



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 4, 2019 – 12:24 PM EDT

PDB ID : 6K5P
Title : Structure of mosquito-larvicidal Binary toxin receptor, Cqm1
Authors : Kumar, V.; Sharma, M.
Deposited on : 2019-05-30
Resolution : 1.80 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.8.0 (224370), CSD as540be (2019)
Xtriage (Phenix) : 1.13
EDS : 2.4
buster-report : 1.1.7 (2018)
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.4

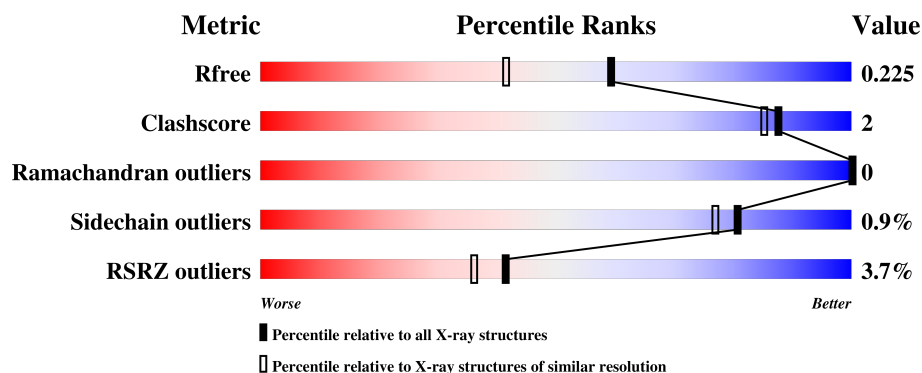
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	111664	5253 (1.80-1.80)
Clashscore	122126	6077 (1.80-1.80)
Ramachandran outliers	120053	6011 (1.80-1.80)
Sidechain outliers	120020	6010 (1.80-1.80)
RSRZ outliers	108989	5157 (1.80-1.80)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	561	<div> <div>2%</div> <div>89% 5% 5%</div> </div>
1	B	561	<div> <div>3%</div> <div>87% 7% 5%</div> </div>
1	C	561	<div> <div>5%</div> <div>90% 5% 5%</div> </div>
1	D	561	<div> <div>3%</div> <div>91% • 6%</div> </div>

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
5	ACT	B	607	-	-	X	-
5	ACT	D	607	-	-	X	-

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 19526 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 Binary toxin receptor protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	531	Total	C	N	O	S	0	0	0
			4310	2740	734	818	18			
1	B	531	Total	C	N	O	S	0	0	0
			4292	2728	727	819	18			
1	C	531	Total	C	N	O	S	0	0	0
			4277	2722	723	814	18			
1	D	529	Total	C	N	O	S	0	0	0
			4299	2733	730	818	18			

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mg	0	0
			1	1		
2	A	1	Total	Mg	0	0
			1	1		
2	D	1	Total	Mg	0	0
			1	1		
2	C	1	Total	Mg	0	0
			1	1		

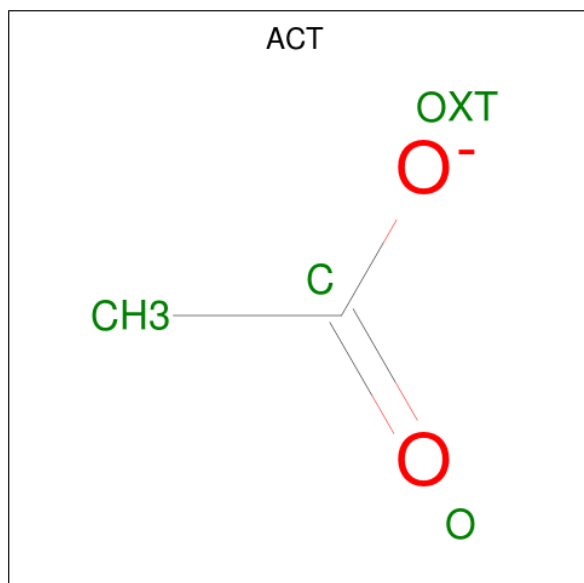
- Molecule 3 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	4	Total	Cd	0	0
			4	4		
3	A	5	Total	Cd	0	0
			5	5		
3	D	3	Total	Cd	0	0
			3	3		
3	C	2	Total	Cd	0	0
			2	2		

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by author).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Ca 1 1	0	0
4	A	1	Total Ca 1 1	0	0
4	D	1	Total Ca 1 1	0	0
4	C	1	Total Ca 1 1	0	0

- Molecule 5 is ACETATE ION (three-letter code: ACT) (formula: C₂H₃O₂).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 4 2 2	0	0
5	B	1	Total C O 4 2 2	0	0
5	C	1	Total C O 4 2 2	0	0
5	C	1	Total C O 4 2 2	0	0
5	D	1	Total C O 4 2 2	0	0

- Molecule 6 is COBALT (II) ION (three-letter code: CO) (formula: Co).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	1	Total 1	Co 1	0	0
6	C	1	Total 1	Co 1	0	0

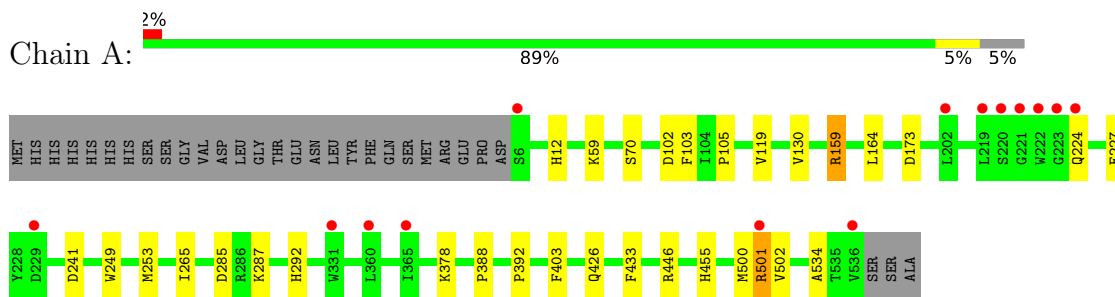
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	616	Total 616	O 616	0	0
7	B	601	Total 601	O 601	0	0
7	C	494	Total 494	O 494	0	0
7	D	593	Total 593	O 593	0	0

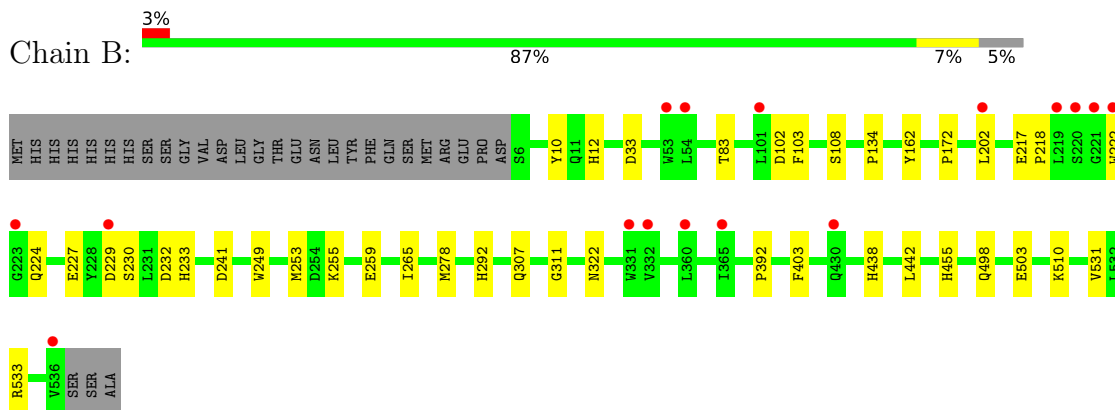
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

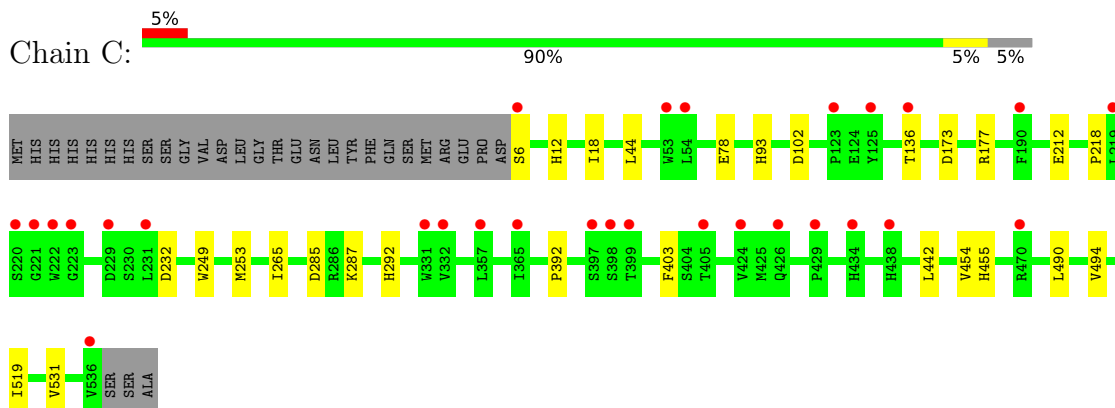
- Molecule 1: Binary toxin receptor protein



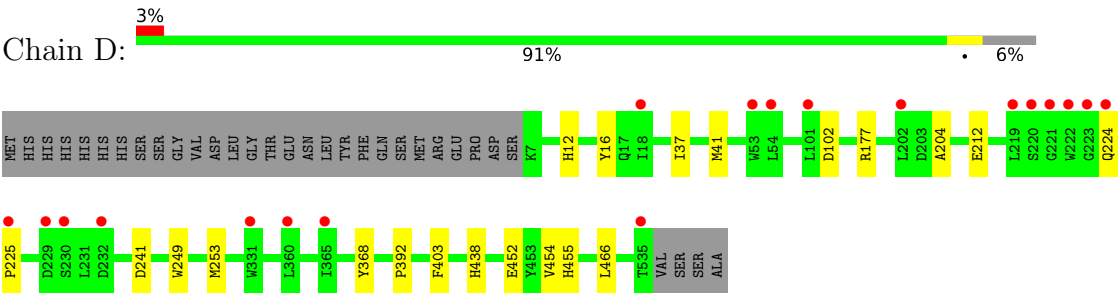
- Molecule 1: Binary toxin receptor protein



- Molecule 1: Binary toxin receptor protein



● Molecule 1: Binary toxin receptor protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	59.18Å 191.49Å 205.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.83 – 1.80 39.23 – 1.81	Depositor EDS
% Data completeness (in resolution range)	93.7 (29.83-1.80) 93.9 (39.23-1.81)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.61 (at 1.81Å)	Xtriage
Refinement program	PHENIX (1.14_3260: ???)	Depositor
R, R_{free}	0.183 , 0.227 0.183 , 0.225	Depositor DCC
R_{free} test set	9883 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å ²)	23.8	Xtriage
Anisotropy	0.350	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 59.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	19526	wwPDB-VP
Average B, all atoms (Å ²)	28.0	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 20.73 % 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 8.1569e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, ACT, CO, CA, CD

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	0.37	0/4438	0.56	0/6032
1	B	0.37	0/4418	0.55	0/6007
1	C	0.33	0/4405	0.52	0/5995
1	D	0.37	0/4427	0.55	0/6017
All	All	0.36	0/17688	0.55	0/24051

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4310	0	4049	22	0
1	B	4292	0	4010	22	0
1	C	4277	0	3978	16	0
1	D	4299	0	4034	14	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	5	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	4	0	0	0	0
3	C	2	0	0	0	0
3	D	3	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
5	A	4	0	3	0	0
5	B	4	0	3	2	0
5	C	8	0	6	2	0
5	D	4	0	3	3	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
7	A	616	0	0	4	0
7	B	601	0	0	3	0
7	C	494	0	0	5	0
7	D	593	0	0	3	0
All	All	19526	0	16086	74	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 2.

The worst 5 of 74 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:D:438:HIS:NE2	5:D:607:ACT:C	2.23	1.02
1:D:454:VAL:HG23	1:D:466:LEU:HB3	1.60	0.83
5:D:607:ACT:OXT	7:D:701:HOH:O	2.01	0.77
1:A:426:GLN:NE2	1:A:433:PHE:CE2	2.57	0.72
1:D:438:HIS:NE2	5:D:607:ACT:OXT	2.25	0.69

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	529/561 (94%)	515 (97%)	14 (3%)	0	100	100
1	B	529/561 (94%)	518 (98%)	11 (2%)	0	100	100
1	C	529/561 (94%)	517 (98%)	12 (2%)	0	100	100
1	D	527/561 (94%)	515 (98%)	12 (2%)	0	100	100
All	All	2114/2244 (94%)	2065 (98%)	49 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	462/495 (93%)	456 (99%)	6 (1%)	71	65
1	B	457/495 (92%)	452 (99%)	5 (1%)	76	71
1	C	453/495 (92%)	450 (99%)	3 (1%)	85	83
1	D	461/495 (93%)	458 (99%)	3 (1%)	85	83
All	All	1833/1980 (93%)	1816 (99%)	17 (1%)	81	77

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	102	ASP
1	B	241	ASP
1	C	455	HIS
1	B	33	ASP
1	D	102	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	426	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 29 ligands modelled in this entry, 24 are monoatomic - leaving 5 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z > 2$	Counts	RMSZ	# $ Z > 2$
5	ACT	A	607	-	1,3,3	6.55	1 (100%)	0,3,3	0.00	-
5	ACT	B	607	3	1,3,3	6.84	1 (100%)	0,3,3	0.00	-
5	ACT	C	606	6	1,3,3	6.63	1 (100%)	0,3,3	0.00	-
5	ACT	C	607	-	1,3,3	3.58	1 (100%)	0,3,3	0.00	-
5	ACT	D	607	6	1,3,3	6.48	1 (100%)	0,3,3	0.00	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	607	ACT	CH3-C	6.84	1.57	1.48
5	C	606	ACT	CH3-C	6.63	1.57	1.48
5	A	607	ACT	CH3-C	6.55	1.57	1.48

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	607	ACT	CH3-C	6.48	1.57	1.48
5	C	607	ACT	CH3-C	3.58	1.53	1.48

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 95th 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	531/561 (94%)	0.04	14 (2%) 56 51	15, 24, 40, 72	0
1	B	531/561 (94%)	0.10	16 (3%) 50 45	15, 23, 42, 70	0
1	C	531/561 (94%)	0.29	29 (5%) 25 20	18, 31, 46, 69	0
1	D	529/561 (94%)	0.11	19 (3%) 42 37	17, 23, 42, 73	0
All	All	2122/2244 (94%)	0.13	78 (3%) 41 36	15, 25, 43, 73	0

The worst 5 of 78 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	222	TRP	9.0
1	D	222	TRP	8.9
1	C	222	TRP	8.2
1	B	222	TRP	8.1
1	C	221	GLY	6.4

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

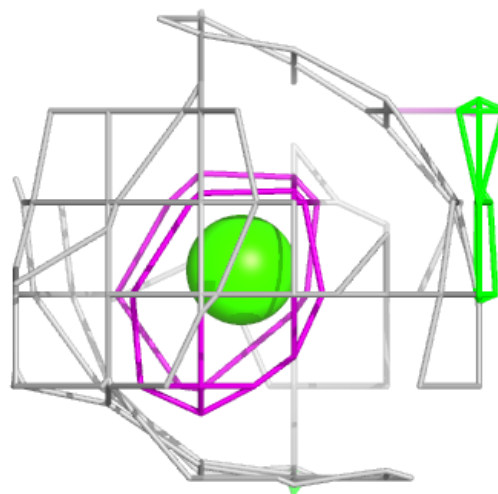
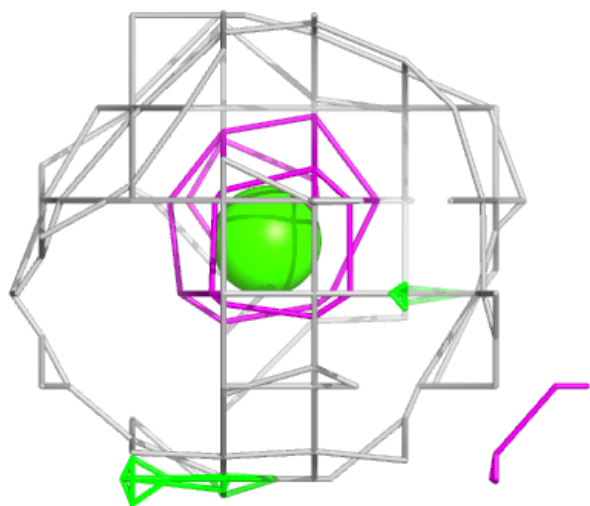
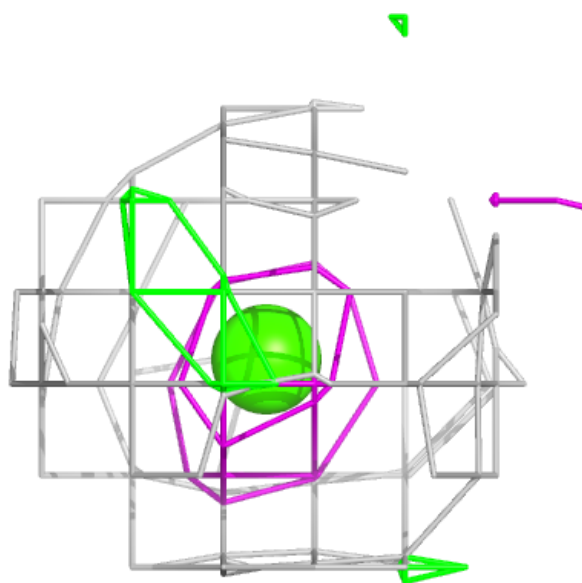
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	CD	C	604	1/1	0.90	0.09	48,48,48,48	0
5	ACT	C	606	4/4	0.91	0.13	27,30,37,37	0
5	ACT	B	607	4/4	0.91	0.15	19,20,28,40	0
5	ACT	C	607	4/4	0.91	0.12	31,39,42,58	0
5	ACT	A	607	4/4	0.92	0.15	47,54,56,59	0
3	CD	B	604	1/1	0.93	0.08	42,42,42,42	0
3	CD	A	608	1/1	0.93	0.06	67,67,67,67	0
5	ACT	D	607	4/4	0.93	0.13	23,25,32,40	0
3	CD	A	604	1/1	0.94	0.06	42,42,42,42	1
3	CD	D	604	1/1	0.94	0.10	39,39,39,39	0
3	CD	B	606	1/1	0.95	0.04	68,68,68,68	0
3	CD	A	606	1/1	0.96	0.05	35,35,35,35	0
3	CD	D	606	1/1	0.97	0.05	35,35,35,35	0
2	MG	C	601	1/1	0.97	0.03	26,26,26,26	0
6	CO	D	605	1/1	0.98	0.04	33,33,33,33	0
6	CO	C	605	1/1	0.98	0.03	36,36,36,36	0
4	CA	C	603	1/1	0.98	0.07	23,23,23,23	0
3	CD	C	602	1/1	0.99	0.03	39,39,39,39	1
2	MG	A	601	1/1	0.99	0.05	20,20,20,20	0
2	MG	B	601	1/1	0.99	0.03	21,21,21,21	0
3	CD	A	605	1/1	0.99	0.03	30,30,30,30	1
4	CA	A	603	1/1	0.99	0.08	19,19,19,19	0
3	CD	B	602	1/1	0.99	0.05	25,25,25,25	1
4	CA	D	603	1/1	0.99	0.08	17,17,17,17	0
3	CD	B	605	1/1	0.99	0.03	31,31,31,31	1
4	CA	B	603	1/1	0.99	0.07	19,19,19,19	0
2	MG	D	601	1/1	0.99	0.07	20,20,20,20	0
3	CD	A	602	1/1	1.00	0.04	29,29,29,29	1
3	CD	D	602	1/1	1.00	0.04	27,27,27,27	1

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.

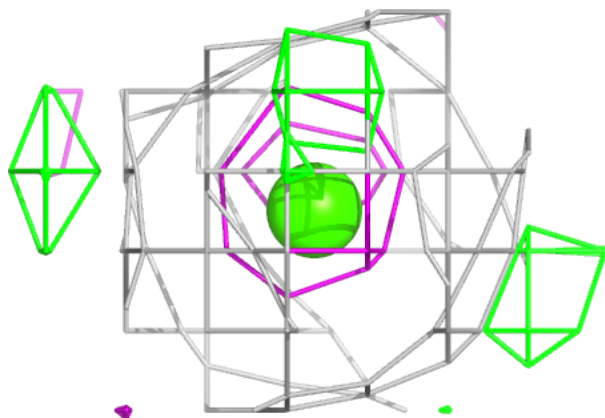
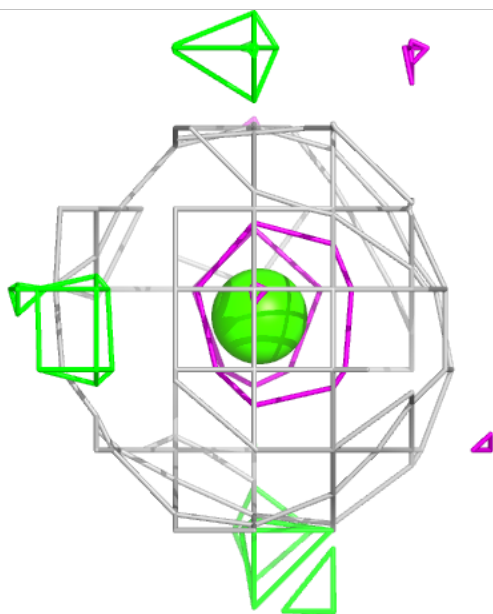
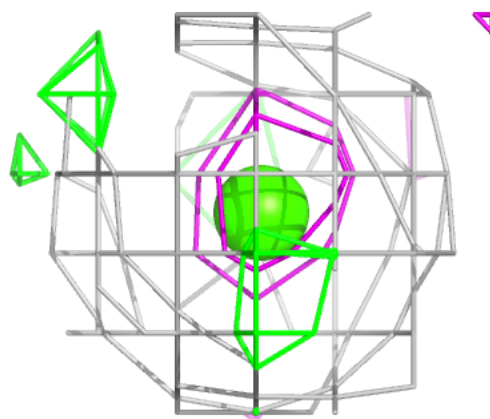
Electron density around CA C 603:

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



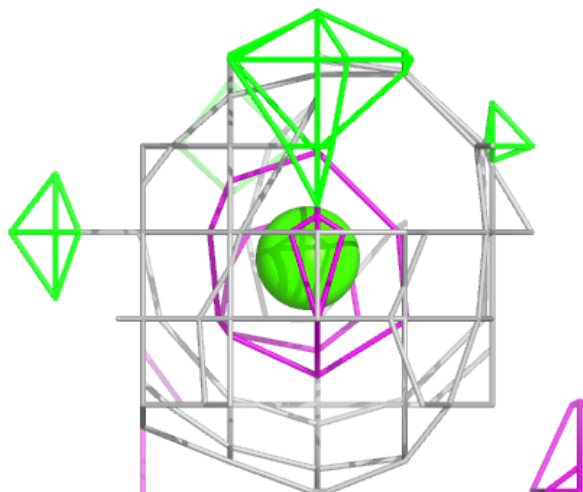
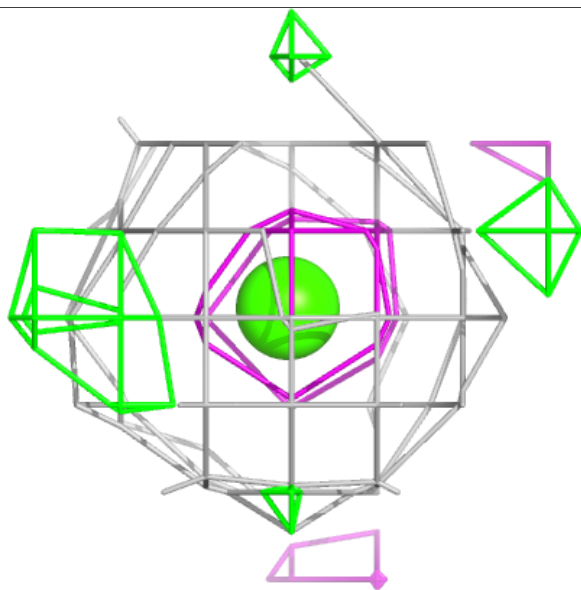
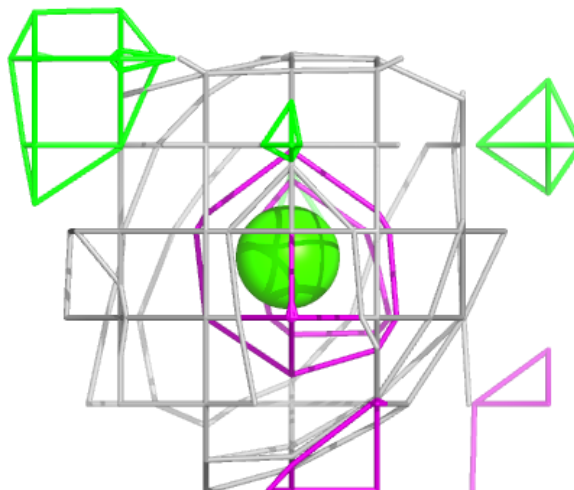
Electron density around CA A 603:

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



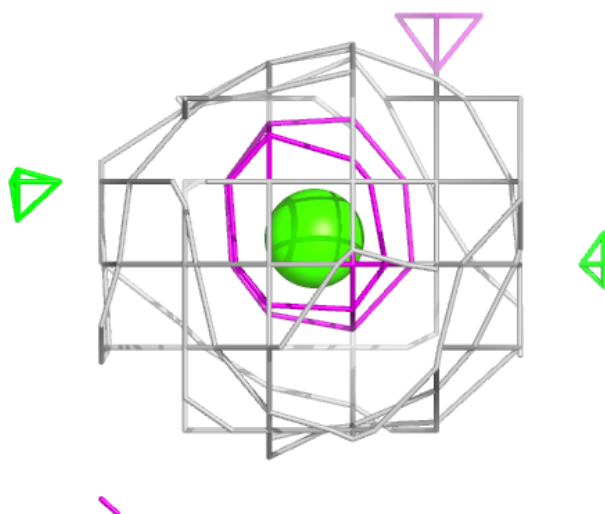
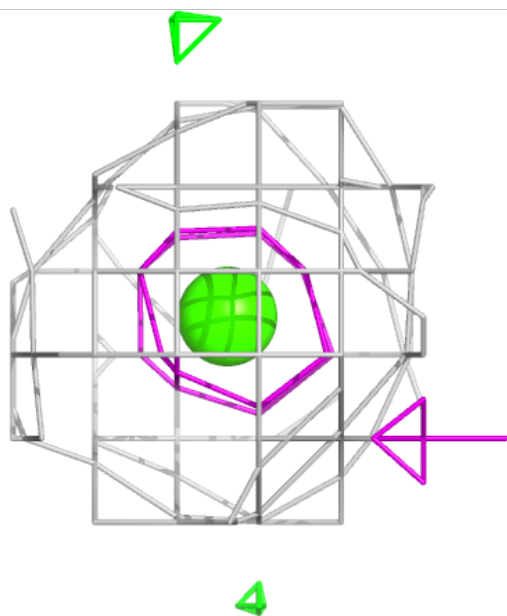
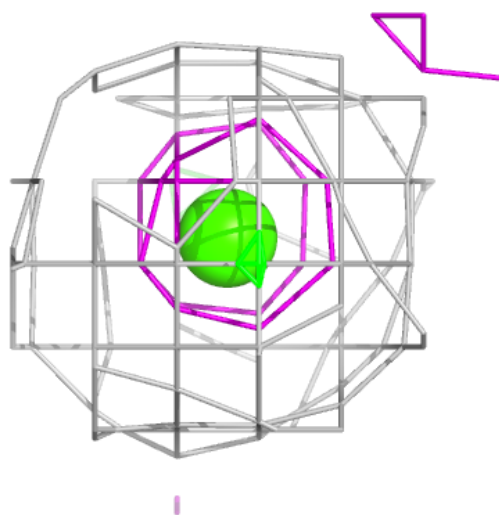
Electron density around CA D 603:

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



Electron density around CA B 603:

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



6.5 Other polymers ⓘ

There are no such residues in this entry.