wwPDB X-ray Structure Validation Summary Report (i)

May 2, 2024 – 02:27 am BST

PDB ID	:	8P1F
Title	:	X-ray structure of acetylcholine-binding protein (AChBP) in complex with FL001909.
Authors	:	Cederfelt, D.; Boronat, P.; Dobritzsch, D.; Hennig, S.; Fitzgerald, E.A.; de Esch, I.J.P.; Danielson, U.H.
Deposited on	:	2023-05-12
Resolution	:	2.10 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

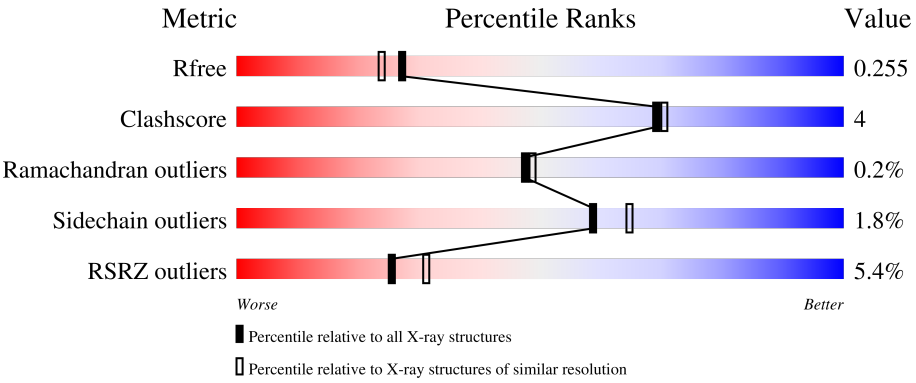
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	237	3% 76% 8% 16%
1	B	237	5% 75% 10% 15%
1	C	237	4% 77% 8% 15%
1	D	237	4% 75% 9% • 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	E	237	6% 80% 5% 15%
1	F	237	3% 73% 10% 14%
1	G	237	4% 78% 7% 14%
1	H	237	2% 81% 15%
1	I	237	% 76% 11% 14%
1	J	237	2% 74% 10% 15%
1	K	237	14% 79% 16%
1	L	237	7% 73% 11% 15%
1	M	237	6% 77% 7% 15%
1	N	237	3% 74% 10% 16%
1	O	237	4% 75% 9% 16%
1	P	237	7% 79% 5% 16%
1	Q	237	5% 74% 11% 14%
1	R	237	4% 79% 5% 15%
1	S	237	4% 78% 8% 13%
1	T	237	6% 79% 6% 14%

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
2	NAG	E	302	-	-	-	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 33433 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 Acetylcholine-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	200	Total	C	N	O	S	0	3	0
			1626	1018	280	323	5			
1	B	201	Total	C	N	O	S	0	3	0
			1632	1021	281	325	5			
1	C	202	Total	C	N	O	S	0	1	0
			1618	1016	276	321	5			
1	D	203	Total	C	N	O	S	0	0	0
			1623	1016	280	322	5			
1	E	202	Total	C	N	O	S	0	0	0
			1612	1010	276	321	5			
1	F	204	Total	C	N	O	S	0	0	0
			1628	1019	278	326	5			
1	G	204	Total	C	N	O	S	0	0	0
			1632	1021	279	327	5			
1	H	202	Total	C	N	O	S	0	0	0
			1613	1011	276	321	5			
1	I	205	Total	C	N	O	S	0	5	0
			1673	1044	289	335	5			
1	J	201	Total	C	N	O	S	0	2	0
			1629	1020	285	321	3			
1	K	200	Total	C	N	O	S	0	3	0
			1619	1018	277	319	5			
1	L	201	Total	C	N	O	S	0	0	0
			1606	1007	275	319	5			
1	M	201	Total	C	N	O	S	0	0	0
			1609	1009	275	320	5			
1	N	200	Total	C	N	O	S	0	4	0
			1633	1023	283	322	5			
1	O	200	Total	C	N	O	S	0	1	0
			1610	1010	277	318	5			
1	P	199	Total	C	N	O	S	0	1	0
			1601	1006	276	314	5			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Q	204	Total	C	N	O	S	0	0	0
			1632	1021	279	327	5			
1	R	201	Total	C	N	O	S	0	0	0
			1609	1009	275	320	5			
1	S	206	Total	C	N	O	S	0	0	0
			1647	1029	284	329	5			
1	T	203	Total	C	N	O	S	0	3	0
			1647	1029	286	327	5			

There are 160 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	230	GLY	-	expression tag	UNP P58154
A	231	SER	-	expression tag	UNP P58154
A	232	HIS	-	expression tag	UNP P58154
A	233	HIS	-	expression tag	UNP P58154
A	234	HIS	-	expression tag	UNP P58154
A	235	HIS	-	expression tag	UNP P58154
A	236	HIS	-	expression tag	UNP P58154
A	237	HIS	-	expression tag	UNP P58154
B	230	GLY	-	expression tag	UNP P58154
B	231	SER	-	expression tag	UNP P58154
B	232	HIS	-	expression tag	UNP P58154
B	233	HIS	-	expression tag	UNP P58154
B	234	HIS	-	expression tag	UNP P58154
B	235	HIS	-	expression tag	UNP P58154
B	236	HIS	-	expression tag	UNP P58154
B	237	HIS	-	expression tag	UNP P58154
C	230	GLY	-	expression tag	UNP P58154
C	231	SER	-	expression tag	UNP P58154
C	232	HIS	-	expression tag	UNP P58154
C	233	HIS	-	expression tag	UNP P58154
C	234	HIS	-	expression tag	UNP P58154
C	235	HIS	-	expression tag	UNP P58154
C	236	HIS	-	expression tag	UNP P58154
C	237	HIS	-	expression tag	UNP P58154
D	230	GLY	-	expression tag	UNP P58154
D	231	SER	-	expression tag	UNP P58154
D	232	HIS	-	expression tag	UNP P58154
D	233	HIS	-	expression tag	UNP P58154
D	234	HIS	-	expression tag	UNP P58154
D	235	HIS	-	expression tag	UNP P58154
D	236	HIS	-	expression tag	UNP P58154

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	237	HIS	-	expression tag	UNP P58154
E	230	GLY	-	expression tag	UNP P58154
E	231	SER	-	expression tag	UNP P58154
E	232	HIS	-	expression tag	UNP P58154
E	233	HIS	-	expression tag	UNP P58154
E	234	HIS	-	expression tag	UNP P58154
E	235	HIS	-	expression tag	UNP P58154
E	236	HIS	-	expression tag	UNP P58154
E	237	HIS	-	expression tag	UNP P58154
F	230	GLY	-	expression tag	UNP P58154
F	231	SER	-	expression tag	UNP P58154
F	232	HIS	-	expression tag	UNP P58154
F	233	HIS	-	expression tag	UNP P58154
F	234	HIS	-	expression tag	UNP P58154
F	235	HIS	-	expression tag	UNP P58154
F	236	HIS	-	expression tag	UNP P58154
F	237	HIS	-	expression tag	UNP P58154
G	230	GLY	-	expression tag	UNP P58154
G	231	SER	-	expression tag	UNP P58154
G	232	HIS	-	expression tag	UNP P58154
G	233	HIS	-	expression tag	UNP P58154
G	234	HIS	-	expression tag	UNP P58154
G	235	HIS	-	expression tag	UNP P58154
G	236	HIS	-	expression tag	UNP P58154
G	237	HIS	-	expression tag	UNP P58154
H	230	GLY	-	expression tag	UNP P58154
H	231	SER	-	expression tag	UNP P58154
H	232	HIS	-	expression tag	UNP P58154
H	233	HIS	-	expression tag	UNP P58154
H	234	HIS	-	expression tag	UNP P58154
H	235	HIS	-	expression tag	UNP P58154
H	236	HIS	-	expression tag	UNP P58154
H	237	HIS	-	expression tag	UNP P58154
I	230	GLY	-	expression tag	UNP P58154
I	231	SER	-	expression tag	UNP P58154
I	232	HIS	-	expression tag	UNP P58154
I	233	HIS	-	expression tag	UNP P58154
I	234	HIS	-	expression tag	UNP P58154
I	235	HIS	-	expression tag	UNP P58154
I	236	HIS	-	expression tag	UNP P58154
I	237	HIS	-	expression tag	UNP P58154
J	230	GLY	-	expression tag	UNP P58154

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
J	231	SER	-	expression tag	UNP P58154
J	232	HIS	-	expression tag	UNP P58154
J	233	HIS	-	expression tag	UNP P58154
J	234	HIS	-	expression tag	UNP P58154
J	235	HIS	-	expression tag	UNP P58154
J	236	HIS	-	expression tag	UNP P58154
J	237	HIS	-	expression tag	UNP P58154
K	230	GLY	-	expression tag	UNP P58154
K	231	SER	-	expression tag	UNP P58154
K	232	HIS	-	expression tag	UNP P58154
K	233	HIS	-	expression tag	UNP P58154
K	234	HIS	-	expression tag	UNP P58154
K	235	HIS	-	expression tag	UNP P58154
K	236	HIS	-	expression tag	UNP P58154
K	237	HIS	-	expression tag	UNP P58154
L	230	GLY	-	expression tag	UNP P58154
L	231	SER	-	expression tag	UNP P58154
L	232	HIS	-	expression tag	UNP P58154
L	233	HIS	-	expression tag	UNP P58154
L	234	HIS	-	expression tag	UNP P58154
L	235	HIS	-	expression tag	UNP P58154
L	236	HIS	-	expression tag	UNP P58154
L	237	HIS	-	expression tag	UNP P58154
M	230	GLY	-	expression tag	UNP P58154
M	231	SER	-	expression tag	UNP P58154
M	232	HIS	-	expression tag	UNP P58154
M	233	HIS	-	expression tag	UNP P58154
M	234	HIS	-	expression tag	UNP P58154
M	235	HIS	-	expression tag	UNP P58154
M	236	HIS	-	expression tag	UNP P58154
M	237	HIS	-	expression tag	UNP P58154
N	230	GLY	-	expression tag	UNP P58154
N	231	SER	-	expression tag	UNP P58154
N	232	HIS	-	expression tag	UNP P58154
N	233	HIS	-	expression tag	UNP P58154
N	234	HIS	-	expression tag	UNP P58154
N	235	HIS	-	expression tag	UNP P58154
N	236	HIS	-	expression tag	UNP P58154
N	237	HIS	-	expression tag	UNP P58154
O	230	GLY	-	expression tag	UNP P58154
O	231	SER	-	expression tag	UNP P58154
O	232	HIS	-	expression tag	UNP P58154

Continued on next page...

Continued from previous page...

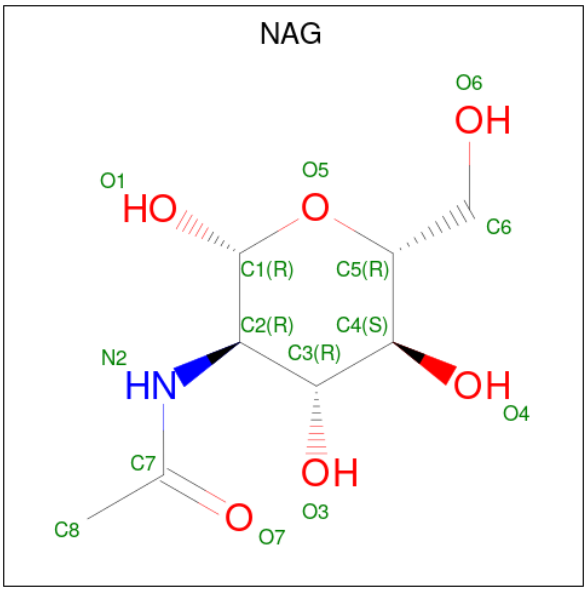
Chain	Residue	Modelled	Actual	Comment	Reference
O	233	HIS	-	expression tag	UNP P58154
O	234	HIS	-	expression tag	UNP P58154
O	235	HIS	-	expression tag	UNP P58154
O	236	HIS	-	expression tag	UNP P58154
O	237	HIS	-	expression tag	UNP P58154
P	230	GLY	-	expression tag	UNP P58154
P	231	SER	-	expression tag	UNP P58154
P	232	HIS	-	expression tag	UNP P58154
P	233	HIS	-	expression tag	UNP P58154
P	234	HIS	-	expression tag	UNP P58154
P	235	HIS	-	expression tag	UNP P58154
P	236	HIS	-	expression tag	UNP P58154
P	237	HIS	-	expression tag	UNP P58154
Q	230	GLY	-	expression tag	UNP P58154
Q	231	SER	-	expression tag	UNP P58154
Q	232	HIS	-	expression tag	UNP P58154
Q	233	HIS	-	expression tag	UNP P58154
Q	234	HIS	-	expression tag	UNP P58154
Q	235	HIS	-	expression tag	UNP P58154
Q	236	HIS	-	expression tag	UNP P58154
Q	237	HIS	-	expression tag	UNP P58154
R	230	GLY	-	expression tag	UNP P58154
R	231	SER	-	expression tag	UNP P58154
R	232	HIS	-	expression tag	UNP P58154
R	233	HIS	-	expression tag	UNP P58154
R	234	HIS	-	expression tag	UNP P58154
R	235	HIS	-	expression tag	UNP P58154
R	236	HIS	-	expression tag	UNP P58154
R	237	HIS	-	expression tag	UNP P58154
S	230	GLY	-	expression tag	UNP P58154
S	231	SER	-	expression tag	UNP P58154
S	232	HIS	-	expression tag	UNP P58154
S	233	HIS	-	expression tag	UNP P58154
S	234	HIS	-	expression tag	UNP P58154
S	235	HIS	-	expression tag	UNP P58154
S	236	HIS	-	expression tag	UNP P58154
S	237	HIS	-	expression tag	UNP P58154
T	230	GLY	-	expression tag	UNP P58154
T	231	SER	-	expression tag	UNP P58154
T	232	HIS	-	expression tag	UNP P58154
T	233	HIS	-	expression tag	UNP P58154
T	234	HIS	-	expression tag	UNP P58154

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
T	235	HIS	-	expression tag	UNP P58154
T	236	HIS	-	expression tag	UNP P58154
T	237	HIS	-	expression tag	UNP P58154

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



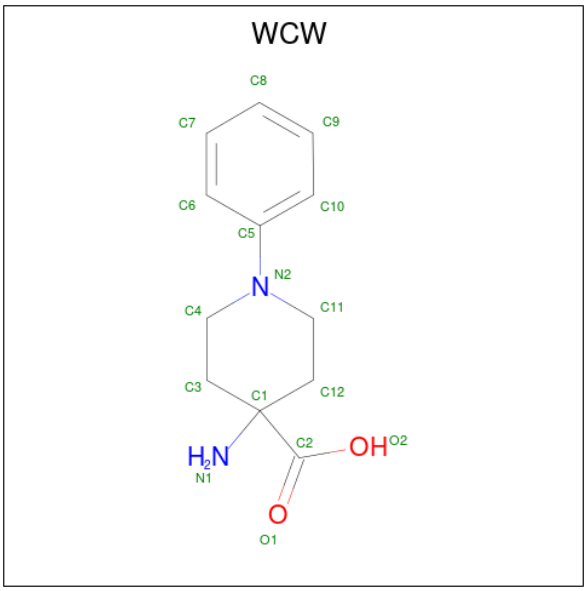
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	C	1	Total	C	N	O	0	0
			14	8	1	5		
2	E	1	Total	C	N	O	0	0
			14	8	1	5		
2	F	1	Total	C	N	O	0	0
			14	8	1	5		
2	I	1	Total	C	N	O	0	0
			14	8	1	5		
2	J	1	Total	C	N	O	0	0
			14	8	1	5		
2	K	1	Total	C	N	O	0	0
			14	8	1	5		
2	L	1	Total	C	N	O	0	0
			14	8	1	5		
2	M	1	Total	C	N	O	0	0
			14	8	1	5		
2	N	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	O	1	Total C N O 14 8 1 5	0	0
2	Q	1	Total C N O 14 8 1 5	0	0
2	R	1	Total C N O 14 8 1 5	0	0
2	T	1	Total C N O 14 8 1 5	0	0

- Molecule 3 is 4-azanyl-1-phenyl-piperidine-4-carboxylic acid (three-letter code: WCW) (formula: $C_{12}H_{16}N_2O_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	E	1	Total C N O 16 12 2 2	0	0
3	G	1	Total C N O 16 12 2 2	0	0
3	I	1	Total C N O 16 12 2 2	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	31	Total O 31 31	0	0
4	B	25	Total O 25 25	0	0

Continued on next page...

Continued from previous page...

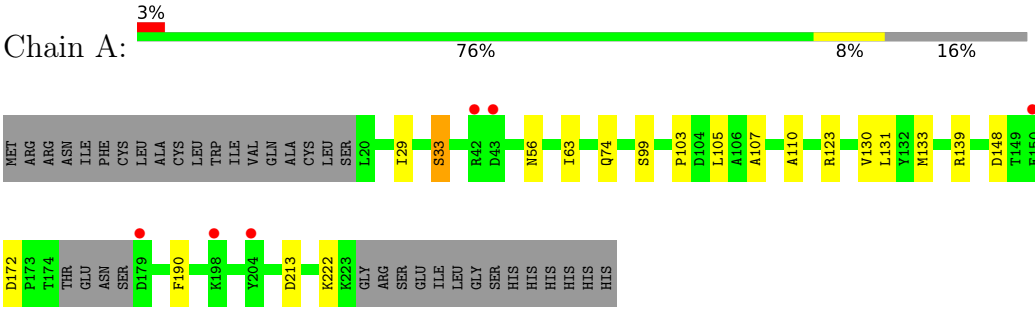
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	C	31	Total O 31 31	0	0
4	D	37	Total O 37 37	0	0
4	E	37	Total O 37 37	0	0
4	F	50	Total O 50 50	0	0
4	G	59	Total O 60 60	0	1
4	H	55	Total O 55 55	0	0
4	I	52	Total O 52 52	0	0
4	J	56	Total O 56 56	0	0
4	K	15	Total O 15 15	0	0
4	L	15	Total O 15 15	0	0
4	M	29	Total O 29 29	0	0
4	N	36	Total O 36 36	0	0
4	O	35	Total O 35 35	0	0
4	P	17	Total O 17 17	0	0
4	Q	28	Total O 28 28	0	0
4	R	37	Total O 37 37	0	0
4	S	31	Total O 31 31	0	0
4	T	27	Total O 27 27	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.

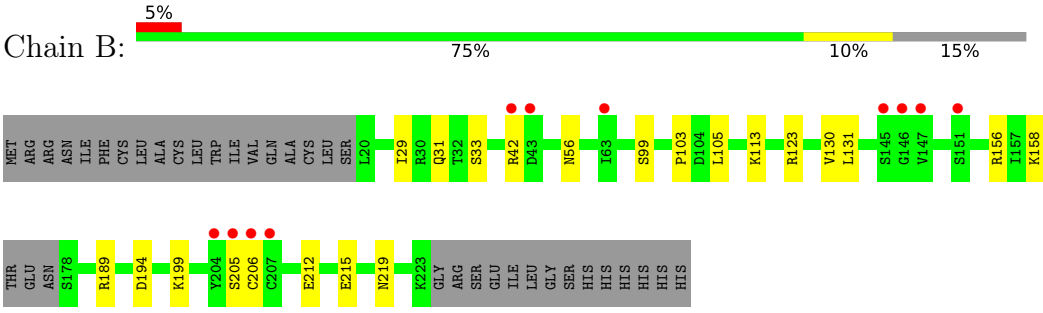
• Molecule 1: Acetylcholine-binding protein

Chain A:



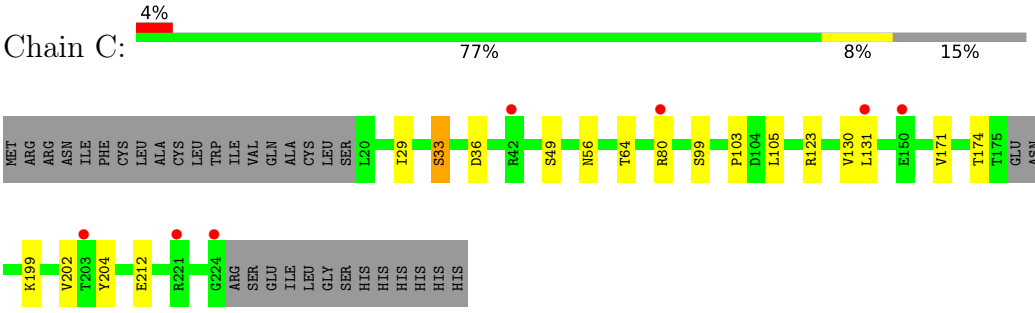
• Molecule 1: Acetylcholine-binding protein

Chain B:



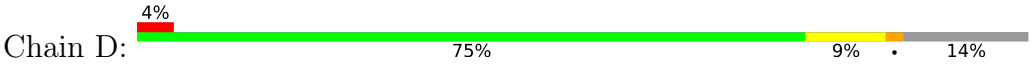
• Molecule 1: Acetylcholine-binding protein

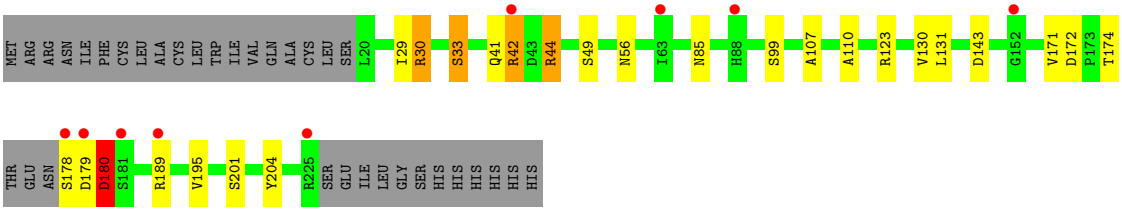
Chain C:



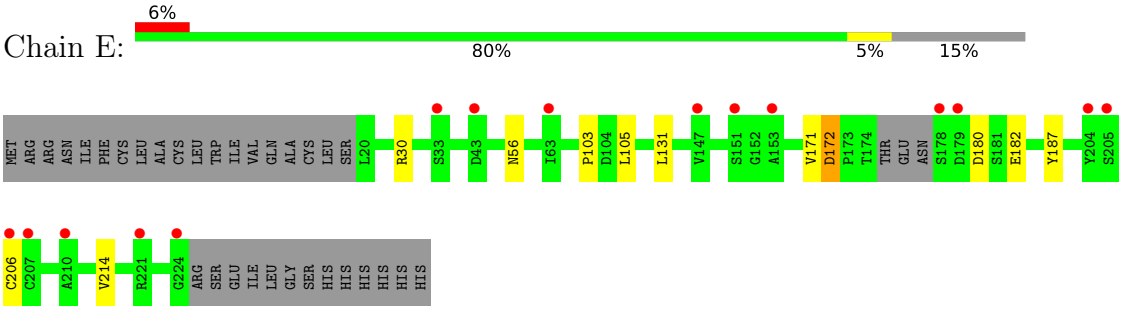
• Molecule 1: Acetylcholine-binding protein

Chain D:

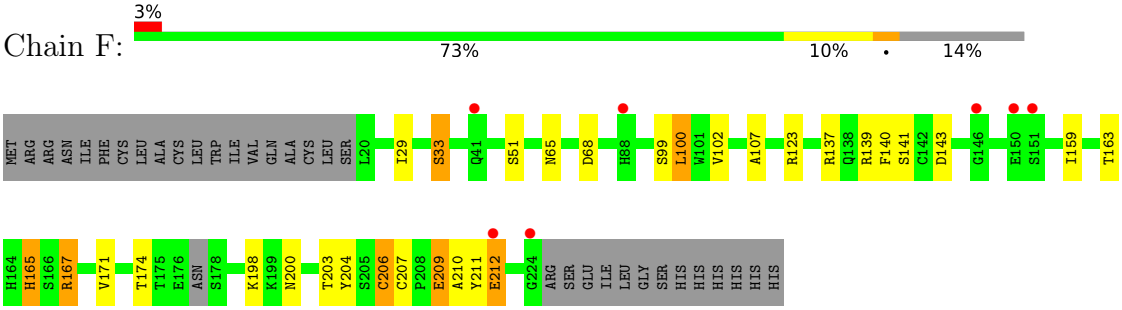




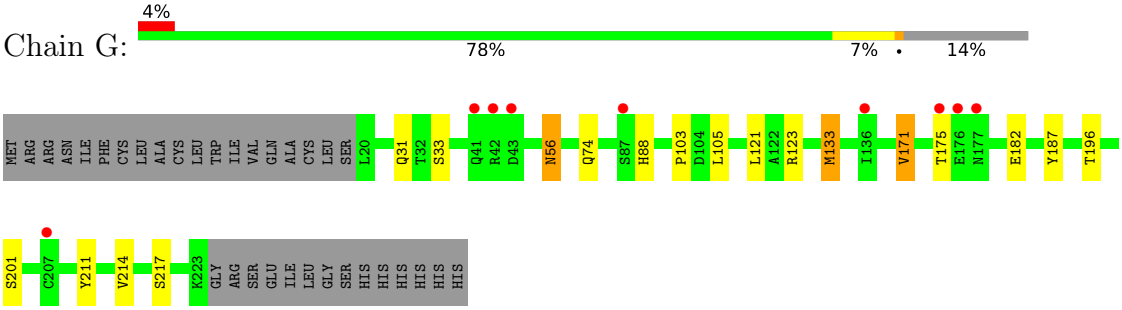
• Molecule 1: Acetylcholine-binding protein



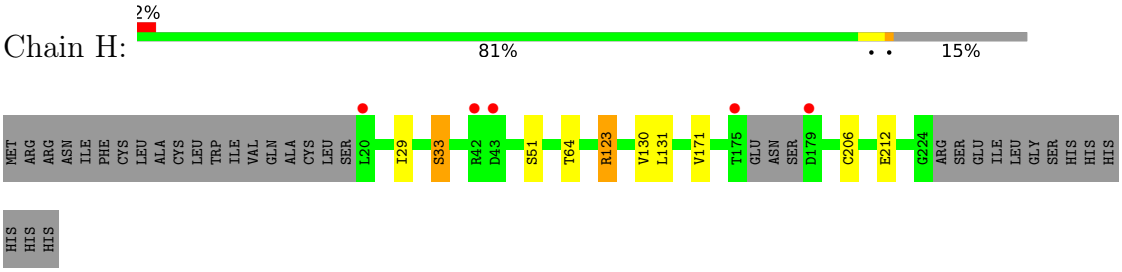
• Molecule 1: Acetylcholine-binding protein



• Molecule 1: Acetylcholine-binding protein

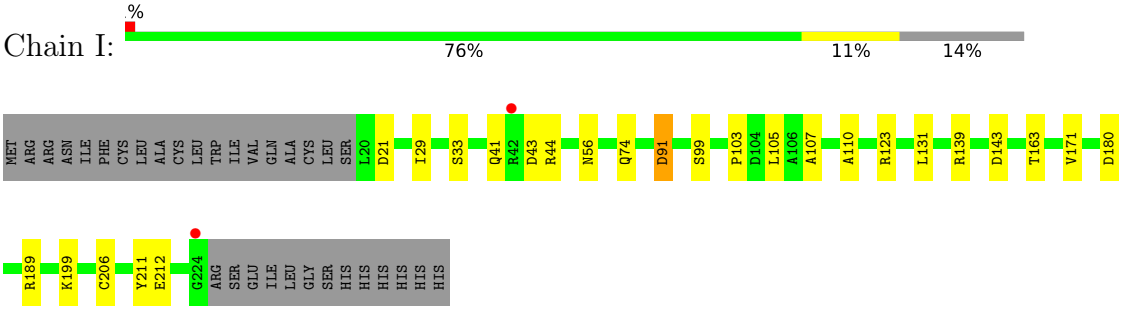


• Molecule 1: Acetylcholine-binding protein



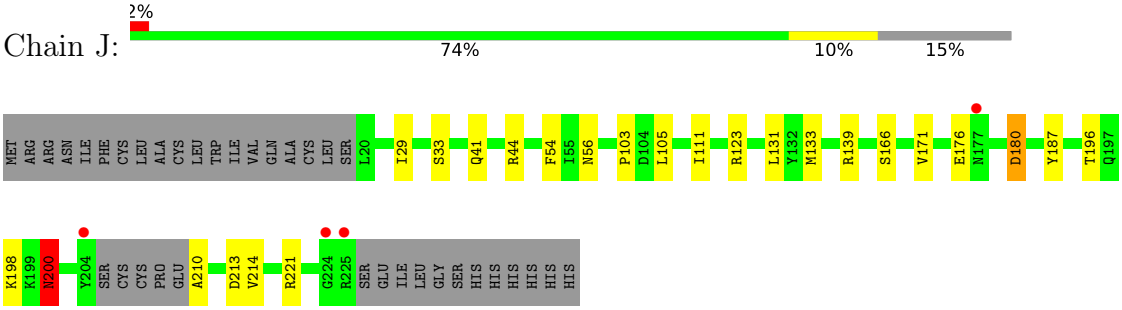
• Molecule 1: Acetylcholine-binding protein

Chain I:



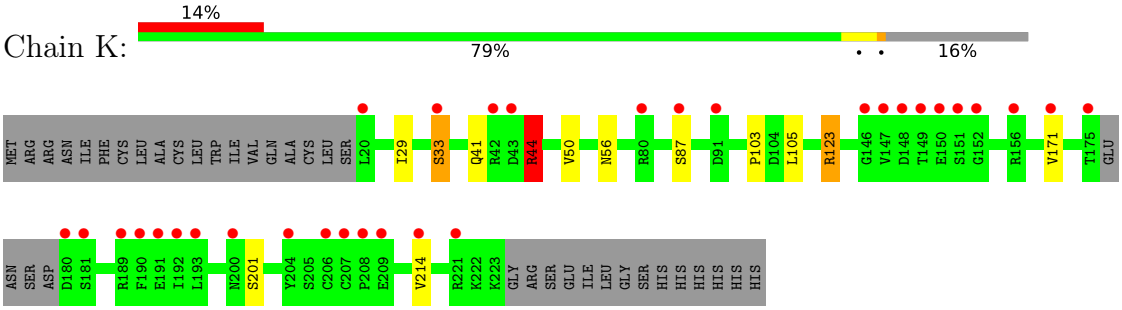
• Molecule 1: Acetylcholine-binding protein

Chain J:



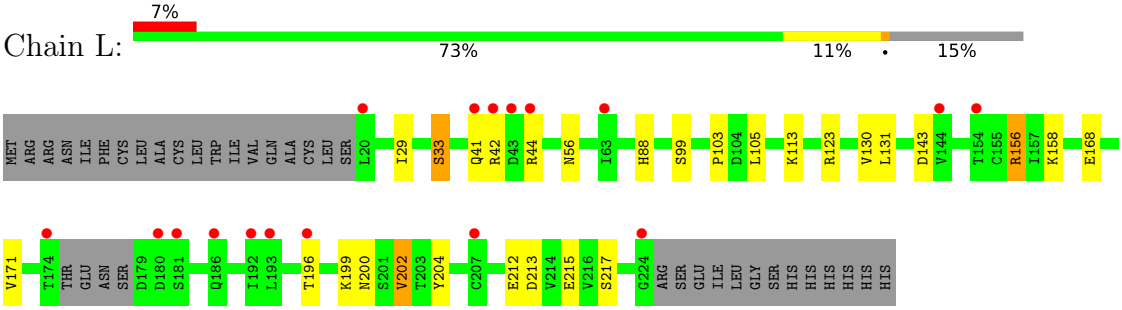
• Molecule 1: Acetylcholine-binding protein

Chain K:



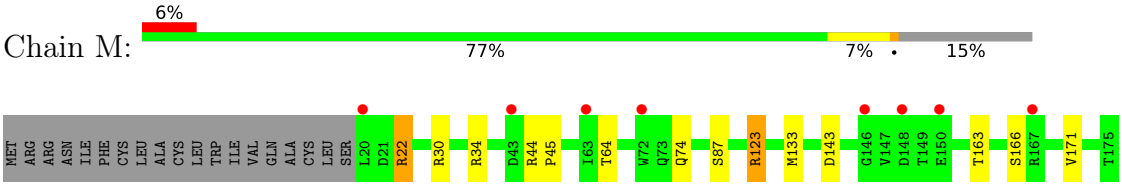
• Molecule 1: Acetylcholine-binding protein

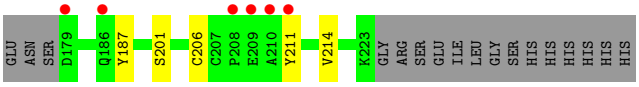
Chain L:



• Molecule 1: Acetylcholine-binding protein

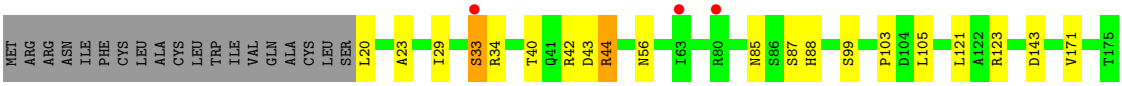
Chain M:

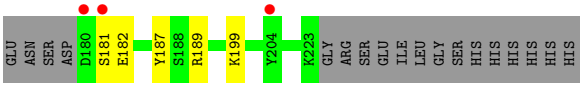




• Molecule 1: Acetylcholine-binding protein

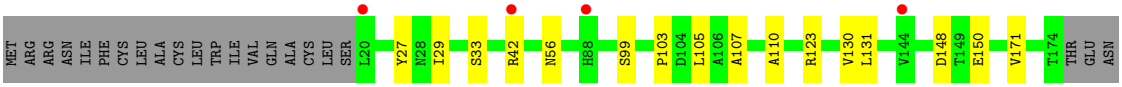
Chain N:

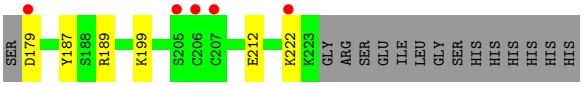




• Molecule 1: Acetylcholine-binding protein

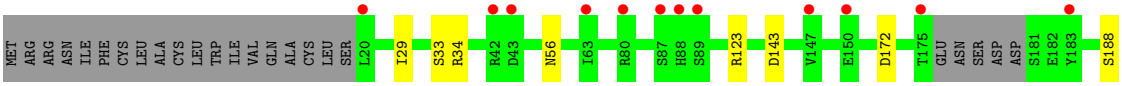
Chain O:

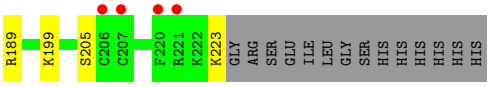




• Molecule 1: Acetylcholine-binding protein

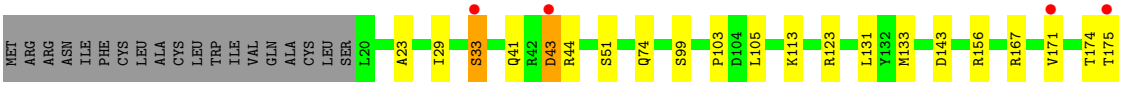
Chain P:

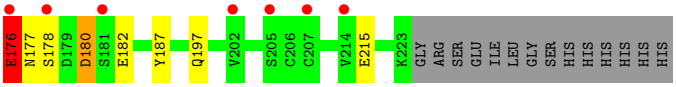




• Molecule 1: Acetylcholine-binding protein

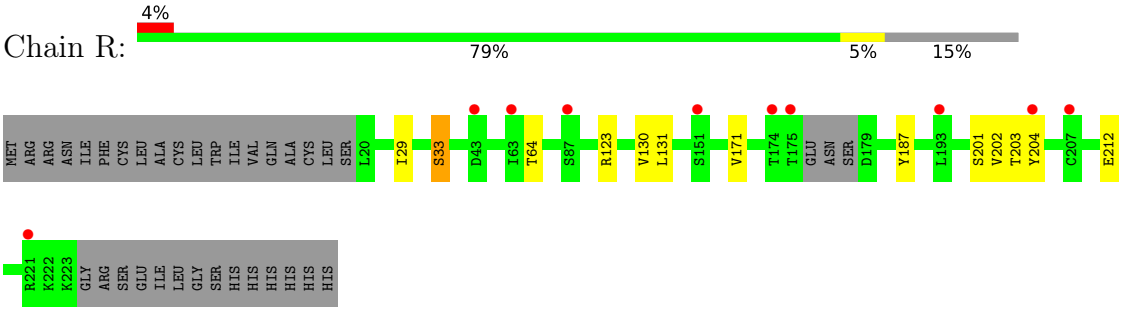
Chain Q:





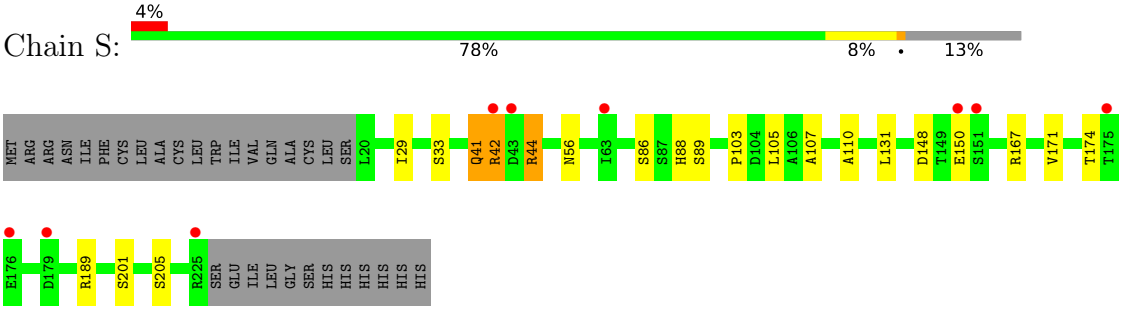
• Molecule 1: Acetylcholine-binding protein

Chain R:



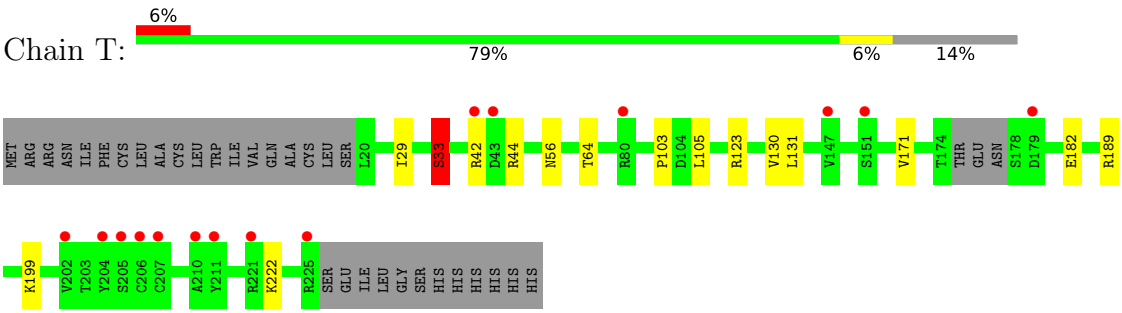
• Molecule 1: Acetylcholine-binding protein

Chain S:



• Molecule 1: Acetylcholine-binding protein

Chain T:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	116.71Å 140.71Å 147.17Å 90.00° 109.68° 90.00°	Depositor
Resolution (Å)	49.37 – 2.10 49.71 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (49.37-2.10) 99.7 (49.71-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.48 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.225 , 0.255 0.228 , 0.255	Depositor DCC
R_{free} test set	12970 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	50.2	Xtriage
Anisotropy	0.075	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 48.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	33433	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.40% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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, WCW

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	0.67	0/1661	0.84	0/2266
1	B	0.69	0/1667	0.88	1/2274 (0.0%)
1	C	0.66	0/1656	0.84	1/2259 (0.0%)
1	D	0.72	1/1658 (0.1%)	0.88	2/2260 (0.1%)
1	E	0.67	0/1647	0.88	3/2246 (0.1%)
1	F	0.81	2/1663 (0.1%)	0.95	3/2268 (0.1%)
1	G	0.78	1/1668 (0.1%)	0.97	5/2277 (0.2%)
1	H	0.70	0/1648	0.83	0/2248
1	I	0.71	0/1715	0.87	1/2340 (0.0%)
1	J	0.72	0/1669	0.90	3/2274 (0.1%)
1	K	0.67	1/1663 (0.1%)	0.83	1/2269 (0.0%)
1	L	0.66	1/1641 (0.1%)	0.84	1/2238 (0.0%)
1	M	0.70	1/1644 (0.1%)	0.88	2/2243 (0.1%)
1	N	0.69	0/1671	0.86	1/2279 (0.0%)
1	O	0.70	0/1648	0.84	0/2248
1	P	0.67	0/1639	0.82	0/2236
1	Q	0.70	0/1668	0.87	4/2277 (0.2%)
1	R	0.67	0/1644	0.82	0/2243
1	S	0.71	1/1683 (0.1%)	0.84	1/2296 (0.0%)
1	T	0.68	2/1682 (0.1%)	0.87	0/2293
All	All	0.70	10/33235 (0.0%)	0.87	29/45334 (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	B	0	2
1	D	0	3
1	E	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
1	F	0	2
1	G	0	1
1	J	0	1
1	M	0	1
1	N	0	2
1	Q	0	1
1	S	0	3
All	All	0	17

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	S	42	ARG	C-O	6.88	1.36	1.23
1	F	212	GLU	C-O	5.99	1.34	1.23
1	M	166	SER	CB-OG	-5.71	1.34	1.42
1	D	42	ARG	C-O	5.52	1.33	1.23
1	G	33	SER	CA-CB	-5.33	1.45	1.52

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	123	ARG	CG-CD-NE	-10.30	90.17	111.80
1	E	172	ASP	CB-CG-OD1	-9.77	109.50	118.30
1	G	123	ARG	CG-CD-NE	-7.89	95.22	111.80
1	F	123	ARG	CG-CD-NE	-7.88	95.25	111.80
1	L	156	ARG	CG-CD-NE	-7.44	96.17	111.80

There are no chirality outliers.

5 of 17 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	205	SER	Mainchain
1	B	206	CYS	Peptide
1	D	41	GLN	Peptide
1	D	44	ARG	Mainchain,Peptide
1	E	206	CYS	Peptide

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	1626	0	1575	15	0
1	B	1632	0	1580	16	0
1	C	1618	0	1570	15	0
1	D	1623	0	1571	17	1
1	E	1612	0	1557	4	0
1	F	1628	0	1570	23	0
1	G	1632	0	1575	15	0
1	H	1613	0	1560	9	0
1	I	1673	0	1619	18	2
1	J	1629	0	1589	16	0
1	K	1619	0	1580	7	0
1	L	1606	0	1552	19	3
1	M	1609	0	1556	12	0
1	N	1633	0	1590	16	2
1	O	1610	0	1562	18	1
1	P	1601	0	1562	8	0
1	Q	1632	0	1574	19	0
1	R	1609	0	1556	10	2
1	S	1647	0	1591	15	2
1	T	1647	0	1595	14	1
2	C	14	0	13	0	0
2	E	14	0	13	0	0
2	F	14	0	13	1	0
2	I	14	0	13	0	0
2	J	14	0	13	0	0
2	K	14	0	13	2	0
2	L	14	0	13	0	0
2	M	14	0	13	0	0
2	N	14	0	13	0	0
2	O	14	0	13	0	0
2	Q	14	0	13	0	0
2	R	14	0	13	0	0
2	T	14	0	13	0	0
3	E	16	0	0	0	0
3	G	16	0	0	0	0
3	I	16	0	0	2	0
4	A	31	0	0	1	0
4	B	25	0	0	0	0
4	C	31	0	0	0	0
4	D	37	0	0	0	0
4	E	37	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	F	50	0	0	1	0
4	G	60	0	0	0	0
4	H	55	0	0	0	0
4	I	52	0	0	2	0
4	J	56	0	0	2	0
4	K	15	0	0	0	0
4	L	15	0	0	1	0
4	M	29	0	0	0	0
4	N	36	0	0	2	0
4	O	35	0	0	0	0
4	P	17	0	0	1	0
4	Q	28	0	0	1	0
4	R	37	0	0	0	0
4	S	31	0	0	1	0
4	T	27	0	0	0	0
All	All	33433	0	31653	238	7

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 4.

The worst 5 of 238 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:F:203:THR:HG21	1:N:182:GLU:HG3	1.30	1.11
1:Q:74:GLN:HB2	1:Q:133:MET:HE3	1.29	1.07
1:Q:74:GLN:HB2	1:Q:133:MET:CE	1.89	1.02
1:G:74:GLN:HB2	1:G:133:MET:HE2	1.40	1.02
1:O:150:GLU:HB2	1:S:150:GLU:HG3	1.50	0.93

The worst 5 of 7 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:L:200:ASN:OD1	1:R:201:SER:OG[2_657]	1.74	0.46
1:L:199:LYS:O	1:R:203:THR:OG1[2_657]	1.91	0.29
1:N:88:HIS:ND1	1:T:42:ARG:NH2[1_455]	1.92	0.28
1:I:43:ASP:OD2	1:N:42:ARG:O[2_556]	1.97	0.23
1:D:201:SER:OG	1:S:201:SER:OG[2_656]	2.02	0.18

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	199/237 (84%)	197 (99%)	2 (1%)	0	100	100
1	B	200/237 (84%)	194 (97%)	6 (3%)	0	100	100
1	C	199/237 (84%)	197 (99%)	2 (1%)	0	100	100
1	D	199/237 (84%)	196 (98%)	2 (1%)	1 (0%)	29	26
1	E	198/237 (84%)	196 (99%)	2 (1%)	0	100	100
1	F	200/237 (84%)	197 (98%)	3 (2%)	0	100	100
1	G	202/237 (85%)	199 (98%)	2 (1%)	1 (0%)	29	26
1	H	198/237 (84%)	196 (99%)	2 (1%)	0	100	100
1	I	208/237 (88%)	205 (99%)	3 (1%)	0	100	100
1	J	199/237 (84%)	196 (98%)	3 (2%)	0	100	100
1	K	199/237 (84%)	196 (98%)	3 (2%)	0	100	100
1	L	197/237 (83%)	195 (99%)	2 (1%)	0	100	100
1	M	197/237 (83%)	194 (98%)	3 (2%)	0	100	100
1	N	200/237 (84%)	199 (100%)	1 (0%)	0	100	100
1	O	197/237 (83%)	196 (100%)	1 (0%)	0	100	100
1	P	196/237 (83%)	195 (100%)	1 (0%)	0	100	100
1	Q	202/237 (85%)	197 (98%)	3 (2%)	2 (1%)	15	11
1	R	197/237 (83%)	195 (99%)	2 (1%)	0	100	100
1	S	204/237 (86%)	202 (99%)	0	2 (1%)	15	11
1	T	202/237 (85%)	199 (98%)	3 (2%)	0	100	100
All	All	3993/4740 (84%)	3941 (99%)	46 (1%)	6 (0%)	47	49

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	175	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	S	174	THR
1	D	42	ARG
1	S	42	ARG
1	Q	180	ASP

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	190/220 (86%)	185 (97%)	5 (3%)	46	50
1	B	191/220 (87%)	188 (98%)	3 (2%)	62	69
1	C	189/220 (86%)	187 (99%)	2 (1%)	73	79
1	D	189/220 (86%)	181 (96%)	8 (4%)	30	30
1	E	188/220 (86%)	184 (98%)	4 (2%)	53	59
1	F	190/220 (86%)	185 (97%)	5 (3%)	46	50
1	G	191/220 (87%)	191 (100%)	0	100	100
1	H	188/220 (86%)	185 (98%)	3 (2%)	62	69
1	I	196/220 (89%)	193 (98%)	3 (2%)	65	71
1	J	189/220 (86%)	186 (98%)	3 (2%)	62	69
1	K	190/220 (86%)	185 (97%)	5 (3%)	46	50
1	L	187/220 (85%)	183 (98%)	4 (2%)	53	59
1	M	188/220 (86%)	185 (98%)	3 (2%)	62	69
1	N	191/220 (87%)	185 (97%)	6 (3%)	40	43
1	O	188/220 (86%)	187 (100%)	1 (0%)	88	92
1	P	187/220 (85%)	183 (98%)	4 (2%)	53	59
1	Q	191/220 (87%)	186 (97%)	5 (3%)	46	50
1	R	188/220 (86%)	186 (99%)	2 (1%)	73	79
1	S	192/220 (87%)	190 (99%)	2 (1%)	76	82
1	T	192/220 (87%)	189 (98%)	3 (2%)	62	69

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	3795/4400 (86%)	3724 (98%)	71 (2%)	59 63

5 of 71 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	P	172	ASP
1	Q	33	SER
1	R	123	ARG
1	F	159	ILE
1	F	100	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	Q	74	GLN
1	Q	219	ASN
1	G	56	ASN
1	G	88	HIS
1	J	41	GLN

5.3.3 RNA (i)

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

16 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	$\#\|Z\| > 2$
2	NAG	Q	301	1	14,14,15	0.64	0	17,19,21	1.50	4 (23%)
2	NAG	T	301	1	14,14,15	0.71	0	17,19,21	1.44	3 (17%)
3	WCW	I	301	-	15,17,17	0.69	1 (6%)	19,24,24	1.31	5 (26%)
2	NAG	K	301	1	14,14,15	0.93	1 (7%)	17,19,21	2.81	7 (41%)
2	NAG	R	301	1	14,14,15	0.64	0	17,19,21	1.44	1 (5%)
3	WCW	G	301	-	15,17,17	0.76	1 (6%)	19,24,24	1.37	2 (10%)
2	NAG	J	301	1	14,14,15	0.57	0	17,19,21	0.83	1 (5%)
2	NAG	I	302	1	14,14,15	0.71	0	17,19,21	1.69	5 (29%)
2	NAG	C	301	1	14,14,15	0.60	0	17,19,21	1.32	2 (11%)
2	NAG	E	302	1	14,14,15	0.66	0	17,19,21	1.27	2 (11%)
2	NAG	M	301	1	14,14,15	0.86	0	17,19,21	2.18	5 (29%)
2	NAG	O	301	1	14,14,15	0.79	0	17,19,21	1.78	3 (17%)
3	WCW	E	301	-	15,17,17	0.94	2 (13%)	19,24,24	1.21	2 (10%)
2	NAG	L	301	1	14,14,15	0.41	0	17,19,21	1.57	3 (17%)
2	NAG	N	301	1	14,14,15	0.52	0	17,19,21	1.41	2 (11%)
2	NAG	F	301	1	14,14,15	0.57	0	17,19,21	1.80	3 (17%)

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	Q	301	1	-	0/6/23/26	0/1/1/1
2	NAG	T	301	1	-	2/6/23/26	0/1/1/1
3	WCW	I	301	-	-	4/9/22/22	0/2/2/2
2	NAG	K	301	1	-	3/6/23/26	0/1/1/1
2	NAG	R	301	1	-	0/6/23/26	0/1/1/1
3	WCW	G	301	-	-	5/9/22/22	0/2/2/2
2	NAG	J	301	1	-	0/6/23/26	0/1/1/1
2	NAG	I	302	1	-	2/6/23/26	0/1/1/1
2	NAG	C	301	1	-	0/6/23/26	0/1/1/1
2	NAG	E	302	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
2	NAG	M	301	1	-	1/6/23/26	0/1/1/1
2	NAG	O	301	1	-	2/6/23/26	0/1/1/1
3	WCW	E	301	-	-	0/9/22/22	0/2/2/2
2	NAG	L	301	1	-	0/6/23/26	0/1/1/1
2	NAG	N	301	1	-	2/6/23/26	0/1/1/1
2	NAG	F	301	1	-	0/6/23/26	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	301	WCW	O2-C2	-2.37	1.21	1.30
3	E	301	WCW	C4-N2	2.14	1.50	1.46
3	G	301	WCW	O2-C2	-2.10	1.22	1.30
3	I	301	WCW	O2-C2	-2.02	1.22	1.30
2	K	301	NAG	C1-C2	2.02	1.55	1.52

The worst 5 of 50 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	K	301	NAG	C2-N2-C7	7.04	132.93	122.90
2	F	301	NAG	C1-O5-C5	5.93	120.22	112.19
2	M	301	NAG	C2-N2-C7	4.96	129.96	122.90
2	K	301	NAG	C8-C7-N2	4.96	124.49	116.10
2	N	301	NAG	C1-O5-C5	4.32	118.04	112.19

There are no chirality outliers.

5 of 21 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	K	301	NAG	C3-C2-N2-C7
2	M	301	NAG	C3-C2-N2-C7
3	G	301	WCW	N1-C1-C2-O2
3	I	301	WCW	N1-C1-C2-O2
2	N	301	NAG	C4-C5-C6-O6

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	I	301	WCW	2	0

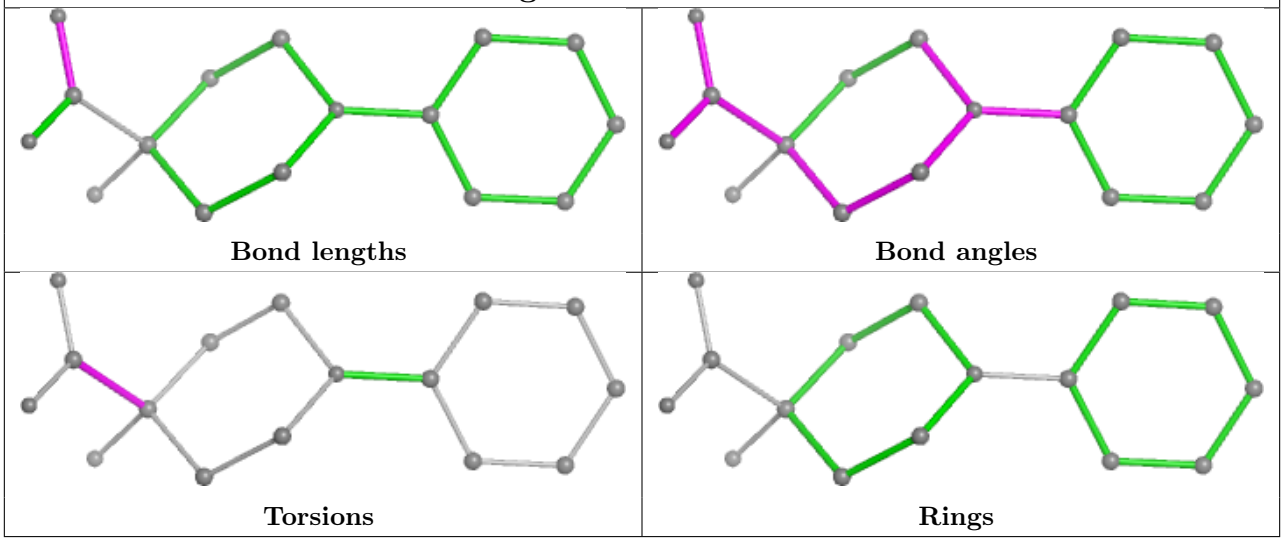
Continued on next page...

Continued from previous page...

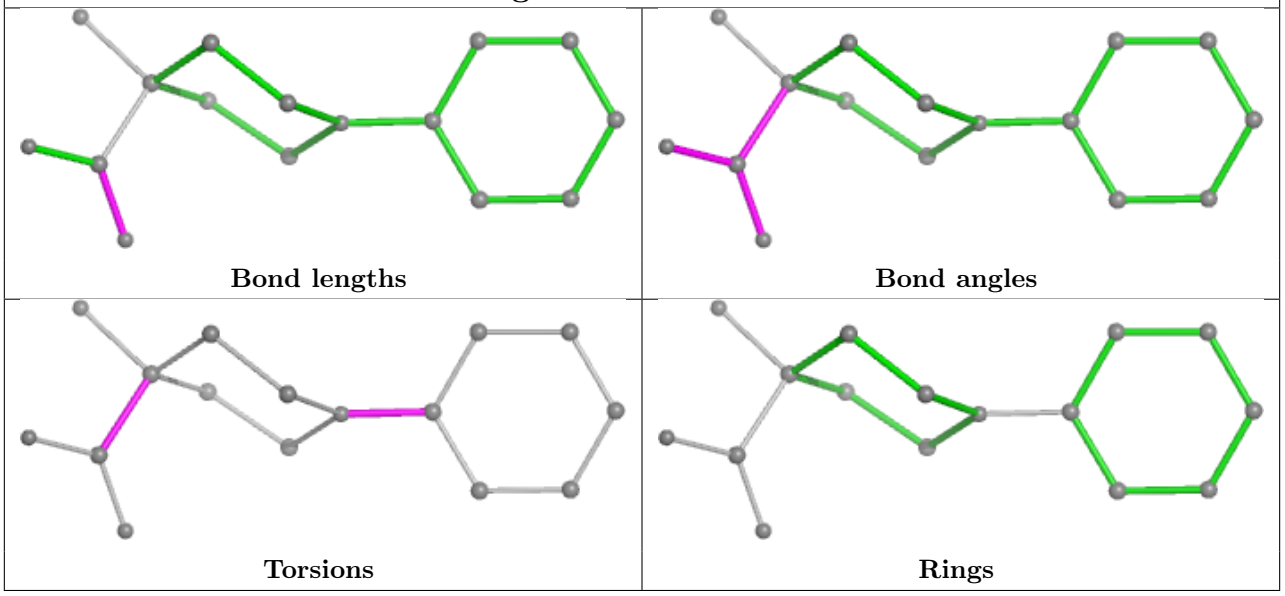
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	K	301	NAG	2	0
2	F	301	NAG	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

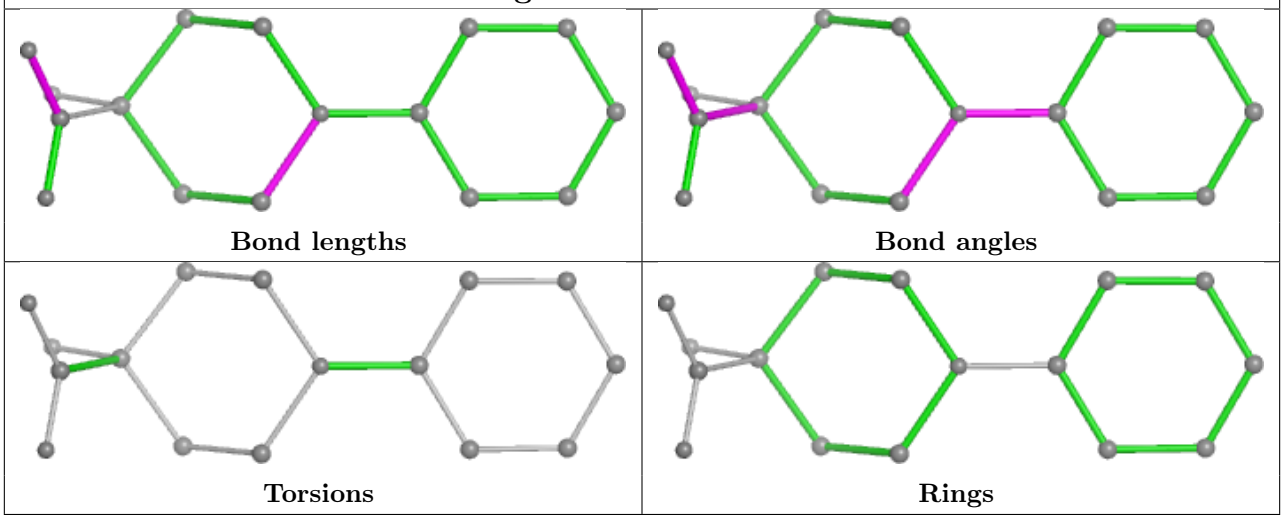
Ligand WCW I 301



Ligand WCW G 301



Ligand WCW E 301



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^{th} 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^2)	Q<0.9
1	A	200/237 (84%)	0.30	7 (3%) 44 50	40, 56, 96, 117	0
1	B	201/237 (84%)	0.62	11 (5%) 25 31	43, 60, 104, 145	0
1	C	202/237 (85%)	0.37	9 (4%) 33 38	43, 59, 94, 120	0
1	D	203/237 (85%)	0.40	9 (4%) 34 40	38, 58, 105, 148	0
1	E	202/237 (85%)	0.40	15 (7%) 14 18	39, 56, 97, 143	0
1	F	204/237 (86%)	0.38	7 (3%) 45 51	38, 53, 89, 106	0
1	G	204/237 (86%)	0.47	9 (4%) 34 40	37, 51, 87, 126	0
1	H	202/237 (85%)	0.20	5 (2%) 57 62	35, 49, 79, 112	0
1	I	205/237 (86%)	0.24	2 (0%) 82 85	34, 48, 78, 142	0
1	J	201/237 (84%)	0.25	4 (1%) 65 69	36, 49, 85, 99	0
1	K	200/237 (84%)	0.86	32 (16%) 1 2	44, 68, 119, 185	0
1	L	201/237 (84%)	0.73	17 (8%) 10 13	44, 70, 110, 164	0
1	M	201/237 (84%)	0.55	14 (6%) 16 20	40, 62, 101, 128	0
1	N	200/237 (84%)	0.33	6 (3%) 50 56	38, 55, 89, 116	0
1	O	200/237 (84%)	0.44	9 (4%) 33 38	37, 58, 106, 137	0
1	P	199/237 (83%)	0.57	16 (8%) 12 16	43, 65, 105, 142	0
1	Q	204/237 (86%)	0.49	11 (5%) 25 31	42, 60, 114, 144	0
1	R	201/237 (84%)	0.35	10 (4%) 28 34	40, 59, 92, 130	0
1	S	206/237 (86%)	0.31	9 (4%) 34 40	39, 54, 100, 128	0
1	T	203/237 (85%)	0.46	15 (7%) 14 18	41, 59, 109, 152	0
All	All	4039/4740 (85%)	0.44	217 (5%) 25 31	34, 57, 100, 185	0

The worst 5 of 217 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	42	ARG	9.4
1	B	204	TYR	7.8
1	T	43	ASP	5.8
1	I	42	ARG	5.5
1	S	42	ARG	5.5

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^{th} 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.

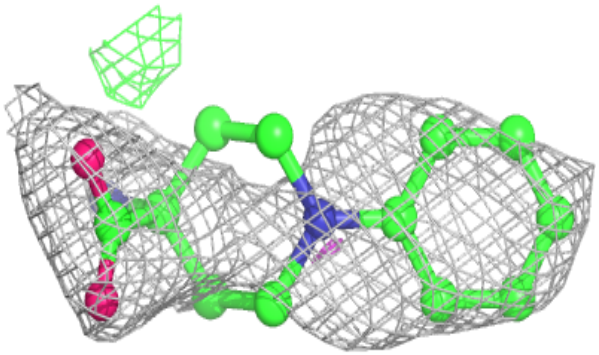
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAG	T	301	14/15	0.34	0.37	96,125,133,135	0
2	NAG	E	302	14/15	0.36	0.45	93,121,133,138	0
2	NAG	O	301	14/15	0.51	0.36	84,119,130,131	0
2	NAG	C	301	14/15	0.55	0.33	107,122,129,130	0
2	NAG	M	301	14/15	0.60	0.30	96,107,115,119	0
2	NAG	R	301	14/15	0.64	0.32	88,118,124,125	0
2	NAG	K	301	14/15	0.66	0.29	91,109,113,117	0
2	NAG	J	301	14/15	0.70	0.32	121,133,139,140	0
2	NAG	N	301	14/15	0.72	0.29	113,123,127,128	0
2	NAG	L	301	14/15	0.75	0.28	100,119,125,125	0
2	NAG	I	302	14/15	0.75	0.25	82,106,111,119	0
2	NAG	F	301	14/15	0.76	0.26	84,96,107,108	0
2	NAG	Q	301	14/15	0.79	0.22	100,117,121,124	0
3	WCW	E	301	16/16	0.79	0.27	72,95,114,115	0
3	WCW	I	301	16/16	0.83	0.37	72,87,97,117	0
3	WCW	G	301	16/16	0.91	0.20	78,85,95,104	0

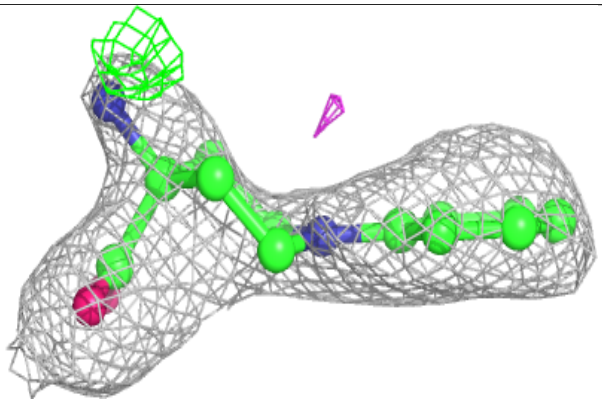
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers

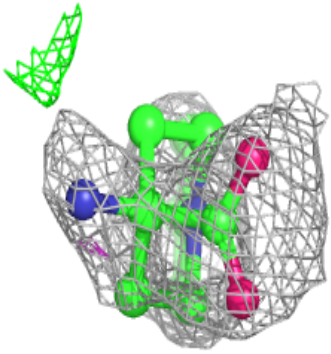
as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around WCW E 301:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

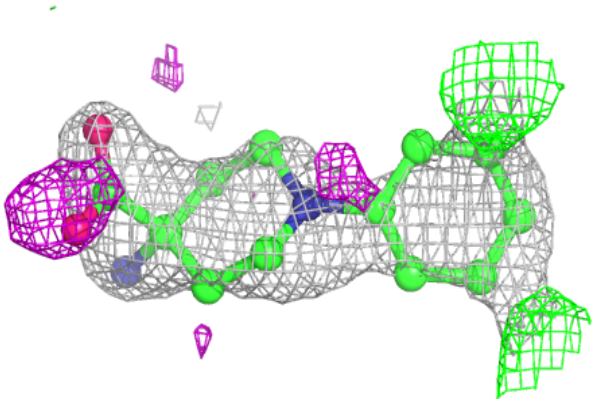


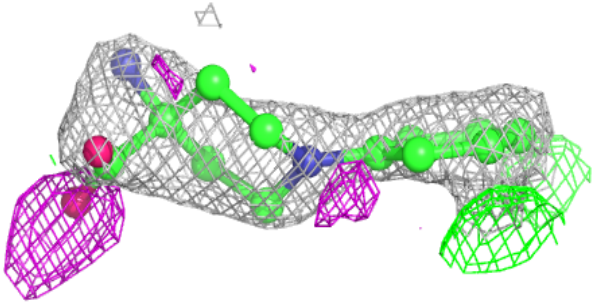


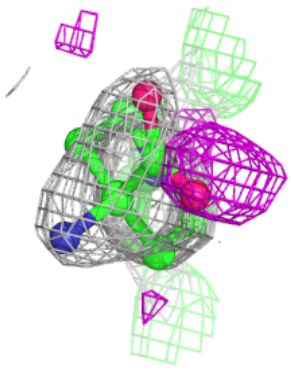


Electron density around WCW I 301:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

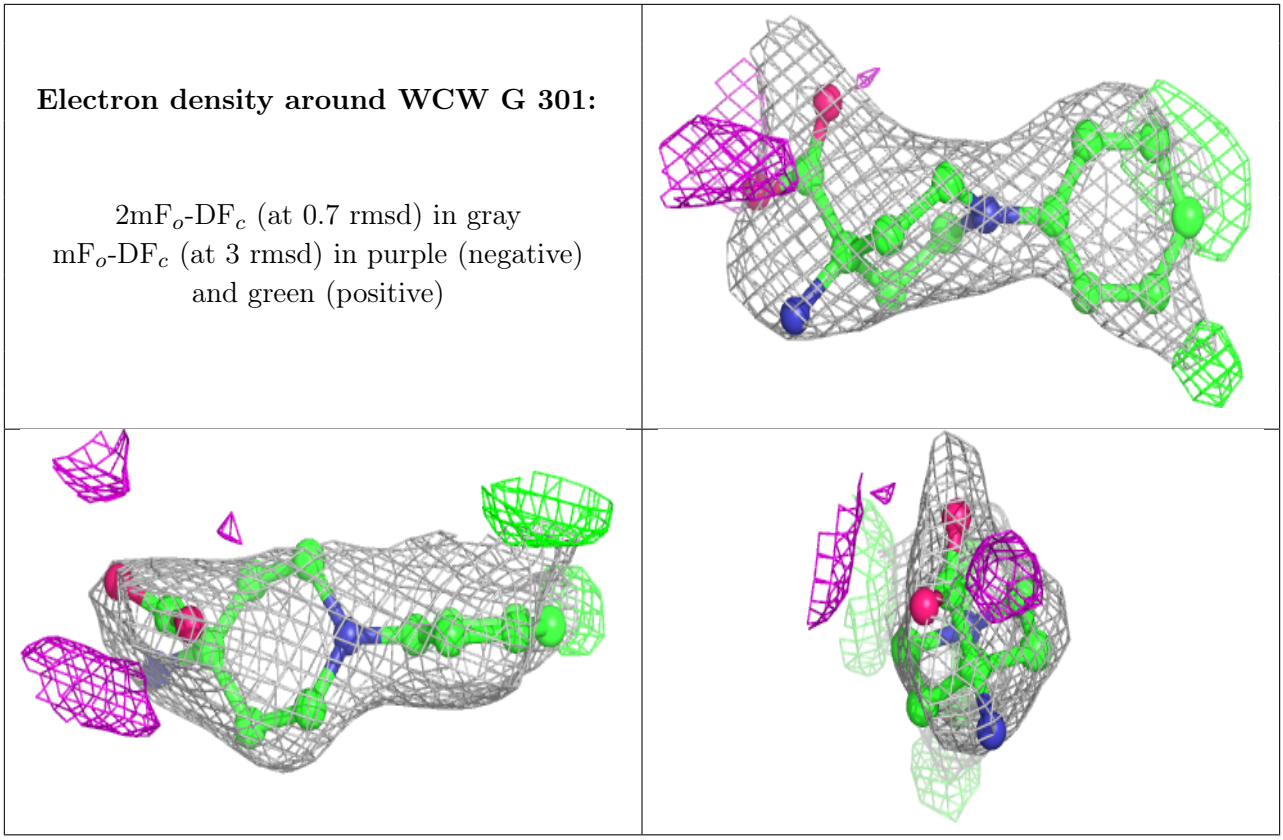






Electron density around WCW G 301:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.