

## Diversity and catalytic potential of PAH-specific ring-hydroxylating dioxygenases from a hydrocarbon-contaminated soil

Florence Martin<sup>1,2</sup>, Laure Malagnoux<sup>1,2</sup>, Fabien Violet<sup>1,2</sup>, Jean Jakoncic<sup>3</sup> and Yves Jouanneau<sup>1,2</sup>

<sup>1</sup>CEA, DSV, Laboratoire de Chimie et Biologie des Métaux, F-38054 Grenoble Cedex 9, France.

<sup>2</sup>CNRS UMR 5249, Grenoble, France.

<sup>3</sup>Brookhaven National Laboratory, National Synchrotron Light Source, Upton, NY, USA

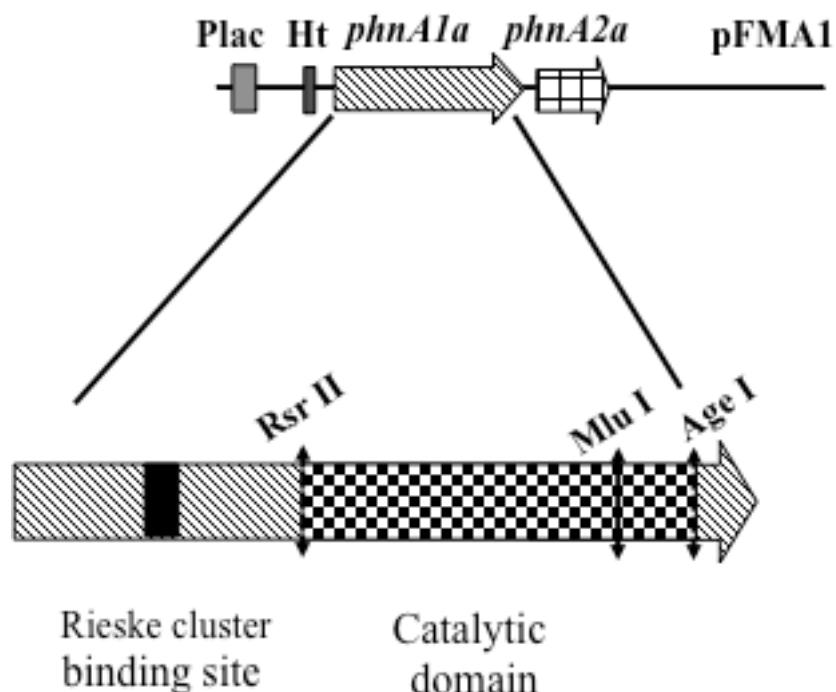


Figure S1 : Scheme illustrating the strategies employed to construct hybrid genes by replacing a portion of the *phnA1a* gene from strain CHY-1 by corresponding sequences amplified from soil DNA. Restriction sites used for hybrid gene construction are indicated. Ht refers to the sequence encoding a 6-histidine tag

phnA1a-CHY1	LLGPPPMKSLLCNWKVPTENFVGDGYHVGVTHAAALQOIGGPLAGLAGNRADIPFDDLGL 240
U3-116hyb	<b>LLGPPPMKSILACNWKVP</b> TENFVGDGYHVGVTHAGALLGIGGELAGLAGNRADIPFEDLGY 240
U3-18hyb	LLGPPPLKSTLACNWKVP <b>TENFVGDGYHVGVTHAAALQOIGGELAGL</b> SGNRADMPFDDLGL 240
U3-60hyb	LLGPPPLKSTLACNWKVP <b>TENFVGDGYHVGVTHAAALQOIGGELAGL</b> SGNRADMPFDDLGL 240
U3-72hyb	LLGPPPLKSTLACNWKVP <b>TENFVGDGYHVGVTHAAALQOIGGELAGL</b> SGNRADMPFDDLGL 240
U3-112hyb	<b>LLGPPPLKSTLACNWEVPTENFVGDGYHVGVTHAAALQOIGGELAGL</b> SGNRADMPFDDLGL 240
Nah1f-LH128	LVGPPPLKSTLACNWKVP <b>TENFVGDGYHVGVTHAAALQOIGGELAGL</b> SGNRADMPFDDLGL 240
U3-89hyb	<b>LLGPPPMKSILRCNWKVPSENFIGDGYHVGVTHAAAIKV</b> LGGPLSLAGN-AEIPFDDAGL 239
	*:****:*** * ***:***:*****:*. : *:***:*** *:***:*
phnA1a-CHY1	QFTTRHGHGFGVIDNAAAAIHRKGDGWNKYLEDTRGEVRRKGADRERLYVGHWNGAIFP 300
U3-116hyb	QFTTRHGHGFGVINNAAAAIHVKRDGWAKYLADTRAEVGRKFPEREELYVGHWNCISIFP 300
U3-18hyb	QFTMRHGHGFG <b>LIDNAATAI</b> HVKRDGYVKYLEETRGGI <b>REKFGPERERLYVGHWNTS</b> IFP 300
U3-60hyb	QFTMRHGHGFG <b>LIDNAATAI</b> HVKRDGYVKYLEETRGGI <b>REKFGPERERLYVGHWNTS</b> IFP 300
U3-72hyb	QFTMRHGHGFG <b>LIDNAATAI</b> HVKRDGYVKYLE <b>QTRGGIREKFGPERERLYVGHWNTS</b> IFP 300
U3-112hyb	QFTMRHGHGFG <b>LIDNAATAI</b> HVKRDGYVKYLEETRGGI <b>REKFGPERERLYVGHWNTS</b> IFP 300
Nah1f-LH128	QFTMRHGHGFG <b>LIDNAATAI</b> HVKRDGYVKYLE <b>QTRGGIREKFGPERERLYVGHWNTS</b> IFP 300
U3-89hyb	QVTSRHGHGFGVIWDGLLIH-DDPAYRDYAYANLPAVAEKMPGPWRAKLYTGHWNAGIFP 298
	*.* *****:*. : . ** . . : * . : . :*. * . *.**** .***
phnA1a-CHY1	NCSFLYGTNTFKIWHPRGPHEIEVWTYTMVPSDADPATKSAIQREATRTFGTAGTLESDD 360
U3-116hyb	SCSFLYGTNTFKVWHPRGPHEIEVWTYTMVPRDADAETKSMIQREATRSFGTAGTLESDD 360
U3-18hyb	NCSFLYGTNTFKIWHPRGPHEIEVWTYTMVPKNADETETKRSIQREAIRSFGTAGTLESDD 360
U3-60hyb	NCSFLYGTNTFKIWHPRGPHEIEVWTYTMVPKNADETETKRSIQREAIRSFGTAGTLESDD 360
U3-72hyb	NCSFLYGTNTFKIWHPRGPHEIEVWTYTMVPKNADETETKRSIQREAIRSFGTAGTLESDD 360
U3-112hyb	NCSFLYGTNTFKIWHPRGPHEIEVWTYTMVPKNADETETKRSIQREAIRSFGTAGTLESDD 360
Nah1f-LH128	NCSFLYGTNTFKIWHPRGPHEIEVWTYTMVPKNADETETKRSIQREAIRSFGTAGTLESDD 360
U3-89hyb	NCSFLYGTNTFKIWHPRGPHEIEVWTWTLVEKAMSP <del>E</del> LAEVVKQAIHSFGTAGTLESDD 358
	*:*****:*****:*****:*****:*. : .. * : :* : :*****:*****
phnA1a-CHY1	GENMSSATYVNRGVITRDGMNSTMGVGYEGPHPVYPGIVGISFIGETSYRGFYRFW <b>KEM</b> 420
U3-116hyb	GENMSSSTFVNRSITRDGRINSSMGAGFEGPHPVPGIVGSSFIGET <b>CYRGFYRFW</b> KEM 420
U3-18hyb	GENMSSATYINNGIITRKGRMNSSMGKDREGPHPVYPGIVGVFIGET <b>CYRGFYRF</b> KEM 420
U3-60hyb	GENMSSATYINNGIITRKGRMNSSMGKDREGPHPVYPGIVGVFIGETTRYRGFYRF <b>KEM</b> 420
U3-72hyb	GENMSSATYNNNGIITRKGRMNSSMGKDREGPHPVYPGIVGVFIGETSYRGFYRF <b>KEM</b> 420
U3-112hyb	GENMSSATYNNNGIITRKGRMNSSMGKDREGPHPVYPGIVGVFIGETSYRGFYRF <b>WOEI</b> 420
Nah1f-LH128	GENMETCTWSNRPQTRKGVMNSQMGQINDAEHPELPGIIGKNFIGETWYRGFYRF <b>KEM</b> 418
U3-89hyb	*****.:.*: *.* ***. :** ** :. ** ***;* .***** *****:*

### Figure S2 : Comparison of alpha subunit sequences of hybrid RHDs constructed in this study.

The alignment was performed with Clustal W2. Only part of the sequences starting at residue 180 (PhnA1a numbering) and corresponding to the catalytic domain are shown. Amino acids highlighted in grey are part of the hybrids that are identical to PhnA1a of strain CHY-1. Residues highlighted in red were shown to be the ligands of the ferrous iron atom at the active site. Cys409, which is present only in two hybrids (see text), is highlighted in blue. The NahA1f sequence of the LH128 RHD has been included in the alignment because it is closely similar to four of the hybrid sequences.