Full wwPDB X-ray Structure Validation Report (i)

Sep 24, 2023 – 03:06 AM EDT

PDB ID	:	5T0P
Title	:	c-Src kinase domain in complex with Rao-IV-151
Authors	:	Dieter, E.M.; Merritt, E.A.; Maly, D.J.
Deposited on	:	2016-08-16
Resolution	:	2.50 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 (i)) were used in the production of this report:

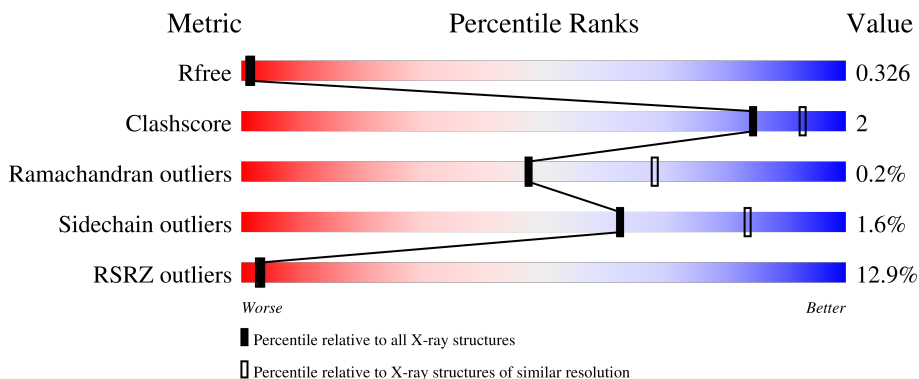|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 2.50 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		0.326
Clashscore		2
Ramachandran outliers		0.2%
Sidechain outliers		1.6%
RSRZ outliers		12.9%

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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	A	286	15% / 84% 8% · 8%
1	B	286	8% / 87% 5% 8%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4329 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 Proto-oncogene tyrosine-protein kinase Src.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	264	Total	C	N	O	S	0	0	0
			2104	1350	350	386	18			
1	B	263	Total	C	N	O	S	0	0	0
			2095	1344	348	385	18			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	248	GLY	-	expression tag	UNP P00523
A	249	HIS	-	expression tag	UNP P00523
A	250	MET	-	expression tag	UNP P00523
A	313	CYS	VAL	engineered mutation	UNP P00523
B	248	GLY	-	expression tag	UNP P00523
B	249	HIS	-	expression tag	UNP P00523
B	250	MET	-	expression tag	UNP P00523
B	313	CYS	VAL	engineered mutation	UNP P00523

- Molecule 2 is (2E)-3-[4-amino-5-(4-chlorophenyl)-7-(2-methoxyethyl)-7H-pyrrolo[2,3-d]pyrimidin-6-yl]-2-cyano-N-(propan-2-yl)prop-2-enamide (three-letter code: 73A) (formula: $C_{22}H_{23}ClN_6O_2$).
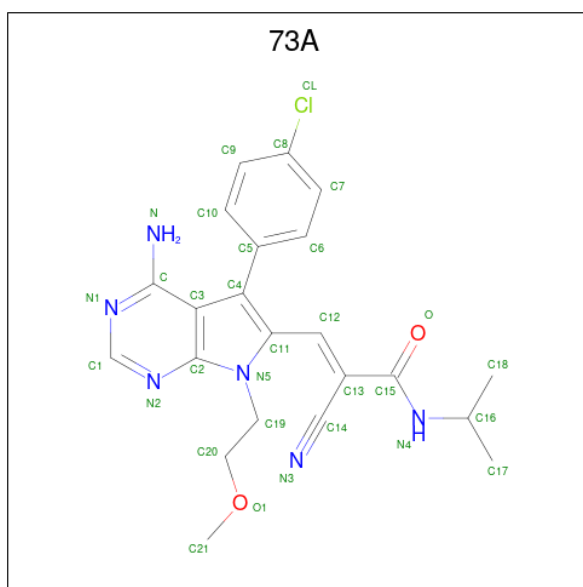
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Cl	N	O	0	0
			31	22	1	6	2		
2	B	1	Total	C	Cl	N	O	0	0
			31	22	1	6	2		

• Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	38	Total	O	0	0
			38	38		
3	B	30	Total	O	0	0
			30	30		

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: Proto-oncogene tyrosine-protein kinase Src

Chain A:

• Molecule 1: Proto-oncogene tyrosine-protein kinase Src

Chain B:
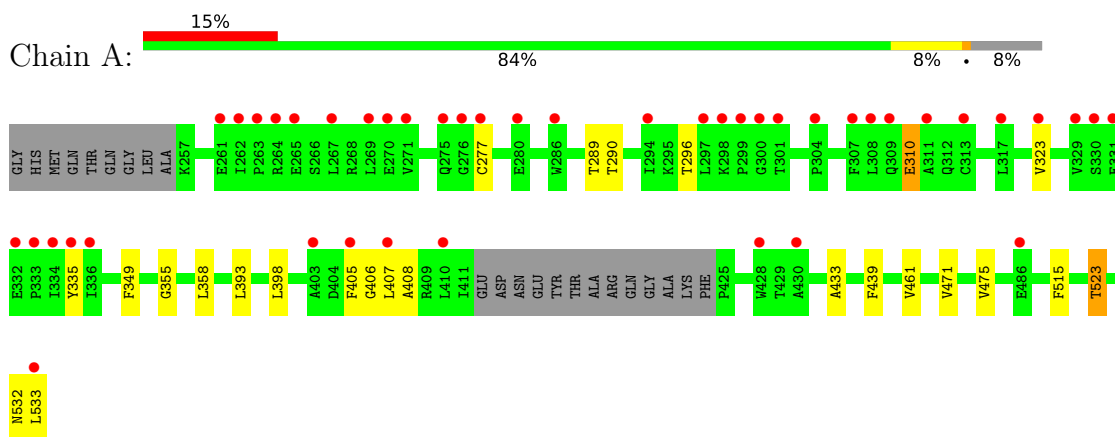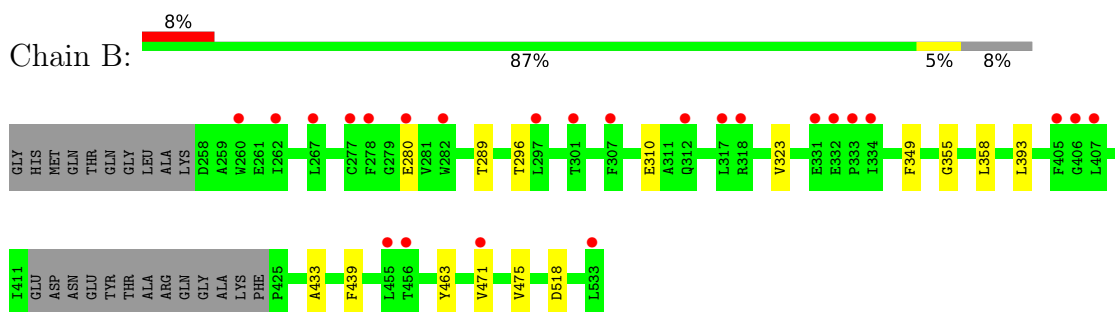
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	41.67Å 63.92Å 74.99Å 102.08° 90.36° 90.14°	Depositor
Resolution (Å)	73.33 – 2.50 43.31 – 2.50	Depositor EDS
% Data completeness (in resolution range)	76.2 (73.33-2.50) 76.2 (43.31-2.50)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.63 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, R_{free}	0.264 , 0.323 0.267 , 0.326	Depositor DCC
R_{free} test set	1045 reflections (5.24%)	wwPDB-VP
Wilson B-factor (Å^2)	46.2	Xtriage
Anisotropy	0.911	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 42.6	EDS
L-test for twinning[2]	$< |L| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.012 for h,-k,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	4329	wwPDB-VP
Average B, all atoms (Å^2)	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 17.26% 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: 73A

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	0.53	0/2155	0.72	0/2918
1	B	0.53	0/2146	0.72	0/2907
All	All	0.53	0/4301	0.72	0/5825

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	2104	0	2082	12	0
1	B	2095	0	2069	8	0
2	A	31	0	0	0	0
2	B	31	0	0	0	0
3	A	38	0	0	1	0
3	B	30	0	0	1	1
All	All	4329	0	4151	19	1

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 2.

All (19) 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:323:VAL:HG21	1:A:393:LEU:HD12	1.71	0.72
1:B:323:VAL:HG21	1:B:393:LEU:HD12	1.74	0.69
1:A:398:LEU:HD23	3:A:715:HOH:O	1.98	0.62
1:A:405:PHE:HB2	1:A:408:ALA:HB2	1.86	0.55
1:A:296:THR:HG22	1:A:335:TYR:CD2	2.48	0.49
1:B:280:GLU:O	1:B:296:THR:HG22	2.12	0.49
1:B:349:PHE:CE2	1:B:358:LEU:HD11	2.48	0.48
1:B:433:ALA:HB1	1:B:439:PHE:CE1	2.49	0.47
1:A:349:PHE:CE2	1:A:358:LEU:HD11	2.50	0.47
1:B:355:GLY:HA3	1:B:358:LEU:HD12	1.97	0.46
1:A:433:ALA:HB1	1:A:439:PHE:CE1	2.51	0.46
1:B:463:TYR:HA	3:B:723:HOH:O	2.16	0.46
1:A:532:ASN:O	1:A:533:LEU:HD23	2.17	0.45
1:A:355:GLY:HA3	1:A:358:LEU:HD12	1.98	0.44
1:B:471:VAL:O	1:B:475:VAL:HG23	2.19	0.43
1:A:471:VAL:O	1:A:475:VAL:HG23	2.19	0.42
1:A:310:GLU:HG2	1:A:406:GLY:HA3	2.04	0.40
1:A:461:VAL:HG23	1:B:518:ASP:HB3	2.02	0.40
1:A:515:PHE:CE2	1:A:523:THR:HG21	2.56	0.40

All (1) 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)
3:B:704:HOH:O	3:B:716:HOH:O[1_455]	2.11	0.09

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	260/286 (91%)	241 (93%)	18 (7%)	1 (0%)	34	54

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	259/286 (91%)	241 (93%)	18 (7%)	0	100	100
All	All	519/572 (91%)	482 (93%)	36 (7%)	1 (0%)	47	68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	407	LEU

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	225/245 (92%)	220 (98%)	5 (2%)	52	77
1	B	224/245 (91%)	222 (99%)	2 (1%)	78	92
All	All	449/490 (92%)	442 (98%)	7 (2%)	62	84

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	277	CYS
1	A	289	THR
1	A	290	THR
1	A	310	GLU
1	A	523	THR
1	B	289	THR
1	B	310	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	381	ASN
1	B	381	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

2 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	#$|Z| > 2$
2	73A	B	600	-	30,33,33	0.56	0	32,46,46	1.19	4 (12%)
2	73A	A	600	-	30,33,33	0.64	1 (3%)	32,46,46	0.88	1 (3%)

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	73A	B	600	-	-	7/20/38/38	0/3/3/3
2	73A	A	600	-	-	13/20/38/38	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	600	73A	C11-C12	-2.10	1.39	1.45

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	600	73A	C4-C11-N5	3.82	112.01	108.27
2	B	600	73A	C12-C13-C14	-3.77	118.04	123.57
2	B	600	73A	C4-C11-N5	3.44	111.64	108.27
2	B	600	73A	C12-C13-C15	2.73	123.32	119.19
2	B	600	73A	C4-C3-C2	-2.03	105.92	107.54

There are no chirality outliers.
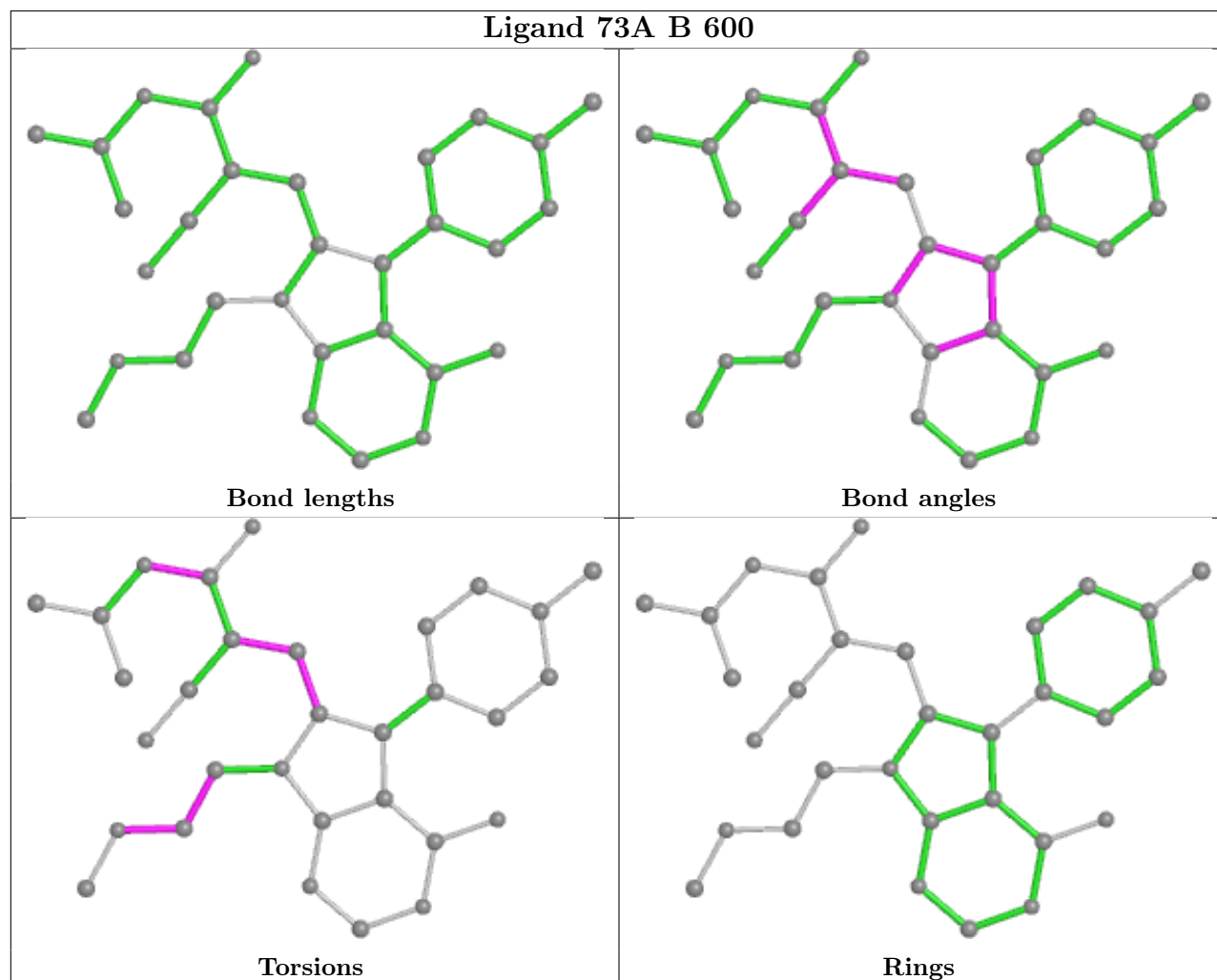
All (20) torsion outliers are listed below:

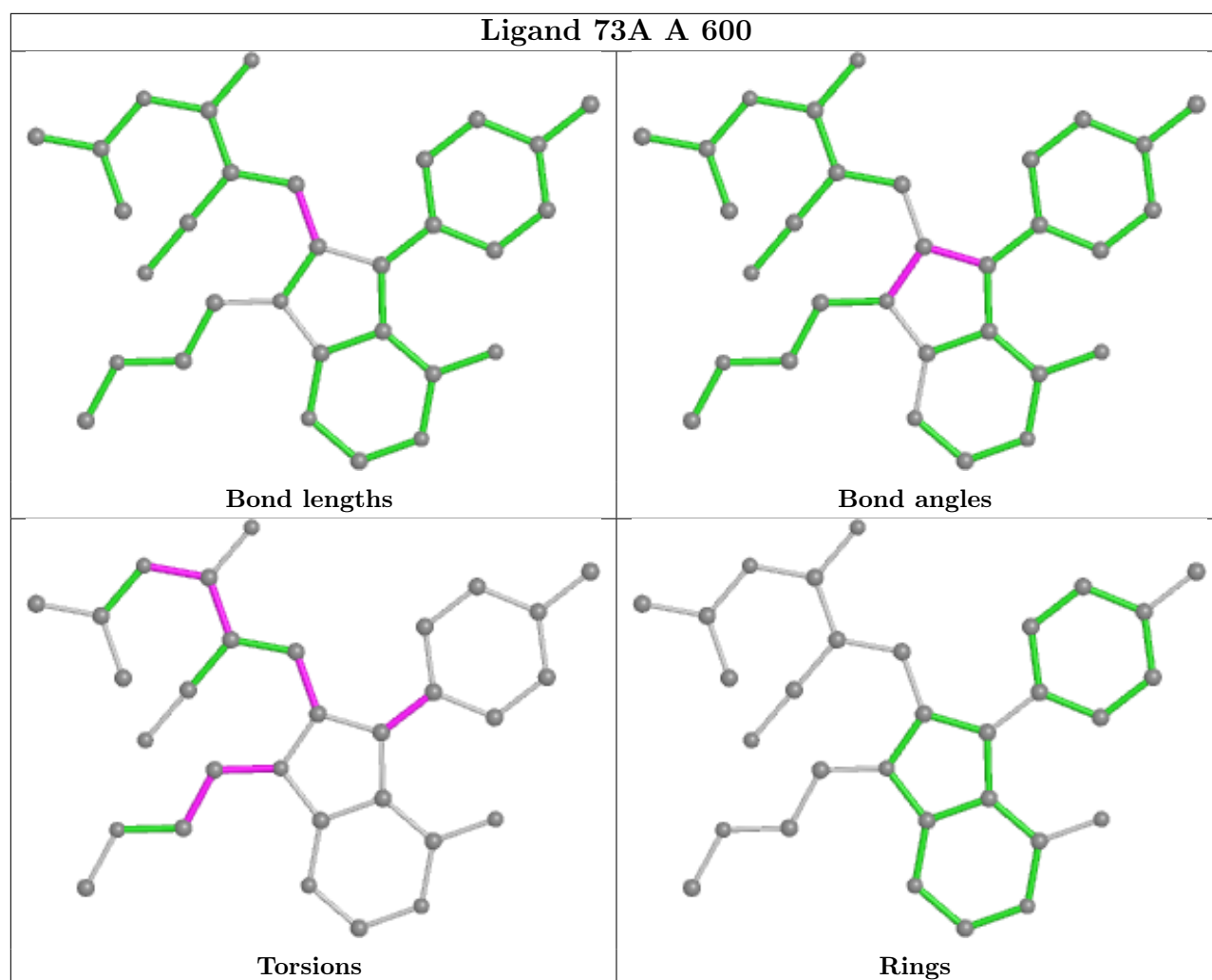| Mol	Chain	Res	Type	Atoms
2	A	600	73A	C3-C4-C5-C6
2	A	600	73A	C3-C4-C5-C10
2	A	600	73A	C13-C15-N4-C16
2	A	600	73A	O-C15-N4-C16
2	A	600	73A	C4-C11-C12-C13
2	A	600	73A	C14-C13-C15-N4
2	A	600	73A	C14-C13-C15-O
2	A	600	73A	C12-C13-C15-N4
2	A	600	73A	C12-C13-C15-O
2	B	600	73A	C13-C15-N4-C16
2	B	600	73A	O-C15-N4-C16
2	B	600	73A	C4-C11-C12-C13
2	A	600	73A	N5-C19-C20-O1
2	B	600	73A	N5-C19-C20-O1
2	B	600	73A	C11-C12-C13-C14
2	B	600	73A	C11-C12-C13-C15
2	A	600	73A	C20-C19-N5-C2
2	B	600	73A	C19-C20-O1-C21
2	A	600	73A	C11-C4-C5-C10
2	A	600	73A	C11-C4-C5-C6

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 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 73A B 600

Bond lengths

Bond angles

Torsions

Rings

Ligand 73A A 600

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	A	264/286 (92%)	0.89	44 (16%)	1	1	39, 74, 145, 184	0
1	B	263/286 (91%)	0.76	24 (9%)	9	9	40, 74, 132, 156	0
All	All	527/572 (92%)	0.82	68 (12%)	3	3	39, 74, 140, 184	0

All (68) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	335	TYR	6.8
1	A	297	LEU	6.5
1	B	262	ILE	5.9
1	B	307	PHE	5.9
1	A	332	GLU	5.3
1	B	407	LEU	5.2
1	B	405	PHE	5.0
1	A	334	ILE	4.9
1	A	267	LEU	4.5
1	A	276	GLY	4.5
1	B	334	ILE	4.5
1	A	280	GLU	4.5
1	A	308	LEU	4.5
1	A	307	PHE	4.4
1	B	533	LEU	4.3
1	B	317	LEU	4.2
1	B	301	THR	4.1
1	B	318	ARG	4.1
1	A	277	CYS	4.0
1	A	533	LEU	4.0
1	A	271	VAL	4.0
1	A	301	THR	3.7
1	A	300	GLY	3.5
1	A	299	PRO	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	486	GLU	3.5
1	A	317	LEU	3.4
1	A	330	SER	3.2
1	A	331	GLU	3.2
1	A	407	LEU	3.2
1	A	265	GLU	3.0
1	A	309	GLN	3.0
1	A	323	VAL	3.0
1	A	262	ILE	2.9
1	A	263	PRO	2.9
1	A	336	ILE	2.8
1	A	329	VAL	2.8
1	A	275	GLN	2.8
1	B	278	PHE	2.8
1	B	260	TRP	2.8
1	A	311	ALA	2.7
1	A	405	PHE	2.7
1	B	297	LEU	2.7
1	A	264	ARG	2.6
1	B	277	CYS	2.6
1	B	331	GLU	2.6
1	B	332	GLU	2.6
1	A	403	ALA	2.6
1	B	280	GLU	2.5
1	B	267	LEU	2.5
1	A	261	GLU	2.5
1	A	298	LYS	2.5
1	A	286	TRP	2.5
1	A	270	GLU	2.5
1	A	269	LEU	2.4
1	A	410	LEU	2.4
1	B	282	TRP	2.4
1	A	304	PRO	2.3
1	B	406	GLY	2.3
1	A	313	CYS	2.3
1	A	430	ALA	2.3
1	B	333	PRO	2.3
1	A	428	TRP	2.2
1	A	294	ILE	2.1
1	A	333	PRO	2.1
1	B	456	THR	2.1
1	B	312	GLN	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	471	VAL	2.0
1	B	455	LEU	2.0

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.
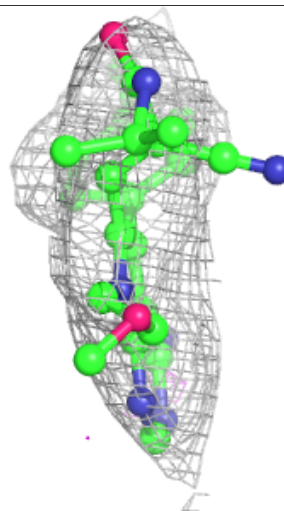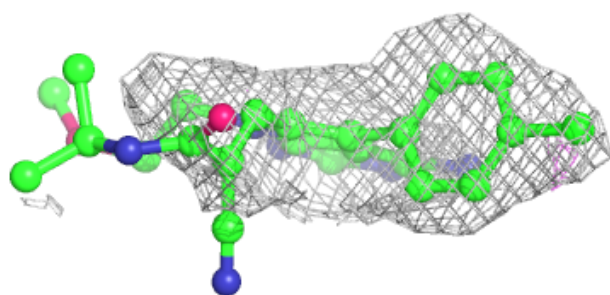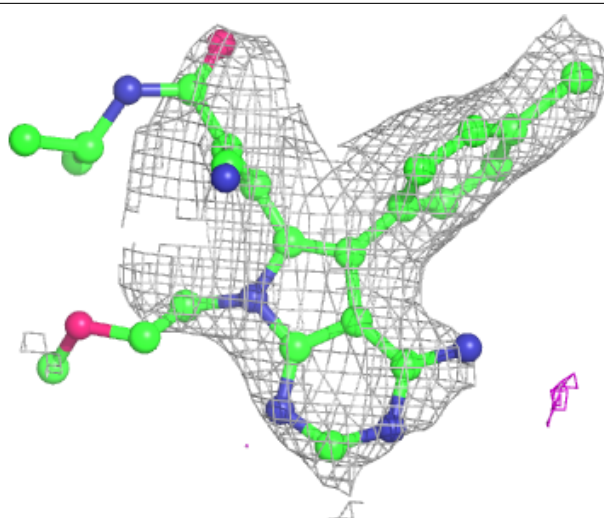
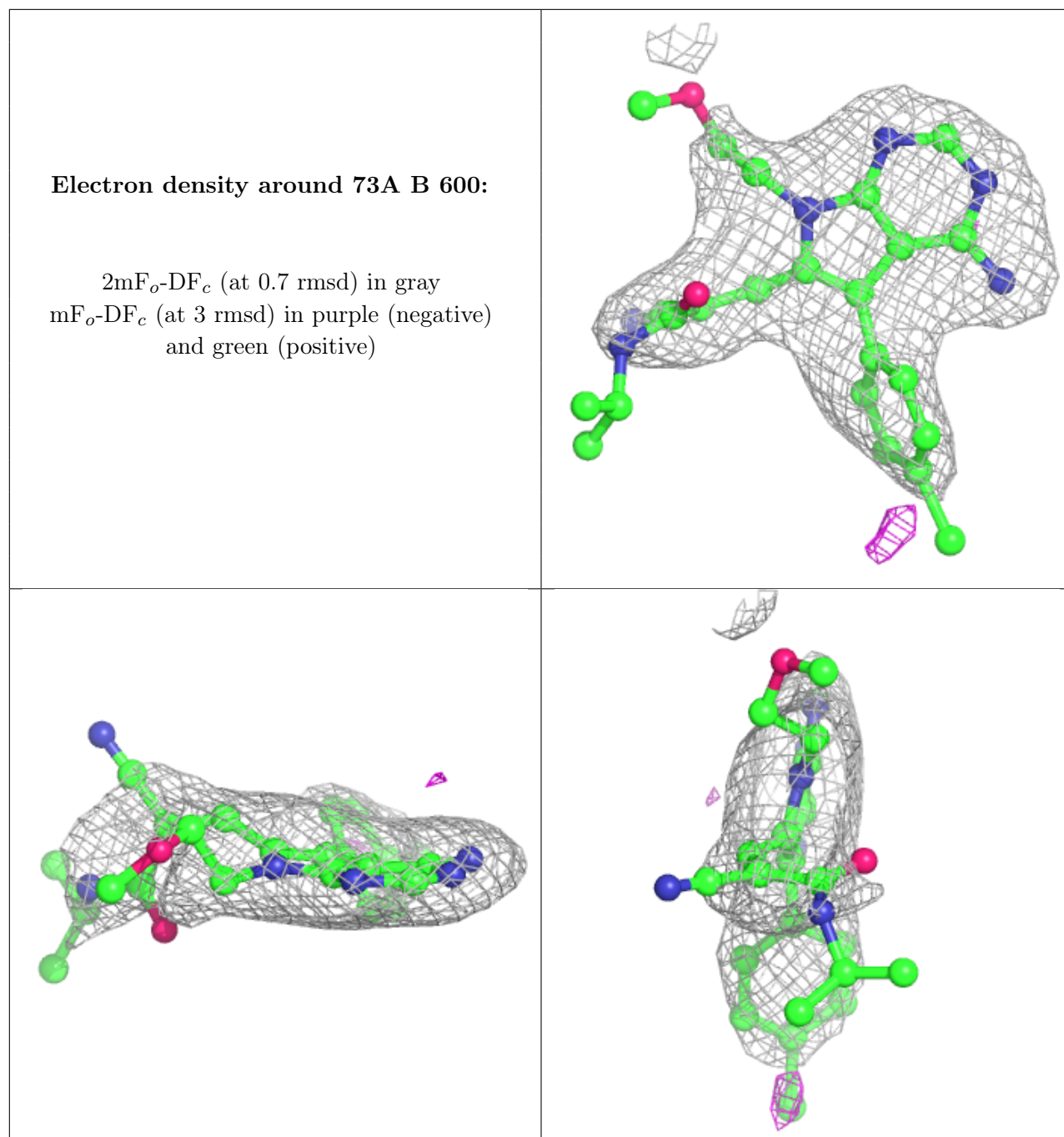| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	73A	A	600	31/31	0.84	0.33	90,111,162,163	0
2	73A	B	600	31/31	0.92	0.30	45,89,140,148	0

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 73A A 600:

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

Electron density around 73A B 600:

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