wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 23, 2024 – 01:29 PM EDT

PDB ID	:	5ACB
Title	:	Crystal Structure of the Human Cdk12-Cyclink Complex
Authors	:	Dixon Clarke, S.E.; Elkins, J.M.; Pike, A.C.W.; Mackenzie, A.; Goubin, S.; Strain-Damerell, C.; Mahajan, P.; Tallant, C.; Chalk, R.; Wiggers, H.; Kopec, J.; Fitzpatrick, F.; Burgess-Brown, N.; Carpenter, E.P.; von Delft, F.; Arrowsmith, C.H.; Edwards, A.M.; Bountra, C.; Bullock, A.
Deposited on	:	2015-08-14
Resolution	:	2.70 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 ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

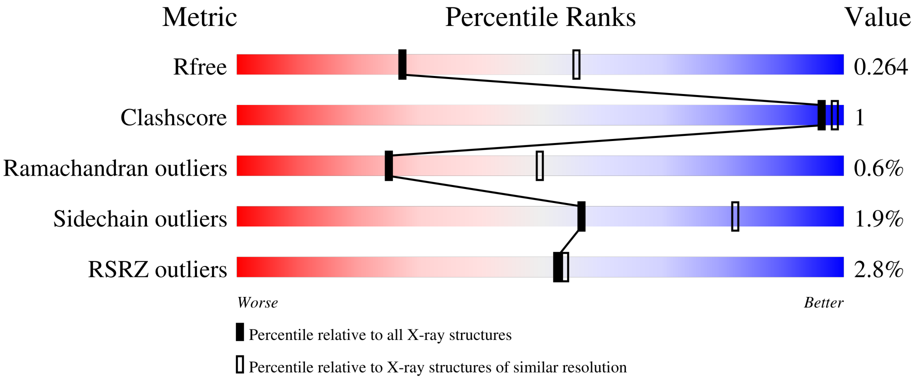
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance ⓘ

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

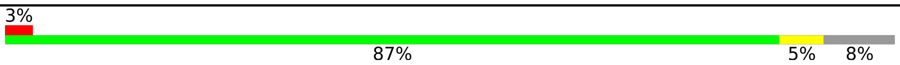
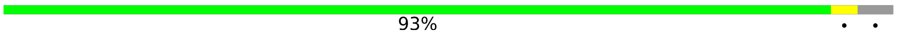
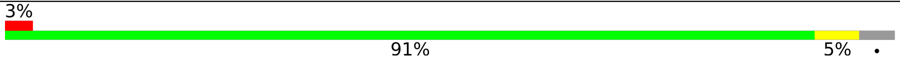
The reported resolution of this entry is 2.70 Å.

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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 >=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	259	
1	B	259	
2	C	340	
2	D	340	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9073 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called CYCLIN-K.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	237	1907	1245	307	344	11	0	0	0
1	B	248	2038	1325	337	363	13	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	SER	-	expression tag	UNP G3V5E1
A	10	MET	-	expression tag	UNP G3V5E1
B	9	SER	-	expression tag	UNP G3V5E1
B	10	MET	-	expression tag	UNP G3V5E1

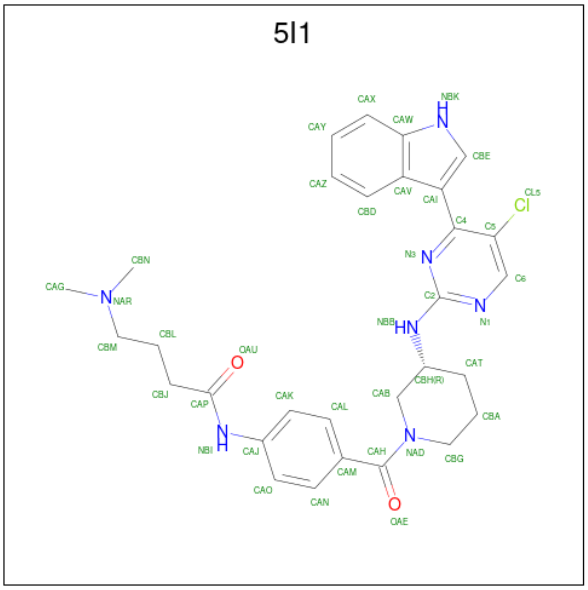
- Molecule 2 is a protein called CYCLIN-DEPENDENT KINASE 12.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
2	C	325	2571	1654	424	475	1	17	0	0	0
2	D	320	2477	1601	403	456	1	16	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	713	SER	-	expression tag	UNP Q9NYV4
C	714	MET	-	expression tag	UNP Q9NYV4
D	713	SER	-	expression tag	UNP Q9NYV4
D	714	MET	-	expression tag	UNP Q9NYV4

- Molecule 3 is N-[4-[(3R)-3-[[5-chloranyl-4-(1H-indol-3-yl)pyrimidin-2-yl]amino]piperidin-1-yl]carbonylphenyl]-4-(dimethylamino)butanamide (three-letter code: 5I1) (formula: $C_{30}H_{34}ClN_7O_2$).

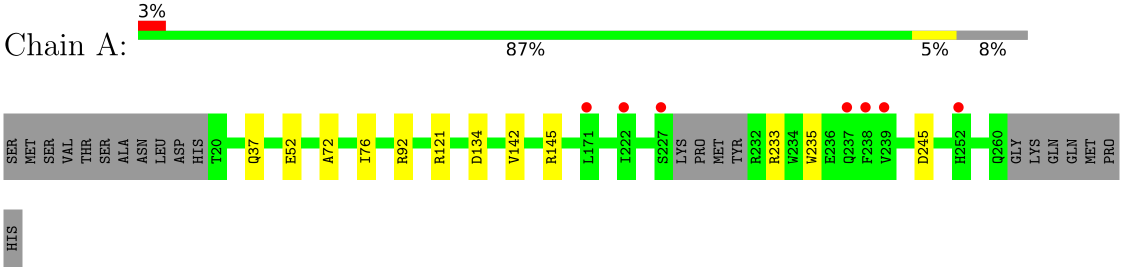


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Cl	N	O		
3	C	1	40	30	1	7	2	0	0
3	D	1	40	30	1	7	2	0	0

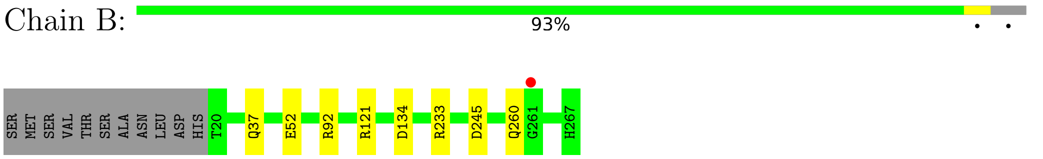
3 Residue-property plots ⓘ

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.

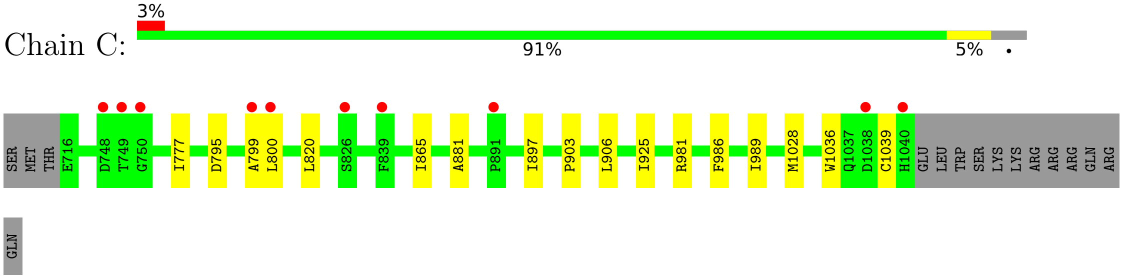
- Molecule 1: CYCLIN-K



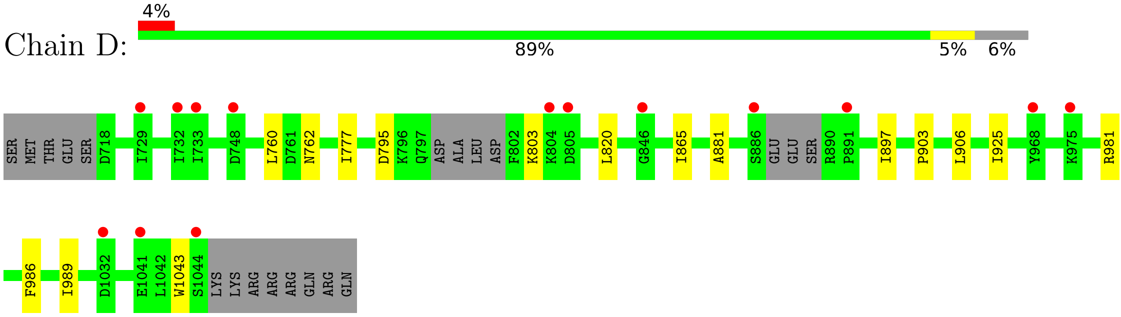
- Molecule 1: CYCLIN-K



- Molecule 2: CYCLIN-DEPENDENT KINASE 12



- Molecule 2: CYCLIN-DEPENDENT KINASE 12



4 Data and refinement statistics

Property	**Value**	**Source**
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	49.78Å 148.69Å 91.58Å 90.00° 93.81° 90.00°	Depositor
Resolution (Å)	91.38 – 2.70 41.30 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.6 (91.38-2.70) 98.7 (41.30-2.70)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.97 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0123	Depositor
R, R_{free}	0.221 , 0.262 0.223 , 0.264	Depositor DCC
R_{free} test set	1756 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å^2)	54.3	Xtriage
Anisotropy	0.590	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 29.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	9073	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.60% 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: 5I1, 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	0.38	0/1959	0.54	0/2667
1	B	0.38	0/2095	0.58	0/2842
2	C	0.38	0/2619	0.58	0/3546
2	D	0.37	0/2524	0.57	0/3429
All	All	0.38	0/9197	0.57	0/12484

There are no bond length outliers.

There are no bond angle outliers.

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	1907	0	1803	5	0
1	B	2038	0	1981	1	0
2	C	2571	0	2494	5	0
2	D	2477	0	2336	7	0
3	C	40	0	0	1	0
3	D	40	0	0	1	0
All	All	9073	0	8614	18	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 1.

The worst 5 of 18 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:B:52:GLU:OE1	1:B:92:ARG:NH1	2.38	0.57
1:A:52:GLU:OE1	1:A:92:ARG:NH1	2.37	0.56
2:C:865:ILE:CD1	2:C:925:ILE:HD13	2.39	0.53
2:D:865:ILE:CD1	2:D:925:ILE:HD13	2.40	0.52
2:D:986:PHE:O	2:D:989:ILE:HG22	2.12	0.49

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	233/259 (90%)	229 (98%)	4 (2%)	0	100	100
1	B	246/259 (95%)	243 (99%)	3 (1%)	0	100	100
2	C	322/340 (95%)	300 (93%)	17 (5%)	5 (2%)	9	24
2	D	313/340 (92%)	293 (94%)	18 (6%)	2 (1%)	25	50
All	All	1114/1198 (93%)	1065 (96%)	42 (4%)	7 (1%)	25	50

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	1036	TRP
2	D	795	ASP
2	C	795	ASP
2	C	799	ALA
2	C	897	ILE

5.3.2 Protein sidechains ⓘ

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	195/233 (84%)	191 (98%)	4 (2%)	53	80
1	B	216/233 (93%)	210 (97%)	6 (3%)	43	73
2	C	270/307 (88%)	266 (98%)	4 (2%)	65	86
2	D	249/307 (81%)	245 (98%)	4 (2%)	62	85
All	All	930/1080 (86%)	912 (98%)	18 (2%)	57	82

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	762	ASN
2	D	1043	TRP
2	D	981	ARG
1	B	245	ASP
2	C	1028	MET

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	C	783	HIS
2	C	977	GLN
2	D	783	HIS
2	D	762	ASN
1	B	156	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 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	D	893	2	8,10,11	0.92	0	10,14,16	1.02	0
2	TPO	C	893	2	8,10,11	0.77	0	10,14,16	1.05	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
2	TPO	D	893	2	-	0/9/11/13	-
2	TPO	C	893	2	-	0/9/11/13	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

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)

2 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	$\#\|Z\| > 2$
3	5I1	C	2000	2	41,44,44	1.89	9 (21%)	52,61,61	1.97	11 (21%)
3	5I1	D	2000	2	41,44,44	1.84	8 (19%)	52,61,61	1.74	11 (21%)

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
3	5I1	C	2000	2	-	3/22/36/36	0/5/5/5
3	5I1	D	2000	2	-	5/22/36/36	0/5/5/5

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	2000	5I1	CAM-CAH	-6.64	1.39	1.50
3	D	2000	5I1	CAM-CAH	-6.10	1.39	1.50
3	D	2000	5I1	CAX-CAW	-4.32	1.34	1.41
3	C	2000	5I1	CAX-CAW	-4.23	1.34	1.41
3	C	2000	5I1	CAJ-NBI	-3.63	1.34	1.41

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	2000	5I1	N1-C2-N3	-6.87	120.04	126.55
3	D	2000	5I1	N1-C2-N3	-6.25	120.63	126.55
3	C	2000	5I1	CAM-CAH-NAD	5.68	125.93	118.72
3	C	2000	5I1	CAT-CBA-CBG	-4.53	104.51	110.85
3	D	2000	5I1	CBG-NAD-CAB	3.88	120.82	113.06

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	2000	5I1	CAP-CBJ-CBL-CBM
3	D	2000	5I1	CBJ-CBL-CBM-NAR
3	D	2000	5I1	CAP-CBJ-CBL-CBM
3	C	2000	5I1	OAU-CAP-CBJ-CBL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	C	2000	5I1	NBI-CAP-CBJ-CBL

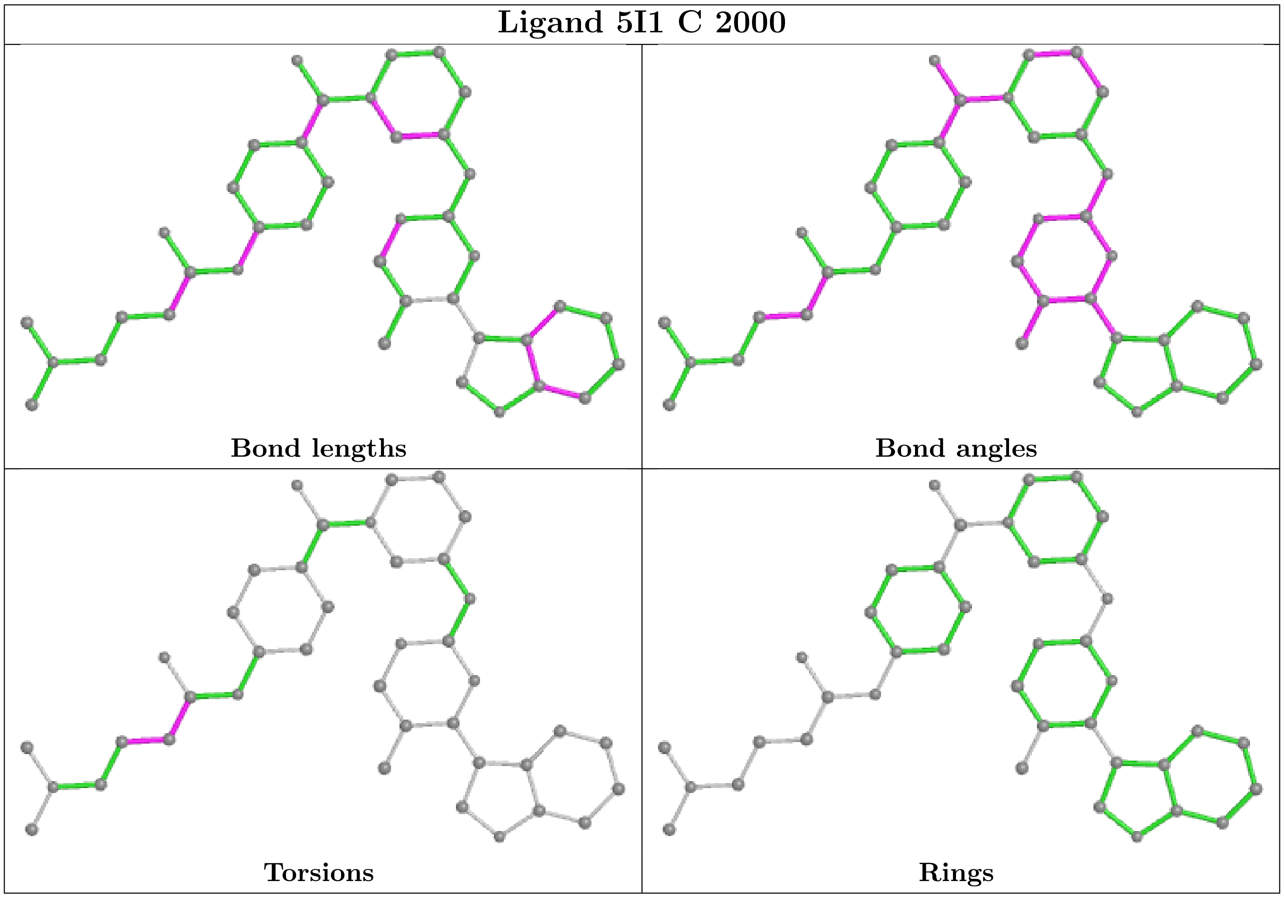
There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	2000	5I1	1	0
3	D	2000	5I1	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 5I1 C 2000



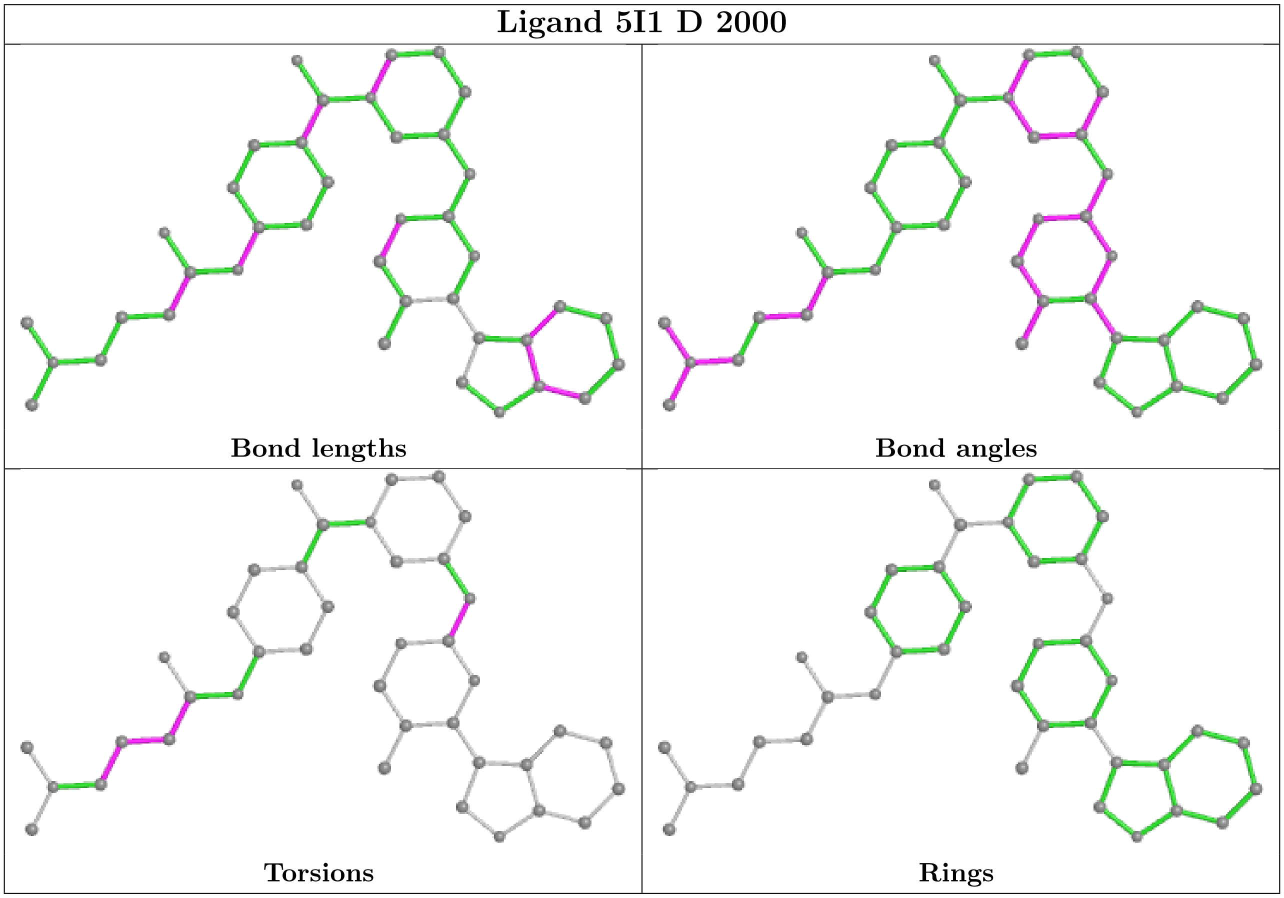
Bond lengths

Bond angles

Torsions

Rings

Ligand 5I1 D 2000



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	237/259 (91%)	0.08	7 (2%)	50	51	37, 58, 87, 108	0
1	B	248/259 (95%)	-0.06	1 (0%)	92	93	31, 47, 78, 129	0
2	C	324/340 (95%)	0.19	10 (3%)	49	49	39, 60, 95, 141	0
2	D	319/340 (93%)	0.26	14 (4%)	34	33	45, 69, 106, 129	0
All	All	1128/1198 (94%)	0.13	32 (2%)	53	54	31, 60, 97, 141	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	826	SER	4.7
2	C	800	LEU	4.4
2	D	1044	SER	4.0
2	D	805	ASP	4.0
2	C	799	ALA	4.0

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	C	893	11/12	0.92	0.14	74,78,82,85	0
2	TPO	D	893	11/12	0.95	0.13	76,78,79,82	0

6.3 Carbohydrates ⓘ

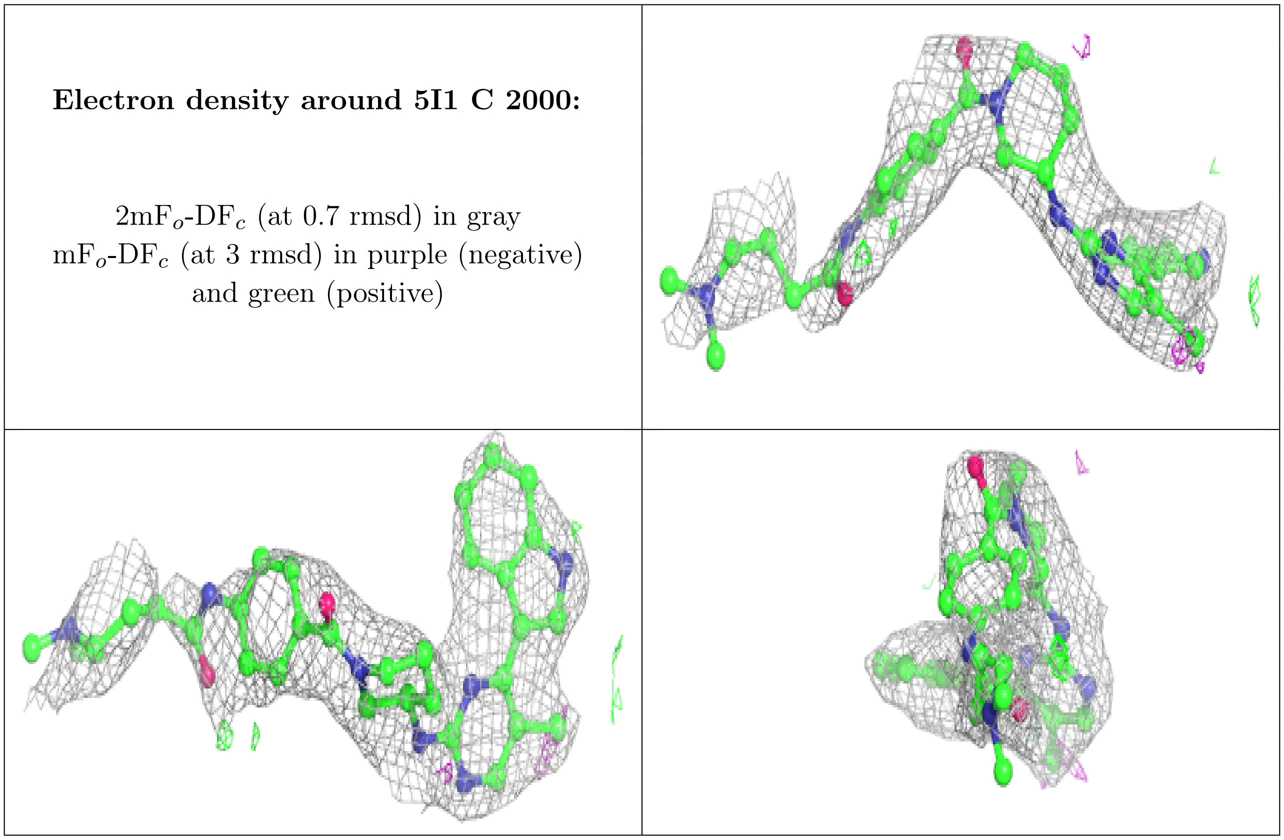
There are no monosaccharides in this entry.

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, 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
3	5I1	C	2000	40/40	0.89	0.26	68,76,120,121	0
3	5I1	D	2000	40/40	0.91	0.33	71,77,112,113	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.

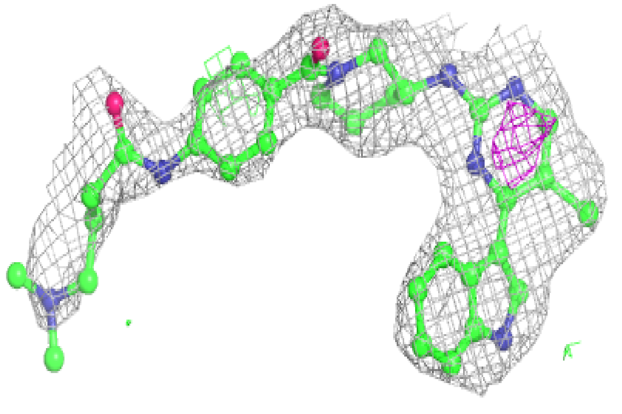


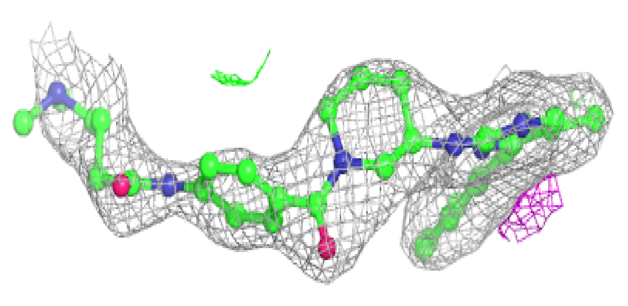
Electron density around 5I1 C 2000:

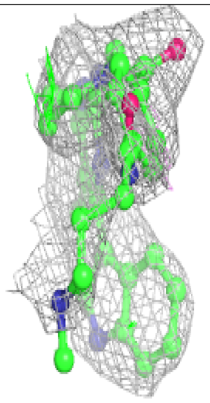
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around 5I1 D 2000:

$2mF_o$-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.