



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 8, 2020 – 04:37 AM BST

PDB ID : 2UZ6  
Title : AChBP-targeted  $\alpha$ -conotoxin correlates distinct binding orientations with nAChR subtype selectivity.  
Authors : Ulens, C.; Dutertre, S.; Buttner, R.; Fish, A.; van Elk, R.; Kendel, Y.; Hop-ping, G.; Alewood, P.F.; Schroeder, C.; Nicke, A.; Smit, A.B.; Sixma, T.K.; Lewis, R.J.  
Deposited on : 2007-04-25  
Resolution : 2.40 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

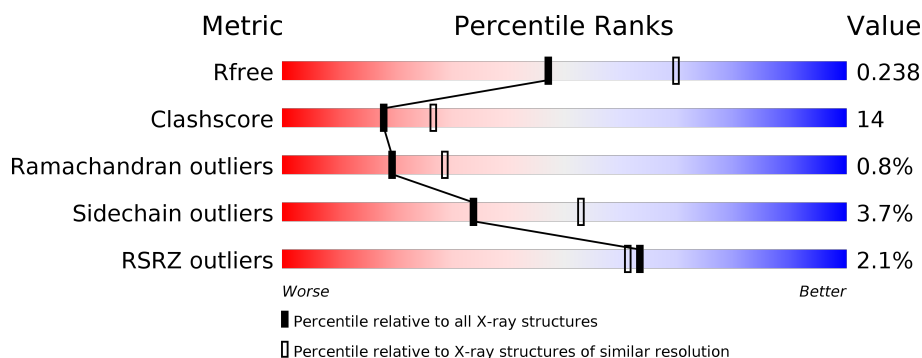
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	217	<div> <div>67%</div> <div>23%</div> <div>5%</div> <div>6%</div> </div>
1	B	217	<div> <div>74%</div> <div>19%</div> <div>6%</div> </div>
1	C	217	<div> <div>70%</div> <div>22%</div> <div>6%</div> </div>
1	D	217	<div> <div>70%</div> <div>22%</div> <div>6%</div> </div>
1	E	217	<div> <div>76%</div> <div>18%</div> <div>6%</div> </div>
1	F	217	<div> <div>71%</div> <div>21%</div> <div>6%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	217	
1	H	217	
1	I	217	
1	J	217	
2	K	17	
2	L	17	
2	M	17	
2	N	17	
2	O	17	
2	P	17	
2	Q	17	
2	R	17	
2	S	17	
2	T	17	

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	NAG	D	1206	X	-	-	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 18157 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 SOLUBLE ACETYLCHOLINE RECEPTOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	205	Total	C	N	O	S	0	0	0
			1636	1036	266	325	9			
1	B	205	Total	C	N	O	S	0	0	0
			1636	1036	266	325	9			
1	C	205	Total	C	N	O	S	0	0	0
			1636	1036	266	325	9			
1	D	205	Total	C	N	O	S	0	0	0
			1636	1036	266	325	9			
1	E	205	Total	C	N	O	S	0	0	0
			1636	1036	266	325	9			
1	F	205	Total	C	N	O	S	0	0	0
			1636	1036	266	325	9			
1	G	205	Total	C	N	O	S	0	0	0
			1636	1036	266	325	9			
1	H	205	Total	C	N	O	S	0	0	0
			1636	1036	266	325	9			
1	I	205	Total	C	N	O	S	0	0	0
			1636	1036	266	325	9			
1	J	205	Total	C	N	O	S	0	0	0
			1636	1036	266	325	9			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	41	VAL	ALA	conflict	UNP Q8WSF8
A	136	VAL	ALA	conflict	UNP Q8WSF8
B	41	VAL	ALA	conflict	UNP Q8WSF8
B	136	VAL	ALA	conflict	UNP Q8WSF8
C	41	VAL	ALA	conflict	UNP Q8WSF8
C	136	VAL	ALA	conflict	UNP Q8WSF8
D	41	VAL	ALA	conflict	UNP Q8WSF8
D	136	VAL	ALA	conflict	UNP Q8WSF8
E	41	VAL	ALA	conflict	UNP Q8WSF8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	136	VAL	ALA	conflict	UNP Q8WSF8
F	41	VAL	ALA	conflict	UNP Q8WSF8
F	136	VAL	ALA	conflict	UNP Q8WSF8
G	41	VAL	ALA	conflict	UNP Q8WSF8
G	136	VAL	ALA	conflict	UNP Q8WSF8
H	41	VAL	ALA	conflict	UNP Q8WSF8
H	136	VAL	ALA	conflict	UNP Q8WSF8
I	41	VAL	ALA	conflict	UNP Q8WSF8
I	136	VAL	ALA	conflict	UNP Q8WSF8
J	41	VAL	ALA	conflict	UNP Q8WSF8
J	136	VAL	ALA	conflict	UNP Q8WSF8

- Molecule 2 is a protein called ALPHA-CONOTOXIN TXIA(A10L).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	K	17	Total 115	C 68	N 22	O 21	S 4	0	0	1
2	L	17	Total 115	C 68	N 22	O 21	S 4	0	0	1
2	M	17	Total 115	C 68	N 22	O 21	S 4	0	0	1
2	N	17	Total 115	C 68	N 22	O 21	S 4	0	0	1
2	O	17	Total 115	C 68	N 22	O 21	S 4	0	0	1
2	P	17	Total 115	C 68	N 22	O 21	S 4	0	0	1
2	Q	17	Total 115	C 68	N 22	O 21	S 4	0	0	1
2	R	17	Total 115	C 68	N 22	O 21	S 4	0	0	1
2	S	17	Total 115	C 68	N 22	O 21	S 4	0	0	1
2	T	17	Total 115	C 68	N 22	O 21	S 4	0	0	1

- 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	D	1	Total	C	N	O	0	0
			14	8	1	5		
3	H	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	E	1	Total	C	O	0	0
			6	3	3		
4	G	1	Total	C	O	0	0
			6	3	3		

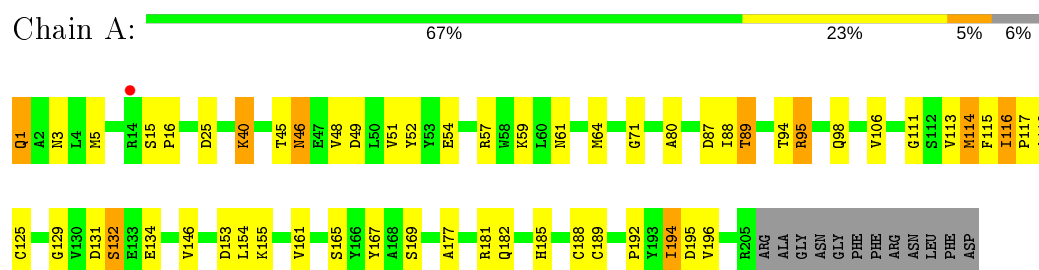
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	64	Total O 64 64	0	0
5	B	54	Total O 54 54	0	0
5	C	67	Total O 67 67	0	0
5	D	52	Total O 52 52	0	0
5	E	55	Total O 55 55	0	0
5	F	60	Total O 60 60	0	0
5	G	42	Total O 42 42	0	0
5	H	63	Total O 63 63	0	0
5	I	63	Total O 63 63	0	0
5	J	57	Total O 57 57	0	0
5	K	2	Total O 2 2	0	0
5	L	3	Total O 3 3	0	0
5	M	3	Total O 3 3	0	0
5	N	1	Total O 1 1	0	0
5	O	3	Total O 3 3	0	0
5	P	3	Total O 3 3	0	0
5	Q	1	Total O 1 1	0	0
5	R	7	Total O 7 7	0	0
5	S	6	Total O 6 6	0	0
5	T	1	Total O 1 1	0	0

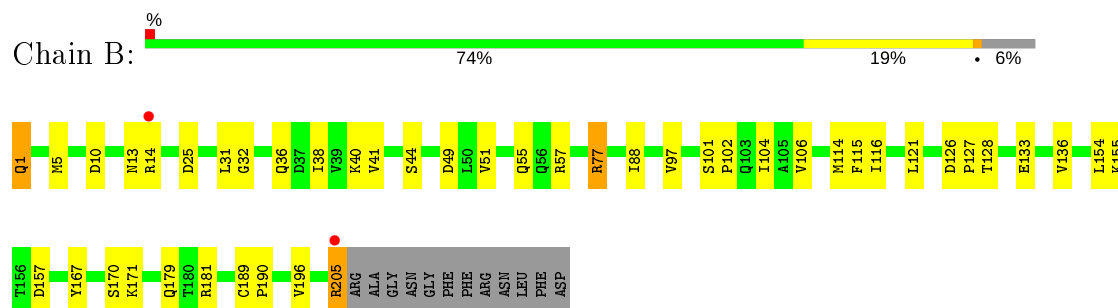
### 3 Residue-property plots [i](#)

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.

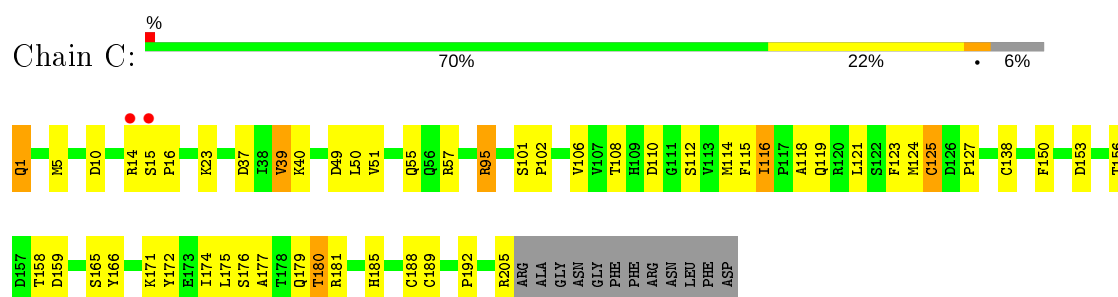
#### • Molecule 1: SOLUBLE ACETYLCHOLINE RECEPTOR



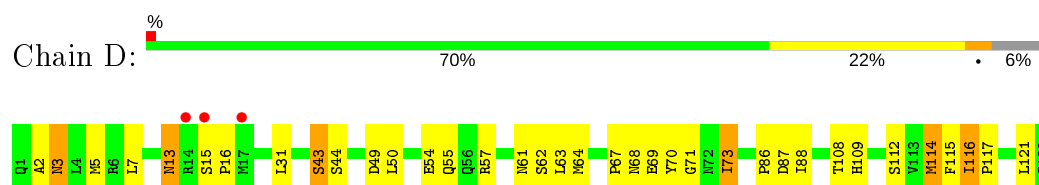
#### • Molecule 1: SOLUBLE ACETYLCHOLINE RECEPTOR



#### • Molecule 1: SOLUBLE ACETYLCHOLINE RECEPTOR



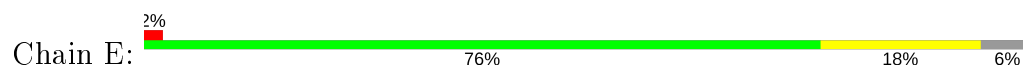
#### • Molecule 1: SOLUBLE ACETYLCHOLINE RECEPTOR



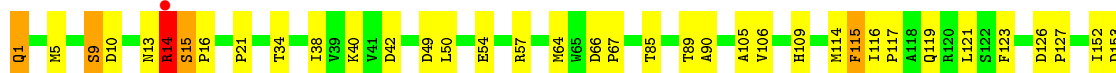




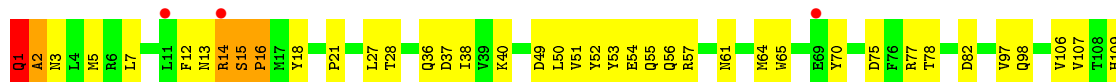
• Molecule 1: SOLUBLE ACETYLCHOLINE RECEPTOR



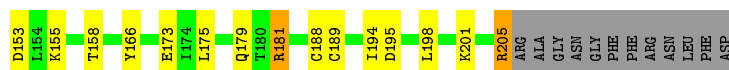
• Molecule 1: SOLUBLE ACETYLCHOLINE RECEPTOR



• Molecule 1: SOLUBLE ACETYLCHOLINE RECEPTOR

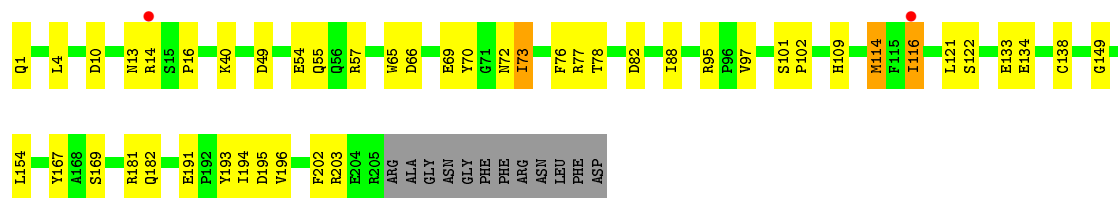


• Molecule 1: SOLUBLE ACETYLCHOLINE RECEPTOR

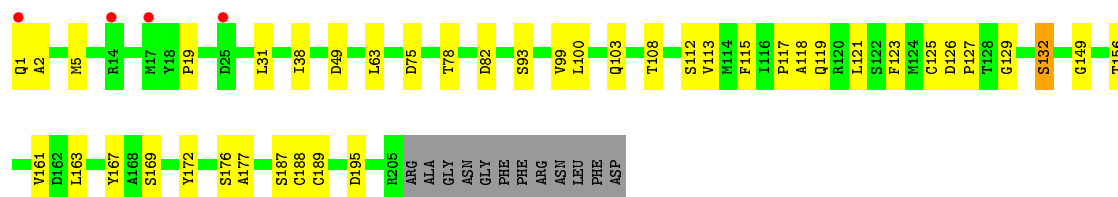
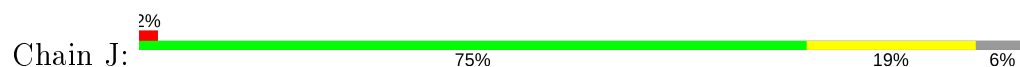


• Molecule 1: SOLUBLE ACETYLCHOLINE RECEPTOR





- Molecule 1: SOLUBLE ACETYLCHOLINE RECEPTOR



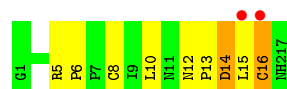
- Molecule 2: ALPHA-CONOTOXIN TXIA(A10L)



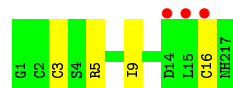
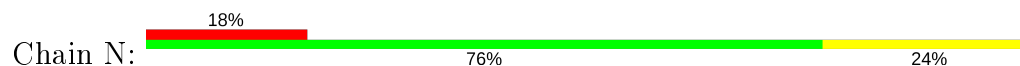
- Molecule 2: ALPHA-CONOTOXIN TXIA(A10L)



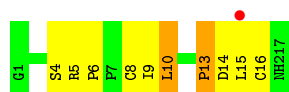
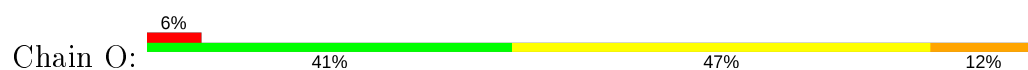
- Molecule 2: ALPHA-CONOTOXIN TXIA(A10L)



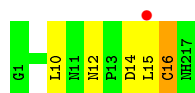
- Molecule 2: ALPHA-CONOTOXIN TXIA(A10L)



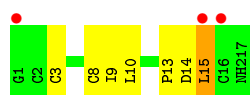
- Molecule 2: ALPHA-CONOTOXIN TXIA(A10L)



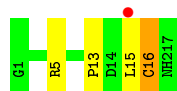
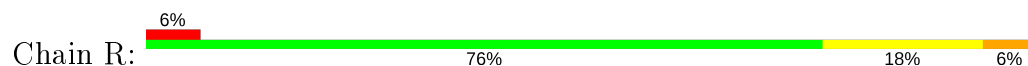
- Molecule 2: ALPHA-CONOTOXIN TXIA(A10L)



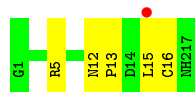
- Molecule 2: ALPHA-CONOTOXIN TXIA(A10L)



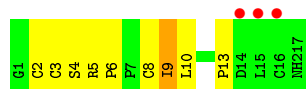
- Molecule 2: ALPHA-CONOTOXIN TXIA(A10L)



- Molecule 2: ALPHA-CONOTOXIN TXIA(A10L)



- Molecule 2: ALPHA-CONOTOXIN TXIA(A10L)



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	72.54Å 85.75Å 121.67Å 90.14° 80.01° 70.64°	Depositor
Resolution (Å)	67.42 – 2.40 67.26 – 2.40	Depositor EDS
% Data completeness (in resolution range)	95.9 (67.42-2.40) 95.8 (67.26-2.40)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.27 (at 2.40Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.236 , 0.252 0.225 , 0.238	Depositor DCC
$R_{free}$ test set	5036 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.8	Xtriage
Anisotropy	0.864	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 54.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	18157	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 4.32% 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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, NAG, NH2

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	0.54	3/1676 (0.2%)	0.78	1/2287 (0.0%)
1	B	0.46	0/1676	0.76	1/2287 (0.0%)
1	C	0.56	1/1676 (0.1%)	0.75	0/2287
1	D	0.54	2/1676 (0.1%)	0.78	0/2287
1	E	0.50	0/1676	0.75	0/2287
1	F	0.53	2/1676 (0.1%)	0.76	2/2287 (0.1%)
1	G	0.51	1/1676 (0.1%)	0.86	4/2287 (0.2%)
1	H	0.46	0/1676	0.74	0/2287
1	I	0.53	2/1676 (0.1%)	0.75	0/2287
1	J	0.49	2/1676 (0.1%)	0.76	1/2287 (0.0%)
2	K	0.39	0/116	0.79	0/158
2	L	0.44	0/116	0.70	0/158
2	M	0.84	0/116	0.87	0/158
2	N	0.40	0/116	0.59	0/158
2	O	0.38	0/116	0.66	0/158
2	P	0.40	0/116	0.80	0/158
2	Q	0.39	0/116	0.69	0/158
2	R	0.41	0/116	0.65	0/158
2	S	0.41	0/116	0.67	0/158
2	T	0.38	0/116	0.65	0/158
All	All	0.51	13/17920 (0.1%)	0.77	9/24450 (0.0%)

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	G	0	1
2	O	1	0
2	T	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
All	All	2	1

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	39	VAL	C-N	-11.12	1.08	1.34
1	I	72	ASN	C-N	-8.60	1.14	1.34
1	D	43	SER	CB-OG	6.55	1.50	1.42
1	D	44	SER	CB-OG	6.54	1.50	1.42
1	I	73	ILE	C-N	-6.18	1.19	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	1	GLN	CA-C-N	-9.92	95.37	117.20
1	G	1	GLN	O-C-N	9.81	138.39	122.70
1	F	14	ARG	N-CA-C	7.39	130.96	111.00
1	A	194	ILE	CB-CA-C	-5.61	100.38	111.60
1	F	38	ILE	N-CA-C	-5.55	96.00	111.00

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	O	9	ILE	CB
2	T	9	ILE	CB

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	G	1	GLN	Mainchain

## 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	1636	0	1572	53	0
1	B	1636	0	1572	44	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1636	0	1571	56	0
1	D	1636	0	1572	42	0
1	E	1636	0	1572	30	0
1	F	1636	0	1572	38	0
1	G	1636	0	1572	64	0
1	H	1636	0	1572	52	0
1	I	1636	0	1570	54	0
1	J	1636	0	1572	22	0
2	K	115	0	109	7	0
2	L	115	0	109	5	0
2	M	115	0	109	6	0
2	N	115	0	109	2	0
2	O	115	0	109	12	0
2	P	115	0	109	4	0
2	Q	115	0	109	7	0
2	R	115	0	109	7	0
2	S	115	0	109	3	0
2	T	115	0	109	10	0
3	D	14	0	13	0	0
3	H	14	0	13	1	0
4	E	6	0	8	0	0
4	G	6	0	8	0	0
5	A	64	0	0	1	0
5	B	54	0	0	2	0
5	C	67	0	0	2	0
5	D	52	0	0	2	0
5	E	55	0	0	2	0
5	F	60	0	0	2	0
5	G	42	0	0	2	0
5	H	63	0	0	1	0
5	I	63	0	0	0	0
5	J	57	0	0	1	0
5	K	2	0	0	0	0
5	L	3	0	0	0	0
5	M	3	0	0	1	0
5	N	1	0	0	0	0
5	O	3	0	0	1	0
5	P	3	0	0	1	0
5	Q	1	0	0	0	0
5	R	7	0	0	2	0
5	S	6	0	0	1	0
5	T	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	18157	0	16849	465	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 14.

The worst 5 of 465 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:I:55:GLN:CA	1:I:116:ILE:HG23	1.67	1.23
1:I:55:GLN:HA	1:I:116:ILE:HG23	1.17	1.14
1:F:10:ASP:O	1:F:14:ARG:HB3	1.47	1.13
1:C:10:ASP:O	1:C:14:ARG:HB2	1.48	1.12
1:H:205:ARG:HH11	1:H:205:ARG:HG3	1.03	1.11

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	203/217 (94%)	194 (96%)	7 (3%)	2 (1%)	15	23
1	B	203/217 (94%)	193 (95%)	10 (5%)	0	100	100
1	C	203/217 (94%)	192 (95%)	11 (5%)	0	100	100
1	D	203/217 (94%)	193 (95%)	7 (3%)	3 (2%)	10	14
1	E	203/217 (94%)	193 (95%)	10 (5%)	0	100	100
1	F	203/217 (94%)	195 (96%)	8 (4%)	0	100	100
1	G	203/217 (94%)	187 (92%)	11 (5%)	5 (2%)	5	6
1	H	203/217 (94%)	190 (94%)	12 (6%)	1 (0%)	29	41
1	I	203/217 (94%)	190 (94%)	13 (6%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	J	203/217 (94%)	185 (91%)	16 (8%)	2 (1%)	15	23
2	K	15/17 (88%)	12 (80%)	2 (13%)	1 (7%)	1	0
2	L	15/17 (88%)	14 (93%)	1 (7%)	0	100	100
2	M	15/17 (88%)	12 (80%)	2 (13%)	1 (7%)	1	0
2	N	15/17 (88%)	12 (80%)	3 (20%)	0	100	100
2	O	15/17 (88%)	11 (73%)	3 (20%)	1 (7%)	1	0
2	P	15/17 (88%)	13 (87%)	2 (13%)	0	100	100
2	Q	15/17 (88%)	13 (87%)	1 (7%)	1 (7%)	1	0
2	R	15/17 (88%)	12 (80%)	3 (20%)	0	100	100
2	S	15/17 (88%)	13 (87%)	2 (13%)	0	100	100
2	T	15/17 (88%)	14 (93%)	1 (7%)	0	100	100
All	All	2180/2340 (93%)	2038 (94%)	125 (6%)	17 (1%)	19	29

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	14	ARG
1	G	16	PRO
1	D	3	ASN
2	O	13	PRO
1	D	62	SER

### 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	188/197 (95%)	180 (96%)	8 (4%)	29	46
1	B	188/197 (95%)	181 (96%)	7 (4%)	34	53
1	C	188/197 (95%)	181 (96%)	7 (4%)	34	53
1	D	188/197 (95%)	181 (96%)	7 (4%)	34	53
1	E	188/197 (95%)	184 (98%)	4 (2%)	53	72

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	F	188/197 (95%)	180 (96%)	8 (4%)	29	46
1	G	188/197 (95%)	182 (97%)	6 (3%)	39	59
1	H	188/197 (95%)	181 (96%)	7 (4%)	34	53
1	I	188/197 (95%)	180 (96%)	8 (4%)	29	46
1	J	188/197 (95%)	187 (100%)	1 (0%)	88	95
2	K	15/15 (100%)	14 (93%)	1 (7%)	16	26
2	L	15/15 (100%)	15 (100%)	0	100	100
2	M	15/15 (100%)	12 (80%)	3 (20%)	1	1
2	N	15/15 (100%)	14 (93%)	1 (7%)	16	26
2	O	15/15 (100%)	12 (80%)	3 (20%)	1	1
2	P	15/15 (100%)	14 (93%)	1 (7%)	16	26
2	Q	15/15 (100%)	15 (100%)	0	100	100
2	R	15/15 (100%)	14 (93%)	1 (7%)	16	26
2	S	15/15 (100%)	14 (93%)	1 (7%)	16	26
2	T	15/15 (100%)	14 (93%)	1 (7%)	16	26
All	All	2030/2120 (96%)	1955 (96%)	75 (4%)	34	53

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	1	GLN
1	G	1	GLN
2	O	10	LEU
1	F	14	ARG
1	F	178	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	185	HIS
1	E	56	GLN
1	I	182	GLN
2	P	12	ASN
2	T	12	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	D	1206	-	14,14,15	1.68	2 (14%)	17,19,21	2.43	7 (41%)
4	GOL	G	1206	-	5,5,5	0.40	0	5,5,5	0.50	0
4	GOL	E	3147	-	5,5,5	0.36	0	5,5,5	0.24	0
3	NAG	H	1206	-	14,14,15	1.57	2 (14%)	17,19,21	2.59	5 (29%)

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	D	1206	-	1/1/5/7	4/6/23/26	0/1/1/1
4	GOL	G	1206	-	-	3/4/4/4	-
4	GOL	E	3147	-	-	2/4/4/4	-
3	NAG	H	1206	-	-	3/6/23/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	1206	NAG	O3-C3	4.62	1.53	1.43
3	H	1206	NAG	C1-C2	4.14	1.58	1.52
3	H	1206	NAG	C3-C2	3.64	1.60	1.52
3	D	1206	NAG	C1-C2	3.41	1.57	1.52

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	H	1206	NAG	C2-N2-C7	6.34	131.93	122.90
3	H	1206	NAG	O5-C1-C2	-5.04	103.33	111.29
3	D	1206	NAG	C3-C4-C5	-4.84	101.61	110.24
3	H	1206	NAG	C1-C2-N2	4.73	118.56	110.49
3	D	1206	NAG	O3-C3-C2	4.61	119.00	109.47

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	D	1206	NAG	C5

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	1206	NAG	C1-C2-N2-C7
4	G	1206	GOL	O1-C1-C2-C3
4	E	3147	GOL	O1-C1-C2-C3
3	H	1206	NAG	C4-C5-C6-O6
4	E	3147	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	H	1206	NAG	1	0

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	I	2
1	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	I	73:ILE	C	74:THR	N	1.19
1	I	72:ASN	C	73:ILE	N	1.14
1	C	39:VAL	C	40:LYS	N	1.08

## 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	205/217 (94%)	0.09	1 (0%) 91 89	15, 25, 43, 56	0
1	B	205/217 (94%)	0.07	2 (0%) 82 80	10, 26, 44, 55	0
1	C	205/217 (94%)	0.07	2 (0%) 82 80	13, 27, 48, 61	0
1	D	205/217 (94%)	0.14	3 (1%) 73 72	11, 24, 50, 59	0
1	E	205/217 (94%)	0.14	5 (2%) 59 57	11, 28, 50, 60	0
1	F	205/217 (94%)	0.11	1 (0%) 91 89	12, 25, 50, 57	0
1	G	205/217 (94%)	0.20	6 (2%) 51 50	12, 30, 52, 63	0
1	H	205/217 (94%)	0.09	2 (0%) 82 80	15, 26, 45, 57	0
1	I	205/217 (94%)	0.09	2 (0%) 82 80	11, 25, 43, 56	0
1	J	205/217 (94%)	0.09	4 (1%) 65 63	10, 27, 44, 61	0
2	K	16/17 (94%)	0.86	1 (6%) 20 18	31, 36, 45, 46	0
2	L	16/17 (94%)	1.13	2 (12%) 3 3	33, 39, 50, 51	0
2	M	16/17 (94%)	1.00	2 (12%) 3 3	27, 31, 45, 47	0
2	N	16/17 (94%)	0.88	3 (18%) 1 1	32, 35, 48, 49	0
2	O	16/17 (94%)	0.83	1 (6%) 20 18	38, 43, 51, 51	0
2	P	16/17 (94%)	0.89	1 (6%) 20 18	31, 36, 46, 47	0
2	Q	16/17 (94%)	1.17	3 (18%) 1 1	39, 44, 55, 55	0
2	R	16/17 (94%)	1.00	1 (6%) 20 18	27, 33, 44, 45	0
2	S	16/17 (94%)	1.10	1 (6%) 20 18	34, 39, 51, 52	0
2	T	16/17 (94%)	1.23	3 (18%) 1 1	27, 31, 44, 44	0
All	All	2210/2340 (94%)	0.17	46 (2%) 63 61	10, 27, 49, 63	0

The worst 5 of 46 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	M	15	LEU	7.5
2	K	15	LEU	5.2
2	R	15	LEU	5.0
2	L	15	LEU	4.8
2	S	15	LEU	4.7

## 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	H	1206	14/15	0.51	0.23	50,51,53,53	0
3	NAG	D	1206	14/15	0.79	0.17	53,54,54,54	0
4	GOL	G	1206	6/6	0.80	0.13	22,23,23,23	0
4	GOL	E	3147	6/6	0.84	0.23	24,25,25,25	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.