Full wwPDB X-ray Structure Validation Report (i)

May 13, 2020 – 03:54 am BST

PDB ID	:	5M5G
Title	:	Crystal structure of the Chaetomium Thermophilum polycomb repressive complex 2 (PRC2)
Authors	:	Zhang, Y.; Justin, N.; Wilson, J.; Gamblin, S.
Deposited on	:	2016-10-21
Resolution	:	2.27 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

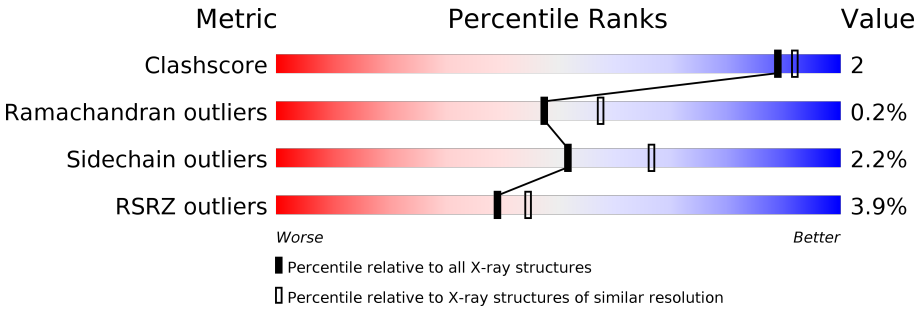
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.27 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))
Clashscore	141614	7711 (2.30-2.26)
Ramachandran outliers	138981	7597 (2.30-2.26)
Sidechain outliers	138945	7598 (2.30-2.26)
RSRZ outliers	127900	6849 (2.30-2.26)

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 on 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	605	74% 5% 20%
2	B	937	5% 78% 7% 15%
3	D	11	45% 18% 9% 27%
4	E	9	22% 89% 11%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 10893 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 putative polycomb protein EED.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	481	Total	C	N	O	S	0	0	0
			3782	2417	647	699	19			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-39	MET	-	initiating methionine	UNP G0S8H7
A	-38	ALA	-	expression tag	UNP G0S8H7
A	-37	SER	-	expression tag	UNP G0S8H7
A	-36	ALA	-	expression tag	UNP G0S8H7
A	-35	TRP	-	expression tag	UNP G0S8H7
A	-34	SER	-	expression tag	UNP G0S8H7
A	-33	HIS	-	expression tag	UNP G0S8H7
A	-32	PRO	-	expression tag	UNP G0S8H7
A	-31	GLN	-	expression tag	UNP G0S8H7
A	-30	PHE	-	expression tag	UNP G0S8H7
A	-29	GLU	-	expression tag	UNP G0S8H7
A	-28	LYS	-	expression tag	UNP G0S8H7
A	-27	GLY	-	expression tag	UNP G0S8H7
A	-26	GLY	-	expression tag	UNP G0S8H7
A	-25	GLY	-	expression tag	UNP G0S8H7
A	-24	SER	-	expression tag	UNP G0S8H7
A	-23	GLY	-	expression tag	UNP G0S8H7
A	-22	GLY	-	expression tag	UNP G0S8H7
A	-21	GLY	-	expression tag	UNP G0S8H7
A	-20	SER	-	expression tag	UNP G0S8H7
A	-19	GLY	-	expression tag	UNP G0S8H7
A	-18	GLY	-	expression tag	UNP G0S8H7
A	-17	SER	-	expression tag	UNP G0S8H7
A	-16	ALA	-	expression tag	UNP G0S8H7
A	-15	TRP	-	expression tag	UNP G0S8H7
A	-14	SER	-	expression tag	UNP G0S8H7
A	-13	HIS	-	expression tag	UNP G0S8H7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-12	PRO	-	expression tag	UNP G0S8H7
A	-11	GLN	-	expression tag	UNP G0S8H7
A	-10	PHE	-	expression tag	UNP G0S8H7
A	-9	GLU	-	expression tag	UNP G0S8H7
A	-8	LYS	-	expression tag	UNP G0S8H7
A	-7	LEU	-	expression tag	UNP G0S8H7
A	-6	GLU	-	expression tag	UNP G0S8H7
A	-5	VAL	-	expression tag	UNP G0S8H7
A	-4	LEU	-	expression tag	UNP G0S8H7
A	-3	PHE	-	expression tag	UNP G0S8H7
A	-2	GLN	-	expression tag	UNP G0S8H7
A	-1	GLY	-	expression tag	UNP G0S8H7
A	0	PRO	-	expression tag	UNP G0S8H7

- Molecule 2 is a protein called Putative uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	797	Total	C	N	O	S	0	0	0
			6426	4058	1155	1173	40			

There are 177 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	182	SER	-	expression tag	UNP G0SDW4
B	183	ASN	-	expression tag	UNP G0SDW4
B	184	HIS	-	expression tag	UNP G0SDW4
B	185	HIS	-	expression tag	UNP G0SDW4
B	186	HIS	-	expression tag	UNP G0SDW4
B	187	HIS	-	expression tag	UNP G0SDW4
B	188	HIS	-	expression tag	UNP G0SDW4
B	189	HIS	-	expression tag	UNP G0SDW4
B	190	ALA	-	expression tag	UNP G0SDW4
B	2524	LEU	ALA	conflict	UNP G0SDW4
B	2525	VAL	ALA	conflict	UNP G0SDW4
B	2526	PRO	-	expression tag	UNP G0SDW4
B	2527	ARG	-	expression tag	UNP G0SDW4
B	2528	GLY	-	expression tag	UNP G0SDW4
B	2529	SER	-	expression tag	UNP G0SDW4
B	2530	GLU	-	expression tag	UNP G0SDW4
B	2531	VAL	-	expression tag	UNP G0SDW4
B	2532	MET	-	expression tag	UNP G0SDW4
B	2533	LEU	-	expression tag	UNP G0SDW4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	2534	PRO	-	expression tag	UNP G0SDW4
B	2535	GLY	-	expression tag	UNP G0SDW4
B	2536	ARG	-	expression tag	UNP G0SDW4
B	2537	GLY	-	expression tag	UNP G0SDW4
B	2538	VAL	-	expression tag	UNP G0SDW4
B	2539	PRO	-	expression tag	UNP G0SDW4
B	2540	LYS	-	expression tag	UNP G0SDW4
B	2541	LYS	-	expression tag	UNP G0SDW4
B	2542	PRO	-	expression tag	UNP G0SDW4
B	2543	LEU	-	expression tag	UNP G0SDW4
B	2544	ARG	-	expression tag	UNP G0SDW4
B	2545	ARG	-	expression tag	UNP G0SDW4
B	2546	PRO	-	expression tag	UNP G0SDW4
B	2547	LYS	-	expression tag	UNP G0SDW4
B	2548	ARG	-	expression tag	UNP G0SDW4
B	2549	ARG	-	expression tag	UNP G0SDW4
B	2550	PRO	-	expression tag	UNP G0SDW4
B	2551	LEU	-	expression tag	UNP G0SDW4
B	2552	LEU	-	expression tag	UNP G0SDW4
B	2553	VAL	-	expression tag	UNP G0SDW4
B	2554	PRO	-	expression tag	UNP G0SDW4
B	2555	LYS	-	expression tag	UNP G0SDW4
B	2556	THR	-	expression tag	UNP G0SDW4
B	2557	THR	-	expression tag	UNP G0SDW4
B	2558	GLN	-	expression tag	UNP G0SDW4
B	2559	PRO	-	expression tag	UNP G0SDW4
B	2560	LEU	-	expression tag	UNP G0SDW4
B	2561	PHE	-	expression tag	UNP G0SDW4
B	2562	ASP	-	expression tag	UNP G0SDW4
B	2563	PRO	-	expression tag	UNP G0SDW4
B	2564	LEU	-	expression tag	UNP G0SDW4
B	2565	SER	-	expression tag	UNP G0SDW4
B	2566	LYS	-	expression tag	UNP G0SDW4
B	2567	VAL	-	expression tag	UNP G0SDW4
B	2568	GLN	-	expression tag	UNP G0SDW4
B	2569	LEU	-	expression tag	UNP G0SDW4
B	2570	LEU	-	expression tag	UNP G0SDW4
B	2571	PRO	-	expression tag	UNP G0SDW4
B	2572	GLY	-	expression tag	UNP G0SDW4
B	2573	GLN	-	expression tag	UNP G0SDW4
B	2574	PRO	-	expression tag	UNP G0SDW4
B	2575	LEU	-	expression tag	UNP G0SDW4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	2576	PRO	-	expression tag	UNP G0SDW4
B	2577	GLN	-	expression tag	UNP G0SDW4
B	2578	HIS	-	expression tag	UNP G0SDW4
B	2579	PRO	-	expression tag	UNP G0SDW4
B	2580	ILE	-	expression tag	UNP G0SDW4
B	2581	ASP	-	expression tag	UNP G0SDW4
B	2582	ASP	-	expression tag	UNP G0SDW4
B	2583	SER	-	expression tag	UNP G0SDW4
B	2584	TRP	-	expression tag	UNP G0SDW4
B	2585	LEU	-	expression tag	UNP G0SDW4
B	2586	LEU	-	expression tag	UNP G0SDW4
B	2587	LEU	-	expression tag	UNP G0SDW4
B	2588	LYS	-	expression tag	UNP G0SDW4
B	2589	HIS	-	expression tag	UNP G0SDW4
B	2590	ARG	-	expression tag	UNP G0SDW4
B	2591	ASP	-	expression tag	UNP G0SDW4
B	2592	ASN	-	expression tag	UNP G0SDW4
B	2593	LEU	-	expression tag	UNP G0SDW4
B	2594	GLN	-	expression tag	UNP G0SDW4
B	2595	ASP	-	expression tag	UNP G0SDW4
B	2596	PHE	-	expression tag	UNP G0SDW4
B	2597	ILE	-	expression tag	UNP G0SDW4
B	2598	ASP	-	expression tag	UNP G0SDW4
B	2599	LEU	-	expression tag	UNP G0SDW4
B	2600	ARG	-	expression tag	UNP G0SDW4
B	2601	PRO	-	expression tag	UNP G0SDW4
B	2602	GLU	-	expression tag	UNP G0SDW4
B	2603	GLU	-	expression tag	UNP G0SDW4
B	2604	LYS	-	expression tag	UNP G0SDW4
B	2605	GLU	-	expression tag	UNP G0SDW4
B	2606	PHE	-	expression tag	UNP G0SDW4
B	2607	LEU	-	expression tag	UNP G0SDW4
B	2608	GLN	-	expression tag	UNP G0SDW4
B	2609	GLU	-	expression tag	UNP G0SDW4
B	2610	TRP	-	expression tag	UNP G0SDW4
B	2611	ASP	-	expression tag	UNP G0SDW4
B	2612	ALA	-	expression tag	UNP G0SDW4
B	2613	PHE	-	expression tag	UNP G0SDW4
B	2614	ILE	-	expression tag	UNP G0SDW4
B	2615	LEU	-	expression tag	UNP G0SDW4
B	2616	ARG	-	expression tag	UNP G0SDW4
B	2617	ARG	-	expression tag	UNP G0SDW4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	2618	HIS	-	expression tag	UNP G0SDW4
B	2619	ILE	-	expression tag	UNP G0SDW4
B	2620	SER	-	expression tag	UNP G0SDW4
B	2621	SER	-	expression tag	UNP G0SDW4
B	2622	GLU	-	expression tag	UNP G0SDW4
B	2623	GLN	-	expression tag	UNP G0SDW4
B	2624	TYR	-	expression tag	UNP G0SDW4
B	2625	LEU	-	expression tag	UNP G0SDW4
B	2626	PRO	-	expression tag	UNP G0SDW4
B	2627	ARG	-	expression tag	UNP G0SDW4
B	2628	TYR	-	expression tag	UNP G0SDW4
B	2629	PHE	-	expression tag	UNP G0SDW4
B	2630	LEU	-	expression tag	UNP G0SDW4
B	2631	ARG	-	expression tag	UNP G0SDW4
B	2632	PHE	-	expression tag	UNP G0SDW4
B	2633	VAL	-	expression tag	UNP G0SDW4
B	2634	ARG	-	expression tag	UNP G0SDW4
B	2635	GLU	-	expression tag	UNP G0SDW4
B	2636	LYS	-	expression tag	UNP G0SDW4
B	2637	ALA	-	expression tag	UNP G0SDW4
B	2638	ASP	-	expression tag	UNP G0SDW4
B	2639	TRP	-	expression tag	UNP G0SDW4
B	2640	LEU	-	expression tag	UNP G0SDW4
B	2641	VAL	-	expression tag	UNP G0SDW4
B	2642	SER	-	expression tag	UNP G0SDW4
B	2643	LYS	-	expression tag	UNP G0SDW4
B	2644	ARG	-	expression tag	UNP G0SDW4
B	2645	SER	-	expression tag	UNP G0SDW4
B	2646	ARG	-	expression tag	UNP G0SDW4
B	2647	GLY	-	expression tag	UNP G0SDW4
B	2648	GLU	-	expression tag	UNP G0SDW4
B	2649	GLU	-	expression tag	UNP G0SDW4
B	2650	PHE	-	expression tag	UNP G0SDW4
B	2651	SER	-	expression tag	UNP G0SDW4
B	2652	LYS	-	expression tag	UNP G0SDW4
B	2653	LEU	-	expression tag	UNP G0SDW4
B	2654	VAL	-	expression tag	UNP G0SDW4
B	2655	ALA	-	expression tag	UNP G0SDW4
B	2656	THR	-	expression tag	UNP G0SDW4
B	2657	LEU	-	expression tag	UNP G0SDW4
B	2658	LEU	-	expression tag	UNP G0SDW4
B	2659	ALA	-	expression tag	UNP G0SDW4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	2660	ARG	-	expression tag	UNP G0SDW4
B	2661	ARG	-	expression tag	UNP G0SDW4
B	2662	VAL	-	expression tag	UNP G0SDW4
B	2663	LEU	-	expression tag	UNP G0SDW4
B	2664	PRO	-	expression tag	UNP G0SDW4
B	2665	GLU	-	expression tag	UNP G0SDW4
B	2666	ARG	-	expression tag	UNP G0SDW4
B	2667	VAL	-	expression tag	UNP G0SDW4
B	2668	VAL	-	expression tag	UNP G0SDW4
B	2669	ILE	-	expression tag	UNP G0SDW4
B	2670	GLU	-	expression tag	UNP G0SDW4
B	2671	ALA	-	expression tag	UNP G0SDW4
B	2672	THR	-	expression tag	UNP G0SDW4
B	2673	GLN	-	expression tag	UNP G0SDW4
B	2674	VAL	-	expression tag	UNP G0SDW4
B	2675	LEU	-	expression tag	UNP G0SDW4
B	2676	ASN	-	expression tag	UNP G0SDW4
B	2677	ASP	-	expression tag	UNP G0SDW4
B	2678	ALA	-	expression tag	UNP G0SDW4
B	2679	ARG	-	expression tag	UNP G0SDW4
B	2680	GLY	-	expression tag	UNP G0SDW4
B	2681	ARG	-	expression tag	UNP G0SDW4
B	2682	LEU	-	expression tag	UNP G0SDW4
B	2683	ARG	-	expression tag	UNP G0SDW4
B	2684	GLU	-	expression tag	UNP G0SDW4
B	2685	GLN	-	expression tag	UNP G0SDW4
B	2686	GLY	-	expression tag	UNP G0SDW4
B	2687	GLY	-	expression tag	UNP G0SDW4
B	2688	VAL	-	expression tag	UNP G0SDW4
B	2689	ILE	-	expression tag	UNP G0SDW4
B	2690	GLU	-	expression tag	UNP G0SDW4
B	2691	GLY	-	expression tag	UNP G0SDW4

- Molecule 3 is a protein called HISTONE H3 11-Mer peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	8	Total	C	N	O	0	0	0
			60	38	13	9			

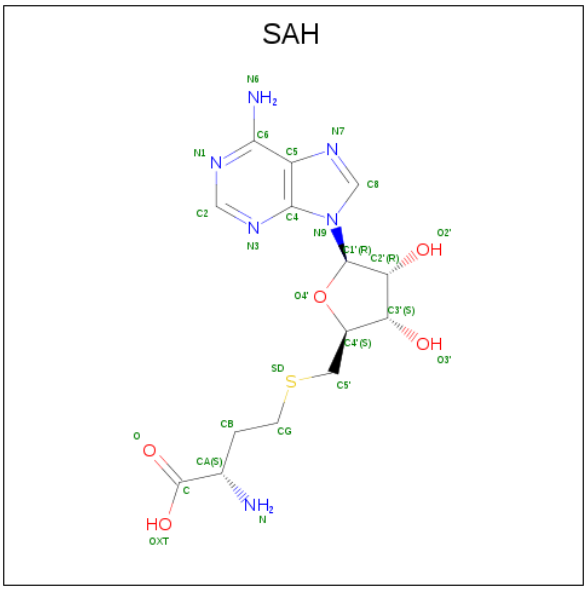
- Molecule 4 is a protein called Fragment from molecular 2 (region containing putative polycomb protein Suz12).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	9	Total	C	N	O	S	0	0	0
			63	41	12	9	1			

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	8	Total	Zn	0	0
			8	8		

- Molecule 6 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula: $C_{14}H_{20}N_6O_5S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	B	1	Total	C	N	O	S	0	0
			26	14	6	5	1		

- Molecule 7 is water.

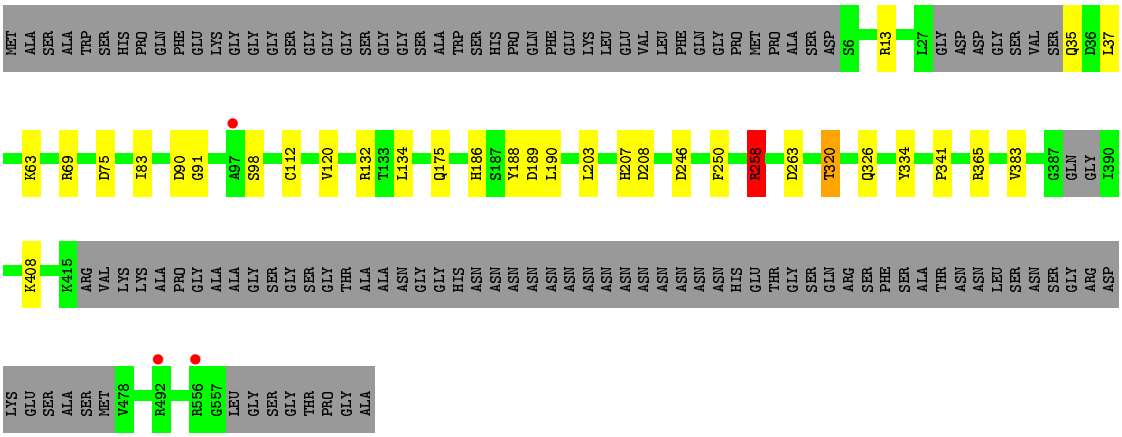
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	246	Total	O	0	0
			246	246		
7	B	274	Total	O	0	0
			274	274		
7	D	4	Total	O	0	0
			4	4		
7	E	4	Total	O	0	0
			4	4		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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.

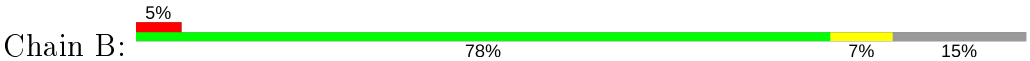
- Molecule 1: putative polycomb protein EED

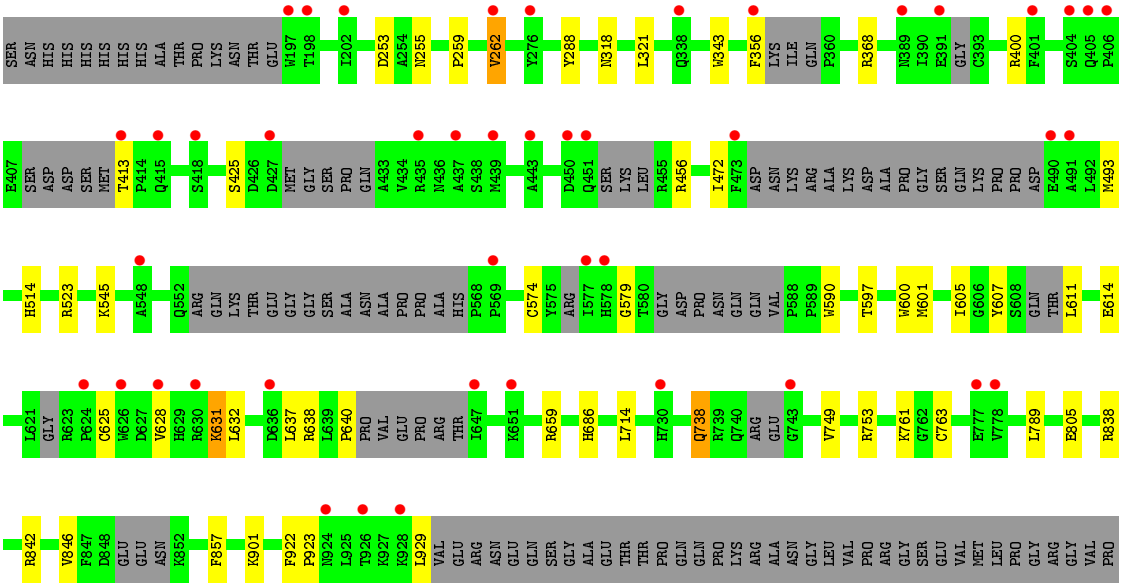
Chain A:



- Molecule 2: Putative uncharacterized protein

Chain B:





• Molecule 3: HISTONE H3 11-Mer peptide

Chain D:

• Molecule 4: Fragment from molecular 2 (region containing putative polycomb protein Suz12)

Chain E:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	81.55Å 74.61Å 128.51Å 90.00° 107.70° 90.00°	Depositor
Resolution (Å)	45.84 – 2.27 45.84 – 2.27	Depositor EDS
% Data completeness (in resolution range)	90.7 (45.84-2.27) 90.7 (45.84-2.27)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.24 (at 2.27Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.171 , 0.221 0.178 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	30.6	Xtriage
Anisotropy	0.130	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 51.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.014 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	10893	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.39% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

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

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, SAH, M3L

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	0.80	0/3894	0.88	8/5306 (0.2%)
2	B	0.76	0/6570	0.85	9/8870 (0.1%)
3	D	0.92	0/48	1.12	1/63 (1.6%)
4	E	0.81	0/64	0.92	0/86
All	All	0.78	0/10576	0.86	18/14325 (0.1%)

There are no bond length outliers.

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	258	ARG	NE-CZ-NH1	8.28	124.44	120.30
2	B	753	ARG	NE-CZ-NH1	-7.88	116.36	120.30
1	A	258	ARG	NE-CZ-NH2	-7.33	116.63	120.30
2	B	753	ARG	NE-CZ-NH2	7.11	123.85	120.30
1	A	208	ASP	CB-CG-OD1	6.21	123.89	118.30
1	A	132	ARG	NE-CZ-NH2	-6.21	117.20	120.30
3	D	26	ARG	NE-CZ-NH1	5.82	123.21	120.30
2	B	400	ARG	NE-CZ-NH1	5.62	123.11	120.30
2	B	2600	ARG	NE-CZ-NH2	5.52	123.06	120.30
1	A	13	ARG	NE-CZ-NH1	-5.47	117.57	120.30
2	B	523	ARG	NE-CZ-NH1	5.46	123.03	120.30
1	A	69	ARG	NE-CZ-NH2	-5.38	117.61	120.30
2	B	763	CYS	CA-C-N	5.35	126.90	116.20
1	A	263	ASP	CB-CG-OD1	5.31	123.08	118.30
2	B	368	ARG	NE-CZ-NH2	-5.31	117.65	120.30
1	A	75	ASP	CB-CG-OD1	5.10	122.89	118.30
2	B	456	ARG	NE-CZ-NH1	5.06	122.83	120.30
2	B	2590	ARG	NE-CZ-NH1	5.03	122.81	120.30

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	3782	0	3634	13	0
2	B	6426	0	6304	24	0
3	D	60	0	70	2	0
4	E	63	0	70	1	0
5	B	8	0	0	0	0
6	B	26	0	19	0	0
7	A	246	0	0	0	0
7	B	274	0	0	1	0
7	D	4	0	0	0	0
7	E	4	0	0	1	0
All	All	10893	0	10097	38	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 2.

All (38) 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:246:ASP:OD1	1:A:258:ARG:HD3	1.86	0.75
2:B:321:LEU:HD13	2:B:842:ARG:HD2	1.79	0.63
2:B:574:CYS:HA	2:B:625:CYS:HB3	1.86	0.56
1:A:63:LYS:NZ	1:A:91:GLY:O	2.39	0.56
1:A:320:THR:CG2	1:A:320:THR:O	2.56	0.54
2:B:632:LEU:HD12	2:B:637:LEU:HD12	1.90	0.52
3:D:26:ARG:HD2	3:D:28:SER:O	2.11	0.51
2:B:318:ASN:HB3	2:B:838:ARG:HH12	1.76	0.49
2:B:607:TYR:OH	2:B:640:PRO:O	2.28	0.49
1:A:35:GLN:OE1	1:A:37:LEU:HD23	2.12	0.49
1:A:189:ASP:HB2	1:A:207:HIS:HD2	1.79	0.48
4:E:21:VAL:N	7:E:101:HOH:O	2.47	0.48
2:B:597:THR:HA	2:B:600:TRP:CE2	2.49	0.47
2:B:2622:GLU:HB2	2:B:2662:VAL:HG13	1.97	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:413:THR:O	2:B:413:THR:HG23	2.15	0.46
2:B:600:TRP:CH2	2:B:2652:LYS:HE2	2.50	0.46
1:A:326:GLN:OE1	3:D:26:ARG:NH1	2.44	0.44
1:A:186:HIS:CG	1:A:190:LEU:HD21	2.52	0.44
2:B:253:ASP:HB3	2:B:255:ASN:OD1	2.18	0.44
1:A:334:TYR:O	1:A:341:PRO:HA	2.18	0.44
1:A:98:SER:HA	1:A:112:CYS:O	2.18	0.44
2:B:545:LYS:HG3	2:B:2661:ARG:HG2	2.00	0.44
1:A:203:LEU:HG	1:A:250:PHE:CE2	2.53	0.43
2:B:2664:PRO:HB2	2:B:2666:ARG:HG2	1.99	0.43
1:A:120:VAL:HB	1:A:134:LEU:HB2	2.00	0.43
2:B:259:PRO:HB2	2:B:262:VAL:HG13	2.00	0.43
1:A:83:ILE:HB	2:B:288:TYR:CE2	2.53	0.43
2:B:472:ILE:O	2:B:472:ILE:HG22	2.19	0.42
2:B:761:LYS:NZ	7:B:8106:HOH:O	2.52	0.42
2:B:605:ILE:CD1	2:B:611:LEU:HD21	2.50	0.41
2:B:514:HIS:CG	2:B:2607:LEU:HD22	2.56	0.41
1:A:383:VAL:CG1	1:A:408:LYS:HB2	2.50	0.41
2:B:601:MET:O	2:B:605:ILE:HG12	2.21	0.41
2:B:686:HIS:HB2	2:B:714:LEU:HD12	2.04	0.40
2:B:590:TRP:CG	2:B:631:LYS:HD3	2.56	0.40
2:B:343:TRP:CD1	2:B:846:VAL:HG13	2.56	0.40
2:B:789:LEU:HD11	2:B:901:LYS:HB3	2.03	0.40
2:B:738:GLN:HA	2:B:749:VAL:HG11	2.03	0.40

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	473/605 (78%)	459 (97%)	14 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	B	765/937 (82%)	725 (95%)	38 (5%)	2 (0%)	41 49
3	D	5/11 (46%)	5 (100%)	0	0	100 100
4	E	7/9 (78%)	7 (100%)	0	0	100 100
All	All	1250/1562 (80%)	1196 (96%)	52 (4%)	2 (0%)	47 57

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	579	GLY
2	B	923	PRO

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	406/495 (82%)	400 (98%)	6 (2%)	65 77
2	B	699/816 (86%)	682 (98%)	17 (2%)	49 63
3	D	4/6 (67%)	3 (75%)	1 (25%)	0 0
4	E	7/7 (100%)	7 (100%)	0	100 100
All	All	1116/1324 (84%)	1092 (98%)	24 (2%)	52 66

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	90	ASP
1	A	175	GLN
1	A	188	TYR
1	A	258	ARG
1	A	320	THR
1	A	365	ARG
2	B	262	VAL
2	B	356	PHE
2	B	425	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	493	MET
2	B	614	GLU
2	B	628	VAL
2	B	631	LYS
2	B	638	ARG
2	B	659	ARG
2	B	738	GLN
2	B	805	GLU
2	B	857	PHE
2	B	922	PHE
2	B	929	LEU
2	B	2567	VAL
2	B	2644	ARG
2	B	2683	ARG
3	D	23	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	175	GLN
2	B	738	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)

1 non-standard protein/DNA/RNA residue 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	M3L	D	27	3	10,11,12	0.61	0	9,14,16	0.50	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
3	M3L	D	27	3	-	0/9/10/12	-

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 8 are monoatomic - leaving 1 for Mogul analysis.

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	SAH	B	8009	-	21,28,28	1.50	4 (19%)	20,40,40	1.70	3 (15%)

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	SAH	B	8009	-	-	1/7/31/31	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	8009	SAH	C2'-C1'	-2.76	1.49	1.53
6	B	8009	SAH	C4-N3	2.69	1.39	1.35
6	B	8009	SAH	C2-N3	2.67	1.36	1.32
6	B	8009	SAH	C5-C4	2.27	1.46	1.40

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	B	8009	SAH	N3-C2-N1	-4.95	120.94	128.68
6	B	8009	SAH	C2-N1-C6	3.03	123.93	118.75
6	B	8009	SAH	N6-C6-N1	2.09	122.90	118.57

There are no chirality outliers.

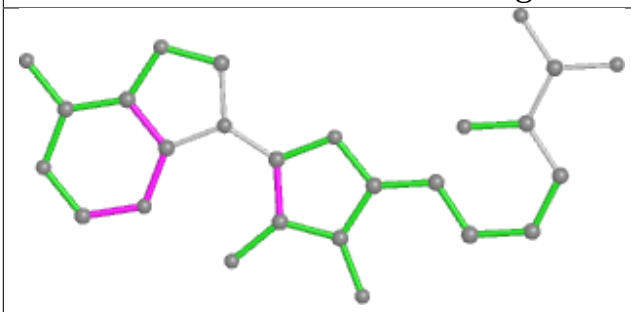
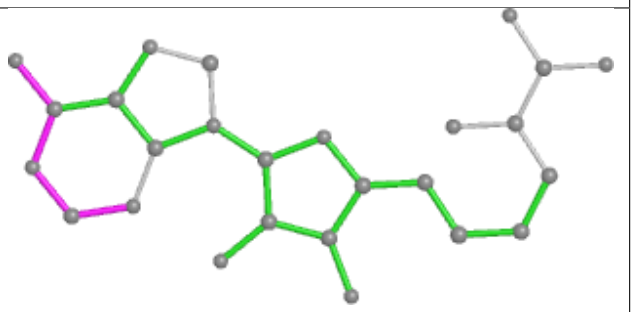
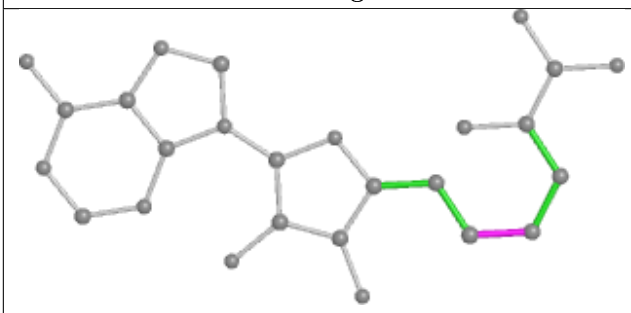
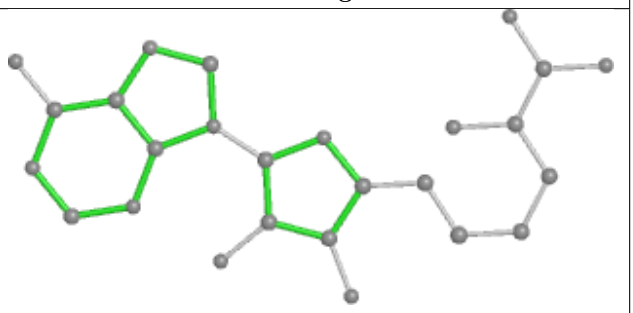
All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	8009	SAH	CB-CG-SD-C5'

There are no ring outliers.

No monomer is involved in short contacts.

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 SAH B 8009	
 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, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	481/605 (79%)	-0.11	3 (0%) 89 91	14, 28, 60, 88	0
2	B	797/937 (85%)	0.17	46 (5%) 23 28	18, 43, 83, 105	0
3	D	7/11 (63%)	0.15	0 100 100	32, 41, 51, 83	0
4	E	9/9 (100%)	0.65	2 (22%) 0 1	31, 47, 83, 85	0
All	All	1294/1562 (82%)	0.07	51 (3%) 39 44	14, 36, 78, 105	0

All (51) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	451	GLN	4.2
2	B	473	PHE	4.1
2	B	577	ILE	3.9
2	B	490	GLU	3.6
2	B	197	TRP	3.6
4	E	21	VAL	3.6
2	B	636	ASP	3.4
2	B	391	GLU	3.3
2	B	276	TYR	3.2
2	B	778	VAL	3.2
2	B	443	ALA	3.2
2	B	628	VAL	3.1
2	B	439	MET	3.0
2	B	491	ALA	3.0
2	B	578	HIS	3.0
2	B	356	PHE	3.0
2	B	569	PRO	2.9
2	B	647	ILE	2.9
2	B	262	VAL	2.9
2	B	404	SER	2.9
2	B	401	PHE	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	389	ASN	2.8
1	A	492	ARG	2.8
2	B	413	THR	2.8
2	B	2673	GLN	2.7
2	B	626	TRP	2.7
2	B	450	ASP	2.7
2	B	926	THR	2.6
2	B	427	ASP	2.6
1	A	556	ARG	2.5
4	E	22	MET	2.5
2	B	198	THR	2.4
2	B	435	ARG	2.4
2	B	338	GLN	2.4
2	B	406	PRO	2.4
2	B	630	ARG	2.3
2	B	548	ALA	2.3
2	B	924	ASN	2.3
2	B	777	GLU	2.3
2	B	405	GLN	2.2
2	B	928	LYS	2.2
2	B	624	PRO	2.2
2	B	437	ALA	2.1
2	B	415	GLN	2.1
1	A	97	ALA	2.1
2	B	651	LYS	2.1
2	B	743	GLY	2.1
2	B	730	HIS	2.1
2	B	2681	ARG	2.1
2	B	418	SER	2.1
2	B	202	ILE	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^{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	M3L	D	27	12/13	0.96	0.16	21,24,26,28	0

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

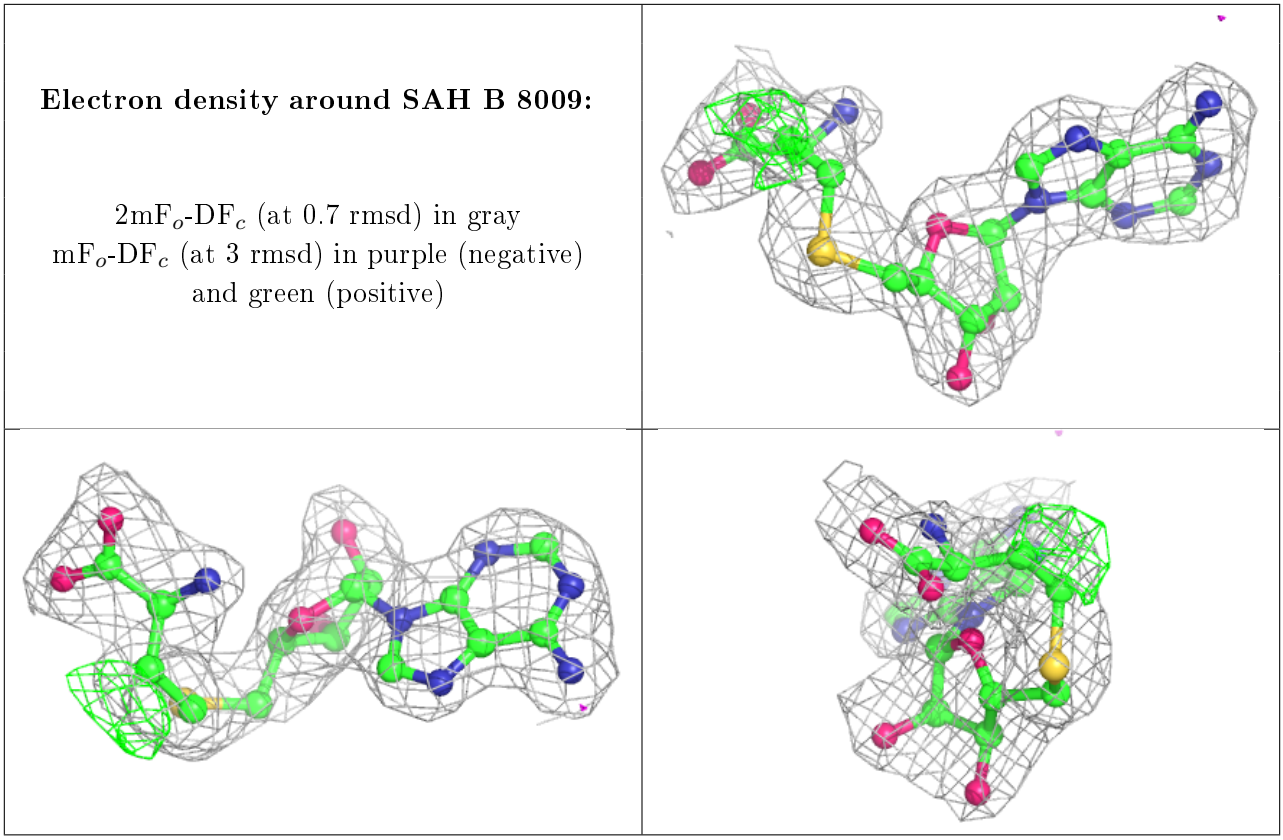
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	ZN	B	8008	1/1	0.91	0.06	73,73,73,73	0
6	SAH	B	8009	26/26	0.95	0.14	13,24,27,30	26
5	ZN	B	8005	1/1	0.99	0.09	29,29,29,29	0
5	ZN	B	8001	1/1	0.99	0.12	31,31,31,31	0
5	ZN	B	8003	1/1	0.99	0.10	34,34,34,34	0
5	ZN	B	8002	1/1	0.99	0.11	35,35,35,35	0
5	ZN	B	8004	1/1	0.99	0.08	28,28,28,28	0
5	ZN	B	8006	1/1	0.99	0.10	27,27,27,27	0
5	ZN	B	8007	1/1	1.00	0.12	24,24,24,24	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 SAH B 8009:

$2mF_o$-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.