W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 23, 2020 – 03:48 am BST

PDB ID	:	1HNV
Title	:	STRUCTURE OF HIV-1 RT(SLASH)TIBO R 86183 COMPLEX REVEALS SIMILARITY IN THE BINDING OF DIVERSE NONNUCLEOSIDE IN-HIBITORS
Authors	:	Das, K.; Ding, J.; Arnold, E.
Deposited on	:	1995-03-30
Resolution	:	3.00 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

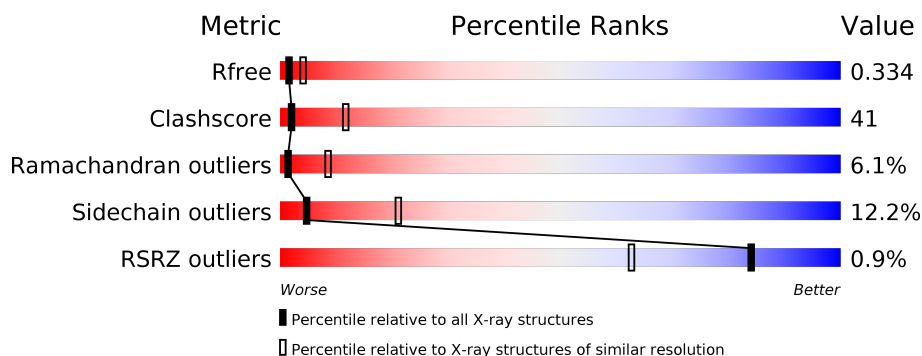|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 3.00 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		0.334
Clashscore		41
Ramachandran outliers		6.1%
Sidechain outliers		12.2%
RSRZ outliers		0.9%

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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	A	558	38% 53% 8%
2	B	427	35% 49% 14%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7833 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 HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P66).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	558	Total 4370	C 2827	N 725	O 812	S 6	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	280	SER	CYS	CONFLICT	UNP P03366

- Molecule 2 is a protein called HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P51).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	427	Total 3442	C 2240	N 567	O 630	S 5	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	280	SER	CYS	CONFLICT	UNP P03366

- Molecule 3 is 5-CHLORO-8-METHYL-7-(3-METHYL-BUT-2-ENYL)-6,7,8,9-TETRAH YDRO-2H-2,7,9A-TRIAZA-BENZO[CD]AZULENE-1-THIONE (three-letter code: TBO) (formula: $C_{16}H_{20}ClN_3S$).
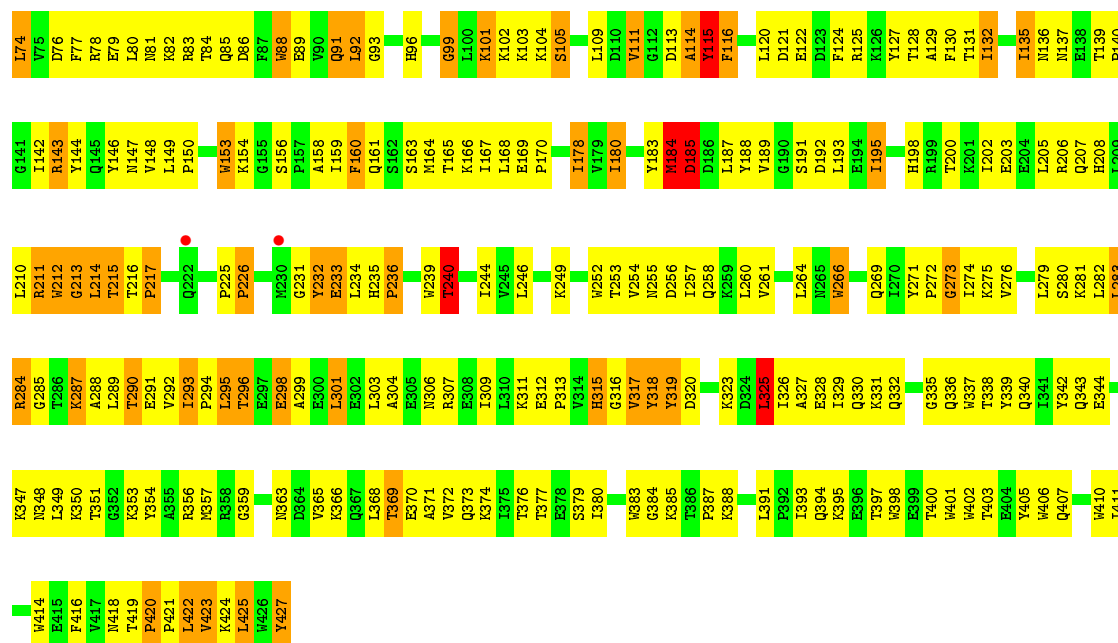
TBO

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	Cl	N	S	0	0
			21	16	1	3	1		

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: HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P66)

- Molecule 2: HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P51)

L74 V75 D76 F77 R78 E79 L80 N81 K82 R83 T84 Q85 D86 F87 W88 E89 V90 Q91 L92 G93 H96 G99 L100 K101 K102 K103 K104 S105 L109 D110 V111 G112 D113 A114 Y115 F116 L120 D121 E122 D123 F124 R125 K126 Y127 T128 A129 F130 T131 I132 I135 N136 N137 E138 T139 P140

G141 I142 R143 Y144 Q145 Y146 N147 V148 L149 P150 W153 K154 G155 S156 P157 A158 I159 F160 Q161 S162 S163 M164 T165 K166 I167 L168 E169 P170 I178 V179 I180 Y183 M184 D185 D186 L187 Y188 V189 G190 S191 D192 L193 E194 I195 H198 R199 T200 K201 I202 E203 E204 L205 R206 Q207 H208 L209

L210 R211 W212 G213 L214 T215 T216 P217 Q222 P225 P226 M230 G231 Y232 E233 L234 H235 P236 W239 T240 I244 V245 L246 K249 W252 T253 V254 N255 D256 I257 Q258 K259 L260 V261 L264 N265 W266 Q269 I270 Y271 P272 G273 I274 K275 V276 L279 S280 K281 L282 L283

R284 G285 T286 K287 A288 L289 T290 E291 V292 I293 P294 L295 T296 E297 E298 A299 E300 L301 E302 L303 A304 E305 N306 R307 E308 I309 L310 K311 E312 P313 V314 H315 G316 V317 Y318 Y319 D320 K323 D324 L325 I326 A327 E328 I329 Q330 K331 Q332 G335 Q336 W337 T338 Y339 Q340 I341 Y342 Q343 E344

K347 N348 L349 K350 T351 G352 K353 Y354 A355 R356 M357 R358 G359 N363 D364 V365 K366 Q367 L368 T369 E370 A371 V372 Q373 K374 I375 T376 T377 E378 S379 I380 W383 G384 K385 T386 P387 K388 L391 P392 I393 Q394 K395 E396 T397 W398 E399 T400 W401 W402 T403 E404 Y405 W406 Q407 W410 I411

W414 E415 F416 V417 N418 T419 P420 P421 L422 V423 K424 L425 W426 Y427

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	227.20Å 70.20Å 105.70Å 90.00° 105.60° 90.00°	Depositor
Resolution (Å)	10.00 – 3.00 14.97 – 2.80	Depositor EDS
% Data completeness (in resolution range)	83.2 (10.00-3.00) 81.1 (14.97-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.75 (at 2.81Å)	Xtriage
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.249 , 0.356 0.248 , 0.334	Depositor DCC
R_{free} test set	1594 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	55.2	Xtriage
Anisotropy	0.023	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.19 , 62.6	EDS
L-test for twinning[2]	$< |L| > = 0.46, < L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	7833	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.50% 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: TBO

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	0.64	0/4484	0.85	4/6113 (0.1%)
2	B	0.72	1/3541 (0.0%)	0.91	7/4822 (0.1%)
All	All	0.68	1/8025 (0.0%)	0.88	11/10935 (0.1%)

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
2	B	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	153	TRP	CB-CG	-6.62	1.38	1.50

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	134	SER	N-CA-C	6.52	128.60	111.00
2	B	184	MET	CG-SD-CE	6.40	110.45	100.20
2	B	215	THR	N-CA-C	-5.80	95.35	111.00
1	A	288	ALA	N-CA-C	-5.74	95.50	111.00
2	B	225	PRO	N-CA-CB	5.67	110.11	103.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	56	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 the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4370	0	4265	368	0
2	B	3442	0	3405	298	0
3	A	21	0	20	7	0
All	All	7833	0	7690	632	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 41.

The worst 5 of 632 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:92:LEU:HD11	2:B:22:LYS:NZ	1.25	1.47
1:A:13:LYS:HB2	1:A:16:MET:HG3	1.29	1.10
1:A:92:LEU:CD1	2:B:22:LYS:NZ	2.16	1.08
2:B:139:THR:HG22	2:B:140:PRO:HD2	1.36	1.08
1:A:92:LEU:CD1	2:B:22:LYS:HZ3	1.68	1.07

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	554/558 (99%)	426 (77%)	104 (19%)	24 (4%)	2	15
2	B	425/427 (100%)	322 (76%)	67 (16%)	36 (8%)	1	4
All	All	979/985 (99%)	748 (76%)	171 (18%)	60 (6%)	1	8

5 of 60 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	52	PRO
1	A	90	VAL
1	A	345	PRO
1	A	358	ARG
1	A	420	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	454/498 (91%)	398 (88%)	56 (12%)	4	21
2	B	367/389 (94%)	323 (88%)	44 (12%)	5	22
All	All	821/887 (93%)	721 (88%)	100 (12%)	5	21

5 of 100 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	491	LEU
1	A	553	SER
2	B	369	THR
1	A	497	THR
1	A	536	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	474	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	524	GLN
2	B	330	GLN
1	A	373	GLN
1	A	407	GLN

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 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	TBO	A	559	-	19,23,23	4.42	9 (47%)	16,34,34	4.67	11 (68%)

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	TBO	A	559	-	-	1/4/17/17	0/2/3/3

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	559	TBO	C7A-C3A	-10.87	1.40	1.54
3	A	559	TBO	C9-C8	-7.98	1.42	1.52
3	A	559	TBO	C2-N1	6.75	1.43	1.34
3	A	559	TBO	C4-N3	-6.20	1.42	1.46
3	A	559	TBO	C3A-N3	-6.16	1.39	1.47

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	559	TBO	C9-C8-CL8	10.55	120.64	109.51
3	A	559	TBO	C1A-C3A-N3	8.17	113.44	102.58
3	A	559	TBO	C16-C14-C15	7.48	131.13	114.60
3	A	559	TBO	C15-C14-C13	-5.32	107.28	122.65
3	A	559	TBO	C10-C9-C8	4.25	119.74	112.22

There are no chirality outliers.

All (1) torsion outliers are listed below:

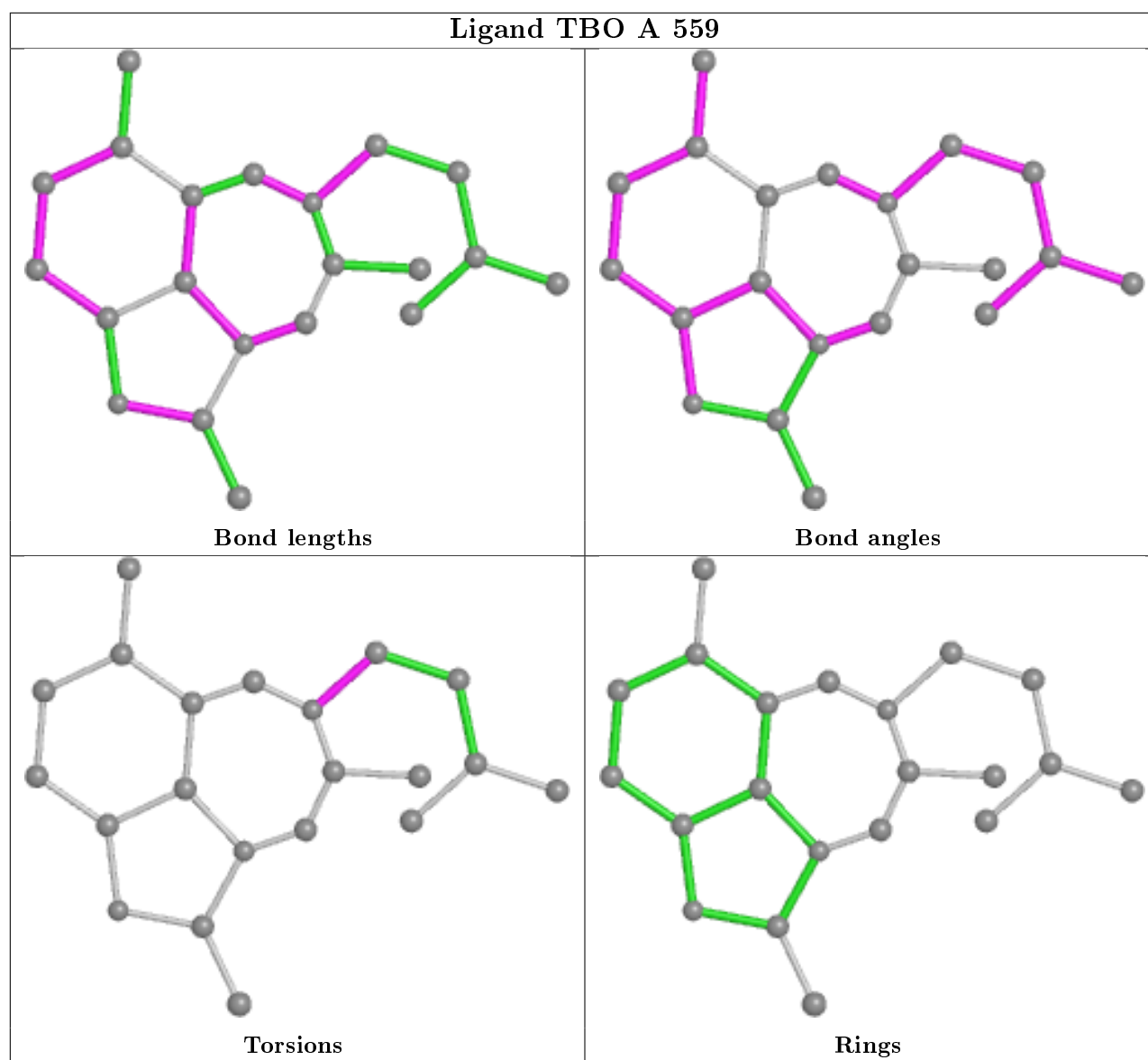| Mol	Chain	Res	Type	Atoms
3	A	559	TBO	C13-C12-N6-C7

There are no ring outliers.

1 monomer is involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	559	TBO	7	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 TBO A 559

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 (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	558/558 (100%)	-0.56	4 (0%)	87 69	10, 43, 60, 67	0
2	B	427/427 (100%)	-0.59	5 (1%)	79 54	10, 35, 61, 69	0
All	All	985/985 (100%)	-0.57	9 (0%)	84 63	10, 41, 60, 69	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	66	LYS	3.4
2	B	3	SER	3.4
2	B	222	GLN	3.1
2	B	2	ILE	3.0
1	A	141	GLY	2.9

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

There are no non-standard protein/DNA/RNA residues in this entry.
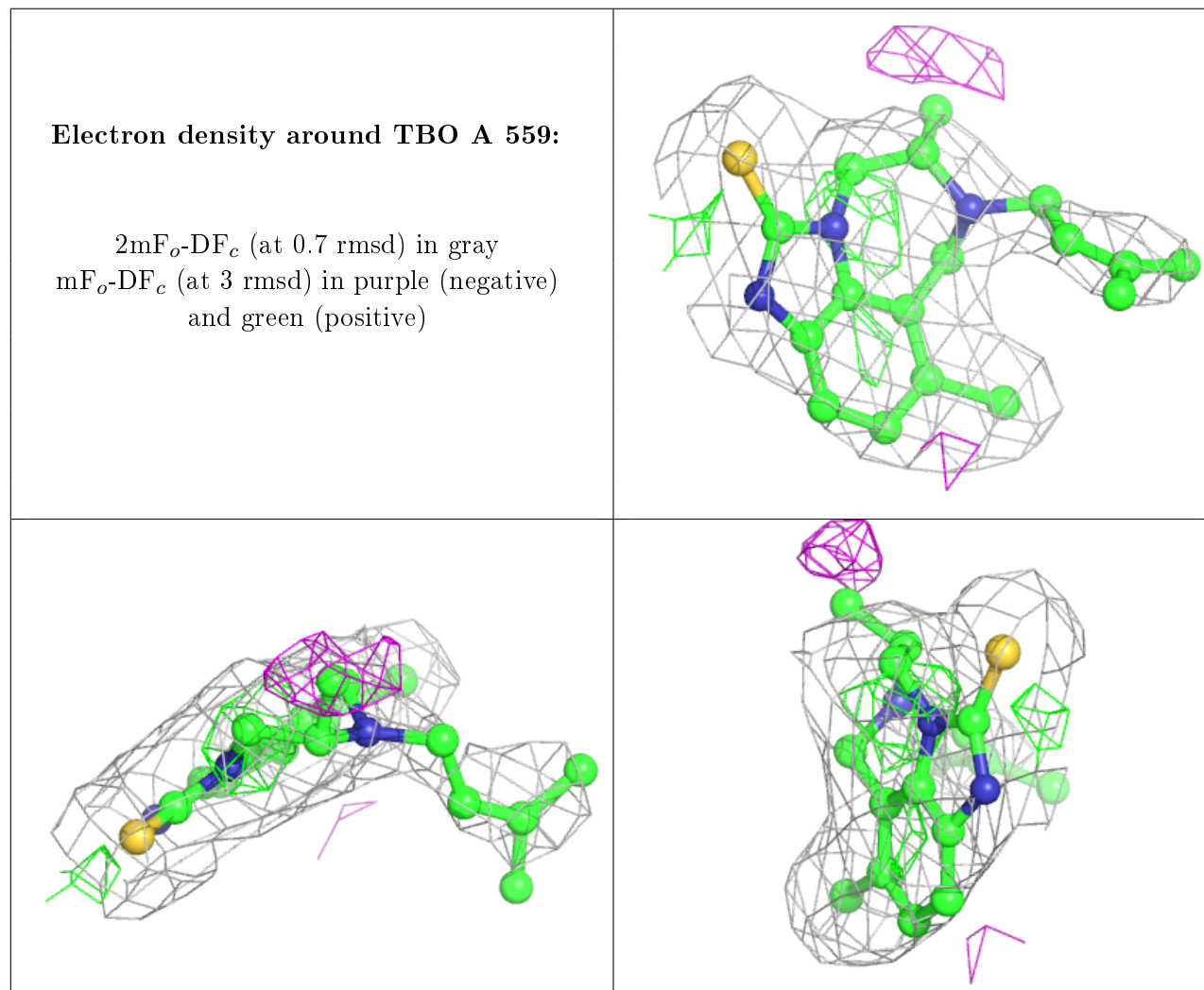
6.3 Carbohydrates (i)

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(Å^2)	Q<0.9
3	TBO	A	559	21/21	0.93	0.24	45,45,45,45	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 TBO A 559:

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