



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 2, 2020 – 03:02 pm BST

PDB ID : 5YRV
Title : Diol dehydratase, AdoCbl/1,2-propanediol, anaerobically-prepared crystal
Authors : Shibata, N.
Deposited on : 2017-11-10
Resolution : 1.55 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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

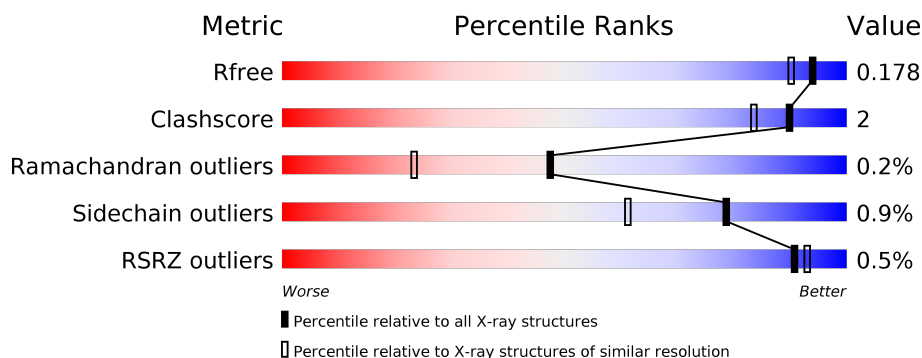
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.55 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	1483 (1.56-1.56)
Clashscore	141614	1529 (1.56-1.56)
Ramachandran outliers	138981	1498 (1.56-1.56)
Sidechain outliers	138945	1495 (1.56-1.56)
RSRZ outliers	127900	1465 (1.56-1.56)

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

Mol	Chain	Length	Quality of chain
1	A	554	<div><div></div><div>94%6%</div></div>
1	D	554	<div><div>%</div><div>96%.</div></div>
1	G	554	<div><div></div><div>97%.</div></div>
1	J	554	<div><div></div><div>95%5%</div></div>
2	B	200	<div><div>%</div><div>85%7%9%</div></div>
2	E	200	<div><div>%</div><div>88%.11%</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	200	
2	K	200	
3	C	137	
3	F	137	
3	I	137	
3	L	137	

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
8	CL	C	201	-	-	X	-

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 30656 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 Diol dehydrase alpha subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	551	Total	C	N	O	S	0	6	0
			4228	2639	728	831	30			
1	D	551	Total	C	N	O	S	0	5	0
			4222	2635	729	829	29			
1	G	554	Total	C	N	O	S	0	5	0
			4250	2651	731	838	30			
1	J	551	Total	C	N	O	S	0	6	0
			4231	2640	731	830	30			

- Molecule 2 is a protein called Diol dehydrase beta subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	182	Total	C	N	O	S	0	3	0
			1397	884	249	262	2			
2	E	178	Total	C	N	O	S	0	2	0
			1372	870	245	255	2			
2	H	179	Total	C	N	O	S	0	1	0
			1372	869	246	255	2			
2	K	182	Total	C	N	O	S	0	1	0
			1385	876	248	259	2			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	25	MET	-	expression tag	UNP Q59471
B	26	SER	-	expression tag	UNP Q59471
B	27	SER	-	expression tag	UNP Q59471
B	28	HIS	-	expression tag	UNP Q59471
B	29	HIS	-	expression tag	UNP Q59471
B	30	HIS	-	expression tag	UNP Q59471
B	31	HIS	-	expression tag	UNP Q59471
B	32	HIS	-	expression tag	UNP Q59471

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	33	HIS	-	expression tag	UNP Q59471
B	34	SER	-	expression tag	UNP Q59471
B	35	ALA	-	expression tag	UNP Q59471
B	36	ALA	-	expression tag	UNP Q59471
B	37	LEU	-	expression tag	UNP Q59471
B	38	GLU	-	expression tag	UNP Q59471
B	39	VAL	-	expression tag	UNP Q59471
B	40	LEU	-	expression tag	UNP Q59471
B	41	PHE	-	expression tag	UNP Q59471
B	42	GLN	-	expression tag	UNP Q59471
B	43	GLY	-	expression tag	UNP Q59471
B	44	PRO	-	expression tag	UNP Q59471
B	45	GLY	-	expression tag	UNP Q59471
E	25	MET	-	expression tag	UNP Q59471
E	26	SER	-	expression tag	UNP Q59471
E	27	SER	-	expression tag	UNP Q59471
E	28	HIS	-	expression tag	UNP Q59471
E	29	HIS	-	expression tag	UNP Q59471
E	30	HIS	-	expression tag	UNP Q59471
E	31	HIS	-	expression tag	UNP Q59471
E	32	HIS	-	expression tag	UNP Q59471
E	33	HIS	-	expression tag	UNP Q59471
E	34	SER	-	expression tag	UNP Q59471
E	35	ALA	-	expression tag	UNP Q59471
E	36	ALA	-	expression tag	UNP Q59471
E	37	LEU	-	expression tag	UNP Q59471
E	38	GLU	-	expression tag	UNP Q59471
E	39	VAL	-	expression tag	UNP Q59471
E	40	LEU	-	expression tag	UNP Q59471
E	41	PHE	-	expression tag	UNP Q59471
E	42	GLN	-	expression tag	UNP Q59471
E	43	GLY	-	expression tag	UNP Q59471
E	44	PRO	-	expression tag	UNP Q59471
E	45	GLY	-	expression tag	UNP Q59471
H	25	MET	-	expression tag	UNP Q59471
H	26	SER	-	expression tag	UNP Q59471
H	27	SER	-	expression tag	UNP Q59471
H	28	HIS	-	expression tag	UNP Q59471
H	29	HIS	-	expression tag	UNP Q59471
H	30	HIS	-	expression tag	UNP Q59471
H	31	HIS	-	expression tag	UNP Q59471
H	32	HIS	-	expression tag	UNP Q59471

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	33	HIS	-	expression tag	UNP Q59471
H	34	SER	-	expression tag	UNP Q59471
H	35	ALA	-	expression tag	UNP Q59471
H	36	ALA	-	expression tag	UNP Q59471
H	37	LEU	-	expression tag	UNP Q59471
H	38	GLU	-	expression tag	UNP Q59471
H	39	VAL	-	expression tag	UNP Q59471
H	40	LEU	-	expression tag	UNP Q59471
H	41	PHE	-	expression tag	UNP Q59471
H	42	GLN	-	expression tag	UNP Q59471
H	43	GLY	-	expression tag	UNP Q59471
H	44	PRO	-	expression tag	UNP Q59471
H	45	GLY	-	expression tag	UNP Q59471
K	25	MET	-	expression tag	UNP Q59471
K	26	SER	-	expression tag	UNP Q59471
K	27	SER	-	expression tag	UNP Q59471
K	28	HIS	-	expression tag	UNP Q59471
K	29	HIS	-	expression tag	UNP Q59471
K	30	HIS	-	expression tag	UNP Q59471
K	31	HIS	-	expression tag	UNP Q59471
K	32	HIS	-	expression tag	UNP Q59471
K	33	HIS	-	expression tag	UNP Q59471
K	34	SER	-	expression tag	UNP Q59471
K	35	ALA	-	expression tag	UNP Q59471
K	36	ALA	-	expression tag	UNP Q59471
K	37	LEU	-	expression tag	UNP Q59471
K	38	GLU	-	expression tag	UNP Q59471
K	39	VAL	-	expression tag	UNP Q59471
K	40	LEU	-	expression tag	UNP Q59471
K	41	PHE	-	expression tag	UNP Q59471
K	42	GLN	-	expression tag	UNP Q59471
K	43	GLY	-	expression tag	UNP Q59471
K	44	PRO	-	expression tag	UNP Q59471
K	45	GLY	-	expression tag	UNP Q59471

- Molecule 3 is a protein called Diol dehydrase gamma subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	136	Total	C	N	O	S	0	5	0
			1111	694	195	218	4			
3	F	136	Total	C	N	O	S	0	2	0
			1097	684	195	215	3			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	136	Total	C	N	O	S	0	4	0
			1108	692	198	214	4			
3	L	136	Total	C	N	O	S	0	4	0
			1109	693	195	217	4			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	37	MET	-	expression tag	UNP Q59472
F	37	MET	-	expression tag	UNP Q59472
I	37	MET	-	expression tag	UNP Q59472
L	37	MET	-	expression tag	UNP Q59472

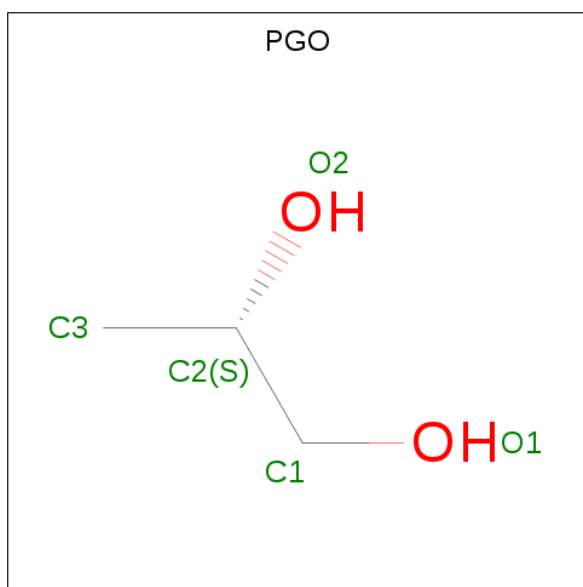
- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	G	1	Total	Ca	0	0
			1	1		
4	J	1	Total	Ca	0	0
			1	1		
4	A	1	Total	Ca	0	0
			1	1		
4	D	1	Total	Ca	0	0
			1	1		

- Molecule 5 is POTASSIUM ION (three-letter code: K) (formula: K).

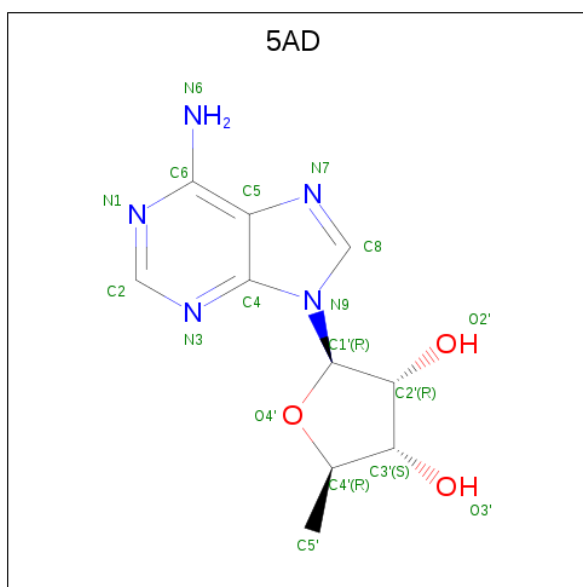
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	G	2	Total	K	0	0
			2	2		
5	J	2	Total	K	0	0
			2	2		
5	A	2	Total	K	0	0
			2	2		
5	D	2	Total	K	0	0
			2	2		

- Molecule 6 is S-1,2-PROPANEDIOL (three-letter code: PGO) (formula: C₃H₈O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			5	3	2		
6	D	1	Total	C	O	0	0
			5	3	2		
6	G	1	Total	C	O	0	0
			5	3	2		

- Molecule 7 is 5'-DEOXYADENOSINE (three-letter code: 5AD) (formula: $C_{10}H_{13}N_5O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	1	Total	C	N	O	0	0
			18	10	5	3		

Continued on next page...

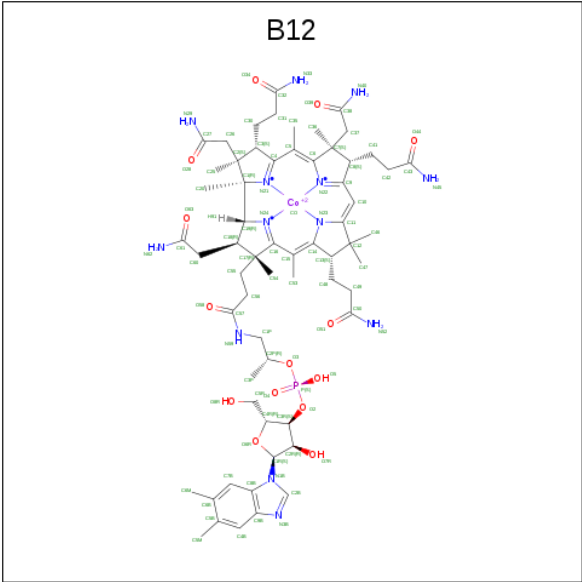
Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	D	1	Total	C	N	O	0	0
			18	10	5	3		
7	G	1	Total	C	N	O	0	0
			18	10	5	3		
7	J	1	Total	C	N	O	0	0
			18	10	5	3		

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	G	2	Total	Cl	0	0
			2	2		
8	J	2	Total	Cl	0	0
			2	2		
8	D	2	Total	Cl	0	0
			2	2		
8	I	1	Total	Cl	0	0
			1	1		
8	C	1	Total	Cl	0	0
			1	1		
8	A	2	Total	Cl	0	0
			2	2		
8	L	1	Total	Cl	0	0
			1	1		
8	F	1	Total	Cl	0	0
			1	1		

- Molecule 9 is COBALAMIN (three-letter code: B12) (formula: C₆₂H₈₉CoN₁₃O₁₄P).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
9	B	1	Total	C	Co	N	O	P	0	0
			91	62	1	13	14	1		
9	E	1	Total	C	Co	N	O	P	0	0
			91	62	1	13	14	1		
9	H	1	Total	C	Co	N	O	P	0	0
			91	62	1	13	14	1		
9	K	1	Total	C	Co	N	O	P	0	0
			91	62	1	13	14	1		

- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	509	Total	O	0	0
			509	509		
10	B	192	Total	O	0	0
			192	192		
10	C	194	Total	O	0	0
			194	194		
10	D	485	Total	O	0	0
			485	485		
10	E	158	Total	O	0	0
			158	158		
10	F	119	Total	O	0	0
			119	119		
10	G	491	Total	O	0	0
			491	491		
10	H	172	Total	O	0	0
			172	172		

Continued on next page...

Continued from previous page...

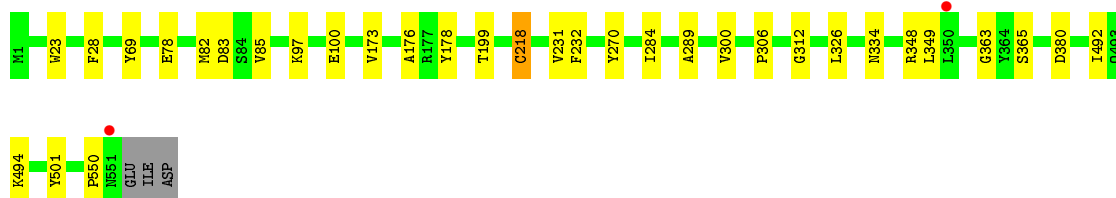
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	I	152	Total 152	O 152	0	0
10	J	478	Total 478	O 478	0	0
10	K	197	Total 197	O 197	0	0
10	L	152	Total 152	O 152	0	0

3 Residue-property plots [i](#)

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: Diol dehydrase alpha subunit

Chain A: 



- Molecule 1: Diol dehydrase alpha subunit

Chain D: 



- Molecule 1: Diol dehydrase alpha subunit

Chain G: 




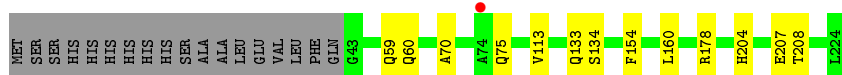
- Molecule 1: Diol dehydrase alpha subunit

Chain J: 

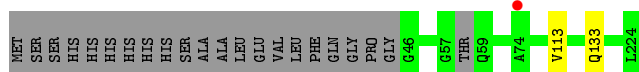
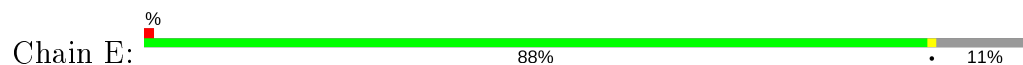


- Molecule 2: Diol dehydrase beta subunit

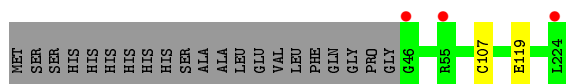
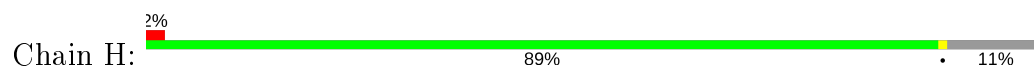
Chain B: 



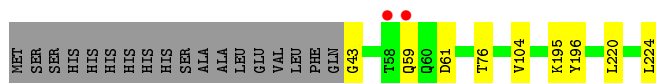
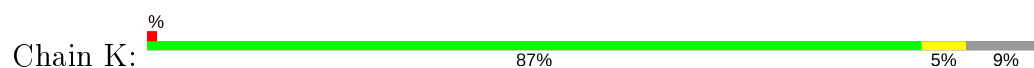
- Molecule 2: Diol dehydrase beta subunit



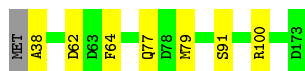
- Molecule 2: Diol dehydrase beta subunit



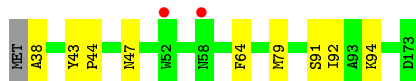
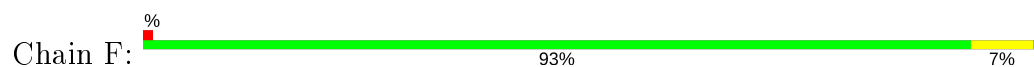
- Molecule 2: Diol dehydrase beta subunit



- Molecule 3: Diol dehydrase gamma subunit



- Molecule 3: Diol dehydrase gamma subunit



- Molecule 3: Diol dehydrase gamma subunit



- Molecule 3: Diol dehydrase gamma subunit



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	72.78Å 95.67Å 114.51Å 80.61° 87.77° 81.25°	Depositor
Resolution (Å)	44.00 – 1.55 46.77 – 1.55	Depositor EDS
% Data completeness (in resolution range)	93.7 (44.00-1.55) 96.2 (46.77-1.55)	Depositor EDS
R_{merge}	0.03	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.74 (at 1.55Å)	Xtriage
Refinement program	PHENIX (1.10.1 _2155: ???)	Depositor
R, R_{free}	0.155 , 0.181 0.152 , 0.178	Depositor DCC
R_{free} test set	4284 reflections (1.02%)	wwPDB-VP
Wilson B-factor (Å ²)	18.6	Xtriage
Anisotropy	0.351	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 46.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	30656	wwPDB-VP
Average B, all atoms (Å ²)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.36% of the height of the origin peak. No significant pseudotranslation is detected.*

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: CL, PGO, K, B12, CA, 5AD

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	1/4318 (0.0%)	0.57	0/5847
1	D	0.34	0/4309	0.57	0/5836
1	G	0.34	0/4337	0.56	0/5871
1	J	0.35	0/4321	0.57	0/5850
2	B	0.35	0/1429	0.53	0/1933
2	E	0.33	0/1399	0.52	0/1889
2	H	0.30	0/1397	0.51	0/1891
2	K	0.32	0/1411	0.54	0/1910
3	C	0.32	0/1141	0.53	0/1541
3	F	0.29	0/1118	0.47	0/1511
3	I	0.30	0/1135	0.49	0/1532
3	L	0.29	0/1136	0.50	0/1534
All	All	0.34	1/27451 (0.0%)	0.55	0/37145

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	218	CYS	CB-SG	-6.75	1.70	1.82

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	4228	0	4179	16	0
1	D	4222	0	4172	10	0
1	G	4250	0	4193	8	0
1	J	4231	0	4183	12	0
2	B	1397	0	1451	7	0
2	E	1372	0	1430	1	0
2	H	1372	0	1427	1	0
2	K	1385	0	1437	5	0
3	C	1111	0	1126	4	0
3	F	1097	0	1106	5	0
3	I	1108	0	1129	3	0
3	L	1109	0	1123	4	0
4	A	1	0	0	0	0
4	D	1	0	0	0	0
4	G	1	0	0	0	0
4	J	1	0	0	0	0
5	A	2	0	0	0	0
5	D	2	0	0	0	0
5	G	2	0	0	0	0
5	J	2	0	0	0	0
6	A	5	0	6	2	0
6	D	5	0	6	2	0
6	G	5	0	6	2	0
7	A	18	0	13	2	0
7	D	18	0	13	3	0
7	G	18	0	13	2	0
7	J	18	0	13	0	0
8	A	2	0	0	0	0
8	C	1	0	0	2	0
8	D	2	0	0	0	0
8	F	1	0	0	0	0
8	G	2	0	0	0	0
8	I	1	0	0	0	0
8	J	2	0	0	0	0
8	L	1	0	0	1	0
9	B	91	0	88	4	0
9	E	91	0	88	7	0
9	H	91	0	88	6	0
9	K	91	0	88	4	0
10	A	509	0	0	1	0
10	B	192	0	0	1	0
10	C	194	0	0	2	0
10	D	485	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	E	158	0	0	0	0
10	F	119	0	0	0	0
10	G	491	0	0	1	0
10	H	172	0	0	0	0
10	I	152	0	0	2	0
10	J	478	0	0	0	0
10	K	197	0	0	0	0
10	L	152	0	0	0	0
All	All	30656	0	27378	98	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 98 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)
6:A:604:PGO:H12	7:A:605:5AD:H5'1	1.50	0.91
6:G:604:PGO:H12	7:G:605:5AD:H5'1	1.57	0.87
6:D:604:PGO:H12	7:D:605:5AD:H5'1	1.65	0.78
1:D:551:ASN:ND2	10:D:702:HOH:O	2.19	0.74
9:H:1601:B12:H552	9:H:1601:B12:H531	1.76	0.66

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	555/554 (100%)	537 (97%)	16 (3%)	2 (0%)	34	14
1	D	554/554 (100%)	538 (97%)	14 (2%)	2 (0%)	34	14
1	G	557/554 (100%)	539 (97%)	16 (3%)	2 (0%)	34	14

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	J	555/554 (100%)	537 (97%)	16 (3%)	2 (0%)	34	14
2	B	183/200 (92%)	177 (97%)	6 (3%)	0	100	100
2	E	176/200 (88%)	174 (99%)	2 (1%)	0	100	100
2	H	178/200 (89%)	173 (97%)	5 (3%)	0	100	100
2	K	181/200 (90%)	177 (98%)	4 (2%)	0	100	100
3	C	139/137 (102%)	138 (99%)	1 (1%)	0	100	100
3	F	136/137 (99%)	135 (99%)	1 (1%)	0	100	100
3	I	138/137 (101%)	136 (99%)	2 (1%)	0	100	100
3	L	138/137 (101%)	137 (99%)	1 (1%)	0	100	100
All	All	3490/3564 (98%)	3398 (97%)	84 (2%)	8 (0%)	47	23

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	363	GLY
1	A	300	VAL
1	D	300	VAL
1	D	363	GLY
1	G	300	VAL

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	456/453 (101%)	451 (99%)	5 (1%)	73	53
1	D	455/453 (100%)	449 (99%)	6 (1%)	69	44
1	G	458/453 (101%)	452 (99%)	6 (1%)	69	44
1	J	456/453 (101%)	450 (99%)	6 (1%)	69	44
2	B	151/164 (92%)	149 (99%)	2 (1%)	69	44
2	E	148/164 (90%)	148 (100%)	0	100	100
2	H	148/164 (90%)	148 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	K	149/164 (91%)	149 (100%)	0	100	100
3	C	120/116 (103%)	118 (98%)	2 (2%)	60	32
3	F	117/116 (101%)	117 (100%)	0	100	100
3	I	119/116 (103%)	119 (100%)	0	100	100
3	L	119/116 (103%)	119 (100%)	0	100	100
All	All	2896/2932 (99%)	2869 (99%)	27 (1%)	78	61

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	348	ARG
1	G	28	PHE
1	J	348	ARG
1	D	492	ILE
1	A	501	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	74	ASN
3	L	47	ASN
2	H	56	GLN
2	B	60	GLN
2	E	218	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 carbohydrates in this entry.

5.6 Ligand geometry

Of 35 ligands modelled in this entry, 24 are monoatomic - leaving 11 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
7	5AD	D	605	-	17,20,20	1.10	1 (5%)	15,30,30	2.10	2 (13%)
9	B12	K	1601	-	80,101,101	1.18	5 (6%)	101,166,166	1.30	13 (12%)
7	5AD	J	604	-	17,20,20	1.04	1 (5%)	15,30,30	1.82	4 (26%)
9	B12	B	1601	-	80,101,101	1.18	6 (7%)	101,166,166	1.29	13 (12%)
7	5AD	A	605	-	17,20,20	1.03	2 (11%)	15,30,30	2.07	4 (26%)
7	5AD	G	605	-	17,20,20	1.02	1 (5%)	15,30,30	1.85	3 (20%)
9	B12	E	1601	-	80,101,101	1.21	6 (7%)	101,166,166	1.32	14 (13%)
6	PGO	A	604	4	3,4,4	0.22	0	1,4,4	0.14	0
6	PGO	G	604	4	3,4,4	0.22	0	1,4,4	0.01	0
6	PGO	D	604	4	3,4,4	0.20	0	1,4,4	0.10	0
9	B12	H	1601	-	80,101,101	1.20	5 (6%)	101,166,166	1.35	14 (13%)

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
7	5AD	D	605	-	-	0/0/20/20	0/3/3/3
9	B12	K	1601	-	-	5/51/223/223	0/3/11/11
7	5AD	J	604	-	-	0/0/20/20	0/3/3/3
9	B12	B	1601	-	-	4/51/223/223	0/3/11/11
7	5AD	A	605	-	-	0/0/20/20	0/3/3/3
7	5AD	G	605	-	-	0/0/20/20	0/3/3/3
9	B12	E	1601	-	-	6/51/223/223	0/3/11/11
6	PGO	A	604	4	-	0/2/2/2	-
6	PGO	G	604	4	-	0/2/2/2	-
6	PGO	D	604	4	-	0/2/2/2	-
9	B12	H	1601	-	-	7/51/223/223	0/3/11/11

The worst 5 of 27 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	E	1601	B12	C11-C10	-4.98	1.32	1.40
9	K	1601	B12	C11-C10	-4.95	1.32	1.40
9	H	1601	B12	C11-C10	-4.87	1.32	1.40
9	B	1601	B12	C8B-C9B	4.55	1.49	1.40
9	E	1601	B12	C8B-C9B	4.42	1.49	1.40

The worst 5 of 67 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	D	605	5AD	C5'-C4'-C3'	-6.36	109.02	115.70
7	A	605	5AD	C5'-C4'-C3'	-5.86	109.55	115.70
7	G	605	5AD	C5'-C4'-C3'	-5.23	110.21	115.70
9	H	1601	B12	C20-C1-C19	-4.93	104.61	109.36
9	E	1601	B12	C20-C1-C19	-4.45	105.07	109.36

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	H	1601	B12	C1-C2-C26-C27
9	H	1601	B12	C25-C2-C26-C27
9	B	1601	B12	C1-C2-C26-C27
9	B	1601	B12	C25-C2-C26-C27
9	K	1601	B12	C1-C2-C26-C27

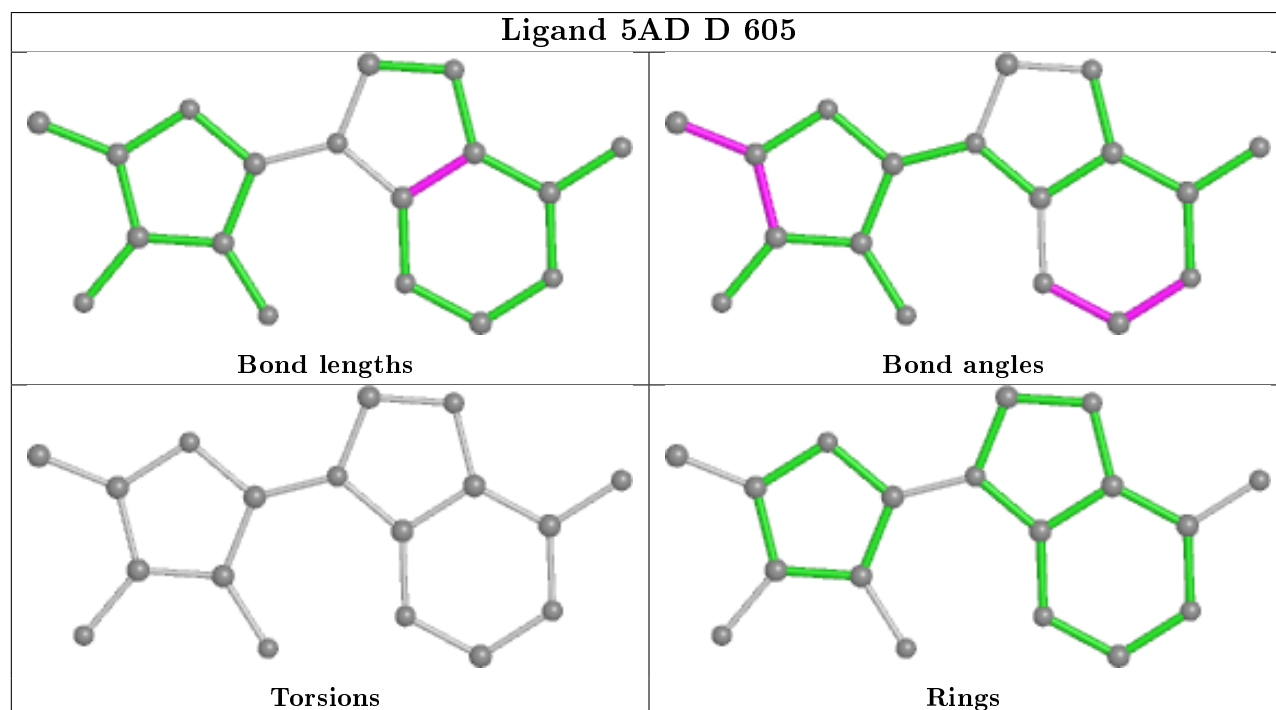
There are no ring outliers.

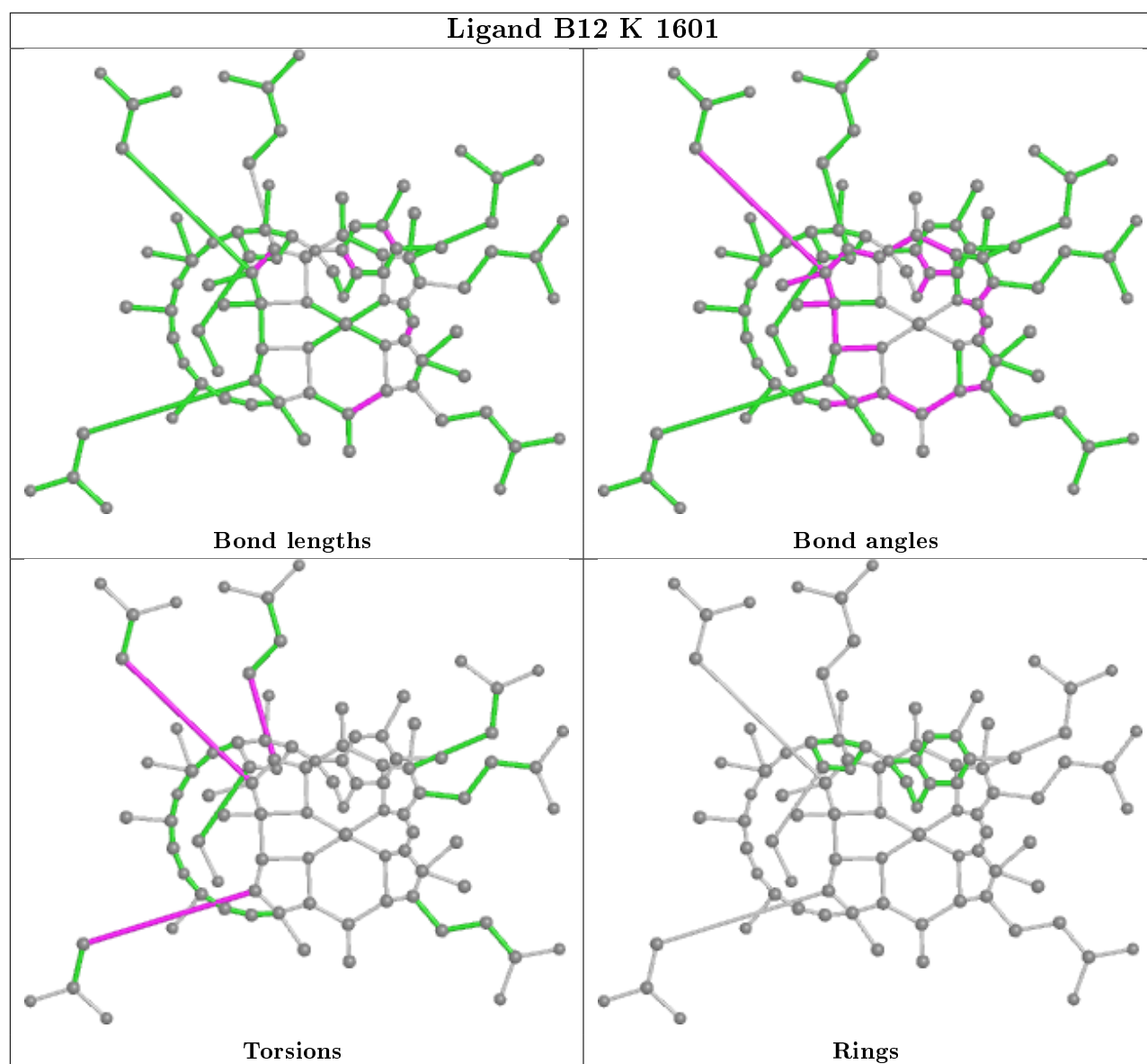
10 monomers are involved in 27 short contacts:

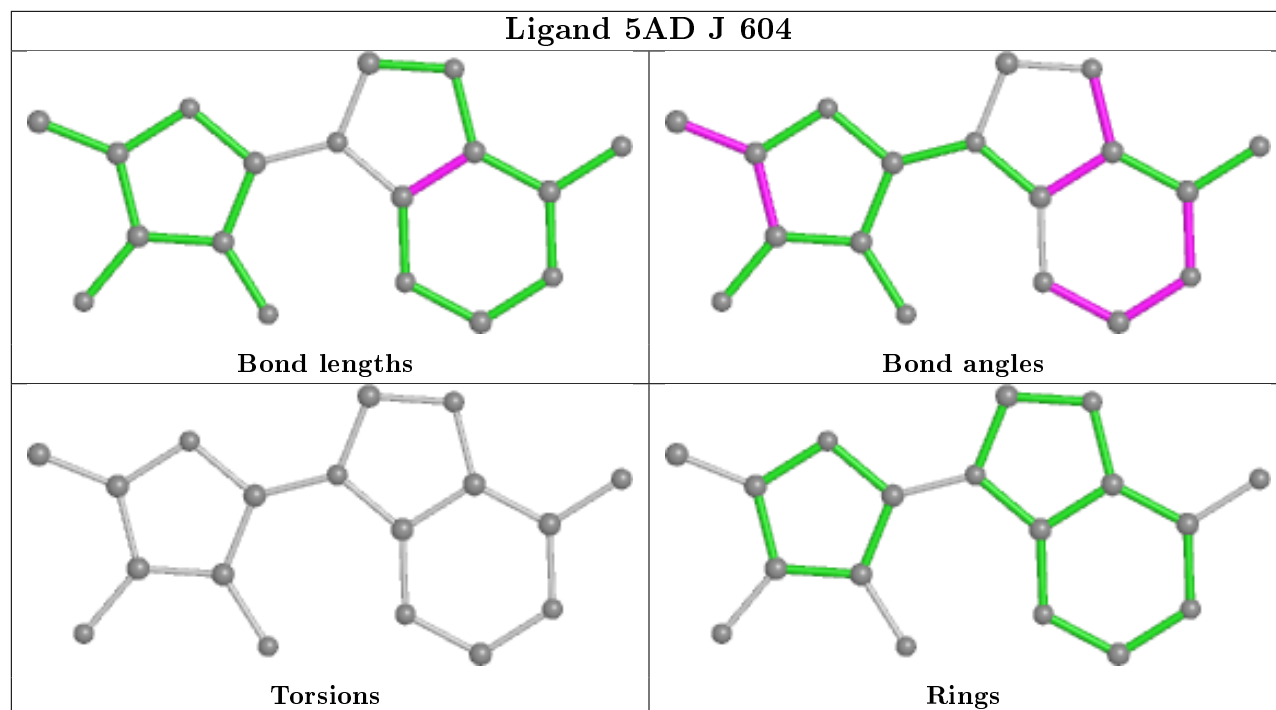
Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	D	605	5AD	3	0
9	K	1601	B12	4	0
9	B	1601	B12	4	0
7	A	605	5AD	2	0
7	G	605	5AD	2	0
9	E	1601	B12	7	0
6	A	604	PGO	2	0
6	G	604	PGO	2	0
6	D	604	PGO	2	0
9	H	1601	B12	6	0

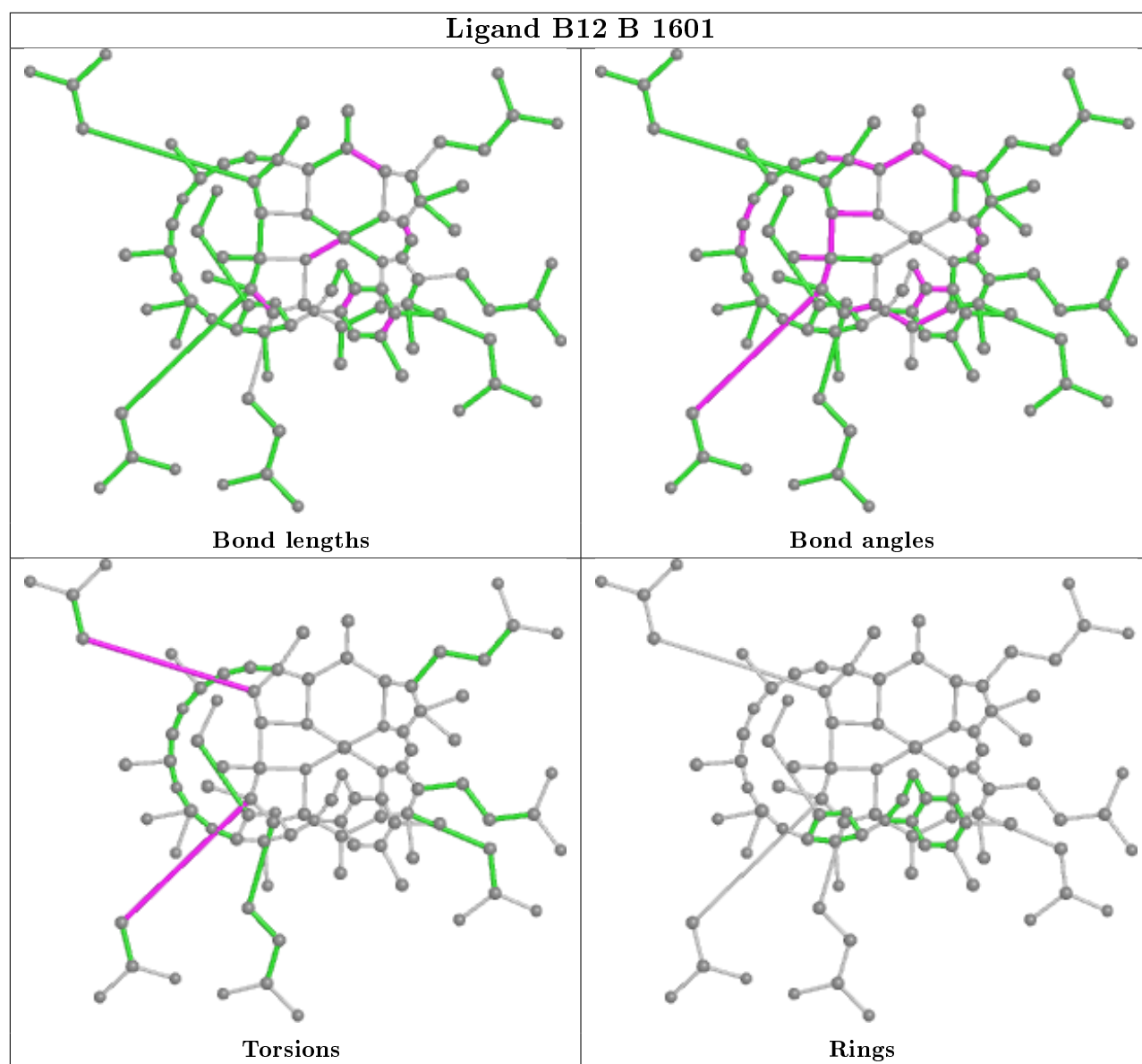
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths,

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 than 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.

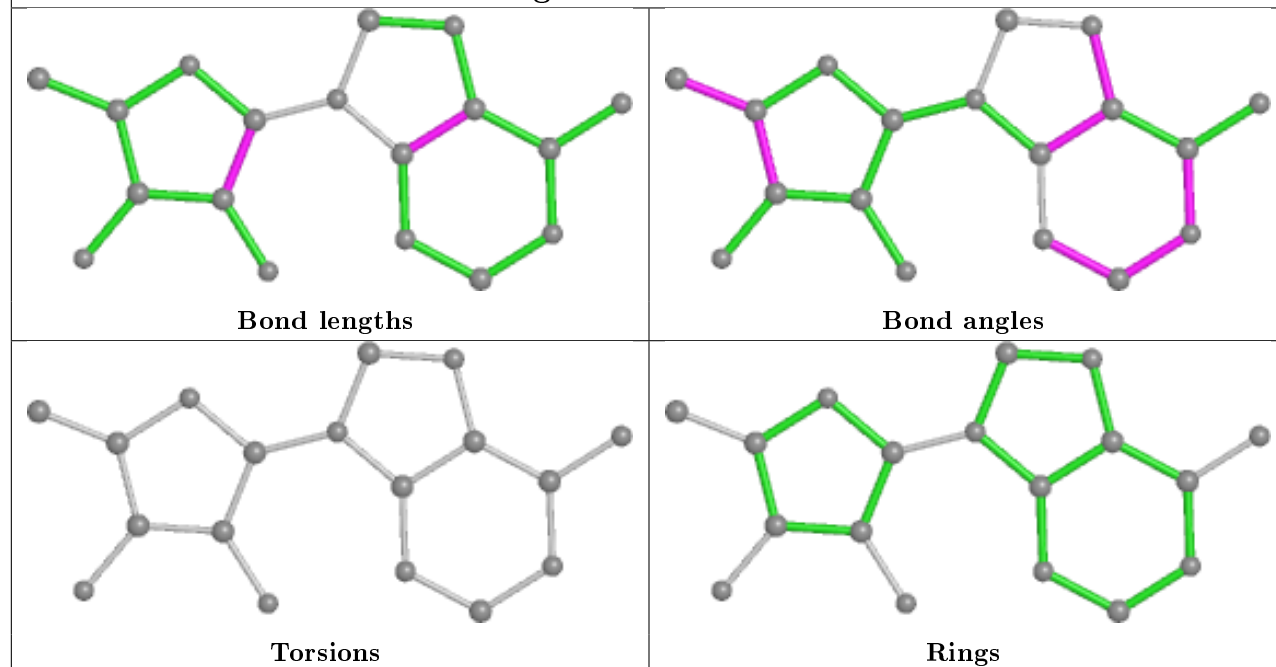




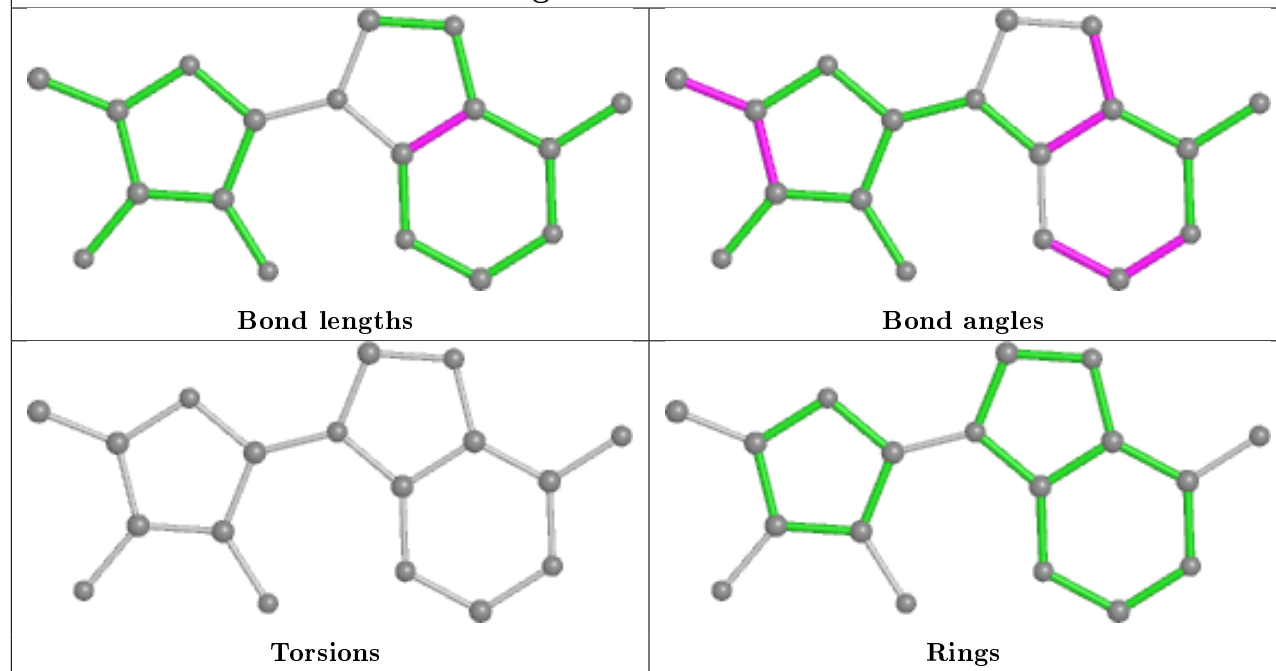




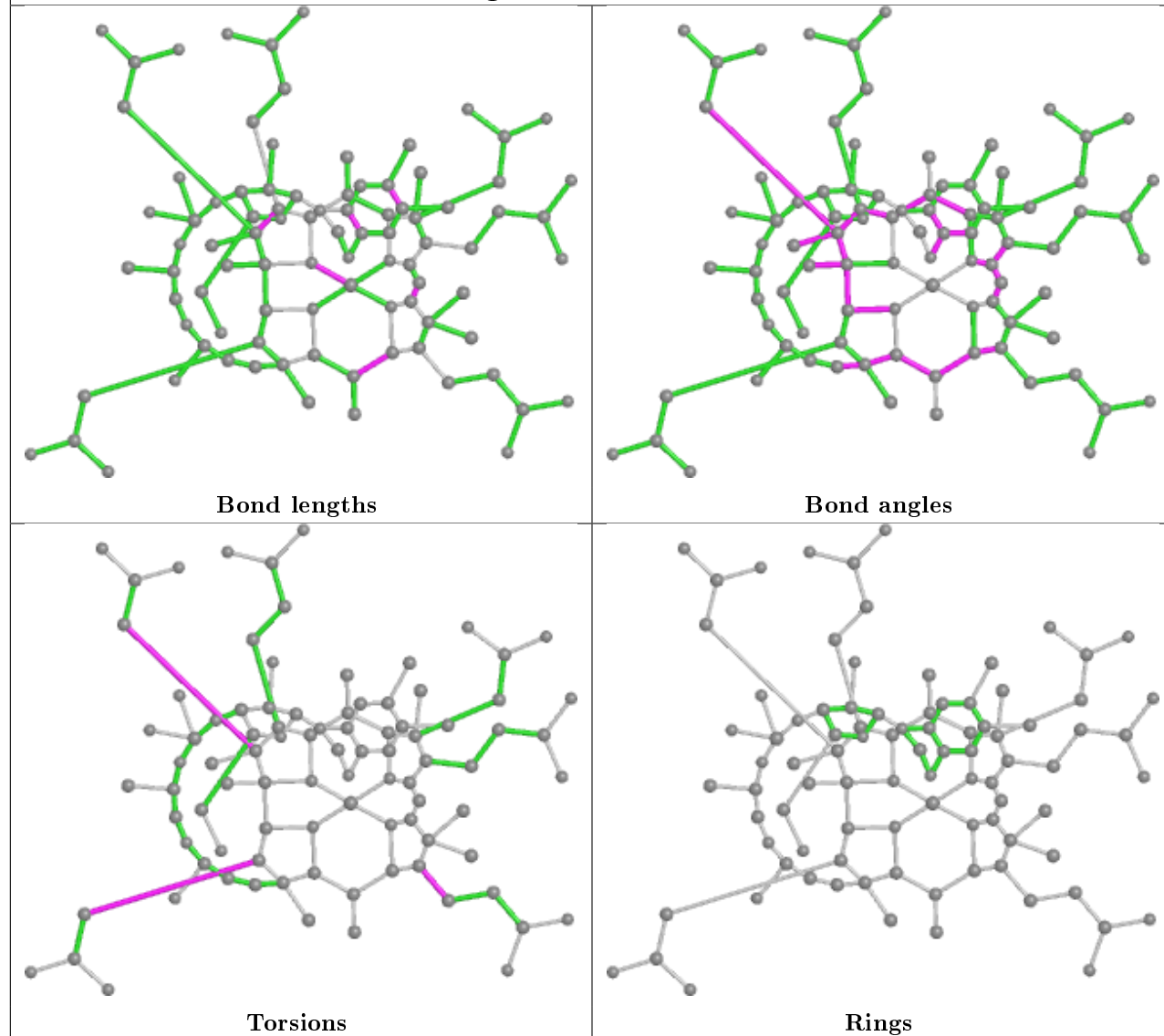
Ligand 5AD A 605

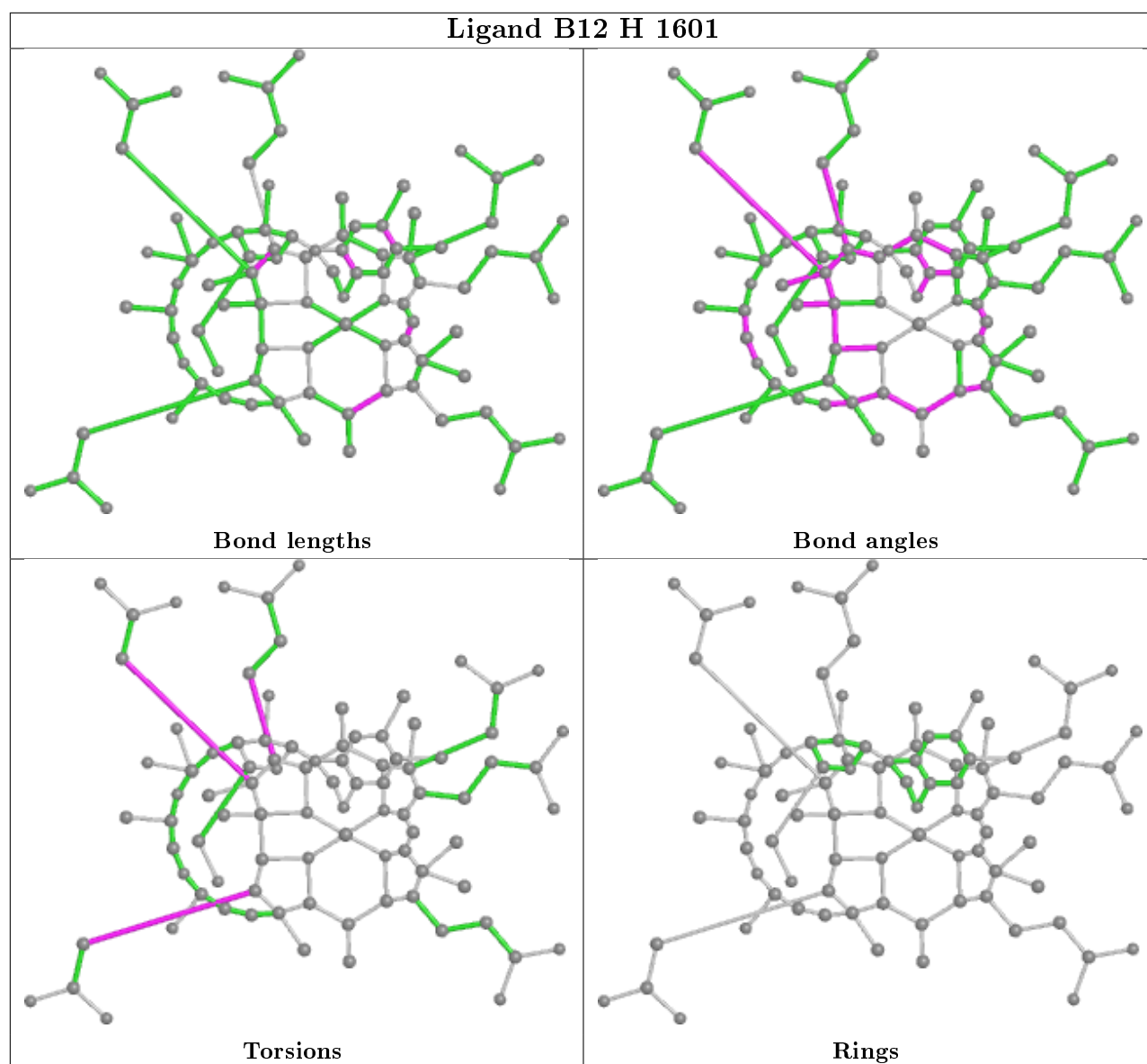


Ligand 5AD G 605



Ligand B12 E 1601





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, 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	551/554 (99%)	-0.45	2 (0%) 92 94	12, 18, 31, 71	0
1	D	551/554 (99%)	-0.40	4 (0%) 87 90	14, 21, 36, 58	0
1	G	554/554 (100%)	-0.40	2 (0%) 92 94	15, 20, 32, 49	0
1	J	551/554 (99%)	-0.42	1 (0%) 95 95	13, 20, 33, 66	1 (0%)
2	B	182/200 (91%)	-0.48	1 (0%) 91 93	15, 24, 42, 61	0
2	E	178/200 (89%)	-0.26	1 (0%) 89 92	18, 30, 48, 67	0
2	H	179/200 (89%)	-0.12	3 (1%) 70 75	20, 28, 45, 72	0
2	K	182/200 (91%)	-0.57	2 (1%) 80 84	16, 22, 38, 67	0
3	C	136/137 (99%)	-0.50	0 100 100	16, 26, 41, 53	1 (0%)
3	F	136/137 (99%)	-0.23	2 (1%) 73 78	23, 32, 49, 63	0
3	I	136/137 (99%)	-0.41	0 100 100	19, 26, 44, 56	1 (0%)
3	L	136/137 (99%)	-0.42	0 100 100	18, 28, 43, 54	0
All	All	3472/3564 (97%)	-0.40	18 (0%) 91 93	12, 22, 40, 72	3 (0%)

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	224	LEU	5.2
2	K	58	THR	4.2
1	J	1	MET	4.2
1	A	551	ASN	3.7
2	E	74	ALA	3.4

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 carbohydrates 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, 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
8	CL	C	201	1/1	0.96	0.08	33,33,33,33	0
7	5AD	D	605	18/18	0.96	0.08	16,19,21,22	0
7	5AD	G	605	18/18	0.98	0.06	15,18,22,22	0
9	B12	H	1601	91/91	0.98	0.07	15,18,21,23	0
7	5AD	J	604	18/18	0.98	0.08	16,18,21,21	0
8	CL	J	606	1/1	0.98	0.03	32,32,32,32	0
8	CL	D	606	1/1	0.98	0.03	32,32,32,32	0
7	5AD	A	605	18/18	0.98	0.09	14,16,18,19	0
9	B12	B	1601	91/91	0.98	0.07	13,16,20,23	0
6	PGO	D	604	5/5	0.98	0.10	19,21,29,30	0
6	PGO	G	604	5/5	0.98	0.08	19,22,27,33	0
9	B12	K	1601	91/91	0.98	0.07	13,17,19,22	0
9	B12	E	1601	91/91	0.98	0.07	15,18,21,31	0
6	PGO	A	604	5/5	0.98	0.07	16,17,28,29	0
8	CL	F	201	1/1	0.99	0.04	27,27,27,27	0
8	CL	L	201	1/1	0.99	0.13	31,31,31,31	0
8	CL	G	607	1/1	0.99	0.06	27,27,27,27	0
8	CL	A	606	1/1	0.99	0.03	23,23,23,23	0
8	CL	J	605	1/1	0.99	0.05	28,28,28,28	0
8	CL	D	607	1/1	0.99	0.07	28,28,28,28	0
5	K	J	602	1/1	1.00	0.06	17,17,17,17	0
5	K	J	603	1/1	1.00	0.05	17,17,17,17	0
8	CL	G	606	1/1	1.00	0.05	29,29,29,29	0
5	K	D	603	1/1	1.00	0.06	19,19,19,19	0
8	CL	A	607	1/1	1.00	0.06	29,29,29,29	0
5	K	G	602	1/1	1.00	0.03	18,18,18,18	0
4	CA	A	601	1/1	1.00	0.07	15,15,15,15	0
5	K	A	603	1/1	1.00	0.07	17,17,17,17	0
4	CA	G	601	1/1	1.00	0.05	18,18,18,18	0
8	CL	I	201	1/1	1.00	0.02	21,21,21,21	0
5	K	D	602	1/1	1.00	0.04	18,18,18,18	0
4	CA	D	601	1/1	1.00	0.04	16,16,16,16	0

Continued on next page...

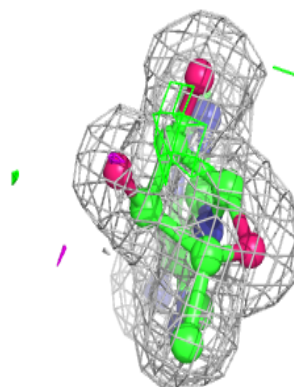
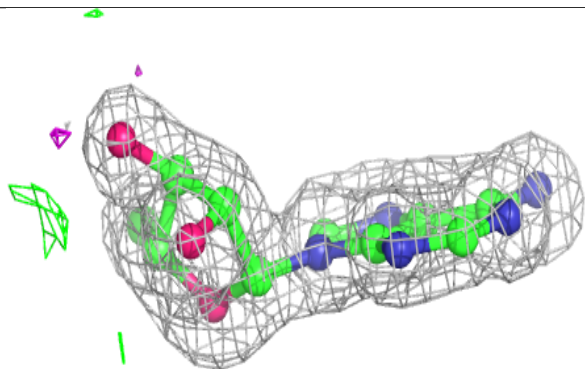
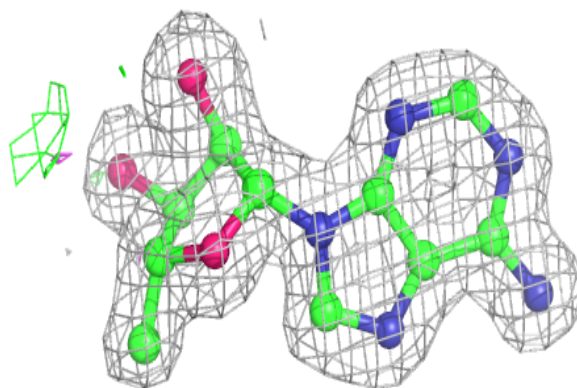
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	CA	J	601	1/1	1.00	0.05	16,16,16,16	0
5	K	A	602	1/1	1.00	0.06	14,14,14,14	0
5	K	G	603	1/1	1.00	0.07	18,18,18,18	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.

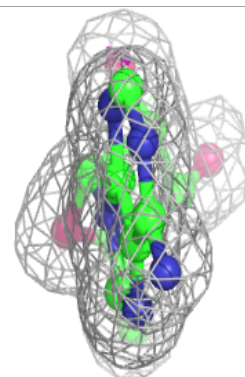
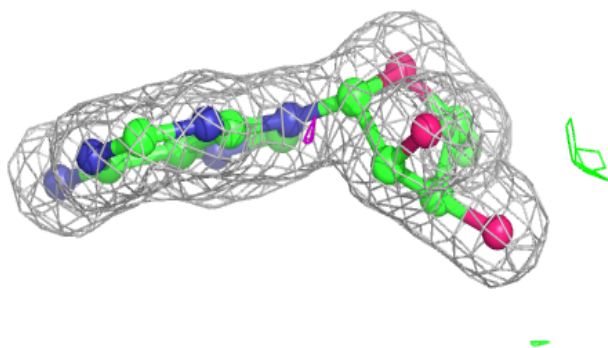
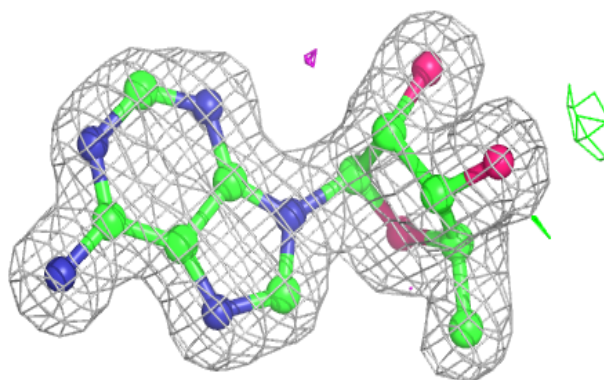
Electron density around 5AD D 605:

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



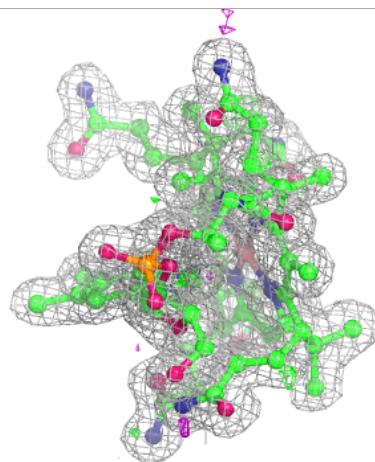
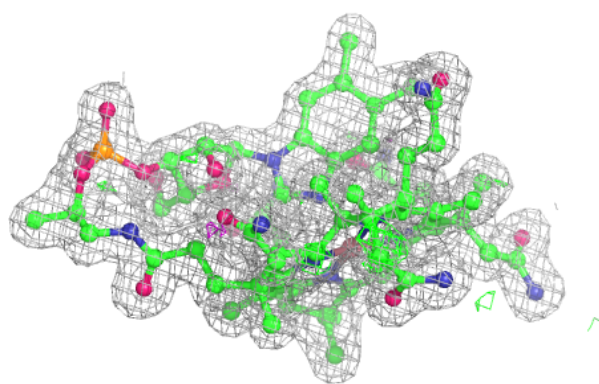
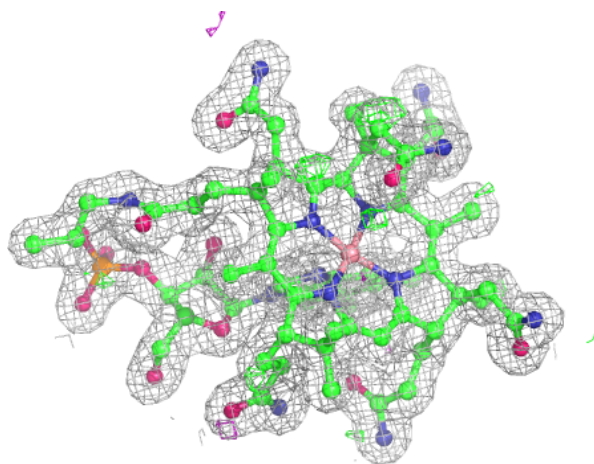
Electron density around 5AD G 605:

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



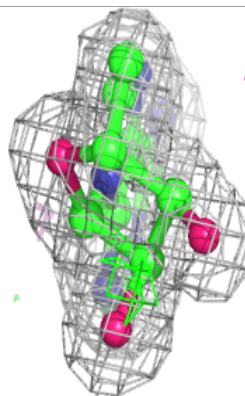
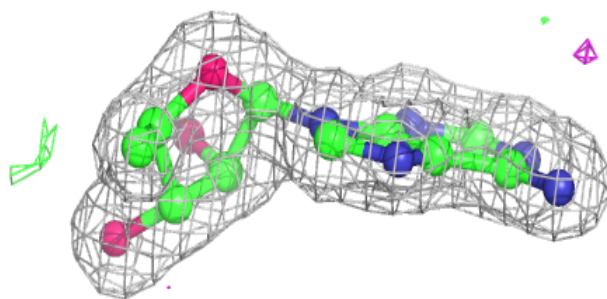
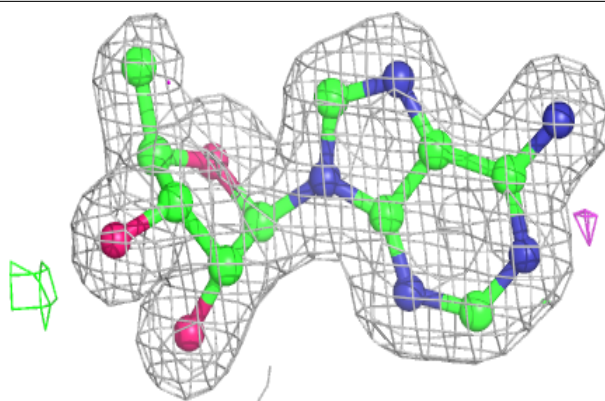
Electron density around B12 H 1601:

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

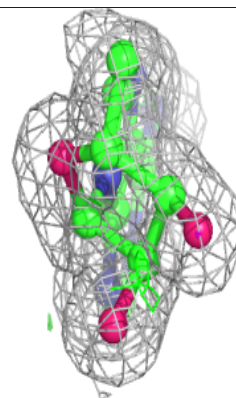
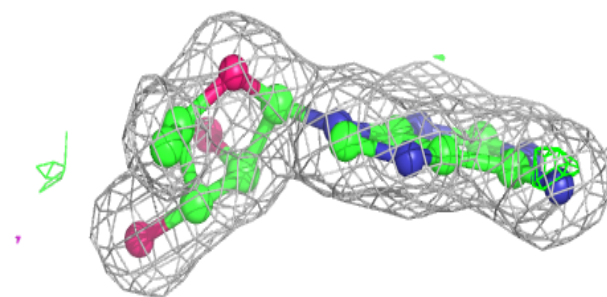
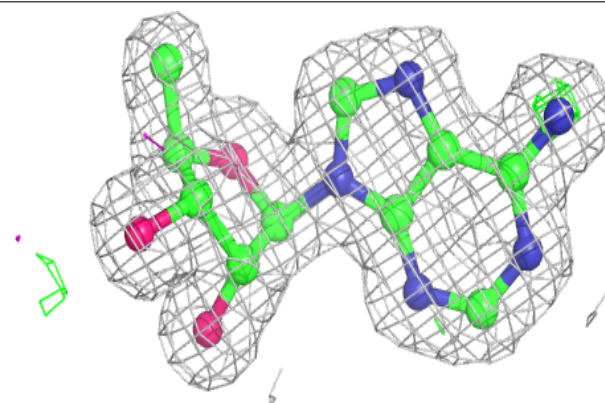


Electron density around 5AD J 604:

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

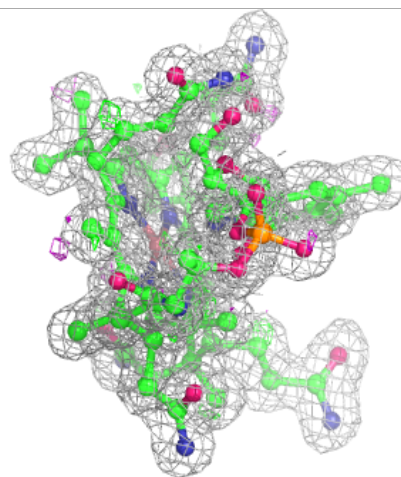
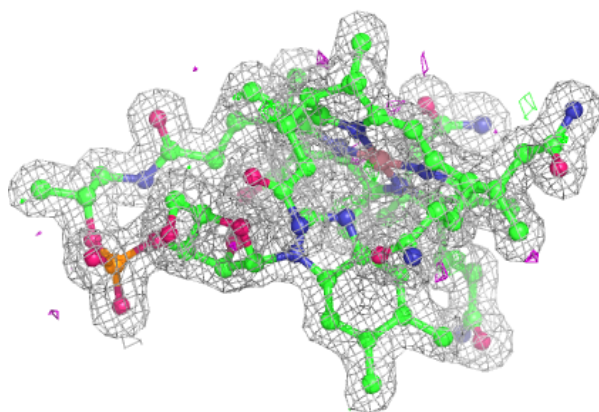
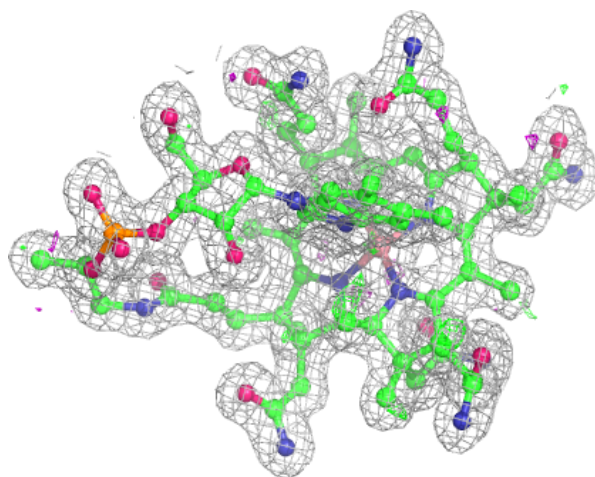
**Electron density around 5AD A 605:**

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



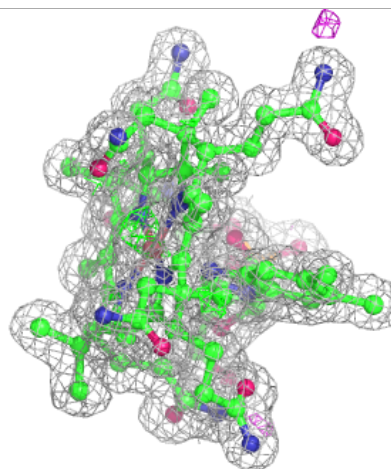
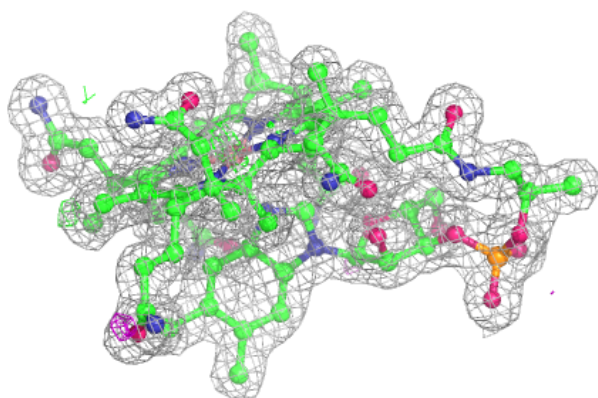
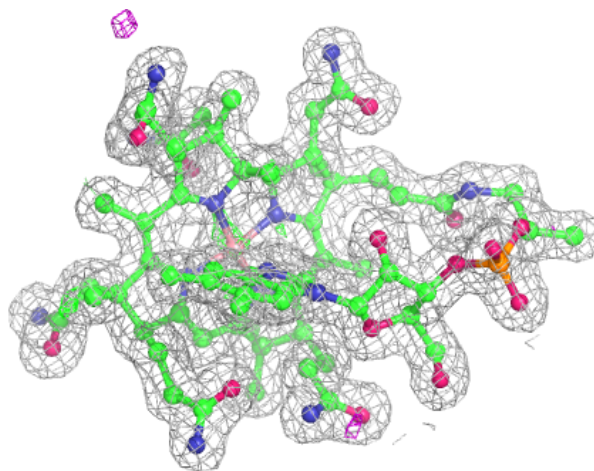
Electron density around B12 B 1601:

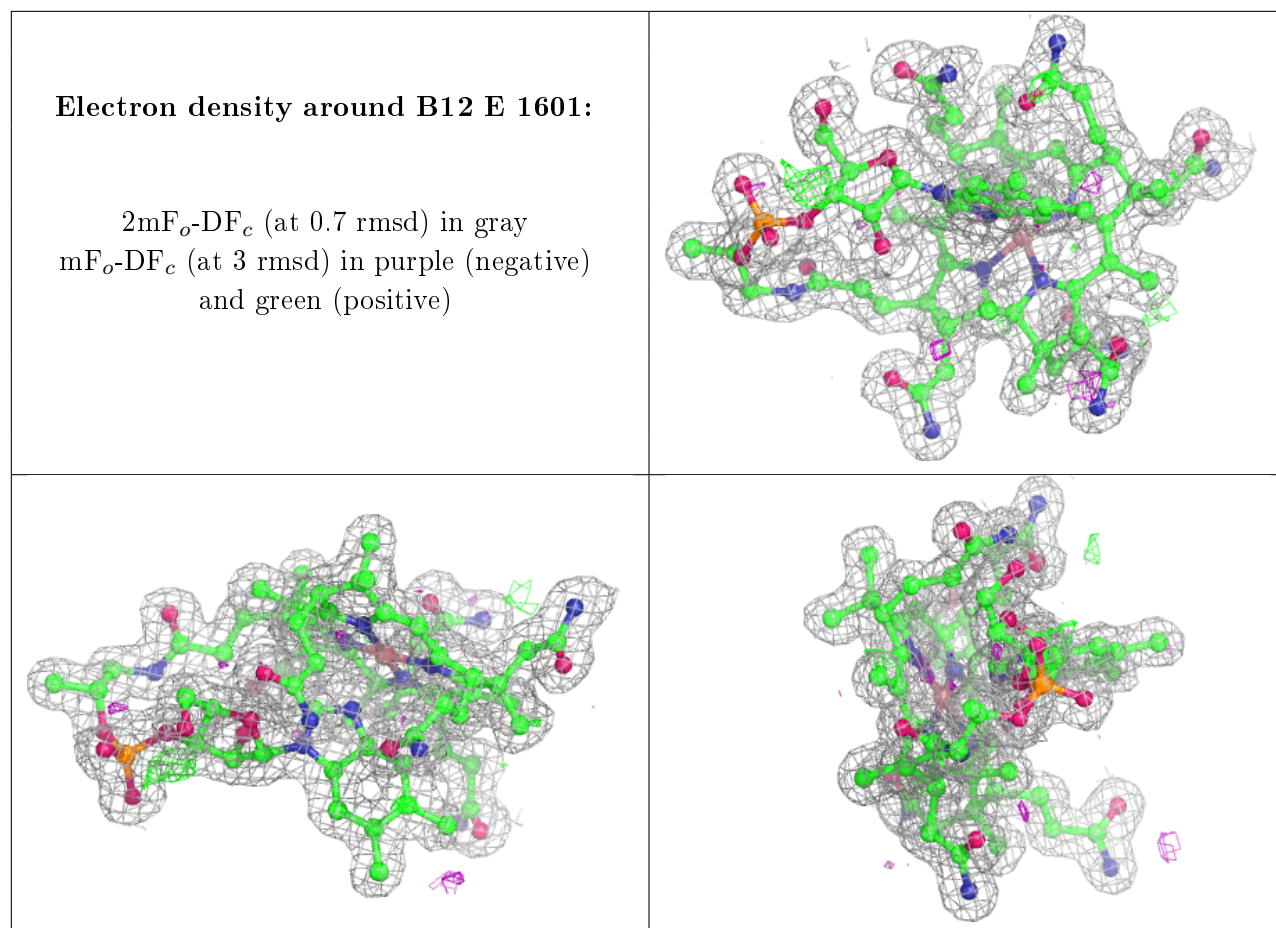
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around B12 K 1601:

$2mF_o-DF_c$ (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.