Biochimica et Biophysica Acta, 1144 (1993) 221–224
© 1993 Elsevier Science Publishers B.V. All rights reserved 0005-2728/93/$06.00

BBABIO 40276

Rapid Report

Intimate relationships of the large and the small subunits of all nickel hydrogenases with two nuclear-encoded subunits of mitochondrial NADH : ubiquinone oxidoreductase

Simon P.J. Albracht

E.C. Slater Institute, BioCentrum, University of Amsterdam, Amsterdam (The Netherlands)

(Received 21 June 1993)

Key words: Nickel hydrogenase; Complex I; Subunit; Sequence homology; Relationship

The sequence pattern $CxxCx_nGxCxxxGx_mGCPP$, thus far found in the small subunits from 21 different nickel hydrogenases, appears also to be present in the PSST polypeptide from NADH:ubiquinone oxidoreductase (Complex I) of beef-heart mitochondria. There is only one difference: the first cysteine residue is a leucine in the PSST subunit. The large nickel-binding subunit of nickel hydrogenases shows a surprising homology with the 49 kDa subunit of mitochondrial Complex I.

All but one of the nickel hydrogenases described so far are believed to consist of at least two subunits. At the time of writing, the structural genes coding for these subunits have been sequenced for at least 24 different enzymes (*Azotobacter chroococcum* [1]; *Alcaligenes eutrophus* H16 (membrane-bound) [2]; *A. eutrophus* H16 (soluble) [3]; *Azotobacter vinelandii* [4]; *Bradyrhizobium japonicum* [5]; *Desulfovibrio baculatus* [6,7]; *Desulfovibrio fructovorans* [8]; *Desulfovibrio gigas* [7,9]; *Desulfovibrio vulgaris* Miyazaki F [10]; *D. vulgaris* (selenium enzyme) [11]; hydrogenase-1 [12], hydrogenase-2 [11] and hydrogenase-3 (formate-hydrogen-lyase) [13] from *Escherichia coli*; *Methanothermus fervidus* (F_{420}-nonreducing) [14]; the F_{420}-reducing hydrogenase [15] and the F_{420}-nonreducing hydrogenase [16] from *Methanobacterium thermoautotrophicum* 'gr – delta – uc'H; *M. thermoautotrophicum*, strain Winter [16]; the four hydrogenases from *Methanococcus voltae* (F_{420}-reducing, selenium enzyme; F_{420}-reducing; F_{420}-nonreducing, selenium enzyme; F_{420}-reducing) [17]; *Rhodobacter capsulatus* [18]; *Rhodocyclus gelatinosus* [19]; *Rhizobium leguminosarum* [20] and *Wolinella succinogenes* [21]). Predicted sequences of many of them have been compared in recent reviews [11,22,23]. The sequences that became known after these reviews had appeared did not present any surprises.

As indicated already by the first reviewers [16,22], the large subunit shows a number of conservative amino acids. In the N-terminal region there is a strictly conserved pattern $RGxEx_{16}RxCGxCxxxH$ in all but two of the 22 sequences that I could find: in the *E. coli* hydrogenase-3, which is part of the formate-hydrogen-lyase complex, the 16-amino-acid spacing is 17 amino acids long and for the *A. chroococcum* enzyme a threonine was reported at the histidine position. This conservative pattern is called here pattern 1L. Roughly in the middle of the polypeptide chain there is an Hx_6L pattern (2L) in all but one of the sequences: the *E. coli* hydrogenase-3 has an alanine instead of the leucine residue. Towards the carboxy-terminus, two further conserved sequences were reported: GxxxxPR-GxxxH (3L) and DPCxxCxxH (4L). It is now quite certain that the latter pattern is involved in the direct coordination of nickel in these enzymes. The first cysteine residue in pattern 4L was predicted to be selenocysteine in a number of enzymes [6,17,24]. Indeed, Se has been proven to be a ligand to nickel in the hydrogenase from *D. baculatus* [25,26], whereas the presence of selenocysteine in an unusually short polypeptide of the selenium-containing hydrogenase from *M. voltae* has recently been verified by mass spectrometry [27].

The small subunits in nickel hydrogenases are generally considered to host most, if not all, of the Fe-S clusters [11,23]. Also here, some definite conservative patterns can be observed. In my inspection I have included 21 different sequences. In the N-terminus there is a virtually completely conserved pattern $CxxCx_nGxCxxxGx_mGCPP$, here called pattern 1S, in all published sequence predictions, where n is between 66 and 106 and m is between 25 and 61. In most, but not

Correspondence to: S.P.J. Albracht, E.C. Slater Institute, BioCentrum, University of Amsterdam, Plantage Muidergracht 12, NL-1018 TV Amsterdam, The Netherlands. Fax: +31 20 5255124.

all, sequences there are three additional conserved groups of residues, a rather conserved pattern around a single conserved cysteine, a pattern CxxxxxCxG-PxxxxxC, and a CxxC pattern. In at least four published sequences, these last three groups are absent. The small subunit, called subunit δ, of the soluble NAD^+-reducing nickel hydrogenase from *A. eutrophus* simply lacks the peptide piece, hosting these patterns in other hydrogenases. In the F_{420}-reducing hydrogenase from *M. thermoautotrophicum*, strain 'gr – delta – uc'H and in both the F_{420}-reducing hydrogenases (normal one and the selenium-containing one) from *M. voltae*, two classical cysteine patterns $CxxCxxCx_nCP$, characteristic for the coordination of two normal $[4Fe\text{-}4S]^{2+(2+:1+)}$ clusters, are found instead.

While carrying out this inspection of the hydrogenase sequences, I found it curious to note that for all but one of the nickel hydrogenases the gene coding for the small subunit has always been recognized, usually close to the gene coding for the large subunit. Böhm et al. [13], when sequencing part of the operon for the formate-hydrogen-lyase complex of *E. coli*, did recognize the structural gene, called ORF5, coding for the large subunit of the hydrogenase activity in this complex, but failed to find a candidate for a possible small subunit (in retrospect presumably due to the limited number of known reference sequences at that time). When I inspected the published sequences by Böhm et al. [13], and having the highly conservative pattern 1S in the small subunit of other nickel hydrogenases at hand, it was not difficult to find that a gene called ORF7 in fact must be coding for the small subunit of hydrogenase-3 in *E. coli*. Böhm et al. [13] mentioned that their ORF7 showed some structural homology

PSST:	T	F	G	L	A	C	C	A	V	E
Motif 1	.	.	.	C	.	.	C	.	.	.
	G_9	L_8	E_{12}		T_{15}	G_{13}	*	S_{11}	E_{12}	S_{18}
	L_5	A_4	G_7		S_3	C_8		T_3	I_4	A
	N_3	G_3	C		C			H_3	L_2	E
	M_2	S_3	S		N			L_2	V_2	*
	C	C_2			W			E	*	
	R	Q						*		

PSST:	V	S	M	G	S	C	A	N	Q	G	G	Y	Y
Motif 2	.	.	.	G	.	C	.	.	.	G	.	.	.
	I_{16}	A_{15}	W_9	$	A_7	$	A_{17}	S_9	W_8		G_{11}	V_{14}	Q_{12}
	V_3	S_5	Y_3		T_7		S_2	V_4	Y_6		C_9	I_6	P_5
	C	$	I_2		S_6		G	A_3	S_3		N	F	T_3
	F		V_2		$		$	T_2	F_2				F
			F_2					N	T				
			L_2					Q	$				
			$					$					

PSST:	Y	V	P	G	C	P	P	T	A	E
Motif 3	.	.	.	G	C	P	P	.	.	.
	K_7	V_{11}	P_{18}					I_8	P_{10}	E_{12}
	N_5	I_8	A_2					S_4	A_7	D_4
	A_3	F	Q					N_3	E	A
	E	L						R_2	G	I
	F							D	S	K
	I							H	T	M
	P							K		Y
	Y							T		
	V									

Fig. 1. Comparison of the conservative pattern 1S of the small subunits of 21 different nickel hydrogenases with the sequence of the PSST subunit of mitochondrial Complex I. The dots in the middle line of the motifs 1–3, that make up the pattern 1S, indicate non-conservative amino acids in the small subunits of hydrogenase. The rows below this line show how frequently a residue occurs at that position. An asterisk (*) or a dollar ($) sign points to possible sequence errors encountered in published sequences of two different enzymes. The upper line shows the relevant sequence stretches of the PSST polypeptide.

with the predicted sequence of protein G from chloroplasts of *Zea mais* and *Marchantia polymorpha*. They had already noted, however, that according to Nixon et al. [28] the *psbG* gene is rather a *ndh* gene (the name *ndh* stems from NADH dehydrogenase) for a putative enzyme in chloroplasts similar to the mitochondrial Complex I. Further evidence for this has been recently provided by Arizmendi et al. [29], who sequenced a 20 kDa polypeptide, named PSST after the first four amino acids. They clearly showed the homology to the product of the former *psbG* gene. They renamed the latter gene accordingly as *ndh*K. A summary of further homologies of the PSST gene can be found in a recent review [30]. Not surprisingly, Fearnley and Walker [30] could confirm the sequence relationships with the ORF7 gene of the *E. coli* formate-hydrogen-lyase system. Once I realized the possible implications of these links, it was only logical to compare the sequence of the PSST gene [29] with those of the small subunits of nickel hydrogenases. The result is summarized in Fig. 1.

The PSST polypeptide contains nearly all of the motifs of the highly conservative pattern 1S of the small subunit of nickel hydrogenases: $LxxCx_{61}GxCxx$-$xGx_{25}GCPP$. The only difference is that the first conservative cysteine residue in the 1S pattern of nickel hydrogenases is replaced by a leucine residue in the PSST subunit. The polypeptide chain of PSST stops some distance after the described pattern, just as is the case in the small subunits of the soluble hydrogenase of *A. eutrophus* and the *E. coli* hydrogenase-3. For me, working on Complex I since the early 1970's and exploring nickel hydrogenases since the late 1970's, this was a most amazing and also an amusing result. It indicates that one of the subunits of mitochondrial Complex I owes its structural information to the strictly conservative part present in the small subunit of all nickel hydrogenases.

As discovered by Pilkington et al. [31], the sequences of the subunits responsible for the diaphorase activity of the soluble NAD^+-reducing nickel hydrogenase of *A. eutrophus* [3] have much in common with some subunits of mitochondrial Complex I. The α subunit appeared to be a fusion product of the 24 kDa subunit and the 51 kDa subunit of Complex I. The 'gr – gamma – lc' subunit contains the sequence of the first 200 amino acids of the 75 kDa subunit of Complex I. It can now be added that the δ subunit of the hydrogenase part of the *Alcaligenes* enzyme is in fact very homologous to the PSST subunit of Complex I. Hence, four of the subunits of Complex I carry a considerable amount of structural information also found in three of the four subunits of this hydrogenase. For the NAD^+-reducing hydrogenase from *Alcaligenes*, this means that in fact only the β subunit appears unique for this enzyme.

Even in the large subunit of nickel hydrogenases, however, there appear to be some surprising correlations with Complex I. In an alignment of the sequence of the 49 kDa subunit of mitochondrial Complex I, Fearnley and Walker [30] noted the relationship with the ORF5 gene of the *E. coli* formate-hydrogen-lyase system. Böck and coworkers had already indicated in 1990 [13] that this gene most likely is the structural gene for the large subunit of the nickel hydrogenase of this complex, and comparison with the sequences of other nickel hydrogenases [11,23] firmly establishes this prediction. Also, the finding reported here about the strong relationship of ORF7 with the small subunit of nickel hydrogenases is completely in line with this proposal. It thus appears that the 49 kDa subunit of mitochondrial complex I shows quite some homology with the large subunit of nickel hydrogenases. Upon inspecting the sequence of the 49 kDa subunit, one can see that the subunit contains the sequence $RGxEx_{16}R$ from the pattern 1L, but the following CGxCxxxH residues from pattern 1L are not present. To my surprise, pattern 2L, Hx_6L, is also present in the 49 kDa subunit, but this is probably a coincidence, as it does not line up with the conservative histidine residue in the sequence of the large subunit of *E. coli* hydrogenase-3 (concluded from the sequence alignment published in Ref. 30). Patterns 3L and 4L are absent.

All the sequence comparisons in this report were carried out by hand. Submitochondrial particles did not show any trace of activity in our standard hydrogenase assays.

The findings reported here can be of help in the further elucidation of the function of the relevant subunits both in Complex I and in nickel hydrogenases.

I am indebted to Mrs. M. van Veenhuizen for the hydrogenase activity measurements, to Mr. O. Sorgenfrei for the print-out of the *M. voltae* sequences and to Prof. A. Böck for critically reading the manuscript.

References

1 Ford, C.M., Garg, N., Garg, R.P., Tibelius, K.H., Yates, M.G., Arp, D.J. and Seefeldt, L.C. (1990) Mol. Microbiol. 4, 999–1008.
2 Kortlücke, C., Horstmann, K., Schwartz, E., Rohde, M., Binsack, R. and Friedrich, B. (1992) J. Bacteriol. 174, 6277–6289.
3 Tran-Betcke, A., Warnecke, U., Böcker, C., Zaborosch, C. and Friedrich, B. (1990) J. Bacteriol. 172, 2920–2929.
4 Menon, A.L., Stults, L.W., Robson, R.L. and Mortenson, L.E. (1990) Gene 96, 67–74.
5 Sayavedra-Soto, L.A., Powell, G.K., Evans, H.J. and Morris, R.O. (1988) Proc. Natl. Acad. Sci. USA 85, 8395–8399.
6 Menon, N.K., Peck, H.D., Jr., LeGall, J. and Przybyla, A.E. (1987) J. Bacteriol. 169, 5401–5407.
7 Voordouw, G., Menon, N.K., LeGall, J., Choi, E.-S., Peck, H.D., Jr., and Przybyla, A.E. (1989) J. Bacteriol. 171, 2894–2899.
8 Rousset, M., Dermoun, Z., Hatchikian, C.E. and Bélaich, J.-P. (1990) Gene 94, 95–101.
9 Li, C., Peck, H.D., Jr., LeGall, J. and Przybyla, A. (1987) DNA 6, 539–551.
10 Deckers, H.M., Wilson, F.R. and Voordouw, G. (1990) J. Gen. Microbiol. 136, 2021–2028.

11 Przybyla, A.E., Robbins, J., Menon, N., & Peck, H.D., Jr. (1992) FEMS Microbiol. Rev. 88, 109–136.
12 Menon, N.K., Robbins, J., Peck, H.D., Jr., Chatelus, C.Y., Choi, E.-S. and Przybyla, A.E. (1990) J. Bacteriol. 172, 1969–1977.
13 Böhm, R., Sauter, M. and Böck, A. (1990) Mol. Microbiol. 4, 231–243.
14 Steigerwald, V.J., Beckler, G.S. and Reeve, J.N. (1990) J. Bacteriol. 172, 4715–4718.
15 Alex, L.A., Reeve, J.N., Orme-Johnson, W.H. and Walsh, C.T. (1990) Biochemistry 29, 7237–7244.
16 Reeve, J.N., Beckler, G.S., Cram, D.S., Hamilton, P.T., Brown, J.W., Krzycki, J.A. Kolodziej, A.F., Alex, L., Orme-Johnson, W.H. and Walsh, C.T. (1989) Proc. Natl. Acad. Sci. USA 86, 3031–3035.
17 Halboth, S. and Klein, A. (1992) Mol. Gen. Genet. 233, 217–224.
18 Leclerc, M., Colbeau, A., Cauvin, B. and Vignais, P.M. (1988) Mol. Gen. Genet. 214, 97–107.
19 Uffen, R.L., Colbeau, A., Richaud, P. and Vignais, P.M. (1990) Mol. Gen. Genet. 221, 49–58.
20 Hidalgo, E., Leyva, A. and Ruiz-Argüeso, T., J. (1990) Plant Mol. Biol. 15, 367–370.
21 Dross, F., Geisler, V., Lenger, R., Theis, F., Krafft, T., Fahrenholz, F., Kojro, E., Duchêne, Tripier, D., Juvenal, K. and Kröger, A. (1992) Eur. J. Biochem. 206, 93–102.
22 Reeve, J.N. and Beckler, G.S. (1990) FEMS Microbiol. Rev. 87, 419–424.
23 Voordouw, G. (1992) in Advances in Inorganic Chemistry (Cammack, R. and Sykes, A.G., eds.), Vol 38, pp. 397–422, Academic Press, New York.
24 Böck, A., Forchhammer, K., Heider, J., Leinfelder, W., Sawers, G., Veprek, B. and Zinoni, F. (1991) Mol. Microbiol. 5, 515–520.
25 He, S.H., Teixeira, M., LeGall, J., Patil, D.S., Moura, I., Moura, J.J.G., DerVartanian, D.V., Huynh, B.H. and Peck, H.D., Jr. (1989) J. Biol. Chem. 264, 2678–2682.
26 Eidsness, M.K., Scott, R.A., Prickril, B.C., DerVartanian, D.V., LeGall, J., Moura, I., Moura, J.J.G. and Peck, H.D., Jr. (1989) Proc. Natl. Acad. Sci. USA 86, 147–151.
27 Sorgenfrei, O., Linder, D., Karas, M. and Klein, A. (1993) Eur. J. Biochem. 213, 1355–1358.
28 Nixon, P.J., Gounaris, K., Coomber, S.A., Hunter, C.N., Dyer, T.A. and Barber, J. (1989) J. Biol. Chem. 264, 14129–14135.
29 Arizmendi, J.M., Runswick, M.J., Skehel, J.M. and Walker, J.E. (1992) FEBS Lett. 301, 237–242.
30 Fearnley, I.M. and Walker, J.E. (1992) Biochim. Biophys. Acta, 1140, 105–134.
31 Pilkington, S.J., Skehel, J.M., Gennis, R.B. and Walker, J.E. (1991) Biochemistry 30, 2166–2175.