Full wwPDB X-ray Structure Validation Report (i)

May 5, 2025 – 08:48 PM EDT

PDB ID	:	9C4W / pdb_00009c4w
Title	:	Menin mutant - G331R
Authors	:	Clegg, B.D.; Cierpicki, T.; Grembecka, J.
Deposited on	:	2024-06-05
Resolution	:	1.40 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

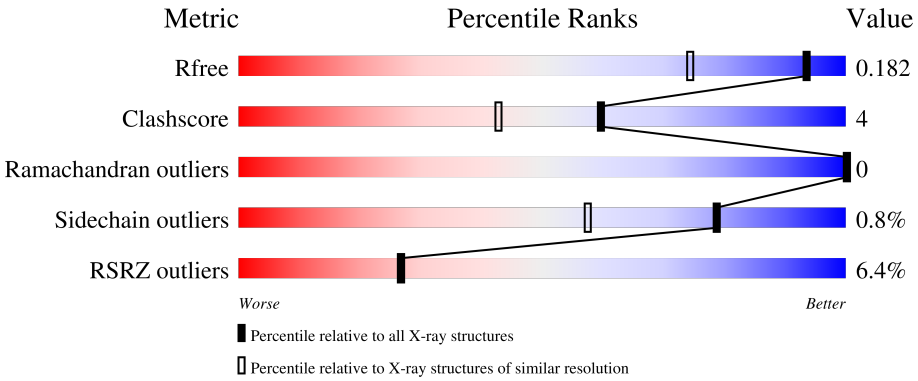
MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.40 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}	164625	2247 (1.40-1.40)
Clashscore	180529	2446 (1.40-1.40)
Ramachandran outliers	177936	2398 (1.40-1.40)
Sidechain outliers	177891	2397 (1.40-1.40)
RSRZ outliers	164620	2246 (1.40-1.40)

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 of 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	489	6% 88% 7% .

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8347 atoms, of which 3918 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 Menin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	468	7737	2465	3864	674	717	17	131	45	0

There are 39 discrepancies between the modelled and reference sequences:

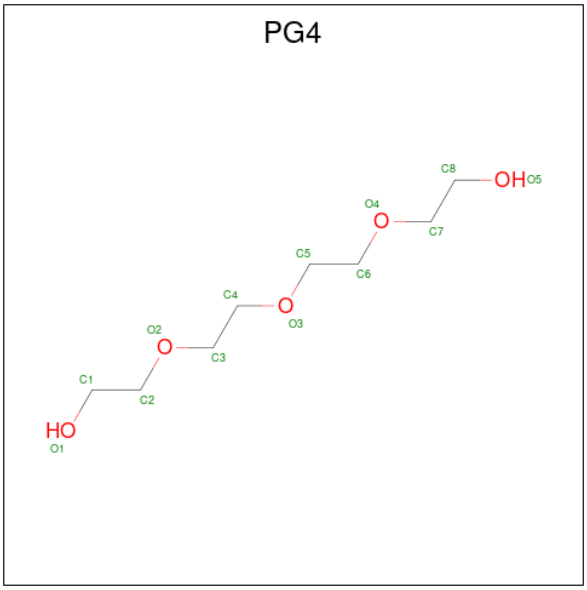
Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	expression tag	UNP O00255
A	-3	GLY	-	expression tag	UNP O00255
A	-2	SER	-	expression tag	UNP O00255
A	-1	SER	-	expression tag	UNP O00255
A	0	SER	-	expression tag	UNP O00255
A	?	-	ILE	deletion	UNP O00255
A	?	-	PRO	deletion	UNP O00255
A	?	-	THR	deletion	UNP O00255
A	?	-	ASN	deletion	UNP O00255
A	?	-	VAL	deletion	UNP O00255
A	?	-	PRO	deletion	UNP O00255
A	?	-	GLU	deletion	UNP O00255
A	?	-	LEU	deletion	UNP O00255
A	?	-	THR	deletion	UNP O00255
A	?	-	PHE	deletion	UNP O00255
A	?	-	GLN	deletion	UNP O00255
A	?	-	PRO	deletion	UNP O00255
A	?	-	SER	deletion	UNP O00255
A	?	-	PRO	deletion	UNP O00255
A	?	-	ALA	deletion	UNP O00255
A	?	-	PRO	deletion	UNP O00255
A	?	-	ASP	deletion	UNP O00255
A	?	-	PRO	deletion	UNP O00255
A	?	-	PRO	deletion	UNP O00255
A	?	-	GLY	deletion	UNP O00255
A	326	ARG	GLY	engineered mutation	UNP O00255
A	?	-	GLU	deletion	UNP O00255

Continued on next page...

Continued from previous page...

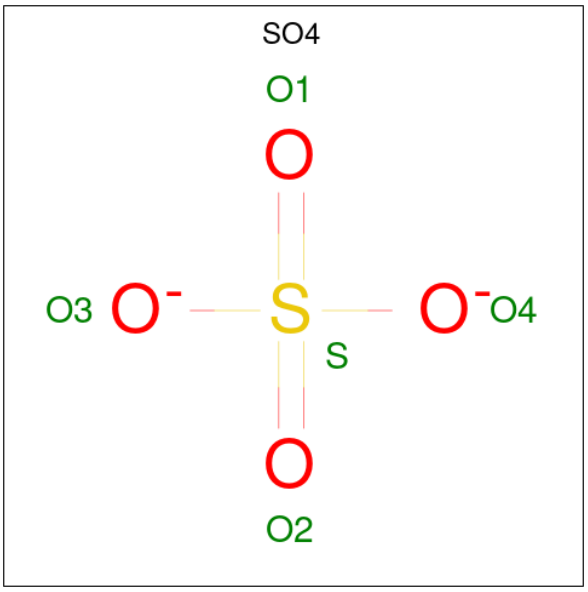
Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	GLU	deletion	UNP O00255
A	?	-	ARG	deletion	UNP O00255
A	?	-	PRO	deletion	UNP O00255
A	?	-	GLY	deletion	UNP O00255
A	?	-	GLU	deletion	UNP O00255
A	?	-	GLN	deletion	UNP O00255
A	?	-	SER	deletion	UNP O00255
A	?	-	GLN	deletion	UNP O00255
A	?	-	GLY	deletion	UNP O00255
A	?	-	THR	deletion	UNP O00255
A	?	-	GLN	deletion	UNP O00255
A	541	ALA	THR	variant	UNP O00255

- Molecule 2 is TETRAETHYLENE GLYCOL (CCD ID: PG4) (formula: $C_8H_{18}O_5$).



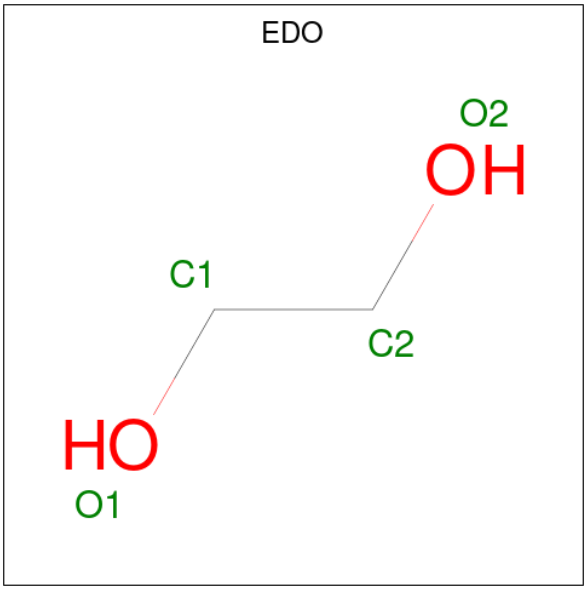
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
2	A	1	31	8	18	5	2	0

- Molecule 3 is SULFATE ION (CCD ID: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0

- Molecule 4 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C H O 10 2 6 2	2	0
4	A	1	Total C H O 10 2 6 2	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	2	0
			10	2	6	2		
4	A	1	Total	C	H	O	2	0
			10	2	6	2		
4	A	1	Total	C	H	O	2	0
			10	2	6	2		
4	A	1	Total	C	H	O	2	0
			10	2	6	2		

- Molecule 5 is water.

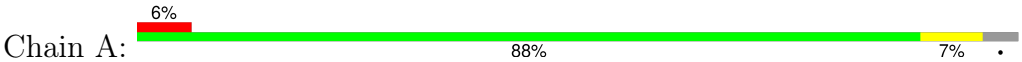
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	509	Total	O	0	0
			509	509		

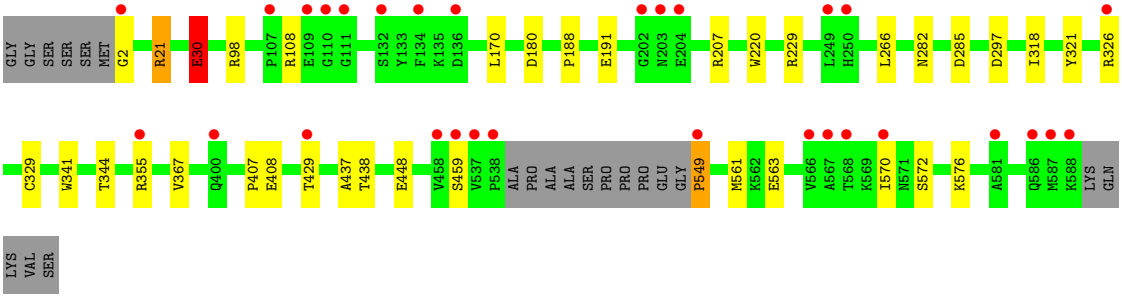
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Menin

Chain A:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	48.71Å 80.01Å 124.62Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	67.33 – 1.40 67.33 – 1.40	Depositor EDS
% Data completeness (in resolution range)	99.4 (67.33-1.40) 99.4 (67.33-1.40)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.28 (at 1.40Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.158 , 0.181 0.158 , 0.182	Depositor DCC
R_{free} test set	4937 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	15.1	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 37.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	8347	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.29% 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: SO4, PG4, EDO

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	0.71	0/4158	1.09	8/5637 (0.1%)

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	A	0	2

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	549	PRO	N-CA-CB	7.11	110.82	103.00
1	A	297	ASP	CA-CB-CG	6.61	119.21	112.60
1	A	355	ARG	N-CA-C	6.39	120.18	112.38
1	A	21	ARG	CD-NE-CZ	5.84	132.58	124.40
1	A	30	GLU	CB-CG-CD	5.74	122.35	112.60
1	A	282	ASN	CA-CB-CG	5.62	118.22	112.60
1	A	438	THR	CA-CB-OG1	-5.27	101.70	109.60
1	A	285	ASP	CA-CB-CG	5.08	117.68	112.60

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	207	ARG	Sidechain
1	A	21	ARG	Sidechain

5.2 Too-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 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	3873	3864	3623	32	0
2	A	13	18	18	0	0
3	A	10	0	0	0	0
4	A	24	36	36	1	0
5	A	509	0	0	10	0
All	All	4429	3918	3677	33	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 4.

All (33) 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:329[B]:CYS:SG	1:A:341:TRP:CH2	2.40	1.14
1:A:329[B]:CYS:SG	1:A:341:TRP:HH2	1.84	1.00
1:A:563[B]:GLU:H	1:A:563[B]:GLU:CD	1.89	0.81
1:A:448:GLU:HG2	5:A:847:HOH:O	1.81	0.81
1:A:329[B]:CYS:SG	1:A:341:TRP:CZ2	2.77	0.77
1:A:266:LEU:HD12	5:A:1153:HOH:O	1.92	0.69
1:A:408:GLU:HG2	5:A:1201:HOH:O	2.04	0.57
1:A:321:TYR:HB2	1:A:344[A]:THR:HG22	1.86	0.57
1:A:2:GLY:HA2	5:A:886:HOH:O	2.06	0.55
1:A:229[A]:ARG:HD3	5:A:1062:HOH:O	2.05	0.55
1:A:326[C]:ARG:HG3	5:A:807:HOH:O	2.07	0.55
1:A:30:GLU:CD	1:A:30:GLU:H	2.18	0.51
1:A:407:PRO:HB2	1:A:549:PRO:HG2	1.92	0.51
1:A:563[B]:GLU:CD	1:A:563[B]:GLU:N	2.63	0.51
1:A:108[A]:ARG:HD2	1:A:170:LEU:CD2	2.43	0.49
4:A:707:EDO:H11	5:A:1176:HOH:O	2.12	0.49
1:A:108[A]:ARG:HD2	1:A:170:LEU:HD22	1.94	0.48
1:A:570:ILE:HG12	5:A:1127:HOH:O	2.16	0.46
1:A:326[B]:ARG:HH22	1:A:367:VAL:HG22	1.83	0.44
1:A:180:ASP:HB2	1:A:220:TRP:CH2	2.53	0.43
1:A:326[B]:ARG:HG2	5:A:807:HOH:O	2.19	0.42
1:A:188:PRO:O	1:A:191[B]:GLU:HG2	2.20	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:437:ALA:HB1	1:A:576:LYS:HG2	2.03	0.41
1:A:572[A]:SER:O	1:A:576:LYS:HG3	2.20	0.41
1:A:108[A]:ARG:HH11	1:A:108[A]:ARG:HD3	1.71	0.40
1:A:318:ILE:HG23	1:A:344[B]:THR:HG23	2.02	0.40
1:A:98:ARG:HD2	5:A:1185:HOH:O	2.21	0.40

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	511/489 (104%)	496 (97%)	15 (3%)	0	100 100

There are no Ramachandran outliers to report.

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	426/411 (104%)	423 (99%)	3 (1%)	81 61

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	30	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	429	THR
1	A	459	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	212	ASN
1	A	260	GLN
1	A	282	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

9 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	$\#\|Z\| > 2$
4	EDO	A	707	-	3,3,3	0.12	0	2,2,2	0.30	0
4	EDO	A	709	-	3,3,3	0.14	0	2,2,2	0.16	0
4	EDO	A	705	-	3,3,3	0.07	0	2,2,2	0.12	0
4	EDO	A	706	-	3,3,3	0.12	0	2,2,2	0.29	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	EDO	A	708	-	3,3,3	0.17	0	2,2,2	0.31	0
3	SO4	A	702	-	4,4,4	0.31	0	6,6,6	0.07	0
2	PG4	A	701	-	12,12,12	0.14	0	11,11,11	0.17	0
4	EDO	A	703	-	3,3,3	0.15	0	2,2,2	0.20	0
3	SO4	A	704	-	4,4,4	0.32	0	6,6,6	0.59	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
4	EDO	A	707	-	-	0/1/1/1	-
4	EDO	A	709	-	-	0/1/1/1	-
4	EDO	A	705	-	-	1/1/1/1	-
4	EDO	A	706	-	-	1/1/1/1	-
4	EDO	A	708	-	-	1/1/1/1	-
2	PG4	A	701	-	-	1/10/10/10	-
4	EDO	A	703	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	701	PG4	O1-C1-C2-O2
4	A	706	EDO	O1-C1-C2-O2
4	A	708	EDO	O1-C1-C2-O2
4	A	705	EDO	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	707	EDO	1	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	468/489 (95%)	0.08	30 (6%) 27 27	5, 15, 32, 44	23 (4%)

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	537	VAL	6.1
1	A	538	PRO	6.0
1	A	326[A]	ARG	5.0
1	A	581	ALA	4.3
1	A	136	ASP	4.1
1	A	587	MET	3.9
1	A	111	GLY	3.3
1	A	134	PHE	3.2
1	A	203	ASN	3.0
1	A	566	VAL	3.0
1	A	429	THR	2.9
1	A	109[A]	GLU	2.8
1	A	568	THR	2.8
1	A	110	GLY	2.8
1	A	202	GLY	2.8
1	A	249	LEU	2.6
1	A	2	GLY	2.6
1	A	567	ALA	2.5
1	A	204	GLU	2.5
1	A	588	LYS	2.5
1	A	458	VAL	2.4
1	A	570	ILE	2.4
1	A	586	GLN	2.4
1	A	355	ARG	2.3
1	A	400	GLN	2.3
1	A	107	PRO	2.3
1	A	250	HIS	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	132	SER	2.2
1	A	459	SER	2.2
1	A	549	PRO	2.1

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 monosaccharides 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^{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	SO4	A	702	5/5	0.75	0.16	65,69,73,74	0
4	EDO	A	707	4/4	0.76	0.14	44,45,48,48	2
4	EDO	A	708	4/4	0.78	0.17	45,51,53,53	2
4	EDO	A	709	4/4	0.79	0.14	39,42,47,50	2
4	EDO	A	706	4/4	0.81	0.17	46,47,49,49	2
3	SO4	A	704	5/5	0.86	0.15	27,29,34,47	0
4	EDO	A	705	4/4	0.86	0.10	37,38,40,40	2
4	EDO	A	703	4/4	0.89	0.16	38,41,44,44	2
2	PG4	A	701	13/13	0.95	0.08	19,23,36,39	2

6.5 Other polymers (i)

There are no such residues in this entry.