Full wwPDB X-ray Structure Validation Report ⓘ

Nov 22, 2023 – 06:04 PM EST

PDB ID	:	8STT
Title	:	Crystal Structure of HIV-1 Reverse Transcriptase (Y181C, V106A) varient in Complex with 8-(2-(2-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)ethoxy)phenoxy)indolizine-2-carbonitrile (JLJ555), a non-nucleoside inhibitor
Authors	:	Hollander, K.; Jorgensen, W.L.; Anderson, K.S.
Deposited on	:	2023-05-11
Resolution	:	2.62 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

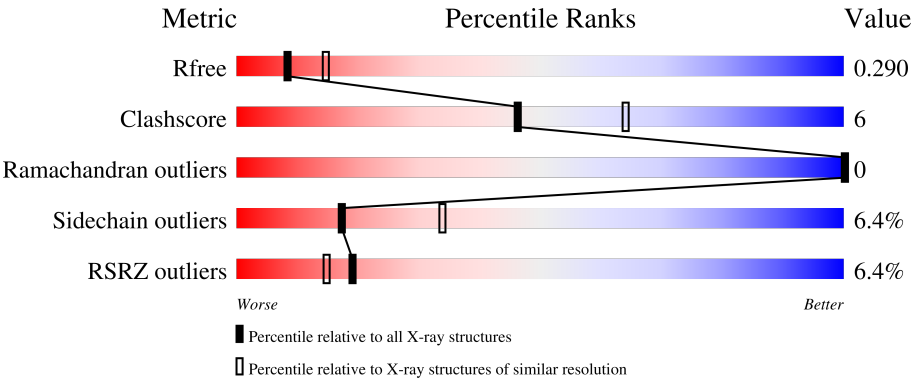
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.62 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	3797 (2.64-2.60)
Clashscore	141614	4168 (2.64-2.60)
Ramachandran outliers	138981	4093 (2.64-2.60)
Sidechain outliers	138945	4093 (2.64-2.60)
RSRZ outliers	127900	3731 (2.64-2.60)

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	558	4% 81% 16%
1	C	558	8% 77% 19%
2	B	428	2% 75% 18% 5%
2	D	428	10% 77% 15% 6%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 14881 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
1	A	547	Total	C	N	O	S	0	1	0
			4184	2692	686	799	7			
1	C	544	Total	C	N	O	S	0	3	0
			4225	2706	704	808	7			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	expression tag	UNP P03366
A	0	VAL	-	expression tag	UNP P03366
A	106	ALA	VAL	engineered mutation	UNP P03366
A	172	ALA	LYS	engineered mutation	UNP P03366
A	173	ALA	LYS	engineered mutation	UNP P03366
A	181	CYS	TYR	engineered mutation	UNP P03366
A	280	SER	CYS	engineered mutation	UNP P03366
C	-1	MET	-	expression tag	UNP P03366
C	0	VAL	-	expression tag	UNP P03366
C	106	ALA	VAL	engineered mutation	UNP P03366
C	172	ALA	LYS	engineered mutation	UNP P03366
C	173	ALA	LYS	engineered mutation	UNP P03366
C	181	CYS	TYR	engineered mutation	UNP P03366
C	280	SER	CYS	engineered mutation	UNP P03366

- Molecule 2 is a protein called p51 RT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	405	Total	C	N	O	S	23	2	0
			3270	2123	532	609	6			
2	D	401	Total	C	N	O	S	14	1	0
			3078	1989	508	577	4			

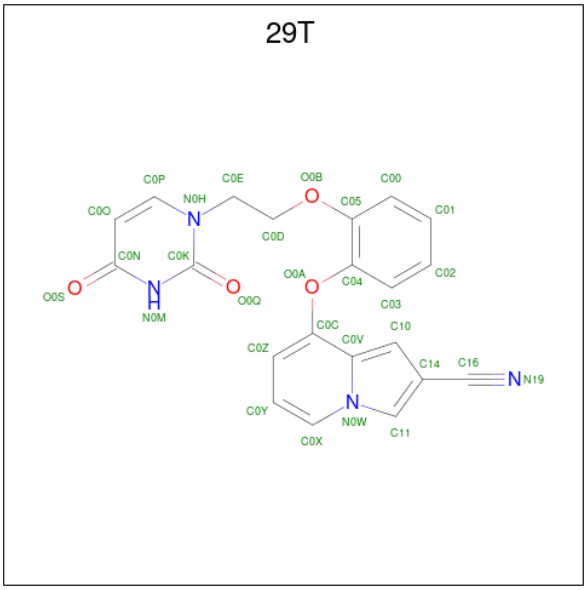
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	280	SER	CYS	engineered mutation	UNP P03366
D	280	SER	CYS	engineered mutation	UNP P03366

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Mg 1 1	0	0
3	C	1	Total Mg 1 1	0	0

- Molecule 4 is 8-{2-[2-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)ethoxy]phenoxy}indolizine-2-carbonitrile (three-letter code: 29T) (formula: $C_{21}H_{16}N_4O_4$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 29 21 4 4	0	0
4	C	1	Total C N O 29 21 4 4	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	26	Total O 26 26	0	0

Continued on next page...

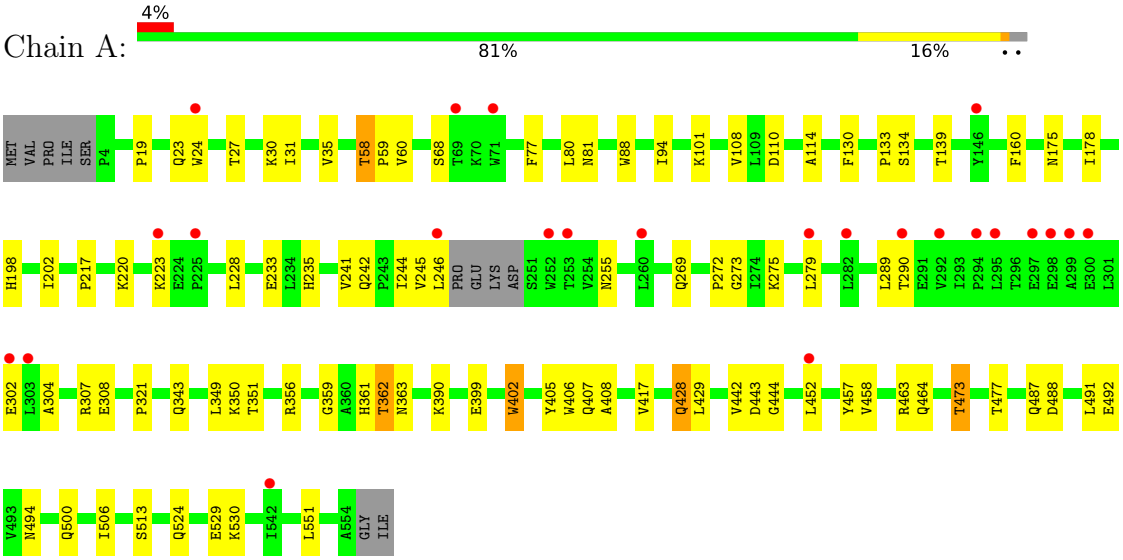
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	27	Total O 27 27	0	0
5	C	10	Total O 10 10	0	0
5	D	1	Total O 1 1	0	0

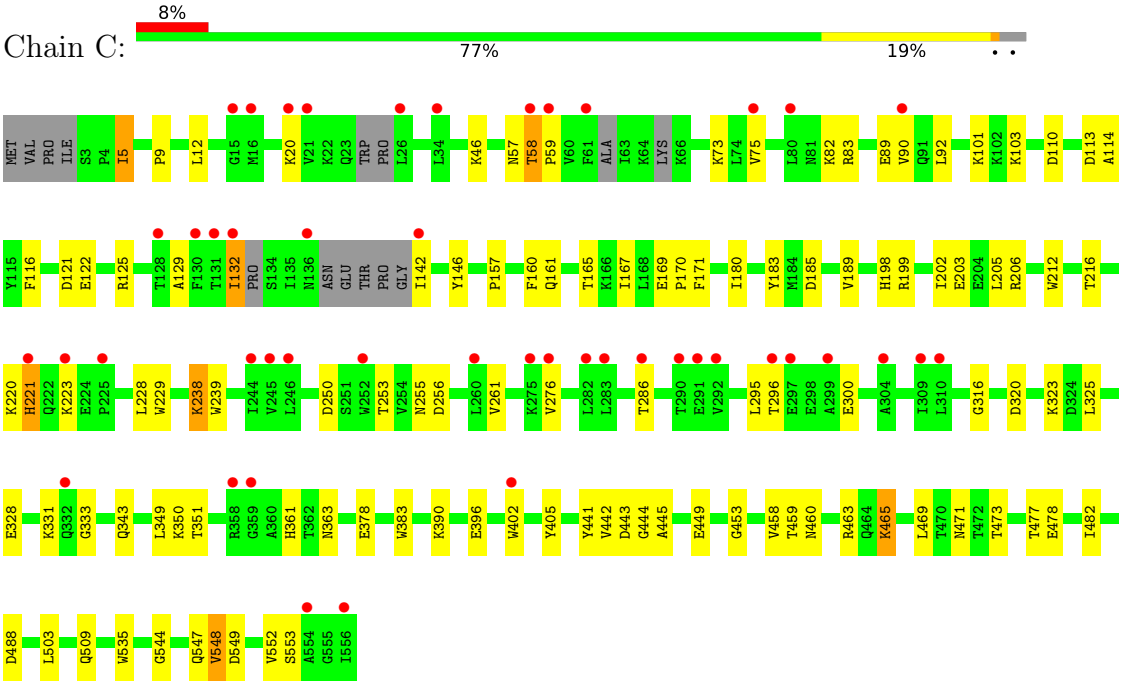
3 Residue-property plots (i)

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/ribonuclease H

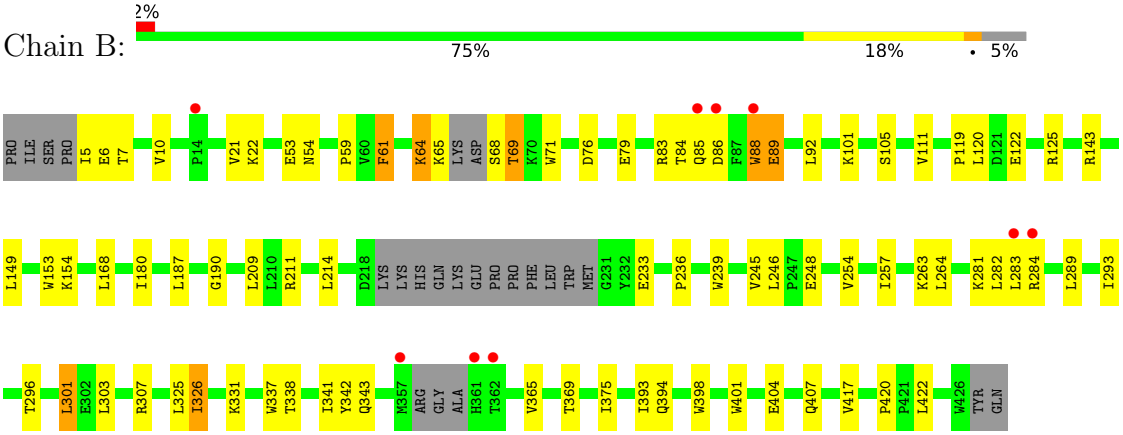


• Molecule 1: Reverse transcriptase/ribonuclease H



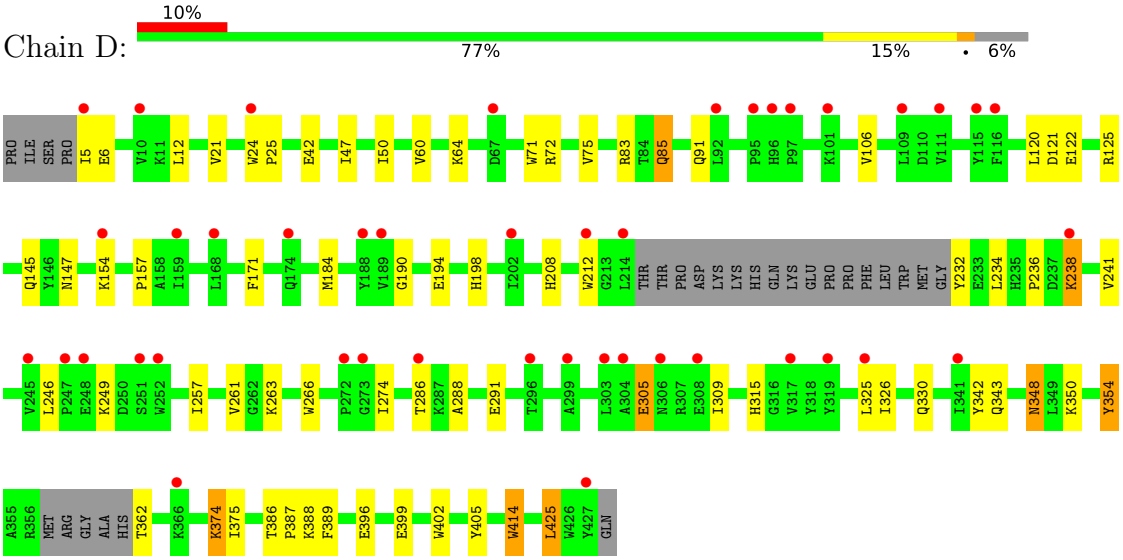
• Molecule 2: p51 RT

Chain B:



• Molecule 2: p51 RT

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	111.69Å 73.01Å 170.47Å 90.00° 97.60° 90.00°	Depositor
Resolution (Å)	54.79 – 2.62 54.79 – 2.62	Depositor EDS
% Data completeness (in resolution range)	99.1 (54.79-2.62) 99.1 (54.79-2.62)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.07 (at 2.61Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.263 , 0.291 0.261 , 0.290	Depositor DCC
R_{free} test set	4011 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å^2)	77.9	Xtriage
Anisotropy	0.071	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 59.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	14881	wwPDB-VP
Average B, all atoms (Å^2)	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 29T, MG

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	0/4296	0.49	0/5880
1	C	0.29	0/4328	0.47	0/5894
2	B	0.32	0/3366	0.51	0/4592
2	D	0.27	0/3169	0.45	0/4343
All	All	0.30	0/15159	0.48	0/20709

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	4184	0	3930	51	0
1	C	4225	0	4032	46	0
2	B	3270	0	3199	52	0
2	D	3078	0	2821	40	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
4	A	29	0	16	0	0
4	C	29	0	16	1	0
5	A	26	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	27	0	0	2	0
5	C	10	0	0	0	0
5	D	1	0	0	0	0
All	All	14881	0	14014	180	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 6.

All (180) 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:C:328:GLU:HG2	1:C:390:LYS:HB2	1.76	0.68
1:C:132:ILE:HB	1:C:142:ILE:HD12	1.76	0.67
2:B:85:GLN:HG3	2:B:154:LYS:HB3	1.77	0.66
1:C:460:ASN:HD22	2:D:288:ALA:HB2	1.63	0.64
2:B:84:THR:HG21	2:B:153:TRP:HZ2	1.64	0.63
1:A:175:ASN:HB3	1:A:178:ILE:HD13	1.80	0.63
1:C:90:VAL:HG21	1:C:157:PRO:HB2	1.83	0.60
2:D:157:PRO:HG2	2:D:184:MET:HA	1.83	0.60
1:C:441:TYR:CD2	1:C:544:GLY:HA3	2.36	0.60
2:B:85:GLN:O	2:B:89:GLU:N	2.34	0.60
1:A:58:THR:HG22	1:A:59:PRO:HD2	1.83	0.60
2:B:64:LYS:HD2	2:B:69:THR:O	2.01	0.60
2:D:348:ASN:N	2:D:348:ASN:OD1	2.34	0.60
1:A:356:ARG:NH1	1:A:359:GLY:O	2.34	0.59
2:B:10:VAL:HA	2:B:88:TRP:CH2	2.37	0.59
1:C:253:THR:HG23	1:C:256:ASP:H	1.67	0.59
2:D:194:GLU:O	2:D:198:HIS:N	2.32	0.59
1:A:390:LYS:HB3	1:A:417:VAL:HG21	1.86	0.57
2:B:7:THR:HG22	2:B:119:PRO:HG2	1.87	0.57
1:C:444:GLY:HA2	1:C:552:VAL:HG21	1.86	0.56
1:C:331:LYS:NZ	1:C:333:GLY:O	2.38	0.56
2:B:257:ILE:HB	2:B:283:LEU:HD21	1.87	0.56
1:C:89:GLU:HG2	1:C:92:LEU:HG	1.87	0.56
2:B:254:VAL:HG13	2:B:283:LEU:HD22	1.87	0.56
1:A:255:ASN:HB2	1:A:289:LEU:HB3	1.89	0.56
2:B:65:LYS:HA	2:B:407:GLN:OE1	2.06	0.55
1:A:23:GLN:HE22	1:A:133:PRO:HD3	1.70	0.55
1:C:445:ALA:HB3	1:C:553:SER:HB3	1.89	0.55
2:B:64:LYS:HD3	2:B:71:TRP:CE2	2.43	0.54
1:A:402:TRP:C	1:A:402:TRP:CD1	2.81	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:31:ILE:HG12	1:A:133:PRO:HG2	1.89	0.53
1:A:399:GLU:HA	1:A:402:TRP:CE3	2.44	0.53
1:A:241:VAL:HG12	1:A:242:GLN:O	2.09	0.53
1:C:343:GLN:HG3	1:C:349:LEU:HD21	1.91	0.52
1:A:442:VAL:HG12	1:A:457:TYR:HB3	1.91	0.52
2:B:369:THR:HG22	2:B:398:TRP:CH2	2.45	0.52
2:D:325:LEU:HD23	2:D:343:GLN:HG3	1.90	0.52
1:C:503:LEU:HD22	1:C:535:TRP:HB2	1.90	0.52
2:D:12:LEU:HD22	2:D:83:ARG:HB2	1.92	0.52
2:D:125:ARG:HD3	2:D:147:ASN:HA	1.91	0.52
1:A:101:LYS:HE3	1:A:321:PRO:HG3	1.90	0.51
2:D:60:VAL:HG12	2:D:75:VAL:HG22	1.92	0.51
2:B:54:ASN:O	2:B:143:ARG:NH2	2.44	0.51
2:B:21:VAL:HB	2:B:59:PRO:HD3	1.93	0.51
2:B:111:VAL:HG11	2:B:187:LEU:HD22	1.92	0.50
1:C:46:LYS:HD2	1:C:116:PHE:HB3	1.94	0.50
1:C:58:THR:HG22	1:C:59:PRO:HD2	1.93	0.50
1:A:110:ASP:O	1:A:217:PRO:HD3	2.12	0.49
2:D:375:ILE:HB	2:D:389:PHE:HZ	1.76	0.49
2:D:24[B]:TRP:HZ3	2:D:402:TRP:HE1	1.60	0.49
1:A:443:ASP:OD1	1:A:444:GLY:N	2.46	0.49
2:B:5:ILE:HD12	2:B:6:GLU:HG3	1.95	0.49
2:B:233:GLU:HG3	5:B:527:HOH:O	2.12	0.49
1:C:453:GLY:HA3	1:C:469:LEU:HB2	1.94	0.49
1:A:457:TYR:HE1	1:A:463:ARG:HG2	1.78	0.49
2:B:245:VAL:O	2:B:263:LYS:NZ	2.38	0.49
1:A:500:GLN:HG3	2:B:422:LEU:HD13	1.94	0.48
2:D:388:LYS:O	2:D:414:TRP:HA	2.14	0.48
1:C:114:ALA:HB1	1:C:160:PHE:CE1	2.48	0.47
1:A:492:GLU:HG2	1:A:530:LYS:HB2	1.95	0.47
1:C:180:ILE:HG23	1:C:189:VAL:HG22	1.95	0.47
1:A:491:LEU:HD13	1:A:529:GLU:OE2	2.14	0.47
1:A:31:ILE:O	1:A:35:VAL:HG23	2.15	0.47
2:B:187:LEU:HD12	2:B:187:LEU:HA	1.72	0.47
2:B:365:VAL:O	2:B:369:THR:HG23	2.15	0.47
1:A:405:TYR:CE2	1:A:407:GLN:HB2	2.50	0.47
1:C:167:ILE:O	1:C:170:PRO:HD2	2.15	0.47
1:C:458:VAL:HG23	1:C:548:VAL:HG13	1.96	0.47
1:A:390:LYS:HB3	1:A:417:VAL:CG2	2.45	0.47
2:B:120:LEU:HD13	2:B:149:LEU:HD23	1.97	0.47
2:D:64:LYS:HE2	2:D:71:TRP:CZ2	2.50	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:387:PRO:HD2	2:D:389:PHE:HE2	1.80	0.46
2:B:79:GLU:OE1	2:B:83:ARG:NH1	2.48	0.46
2:B:211:ARG:HB3	1:C:212:TRP:CZ2	2.51	0.46
2:B:331:LYS:HB2	2:B:337:TRP:CZ3	2.51	0.46
1:A:428:GLN:HE21	1:A:428:GLN:HA	1.80	0.46
2:B:303:LEU:O	2:B:307:ARG:HG3	2.15	0.46
2:B:325:LEU:HD23	2:B:343:GLN:HG3	1.97	0.46
2:D:50:ILE:HG21	2:D:145:GLN:HB3	1.96	0.46
2:D:238:LYS:HA	2:D:238:LYS:HD2	1.59	0.46
2:D:396:GLU:OE1	2:D:396:GLU:N	2.45	0.46
2:B:84:THR:HG21	2:B:153:TRP:CZ2	2.48	0.46
1:C:203:GLU:OE2	1:C:206:ARG:NH1	2.47	0.46
1:A:494:ASN:HB3	2:B:289:LEU:HD12	1.98	0.46
2:B:246:LEU:HD11	2:B:264:LEU:HD21	1.97	0.46
1:C:9:PRO:HA	1:C:121:ASP:OD2	2.15	0.46
2:D:171:PHE:HB2	2:D:208:HIS:ND1	2.31	0.46
2:B:61:PHE:HE1	2:B:76:ASP:HB2	1.81	0.45
2:B:394:GLN:NE2	5:B:507:HOH:O	2.49	0.45
1:A:60:VAL:CG2	1:A:130:PHE:HB2	2.46	0.45
1:A:473:THR:O	1:A:477:THR:HG23	2.16	0.45
2:D:305:GLU:O	2:D:309:ILE:HG12	2.16	0.45
1:A:134:SER:HB2	1:A:139:THR:HB	1.97	0.45
1:C:465:LYS:HE3	1:C:488:ASP:OD2	2.17	0.45
1:A:114:ALA:HB1	1:A:160:PHE:CE1	2.51	0.45
1:A:27:THR:O	1:A:31:ILE:HG13	2.17	0.45
1:A:273:GLY:O	1:A:275:LYS:NZ	2.44	0.45
1:A:487:GLN:HG2	1:A:524:GLN:NE2	2.32	0.45
1:C:296:THR:O	1:C:300:GLU:HG2	2.17	0.44
2:D:266:TRP:CG	2:D:425:LEU:HD13	2.52	0.44
2:D:24[A]:TRP:CD2	2:D:25:PRO:HD2	2.52	0.44
1:C:12:LEU:HD22	1:C:83:ARG:HB3	1.99	0.44
2:D:326:ILE:HG22	2:D:388:LYS:HB3	2.00	0.44
1:C:171:PHE:CZ	1:C:205:LEU:HB2	2.52	0.44
2:B:168:LEU:HD13	2:B:180:ILE:HG21	1.99	0.44
2:D:274:ILE:HD12	2:D:274:ILE:HA	1.90	0.44
1:A:272:PRO:HA	1:A:351:THR:HG21	1.98	0.44
1:A:406:TRP:CD1	2:B:420:PRO:HB3	2.53	0.43
1:A:429:LEU:HD11	1:A:506:ILE:HG22	2.00	0.43
1:C:320:ASP:OD2	1:C:323:LYS:NZ	2.51	0.43
2:D:326:ILE:HG12	2:D:342:TYR:O	2.18	0.43
1:A:23:GLN:NE2	1:A:60:VAL:HB	2.33	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:60:VAL:HG22	1:A:130:PHE:HB2	1.99	0.43
2:B:209:LEU:HB3	2:B:214:LEU:HB2	2.00	0.43
1:C:59:PRO:O	1:C:75:VAL:HG13	2.17	0.43
2:B:101:LYS:O	2:B:236:PRO:HB2	2.18	0.43
1:C:325:LEU:HD21	1:C:383:TRP:CE3	2.54	0.43
2:D:120:LEU:HD23	2:D:125:ARG:HG2	1.99	0.43
1:C:82:LYS:HB2	1:C:82:LYS:HE3	1.81	0.43
2:B:248:GLU:HB2	2:B:307:ARG:HH22	1.84	0.43
1:A:233:GLU:HB3	1:A:235:HIS:CE1	2.54	0.43
2:B:281:LYS:HA	2:B:284:ARG:HG3	2.01	0.43
2:B:341:ILE:HD11	2:B:375:ILE:HG23	2.01	0.43
1:C:473:THR:O	1:C:477:THR:HG23	2.18	0.43
2:B:64:LYS:HG3	2:B:71:TRP:HA	2.01	0.43
1:A:279:LEU:HG	1:A:302:GLU:OE1	2.19	0.42
2:D:234:LEU:O	2:D:236:PRO:HD3	2.19	0.42
1:C:221:HIS:CD2	1:C:221:HIS:H	2.35	0.42
2:D:24[A]:TRP:CG	2:D:25:PRO:HD2	2.53	0.42
1:A:88:TRP:CE2	2:B:143:ARG:HD3	2.53	0.42
1:A:408:ALA:O	2:B:393:ILE:HG13	2.19	0.42
2:B:248:GLU:HB2	2:B:307:ARG:NH2	2.35	0.42
1:A:198:HIS:NE2	1:A:202:ILE:HD11	2.35	0.42
1:A:343:GLN:HG3	1:A:349:LEU:HD11	2.01	0.42
2:D:354:TYR:HB2	2:D:374:LYS:HE3	2.00	0.42
1:C:229:TRP:CD2	4:C:602:29T:H2	2.55	0.42
1:C:239:TRP:NE1	1:C:316:GLY:HA3	2.34	0.42
2:B:282:LEU:HD12	2:B:293:ILE:CG2	2.49	0.42
1:C:471:ASN:O	1:C:471:ASN:ND2	2.52	0.42
2:B:401:TRP:HD1	2:B:404[A]:GLU:HG3	1.84	0.42
2:D:122:GLU:HA	2:D:125:ARG:NE	2.35	0.42
2:D:375:ILE:HB	2:D:389:PHE:CZ	2.54	0.42
2:B:122:GLU:HA	2:B:125:ARG:HG3	2.02	0.42
2:B:326:ILE:HG12	2:B:342:TYR:CE2	2.55	0.42
1:C:90:VAL:HG13	1:C:161:GLN:HG3	2.02	0.42
1:C:165:THR:O	1:C:169:GLU:HG3	2.20	0.42
1:C:5:ILE:HD13	1:C:5:ILE:HA	1.95	0.42
1:C:238:LYS:HE3	1:C:238:LYS:HB2	1.83	0.42
1:A:269:GLN:O	1:A:351:THR:N	2.48	0.41
1:A:304:ALA:O	1:A:308[A]:GLU:HG2	2.20	0.41
1:A:361:HIS:CE1	1:A:513:SER:HG	2.38	0.41
2:D:241:VAL:HG12	2:D:350:LYS:HG2	2.01	0.41
1:A:19:PRO:HG3	1:A:80:LEU:HB2	2.00	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:478:GLU:O	1:C:482:ILE:HG12	2.19	0.41
2:D:12:LEU:HD23	2:D:12:LEU:HA	1.90	0.41
1:C:261:VAL:HG13	1:C:276:VAL:HB	2.01	0.41
1:C:122:GLU:HA	1:C:125:ARG:HD2	2.03	0.41
2:D:106:VAL:HA	2:D:190:GLY:HA2	2.03	0.41
1:A:246:LEU:CB	1:A:307:ARG:HE	2.34	0.41
1:C:58:THR:N	1:C:129:ALA:O	2.54	0.41
1:C:350:LYS:HE3	1:C:378:GLU:OE2	2.20	0.41
2:D:85:GLN:NE2	2:D:154:LYS:HB2	2.36	0.41
1:A:457:TYR:OH	1:A:488:ASP:OD2	2.31	0.41
2:D:386:THR:HA	2:D:387:PRO:HD3	1.86	0.41
2:B:53:GLU:OE1	2:B:53:GLU:N	2.45	0.41
2:B:301:LEU:HD22	2:B:301:LEU:HA	1.95	0.41
1:A:362:THR:HG22	1:A:363:ASN:H	1.86	0.40
2:B:105:SER:O	2:B:190:GLY:HA2	2.21	0.40
1:C:198:HIS:O	1:C:202:ILE:HG12	2.21	0.40
2:D:257:ILE:O	2:D:261:VAL:HG23	2.21	0.40
1:C:460:ASN:ND2	2:D:288:ALA:HB2	2.33	0.40
2:D:24[B]:TRP:CH2	2:D:399:GLU:HG2	2.56	0.40
2:D:263:LYS:HG2	2:D:425:LEU:CD2	2.50	0.40
1:A:77:PHE:O	1:A:81:ASN:N	2.40	0.40
1:A:458:VAL:HG22	1:A:464:GLN:HG2	2.03	0.40
2:B:236:PRO:HA	2:B:239:TRP:CD2	2.56	0.40
1:A:406:TRP:CG	2:B:420:PRO:HB3	2.57	0.40
2:B:282:LEU:HD11	2:B:296:THR:HG23	2.02	0.40
2:D:42:GLU:HA	2:D:47:ILE:O	2.21	0.40

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	544/558 (98%)	533 (98%)	11 (2%)	0	100	100
1	C	535/558 (96%)	525 (98%)	10 (2%)	0	100	100
2	B	399/428 (93%)	392 (98%)	7 (2%)	0	100	100
2	D	396/428 (92%)	387 (98%)	9 (2%)	0	100	100
All	All	1874/1972 (95%)	1837 (98%)	37 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	421/495 (85%)	402 (96%)	19 (4%)	27	50
1	C	431/495 (87%)	390 (90%)	41 (10%)	8	15
2	B	351/390 (90%)	338 (96%)	13 (4%)	34	58
2	D	300/390 (77%)	276 (92%)	24 (8%)	12	23
All	All	1503/1770 (85%)	1406 (94%)	97 (6%)	17	33

All (97) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	24	TRP
1	A	30	LYS
1	A	58	THR
1	A	68	SER
1	A	94	ILE
1	A	108	VAL
1	A	220	LYS
1	A	223	LYS
1	A	228	LEU
1	A	244	ILE
1	A	245	VAL
1	A	290	THR
1	A	350	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	362	THR
1	A	402	TRP
1	A	428	GLN
1	A	452	LEU
1	A	473	THR
1	A	551	LEU
2	B	22	LYS
2	B	61	PHE
2	B	64	LYS
2	B	68	SER
2	B	69	THR
2	B	86	ASP
2	B	88	TRP
2	B	89	GLU
2	B	92	LEU
2	B	301	LEU
2	B	326	ILE
2	B	338	THR
2	B	417	VAL
1	C	5	ILE
1	C	20	LYS
1	C	57	ASN
1	C	58	THR
1	C	73	LYS
1	C	101	LYS
1	C	103	LYS
1	C	110[A]	ASP
1	C	110[B]	ASP
1	C	113	ASP
1	C	132	ILE
1	C	146	TYR
1	C	183	TYR
1	C	185	ASP
1	C	199	ARG
1	C	216	THR
1	C	220	LYS
1	C	221	HIS
1	C	223	LYS
1	C	228	LEU
1	C	238	LYS
1	C	250	ASP
1	C	255	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	286	THR
1	C	295	LEU
1	C	351	THR
1	C	361	HIS
1	C	363	ASN
1	C	396	GLU
1	C	402	TRP
1	C	405	TYR
1	C	442	VAL
1	C	443	ASP
1	C	449	GLU
1	C	459	THR
1	C	463	ARG
1	C	465	LYS
1	C	509	GLN
1	C	547	GLN
1	C	548	VAL
1	C	549	ASP
2	D	5	ILE
2	D	6	GLU
2	D	21	VAL
2	D	72	ARG
2	D	85	GLN
2	D	91	GLN
2	D	121	ASP
2	D	212	TRP
2	D	232	TYR
2	D	238	LYS
2	D	246	LEU
2	D	249	LYS
2	D	286	THR
2	D	291	GLU
2	D	305	GLU
2	D	315	HIS
2	D	330	GLN
2	D	348	ASN
2	D	354	TYR
2	D	362	THR
2	D	374	LYS
2	D	405	TYR
2	D	414	TRP
2	D	425	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	235	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

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	$\#\|Z\| > 2$
4	29T	A	602	-	30,32,32	0.47	0	35,44,44	0.67	0
4	29T	C	602	-	30,32,32	0.75	1 (3%)	35,44,44	0.62	0

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
4	29T	A	602	-	-	0/10/12/12	0/4/4/4
4	29T	C	602	-	-	1/10/12/12	0/4/4/4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	602	29T	C11-C14	2.82	1.40	1.38

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	602	29T	C05-C04-O0A-C0C

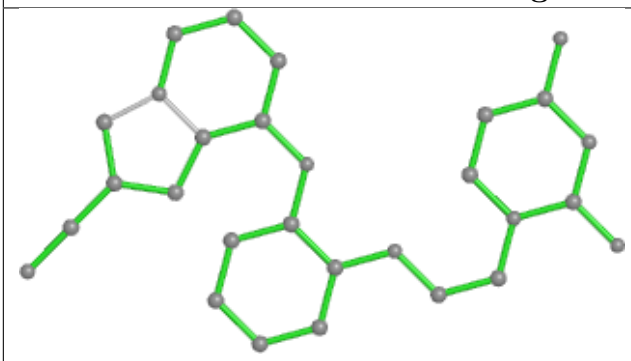
There are no ring outliers.

1 monomer is involved in 1 short contact:

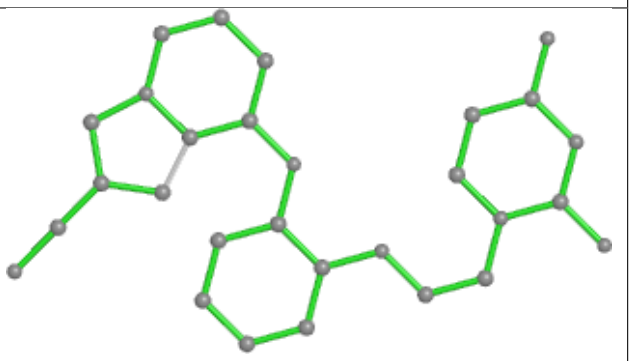
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	602	29T	1	0

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.

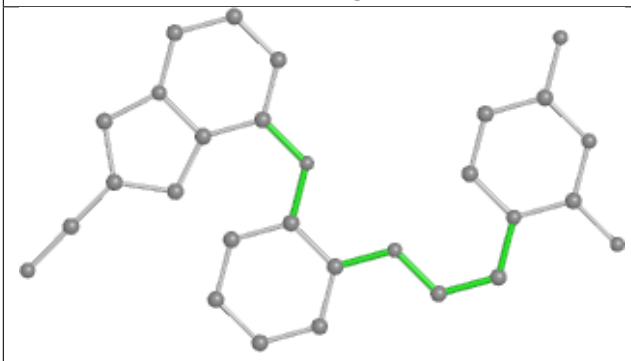
Ligand 29T A 602



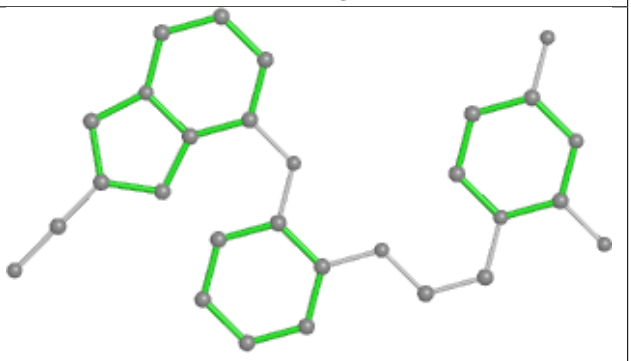
Bond lengths



Bond angles

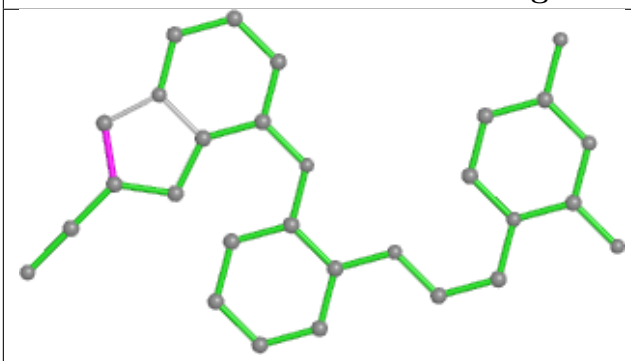


Torsions

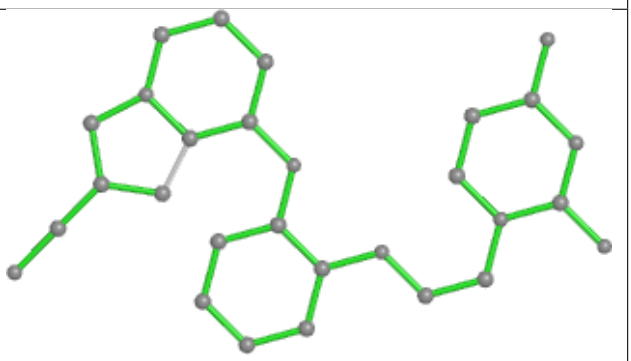


Rings

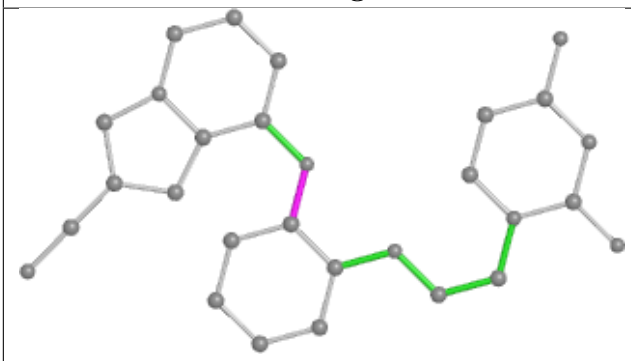
Ligand 29T C 602



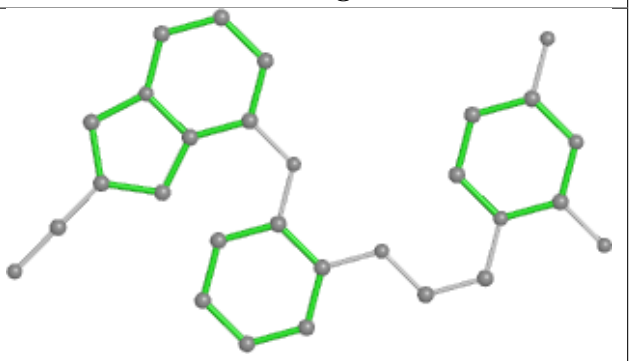
Bond lengths



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

6.1 Protein, DNA and RNA chains

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, 95^{th} 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	547/558 (98%)	0.23	24 (4%) 34 28	39, 76, 139, 160	0
1	C	544/558 (97%)	0.58	46 (8%) 10 7	68, 100, 134, 159	0
2	B	405/428 (94%)	0.18	9 (2%) 62 57	43, 69, 105, 134	7 (1%)
2	D	401/428 (93%)	0.68	43 (10%) 6 4	76, 114, 141, 155	5 (1%)
All	All	1897/1972 (96%)	0.42	122 (6%) 19 15	39, 92, 137, 160	12 (0%)

All (122) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	245	VAL	4.9
1	A	225	PRO	4.9
1	C	246	LEU	4.8
1	C	556	ILE	4.7
2	D	212	TRP	4.5
1	C	286	THR	4.4
1	A	223	LYS	4.4
1	C	26	LEU	4.3
1	C	290	THR	4.3
1	C	276	VAL	4.2
1	C	225	PRO	4.1
2	D	24[A]	TRP	4.1
2	B	362	THR	4.1
2	D	272	PRO	4.0
1	A	294	PRO	3.9
1	C	244	ILE	3.9
2	D	174	GLN	3.7
1	C	297	GLU	3.7
1	C	131	THR	3.6
2	D	303	LEU	3.6
1	A	282	LEU	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	245	VAL	3.6
2	D	67	ASP	3.5
1	C	21	VAL	3.5
2	D	168	LEU	3.4
2	D	10	VAL	3.4
2	D	116	PHE	3.4
1	C	260	LEU	3.3
2	D	214	LEU	3.3
2	D	299	ALA	3.3
1	C	132	ILE	3.3
1	C	402	TRP	3.3
1	A	290	THR	3.2
1	C	75	VAL	3.2
1	A	71	TRP	3.2
2	B	361	HIS	3.2
1	C	304	ALA	3.2
1	A	303	LEU	3.2
1	C	136	ASN	3.2
2	D	95	PRO	3.2
1	A	246	LEU	3.2
2	D	5	ILE	3.1
1	C	61	PHE	3.1
1	C	128	THR	3.1
1	C	142	ILE	3.1
1	C	252	TRP	3.0
2	D	252	TRP	3.0
1	C	359	GLY	3.0
1	C	299	ALA	3.0
2	D	304	ALA	3.0
2	D	238	LYS	3.0
1	C	292	VAL	3.0
2	D	115	TYR	2.9
2	D	251	SER	2.9
1	A	299	ALA	2.9
2	D	306	ASN	2.9
1	A	300	GLU	2.9
1	A	452	LEU	2.9
1	C	130	PHE	2.8
1	A	253	THR	2.8
1	C	15	GLY	2.8
1	A	297	GLU	2.8
2	D	96	HIS	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	427	TYR	2.8
1	A	24	TRP	2.8
1	A	279	LEU	2.8
2	D	109	LEU	2.8
2	D	319	TYR	2.8
1	A	542	ILE	2.8
2	D	111	VAL	2.7
2	D	308	GLU	2.7
1	A	69	THR	2.7
1	A	252	TRP	2.7
1	C	309	ILE	2.7
1	C	223	LYS	2.7
2	B	14	PRO	2.6
1	C	283	LEU	2.6
1	C	310	LEU	2.6
1	C	58	THR	2.5
1	A	292	VAL	2.5
1	C	34	LEU	2.5
2	D	317	VAL	2.5
2	B	284	ARG	2.5
1	A	295	LEU	2.5
1	C	221	HIS	2.5
1	C	332[A]	GLN	2.4
1	C	20	LYS	2.4
1	C	554	ALA	2.4
1	A	260	LEU	2.4
2	D	286	THR	2.4
2	B	88	TRP	2.4
2	D	189	VAL	2.4
2	D	248	GLU	2.4
1	C	80	LEU	2.4
2	B	86	ASP	2.4
2	D	247	PRO	2.3
2	D	92	LEU	2.3
1	C	90	VAL	2.3
2	D	159	ILE	2.3
1	A	146	TYR	2.3
1	C	358	ARG	2.3
2	D	325	LEU	2.3
1	C	59	PRO	2.2
1	C	296	THR	2.2
1	C	282	LEU	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	366	LYS	2.2
1	C	275	LYS	2.2
2	B	85	GLN	2.2
2	D	296	THR	2.2
2	D	202	ILE	2.1
2	D	273	GLY	2.1
2	B	357	MET	2.1
2	D	154	LYS	2.1
2	D	101	LYS	2.1
2	D	341	ILE	2.1
1	A	302	GLU	2.1
2	D	97	PRO	2.1
1	C	291	GLU	2.1
2	B	283	LEU	2.0
1	A	298	GLU	2.0
1	C	16	MET	2.0
2	D	188	TYR	2.0

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, 95^{th} 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.

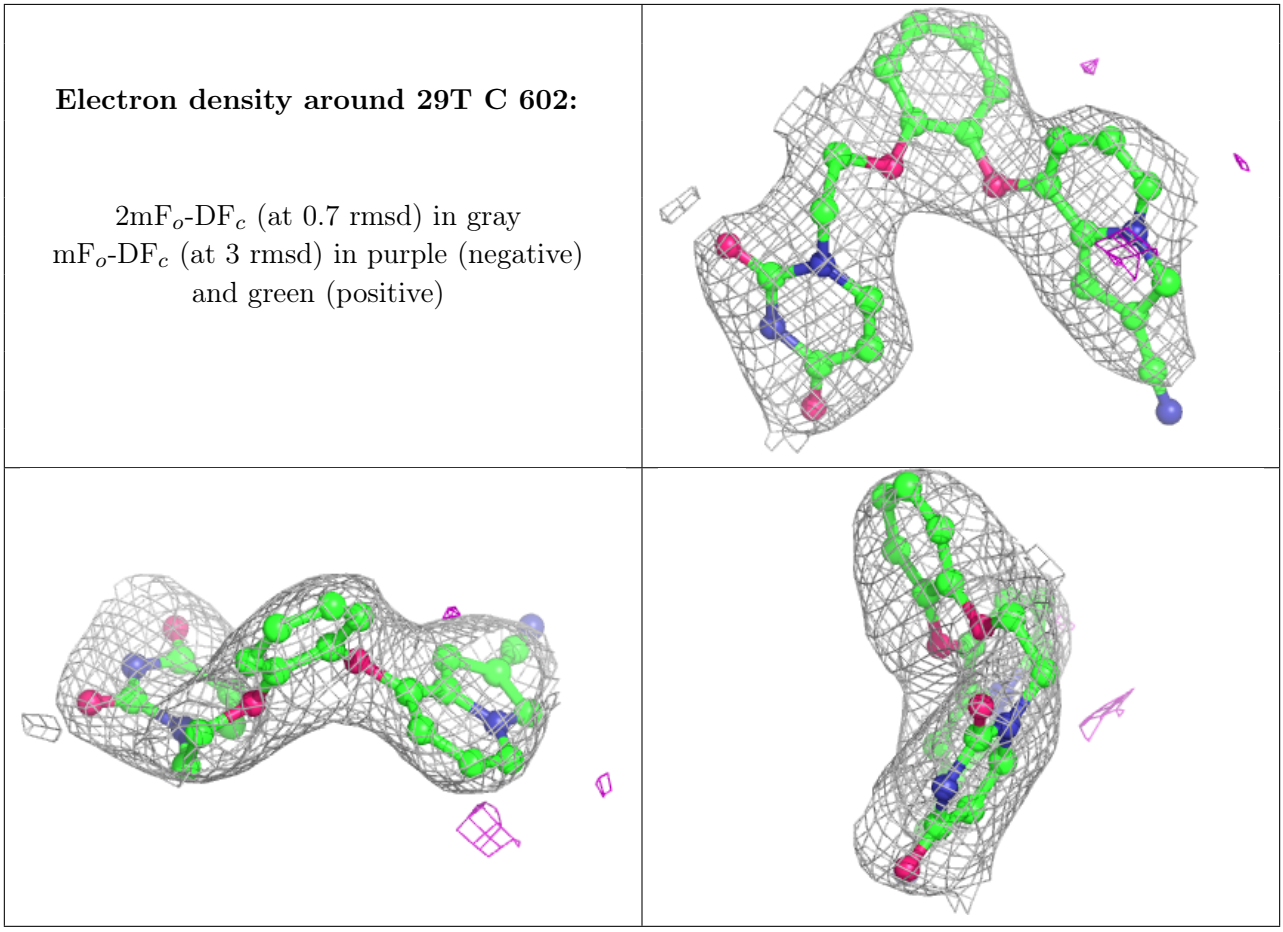
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	MG	C	601	1/1	0.85	0.25	86,86,86,86	0
3	MG	A	601	1/1	0.90	0.23	85,85,85,85	0
4	29T	C	602	29/29	0.92	0.26	75,82,95,98	0
4	29T	A	602	29/29	0.97	0.18	48,58,66,69	0

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.

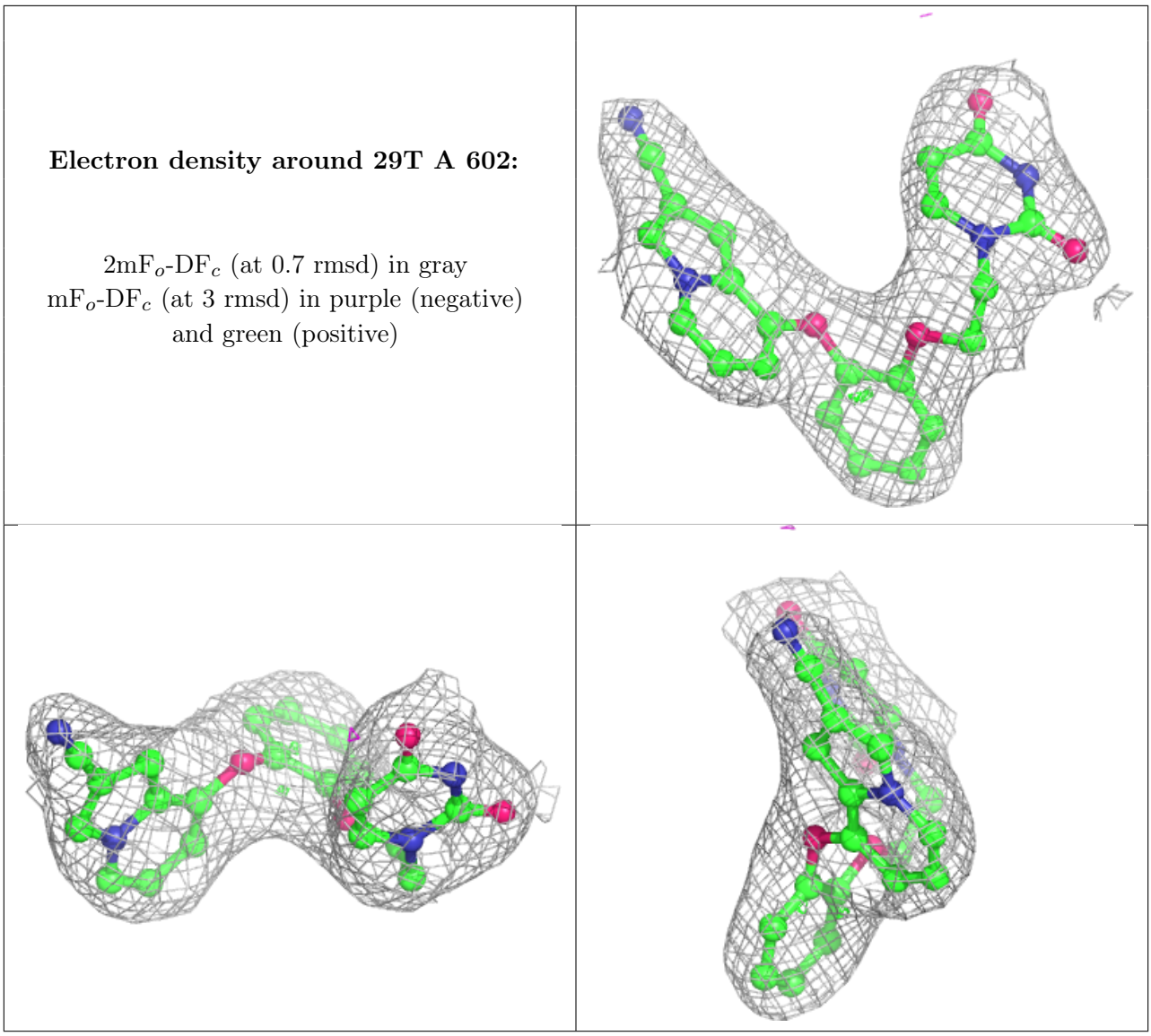
Electron density around 29T C 602:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



Electron density around 29T A 602:

$2mF_o$-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.