



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 05:27 AM UTC

PDB ID : 8KAH / pdb_00008kah
Title : Crystal structure of SpyCas9-crRNA-tracrRNA complex bound to 18nt target DNA
Authors : Chen, Y.; Chen, J.; Liu, L.
Deposited on : 2023-08-03
Resolution : 3.36 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
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

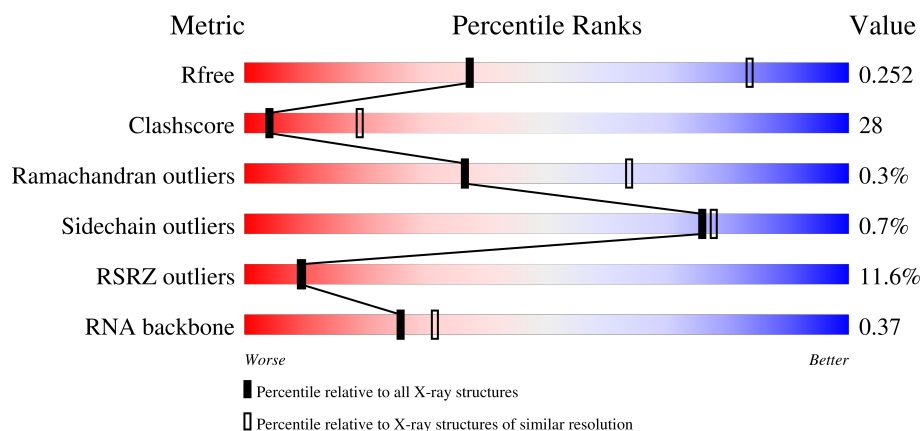
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.36 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	1099 (3.40-3.32)
Clashscore	190562	1116 (3.40-3.32)
Ramachandran outliers	187476	1101 (3.40-3.32)
Sidechain outliers	187428	1101 (3.40-3.32)
RSRZ outliers	180081	1099 (3.40-3.32)
RNA backbone	3983	1110 (3.72-3.00)

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	34	3% 29% 41% 29%
1	E	34	3% 21% 53% 21% 6%
2	B	1368	8% 53% 43% ..
2	F	1368	16% 51% 45% ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	26	4% 46% 54%
3	G	26	4% 42% 58%
4	D	11	18% 55% 45%
4	H	11	9% 45% 55%
5	I	65	2% 38% 40% 17% • •
5	J	65	5% 28% 52% 17% •

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 27025 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 (34-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	34	Total	C	N	O	P	0	0	0
			725	325	127	239	34			
1	E	32	Total	C	N	O	P	0	0	0
			685	307	123	223	32			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1322	Total	C	N	O	S	0	0	0
			10732	6827	1864	2019	22			
2	F	1327	Total	C	N	O	S	0	0	0
			10695	6814	1845	2014	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
F	10	ALA	ASP	engineered mutation	UNP Q99ZW2
F	840	ALA	HIS	engineered mutation	UNP Q99ZW2

- Molecule 3 is a DNA chain called DNA (26-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	26	Total	C	N	O	P	0	0	0
			521	254	85	157	25			
3	G	26	Total	C	N	O	P	0	0	0
			521	254	85	157	25			

- 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	H	11	Total	C	N	O	P	0	0	0
			225	110	37	68	10			

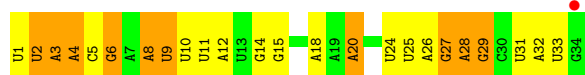
- Molecule 5 is a RNA chain called RNA (65-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	I	63	Total	C	N	O	P	0	0	0
			1348	603	245	437	63			
5	J	63	Total	C	N	O	P	0	0	0
			1348	603	245	437	63			

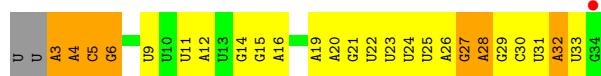
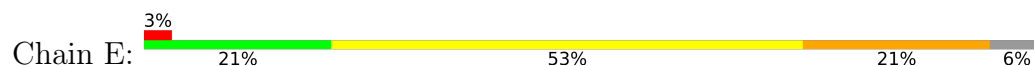
3 Residue-property plots

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.

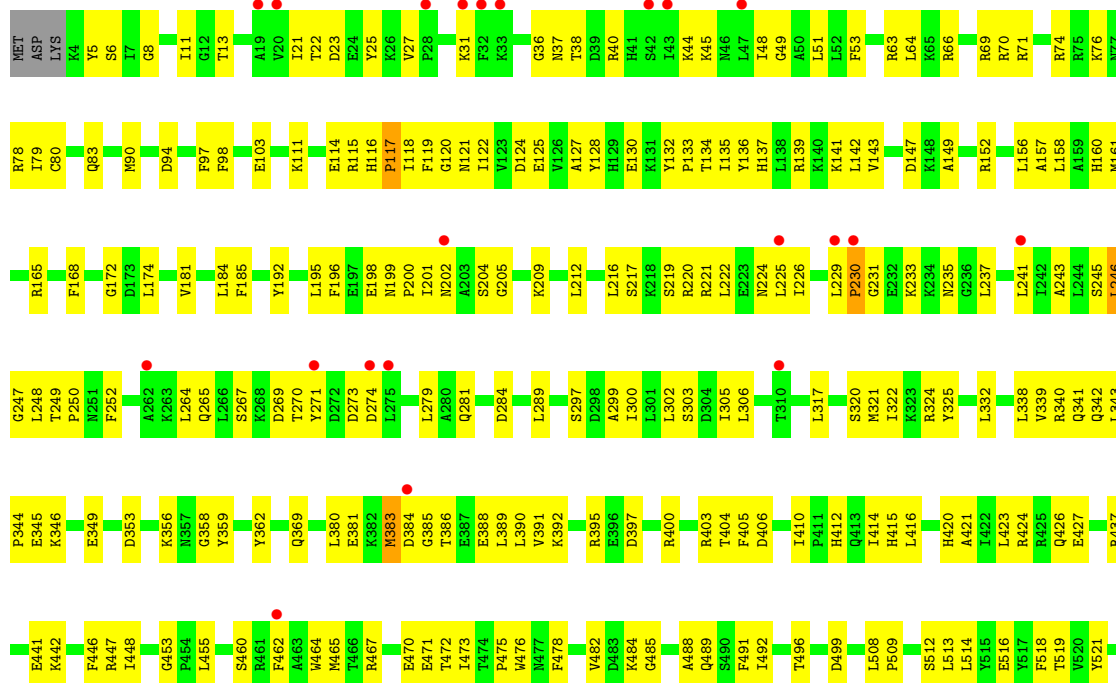
- Molecule 1: RNA (34-MER)

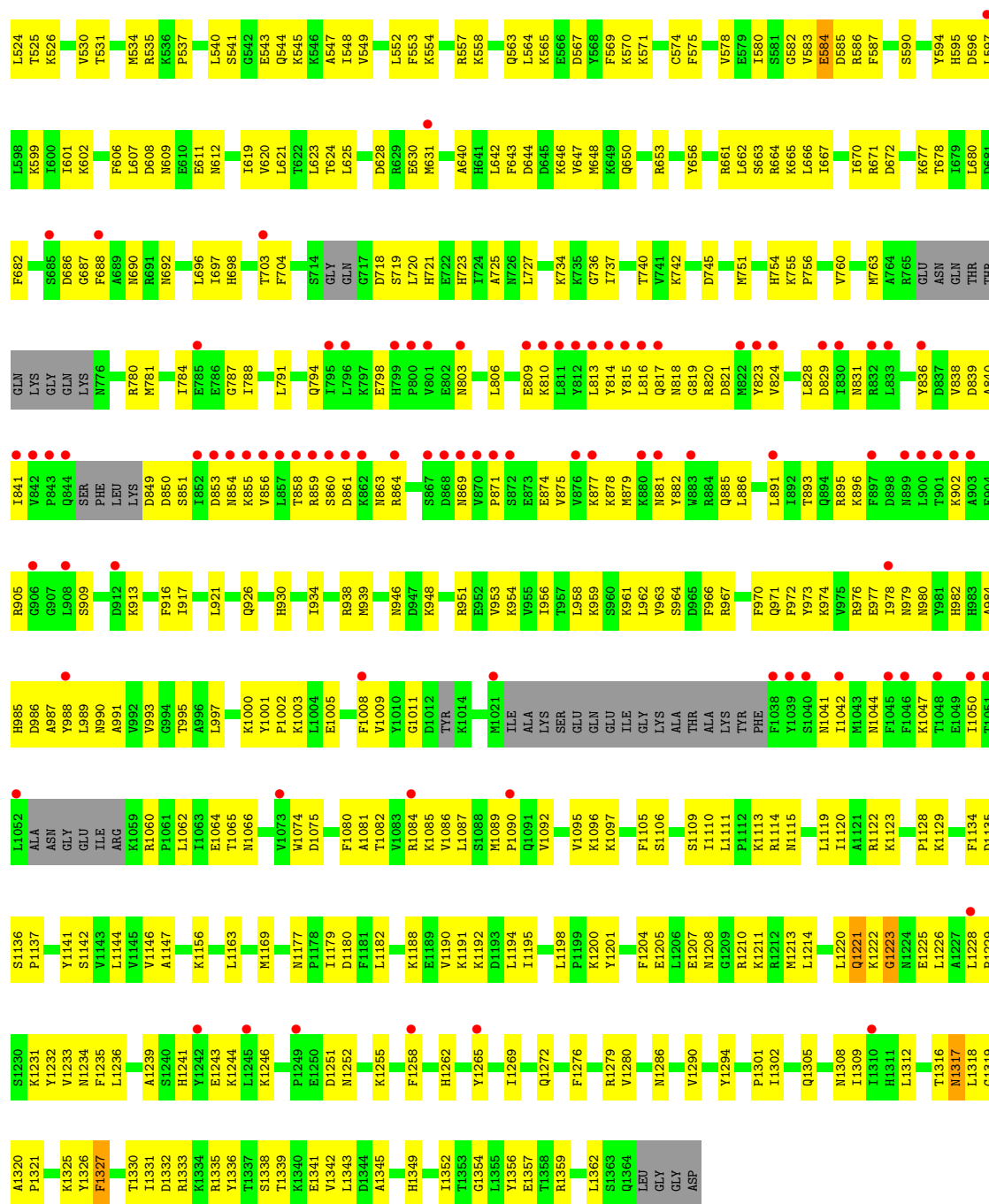


- Molecule 1: RNA (34-MER)

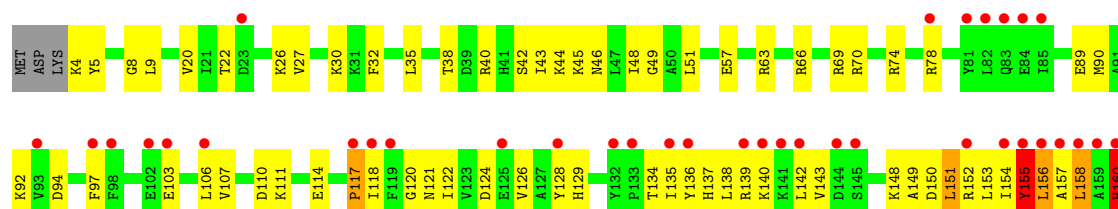


- Molecule 2: CRISPR-associated endonuclease Cas9/Csn1

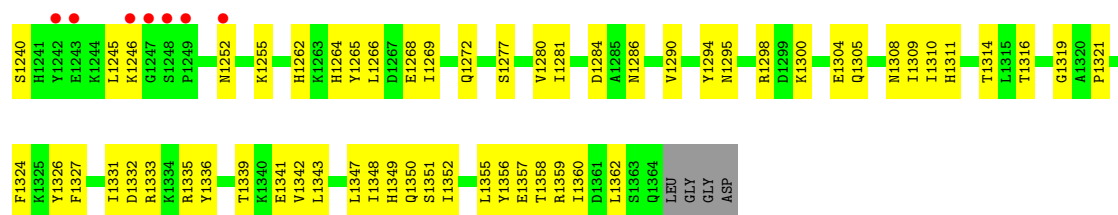




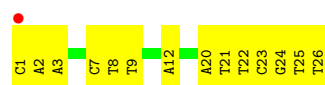
• Molecule 2: CRISPR-associated endonuclease Cas9/Csn1



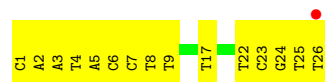
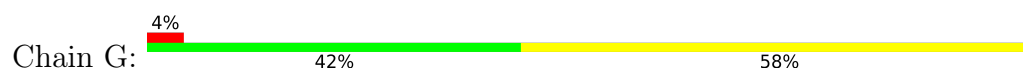
E1162	T1065	D912	Y836	GLN	I679	V527	E441	S368	I300	P230	M161
L1163		K913	I841	LYS	L880	K528	K442	Q369	L301	G231	I162
I1166	V1073	F916	V842	GLY	D681	Y529	I443	E370	L302	E232	F163
T1167	W1074	K999	P843	GLN	F682	V530	L444	E371	S303	K233	K164
K1168	D1075	K917	P843	LYS	L683	T531	T445	F372	K304	K234	R165
M1169	K1000	K918	K844		K684		F446	Y373	L305	G236	G166
E1170	Y1001	K919	S845	N776	S685	P537	R447	K374	L306	L237	H167
	P1002	G920	F846	S777	D886	A536	I448	F375	K307	L238	F168
		L921	K847	R780	G687	F539	Y451	I376	V308	L239	L169
		V922	K848	M781	L615	L540		L380	N309	G239	I170
		E923	D849	F688	L616	S541		E381		G240	
		E924	D850	I784	I619	G542	P454	K382	L312	L241	L174
		G926	S851	I788	I620		L455	M383	T313	T242	
		I927	I852		L621	K545			K314	A243	
		T928	D853	L791	T622	K546	M465	L389	A315	L244	D177
		T928	D853	G792	L625	V549	M465	P316	P316	L245	M178
		G929	K855	L796	L626	D550	T466	L390	L317	L246	S179
		H930	V856	S793	F626	L551	R467	K467	S318	G247	D180
		D936	R859	G794	E627	L552	K468	V391	A319	D248	V181
		M939	S860	I795	D628	F553	S469	D397	S320	T249	D182
		N940	D861	L797	M631	K554	E471	L398	K321	P250	K183
		Y943	N863	E798	L702	T555	N477	L399	K322	K253	F185
		L949	S867	E799	T703	N556	F478	K401	R324		I186
		T950	D868	P800	F704	R557		Q402	Y325	F256	Q187
		G951	N869	V801		K558	V482	R403	D326	D257	V188
		E952	P871	L806	GLY		D482	T404	E327	L258	V189
		G953	E374	Q807	GLN		D483	F405	L334	L259	
		K954	K878	K810	G717		F405	H411	L335	L260	
		Y955	M879	L811	I737		D484	P412	K336	S267	N199
		I956	Y882	L812	L738		D485	H412	A337	K268	P200
		V963	W853	L813	Q739		A488	I413	L338	D269	L201
		R967	L883	Y814	T740		Q489	I414	L339	Y271	L202
		Q971	L886	Y815	I744		I492	I415	V339	D270	A203
		F972	L887	L816	V744		M495	H416	R340	D271	S204
		G973	L891	Q817	K749		T496	G417	L343	D272	GLY
		K974	I892	G818	V750		D499	E418		D273	VAL
		F975	T893	G819	K755		E505	H419	K346	D274	D207
		R976	K894	R820	M751		K506	A421	Y347	A208	A209
		E977	R895	R821	K755		P509	I422	K348	A210	L211
		T978	K896	Y823	N758		L513	L423	E349	L275	L212
		G979	F937	D825	I759		L514	R424	I350	L276	S213
		N980	D896	Q826	V760		F515	Q426	F352	L277	ALA
		Y981	D899	Q827	I761		E516	E427	D353	L278	ARG
		H982	S899	E927	E762		Y517	D428	Y359	L279	LEU
		A984	T901	L828	R671		F518	F429	A360		SER
		H985	G907	R830	D672		T519	F432	G361	I282	LYS
		D986	L968	R831	R765		V520	K432	Y362	Y286	SER
		A987	S909	R832	GLU		N521	L433	A292	A287	R220
		Y988	E910	L833	ASN		E522	K434	N295	D288	R221
		N990	L911	D835	THR		E523	R437	G365	L289	L222
							L524	I440	D364		E223
							T525		G366		N224
							K526		A367		L225
											L226
											A227
											Q228
											L229



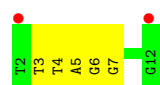
• Molecule 3: DNA (26-MER)



• Molecule 3: DNA (26-MER)



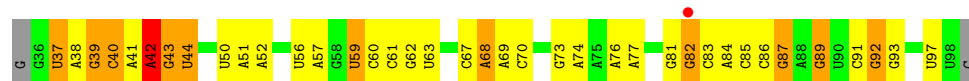
• Molecule 4: DNA (5'-D(*TP*TP*TP*AP*GP*GP*TP*AP*TP*TP*G)-3')



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

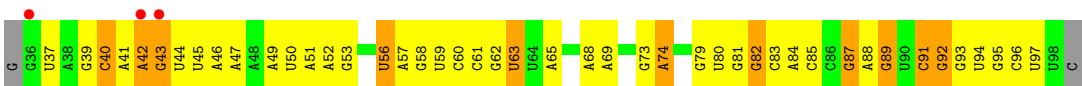


• Molecule 5: RNA (65-MER)



• Molecule 5: RNA (65-MER)





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	145.21Å 130.60Å 148.73Å 90.00° 104.08° 90.00°	Depositor
Resolution (Å)	48.23 – 3.36 48.23 – 3.36	Depositor EDS
% Data completeness (in resolution range)	59.4 (48.23-3.36) 73.8 (48.23-3.36)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.34 (at 3.33Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.230 , 0.255 0.229 , 0.252	Depositor DCC
R_{free} test set	1990 reflections (2.58%)	wwPDB-VP
Wilson B-factor (Å ²)	56.7	Xtriage
Anisotropy	0.207	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 47.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	0.069 for l,-k,h	Xtriage
F_o, F_c correlation	0.84	EDS
Total number of atoms	27025	wwPDB-VP
Average B, all atoms (Å ²)	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.53% of the height of the origin peak. No significant pseudotranslation is detected.*

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](#)

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	0.59	0/811	0.92	0/1261
1	E	0.54	0/767	0.83	0/1193
2	B	0.56	2/10915 (0.0%)	0.85	6/14673 (0.0%)
2	F	0.56	1/10879 (0.0%)	0.84	7/14636 (0.0%)
3	C	0.67	0/581	0.89	1/893 (0.1%)
3	G	0.59	0/581	0.86	0/893
4	D	0.74	0/251	0.85	0/387
4	H	0.70	0/251	0.82	0/387
5	I	0.57	0/1509	0.82	1/2350 (0.0%)
5	J	0.52	0/1509	0.80	0/2350
All	All	0.57	3/28054 (0.0%)	0.84	15/39023 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	1089	MET	C-N	-9.60	1.24	1.34
2	B	1320	ALA	C-N	5.30	1.40	1.33
2	B	1321	PRO	N-CD	5.21	1.55	1.47

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1223	GLY	N-CA-C	9.95	125.51	112.77
2	B	246	LEU	N-CA-C	-7.58	99.81	110.50
2	F	211	ILE	N-CA-C	7.55	118.32	110.62
2	B	1320	ALA	CA-C-N	-6.56	113.00	119.76
2	B	1320	ALA	C-N-CA	-6.56	113.00	119.76

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	725	0	362	40	0
1	E	685	0	342	44	0
2	B	10732	0	10828	538	1
2	F	10695	0	10736	719	0
3	C	521	0	299	20	0
3	G	521	0	299	20	0
4	D	225	0	129	3	0
4	H	225	0	129	10	0
5	I	1348	0	678	40	0
5	J	1348	0	678	62	0
All	All	27025	0	24480	1411	1

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 28.

The worst 5 of 1411 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:F:142:LEU:CD1	2:F:154:ILE:HG12	1.37	1.54
2:F:142:LEU:HD11	2:F:154:ILE:CG1	1.39	1.51
2:F:142:LEU:CD2	2:F:154:ILE:HD11	1.41	1.48
2:F:142:LEU:CD2	2:F:154:ILE:CD1	1.94	1.45
2:B:1222:LYS:O	2:B:1318:LEU:CD2	1.66	1.42

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:202:ASN:OD1	2:B:541:SER:OG[2_545]	2.07	0.13

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	
2	B	1308/1368 (96%)	1268 (97%)	36 (3%)	4 (0%)	36	63
2	F	1313/1368 (96%)	1267 (96%)	41 (3%)	5 (0%)	30	58
All	All	2621/2736 (96%)	2535 (97%)	77 (3%)	9 (0%)	36	63

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	1042	ILE
2	F	585	ASP
2	F	1011	GLY
2	F	1020	LYS
2	B	1327	PHE

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	
2	B	1170/1225 (96%)	1162 (99%)	8 (1%)	76	77
2	F	1155/1225 (94%)	1146 (99%)	9 (1%)	73	76
All	All	2325/2450 (95%)	2308 (99%)	17 (1%)	76	77

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	686	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	977	GLU
2	B	1317	ASN
2	F	150	ASP
2	F	151	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	178	ASN
2	F	1252	ASN
2	F	501	ASN
2	F	1317	ASN
2	F	982	HIS

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	33/34 (97%)	11 (33%)	4 (12%)
1	E	32/34 (94%)	11 (34%)	2 (6%)
5	I	62/65 (95%)	19 (30%)	1 (1%)
5	J	62/65 (95%)	18 (29%)	1 (1%)
All	All	189/198 (95%)	59 (31%)	8 (4%)

5 of 59 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	2	U
1	A	3	A
1	A	4	A
1	A	6	G
1	A	9	U

5 of 8 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
5	J	42	A
5	I	42	A
1	E	3	A
1	A	28	A
1	E	27	G

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](#)

There are no ligands in this entry.

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	34/34 (100%)	-0.21	1 (2%) 53 39	18, 45, 134, 169	0
1	E	32/34 (94%)	0.54	1 (3%) 51 38	38, 77, 181, 216	0
2	B	1322/1368 (96%)	0.51	111 (8%) 17 14	14, 71, 195, 218	0
2	F	1327/1368 (97%)	0.87	217 (16%) 4 5	13, 92, 149, 193	0
3	C	26/26 (100%)	0.01	1 (3%) 44 31	27, 50, 107, 115	0
3	G	26/26 (100%)	0.17	1 (3%) 44 31	45, 63, 112, 124	0
4	D	11/11 (100%)	0.45	2 (18%) 3 4	41, 53, 134, 168	0
4	H	11/11 (100%)	0.66	1 (9%) 15 13	41, 59, 110, 175	0
5	I	63/65 (96%)	-0.15	1 (1%) 70 55	20, 74, 126, 157	0
5	J	63/65 (96%)	-0.04	3 (4%) 35 25	36, 56, 156, 192	0
All	All	2915/3008 (96%)	0.63	339 (11%) 9 9	13, 77, 168, 218	0

The worst 5 of 339 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	305	ILE	9.8
1	E	34	G	8.9
2	F	1243	GLU	8.9
2	F	362	TYR	7.6
2	B	900	LEU	6.8

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.