



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 03:59 PM UTC

PDB ID : 8KAL / pdb_00008kal
Title : Crystal structure of SpyCas9 in complex with sgRNA and 17nt target DNA
Authors : Chen, Y.; Chen, J.; Liu, L.
Deposited on : 2023-08-03
Resolution : 3.16 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
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

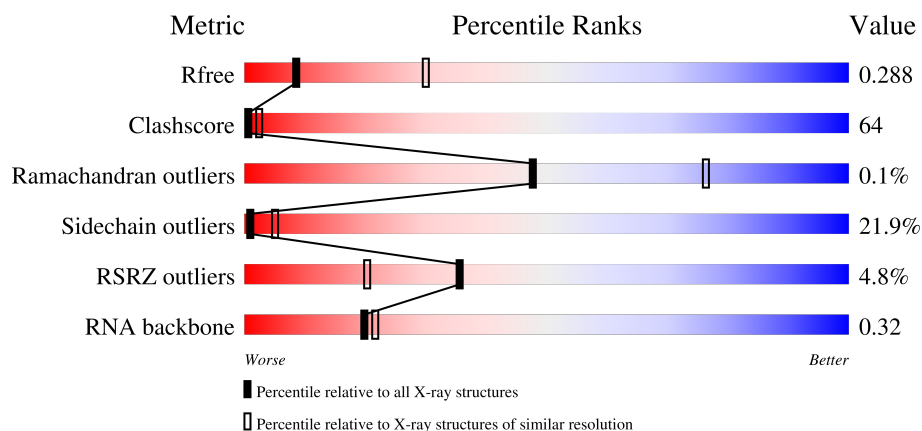
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.16 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.



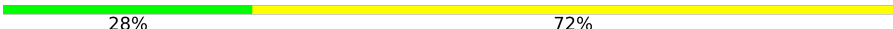
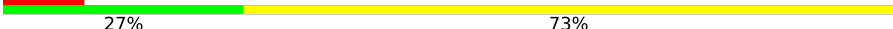
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	2361 (3.20-3.12)
Clashscore	190562	2486 (3.20-3.12)
Ramachandran outliers	187476	2405 (3.20-3.12)
Sidechain outliers	187428	2404 (3.20-3.12)
RSRZ outliers	180081	2361 (3.20-3.12)
RNA backbone	3983	1113 (3.42-2.90)

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 $\leq 5\%$. 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	98	15% 56% 24% .
1	E	98	2% 22% 54% 20% .
2	B	1368	5% 29% 49% 18% ..
2	G	1368	5% 31% 49% 16% ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	25	 24% 76%
3	H	25	 28% 72%
4	D	11	 9% 27% 73%
4	J	11	 9% 55% 45%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 27214 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 RNA chain called RNA (98-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	94	Total	C	N	O	P	0	0	0
			2009	899	362	654	94			
1	E	95	Total	C	N	O	P	0	0	0
			2029	908	365	661	95			

- Molecule 2 is a protein called CRISPR-associated endonuclease Cas9/Csn1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1326	Total	C	N	O	S	0	0	0
			10821	6892	1877	2030	22			
2	G	1326	Total	C	N	O	S	0	0	0
			10822	6892	1879	2029	22			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	10	ALA	ASP	engineered mutation	UNP Q99ZW2
B	840	ALA	HIS	engineered mutation	UNP Q99ZW2
G	10	ALA	ASP	engineered mutation	UNP Q99ZW2
G	840	ALA	HIS	engineered mutation	UNP Q99ZW2

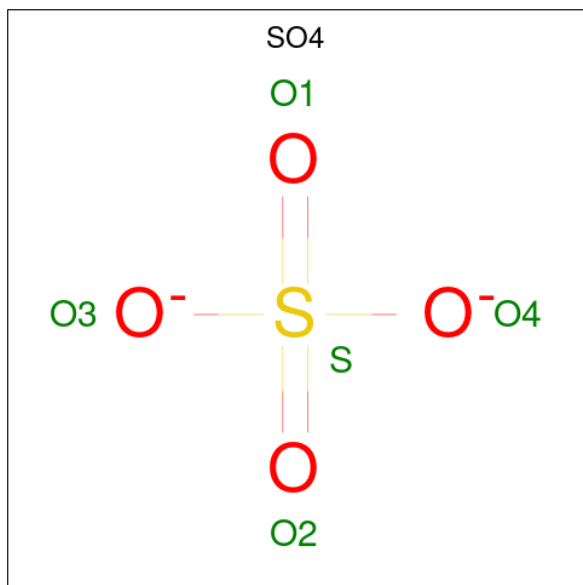
- Molecule 3 is a DNA chain called DNA (25-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	25	Total	C	N	O	P	0	0	0
			505	243	93	145	24			
3	H	25	Total	C	N	O	P	0	0	0
			505	243	93	145	24			

- Molecule 4 is a DNA chain called DNA (5'-D(*TP*TP*TP*AP*GP*GP*TP*AP*TP*TP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	11	Total	C	N	O	P	0	0	0
			225	110	37	68	10			
4	J	11	Total	C	N	O	P	0	0	0
			225	110	37	68	10			

- Molecule 5 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



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

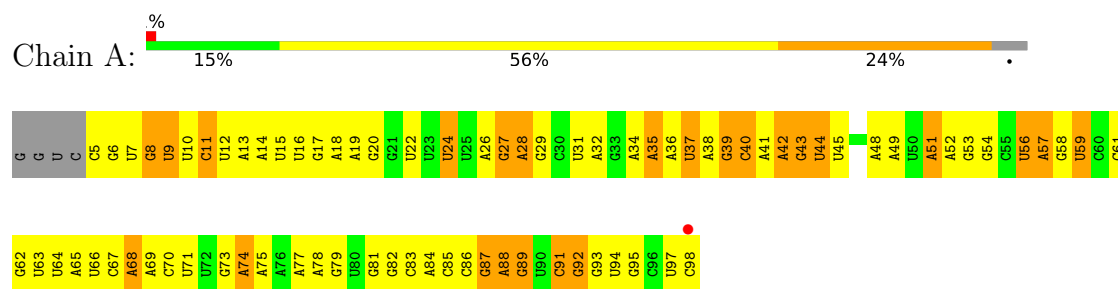
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	3	Total	O	0	0
			3	3		
6	B	26	Total	O	0	0
			26	26		
6	C	1	Total	O	0	0
			1	1		
6	E	3	Total	O	0	0
			3	3		
6	G	29	Total	O	0	0
			29	29		
6	H	1	Total	O	0	0
			1	1		

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.

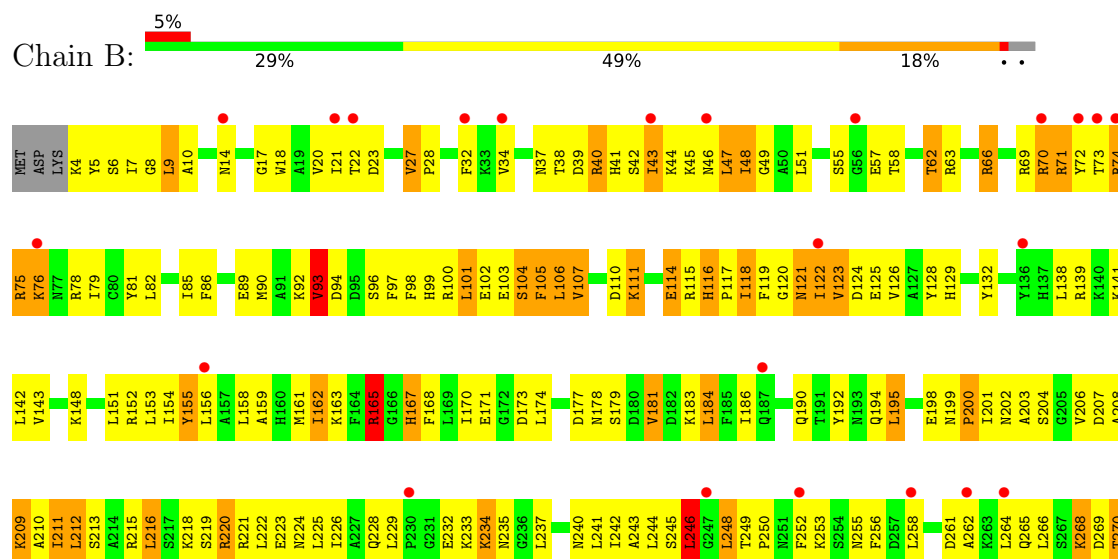
• Molecule 1: RNA (98-MER)



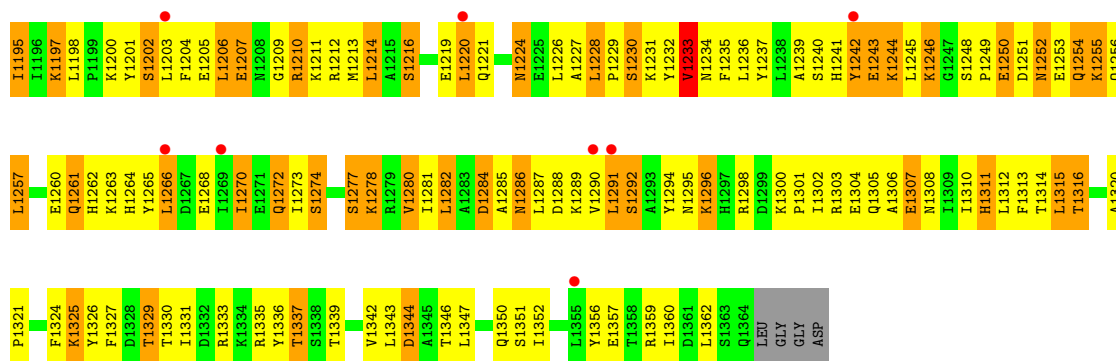
• Molecule 1: RNA (98-MER)



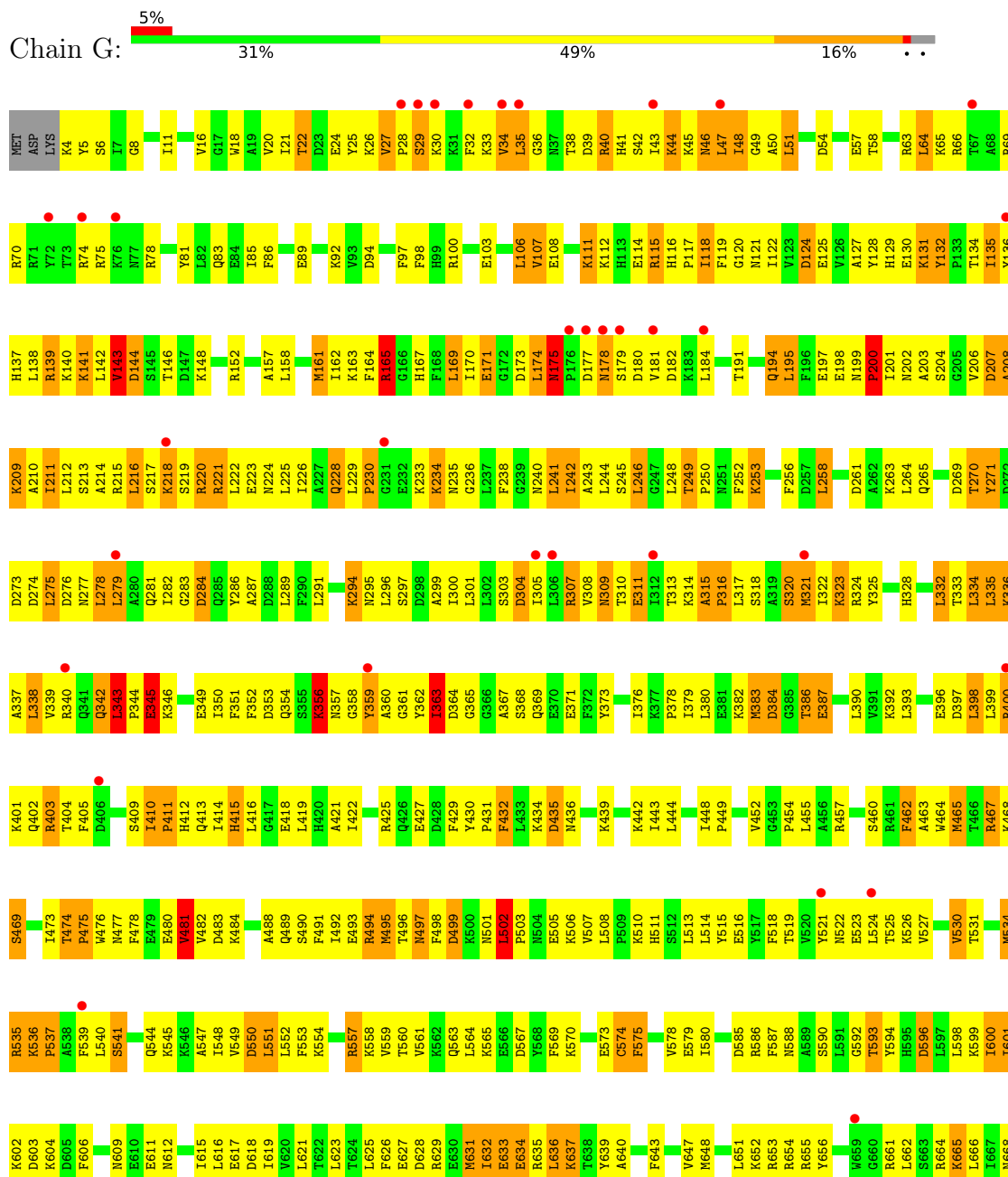
• Molecule 2: CRISPR-associated endonuclease Cas9/Csn1

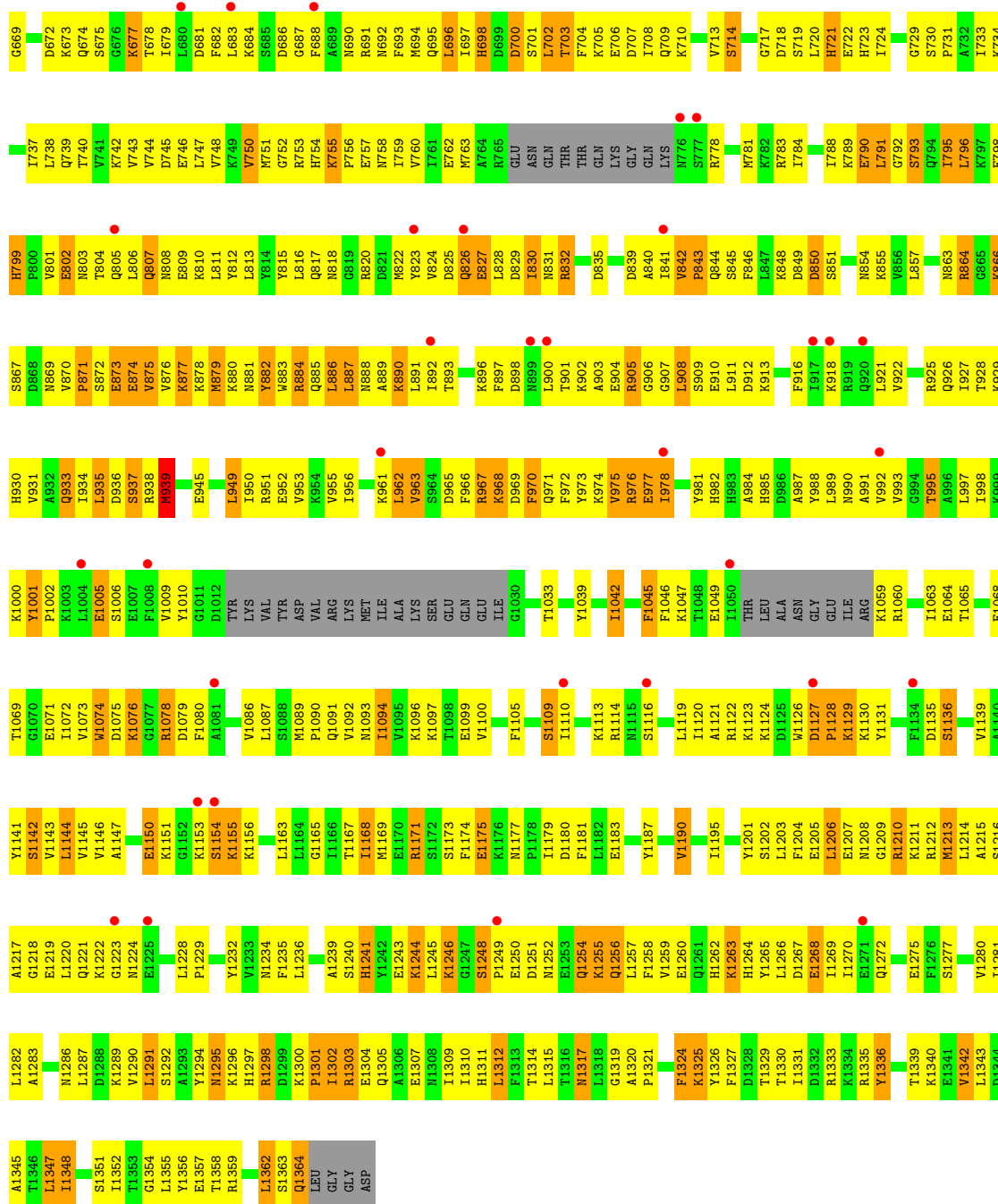


K1130	Y1131	F1134	D1135	S1136	P1137	T1138	V1139	A1140	Y1141	S1142	V1143	L1144	V1145	V1146	A1147	K1148	V1149	E1150	K1151	G1152	K1153	S1154	S1155	K1156	L1157	K1158	S1159	V1160	K1161	L1164	G1165	I1166	T1167	L1168	M1169	E1170	R1171	F1174	E1175	K1176	M1177	P1178	I1179	D1180	F1181	L1182	E1183	A1184	K1185	G1186	Y1187	K1191	L1192	D1193	L1194			
THR	LEU	ALA	ASN	GLY	GLU	ILE	ARG	K1069	P1061	L1062	I1063	E1064	G1067	V1074	D1075	K1076	G1077	R1078	D1079	F1080	A1081	T1082	V1083	V1086	V1087	S1088	R1089	Q1091	N1092	L1093	I1094	V1095	V1100	Q1101	T1102	F1105	S1106	K1107	E1108	S1109	I1110	K1113	R1114	A1121	R1122	K1123	D1127	P1128	K1129									
H983	A984	A987	Y988	L989	N990	P991	V992	T995	Y1001	P1002	E1005	S1006	E1007	F1008	V1009	D1012	TTR	VAL	ASP	VAL	ARG	LYS	MET	ILE	P1087	A1088	Q1091	SER	GLU	GLN	GLU	K961	L962	V963	S964	D965	R967	K968	F970	Q971	F972	Y973	K974	V975	R976	E977	F1045	F1046	K1047	T1048	E1049	K1129						
K848	D849	S851	N854	K855	V856	L857	T858	R859	R864	G865	K866	K867	D868	N869	V870	P871	R872	M873	N940	D944	E945	N946	D947	Y882	W883	L949	R884	Q885	L886	L887	N888	A889	K890	T893	K896	F897	S901	K902	A903	E904	R905	G907	L908	S909	E910	L911	D912	K913	A914	G915	I917							
K918	R919	Q920	E923	T924	R925	Q926	T927	T928	K929	H930	V931	A932	Q933	T934	L935	D936	S937	R938	M939	N940	E945	N946	D947	Y882	W883	L949	R884	Q885	L886	L887	N888	A889	K890	T893	K896	F897	S901	K902	A903	E904	R905	G907	L908	S909	E910	L911	D912	K913	A914	G915	I917							
K918	R919	Q920	E923	T924	R925	Q926	T927	T928	K929	H930	V931	A932	Q933	T934	L935	D936	S937	R938	M939	N940	E945	N946	D947	Y882	W883	L949	R884	Q885	L886	L887	N888	A889	K890	T893	K896	F897	S901	K902	A903	E904	R905	G907	L908	S909	E910	L911	D912	K913	A914	G915	I917							
H983	A984	A987	Y988	L989	N990	P991	V992	T995	Y1001	P1002	E1005	S1006	E1007	F1008	V1009	D1012	TTR	VAL	ASP	VAL	ARG	LYS	MET	ILE	P1087	A1088	Q1091	SER	GLU	GLN	GLU	K961	L962	V963	S964	D965	R967	K968	F970	Q971	F972	Y973	K974	V975	R976	E977	F1045	F1046	K1047	T1048	E1049	K1129						
THR	LEU	ALA	ASN	GLY	GLU	ILE	ARG	K1069	P1061	L1062	I1063	E1064	G1067	V1074	D1075	K1076	G1077	R1078	D1079	F1080	A1081	T1082	V1083	V1086	V1087	S1088	R1089	Q1091	N1092	L1093	I1094	V1095	V1100	Q1101	T1102	F1105	S1106	K1107	E1108	S1109	I1110	K1113	R1114	A1121	R1122	K1123	D1127	P1128	K1129									
K336	A337	L338	V339	R340	D341	T342	L343	P344	E345	K346	Y347	K348	E349	I350	F351	F352	K356	N357	G358	Y359	A360	C361	V362	I363	D364	Q369	F372	L373	K374	F376	L377	P378	L379	L380	E381	K382	N383	D384	G385	T386	E387	E388	L389	L390	V391	K392	K393	R394	R395	E396	H397	L398	T399	R400	A401			
Q402	R403	T404	F405	N406	D407	G408	S409	T410	A411	H412	Q413	T414	H415	L416	G417	E418	H420	A421	L422	L423	R424	R425	D426	F429	Y430	P431	F432	L433	K434	N501	L502	P503	N504	E505	K506	V507	L508	P509	K510	T512	H511	S512	L513	Y514	E515	V516	Y517	F518	T519	V520	E523	L524	T525	K526	N527	F528	A529	V464
V530	T531	G532	E533	M534	R535	K536	P537	L538	I539	F540	S541	G542	E543	Q544	K545	K546	I548	V549	D550	A551	L552	F553	K554	T555	N556	L557	K558	V559	T560	V561	K562	Q563	E566	D567	Y568	F569	K570	M571	I572	E573	E574	F575	D576	S577	V578	E579	I580	S581	G582	E583	E584	D585	R586	F587	N588	A589	S590	
L591	G592	T593	Y594	H595	D596	L597	L598	K599	I600	O601	K602	D603	K604	D605	F606	L607	D608	V609	E610	E611	N612	L613	D614	I615	L616	E617	D618	P619	V620	L621	T622	L623	F626	E627	D628	R629	E630	M631	I632	E633	E634	C635	R636	L637	L638	T639	A640	H641	L642	F643	E644	V647	K648	K649	Q650	L651		
K652	R653	G654	T655	V656	L657	G658	M659	G660	S663	R664	K665	L666	L667	F668	G669	L670	R671	E672	E673	Q674	S675	D676	K677	T678	L679	L680	D681	F682	L683	K684	S685	F686	A689	R690	R691	M692	F693	M694	Q695	L696	L697	H698	D699	T700	S701	L702	T703	F704	K705	E706	D707	L708	Q709	A710	R711	Q712		
S719	H721	T722	H723	L724	T725	L727	S730	F731	A732	L733	K734	G735	T737	L738	Q739	T740	V743	L744	D745	E746	L747	V748	V750	H751	G752	R753	H754	K755	P756	E757	N758	I759	V760	L761	E762	M763	A764	R765	GLU	ASN	GLN	THR	GLN	THR	GLN	L776	S777	R778	E779									
R780	M781	K782	L783	T784	E785	G786	T787	L788	K789	G790	L791	G792	S793	Q794	T795	L796	P800	V801	L806	K810	L811	Y812	L813	Y814	Y815	L816	Q817	N818	G819	R820	D821	M822	Y823	V824	E827	L828	D829	L830	N831	R832	L833	S834	D835	Y836	D837	V838	D839	A840	L841	P842	Q844	S845	F846	L847				
K848	D849	S851	N854	K855	V856	L857	T858	R859	R864	G865	K866	K867	D868	N869	V870	P871	R872	M873	N940	D944	E945	N946	D947	Y882	W883	L949	R884	Q885	L886	L887	N888	A889	K890	T893	K896	F897	S901	K902	A903	E904	R905	G907	L908	S909	E910	L911	D912	K913	A914	G915	I917							
K918	R919	Q920	E923	T924	R925	Q926	T927	T928	K929	H930	V931	A932	Q933	T934	L935	D936	S937	R938	M939	N940	E945	N946	D947	Y882	W883	L949	R884	Q885	L886	L887	N888	A889	K890	T893	K896	F897	S901	K902	A903	E904	R905	G907	L908	S909	E910	L911	D912	K913	A914	G915	I917							
H983	A984	A987	Y988	L989	N990	P991	V992	T995	Y1001	P1002	E1005	S1006	E1007	F1008	V1009	D1012	TTR	VAL	ASP	VAL	ARG	LYS	MET	ILE	P1087	A1088	Q1091	SER	GLU	GLN	GLU	K961	L962	V963	S964	D965	R967	K968	F970	Q971	F972	Y973	K974	V975	R976	E977	F1045	F1046	K1047	T1048	E1049	K1129						
THR	LEU	ALA	ASN	GLY	GLU	ILE	ARG	K1069	P1061	L1062	I1063	E1064	G1067	V1074	D1075	K1076	G1077	R1078	D1079	F1080	A1081	T1082	V1083	V1086	V1087	S1088	R1089	Q1091	N1092	L1093	I1094	V1095	V1100	Q1101	T1102	F1105	S1106	K1107	E1108	S1109	I1110	K1113	R1114	A1121	R1122	K1123	D1127	P1128	K1129									
K1130	Y1131	F1134	D1135	S1136	P1137	T1138	V1139	A1140	Y1141	S1142	V1143	L1144	V1145	V1146	A1147	K1148	V1149	E1150	K1151	G1152	K1153	S1154	S1155	K1156	L1157	K1158	S1159	V1160	K1161	L1164	G1165	I1166	T1167	L1168	M1169	E1170	R1171	F1174	E1175	K1176	M1177	P1178	I1179	D1180	F1181	L1182	E1183	A1184	K1185	G1186	Y1187	K1191	L1192	D1193	L1194			



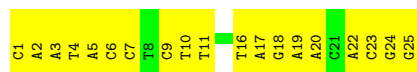
• Molecule 2: CRISPR-associated endonuclease Cas9/Csn1





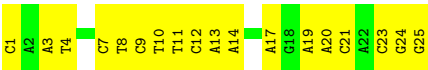
• Molecule 3: DNA (25-MER)

Chain C: 24% 76%

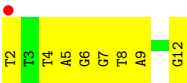


• Molecule 3: DNA (25-MER)

Chain H: 28% 72%



● Molecule 4: DNA (5'-D(*TP*TP*TP*AP*GP*GP*TP*AP*TP*TP*G)-3')



● Molecule 4: DNA (5'-D(*TP*TP*TP*AP*GP*GP*TP*AP*TP*TP*G)-3')



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	362.05Å 70.96Å 200.10Å 90.00° 101.52° 90.00°	Depositor
Resolution (Å)	50.17 – 3.16 50.17 – 3.16	Depositor EDS
% Data completeness (in resolution range)	65.6 (50.17-3.16) 65.6 (50.17-3.16)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.79 (at 3.19Å)	Xtriage
Refinement program	PHENIX 1.17.1 _3660	Depositor
R, R_{free}	0.251 , 0.288 0.248 , 0.288	Depositor DCC
R_{free} test set	2797 reflections (3.24%)	wwPDB-VP
Wilson B-factor (Å ²)	58.3	Xtriage
Anisotropy	0.329	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 58.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.39$, $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.86	EDS
Total number of atoms	27214	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 43.87 % 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 1.6562e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical 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

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	0.47	0/2249	0.69	0/3503
1	E	0.47	0/2271	0.71	0/3537
2	B	0.81	32/11013 (0.3%)	1.08	95/14802 (0.6%)
2	G	0.79	22/11013 (0.2%)	1.03	71/14799 (0.5%)
3	C	0.63	0/566	0.77	0/870
3	H	0.67	0/566	0.76	0/870
4	D	0.63	0/251	0.74	0/387
4	J	0.60	0/251	0.65	0/387
All	All	0.75	54/28180 (0.2%)	0.98	166/39155 (0.4%)

The worst 5 of 54 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	27	VAL	CA-C	13.89	1.67	1.52
2	G	502	LEU	C-N	8.76	1.40	1.33
2	G	1001	TYR	C-N	7.79	1.40	1.33
2	G	430	TYR	C-N	6.59	1.40	1.34
2	B	755	LYS	C-N	6.26	1.40	1.33

The worst 5 of 166 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	684	LYS	N-CA-C	-14.17	86.15	109.96
2	G	506	LYS	N-CA-C	-14.14	86.20	109.96
2	G	507	VAL	N-CA-CB	-10.37	91.67	110.95
2	B	906	GLY	N-CA-C	-10.07	100.30	112.48
2	B	462	PHE	N-CA-C	8.96	121.13	111.36

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

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	2009	0	1009	111	0
1	E	2029	0	1020	126	0
2	B	10821	0	10948	1541	0
2	G	10822	0	10967	1637	0
3	C	505	0	283	31	0
3	H	505	0	283	23	0
4	D	225	0	129	18	0
4	J	225	0	129	6	0
5	B	5	0	0	0	0
5	G	5	0	0	1	0
6	A	3	0	0	0	0
6	B	26	0	0	7	0
6	C	1	0	0	0	0
6	E	3	0	0	1	0
6	G	29	0	0	10	0
6	H	1	0	0	0	0
All	All	27214	0	24768	3310	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 64.

The worst 5 of 3310 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)
2:G:530:VAL:CG1	2:G:537:PRO:HB3	1.21	1.65
2:B:181:VAL:HG11	2:B:300:ILE:CD1	1.14	1.59
2:G:279:LEU:CD1	2:G:287:ALA:HB2	1.28	1.59
2:G:279:LEU:HD11	2:G:287:ALA:CB	1.21	1.59
2:G:870:VAL:CG2	2:G:908:LEU:CD2	1.77	1.59

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	
2	B	1318/1368 (96%)	1279 (97%)	38 (3%)	1 (0%)	48	76
2	G	1318/1368 (96%)	1273 (97%)	43 (3%)	2 (0%)	43	71
All	All	2636/2736 (96%)	2552 (97%)	81 (3%)	3 (0%)	48	76

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	200	PRO
2	G	200	PRO
2	G	230	PRO

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	
2	B	1185/1225 (97%)	918 (78%)	267 (22%)	1	4
2	G	1186/1225 (97%)	933 (79%)	253 (21%)	1	5
All	All	2371/2450 (97%)	1851 (78%)	520 (22%)	1	5

5 of 520 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	G	1073	VAL
2	G	1150	GLU
2	G	1069	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	979	ASN
2	B	964	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 55 such sidechains are listed below:

Mol	Chain	Res	Type
2	G	235	ASN
2	G	497	ASN
2	G	1350	GLN
2	G	1256	GLN
2	G	281	GLN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	93/98 (94%)	31 (33%)	4 (4%)
1	E	94/98 (95%)	24 (25%)	3 (3%)
All	All	187/196 (95%)	55 (29%)	7 (3%)

5 of 55 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	9	U
1	A	11	C
1	A	20	G
1	A	24	U
1	A	27	G

5 of 7 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	68	A
1	E	8	G
1	E	42	A
1	E	27	G
1	A	42	A

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

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

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

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	$\# Z > 2$
5	SO4	B	1401	-	4,4,4	1.04	0	6,6,6	1.39	1 (16%)
5	SO4	G	1401	-	4,4,4	1.04	0	6,6,6	1.39	1 (16%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	1401	SO4	O4-S-O3	3.06	125.39	108.54
5	G	1401	SO4	O4-S-O3	3.06	125.39	108.54

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	G	1401	SO4	1	0

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, 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	94/98 (95%)	-0.07	1 (1%) 78 60	27, 56, 107, 151	0
1	E	95/98 (96%)	-0.20	2 (2%) 63 42	31, 56, 109, 125	0
2	B	1326/1368 (96%)	0.26	69 (5%) 33 19	24, 57, 90, 137	0
2	G	1326/1368 (96%)	0.27	66 (4%) 34 20	25, 57, 89, 123	0
3	C	25/25 (100%)	0.18	0 100 100	32, 47, 96, 107	0
3	H	25/25 (100%)	0.05	0 100 100	34, 47, 99, 112	0
4	D	11/11 (100%)	0.38	1 (9%) 15 8	48, 65, 121, 133	0
4	J	11/11 (100%)	0.56	1 (9%) 15 8	35, 60, 114, 132	0
All	All	2913/3004 (96%)	0.24	140 (4%) 35 20	24, 57, 93, 151	0

The worst 5 of 140 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	G	1134	PHE	5.7
2	G	76	LYS	4.3
2	B	539	PHE	4.3
2	B	46	ASN	4.2
2	G	305	ILE	4.2

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 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, 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($\text{\AA}^2$)	Q<0.9
5	SO4	G	1401	5/5	0.80	0.10	78,81,95,107	0
5	SO4	B	1401	5/5	0.84	0.24	30,30,30,30	0

6.5 Other polymers [i](#)

There are no such residues in this entry.