Full wwPDB X-ray Structure Validation Report ⓘ

Nov 22, 2023 – 06:09 PM EST

PDB ID	:	8STV
Title	:	Crystal Structure of HIV-1 Reverse Transcriptase (Y181C, V106A) variant in Complex with 5-(2-(2-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)ethoxy)phenoxy)-2-naphthonitrile (JLJ600), a non-nucleoside inhibitor
Authors	:	Hollander, K.; Jorgensen, W.L.; Anderson, K.S.
Deposited on	:	2023-05-11
Resolution	:	2.78 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:

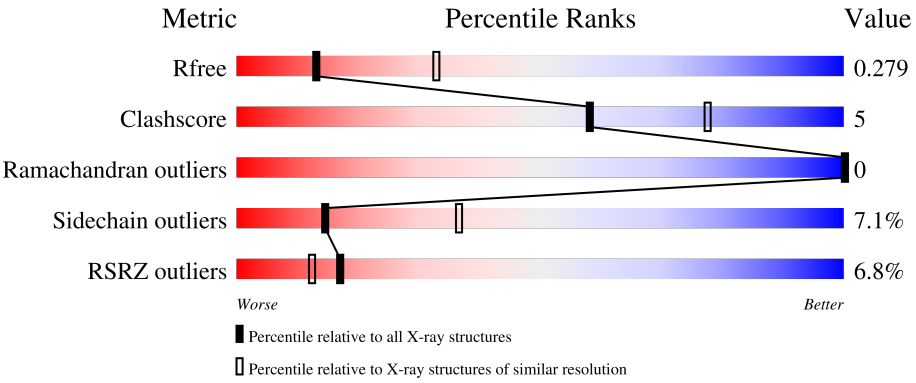
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.78 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	4107 (2.80-2.76)
Clashscore	141614	4575 (2.80-2.76)
Ramachandran outliers	138981	4487 (2.80-2.76)
Sidechain outliers	138945	4489 (2.80-2.76)
RSRZ outliers	127900	4027 (2.80-2.76)

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	5% 80% 16%
1	C	558	8% 79% 18%
2	B	428	3% 83% 11% 5%
2	D	428	11% 80% 13% 6%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 14571 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	544	Total	C	N	O	S	0	1	0
			4195	2703	695	790	7			
1	C	542	Total	C	N	O	S	0	2	0
			4072	2609	674	782	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

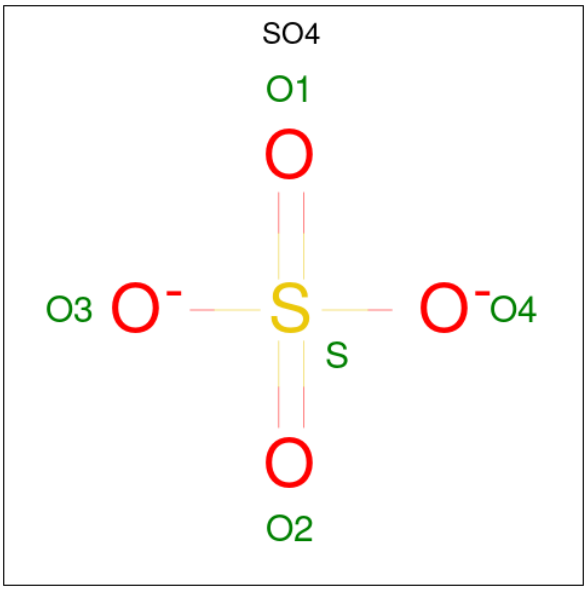
- Molecule 2 is a protein called p51 RT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	408	Total	C	N	O	S	19	1	0
			3285	2137	532	611	5			
2	D	401	Total	C	N	O	S	10	1	0
			2869	1838	488	540	3			

There are 2 discrepancies between the modelled and reference sequences:

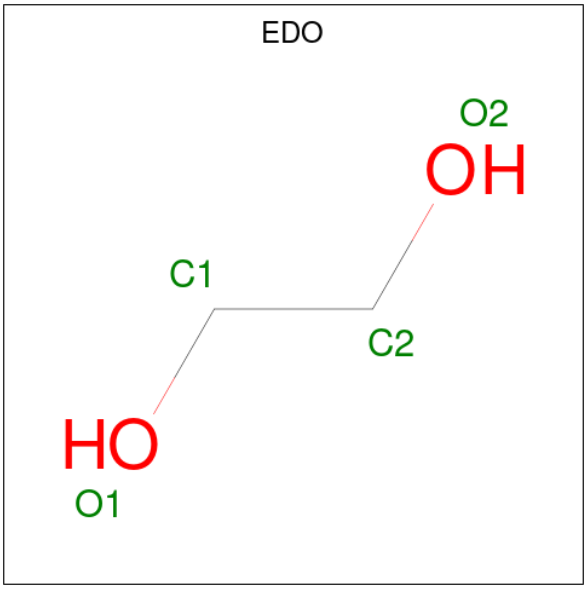
Chain	Residue	Modelled	Actual	Comment	Reference
B	280	SER	CYS	engineered mutation	UNP P03366
D	280	SER	CYS	engineered mutation	UNP P03366

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



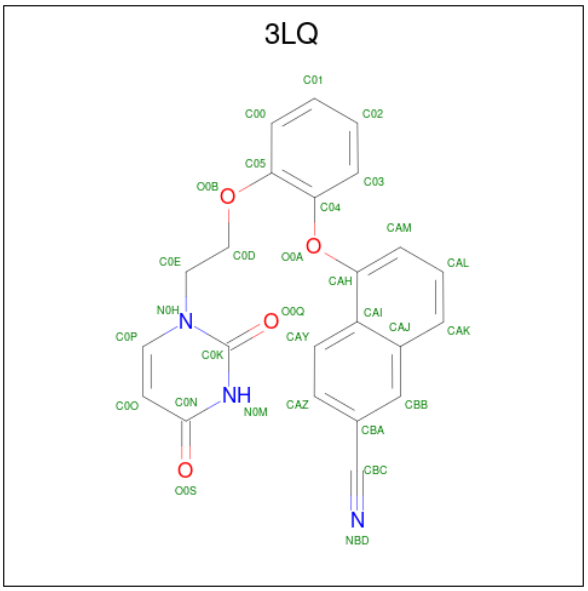
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	C	1	Total C O 4 2 2	0	0

- Molecule 5 is 5-{2-[2-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)ethoxy]phenoxy}naphthalene-2-carbonitrile (three-letter code: 3LQ) (formula: $C_{23}H_{17}N_3O_4$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C N O 30 23 3 4	0	0
5	C	1	Total C N O 30 23 3 4	0	0

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	1	Total Mg 1 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	31	Total O 31 31	0	0

Continued on next page...

Continued from previous page...

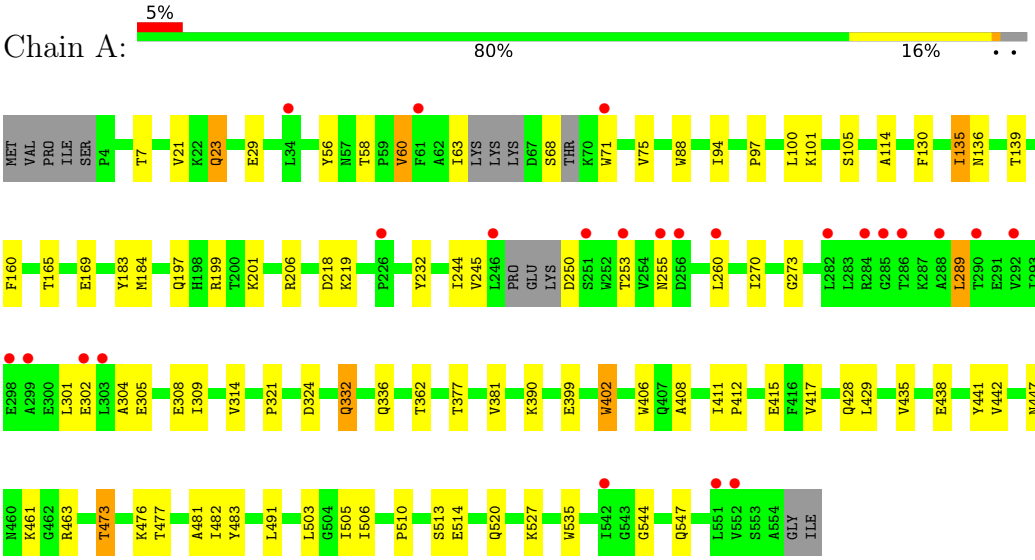
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	29	Total O 29 29	0	0
7	C	3	Total O 3 3	0	0
7	D	8	Total O 8 8	0	0

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.

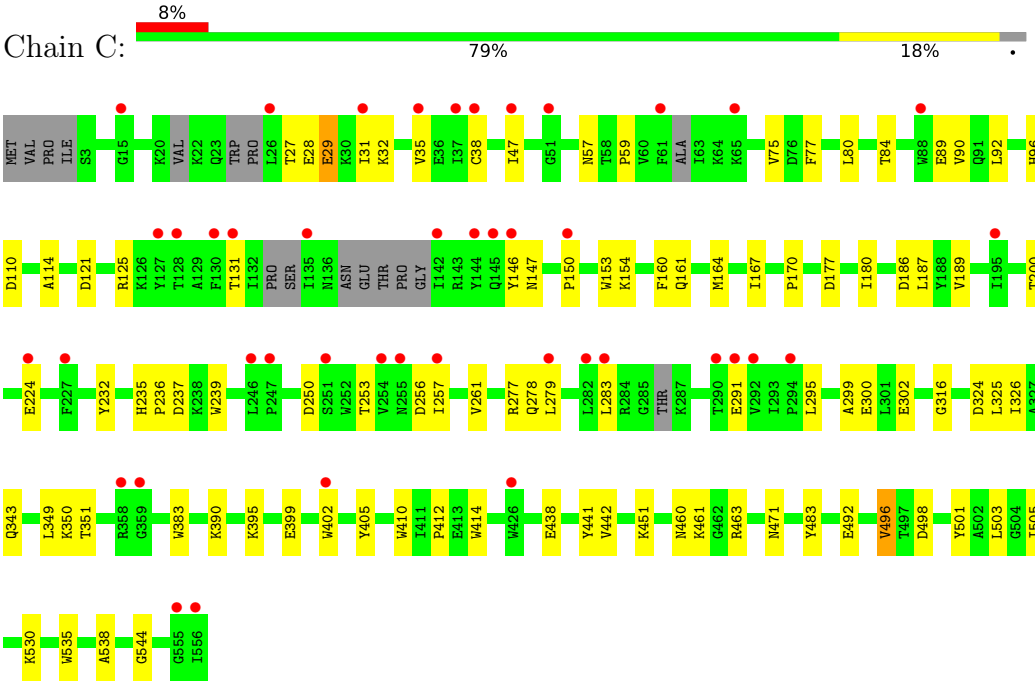
• Molecule 1: Reverse transcriptase/ribonuclease H

Chain A:



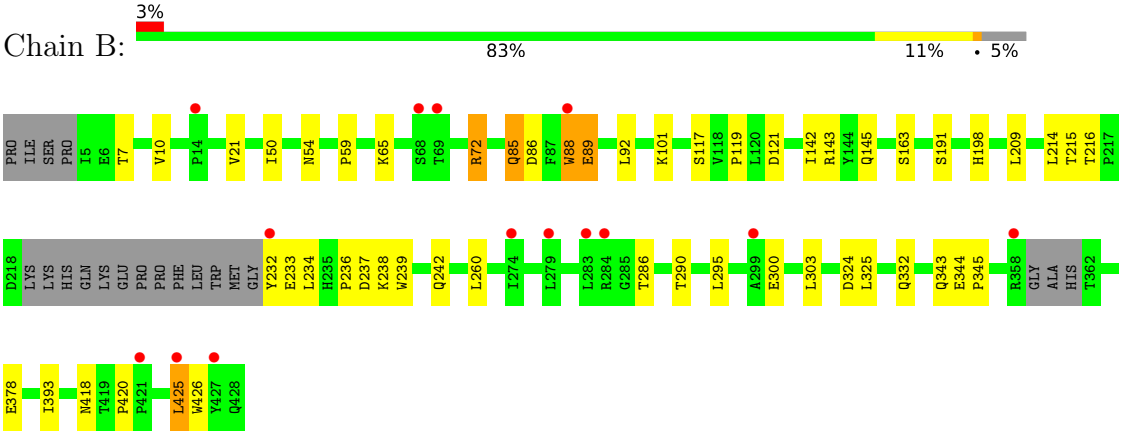
• Molecule 1: Reverse transcriptase/ribonuclease H

Chain C:



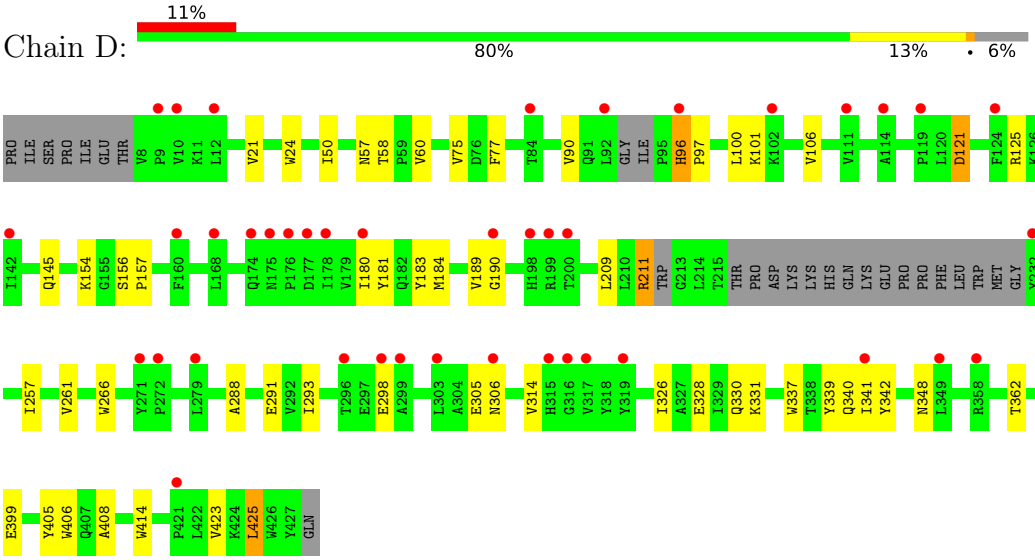
• Molecule 2: p51 RT

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	111.59Å 73.22Å 170.66Å 90.00° 97.73° 90.00°	Depositor
Resolution (Å)	37.83 – 2.78 37.83 – 2.78	Depositor EDS
% Data completeness (in resolution range)	99.8 (37.83-2.78) 99.8 (37.83-2.78)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.26 (at 2.77Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.232 , 0.276 0.235 , 0.279	Depositor DCC
R_{free} test set	3390 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å^2)	86.8	Xtriage
Anisotropy	0.049	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 79.4	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	0.94	EDS
Total number of atoms	14571	wwPDB-VP
Average B, all atoms (Å^2)	99.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.09% 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: SO4, MG, EDO, 3LQ

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	0.25	0/4306	0.45	0/5884
1	C	0.25	0/4175	0.43	0/5710
2	B	0.25	0/3381	0.45	0/4617
2	D	0.26	0/2949	0.45	0/4065
All	All	0.25	0/14811	0.45	0/20276

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	4195	0	3982	45	0
1	C	4072	0	3724	43	0
2	B	3285	0	3200	24	0
2	D	2869	0	2419	25	0
3	A	10	0	0	0	0
4	A	4	0	6	0	0
4	C	4	0	6	1	0
5	A	30	0	17	1	0
5	C	30	0	17	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	C	1	0	0	0	0
7	A	31	0	0	1	0
7	B	29	0	0	0	0
7	C	3	0	0	0	0
7	D	8	0	0	0	0
All	All	14571	0	13371	132	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 5.

All (132) 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:451:LYS:HA	1:C:471:ASN:H	1.64	0.63
1:C:328:GLU:HG2	1:C:390:LYS:HB2	1.79	0.63
1:A:135:ILE:HD12	1:A:136:ASN:H	1.63	0.62
2:B:209:LEU:HB3	2:B:214:LEU:HB2	1.82	0.61
1:A:399:GLU:HA	1:A:402:TRP:CE3	2.39	0.57
1:A:438:GLU:OE2	1:A:459:THR:HG21	2.04	0.55
1:C:235:HIS:HB3	1:C:236:PRO:HD2	1.88	0.55
2:B:85:GLN:O	2:B:89:GLU:N	2.39	0.55
1:C:114:ALA:HB1	1:C:160:PHE:CE1	2.42	0.54
1:C:343:GLN:HG3	1:C:349:LEU:HD11	1.90	0.53
1:A:23:GLN:HE21	1:A:60:VAL:H	1.57	0.53
1:A:459:THR:HG23	1:A:461:LYS:H	1.73	0.53
1:A:429:LEU:HD11	1:A:506:ILE:HG22	1.92	0.52
2:B:10:VAL:HA	2:B:88:TRP:CH2	2.44	0.52
1:A:332:GLN:O	1:A:336:GLN:HB2	2.10	0.52
1:C:89:GLU:HG2	1:C:92:LEU:HG	1.92	0.52
2:D:50:ILE:HG21	2:D:145:GLN:HB3	1.92	0.52
1:C:460:ASN:HD22	2:D:288:ALA:HB2	1.74	0.51
1:A:60:VAL:HG13	1:A:75:VAL:HG22	1.93	0.51
1:A:56:TYR:O	1:A:143:ARG:NH2	2.44	0.51
2:B:7:THR:HG22	2:B:119:PRO:HG2	1.93	0.51
2:B:232:TYR:CG	2:B:233:GLU:N	2.78	0.51
1:C:150:PRO:HB2	1:C:153:TRP:HB2	1.92	0.51
2:D:180:ILE:HD13	2:D:189:VAL:HA	1.92	0.51
2:D:406:TRP:NE1	2:D:408:ALA:HB3	2.26	0.51
1:C:498:ASP:HB2	1:C:538:ALA:HB2	1.93	0.51
1:A:459:THR:HG22	1:A:463:ARG:H	1.76	0.50
1:C:438:GLU:OE2	1:C:463:ARG:HD3	2.12	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:441:TYR:CD2	1:C:544:GLY:HA3	2.46	0.50
2:D:156:SER:HB2	2:D:157:PRO:HD3	1.94	0.50
2:B:21:VAL:HB	2:B:59:PRO:HD3	1.94	0.49
1:A:101:LYS:HE3	1:A:321:PRO:HG3	1.94	0.49
1:A:273:GLY:N	7:A:703:HOH:O	2.44	0.49
1:C:239:TRP:O	1:C:316:GLY:N	2.45	0.49
2:D:106:VAL:HA	2:D:190:GLY:HA2	1.94	0.49
1:A:114:ALA:HB1	1:A:160:PHE:CE1	2.47	0.49
1:A:206:ARG:CZ	1:A:218:ASP:HB2	2.43	0.49
2:D:209:LEU:C	2:D:211:ARG:N	2.63	0.49
1:A:441:TYR:CD2	1:A:544:GLY:HA3	2.48	0.48
1:A:503:LEU:HD22	1:A:535:TRP:HB2	1.95	0.48
1:A:100:LEU:HB3	5:A:604:3LQ:H6	1.96	0.48
2:B:191:SER:OG	2:B:198:HIS:ND1	2.40	0.47
1:A:402:TRP:C	1:A:402:TRP:CD1	2.87	0.47
1:C:57:ASN:HD21	1:C:131:THR:HB	1.79	0.47
2:D:60:VAL:HG12	2:D:75:VAL:HG22	1.96	0.47
1:A:447:ASN:HB3	1:A:450:THR:OG1	2.14	0.47
2:B:54:ASN:O	2:B:143:ARG:NH2	2.47	0.47
1:C:253:THR:HG23	1:C:256:ASP:H	1.80	0.47
1:C:59:PRO:O	1:C:75:VAL:HG13	2.15	0.47
1:C:77:PHE:HB3	1:C:80:LEU:HB3	1.97	0.47
1:C:89:GLU:OE1	1:C:90:VAL:N	2.47	0.47
2:D:96:HIS:N	2:D:97:PRO:HD2	2.30	0.47
2:D:266:TRP:CD2	2:D:425:LEU:HD13	2.50	0.46
1:A:114:ALA:HB1	1:A:160:PHE:CZ	2.50	0.46
2:B:295:LEU:HB3	2:B:300:GLU:HG2	1.97	0.46
2:B:332:GLN:HE22	2:B:425:LEU:HB2	1.80	0.46
1:C:164:MET:SD	1:C:187:LEU:HD21	2.55	0.46
1:A:505:ILE:O	1:A:510:PRO:HD3	2.16	0.46
2:D:154:LYS:O	2:D:157:PRO:HD2	2.17	0.45
1:A:428:GLN:HE21	1:A:428:GLN:HB3	1.63	0.44
1:C:517:LEU:H	4:C:602:EDO:H22	1.81	0.44
2:D:58:THR:HG21	2:D:77:PHE:CD1	2.53	0.44
2:B:50:ILE:HG21	2:B:145:GLN:HB3	2.00	0.44
1:C:257:ILE:O	1:C:261:VAL:HG23	2.17	0.44
1:C:325:LEU:HD21	1:C:383:TRP:CE3	2.53	0.44
1:A:305:GLU:O	1:A:309:ILE:HG13	2.18	0.44
1:A:473:THR:O	1:A:477:THR:HG23	2.18	0.44
1:C:278:GLN:H	1:C:302:GLU:CD	2.21	0.44
1:C:501:TYR:CZ	1:C:505:ILE:HD11	2.53	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:503:LEU:HD22	1:C:535:TRP:HB2	1.99	0.44
1:A:88:TRP:CZ2	2:B:143:ARG:HD3	2.53	0.44
1:A:408:ALA:O	2:B:393:ILE:HG13	2.18	0.44
1:A:60:VAL:HG22	1:A:130:PHE:HB2	2.00	0.43
1:A:304:ALA:O	1:A:308[A]:GLU:HG2	2.16	0.43
1:A:411:ILE:HG22	1:A:412:PRO:O	2.18	0.43
1:A:442:VAL:HB	1:A:481:ALA:HB1	2.01	0.43
2:B:260:LEU:HD21	2:B:303:LEU:HD13	1.99	0.43
2:B:324:ASP:O	2:B:343:GLN:HG2	2.18	0.43
1:C:97:PRO:HG2	1:C:232:TYR:CD1	2.53	0.43
2:B:344:GLU:HB3	2:B:345:PRO:HD2	2.01	0.43
1:C:180:ILE:HG12	1:C:189:VAL:HG22	1.99	0.43
1:C:483:TYR:HB2	1:C:521:ILE:HG12	2.00	0.43
2:D:24:TRP:CH2	2:D:399:GLU:HG2	2.54	0.43
1:C:492:GLU:OE2	1:C:530:LYS:HD3	2.19	0.43
2:D:331:LYS:HB2	2:D:337:TRP:CZ3	2.53	0.43
1:C:38:CYS:HA	1:C:47:ILE:HD11	2.01	0.43
2:B:101:LYS:O	2:B:236:PRO:HB2	2.19	0.42
1:A:301:LEU:O	1:A:305:GLU:HG3	2.18	0.42
1:C:96:HIS:CE1	1:C:350:LYS:HD2	2.54	0.42
1:A:165:THR:O	1:A:169:GLU:HG3	2.20	0.42
1:C:29:GLU:H	1:C:29:GLU:HG3	1.43	0.42
1:A:255:ASN:HB2	1:A:289:LEU:HB3	2.01	0.42
1:A:406:TRP:NE1	2:B:420:PRO:HG3	2.35	0.42
2:B:325:LEU:HD23	2:B:343:GLN:HG3	2.00	0.42
1:C:501:TYR:CE2	1:C:505:ILE:HD11	2.54	0.42
1:A:270:ILE:HG13	1:A:314:VAL:HG12	2.02	0.42
1:C:167:ILE:O	1:C:170:PRO:HD2	2.19	0.42
2:D:135:ILE:HD13	2:D:135:ILE:HA	1.84	0.42
1:C:410:TRP:CZ2	1:C:412:PRO:HA	2.55	0.42
2:D:342:TYR:HB3	2:D:348:ASN:HA	2.02	0.42
1:A:97:PRO:HG2	1:A:232:TYR:CD1	2.55	0.41
2:D:157:PRO:HG2	2:D:184:MET:HA	2.01	0.41
1:A:135:ILE:H	1:A:135:ILE:HG13	1.49	0.41
2:D:305:GLU:HG2	2:D:306:ASN:N	2.36	0.41
1:C:84:THR:HB	1:C:154:LYS:HD3	2.02	0.41
1:C:121:ASP:O	1:C:125:ARG:HG3	2.21	0.41
1:A:197:GLN:O	1:A:201:LYS:HG2	2.20	0.41
2:D:181:TYR:CZ	2:D:183:TYR:HB2	2.55	0.41
2:D:209:LEU:C	2:D:211:ARG:H	2.24	0.41
1:A:94:ILE:HG22	1:A:183:TYR:CE1	2.55	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:390:LYS:HB3	1:A:417:VAL:HG21	2.02	0.41
1:A:476:LYS:HE3	1:A:476:LYS:HB2	1.82	0.41
1:A:483:TYR:CE1	1:A:520:GLN:HB3	2.56	0.41
2:B:237:ASP:OD1	2:B:238:LYS:N	2.54	0.41
1:A:377:THR:O	1:A:381:VAL:HG23	2.21	0.41
2:B:65:LYS:HD2	2:B:72:ARG:HD2	2.03	0.41
2:B:239:TRP:CH2	2:B:378:GLU:HG2	2.56	0.41
1:C:279:LEU:HD13	1:C:299:ALA:HB1	2.02	0.41
1:C:438:GLU:HG3	1:C:461:LYS:HG2	2.03	0.41
1:C:441:TYR:HA	1:C:496:VAL:HG22	2.02	0.41
2:D:328:GLU:O	2:D:339:TYR:HA	2.21	0.41
2:D:121:ASP:O	2:D:125:ARG:HG3	2.21	0.40
2:D:326:ILE:O	2:D:341:ILE:HA	2.22	0.40
1:A:60:VAL:CG2	1:A:130:PHE:HB2	2.51	0.40
2:D:21:VAL:O	2:D:57:ASN:ND2	2.54	0.40
2:D:257:ILE:O	2:D:261:VAL:HG23	2.21	0.40
1:C:295:LEU:HB3	1:C:300:GLU:OE1	2.22	0.40
1:A:435:VAL:HA	2:B:290:THR:HG21	2.04	0.40
2:B:142:ILE:HD12	2:B:142:ILE:H	1.86	0.40
1:C:295:LEU:HB3	1:C:300:GLU:CD	2.42	0.40
1:C:326:ILE:O	1:C:341:ILE:HA	2.22	0.40
1:C:395:LYS:HD3	1:C:414:TRP:CZ2	2.56	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	537/558 (96%)	530 (99%)	7 (1%)	0	100	100
1	C	530/558 (95%)	524 (99%)	6 (1%)	0	100	100
2	B	403/428 (94%)	397 (98%)	6 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	394/428 (92%)	386 (98%)	8 (2%)	0	100	100
All	All	1864/1972 (94%)	1837 (99%)	27 (1%)	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	425/495 (86%)	389 (92%)	36 (8%)	10	28
1	C	393/495 (79%)	363 (92%)	30 (8%)	13	33
2	B	350/390 (90%)	333 (95%)	17 (5%)	25	54
2	D	242/390 (62%)	224 (93%)	18 (7%)	13	34
All	All	1410/1770 (80%)	1309 (93%)	101 (7%)	14	36

All (101) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	THR
1	A	21	VAL
1	A	23	GLN
1	A	29	GLU
1	A	58	THR
1	A	60	VAL
1	A	63	ILE
1	A	68	SER
1	A	71	TRP
1	A	105	SER
1	A	135	ILE
1	A	139	THR
1	A	151	GLN
1	A	184	MET
1	A	199	ARG
1	A	219	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	244	ILE
1	A	245	VAL
1	A	250	ASP
1	A	253	THR
1	A	260	LEU
1	A	289	LEU
1	A	302	GLU
1	A	324	ASP
1	A	332	GLN
1	A	362	THR
1	A	402	TRP
1	A	415	GLU
1	A	459	THR
1	A	473	THR
1	A	482	ILE
1	A	491	LEU
1	A	513	SER
1	A	514	GLU
1	A	527	LYS
1	A	547	GLN
2	B	72	ARG
2	B	85	GLN
2	B	86	ASP
2	B	88	TRP
2	B	89	GLU
2	B	92	LEU
2	B	117	SER
2	B	121	ASP
2	B	163	SER
2	B	215	THR
2	B	216	THR
2	B	234	LEU
2	B	242	GLN
2	B	286	THR
2	B	418	ASN
2	B	425	LEU
2	B	426	TRP
1	C	27	THR
1	C	28	GLU
1	C	29	GLU
1	C	31	ILE
1	C	32	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	35	VAL
1	C	101	LYS
1	C	110[A]	ASP
1	C	110[B]	ASP
1	C	146	TYR
1	C	147	ASN
1	C	161	GLN
1	C	177	ASP
1	C	186	ASP
1	C	200	THR
1	C	216	THR
1	C	221	HIS
1	C	224	GLU
1	C	237	ASP
1	C	250	ASP
1	C	277	ARG
1	C	283	LEU
1	C	291	GLU
1	C	324	ASP
1	C	351	THR
1	C	399	GLU
1	C	402	TRP
1	C	405	TYR
1	C	442	VAL
1	C	496	VAL
2	D	90	VAL
2	D	96	HIS
2	D	100	LEU
2	D	101	LYS
2	D	121	ASP
2	D	211	ARG
2	D	291	GLU
2	D	293	ILE
2	D	298	GLU
2	D	314	VAL
2	D	330	GLN
2	D	340	GLN
2	D	362	THR
2	D	377	THR
2	D	405	TYR
2	D	414	TRP
2	D	423	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
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
2	D	394	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 1 is monoatomic - leaving 6 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	$\#\|Z\| > 2$
5	3LQ	C	603	-	33,33,33	0.48	0	44,45,45	0.60	0
3	SO4	A	602	-	4,4,4	0.14	0	6,6,6	0.05	0
3	SO4	A	601	-	4,4,4	0.14	0	6,6,6	0.05	0
4	EDO	C	602	-	3,3,3	0.43	0	2,2,2	0.39	0
4	EDO	A	603	-	3,3,3	0.47	0	2,2,2	0.25	0
5	3LQ	A	604	-	33,33,33	0.45	0	44,45,45	0.52	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
5	3LQ	C	603	-	-	5/12/12/12	0/4/4/4
4	EDO	A	603	-	-	1/1/1/1	-
4	EDO	C	602	-	-	0/1/1/1	-
5	3LQ	A	604	-	-	0/12/12/12	0/4/4/4

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	603	3LQ	O0B-C0D-C0E-N0H
5	C	603	3LQ	C0E-C0D-O0B-C05
5	C	603	3LQ	C00-C05-O0B-C0D
5	C	603	3LQ	C04-C05-O0B-C0D
5	C	603	3LQ	C05-C04-O0A-CAH
4	A	603	EDO	O1-C1-C2-O2

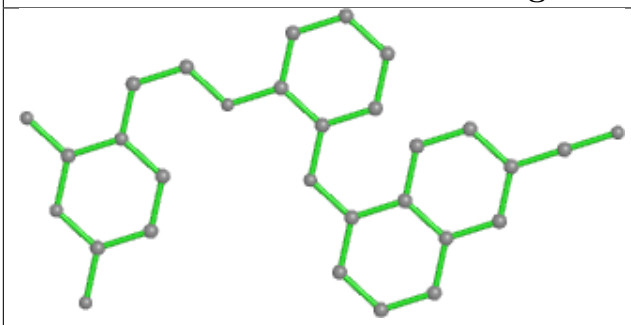
There are no ring outliers.

2 monomers are involved in 2 short contacts:

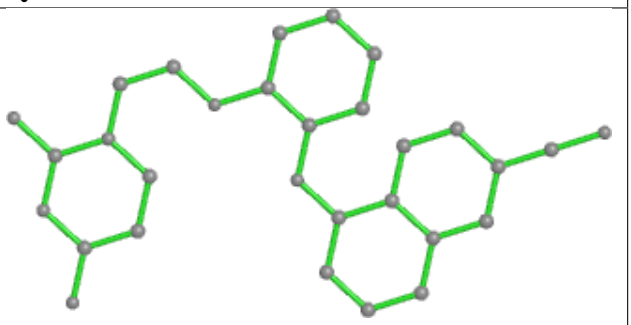
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	602	EDO	1	0
5	A	604	3LQ	1	0

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.

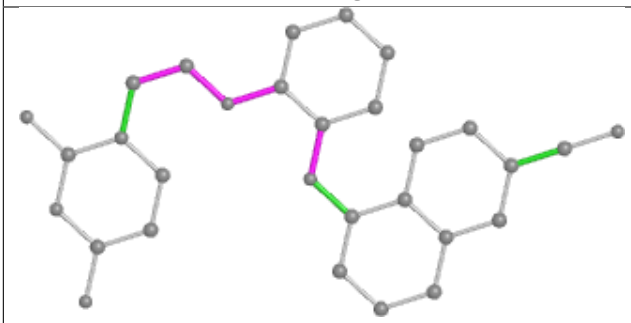
Ligand 3LQ C 603



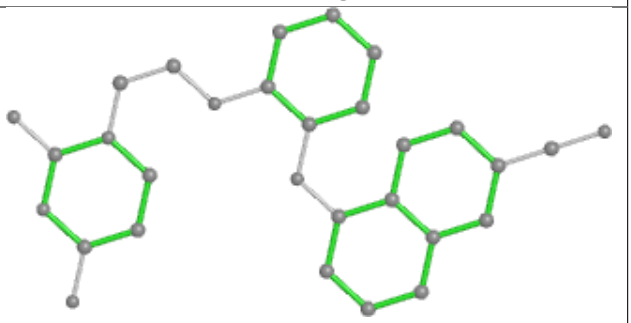
Bond lengths



Bond angles

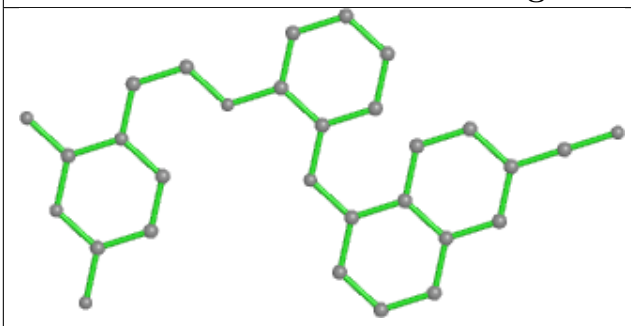


Torsions

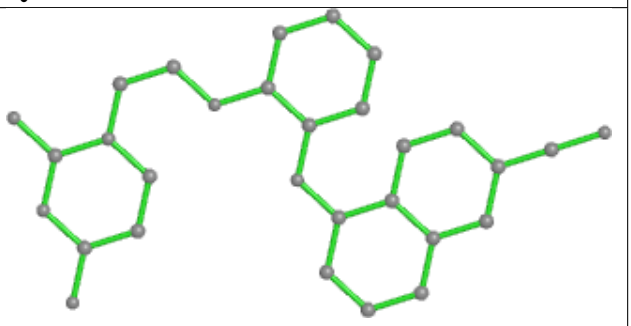


Rings

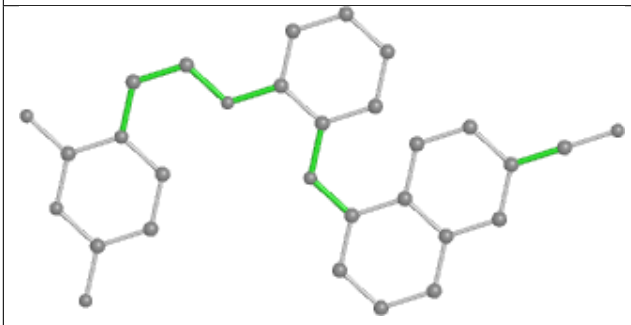
Ligand 3LQ A 604



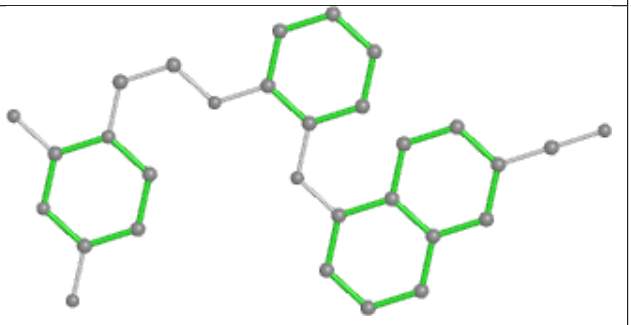
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers

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

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	544/558 (97%)	0.30	26 (4%)	30	24	47, 84, 134, 161	0
1	C	542/558 (97%)	0.50	43 (7%)	12	9	74, 111, 148, 171	0
2	B	408/428 (95%)	0.33	14 (3%)	45	39	48, 78, 121, 139	6 (1%)
2	D	401/428 (93%)	0.55	45 (11%)	5	3	81, 125, 159, 170	4 (0%)
All	All	1895/1972 (96%)	0.41	128 (6%)	17	12	47, 100, 150, 171	10 (0%)

All (128) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	283	LEU	4.6
2	D	272	PRO	4.6
2	D	299	ALA	4.5
2	D	232	TYR	4.5
1	A	292	VAL	4.5
1	C	247	PRO	4.4
1	C	142	ILE	4.4
1	A	552	VAL	4.3
1	C	290	THR	4.2
1	A	286	THR	4.2
1	C	282	LEU	4.2
2	D	200	THR	4.1
1	C	144	TYR	4.0
1	C	61	PHE	4.0
1	C	224	GLU	3.8
1	C	131	THR	3.8
1	A	282	LEU	3.6
1	C	402	TRP	3.6
1	A	251	SER	3.4
2	D	316	GLY	3.4
1	C	31	ILE	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	358	ARG	3.4
1	A	255	ASN	3.3
1	C	292	VAL	3.3
1	C	127	TYR	3.3
2	D	178	ILE	3.3
2	D	317	VAL	3.2
1	C	283	LEU	3.2
1	C	246	LEU	3.2
2	D	10	VAL	3.2
2	D	349	LEU	3.1
2	D	319	TYR	3.1
1	A	34	LEU	3.1
1	C	254	VAL	3.0
2	B	274	ILE	3.0
2	D	271	TYR	3.0
2	D	252	TRP	3.0
2	D	160	PHE	2.9
1	C	146	TYR	2.9
1	C	26	LEU	2.9
1	A	294	PRO	2.9
1	A	256	ASP	2.9
2	D	176	PRO	2.9
1	A	302	GLU	2.9
2	B	232	TYR	2.9
1	A	542	ILE	2.9
2	D	177	ASP	2.9
1	C	251	SER	2.8
1	C	130	PHE	2.8
1	C	279	LEU	2.8
1	C	37	ILE	2.8
2	D	421	PRO	2.8
1	A	290	THR	2.8
1	C	35	VAL	2.8
2	D	102	LYS	2.8
2	D	245	VAL	2.7
2	D	96	HIS	2.7
2	D	111	VAL	2.7
2	B	284	ARG	2.7
2	B	14	PRO	2.7
2	D	127	TYR	2.6
1	C	359	GLY	2.6
2	D	315	HIS	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	227	PHE	2.6
2	D	92	LEU	2.6
2	D	174	GLN	2.6
1	C	195	ILE	2.6
2	D	84	THR	2.6
1	A	260	LEU	2.5
2	D	114	ALA	2.5
2	D	142	ILE	2.5
2	D	341	ILE	2.5
2	D	199	ARG	2.5
2	D	296	THR	2.5
1	A	551	LEU	2.5
2	D	180	ILE	2.4
1	C	426	TRP	2.4
1	C	291	GLU	2.4
1	A	71	TRP	2.4
2	D	298	GLU	2.4
1	C	555	GLY	2.4
1	A	226	PRO	2.4
2	D	358	ARG	2.4
2	D	124	PHE	2.4
2	D	190	GLY	2.4
1	C	255	ASN	2.4
1	C	38	CYS	2.4
1	A	61	PHE	2.4
2	B	69	THR	2.3
1	C	257	ILE	2.3
1	C	15	GLY	2.3
1	C	556	ILE	2.3
1	A	285	GLY	2.3
1	C	135	ILE	2.3
2	D	168	LEU	2.3
1	A	246	LEU	2.2
1	C	294	PRO	2.2
1	C	145	GLN	2.2
1	C	88	TRP	2.2
1	A	288	ALA	2.2
2	B	299	ALA	2.2
1	C	51	GLY	2.2
2	D	12	LEU	2.2
2	D	303	LEU	2.2
1	A	296	THR	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	284	ARG	2.2
1	C	65	LYS	2.1
1	C	358	ARG	2.1
2	D	383	TRP	2.1
1	C	47	ILE	2.1
2	B	88	TRP	2.1
2	D	9	PRO	2.1
2	D	119	PRO	2.1
1	A	253	THR	2.1
2	D	306	ASN	2.1
2	B	279	LEU	2.1
2	B	425	LEU	2.1
2	B	427	TYR	2.1
1	C	150	PRO	2.0
2	B	421	PRO	2.0
1	A	299	ALA	2.0
1	A	303	LEU	2.0
1	A	298	GLU	2.0
2	D	175	ASN	2.0
1	C	128	THR	2.0
2	D	279	LEU	2.0
2	D	198	HIS	2.0
2	B	68	SER	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.

Continued on next page...

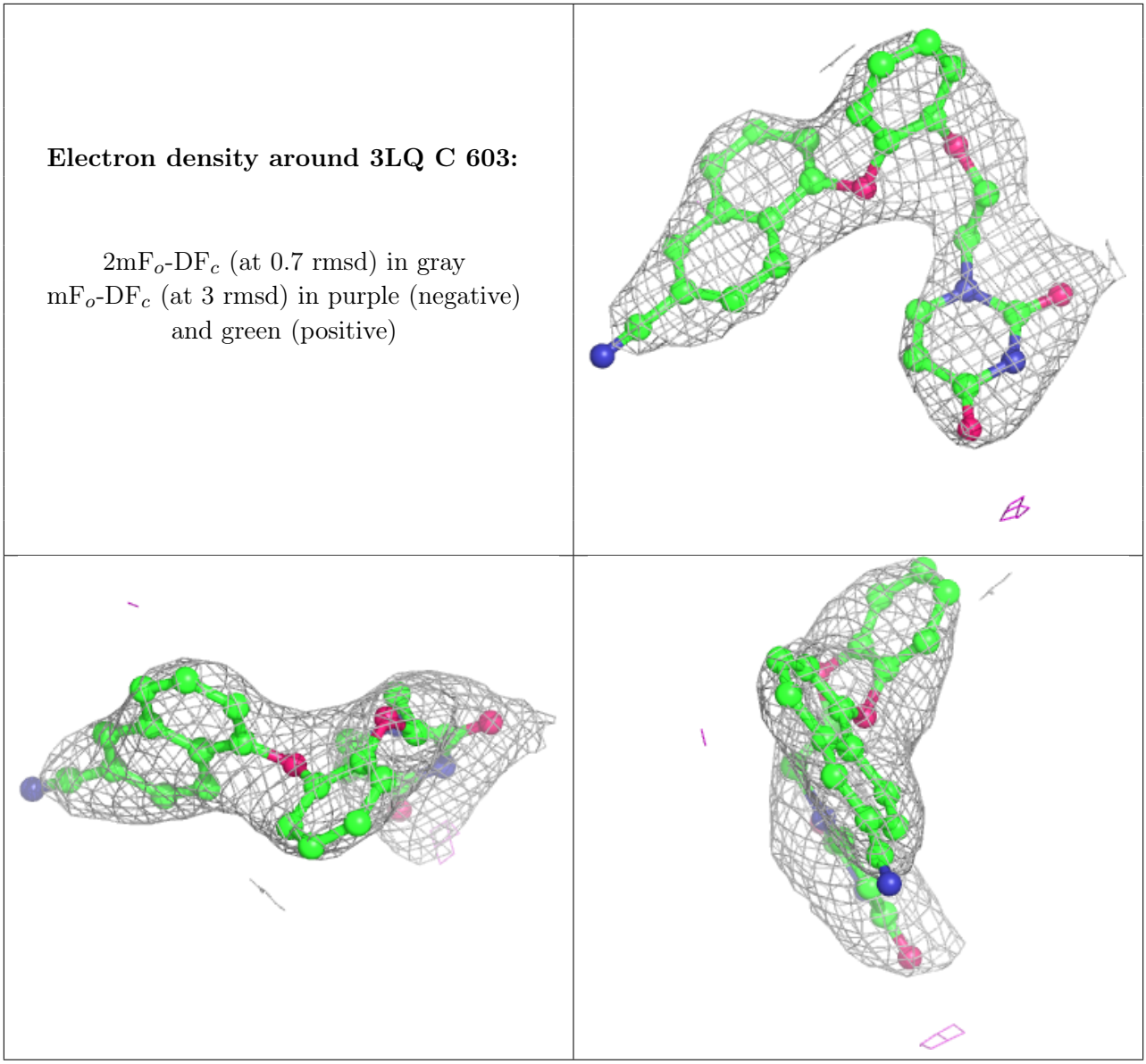
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	EDO	A	603	4/4	0.69	0.27	75,81,85,87	0
6	MG	C	601	1/1	0.84	0.30	100,100,100,100	0
4	EDO	C	602	4/4	0.88	0.19	99,100,105,117	0
3	SO4	A	602	5/5	0.92	0.12	107,119,130,140	0
5	3LQ	C	603	30/30	0.94	0.26	85,100,109,118	0
3	SO4	A	601	5/5	0.94	0.15	106,107,115,130	0
5	3LQ	A	604	30/30	0.97	0.20	51,64,77,86	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.

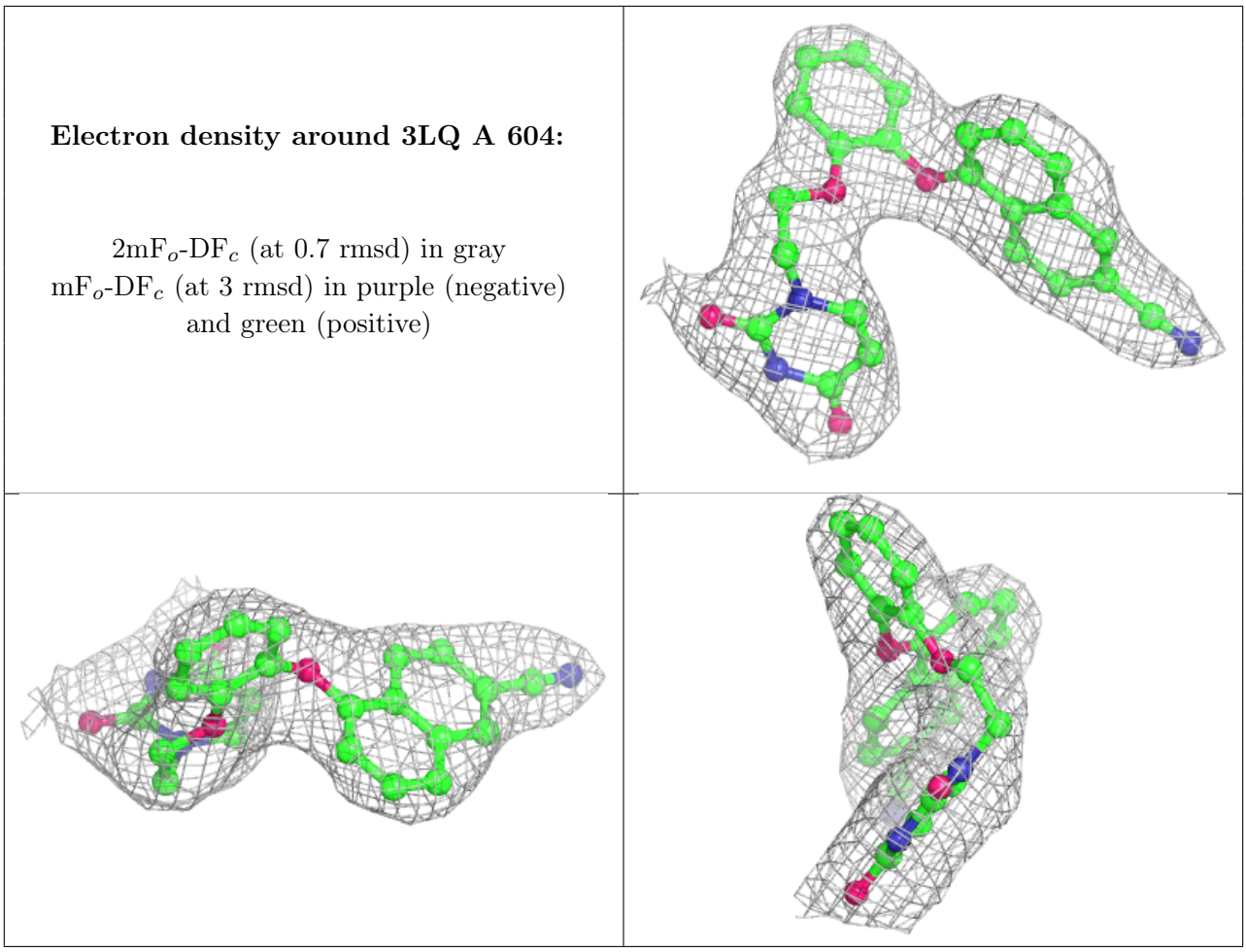
Electron density around 3LQ C 603:

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



Electron density around 3LQ A 604:

$2mF_o$-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

There are no such residues in this entry.