



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 9, 2020 – 03:01 PM BST

PDB ID : 4UY2  
Title : Crystal structure of the complex of the extracellular domain of human alpha9 nAChR with alpha-bungarotoxin.  
Authors : Giastas, P.; Zouridakis, M.; Zarkadas, E.; Tzartos, S.J.  
Deposited on : 2014-08-28  
Resolution : 2.70 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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

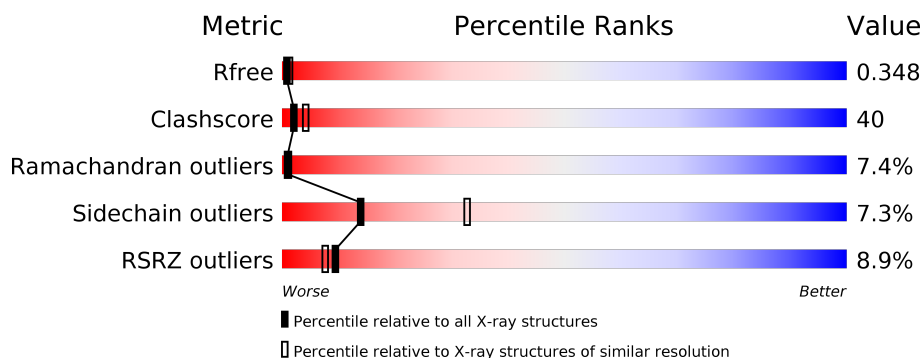
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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))
$R_{free}$	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 respectively. 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	218	<div> <div>9%</div> <div>47% 39% 5% • 7%</div> </div>
1	B	218	<div> <div>9%</div> <div>46% 36% 8% • 7%</div> </div>
2	C	74	<div> <div>8%</div> <div>43% 38% 15% • •</div> </div>
2	D	74	<div> <div>5%</div> <div>47% 35% 14% • •</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4468 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 NEURONAL ACETYLCHOLINE RECEPTOR SUBUNIT ALPHA-9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	203	Total	C	N	O	S	0	0	0
			1658	1061	270	321	6			
1	B	203	Total	C	N	O	S	0	0	0
			1658	1061	270	321	6			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	213	HIS	-	expression tag	UNP Q9UGM1
A	214	HIS	-	expression tag	UNP Q9UGM1
A	215	HIS	-	expression tag	UNP Q9UGM1
A	216	HIS	-	expression tag	UNP Q9UGM1
A	217	HIS	-	expression tag	UNP Q9UGM1
A	218	HIS	-	expression tag	UNP Q9UGM1
B	213	HIS	-	expression tag	UNP Q9UGM1
B	214	HIS	-	expression tag	UNP Q9UGM1
B	215	HIS	-	expression tag	UNP Q9UGM1
B	216	HIS	-	expression tag	UNP Q9UGM1
B	217	HIS	-	expression tag	UNP Q9UGM1
B	218	HIS	-	expression tag	UNP Q9UGM1

- Molecule 2 is a protein called ALPHA-BUNGAROTOXIN ISOFORM V31.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	73	Total	C	N	O	S	0	0	0
			548	338	96	103	11			
2	D	73	Total	C	N	O	S	0	0	0
			548	338	96	103	11			

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



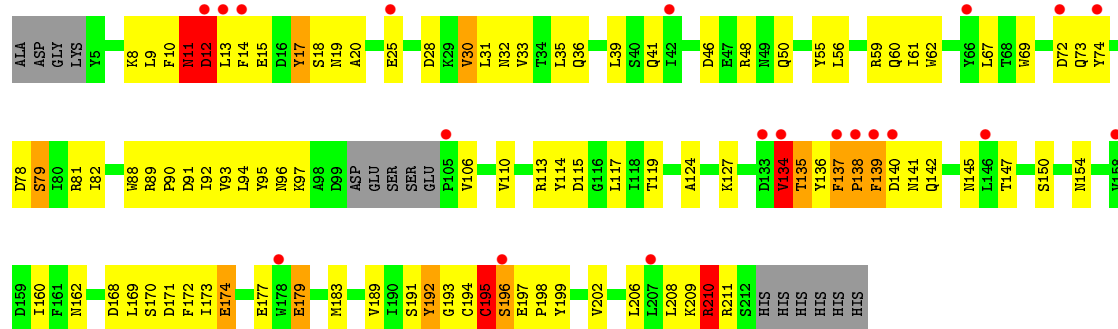
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		

### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

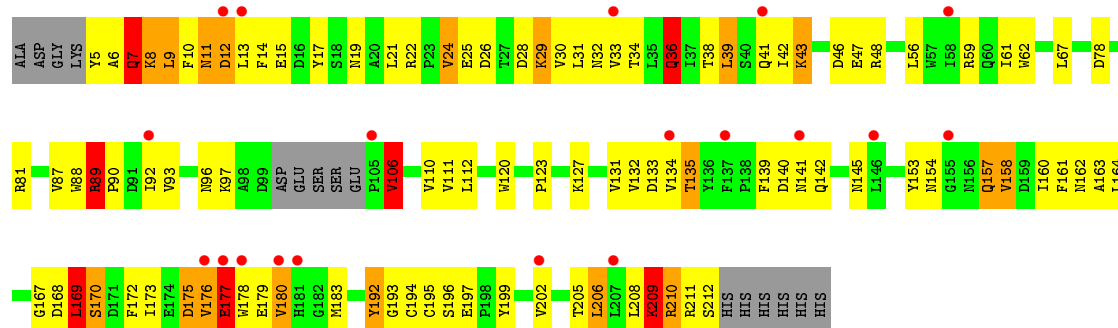
#### • Molecule 1: NEURONAL ACETYLCHOLINE RECEPTOR SUBUNIT ALPHA-9

Chain A: 



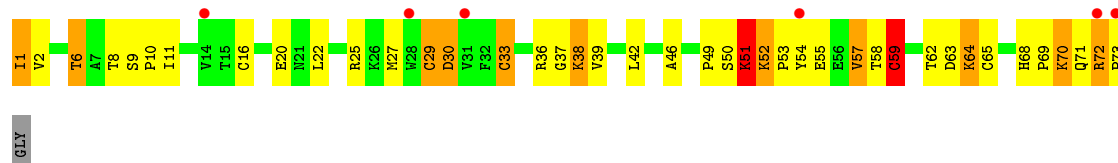
#### • Molecule 1: NEURONAL ACETYLCHOLINE RECEPTOR SUBUNIT ALPHA-9

Chain B: 

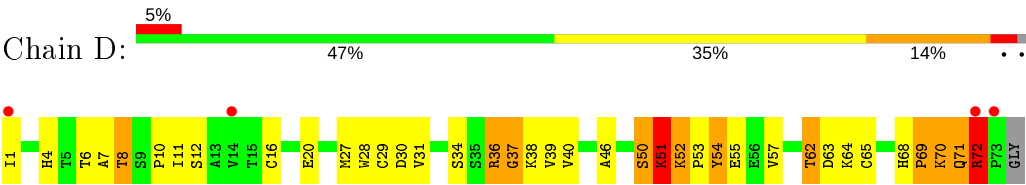


#### • Molecule 2: ALPHA-BUNGAROTOXIN ISOFORM V31

Chain C: 



● Molecule 2: ALPHA-BUNGAROTOXIN ISOFORM V31



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	42.32Å 119.35Å 79.11Å 90.00° 103.02° 90.00°	Depositor
Resolution (Å)	47.19 – 2.70 47.18 – 2.70	Depositor EDS
% Data completeness (in resolution range)	97.8 (47.19-2.70) 99.5 (47.18-2.70)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.44 (at 2.69Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.258 , 0.331 0.279 , 0.348	Depositor DCC
$R_{free}$ test set	1043 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	67.8	Xtriage
Anisotropy	0.727	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 90.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4468	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	110.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 41.59 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.3104e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG

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	0.75	0/1700	1.13	8/2323 (0.3%)
1	B	0.75	1/1700 (0.1%)	1.13	13/2323 (0.6%)
2	C	0.90	3/562 (0.5%)	1.17	4/765 (0.5%)
2	D	0.84	0/562	1.29	5/765 (0.7%)
All	All	0.78	4/4524 (0.1%)	1.16	30/6176 (0.5%)

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	4
1	B	0	6
2	C	0	3
2	D	0	2
All	All	0	15

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	51	LYS	CG-CD	-5.63	1.33	1.52
2	C	57	VAL	CB-CG1	5.63	1.64	1.52
1	B	180	VAL	CB-CG1	5.31	1.64	1.52
2	C	57	VAL	CA-CB	-5.18	1.43	1.54

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	210	ARG	NE-CZ-NH1	11.18	125.89	120.30

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	89	ARG	NE-CZ-NH1	10.54	125.57	120.30
2	D	51	LYS	CD-CE-NZ	10.50	135.85	111.70
1	A	137	PHE	C-N-CD	-9.54	99.61	120.60
1	B	164	LEU	CA-CB-CG	8.81	135.55	115.30

There are no chirality outliers.

5 of 15 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	11	ASN	Peptide
1	A	134	VAL	Peptide
1	A	135	THR	Peptide
1	A	195	CYS	Peptide
1	B	7	GLN	Peptide

## 5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1658	0	1577	118	0
1	B	1658	0	1577	149	0
2	C	548	0	533	51	0
2	D	548	0	533	52	0
3	A	28	0	26	1	0
3	B	28	0	26	2	0
All	All	4468	0	4272	352	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 40.

The worst 5 of 352 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:B:142:GLN:H	1:B:209:LYS:HE3	1.17	1.04
1:A:210:ARG:HD3	1:A:211:ARG:N	1.77	0.99

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:177:GLU:HA	1:A:210:ARG:HE	1.28	0.97
1:A:179:GLU:OE2	1:A:210:ARG:NH2	1.96	0.96
1:A:210:ARG:HH11	1:A:211:ARG:H	1.07	0.95

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	199/218 (91%)	166 (83%)	19 (10%)	14 (7%)	1	1
1	B	199/218 (91%)	168 (84%)	19 (10%)	12 (6%)	1	2
2	C	71/74 (96%)	56 (79%)	9 (13%)	6 (8%)	1	1
2	D	71/74 (96%)	58 (82%)	5 (7%)	8 (11%)	0	0
All	All	540/584 (92%)	448 (83%)	52 (10%)	40 (7%)	1	1

5 of 40 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	19	ASN
1	A	97	LYS
1	A	106	VAL
1	A	138	PRO
1	A	170	SER

### 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	186/200 (93%)	177 (95%)	9 (5%)	25	53
1	B	186/200 (93%)	171 (92%)	15 (8%)	11	27
2	C	66/66 (100%)	58 (88%)	8 (12%)	5	11
2	D	66/66 (100%)	61 (92%)	5 (8%)	13	30
All	All	504/532 (95%)	467 (93%)	37 (7%)	14	33

5 of 37 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	106	VAL
1	B	158	VAL
2	D	51	LYS
1	B	127	LYS
1	B	135	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	36	GLN
2	D	21	ASN
1	B	50	GLN
1	A	50	GLN
1	B	141	ASN

### 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 ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

4 ligands 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	1215	1	14,14,15	0.94	1 (7%)	17,19,21	0.86	0
3	NAG	B	1215	1	14,14,15	0.55	0	17,19,21	0.83	1 (5%)
3	NAG	A	1214	1	14,14,15	0.49	0	17,19,21	1.15	1 (5%)
3	NAG	B	1214	1	14,14,15	0.86	1 (7%)	17,19,21	1.19	1 (5%)

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	1215	1	-	2/6/23/26	0/1/1/1
3	NAG	B	1215	1	-	2/6/23/26	0/1/1/1
3	NAG	A	1214	1	-	2/6/23/26	0/1/1/1
3	NAG	B	1214	1	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1215	NAG	O5-C1	-3.28	1.38	1.43
3	B	1214	NAG	O5-C1	3.00	1.48	1.43

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1214	NAG	C1-O5-C5	4.07	117.70	112.19
3	B	1214	NAG	C1-O5-C5	3.13	116.43	112.19
3	B	1215	NAG	O5-C1-C2	-2.13	107.93	111.29

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1215	NAG	O5-C5-C6-O6
3	B	1215	NAG	O5-C5-C6-O6
3	A	1215	NAG	C4-C5-C6-O6
3	B	1215	NAG	C4-C5-C6-O6
3	A	1214	NAG	C1-C2-N2-C7

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1214	NAG	1	0
3	B	1214	NAG	2	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	203/218 (93%)	0.85	20 (9%) 7 5	47, 97, 173, 220	0
1	B	203/218 (93%)	0.81	19 (9%) 8 6	60, 107, 171, 232	0
2	C	73/74 (98%)	0.62	6 (8%) 11 9	68, 97, 174, 192	0
2	D	73/74 (98%)	0.52	4 (5%) 25 24	76, 103, 170, 198	0
All	All	552/584 (94%)	0.76	49 (8%) 9 7	47, 101, 177, 232	0

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	73	PRO	7.9
1	A	13	LEU	5.9
1	B	12	ASP	5.7
1	A	139	PHE	5.3
1	B	178	TRP	5.3

### 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 monosaccharides 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( $\text{\AA}^2$ )	Q<0.9
3	NAG	B	1214	14/15	0.85	0.17	106,123,150,151	0
3	NAG	A	1215	14/15	0.89	0.18	104,114,129,149	0
3	NAG	B	1215	14/15	0.90	0.17	99,114,124,133	0
3	NAG	A	1214	14/15	0.91	0.14	75,105,115,123	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.