Full wwPDB X-ray Structure Validation Report (i)

Oct 16, 2023 – 12:20 AM EDT

PDB ID	:	8FO3
Title	:	Structure of full-length amyloidogenic immunoglobulin light chain H9 in complex with (E)-3-nitro-4-(2-(2-phenylpropylidene)hydrazineyl)benzenesulfonamide
Authors	:	Yan, N.L.; Wilson, I.A.; Kelly, J.W.
Deposited on	:	2022-12-29
Resolution	:	2.00 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 (i) 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 (1)) were used in the production of this report:

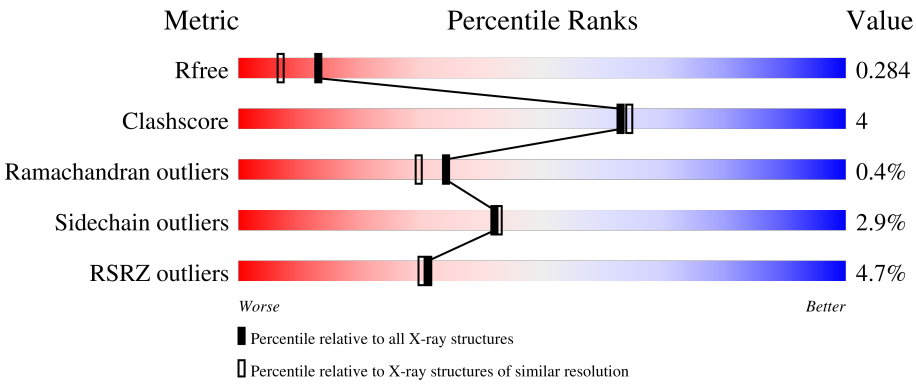
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	216	3% 94% 5%
1	B	216	2% 94% 6%
1	C	216	3% 88% 11% •
1	D	216	% 93% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	E	216	3% 87% 10%
1	F	216	4% 87% 12%
1	G	216	6% 86% 10%
1	H	216	3% 92% 6%
1	I	216	4% 92% 6%
1	J	216	6% 87% 11%
1	K	216	15% 71% 24%
1	L	216	4% 90% 8%

2 Entry composition (i)

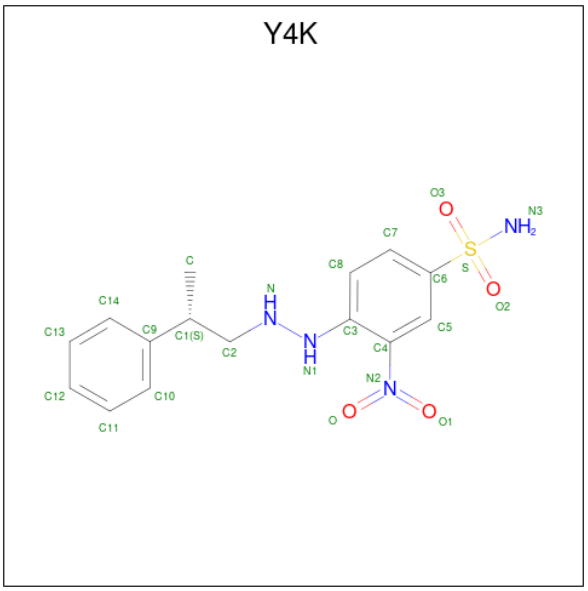
There are 4 unique types of molecules in this entry. The entry contains 20516 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 H9 immunoglobulin light chain.

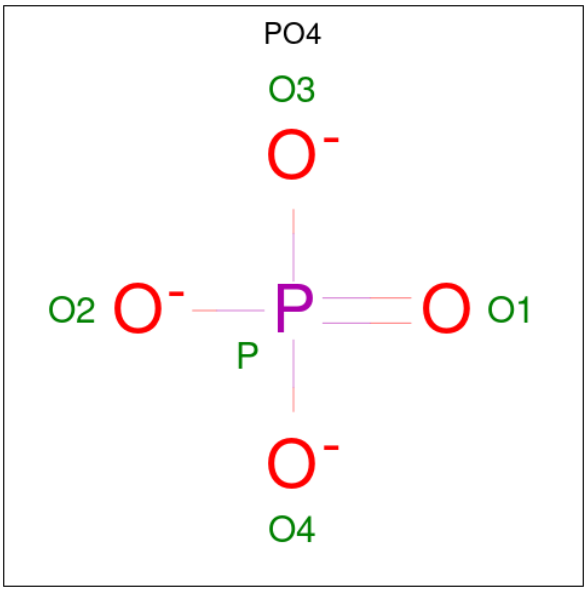
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	215	1600	990	269	336	5	0	3	0
1	B	215	1585	983	265	332	5	0	1	0
1	C	214	1605	994	269	337	5	0	5	0
1	D	215	1584	981	265	333	5	0	1	0
1	G	211	1572	976	262	330	4	0	4	0
1	H	212	1554	965	260	325	4	0	0	0
1	I	211	1593	988	268	333	4	0	7	0
1	J	214	1579	980	264	331	4	0	1	0
1	E	209	1570	974	264	328	4	0	5	0
1	F	213	1569	973	263	329	4	0	1	0
1	K	210	1563	970	263	326	4	0	2	0
1	L	214	1578	978	264	332	4	0	1	0

- Molecule 2 is 3-nitro-4-{2-[(2S)-2-phenylpropyl]hydrazinyl}benzene-1-sulfonamide (three-letter code: Y4K) (formula: $C_{15}H_{18}N_4O_4S$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			24	15	4	4	1		
2	C	1	Total	C	N	O	S	0	0
			24	15	4	4	1		
2	G	1	Total	C	N	O	S	0	0
			24	15	4	4	1		
2	I	1	Total	C	N	O	S	0	0
			24	15	4	4	1		
2	E	1	Total	C	N	O	S	0	0
			24	15	4	4	1		
2	K	1	Total	C	N	O	S	0	0
			24	15	4	4	1		

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O P 5 4 1	0	0
3	B	1	Total O P 5 4 1	0	0
3	G	1	Total O P 5 4 1	0	0
3	H	1	Total O P 5 4 1	0	0
3	E	1	Total O P 5 4 1	0	0
3	K	1	Total O P 5 4 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	165	Total O 165 165	0	0
4	B	149	Total O 149 149	0	0
4	C	142	Total O 142 142	0	0
4	D	153	Total O 153 153	0	0
4	G	77	Total O 77 77	0	0
4	H	104	Total O 104 104	0	0

Continued on next page...

Continued from previous page...

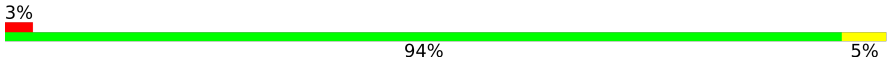
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	I	123	123	123	0	0
			Total	O		
4	J	105	105	105	0	0
			Total	O		
4	E	105	105	105	0	0
			Total	O		
4	F	104	104	104	0	0
			Total	O		
4	K	64	64	64	0	0
			Total	O		
4	L	99	99	99	0	0

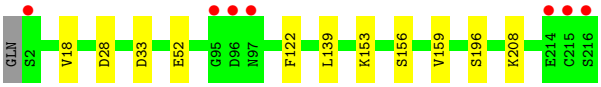
3 Residue-property plots (i)

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: H9 immunoglobulin light chain

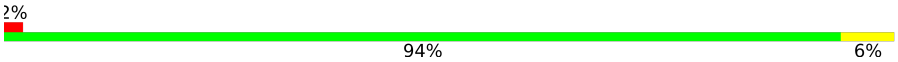
Chain A:

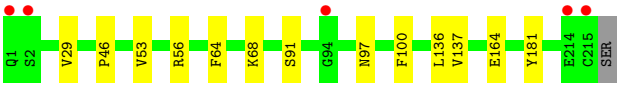




• Molecule 1: H9 immunoglobulin light chain

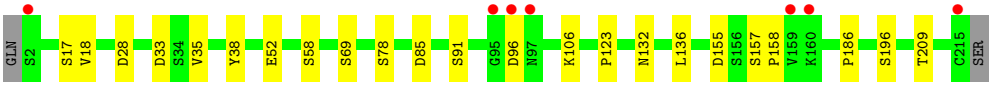
Chain B:





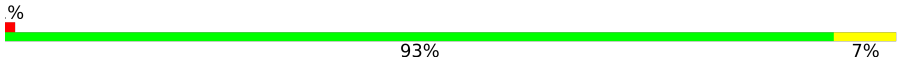
• Molecule 1: H9 immunoglobulin light chain

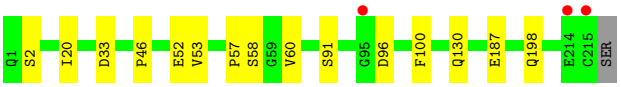
Chain C:



• Molecule 1: H9 immunoglobulin light chain

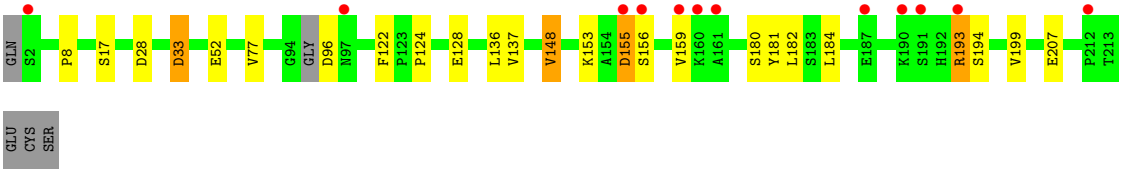
Chain D:





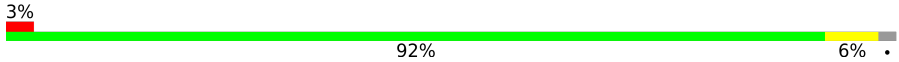
• Molecule 1: H9 immunoglobulin light chain

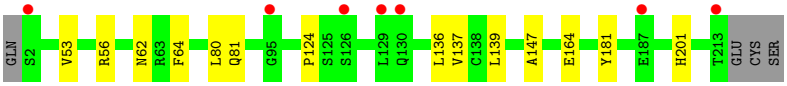
Chain G:



• Molecule 1: H9 immunoglobulin light chain

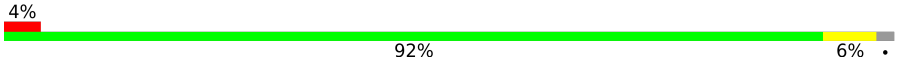
Chain H:





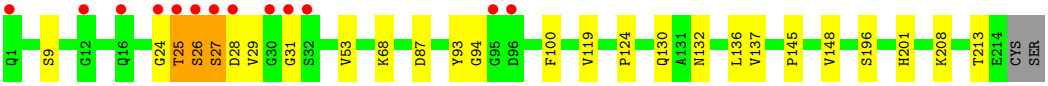
• Molecule 1: H9 immunoglobulin light chain

Chain I:



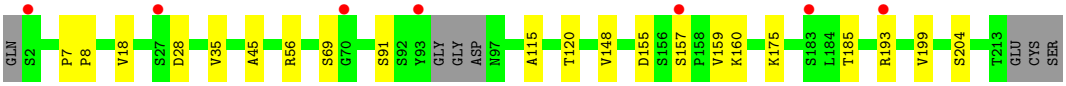
• Molecule 1: H9 immunoglobulin light chain

Chain J:



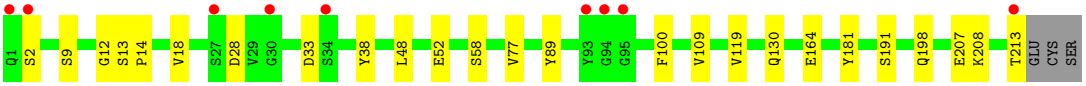
• Molecule 1: H9 immunoglobulin light chain

Chain E:



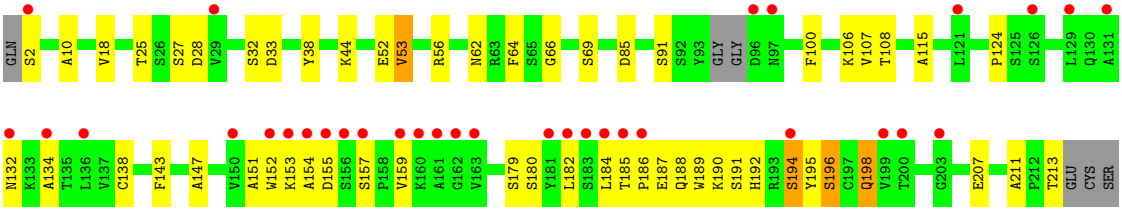
• Molecule 1: H9 immunoglobulin light chain

Chain F:



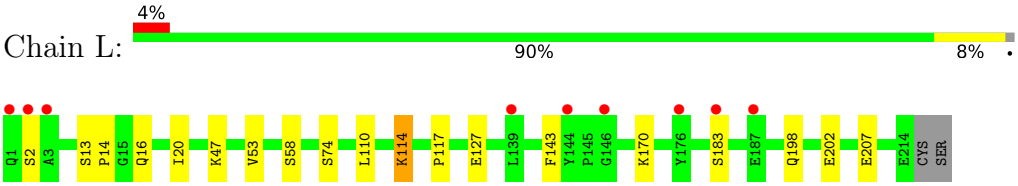
• Molecule 1: H9 immunoglobulin light chain

Chain K:



• Molecule 1: H9 immunoglobulin light chain

Chain L:


4%
90%
8%
Q1
S2
A3
S13
P14
G15
Q16
I20
K47
V53
S58
S74
L110
K114
P117
E127
L139
F143
Y144
P145
G146
K170
Y176
S183
E187
Q198
E202
E207
E214
CYS
SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	63.14Å 95.66Å 125.92Å 106.31° 92.93° 90.16°	Depositor
Resolution (Å)	47.62 – 2.00 47.57 – 2.00	Depositor EDS
% Data completeness (in resolution range)	96.4 (47.62-2.00) 96.4 (47.57-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	3.30 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.229 , 0.280 0.234 , 0.284	Depositor DCC
R_{free} test set	9361 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å^2)	28.0	Xtriage
Anisotropy	0.083	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 42.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	20516	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 40.27 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.8048e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: Y4K, PO4

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	0.71	0/1638	0.86	0/2234
1	B	0.74	0/1623	0.86	0/2216
1	C	0.71	0/1646	0.87	0/2247
1	D	0.73	0/1622	0.86	0/2214
1	E	0.71	0/1610	0.84	0/2197
1	F	0.74	0/1607	0.84	0/2194
1	G	0.69	0/1612	0.83	0/2200
1	H	0.72	0/1592	0.85	0/2174
1	I	0.70	0/1637	0.87	0/2235
1	J	0.71	0/1617	0.85	0/2208
1	K	0.72	0/1600	0.85	0/2184
1	L	0.72	0/1616	0.86	0/2206
All	All	0.72	0/19420	0.85	0/26509

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 (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	1600	0	1527	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1585	0	1517	8	0
1	C	1605	0	1533	9	0
1	D	1584	0	1513	7	0
1	E	1570	0	1507	9	0
1	F	1569	0	1503	15	0
1	G	1572	0	1505	13	0
1	H	1554	0	1488	10	0
1	I	1593	0	1529	14	0
1	J	1579	0	1513	29	0
1	K	1563	0	1498	46	0
1	L	1578	0	1509	8	0
2	A	24	0	0	1	0
2	C	24	0	0	1	0
2	E	24	0	0	0	0
2	G	24	0	0	0	0
2	I	24	0	0	0	0
2	K	24	0	0	1	0
3	A	5	0	0	0	0
3	B	5	0	0	0	0
3	E	5	0	0	0	0
3	G	5	0	0	0	0
3	H	5	0	0	0	0
3	K	5	0	0	0	0
4	A	165	0	0	0	0
4	B	149	0	0	1	0
4	C	142	0	0	0	0
4	D	153	0	0	1	0
4	E	105	0	0	3	0
4	F	104	0	0	4	0
4	G	77	0	0	1	0
4	H	104	0	0	1	0
4	I	123	0	0	1	0
4	J	105	0	0	2	0
4	K	64	0	0	4	0
4	L	99	0	0	0	0
All	All	20516	0	18142	161	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 4.

All (161) 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:J:213:THR:HA	4:J:386:HOH:O	1.35	1.25
1:K:186:PRO:O	1:K:190:LYS:HB2	1.75	0.87
1:K:185:THR:O	1:K:189:TRP:HB3	1.78	0.84
1:F:119:VAL:O	1:F:208:LYS:HE3	1.85	0.76
1:K:188:GLN:O	1:K:192:HIS:HD2	1.70	0.75
1:G:155:ASP:OD1	1:G:193:ARG:HB2	1.87	0.75
1:K:186:PRO:O	1:K:190:LYS:CB	2.35	0.75
1:K:56[B]:ARG:HB2	1:K:56[B]:ARG:NH1	2.02	0.74
1:B:56:ARG:HD3	1:B:64:PHE:O	1.88	0.73
1:D:20:ILE:HB	4:D:314:HOH:O	1.91	0.71
1:E:56[A]:ARG:NH2	4:E:402:HOH:O	2.28	0.67
1:K:25:THR:HG1	1:K:27:SER:HG	1.43	0.64
1:K:188:GLN:O	1:K:192:HIS:CD2	2.53	0.61
1:J:24:GLY:HA3	1:J:29:VAL:CG2	2.31	0.60
1:K:106:LYS:HG2	1:K:147:ALA:HB2	1.83	0.60
1:D:33:ASP:O	1:D:52:GLU:HA	2.02	0.60
1:K:33:ASP:O	1:K:52:GLU:HA	2.03	0.58
1:I:42:PRO:HD2	4:I:520:HOH:O	2.03	0.58
1:A:139:LEU:HD22	1:B:137[A]:VAL:HG22	1.87	0.57
1:B:97:ASN:ND2	4:B:402:HOH:O	2.34	0.57
1:J:28:ASP:OD2	1:J:100:PHE:CD2	2.57	0.57
1:F:13:SER:HB2	1:F:14:PRO:HD2	1.87	0.57
1:K:154:ALA:HB2	1:K:159:VAL:HG11	1.86	0.57
1:G:153:LYS:HD3	1:G:207:GLU:OE2	2.05	0.56
1:F:213:THR:HA	4:F:333:HOH:O	2.04	0.56
1:K:138:CYS:HB3	1:K:180:SER:HB3	1.87	0.56
1:E:175:LYS:HE2	4:E:495:HOH:O	2.04	0.55
1:K:56[B]:ARG:HH11	1:K:56[B]:ARG:CB	2.19	0.55
1:J:25:THR:OG1	1:J:27:SER:N	2.37	0.55
1:K:56[B]:ARG:HB2	1:K:56[B]:ARG:HH11	1.72	0.54
1:C:196:SER:HG	1:C:209:THR:HG1	1.41	0.54
1:G:137[B]:VAL:HG22	1:H:139:LEU:HD22	1.89	0.54
1:C:157:SER:O	1:C:158:PRO:C	2.44	0.54
1:K:185:THR:O	1:K:189:TRP:CB	2.54	0.54
1:J:28:ASP:HB3	1:J:94:GLY:CA	2.38	0.53
1:H:80:LEU:O	1:H:81:GLN:NE2	2.42	0.53
1:J:24:GLY:HA3	1:J:29:VAL:HG21	1.90	0.53
1:K:124:PRO:HB3	1:K:134:ALA:HA	1.91	0.52
1:L:117:PRO:HA	1:L:143:PHE:HB3	1.91	0.52
1:I:159:VAL:O	1:I:159:VAL:HG22	2.07	0.52
1:A:33:ASP:O	1:A:52:GLU:HA	2.09	0.52
1:J:25:THR:HG23	1:J:28:ASP:OD1	2.09	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:137[B]:VAL:CG2	1:H:139:LEU:HD22	2.38	0.52
1:E:155:ASP:OD1	1:E:193:ARG:HB3	2.09	0.52
1:H:124:PRO:HD3	1:H:136:LEU:HD11	1.92	0.52
1:K:152:TRP:CE3	1:K:196:SER:O	2.62	0.52
1:L:198:GLN:OE1	1:L:207:GLU:OE1	2.27	0.52
1:K:91:SER:HA	1:K:100:PHE:O	2.11	0.51
1:J:28:ASP:HA	1:J:93:TYR:O	2.09	0.51
1:K:185:THR:OG1	1:K:188:GLN:N	2.44	0.51
1:I:139:LEU:HD22	1:J:137[A]:VAL:HG22	1.93	0.51
1:K:132:ASN:O	1:K:132:ASN:CG	2.50	0.51
1:E:148:VAL:CG2	1:E:199:VAL:HG13	2.41	0.50
1:C:17:SER:OG	1:C:78:SER:HA	2.11	0.50
1:K:159:VAL:CG2	4:K:409:HOH:O	2.59	0.50
1:F:198:GLN:NE2	1:F:207:GLU:OE1	2.42	0.50
1:J:28:ASP:OD2	1:J:100:PHE:CE2	2.64	0.49
1:H:62:ASN:HB2	4:H:457:HOH:O	2.13	0.49
1:A:122:PHE:CZ	1:B:137[A]:VAL:HG23	2.47	0.49
1:K:186:PRO:O	1:K:190:LYS:N	2.43	0.49
1:K:159:VAL:HG22	4:K:409:HOH:O	2.13	0.49
1:J:29:VAL:O	1:J:68:LYS:HD2	2.12	0.49
1:J:25:THR:HG23	1:J:28:ASP:CG	2.34	0.49
1:C:132:ASN:HA	1:C:186:PRO:HG2	1.95	0.49
1:J:25:THR:HG1	1:J:27:SER:H	1.59	0.49
1:F:12:GLY:O	1:F:109:VAL:HA	2.14	0.48
1:I:122:PHE:CE1	1:J:137[A]:VAL:HG21	2.48	0.48
1:G:148:VAL:HG22	1:G:199:VAL:HG13	1.96	0.48
1:I:122:PHE:HB2	1:I:137[A]:VAL:HG13	1.96	0.48
1:K:154:ALA:HB2	1:K:159:VAL:CG1	2.43	0.48
1:I:4:LEU:O	1:I:102:GLY:HA2	2.14	0.48
1:G:182:LEU:HG	1:G:184:LEU:HG	1.96	0.47
1:J:26:SER:O	1:J:31:GLY:CA	2.62	0.47
1:C:33:ASP:O	1:C:52:GLU:HA	2.15	0.47
1:J:26:SER:O	1:J:31:GLY:HA2	2.15	0.47
1:J:24:GLY:HA3	1:J:29:VAL:HG23	1.96	0.47
1:F:213:THR:CA	4:F:333:HOH:O	2.63	0.47
1:C:85:ASP:OD2	1:C:106:LYS:NZ	2.48	0.46
1:J:27:SER:O	1:J:94:GLY:HA2	2.16	0.46
1:F:213:THR:C	4:F:333:HOH:O	2.54	0.46
1:E:45:ALA:HA	1:F:89:TYR:CE1	2.51	0.46
1:D:198:GLN:OE1	1:I:193:ARG:NH2	2.49	0.46
1:K:153:LYS:HD3	1:K:207:GLU:OE2	2.16	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:K:159:VAL:O	1:K:159:VAL:HG23	2.16	0.45
1:G:148:VAL:CG2	1:G:199:VAL:HG13	2.47	0.45
1:F:191:SER:O	4:F:301:HOH:O	2.21	0.45
1:B:29:VAL:O	1:B:68:LYS:HD2	2.17	0.45
1:L:114:LYS:HE3	1:L:202:GLU:CG	2.47	0.45
1:H:164:GLU:HB2	1:H:181:TYR:CZ	2.51	0.45
1:I:122:PHE:CE1	1:J:137[A]:VAL:CG2	2.99	0.45
1:I:122:PHE:CZ	1:J:137[A]:VAL:HG23	2.52	0.45
1:I:198[B]:GLN:CG	1:I:205:THR:CG2	2.95	0.45
1:J:119:VAL:O	1:J:208:LYS:HE3	2.17	0.45
1:J:124:PRO:HD3	1:J:136:LEU:HD11	1.98	0.45
1:A:196:SER:HA	1:A:208:LYS:O	2.16	0.45
1:J:28:ASP:HA	1:J:94:GLY:HA2	1.99	0.44
1:J:145:PRO:O	1:J:201:HIS:HE1	2.00	0.44
1:F:38:TYR:CE2	1:F:48:LEU:HD13	2.52	0.44
1:L:14:PRO:HD3	1:L:110:LEU:O	2.17	0.44
1:I:198[B]:GLN:HG3	1:I:205:THR:CG2	2.47	0.44
1:E:35:VAL:HA	1:E:91:SER:O	2.17	0.44
1:G:17:SER:HA	1:G:77:VAL:O	2.17	0.44
1:F:52:GLU:OE1	1:K:56[B]:ARG:NH2	2.44	0.44
2:C:301:Y4K:C10	1:D:46:PRO:HG3	2.48	0.44
1:G:136:LEU:O	1:G:181:TYR:HA	2.17	0.44
1:K:53:VAL:O	1:K:66:GLY:HA3	2.18	0.44
1:K:187:GLU:O	1:K:191:SER:N	2.28	0.44
1:K:56[A]:ARG:HD3	1:K:62:ASN:HA	1.99	0.44
1:K:182:LEU:HG	1:K:184:LEU:HG	1.99	0.44
1:L:114:LYS:HE3	1:L:202:GLU:HG3	1.99	0.44
1:G:124:PRO:HA	1:G:128:GLU:OE2	2.18	0.44
1:K:185:THR:HG1	1:K:188:GLN:H	1.66	0.44
1:F:18:VAL:HG12	1:F:77:VAL:HB	2.00	0.44
1:H:56:ARG:HD3	1:H:64:PHE:O	2.19	0.43
1:H:136:LEU:O	1:H:181:TYR:HA	2.18	0.43
1:E:185:THR:HG21	4:E:431:HOH:O	2.18	0.43
1:K:138:CYS:N	1:K:180:SER:O	2.34	0.43
1:K:194:SER:HB3	1:K:211:ALA:HB2	2.00	0.43
1:D:91:SER:HA	1:D:100:PHE:O	2.18	0.43
1:L:13:SER:O	1:L:16:GLN:HB2	2.18	0.43
1:J:28:ASP:HB3	1:J:94:GLY:HA2	2.00	0.43
1:J:132:ASN:ND2	4:J:301:HOH:O	2.30	0.43
1:K:184:LEU:HD13	1:K:195:TYR:CE1	2.54	0.43
1:I:153:LYS:HD2	1:I:198[A]:GLN:NE2	2.34	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:K:44:LYS:HG3	4:K:454:HOH:O	2.19	0.43
1:K:182:LEU:HD21	1:K:195:TYR:CE2	2.54	0.43
1:C:35:VAL:HA	1:C:91:SER:O	2.19	0.43
1:J:196:SER:HA	1:J:208:LYS:O	2.18	0.43
1:K:108:THR:HG23	4:K:401:HOH:O	2.18	0.43
1:C:123:PRO:HA	1:C:136:LEU:HD22	2.01	0.42
1:K:138:CYS:O	1:K:180:SER:N	2.49	0.42
1:D:57:PRO:HD2	1:D:60:VAL:HG21	2.01	0.42
1:B:91:SER:HA	1:B:100:PHE:O	2.20	0.42
1:K:56[B]:ARG:HD3	1:K:64:PHE:O	2.20	0.42
1:D:187:GLU:H	1:D:187:GLU:CD	2.23	0.42
1:F:33:ASP:O	1:F:52:GLU:HA	2.19	0.42
1:G:33:ASP:O	1:G:52:GLU:HA	2.19	0.41
1:E:115:ALA:CB	1:E:175:LYS:HE3	2.50	0.41
1:K:2:SER:CB	2:K:301:Y4K:N3	2.83	0.41
1:C:196:SER:OG	1:C:209:THR:OG1	2.23	0.41
1:I:122:PHE:HB2	1:I:137[A]:VAL:CG1	2.50	0.41
1:F:2:SER:O	1:F:100:PHE:HD1	2.04	0.41
1:G:122:PHE:CE1	1:H:137:VAL:CG2	3.04	0.41
1:I:153:LYS:HA	1:I:157:SER:O	2.21	0.41
1:J:148:VAL:HG12	1:J:201:HIS:HB2	2.03	0.41
1:K:56[B]:ARG:NH1	1:K:56[B]:ARG:CB	2.74	0.41
1:A:153:LYS:HE3	1:A:156:SER:HA	2.03	0.41
2:A:301:Y4K:C10	1:B:46:PRO:HG3	2.51	0.41
1:K:115:ALA:O	1:K:143:PHE:HA	2.21	0.41
1:L:198:GLN:HG2	1:L:207:GLU:HG3	2.03	0.41
1:B:164:GLU:HB2	1:B:181:TYR:CZ	2.56	0.41
1:G:8:PRO:HD2	4:G:444:HOH:O	2.19	0.41
1:H:147:ALA:O	1:H:201:HIS:HD2	2.04	0.40
1:K:10:ALA:O	1:K:107:VAL:HA	2.21	0.40
1:K:151:ALA:HB3	1:K:198:GLN:HG3	2.03	0.40
1:E:7:PRO:HA	1:E:8:PRO:HD3	1.91	0.40
1:K:85:ASP:OD2	1:K:106:LYS:NZ	2.54	0.40
1:J:25:THR:HG23	1:J:28:ASP:OD2	2.21	0.40
1:F:164:GLU:HB2	1:F:181:TYR:CZ	2.57	0.40
1:L:20:ILE:O	1:L:74:SER:HA	2.22	0.40
1:K:152:TRP:HA	1:K:196:SER:O	2.22	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	216/216 (100%)	207 (96%)	9 (4%)	0	100	100
1	B	214/216 (99%)	203 (95%)	10 (5%)	1 (0%)	29	23
1	C	217/216 (100%)	205 (94%)	11 (5%)	1 (0%)	29	23
1	D	214/216 (99%)	207 (97%)	6 (3%)	1 (0%)	29	23
1	E	210/216 (97%)	200 (95%)	10 (5%)	0	100	100
1	F	212/216 (98%)	202 (95%)	10 (5%)	0	100	100
1	G	211/216 (98%)	204 (97%)	6 (3%)	1 (0%)	29	23
1	H	210/216 (97%)	201 (96%)	8 (4%)	1 (0%)	29	23
1	I	216/216 (100%)	206 (95%)	10 (5%)	0	100	100
1	J	213/216 (99%)	201 (94%)	11 (5%)	1 (0%)	29	23
1	K	208/216 (96%)	194 (93%)	12 (6%)	2 (1%)	15	9
1	L	213/216 (99%)	203 (95%)	9 (4%)	1 (0%)	29	23
All	All	2554/2592 (98%)	2433 (95%)	112 (4%)	9 (0%)	34	30

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	155	ASP
1	C	155	ASP
1	L	53	VAL
1	B	53	VAL
1	D	53	VAL
1	J	53	VAL
1	K	155	ASP
1	H	53	VAL
1	K	53	VAL

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	182/180 (101%)	179 (98%)	3 (2%)	62	67
1	B	180/180 (100%)	179 (99%)	1 (1%)	86	90
1	C	183/180 (102%)	177 (97%)	6 (3%)	38	37
1	D	180/180 (100%)	176 (98%)	4 (2%)	52	55
1	E	180/180 (100%)	172 (96%)	8 (4%)	28	25
1	F	178/180 (99%)	174 (98%)	4 (2%)	52	55
1	G	180/180 (100%)	171 (95%)	9 (5%)	24	20
1	H	176/180 (98%)	176 (100%)	0	100	100
1	I	182/180 (101%)	180 (99%)	2 (1%)	73	78
1	J	179/180 (99%)	173 (97%)	6 (3%)	37	36
1	K	178/180 (99%)	167 (94%)	11 (6%)	18	13
1	L	179/180 (99%)	172 (96%)	7 (4%)	32	30
All	All	2157/2160 (100%)	2096 (97%)	61 (3%)	42	44

All (61) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	VAL
1	A	28	ASP
1	A	159	VAL
1	B	136	LEU
1	C	18	VAL
1	C	28	ASP
1	C	38	TYR
1	C	58	SER
1	C	69	SER
1	C	96	ASP
1	D	2	SER
1	D	58	SER
1	D	96	ASP
1	D	130	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	G	28	ASP
1	G	33	ASP
1	G	96	ASP
1	G	148	VAL
1	G	156	SER
1	G	159	VAL
1	G	180	SER
1	G	193	ARG
1	G	194	SER
1	I	159	VAL
1	I	206	VAL
1	J	9	SER
1	J	25	THR
1	J	26	SER
1	J	27	SER
1	J	87	ASP
1	J	130	GLN
1	E	18	VAL
1	E	28	ASP
1	E	69	SER
1	E	120	THR
1	E	157	SER
1	E	159	VAL
1	E	160	LYS
1	E	204	SER
1	F	9	SER
1	F	28	ASP
1	F	58	SER
1	F	130	GLN
1	K	18	VAL
1	K	28	ASP
1	K	32	SER
1	K	38	TYR
1	K	69	SER
1	K	157	SER
1	K	179	SER
1	K	194	SER
1	K	196	SER
1	K	198	GLN
1	K	213	THR
1	L	2	SER
1	L	47	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	L	58	SER
1	L	114	LYS
1	L	127	GLU
1	L	170	LYS
1	L	183	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	132	ASN
1	H	201	HIS
1	J	41	HIS
1	K	192	HIS

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)

12 ligands are 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	PO4	K	302	-	4,4,4	0.49	0	6,6,6	0.42	0
2	Y4K	A	301	-	23,25,25	1.32	2 (8%)	32,35,35	0.82	1 (3%)
3	PO4	E	302	-	4,4,4	0.81	0	6,6,6	0.62	0
2	Y4K	E	301	-	23,25,25	0.60	1 (4%)	32,35,35	0.87	1 (3%)
2	Y4K	C	301	-	23,25,25	1.21	1 (4%)	32,35,35	0.83	1 (3%)
2	Y4K	I	301	-	23,25,25	1.10	1 (4%)	32,35,35	0.97	2 (6%)
3	PO4	A	302	-	4,4,4	1.52	1 (25%)	6,6,6	0.40	0
2	Y4K	G	301	-	23,25,25	1.21	1 (4%)	32,35,35	0.75	1 (3%)
3	PO4	B	301	-	4,4,4	1.05	0	6,6,6	0.36	0
3	PO4	G	302	-	4,4,4	0.91	0	6,6,6	0.43	0
3	PO4	H	301	-	4,4,4	0.85	0	6,6,6	0.49	0
2	Y4K	K	301	-	23,25,25	0.63	1 (4%)	32,35,35	0.92	1 (3%)

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
2	Y4K	A	301	-	-	3/20/20/20	0/2/2/2
2	Y4K	E	301	-	-	3/20/20/20	0/2/2/2
2	Y4K	C	301	-	-	3/20/20/20	0/2/2/2
2	Y4K	I	301	-	-	3/20/20/20	0/2/2/2
2	Y4K	G	301	-	-	3/20/20/20	0/2/2/2
2	Y4K	K	301	-	-	2/20/20/20	0/2/2/2

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	301	Y4K	N1-N	-5.33	1.36	1.41
2	A	301	Y4K	N1-N	-5.25	1.36	1.41
2	C	301	Y4K	N1-N	-5.22	1.36	1.41
2	I	301	Y4K	N1-N	-4.47	1.37	1.41
2	K	301	Y4K	N1-N	-2.71	1.38	1.41
2	E	301	Y4K	N1-N	-2.53	1.39	1.41
2	A	301	Y4K	C3-N1	-2.37	1.32	1.38
3	A	302	PO4	P-O4	-2.08	1.48	1.54

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	K	301	Y4K	C2-N-N1	4.55	115.93	111.27
2	I	301	Y4K	C2-N-N1	4.21	115.58	111.27
2	E	301	Y4K	C2-N-N1	4.19	115.56	111.27
2	C	301	Y4K	C2-N-N1	3.95	115.31	111.27
2	G	301	Y4K	C2-N-N1	3.37	114.72	111.27
2	A	301	Y4K	C2-N-N1	3.23	114.58	111.27
2	I	301	Y4K	C8-C3-N1	-2.25	119.06	121.06

There are no chirality outliers.

All (17) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	301	Y4K	C9-C1-C2-N
2	C	301	Y4K	C2-N-N1-C3
2	C	301	Y4K	C9-C1-C2-N
2	G	301	Y4K	C2-N-N1-C3
2	G	301	Y4K	C9-C1-C2-N
2	I	301	Y4K	C2-N-N1-C3
2	I	301	Y4K	C9-C1-C2-N
2	E	301	Y4K	C2-N-N1-C3
2	E	301	Y4K	C9-C1-C2-N
2	K	301	Y4K	C2-N-N1-C3
2	K	301	Y4K	C9-C1-C2-N
2	A	301	Y4K	C-C1-C2-N
2	C	301	Y4K	C-C1-C2-N
2	G	301	Y4K	C-C1-C2-N
2	I	301	Y4K	C-C1-C2-N
2	E	301	Y4K	C-C1-C2-N
2	A	301	Y4K	C2-N-N1-C3

There are no ring outliers.

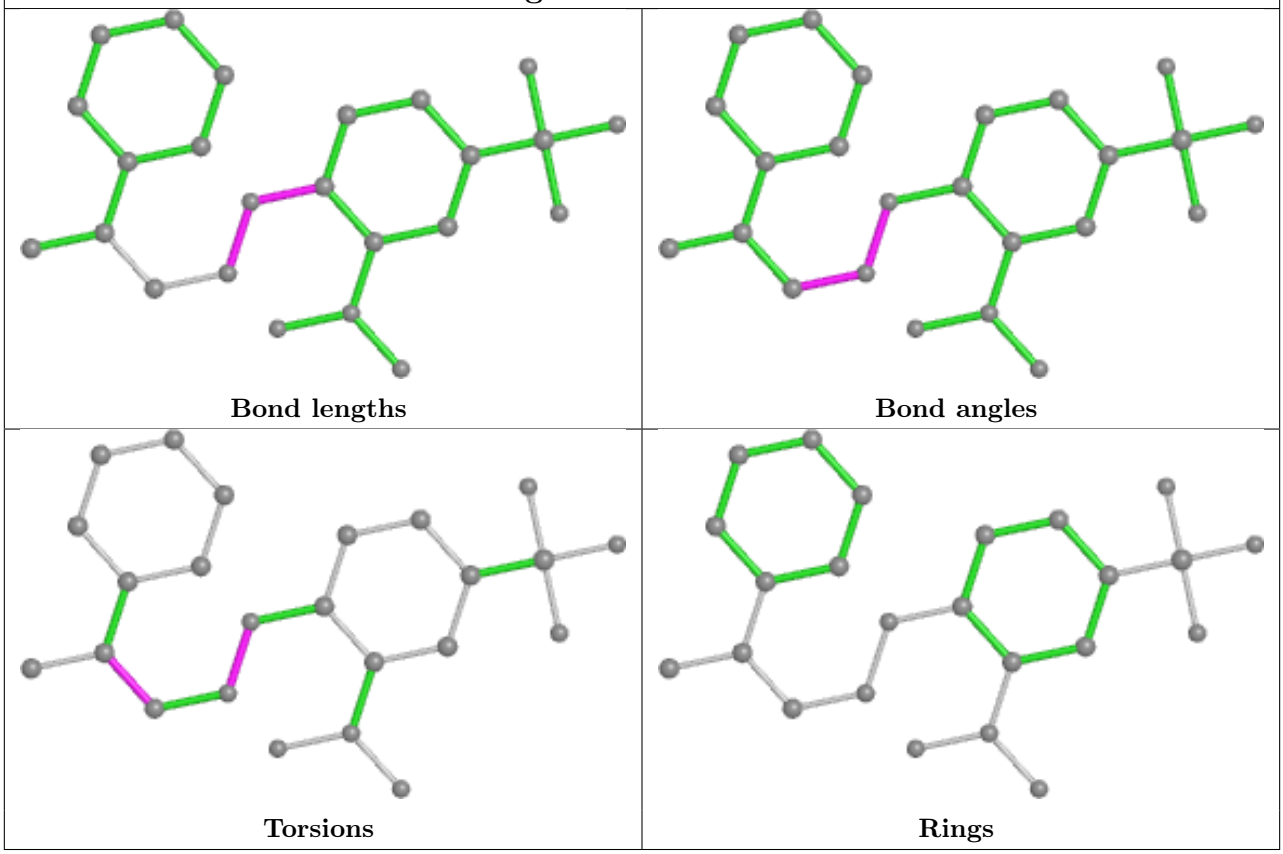
3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	Y4K	1	0
2	C	301	Y4K	1	0
2	K	301	Y4K	1	0

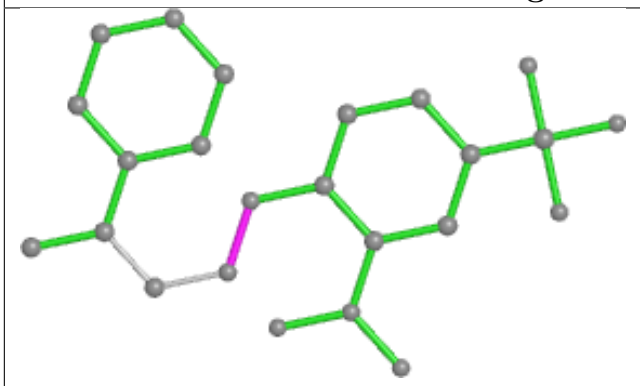
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

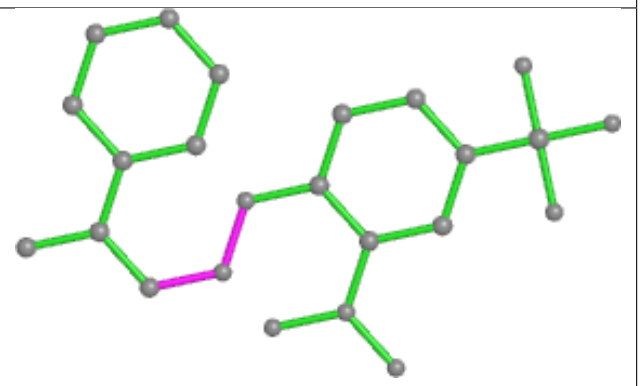
Ligand Y4K A 301



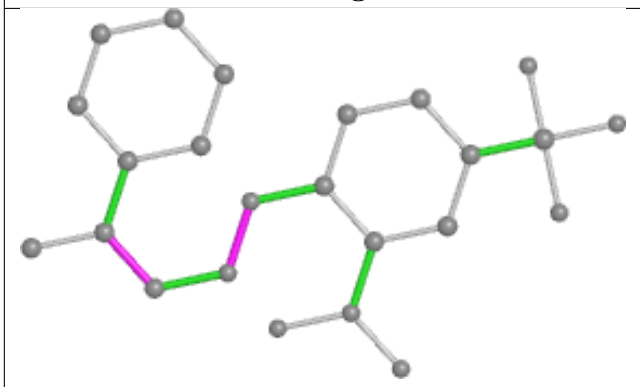
Ligand Y4K E 301



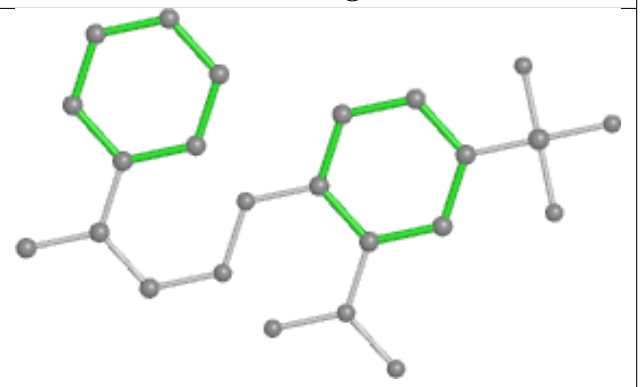
Bond lengths



Bond angles

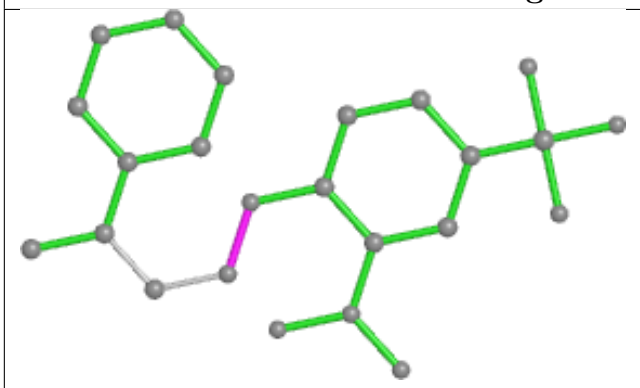


Torsions

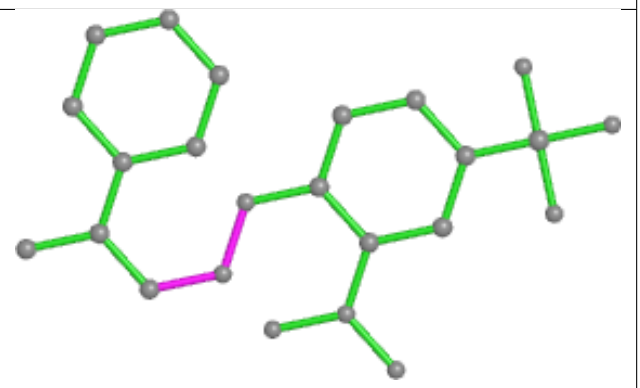


Rings

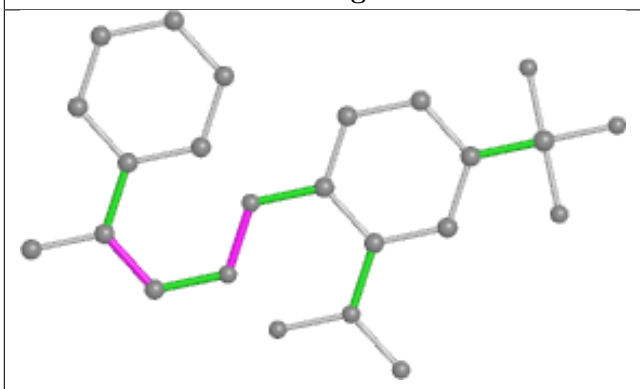
Ligand Y4K C 301



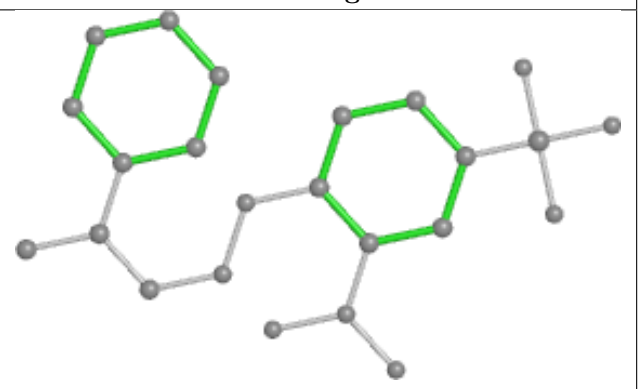
Bond lengths



Bond angles

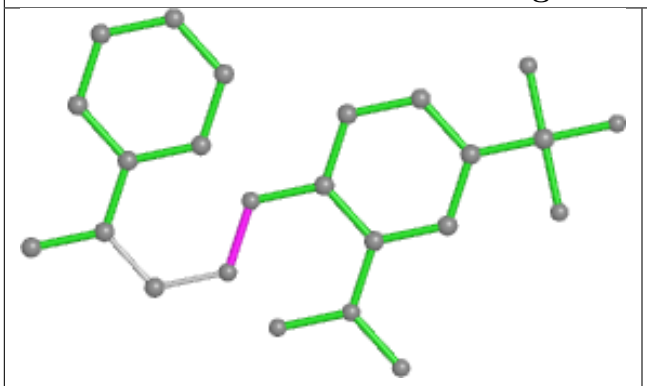


Torsions

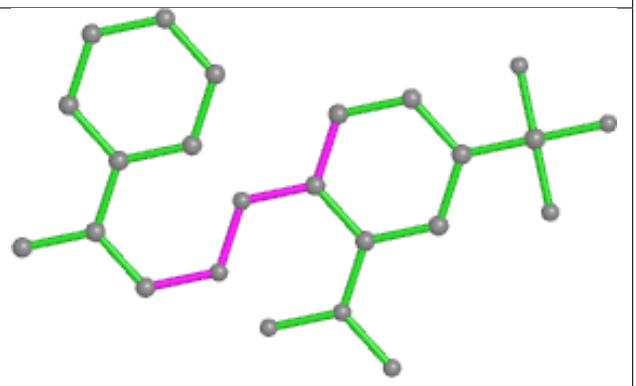


Rings

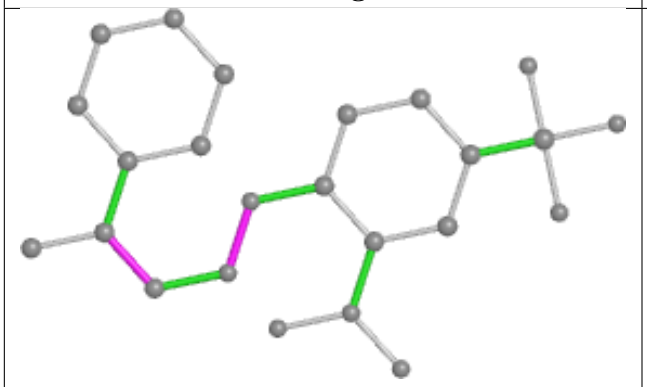
Ligand Y4K I 301



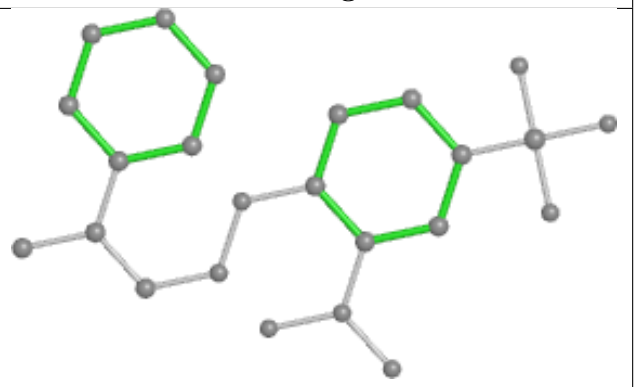
Bond lengths



Bond angles

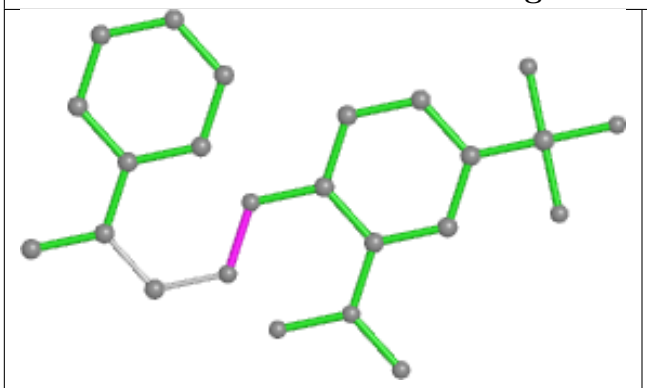


Torsions

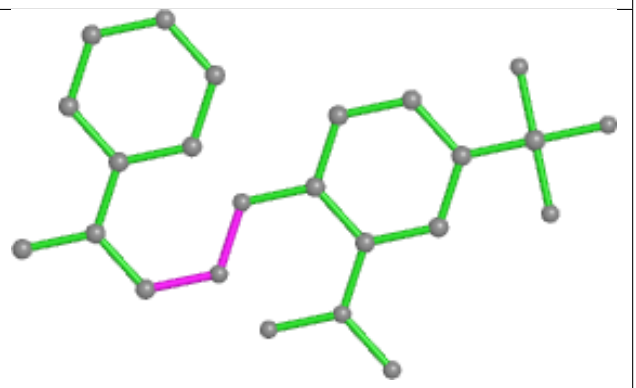


Rings

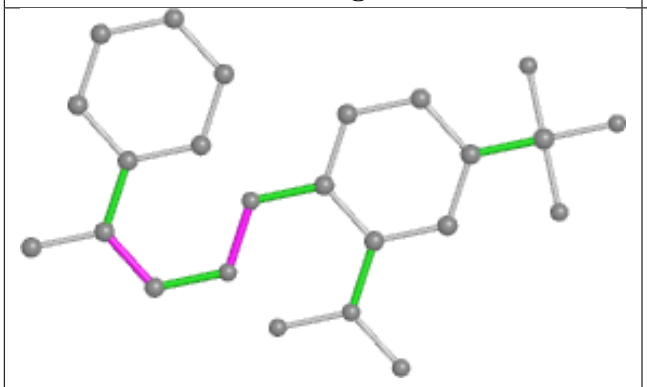
Ligand Y4K G 301



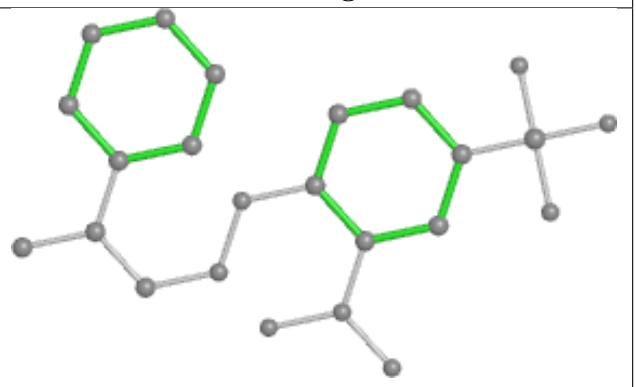
Bond lengths



Bond angles

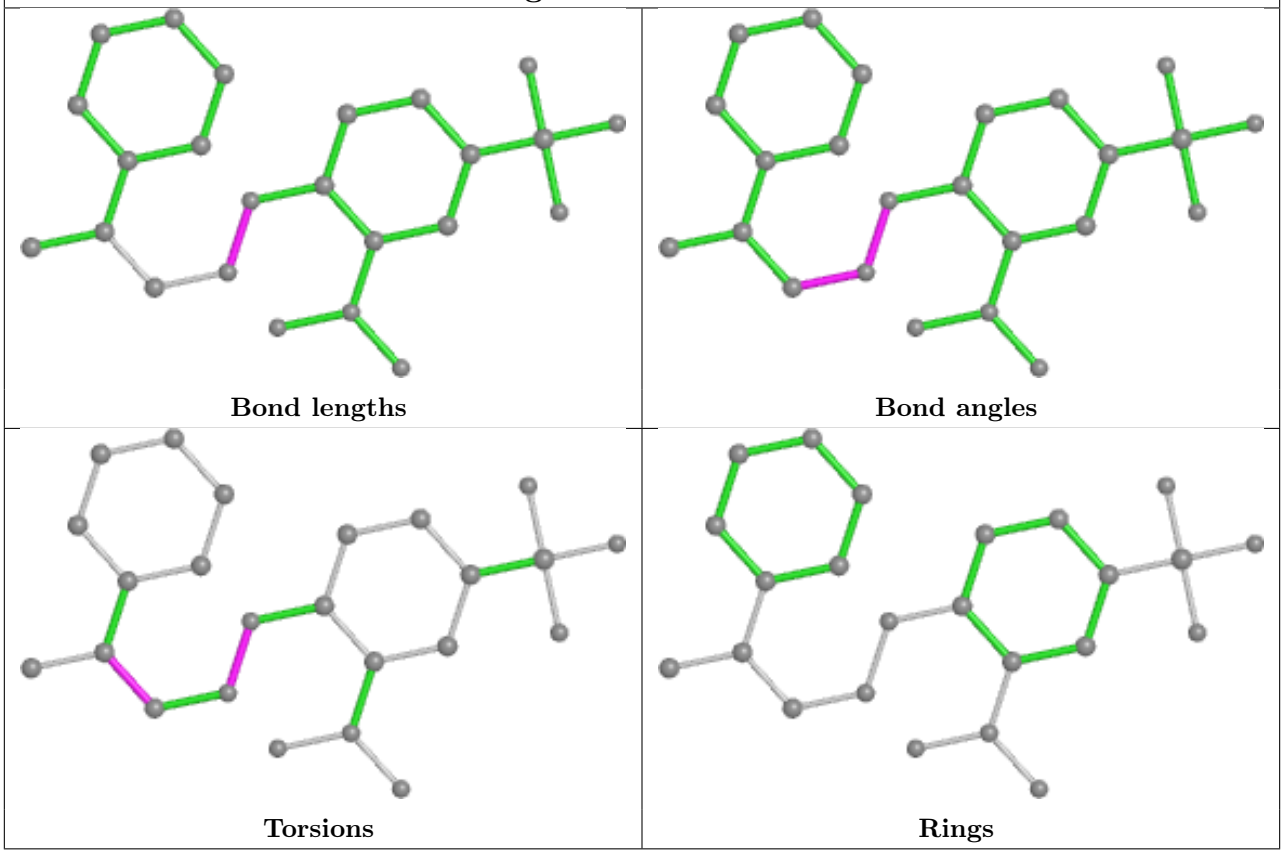


Torsions



Rings

Ligand Y4K K 301



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	215/216 (99%)	0.22	7 (3%)	46	45	15, 25, 44, 86	0
1	B	215/216 (99%)	0.05	5 (2%)	60	59	13, 27, 48, 79	0
1	C	214/216 (99%)	0.23	7 (3%)	46	45	15, 26, 52, 85	0
1	D	215/216 (99%)	0.04	3 (1%)	75	74	14, 27, 45, 83	0
1	E	209/216 (96%)	0.27	7 (3%)	46	45	21, 34, 60, 77	0
1	F	213/216 (98%)	0.31	9 (4%)	36	35	22, 36, 59, 75	0
1	G	211/216 (97%)	0.35	12 (5%)	23	23	23, 36, 73, 86	0
1	H	212/216 (98%)	0.12	7 (3%)	46	45	21, 31, 54, 78	0
1	I	211/216 (97%)	0.25	9 (4%)	35	34	20, 31, 58, 88	0
1	J	214/216 (99%)	0.30	13 (6%)	21	20	22, 35, 61, 85	0
1	K	210/216 (97%)	0.98	33 (15%)	2	1	24, 45, 85, 112	0
1	L	214/216 (99%)	0.46	9 (4%)	36	35	22, 38, 66, 76	0
All	All	2553/2592 (98%)	0.30	121 (4%)	31	30	13, 32, 63, 112	0

All (121) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	146	GLY	6.6
1	A	215	CYS	6.1
1	C	2	SER	6.0
1	A	95	GLY	5.9
1	A	216	SER	5.7
1	F	95	GLY	5.7
1	I	95	GLY	5.6
1	K	155	ASP	5.6
1	K	153	LYS	5.6
1	K	156	SER	5.5
1	C	215	CYS	5.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	H	2	SER	5.2
1	K	184	LEU	5.1
1	K	121	LEU	5.0
1	J	31	GLY	5.0
1	K	97	ASN	4.8
1	F	93	TYR	4.6
1	E	93	TYR	4.6
1	K	157	SER	4.6
1	L	139	LEU	4.6
1	G	191	SER	4.6
1	K	160	LYS	4.4
1	G	2	SER	4.4
1	I	193	ARG	4.3
1	I	160	LYS	4.2
1	I	96	ASP	4.1
1	L	176	TYR	4.0
1	K	159	VAL	4.0
1	K	152	TRP	3.9
1	K	161	ALA	3.9
1	E	2	SER	3.9
1	K	2	SER	3.8
1	D	215	CYS	3.8
1	A	96	ASP	3.8
1	J	26	SER	3.8
1	L	1	GLN	3.8
1	G	190	LYS	3.6
1	J	25	THR	3.6
1	G	160	LYS	3.6
1	K	200	THR	3.5
1	F	1	GLN	3.4
1	L	2	SER	3.4
1	F	2	SER	3.4
1	K	150	VAL	3.4
1	K	182	LEU	3.3
1	K	132	ASN	3.3
1	A	214	GLU	3.2
1	K	186	PRO	3.2
1	C	95	GLY	3.2
1	B	214	GLU	3.2
1	G	155	ASP	3.1
1	B	94	GLY	3.1
1	J	96	ASP	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	H	95	GLY	3.0
1	K	129	LEU	3.0
1	J	1	GLN	3.0
1	C	160	LYS	3.0
1	K	136	LEU	3.0
1	B	2	SER	3.0
1	A	2	SER	3.0
1	J	28	ASP	2.9
1	C	159	VAL	2.9
1	J	95	GLY	2.9
1	F	94	GLY	2.9
1	C	97	ASN	2.8
1	G	156	SER	2.8
1	K	185	THR	2.8
1	F	34	SER	2.8
1	K	181	TYR	2.8
1	E	27	SER	2.8
1	K	194	SER	2.8
1	K	163	VAL	2.7
1	J	24	GLY	2.7
1	B	215	CYS	2.7
1	D	214	GLU	2.7
1	L	3	ALA	2.7
1	B	1	GLN	2.6
1	H	213	THR	2.6
1	H	129	LEU	2.6
1	A	97	ASN	2.6
1	H	126	SER	2.6
1	G	97	ASN	2.6
1	I	97	ASN	2.6
1	K	29	VAL	2.6
1	G	187	GLU	2.6
1	G	193	ARG	2.6
1	I	93	TYR	2.6
1	K	96	ASP	2.6
1	I	157	SER	2.6
1	F	30	GLY	2.5
1	E	183	SER	2.5
1	G	212	PRO	2.5
1	J	30	GLY	2.5
1	K	134	ALA	2.5
1	C	96	ASP	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	F	27	SER	2.4
1	J	12	GLY	2.4
1	K	154	ALA	2.4
1	H	187	GLU	2.4
1	L	144	TYR	2.4
1	J	16	GLN	2.3
1	K	131	ALA	2.3
1	K	162	GLY	2.3
1	F	213	THR	2.2
1	D	95	GLY	2.2
1	J	27	SER	2.2
1	L	187	GLU	2.2
1	I	94	GLY	2.1
1	K	126	SER	2.1
1	K	183	SER	2.1
1	E	193	ARG	2.1
1	H	130	GLN	2.1
1	E	70	GLY	2.1
1	I	33	ASP	2.1
1	G	161	ALA	2.1
1	E	157	SER	2.1
1	L	183	SER	2.1
1	G	159	VAL	2.1
1	K	203	GLY	2.1
1	K	199	VAL	2.0
1	J	32	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)

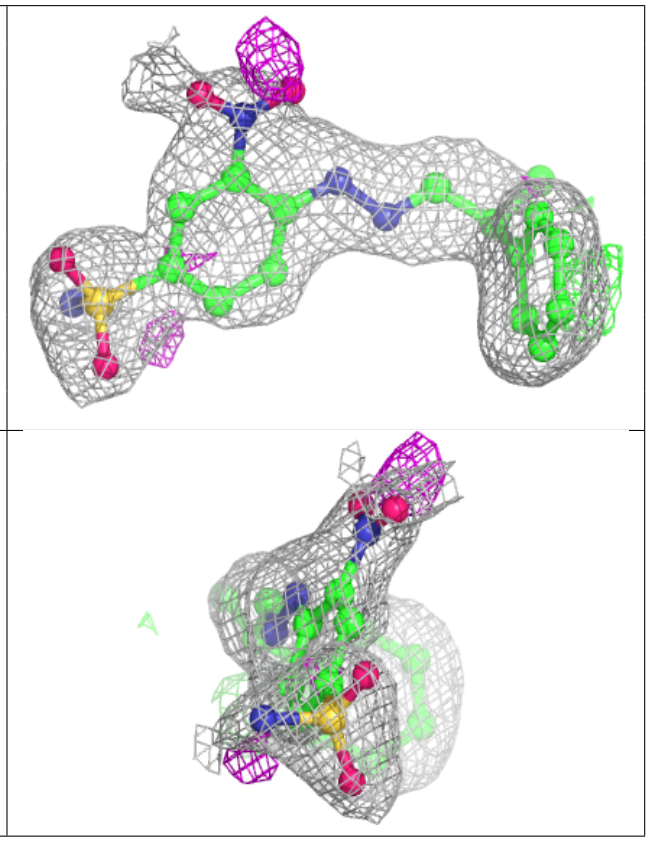
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	Y4K	E	301	24/24	0.85	0.17	33,40,48,63	0
2	Y4K	I	301	24/24	0.90	0.15	23,35,42,46	0
2	Y4K	A	301	24/24	0.90	0.14	20,27,37,44	0
2	Y4K	K	301	24/24	0.91	0.14	23,35,43,51	0
2	Y4K	G	301	24/24	0.92	0.15	21,34,44,54	0
3	PO4	K	302	5/5	0.92	0.13	47,54,58,59	0
2	Y4K	C	301	24/24	0.93	0.13	23,30,39,47	0
3	PO4	B	301	5/5	0.98	0.16	35,35,40,49	0
3	PO4	G	302	5/5	0.98	0.10	42,42,44,50	0
3	PO4	H	301	5/5	0.98	0.10	40,42,43,45	0
3	PO4	A	302	5/5	0.98	0.14	35,36,39,40	0
3	PO4	E	302	5/5	0.99	0.11	37,38,40,40	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.

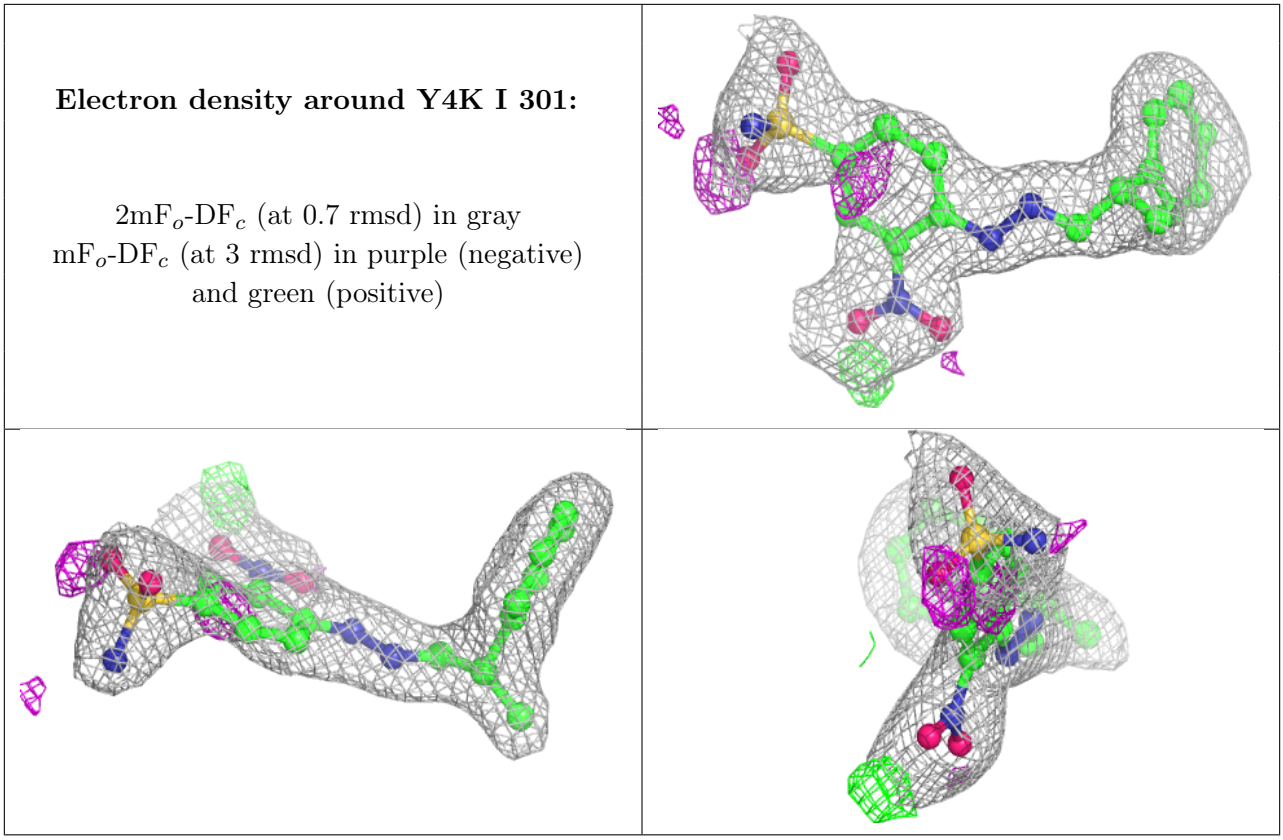
Electron density around Y4K E 301:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)



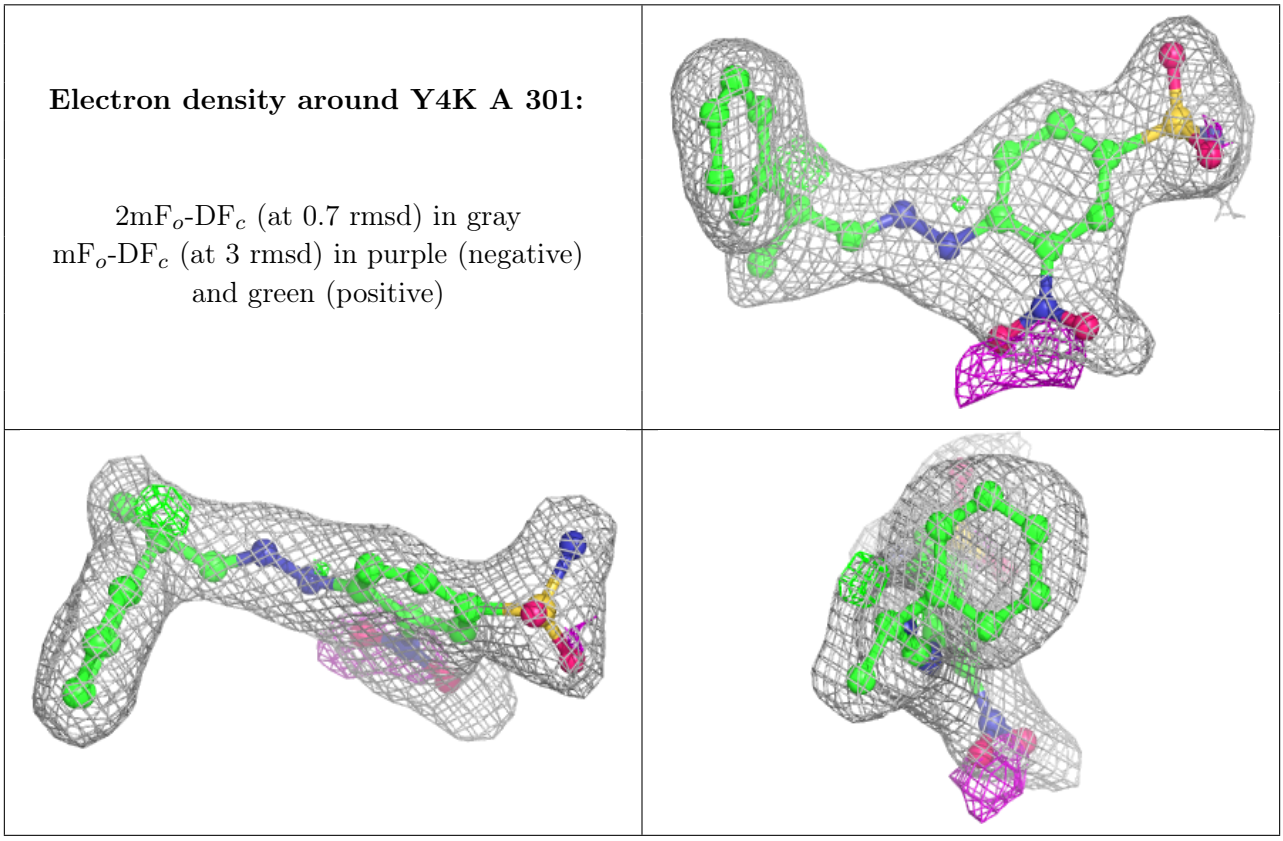
Electron density around Y4K I 301:

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



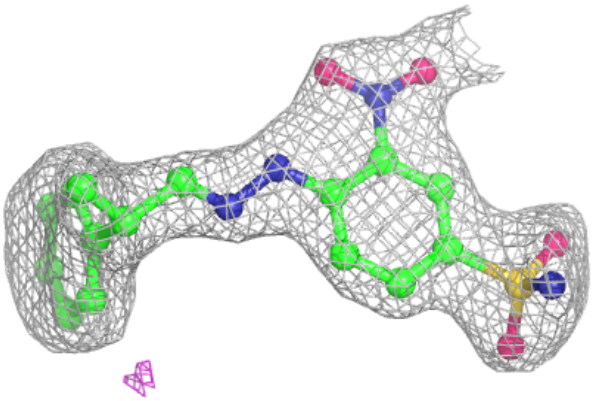
Electron density around Y4K A 301:

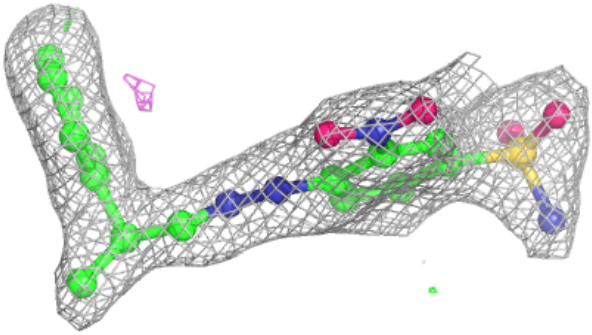
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

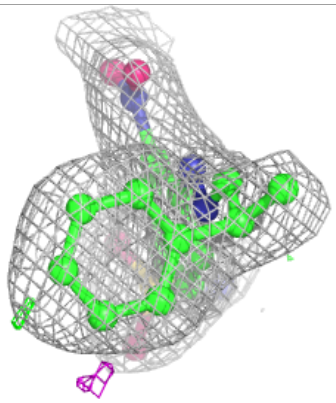


Electron density around Y4K K 301:

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

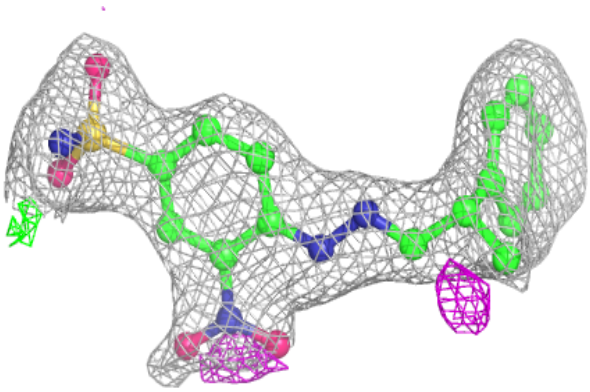


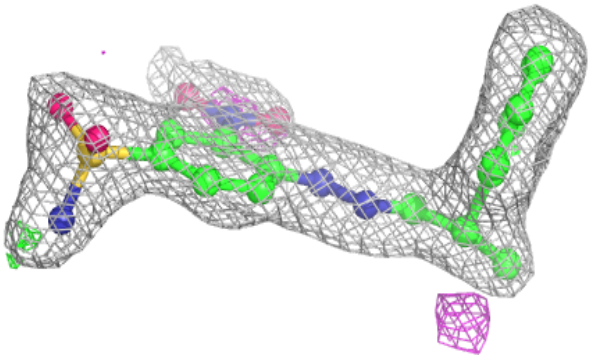


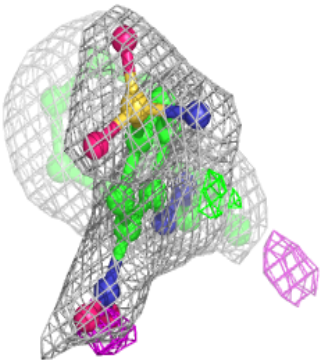


Electron density around Y4K G 301:

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

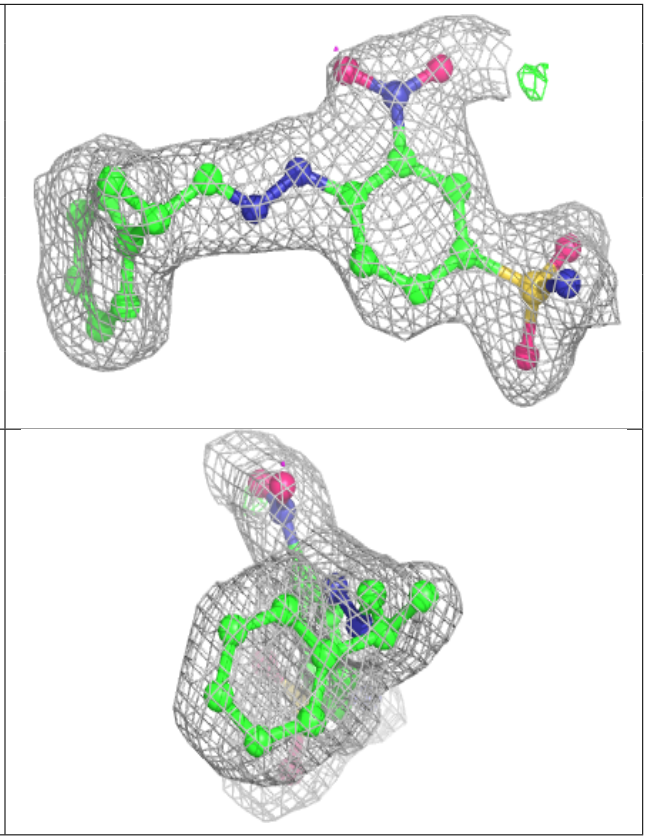






Electron density around Y4K C 301:

$2mF_o$-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.