



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 6, 2026 – 12:18 pm GMT

PDB ID : 9SDM / pdb_00009sdm
Title : Crystal structure of SARS-CoV-2 main protease (MPro) in complex with the covalently bound inhibitor GUE-4303 (compound 12 in publication)
Authors : Yang, C.-C.; Strater, N.; Sylvester, K.; Muller, C.E.; Voget, R.; Guetschow, M.
Deposited on : 2025-08-14
Resolution : 1.55 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	:	?.? (???), CSD ??CSD?? (????)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
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)	:	2.47

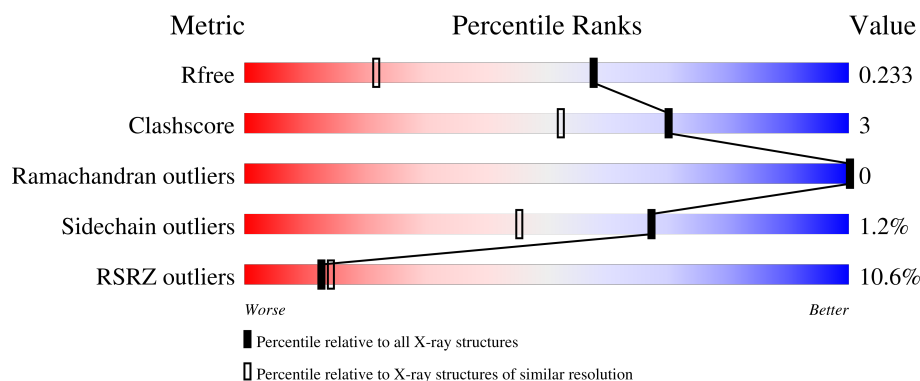
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.55 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1935 (1.56-1.56)
Clashscore	180529	2073 (1.56-1.56)
Ramachandran outliers	177936	2037 (1.56-1.56)
Sidechain outliers	177891	2034 (1.56-1.56)
RSRZ outliers	164620	1935 (1.56-1.56)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments 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	2% 96% .
1	B	306	19% 87% 11% .

2 Entry composition [i](#)

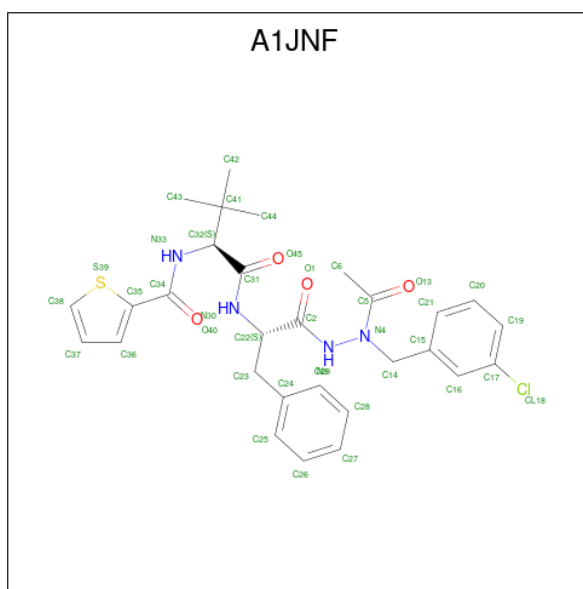
There are 5 unique types of molecules in this entry. The entry contains 9958 atoms, of which 4671 are hydrogens and 0 are deuteriums.

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 nsp5.

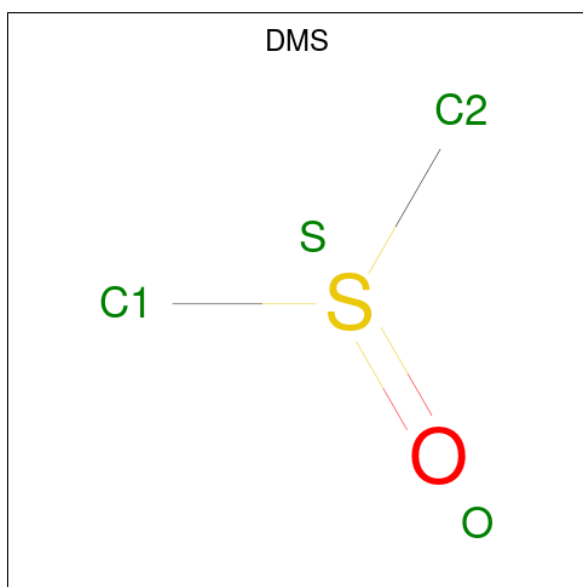
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	305	Total	C	H	N	O	S	0	5	0
			4681	1509	2305	400	444	23			
1	B	301	Total	C	H	N	O	S	0	2	0
			4621	1479	2284	397	438	23			

- Molecule 2 is N-[(2S)-1-[[[(2S)-1-[2-[(3-chlorophenyl)methyl]-2-ethanoyl-hydrazinyl]-1-oxidan ylidene-3-phenyl-propan-2-yl]amino]-3,3-dimethyl-1-oxidanylidene-butan-2-yl]thiophene-2-carboxamide (CCD ID: A1JNF) (formula: C₂₉H₃₃ClN₄O₄S) (labeled as "Ligand of Interest" by depositor).



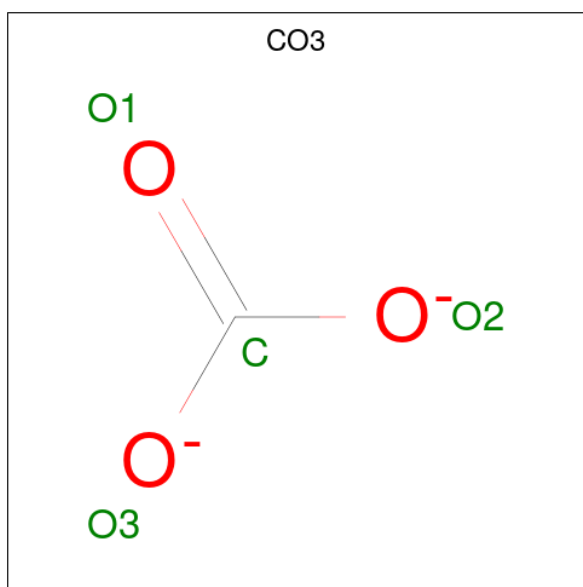
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	Cl	H	N	O	S	0
			71	29	1	32	4	4	1	0
2	B	1	Total	C	Cl	H	N	O	S	0
			71	29	1	32	4	4	1	0

- Molecule 3 is DIMETHYL SULFOXIDE (CCD ID: DMS) (formula: C₂H₆OS).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	H	O	S	0	0
			10	2	6	1	1		
3	A	1	Total	C	H	O	S	0	0
			10	2	6	1	1		
3	B	1	Total	C	H	O	S	0	0
			10	2	6	1	1		

- Molecule 4 is CARBONATE ION (CCD ID: CO3) (formula: CO₃).



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

Continued on next page...

Continued from previous page...

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

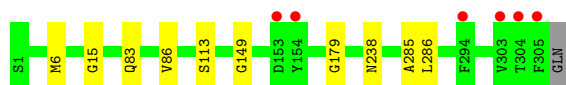
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	290	Total	O	0	0
			290	290		
5	B	182	Total	O	0	0
			182	182		

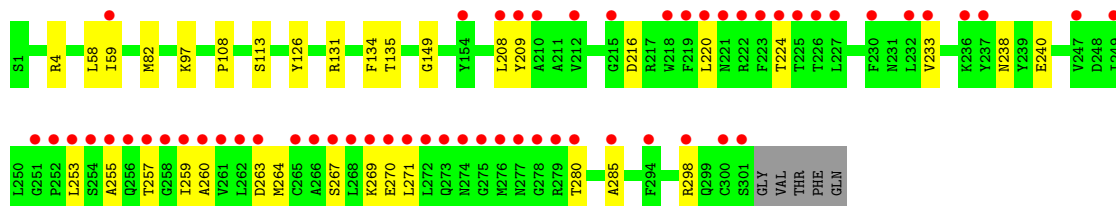
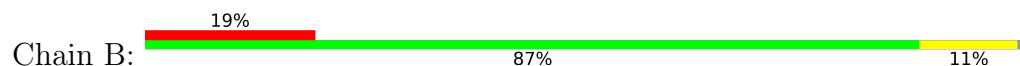
3 Residue-property plots [i](#)

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: 3C-like proteinase nsp5



- Molecule 1: 3C-like proteinase nsp5



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	67.84Å 101.73Å 104.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.37 – 1.55 36.37 – 1.55	Depositor EDS
% Data completeness (in resolution range)	64.1 (36.37-1.55) 64.0 (36.37-1.55)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.15 (at 1.55Å)	Xtriage
Refinement program	PHENIX (1.21.2_5419: ???)	Depositor
R, R_{free}	0.187 , 0.231 0.189 , 0.233	Depositor DCC
R_{free} test set	3221 reflections (3.10%)	wwPDB-VP
Wilson B-factor (Å ²)	21.8	Xtriage
Anisotropy	0.023	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 33.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.017 for -h,l,k	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	9958	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.41% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: DMS, A1JNF, CO3

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	0.58	0/2444	0.65	0/3322
1	B	0.53	0/2395	0.64	0/3255
All	All	0.55	0/4839	0.65	0/6577

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	2376	2305	2340	13	0
1	B	2337	2284	2288	23	0
2	A	39	32	0	0	0
2	B	39	32	0	0	0
3	A	8	12	12	1	0
3	B	4	6	6	1	0
4	A	4	0	0	0	0
4	B	8	0	0	0	0
5	A	290	0	0	1	0
5	B	182	0	0	5	0
All	All	5287	4671	4646	31	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 (31) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:286[A]:LEU:HD22	5:B:675:HOH:O	1.86	0.73
1:B:58:LEU:HD22	1:B:82:MET:HE3	1.70	0.72
1:A:86[A]:VAL:HG13	1:A:179:GLY:HA2	1.73	0.69
1:B:298:ARG:NH1	3:B:403:DMS:H21	2.13	0.63
1:B:240:GLU:HG3	5:B:581:HOH:O	2.00	0.61
1:A:15:GLY:O	3:A:402:DMS:H21	2.01	0.60
1:A:286[A]:LEU:HD23	5:A:533:HOH:O	2.02	0.58
1:B:257:THR:OG1	1:B:259:ILE:HD12	2.04	0.56
1:B:220:LEU:HD22	1:B:259:ILE:HG12	1.86	0.56
1:B:108:PRO:HG3	1:B:134:PHE:CE1	2.43	0.53
1:A:6:MET:HE2	1:B:126:TYR:CE2	2.44	0.52
1:B:267:SER:O	1:B:271:LEU:HD12	2.10	0.52
1:A:285:ALA:HB3	1:B:285:ALA:HB3	1.92	0.52
1:B:260:ALA:HB3	1:B:263:ASP:HB2	1.92	0.51
1:A:113:SER:O	1:A:149:GLY:HA2	2.11	0.50
1:A:86[A]:VAL:HG13	1:A:179:GLY:CA	2.40	0.50
1:B:113:SER:O	1:B:149:GLY:HA2	2.14	0.48
1:B:4:ARG:NH1	5:B:512:HOH:O	2.50	0.44
1:B:255:ALA:O	1:B:257:THR:O	2.35	0.44
1:B:240:GLU:CG	5:B:581:HOH:O	2.64	0.44
1:A:286[A]:LEU:HD11	1:B:280:THR:HG22	1.98	0.44
1:B:208:LEU:HB3	1:B:264:MET:HE2	2.00	0.44
1:A:83:GLN:O	1:A:86[A]:VAL:HG22	2.17	0.44
1:B:97:LYS:HD2	5:B:571:HOH:O	2.19	0.42
1:B:233:VAL:HG21	1:B:269:LYS:HD3	2.01	0.42
1:A:286[A]:LEU:HD11	1:B:280:THR:CG2	2.49	0.42
1:B:209:TYR:HB3	1:B:253:LEU:HB3	2.02	0.42
1:B:257:THR:CB	1:B:259:ILE:HD12	2.51	0.41
1:A:6:MET:HE2	1:B:126:TYR:CD2	2.57	0.40
1:B:131:ARG:HG2	1:B:135:THR:O	2.22	0.40

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	308/306 (101%)	300 (97%)	8 (3%)	0	100	100
1	B	301/306 (98%)	287 (95%)	14 (5%)	0	100	100
All	All	609/612 (100%)	587 (96%)	22 (4%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	267/263 (102%)	266 (100%)	1 (0%)	89	82
1	B	261/263 (99%)	256 (98%)	5 (2%)	52	24
All	All	528/526 (100%)	522 (99%)	6 (1%)	67	49

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	238	ASN
1	B	59	ILE
1	B	216	ASP
1	B	224	THR
1	B	238	ASN
1	B	270	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	72	ASN
1	A	107	GLN
1	A	180	ASN
1	B	19	GLN
1	B	84	ASN
1	B	119	ASN
1	B	238	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

8 ligands are modelled in this entry.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

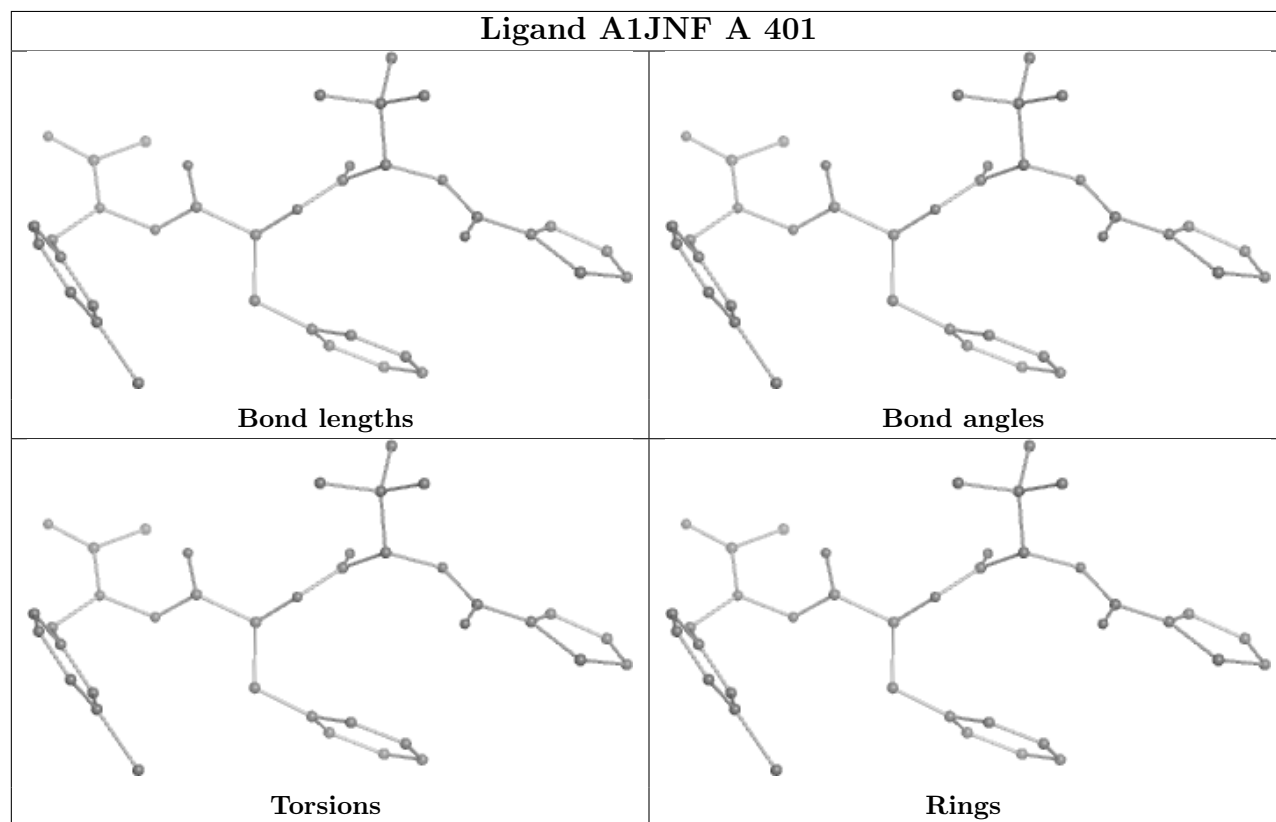
There are no torsion outliers.

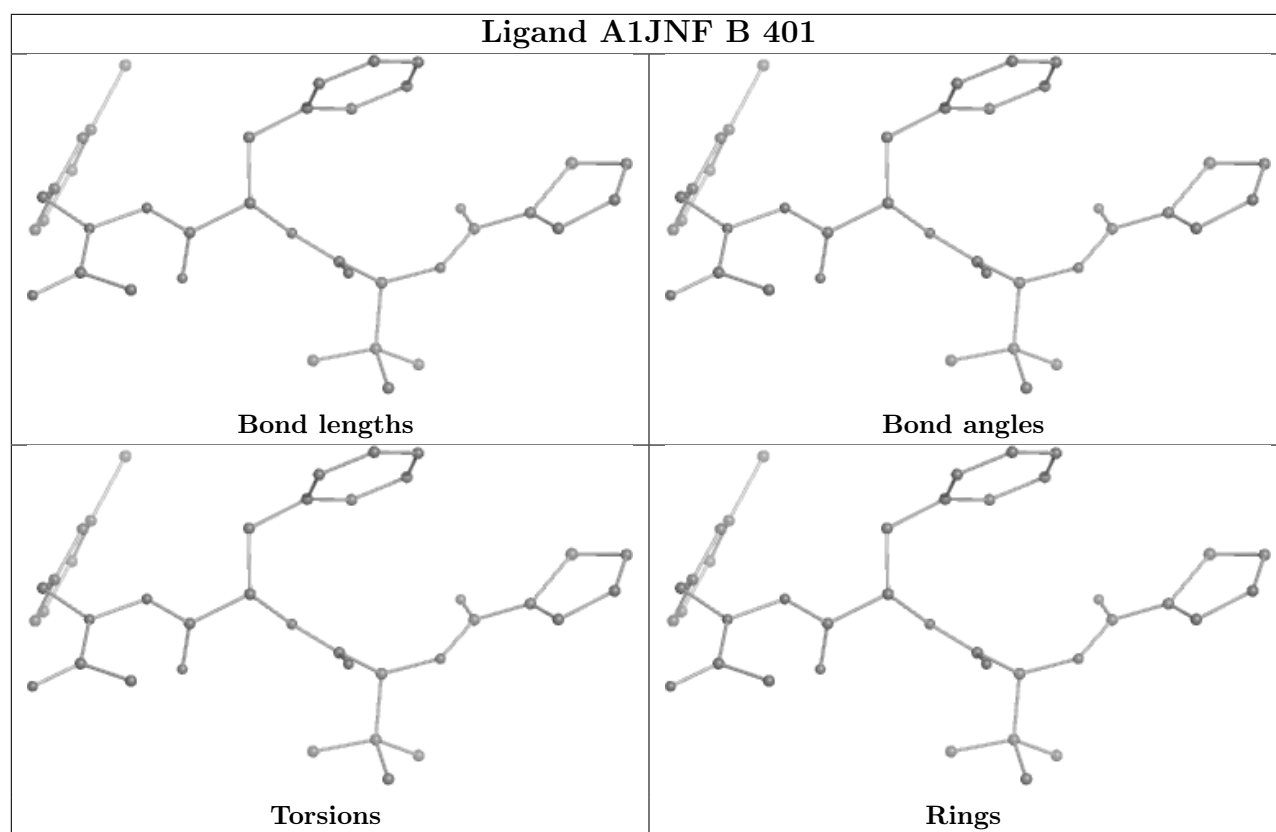
There are no ring outliers.

No monomer is involved in short contacts.

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	305/306 (99%)	-0.03	6 (1%) 64 73	12, 27, 46, 76	5 (1%)
1	B	301/306 (98%)	0.79	58 (19%) 4 4	14, 34, 67, 78	2 (0%)
All	All	606/612 (99%)	0.38	64 (10%) 13 14	12, 29, 62, 78	7 (1%)

All (64) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	259	ILE	7.0
1	B	220	LEU	6.1
1	B	218	TRP	5.5
1	B	257	THR	4.9
1	A	154	TYR	4.8
1	B	223	PHE	4.8
1	B	266	ALA	4.4
1	B	260	ALA	4.4
1	B	278	GLY	4.3
1	B	154	TYR	4.1
1	B	280	THR	4.1
1	B	265	CYS	4.1
1	A	305	PHE	4.0
1	B	219	PHE	4.0
1	B	276	MET	4.0
1	B	275	GLY	3.9
1	B	226	THR	3.9
1	B	258	GLY	3.9
1	B	272	LEU	3.9
1	A	304	THR	3.9
1	B	230	PHE	3.8
1	B	256	GLN	3.8
1	A	303	VAL	3.7
1	B	209	TYR	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	269	LYS	3.7
1	B	263	ASP	3.6
1	B	254	SER	3.5
1	B	262	LEU	3.4
1	B	277	ASN	3.4
1	B	285	ALA	3.4
1	B	255	ALA	3.3
1	B	267	SER	3.3
1	B	268	LEU	3.2
1	B	301	SER	3.2
1	B	233	VAL	3.2
1	B	225	THR	3.2
1	B	261	VAL	3.1
1	B	251	GLY	3.0
1	B	294	PHE	3.0
1	B	249	ILE	2.9
1	B	237	TYR	2.9
1	B	253	LEU	2.9
1	B	271	LEU	2.9
1	B	252	PRO	2.9
1	B	227	LEU	2.8
1	B	215	GLY	2.7
1	B	232	LEU	2.7
1	B	208	LEU	2.6
1	B	59	ILE	2.6
1	B	212	VAL	2.6
1	B	279	ARG	2.6
1	B	300	CYS	2.5
1	A	294	PHE	2.5
1	B	247	VAL	2.5
1	B	222	ARG	2.5
1	B	273	GLN	2.4
1	B	274	ASN	2.4
1	B	210	ALA	2.3
1	B	221	ASN	2.2
1	B	224	THR	2.2
1	B	298	ARG	2.2
1	A	153	ASP	2.1
1	B	236	LYS	2.1
1	B	270	GLU	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](#)

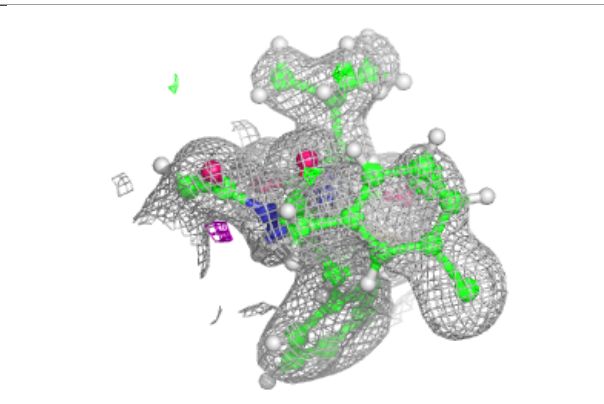
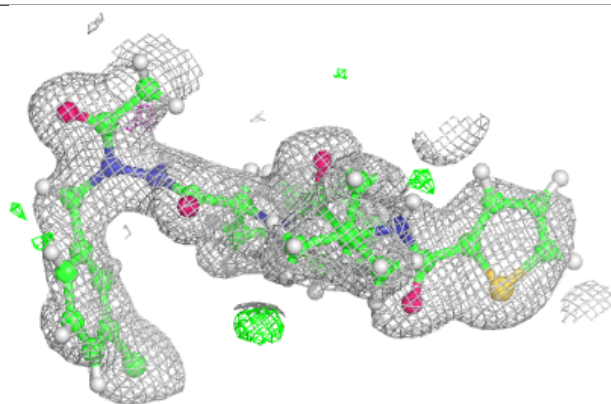
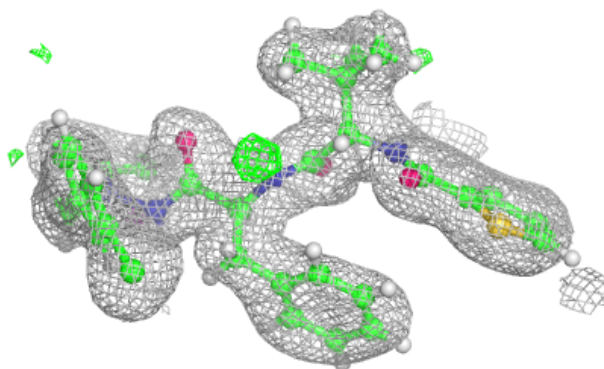
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
3	DMS	A	402	4/4	0.86	0.15	41,53,64,81	0
3	DMS	B	403	4/4	0.91	0.11	28,38,46,46	0
4	CO3	A	404	4/4	0.92	0.09	25,34,39,46	0
3	DMS	A	403	4/4	0.93	0.09	40,48,57,60	0
4	CO3	B	404	4/4	0.93	0.15	28,31,45,56	0
4	CO3	B	402	4/4	0.94	0.10	25,45,46,48	0
2	A1JNF	B	401	39/39	0.95	0.08	21,33,55,68	0
2	A1JNF	A	401	39/39	0.96	0.07	18,27,43,62	0

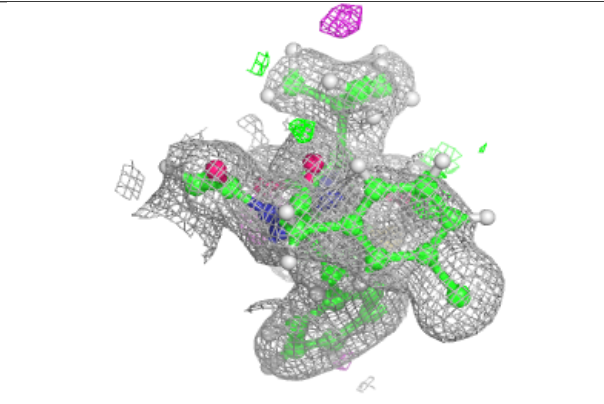
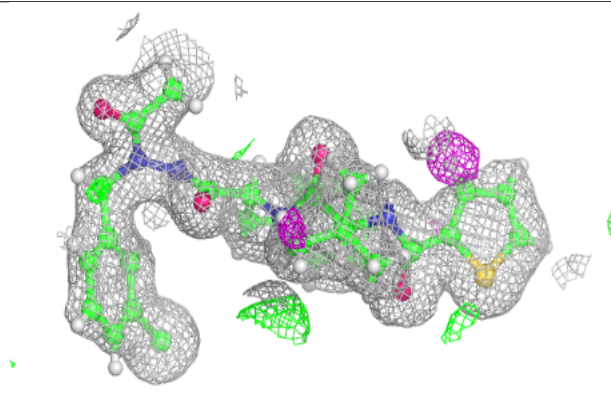
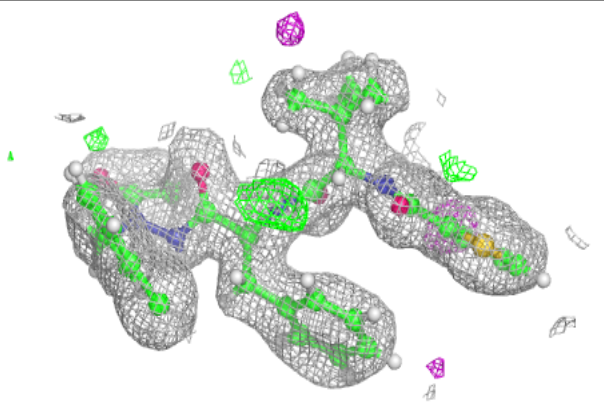
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around A1JNF 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)

**Electron density around A1JNF 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)



6.5 Other polymers [i](#)

There are no such residues in this entry.