



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 04:06 PM UTC

PDB ID : 5WTJ / pdb\_00005wtj  
Title : Crystal structure of an endonuclease  
Authors : Liu, L.; Wang, Y.  
Deposited on : 2016-12-13  
Resolution : 3.50 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 : 4-5-2 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

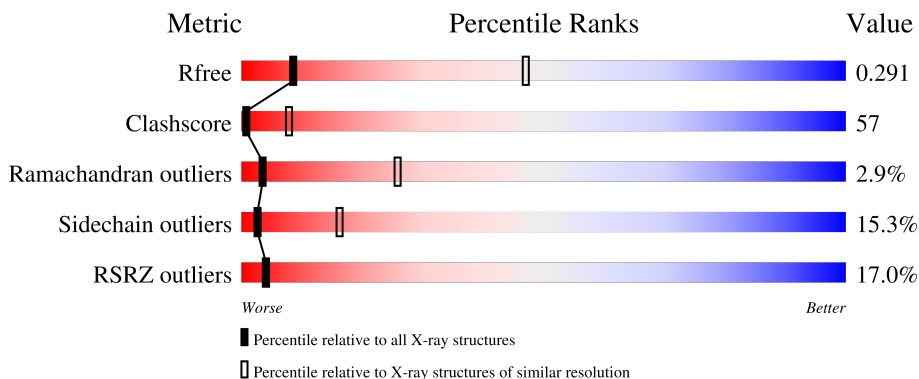
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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	1085 (3.54-3.46)
Clashscore	190562	1140 (3.54-3.46)
Ramachandran outliers	187476	1113 (3.54-3.46)
Sidechain outliers	187428	1114 (3.54-3.46)
RSRZ outliers	180081	1084 (3.54-3.46)

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	1397	
1	B	1397	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 15952 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 CRISPR-associated endoribonuclease C2c2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	1019	8013	5109	1350	1538	5	11	0	0	0
1	B	1012	7938	5055	1326	1541	6	10	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1390	LEU	-	expression tag	UNP P0DOC6
A	1391	GLU	-	expression tag	UNP P0DOC6
A	1392	HIS	-	expression tag	UNP P0DOC6
A	1393	HIS	-	expression tag	UNP P0DOC6
A	1394	HIS	-	expression tag	UNP P0DOC6
A	1395	HIS	-	expression tag	UNP P0DOC6
A	1396	HIS	-	expression tag	UNP P0DOC6
A	1397	HIS	-	expression tag	UNP P0DOC6
B	1390	LEU	-	expression tag	UNP P0DOC6
B	1391	GLU	-	expression tag	UNP P0DOC6
B	1392	HIS	-	expression tag	UNP P0DOC6
B	1393	HIS	-	expression tag	UNP P0DOC6
B	1394	HIS	-	expression tag	UNP P0DOC6
B	1395	HIS	-	expression tag	UNP P0DOC6
B	1396	HIS	-	expression tag	UNP P0DOC6
B	1397	HIS	-	expression tag	UNP P0DOC6

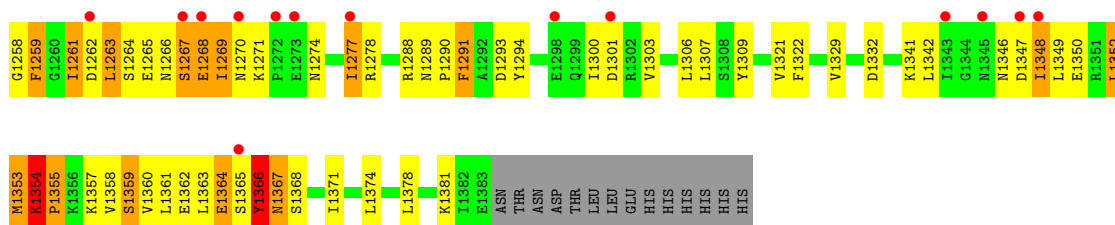
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	O	0	0
			1	1		





K1164	K1165	R1166	F1170	F1171	L1174	N1175	K1176	L1177	E1178	L1188	Q1191	M1192	A1193	R1194	F1195	E1196	R1197	D1198	M1199	H1200	Y1201	I1202	V1203	M1204	G1205	L1206	R1207	E1208	I1212	K1213	L1214	T1219	S1222	R1223	A1224	K1227	G1230	S1231	Y1241	Y1249	K1250	K1251	F1252	E1253	K1254	I1255	C1256	Y1257														
I1099	M1102	K1103	E1036	S1037	E1038	N1039	E1040	M1041	A1109	L1110	F1043	L1044	E1045	L1046	Y1047	R1048	P1049	K1050	E1051	R1052	K1053	M1054	E1055	L1056	Y1057	I1058	Y1059	K1060	L1061	M1062	L1063	I1067	G1068	M1069	P1070	M1071	F1072	I1075	Y1076	G1077	L1078	I1079	I1083	K1084	L1087	K1088	F1090	L1091	F1092	M1093	L1094	D1095	G1096	K1097	L1031	M1098						
K954	L955	E956	K957	I958	VAL	ILE	PHE	ASP	GLU	THR	LYS	PHE	E968	R969	D970	I975	L976	Q977	D978	E979	Q980	R981	K982	L983	S984	N985	R986	N987	D990	L991	V995	D996	I999	D1003	N935	I1006	K1007	L1011	I1015	R941	K942	D943	L1021	K1022	K1023	K1026	E1027	I1028	D1029	M1030	L1031	I1032										
E1033	D1034	M1035	E1036	S1037	E1038	N1039	E1040	M1041	A1109	L1110	F1043	L1044	E1045	L1046	Y1047	R1048	P1049	K1050	E1051	R1052	K1053	M1054	E1055	L1056	Y1057	I1058	Y1059	K1060	L1061	M1062	L1063	I1067	G1068	M1069	P1070	M1071	F1072	I1075	Y1076	G1077	L1078	I1079	I1083	K1084	L1087	K1088	F1090	L1091	F1092	M1093	L1094	D1095	G1096	K1097	L1031	M1098						
R888	M889	E890	C891	E892	T893	E894	M895	M896	N897	L898	GLU	THR	LYS	PHE	E968	R969	D970	I975	L976	Q977	D978	E979	Q980	R981	K982	L983	S984	N985	R986	N987	D990	L991	V995	D996	I999	D1003	N935	I1006	K1007	L1011	I1015	R941	K942	D943	L1021	K1022	K1023	K1026	E1027	I1028	D1029	M1030	L1031	I1032								
K816	T817	E818	Y818	I821	T822	V823	R824	T825	K828	T829	L830	V831	I832	F833	N834	D835	F836	Y838	I842	L845	L846	N847	R848	F849	D850	L851	L852	K854	I855	R856	T862	M865	L866	S869	N873	I874	I875	D876	I877	E880	I881	K882	M883	R884	M885	T886	K887															
R749	I750	E751	E752	N753	Y754	Y755	K756	R757	A758	Q759	I760	S761	A762	S763	K764	L765	M766	N767	A768	I770	K771	K772	L773	Y774	A775	I776	G778	I779	L782	G783	Y784	N788	E789	E790	E791	L792	F793	D794	F795	S796	D797	PHE	LYS	ASN	ILE	Q803	E804	I805	K806	R807	C808	L809	K810	D811	I812							
M686	F693	D694	T695	I696	E697	T698	V699	K700	X701	I702	L703	M704	A705	L706	I707	N710	K711	E712	L713	Y714	K715	L716	L717	L718	L719	D720	D721	R722	L723	E724	E725	N726	E727	S728	K729	N730	I731	F732	L733	Y734	E735	L736	K737	K738	T739	L740	G741	M742	I743	D744	E745	L746	D747	E748								
Q625	M626	L627	K628	L629	S630	D631	E632	E633	V634	S635	K636	A637	M638	L639	D640	D641	V642	V643	F644	K645	D646	K647	K648	M649	I650	I651	L654	M655	D656	I657	R658	L659	S660	E661	E662	M664	M665	D666	L667	P668	L669	P670	M671	S672	F673	S674	K675	V676	L677	P678	L681	M682	L683	Y684	R685							
Y486	L487	G488	K489	L490	R491	H492	M493	D494	I495	L496	M497	T498	T499	V500	F501	D502	D503	D504	S506	K507	L508	H509	A510	K511	E512	L514	D515	L516	I519	T520	F521	S524	M527	E528	K531	I532	F533	S534	R536	E536	D537	L538	I539	Q540	N543	D549	G550	L551	R552	K553	N554											
Y555	V556	L557	K558	K559	I561	L562	M563	E564	V565	I568	I569	R570	F574	I575	D576	M577	K578	N579	N580	I581	T582	M583	I586	E587	K588	L589	L590	T594	M595	I600	L601	H602	H602	M604	D605	L606	D609	L610	E611	G612	T613	Q614	D615	D616	Y617	M618	K619	V620	I623													
E423	E424	K425	E426	L427	Y428	K429	L430	L431	R432	R433	Y434	L435	K436	G437	R438	F439	E440	K441	L442	L443	Y444	N445	E446	Q447	K448	VAL	ARG	L457	C458	T459	E460	K461	L462	L463	N464	E465	S466	L467	L468	K409	L470	L472	L473	N411	F412	D413	S414	K415	L416	Y417	L480	T481	L481	E482	H483	L484	N485					
I362	K363	N364	N365	S366	L367	K368	E369	L370	E371	K373	L374	E375	A376	E377	F378	K379	L380	D381	E382	L383	L384	K385	E388	K389	E390	L391	N392	K393	K394	N395	C396	T397	T398	E399	L400	F401	G402	L403	F404	K405	K406	H407	K409	L410	N411	F412	D413	S414	K415	L416	Y417	L480	T481	L481	E482	K419	K420	D422				
GLU	LEU	GLU	PHE	TRP	ASN	ILE	THR	LYS	ARG	ILE	GLU	LYS	VAL	LYS	LYS	VAL	ASN	ASN	GLU	PHE	LEU	GLU	LYS	ARG	ARG	ASN	ASN	ARG	THR	TYR	ILE	ILE	LYS	SER	TYR	VAL	LEU	LEU	ASP	LYS	HIS	GLU	LYS	LYS	PHE	LYS	ILE	GLU	ARG	GLU	GLY	ASN	K350	K351	D352	K353	I354	V355	S418	F357	F358	V359



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	91.66Å 94.23Å 338.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.38 – 3.50 48.38 – 3.50	Depositor EDS
% Data completeness (in resolution range)	81.0 (48.38-3.50) 81.0 (48.38-3.50)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.71 (at 3.48Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.263 , 0.291 0.262 , 0.291	Depositor DCC
$R_{free}$ test set	1546 reflections (4.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.6	Xtrriage
Anisotropy	0.648	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 44.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.26$	Xtrriage
Estimated twinning fraction	0.053 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.78	EDS
Total number of atoms	15952	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.41% of the height of the origin peak. No significant pseudotranslation is detected.*

<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

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	0.64	10/8117 (0.1%)	0.98	60/10925 (0.5%)
1	B	0.53	4/8038 (0.0%)	0.94	53/10824 (0.5%)
All	All	0.59	14/16155 (0.1%)	0.96	113/21749 (0.5%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
1	B	0	7
All	All	0	9

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	687	ASN	C-N	7.79	1.40	1.33
1	A	896	TRP	CA-C	-6.25	1.44	1.52
1	B	607	GLU	C-O	-6.03	1.18	1.23
1	A	896	TRP	N-CA	-5.69	1.42	1.46
1	B	1354	LYS	C-N	5.55	1.40	1.33

The worst 5 of 113 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	501	ASN	N-CA-C	13.72	125.75	111.07
1	B	634	VAL	CB-CA-C	-12.86	90.21	111.29
1	A	1358	VAL	N-CA-C	12.13	122.99	110.62
1	B	618	ASN	N-CA-C	-10.81	101.13	114.75
1	B	793	PHE	N-CA-C	10.55	122.86	111.36

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1051	GLU	Peptide
1	A	1356	LYS	Peptide
1	B	412	PHE	Peptide
1	B	605	SER	Peptide
1	B	644	PHE	Peptide

## 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	8013	0	7662	857	3
1	B	7938	0	7580	910	1
2	A	1	0	0	0	0
All	All	15952	0	15242	1767	4

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 57.

The worst 5 of 1767 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:976:LEU:HD23	1:B:977:GLN:N	1.22	1.45
1:A:696:ILE:CB	1:A:699:GLU:HB3	1.55	1.35
1:B:817:THR:CG2	1:B:818:TYR:HA	1.54	1.35
1:B:914:PHE:HA	1:B:915:ASP:CB	1.47	1.34
1:B:914:PHE:CA	1:B:915:ASP:HB2	1.54	1.32

All (4) 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)
1:B:426:GLU:OE1	1:B:1037:SER:OG[4_446]	1.51	0.69
1:A:776:LYS:CB	1:A:1348:ILE:CD1[3_655]	1.91	0.29
1:A:398:THR:CB	1:A:907:MSE:CE[3_655]	2.08	0.12

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:776:LYS:CD	1:A:1348:ILE:CD1[3_655]	2.10	0.10

### 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	1013/1397 (72%)	908 (90%)	79 (8%)	26 (3%)	4	27
1	B	1004/1397 (72%)	888 (88%)	84 (8%)	32 (3%)	3	24
All	All	2017/2794 (72%)	1796 (89%)	163 (8%)	58 (3%)	3	26

5 of 58 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	392	LYS
1	A	410	VAL
1	A	458	GLU
1	A	460	GLU
1	A	461	LYS

#### 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	824/1316 (63%)	689 (84%)	135 (16%)	2	14
1	B	822/1316 (62%)	705 (86%)	117 (14%)	3	18

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1646/2632 (62%)	1394 (85%)	252 (15%)	<b>3</b> <b>16</b>

5 of 252 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1254	LYS
1	B	985	ASN
1	B	410	VAL
1	B	980	GLN
1	B	1122	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 40 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	788	ASN
1	B	1119	ASN
1	B	936	ASN
1	B	1017	ASN
1	B	1200	HIS

### 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](#)

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, 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	1008/1397 (72%)	0.79	116 (11%) <b>9</b> <b>6</b>	2, 42, 106, 153	0
1	B	1002/1397 (71%)	1.33	226 (22%) <b>2</b> <b>2</b>	12, 67, 134, 199	0
All	All	2010/2794 (71%)	1.06	342 (17%) <b>4</b> <b>4</b>	2, 55, 128, 199	0

The worst 5 of 342 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	500	VAL	7.1
1	B	407	HIS	6.7
1	B	396	CYS	6.5
1	B	1345	ASN	6.1
1	B	726	ASN	5.9

### 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.