



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2016 – 09:26 PM GMT

PDB ID : 5A6B  
Title : GH20C, Beta-hexosaminidase from Streptococcus pneumoniae in complex with PUGNAc  
Authors : Cid, M.; Robb, C.S.; Higgins, M.A.; Boraston, A.B.  
Deposited on : 2015-06-24  
Resolution : 1.77 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  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20026982

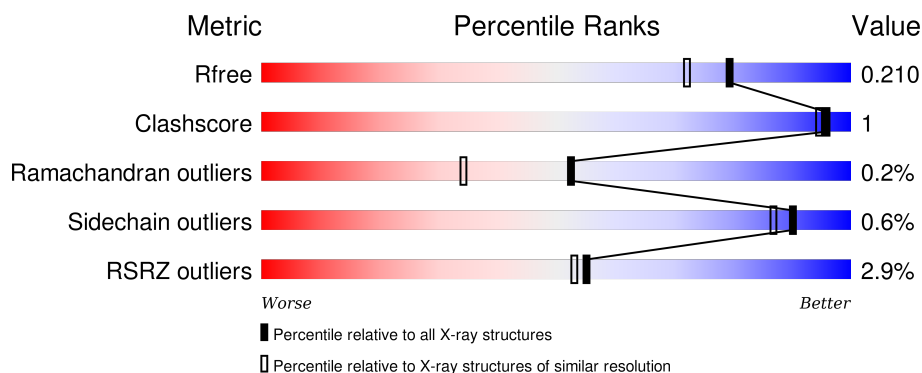
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.77 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))
$R_{free}$	91344	6655 (1.80-1.76)
Clashscore	102246	7658 (1.80-1.76)
Ramachandran outliers	100387	7570 (1.80-1.76)
Sidechain outliers	100360	7569 (1.80-1.76)
RSRZ outliers	91569	6671 (1.80-1.76)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	649	<div> <div>4%</div> <div>89% 5% 6%</div> </div>
1	B	649	<div> <div>2%</div> <div>90% 5% 5%</div> </div>
1	C	649	<div> <div>2%</div> <div>89% 5% 6%</div> </div>
2	D	648	<div> <div>2%</div> <div>91% • 5%</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 21421 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 N-ACETYL-BETA-D-GLUCOSAMINIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	612	Total	C	N	O	S	0	0	0
			4778	3048	800	911	19			
1	B	617	Total	C	N	O	S	0	4	0
			4872	3115	815	922	20			
1	C	610	Total	C	N	O	S	0	1	0
			4822	3075	812	916	19			

There are 75 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-22	MET	-	EXPRESSION TAG	UNP C1CH75
A	-21	GLY	-	EXPRESSION TAG	UNP C1CH75
A	-20	SER	-	EXPRESSION TAG	UNP C1CH75
A	-19	SER	-	EXPRESSION TAG	UNP C1CH75
A	-18	HIS	-	EXPRESSION TAG	UNP C1CH75
A	-17	HIS	-	EXPRESSION TAG	UNP C1CH75
A	-16	HIS	-	EXPRESSION TAG	UNP C1CH75
A	-15	HIS	-	EXPRESSION TAG	UNP C1CH75
A	-14	HIS	-	EXPRESSION TAG	UNP C1CH75
A	-13	HIS	-	EXPRESSION TAG	UNP C1CH75
A	-12	SER	-	EXPRESSION TAG	UNP C1CH75
A	-11	SER	-	EXPRESSION TAG	UNP C1CH75
A	-10	GLY	-	EXPRESSION TAG	UNP C1CH75
A	-9	LEU	-	EXPRESSION TAG	UNP C1CH75
A	-8	VAL	-	EXPRESSION TAG	UNP C1CH75
A	-7	PRO	-	EXPRESSION TAG	UNP C1CH75
A	-6	ARG	-	EXPRESSION TAG	UNP C1CH75
A	-5	GLY	-	EXPRESSION TAG	UNP C1CH75
A	-4	SER	-	EXPRESSION TAG	UNP C1CH75
A	-3	SER	-	EXPRESSION TAG	UNP C1CH75
A	-2	MET	-	EXPRESSION TAG	UNP C1CH75
A	-1	ALA	-	EXPRESSION TAG	UNP C1CH75
A	0	SER	-	EXPRESSION TAG	UNP C1CH75

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	LEU	PRO	CONFLICT	UNP C1CH75
A	158	VAL	MET	CONFLICT	UNP C1CH75
B	-22	MET	-	EXPRESSION TAG	UNP C1CH75
B	-21	GLY	-	EXPRESSION TAG	UNP C1CH75
B	-20	SER	-	EXPRESSION TAG	UNP C1CH75
B	-19	SER	-	EXPRESSION TAG	UNP C1CH75
B	-18	HIS	-	EXPRESSION TAG	UNP C1CH75
B	-17	HIS	-	EXPRESSION TAG	UNP C1CH75
B	-16	HIS	-	EXPRESSION TAG	UNP C1CH75
B	-15	HIS	-	EXPRESSION TAG	UNP C1CH75
B	-14	HIS	-	EXPRESSION TAG	UNP C1CH75
B	-13	HIS	-	EXPRESSION TAG	UNP C1CH75
B	-12	SER	-	EXPRESSION TAG	UNP C1CH75
B	-11	SER	-	EXPRESSION TAG	UNP C1CH75
B	-10	GLY	-	EXPRESSION TAG	UNP C1CH75
B	-9	LEU	-	EXPRESSION TAG	UNP C1CH75
B	-8	VAL	-	EXPRESSION TAG	UNP C1CH75
B	-7	PRO	-	EXPRESSION TAG	UNP C1CH75
B	-6	ARG	-	EXPRESSION TAG	UNP C1CH75
B	-5	GLY	-	EXPRESSION TAG	UNP C1CH75
B	-4	SER	-	EXPRESSION TAG	UNP C1CH75
B	-3	SER	-	EXPRESSION TAG	UNP C1CH75
B	-2	MET	-	EXPRESSION TAG	UNP C1CH75
B	-1	ALA	-	EXPRESSION TAG	UNP C1CH75
B	0	SER	-	EXPRESSION TAG	UNP C1CH75
B	9	LEU	PRO	CONFLICT	UNP C1CH75
B	158	VAL	MET	CONFLICT	UNP C1CH75
C	-22	MET	-	EXPRESSION TAG	UNP C1CH75
C	-21	GLY	-	EXPRESSION TAG	UNP C1CH75
C	-20	SER	-	EXPRESSION TAG	UNP C1CH75
C	-19	SER	-	EXPRESSION TAG	UNP C1CH75
C	-18	HIS	-	EXPRESSION TAG	UNP C1CH75
C	-17	HIS	-	EXPRESSION TAG	UNP C1CH75
C	-16	HIS	-	EXPRESSION TAG	UNP C1CH75
C	-15	HIS	-	EXPRESSION TAG	UNP C1CH75
C	-14	HIS	-	EXPRESSION TAG	UNP C1CH75
C	-13	HIS	-	EXPRESSION TAG	UNP C1CH75
C	-12	SER	-	EXPRESSION TAG	UNP C1CH75
C	-11	SER	-	EXPRESSION TAG	UNP C1CH75
C	-10	GLY	-	EXPRESSION TAG	UNP C1CH75
C	-9	LEU	-	EXPRESSION TAG	UNP C1CH75
C	-8	VAL	-	EXPRESSION TAG	UNP C1CH75

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-7	PRO	-	EXPRESSION TAG	UNP C1CH75
C	-6	ARG	-	EXPRESSION TAG	UNP C1CH75
C	-5	GLY	-	EXPRESSION TAG	UNP C1CH75
C	-4	SER	-	EXPRESSION TAG	UNP C1CH75
C	-3	SER	-	EXPRESSION TAG	UNP C1CH75
C	-2	MET	-	EXPRESSION TAG	UNP C1CH75
C	-1	ALA	-	EXPRESSION TAG	UNP C1CH75
C	0	SER	-	EXPRESSION TAG	UNP C1CH75
C	9	LEU	PRO	CONFLICT	UNP C1CH75
C	158	VAL	MET	CONFLICT	UNP C1CH75

- Molecule 2 is a protein called N-ACETYL-BETA-D-GLUCOSAMINIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	616	Total	C	N	O	S	0	3	0
			4888	3118	820	931	19			

There are 26 discrepancies between the modelled and reference sequences:

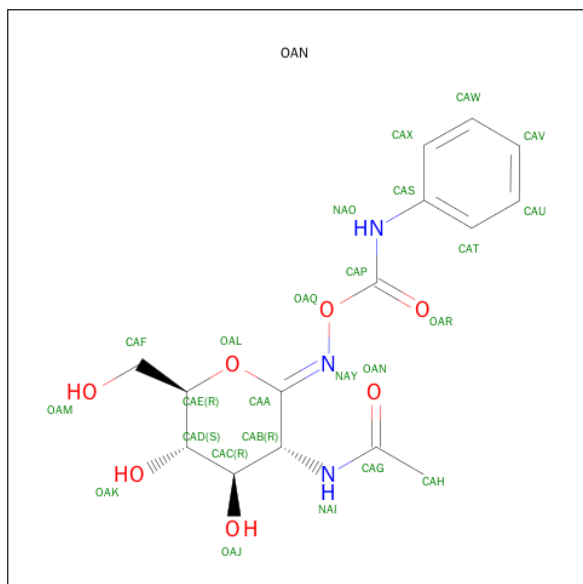
Chain	Residue	Modelled	Actual	Comment	Reference
D	-22	MET	-	EXPRESSION TAG	UNP C1CH75
D	-21	GLY	-	EXPRESSION TAG	UNP C1CH75
D	-20	SER	-	EXPRESSION TAG	UNP C1CH75
D	-19	SER	-	EXPRESSION TAG	UNP C1CH75
D	-18	HIS	-	EXPRESSION TAG	UNP C1CH75
D	-17	HIS	-	EXPRESSION TAG	UNP C1CH75
D	-16	HIS	-	EXPRESSION TAG	UNP C1CH75
D	-15	HIS	-	EXPRESSION TAG	UNP C1CH75
D	-14	HIS	-	EXPRESSION TAG	UNP C1CH75
D	-13	HIS	-	EXPRESSION TAG	UNP C1CH75
D	-12	SER	-	EXPRESSION TAG	UNP C1CH75
D	-11	SER	-	EXPRESSION TAG	UNP C1CH75
D	-10	GLY	-	EXPRESSION TAG	UNP C1CH75
D	-9	LEU	-	EXPRESSION TAG	UNP C1CH75
D	-8	VAL	-	EXPRESSION TAG	UNP C1CH75
D	-7	PRO	-	EXPRESSION TAG	UNP C1CH75
D	-6	ARG	-	EXPRESSION TAG	UNP C1CH75
D	-5	GLY	-	EXPRESSION TAG	UNP C1CH75
D	-4	SER	-	EXPRESSION TAG	UNP C1CH75
D	-3	SER	-	EXPRESSION TAG	UNP C1CH75
D	-2	MET	-	EXPRESSION TAG	UNP C1CH75
D	-1	ALA	-	EXPRESSION TAG	UNP C1CH75

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	SER	-	EXPRESSION TAG	UNP C1CH75
D	9	LEU	PRO	CONFLICT	UNP C1CH75
D	158	VAL	MET	CONFLICT	UNP C1CH75
D	.	-	LEU	DELETION	UNP C1CH75

- Molecule 3 is O-(2-ACETAMIDO-2-DEOXY D-GLUCOPYRANOSYLIDENE) AMINO-N-PHENYLCARBAMATE (three-letter code: OAN) (formula: C<sub>15</sub>H<sub>19</sub>N<sub>3</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			25	15	3	7		
3	B	1	Total	C	N	O	0	0
			25	15	3	7		
3	C	1	Total	C	N	O	0	0
			25	15	3	7		
3	D	1	Total	C	N	O	0	0
			25	15	3	7		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Mg	0	0
			1	1		
4	A	1	Total	Mg	0	0
			1	1		
4	C	1	Total	Mg	0	0
			1	1		

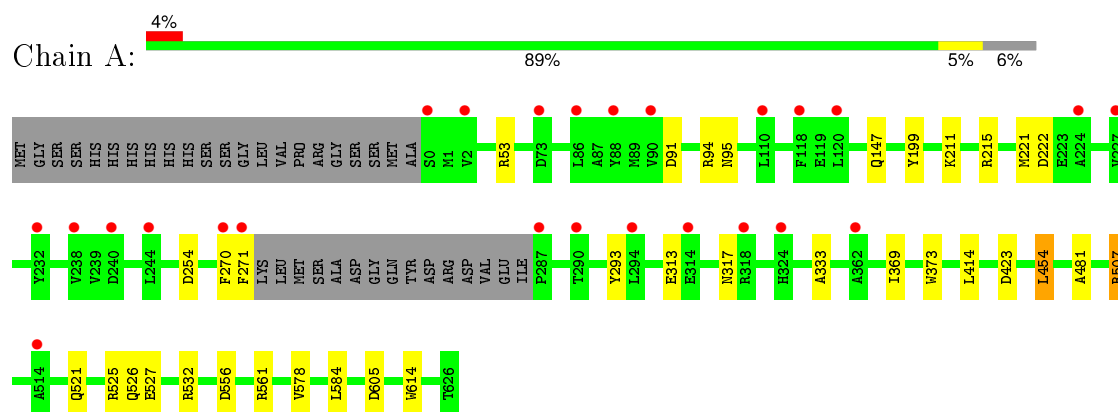
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	425	Total 425	O 425	0	0
5	B	492	Total 492	O 492	0	0
5	C	533	Total 533	O 533	0	0
5	D	508	Total 508	O 508	0	0

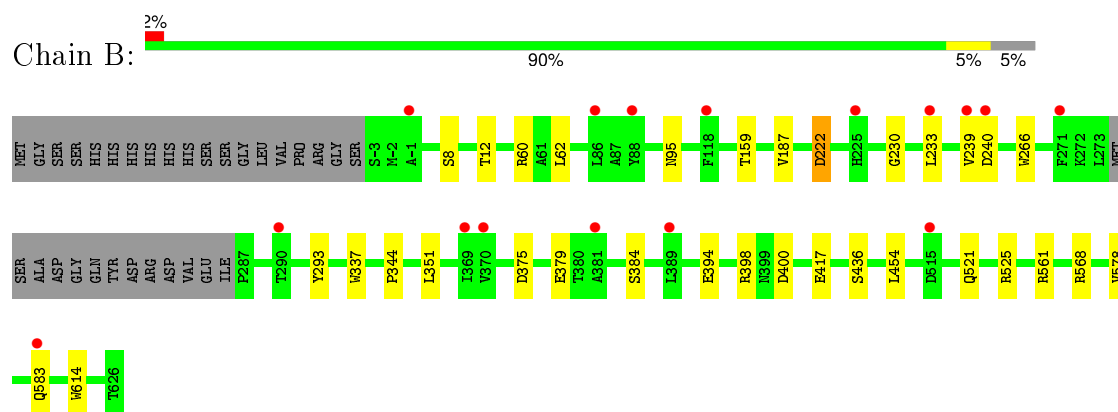
### 3 Residue-property plots

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.

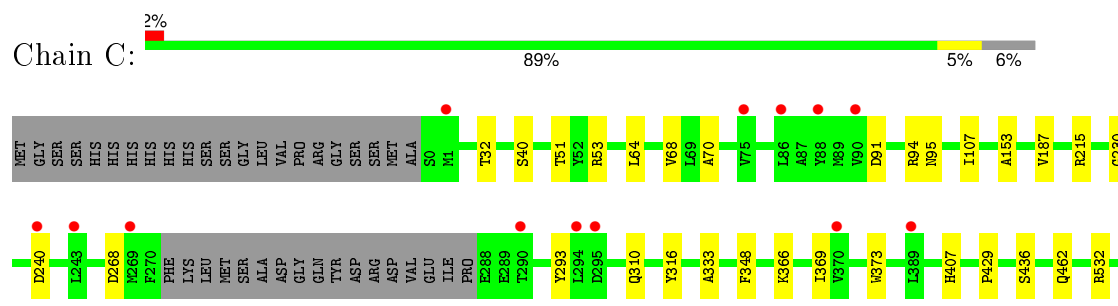
#### • Molecule 1: N-ACETYL-BETA-D-GLUCOSAMINIDASE



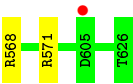
#### • Molecule 1: N-ACETYL-BETA-D-GLUCOSAMINIDASE



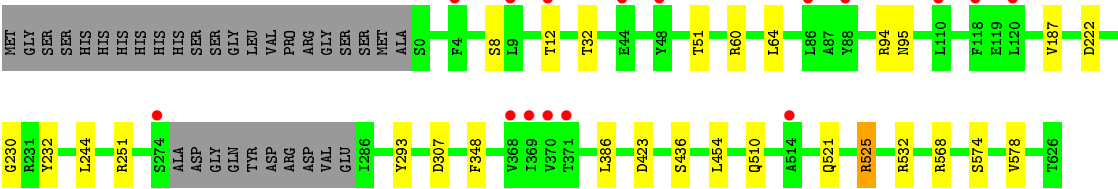
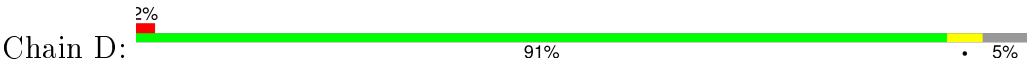
#### • Molecule 1: N-ACETYL-BETA-D-GLUCOSAMINIDASE







● Molecule 2: N-ACETYL-BETA-D-GLUCOSAMINIDASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	93.70Å 136.94Å 201.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.69 – 1.77 38.65 – 1.78	Depositor EDS
% Data completeness (in resolution range)	99.6 (38.69-1.77) 99.6 (38.65-1.78)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.36 (at 1.77Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.164 , 0.203 0.175 , 0.210	Depositor DCC
$R_{free}$ test set	12474 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	23.2	Xtriage
Anisotropy	0.036	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 48.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 248366 reflections (0.000%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	21421	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 16.45% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: OAN, MG

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	0.96	1/4879 (0.0%)	0.98	12/6626 (0.2%)
1	B	0.98	1/4994 (0.0%)	0.93	7/6780 (0.1%)
1	C	0.95	1/4926 (0.0%)	0.93	6/6683 (0.1%)
2	D	0.98	0/5003	0.95	8/6787 (0.1%)
All	All	0.97	3/19802 (0.0%)	0.95	33/26876 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	568	ARG	CZ-NH1	-5.76	1.25	1.33
1	B	417	GLU	CD-OE2	5.35	1.31	1.25
1	A	199	TYR	CB-CG	-5.12	1.44	1.51

All (33) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	507	ARG	NE-CZ-NH2	-16.65	111.98	120.30
1	A	507	ARG	NE-CZ-NH1	12.86	126.73	120.30
2	D	532	ARG	NE-CZ-NH1	11.14	125.87	120.30
1	A	53	ARG	NE-CZ-NH1	8.01	124.30	120.30
1	C	532	ARG	NE-CZ-NH1	7.97	124.29	120.30
1	A	532	ARG	NE-CZ-NH1	7.58	124.09	120.30
1	A	53	ARG	NE-CZ-NH2	-7.41	116.59	120.30
1	C	571	ARG	NE-CZ-NH2	-7.18	116.71	120.30
1	C	215	ARG	NE-CZ-NH1	6.97	123.79	120.30
2	D	532	ARG	NE-CZ-NH2	-6.81	116.90	120.30
1	A	423	ASP	CB-CG-OD1	6.64	124.27	118.30
1	A	215	ARG	NE-CZ-NH1	6.28	123.44	120.30
1	B	222	ASP	CB-CG-OD2	-6.25	112.67	118.30
1	B	568	ARG	NE-CZ-NH1	6.24	123.42	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	525	ARG	NE-CZ-NH1	6.08	123.34	120.30
2	D	386	LEU	CB-CG-CD2	5.95	121.12	111.00
1	C	53	ARG	NE-CZ-NH1	5.94	123.27	120.30
1	B	60	ARG	NE-CZ-NH2	-5.84	117.38	120.30
1	A	254	ASP	CB-CG-OD1	5.79	123.52	118.30
2	D	307	ASP	CB-CG-OD2	-5.79	113.09	118.30
2	D	423	ASP	CB-CG-OD2	-5.52	113.33	118.30
1	A	222	ASP	CB-CG-OD1	5.49	123.24	118.30
1	A	556	ASP	CB-CG-OD1	5.48	123.23	118.30
2	D	251	ARG	NE-CZ-NH1	5.45	123.03	120.30
1	B	394	GLU	OE1-CD-OE2	-5.44	116.77	123.30
1	A	454	LEU	CB-CG-CD2	5.35	120.09	111.00
1	A	556	ASP	CB-CG-OD2	-5.28	113.55	118.30
1	B	398	ARG	NE-CZ-NH2	-5.21	117.70	120.30
1	C	268	ASP	CB-CG-OD2	5.18	122.96	118.30
2	D	568	ARG	NE-CZ-NH1	5.16	122.88	120.30
1	B	400	ASP	CB-CG-OD1	5.14	122.92	118.30
1	B	379	GLU	OE1-CD-OE2	-5.02	117.28	123.30
1	C	240	ASP	CB-CG-OD1	5.02	122.82	118.30

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	4778	0	4548	13	0
1	B	4872	0	4677	16	0
1	C	4822	0	4659	11	0
2	D	4888	0	4705	11	0
3	A	25	0	19	2	0
3	B	25	0	19	2	0
3	C	25	0	19	0	0
3	D	25	0	19	1	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	1	0	0	0	0
5	A	425	0	0	1	0
5	B	492	0	0	4	1
5	C	533	0	0	0	0
5	D	508	0	0	1	1
All	All	21421	0	18665	56	1

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 1.

All (56) 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)
2:D:521:GLN:HG3	2:D:525:ARG:HD2	1.63	0.79
3:A:1627:OAN:HAT	3:A:1627:OAN:OAR	1.89	0.70
1:B:240:ASP:CB	5:B:2284:HOH:O	2.41	0.68
2:D:8:SER:O	2:D:12:THR:HG23	1.95	0.67
1:A:507:ARG:NH2	1:A:527:GLU:OE1	2.32	0.59
1:B:240:ASP:CB	5:B:2262:HOH:O	2.51	0.57
1:B:521:GLN:CG	1:B:525:ARG:HD2	2.35	0.57
3:B:1627:OAN:HAX	3:B:1627:OAN:OAR	2.06	0.56
1:B:222:ASP:HB3	1:B:266:TRP:CD1	2.42	0.55
1:A:270:PHE:O	1:A:271:PHE:CB	2.56	0.54
1:A:526:GLN:NE2	5:A:2345:HOH:O	2.41	0.54
2:D:521:GLN:HG3	2:D:525:ARG:HH11	1.73	0.53
2:D:521:GLN:OE1	2:D:525:ARG:NH1	2.42	0.53
2:D:232:TYR:CZ	2:D:244:LEU:HD11	2.43	0.53
1:B:233:LEU:C	1:B:233:LEU:HD23	2.30	0.52
1:A:578:VAL:HG12	1:A:584:LEU:HG	1.93	0.51
1:A:313:GLU:HG2	1:A:317:ASN:ND2	2.27	0.49
1:C:68:VAL:O	1:C:70:ALA:O	2.31	0.48
2:D:187:VAL:HB	2:D:230:GLY:HA3	1.96	0.48
3:D:1627:OAN:HAT	3:D:1627:OAN:OAR	2.14	0.48
1:B:8:SER:O	1:B:12:THR:HG23	2.15	0.47
1:C:64:LEU:CD2	1:C:407:HIS:CE1	2.97	0.47
1:C:429:PRO:HG2	1:C:462:GLN:HG3	1.95	0.46
1:B:344:PRO:HG3	1:B:384:SER:HB2	1.97	0.46
1:B:521:GLN:HG2	1:B:525:ARG:HD2	1.97	0.46
1:B:351:LEU:HD23	1:B:351:LEU:C	2.36	0.46
1:B:561:ARG:HG2	1:B:614:TRP:CE2	2.51	0.46
1:C:187:VAL:HB	1:C:230:GLY:HA3	1.97	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:337:TRP:CH2	1:B:436:SER:HB2	2.50	0.45
2:D:60:ARG:O	2:D:64:LEU:HG	2.16	0.45
1:B:187:VAL:HB	1:B:230:GLY:HA3	1.99	0.45
1:C:333:ALA:HA	1:C:369:ILE:O	2.17	0.45
1:C:91:ASP:HB2	1:C:373:TRP:HA	2.00	0.44
2:D:32:THR:O	2:D:51:THR:HA	2.17	0.44
1:A:147:GLN:NE2	1:A:211:LYS:HB3	2.33	0.43
1:C:32:THR:O	1:C:51:THR:HA	2.18	0.43
1:C:40:SER:HB3	1:C:51:THR:OG1	2.18	0.43
1:A:91:ASP:HB2	1:A:373:TRP:HA	2.01	0.43
1:A:605:ASP:CG	1:A:605:ASP:O	2.57	0.43
3:A:1627:OAN:OAR	3:A:1627:OAN:CAT	2.55	0.43
1:C:348:PHE:CZ	1:C:436:SER:HB3	2.53	0.43
2:D:348:PHE:CZ	2:D:436:SER:HB3	2.53	0.43
3:B:1627:OAN:HAU	5:B:2249:HOH:O	2.19	0.42
1:C:310:GLN:HB2	1:C:316:TYR:CZ	2.54	0.42
1:A:561:ARG:HG2	1:A:614:TRP:CE2	2.54	0.42
1:B:578:VAL:HG13	1:B:583:GLN:HB2	2.02	0.41
1:B:62:LEU:HD23	1:B:62:LEU:HA	1.85	0.41
1:B:239:VAL:HG12	1:B:240:ASP:O	2.20	0.41
1:A:521:GLN:HG3	1:A:525:ARG:HD2	2.03	0.41
1:A:147:GLN:HE21	1:A:211:LYS:HB3	1.86	0.41
1:A:414:LEU:HD21	1:A:481:ALA:HA	2.03	0.41
2:D:574:SER:O	2:D:578:VAL:HG23	2.21	0.41
1:B:159:THR:HG23	5:B:2148:HOH:O	2.21	0.41
2:D:510:GLN:HG3	5:D:2419:HOH:O	2.20	0.41
1:A:333:ALA:HA	1:A:369:ILE:O	2.21	0.40
1:C:107:ILE:HG21	1:C:153:ALA:HA	2.03	0.40

All (1) 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)
5:B:2137:HOH:O	5:D:2042:HOH:O[3_655]	2.13	0.07

## 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	608/649 (94%)	588 (97%)	19 (3%)	1 (0%)	52	34
1	B	617/649 (95%)	598 (97%)	17 (3%)	2 (0%)	46	28
1	C	607/649 (94%)	587 (97%)	19 (3%)	1 (0%)	52	34
2	D	615/648 (95%)	597 (97%)	16 (3%)	2 (0%)	46	28
All	All	2447/2595 (94%)	2370 (97%)	71 (3%)	6 (0%)	52	34

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	95	ASN
1	B	95	ASN
1	C	95	ASN
2	D	95	ASN
1	B	375	ASP
2	D	222	ASP

### 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	486/557 (87%)	482 (99%)	4 (1%)	86	81
1	B	505/557 (91%)	503 (100%)	2 (0%)	93	91
1	C	501/557 (90%)	498 (99%)	3 (1%)	90	87
2	D	508/556 (91%)	505 (99%)	3 (1%)	90	87

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	2000/2227 (90%)	1988 (99%)	12 (1%)	90	87

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	94	ARG
1	A	221	MET
1	A	293	TYR
1	A	454	LEU
1	B	293	TYR
1	B	454	LEU
1	C	94	ARG
1	C	293	TYR
1	C	366	LYS
2	D	94	ARG
2	D	293	TYR
2	D	454	LEU

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

Mol	Chain	Res	Type
1	A	526	GLN
1	B	154	GLN
1	C	324	HIS
1	C	522	GLN
2	D	567	GLN

### 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 7 ligands modelled in this entry, 3 are monoatomic - leaving 4 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	OAN	A	1627	-	22,26,26	1.70	3 (13%)	22,35,35	0.79	1 (4%)
3	OAN	B	1627	-	22,26,26	1.43	1 (4%)	22,35,35	1.08	1 (4%)
3	OAN	C	1627	-	22,26,26	1.54	3 (13%)	22,35,35	0.89	0
3	OAN	D	1627	-	22,26,26	1.44	1 (4%)	22,35,35	0.99	1 (4%)

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	OAN	A	1627	-	-	0/12/35/35	0/2/2/2
3	OAN	B	1627	-	-	0/12/35/35	0/2/2/2
3	OAN	C	1627	-	-	0/12/35/35	0/2/2/2
3	OAN	D	1627	-	-	0/12/35/35	0/2/2/2

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1627	OAN	OAQ-NAY	-6.36	1.28	1.44
3	B	1627	OAN	OAQ-NAY	-5.57	1.30	1.44
3	D	1627	OAN	OAQ-NAY	-5.26	1.31	1.44
3	C	1627	OAN	OAQ-NAY	-4.37	1.33	1.44
3	C	1627	OAN	CAS-NAO	-2.77	1.36	1.41
3	A	1627	OAN	CAS-NAO	-2.53	1.36	1.41
3	C	1627	OAN	OAL-CAE	-2.33	1.43	1.46
3	A	1627	OAN	CAP-NAO	-2.01	1.31	1.36

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
3	A	1627	OAN	CAS-NAO-CAP	-2.27	122.79	126.36
3	B	1627	OAN	OAK-CAD-CAC	-2.16	105.48	110.36
3	D	1627	OAN	CAF-CAE-CAD	2.17	118.43	112.99

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1627	OAN	2	0
3	B	1627	OAN	2	0
3	D	1627	OAN	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	612/649 (94%)	0.27	25 (4%) 41 39	13, 27, 45, 59	0
1	B	617/649 (95%)	0.04	16 (2%) 59 58	12, 22, 38, 57	1 (0%)
1	C	610/649 (93%)	0.07	14 (2%) 64 63	13, 22, 41, 58	2 (0%)
2	D	616/648 (95%)	0.06	16 (2%) 59 58	13, 22, 38, 56	0
All	All	2455/2595 (94%)	0.11	71 (2%) 55 53	12, 23, 42, 59	3 (0%)

All (71) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	240	ASP	4.1
1	B	583	GLN	4.0
1	C	294	LEU	3.8
1	A	290	THR	3.6
1	B	240	ASP	3.6
2	D	369	ILE	3.4
1	A	238	VAL	3.3
1	A	224	ALA	3.2
1	B	-1	ALA	3.1
1	B	239	VAL	3.1
1	C	90	VAL	3.1
1	B	370	VAL	3.1
1	C	269	MET	3.0
2	D	370	VAL	3.0
1	C	240	ASP	3.0
2	D	4	PHE	2.9
1	A	270	PHE	2.8
2	D	86	LEU	2.7
1	A	90	VAL	2.7
2	D	9	LEU	2.7
1	C	290	THR	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	225	HIS	2.6
1	A	88	TYR	2.6
1	A	294	LEU	2.6
2	D	48	TYR	2.6
1	A	362	ALA	2.6
2	D	368	VAL	2.6
2	D	120	LEU	2.6
1	A	118	PHE	2.5
2	D	118	PHE	2.5
1	A	514	ALA	2.5
2	D	274	SER	2.5
1	A	287	PRO	2.5
2	D	110	LEU	2.5
1	A	324	HIS	2.5
1	A	318	ARG	2.5
1	A	110	LEU	2.5
1	C	605	ASP	2.5
1	B	233	LEU	2.5
1	B	290	THR	2.4
1	A	244	LEU	2.4
1	A	73	ASP	2.4
1	B	381	ALA	2.4
1	B	86	LEU	2.4
1	C	88	TYR	2.4
2	D	88	TYR	2.4
1	A	227	VAL	2.4
2	D	44	GLU	2.4
1	C	1	MET	2.4
1	A	120	LEU	2.3
1	C	243	LEU	2.3
1	B	515	ASP	2.3
1	C	389	LEU	2.3
2	D	371	THR	2.3
1	C	295	ASP	2.3
1	C	75	VAL	2.3
2	D	12	THR	2.2
1	B	389	LEU	2.2
1	C	86	LEU	2.2
1	A	0	SER	2.2
1	A	86	LEU	2.2
1	A	314	GLU	2.2
1	B	88	TYR	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	232	TYR	2.1
1	B	118	PHE	2.1
1	B	271	PHE	2.1
1	A	2	VAL	2.1
1	C	370	VAL	2.1
1	B	369	ILE	2.1
2	D	514	ALA	2.0
1	A	271	PHE	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 carbohydrates 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. 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, 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	OAN	C	1627	25/25	0.96	0.12	0.57	13,18,37,38	8
3	OAN	A	1627	25/25	0.96	0.12	0.50	18,24,34,36	9
3	OAN	D	1627	25/25	0.94	0.11	-0.09	15,18,42,44	0
3	OAN	B	1627	25/25	0.93	0.11	-0.15	15,19,48,49	0
4	MG	C	1628	1/1	0.99	0.16	-	21,21,21,21	0
4	MG	A	1628	1/1	0.98	0.19	-	27,27,27,27	0
4	MG	B	1628	1/1	0.99	0.15	-	22,22,22,22	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.