wwPDB X-ray Structure Validation Summary Report (i)

Sep 7, 2023 – 01:44 pm BST

PDB ID	:	8BU1
Title	:	Structure of DDB1 bound to DS17-engaged CDK12-cyclin K
Authors	:	Kozicka, Z.; Kempf, G.; Petzold, G.; Thoma, N.H.
Deposited on	:	2022-11-30
Resolution	:	2.98 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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:

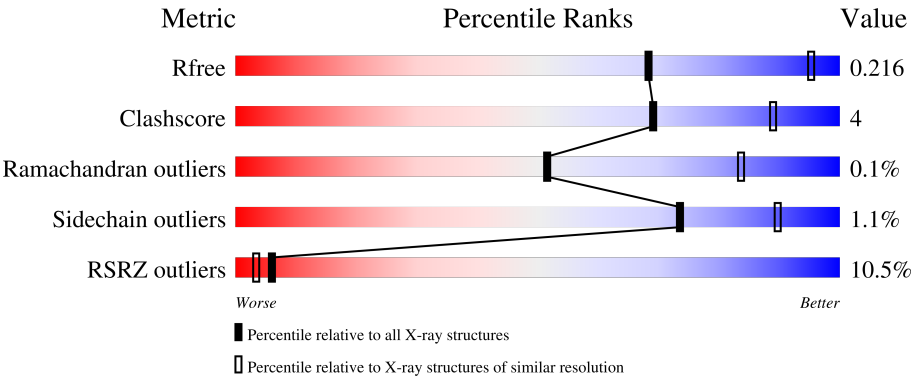
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.98 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.



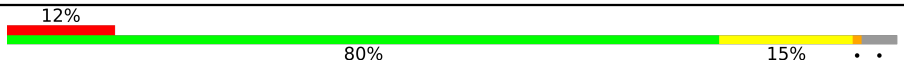
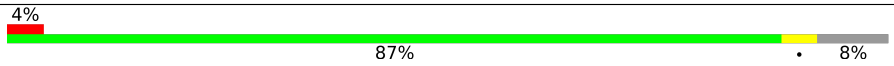
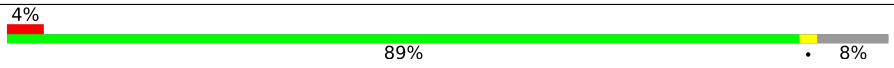
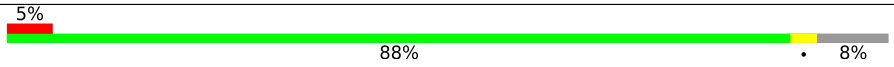
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2754 (3.00-2.96)
Clashscore	141614	3103 (3.00-2.96)
Ramachandran outliers	138981	2993 (3.00-2.96)
Sidechain outliers	138945	2996 (3.00-2.96)
RSRZ outliers	127900	2644 (3.00-2.96)

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	840	9% 91% 7% ·
1	D	840	11% 89% 9% ·
1	G	840	11% 88% 10% ·
2	B	344	18% 83% 12% · ·
2	E	344	12% 83% 12% · ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	344	
3	C	271	
3	F	271	
3	I	271	

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
4	GOL	C	302	-	-	-	X
4	GOL	H	1102	-	-	-	X
4	GOL	H	1103	-	-	-	X
4	GOL	I	302	-	-	-	X

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 67705 atoms, of which 33786 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 DNA damage-binding protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	826	Total	C	H	N	O	S	6449	0	0
			12936	4105	6449	1094	1252	36			
1	D	827	Total	C	H	N	O	S	6462	0	0
			12957	4111	6462	1095	1253	36			
1	G	826	Total	C	H	N	O	S	6452	0	0
			12938	4106	6452	1093	1251	36			

There are 39 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP Q16531
A	-2	GLY	-	expression tag	UNP Q16531
A	-1	GLY	-	expression tag	UNP Q16531
A	0	ARG	-	expression tag	UNP Q16531
A	700	GLY	-	linker	UNP Q16531
A	701	ASN	-	linker	UNP Q16531
A	702	GLY	-	linker	UNP Q16531
A	703	ASN	-	linker	UNP Q16531
A	704	SER	-	linker	UNP Q16531
A	705	GLY	-	linker	UNP Q16531
A	706	GLU	-	linker	UNP Q16531
A	707	ILE	-	linker	UNP Q16531
A	708	GLN	-	linker	UNP Q16531
D	-3	GLY	-	expression tag	UNP Q16531
D	-2	GLY	-	expression tag	UNP Q16531
D	-1	GLY	-	expression tag	UNP Q16531
D	0	ARG	-	expression tag	UNP Q16531
D	700	GLY	-	linker	UNP Q16531
D	701	ASN	-	linker	UNP Q16531
D	702	GLY	-	linker	UNP Q16531
D	703	ASN	-	linker	UNP Q16531
D	704	SER	-	linker	UNP Q16531
D	705	GLY	-	linker	UNP Q16531

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	706	GLU	-	linker	UNP Q16531
D	707	ILE	-	linker	UNP Q16531
D	708	GLN	-	linker	UNP Q16531
G	-3	GLY	-	expression tag	UNP Q16531
G	-2	GLY	-	expression tag	UNP Q16531
G	-1	GLY	-	expression tag	UNP Q16531
G	0	ARG	-	expression tag	UNP Q16531
G	700	GLY	-	linker	UNP Q16531
G	701	ASN	-	linker	UNP Q16531
G	702	GLY	-	linker	UNP Q16531
G	703	ASN	-	linker	UNP Q16531
G	704	SER	-	linker	UNP Q16531
G	705	GLY	-	linker	UNP Q16531
G	706	GLU	-	linker	UNP Q16531
G	707	ILE	-	linker	UNP Q16531
G	708	GLN	-	linker	UNP Q16531

- Molecule 2 is a protein called Cyclin-dependent kinase 12.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
2	B	330	Total	C	H	N	O	P	S	2702	0	0
			5395	1726	2702	455	494	1	17			
2	E	330	Total	C	H	N	O	P	S	2702	0	0
			5395	1726	2702	455	494	1	17			
2	H	330	Total	C	H	N	O	P	S	2701	0	0
			5394	1726	2701	455	494	1	17			

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	709	GLY	-	expression tag	UNP Q9NYV4
B	710	GLY	-	expression tag	UNP Q9NYV4
B	711	GLY	-	expression tag	UNP Q9NYV4
B	965	ARG	LYS	engineered mutation	UNP Q9NYV4
B	1052	GLN	-	expression tag	UNP Q9NYV4
E	709	GLY	-	expression tag	UNP Q9NYV4
E	710	GLY	-	expression tag	UNP Q9NYV4
E	711	GLY	-	expression tag	UNP Q9NYV4
E	965	ARG	LYS	engineered mutation	UNP Q9NYV4
E	1052	GLN	-	expression tag	UNP Q9NYV4
H	709	GLY	-	expression tag	UNP Q9NYV4
H	710	GLY	-	expression tag	UNP Q9NYV4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	711	GLY	-	expression tag	UNP Q9NYV4
H	965	ARG	LYS	engineered mutation	UNP Q9NYV4
H	1052	GLN	-	expression tag	UNP Q9NYV4

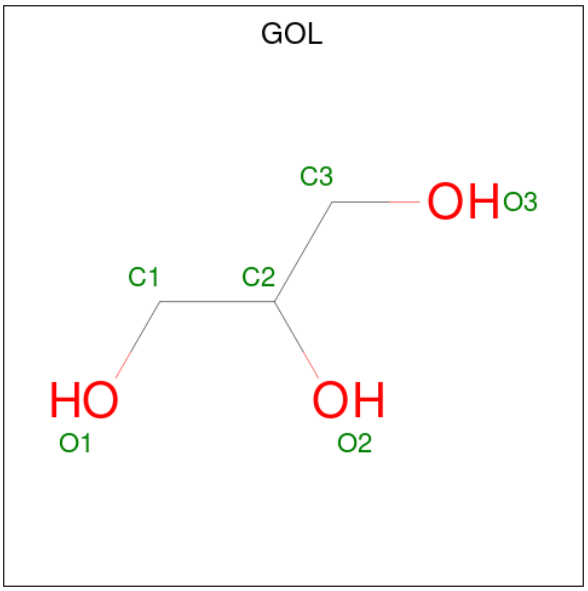
- Molecule 3 is a protein called Cyclin-K.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	248	Total	C	H	N	O	S	2048	0	0
			4111	1341	2048	346	363	13			
3	F	248	Total	C	H	N	O	S	2048	0	0
			4111	1341	2048	346	363	13			
3	I	248	Total	C	H	N	O	S	2048	0	0
			4111	1341	2048	346	363	13			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	GLY	-	expression tag	UNP O75909
C	-2	GLY	-	expression tag	UNP O75909
C	-1	GLY	-	expression tag	UNP O75909
C	0	ARG	-	expression tag	UNP O75909
F	-3	GLY	-	expression tag	UNP O75909
F	-2	GLY	-	expression tag	UNP O75909
F	-1	GLY	-	expression tag	UNP O75909
F	0	ARG	-	expression tag	UNP O75909
I	-3	GLY	-	expression tag	UNP O75909
I	-2	GLY	-	expression tag	UNP O75909
I	-1	GLY	-	expression tag	UNP O75909
I	0	ARG	-	expression tag	UNP O75909

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



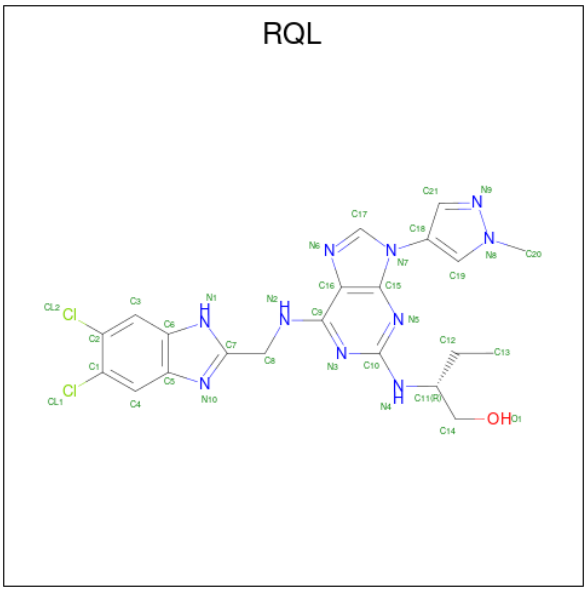
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	8	0
			14	3	8	3		
4	A	1	Total	C	H	O	8	0
			14	3	8	3		
4	A	1	Total	C	H	O	8	0
			14	3	8	3		
4	C	1	Total	C	H	O	8	0
			14	3	8	3		
4	C	1	Total	C	H	O	8	0
			14	3	8	3		
4	D	1	Total	C	H	O	8	0
			14	3	8	3		
4	D	1	Total	C	H	O	8	0
			14	3	8	3		
4	D	1	Total	C	H	O	8	0
			14	3	8	3		
4	F	1	Total	C	H	O	8	0
			14	3	8	3		
4	F	1	Total	C	H	O	8	0
			14	3	8	3		
4	G	1	Total	C	H	O	8	0
			14	3	8	3		
4	H	1	Total	C	H	O	8	0
			14	3	8	3		
4	H	1	Total	C	H	O	8	0
			14	3	8	3		
4	I	1	Total	C	H	O	8	0
			14	3	8	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	I	1	Total	C	H	O	8	0
			14	3	8	3		

- Molecule 5 is (2 {R})-2-[[6-[[5,6-bis(chloranyl)-1 {H}-benzimidazol-2-yl]methylamino] -9-(1-methylpyrazol-4-yl)purin-2-yl]amino]butan-1-ol (three-letter code: RQL) (formula: $C_{21}H_{22}Cl_2N_{10}O$) (labeled as "Ligand of Interest" by depositor).



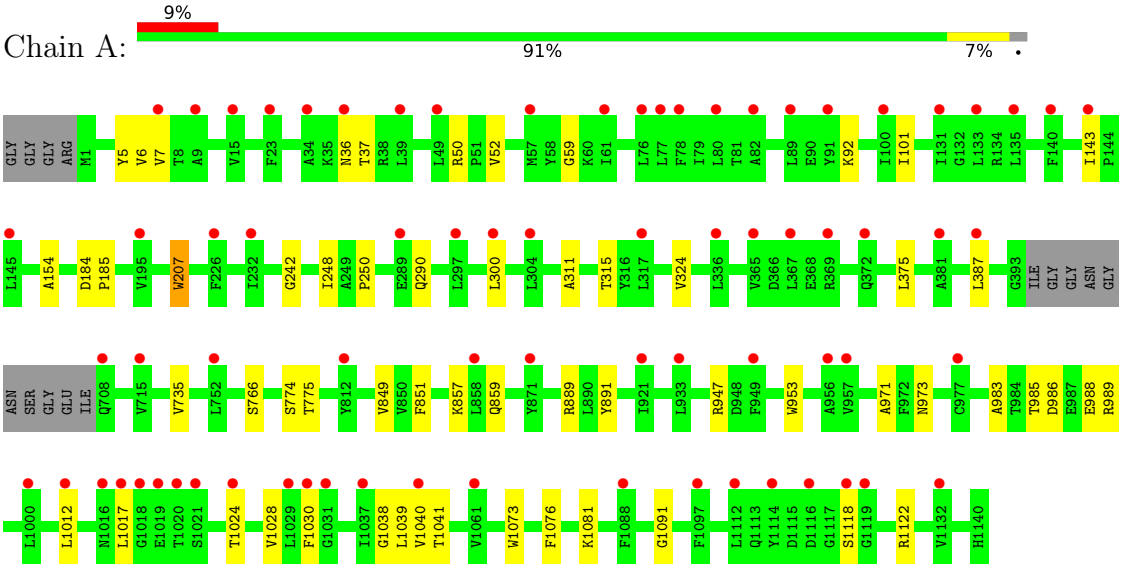
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
5	B	1	Total	C	Cl	H	N	O	18	0
			49	20	2	18	8	1		
5	E	1	Total	C	Cl	H	N	O	18	0
			49	20	2	18	8	1		
5	H	1	Total	C	Cl	H	N	O	18	0
			49	20	2	18	8	1		

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.

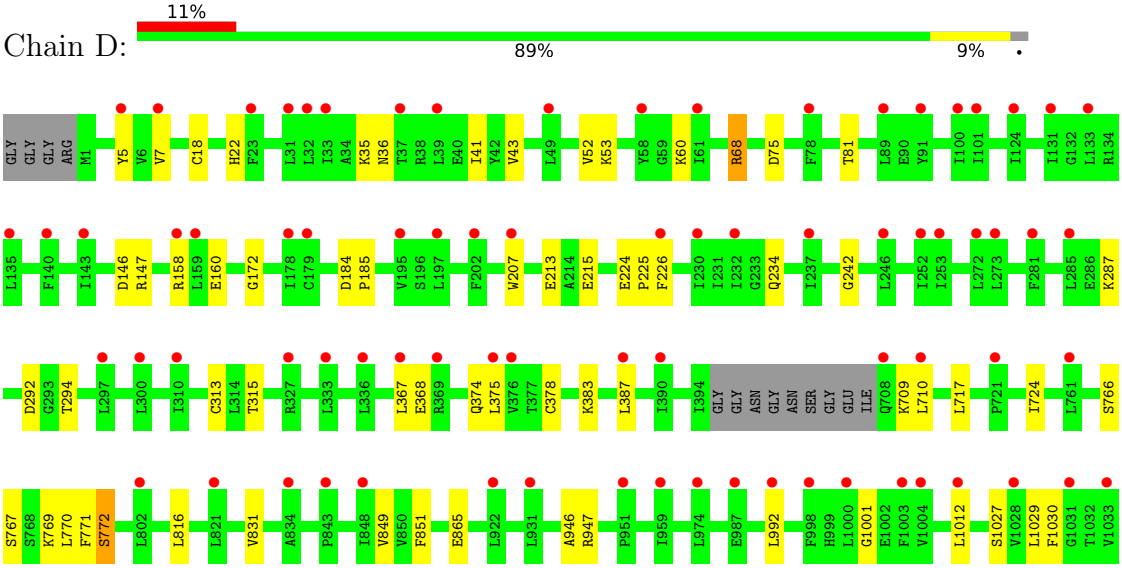
• Molecule 1: DNA damage-binding protein 1

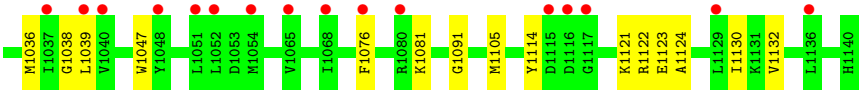
Chain A:



• Molecule 1: DNA damage-binding protein 1

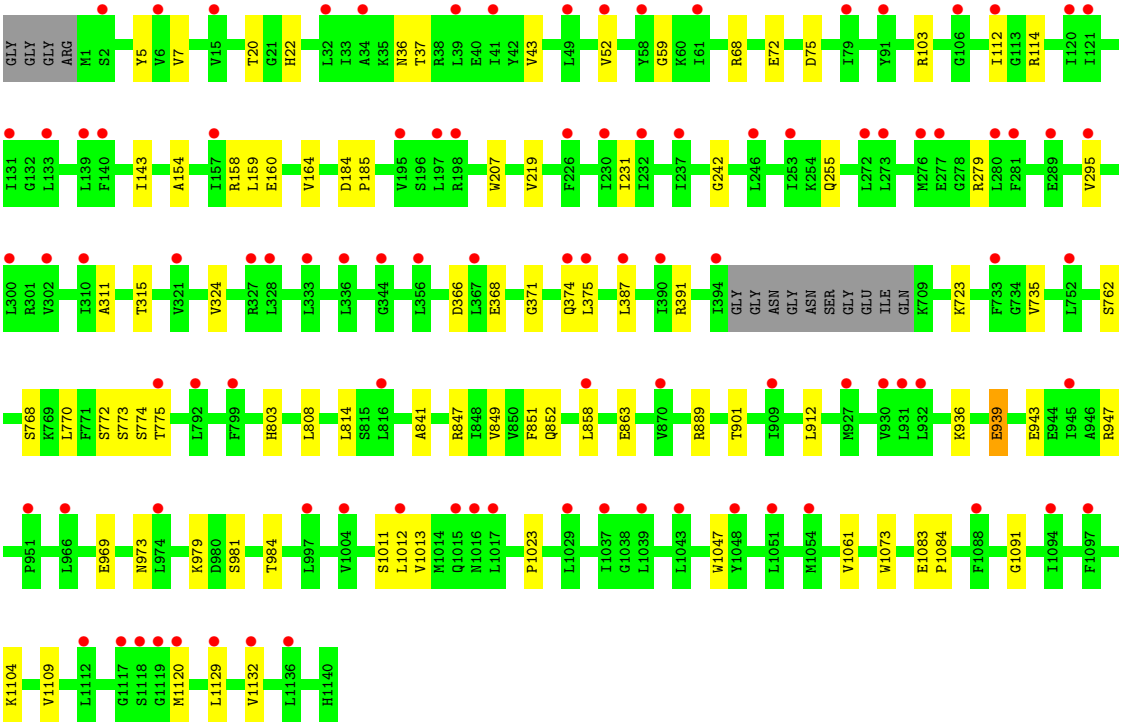
Chain D:





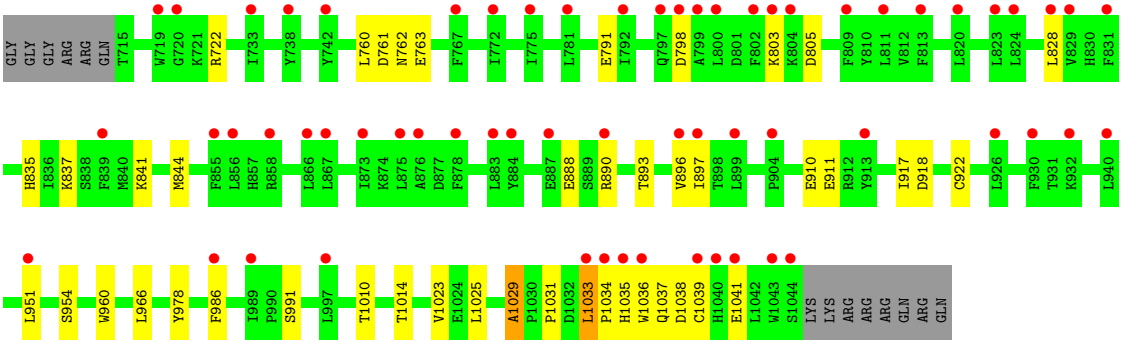
• Molecule 1: DNA damage-binding protein 1

Chain G:



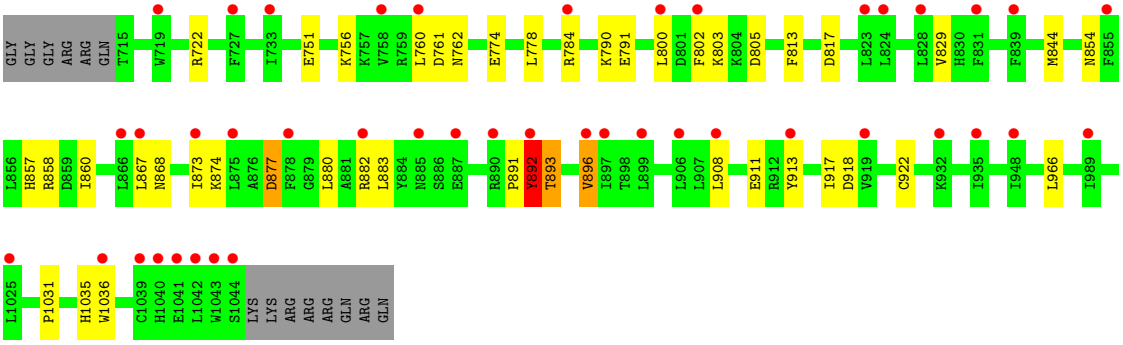
• Molecule 2: Cyclin-dependent kinase 12

Chain B:

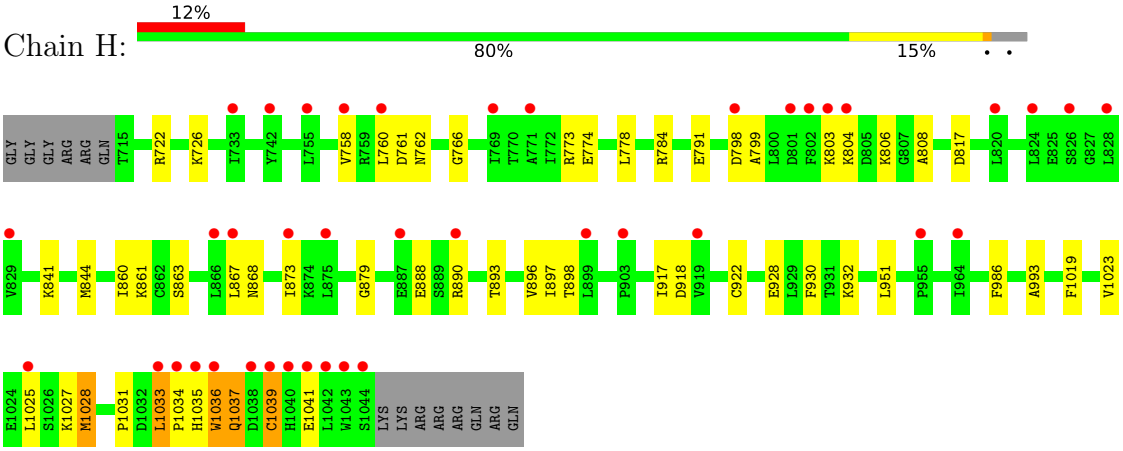


• Molecule 2: Cyclin-dependent kinase 12

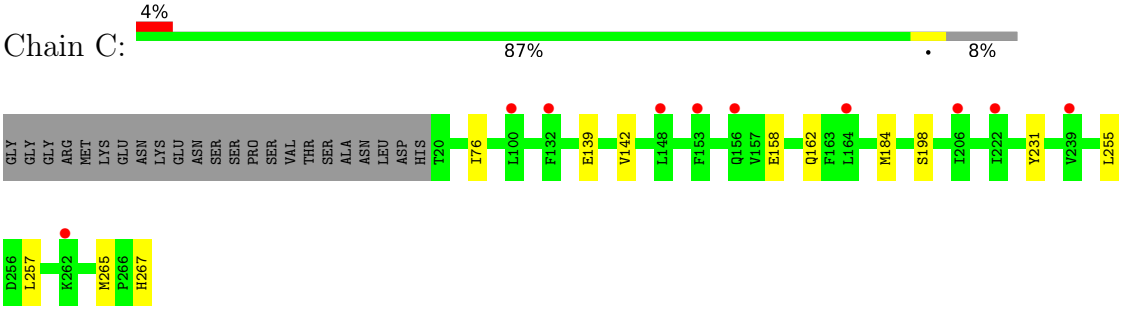
Chain E:



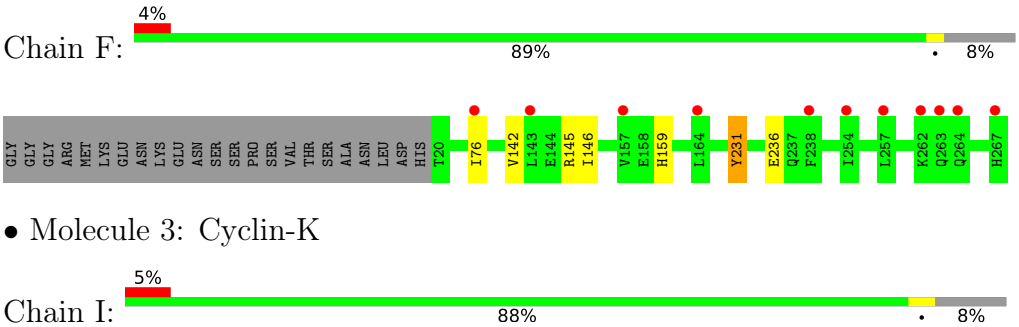
• Molecule 2: Cyclin-dependent kinase 12



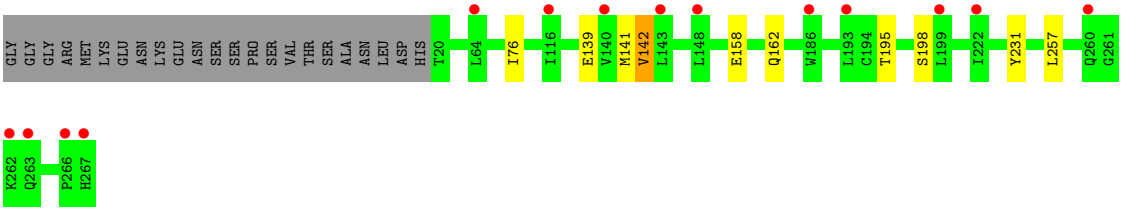
• Molecule 3: Cyclin-K



• Molecule 3: Cyclin-K



• Molecule 3: Cyclin-K


GLY
GLY
GLY
ARG
MET
LYS
GLU
ASN
LYS
GLU
ASN
SER
SER
PRO
SER
VAL
THR
SER
ALA
ASN
LEU
ASP
HIS
T20
L64
I76
I116
E139
V140
M141
V142
L143
L148
E158
Q162
W186
L193
C194
T195
S198
L199
I222
Y231
L257
Q260
G261
K262
Q263
P266
H267

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	249.45Å 249.45Å 218.45Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	60.12 – 2.98 216.03 – 2.98	Depositor EDS
% Data completeness (in resolution range)	87.6 (60.12-2.98) 87.7 (216.03-2.98)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.58 (at 2.96Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.193 , 0.218 0.192 , 0.216	Depositor DCC
R_{free} test set	6883 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	102.7	Xtriage
Anisotropy	0.005	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 63.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.009 for -h,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	67705	wwPDB-VP
Average B, all atoms (Å^2)	113.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.68% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, TPO, RQL

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	0.28	0/6604	0.55	0/8931
1	D	0.29	0/6612	0.55	1/8942 (0.0%)
1	G	0.29	0/6603	0.55	0/8930
2	B	0.28	0/2743	0.54	0/3699
2	E	0.51	5/2743 (0.2%)	0.60	2/3699 (0.1%)
2	H	0.29	0/2743	0.53	0/3699
3	C	0.27	0/2120	0.48	0/2868
3	F	0.28	0/2120	0.49	0/2868
3	I	0.28	0/2120	0.49	0/2868
All	All	0.31	5/34408 (0.0%)	0.54	3/46504 (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
2	E	0	1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	892	TYR	CE2-CZ	-14.09	1.20	1.38
2	E	892	TYR	CG-CD2	9.38	1.51	1.39
2	E	892	TYR	CA-CB	7.41	1.70	1.53
2	E	892	TYR	CB-CG	7.23	1.62	1.51
2	E	892	TYR	CG-CD1	-5.17	1.32	1.39

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	892	TYR	CB-CG-CD2	9.71	126.82	121.00
2	E	892	TYR	CD1-CE1-CZ	8.83	127.74	119.80
1	D	68	ARG	NE-CZ-NH1	5.14	122.87	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	E	892	TYR	Sidechain

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	6487	6449	6451	35	0
1	D	6495	6462	6464	56	0
1	G	6486	6452	6454	57	0
2	B	2693	2702	2701	27	0
2	E	2693	2702	2702	36	0
2	H	2693	2701	2701	36	0
3	C	2063	2048	2048	8	0
3	F	2063	2048	2048	6	0
3	I	2063	2048	2048	8	0
4	A	18	24	24	0	0
4	C	12	16	16	0	0
4	D	18	24	24	0	0
4	F	12	16	16	0	0
4	G	6	8	8	0	0
4	H	12	16	16	1	0
4	I	12	16	16	0	0
5	B	31	18	0	0	0
5	E	31	18	0	0	0
5	H	31	18	0	0	0
All	All	33919	33786	33737	253	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 4.

The worst 5 of 253 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:986:ASP:OD1	1:A:989:ARG:NH2	1.91	1.02
1:G:723:LYS:NZ	1:G:814:LEU:O	1.93	1.02
2:E:882:ARG:HH12	2:E:892:TYR:HA	1.23	1.00
3:F:231:TYR:OH	3:F:236:GLU:OE1	1.88	0.92
2:E:858:ARG:NH1	2:E:882:ARG:HG3	1.91	0.85

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	822/840 (98%)	799 (97%)	23 (3%)	0	100	100
1	D	823/840 (98%)	806 (98%)	16 (2%)	1 (0%)	51	83
1	G	822/840 (98%)	799 (97%)	21 (3%)	2 (0%)	47	80
2	B	327/344 (95%)	315 (96%)	11 (3%)	1 (0%)	41	74
2	E	327/344 (95%)	318 (97%)	8 (2%)	1 (0%)	41	74
2	H	327/344 (95%)	317 (97%)	9 (3%)	1 (0%)	41	74
3	C	246/271 (91%)	242 (98%)	4 (2%)	0	100	100
3	F	246/271 (91%)	242 (98%)	4 (2%)	0	100	100
3	I	246/271 (91%)	242 (98%)	4 (2%)	0	100	100
All	All	4186/4365 (96%)	4080 (98%)	100 (2%)	6 (0%)	51	83

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	772	SER
2	E	877	ASP
1	G	774	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	G	772	SER
2	H	1028	MET

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	721/728 (99%)	717 (99%)	4 (1%)	86	94
1	D	722/728 (99%)	718 (99%)	4 (1%)	86	94
1	G	721/728 (99%)	717 (99%)	4 (1%)	86	94
2	B	297/308 (96%)	289 (97%)	8 (3%)	44	75
2	E	297/308 (96%)	292 (98%)	5 (2%)	60	84
2	H	297/308 (96%)	285 (96%)	12 (4%)	31	66
3	C	223/242 (92%)	221 (99%)	2 (1%)	78	91
3	F	223/242 (92%)	222 (100%)	1 (0%)	91	97
3	I	223/242 (92%)	221 (99%)	2 (1%)	78	91
All	All	3724/3834 (97%)	3682 (99%)	42 (1%)	73	90

5 of 42 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	860	ILE
2	H	1035	HIS
2	H	863	SER
2	H	898	THR
2	H	1037	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	189	HIS
1	A	859	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

3 non-standard protein/DNA/RNA residues are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	TPO	E	893	2	8,10,11	1.63	1 (12%)	10,14,16	1.32	2 (20%)
2	TPO	H	893	2	8,10,11	1.79	2 (25%)	10,14,16	1.00	0
2	TPO	B	893	2	8,10,11	1.59	1 (12%)	10,14,16	1.53	1 (10%)

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
2	TPO	E	893	2	-	4/9/11/13	-
2	TPO	H	893	2	-	4/9/11/13	-
2	TPO	B	893	2	-	1/9/11/13	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	893	TPO	P-O1P	3.40	1.61	1.50
2	H	893	TPO	P-O1P	3.38	1.61	1.50
2	B	893	TPO	P-O1P	3.36	1.61	1.50
2	H	893	TPO	CB-CA	2.57	1.59	1.53

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	893	TPO	P-OG1-CB	-4.19	110.56	123.21
2	E	893	TPO	P-OG1-CB	-2.71	115.04	123.21
2	E	893	TPO	CG2-CB-CA	-2.18	108.86	113.16

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	E	893	TPO	N-CA-CB-CG2
2	E	893	TPO	N-CA-CB-OG1
2	E	893	TPO	C-CA-CB-CG2
2	E	893	TPO	CG2-CB-OG1-P
2	H	893	TPO	N-CA-CB-CG2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	893	TPO	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

18 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	GOL	A	1201	-	5,5,5	0.75	0	5,5,5	0.99	0
4	GOL	G	1201	-	5,5,5	0.81	0	5,5,5	1.08	0
4	GOL	H	1103	-	5,5,5	0.77	0	5,5,5	1.04	0
5	RQL	E	1101	-	28,34,38	0.84	0	27,49,55	1.61	4 (14%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	GOL	D	1201	-	5,5,5	0.70	0	5,5,5	0.91	0
4	GOL	A	1202	-	5,5,5	0.76	0	5,5,5	1.00	0
5	RQL	H	1101	-	28,34,38	0.86	0	27,49,55	1.32	4 (14%)
4	GOL	A	1203	-	5,5,5	0.75	0	5,5,5	1.09	0
4	GOL	D	1202	-	5,5,5	0.76	0	5,5,5	1.07	0
4	GOL	F	302	-	5,5,5	0.72	0	5,5,5	0.89	0
4	GOL	C	301	-	5,5,5	0.90	0	5,5,5	1.10	0
4	GOL	H	1102	-	5,5,5	0.51	0	5,5,5	0.58	0
4	GOL	F	301	-	5,5,5	0.73	0	5,5,5	0.95	0
4	GOL	I	302	-	5,5,5	0.78	0	5,5,5	1.02	0
4	GOL	C	302	-	5,5,5	0.81	0	5,5,5	1.11	0
5	RQL	B	1101	-	28,34,38	0.85	0	27,49,55	1.36	4 (14%)
4	GOL	D	1203	-	5,5,5	0.80	0	5,5,5	1.06	0
4	GOL	I	301	-	5,5,5	0.68	0	5,5,5	0.90	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	GOL	A	1201	-	-	0/4/4/4	-
4	GOL	G	1201	-	-	0/4/4/4	-
4	GOL	H	1103	-	-	2/4/4/4	-
5	RQL	E	1101	-	-	4/11/17/17	0/4/4/5
4	GOL	D	1201	-	-	0/4/4/4	-
4	GOL	A	1202	-	-	1/4/4/4	-
5	RQL	H	1101	-	-	3/11/17/17	0/4/4/5
4	GOL	A	1203	-	-	0/4/4/4	-
4	GOL	D	1202	-	-	0/4/4/4	-
4	GOL	F	302	-	-	0/4/4/4	-
4	GOL	C	301	-	-	0/4/4/4	-
4	GOL	H	1102	-	-	2/4/4/4	-
4	GOL	F	301	-	-	0/4/4/4	-
4	GOL	I	302	-	-	1/4/4/4	-
4	GOL	C	302	-	-	2/4/4/4	-
5	RQL	B	1101	-	-	4/11/17/17	0/4/4/5
4	GOL	D	1203	-	-	0/4/4/4	-
4	GOL	I	301	-	-	0/4/4/4	-

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	E	1101	RQL	C12-C11-C14	5.64	120.81	111.72
5	B	1101	RQL	C12-C11-C14	3.57	117.47	111.72
5	E	1101	RQL	C16-C9-N3	-2.82	118.46	120.81
5	B	1101	RQL	C16-C9-N3	-2.78	118.50	120.81
5	H	1101	RQL	C2-C3-C6	-2.61	117.16	119.70

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	302	GOL	O1-C1-C2-O2
4	C	302	GOL	O1-C1-C2-C3
5	B	1101	RQL	N3-C10-N4-C11
5	B	1101	RQL	N5-C10-N4-C11
5	B	1101	RQL	C14-C11-C12-C13

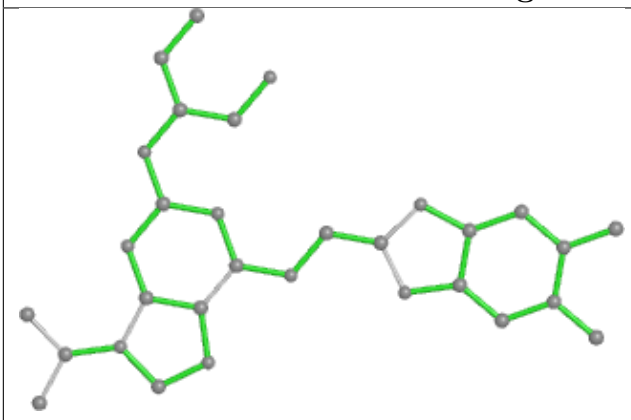
There are no ring outliers.

1 monomer is involved in 1 short contact:

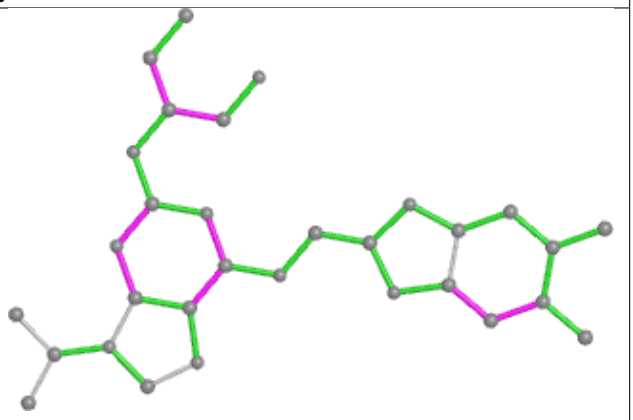
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	H	1102	GOL	1	0

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.

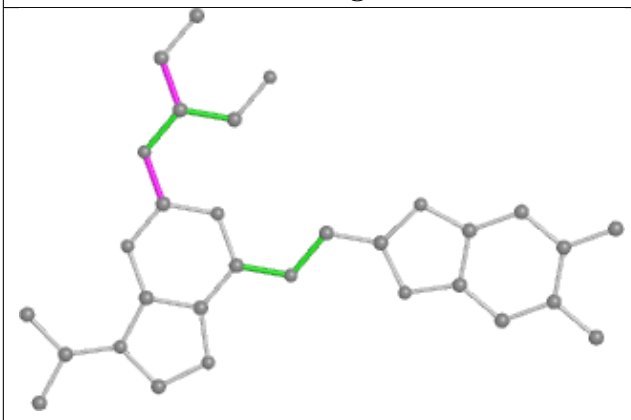
Ligand RQL E 1101



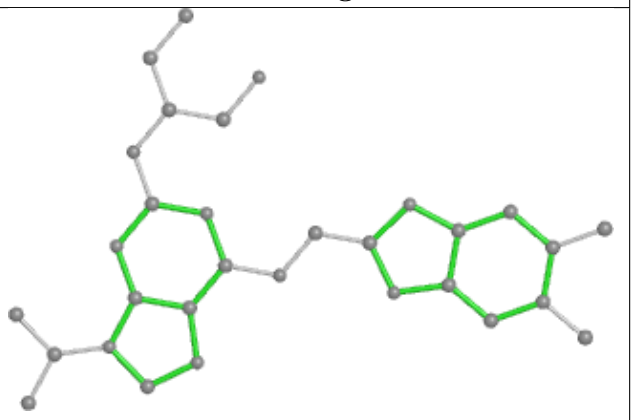
Bond lengths



Bond angles

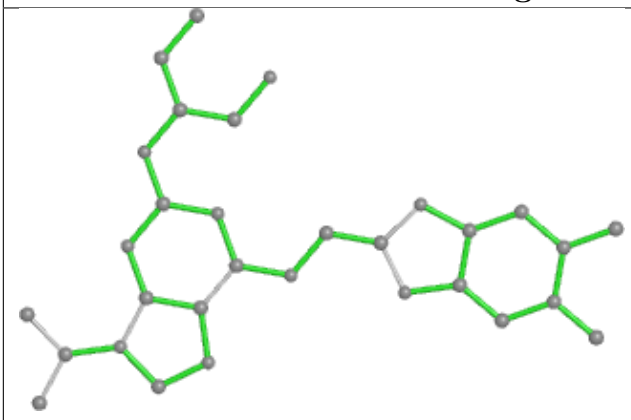


Torsions

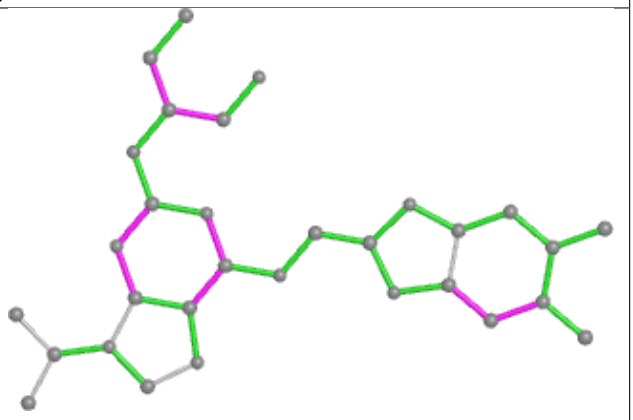


Rings

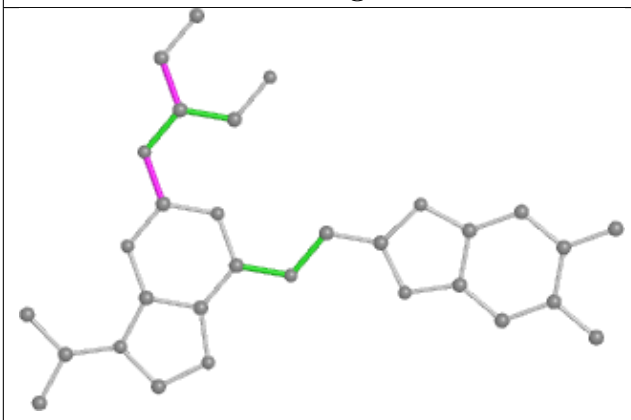
Ligand RQL H 1101



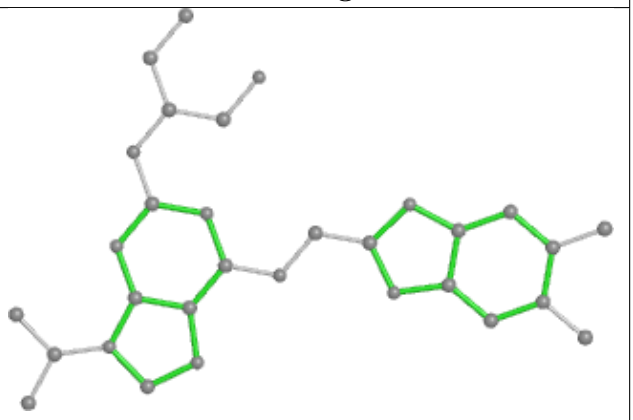
Bond lengths



Bond angles

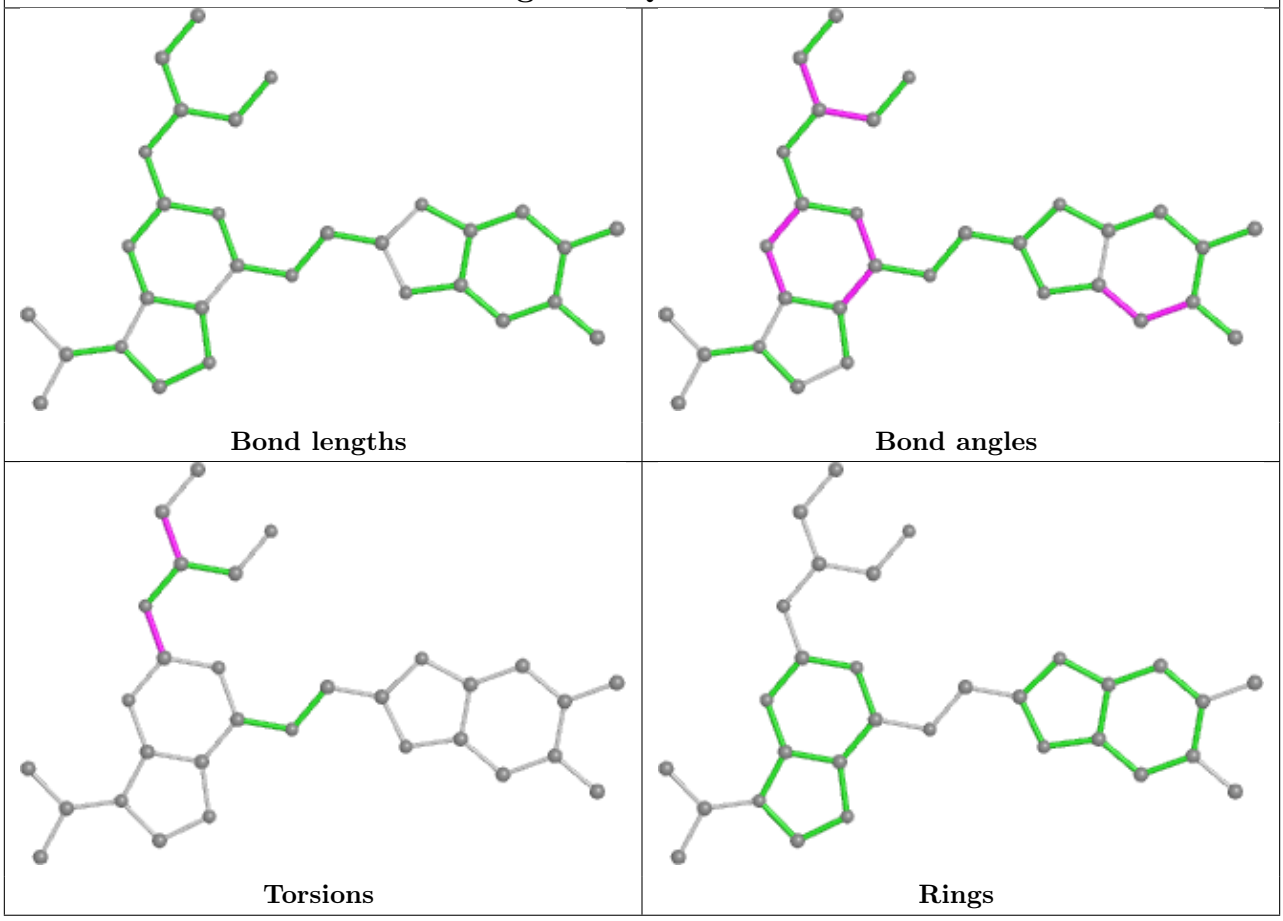


Torsions



Rings

Ligand RQL B 1101



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

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} 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	826/840 (98%)	0.86	74 (8%) 9 5	69, 99, 175, 319	0
1	D	827/840 (98%)	0.82	93 (11%) 5 3	68, 102, 169, 260	0
1	G	826/840 (98%)	0.83	96 (11%) 4 2	70, 102, 171, 233	0
2	B	329/344 (95%)	1.18	62 (18%) 1 0	77, 101, 169, 253	0
2	E	329/344 (95%)	1.08	43 (13%) 3 1	75, 99, 173, 204	0
2	H	329/344 (95%)	1.09	40 (12%) 4 2	61, 81, 165, 205	0
3	C	248/271 (91%)	0.77	10 (4%) 38 23	74, 91, 130, 185	0
3	F	248/271 (91%)	0.85	11 (4%) 34 20	62, 79, 116, 186	0
3	I	248/271 (91%)	0.86	14 (5%) 24 13	64, 82, 122, 189	0
All	All	4210/4365 (96%)	0.90	443 (10%) 6 3	61, 96, 168, 319	0

The worst 5 of 443 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	1043	TRP	11.8
2	H	1043	TRP	10.0
2	B	1044	SER	9.6
2	B	1043	TRP	9.3
2	B	800	LEU	9.3

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

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^{th} 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	TPO	E	893	11/12	0.73	0.26	153,172,207,215	6
2	TPO	H	893	11/12	0.84	0.26	99,122,147,149	5
2	TPO	B	893	11/12	0.86	0.20	148,167,200,202	6

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

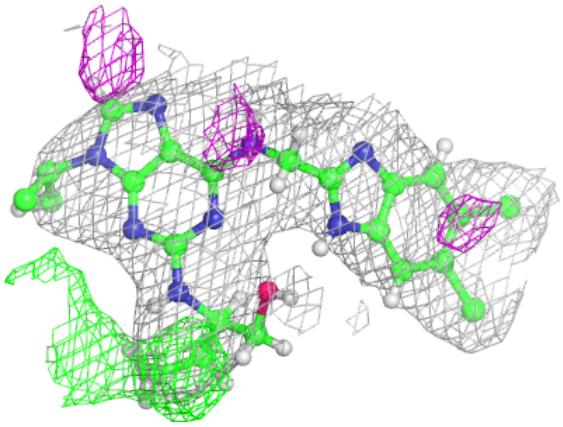
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

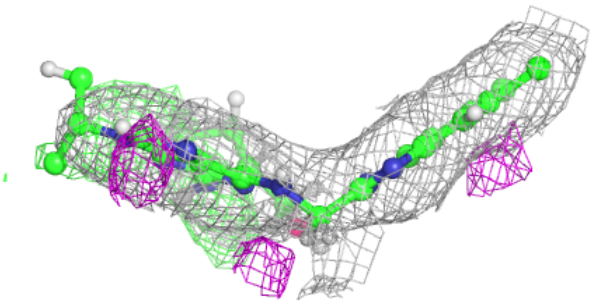
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	GOL	H	1102	6/6	0.35	0.53	90,113,127,136	8
4	GOL	I	302	6/6	0.55	0.48	96,116,124,126	8
4	GOL	H	1103	6/6	0.70	0.54	89,108,136,136	8
4	GOL	A	1201	6/6	0.72	0.21	108,130,136,140	8
4	GOL	G	1201	6/6	0.73	0.27	99,119,135,136	8
4	GOL	C	302	6/6	0.76	0.41	87,105,122,122	8
4	GOL	D	1201	6/6	0.76	0.23	106,128,143,149	8
4	GOL	C	301	6/6	0.84	0.22	96,116,130,130	8
4	GOL	F	302	6/6	0.84	0.55	101,121,146,146	8
4	GOL	D	1202	6/6	0.86	0.59	91,109,119,130	8
4	GOL	I	301	6/6	0.88	0.16	85,103,117,122	8
4	GOL	A	1202	6/6	0.88	0.48	88,111,117,135	8
4	GOL	F	301	6/6	0.90	0.16	92,111,121,124	8
4	GOL	A	1203	6/6	0.91	0.25	82,99,116,121	8
4	GOL	D	1203	6/6	0.93	0.21	89,107,116,120	8
5	RQL	E	1101	31/34	0.93	0.32	79,91,114,119	18
5	RQL	H	1101	31/34	0.94	0.30	69,83,103,104	18
5	RQL	B	1101	31/34	0.95	0.27	72,87,105,114	18

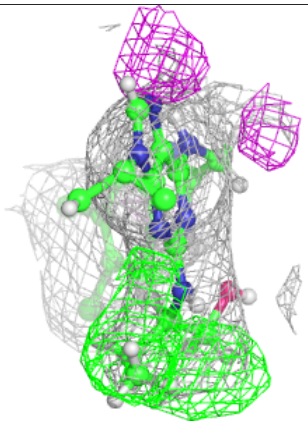
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around RQL E 1101:

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

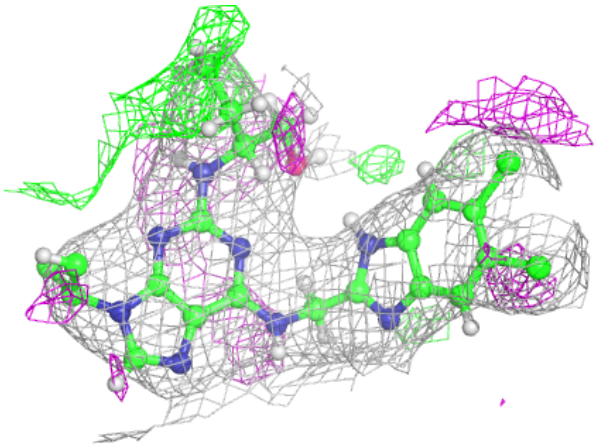


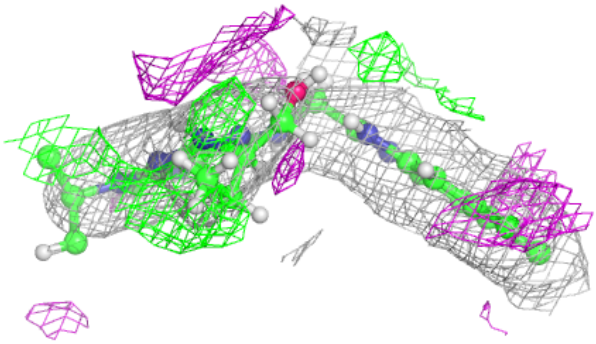


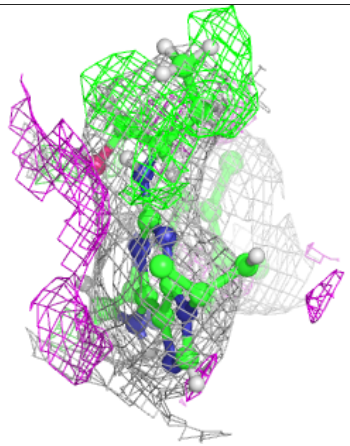


Electron density around RQL H 1101:

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

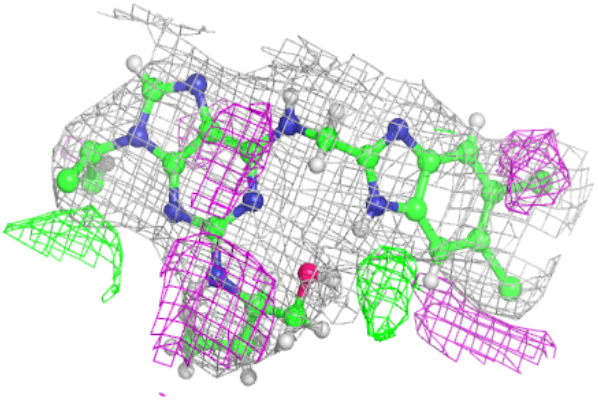


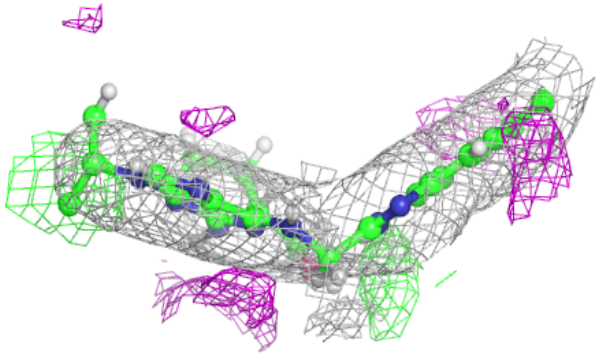


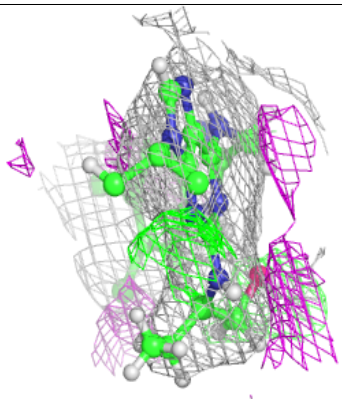


Electron density around RQL B 1101:

$2mF_o$-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.