



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 05:56 pm BST

PDB ID : 2INN  
Title : Structure of the Phenol Hydroxylase-Regulatory Protein Complex  
Authors : Sazinsky, M.H.; Dunten, P.W.; McCormick, M.S.; Lippard, S.J.  
Deposited on : 2006-10-08  
Resolution : 2.70 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

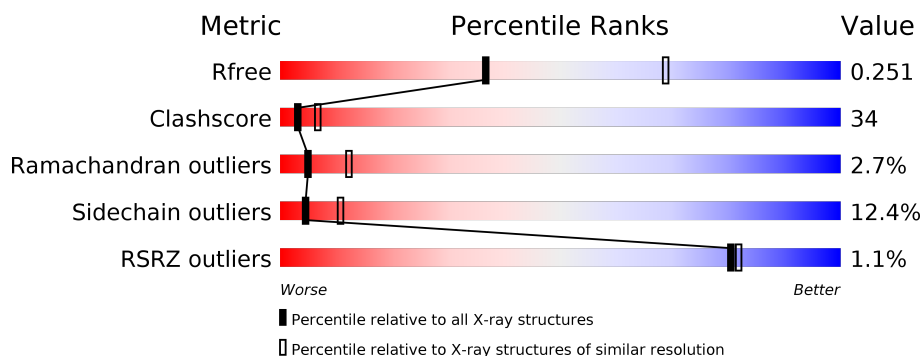
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	511	
1	B	511	
2	C	333	
2	D	333	
3	E	119	
3	F	119	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	L	89	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	MOO	B	515	-	-	X	-

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 16389 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Phenol hydroxylase component pHN.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	496	Total	C	N	O	S	Se	0	0	0
			4129	2655	694	756	8	16			
1	B	493	Total	C	N	O	S	Se	0	3	0
			4140	2665	694	757	8	16			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	21	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	58	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	133	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	149	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	165	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	211	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	220	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	237	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	278	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	279	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	280	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	283	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	289	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	358	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	361	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	416	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
A	510	ASP	ALA	CONFLICT	UNP Q84AQ2
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	21	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	58	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	133	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	149	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	165	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	211	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	220	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	237	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	278	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	279	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	280	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	283	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	289	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	358	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	361	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	416	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ2
B	510	ASP	ALA	CONFLICT	UNP Q84AQ2

- Molecule 2 is a protein called Phenol hydroxylase component pH<sub>L</sub>.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	319	Total	C	N	O	S	Se	0	0	0
			2593	1632	450	493	2	16			
2	D	325	Total	C	N	O	S	Se	0	0	0
			2643	1663	458	504	2	16			

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	1	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	67	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	91	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	135	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	152	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	158	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	173	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	190	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	194	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	198	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	225	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	226	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	234	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	248	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	253	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	267	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
C	268	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	1	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	67	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	91	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	135	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	152	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	158	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	173	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	190	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	194	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	198	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	225	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	226	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	234	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	248	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	253	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	267	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4
D	268	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ4

- Molecule 3 is a protein called Phenol hydroxylase component phO.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	E	118	Total	C	N	O	S	Se	0	0	0
			925	602	145	173	1	4			
3	F	118	Total	C	N	O	S	Se	0	0	0
			925	602	145	173	1	4			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	1	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ1
E	23	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ1
E	48	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ1
E	103	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ1
E	114	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ1
F	1	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ1
F	23	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ1
F	48	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ1
F	103	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ1
F	114	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ1

- Molecule 4 is a protein called Phenol hydroxylase component phM.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	87	Total	C	N	O	Se	0	0	0
			719	446	122	146	5			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	1	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ3
L	23	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ3
L	31	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ3
L	37	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ3
L	53	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ3
L	66	MSE	MET	MODIFIED RESIDUE	UNP Q84AQ3

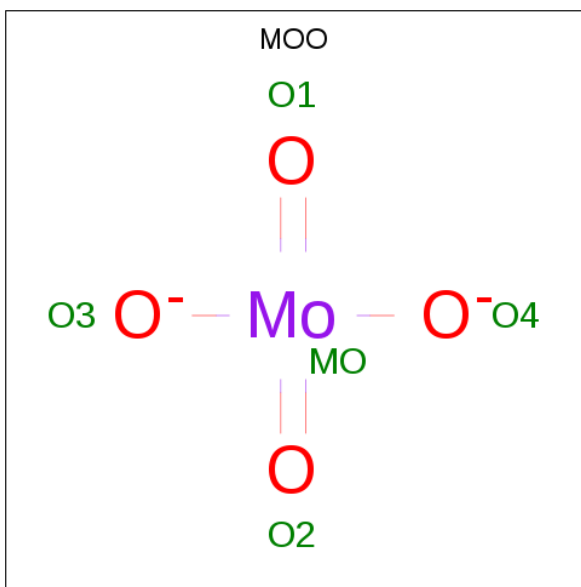
- Molecule 5 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	2	Total	Fe	0	0
			2	2		
5	A	2	Total	Fe	0	0
			2	2		

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total	Zn	0	0
			1	1		
6	A	1	Total	Zn	0	0
			1	1		

- Molecule 7 is MOLYBDATE ION (three-letter code: MOO) (formula: MoO<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	Mo	O	0	0
			5	1	4		

- Molecule 8 is water.

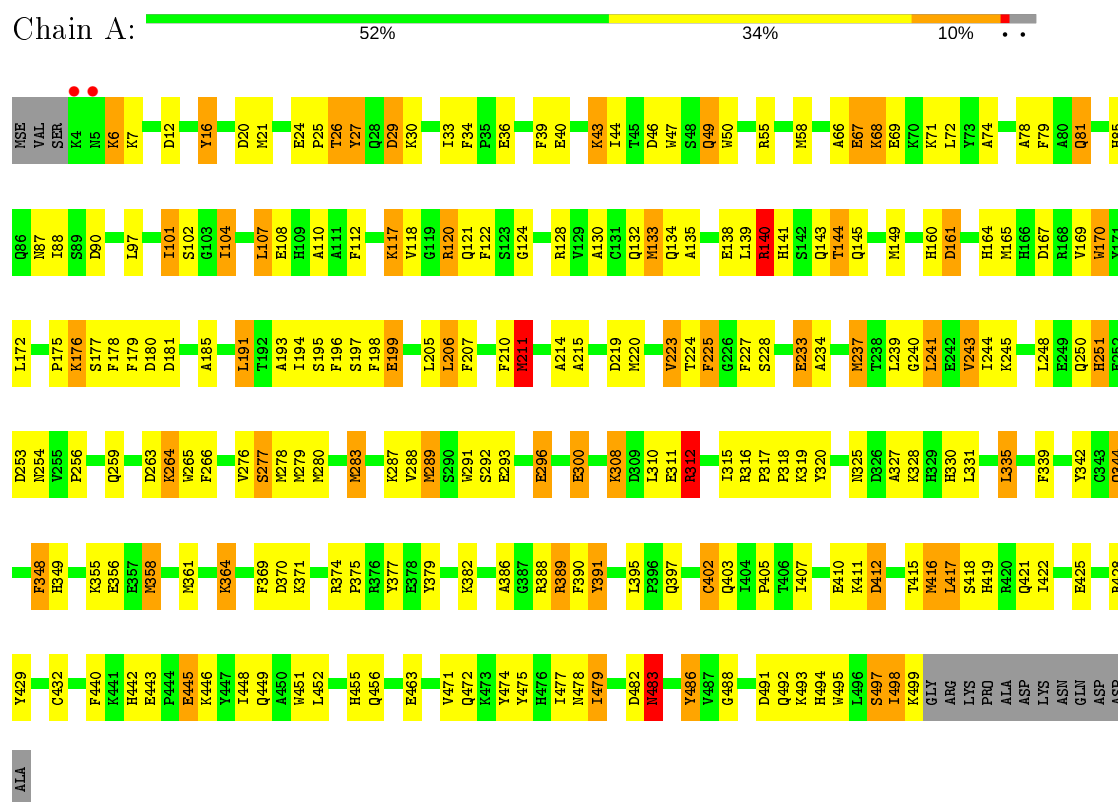
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	80	Total	O	0	0
			80	80		
8	B	71	Total	O	0	0
			71	71		
8	C	53	Total	O	0	0
			53	53		
8	D	60	Total	O	0	0
			60	60		
8	E	23	Total	O	0	0
			23	23		
8	F	13	Total	O	0	0
			13	13		
8	L	4	Total	O	0	0
			4	4		



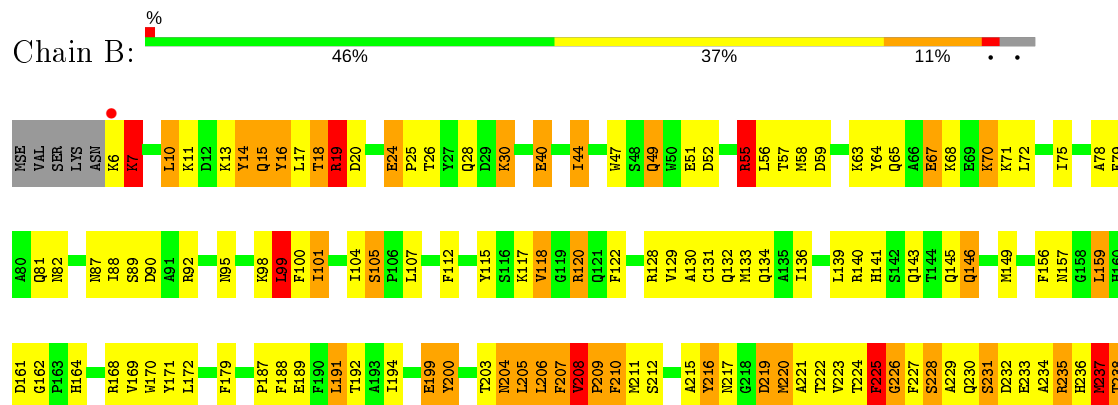
### 3 Residue-property plots

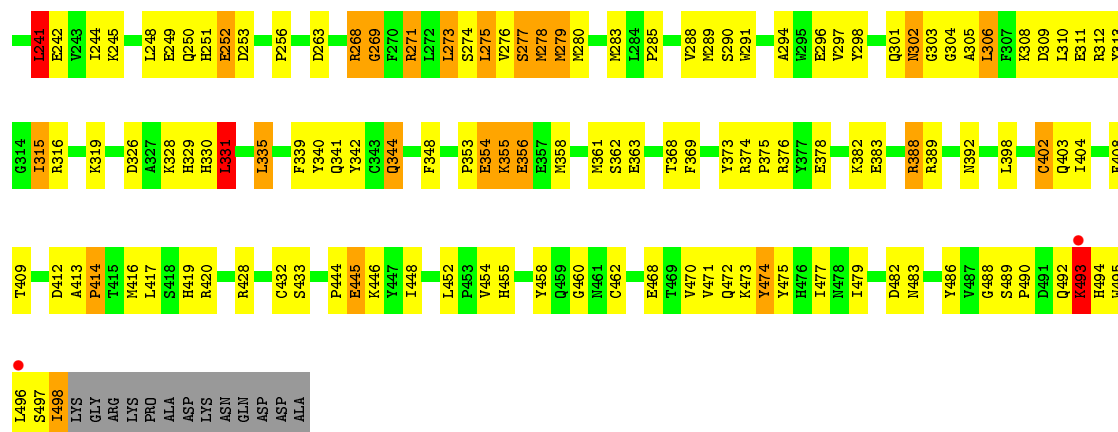
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Phenol hydroxylase component pHN

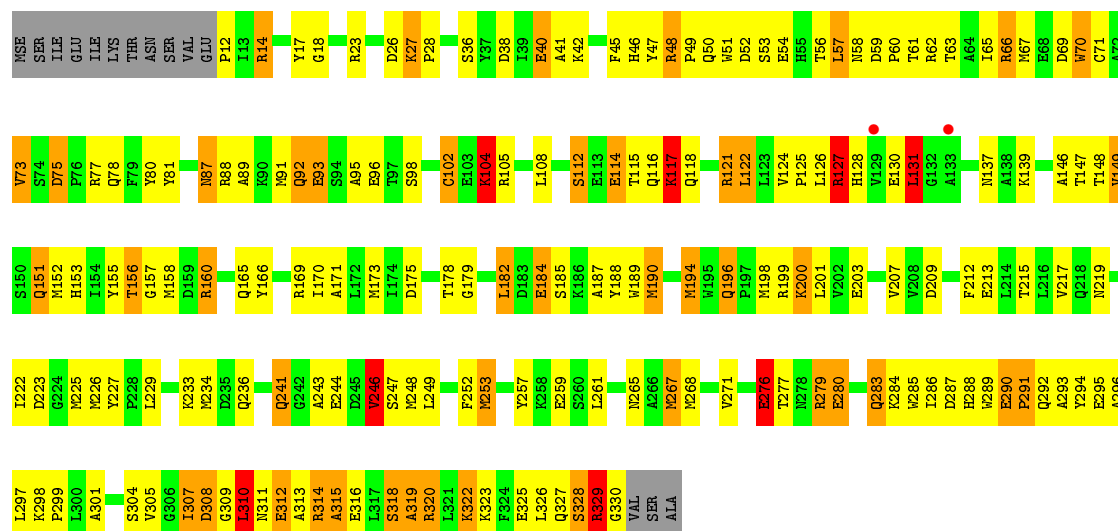
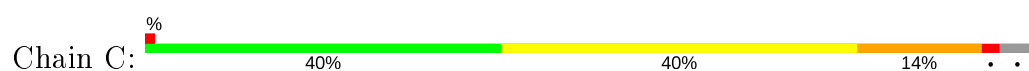


- Molecule 1: Phenol hydroxylase component pHN

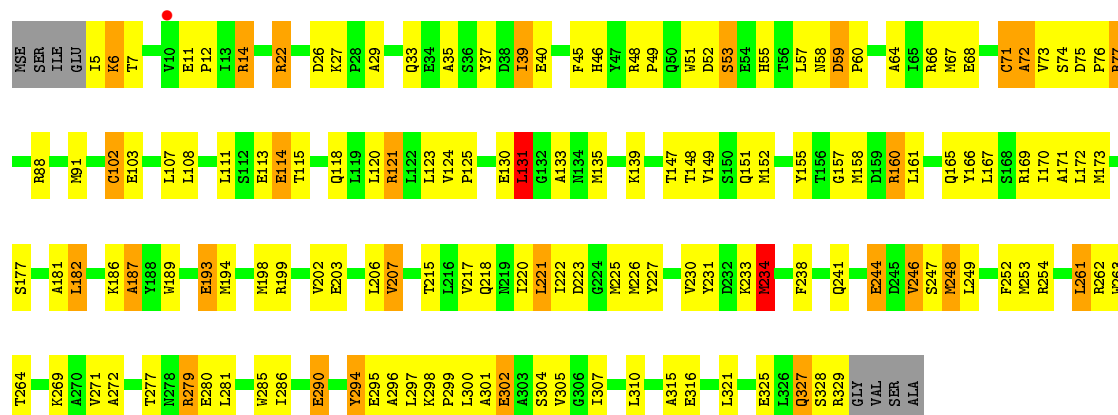




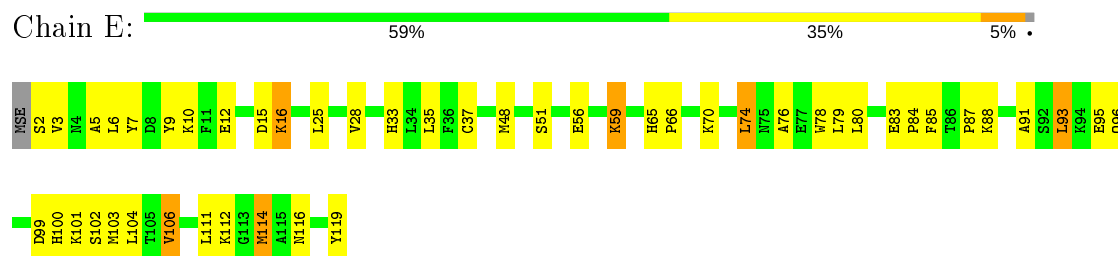
• Molecule 2: Phenol hydroxylase component pH<sub>L</sub>



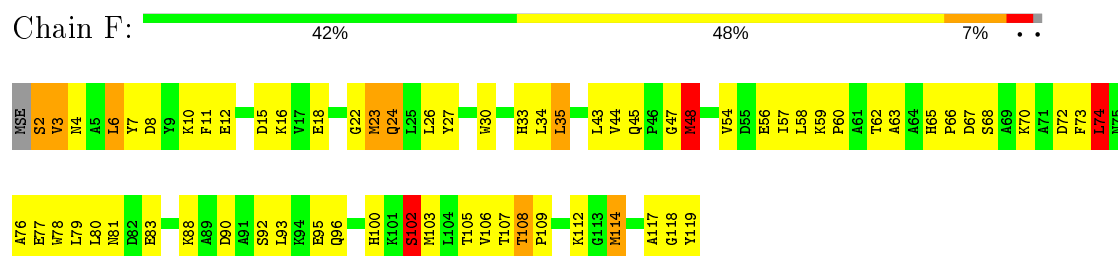
• Molecule 2: Phenol hydroxylase component pH<sub>L</sub>



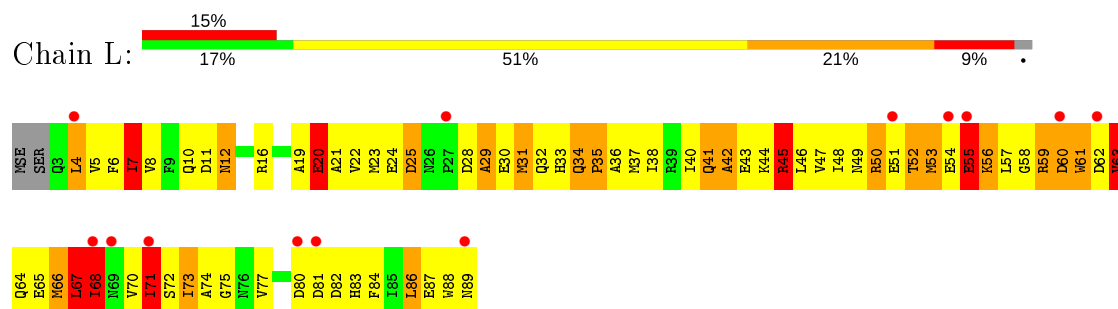
- Molecule 3: Phenol hydroxylase component phO



- Molecule 3: Phenol hydroxylase component phO



- Molecule 4: Phenol hydroxylase component phM



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	87.45Å 146.93Å 189.79Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.70 29.86 – 2.70	Depositor EDS
% Data completeness (in resolution range)	86.3 (30.00-2.70) 86.3 (29.86-2.70)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.37 (at 2.68Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.202 , 0.252 0.164 , 0.251	Depositor DCC
$R_{free}$ test set	2973 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.5	Xtriage
Anisotropy	0.250	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 45.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	16389	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.24% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, MOO, FE

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	1.53	38/4247 (0.9%)	1.29	33/5728 (0.6%)
1	B	1.54	45/4260 (1.1%)	1.33	37/5746 (0.6%)
2	C	1.58	24/2636 (0.9%)	1.38	24/3537 (0.7%)
2	D	1.43	17/2686 (0.6%)	1.33	24/3606 (0.7%)
3	E	1.54	5/949 (0.5%)	1.32	5/1285 (0.4%)
3	F	1.42	5/949 (0.5%)	1.23	4/1285 (0.3%)
4	L	1.03	0/725	1.15	4/971 (0.4%)
All	All	1.50	134/16452 (0.8%)	1.31	131/22158 (0.6%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1

The worst 5 of 134 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	102	CYS	CB-SG	-12.86	1.60	1.82
1	B	210	PHE	CD1-CE1	10.94	1.61	1.39
2	C	71	CYS	CB-SG	-9.38	1.66	1.82
1	A	233	GLU	CG-CD	9.30	1.65	1.51
1	B	118	VAL	CB-CG2	-9.28	1.33	1.52

The worst 5 of 131 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	273	LEU	CA-CB-CG	-12.75	85.98	115.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	66	ARG	NE-CZ-NH1	12.32	126.46	120.30
2	D	77	ARG	NE-CZ-NH1	-11.38	114.61	120.30
2	D	88	ARG	NE-CZ-NH1	10.85	125.73	120.30
1	B	99	LEU	CB-CG-CD1	-10.52	93.12	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	482	ASP	Peptide

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4129	0	3860	228	0
1	B	4140	0	3863	334	0
2	C	2593	0	2504	209	0
2	D	2643	0	2557	123	0
3	E	925	0	898	28	0
3	F	925	0	898	66	0
4	L	719	0	677	158	0
5	A	2	0	0	0	0
5	B	2	0	0	0	0
6	A	1	0	0	1	0
6	B	1	0	0	0	0
7	B	5	0	0	9	0
8	A	80	0	0	14	0
8	B	71	0	0	6	0
8	C	53	0	0	8	0
8	D	60	0	0	1	0
8	E	23	0	0	0	0
8	F	13	0	0	3	0
8	L	4	0	0	0	0
All	All	16389	0	15257	1052	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including

hydrogen atoms). The all-atom clashscore for this structure is 34.

The worst 5 of 1052 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:99:LEU:CG	1:B:99:LEU:CD2	1.80	1.57
4:L:34:GLN:HB3	4:L:35:PRO:CD	1.22	1.52
2:C:190:MSE:CE	2:C:190:MSE:SE	2.14	1.46
1:A:289:MSE:SE	1:A:289:MSE:CE	2.16	1.44
2:D:135:MSE:CE	2:D:135:MSE:SE	2.15	1.44

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	494/511 (97%)	444 (90%)	45 (9%)	5 (1%)	15	37
1	B	494/511 (97%)	416 (84%)	60 (12%)	18 (4%)	3	7
2	C	317/333 (95%)	271 (86%)	38 (12%)	8 (2%)	5	14
2	D	323/333 (97%)	298 (92%)	21 (6%)	4 (1%)	13	32
3	E	116/119 (98%)	107 (92%)	8 (7%)	1 (1%)	17	40
3	F	116/119 (98%)	95 (82%)	18 (16%)	3 (3%)	5	13
4	L	85/89 (96%)	55 (65%)	15 (18%)	15 (18%)	0	0
All	All	1945/2015 (96%)	1686 (87%)	205 (10%)	54 (3%)	5	11

5 of 54 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	251	HIS
1	A	498	ILE
1	B	7	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	228	SER
2	C	329	ARG

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	433/431 (100%)	389 (90%)	44 (10%)	7	17
1	B	435/431 (101%)	386 (89%)	49 (11%)	6	13
2	C	272/268 (102%)	230 (85%)	42 (15%)	2	7
2	D	279/268 (104%)	250 (90%)	29 (10%)	7	16
3	E	98/94 (104%)	87 (89%)	11 (11%)	6	13
3	F	98/94 (104%)	82 (84%)	16 (16%)	2	6
4	L	79/75 (105%)	58 (73%)	21 (27%)	0	1
All	All	1694/1661 (102%)	1482 (88%)	212 (12%)	4	10

5 of 212 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	40	GLU
2	C	246	VAL
4	L	30	GLU
2	C	87	ASN
2	C	139	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 56 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	400	GLN
2	C	50	GLN
3	F	96	GLN
1	B	455	HIS
1	B	492	GLN



### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 6 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
7	MOO	B	515	-	2,4,4	14.03	2 (100%)	-		

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	515	MOO	O2-MO	-14.51	1.43	1.73
7	B	515	MOO	O1-MO	-13.52	1.45	1.73

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	515	MOO	9	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	480/511 (93%)	-0.61	2 (0%) 92 93	17, 32, 47, 61	0
1	B	477/511 (93%)	-0.57	3 (0%) 89 91	18, 32, 57, 69	0
2	C	303/333 (90%)	-0.54	2 (0%) 87 89	17, 34, 55, 74	0
2	D	309/333 (92%)	-0.60	1 (0%) 94 95	18, 34, 52, 73	0
3	E	114/119 (95%)	-0.44	0 100 100	20, 34, 45, 50	0
3	F	114/119 (95%)	-0.34	0 100 100	30, 48, 62, 64	0
4	L	82/89 (92%)	1.02	13 (15%) 1 1	23, 39, 48, 50	82 (100%)
All	All	1879/2015 (93%)	-0.49	21 (1%) 80 82	17, 34, 55, 74	82 (4%)

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	L	81	ASP	3.1
4	L	27	PRO	3.0
4	L	54	GLU	2.9
4	L	4	LEU	2.9
4	L	60	ASP	2.9

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
7	MOO	B	515	5/5	0.84	0.60	84,86,87,88	0
5	FE	A	513	1/1	0.98	0.11	37,37,37,37	0
5	FE	B	512	1/1	0.99	0.10	28,28,28,28	0
5	FE	A	512	1/1	0.99	0.07	35,35,35,35	0
6	ZN	B	514	1/1	1.00	0.10	36,36,36,36	0
6	ZN	A	514	1/1	1.00	0.09	32,32,32,32	0
5	FE	B	513	1/1	1.00	0.08	47,47,47,47	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.