Full wwPDB X-ray Structure Validation Report (i)

Jun 22, 2024 – 09:01 PM EDT

PDB ID	:	4R7U
Title	:	Structure of UDP-N-acetylglucosamine 1-carboxyvinyltransferase from Vibrio cholerae in complex with substrate UDP-N-acetylglucosamine and the drug fosfomycin
Authors	:	Nocek, B.; Maltseva, N.; Anderson, W.; Joachimiak, A.; Center for Structural Genomics of Infectious Diseases (CSGID)
Deposited on	:	2014-08-28
Resolution	:	2.45 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 (i)) were used in the production of this report:

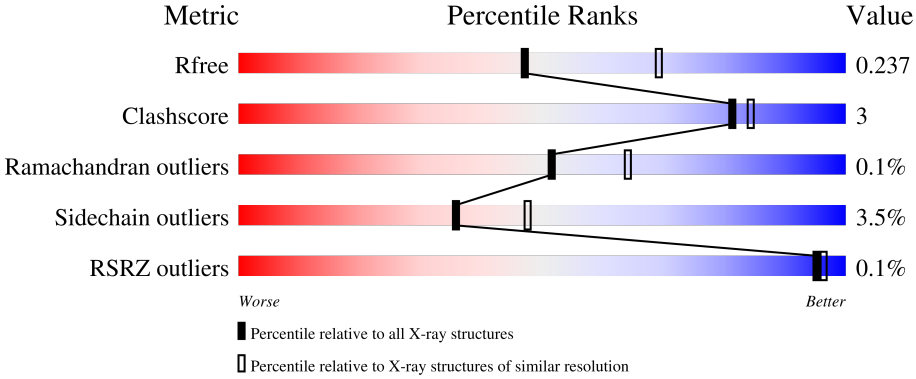
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.45 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	1544 (2.48-2.44)
Clashscore	141614	1613 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)
RSRZ outliers	127900	1523 (2.48-2.44)

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	422	92% 6% ••
1	B	422	91% 7% •
1	C	422	92% 7% •
1	D	422	91% 7% ••

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	FFQ	A	502	-	-	X	-
3	FFQ	B	502	-	-	X	-
3	FFQ	C	502	-	-	X	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 12961 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.

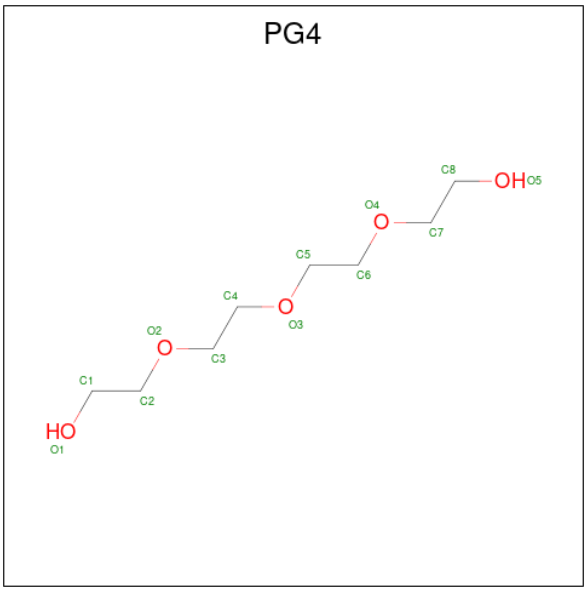
- Molecule 1 is a protein called UDP-N-acetylglucosamine 1-carboxyvinyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	418	3114	1950	554	591	19	0	0	0
1	B	418	3110	1948	551	591	20	0	0	0
1	C	419	3125	1956	555	594	20	0	0	0
1	D	418	3111	1948	551	593	19	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

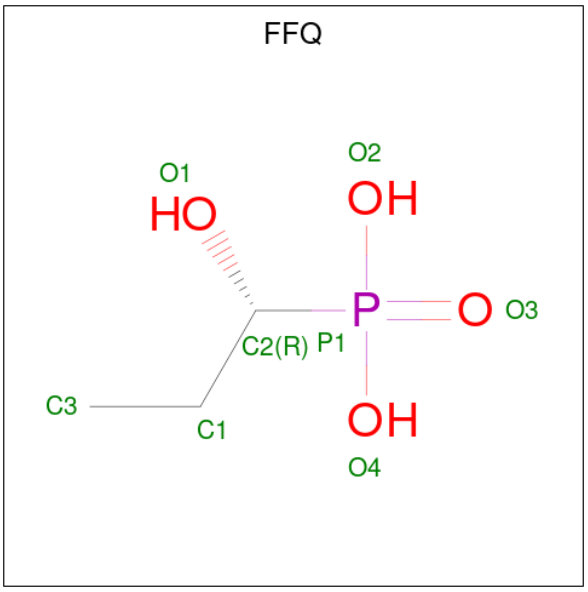
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	EXPRESSION TAG	UNP Q9KP62
A	-1	ASN	-	EXPRESSION TAG	UNP Q9KP62
A	0	ALA	-	EXPRESSION TAG	UNP Q9KP62
B	-2	SER	-	EXPRESSION TAG	UNP Q9KP62
B	-1	ASN	-	EXPRESSION TAG	UNP Q9KP62
B	0	ALA	-	EXPRESSION TAG	UNP Q9KP62
C	-2	SER	-	EXPRESSION TAG	UNP Q9KP62
C	-1	ASN	-	EXPRESSION TAG	UNP Q9KP62
C	0	ALA	-	EXPRESSION TAG	UNP Q9KP62
D	-2	SER	-	EXPRESSION TAG	UNP Q9KP62
D	-1	ASN	-	EXPRESSION TAG	UNP Q9KP62
D	0	ALA	-	EXPRESSION TAG	UNP Q9KP62

- Molecule 2 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	13	8	5	0	0

- Molecule 3 is [(1R)-1-hydroxypropyl]phosphonic acid (three-letter code: FFQ) (formula: $C_3H_9O_4P$).



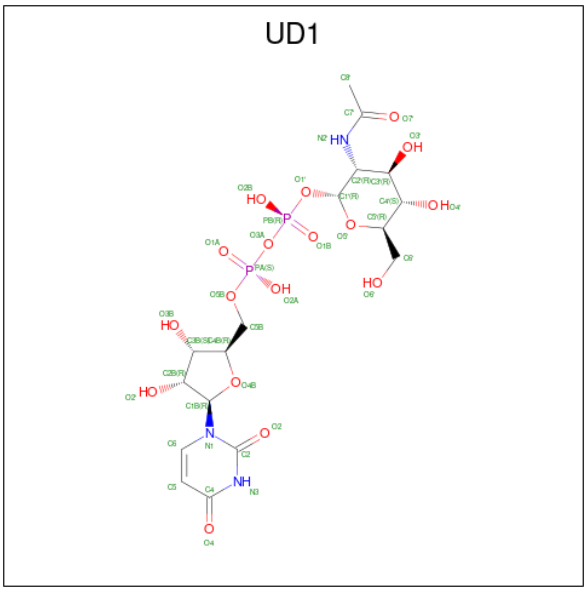
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	P		
3	A	1	8	3	4	1	0	0
3	B	1	8	3	4	1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	P		
3	C	1	8	3	4	1	0	0
			Total	C	O	P		
3	D	1	8	3	4	1	0	0

- Molecule 4 is URIDINE-DIPHOSPHATE-N-ACETYLGLUCOSAMINE (three-letter code: UD1) (formula: $C_{17}H_{27}N_3O_{17}P_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	A	1	39	17	3	17	2	0	0
			Total	C	N	O	P		
4	B	1	39	17	3	17	2	0	0
			Total	C	N	O	P		
4	C	1	39	17	3	17	2	0	0
			Total	C	N	O	P		
4	D	1	39	17	3	17	2	0	0

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
5	A	1	1	1	0	0
			Total	Na		
5	B	1	1	1	0	0
			Total	Na		
5	C	2	2	2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
5	D	1	1	1	0	0

- Molecule 6 is water.

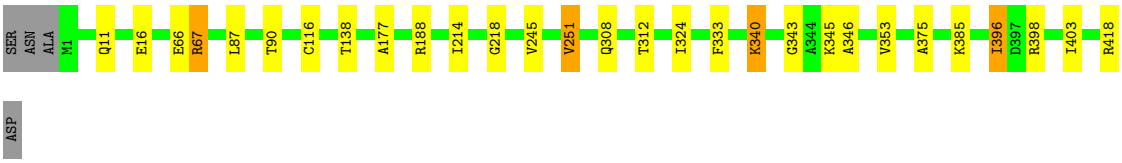
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	77	77	77	0	0
			Total	O		
6	B	61	61	61	0	0
			Total	O		
6	C	67	67	67	0	0
			Total	O		
6	D	90	90	90	0	0

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.

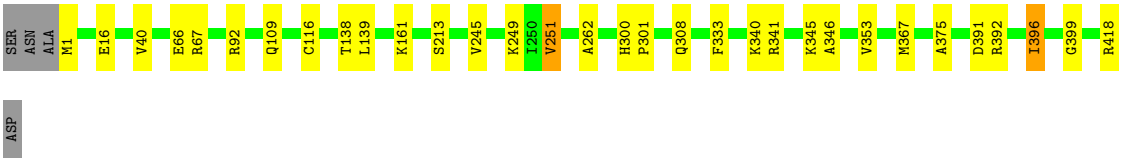
• Molecule 1: UDP-N-acetylglucosamine 1-carboxyvinyltransferase

Chain A:



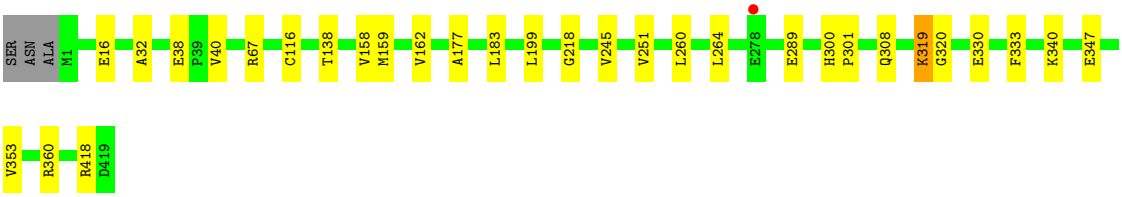
• Molecule 1: UDP-N-acetylglucosamine 1-carboxyvinyltransferase

Chain B:



• Molecule 1: UDP-N-acetylglucosamine 1-carboxyvinyltransferase

Chain C:



• Molecule 1: UDP-N-acetylglucosamine 1-carboxyvinyltransferase

Chain D:

I396
R418
ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	204.28Å 204.28Å 118.88Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	38.64 – 2.45 38.61 – 2.45	Depositor EDS
% Data completeness (in resolution range)	86.7 (38.64-2.45) 86.7 (38.61-2.45)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.20 (at 2.45Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.178 , 0.234 0.181 , 0.237	Depositor DCC
R_{free} test set	2986 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	28.8	Xtriage
Anisotropy	0.090	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 24.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.035 for -2/3*h-1/3*k-4/3*l,-1/3*h-2/3*k+4/3*l,-1/3*h+1/3*k+1/3*l 0.024 for -h,1/3*h-1/3*k-4/3*l,-1/3*h-2/3*k+1/3*l 0.024 for -1/3*h+1/3*k+4/3*l,-k,2/3*h+1/3*k+1/3*l 0.035 for -h,2/3*h+1/3*k+4/3*l,1/3*h+2/3*k-1/3*l 0.055 for -1/3*h-2/3*k+4/3*l,-2/3*h-1/3*k-4/3*l,1/3*h-1/3*k-1/3*l 0.034 for 1/3*h+2/3*k-4/3*l,-k,-2/3*h-1/3*k-1/3*l 0.036 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	12961	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.31% 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: PG4, NA, UD1, FFQ

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	0.67	0/3159	0.74	1/4278 (0.0%)
1	B	0.66	0/3155	0.74	1/4273 (0.0%)
1	C	0.67	0/3170	0.77	1/4292 (0.0%)
1	D	0.67	0/3156	0.74	2/4275 (0.0%)
All	All	0.67	0/12640	0.75	5/17118 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
1	C	0	1
1	D	0	1
All	All	0	4

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	67	ARG	NE-CZ-NH1	5.83	123.22	120.30
1	D	296	ARG	NE-CZ-NH2	-5.50	117.55	120.30
1	C	360	ARG	NE-CZ-NH1	-5.31	117.65	120.30
1	A	398	ARG	NE-CZ-NH2	-5.18	117.71	120.30
1	B	92	ARG	NE-CZ-NH2	-5.08	117.76	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	67	ARG	Peptide
1	B	67	ARG	Peptide
1	C	67	ARG	Peptide
1	D	67	ARG	Peptide

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	3114	0	3163	16	0
1	B	3110	0	3157	16	0
1	C	3125	0	3174	15	0
1	D	3111	0	3154	15	0
2	A	13	0	18	0	0
3	A	8	0	7	5	0
3	B	8	0	7	4	0
3	C	8	0	7	5	0
3	D	8	0	7	3	0
4	A	39	0	25	0	0
4	B	39	0	25	0	0
4	C	39	0	25	0	0
4	D	39	0	25	2	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	C	2	0	0	0	0
5	D	1	0	0	0	0
6	A	77	0	0	2	0
6	B	61	0	0	0	0
6	C	67	0	0	1	0
6	D	90	0	0	1	0
All	All	12961	0	12794	61	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 (61) 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:116:CYS:SG	3:B:502:FFQ:H3B	1.69	1.31
1:D:116:CYS:SG	3:D:502:FFQ:H3B	1.91	1.11
1:C:116:CYS:SG	3:C:502:FFQ:H3B	2.05	0.96
1:B:116:CYS:HG	3:B:502:FFQ:H3B	1.10	0.91
1:A:116:CYS:SG	3:A:502:FFQ:HO2	2.26	0.75
1:A:116:CYS:SG	3:A:502:FFQ:H3B	2.28	0.72
1:B:341:ARG:HH11	1:B:341:ARG:HG2	1.57	0.68
1:D:116:CYS:SG	3:D:502:FFQ:C2	2.84	0.65
1:A:116:CYS:SG	3:A:502:FFQ:C3	2.86	0.63
1:B:333:PHE:CE1	1:B:353:VAL:HG23	2.39	0.58
1:A:116:CYS:SG	3:A:502:FFQ:H1	2.46	0.56
1:A:340:LYS:NZ	1:A:346:ALA:O	2.38	0.55
1:B:116:CYS:SG	3:B:502:FFQ:C3	2.93	0.55
1:A:214:ILE:HD11	6:A:661:HOH:O	2.07	0.54
1:B:341:ARG:HG2	1:B:341:ARG:NH1	2.24	0.53
1:B:375:ALA:HA	1:B:396:ILE:HD11	1.91	0.52
1:C:116:CYS:SG	3:C:502:FFQ:C2	2.97	0.52
1:B:333:PHE:CE1	1:B:353:VAL:CG2	2.93	0.52
1:A:116:CYS:SG	3:A:502:FFQ:C2	2.97	0.52
1:B:116:CYS:SG	3:B:502:FFQ:C2	2.98	0.51
1:D:273:GLU:OE2	1:D:288:ARG:NH2	2.43	0.51
1:D:138:THR:HG22	1:D:139:LEU:N	2.27	0.50
3:D:502:FFQ:O4	6:D:619:HOH:O	2.20	0.49
1:C:116:CYS:SG	3:C:502:FFQ:C3	3.01	0.49
1:C:116:CYS:SG	3:C:502:FFQ:H1	2.53	0.48
1:A:312:THR:HA	1:A:324:ILE:CD1	2.43	0.48
1:D:96:TRP:HH2	4:D:501:UD1:H8'1	1.78	0.48
1:C:116:CYS:SG	3:C:502:FFQ:HO2	2.55	0.47
1:C:177:ALA:O	1:C:218:GLY:HA3	2.13	0.47
1:C:16:GLU:HG2	1:C:251:VAL:HG13	1.97	0.47
1:A:343:GLY:HA2	1:B:139:LEU:HD23	1.96	0.46
1:D:312:THR:O	1:D:316:MET:HG3	2.16	0.46
1:B:333:PHE:CZ	1:B:353:VAL:HG23	2.51	0.46
1:D:56:GLU:O	1:D:60:ARG:HG2	2.15	0.46
1:C:300:HIS:CG	1:C:301:PRO:HA	2.52	0.45
1:B:391:ASP:OD1	1:B:392:ARG:N	2.46	0.45
1:A:396:ILE:HD13	1:A:403:ILE:HG21	1.99	0.45
1:B:340:LYS:HE2	1:B:346:ALA:O	2.17	0.44
1:D:127:ILE:HG23	1:D:137:ILE:HG21	1.98	0.44
1:A:177:ALA:O	1:A:218:GLY:HA3	2.17	0.44
1:C:159:MET:SD	1:C:183:LEU:HD11	2.58	0.44
1:B:262:ALA:HA	1:C:158:VAL:HG11	1.99	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:260:LEU:O	1:C:264:LEU:HG	2.17	0.44
1:D:374:SER:HB3	1:D:396:ILE:HG13	1.99	0.43
1:C:32:ALA:HB1	1:C:199:LEU:HD21	2.00	0.43
1:B:16:GLU:HG2	1:B:251:VAL:HG13	2.00	0.43
1:A:67:ARG:NH1	6:A:601:HOH:O	2.52	0.42
1:A:333:PHE:CE1	1:A:353:VAL:HG23	2.54	0.42
1:A:375:ALA:HA	1:A:396:ILE:HD11	2.02	0.42
1:A:16:GLU:HG2	1:A:251:VAL:HG13	2.01	0.42
1:B:300:HIS:CG	1:B:301:PRO:HA	2.54	0.42
1:D:48:LEU:HD12	1:D:51:ILE:HD12	2.02	0.41
1:C:162:VAL:HA	6:C:648:HOH:O	2.19	0.41
1:C:319:LYS:HG3	1:C:320:GLY:N	2.35	0.41
1:C:333:PHE:CE1	1:C:353:VAL:CG2	3.04	0.41
1:D:86:ASP:HA	1:D:89:LYS:HD3	2.02	0.41
1:A:87:LEU:O	1:A:90:THR:HB	2.21	0.41
1:D:312:THR:HA	1:D:324:ILE:CD1	2.51	0.41
1:D:24:ASN:HB3	4:D:501:UD1:H8'2	2.03	0.40
1:D:76:GLY:N	1:D:77:PRO:CD	2.85	0.40
1:D:250:ILE:HG12	1:D:251:VAL:N	2.36	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	416/422 (99%)	402 (97%)	14 (3%)	0	100	100
1	B	416/422 (99%)	404 (97%)	11 (3%)	1 (0%)	47	57
1	C	417/422 (99%)	407 (98%)	10 (2%)	0	100	100
1	D	416/422 (99%)	401 (96%)	15 (4%)	0	100	100
All	All	1665/1688 (99%)	1614 (97%)	50 (3%)	1 (0%)	51	64

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	399	GLY

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	324/329 (98%)	312 (96%)	12 (4%)	34	45
1	B	324/329 (98%)	309 (95%)	15 (5%)	27	35
1	C	326/329 (99%)	315 (97%)	11 (3%)	37	48
1	D	324/329 (98%)	316 (98%)	8 (2%)	47	60
All	All	1298/1316 (99%)	1252 (96%)	46 (4%)	36	47

All (46) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	11	GLN
1	A	66	GLU
1	A	138	THR
1	A	188	ARG
1	A	245	VAL
1	A	251	VAL
1	A	308	GLN
1	A	340	LYS
1	A	345	LYS
1	A	385	LYS
1	A	396	ILE
1	A	418	ARG
1	B	1	MET
1	B	40	VAL
1	B	66	GLU
1	B	109	GLN
1	B	138	THR
1	B	161	LYS
1	B	213	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	245	VAL
1	B	249	LYS
1	B	251	VAL
1	B	308	GLN
1	B	345	LYS
1	B	367	MET
1	B	396	ILE
1	B	418	ARG
1	C	38	GLU
1	C	40	VAL
1	C	138	THR
1	C	245	VAL
1	C	289	GLU
1	C	308	GLN
1	C	319	LYS
1	C	330	GLU
1	C	340	LYS
1	C	347	GLU
1	C	418	ARG
1	D	92	ARG
1	D	95	ILE
1	D	161	LYS
1	D	251	VAL
1	D	260	LEU
1	D	308	GLN
1	D	396	ILE
1	D	418	ARG

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

Mol	Chain	Res	Type
1	D	147	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)

Of 14 ligands modelled in this entry, 5 are monoatomic - leaving 9 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	FFQ	A	502	1	5,7,7	3.14	3 (60%)	7,10,10	1.35	1 (14%)
3	FFQ	C	502	1	5,7,7	3.37	2 (40%)	7,10,10	1.10	0
4	UD1	C	501	-	38,41,41	1.03	3 (7%)	57,62,62	1.55	10 (17%)
3	FFQ	B	502	1	5,7,7	2.89	3 (60%)	7,10,10	1.61	1 (14%)
4	UD1	A	503	-	38,41,41	1.14	5 (13%)	57,62,62	1.53	11 (19%)
4	UD1	B	501	-	38,41,41	1.02	4 (10%)	57,62,62	1.28	8 (14%)
4	UD1	D	501	-	38,41,41	0.98	2 (5%)	57,62,62	1.55	8 (14%)
3	FFQ	D	502	1	5,7,7	2.49	2 (40%)	7,10,10	1.43	2 (28%)
2	PG4	A	501	-	12,12,12	0.54	0	11,11,11	0.40	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	FFQ	A	502	1	-	0/7/8/8	-
3	FFQ	C	502	1	-	0/7/8/8	-
4	UD1	C	501	-	-	5/26/63/63	0/3/3/3
3	FFQ	B	502	1	-	1/7/8/8	-
4	UD1	A	503	-	-	7/26/63/63	0/3/3/3
4	UD1	B	501	-	-	5/26/63/63	0/3/3/3
4	UD1	D	501	-	-	6/26/63/63	0/3/3/3
3	FFQ	D	502	1	-	0/7/8/8	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PG4	A	501	-	-	4/10/10/10	-

All (24) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	502	FFQ	P1-O4	5.71	1.63	1.54
3	B	502	FFQ	P1-O2	4.59	1.62	1.54
3	A	502	FFQ	P1-O2	4.48	1.62	1.54
3	C	502	FFQ	P1-O2	4.44	1.61	1.54
3	A	502	FFQ	P1-O4	4.20	1.61	1.54
3	D	502	FFQ	P1-O2	3.89	1.61	1.54
3	D	502	FFQ	P1-O4	3.59	1.60	1.54
3	B	502	FFQ	O1-C2	3.41	1.45	1.41
3	A	502	FFQ	O1-C2	3.36	1.45	1.41
4	D	501	UD1	C2-N3	-3.05	1.32	1.38
3	B	502	FFQ	P1-O4	3.00	1.59	1.54
4	A	503	UD1	C2-N3	-2.93	1.32	1.38
4	A	503	UD1	C4-N3	-2.93	1.33	1.38
4	D	501	UD1	C4-N3	-2.65	1.33	1.38
4	C	501	UD1	C5-C4	-2.39	1.38	1.43
4	C	501	UD1	C4-N3	-2.39	1.34	1.38
4	B	501	UD1	C6-C5	2.37	1.40	1.35
4	B	501	UD1	C2-N1	2.30	1.42	1.38
4	C	501	UD1	C6-N1	-2.19	1.32	1.38
4	A	503	UD1	C1'-C2'	-2.18	1.49	1.53
4	B	501	UD1	O2-C2	2.06	1.26	1.23
4	B	501	UD1	O5'-C5'	-2.06	1.39	1.44
4	A	503	UD1	C5-C4	-2.05	1.39	1.43
4	A	503	UD1	C2-N1	2.05	1.41	1.38

All (41) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	501	UD1	N3-C2-N1	4.67	121.09	114.89
4	C	501	UD1	O4-C4-C5	-4.14	117.87	125.16
4	C	501	UD1	C4-N3-C2	-4.14	121.12	126.58
4	C	501	UD1	O2-C2-N1	-4.04	117.42	122.79
4	D	501	UD1	C4-N3-C2	-4.01	121.28	126.58
4	A	503	UD1	C5-C4-N3	3.91	120.68	114.84
4	C	501	UD1	N3-C2-N1	3.73	119.84	114.89
4	D	501	UD1	C5-C4-N3	3.67	120.34	114.84
4	A	503	UD1	C1'-C2'-N2'	-3.45	105.07	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	503	UD1	C4-N3-C2	-3.37	122.14	126.58
4	B	501	UD1	N3-C2-N1	3.12	119.04	114.89
4	A	503	UD1	PB-O3A-PA	-3.11	122.14	132.83
4	C	501	UD1	C5-C4-N3	3.07	119.43	114.84
4	A	503	UD1	N3-C2-N1	3.04	118.92	114.89
3	B	502	FFQ	O4-P1-O3	-2.96	106.01	113.45
4	B	501	UD1	C4-N3-C2	-2.88	122.78	126.58
4	C	501	UD1	C4'-C3'-C2'	-2.80	106.24	110.34
3	D	502	FFQ	O2-P1-O3	-2.78	106.47	113.45
4	B	501	UD1	O4'-C4'-C3'	-2.75	103.99	110.35
4	B	501	UD1	PB-O3A-PA	-2.70	123.56	132.83
4	B	501	UD1	C5-C4-N3	2.64	118.79	114.84
4	A	503	UD1	O2B-PB-O1B	2.61	125.13	112.24
3	A	502	FFQ	O3-P1-C2	-2.51	106.84	112.94
4	B	501	UD1	O2-C2-N1	-2.49	119.48	122.79
4	D	501	UD1	O3'-C3'-C2'	2.46	114.63	109.66
4	A	503	UD1	O6'-C6'-C5'	-2.44	102.91	111.29
4	C	501	UD1	O4'-C4'-C3'	-2.44	104.71	110.35
4	D	501	UD1	C1'-C2'-N2'	-2.41	106.85	111.00
4	D	501	UD1	O4B-C1B-N1	2.40	113.84	108.36
4	B	501	UD1	O4-C4-C5	-2.39	120.95	125.16
4	C	501	UD1	C1'-C2'-N2'	-2.35	106.95	111.00
4	C	501	UD1	O4-C4-N3	2.31	122.69	119.31
4	D	501	UD1	O4-C4-C5	-2.27	121.17	125.16
4	D	501	UD1	PB-O3A-PA	-2.23	125.19	132.83
4	A	503	UD1	C4'-C3'-C2'	-2.17	107.17	110.34
3	D	502	FFQ	O4-P1-O3	-2.11	108.14	113.45
4	A	503	UD1	O4-C4-C5	-2.09	121.48	125.16
4	A	503	UD1	O4B-C1B-N1	2.07	113.10	108.36
4	B	501	UD1	O3A-PB-O1'	-2.03	98.39	102.48
4	A	503	UD1	O5'-C1'-C2'	-2.01	106.66	110.58
4	C	501	UD1	PB-O3A-PA	-2.01	125.94	132.83

There are no chirality outliers.

All (28) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	503	UD1	C5B-O5B-PA-O2A
4	B	501	UD1	C5B-O5B-PA-O1A
4	B	501	UD1	C5B-O5B-PA-O2A
4	C	501	UD1	C5B-O5B-PA-O1A
4	C	501	UD1	C5B-O5B-PA-O2A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	D	501	UD1	C5B-O5B-PA-O1A
4	D	501	UD1	C5B-O5B-PA-O2A
4	A	503	UD1	O7'-C7'-N2'-C2'
4	A	503	UD1	C8'-C7'-N2'-C2'
4	A	503	UD1	C4'-C5'-C6'-O6'
4	B	501	UD1	C4'-C5'-C6'-O6'
4	D	501	UD1	C8'-C7'-N2'-C2'
4	D	501	UD1	O7'-C7'-N2'-C2'
4	A	503	UD1	O5'-C5'-C6'-O6'
4	B	501	UD1	O5'-C5'-C6'-O6'
2	A	501	PG4	O2-C3-C4-O3
2	A	501	PG4	O4-C7-C8-O5
2	A	501	PG4	O1-C1-C2-O2
4	A	503	UD1	C5B-O5B-PA-O3A
4	B	501	UD1	C5B-O5B-PA-O3A
4	D	501	UD1	C5B-O5B-PA-O3A
4	D	501	UD1	PB-O3A-PA-O1A
2	A	501	PG4	C6-C5-O3-C4
3	B	502	FFQ	C1-C2-P1-O3
4	C	501	UD1	C5B-O5B-PA-O3A
4	C	501	UD1	PB-O3A-PA-O1A
4	C	501	UD1	PB-O3A-PA-O2A
4	A	503	UD1	C5B-O5B-PA-O1A

There are no ring outliers.

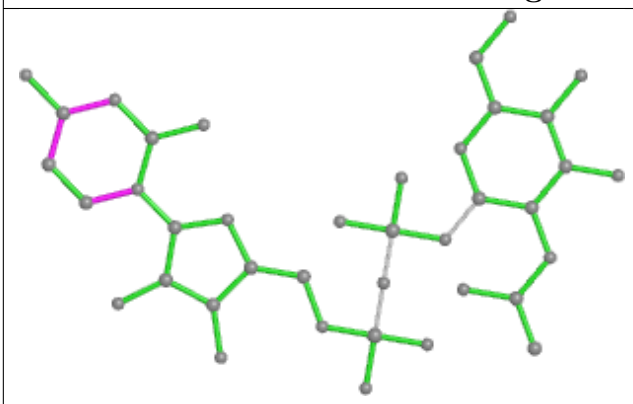
5 monomers are involved in 19 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	502	FFQ	5	0
3	C	502	FFQ	5	0
3	B	502	FFQ	4	0
4	D	501	UD1	2	0
3	D	502	FFQ	3	0

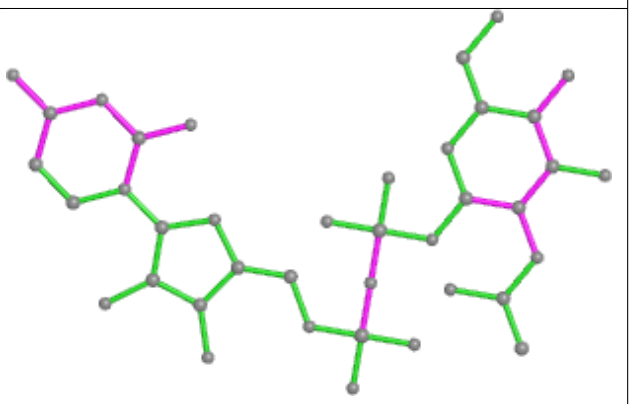
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

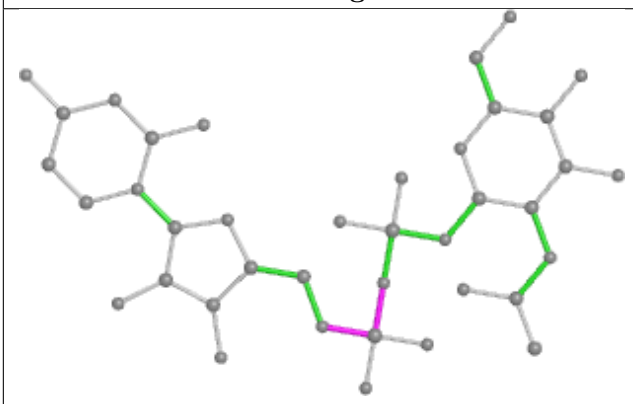
Ligand UD1 C 501



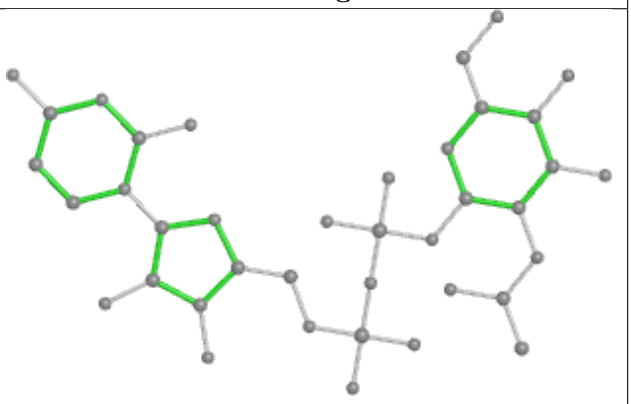
Bond lengths



Bond angles

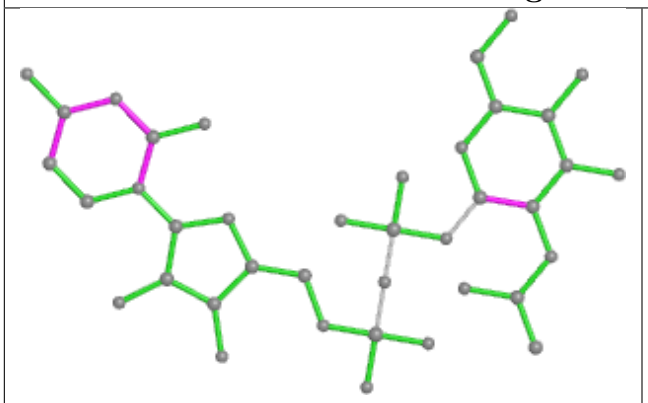


Torsions

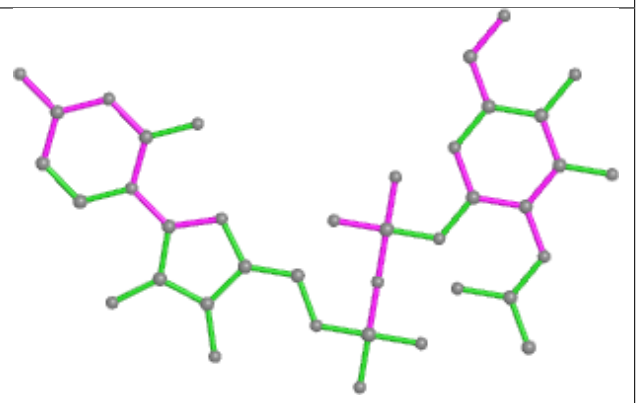


Rings

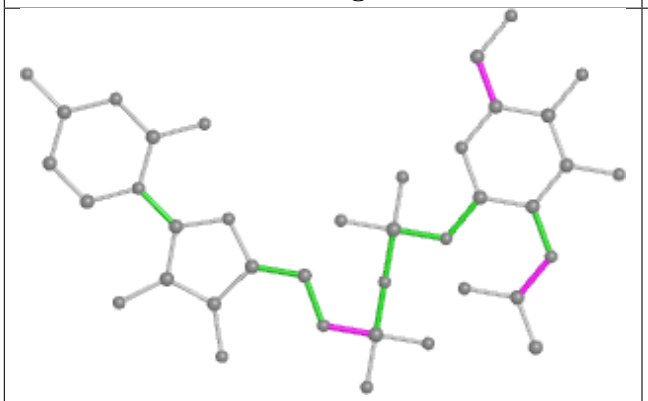
Ligand UD1 A 503



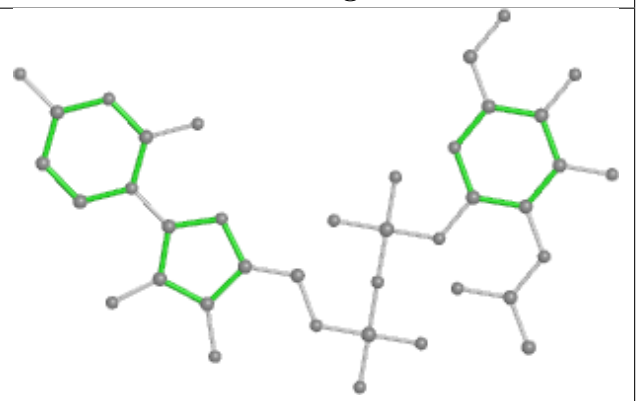
Bond lengths



Bond angles

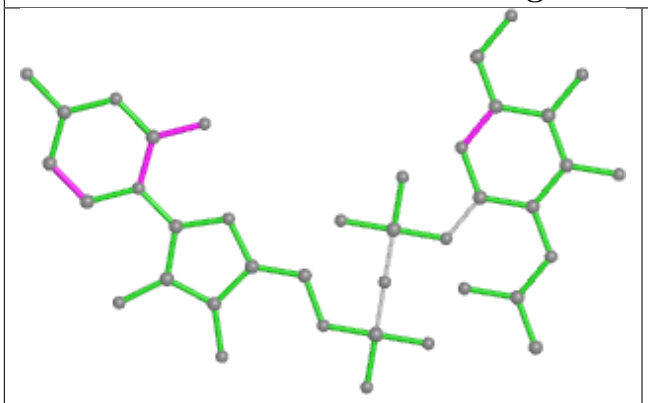


Torsions

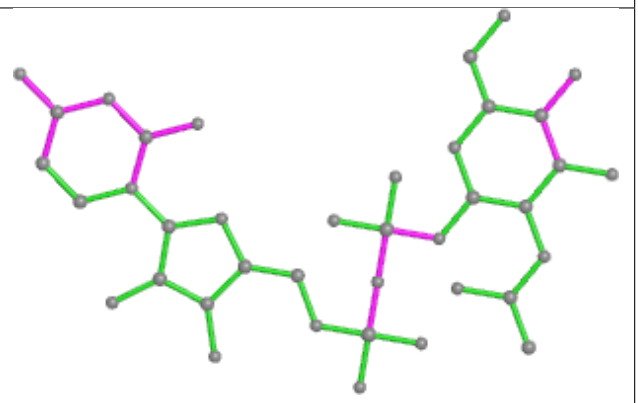


Rings

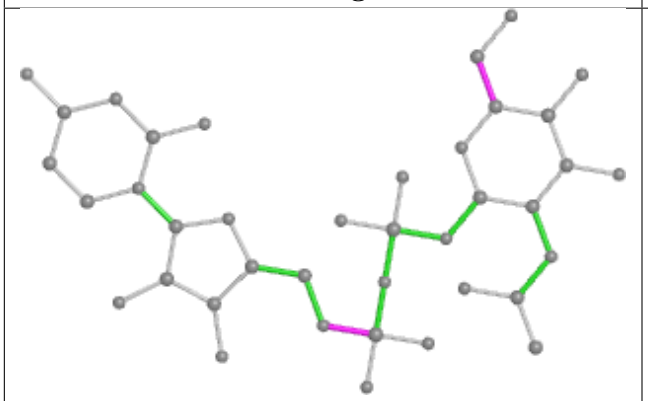
Ligand UD1 B 501



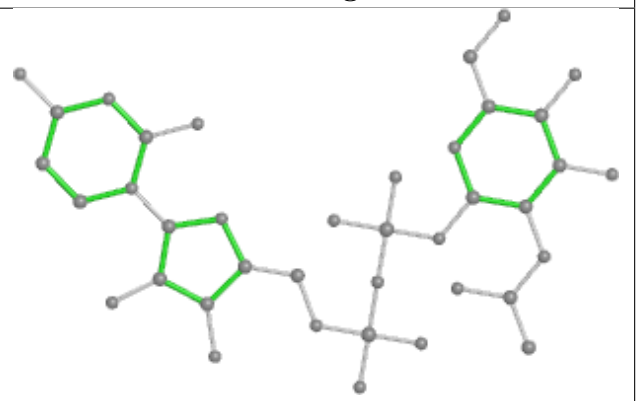
Bond lengths



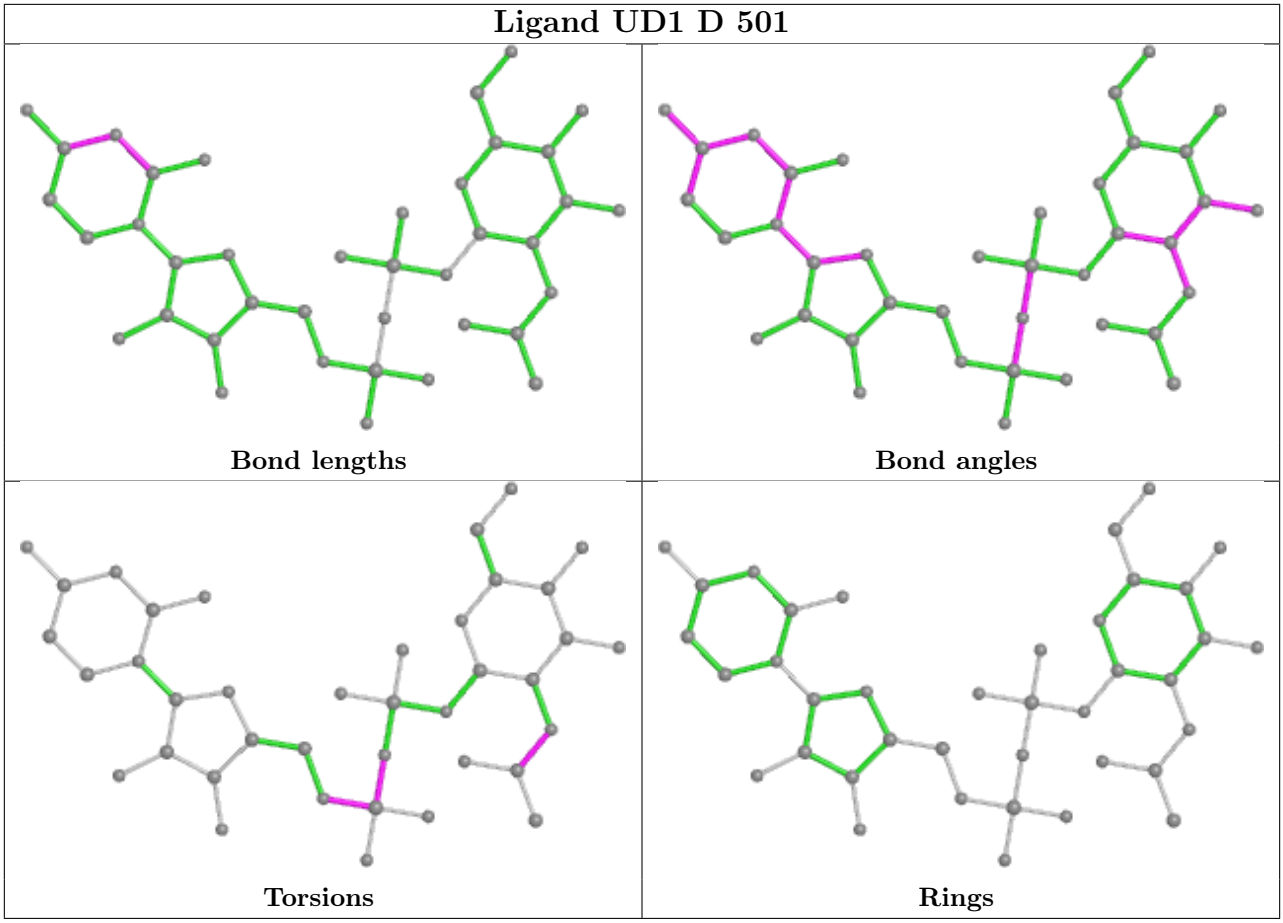
Bond angles



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	418/422 (99%)	-0.72	0	100	100	18, 29, 45, 62	0
1	B	418/422 (99%)	-0.67	0	100	100	20, 31, 48, 64	0
1	C	419/422 (99%)	-0.66	1 (0%)	95	95	20, 31, 47, 68	0
1	D	418/422 (99%)	-0.74	0	100	100	19, 29, 42, 68	0
All	All	1673/1688 (99%)	-0.70	1 (0%)	95	96	18, 30, 46, 68	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	278	GLU	2.4

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	PG4	A	501	13/13	0.88	0.21	56,58,64,66	0

Continued on next page...

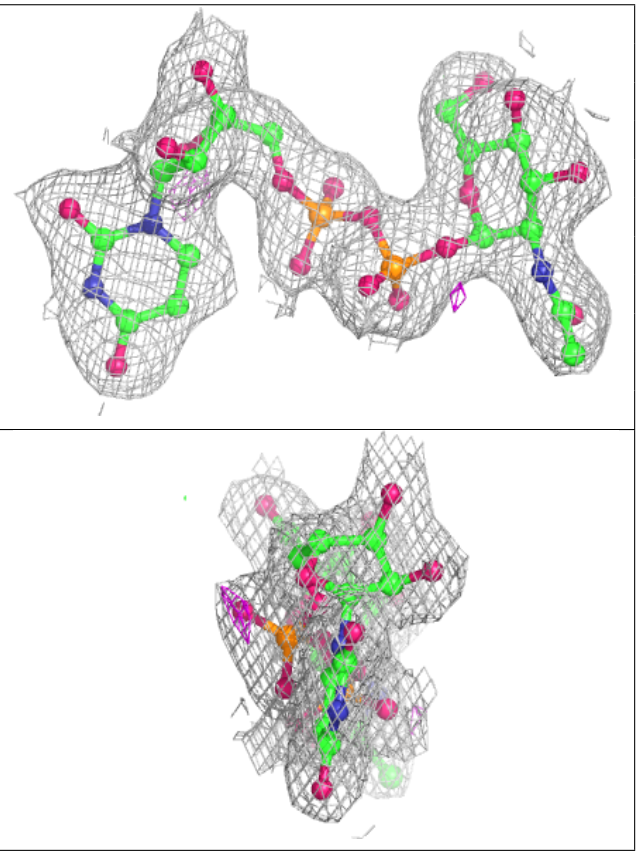
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	FFQ	C	502	8/8	0.90	0.16	49,53,54,54	0
3	FFQ	D	502	8/8	0.91	0.16	46,47,51,52	0
5	NA	C	504	1/1	0.93	0.10	54,54,54,54	0
3	FFQ	B	502	8/8	0.95	0.14	38,47,49,50	0
3	FFQ	A	502	8/8	0.95	0.13	47,49,53,53	0
4	UD1	B	501	39/39	0.98	0.10	19,24,27,28	0
4	UD1	C	501	39/39	0.98	0.11	17,23,28,30	0
4	UD1	D	501	39/39	0.98	0.11	18,23,26,30	0
5	NA	B	503	1/1	0.98	0.04	29,29,29,29	0
5	NA	C	503	1/1	0.98	0.07	18,18,18,18	0
4	UD1	A	503	39/39	0.98	0.11	16,20,25,30	0
5	NA	D	503	1/1	0.98	0.25	24,24,24,24	0
5	NA	A	504	1/1	0.99	0.04	28,28,28,28	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.

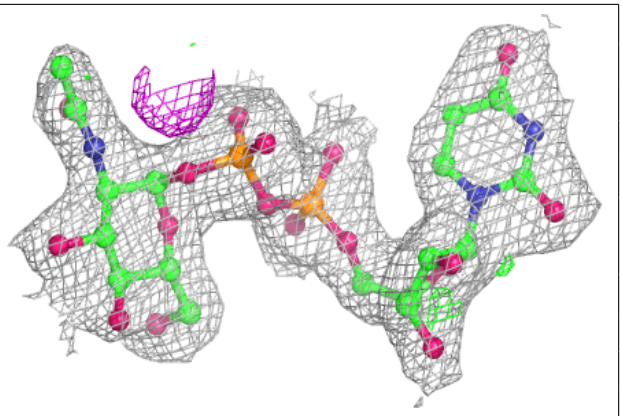
Electron density around UD1 B 501:

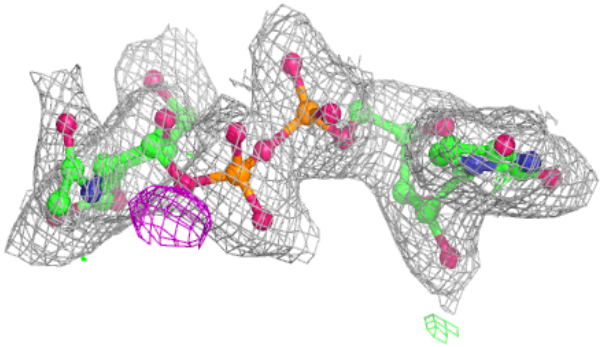
2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)

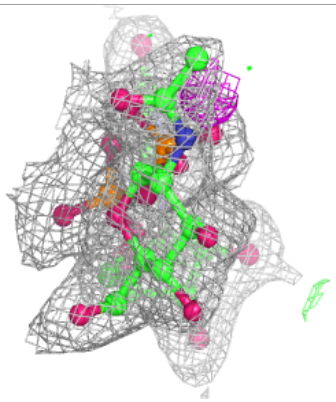


Electron density around UD1 C 501:

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

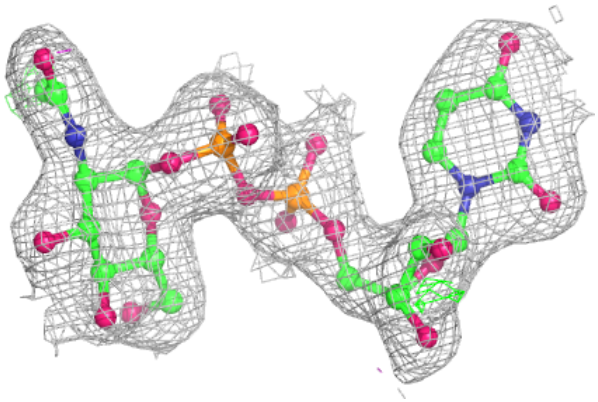


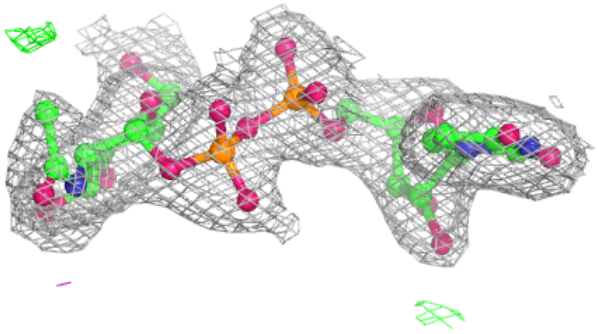


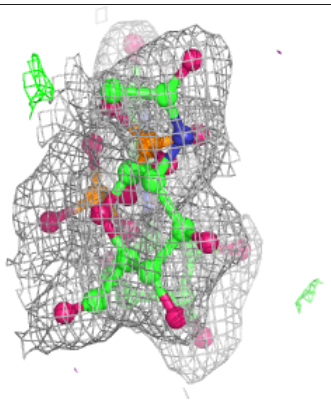


Electron density around UD1 D 501:

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

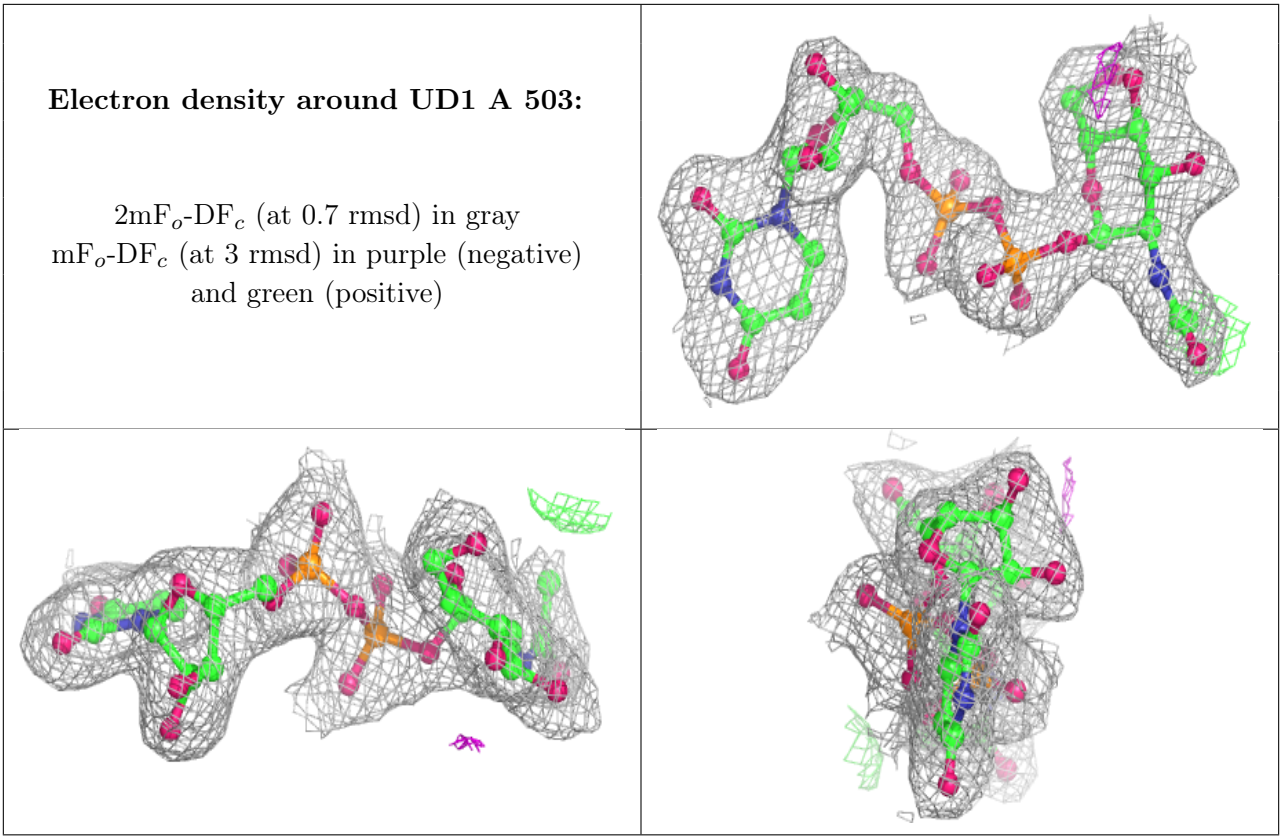






Electron density around UD1 A 503:

$2mF_o$-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.