Full wwPDB X-ray Structure Validation Report ⓘ

Jun 24, 2024 – 08:26 AM EDT

PDB ID	:	6ARV
Title	:	Crystal structure of CARM1 with Compound 2 and SAH
Authors	:	Boriack-Sjodin, P.A.; Jin, L.
Deposited on	:	2017-08-23
Resolution	:	2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

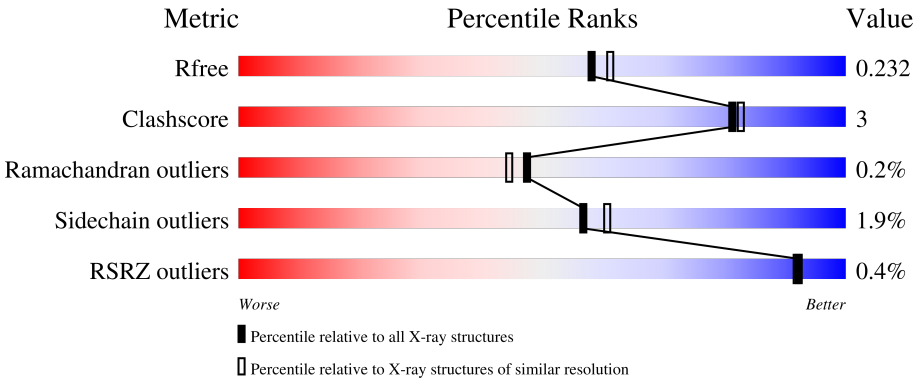
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	349	91% 7% •
1	B	349	91% 7% ••
1	C	349	% 91% 7% ••
1	D	349	% 92% 6% ••

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 12189 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.

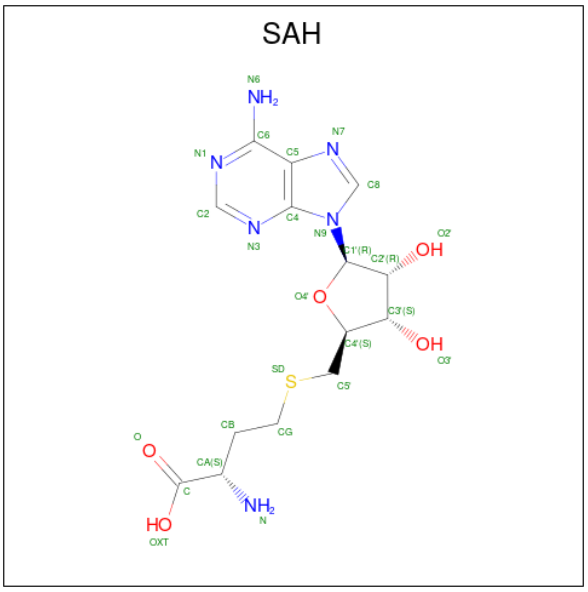
- Molecule 1 is a protein called Histone-arginine methyltransferase CARM1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	344	Total	C	N	O	S	0	7	0
			2810	1813	464	516	17			
1	B	344	Total	C	N	O	S	0	8	0
			2806	1814	462	515	15			
1	C	344	Total	C	N	O	S	0	3	0
			2776	1792	459	509	16			
1	D	344	Total	C	N	O	S	0	3	0
			2784	1794	462	513	15			

There are 12 discrepancies between the modelled and reference sequences:

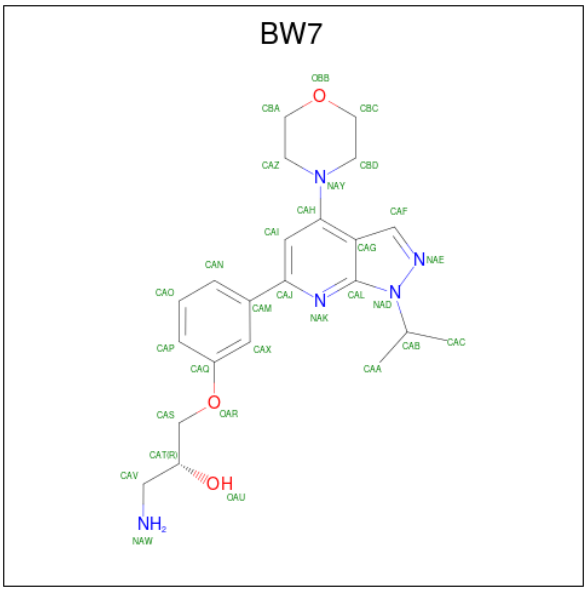
Chain	Residue	Modelled	Actual	Comment	Reference
A	131	SER	-	expression tag	UNP Q86X55
A	132	ILE	-	expression tag	UNP Q86X55
A	133	ALA	-	expression tag	UNP Q86X55
B	131	SER	-	expression tag	UNP Q86X55
B	132	ILE	-	expression tag	UNP Q86X55
B	133	ALA	-	expression tag	UNP Q86X55
C	131	SER	-	expression tag	UNP Q86X55
C	132	ILE	-	expression tag	UNP Q86X55
C	133	ALA	-	expression tag	UNP Q86X55
D	131	SER	-	expression tag	UNP Q86X55
D	132	ILE	-	expression tag	UNP Q86X55
D	133	ALA	-	expression tag	UNP Q86X55

- Molecule 2 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula: $C_{14}H_{20}N_6O_5S$).



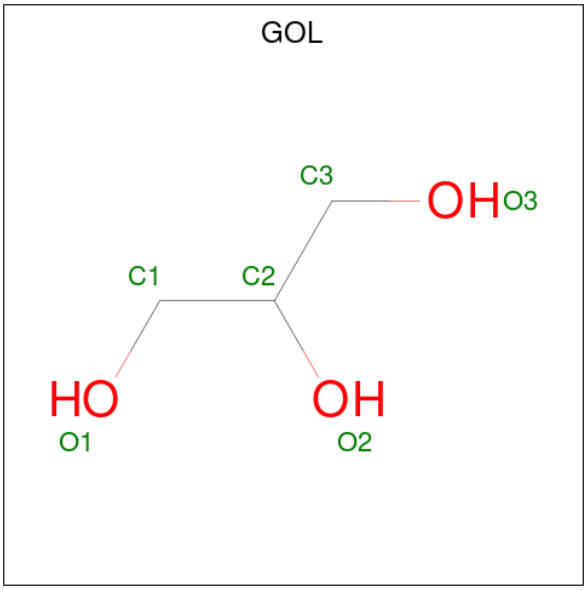
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			26	14	6	5	1		
2	B	1	Total	C	N	O	S	0	0
			26	14	6	5	1		
2	C	1	Total	C	N	O	S	0	0
			26	14	6	5	1		
2	D	1	Total	C	N	O	S	0	0
			26	14	6	5	1		

- Molecule 3 is (2R)-1-amino-3-{3-[4-(morpholin-4-yl)-1-(propan-2-yl)-1H-pyrazolo[3,4-b]pyridin-6-yl]phenoxy}propan-2-ol (three-letter code: BW7) (formula: $C_{22}H_{29}N_5O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O 30 22 5 3	0	0
3	B	1	Total C N O 30 22 5 3	0	0
3	C	1	Total C N O 30 22 5 3	0	0
3	D	1	Total C N O 30 22 5 3	0	0

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



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0
4	C	1	Total C O 6 3 3	0	0
4	C	1	Total C O 6 3 3	0	0

Continued on next page...

Continued from previous page...

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

- Molecule 5 is water.

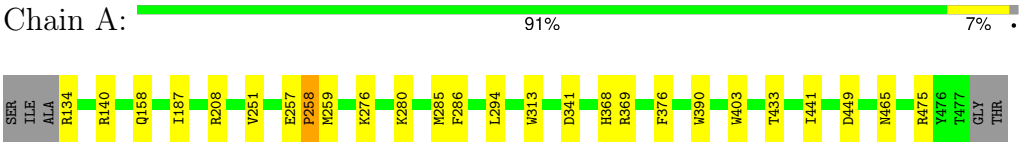
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	203	Total O 203 203	0	0
5	B	221	Total O 221 221	0	0
5	C	139	Total O 139 139	0	0
5	D	160	Total O 160 160	0	0

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.

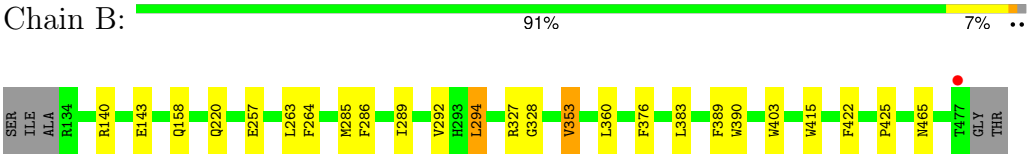
• Molecule 1: Histone-arginine methyltransferase CARM1

Chain A:



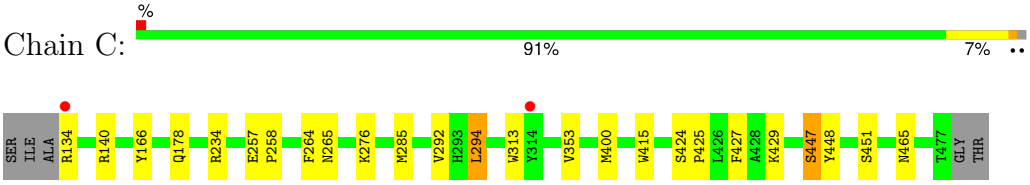
• Molecule 1: Histone-arginine methyltransferase CARM1

Chain B:



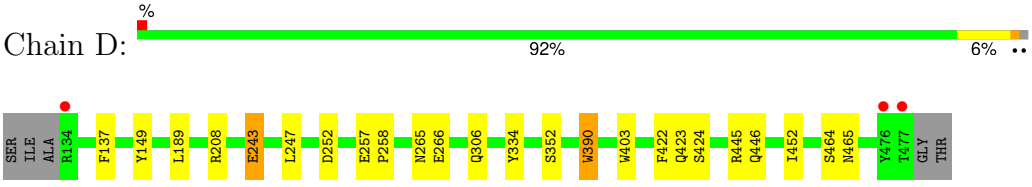
• Molecule 1: Histone-arginine methyltransferase CARM1

Chain C:



• Molecule 1: Histone-arginine methyltransferase CARM1

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	75.23Å 99.12Å 208.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	70.78 – 2.00 70.78 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.2 (70.78-2.00) 99.2 (70.78-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	2.17 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.189 , 0.232 0.189 , 0.232	Depositor DCC
R_{free} test set	5266 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	29.7	Xtriage
Anisotropy	0.219	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 45.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	12189	wwPDB-VP
Average B, all atoms (Å^2)	30.0	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 34.61 % 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 6.6419e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: SAH, GOL, BW7

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	0.61	3/2893 (0.1%)	0.67	0/3917
1	B	0.63	3/2901 (0.1%)	0.65	0/3929
1	C	0.57	2/2855 (0.1%)	0.63	0/3867
1	D	0.58	2/2854 (0.1%)	0.62	0/3866
All	All	0.60	10/11503 (0.1%)	0.64	0/15579

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	390	TRP	CD2-CE2	5.83	1.48	1.41
1	A	313	TRP	CD2-CE2	5.76	1.48	1.41
1	D	403	TRP	CD2-CE2	5.51	1.48	1.41
1	B	403	TRP	CD2-CE2	5.51	1.48	1.41
1	D	390	TRP	CD2-CE2	5.32	1.47	1.41
1	A	403	TRP	CD2-CE2	5.21	1.47	1.41
1	C	415	TRP	CD2-CE2	5.20	1.47	1.41
1	B	390	TRP	CD2-CE2	5.16	1.47	1.41
1	C	313	TRP	CD2-CE2	5.14	1.47	1.41
1	B	415	TRP	CD2-CE2	5.13	1.47	1.41

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen

atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2810	0	2761	12	0
1	B	2806	0	2771	16	0
1	C	2776	0	2734	11	0
1	D	2784	0	2730	16	0
2	A	26	0	19	0	0
2	B	26	0	19	0	0
2	C	26	0	19	0	0
2	D	26	0	19	0	0
3	A	30	0	0	1	0
3	B	30	0	0	1	0
3	C	30	0	0	1	0
3	D	30	0	0	1	0
4	A	24	0	32	0	0
4	B	12	0	16	3	0
4	C	12	0	16	1	0
4	D	18	0	24	3	0
5	A	203	0	0	4	0
5	B	221	0	0	2	0
5	C	139	0	0	0	0
5	D	160	0	0	2	0
All	All	12189	0	11160	59	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

All (59) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:158[A]:GLN:HG2	5:A:760:HOH:O	1.77	0.84
1:A:369:ARG:HG3	1:A:441:ILE:CD1	2.18	0.73
1:D:465[B]:ASN:ND2	1:D:465[B]:ASN:N	2.40	0.69
1:D:352:SER:H	4:D:504:GOL:H2	1.61	0.65
1:B:422:PHE:HA	1:B:465:ASN:ND2	2.13	0.64
1:D:465[B]:ASN:N	1:D:465[B]:ASN:HD22	1.95	0.64
3:A:502:BW7:CBD	3:A:502:BW7:CAF	2.75	0.64
3:C:502:BW7:CAF	3:C:502:BW7:CBD	2.76	0.63
1:D:423:GLN:N	1:D:465[B]:ASN:OD1	2.32	0.62
1:B:353:VAL:HG13	1:B:376:PHE:CE1	2.35	0.62
1:D:422:PHE:HA	1:D:465[B]:ASN:OD1	2.02	0.60

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:449:ASP:OD2	5:A:601:HOH:O	2.17	0.59
1:D:445:ARG:HD3	5:D:710:HOH:O	2.02	0.58
3:D:502:BW7:CAZ	3:D:502:BW7:CAF	2.82	0.58
1:A:368:HIS:HE1	5:A:613:HOH:O	1.87	0.57
3:B:502:BW7:CAZ	3:B:502:BW7:CAF	2.84	0.56
1:D:208:ARG:HD2	5:D:736:HOH:O	2.05	0.56
1:A:276:LYS:HD3	1:A:285[B]:MET:SD	2.46	0.56
1:A:341[A]:ASP:OD1	1:A:475:ARG:HD3	2.07	0.55
1:A:258:PRO:HD2	1:A:259:MET:SD	2.47	0.55
1:C:166:TYR:CE2	4:C:503:GOL:H11	2.43	0.54
1:D:464[B]:SER:C	1:D:465[B]:ASN:ND2	2.62	0.53
1:A:251:VAL:O	1:A:280:LYS:HG3	2.09	0.53
1:B:294:LEU:HD23	1:B:389:PHE:CE2	2.44	0.52
1:C:276:LYS:HD3	1:C:285[B]:MET:SD	2.51	0.51
1:A:158[A]:GLN:CG	5:A:760:HOH:O	2.49	0.51
1:D:422:PHE:HE1	1:D:452:ILE:HG21	1.77	0.50
1:D:252:ASP:OD1	4:D:505:GOL:H12	2.11	0.50
1:D:306:GLN:HE21	1:D:334:TYR:HB3	1.78	0.49
1:C:264:PHE:CE1	1:C:292:VAL:HG21	2.48	0.48
1:B:285[B]:MET:HG2	1:B:360:LEU:HD23	1.96	0.48
1:D:137:PHE:HB2	1:D:243:GLU:CG	2.43	0.48
1:B:158[B]:GLN:NE2	1:B:158[B]:GLN:HA	2.28	0.48
1:B:220:GLN:NE2	5:B:605:HOH:O	2.46	0.48
1:B:328:GLY:HA2	4:B:504:GOL:C1	2.44	0.47
1:C:427:PHE:CE2	1:C:429:LYS:HE2	2.49	0.47
1:C:427:PHE:HE2	1:C:429:LYS:HE2	1.78	0.47
1:D:390:TRP:HZ2	4:D:504:GOL:H11	1.79	0.47
1:C:265:ASN:ND2	1:C:447:SER:OG	2.48	0.46
1:A:369:ARG:HG3	1:A:441:ILE:HD11	1.96	0.45
1:C:294:LEU:HD12	1:C:353:VAL:HB	1.99	0.45
1:A:187:ILE:HD11	1:A:208:ARG:HH11	1.81	0.45
1:B:383:LEU:HD11	1:B:425:PRO:HB2	1.98	0.45
1:A:376:PHE:O	1:A:433:THR:HA	2.17	0.45
1:B:264:PHE:CE1	1:B:292:VAL:HG21	2.52	0.45
1:B:328:GLY:CA	4:B:504:GOL:H11	2.47	0.44
1:D:189:LEU:HD13	1:D:247:LEU:HD21	1.99	0.44
1:B:220:GLN:HG3	5:B:744:HOH:O	2.18	0.44
1:B:263:LEU:HD23	1:B:292:VAL:CG2	2.48	0.44
1:B:383:LEU:CD1	1:B:425:PRO:HB2	2.48	0.43
1:B:289:ILE:HD12	1:B:289:ILE:C	2.38	0.43
1:C:264:PHE:HB3	1:C:448:TYR:CE2	2.55	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:294:LEU:CD1	1:B:353:VAL:HG22	2.50	0.41
1:D:149:TYR:HB2	1:D:446:GLN:HE22	1.85	0.41
1:C:178:GLN:NE2	1:C:400:MET:SD	2.93	0.41
1:C:424:SER:HA	1:C:425:PRO:HD2	1.85	0.41
1:D:265:ASN:O	1:D:266:GLU:HB2	2.20	0.41
1:C:264:PHE:HE1	1:C:292:VAL:HG21	1.85	0.41
1:B:327:ARG:HB3	4:B:504:GOL:H32	2.04	0.40

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	349/349 (100%)	337 (97%)	11 (3%)	1 (0%)	41	37
1	B	350/349 (100%)	338 (97%)	12 (3%)	0	100	100
1	C	345/349 (99%)	333 (96%)	11 (3%)	1 (0%)	41	37
1	D	345/349 (99%)	336 (97%)	8 (2%)	1 (0%)	41	37
All	All	1389/1396 (100%)	1344 (97%)	42 (3%)	3 (0%)	47	44

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	258	PRO
1	D	258	PRO
1	C	258	PRO

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	307/303 (101%)	301 (98%)	6 (2%)	55	58
1	B	308/303 (102%)	302 (98%)	6 (2%)	57	61
1	C	303/303 (100%)	295 (97%)	8 (3%)	46	48
1	D	303/303 (100%)	300 (99%)	3 (1%)	76	81
All	All	1221/1212 (101%)	1198 (98%)	23 (2%)	57	61

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	134	ARG
1	A	140	ARG
1	A	257	GLU
1	A	286	PHE
1	A	294	LEU
1	A	465	ASN
1	B	140	ARG
1	B	143	GLU
1	B	257	GLU
1	B	286	PHE
1	B	294	LEU
1	B	353	VAL
1	C	134	ARG
1	C	140	ARG
1	C	234	ARG
1	C	257	GLU
1	C	294	LEU
1	C	447	SER
1	C	451	SER
1	C	465	ASN
1	D	243	GLU
1	D	257	GLU
1	D	424	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (22) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	148	GLN
1	A	265	ASN
1	A	417	GLN
1	A	446	GLN
1	B	148	GLN
1	B	160	GLN
1	B	178	GLN
1	B	220	GLN
1	B	265	ASN
1	B	301	GLN
1	B	446	GLN
1	B	465	ASN
1	C	148	GLN
1	C	160	GLN
1	C	164	GLN
1	C	265	ASN
1	C	311	ASN
1	D	151	GLN
1	D	178	GLN
1	D	265	ASN
1	D	306	GLN
1	D	446	GLN

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

19 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	$\#\|Z\| > 2$
4	GOL	A	506	-	5,5,5	0.13	0	5,5,5	0.56	0
4	GOL	D	503	-	5,5,5	0.37	0	5,5,5	0.22	0
4	GOL	C	504	-	5,5,5	0.26	0	5,5,5	0.33	0
3	BW7	C	502	-	30,33,33	1.64	5 (16%)	35,46,46	1.44	5 (14%)
4	GOL	D	504	-	5,5,5	0.28	0	5,5,5	0.58	0
4	GOL	D	505	-	5,5,5	0.31	0	5,5,5	0.41	0
2	SAH	A	501	-	23,28,28	1.09	1 (4%)	22,40,40	1.68	3 (13%)
3	BW7	A	502	-	30,33,33	1.61	5 (16%)	35,46,46	1.54	5 (14%)
2	SAH	B	501	-	23,28,28	0.98	0	22,40,40	1.46	3 (13%)
4	GOL	A	503	-	5,5,5	0.33	0	5,5,5	0.49	0
4	GOL	A	504	-	5,5,5	0.22	0	5,5,5	0.47	0
3	BW7	D	502	-	30,33,33	1.65	4 (13%)	35,46,46	1.36	6 (17%)
4	GOL	B	503	-	5,5,5	0.14	0	5,5,5	0.47	0
4	GOL	A	505	-	5,5,5	0.19	0	5,5,5	0.61	0
4	GOL	B	504	-	5,5,5	0.33	0	5,5,5	0.49	0
2	SAH	D	501	-	23,28,28	1.08	2 (8%)	22,40,40	1.37	3 (13%)
3	BW7	B	502	-	30,33,33	1.63	4 (13%)	35,46,46	1.53	5 (14%)
4	GOL	C	503	-	5,5,5	0.31	0	5,5,5	0.47	0
2	SAH	C	501	-	23,28,28	1.06	3 (13%)	22,40,40	1.40	4 (18%)

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
4	GOL	A	506	-	-	4/4/4/4	-
4	GOL	D	503	-	-	0/4/4/4	-
4	GOL	C	504	-	-	1/4/4/4	-
3	BW7	C	502	-	-	3/19/27/27	0/4/4/4
4	GOL	D	504	-	-	1/4/4/4	-
4	GOL	D	505	-	-	2/4/4/4	-
2	SAH	A	501	-	-	1/11/31/31	0/3/3/3
3	BW7	A	502	-	-	2/19/27/27	0/4/4/4

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SAH	B	501	-	-	0/11/31/31	0/3/3/3
4	GOL	A	503	-	-	2/4/4/4	-
4	GOL	A	504	-	-	0/4/4/4	-
3	BW7	D	502	-	-	3/19/27/27	0/4/4/4
4	GOL	B	503	-	-	0/4/4/4	-
4	GOL	A	505	-	-	1/4/4/4	-
4	GOL	B	504	-	-	3/4/4/4	-
2	SAH	D	501	-	-	0/11/31/31	0/3/3/3
3	BW7	B	502	-	-	4/19/27/27	0/4/4/4
4	GOL	C	503	-	-	2/4/4/4	-
2	SAH	C	501	-	-	0/11/31/31	0/3/3/3

All (24) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	502	BW7	CAM-CAJ	-5.69	1.40	1.49
3	D	502	BW7	CAM-CAJ	-5.53	1.40	1.49
3	C	502	BW7	CAM-CAJ	-5.25	1.40	1.49
3	A	502	BW7	CAM-CAJ	-5.07	1.41	1.49
3	D	502	BW7	CAH-CAG	-3.78	1.34	1.42
3	D	502	BW7	CAG-CAL	-3.55	1.34	1.43
3	C	502	BW7	CAH-CAG	-3.54	1.35	1.42
3	B	502	BW7	CAH-CAG	-3.36	1.35	1.42
3	C	502	BW7	CAG-CAL	-3.36	1.35	1.43
3	A	502	BW7	CAH-CAG	-3.31	1.35	1.42
3	A	502	BW7	CAG-CAL	-3.25	1.35	1.43
3	B	502	BW7	CAG-CAL	-3.17	1.35	1.43
2	A	501	SAH	C2-N3	2.87	1.36	1.32
3	A	502	BW7	CAF-CAG	-2.70	1.33	1.40
3	B	502	BW7	CAF-CAG	-2.69	1.33	1.40
3	D	502	BW7	CAF-CAG	-2.65	1.33	1.40
3	C	502	BW7	CAF-CAG	-2.55	1.34	1.40
2	C	501	SAH	C2-N3	2.40	1.35	1.32
3	A	502	BW7	CAZ-NAY	2.30	1.50	1.46
2	D	501	SAH	C2-N3	2.21	1.35	1.32
3	C	502	BW7	CBD-NAY	2.15	1.50	1.46
2	D	501	SAH	OXT-C	-2.14	1.23	1.30
2	C	501	SAH	OXT-C	-2.14	1.23	1.30
2	C	501	SAH	O4'-C1'	2.11	1.43	1.40

All (34) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	501	SAH	N3-C2-N1	-5.09	121.76	128.67
2	B	501	SAH	N3-C2-N1	-4.51	122.55	128.67
3	A	502	BW7	CBD-NAY-CAZ	4.32	121.29	111.57
2	D	501	SAH	N3-C2-N1	-4.32	122.81	128.67
3	C	502	BW7	CBD-NAY-CAZ	4.13	120.85	111.57
2	C	501	SAH	N3-C2-N1	-3.93	123.34	128.67
3	D	502	BW7	CAI-CAJ-NAK	-3.90	118.62	122.17
3	B	502	BW7	CBD-NAY-CAZ	3.85	120.23	111.57
3	B	502	BW7	CAI-CAJ-NAK	-3.85	118.67	122.17
3	C	502	BW7	CAI-CAJ-NAK	-3.76	118.75	122.17
3	A	502	BW7	CAC-CAB-NAD	3.56	113.66	109.48
3	A	502	BW7	CAI-CAJ-NAK	-3.44	119.04	122.17
3	B	502	BW7	CAM-CAJ-NAK	3.14	121.23	117.05
2	A	501	SAH	O4'-C1'-N9	-2.93	104.86	108.75
2	A	501	SAH	OXT-C-O	-2.86	117.59	124.08
3	D	502	BW7	CBD-NAY-CAZ	2.83	117.93	111.57
3	D	502	BW7	CAM-CAJ-NAK	2.80	120.77	117.05
3	B	502	BW7	CAC-CAB-NAD	2.72	112.67	109.48
3	D	502	BW7	CAJ-NAK-CAL	2.58	121.36	117.98
3	C	502	BW7	CAC-CAB-NAD	2.58	112.51	109.48
3	D	502	BW7	CAC-CAB-NAD	2.55	112.47	109.48
3	B	502	BW7	CAJ-NAK-CAL	2.53	121.30	117.98
2	D	501	SAH	OXT-C-O	-2.48	118.45	124.08
3	C	502	BW7	CAJ-NAK-CAL	2.47	121.22	117.98
3	C	502	BW7	CBA-CAZ-NAY	2.46	114.59	109.93
2	C	501	SAH	OXT-C-O	-2.39	118.65	124.08
3	A	502	BW7	CBC-CBD-NAY	2.36	114.39	109.93
2	C	501	SAH	C4-C5-N7	-2.30	106.91	109.34
2	B	501	SAH	O4'-C1'-N9	-2.25	105.76	108.75
2	B	501	SAH	OXT-C-O	-2.15	119.19	124.08
3	A	502	BW7	CBA-CAZ-NAY	2.12	113.94	109.93
2	D	501	SAH	C4-C5-N7	-2.10	107.12	109.34
3	D	502	BW7	CAZ-NAY-CAH	2.05	122.38	116.32
2	C	501	SAH	N6-C6-N1	2.05	122.71	118.33

There are no chirality outliers.

All (29) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	502	BW7	CAI-CAH-NAY-CBD
3	A	502	BW7	CAG-CAH-NAY-CBD
3	B	502	BW7	CAI-CAH-NAY-CAZ
3	B	502	BW7	CAG-CAH-NAY-CAZ

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	C	502	BW7	CAI-CAH-NAY-CBD
3	C	502	BW7	CAG-CAH-NAY-CBD
3	D	502	BW7	CAI-CAH-NAY-CAZ
3	D	502	BW7	CAG-CAH-NAY-CAZ
4	A	506	GOL	O1-C1-C2-C3
4	B	504	GOL	C1-C2-C3-O3
4	B	504	GOL	O2-C2-C3-O3
4	C	503	GOL	O1-C1-C2-C3
4	D	505	GOL	O1-C1-C2-C3
4	C	503	GOL	O1-C1-C2-O2
4	A	503	GOL	O1-C1-C2-C3
4	A	506	GOL	O1-C1-C2-O2
4	D	505	GOL	O1-C1-C2-O2
4	A	503	GOL	O1-C1-C2-O2
3	B	502	BW7	CAC-CAB-NAD-CAL
4	A	506	GOL	O2-C2-C3-O3
4	A	506	GOL	C1-C2-C3-O3
4	D	504	GOL	O2-C2-C3-O3
3	C	502	BW7	NAK-CAJ-CAM-CAN
4	A	505	GOL	O2-C2-C3-O3
4	C	504	GOL	O1-C1-C2-O2
3	B	502	BW7	CAA-CAB-NAD-NAE
4	B	504	GOL	O1-C1-C2-O2
3	D	502	BW7	CAC-CAB-NAD-CAL
2	A	501	SAH	OXT-C-CA-N

There are no ring outliers.

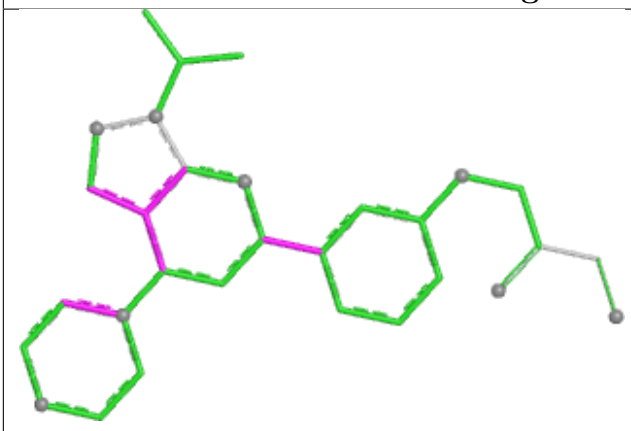
8 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	502	BW7	1	0
4	D	504	GOL	2	0
4	D	505	GOL	1	0
3	A	502	BW7	1	0
3	D	502	BW7	1	0
4	B	504	GOL	3	0
3	B	502	BW7	1	0
4	C	503	GOL	1	0

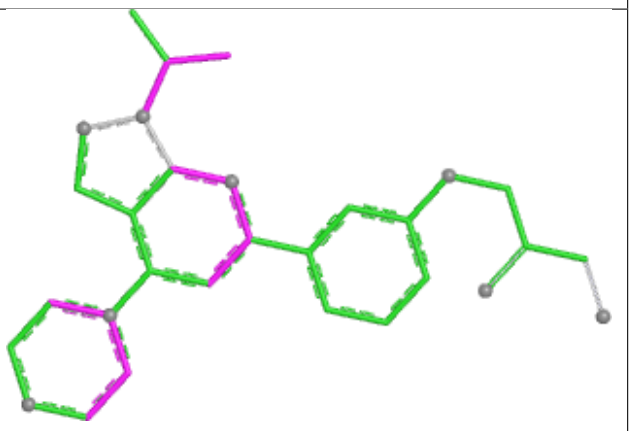
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will

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.

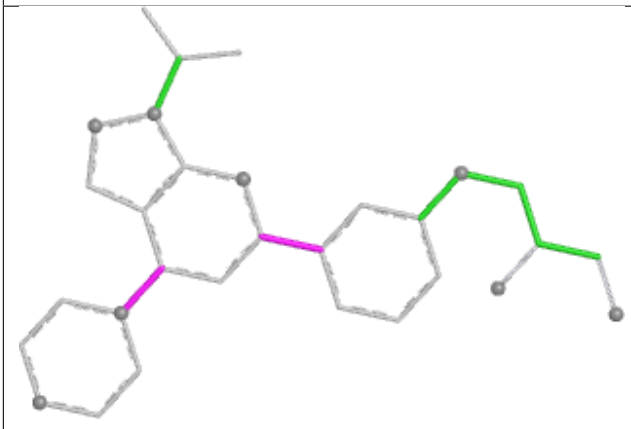
Ligand BW7 C 502



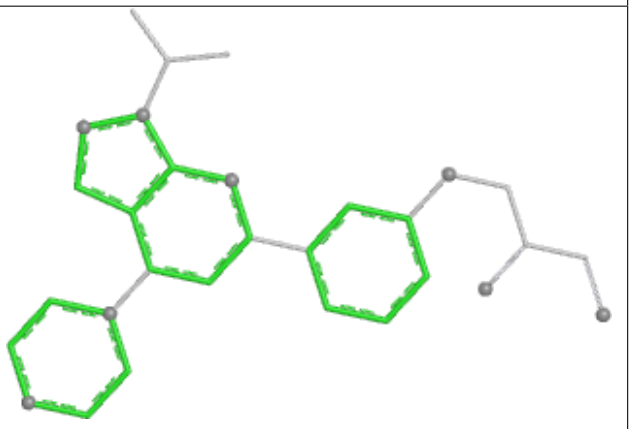
Bond lengths



Bond angles

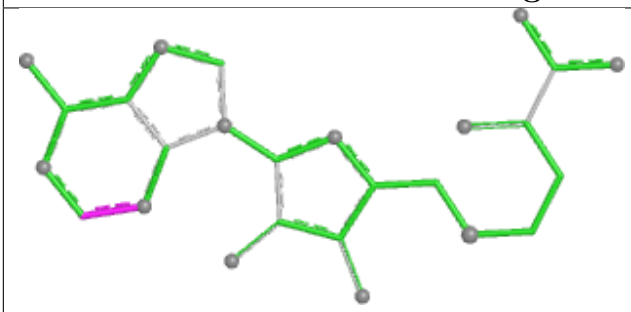


Torsions

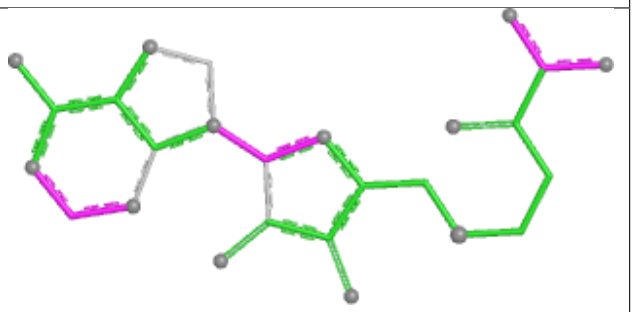


Rings

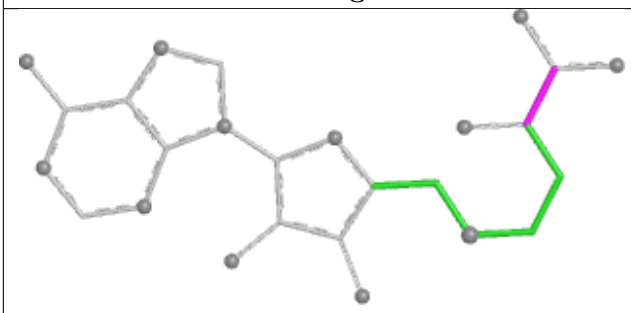
Ligand SAH A 501



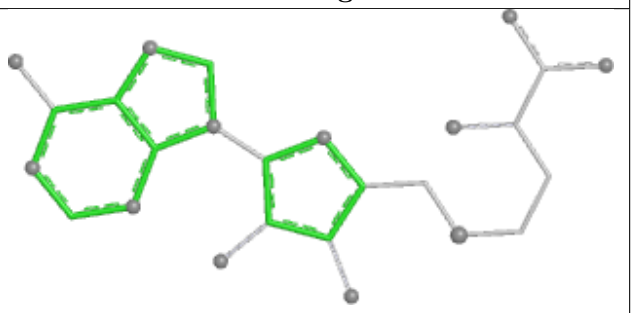
Bond lengths



Bond angles

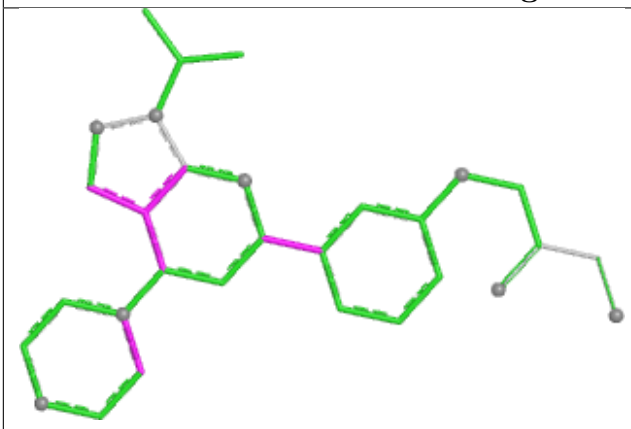


Torsions

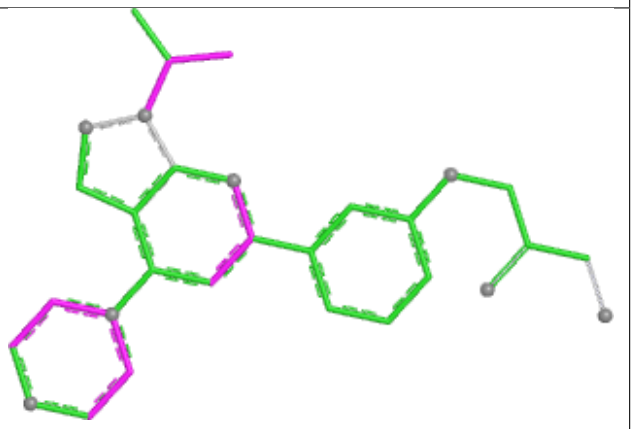


Rings

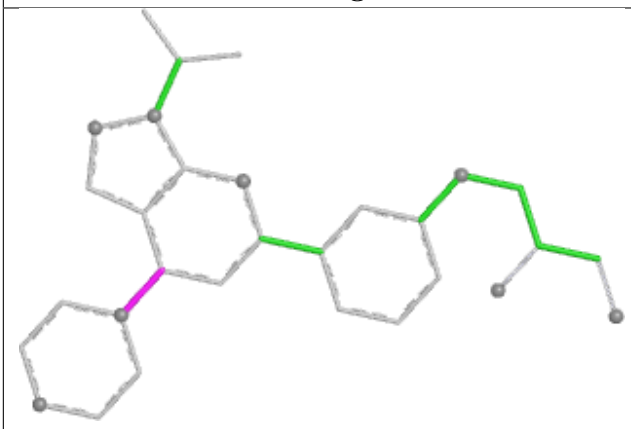
Ligand BW7 A 502



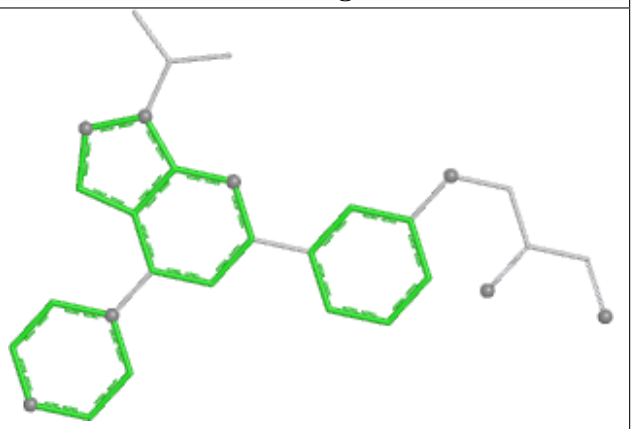
Bond lengths



Bond angles

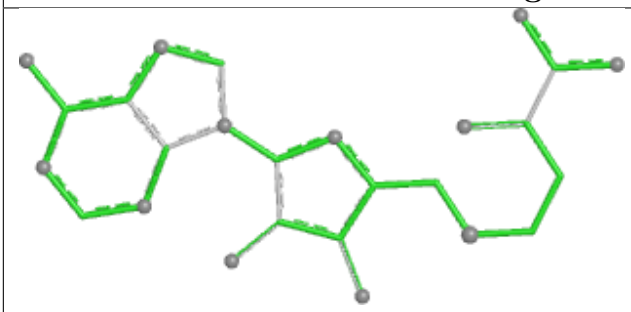


Torsions

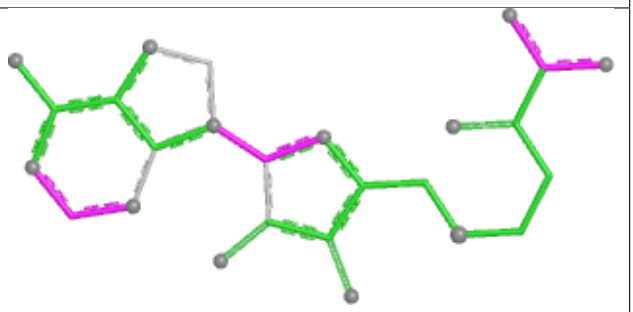


Rings

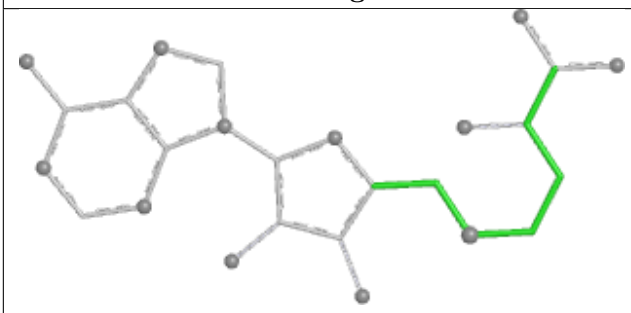
Ligand SAH B 501



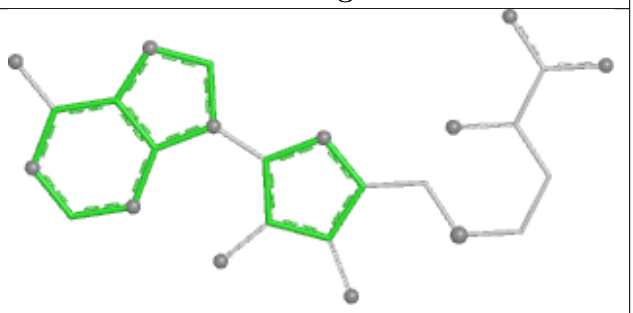
Bond lengths



Bond angles

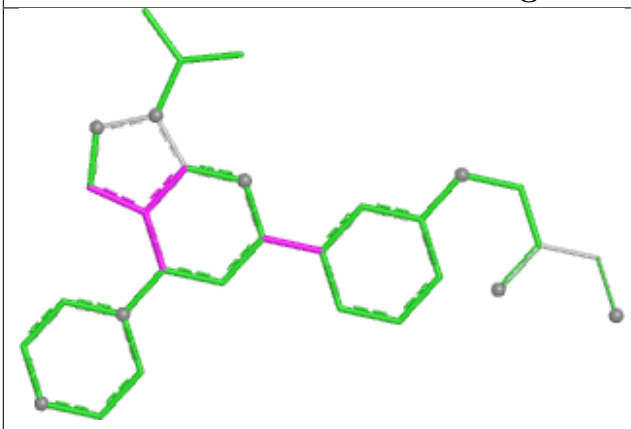


Torsions

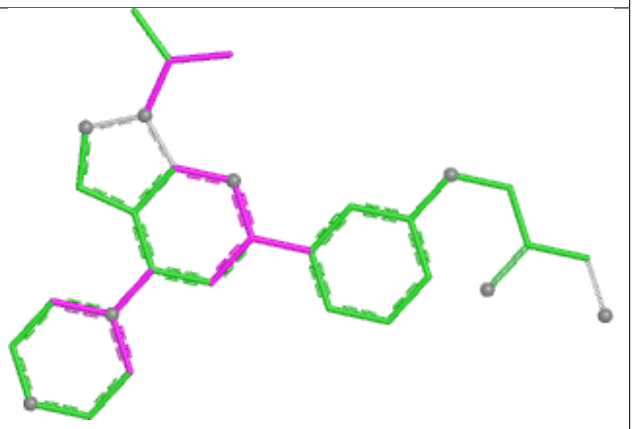


Rings

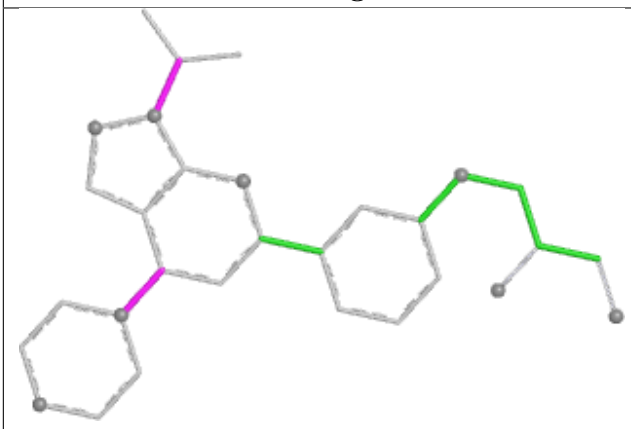
Ligand BW7 D 502



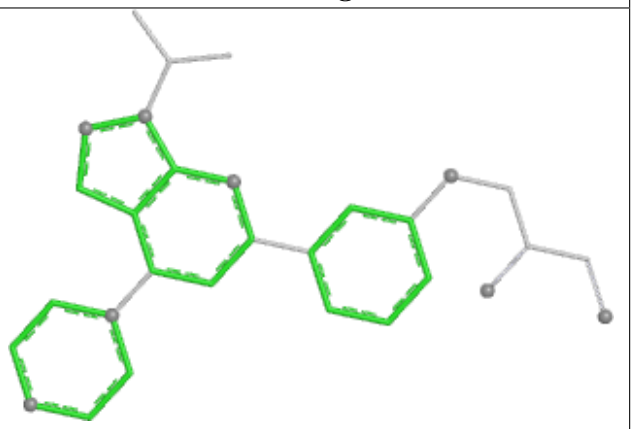
Bond lengths



Bond angles

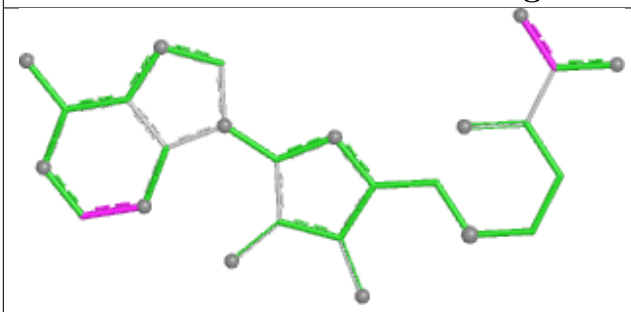


Torsions

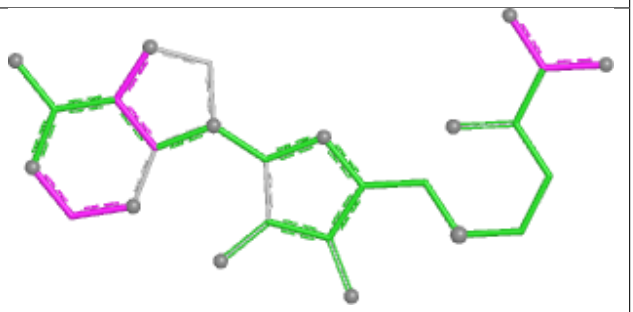


Rings

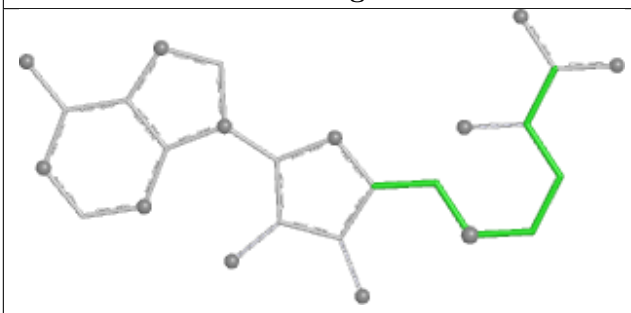
Ligand SAH D 501



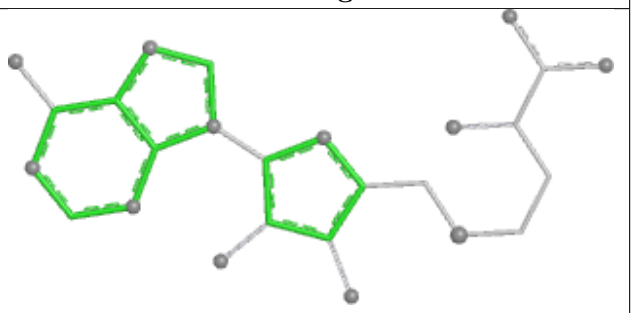
Bond lengths



Bond angles

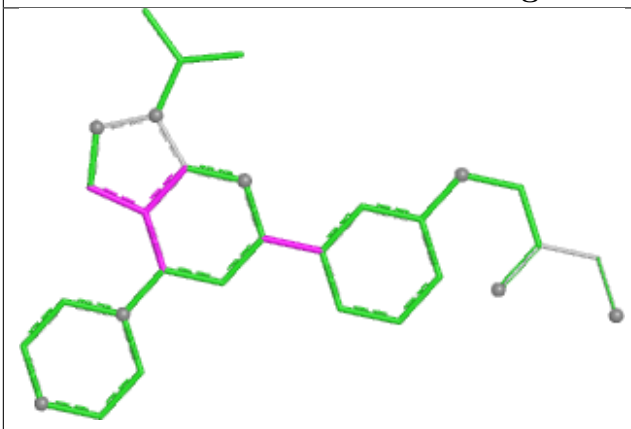


Torsions

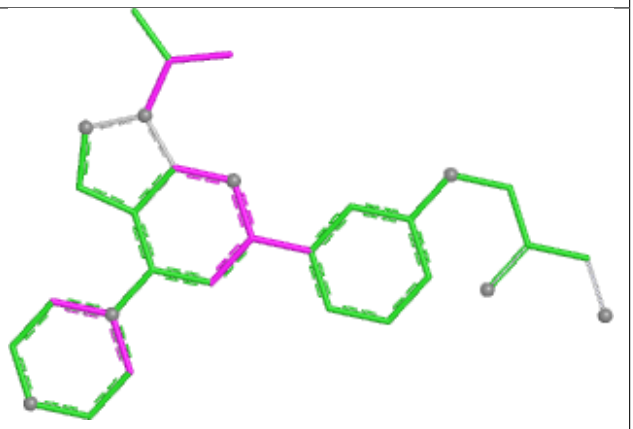


Rings

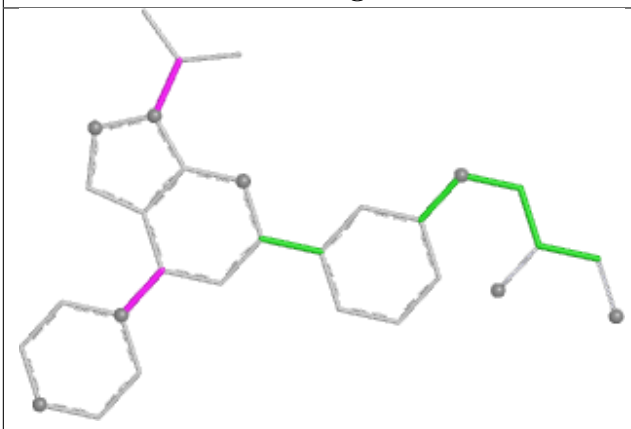
Ligand BW7 B 502



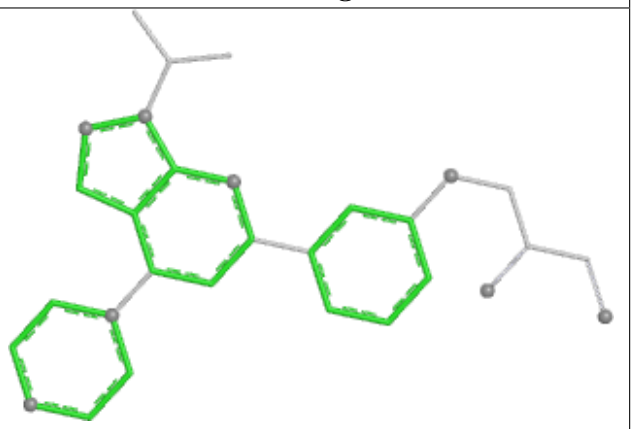
Bond lengths



Bond angles

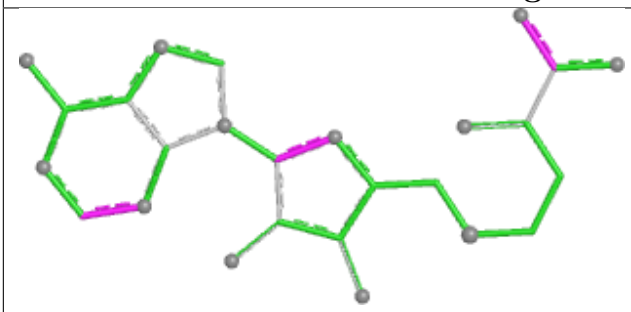


Torsions

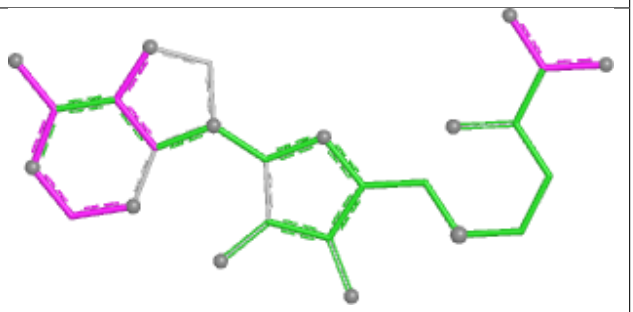


Rings

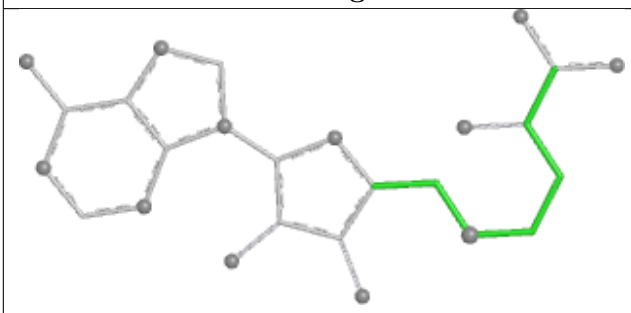
Ligand SAH C 501



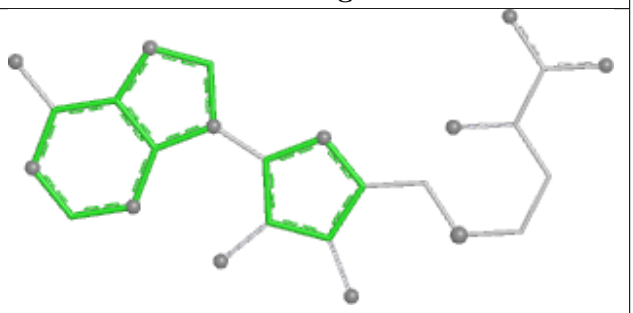
Bond lengths



Bond angles



Torsions



Rings

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	344/349 (98%)	-0.35	0	100	100	17, 26, 37, 62	0
1	B	344/349 (98%)	-0.30	1 (0%)	94	93	15, 24, 37, 67	0
1	C	344/349 (98%)	-0.31	2 (0%)	89	88	23, 31, 43, 74	0
1	D	344/349 (98%)	-0.31	3 (0%)	84	83	22, 32, 45, 84	0
All	All	1376/1396 (98%)	-0.32	6 (0%)	92	92	15, 29, 42, 84	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	134	ARG	3.9
1	C	134	ARG	2.8
1	B	477	THR	2.2
1	D	476	TYR	2.2
1	C	314	TYR	2.2
1	D	477	THR	2.0

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

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,

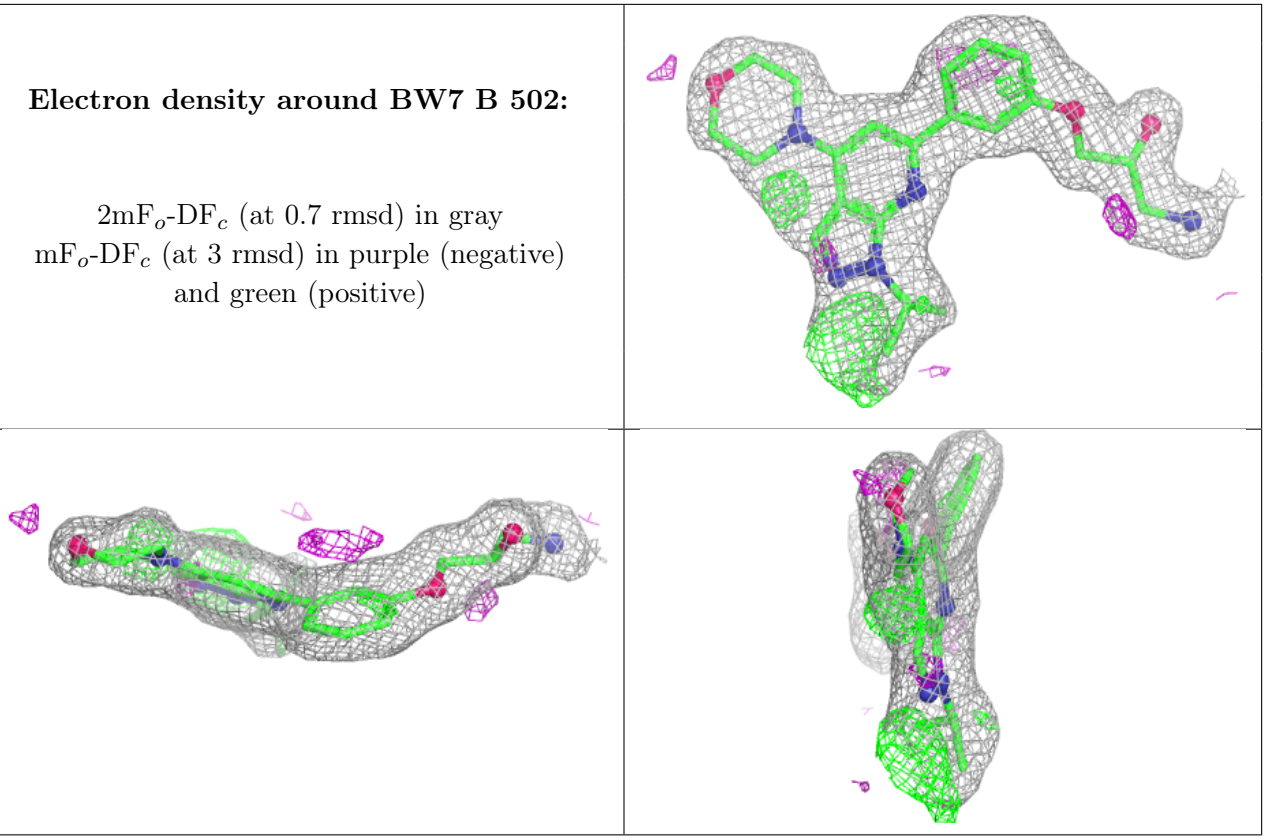
median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	GOL	A	506	6/6	0.76	0.32	58,61,62,66	0
4	GOL	D	504	6/6	0.77	0.24	49,52,54,59	0
4	GOL	A	505	6/6	0.79	0.27	53,57,60,60	0
4	GOL	D	503	6/6	0.82	0.23	49,51,52,54	0
4	GOL	D	505	6/6	0.82	0.27	68,72,73,73	0
4	GOL	A	504	6/6	0.86	0.14	43,43,45,50	0
4	GOL	C	503	6/6	0.87	0.17	45,48,49,51	0
4	GOL	B	504	6/6	0.87	0.19	30,38,43,44	0
4	GOL	A	503	6/6	0.90	0.16	43,45,46,50	0
4	GOL	B	503	6/6	0.91	0.13	47,49,52,53	0
4	GOL	C	504	6/6	0.92	0.14	48,53,54,56	0
3	BW7	B	502	30/30	0.92	0.14	21,29,39,39	0
3	BW7	D	502	30/30	0.93	0.13	27,36,46,48	0
3	BW7	C	502	30/30	0.94	0.12	27,35,44,46	0
3	BW7	A	502	30/30	0.95	0.11	17,26,35,40	0
2	SAH	C	501	26/26	0.96	0.09	25,26,28,28	0
2	SAH	B	501	26/26	0.97	0.09	18,20,21,22	0
2	SAH	D	501	26/26	0.98	0.08	27,29,32,32	0
2	SAH	A	501	26/26	0.98	0.09	16,19,19,21	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.

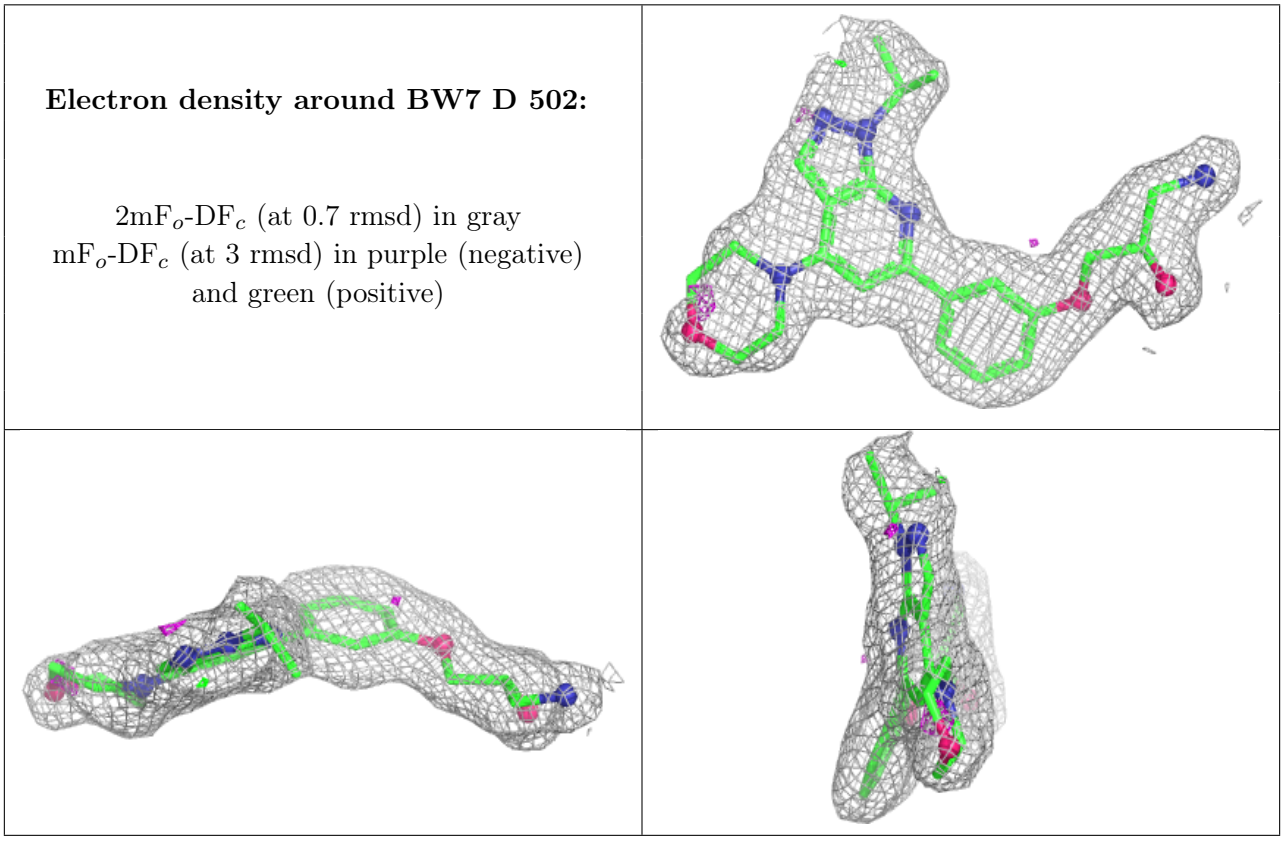
Electron density around BW7 B 502:

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



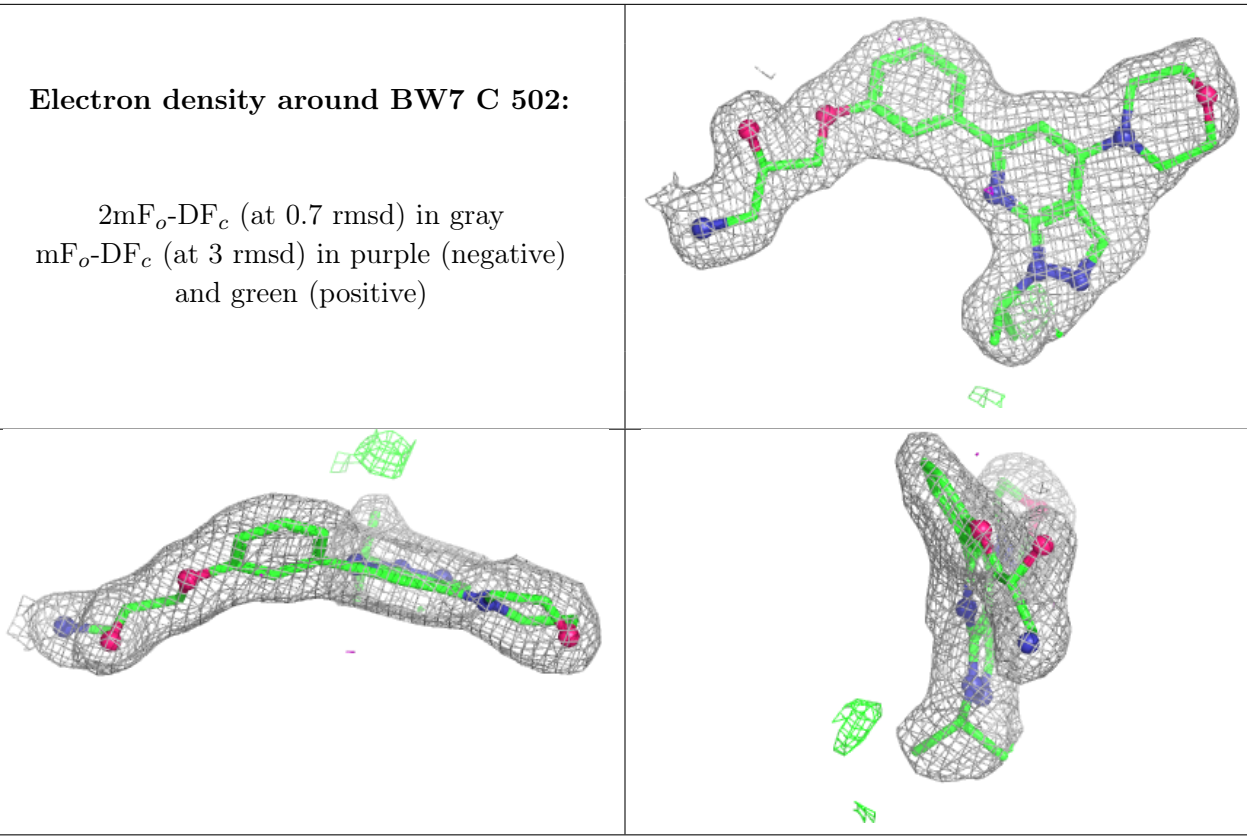
Electron density around BW7 D 502:

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



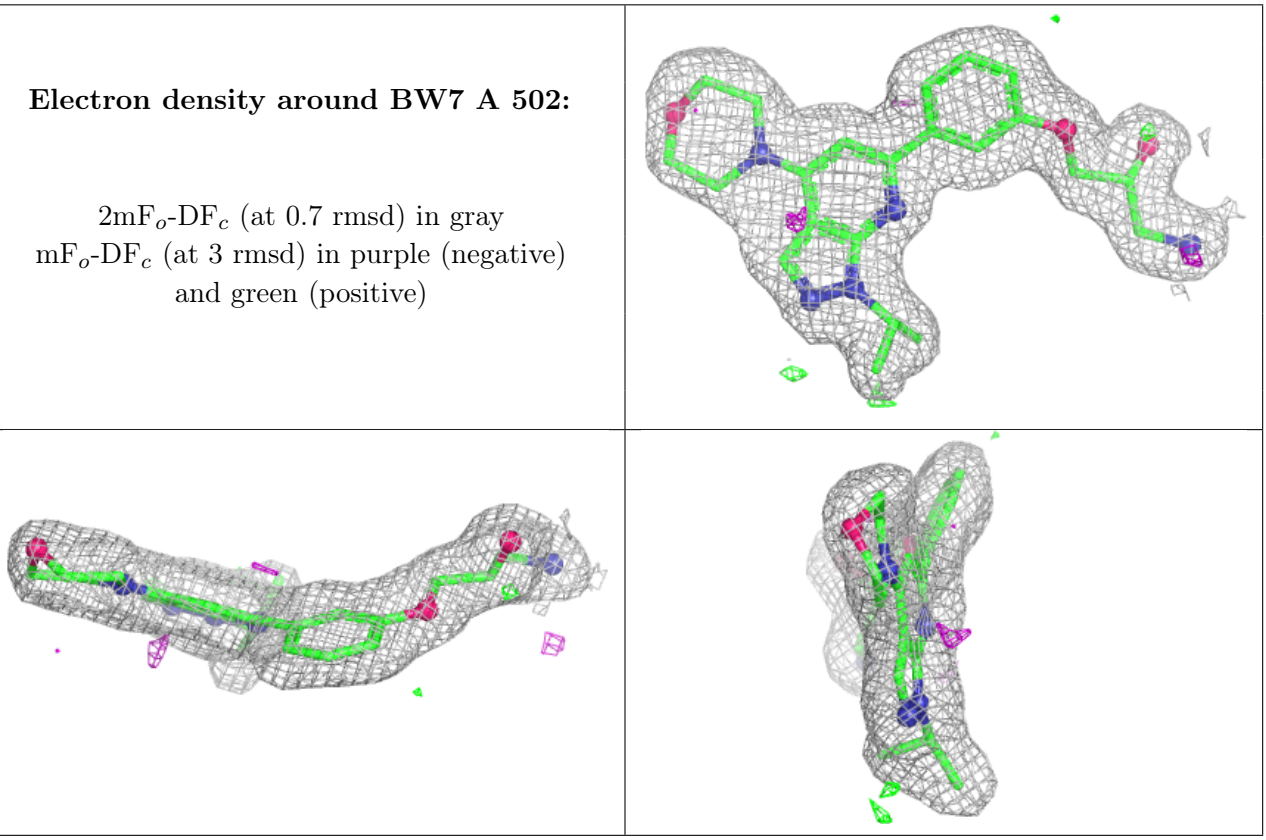
Electron density around BW7 C 502:

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



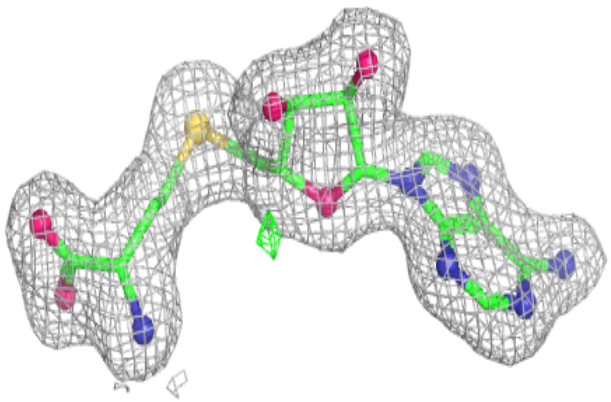
Electron density around BW7 A 502:

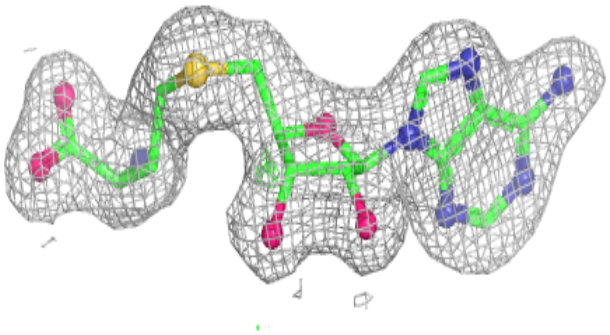
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

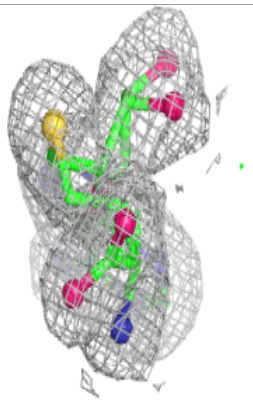


Electron density around SAH C 501:

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

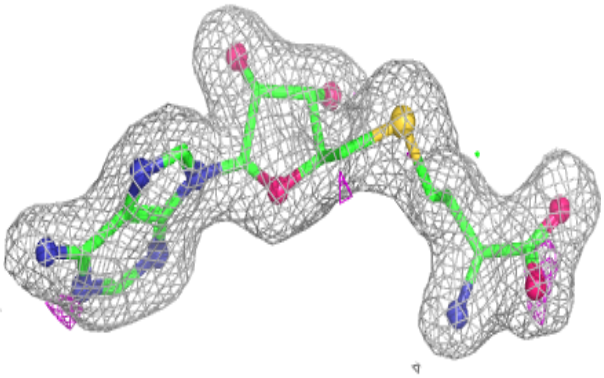


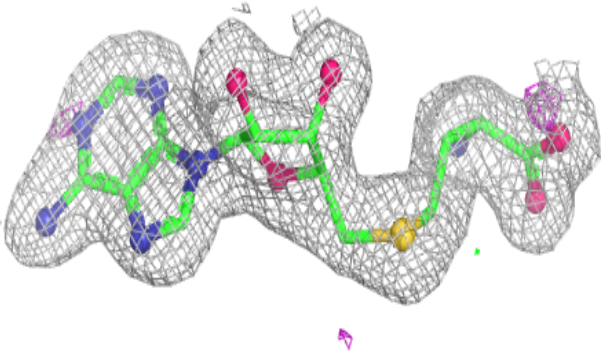


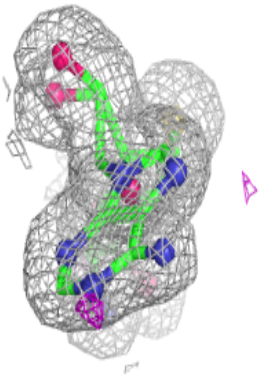


Electron density around SAH B 501:

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

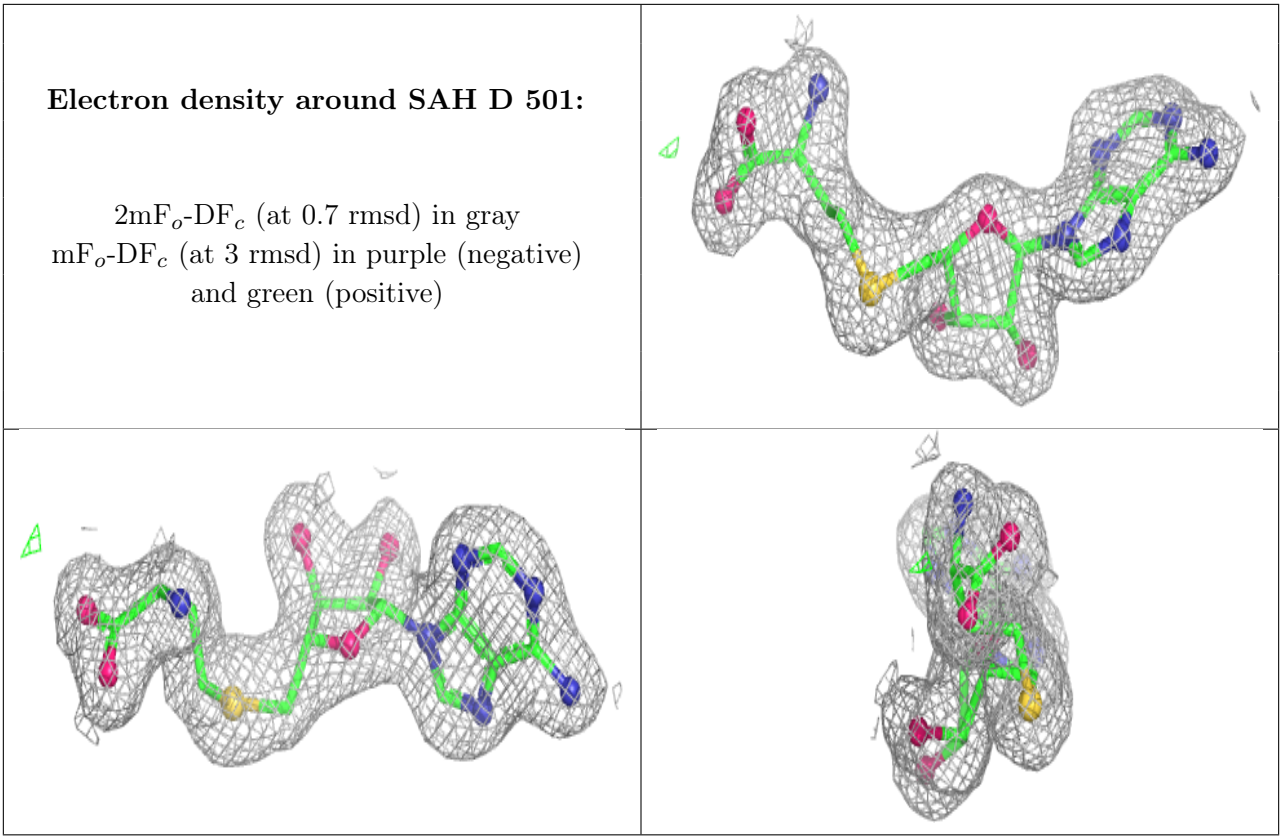






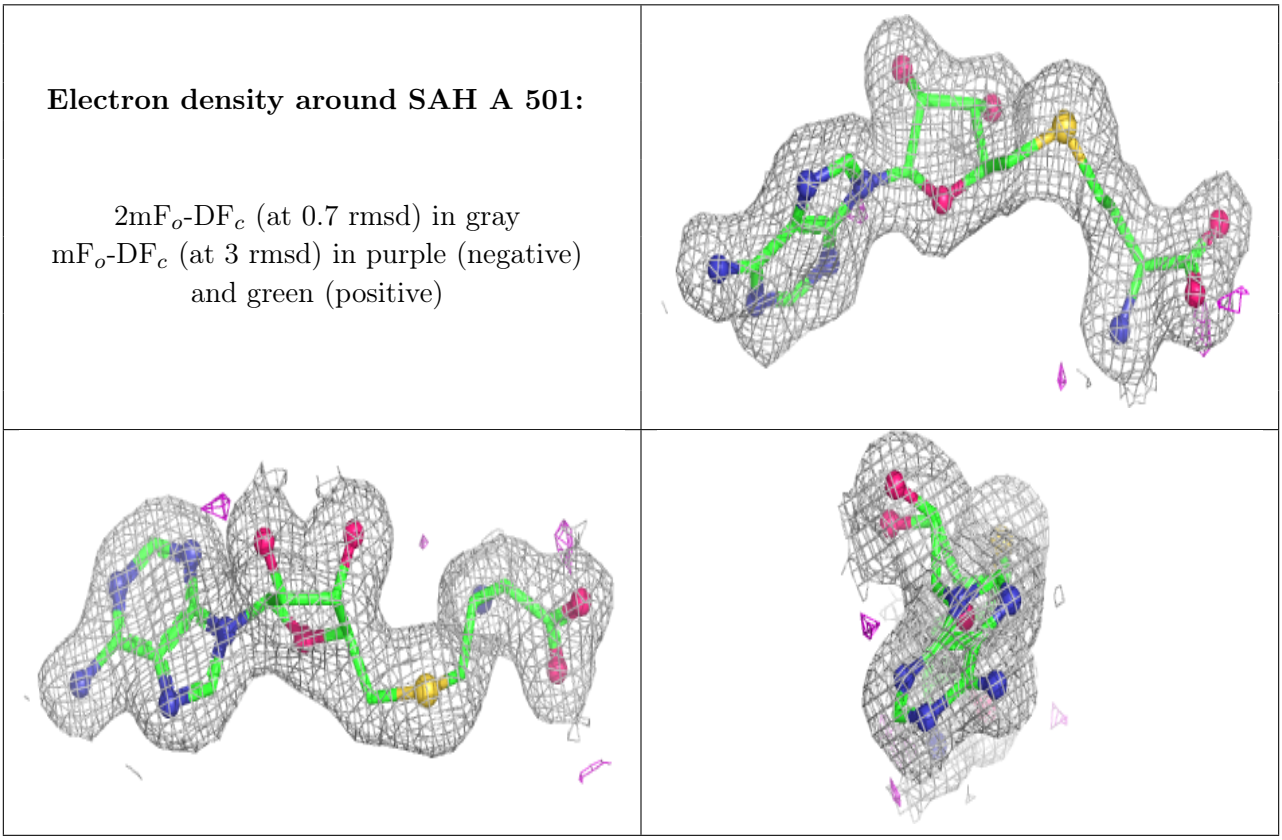
Electron density around SAH D 501:

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



Electron density around SAH A 501:

$2mF_o$-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.