



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 18, 2019 – 02:29 PM EDT

PDB ID : 6FE5  
Title : X-ray structure of human glutamate carboxypeptidase II (GCPII) - the E424M inactive mutant, in complex with a inhibitor JHU 2249  
Authors : Barinka, C.; Novakova, Z.; Motlova, L.  
Deposited on : 2017-12-29  
Resolution : 1.52 Å(reported)

This is a wwPDB X-ray Structure 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.0 (224370), CSD as540be (2019)  
Xtriage (Phenix) : 1.13  
EDS : rb-20031633  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20031633

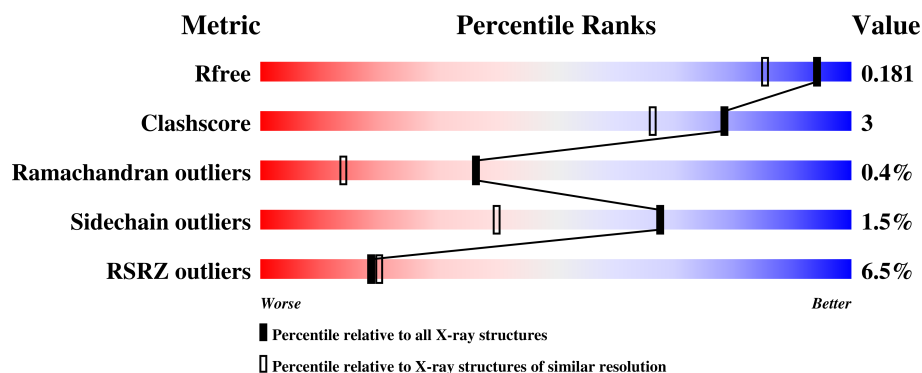
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


The reported resolution of this entry is 1.52 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))
$R_{free}$	111664	3365 (1.54-1.50)
Clashscore	122126	3586 (1.54-1.50)
Ramachandran outliers	120053	3504 (1.54-1.50)
Sidechain outliers	120020	3502 (1.54-1.50)
RSRZ outliers	108989	3301 (1.54-1.50)

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	

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 6638 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.

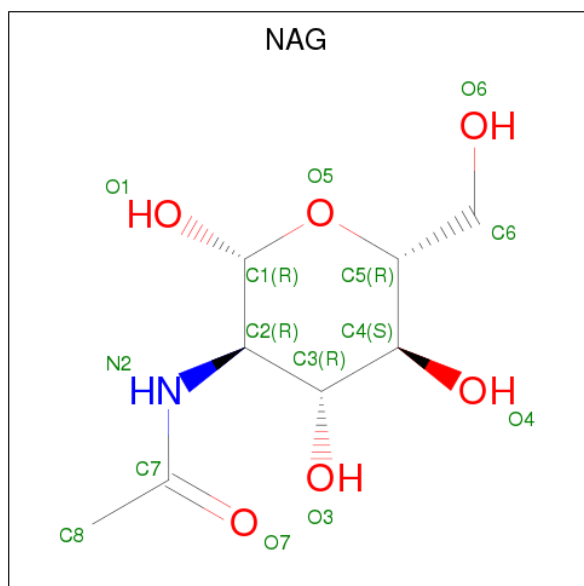
- Molecule 1 is a protein called Glutamate carboxypeptidase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	693	Total	C	N	O	S	0	50	0
			5824	3733	979	1089	23			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	424	MET	GLU	engineered mutation	UNP Q04609

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



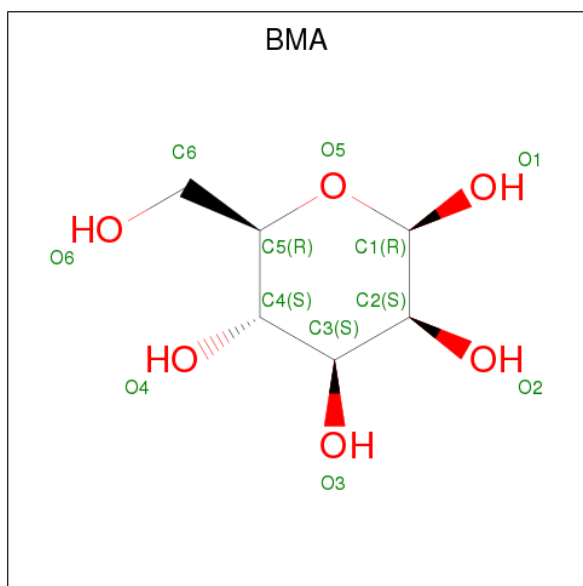
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

Continued from previous page...

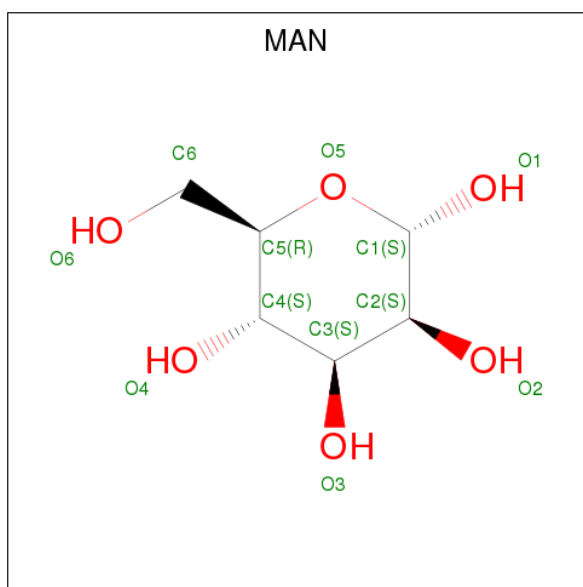
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

- 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		
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		
4	A	1	Total	C	O	0	0
			11	6	5		
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		

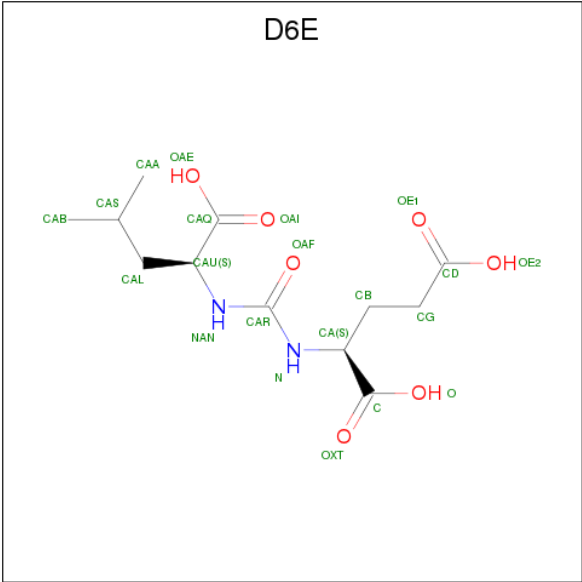
- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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 (2 {S})-2-[[ (2 {S})-4-methyl-1-oxidanyl-1-oxidanylidene-pentan-2-yl]carbamoylamino]pentanedioic acid (three-letter code: D6E) (formula: C<sub>12</sub>H<sub>20</sub>N<sub>2</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	A	1	Total	C	N	O	0	0
			21	12	2	7		

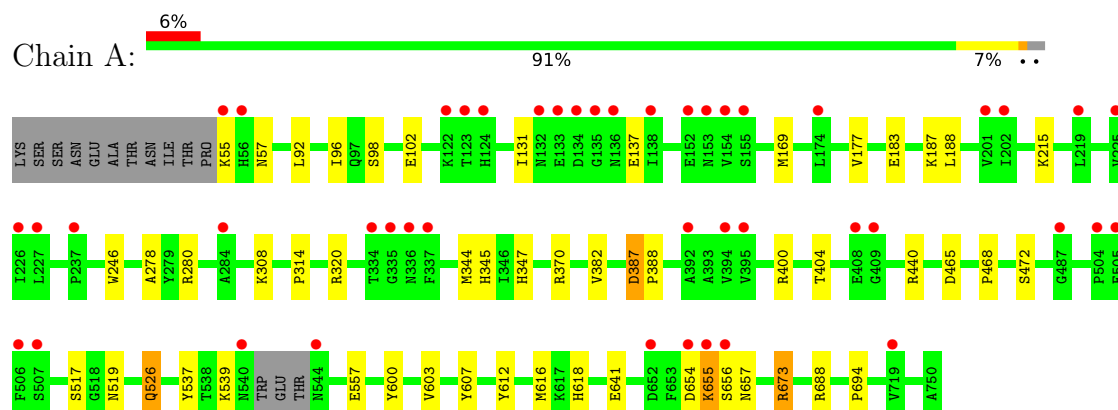
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	579	Total	O	0	4
			580	580		

### 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, $\alpha$ , $\beta$ , $\gamma$	101.70Å 130.69Å 159.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.52 29.57 – 1.52	Depositor EDS
% Data completeness (in resolution range)	99.5 (30.00-1.52) 99.5 (29.57-1.52)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.63 (at 1.52Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.158 , 0.179 0.162 , 0.181	Depositor DCC
$R_{free}$ test set	4964 reflections (3.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.6	Xtriage
Anisotropy	0.553	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 48.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	6638	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.30% 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, D6E, 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	0.91	2/6072 (0.0%)	0.90	6/8219 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	183	GLU	CD-OE1	5.32	1.31	1.25
1	A	557	GLU	CD-OE2	-5.04	1.20	1.25

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	673	ARG	NE-CZ-NH1	9.13	124.86	120.30
1	A	673	ARG	NE-CZ-NH2	-7.83	116.39	120.30
1	A	387	ASP	CB-CG-OD1	-7.04	111.96	118.30
1	A	465	ASP	CB-CG-OD1	6.44	124.09	118.30
1	A	370	ARG	NE-CZ-NH1	6.21	123.41	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	5824	0	5680	36	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	154	0	137	1	0
3	A	22	0	17	0	0
4	A	33	0	30	1	0
5	A	2	0	0	0	0
6	A	1	0	0	0	0
7	A	1	0	0	0	0
8	A	21	0	0	0	0
9	A	580	0	0	10	0
All	All	6638	0	5864	37	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.

The worst 5 of 37 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:654[B]:ASP:O	1:A:655[B]:LYS:O	1.89	0.90
1:A:654[B]:ASP:HB2	9:A:1208:HOH:O	1.79	0.82
1:A:655[B]:LYS:O	1:A:657[B]:ASN:N	2.12	0.81
1:A:603[B]:VAL:HG13	1:A:607:TYR:CE2	2.22	0.75
1:A:472[A]:SER:OG	9:A:902:HOH:O	2.05	0.74

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	739/707 (104%)	719 (97%)	15 (2%)	5 (1%)	24 6

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	655[A]	LYS
1	A	655[B]	LYS
1	A	656[A]	SER
1	A	656[B]	SER
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	640/603 (106%)	630 (98%)	10 (2%)	65 37

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	526[A]	GLN
1	A	526[B]	GLN
1	A	539	LYS
1	A	519	ASN
1	A	537	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	303	GLN
1	A	651	GLN
1	A	347	HIS
1	A	216	ASN
1	A	345	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 21 ligands modelled in this entry, 4 are monoatomic - leaving 17 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	0.48	0	17,19,21	1.11	2 (11%)
2	NAG	A	802	2	14,14,15	0.55	0	17,19,21	1.14	2 (11%)
2	NAG	A	803	1	14,14,15	0.60	0	17,19,21	1.47	3 (17%)
2	NAG	A	804	1,2	14,14,15	0.67	0	17,19,21	1.10	1 (5%)
2	NAG	A	805	2	14,14,15	0.63	0	17,19,21	0.97	1 (5%)
2	NAG	A	806	1	14,14,15	0.67	0	17,19,21	1.91	4 (23%)
2	NAG	A	807	1,2	14,14,15	0.62	0	17,19,21	1.07	2 (11%)
2	NAG	A	808	3,2	14,14,15	0.35	0	17,19,21	1.08	1 (5%)
3	BMA	A	809	2,4	11,11,12	0.71	0	15,15,17	1.18	1 (6%)
4	MAN	A	810	3	11,11,12	0.76	0	15,15,17	0.98	1 (6%)
2	NAG	A	811	1,2	14,14,15	0.53	0	17,19,21	1.21	2 (11%)
2	NAG	A	812	3,2	14,14,15	0.52	0	17,19,21	1.30	3 (17%)
3	BMA	A	813	2,4	11,11,12	0.41	0	15,15,17	0.91	1 (6%)
4	MAN	A	814	3	11,11,12	0.61	0	15,15,17	1.13	2 (13%)
4	MAN	A	815	3	11,11,12	0.63	0	15,15,17	1.13	1 (6%)
2	NAG	A	820	1	14,14,15	0.96	1 (7%)	17,19,21	1.82	4 (23%)
8	D6E	A	821	-	11,20,20	1.24	2 (18%)	13,26,26	1.39	3 (23%)

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	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	808	3,2	-	0/6/23/26	0/1/1/1
3	BMA	A	809	2,4	-	0/2/19/22	0/1/1/1
4	MAN	A	810	3	-	0/2/19/22	0/1/1/1
2	NAG	A	811	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	812	3,2	-	0/6/23/26	0/1/1/1
3	BMA	A	813	2,4	-	0/2/19/22	0/1/1/1
4	MAN	A	814	3	-	0/2/19/22	0/1/1/1
4	MAN	A	815	3	-	0/2/19/22	0/1/1/1
2	NAG	A	820	1	-	0/6/23/26	0/1/1/1
8	D6E	A	821	-	-	0/15/25/25	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	820	NAG	C1-C2	2.06	1.55	1.52
8	A	821	D6E	CAL-CAU	2.33	1.56	1.53
8	A	821	D6E	CB-CA	2.45	1.56	1.53

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	803	NAG	C1-C2-N2	-3.14	105.12	110.49
2	A	820	NAG	O3-C3-C4	-3.04	103.28	110.34
2	A	820	NAG	O5-C1-C2	-2.69	107.13	111.36
8	A	821	D6E	C-CA-N	-2.67	106.61	112.39
3	A	813	BMA	O3-C3-C2	-2.62	105.17	110.02

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	815	MAN	1	0
2	A	820	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	693/707 (98%)	0.11	45 (6%)	19 20	19, 29, 46, 69	0

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	155[A]	SER	5.4
1	A	55	LYS	5.2
1	A	153	ASN	4.5
1	A	506	PHE	4.4
1	A	335	GLY	4.3

### 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](#)

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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
2	NAG	A	802	14/15	0.69	0.27	43,54,58,64	0
2	NAG	A	803	14/15	0.73	0.28	47,56,63,67	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAG	A	806	14/15	0.79	0.24	64,72,78,79	0
4	MAN	A	810	11/12	0.80	0.31	48,55,57,60	0
2	NAG	A	805	14/15	0.81	0.28	52,59,67,69	0
2	NAG	A	820	14/15	0.82	0.21	33,50,56,60	0
4	MAN	A	815	11/12	0.84	0.32	56,63,67,70	0
3	BMA	A	809	11/12	0.86	0.33	50,52,58,67	0
2	NAG	A	804	14/15	0.87	0.17	38,43,49,51	0
2	NAG	A	808	14/15	0.88	0.16	34,43,52,61	0
2	NAG	A	811	14/15	0.88	0.10	26,33,41,49	0
2	NAG	A	812	14/15	0.89	0.22	38,41,52,60	0
3	BMA	A	813	11/12	0.91	0.18	40,42,45,47	0
4	MAN	A	814	11/12	0.93	0.19	43,45,48,48	0
2	NAG	A	807	14/15	0.93	0.10	28,32,37,44	0
2	NAG	A	801	14/15	0.94	0.11	34,41,48,53	0
8	D6E	A	821	21/21	0.97	0.07	22,24,27,29	0
7	CL	A	819	1/1	0.99	0.08	25,25,25,25	0
5	ZN	A	816	1/1	1.00	0.06	23,23,23,23	0
6	CA	A	818	1/1	1.00	0.07	20,20,20,20	0
5	ZN	A	817	1/1	1.00	0.06	22,22,22,22	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.