Full wwPDB X-ray Structure Validation Report (i)

May 14, 2020 – 11:33 pm BST

PDB ID	:	6MVO
Title	:	HCV NS5B 1A Y316 bound to Compound 49
Authors	:	Williams, S.P.; Kahler, K.; Price, D.J.; Peat, A.J.
Deposited on	:	2018-10-26
Resolution	:	1.95 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 following versions of software and data (see references (i)) were used in the production of this report:

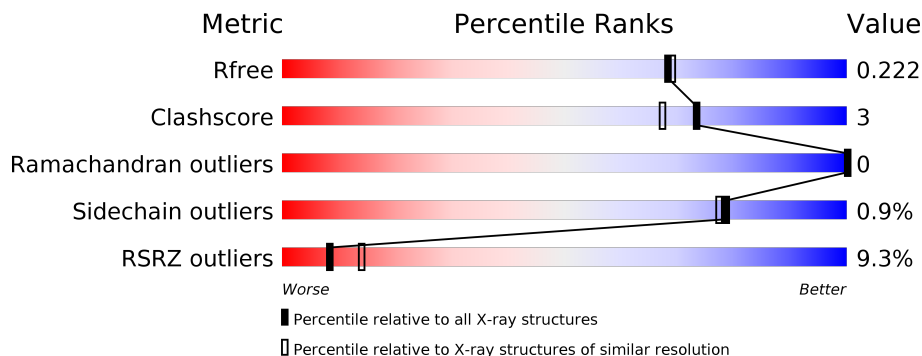|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.95 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	Percentile Ranks	Value
Rfree		0.222
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		0.9%
RSRZ outliers		9.3%

Worse — Better

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=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	562	7% / 93% / 7%
1	B	562	11% / 93% / 7%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 9065 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 RNA-directed RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	562	4298	2711	761	795	31	0	1	0
1	B	562	Total	C	N	O	S	0	3	0
			4257	2690	745	791	31			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	98	LYS	ARG	conflict	UNP Q66N85
A	101	TYR	PHE	engineered mutation	UNP Q66N85
A	110	SER	CYS	engineered mutation	UNP Q66N85
A	113	SER	ARG	engineered mutation	UNP Q66N85
A	114	ARG	LYS	engineered mutation	UNP Q66N85
A	184	LEU	VAL	conflict	UNP Q66N85
A	316	TYR	CYS	engineered mutation	UNP Q66N85
A	546	ALA	ASP	conflict	UNP Q66N85
B	98	LYS	ARG	conflict	UNP Q66N85
B	101	TYR	PHE	engineered mutation	UNP Q66N85
B	110	SER	CYS	engineered mutation	UNP Q66N85
B	113	SER	ARG	engineered mutation	UNP Q66N85
B	114	ARG	LYS	engineered mutation	UNP Q66N85
B	184	LEU	VAL	conflict	UNP Q66N85
B	316	TYR	CYS	engineered mutation	UNP Q66N85
B	546	ALA	ASP	conflict	UNP Q66N85

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

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

- Molecule 3 is 6-[(7-chloro-1-hydroxy-1,3-dihydro-2,1-benzoxaborol-5-yl)(methylsulfonyl)am ino]-5-cyclopropyl-2-(4-fluorophenyl)-N-methyl-1-benzofuran-3-carboxamide (three-letter code: K4P) (formula: $C_{27}H_{23}BClFN_2O_6S$) (labeled as "Ligand of Interest" by author).
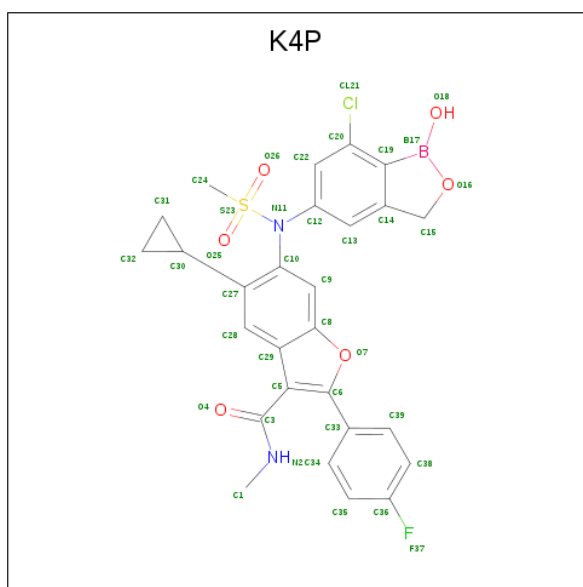
K4P

Mol	Chain	Residues	Atoms								ZeroOcc	AltConf
3	A	1	Total	B	C	Cl	F	N	O	S	0	0
			39	1	27	1	1	2	6	1		
3	B	1	Total	B	C	Cl	F	N	O	S	0	0
			39	1	27	1	1	2	6	1		

• Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	213	Total	O	0	0
			213	213		
4	B	184	Total	O	0	0
			184	184		

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: RNA-directed RNA polymerase
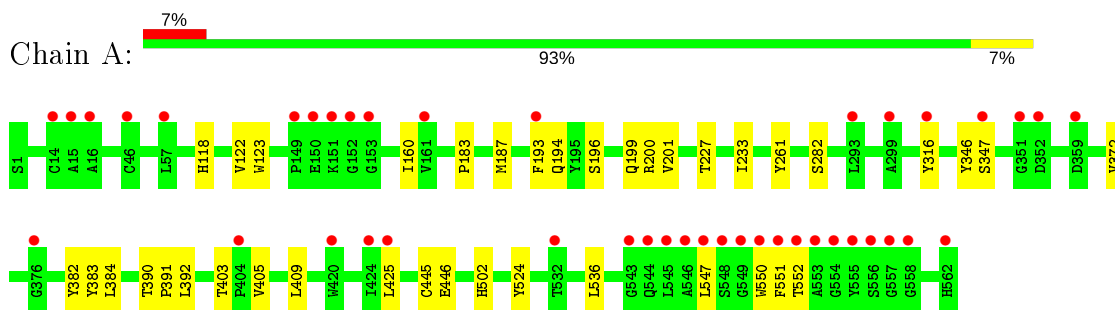
- Molecule 1: RNA-directed RNA polymerase
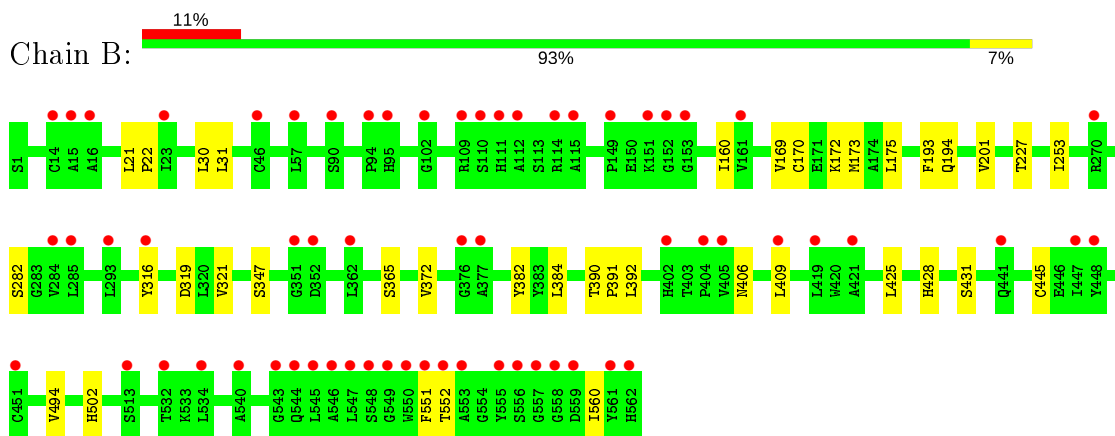
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	53.23Å 61.20Å 91.77Å 89.70° 86.98° 80.92°	Depositor
Resolution (Å)	29.28 – 1.95 29.28 – 1.95	Depositor EDS
% Data completeness (in resolution range)	97.8 (29.28-1.95) 97.9 (29.28-1.95)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	4.25 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.6.0081	Depositor
R, R_{free}	0.217 , 0.235 0.205 , 0.222	Depositor DCC
R_{free} test set	5906 reflections (7.24%)	wwPDB-VP
Wilson B-factor (Å^2)	21.7	Xtriage
Anisotropy	0.256	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 45.5	EDS
L-test for twinning[2]	$<|L|> = 0.49, <L^2> = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9065	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.38% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: K4P, SO4

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	0.47	1/4397 (0.0%)	0.58	0/5980
1	B	0.47	0/4363	0.57	0/5943
All	All	0.47	1/8760 (0.0%)	0.58	0/11923

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	123	TRP	CD2-CE2	5.01	1.47	1.41

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	4298	0	4256	28	0
1	B	4257	0	4168	28	0
2	A	15	0	0	0	0
2	B	20	0	0	0	0
3	A	39	0	0	0	0
3	B	39	0	0	0	0
4	A	213	0	0	3	0
4	B	184	0	0	0	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	9065	0	8424	56	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 3.

All (56) 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:B:193:PHE:CE1	1:B:316:TYR:CE1	2.11	1.39
1:A:193:PHE:CE1	1:A:316:TYR:CE1	2.17	1.32
1:B:170:CYS:HA	1:B:173:MET:HE3	1.33	1.05
1:B:193:PHE:CE1	1:B:316:TYR:HE1	1.73	0.99
1:A:193:PHE:CE1	1:A:316:TYR:HE1	1.83	0.94
1:B:193:PHE:CD1	1:B:316:TYR:HE1	1.87	0.91
1:A:403:THR:CG2	4:A:899:HOH:O	2.17	0.90
1:A:160:ILE:HD12	1:A:282:SER:OG	1.71	0.90
1:A:193:PHE:CD1	1:A:316:TYR:HE1	1.93	0.86
1:B:193:PHE:CD1	1:B:316:TYR:CE1	2.62	0.86
1:A:193:PHE:CD1	1:A:316:TYR:CE1	2.65	0.83
1:B:319:ASP:OD1	1:B:365:SER:OG	1.98	0.81
1:B:193:PHE:CZ	1:B:316:TYR:CE1	2.70	0.80
1:B:201:VAL:HG22	1:B:384:LEU:HG	1.70	0.72
1:A:193:PHE:CZ	1:A:316:TYR:CE1	2.76	0.71
1:A:409:LEU:HD23	1:A:445:CYS:SG	2.32	0.70
1:B:170:CYS:CA	1:B:173:MET:HE3	2.20	0.68
1:B:193:PHE:CE1	1:B:316:TYR:CZ	2.78	0.66
1:A:193:PHE:CE1	1:A:316:TYR:CZ	2.80	0.66
1:B:201:VAL:CG2	1:B:384:LEU:HG	2.29	0.61
1:B:193:PHE:HE1	1:B:316:TYR:CE1	2.09	0.61
1:B:392:LEU:HD23	1:B:425:LEU:HD21	1.84	0.58
1:A:425:LEU:HD13	4:A:832:HOH:O	2.03	0.57
1:A:227:THR:HB	1:A:347:SER:O	2.04	0.56
1:B:409:LEU:HD23	1:B:445:CYS:SG	2.45	0.55
1:B:193:PHE:CZ	1:B:316:TYR:CD1	2.94	0.55
1:B:227:THR:HB	1:B:347:SER:O	2.06	0.55
1:A:403:THR:HG22	4:A:899:HOH:O	1.97	0.54
1:B:193:PHE:CE1	1:B:316:TYR:CD1	2.89	0.54
1:B:172:LYS:HE3	1:B:560:ILE:HD13	1.91	0.53
1:B:194:GLN:HA	1:B:551:PHE:O	2.10	0.52
1:A:193:PHE:HE1	1:A:316:TYR:CE1	2.16	0.51
1:A:194:GLN:HA	1:A:551:PHE:O	2.11	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:196:SER:H	1:A:199:GLN:HE21	1.59	0.49
1:A:193:PHE:CZ	1:A:316:TYR:CD1	3.00	0.49
1:B:175:LEU:HD21	1:B:253:ILE:HG12	1.94	0.49
1:B:428:HIS:O	1:B:431[B]:SER:OG	2.22	0.49
1:A:200:ARG:HD2	1:A:384:LEU:HD11	1.95	0.48
1:B:372:VAL:HG22	1:B:382:TYR:CD1	2.48	0.48
1:A:392:LEU:HD23	1:A:425:LEU:HD21	1.97	0.47
1:A:200:ARG:HD2	1:A:384:LEU:CD1	2.46	0.45
1:B:160:ILE:HD12	1:B:282:SER:OG	2.18	0.44
1:B:30:LEU:O	1:B:494:VAL:HG22	2.19	0.43
1:A:372:VAL:HG22	1:A:382:TYR:CD1	2.54	0.42
1:A:118:HIS:O	1:A:122:VAL:HG23	2.19	0.42
1:A:183:PRO:HB3	1:A:187:MET:CE	2.49	0.42
1:B:390:THR:HB	1:B:391:PRO:HD3	2.02	0.41
1:A:390:THR:HB	1:A:391:PRO:HD3	2.01	0.41
1:A:524:TYR:CD2	1:A:536:LEU:HD22	2.55	0.41
1:B:21:LEU:HD12	1:B:22:PRO:HD2	2.01	0.41
1:A:346:TYR:O	1:A:347:SER:HB3	2.20	0.41
1:A:547:LEU:O	1:A:550:TRP:HB2	2.21	0.41
1:A:201:VAL:HG21	1:A:383:TYR:HA	2.02	0.41
1:B:169:VAL:HG12	1:B:173:MET:HE2	2.03	0.41
1:B:321:VAL:CG2	1:B:365:SER:HB2	2.51	0.41
1:A:233:ILE:HD13	1:A:261:TYR: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	561/562 (100%)	557 (99%)	4 (1%)	0	100	100
1	B	563/562 (100%)	559 (99%)	4 (1%)	0	100	100

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1124/1124 (100%)	1116 (99%)	8 (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	454/467 (97%)	450 (99%)	4 (1%)	78 77
1	B	444/467 (95%)	440 (99%)	4 (1%)	78 77
All	All	898/934 (96%)	890 (99%)	8 (1%)	78 77

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	405	VAL
1	A	446	GLU
1	A	502	HIS
1	A	552	THR
1	B	31	LEU
1	B	406	ASN
1	B	502	HIS
1	B	552	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	35	ASN
1	A	199	GLN
1	A	441	GLN
1	B	35	ASN
1	B	142	ASN
1	B	406	ASN
1	B	411	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	438	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 carbohydrates in this entry.

5.6   Ligand geometry (i)

9 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	$\#|Z| > 2$
2	SO4	B	602	-	4,4,4	0.32	0	6,6,6	0.06	0
2	SO4	A	602	-	4,4,4	0.33	0	6,6,6	0.07	0
2	SO4	A	601	-	4,4,4	0.32	0	6,6,6	0.08	0
2	SO4	B	604	-	4,4,4	0.34	0	6,6,6	0.14	0
3	K4P	A	604	-	36,44,44	2.54	7 (19%)	50,68,68	2.91	16 (32%)
2	SO4	B	603	-	4,4,4	0.36	0	6,6,6	0.05	0
2	SO4	A	603	-	4,4,4	0.33	0	6,6,6	0.05	0
3	K4P	B	605	-	36,44,44	2.61	7 (19%)	50,68,68	2.73	16 (32%)
2	SO4	B	601	-	4,4,4	0.27	0	6,6,6	0.15	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
3	K4P	B	605	-	-	6/22/39/39	0/6/6/6
3	K4P	A	604	-	-	5/22/39/39	0/6/6/6

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	605	K4P	B17-O18	12.76	1.53	1.35
3	A	604	K4P	B17-O18	12.28	1.53	1.35
3	A	604	K4P	C24-S23	-4.17	1.66	1.75
3	B	605	K4P	C24-S23	-4.10	1.66	1.75
3	B	605	K4P	B17-O16	3.86	1.44	1.39
3	A	604	K4P	B17-O16	3.56	1.44	1.39
3	A	604	K4P	C28-C27	2.98	1.41	1.36
3	B	605	K4P	C20-CL21	2.81	1.80	1.73
3	A	604	K4P	C29-C8	-2.80	1.37	1.43
3	B	605	K4P	C28-C27	2.79	1.41	1.36
3	B	605	K4P	C29-C8	-2.69	1.37	1.43
3	A	604	K4P	C20-CL21	2.55	1.79	1.73
3	B	605	K4P	C27-C30	-2.53	1.47	1.52
3	A	604	K4P	C27-C30	-2.03	1.48	1.52

All (32) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	604	K4P	O25-S23-N11	9.51	116.73	107.16
3	B	605	K4P	O25-S23-N11	9.02	116.23	107.16
3	A	604	K4P	C9-C10-N11	-8.17	109.70	119.69
3	B	605	K4P	O25-S23-O26	-7.07	108.51	118.59
3	A	604	K4P	O25-S23-O26	-6.39	109.48	118.59
3	B	605	K4P	C9-C10-N11	-6.27	112.02	119.69
3	A	604	K4P	C22-C12-N11	5.84	125.64	119.66
3	B	605	K4P	C5-C6-C33	4.94	134.19	127.13
3	A	604	K4P	C5-C6-C33	4.89	134.12	127.13
3	A	604	K4P	O16-C15-C14	-4.79	103.39	105.79
3	A	604	K4P	C5-C29-C8	-4.57	100.58	107.77
3	B	605	K4P	O16-C15-C14	-4.41	103.59	105.79
3	B	605	K4P	C5-C29-C8	-4.33	100.96	107.77
3	A	604	K4P	C13-C12-N11	-4.23	115.33	119.66
3	B	605	K4P	B17-C19-C14	4.11	107.47	105.19
3	A	604	K4P	C10-N11-S23	-4.04	112.84	118.57

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	605	K4P	C22-C12-N11	3.95	123.71	119.66
3	B	605	K4P	C10-N11-S23	-3.81	113.17	118.57
3	B	605	K4P	O16-B17-C19	-3.47	104.17	108.06
3	A	604	K4P	O16-B17-C19	-3.27	104.39	108.06
3	B	605	K4P	C32-C30-C27	-3.19	111.26	120.68
3	A	604	K4P	C15-C14-C19	3.04	112.09	110.53
3	A	604	K4P	C32-C30-C27	-2.73	112.62	120.68
3	A	604	K4P	C29-C5-C3	-2.70	121.57	125.83
3	A	604	K4P	C5-C3-N2	2.54	119.27	114.80
3	A	604	K4P	B17-C19-C14	2.51	106.59	105.19
3	B	605	K4P	C15-C14-C19	2.45	111.78	110.53
3	B	605	K4P	B17-O16-C15	2.37	112.36	110.48
3	B	605	K4P	O26-S23-N11	2.34	109.52	107.16
3	B	605	K4P	C29-C5-C3	-2.34	122.13	125.83
3	A	604	K4P	C38-C36-C35	-2.17	119.95	122.83
3	B	605	K4P	C38-C36-C35	-2.16	119.95	122.83

There are no chirality outliers.
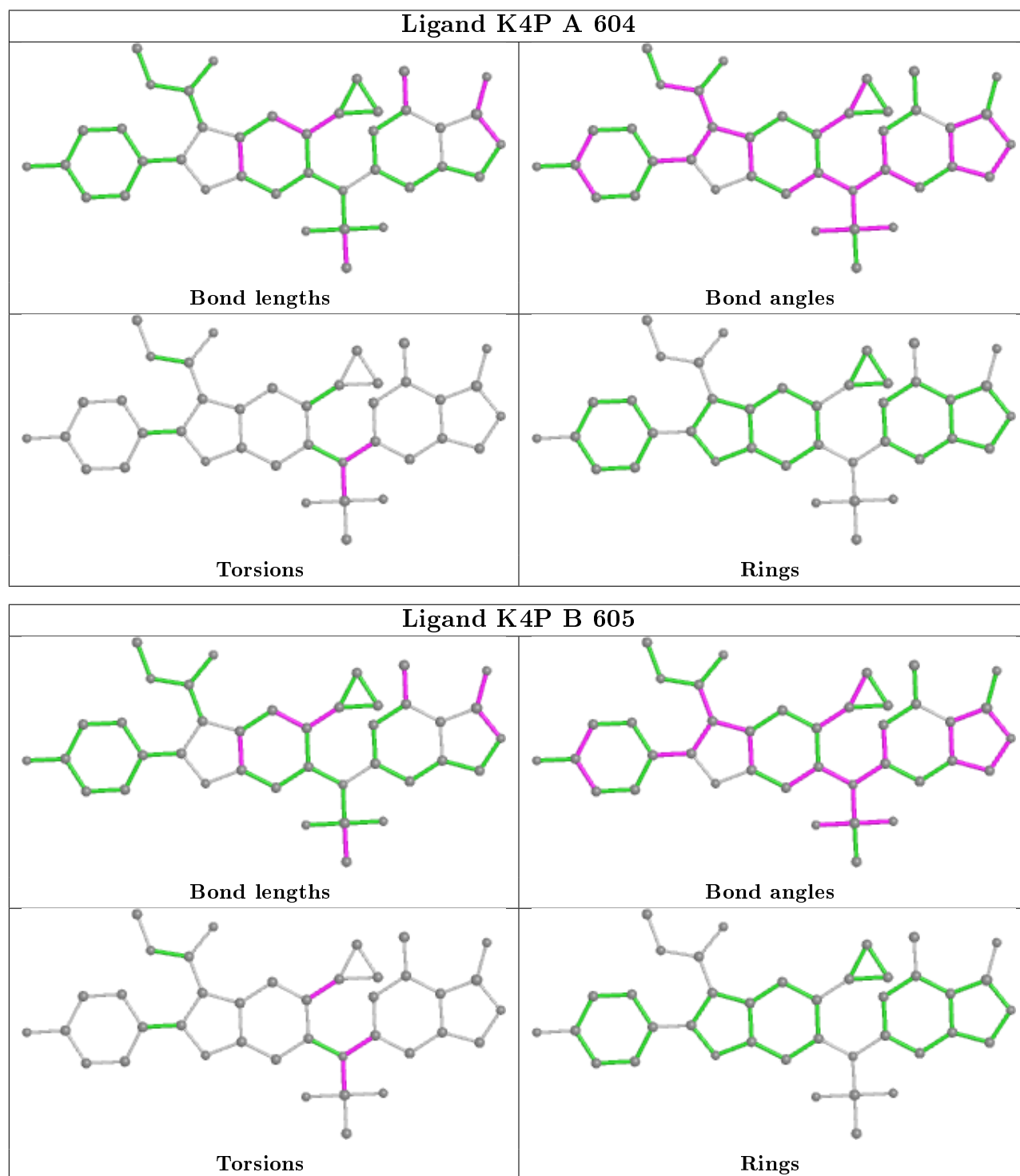
All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	605	K4P	C22-C12-N11-S23
3	A	604	K4P	C22-C12-N11-S23
3	B	605	K4P	C13-C12-N11-S23
3	A	604	K4P	C13-C12-N11-S23
3	A	604	K4P	C10-N11-S23-O25
3	B	605	K4P	C10-N11-S23-O25
3	A	604	K4P	C10-N11-S23-C24
3	B	605	K4P	C10-N11-S23-C24
3	B	605	K4P	C10-C27-C30-C32
3	A	604	K4P	C13-C12-N11-C10
3	B	605	K4P	C13-C12-N11-C10

There are no ring outliers.

No monomer is involved in short contacts.

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 K4P A 604

Bond lengths

Bond angles

Torsions

Rings

Ligand K4P B 605

Bond lengths

Bond angles

Torsions

Rings

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	562/562 (100%)	0.68	42 (7%)	14	22	16, 22, 32, 42	1 (0%)
1	B	562/562 (100%)	0.78	63 (11%)	5	8	17, 23, 36, 44	0
All	All	1124/1124 (100%)	0.73	105 (9%)	8	14	16, 22, 35, 44	1 (0%)

All (105) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	548	SER	8.8
1	B	548	SER	8.0
1	B	552	THR	7.5
1	A	552	THR	7.0
1	A	545	LEU	6.7
1	A	554	GLY	6.5
1	B	547	LEU	6.5
1	A	550	TRP	6.0
1	A	557	GLY	6.0
1	A	558	GLY	6.0
1	B	15	ALA	5.9
1	A	547	LEU	5.8
1	A	553	ALA	5.7
1	B	14	CYS	5.7
1	A	14	CYS	5.4
1	B	153	GLY	5.4
1	B	57	LEU	5.0
1	B	545	LEU	4.9
1	A	549	GLY	4.8
1	B	16	ALA	4.7
1	A	149	PRO	4.7
1	A	150	GLU	4.5
1	A	57	LEU	4.5
1	A	556	SER	4.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	562	HIS	4.5
1	B	151	LYS	4.5
1	B	546	ALA	4.4
1	A	376	GLY	4.4
1	B	149	PRO	4.3
1	A	532	THR	4.3
1	B	558	GLY	4.1
1	B	543	GLY	4.0
1	A	15	ALA	4.0
1	B	553	ALA	4.0
1	A	16	ALA	3.8
1	A	153	GLY	3.8
1	A	543	GLY	3.8
1	A	352	ASP	3.7
1	A	546	ALA	3.7
1	B	549	GLY	3.7
1	A	562	HIS	3.7
1	A	555	TYR	3.7
1	B	352	ASP	3.7
1	B	557	GLY	3.7
1	B	559	ASP	3.5
1	B	550	TRP	3.5
1	B	534	LEU	3.4
1	B	95	HIS	3.4
1	B	532	THR	3.3
1	B	376	GLY	3.3
1	B	556	SER	3.2
1	B	404	PRO	3.2
1	B	555	TYR	3.2
1	B	448	TYR	3.1
1	A	151	LYS	3.1
1	B	46[A]	CYS	3.1
1	B	110	SER	3.1
1	B	513	SER	3.1
1	B	402	HIS	3.1
1	B	561	TYR	3.0
1	A	152	GLY	3.0
1	B	161	VAL	3.0
1	B	102	GLY	2.9
1	B	316	TYR	2.9
1	A	193	PHE	2.9
1	B	114	ARG	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	551	PHE	2.8
1	B	540	ALA	2.7
1	B	152	GLY	2.7
1	A	293	LEU	2.7
1	A	425	LEU	2.7
1	B	377	ALA	2.6
1	B	405	VAL	2.6
1	A	551	PHE	2.5
1	B	111	HIS	2.5
1	A	316	TYR	2.5
1	A	351	GLY	2.5
1	B	90	SER	2.5
1	B	94	PRO	2.4
1	A	544	GLN	2.4
1	B	351	GLY	2.4
1	B	23	ILE	2.4
1	B	419	LEU	2.4
1	B	109	ARG	2.4
1	A	424	ILE	2.4
1	B	112	ALA	2.4
1	A	359	ASP	2.3
1	B	441	GLN	2.3
1	A	161	VAL	2.3
1	A	46[A]	CYS	2.2
1	B	421	ALA	2.2
1	B	284	VAL	2.2
1	A	404	PRO	2.1
1	B	285	LEU	2.1
1	A	347	SER	2.1
1	B	362	LEU	2.1
1	B	447	ILE	2.1
1	B	293	LEU	2.1
1	B	409	LEU	2.1
1	B	451	CYS	2.1
1	B	544	GLN	2.1
1	B	270	ARG	2.0
1	A	299	ALA	2.0
1	A	420	TRP	2.0
1	B	115	ALA	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 carbohydrates in this entry.

6.4 Ligands (i)

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

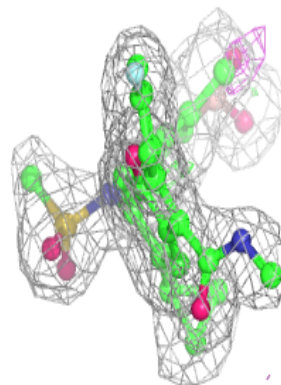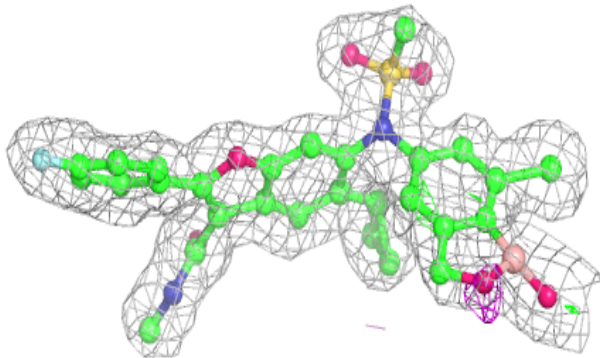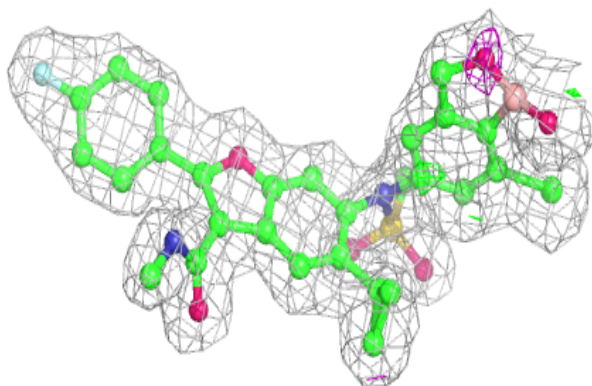| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	SO4	B	604	5/5	0.88	0.20	43,43,44,45	0
2	SO4	B	603	5/5	0.92	0.29	40,41,42,42	0
3	K4P	B	605	39/39	0.92	0.14	25,27,29,30	0
2	SO4	B	602	5/5	0.93	0.16	52,52,53,53	0
3	K4P	A	604	39/39	0.94	0.12	20,21,24,25	0
2	SO4	B	601	5/5	0.94	0.12	26,27,28,28	0
2	SO4	A	601	5/5	0.95	0.19	37,37,38,38	0
2	SO4	A	602	5/5	0.96	0.23	35,36,36,37	0
2	SO4	A	603	5/5	0.97	0.21	42,43,44,44	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 K4P B 605:

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
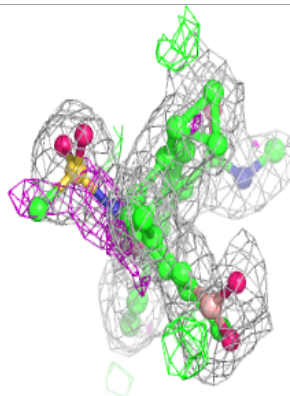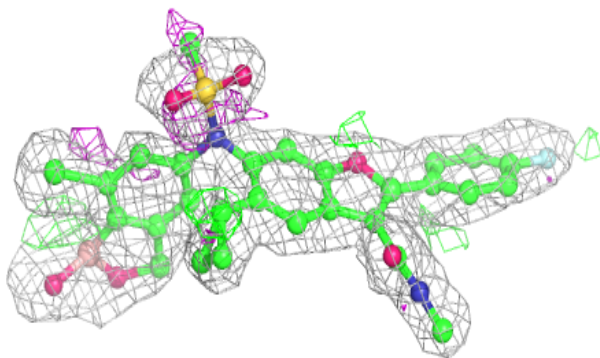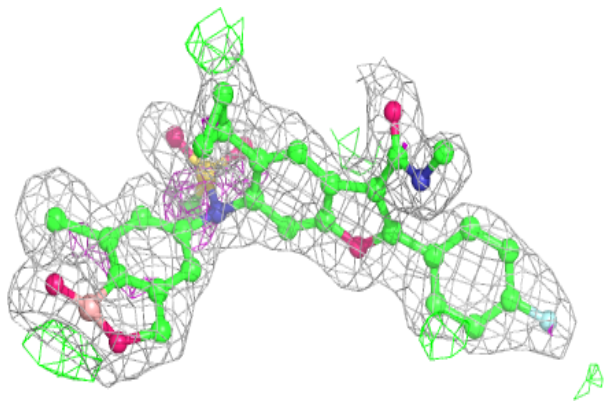
Electron density around K4P A 604:

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

6.5 Other polymers (i)

There are no such residues in this entry.