



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 6, 2016 – 07:21 PM EST

PDB ID : 5SWB  
Title : Crystal structure of N-glycan transport solute binding protein (NgtS) from Streptococcus pneumoniae in complex with Man5GlcNAc  
Authors : Robb, M.; Boraston, A.B.  
Deposited on : 2016-08-08  
Resolution : 1.73 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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-20028442
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-20028442

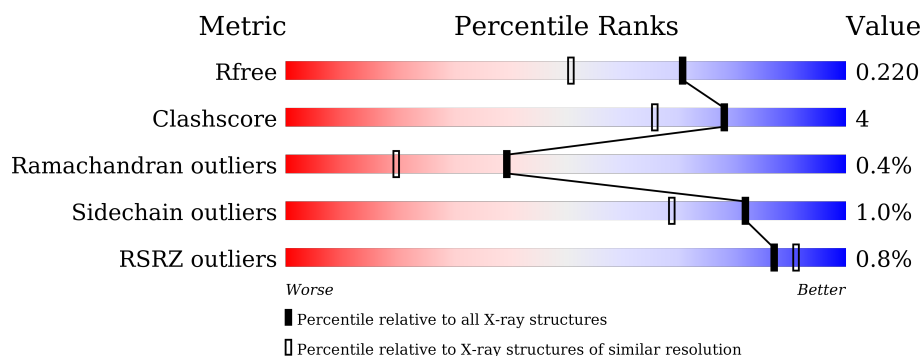
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.73 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	2417 (1.76-1.72)
Clashscore	102246	2570 (1.76-1.72)
Ramachandran outliers	100387	2544 (1.76-1.72)
Sidechain outliers	100360	2544 (1.76-1.72)
RSRZ outliers	91569	2420 (1.76-1.72)

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	476	<div> <div></div> <div>88% 8% 5%</div> </div>
1	C	476	<div> <div></div> <div>89% 5% 5%</div> </div>
1	D	476	<div> <div></div> <div>86% 9% 5%</div> </div>
1	G	476	<div> <div></div> <div>86% 8% 5%</div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 16745 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 Extracellular solute-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	454	Total	C	N	O	S	0	8	0
			3583	2288	584	704	7			
1	C	451	Total	C	N	O	S	0	5	0
			3552	2263	581	701	7			
1	D	450	Total	C	N	O	S	0	5	0
			3548	2266	578	697	7			
1	G	454	Total	C	N	O	S	0	8	0
			3600	2300	587	706	7			

There are 36 discrepancies between the modelled and reference sequences:

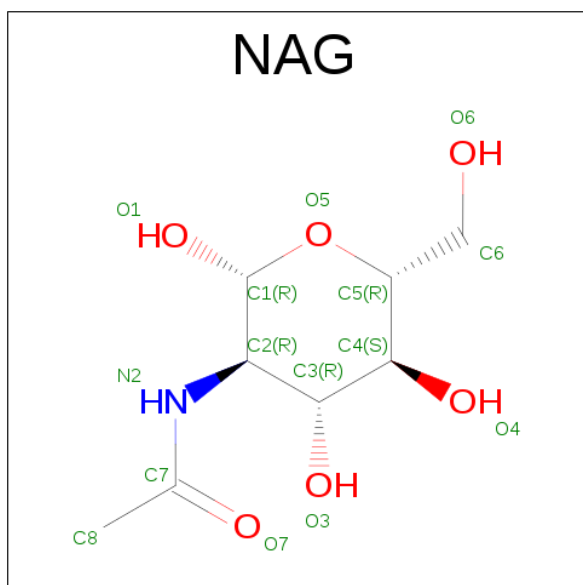
Chain	Residue	Modelled	Actual	Comment	Reference
A	24	MET	-	initiating methionine	UNP A0A0T7JX40
A	492	LEU	-	expression tag	UNP A0A0T7JX40
A	493	GLU	-	expression tag	UNP A0A0T7JX40
A	494	HIS	-	expression tag	UNP A0A0T7JX40
A	495	HIS	-	expression tag	UNP A0A0T7JX40
A	496	HIS	-	expression tag	UNP A0A0T7JX40
A	497	HIS	-	expression tag	UNP A0A0T7JX40
A	498	HIS	-	expression tag	UNP A0A0T7JX40
A	499	HIS	-	expression tag	UNP A0A0T7JX40
C	24	MET	-	initiating methionine	UNP A0A0T7JX40
C	492	LEU	-	expression tag	UNP A0A0T7JX40
C	493	GLU	-	expression tag	UNP A0A0T7JX40
C	494	HIS	-	expression tag	UNP A0A0T7JX40
C	495	HIS	-	expression tag	UNP A0A0T7JX40
C	496	HIS	-	expression tag	UNP A0A0T7JX40
C	497	HIS	-	expression tag	UNP A0A0T7JX40
C	498	HIS	-	expression tag	UNP A0A0T7JX40
C	499	HIS	-	expression tag	UNP A0A0T7JX40
D	24	MET	-	initiating methionine	UNP A0A0T7JX40
D	492	LEU	-	expression tag	UNP A0A0T7JX40
D	493	GLU	-	expression tag	UNP A0A0T7JX40

*Continued on next page...*

Continued from previous page...

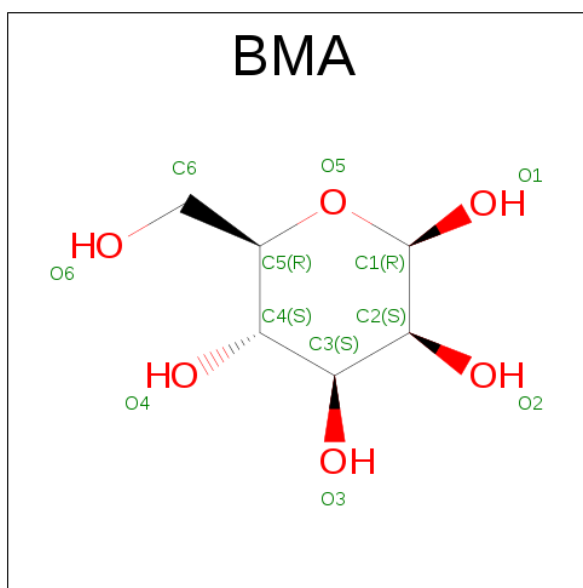
Chain	Residue	Modelled	Actual	Comment	Reference
D	494	HIS	-	expression tag	UNP A0A0T7JX40
D	495	HIS	-	expression tag	UNP A0A0T7JX40
D	496	HIS	-	expression tag	UNP A0A0T7JX40
D	497	HIS	-	expression tag	UNP A0A0T7JX40
D	498	HIS	-	expression tag	UNP A0A0T7JX40
D	499	HIS	-	expression tag	UNP A0A0T7JX40
G	24	MET	-	initiating methionine	UNP A0A0T7JX40
G	492	LEU	-	expression tag	UNP A0A0T7JX40
G	493	GLU	-	expression tag	UNP A0A0T7JX40
G	494	HIS	-	expression tag	UNP A0A0T7JX40
G	495	HIS	-	expression tag	UNP A0A0T7JX40
G	496	HIS	-	expression tag	UNP A0A0T7JX40
G	497	HIS	-	expression tag	UNP A0A0T7JX40
G	498	HIS	-	expression tag	UNP A0A0T7JX40
G	499	HIS	-	expression tag	UNP A0A0T7JX40

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



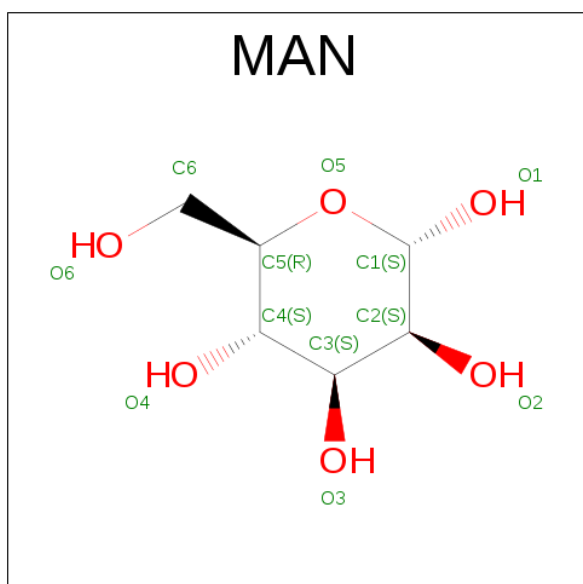
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			15	8	1	6		
2	C	1	Total	C	N	O	0	0
			15	8	1	6		
2	D	1	Total	C	N	O	0	0
			15	8	1	6		
2	G	1	Total	C	N	O	0	0
			15	8	1	6		

- Molecule 3 is BETA-D-MANNOSE (three-letter code: BMA) (formula:  $C_6H_{12}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			11	6	5		
3	C	1	Total	C	O	0	0
			11	6	5		
3	D	1	Total	C	O	0	0
			11	6	5		
3	G	1	Total	C	O	0	0
			11	6	5		

- Molecule 4 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula:  $C_6H_{12}O_6$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 11 6 5	0	0
4	A	1	Total C O 11 6 5	0	0
4	A	1	Total C O 11 6 5	0	0
4	A	1	Total C O 11 6 5	0	0
4	C	1	Total C O 11 6 5	0	0
4	C	1	Total C O 9 5 4	0	0
4	C	1	Total C O 11 6 5	0	0
4	C	1	Total C O 12 6 6	0	0
4	D	1	Total C O 11 6 5	0	0
4	D	1	Total C O 11 6 5	0	0
4	D	1	Total C O 11 6 5	0	0
4	D	1	Total C O 11 6 5	0	0
4	G	1	Total C O 11 6 5	0	0
4	G	1	Total C O 11 6 5	0	0
4	G	1	Total C O 11 6 5	0	0
4	G	1	Total C O 11 6 5	0	0

- Molecule 5 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	G	7	Total Cd 7 7	0	0
5	A	11	Total Cd 11 11	0	0
5	D	6	Total Cd 6 6	0	0
5	C	10	Total Cd 10 10	0	0

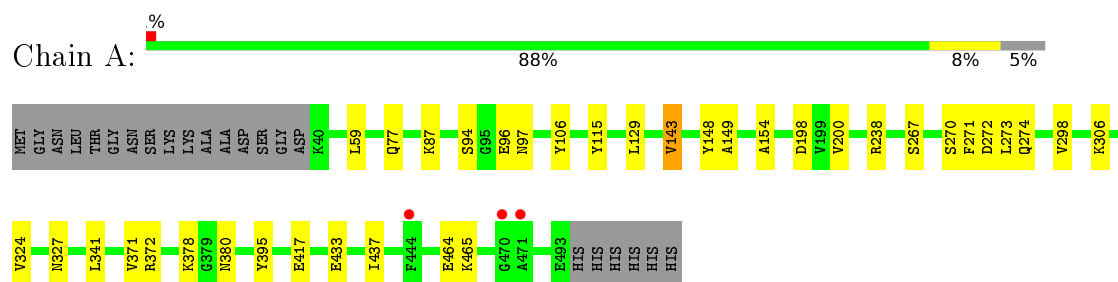
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	549	Total 549	O 549	0	0
6	C	500	Total 500	O 500	0	0
6	D	555	Total 555	O 555	0	0
6	G	545	Total 545	O 545	0	0

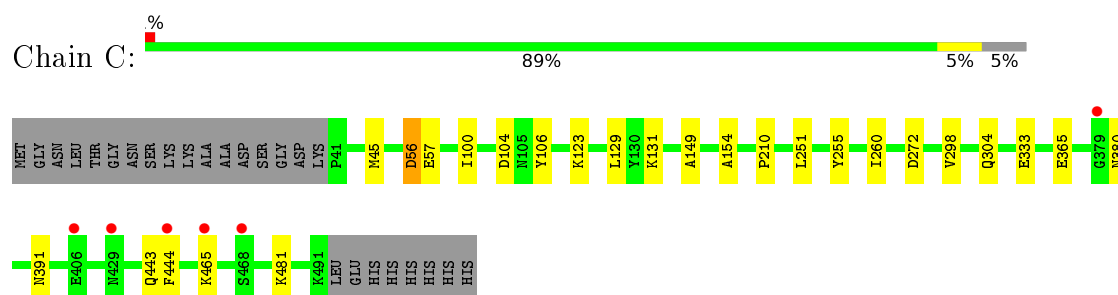
### 3 Residue-property plots [i](#)

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.

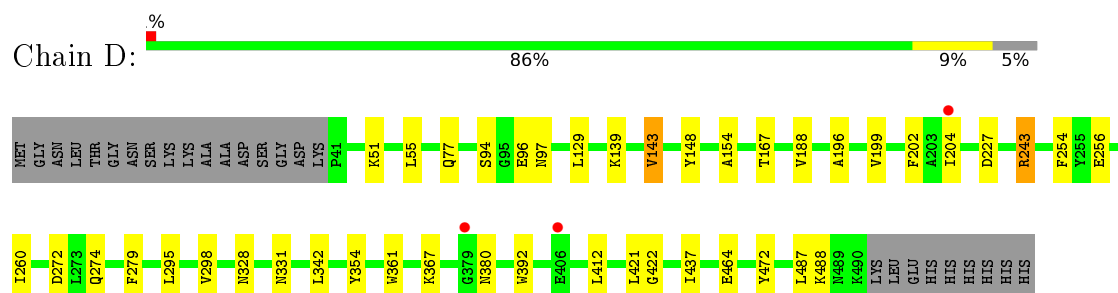
- Molecule 1: Extracellular solute-binding protein



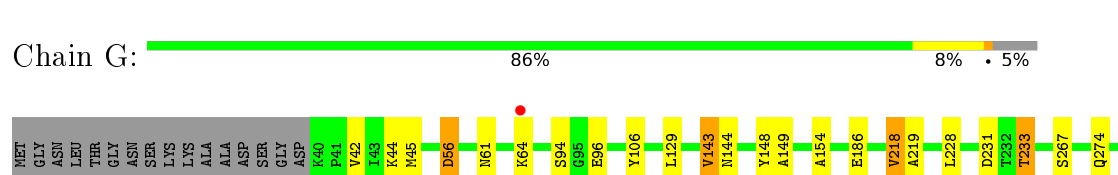
- Molecule 1: Extracellular solute-binding protein



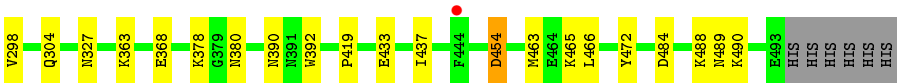
- Molecule 1: Extracellular solute-binding protein



- Molecule 1: Extracellular solute-binding protein







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	104.18Å 110.60Å 105.84Å 90.00° 109.52° 90.00°	Depositor
Resolution (Å)	99.76 – 1.73 48.18 – 1.73	Depositor EDS
% Data completeness (in resolution range)	99.9 (99.76-1.73) 99.9 (48.18-1.73)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.28 (at 1.73Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.180 , 0.220 0.180 , 0.220	Depositor DCC
$R_{free}$ test set	11777 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.0	Xtriage
Anisotropy	0.275	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 49.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.019 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	16745	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 36.14 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 5.2350e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: CD, BMA, NAG, MAN

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	0.94	1/3682 (0.0%)	0.95	5/4994 (0.1%)
1	C	0.92	2/3641 (0.1%)	0.91	5/4938 (0.1%)
1	D	0.92	1/3635 (0.0%)	0.97	4/4930 (0.1%)
1	G	0.95	0/3701	0.97	6/5019 (0.1%)
All	All	0.93	4/14659 (0.0%)	0.95	20/19881 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	365	GLU	CD-OE1	6.21	1.32	1.25
1	A	417	GLU	CD-OE2	6.00	1.32	1.25
1	C	333	GLU	CD-OE1	-5.19	1.20	1.25
1	D	392	TRP	CB-CG	5.07	1.59	1.50

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	454	ASP	CB-CG-OD2	19.06	135.46	118.30
1	D	243	ARG	NE-CZ-NH2	-16.20	112.20	120.30
1	D	243	ARG	NE-CZ-NH1	14.70	127.65	120.30
1	G	454	ASP	OD1-CG-OD2	-8.03	108.05	123.30
1	C	56[A]	ASP	CB-CG-OD1	7.48	125.03	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	3583	0	3526	21	0
1	C	3552	0	3482	20	0
1	D	3548	0	3501	32	0
1	G	3600	0	3543	38	1
2	A	15	0	14	0	0
2	C	15	0	14	0	0
2	D	15	0	14	0	0
2	G	15	0	14	0	0
3	A	11	0	8	0	0
3	C	11	0	8	0	0
3	D	11	0	8	0	0
3	G	11	0	8	0	0
4	A	44	0	38	0	0
4	C	43	0	36	0	0
4	D	44	0	38	0	0
4	G	44	0	38	0	0
5	A	11	0	0	0	1
5	C	10	0	0	1	0
5	D	6	0	0	1	1
5	G	7	0	0	0	1
6	A	549	0	0	6	1
6	C	500	0	0	6	1
6	D	555	0	0	16	2
6	G	545	0	0	14	1
All	All	16745	0	14290	105	5

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.

The worst 5 of 105 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:C:56[B]:ASP:OD1	1:G:56[B]:ASP:CG	1.91	1.09
1:D:367:LYS:HE2	6:D:686:HOH:O	1.57	1.04
1:C:56[B]:ASP:OD1	1:G:56[B]:ASP:OD2	1.77	1.02

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:463:MET:HE2	1:G:466:LEU:HD12	1.42	0.97
1:G:463:MET:CE	1:G:466:LEU:HD12	1.96	0.96

All (5) 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:G:454:ASP:OD2	5:G:511:CD:CD[2_555]	1.42	0.78
5:D:510:CD:CD	6:D:694:HOH:O[2_544]	1.53	0.67
5:A:511:CD:CD	6:C:612:HOH:O[2_454]	1.74	0.46
6:D:615:HOH:O	6:D:976:HOH:O[2_554]	2.00	0.20
6:A:823:HOH:O	6:G:1053:HOH:O[1_554]	2.09	0.11

## 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	460/476 (97%)	449 (98%)	8 (2%)	3 (1%)	26	9
1	C	454/476 (95%)	442 (97%)	11 (2%)	1 (0%)	52	32
1	D	453/476 (95%)	441 (97%)	11 (2%)	1 (0%)	52	32
1	G	460/476 (97%)	445 (97%)	13 (3%)	2 (0%)	39	20
All	All	1827/1904 (96%)	1777 (97%)	43 (2%)	7 (0%)	39	20

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	154	ALA
1	C	154	ALA
1	D	154	ALA
1	G	154	ALA
1	G	380	ASN

### 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	389/408 (95%)	385 (99%)	4 (1%)	82	69
1	C	385/408 (94%)	382 (99%)	3 (1%)	86	77
1	D	385/408 (94%)	383 (100%)	2 (0%)	92	87
1	G	391/408 (96%)	383 (98%)	8 (2%)	63	39
All	All	1550/1632 (95%)	1533 (99%)	17 (1%)	82	66

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	129	LEU
1	D	143	VAL
1	G	233	THR
1	C	481	LYS
1	G	327[A]	ASN

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

Mol	Chain	Res	Type
1	A	77	GLN
1	C	369	ASN
1	G	489	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 58 ligands modelled in this entry, 34 are monoatomic - leaving 24 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	NAG	A	501	3	15,15,15	1.02	1 (6%)	17,21,21	0.90	0
3	BMA	A	502	2,4	11,11,12	1.23	2 (18%)	15,15,17	1.28	2 (13%)
4	MAN	A	503	3	11,11,12	0.86	0	15,15,17	1.01	1 (6%)
4	MAN	A	504	3,4	11,11,12	0.83	0	15,15,17	1.64	2 (13%)
4	MAN	A	505	4	11,11,12	0.74	0	15,15,17	1.55	4 (26%)
4	MAN	A	506	4	11,11,12	1.17	2 (18%)	15,15,17	1.44	2 (13%)
2	NAG	C	501	3	15,15,15	0.93	0	17,21,21	1.20	1 (5%)
3	BMA	C	502	2,4	11,11,12	1.31	1 (9%)	15,15,17	1.08	0
4	MAN	C	503	3	11,11,12	1.65	3 (27%)	15,15,17	1.31	2 (13%)
4	MAN	C	504	3,4	9,9,12	2.08	2 (22%)	12,12,17	4.63	6 (50%)
4	MAN	C	505	4	11,11,12	1.45	2 (18%)	15,15,17	1.18	2 (13%)
4	MAN	C	506	4	12,12,12	1.18	1 (8%)	17,17,17	1.77	4 (23%)
2	NAG	D	502	3	15,15,15	0.64	0	17,21,21	1.55	3 (17%)
3	BMA	D	503	2,4	11,11,12	1.12	1 (9%)	15,15,17	1.62	4 (26%)
4	MAN	D	504	3	11,11,12	1.41	2 (18%)	15,15,17	1.32	2 (13%)
4	MAN	D	505	3,4	11,11,12	1.24	1 (9%)	15,15,17	1.55	2 (13%)
4	MAN	D	506	4	11,11,12	0.77	0	15,15,17	1.61	3 (20%)
4	MAN	D	507	4	11,11,12	1.32	2 (18%)	15,15,17	1.20	2 (13%)
2	NAG	G	501	3	15,15,15	1.07	0	17,21,21	1.15	1 (5%)
3	BMA	G	502	2,4	11,11,12	1.29	1 (9%)	15,15,17	1.27	2 (13%)
4	MAN	G	503	3	11,11,12	1.43	2 (18%)	15,15,17	1.36	3 (20%)
4	MAN	G	504	3,4	11,11,12	1.09	1 (9%)	15,15,17	2.09	3 (20%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	MAN	G	505	4	11,11,12	0.85	0	15,15,17	1.45	3 (20%)
4	MAN	G	506	4	11,11,12	1.01	0	15,15,17	1.86	4 (26%)

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	NAG	A	501	3	-	0/6/26/26	0/1/1/1
3	BMA	A	502	2,4	-	0/2/19/22	0/1/1/1
4	MAN	A	503	3	-	0/2/19/22	0/1/1/1
4	MAN	A	504	3,4	-	0/2/19/22	0/1/1/1
4	MAN	A	505	4	-	0/2/19/22	0/1/1/1
4	MAN	A	506	4	-	0/2/19/22	0/1/1/1
2	NAG	C	501	3	-	0/6/26/26	0/1/1/1
3	BMA	C	502	2,4	-	0/2/19/22	0/1/1/1
4	MAN	C	503	3	-	0/2/19/22	0/1/1/1
4	MAN	C	504	3,4	-	0/0/14/22	0/1/1/1
4	MAN	C	505	4	-	0/2/19/22	0/1/1/1
4	MAN	C	506	4	-	0/2/22/22	0/1/1/1
2	NAG	D	502	3	-	0/6/26/26	0/1/1/1
3	BMA	D	503	2,4	-	0/2/19/22	0/1/1/1
4	MAN	D	504	3	-	0/2/19/22	0/1/1/1
4	MAN	D	505	3,4	-	0/2/19/22	0/1/1/1
4	MAN	D	506	4	-	0/2/19/22	0/1/1/1
4	MAN	D	507	4	-	0/2/19/22	0/1/1/1
2	NAG	G	501	3	-	0/6/26/26	0/1/1/1
3	BMA	G	502	2,4	-	0/2/19/22	0/1/1/1
4	MAN	G	503	3	-	0/2/19/22	0/1/1/1
4	MAN	G	504	3,4	-	0/2/19/22	0/1/1/1
4	MAN	G	505	4	-	0/2/19/22	0/1/1/1
4	MAN	G	506	4	-	0/2/19/22	0/1/1/1

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	G	504	MAN	O5-C5	-2.95	1.37	1.43
4	D	507	MAN	O5-C1	-2.87	1.39	1.43
4	A	506	MAN	O2-C2	-2.63	1.37	1.43
4	C	505	MAN	O5-C1	-2.56	1.39	1.43
4	A	506	MAN	O5-C1	-2.37	1.39	1.43



The worst 5 of 58 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	504	MAN	O3-C3-C2	-4.80	101.22	110.01
4	C	504	MAN	C1-C2-C3	-4.12	104.56	109.55
4	A	504	MAN	O3-C3-C2	-3.84	102.97	110.01
4	C	504	MAN	O5-C5-C4	-3.79	102.50	110.49
4	G	506	MAN	C1-C2-C3	-3.76	105.00	109.55

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	454/476 (95%)	-0.34	3 (0%) 89 92	16, 26, 42, 55	0
1	C	451/476 (94%)	-0.27	6 (1%) 79 85	18, 29, 46, 61	0
1	D	450/476 (94%)	-0.33	3 (0%) 89 92	17, 27, 44, 61	0
1	G	454/476 (95%)	-0.30	2 (0%) 93 95	16, 27, 44, 63	1 (0%)
All	All	1809/1904 (95%)	-0.31	14 (0%) 87 91	16, 27, 44, 63	1 (0%)

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	468	SER	4.5
1	C	444	PHE	2.8
1	C	379	GLY	2.7
1	D	379	GLY	2.5
1	C	406	GLU	2.4

### 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
2	NAG	D	502	15/15	0.97	0.12	1.43	15,17,18,19	0
2	NAG	A	501	15/15	0.98	0.12	1.32	14,17,19,20	0
2	NAG	C	501	15/15	0.98	0.12	1.23	17,18,21,23	0
4	MAN	D	504	11/12	0.96	0.12	1.21	16,19,21,22	0
2	NAG	G	501	15/15	0.98	0.11	0.94	13,16,18,18	0
4	MAN	A	503	11/12	0.97	0.11	0.92	17,19,21,21	0
4	MAN	G	503	11/12	0.97	0.09	0.42	15,17,20,20	0
5	CD	A	507	1/1	0.99	0.06	-0.11	28,28,28,28	0
5	CD	C	511	1/1	0.95	0.08	-0.14	42,42,42,42	1
4	MAN	C	506	12/12	0.95	0.10	-0.16	21,23,29,31	0
5	CD	C	514	1/1	0.96	0.10	-0.32	46,46,46,46	1
4	MAN	C	503	11/12	0.96	0.07	-0.39	18,20,23,24	0
4	MAN	D	507	11/12	0.98	0.10	-0.43	19,22,25,26	0
4	MAN	D	506	11/12	0.96	0.10	-0.59	25,27,37,45	0
5	CD	D	511	1/1	0.98	0.06	-0.68	31,31,31,31	1
5	CD	A	516	1/1	0.98	0.06	-0.83	34,34,34,34	1
5	CD	D	508	1/1	0.98	0.05	-0.86	26,26,26,26	0
5	CD	G	508	1/1	0.97	0.07	-1.03	27,27,27,27	1
5	CD	G	513	1/1	0.96	0.07	-1.14	23,23,23,23	1
5	CD	A	515	1/1	0.97	0.05	-1.26	31,31,31,31	1
5	CD	G	510	1/1	0.98	0.06	-1.34	37,37,37,37	1
5	CD	A	510	1/1	0.99	0.05	-1.40	29,29,29,29	1
5	CD	C	516	1/1	0.99	0.05	-1.45	48,48,48,48	1
5	CD	C	508	1/1	0.99	0.05	-1.48	29,29,29,29	0
4	MAN	A	505	11/12	0.95	0.07	-1.62	22,25,32,38	0
4	MAN	G	505	11/12	0.97	0.07	-1.68	20,21,26,31	0
4	MAN	C	505	11/12	0.98	0.09	-1.78	20,22,27,29	0
5	CD	A	512	1/1	0.98	0.03	-1.90	40,40,40,40	1
4	MAN	A	506	11/12	0.98	0.09	-2.09	19,21,22,24	0
5	CD	C	507	1/1	0.99	0.05	-2.15	24,24,24,24	1
5	CD	G	511	1/1	0.99	0.04	-2.24	25,25,25,25	1
4	MAN	G	506	11/12	0.97	0.08	-2.57	19,19,22,24	0
5	CD	G	507	1/1	1.00	0.05	-2.72	22,22,22,22	1
5	CD	D	501	1/1	1.00	0.03	-3.64	24,24,24,24	1
5	CD	C	513	1/1	0.99	0.03	-5.51	27,27,27,27	1
5	CD	A	513	1/1	0.99	0.02	-6.69	32,32,32,32	1
4	MAN	C	504	9/12	0.95	0.11	-	19,20,23,27	0
5	CD	C	515	1/1	0.98	0.07	-	42,42,42,42	1
5	CD	A	511	1/1	0.98	0.02	-	43,43,43,43	1
5	CD	A	514	1/1	0.82	0.10	-	75,75,75,75	1
5	CD	D	509	1/1	0.99	0.02	-	31,31,31,31	1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	CD	C	509	1/1	0.95	0.05	-	40,40,40,40	1
5	CD	D	512	1/1	0.98	0.05	-	41,41,41,41	1
3	BMA	D	503	11/12	0.98	0.12	-	15,17,18,18	0
5	CD	C	512	1/1	0.99	0.02	-	35,35,35,35	1
5	CD	A	508	1/1	0.94	0.07	-	49,49,49,49	1
4	MAN	D	505	11/12	0.97	0.11	-	18,18,21,22	0
3	BMA	A	502	11/12	0.99	0.12	-	14,16,18,18	0
5	CD	G	509	1/1	0.95	0.08	-	115,115,115,115	0
5	CD	A	509	1/1	0.99	0.04	-	23,23,23,23	1
5	CD	D	510	1/1	0.99	0.02	-	28,28,28,28	1
5	CD	G	512	1/1	0.75	0.11	-	136,136,136,136	0
5	CD	C	510	1/1	0.99	0.06	-	25,25,25,25	1
3	BMA	G	502	11/12	0.98	0.12	-	13,15,17,18	0
5	CD	A	517	1/1	0.90	0.09	-	47,47,47,47	1
3	BMA	C	502	11/12	0.98	0.12	-	17,18,18,18	0
4	MAN	A	504	11/12	0.98	0.11	-	18,19,20,22	0
4	MAN	G	504	11/12	0.98	0.10	-	17,18,20,20	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.