



wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 06:01 am BST

PDB ID : 1HNI
Title : STRUCTURE OF HIV-1 REVERSE TRANSCRIPTASE IN A COMPLEX
WITH THE NONNUCLEOSIDE INHIBITOR ALPHA-APA R 95845 AT 2.8
ANGSTROMS RESOLUTION
Authors : Ding, J.; Das, K.; Arnold, E.
Deposited on : 1995-02-28
Resolution : 2.80 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

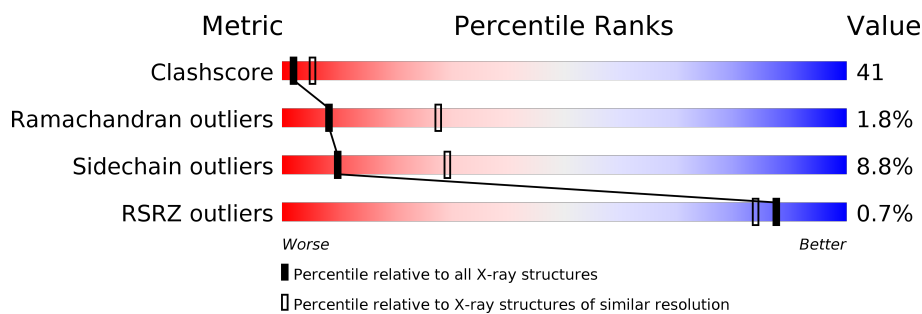
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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))
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 $\leq 5\%$. 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	558	<div> <div>%</div> <div> <div></div> <div>39%</div> <div>53%</div> <div>7%</div> </div> </div>
2	B	427	<div> <div>%</div> <div> <div></div> <div>41%</div> <div>54%</div> <div>5%</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	AAA	A	559	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7819 atoms, of which 4 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	C	N	O	S	0	0	0
			4368	2825	724	813	6			

There is a discrepancy between the modelled and reference sequences:

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

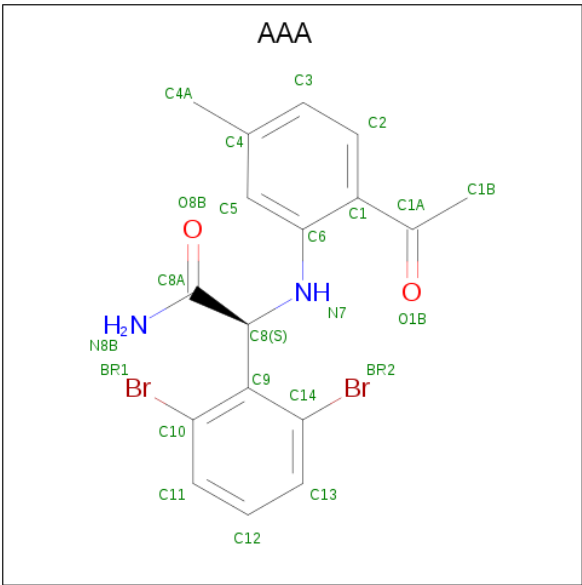
- Molecule 2 is a protein called HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P51).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	427	Total	C	N	O	S	0	0	0
			3424	2231	564	624	5			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	172	ARG	LYS	CONFLICT	UNP P03367
B	280	SER	CYS	CONFLICT	UNP P03367
B	357	THR	MET	CONFLICT	UNP P03367

- Molecule 3 is (2-ACETYL-5-METHYLANILINO)(2,6-DIBROMOPHENYL)ACETAMIDE (three-letter code: AAA) (formula: C₁₇H₁₆Br₂N₂O₂).

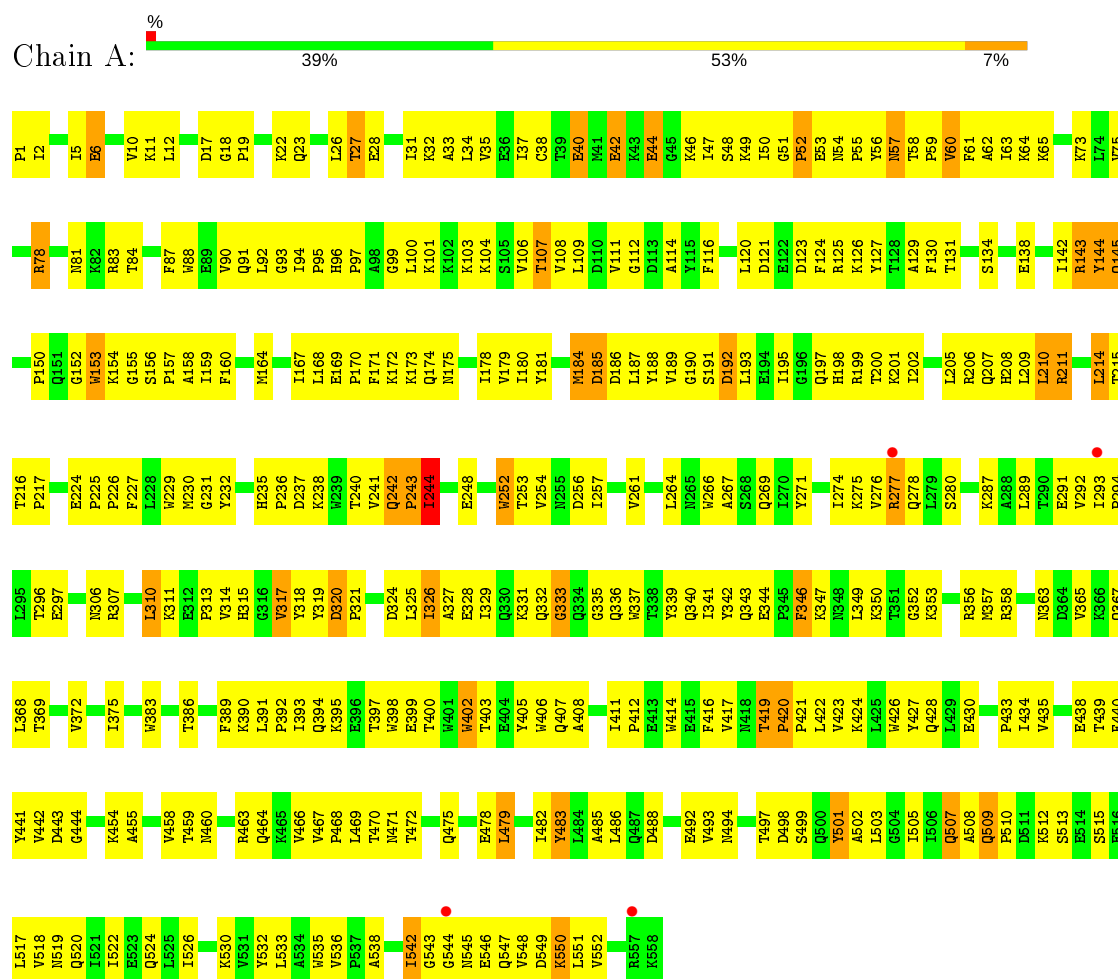


Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	Br	C	H	N	O	0	0
			27	2	17	4	2	2		

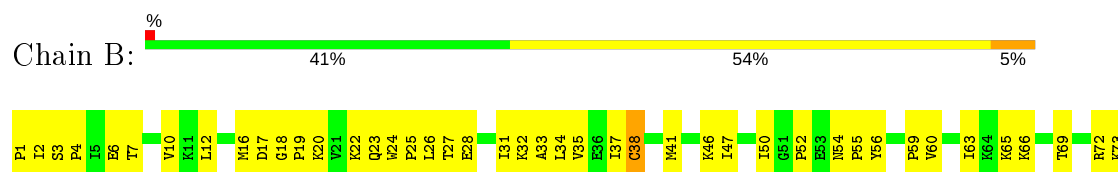
3 Residue-property plots

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: HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P66)



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



L74	P150	L214	T290	D864
V75	Q151	T215	E291	V365
D76	G152	T216	V292	K366
F77	W153	P217	I293	Q367
R78	K154	K220	P294	L368
E79	G155	H221	L295	T369
L80	S156	Q222	T296	E370
N81	P157	K223	E297	A371
K82	A158	E224	E298	V372
R83	I159	E225	A299	Q373
T84	F160	P226	E300	T376
Q85	Q161	M230	L301	S379
D86	S162	G231	E302	I380
F87	S163	Y232	L303	V383
W88	M164	E233	M306	G384
E89	T165	P236	R307	I389
V90	K166	W239	E308	Q394
Q91	I167	V241	I309	K395
I94	L168	Q242	L310	E396
K103	E169	P243	V314	T397
K104	P170	V245	H315	W398
L109	K173	I244	G316	E399
D110	Q174	L246	V317	W402
V111	P176	D250	Y318	T403
G112	D177	S251	D320	E404
D113	I178	W252	P321	Y405
A114	V179	T253	S322	W406
Y115	I180	V254	K323	Q407
F116	Y181	N255	D324	A408
S117	Q182	D256	L325	I411
V118	Y183	I257	E329	P412
P119	M184	Q258	Q336	E413
L120	D185	K259	W337	W414
L121	D186	L260	T338	W418
E122	L187	V261	Y339	T419
D123	Y188	G262	K263	P420
F124	G190	L264	N265	K424
Y127	S191	M265	I274	L425
F130	D192	K275	K276	W426
T131	L193	V276	R277	Y427
P133	E194	Q278	K347	
S134	Q197	L279	M348	
I135	T200	S280	L349	
N136	K201	K281	K350	
N137	I202	L282	T351	
E138	E203	L283	G352	
I142	L205	R284	K353	
R143	R206	L289	Y354	
Y144	Q207		A355	
Q145	H208		R356	
Y146	L209			
Y149	L210			
	R211			

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	223.30 Å 69.90 Å 106.50 Å 90.00° 105.40° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.80 36.66 – 2.60	Depositor EDS
% Data completeness (in resolution range)	(Not available) ((Not available)-2.80) 75.7 (36.66-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.91 (at 2.61 Å)	Xtriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.255 , (Not available) 0.272 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	30.3	Xtriage
Anisotropy	0.172	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 596.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.80	EDS
Total number of atoms	7819	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.93% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: AAA

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	0.67	0/4483	0.50	1/6115 (0.0%)
2	B	1.41	4/3523 (0.1%)	1.25	8/4800 (0.2%)
All	All	1.06	4/8006 (0.0%)	0.91	9/10915 (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
1	A	0	2
2	B	0	1
All	All	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	172	ARG	CD-NE	60.74	2.49	1.46
2	B	22	LYS	CE-NZ	28.75	2.21	1.49
2	B	172	ARG	NE-CZ	12.94	1.49	1.33
2	B	22	LYS	CD-CE	-10.15	1.25	1.51

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	172	ARG	CD-NE-CZ	-63.77	34.32	123.60
2	B	22	LYS	CD-CE-NZ	-36.13	28.59	111.70
2	B	172	ARG	CG-CD-NE	15.56	144.48	111.80
2	B	78	ARG	NE-CZ-NH2	7.46	124.03	120.30
2	B	83	ARG	NE-CZ-NH2	7.32	123.96	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	483	TYR	Sidechain
1	A	501	TYR	Sidechain
2	B	127	TYR	Sidechain

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	4368	0	4256	388	0
2	B	3424	0	3372	267	0
3	A	23	4	16	12	0
All	All	7815	4	7644	631	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 631 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:85:GLN:HG3	2:B:154:LYS:CB	1.67	1.25
1:A:420:PRO:HB2	1:A:421:PRO:HD3	1.27	1.11
1:A:112:GLY:HA2	1:A:185:ASP:OD2	1.56	1.04
2:B:59:PRO:HG2	2:B:76:ASP:HB3	1.35	1.02
1:A:536:VAL:HG21	1:A:542:ILE:HG12	1.41	1.01

There are no symmetry-related clashes.

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	556/558 (100%)	443 (80%)	99 (18%)	14 (2%)	5	19
2	B	425/427 (100%)	354 (83%)	67 (16%)	4 (1%)	17	46
All	All	981/985 (100%)	797 (81%)	166 (17%)	18 (2%)	8	28

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	184	MET
1	A	53	GLU
1	A	153	TRP
1	A	242	GLN
1	A	287	LYS

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	453/498 (91%)	409 (90%)	44 (10%)	8	24
2	B	362/389 (93%)	334 (92%)	28 (8%)	13	35
All	All	815/887 (92%)	743 (91%)	72 (9%)	10	29

5 of 72 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	386	THR
1	A	509	GLN
2	B	340	GLN
1	A	402	TRP
1	A	472	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	507	GLN
1	A	520	GLN
2	B	315	HIS
1	A	474	ASN
2	B	343	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	AAA	A	559	-	24,24,24	1.35	3 (12%)	29,34,34	1.12	2 (6%)

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	AAA	A	559	-	-	3/16/16/16	0/2/2/2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	559	AAA	C5-C6	3.82	1.45	1.39
3	A	559	AAA	C6-N7	2.44	1.43	1.37
3	A	559	AAA	C1-C6	2.43	1.44	1.41

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	559	AAA	O1B-C1A-C1	2.35	123.66	120.56
3	A	559	AAA	C10-C9-C8	-2.26	119.76	122.34

There are no chirality outliers.

All (3) torsion outliers are listed below:

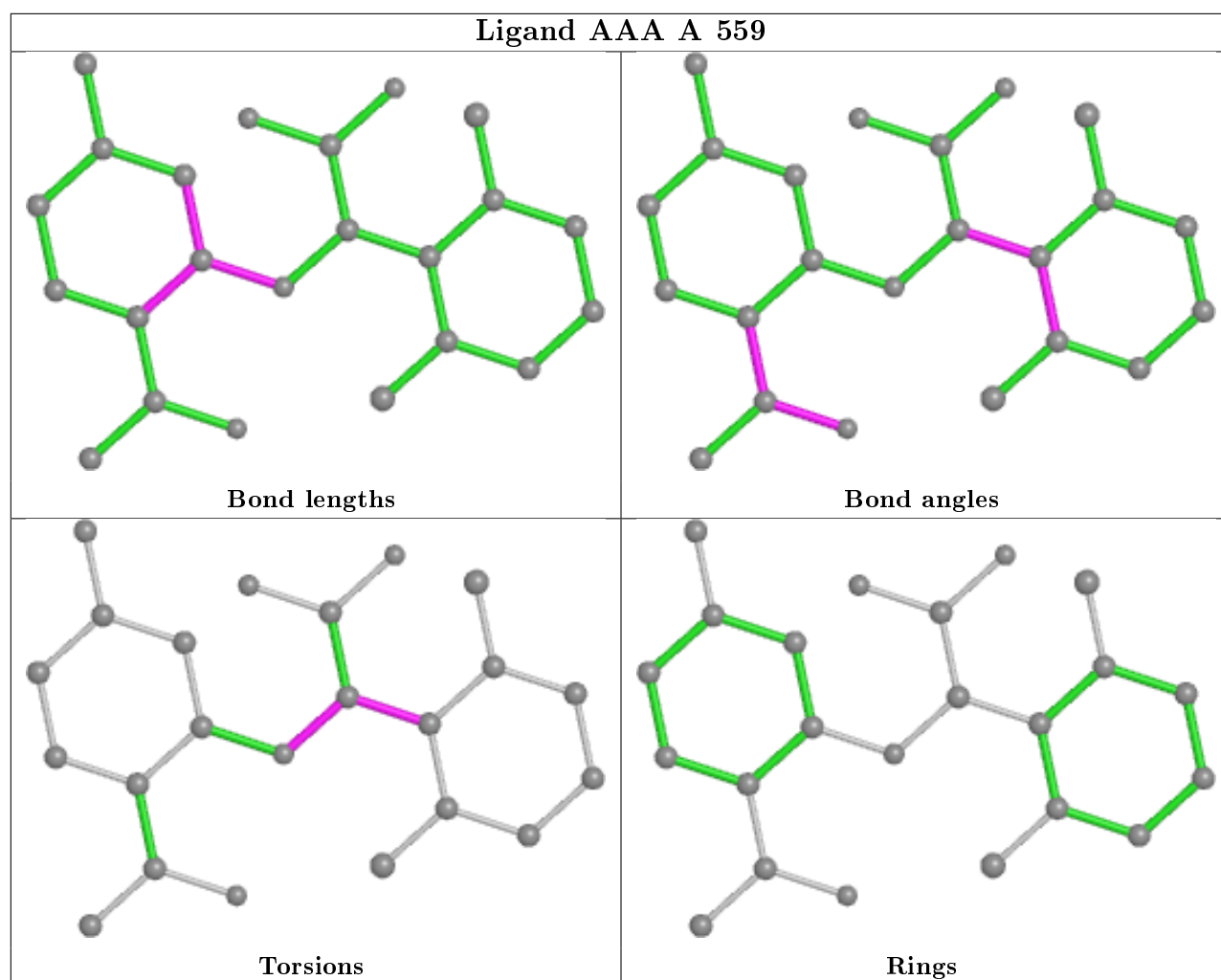
Mol	Chain	Res	Type	Atoms
3	A	559	AAA	C8A-C8-N7-C6
3	A	559	AAA	C9-C8-N7-C6
3	A	559	AAA	C8A-C8-C9-C14

There are no ring outliers.

1 monomer is involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	559	AAA	12	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	558/558 (100%)	-0.43	4 (0%) 87 84	8, 44, 66, 72	0
2	B	427/427 (100%)	-0.47	3 (0%) 87 84	5, 33, 64, 72	0
All	All	985/985 (100%)	-0.45	7 (0%) 87 84	5, 41, 65, 72	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	231	GLY	3.8
2	B	282	LEU	3.6
1	A	293	ILE	3.2
1	A	557	ARG	3.0
1	A	544	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](#)

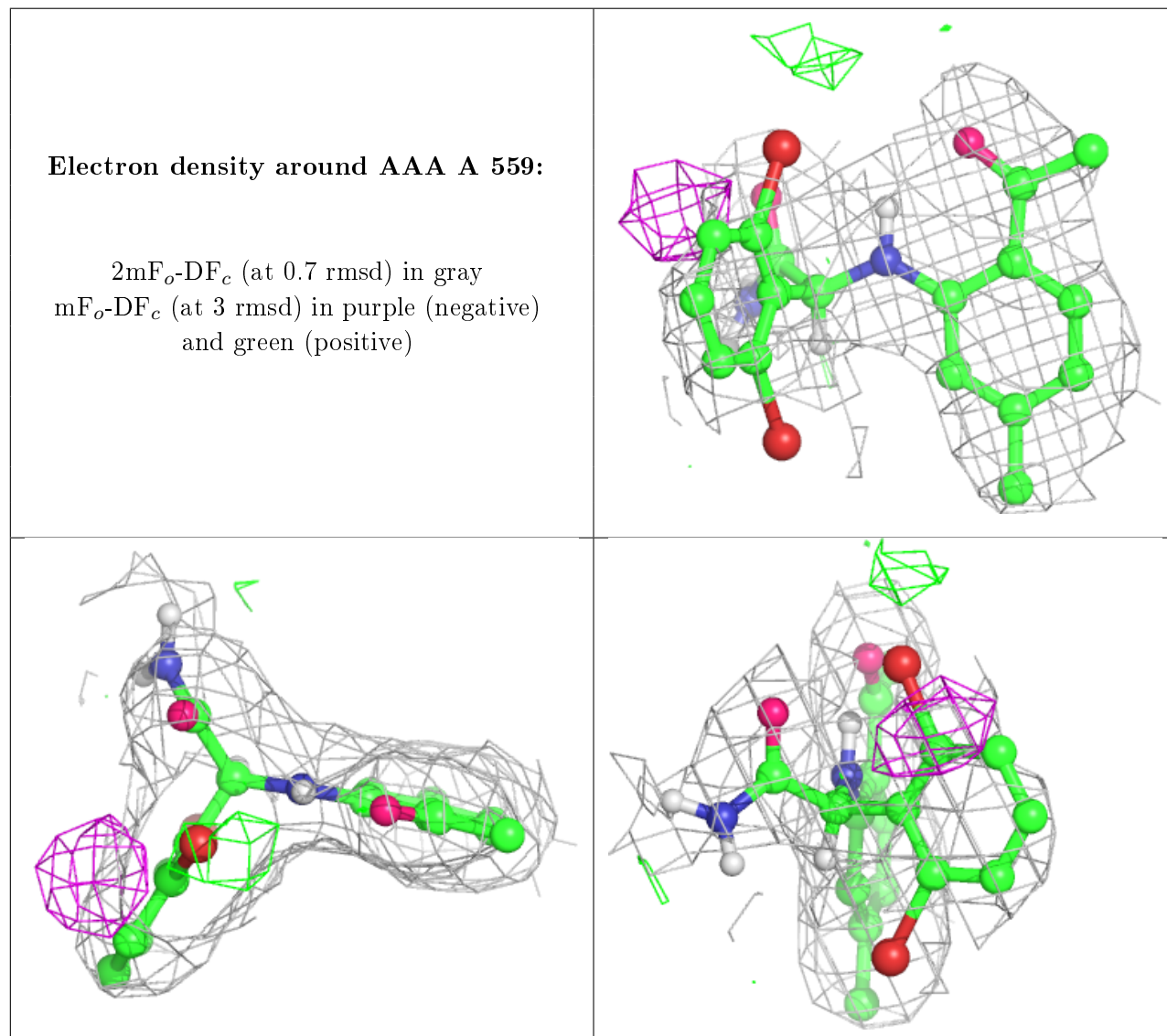
There are no carbohydrates in this entry.

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(\AA^2)	Q<0.9
3	AAA	A	559	23/23	0.94	0.15	10,20,28,32	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.