wwPDB X-ray Structure Validation Summary Report (i)

Sep 7, 2023 – 10:25 pm BST

PDB ID	:	8BUB
Title	:	Structure of DDB1 bound to dCeMM4-engaged CDK12-cyclin K
Authors	:	Kozicka, Z.; Kempf, G.; Petzold, G.; Thoma, N.H.
Deposited on	:	2022-11-30
Resolution	:	3.42 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:

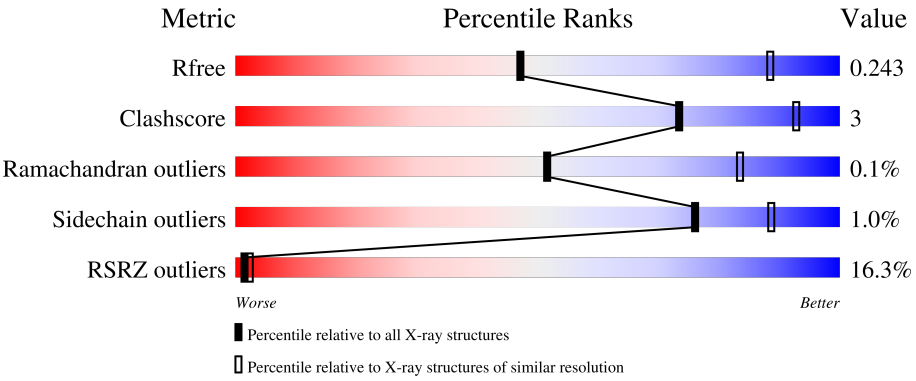
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
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 3.42 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.



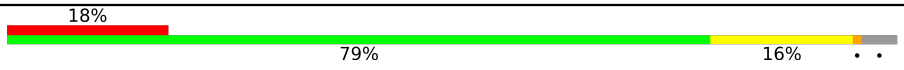
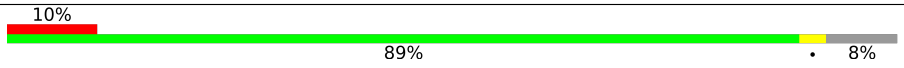
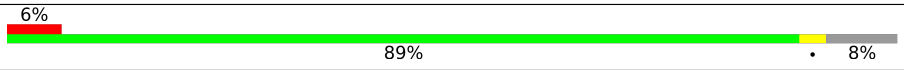
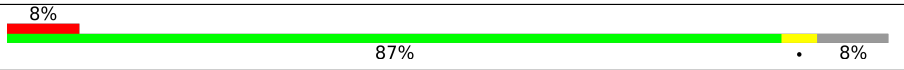
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1486 (3.50-3.34)
Clashscore	141614	1572 (3.50-3.34)
Ramachandran outliers	138981	1534 (3.50-3.34)
Sidechain outliers	138945	1535 (3.50-3.34)
RSRZ outliers	127900	1395 (3.50-3.34)

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	17% 87% 12% •
1	D	840	14% 90% 9% •
1	G	840	19% 87% 11% •
2	B	344	21% 83% 11% • 5%
2	E	344	16% 87% 8% • •

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
2	TPO	E	893	-	-	-	X
2	TPO	H	893	-	-	-	X
4	CIT	A	1201	-	-	-	X
4	CIT	F	301	-	-	-	X
4	CIT	F	302	-	-	-	X
4	CIT	H	1101	-	-	-	X

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 67518 atoms, of which 33659 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	6460	0	0
			12955	4111	6460	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	327	Total	C	H	N	O	P	S	2777	0	0
			5341	1706	2676	451	490	1	17			
2	E	329	Total	C	H	N	O	P	S	2777	0	0
			5384	1723	2697	454	492	1	17			
2	H	330	Total	C	H	N	O	P	S	2797	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

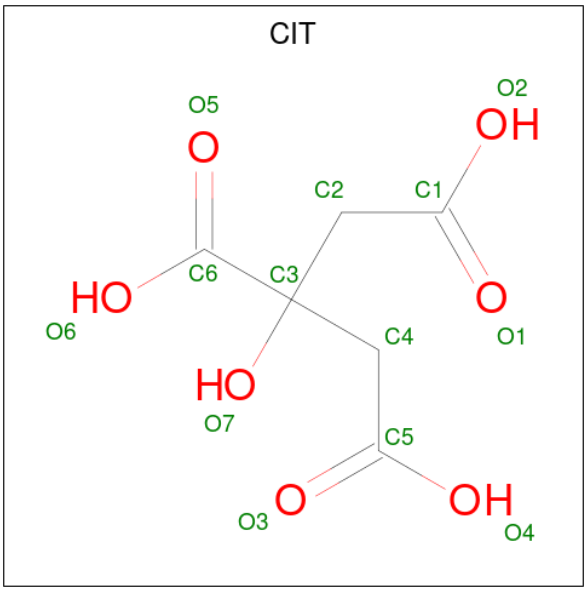
- Molecule 3 is a protein called Cyclin-K.

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:

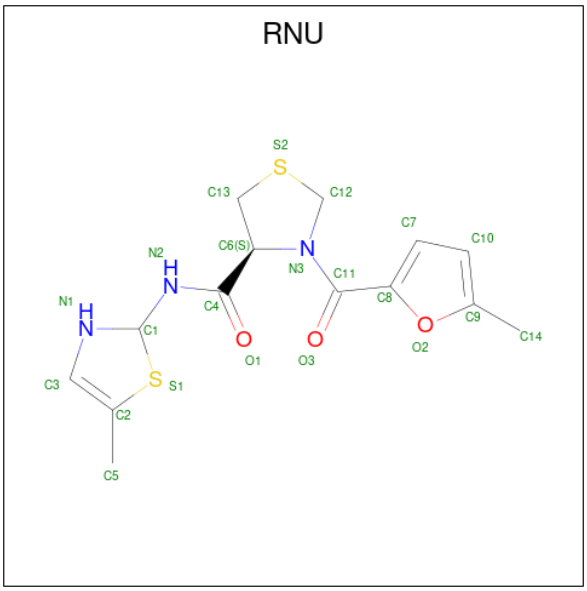
Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	GLY	-	expression tag	UNP O75909
C	-2	GLY	-	expression tag	UNP O75909
C	-1	GLY	-	expression tag	UNP O75909
C	0	ARG	-	expression tag	UNP O75909
F	-3	GLY	-	expression tag	UNP O75909
F	-2	GLY	-	expression tag	UNP O75909
F	-1	GLY	-	expression tag	UNP O75909
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 CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C H O 18 6 5 7	5	0
4	C	1	Total C H O 18 6 5 7	5	0
4	D	1	Total C H O 18 6 5 7	5	0
4	F	1	Total C H O 18 6 5 7	5	0
4	F	1	Total C H O 18 6 5 7	5	0
4	G	1	Total C H O 18 6 5 7	5	0
4	H	1	Total C H O 18 6 5 7	5	0

- Molecule 5 is {N}-(5-methyl-2,3-dihydro-1,3-thiazol-2-yl)-3-(5-methylfuran-2-yl)carbonyl-1,3-thiazolidine-4-carboxamide (three-letter code: RNU) (formula: $C_{14}H_{17}N_3O_3S_2$) (labeled as "Ligand of Interest" by depositor).



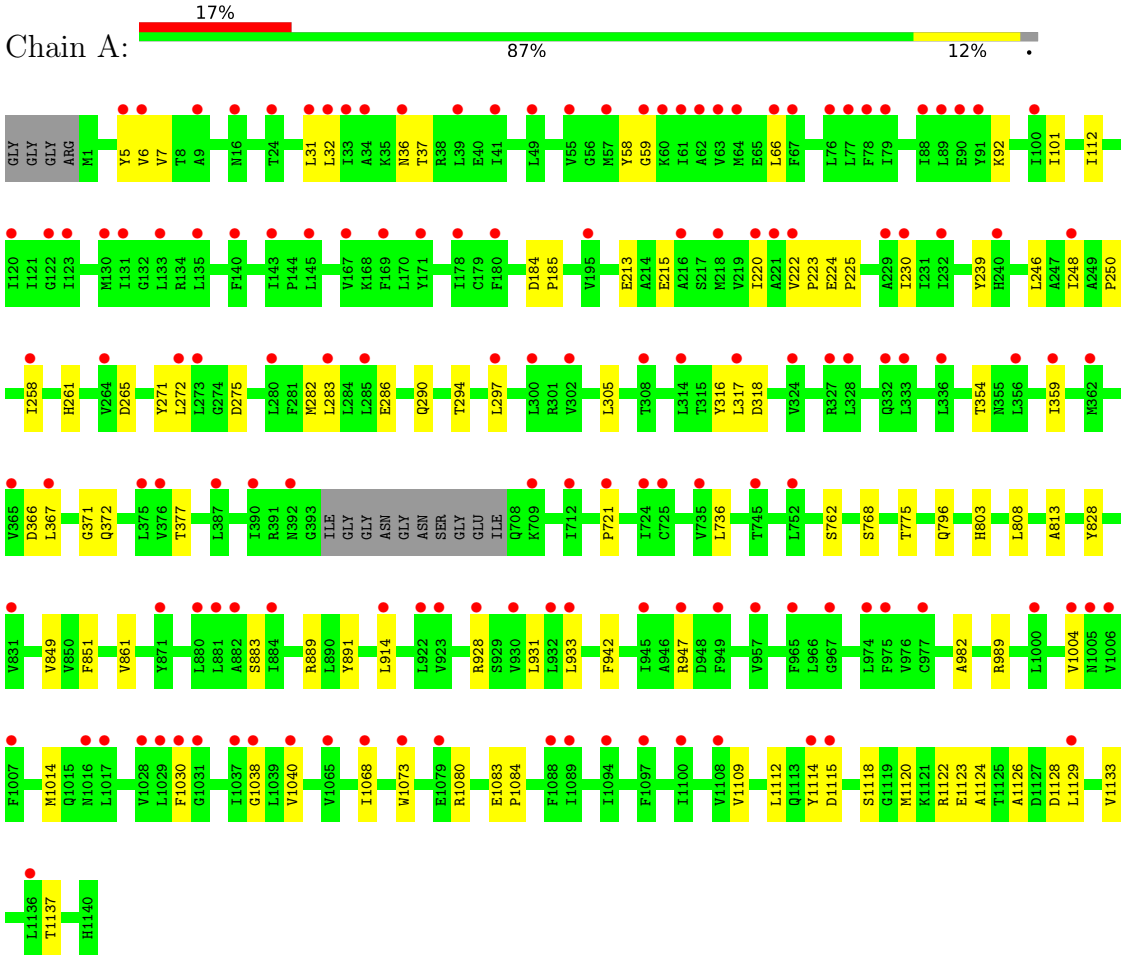
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
5	B	1	Total	C	H	N	O	S	15	0
			37	14	15	3	3	2		
5	E	1	Total	C	H	N	O	S	15	0
			37	14	15	3	3	2		
5	H	1	Total	C	H	N	O	S	15	0
			37	14	15	3	3	2		

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.

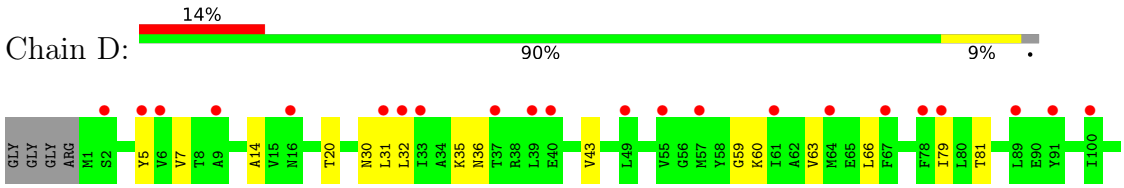
• Molecule 1: DNA damage-binding protein 1

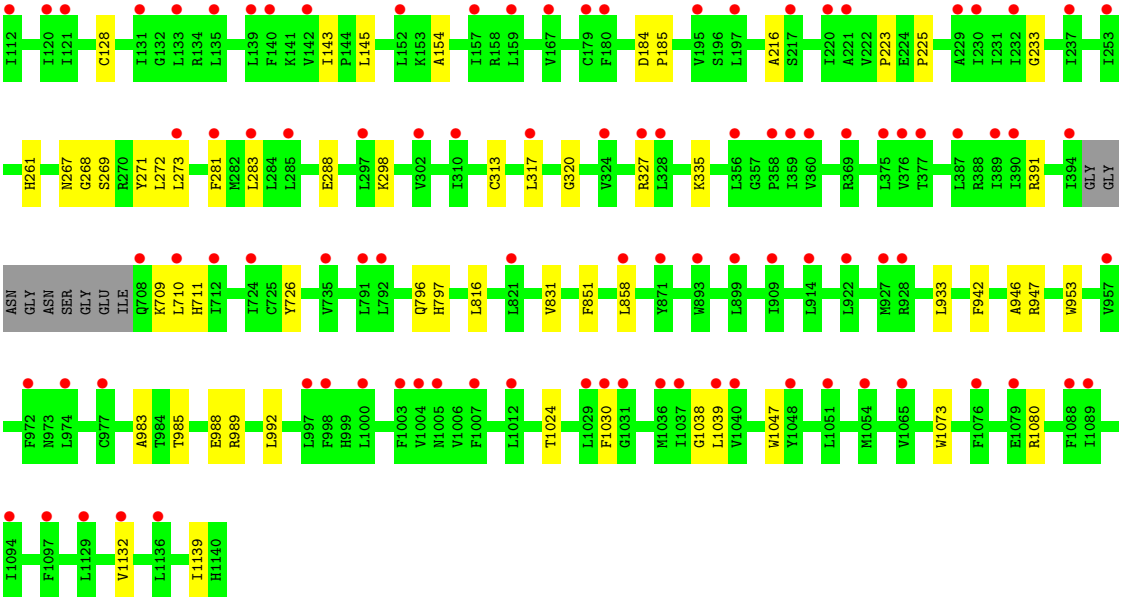
Chain A:



• Molecule 1: DNA damage-binding protein 1

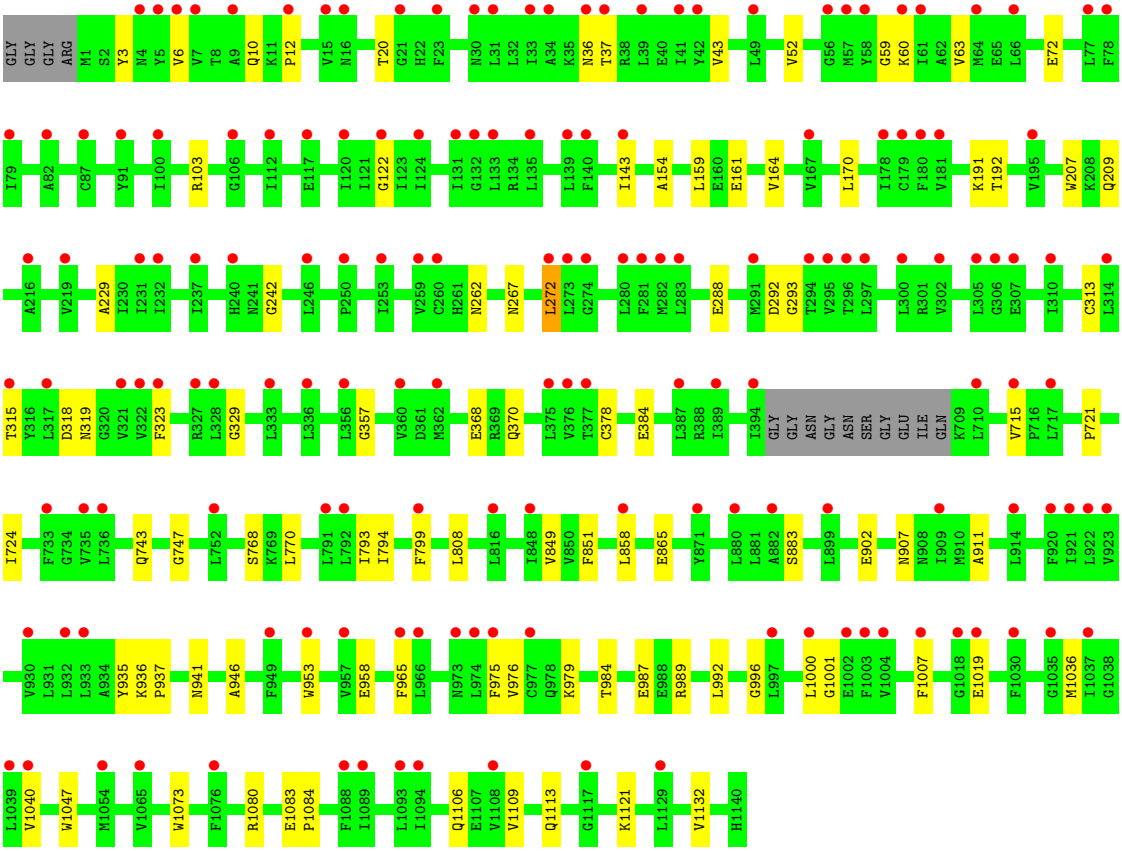
Chain D:





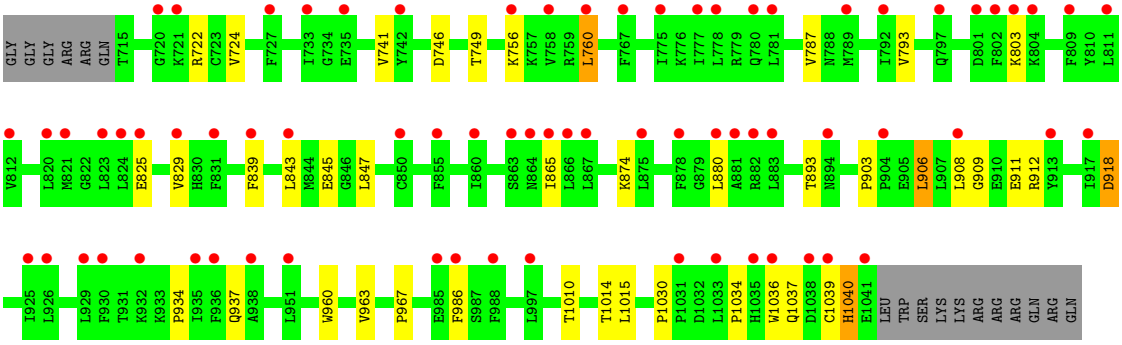
• Molecule 1: DNA damage-binding protein 1

Chain G:



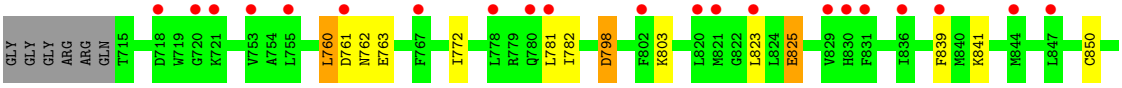
• Molecule 2: Cyclin-dependent kinase 12

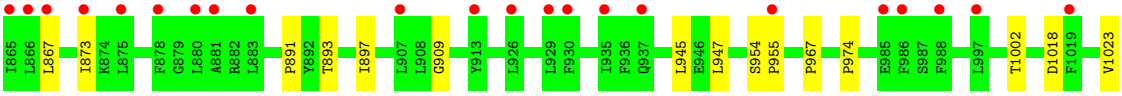
Chain B:

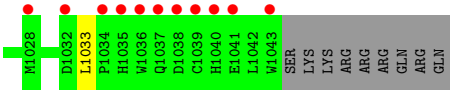


• Molecule 2: Cyclin-dependent kinase 12

Chain E:

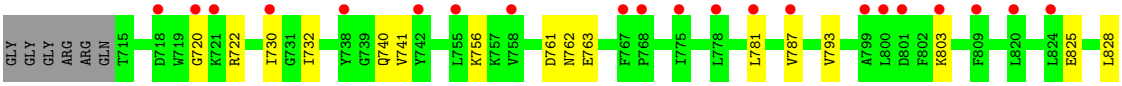


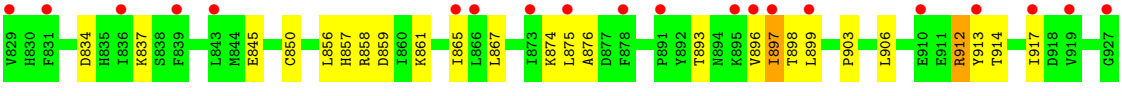


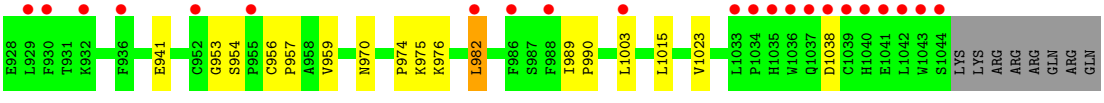


• Molecule 2: Cyclin-dependent kinase 12

Chain H:

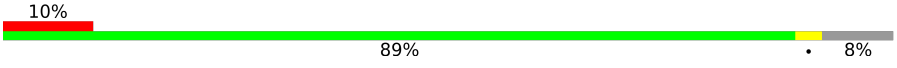


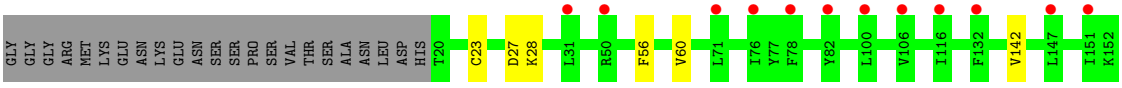


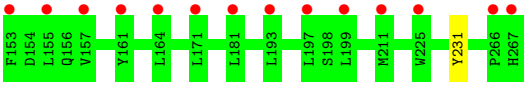


• Molecule 3: Cyclin-K

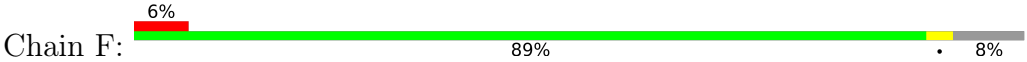
Chain C:

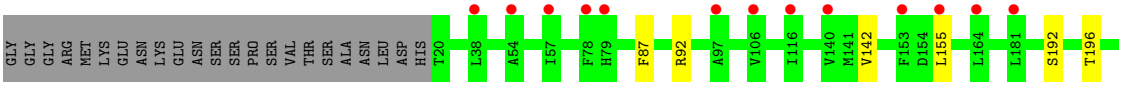


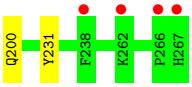




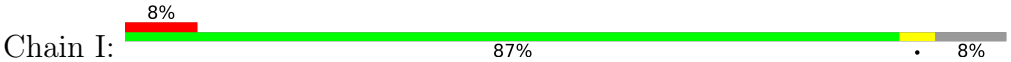
• Molecule 3: Cyclin-K

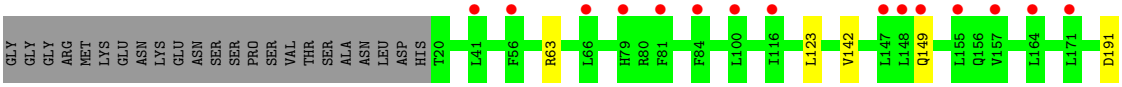


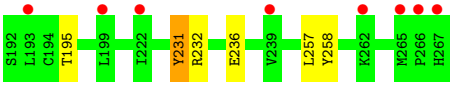




• Molecule 3: Cyclin-K







4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	247.86Å 247.86Å 220.72Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	60.04 – 3.42 214.65 – 3.42	Depositor EDS
% Data completeness (in resolution range)	84.6 (60.04-3.42) 84.6 (214.65-3.42)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.45 (at 3.41Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.212 , 0.237 0.220 , 0.243	Depositor DCC
R_{free} test set	4460 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	147.3	Xtriage
Anisotropy	0.004	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 105.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.046 for -h,-k,l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	67518	wwPDB-VP
Average B, all atoms (Å^2)	172.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.69% 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: RNU, CIT, TPO

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	0.51	0/8931
1	D	0.27	0/6612	0.51	0/8942
1	G	0.28	0/6603	0.51	0/8930
2	B	0.28	0/2713	0.48	0/3657
2	E	0.29	0/2737	0.49	0/3691
2	H	0.31	0/2743	0.54	2/3699 (0.1%)
3	C	0.26	0/2120	0.45	0/2868
3	F	0.28	0/2120	0.47	0/2868
3	I	0.26	0/2120	0.45	0/2868
All	All	0.28	0/34372	0.50	2/46454 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	H	912	ARG	CG-CD-NE	-6.60	97.94	111.80
2	H	912	ARG	NE-CZ-NH1	6.17	123.38	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	6487	6449	6451	53	0
1	D	6495	6460	6462	40	0
1	G	6486	6452	6454	54	0
2	B	2665	2676	2676	24	0
2	E	2687	2697	2697	14	1
2	H	2693	2701	2701	33	0
3	C	2063	2048	2048	3	0
3	F	2063	2048	2048	5	0
3	I	2063	2048	2048	7	1
4	A	13	5	5	0	0
4	C	13	5	5	0	0
4	D	13	5	5	0	0
4	F	26	10	10	0	0
4	G	13	5	5	0	0
4	H	13	5	5	0	0
5	B	22	15	0	1	0
5	E	22	15	0	0	0
5	H	22	15	0	0	0
All	All	33859	33659	33620	222	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

The worst 5 of 222 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:1118:SER:HB2	1:A:1122:ARG:NH1	2.10	0.66
2:H:865:ILE:HD13	2:H:875:LEU:HD13	1.78	0.65
2:H:899:LEU:HD11	2:H:941:GLU:HA	1.77	0.65
3:F:142:VAL:HG11	2:E:803:LYS:HG3	1.80	0.64
1:A:1114:TYR:HE2	1:A:1124:ALA:HB2	1.62	0.64

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:I:232:ARG:HE	2:E:1018:ASP:OD2[5_554]	1.42	0.18

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%)	792 (96%)	29 (4%)	1 (0%)	51	83
1	D	823/840 (98%)	803 (98%)	20 (2%)	0	100	100
1	G	822/840 (98%)	795 (97%)	26 (3%)	1 (0%)	51	83
2	B	324/344 (94%)	317 (98%)	6 (2%)	1 (0%)	41	74
2	E	326/344 (95%)	318 (98%)	7 (2%)	1 (0%)	41	74
2	H	327/344 (95%)	320 (98%)	7 (2%)	0	100	100
3	C	246/271 (91%)	241 (98%)	5 (2%)	0	100	100
3	F	246/271 (91%)	243 (99%)	3 (1%)	0	100	100
3	I	246/271 (91%)	243 (99%)	3 (1%)	0	100	100
All	All	4182/4365 (96%)	4072 (97%)	106 (2%)	4 (0%)	51	83

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	1040	HIS
1	G	357	GLY
1	A	721	PRO
2	E	891	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	721/728 (99%)	719 (100%)	2 (0%)	92	97
1	D	722/728 (99%)	718 (99%)	4 (1%)	86	94
1	G	721/728 (99%)	715 (99%)	6 (1%)	81	92
2	B	294/308 (96%)	289 (98%)	5 (2%)	60	82
2	E	296/308 (96%)	288 (97%)	8 (3%)	44	73
2	H	297/308 (96%)	290 (98%)	7 (2%)	49	76
3	C	223/242 (92%)	221 (99%)	2 (1%)	78	90
3	F	223/242 (92%)	221 (99%)	2 (1%)	78	90
3	I	223/242 (92%)	222 (100%)	1 (0%)	91	96
All	All	3720/3834 (97%)	3683 (99%)	37 (1%)	76	88

5 of 37 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	947	LEU
2	H	982	LEU
2	E	1002	THR
2	H	867	LEU
1	D	953	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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	B	893	2	8,10,11	1.62	1 (12%)	10,14,16	1.33	1 (10%)
2	TPO	H	893	2	8,10,11	1.69	1 (12%)	10,14,16	1.53	2 (20%)
2	TPO	E	893	2	8,10,11	1.64	1 (12%)	10,14,16	1.27	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	B	893	2	-	5/9/11/13	-
2	TPO	H	893	2	-	5/9/11/13	-
2	TPO	E	893	2	-	3/9/11/13	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	893	TPO	P-O1P	3.56	1.62	1.50
2	B	893	TPO	P-O1P	3.52	1.61	1.50
2	E	893	TPO	P-O1P	3.51	1.61	1.50

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	893	TPO	P-OG1-CB	-3.67	112.13	123.21
2	H	893	TPO	P-OG1-CB	-3.43	112.83	123.21
2	E	893	TPO	P-OG1-CB	-3.18	113.59	123.21
2	H	893	TPO	CG2-CB-CA	-2.24	108.74	113.16

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

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

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

10 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	CIT	A	1201	-	12,12,12	1.05	0	17,17,17	1.59	2 (11%)
4	CIT	D	1201	-	12,12,12	1.09	0	17,17,17	1.58	2 (11%)
5	RNU	E	1101	-	20,24,24	1.88	4 (20%)	21,34,34	1.80	3 (14%)
5	RNU	H	1102	-	20,24,24	1.88	4 (20%)	21,34,34	1.79	4 (19%)
4	CIT	F	301	-	12,12,12	1.06	0	17,17,17	1.61	3 (17%)
4	CIT	H	1101	-	12,12,12	1.06	0	17,17,17	1.43	1 (5%)
4	CIT	G	1201	-	12,12,12	1.05	0	17,17,17	1.68	3 (17%)
4	CIT	F	302	-	12,12,12	1.06	0	17,17,17	1.55	3 (17%)
5	RNU	B	1101	-	20,24,24	1.90	4 (20%)	21,34,34	1.76	4 (19%)
4	CIT	C	301	-	12,12,12	1.05	0	17,17,17	1.59	2 (11%)

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	CIT	A	1201	-	-	3/16/16/16	-
4	CIT	D	1201	-	-	8/16/16/16	-
5	RNU	E	1101	-	-	0/10/35/35	0/3/3/3
5	RNU	H	1102	-	-	1/10/35/35	0/3/3/3

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	CIT	F	301	-	-	3/16/16/16	-
4	CIT	H	1101	-	-	0/16/16/16	-
4	CIT	G	1201	-	-	0/16/16/16	-
4	CIT	F	302	-	-	0/16/16/16	-
5	RNU	B	1101	-	-	1/10/35/35	0/3/3/3
4	CIT	C	301	-	-	6/16/16/16	-

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	1101	RNU	C1-N1	-5.29	1.30	1.44
5	H	1102	RNU	C1-N1	-5.28	1.30	1.44
5	B	1101	RNU	C1-N1	-5.28	1.30	1.44
5	E	1101	RNU	C3-N1	4.06	1.39	1.33
5	H	1102	RNU	C3-N1	4.04	1.39	1.33

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	E	1101	RNU	C3-C2-S1	-5.03	107.47	110.16
5	H	1102	RNU	C3-C2-S1	-4.90	107.54	110.16
5	B	1101	RNU	C3-C2-S1	-4.87	107.55	110.16
4	G	1201	CIT	O6-C6-C3	4.49	120.84	113.05
5	H	1102	RNU	C14-C9-C10	4.26	134.68	128.25

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1201	CIT	C1-C2-C3-O7
4	A	1201	CIT	C1-C2-C3-C4
4	A	1201	CIT	C1-C2-C3-C6
4	C	301	CIT	C1-C2-C3-O7
4	C	301	CIT	C1-C2-C3-C4

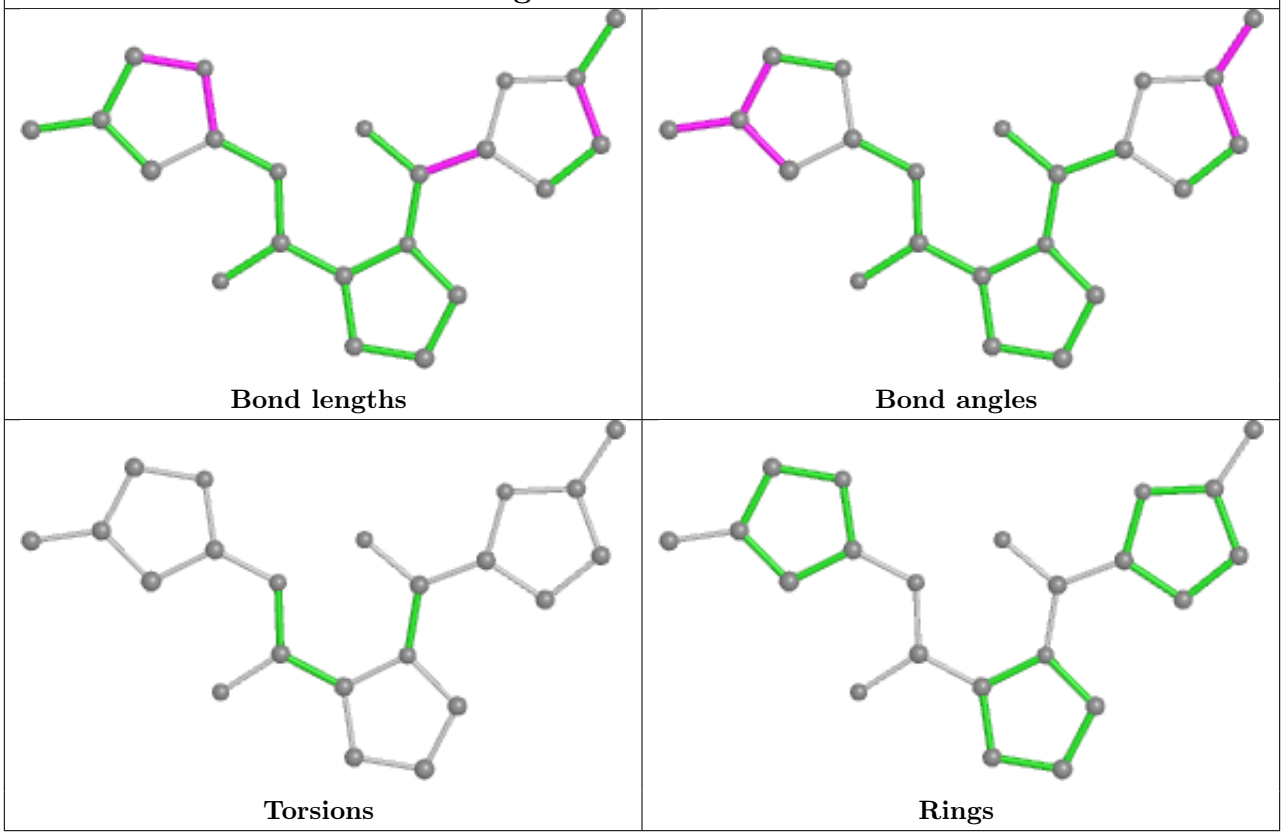
There are no ring outliers.

1 monomer is involved in 1 short contact:

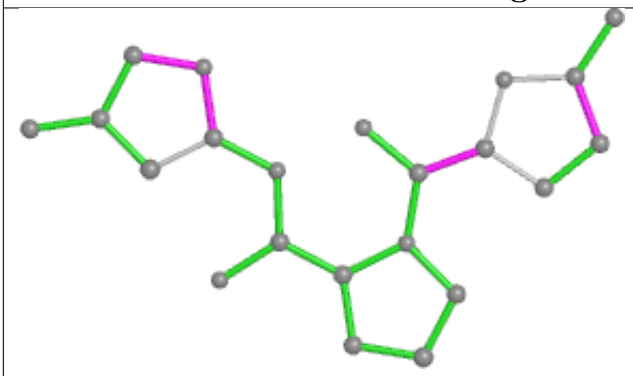
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	1101	RNU	1	0

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.

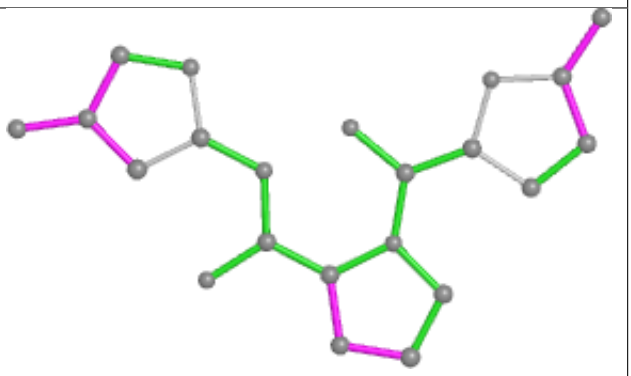
Ligand RNU E 1101



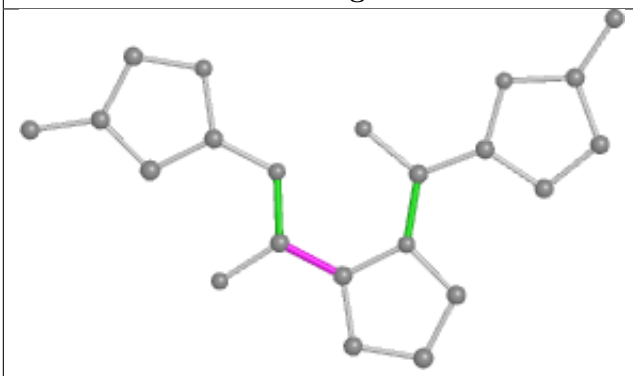
Ligand RNU H 1102



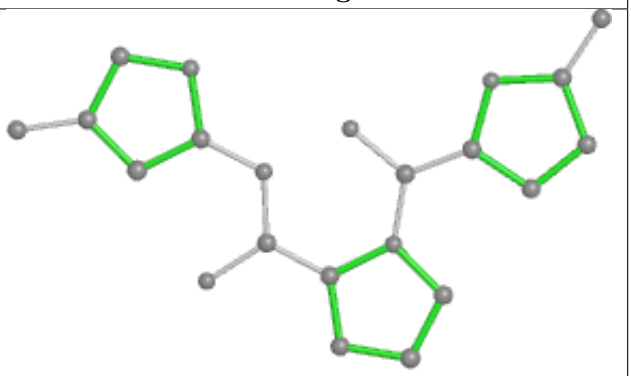
Bond lengths



Bond angles

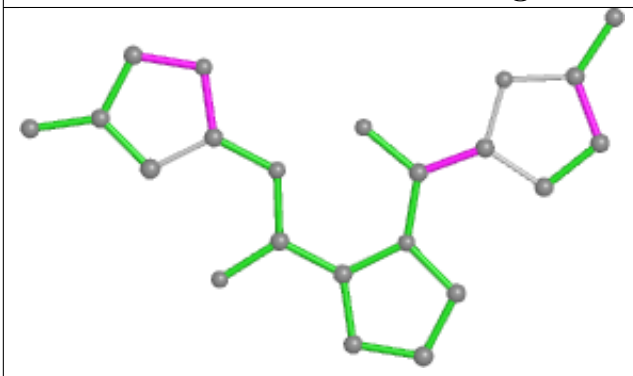


Torsions

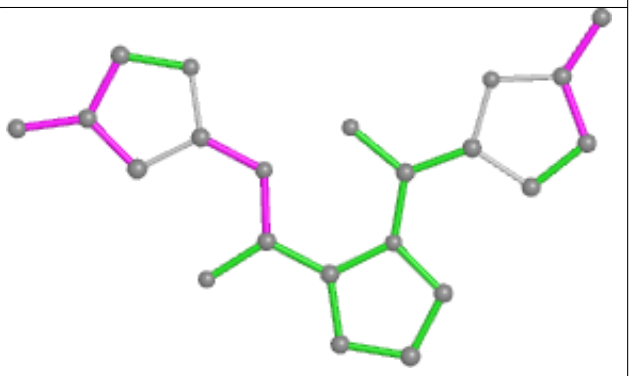


Rings

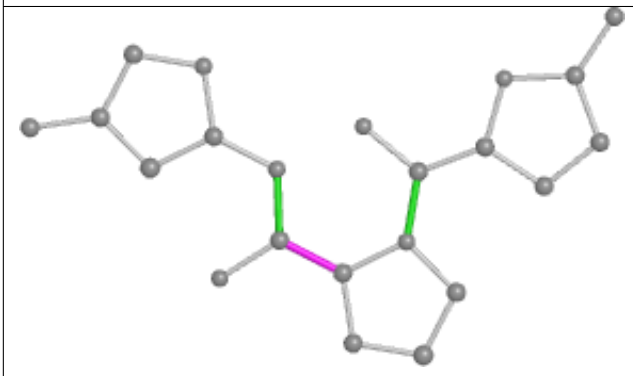
Ligand RNU B 1101



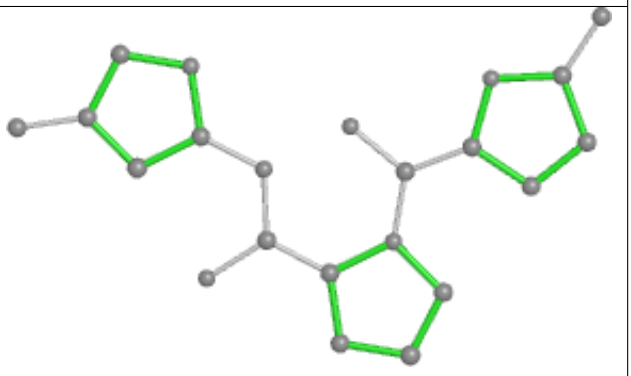
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

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%)	0.83	145 (17%) 1 2	106, 157, 224, 362	0
1	D	827/840 (98%)	0.73	119 (14%) 2 4	103, 147, 219, 301	0
1	G	826/840 (98%)	0.88	159 (19%) 1 1	113, 161, 241, 322	0
2	B	314/344 (91%)	1.10	73 (23%) 0 1	128, 159, 209, 262	0
2	E	319/344 (92%)	1.04	54 (16%) 1 2	109, 150, 222, 277	0
2	H	319/344 (92%)	1.21	63 (19%) 1 1	100, 136, 222, 328	1 (0%)
3	C	248/271 (91%)	0.71	26 (10%) 6 9	115, 145, 190, 269	0
3	F	248/271 (91%)	0.80	17 (6%) 16 20	90, 117, 156, 255	0
3	I	248/271 (91%)	0.78	23 (9%) 8 11	109, 137, 179, 237	0
All	All	4175/4365 (95%)	0.87	679 (16%) 1 2	90, 150, 223, 362	1 (0%)

The worst 5 of 679 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	1040	HIS	12.4
2	E	1043	TRP	12.0
2	H	891	PRO	10.5
2	B	802	PHE	9.4
2	H	1043	TRP	7.7

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	H	893	11/12	0.46	0.54	191,206,247,248	5
2	TPO	E	893	11/12	0.63	0.59	219,240,275,294	6
2	TPO	B	893	11/12	0.84	0.20	224,235,278,282	6

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

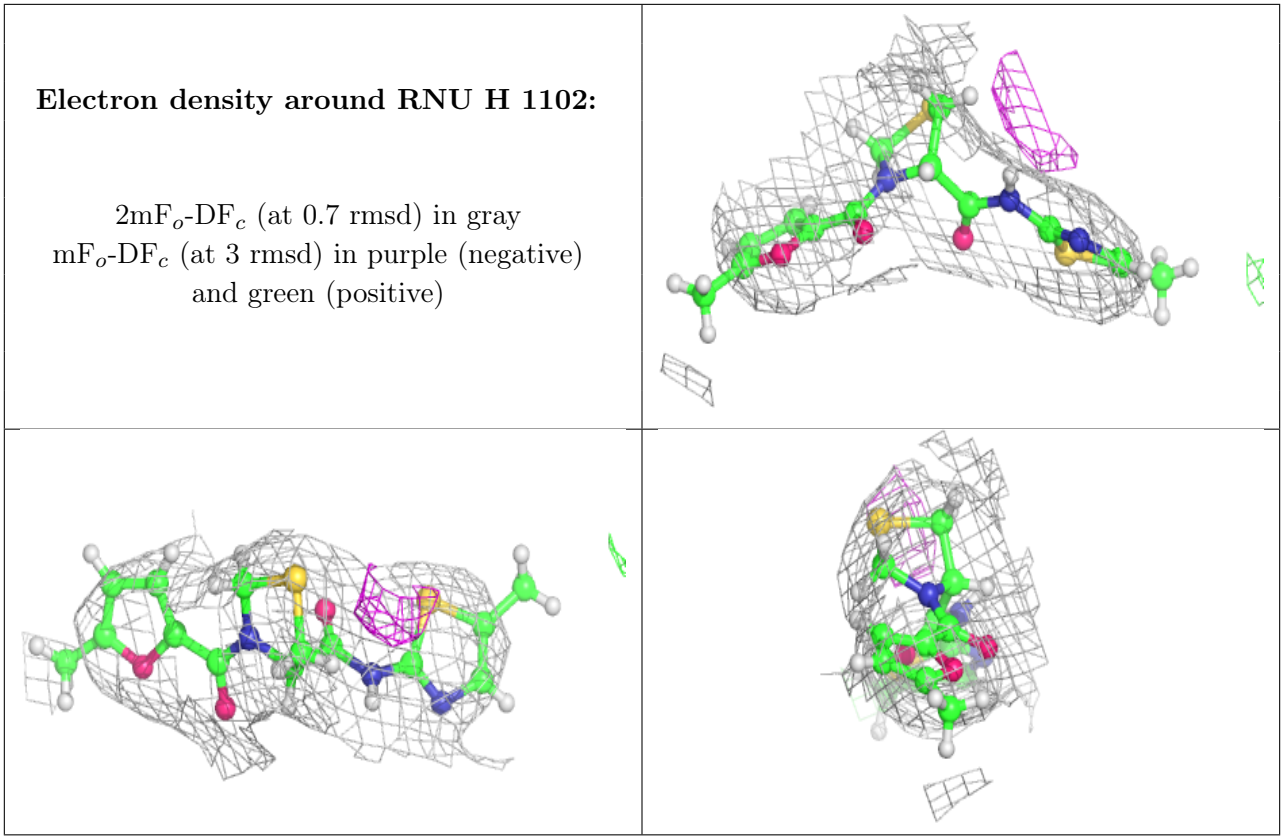
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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	CIT	F	301	13/13	0.59	0.47	155,177,228,228	5
4	CIT	F	302	13/13	0.61	1.00	158,200,227,234	5
4	CIT	A	1201	13/13	0.69	0.42	178,206,245,245	5
4	CIT	D	1201	13/13	0.72	0.32	174,202,235,235	5
4	CIT	H	1101	13/13	0.73	0.41	191,218,248,268	5
4	CIT	C	301	13/13	0.78	0.23	161,170,204,205	5
4	CIT	G	1201	13/13	0.84	0.42	181,197,235,264	5
5	RNU	H	1102	22/22	0.92	0.38	128,146,179,190	15
5	RNU	B	1101	22/22	0.93	0.44	135,144,176,177	15
5	RNU	E	1101	22/22	0.94	0.51	134,147,177,177	15

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.

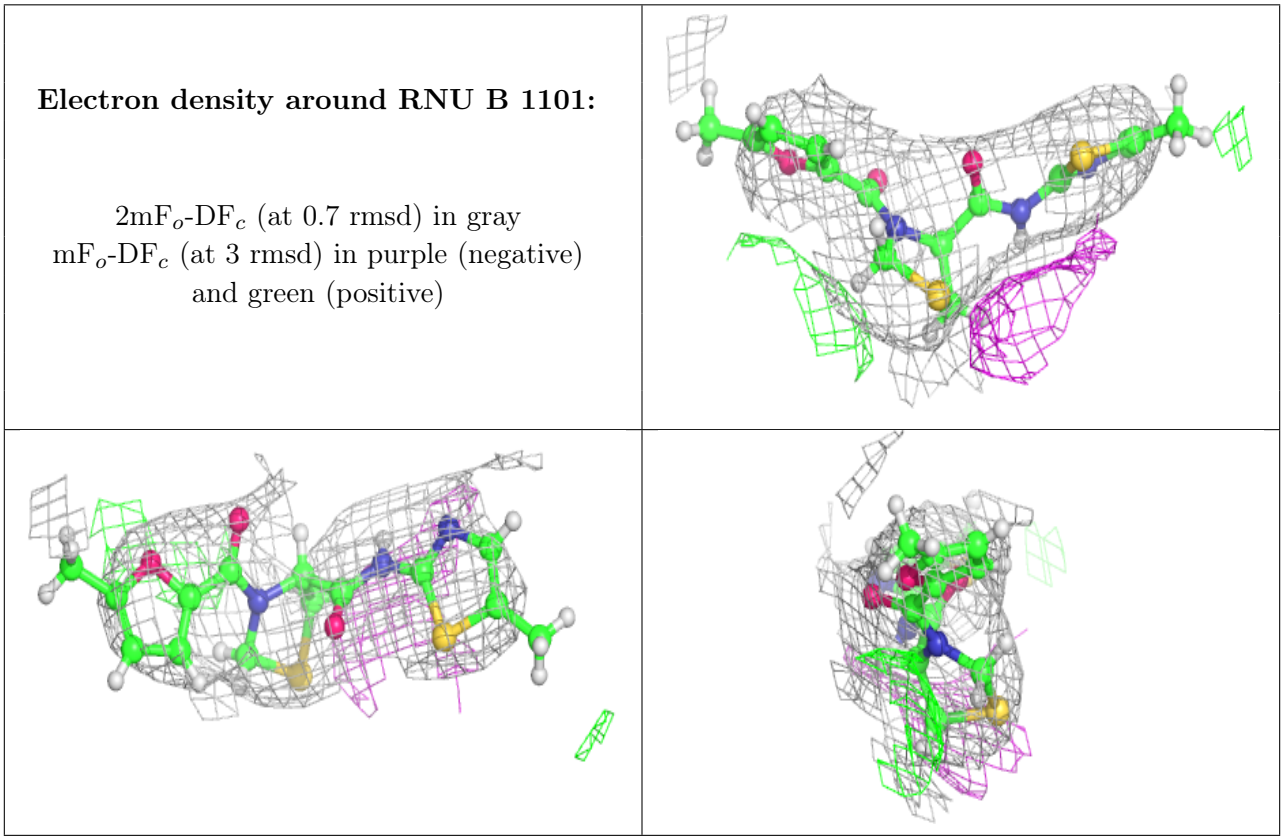
Electron density around RNU H 1102:

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



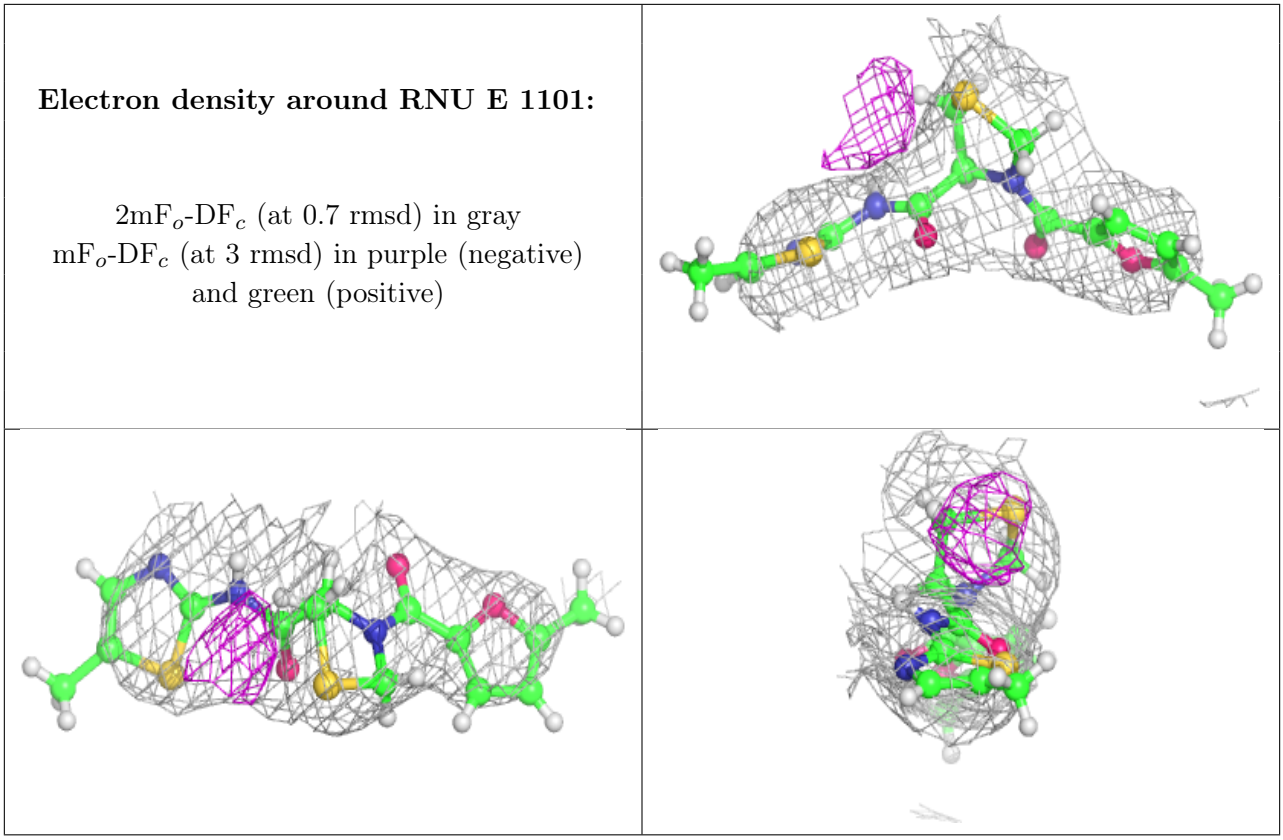
Electron density around RNU B 1101:

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



Electron density around RNU 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)



6.5 Other polymers

There are no such residues in this entry.