total cyanogen bromide cleavages, and proteolysis with *Staphylococcus aureus* V8 protease, were performed on purified 51-kDa subunit as described before for the 75-kDa subunit of complex I (Runswick & 1989). Digestion of the 51-kDa subunit (10–20 μg) with endoproteinase Lys-C (0.25 μg) was carried out for 2 h at 37 °C in 0.1 M ammonium bicarbonate buffer (40 μL). Samples of the protein were cleaved with 70% or 98% formic acid for 12 h at room temperature. Intact complex I and the fragments in these digests were fractionated on minigels (Runswick et al., 1989) and then transferred by the semidry electrophoresis method to a poly(vinylidene difluoride) (PVDF) membrane [Immobilon, from Millipore (U.K.) Ltd., Harrow, U.K.; Matsuidara, 1987]. Between the gel and cathode, and between the membrane and anode, are placed one and two pieces of Whatman 3MM paper, respectively, each wetted with a different buffer. These are the following: 25 mM Tris-HCl (pH 9.4)/10% methanol (sheet next to cathode); 0.3 M Tris-HCl (pH 10.4)/5% methanol (sheet next to anode); 25 mM Tris-HCl (pH 10.4)/5% methanol (sheet next to membrane). The transfer of proteins is carried out for 30 min with a current of 250 mA. Proteins adsorbed to the membrane are detected by staining with a solution of 0.2% Page Blue 83 dye in 50% methanol and 1% acetic acid, and destained with 50% methanol. Stained bands were excised and analyzed in a modified Applied Biosystems 470 gas-phase protein sequencer (Fearnley et al., 1989). No attempt was made to detect cysteine residues. The sequence from the N-terminus of the intact 51-kDa subunit was extended by the introduction into the Edman degradation cycles of two consecutive acid cleavage steps for glycine and proline residues.

Oligonucleotide Synthesis. Synthesis were performed in an Applied Biosystems 320B DNA synthesizer. Oligonucleotides were used as sequencing primers, as primers in polymerase chain reactions, and as hybridization probes to investigate the products of these reactions. Forward and reverse primers in polymerase chain reactions were made with *Eco*RI or *Hin*dIII linkers, respectively, on their 5′ ends. In all cases, synthetic oligonucleotides were purified by absorption on C_{18} Sep-Pak columns (Waters Associates, Milford, MA) followed by elution with 60% methanol. Conditions for radiolabeling oligonucleotides and for their use as hybridization probes have been described before (Powell et al., 1989; Walker et al., 1989).

Polymerase Chain Reactions. Single-stranded and double-stranded cDNAs were made from poly(A^+) mRNA (Viñas et al., 1990), using a cDNA synthesis kit (Amersham International, U.K.). Segments of single-stranded cDNA were amplified under conditions described in detail below, using a variety of unique and mixed synthetic oligonucleotide primers. In order to extend regions of known DNA sequence to the 3′ poly(A) tail, single-stranded DNA template was prepared by priming on mRNA with a 36-base oligonucleotide containing *Eco*RI, *Bam*HI, and *Hin*dIII restriction sites followed by 15 T nucleotides (A. Dupuis and J. E. Walker, unpublished results). Two kinds of template were used in experiments to extend known sequences to the 5′ extremity of the cDNA. First, single-stranded cDNA (derived from 1.5 μg of mRNA) was primed with random hexanucleotides, and the product was tailed at its 5′ end with a run of A residues using terminal deoxynucleotide transferase (Runswick et al., 1990). Second, circular cDNA was prepared as follows: the ends of double-stranded cDNA (100 ng) were repaired, and the product was circularized by incubation at 15 °C for 12 h in a buffer (30 μL) containing 33 mM Tris–acetate, pH 7.5, 66 mM potassium acetate, 10 mM magnesium acetate, 3 mM spermidine, bovine serum albumin (0.1 mg/mL), 0.75 mM ATP, 240 units of T4 DNA ligase, and 160 units of T4 RNA ligase. A portion (5 μL) of this solution was used in an amplification reaction.

Polymerase chain reactions were carried out by using a Techne programmable Dri-Block, Model PHC-1. The following schedule was followed when degenerate primers were employed; denaturation at 94 °C for 1 min; annealing for 2.0 min, either at 5 °C below the minimum dissociation temperature (Suggs et al., 1981) or at 40 °C, whichever was the lower; synthesis for 2 min at 70 °C for 35 cycles; and finally a single incubation at 70 °C for 7 min. With pairs of specific primers, and also with a specific primer and one annealing to the 5′ poly(dA) tail, or 3′ [polylinker plus poly(dA)] tail, an annealing temperature of 55 °C was used. The reactions were performed in a solution (100 μL) containing 50 mM potassium chloride, 10 mM Tris-HCl, pH 8.3, 1.5 mM magnesium chloride, each dNTP (200 μM), both primers (for concentrations see below), cDNA (10 ng), and *Thermus aquaticus* DNA polymerase (3 units; Perkin Elmer Cetus Corp.). The reaction mixture was prepared without enzyme and preincubated at 94 °C for 2.5 min. Then the enzyme was added and the reaction mixture covered with mineral oil (100 μL; Sigma Chemical Co. Ltd., Poole, Dorset, U.K.). When two unique primers were to be used in a polymerase chain reaction on linear cDNA, each was added to a concentration of 1 μM. Reactions using degenerate primers, or specific primers on circularized cDNA templates, were carried out in two stages; first, 20 cycles were performed with primer concentrations of 200 pM; second, the primer concentrations were increased to 2 μM, and amplification was continued for an additional 20 or 30 cycles. Amplifications between a specific primer and the 5′ poly(dA) tail or the 3′ [polylinker plus poly(dA)] tail were carried out with a primer concentrations of 25 nM for the first 15 cycles, followed by 20 more cycles at primer concentrations of 1 μM.

The products of polymerase chain reactions were analyzed on 1–1.4% high-melting agarose gels, in a buffer containing 90 mM Tris, 90 mM boric acid, 2.5 mM EDTA, and ethidium bromide (1 μg/mL). DNA fragments were transferred by capillary blotting to Hybond-N membranes (Amersham International, U.K.), cross-linked to the membrane by irradiation with UV light, and hybridized with radioactively labeled synthetic oligonucleotides at 5 °C below the minimum dissociation temperature. Fragments that hybridized with the probe were recovered from a duplicate gel by using the method

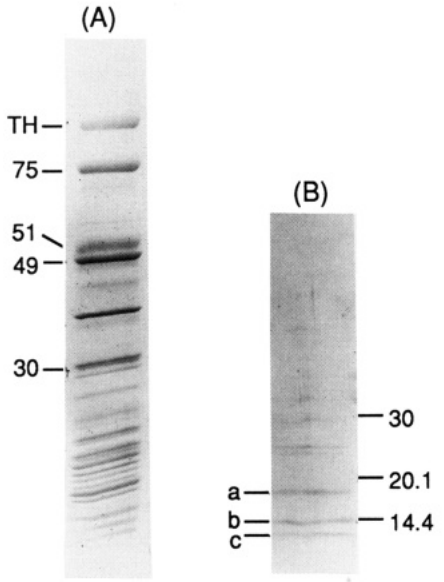

FIGURE 1: Separation of the subunits of complex I from bovine heart mitochondria and of cyanogen bromide fragments of the 51-kDa subunit. (A) Complex I; the positions of some of its subunits and of a contaminant, transhydrogenase (TH), are shown. (B) Cyanogen bromide cleavage products (a, b, and c; see Table I) of the 51-kDa subunit. The positions of molecular weight markers ($\times 10^{-3}$) are shown at the right-hand side.

of Vogelstein and Gillespie (1979) with a "Geneclean" Kit (Bio 101 Inc., La Jolla, CA). They were digested with *Eco*RI and *Hin*dIII, and the products were cloned into M13mp18 and M13mp19 vectors already cleaved with the same restriction enzymes. The products of amplifications between primers 5F and 5R were cleaved in addition with *Pst*I and *Hin*dIII, and cloned into appropriate M13 vectors. In experiments with pairs of specific primers, the gel blotting and DNA hybridization steps were omitted if a clear band of the expected size was detected on agarose gels. When the desired product was less than 100 bp in length, 10 μL of the reaction mixture was subjected to the "Geneclean" procedure and then digested with *Eco*RI and *Hin*dIII. The enzymes were inactivated by heating at 70 °C for 30 min, and the fragments were cloned into M13 vectors as described above.

Data Analysis. DNA sequences were determined by the modified dideoxy chain termination method (Sanger et al., 1977; Biggin et al., 1983). Data were compiled and analyzed with the computer programs DBUTIL (Staden, 1982a) and ANALYSEQ (Staden, 1985). The protein sequence was compared with those in the PIR, PIRNEW, and EMBL databases with the program FASTP (Lipman & Pearson, 1982). Pairwise comparisons of protein sequences were made with DIAGON (Staden, 1982b). Hydrophobicity profiles were calculated by using HYDROPLOT, a version of SOAP (Kyte & Doolittle, 1982), and structural predictions were made by the method of Garnier et al. (1978).

Results and Discussion

Partial Protein Sequence Analysis of the 51-kDa Subunit. The 51-kDa subunit of complex I can be separated from others by one-dimensional polyacrylamide gel electrophoresis (see Figure 1A). The purified protein has a unique N-terminal sequence, but the sequence data could not be interpreted beyond the 14th residue. This was due, at least in part, to the incomplete cleavage of proline-7 and glycine-13 in the Edman degradation. Therefore, a double-cleavage reaction was performed in these cycles, and then the sequence could be read out to residue 21 (see Table I). Unique internal protein

Table I: Protein Sequences Determined on the 51-kDa Subunit of Complex I and on Its Digestion Products

sample	apparent molecular mass (kDa)	determined sequence	position in protein
intact protein	51	SGDTTAPKKTSF-GSLKDEDXI	1–21
cyanogen bromide (a)	18	GARAAYIYIRGE-FYNEAXNLQV	130–151
cyanogen bromide (b)	12	NKPSDGRPKYL-VVNA	83–97
V8 protease	14	TVLMDFRAL-IQAQTG	314–328
Lys-C protease	30	QXKPRLKPPFP	200–210
98% formic acid	14	VGVFGXPTTVA	213–223

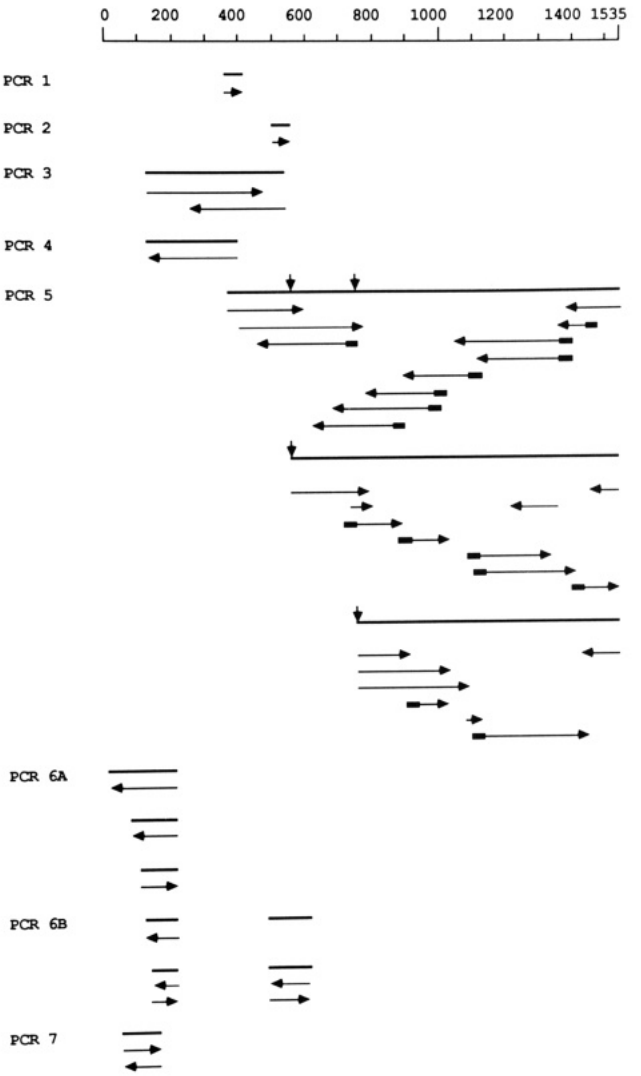

FIGURE 2: Generation by polymerase chain reactions and sequence analysis of clones encoding the 51-kDa subunit of complex I from bovine heart mitochondria. PCRs 1–7 are the cloned partial cDNAs generated by polymerase chain reactions (see the text and Figure 2 for details). The heavy lines represent these cDNAs, and the arrows indicate the directions and extents of sequences obtained from them. Boxes attached to these arrows indicate that internal synthetic oligonucleotide primers were used in the sequencing reactions. Vertical arrows indicate internal *Pst*I sites used in the sequence analysis. The scale is in bases.

sequences were obtained from fragments produced by cyanogen bromide (see Figure 1B), *Staph. aureus* V8 protease, endoproteinase Lys-C, and acid cleavages of the protein (Table I). Some of them were used in the design of cloning exper-

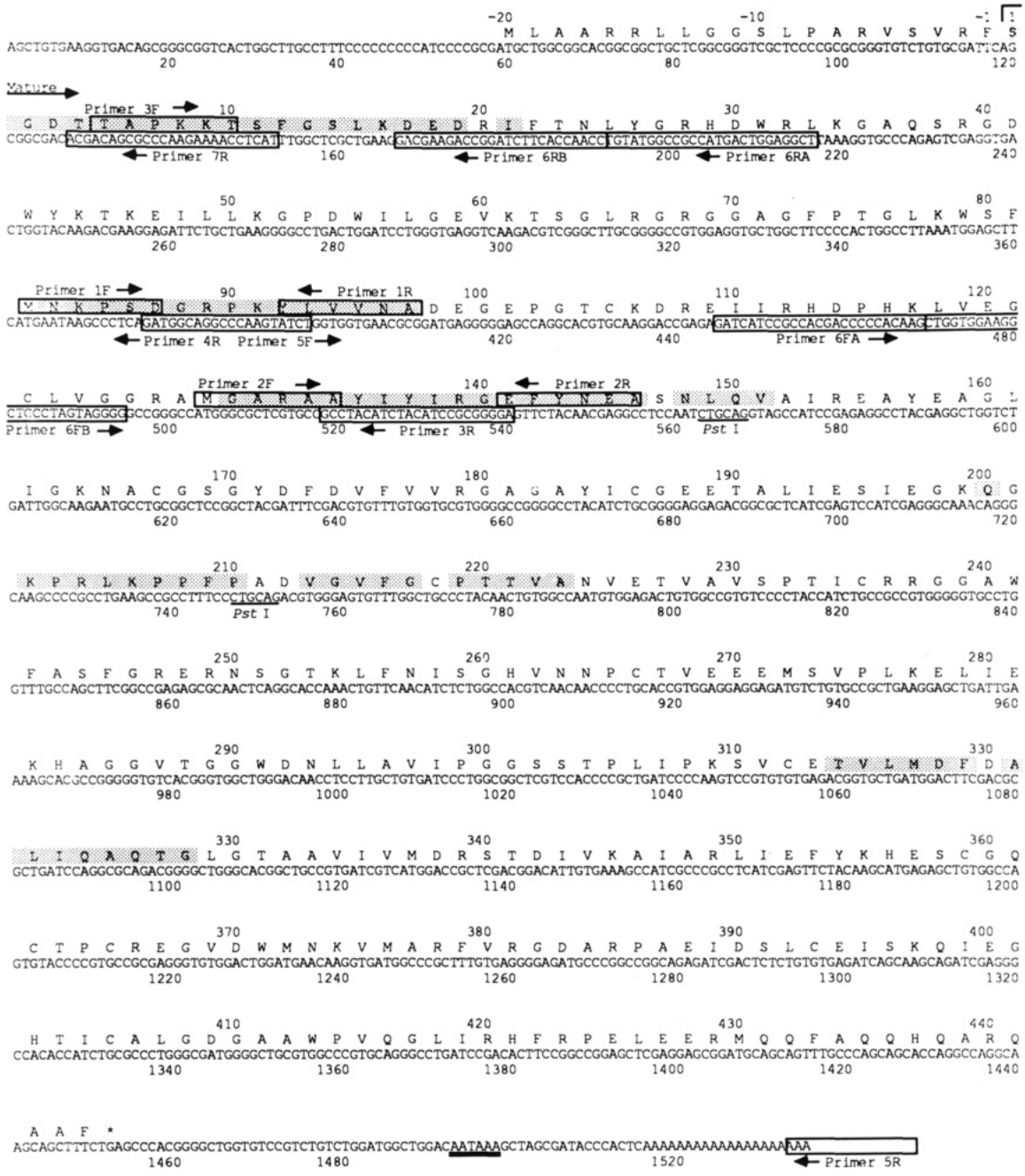

FIGURE 3: Complete cDNA sequence, compiled from partial cDNAs, encoding the 51-kDa subunit of bovine heart mitochondrial complex I. The position of the N-terminus of the mature protein is shown. Partial protein sequences obtained from the intact and fragmented subunit are shaded (see Table I for further details). Boxed protein sequences were used to design degenerate oligonucleotides for use as primers in polymerase chain reactions. Unique primers used in polymerase chain reactions are denoted by boxed nucleotide regions. The arrows pointing to the right or left indicate that the primers were synthesized with the sequence shown, or its complement, respectively. The primer complementary to a synthetic 5′ oligo(dA) tail, which was used to amplify the 5′ region of the cDNA, is not shown. A potential polyadenylation signal is underlined twice. Restriction enzyme sites used in the sequence analysis are indicated.

iments (see Figures 2 and 3). In other experiments, single bands from polyacrylamide gels were found to contain more than one component, but once the protein sequence of the 51-kDa subunit was known from the cDNA sequence, these mixed sequences could be deconvoluted, and all of them could be attributed to the 51-kDa subunit (data not shown).

Isolation and Sequence Analysis of Complementary DNA Clones Encoding the 51-kDa Subunit. On the basis of the partial N-terminal and internal protein sequences of the 51-kDa protein, a number of mixtures of oligonucleotides 17 bases in length were synthesized and employed as hybridization probes in attempts to isolate clones encoding the subunit from a bovine cDNA library (Gay & Walker, 1985). This library has been the source of clones of many components of the inner mitochondrial membrane, including the 75-, 49-, 30-, and 24-kDa components of complex I (Runswick et al., 1989; Fearnley et al., 1989; Pilkington et al., 1991; Pilkington & Walker, 1989). However, no clones for the 51-kDa subunit could be found in it.

Therefore, it was decided to adopt a different strategy based on the polymerase chain reaction, which has now been used to clone 20 proteins of mitochondrial origin (A. Dupuis, I. M. Fearnley, M. Finel, S. M. Medd, S. J. Pilkington, M. J. Runswick, J. M. Skehel, and J. E. Walker, unpublished work). The minimal requirement of this strategy is for a sequence of 17 consecutive amino acids in the protein of interest. In brief,

the procedure consists of synthesizing a short cDNA with the polymerase chain reaction, employing as forward and reverse primers synthetic oligonucleotide mixtures based on the N- and C-terminal extremities, respectively, of the available sequence. Then the reaction products are cloned into an M13 vector, and clones containing the sequence of interest are recognized with a synthetic hybridization probe based upon the central region of the known protein sequence. Both primers and probes are usually 17 bases long (plus appropriate linker sequences for primers; see Materials and Methods), and mixtures with complexities ranging from 48 to 4096 have been used successfully. The sequence of the short cDNA obtained in this way is accurate only in the central region between the mixed primers, and this region is used to design unique primers of 17 bases or more (excluding the linker) to generate more extensive cDNA clones in further polymerase chain reactions. If no other protein sequence is available, this unique oligonucleotide can be used with a "universal" primer complementary either to a synthetic 5′ (dA) tail added with terminal deoxynucleotide transfer or to the 3′ poly(dA) tail. Thereby, products covering the whole of the cDNA can be generated.

In the case of the 51-kDa subunit, only a limited N-terminal sequence was available, whereas several segments of internal protein sequence were known. Therefore, a modified strategy was used (see Figure 2), and two segments of protein sequence determined from cyanogen bromide cleavage products were used to generate short cDNA clones as described above (reactions PCR 1 and PCR 2 in Figure 2). Then two specific reverse primers derived from these clones were used separately with a degenerate primer based on the protein sequence near the N-terminus of the protein. These polymerase chain reactions (PCR 3 and PCR 4) each yielded a single product of about 250 and 420 bases, respectively. Their sequences encode amino acids 11–140 and 11–92 and established the order of the two cyanogen bromide fragments in the protein.

In order to extend the sequence to the 3′ end of the cDNA, primers 5F and 5R were used in further polymerase chain reactions (reaction PCR 5; see Figures 2 and 3) to make a product of about 1.2 kb. This was cloned and proved to extend from nucleotide 377 to the 3′ poly(A) tail. This sequence contains two *Pst*I sites, and fragments from these sites up to the 3′ end were used in establishing its sequence (Figures 2 and 3). Finally, two methods were used to extend the sequence in the 5′ direction. The nested primers 6RA and 6RB were used with primer 7F, which is complementary to the synthetic 5′ poly(dA) tail (experiment PCR 6A). The same nested primers (6RA and 6RB), and nested primers 6FA and 6FB, were used with circularized cDNA as template (experiment PCR 6B). Confirmation of the 5′ sequence was obtained by amplifying between primers 7R and 7F.

The complete cDNA sequence shown in Figure 2 is compiled from the shorter overlapping cDNA clones. It is 1535 nucleotides in length and is terminated at its 3′ end by a run of A residues separated by 17 nucleotides from the preceding sequence, AATAAA, a potential polyadenylation signal (Proudfoot & Brownlee, 1976).

Sequence of the 51-kDa Subunit of Complex I. The mature 51-kDa subunit of complex I encoded in the cDNA sequence is 444 amino acids long. Its molecular weight calculated from the sequence is 48 416 and agrees with the value of 51–53K estimated by polyacrylamide gel electrophoresis. The amino acid composition calculated from the sequence is in excellent agreement with that determined earlier by analysis of the bovine 51-kDa subunit (Galante & Hatefi, 1979). The mature protein is preceded by a 20 amino acid presequence, which contains 4 basic and no acidic amino acids, and is predicted by the method of Garnier et al. (1978) to be able to form an α-helical structure. These are characteristic features of mitochondrial import sequences, and, as in the present case, they often have an arginine residue close to the site of cleavage (von Heijne, 1986; Hendrick, 1989). The initiation codon cannot be identified with complete certainty, as there is no termination codon in the 5′ sequence of the cDNA in phase with the coding sequence, but it is reasonable to assume that nucleotides 59–61 serve this function.

Relationship between Complex I and Bacterial NAD-Reducing Hydrogenase. The protein sequences of the 51-, 24-, and 75-kDa subunits of complex I were compared with those of the four subunits of the soluble NAD^+-reducing hydrogenase from *A. eutrophus* (Tran-Betcke et al., 1990). Most of a region in HoxF extending from amino acids 4–186 is homologous to amino acids 30–209 of the 24-kDa subunit, and residues 189–565 of HoxF are closely related to the bovine 51-kDa subunit of complex I throughout the length of its polypeptide chain (Figure 5A,Bb). The HoxU (γ) sequence corresponds to amino acids 1–240 of the mitochondrial 75-kDa subunit (Figures 4B and 5B). No region of significant homology was found between the 49- and 30-kDa subunits of complex I (Fearnley et al., 1989; Pilkington et al., 1991) and the four subunits of the NAD^+-reducing hydrogenase.

Possible Nucleotide Binding Sites in the 51-kDa Subunit of Complex I and the HoxF Protein. Neither HoxF nor the 51-kDa subunit is related extensively to any other known sequence, including the soluble NADH dehydrogenase from *Escherichia coli* (Young et al, 1981). So, in order to try and identify sequences that could be involved in binding NADH in both proteins, their sequences were analyzed by the method proposed by Wierenga et al. (1986). This gives an assessment of the likelihood of a sequence folding into a β–α–β adenine nucleotide binding structure (Rossman et al., 1975). The sequence is compared with a "core" of four positions, occupied by three invariant glycine residues and an invariant acidic residue in NADH binding sites in known structures, and seven other positions are examined at which a particular class of amino acid is usually found. Glycines-1 and -3 allow a sharp turn at the β–α junctions, and glycine-2 permits close proximity between the end of the α-helix and the pyrophosphate of the nucleotide. The invariant acidic residues is at the end of the second β-strand, and forms a hydrogen bond with a 2′-OH, but would repel a phosphate group at this position, and so could be responsible for discrimination between NADH and NADPH [see Scrutton et al. (1990)]. The same "fingerprint" sequence has also been noted in other NADH binding enzymes (Scrutton et al., 1990). The region between the α-helix and the second β-strand is a loop of variable length in different proteins.

The sequences present in the 51-kDa subunit of complex I and HoxF each contain one sequence with three glycines separated by one and two amino acids, respectively (essential glycines-1, -2, and -3 in the "fingerprint"). These are summarized in Table II and are compared with sequences involved in NAD(H) binding in three dehydrogenases of known structure. The consensus requirements for amino acids at particular positions are indicated, and their presence or absence in the 51-kDa and HoxF sequences is scored. The score for each sequence is the number of positions at which the correct type of residue is present. The maximum attainable score is 11, and the sequences identified in the 51-kDa subunit and HoxF have scores of 9 and 8, respectively. Both proteins contain other sequences with two glycines separated by four

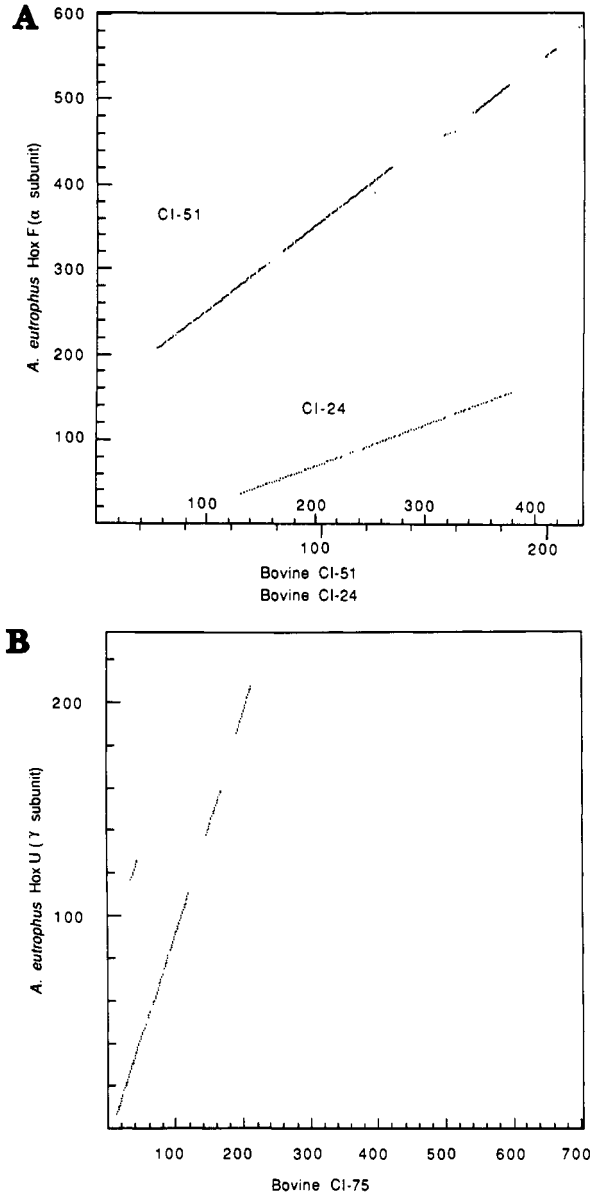

FIGURE 4: Comparisons of the sequences of the 51-, 24-, and 75-kDa subunits of bovine complex I (abscissas) with the HoxF and HoxU components of the NAD-reducing hydrogenase from *A. eutrophus* (ordinates). The comparisons were made by using the computer program DIAGON (Staden, 1982b) with a window of 25 amino acids and a score of 280. The scales are in amino acids.

residues (glycines-1 and -2 in the "fingerprint"), but none had a score >7 (data not shown). Therefore, residues 213–250 of HoxF and 62–99 of the 51-kDa subunit appear to be the most likely sequences to be involved in binding the ADP moiety of NAD(H). The two sites of deviation of the 51-kDa subunit sequence from the "fingerprint" (amino acids 76 and 79), where "small hydrophobic" residues are expected, have a glycine and a tryptophan residue, respectively. A glycine residue is present in the equivalent of residue 76 in the ADP part of the FAD binding site of *p*-hydroxybenzoate hydroxylase (Wierenga et al., 1979; Weijer et al., 1982), and dihydrolipoamide dehydrogenase (Scrutton et al., 1990) isolated from several sources has a glycine and a tryptophan, respectively, at the same two positions of its NADH binding site. The only difference between the 51-kDa subunit and HoxF among the "fingerprint" residues occurs at the first position, where a basic or hydrophilic residue is usually present, and HoxF has an aspartic acid residue. Glutathione reductase from human also

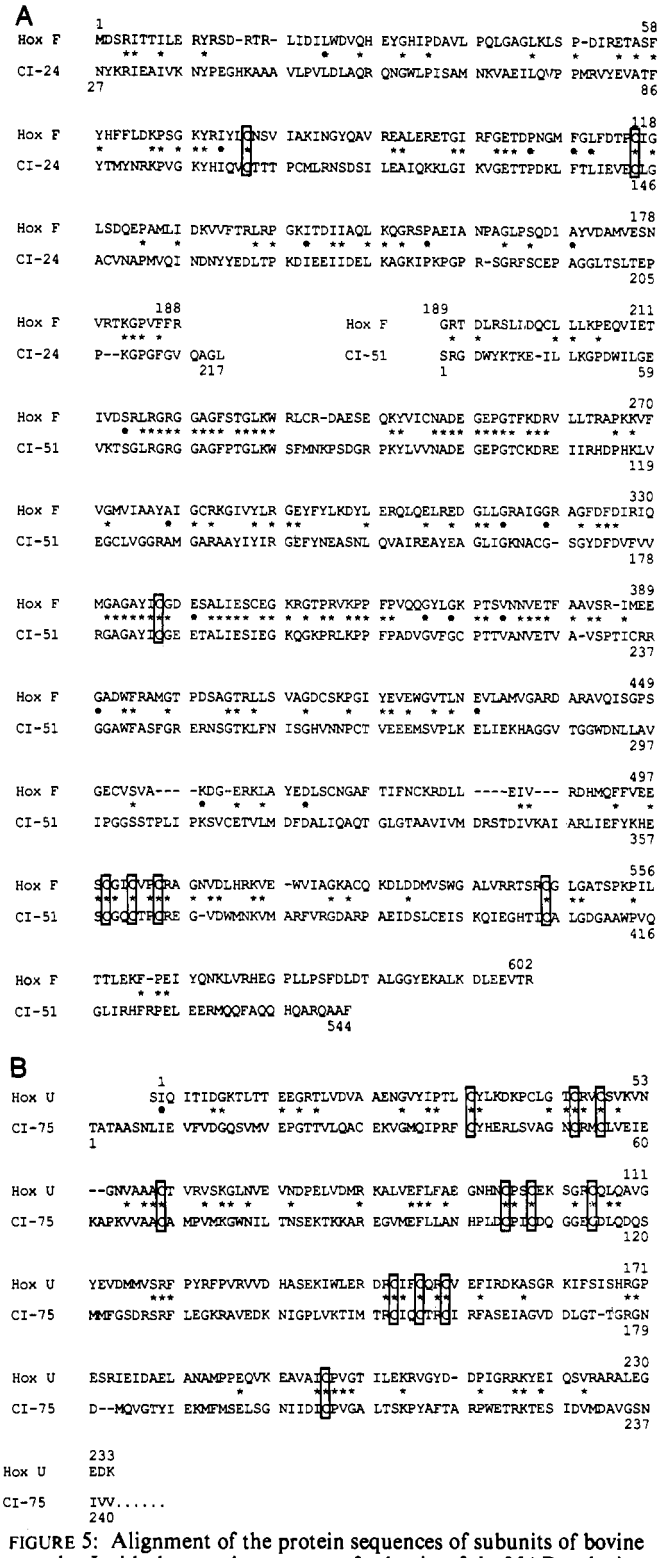

FIGURE 5: Alignment of the protein sequences of subunits of bovine complex I with the protein sequence of subunits of the NAD-reducing hydrogenase from *A. eutrophus*. The alignments were based upon the DIAGON calculations (Figure 4). (A) Alignment of the bovine 24- and 51-kDa subunits with residues 1–188 and 189–602, respectively, of the HoxF subunit from *A. eutrophus*. (B) The bovine 75-kDa subunit of complex I (residues 1–240) aligned with the HoxU protein from *A. eutrophus*. Identical residues are indicated by asterisks, and conserved cysteine residues are boxed. Insertions to improve the alignments are indicated by dashes. The sequences are numbered.

has an aspartic acid at this position of its binding site for the ADP portion of FAD (Krauth-Siegel et al., 1982; Schulz et al., 1982), so this difference is not incompatible with dinucleotide binding.

Table II: Alignment of Sequences Known To Bind the ADP Portion of NADH with Possible ADP Binding Regions from the 51-kDa Subunit of Bovine Complex I and the HoxF (α) Subunit of the NAD Reducing Dehydrogenase from *A. eutrophus* H 16[a]

Protein	Amino acid sequence	Score
Structure	____ βA ____ ____ αB ____ ____ βB ____	
LDH (21-52)	N K I T V V G V G A V G M A C A I S I L M K D - - - - - - - L A D E V A L V D	11
ADH (193-223)	S T C A V F G L G G V G L S V I M G C K A A G - - - - - - - - - A A R I I G V D	10
G3PDH L (1-31)	S K I G I N G F G R I G R L V L R A A L S C G - - - - - - - - - A Q V V A V N D	11
Consensus	Δ * * • • • * * * * * ^	
51 (61-99)	K T S G L R G R G G A G F P T G L K W S F M N K P S D G R P K Y L V V N A D E	9
Hox F (213-250)	V D S R L R G R G G A G F S T G L K W R L C R - D A E S E Q K Y V I C N A D E	8

[a] Abbreviations: LDH, dogfish lactate dehydrogenase (White et al., 1976; Taylor, 1977); ADH, horse liver alcohol dehydrogenase (Jörnval, 1970; Eklund et al., 1981); G3PDH L, lobster glyceraldehyde-3-phosphate dehydrogenase (Davidson et al., 1967; Moras et al., 1975). The consensus "fingerprint" is based on that described by Wierenga and Hol (1983). The symbols are as follows: (Δ) basic or hydrophilic residue (K, R, H, S, T, Q, N); (*) small and hydrophobic (A, I, L, V, M, C); (●) in variant glycine; (circumflex) acidic residue (D, E.). The score indicates the number of "fingerprint" positions at which the correct type of residue is found. Residues in agreement with the "fingerprint" are underlined. The sequences shown from the 51-kD subunit and the HoxF protein are the only ones in which all three essential glycines are present. None of the sequences in these two proteins which had two glycines separated by four residues (corresponding to essential glycines-1 and -3) had a score >7. The "fingerprint" is also present in other NADH binding proteins of unknown structure [see Scrutton et al. (1990)].

In NADH binding proteins of known structure, the nicotinamide–ribose moiety is bound by a β–α–β fold, which is very similar to those binding ADP. However, the sequences in this regions are not evidently conserved, and their structures are more variable than those of ADP binding sites (Rossman et al., 1975). Thus, it is not possible to predict the parts of the polypeptide chain of the 51-kDa subunit of HoxF that form the nicotinamide–ribose binding pocket from the sequences alone. However, it should be noted that in all NADH binding proteins of known structure, the site is formed by amino acid residues that lie to the C-terminal side of those forming the ADP binding pocket, and so in the 51-kDa subunit of complex I and HoxF these residues are likely to be found after residues 110 and 262, respectively.

The dimer of the α- and γ-subunits of bacterial enzyme, which constitute an NADH oxidoreductase activity, contains one molecule of bound flavin. Given that the γ-subunit seems to be involved in providing the structure for Fe–S clusters (see below), it is therefore reasonable to suppose that the FMN binding site will be in the α-subunit (HoxF). The evidence for the location of the FMN binding site in the 51-kDa subunit of complex I is similarly indirect and uncertain, as the prosthetic group is lost on resolution of the FP fraction into its constituent subunits. However, it is the probable acceptor of electrons from NADH (Ingledew & Ohnishi, 1980), and so its binding to the 51-kDa protein is argued on that basis.

Two different types of FMN binding sites have been described in other proteins. The flavodoxins, small bacterial proteins which transfer electrons between other redox proteins using FMN as their only prosthetic group, bind FMN in a similar β–α–β fold to the ADP fold described above (Mayhew & Ludwig, 1975; Smith et al., 1983). However, they do not have the pattern of amino acid residues associated with the ADP binding site. A quite different FMN binding site is found in spinach glycolate oxidase (Lindqvist & Brändén, 1985), in yeast flavocytochrome b_2 (Xia et al., 1987), and in trimethylamine dehydrogenase from the methylotrophic bacterium W3A1 (Lim et al., 1986). In these enzymes, the FMN is positioned between the end of an eight-stranded α–β barrel and five α-helices in the first two enzymes, and three α-helices in the last.

The sequences of flavodoxins are conserved and are related distantly to that of rat NADPH–cytochrome P-450 oxidoreductase (Porter & Kasper, 1985), which also uses FMN as one of its cofactors. However, no regions of significant homology were found between the 51-kDa subunit of complex I or the HoxF protein of *A. eutrophus* and these proteins. Nor are the sequences of the 51-kDa and HoxF proteins evidently related to spinach glycolate oxidase (Volokita & Somerville, 1987) or yeast flavocytochrome b_2 (Lederer et al., 1985; Guiard, 1985), and there are no clear indications from secondary structure prediction of an eight-stranded α–β barrel in either protein. The essential features of both types of FMN binding site are poorly defined, and so the presence of an FMN binding site in the 51-kDa subunit or HoxF is neither proved nor excluded by these comparisons. No sequences related to either NAD(H) binding sites or FMN binding proteins were found in the 24-kDa subunit. The sequence of the 10-kDa subunit has not been determined, but the protein may be too small to contain an FMN binding site, since the smallest known flavodoxins have molecular weights around 15K (Mayhew & Ludwig, 1975). Also, nearly all of the sequence of the α- and γ-subunits of the NAD-reducing hydrogenase is homologous to known subunits of complex I, and so it is unlikely that there is a homologue of the 10-kDa subunit of complex I in the FMN-containing fraction of the bacterial NAD-reducing hydrogenase. In our present state of knowledge, the region most likely to be involved in the FMN binding sites is the highly conserved sequence encompassing residues 332–386 and 180–234, respectively, in the HoxF and 51-kDa proteins (see Figure 5A).

Iron–Sulfur Centers. Various estimates have been made of the number and type of Fe–S clusters in the bacterial hydrogenase. The presence in the enzyme of two binuclear and two tetranuclear centers was proposed originally (Schneider et al., 1979), and subsequently this was amended to one binuclear and three tetranuclear centers (Schneider et al., 1984). The NADH oxidoreductase ($\alpha\gamma$) dimer of the NAD-reducing hydrogenase from *Nocardia opaca*, an enzyme which is very similar to the enzyme from *A. eutrophus*, contains two 4Fe–4S clusters and a 2Fe–2S cluster (Schneider et al., 1984), but the Fe–S clusters were not assigned to specific components of the dimer. In bovine complex I, electron paramagnetic resonance spectra and optical spectra of the 51-kDa subunit show that its Fe–S center loses its characteristic spectra on isolation from the FP fraction. Consequently, the type of cluster cannot be identified unambiguously but was considered by Ohnishi et al. (1985) to be a tetranuclear cluster, designated N-3. Spin–spin interaction between this cluster and the FMN free radical has also been reported (Salerno et al., 1977), which is consistent with the assignment of both these prosthetic groups to the 51-kDa subunit.

On the basis of the aligned protein sequences from complex I and *A. eutrophus*, a number of proposals about these clusters can now be made. The conserved sequence motif CysXX-CysXXCys in the HoxF and 51-kDa proteins is good evidence

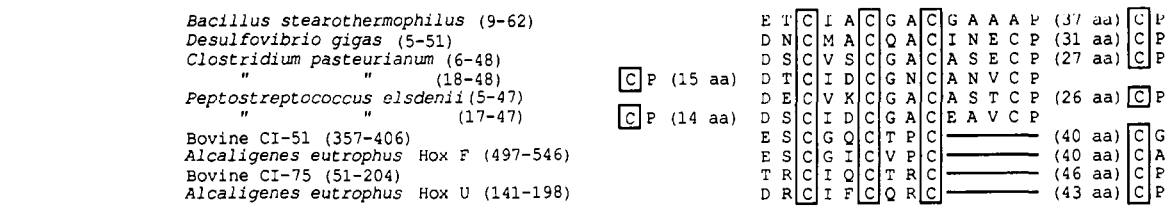

FIGURE 6: Alignment of amino acid sequences surrounding 4Fe–4S iron–sulfur centers in bacterial ferredoxins with sequences present in the 51- and 75-kDa subunits of complex I, and in the HoxF and HoxU proteins from *A. eutrophus*. Numbers in parentheses indicate the residue positions within each sequence. The four cysteine residues which ligand each iron–sulfur center in the ferredoxins, and possible iron–sulfur center ligands in the other protein sequences, are boxed. In each ferredoxin, this consists of the three cysteines in the CysXXCysXXCys motif, and of a fourth more distant cysteine residue (also boxed), which is followed by a proline residue. The ferredoxin sequences and their alignments are taken from Yasunobu and Tanaka (1980).

for these residues being involved in a tetranuclear center in each protein, as this arrangement has been found in 4Fe–4S centers in ferredoxins (see Figure 5). In the ferredoxins, the fourth cysteine ligand is followed by a proline residue, and is usually some distance away in the primary structure, either to the N-terminal or to the C-terminal side of the motif. Two other cysteines are conserved in HoxF and the 51-kDa subunit at residues 186 and 405 in the latter sequence, and it is possible that either could provide the fourth ligand. Cysteine-186 in the region that was proposed above as forming part of the FMN binding site, and on that basis alone, cysteine-405 is perhaps the more likely participant in the Fe–S cluster. The sequence Cys-Pro occurs at amino acids 218–219 in the 51-kDa protein, but the cysteine is not conserved in HoxF. None of the other cysteines in either protein is clustered with others, and the cysteine motif associated with 2Fe–2S clusters is absent from both proteins (Yasunobu & Tanaka, 1980; Cammack, 1983). Thus, it appers that these regions of the 51-kDa and HoxF proteins could contain one 4Fe–4S center but are unlikely to contain 2Fe–2S centers. This interpretation differs from the proposal based upon the sequence that HoxF contains a 2Fe–2S and not a 4Fe–4S center (Tran-Beckte et al., 1990).

Ligands for a second 4Fe–4S center appear to be present in the region of homology between HoxU and the 75-kDa subunit of complex I (see Figure 6) in which the characteristic CysXXCysXXCys sequence is conserved. The fourth cysteine ligand could be provided by the conserved cysteine-203 (75-kDa subunit numbering), which is followed by proline in both sequences, although among the seven other cysteines that are also conserved (see Figure 4B) cysteines-41, -69, and possibly -115 (75-kDa numbering) could also possibly fulfill this role. The cysteine residue at position 41 in HoxU, which was suggested to be involved in a 4Fe–4S center (Tran-Betcke et al., 1990), is not conserved in the 75-kDa subunit. None of the remaining conserved cysteine residues belong to cysteine-rich motifs characteristic of 2Fe–2S and 4Fe–4S centers. Nevertheless, they could be involved in liganding a 2Fe–2S or a high-potential 4Fe–4S iron–sulfur center.

The 24- or 10-kDa subunits of complex I contain a 2Fe–2S center designated N-1b (Ohnishi et al., 1985). The 24-kDa subunit has no characteristic cysteine-rich motifs, but two of its five cysteines are conserved in HoxF (Figure 5A). Here a number of interpretations are possible. First, an Fe–S center could be formed in HoxF using cysteines from the region related to the 51-kDa subunit as well as those in the region related to the 24-kDa protein. Second, the number or location of Fe–S centers involved in electron transport in the two enzymes may be different with a 2Fe–2S center present in the 24-kDa protein, but not in the equivalent part of HoxF. Third, the 2Fe–2S center may be located in the 10-kDa subunit of complex I and have no equivalent in the NAD-reducing dehydrogenase.

It should be emphasized that the assignment of parts of the sequence of the 51-kDa subunit and HoxF as being involved in NAD(H), FMN, and Fe–S binding sites must be considered as tentative in the absence of corroborative structural evidence, and this is particularly so in the case of the FMN binding site. Nonetheless, a final point of note is that these assignments place the three domains for NAD(H), FMN, and the Fe–S cluster in that order, N- to C-terminus in the proteins.

Implications for the Mechanism of Complex I. The path of electron transfer through complex I is far from clear. Various schemes have been proposed (Krishnamoorthy & Hinkle, 1988; van Belzen & Albracht, 1989). The common features are that FMN acts as the immediate acceptor of electrons from NADH and that the tetranuclear cluster N-2 transfers electrons to ubiquinone. Between these points, the tetranuclear clusters N-3 (g_{xyz} = 1.86, 1.93, and 2.04) and N-4 (g_{xyz} = 1.88, 1.94, and 2.10) and the binuclear cluster N-1b may be involved. The binuclear cluster N-1a reported by Ohnishi (1975, 1976) does not appear to be reduced by NADH, and has not been detected by others (Beinert & Albracht 1982). The order of redox centers in the chain is difficult to determine because all of the above four Fe–S clusters are fully reduced 5 ms after mixing NADH with the enzyme, but the use of various inhibitors and chemical modifiers (Krishnamoorthy & Hinkle, 1988) indicates that N-1b may come before N-3 and N-4. The N-1b cluster has been assigned to the 24- or 10-kDa subunits, and the N-3 cluster is considered to be in the 51-kDa subunit. The locations of Fe–S centers in the IP fraction are less clear (Ragan et al., 1982, Ragan, 1987), but the 49-, 30-, and 13-kDa subunits may contain a 4Fe–4S and a 2Fe–2S cluster between them, and the 75-kDa subunit is thought to contain the N-1a 2Fe–2S center. The presence of sequences homologous to the 51-, 24-, and 75-kDa subunits in a bacterial enzyme that can transfer electrons between NAD(H) and a variety of electron acceptors/donors may help to resolve the Fe–S centers involved, and place them in order. The 4Fe–4S (possibly N-3) is present in the 51-kDa subunit and its homologue, but the location of center N-1b is less clear from the sequence. The presence of a conserved 4Fe–4S center in the 75-kDa protein and in its bacterial homologue suggests that this center may be an early component of the electron-transfer path in complex I.

Phylogenetic Implications. It is evident that a portion of complex I is phylogenetically related to the water-soluble bacterial NAD^+-reducing hydrogenase. Whereas complex I is a relatively ubiquitous enzyme, the soluble NAD^+-reducing hydrogenase is found only in a few bacterial species. Aerobic hydrogen-oxidizing bacteria form a diverse group. These bacteria contain uptake hydrogenases which provide the reducing power for CO_2 fixation and which also feed electrons into the respiratory chain to drive oxidative phosphorylation (Bowien & Schlegel, 1981). In all but a few species, these bacteria rely on a membrane-bound hydrogenase that can

reduce NAD^+ only by reverse electron transport via the respiratory chain, presumably reversing electron flow through complex I. *A. eutrophus* is unusual in that it contains a soluble hydrogenase which directly reduces NAD^+. Perhaps it should not be surprising that two of the four subunits of this hydrogenase are closely related to those subunits of complex I which appear to be involved with the NAD^+/NADH oxidoreductase activity.

Recent work has pointed out strong sequence similarities of several other subunits of complex I with gene products that are probably subunits of the hydrogenase portion of *E. coli* formate hydrogenlyase (Böhm et al., 1990). This hydrogenase uses electrons from the oxidation of formate to reduce protons and release H_2. The overall reaction is to split formate into H_2 plus CO_2. Formate hydrogenlyase is used during some fermentative growth conditions to eliminate excess formate. Of particular interest is the presence of a homologue of the ND1 subunit of complex I in the formate hydrogenlyase. The ND1 subunit has been implicated in the binding of the inhibitors piericidin and rotenone to complex I (Earley et al., 1987). These inhibitors appear to function at the quinone-reducing portion of complex I. Whether complex I evolved by co-opting portions of these fermentative and lithiotropic enzymes or vice versa is not clear.

ACKNOWLEDGMENTS

We thank Prof. C.-I. Brändén and Prof. D. Rice for discussions about FMN and NAD binding sites, respectively, Dr. I. M. Fearnley for the preparation of complex I, and Mr. T. V. Smith and Mrs. J. Fogg for the synthesis of oligonucleotides.

REFERENCES

Beinert, H., & Albracht, S. P. J. (1982) *Biochim. Biophys. Acta 683*, 245–277.
Biggin, M. D., Gibson, T. J., & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3963–3965.
Böhm, R., Sauter, M., & Böck, A. (1990) *Mol. Microbiol. 4*, 231–243.
Bowien, B., & Schlegel, H. G. (1981) *Annu. Rev. Microbiol. 35*, 405–452.
Cammack, R. (1983) *Chem. Scr. 21*, 87–95.
Chen, S., & Guillory, R. J. (1981) *J. Biol. Chem. 256*, 8318–8323.
Chomyn, A., Mariottini, P., Cleeter, M. W. J., Ragan, C. I., Matsuno-Yagi, A., Hatefi, Y., Doolittle, R. F., & Attardi, G. (1985) *Nature 314*, 592–597.
Chomyn, A., Cleeter, M. W. J., Ragan, C. I., Riley, M., Doolittle, R. F., & Attardi, G. (1986) *Science 234*, 614–618.
Davidson, B. E., Sajgo, M., Noller, H. F., & Harris, J. I. (1967) *Nature 216*, 1181–1185.
Deng, P. S. K., Hatefi, Y., & Chen, S. (1990) *Biochemistry 29*, 1094–1098.
Dooijewaard, G., & Slater, E. C. (1976a) *Biochim. Biophys. Acta 440*, 1–15.
Dooijewaard, G., & Slater, E. C. (1976b) *Biochim. Biophys. Acta 440*, 16–35.
Earley, F. G. P., Patel, S. D., Ragan, C. I., & Attardi, G. (1987) *FEBS Lett. 219*, 108–113.
Eklund, H., Samana, J. P., Wallén, L., Brändén, C. I., Åkeson, Å., & Jones, T. A. (1981) *J. Mol. Biol. 146*, 561–587.
Fearnley, I. M., Runswick, M. J., & Walker, J. E. (1989) *EMBO J. 8*, 665–672.
Galante, Y. M., & Hatefi, Y. (1979) *Arch. Biochem. Biophys. 192*, 559–568.
Garnier, J., Osguthorpe, D. J., & Robson, B. (1978) *J. Mol. Biol. 120*, 97–120.
Gay, N. J., & Walker, J. E. (1985) *Biochem. J. 225*, 707–712.
Guiard, B. (1985) *EMBO J. 4*, 3265–3272.
Hatefi, Y., & Stempel, K. E. (1969) *J. Biol. Chem. 244*, 2350–2357.
Hendrick, J. P., Hodges, P. E., & Rosenberg, L. E. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4056–4060.
Ingledew, W. J., & Ohnishi, T. (1980) *Biochem. J. 186*, 111–117.
Jörnvall, H. (1970) *Eur. J. Biochem. 16*, 41–49.
Krauth-Siegel, R. L., Blatterspiel, R., Saleh, M., Schiltz, E., Schirmer, R. H., & Untucht-Grau, R. (1982) *Eur. J. Biochem. 121*, 259–267.
Krishnamoorthy, G., & Hinkle, P. C. (1988) *J. Biol. Chem. 263*, 17566–17575.
Kyte, J., & Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105–132.
Laemmli, U. K. (1970) *Nature 227*, 680–685.
Lederer, F., Cortial, S., Becam, A. M., Haumont, P. Y., & Perez, L. (1985) *Eur. J. Biochem. 152*, 419–428.
Lim, L. W., Shamala, N., Mathews, F. S., Steenkamp, D. J., Hamlin, R., & huu Xuong, N. (1986) *J. Biol. Chem. 261*, 15140–15146.
Lindqvist, Y., & Brändén, C. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 6855–6859.
Lipman, D. J., & Pearson, W. R. (1985) *Science 227*, 1435–1441.
Matsudaira, P. (1987) *J. Biol. Chem. 262*, 10035–10038.
Mayhew, S. G., & Ludwig, M. L. (1975) *Enzymes (3rd Ed.) 12*, 57–118.
Moras, D., Olsen, K. W., Sabesan, M. N., Buehner, M., Ford, G. C., & Rossmann, M. G. (1975) *J. Biol. Chem. 290*, 9137–9162.
Ohnishi, T. (1975) *Biochim. Biophys. Acta 387*, 475–490.
Ohnishi, T. (1976) *Eur. J. Biochem. 64*, 91–103.
Ohnishi, T., Ragan, C. I., & Hatefi, Y. (1985) *J. Biol. Chem. 260*, 2782–2788.
Pilkington, S. J., & Walker, J. E. (1989) *Biochemistry 28*, 3257–3264.
Pilkington, S. J., Skehel, J. M., & Walker, J. E. (1991) *Biochemistry 30*, 1901–1908.
Porter, T. D., & Kasper, C. B. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 973–977.
Powell, S. J., Medd, S. M., Runswick, M. J., & Walker, J. E. (1989) *Biochemistry 28*, 866–873.
Proudfoot, N. J., & Brownlee, G. G. (1976) *Nature 263*, 211–214.
Ragan, C. I. (1987) *Curr. Top. Bioenerg. 15*, 1–36.
Ragan, C. I., Galante, Y. M., & Hatefi, Y. (1982) *Biochemistry 21*, 2518–2524.
Rossman, M. G., Liljas, A., Brändén, C. I., & Banaszak, L. J. (1975) *Enzymes (3rd Ed.) 11*, 61–102.
Runswick, M. J., Gennis, R. B., Fearnley, I. M., & Walker, J. E. (1989) *Biochemistry 28*, 9452–9459.
Runswick, M. J., Walker, J. E., Bisaccia, F., Iacobazzi, V., & Palmieri, F. (1990) *Biochemistry 29*, 11033–11040.
Salerno, J. C., Ohnishi, T., Lim, J., Widger, R., & King, T. E. (1977) *Biochem. Biophys. Res. Commun. 75*, 618–624.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Schneider, K., Cammack, R., Schlegel, H. G., & Hall, D. O. (1979) *Biochim. Biophys. Acta 578*, 445–461.
Schneider, K., Cammack, R., & Schlegel, H. G. (1984) *Eur. J. Biochem. 142*, 75–84.
Schulz, G. E., Schirmer, R. H., & Pai, E. F. (1982) *J. Mol. Biol. 160*, 287–308.
Scrutton, N. S., Berry, A., & Perham, R. N. (1990) *Nature 343*, 38–43.

Smith, W. W., Pattridge, K. A., & Ludwig, M. L. (1983) *J. Mol. Biol. 165*, 737–755.
Staden, R. (1982a) *Nucleic Acids Res. 10*, 4731–4751.
Staden, R. (1982b) *Nucleic Acids Res. 10*, 2951–2961.
Staden, R. (1985) in *Genetic Engineering: Principles and Methods* (Setlow, J. K., & Hollaender, A., Eds.) pp 67–114, Plenum Publishing Corp., New York and London.
Suggs, S. V., Hirose, T., Miyake, T., Kawashima, E. H., Johnson, M. J., Itakura, K., & Wallace, R. B. (1981) in *Developmental Biology Using Purified Genes* (Brown, D. D., Ed.) pp 683–693, Academic Press, New York and London.
Taylor, S. S. (1977) *J. Biol. Chem. 252*, 1799–1806.
Tran-Betcke, A., Warnecke, U., Böcker, C., Zabarosch, C., & Friedrich, B. (1990) *J. Bacteriol. 172*, 2920–2929.
van Belzen, R., & Albracht, S. P. J. (1989) *Biochim. Biophys. Acta 974*, 311–320.
Viñas, O., Powell, S. J., Runswick, M. J., Iacobazzi, V., & Walker, J. E. (1990) *Biochem. J. 265*, 321–326.
Vogelstein, B., & Gillespie, D. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 615–619.
Volokita, M., & Somerville, C. R. (1987) *J. Biol. Chem. 262*, 15825–15828.
von Heijne, G. (1986) *EMBO J. 5*, 1335–1342.
Walker, J. E., Powell, S. J., Viñas, O., & Runswick, M. J. (1989) *Biochemistry 28*, 4702–4708.
Weijer, W. J., Hofsteenge, J., Vereijken, J. M., Jekel, P. A., & Beintema, J. J. (1982) *Biochim. Biophys. Acta 704*, 385–388.
White, J. L., Hackert, M. L., Buehner, M., Adams, M. J., Ford, G. C., Lentz, P. J., Jr., Smiley, I. E., Steindel, S. J., & Rossmann, M. G. (1978) *J. Mol. Biol. 102*, 759–779.
Wierenga, R. K., & Hol, W. G. (1983) *Nature 302*, 842–844.
Wierenga, R. K., De Jong, R. J., Kalk, K. H., Hol, W. G. J., & Drenth, J. (1979) *J. Mol. Biol. 131*, 55–73.
Wierenga, R. K., Terpstra, P., & Hol, W. G. J. (1986) *J. Mol. Biol. 187*, 101–107.
Wikström, M. (1984) *FEBS Lett. 169*, 300–304.
Xia, Z.-X., Shamala, N., Bethge, P. H., Lim, L. W., Bellamy, H. D., Xuong, N. H., Lederer, F., & Mathews, F. S. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 2629–2633.
Yasunobu, K., & Tanaka, M. (1980) *Methods Enzymol. 69*, 228–238.
Young, I. G., Rogers, B. L., Campbell, H. D., Jaworowski, A., & Shaw, D. C. (1981) *Eur. J. Biochem. 116*, 165–170.

Structural Study of Porcine Pancreatic Elastase Complexed with 7-Amino-3-(2-bromoethoxy)-4-chloroisocoumarin as a Nonreactivatable Doubly Covalent Enzyme–Inhibitor Complex[†,‡]

J. Vijayalakshmi,[§] Edgar F. Meyer, Jr.,[*,§] Chih-Min Kam,[∥] and James C. Powers[∥]
Department of Bichemistry and Biophysics, Texas A&M University, College Station, Texas 77843-2128, and School of Chemistry, Georgia Institute of Technology, Atlanta, Georgia 30332-0400
Received August 6, 1990; Revised Manuscript Received October 17, 1990

ABSTRACT: The complex of porcine pancreatic elastase (PPE) with 7-amino-3-(2-bromoethoxy)-4-chloroisocoumarin, a potent mechanism-based inhibitor, was crystallized and the crystal structure determined at 1.9-Å resolution with a final *R* factor of 17.1%. The unbiased difference Fourier electron density map showed continuous density from Oγ of Ser 195 to the benzoyl carbonyl carbon atom and from Nε2 of His 57 to the carbon atom at the 4-position of the isocoumarin ring in the inhibitor. This suggested unambiguously that the inhibitor was doubly covalently bound to the enzyme. It represents the first structural evidence for irreversible binding of an isocoumarin inhibitor to PPE through both Ser 195 and His 57 in the active site. The PPE–inhibitor complex is only partially activated in solution by hydroxylamine and confirms the existence of the doubly covalently bound complex along with the acyl enzyme. The benzoyl carbonyl oxygen atom of the inhibitor is not situated in the oxyanion hole formed by the amide (>NH) groups of Gly 193 and Ser 195. The complex is stabilized by the hydrogen-bonding interactions in the active site (from the Nε2 of Gln 192 to the bromine atom in the inhibitor and the amino group at the 7-position of the isocoumarin ring to the carbonyl oxygen of Thr 41) and by van der Waals interactions. The inhibition rates of several 7-substituted 4-chloro-3-(bromoalkoxy)isocoumarins toward PPE were measured. The *N*-alkylureido and *N*-(arylalkyl)ureido derivatives are more potent inhibitors of PPE than 7-amino-3-(2-bromoethoxy)-4-chloroisocoumarin due to a proposed favorable interaction between the *N*-alkyl group of the inhibitors and the hydrophobic subsites S_2' or S_3' subsites.

Serine proteases are capable of cleaving connective tissue proteins such as elastin as well as destroying invading bacteria. Excessively high levels of these proteases in the system or a functional deficiency of natural inhibitors to these proteases, such as α1-protease inhibitor, can cause inflammatory diseases such as pulmonary emphysema (Lungarella et al., 1985; Powers, 1976), arthritis (Janoff, 1973, 1978), pancreatitis (Geokas et al., 1968), adult respiratory distress syndrome (Burchardi et al., 1984), and certain degenerative skin disorders. The importance of serine proteases as pathogenic agents in a variety of diseases has stimulated considerable

[†]This work was supported by grants from the Robert A. Welch Foundation and the Texas Agricultural Experiment Station to E.F.M. and from the National Institutes of Health to J.C.P. (HL 29307).
[‡]Coordinates of the title complex are being submitted to the Brookhaven Protein Data Bank and are designated as 9EST.
[*]To whom correspondence should be addressed.
[§]Texas A&M University.
[∥]Georgia Institute of Technology.

0006-2960/91/0430-2175$02.50/0