



# wwPDB X-ray Structure Validation Summary Report

Jun 12, 2014 – 07:13 AM EDT

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 wwPDB validation summary 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>

---

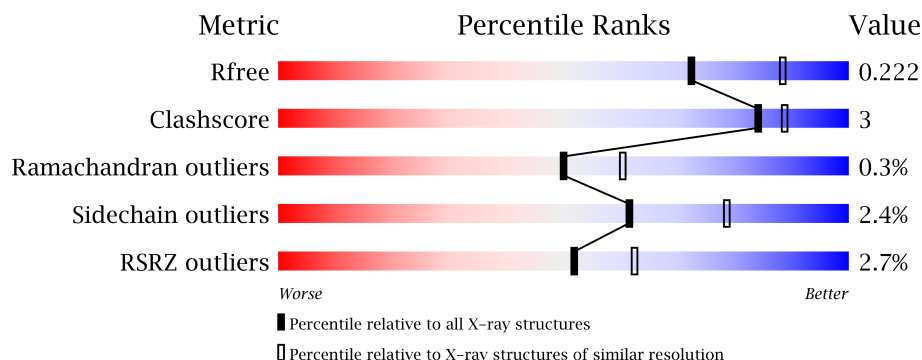
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable23161  
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) : stable23161

# 1 Overall quality at a glance

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))
$R_{free}$	66092	3293 (2.34-2.30)
Clashscore	79885	4097 (2.34-2.30)
Ramachandran outliers	78287	4055 (2.34-2.30)
Sidechain outliers	78261	4054 (2.34-2.30)
RSRZ outliers	66119	3294 (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. 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	739	

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

Mol	Type	Chain	Res	Geometry	Electron density
2	NAG	A	801	-	X

## 2 Entry composition i

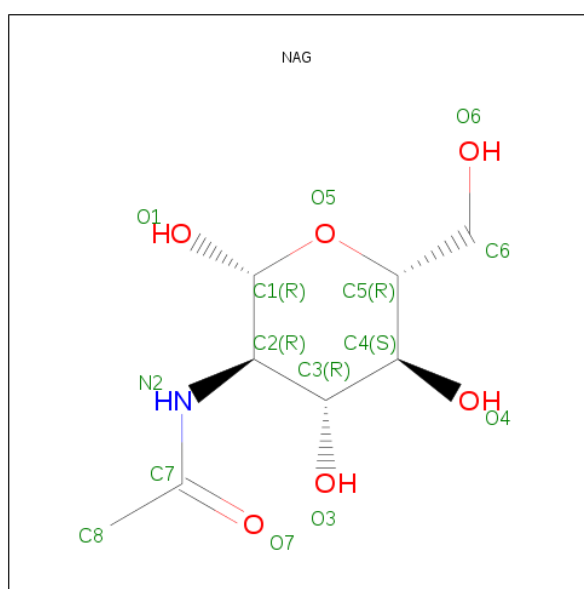
There are 10 unique types of molecules in this entry. The entry contains 5985 atoms, of which 0 are hydrogen and 0 are deuterium.

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
			Total	C	N	O	S			
1	A	692	5494	3533	921	1022	18	0	1	0

- 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		

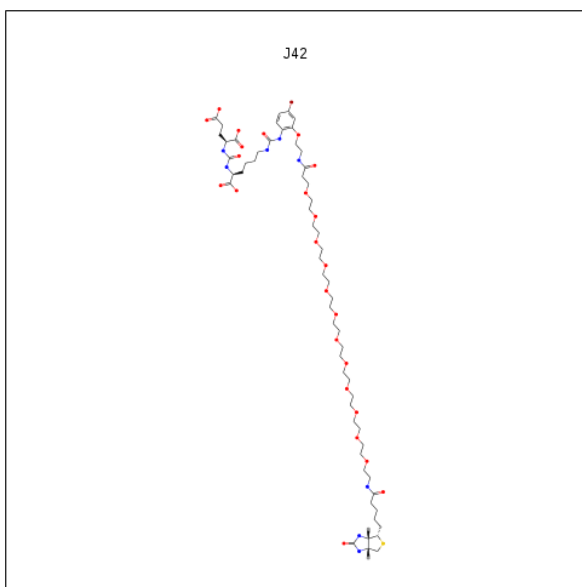
- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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-GLUTAMICACID (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

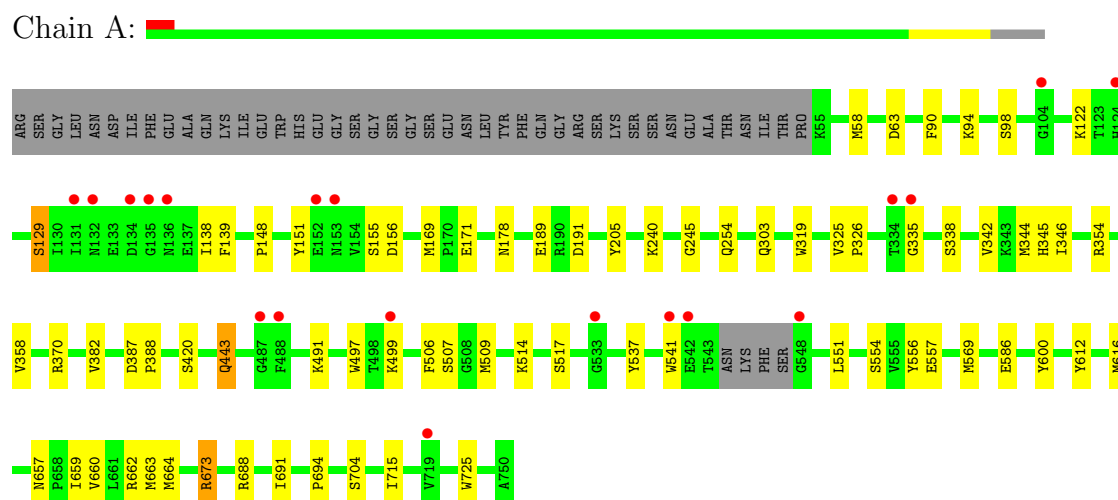
- Molecule 10 is water.

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

### 3 Residue-property plots

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 ( $\text{RSRZ} > 2$ ). 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 0.177 , 0.222	Depositor DCC
$R_{free}$ test set	2316 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> )	0.37 , 34.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 46306 reflections	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.

## 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

The worst 5 of 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

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	129	SER	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	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	0	5	0
10	A	285	0	0	1	0
All	All	5985	0	5435	32	0

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 3.

The worst 5 of 32 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:443:GLN:HG2	10:A:1066:HOH:O	1.68	0.92
9:A:815[A]:J42:O7	9:A:815[A]:J42:C2	2.33	0.75
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
1:A:541:TRP:HD1	9:A:815[B]:J42:BR	2.37	0.63

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	689/739 (93%)	669 (97%)	18 (3%)	2 (0%)	50 60

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%)	61 78

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	443	GLN
1	A	499	LYS
1	A	557	GLU
1	A	338	SER
1	A	554	SER

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 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 ⓘ

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	12,14,15	1.75	4 (33%)	15,19,21	3.67	9 (60%)
3	NAG	A	806	3	12,14,15	0.91	1 (8%)	15,19,21	2.38	9 (60%)
4	NAG	A	807	1,4	12,14,15	1.43	1 (8%)	15,19,21	1.35	2 (13%)
4	NAG	A	808	4	12,14,15	1.12	1 (8%)	15,19,21	2.12	4 (26%)
4	BMA	A	809	4	10,11,12	1.18	1 (10%)	11,15,17	2.34	6 (54%)

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

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	807	NAG	C2-N2	-3.88	1.42	1.46
3	A	805	NAG	C2-N2	-3.14	1.43	1.46
3	A	805	NAG	O5-C5	-3.13	1.40	1.45
4	A	808	NAG	O5-C5	-3.08	1.40	1.45
4	A	809	BMA	C4-C5	2.48	1.58	1.53

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	805	NAG	O5-C5-C6	-8.33	98.24	106.98
3	A	805	NAG	O5-C5-C4	6.53	118.94	110.65
4	A	808	NAG	C3-C2-N2	-6.35	102.09	111.62
3	A	806	NAG	O5-C5-C4	5.07	117.09	110.65
3	A	805	NAG	O6-C6-C5	-4.28	96.47	111.37

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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	12,14,15	0.68	0	15,19,21	1.91	5 (33%)
2	NAG	A	802	1	12,14,15	1.04	1 (8%)	15,19,21	3.25	7 (46%)
2	NAG	A	803	1	12,14,15	1.13	1 (8%)	15,19,21	2.35	7 (46%)
2	NAG	A	804	1	12,14,15	1.09	1 (8%)	15,19,21	3.61	6 (40%)
9	J42	A	815[A]	5	38,39,94	1.09	3 (7%)	48,50,112	1.57	11 (22%)
9	J42	A	815[B]	5	38,39,94	1.05	3 (7%)	48,50,112	1.36	6 (12%)

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

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	J42	A	815[A]	5	-	0/40/40/110	0/1/1/3
9	J42	A	815[B]	5	-	0/40/40/110	0/1/1/3

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	A	815[A]	J42	O1-C10	3.00	1.32	1.22
9	A	815[B]	J42	O1-C10	2.97	1.32	1.22
9	A	815[B]	J42	C3-N	-2.80	1.36	1.41
2	A	804	NAG	C2-N2	-2.65	1.43	1.46
9	A	815[A]	J42	C3-N	-2.57	1.36	1.41

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	802	NAG	O5-C5-C6	9.57	117.02	106.98
2	A	804	NAG	O5-C5-C6	8.72	116.14	106.98
2	A	804	NAG	O3-C3-C4	-5.63	97.79	110.36
2	A	804	NAG	O5-C5-C4	5.62	117.79	110.65
2	A	804	NAG	C6-C5-C4	-5.56	99.45	113.04

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	692/739 (93%)	-0.18	19 (2%) 52 62	13, 25, 53, 90	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	548	GLY	5.3
1	A	541	TRP	3.9
1	A	136	ASN	3.4
1	A	487	GLY	3.0
1	A	335	GLY	2.8

### 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
3	NAG	A	805	14/15	0.12	1.15	26,33,36,36	0
4	BMA	A	809	11/12	0.13	0.00	43,45,50,50	0
4	NAG	A	807	14/15	0.09	-0.17	21,31,42,42	0
4	NAG	A	808	14/15	0.12	-0.21	42,48,54,63	0
3	NAG	A	806	14/15	0.13	-	33,41,49,54	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
2	NAG	A	801	14/15	0.18	6.06	34,42,46,49	0
2	NAG	A	804	14/15	0.12	1.49	28,50,59,65	0
2	NAG	A	802	14/15	0.28	1.35	49,57,73,77	0
2	NAG	A	803	14/15	0.21	0.85	39,46,51,51	0
9	J42	A	815[B]	39/92	0.17	0.37	17,30,41,48	39
9	J42	A	815[A]	39/92	0.17	0.33	17,35,80,85	39
8	NA	A	814[B]	1/1	0.10	-0.63	29,29,29,29	1
6	CL	A	812	1/1	0.12	-0.69	19,19,19,19	0
7	CA	A	813	1/1	0.08	-0.93	14,14,14,14	0
5	ZN	A	811	1/1	0.07	-1.34	20,20,20,20	0
5	ZN	A	810	1/1	0.06	-3.87	21,21,21,21	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.