WORLDWIDE PDB

PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Sep 14, 2023 – 01:41 pm BST

PDB ID	:	8BO5
Title	:	COAGULATION FACTOR XI PROTEASE DOMAIN IN COMPLEX WITH ACTIVE SITE INHIBITOR 3
Authors	:	Schaefer, M.; Roehrig, S.; Ackerstaff, J.; Nunez, E.J.; Gericke, K.M.; Meier, K.; Tersteegen, A.; Stampfuss, J.; Ellerbrock, P.; Meibom, D.; Lang, D.; Heitmeier, S.; Hillisch, A.
Deposited on	:	2022-11-14
Resolution	:	1.70 Å(reported)

The following versions of software and data (see references ①) were used in the production of this report:

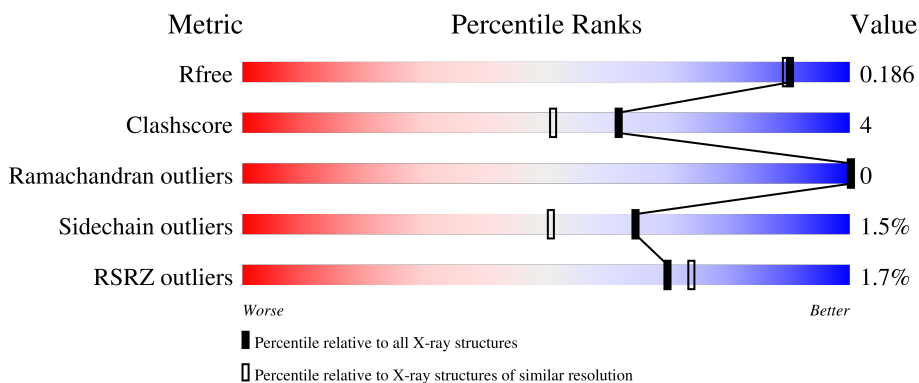|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 1.70 Å.

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

Metric	Percentile Ranks	Value
Rfree		0.186
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		1.5%
RSRZ outliers		1.7%

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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=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	238	2% ... 90% ... 8% ..

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:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	GOL	AAA	704[B]	-	-	X	-

WORLDWIDE
PDB
PROTEIN DATA BANK

2 Entity composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 2192 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 Coagulation factor XIa light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	236	Total 1940	C 1222	N 341	O 366	S 11	22	9	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	500	SER	CYS	engineered mutation	UNP P03951

- Molecule 2 is 2-[4-(3-chlorophenyl)-2,5-bis(oxidanylidene)pyrrolo[3,4-b]pyridin-1-yl]- {N}-[4-(1 {H}-1,2,3,4-tetrazol-5-yl)phenyl]ethanamide (three-letter code: QVO) (formula: $C_{22}H_{14}ClN_7O_3$) (labeled as "Ligand of Interest" by depositor).
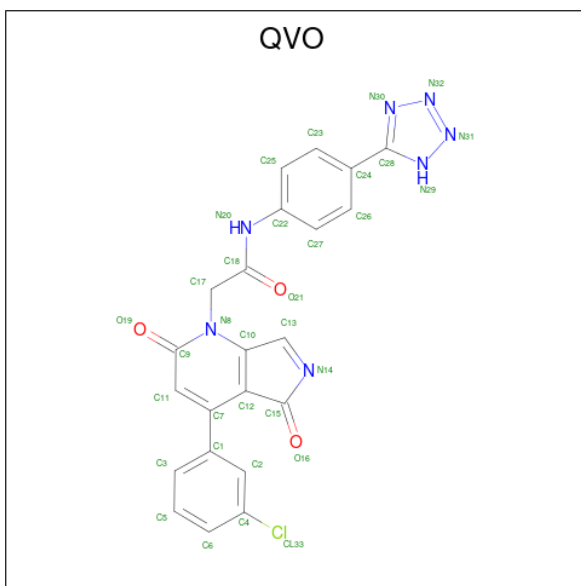
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	AAA	1	Total 33	C 22	Cl 1	N 7	O 3	0	0

- Molecule 3 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).

CIT

O5

O2
OH

C2 C1

C6 C3
HO
O6 C4 O1

HO
O7

C5

O3 O OH
O4
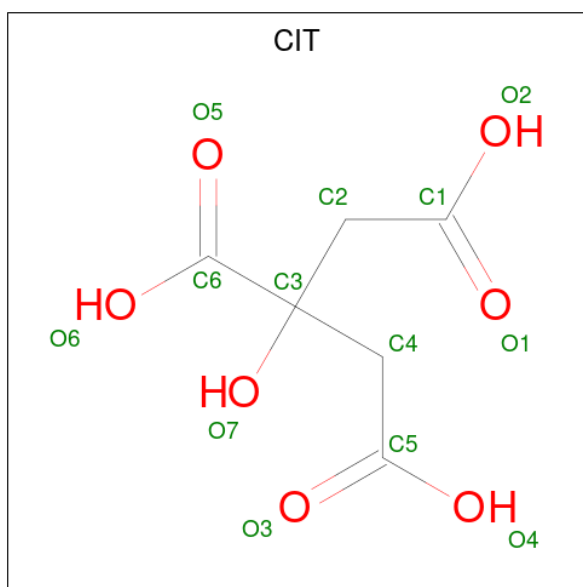
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	AAA	1	Total 13	C 6	O 7	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

GOL

C3
OH O3

C1 C2

HO OH
O1 O2
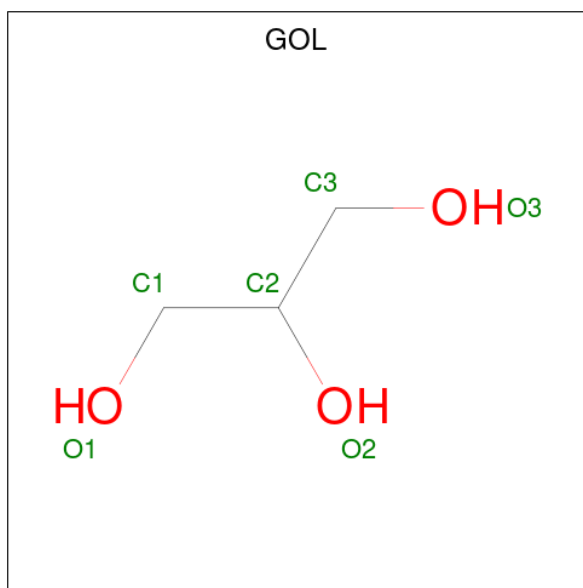
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	AAA	1	Total 12	C 6	O 6	0	1
4	AAA	1	Total 12	C 6	O 6	0	1

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	AAA	179	Total 182	O 182	0	7

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: Coagulation factor XIa light chain

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	58.76Å 59.74Å 66.93Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.59 – 1.70 44.57 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.3 (44.59-1.70) 99.3 (44.57-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.25 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.154 , 0.174 0.167 , 0.186	Depositor DCC
R_{free} test set	884 reflections (3.35%)	wwPDB-VP
Wilson B-factor (Å^2)	19.2	Xtriage
Anisotropy	0.242	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 48.8	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.023 for k,h,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	2192	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.11% 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: GOL, QVO, CIT

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	0.72	2/1988 (0.1%)	0.92	5/2692 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	AAA	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	AAA	550	ARG	NE-CZ	-7.15	1.23	1.33
1	AAA	437	GLU	CD-OE1	5.05	1.31	1.25

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	AAA	550	ARG	NE-CZ-NH1	-16.72	111.94	120.30
1	AAA	550	ARG	NE-CZ-NH2	13.82	127.21	120.30
1	AAA	550	ARG	CD-NE-CZ	6.03	132.04	123.60
1	AAA	543	GLU	CG-CD-OE2	5.31	128.93	118.30
1	AAA	543	GLU	CG-CD-OE1	-5.21	107.88	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	AAA	550	ARG	Sidechain

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	AAA	1940	0	1890	17	0
2	AAA	33	0	0	2	0
3	AAA	13	0	5	0	0
4	AAA	24	0	32	5	0
5	AAA	182	0	0	3	0
All	All	2192	0	1927	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 4.

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:476[A]:GLU:OE2	5:AAA:801:HOH:O	1.64	1.12
1:AAA:543:GLU:HG3	5:AAA:863:HOH:O	1.75	0.85
1:AAA:575[B]:SER:OG	2:AAA:701:QVO:C9	2.26	0.84
1:AAA:584:ASN:N	4:AAA:704[B]:GOL:H11	1.97	0.79
1:AAA:431:HIS:HD2	1:AAA:480:ASP:OD2	1.66	0.77
1:AAA:575[B]:SER:OG	2:AAA:701:QVO:O19	2.18	0.60
1:AAA:522:ARG:HE	1:AAA:529:GLN:CD	2.08	0.56
1:AAA:584:ASN:H	4:AAA:704[B]:GOL:H11	1.67	0.56
1:AAA:540[A]:VAL:HG22	1:AAA:560:CYS:HB2	1.89	0.54
1:AAA:400:PRO:HA	4:AAA:703[B]:GOL:H12	1.92	0.50
1:AAA:540[A]:VAL:CG2	1:AAA:560:CYS:HB2	2.44	0.47
1:AAA:585:GLU:H	4:AAA:704[B]:GOL:H11	1.81	0.45
1:AAA:507:ARG:NE	1:AAA:613:GLU:OE2	2.50	0.44
1:AAA:572:LYS:NZ	5:AAA:812:HOH:O	2.52	0.43
1:AAA:585:GLU:H	4:AAA:704[B]:GOL:C1	2.34	0.41
1:AAA:522:ARG:HE	1:AAA:529:GLN:NE2	2.18	0.41
1:AAA:539:LEU:HD13	1:AAA:559:ILE:HD11	2.03	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	243/238 (102%)	236 (97%)	7 (3%)	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	212/204 (104%)	209 (99%)	3 (1%)	67	53

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	507	ARG
1	AAA	523	LYS
1	AAA	610	ASN

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 (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)

6 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	CIT	AAA	702	-	12,12,12	1.15	2 (16%)	17,17,17	1.43	3 (17%)
4	GOL	AAA	703[B]	-	5,5,5	0.10	0	5,5,5	0.24	0
2	QVO	AAA	701	-	36,37,37	2.12	9 (25%)	44,53,53	1.57	11 (25%)
4	GOL	AAA	704[B]	-	5,5,5	0.07	0	5,5,5	0.29	0
4	GOL	AAA	704[A]	-	5,5,5	0.08	0	5,5,5	0.22	0
4	GOL	AAA	703[A]	-	5,5,5	0.13	0	5,5,5	0.34	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
3	CIT	AAA	702	-	-	0/16/16/16	-
4	GOL	AAA	703[B]	-	-	2/4/4/4	-
2	QVO	AAA	701	-	-	2/16/25/25	0/5/5/5
4	GOL	AAA	704[B]	-	-	2/4/4/4	-
4	GOL	AAA	704[A]	-	-	3/4/4/4	-
4	GOL	AAA	703[A]	-	-	3/4/4/4	-

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	AAA	701	QVO	C24-C28	-5.34	1.34	1.48
2	AAA	701	QVO	C13-N14	4.93	1.47	1.33
2	AAA	701	QVO	C17-N8	4.17	1.52	1.46
2	AAA	701	QVO	C1-C7	-4.12	1.42	1.48
2	AAA	701	QVO	C9-N8	-4.01	1.34	1.40
2	AAA	701	QVO	C4-CL33	3.45	1.82	1.74
2	AAA	701	QVO	C22-N20	-3.03	1.35	1.41
2	AAA	701	QVO	C10-N8	-2.41	1.35	1.41
2	AAA	701	QVO	C15-N14	-2.32	1.33	1.37
3	AAA	702	CIT	C3-C6	2.07	1.55	1.53
3	AAA	702	CIT	O4-C5	-2.01	1.24	1.30

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	AAA	702	CIT	O5-C6-C3	-3.90	116.73	122.25
2	AAA	701	QVO	C28-N29-N31	3.38	107.88	104.87
2	AAA	701	QVO	O16-C15-N14	-2.99	118.40	124.41
2	AAA	701	QVO	C22-N20-C18	-2.89	122.43	127.50
2	AAA	701	QVO	N29-N31-N32	-2.84	107.68	109.53
2	AAA	701	QVO	C1-C2-C4	2.77	122.36	119.19
3	AAA	702	CIT	O6-C6-C3	2.76	117.85	113.05
2	AAA	701	QVO	O21-C18-N20	2.69	128.53	123.63
2	AAA	701	QVO	C11-C9-N8	2.47	119.00	115.08
2	AAA	701	QVO	C28-N30-N32	2.39	106.99	104.87
2	AAA	701	QVO	N30-N32-N31	-2.28	108.05	109.53
2	AAA	701	QVO	O21-C18-C17	-2.27	117.10	121.08
2	AAA	701	QVO	C1-C7-C11	-2.14	115.12	120.35
3	AAA	702	CIT	O1-C1-C2	-2.07	116.90	122.94

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	AAA	703[A]	GOL	O1-C1-C2-C3
4	AAA	703[A]	GOL	C1-C2-C3-O3
4	AAA	703[B]	GOL	C1-C2-C3-O3
4	AAA	704[A]	GOL	O1-C1-C2-C3
4	AAA	704[B]	GOL	O1-C1-C2-C3
4	AAA	703[A]	GOL	O1-C1-C2-O2
4	AAA	704[A]	GOL	O1-C1-C2-O2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	AAA	703[B]	GOL	O2-C2-C3-O3
4	AAA	704[B]	GOL	O1-C1-C2-O2
2	AAA	701	QVO	C2-C1-C7-C11
4	AAA	704[A]	GOL	O2-C2-C3-O3
2	AAA	701	QVO	C3-C1-C7-C11

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	AAA	703[B]	GOL	1	0
2	AAA	701	QVO	2	0
4	AAA	704[B]	GOL	4	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 QVO AAA 701

Bond lengths

Bond angles

Torsions

Rings
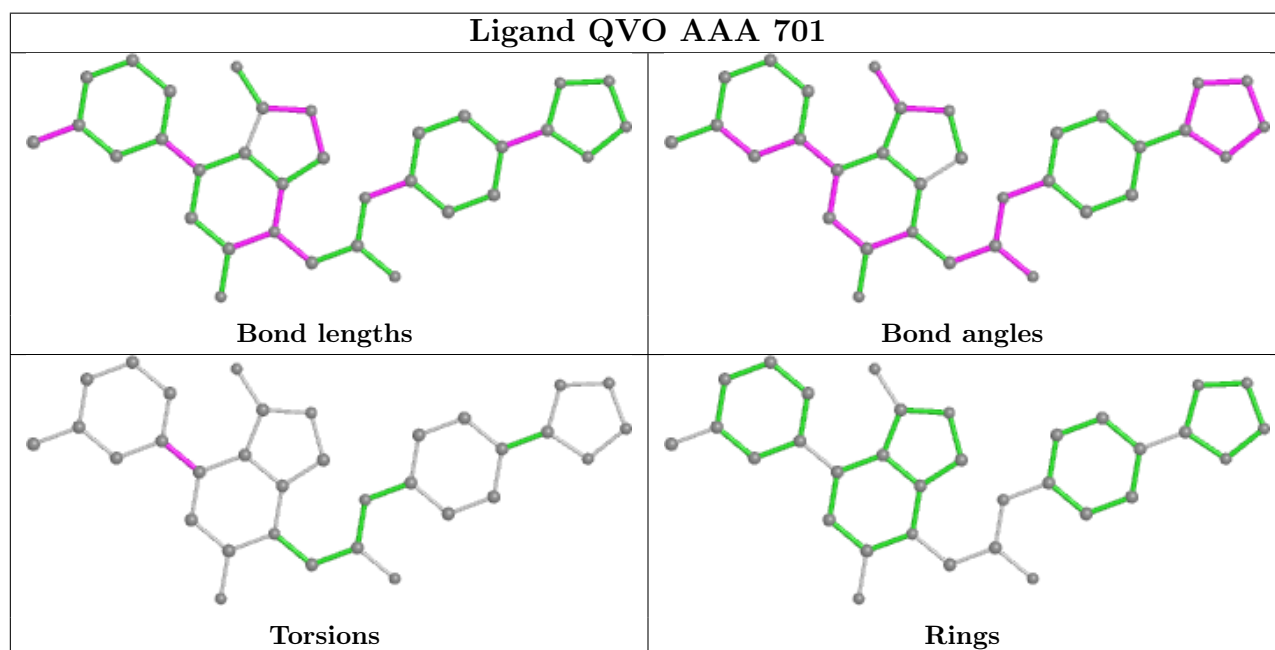
5.7 Other polymers ⓘ

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 (i)

6.1 Protein, DNA and RNA chains (i)

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	236/238 (99%)	-0.30	4 (1%)	70	74	19, 28, 46, 69	9 (3%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	504	LYS	3.2
1	AAA	524	LEU	2.3
1	AAA	505	GLY	2.2
1	AAA	508	ASN	2.1

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.
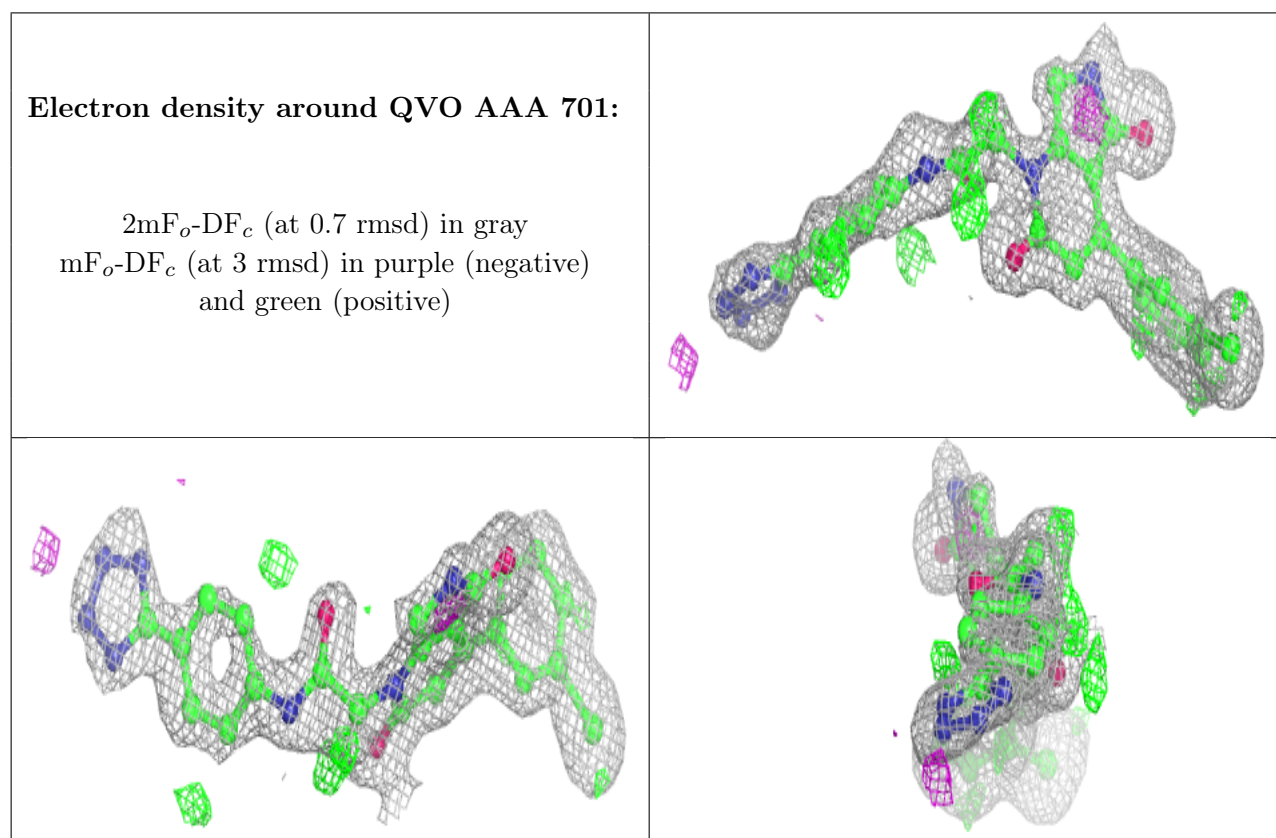
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	QVO	AAA	701	33/33	0.83	0.17	29,36,49,51	33
4	GOL	AAA	704[A]	6/6	0.84	0.15	57,59,59,59	6
4	GOL	AAA	704[B]	6/6	0.84	0.15	33,35,38,40	6

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å²)	Q<0.9
4	GOL	AAA	703[B]	6/6	0.86	0.16	54,57,58,61	6
3	CIT	AAA	702	13/13	0.86	0.17	41,44,54,55	0
4	GOL	AAA	703[A]	6/6	0.86	0.16	31,33,34,36	6

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 QVO AAA 701:

$2mF_o\text{-}DF_c$ (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.