



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 01:34 am GMT

PDB ID : 4MCR  
Title : A high resolution structure of human glutamate carboxypeptidase II (GCP II) in complex with folyltri-gamma-L-glutamic acid (pteroyltetra-gamma-L-glutamic acid)  
Authors : Navratil, M.; Barinka, C.; Lubkowski, J.  
Deposited on : 2013-08-21  
Resolution : 1.65 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

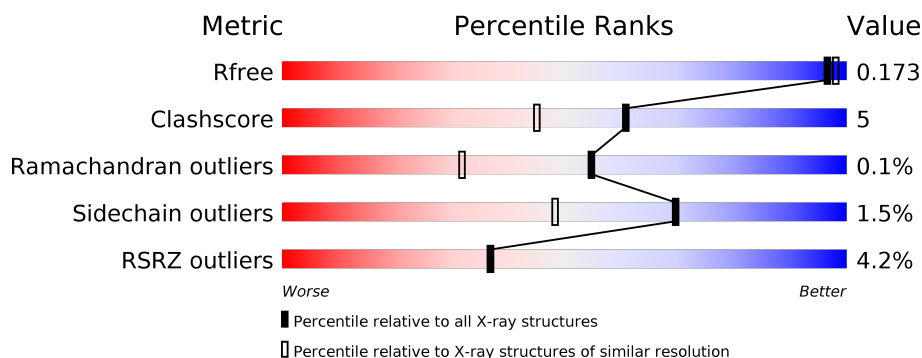
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.65 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	1368 (1.66-1.66)
Clashscore	112137	1468 (1.66-1.66)
Ramachandran outliers	110173	1438 (1.66-1.66)
Sidechain outliers	110143	1438 (1.66-1.66)
RSRZ outliers	101464	1371 (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	757	<div> <div>4%</div> <div>82%</div> <div>9%</div> <div>8%</div> </div>

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
3	CA	A	803	-	-	-	X
5	NAG	A	806	-	-	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	NAG	A	808	-	-	-	X
5	NAG	A	812	-	-	-	X
6	NAG	A	807	-	-	-	X
7	MAN	A	817	-	-	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 6681 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	693	Total	C	N	O	S	0	56	0
			5827	3740	976	1089	22			

There are 51 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MET	-	INITIATING METHIONINE	UNP Q04609
A	-5	LYS	-	EXPRESSION TAG	UNP Q04609
A	-4	LEU	-	EXPRESSION TAG	UNP Q04609
A	-3	CYS	-	EXPRESSION TAG	UNP Q04609
A	-2	ILE	-	EXPRESSION TAG	UNP Q04609
A	-1	LEU	-	EXPRESSION TAG	UNP Q04609
A	0	LEU	-	EXPRESSION TAG	UNP Q04609
A	1	ALA	-	EXPRESSION TAG	UNP Q04609
A	2	VAL	-	EXPRESSION TAG	UNP Q04609
A	3	VAL	-	EXPRESSION TAG	UNP Q04609
A	4	ALA	-	EXPRESSION TAG	UNP Q04609
A	5	PHE	-	EXPRESSION TAG	UNP Q04609
A	6	VAL	-	EXPRESSION TAG	UNP Q04609
A	7	GLY	-	EXPRESSION TAG	UNP Q04609
A	8	LEU	-	EXPRESSION TAG	UNP Q04609
A	9	SER	-	EXPRESSION TAG	UNP Q04609
A	10	LEU	-	EXPRESSION TAG	UNP Q04609
A	11	GLY	-	EXPRESSION TAG	UNP Q04609
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

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
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
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
A	424	ALA	GLU	ENGINEERED MUTATION	UNP Q04609

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

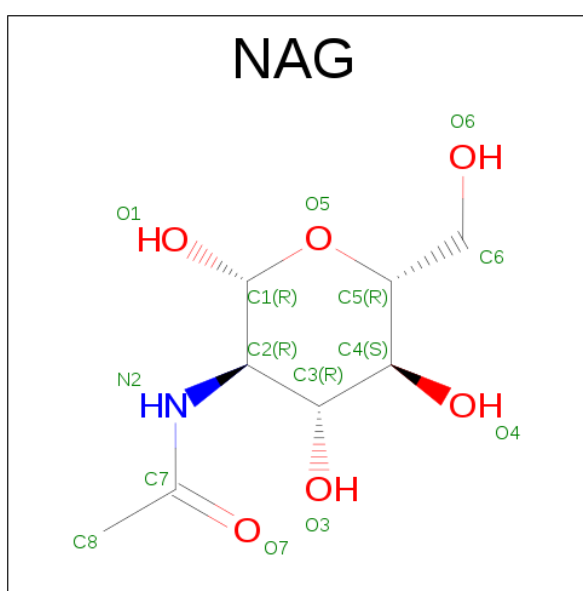
- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 5 is a polymer of unknown type called SUGAR (2-MER).

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

- Molecule 6 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



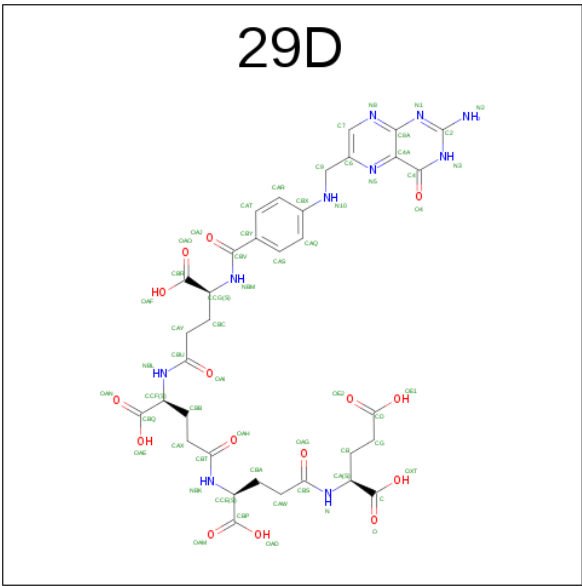
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 7 is a polymer of unknown type called SUGAR (4-MER).

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

- Molecule 8 is N-(4-[(2-AMINO-4-OXO-3,4-DIHYDROPTERIDIN-6-YL)METHYL]AMINO}BENZOYL)-L-GAMMA-GLUTAMYL-L-GAMMA-GLUTAMYL-L-GAMMA-GLUTA

MYL-L-GLUTAMIC ACID (three-letter code: 29D) (formula: C<sub>34</sub>H<sub>40</sub>N<sub>10</sub>O<sub>15</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	A	1	Total	C	N	O	0	0
			59	34	10	15		

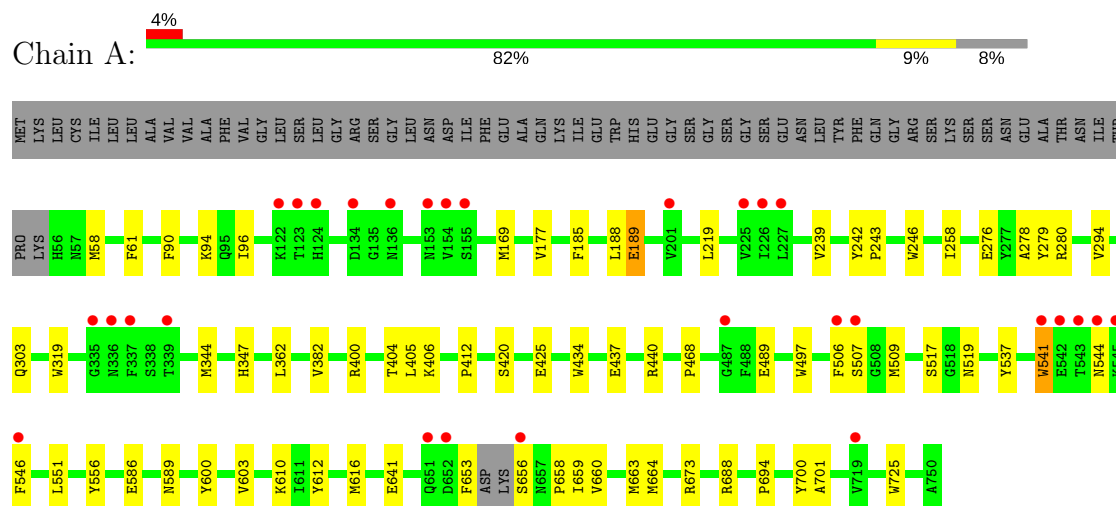
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	610	Total	O	0	5
			615	615		

### 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 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, $\alpha$ , $\beta$ , $\gamma$	101.49Å 129.83Å 158.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.48 – 1.65 29.48 – 1.65	Depositor EDS
% Data completeness (in resolution range)	91.2 (29.48-1.65) 91.2 (29.48-1.65)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.93 (at 1.65Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.133 , 0.167 0.137 , 0.173	Depositor DCC
$R_{free}$ test set	1140 reflections (1.01%)	DCC
Wilson B-factor (Å <sup>2</sup> )	21.4	Xtriage
Anisotropy	0.043	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 56.6	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	0.98	EDS
Total number of atoms	6681	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.75% 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, 29D, 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	0.65	5/6114 (0.1%)	0.67	2/8282 (0.0%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	653	PHE	C-O	5.59	1.33	1.23
1	A	434	TRP	CD2-CE2	5.42	1.47	1.41
1	A	319	TRP	CD2-CE2	5.40	1.47	1.41
1	A	725	TRP	CD2-CE2	5.40	1.47	1.41
1	A	497	TRP	CD2-CE2	5.14	1.47	1.41

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	425	GLU	OE1-CD-OE2	-8.11	113.57	123.30
1	A	440	ARG	NE-CZ-NH2	-6.46	117.07	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	5827	0	5657	53	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	2	0	0	0	0
3	A	1	0	0	0	0
4	A	1	0	0	0	0
5	A	84	0	75	2	0
6	A	42	0	39	2	0
7	A	50	0	43	0	0
8	A	59	0	35	10	0
9	A	615	0	0	14	0
All	All	6681	0	5849	63	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 5.

All (63) 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:90[B]:PHE:CE2	1:A:94:LYS:HE2	1.65	1.29
1:A:660[A]:VAL:O	1:A:664[A]:MET:HG2	1.33	1.24
1:A:90[B]:PHE:CE2	1:A:94:LYS:CE	2.35	1.08
1:A:90[B]:PHE:CZ	1:A:94:LYS:HE2	1.90	1.05
1:A:641[A]:GLU:HG3	9:A:1308:HOH:O	1.65	0.96
8:A:818:29D:H1	8:A:818:29D:CAQ	1.98	0.93
1:A:610:LYS:HE2	9:A:1075:HOH:O	1.69	0.91
1:A:90[B]:PHE:CE2	1:A:94:LYS:NZ	2.39	0.89
8:A:818:29D:CBX	8:A:818:29D:H1	2.04	0.86
1:A:412:PRO:HA	1:A:589[B]:ASN:HD21	1.38	0.86
1:A:58:MET:CE	1:A:586:GLU:HG2	2.08	0.84
1:A:58:MET:HE1	1:A:586:GLU:HG2	1.61	0.82
1:A:90[B]:PHE:CZ	1:A:94:LYS:CE	2.61	0.78
8:A:818:29D:CBX	8:A:818:29D:C7	2.62	0.76
1:A:541[A]:TRP:O	1:A:541[A]:TRP:CE3	2.38	0.76
1:A:90[B]:PHE:HE2	1:A:94:LYS:NZ	1.81	0.75
1:A:90[B]:PHE:CD2	1:A:94:LYS:HE2	2.25	0.70
1:A:400:ARG:O	1:A:404[B]:THR:HG23	1.92	0.70
1:A:185:PHE:O	1:A:189[A]:GLU:HG2	1.94	0.68
1:A:541[B]:TRP:CE2	8:A:818:29D:H8	2.35	0.62
1:A:90[B]:PHE:HE2	1:A:94:LYS:HZ3	1.35	0.61
1:A:541[A]:TRP:CD2	1:A:541[A]:TRP:O	2.55	0.60
5:A:813:NAG:H82	9:A:1275:HOH:O	2.02	0.59
1:A:701:ALA:HA	8:A:818:29D:H17	1.86	0.58
1:A:420[B]:SER:O	1:A:420[B]:SER:OG	2.23	0.57

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:A:813:NAG:C8	9:A:1275:HOH:O	2.53	0.57
1:A:276[A]:GLU:O	1:A:276[A]:GLU:HG3	2.05	0.55
1:A:688[B]:ARG:NH2	9:A:969:HOH:O	2.05	0.54
1:A:506:PHE:HB2	1:A:509:MET:HG3	1.91	0.53
1:A:278:ALA:HB3	1:A:280[A]:ARG:NH1	2.23	0.53
1:A:412:PRO:HA	1:A:589[B]:ASN:ND2	2.17	0.53
1:A:612:TYR:CZ	1:A:616[A]:MET:HG3	2.45	0.51
1:A:362:LEU:HD11	1:A:406:LYS:HG3	1.94	0.50
1:A:541[A]:TRP:CE3	1:A:541[A]:TRP:C	2.84	0.49
6:A:807:NAG:C8	9:A:1294:HOH:O	2.60	0.49
1:A:656[B]:SER:O	1:A:658[B]:PRO:HD3	2.14	0.47
8:A:818:29D:H15	9:A:1494:HOH:O	2.14	0.47
1:A:58:MET:HE3	1:A:586:GLU:HG2	1.96	0.47
1:A:659[B]:ILE:O	1:A:663[B]:MET:HG3	2.15	0.47
1:A:517:SER:HB2	1:A:694:PRO:HG3	1.97	0.46
8:A:818:29D:CBS	9:A:1339:HOH:O	2.64	0.46
1:A:276[B]:GLU:HB2	9:A:1065:HOH:O	2.16	0.45
1:A:610:LYS:CE	9:A:1075:HOH:O	2.44	0.45
1:A:169:MET:HA	1:A:344:MET:O	2.18	0.44
1:A:700:TYR:HB3	8:A:818:29D:H23	2.01	0.43
8:A:818:29D:C7	8:A:818:29D:CAR	2.96	0.43
1:A:242:TYR:CG	1:A:243:PRO:HA	2.54	0.43
1:A:219:LEU:HD12	9:A:1397:HOH:O	2.19	0.42
1:A:61:PHE:HB2	1:A:405[A]:LEU:HD11	2.02	0.42
1:A:258:ILE:HD13	1:A:294:VAL:HB	2.01	0.42
1:A:437:GLU:OE1	9:A:1358:HOH:O	2.21	0.42
1:A:96[B]:ILE:HD13	9:A:985:HOH:O	2.19	0.41
1:A:468:PRO:HG2	1:A:603:VAL:HG21	2.03	0.41
1:A:544:ASN:HB3	1:A:546:PHE:CE1	2.54	0.41
1:A:177:VAL:HG12	1:A:188:LEU:HD11	2.02	0.41
1:A:347:HIS:HE1	9:A:1289:HOH:O	2.02	0.41
1:A:551:LEU:HD22	1:A:556:TYR:HB2	2.02	0.41
1:A:246:TRP:CD1	6:A:811:NAG:H83	2.56	0.41
1:A:489:GLU:H	1:A:489:GLU:CD	2.25	0.40
8:A:818:29D:H19	8:A:818:29D:H15	1.58	0.40
1:A:468:PRO:CG	1:A:603:VAL:HG21	2.51	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	742/757 (98%)	724 (98%)	17 (2%)	1 (0%)	55	32

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	643/642 (100%)	631 (98%)	12 (2%)	62	38

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

Mol	Chain	Res	Type
1	A	189[A]	GLU
1	A	189[B]	GLU
1	A	239[A]	VAL
1	A	239[B]	VAL
1	A	303	GLN
1	A	507	SER
1	A	519	ASN
1	A	537	TYR
1	A	541[A]	TRP
1	A	541[B]	TRP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	600	TYR
1	A	673	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	544	ASN
1	A	651	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 ⓘ

10 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	$\# Z  > 2$
5	NAG	A	805	1,5	14,14,15	0.61	0	15,19,21	0.95	0
5	NAG	A	806	5	14,14,15	0.48	0	15,19,21	1.16	1 (6%)
5	NAG	A	808	1,5	14,14,15	0.49	0	15,19,21	1.14	1 (6%)
5	NAG	A	809	5	14,14,15	0.48	0	15,19,21	0.78	0
5	NAG	A	812	1,5	14,14,15	0.62	0	15,19,21	1.23	2 (13%)
5	NAG	A	813	5	14,14,15	0.43	0	15,19,21	1.19	2 (13%)
7	NAG	A	814	1,7	14,14,15	0.61	0	15,19,21	1.39	2 (13%)
7	NAG	A	815	7	14,14,15	0.55	0	15,19,21	0.88	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	BMA	A	816	7	11,11,12	0.43	0	13,15,17	0.69	0
7	MAN	A	817	7	11,11,12	0.72	0	13,15,17	0.60	0

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
5	NAG	A	805	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	806	5	-	0/6/23/26	0/1/1/1
5	NAG	A	808	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	809	5	-	0/6/23/26	0/1/1/1
5	NAG	A	812	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	813	5	-	0/6/23/26	0/1/1/1
7	NAG	A	814	1,7	-	0/6/23/26	0/1/1/1
7	NAG	A	815	7	-	0/6/23/26	0/1/1/1
7	BMA	A	816	7	-	0/2/19/22	0/1/1/1
7	MAN	A	817	7	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	814	NAG	O5-C1-C2	-2.84	107.53	111.47
5	A	813	NAG	O7-C7-C8	-2.04	118.35	122.06
5	A	808	NAG	C1-O5-C5	2.26	115.28	112.17
5	A	812	NAG	C1-O5-C5	2.33	115.38	112.17
5	A	812	NAG	C2-N2-C7	2.46	126.54	122.94
5	A	806	NAG	C2-N2-C7	2.49	126.58	122.94
5	A	813	NAG	C8-C7-N2	2.63	120.86	116.11
7	A	814	NAG	C1-O5-C5	3.49	116.97	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	813	NAG	2	0

## 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$
6	NAG	A	807	1	14,14,15	0.62	0	15,19,21	1.09	1 (6%)
6	NAG	A	810	1	14,14,15	0.36	0	15,19,21	2.54	2 (13%)
6	NAG	A	811	1	14,14,15	0.54	0	15,19,21	1.16	2 (13%)
8	29D	A	818	2	45,61,61	1.20	2 (4%)	53,83,83	2.12	16 (30%)

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
6	NAG	A	807	1	-	0/6/23/26	0/1/1/1
6	NAG	A	810	1	-	0/6/23/26	0/1/1/1
6	NAG	A	811	1	-	0/6/23/26	0/1/1/1
8	29D	A	818	2	-	0/43/61/61	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	A	818	29D	C4A-N5	2.63	1.37	1.33
8	A	818	29D	O4-C4	4.90	1.36	1.24

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	A	818	29D	CBB-CCF-NBL	-5.82	101.38	110.22
8	A	818	29D	C7-C6-N5	-5.10	117.45	120.80
8	A	818	29D	N1-C2-N3	-4.78	120.48	127.46
8	A	818	29D	CBA-CCE-CBP	-3.36	107.41	112.28
8	A	818	29D	CBA-CAW-CBS	-3.28	105.76	113.18

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	A	818	29D	C4A-C4-N3	-2.85	119.42	123.48
8	A	818	29D	CB-CA-C	-2.55	108.58	112.28
6	A	807	NAG	C1-C2-N2	-2.44	106.31	110.49
8	A	818	29D	OAJ-CBV-CBY	-2.39	116.71	120.94
8	A	818	29D	C4-C4A-C8A	-2.37	118.04	119.96
8	A	818	29D	CAW-CBS-N	-2.24	111.88	115.82
6	A	811	NAG	O5-C1-C2	-2.04	108.64	111.47
6	A	810	NAG	O5-C1-C2	2.08	114.36	111.47
6	A	811	NAG	C1-O5-C5	2.67	115.84	112.17
8	A	818	29D	C9-C6-N5	2.72	121.71	116.58
8	A	818	29D	N8-C8A-N1	2.73	119.56	116.04
8	A	818	29D	C4-C4A-N5	2.87	121.83	118.68
8	A	818	29D	C4-N3-C2	3.62	121.26	116.06
8	A	818	29D	CCG-NBM-CBV	3.66	127.16	122.15
8	A	818	29D	C2-N1-C8A	4.30	120.18	115.16
6	A	810	NAG	C1-O5-C5	9.01	124.59	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	807	NAG	1	0
6	A	811	NAG	1	0
8	A	818	29D	10	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	693/757 (91%)	-0.24	29 (4%) 37 37	13, 22, 40, 61	11 (1%)

All (29) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	541[A]	TRP	4.6
1	A	543	THR	4.1
1	A	155[A]	SER	3.9
1	A	544	ASN	3.7
1	A	506	PHE	3.7
1	A	507	SER	3.6
1	A	123	THR	3.4
1	A	153	ASN	3.3
1	A	136[A]	ASN	3.2
1	A	546	PHE	3.2
1	A	719	VAL	3.2
1	A	336	ASN	3.1
1	A	134	ASP	2.9
1	A	335	GLY	2.9
1	A	154	VAL	2.9
1	A	201[A]	VAL	2.8
1	A	225	VAL	2.7
1	A	656[A]	SER	2.7
1	A	124	HIS	2.4
1	A	652	ASP	2.4
1	A	487	GLY	2.4
1	A	339	THR	2.3
1	A	651	GLN	2.2
1	A	542	GLU	2.2
1	A	227	LEU	2.2
1	A	337	PHE	2.2
1	A	545	LYS	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	122	LYS	2.1
1	A	226	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
5	NAG	A	806	14/15	0.85	0.22	4.53	35,44,55,58	0
7	MAN	A	817	11/12	0.94	0.21	4.18	39,46,52,53	0
5	NAG	A	812	14/15	0.93	0.10	2.84	25,28,35,37	0
5	NAG	A	808	14/15	0.93	0.15	2.03	38,41,46,50	0
7	NAG	A	814	14/15	0.94	0.06	0.55	20,26,33,45	0
5	NAG	A	805	14/15	0.97	0.08	-	28,36,41,49	0
7	NAG	A	815	14/15	0.92	0.23	-	32,36,57,63	0
5	NAG	A	813	14/15	0.93	0.20	-	32,39,49,53	0
5	NAG	A	809	14/15	0.85	0.28	-	46,55,64,69	0
7	BMA	A	816	11/12	0.92	0.19	-	31,35,38,40	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
3	CA	A	803	1/1	1.00	0.08	2.57	15,15,15,15	0
6	NAG	A	807	14/15	0.77	0.28	2.21	38,46,67,68	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
8	29D	A	818	59/59	0.92	0.19	1.84	16,72,120,127	0
4	CL	A	804	1/1	1.00	0.08	-1.33	22,22,22,22	0
2	ZN	A	802	1/1	1.00	0.04	-3.06	18,18,18,18	0
2	ZN	A	801	1/1	1.00	0.05	-5.10	17,17,17,17	0
6	NAG	A	811	14/15	0.93	0.14	-	27,39,47,48	0
6	NAG	A	810	14/15	0.90	0.19	-	57,67,80,81	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.