



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 07:49 am BST

PDB ID : 2VG5
Title : Crystal structures of HIV-1 reverse transcriptase complexes with thiocarbamate non-nucleoside inhibitors
Authors : Spallarossa, A.; Cesarini, S.; Ranise, A.; Ponassi, M.; Unge, T.; Bolognesi, M.
Deposited on : 2007-11-08
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

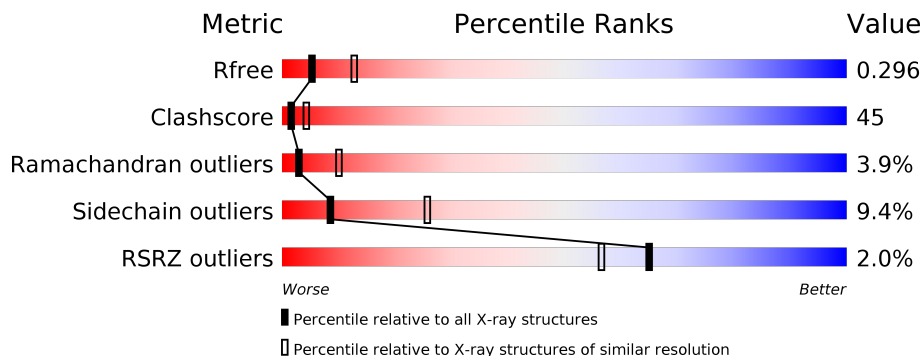
1 Overall quality at a glance

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))
R_{free}	130704	3140 (2.80-2.80)
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	557	 43% 47% 9%
2	B	428	 39% 45% 9% 6%

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	NNC	A	1551	-	-	X	-

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7872 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.

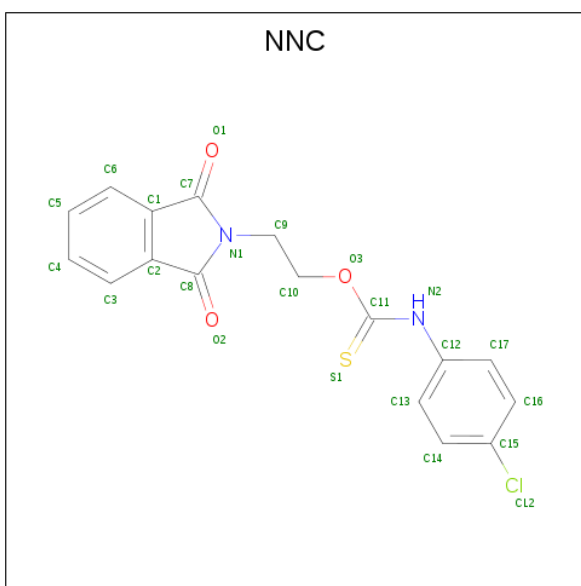
- Molecule 1 is a protein called REVERSE TRANSCRIPTASE/RIBONUCLEASE H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	551	4475	2898	744	825	8	0	0	0

- Molecule 2 is a protein called P51 RT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	401	3318	2163	543	605	7	0	0	0

- Molecule 3 is O-[2-(1,3-dioxo-1,3-dihydro-2H-isoindol-2-yl)ethyl] (4-chlorophenyl)thiocarbamate (three-letter code: NNC) (formula: C₁₇H₁₃ClN₂O₃S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	Cl	N	O			S
3	A	1	24	17	1	2	3	1	0	0

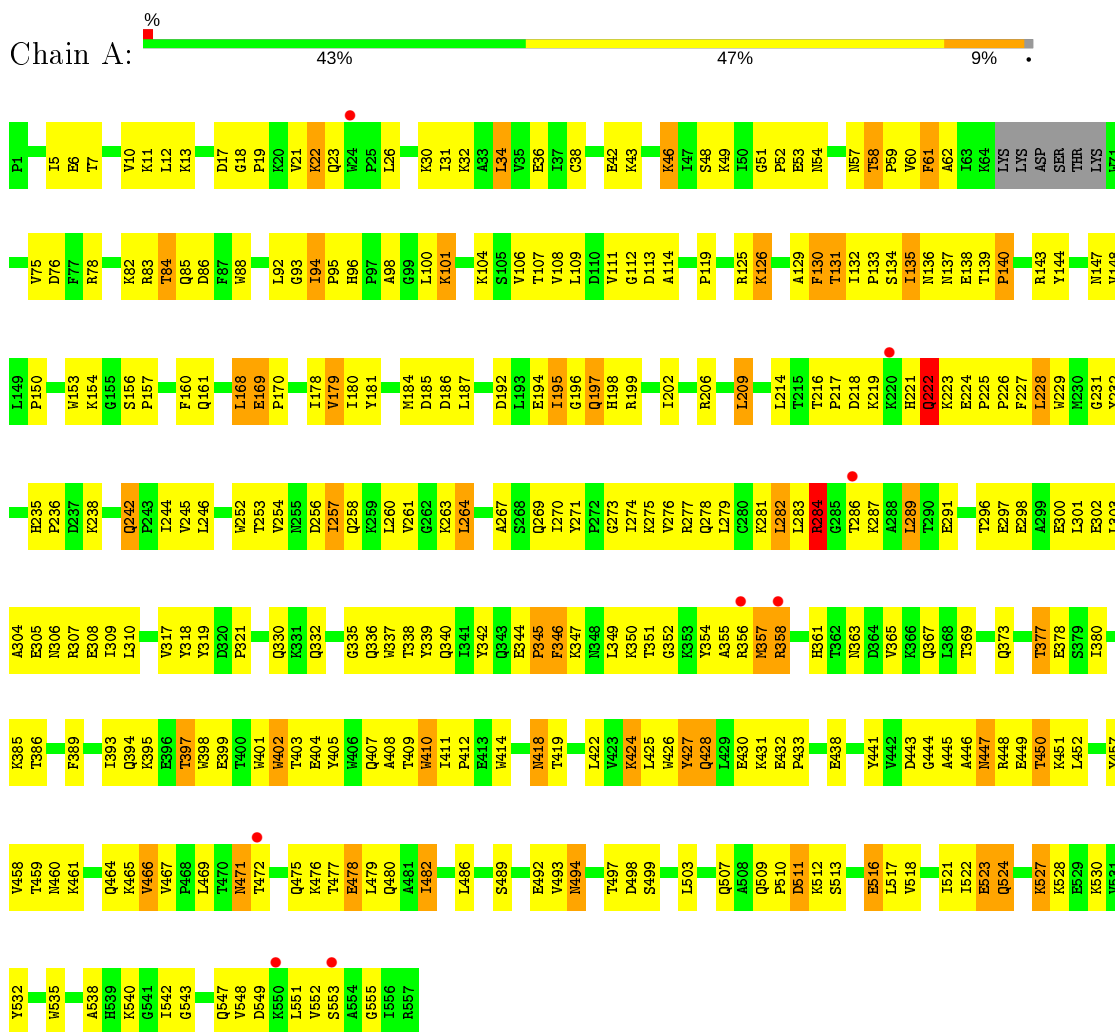
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	30	Total 30	O 30	0	0
4	B	25	Total 25	O 25	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: REVERSE TRANSCRIPTASE/RIBONUCLEASE H



D1364	V1365	K1366	Q1367	L1368	T1369	E1370	A1371	V1372	I1375	I1380	V1381	I1382	K1385	K1388	F1389	K1390	L1391	P1392	I1393	Q1394	K1395	K1396	T1397	W1398	W1401	V1402	T1403	E1404	Y1405	W1406	Q1407	E1413	W1414	N1418	T1419	P1420	P1421	L1422	V1423	K1424	L1425	W1426	Y1427	Q1428												
T1296	E1297	E1298	A1299	E1300	L1301	E1302	L1303	A1304	E1305	N1306	R1307	E1308	I1309	L1310	K1311	E1312	P1313	V1314	V1317	Y1318	Y1319	K1323	D1324	L1325	I1326	Q1330	K1331	Q1332	G1333	Q1334	G1335	Q1336	W1337	T1338	Y1342	Q1343	E1344	P1345	F1346	K1347	N1348	L1349	K1350	T1351	Y1354	A1355	ARG	MET	ARG	GLY	ALA	H1361	T1362	N1363		
Y1282	E1283	L1284	H1285	P1286	D1287	K1288	W1289	T1290	V1241	Q1242	P1243	I1244	V1245	L1246	P1247	E1248	K1249	D1250	T1253	V1254	N1255	D1256	I1257	Q1258	K1259	L1260	V1261	G1262	K1263	L1264	M1265	W1266	A1267	S1268	Q1269	I1270	Y1271	I1274	K1275	V1276	R1277	Q1278	L1279	C1280	K1281	L1282	LEU	ARG	GLY	T1286	K1287	A1288	L1289	I1293	P1294	L1295
K1154	G1155	S1156	P1157	A1158	I1159	F1160	Q1161	K1166	I1167	L1168	E1169	P1170	N1175	I1178	I1094	V1179	I1180	Y1183	M1184	D1185	D1186	L1187	Y1188	V1189	L1190	S1191	I1111	G1112	D1113	A1114	Y1115	F1116	S1117	V1118	E1122	W1212	G1213	L1214	THR	THR	PRO	ASP	LYS	LYS	HIS	GLN	LYS	GLU	Q1145	PRO	PRO	PHE	LEU	M1230	G1231	
F1077	R1078	E1079	L1080	N1081	K1082	R1083	T1084	Q1085	D1086	F1087	W1088	E1089	V1090	Q1091	L1092	G1093	I1094	P1095	H1096	K1101	K1104	S1105	V1106	T1107	V1108	G1109	D1110	V1111	G1112	D1113	A1114	Y1115	F1116	S1117	V1118	E1122	R1125	K1126	F1130	T1131	I1132	W1136	N1137	Y1144	Q1145	Y1146	L1149	F1150	Q1151	W1153						

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	118.86Å 156.13Å 154.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.80 19.95 – 2.80	Depositor EDS
% Data completeness (in resolution range)	89.0 (20.00-2.80) 89.0 (19.95-2.80)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.60 (at 2.79Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.237 , 0.297 0.235 , 0.296	Depositor DCC
R_{free} test set	1589 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	52.5	Xtrriage
Anisotropy	0.004	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 46.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	7872	wwPDB-VP
Average B, all atoms (Å ²)	45.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.83% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical 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: NNC

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	0.26	1/4592 (0.0%)	0.40	0/6240
2	B	0.24	0/3411	0.40	0/4632
All	All	0.25	1/8003 (0.0%)	0.40	0/10872

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	478	GLU	CD-OE2	6.97	1.33	1.25

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	A	4475	0	4521	428	0
2	B	3318	0	3341	292	0
3	A	24	0	13	12	0
4	A	30	0	0	9	0
4	B	25	0	0	6	0
All	All	7872	0	7875	705	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 45.

The worst 5 of 705 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:107:THR:HG21	1:A:202:ILE:CD1	1.42	1.44
1:A:450:THR:CG2	1:A:452:LEU:HD22	1.51	1.38
1:A:357:MET:H	1:A:357:MET:CE	1.51	1.24
1:A:450:THR:HG21	1:A:452:LEU:HD22	1.22	1.18
2:B:1428:GLN:O	2:B:1428:GLN:NE2	1.77	1.17

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	547/557 (98%)	480 (88%)	48 (9%)	19 (4%)	3	12
2	B	393/428 (92%)	331 (84%)	44 (11%)	18 (5%)	2	7
All	All	940/985 (95%)	811 (86%)	92 (10%)	37 (4%)	3	10

5 of 37 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	84	THR
1	A	140	PRO
1	A	222	GLN
2	B	1065	LYS
2	B	1085	GLN

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	489/497 (98%)	445 (91%)	44 (9%)	9	28
2	B	366/390 (94%)	330 (90%)	36 (10%)	8	24
All	All	855/887 (96%)	775 (91%)	80 (9%)	8	26

5 of 80 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	471	ASN
2	B	1011	LYS
2	B	1336	GLN
1	A	482	ILE
1	A	516	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	418	ASN
1	A	507	GLN
2	B	1340	GLN
1	A	428	GLN
1	A	447	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 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	NNC	A	1551	-	26,26,26	2.44	6 (23%)	36,36,36	3.10	12 (33%)

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	NNC	A	1551	-	-	0/10/26/26	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	1551	NNC	C1-C7	-5.88	1.39	1.48
3	A	1551	NNC	C2-C8	-5.76	1.39	1.48
3	A	1551	NNC	C8-N1	-5.44	1.33	1.39
3	A	1551	NNC	C7-N1	-5.14	1.33	1.39
3	A	1551	NNC	C12-N2	-4.08	1.33	1.41

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1551	NNC	C10-O3-C11	-11.77	109.55	119.11
3	A	1551	NNC	O3-C11-S1	-7.64	119.88	125.10
3	A	1551	NNC	C12-N2-C11	-5.19	120.73	130.00
3	A	1551	NNC	C1-C7-N1	4.74	109.25	105.88

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1551	NNC	C2-C8-N1	4.71	109.22	105.88

There are no chirality outliers.

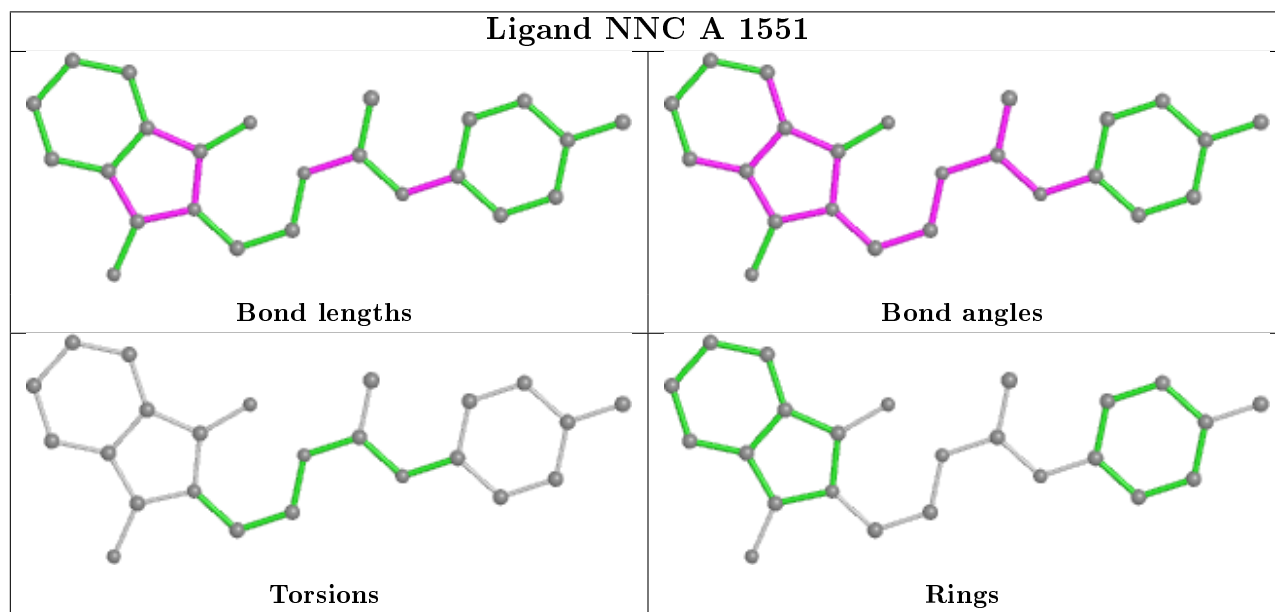
There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1551	NNC	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	551/557 (98%)	-0.09	8 (1%) 73 68	20, 44, 67, 85	0
2	B	401/428 (93%)	-0.17	11 (2%) 54 44	23, 40, 75, 90	0
All	All	952/985 (96%)	-0.13	19 (1%) 65 56	20, 43, 72, 90	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	286	THR	6.0
1	A	24	TRP	4.4
2	B	1067	ASP	3.7
2	B	1230	MET	3.6
1	A	472	THR	3.5

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](#)

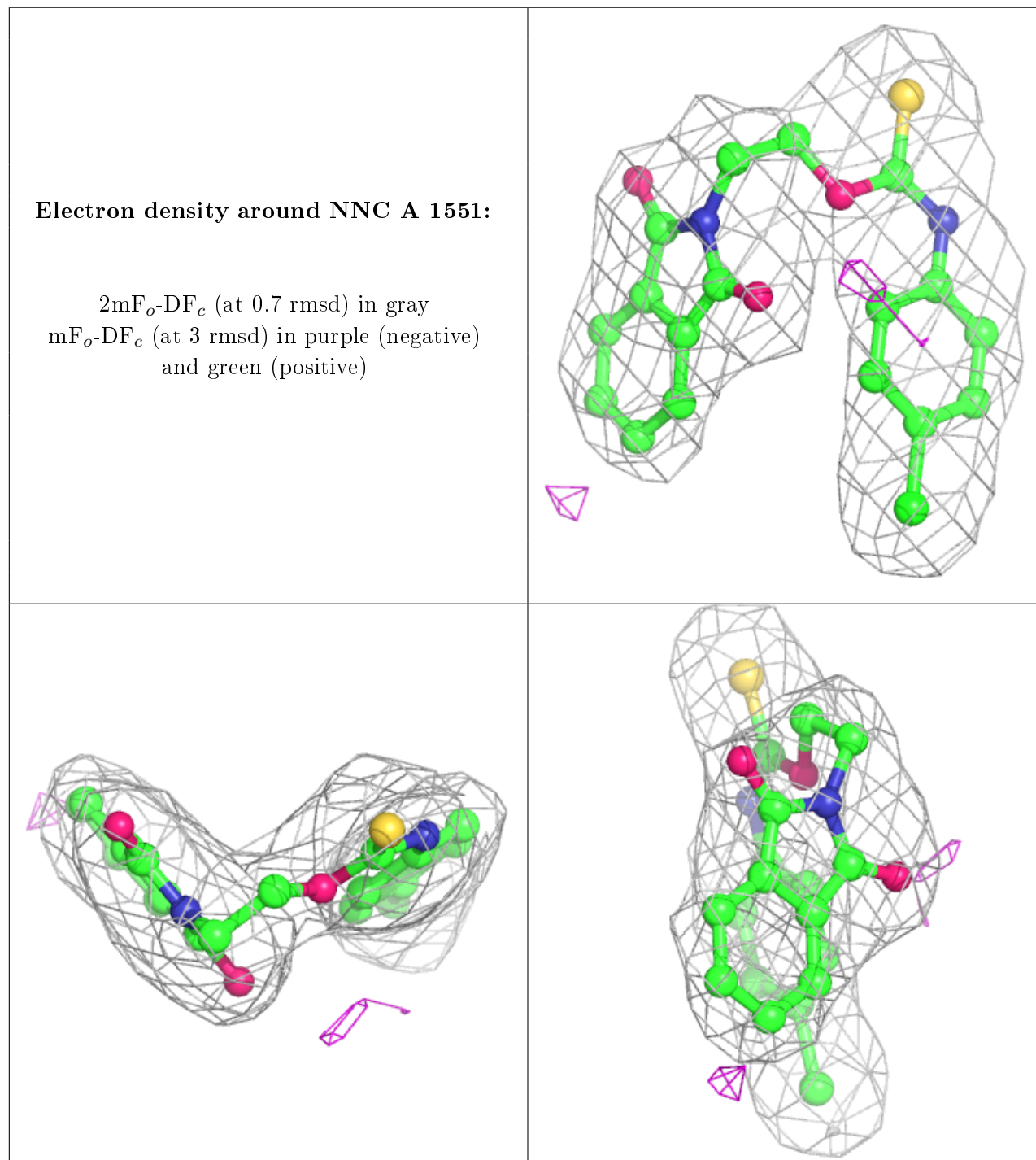
There are no carbohydrates in this entry.

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	NNC	A	1551	24/24	0.97	0.14	26,29,36,38	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 [i](#)

There are no such residues in this entry.