



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 5, 2024 – 04:14 pm BST

PDB ID : 8RO9  
Title : Human cohesin SMC1A-HD(longCC-EQ)/RAD21-C complex - Open/closed  
P-loop conformation  
Authors : Vitoria Gomes, M.; Romier, C.  
Deposited on : 2024-01-11  
Resolution : 1.77 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 : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.002 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.38.2

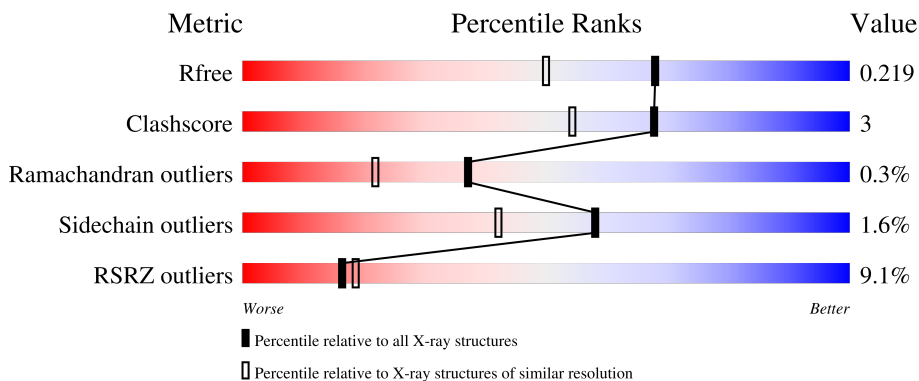
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.77 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))
$R_{free}$	164625	2888 (1.76-1.76)
Clashscore	180529	3097 (1.76-1.76)
Ramachandran outliers	177936	3072 (1.76-1.76)
Sidechain outliers	177891	3072 (1.76-1.76)
RSRZ outliers	164620	2887 (1.76-1.76)

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	456	
1	C	456	
2	B	81	
2	D	81	

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 6877 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 Structural maintenance of chromosomes protein 1A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	363	2912	1858	497	548	9	0	9	0
1	C	324	2573	1642	436	487	8	0	7	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	978	GLU	-	linker	UNP Q14683
A	979	SER	-	linker	UNP Q14683
A	980	SER	-	linker	UNP Q14683
A	981	LYS	-	linker	UNP Q14683
A	982	HIS	-	linker	UNP Q14683
A	983	PRO	-	linker	UNP Q14683
A	984	THR	-	linker	UNP Q14683
A	985	SER	-	linker	UNP Q14683
A	986	LEU	-	linker	UNP Q14683
A	987	VAL	-	linker	UNP Q14683
A	988	PRO	-	linker	UNP Q14683
A	989	ARG	-	linker	UNP Q14683
A	990	GLY	-	linker	UNP Q14683
A	991	SER	-	linker	UNP Q14683
A	1157	GLN	GLU	engineered mutation	UNP Q14683
C	978	GLU	-	linker	UNP Q14683
C	979	SER	-	linker	UNP Q14683
C	980	SER	-	linker	UNP Q14683
C	981	LYS	-	linker	UNP Q14683
C	982	HIS	-	linker	UNP Q14683
C	983	PRO	-	linker	UNP Q14683
C	984	THR	-	linker	UNP Q14683
C	985	SER	-	linker	UNP Q14683
C	986	LEU	-	linker	UNP Q14683
C	987	VAL	-	linker	UNP Q14683

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	988	PRO	-	linker	UNP Q14683
C	989	ARG	-	linker	UNP Q14683
C	990	GLY	-	linker	UNP Q14683
C	991	SER	-	linker	UNP Q14683
C	1157	GLN	GLU	engineered mutation	UNP Q14683

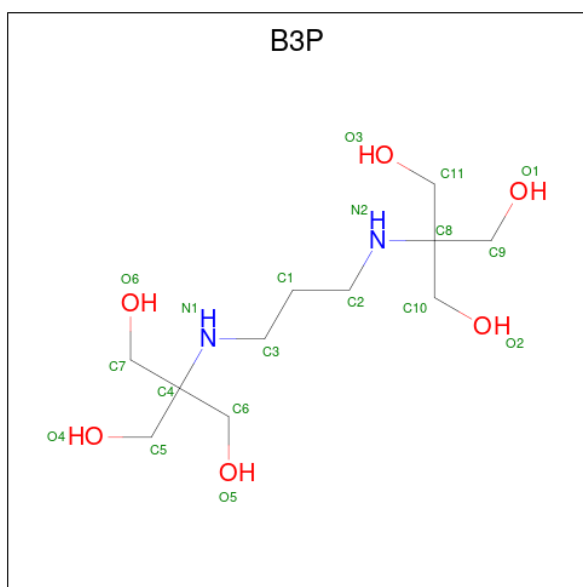
- Molecule 2 is a protein called 64-kDa C-terminal product.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	60	Total	C	N	O	S	0	1	0
			488	314	81	92	1			
2	D	62	Total	C	N	O	S	0	0	0
			496	318	84	93	1			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	557	MET	-	initiating methionine	UNP O60216
B	630	GLY	-	expression tag	UNP O60216
B	631	SER	-	expression tag	UNP O60216
B	632	LEU	-	expression tag	UNP O60216
B	633	GLU	-	expression tag	UNP O60216
B	634	VAL	-	expression tag	UNP O60216
B	635	LEU	-	expression tag	UNP O60216
B	636	PHE	-	expression tag	UNP O60216
B	637	GLN	-	expression tag	UNP O60216
D	557	MET	-	initiating methionine	UNP O60216
D	630	GLY	-	expression tag	UNP O60216
D	631	SER	-	expression tag	UNP O60216
D	632	LEU	-	expression tag	UNP O60216
D	633	GLU	-	expression tag	UNP O60216
D	634	VAL	-	expression tag	UNP O60216
D	635	LEU	-	expression tag	UNP O60216
D	636	PHE	-	expression tag	UNP O60216
D	637	GLN	-	expression tag	UNP O60216

- Molecule 3 is 2-[3-(2-HYDROXY-1,1-DIHYDROXYMETHYL-ETHYLAMINO)-PROPYL AMINO]-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: B3P) (formula:  $C_{11}H_{26}N_2O_6$ ).

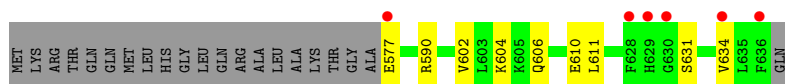


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	19	11	2	6	0	0

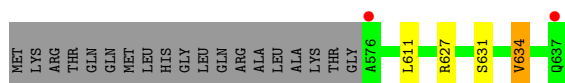
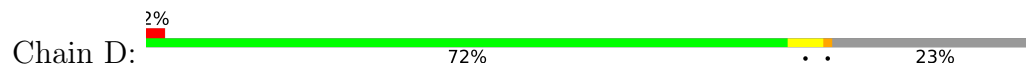
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	170	Total	O	0	0
			170	170		
4	B	33	Total	O	0	0
			33	33		
4	C	151	Total	O	0	0
			151	151		
4	D	35	Total	O	0	0
			35	35		





- Molecule 2: 64-kDa C-terminal product



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	187.92Å 64.60Å 89.95Å 90.00° 99.74° 90.00°	Depositor
Resolution (Å)	44.65 – 1.77 44.65 – 1.77	Depositor EDS
% Data completeness (in resolution range)	99.7 (44.65-1.77) 99.7 (44.65-1.77)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.30 (at 1.77Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.190 , 0.218 0.188 , 0.219	Depositor DCC
$R_{free}$ test set	5157 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.2	Xtrriage
Anisotropy	0.278	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 49.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6877	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 20.86 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 7.9381e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: B3P

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	0/2983	0.56	0/4009
1	C	0.39	0/2633	0.58	0/3548
2	B	0.43	0/500	0.54	0/674
2	D	0.40	0/505	0.57	0/681
All	All	0.39	0/6621	0.57	0/8912

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	2912	0	2940	22	0
1	C	2573	0	2586	15	0
2	B	488	0	494	7	0
2	D	496	0	501	3	0
3	A	19	0	26	2	0
4	A	170	0	0	6	0
4	B	33	0	0	2	0
4	C	151	0	0	1	0
4	D	35	0	0	1	0
All	All	6877	0	6547	45	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 3.

The worst 5 of 45 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1123:ARG:NH1	4:A:1401:HOH:O	2.09	0.84
1:A:92[B]:ARG:NH1	4:A:1402:HOH:O	2.12	0.72
2:B:606:GLN:OE1	4:B:701:HOH:O	2.10	0.69
1:A:1216:LEU:HD11	2:B:611:LEU:HD13	1.73	0.69
1:C:1216:LEU:HD11	2:D:611:LEU:HD13	1.75	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	368/456 (81%)	358 (97%)	8 (2%)	2 (0%)	25	11
1	C	327/456 (72%)	317 (97%)	9 (3%)	1 (0%)	37	22
2	B	59/81 (73%)	58 (98%)	1 (2%)	0	100	100
2	D	60/81 (74%)	60 (100%)	0	0	100	100
All	All	814/1074 (76%)	793 (97%)	18 (2%)	3 (0%)	37	16

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	1130	GLY
1	A	37[A]	GLY
1	A	37[B]	GLY

### 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	311/388 (80%)	308 (99%)	3 (1%)	73	61
1	C	278/388 (72%)	271 (98%)	7 (2%)	42	22
2	B	54/69 (78%)	54 (100%)	0	100	100
2	D	54/69 (78%)	53 (98%)	1 (2%)	52	34
All	All	697/914 (76%)	686 (98%)	11 (2%)	58	42

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	1158	ILE
1	C	1199	SER
2	D	634	VAL
1	C	1210	CYS
1	C	143	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	137	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

1 ligand is 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	B3P	A	1300	-	18,18,18	0.48	0	21,23,23	1.21	2 (9%)

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	B3P	A	1300	-	-	9/28/28/28	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1300	B3P	C3-N1-C4	-3.88	110.58	116.08
3	A	1300	B3P	O2-C10-C8	-2.06	107.46	111.63

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

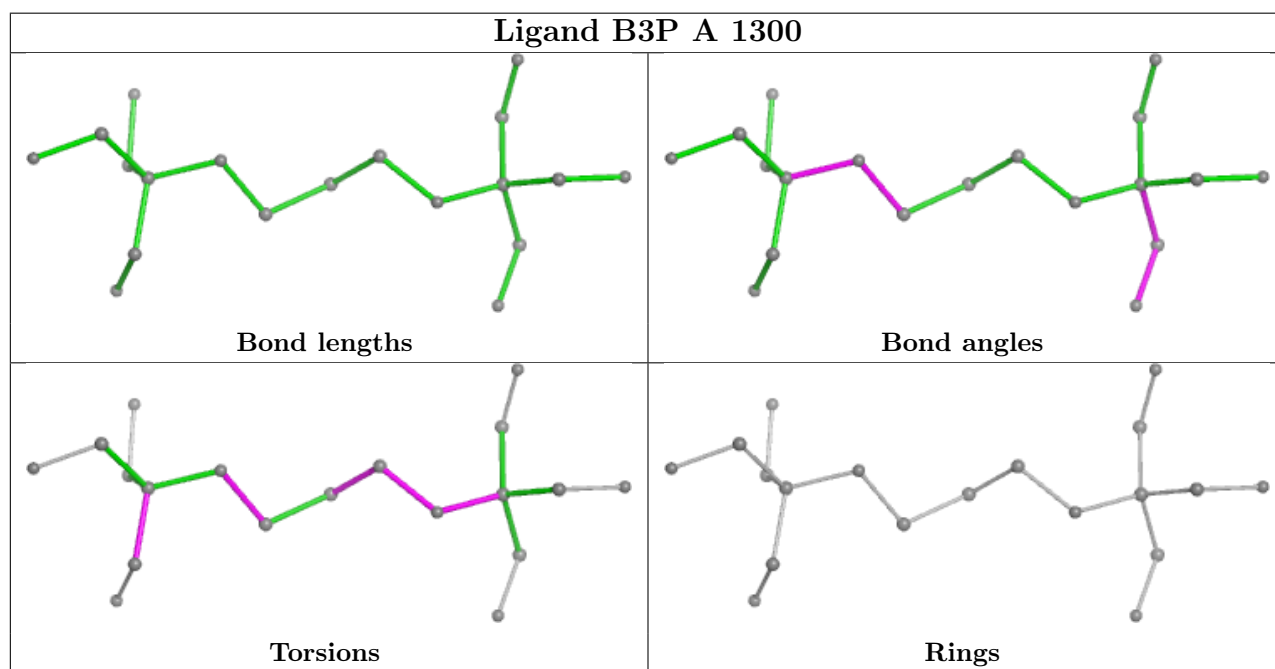
Mol	Chain	Res	Type	Atoms
3	A	1300	B3P	N1-C4-C7-O6
3	A	1300	B3P	C5-C4-C7-O6
3	A	1300	B3P	C6-C4-C7-O6
3	A	1300	B3P	C9-C8-N2-C2
3	A	1300	B3P	C10-C8-N2-C2

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1300	B3P	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 [i](#)

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	363/456 (79%)	0.64	30 (8%) 19 21	17, 43, 84, 134	9 (2%)
1	C	324/456 (71%)	0.66	36 (11%) 12 13	17, 42, 91, 147	7 (2%)
2	B	60/81 (74%)	0.58	6 (10%) 14 16	25, 37, 86, 123	1 (1%)
2	D	62/81 (76%)	0.40	2 (3%) 50 57	30, 39, 65, 113	0
All	All	809/1074 (75%)	0.62	74 (9%) 16 18	17, 42, 87, 147	17 (2%)

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1045	PHE	8.5
1	C	35[A]	GLY	7.2
1	C	1067	PHE	6.8
1	A	35[A]	GLY	6.6
1	A	1226	ALA	6.5

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

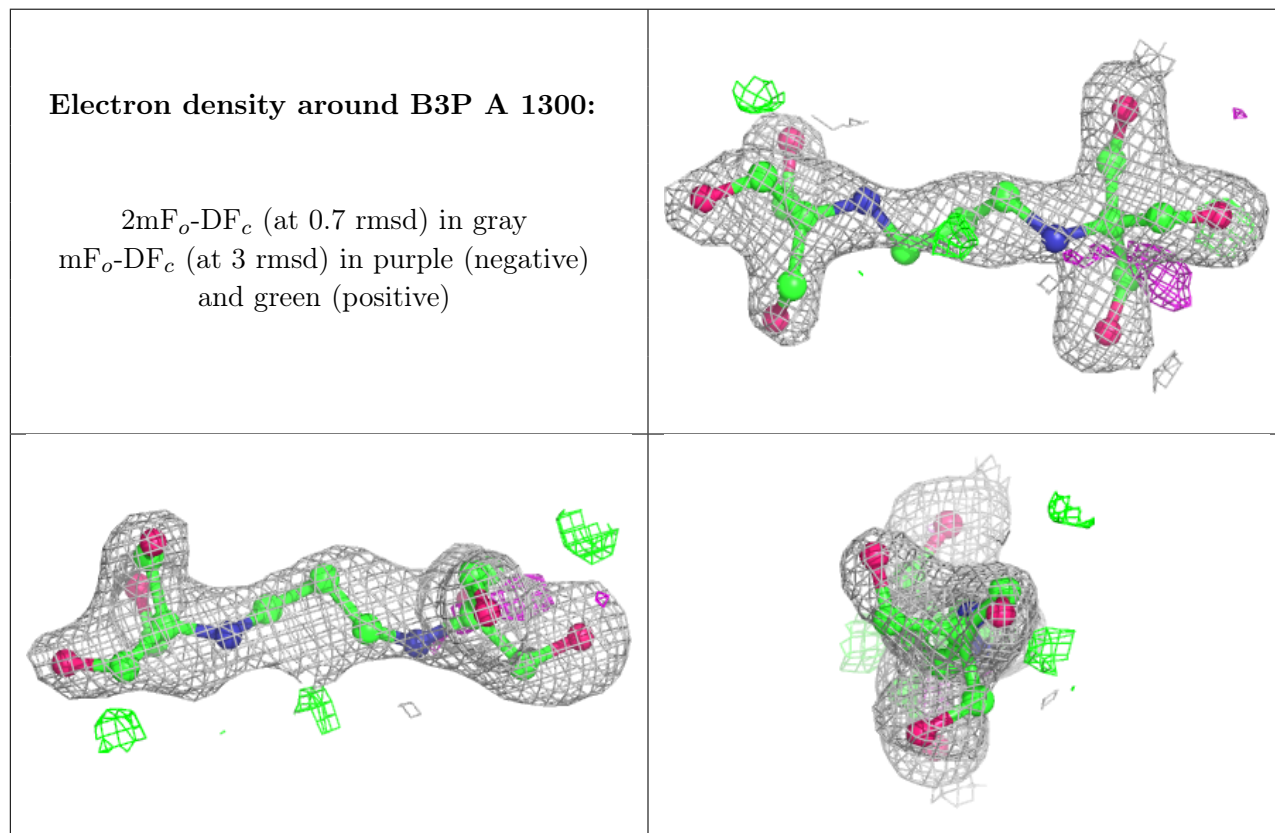
There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<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
3	B3P	A	1300	19/19	0.83	0.14	47,72,80,82	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.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.