



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 12:25 am GMT

PDB ID : 4NGT  
Title : Crystal Structure of Glutamate Carboxypeptidase II in a complex with urea-based inhibitor  
Authors : Tykvart, J.; Pachl, P.  
Deposited on : 2013-11-02  
Resolution : 2.31 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

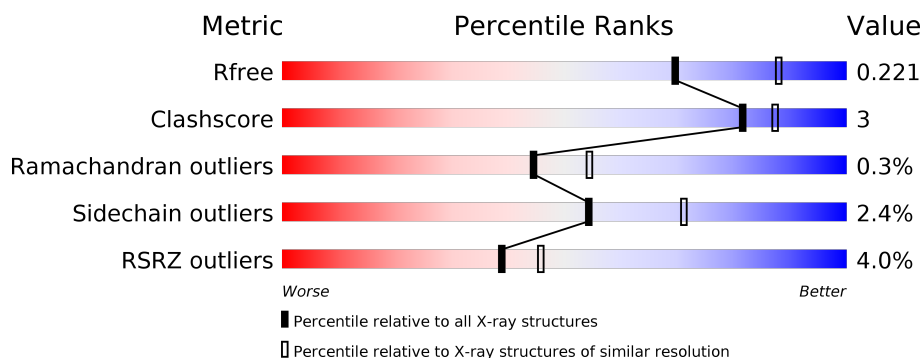
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.31 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))
$R_{free}$	100719	4787 (2.34-2.30)
Clashscore	112137	5439 (2.34-2.30)
Ramachandran outliers	110173	5386 (2.34-2.30)
Sidechain outliers	110143	5385 (2.34-2.30)
RSRZ outliers	101464	4814 (2.34-2.30)

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	739	<div> <div>4%</div> <div>84%</div> <div>9%</div> <div>6%</div> </div>

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 5985 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 Glutamate carboxypeptidase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	692	5494	3533	921	1022	18	0	1	0

There are 32 discrepancies between the modelled and reference sequences:

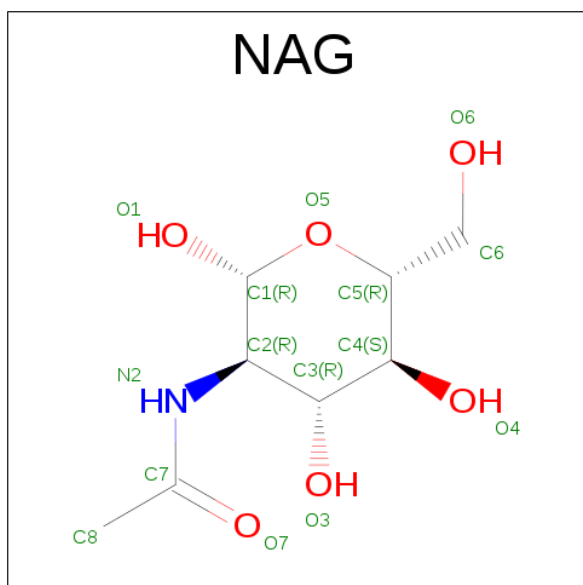
Chain	Residue	Modelled	Actual	Comment	Reference
A	12	ARG	-	EXPRESSION TAG	UNP Q04609
A	13	SER	-	EXPRESSION TAG	UNP Q04609
A	14	GLY	-	EXPRESSION TAG	UNP Q04609
A	15	LEU	-	EXPRESSION TAG	UNP Q04609
A	16	ASN	-	EXPRESSION TAG	UNP Q04609
A	17	ASP	-	EXPRESSION TAG	UNP Q04609
A	18	ILE	-	EXPRESSION TAG	UNP Q04609
A	19	PHE	-	EXPRESSION TAG	UNP Q04609
A	20	GLU	-	EXPRESSION TAG	UNP Q04609
A	21	ALA	-	EXPRESSION TAG	UNP Q04609
A	22	GLN	-	EXPRESSION TAG	UNP Q04609
A	23	LYS	-	EXPRESSION TAG	UNP Q04609
A	24	ILE	-	EXPRESSION TAG	UNP Q04609
A	25	GLU	-	EXPRESSION TAG	UNP Q04609
A	26	TRP	-	EXPRESSION TAG	UNP Q04609
A	27	HIS	-	EXPRESSION TAG	UNP Q04609
A	28	GLU	-	EXPRESSION TAG	UNP Q04609
A	29	GLY	-	EXPRESSION TAG	UNP Q04609
A	30	SER	-	EXPRESSION TAG	UNP Q04609
A	31	GLY	-	EXPRESSION TAG	UNP Q04609
A	32	SER	-	EXPRESSION TAG	UNP Q04609
A	33	GLY	-	EXPRESSION TAG	UNP Q04609
A	34	SER	-	EXPRESSION TAG	UNP Q04609
A	35	GLU	-	EXPRESSION TAG	UNP Q04609
A	36	ASN	-	EXPRESSION TAG	UNP Q04609
A	37	LEU	-	EXPRESSION TAG	UNP Q04609
A	38	TYR	-	EXPRESSION TAG	UNP Q04609

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	39	PHE	-	EXPRESSION TAG	UNP Q04609
A	40	GLN	-	EXPRESSION TAG	UNP Q04609
A	41	GLY	-	EXPRESSION TAG	UNP Q04609
A	42	ARG	-	EXPRESSION TAG	UNP Q04609
A	43	SER	-	EXPRESSION TAG	UNP Q04609

- 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	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		

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

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

- 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		

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

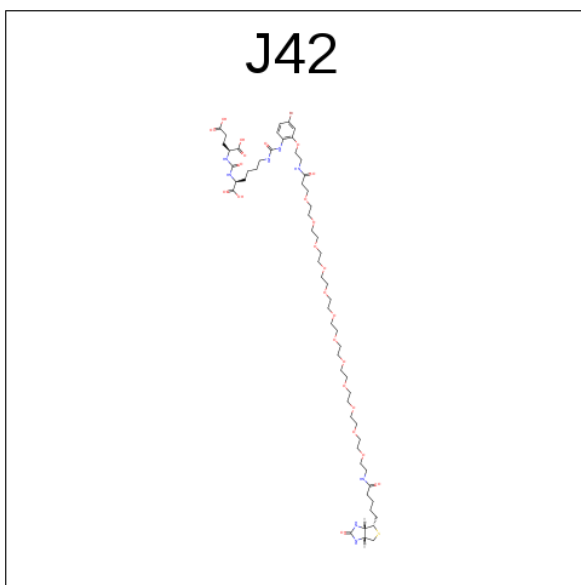
- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 8 is SODIUM ION (three-letter code: NA) (formula: Na).

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

- Molecule 9 is N-{[(1S)-5-({[4-BROMO-2-({4,44-DIOXO-48-[(3AS,4S,6AR)-2-OXOHEXA HYDRO-1H-THIENO[3,4-D]IMIDAZOL-4-YL]-7,10,13,16,19,22,25,28,31,34,37,40-DODE CAOXA-3,43-DIAZAOCTATETRACONT-1-YL}OXY)PHENYL]CARBAMOYL}AMIN O)-1-CARBOXYPENTYL]CARBAMOYL}-L-GLUTAMIC ACID (three-letter code: J42) (formula: C<sub>58</sub>H<sub>97</sub>BrN<sub>8</sub>O<sub>24</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	Br	C	N	O		
9	A	1	78	2	46	10	20	0	1

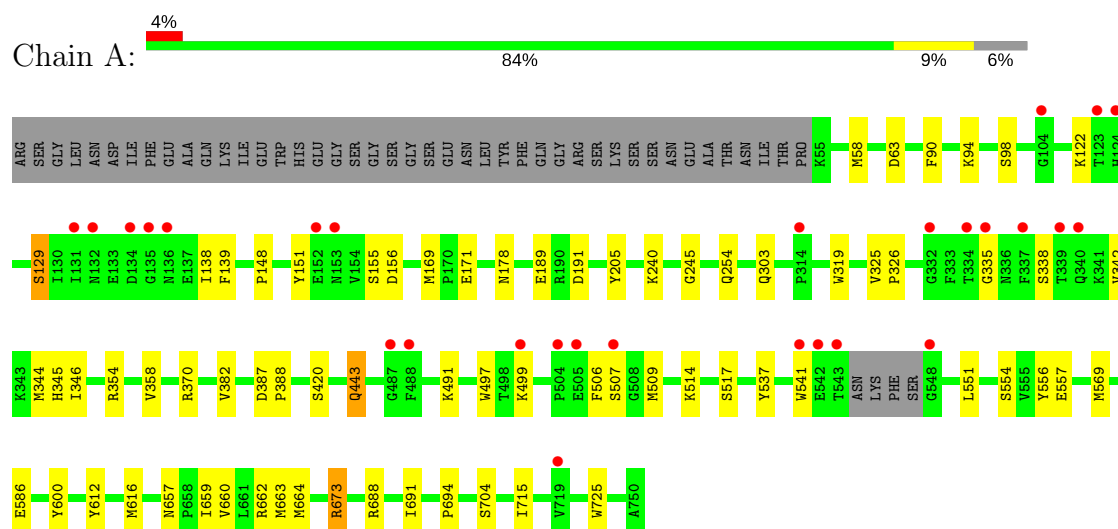
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	285	Total	O	0	0
			285	285		

### 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 the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Glutamate carboxypeptidase 2



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	101.54Å 130.30Å 159.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.19 – 2.31 47.19 – 2.31	Depositor EDS
% Data completeness (in resolution range)	98.9 (47.19-2.31) 98.9 (47.19-2.31)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.21 (at 2.32Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.176 , 0.223 0.176 , 0.221	Depositor DCC
$R_{free}$ test set	2315 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	28.5	Xtriage
Anisotropy	0.205	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 51.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5985	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.19% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, BMA, NAG, CL, NA, CA, J42

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	1.01	1/5651 (0.0%)	0.98	10/7663 (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	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	497	TRP	CG-CD1	-5.17	1.29	1.36

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	688	ARG	NE-CZ-NH2	-10.28	115.16	120.30
1	A	370	ARG	NE-CZ-NH1	7.97	124.28	120.30
1	A	688	ARG	NE-CZ-NH1	7.36	123.98	120.30
1	A	673	ARG	NE-CZ-NH1	6.50	123.55	120.30
1	A	662	ARG	NE-CZ-NH2	-6.37	117.11	120.30
1	A	569	MET	CG-SD-CE	5.96	109.73	100.20
1	A	354	ARG	NE-CZ-NH1	5.41	123.00	120.30
1	A	514	LYS	CD-CE-NZ	-5.38	99.31	111.70
1	A	664	MET	CG-SD-CE	5.25	108.59	100.20
1	A	63	ASP	CB-CG-OD1	5.11	122.90	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	129	SER	Peptide

## 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	5494	0	5325	29	0
2	A	56	0	52	0	0
3	A	28	0	24	0	0
4	A	39	0	34	0	0
5	A	2	0	0	0	0
6	A	1	0	0	0	0
7	A	1	0	0	0	0
8	A	1	0	0	0	0
9	A	78	0	52	6	0
10	A	285	0	0	1	0
All	All	5985	0	5487	33	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 3.

All (33) 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)
9:A:815[A]:J42:O7	9:A:815[A]:J42:H6	1.62	0.97
1:A:443:GLN:HG2	10:A:1066:HOH:O	1.68	0.92
1:A:659:ILE:O	1:A:663[B]:MET:HG3	1.88	0.73
1:A:58:MET:CE	1:A:586:GLU:HG2	2.24	0.68
9:A:815[A]:J42:O7	9:A:815[A]:J42:C2	2.33	0.67
1:A:541:TRP:HD1	9:A:815[B]:J42:BR	2.37	0.63
9:A:815[B]:J42:O7	9:A:815[B]:J42:H6	1.98	0.62
1:A:169:MET:HA	1:A:344:MET:O	2.08	0.54
1:A:58:MET:HE1	1:A:586:GLU:HG2	1.88	0.54
1:A:129:SER:HA	1:A:139:PHE:O	2.10	0.51
1:A:506:PHE:CB	1:A:509:MET:HG3	2.40	0.51

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:148:PRO:HG2	1:A:151:TYR:CD1	2.46	0.51
1:A:155:SER:O	1:A:156:ASP:HB2	2.12	0.49
9:A:815[A]:J42:C16	9:A:815[A]:J42:O6	2.60	0.49
1:A:358:VAL:HB	1:A:420:SER:HB3	1.95	0.48
1:A:387:ASP:HA	1:A:388:PRO:HA	1.66	0.48
1:A:191:ASP:N	1:A:191:ASP:OD1	2.45	0.48
1:A:240:LYS:O	1:A:245:GLY:HA3	2.13	0.48
1:A:58:MET:HE2	1:A:586:GLU:HG2	1.94	0.47
1:A:541:TRP:CD1	9:A:815[B]:J42:BR	3.23	0.45
1:A:612:TYR:CZ	1:A:616:MET:HG3	2.52	0.45
1:A:551:LEU:HD22	1:A:556:TYR:HB2	2.01	0.43
1:A:517:SER:HB2	1:A:694:PRO:HG3	2.01	0.43
1:A:345:HIS:O	1:A:346:ILE:HD13	2.19	0.42
1:A:325:VAL:HB	1:A:326:PRO:HD2	2.00	0.42
1:A:691:ILE:O	1:A:704:SER:HA	2.19	0.42
1:A:205:TYR:CE2	1:A:254:GLN:HB3	2.55	0.42
1:A:178:ASN:HB2	1:A:319:TRP:CE3	2.54	0.42
1:A:715:ILE:HD13	1:A:725:TRP:CE2	2.55	0.42
1:A:491:LYS:HD3	1:A:491:LYS:N	2.33	0.42
1:A:90:PHE:CZ	1:A:94:LYS:HD3	2.55	0.41
1:A:657:ASN:ND2	1:A:660:VAL:HG23	2.37	0.40
1:A:171:GLU:HA	1:A:342:VAL:O	2.22	0.40

There are no symmetry-related clashes.

## 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	689/739 (93%)	669 (97%)	18 (3%)	2 (0%)	44 54

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	335	GLY
1	A	382	VAL

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	583/629 (93%)	569 (98%)	14 (2%)	54 71

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	98	SER
1	A	122	LYS
1	A	138	ILE
1	A	189	GLU
1	A	303	GLN
1	A	338	SER
1	A	443	GLN
1	A	499	LYS
1	A	507	SER
1	A	537	TYR
1	A	554	SER
1	A	557	GLU
1	A	600	TYR
1	A	673	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 ⓘ

5 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	NAG	A	805	1,3	14,14,15	1.68	3 (21%)	15,19,21	3.22	8 (53%)
3	NAG	A	806	3	14,14,15	0.97	0	15,19,21	2.36	7 (46%)
4	NAG	A	807	1,4	14,14,15	1.07	1 (7%)	15,19,21	1.68	2 (13%)
4	NAG	A	808	4	14,14,15	0.92	1 (7%)	15,19,21	1.60	4 (26%)
4	BMA	A	809	4	11,11,12	1.11	1 (9%)	13,15,17	2.14	6 (46%)

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	NAG	A	805	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	806	3	-	0/6/23/26	0/1/1/1
4	NAG	A	807	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	808	4	-	0/6/23/26	0/1/1/1
4	BMA	A	809	4	-	0/2/19/22	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	805	NAG	O5-C1	-3.89	1.37	1.43
3	A	805	NAG	O7-C7	-2.53	1.17	1.23
4	A	808	NAG	O5-C1	-2.43	1.39	1.43
3	A	805	NAG	O3-C3	-2.35	1.37	1.43

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	807	NAG	C2-N2	-2.34	1.42	1.46
4	A	809	BMA	C4-C5	2.55	1.58	1.53

All (27) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	805	NAG	O6-C6-C5	-4.42	96.47	111.34
3	A	805	NAG	O3-C3-C4	-4.32	100.96	110.36
4	A	807	NAG	O5-C1-C2	-3.50	106.60	111.47
3	A	805	NAG	O5-C1-C2	-3.39	106.76	111.47
3	A	806	NAG	O6-C6-C5	-3.32	100.18	111.34
3	A	805	NAG	C1-C2-N2	-3.23	104.98	110.49
3	A	805	NAG	O7-C7-N2	-3.20	115.75	121.92
3	A	805	NAG	C6-C5-C4	-3.20	105.52	113.00
3	A	806	NAG	O3-C3-C4	-3.10	103.62	110.36
4	A	809	BMA	O4-C4-C3	-2.92	104.01	110.36
3	A	806	NAG	C6-C5-C4	-2.55	107.03	113.00
3	A	806	NAG	O7-C7-C8	-2.48	117.54	122.06
4	A	808	NAG	C1-C2-N2	-2.11	106.89	110.49
4	A	808	NAG	O4-C4-C3	2.16	115.06	110.36
4	A	809	BMA	O6-C6-C5	2.39	119.39	111.34
3	A	806	NAG	C8-C7-N2	2.57	120.74	116.11
4	A	809	BMA	C1-C2-C3	2.63	112.98	109.65
4	A	809	BMA	C6-C5-C4	2.72	119.36	113.00
4	A	809	BMA	C3-C4-C5	2.75	115.06	110.22
4	A	808	NAG	O3-C3-C4	2.84	116.55	110.36
4	A	808	NAG	C1-O5-C5	3.09	116.42	112.17
3	A	806	NAG	O4-C4-C5	3.32	117.66	109.28
4	A	809	BMA	O3-C3-C4	3.44	117.85	110.36
3	A	805	NAG	C8-C7-N2	3.81	122.99	116.11
4	A	807	NAG	C1-O5-C5	4.13	117.86	112.17
3	A	806	NAG	C1-O5-C5	4.96	119.01	112.17
3	A	805	NAG	C1-O5-C5	7.12	121.98	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.6 Ligand geometry

Of 11 ligands modelled in this entry, 5 are monoatomic - leaving 6 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	801	1	14,14,15	0.80	1 (7%)	15,19,21	1.60	5 (33%)
2	NAG	A	802	1	14,14,15	0.98	1 (7%)	15,19,21	2.58	7 (46%)
2	NAG	A	803	1	14,14,15	1.02	1 (7%)	15,19,21	2.33	6 (40%)
2	NAG	A	804	1	14,14,15	0.98	0	15,19,21	3.56	5 (33%)
9	J42	A	815[A]	5	30,39,94	1.00	1 (3%)	36,50,112	1.64	9 (25%)
9	J42	A	815[B]	5	30,39,94	0.95	1 (3%)	36,50,112	1.38	4 (11%)

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	801	1	-	0/6/23/26	0/1/1/1
2	NAG	A	802	1	-	0/6/23/26	0/1/1/1
2	NAG	A	803	1	-	0/6/23/26	0/1/1/1
2	NAG	A	804	1	-	0/6/23/26	0/1/1/1
9	J42	A	815[A]	5	-	0/30/40/110	0/1/1/3
9	J42	A	815[B]	5	-	0/30/40/110	0/1/1/3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	A	815[B]	J42	C3-N	-2.77	1.36	1.41
9	A	815[A]	J42	C3-N	-2.54	1.36	1.41
2	A	801	NAG	O5-C1	-2.02	1.40	1.43
2	A	802	NAG	C6-C5	2.27	1.59	1.51
2	A	803	NAG	C4-C5	2.43	1.58	1.53

All (36) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	804	NAG	C6-C5-C4	-5.79	99.45	113.00
2	A	804	NAG	O3-C3-C4	-5.78	97.79	110.36
2	A	803	NAG	O3-C3-C2	-3.35	102.22	109.39
2	A	802	NAG	O5-C1-C2	-3.32	106.85	111.47
2	A	801	NAG	O5-C1-C2	-3.00	107.30	111.47
9	A	815[B]	J42	C13-C12-C16	-2.90	108.07	112.28
9	A	815[A]	J42	C13-C12-C16	-2.73	108.32	112.28
9	A	815[B]	J42	C8-C9-C10	-2.69	108.38	112.28
2	A	802	NAG	O7-C7-C8	-2.66	117.22	122.06
2	A	803	NAG	O7-C7-C8	-2.65	117.23	122.06
9	A	815[A]	J42	C3-N-C4	-2.54	119.76	125.42
2	A	802	NAG	O3-C3-C2	-2.47	104.10	109.39
2	A	804	NAG	O7-C7-C8	-2.43	117.64	122.06
2	A	801	NAG	C2-N2-C7	-2.19	119.75	122.94
2	A	801	NAG	O3-C3-C4	-2.11	105.77	110.36
9	A	815[A]	J42	O7-C4-N	-2.02	120.14	123.58
9	A	815[A]	J42	C17-C3-N	2.05	120.47	116.70
9	A	815[A]	J42	C17-C57-C	2.23	121.24	118.72
9	A	815[B]	J42	C13-C12-N3	2.25	113.63	110.22
9	A	815[A]	J42	C13-C12-N3	2.33	113.76	110.22
2	A	803	NAG	O7-C7-N2	2.44	126.61	121.92
2	A	801	NAG	O4-C4-C5	2.44	115.44	109.28
2	A	804	NAG	O4-C4-C3	2.56	115.92	110.36
9	A	815[A]	J42	C18-O8-C17	2.59	124.01	117.70
2	A	801	NAG	C1-O5-C5	2.74	115.95	112.17
9	A	815[A]	J42	O8-C17-C3	2.98	120.21	114.98
2	A	803	NAG	O4-C4-C3	3.00	116.89	110.36
2	A	802	NAG	C3-C4-C5	3.13	115.74	110.22
2	A	803	NAG	C2-N2-C7	3.36	127.85	122.94
9	A	815[A]	J42	N-C4-N1	3.78	120.17	113.91
2	A	802	NAG	O6-C6-C5	4.06	125.00	111.34
2	A	802	NAG	C4-C3-C2	4.19	117.16	111.02
9	A	815[B]	J42	C18-O8-C17	4.90	129.63	117.70
2	A	802	NAG	C1-O5-C5	4.99	119.05	112.17
2	A	803	NAG	C1-O5-C5	5.38	119.58	112.17
2	A	804	NAG	C1-O5-C5	9.99	125.93	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 6 short contacts:



Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	A	815[A]	J42	3	0
9	A	815[B]	J42	3	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	692/739 (93%)	-0.06	28 (4%)	39 46	13, 25, 53, 90	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	548	GLY	6.4
1	A	541	TRP	4.0
1	A	136	ASN	3.4
1	A	487	GLY	3.3
1	A	132	ASN	3.0
1	A	153	ASN	2.9
1	A	719	VAL	2.9
1	A	542	GLU	2.9
1	A	334	THR	2.8
1	A	134	ASP	2.8
1	A	340	GLN	2.8
1	A	152	GLU	2.8
1	A	335	GLY	2.7
1	A	124	HIS	2.7
1	A	505	GLU	2.6
1	A	135	GLY	2.5
1	A	499	LYS	2.5
1	A	131	ILE	2.5
1	A	504	PRO	2.4
1	A	104	GLY	2.3
1	A	488	PHE	2.3
1	A	332	GLY	2.2
1	A	507	SER	2.1
1	A	314	PRO	2.1
1	A	339	THR	2.1
1	A	337	PHE	2.0
1	A	123	THR	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	543	THR	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	A	807	14/15	0.94	0.11	0.27	21,31,42,42	0
4	NAG	A	808	14/15	0.94	0.13	-	42,48,54,63	0
3	NAG	A	805	14/15	0.95	0.13	-	26,33,36,36	0
4	BMA	A	809	11/12	0.81	0.14	-	43,45,50,50	0
3	NAG	A	806	14/15	0.93	0.13	-	33,41,49,54	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	A	803	14/15	0.80	0.24	1.22	39,46,51,51	0
8	NA	A	814[B]	1/1	0.98	0.14	0.65	29,29,29,29	1
9	J42	A	815[B]	39/92	0.95	0.20	0.45	17,30,41,48	39
9	J42	A	815[A]	39/92	0.95	0.20	0.41	17,35,80,85	39
6	CL	A	812	1/1	1.00	0.18	0.23	19,19,19,19	0
7	CA	A	813	1/1	1.00	0.08	-1.83	14,14,14,14	0
5	ZN	A	811	1/1	0.99	0.09	-2.01	20,20,20,20	0
5	ZN	A	810	1/1	1.00	0.07	-3.09	21,21,21,21	0
2	NAG	A	801	14/15	0.90	0.18	-	34,42,46,49	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAG	A	802	14/15	0.80	0.30	-	49,57,73,77	0
2	NAG	A	804	14/15	0.91	0.12	-	28,50,59,65	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.