



Full wwPDB X-ray Structure Validation Report ⓘ

Apr 20, 2021 – 08:14 pm BST

PDB ID : 6FV1
Title : Structure of human coronavirus NL63 main protease in complex with the alpha-ketoamide (S)-N-((S)-4-(benzylamino)-3,4-dioxo-1-((S)-2-oxopyrrolidin-3-yl)butan-2-yl)-2-cinnamamido-4-methylpentanamide (cinnamoyl-leucine-GlnLactam-CO-CO-NH-benzyl)
Authors : Zhang, L.; Hilgenfeld, R.
Deposited on : 2018-02-28
Resolution : 2.30 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 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 : 2.18
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) : 2.18

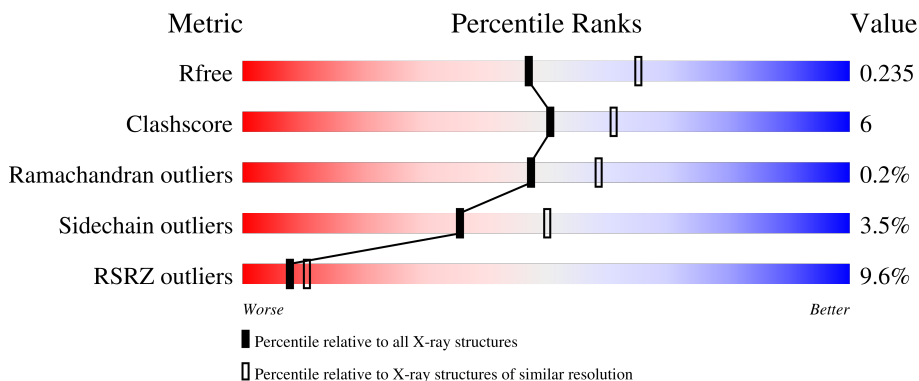
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	306	 89% 8% ..
1	B	306	 15% 80% 15% ..
1	C	306	 12% 83% 13% ..

2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 7438 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.

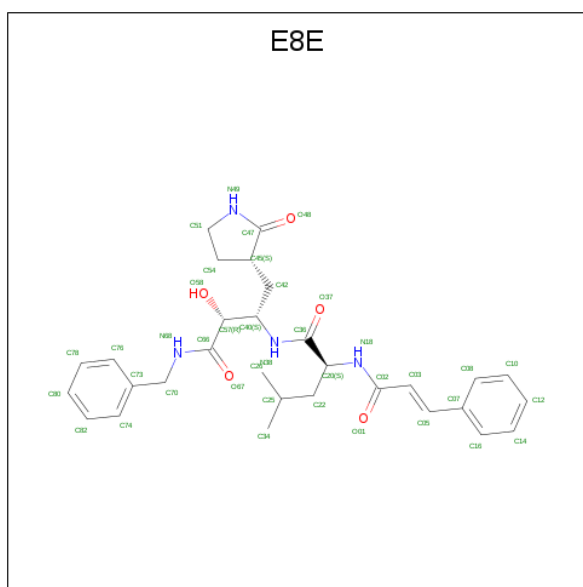
- Molecule 1 is a protein called 3C-like proteinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	300	2272	1439	389	427	17	0	0	0
1	B	299	2266	1434	388	427	17	0	0	0
1	C	299	2266	1434	388	427	17	0	0	0

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	302	HIS	-	expression tag	UNP P0C6X5
A	303	HIS	-	expression tag	UNP P0C6X5
A	304	HIS	-	expression tag	UNP P0C6X5
A	305	HIS	-	expression tag	UNP P0C6X5
A	306	HIS	-	expression tag	UNP P0C6X5
B	302	HIS	-	expression tag	UNP P0C6X5
B	303	HIS	-	expression tag	UNP P0C6X5
B	304	HIS	-	expression tag	UNP P0C6X5
B	305	HIS	-	expression tag	UNP P0C6X5
B	306	HIS	-	expression tag	UNP P0C6X5
C	302	HIS	-	expression tag	UNP P0C6X5
C	303	HIS	-	expression tag	UNP P0C6X5
C	304	HIS	-	expression tag	UNP P0C6X5
C	305	HIS	-	expression tag	UNP P0C6X5
C	306	HIS	-	expression tag	UNP P0C6X5

- Molecule 2 is (2 {S})-4-methyl- {N}-[(2 {S},3 {R})-3-oxidanyl-4-oxidanylidene-1-[(3 {S})-2-oxidanylidene-pyrrolidin-3-yl]-4-[(phenylmethyl)amino]butan-2-yl]-2-[[({E})-3-phenylprop-2-enoyl]amino]pentanamide (three-letter code: E8E) (formula: C₃₀H₃₈N₄O₅).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
			Total	C	N			O
2	A	1	39	30	4	5	0	0
2	B	1	39	30	4	5	0	0
2	C	1	39	30	4	5	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



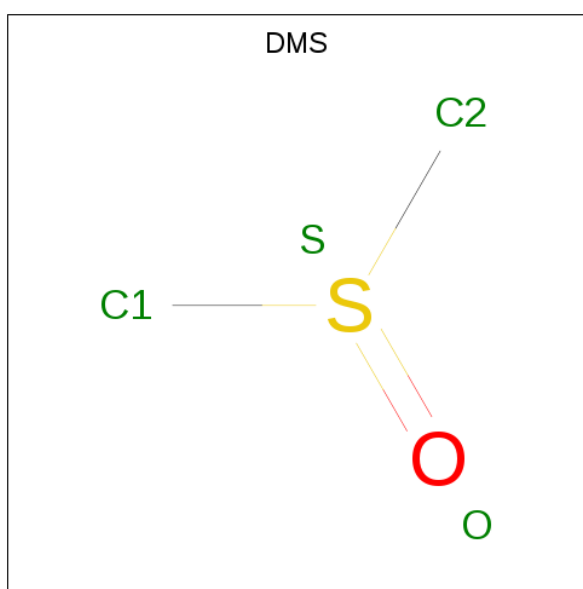
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	A	1	6	3	3	0	0

Continued on next page...

Continued from previous page...

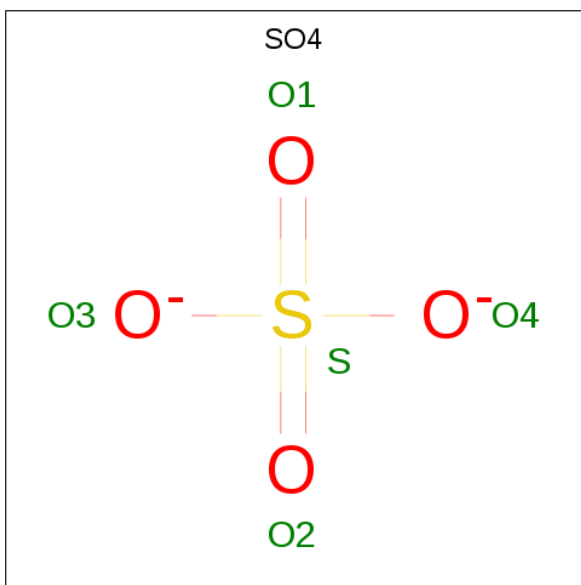
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C₂H₆OS).



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

- Molecule 5 is SULFATE ION (three-letter code: SO₄) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	S	0	0
			5	4	1		


- Molecule 6 is water.

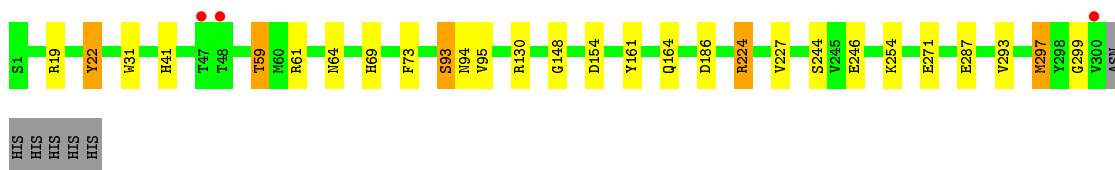
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	214	Total	O	0	0
			214	214		
6	B	131	Total	O	0	0
			131	131		
6	C	129	Total	O	0	0
			129	129		

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.

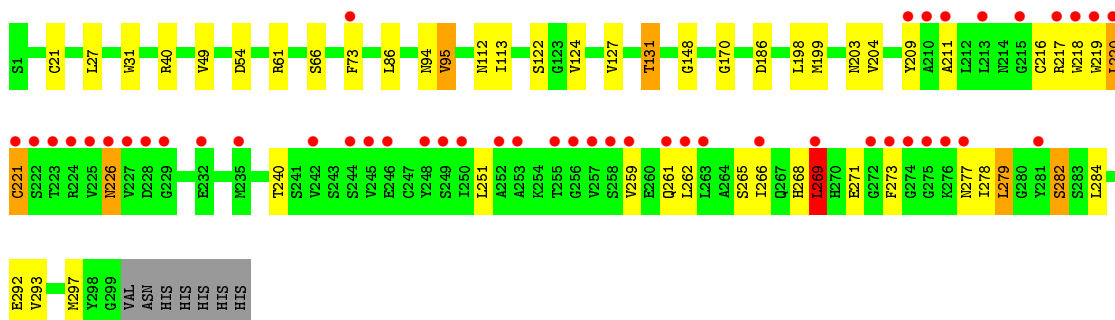
- Molecule 1: 3C-like proteinase

Chain A: 




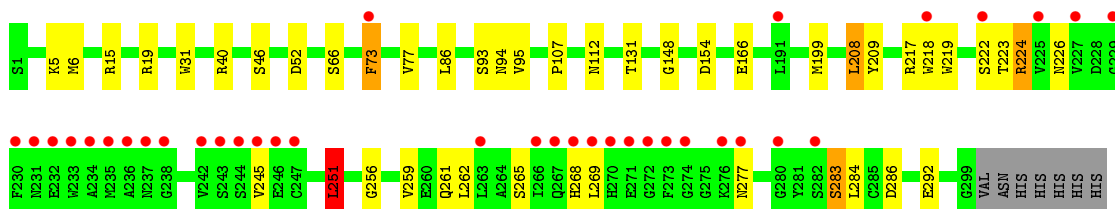
- Molecule 1: 3C-like proteinase

Chain B: 



- Molecule 1: 3C-like proteinase

Chain C: 



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	133.86Å 211.33Å 118.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.24 – 2.30 44.33 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.24-2.30) 99.9 (44.33-2.30)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.89 (at 2.29Å)	Xtrriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.195 , 0.231 0.200 , 0.235	Depositor DCC
R_{free} test set	3637 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å ²)	38.7	Xtrriage
Anisotropy	0.583	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 45.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7438	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.89% 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: DMS, SO4, E8E, GOL

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	1.16	7/2323 (0.3%)	1.10	13/3156 (0.4%)
1	B	1.03	6/2317 (0.3%)	1.05	7/3147 (0.2%)
1	C	1.08	1/2317 (0.0%)	1.11	13/3147 (0.4%)
All	All	1.09	14/6957 (0.2%)	1.09	33/9450 (0.3%)

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	73	PHE	CB-CG	-7.15	1.39	1.51
1	A	73	PHE	CB-CG	-7.05	1.39	1.51
1	B	73	PHE	CB-CG	-6.92	1.39	1.51
1	A	224	ARG	CZ-NH2	-5.94	1.25	1.33
1	A	287	GLU	CD-OE1	-5.70	1.19	1.25
1	A	22	TYR	CG-CD2	-5.66	1.31	1.39
1	B	122	SER	CB-OG	-5.55	1.35	1.42
1	B	127	VAL	CB-CG2	-5.54	1.41	1.52
1	B	131	THR	CB-CG2	-5.47	1.34	1.52
1	A	22	TYR	CE1-CZ	-5.39	1.31	1.38
1	A	22	TYR	CG-CD1	-5.36	1.32	1.39
1	B	271	GLU	CD-OE1	-5.23	1.19	1.25
1	A	224	ARG	CD-NE	-5.22	1.37	1.46
1	B	170	GLY	N-CA	5.19	1.53	1.46

All (33) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	286	ASP	CB-CG-OD1	8.83	126.25	118.30
1	B	61	ARG	NE-CZ-NH2	8.57	124.59	120.30
1	A	61	ARG	NE-CZ-NH1	7.33	123.96	120.30
1	A	19	ARG	NE-CZ-NH1	7.11	123.86	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	130	ARG	NE-CZ-NH2	-7.05	116.77	120.30
1	B	269	LEU	CB-CG-CD2	-6.99	99.12	111.00
1	B	269	LEU	CB-CG-CD1	6.95	122.82	111.00
1	C	19	ARG	NE-CZ-NH2	6.89	123.75	120.30
1	B	271	GLU	OE1-CD-OE2	-6.88	115.05	123.30
1	C	6	MET	CG-SD-CE	-6.82	89.29	100.20
1	C	251	LEU	CB-CG-CD2	-6.74	99.55	111.00
1	C	166	GLU	OE1-CD-OE2	-6.44	115.57	123.30
1	B	186	ASP	CB-CG-OD1	6.17	123.85	118.30
1	C	154	ASP	CB-CG-OD1	5.97	123.67	118.30
1	C	73	PHE	CB-CG-CD1	-5.94	116.64	120.80
1	C	224	ARG	NE-CZ-NH2	5.90	123.25	120.30
1	A	227	VAL	CG1-CB-CG2	-5.86	101.52	110.90
1	B	73	PHE	CB-CG-CD1	-5.79	116.75	120.80
1	C	52	ASP	CB-CG-OD1	5.68	123.41	118.30
1	A	224	ARG	CB-CA-C	-5.63	99.14	110.40
1	A	297	MET	CG-SD-CE	5.47	108.95	100.20
1	C	6	MET	CB-CA-C	-5.38	99.65	110.40
1	A	154	ASP	CB-CG-OD2	-5.36	113.48	118.30
1	B	226	ASN	N-CA-CB	-5.36	100.95	110.60
1	A	186	ASP	CB-CG-OD1	5.32	123.09	118.30
1	A	271	GLU	OE1-CD-OE2	-5.25	117.00	123.30
1	C	77	VAL	CG1-CB-CG2	-5.24	102.52	110.90
1	C	286	ASP	CB-CG-OD2	-5.22	113.60	118.30
1	A	59	THR	CB-CA-C	-5.20	97.56	111.60
1	A	61	ARG	NH1-CZ-NH2	-5.14	113.74	119.40
1	A	154	ASP	CB-CG-OD1	5.12	122.91	118.30
1	A	130	ARG	NE-CZ-NH1	5.05	122.83	120.30
1	C	15	ARG	CG-CD-NE	5.03	122.35	111.80

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	2272	0	2212	12	1
1	B	2266	0	2206	37	1
1	C	2266	0	2206	32	1
2	A	39	0	0	1	0
2	B	39	0	0	0	0
2	C	39	0	0	0	0
3	A	6	0	8	0	0
3	B	18	0	24	1	0
3	C	6	0	8	0	0
4	A	8	0	12	0	0
5	A	5	0	0	0	0
6	A	214	0	0	4	1
6	B	131	0	0	4	0
6	C	129	0	0	3	0
All	All	7438	0	6676	78	3

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 6.

All (78) 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:C:277:ASN:ND2	6:C:501:HOH:O	1.86	1.08
1:C:209:TYR:HB3	1:C:251:LEU:CD2	2.04	0.88
1:B:211:ALA:HA	1:B:279:LEU:HD11	1.58	0.86
1:B:221:CYS:SG	1:B:268:HIS:CD2	2.80	0.75
1:C:224:ARG:HH21	1:C:224:ARG:HG3	1.51	0.74
1:C:209:TYR:CB	1:C:251:LEU:CD2	2.68	0.71
1:C:209:TYR:HB3	1:C:251:LEU:HD23	1.71	0.71
1:C:261:GLN:NE2	6:C:502:HOH:O	2.22	0.71
1:B:209:TYR:HB3	1:B:251:LEU:HD23	1.73	0.69
1:C:208:LEU:HD13	1:C:262:LEU:HD22	1.74	0.68
1:B:49:VAL:CG2	6:B:615:HOH:O	2.42	0.67
1:B:203:ASN:OD1	6:B:501:HOH:O	2.12	0.67
1:B:218:TRP:CE3	1:B:219:TRP:N	2.65	0.64
1:C:5:LYS:NZ	1:C:292:GLU:OE2	2.30	0.64
1:B:211:ALA:CA	1:B:279:LEU:HD11	2.27	0.63
1:A:69:HIS:HE1	6:A:508:HOH:O	1.82	0.63
1:B:266:ILE:HA	1:B:269:LEU:HD23	1.81	0.62
1:C:208:LEU:HD12	1:C:262:LEU:HD13	1.82	0.60
1:C:209:TYR:CG	1:C:251:LEU:HD23	2.37	0.60
1:C:209:TYR:HB3	1:C:251:LEU:HD21	1.84	0.60

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:209:TYR:CG	1:C:251:LEU:CD2	2.84	0.59
1:B:218:TRP:CE3	1:B:219:TRP:CA	2.87	0.57
1:B:218:TRP:CE3	1:B:219:TRP:HA	2.40	0.56
1:B:269:LEU:HB3	1:B:273:PHE:HE1	1.69	0.56
1:B:220:LEU:HD11	1:B:261:GLN:OE1	2.06	0.56
1:B:131:THR:HG21	1:B:198:LEU:HB2	1.87	0.56
1:A:293:VAL:O	1:A:297:MET:HG2	2.05	0.56
1:C:95:VAL:HG12	6:C:551:HOH:O	2.07	0.55
1:B:21:CYS:HB3	1:B:66:SER:HB3	1.89	0.54
1:A:93:SER:HB3	1:C:256:GLY:O	2.07	0.54
1:B:269:LEU:H	1:B:269:LEU:HD22	1.71	0.52
1:A:22:TYR:CE1	1:A:64:ASN:HB2	2.45	0.52
1:B:282:SER:OG	1:C:283:SER:N	2.44	0.51
1:C:209:TYR:CB	1:C:251:LEU:HD23	2.36	0.50
1:C:224:ARG:HG3	1:C:224:ARG:NH2	2.24	0.50
1:B:95:VAL:HG12	6:B:565:HOH:O	2.11	0.49
1:A:148:GLY:HA3	1:A:161:TYR:HB3	1.95	0.49
1:B:292:GLU:OE2	3:B:402:GOL:H11	2.13	0.49
1:B:31:TRP:CE2	1:B:94:ASN:HB2	2.48	0.49
1:B:278:ILE:HG22	1:B:279:LEU:HD12	1.95	0.49
1:B:209:TYR:CE1	1:B:262:LEU:HD11	2.49	0.47
1:A:95:VAL:O	1:C:217:ARG:NH1	2.48	0.47
1:B:40:ARG:HA	1:B:86:LEU:HG	1.96	0.46
1:C:66:SER:HB2	1:C:73:PHE:HE1	1.80	0.46
1:C:107:PRO:HB3	1:C:131:THR:HA	1.98	0.46
1:B:266:ILE:CA	1:B:269:LEU:HD23	2.45	0.46
1:C:224:ARG:HH21	1:C:224:ARG:CG	2.23	0.45
1:C:31:TRP:CE2	1:C:94:ASN:HB2	2.52	0.45
1:A:254:LYS:HE3	6:A:687:HOH:O	2.17	0.44
1:B:278:ILE:CG2	1:B:279:LEU:HD12	2.48	0.44
1:C:265:SER:O	1:C:269:LEU:HG	2.17	0.44
1:B:265:SER:CB	1:B:269:LEU:HD21	2.47	0.44
1:C:224:ARG:NH2	1:C:224:ARG:CG	2.80	0.44
1:C:208:LEU:HD11	1:C:262:LEU:HB3	1.99	0.44
1:B:218:TRP:CZ3	1:B:219:TRP:HB3	2.53	0.44
1:A:244:SER:HB2	1:A:246:GLU:OE2	2.18	0.44
1:C:209:TYR:CD2	1:C:251:LEU:HD22	2.52	0.44
1:A:41:HIS:NE2	2:A:401:E8E:O58	2.51	0.44
1:B:112:ASN:O	1:B:148:GLY:HA2	2.17	0.44
1:C:31:TRP:CD2	1:C:94:ASN:HB2	2.53	0.44
1:C:40:ARG:HA	1:C:86:LEU:HG	1.99	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:246:GLU:HG2	6:A:609:HOH:O	2.18	0.43
1:B:49:VAL:HG22	6:B:615:HOH:O	2.09	0.43
1:B:293:VAL:O	1:B:297:MET:HG2	2.18	0.43
1:B:31:TRP:CD2	1:B:94:ASN:HB2	2.54	0.43
1:C:218:TRP:CE3	1:C:219:TRP:HA	2.54	0.43
1:B:221:CYS:SG	1:B:268:HIS:HD2	2.39	0.42
1:C:112:ASN:O	1:C:148:GLY:HA2	2.20	0.42
1:A:31:TRP:CD2	1:A:94:ASN:HB2	2.55	0.42
1:B:218:TRP:CD2	1:B:219:TRP:N	2.88	0.41
1:C:209:TYR:CG	1:C:251:LEU:HD22	2.53	0.41
1:B:113:ILE:O	1:B:124:VAL:HA	2.20	0.41
1:B:131:THR:HG21	1:B:198:LEU:H	1.85	0.41
1:B:266:ILE:CA	1:B:269:LEU:CD2	2.99	0.41
1:A:69:HIS:CE1	6:A:508:HOH:O	2.64	0.41
1:B:204:VAL:CG1	1:B:284:LEU:HD22	2.51	0.40
1:B:273:PHE:CD2	1:B:282:SER:O	2.75	0.40
1:C:245:VAL:CG2	1:C:259:VAL:HG21	2.51	0.40

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:A:629:HOH:O	6:A:706:HOH:O[3_655]	1.85	0.35
1:C:268:HIS:ND1	1:C:268:HIS:ND1[3_655]	1.91	0.29
1:A:224:ARG:NH2	1:B:54:ASP:OD1[4_555]	1.92	0.28

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	298/306 (97%)	291 (98%)	6 (2%)	1 (0%)	41 50

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	297/306 (97%)	291 (98%)	5 (2%)	1 (0%)	41	50
1	C	297/306 (97%)	286 (96%)	11 (4%)	0	100	100
All	All	892/918 (97%)	868 (97%)	22 (2%)	2 (0%)	47	58

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	279	LEU
1	A	299	GLY

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	250/257 (97%)	247 (99%)	3 (1%)	71	84
1	B	250/257 (97%)	237 (95%)	13 (5%)	23	32
1	C	250/257 (97%)	240 (96%)	10 (4%)	31	44
All	All	750/771 (97%)	724 (96%)	26 (4%)	36	50

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	59	THR
1	A	93	SER
1	A	164	GLN
1	B	27	LEU
1	B	95	VAL
1	B	199	MET
1	B	216	CYS
1	B	217	ARG
1	B	220	LEU
1	B	221	CYS
1	B	226	ASN
1	B	240	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	259	VAL
1	B	269	LEU
1	B	277	ASN
1	B	282	SER
1	C	46	SER
1	C	93	SER
1	C	199	MET
1	C	208	LEU
1	C	222	SER
1	C	223	THR
1	C	226	ASN
1	C	251	LEU
1	C	283	SER
1	C	284	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	164	GLN
1	A	268	HIS
1	B	268	HIS
1	B	270	HIS

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](#)

11 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	GOL	B	403	-	5,5,5	0.73	0	5,5,5	0.60	0
3	GOL	B	402	-	5,5,5	0.53	0	5,5,5	0.89	0
2	E8E	C	401	1	41,41,41	2.27	6 (14%)	49,54,54	1.72	12 (24%)
5	SO4	A	405	-	4,4,4	0.62	0	6,6,6	0.51	0
3	GOL	A	402	-	5,5,5	0.47	0	5,5,5	0.73	0
3	GOL	B	404	-	5,5,5	0.66	0	5,5,5	0.55	0
3	GOL	C	402	-	5,5,5	0.90	0	5,5,5	0.66	0
2	E8E	A	401	1	41,41,41	1.93	10 (24%)	49,54,54	1.46	6 (12%)
4	DMS	A	403	-	3,3,3	0.46	0	3,3,3	1.06	0
4	DMS	A	404	-	3,3,3	0.32	0	3,3,3	1.19	0
2	E8E	B	401	1	41,41,41	2.32	11 (26%)	49,54,54	1.33	6 (12%)

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
3	GOL	B	403	-	-	2/4/4/4	-
3	GOL	B	402	-	-	2/4/4/4	-
2	E8E	C	401	1	-	0/38/48/48	0/3/3/3
3	GOL	A	402	-	-	1/4/4/4	-
3	GOL	B	404	-	-	0/4/4/4	-
3	GOL	C	402	-	-	2/4/4/4	-
2	E8E	A	401	1	-	0/38/48/48	0/3/3/3
2	E8E	B	401	1	-	3/38/48/48	0/3/3/3

All (27) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	401	E8E	C45-C47	-7.87	1.42	1.52
2	C	401	E8E	C03-C05	7.60	1.52	1.33
2	B	401	E8E	C03-C05	6.83	1.50	1.33

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	401	E8E	C45-C47	-6.47	1.44	1.52
2	A	401	E8E	C03-C05	6.14	1.49	1.33
2	C	401	E8E	C42-C40	5.09	1.60	1.52
2	C	401	E8E	C70-C73	-4.97	1.40	1.51
2	A	401	E8E	C70-C73	-4.36	1.42	1.51
2	B	401	E8E	C51-N49	4.13	1.54	1.46
2	B	401	E8E	C57-C40	3.90	1.58	1.54
2	A	401	E8E	C45-C47	-3.68	1.47	1.52
2	A	401	E8E	O58-C57	3.62	1.49	1.42
2	B	401	E8E	C03-C02	3.26	1.55	1.48
2	B	401	E8E	C70-C73	-3.23	1.44	1.51
2	B	401	E8E	C47-N49	3.20	1.37	1.33
2	B	401	E8E	C70-N68	3.18	1.52	1.46
2	A	401	E8E	C03-C02	2.85	1.54	1.48
2	A	401	E8E	C07-C05	-2.80	1.39	1.47
2	C	401	E8E	C07-C05	-2.77	1.39	1.47
2	B	401	E8E	C42-C40	2.75	1.57	1.52
2	A	401	E8E	C20-C36	-2.74	1.45	1.52
2	A	401	E8E	C51-N49	2.60	1.51	1.46
2	A	401	E8E	C42-C40	2.49	1.56	1.52
2	A	401	E8E	O01-C02	2.39	1.29	1.24
2	B	401	E8E	C07-C05	-2.29	1.40	1.47
2	B	401	E8E	C57-C66	2.26	1.57	1.52
2	C	401	E8E	C02-N18	2.11	1.39	1.34

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	401	E8E	O01-C02-N18	6.19	130.54	122.35
2	B	401	E8E	C70-N68-C66	4.27	128.48	122.34
2	A	401	E8E	C03-C02-N18	-3.48	107.87	114.56
2	C	401	E8E	C22-C20-C36	-3.39	102.50	110.57
2	A	401	E8E	C57-C40-N38	-3.25	103.91	110.03
2	C	401	E8E	C07-C05-C03	-3.21	119.57	126.91
2	A	401	E8E	C10-C08-C07	-2.81	117.24	120.65
2	A	401	E8E	C70-N68-C66	2.80	126.37	122.34
2	A	401	E8E	C40-N38-C36	-2.63	118.42	123.07
2	B	401	E8E	C26-C25-C22	2.61	120.72	111.11
2	B	401	E8E	C51-C54-C45	-2.56	101.61	105.75
2	C	401	E8E	C73-C70-N68	-2.52	107.65	113.05
2	C	401	E8E	O58-C57-C66	-2.50	105.24	110.63
2	B	401	E8E	C54-C45-C47	2.40	106.01	102.88

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	401	E8E	C57-C40-N38	-2.32	105.66	110.03
2	C	401	E8E	C03-C02-N18	-2.32	110.10	114.56
2	C	401	E8E	C70-N68-C66	2.30	125.64	122.34
2	C	401	E8E	C22-C20-N18	2.27	115.80	110.58
2	C	401	E8E	C42-C45-C54	-2.19	109.27	117.31
2	B	401	E8E	C20-C36-N38	-2.16	111.96	116.70
2	B	401	E8E	C07-C05-C03	-2.16	121.97	126.91
2	C	401	E8E	C57-C66-N68	-2.13	112.22	116.66
2	A	401	E8E	C57-C66-N68	-2.09	112.31	116.66
2	C	401	E8E	C70-C73-C76	-2.04	116.68	120.91

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	402	GOL	O1-C1-C2-C3
3	B	403	GOL	C1-C2-C3-O3
3	C	402	GOL	O1-C1-C2-C3
2	B	401	E8E	C20-C22-C25-C26
3	B	402	GOL	O1-C1-C2-O2
3	B	403	GOL	O2-C2-C3-O3
3	C	402	GOL	O1-C1-C2-O2
2	B	401	E8E	C20-C22-C25-C34
3	A	402	GOL	C1-C2-C3-O3
2	B	401	E8E	N18-C20-C22-C25

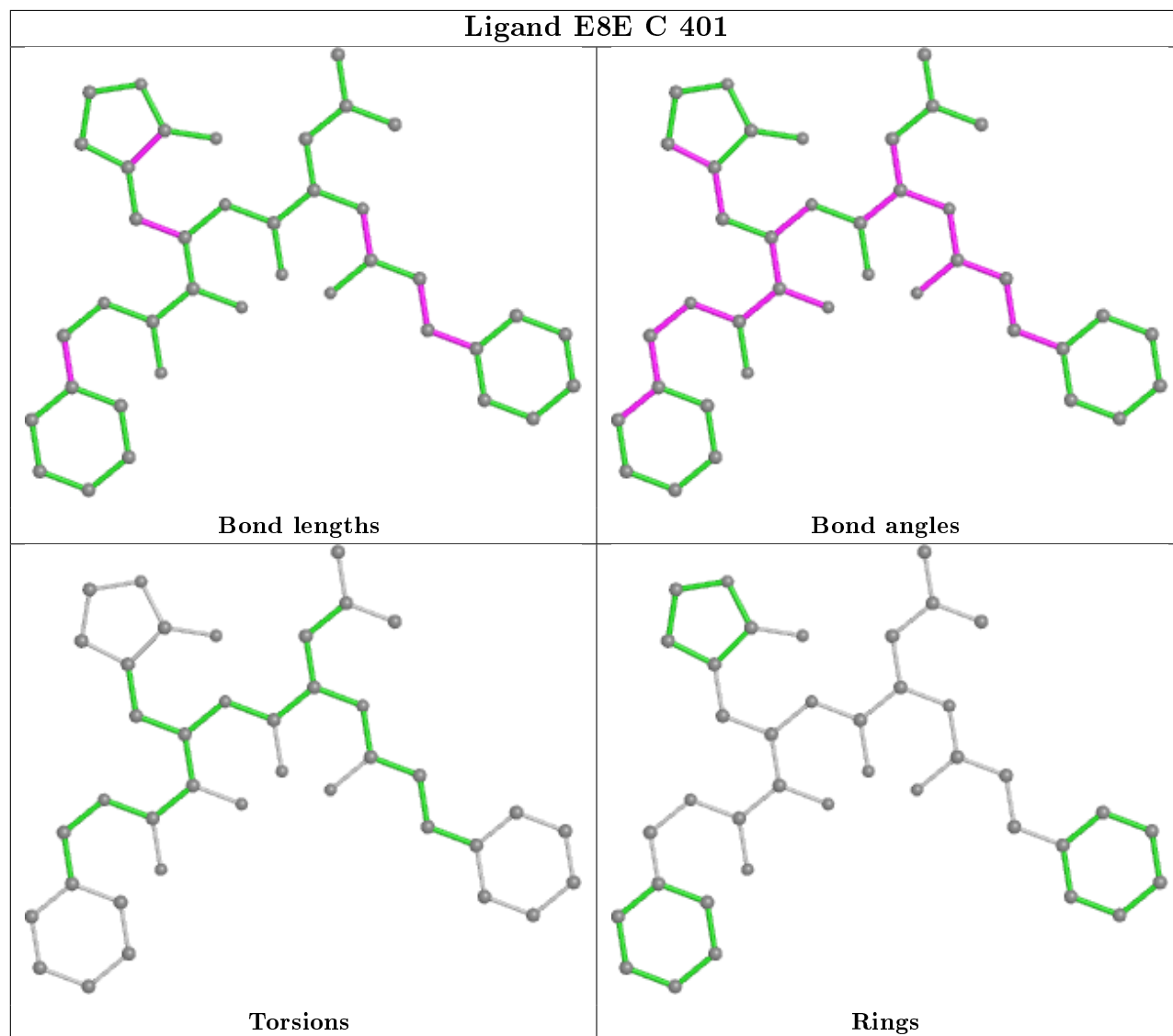
There are no ring outliers.

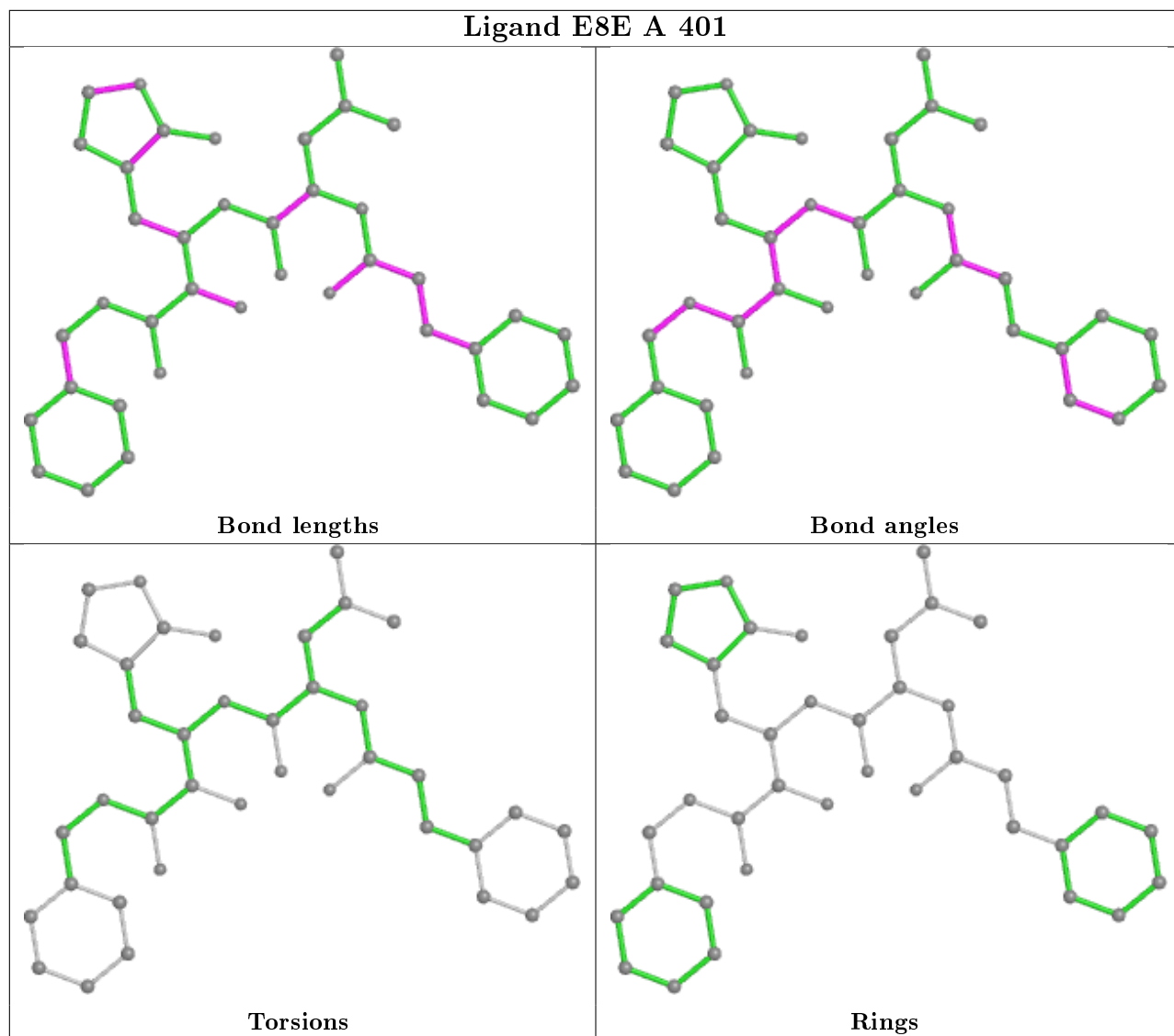
2 monomers are involved in 2 short contacts:

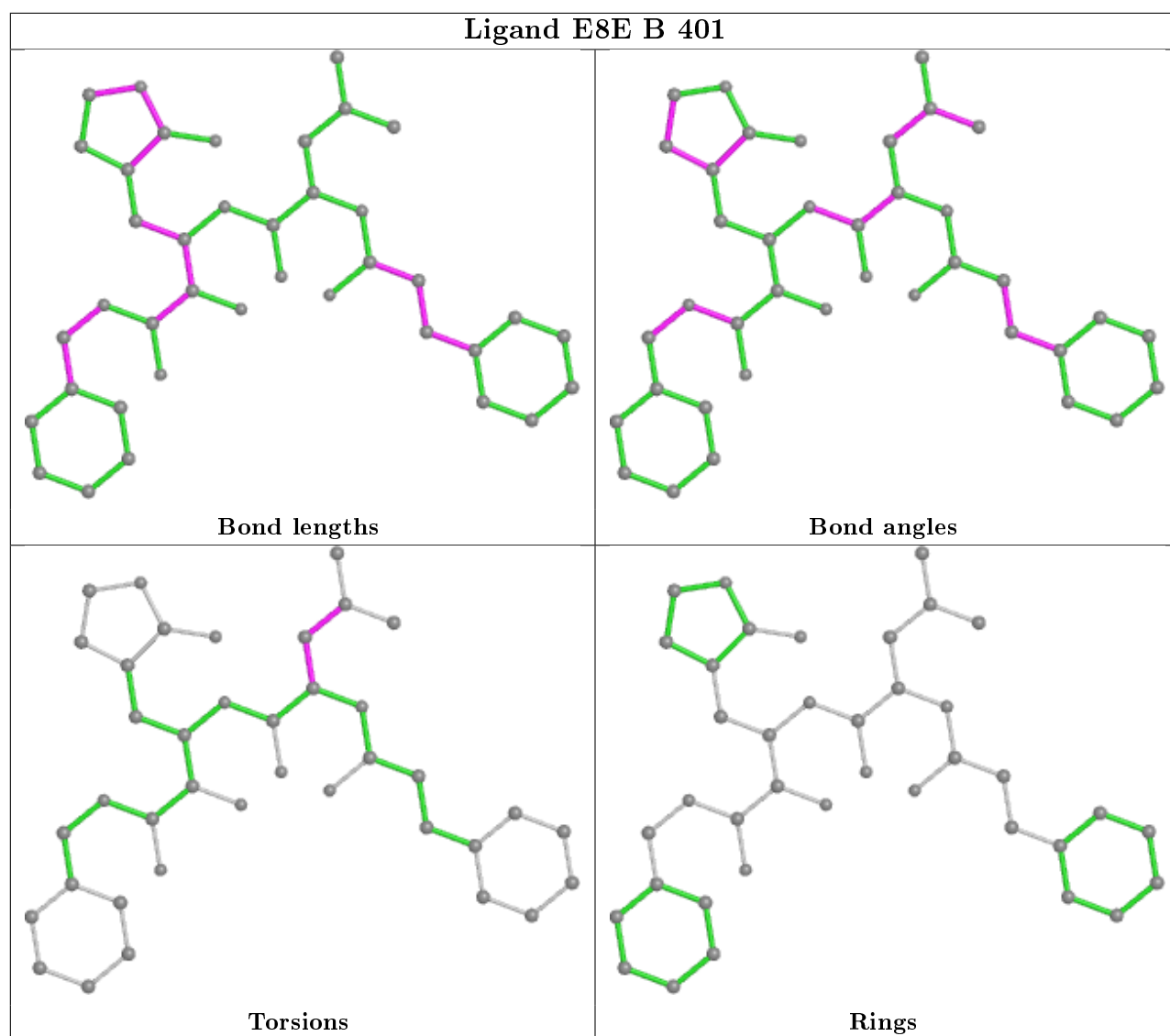
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	402	GOL	1	0
2	A	401	E8E	1	0

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 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	300/306 (98%)	-0.30	3 (1%) 82 86	20, 34, 64, 98	0
1	B	299/306 (97%)	0.56	47 (15%) 2 2	26, 48, 127, 160	0
1	C	299/306 (97%)	0.54	36 (12%) 4 6	27, 52, 107, 167	0
All	All	898/918 (97%)	0.26	86 (9%) 8 10	20, 44, 112, 167	0

All (86) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	274	GLY	10.3
1	C	272	GLY	8.9
1	C	273	PHE	7.8
1	B	273	PHE	7.0
1	B	245	VAL	6.6
1	B	221	CYS	6.4
1	C	271	GLU	5.7
1	B	257	VAL	5.6
1	B	258	SER	5.1
1	B	246	GLU	4.8
1	B	222	SER	4.7
1	C	244	SER	4.7
1	C	268	HIS	4.6
1	B	217	ARG	4.5
1	C	230	PHE	4.4
1	B	272	GLY	4.4
1	B	218	TRP	4.3
1	B	227	VAL	4.3
1	B	256	GLY	4.3
1	B	274	GLY	4.3
1	C	245	VAL	4.2
1	C	282	SER	4.1
1	B	276	LYS	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	219	TRP	4.0
1	B	262	LEU	3.8
1	B	244	SER	3.8
1	C	227	VAL	3.8
1	B	281	TYR	3.8
1	C	235	MET	3.8
1	B	259	VAL	3.7
1	B	224	ARG	3.6
1	B	275	GLY	3.6
1	B	248	TYR	3.5
1	B	252	ALA	3.5
1	C	280	GLY	3.5
1	B	249	SER	3.4
1	C	238	GLY	3.4
1	B	253	ALA	3.4
1	B	220	LEU	3.4
1	B	211	ALA	3.3
1	A	47	THR	3.3
1	B	250	ILE	3.3
1	B	255	THR	3.3
1	A	300	VAL	3.3
1	C	247	CYS	3.2
1	C	222	SER	3.2
1	C	225	VAL	3.2
1	C	237	ASN	3.2
1	C	243	SER	3.2
1	B	242	VAL	3.1
1	B	266	ILE	3.0
1	C	234	ALA	2.9
1	B	223	THR	2.9
1	C	242	VAL	2.8
1	C	267	GLN	2.8
1	C	232	GLU	2.8
1	B	73	PHE	2.8
1	B	235	MET	2.8
1	B	209	TYR	2.8
1	B	263	LEU	2.7
1	C	270	HIS	2.7
1	B	213	LEU	2.7
1	B	269	LEU	2.7
1	C	277	ASN	2.6
1	C	218	TRP	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	232	GLU	2.6
1	C	233	TRP	2.5
1	B	261	GLN	2.5
1	C	231	ASN	2.5
1	C	246	GLU	2.4
1	B	277	ASN	2.4
1	C	191	LEU	2.3
1	C	229	GLY	2.3
1	B	228	ASP	2.3
1	C	276	LYS	2.3
1	C	263	LEU	2.2
1	B	225	VAL	2.2
1	C	236	ALA	2.2
1	B	229	GLY	2.1
1	B	226	ASN	2.1
1	C	266	ILE	2.1
1	A	48	THR	2.1
1	B	210	ALA	2.1
1	B	215	GLY	2.1
1	C	73	PHE	2.0
1	C	269	LEU	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, 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.

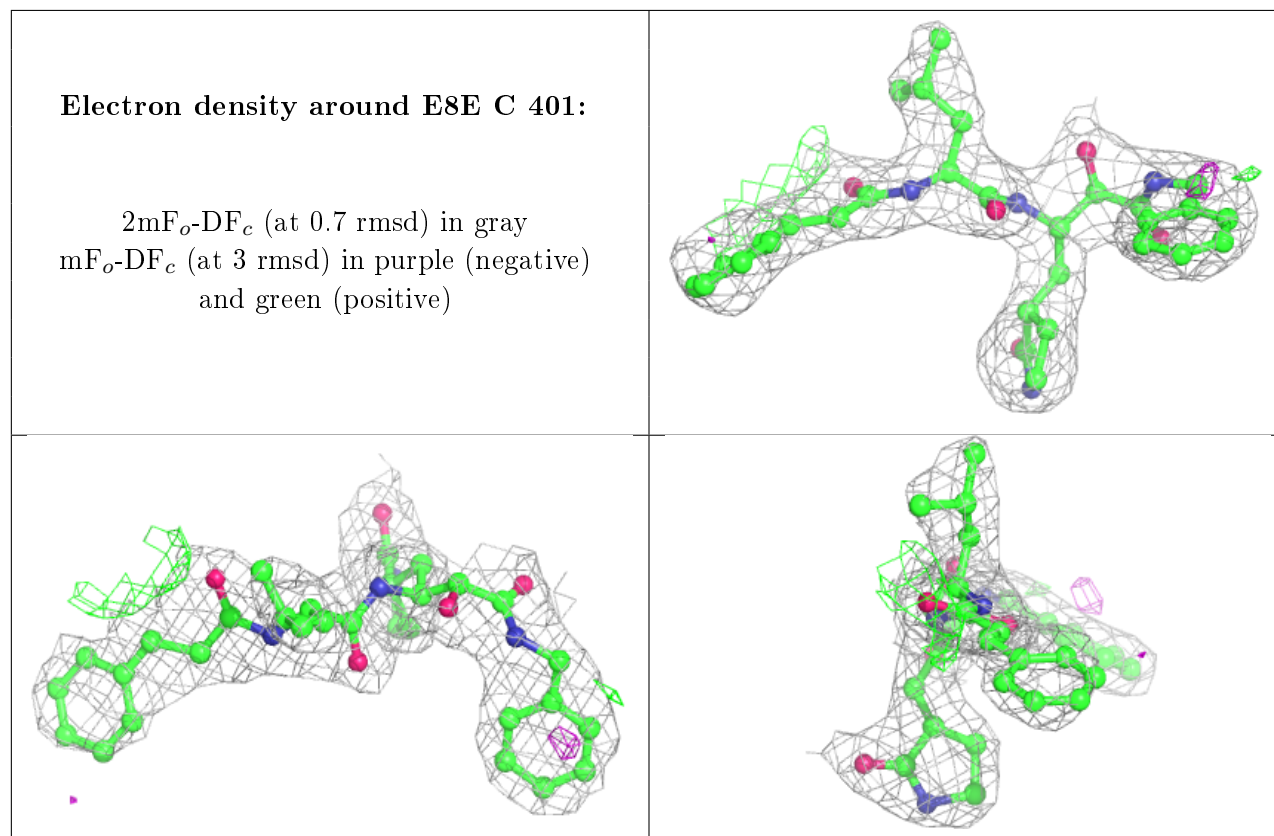
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	GOL	C	402	6/6	0.75	0.22	62,77,81,93	0
4	DMS	A	404	4/4	0.83	0.22	90,94,96,101	0

Continued on next page...

Continued from previous page...

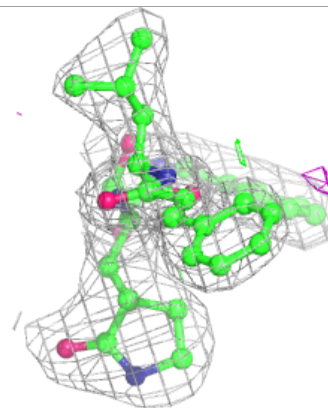
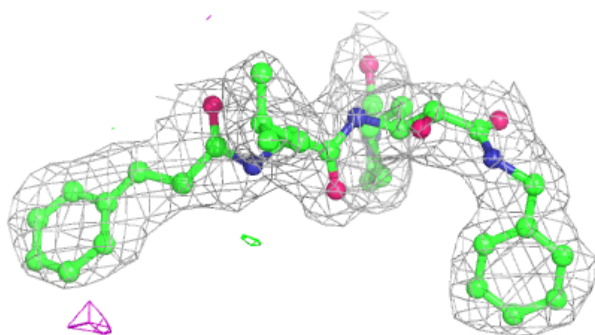
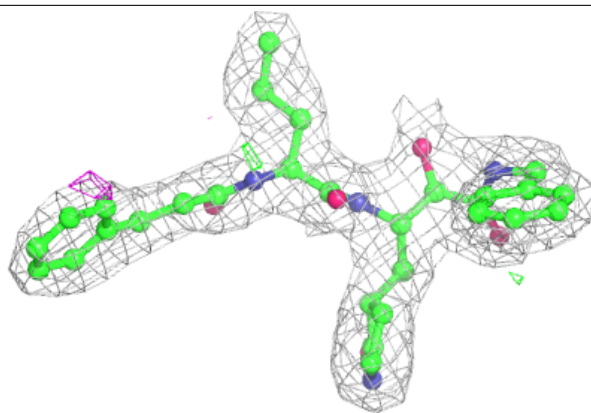
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
5	SO4	A	405	5/5	0.85	0.17	76,84,99,101	0
3	GOL	B	404	6/6	0.90	0.16	59,67,71,73	0
3	GOL	B	402	6/6	0.92	0.23	42,57,59,82	0
3	GOL	B	403	6/6	0.92	0.15	64,69,75,78	0
4	DMS	A	403	4/4	0.93	0.18	72,81,82,83	0
2	E8E	C	401	39/39	0.94	0.14	40,53,78,85	0
3	GOL	A	402	6/6	0.95	0.14	34,37,39,40	0
2	E8E	A	401	39/39	0.95	0.12	28,34,55,57	0
2	E8E	B	401	39/39	0.96	0.11	28,35,47,50	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.

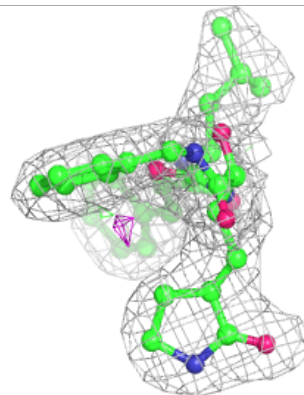
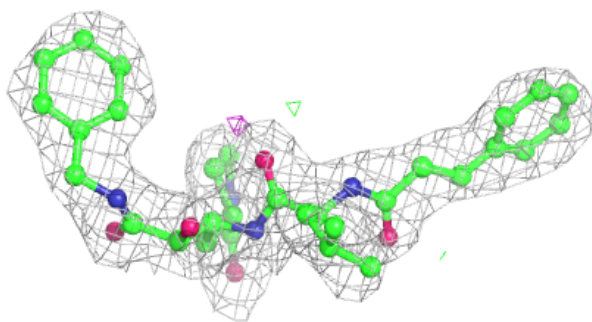
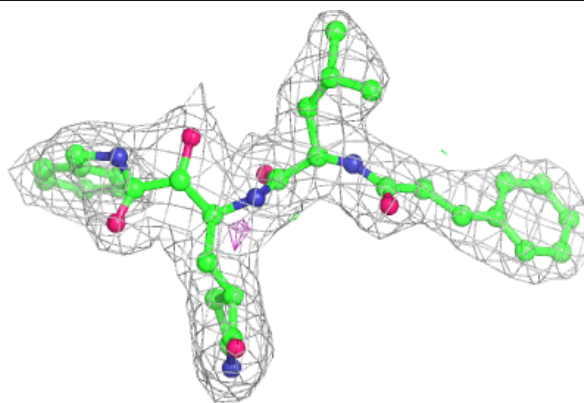


Electron density around E8E A 401:

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

**Electron density around E8E B 401:**

$2mF_o-DF_c$ (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.