Comparison of the subunit sequences from the *Acidithiobacillus ferrooxidans* [NiFe]-hydrogenase with those of some cyanobacterial [NiFe]-hydrogenases, H₂-sensor proteins and standard [NiFe]-hydrogenases

Below, some results from the sequence comparisons are reported that were not, or only briefly, described in the main text.

The HoxG (large) subunit

In addition to the extremely high similarity of the HoxG sequence to those of cyanobacterial [NiFe]-hydrogenases described in the main text (the Af group), there was a rather high similarity with two additional groups within the 60 sequences most similar to HoxG. The first group comprised 14 sequences (here termed the ‘sensor group’) and had expectation values (E values) between 3·10⁻⁶⁹ and 1·10⁻⁵⁴. It contained the large subunits of the well characterized H₂-sensor protein of *Ralstonia eutropha* [1-5] and that of *Rhodobacter capsulatus* (HupV [6-8]), along with proteins (some annotated as HupV) from other bacteria. A close inspection of a separate alignment of this set of 14 sequences showed beyond doubt that all belong to the H₂-sensor family. One characteristic of this group is e.g. that the C-terminal extension on the large subunit as present in standard [NiFe]-hydrogenases (see Fig. 1S), which is removed during maturation, [9, 10], is lacking in H₂-sensor proteins [1, 3, 6, 11, 12].

The second group (‘residual group’), being the remaining sequences in the alignment of the 60 sequences, had E values of 9·10⁻⁵¹ and higher. Closest were some Se-containing [NiFe]-hydrogenases, a.o. the well characterized enzyme from *Desulfomicrobium baculatum* [13], as well as some standard [NiFe]-hydrogenases, a.o. the periplasmic enzyme *Desulfovibrio fructosovorans* [14]. A little further down the list were F₄₂₀-non reducing hydrogenases from *Methanosarcina Barkeri* and *Methanosarcina mazei*. To illustrate the similarities and differences between the Af group, the sensor group and standard [NiFe]-hydrogenases, we have depicted an alignment of some typical members of each group in Fig. 1S.

In view of the fact that, like the 'Knallgas bacteria', Af can grow on H₂ and O₂ we have inspected the sequence of the HoxG subunit for motifs that other workers have identified as possibly responsible
for the O₂-insensitivity of H₂-sensor proteins. Volbeda et al. [15] concluded from the known structures of five [NiFe]-hydrogenases (D. baculatum, Desulfovibrio desulfuricans, Desulfovibrio vulgaris Miyazaki, Desulfovibrio gigas and D. fructosovorans) that from the amino acids lining the putative principal gas channel [16], five of the six amino acids that are closest to the Ni ion in the active site are conserved in these structures (Val74, Leu122, Arg476 in the large subunit of the D. fructosovorans enzyme and Thr18, Thr47 in the small subunit). Except for the Arg residue, all these residues are different in the sequences of the H₂-sensor proteins of R. eutropha and R. capsulatus. Noticeably, two of the residues, Val74 and Leu122 in the D. fructosovorans protein, are replaced by the bulkier residues Ile and Phe, respectively, in the H₂ sensors [15]. It was shown that the calculated gas access to the active site of the D. fructosovorans enzyme was significantly decreased when both residues were substituted in the crystal structure. Hence, it was suggested that the O₂ insensitivity of these sensors may be explained by partial blocking of the gas access to the active site.

This hypothesis was experimentally tested by Buhrke et al. [17], who substituted these bulky residues by smaller ones in the large subunit (HoxC) of the R. eutropha H₂-sensor protein. Aerobic purification of the single mutants I62V and F110L resulted in enzyme with an activity of 4% and 18%, respectively, of that of the wild-type enzyme. The double-mutant protein (I62V + F110L) was completely inactive, but could be reactivated upon reduction with dithionite. The anaerobically-purified enzymes were active (52%, 71% and 47% of the wild-type activity for the I62V, F110L and I62V + F110L mutant enzymes, respectively). Exposure to air resulted in a decrease of activity of about 30%, 30% and 60%, respectively, within 40 min, whereas the activity of the wild-type enzyme was not affected [17]. Similar experiments were also reported for the R. capsulatus H₂ sensor [8].

Inspection of the 60 aligned sequences most similar to HoxG showed that all members of the Af group and sensor group have a PRxCGIC motif (positions 80-86 in Fig. 1S), where the underlined residues are strictly conserved in all [NiFe]-hydrogenases and the italic residue is the Ile residue discussed above. The Ile residue is a Val residue in standard [NiFe]-hydrogenases. Hence, the presence of the Ile residue in the 14 H₂-sensor proteins is in agreement with Volbeda's proposal and Buhrke's experiments that this residues may be of importance for the O₂-resistance of the sensors. The presence of this conserved residue in the Af group is in agreement with the observation that the Af hydrogenase in cells functions in the presence of oxygen. It has been reported that the uptake hydrogenase from Anabaena sp. 7120, another member of the Af group, can be assayed in cells via the oxidation of H₂ by O₂ catalyzed by hydrogenase and oxidases in the heterocysts [18], confirming earlier reports on the tolerance of the uptake hydrogenase to oxygen [19]. In Anabaena the enzyme is apparently bound to the heterocyst membrane [18]. However, once the enzyme was solubilized, it was irreversibly inactivated by O₂ [20].

The replacement of Leu (Leu122 in the D. fructosovorans enzyme) by Phe (Phe110 in the R. eutropha sensor) appeared to be specific for the sensor group, which had a HFxxFFMxD motif (the underlined Asp residue is conserved in all [NiFe]-hydrogenases; the italic Phe residue is the one under
The Af group had a xFYxLxxD motif (positions 129-137 in Fig. 1S; the Phe residue in this motif was replaced by a His residue only in the Af sequence), while most members of the residual group had HFYHLxxxAxD. Hence, the presence of the Phe residue (italic in the HFxxFFMxxD motif) is specific for the H2-sensor proteins and this may be the most important reason for their O2 resistance.

The HoxK (small) subunit

In addition to the group of highly similar cyanobacterial small subunits (Af group), two more groups were recognized within the 60 sequences most similar to the HoxK subunit. One group (sensor group; E values of 3·10^{-71} to 2·10^{-54}) consisted of 16 sequences, e.g. from the H2 sensors from R. eutropha and R. capsulatus. Again, two sequences assigned to the 20 kDa subunit of NADH:ubiquinone were amongst them. This must be in error because a separate CLUSTAL-X alignment of the 16 sequences showed beyond doubt that all of them belonged to the small subunit of H2 sensors, e.g. all showed a C-terminal extension with a conserved Pro residue (see Fig. 2S), which is characteristic for H2-sensor proteins [12].

The other group (residual group, E values of 7·10^{-44} and higher) contained the small subunits of standard hydrogenases from D. gigas, D. fructosovorans, D. vulgaris Miyazaki and Thiocapsa roseopersicina. Also in this case four sequences ascribed to the 20 kDa subunit of NADH:ubiquinone oxidoreductase appeared, but also here this annotation cannot be correct. Fig. 2S depicts the similarities and differences between some typical members of the Af group, the sensor group and standard [NiFe]-hydrogenases.

Iron-sulphur clusters in the small subunit

The crystal structures of five standard [NiFe]-hydrogenases have established the Cys ligands of the three Fe-S clusters in the small subunit, the so-called proximal [4Fe-4S] cluster, the distal [4Fe-4S] cluster and the medial [3Fe-4S] cluster. With this information at hand, the following can be predicted for the inspected HoxK-like sequences. The first Cys ligand (C1) in the sequence motif binding the proximal cluster was conserved in all 60 sequences but one (position 67 in Fig. 2S). As noticed before [21] all [NiFe]-hydrogenases known until the end of the 1990's contain the four Cys residues that are ligands to the proximal [4Fe-4S] cluster. In agreement with this information, the second Cys ligand (C2) was conserved in the sensor group and all but eight of the sequences of the residual group (position 70 in Fig. 2S). In the Af group an Asn residue was found at the position of C2, but an extra Cys residue, conserved in most of the members of this group, occurred 13 residues after the Asn residue (at position 83 in Fig. 2S). The third (position 170 in Fig. 2S; note that C3 shifted 4 places to the N-terminus in the D. fructosovorans sequence) and fourth (positions 218 in Fig. 2S) Cys residues, typical for binding of the proximal cluster, were found in all (but one) of the 60 sequences. This
suggests that all enzymes in the Af and sensor groups have a proximal cluster, like that of standard [NiFe]-hydrogenases, but that within the Af group the position in the sequence of the second co-ordinating Cys residue is different.

The first ligand in the sequence motif binding the distal cluster in standard [NiFe]-hydrogenases is a His residue. This residue was found to be conserved in the sensor and residual groups (position 257 in Fig. 2S). However, all members of the Af group had a Gln residue at this position and there were no nearby conserved His or Cys residues. The C2 (position 260 in Fig. 2S), C3 (Fig. 2S, position 284 in Af and sensor groups; position 285 in standard hydrogenases) and C4 (Fig. 2S, position 291) residues typical for binding of the distal cluster were found in all 60 sequences. This finding suggests that the enzymes in the Af group may contain an unusual distal [4Fe-4S] cluster or that they may not posses such a cluster.

The first ligand in the sequence motif of standard [NiFe]-hydrogenases binding the medial [3Fe-4S] cluster is a Pro residue. This was found in sequences of all members of the Af and residual groups (Fig. 2S, position 294). In the sensor group the Pro residue was replaced by a Thr residue. The C2 residue, binding the medial cluster, was conserved in all 60 sequences (Fig. 2S, position 300). The sensor group had an additional Cys residue, conserved within this group, 11 residues after C2 (Fig. 2, position 312). The C3 (Fig. 2S, position 319) and C4 (Fig. 2S, position 322) ligands binding the medial cluster were found in all 60 sequences. This suggests that the H₂-sensor proteins do not have a [3Fe-4S] cluster; the precise nature of the Fe-S clusters in the best studied representative of these proteins, the regulatory hydrogenase from R. eutropha, is not yet known [22].

This survey suggests that the small subunit of the Af hydrogenase probably harbours a proximal [4Fe-4S] cluster as well as a medial [3Fe-4S] cluster, but that it may lack a distal [4Fe-4S] cluster (or contains an unusual one). The absence of a distal cluster in the purified enzyme would be in approximate agreement with the previously determined metal content of 8.8 Fe per Ni.

Legends to figures

Fig. 1S Comparison of the sequence of the large subunit (HoxG) from the Acidithiobacillus ferrooxidans hydrogenase with those of the large subunits from some typical cyanobacterial [NiFe]-hydrogenases (a-e), some H₂-sensor proteins (f-j) and some standard [NiFe]-hydrogenases (k-p). The sequences were aligned with CLUSTAL-X. Residues marked with an asterisk are identical in all sequences. Stretches discussed in the main or supplementary text are underlined with a red bar.

Abbreviations: A.ferrooxidans_Large, large subunit of the [NiFe]-hydrogenase from A. ferrooxidans; a._C.watsonii, Crocosphaera watsonii; b._N.punctiforme, Nostoc punctiforme; c._A.variabilis, Anabaena variabilis ATCC 29413; d._Gloeothece_sp., Gloeothece sp. PCC 6909; e._L.majuscula, Lyngbya majuscula CCAP 1446/4; f._R.capsulatus, Rhodobacter capsulatus; g._A.caulinodans, Azorhizobium caulindans; h._T.roseopersicina, Thiocapsa roseopersicina; i._R.eutropha, Ralstonia eutropha; j._B.japonicum, Bradyrhizobium japonicum; k._D.gigas, Desulfovibrio gigas;
1. _D.vulgaris_M, *Desulfovibrio vulgaris* (strain Miyazaki); m. _D.fructosovorans,* Desulfovibrio fructosovorans* (periplasmic enzyme); n. _A.vinosum,* Allochromatium vinosum (membrane-bound enzyme); o. _T.roseopersicina,* Thiocapsa roseopersicina (membrane-bound enzyme); p. _R.eutropha,* Ralstonia eutropha (membrane-bound enzyme).

**Fig. 2S** Comparison of the sequence of the small subunit (HoxK) from the *Af* hydrogenase with those of the small subunits from some typical cyanobacterial [NiFe]-hydrogenases (a-e), some H₂-sensor proteins (f-j) and some standard [NiFe]-hydrogenases (k-p). Markings and abbreviations are as in Fig. 1S. For the *A. vinosum* enzyme there is no information on a possible leader sequence of the small subunit [23].

**References**


Fig. 2S (Schröder et al.)