

Full wwPDB X-ray Structure Validation Report (i)

Jun 24, 2024 – 07:13 PM EDT

PDB ID	:	6X3O
Title	:	Co-structure of BTK kinase domain with L-005191930 inhibitor
Authors	:	Fischmann, T.O.
Deposited on		
Resolution	:	1.90 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 (i) 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 (1)) were used in the production of this report:

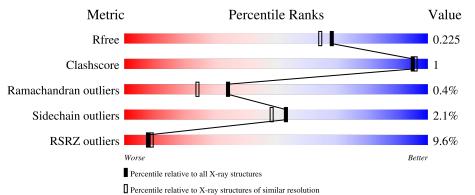
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
buster-report	:	1.1.7(2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY \, DIFFRACTION$

The reported resolution of this entry is 1.90 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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R_{free}	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	А	271	93%	••					
1	В	271	94%	•					



6X3O

2 Entry composition (i)

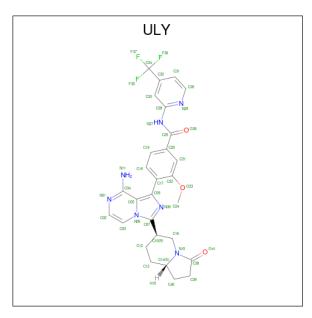
There are 4 unique types of molecules in this entry. The entry contains 4889 atoms, of which 0 are hydrogens and 0 are deuteriums.

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.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
1	Λ	265	Total	С	Ν	0	S	0	2	1
	A	205	2186	1401	360	405	20	0	2	
1	В	267	Total	С	Ν	0	S	0	9	1
	D	207	2203	1412	362	409	20	0)	

• Molecule 1 is a protein called Tyrosine-protein kinase BTK.

Molecule 2 is 4-{8-amino-3-[(6R,8aS)-3-oxo-3,5,6,7,8,8a-hexahydroindolizin-6-yl]imidazo[1,5 -a]pyrazin-1-yl}-3-methoxy-N-[4-(trifluoromethyl)pyridin-2-yl]benzamide (three-letter code: ULY) (formula: C₂₈H₂₆F₃N₇O₃) (labeled as "Ligand of Interest" by depositor).

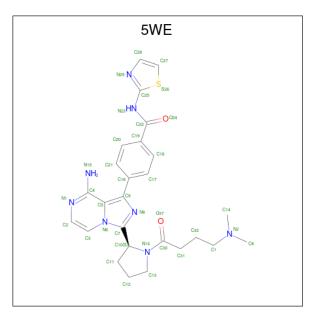


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
2	А	1	Total 41	C 28				0	0
2	В	1	Total 41	C 28				0	0

• Molecule 3 is 4-[8-azanyl-3-[(2 {S})-1-[4-(dimethylamino)butanoyl]pyrrolidin-2-yl]imidaz o[1,5-a]pyrazin-1-yl]- {N}-(1,3-thiazol-2-yl)benzamide (three-letter code: 5WE) (formula:



 ${\rm C}_{26}{\rm H}_{30}{\rm N}_{8}{\rm O}_{2}{\rm S}\big).$



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
3	Δ	1	Total					0	0
5	Л	1	37	26	8	2	1	0	0
9	D	1	Total	С	Ν	Ο	S	0	0
J	D	1	37	26	8	2	1	0	

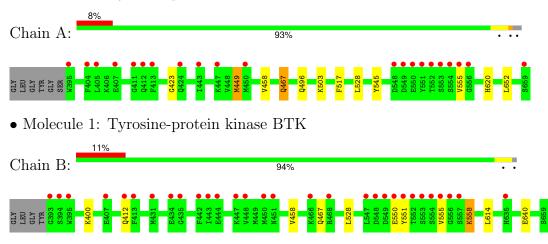
• Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	169	Total O 169 169	0	0
4	В	175	Total O 175 175	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: Tyrosine-protein kinase BTK



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
$\begin{array}{c} \text{Cell constants} \\ \text{a, b, c, } \alpha, \beta, \gamma \end{array}$	$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$	Depositor
Resolution (Å)	$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$	Depositor EDS
% Data completeness	99.0(34.79-1.90)	Depositor
(in resolution range)	99.0(34.79-1.90)	EDS
R _{merge}	0.07	Depositor
R _{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	1.99 (at 1.89 Å)	Xtriage
Refinement program	BUSTER	Depositor
R, R_{free}	$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$	Depositor DCC
R_{free} test set	2215 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	24.0	Xtriage
Anisotropy	0.129	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.33, 50.0	EDS
L-test for twinning ²	$< L > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.067 for h,-k,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4889	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 8.10% of the height of the origin peak. No significant pseudotranslation is detected.

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.



¹Intensities estimated from amplitudes.

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: ULY, $5\mathrm{WE}$

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	Bo	nd lengths	Bond angles		
	Unam	RMSZ	# Z > 5	RMSZ	# Z > 5	
1	А	0.54	1/2237~(0.0%)	0.62	0/3015	
1	В	0.53	0/2254	0.62	0/3038	
All	All	0.53	1/4491~(0.0%)	0.62	0/6053	

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	А	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	А	449	MET	SD-CE	-5.06	1.49	1.77

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	А	545	TYR	Sidechain

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



Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	А	2186	0	2149	5	0
1	В	2203	0	2165	3	0
2	А	41	0	0	0	0
2	В	41	0	0	0	0
3	А	37	0	30	2	0
3	В	37	0	30	2	0
4	А	169	0	0	0	0
4	В	175	0	0	0	0
All	All	4889	0	4374	11	0

the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

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 1.

All (11) 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)
3:B:702:5WE:H9	3:B:702:5WE:H7	1.24	1.00
3:A:702:5WE:H9	3:A:702:5WE:H7	1.31	0.96
1:A:620:HIS:HA	1:B:400:LYS:HD2	1.78	0.64
3:B:702:5WE:H9	3:B:702:5WE:N15	2.07	0.61
1:A:449:MET:CE	1:A:517:PHE:CZ	2.94	0.50
3:A:702:5WE:H9	3:A:702:5WE:N15	2.14	0.48
1:A:449:MET:CE	1:A:517:PHE:HZ	2.27	0.47
1:A:503:LYS:HE2	1:A:652:LEU:HD21	1.97	0.47
1:B:551:TYR:CZ	1:B:558:LYS:HD2	2.53	0.44
1:A:458:VAL:HG21	1:A:528:LEU:HD12	2.01	0.43
1:B:458:VAL:HG21	1:B:528:LEU:HD12	2.01	0.43

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	Percer	ntiles
1	А	265/271~(98%)	259~(98%)	4(2%)	2(1%)	19	9
1	В	268/271~(99%)	263~(98%)	5(2%)	0	100	100
All	All	533/542~(98%)	522 (98%)	9~(2%)	2 (0%)	34	24

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	А	423	GLY
1	А	467	GLN

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent side chain 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	А	241/243~(99%)	238~(99%)	3~(1%)	71 70		
1	В	243/243~(100%)	236~(97%)	7(3%)	42 35		
All	All	484/486 (100%)	474 (98%)	10 (2%)	53 48		

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	А	467	GLN
1	А	496	GLN
1	А	555	VAL
1	В	412	GLN
1	В	467	GLN
1	В	550	GLU
1	В	555	VAL
1	В	558	LYS
1	В	614	LEU
1	В	640	GLU

Sometimes side chains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such side chains are listed below:



Mol	Chain	Res	Type
1	А	650	ASN

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)

4 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	Turne	pe Chain Res Lir		Link	Bond lengths				Bond angles		
	Type	Chain	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z > 2	
3	5WE	В	702	-	33,41,41	1.46	5 (15%)	$35,\!58,\!58$	1.09	1 (2%)	
3	5WE	А	702	-	33,41,41	1.49	8 (24%)	$35,\!58,\!58$	1.12	2(5%)	
2	ULY	В	701	-	41,46,46	0.40	0	53,69,69	0.76	2 (3%)	
2	ULY	А	701	-	41,46,46	0.41	0	53,69,69	0.71	1 (1%)	

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	5WE	В	702	-	-	2/20/36/36	0/5/5/5

Continued on next page...



Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	5WE	А	702	-	-	6/20/36/36	0/5/5/5
2	ULY	В	701	-	-	2/21/46/46	0/6/6/6
2	ULY	А	701	-	-	2/21/46/46	0/6/6/6

Continued from previous page...

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Ζ	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
3	А	702	5WE	C30-N14	3.80	1.46	1.35
3	В	702	5WE	C30-N14	3.61	1.45	1.35
3	В	702	5WE	C3-N6	2.83	1.42	1.38
3	А	702	5WE	C3-N6	2.78	1.42	1.38
3	В	702	5WE	C2-N1	2.54	1.39	1.34
3	В	702	5WE	C22-N23	2.47	1.42	1.35
3	А	702	5WE	C18-C17	2.30	1.42	1.38
3	А	702	5WE	C18-C19	2.18	1.42	1.39
3	А	702	5WE	C2-N1	2.18	1.39	1.34
3	В	702	5WE	C25-N23	2.12	1.40	1.36
3	А	702	5WE	C22-N23	2.08	1.41	1.35
3	А	702	5WE	C25-N23	2.07	1.40	1.36
3	А	702	5WE	C20-C19	2.00	1.42	1.39

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
3	В	702	5WE	C11-C10-N14	4.02	106.40	101.93
3	А	702	5WE	C11-C10-N14	3.96	106.34	101.93
2	В	701	ULY	C32-C33-C28	-3.00	118.54	122.65
2	А	701	ULY	C32-C33-C28	-2.85	118.75	122.65
2	В	701	ULY	C16-C10-C07	-2.02	106.87	111.89
3	А	702	5WE	C13-N14-C10	-2.01	109.61	111.88

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	А	701	ULY	N06-C07-C10-C12
2	А	701	ULY	N08-C09-C17-C18
2	В	701	ULY	N06-C07-C10-C12
2	В	701	ULY	N08-C09-C17-C18
3	А	702	5WE	C30-C31-C32-C1

Continued on next page...



\mathbf{Mol}	Chain	Res	Type	Atoms
3	А	702	5WE	C32-C1-N2-C8
3	А	702	5WE	C32-C1-N2-C14
3	А	702	5WE	N2-C1-C32-C31
3	А	702	5WE	C17-C16-C9-C5
3	В	702	5WE	C17-C16-C9-C5
3	А	702	5WE	C21-C16-C9-C5
3	В	702	5WE	C21-C16-C9-C5

Continued from previous page...

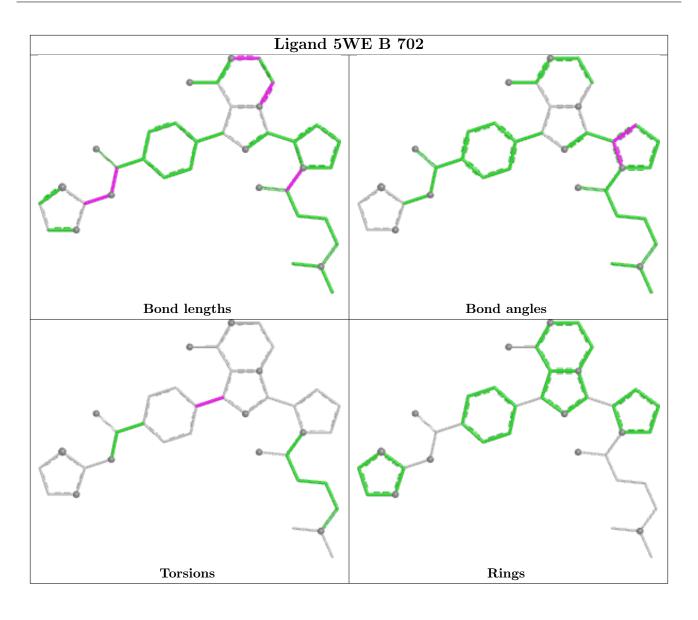
There are no ring outliers.

2 monomers are involved in 4 short contacts:

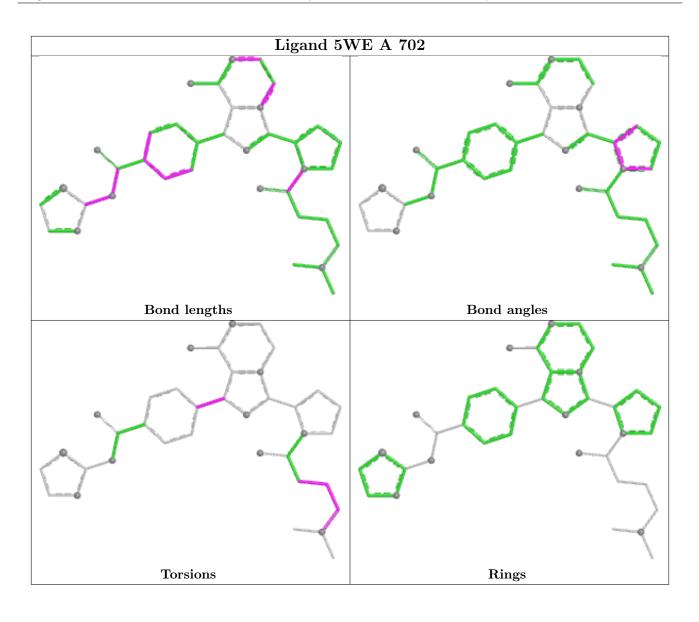
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	В	702	5WE	2	0
3	А	702	5WE	2	0

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.

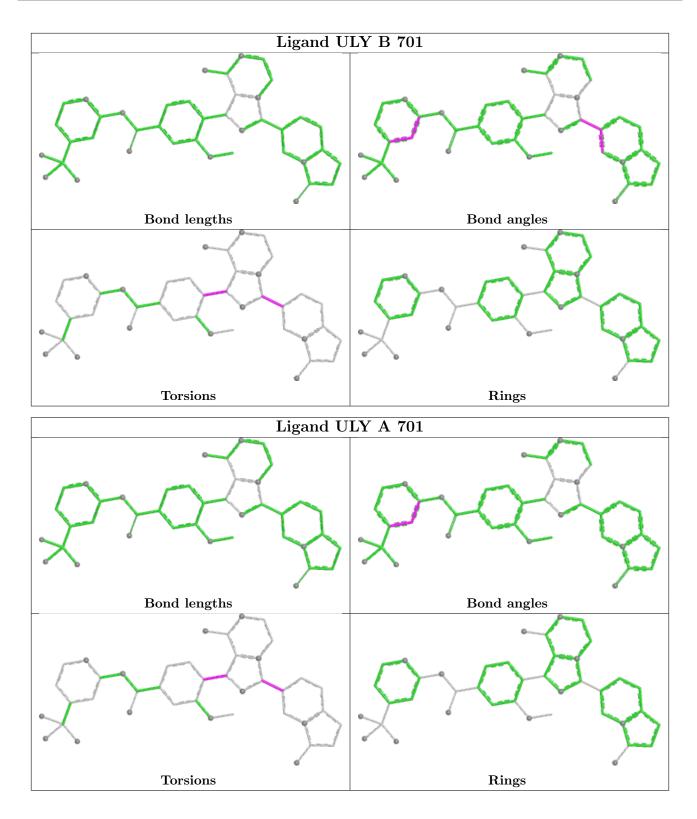












5.7 Other polymers (i)

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 (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, 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	$\langle RSRZ \rangle$	#RSRZ>2	$OWAB(Å^2)$	Q<0.9
1	А	265/271~(97%)	0.47	21 (7%) 12 14	14, 29, 62, 79	0
1	В	267/271 (98%)	0.60	30 (11%) 5 6	12, 27, 69, 90	0
All	All	532/542~(98%)	0.53	51 (9%) 8 9	12, 28, 66, 90	0

All (51) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	В	555	VAL	11.1
1	В	393	GLY	9.9
1	В	395	TRP	8.6
1	А	555	VAL	7.8
1	А	395	TRP	7.4
1	В	549	ASP	6.7
1	В	413	PHE	6.0
1	В	447	LYS	5.9
1	А	413	PHE	5.4
1	В	551	TYR	5.1
1	В	548	ASP	5.0
1	В	448	VAL	4.9
1	В	550	GLU	4.3
1	А	554	SER	4.2
1	А	412	GLN	4.1
1	А	549	ASP	3.8
1	В	394	SER	3.7
1	В	556	GLY	3.7
1	А	553	SER	3.7
1	В	434	GLU	3.6
1	В	435	GLY	3.6
1	В	557	SER	3.5
1	А	405	LEU	3.4
1	B	553	SER	3.4

Continued on next page...



Mol	Chain Res		Type	RSRZ
1	А	659	SER	3.3
1	В	412	GLN	3.2
1	В	450	MET	3.1
1	А	443	ILE	2.9
1	А	548	ASP	2.8
1	В	554	SER	2.7
1	В	407	GLU	2.7
1	А	552	THR	2.7
1	В	466	LYS	2.5
1	А	551	TYR	2.4
1	А	550	GLU	2.4
1	А	407	GLU	2.4
1	В	443	ILE	2.3
1	В	431	MET	2.2
1	В	444	GLU	2.2
1	В	468	ARG	2.2
1	А	450	MET	2.2
1	В	552	THR	2.2
1	В	451	ASN	2.2
1	А	556	GLY	2.2
1	В	442	PHE	2.2
1	В	635	HIS	2.1
1	В	547	LEU	2.1
1	А	411	GLY	2.1
1	А	424	GLN	2.1
1	А	447	LYS	2.1
1	А	404	PHE	2.0

Continued from previous page...

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)

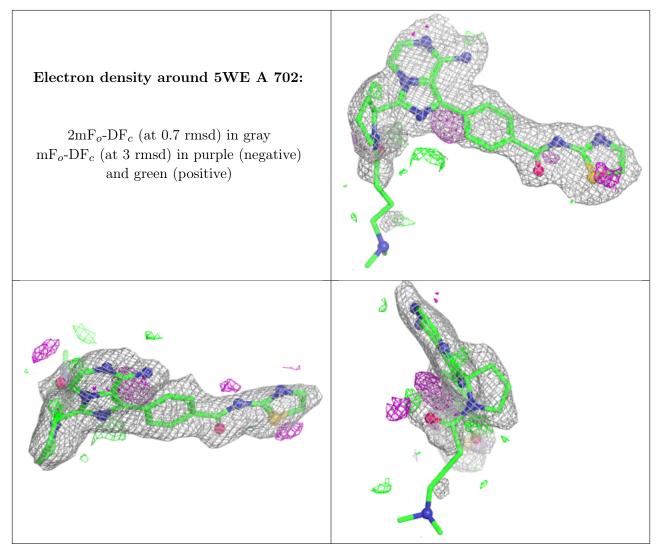
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,



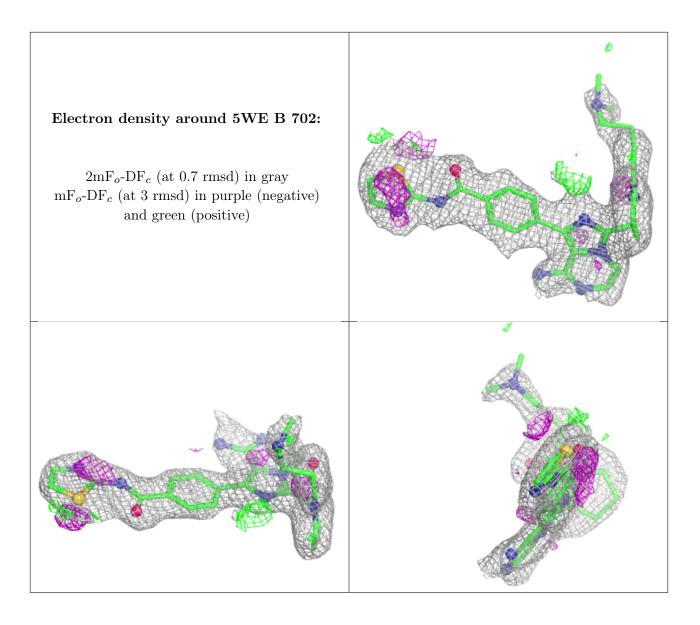
Mol	Type	Chain	Res	Atoms	RSCC	RSR	$B-factors(A^2)$	Q<0.9
3	5WE	А	702	37/37	0.81	0.23	45,57,76,76	0
3	5WE	В	702	37/37	0.85	0.19	46,53,61,62	0
2	ULY	А	701	41/41	0.92	0.11	20,34,48,51	0
2	ULY	В	701	41/41	0.94	0.11	16,28,39,45	0

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.

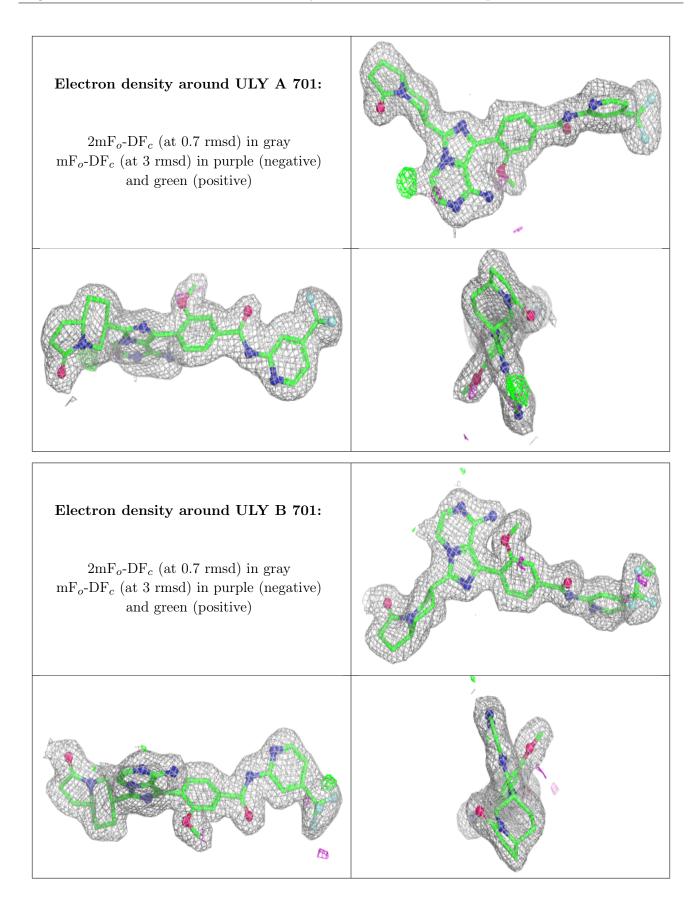
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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













6.5 Other polymers (i)

There are no such residues in this entry.

