



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 06:44 AM GMT

PDB ID : 3D7G  
Title : A high resolution crystal structure of human glutamate carboxypeptidase II (GCPII) in a complex with DCMC, a urea-based inhibitor  
Authors : Lubkowski, J.; Barinka, C.  
Deposited on : 2008-05-21  
Resolution : 1.75 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

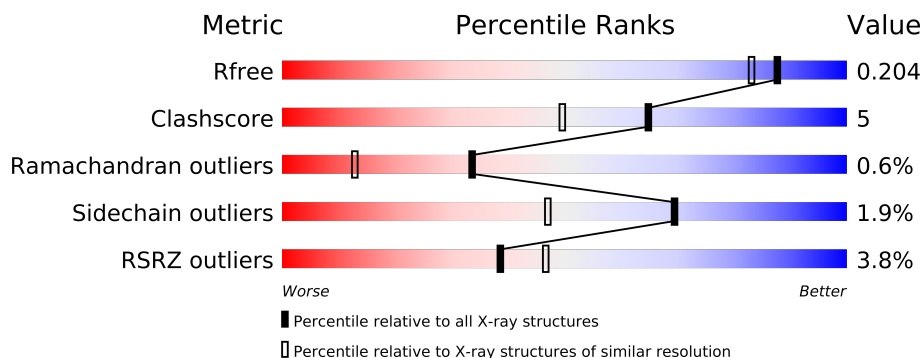
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance


The reported resolution of this entry is 1.75 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))
$R_{free}$	66092	1134 (1.76-1.76)
Clashscore	79885	1304 (1.76-1.76)
Ramachandran outliers	78287	1288 (1.76-1.76)
Sidechain outliers	78261	1288 (1.76-1.76)
RSRZ outliers	66119	1135 (1.76-1.76)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	709	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
3	NAG	A	1757	-	X
3	NAG	A	1759	-	X
3	NAG	A	1760	-	X

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 6552 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Glutamate carboxypeptidase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	691	Total	C	N	O	S	0	60	0
			5724	3686	956	1060	22			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	42	ARG	-	EXPRESSION TAG	UNP Q04609
A	43	SER	-	EXPRESSION TAG	UNP Q04609

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

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

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	42	ARG	-	EXPRESSION TAG	UNP Q04609
A	43	SER	-	EXPRESSION TAG	UNP Q04609
A	42	ARG	-	EXPRESSION TAG	UNP Q04609
A	43	SER	-	EXPRESSION TAG	UNP Q04609

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is a polymer of unknown type called SUGAR (3-MER).

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

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	42	ARG	-	EXPRESSION TAG	UNP Q04609
A	43	SER	-	EXPRESSION TAG	UNP Q04609

- Molecule 5 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	4	Total	C	N	O	0	0
			50	28	2	20		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	42	ARG	-	EXPRESSION TAG	UNP Q04609
A	43	SER	-	EXPRESSION TAG	UNP Q04609

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	2	Total	Zn	0	0
			2	2		

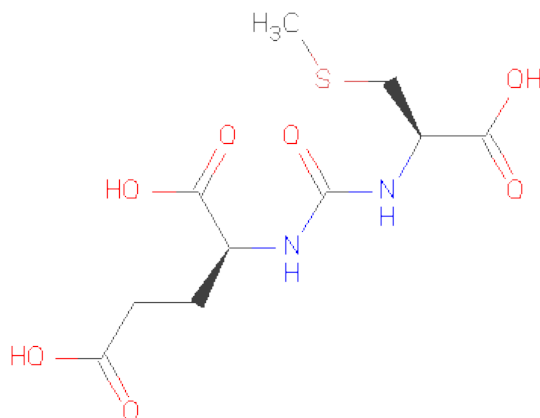
- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	Ca	0	0
			1	1		

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	1	Total	Cl	0	0
			1	1		

- Molecule 9 is N-{[(1R)-1-CARBOXY-2-(METHYLSULFANYL)ETHYL]CARBAMOYL}-L-GLUTAMICACID (three-letter code: MUD) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>2</sub>O<sub>7</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
9	A	1	Total	C	N	O	S	0	0
			20	10	2	7	1		

- Molecule 10 is water.

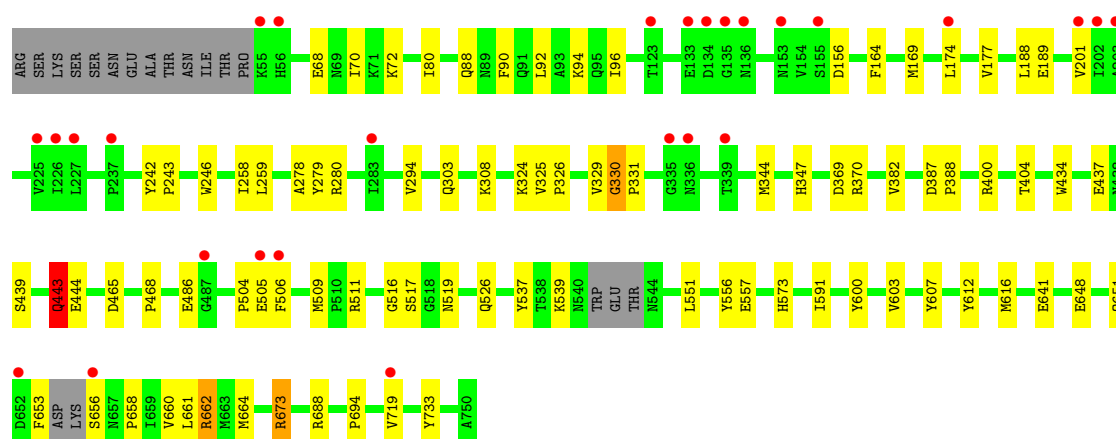
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	617	Total 617	O 617	0	1

### 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: Glutamate carboxypeptidase 2

Chain A: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	101.79Å 129.69Å 159.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 1.75 28.59 – 1.75	Depositor EDS
% Data completeness (in resolution range)	99.1 (15.00-1.75) 99.1 (28.59-1.75)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.52 (at 1.75Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.170 , 0.204 0.170 , 0.204	Depositor DCC
$R_{free}$ test set	1552 reflections (1.50%)	DCC
Wilson B-factor (Å <sup>2</sup> )	23.0	Xtriage
Anisotropy	0.472	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 45.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 105077 reflections	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	6552	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.66% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.



## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, BMA, NAG, CL, CA, MUD, 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	1.02	3/6103 (0.0%)	0.91	8/8256 (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

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	434	TRP	CE3-CZ3	7.14	1.50	1.38
1	A	443[A]	GLN	CB-CG	-5.93	1.36	1.52
1	A	557	GLU	CB-CG	-5.40	1.41	1.52

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	330	GLY	C-N-CD	-16.42	84.48	120.60
1	A	330	GLY	C-N-CA	10.30	165.25	122.00
1	A	370	ARG	NE-CZ-NH1	6.49	123.54	120.30
1	A	174[A]	LEU	CA-CB-CG	5.79	128.62	115.30
1	A	174[B]	LEU	CA-CB-CG	5.79	128.62	115.30
1	A	673	ARG	NE-CZ-NH2	-5.18	117.71	120.30
1	A	465	ASP	CB-CG-OD1	5.10	122.89	118.30
1	A	370	ARG	NE-CZ-NH2	-5.09	117.75	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	242	TYR	Mainchain,Peptide
1	A	330	GLY	Peptide
1	A	387	ASP	Peptide

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5724	0	5552	57	2
2	A	56	0	50	1	0
3	A	42	0	39	2	0
4	A	39	0	34	3	0
5	A	50	0	43	0	0
6	A	2	0	0	0	0
7	A	1	0	0	0	0
8	A	1	0	0	0	0
9	A	20	0	10	0	0
10	A	617	0	0	19	4
All	All	6552	0	5728	62	4

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 5.

All (62) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:660[A]:VAL:O	1:A:664[A]:MET:HG2	1.37	1.24
1:A:80:ILE:HD12	1:A:88[B]:GLN:HG2	1.51	0.90
1:A:658[A]:PRO:HB3	10:A:2137:HOH:O	1.82	0.79
1:A:733:TYR:HE2	10:A:2158:HOH:O	1.64	0.79
1:A:400:ARG:O	1:A:404[B]:THR:HG23	1.87	0.75
1:A:641:GLU:HG3	10:A:2171:HOH:O	1.90	0.70
1:A:90[B]:PHE:HE2	1:A:94:LYS:HZ3	1.43	0.67
1:A:468:PRO:HG2	1:A:603[A]:VAL:HG21	1.81	0.62
1:A:603[B]:VAL:CG1	1:A:607:TYR:CE2	2.85	0.60
1:A:177:VAL:HG12	1:A:188:LEU:HD11	1.83	0.60

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:603[B]:VAL:HG13	1:A:607:TYR:CZ	2.37	0.59
1:A:278:ALA:HB3	1:A:280[A]:ARG:NH1	2.17	0.59
1:A:660[A]:VAL:O	1:A:664[A]:MET:CG	2.31	0.58
1:A:653:PHE:HZ	1:A:661[A]:LEU:HD13	1.70	0.57
1:A:591[A]:ILE:HG23	1:A:661[A]:LEU:HD21	1.87	0.55
4:A:1762:NAG:C8	10:A:1985:HOH:O	2.55	0.55
1:A:443[A]:GLN:CG	10:A:2376:HOH:O	2.55	0.53
1:A:369:ASP:OD2	1:A:662[B]:ARG:NH2	2.30	0.53
1:A:486:GLU:OE1	10:A:2025:HOH:O	2.19	0.53
1:A:506:PHE:HB2	1:A:509:MET:HG3	1.89	0.53
2:A:1756:NAG:H81	10:A:2023:HOH:O	2.08	0.52
1:A:90[B]:PHE:CZ	1:A:94:LYS:HD3	2.45	0.52
1:A:517:SER:HB2	1:A:694:PRO:HG3	1.91	0.52
1:A:504:PRO:HG3	1:A:511[A]:ARG:NH1	2.26	0.51
1:A:648:GLU:O	1:A:651[B]:GLN:HG2	2.10	0.51
1:A:661[A]:LEU:O	1:A:664[A]:MET:HB2	2.11	0.51
1:A:468:PRO:CG	1:A:603[A]:VAL:HG21	2.40	0.50
1:A:437:GLU:OE1	10:A:2053:HOH:O	2.19	0.50
1:A:603[B]:VAL:CG1	1:A:607:TYR:CZ	2.94	0.50
4:A:1761:NAG:H83	10:A:2024:HOH:O	2.12	0.49
1:A:551:LEU:HD22	1:A:556:TYR:HB2	1.93	0.49
1:A:443[A]:GLN:HG3	10:A:2376:HOH:O	2.13	0.48
1:A:258:ILE:HD13	1:A:294:VAL:HB	1.95	0.48
1:A:443[A]:GLN:HG2	10:A:2376:HOH:O	2.12	0.48
4:A:1762:NAG:H82	10:A:1985:HOH:O	2.13	0.48
1:A:612:TYR:CZ	1:A:616:MET:HG3	2.49	0.48
1:A:188:LEU:CD2	1:A:329:VAL:HG11	2.44	0.47
1:A:506:PHE:HB2	1:A:509:MET:CG	2.46	0.46
1:A:733:TYR:CE2	10:A:2158:HOH:O	2.51	0.46
1:A:591[B]:ILE:HG12	10:A:1864:HOH:O	2.15	0.45
1:A:88[B]:GLN:NE2	10:A:2141:HOH:O	2.50	0.45
1:A:90[B]:PHE:CE2	1:A:94:LYS:CE	3.00	0.45
1:A:325:VAL:HB	1:A:326:PRO:HD2	1.97	0.44
1:A:516:GLY:O	1:A:526[B]:GLN:NE2	2.51	0.43
1:A:164:PHE:CG	1:A:259:LEU:HD21	2.52	0.43
1:A:177:VAL:CG2	1:A:201[B]:VAL:HG13	2.48	0.43
1:A:90[B]:PHE:CE2	1:A:94:LYS:NZ	2.84	0.43
1:A:347:HIS:HE1	10:A:2121:HOH:O	2.01	0.43
1:A:246:TRP:CD1	3:A:1760:NAG:H83	2.53	0.43
1:A:444:GLU:HA	1:A:444:GLU:OE1	2.19	0.43
1:A:308:LYS:HB2	10:A:2119:HOH:O	2.19	0.42
1:A:656[B]:SER:O	1:A:658[B]:PRO:HD3	2.19	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:688[B]:ARG:CZ	10:A:1878:HOH:O	2.67	0.42
1:A:70:ILE:HG21	1:A:573:HIS:HB3	2.02	0.41
1:A:92:LEU:O	1:A:96[A]:ILE:HG12	2.20	0.41
1:A:169:MET:HA	1:A:344:MET:O	2.20	0.41
1:A:68[A]:GLU:OE2	1:A:72:LYS:HE3	2.22	0.40
3:A:1757:NAG:H81	10:A:2090:HOH:O	2.19	0.40
1:A:439:SER:O	1:A:443[A]:GLN:HB2	2.21	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	Distance(Å)	Clash(Å)
10:A:1990:HOH:O	10:A:2344:HOH:O[2_565]	1.88	0.32
1:A:656[B]:SER:OG	10:A:2137:HOH:O[4_566]	2.01	0.19
10:A:2039:HOH:O	10:A:2173:HOH:O[2_565]	2.10	0.10
1:A:656[B]:SER:O	10:A:2137:HOH:O[4_566]	2.13	0.07

## 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	740/709 (104%)	720 (97%)	16 (2%)	4 (0%)	38 15

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	331	PRO
1	A	388	PRO
1	A	382	VAL
1	A	243	PRO

### 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	644/605 (106%)	632 (98%)	12 (2%)	69	47

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

Mol	Chain	Res	Type
1	A	189	GLU
1	A	303	GLN
1	A	443[A]	GLN
1	A	505	GLU
1	A	519	ASN
1	A	537	TYR
1	A	539	LYS
1	A	600	TYR
1	A	662[A]	ARG
1	A	662[B]	ARG
1	A	673	ARG
1	A	719	VAL

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	91	GLN
1	A	136	ASN
1	A	618	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

11 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	$\# Z  > 2$
2	NAG	A	1755	1,2	12,14,15	0.59	0	15,19,21	1.76	2 (13%)
2	NAG	A	1756	2	12,14,15	0.73	0	15,19,21	0.89	0
2	NAG	A	1758	1,2	12,14,15	0.73	0	15,19,21	1.20	2 (13%)
4	NAG	A	1761	1,4	12,14,15	1.10	1 (8%)	15,19,21	1.68	2 (13%)
4	NAG	A	1762	4	12,14,15	0.62	0	15,19,21	1.30	3 (20%)
5	NAG	A	1763	1,5	12,14,15	0.94	0	15,19,21	1.19	2 (13%)
5	NAG	A	1764	5	12,14,15	0.73	1 (8%)	15,19,21	1.64	3 (20%)
5	BMA	A	1765	5	10,11,12	0.46	0	11,15,17	1.15	1 (9%)
5	MAN	A	1766	5	10,11,12	0.81	0	11,15,17	1.71	1 (9%)
2	NAG	A	1767	2	12,14,15	0.51	0	15,19,21	1.28	1 (6%)
4	BMA	A	1768	4	10,11,12	0.58	0	11,15,17	1.30	1 (9%)

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	1755	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	1756	2	-	0/6/23/26	0/1/1/1
2	NAG	A	1758	1,2	-	0/6/23/26	0/1/1/1
4	NAG	A	1761	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1762	4	-	0/6/23/26	0/1/1/1
5	NAG	A	1763	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	1764	5	-	0/6/23/26	0/1/1/1
5	BMA	A	1765	5	-	0/2/19/22	0/1/1/1
5	MAN	A	1766	5	-	0/2/19/22	0/1/1/1
2	NAG	A	1767	2	-	0/6/23/26	0/1/1/1
4	BMA	A	1768	4	-	0/2/19/22	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1761	NAG	O7-C7	2.63	1.28	1.23
5	A	1764	NAG	O5-C5	-2.02	1.41	1.45

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1761	NAG	O5-C5-C4	-5.00	104.31	110.65
5	A	1766	MAN	O5-C5-C6	4.78	112.00	106.98
2	A	1755	NAG	O5-C5-C6	4.42	111.61	106.98
5	A	1764	NAG	O5-C5-C6	3.60	110.76	106.98
2	A	1755	NAG	C2-N2-C7	-3.31	117.53	123.09
2	A	1767	NAG	O5-C5-C6	3.22	110.36	106.98
2	A	1758	NAG	C3-C2-N2	-2.80	107.50	111.76
4	A	1768	BMA	O5-C5-C4	2.78	114.17	110.65
4	A	1761	NAG	C3-C2-N2	-2.73	107.60	111.76
4	A	1762	NAG	C8-C7-N2	2.70	121.39	116.11
5	A	1764	NAG	O5-C5-C4	-2.63	107.31	110.65
2	A	1758	NAG	O5-C5-C6	2.55	109.66	106.98
4	A	1762	NAG	O7-C7-C8	-2.50	117.17	122.04
5	A	1765	BMA	O5-C5-C4	-2.41	107.59	110.65
5	A	1763	NAG	O4-C4-C5	-2.28	103.28	109.28
5	A	1763	NAG	O5-C5-C6	2.26	109.35	106.98
5	A	1764	NAG	C3-C4-C5	-2.23	106.22	110.20
4	A	1762	NAG	O5-C5-C6	2.14	109.23	106.98

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.6 Ligand geometry

Of 8 ligands modelled in this entry, 4 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
9	MUD	A	1	-	19,19,19	0.93	1 (5%)	24,24,24	2.16	7 (29%)
3	NAG	A	1757	1	12,14,15	0.64	0	15,19,21	1.92	3 (20%)
3	NAG	A	1759	1	12,14,15	0.58	0	15,19,21	1.48	1 (6%)
3	NAG	A	1760	1	12,14,15	0.58	0	15,19,21	2.08	3 (20%)

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
9	MUD	A	1	-	-	0/24/24/24	0/0/0/0
3	NAG	A	1757	1	-	0/6/23/26	0/1/1/1
3	NAG	A	1759	1	-	0/6/23/26	0/1/1/1
3	NAG	A	1760	1	-	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)
9	A	1	MUD	O8-C7	2.05	1.29	1.22

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	A	1	MUD	C16-S15-C14	-6.90	89.52	101.14
3	A	1760	NAG	O5-C5-C6	5.08	112.31	106.98
3	A	1759	NAG	O5-C5-C6	4.48	111.68	106.98
3	A	1757	NAG	O5-C5-C4	3.82	115.51	110.65
9	A	1	MUD	C14-C13-N12	-3.78	100.72	110.95
3	A	1757	NAG	C3-C4-C5	3.78	116.95	110.20
3	A	1760	NAG	C2-N2-C7	-3.66	116.95	123.09
3	A	1760	NAG	C3-C4-C5	-3.28	104.35	110.20
3	A	1757	NAG	C6-C5-C4	-3.25	105.15	113.00
9	A	1	MUD	C5-C6-C7	-2.96	104.34	110.71
9	A	1	MUD	C7-C6-N10	-2.83	103.25	110.53
9	A	1	MUD	O9-C7-O8	-2.45	118.54	124.07
9	A	1	MUD	N12-C11-N10	2.35	118.81	114.84
9	A	1	MUD	O20-C11-N10	-2.03	118.06	122.82

There are no chirality outliers.

There are no torsion outliers.



There are no ring outliers.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	691/709 (97%)	-0.01	27 (3%)	37 45	14, 24, 42, 61	0

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	55	LYS	4.8
1	A	201[A]	VAL	4.1
1	A	505	GLU	3.9
1	A	153	ASN	3.9
1	A	719	VAL	3.7
1	A	135	GLY	3.6
1	A	506	PHE	3.6
1	A	336	ASN	3.3
1	A	227	LEU	3.1
1	A	225	VAL	3.0
1	A	226	ILE	3.0
1	A	202	ILE	3.0
1	A	335	GLY	2.9
1	A	56	HIS	2.8
1	A	237	PRO	2.8
1	A	652	ASP	2.6
1	A	487	GLY	2.6
1	A	174[A]	LEU	2.6
1	A	656[A]	SER	2.5
1	A	203	ALA	2.4
1	A	123	THR	2.3
1	A	134	ASP	2.2
1	A	283[A]	ILE	2.2
1	A	136	ASN	2.2
1	A	155[A]	SER	2.1
1	A	133	GLU	2.0
1	A	339	THR	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	NAG	A	1756	14/15	0.26	5.62	40,49,52,55	0
5	NAG	A	1764	14/15	0.20	3.39	36,38,46,52	0
4	BMA	A	1768	11/12	0.28	1.73	54,57,60,62	0
4	NAG	A	1761	14/15	0.09	1.62	28,30,34,38	0
5	MAN	A	1766	11/12	0.15	1.18	43,45,49,50	0
2	NAG	A	1755	14/15	0.08	1.16	32,38,42,44	0
2	NAG	A	1767	14/15	0.23	0.83	55,57,63,64	0
2	NAG	A	1758	14/15	0.12	0.74	39,42,45,50	0
5	NAG	A	1763	14/15	0.07	0.29	22,29,36,46	0
4	NAG	A	1762	14/15	0.17	0.20	35,42,49,50	0
5	BMA	A	1765	11/12	0.13	0.11	37,39,41,41	0

## 6.4 Ligands ⓘ

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	NAG	A	1759	14/15	0.19	6.08	60,65,68,69	0
3	NAG	A	1760	14/15	0.16	3.15	28,48,52,56	0
3	NAG	A	1757	14/15	0.28	3.08	45,51,60,64	0
7	CA	A	1753	1/1	0.06	-0.26	16,16,16,16	0
9	MUD	A	1	20/20	0.08	-0.33	17,22,37,44	0
8	CL	A	1754	1/1	0.04	-1.87	24,24,24,24	0
6	ZN	A	1752	1/1	0.05	-2.39	20,20,20,20	0
6	ZN	A	1751	1/1	0.05	-4.44	20,20,20,20	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.