



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 07:29 pm BST

PDB ID : 3I0R
Title : crystal structure of HIV reverse transcriptase in complex with inhibitor 3
Authors : Yan, Y.; Prasad, S.
Deposited on : 2009-06-25
Resolution : 2.98 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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

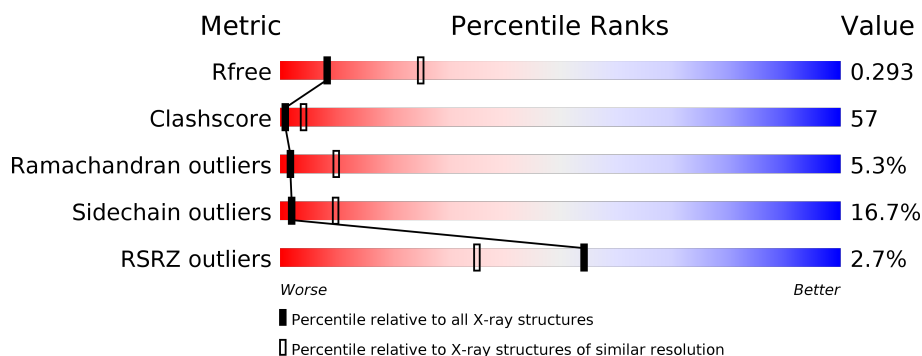
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.98 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2754 (3.00-2.96)
Clashscore	141614	3103 (3.00-2.96)
Ramachandran outliers	138981	2993 (3.00-2.96)
Sidechain outliers	138945	2996 (3.00-2.96)
RSRZ outliers	127900	2644 (3.00-2.96)

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 $\leq 5\%$. 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	563	<div> <div>2%</div> <div> <div></div> <div>33%</div> <div>49%</div> <div>13%</div> <div>• •</div> </div> </div>
2	B	443	<div> <div>4%</div> <div> <div></div> <div>38%</div> <div>39%</div> <div>14%</div> <div>• 9%</div> </div> </div>

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
3	RT3	A	601	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7823 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 Reverse transcriptase/ribonuclease H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	545	Total	C	N	O	S	0	0	0
			4442	2874	741	819	8			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MET	-	EXPRESSION TAG	UNP P04585
A	-1	ASN	-	EXPRESSION TAG	UNP P04585
A	0	SER	-	EXPRESSION TAG	UNP P04585

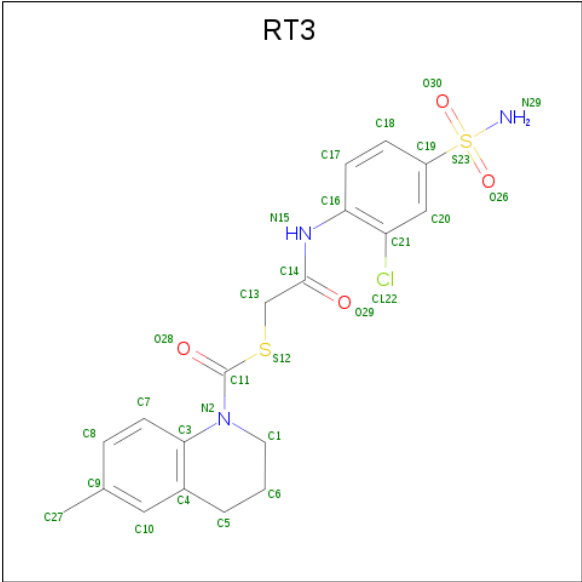
- Molecule 2 is a protein called p51 RT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	405	Total	C	N	O	S	0	0	0
			3352	2182	555	609	6			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	MET	-	EXPRESSION TAG	UNP P04585
B	-1	ASN	-	EXPRESSION TAG	UNP P04585
B	0	SER	-	EXPRESSION TAG	UNP P04585

- Molecule 3 is S-{2-[(2-chloro-4-sulfamoylphenyl)amino]-2-oxoethyl} 6-methyl-3,4-dihydroquinoline-1(2H)-carbothioate (three-letter code: RT3) (formula: C₁₉H₂₀ClN₃O₄S₂).

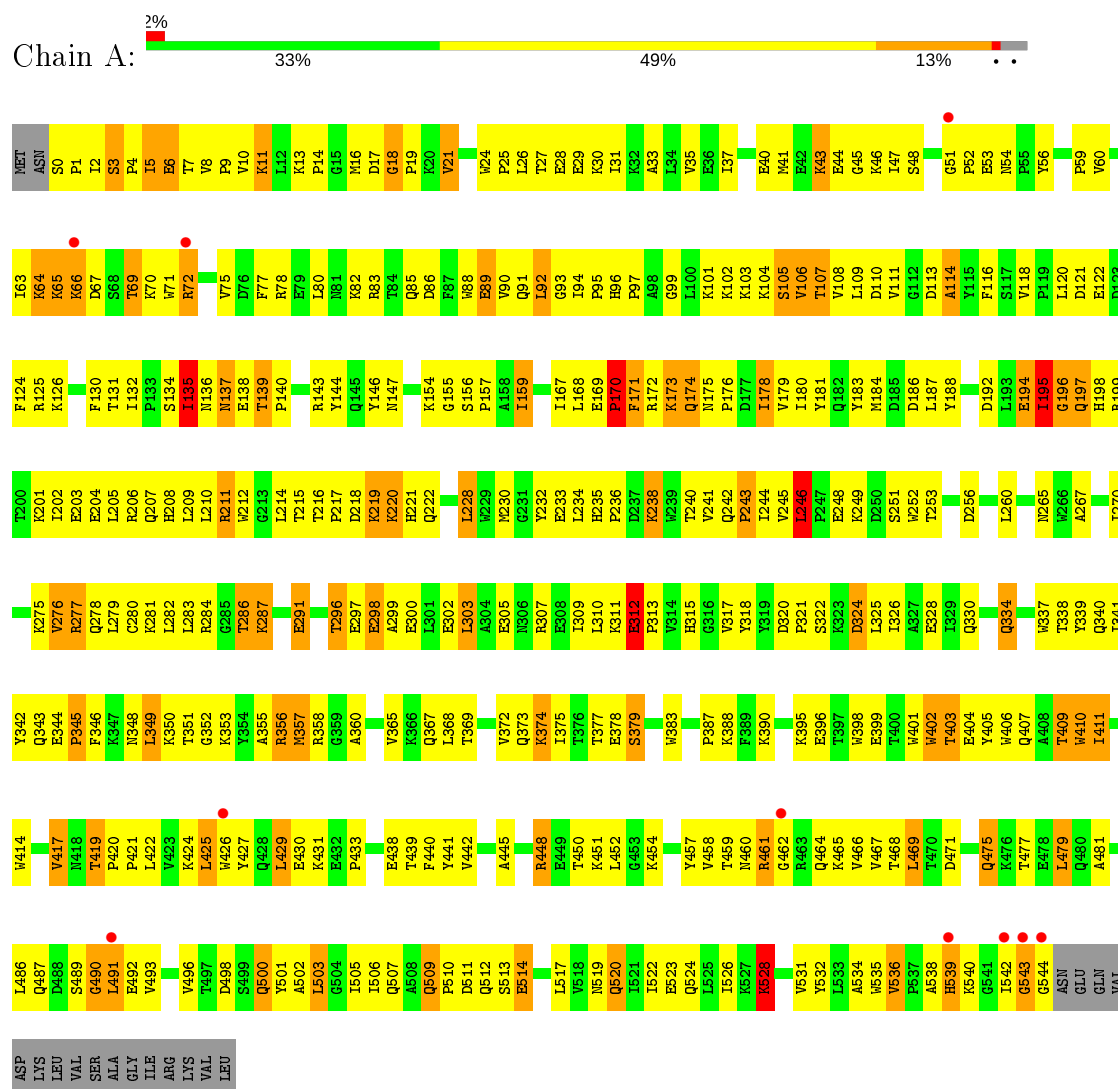


Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	C	Cl	N	O	S	0	0
			29	19	1	3	4	2		

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: Reverse transcriptase/ribonuclease H



• Molecule 2: p51 RT



K374	I375	T376	T377	E378	S379	I382	K385	T386	K387	K388	F389	K390	L391	E396	E399	T400	T403	E404	Y405	W406	Q407	W410	I411	P412	E413	W414	E415	N418	T419	P420	P421	L422	V423	K424	L425	W426	Y427	Q428	LEU	GLU	LYS	GLU	PRO	ILE	VAL	GLY	ALA	GLU	THR	PHE								
K249	D250	S251	W252	T253	V254	N255	D256	I257	Q258	K259	L260	V261	G262	K263	L264	N265	W266	L267	S268	Q269	G270	I271	P272	G273	I274	K275	V276	R277	Q278	L279	C280	K281	L282	L283	R284	G285	T286	L289	T290	E291	V292	L293	P294	L295	T296	E297	E298	L299	E300	L301	E302	L303	A304	E305	N306	R307	E308	I309
L310	K311	E312	P313	V314	V317	Y318	Y319	D320	P321	S322	K323	D324	L325	E328	I329	Q330	K331	Q332	G333	Q334	G335	Q336	W337	T338	Y339	Q340	I341	Y342	Q343	E344	P345	F346	K347	N348	L349	K350	T351	G352	K353	Y354	A355	R356	MET	ARG	GLY	ALA	H361	T362	N363	K366	Q367	L368	T369	V372	Q373			
V179	I180	Y181	Q182	Y183	M184	L187	Y188	V189	G190	I195	G196	Q197	H198	R199	T200	K201	L202	E203	E204	L205	L209	I214	T215	THR	PRO	ASP	L219	L120	D121	K126	N136	M137	G143	W153	K154	G155	S156	P157	A158	I159	F160	Q161	S162	T165	K166	I167	L168	E169	P170	K173	Q174	N175	I178	E179				
L80	R81	K82	R83	T84	Q85	D86	F87	W88	E89	V90	D91	L92	K102	K103	R104	S105	V108	V111	G112	D113	A114	P119	L120	D121	K126	N136	M137	G143	W153	K154	G155	S156	P157	A158	I159	F160	Q161	S162	T165	K166	I167	L168	E169	P170	K173	Q174	N175	I178	E179									
MET	ASN	SER	PRO	ILE	SER	PRO	I5	E6	T7	V8	K11	L12	K13	D17	G18	K22	Q23	W24	P25	L26	E29	K30	I37	M41	K46	I47	S48	K49	I50	O51	P52	E53	N54	V60	I63	K64	K65	K66	D67	S68	T69	R72	K73	L74	V75	D76	F77	R78	E79									

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	119.33Å 155.31Å 154.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.98 47.49 – 2.98	Depositor EDS
% Data completeness (in resolution range)	99.3 (50.00-2.98) 99.3 (47.49-2.98)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.81 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.5.0066	Depositor
R, R_{free}	0.220 , 0.295 0.222 , 0.293	Depositor DCC
R_{free} test set	1494 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	76.2	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 44.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	7823	wwPDB-VP
Average B, all atoms (Å ²)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.22% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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.

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: RT3

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	0.60	0/4559	0.72	0/6196
2	B	0.63	1/3446 (0.0%)	0.72	0/4682
All	All	0.61	1/8005 (0.0%)	0.72	0/10878

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	A	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	24	TRP	CB-CG	-5.02	1.41	1.50

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	539	HIS	Peptide

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	A	4442	0	4489	550	0
2	B	3352	0	3380	383	0
3	A	29	0	20	17	0
All	All	7823	0	7889	899	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 57.

The worst 5 of 899 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)
2:B:244:ILE:CD1	2:B:266:TRP:CZ3	2.03	1.41
2:B:244:ILE:CD1	2:B:266:TRP:CH2	2.21	1.23
1:A:1:PRO:O	1:A:2:ILE:HD13	1.36	1.21
1:A:206:ARG:NH2	1:A:218:ASP:HB3	1.57	1.19
2:B:244:ILE:HD13	2:B:266:TRP:CZ3	1.69	1.18

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	A	543/563 (96%)	442 (81%)	72 (13%)	29 (5%)	2	10
2	B	399/443 (90%)	331 (83%)	47 (12%)	21 (5%)	2	10
All	All	942/1006 (94%)	773 (82%)	119 (13%)	50 (5%)	2	10

5 of 50 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	65	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	90	VAL
1	A	195	ILE
1	A	345	PRO
1	A	356	ARG

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	486/503 (97%)	405 (83%)	81 (17%)	2	10
2	B	369/403 (92%)	307 (83%)	62 (17%)	2	10
All	All	855/906 (94%)	712 (83%)	143 (17%)	2	10

5 of 143 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	429	LEU
1	A	536	VAL
2	B	349	LEU
1	A	448	ARG
1	A	479	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	407	GLN
1	A	519	ASN
2	B	373	GLN
1	A	480	GLN
1	A	524	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

1 ligand 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	RT3	A	601	-	29,31,31	2.39	6 (20%)	39,45,45	1.79	11 (28%)

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	RT3	A	601	-	-	12/18/29/29	0/3/3/3

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	601	RT3	C19-S23	-10.28	1.61	1.77
3	A	601	RT3	C3-C4	4.58	1.46	1.40
3	A	601	RT3	C16-C21	4.02	1.49	1.39
3	A	601	RT3	C16-N15	-2.64	1.36	1.41
3	A	601	RT3	C13-S12	-2.34	1.77	1.80

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	601	RT3	O26-S23-O30	-5.95	108.97	118.76
3	A	601	RT3	C4-C3-N2	3.20	121.83	117.73
3	A	601	RT3	O26-S23-N29	2.70	111.37	107.36
3	A	601	RT3	O26-S23-C19	2.69	110.35	107.35
3	A	601	RT3	C13-C14-N15	2.53	117.87	114.41

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

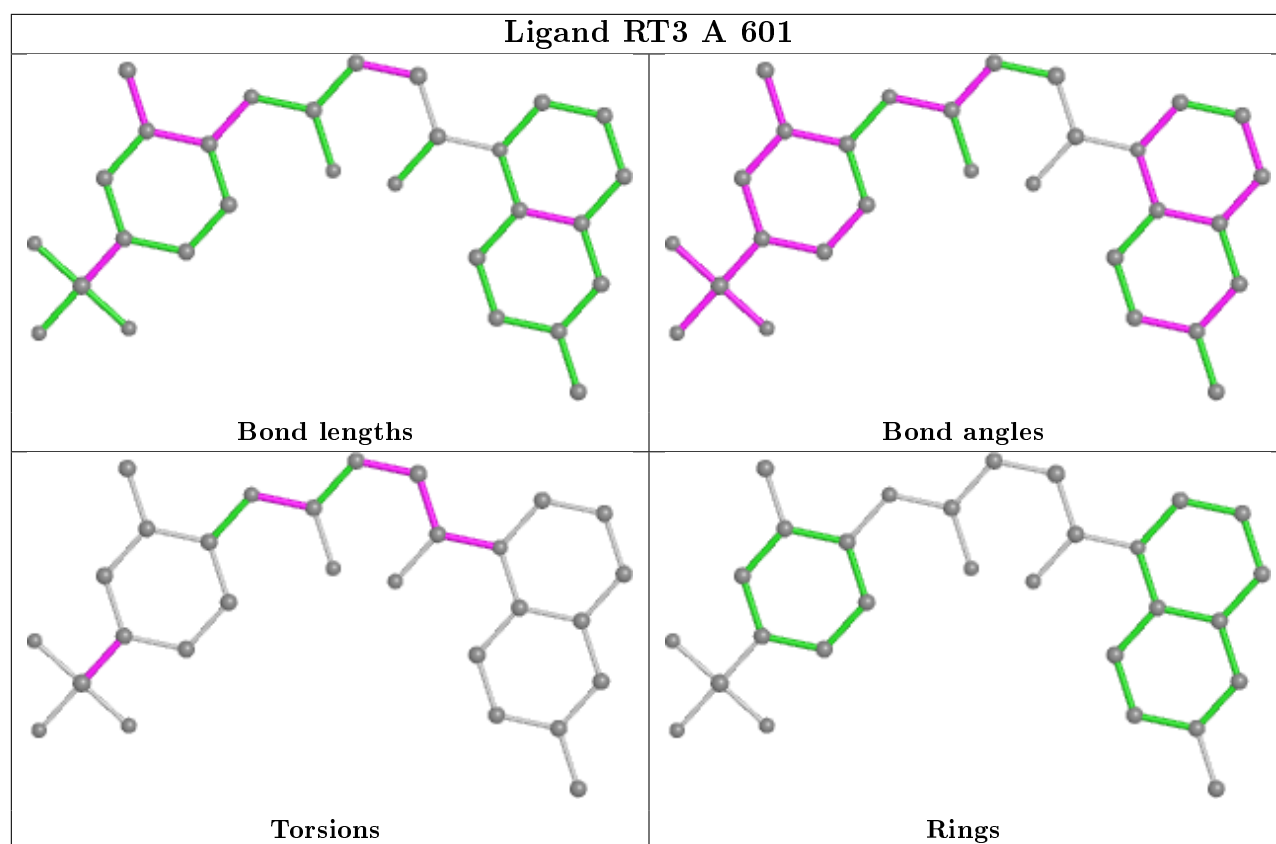
Mol	Chain	Res	Type	Atoms
3	A	601	RT3	O29-C14-N15-C16
3	A	601	RT3	C13-C14-N15-C16
3	A	601	RT3	O28-C11-S12-C13
3	A	601	RT3	N2-C11-S12-C13
3	A	601	RT3	S12-C11-N2-C1

There are no ring outliers.

1 monomer is involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	601	RT3	17	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 than 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, 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	545/563 (96%)	0.15	10 (1%) 68 48	44, 71, 91, 113	0
2	B	405/443 (91%)	0.22	16 (3%) 38 23	45, 64, 108, 114	0
All	All	950/1006 (94%)	0.18	26 (2%) 54 35	44, 69, 104, 114	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	66	LYS	4.5
2	B	252	TRP	4.5
1	A	542	ILE	4.2
2	B	270	ILE	4.2
1	A	426	TRP	3.7

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](#)

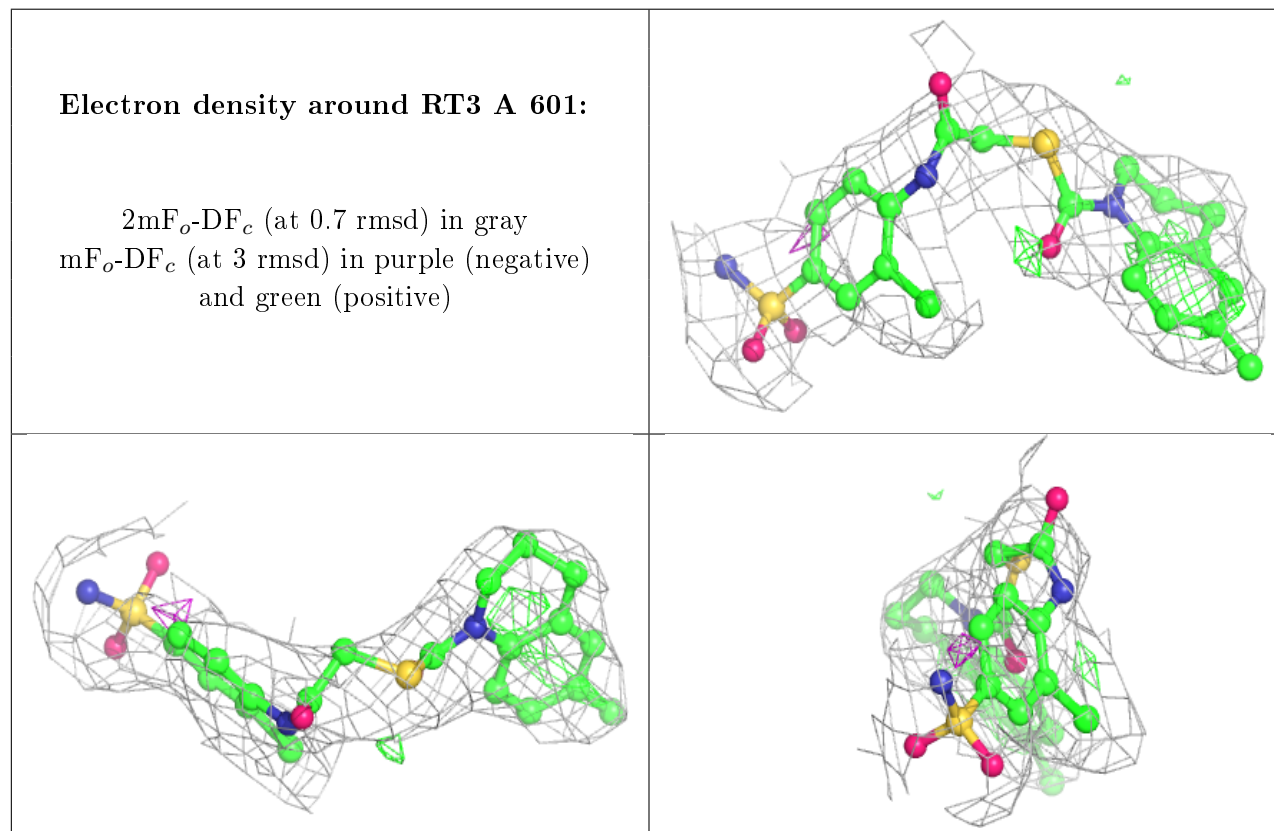
There are no carbohydrates in this entry.

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(\AA^2)	Q<0.9
3	RT3	A	601	29/29	0.86	0.26	86,100,107,107	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.



6.5 Other polymers ⓘ

There are no such residues in this entry.