Supplementary Information

A microbial sensor for organophosphate hydrolysis exploiting an engineered specificity switch in a transcription factor

Ramesh K. Jha[1,*], Theresa L. Kern[1,a], Youngchang Kim[2,a], Christine Tesar[2], Robert Jedrzejczak[2], Andrzej Joachimiak[2,3] and Charlie E. M. Strauss[1,*]

[1] Bioscience Division, PO Box 1663, Los Alamos National Laboratory, Los Alamos NM 87545, USA
[2] The Midwest Center for Structural Genomics and Structural Biology Center, Biosciences, Argonne National Laboratory, Argonne, IL 60439, USA
[3] Department of Biochemistry and Molecular Biology, University of Chicago, Illinois 60637, USA.

* To whom correspondence should be addressed. Tel: +1-505-431-6193; Fax: +1-505-665-3024; Email: cems@lanl.gov

Correspondence may also be addressed to: Tel: +1-505-412-1666; Fax: +1-505-665-3024; Email: rjha@lanl.gov

[a] Equal contributions

SECTION 1: FASTA SEQUENCES

A) pobR intergenic region in Acinetobacter sp ADP1

← (*pobR*) CATaacattcaaatccaaaatggttttgtccgatcatcggacagttgtaatgctaatcggataattttga
gccttgattatagatgtctttttaatgaggcggtactttaaaaatagaaaatagcaaggatgatgttATG (*pobA*) →

-PobR operator region is underlined. Inverted repeats are marked with arrows (1).

B) pobR intergenic region adapted for E. coli

← (*pobR*) CATaacattcaaatccaaaatggttttgtccgatcatcggacagttgtaatgctaatcggataattttga
gccttgattatagatgtctttttaatgaggcggtactttaaaaatagaaaatagcaaggaGATATACATATG (*gfp*)
→

-Mutated bases are shown in bold and uppercase letters. *E. coli RBS* is highlighted in yellow. *RBS* in *pobR-pobA* has high sequence identity with *E. coli* consensus *RBS* sequence.

C) pobR-wt sequence adapted for E. coli

ATGGAACAACACCATCAGTATTTGGCGCATCCGCATTCGAGCGAAGAGATTCGTACAGAGGACTACATCGCGGGACT
GGCAAAAGGTCTGGCGTTACTGGAAGCGTTTGGCATTGATCGGCAGCGCTTAAACGTGACACAGGTAGCTGAACGCA
CCGGCATTAGCCGTACAGCAGCTCGTCGCTATCTGAAAACCCTGAAGTTTCTGGGTTACCTGGATACTGACGAACAC
TACTTCTGGTTAACCCATCGTGTTTTGCGCTTCTCTTCGAGCTATCTGAGTTCAGCGCATTTGCCGAAAGTGGCCCA
ATCTTTCCTCAATCTGCTGTGTGCGCAGACGAGTCTGACGTTTAGCATTGTGGTCCTGGATGAACACGAGGTGGTTC
CAGTCGCCCGTTCCTATCTGCCTCAGCAAGACAACTTGCGCGTCAGTCCGTATGGCATGCACTTAGGGAATCGTCTG
CCTGCGCATGCTACCTCAACCGGTAAAGTGCTTCTTAGCGTGCTTGATCGCGAAGTACAGATCGAGTGGATCGAGAA
GTATGGCCTGAAACGCCTGACGCCGTATACCATCACCGATGAACACACCTTTCTTGAAACCCTGGATGCCGTTCGTC
AGTCGGATTACTGCTTATCCACGGAAGAACATGAGCTCGGTCTGATTGCCATTGCGGTTCCAGTCCTCAACGCACAA
GGGCTGACGATTGCAGCCCTGAACTGCATGTCCCAGACTAATCGGGTTCAACCCCAGTACCTCATCGACCAGGTGTT
ACCGTTGCTGCGCAACACTGCCAACGAACTGCGCAATCTGGTATAA

D) DoubleMut sequence

ATGGAACAACACCATCAGTATTTGGCGCATCCGCATTCGAGCGAAGAGATTCGTACAGAGGACTACATCGCGGGACT
GGCAAAAGGTCTGGCGTTACTGGAAGCGTTTGGCATTGATCGGCAGCGCTTAAACGTGACACAGGTAGCTGAACGCA
CCGGCATTAGCCGTACAGCAGCTCGTCGCTATCTGAAAACCCTGAAGTTTCTGGGTTACCTGGATACTGACGAACAC
TACTTCTGGTTAACCCATCGTGTTTTGCGCTTCTCTTCGAGCTATCTGAGTTCAGCGCATTTGCCGAAAGTGGCCCA
ATCTTTCCTCAATCTGCTGTGTGCGCAGACGAGTCTGACGTTTAGCATTGTGGTCCTGGATGAACACGAGGTGGTTC
CAGTCGCCCGTTCCTATCTGCCTCAGCAAGACAACCGCGTCAGTCCGTATGGCATGCACTTAGGGAATCGTCTGCCT
GCGCATGCTACCTCAACCGGTAAAGTGCTTCTTAGCGTGCTTGATCGCGAAGTACAGATCGAGTGGATCGAGAAGTA
TGGCCTGAAACGCCTGACGCCGTATACCATCACCGATGAACACACCTTTCTTGAAACCCTGGATGCCGTTCGTCAGT
CGGATTACTGCTTATCCACGGAAGAACATGAGCTCGGTGTGATTGCCATTGCGGTTCCAGTCCTCAACGCACAAGGG
CTGACGATTGCAGCCCTGAACTGCATGTCCCAGACTAATCGGGTTCAACCCCAGTACCTCATCGACCAGGTGTTACC
GTTGCTGCGCAACACTGCCAACGAACTGCGCAATCTGGTATAA

-The codon TTG originally between AAC and CGC (highlighted yellow) in *pobR-wt* is deleted to represent deleted L141 residue in DoubleMut.
-The underlined codon represents L220V mutation in DoubleMut.

E) pNPmut1 sequence

ATGGAACAACACCATCAGTATTTGGCGCATCCGCATTCGAGCGAAGAGATTCGTACAGAGGACTACATCGCGGGACT
GGCAAAAGGTCTGGCGTTACTGGAAGCGTTTGGCATTGATCGGCAGCGCTTAAACGTGACACAGGTAGCTGAACGCA
CCGGCATTAGCCGTACAGCAGCTCGTCGCTATCTGAAAACCCTGAAGTTTCTGGGTTACCTGGATACTGACGAACAC
TACTTCTGGTTAACCCATCGTGTTTTGCGCTTCTCTTCGAGCTATCTGAGTTCAGCGCATTTGCCGAAAGTGGCCCA
ATCTTTCCTCAATCTGCTGTGTGCGCAGACGAGTCTGACGTTTACCATTGTGGTCCTGGATGAACACGAGGTGGTTC
CAGTCGCCCGTTCCTATCTGCCTCAGCAAGACAACCGCCTCAGTCCGTATGGCATGCACTTAGGGAATCGTCTGCCT
GCGCATGCTATCGCAACCGGTAAAGTGCTTCTTAGCGTGCTTGATCGCGAAGTACAGATCGAGTGGATCGAGAAGTA
TGGCCTGAAACGCCTGACGCCGTATACCATCACCGATGAACACACCTTTCTTGAAACCCTGGATGCCGTTCGTCAGT
CGGATTACTGCTTAACCACGGAAGAATATGAGCTCGGTGTGATTGGCATTGCGGTTCCAGTCCTCAACGCACAAGGG
CTGACGATTGCAGCCCTGAACTGCCTGTCCCAGACTAATCGGGTTCAACCCCAGTACCTCATCGACCAGGTGTTACC
GTTGCTGCGCAACACTGCCAACGAACTGCGCAATCTGGTATAA

-Codons mutated from DoubleMut are underlined.

F) Insert consisting of regulatory region and pNPmut1 gene for Nhe1/Cla1 digested pGLO vector

← *(gfp)*

GCTAGCCATatgtatatctccttgctattttctatttttaaagtaccgcctcattaaaaagacatctataatcaagg
ctcaaaattatccgattagcattacaactgtccgatgatcggacaaaaccattttggatttgaatgttATGGAACAA
CACCATCAGTATTTGGCGCATCCGCATTCGAGCGAAGAGATTCGTACAGAGGACTACATCGCGGGACTGGCAAAAGG
TCTGGCGTTACTGGAAGCGTTTGGCATTGATCGGCAGCGCTTAAACGTGACACAGGTAGCTGAACGCACCGGCATTA
GCCGTACAGCAGCTCGTCGCTATCTGAAAACCCTGAAGTTTCTGGGTTACCTGGATACTGACGAACACTACTTCTGG
TTAACCCATCGTGTTTTGCGCTTCTCTTCGAGCTATCTGAGTTCAGCGCATTTGCCGAAAGTGGCCCAATCTTTCCT
CAATCTGCTGTGTGCGCAGACGAGTCTGACGTTTACCATTGTGGTCCTGGATGAACACGAGGTGGTTCCAGTCGCCC
GTTCCTATCTGCCTCAGCAAGACAACCGCCTCAGTCCGTATGGCATGCACTTAGGGAATCGTCTGCCTGCGCATGCT
ATCGCAACCGGTAAAGTGCTTCTTAGCGTGCTTGATCGCGAAGTACAGATCGAGTGGATCGAGAAGTATGGCCTGAA
ACGCCTGACGCCGTATACCATCACCGATGAACACACCTTTCTTGAAACCCTGGATGCCGTTCGTCAGTCGGATTACT
GCTTAACCACGGAAGAATATGAGCTCGGTGTGATTGGCATTGCGGTTCCAGTCCTCAACGCACAAGGGCTGACGATT
GCAGCCCTGAACTGCCTGTCCCAGACTAATCGGGTTCAACCCCAGTACCTCATCGACCAGGTGTTACCGTTGCTGCG
CAACACTGCCAACGAACTGCGCAATCTGGTATAA*ttggtaacgaatcagacaattgacggcttgacggagtagcata*
gggtttgcagaatccctgcttcgtccatttgacaggcacattatgcatcgat

-Nhe1 and Cla1 restriction sites are highlighted grey.
-Start codons for *gfp* and *pNPmut1* are highlighted green.
-Stop codon for *pNPmut1* is highlighted red.
-PobR operator is underlined, with inverted repeats marked with arrows.
-Promoter proximal inverted repeat sequence of the operator is highlighted yellow.

G) pNPmut1-1 based insert for sensor plasmid highlighting two mutational differences from pNPmut1

← *(gfp)*

GCTAGCCATatgtatatctccttgctattttctatttttaaagtaccgcctcattaaaaagacatctataatcaagg
ctcaaaattatccg**C**ttagcattacaactgtccgatgatcggacaaaaccattttggatttgaatgttATGGAACAA
CACCATCAGTATTTGGCGCATCCGCATTCGAGCGAAGAGATTCGTACAGAGGACTACATCGCGGGACTGGCAAAAGG
TCTGGCGTTACTGGAAGCGTTTGGCATTGATCGGCAGCGCTTAAACGTGACACAGGTAGCTGAACGCACCGGCATTA
GCCGTACAGCAGCTCGTCGCTATCTGAAAACCCTGAAGTTTCTGGGTTACCTGGATACTGACGAACACTACTTCTGG
TTAACCCATCGTGTTTTGCGCTTCTCTTCGAGCTATCTGAGTTCAGCGCATTTGCCGAAAGTGGCCCAATCTTTCCT
CAATCTGCTGTGTGCGCAGACGAGTCTGACGTTTACCATTGTGGTCCTGGATGAACACGAGGTGGTTCCAGTCGCCC
GTTCCTATCTGCCTCAGCAA**G**ACAACCGCCTCAGTCCGTATGGCATGCACTTAGGGAATCGTCTGCCTGCGCATGCT
ATCGCAACCGGTAAAGTGCTTCTTAGCGTGCTTGATCGCGAAGTACAGATCGAGTGGATCGAGAAGTATGGCCTGAA
ACGCCTGACGCCGTATACCATCACCGATGAACACACCTTTCTTGAAACCCTGGATGCCGTTCGTCAGTCGGATTACT
GCTTAACCACGGAAGAATATGAGCTCGGTGTGATTGGCATTGCGGTTCCAGTCCTCAACGCACAAGGGCTGACGATT
GCAGCCCTGAACTGCCTGTCCCAGACTAATCGGGTTCAACCCCAGTACCTCATCGACCAGGTGTTACCGTTGCTGCG

CAACACTGCCAACGAACTGCGCAATCTGGTATAA*ttggtaacgaatcagacaattgacggcttgacggagtagcata*
gggtttgcagaatccctgcttcgtccatttgacaggcacattatgcatcgat

-Two mutations occurring in the operator and in *pNPmut* gene are shown in uppercase bold characters and marked with bold underline. The first mutation results in subtle decrease in constitutive response of the transcription factor while the second mutation converts Aspartate 139 to Asparagine residue and enhances the response amplitude and contrast ratio significantly.

SECTION 2: TABLES & FIGURES

Supplementary Table S1. Data collection and refinement statistics

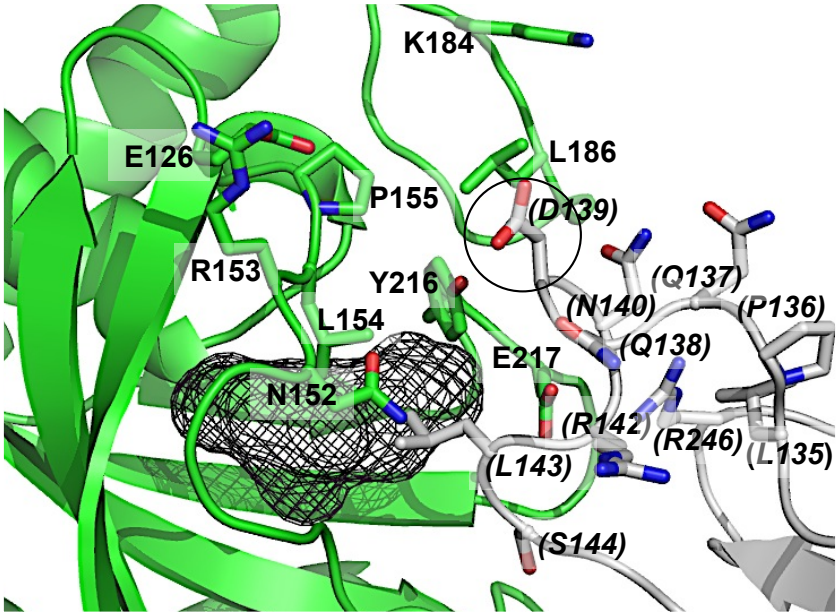
	DoubleMut-IBD	DoubleMut-IBD •3HB
Data collection		
Space group	I422	$C222_1$
Cell dimensions		
a, *b*, *c* (Å)	111.35, 111.35, 282.17	83.49, 125.66, 156.67
Protein molecules/ASU	3	4
Temperature (K)	100	100
Radiation source	APS, 19-ID	APS, 19-ID
Wavelength (Å)	0.97934	0.97934
Resolution (Å)	40.38-2.30 (2.34-2.30)[a]	49.0-2.95 (3.00-2.95)
Unique reflections	37751 (1897)	16803 (777)
R_{merge}[b]	0.119 (0.753)	0.157 (0.827)
$\langle I \rangle$ / $\langle \sigma I \rangle$	23.1(3.1)	7.5(1.3)
Completeness (%)	96.4 (98.0)	96.2 (88.9)
Redundancy	7.8 (7.8)	4.3 (4.0)
Refinement		
Resolution (Å)	2.31 (2.37-2.31)	2.96 (3.15-2.96)
Reflections: work/test set	35868/1874 (2548/118)	15834/817(2318/138)
R_{work}/R_{free}[c]	0.170/0.215 (0.225/0.279)	0.197/0.258 (0.269/0.311)
No. of atoms protein/ligands[d]/water	4138/57/118	5467/77/13
Average *B* factor (Å^2) protein/ligands/water	58.7/96.0/54.0	76.5/80.0/41.9
Bond lengths (Å)	0.008	0.009
Bond angles (°)	0.925	0.977
Most favored	97.9	94.8
Outliers	0.39	0.15
PDB ID	5HPF	5HPI

[a]All numbers in parentheses are from the highest resolution shells, [b] $R_{merge} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i|\langle I_i(hkl)\rangle|$, where I_i(hkl) is the intensity for the *i*th measurement of an equivalent reflection with indices *h*, *k*, and *l*, [c] $R = \Sigma_{hkl}\big||F_{obs}| - |F_{calc}|\big| / \Sigma_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are observed and calculated structure factors, respectively.

Supplementary Table S2. Diversification of promoter proximal inverted repeat of PobR operator and insertion and deletion to modify the distance between operator and promoter

Native intergenic sequence	…CATCTATAATCAAGGCTCAAAATT**ATCCGAT**TAGCATTACAACTGTCCGATGATC…
Randomization (Primer 1)	CATCTATAATCAAGGCTCAAAATT**ATNNNNN**TAGCATTACAACTGTCCGATGATC
Randomization with single deletion (Primer 2)	CATCTATAATCAAGGCTCAAAATT**ANNNNN**TAGCATTACAACTGTCCGATGATC
Randomization with double deletion (Primer 3)	CATCTATAATCAAGGCTCAAAATT**NNNNN**TAGCATTACAACTGTCCGATGATC
Randomization with single insertion (Primer 4)	CATCTATAATCAAGGCTCAAAATT**ATANNNNN**TAGCATTACAACTGTCCGATGATC
Randomization with double insertion (Primer 5)	CATCTATAATCAAGGCTCAAAATT**ATATNNNNN**TAGCATTACAACTGTCCGATGATC

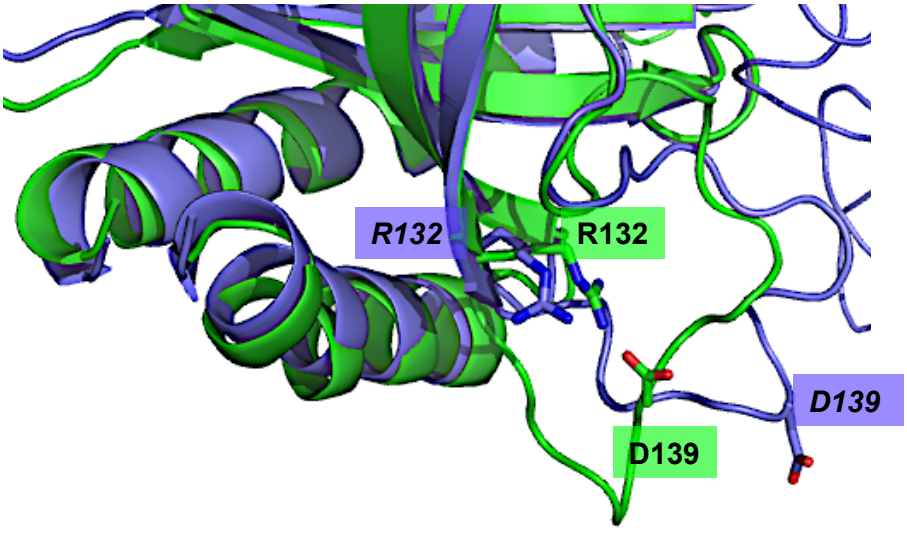
-Equimolar mix of primers 1 to 5 was used with an appropriate reverse primer to amplify the native intergenic sequence and then assemble it using overlapping oligonucleotide PCR assembly method (2).
- Theoretical diversity of the library was 5×10^3.



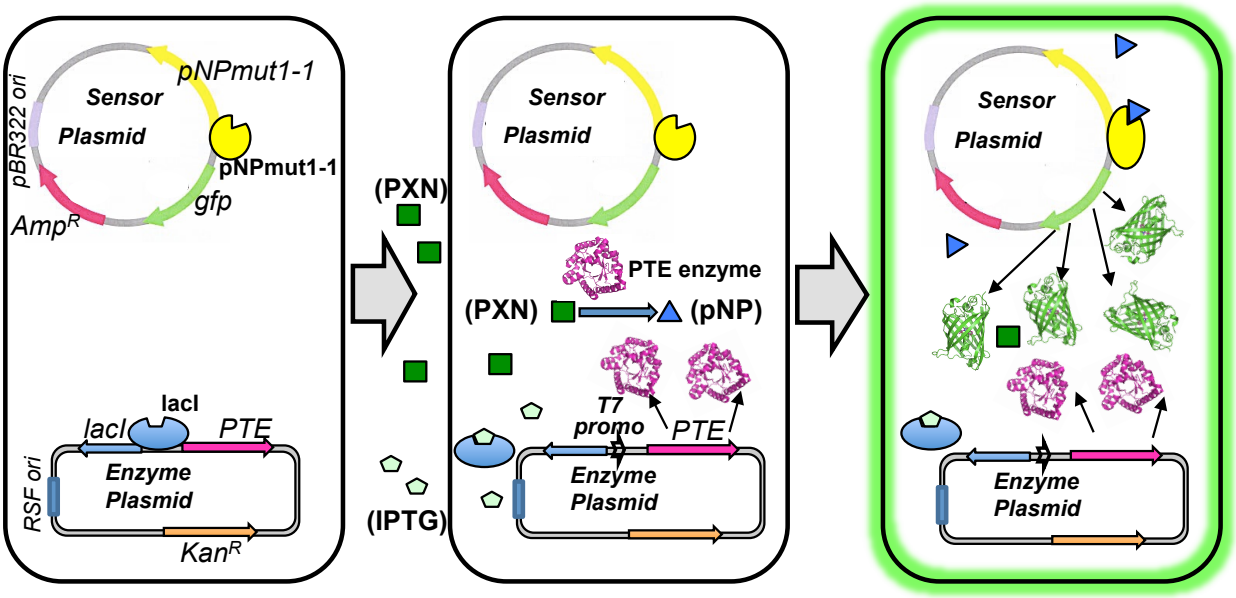
Supplementary Figure S1. Dimer model of pNPmut1 based on symmetric homodimer model of DoubleMut inducer binding domain. Single base change resulted in mutation of D139 (circled) to Asparagine. Residues in 8 Å radius, which could be affected by the mutation directly or indirectly are shown as sticks and labeled. Intrachain neighbors of D139 are labeled in parentheses and shown in *italics*. Possible positioning of pNP in pNPmut1 model is depicted as a 'mesh' and is based on top 10 unique conformations of pNP in the IBD pocket.

```
                 *   :.: : : **: .**:***::..*.  .** .:::**. *.::*::***:* **  ***: **   **** ***.:. **
PobR             MEQHHQYLAHPHSSEEIRTEDYIAGLAKGLALLEAFGIDRQRLNVTQVAERTGISRTAARRYLKTLKFLGYLDTDEHYFWLTHRVLRFSSSY    92
DoubleMut        MEQHHQYLAHPHSSEEIRTEDYIAGLAKGLALLEAFGIDRQRLNVTQVAERTGISRTAARRYLKTLKFLGYLDTDEHYFWLTHRVLRFSSSY    92
2IA2             ------MTATEPTEKILPSPDYVQSLARGLAVIRCFDHRNQRRTLSDVARATDLTRATARRFLLTLVELGYVATDGSAFWLTPRVLELGYSY    92
                 1.......10........20........30........40........50........60........70........80........90..

                 ***  **:***..*: *. :.  : *: :**  ::* ***  :.   *: . *: :*.****:*** *:***: *  :    ::** .::
PobR             LSSAHLPKVAQSFLNLLCAQTSLTFSIVVLDEHEVVPVARSYLPQQDNLRVSPYGMHLGNRLPAHATSTGKVLLSVLDREVQIEWIEKYGLK   184
DoubleMut        LSSAHLPKVAQSFLNLLCAQTSLTFSIVVLDEHEVVPVARSYLPQQDN-RVSPYGMHLGNRLPAHATSTGKVLLSVLDREVQIEWIEKYGLK   184
2IA2             LSSLSLPEVAQPHLEKLSHKVHESSSVSILDGADIVYVARVPVS----RIMTVGITIGTRLPAYATSMGRVLLAGLPDDELDAYLEKLDIQ   184
                 .....100.......110.......120.......130.......140.......150.......160.......170.......180....

                 ***  *** .. :  :: *** .. *:  :* * *: ::*.*: .*.***:**:*  : : * . : * .:::* ** **.::.: :
PobR             RLTPYTITDEHTFLETLDAVRQSDYCLSTEEHELGLIAIAVPVLNAQGLTIAALNCMSQTNRVQPQYLIDQVLPLLRNTANELRNLV-----   276
DoubleMut        RLTPYTITDEHTFLETLDAVRQSDYCLSTEEHELGVIAIAVPVLNAQGLTIAALNCMSQTNRVQPQYLIDQVLPLLRNTANELRNLV-----   276
2IA2             RLTERTITARDELKAAILAVRADGICVLDQELEAGLRSMAAPIRGASGLTVAAVNISTPAARYSLEDLHSDLIPSLRVTATDIEQDLATVNR   276
                 ...190.......200.......210.......220.......230.......240.......250.......260.......270......
```

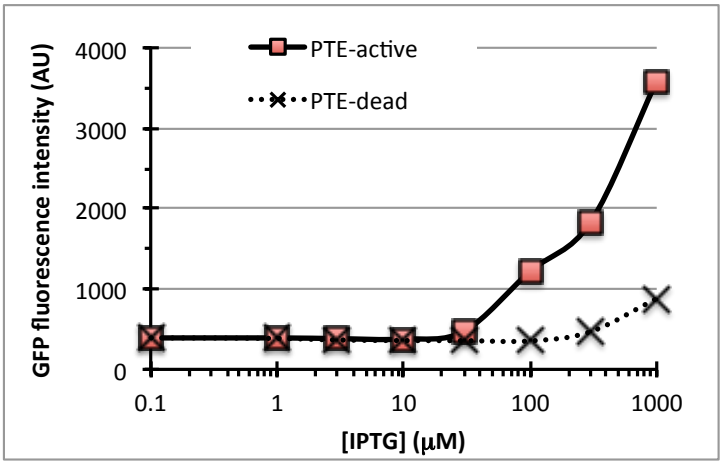
Supplementary Figure S2. Multiple sequence alignment of PobR, DoubleMut and the IclR template (PDB code 2IA2). Highlighted D139 is present on a loop (L1), which is shorter in many IclR transcription factors including the template used for homology modeling.



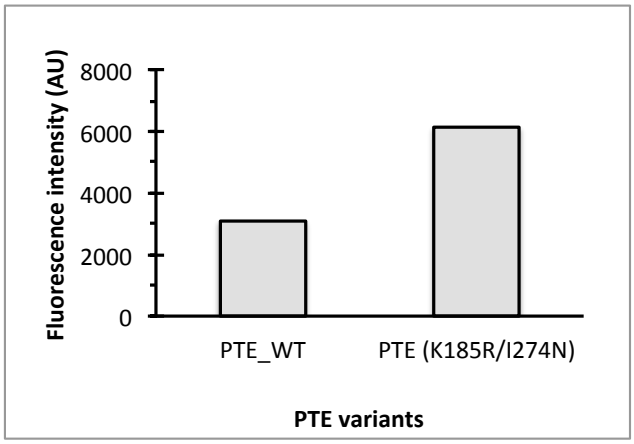
Supplementary Figure S3. A D139/R132 intradomain salt bridge observed in the crystal structure of DoubleMut (green). This salt bridge is absent in the DoubleMut homology model (blue). D139N mutation will affect this interaction resulting in an alternate conformation of loop L1. A similar salt bridge is also observed in another IclR transcription factor (PDB code 2XRO).



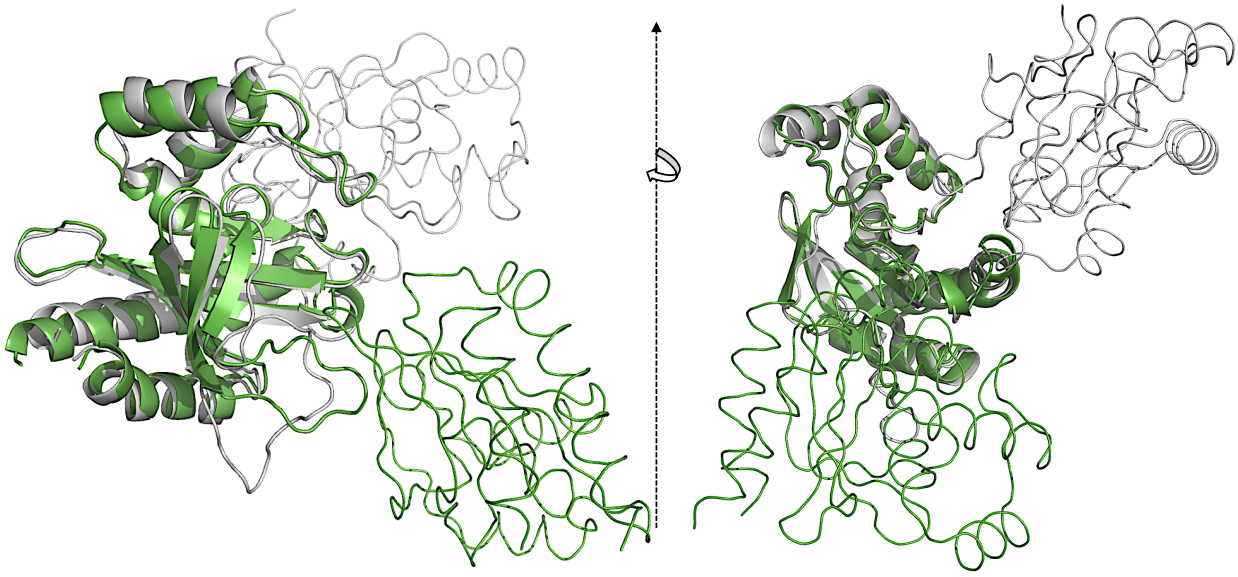
Supplementary Figure S4. A smart microbial cell and a workflow showing reagent addition, intracellular enzyme expression and catalytic activity resulting in activation of the pNP specific transcription factor. Activation of pNPmut1-1 results in expression of green fluorescent protein (GFP), which ultimately makes the cells fluorescent. The figure has been recreated from our earlier work (3).



Supplementary Figure S5. Fluorescence intensity of smart microbial cells in the presence of varying IPTG concentration that controls the expression level of phosphotriesterase (PTE) enzyme. Two variants of PTE (PTE-active consisting of $PTE^{K185R/I274N}$ and PTE-dead consisting of $PTE^{K185R/I274N/W131K/F306E}$) expressed in *E. coli* result in different level of cell fluorescence. PTE-active is similar to wild-type PTE (PTE_WT) in terms of catalytic efficiency (kcat/Km), but is expected to be a higher expressing variant than PTE_WT. W131K and F306E mutations to PTE result in 4- and 3-orders of magnitude weaker catalytic efficiency independently (when Co^{2+} is used as a cofactor) (4) and these mutations were combined in a construct to create PTE-dead version.



Supplementary Figure S6. Effect of K185R and I274N mutations on the wild-type PTE activity. The variants have similar catalytic efficiency but earlier work showed mutations resulted in increase in cell lysate activity (5) possibly due to increase in expression level. Our smart microbial cell format also confirmed the previous observation by showing increase in cell fluorescence intensity when the mutant PTE was expressed inside the cell.



Supplementary Figure S7. Alternate views of DoubleMut-IBD crystal structure dimer (light grey) overlaid on IclR dimer template for homology modeling (green, PDB code 2IA2). While inducer binding domains (IBD) of the template 2IA2 are in C2 symmetry, the crystal structure is asymmetric. One of the possibilities for such an observation is that in the absence of DNA binding domains (which form tighter dimers) and DNA operator, the canonical weak IBD homodimer interface failed to form in the crystal structure.

SECTION 3: COMPUTATIONAL MODELING

Rosetta scripts and step-by-step protocol for homology modeling and ligand-docking related to the present work has been also discussed earlier (6).

```
1. Structural template for homology modeling: 2ia2.pdb
Create a dimer based on chains A and C each from Residue 91-263. The domains
are in C2 symmetry. The dimer is called 2ia2_dimer.pdb

2. Using a perl script in Rosetta version 3.4 or later, create symmetry
definition file

>perl
~/rosetta3.4/rosetta_source/src/apps/public/symmetry/make_symmdef_file.pl -
mode NCS -a A -i B -p input_files/2ia2_dimer.pdb >
input_files/symm_def_2ia2_dimer.dat
```

symm_def_2ia2_dimer.dat is shown below:

```
symmetry_name input_files/2ia2_dimer__2
E = 2*VRT0_base + 1*(VRT0_base:VRT1_base)
anchor_residue 71
virtual_coordinates_start
xyz VRT0  0.1937758,-0.2263170,0.9545845  0.7596052,0.6503845,0.0000000
10.3784383,11.2616803,48.4682588
xyz VRT0_base  0.1937758,-0.2263170,0.9545845  0.7596052,0.6503845,0.0000000
6.9406243,15.2768150,31.5327861
xyz VRT1  -0.1937758,0.2263170,-0.9545845  0.9804343,0.0790260,-0.1802873
10.3784383,11.2616803,48.4682588
xyz VRT1_base  -0.1937758,0.2263170,-0.9545845  0.9804343,0.0790260,-
0.1802873  13.8162524,7.2465455,65.4037314
xyz VRT  1.0000000,0.0000000,0.0000000  0.0000000,1.0000000,0.0000000
11.3784383,11.2616803,48.4682588
virtual_coordinates_stop
connect_virtual JUMP0_to_com VRT0 VRT0_base
connect_virtual JUMP0_to_subunit VRT0_base SUBUNIT
connect_virtual JUMP1_to_com VRT1 VRT1_base
connect_virtual JUMP1_to_subunit VRT1_base SUBUNIT
connect_virtual JUMP0 VRT VRT0
connect_virtual JUMP1 VRT0 VRT1
set_dof JUMP0_to_com x(17.7411980103447) angle_x
set_dof JUMP0_to_subunit angle_x angle_y angle_z
set_jump_group JUMPGROUP2 JUMP0_to_com JUMP1_to_com
set_jump_group JUMPGROUP3 JUMP1_to_subunit JUMP0_to_subunit

3. Create an alignment file for query sequence and template sequence. The
file name is 2ia2_align.grishin, where first sequence is the query sequence
for comparative modeling or DoubleMut (pobR ΔL141/L220V) and second sequence
is the template (2ia2)

1
HLPKVAQSFLNLLCAQTSLTFSIVVLDEHEVVPVARSYLPQQDNRVSPYGMHLGNRLPAHATSTGKVLLSVLDREVQ
IEWIEKYGLKRLTPYTITDEHTFLETLDAVRQSDYCLSTEEHELGVIAIAVPVLNAQGLTIAALNCMSQTNRVQPQY
LIDQVLPLLRNTANELRNLV-----
```

```
0 SLPEVAQPHLEKLSHKVHESSSVSILDGADIVYVARVPVS----
RIMTVGITIGTRLPAYATSMGRVLLAGLPDDELDAYLEKLDIQRLTERTITARDELKAAILAVRADGICVLDQELEA
GLRSMAAPIRGASGLTVAAVNISTPAARYSLEDLHSDLIPSLRVTATDIEQDLATVNR
--
```

4. Using Robetta server (www.robetta.org), to get following files for the query sequence

```
fasta file : t000_.fasta
secondary structure file: t000_.psipred_ss2
fragment files: aat000_09_05.200_v1_3, aat000_03_05.200_v1_3
```

5. Run command (Rosetta compilation for mpi runs required)

```
mpirun -np 159 /mypath/minirosetta.mpi.linuxgccrelease @options_symm_homology
```

where, **options_symm_homology** consist of following options:

```
-database /mypath/rosetta_database
-run:protocol threading
-in:file:alignment input_files/2ia2_align.grishin
-in:file:template_pdb input_files/2ia2_input.pdb
-symmetry_definition input_files/symm_def_2ia2_dimer.dat
-nstruct 20000   # a suitable number of trajectories
-in:file:extended_pose 1
-in:file:fasta input_files/t000_.fasta
-in:file:fullatom
-in:file:psipred_ss2 input_files/t000_.psipred_ss2
-loops:frag_sizes 9 3 1
-loops:frag_files input_files/aat000_09_05.200_v1_3
input_files/aat000_03_05.200_v1_3 none
-loops:random_order
-loops:random_grow_loops_by 4
-loops:remodel quick_ccd
-loops:relax relax
-relax:default_repeats 1
-relax:jump_move true
-loops:constrain_rigid_segments 0.3
-relax:coord_cst_width 2.0
-relax:coord_cst_stdev 1.0
-cm:loop_rebuild_filter 500
-cm:aln_format grishin
-cm:max_loop_rebuild 10
-cm:min_loop_size 4
-out:file:silent silent.out
```

6. Command line for cluster application and extracting PDB from silent file

```
/mypath/cluster.macosgccrelease @cluster_flags
```

where **cluster_flags** consist of following options:

```
-database mypath/rosetta_database
-in:file:silent silent.out
-score:weights score13_env_hb
-in:file:fullatom
-input_score_filter 0
-in:file:silent_struct_type binary
```

```
-limit_cluster_size 5
-limit_clusters 5
-limit_total_structures 25
-cluster:radius 4.0
-sort_groups_by_energy
-out:prefix output/
```

7. Create ligands and their respective params file
A ligand file for 4HB, pNP and 34DHB were created using Avogadro software (7) and partial charges added using AM1-BCC application in QuacPac package, version 1.3.1 of OpenEye Scientific Software (Santa Fe, NM) (http://www.eyesopen.com)

4HB ligand file (phx.mol2)

```
@<TRIPOS>MOLECULE
*****
   16    16     0     0     0
SMALL
USER_CHARGES

@<TRIPOS>ATOM
      1 C1          0.0116    0.6548    0.0000 C.ar      1 <0>         -0.1482
      2 C2         -1.1946   -0.0669   -0.0000 C.ar      1 <0>         -0.0770
      3 C3         -1.1986   -1.4644   -0.0000 C.ar      1 <0>         -0.2387
      4 C4          0.0118   -2.1512   -0.0000 C.ar      1 <0>          0.0827
      5 C5          1.2193   -1.4646   -0.0000 C.ar      1 <0>         -0.1853
      6 C6          1.2179   -0.0671    0.0000 C.ar      1 <0>         -0.0817
      7 C7          0.0087    2.1719    0.0000 C.2       1 <0>          0.9107
      8 O1          1.1388    2.7554    0.0000 O.co2     1 <0>         -0.8323
      9 O2         -1.1258    2.7479    0.0000 O.co2     1 <0>         -0.8353
     10 O3          0.0484   -3.5164   -0.0000 O.3       1 <0>         -0.5241
     11 H1         -2.1358    0.4811   -0.0000 H         1 <0>          0.1499
     12 H2         -2.1493   -1.9844   -0.0000 H         1 <0>          0.1064
     13 H3          2.1593   -2.0051    0.0000 H         1 <0>          0.1219
     14 H4          2.1605    0.4780    0.0000 H         1 <0>          0.1517
     15 H5         -0.8698   -3.8295   -0.0000 H         1 <0>          0.3993
     16 X1          0.0117   -0.7482    0.0000 Du        1 <0>          0.0000
@<TRIPOS>BOND
     1    1    2 ar
     2    1    6 ar
     3    1    7 1
     4    2    3 ar
     5    2   11 1
     6    3    4 ar
     7    3   12 1
     8    4    5 ar
     9    4   10 1
    10    5    6 ar
    11    5   13 1
    12    6   14 1
    13    7    8 2
    14    7    9 1
    15   10   15 1
    16    1   16 1
---
```

pNP ligand file (pnx.mol2)

```
@<TRIPOS>MOLECULE
*****
   15    15     0     0     0
SMALL
USER_CHARGES

@<TRIPOS>ATOM
      1 C1          -3.6990    1.6517    0.1014 C.ar      1 <0>         -0.3255
      2 C2          -3.4730    0.2716    0.0447 C.ar      1 <0>         -0.0005
      3 C3          -2.1635   -0.1957   -0.0492 C.ar      1 <0>         -0.3828
      4 C4          -1.0967    0.7002   -0.0860 C.ar      1 <0>          0.0003
      5 C5          -1.3610    2.0734   -0.0272 C.ar      1 <0>         -0.3261
      6 C6          -2.6620    2.6039    0.0681 C.ar      1 <0>          0.5742
      7 N1          -1.9071   -1.6355   -0.1095 N.pl3     1 <0>          0.3401
      8 O1          -2.8809   -2.4056   -0.0771 O.2       1 <0>         -0.3088
      9 O2          -0.7282   -2.0175   -0.1902 O.2       1 <0>         -0.3091
     10 O3          -2.8849    3.8558    0.1206 O.3       1 <0>         -0.7327
     11 H1          -4.7250    2.0136    0.1749 H         1 <0>          0.1136
     12 H2          -4.3241   -0.4001    0.0754 H         1 <0>          0.1215
     13 H3          -0.0668    0.3678   -0.1588 H         1 <0>          0.1202
     14 H4          -0.5237    2.7713   -0.0563 H         1 <0>          0.1155
     15 X1          -2.3978    1.1759   -0.0077 Du        1 <0>          0.0000
@<TRIPOS>BOND
     1    1    2 ar
     2    2    3 ar
     3    3    4 ar
     4    4    5 ar
     5    5    6 ar
     6    1    6 ar
     7    3    7 1
     8    7    8 2
     9    7    9 2
    10    6   10 1
    11    1   11 1
    12    2   12 1
    13    4   13 1
    14    5   14 1
    15    1   15 1
---
```

34DHB ligand file (dhx.mol2)

```
@<TRIPOS>MOLECULE
*****
   17    17     0     0     0
SMALL
USER_CHARGES

@<TRIPOS>ATOM
      1 C1          -2.7452    1.7819   -0.1371 C.ar      1 <0>         -0.1065
      2 C2          -2.7630    0.3861   -0.2071 C.ar      1 <0>         -0.1996
      3 C3          -1.5614   -0.3065   -0.1366 C.ar      1 <0>          0.0018
      4 C4          -0.3566    0.3803    0.0019 C.ar      1 <0>          0.0732
      5 C5          -0.3330    1.7670    0.0720 C.ar      1 <0>         -0.1038
      6 C6          -1.5358    2.4935    0.0030 C.ar      1 <0>         -0.1214
```

```
      7 C7         -1.5399    4.0084    0.0770 C.2       1 <0>         0.9101
      8 O1          0.8146   -0.3231    0.0693 O.3       1 <0>        -0.5121
      9 O2         -1.5165   -1.6744   -0.1998 O.3       1 <0>        -0.5386
     10 O3         -0.4173    4.5897    0.2032 O.co2     1 <0>        -0.8261
     11 O4         -2.6721    4.5840    0.0068 O.co2     1 <0>        -0.8321
     12 H1         -3.6784    2.3410   -0.1908 H         1 <0>         0.1561
     13 H2         -3.7138   -0.1222   -0.3148 H         1 <0>         0.1066
     14 H3          0.6099    2.2967    0.1801 H         1 <0>         0.1689
     15 H4          0.5295   -1.2536   -0.0011 H         1 <0>         0.4137
     16 H5         -2.4335   -1.9815   -0.2946 H         1 <0>         0.4096
     17 X1         -1.5480    1.0765   -0.0675 Du        1 <0>         0.0000
@<TRIPOS>BOND
     1    1    2 ar
     2    2    3 ar
     3    3    4 ar
     4    4    5 ar
     5    5    6 ar
     6    1    6 ar
     7    6    7 1
     8    4    8 1
     9    3    9 1
    10    7   10 2
    11    7   11 1
    12    1   12 1
    13    2   13 1
    14    5   14 1
    15    8   15 1
    16    9   16 1
    17    2   17 1
---

Params file created using following command line:
/mypath/molfile_to_params.py lig.mol2 -n lig

(where lig refers to phx, pnx and dhx for 4HB, pNP and 34DHB respectively)

NAME phx
IO_STRING phx Z
TYPE LIGAND
AA UNK
ATOM  C6  aroC  X   -0.08
ATOM  C1  aroC  X   -0.15
ATOM  C2  aroC  X   -0.08
ATOM  C3  aroC  X   -0.24
ATOM  C4  aroC  X   0.08
ATOM  C5  aroC  X   -0.19
ATOM  H3  Haro  X   0.12
ATOM  O3  OH    X   -0.52
ATOM  H5  Hpol  X   0.40
ATOM  H2  Haro  X   0.11
ATOM  H1  Haro  X   0.15
ATOM  C7  COO   X   0.91
ATOM  O1  OOC   X   -0.83
ATOM  O2  OOC   X   -0.84
ATOM  X1  VIRT VIRT 0.00
ATOM  H4  Haro  X   0.15
BOND  C1   C2
```

```
BOND  C1   C6
BOND  C1   C7
BOND  C2   C3
BOND  C2   H1
BOND  C3   C4
BOND  C3   H2
BOND  C4   C5
BOND  C4   O3
BOND  C5   C6
BOND  C5   H3
BOND  C6   H4
BOND  C7   O1
BOND  C7   O2
BOND  O3   H5
BOND  C1   X1
CHI 1  C3   C4   O3   H5
PROTON_CHI 1 SAMPLES 2 0 180 EXTRA 1 20
CHI 2  C6   C1   C7   O1
NBR_ATOM  C6
NBR_RADIUS 4.743740
ICOOR_INTERNAL    C6     0.000000    0.000000    0.000000   C6    C1    C2
ICOOR_INTERNAL    C1     0.000000  180.000000    1.405809   C6    C1    C2
ICOOR_INTERNAL    C2     0.000000   61.791205    1.405621   C1    C6    C2
ICOOR_INTERNAL    C3     0.000000   58.942854    1.397506   C2    C1    C6
ICOOR_INTERNAL    C4     0.000000   60.592736    1.391676   C3    C2    C1
ICOOR_INTERNAL    C5     0.000000   59.194374    1.389056   C4    C3    C2
ICOOR_INTERNAL    H3   180.000000   59.522018    1.084316   C5    C4    C3
ICOOR_INTERNAL    O3   180.000000   58.893052    1.365691   C4    C3    C5
ICOOR_INTERNAL    H5     0.000000   72.706660    0.970115   O3    C4    C3
ICOOR_INTERNAL    H2   180.000000   61.158872    1.083619   C3    C2    C4
ICOOR_INTERNAL    H1   180.000000   61.102632    1.089110   C2    C1    C3
ICOOR_INTERNAL    C7   180.000000   58.992423    1.517103   C1    C6    C2
ICOOR_INTERNAL    O1     0.000000   62.801041    1.271848   C7    C1    C6
ICOOR_INTERNAL    O2   180.000000   62.972996    1.272347   C7    C1    O1
ICOOR_INTERNAL    X1   180.000000  120.902137    1.403000   C1    C6    C7
ICOOR_INTERNAL    H4   180.000000   60.938594    1.088866   C6    C1    C2
PDB_ROTAMERS phx.rotlib18.pdb
----------

NAME pnx
IO_STRING pnx Z
TYPE LIGAND
AA UNK
ATOM  C2  aroC  X   -0.00
ATOM  C1  aroC  X   -0.33
ATOM  C6  aroC  X   0.57
ATOM  C5  aroC  X   -0.33
ATOM  C4  aroC  X   0.00
ATOM  C3  aroC  X   -0.38
ATOM  N1  Npro  X   0.34
ATOM  O1  ONH2  X   -0.31
ATOM  O2  ONH2  X   -0.31
ATOM  H3  Haro  X   0.12
ATOM  H4  Haro  X   0.12
ATOM  O3  OOC   X   -0.73
ATOM  X1  VIRT VIRT 0.00
ATOM  H1  Haro  X   0.11
```

```
ATOM  H2  Haro  X   0.12
BOND  C1   C2
BOND  C2   C3
BOND  C3   C4
BOND  C4   C5
BOND  C5   C6
BOND  C1   C6
BOND  C3   N1
BOND  N1   O1
BOND  N1   O2
BOND  C6   O3
BOND  C1   H1
BOND  C2   H2
BOND  C4   H3
BOND  C5   H4
BOND  C1   X1
CHI 1  C4   C3   N1   O1
NBR_ATOM  C2
NBR_RADIUS 4.099173
ICOOR_INTERNAL    C2     0.000000    0.000000    0.000000   C2    C1    C6
ICOOR_INTERNAL    C1     0.000000  180.000000    1.399631   C2    C1    C6
ICOOR_INTERNAL    C6     0.000000   56.847968    1.408248   C1    C2    C6
ICOOR_INTERNAL    C5    -0.009075   64.710905    1.408231   C6    C1    C2
ICOOR_INTERNAL    C4     0.008738   56.845661    1.399639   C5    C6    C1
ICOOR_INTERNAL    C3    -0.006368   60.982937    1.393576   C4    C5    C6
ICOOR_INTERNAL    N1  -179.995716   60.185166    1.463694   C3    C4    C5
ICOOR_INTERNAL    O1  -179.860391   61.866203    1.241930   N1    C3    C4
ICOOR_INTERNAL    O2   179.997235   61.865234    1.241870   N1    C3    O1
ICOOR_INTERNAL    H3  -179.998885   61.132361    1.084658   C4    C5    C3
ICOOR_INTERNAL    H4   179.990544   61.958932    1.090405   C5    C6    C4
ICOOR_INTERNAL    O3  -179.991952   57.644121    1.272672   C6    C1    C5
ICOOR_INTERNAL    X1     0.722342  119.468740    1.389752   C1    C2    C6
ICOOR_INTERNAL    H1   179.274258   61.189479    1.090436   C1    C2    X1
ICOOR_INTERNAL    H2  -179.996995   61.132795    1.084663   C2    C1    C6
PDB_ROTAMERS pnx.rotlib18.pdb
----------

NAME dhx
IO_STRING dhx Z
TYPE LIGAND
AA UNK
ATOM  C5  aroC  X   -0.10
ATOM  C4  aroC  X   0.07
ATOM  C3  aroC  X   0.00
ATOM  C2  aroC  X   -0.20
ATOM  C1  aroC  X   -0.11
ATOM  C6  aroC  X   -0.12
ATOM  C7  COO   X   0.91
ATOM  O3  OOC   X   -0.83
ATOM  O4  OOC   X   -0.83
ATOM  H1  Haro  X   0.16
ATOM  X1  VIRT VIRT 0.00
ATOM  H2  Haro  X   0.11
ATOM  O2  OH    X   -0.54
ATOM  H5  Hpol  X   0.41
ATOM  O1  OH    X   -0.51
ATOM  H4  Hpol  X   0.41
```

```
ATOM  H3  Haro  X   0.17
BOND  C1   C2
BOND  C2   C3
BOND  C3   C4
BOND  C4   C5
BOND  C5   C6
BOND  C1   C6
BOND  C6   C7
BOND  C4   O1
BOND  C3   O2
BOND  C7   O3
BOND  C7   O4
BOND  C1   H1
BOND  C2   H2
BOND  C5   H3
BOND  O1   H4
BOND  O2   H5
BOND  C2   X1
CHI 1  C5   C4   O1   H4
PROTON_CHI 1 SAMPLES 2 0 180 EXTRA 1 20
CHI 2  C4   C3   O2   H5
PROTON_CHI 2 SAMPLES 2 0 180 EXTRA 1 20
CHI 3  C1   C6   C7   O3
NBR_ATOM  C5
NBR_RADIUS 4.760854
ICOOR_INTERNAL    C5     0.000000    0.000000    0.000000   C5    C4    C3
ICOOR_INTERNAL    C4     0.000000    0.000000    1.388671   C5    C4    C3
ICOOR_INTERNAL    C3     0.000000   59.216562    1.393707   C4    C5    C3
ICOOR_INTERNAL    C2    -0.001557   59.518600    1.388707   C3    C4    C5
ICOOR_INTERNAL    C1     0.003494   61.019711    1.397668   C2    C3    C4
ICOOR_INTERNAL    C6    -0.002719   58.678849    1.410195   C1    C2    C3
ICOOR_INTERNAL    C7  -179.997200   59.566566    1.516712   C6    C1    C2
ICOOR_INTERNAL    O3   179.997782   62.647581    1.270459   C7    C6    C1
ICOOR_INTERNAL    O4  -179.997467   63.149145    1.272054   C7    C6    O3
ICOOR_INTERNAL    H1  -179.996523   60.049625    1.089192   C1    C2    C6
ICOOR_INTERNAL    X1    -0.005198  120.559891    1.404409   C2    C3    C1
ICOOR_INTERNAL    H2   179.998362   57.969939    1.083508   C2    C3    X1
ICOOR_INTERNAL    O2  -179.997544   62.079148    1.370095   C3    C4    C2
ICOOR_INTERNAL    H5   179.997751   73.195106    0.971693   O2    C3    C4
ICOOR_INTERNAL    O1  -179.998003   60.233430    1.367854   C4    C5    C3
ICOOR_INTERNAL    H4   179.995830   76.310650    0.975740   O1    C4    C5
ICOOR_INTERNAL    H3   179.994519   59.573924    1.086889   C5    C4    C3
PDB_ROTAMERS dhx.rotlib18.pdb
```

8. Command line for ligand docking

```
/mypath/ligand_dock.macosgccrelease @dock_flags
```

where, dock_flags consists of following options:

```
-database /mypath/rosetta_database
-in:file:s DoubleMut.pdb
-in:file:extra_res_fa dhx.params
-out:file:renumber_pdb
-packing:no_optH
-packing:use_input_sc
-packing:ex1
-packing:ex2
-packing:extrachi_cutoff 1
-docking:randomize2
```

```
-docking:uniform_trans 5
-docking:ligand:minimize_ligand
-docking:ligand:harmonic_torsions 10
-docking:ligand:minimize_backbone
-docking:ligand:harmonic_Calphas 0.3
-docking:ligand:soft_rep
-docking:ligand:protocol abbrev2
-mute core.util.prof
-mute core.io.database
-flip_HNQ
-nstruct 5000
-docking:ligand:start_from 8.4075 12.37 36.728
```

References:

1. Molina-Henares AJ, Krell T, Eugenia Guazzaroni M, Segura A, Ramos JL (2006) Members of the IclR family of bacterial transcriptional regulators function as activators and/or repressors. *FEMS Microbiol Rev* 30(2):157–186.

2. Stemmer WPC, Crameri A, Ha KD, Brennan TM, Heyneker HL (1995) Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene* 164(1):49–53.

3. Jha RK, Kern TL, Fox DT, M. Strauss CE (2014) Engineering an *Acinetobacter* regulon for biosensing and high-throughput enzyme screening in *E. coli* via flow cytometry. *Nucleic Acids Res* 42(12):8150–8160.

4. Watkins LM, Mahoney HJ, McCulloch JK, Raushel FM (1997) Augmented hydrolysis of diisopropyl fluorophosphate in engineered mutants of phosphotriesterase. *J Biol Chem* 272(41):25596–25601.

5. Cho CM-H, Mulchandani A, Chen W (2006) Functional analysis of organophosphorus hydrolase variants with high degradation activity towards organophosphate pesticides. *Protein Eng Des Sel* 19(3):99–105.

6. Jha RK, Chakraborti S, Kern TL, Fox DT, Strauss CEM (2015) Rosetta comparative modeling for library design: Engineering alternative inducer specificity in a transcription factor. *Proteins Struct Funct Bioinforma* 83(7):1327–1340.

7. Hanwell MD, et al. (2012) Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J Cheminformatics* 4(1):17.