



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2026 – 11:10 AM UTC

PDB ID : 8CBS / pdb_00008cbs
Title : HIV-1 Integrase Catalytic Core Domain and C-Terminal Domain in Complex with Allosteric Integrase Inhibitor MUT871
Authors : Singer, M.R.; Pye, V.E.; Yu, Z.; Cherepanov, P.
Deposited on : 2023-01-25
Resolution : 1.70 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

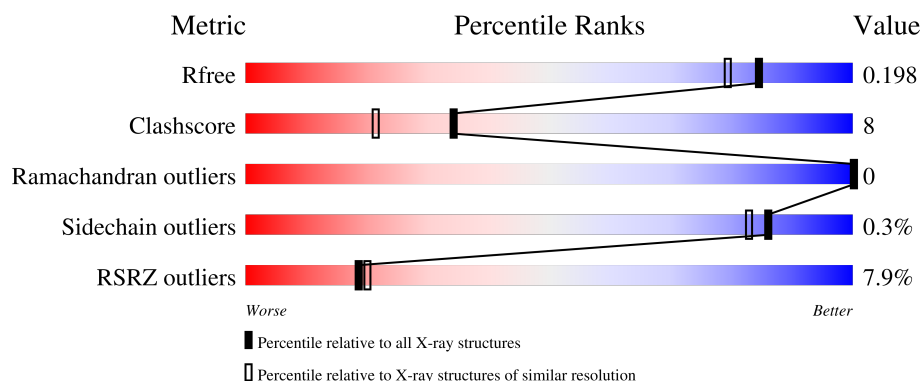
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.70 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}	180053	5551 (1.70-1.70)
Clashscore	190562	5924 (1.70-1.70)
Ramachandran outliers	187476	5846 (1.70-1.70)
Sidechain outliers	187428	5846 (1.70-1.70)
RSRZ outliers	180081	5554 (1.70-1.70)

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 $\leq 5\%$. 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	233	3% (red) 19% (orange) • (yellow) 77% (grey)
1	B	233	2% (red) 52% (orange) 8% (yellow) 39% (grey)
1	C	233	6% (red) 18% (orange) • (yellow) 78% (grey)
1	D	233	3% (red) 54% (orange) 8% (yellow) 38% (grey)

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 3490 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 Integrase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	53	Total	C	N	O		0	2	0
			449	287	85	77				
1	B	141	Total	C	N	O	S	0	5	0
			1125	723	195	202	5			
1	C	52	Total	C	N	O		0	2	0
			439	280	82	77				
1	D	144	Total	C	N	O	S	0	3	0
			1132	726	196	205	5			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	219	SER	-	expression tag	UNP P12497
A	243	GLU	TRP	engineered mutation	UNP P12497
A	424	LYS	PHE	engineered mutation	UNP P12497
B	-20	SER	-	expression tag	UNP P12497
B	4	GLU	TRP	engineered mutation	UNP P12497
B	185	LYS	PHE	engineered mutation	UNP P12497
C	219	SER	-	expression tag	UNP P12497
C	243	GLU	TRP	engineered mutation	UNP P12497
C	424	LYS	PHE	engineered mutation	UNP P12497
D	-20	SER	-	expression tag	UNP P12497
D	4	GLU	TRP	engineered mutation	UNP P12497
D	185	LYS	PHE	engineered mutation	UNP P12497

- Molecule 2 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: C₂H₆O₂).

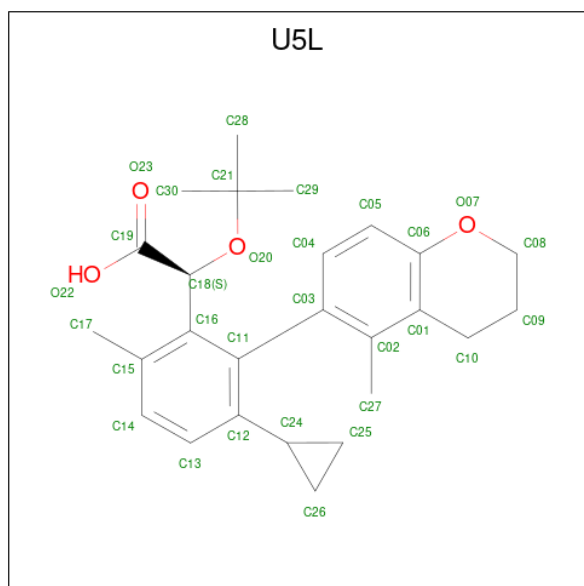


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		
2	B	1	Total	C	O	0	0
			4	2	2		
2	B	1	Total	C	O	0	0
			4	2	2		
2	B	1	Total	C	O	0	0
			4	2	2		
2	B	1	Total	C	O	0	0
			4	2	2		
2	C	1	Total	C	O	0	0
			4	2	2		
2	C	1	Total	C	O	0	0
			4	2	2		
2	D	1	Total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

- Molecule 3 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Cl 1 1	0	0

- Molecule 4 is (2 {S})-2-[3-cyclopropyl-6-methyl-2-(5-methyl-3,4-dihydro-2 {H}-chromen-6-yl)phenyl]-2-[(2-methylpropan-2-yl)oxy]ethanoic acid (CCD ID: U5L) (formula: C₂₆H₃₂O₄) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C O 30 26 4	0	0
4	D	1	Total C O 30 26 4	0	0

- Molecule 5 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total Mg 1 1	0	0
5	D	1	Total Mg 1 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	23	Total O 23 23	0	0

Continued on next page...

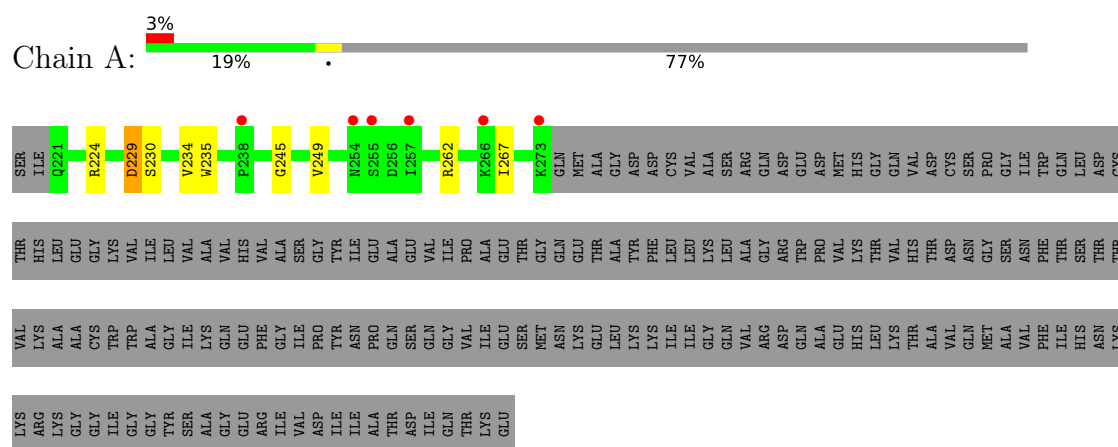
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	93	Total 93	O 93	0	0
6	C	22	Total 22	O 22	0	0
6	D	88	Total 88	O 88	0	0

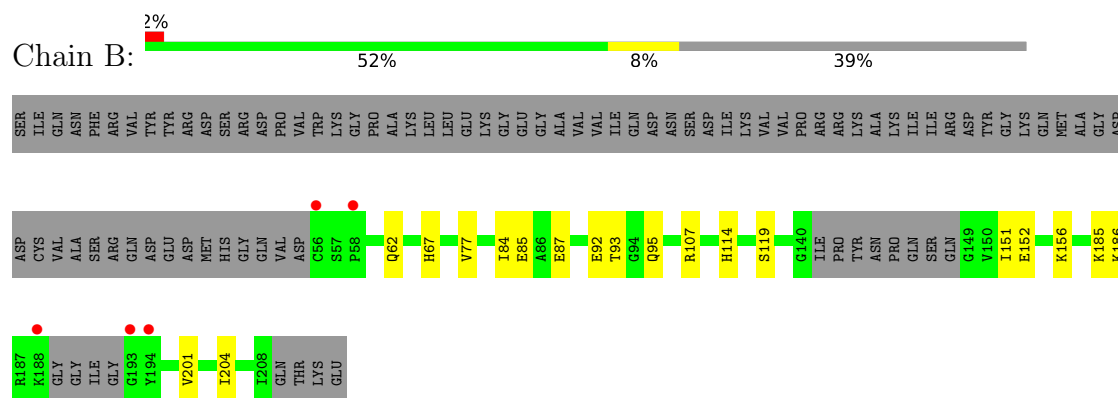
3 Residue-property plots

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.

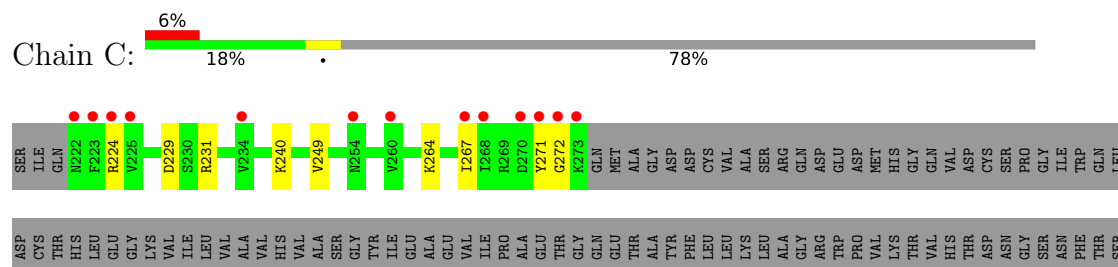
• Molecule 1: Integrase



• Molecule 1: Integrase



• Molecule 1: Integrase



THR	ASN
THR	LYS
VAL	LYS
LYS	ARG
ALA	LYS
ALA	GLY
CYS	GLY
TRP	ILE
TRP	GLY
ALA	TYR
GLY	SER
ILE	ILE
LYS	ALA
GLN	GLY
GLU	ARG
PHE	ARG
GLY	ILE
ILE	VAL
ILE	ASP
PRO	ASP
TYR	ILE
	ILE
	ALA
	THR
	ASP
	SER
	GLN
	ILE
	GLN
	THR
	LYS
	GLU
	GLY
	ILE
	GLN
	VAL
	ARG
	ASP
	GLN
	ALA
	GLU
	HIS
	LEU
	LYS
	THR
	ALA
	GLN
	MET
	VAL
	PHE
	ILE
	HIS

● Molecule 1: Integrase



SER	ILE	GLN	ASN	PHE	ARG	VAL	TRP	TYR	ASP	SER	ARG	ASP	PRO	VAL	ASP	TRP	LYS	GLY	PRO	ALA	ASN	LYS	LEU	LEU	GLU	GLY	VAL	ALA	VAL	ILE	GLN	ASP	ASN	SER	ASP	ILE	LYS	VAL	GLN	VAL	PRO	ARG	ARG	LYS	ALA	LYS	ILE	ILE	ARG	ASP	TYR	GLY	LYS	MET	ALA	GLY	ASP
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

ASP	CYS	VAL	ALA	SER	GLN	ARG	ASP	GLU	ASP	MET	HIS	GLY	GLN	VAL	ASP	C56	V77	I84	E85	T93	R107	M120	W131	W132	G140	ILE	PRO	TYR	ASN	PRO	GLN	SER	GLN	G149	V150	I151	M154	R166	K173	R196	R187	K188	GLY	GLY	ILE	GLY	G193	Y194
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	------	------

E196	V201	A205	T206	D207	T208	K211	GLU
------	------	------	------	------	------	------	-----

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	53.23Å 64.85Å 70.44Å 90.00° 101.50° 90.00°	Depositor
Resolution (Å)	52.16 – 1.70 52.16 – 1.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (52.16-1.70) 100.0 (52.16-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.03 (at 1.70Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.167 , 0.198 0.167 , 0.198	Depositor DCC
R_{free} test set	2574 reflections (4.69%)	wwPDB-VP
Wilson B-factor (Å ²)	38.0	Xtriage
Anisotropy	0.232	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 49.6	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	0.97	EDS
Total number of atoms	3490	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.94% 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: EDO, U5L, CL, MG

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	0.56	0/460	0.77	0/616
1	B	0.76	0/1156	0.76	0/1563
1	C	0.63	0/450	0.75	0/605
1	D	0.79	0/1162	0.80	0/1570
All	All	0.73	0/3228	0.78	0/4354

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	449	0	473	9	0
1	B	1125	0	1145	18	0
1	C	439	0	456	8	0
1	D	1132	0	1156	18	0
2	A	8	0	12	6	0
2	B	16	0	24	4	0
2	C	8	0	12	3	0
2	D	24	0	36	6	0
3	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	30	0	0	0	0
4	D	30	0	0	0	0
5	B	1	0	0	0	0
5	D	1	0	0	0	0
6	A	23	0	0	0	0
6	B	93	0	0	2	0
6	C	22	0	0	2	0
6	D	88	0	0	0	0
All	All	3490	0	3314	50	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 8.

All (50) 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:B:93:THR:HA	2:B:601:EDO:H11	1.67	0.76
1:B:77[B]:VAL:HG22	1:B:84:ILE:HG22	1.70	0.72
1:C:249[A]:VAL:HG11	1:C:267:ILE:HD11	1.74	0.69
1:D:173:LYS:HE2	2:D:308:EDO:H22	1.77	0.67
1:D:187:ARG:NH1	1:D:193:GLY:O	2.31	0.64
1:D:166:ARG:HB3	2:D:305:EDO:H21	1.79	0.63
1:C:271:TYR:CE2	2:C:501:EDO:H11	2.33	0.63
1:A:224:ARG:HG2	2:A:502:EDO:H21	1.81	0.63
1:B:77[B]:VAL:HG23	1:B:151:ILE:HD12	1.81	0.63
1:D:77[B]:VAL:HG22	1:D:84:ILE:HG22	1.82	0.61
1:B:185:LYS:HB2	2:C:501:EDO:H12	1.82	0.61
1:A:235:TRP:HD1	2:A:501:EDO:H21	1.66	0.60
1:C:272:GLY:HA3	1:D:132:TRP:O	2.02	0.59
1:C:264:LYS:NZ	6:C:602:HOH:O	2.35	0.58
1:D:187:ARG:NH2	1:D:198:GLU:OE2	2.36	0.58
1:D:85:GLU:OE1	1:D:107:ARG:NH1	2.37	0.57
1:B:67[B]:HIS:HE1	1:B:92:GLU:OE1	1.91	0.53
1:D:93:THR:HA	2:D:304:EDO:H22	1.92	0.50
1:B:77[B]:VAL:HG23	1:B:151:ILE:CD1	2.41	0.50
1:B:185:LYS:HA	2:B:602:EDO:H21	1.93	0.50
1:D:207:ASP:O	1:D:211:LYS:HG2	2.11	0.49
1:B:77[B]:VAL:CG2	1:B:151:ILE:HD12	2.43	0.49
1:B:87:GLU:OE1	6:B:701:HOH:O	2.20	0.48
1:D:173:LYS:CE	2:D:308:EDO:H22	2.43	0.48
1:A:234:VAL:HA	2:A:501:EDO:H22	1.95	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:85:GLU:OE1	1:B:107:ARG:NH1	2.47	0.47
1:A:235:TRP:H	2:A:501:EDO:C2	2.28	0.47
1:B:62:GLN:HG2	1:B:114[B]:HIS:HB3	1.97	0.46
1:A:224:ARG:HE	2:A:502:EDO:H21	1.80	0.46
1:B:201:VAL:HG12	1:D:205:ALA:HB2	1.97	0.46
1:C:240:LYS:HA	6:C:615:HOH:O	2.16	0.46
1:A:224:ARG:NH2	2:A:502:EDO:O1	2.47	0.45
1:D:166:ARG:HB3	2:D:305:EDO:C2	2.46	0.45
1:D:211:LYS:HD3	1:D:211:LYS:HA	1.78	0.44
1:C:229:ASP:O	1:C:231:ARG:N	2.43	0.44
1:D:120:ASN:HA	2:D:304:EDO:H11	1.99	0.44
1:C:224:ARG:HD3	1:D:131:TRP:CE2	2.53	0.44
2:B:601:EDO:C2	2:B:604:EDO:H11	2.49	0.43
1:B:62:GLN:HG2	1:B:114[A]:HIS:HB2	2.00	0.43
1:A:229:ASP:HB3	1:A:230:SER:H	1.63	0.43
1:B:95:GLN:HG3	6:B:773:HOH:O	2.18	0.43
1:D:77[B]:VAL:HG23	1:D:151:ILE:HD12	1.99	0.43
1:B:119:SER:HB2	2:B:604:EDO:H22	2.01	0.42
1:B:204[B]:ILE:HG22	1:D:201:VAL:HG11	2.01	0.42
1:C:271:TYR:HE2	2:C:501:EDO:H11	1.79	0.42
1:D:84:ILE:HG21	1:D:154:MET:HG2	2.00	0.42
1:A:249[A]:VAL:HG11	1:A:267:ILE:HD11	2.00	0.42
1:B:152:GLU:O	1:B:156:LYS:HG3	2.18	0.42
1:A:245:GLY:O	1:A:262:ARG:NH1	2.51	0.42
1:B:186:LYS:HE3	1:B:186:LYS:HB2	1.81	0.42

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	53/233 (23%)	49 (92%)	4 (8%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	140/233 (60%)	139 (99%)	1 (1%)	0	100	100
1	C	52/233 (22%)	48 (92%)	4 (8%)	0	100	100
1	D	141/233 (60%)	139 (99%)	2 (1%)	0	100	100
All	All	386/932 (41%)	375 (97%)	11 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	48/193 (25%)	47 (98%)	1 (2%)	47	30
1	B	120/193 (62%)	120 (100%)	0	100	100
1	C	47/193 (24%)	47 (100%)	0	100	100
1	D	121/193 (63%)	121 (100%)	0	100	100
All	All	336/772 (44%)	335 (100%)	1 (0%)	86	83

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	229	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	62	GLN
1	B	168	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 19 ligands modelled in this entry, 3 are monoatomic - leaving 16 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	$\# Z > 2$
2	EDO	D	302	-	3,3,3	0.59	0	2,2,2	0.38	0
2	EDO	A	501	-	3,3,3	0.63	0	2,2,2	0.28	0
2	EDO	B	602	-	3,3,3	0.41	0	2,2,2	0.34	0
2	EDO	C	502	-	3,3,3	0.54	0	2,2,2	0.26	0
2	EDO	D	306	-	3,3,3	0.57	0	2,2,2	0.15	0
4	U5L	D	301	-	32,33,33	1.95	7 (21%)	40,50,50	1.56	6 (15%)
2	EDO	C	501	-	3,3,3	0.51	0	2,2,2	0.22	0
2	EDO	D	303	-	3,3,3	0.75	0	2,2,2	0.47	0
2	EDO	D	305	-	3,3,3	0.69	0	2,2,2	0.67	0
2	EDO	B	603	-	3,3,3	0.71	0	2,2,2	0.57	0
2	EDO	D	304	-	3,3,3	0.38	0	2,2,2	0.87	0
2	EDO	A	502	-	3,3,3	0.34	0	2,2,2	0.61	0
2	EDO	D	308	-	3,3,3	0.48	0	2,2,2	0.04	0
4	U5L	B	605	-	32,33,33	2.13	8 (25%)	40,50,50	1.66	8 (20%)
2	EDO	B	601	-	3,3,3	0.58	0	2,2,2	0.50	0
2	EDO	B	604	-	3,3,3	0.40	0	2,2,2	0.56	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	EDO	D	302	-	-	0/1/1/1	-
2	EDO	A	501	-	-	1/1/1/1	-
2	EDO	B	602	-	-	0/1/1/1	-
2	EDO	C	502	-	-	0/1/1/1	-
2	EDO	D	306	-	-	1/1/1/1	-
4	U5L	D	301	-	-	3/21/30/30	0/4/4/4
2	EDO	C	501	-	-	0/1/1/1	-
2	EDO	D	303	-	-	0/1/1/1	-
2	EDO	D	305	-	-	1/1/1/1	-
2	EDO	B	603	-	-	1/1/1/1	-
2	EDO	D	304	-	-	0/1/1/1	-
2	EDO	A	502	-	-	0/1/1/1	-
2	EDO	D	308	-	-	0/1/1/1	-
4	U5L	B	605	-	-	3/21/30/30	0/4/4/4
2	EDO	B	601	-	-	1/1/1/1	-
2	EDO	B	604	-	-	1/1/1/1	-

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	301	U5L	C10-C01	-7.32	1.40	1.51
4	B	605	U5L	C10-C01	-7.16	1.40	1.51
4	B	605	U5L	C17-C15	4.46	1.59	1.51
4	B	605	U5L	C05-C04	3.69	1.44	1.38
4	D	301	U5L	C27-C02	3.16	1.57	1.51
4	D	301	U5L	C03-C11	3.03	1.54	1.50
4	B	605	U5L	O23-C19	2.88	1.30	1.22
4	B	605	U5L	C14-C13	2.88	1.43	1.38
4	B	605	U5L	C03-C11	2.82	1.53	1.50
4	D	301	U5L	O23-C19	2.38	1.29	1.22
4	B	605	U5L	C28-C21	2.32	1.57	1.51
4	D	301	U5L	O20-C18	-2.31	1.39	1.43
4	B	605	U5L	C11-C16	2.24	1.44	1.40
4	D	301	U5L	C13-C12	-2.21	1.37	1.39
4	D	301	U5L	C04-C03	2.18	1.43	1.40

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	301	U5L	C25-C24-C12	5.63	132.30	121.75
4	B	605	U5L	C13-C12-C11	3.88	121.49	118.48
4	B	605	U5L	C13-C12-C24	-3.55	114.10	119.96
4	B	605	U5L	C10-C01-C02	-3.49	114.90	121.22

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	301	U5L	O22-C19-O23	-3.33	116.52	124.08
4	B	605	U5L	O07-C06-C01	-3.18	119.74	122.74
4	B	605	U5L	C26-C24-C12	-3.16	115.84	121.75
4	D	301	U5L	O20-C18-C19	-2.75	101.69	109.46
4	B	605	U5L	O22-C19-O23	-2.74	117.87	124.08
4	B	605	U5L	C25-C24-C12	2.70	126.81	121.75
4	B	605	U5L	C10-C09-C08	-2.49	105.25	110.06
4	D	301	U5L	C17-C15-C14	-2.30	115.85	120.28
4	D	301	U5L	O07-C08-C09	-2.29	108.01	112.11
4	D	301	U5L	C13-C14-C15	-2.25	118.61	121.98

There are no chirality outliers.

All (12) torsion outliers are listed below:

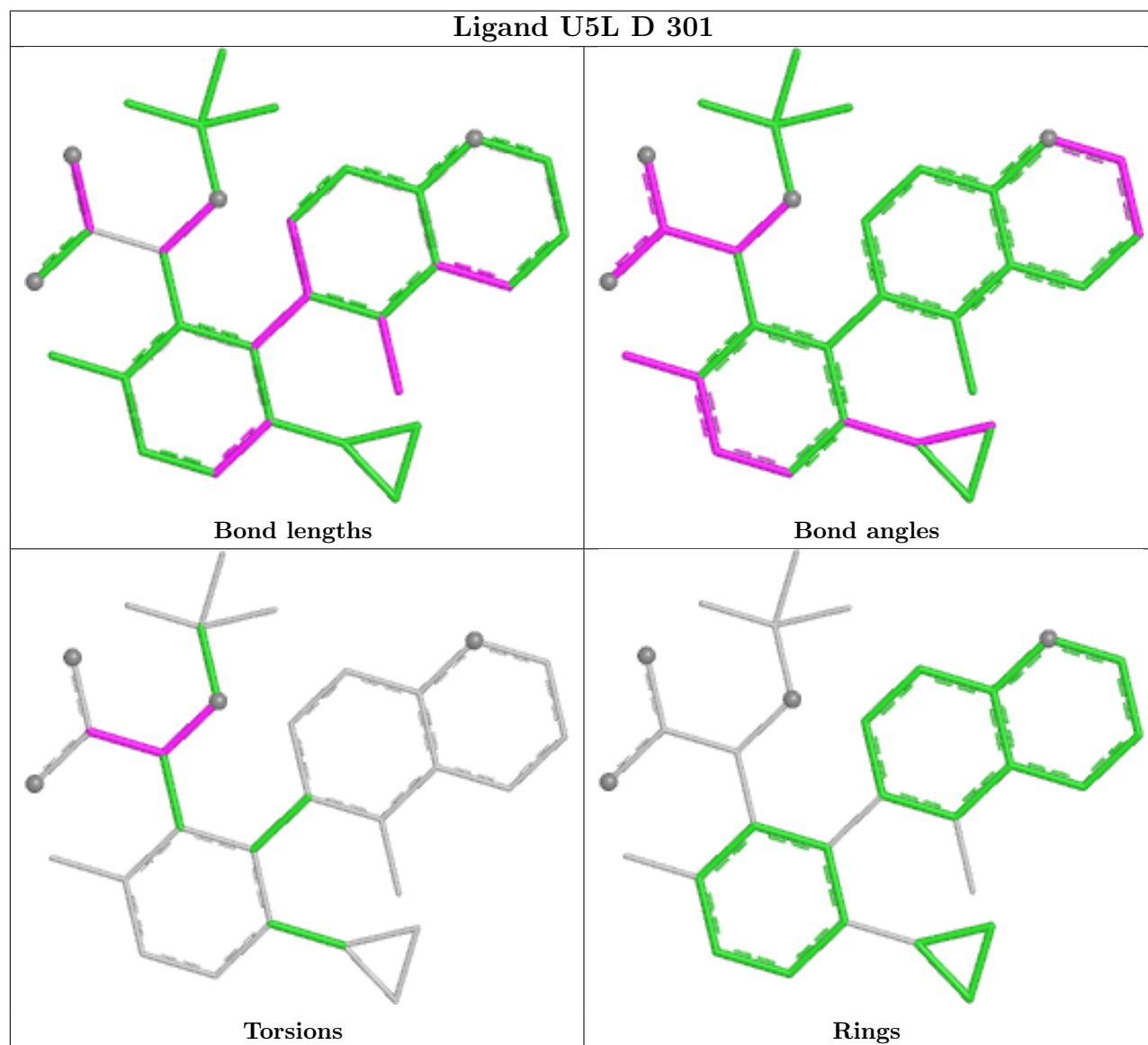
Mol	Chain	Res	Type	Atoms
4	B	605	U5L	C16-C18-C19-O22
4	D	301	U5L	C16-C18-C19-O22
4	B	605	U5L	C16-C18-C19-O23
4	D	301	U5L	C16-C18-C19-O23
2	B	601	EDO	O1-C1-C2-O2
2	B	604	EDO	O1-C1-C2-O2
2	D	305	EDO	O1-C1-C2-O2
2	A	501	EDO	O1-C1-C2-O2
4	D	301	U5L	C19-C18-O20-C21
4	B	605	U5L	C15-C16-C18-O20
2	D	306	EDO	O1-C1-C2-O2
2	B	603	EDO	O1-C1-C2-O2

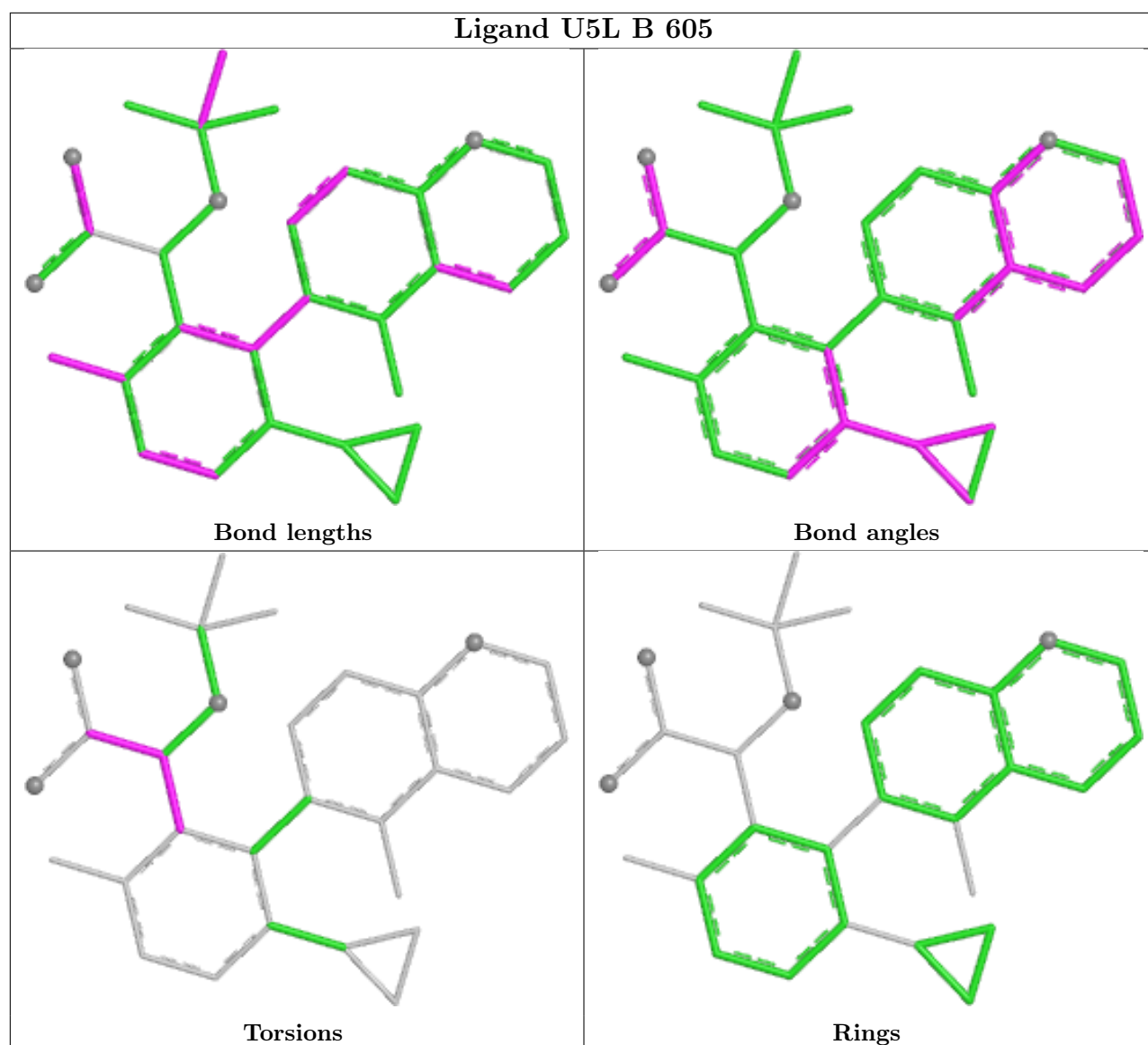
There are no ring outliers.

9 monomers are involved in 19 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	EDO	3	0
2	B	602	EDO	1	0
2	C	501	EDO	3	0
2	D	305	EDO	2	0
2	D	304	EDO	2	0
2	A	502	EDO	3	0
2	D	308	EDO	2	0
2	B	601	EDO	2	0
2	B	604	EDO	2	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.





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	53/233 (22%)	1.07	6 (11%)	10 10	24, 61, 107, 141	2 (3%)
1	B	141/233 (60%)	0.25	5 (3%)	47 52	19, 40, 74, 96	5 (3%)
1	C	52/233 (22%)	1.34	13 (25%)	2 1	21, 61, 106, 114	2 (3%)
1	D	144/233 (61%)	0.30	7 (4%)	35 38	23, 42, 74, 99	3 (2%)
All	All	390/932 (41%)	0.53	31 (7%)	18 20	19, 45, 91, 141	12 (3%)

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	271	TYR	4.4
1	B	56	CYS	4.3
1	D	56	CYS	3.8
1	B	58	PRO	3.6
1	C	272	GLY	3.5
1	D	194	TYR	3.5
1	C	223	PHE	3.3
1	C	268	ILE	3.1
1	D	208	ILE	2.9
1	A	238	PRO	2.7
1	A	273	LYS	2.7
1	C	222	ASN	2.7
1	C	267	ILE	2.6
1	C	270	ASP	2.5
1	C	254	ASN	2.5
1	A	266[A]	LYS	2.4
1	C	273	LYS	2.4
1	D	131	TRP	2.4
1	B	188	LYS	2.4
1	A	255	SER	2.3
1	D	186	LYS	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	193	GLY	2.3
1	C	234	VAL	2.2
1	C	224	ARG	2.2
1	A	254	ASN	2.2
1	A	257	ILE	2.1
1	B	194	TYR	2.1
1	D	149	GLY	2.0
1	C	225	VAL	2.0
1	C	260	VAL	2.0
1	D	188	LYS	2.0

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 oligosaccharides 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(Å ²)	Q<0.9
2	EDO	A	502	4/4	0.75	0.14	64,65,66,85	0
2	EDO	C	502	4/4	0.76	0.16	72,72,75,82	0
2	EDO	B	601	4/4	0.77	0.20	49,58,63,64	0
2	EDO	D	303	4/4	0.78	0.18	44,58,63,68	0
2	EDO	D	305	4/4	0.81	0.14	59,63,64,68	0
2	EDO	C	501	4/4	0.84	0.15	55,55,68,73	0
2	EDO	A	501	4/4	0.84	0.17	52,57,71,72	0
2	EDO	B	602	4/4	0.86	0.13	59,68,71,72	0
2	EDO	B	603	4/4	0.87	0.15	54,58,58,70	0
2	EDO	D	304	4/4	0.89	0.12	53,60,65,67	0
2	EDO	B	604	4/4	0.89	0.11	48,54,59,64	0
2	EDO	D	306	4/4	0.90	0.10	59,61,66,72	0
2	EDO	D	308	4/4	0.90	0.10	61,68,78,84	0

Continued on next page...

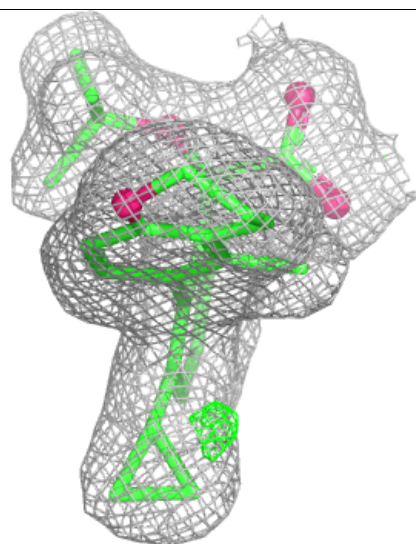
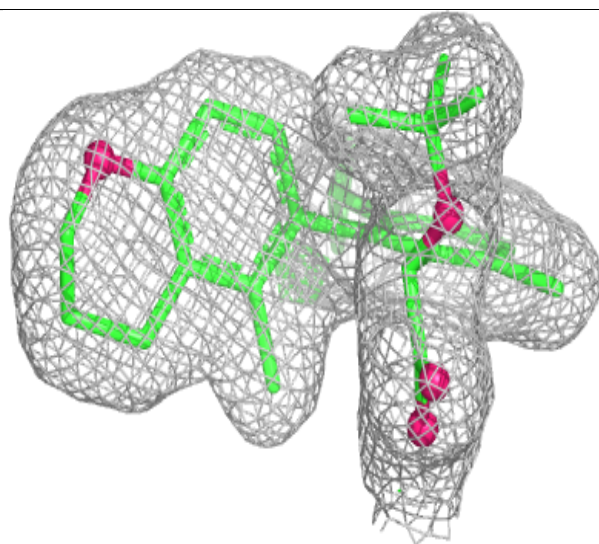
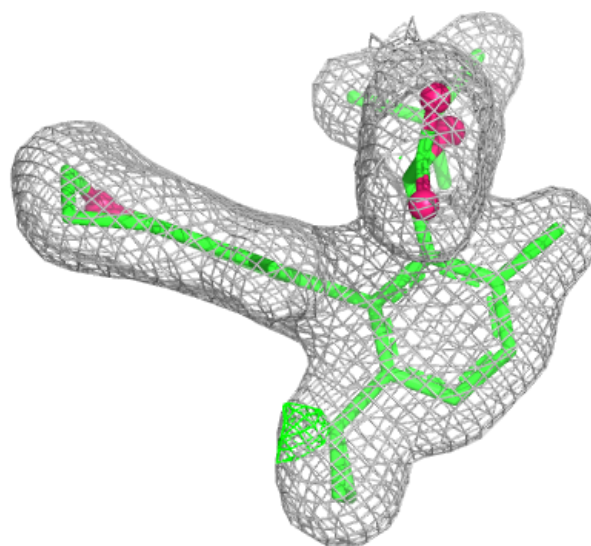
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
2	EDO	D	302	4/4	0.91	0.10	45,47,48,49	0
3	CL	A	503	1/1	0.96	0.10	65,65,65,65	0
4	U5L	B	605	30/30	0.97	0.07	31,34,38,41	0
4	U5L	D	301	30/30	0.98	0.06	32,35,39,42	0
5	MG	B	606	1/1	0.99	0.02	36,36,36,36	0
5	MG	D	307	1/1	0.99	0.04	44,44,44,44	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.

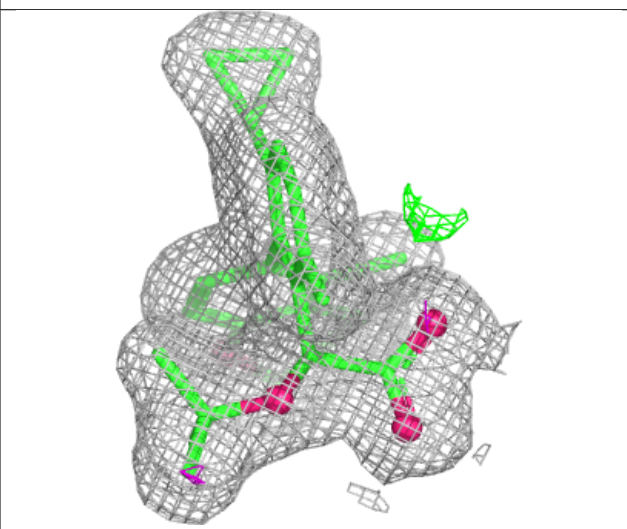
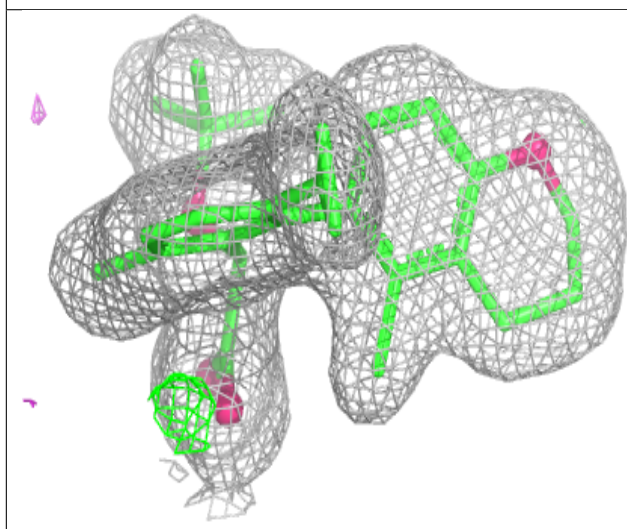
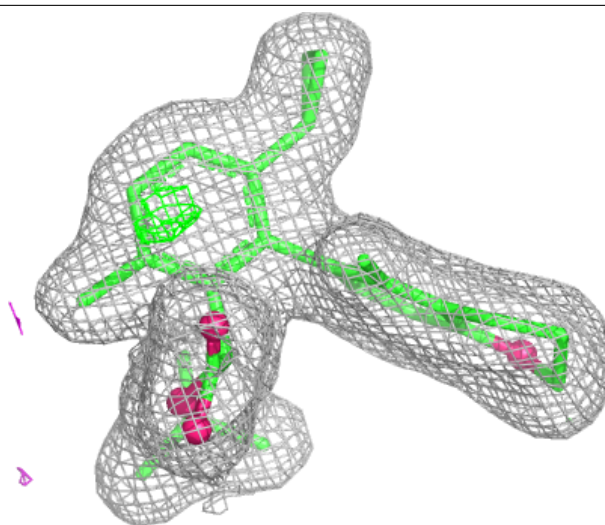
Electron density around U5L B 605:

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 U5L D 301:

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



6.5 Other polymers [i](#)

There are no such residues in this entry.