



Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2020 – 10:36 pm BST

PDB ID : 1O5O
Title : Crystal structure of Uracil phosphoribosyltransferase (TM0721) from *Thermotoga maritima* at 2.30 Å resolution
Authors : Joint Center for Structural Genomics (JCSG)
Deposited on : 2003-10-03
Resolution : 2.30 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 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	:	2.11
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

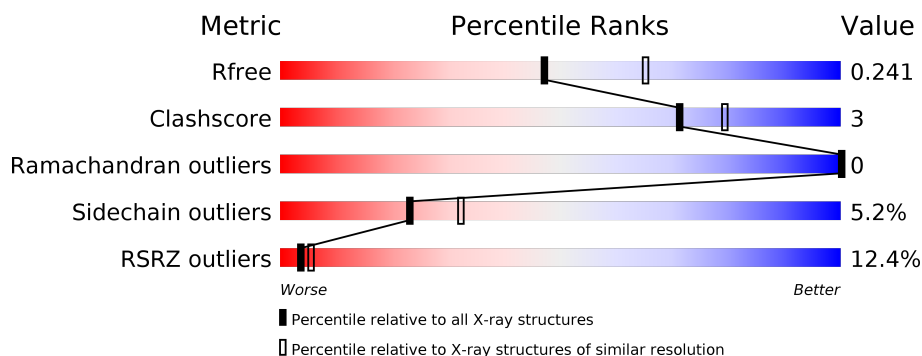
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	221	<div> <div>10%</div> <div> <div></div> <div>81%</div> <div>11%</div> <div>• 5%</div> </div> </div>
1	B	221	<div> <div>15%</div> <div> <div></div> <div>79%</div> <div>15%</div> <div>• 5%</div> </div> </div>
1	C	221	<div> <div>12%</div> <div> <div></div> <div>85%</div> <div>8%</div> <div>• 5%</div> </div> </div>
1	D	221	<div> <div>10%</div> <div> <div></div> <div>81%</div> <div>13%</div> <div>5%</div> </div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7038 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 Uracil phosphoribosyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	210	Total	C	N	O	S	0	0	0
			1636	1055	274	302	5			
1	B	209	Total	C	N	O	S	0	0	0
			1629	1048	273	303	5			
1	C	210	Total	C	N	O	S	0	0	0
			1643	1056	276	306	5			
1	D	210	Total	C	N	O	S	0	0	0
			1646	1058	277	306	5			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	LEADER SEQUENCE	UNP Q9WZI0
A	-10	GLY	-	LEADER SEQUENCE	UNP Q9WZI0
A	-9	SER	-	LEADER SEQUENCE	UNP Q9WZI0
A	-8	ASP	-	LEADER SEQUENCE	UNP Q9WZI0
A	-7	LYS	-	LEADER SEQUENCE	UNP Q9WZI0
A	-6	ILE	-	LEADER SEQUENCE	UNP Q9WZI0
A	-5	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
A	-4	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
A	-3	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
A	-2	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
A	-1	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
A	0	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
B	-11	MET	-	LEADER SEQUENCE	UNP Q9WZI0
B	-10	GLY	-	LEADER SEQUENCE	UNP Q9WZI0
B	-9	SER	-	LEADER SEQUENCE	UNP Q9WZI0
B	-8	ASP	-	LEADER SEQUENCE	UNP Q9WZI0
B	-7	LYS	-	LEADER SEQUENCE	UNP Q9WZI0
B	-6	ILE	-	LEADER SEQUENCE	UNP Q9WZI0
B	-5	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
B	-4	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
B	-3	HIS	-	LEADER SEQUENCE	UNP Q9WZI0

Continued on next page...

Continued from previous page...

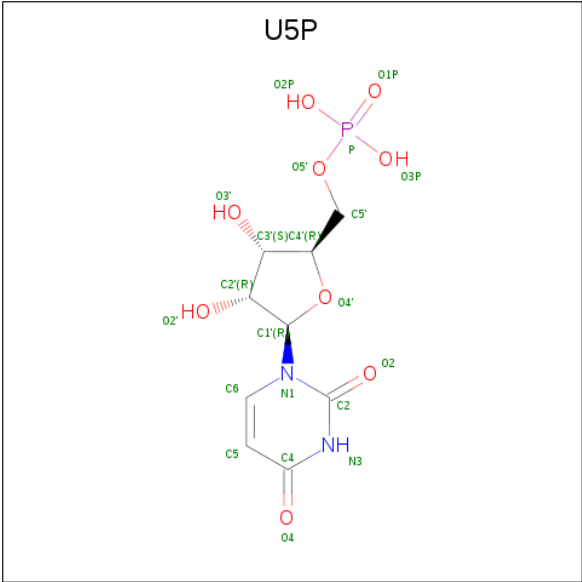
Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
B	-1	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
B	0	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
C	-11	MET	-	LEADER SEQUENCE	UNP Q9WZI0
C	-10	GLY	-	LEADER SEQUENCE	UNP Q9WZI0
C	-9	SER	-	LEADER SEQUENCE	UNP Q9WZI0
C	-8	ASP	-	LEADER SEQUENCE	UNP Q9WZI0
C	-7	LYS	-	LEADER SEQUENCE	UNP Q9WZI0
C	-6	ILE	-	LEADER SEQUENCE	UNP Q9WZI0
C	-5	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
C	-4	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
C	-3	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
C	-2	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
C	-1	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
C	0	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
D	-11	MET	-	LEADER SEQUENCE	UNP Q9WZI0
D	-10	GLY	-	LEADER SEQUENCE	UNP Q9WZI0
D	-9	SER	-	LEADER SEQUENCE	UNP Q9WZI0
D	-8	ASP	-	LEADER SEQUENCE	UNP Q9WZI0
D	-7	LYS	-	LEADER SEQUENCE	UNP Q9WZI0
D	-6	ILE	-	LEADER SEQUENCE	UNP Q9WZI0
D	-5	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
D	-4	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
D	-3	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
D	-2	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
D	-1	HIS	-	LEADER SEQUENCE	UNP Q9WZI0
D	0	HIS	-	LEADER SEQUENCE	UNP Q9WZI0

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



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

- Molecule 3 is URIDINE-5'-MONOPHOSPHATE (three-letter code: U5P) (formula: $C_9H_{13}N_2O_9P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			21	9	2	9	1		
3	B	1	Total	C	N	O	P	0	0
			21	9	2	9	1		
3	C	1	Total	C	N	O	P	0	0
			21	9	2	9	1		
3	D	1	Total	C	N	O	P	0	0
			21	9	2	9	1		

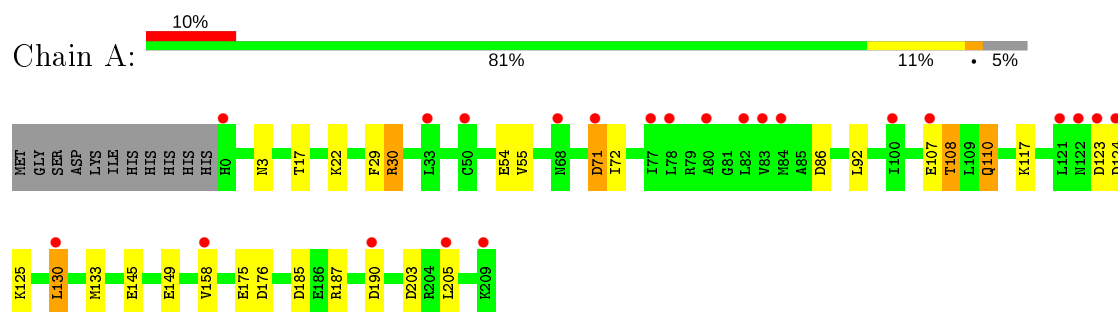
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	67	Total	O	0	0
			67	67		
4	B	94	Total	O	0	0
			94	94		
4	C	85	Total	O	0	0
			85	85		
4	D	109	Total	O	0	0
			109	109		

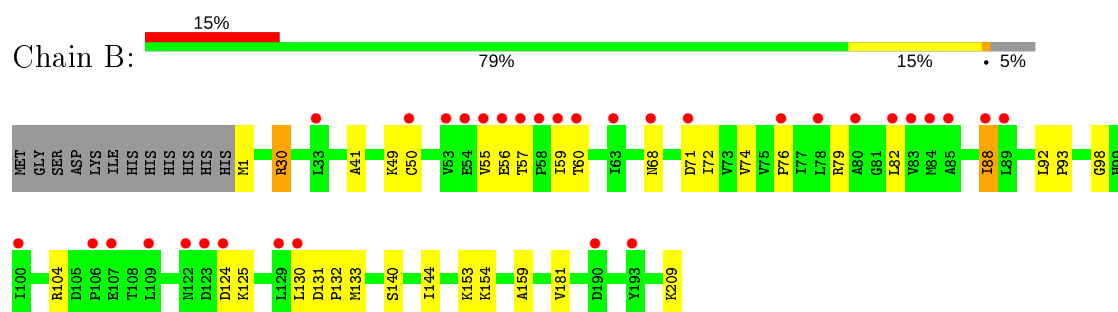
3 Residue-property plots [i](#)

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.

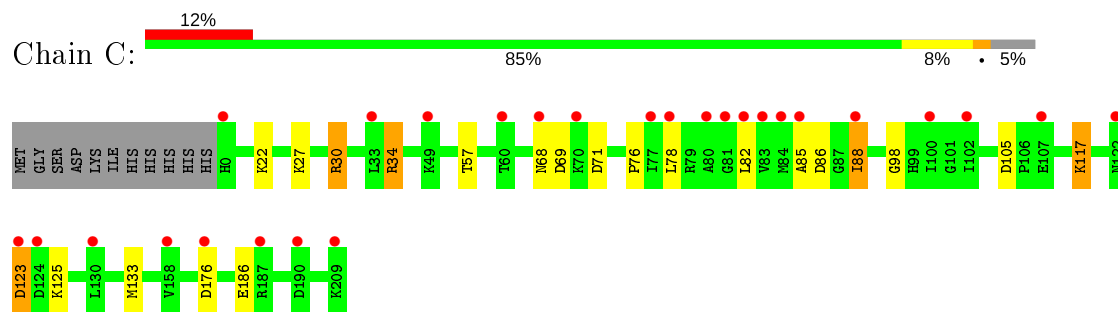
• Molecule 1: Uracil phosphoribosyltransferase



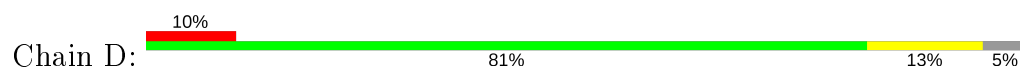
• Molecule 1: Uracil phosphoribosyltransferase

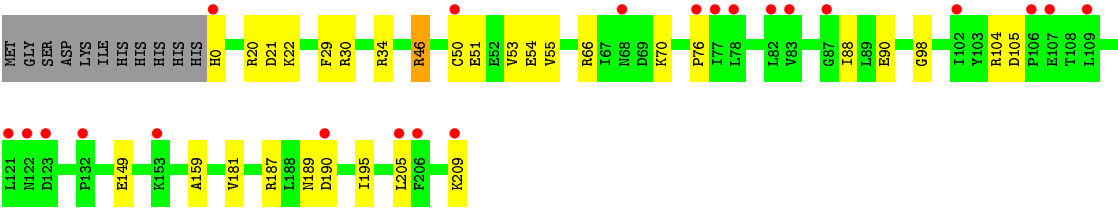


• Molecule 1: Uracil phosphoribosyltransferase



• Molecule 1: Uracil phosphoribosyltransferase





4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	134.08Å 87.41Å 90.83Å 90.00° 115.29° 90.00°	Depositor
Resolution (Å)	30.93 – 2.30 30.93 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.7 (30.93-2.30) 99.7 (30.93-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.29 (at 2.31Å)	Xtriage
Refinement program	REFMAC 5.1.9999	Depositor
R, R_{free}	0.167 , 0.236 0.179 , 0.241	Depositor DCC
R_{free} test set	2126 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	38.5	Xtriage
Anisotropy	0.119	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 40.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7038	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.77% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: U5P, 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	0.82	0/1663	0.97	9/2255 (0.4%)
1	B	0.87	1/1655 (0.1%)	0.96	5/2245 (0.2%)
1	C	0.85	0/1670	0.93	7/2265 (0.3%)
1	D	0.90	1/1673 (0.1%)	0.95	4/2268 (0.2%)
All	All	0.86	2/6661 (0.0%)	0.95	25/9033 (0.3%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	50	CYS	CB-SG	-5.11	1.73	1.81
1	D	50	CYS	CB-SG	-5.11	1.73	1.81

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	30	ARG	NE-CZ-NH2	-9.61	115.50	120.30
1	A	30	ARG	NE-CZ-NH2	-8.77	115.92	120.30
1	D	190	ASP	CB-CG-OD2	8.13	125.62	118.30
1	A	86	ASP	CB-CG-OD2	7.65	125.18	118.30
1	C	176	ASP	CB-CG-OD2	7.11	124.70	118.30
1	B	30	ARG	NE-CZ-NH1	6.99	123.80	120.30
1	A	30	ARG	NE-CZ-NH1	6.89	123.75	120.30
1	C	105	ASP	CB-CG-OD1	6.84	124.46	118.30
1	A	190	ASP	CB-CG-OD2	6.62	124.26	118.30
1	C	86	ASP	CB-CG-OD2	6.26	123.94	118.30
1	D	30	ARG	NE-CZ-NH2	-6.22	117.19	120.30
1	B	71	ASP	CB-CG-OD2	6.15	123.83	118.30
1	C	30	ARG	NE-CZ-NH2	-6.11	117.24	120.30
1	B	104	ARG	NE-CZ-NH2	-6.08	117.26	120.30
1	A	185	ASP	CB-CG-OD2	5.75	123.47	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	71	ASP	CB-CG-OD2	5.51	123.26	118.30
1	A	203	ASP	CB-CG-OD1	5.43	123.19	118.30
1	A	71	ASP	CB-CG-OD2	5.43	123.19	118.30
1	A	124	ASP	CB-CG-OD2	5.40	123.16	118.30
1	C	34	ARG	NE-CZ-NH1	5.33	122.96	120.30
1	D	21	ASP	CB-CG-OD1	5.30	123.07	118.30
1	A	176	ASP	CB-CG-OD2	5.26	123.04	118.30
1	C	30	ARG	NE-CZ-NH1	5.25	122.93	120.30
1	D	105	ASP	CB-CG-OD1	5.17	122.96	118.30
1	B	124	ASP	CB-CG-OD2	5.08	122.87	118.30

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	1636	0	1707	13	0
1	B	1629	0	1698	18	0
1	C	1643	0	1709	9	0
1	D	1646	0	1718	12	0
2	A	15	0	0	0	0
2	B	10	0	0	0	0
2	C	5	0	0	0	0
2	D	15	0	0	0	0
3	A	21	0	11	1	0
3	B	21	0	11	1	0
3	C	21	0	11	1	0
3	D	21	0	11	0	0
4	A	67	0	0	1	0
4	B	94	0	0	3	0
4	C	85	0	0	2	0
4	D	109	0	0	3	0
All	All	7038	0	6876	47	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 (47) 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:D:46:ARG:HH11	1:D:46:ARG:HG3	1.49	0.78
1:D:70:LYS:HB3	4:D:819:HOH:O	1.92	0.69
1:D:51:GLU:OE1	1:D:66:ARG:NH2	2.32	0.62
1:A:108:THR:HG22	1:A:110:GLN:H	1.62	0.61
1:A:17:THR:HG22	1:B:55:VAL:HG11	1.85	0.59
1:A:130:LEU:HD23	1:A:158:VAL:HB	1.85	0.58
1:C:133:MET:HB2	3:C:602:U5P:H2'	1.86	0.58
1:B:59:ILE:HG22	1:B:60:THR:HG23	1.85	0.57
1:B:49:LYS:O	4:B:756:HOH:O	2.18	0.55
4:B:775:HOH:O	1:C:117:LYS:HE3	2.07	0.55
1:D:46:ARG:HH11	1:D:46:ARG:CG	2.18	0.54
1:C:76:PRO:HD2	1:C:98:GLY:O	2.09	0.52
1:B:79:ARG:O	1:B:82:LEU:HG	2.10	0.52
1:A:133:MET:HB2	3:A:600:U5P:H2'	1.90	0.52
1:B:133:MET:HB2	3:B:601:U5P:H2'	1.92	0.51
1:A:110:GLN:HA	1:A:110:GLN:HE21	1.74	0.51
1:C:30:ARG:NH2	4:C:731:HOH:O	2.37	0.50
1:A:29:PHE:HD2	1:A:205:LEU:HD23	1.77	0.50
1:C:78:LEU:HD22	1:C:82:LEU:HD21	1.95	0.48
1:D:189:ASN:HB3	1:D:195:ILE:HD11	1.96	0.47
1:B:159:ALA:O	1:B:181:VAL:HA	2.15	0.47
1:A:117:LYS:HE3	4:D:788:HOH:O	2.15	0.47
1:A:72:ILE:HG21	1:A:92:LEU:HD13	1.97	0.46
1:B:76:PRO:HD2	1:B:98:GLY:O	2.15	0.46
1:A:3:ASN:ND2	1:A:175:GLU:O	2.43	0.46
1:D:209:LYS:O	4:D:788:HOH:O	2.20	0.46
1:D:34:ARG:NE	1:D:90:GLU:OE2	2.47	0.46
1:A:30:ARG:NH2	4:A:752:HOH:O	2.41	0.45
1:D:46:ARG:CG	1:D:46:ARG:NH1	2.77	0.45
1:B:93:PRO:HB2	1:C:27:LYS:HD2	2.00	0.44
1:B:74:VAL:HG13	1:B:130:LEU:HD11	2.00	0.44
1:D:29:PHE:HD2	1:D:205:LEU:HD23	1.84	0.43
1:D:159:ALA:O	1:D:181:VAL:HA	2.18	0.43
1:B:140:SER:O	1:B:144:ILE:HG13	2.20	0.42
1:B:30:ARG:NH2	4:C:713:HOH:O	2.44	0.42
1:B:72:ILE:HD13	1:B:92:LEU:HD22	2.01	0.42
1:C:85:ALA:HA	1:C:88:ILE:HG22	2.02	0.42
1:B:41:ALA:HB2	1:B:88:ILE:HD12	2.02	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:145:GLU:O	1:A:149:GLU:HB2	2.21	0.41
1:B:131:ASP:OD1	1:B:132:PRO:HD2	2.21	0.41
1:C:57:THR:HA	1:D:20:ARG:O	2.21	0.41
1:D:76:PRO:HD2	1:D:98:GLY:O	2.21	0.41
1:B:30:ARG:HD2	4:B:726:HOH:O	2.20	0.41
1:A:17:THR:CG2	1:B:55:VAL:HG11	2.49	0.40
1:A:22:LYS:HE2	1:B:56:GLU:OE2	2.21	0.40
1:B:59:ILE:HG22	1:B:60:THR:CG2	2.51	0.40
1:C:123:ASP:N	1:C:123:ASP:OD1	2.54	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	208/221 (94%)	202 (97%)	6 (3%)	0	100	100
1	B	207/221 (94%)	201 (97%)	6 (3%)	0	100	100
1	C	208/221 (94%)	202 (97%)	6 (3%)	0	100	100
1	D	208/221 (94%)	201 (97%)	7 (3%)	0	100	100
All	All	831/884 (94%)	806 (97%)	25 (3%)	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	176/190 (93%)	166 (94%)	10 (6%)	20	28
1	B	176/190 (93%)	168 (96%)	8 (4%)	27	39
1	C	178/190 (94%)	169 (95%)	9 (5%)	24	33
1	D	179/190 (94%)	169 (94%)	10 (6%)	21	29
All	All	709/760 (93%)	672 (95%)	37 (5%)	23	32

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	54	GLU
1	A	55	VAL
1	A	71	ASP
1	A	107	GLU
1	A	108	THR
1	A	110	GLN
1	A	123	ASP
1	A	125	LYS
1	A	130	LEU
1	A	187	ARG
1	B	1	MET
1	B	57	THR
1	B	68	ASN
1	B	88	ILE
1	B	125	LYS
1	B	153	LYS
1	B	154	LYS
1	B	209	LYS
1	C	22	LYS
1	C	34	ARG
1	C	68	ASN
1	C	69	ASP
1	C	88	ILE
1	C	117	LYS
1	C	123	ASP
1	C	125	LYS
1	C	186	GLU
1	D	0	HIS
1	D	22	LYS
1	D	46	ARG
1	D	53	VAL
1	D	54	GLU
1	D	55	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	88	ILE
1	D	104	ARG
1	D	149	GLU
1	D	187	ARG

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

Mol	Chain	Res	Type
1	A	0	HIS
1	A	110	GLN
1	C	68	ASN
1	D	110	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](#)

13 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	U5P	D	603	-	18,22,22	1.37	3 (16%)	21,33,33	1.28	3 (14%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	U5P	C	602	-	18,22,22	1.31	3 (16%)	21,33,33	1.11	2 (9%)
2	SO4	A	701	-	4,4,4	0.38	0	6,6,6	0.73	0
2	SO4	A	705	-	4,4,4	0.17	0	6,6,6	0.39	0
2	SO4	B	702	-	4,4,4	0.21	0	6,6,6	0.83	0
3	U5P	B	601	-	18,22,22	1.26	2 (11%)	21,33,33	1.39	2 (9%)
2	SO4	B	706	-	4,4,4	0.22	0	6,6,6	0.17	0
2	SO4	D	704	-	4,4,4	0.23	0	6,6,6	0.29	0
3	U5P	A	600	-	18,22,22	1.04	1 (5%)	21,33,33	1.07	2 (9%)
2	SO4	D	710	-	4,4,4	0.14	0	6,6,6	0.59	0
2	SO4	D	708	-	4,4,4	0.25	0	6,6,6	0.70	0
2	SO4	A	709	-	4,4,4	0.24	0	6,6,6	0.42	0
2	SO4	C	703	-	4,4,4	0.31	0	6,6,6	0.61	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	U5P	B	601	-	-	3/8/26/26	0/2/2/2
3	U5P	D	603	-	-	3/8/26/26	0/2/2/2
3	U5P	C	602	-	-	0/8/26/26	0/2/2/2
3	U5P	A	600	-	-	0/8/26/26	0/2/2/2

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	602	U5P	O4'-C1'	3.17	1.45	1.41
3	D	603	U5P	O4'-C1'	3.07	1.45	1.41
3	B	601	U5P	O4'-C1'	2.94	1.45	1.41
3	C	602	U5P	C4-N3	2.50	1.37	1.33
3	D	603	U5P	C4-N3	2.36	1.37	1.33
3	A	600	U5P	O4'-C1'	2.15	1.44	1.41
3	D	603	U5P	P-O2P	-2.08	1.46	1.54
3	C	602	U5P	O2'-C2'	2.06	1.47	1.43
3	B	601	U5P	P-O2P	-2.03	1.47	1.54

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	601	U5P	C3'-C2'-C1'	2.88	105.31	100.98

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	603	U5P	C3'-C2'-C1'	2.82	105.22	100.98
3	A	600	U5P	O3P-P-O5'	-2.64	99.70	106.73
3	D	603	U5P	P-O5'-C5'	2.47	125.11	118.30
3	C	602	U5P	C5-C4-N3	-2.35	118.14	123.31
3	A	600	U5P	O3P-P-O1P	2.17	119.19	110.68
3	C	602	U5P	O5'-P-O1P	-2.12	100.52	106.47
3	D	603	U5P	O4'-C4'-C3'	2.05	109.16	105.11
3	B	601	U5P	O2P-P-O5'	-2.02	101.36	106.73

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	603	U5P	C5'-O5'-P-O1P
3	D	603	U5P	C5'-O5'-P-O2P
3	D	603	U5P	C5'-O5'-P-O3P
3	B	601	U5P	C5'-O5'-P-O1P
3	B	601	U5P	C5'-O5'-P-O2P
3	B	601	U5P	C5'-O5'-P-O3P

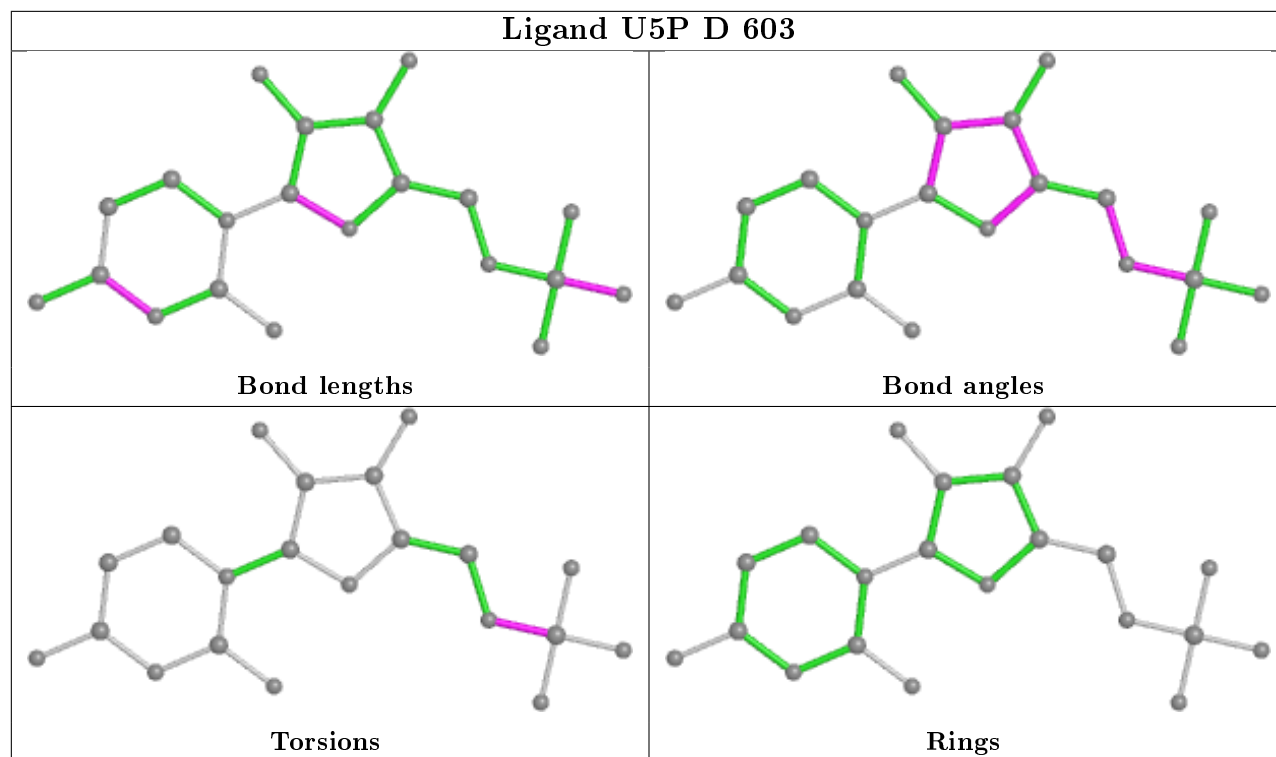
There are no ring outliers.

3 monomers are involved in 3 short contacts:

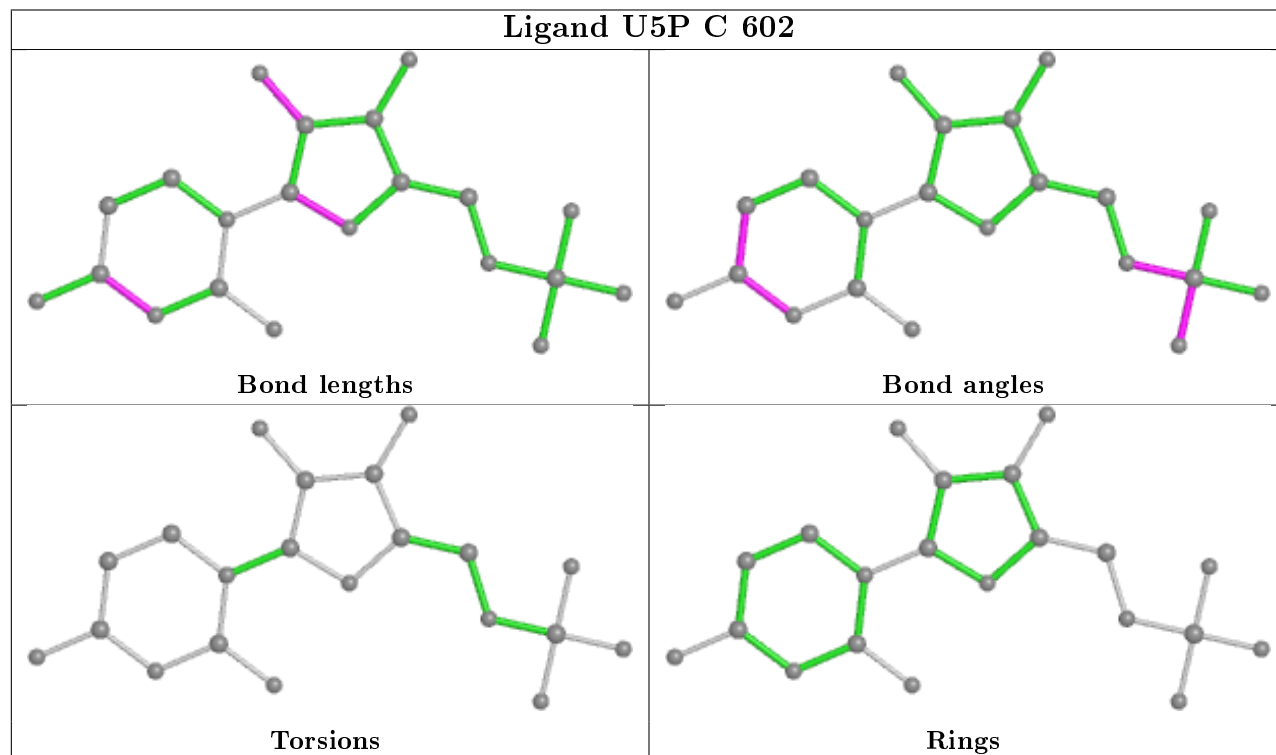
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	602	U5P	1	0
3	B	601	U5P	1	0
3	A	600	U5P	1	0

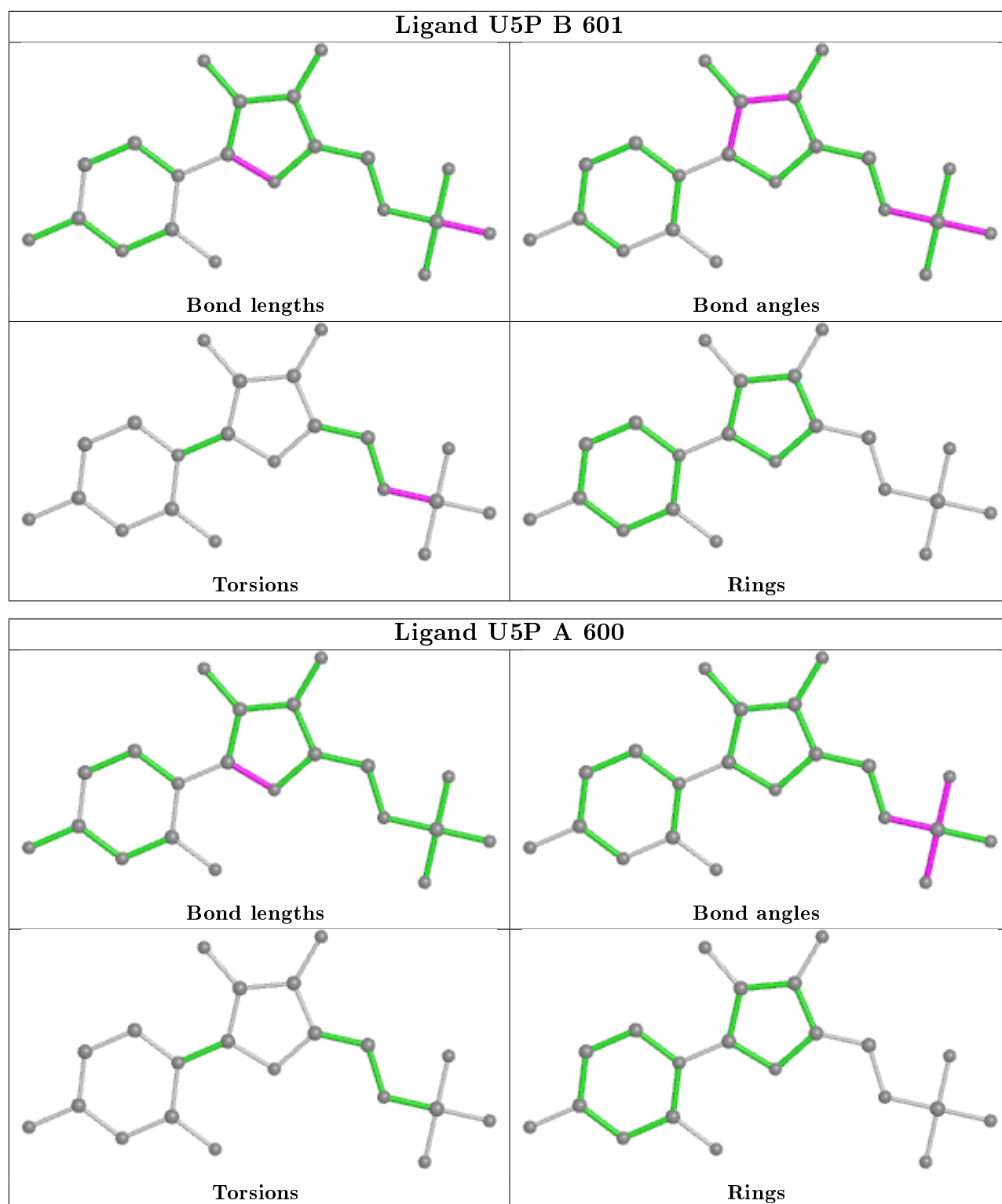
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand U5P D 603



Ligand U5P C 602





5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	210/221 (95%)	0.84	22 (10%) 6 8	42, 48, 56, 64	0
1	B	209/221 (94%)	0.93	33 (15%) 2 2	40, 48, 59, 66	0
1	C	210/221 (95%)	0.82	27 (12%) 3 5	41, 48, 58, 67	0
1	D	210/221 (95%)	0.76	22 (10%) 6 8	40, 48, 59, 72	0
All	All	839/884 (94%)	0.84	104 (12%) 4 5	40, 48, 59, 72	0

All (104) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	68	ASN	6.4
1	B	60	THR	6.2
1	A	124	ASP	4.8
1	B	50	CYS	4.6
1	A	0	HIS	4.6
1	C	68	ASN	4.6
1	C	82	LEU	4.3
1	B	82	LEU	4.2
1	D	68	ASN	4.1
1	A	82	LEU	3.9
1	D	82	LEU	3.9
1	D	107	GLU	3.8
1	C	0	HIS	3.7
1	A	122	ASN	3.7
1	A	83	VAL	3.7
1	B	123	ASP	3.6
1	C	122	ASN	3.6
1	C	123	ASP	3.5
1	D	190	ASP	3.5
1	D	209	LYS	3.5
1	C	77	ILE	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	55	VAL	3.4
1	D	78	LEU	3.3
1	A	123	ASP	3.3
1	A	121	LEU	3.2
1	A	77	ILE	3.1
1	B	78	LEU	3.1
1	A	78	LEU	3.1
1	C	78	LEU	3.1
1	A	107	GLU	3.1
1	D	122	ASN	3.0
1	D	106	PRO	3.0
1	C	176	ASP	3.0
1	D	121	LEU	2.9
1	D	102	ILE	2.9
1	C	60	THR	2.9
1	C	124	ASP	2.9
1	D	109	LEU	2.8
1	B	107	GLU	2.8
1	D	83	VAL	2.8
1	D	77	ILE	2.8
1	B	84	MET	2.7
1	A	100	ILE	2.7
1	B	193	TYR	2.7
1	C	83	VAL	2.7
1	B	57	THR	2.7
1	C	49	LYS	2.7
1	B	130	LEU	2.7
1	B	88	ILE	2.6
1	A	50	CYS	2.6
1	A	33	LEU	2.5
1	A	84	MET	2.5
1	B	53	VAL	2.5
1	C	70	LYS	2.5
1	A	68	ASN	2.5
1	B	190	ASP	2.5
1	C	107	GLU	2.5
1	B	58	PRO	2.5
1	B	100	ILE	2.5
1	B	106	PRO	2.5
1	D	76	PRO	2.5
1	B	80	ALA	2.5
1	B	59	ILE	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	109	LEU	2.5
1	D	205	LEU	2.5
1	C	209	LYS	2.4
1	D	153	LYS	2.4
1	D	206	PHE	2.4
1	A	130	LEU	2.4
1	C	130	LEU	2.4
1	B	76	PRO	2.4
1	A	190	ASP	2.3
1	B	85	ALA	2.3
1	B	124	ASP	2.3
1	B	89	LEU	2.3
1	B	129	LEU	2.3
1	C	33	LEU	2.3
1	C	81	GLY	2.2
1	A	80	ALA	2.2
1	B	54	GLU	2.2
1	C	100	ILE	2.2
1	C	187	ARG	2.2
1	D	132	PRO	2.2
1	C	190	ASP	2.2
1	A	205	LEU	2.2
1	C	84	MET	2.2
1	A	158	VAL	2.2
1	D	50	CYS	2.2
1	B	71	ASP	2.2
1	C	88	ILE	2.2
1	B	83	VAL	2.1
1	D	0	HIS	2.1
1	B	63	ILE	2.1
1	B	33	LEU	2.1
1	C	85	ALA	2.1
1	B	56	GLU	2.1
1	C	102	ILE	2.1
1	A	71	ASP	2.1
1	B	122	ASN	2.1
1	D	123	ASP	2.1
1	D	87	GLY	2.1
1	C	158	VAL	2.0
1	A	209	LYS	2.0
1	C	80	ALA	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

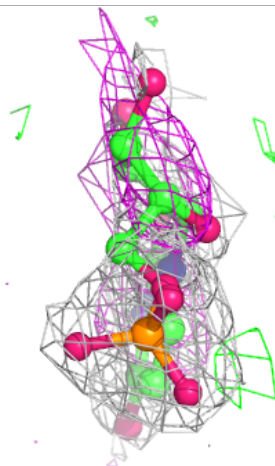
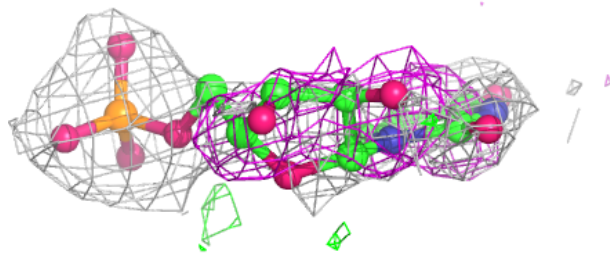
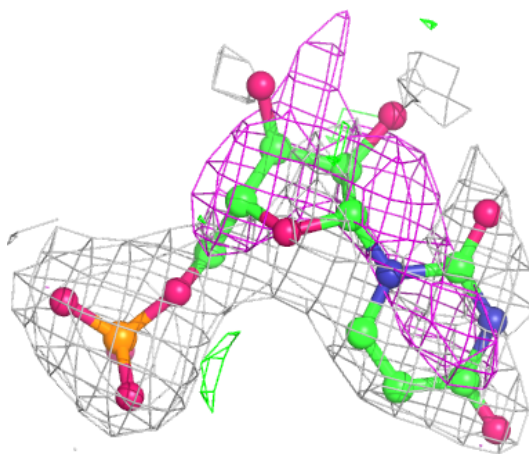
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	SO4	D	708	5/5	0.86	0.45	39,40,44,44	5
2	SO4	B	706	5/5	0.87	0.43	47,47,49,50	5
3	U5P	D	603	21/21	0.92	0.27	45,81,83,84	0
2	SO4	B	702	5/5	0.92	0.13	65,70,71,73	0
2	SO4	A	705	5/5	0.93	0.29	42,43,47,47	5
3	U5P	B	601	21/21	0.94	0.27	43,76,78,79	0
2	SO4	D	710	5/5	0.96	0.22	37,38,40,40	5
3	U5P	C	602	21/21	0.96	0.12	40,51,55,57	0
2	SO4	A	709	5/5	0.96	0.15	35,36,37,38	5
2	SO4	D	704	5/5	0.97	0.10	76,76,78,78	0
3	U5P	A	600	21/21	0.97	0.16	43,52,54,55	0
2	SO4	C	703	5/5	0.98	0.09	48,48,52,54	0
2	SO4	A	701	5/5	0.98	0.08	44,45,47,49	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.

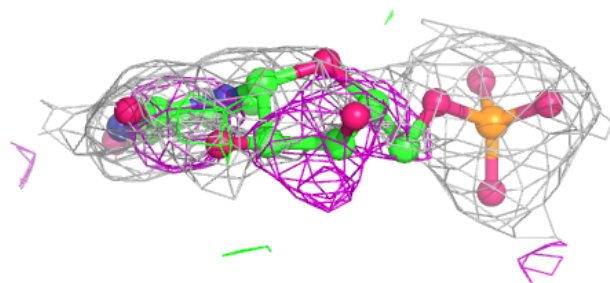
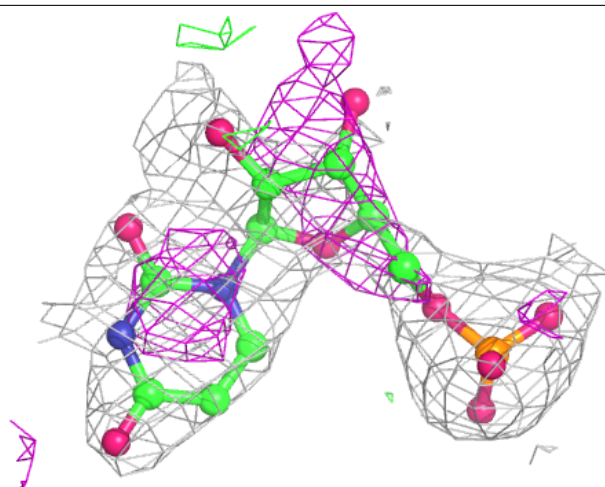
Electron density around U5P D 603:

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



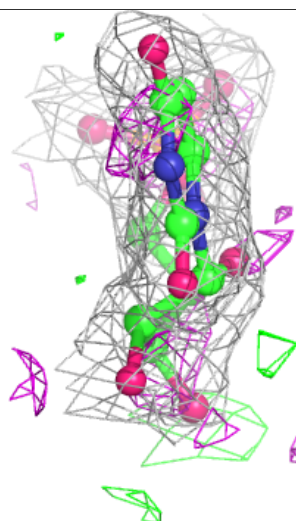
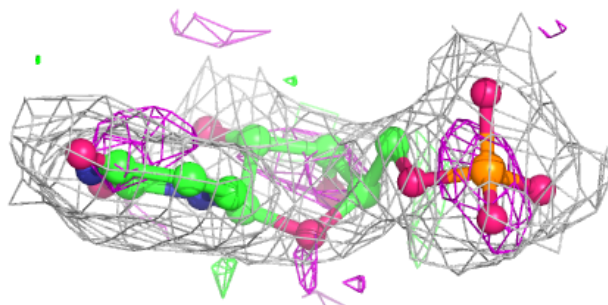
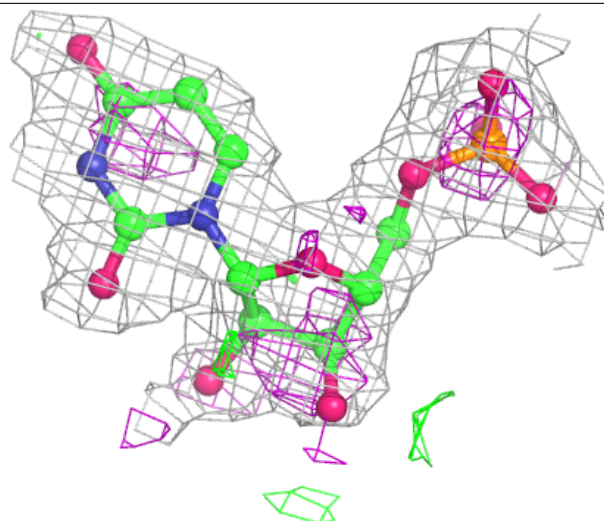
Electron density around U5P B 601:

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



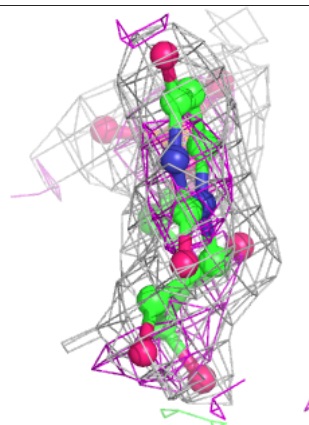
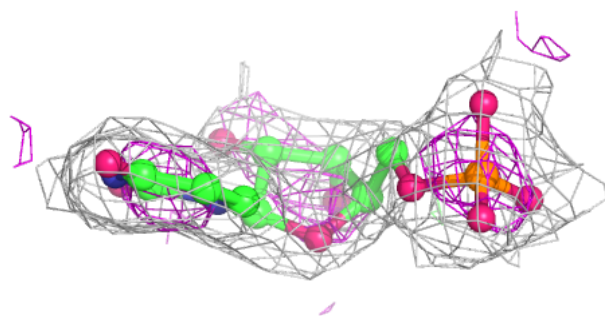
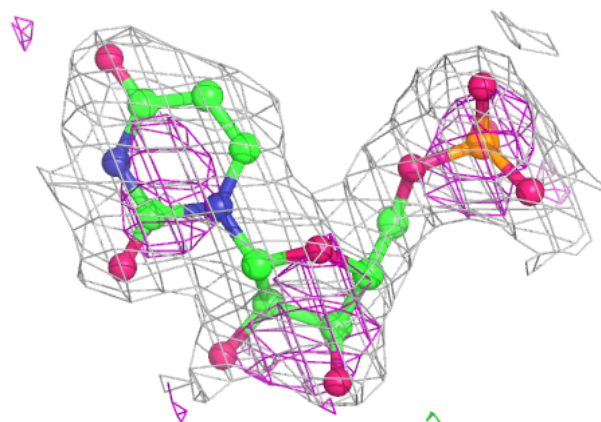
Electron density around U5P C 602:

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



Electron density around U5P A 600:

$2mF_o-DF_c$ (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.