Cloning and sequencing of the genes for the proton-translocating nicotinamide nucleotide transhydrogenase from *Rhodospirillum rubrum* and the implications for the domain structure of the enzyme

Ross Williams,1† Nick P. J. Cotton,1 Christopher M. Thomas2 and J. Baz Jackson1

Author for correspondence: J. Baz Jackson. Tel: +44 21 414 5423. Fax: +44 21 414 3982.

The genes for the proton-translocating nicotinamide nucleotide transhydrogenase from *Rhodospirillum rubrum* have been cloned using a probe constructed with the polymerase chain reaction, genomic DNA as target and oligonucleotide primers corresponding to amino acid sequence obtained from the purified soluble subunit. There is a cluster of three genes, designated pntAA, pntAB and pntB, whose translation products indicate polypeptides of 384, 139 and 464 amino acids, respectively. This contrasts with the situation in the enzymes from *Escherichia coli* (two polypeptides) and bovine mitochondria (one polypeptide) but there is close similarity between the sequences. PntAA is the soluble subunit of the enzyme from *R. rubrum*, equivalent to the relatively hydrophilic domain I that forms the N-terminal part of the α polypeptide of *E. coli* transhydrogenase and which probably contains the NAD(H)-binding site. PntAB corresponds to the strongly hydrophobic domain IIa at the C-terminus of the α polypeptide of the *E. coli* transhydrogenase. PntB corresponds to the *E. coli* β polypeptide, which comprises the strongly hydrophobic domain IIb and the relatively hydrophilic domain III, thought to contain the NADP(H)-binding site. The peptide bond between PntAA-Lys237 and -Glu238 of both the denatured and the native soluble subunit is very sensitive to proteolysis by trypsin and the neighbouring peptide bond Lys227-Thr228 to cleavage by the endoproteinase Lys-C. Related sites have been reported to be sensitive to trypsin in the *E. coli* and bovine mitochondrial enzymes. The two tryptic fragments from the native *R. rubrum* soluble subunit are unable to reconstitute transhydrogenase activity to membranes depleted of the soluble subunit but they can block reconstitution by intact soluble subunit. It is suggested that this protease-sensitive region separates two subdomains and that, after trypsinolysis, at least one retains structural integrity and can dock with domains II and/or III.

Keywords: transhydrogenase, gene cloning, gene sequencing, proton translocation, *Rhodospirillum rubrum*

INTRODUCTION

Transhydrogenase (H⁺-Thase), from the inner membranes of animal mitochondria and the cytoplasmic membranes of many bacteria, couples the transfer of reducing equivalents between NAD(H) and NADP(H) to the translocation of protons:

\[ \text{NADH} + \text{NADP}^+ + \text{H}^+_{\text{out}} \leftrightarrow \text{NAD}^+ + \text{NADPH} + \text{H}^+_{\text{in}} \]
Thus, in intact mitochondria and bacteria, the imposition of a transmembrane proton electrochemical gradient, for example through the action of a respiratory or photosynthetic electron-transport chain, drives the reaction to the right (Rydstrom et al., 1987; Jackson, 1991; Hatefi & Yamaguchi, 1992).

The amino acid sequences of H+-Thase from *Escherichia coli* (Clark et al., 1986; Ahmad et al., 1992) and bovine mitochondria (Yamaguchi et al., 1988) have been predicted from the nucleotide sequences of the genes (*pntA* and *pntB*) and the cDNA, respectively. The *E. coli* enzyme is composed of two polypeptides (molecular masses 54 kDa and 49 kDa). When their amino acid sequences are lined up contiguously (the C-terminus of α to the N-terminus of β), there is close similarity with the sequence of the single polypeptide (109 kDa) of the bovine enzyme (approximately 52% identity). Hydrophatic profiles (Clarke et al., 1986; Yamaguchi et al., 1988) indicate that H⁺-Thase has three large domains. In the bovine enzyme the approximate positions of these domains are from residue 1 to 431 (domain I), from residue 432 to 836 (domain II) and from residue 837 to 1043 (domain III). In *E. coli* H⁺-Thase the 'break' between the two polypeptides is located in domain II. For reasons that will become clear, it is appropriate to consider that domain II is composed of two subdomains; the part of domain II on the α polypeptide of *E. coli* (or its homologues), will be denoted domain IIA, and the part of the β polypeptide (or its homologues) as domain IIB.

Domain II (both IIA and IIB) of H⁺-Thase is strongly hydrophilic and might comprise between 12 and 14 transmembrane helices (Clarke et al., 1986; Yamaguchi et al., 1988; Tong et al., 1991). Domains I and III are relatively hydrophilic; they probably contain the NAD⁺/NADH-binding site and the NADP⁺/NADPH-binding site, respectively (Clarke et al., 1986; Wakabayashi & Hatefi, 1987a; Yamaguchi et al., 1988; Yamaguchi & Hatefi, 1993). The susceptibility of H⁺-Thase to proteases and the nature of the reaction with antibodies suggest that domains I and III both protrude on the matrix side of the membrane in bovine mitochondria (the cytoplasmic side in *E. coli*: Yamaguchi et al., 1990; Yamaguchi & Hatefi, 1991; Tong et al., 1991). It was shown (Cunningham et al., 1992a) that there are strong sequence similarities (approximately 30% identity) between domain I of H⁺-Thase and the enzyme, alanine dehydrogenase (Kuroda et al., 1990), which catalyses the reversible, reductive amination of pyruvate to alanine.

Recently, the translation of a sequence of cDNA isolated from human brain was found to be homologous with a part of domain IIB of bovine H⁺-Thase (Adams et al., 1992; Kramer et al., 1993). Sequence identity extends by two amino acid residues into the region which, in the bovine enzyme, forms the short polypeptide link between domains IIA and IIB. This indicates that the gene and polypeptide composition are similar in the two organisms. In contrast, the translation of a polymerase chain reaction product of genomic DNA from the protozoan parasite *Eimeria tenella* indicates that this organism has an H⁺-Thase with a different polypeptide and gene structure (Kramer et al., 1993). The full-length coding sequence establishes that H⁺-Thase from *Ei. tenella* comprises a single polypeptide, but that the domain order is different to that in the bovine enzyme; from the N-terminus of the protein it proceeds IIB > III > I > IIA. Thus, to observe sequence homology between *Ei. tenella* and *E. coli*, the two polypeptides of the bacterial enzyme must be aligned such that the C-terminus of the β polypeptide is contiguous with the N-terminus of the α. In that arrangement there is 45% amino acid identity between the sequences of H⁺-Thase from *E. coli* and *Ei. tenella* (Kramer et al., 1993).

Another different polypeptide composition is suggested by analysis of H⁺-Thase from the photosynthetic bacterium *Rhodospirillum rubrum* (Cunningham et al., 1992a). In this organism, the enzyme can be resolved, under mild conditions, into a water-soluble component (formerly known as Th₉) and a membrane-bound component (Fisher & Guillory, 1971a, b; Cunningham et al., 1992a). Separately, neither component has transhydrogenase activity but the protein can be fully reconstituted in the presence of nucleotides (Cunningham et al., 1992b). The N-terminal amino acid sequence and the molecular mass of the purified soluble transhydrogenase subunit indicated that it probably corresponds to domain I of H⁺-Thase (Cunningham et al., 1992a). In this report we describe the sensitivity of Th₉ to proteases, the cloning and sequencing of the genes for H⁺-Thase from *R. rubrum* and the implications for the domain structure of the enzyme.

**METHODS**

**Bacterial strains and growth conditions.** *R. rubrum* strain S1, from Dr L. Slooten, Vrije Universiteit, Brussels, Belgium, was grown anaerobically under phototrophic conditions on RCV medium (Weaver et al., 1975) at 30°C in completely sealed bottles, as described by Cunningham et al. (1992a). *E. coli* strain MC1061 (*bsdR merB araD139 A(araABC-leta)7679 AlacX74 galU galK rpsL tbi; Meissner et al., 1987) was grown with shaking in L-broth, or on L-agar (1.5%, w/v, agar added to L broth) at 37°C, as described by Sambrook et al. (1989). Antibiotic resistance was selected by addition of 100 µg ampicillin ml⁻¹ to the solid medium.

**Biochemical preparations and procedures.** Chromatophores (everted membrane vesicles) of *R. rubrum* and *C₇*-particles (chromatophores washed to remove soluble transhydrogenase subunit) were prepared, and the soluble transhydrogenase subunit was purified and assayed, as described by Cunningham et al. (1992a). Crude soluble subunit (type II Th₉) was prepared as in Fisher & Guillory (1971a). Bacteriochlorophyll was assayed using the *in vivo* absorption coefficient (Clayton, 1963). Light-driven transhydrogenation was measured by following the reduction of thio-NAD⁺ with NADH at 395 nm using known absorption coefficients (Palmer & Jackson, 1992).

When N-terminal amino acid sequences of fragments were required, purified soluble transhydrogenase subunit was first subjected to SDS-PAGE. The 43 kDa polypeptide was excised, electro-eluted in 0.05% (w/v) SDS, diluted and incubated

1596
overnight at 18 °C with protease (approximately 100:1 by mass). The samples were then concentrated, again subjected to SDS-PAGE and electro-blotted on to poly(vinylidene difluoride) membrane. Peptides were visualized with PAGE Blue-83, cut out and analysed with an Applied Biosystems 473A gas-phase sequencer. Further details of these procedures are described by Palmer et al. (1993). For the reconstitution experiments shown in Fig. 6, purified soluble subunit of R. rubrum H+-Thase (21 μg protein), dialysed against 1 mM dithiothreitol, 10 mM Tris/HCl pH 8.0, to remove salt and phenylmethylsulphonyl fluoride, was incubated for 5 h with 0·6 μg trypsin in a volume of 0·9 ml. The reaction was stopped with 6 μg trypsin inhibitor.

**In vitro manipulation and analysis of DNA.** Genomic DNA from *R. rubrum* was prepared by a modification of the method of Fitzmaurice et al. (1989), with the following exceptions. The 80 ml overnight culture was grown phototrophically (see above). The bacterial cells, harvested by centrifugation, were washed in 20 mM EDTA, 50 mM Tris/HCl pH 8.0, resuspended in 10 ml of the same medium supplemented with 0·5% SDS and 2·5 mg RNase, and incubated at 37 °C for 1 h. Proteinase K (25 mg) was added and the incubation was continued at 65 °C for 3 h. The lysate was extracted twice with phenol/chloroform/isoamyl alcohol (25:24:1, by vol.) and then twice with chloroform/isoamyl alcohol (24:1, v/v) followed by ethanol. The DNA was precipitated with ethanol, resuspended in sodium acetate, reprecipitated and washed in ethanol.

Plasmid DNA was prepared on both small and large scales by the alkaline-SDS method (Birnboim & Doly, 1979) and purified, where necessary, by caesium chloride/ethidium bromide equilibrium density-gradient centrifugation. DNA manipulations and the handling of bacteria were carried out by standard techniques (Sambrook et al., 1989). Restriction enzymes were used according to the manufacturers' instructions. Oligonucleotide probes were labelled with [γ-32P]dCTP, as described by Feinberg & Vogelstein (1984). DNA sequencing was performed both by the dideoxy-chain-termination method (Sanger et al., 1977) with a Sequenase version 2.0 kit from United States Biochemicals, and with an Applied Biosystems Incorporated 373A automated sequencer using the manufacturer's kit employing *Taq* polymerase and dye terminators according to their instructions.

Computer alignments, using the University of Wisconsin GCG package (Devereux et al., 1984), and searches, were carried out on the Owl database at the Science and Engineering Research Council Facility at Daresbury, and through the Biocomputing Unit at the University of Edinburgh. Hydropathic profiles were determined using a program written by Dr D. A. Rouch, School of Biological Sciences, University of Birmingham, based on the parameters described by Von Heijne (1992).

**Cloning of genes for H+-Thase from *R. rubrum***. On the basis of N-terminal sequence of Thα and peptides prepared by proteolysis of Thα, (Cunningham et al., 1992a), degenerate oligonucleotides were synthesized as primers for the polymerase chain reaction (PCR) using genomic DNA from *R. rubrum* as a target. The products were re-amplified and sequenced. Comparison with published sequences confirmed that the PCR products were derived from H+-Thase. A restriction map of *R. rubrum* DNA, using labelled PCR product as a probe of Southern blots, is shown in Fig. 1. *Bam*HI fragments of *R. rubrum* DNA between 4 and 7 kb were cloned into pBR322 and used to transform *E. coli* MC1061. Clones bearing inserts were selected, by colony hybridization, again using labelled PCR product as a probe. Four positive clones were selected, and shown by restriction analysis to contain DNA equivalent to the PCR product. One of these, pNIC1, was used for nucleotide sequencing (see Fig. 1). On the basis of the restriction map of this construct, DNA fragments were subcloned into pUC18 and nucleotide sequences of the 5′ and 3′ ends of the fragments were determined using universal forward and reverse primers. From these sequences oligonucleotides were constructed to serve as

---

**Fig. 1.** (a) Restriction map of the 6·1 kb *Bam*HI insert of *R. rubrum* DNA obtained by colony hybridization, and (b) the sequencing strategy: E, *Eco*RI; B, *Bam*HI; S, *Sal*I; H, *Hind*III. The solid arrows to the left and right show sequences obtained for the top and bottom DNA strands, respectively. The direction of the arrows show the direction of sequencing and the position of the arrows indicates regions over which sequence was obtained using universal forward and reverse, or newly-synthesized primers (see Methods). The dashed arrows indicate the position of the three H+-Thase genes. The numbers of base pairs between the structural genes and the *Bam*HI sites are only approximate.
primers for further sequencing using either pNIC1 or appropriate subclones in pUC18.

Materials. Nicotinamide nucleotides, TPKC-treated trypsin (T8642) and soybean trypsin inhibitor (T9003) were obtained from Sigma, sequencing-grade Lys-C and Asp-N endoproteinas from Boehringer, restriction enzymes and ligases from Northumbria Biologicals, radiolabelled nucleotides from Amersham, and oligonucleotides from Alta Bioscience.

RESULTS AND DISCUSSION

Nucleotide sequence of R. rubrum pnt genes reveals novel gene organization

The nucleotide sequence from the cloned BamHI fragment, and its translation product, are shown in Fig. 2. The pnt genes were assigned by locating regions coding for polypeptides containing sequence identical to that of proteolytic fragments of the soluble subunit of R. rubrum H+-Thase, and by similarity to the E. coli pnt genes. There are three ORFs, which we have designated pntAA, pntAB and pntB (reasons for this nomenclature will become evident below). Each is preceded by reasonable translational signals and is separated from the adjacent ORF by only 3 and 13 nucleotides, respectively. Downstream of pntB are two IVRs, both of which are potential rho-independent transcriptional terminators, since they consist of a G+C-rich hairpin, followed by a run of A+T-rich sequence (Platt, 1986). Upstream of pntAA is a region of at least 103 bp which is unlikely to be translated because of the presence of stop codons in all three frames. This region also contains two IVRs which may be involved with transcriptional termination (Platt, 1986) but could also be binding sites for regulatory proteins. These upstream and downstream features probably thus define the pnt genes as a translational unit containing three cistrons. Upstream of pntAA there is no obvious sequence to fit a consensus for either Ee70 or Ee64 (Harley & Reynolds, 1987; Kustu et al., 1989) which can identify the likely promoter, although there clearly would be space for an RNA polymerase binding site. The expression and regulation of these genes will form the basis of a separate study.

The predicted molecular masses of the gene products PntAA, PntAB and PntB are 40-3, 14-9 and 47-8 kDa, respectively. Analysis and comparison of the predicted amino acid sequences clearly reveal that, in contrast to the one and two polypeptides of bovine mitochondrial and E. coli transhydrogenases, respectively, the enzyme from R. rubrum consists of three polypeptides (Fig. 3). Thus, the α polypeptide of H+-Thase of E. coli can be viewed as a fusion of the PntAA and PntAB polypeptides of the R. rubrum enzyme (from the C-terminus of PntAA to the N-terminus of PntAB); the β polypeptide of E. coli and the PntB polypeptide of R. rubrum are equivalent. On this basis, and on the basis of the previously established relationship between the E. coli and bovine enzymes, the alignment shown in Fig. 4 indicates that there is 49.6% amino acid identity between the E. coli and R. rubrum H+-Thases and 43.4% between the R. rubrum and bovine mitochondrial proteins (not shown). There is greater sequence divergence between the R. rubrum enzyme and the E. coli and the bovine enzymes than between the E. coli and bovine enzymes (52% identity). This divergence arises mainly in the region of the PntAA polypeptide of R. rubrum.

Subunit properties predicted from the Pnt primary amino acid sequences

Hydropathic profiles of the predicted polypeptides of R. rubrum H+-Thase are shown in Fig. 5. Except for the region of the NAS-binding site (see below), PntAA generally has a rather hydrophilic character. Comparison of the N-terminal and internal sequences with the amino acid sequence predicted from the gene establishes unequivocally that this gene product is Thα. The possibility that Thβ is a proteolytic breakdown product of H+-Thase, which formerly could not be eliminated (Cunningham et al., 1992a) can now be discounted. Thus, uniquely in R. rubrum, domain I can be displaced as a discrete, water-soluble protein by washing from the chromatophore membrane. Interestingly, the region between domains I and II in H+-Thase of E. coli and bovine mitochondria is sensitive to proteolysis by trypsin (Tong et al., 1991; Yamaguchi et al., 1990) and it is possible that, in those systems under non-denaturing conditions, the resulting 43 kDa polypeptides might reconstitute transhydrogenase activity to depleted membranes.

PntAB is a small hydrophobic polypeptide. In the equivalent region (domain IIa) of the bovine and E. coli enzymes, four transmembrane helices were predicted (Yamaguchi et al., 1988; Yamaguchi & Hatefi, 1991; Tong et al., 1991). Four helices are also indicated in a hydropathy analysis of the sequence from Kramer et al. (1993) of Ei. tenella transhydrogenase (data not shown). However, the hydropathy profile shown in Fig. 5, which uses the scale of Von Heijne (1992), predicts, with reasonable confidence, that only the last three of these segments are transmembrane in R. rubrum PntAB. In an analysis using the amalgamated hydropathy scale of Kyte & Doolittle (1982), the first putative transmembrane helix (from Ala25 to Thr43) is again only weakly predicted in PntAB (data not shown). If it is assumed that the arrangement of transmembrane helices in the four transhydrogenases is similar, then, in view of the well-conserved cluster of positive charges at the C-terminus of PntAB (residues 129–139), which, by the ‘positive-inside rule’ should be located on the cytoplasmic side of the membrane (Von Heijne, 1992), and the fact that domain I is assuredly cytoplasmic (Cunningham et al., 1992a), an even number of transmembrane helices in the polypeptide is most likely. Provisionally therefore, it is suggested that the first, weakly predicted transmembrane helix of PntAB should be accepted. It cannot be excluded that the transmembrane helix composition of PntAB is different to the equivalent region in other species but, if there were only three transmembrane helices, the N-terminus of PntAB would probably be located in the periplasm, which

1598
Fig. 2. Nucleotide sequence of the pntA gene. Probable Shine-Dalgarno sequences are indicated by horizontal lines above the nucleotide sequence. Inverted repeats are indicated by pairs of arrows. The N-terminal sequences of the soluble subunit of H+-Thase and its proteolytic fragments (Cunningham et al., 1992a) are shown with solid lines under the amino acid sequence. The SD sequences of one of the proteolytic fragments, DILSSQSNLAGYRAV, obtained with endoproteinase Asp-N, has not been previously published.
seems unlikely as it does not possess a recognizable signal sequence (see Von Heijne, 1986).

PntB, homologous to the β subunit of E. coli H+-Thase, has a strongly hydrophobic domain IIb (Fig. 5), and a relatively hydrophilic domain III, which, by analogy, probably protrudes from the membrane on the cytoplasmic side (Yamaguchi et al., 1990; Yamaguchi & Hatefi, 1991; Tong et al., 1991). Using the analysis of Kyte & Doolittle (1982), eight transmembrane helices were predicted for the E. coli β subunit (Tong et al., 1991). The equivalent helices are clearly recognizable in the hydropathy profile of PntB of R. rubrum (Fig. 5). If the conclusion (above) is correct, that PntAB has four transmembrane helices, and if the assumption, that the bovine, E. coli and R. rubrum enzymes have a similar arrangement of transmembrane helices, is to be maintained, then an even number of helices for PntB (and E. coli β subunit) is indeed to be expected. It was predicted that the homologous region of bovine mitochondrial H+-Thase has ten transmembrane helices (Yamaguchi et al., 1988; Yamaguchi & Hatefi, 1991). One of these was proposed within the sequence, unique to bovine H+-Thase, that forms the link between the equivalent of R. rubrum PntAB and PntB (or the α and β subunits of the E. coli enzyme). The sequence is also absent in H+-Thase from Ei. tenella (Kramer et al., 1993) and the suggestion that it constitutes a transmembrane helix should perhaps be treated with some reservation. The other ‘extra’ putative transmembrane helix in mitochondrial H+-Thase (Yamaguchi et al., 1988; Yamaguchi & Hatefi, 1991) can, in fact, be accommodated in the hydropathic profiles of the E. coli β and the R. rubrum PntB polypeptides (e.g. residues 215–233 in the latter). However, an odd number of transmembrane helices would challenge our assumption that the arrangement is similar in all the known enzymes. The dilemma is heightened by a consideration of the charged residues in predicted loops lying between putative transmembrane helices. Thus, only one (148–164 in PntB) has a preponderance of Arg and Lys residues in H+-Thase from all four species. Taking PntB from R. rubrum in isolation, only two of the possible loops in domain IIa have net positive charge (one, adjacent to the two Lys residues at positions 76 and 77, and the other at residues 148–164). Assuming that these loops, the N-terminus and the C-terminus of PntB are all cytoplasmic, this fits comfortably with the inclusion (at residues 215–233) of the ‘extra’ transmembrane helix described by Hatefi and colleagues in the mitochondrial enzyme (see above). However, it would also require either that one of the first four predicted transmembrane helices, identified in the hydropathy profile, is in error, or that the N-terminus of the polypeptide is periplasmic rather than cytoplasmic (though it has no recognizable signal sequence). Clearly more evidence is required before firm conclusions can be drawn.

The membrane-bound component of H+-Thase from R. rubrum proved difficult to purify (Palmer et al., 1993), partly because it is a relatively minor protein in chromatophore membranes, but also because of the very stringent detergent requirement to achieve solubilization without inactivation (the ability to reconstitute transhydrogenase activity with the water-soluble subunit); possibly PntAB is easily separated from PntB by many of the commonly used detergents. In routine preparations of the membrane-bound component of R. rubrum H+-Thase, polypeptides with apparent molecular mass of 50 kDa, 52 kDa and 56 kDa were observed on SDS-PAGE (Palmer et al., 1993), and are likely candidates for PntB, but they were

---

**Fig. 3.** Comparison of the domain and polypeptide structures of H+-Thase from R. rubrum, E. coli, bovine mitochondria and Ei. tenella and with alanine dehydrogenase from Bacillus stearothermophilus. Unshaded domains are relatively hydrophilic; shaded domains are strongly hydrophobic. The dashed line linking domains III and I of the Ei. tenella protein represents a linker of 38 amino acid residues (Kramer et al., 1993).
obtained in insufficient quantities for N-terminal sequencing. Polypeptides with a molecular mass corresponding to that of PntAB were not detected on SDS-PAGE, again indicating that this component might be easily lost during purification. Although replacing the equivalent residue in E. coli by site-directed mutagenesis indicates that it is not essential (Olsson et al., 1993), it is possible that this region of Glu257 in the H+-Thase enzyme, this Glu is replaced with a Gln (PntAA, position 247) and is unlikely to be involved in catalysis. Changing the equivalent residue in Trans hydrogenase from Rhodospirillum rubrum

The nucleotide-binding sites of H+-Thase

There is presently considerable interest in the nature and location of the NAD(H)- and the NADP(H)-binding sites on H+-Thase. The b, b Satellite fold, characteristic of ADP/FAD binding sites (Wierenga et al., 1986), and thought to represent the NAD(H)-binding site in H+-Thase (Clarke et al., 1986; Yamaguchi et al., 1993), it is possible that this region of Glu257 in the H+-Thase enzyme, this Glu is replaced with a Gln (PntAA, position 247) and is unlikely to be involved in catalysis. Changing the equivalent Glu in the E. coli enzyme by site-directed mutagenesis led to the same conclusion (Glavas et al., 1993). Similarly, it has been argued (Wakabayashi & Hatefi, 1987b) that Glu257 in the mitochondrial H+-Thase, which, in the absence of NAD(H), reacts with N,N-dicyclohexylcarbodiimide, might lie at or near the NAD(H)-binding site. However, the gene sequence (Fig. 2) confirms an earlier suspicion based on the amino acid sequence of peptide fragments (Palmer et al., 1993), that, in the absence of NAD(H), it lies at or near the NAD(H)-binding site. However, the gene sequence (Fig. 2) confirms an earlier suspicion based on the amino acid sequence of peptide fragments (Palmer et al., 1993), that, in the absence of NAD(H), it lies at or near the NAD(H)-binding site. However, the gene sequence (Fig. 2) confirms an earlier suspicion based on the amino acid sequence of peptide fragments (Palmer et al., 1993), that, in the absence of NAD(H), it lies at or near the NAD(H)-binding site.
domain I, just downstream of the NAD(H)-binding site and rich in polar and charged residues (Fig. 5), corresponds to a flexible loop which is susceptible to modification by proteases and, in some conditions, by dicyclohexylcarbodiimide and [p-(fluorosulphonyl)benzoyl]-5'-adenosine. Thus, the peptide bond C-terminal to a domain I, just downstream of the NAD(H)-binding site (Yamaguchi et al., 1992a), E. coli (Tong et al., 1991) and bovine mitochondria (Yamaguchi et al., 1990). Furthermore, in PntAA a neighbouring peptide bond, Lys227-Thr228, is very sensitive to the endoprotease Lys-C (apparently not investigated in the E. coli and bovine enzymes). When the native purified soluble subunit of R. rubrum H+-Thase was treated with concentrations of trypsin low enough to give only two polypeptides on SDS-PAGE (apparent molecular mass 29 kDa and 19 kDa, thus corresponding to cleavage at Lys237-Glu238), with only minimal contamination from other fragments, its capacity to reconstitute transhydrogenase activity to depleted chromatophore membranes was severely inhibited. However, the data of Fig. 6 show that the trypsin-treated protein blocked reconstitution by untreated material. Thus, a crude preparation of soluble subunit used to titrate light-driven transhydrogenation at 30 °C. The incubation conditions indicated by the arrow were used as a basis for the experiments shown in (b). Here, additional purified soluble subunit (30 µg protein ml⁻¹), either trypsin-treated (see Methods) or untreated, was added, as shown. In control experiments, (i) addition of trypsin and trypsin-inhibitor to C₇-particles in the presence of crude soluble protein to give 90% reconstitution, but in the absence of purified protein, decreased the rate of transhydrogenation by only 7%, (ii) 50 µl of untreated, purified soluble subunit added to C₇-particles (in the absence of crude protein) gave transhydrogenation rates of 0.90 pmol (µmol bacteriochlorophyll⁻¹ min⁻¹), and (iii) 50 µl of trypsin-treated, purified soluble subunit added to C₇-particles (in the absence of pure protein) gave transhydrogenation rates of 0.37 pmol (µmol⁻¹ bacteriochlorophyll⁻¹ min⁻¹).

Information on the location of the NADP(H)-binding site of H+-Thase from R. rubrum is rather less clear, although it seems to be in domain III of the enzyme (Yamaguchi & Hatefi, 1993). The Gly residue at #314, which is replaced by a Glu in the RH1 mutant of E. coli H⁺-Thase (Ahmad et al., 1993), is conserved in the R. rubrum enzyme (PntB-Gly315). In the RH1 mutant, the influence of NADP(H) on the pattern of polypeptides generated by trypsin treatment is lost (Ahmad et al., 1993). The Tyr residue, which has been implicated in NADP(H) binding because of its reactivity in the mitochondrial enzyme with [p-(fluorosulphonyl)benzoyl]-5'-adenosine (Wakabayashi & Hatefi, 1987a) and 8-azido-AMP (Hu et al., 1992), despite the fact that, strictly, both are NAD(H) analogues, is conserved in
R. rubrum H\textsuperscript{+}-Thase (PntB-432). Note that replacement of this residue by site-directed mutagenesis shows that it is not essential (Olausson et al., 1993).

Evolutionary relationships between alanine dehydrogenase and H\textsuperscript{+}-Thase

A strong sequence similarity between the alanine dehydrogenases of two species of Bacillus and domain I of the H\textsuperscript{+}-Thases of E. coli and bovine mitochondria was previously reported (Cunningham et al., 1992a). There is 29% identity between PntAA from R. rubrum and alanine dehydrogenase from Bacillus stearothermophilus (Kuroda et al., 1990 – see Fig. 3).

It is evident that, despite the different polypeptide composition, the order of translation of the domains within the transcriptional unit of H\textsuperscript{+}-Thase is the same in the R. rubrum, E. coli and bovine systems (I > IIa > IIb > III) but is different in Ei. tenella (IIb > III > I > IIa; Kramer et al., 1993). It is likely that gene fusions and rearrangements have given rise to the different organizations. In view of the sequence similarity between domain I and alanine dehydrogenase, it is possible that the R. rubrum structure represents the most primitive of those studied so far. Thus, following suggestions for the evolution of other membrane proteins (Walker, 1992), we suppose that an R. rubrum-like H\textsuperscript{+}-Thase would have originated from an association between an ancestor of alanine dehydrogenase (becoming domain I) with an ancestor of the membrane-located domain IIa and domain IIb/III proteins. Subsequent clustering of the genes and fusion between domains I and the IIa protein would have given the E. coli organization, and further fusion between domains IIa and IIb would have given the bovine organization. A different gene clustering arrangement or a transposition of the gene segments corresponding to domain I-IIa and domain IIb-III followed by fusion would have given an Ei. tenella-like protein.

ACKNOWLEDGEMENTS

We are grateful to the Science and Engineering Research Council for financial support and to Dr John Fox and his colleagues in Alta Bioscience for assistance with the N-terminal amino acid sequencing, the synthesis of oligonucleotides and the automated DNA sequencing. The automatic DNA sequencer was purchased with a grant from the Wellcome Trust.

REFERENCES


Received 21 January 1994; revised 21 February 1994; accepted 28 February 1994.