Full wwPDB X-ray Structure Validation Report (i)

Jan 2, 2024 – 11:02 pm GMT

PDB ID	:	5FOE
Title	:	Crystal structure of the C. elegans Protein O-fucosyltransferase 2 (CePO-FUT2) double mutant (R298K-R299K) in complex with GDP and the human TSR1 from thrombospondin 1
Authors	:	Valero-Gonzalez, J.; Leonhard-Melief, C.; Lira-Navarrete, E.; Jimenez-Oses, G.; Hernandez-Ruiz, C.; Pallares, M.C.; Yruela, I.; Vasudevan, D.; Lostao, A.; Corzana, F.; Takeuchi, H.; Haltiwanger, R.S.; Hurtado-Guerrero, R.
Deposited on	:	2015-11-19
Resolution	:	1.98 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:

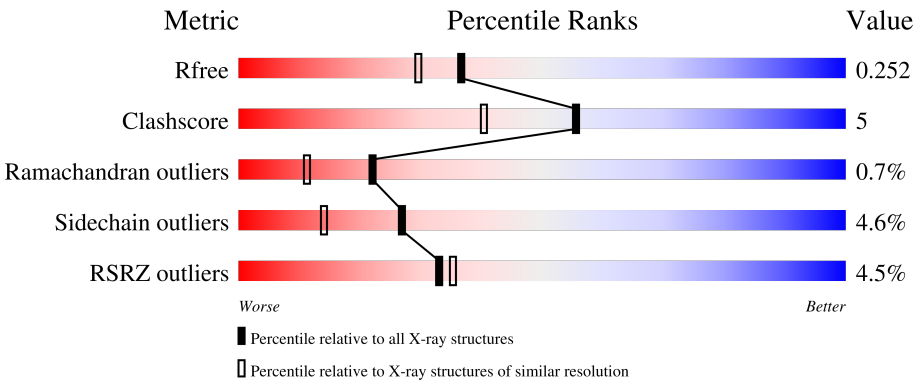
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)

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 1.98 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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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 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	500	4% 74% 11% •• 14%
1	B	500	4% 75% 11% • 13%

Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 7818 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 GDP-fucose protein O-fucosyltransferase 2,Thrombospondin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	429	3556	2256	614	671	15	0	2	0
1	B	435	3607	2291	625	675	16	0	0	0

There are 52 discrepancies between the modelled and reference sequences:

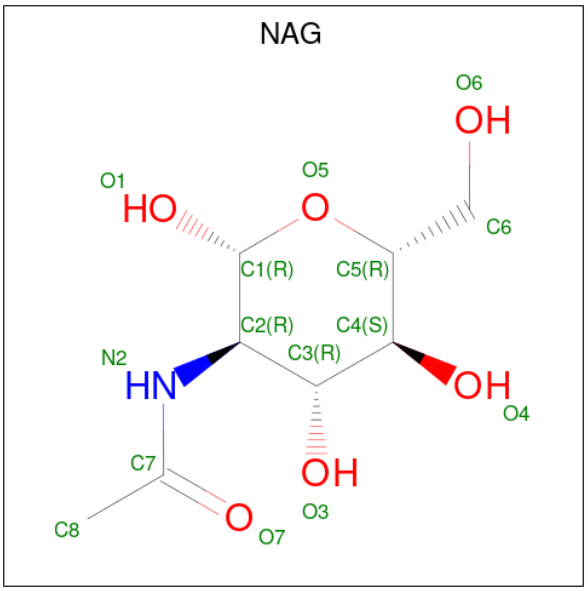
Chain	Residue	Modelled	Actual	Comment	Reference
A	298	LYS	ARG	engineered mutation	UNP Q8WR51
A	299	LYS	ARG	engineered mutation	UNP Q8WR51
A	425	SER	-	linker	UNP Q8WR51
A	426	SER	-	linker	UNP Q8WR51
A	979	SER	-	linker	UNP Q8WR51
A	980	SER	-	linker	UNP Q8WR51
A	981	GLY	-	linker	UNP Q8WR51
A	982	GLY	-	linker	UNP Q8WR51
A	983	GLY	-	linker	UNP Q8WR51
A	984	GLY	-	linker	UNP Q8WR51
A	985	SER	-	linker	UNP Q8WR51
A	986	GLY	-	linker	UNP Q8WR51
A	987	GLY	-	linker	UNP Q8WR51
A	988	GLY	-	linker	UNP Q8WR51
A	989	GLY	-	linker	UNP Q8WR51
A	990	SER	-	linker	UNP Q8WR51
A	991	GLY	-	linker	UNP Q8WR51
A	992	GLY	-	linker	UNP Q8WR51
A	993	GLY	-	linker	UNP Q8WR51
A	994	GLY	-	linker	UNP Q8WR51
A	995	GLY	-	linker	UNP Q8WR51
A	996	SER	-	linker	UNP Q8WR51
A	997	SER	-	linker	UNP Q8WR51
A	998	GLY	-	linker	UNP Q8WR51
A	999	SER	-	linker	UNP Q8WR51

Continued on next page...

Continued from previous page...

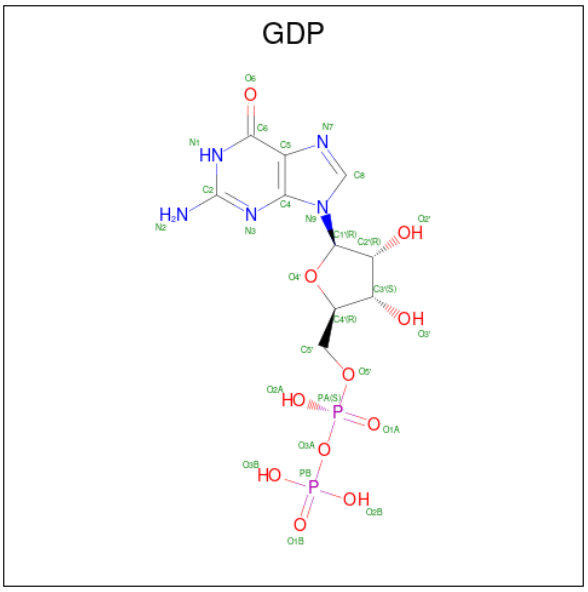
Chain	Residue	Modelled	Actual	Comment	Reference
A	1000	SER	-	linker	UNP Q8WR51
B	298	LYS	ARG	engineered mutation	UNP Q8WR51
B	299	LYS	ARG	engineered mutation	UNP Q8WR51
B	425	SER	-	linker	UNP Q8WR51
B	978	SER	-	linker	UNP Q8WR51
B	979	SER	-	linker	UNP Q8WR51
B	980	SER	-	linker	UNP Q8WR51
B	981	GLY	-	linker	UNP Q8WR51
B	982	GLY	-	linker	UNP Q8WR51
B	983	GLY	-	linker	UNP Q8WR51
B	984	GLY	-	linker	UNP Q8WR51
B	985	SER	-	linker	UNP Q8WR51
B	986	GLY	-	linker	UNP Q8WR51
B	987	GLY	-	linker	UNP Q8WR51
B	988	GLY	-	linker	UNP Q8WR51
B	989	GLY	-	linker	UNP Q8WR51
B	990	SER	-	linker	UNP Q8WR51
B	991	GLY	-	linker	UNP Q8WR51
B	992	GLY	-	linker	UNP Q8WR51
B	993	GLY	-	linker	UNP Q8WR51
B	994	GLY	-	linker	UNP Q8WR51
B	995	GLY	-	linker	UNP Q8WR51
B	996	SER	-	linker	UNP Q8WR51
B	997	SER	-	linker	UNP Q8WR51
B	998	GLY	-	linker	UNP Q8WR51
B	999	SER	-	linker	UNP Q8WR51
B	1000	SER	-	linker	UNP Q8WR51

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	14	8	1	5	0	0
			Total	C	N	O		
2	B	1	14	8	1	5	0	0

- Molecule 3 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: $C_{10}H_{15}N_5O_{11}P_2$).



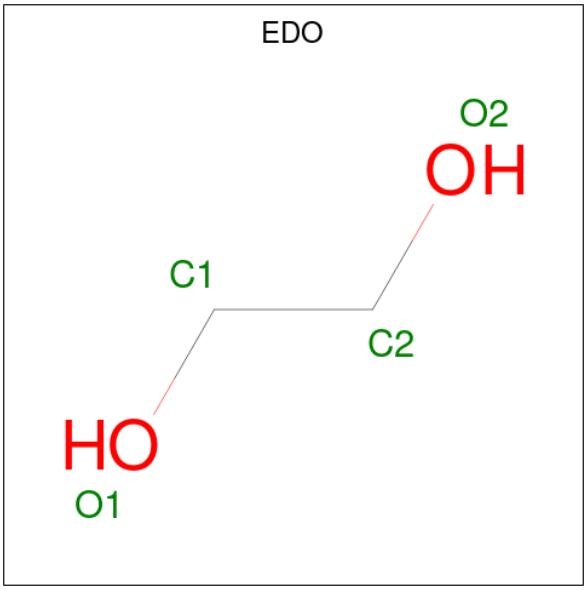
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	A	1	28	10	5	11	2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	B	1	28	10	5	11	2	0	0

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



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

- Molecule 5 is water.

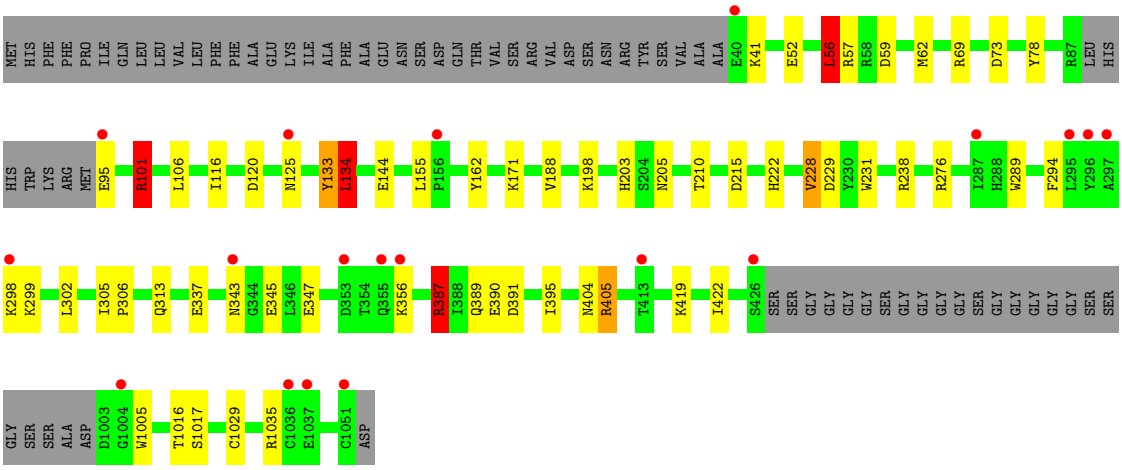
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	290	290	290	0	0
			Total	O		
5	B	241	241	241	0	0

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.

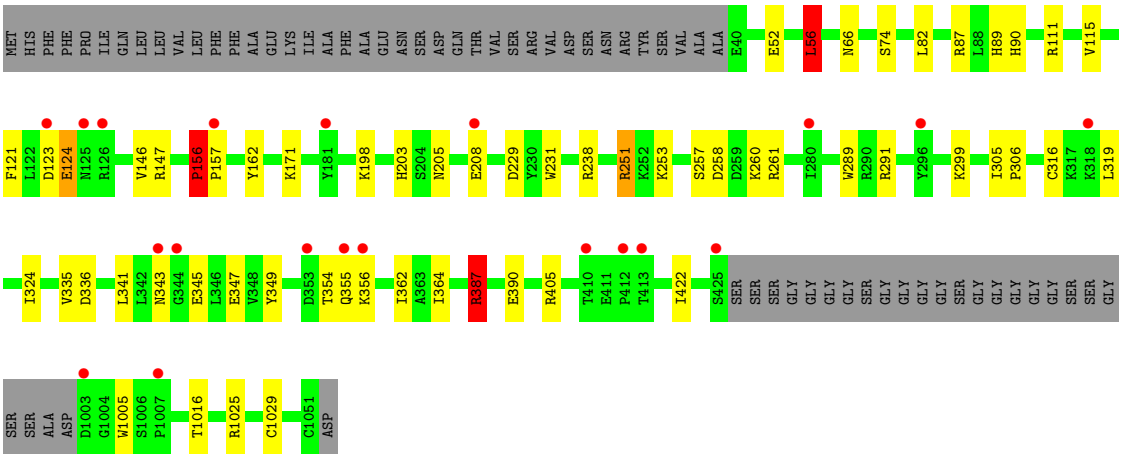
• Molecule 1: GDP-fucose protein O-fucosyltransferase 2,Thrombospondin-1

Chain A:



• Molecule 1: GDP-fucose protein O-fucosyltransferase 2,Thrombospondin-1

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	88.03Å 67.45Å 90.41Å 90.00° 116.07° 90.00°	Depositor
Resolution (Å)	81.21 – 1.98 20.00 – 1.98	Depositor EDS
% Data completeness (in resolution range)	99.8 (81.21-1.98) 99.9 (20.00-1.98)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.00 (at 1.97Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.200 , 0.246 0.206 , 0.252	Depositor DCC
R_{free} test set	1852 reflections (2.79%)	wwPDB-VP
Wilson B-factor (Å^2)	28.5	Xtriage
Anisotropy	0.183	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 43.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.034 for l,-k,h	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7818	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 31.10 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.1701e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, GDP, EDO

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	0.63	1/3650 (0.0%)	0.83	13/4935 (0.3%)
1	B	0.57	0/3700	0.80	8/5004 (0.2%)
All	All	0.60	1/7350 (0.0%)	0.81	21/9939 (0.2%)

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	B	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	134	LEU	N-CA	5.48	1.57	1.46

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	156	PRO	C-N-CD	-11.59	95.10	120.60
1	A	387	ARG	NE-CZ-NH2	-9.91	115.34	120.30
1	B	387	ARG	NE-CZ-NH2	-8.63	115.98	120.30
1	A	238	ARG	NE-CZ-NH2	-8.52	116.04	120.30
1	B	238	ARG	NE-CZ-NH2	-8.45	116.08	120.30
1	B	238	ARG	NE-CZ-NH1	8.11	124.35	120.30
1	A	387	ARG	NE-CZ-NH1	8.06	124.33	120.30
1	A	134	LEU	N-CA-CB	7.76	125.93	110.40
1	B	387	ARG	NE-CZ-NH1	7.48	124.04	120.30
1	A	238	ARG	NE-CZ-NH1	7.37	123.99	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	134	LEU	CA-CB-CG	7.30	132.09	115.30
1	A	405	ARG	NE-CZ-NH2	-7.29	116.66	120.30
1	A	133	TYR	O-C-N	-7.03	111.44	122.70
1	A	101	ARG	CB-CA-C	-6.89	96.62	110.40
1	A	405	ARG	NE-CZ-NH1	6.88	123.74	120.30
1	B	251	ARG	NE-CZ-NH1	6.55	123.58	120.30
1	A	228	VAL	CB-CA-C	-6.07	99.87	111.40
1	B	156	PRO	C-N-CA	5.55	145.30	122.00
1	B	56	LEU	CA-CB-CG	-5.27	103.18	115.30
1	A	56	LEU	CA-CB-CG	-5.20	103.34	115.30
1	A	57	ARG	NE-CZ-NH2	-5.13	117.73	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	156	PRO	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	3556	0	3417	40	0
1	B	3607	0	3464	33	0
2	A	14	0	13	0	0
2	B	14	0	13	0	0
3	A	28	0	12	0	0
3	B	28	0	12	0	0
4	A	20	0	30	0	0
4	B	20	0	30	1	0
5	A	290	0	0	10	0
5	B	241	0	0	4	0
All	All	7818	0	6991	73	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

All (73) 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:A:59:ASP:HA	1:A:62:MET:CE	1.94	0.98
1:B:316:CYS:SG	1:B:324:ILE:HD11	2.04	0.98
1:A:59:ASP:HA	1:A:62:MET:HE2	1.57	0.83
1:A:116:ILE:CG2	1:A:120:ASP:HB2	2.13	0.78
1:A:133:TYR:O	1:A:134:LEU:CB	2.33	0.73
1:B:258:ASP:OD2	1:B:261:ARG:NH1	2.21	0.73
1:A:116:ILE:HG22	1:A:120:ASP:HB2	1.72	0.71
1:B:89:HIS:CD2	1:B:90:HIS:HD2	2.10	0.70
1:A:305:ILE:HB	1:A:306:PRO:HD3	1.77	0.67
1:A:222:HIS:HD2	5:A:2021:HOH:O	1.78	0.66
1:A:101:ARG:HG2	5:A:2035:HOH:O	1.99	0.62
1:A:133:TYR:O	1:A:134:LEU:HB2	1.99	0.61
1:B:89:HIS:CD2	1:B:90:HIS:CD2	2.88	0.61
1:B:82:LEU:HD11	1:B:115:VAL:CG1	2.31	0.60
1:A:203:HIS:HD2	1:A:205:ASN:H	1.49	0.59
1:A:59:ASP:HA	1:A:62:MET:HE1	1.80	0.59
1:A:62:MET:HE1	5:A:2017:HOH:O	2.04	0.58
1:A:203:HIS:CD2	1:A:205:ASN:H	2.23	0.57
1:B:316:CYS:SG	1:B:324:ILE:CD1	2.88	0.57
1:A:276:ARG:NH1	5:A:2176:HOH:O	2.26	0.56
1:B:203:HIS:HD2	1:B:205:ASN:H	1.54	0.55
1:A:69[B]:ARG:HD2	1:A:73:ASP:OD2	2.06	0.55
1:A:389:GLN:NE2	5:A:2232:HOH:O	2.41	0.54
1:B:316:CYS:HG	1:B:324:ILE:HD11	1.71	0.54
1:B:123:ASP:O	1:B:124:GLU:HB2	2.09	0.53
1:B:203:HIS:CD2	1:B:205:ASN:H	2.28	0.52
1:B:82:LEU:HD11	1:B:115:VAL:HG11	1.92	0.52
1:A:52:GLU:CD	1:A:56:LEU:HD22	2.31	0.50
1:B:251:ARG:HD2	1:B:257:SER:OG	2.10	0.50
1:A:62:MET:HE1	5:A:2015:HOH:O	2.12	0.49
1:B:289:TRP:CE2	1:B:291:ARG:HD3	2.47	0.49
1:A:59:ASP:OD2	1:A:222:HIS:HE1	1.96	0.48
1:B:52:GLU:CD	1:B:56:LEU:HD22	2.34	0.48
1:A:171:LYS:HD2	5:A:2117:HOH:O	2.13	0.48
1:A:419:LYS:HE2	5:A:2240:HOH:O	2.13	0.47
1:A:133:TYR:O	1:A:215:ASP:O	2.34	0.46
1:B:1005:TRP:CE3	1:B:1029:CYS:HB2	2.51	0.46
1:B:305:ILE:HB	1:B:306:PRO:HD3	1.98	0.46
1:A:294:PHE:O	1:A:298:LYS:HB2	2.16	0.46
1:A:1005:TRP:CE3	1:A:1029:CYS:HB2	2.51	0.46
1:B:354:THR:HB	1:B:362:ILE:HD11	1.98	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:355:GLN:HA	1:B:355:GLN:OE1	2.16	0.45
1:A:313:GLN:HE21	1:A:343:ASN:H	1.64	0.45
1:B:336:ASP:OD2	5:B:2174:HOH:O	2.21	0.45
1:A:69[B]:ARG:HH12	1:A:229:ASP:CG	2.21	0.45
1:A:356:LYS:N	1:A:356:LYS:HD2	2.33	0.44
1:A:387:ARG:CD	1:A:390:GLU:OE1	2.66	0.44
1:A:134:LEU:HA	1:A:188:VAL:O	2.17	0.44
1:B:198:LYS:HE3	1:B:198:LYS:HB3	1.76	0.44
1:B:111:ARG:NH1	5:B:2047:HOH:O	2.50	0.43
1:A:101:ARG:CG	5:A:2035:HOH:O	2.59	0.43
1:B:147:ARG:NH2	5:B:2080:HOH:O	2.52	0.43
1:A:391:ASP:O	1:A:395:ILE:HG12	2.17	0.43
1:B:231:TRP:CE2	1:B:422:ILE:HA	2.54	0.43
1:B:121:PHE:O	1:B:124:GLU:HB3	2.19	0.43
1:A:231:TRP:CE2	1:A:422:ILE:HA	2.54	0.43
1:B:291:ARG:HD2	5:B:2159:HOH:O	2.19	0.43
1:A:56:LEU:HD13	1:A:1017:SER:N	2.34	0.42
1:B:66:ASN:HD21	1:B:229:ASP:HB3	1.84	0.42
1:B:82:LEU:HG	1:B:115:VAL:HG13	2.01	0.42
1:B:146:VAL:HB	4:B:1430:EDO:H22	2.02	0.42
1:B:387:ARG:CD	1:B:390:GLU:OE1	2.68	0.42
1:A:69[B]:ARG:NH1	1:A:229:ASP:CG	2.73	0.42
1:A:387:ARG:HD3	1:A:390:GLU:OE1	2.20	0.42
1:B:56:LEU:O	1:B:56:LEU:HG	2.18	0.42
1:A:198:LYS:HE3	1:A:198:LYS:HB3	1.80	0.42
1:B:387:ARG:HD3	1:B:390:GLU:OE1	2.21	0.41
1:A:56:LEU:O	1:A:56:LEU:HG	2.20	0.41
1:B:162:TYR:HA	1:B:171:LYS:O	2.21	0.41
1:A:162:TYR:HA	1:A:171:LYS:O	2.21	0.41
1:A:404:ASN:HB3	5:A:2241:HOH:O	2.21	0.40
1:B:347:GLU:HG2	1:B:349:TYR:CE1	2.56	0.40
1:A:289:TRP:CH2	1:A:305:ILE:HD12	2.56	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries

of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	425/500 (85%)	411 (97%)	13 (3%)	1 (0%)	47	38
1	B	431/500 (86%)	413 (96%)	13 (3%)	5 (1%)	13	4
All	All	856/1000 (86%)	824 (96%)	26 (3%)	6 (1%)	22	11

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	124	GLU
1	B	157	PRO
1	B	74	SER
1	A	134	LEU
1	B	87	ARG
1	B	156	PRO

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	388/437 (89%)	368 (95%)	20 (5%)	23	11
1	B	392/437 (90%)	376 (96%)	16 (4%)	30	18
All	All	780/874 (89%)	744 (95%)	36 (5%)	27	14

All (36) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	41	LYS
1	A	56	LEU
1	A	78	TYR
1	A	95	GLU
1	A	101	ARG
1	A	106	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	125	ASN
1	A	144	GLU
1	A	155	LEU
1	A	210	THR
1	A	228	VAL
1	A	299	LYS
1	A	302	LEU
1	A	337	GLU
1	A	345	GLU
1	A	347	GLU
1	A	387	ARG
1	A	405	ARG
1	A	1016	THR
1	A	1035	ARG
1	B	56	LEU
1	B	208	GLU
1	B	253	LYS
1	B	260	LYS
1	B	299	LYS
1	B	319	LEU
1	B	335	VAL
1	B	341	LEU
1	B	343	ASN
1	B	345	GLU
1	B	356	LYS
1	B	364	ILE
1	B	387	ARG
1	B	405	ARG
1	B	1016	THR
1	B	1025	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (17) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	66	ASN
1	A	191	GLN
1	A	203	HIS
1	A	222	HIS
1	A	313	GLN
1	A	322	GLN
1	A	389	GLN
1	B	89	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	90	HIS
1	B	110	ASN
1	B	161	HIS
1	B	203	HIS
1	B	241	ASN
1	B	404	ASN
1	B	1020	ASN
1	B	1033	ASN
1	B	1047	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)

14 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	$\#\|Z\| > 2$
4	EDO	B	1431	-	3,3,3	0.73	0	2,2,2	0.58	0
4	EDO	A	1433	-	3,3,3	0.30	0	2,2,2	0.58	0
3	GDP	B	1426	-	24,30,30	1.04	1 (4%)	30,47,47	1.16	4 (13%)
3	GDP	A	1428	-	24,30,30	0.90	0	30,47,47	1.38	6 (20%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	NAG	B	1427	1	14,14,15	0.73	0	17,19,21	1.81	4 (23%)
4	EDO	A	1431	-	3,3,3	0.39	0	2,2,2	0.45	0
2	NAG	A	1427	1	14,14,15	0.61	0	17,19,21	1.88	2 (11%)
4	EDO	A	1432	-	3,3,3	0.48	0	2,2,2	0.57	0
4	EDO	B	1429	-	3,3,3	0.40	0	2,2,2	0.23	0
4	EDO	B	1430	-	3,3,3	0.33	0	2,2,2	0.63	0
4	EDO	A	1430	-	3,3,3	0.48	0	2,2,2	0.17	0
4	EDO	B	1428	-	3,3,3	0.43	0	2,2,2	0.49	0
4	EDO	B	1432	-	3,3,3	0.42	0	2,2,2	0.59	0
4	EDO	A	1429	-	3,3,3	0.41	0	2,2,2	0.51	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
4	EDO	B	1431	-	-	1/1/1/1	-
4	EDO	A	1433	-	-	0/1/1/1	-
3	GDP	B	1426	-	-	3/12/32/32	0/3/3/3
3	GDP	A	1428	-	-	3/12/32/32	0/3/3/3
2	NAG	B	1427	1	-	0/6/23/26	0/1/1/1
4	EDO	A	1431	-	-	0/1/1/1	-
2	NAG	A	1427	1	-	0/6/23/26	0/1/1/1
4	EDO	A	1432	-	-	1/1/1/1	-
4	EDO	B	1429	-	-	0/1/1/1	-
4	EDO	B	1430	-	-	1/1/1/1	-
4	EDO	A	1430	-	-	1/1/1/1	-
4	EDO	B	1428	-	-	1/1/1/1	-
4	EDO	B	1432	-	-	0/1/1/1	-
4	EDO	A	1429	-	-	0/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1426	GDP	C6-N1	-2.15	1.34	1.37

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1427	NAG	C1-O5-C5	4.82	118.72	112.19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1427	NAG	O5-C1-C2	-4.37	104.38	111.29
2	A	1427	NAG	C6-C5-C4	-4.18	103.22	113.00
2	B	1427	NAG	C4-C3-C2	-2.79	106.92	111.02
3	A	1428	GDP	O3B-PB-O3A	-2.73	95.47	104.64
3	A	1428	GDP	C8-N7-C5	2.70	108.14	102.99
3	A	1428	GDP	O6-C6-C5	-2.65	119.19	124.37
3	B	1426	GDP	C5-C6-N1	2.47	118.31	113.95
3	B	1426	GDP	C8-N7-C5	2.46	107.67	102.99
2	B	1427	NAG	C1-C2-N2	2.35	114.51	110.49
2	B	1427	NAG	C2-N2-C7	-2.27	119.67	122.90
3	B	1426	GDP	O3B-PB-O3A	-2.26	97.06	104.64
3	A	1428	GDP	O3'-C3'-C4'	2.19	117.39	111.05
3	A	1428	GDP	O3B-PB-O1B	2.17	119.17	110.68
3	A	1428	GDP	C5-C6-N1	2.11	117.68	113.95
3	B	1426	GDP	O3'-C3'-C4'	2.00	116.83	111.05

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1428	GDP	PB-O3A-PA-O5'
4	B	1428	EDO	O1-C1-C2-O2
4	B	1431	EDO	O1-C1-C2-O2
3	B	1426	GDP	C5'-O5'-PA-O3A
3	A	1428	GDP	C4'-C5'-O5'-PA
3	B	1426	GDP	C4'-C5'-O5'-PA
3	B	1426	GDP	O4'-C4'-C5'-O5'
4	A	1430	EDO	O1-C1-C2-O2
4	B	1430	EDO	O1-C1-C2-O2
3	A	1428	GDP	O4'-C4'-C5'-O5'
4	A	1432	EDO	O1-C1-C2-O2

There are no ring outliers.

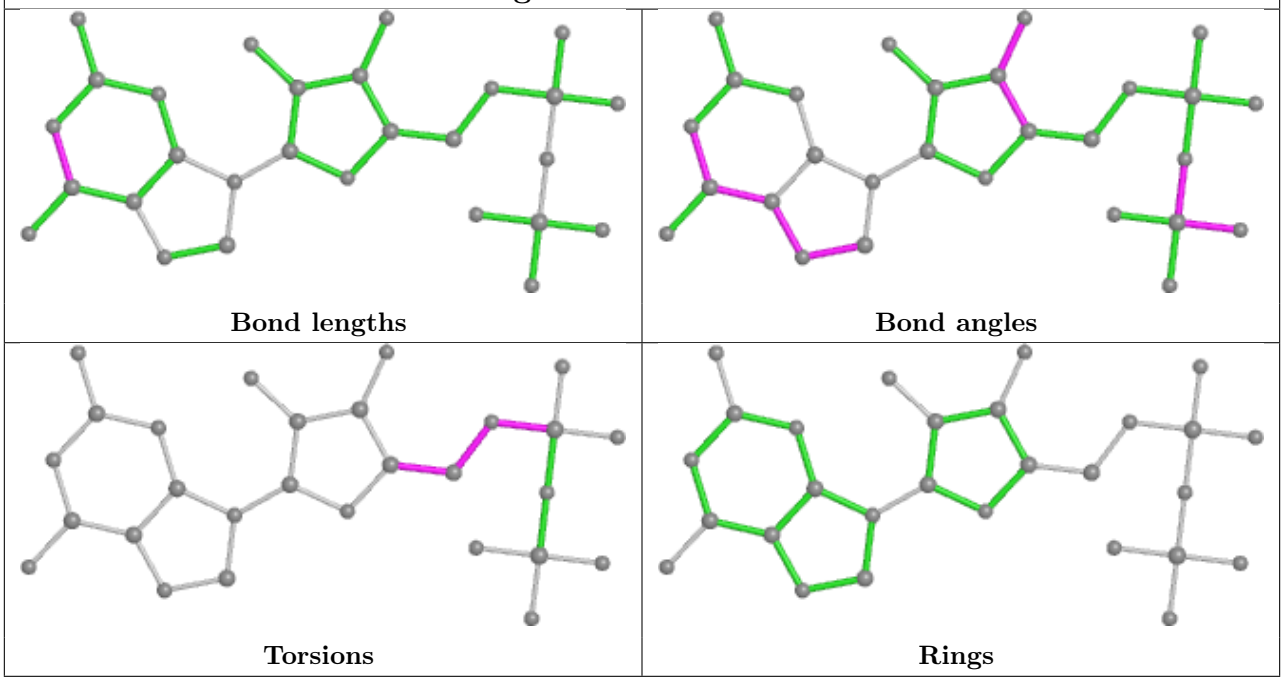
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	1430	EDO	1	0

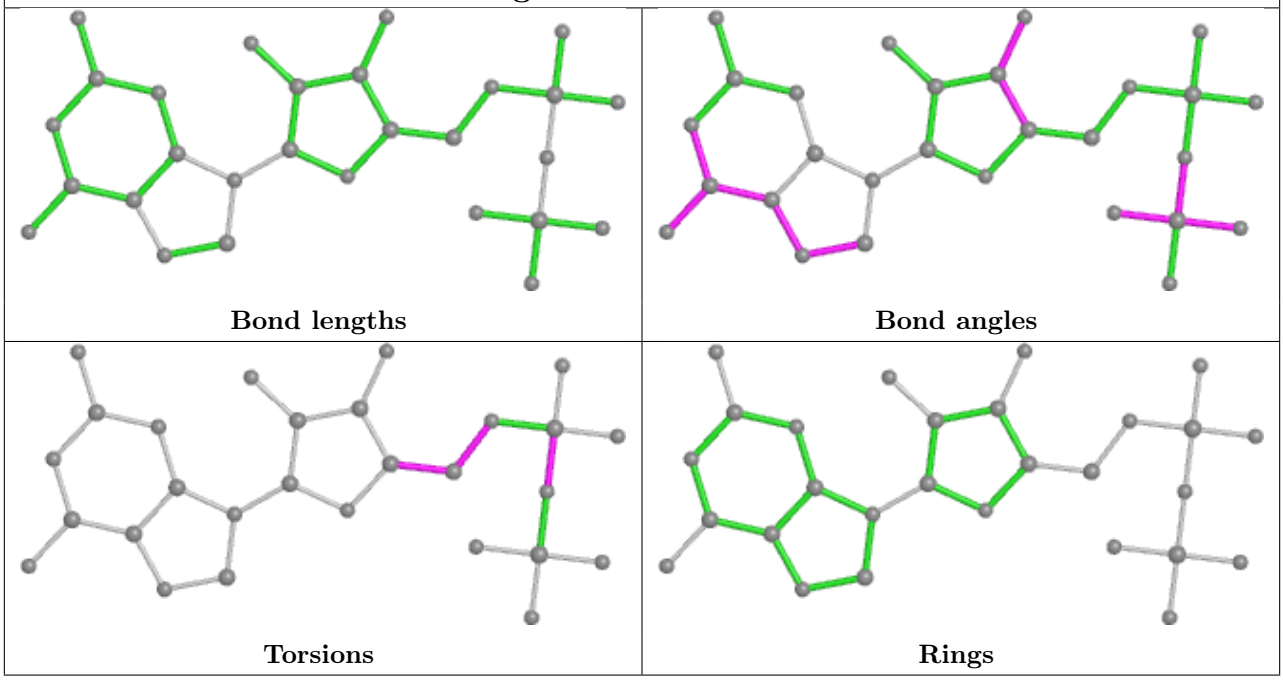
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will

also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand GDP B 1426



Ligand GDP A 1428



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	429/500 (85%)	0.17	19 (4%)	34	36	19, 30, 51, 99	0
1	B	435/500 (87%)	0.34	20 (4%)	32	34	22, 37, 57, 75	0
All	All	864/1000 (86%)	0.26	39 (4%)	33	35	19, 34, 56, 99	0

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	296	TYR	10.8
1	B	296	TYR	7.2
1	A	95	GLU	5.2
1	B	1003	ASP	5.0
1	A	413	THR	4.2
1	A	295	LEU	3.8
1	B	355	GLN	3.8
1	A	356	LYS	3.6
1	A	297	ALA	3.5
1	B	413	THR	3.4
1	B	344	GLY	3.4
1	B	343	ASN	3.4
1	B	123	ASP	3.4
1	A	426	SER	3.3
1	B	356	LYS	3.0
1	A	156	PRO	2.8
1	B	318	LYS	2.8
1	A	1051	CYS	2.8
1	A	125	ASN	2.7
1	B	425	SER	2.6
1	A	355	GLN	2.6
1	A	40	GLU	2.6
1	B	412	PRO	2.5
1	A	353	ASP	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	125	ASN	2.4
1	A	287	ILE	2.4
1	A	1004	GLY	2.4
1	B	181	TYR	2.3
1	B	157	PRO	2.3
1	B	280	ILE	2.3
1	A	298	LYS	2.3
1	B	126	ARG	2.2
1	B	1007	PRO	2.2
1	B	353	ASP	2.2
1	A	1037	GLU	2.1
1	B	208	GLU	2.1
1	A	343	ASN	2.1
1	A	1036	CYS	2.1
1	B	410	THR	2.0

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	EDO	B	1428	4/4	0.70	0.33	58,59,59,64	0
4	EDO	A	1431	4/4	0.81	0.19	50,50,52,57	0
4	EDO	B	1429	4/4	0.83	0.14	60,60,60,61	0
4	EDO	A	1432	4/4	0.84	0.19	42,46,48,50	0
4	EDO	A	1429	4/4	0.86	0.14	46,47,48,51	0
4	EDO	B	1430	4/4	0.86	0.23	38,42,45,47	0
4	EDO	B	1431	4/4	0.86	0.16	38,41,42,43	0
4	EDO	A	1433	4/4	0.88	0.18	44,45,49,50	0

Continued on next page...

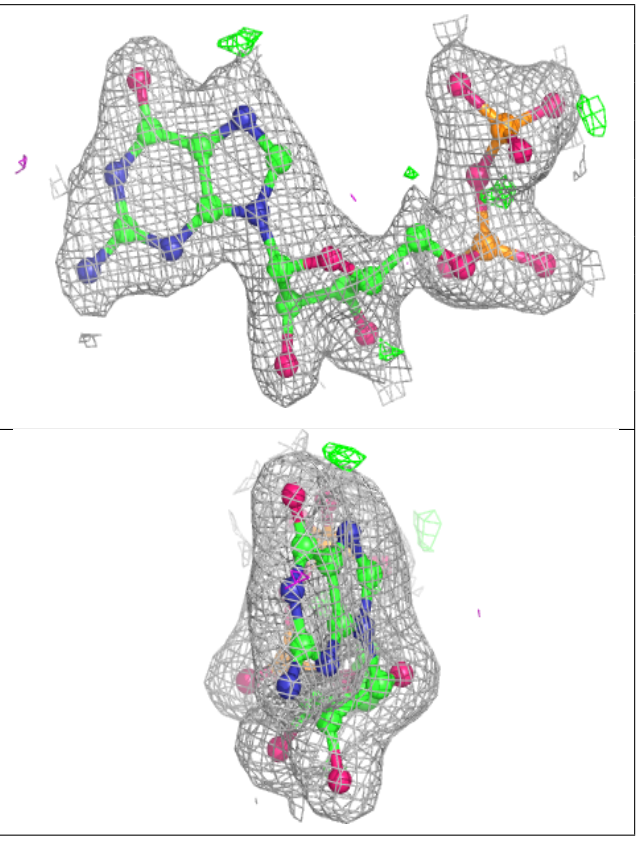
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAG	B	1427	14/15	0.89	0.20	43,52,58,59	0
4	EDO	A	1430	4/4	0.90	0.19	54,54,54,55	0
2	NAG	A	1427	14/15	0.91	0.11	32,37,40,42	0
4	EDO	B	1432	4/4	0.95	0.10	32,36,40,40	0
3	GDP	B	1426	28/28	0.96	0.09	22,29,31,31	0
3	GDP	A	1428	28/28	0.97	0.08	23,27,29,31	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.

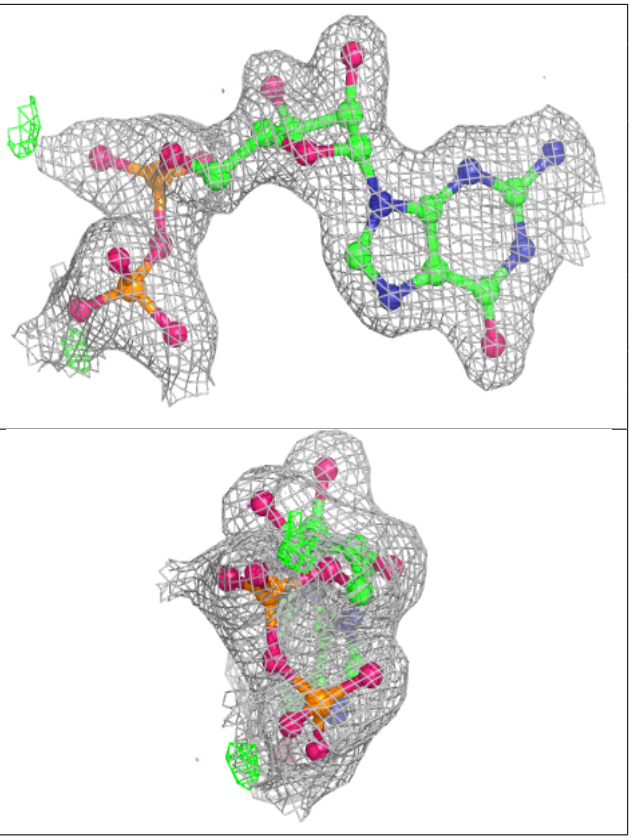
Electron density around GDP B 1426:

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 GDP A 1428:

$2mF_o$-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.