Supplemental Data

Location of CYP genes

N. aromaticivorans contains two plasmids, pNL1 (180 kbp) and pNL2 (480 kbp), alongside its chromosomal genomic DNA (3.56 Mbp). CYP101C1, CYP101D1, CYP101D2, CYP111A2, ArR and Arx are located on the chromosomal DNA while CYP101B1 is located on pNL2. No *CYP* genes are located on pNL1 (1).

CYP101B1	Saro_3533	pNL2
CYP101C1	Saro_2249	chromosomal
CYP101D1	Saro_0514	chromosomal
CYP101D2	Saro_1478	chromosomal
CYP111A2	Saro_0759	chromosomal
ArR	Saro_0216	chromosomal
Arx	Saro_1477	chromosomal

The organisation of the genes in *Sphingomonas* bacteria are complex in that the genes necessary for degradation of one type of compound may be distributed into multiple operons that also possess genes for the degradation of other compounds. The lateral transfer of genes between *Sphingomonas* and *Pseudomonas* species has previously been proposed (1). Full details on the cloning of the genes have been reported elsewhere (2,3).

Product formation

After all of the NADH had been consumed in an incubation reaction, 10 µl of a 25 mM stock of 9-fluorenol was added to the mixture to act as an internal standard for the extraction and GC analysis. 1 ml of the turnover was extracted with 400 µl ethyl acetate and the organic phase was removed and analysed via gas chromatography. Analyses were performed on a ThermoFinnegan Trace GC instrument equipped with an auto-sampler and a DB-1 fused silica column (7 m × 0.25 mm) using helium as the carrier gas and flame ionization detection. The injector and flame

ionization detector were held at 200°C and 250°C, respectively. The temperature of the DB-1 fused silica column (7m x 0.25m) was held at 60 ºC for 1 min and then increased at 15 ºC/min up to 150 ºC. The retention times of substrate and products were as follows: camphor; 3.78 min, 5-*exo*-hydroxycamphor; 5.74 min, 6-*exo*-hydroxycamphor; 5.50 min, 5-*endo*-hydroxycamphor; 5.96 min, 9-hydroxycamphor; 6.14 min, β-ionone; 6.65 min, 4-hydroxy-β-ionone; 8.39 min, 3-hydroxy-β-ionone; 8.71 min, 4-(1,3,3-trimethyl-7-oxabicyclo[4.1.0]hept-2-yl)-3-buten-2-one (α-ionone-epoxide); 7.05 min, linalool; 3.38 min, 8-hydroxylinalool; 5.67 min, 2-adamantanone; 5.30 min, 5-hydroxy-2-adamantanone; 6.52 min, unidentified minor 2-adamantanone oxidation product; 6.62 min and 9-hydroxyflourenol; 9.35 min.

The product concentration in incubation mixtures was calculated by calibrating the concentration response of the GC detector to an authentic sample or a structural isomer (e.g. 2-hydroxy-3-pinanone for 5-*exo*-hydroxycamphor). Mixtures containing different concentrations of the product and the internal standard were extracted and analyzed as for normal incubations. The plot of the ratio of the peak area for the product to the 9-fluorenol internal standard against the product concentration gave a calibration plot from which the absolute concentration of products produced by enzymatic turnover in an incubation mixture could be determined. Where more than one product was produced the detector response was assumed to be equal, e. g. 4-hydroxy-β-ionone and 3-hydroxy-β-ionone. The coupling efficiency was the percentage of NADH consumed that led to product formation.

The NAD-binding site in ArR

In ArR the solvent accessible cavity above the *re* side of the isoalloxazine ring is believed to accommodate the nicotinamide ring of NAD (Fig. S6). The structure of BphA4, a related ONFR protein, with NAD bound has been determined. The position of the loops that constitute the NAD binding site in both proteins are similar (Fig. S6). The conserved acidic residues Glu165 and Glu301 form hydrogen bonds with the carboxamide group of the nicotinamide ring and ribose (δ-ring) oxygen, respectively. Another conserved acidic residue Glu181 may stabilize the adenine ribose through hydrogen bonds with the ribose O2 and O3 atoms. In BphA4, residues Gln177 and Ser182 also form hydrogen bonds with these adenine ribose oxygens. However, the corresponding

residues in ArR are nonpolar Leu183 and Ala188 and thus NAD binding might be weakened. Another potentially significant difference between ArR and BphA4 is the orientation of Arg189 (Arg183 in BphA4). Arg183 in BphA4 may act a wall in the channel of NAD binding site and stabilize the ribose and pyrophosphate group but Arg189 in ArR is located at the surface of the NAD binding site and may form hydrogen bonds with the adenine ribose.

The C-terminal arm of Arx

The hydrogen-bonding network at the C-terminus of Arx is different from other vertebrate-type ferredoxins due to Arg8 which forms hydrogen bonds with the carbonyl and side chain oxygens of Glu34, the carbonyl oxygen of Ala101, and two water molecules (Wat20 and Wat11), as well as a salt bridge with Glu103. The water molecules bridge Arg8 and the carbonyl oxygens of Thr47 and His49 and the main chain amide nitrogen atoms of Leu36 and His49 (Supplemental Fig. S10c). This network of hydrogen bonds and salt bridges around Arg8 links the N-terminus, C-terminus and the cluster loop of Arx. The corresponding residue to Arg8 in other ferredoxins is a histidine whose side-chain points towards the enzyme surface.

The overall structure of CYP101D1

In the native form, one 1,4-dioxane molecule, six phosphate ions, 466 water molecules, and one unidentified density were found. In the camphor-bound form, two camphor molecules, 817 water molecules and four phosphate ions were identified.

Figure S1. SDS page gels of ArR, Arx and CYP101D1.

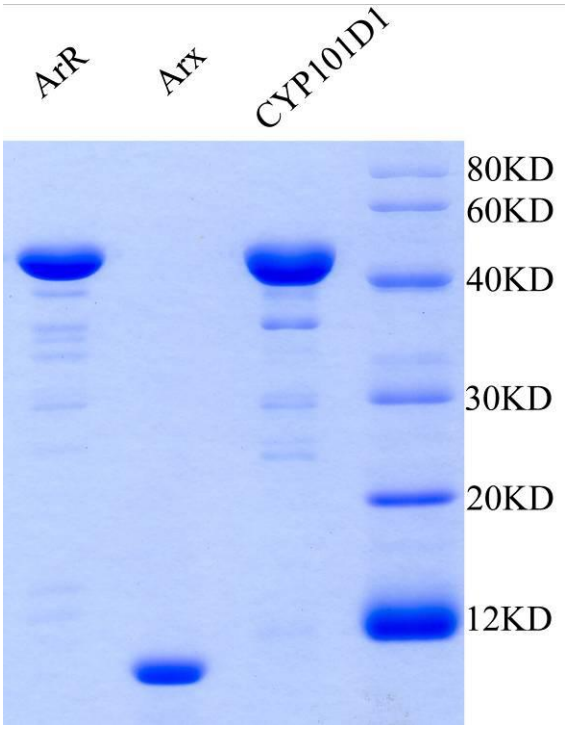
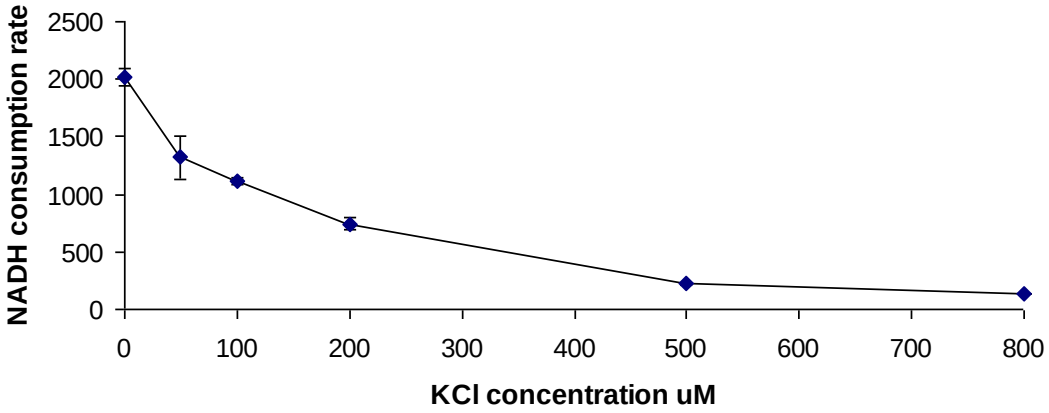


Figure S2. (a) The effect of KCl on the NADH consumption rate of the ArR/Arx/CYP101D2 system (b) The effect of pH on the NADH consumption rate of the ArR/Arx/CYP101D2 system. The coupling of product formation to NADH consumption (≥ 95%) was unaffected by the presence of ≤ 500 mM KCl or the pH values studied. The optimal activity was found at pH 7.4.



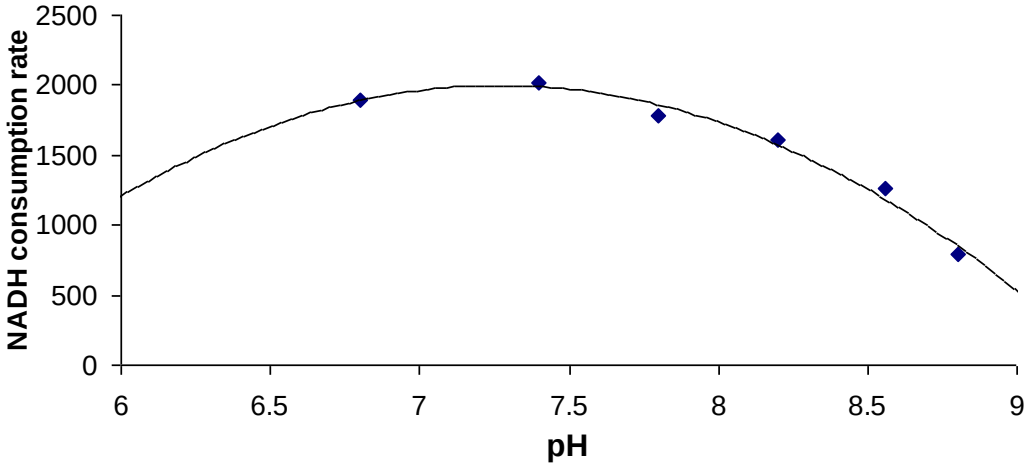
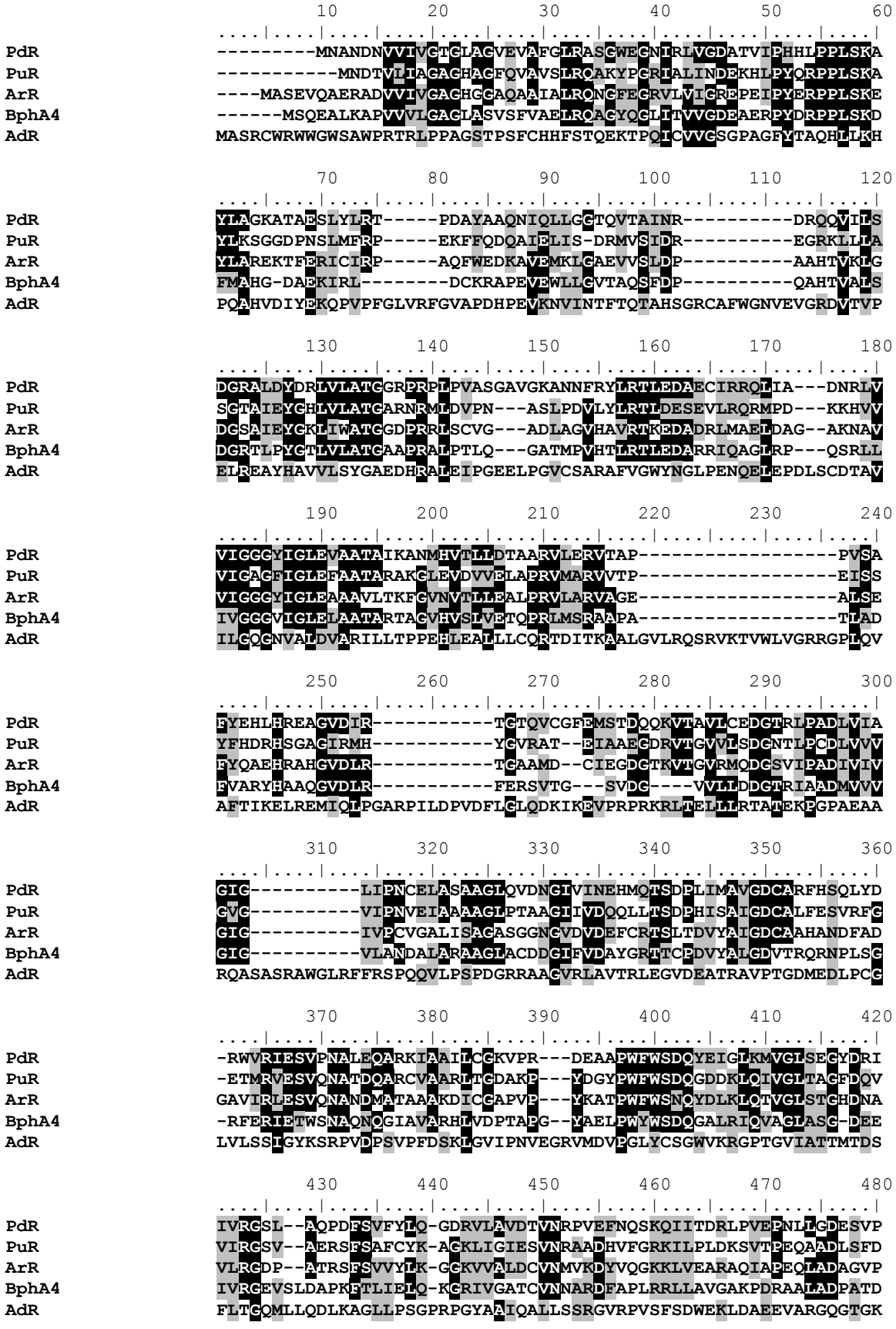


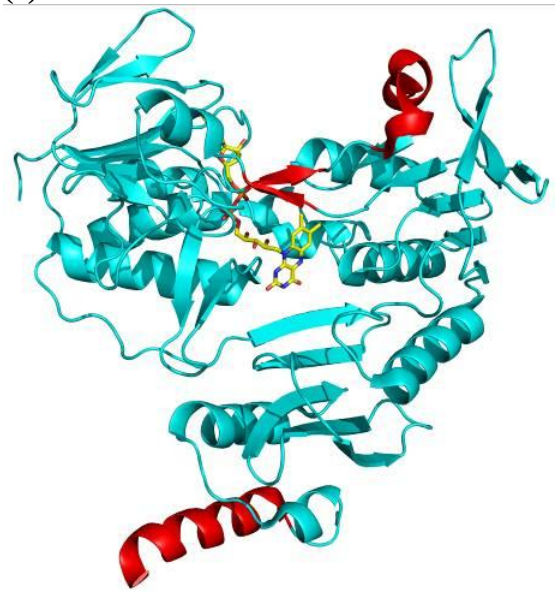
Figure S3. Protein sequence alignments of the ONFR proteins PdR, PuR, ArR, BphA4 and AdR.



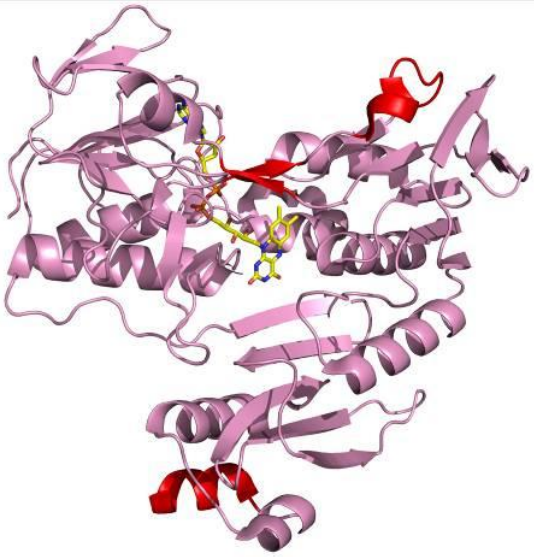
Sequence alignments were performed using ClustalW (http://www.ebi.ac.uk/clustalw/) and annotated using BioEdit (4,5).

Figure S4 a-c) Overall structure comparison of PdR (a), BphA4 (b) and PuR (c) colored in cyan, pink and orange, respectively. The FAD molecules are shown in yellow stick representation. Differences in the secondary structures are colored in red.

(a)



(b)



(c)

Figure S5. The surface of PdR (a) and PuR (b) at the top of the adenine ribose moiety, showing the more open conformations of these enzymes compared to ArR (Fig. 2c).

(a) (b)

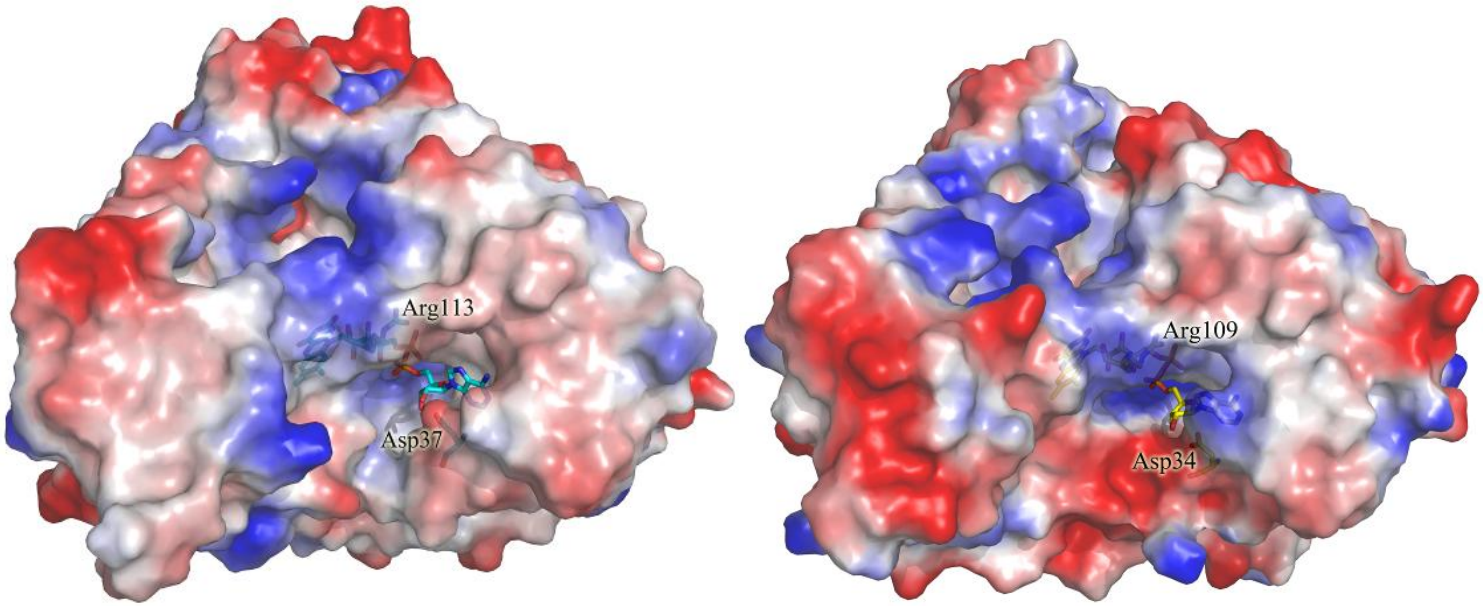


Figure S6. Overlay of the NAD-binding domain of the NAD^+ (in yellow) complex of BphA4 (in cyan) with ArR (in orange) without NAD^+ bound. The structurally conserved loops (I, II and III) involved in cofactor binding and the location of potential key residues such as Glu165, Glu181 and Glu301 are shown.

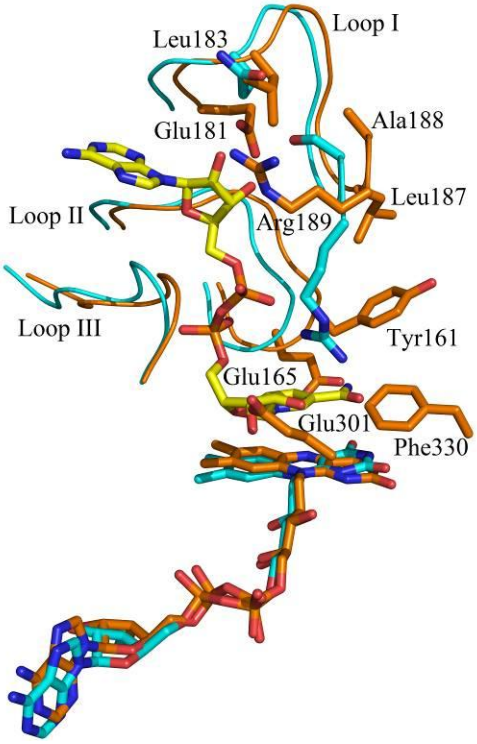
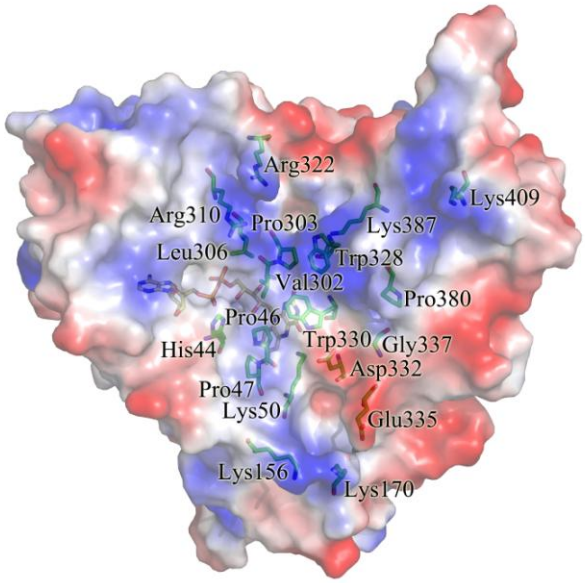


Figure S7. The electrostatic potentials of the FAD *si* side surface of PdR (a) and PuR (b) compared to ArR (Fig. 3a). Negatively and positively charged surface areas are colored in red and blue, respectively. Residues that contribute to the different surface potential distributions and that may be involved in ferredoxin binding are labeled.

(a)



(b)

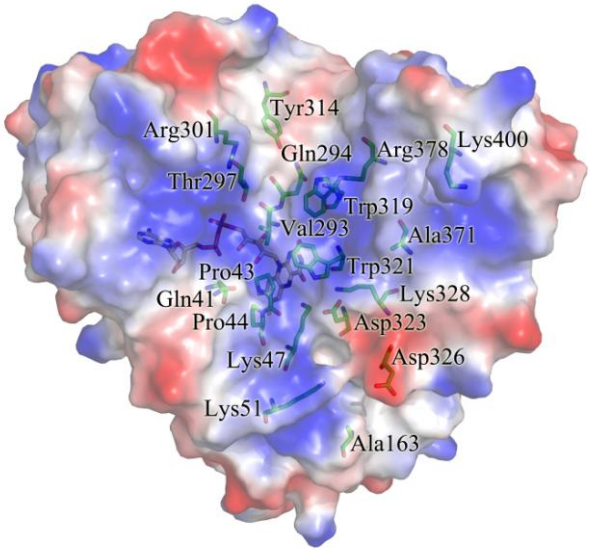
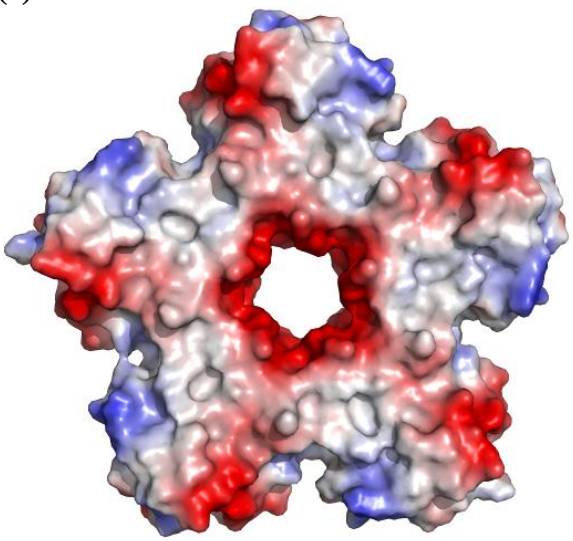


Figure S8. (a) and (b) The electrostatic potential surfaces of the Arx pentamers. Positively charged regions are colored in blue and negatively charged regions in red. The more neutral surface is the interaction region in the back-to-back packing of pentamers in the crystals. The intermolecular disulfide bonds are also located within this neutral region.

(a)



(b)

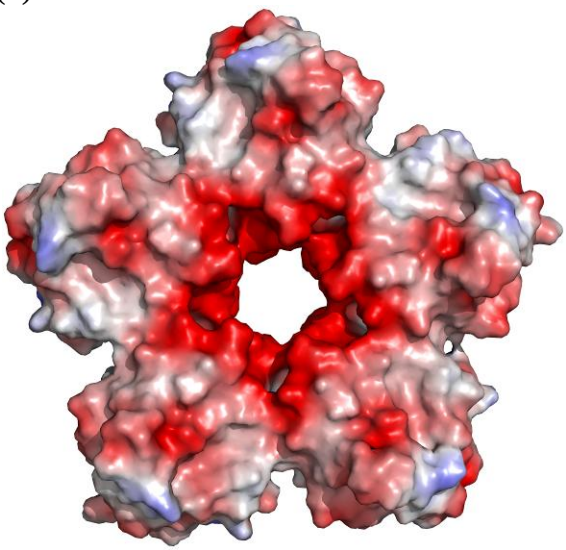


Figure S9. Gel filtration of the Arx protein (Superdex-200)

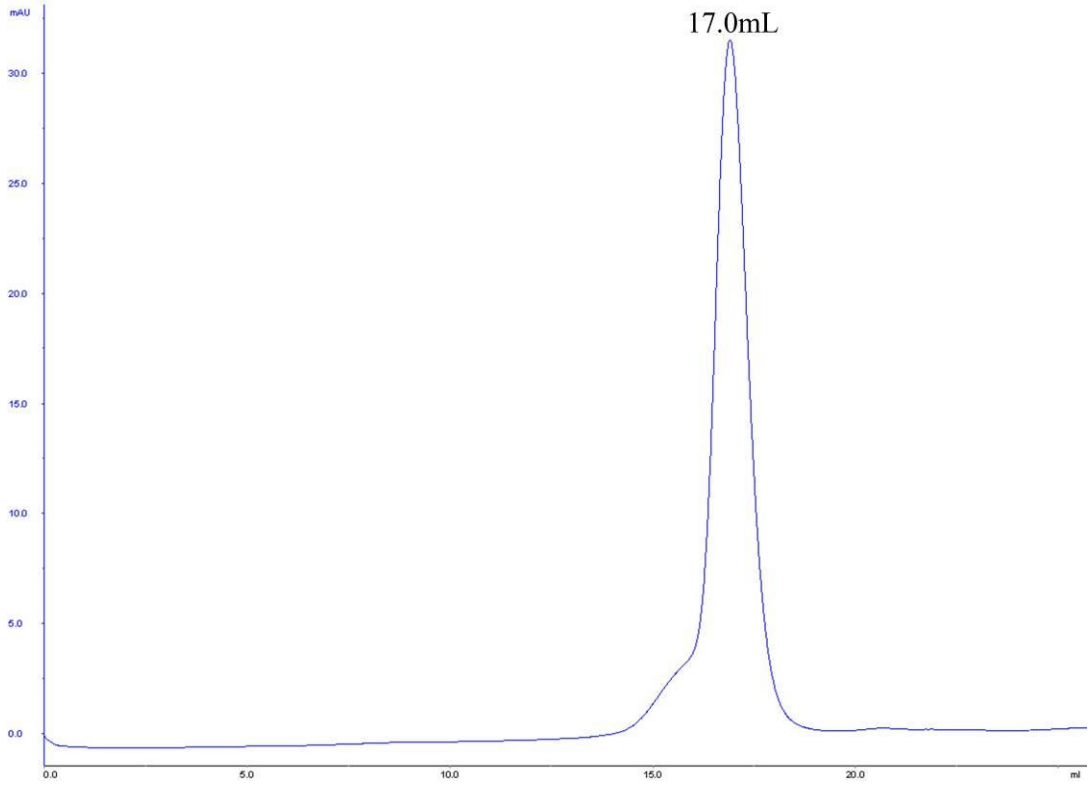
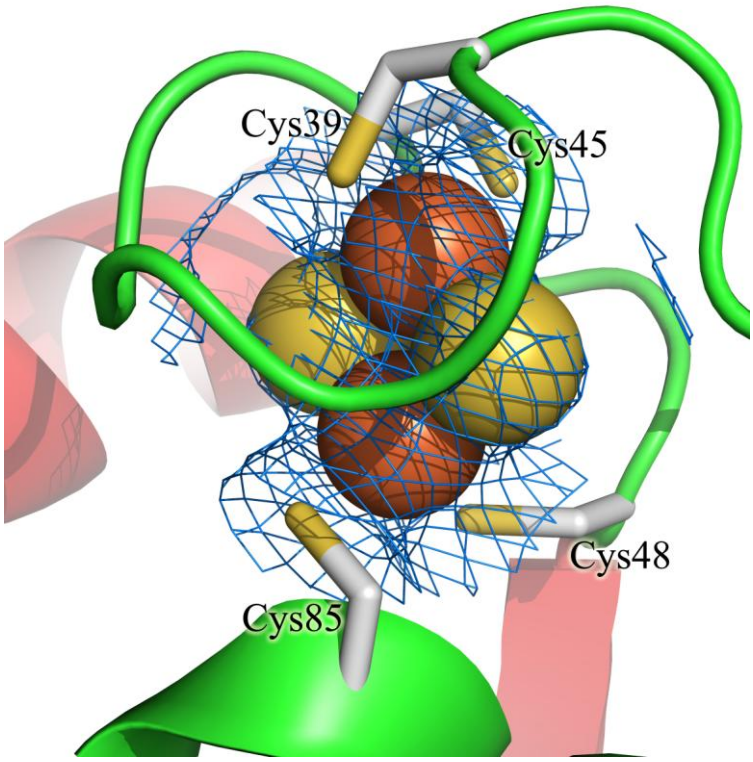
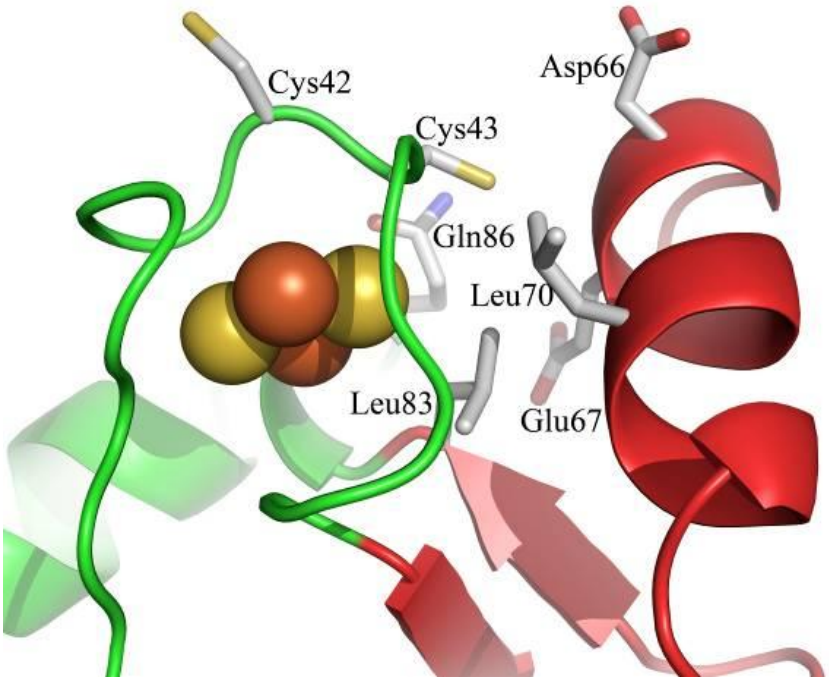


Figure S10. (a) Stereoview of the electron density and model for the [2Fe-2S] cluster in Arx. The electron density of [2Fe-2S] cluster ($2F_o$-F_c contoured at 1σ) and the side chains of the four cysteine residues are colored in marine and grey, respectively.



(b) The location of two non-ligand cysteines (Cys42 and Cys43) in the cluster binding loop of Arx. The [2Fe-2S] cluster, the secondary structure of the core domain (green) and the interaction domain (red), and the side chains of the residues (grey) surrounding Cys43 are shown.



(c) The hydrogen bond network around His49 that connects the Arx C-terminal arm, N-terminal and the cluster binding loop. The relevant residues (green), the hydrogen bonds interactions (black dashed lines) and the water molecules (red) are shown.

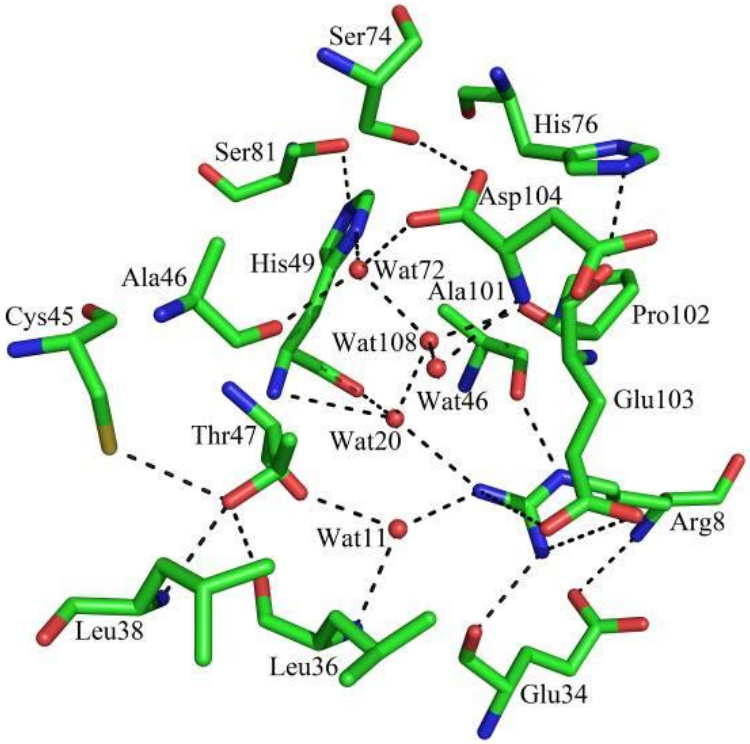
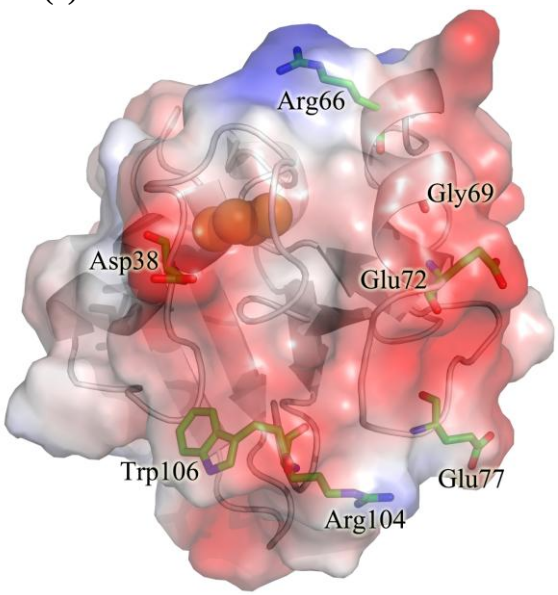


Figure S11. The electrostatic surface potentials of the interaction face of Pdx (a) and Adx (b) surrounding the [2Fe-2S] cluster regions compared to Arx (Fig. 3b). Negatively and positively charged surface areas are colored in red and blue, respectively. The electrostatic potential surfaces of the cluster binding loop in Arx and Adx are more neutral than that of Pdx. Residues that contribute to the different surface potential distributions and that may be involved in protein recognition are labeled.

(a)



(b)

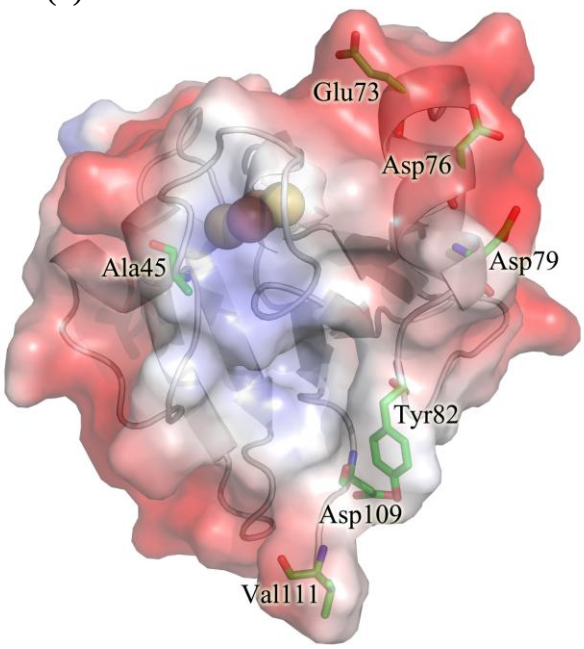


Figure S12. Gel filtration of CYP101D1 (Superdex-200)

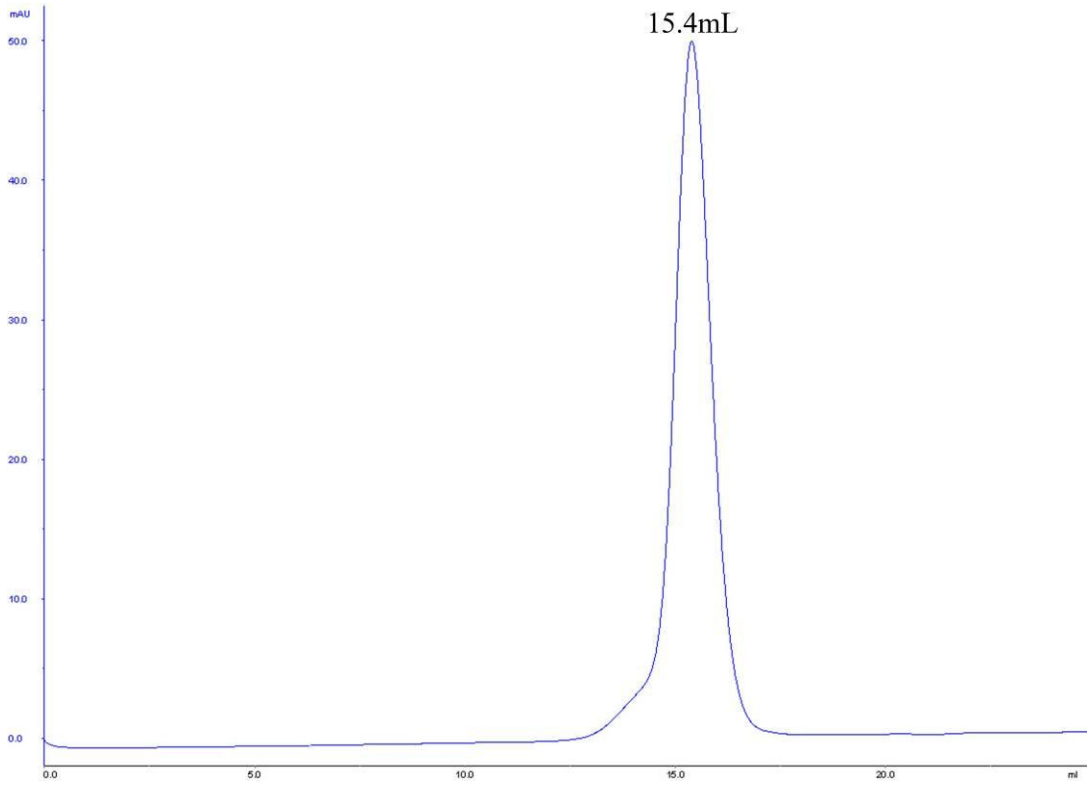


Figure S13. Structure of (a) CYP101D1 (green) and (b) CYP101A1 (P450cam, cyan). The more extended N-terminus of the G helix and the different position of the F-G loop in CYP101D1 are shown in red. Other differences were found in the N-terminal helical region, the β2 sheets and the C-terminal β sheets and are also highlighted in red.

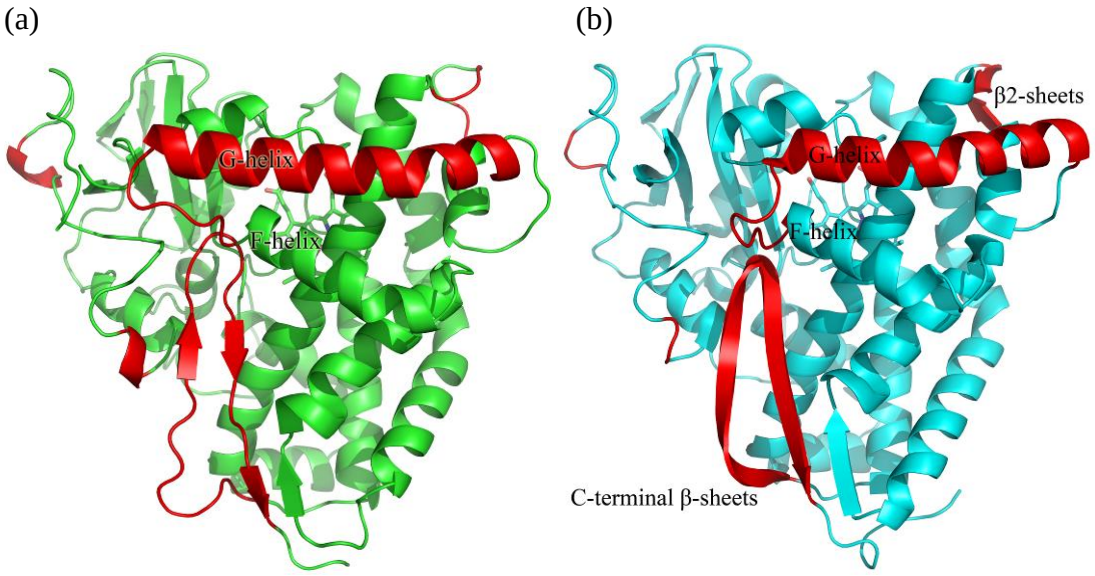
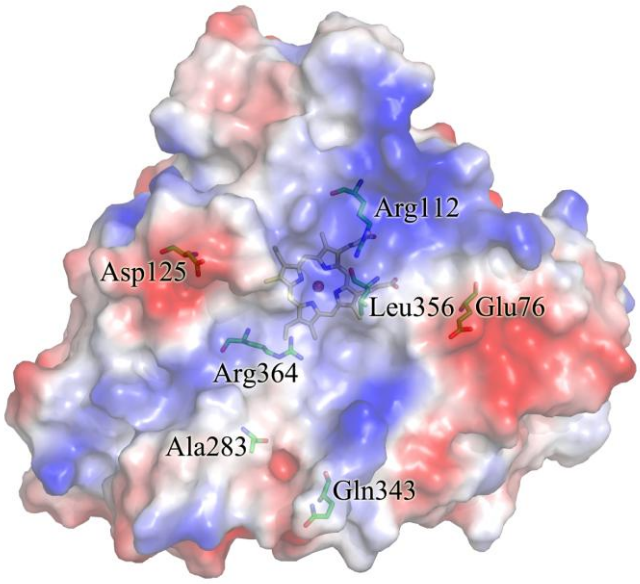
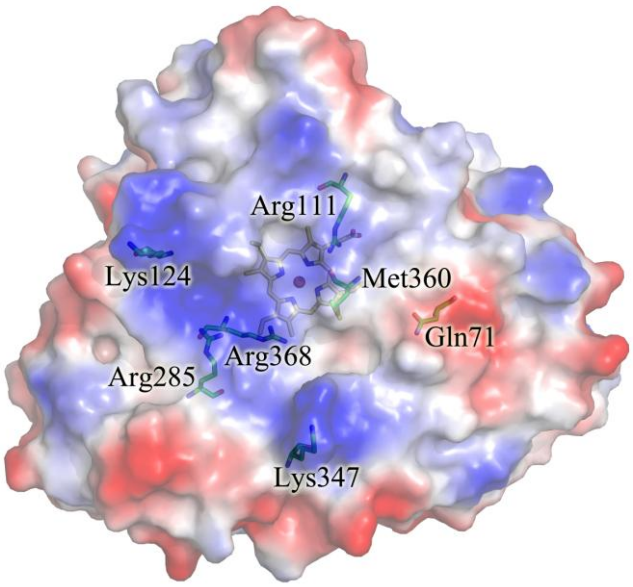


Figure S14. The electrostatic surface potentials of the proximal face of CYP101A1 (a), CYP199A2 from Rhodopseudomonas palustris CGA009 (b) and CYP24A1 from rat mitochondria (c) compared to CYP101D1 (Fig. 3c). Negatively and positively charged surface areas are colored in red and blue, respectively. Residues that contribute to the different surface potential distributions and that may be involved in ferredoxin binding are labeled.

(a)



(b)



(c)

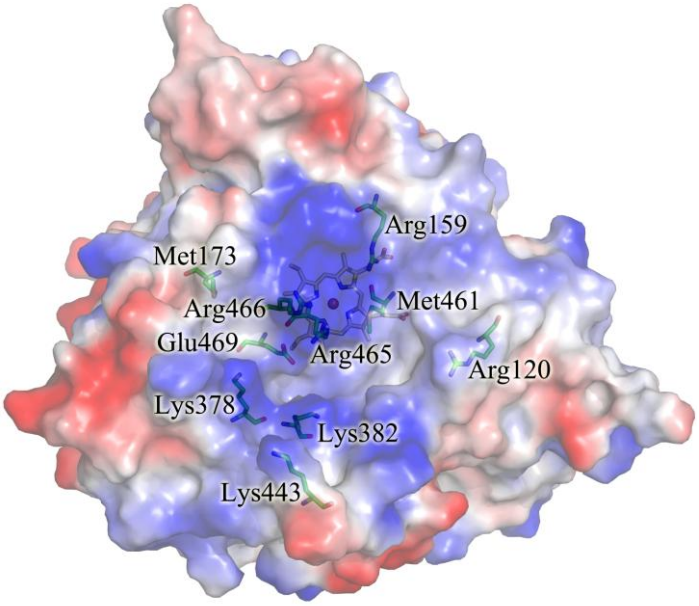
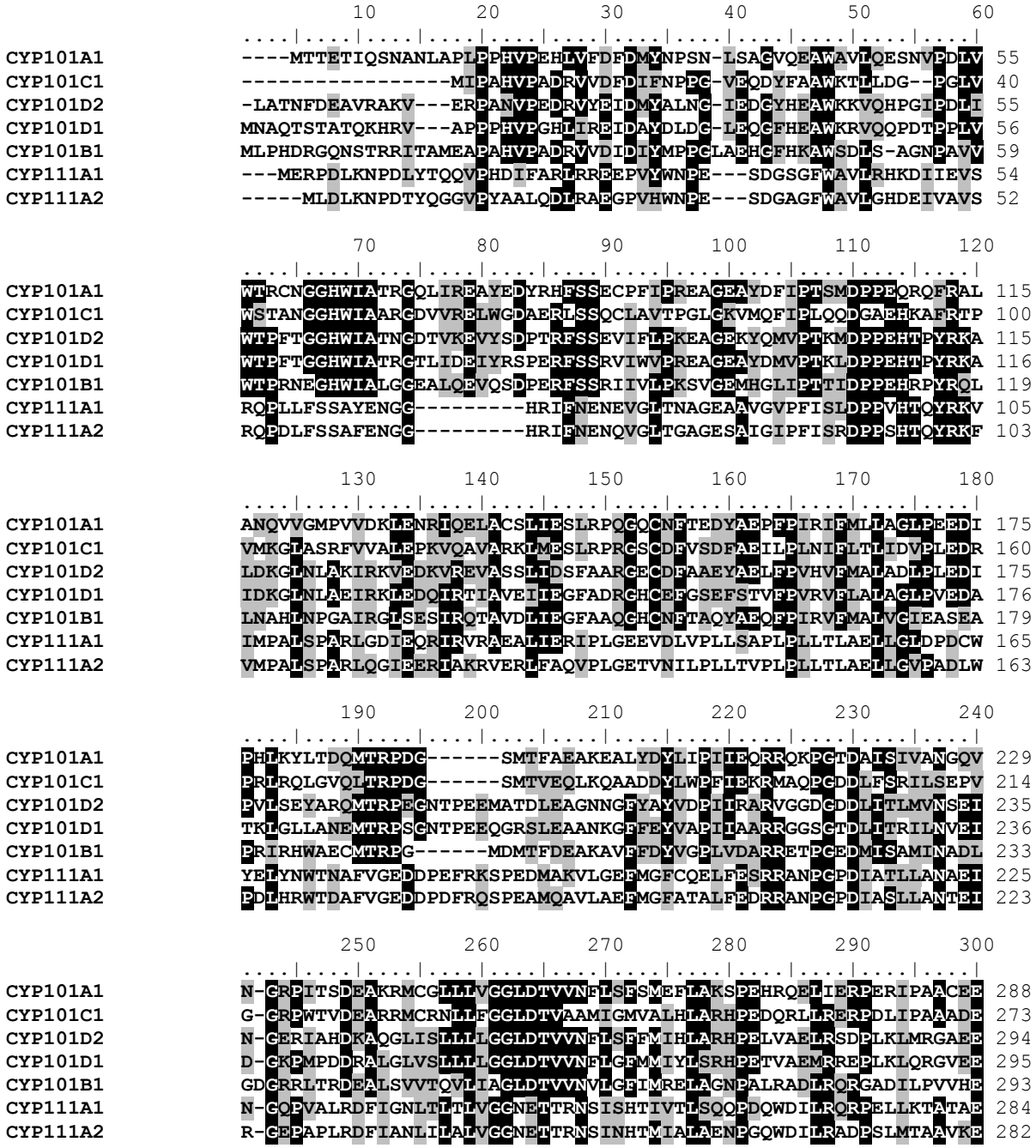
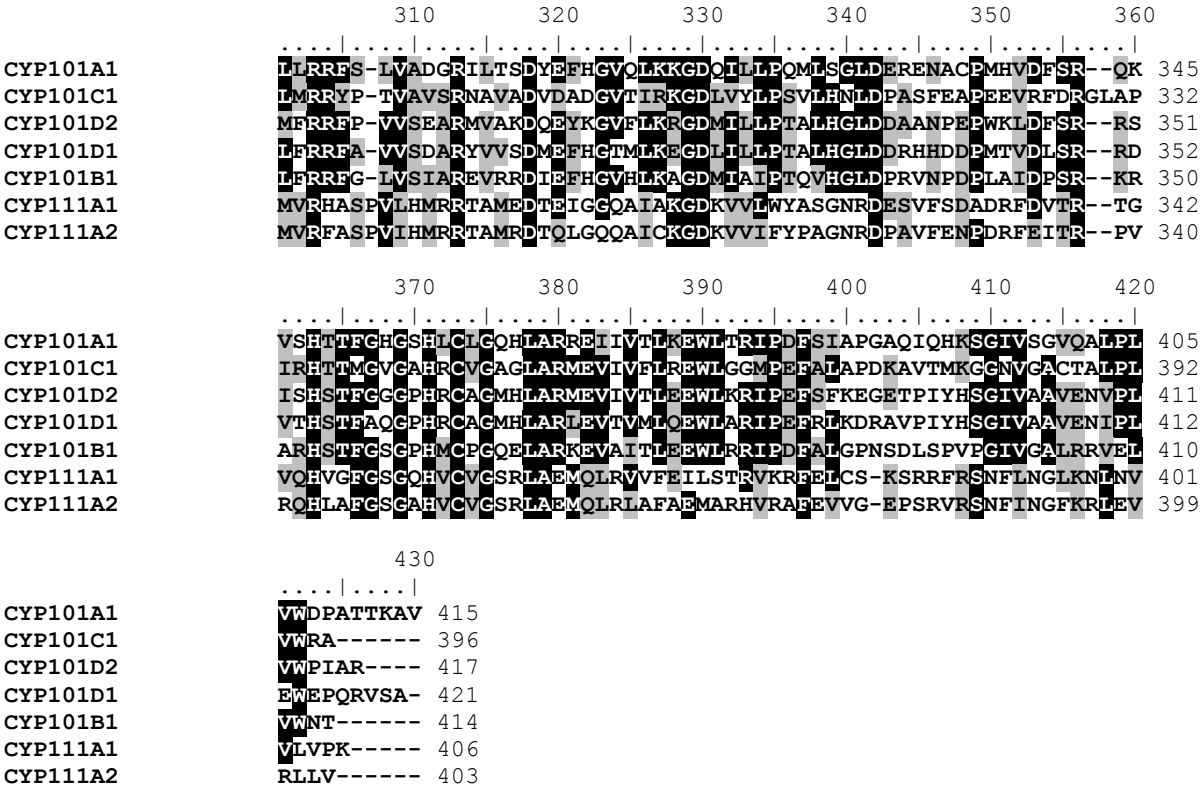

Arg159
Met173
Arg466
Met461
Glu469
Arg465
Arg120
Lys378
Lys382
Lys443

Figure S15. Protein sequence alignments of the CYP101 and CYP111 family proteins from *P. putida*, *P. putida* strain PpG777 and *N. aromaticivorans*.





Sequence alignments were performed using ClustalW (http://www.ebi.ac.uk/clustalw/) and annotated using BioEdit (4,5).

Table S1. Steady state NADH turnover activity of CYP101D2 and CYP101A1 with camphor in the presence or absence of 200 mM KCl, with different electron transfer proteins from *Pseudomonas putida* and *Novosphingobium aromaticivorans*. Turnover rates are given in nmol.(nmol CYP)$^{-1}$.min^{-1} and the data are reported as mean±S.D. ($n \geq 3$). Product formation rates and coupling were estimated using the 5-*exo*-hydroxycamphor analogue 2-hydroxy-3-pinanone as a standard and adjusted assuming 100% coupling for the PdR/Pdx/CYP101A1 turnover with camphor. Leak rates were obtained in the absence of substrate and were 25.7 ± 2.7 for ArR/Arx and 11.2 ± 0.9 for PdR/Pdx. ArR: Aromaticivoransredoxin reductase; Arx: Aromaticivoransredoxin; PdR: putidaredoxin reductase; Pdx: putidaredoxin.

	CYP101D2		
Electron transfer chain	NADH consumption rate	Product formation rate	Coupling %
ArR/Arx (200mM KCl)	745±46	741±28	99
ArR/Arx	2020±78	2010±100	99
PdR/Arx (200mM KCl)	156±4.6	153±4.5	98
PdR/Arx	279±5.6	259±23	93
PdR/Pdx (200mM KCl)	16.0±4.7	11.1±3.4	70

	CYP101A1		
Electron transfer chain	NADH consumption rate	Product formation rate	Coupling %
ArR/Arx (200 mM KCl)	24.6±3.1	0.7±0.1	3
ArR/Pdx (200 mM KCl)	17.9±0.9	6.3±0.3	35
PdR/Pdx (200 mM KCl)	753±16	753±16	100

Table S2. Comparison of the active site residues of CYP101D1 and CYP101A1. Differences in CYP101D1 are underlined.

P. putida	***N. aromaticivorans***
CYP101A1	**CYP101D1**
E84	R85
F87	W88
G93	G94
E94	E95
Y96	Y97
F98	M99
T101	T102
T185	T186
L244	L251
V247	L254
V295	V302
D297	D304
I395	I402
V396	V403

Table S3. Protein-camphor interactions in CYP101D1.

O	Tyr97(2.6 Å)
C2	Leu251$C_{\delta1}$(4.0 Å), Trp88C_{H2}(4.0 Å)
C3	Leu251$C_{\delta1}$(4.1 Å), Thr102$C_{\gamma2}$(4.0 Å)
C5	Leu251$C_{\delta1}$(4.0 Å), Gly255C_{α}(4.0 Å)
C6	Leu251$C_{\delta1}$(4.2 Å), Gly255C_{α}(3.5 Å), Leu254C_{γ}(4.2 Å)
C8	Val302$C_{\gamma1}$(3.7 Å)
C9	Val302$C_{\gamma2}$(3.6 Å), Thr259$C_{\gamma2}$(3.6 Å),
C10	Leu254C_{γ}(3.9 Å), Leu254$C_{\delta1}$(3.8 Å), Thr186$C_{\gamma2}$(3.7 Å), Val403$C_{\gamma2}$(3.9 Å)

Table S4. A comparison of selected residues in Adx, Arx and Pdx.

Arx	**Adx**	**Pdx**
Leu38	Ala45	Asp38
Gly65	Asp72	Glu65
Asp66	Glu73	Arg66
Asp69	Asp76	Gly69
Asp72	Asp79	Glu72
Asp75	Tyr82	Thr75

References

1. Romine, M. F., Stillwell, L. C., Wong, K. K., Thurston, S. J., Sisk, E. C., Sensen, C., Gaasterland, T., Fredrickson, J. K., and Saffer, J. D. (1999) *J. Bacteriol.* **181**(5), 1585-1602
2. Bell, S. G., Dale, A., Rees, N. H., and Wong, L. L. (2010) *Appl. Microbiol. Biotechnol.* **86**, 163-175
3. Bell, S. G., and Wong, L. L. (2007) *Biochem. Biophys. Res. Commun.* **360**(3), 666-672
4. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., and Thompson, J. D. (2003) *Nucleic Acids Res.* **31**(13), 3497-3500
5. Hall, T. A. (1999) *Nucleic Acids Symposium Series* **41**(Symposium on RNA Biology III: RNA, Tool & Target), 95-98