



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 09:52 AM GMT

PDB ID : 3JYH  
Title : Human dipeptidyl peptidase DPP7  
Authors : Dobrovetsky, E.; Dong, A.; Seitova, A.; Crombett, L.; Paganon, S.; Cossar, D.; Edwards, A.M.; Arrowsmith, C.H.; Bountra, C.; Weigelt, J.; Hassel, A.; Shewchuk, L.; Bochkarev, A.; Structural Genomics Consortium (SGC)  
Deposited on : 2009-09-21  
Resolution : 2.19 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

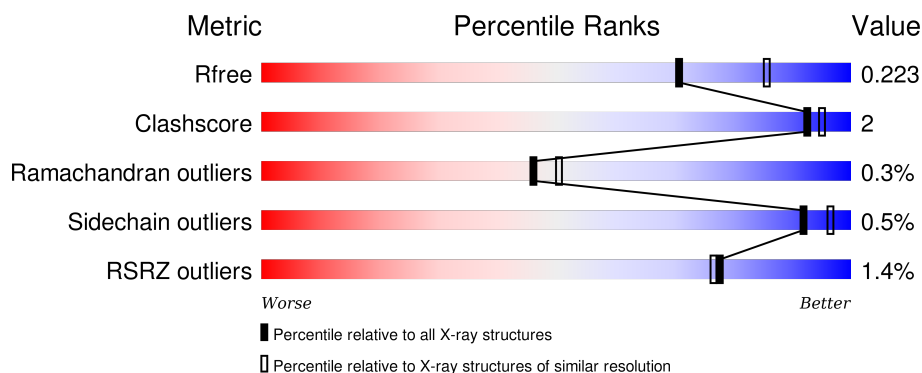
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.19 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	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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	469	<div> <div></div> <div>92%</div> <div>• •</div> </div>
1	B	469	<div> <div></div> <div>92%</div> <div>• •</div> </div>
1	C	469	<div> <div>%</div> <div>89%</div> <div>7% •</div> </div>
1	D	469	<div> <div>4%</div> <div>90%</div> <div>6% •</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 15144 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.

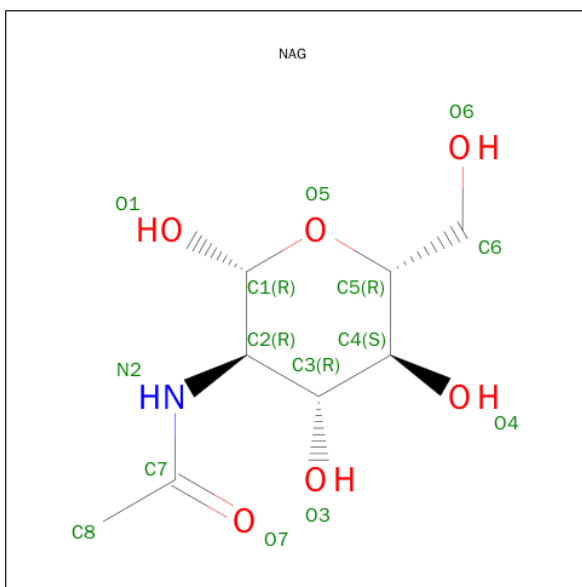
- Molecule 1 is a protein called Dipeptidyl-peptidase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	452	Total	C	N	O	S	0	1	0
			3472	2228	590	639	15			
1	B	452	Total	C	N	O	S	0	0	0
			3471	2225	595	637	14			
1	C	451	Total	C	N	O	S	0	0	0
			3467	2220	599	634	14			
1	D	450	Total	C	N	O	S	0	0	0
			3447	2206	596	631	14			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	24	GLY	-	EXPRESSION TAG	UNP Q9UHL4
A	25	ALA	-	EXPRESSION TAG	UNP Q9UHL4
A	26	MET	-	EXPRESSION TAG	UNP Q9UHL4
A	27	ASP	-	EXPRESSION TAG	UNP Q9UHL4
B	24	GLY	-	EXPRESSION TAG	UNP Q9UHL4
B	25	ALA	-	EXPRESSION TAG	UNP Q9UHL4
B	26	MET	-	EXPRESSION TAG	UNP Q9UHL4
B	27	ASP	-	EXPRESSION TAG	UNP Q9UHL4
C	24	GLY	-	EXPRESSION TAG	UNP Q9UHL4
C	25	ALA	-	EXPRESSION TAG	UNP Q9UHL4
C	26	MET	-	EXPRESSION TAG	UNP Q9UHL4
C	27	ASP	-	EXPRESSION TAG	UNP Q9UHL4
D	24	GLY	-	EXPRESSION TAG	UNP Q9UHL4
D	25	ALA	-	EXPRESSION TAG	UNP Q9UHL4
D	26	MET	-	EXPRESSION TAG	UNP Q9UHL4
D	27	ASP	-	EXPRESSION TAG	UNP Q9UHL4

- Molecule 2 is SUGAR (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
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	C	1	Total	C	N	O	0	0
			14	8	1	5		
2	C	1	Total	C	N	O	0	0
			14	8	1	5		
2	D	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	3	Total	C	N	O	0	0
			39	22	2	15		
3	B	3	Total	C	N	O	0	0
			39	22	2	15		
3	C	3	Total	C	N	O	0	0
			39	22	2	15		

- Molecule 4 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	2	Total	C	N	O	0	0
			28	16	2	10		
4	B	2	Total	C	N	O	0	0
			28	16	2	10		
4	C	2	Total	C	N	O	0	0
			28	16	2	10		
4	D	2	Total	C	N	O	0	0
			28	16	2	10		
4	D	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	292	Total	O	0	0
			292	292		
5	B	286	Total	O	0	0
			286	286		
5	C	195	Total	O	0	0
			195	195		
5	D	173	Total	O	0	0
			173	173		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	80.22Å 130.22Å 124.44Å 90.00° 102.36° 90.00°	Depositor
Resolution (Å)	100.00 – 2.19 48.71 – 2.19	Depositor EDS
% Data completeness (in resolution range)	98.8 (100.00-2.19) 98.9 (48.71-2.19)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.01 (at 2.18Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.180 , 0.222 0.180 , 0.223	Depositor DCC
$R_{free}$ test set	6397 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	31.2	Xtriage
Anisotropy	0.026	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 54.5	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 127511 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	15144	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.85% 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: BMA, NAG

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	0.69	0/3573	0.68	2/4865 (0.0%)
1	B	0.71	0/3569	0.69	3/4861 (0.1%)
1	C	0.61	0/3565	0.64	1/4856 (0.0%)
1	D	0.59	0/3545	0.62	1/4832 (0.0%)
All	All	0.65	0/14252	0.66	7/19414 (0.0%)

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	343	ASP	CB-CG-OD1	5.96	123.66	118.30
1	B	137	ASP	CB-CG-OD1	5.89	123.60	118.30
1	B	142	LEU	CA-CB-CG	5.65	128.31	115.30
1	A	247	GLN	CB-CA-C	-5.49	99.42	110.40
1	C	343	ASP	CB-CG-OD1	5.48	123.23	118.30
1	B	203	ARG	NE-CZ-NH2	-5.35	117.62	120.30
1	D	203	ARG	NE-CZ-NH1	-5.25	117.67	120.30

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	3472	0	3243	11	0
1	B	3471	0	3246	9	0
1	C	3467	0	3247	21	0
1	D	3447	0	3214	19	0
2	A	28	0	26	0	0
2	B	14	0	13	0	0
2	C	28	0	26	1	0
2	D	14	0	13	0	0
3	A	39	0	34	0	0
3	B	39	0	34	0	0
3	C	39	0	34	1	0
4	A	28	0	25	0	0
4	B	28	0	25	0	0
4	C	28	0	25	0	0
4	D	56	0	50	0	0
5	A	292	0	0	1	0
5	B	286	0	0	1	0
5	C	195	0	0	2	0
5	D	173	0	0	2	0
All	All	15144	0	13255	54	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 2.

All (54) 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:D:247:GLN:NE2	1:D:324:ASP:OD1	2.18	0.77
1:C:127:GLU:HG2	1:D:44:ASN:HD21	1.50	0.74
1:C:127:GLU:HG2	1:D:44:ASN:ND2	2.09	0.67
1:C:134:ALA:O	1:C:137:ASP:HB2	2.00	0.61
1:A:216:CYS:O	1:A:220:VAL:HG23	2.07	0.55
1:C:400:TRP:HA	1:D:177:HIS:CD2	2.44	0.53
1:C:236:TYR:HB3	1:C:249:LEU:O	2.07	0.53
1:A:260[A]:MET:HE1	5:A:646:HOH:O	2.09	0.51
1:A:177:HIS:CD2	1:B:400:TRP:HA	2.45	0.51
1:D:65:VAL:O	1:D:67:GLY:N	2.44	0.50
1:B:198:SER:HB2	1:B:358:THR:O	2.11	0.50
1:D:29:PRO:HB3	1:D:63:PHE:CE2	2.48	0.48
1:C:177:HIS:CD2	1:D:400:TRP:HA	2.48	0.48
1:D:203:ARG:NH1	5:D:1225:HOH:O	2.23	0.48
1:B:69:GLY:HA3	1:B:99:GLY:O	2.13	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:273:ASP:HA	1:B:288:PRO:HD2	1.95	0.48
1:C:247:GLN:NE2	1:C:324:ASP:OD1	2.48	0.47
1:C:127:GLU:CG	1:D:44:ASN:HD21	2.25	0.47
1:C:92:ALA:HA	1:C:102:LEU:HD11	1.97	0.47
1:C:48:PHE:CE2	1:C:128:LEU:HD21	2.50	0.47
1:D:415:GLY:HA3	1:D:441:GLY:O	2.15	0.47
3:C:1:BMA:O2	3:C:493:NAG:H4	2.13	0.47
1:A:97:GLU:OE2	1:A:98:ARG:NH2	2.48	0.46
1:C:200:GLN:HB3	5:C:509:HOH:O	2.16	0.46
1:D:330:HIS:CE1	1:D:340:THR:HG22	2.50	0.46
1:C:275:PRO:HG3	1:C:455:PRO:HD3	1.96	0.46
1:A:40:LEU:HD12	1:A:51:LYS:HB2	1.96	0.46
1:A:400:TRP:HA	1:B:177:HIS:CD2	2.51	0.46
1:C:332:CYS:HB2	1:C:348:ASP:OD1	2.16	0.45
1:A:247:GLN:NE2	1:A:324:ASP:OD1	2.49	0.45
1:B:247:GLN:NE2	1:B:324:ASP:OD1	2.50	0.44
1:B:109:TYR:CE2	1:B:115:PRO:HG2	2.53	0.44
1:D:29:PRO:HB3	1:D:63:PHE:HE2	1.82	0.44
1:D:203:ARG:NE	5:D:1225:HOH:O	2.37	0.43
1:A:185:ALA:HA	1:A:413:SER:OG	2.19	0.42
1:D:108:ARG:O	1:D:109:TYR:HB2	2.20	0.42
1:B:203:ARG:HD3	5:B:613:HOH:O	2.19	0.42
1:A:356:ASN:OD1	1:A:392:PRO:HA	2.19	0.42
1:C:63:PHE:O	1:C:99:GLY:HA2	2.20	0.42
1:C:425:ILE:HD12	1:C:429:LEU:HD11	2.02	0.42
1:D:188:PRO:O	1:D:192:VAL:HG22	2.20	0.42
1:C:176:PRO:HD2	1:D:400:TRP:CE2	2.55	0.41
1:A:149:LEU:HD23	1:A:149:LEU:HA	1.82	0.41
1:D:134:ALA:O	1:D:137:ASP:HB2	2.21	0.41
1:A:95:ALA:HA	1:A:100:ALA:HB3	2.02	0.41
1:C:213:SER:HB3	1:C:216:CYS:HB3	2.03	0.41
1:D:269:LEU:HD22	1:D:288:PRO:HB2	2.02	0.41
2:C:495:NAG:H83	5:C:1252:HOH:O	2.20	0.41
1:C:269:LEU:HD22	1:C:288:PRO:HB2	2.03	0.41
1:C:273:ASP:HA	1:C:288:PRO:HD2	2.02	0.41
1:C:213:SER:HB3	1:C:216:CYS:CB	2.51	0.40
1:D:166:MET:CE	1:D:192:VAL:HG21	2.51	0.40
1:B:231:PHE:CD2	1:B:255:LEU:HD13	2.57	0.40
1:C:411:ILE:HG13	1:C:472:TRP:CH2	2.57	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	451/469 (96%)	433 (96%)	17 (4%)	1 (0%)	52	59
1	B	450/469 (96%)	434 (96%)	15 (3%)	1 (0%)	52	59
1	C	449/469 (96%)	426 (95%)	23 (5%)	0	100	100
1	D	448/469 (96%)	424 (95%)	20 (4%)	4 (1%)	21	19
All	All	1798/1876 (96%)	1717 (96%)	75 (4%)	6 (0%)	46	50

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	49	GLY
1	D	66	ARG
1	A	66	ARG
1	D	442	ALA
1	D	444	HIS
1	B	162	SER

### 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	338/371 (91%)	337 (100%)	1 (0%)	94	98
1	B	338/371 (91%)	336 (99%)	2 (1%)	90	95
1	C	339/371 (91%)	338 (100%)	1 (0%)	94	98
1	D	336/371 (91%)	333 (99%)	3 (1%)	84	92

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1351/1484 (91%)	1344 (100%)	7 (0%)	92 96

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	198	SER
1	B	142	LEU
1	B	427	ARG
1	C	152	GLN
1	D	162	SER
1	D	213	SER
1	D	231	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	37	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 ⓘ

19 carbohydrates 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
3	BMA	A	493	3	11,11,12	0.73	0	14,15,17	1.04	1 (7%)
3	NAG	A	494	3	14,14,15	0.55	0	15,19,21	0.98	0
4	NAG	A	496	1,4	14,14,15	0.78	0	15,19,21	1.50	3 (20%)
4	NAG	A	497	4	14,14,15	0.53	0	15,19,21	1.37	2 (13%)
3	NAG	A	498	1,3	14,14,15	0.87	1 (7%)	15,19,21	1.43	1 (6%)
3	BMA	B	493	3	11,11,12	0.84	0	14,15,17	1.04	1 (7%)
3	NAG	B	494	3	14,14,15	0.67	0	15,19,21	1.47	3 (20%)
4	NAG	B	495	1,4	14,14,15	0.39	0	15,19,21	0.80	0
4	NAG	B	496	4	14,14,15	0.53	0	15,19,21	1.80	3 (20%)
3	NAG	B	497	1,3	14,14,15	0.72	0	15,19,21	1.15	0
3	BMA	C	1	3	11,11,12	0.76	0	14,15,17	1.83	2 (14%)
3	NAG	C	493	3	14,14,15	0.66	0	15,19,21	2.15	2 (13%)
4	NAG	C	496	1,4	14,14,15	0.62	0	15,19,21	1.14	1 (6%)
4	NAG	C	497	4	14,14,15	0.59	0	15,19,21	1.56	1 (6%)
3	NAG	C	498	1,3	14,14,15	0.57	0	15,19,21	1.29	1 (6%)
4	NAG	D	493	1,4	14,14,15	0.53	0	15,19,21	1.92	2 (13%)
4	NAG	D	494	4	14,14,15	0.65	0	15,19,21	1.20	2 (13%)
4	NAG	D	495	1,4	14,14,15	0.48	0	15,19,21	1.13	1 (6%)
4	NAG	D	496	4	14,14,15	0.62	0	15,19,21	1.21	2 (13%)

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	BMA	A	493	3	-	0/2/19/22	0/1/1/1
3	NAG	A	494	3	-	0/6/23/26	0/1/1/1
4	NAG	A	496	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	497	4	-	0/6/23/26	0/1/1/1
3	NAG	A	498	1,3	-	0/6/23/26	0/1/1/1
3	BMA	B	493	3	-	0/2/19/22	0/1/1/1
3	NAG	B	494	3	-	0/6/23/26	0/1/1/1
4	NAG	B	495	1,4	-	0/6/23/26	0/1/1/1
4	NAG	B	496	4	-	0/6/23/26	0/1/1/1
3	NAG	B	497	1,3	-	0/6/23/26	0/1/1/1
3	BMA	C	1	3	-	0/2/19/22	0/1/1/1
3	NAG	C	493	3	-	0/6/23/26	0/1/1/1
4	NAG	C	496	1,4	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	C	497	4	-	0/6/23/26	0/1/1/1
3	NAG	C	498	1,3	-	0/6/23/26	0/1/1/1
4	NAG	D	493	1,4	-	0/6/23/26	0/1/1/1
4	NAG	D	494	4	-	0/6/23/26	0/1/1/1
4	NAG	D	495	1,4	-	0/6/23/26	0/1/1/1
4	NAG	D	496	4	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	498	NAG	C1-C2	2.15	1.55	1.52

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	1	BMA	C3-C4-C5	-3.53	104.05	110.20
4	D	494	NAG	C2-N2-C7	-2.77	119.48	123.04
4	B	496	NAG	C4-C3-C2	-2.75	106.96	111.23
4	A	496	NAG	C3-C4-C5	-2.23	106.32	110.20
3	B	494	NAG	C2-N2-C7	-2.19	120.22	123.04
3	B	494	NAG	C4-C3-C2	2.08	114.47	111.23
4	A	496	NAG	O4-C4-C3	2.12	115.11	110.34
3	B	493	BMA	O3-C3-C2	2.19	113.96	110.00
4	D	496	NAG	C4-C3-C2	2.23	114.70	111.23
4	D	496	NAG	C1-O5-C5	2.34	115.22	112.25
4	D	495	NAG	O4-C4-C3	2.36	115.66	110.34
4	C	496	NAG	C1-O5-C5	2.47	115.39	112.25
4	D	494	NAG	C4-C3-C2	2.51	115.13	111.23
4	D	493	NAG	C2-N2-C7	2.69	126.50	123.04
3	A	493	BMA	C1-C2-C3	2.71	112.75	109.54
4	A	497	NAG	O5-C5-C6	2.73	113.26	107.35
4	B	496	NAG	O5-C5-C6	2.99	113.82	107.35
4	A	497	NAG	C1-O5-C5	3.25	116.38	112.25
3	C	493	NAG	C4-C3-C2	3.49	116.66	111.23
3	C	498	NAG	C1-O5-C5	3.63	116.86	112.25
3	B	494	NAG	C1-O5-C5	3.91	117.21	112.25
4	A	496	NAG	C1-O5-C5	4.13	117.49	112.25
3	C	1	BMA	O5-C5-C6	4.21	116.47	107.35
4	B	496	NAG	C1-O5-C5	4.44	117.88	112.25
3	A	498	NAG	C1-O5-C5	4.59	118.07	112.25
4	C	497	NAG	C1-O5-C5	4.93	118.51	112.25
3	C	493	NAG	C1-O5-C5	5.90	119.73	112.25

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	493	NAG	C1-O5-C5	6.32	120.27	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	1	BMA	1	0
3	C	493	NAG	1	0

## 5.6 Ligand geometry

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	$\# Z  > 2$
2	NAG	A	1	1	14,14,15	0.61	0	15,19,21	0.96	0
2	NAG	A	495	1	14,14,15	0.42	0	15,19,21	1.57	1 (6%)
2	NAG	B	1	1	14,14,15	0.61	0	15,19,21	1.04	2 (13%)
2	NAG	C	494	1	14,14,15	0.57	0	15,19,21	1.21	1 (6%)
2	NAG	C	495	1	14,14,15	0.50	0	15,19,21	1.04	1 (6%)
2	NAG	D	1	1	14,14,15	0.61	0	15,19,21	1.10	2 (13%)

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	1	1	-	0/6/23/26	0/1/1/1
2	NAG	A	495	1	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	B	1	1	-	0/6/23/26	0/1/1/1
2	NAG	C	494	1	-	0/6/23/26	0/1/1/1
2	NAG	C	495	1	-	0/6/23/26	0/1/1/1
2	NAG	D	1	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	1	NAG	C3-C4-C5	-2.18	106.39	110.20
2	C	495	NAG	O7-C7-C8	-2.06	118.28	122.06
2	B	1	NAG	C1-O5-C5	2.15	114.97	112.25
2	B	1	NAG	O4-C4-C5	2.18	115.00	109.24
2	D	1	NAG	O5-C5-C6	2.57	112.91	107.35
2	C	494	NAG	C1-O5-C5	3.37	116.52	112.25
2	A	495	NAG	C1-O5-C5	4.97	118.56	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	495	NAG	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	452/469 (96%)	-0.13	1 (0%) 95 95	16, 26, 39, 56	0
1	B	452/469 (96%)	-0.10	1 (0%) 95 95	16, 26, 39, 58	0
1	C	451/469 (96%)	0.14	5 (1%) 82 82	23, 37, 50, 61	0
1	D	450/469 (95%)	0.34	18 (4%) 42 41	23, 39, 52, 63	0
All	All	1805/1876 (96%)	0.06	25 (1%) 78 77	16, 32, 48, 63	0

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	230	LEU	4.5
1	D	236	TYR	4.4
1	D	233	GLN	3.7
1	D	231	PHE	3.6
1	D	232	LEU	3.3
1	D	29	PRO	3.3
1	C	431	ALA	3.2
1	D	389	TRP	3.0
1	C	213	SER	3.0
1	D	234	GLY	3.0
1	D	476	ALA	3.0
1	D	249	LEU	2.9
1	D	68	GLU	2.6
1	B	232	LEU	2.6
1	D	475	ALA	2.4
1	C	389	TRP	2.4
1	D	298	SER	2.4
1	D	65	VAL	2.4
1	D	297	LEU	2.3
1	C	304	THR	2.3
1	A	68	GLU	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	301	GLN	2.1
1	D	301	GLN	2.1
1	D	312	LEU	2.1
1	D	303	ILE	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](#)

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(Å <sup>2</sup> )	Q<0.9
4	NAG	D	495	14/15	0.90	0.18	1.00	43,48,51,53	0
3	NAG	C	498	14/15	0.93	0.13	0.22	31,39,41,44	0
4	NAG	B	495	14/15	0.98	0.12	-0.07	27,31,37,40	0
4	NAG	C	496	14/15	0.94	0.12	-0.31	37,41,44,46	0
4	NAG	A	496	14/15	0.96	0.10	-0.33	24,29,36,37	0
3	NAG	A	498	14/15	0.95	0.10	-0.40	21,30,34,39	0
3	NAG	B	497	14/15	0.96	0.11	-0.57	22,30,40,41	0
4	NAG	D	493	14/15	0.96	0.12	-0.76	36,40,45,46	0
4	NAG	C	497	14/15	0.91	0.25	-	50,53,55,55	0
3	NAG	A	494	14/15	0.96	0.16	-	38,41,45,50	0
3	NAG	B	494	14/15	0.88	0.20	-	43,46,50,55	0
3	BMA	C	1	11/12	0.64	0.22	-	57,59,62,62	0
4	NAG	D	494	14/15	0.89	0.23	-	50,53,58,59	0
4	NAG	A	497	14/15	0.86	0.13	-	41,45,48,49	0
3	BMA	A	493	11/12	0.73	0.18	-	53,55,57,57	0
3	BMA	B	493	11/12	0.78	0.21	-	58,61,63,65	0
4	NAG	D	496	14/15	0.84	0.21	-	57,60,65,65	0
3	NAG	C	493	14/15	0.93	0.14	-	43,46,49,53	0
4	NAG	B	496	14/15	0.89	0.13	-	44,49,53,54	0

## 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(Å <sup>2</sup> )	Q<0.9
2	NAG	C	494	14/15	0.92	0.22	-	51,56,59,60	0
2	NAG	C	495	14/15	0.93	0.19	-	50,56,57,59	0
2	NAG	A	495	14/15	0.94	0.14	-	46,48,52,52	0
2	NAG	B	1	14/15	0.79	0.22	-	45,50,55,58	0
2	NAG	A	1	14/15	0.89	0.16	-	44,48,54,57	0
2	NAG	D	1	14/15	0.88	0.26	-	58,63,64,65	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.