W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 29, 2020 – 06:58 am BST

PDB ID	:	4Q9R
Title	:	Crystal structure of an RNA aptamer bound to trifluoroethyl-ligand analog in complex with Fab
Authors	:	Huang, H.; Suslov, N.B.; Li, N.-S.; Shelke, S.A.; Evans, M.E.; Koldobskaya, Y.; Rice, P.A.; Piccirilli, J.A.
Deposited on	:	2014-05-01
Resolution	:	3.12 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 following versions of software and data (see references (i)) were used in the production of this report:

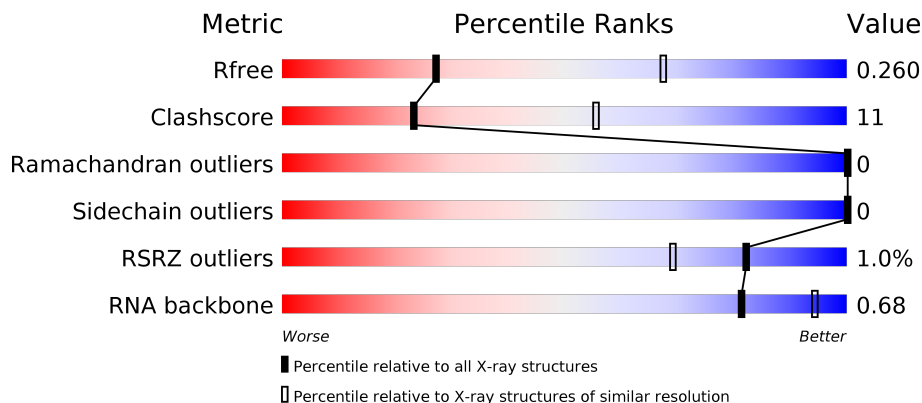|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.12 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		0.260
Clashscore		11
Ramachandran outliers		0
Sidechain outliers		0
RSRZ outliers		1.0%
RNA backbone		0.68

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	1292 (3.14-3.10)
Clashscore	141614	1389 (3.14-3.10)
Ramachandran outliers	138981	1337 (3.14-3.10)
Sidechain outliers	138945	1337 (3.14-3.10)
RSRZ outliers	127900	1260 (3.14-3.10)
RNA backbone	3102	1134 (3.44-2.80)

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 on 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	H	223	2% 76% 24%
2	L	214	76% 24%
3	R	84	% 62% 31% 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:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	2ZY	R	102	-	X	-	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 5174 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 Fab BL3-6, HEAVY CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	223	Total	C	N	O	S	0	0	0
			1666	1049	285	327	5			

- Molecule 2 is a protein called Fab BL3-6, LIGHT CHAIN.

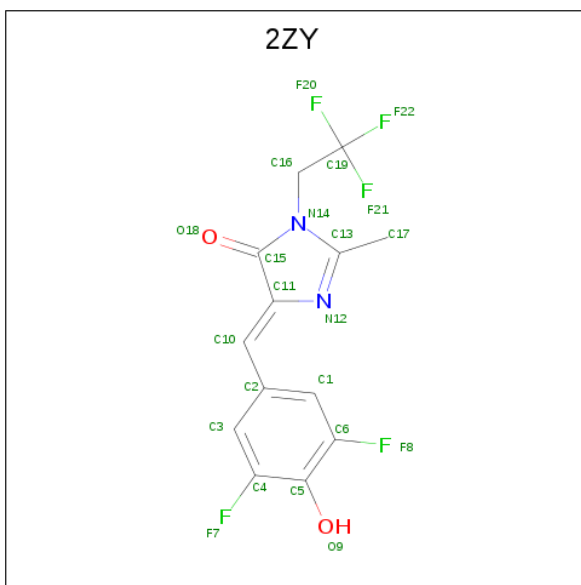| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	214	Total	C	N	O	S	0	0	0
			1637	1022	274	336	5			

- Molecule 3 is a RNA chain called Spinach RNA aptamer.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	R	84	Total	C	N	O	P	0	0	0
			1815	806	337	587	85			

- Molecule 4 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	R	1	Total	K	0	0
			1	1		

- Molecule 5 is (5Z)-5-(3,5-difluoro-4-hydroxybenzylidene)-2-methyl-3-(2,2,2-trifluoroethyl)-3, 5-dihydro-4H-imidazol-4-one (three-letter code: 2ZY) (formula: $C_{13}H_9F_5N_2O_2$).

| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	R	1	Total 22	C 13	F 5	N 2	O 2	0	0

• Molecule 6 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	H	11	Total 11	O 11	0	0
6	L	14	Total 14	O 14	0	0
6	R	8	Total 8	O 8	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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.
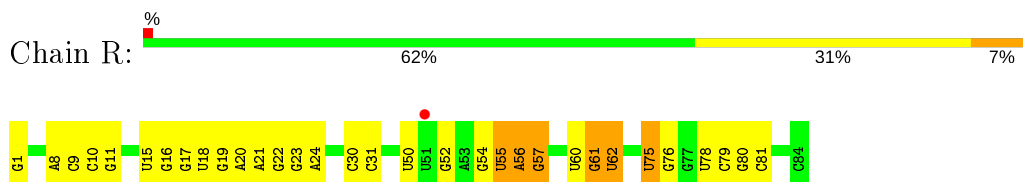
- Molecule 1: Fab BL3-6, HEAVY CHAIN
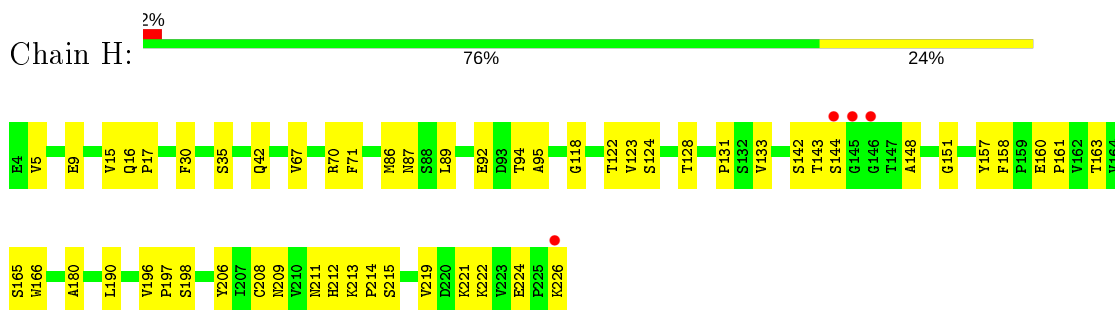
- Molecule 2: Fab BL3-6, LIGHT CHAIN
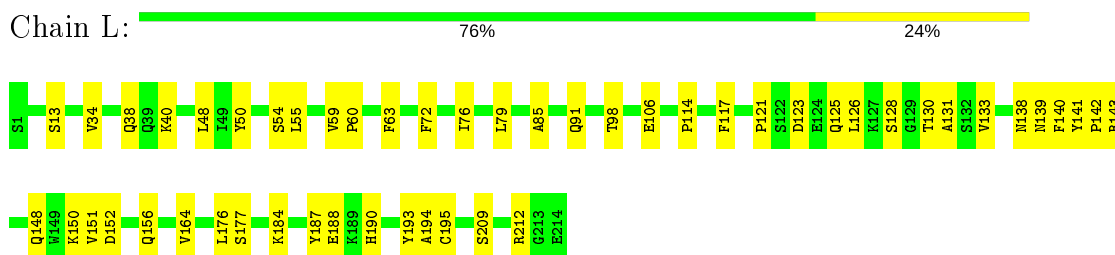
- Molecule 3: Spinach RNA aptamer

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	147.51Å 78.99Å 94.36Å 90.00° 112.42° 90.00°	Depositor
Resolution (Å)	67.68 – 3.12 87.22 – 3.12	Depositor EDS
% Data completeness (in resolution range)	98.1 (67.68-3.12) 87.0 (87.22-3.12)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.45 (at 3.13Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1678)	Depositor
R, R_{free}	0.203 , 0.253 0.223 , 0.260	Depositor DCC
R_{free} test set	882 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å^2)	55.9	Xtriage
Anisotropy	0.516	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 60.4	EDS
L-test for twinning[2]	$<|L|> = 0.47$, $<L^2> = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	5174	wwPDB-VP
Average B, all atoms (Å^2)	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.31% 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: GDP, K, 2ZY

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	H	0.26	0/1707	0.41	0/2323
2	L	0.24	0/1672	0.42	0/2269
3	R	0.17	0/2001	0.67	0/3121
All	All	0.23	0/5380	0.53	0/7713

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 (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	H	1666	0	1628	43	0
2	L	1637	0	1590	35	0
3	R	1815	0	910	27	0
4	R	1	0	0	0	0
5	R	22	0	8	0	0
6	H	11	0	0	1	0
6	L	14	0	0	0	0
6	R	8	0	0	0	0
All	All	5174	0	4136	102	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 11.

The worst 5 of 102 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:R:56:A:H3'	3:R:57:G:H4'	1.36	1.05
1:H:131:PRO:HB3	1:H:157:TYR:HB3	1.53	0.89
2:L:194:ALA:HA	2:L:209:SER:OG	1.75	0.86
1:H:86:MET:HB3	1:H:89:LEU:HD21	1.63	0.81
3:R:22:G:H1'	3:R:23:G:C8	2.16	0.81

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	Percentiles	
1	H	221/223 (99%)	210 (95%)	11 (5%)	0	100	100
2	L	212/214 (99%)	203 (96%)	9 (4%)	0	100	100
All	All	433/437 (99%)	413 (95%)	20 (5%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

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	H	184/184 (100%)	184 (100%)	0	100	100
2	L	189/189 (100%)	189 (100%)	0	100	100
All	All	373/373 (100%)	373 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	R	83/84 (98%)	7 (8%)	1 (1%)

5 of 7 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	R	55	U
3	R	56	A
3	R	57	G
3	R	60	U
3	R	61	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
3	R	1	GDP

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

1 non-standard protein/DNA/RNA residue is 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	GDP	R	1	3	24,30,30	1.20	2 (8%)	31,47,47	1.97	8 (25%)

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	GDP	R	1	3	-	3/12/32/32	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	R	1	GDP	C6-C5	4.26	1.48	1.41
3	R	1	GDP	C5-C4	2.50	1.47	1.40

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	R	1	GDP	C2-N3-C4	4.91	120.97	115.36
3	R	1	GDP	C5-C6-N1	-4.08	117.84	123.43
3	R	1	GDP	C6-N1-C2	4.01	122.30	115.93
3	R	1	GDP	C6-C5-C4	-3.50	117.45	120.80
3	R	1	GDP	N3-C2-N1	-3.17	122.99	127.22

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	R	1	GDP	C3'-C4'-C5'-O5'
3	R	1	GDP	O4'-C4'-C5'-O5'
3	R	1	GDP	C5'-O5'-PA-O3A

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	2ZY	R	102	-	23,23,23	5.64	11 (47%)	34,35,35	8.12	20 (58%)

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
5	2ZY	R	102	-	-	5/9/25/25	0/2/2/2

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	R	102	2ZY	C10-C11	19.29	1.51	1.35
5	R	102	2ZY	C5-C6	8.80	1.49	1.39
5	R	102	2ZY	C13-N12	8.20	1.43	1.32
5	R	102	2ZY	C16-C19	-7.92	1.39	1.50
5	R	102	2ZY	C11-C15	6.89	1.55	1.48

The worst 5 of 20 bond angle outliers are listed below:

| Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	R	102	2ZY	C11-C15-N14	24.75	115.08	103.37
5	R	102	2ZY	C15-N14-C13	-15.07	100.34	107.97
5	R	102	2ZY	C16-N14-C13	13.78	139.66	127.35
5	R	102	2ZY	C10-C11-N12	13.57	147.64	128.83
5	R	102	2ZY	C2-C10-C11	11.95	144.59	129.94

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	R	102	2ZY	C19-C16-N14-C15
5	R	102	2ZY	C19-C16-N14-C13
5	R	102	2ZY	N14-C16-C19-F20
5	R	102	2ZY	N14-C16-C19-F21
5	R	102	2ZY	N14-C16-C19-F22

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 2ZY R 102

Bond lengths

Bond angles

Torsions

Rings
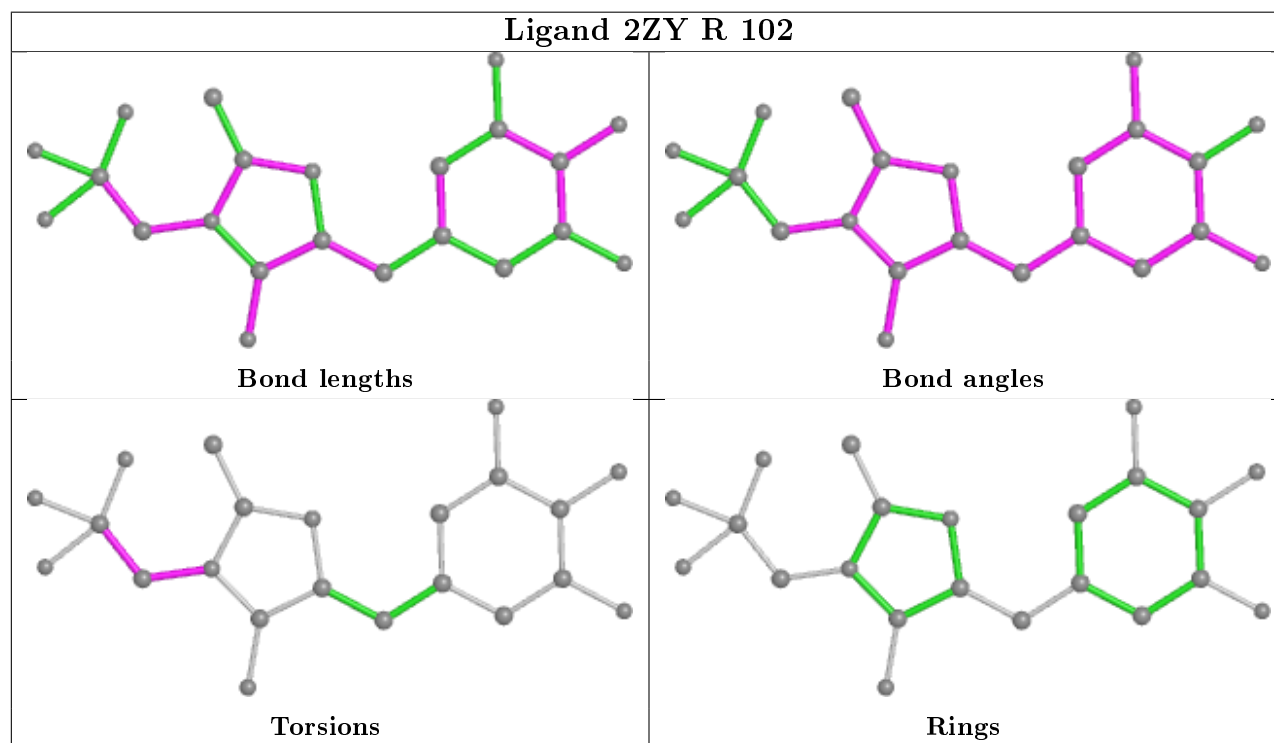
5.7 Other polymers ⓘ

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, 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(Å²)	Q<0.9
1	H	223/223 (100%)	0.00	4 (1%)	68	48	41, 67, 110, 165	0
2	L	214/214 (100%)	-0.04	0	100	100	39, 64, 91, 124	0
3	R	83/84 (98%)	-0.07	1 (1%)	79	63	56, 118, 171, 232	0
All	All	520/521 (99%)	-0.02	5 (0%)	82	69	39, 68, 146, 232	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	144	SER	6.2
1	H	145	GLY	5.1
3	R	51	U	2.5
1	H	226	LYS	2.4
1	H	146	GLY	2.1

6.2 Non-standard residues in protein, DNA, RNA chains (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.

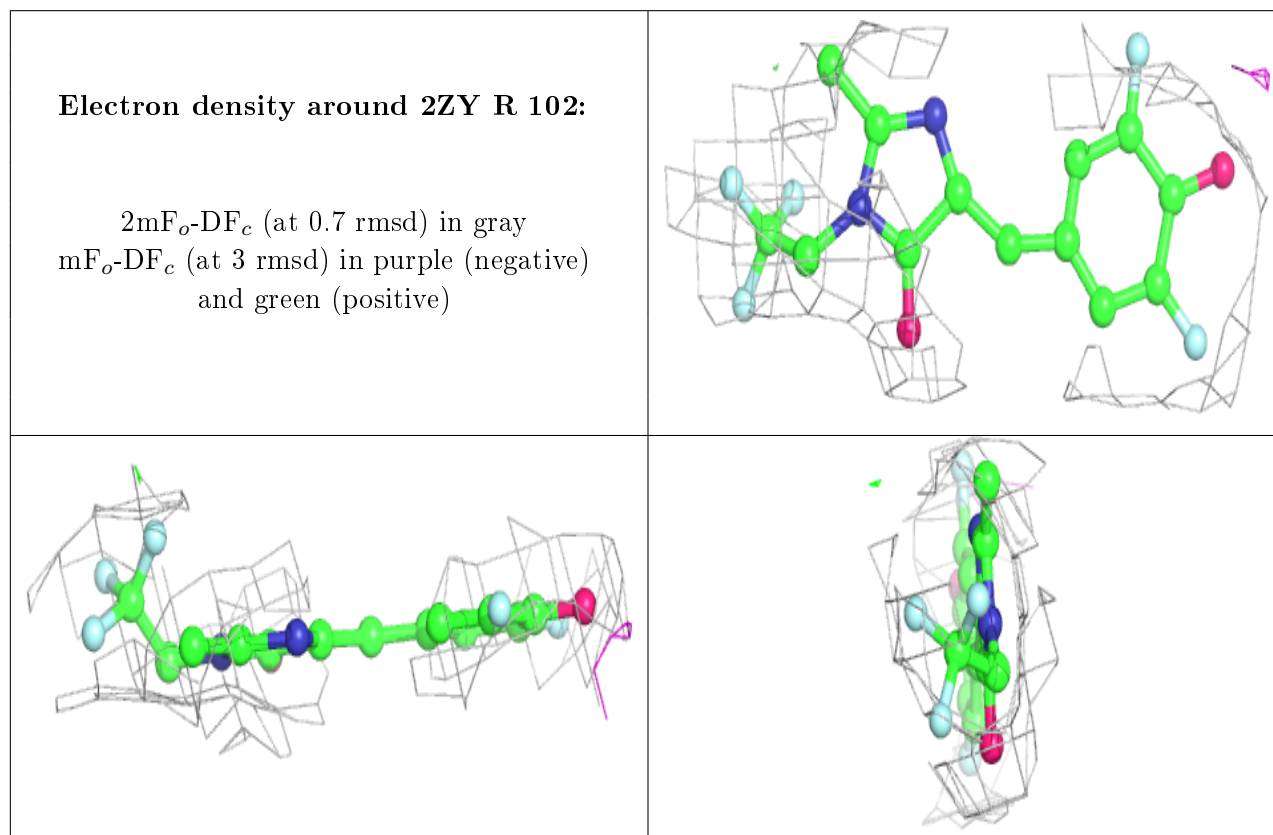| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å²)	Q<0.9
3	GDP	R	1	28/28	0.77	0.21	146,164,180,219	0

6.3 Carbohydrates (i)

There are no carbohydrates 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
4	K	R	101	1/1	-0.16	0.13	151,151,151,151	0
5	2ZY	R	102	22/22	0.77	0.16	129,157,170,177	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 2ZY R 102:

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

6.5 Other polymers (i)

There are no such residues in this entry.