



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 26, 2026 – 11:50 AM EDT

PDB ID : 8BUC / pdb\_00008buc  
Title : Structure of DDB1 bound to dCeMM3-engaged CDK12-cyclin K  
Authors : Kozicka, Z.; Kempf, G.; Focht, V.; Thoma, N.H.  
Deposited on : 2022-11-30  
Resolution : 3.85 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Buster-report : wwPDB partial adaption of 1.1.7 (2018)  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

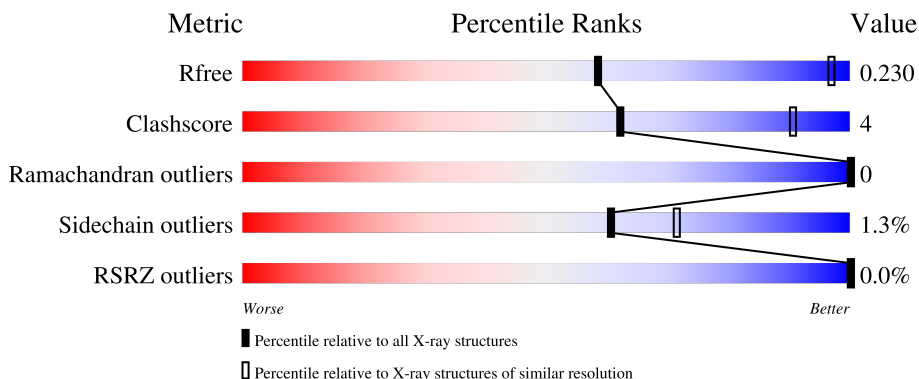
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.85 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))
$R_{free}$	180053	1165 (4.02-3.70)
Clashscore	190562	1207 (4.02-3.70)
Ramachandran outliers	187476	1149 (4.02-3.70)
Sidechain outliers	187428	1142 (4.02-3.70)
RSRZ outliers	180081	1164 (4.02-3.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	840	
1	D	840	
1	G	840	
2	B	344	
2	E	344	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	344	 79% 15% ••
3	C	271	 86% 6% 8% •
3	F	271	 88% • 8%
3	I	271	 88% • 8%

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 67386 atoms, of which 33591 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
			Total	C	H	N	O	S			
1	A	826	12936	4105	6449	1094	1252	36	6449	0	0
1	D	827	12957	4111	6462	1095	1253	36	6462	0	0
1	G	826	12938	4106	6452	1093	1251	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	327	Total	C	H	N	O	P	S	2676	0	0
			5341	1706	2676	451	490	1	17			
2	E	325	Total	C	H	N	O	P	S	2663	0	0
			5309	1695	2663	447	486	1	17			
2	H	330	Total	C	H	N	O	P	S	2702	0	0
			5395	1726	2702	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:

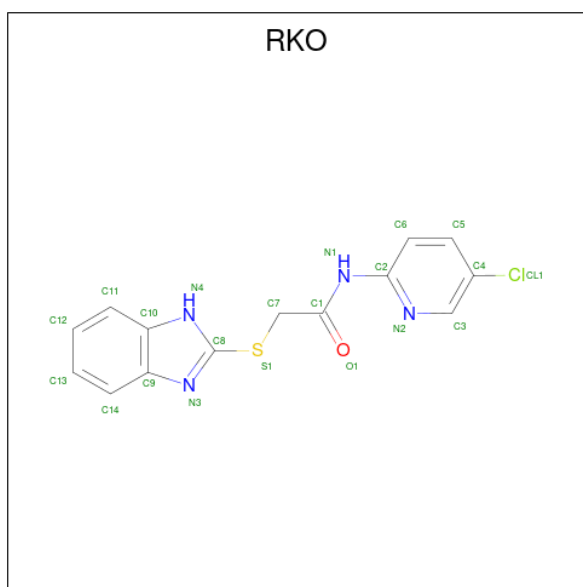
Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	GLY	-	expression tag	UNP O75909
C	-2	GLY	-	expression tag	UNP O75909
C	-1	GLY	-	expression tag	UNP O75909
C	0	ARG	-	expression tag	UNP O75909
F	-3	GLY	-	expression tag	UNP O75909
F	-2	GLY	-	expression tag	UNP O75909
F	-1	GLY	-	expression tag	UNP O75909
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 SULFATE ION (CCD ID: SO4) (formula: O<sub>4</sub>S).



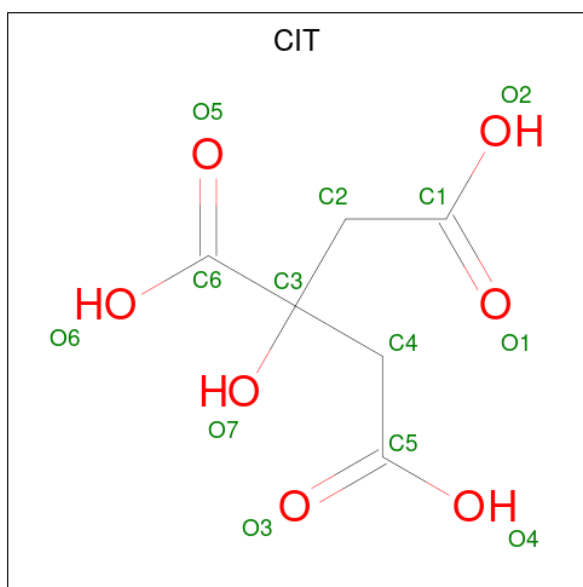
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	C	1	Total O S 5 4 1	0	0
4	D	1	Total O S 5 4 1	0	0
4	G	1	Total O S 5 4 1	0	0
4	G	1	Total O S 5 4 1	0	0
4	G	1	Total O S 5 4 1	0	0
4	I	1	Total O S 5 4 1	0	0

- Molecule 5 is 2-(1 {H}-benzimidazol-2-ylsulfanyl)- {N}-(5-chloranylpyridin-2-yl)ethanamide (CCD ID: RKO) (formula: C<sub>14</sub>H<sub>11</sub>ClN<sub>4</sub>OS) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	
			Total	C	Cl	H	N	O			S
5	B	1	Total	C	Cl	H	N	O	S	11	0
			32	14	1	11	4	1	1		
5	E	1	Total	C	Cl	H	N	O	S	11	0
			32	14	1	11	4	1	1		
5	H	1	Total	C	Cl	H	N	O	S	11	0
			32	14	1	11	4	1	1		

- Molecule 6 is CITRIC ACID (CCD ID: CIT) (formula:  $C_6H_8O_7$ ).



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

*Continued on next page...*

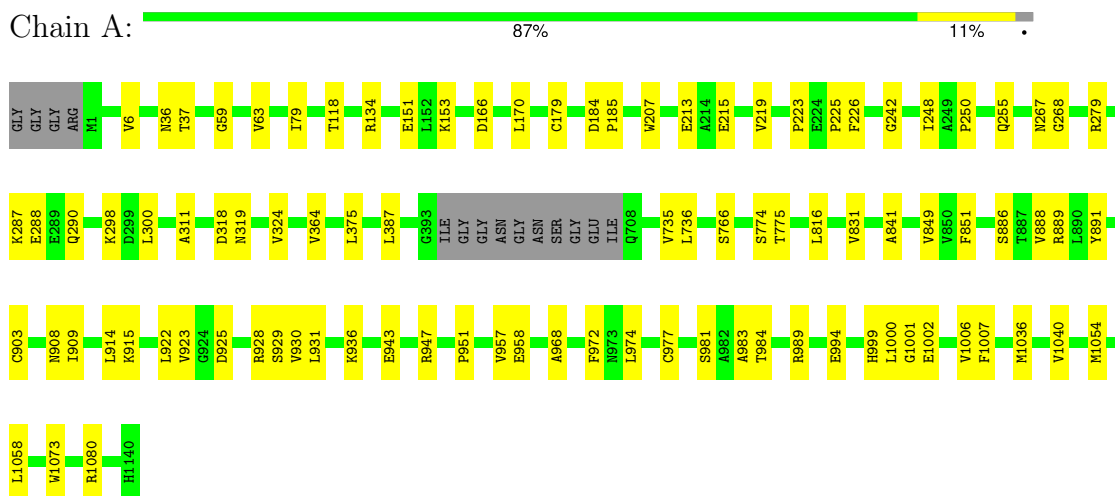
*Continued from previous page...*

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

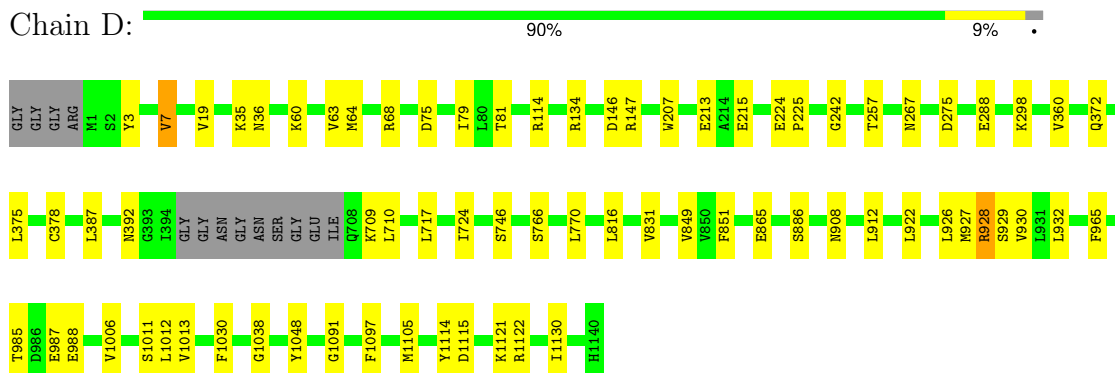
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

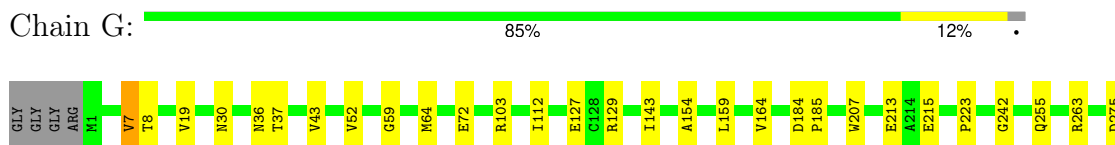
- Molecule 1: DNA damage-binding protein 1

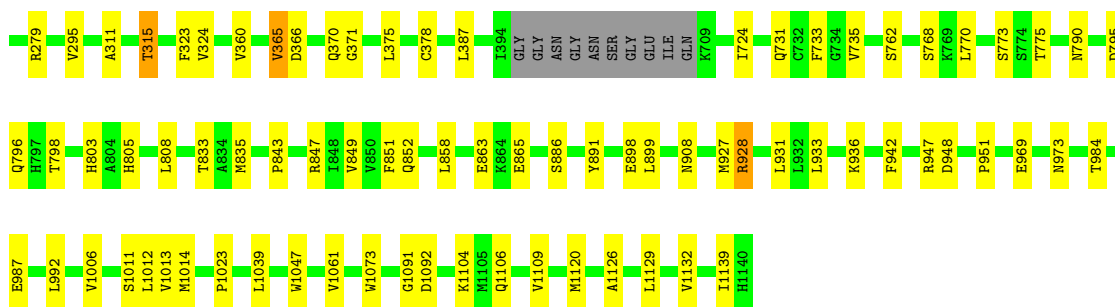


- Molecule 1: DNA damage-binding protein 1



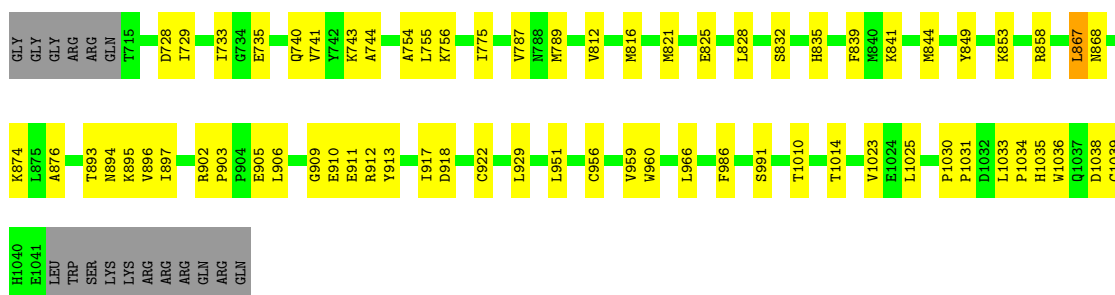
- Molecule 1: DNA damage-binding protein 1





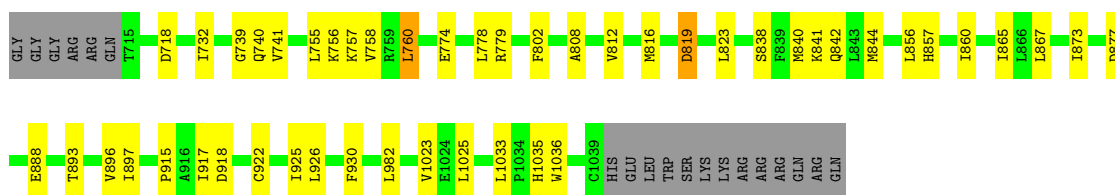
- Molecule 2: Cyclin-dependent kinase 12

Chain B: 75% 19% 5%



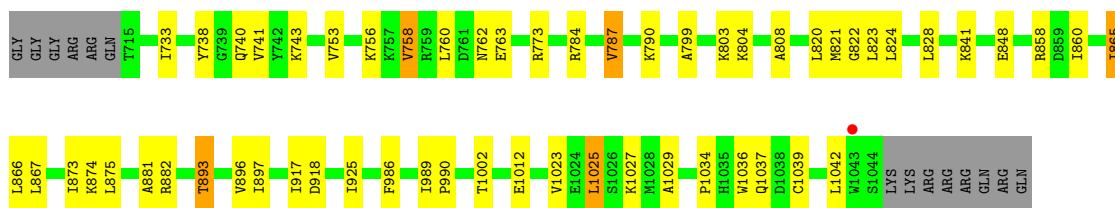
- Molecule 2: Cyclin-dependent kinase 12

Chain E: 81% 13% 6%



- Molecule 2: Cyclin-dependent kinase 12

Chain H: 79% 15% 6%

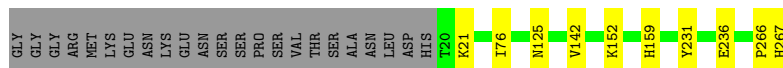
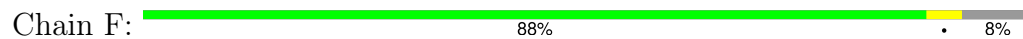


- Molecule 3: Cyclin-K

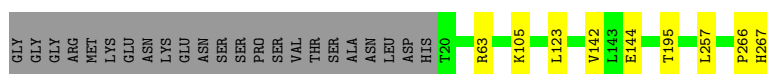
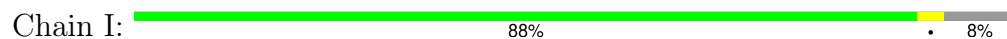
Chain C: 86% 6% 8%



- 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, $\alpha$ , $\beta$ , $\gamma$	249.77Å 249.77Å 218.03Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	76.78 – 3.85 76.78 – 3.85	Depositor EDS
% Data completeness (in resolution range)	93.4 (76.78-3.85) 93.4 (76.78-3.85)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.53 (at 3.89Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.181 , 0.224 0.191 , 0.230	Depositor DCC
$R_{free}$ test set	3442 reflections (4.63%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	130.4	Xtrriage
Anisotropy	0.021	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 142.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.41$ , $\langle L^2 \rangle = 0.23$	Xtrriage
Estimated twinning fraction	0.077 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	67386	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	161.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.54% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, TPO, RKO, CIT

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	0.23	0/6604	0.50	0/8931
1	D	0.23	0/6612	0.49	0/8942
1	G	0.24	0/6603	0.49	0/8930
2	B	0.23	0/2713	0.48	1/3657 (0.0%)
2	E	0.24	0/2693	0.49	0/3630
2	H	0.22	0/2743	0.40	0/3699
3	C	0.21	0/2120	0.42	0/2868
3	F	0.22	0/2120	0.46	0/2868
3	I	0.22	0/2120	0.44	0/2868
All	All	0.23	0/34328	0.47	1/46393 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	B	876	ALA	N-CA-C	5.22	121.91	110.80

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	6449	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	43	0
1	G	6486	6452	6454	67	0
2	B	2665	2676	2676	38	0
2	E	2646	2663	2663	32	0
2	H	2693	2702	2702	37	0
3	C	2063	2048	2048	10	0
3	F	2063	2048	2048	8	0
3	I	2063	2048	2048	6	0
4	A	15	0	0	0	0
4	C	5	0	0	0	0
4	D	5	0	0	0	0
4	G	15	0	0	0	0
4	I	5	0	0	0	0
5	B	21	11	0	2	0
5	E	21	11	0	0	0
5	H	21	11	0	0	0
6	D	13	5	5	0	0
6	F	13	5	5	0	0
All	All	33795	33591	33564	278	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 278 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	1.88	0.90
2:H:803:LYS:HA	3:I:142:VAL:HG11	1.55	0.88
1:A:974:LEU:HD11	1:A:1000:LEU:HD22	1.75	0.69
1:A:922:LEU:HD11	1:A:930:VAL:HB	1.75	0.68
1:D:927:MET:HE2	1:D:928:ARG:NH1	2.10	0.67

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%)	803 (98%)	20 (2%)	0	100	100
1	G	822/840 (98%)	803 (98%)	19 (2%)	0	100	100
2	B	324/344 (94%)	313 (97%)	11 (3%)	0	100	100
2	E	322/344 (94%)	312 (97%)	10 (3%)	0	100	100
2	H	327/344 (95%)	317 (97%)	10 (3%)	0	100	100
3	C	246/271 (91%)	241 (98%)	5 (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	4178/4365 (96%)	4072 (98%)	106 (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%)	76	78
1	D	722/728 (99%)	716 (99%)	6 (1%)	73	77
1	G	721/728 (99%)	712 (99%)	9 (1%)	63	72
2	B	294/308 (96%)	284 (97%)	10 (3%)	32	56
2	E	292/308 (95%)	285 (98%)	7 (2%)	43	63
2	H	297/308 (96%)	286 (96%)	11 (4%)	30	54
3	C	223/242 (92%)	222 (100%)	1 (0%)	84	83
3	F	223/242 (92%)	223 (100%)	0	100	100
3	I	223/242 (92%)	223 (100%)	0	100	100
All	All	3716/3834 (97%)	3667 (99%)	49 (1%)	61	71

5 of 49 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	1036	TRP
1	G	928	ARG
1	G	7	VAL
1	G	360	VAL
1	G	1120	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 44 such sidechains are listed below:

Mol	Chain	Res	Type
2	E	939	ASN
1	G	826	ASN
2	E	977	GLN
1	G	290	GLN
1	G	877	ASN

### 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	1.72	2 (25%)	10,14,16	1.32	1 (10%)
2	TPO	B	893	2	8,10,11	1.51	1 (12%)	10,14,16	1.51	1 (10%)
2	TPO	H	893	2	8,10,11	1.68	1 (12%)	10,14,16	1.19	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	B	893	2	-	0/9/11/13	-
2	TPO	H	893	2	-	0/9/11/13	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	893	TPO	P-O1P	3.40	1.61	1.50
2	B	893	TPO	P-O1P	3.26	1.60	1.50
2	E	893	TPO	P-O1P	3.19	1.60	1.50
2	E	893	TPO	P-OG1	2.09	1.63	1.59

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	893	TPO	P-OG1-CB	-4.09	112.21	123.33
2	E	893	TPO	P-OG1-CB	-2.52	116.48	123.33
2	H	893	TPO	P-OG1-CB	-2.36	116.91	123.33

There are no chirality outliers.

All (4) 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

There are no ring outliers.

1 monomer is involved in 1 short contact:

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

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry

14 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	CIT	D	1201	-	12,12,12	1.19	0	17,17,17	1.39	1 (5%)
4	SO4	G	1203	-	4,4,4	0.34	0	6,6,6	0.12	0
5	RKO	B	1101	-	23,23,23	0.47	0	31,31,31	0.58	0
4	SO4	I	301	-	4,4,4	0.31	0	6,6,6	0.09	0
5	RKO	H	1101	-	23,23,23	0.52	0	31,31,31	0.67	1 (3%)
5	RKO	E	1101	-	23,23,23	0.45	0	31,31,31	0.62	0
4	SO4	A	1201	-	4,4,4	0.30	0	6,6,6	0.09	0
4	SO4	A	1203	-	4,4,4	0.38	0	6,6,6	0.12	0
4	SO4	D	1202	-	4,4,4	0.28	0	6,6,6	0.05	0
6	CIT	F	301	-	12,12,12	1.15	0	17,17,17	1.51	2 (11%)
4	SO4	C	301	-	4,4,4	0.32	0	6,6,6	0.08	0
4	SO4	G	1201	-	4,4,4	0.27	0	6,6,6	0.11	0
4	SO4	A	1202	-	4,4,4	0.28	0	6,6,6	0.10	0
4	SO4	G	1202	-	4,4,4	0.34	0	6,6,6	0.05	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
6	CIT	D	1201	-	-	2/16/16/16	-
5	RKO	H	1101	-	-	1/9/9/9	0/3/3/3
5	RKO	E	1101	-	-	2/9/9/9	0/3/3/3
6	CIT	F	301	-	-	3/16/16/16	-
5	RKO	B	1101	-	-	0/9/9/9	0/3/3/3

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	F	301	CIT	O6-C6-C3	4.13	121.05	113.14
6	D	1201	CIT	O6-C6-C3	3.36	119.58	113.14
6	F	301	CIT	O4-C5-C4	2.19	121.28	114.35
5	H	1101	RKO	C9-N3-C8	-2.06	103.77	108.66

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

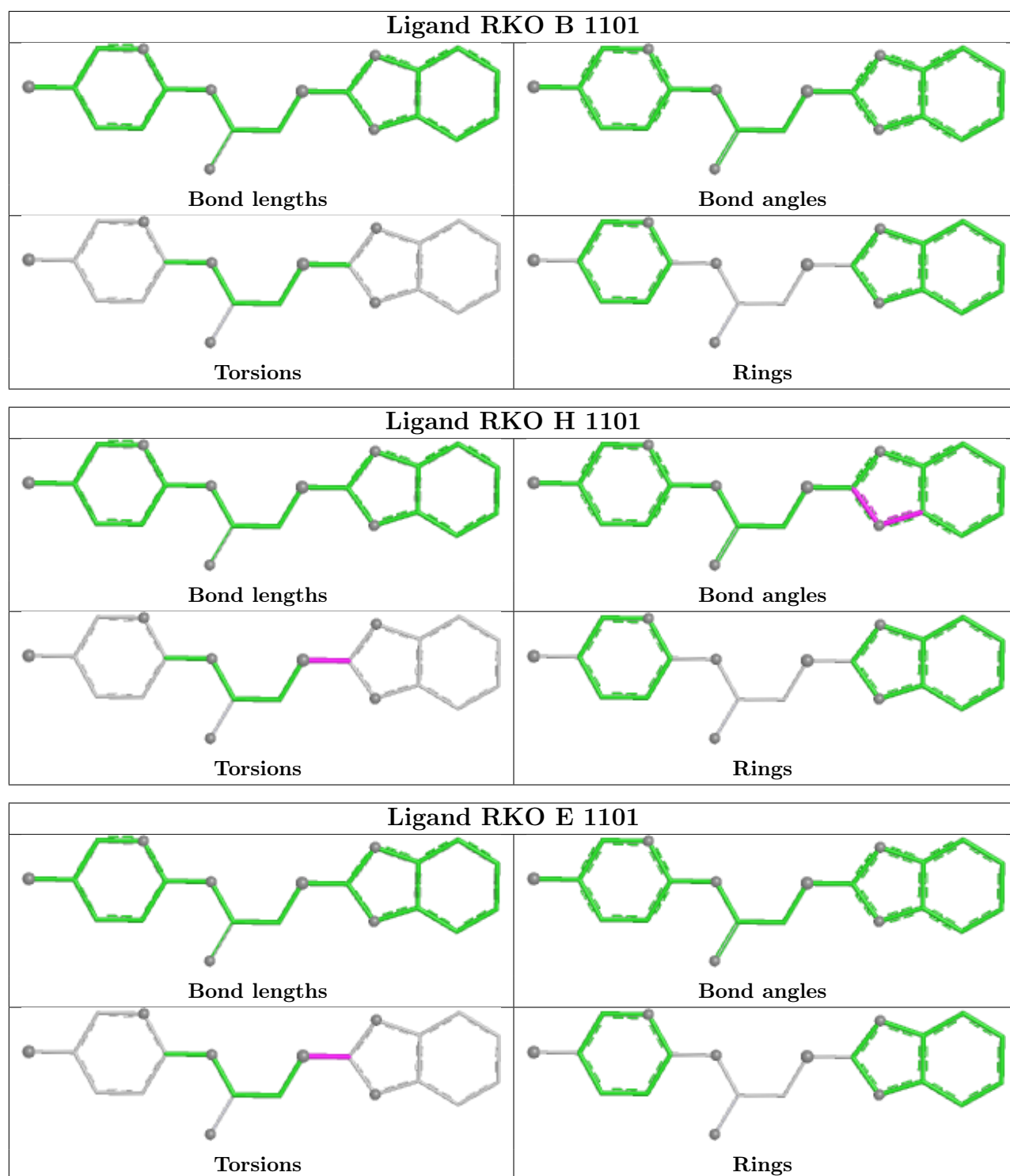
Mol	Chain	Res	Type	Atoms
6	F	301	CIT	C2-C3-C4-C5
6	F	301	CIT	C6-C3-C4-C5
6	D	1201	CIT	C1-C2-C3-C6
6	D	1201	CIT	C1-C2-C3-O7
5	E	1101	RKO	N4-C8-S1-C7

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	1101	RKO	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 5.7 Other polymers [\(i\)](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	826/840 (98%)	-0.65	0 100 100	42, 74, 125, 201	0
1	D	827/840 (98%)	-0.68	0 100 100	39, 72, 122, 190	0
1	G	826/840 (98%)	-0.64	0 100 100	44, 75, 127, 201	0
2	B	326/344 (94%)	-0.58	0 100 100	47, 79, 133, 192	0
2	E	324/344 (94%)	-0.61	0 100 100	45, 75, 125, 194	0
2	H	329/344 (95%)	-0.58	1 (0%) 90 77	38, 66, 122, 194	0
3	C	248/271 (91%)	-0.70	0 100 100	42, 70, 113, 150	0
3	F	248/271 (91%)	-0.73	0 100 100	39, 60, 103, 133	0
3	I	248/271 (91%)	-0.79	0 100 100	39, 63, 102, 140	0
All	All	4202/4365 (96%)	-0.65	1 (0%) 100 100	38, 72, 123, 201	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	1043	TRP	2.0

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

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	TPO	E	893	11/12	0.61	0.08	236,253,321,321	6
2	TPO	B	893	11/12	0.72	0.09	240,245,296,296	6
2	TPO	H	893	11/12	0.80	0.13	164,198,242,242	6

### 6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

### 6.4 Ligands [i](#)

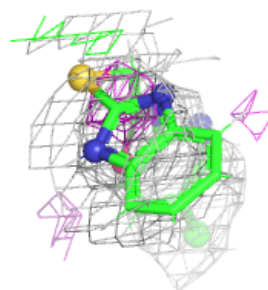
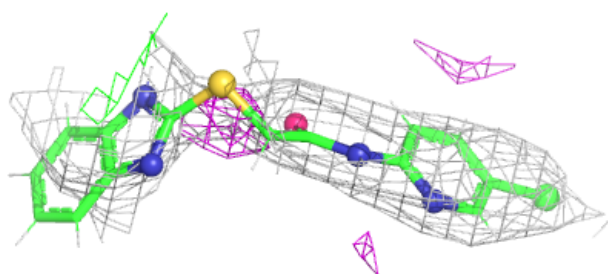
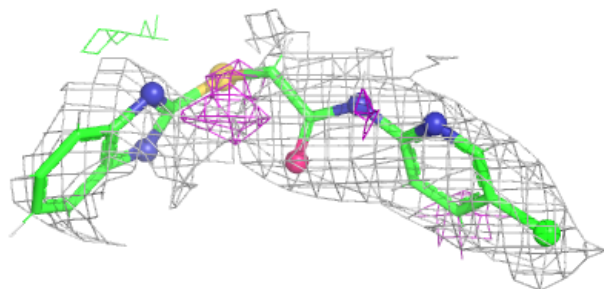
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	SO4	A	1203	5/5	0.78	0.05	146,161,214,219	0
4	SO4	C	301	5/5	0.78	0.10	163,166,172,191	0
4	SO4	I	301	5/5	0.78	0.14	132,158,177,178	0
6	CIT	D	1201	13/13	0.78	0.07	123,163,218,218	5
6	CIT	F	301	13/13	0.78	0.07	141,161,183,211	5
4	SO4	G	1203	5/5	0.79	0.10	176,177,192,217	0
4	SO4	G	1202	5/5	0.89	0.10	148,158,182,195	0
5	RKO	B	1101	21/21	0.90	0.10	104,156,198,220	11
5	RKO	H	1101	21/21	0.90	0.10	102,148,213,215	11
5	RKO	E	1101	21/21	0.91	0.10	97,146,221,228	11
4	SO4	A	1202	5/5	0.91	0.08	157,172,185,208	0
4	SO4	G	1201	5/5	0.92	0.08	141,142,162,167	0
4	SO4	D	1202	5/5	0.94	0.06	133,142,164,166	0
4	SO4	A	1201	5/5	0.94	0.05	136,144,164,167	0

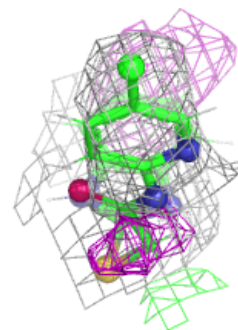
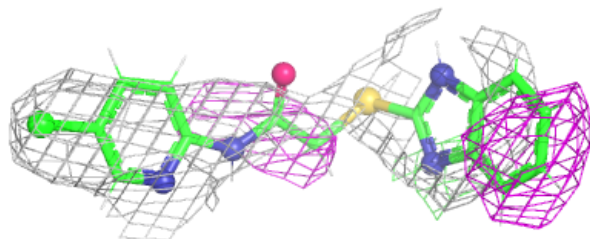
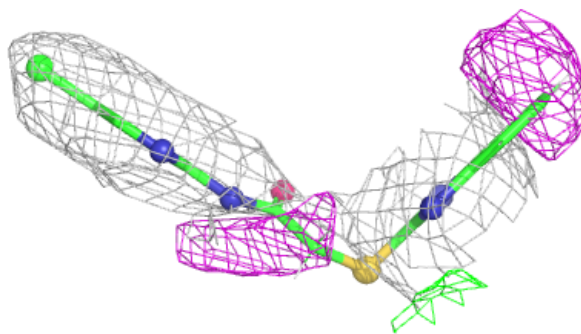
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

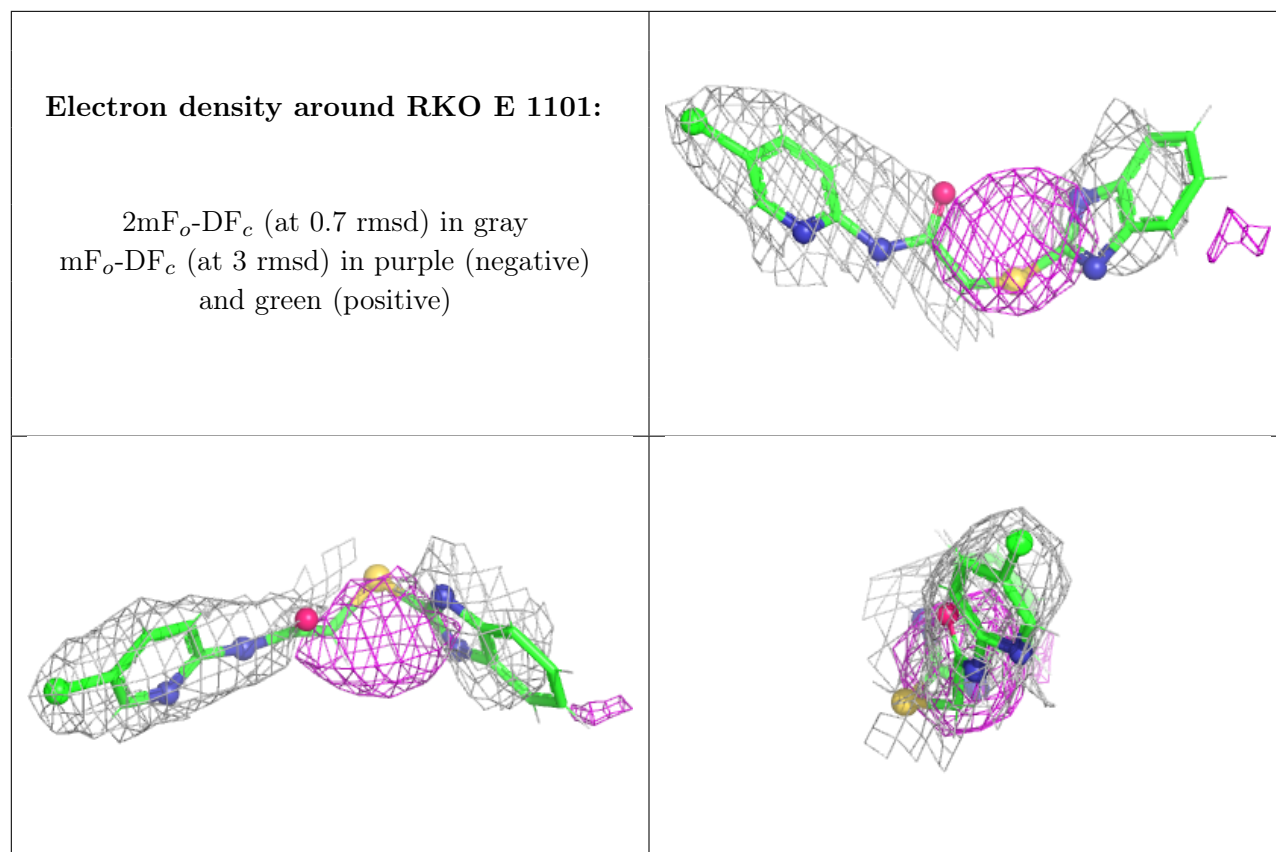
**Electron density around RKO B 1101:**

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

**Electron density around RKO 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)





## 6.5 Other polymers [i](#)

There are no such residues in this entry.