WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Apr 7, 2022 – 04:14 pm BST

PDB ID : 7OFU
Title : Structure of SARS-CoV-2 Papain-like protease PLpro in complex with 3, 4-Dihydroxybenzoic acid, methyl ester
Authors : Srinivasan, V.; Ewert, W.; Werner, N.; Falke, S.; Guenther, S.; Reinke, P.; Sprenger, J.; Brognaro, H.; Ullah, N.; Andaleeb, H.; Perbandt, M.; Alves Franca, B.; Schwinzer, M.; Wang, M.; Wolf, M.; Lieske, J.; Koua, F.; Ginn, H.; Lane, T.J.; Yefanov, O.; Gelisio, L.; Hakanpaeae, J.; Saouane, S.; Tolstikova, A.; Groessler, M.; Fleckenstein, H.; Trost, F.; Lorenzen, K.; Schubert, R.; Han, H.; Schmidt, C.; Brings, L.; Galchenkova, M.; Gevorkov, Y.; Li, C.; Perk, A.; Awel, S.; Wahab, A.; Choudary, I.; Turk, D.; Hinrichs, W.; Chapman, H.N.; Meents, A.; Betzel, C.
Deposited on : 2021-05-05
Resolution : 1.72 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.27
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
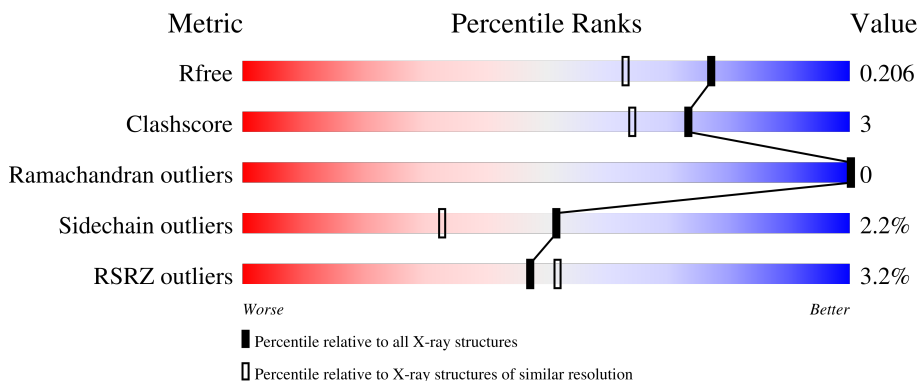Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)

1   Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
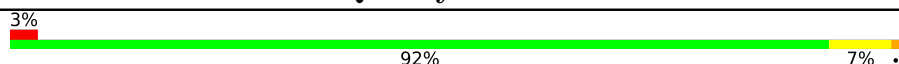*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.72 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	Percentile Ranks	Value
Rfree		0.206
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		2.2%
RSRZ outliers		3.2%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5722 (1.74-1.70)
Clashscore	141614	6152 (1.74-1.70)
Ramachandran outliers	138981	6051 (1.74-1.70)
Sidechain outliers	138945	6051 (1.74-1.70)
RSRZ outliers	127900	5629 (1.74-1.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=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	AAA	315	3% ▮ 92% 7% •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

| Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	PO4	AAA	605	-	-	X	-

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 2975 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 Papain-like protease nsp3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	315	Total	C	N	O	S	0	11	0
			2599	1646	435	499	19			

- Molecule 2 is methyl 3,4-bis(oxidanyl)benzoate (three-letter code: HE9) (formula: $C_8H_8O_4$) (labeled as "Ligand of Interest" by depositor).

HE9
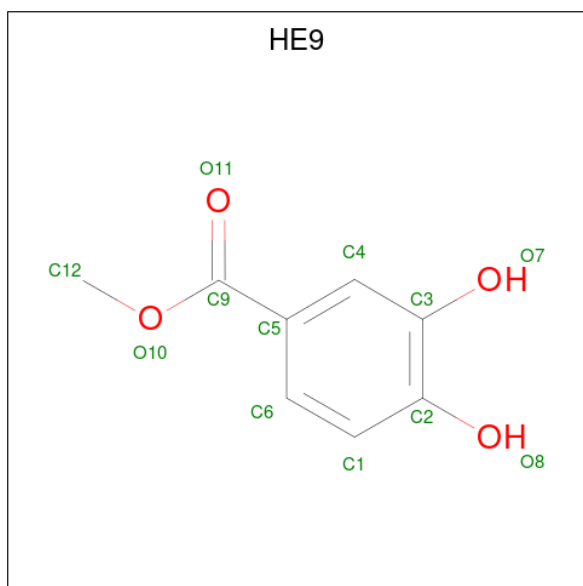
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	AAA	1	Total	C	O	0	0
			12	8	4		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

GOL

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

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	1	Total 1	Zn 1	0	0

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).

PO4
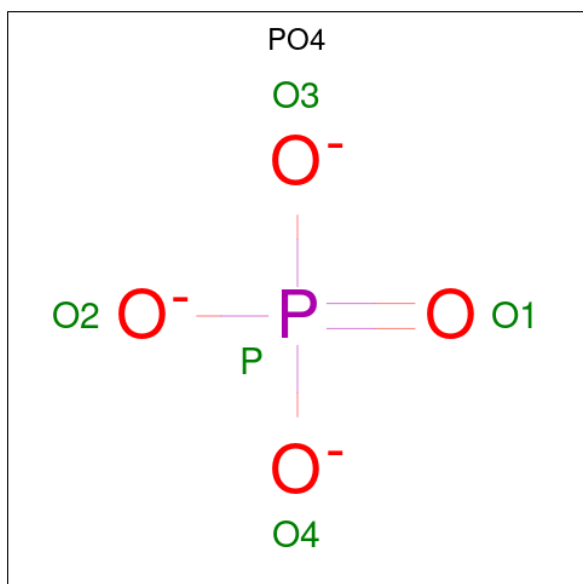
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	AAA	1	Total 5	O 4	P 1	0	0
5	AAA	1	Total 5	O 4	P 1	0	0
5	AAA	1	Total 5	O 4	P 1	0	0
5	AAA	1	Total 5	O 4	P 1	0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	AAA	1	Total 1	Cl 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	AAA	330	Total 330	O 330	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Papain-like protease nsp3

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	82.39Å　82.39Å　134.13Å 90.00°　90.00°　120.00°	Depositor
Resolution (Å)	48.87　–　1.72 48.87　–　1.72	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.87-1.72) 99.9 (48.87-1.72)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.82 (at 1.72Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.175　,　0.202 0.183　,　0.206	Depositor DCC
R_{free} test set	2834 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	28.8	Xtriage
Anisotropy	0.018	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.026 for -h,-k,l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	2975	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.16% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: CL, HE9, GOL, ZN, PO4

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	AAA	0.91	2/2659 (0.1%)	1.01	6/3604 (0.2%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	AAA	70	GLU	CD-OE2	6.67	1.32	1.25
1	AAA	280	GLU	CD-OE2	5.72	1.31	1.25

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	AAA	166	ARG	NE-CZ-NH2	-7.71	116.44	120.30
1	AAA	140	ARG	NE-CZ-NH2	-6.62	116.99	120.30
1	AAA	82	ARG	NE-CZ-NH2	-6.50	117.05	120.30
1	AAA	82	ARG	NE-CZ-NH1	6.42	123.51	120.30
1	AAA	140	ARG	NE-CZ-NH1	6.33	123.47	120.30
1	AAA	166	ARG	NE-CZ-NH1	6.05	123.33	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	AAA	2599	0	2522	13	0
2	AAA	12	0	0	2	0
3	AAA	12	0	16	4	0
4	AAA	1	0	0	0	0
5	AAA	20	0	0	4	0
6	AAA	1	0	0	0	0
7	AAA	330	0	0	3	0
All	All	2975	0	2538	17	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 (17) 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:AAA:89:HIS:NE2	3:AAA:602:GOL:H12	1.64	1.13
1:AAA:60[A]:ASN:ND2	7:AAA:702:HOH:O	1.94	0.99
5:AAA:607:PO4:O4	7:AAA:703:HOH:O	2.08	0.70
1:AAA:255:HIS:ND1	5:AAA:608:PO4:O2	2.30	0.64
2:AAA:601:HE9:O11	5:AAA:605:PO4:O2	2.17	0.63
1:AAA:269[A]:GLN:O	1:AAA:270:CYS:HB3	2.02	0.59
1:AAA:57:VAL:HG21	3:AAA:603:GOL:H31	1.84	0.58
1:AAA:138[A]:ARG:HH11	1:AAA:138[A]:ARG:HG2	1.72	0.53
1:AAA:61:ASP:O	1:AAA:65[B]:ARG:HG3	2.10	0.51
1:AAA:268[B]:TYR:H	1:AAA:268[B]:TYR:HD1	1.60	0.49
1:AAA:188:VAL:HG22	1:AAA:194:GLN:HG2	1.94	0.48
2:AAA:601:HE9:C12	5:AAA:605:PO4:O3	2.65	0.45
1:AAA:263:GLU:O	1:AAA:273:TYR:HA	2.17	0.45
1:AAA:60[A]:ASN:HD21	3:AAA:603:GOL:H2	1.82	0.44
3:AAA:602:GOL:H11	7:AAA:930:HOH:O	2.20	0.42
1:AAA:263:GLU:OE1	1:AAA:296:TYR:OH	2.35	0.41
1:AAA:186:ASN:HD22	1:AAA:186:ASN:HA	1.67	0.41

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	AAA	324/315 (103%)	312 (96%)	12 (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	AAA	286/275 (104%)	278 (97%)	8 (3%)	43	23

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	60[A]	ASN
1	AAA	60[B]	ASN
1	AAA	174[A]	GLN
1	AAA	174[B]	GLN
1	AAA	236	GLN
1	AAA	270	CYS
1	AAA	297	LYS
1	AAA	315	LYS

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

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 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	#$|Z| > 2$
5	PO4	AAA	608	-	4,4,4	0.85	0	6,6,6	0.51	0
3	GOL	AAA	603	-	5,5,5	0.42	0	5,5,5	0.79	0
2	HE9	AAA	601	-	12,12,12	1.94	2 (16%)	16,16,16	3.81	7 (43%)
3	GOL	AAA	602	-	5,5,5	0.57	0	5,5,5	0.98	0
5	PO4	AAA	606	-	4,4,4	1.55	1 (25%)	6,6,6	0.43	0
5	PO4	AAA	607	-	4,4,4	0.89	0	6,6,6	0.33	0
5	PO4	AAA	605	-	4,4,4	0.62	0	6,6,6	0.60	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	HE9	AAA	601	-	-	1/6/6/6	0/1/1/1
3	GOL	AAA	603	-	-	4/4/4/4	-
3	GOL	AAA	602	-	-	0/4/4/4	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	AAA	601	HE9	O10-C9	5.20	1.44	1.33
5	AAA	606	PO4	P-O1	2.75	1.57	1.50
2	AAA	601	HE9	C3-C2	-2.20	1.36	1.40

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	AAA	601	HE9	O10-C9-C5	11.00	129.01	112.34
2	AAA	601	HE9	O10-C9-O11	-6.51	110.72	123.45
2	AAA	601	HE9	C12-O10-C9	3.87	123.28	115.83
2	AAA	601	HE9	C6-C5-C4	-3.85	114.68	119.24
2	AAA	601	HE9	C1-C6-C5	3.75	125.14	120.78
2	AAA	601	HE9	C4-C3-C2	2.76	122.30	119.86
2	AAA	601	HE9	C6-C5-C9	2.07	125.08	120.40

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	AAA	603	GOL	O1-C1-C2-C3
3	AAA	603	GOL	C1-C2-C3-O3
3	AAA	603	GOL	O1-C1-C2-O2
3	AAA	603	GOL	O2-C2-C3-O3
2	AAA	601	HE9	C5-C9-O10-C12

There are no ring outliers.

6 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	AAA	608	PO4	1	0
3	AAA	603	GOL	2	0
2	AAA	601	HE9	2	0
3	AAA	602	GOL	2	0
5	AAA	607	PO4	1	0
5	AAA	605	PO4	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less 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 HE9 AAA 601

Bond lengths

Bond angles

Torsions

Rings
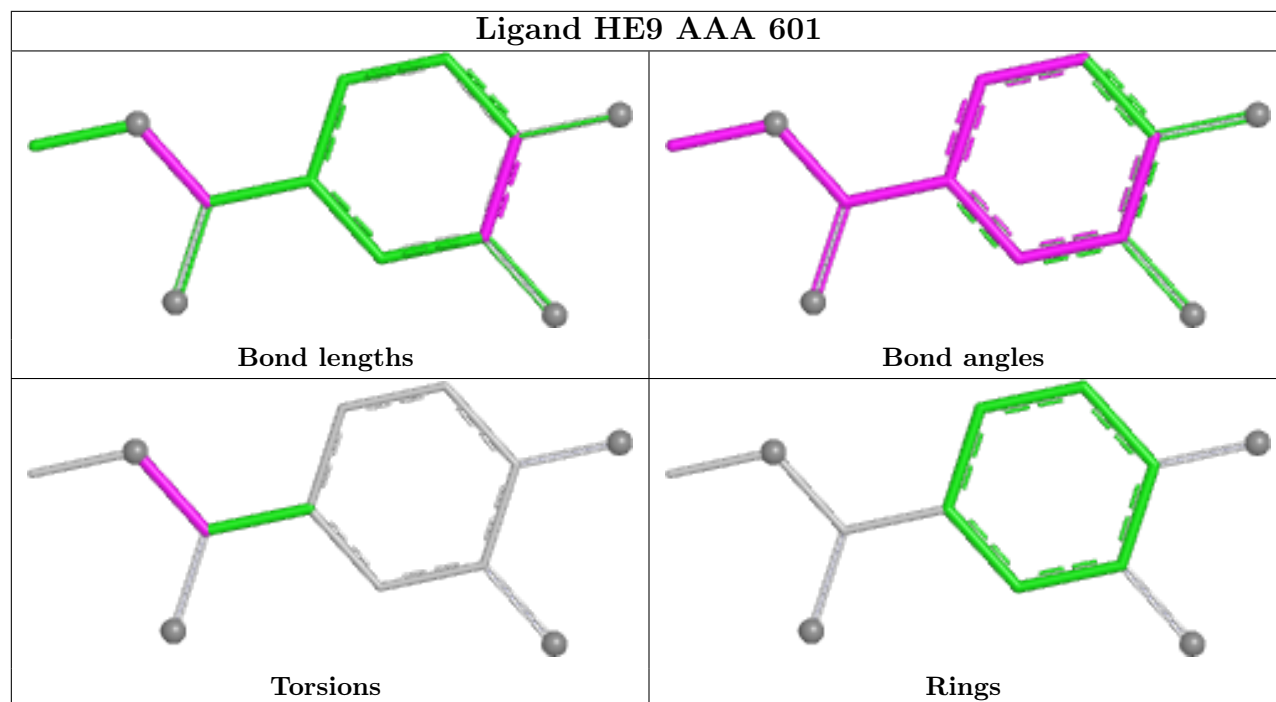
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	AAA	315/315 (100%)	0.08	10 (3%)	47	52	22, 32, 54, 92	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	315	LYS	7.3
1	AAA	2	VAL	4.9
1	AAA	314	ILE	4.2
1	AAA	1	GLU	3.9
1	AAA	268[A]	TYR	3.6
1	AAA	3	ARG	3.3
1	AAA	313	THR	2.5
1	AAA	44	ILE	2.2
1	AAA	187	VAL	2.2
1	AAA	24	SER	2.1

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

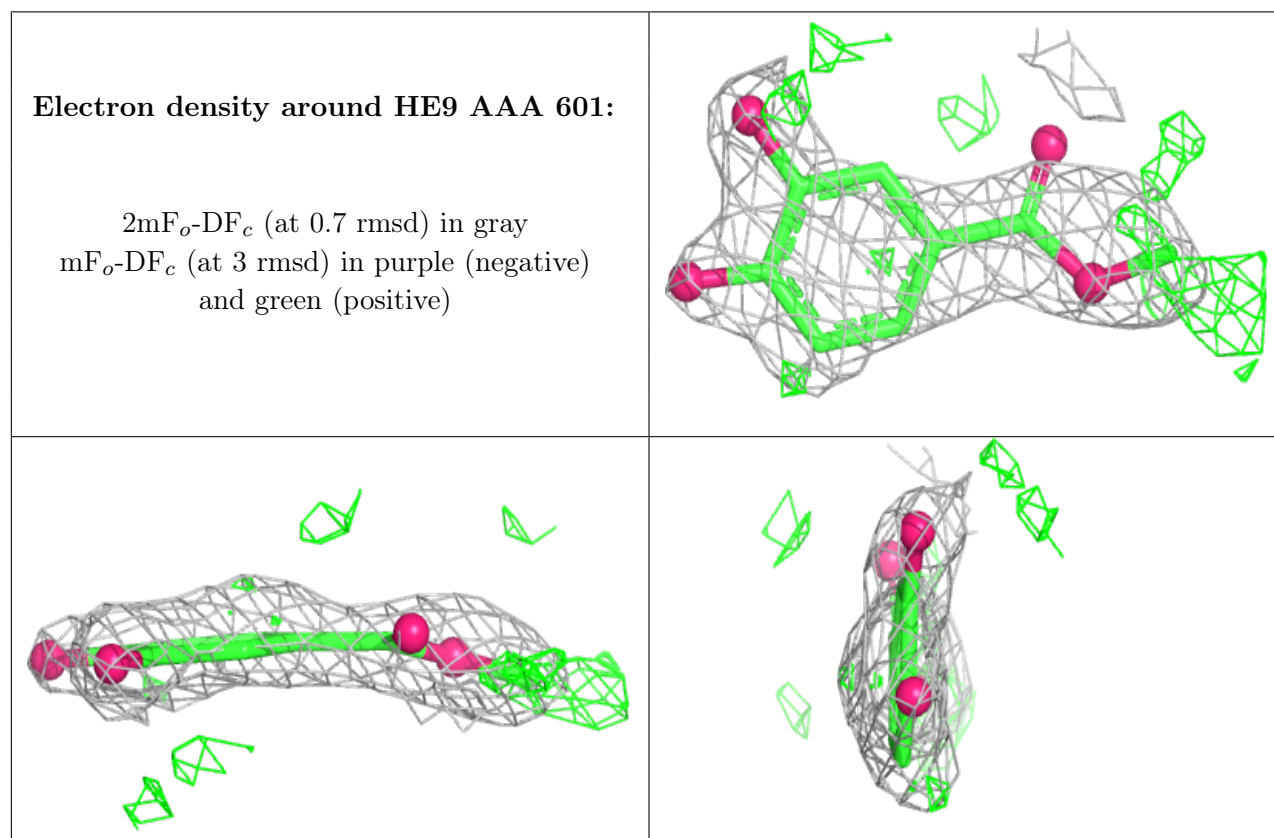There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	PO4	AAA	608	5/5	0.74	0.40	91,98,103,112	0
2	HE9	AAA	601	12/12	0.75	0.39	38,42,45,57	12
5	PO4	AAA	606	5/5	0.77	0.17	69,69,74,86	5
3	GOL	AAA	602	6/6	0.78	0.18	36,54,56,63	0
3	GOL	AAA	603	6/6	0.80	0.25	59,68,70,73	0
5	PO4	AAA	607	5/5	0.85	0.42	78,86,99,101	0
5	PO4	AAA	605	5/5	0.98	0.14	37,40,44,46	5
4	ZN	AAA	604	1/1	0.99	0.02	48,48,48,48	0
6	CL	AAA	609	1/1	1.00	0.08	27,27,27,27	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 HE9 AAA 601:

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

6.5　Other polymers　ⓘ

There are no such residues in this entry.