



wwPDB X-ray Structure Validation Summary Report

Mar 5, 2026 – 04:45 PM UTC

PDB ID : 2VG5 / pdb_00002vg5
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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtrriage (Phenix) : 2.0
EDS : 3.0
Buster-report : wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

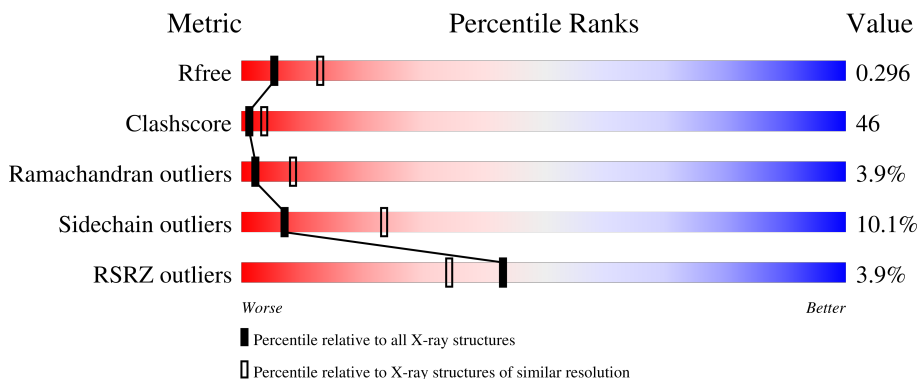
1 Overall quality at a glance

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}	180053	3866 (2.80-2.80)
Clashscore	190562	4276 (2.80-2.80)
Ramachandran outliers	187476	4196 (2.80-2.80)
Sidechain outliers	187428	4198 (2.80-2.80)
RSRZ outliers	180081	3869 (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 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 $\leq 5\%$. 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	
2	B	428	

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.

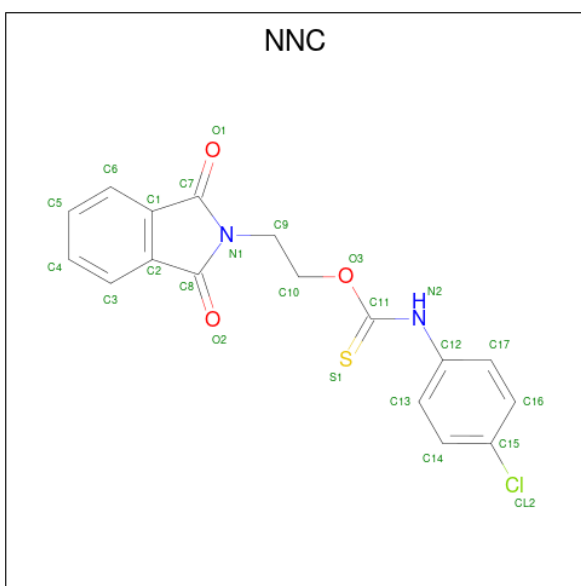
- Molecule 1 is a protein called REVERSE TRANSCRIPTASE/RIBONUCLEASE H.

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 (CCD ID: 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

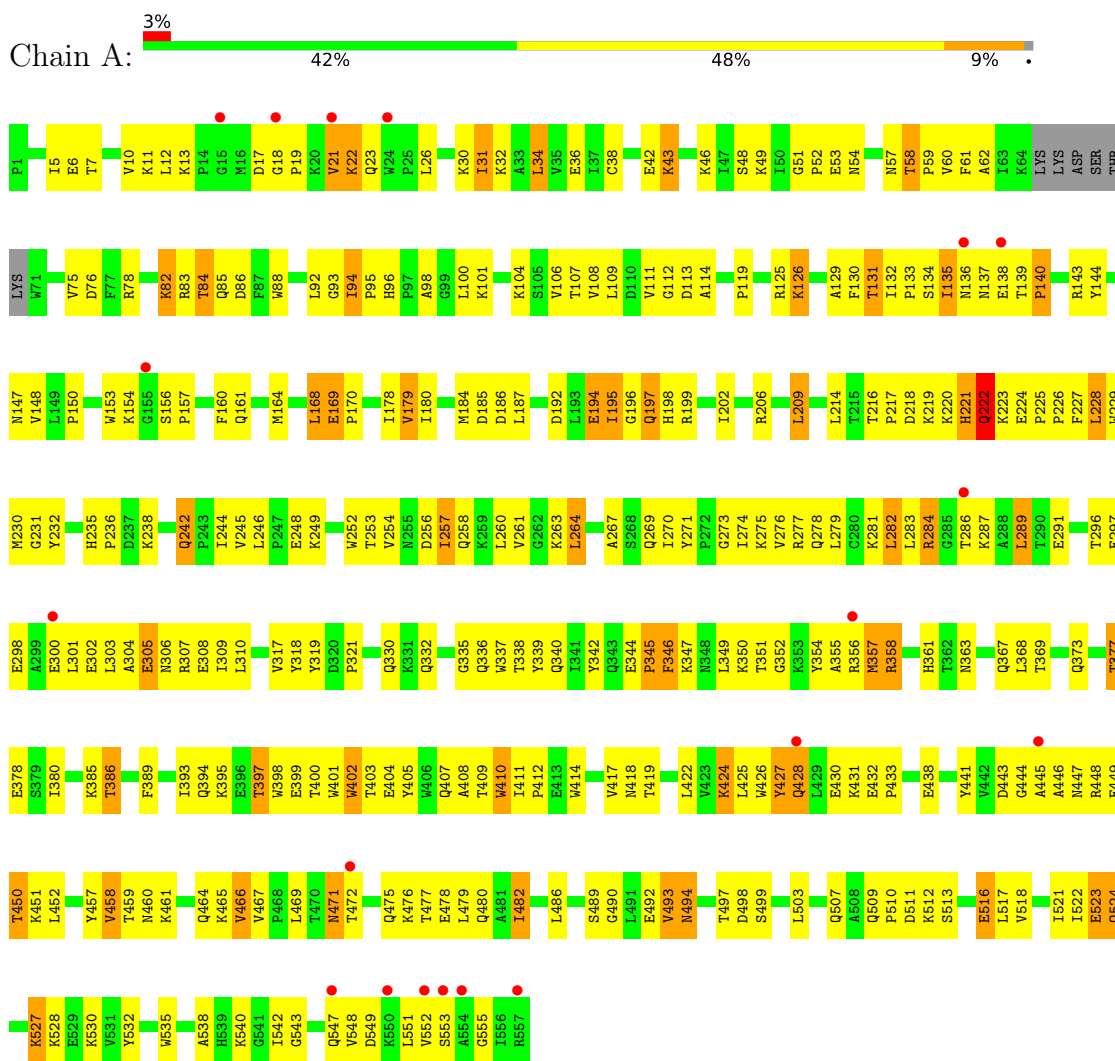
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	30	Total 30	O 30	0	0
4	B	25	Total 25	O 25	0	0

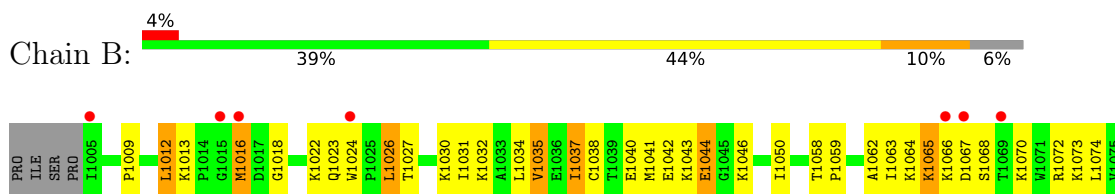
3 Residue-property plots

These plots are drawn for all protein, RNA, 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/RIBONUCLEASE H



• Molecule 2: P51 RT



D1076	G1152	W1230	P1294	H1363
F1077	W1153	G1231	L1295	T1362
R1078	K1154	Y1232	E1297	D1364
E1079	G1155	E1233	E1298	V1365
L1080	S1156	L1234	A1299	K1366
N1081	P1157	H1235	E1300	Q1367
K1082	A1158	P1236	L1301	L1368
R1083	I1159	D1237	E1302	T1369
T1084	F1160	K1238	L1303	A1370
Q1085	Q1161	W1239	A1304	A1371
D1086	K1166	T1240	E1305	V1372
F1087	I1167	V1241	M1306	I1375
H1088	L1168	Q1242	R1307	I1380
E1089	E1169	P1243	E1308	V1381
V1090	P1170	I1244	I1309	L1382
Q1091	M1175	V1245	L1310	K1385
L1092	N1175	L1246	K1311	K1388
G1093	L1178	P1247	E1312	F1389
I1094	I1178	E1248	P1313	K1390
F1095	V1179	K1249	V1314	L1391
H1096	I1180	D1250	V1317	F1392
K1101	I1183	T1253	Y1318	I1393
K1104	M1184	V1254	Y1319	Q1394
S1105	D1185	N1255	D1320	K1395
V1106	D1186	I1257	K1323	E1396
T1107	L1187	Q1258	D1324	T1397
V1108	Y1188	K1259	L1325	W1398
L1109	V1189	L1260	I1326	W1401
D1110	G1190	V1261	Q1330	T1402
V1111	S1191	G1262	K1331	L1403
G1112	I1195	K1263	Q1332	E1404
D1113	A1114	L1264	G1333	Y1405
A1114	I1202	N1265	Q1334	W1406
Y1115	L1205	W1266	G1335	Q1407
F1116	R1206	A1267	Q1336	E1413
S1117	Q1207	S1268	W1337	W1414
V1118	Y1271	Q1269	T1338	M1418
D1121	R1211	Y1274	Y1342	T1419
E1122	W1212	I1274	Q1343	P1420
R1125	G1213	K1275	E1344	P1421
K1126	L1214	V1276	P1345	L1422
F1130	THR	R1277	F1346	V1423
PRO	PRO	Q1278	K1347	K1424
T1131	ASP	L1279	N1348	L1425
I1132	LYS	C1280	L1349	W1426
L1136	LYS	K1281	K1350	Y1427
N1137	LYS	L1282	T1351	Q1428
Y1144	HIS	LEU	Y1354	
Q1145	GLN	ARG	A1355	
V1148	LYS	GLY	ARG	
L1149	GLU	T1286	MET	
P1150	PRO	K1287	ARG	
F1151	PRO	A1288	GLY	
Q1151	LEU	L1289	ALA	
	TRP	I1293		

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 20.00 – 2.80	Depositor EDS
% Data completeness (in resolution range)	89.0 (20.00-2.80) 88.7 (20.00-2.80)	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)	0.33 , 39.2	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.*

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: 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	0.30	0/4592	0.69	0/6240
2	B	0.29	0/3411	0.66	0/4632
All	All	0.29	0/8003	0.68	0/10872

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	4475	0	4521	442	0
2	B	3318	0	3341	301	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	728	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 46.

The worst 5 of 728 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:HE3	1:A:357:MET:N	1.37	1.33
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

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	10
2	B	393/428 (92%)	331 (84%)	44 (11%)	18 (5%)	2	6
All	All	940/985 (95%)	811 (86%)	92 (10%)	37 (4%)	2	8

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%)	441 (90%)	48 (10%)	7	25
2	B	366/390 (94%)	328 (90%)	38 (10%)	7	22
All	All	855/887 (96%)	769 (90%)	86 (10%)	7	23

5 of 86 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	1090	VAL
2	B	1280	CYS
2	B	1109	LEU
2	B	1240	THR
2	B	1314	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	1269	GLN
2	B	1334	GLN
2	B	1367	GLN
1	A	361	HIS
1	A	340	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 oligosaccharides 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	2.45	6 (23%)	36,36,36	3.05	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	-5.95	1.39	1.48
3	A	1551	NNC	C2-C8	-5.83	1.39	1.48
3	A	1551	NNC	C8-N1	-5.43	1.33	1.39
3	A	1551	NNC	C7-N1	-5.14	1.33	1.39
3	A	1551	NNC	C12-N2	-4.09	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	-12.11	109.55	119.04
3	A	1551	NNC	O3-C11-S1	-6.45	119.88	125.07
3	A	1551	NNC	C12-N2-C11	-5.41	120.73	130.07
3	A	1551	NNC	C1-C7-N1	4.55	109.25	105.88
3	A	1551	NNC	C2-C8-N1	4.52	109.22	105.88

There are no chirality outliers.

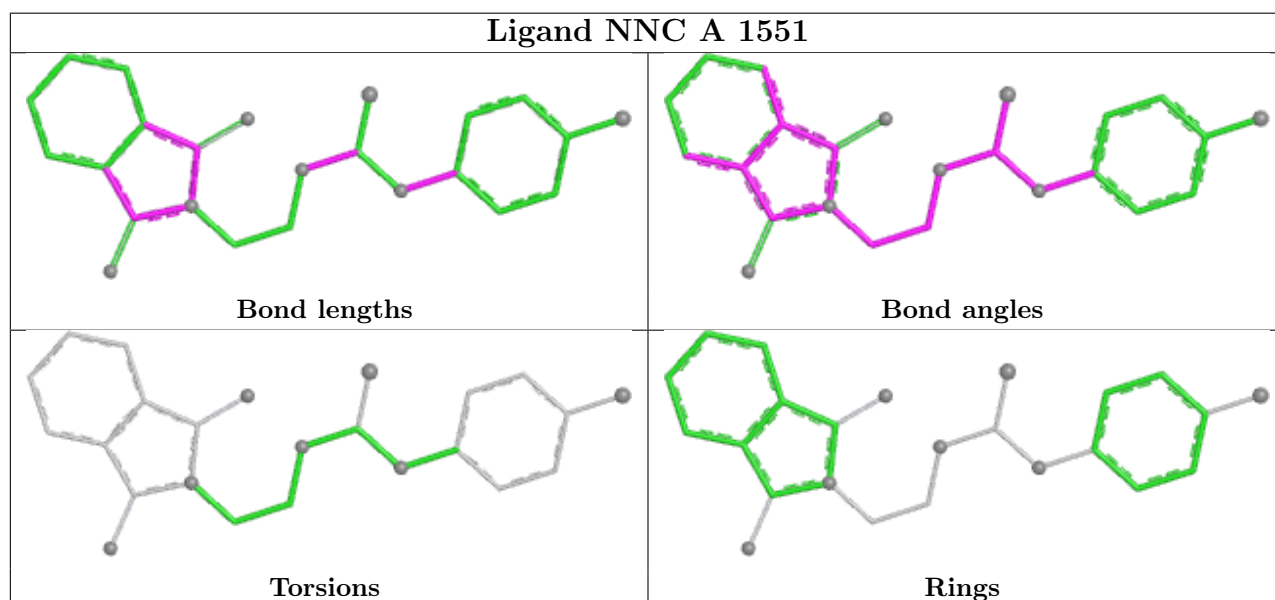
There are no torsion outliers.

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%)	0.21	19 (3%) 48 39	20, 44, 67, 85	0
2	B	401/428 (93%)	0.10	18 (4%) 38 30	23, 40, 75, 90	0
All	All	952/985 (96%)	0.16	37 (3%) 43 34	20, 43, 72, 90	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	24	TRP	5.7
1	A	472	THR	5.1
2	B	1066	LYS	4.1
1	A	136	ASN	3.6
1	A	356	ARG	3.4

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](#)

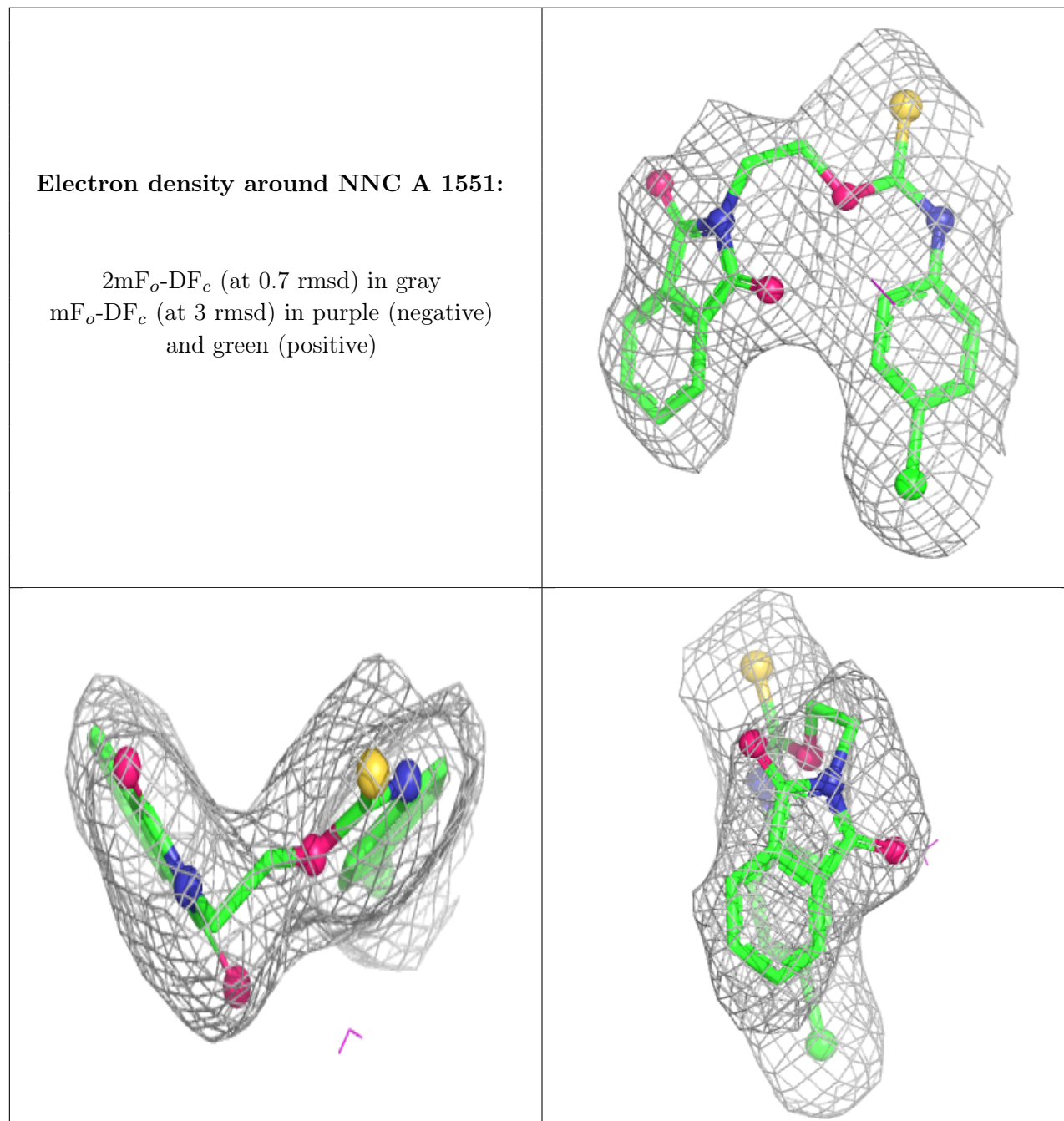
There are no oligosaccharides in this entry.

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	0.06	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.