Full wwPDB X-ray Structure Validation Report (i)

Sep 26, 2014 – 03:36 PM EDT

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 2.06 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	Percentile Ranks	Value
Rfree		0.273
Clashscore		1
Ramachandran outliers		0.8%
Sidechain outliers		3.8%
RSRZ outliers		5.3%

Worse   *Better*

█ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
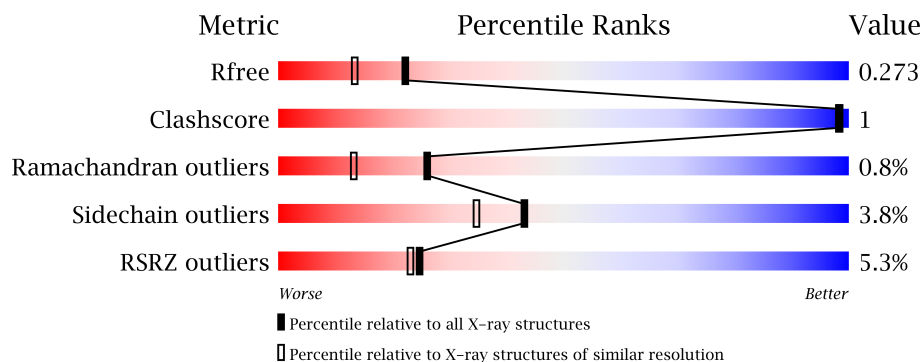
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	77520	1463 (2.08-2.04)
Clashscore	88313	1571 (2.08-2.04)
Ramachandran outliers	86584	1561 (2.08-2.04)
Cα geometry	86677	1559 (2.08-2.04)
Sidechain outliers	86556	1561 (2.08-2.04)
RSRZ outliers	77580	1465 (2.08-2.04)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	326	
1	B	326	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

| Mol	Type	Chain	Res	Geometry	Electron density
2	W9L	A	1548	-	X

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 5758 atoms, of which 0 are hydrogen and 0 are deuterium.

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 BIFUNCTIONAL EPOXIDE HYDROLASE 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	318	Total	C	N	O	S	0	0	0
			2562	1645	431	463	23			
1	B	318	Total	C	N	O	S	0	0	0
			2562	1645	431	463	23			

- Molecule 2 is 1-ETHYL-3-NAPHTHALEN-1-YLUREA (three-letter code: W9L) (formula: $C_{13}H_{14}N_2O$).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			16	13	2	1		
2	B	1	Total	C	N	O	0	0
			16	13	2	1		

- Molecule 3 is water.
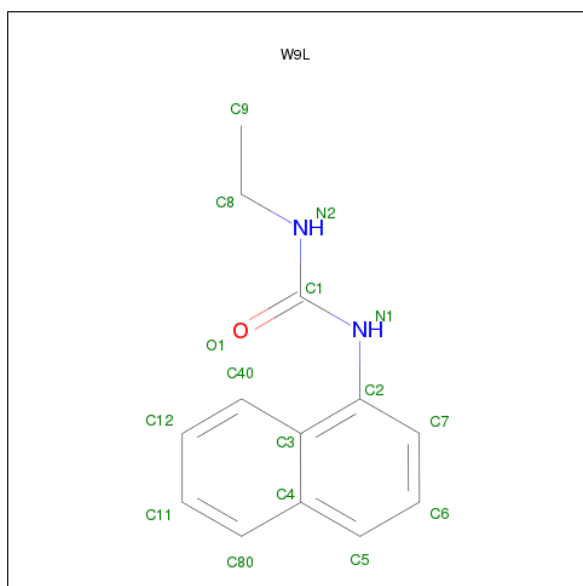
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	334	Total O 334 334	0	0
3	B	268	Total O 268 268	0	0

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 errors 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.
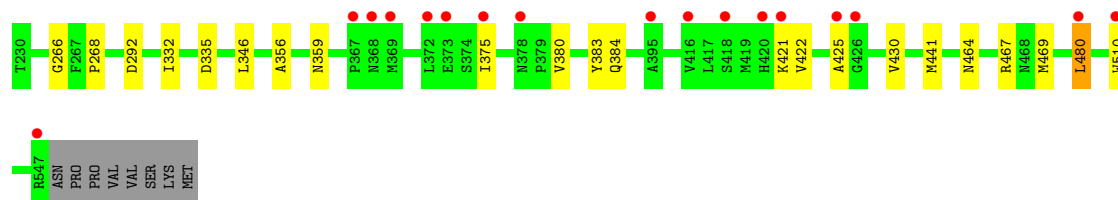
• Molecule 1: BIFUNCTIONAL EPOXIDE HYDROLASE 2

Chain A:
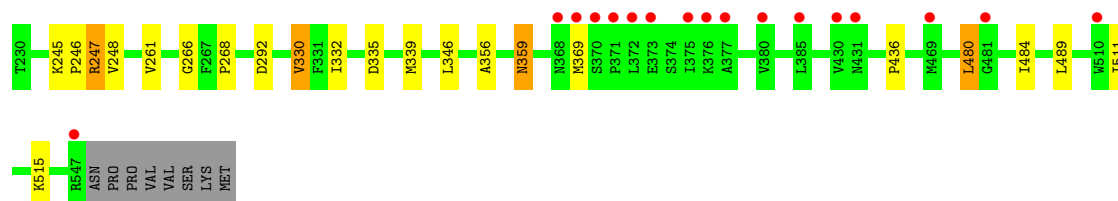
• Molecule 1: BIFUNCTIONAL EPOXIDE HYDROLASE 2

Chain B:

4   Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	46.49Å 79.85Å 87.22Å 90.00° 89.38° 90.00°	Depositor
Resolution (Å)	28.67 – 2.06 36.30 – 2.06	Depositor EDS
% Data completeness (in resolution range)	89.9 (28.67-2.06) 90.0 (36.30-2.06)	Depositor EDS
R_{merge}	0.03	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.87 (at 2.06Å)	Xtriage
Refinement program	BUSTER 2.11.2	Depositor
R, R_{free}	0.220 , 0.253 0.233 , 0.273	Depositor DCC
R_{free} test set	2549 reflections (7.72%)	DCC
Wilson B-factor (Å^2)	20.4	Xtriage
Anisotropy	0.703	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 32.7	EDS
Estimated twinning fraction	0.074 for h,-k,-l	Xtriage
L-test for twinning[2]	$< |L| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Outliers	2 of 35549 reflections (0.006%)	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	5758	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 5.72% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: W9L

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	0.36	0/2638	0.58	0/3578
1	B	0.36	0/2638	0.56	0/3578
All	All	0.36	0/5276	0.57	0/7156

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Close contacts (i)

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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2562	0	2491	7	0
1	B	2562	0	2491	5	0
2	A	16	0	0	0	0
2	B	16	0	0	0	0
3	A	334	0	0	0	0
3	B	268	0	0	0	0
All	All	5758	0	4982	12	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 1.

All (12) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:266:GLY:HA3	1:A:335:ASP:HB3	1.79	0.64
1:A:246:PRO:O	1:A:247:ARG:HB2	2.00	0.62
1:B:266:GLY:HA3	1:B:335:ASP:HB3	1.82	0.61
1:B:332:ILE:HG12	1:B:356:ALA:HB3	1.87	0.55
1:A:359:ASN:N	1:A:359:ASN:HD22	2.10	0.50
1:A:332:ILE:HG12	1:A:356:ALA:HB3	1.95	0.49
1:A:484:ILE:HB	1:A:511:ILE:HG12	1.93	0.48
1:B:425:ALA:HB1	1:B:430:VAL:HG11	1.99	0.45
1:A:261:VAL:HG22	1:A:330:VAL:HG13	2.01	0.42
1:B:464:ASN:HA	1:B:467:ARG:HD2	2.01	0.42
1:A:245:LYS:HB2	1:A:248:VAL:HB	2.01	0.41
1:B:383:TYR:HB3	1:B:422:VAL:HG21	2.03	0.40

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	316/326 (97%)	305 (96%)	8 (2%)	3 (1%)	23	11
1	B	316/326 (97%)	305 (96%)	9 (3%)	2 (1%)	32	19
All	All	632/652 (97%)	610 (96%)	17 (3%)	5 (1%)	26	13

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	480	LEU
1	A	247	ARG
1	B	480	LEU
1	A	268	PRO
1	B	268	PRO

5.3.2 Protein sidechains (i)

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	279/287 (97%)	269 (96%)	10 (4%)	45	37
1	B	279/287 (97%)	268 (96%)	11 (4%)	42	33
All	All	558/574 (97%)	537 (96%)	21 (4%)	43	35

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	292	ASP
1	A	330	VAL
1	A	339	MET
1	A	346	LEU
1	A	359	ASN
1	A	369	MET
1	A	436	PRO
1	A	480	LEU
1	A	489	LEU
1	A	515	LYS
1	B	292	ASP
1	B	346	LEU
1	B	359	ASN
1	B	375	ILE
1	B	380	VAL
1	B	384	GLN
1	B	421	LYS
1	B	441	MET
1	B	469	MET
1	B	480	LEU
1	B	510	TRP

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	359	ASN
1	A	384	GLN
1	A	399	GLN

Continued on next page...

Mol	Chain	Res	Type
1	B	399	GLN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

2 ligands are modelled in this entry.

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	$\#|Z| > 2$
2	W9L	A	1548	-	17,17,17	1.49	4 (23%)	22,22,22	0.55	0
2	W9L	B	1548	-	17,17,17	1.51	4 (23%)	22,22,22	0.53	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
2	W9L	A	1548	-	-	0/7/7/7	0/2/2/2
2	W9L	B	1548	-	-	0/7/7/7	0/2/2/2

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1548	W9L	C1-N1	2.96	1.42	1.37
2	B	1548	W9L	C1-N1	2.94	1.42	1.37
2	B	1548	W9L	C1-N2	2.87	1.41	1.35
2	A	1548	W9L	C1-N2	2.79	1.41	1.35
2	B	1548	W9L	C12-C11	2.25	1.43	1.38
2	A	1548	W9L	C12-C11	2.21	1.43	1.38
2	B	1548	W9L	C6-C7	2.19	1.43	1.38
2	A	1548	W9L	C6-C7	2.14	1.43	1.38

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues

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, 95th 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	318/326 (97%)	0.24	17 (5%)	25	23	12, 20, 49, 75	0
1	B	318/326 (97%)	0.37	17 (5%)	25	23	11, 25, 53, 73	0
All	All	636/652 (97%)	0.31	34 (5%)	25	23	11, 22, 52, 75	0

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	377	ALA	4.6
1	A	372	LEU	4.5
1	B	510	TRP	4.4
1	B	372	LEU	4.2
1	B	480	LEU	3.8
1	B	421	LYS	3.4
1	A	370	SER	3.4
1	B	426	GLY	3.4
1	A	375	ILE	3.3
1	A	376	LYS	3.3
1	A	380	VAL	3.2
1	A	481	GLY	3.2
1	A	430	VAL	3.1
1	B	373	GLU	3.1
1	B	416	VAL	3.1
1	B	420	HIS	3.0
1	B	378	ASN	2.9
1	A	368	ASN	2.8
1	B	368	ASN	2.8
1	B	375	ILE	2.7
1	B	418	SER	2.6
1	A	371	PRO	2.6
1	B	547	ARG	2.5
1	B	367	PRO	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	369	MET	2.4
1	A	431	ASN	2.4
1	A	369	MET	2.3
1	A	385	LEU	2.3
1	B	425	ALA	2.2
1	A	469	MET	2.2
1	A	373	GLU	2.2
1	A	547	ARG	2.2
1	B	395	ALA	2.2
1	A	510	TRP	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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th 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	RSR	LLDF	B-factors(Å²)	Q<0.9
2	W9L	A	1548	16/16	0.20	2.05	33,33,33,34	0
2	W9L	B	1548	16/16	0.16	0.99	38,39,39,39	0

6.5 Other polymers ⓘ

There are no such residues in this entry.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK