



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:30 PM GMT

PDB ID : 4P4J  
Title : X-ray structure of human glutamate carboxypeptidase II (GCP II) in complex with a phosphoramidate inhibitor T33D  
Authors : Barinka, C.  
Deposited on : 2014-03-12  
Resolution : 1.66 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

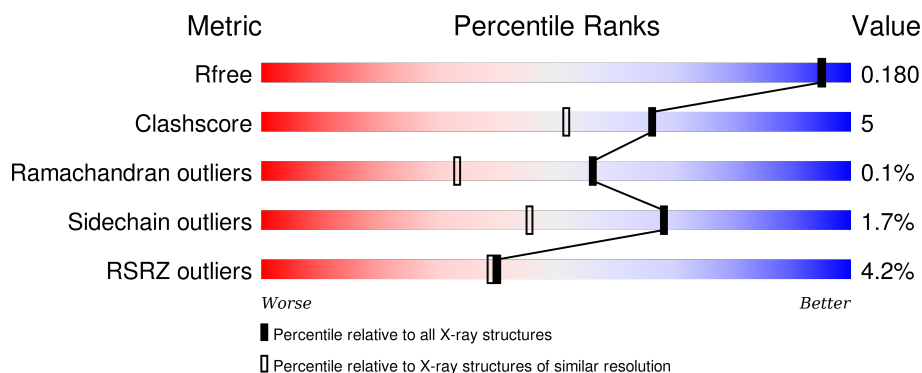
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*


The reported resolution of this entry is 1.66 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	1226 (1.66-1.66)
Clashscore	102246	1323 (1.66-1.66)
Ramachandran outliers	100387	1295 (1.66-1.66)
Sidechain outliers	100360	1295 (1.66-1.66)
RSRZ outliers	91569	1227 (1.66-1.66)

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	707	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAG	A	802	-	-	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAG	A	803	-	-	-	X
4	MAN	A	813	-	-	-	X

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 6624 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
1	A	686	Total	C	N	O	S	0	65	0
			5805	3721	973	1087	24			

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



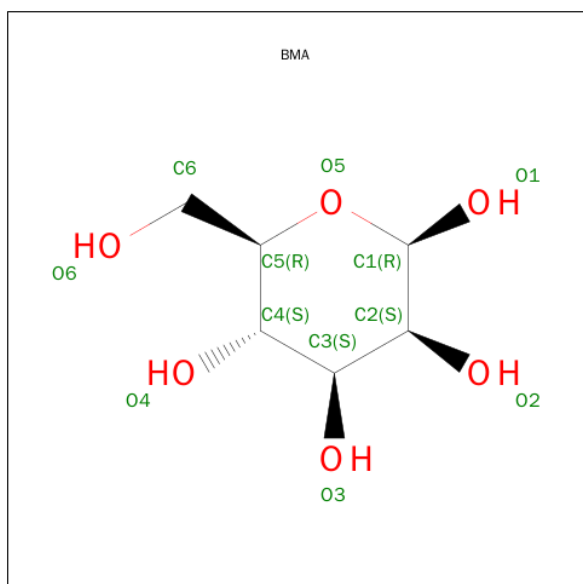
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

Continued from previous page...

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		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

- 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		

- 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	0	0
			11	6	5		

- 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		

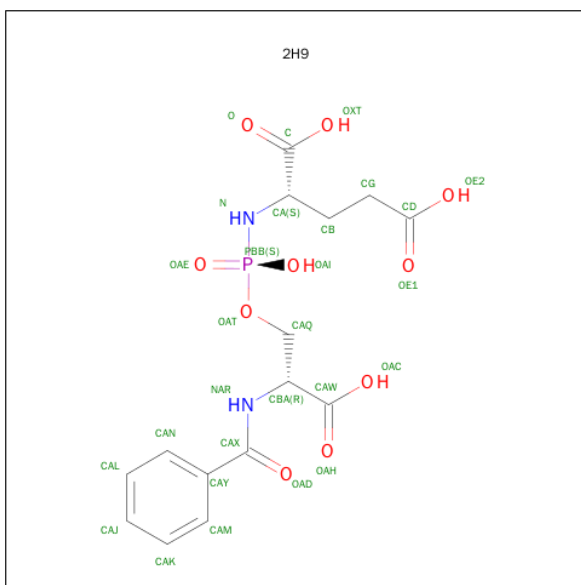
- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 8 is N-[(S)-[(2R)-2-(benzoylamino)-2-carboxyethoxy](hydroxy)phosphoryl]-L-glutamic acid (three-letter code: 2H9) (formula: C<sub>15</sub>H<sub>19</sub>N<sub>2</sub>O<sub>10</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
8	A	1	Total	C	N	O	P	0	0
			28	15	2	10	1		

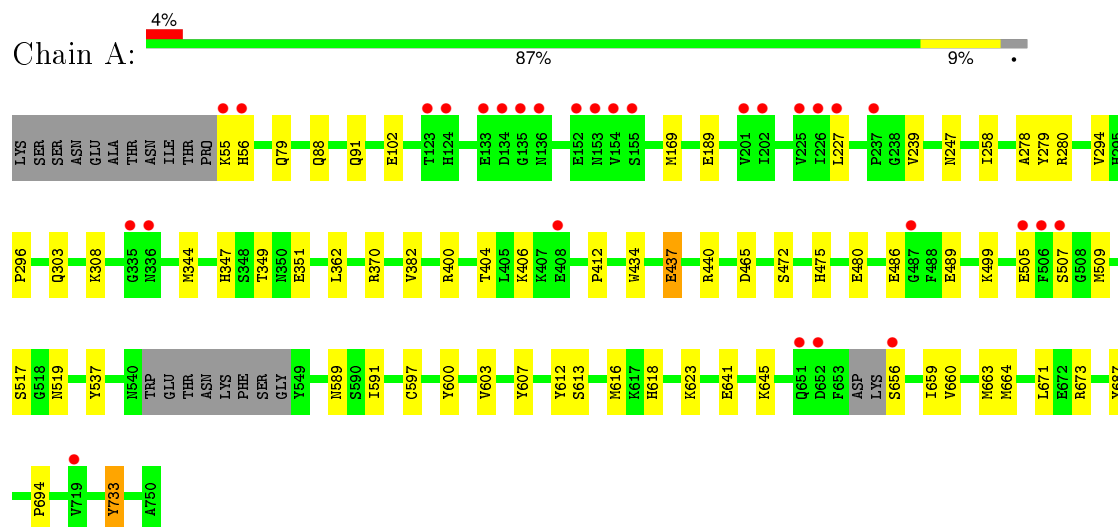
- Molecule 9 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	611	Total O 611 611	0	0

### 3 Residue-property plots

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

#### • Molecule 1: Glutamate carboxypeptidase 2





## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	102.32Å 130.03Å 159.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.68 – 1.66 29.68 – 1.66	Depositor EDS
% Data completeness (in resolution range)	98.1 (29.68-1.66) 98.1 (29.68-1.66)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.58 (at 1.66Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.156 , 0.179 0.156 , 0.180	Depositor DCC
$R_{free}$ test set	1227 reflections (1.01%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.1	Xtriage
Anisotropy	0.006	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 47.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 122265 reflections	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	6624	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.96% 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 [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: 2H9, ZN, BMA, NAG, CL, CA, 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	0.89	4/6104 (0.1%)	0.80	3/8261 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	733	TYR	CD1-CE1	5.22	1.47	1.39
1	A	437[A]	GLU	CD-OE2	-5.09	1.20	1.25
1	A	437[B]	GLU	CD-OE2	-5.09	1.20	1.25
1	A	434	TRP	CE3-CZ3	5.06	1.47	1.38

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	465	ASP	CB-CG-OD1	6.47	124.12	118.30
1	A	440	ARG	NE-CZ-NH2	-5.97	117.31	120.30
1	A	370	ARG	NE-CZ-NH1	5.42	123.01	120.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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	5805	0	5660	54	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	154	0	138	5	0
3	A	11	0	9	0	0
4	A	11	0	10	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	28	0	15	0	0
9	A	611	0	0	26	1
All	All	6624	0	5832	58	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

All (58) 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:A:660[A]:VAL:O	1:A:664[A]:MET:HG2	1.25	1.27
1:A:88[B]:GLN:NE2	1:A:91[B]:GLN:OE1	1.70	1.25
1:A:88[B]:GLN:OE1	9:A:1477:HOH:O	1.76	1.03
1:A:437[A]:GLU:OE1	9:A:901:HOH:O	1.84	0.94
1:A:597[B]:CYS:SG	1:A:671:LEU:HD22	2.14	0.88
1:A:412:PRO:HA	1:A:589[B]:ASN:HD21	1.35	0.88
1:A:88[B]:GLN:CD	1:A:91[B]:GLN:OE1	2.15	0.85
1:A:733:TYR:HE2	9:A:963:HOH:O	1.63	0.80
1:A:645[A]:LYS:HD2	9:A:1449:HOH:O	1.83	0.79
1:A:597[B]:CYS:SG	1:A:671:LEU:CD2	2.70	0.79
1:A:88[A]:GLN:HG2	9:A:1178:HOH:O	1.81	0.79
1:A:349[B]:THR:HG21	1:A:351[B]:GLU:OE1	1.85	0.77
1:A:400:ARG:O	1:A:404[B]:THR:HG23	1.85	0.77
1:A:189[B]:GLU:HG3	9:A:1479:HOH:O	1.87	0.74
1:A:641[A]:GLU:HG3	9:A:1320:HOH:O	1.89	0.71
1:A:480:GLU:HG3	9:A:1472:HOH:O	1.95	0.67
1:A:362:LEU:CD1	1:A:406:LYS:HD2	2.27	0.64
2:A:802:NAG:H81	9:A:1209:HOH:O	1.99	0.63
1:A:612:TYR:CZ	1:A:616:MET:HG3	2.36	0.61
1:A:55:LYS:HD3	1:A:56:HIS:H	1.65	0.61
1:A:349[B]:THR:CG2	1:A:351[B]:GLU:OE1	2.50	0.59
1:A:239[A]:VAL:HG22	1:A:247:ASN:ND2	2.20	0.56
1:A:623:LYS:NZ	9:A:904:HOH:O	2.39	0.56
1:A:412:PRO:HA	1:A:589[B]:ASN:ND2	2.16	0.55
1:A:733:TYR:CE2	9:A:963:HOH:O	2.47	0.55

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:802:NAG:H83	9:A:1420:HOH:O	2.08	0.54
1:A:362:LEU:HD13	1:A:406:LYS:HD2	1.91	0.52
1:A:278:ALA:HB3	1:A:280[B]:ARG:NH1	2.27	0.49
1:A:278:ALA:HB3	1:A:280[B]:ARG:CZ	2.43	0.49
2:A:803:NAG:H83	9:A:1179:HOH:O	2.12	0.49
1:A:656[B]:SER:CB	9:A:902:HOH:O	2.59	0.49
1:A:88[B]:GLN:HE22	1:A:91[B]:GLN:CD	2.12	0.49
1:A:603[B]:VAL:HG12	1:A:607:TYR:CE2	2.48	0.48
1:A:347:HIS:HE1	9:A:1280:HOH:O	1.95	0.48
1:A:517:SER:HB2	1:A:694:PRO:HG3	1.96	0.47
1:A:88[B]:GLN:OE1	1:A:91[B]:GLN:OE1	2.32	0.47
1:A:603[B]:VAL:CG1	1:A:607:TYR:CE2	2.98	0.46
1:A:591[A]:ILE:HD12	9:A:1007:HOH:O	2.15	0.46
1:A:656[B]:SER:O	9:A:902:HOH:O	2.21	0.46
1:A:308:LYS:HB2	9:A:1278:HOH:O	2.16	0.46
1:A:656[B]:SER:OG	9:A:902:HOH:O	2.20	0.46
1:A:169:MET:HA	1:A:344:MET:O	2.16	0.45
1:A:472[A]:SER:OG	9:A:1508:HOH:O	2.21	0.45
1:A:486[A]:GLU:OE1	9:A:1494:HOH:O	2.21	0.45
1:A:239[A]:VAL:HG22	1:A:247:ASN:CG	2.38	0.44
1:A:618:HIS:HE1	9:A:1128:HOH:O	2.01	0.43
1:A:347:HIS:CE1	9:A:1280:HOH:O	2.71	0.42
1:A:613[B]:SER:OG	9:A:1510:HOH:O	2.22	0.42
1:A:258:ILE:HD13	1:A:294:VAL:HB	2.00	0.42
1:A:597[B]:CYS:SG	1:A:671:LEU:HD23	2.59	0.42
1:A:79[B]:GLN:NE2	2:A:801:NAG:H81	2.36	0.41
1:A:227:LEU:O	1:A:296:PRO:HA	2.20	0.41
2:A:802:NAG:C8	9:A:1209:HOH:O	2.63	0.41
1:A:687:TYR:CZ	1:A:694:PRO:HG2	2.56	0.41
1:A:475:HIS:CE1	1:A:509[B]:MET:CE	3.04	0.40
1:A:641[B]:GLU:HG3	9:A:1449:HOH:O	2.21	0.40
1:A:659[B]:ILE:O	1:A:663[B]:MET:HG3	2.21	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
9:A:903:HOH:O	9:A:951:HOH:O[2_565]	1.81	0.39

## 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	741/707 (105%)	724 (98%)	16 (2%)	1 (0%)	56 33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	382	VAL

### 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	646/603 (107%)	635 (98%)	11 (2%)	68 45

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	102[A]	GLU
1	A	102[B]	GLU
1	A	303	GLN
1	A	489	GLU
1	A	499	LYS
1	A	505	GLU
1	A	507	SER
1	A	519	ASN
1	A	537	TYR
1	A	600	TYR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	673	ARG

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

Mol	Chain	Res	Type
1	A	618	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

12 non-standard protein/DNA/RNA residues 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	801	1,2	14,14,15	0.52	0	15,19,21	0.97	0
2	NAG	A	802	2	14,14,15	0.68	0	15,19,21	1.41	1 (6%)
2	NAG	A	803	1	14,14,15	0.73	0	15,19,21	1.06	1 (6%)
2	NAG	A	804	1,2	14,14,15	0.59	0	15,19,21	1.32	2 (13%)
2	NAG	A	805	2	14,14,15	0.62	0	15,19,21	0.87	0
2	NAG	A	806	1	14,14,15	0.49	0	15,19,21	2.33	2 (13%)
2	NAG	A	807	1	14,14,15	0.80	0	15,19,21	1.33	2 (13%)
2	NAG	A	808	1,2	14,14,15	0.91	1 (7%)	15,19,21	0.84	1 (6%)
2	NAG	A	809	2	14,14,15	0.46	0	15,19,21	1.17	1 (6%)
2	NAG	A	810	1,2	14,14,15	0.76	0	15,19,21	1.04	1 (6%)
2	NAG	A	811	3,2	14,14,15	0.65	1 (7%)	15,19,21	1.30	1 (6%)
4	MAN	A	813	3	11,11,12	0.67	0	14,15,17	1.26	1 (7%)

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	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	802	2	-	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	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	805	2	-	0/6/23/26	0/1/1/1
2	NAG	A	806	1	-	0/6/23/26	0/1/1/1
2	NAG	A	807	1	-	0/6/23/26	0/1/1/1
2	NAG	A	808	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	809	2	-	0/6/23/26	0/1/1/1
2	NAG	A	810	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	811	3,2	-	0/6/23/26	0/1/1/1
4	MAN	A	813	3	-	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)
2	A	811	NAG	O5-C1	-2.13	1.40	1.43
2	A	808	NAG	O7-C7	2.78	1.29	1.23

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	811	NAG	C3-C4-C5	-3.07	104.84	110.20
2	A	810	NAG	O4-C4-C5	-2.62	102.28	109.24
2	A	806	NAG	C6-C5-C4	-2.40	107.10	113.02
2	A	807	NAG	O5-C5-C6	2.06	111.80	107.35
2	A	804	NAG	C4-C3-C2	2.10	114.49	111.23
2	A	808	NAG	C2-N2-C7	2.12	125.77	123.04
2	A	803	NAG	O5-C5-C6	2.18	112.07	107.35
4	A	813	MAN	O5-C5-C6	2.49	112.73	107.35
2	A	809	NAG	C8-C7-N2	2.92	121.69	116.11
2	A	807	NAG	C1-O5-C5	3.40	116.56	112.25
2	A	804	NAG	C1-O5-C5	3.86	117.14	112.25
2	A	802	NAG	C2-N2-C7	3.87	128.01	123.04
2	A	806	NAG	C1-O5-C5	8.01	122.42	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	801	NAG	1	0
2	A	802	NAG	3	0
2	A	803	NAG	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 18 ligands modelled in this entry, 4 are monoatomic - leaving 14 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	1,2	14,14,15	0.52	0	15,19,21	0.97	0
2	NAG	A	802	2	14,14,15	0.68	0	15,19,21	1.41	1 (6%)
2	NAG	A	803	1	14,14,15	0.73	0	15,19,21	1.06	1 (6%)
2	NAG	A	804	1,2	14,14,15	0.59	0	15,19,21	1.32	2 (13%)
2	NAG	A	805	2	14,14,15	0.62	0	15,19,21	0.87	0
2	NAG	A	806	1	14,14,15	0.49	0	15,19,21	2.33	2 (13%)
2	NAG	A	807	1	14,14,15	0.80	0	15,19,21	1.33	2 (13%)
2	NAG	A	808	1,2	14,14,15	0.91	1 (7%)	15,19,21	0.84	1 (6%)
2	NAG	A	809	2	14,14,15	0.46	0	15,19,21	1.17	1 (6%)
2	NAG	A	810	1,2	14,14,15	0.76	0	15,19,21	1.04	1 (6%)
2	NAG	A	811	3,2	14,14,15	0.65	1 (7%)	15,19,21	1.30	1 (6%)
3	BMA	A	812	2,4	11,11,12	0.68	0	14,15,17	1.08	1 (7%)
4	MAN	A	813	3	11,11,12	0.67	0	14,15,17	1.26	1 (7%)
8	2H9	A	818	5	17,28,28	2.67	4 (23%)	21,38,38	1.95	3 (14%)

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	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	802	2	-	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	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	805	2	-	0/6/23/26	0/1/1/1
2	NAG	A	806	1	-	0/6/23/26	0/1/1/1
2	NAG	A	807	1	-	0/6/23/26	0/1/1/1
2	NAG	A	808	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	809	2	-	0/6/23/26	0/1/1/1
2	NAG	A	810	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	811	3,2	-	0/6/23/26	0/1/1/1
3	BMA	A	812	2,4	-	0/2/19/22	0/1/1/1
4	MAN	A	813	3	-	0/2/19/22	0/1/1/1
8	2H9	A	818	5	-	0/21/32/32	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	A	818	2H9	PBB-OAI	-2.85	1.48	1.56
2	A	811	NAG	O5-C1	-2.13	1.40	1.43
2	A	808	NAG	O7-C7	2.78	1.29	1.23
8	A	818	2H9	CB-CA	3.33	1.58	1.53
8	A	818	2H9	PBB-N	5.15	1.66	1.61
8	A	818	2H9	PBB-OAE	7.80	1.55	1.46

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	A	818	2H9	OAT-CAQ-CBA	-7.69	102.65	108.17
2	A	811	NAG	C3-C4-C5	-3.07	104.84	110.20
2	A	810	NAG	O4-C4-C5	-2.62	102.28	109.24
3	A	812	BMA	O3-C3-C2	-2.51	105.46	110.00
2	A	806	NAG	C6-C5-C4	-2.40	107.10	113.02
8	A	818	2H9	CAK-CAM-CAY	-2.10	117.70	120.33
2	A	807	NAG	O5-C5-C6	2.06	111.80	107.35
2	A	804	NAG	C4-C3-C2	2.10	114.49	111.23
2	A	808	NAG	C2-N2-C7	2.12	125.77	123.04
2	A	803	NAG	O5-C5-C6	2.18	112.07	107.35
4	A	813	MAN	O5-C5-C6	2.49	112.73	107.35
8	A	818	2H9	CAM-CAY-CAN	2.58	122.43	118.60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	809	NAG	C8-C7-N2	2.92	121.69	116.11
2	A	807	NAG	C1-O5-C5	3.40	116.56	112.25
2	A	804	NAG	C1-O5-C5	3.86	117.14	112.25
2	A	802	NAG	C2-N2-C7	3.87	128.01	123.04
2	A	806	NAG	C1-O5-C5	8.01	122.42	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	801	NAG	1	0
2	A	802	NAG	3	0
2	A	803	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	686/707 (97%)	-0.12	29 (4%) 40 39	14, 24, 42, 66	1 (0%)

All (29) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	153	ASN	4.5
1	A	656[A]	SER	4.2
1	A	55	LYS	3.9
1	A	506	PHE	3.4
1	A	719[A]	VAL	3.3
1	A	155	SER	3.2
1	A	336	ASN	3.2
1	A	507	SER	3.1
1	A	201[A]	VAL	3.0
1	A	237	PRO	3.0
1	A	652	ASP	3.0
1	A	135	GLY	2.9
1	A	134	ASP	2.8
1	A	123	THR	2.8
1	A	136	ASN	2.8
1	A	133	GLU	2.8
1	A	335	GLY	2.7
1	A	651[A]	GLN	2.6
1	A	487	GLY	2.6
1	A	505	GLU	2.5
1	A	227	LEU	2.5
1	A	154	VAL	2.4
1	A	226	ILE	2.4
1	A	124	HIS	2.4
1	A	225	VAL	2.2
1	A	56	HIS	2.2
1	A	152	GLU	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	408	GLU	2.0
1	A	202	ILE	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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	802	14/15	0.80	0.22	7.51	40,45,52,53	0
2	NAG	A	803	14/15	0.79	0.26	2.25	42,48,55,56	0
4	MAN	A	813	11/12	0.90	0.20	2.23	47,49,52,54	0
2	NAG	A	810	14/15	0.94	0.08	1.91	20,27,34,43	0
2	NAG	A	808	14/15	0.94	0.09	1.30	26,30,35,36	0
2	NAG	A	804	14/15	0.91	0.17	-	42,44,53,59	0
2	NAG	A	809	14/15	0.86	0.18	-	34,41,45,51	0
2	NAG	A	806	14/15	0.78	0.22	-	62,68,71,72	0
2	NAG	A	807	14/15	0.84	0.18	-	28,44,47,48	0
2	NAG	A	811	14/15	0.89	0.23	-	36,38,49,51	0
2	NAG	A	801	14/15	0.96	0.10	-	33,37,44,46	0
2	NAG	A	805	14/15	0.85	0.32	-	49,54,58,60	0

## 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(Å <sup>2</sup> )	Q<0.9
2	NAG	A	802	14/15	0.80	0.22	7.51	40,45,52,53	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	803	14/15	0.79	0.26	2.25	42,48,55,56	0
4	MAN	A	813	11/12	0.90	0.20	2.23	47,49,52,54	0
2	NAG	A	810	14/15	0.94	0.08	1.91	20,27,34,43	0
2	NAG	A	808	14/15	0.94	0.09	1.30	26,30,35,36	0
6	CA	A	816	1/1	1.00	0.07	1.25	15,15,15,15	0
8	2H9	A	818	28/28	0.97	0.08	-0.28	15,18,30,33	0
7	CL	A	817	1/1	1.00	0.08	-0.86	20,20,20,20	0
5	ZN	A	815	1/1	1.00	0.05	-2.48	18,18,18,18	0
5	ZN	A	814	1/1	1.00	0.05	-6.25	18,18,18,18	0
2	NAG	A	807	14/15	0.84	0.18	-	28,44,47,48	0
2	NAG	A	811	14/15	0.89	0.23	-	36,38,49,51	0
3	BMA	A	812	11/12	0.81	0.14	-	38,42,47,47	0
2	NAG	A	806	14/15	0.78	0.22	-	62,68,71,72	0
2	NAG	A	809	14/15	0.86	0.18	-	34,41,45,51	0
2	NAG	A	801	14/15	0.96	0.10	-	33,37,44,46	0
2	NAG	A	804	14/15	0.91	0.17	-	42,44,53,59	0
2	NAG	A	805	14/15	0.85	0.32	-	49,54,58,60	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.