wwPDB X-ray Structure Validation Summary Report (i)

Aug 23, 2023 – 12:34 PM EDT

PDB ID	:	3C6T
Title	:	Crystal Structure of HIV Reverse Transcriptase in complex with inhibitor 14
Authors	:	Yan, Y.; Prasad, S.
Deposited on	:	2008-02-05
Resolution	:	2.70 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 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:

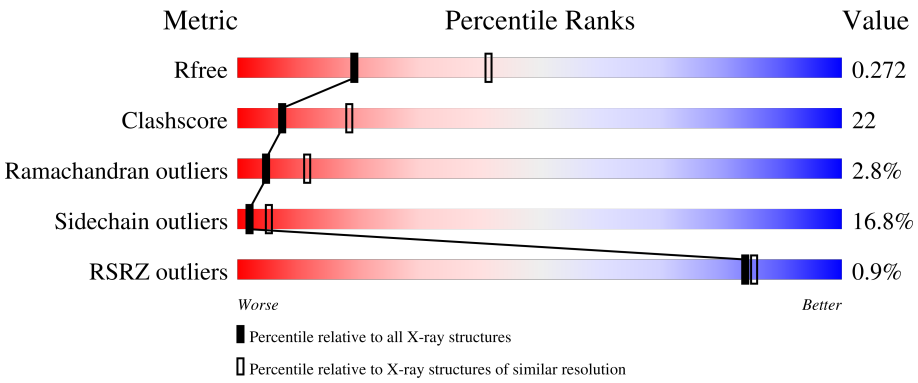
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 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 2.70 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	A	563	% 53% 36% 10% ••
2	B	443	% 47% 37% 7% 9%

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	M14	A	561	-	-	X	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 8293 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	559	4550	2938	762	842	8	0	0	0

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 Reverse transcriptase.

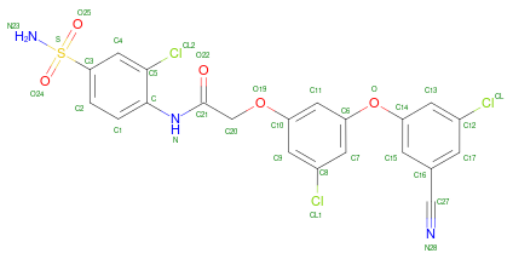
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	404	3344	2176	554	608	6	0	0	0

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 2-[3-chloro-5-(3-chloro-5-cyanophenoxy)phenoxy]-N-(2-chloro-4-sulfamoylphenyl)acetamide (three-letter code: M14) (formula: $C_{21}H_{14}Cl_3N_3O_5S$).

M14



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	Cl	N	O	S		
3	A	1	33	21	3	3	5	1	0	0

- Molecule 4 is water.

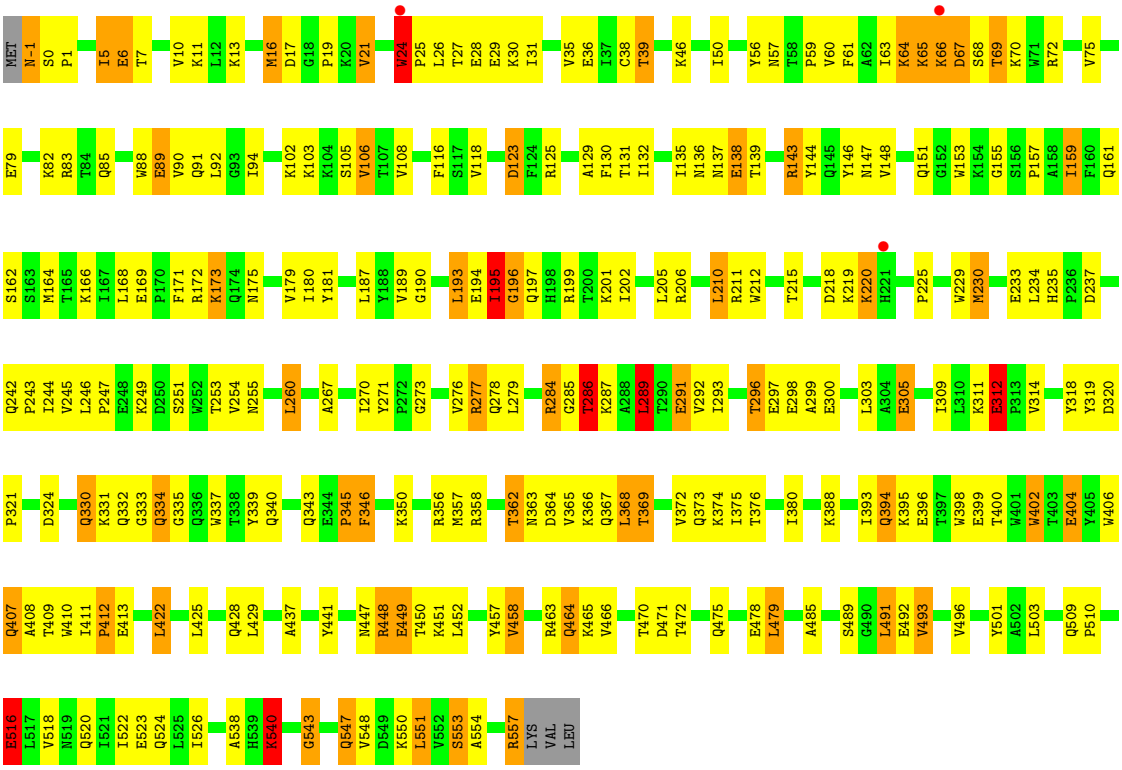
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	209	209	209	0	0
			Total	O		
4	B	157	157	157	0	0

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.

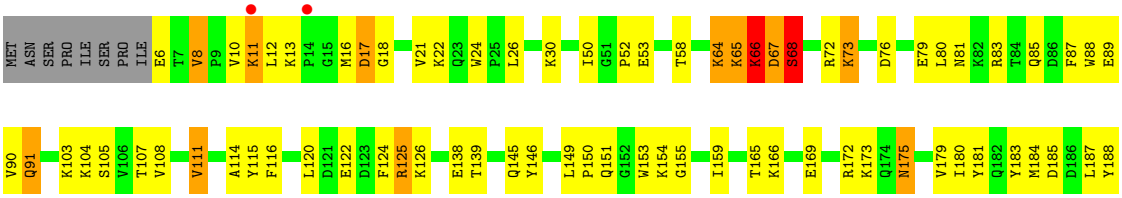
- Molecule 1: Reverse transcriptase

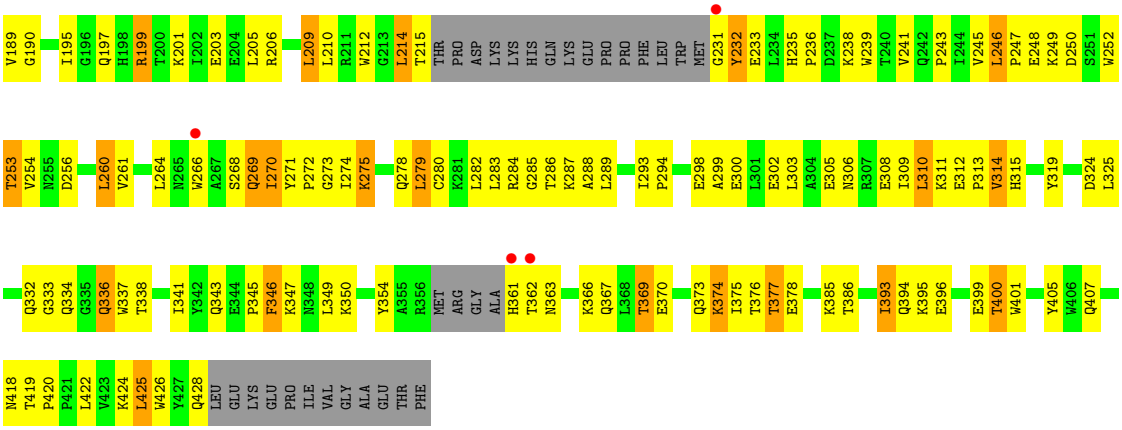
Chain A:



- Molecule 2: Reverse transcriptase

Chain B:




V189 G190 I195 G196 Q197 H198 R199 T200 K201 I202 E203 E204 L205 R206 L209 L210 R211 W212 G213 L214 T215 THR PRO ASP LYS LYS HIS GLN LYS GLU PRO PRO PHE LEU TRP MET G231 Y232 E233 L234 H235 P236 D237 K238 W239 T240 V241 Q242 P243 I244 V245 L246 P247 E248 K249 D250 S251 W252
T253 V254 N255 D256 L260 V261 L264 N265 W266 A267 S268 Q269 I270 Y271 P272 G273 I274 K275 Q278 L279 C280 K281 L282 L283 R284 G285 T286 K287 A288 L289 I293 P294 E298 A299 E300 L301 E302 L303 A304 E305 N306 R307 E308 I309 L310 K311 E312 P313 V314 H315 Y319 D324 L325
Q332 G333 Q334 G335 Q336 W337 T338 I341 Y342 Q343 E344 P345 F346 K347 N348 L349 K350 Y354 A355 R356 MET ARG GLY ALA H361 T362 N363 K366 Q367 L368 T369 E370 Q373 K374 I375 T376 T377 E378 K385 T386 I393 Q394 K395 E396 E399 T400 W401 Y405 W406 Q407
N418 T419 P420 P421 L422 Y423 K424 L425 W426 Y427 Q428 LEU GLU LYS GLU PRO ILE VAL GLY ALA GLU THR PHE

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	118.29Å 154.51Å 155.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.22 – 2.70 47.22 – 2.70	Depositor EDS
% Data completeness (in resolution range)	93.3 (47.22-2.70) 93.4 (47.22-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	2.06 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.202 , 0.281 0.199 , 0.272	Depositor DCC
R_{free} test set	1857 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	53.9	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 61.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	8293	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.00% 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: M14

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	0.92	2/4667 (0.0%)	0.91	5/6341 (0.1%)
2	B	0.90	2/3438 (0.1%)	0.89	0/4671
All	All	0.91	4/8105 (0.0%)	0.90	5/11012 (0.0%)

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 (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	24	TRP	CB-CG	5.80	1.60	1.50
2	B	24	TRP	CB-CG	-5.72	1.40	1.50
1	A	516	GLU	CG-CD	5.22	1.59	1.51
2	B	181	TYR	CD1-CE1	-5.17	1.31	1.39

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	493	VAL	CB-CA-C	-7.12	97.88	111.40
1	A	540	LYS	N-CA-C	-7.04	92.00	111.00
1	A	289	LEU	CA-CB-CG	5.93	128.94	115.30
1	A	193	LEU	CA-CB-CG	5.72	128.45	115.30
1	A	479	LEU	CA-CB-CG	5.71	128.44	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	-1	ASN	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	4550	0	4601	219	0
2	B	3344	0	3369	147	0
3	A	33	0	14	10	0
4	A	209	0	0	21	0
4	B	157	0	0	11	0
All	All	8293	0	7984	354	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 22.

The worst 5 of 354 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:175:ASN:HD21	1:A:201:LYS:CE	1.54	1.20
1:A:175:ASN:ND2	1:A:201:LYS:CE	2.16	1.08
1:A:175:ASN:HD21	1:A:201:LYS:HE3	1.17	1.07
1:A:175:ASN:ND2	1:A:201:LYS:HE2	1.68	1.07
1:A:237:ASP:HB3	4:A:756:HOH:O	1.59	0.99

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	557/563 (99%)	492 (88%)	48 (9%)	17 (3%)	4 9
2	B	398/443 (90%)	353 (89%)	35 (9%)	10 (2%)	5 14
All	All	955/1006 (95%)	845 (88%)	83 (9%)	27 (3%)	5 11

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1	PRO
1	A	67	ASP
1	A	90	VAL
1	A	286	THR
1	A	412	PRO

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	A	499/503 (99%)	405 (81%)	94 (19%)	1 4
2	B	368/403 (91%)	316 (86%)	52 (14%)	3 8
All	All	867/906 (96%)	721 (83%)	146 (17%)	2 5

5 of 146 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	203	GLU
2	B	394	GLN
2	B	214	LEU
2	B	280	CYS
1	A	293	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	373	GLN
2	B	394	GLN
1	A	474	ASN
1	A	428	GLN
2	B	418	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)

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	M14	A	561	-	35,35,35	1.38	3 (8%)	50,50,50	2.27	16 (32%)

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	M14	A	561	-	-	5/21/21/21	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	561	M14	S-N23	5.46	1.71	1.60
3	A	561	M14	C16-C27	-3.20	1.37	1.44
3	A	561	M14	C9-C8	2.57	1.42	1.38

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	561	M14	O24-S-C3	-7.63	98.84	107.35
3	A	561	M14	C5-C-N	-4.58	111.99	119.24
3	A	561	M14	C7-C8-CL1	-4.22	113.88	119.15
3	A	561	M14	C16-C27-N28	-3.95	167.42	177.85
3	A	561	M14	C9-C8-CL1	3.39	123.39	119.15

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	561	M14	O19-C20-C21-O22
3	A	561	M14	C2-C3-S-O24
3	A	561	M14	C4-C3-S-O24
3	A	561	M14	C4-C3-S-N23
3	A	561	M14	C2-C3-S-N23

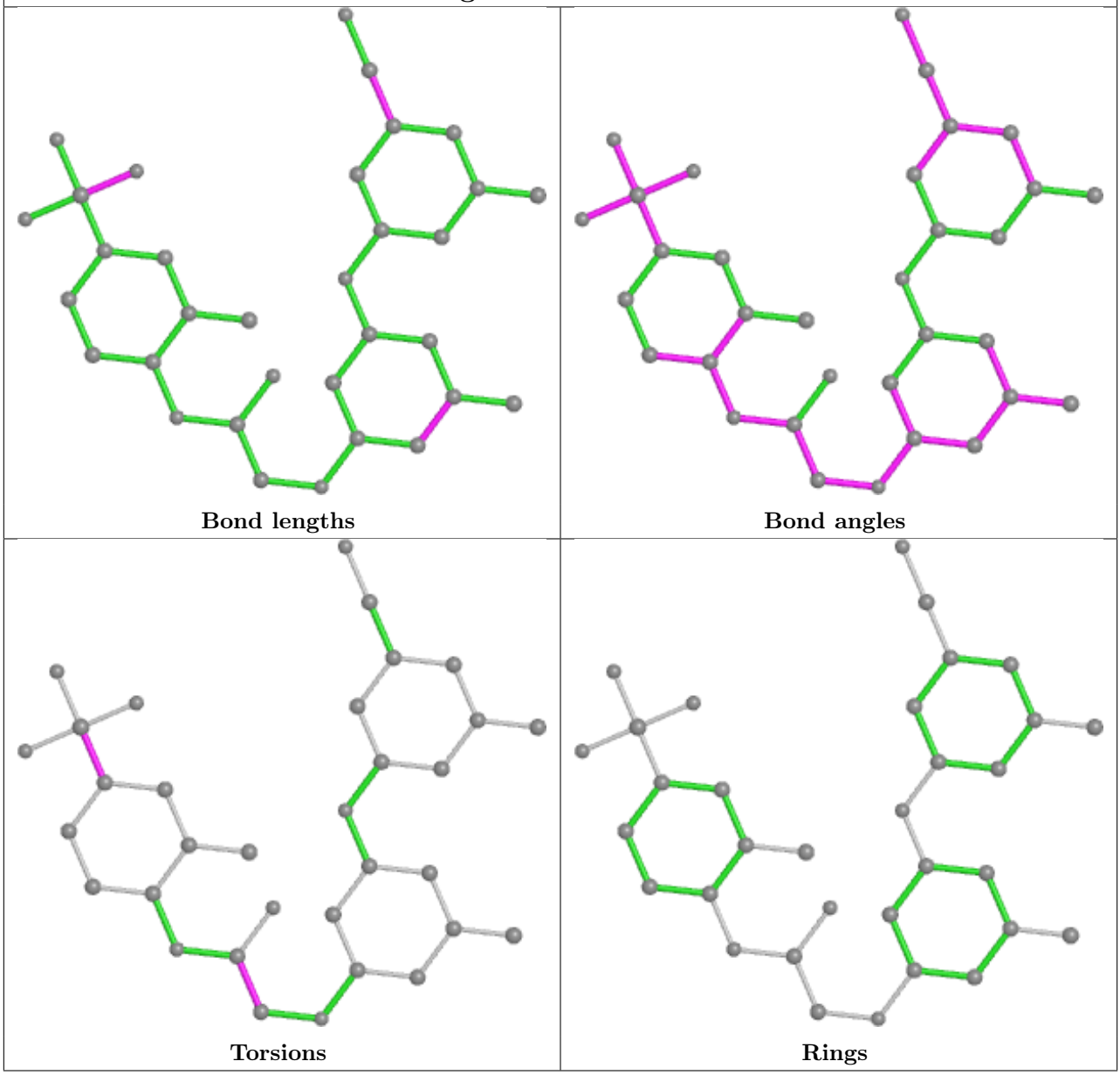
There are no ring outliers.

1 monomer is involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	561	M14	10	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 M14 A 561



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	559/563 (99%)	-0.19	3 (0%)	91	92	24, 46, 68, 91	0
2	B	404/443 (91%)	-0.14	6 (1%)	73	76	24, 44, 82, 96	0
All	All	963/1006 (95%)	-0.17	9 (0%)	84	85	24, 46, 78, 96	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	11	LYS	3.1
1	A	66	LYS	2.9
2	B	14	PRO	2.7
1	A	24	TRP	2.7
1	A	221	HIS	2.6

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)

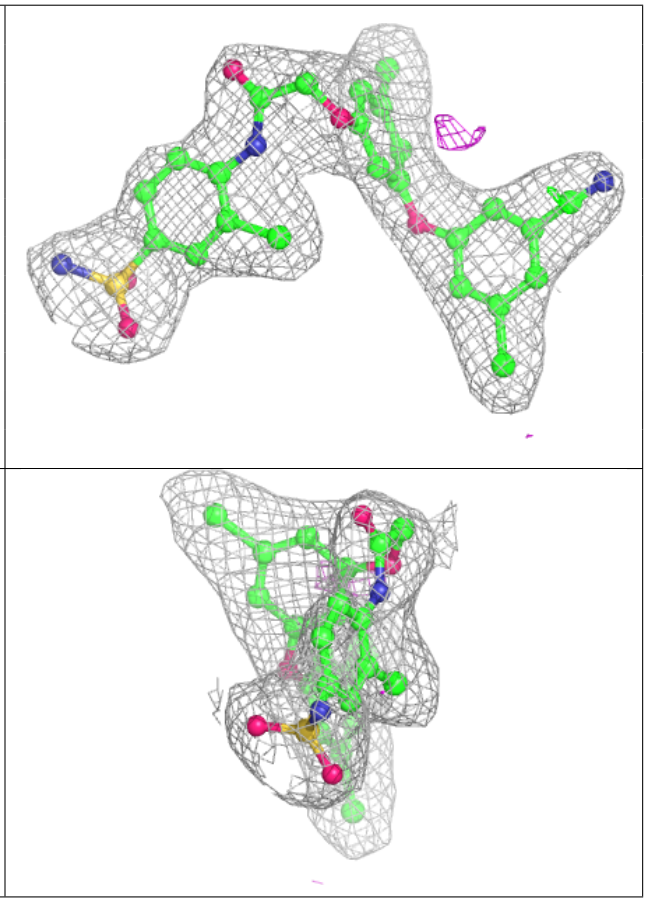
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	M14	A	561	33/33	0.94	0.16	38,48,54,58	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 M14 A 561:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.