



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2026 – 02:51 AM UTC

PDB ID : 9H6H / pdb\_00009h6h  
Title : Bacteroides ovatus GH98 endoxylanase with Seleno-methionine substituents  
Authors : Tomlinson, C.W.E.; Cartmell, A.; Bolam, D.N.  
Deposited on : 2024-10-24  
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](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 Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

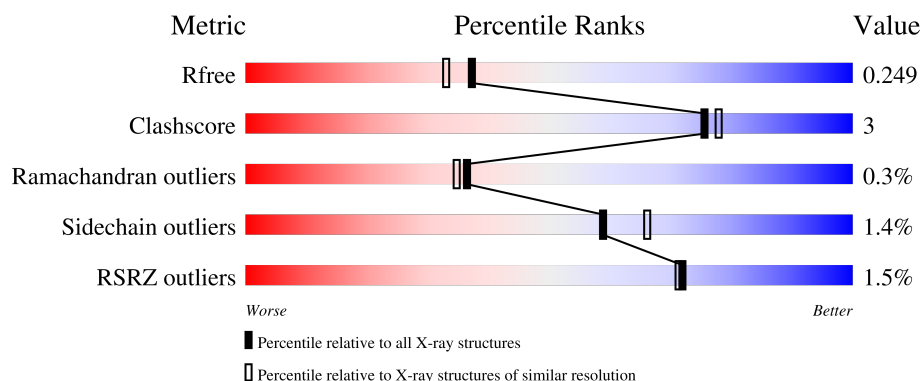
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


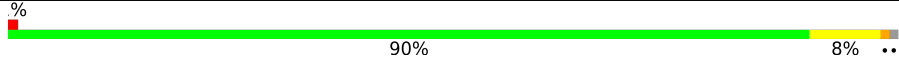
The reported resolution of this entry is 2.00 Å.

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))
$R_{free}$	180053	10052 (2.00-2.00)
Clashscore	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (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 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	906	 2% 89% 9% ..
1	B	906	 % 90% 8% ..

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 29357 atoms, of which 13772 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 Glycosyl hydrolase family 98.

Mol	Chain	Residues	Atoms								ZeroOcc	AltConf	Tra
1	A	897	Total	C	H	N	O	P	S	Se	217	6	0
			14013	4539	6861	1189	1399	2	14	9			
1	B	894	Total	C	H	N	O	P	S	Se	214	2	0
			13962	4524	6841	1187	1386	1	14	9			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	42	MSE	-	initiating methionine	UNP A0A6N3VET1
A	43	ALA	-	expression tag	UNP A0A6N3VET1
A	44	SER	-	expression tag	UNP A0A6N3VET1
A	353	ALA	GLU	engineered mutation	UNP A0A6N3VET1
A	940	LEU	-	expression tag	UNP A0A6N3VET1
A	941	GLU	-	expression tag	UNP A0A6N3VET1
A	942	HIS	-	expression tag	UNP A0A6N3VET1
A	943	HIS	-	expression tag	UNP A0A6N3VET1
A	944	HIS	-	expression tag	UNP A0A6N3VET1
A	945	HIS	-	expression tag	UNP A0A6N3VET1
A	946	HIS	-	expression tag	UNP A0A6N3VET1
A	947	HIS	-	expression tag	UNP A0A6N3VET1
B	42	MSE	-	initiating methionine	UNP A0A6N3VET1
B	43	ALA	-	expression tag	UNP A0A6N3VET1
B	44	SER	-	expression tag	UNP A0A6N3VET1
B	353	ALA	GLU	engineered mutation	UNP A0A6N3VET1
B	940	LEU	-	expression tag	UNP A0A6N3VET1
B	941	GLU	-	expression tag	UNP A0A6N3VET1
B	942	HIS	-	expression tag	UNP A0A6N3VET1
B	943	HIS	-	expression tag	UNP A0A6N3VET1
B	944	HIS	-	expression tag	UNP A0A6N3VET1
B	945	HIS	-	expression tag	UNP A0A6N3VET1
B	946	HIS	-	expression tag	UNP A0A6N3VET1
B	947	HIS	-	expression tag	UNP A0A6N3VET1

- 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		
2	A	1	Total	C	H	O	2	0
			17	4	10	3		
2	B	1	Total	C	H	O	2	0
			17	4	10	3		
2	B	1	Total	C	H	O	2	0
			17	4	10	3		
2	B	1	Total	C	H	O	2	0
			17	4	10	3		
2	B	1	Total	C	H	O	2	0
			17	4	10	3		

- Molecule 3 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ca	0	0
			1	1		
3	B	1	Total	Ca	0	0
			1	1		

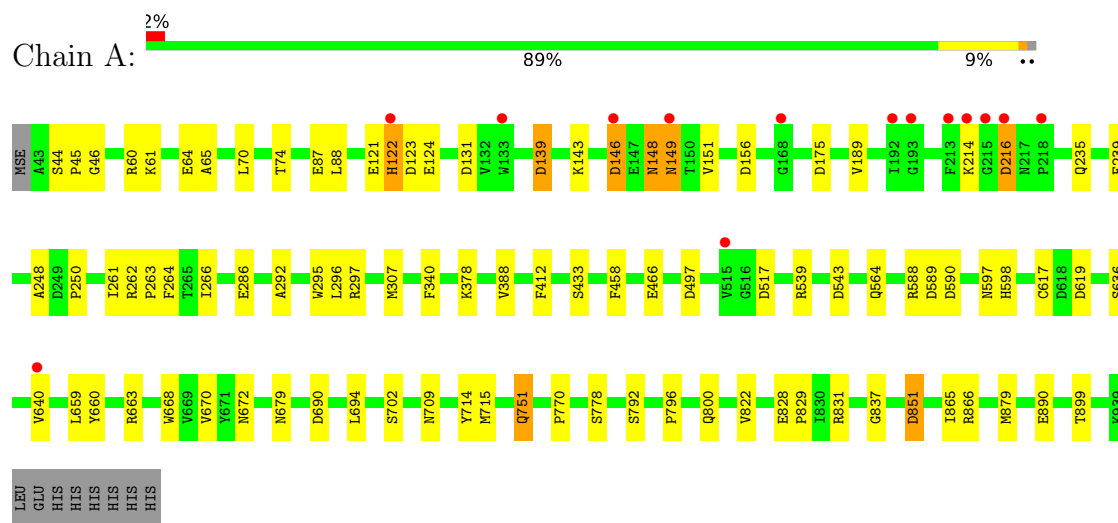
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	564	Total 564	O 564	0	0
4	B	697	Total 697	O 697	0	0

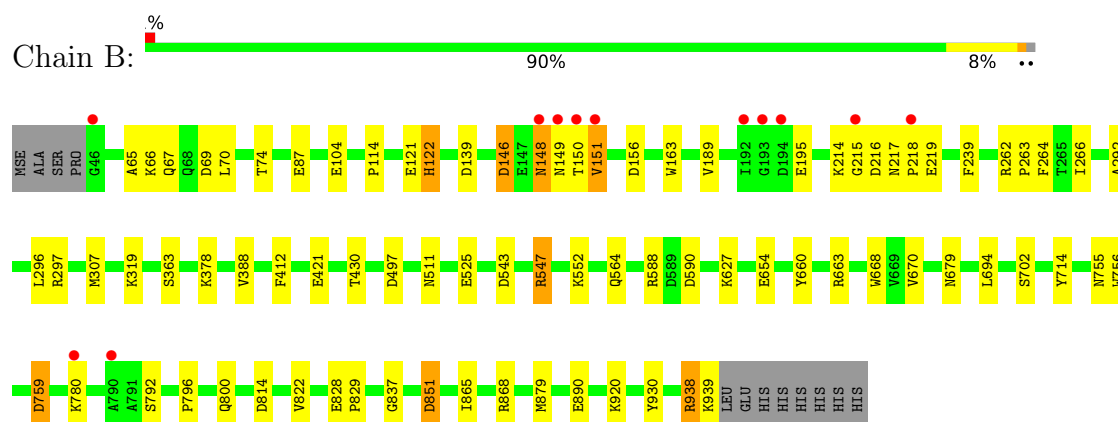
### 3 Residue-property plots [i](#)

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: Glycosyl hydrolase family 98



#### • Molecule 1: Glycosyl hydrolase family 98



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	183.91Å 62.45Å 178.71Å 90.00° 118.06° 90.00°	Depositor
Resolution (Å)	58.40 – 2.00 58.40 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.6 (58.40-2.00) 99.6 (58.40-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.26 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0430 (refmacat 0.4.82)	Depositor
R, $R_{free}$	0.192 , 0.242 0.199 , 0.249	Depositor DCC
$R_{free}$ test set	6155 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.9	Xtriage
Anisotropy	0.164	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 46.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.003 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	29357	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.20% 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: PEG, CA, PHD

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	0.63	0/7318	1.07	15/9940 (0.2%)
1	B	0.66	0/7286	1.07	11/9897 (0.1%)
All	All	0.64	0/14604	1.07	26/19837 (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	4
1	B	0	4
All	All	0	8

There are no bond length outliers.

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	412	PHE	CA-CB-CG	-7.09	106.71	113.80
1	A	412	PHE	CA-CB-CG	-7.05	106.75	113.80
1	B	146	ASP	CA-CB-CG	7.00	119.60	112.60
1	A	800	GLN	CB-CA-C	6.99	121.32	109.72
1	A	590	ASP	CA-CB-CG	6.63	119.23	112.60
1	B	800	GLN	CB-CA-C	6.43	120.39	109.72
1	B	814	ASP	CA-CB-CG	6.03	118.63	112.60
1	A	497	ASP	CA-CB-CG	5.93	118.53	112.60
1	B	590	ASP	CA-CB-CG	5.91	118.51	112.60
1	B	497	ASP	CA-CB-CG	5.88	118.48	112.60
1	A	543	ASP	CA-CB-CG	5.82	118.42	112.60
1	B	195	GLU	CB-CG-CD	5.53	122.00	112.60

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	156	ASP	CA-CB-CG	5.47	118.07	112.60
1	B	104	GLU	N-CA-CB	5.45	118.05	110.26
1	A	146	ASP	CA-CB-CG	5.42	118.02	112.60
1	A	619	ASP	CA-CB-CG	5.40	118.00	112.60
1	A	175	ASP	CA-CB-CG	5.38	117.98	112.60
1	B	543	ASP	CA-CB-CG	5.25	117.85	112.60
1	A	340	PHE	N-CA-CB	5.24	117.92	110.16
1	A	690	ASP	CA-CB-CG	5.22	117.82	112.60
1	A	866	ARG	NE-CZ-NH1	-5.21	116.29	121.50
1	A	458	PHE	CA-CB-CG	-5.20	108.60	113.80
1	B	430	THR	CA-CB-OG1	-5.10	101.95	109.60
1	A	131	ASP	CA-CB-CG	5.09	117.69	112.60
1	A	617	CYS	CB-CA-C	-5.06	100.05	110.38
1	A	589	ASP	CA-CB-CG	5.04	117.64	112.60

There are no chirality outliers.

All (8) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	248	ALA	Peptide
1	A	297	ARG	Sidechain
1	A	588	ARG	Sidechain
1	A	60	ARG	Sidechain
1	B	297	ARG	Sidechain
1	B	547[A]	ARG	Sidechain
1	B	588	ARG	Sidechain
1	B	938	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	7152	6861	6825	42	2
1	B	7121	6841	6817	43	2
2	A	14	20	20	0	0
2	B	35	50	50	1	0
3	A	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	1	0	0	0	0
4	A	564	0	0	8	1
4	B	697	0	0	15	1
All	All	15585	13772	13712	85	4

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 (85) 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:851[A]:PHD:OP1	4:A:1101:HOH:O	1.80	0.97
1:B:547[A]:ARG:NH1	4:B:1101:HOH:O	2.09	0.84
1:B:146:ASP:OD1	1:B:149:ASN:HB2	1.84	0.77
1:B:150:THR:O	1:B:151:VAL:HG13	1.91	0.71
1:B:938:ARG:NH1	1:B:939:LYS:O	2.25	0.70
1:A:851[A]:PHD:P	4:A:1101:HOH:O	2.49	0.69
1:A:865:ILE:HD11	1:A:879:MSE:CE	2.23	0.67
1:B:865:ILE:HD11	1:B:879:MSE:CE	2.25	0.67
1:A:149:ASN:HB2	1:A:151:VAL:HG22	1.78	0.65
1:B:139:ASP:O	4:B:1102:HOH:O	2.15	0.64
1:B:547[B]:ARG:NH2	4:B:1107:HOH:O	2.31	0.63
1:A:851[B]:PHD:OP1	4:A:1101:HOH:O	2.15	0.62
1:A:751:GLN:NE2	1:B:755:ASN:OD1	2.34	0.61
1:B:525:GLU:HG3	4:B:1113:HOH:O	2.04	0.56
1:A:433[B]:SER:OG	1:A:598:HIS:CE1	2.59	0.56
1:A:292:ALA:O	1:A:296:LEU:HB2	2.06	0.55
1:A:663:ARG:HD3	1:A:668:TRP:CZ2	2.41	0.55
1:B:627:LYS:HE2	4:B:1532:HOH:O	2.07	0.54
1:B:307:MSE:HE1	1:B:388:VAL:HG23	1.89	0.54
1:A:61:LYS:HE3	1:A:88:LEU:HD12	1.90	0.54
1:B:121:GLU:O	1:B:122:HIS:C	2.50	0.54
1:B:663:ARG:HD3	1:B:668:TRP:CZ2	2.43	0.54
1:B:851:PHD:P	4:B:1104:HOH:O	2.66	0.53
1:A:44:SER:N	1:A:45:PRO:CD	2.72	0.53
1:A:121:GLU:O	1:A:122:HIS:C	2.52	0.53
1:B:292:ALA:O	1:B:296:LEU:HB2	2.08	0.52
1:A:640:VAL:N	4:A:1118:HOH:O	2.42	0.51
1:B:822:VAL:O	1:B:837:GLY:HA3	2.11	0.51
1:A:286:GLU:OE2	4:A:1102:HOH:O	2.19	0.51
1:B:217:ASN:HB2	1:B:218:PRO:CD	2.41	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:264:PHE:CE2	1:A:796:PRO:HD3	2.48	0.49
1:A:307:MSE:HE1	1:A:388:VAL:HG23	1.94	0.49
1:B:264:PHE:CE2	1:B:796:PRO:HD3	2.48	0.49
1:B:525:GLU:CG	4:B:1113:HOH:O	2.59	0.48
1:A:822:VAL:O	1:A:837:GLY:HA3	2.13	0.48
1:B:868:ARG:HB2	1:B:930:TYR:HB2	1.95	0.48
1:A:235:GLN:CD	4:A:1151:HOH:O	2.56	0.48
1:A:148:ASN:N	1:A:148:ASN:OD1	2.47	0.47
1:B:654:GLU:OE2	1:B:663:ARG:NE	2.43	0.47
1:A:890:GLU:HA	1:A:890:GLU:OE1	2.14	0.47
1:A:433[B]:SER:HG	1:A:598:HIS:CE1	2.31	0.47
1:B:660:TYR:O	1:B:670:VAL:HA	2.14	0.46
1:A:146:ASP:HB3	1:A:149:ASN:OD1	2.14	0.46
1:B:148:ASN:N	1:B:148:ASN:OD1	2.48	0.46
1:B:920:LYS:HE3	4:B:1615:HOH:O	2.15	0.46
1:B:890:GLU:HA	1:B:890:GLU:OE1	2.16	0.45
1:A:261:ILE:HD11	1:A:539:ARG:HG3	1.97	0.45
1:B:756:TRP:CH2	1:B:759:ASP:HA	2.52	0.44
1:B:87:GLU:OE1	4:B:1103:HOH:O	2.20	0.44
1:A:660:TYR:O	1:A:670:VAL:HA	2.17	0.44
1:A:149:ASN:CB	1:A:151:VAL:HG22	2.46	0.44
1:B:239:PHE:HE1	1:B:266:ILE:HD12	1.83	0.44
1:A:239:PHE:HE1	1:A:266:ILE:HD12	1.84	0.43
1:A:828:GLU:HB3	1:A:829:PRO:CD	2.49	0.43
1:B:65:ALA:HB1	1:B:70:LEU:HB2	2.01	0.43
1:A:679:ASN:HA	1:A:694:LEU:O	2.18	0.43
1:B:828:GLU:HB3	1:B:829:PRO:CD	2.49	0.43
1:A:214:LYS:O	1:A:216:ASP:N	2.51	0.43
1:A:61:LYS:NZ	1:A:64:GLU:OE1	2.45	0.42
1:A:831:ARG:HB2	1:A:899:THR:HG22	2.01	0.42
1:B:552:LYS:HE3	4:B:1728:HOH:O	2.18	0.42
1:A:235:GLN:NE2	4:A:1151:HOH:O	2.52	0.42
1:B:679:ASN:HA	1:B:694:LEU:O	2.19	0.42
1:A:65:ALA:HB1	1:A:70:LEU:HB2	2.02	0.42
1:B:214:LYS:O	1:B:216:ASP:N	2.52	0.42
2:B:1004:PEG:C2	4:B:1132:HOH:O	2.67	0.42
1:A:702:SER:HA	1:A:714:TYR:O	2.19	0.42
1:A:262:ARG:N	1:A:263:PRO:CD	2.83	0.42
1:B:114:PRO:HD3	1:B:163:TRP:CD2	2.54	0.42
1:A:124:GLU:OE1	1:A:143:LYS:HE3	2.20	0.42
1:A:715:MSE:O	1:A:770:PRO:HA	2.20	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:149:ASN:HA	4:B:1496:HOH:O	2.19	0.42
1:A:851[B]:PHD:P	4:A:1101:HOH:O	2.77	0.41
1:B:149:ASN:OD1	1:B:151:VAL:N	2.54	0.41
1:B:363:SER:C	4:B:1188:HOH:O	2.64	0.41
1:A:709:ASN:OD1	1:A:778:SER:HB3	2.21	0.41
1:B:74:THR:HG21	4:B:1103:HOH:O	2.21	0.41
1:A:250:PRO:HG3	1:A:295:TRP:CD1	2.56	0.41
1:B:421:GLU:HG3	4:B:1426:HOH:O	2.21	0.41
1:A:74:THR:HG22	1:A:87:GLU:HB3	2.03	0.40
1:B:865:ILE:HD11	1:B:879:MSE:HE2	2.01	0.40
1:B:262:ARG:N	1:B:263:PRO:CD	2.84	0.40
1:A:659:LEU:CD2	1:A:672:ASN:HB2	2.51	0.40
1:B:702:SER:HA	1:B:714:TYR:O	2.21	0.40
1:B:654:GLU:OE2	1:B:663:ARG:NH2	2.53	0.40

All (4) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:378:LYS:HZ1	1:A:792:SER:HG[4_647]	1.25	0.35
1:B:378:LYS:HZ1	1:B:792:SER:HG[4_546]	1.28	0.32
1:A:156:ASP:OD1	1:B:511:ASN:ND2[3_545]	2.06	0.14
4:A:1627:HOH:O	4:B:1702:HOH:O[3_555]	2.17	0.03

## 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	899/906 (99%)	861 (96%)	35 (4%)	3 (0%)	36	35
1	B	893/906 (99%)	854 (96%)	36 (4%)	3 (0%)	36	35
All	All	1792/1812 (99%)	1715 (96%)	71 (4%)	6 (0%)	36	35

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	215	GLY
1	A	139	ASP
1	B	122	HIS
1	A	122	HIS
1	A	46	GLY
1	B	151	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	776/772 (100%)	764 (98%)	12 (2%)	57	64
1	B	773/772 (100%)	763 (99%)	10 (1%)	61	68
All	All	1549/1544 (100%)	1527 (99%)	22 (1%)	59	66

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	123	ASP
1	A	139	ASP
1	A	148	ASN
1	A	149	ASN
1	A	189	VAL
1	A	216	ASP
1	A	466	GLU
1	A	517	ASP
1	A	564	GLN
1	A	597	ASN
1	A	636	SER
1	A	751	GLN
1	B	66	LYS
1	B	67	GLN
1	B	69	ASP
1	B	148	ASN
1	B	189	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	219	GLU
1	B	319	LYS
1	B	564	GLN
1	B	759	ASP
1	B	780	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	68	GLN
1	A	246	ASN
1	A	251	GLN
1	A	302	ASN
1	A	523	ASN
1	A	598	HIS
1	A	622	ASN
1	A	679	ASN
1	A	800	GLN
1	A	817	ASN
1	B	99	GLN
1	B	251	GLN
1	B	302	ASN
1	B	504	GLN
1	B	598	HIS
1	B	840	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

3 non-standard protein/DNA/RNA residues 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
1	PHD	A	851[B]	1	9,11,12	1.02	1 (11%)	9,15,17	1.23	1 (11%)
1	PHD	B	851	1	9,11,12	1.76	2 (22%)	9,15,17	1.08	0
1	PHD	A	851[A]	1	9,11,12	1.21	1 (11%)	9,15,17	1.38	2 (22%)

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
1	PHD	A	851[B]	1	-	4/8/11/13	-
1	PHD	B	851	1	-	5/8/11/13	-
1	PHD	A	851[A]	1	-	5/8/11/13	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	851	PHD	P-OP1	3.69	1.62	1.50
1	A	851[A]	PHD	P-OD1	3.23	1.65	1.59
1	B	851	PHD	P-OD1	2.79	1.64	1.59
1	A	851[B]	PHD	P-OD1	2.60	1.64	1.59

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	851[B]	PHD	OP3-P-OD1	3.02	114.12	105.32
1	A	851[A]	PHD	OD1-CG-CB	2.37	116.52	110.95
1	A	851[A]	PHD	CA-CB-CG	2.13	117.34	112.78

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	851[A]	PHD	CG-OD1-P-OP1
1	A	851[B]	PHD	C-CA-CB-CG
1	B	851	PHD	C-CA-CB-CG
1	A	851[B]	PHD	CA-CB-CG-OD2
1	A	851[B]	PHD	N-CA-CB-CG
1	A	851[B]	PHD	CA-CB-CG-OD1
1	B	851	PHD	CA-CB-CG-OD1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
1	B	851	PHD	CA-CB-CG-OD2
1	A	851[A]	PHD	CA-CB-CG-OD2
1	A	851[A]	PHD	CA-CB-CG-OD1
1	B	851	PHD	N-CA-CB-CG
1	A	851[A]	PHD	C-CA-CB-CG
1	B	851	PHD	CG-OD1-P-OP2
1	A	851[A]	PHD	N-CA-CB-CG

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	851[B]	PHD	2	0
1	B	851	PHD	1	0
1	A	851[A]	PHD	2	0

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 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	$\# Z  > 2$
2	PEG	B	1005	-	6,6,6	0.35	0	5,5,5	0.22	0
2	PEG	B	1001	-	6,6,6	0.26	0	5,5,5	0.19	0
2	PEG	B	1002	-	6,6,6	0.15	0	5,5,5	0.14	0
2	PEG	B	1003	-	6,6,6	0.18	0	5,5,5	0.22	0
2	PEG	A	1002	-	6,6,6	0.22	0	5,5,5	0.22	0
2	PEG	A	1001	-	6,6,6	0.17	0	5,5,5	0.07	0
2	PEG	B	1004	-	6,6,6	0.21	0	5,5,5	0.14	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	PEG	B	1005	-	-	3/4/4/4	-
2	PEG	B	1001	-	-	3/4/4/4	-
2	PEG	B	1002	-	-	3/4/4/4	-
2	PEG	B	1003	-	-	3/4/4/4	-
2	PEG	A	1002	-	-	3/4/4/4	-
2	PEG	A	1001	-	-	0/4/4/4	-
2	PEG	B	1004	-	-	3/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (18) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1003	PEG	O1-C1-C2-O2
2	B	1005	PEG	O2-C3-C4-O4
2	A	1002	PEG	O1-C1-C2-O2
2	A	1002	PEG	O2-C3-C4-O4
2	B	1001	PEG	O2-C3-C4-O4
2	B	1002	PEG	O1-C1-C2-O2
2	B	1003	PEG	O2-C3-C4-O4
2	B	1004	PEG	O1-C1-C2-O2
2	B	1004	PEG	O2-C3-C4-O4
2	B	1005	PEG	O1-C1-C2-O2
2	B	1001	PEG	O1-C1-C2-O2
2	B	1002	PEG	O2-C3-C4-O4
2	B	1004	PEG	C4-C3-O2-C2
2	B	1003	PEG	C4-C3-O2-C2
2	B	1005	PEG	C4-C3-O2-C2
2	A	1002	PEG	C4-C3-O2-C2
2	B	1001	PEG	C1-C2-O2-C3
2	B	1002	PEG	C4-C3-O2-C2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1004	PEG	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<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	887/906 (97%)	0.04	14 (1%) 70 70	10, 28, 53, 104	2 (0%)
1	B	884/906 (97%)	-0.13	12 (1%) 73 73	12, 25, 47, 103	1 (0%)
All	All	1771/1812 (97%)	-0.05	26 (1%) 72 71	10, 26, 51, 104	3 (0%)

All (26) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	150	THR	5.1
1	B	46	GLY	3.9
1	A	215	GLY	3.3
1	B	148	ASN	3.2
1	B	149	ASN	3.2
1	A	193	GLY	3.1
1	B	151	VAL	3.0
1	B	780	LYS	3.0
1	B	193	GLY	3.0
1	B	194	ASP	2.9
1	B	218	PRO	2.9
1	A	213	PHE	2.9
1	B	215	GLY	2.7
1	A	168	GLY	2.6
1	A	146	ASP	2.5
1	A	122	HIS	2.5
1	A	192	ILE	2.5
1	A	218	PRO	2.5
1	B	192	ILE	2.4
1	A	133	TRP	2.3
1	A	214	LYS	2.2
1	A	216	ASP	2.2
1	B	790	ALA	2.1
1	A	149	ASN	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	515	VAL	2.1
1	A	640	VAL	2.1

## 6.2 Non-standard residues in protein, DNA, RNA chains [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
1	PHD	A	851[A]	12/13	0.83	0.10	23,27,27,28	16
1	PHD	A	851[B]	12/13	0.83	0.10	26,28,29,30	16
1	PHD	B	851	12/13	0.88	0.10	20,22,37,42	0

## 6.3 Carbohydrates [i](#)

There are no oligosaccharides 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
2	PEG	B	1005	7/7	0.77	0.19	30,58,64,65	2
2	PEG	A	1001	7/7	0.78	0.16	30,64,68,69	2
2	PEG	B	1004	7/7	0.80	0.14	30,54,56,56	2
2	PEG	B	1001	7/7	0.80	0.13	30,47,50,50	2
2	PEG	A	1002	7/7	0.81	0.14	30,61,65,65	2
2	PEG	B	1002	7/7	0.87	0.10	30,46,47,48	2
2	PEG	B	1003	7/7	0.88	0.11	30,49,55,55	2
3	CA	A	1003	1/1	0.98	0.04	22,22,22,22	0
3	CA	B	1006	1/1	0.98	0.02	17,17,17,17	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.