W O R L D W I D E

wwPDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Sep 7, 2023 – 01:48 pm BST

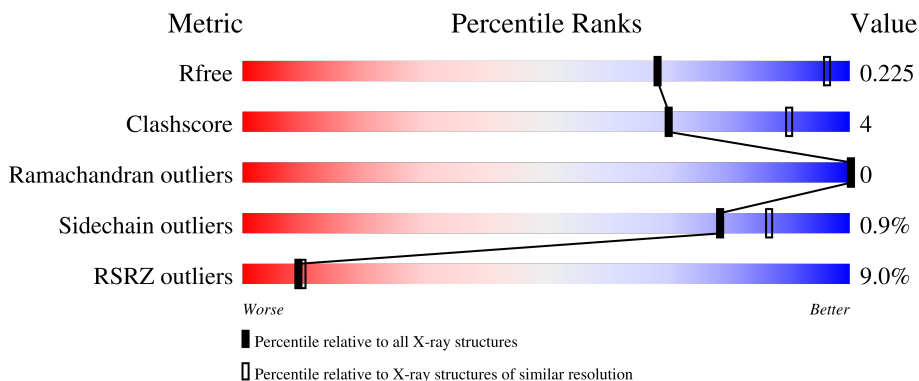|		
PDB ID	:	8BUE
Title	:	Structure of DDB1 bound to Z11-engaged CDK12-cyclin K
Authors	:	Kozicka, Z.; Kempf, G.; Petzold, G.; Thoma, N.H.
Deposited on	:	2022-11-30
Resolution	:	3.25 Å(reported)

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 3.25 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	Percentile Ranks	Value
Rfree		0.225
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		0.9%
RSRZ outliers		9.0%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	1191 (3.30-3.22)
Clashscore	141614	1251 (3.30-3.22)
Ramachandran outliers	138981	1229 (3.30-3.22)
Sidechain outliers	138945	1228 (3.30-3.22)
RSRZ outliers	127900	1154 (3.30-3.22)

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	10% / 87% / 11% / •
1	D	840	8% / 88% / 11% / •
1	G	840	9% / 87% / 11% / •
2	B	344	15% / 83% / 12% / • •
2	E	344	8% / 83% / 11% / • 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	344	9% 83% 12%
3	C	271	6% 85% 6% 8%
3	F	271	4% 86% 8%
3	I	271	6% 87% 8%

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	SO4	F	301	-	-	-	X

2   Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 67598 atoms, of which 33625 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 12937	C 4105	H 6450	N 1094	O 1252	S 36	6450	0	0
1	D	827	Total 12957	C 4111	H 6462	N 1095	O 1253	S 36	6462	0	0
1	G	826	Total 12938	C 4106	H 6452	N 1093	O 1251	S 36	6452	0	0

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
			Total	C	H	N	O	P	S			
2	B	329	5384	1723	2697	454	492	1	17	2697	0	0
2	E	325	Total	C	H	N	O	P	S	2663	0	0
			5309	1695	2663	447	486	1	17			
2	H	329	Total	C	H	N	O	P	S	2697	0	0
			5384	1723	2697	454	492	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 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 SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

O1

O

O3 O⁻ — S — O⁻ O4

S

O

O2

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	B	1	Total 5	O 4	S 1	0	0
4	B	1	Total 5	O 4	S 1	0	0
4	C	1	Total 5	O 4	S 1	0	0
4	C	1	Total 5	O 4	S 1	0	0
4	C	1	Total 5	O 4	S 1	0	0
4	D	1	Total 5	O 4	S 1	0	0
4	D	1	Total 5	O 4	S 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	1	Total 5	O 4	S 1	0	0
4	D	1	Total 5	O 4	S 1	0	0
4	D	1	Total 5	O 4	S 1	0	0
4	D	1	Total 5	O 4	S 1	0	0
4	D	1	Total 5	O 4	S 1	0	0
4	D	1	Total 5	O 4	S 1	0	0
4	D	1	Total 5	O 4	S 1	0	0
4	D	1	Total 5	O 4	S 1	0	0
4	E	1	Total 5	O 4	S 1	0	0
4	E	1	Total 5	O 4	S 1	0	0
4	F	1	Total 5	O 4	S 1	0	0
4	F	1	Total 5	O 4	S 1	0	0
4	F	1	Total 5	O 4	S 1	0	0
4	F	1	Total 5	O 4	S 1	0	0
4	G	1	Total 5	O 4	S 1	0	0
4	G	1	Total 5	O 4	S 1	0	0
4	G	1	Total 5	O 4	S 1	0	0
4	G	1	Total 5	O 4	S 1	0	0
4	G	1	Total 5	O 4	S 1	0	0
4	G	1	Total 5	O 4	S 1	0	0
4	G	1	Total 5	O 4	S 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	G	1	Total O S 5 4 1			0	0
4	H	1	Total O S 5 4 1			0	0
4	H	1	Total O S 5 4 1			0	0
4	H	1	Total O S 5 4 1			0	0
4	H	1	Total O S 5 4 1			0	0
4	I	1	Total O S 5 4 1			0	0
4	I	1	Total O S 5 4 1			0	0
4	I	1	Total O S 5 4 1			0	0

- Molecule 5 is {N}-(1 {H}-benzimidazol-2-yl)-1-(2-methoxy-5-methyl-phenyl)-5-oxidanylidene-pyrrolidine-3-carboxamide (three-letter code: RVU) (formula: $C_{20}H_{20}N_4O_3$) (labeled as "Ligand of Interest" by depositor).

RVU
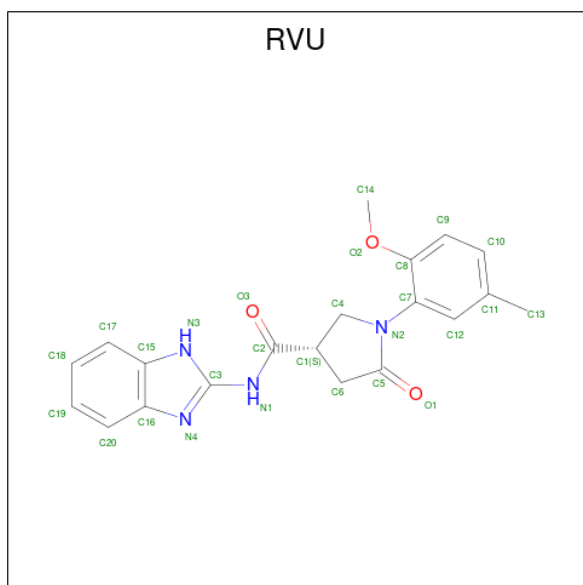
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	B	1	Total 47	C 20	H 20	N 4	O 3	20	0
5	E	1	Total 47	C 20	H 20	N 4	O 3	20	0

Continued from previous page...

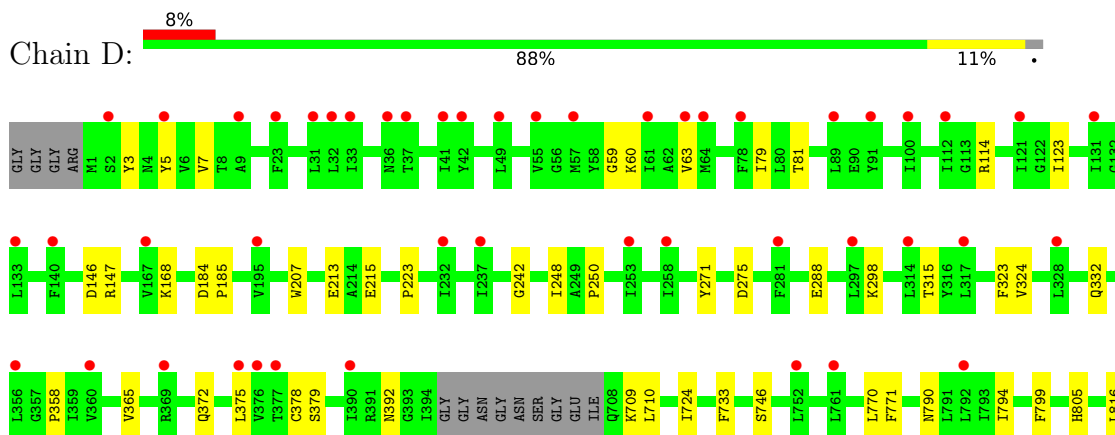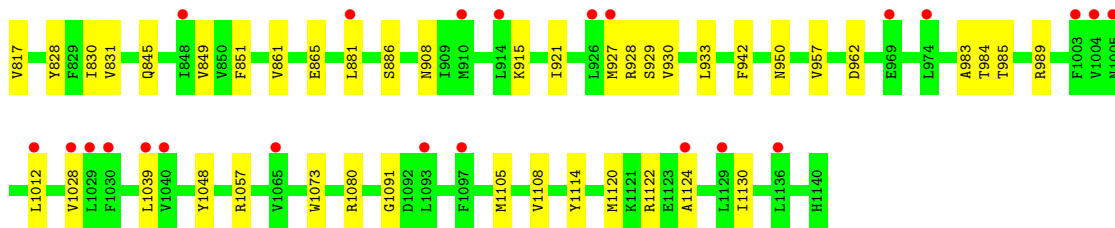| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
5	H	1	47	20	20	4	3	20	0

3 Residue-property plots ⓘ

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.
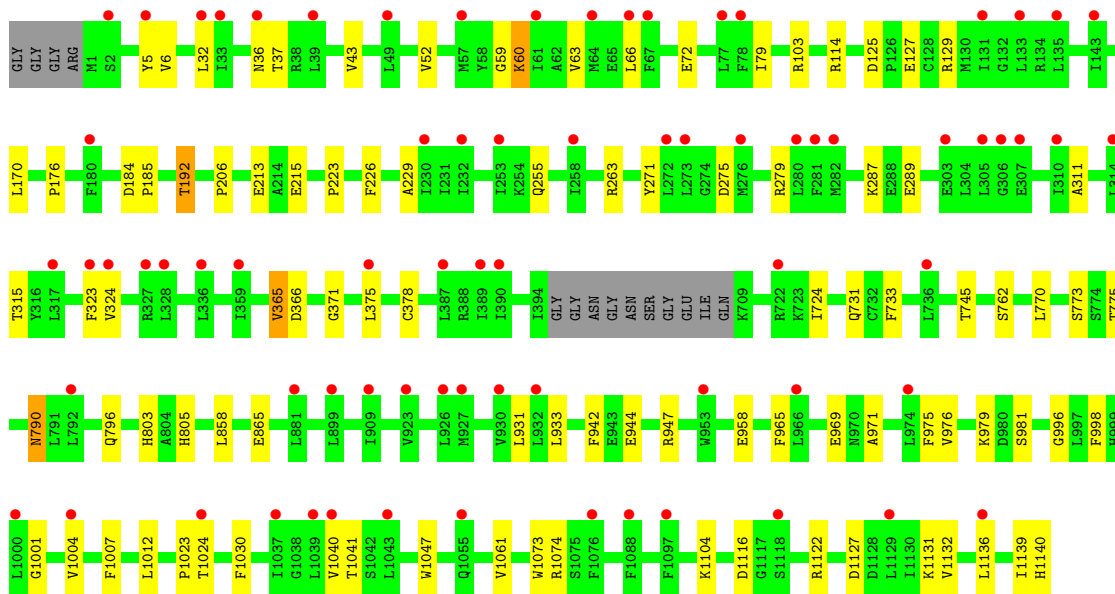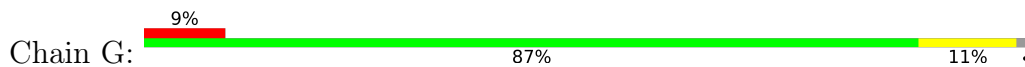
- Molecule 1: DNA damage-binding protein 1

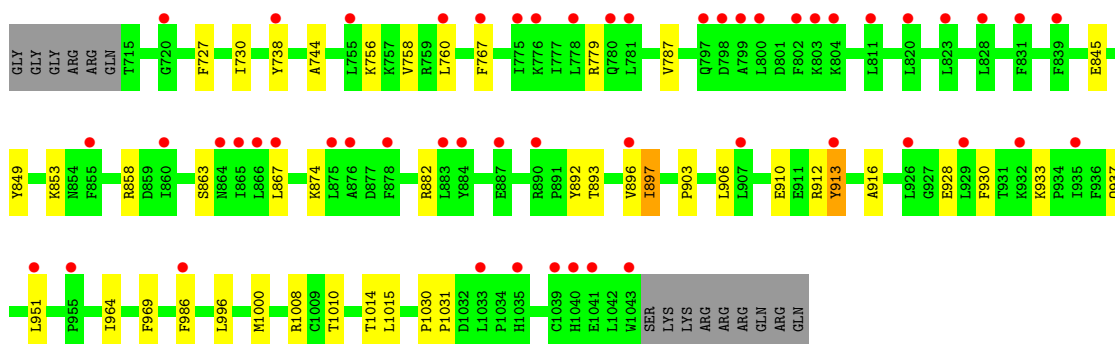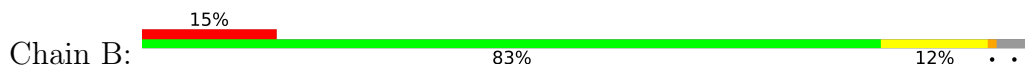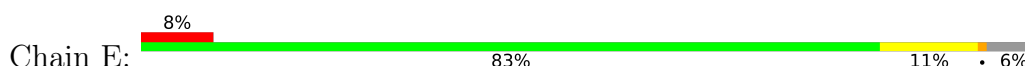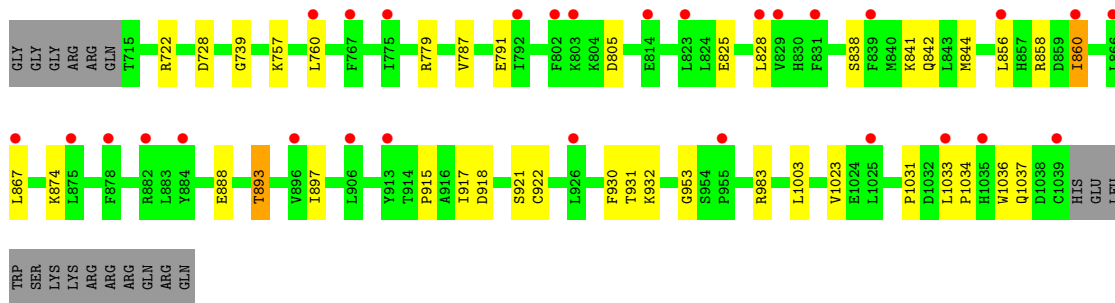- Molecule 1: DNA damage-binding protein 1

V817 Y828 F829 I830 V831 Q845 I848 V849 V850 F851 V861 E865 L881 S886 N908 I909 M910 L914 K915 I921 L926 M927 R928 S929 V930 L933 F942 N950 V957 D962 E969 L974 A983 T984 T985 R989 F1003 V1004 N1005

L1012 V1028 L1029 F1030 L1039 V1040 Y1048 R1057 V1065 W1073 R1080 G1091 D1092 L1093 F1097 M1105 V1108 Y1114 M1120 K1121 R1122 E1123 A1124 L1129 I1130 L1136 H1140

• Molecule 1: DNA damage-binding protein 1

Chain G:

9%

87% 11%

GLY GLY GLY ARG M1 S2 Y5 V6 L32 I33 N36 T37 R38 L39 V43 L49 V52 M57 Y58 G59 K60 I61 A62 V63 M64 E65 L66 F67 E72 L77 F78 I79 R103 R114 D125 P126 E127 C128 R129 M130 I131 G132 L133 R134 L135 I143

L170 P176 F180 D184 P185 T192 P206 E213 A214 E215 P223 F226 A229 I230 I231 I232 I253 K254 Q255 I258 R263 Y271 L272 L273 G274 D275 M276 R279 L280 F281 M282 K287 E288 E289 E303 L304 L305 G306 E307 I310 A311 L314

T315 Y316 L317 F323 V324 R327 L328 L336 I359 V365 D366 G371 L375 C378 L387 R388 I389 I390 I394 GLY GLY ASN GLY ASN SER GLY GLU ILE GLN K709 R722 K723 I724 Q731 C732 F733 L736 T745 S762 L770 S773 S774 T775

N790 L791 L792 Q796 H803 A804 H805 L858 E865 L881 I899 I909 V923 L926 M927 V930 L931 L932 L933 F942 E943 E944 R947 W953 E958 F965 L966 E969 N970 A971 L974 F975 V976 K979 D980 S981 G996 L997 F998 H999

L1000 G1001 V1004 F1007 L1012 P1023 T1024 F1030 I1037 G1038 L1039 V1040 T1041 S1042 L1043 W1047 Q1055 V1061 W1073 R1074 S1075 F1076 F1088 F1097 K1104 D1116 G1117 S1118 R1122 D1127 D1128 L1129 I1130 K1131 V1132 L1136 I1139 H1140
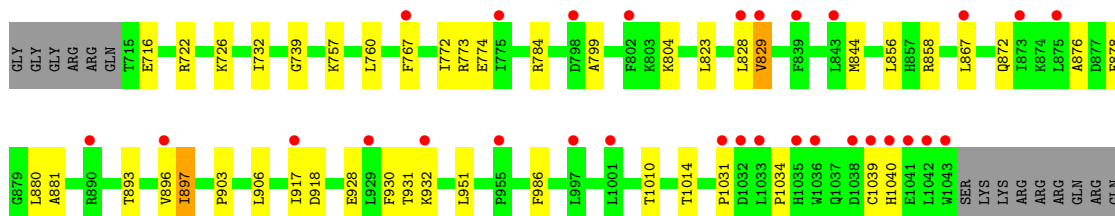
• Molecule 2: Cyclin-dependent kinase 12

Chain B:

15%

83% 12%

GLY GLY GLY ARG ARG GLN T715 G720 F727 I730 Y738 A744 L755 K756 K757 V758 R759 L760 F767 I775 K776 I777 L778 R779 Q780 L781 V787 Q797 D798 A799 L800 D801 F802 K803 K804 L811 L820 L823 L828 F831 F839 E845

Y849 K853 N854 F855 R858 D859 I860 S863 N864 I865 L866 L867 K874 L875 A876 D877 F878 R882 L883 Y884 E887 R890 P891 Y892 T893 V896 I897 P903 L906 L907 E910 E911 R912 Y913 A916 L926 G927 E928 L929 F930 T931 K932 K933 P934 I935 F936 Q937

L951 P955 I964 F969 F986 L996 M1000 R1008 C1009 T1010 T1014 L1015 P1030 P1031 D1032 L1033 P1034 H1035 C1039 H1040 E1041 L1042 W1043 SER LYS LYS ARG ARG GLN ARG GLN
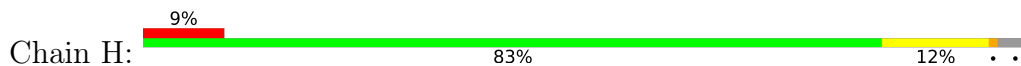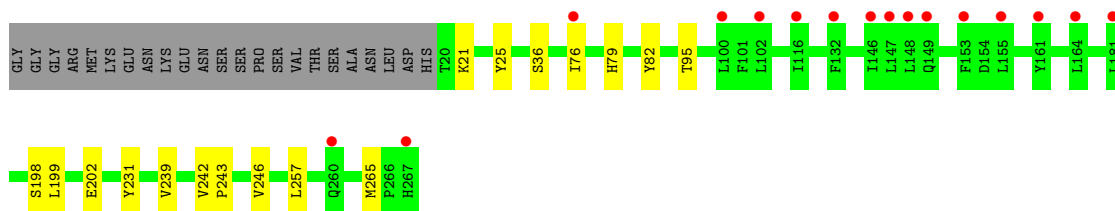
• Molecule 2: Cyclin-dependent kinase 12

Chain E:

8%

83% 11% 6%

GLY GLY GLY ARG ARG GLN T715 R722 D728 G739 K757 L760 F767 I775 R779 V787 E791 I792 F802 K803 K804 D805 E814 L823 L824 E825 L828 V829 H830 F831 S838 F839 M840 K841 Q842 L843 M844 L856 H857 R858 D859 I860 L866

L867 K874 L875 F878 R882 L883 Y884 E888 T893 V896 I897 L906 Y913 T914 P915 A916 I917 D918 S921 C922 L926 F930 T931 K932 G953 S954 P955 R983 L1003 V1023 E1024 L1025 P1031 L1032 L1033 P1034 H1035 W1036 Q1037 D1038 C1039 HIS GLU LEU
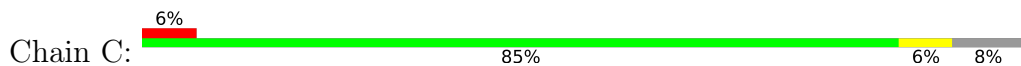
TRP SER LYS ARG ARG GLN ARG GLN

• Molecule 2: Cyclin-dependent kinase 12

Chain H:
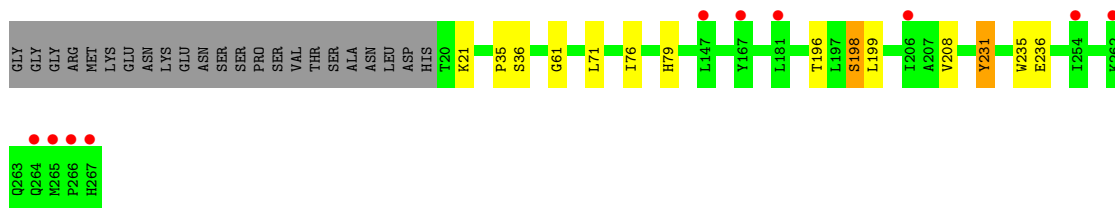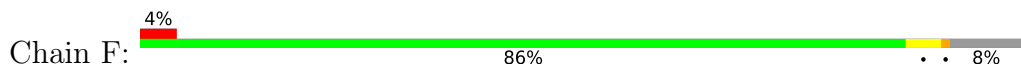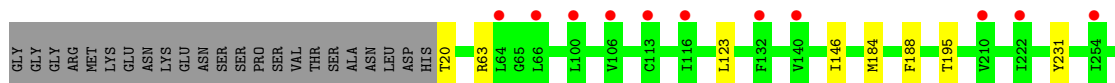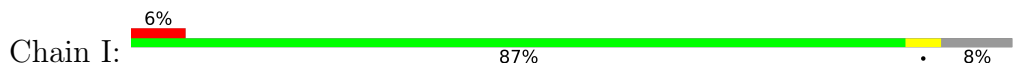
9%

83%　　　　　　　　　　　　　　　　　　12%

GLY GLY GLY ARG ARG GLN T715 E716 R722 K726 I732 G739 K757 L760 F767 I772 R773 E774 I775 R784 D798 A799 F802 K803 K804 L823 L828 V829 F839 L843 M844 L856 H857 R858 L867 Q872 I873 K874 L875 A876 D877 F878

G879 L880 A881 R890 T893 V896 I897 P903 L906 I917 D918 E928 L929 F930 T931 K932 L951 P955 F986 L997 L1001 T1010 T1014 P1031 D1032 L1033 P1034 H1035 W1036 Q1037 D1038 C1039 H1040 E1041 L1042 W1043 SER LYS ARG ARG GLN ARG GLN

• Molecule 3: Cyclin-K

Chain C:

6%

85%　　　　　　　　　　　　　6%　　8%

GLY GLY GLY ARG MET LYS GLU ASN LYS GLU ASN SER SER PRO SER VAL THR SER ALA ASN LEU ASP HIS T20 K21 Y25 S36 I76 H79 Y82 T95 L100 F101 L102 I116 F132 I146 L147 L148 Q149 F153 D154 L155 Y161 L164 L181

S198 L199 E202 Y231 V239 V242 P243 V246 L257 Q260 M265 P266 H267

• Molecule 3: Cyclin-K

Chain F:

4%

86%　　　　　　　　　　　　　　　　•　•　8%

GLY GLY GLY ARG MET LYS GLU ASN LYS GLU ASN SER SER PRO SER VAL THR SER ALA ASN LEU ASP HIS T20 K21 P35 S36 G61 L71 I76 H79 L147 Y167 L181 T196 L197 S198 L199 I206 A207 V208 Y231 W235 E236 I254 K262

Q263 Q264 M265 P266 H267

• Molecule 3: Cyclin-K

Chain I:

6%

87%　　　　　　　　　　　　　　　　　　•　8%

GLY GLY GLY ARG MET LYS GLU ASN LYS GLU ASN SER SER PRO SER VAL THR SER ALA ASN LEU ASP HIS T20 R63 L64 G65 L66 L100 V106 C113 I116 L123 F132 V140 I146 M184 F188 T195 V210 I222 Y231 I254

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	249.33Å 249.33Å 218.68Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	60.41 – 3.25 215.93 – 3.25	Depositor EDS
% Data completeness (in resolution range)	79.3 (60.41-3.25) 79.3 (215.93-3.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.51 (at 3.26Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.184 , 0.213 0.202 , 0.225	Depositor DCC
R_{free} test set	4845 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å^2)	134.4	Xtriage
Anisotropy	0.020	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 92.7	EDS
L-test for twinning[2]	$< |L| > = 0.45, < L^2 > = 0.28$	Xtriage
Estimated twinning fraction	0.035 for -h,-k,l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	67598	wwPDB-VP
Average B, all atoms (Å^2)	156.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.60% of the height of the origin peak. No significant pseudotranslation is detected.*

---

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

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: RVU, SO4, 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	0.36	0/6604	0.61	0/8931
1	D	0.34	0/6612	0.60	0/8942
1	G	0.36	0/6603	0.60	0/8930
2	B	0.36	0/2737	0.60	0/3691
2	E	0.35	0/2693	0.61	0/3630
2	H	0.36	0/2737	0.62	1/3691 (0.0%)
3	C	0.36	0/2120	0.53	0/2868
3	F	0.37	0/2120	0.58	0/2868
3	I	0.35	0/2120	0.56	0/2868
All	All	0.35	0/34346	0.60	1/46419 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	H	876	ALA	N-CA-C	5.43	125.65	111.00

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	6487	6450	6451	52	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	6495	6462	6464	49	0
1	G	6486	6452	6454	56	0
2	B	2687	2697	2697	26	0
2	E	2646	2663	2663	25	0
2	H	2687	2697	2697	27	0
3	C	2063	2048	2048	8	0
3	F	2063	2048	2048	7	0
3	I	2063	2048	2048	6	0
4	A	35	0	0	1	0
4	B	10	0	0	0	0
4	C	15	0	0	0	0
4	D	50	0	0	0	0
4	E	10	0	0	1	0
4	F	20	0	0	0	0
4	G	40	0	0	0	0
4	H	20	0	0	0	0
4	I	15	0	0	0	0
5	B	27	20	0	0	0
5	E	27	20	0	1	0
5	H	27	20	0	0	0
All	All	33973	33625	33570	244	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 244 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)
3:F:231:TYR:OH	3:F:236:GLU:OE1	1.95	0.85
1:A:1022:THR:HB	1:A:1023:PRO:HD2	1.75	0.69
1:G:1047:TRP:HZ3	1:G:1132:VAL:HG13	1.61	0.64
2:E:867:LEU:HD21	2:E:1031:PRO:HG3	1.78	0.64
2:H:828:LEU:O	2:H:1034:PRO:HD2	1.98	0.63

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%)	801 (97%)	22 (3%)	0	100	100
1	G	822/840 (98%)	802 (98%)	20 (2%)	0	100	100
2	B	326/344 (95%)	315 (97%)	11 (3%)	0	100	100
2	E	322/344 (94%)	312 (97%)	10 (3%)	0	100	100
2	H	326/344 (95%)	317 (97%)	9 (3%)	0	100	100
3	C	246/271 (91%)	244 (99%)	2 (1%)	0	100	100
3	F	246/271 (91%)	244 (99%)	2 (1%)	0	100	100
3	I	246/271 (91%)	244 (99%)	2 (1%)	0	100	100
All	All	4179/4365 (96%)	4078 (98%)	101 (2%)	0	100	100

There are no Ramachandran outliers to report.

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%)	716 (99%)	5 (1%)	84	90
1	D	722/728 (99%)	719 (100%)	3 (0%)	91	94
1	G	721/728 (99%)	715 (99%)	6 (1%)	81	89
2	B	296/308 (96%)	292 (99%)	4 (1%)	67	81
2	E	292/308 (95%)	288 (99%)	4 (1%)	67	81

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles		
2	H	296/308 (96%)	294 (99%)	2 (1%)	84	90	
3	C	223/242 (92%)	220 (99%)	3 (1%)	69	82	
3	F	223/242 (92%)	219 (98%)	4 (2%)	59	77	
3	I	223/242 (92%)	221 (99%)	2 (1%)	78	87	
All	All	3717/3834 (97%)	3684 (99%)	33 (1%)	78	87	

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	1140	HIS
2	H	829	VAL
3	I	231	TYR
1	D	746	SER
3	C	257	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	999	HIS
1	D	261	HIS
2	H	999	HIS
1	G	370	GLN
2	B	944	GLN

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

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

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	1.70	1 (12%)	10,14,16	1.36	2 (20%)
2	TPO	B	893	2	8,10,11	1.19	0	10,14,16	1.49	2 (20%)
2	TPO	H	893	2	8,10,11	1.48	1 (12%)	10,14,16	1.13	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	-	6/9/11/13	-
2	TPO	B	893	2	-	0/9/11/13	-
2	TPO	H	893	2	-	1/9/11/13	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	893	TPO	P-O1P	3.45	1.61	1.50
2	H	893	TPO	P-OG1	2.51	1.64	1.59

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	893	TPO	P-OG1-CB	-3.14	113.72	123.21
2	E	893	TPO	P-OG1-CB	-2.46	115.79	123.21
2	E	893	TPO	CG2-CB-CA	-2.41	108.41	113.16
2	H	893	TPO	CG2-CB-CA	-2.39	108.45	113.16
2	B	893	TPO	CG2-CB-CA	-2.04	109.13	113.16

There are no chirality outliers.

5 of 7 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	O-C-CA-CB
2	E	893	TPO	CG2-CB-OG1-P

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 ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

46 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	SO4	D	1207	-	4,4,4	0.18	0	6,6,6	0.09	0
4	SO4	F	304	-	4,4,4	0.18	0	6,6,6	0.16	0
4	SO4	D	1206	-	4,4,4	0.22	0	6,6,6	0.08	0
4	SO4	A	1204	-	4,4,4	0.24	0	6,6,6	0.13	0
4	SO4	G	1202	-	4,4,4	0.19	0	6,6,6	0.04	0
4	SO4	G	1207	-	4,4,4	0.20	0	6,6,6	0.09	0
4	SO4	D	1201	-	4,4,4	0.19	0	6,6,6	0.05	0
4	SO4	G	1205	-	4,4,4	0.22	0	6,6,6	0.10	0
4	SO4	G	1204	-	4,4,4	0.21	0	6,6,6	0.07	0
4	SO4	D	1205	-	4,4,4	0.23	0	6,6,6	0.06	0
4	SO4	D	1204	-	4,4,4	0.25	0	6,6,6	0.05	0
4	SO4	A	1202	-	4,4,4	0.20	0	6,6,6	0.08	0
4	SO4	F	302	-	4,4,4	0.21	0	6,6,6	0.07	0
4	SO4	G	1208	-	4,4,4	0.25	0	6,6,6	0.06	0
4	SO4	C	303	-	4,4,4	0.22	0	6,6,6	0.13	0
4	SO4	D	1208	-	4,4,4	0.25	0	6,6,6	0.08	0
4	SO4	H	1103	-	4,4,4	0.25	0	6,6,6	0.07	0
4	SO4	G	1203	-	4,4,4	0.19	0	6,6,6	0.08	0
4	SO4	D	1209	-	4,4,4	0.27	0	6,6,6	0.06	0
4	SO4	A	1201	-	4,4,4	0.21	0	6,6,6	0.05	0
4	SO4	A	1206	-	4,4,4	0.20	0	6,6,6	0.08	0
4	SO4	H	1105	-	4,4,4	0.23	0	6,6,6	0.07	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
4	SO4	D	1210	-	4,4,4	0.19	0	6,6,6	0.08	0
4	SO4	G	1201	-	4,4,4	0.21	0	6,6,6	0.07	0
4	SO4	D	1202	-	4,4,4	0.20	0	6,6,6	0.07	0
4	SO4	A	1207	-	4,4,4	0.19	0	6,6,6	0.09	0
4	SO4	E	1103	-	4,4,4	0.22	0	6,6,6	0.08	0
4	SO4	F	301	-	4,4,4	0.22	0	6,6,6	0.10	0
4	SO4	B	1102	-	4,4,4	0.24	0	6,6,6	0.08	0
4	SO4	H	1104	-	4,4,4	0.23	0	6,6,6	0.07	0
4	SO4	A	1205	-	4,4,4	0.22	0	6,6,6	0.07	0
4	SO4	A	1203	-	4,4,4	0.22	0	6,6,6	0.12	0
4	SO4	B	1103	-	4,4,4	0.23	0	6,6,6	0.07	0
4	SO4	D	1203	-	4,4,4	0.19	0	6,6,6	0.07	0
4	SO4	C	301	-	4,4,4	0.20	0	6,6,6	0.05	0
4	SO4	H	1102	-	4,4,4	0.22	0	6,6,6	0.05	0
5	RVU	B	1101	-	30,30,30	0.76	0	32,43,43	1.06	3 (9%)
5	RVU	H	1101	-	30,30,30	0.75	0	32,43,43	1.09	3 (9%)
4	SO4	F	303	-	4,4,4	0.16	0	6,6,6	0.06	0
4	SO4	I	303	-	4,4,4	0.22	0	6,6,6	0.12	0
4	SO4	E	1102	-	4,4,4	0.25	0	6,6,6	0.09	0
5	RVU	E	1101	-	30,30,30	0.73	0	32,43,43	1.11	2 (6%)
4	SO4	I	301	-	4,4,4	0.22	0	6,6,6	0.06	0
4	SO4	C	302	-	4,4,4	0.23	0	6,6,6	0.07	0
4	SO4	G	1206	-	4,4,4	0.25	0	6,6,6	0.07	0
4	SO4	I	302	-	4,4,4	0.16	0	6,6,6	0.09	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
5	RVU	B	1101	-	-	2/12/26/26	0/4/4/4
5	RVU	H	1101	-	-	2/12/26/26	0/4/4/4
5	RVU	E	1101	-	-	2/12/26/26	0/4/4/4

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	E	1101	RVU	C3-N1-C2	-3.22	125.41	130.28

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	H	1101	RVU	C3-N1-C2	-2.68	126.22	130.28
5	E	1101	RVU	C18-C17-C15	-2.37	116.68	120.08
5	B	1101	RVU	C3-N1-C2	-2.30	126.80	130.28
5	H	1101	RVU	C18-C17-C15	-2.29	116.79	120.08

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	1101	RVU	C4-C1-C2-N1
5	B	1101	RVU	C4-C1-C2-O3
5	E	1101	RVU	C4-C1-C2-N1
5	E	1101	RVU	C4-C1-C2-O3
5	H	1101	RVU	C4-C1-C2-N1

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1203	SO4	1	0
4	E	1102	SO4	1	0
5	E	1101	RVU	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 RVU B 1101

Bond lengths

Bond angles

Torsions

Rings

Ligand RVU H 1101

Bond lengths

Bond angles

Torsions

Rings

Ligand RVU E 1101

Bond lengths

Bond angles

Torsions

Rings
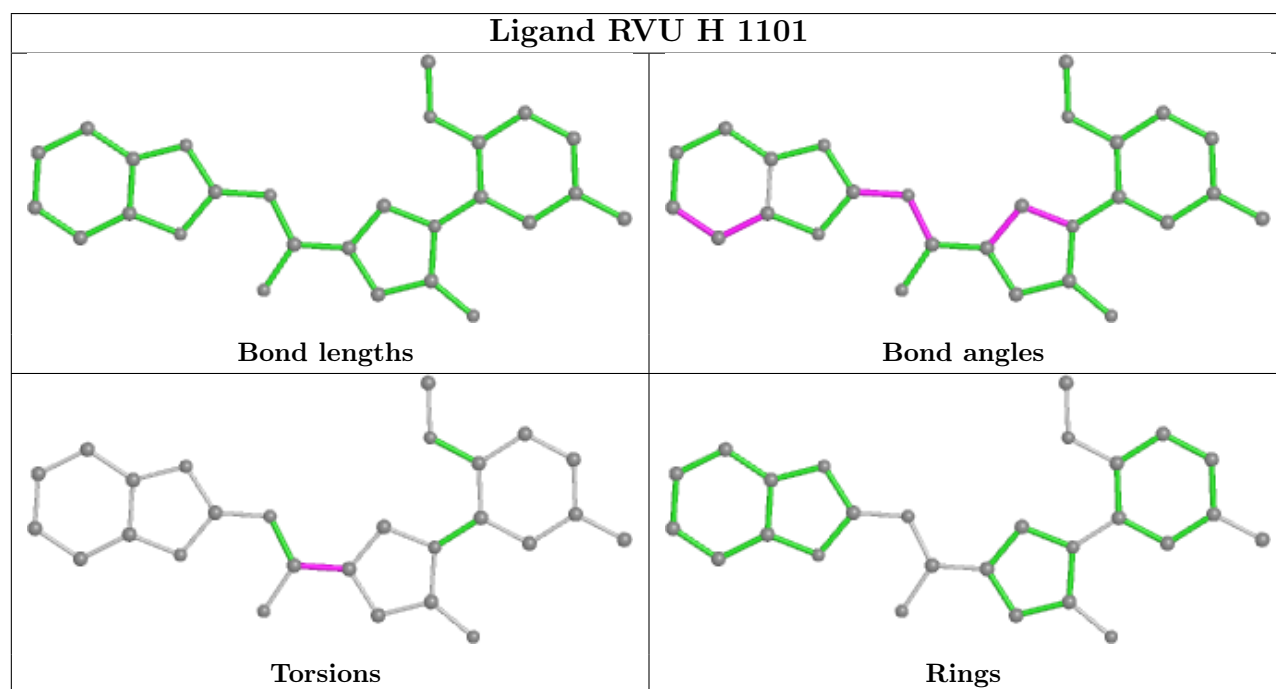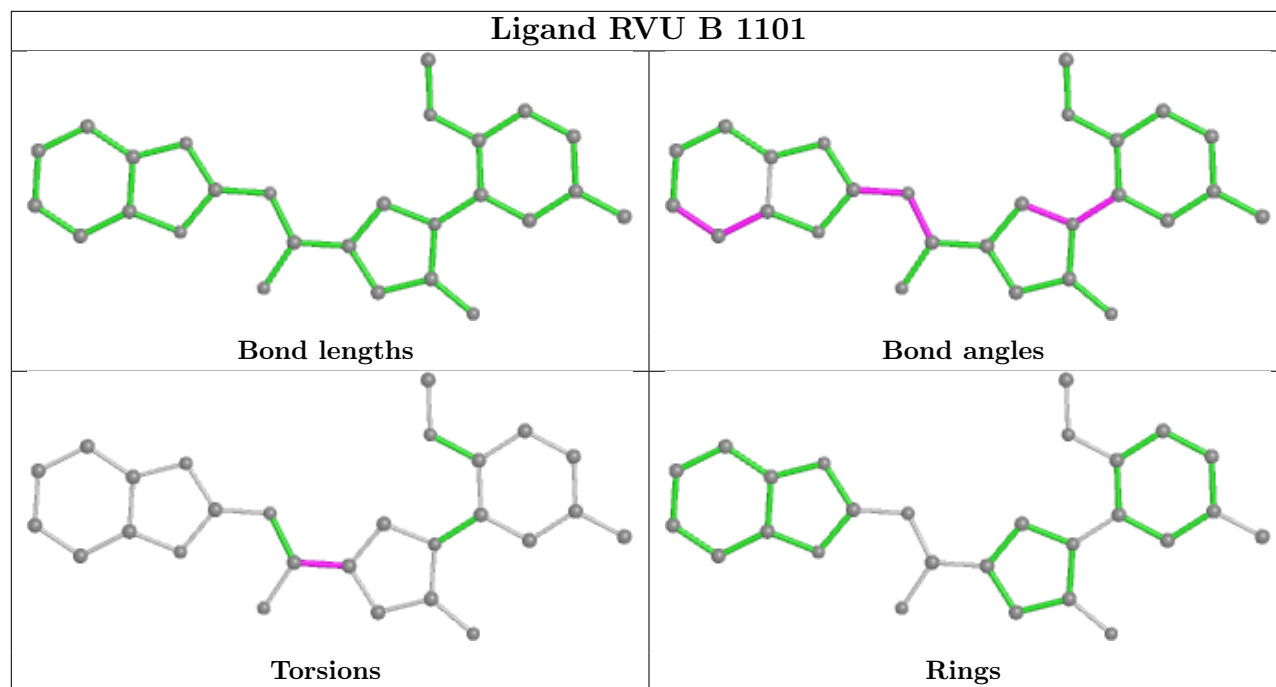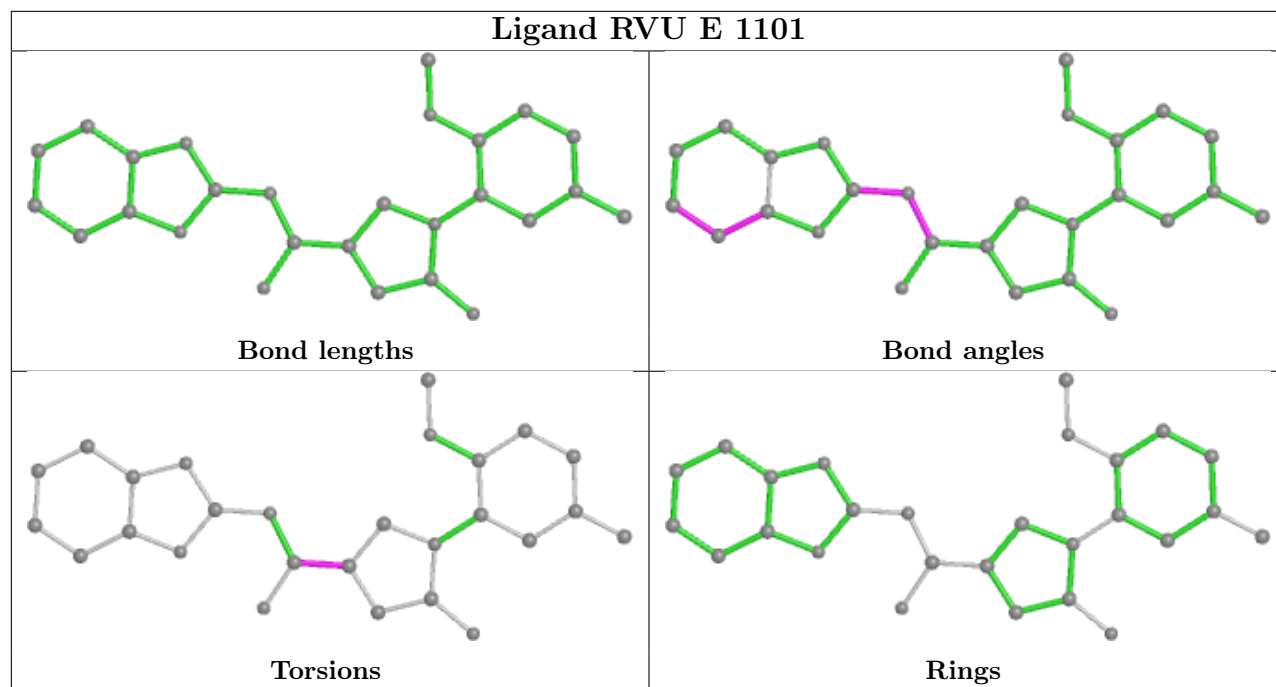
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	826/840 (98%)	0.63	81 (9%)	7	7	105, 141, 210, 292	0
1	D	827/840 (98%)	0.55	70 (8%)	10	10	96, 135, 201, 269	0
1	G	826/840 (98%)	0.55	74 (8%)	9	10	103, 142, 213, 308	0
2	B	328/344 (95%)	0.77	52 (15%)	1	2	112, 146, 215, 239	0
2	E	324/344 (94%)	0.71	29 (8%)	9	10	105, 137, 203, 240	0
2	H	328/344 (95%)	0.76	30 (9%)	9	10	91, 122, 199, 270	0
3	C	248/271 (91%)	0.60	16 (6%)	18	18	109, 135, 173, 246	0
3	F	248/271 (91%)	0.62	10 (4%)	38	35	92, 118, 155, 229	0
3	I	248/271 (91%)	0.64	15 (6%)	21	21	97, 121, 162, 236	0
All	All	4203/4365 (96%)	0.62	377 (8%)	9	10	91, 136, 206, 308	0

The worst 5 of 377 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1016	ASN	10.1
2	H	1039	CYS	7.6
2	H	1043	TRP	7.5
1	A	1017	LEU	6.5
2	B	1035	HIS	6.4

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, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	TPO	E	893	11/12	0.87	0.15	183,196,237,239	6
2	TPO	H	893	11/12	0.87	0.24	141,159,191,194	6
2	TPO	B	893	11/12	0.89	0.15	188,204,247,248	6

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	SO4	I	303	5/5	0.68	0.19	161,162,166,192	0
4	SO4	D	1205	5/5	0.73	0.13	159,162,177,199	0
4	SO4	D	1207	5/5	0.74	0.26	146,154,156,189	0
4	SO4	E	1102	5/5	0.75	0.21	163,172,176,208	0
4	SO4	D	1206	5/5	0.76	0.29	146,157,159,198	0
4	SO4	C	302	5/5	0.77	0.14	154,161,170,190	0
4	SO4	F	304	5/5	0.78	0.19	152,153,156,182	0
4	SO4	A	1205	5/5	0.78	0.26	163,170,174,206	0
4	SO4	F	301	5/5	0.79	0.44	144,161,163,193	0
4	SO4	D	1208	5/5	0.81	0.17	162,163,168,196	0
4	SO4	H	1105	5/5	0.81	0.16	156,162,169,200	0
4	SO4	F	302	5/5	0.81	0.14	157,162,166,193	0
4	SO4	D	1203	5/5	0.82	0.23	141,153,161,193	0
4	SO4	G	1207	5/5	0.83	0.20	149,157,163,195	0
4	SO4	C	303	5/5	0.83	0.11	161,169,170,194	0
4	SO4	D	1204	5/5	0.83	0.20	153,156,174,194	0
4	SO4	G	1208	5/5	0.84	0.37	168,170,176,201	0
4	SO4	G	1206	5/5	0.84	0.09	177,177,183,211	0
4	SO4	G	1204	5/5	0.84	0.14	150,160,164,190	0
4	SO4	A	1204	5/5	0.85	0.16	175,176,179,204	0
4	SO4	D	1210	5/5	0.86	0.16	141,144,154,163	0
4	SO4	G	1201	5/5	0.86	0.17	144,147,151,162	0
4	SO4	A	1201	5/5	0.87	0.16	147,153,155,171	0
4	SO4	A	1202	5/5	0.87	0.16	140,147,147,173	0
4	SO4	A	1203	5/5	0.88	0.30	141,151,170,188	0
4	SO4	D	1202	5/5	0.88	0.19	140,152,161,184	0

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

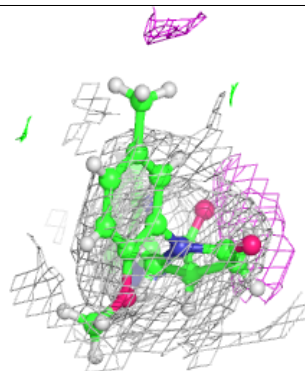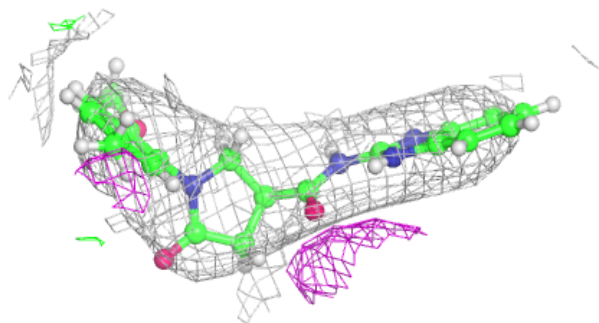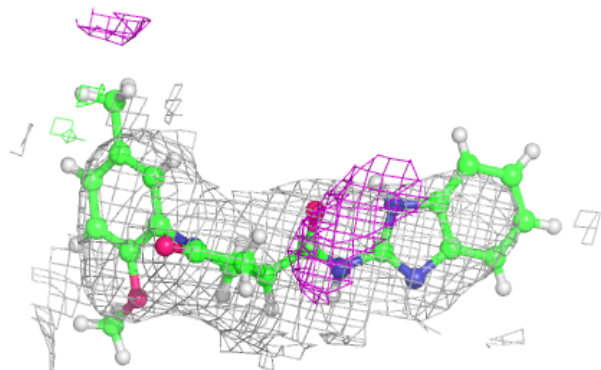| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	SO4	D	1201	5/5	0.89	0.18	135,140,149,151	0
4	SO4	D	1209	5/5	0.89	0.21	180,182,187,212	0
4	SO4	G	1203	5/5	0.89	0.19	132,138,153,163	0
4	SO4	C	301	5/5	0.90	0.23	137,139,150,158	0
4	SO4	A	1206	5/5	0.90	0.15	144,148,152,167	0
4	SO4	G	1205	5/5	0.90	0.17	155,170,181,201	0
4	SO4	A	1207	5/5	0.90	0.14	134,153,171,184	0
4	SO4	G	1202	5/5	0.91	0.14	146,150,156,177	0
4	SO4	H	1103	5/5	0.91	0.10	158,161,177,199	0
4	SO4	E	1103	5/5	0.92	0.12	158,164,167,194	0
4	SO4	B	1102	5/5	0.92	0.12	165,169,173,200	0
4	SO4	I	301	5/5	0.92	0.13	147,154,168,185	0
4	SO4	H	1102	5/5	0.92	0.16	164,164,179,195	0
4	SO4	I	302	5/5	0.93	0.20	124,124,146,147	0
4	SO4	H	1104	5/5	0.93	0.11	150,156,171,190	0
5	RVU	H	1101	27/27	0.94	0.51	102,122,147,148	20
5	RVU	B	1101	27/27	0.95	0.47	111,130,156,159	20
5	RVU	E	1101	27/27	0.95	0.48	104,122,148,158	20
4	SO4	B	1103	5/5	0.95	0.11	163,164,175,200	0
4	SO4	F	303	5/5	0.96	0.17	116,120,136,138	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.
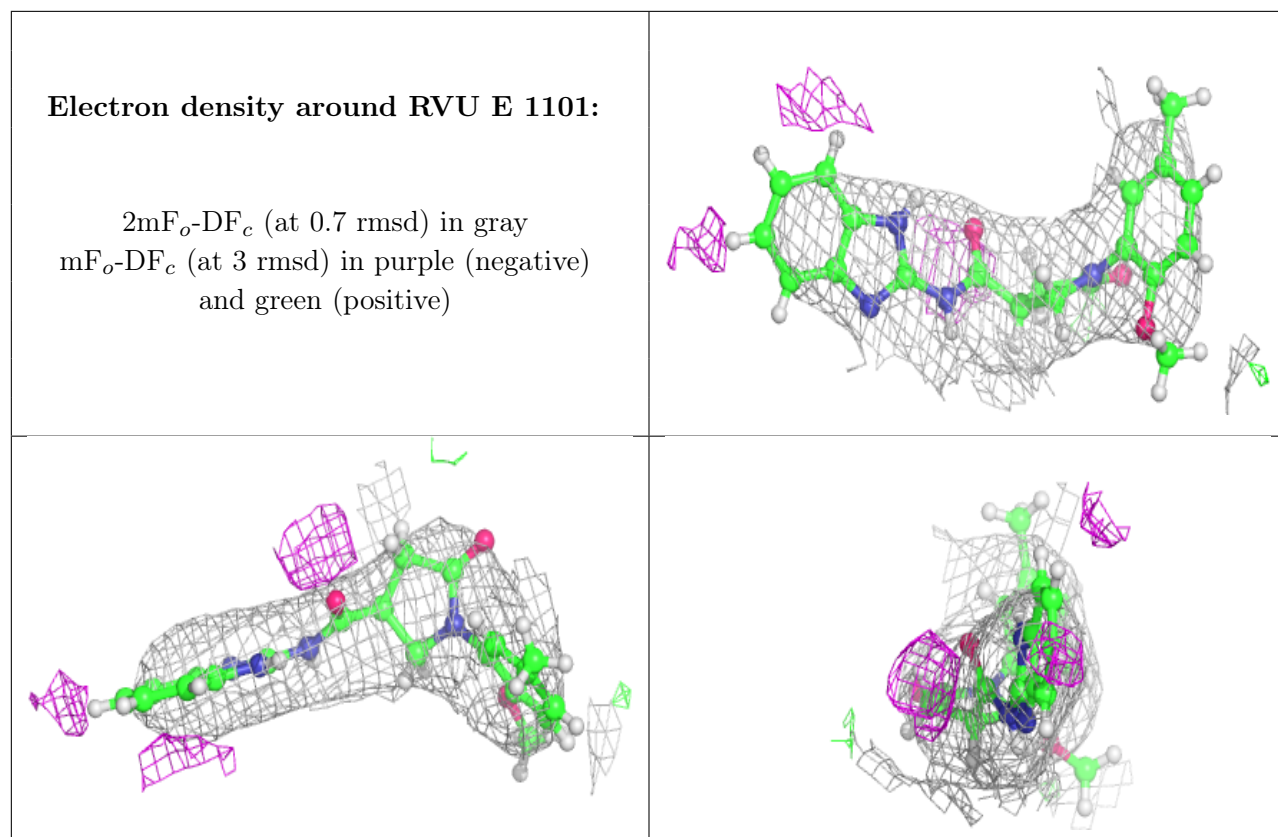
Electron density around RVU H 1101:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
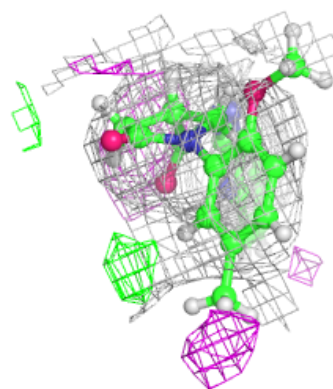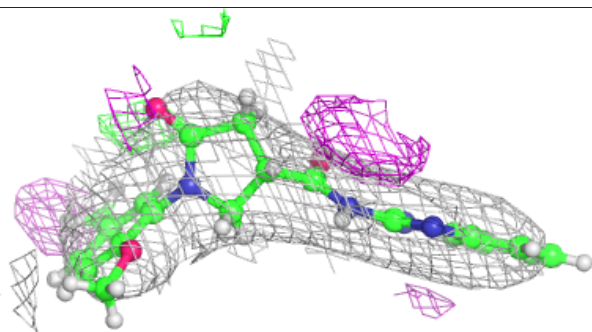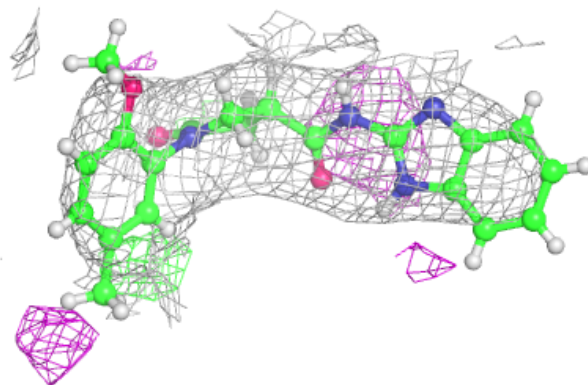
Electron density around RVU B 1101:

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

Electron density around RVU 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.