



# Full wwPDB X-ray Structure Validation Report ⓘ

May 5, 2025 – 08:55 PM EDT

PDB ID : 9C4Z / pdb\_00009c4z  
Title : Menin mutant - G331D  
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](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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) 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

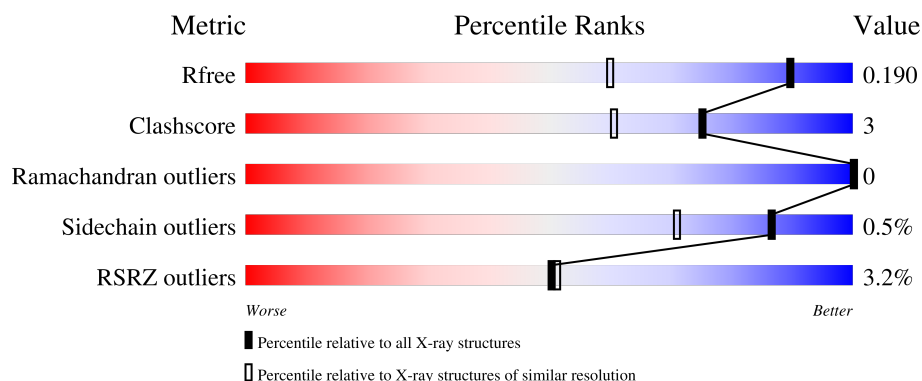
# 1 Overall quality at a glance

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  $\geq 3$ , 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	489	

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 8120 atoms, of which 3797 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
1	A	467	Total	C	H	N	O	S	130	29	0
			7515	2408	3745	641	705	16			

There are 39 discrepancies between the modelled and reference sequences:

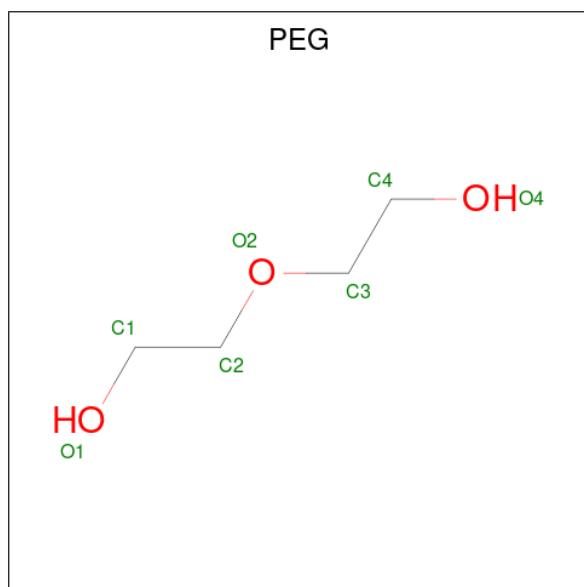
Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	expression tag	UNP O00255
A	-3	GLY	-	expression tag	UNP O00255
A	-2	SER	-	expression tag	UNP O00255
A	-1	SER	-	expression tag	UNP O00255
A	0	SER	-	expression tag	UNP O00255
A	?	-	ILE	deletion	UNP O00255
A	?	-	PRO	deletion	UNP O00255
A	?	-	THR	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	ASP	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 DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula:  $C_4H_{10}O_3$ ).



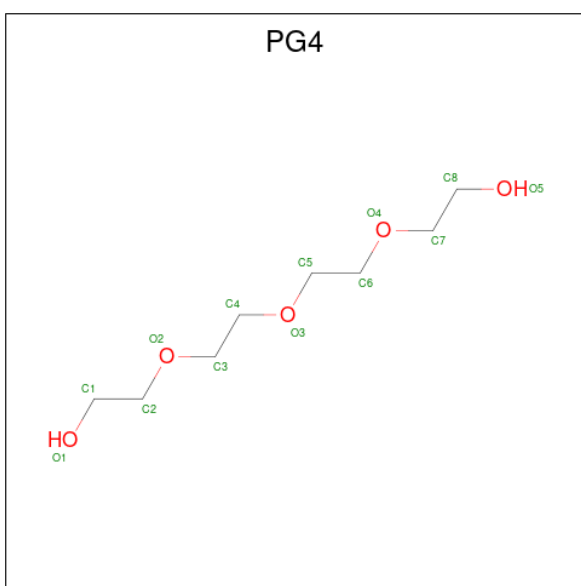
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	O	2	0
			17	4	10	3		

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



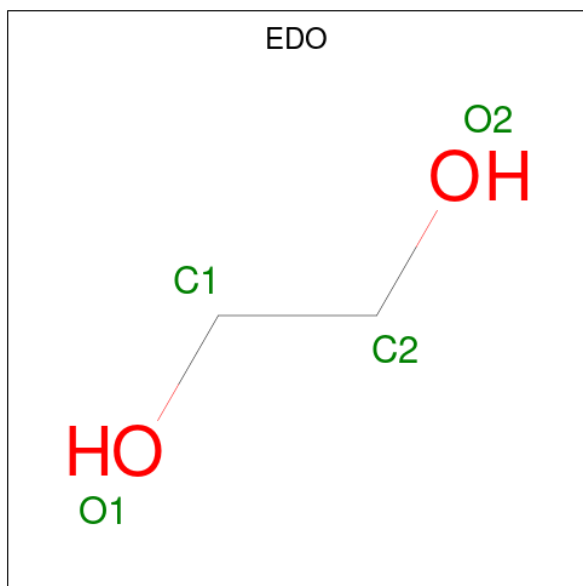
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is TETRAETHYLENE GLYCOL (CCD ID: PG4) (formula:  $C_8H_{18}O_5$ ).



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

- Molecule 5 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula:  $C_2H_6O_2$ ).



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

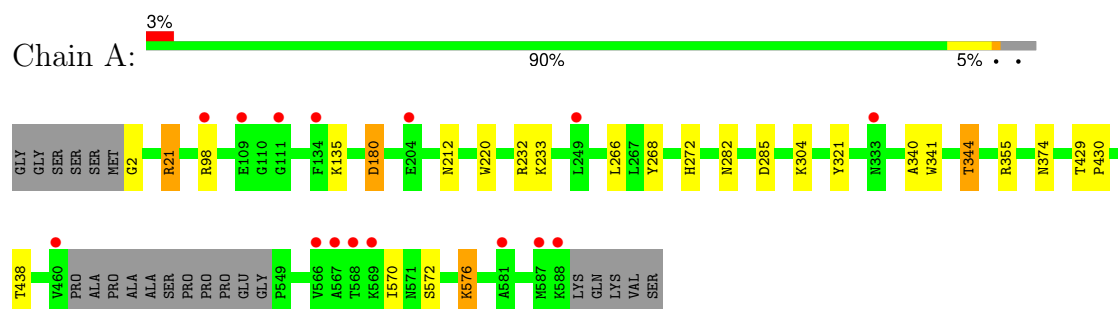
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	506	Total	O	0	0
			506	506		

### 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



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.73Å 79.94Å 124.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.96 – 1.40 24.96 – 1.40	Depositor EDS
% Data completeness (in resolution range)	99.7 (24.96-1.40) 99.7 (24.96-1.40)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.31 (at 1.40Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, $R_{free}$	0.165 , 0.190 0.166 , 0.190	Depositor DCC
$R_{free}$ test set	4820 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	15.3	Xtriage
Anisotropy	0.002	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.43 , 40.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8120	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.19% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PG4, PEG, SO4, 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	0.77	1/3970 (0.0%)	1.15	9/5390 (0.2%)

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	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	232	ARG	NE-CZ	-5.86	1.26	1.33

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	438	THR	CA-CB-OG1	-8.33	97.10	109.60
1	A	282	ASN	CA-CB-CG	7.29	119.89	112.60
1	A	355	ARG	N-CA-C	6.45	120.25	112.38
1	A	285	ASP	CA-CB-CG	5.73	118.33	112.60
1	A	21	ARG	CD-NE-CZ	5.59	132.22	124.40
1	A	180	ASP	CA-CB-CG	5.30	117.91	112.60
1	A	98	ARG	CD-NE-CZ	5.19	131.67	124.40
1	A	212	ASN	OD1-CG-ND2	-5.16	117.44	122.60
1	A	576	LYS	N-CA-CB	5.07	117.36	110.01

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	21	ARG	Sidechain

## 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	3770	3745	3569	21	0
2	A	7	10	10	1	0
3	A	10	0	0	0	0
4	A	26	36	36	1	0
5	A	4	6	6	1	0
6	A	506	0	0	4	0
All	All	4323	3797	3621	21	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 3.

All (21) 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:341:TRP:O	1:A:344[B]:THR:HG22	1.36	1.22
1:A:321:TYR:CB	1:A:344[A]:THR:HG22	2.19	0.72
1:A:321:TYR:HB2	1:A:344[A]:THR:HG22	1.74	0.70
1:A:340:ALA:O	1:A:344[A]:THR:HG23	1.92	0.69
1:A:266:LEU:HD12	6:A:1054:HOH:O	1.97	0.63
1:A:341:TRP:C	1:A:344[B]:THR:HG22	2.21	0.63
1:A:572:SER:O	1:A:576:LYS:HG2	2.01	0.60
1:A:341:TRP:O	1:A:344[B]:THR:CG2	2.31	0.58
1:A:321:TYR:HB3	1:A:344[A]:THR:HG22	1.87	0.55
1:A:268:TYR:OH	1:A:304:LYS:HE3	2.13	0.48
1:A:570:ILE:HG12	6:A:1053:HOH:O	2.15	0.47
1:A:374:ASN:HB3	6:A:1083:HOH:O	2.14	0.46
1:A:572:SER:OG	1:A:576:LYS:HE3	2.17	0.43
1:A:2:GLY:HA2	6:A:867:HOH:O	2.18	0.43
1:A:233:LYS:HG3	1:A:272:HIS:CE1	2.54	0.43
1:A:430:PRO:HB2	5:A:606:EDO:H21	2.00	0.42
1:A:272:HIS:CE1	4:A:603:PG4:H31	2.55	0.42

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:135:LYS:HE2	2:A:601:PEG:H11	2.01	0.41
1:A:180:ASP:HB2	1:A:220:TRP:CH2	2.56	0.41

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	492/489 (101%)	482 (98%)	10 (2%)	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	408/411 (99%)	405 (99%)	3 (1%)	81	61

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

Mol	Chain	Res	Type
1	A	344[A]	THR
1	A	344[B]	THR
1	A	429	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	212	ASN
1	A	260	GLN
1	A	282	ASN
1	A	339	GLN
1	A	584	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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

6 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	$\# Z  > 2$
5	EDO	A	606	-	3,3,3	0.11	0	2,2,2	0.42	0
3	SO4	A	602	-	4,4,4	0.32	0	6,6,6	0.06	0
2	PEG	A	601	-	6,6,6	0.17	0	5,5,5	0.15	0
3	SO4	A	605	-	4,4,4	0.40	0	6,6,6	0.66	0
4	PG4	A	604	-	12,12,12	0.39	0	11,11,11	0.50	0
4	PG4	A	603	-	12,12,12	0.56	0	11,11,11	0.48	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	PG4	A	604	-	-	5/10/10/10	-
4	PG4	A	603	-	-	7/10/10/10	-
5	EDO	A	606	-	-	1/1/1/1	-
2	PEG	A	601	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	603	PG4	O2-C3-C4-O3
2	A	601	PEG	O1-C1-C2-O2
4	A	604	PG4	O4-C7-C8-O5
4	A	603	PG4	C4-C3-O2-C2
4	A	604	PG4	O3-C5-C6-O4
4	A	604	PG4	C4-C3-O2-C2
4	A	603	PG4	C1-C2-O2-C3
4	A	603	PG4	O4-C7-C8-O5
4	A	603	PG4	C8-C7-O4-C6
2	A	601	PEG	C4-C3-O2-C2
4	A	604	PG4	C8-C7-O4-C6
4	A	604	PG4	O2-C3-C4-O3
5	A	606	EDO	O1-C1-C2-O2
4	A	603	PG4	O3-C5-C6-O4
4	A	603	PG4	C3-C4-O3-C5

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	606	EDO	1	0
2	A	601	PEG	1	0
4	A	603	PG4	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 [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	467/489 (95%)	-0.22	15 (3%)	50 51	5, 15, 35, 48	15 (3%)

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	460	VAL	6.5
1	A	111	GLY	3.8
1	A	581	ALA	3.6
1	A	587	MET	3.4
1	A	566	VAL	3.3
1	A	109	GLU	3.1
1	A	333[A]	ASN	3.1
1	A	204	GLU	2.8
1	A	134	PHE	2.4
1	A	569	LYS	2.4
1	A	567	ALA	2.2
1	A	588	LYS	2.1
1	A	568	THR	2.0
1	A	249	LEU	2.0
1	A	98	ARG	2.0

### 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<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(Å <sup>2</sup> )	Q<0.9
4	PG4	A	603	13/13	0.63	0.17	33,37,42,43	2
3	SO4	A	602	5/5	0.81	0.14	69,69,77,79	0
5	EDO	A	606	4/4	0.81	0.12	36,40,49,49	2
3	SO4	A	605	5/5	0.86	0.14	26,26,33,43	0
2	PEG	A	601	7/7	0.87	0.12	42,46,57,57	2
4	PG4	A	604	13/13	0.90	0.10	23,29,38,43	2

## 6.5 Other polymers [i](#)

There are no such residues in this entry.