



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 16, 2023 – 03:44 AM EDT

PDB ID : 1YRO  
Title : Crystal structure of beta14,-galactosyltransferase mutant ARG228Lys in complex with alpha-lactalbumin in the presence of UDP-galactose and Mn  
Authors : Ramakrishnan, B.; Boeggeman, E.; Qasba, P.K.  
Deposited on : 2005-02-04  
Resolution : 1.90 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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.35  
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) : 2.35

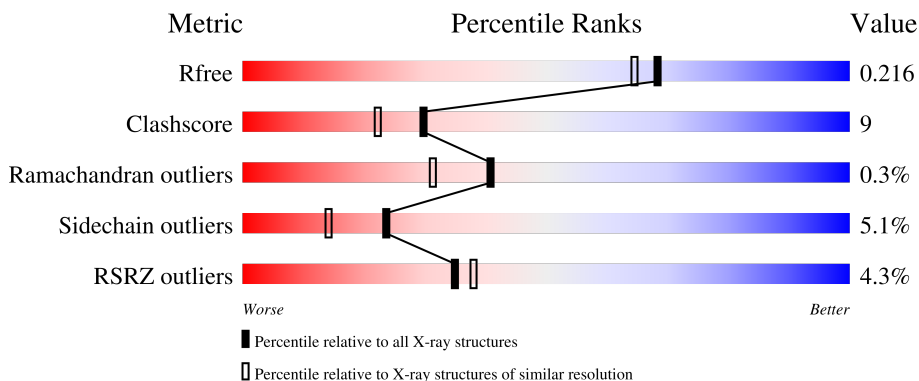
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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  $\geq 3$ , 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	123	 5% (poor fit)   85% (0-1 outliers)   13% (2-3 outliers)   .
1	C	123	 5% (poor fit)   85% (0-1 outliers)   12% (2-3 outliers)   ..
2	B	286	 4% (poor fit)   77% (0-1 outliers)   15% (2-3 outliers)   . 5% (not modelled)
2	D	286	 4% (poor fit)   77% (0-1 outliers)   17% (2-3 outliers)   . 5% (not modelled)

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 7388 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 ALPHA-LACTALBUMIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	123	980	620	156	195	9	0	0	0
1	C	123	980	620	156	195	9	0	0	0

- Molecule 2 is a protein called BETA-1,4-GALACTOSYLTRANSFERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	272	2217	1425	380	399	13	0	0	0
2	D	272	2217	1425	380	399	13	0	0	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	117	ALA	-	SEE REMARK 999	UNP P08037
B	118	SER	-	SEE REMARK 999	UNP P08037
B	119	MET	-	SEE REMARK 999	UNP P08037
B	120	THR	-	SEE REMARK 999	UNP P08037
B	121	GLY	-	SEE REMARK 999	UNP P08037
B	122	GLY	-	SEE REMARK 999	UNP P08037
B	123	GLN	-	SEE REMARK 999	UNP P08037
B	124	GLN	-	SEE REMARK 999	UNP P08037
B	125	MET	-	SEE REMARK 999	UNP P08037
B	126	GLY	-	SEE REMARK 999	UNP P08037
B	127	ARG	-	SEE REMARK 999	UNP P08037
B	128	GLY	-	SEE REMARK 999	UNP P08037
B	129	SER	-	SEE REMARK 999	UNP P08037
B	228	LYS	ARG	engineered mutation	UNP P08037
B	342	THR	CYS	engineered mutation	UNP P08037
D	117	ALA	-	SEE REMARK 999	UNP P08037

*Continued on next page...*

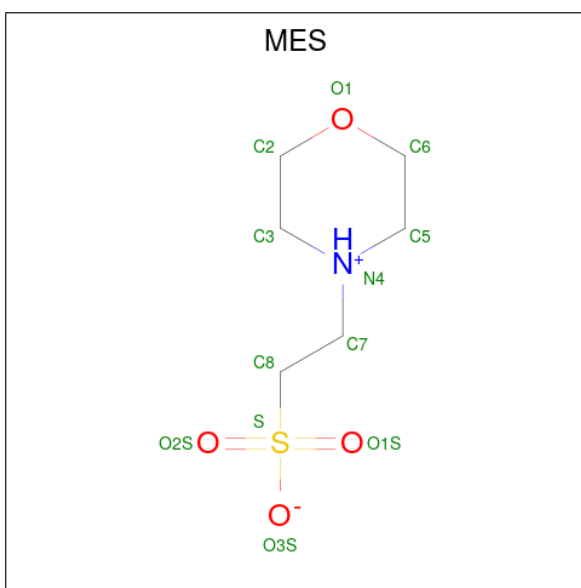
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	118	SER	-	SEE REMARK 999	UNP P08037
D	119	MET	-	SEE REMARK 999	UNP P08037
D	120	THR	-	SEE REMARK 999	UNP P08037
D	121	GLY	-	SEE REMARK 999	UNP P08037
D	122	GLY	-	SEE REMARK 999	UNP P08037
D	123	GLN	-	SEE REMARK 999	UNP P08037
D	124	GLN	-	SEE REMARK 999	UNP P08037
D	125	MET	-	SEE REMARK 999	UNP P08037
D	126	GLY	-	SEE REMARK 999	UNP P08037
D	127	ARG	-	SEE REMARK 999	UNP P08037
D	128	GLY	-	SEE REMARK 999	UNP P08037
D	129	SER	-	SEE REMARK 999	UNP P08037
D	228	LYS	ARG	engineered mutation	UNP P08037
D	342	THR	CYS	engineered mutation	UNP P08037

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Ca 1 1	0	0
3	C	1	Total Ca 1 1	0	0

- Molecule 4 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).



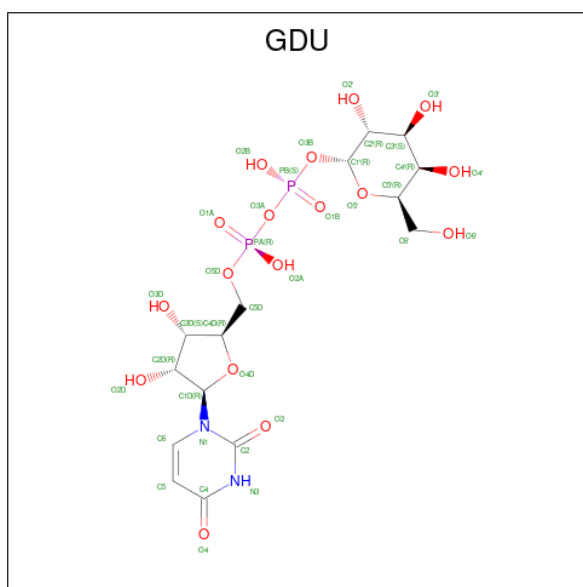
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
4	A	1	12	6	1	4	1	0	0

- Molecule 5 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>5</sub>).



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

- Molecule 6 is GALACTOSE-URIDINE-5'-DIPHOSPHATE (three-letter code: GDU) (formula: C<sub>15</sub>H<sub>24</sub>N<sub>2</sub>O<sub>17</sub>P<sub>2</sub>).

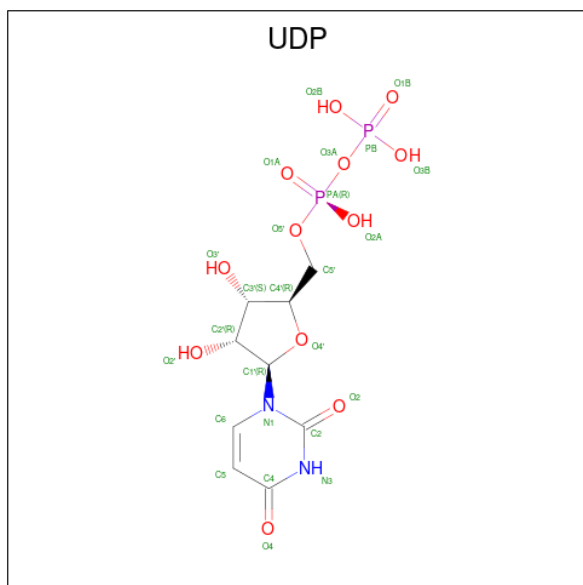


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf		
			Total	C	N	O			P	
6	B	1	Total	36	15	2	17	2	0	0
6	D	1	Total	36	15	2	17	2	0	0

- Molecule 7 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mn		
7	B	1	Total	1	0	0
7	D	1	Total	1	0	0

- Molecule 8 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula: C<sub>9</sub>H<sub>14</sub>N<sub>2</sub>O<sub>12</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
8	B	1	25	9	2	12	2	0	0
8	D	1	25	9	2	12	2	0	0

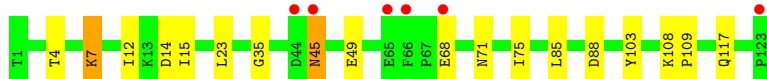
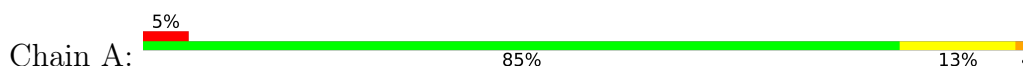
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	156	Total	O	0	0
			156	156		
9	B	253	Total	O	0	0
			253	253		
9	C	151	Total	O	0	0
			151	151		
9	D	270	Total	O	0	0
			270	270		

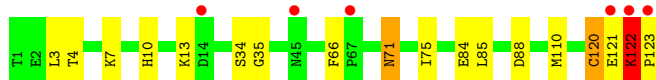
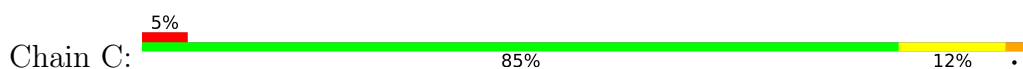
### 3 Residue-property plots [i](#)

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.

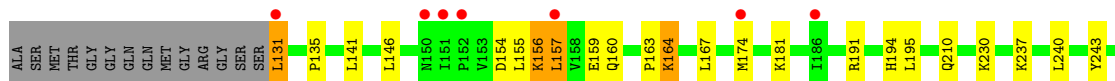
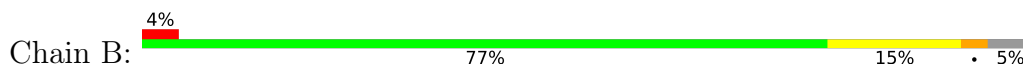
- Molecule 1: ALPHA-LACTALBUMIN



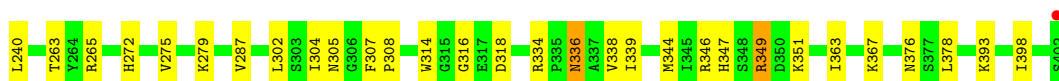
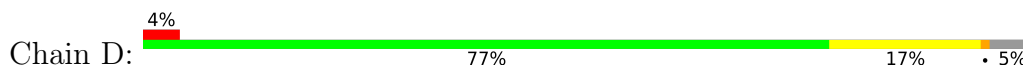
- Molecule 1: ALPHA-LACTALBUMIN



- Molecule 2: BETA-1,4-GALACTOSYLTRANSFERASE



- Molecule 2: BETA-1,4-GALACTOSYLTRANSFERASE





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.33Å 94.09Å 99.70Å 90.00° 101.64° 90.00°	Depositor
Resolution (Å)	32.55 – 1.90 30.76 – 1.88	Depositor EDS
% Data completeness (in resolution range)	97.9 (32.55-1.90) 97.1 (30.76-1.88)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.70 (at 1.88Å)	Xtrriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.189 , 0.225 0.179 , 0.216	Depositor DCC
$R_{free}$ test set	8001 reflections (9.72%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.2	Xtrriage
Anisotropy	0.216	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 49.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7388	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.30% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: UDP, PG4, GDU, CA, MN, MES

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	0.58	0/1001	0.76	0/1350
1	C	0.62	0/1001	0.85	1/1350 (0.1%)
2	B	0.56	0/2277	0.76	1/3084 (0.0%)
2	D	0.57	0/2277	0.76	0/3084
All	All	0.58	0/6556	0.77	2/8868 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	122	LYS	N-CA-C	9.17	135.76	111.00
2	B	296	LEU	CA-CB-CG	5.55	128.07	115.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	980	0	936	14	0
1	C	980	0	936	19	0
2	B	2217	0	2187	46	0
2	D	2217	0	2187	43	0
3	A	1	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	1	0	0	0	0
4	A	12	0	13	2	0
5	A	13	0	18	1	0
5	C	13	0	18	2	0
6	B	36	0	22	1	0
6	D	36	0	22	4	0
7	B	1	0	0	0	0
7	D	1	0	0	0	0
8	B	25	0	11	0	0
8	D	25	0	11	0	0
9	A	156	0	0	3	0
9	B	253	0	0	7	0
9	C	151	0	0	4	0
9	D	270	0	0	7	0
All	All	7388	0	6361	121	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 9.

All (121) 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)
2:B:346:ARG:HB3	2:B:346:ARG:HH11	1.30	0.96
2:D:349:ARG:HH11	2:D:349:ARG:HG2	1.39	0.85
2:D:346:ARG:HD3	9:D:1182:HOH:O	1.78	0.82
1:A:71:ASN:HD21	1:A:75:ILE:H	1.28	0.81
2:B:327:PHE:CZ	2:B:367:LYS:HB2	2.16	0.80
2:B:305:ASN:HD21	2:B:376:ASN:H	1.30	0.77
2:D:151:ILE:H	2:D:151:ILE:HD13	1.53	0.74
2:B:237:LYS:HD2	9:B:1398:HOH:O	1.89	0.73
1:C:4:THR:H	1:C:7:LYS:HE3	1.54	0.72
2:B:277:MET:CE	2:B:279:LYS:HE2	2.21	0.70
2:B:336:ASN:HD22	2:B:338:VAL:H	1.40	0.69
2:B:191:ARG:HH11	2:B:194:HIS:HD2	1.39	0.69
2:B:277:MET:HE1	2:B:279:LYS:HE2	1.72	0.69
2:B:346:ARG:HH11	2:B:346:ARG:CB	2.06	0.68
2:D:336:ASN:HD22	2:D:338:VAL:H	1.41	0.67
2:D:191:ARG:HH11	2:D:194:HIS:HD2	1.43	0.65
1:C:71:ASN:HD21	1:C:75:ILE:H	1.43	0.64
2:B:349:ARG:HD3	2:B:350:ASP:H	1.62	0.64
2:B:131:LEU:N	2:B:131:LEU:HD23	2.13	0.63

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:34:SER:HA	5:C:807:PG4:H51	1.81	0.63
2:D:155:LEU:O	2:D:159:GLU:HG3	1.97	0.63
2:B:352:LYS:HE3	9:B:1361:HOH:O	1.99	0.62
2:D:279:LYS:HD2	2:D:346:ARG:HH12	1.64	0.62
2:D:305:ASN:HD21	2:D:376:ASN:H	1.48	0.62
2:D:349:ARG:HH11	2:D:349:ARG:CG	2.11	0.62
2:B:351:LYS:O	2:B:352:LYS:HB2	1.99	0.61
1:C:7:LYS:HD3	9:C:1216:HOH:O	2.00	0.61
2:B:174:MET:HG2	9:B:1468:HOH:O	2.01	0.60
2:D:316:GLY:HA2	2:D:363:ILE:HD11	1.83	0.60
2:D:151:ILE:H	2:D:151:ILE:CD1	2.15	0.60
1:C:66:PHE:HB2	9:C:996:HOH:O	2.02	0.59
2:B:191:ARG:NH1	2:B:194:HIS:HD2	2.00	0.59
2:B:272:HIS:HB3	2:B:334:ARG:HG2	1.85	0.58
1:A:7:LYS:HE3	9:A:1510:HOH:O	2.03	0.58
2:D:151:ILE:HD13	2:D:151:ILE:N	2.18	0.58
1:C:35:GLY:N	5:C:807:PG4:H32	2.18	0.57
2:B:160:GLN:O	2:B:163:PRO:HD3	2.06	0.56
1:C:122:LYS:HG3	9:C:1270:HOH:O	2.04	0.56
1:A:12:ILE:O	1:A:15:ILE:HG22	2.06	0.56
1:A:49:GLU:OE1	4:A:805:MES:H52	2.07	0.55
2:D:164:LYS:HD2	9:D:1443:HOH:O	2.05	0.55
1:C:4:THR:HG23	1:C:7:LYS:HE2	1.87	0.55
2:B:365:HIS:O	2:B:368:GLU:HG2	2.08	0.54
2:B:191:ARG:HD2	2:B:194:HIS:CD2	2.43	0.54
2:D:132:THR:HG22	2:D:133:ALA:N	2.22	0.53
1:A:108:LYS:HB3	1:A:109:PRO:CD	2.39	0.52
2:D:314:TRP:NE1	6:D:528:GDU:H6'1	2.24	0.52
2:D:279:LYS:HB2	2:D:346:ARG:NH1	2.23	0.52
2:D:344:MET:HE3	2:D:347:HIS:CD2	2.43	0.52
1:C:110:MET:HE3	9:D:1407:HOH:O	2.08	0.52
2:D:367:LYS:HE3	9:D:1726:HOH:O	2.09	0.52
1:C:10:HIS:CD2	1:C:13:LYS:NZ	2.78	0.52
1:A:45:ASN:O	1:A:45:ASN:OD1	2.29	0.51
2:B:327:PHE:CE2	2:B:367:LYS:HD2	2.46	0.51
1:C:10:HIS:O	1:C:13:LYS:HG3	2.11	0.51
1:A:117:GLN:NE2	2:B:287:VAL:HG12	2.25	0.51
2:B:327:PHE:CE1	2:B:367:LYS:HB2	2.45	0.51
2:B:336:ASN:ND2	2:B:339:ILE:H	2.08	0.51
2:D:318:ASP:OD2	6:D:528:GDU:O4'	2.22	0.51
2:B:316:GLY:HA2	2:B:363:ILE:HD11	1.93	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:7:LYS:HG3	9:A:1545:HOH:O	2.11	0.51
1:C:84:GLU:HG3	9:C:1148:HOH:O	2.10	0.50
2:B:164:LYS:HD3	9:B:1026:HOH:O	2.12	0.50
1:A:68:GLU:HG2	1:A:68:GLU:O	2.11	0.50
2:B:164:LYS:HE2	9:B:1341:HOH:O	2.12	0.50
1:C:4:THR:H	1:C:7:LYS:CE	2.23	0.49
2:D:314:TRP:CD1	6:D:528:GDU:H6'1	2.47	0.49
2:B:155:LEU:O	2:B:159:GLU:HG3	2.13	0.49
1:C:120:CYS:O	1:C:122:LYS:N	2.45	0.49
2:D:156:LYS:O	2:D:160:GLN:HG3	2.13	0.49
2:B:336:ASN:HD22	2:B:338:VAL:N	2.08	0.48
2:B:336:ASN:ND2	2:B:338:VAL:H	2.08	0.48
2:B:277:MET:HE1	9:B:964:HOH:O	2.14	0.48
2:D:191:ARG:NH1	2:D:194:HIS:HD2	2.09	0.47
2:D:230:LYS:HD3	2:D:398:ILE:HB	1.96	0.47
1:C:4:THR:N	1:C:7:LYS:HE3	2.27	0.47
2:D:336:ASN:ND2	2:D:339:ILE:H	2.12	0.47
1:A:4:THR:OG1	1:A:7:LYS:HB2	2.15	0.46
2:D:272:HIS:HB3	2:D:334:ARG:HG2	1.97	0.46
2:D:275:VAL:HG11	9:D:1362:HOH:O	2.15	0.46
2:B:194:HIS:HE1	9:B:1132:HOH:O	1.99	0.46
2:D:237:LYS:HD3	2:D:378:LEU:HD23	1.98	0.45
2:D:132:THR:CG2	2:D:133:ALA:N	2.79	0.45
2:B:230:LYS:HD3	2:B:398:ILE:HB	1.98	0.45
2:D:307:PHE:HB3	2:D:308:PRO:HD2	1.99	0.45
2:B:277:MET:HE2	2:B:279:LYS:HE2	1.95	0.45
2:D:314:TRP:CE2	6:D:528:GDU:H6'1	2.51	0.45
2:B:135:PRO:HD2	2:B:210:GLN:HE22	1.81	0.45
1:C:10:HIS:HA	1:C:13:LYS:HE3	1.99	0.45
2:D:187:PRO:HD3	2:D:232:LEU:HD21	1.99	0.45
2:B:277:MET:HE2	2:B:279:LYS:HG2	2.00	0.44
2:D:166:LYS:HE3	9:D:1447:HOH:O	2.16	0.44
2:D:191:ARG:HD2	2:D:194:HIS:CD2	2.53	0.44
1:A:103:TYR:CE2	4:A:805:MES:H51	2.53	0.44
2:D:263:THR:HG22	2:D:265:ARG:HG3	2.00	0.44
1:A:108:LYS:HB3	1:A:109:PRO:HD3	1.99	0.43
1:C:10:HIS:CD2	1:C:13:LYS:HE3	2.54	0.43
2:B:282:PHE:CD2	2:B:341:LYS:HD2	2.54	0.43
2:B:346:ARG:HB3	2:B:346:ARG:NH1	2.14	0.43
2:D:155:LEU:HD23	2:D:155:LEU:HA	1.83	0.43
2:D:151:ILE:HG12	2:D:151:ILE:O	2.19	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:191:ARG:NH1	2:B:194:HIS:CD2	2.85	0.43
2:D:304:ILE:O	2:D:305:ASN:HB3	2.19	0.43
2:B:156:LYS:HA	2:B:156:LYS:HE3	2.01	0.43
1:C:3:LEU:HA	1:C:7:LYS:HE3	1.99	0.43
1:A:35:GLY:N	5:A:806:PG4:H52	2.34	0.42
2:B:349:ARG:NE	2:B:349:ARG:HA	2.33	0.42
2:B:253:VAL:HG12	6:B:403:GDU:H2D	2.01	0.42
2:D:346:ARG:HA	2:D:346:ARG:HD2	1.63	0.42
2:D:351:LYS:HB2	2:D:351:LYS:HE2	1.74	0.42
2:D:233:ASN:HB3	2:D:378:LEU:HD22	2.01	0.42
2:D:393:LYS:HD2	9:D:1163:HOH:O	2.20	0.42
2:B:131:LEU:N	2:B:131:LEU:CD2	2.83	0.42
2:B:368:GLU:HG2	2:B:368:GLU:H	1.61	0.42
2:B:181:LYS:HD3	2:B:243:TYR:CE2	2.55	0.41
2:B:307:PHE:HB3	2:B:308:PRO:HD2	2.01	0.41
1:C:122:LYS:HA	1:C:123:PRO:HD3	1.56	0.41
2:D:336:ASN:ND2	2:D:338:VAL:HB	2.35	0.41
2:D:164:LYS:N	2:D:164:LYS:HD3	2.37	0.40
2:B:154:ASP:HB3	2:B:157:LEU:HB2	2.03	0.40
1:A:14:ASP:HA	9:A:1083:HOH:O	2.22	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	121/123 (98%)	119 (98%)	1 (1%)	1 (1%)	19	9
1	C	121/123 (98%)	116 (96%)	4 (3%)	1 (1%)	19	9
2	B	270/286 (94%)	264 (98%)	6 (2%)	0	100	100
2	D	270/286 (94%)	265 (98%)	5 (2%)	0	100	100
All	All	782/818 (96%)	764 (98%)	16 (2%)	2 (0%)	41	31

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	45	ASN
1	C	122	LYS

### 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	109/109 (100%)	105 (96%)	4 (4%)	34	25
1	C	109/109 (100%)	103 (94%)	6 (6%)	21	12
2	B	245/254 (96%)	230 (94%)	15 (6%)	18	9
2	D	245/254 (96%)	234 (96%)	11 (4%)	27	18
All	All	708/726 (98%)	672 (95%)	36 (5%)	24	14

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

Mol	Chain	Res	Type
1	A	7	LYS
1	A	23	LEU
1	A	85	LEU
1	A	88	ASP
2	B	131	LEU
2	B	141	LEU
2	B	146	LEU
2	B	156	LYS
2	B	157	LEU
2	B	164	LYS
2	B	167	LEU
2	B	195	LEU
2	B	240	LEU
2	B	296	LEU
2	B	302	LEU
2	B	336	ASN
2	B	346	ARG
2	B	349	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	352	LYS
1	C	71	ASN
1	C	85	LEU
1	C	88	ASP
1	C	120	CYS
1	C	121	GLU
1	C	122	LYS
2	D	141	LEU
2	D	151	ILE
2	D	153	VAL
2	D	157	LEU
2	D	167	LEU
2	D	195	LEU
2	D	240	LEU
2	D	287	VAL
2	D	302	LEU
2	D	336	ASN
2	D	349	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (22) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	39	GLN
1	A	45	ASN
1	A	71	ASN
2	B	160	GLN
2	B	161	GLN
2	B	194	HIS
2	B	210	GLN
2	B	305	ASN
2	B	310	ASN
2	B	336	ASN
2	B	365	HIS
1	C	10	HIS
1	C	32	HIS
1	C	71	ASN
2	D	160	GLN
2	D	161	GLN
2	D	194	HIS
2	D	207	GLN
2	D	305	ASN
2	D	310	ASN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
2	D	336	ASN
2	D	365	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 11 ligands modelled in this entry, 4 are monoatomic - leaving 7 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
6	GDU	D	528	7	35,38,38	1.60	5 (14%)	53,58,58	1.77	8 (15%)
8	UDP	B	405	-	24,26,26	1.18	1 (4%)	37,40,40	1.56	8 (21%)
5	PG4	C	807	-	12,12,12	0.50	0	11,11,11	0.33	0
4	MES	A	805	-	12,12,12	1.23	2 (16%)	14,16,16	1.20	3 (21%)
5	PG4	A	806	-	12,12,12	0.49	0	11,11,11	0.39	0
6	GDU	B	403	7	35,38,38	1.37	4 (11%)	53,58,58	1.72	6 (11%)
8	UDP	D	530	-	24,26,26	1.21	1 (4%)	37,40,40	1.52	6 (16%)

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
6	GDU	D	528	7	-	5/23/59/59	0/3/3/3
8	UDP	B	405	-	-	0/16/32/32	0/2/2/2
5	PG4	C	807	-	-	6/10/10/10	-
4	MES	A	805	-	-	2/6/14/14	0/1/1/1
5	PG4	A	806	-	-	7/10/10/10	-
6	GDU	B	403	7	-	5/23/59/59	0/3/3/3
8	UDP	D	530	-	-	5/16/32/32	0/2/2/2

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	528	GDU	PB-O3B	-5.51	1.46	1.60
6	D	528	GDU	O5'-C5'	4.25	1.54	1.44
6	B	403	GDU	PB-O3B	-3.86	1.50	1.60
8	B	405	UDP	PB-O1B	3.49	1.61	1.50
6	D	528	GDU	O5'-C1'	3.43	1.50	1.41
8	D	530	UDP	PB-O1B	3.20	1.60	1.50
6	B	403	GDU	O5'-C5'	2.71	1.50	1.44
6	B	403	GDU	O5'-C1'	2.69	1.48	1.41
4	A	805	MES	C5-N4	2.25	1.53	1.46
6	D	528	GDU	PB-O1B	2.04	1.58	1.50
6	D	528	GDU	PA-O2A	-2.02	1.45	1.55
6	B	403	GDU	C2-N1	2.01	1.41	1.38
4	A	805	MES	C3-N4	2.00	1.52	1.46

All (31) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	D	528	GDU	O5'-C1'-O3B	6.93	120.43	111.36
6	B	403	GDU	O3B-PB-O1B	6.15	132.55	109.47
6	D	528	GDU	O3B-PB-O1B	5.26	129.19	109.47
8	B	405	UDP	O5'-C5'-C4'	4.50	124.49	108.99
6	B	403	GDU	O5'-C1'-O3B	4.46	117.20	111.36
8	D	530	UDP	O3B-PB-O1B	4.04	126.50	110.68
6	B	403	GDU	O5'-C5'-C6'	-4.00	96.48	106.44
8	D	530	UDP	O5'-C5'-C4'	3.99	122.73	108.99
8	B	405	UDP	O3B-PB-O1B	3.83	125.69	110.68
6	B	403	GDU	O4D-C4D-C3D	-3.64	97.92	105.11
6	D	528	GDU	O5'-C5'-C6'	-3.45	97.85	106.44
6	D	528	GDU	O4D-C4D-C3D	-3.40	98.39	105.11

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	D	528	GDU	O3'-C3'-C4'	-2.90	103.65	110.35
6	B	403	GDU	O2-C2-N1	-2.86	118.99	122.79
8	D	530	UDP	O2-C2-N1	-2.79	119.07	122.79
8	B	405	UDP	O2-C2-N1	-2.66	119.25	122.79
6	D	528	GDU	O2-C2-N1	-2.64	119.27	122.79
8	D	530	UDP	C2'-C1'-N1	2.62	120.63	113.22
6	B	403	GDU	O3'-C3'-C4'	-2.59	104.37	110.35
4	A	805	MES	O2S-S-C8	-2.50	103.90	106.92
8	B	405	UDP	C2'-C1'-N1	2.45	120.17	113.22
8	B	405	UDP	O4'-C4'-C3'	-2.44	100.28	105.11
8	B	405	UDP	O4'-C4'-C5'	2.39	117.23	109.37
8	B	405	UDP	O2-C2-N3	2.33	125.84	121.50
8	D	530	UDP	O2-C2-N3	2.32	125.81	121.50
8	D	530	UDP	O3A-PB-O1B	-2.24	98.77	111.19
8	B	405	UDP	O3A-PB-O1B	-2.24	98.79	111.19
4	A	805	MES	O1S-S-C8	2.16	109.52	106.92
4	A	805	MES	C2-C3-N4	-2.16	106.83	110.10
6	D	528	GDU	O2-C2-N3	2.05	125.33	121.50
6	D	528	GDU	O5D-C5D-C4D	2.04	116.02	108.99

There are no chirality outliers.

All (30) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	805	MES	C8-C7-N4-C5
6	B	403	GDU	O5'-C1'-O3B-PB
6	D	528	GDU	O5'-C1'-O3B-PB
8	D	530	UDP	O4'-C4'-C5'-O5'
8	D	530	UDP	C5'-O5'-PA-O2A
8	D	530	UDP	C5'-O5'-PA-O3A
5	C	807	PG4	O2-C3-C4-O3
5	A	806	PG4	O1-C1-C2-O2
8	D	530	UDP	C3'-C4'-C5'-O5'
5	A	806	PG4	O3-C5-C6-O4
5	A	806	PG4	O2-C3-C4-O3
6	D	528	GDU	O5'-C5'-C6'-O6'
6	B	403	GDU	C1'-O3B-PB-O3A
6	B	403	GDU	C2'-C1'-O3B-PB
6	B	403	GDU	O5'-C5'-C6'-O6'
5	C	807	PG4	O1-C1-C2-O2
4	A	805	MES	C8-C7-N4-C3
8	D	530	UDP	PB-O3A-PA-O5'

*Continued on next page...*

*Continued from previous page...*

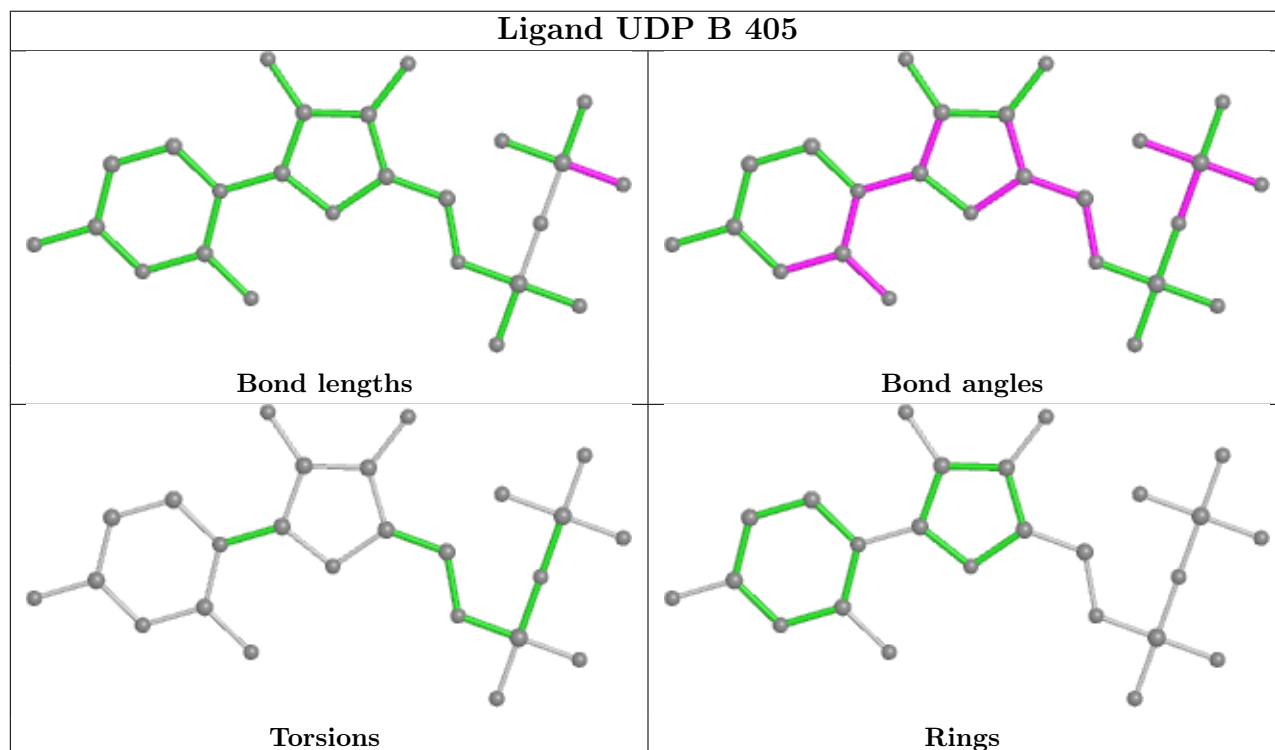
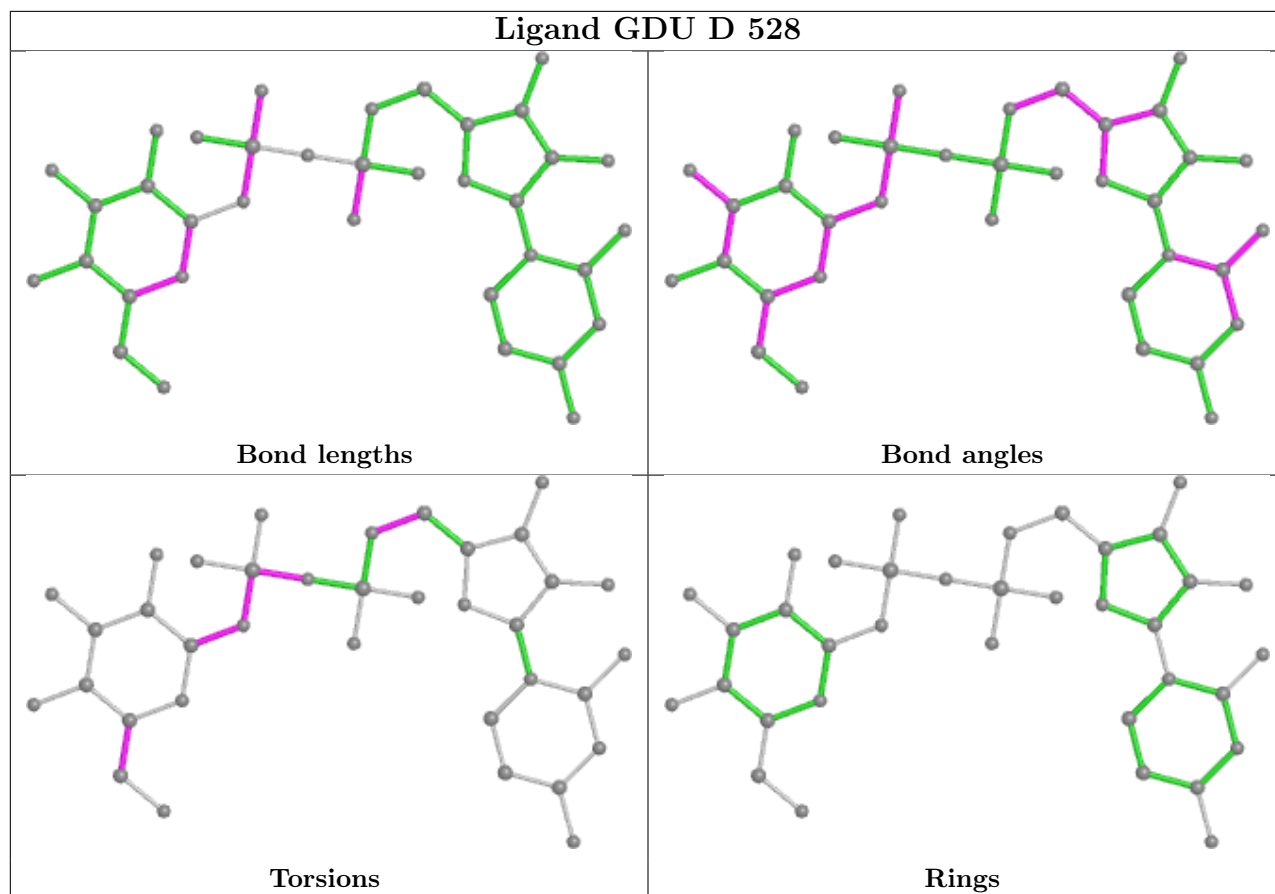
Mol	Chain	Res	Type	Atoms
5	C	807	PG4	C3-C4-O3-C5
5	A	806	PG4	C1-C2-O2-C3
5	A	806	PG4	C4-C3-O2-C2
5	A	806	PG4	C6-C5-O3-C4
5	A	806	PG4	C3-C4-O3-C5
5	C	807	PG4	C5-C6-O4-C7
6	B	403	GDU	PA-O3A-PB-O2B
6	D	528	GDU	C1'-O3B-PB-O3A
6	D	528	GDU	C4D-C5D-O5D-PA
5	C	807	PG4	C8-C7-O4-C6
5	C	807	PG4	C1-C2-O2-C3
6	D	528	GDU	PA-O3A-PB-O2B

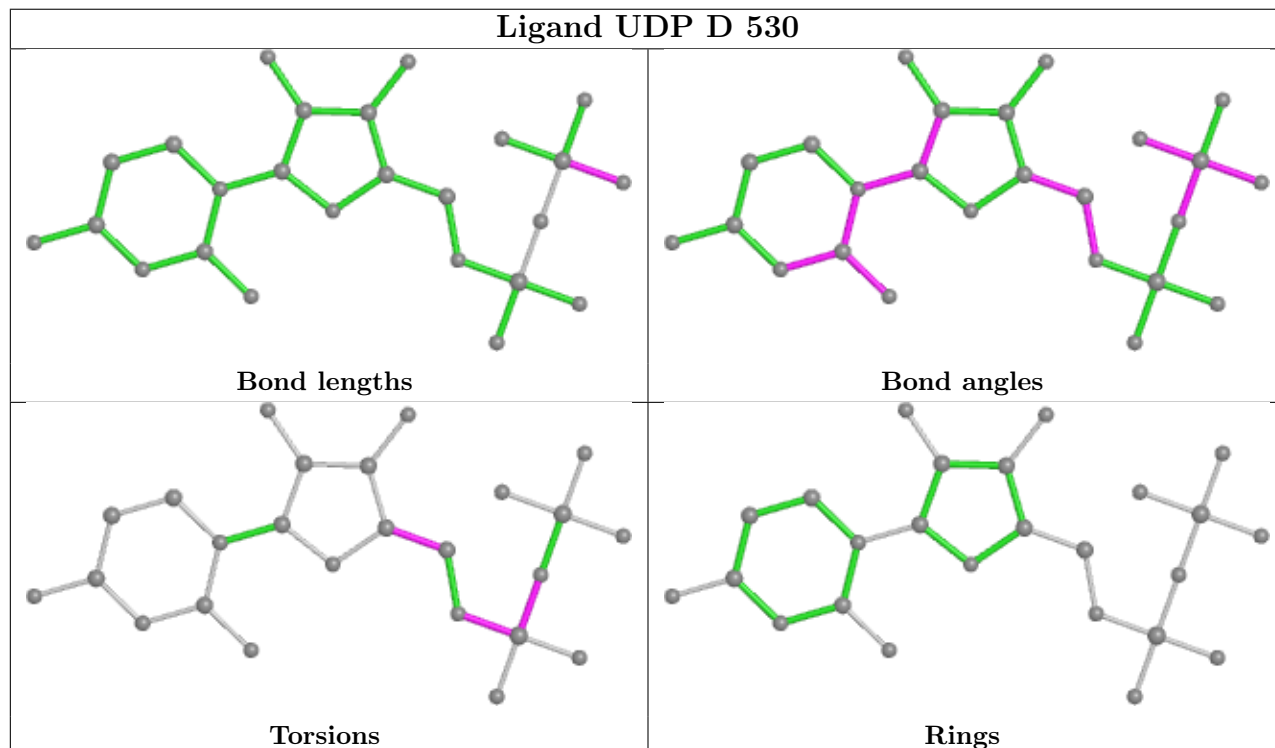
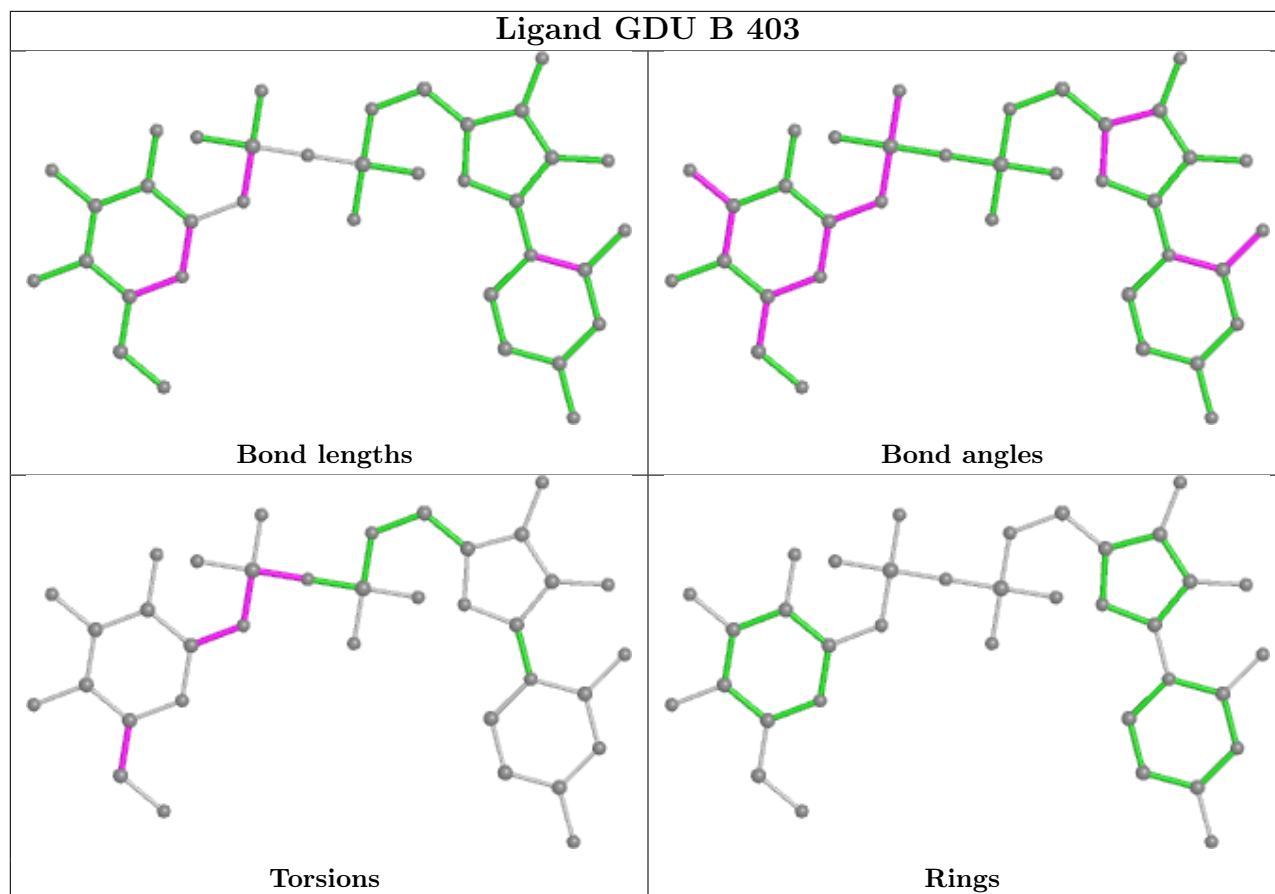
There are no ring outliers.

5 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	D	528	GDU	4	0
5	C	807	PG4	2	0
4	A	805	MES	2	0
5	A	806	PG4	1	0
6	B	403	GDU	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.





## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	123/123 (100%)	-0.09	6 (4%) 29 33	19, 25, 51, 62	0
1	C	123/123 (100%)	-0.00	6 (4%) 29 33	17, 25, 48, 82	0
2	B	272/286 (95%)	0.01	11 (4%) 38 41	18, 28, 49, 60	0
2	D	272/286 (95%)	-0.09	11 (4%) 38 41	18, 27, 44, 58	0
All	All	790/818 (96%)	-0.04	34 (4%) 35 38	17, 27, 48, 82	0

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	123	PRO	13.6
1	C	122	LYS	9.8
2	D	402	SER	8.2
1	A	45	ASN	7.2
1	C	67	PRO	5.6
1	A	123	PRO	5.6
2	B	402	SER	5.2
1	A	66	PHE	4.0
1	A	68	GLU	3.7
2	D	151	ILE	3.5
1	C	45	ASN	3.4
2	B	152	PRO	3.3
2	B	351	LYS	3.3
2	B	151	ILE	3.2
2	D	152	PRO	3.2
2	D	136	GLU	3.2
2	B	174	MET	3.1
1	A	65	GLU	2.9
2	D	132	THR	2.6
2	D	186	ILE	2.5
2	D	185	ILE	2.5

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	186	ILE	2.4
2	B	131	LEU	2.4
2	D	215	ILE	2.4
2	B	150	ASN	2.2
2	D	184	ILE	2.2
2	B	293	VAL	2.2
2	B	157	LEU	2.2
2	D	223	GLU	2.2
1	A	44	ASP	2.1
2	B	349	ARG	2.1
2	D	150	ASN	2.0
1	C	14	ASP	2.0
1	C	121	GLU	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, 95<sup>th</sup> 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.

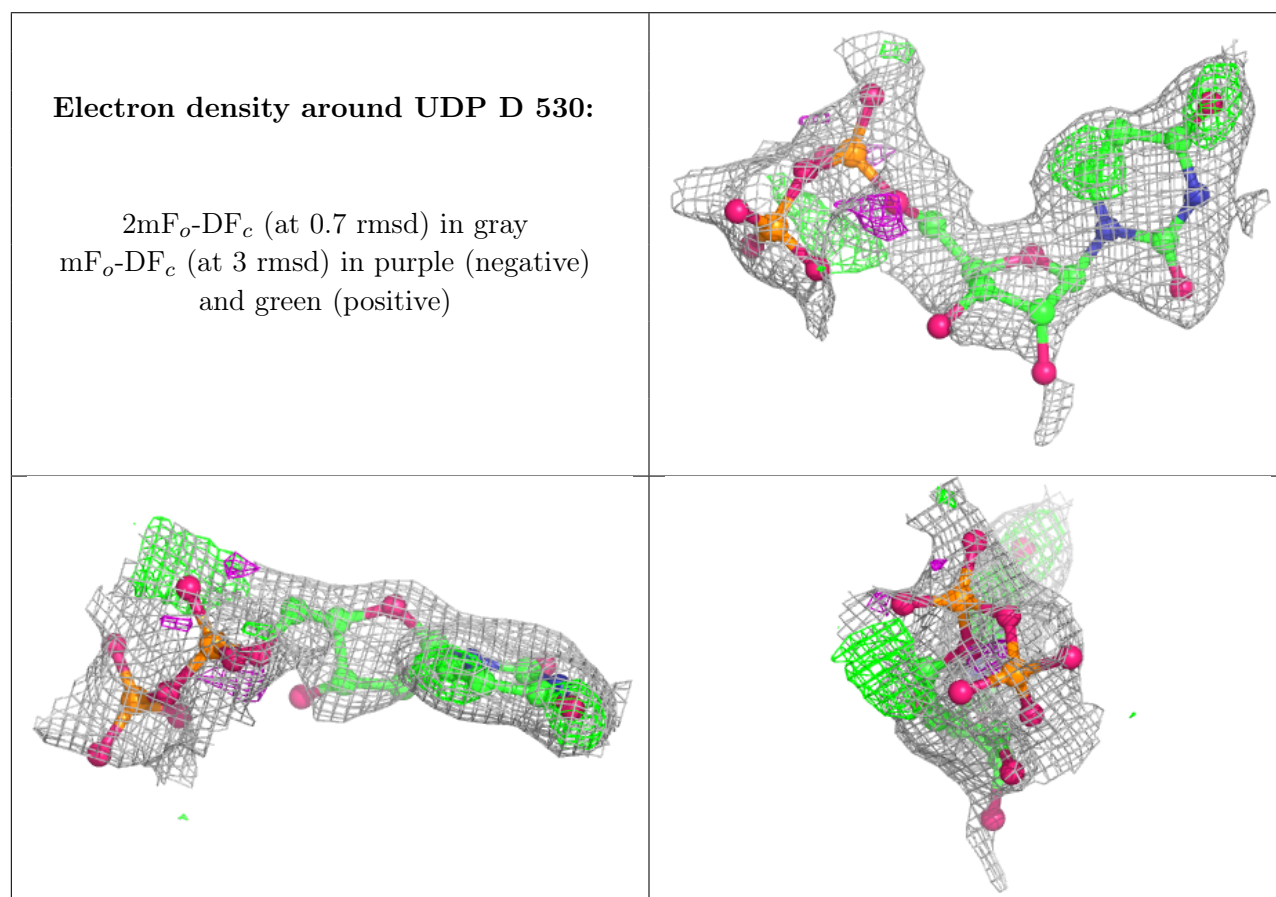
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
8	UDP	D	530	25/25	0.58	0.25	62,80,98,98	0
8	UDP	B	405	25/25	0.63	0.23	70,79,98,99	0
5	PG4	A	806	13/13	0.69	0.30	51,53,58,61	0
5	PG4	C	807	13/13	0.73	0.24	46,50,58,61	0
4	MES	A	805	12/12	0.87	0.16	42,46,50,52	0
6	GDU	B	403	36/36	0.96	0.10	20,25,29,31	0
6	GDU	D	528	36/36	0.96	0.10	19,24,33,35	0
3	CA	C	124	1/1	0.99	0.07	26,26,26,26	0
7	MN	D	529	1/1	1.00	0.07	21,21,21,21	0
3	CA	A	124	1/1	1.00	0.08	21,21,21,21	0

*Continued on next page...*

Continued from previous page...

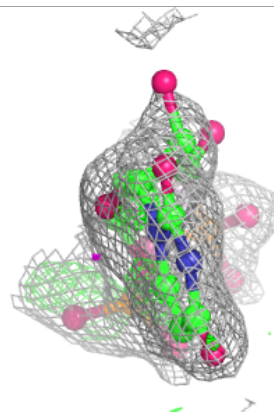
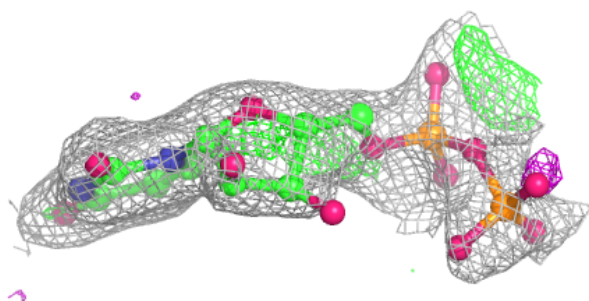
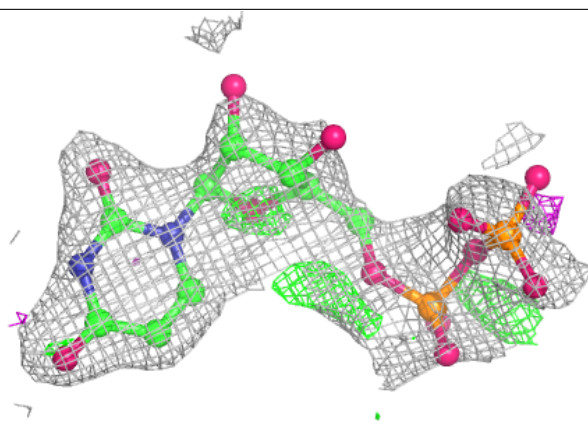
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
7	MN	B	404	1/1	1.00	0.04	25,25,25,25	0

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.

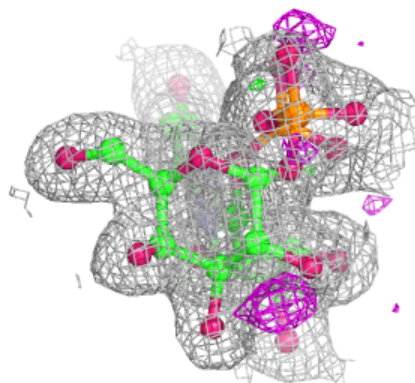
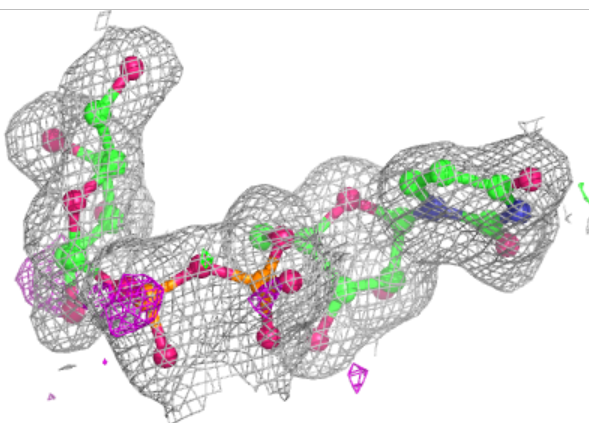
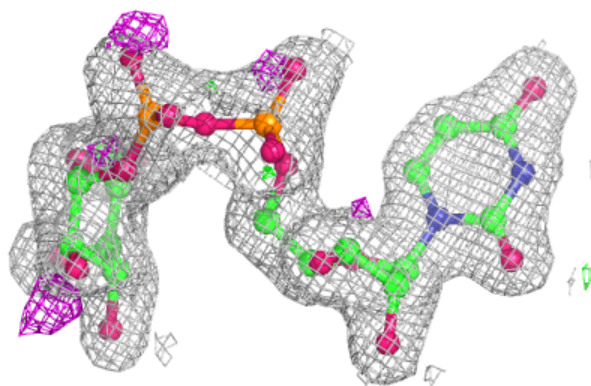


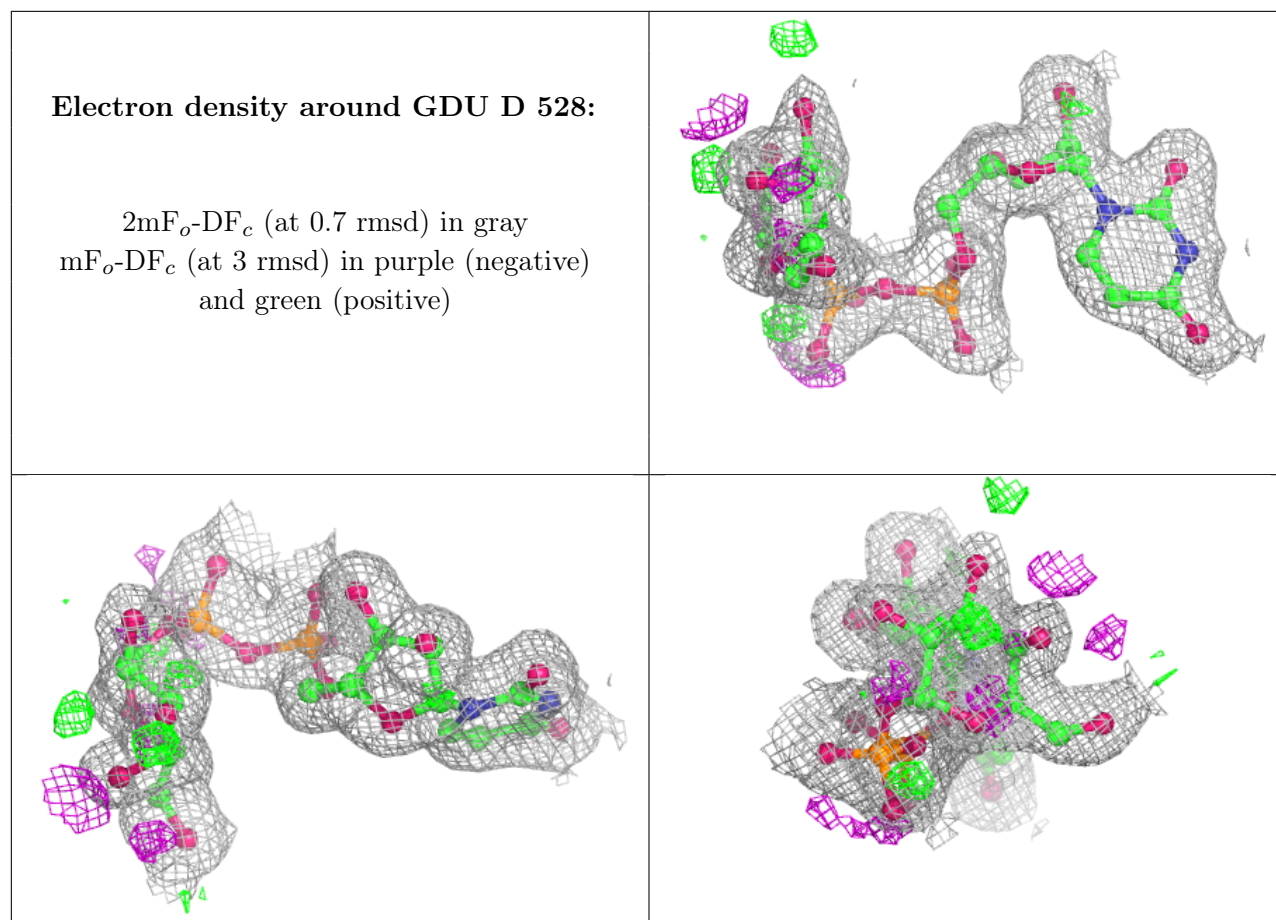
**Electron density around UDP B 405:**

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

**Electron density around GDU B 403:**

$2mF_o-DF_c$  (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.