Full wwPDB X-ray Structure Validation Report ⓘ

Oct 15, 2023 – 04:39 PM EDT

PDB ID	:	8DML
Title	:	Vibrio parahaemolyticus VtrA/VtrC complex bound to the bile salt chenodeoxycholate
Authors	:	Tomchick, D.R.; Orth, K.; Zou, A.J.
Deposited on	:	2022-07-08
Resolution	:	2.08 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:

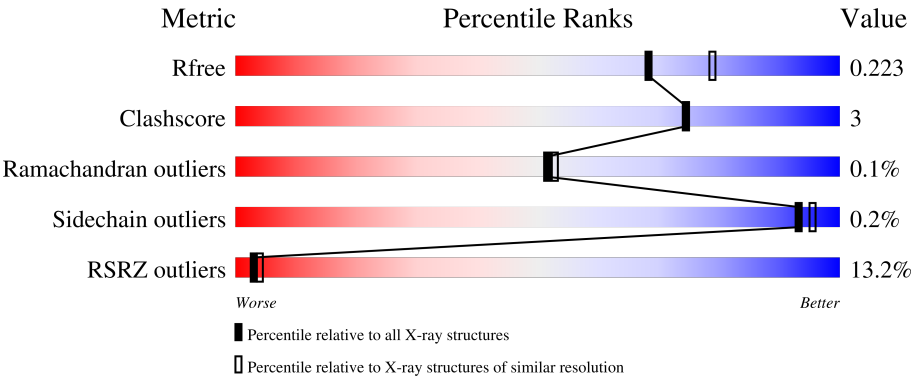
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.08 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.



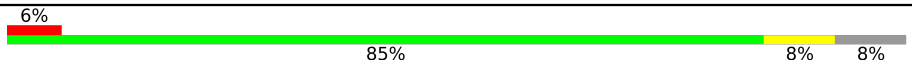
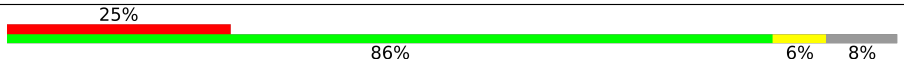
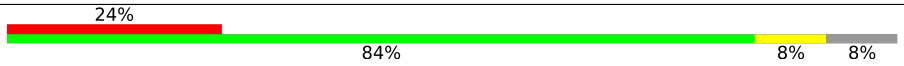
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	6189 (2.10-2.06)
Clashscore	141614	6738 (2.10-2.06)
Ramachandran outliers	138981	6663 (2.10-2.06)
Sidechain outliers	138945	6664 (2.10-2.06)
RSRZ outliers	127900	6057 (2.10-2.06)

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	94	89% 5% 5%
1	C	94	91%
1	E	94	7% 89% 5% 5%
1	G	94	18% 87% 7% 5%
2	B	144	7% 87% 6% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	144	
2	F	144	
2	H	144	

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 15180 atoms, of which 7350 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 VtrA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	89	Total	C	H	N	O	S	0	0	0
			1473	477	737	115	143	1			
1	C	90	Total	C	H	N	O	S	0	1	0
			1497	484	747	117	148	1			
1	E	89	Total	C	H	N	O	S	0	0	0
			1474	477	737	115	144	1			
1	G	89	Total	C	H	N	O	S	0	0	0
			1474	477	737	115	144	1			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	160	MET	-	initiating methionine	UNP Q87GI4
C	160	MET	-	initiating methionine	UNP Q87GI4
E	160	MET	-	initiating methionine	UNP Q87GI4
G	160	MET	-	initiating methionine	UNP Q87GI4

- Molecule 2 is a protein called VtrC.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	133	Total	C	H	N	O	S	0	0	0
			2164	710	1057	181	214	2			
2	D	133	Total	C	H	N	O	S	0	0	0
			2164	710	1057	181	214	2			
2	F	132	Total	C	H	N	O	S	0	0	0
			2147	705	1049	179	212	2			
2	H	132	Total	C	H	N	O	S	0	0	0
			2147	705	1049	179	212	2			

There are 52 discrepancies between the modelled and reference sequences:

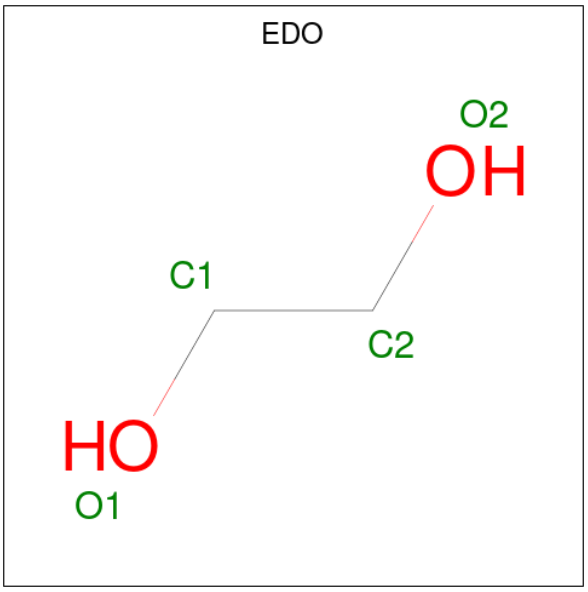
Chain	Residue	Modelled	Actual	Comment	Reference
B	18	MET	-	initiating methionine	UNP Q87GI3
B	19	GLY	-	expression tag	UNP Q87GI3
B	20	SER	-	expression tag	UNP Q87GI3
B	21	SER	-	expression tag	UNP Q87GI3
B	22	HIS	-	expression tag	UNP Q87GI3
B	23	HIS	-	expression tag	UNP Q87GI3
B	24	HIS	-	expression tag	UNP Q87GI3
B	25	HIS	-	expression tag	UNP Q87GI3
B	26	HIS	-	expression tag	UNP Q87GI3
B	27	HIS	-	expression tag	UNP Q87GI3
B	28	SER	-	expression tag	UNP Q87GI3
B	29	GLN	-	expression tag	UNP Q87GI3
B	30	ASP	-	expression tag	UNP Q87GI3
D	18	MET	-	initiating methionine	UNP Q87GI3
D	19	GLY	-	expression tag	UNP Q87GI3
D	20	SER	-	expression tag	UNP Q87GI3
D	21	SER	-	expression tag	UNP Q87GI3
D	22	HIS	-	expression tag	UNP Q87GI3
D	23	HIS	-	expression tag	UNP Q87GI3
D	24	HIS	-	expression tag	UNP Q87GI3
D	25	HIS	-	expression tag	UNP Q87GI3
D	26	HIS	-	expression tag	UNP Q87GI3
D	27	HIS	-	expression tag	UNP Q87GI3
D	28	SER	-	expression tag	UNP Q87GI3
D	29	GLN	-	expression tag	UNP Q87GI3
D	30	ASP	-	expression tag	UNP Q87GI3
F	18	MET	-	initiating methionine	UNP Q87GI3
F	19	GLY	-	expression tag	UNP Q87GI3
F	20	SER	-	expression tag	UNP Q87GI3
F	21	SER	-	expression tag	UNP Q87GI3
F	22	HIS	-	expression tag	UNP Q87GI3
F	23	HIS	-	expression tag	UNP Q87GI3
F	24	HIS	-	expression tag	UNP Q87GI3
F	25	HIS	-	expression tag	UNP Q87GI3
F	26	HIS	-	expression tag	UNP Q87GI3
F	27	HIS	-	expression tag	UNP Q87GI3
F	28	SER	-	expression tag	UNP Q87GI3
F	29	GLN	-	expression tag	UNP Q87GI3
F	30	ASP	-	expression tag	UNP Q87GI3
H	18	MET	-	initiating methionine	UNP Q87GI3
H	19	GLY	-	expression tag	UNP Q87GI3
H	20	SER	-	expression tag	UNP Q87GI3
H	21	SER	-	expression tag	UNP Q87GI3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	22	HIS	-	expression tag	UNP Q87GI3
H	23	HIS	-	expression tag	UNP Q87GI3
H	24	HIS	-	expression tag	UNP Q87GI3
H	25	HIS	-	expression tag	UNP Q87GI3
H	26	HIS	-	expression tag	UNP Q87GI3
H	27	HIS	-	expression tag	UNP Q87GI3
H	28	SER	-	expression tag	UNP Q87GI3
H	29	GLN	-	expression tag	UNP Q87GI3
H	30	ASP	-	expression tag	UNP Q87GI3

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



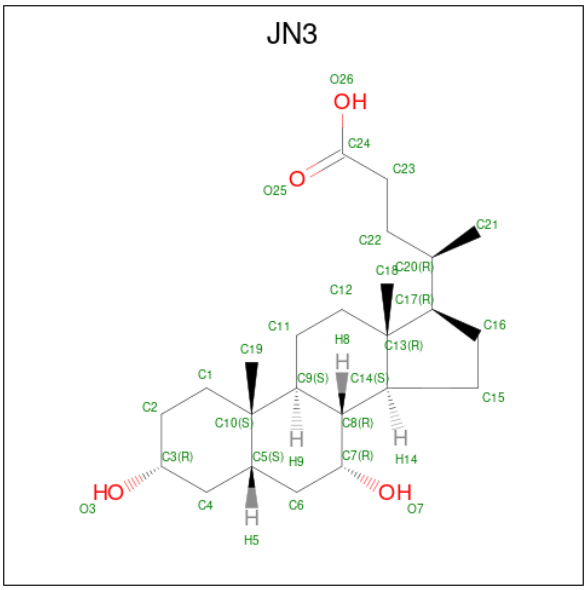
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			7	2	3	2		
3	B	1	Total	C	H	O	0	0
			7	2	3	2		
3	B	1	Total	C	H	O	0	0
			7	2	3	2		
3	B	1	Total	C	H	O	0	0
			7	2	3	2		
3	C	1	Total	C	H	O	0	0
			7	2	3	2		
3	D	1	Total	C	H	O	0	0
			7	2	3	2		
3	D	1	Total	C	H	O	0	0
			7	2	3	2		

Continued on next page...

Continued from previous page...

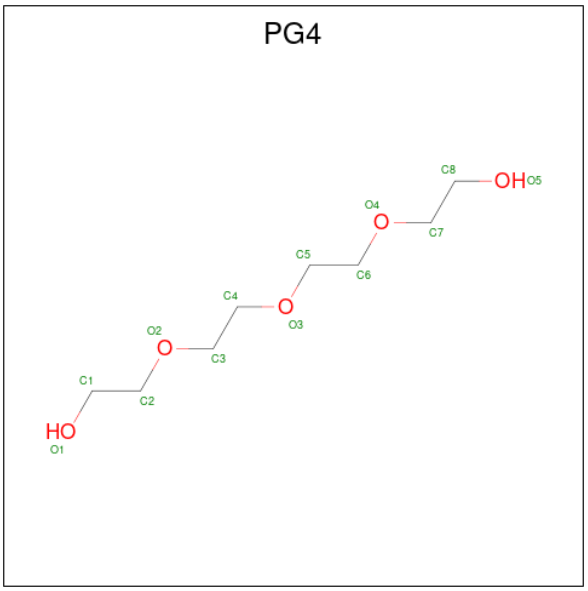
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	H	1	Total C H O 7 2 3 2	0	0

- Molecule 4 is CHENODEOXYCHOLIC ACID (three-letter code: JN3) (formula: $C_{24}H_{40}O_4$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C H O 67 24 39 4	0	0
4	D	1	Total C H O 67 24 39 4	0	0
4	F	1	Total C H O 67 24 39 4	0	0
4	H	1	Total C H O 67 24 39 4	0	0

- Molecule 5 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C O 13 8 5	0	0
5	D	1	Total C O 13 8 5	0	0
5	F	1	Total C O 13 8 5	0	0
5	H	1	Total C O 13 8 5	0	0

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total Ca 1 1	0	0
6	D	1	Total Ca 1 1	0	0
6	F	1	Total Ca 1 1	0	0
6	H	1	Total Ca 1 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	46	Total O 46 46	0	0

Continued on next page...

Continued from previous page...

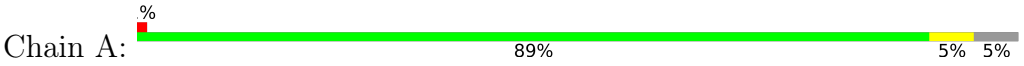
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	62	Total O 62 62	0	0
7	C	51	Total O 51 51	0	0
7	D	52	Total O 52 52	0	0
7	E	9	Total O 9 9	0	0
7	F	20	Total O 20 20	0	0
7	G	4	Total O 4 4	0	0
7	H	16	Total O 16 16	0	0

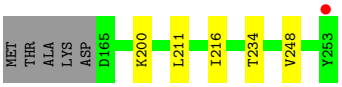
3 Residue-property plots (i)

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: VtrA

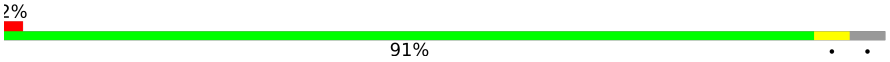
Chain A:





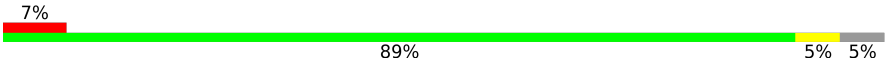
- Molecule 1: VtrA

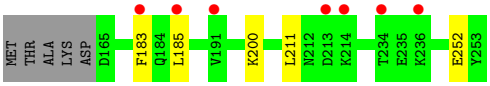
Chain C:



- Molecule 1: VtrA

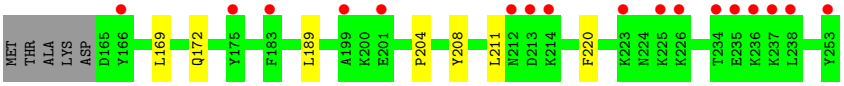
Chain E:





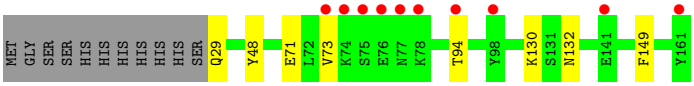
- Molecule 1: VtrA

Chain G:



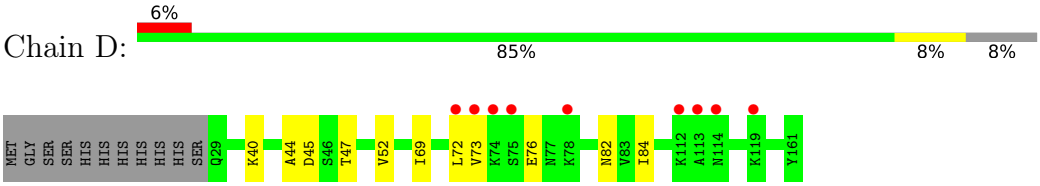
- Molecule 2: VtrC

Chain B:



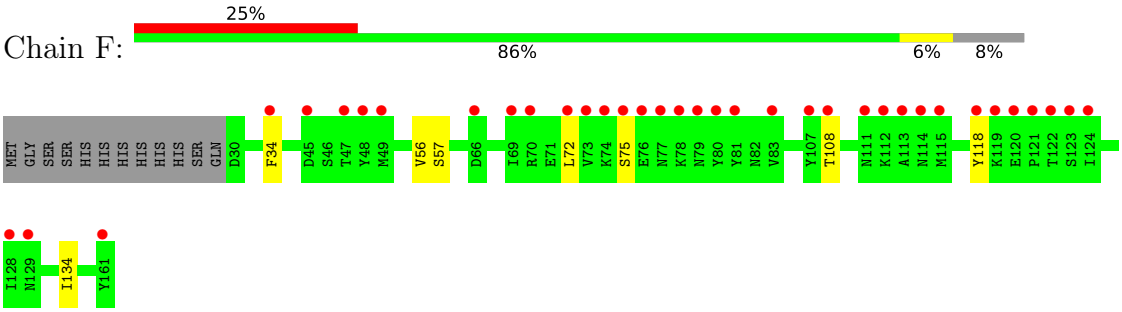
- Molecule 2: VtrC

Chain D:



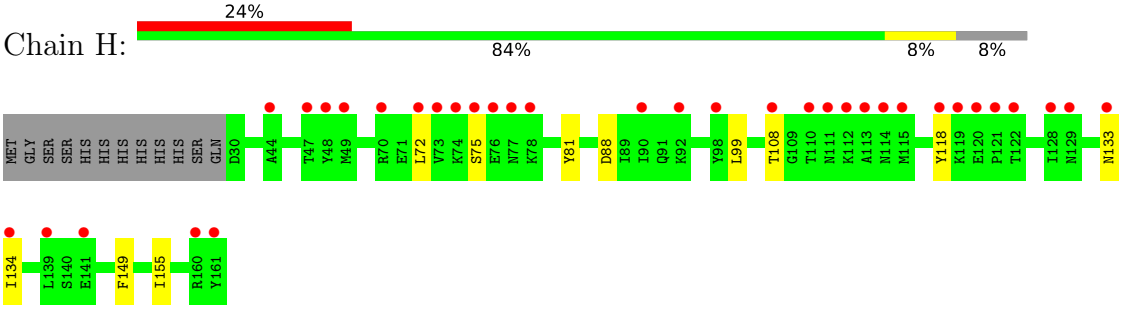
• Molecule 2: VtrC

Chain F:



• Molecule 2: VtrC

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	142.60Å 41.76Å 168.96Å 90.00° 91.57° 90.00°	Depositor
Resolution (Å)	44.78 – 2.08 44.78 – 2.08	Depositor EDS
% Data completeness (in resolution range)	88.9 (44.78-2.08) 88.9 (44.78-2.08)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.82 (at 2.08Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.188 , 0.224 0.187 , 0.223	Depositor DCC
R_{free} test set	2007 reflections (3.70%)	wwPDB-VP
Wilson B-factor (Å^2)	25.8	Xtriage
Anisotropy	0.249	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.42 , 58.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.034 for -h,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	15180	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 11.18% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA, JN3, EDO, PG4

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	0.73	0/750	0.60	0/1010
1	C	0.66	0/767	0.58	0/1032
1	E	0.48	0/751	0.51	0/1010
1	G	0.38	0/751	0.48	0/1010
2	B	0.64	0/1135	0.61	0/1539
2	D	0.60	0/1135	0.61	0/1539
2	F	0.39	0/1126	0.51	0/1527
2	H	0.33	0/1126	0.49	0/1527
All	All	0.54	0/7541	0.55	0/10194

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

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	736	737	737	7	0
1	C	750	747	747	2	0
1	E	737	737	737	4	0
1	G	737	737	737	4	0
2	B	1107	1057	1057	7	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	1107	1057	1057	10	0
2	F	1098	1049	1049	5	0
2	H	1098	1049	1049	9	0
3	A	4	3	6	0	0
3	B	12	9	18	3	0
3	C	4	3	6	0	0
3	D	8	6	12	0	0
3	H	4	3	6	0	0
4	B	28	39	38	0	0
4	D	28	39	37	1	0
4	F	28	39	37	2	0
4	H	28	39	37	1	0
5	B	13	0	18	0	0
5	D	13	0	18	1	0
5	F	13	0	18	2	0
5	H	13	0	18	0	0
6	B	1	0	0	0	0
6	D	1	0	0	0	0
6	F	1	0	0	0	0
6	H	1	0	0	0	0
7	A	46	0	0	0	0
7	B	62	0	0	0	0
7	C	51	0	0	0	0
7	D	52	0	0	0	0
7	E	9	0	0	0	0
7	F	20	0	0	0	0
7	G	4	0	0	0	0
7	H	16	0	0	0	0
All	All	7830	7350	7439	48	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 (48) 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)
4:F:201:JN3:H212	4:F:201:JN3:H183	1.68	0.73
1:A:211:LEU:HD22	1:A:216:ILE:HD12	1.80	0.62
4:D:201:JN3:H183	4:D:201:JN3:H212	1.84	0.58
2:H:99:LEU:HD13	2:H:134:ILE:HG13	1.85	0.57
3:B:206:EDO:H12	2:D:40:LYS:HE2	1.86	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:185:LEU:HD21	1:E:211:LEU:HD13	1.85	0.57
2:H:134:ILE:HD11	2:H:149:PHE:CE1	2.43	0.54
2:F:34:PHE:HD1	2:F:57:SER:HG	1.54	0.53
1:A:234:THR:O	1:A:234:THR:HG23	2.09	0.52
1:A:211:LEU:CD2	1:A:216:ILE:HD12	2.39	0.52
2:H:133:ASN:O	2:H:134:ILE:HD13	2.10	0.52
2:H:99:LEU:HD13	2:H:134:ILE:CG1	2.41	0.51
1:E:185:LEU:CD2	1:E:211:LEU:HD13	2.40	0.50
2:H:81:TYR:HD1	2:H:108:THR:HG22	1.76	0.50
1:C:235:GLU:HG3	2:D:44:ALA:HB1	1.93	0.50
2:D:45:ASP:HB3	2:D:47:THR:HG23	1.92	0.50
2:F:134:ILE:HD13	5:F:202:PG4:H42	1.93	0.50
3:B:206:EDO:H12	2:D:40:LYS:CD	2.42	0.49
2:B:71:GLU:HG2	2:B:73:VAL:HG13	1.95	0.49
2:F:108:THR:HG22	2:F:118:TYR:HB2	1.94	0.49
1:A:200:LYS:HD2	1:A:248:VAL:HG13	1.94	0.48
2:D:73:VAL:O	2:D:73:VAL:HG23	2.14	0.47
2:B:94:THR:HG22	2:B:94:THR:O	2.15	0.47
1:G:204:PRO:HB3	1:G:220:PHE:CE2	2.50	0.47
2:B:48:TYR:CB	2:B:73:VAL:HG12	2.46	0.45
1:E:200:LYS:NZ	1:E:252:GLU:OE1	2.47	0.45
1:A:200:LYS:HZ1	1:A:248:VAL:HG12	1.81	0.45
5:D:203:PG4:H61	5:D:203:PG4:H41	1.62	0.45
2:B:29:GLN:HB3	2:D:72:LEU:HD21	1.98	0.45
4:F:201:JN3:H212	4:F:201:JN3:C18	2.43	0.44
1:G:169:LEU:HD21	1:G:172:GLN:HB2	1.99	0.44
1:A:234:THR:O	1:A:234:THR:CG2	2.67	0.43
2:F:56:VAL:HG21	5:F:202:PG4:H21	1.99	0.43
2:H:72:LEU:HD21	2:H:75:SER:H	1.82	0.43
2:H:134:ILE:HD11	2:H:149:PHE:HE1	1.84	0.43
2:B:94:THR:HG21	1:E:183:PHE:CD2	2.53	0.43
1:A:200:LYS:NZ	1:A:248:VAL:HG12	2.33	0.43
1:C:167:PRO:HD2	1:C:210:MET:HE3	2.00	0.43
2:D:52:VAL:HG22	2:D:69:ILE:HG12	1.99	0.43
2:D:82:ASN:ND2	2:D:84:ILE:HD11	2.34	0.42
2:B:48:TYR:HB2	2:B:73:VAL:HG12	2.01	0.42
3:B:206:EDO:H12	2:D:40:LYS:CE	2.49	0.42
2:D:72:LEU:HD12	2:D:76:GLU:O	2.19	0.42
1:G:208:TYR:HB3	2:H:155:ILE:HD13	2.02	0.42
1:G:189:LEU:HD23	1:G:211:LEU:HD11	2.03	0.41
2:F:72:LEU:HD21	2:F:75:SER:HA	2.01	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:132:ASN:HB3	2:B:149:PHE:CD2	2.56	0.40
2:H:118:TYR:HE2	4:H:201:JN3:H211	1.86	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	87/94 (93%)	85 (98%)	2 (2%)	0	100 100
1	C	89/94 (95%)	85 (96%)	3 (3%)	1 (1%)	14 8
1	E	87/94 (93%)	86 (99%)	1 (1%)	0	100 100
1	G	87/94 (93%)	85 (98%)	2 (2%)	0	100 100
2	B	131/144 (91%)	127 (97%)	4 (3%)	0	100 100
2	D	131/144 (91%)	127 (97%)	4 (3%)	0	100 100
2	F	130/144 (90%)	123 (95%)	7 (5%)	0	100 100
2	H	130/144 (90%)	123 (95%)	7 (5%)	0	100 100
All	All	872/952 (92%)	841 (96%)	30 (3%)	1 (0%)	51 53

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	213	ASP

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	85/89 (96%)	85 (100%)	0	100	100
1	C	87/89 (98%)	87 (100%)	0	100	100
1	E	85/89 (96%)	85 (100%)	0	100	100
1	G	85/89 (96%)	85 (100%)	0	100	100
2	B	124/134 (92%)	123 (99%)	1 (1%)	81	85
2	D	124/134 (92%)	124 (100%)	0	100	100
2	F	123/134 (92%)	123 (100%)	0	100	100
2	H	123/134 (92%)	122 (99%)	1 (1%)	81	85
All	All	836/892 (94%)	834 (100%)	2 (0%)	93	95

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	130	LYS
2	H	88	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 20 ligands modelled in this entry, 4 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	$\#\|Z\| > 2$
4	JN3	F	201	-	31,31,31	1.73	8 (25%)	49,49,49	1.78	12 (24%)
3	EDO	H	204	-	3,3,3	0.50	0	2,2,2	0.40	0
5	PG4	H	202	-	12,12,12	0.12	0	11,11,11	0.66	0
4	JN3	D	201	-	31,31,31	2.07	10 (32%)	49,49,49	2.04	17 (34%)
3	EDO	B	206	-	3,3,3	0.57	0	2,2,2	0.47	0
3	EDO	C	301	-	3,3,3	0.48	0	2,2,2	0.23	0
3	EDO	A	301	-	3,3,3	0.57	0	2,2,2	0.52	0
5	PG4	D	203	-	12,12,12	0.26	0	11,11,11	0.61	0
3	EDO	B	205	-	3,3,3	0.53	0	2,2,2	0.37	0
4	JN3	B	201	-	31,31,31	1.76	7 (22%)	49,49,49	1.59	11 (22%)
5	PG4	B	203	-	12,12,12	0.19	0	11,11,11	0.57	0
5	PG4	F	202	-	12,12,12	0.17	0	11,11,11	0.63	0
3	EDO	D	205	-	3,3,3	0.60	0	2,2,2	0.41	0
3	EDO	B	202	-	3,3,3	0.48	0	2,2,2	0.51	0
4	JN3	H	201	-	31,31,31	1.80	7 (22%)	49,49,49	1.91	12 (24%)
3	EDO	D	202	-	3,3,3	0.58	0	2,2,2	0.49	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
4	JN3	F	201	-	-	8/9/70/70	0/4/4/4
3	EDO	H	204	-	-	0/1/1/1	-
5	PG4	H	202	-	-	3/10/10/10	-
4	JN3	D	201	-	-	6/9/70/70	0/4/4/4
3	EDO	B	206	-	-	0/1/1/1	-
3	EDO	C	301	-	-	0/1/1/1	-
3	EDO	A	301	-	-	1/1/1/1	-
5	PG4	D	203	-	-	3/10/10/10	-
3	EDO	B	205	-	-	0/1/1/1	-
4	JN3	B	201	-	-	2/9/70/70	0/4/4/4

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	PG4	B	203	-	-	6/10/10/10	-
5	PG4	F	202	-	-	6/10/10/10	-
3	EDO	D	205	-	-	0/1/1/1	-
3	EDO	B	202	-	-	0/1/1/1	-
4	JN3	H	201	-	-	0/9/70/70	0/4/4/4
3	EDO	D	202	-	-	0/1/1/1	-

All (32) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	201	JN3	C6-C5	-4.79	1.46	1.53
4	H	201	JN3	O7-C7	-4.60	1.33	1.43
4	D	201	JN3	O7-C7	-4.23	1.34	1.43
4	F	201	JN3	O7-C7	-4.22	1.34	1.43
4	B	201	JN3	C6-C5	-4.15	1.47	1.53
4	B	201	JN3	C13-C14	3.56	1.61	1.55
4	D	201	JN3	C8-C14	-3.51	1.47	1.53
4	D	201	JN3	C12-C11	-3.35	1.46	1.53
4	H	201	JN3	C6-C5	-3.34	1.48	1.53
4	H	201	JN3	C15-C14	-3.27	1.47	1.54
4	B	201	JN3	O7-C7	-3.21	1.36	1.43
4	F	201	JN3	C6-C5	-3.15	1.48	1.53
4	D	201	JN3	C8-C7	3.07	1.58	1.53
4	D	201	JN3	C23-C24	3.07	1.57	1.50
4	F	201	JN3	C8-C7	3.06	1.58	1.53
4	H	201	JN3	C10-C5	-3.03	1.50	1.55
4	B	201	JN3	C18-C13	-2.87	1.49	1.54
4	B	201	JN3	C21-C20	-2.70	1.46	1.53
4	F	201	JN3	C15-C14	-2.68	1.48	1.54
4	B	201	JN3	C15-C14	-2.61	1.49	1.54
4	D	201	JN3	O3-C3	-2.60	1.35	1.43
4	F	201	JN3	C8-C14	-2.50	1.49	1.53
4	B	201	JN3	C10-C5	-2.39	1.51	1.55
4	H	201	JN3	C23-C24	2.39	1.56	1.50
4	F	201	JN3	C21-C20	-2.37	1.47	1.53
4	D	201	JN3	C4-C5	-2.37	1.50	1.53
4	F	201	JN3	O3-C3	-2.27	1.36	1.43
4	H	201	JN3	O3-C3	-2.24	1.36	1.43
4	H	201	JN3	C8-C7	2.24	1.57	1.53
4	D	201	JN3	C21-C20	-2.23	1.47	1.53
4	F	201	JN3	C1-C2	-2.09	1.49	1.53
4	D	201	JN3	C13-C14	2.01	1.58	1.55

All (52) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	201	JN3	C6-C7-C8	-4.99	106.16	111.48
4	H	201	JN3	C6-C7-C8	-4.71	106.45	111.48
4	H	201	JN3	C4-C5-C10	4.53	117.47	112.66
4	D	201	JN3	O7-C7-C8	4.30	119.03	109.43
4	H	201	JN3	O7-C7-C8	4.29	119.02	109.43
4	F	201	JN3	C9-C8-C7	-4.07	107.00	111.88
4	F	201	JN3	C6-C7-C8	-4.06	107.14	111.48
4	D	201	JN3	C21-C20-C22	-4.06	104.00	110.36
4	H	201	JN3	C12-C13-C14	-4.02	101.03	107.27
4	F	201	JN3	O7-C7-C8	3.90	118.14	109.43
4	B	201	JN3	O7-C7-C8	3.89	118.13	109.43
4	F	201	JN3	C4-C5-C10	3.88	116.78	112.66
4	D	201	JN3	C4-C3-C2	-3.70	106.13	110.55
4	D	201	JN3	C10-C9-C8	3.53	115.61	111.82
4	D	201	JN3	C2-C1-C10	3.48	118.75	112.78
4	H	201	JN3	C4-C3-C2	-3.42	106.47	110.55
4	F	201	JN3	C12-C13-C14	-3.35	102.08	107.27
4	B	201	JN3	C4-C5-C10	3.30	116.16	112.66
4	B	201	JN3	C6-C5-C4	-3.22	107.48	111.19
4	D	201	JN3	C9-C8-C7	-3.12	108.14	111.88
4	B	201	JN3	C19-C10-C9	3.03	115.36	111.18
4	F	201	JN3	C13-C17-C20	2.98	124.15	119.49
4	H	201	JN3	C10-C9-C8	2.87	114.90	111.82
4	D	201	JN3	C4-C5-C10	2.86	115.69	112.66
4	B	201	JN3	C6-C5-C10	2.81	115.64	112.66
4	H	201	JN3	C11-C9-C10	2.75	117.09	113.91
4	B	201	JN3	C1-C2-C3	2.73	113.97	110.47
4	H	201	JN3	O7-C7-C6	2.58	116.35	109.94
4	H	201	JN3	C2-C1-C10	2.57	117.18	112.78
4	F	201	JN3	C14-C8-C9	2.51	113.15	109.71
4	D	201	JN3	C11-C12-C13	2.51	117.08	112.78
4	D	201	JN3	C16-C17-C13	-2.47	100.87	103.84
4	F	201	JN3	C5-C4-C3	2.45	116.36	112.76
4	H	201	JN3	C1-C2-C3	2.40	113.55	110.47
4	D	201	JN3	C13-C17-C20	2.39	123.23	119.49
4	F	201	JN3	C2-C1-C10	2.39	116.88	112.78
4	D	201	JN3	C6-C5-C10	-2.37	110.14	112.66
4	H	201	JN3	C11-C12-C13	2.27	116.68	112.78
4	D	201	JN3	O3-C3-C4	2.27	114.37	109.85
4	B	201	JN3	C12-C13-C14	-2.20	103.86	107.27
4	D	201	JN3	C5-C6-C7	2.19	116.88	114.46
4	H	201	JN3	C5-C4-C3	2.18	115.96	112.76

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	201	JN3	C4-C3-C2	-2.17	107.96	110.55
4	D	201	JN3	C12-C11-C9	2.17	116.88	113.11
4	B	201	JN3	C2-C1-C10	2.14	116.46	112.78
4	D	201	JN3	C12-C13-C17	2.11	119.72	116.57
4	B	201	JN3	C12-C13-C17	2.10	119.72	116.57
4	F	201	JN3	C11-C9-C10	2.10	116.34	113.91
4	F	201	JN3	C21-C20-C22	-2.08	107.10	110.36
4	B	201	JN3	C5-C6-C7	2.07	116.75	114.46
4	D	201	JN3	C12-C13-C14	-2.06	104.07	107.27
4	F	201	JN3	O7-C7-C6	2.01	114.92	109.94

There are no chirality outliers.

All (35) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	201	JN3	C13-C17-C20-C21
4	F	201	JN3	C16-C17-C20-C22
4	F	201	JN3	C16-C17-C20-C21
4	F	201	JN3	C13-C17-C20-C21
4	D	201	JN3	C16-C17-C20-C22
4	D	201	JN3	C16-C17-C20-C21
4	D	201	JN3	C13-C17-C20-C22
4	F	201	JN3	C13-C17-C20-C22
5	D	203	PG4	C6-C5-O3-C4
5	F	202	PG4	O3-C5-C6-O4
5	F	202	PG4	C1-C2-O2-C3
5	B	203	PG4	O2-C3-C4-O3
5	B	203	PG4	O4-C7-C8-O5
3	A	301	EDO	O1-C1-C2-O2
5	H	202	PG4	O2-C3-C4-O3
5	D	203	PG4	O3-C5-C6-O4
4	D	201	JN3	C20-C22-C23-C24
5	F	202	PG4	O4-C7-C8-O5
4	F	201	JN3	C17-C20-C22-C23
5	B	203	PG4	C4-C3-O2-C2
5	D	203	PG4	C8-C7-O4-C6
5	H	202	PG4	C3-C4-O3-C5
4	D	201	JN3	C21-C20-C22-C23
5	F	202	PG4	C4-C3-O2-C2
5	H	202	PG4	C1-C2-O2-C3
4	F	201	JN3	C21-C20-C22-C23
5	B	203	PG4	C5-C6-O4-C7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	B	201	JN3	C22-C23-C24-O26
4	F	201	JN3	C22-C23-C24-O26
4	F	201	JN3	C22-C23-C24-O25
4	B	201	JN3	C22-C23-C24-O25
5	B	203	PG4	C3-C4-O3-C5
5	F	202	PG4	C6-C5-O3-C4
5	F	202	PG4	C8-C7-O4-C6
5	B	203	PG4	O3-C5-C6-O4

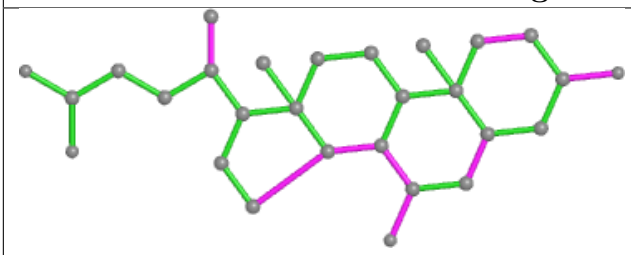
There are no ring outliers.

6 monomers are involved in 10 short contacts:

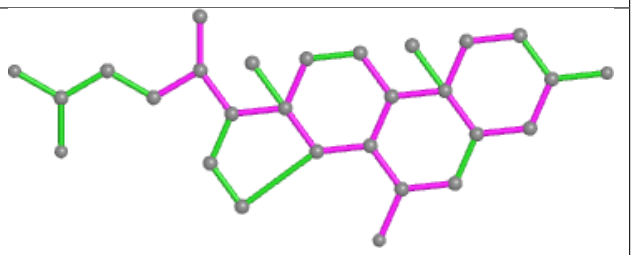
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	201	JN3	2	0
4	D	201	JN3	1	0
3	B	206	EDO	3	0
5	D	203	PG4	1	0
5	F	202	PG4	2	0
4	H	201	JN3	1	0

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.

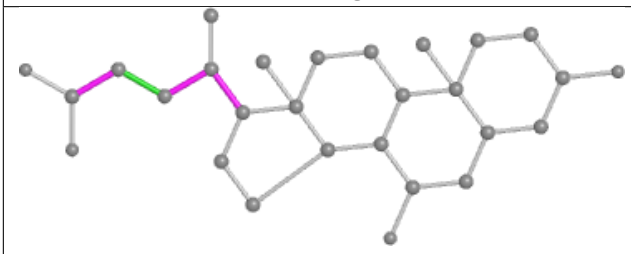
Ligand JN3 F 201



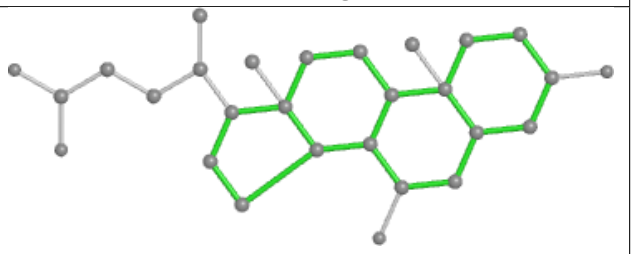
Bond lengths



Bond angles

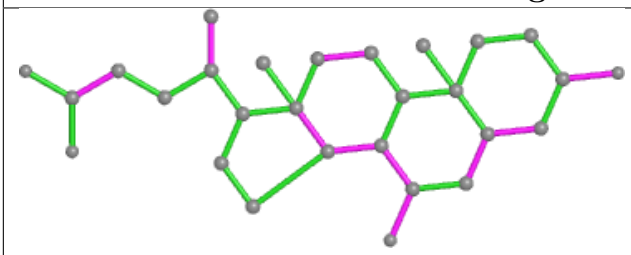


Torsions

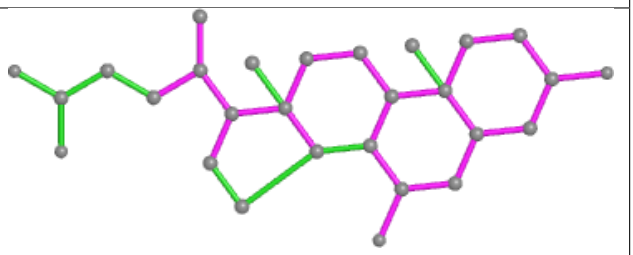


Rings

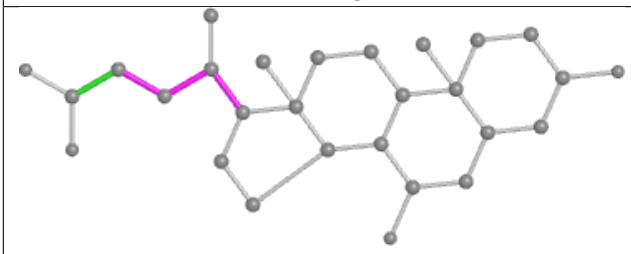
Ligand JN3 D 201



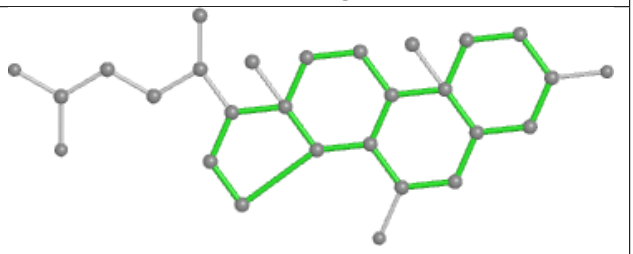
Bond lengths



Bond angles

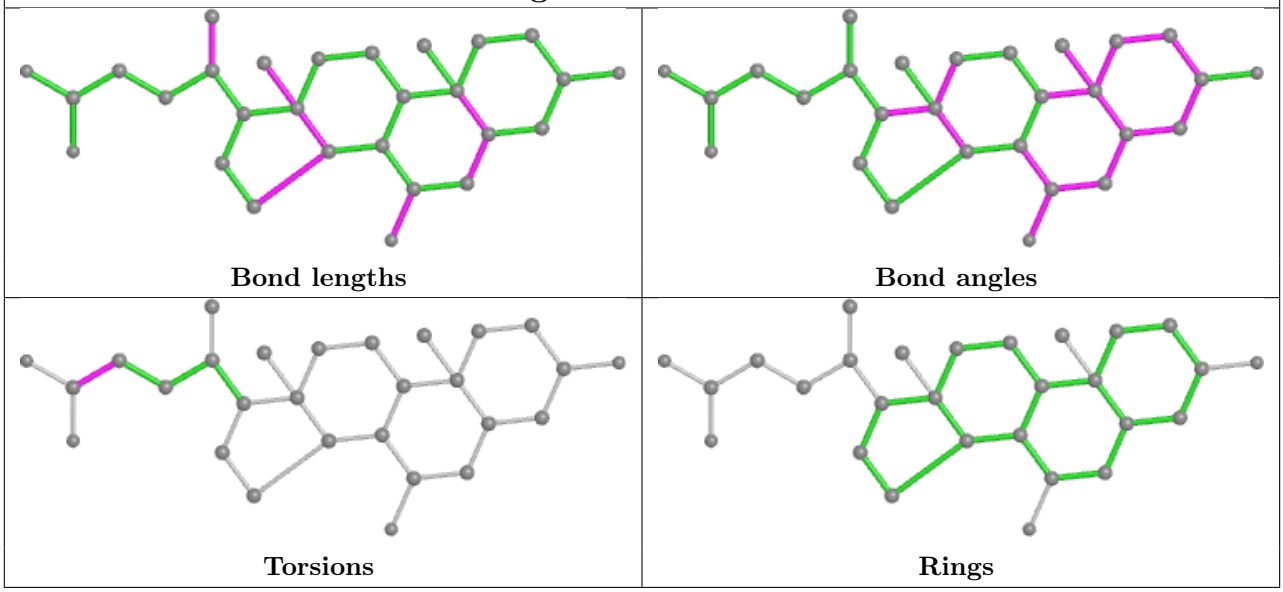


Torsions

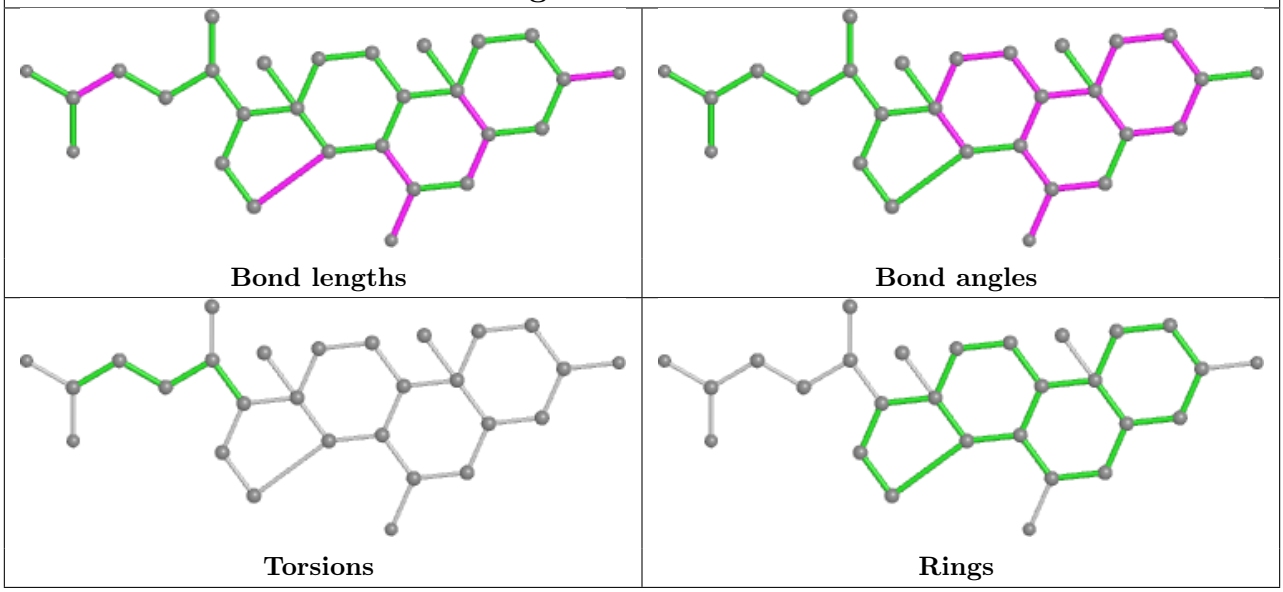


Rings

Ligand JN3 B 201



Ligand JN3 H 201



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	89/94 (94%)	0.38	1 (1%) 80 83	11, 21, 46, 64	0
1	C	90/94 (95%)	0.20	2 (2%) 62 66	13, 22, 51, 71	0
1	E	89/94 (94%)	0.67	7 (7%) 12 15	25, 45, 76, 93	0
1	G	89/94 (94%)	1.17	17 (19%) 1 1	42, 68, 95, 100	0
2	B	133/144 (92%)	0.57	10 (7%) 14 17	11, 28, 61, 104	0
2	D	133/144 (92%)	0.41	9 (6%) 17 21	14, 27, 69, 89	0
2	F	132/144 (91%)	1.22	36 (27%) 0 0	26, 62, 99, 115	0
2	H	132/144 (91%)	1.33	35 (26%) 0 0	38, 65, 109, 124	0
All	All	887/952 (93%)	0.77	117 (13%) 3 4	11, 43, 95, 124	0

All (117) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	113	ALA	9.4
2	H	129	ASN	7.9
2	H	74	LYS	6.7
1	G	234	THR	6.1
2	F	113	ALA	6.0
2	F	74	LYS	5.6
2	B	73	VAL	5.4
2	H	77	ASN	5.3
1	G	238	LEU	5.1
2	B	74	LYS	4.7
2	F	111	ASN	4.6
2	H	139	LEU	4.5
1	G	235	GLU	4.4
2	H	118	TYR	4.4
2	F	72	LEU	4.3
2	F	73	VAL	4.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	161	TYR	4.2
2	D	73	VAL	4.2
2	F	129	ASN	4.2
2	H	121	PRO	4.1
1	G	223	LYS	4.1
2	H	75	SER	4.0
2	B	75	SER	4.0
1	G	213	ASP	4.0
2	F	114	ASN	4.0
2	F	80	TYR	3.8
1	G	226	LYS	3.8
2	F	115	MET	3.7
2	H	122	THR	3.7
2	H	72	LEU	3.7
2	H	128	ILE	3.7
2	F	107	TYR	3.6
2	F	108	THR	3.6
1	C	253	TYR	3.6
2	D	72	LEU	3.6
2	F	120	GLU	3.6
2	F	112	LYS	3.5
2	F	123	SER	3.5
2	H	47	THR	3.5
2	H	110	THR	3.5
2	B	76	GLU	3.5
2	H	133	ASN	3.5
1	G	199	ALA	3.5
2	H	134	ILE	3.4
1	G	201	GLU	3.4
1	G	253	TYR	3.3
2	H	114	ASN	3.3
2	D	78	LYS	3.3
2	H	73	VAL	3.2
2	F	118	TYR	3.2
1	E	213	ASP	3.2
2	H	98	TYR	3.1
2	B	161	TYR	3.1
2	H	141	GLU	3.1
1	G	236	LYS	3.0
2	F	66	ASP	3.0
2	H	111	ASN	2.9
1	E	214	LYS	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	H	120	GLU	2.9
1	G	166	TYR	2.9
2	H	112	LYS	2.9
1	E	183	PHE	2.8
2	F	78	LYS	2.8
2	F	69	ILE	2.8
1	G	237	LYS	2.8
2	F	48	TYR	2.8
2	F	124	ILE	2.7
2	H	119	LYS	2.7
2	B	78	LYS	2.7
2	D	112	LYS	2.7
1	G	183	PHE	2.7
1	A	253	TYR	2.7
2	F	76	GLU	2.7
2	H	76	GLU	2.7
2	F	49	MET	2.7
2	D	75	SER	2.7
2	H	90	ILE	2.7
2	H	49	MET	2.7
2	H	161	TYR	2.6
2	H	48	TYR	2.6
2	F	128	ILE	2.6
1	E	191	VAL	2.6
2	F	75	SER	2.6
2	D	74	LYS	2.6
2	H	70	ARG	2.5
2	H	160	ARG	2.5
2	D	113	ALA	2.4
1	G	212	ASN	2.4
2	B	77	ASN	2.4
2	D	114	ASN	2.4
2	F	70	ARG	2.4
1	E	236	LYS	2.4
2	B	94	THR	2.4
2	B	98	TYR	2.4
1	G	214	LYS	2.4
1	C	214	LYS	2.4
1	E	234	THR	2.4
2	H	92	LYS	2.4
2	H	44	ALA	2.3
2	H	115	MET	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	45	ASP	2.3
2	F	121	PRO	2.3
2	B	141	GLU	2.3
2	F	122	THR	2.2
2	D	119	LYS	2.2
2	F	119	LYS	2.2
2	F	34	PHE	2.2
2	F	81	TYR	2.2
2	F	83	VAL	2.2
2	H	108	THR	2.2
1	E	185	LEU	2.1
2	H	78	LYS	2.1
2	F	77	ASN	2.1
1	G	225	LYS	2.1
2	F	47	THR	2.1
1	G	175	TYR	2.0
2	F	79	ASN	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 monosaccharides in this entry.

6.4 Ligands (i)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	CA	H	203	1/1	0.46	0.10	92,92,92,92	0
3	EDO	B	206	4/4	0.81	0.25	40,43,53,53	0
3	EDO	A	301	4/4	0.83	0.22	36,39,55,55	0
4	JN3	F	201	28/28	0.84	0.23	39,63,88,96	0
3	EDO	D	205	4/4	0.86	0.22	36,43,51,54	0
3	EDO	C	301	4/4	0.87	0.19	36,43,54,54	0

Continued on next page...

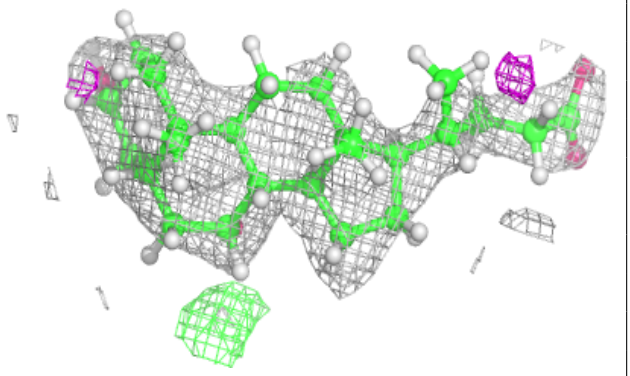
Continued from previous page...

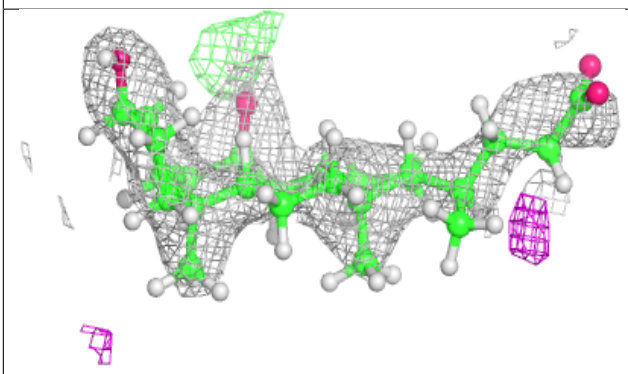
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	PG4	D	203	13/13	0.88	0.16	32,36,41,43	0
4	JN3	H	201	28/28	0.88	0.24	47,64,91,100	0
5	PG4	H	202	13/13	0.89	0.29	48,55,64,70	0
5	PG4	B	203	13/13	0.90	0.16	32,38,46,49	0
6	CA	F	203	1/1	0.90	0.05	70,70,70,70	0
5	PG4	F	202	13/13	0.90	0.25	45,49,58,62	0
3	EDO	B	202	4/4	0.91	0.15	23,45,54,59	0
3	EDO	H	204	4/4	0.91	0.12	62,66,77,79	0
3	EDO	B	205	4/4	0.92	0.25	32,38,46,48	0
3	EDO	D	202	4/4	0.93	0.09	27,33,40,43	0
6	CA	D	204	1/1	0.94	0.12	31,31,31,31	0
4	JN3	B	201	28/28	0.94	0.14	17,24,55,60	0
4	JN3	D	201	28/28	0.94	0.13	19,28,65,75	0
6	CA	B	204	1/1	0.98	0.06	34,34,34,34	0

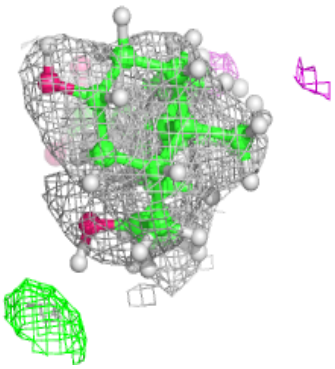
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around JN3 F 201:

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

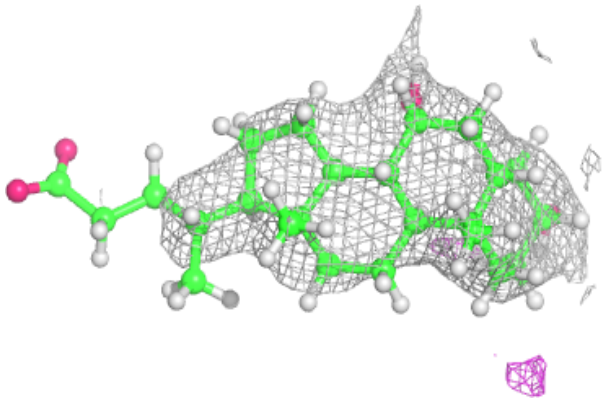


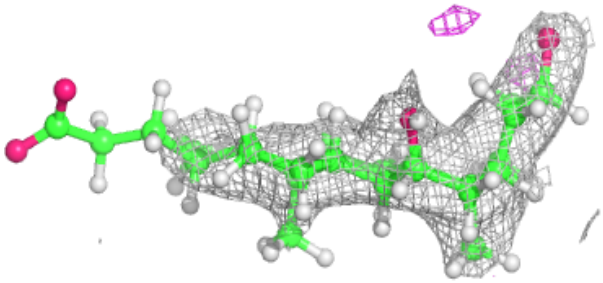


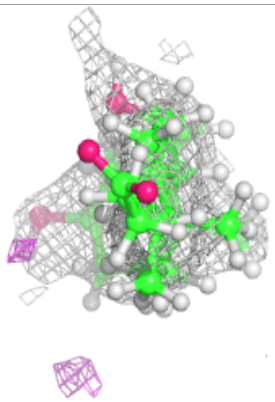


Electron density around JN3 H 201:

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

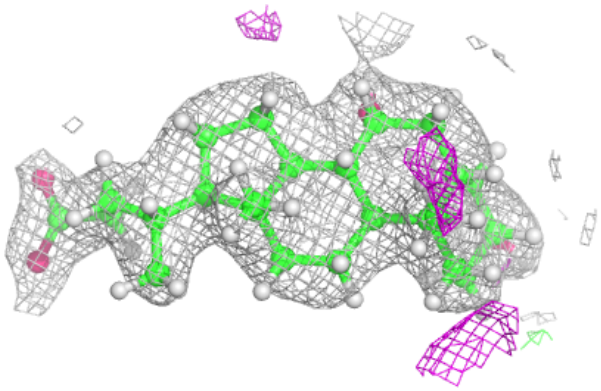


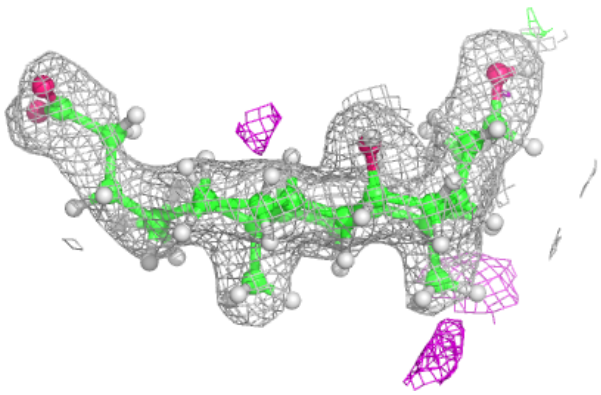


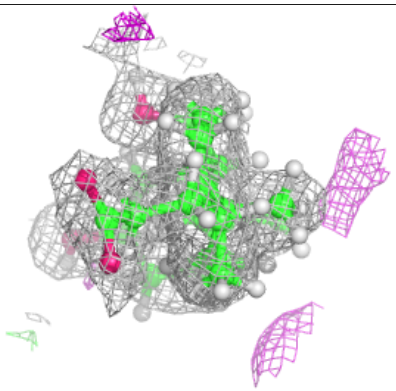


Electron density around JN3 B 201:

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

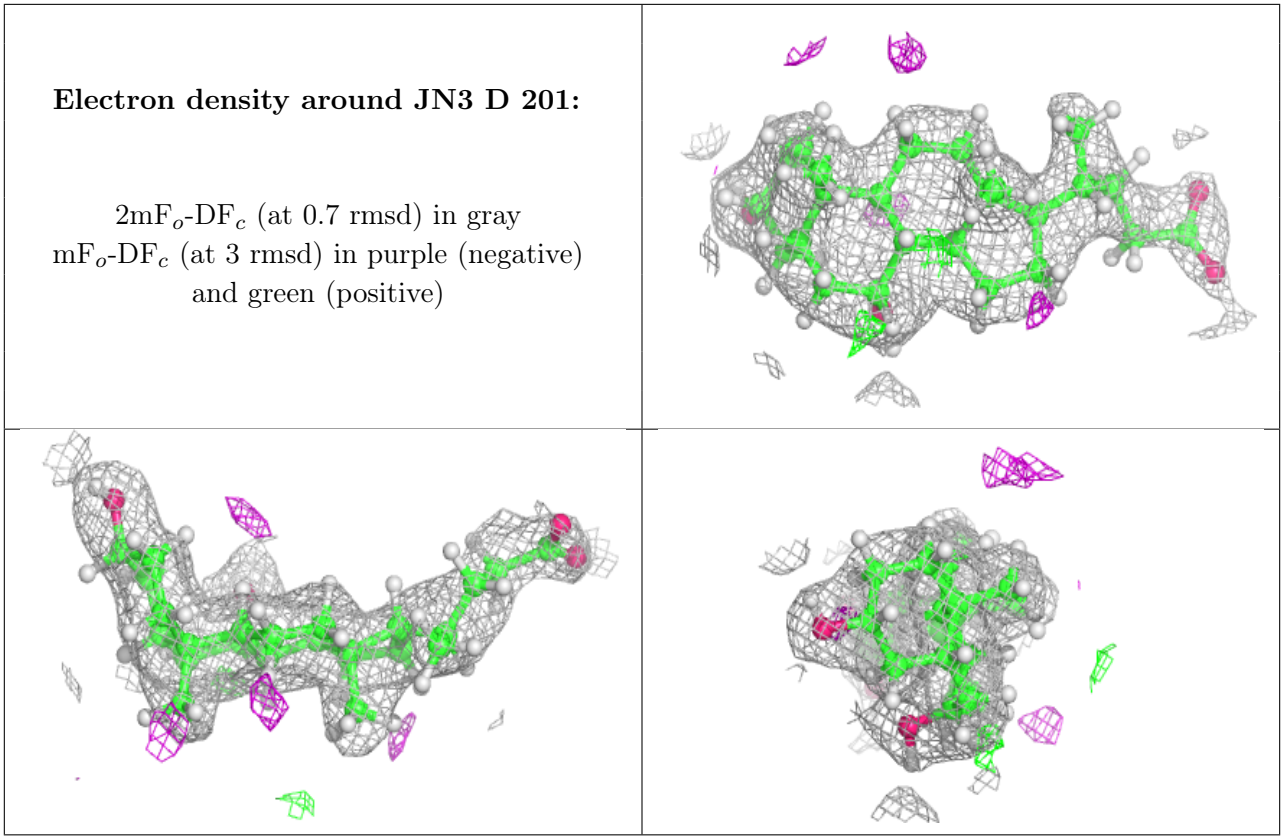






Electron density around JN3 D 201:

$2mF_o$-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.