Full wwPDB X-ray Structure Validation Report

May 22, 2020 – 02:27 am BST

PDB ID	:	4PAH
Title	:	HUMAN PHENYLALANINE HYDROXYLASE CATALYTIC DOMAIN DIMER WITH BOUND NOR-ADRENALINE INHIBITOR
Authors	:	Erlandsen, H.; Flatmark, T.; Stevens, R.C.
Deposited on	:	1998-08-20
Resolution	:	2.00 Å(reported)

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

We welcome your comments at *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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

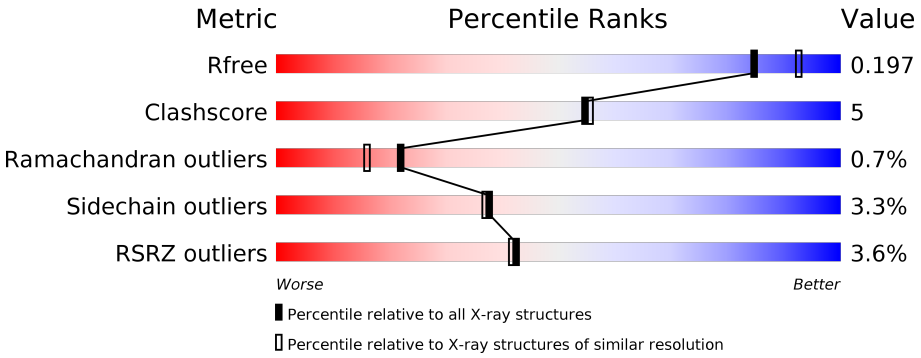
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 >=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 <=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	308	4% 88% 10% •

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
3	LNR	A	600	X	-	-	-

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2665 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.

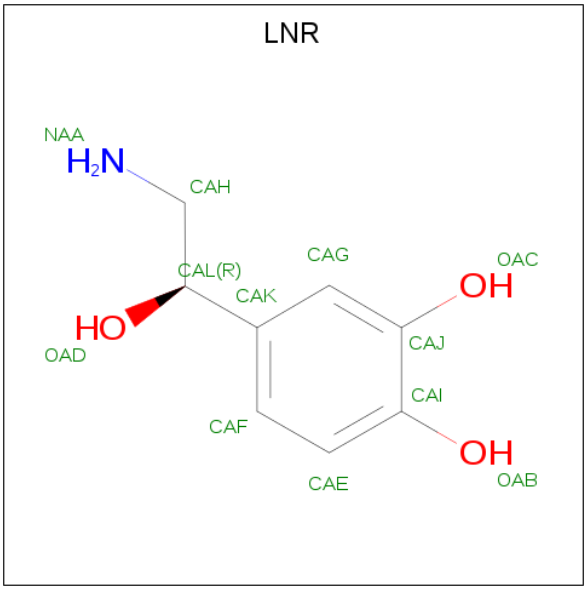
- Molecule 1 is a protein called PHENYLALANINE HYDROXYLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	308	Total	C	N	O	S	0	0	0
			2525	1637	419	460	9			

- Molecule 2 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Fe	0	0
			1	1		

- Molecule 3 is L-NOREPINEPHRINE (three-letter code: LNR) (formula: $C_8H_{11}NO_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			12	8	1	3		

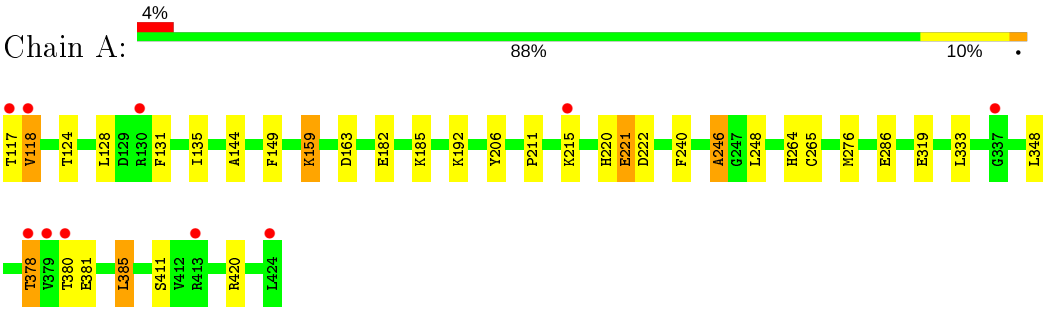
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	127	Total O 127 127	0	0

3 Residue-property plots (i)

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: PHENYLALANINE HYDROXYLASE



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	66.78Å 108.73Å 125.46Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.00 19.98 – 1.91	Depositor EDS
% Data completeness (in resolution range)	98.0 (20.00-2.00) 98.5 (19.98-1.91)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	0.98 (at 1.90Å)	Xtriage
Refinement program	CNS 0.3	Depositor
R, R_{free}	0.169 , 0.208 0.155 , 0.197	Depositor DCC
R_{free} test set	3517 reflections (9.95%)	wwPDB-VP
Wilson B-factor (Å^2)	23.0	Xtriage
Anisotropy	0.213	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 59.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2665	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.86% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: FE, LNR

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	1.10	4/2601 (0.2%)	1.00	1/3527 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	265	CYS	CB-SG	-7.46	1.69	1.82
1	A	182	GLU	CG-CD	5.62	1.60	1.51
1	A	286	GLU	CG-CD	5.13	1.59	1.51
1	A	319	GLU	CG-CD	-5.04	1.44	1.51

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	276	MET	CG-SD-CE	-5.62	91.21	100.20

There are no chirality outliers.

There are no planarity outliers.

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	2525	0	2450	22	0
2	A	1	0	0	0	0
3	A	12	0	9	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	127	0	0	4	0
All	All	2665	0	2459	26	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 5.

All (26) 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:A:192:LYS:HE2	1:A:221:GLU:OE1	1.63	0.97
3:A:600:LNR:HAF	3:A:600:LNR:HNAA	1.41	0.85
1:A:378:THR:HG23	1:A:380:THR:H	1.59	0.67
1:A:385:LEU:HD23	4:A:714:HOH:O	2.00	0.62
3:A:600:LNR:HAF	3:A:600:LNR:NAA	2.14	0.61
1:A:144:ALA:HA	1:A:149:PHE:CD2	2.39	0.57
1:A:220:HIS:CD2	1:A:222:ASP:H	2.26	0.54
1:A:378:THR:CG2	1:A:380:THR:H	2.23	0.51
1:A:192:LYS:HE2	1:A:221:GLU:CD	2.31	0.51
1:A:135:ILE:HD13	1:A:246:ALA:HB3	1.92	0.50
1:A:185:LYS:HE2	4:A:677:HOH:O	2.11	0.50
3:A:600:LNR:CAF	3:A:600:LNR:HNAA	2.21	0.49
1:A:220:HIS:HD2	1:A:222:ASP:H	1.60	0.48
1:A:117:THR:O	1:A:118:VAL:C	2.53	0.47
3:A:600:LNR:CAF	3:A:600:LNR:NAA	2.77	0.46
1:A:385:LEU:O	1:A:385:LEU:HD23	2.17	0.45
1:A:159:LYS:HD3	1:A:159:LYS:O	2.17	0.45
1:A:159:LYS:HE2	1:A:163:ASP:OD1	2.17	0.45
1:A:211:PRO:O	1:A:215:LYS:HG3	2.18	0.44
1:A:124:THR:HG22	1:A:420:ARG:HG2	1.99	0.44
1:A:206:TYR:CD2	1:A:348:LEU:HD13	2.53	0.43
1:A:264:HIS:HE1	4:A:664:HOH:O	2.02	0.43
1:A:264:HIS:HD2	4:A:603:HOH:O	2.01	0.42
1:A:378:THR:HG22	1:A:381:GLU:H	1.84	0.42
1:A:128:LEU:HD22	1:A:131:PHE:CE2	2.55	0.42
1:A:221:GLU:H	1:A:221:GLU:HG3	1.31	0.41

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	306/308 (99%)	299 (98%)	5 (2%)	2 (1%)	22 16

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	246	ALA
1	A	118	VAL

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	270/270 (100%)	261 (97%)	9 (3%)	38 37

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	159	LYS
1	A	221	GLU
1	A	240	PHE
1	A	248	LEU
1	A	333	LEU
1	A	362	PRO
1	A	378	THR
1	A	385	LEU
1	A	411	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	220	HIS
1	A	264	HIS
1	A	375	GLN
1	A	393	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

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

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 2 ligands modelled in this entry, 1 is 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	$\#\|Z\| > 2$
3	LNR	A	600	2	12,12,12	2.80	7 (58%)	15,16,16	0.91	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	LNR	A	600	2	1/1/1/1	2/6/6/6	0/1/1/1

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	600	LNR	CAG-CAK	5.15	1.47	1.39
3	A	600	LNR	CAF-CAK	4.05	1.45	1.39
3	A	600	LNR	CAG-CAJ	3.69	1.44	1.38
3	A	600	LNR	CAK-CAL	3.41	1.57	1.51
3	A	600	LNR	CAI-CAJ	2.99	1.44	1.40
3	A	600	LNR	CAH-CAL	2.51	1.59	1.51
3	A	600	LNR	OAC-CAJ	2.05	1.40	1.36

There are no bond angle outliers.

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	600	LNR	CAL

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	600	LNR	NAA-CAH-CAL-OAD
3	A	600	LNR	NAA-CAH-CAL-CAK

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	600	LNR	4	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

6.1 Protein, DNA and RNA chains

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^{th} 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^2)	Q<0.9
1	A	308/308 (100%)	-0.24	11 (3%) 42 42	14, 24, 46, 61	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	117	THR	8.4
1	A	424	LEU	3.6
1	A	413	ARG	2.7
1	A	215	LYS	2.5
1	A	118	VAL	2.5
1	A	378	THR	2.4
1	A	130	ARG	2.4
1	A	337	GLY	2.3
1	A	370	GLU	2.1
1	A	379	VAL	2.1
1	A	380	THR	2.1

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

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^{th} 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(Å^2)	Q<0.9
3	LNR	A	600	12/12	0.94	0.18	20,36,64,68	0
2	FE	A	425	1/1	1.00	0.03	17,17,17,17	0

6.5 Other polymers (i)

There are no such residues in this entry.