Full wwPDB X-ray Structure Validation Report (i)

Jan 24, 2026 – 08:04 PM EST

PDB ID	:	9PFD / pdb_00009pfd
Title	:	RMI1-RMI2 bound to synthetic peptide P4Ser
Authors	:	Bythell-Douglas, R.; Lau, Y.H.; Alcock, L.J.; Deshpande, C.
Deposited on	:	2025-07-04
Resolution	:	2.01 Å(reported)

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

|		
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
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)	:	2.47

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 2.01 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	Percentile Ranks	Value
Rfree		0.225
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		0.4%
RSRZ outliers		1.0%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
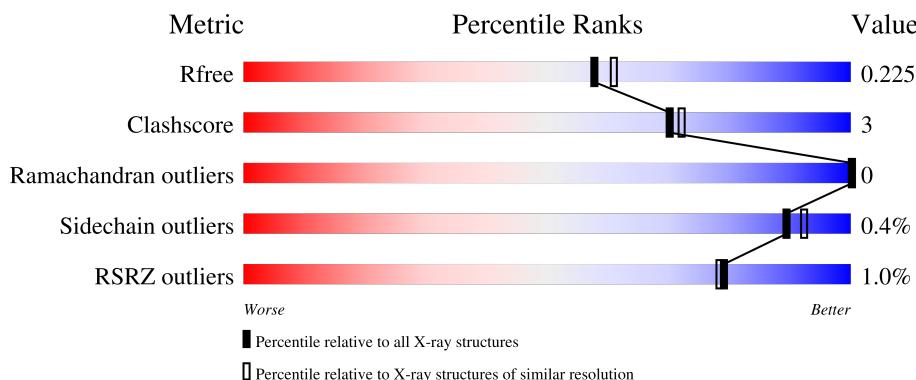
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_free	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (2.00-2.00)

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 <=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	152	% 84% 9% · 6%
2	B	147	% 73% 18% ·· 7%
3	C	15	7% 67% 20% 13%

WORLDWIDE
wwPDB
PROTEIN DATA BANK
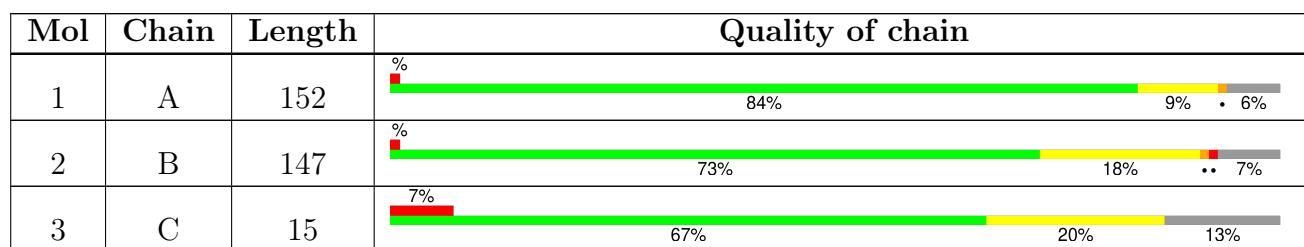
2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 4747 atoms, of which 2312 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 RecQ-mediated genome instability protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	143	Total	C	H	N	O	S	0	1	0
			2273	717	1157	177	213	9			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	474	MET	-	initiating methionine	UNP Q9H9A7

- Molecule 2 is a protein called RecQ-mediated genome instability protein 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	137	Total	C	H	N	O	S	0	1	0
			2105	651	1060	199	186	9			

- Molecule 3 is a protein called P4Ser peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	13	Total	C	H	N	O	0	0	1
			199	74	95	16	14			

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	106	Total	O	0	0
			106	106		
4	B	55	Total	O	0	0
			55	55		
4	C	9	Total	O	0	0
			9	9		

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: RecQ-mediated genome instability protein 1

Chain A:
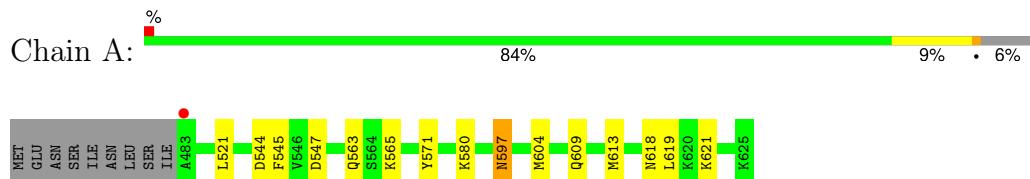
• Molecule 2: RecQ-mediated genome instability protein 2

Chain B:
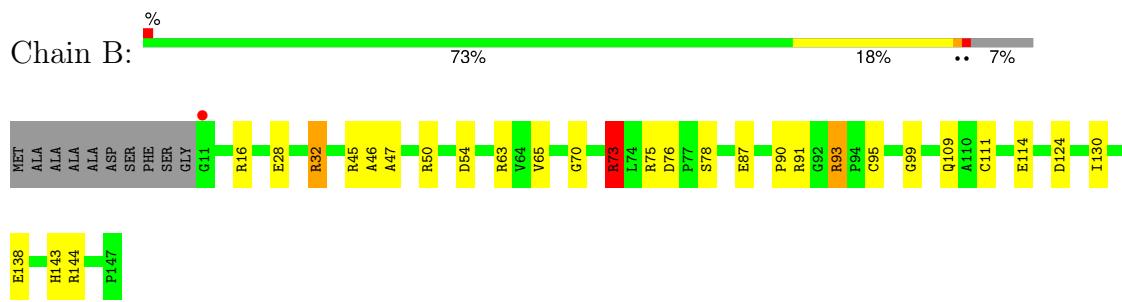
• Molecule 3: P4Ser peptide

Chain C:
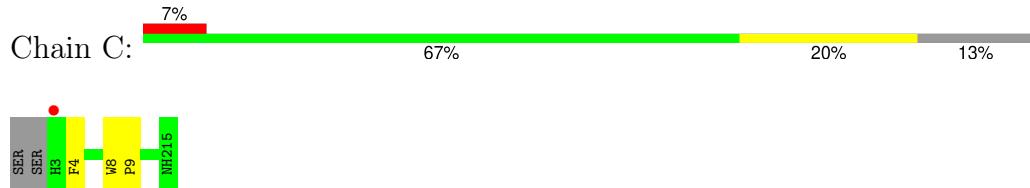
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	117.54Å 117.54Å 53.48Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.72 – 2.01 48.72 – 2.01	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.72-2.01) 99.9 (48.72-2.01)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	7.73 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.181 , 0.224 0.181 , 0.225	Depositor DCC
R_{free} test set	1335 reflections (5.22%)	wwPDB-VP
Wilson B-factor (Å^2)	30.9	Xtriage
Anisotropy	0.001	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 34.5	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4747	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.00% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NH2

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	1.03	1/1134 (0.1%)	1.48	13/1531 (0.8%)
2	B	1.00	1/1068 (0.1%)	1.53	13/1448 (0.9%)
3	C	0.96	0/110	1.46	1/151 (0.7%)
All	All	1.01	2/2312 (0.1%)	1.50	27/3130 (0.9%)

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	1
2	B	0	5
All	All	0	6

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	547	ASP	CG-OD1	5.60	1.35	1.25
2	B	138	GLU	CD-OE1	5.17	1.35	1.25

All (27) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	609	GLN	N-CA-CB	-7.40	97.89	111.13
1	A	597	ASN	CA-CB-CG	-7.07	105.53	112.60
2	B	138	GLU	CG-CD-OE1	7.07	134.65	118.40
2	B	65	VAL	N-CA-C	-7.05	105.75	111.81
1	A	544	ASP	CA-CB-CG	6.56	119.16	112.60
2	B	138	GLU	CG-CD-OE2	-6.27	103.97	118.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	621	LYS	N-CA-CB	6.19	118.98	110.01
2	B	76	ASP	CB-CA-C	6.07	118.35	109.48
2	B	143	HIS	CA-CB-CG	-6.03	107.77	113.80
1	A	580	LYS	CA-CB-CG	-6.01	102.08	114.10
2	B	76	ASP	CA-CB-CG	5.73	118.33	112.60
1	A	580	LYS	CB-CG-CD	5.62	124.22	111.30
2	B	109	GLN	CB-CA-C	-5.59	100.90	110.24
1	A	609	GLN	CB-CA-C	5.46	120.45	109.38
1	A	619	LEU	N-CA-CB	-5.34	102.26	110.16
1	A	545	PHE	N-CA-CB	-5.32	101.38	110.16
2	B	111	CYS	CB-CA-C	-5.30	100.65	109.55
1	A	618	ASN	CA-CB-CG	5.29	117.89	112.60
2	B	45	ARG	NE-CZ-NH2	5.28	123.95	119.20
1	A	547	ASP	CA-CB-CG	5.25	117.85	112.60
2	B	54	ASP	CB-CA-C	5.12	118.41	109.65
3	C	9	PRO	N-CA-CB	-5.12	100.56	103.22
2	B	75	ARG	CD-NE-CZ	5.07	131.49	124.40
2	B	73	ARG	N-CA-CB	-5.06	102.02	110.83
2	B	144	ARG	CA-CB-CG	-5.04	104.01	114.10
1	A	580	LYS	CA-C-N	5.01	126.95	120.44
1	A	580	LYS	C-N-CA	5.01	126.95	120.44

There are no chirality outliers.

All (6) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	521	LEU	Mainchain
2	B	32	ARG	Sidechain
2	B	50	ARG	Sidechain
2	B	73	ARG	Sidechain
2	B	91	ARG	Sidechain
2	B	93	ARG	Sidechain

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	1116	1157	1152	5	0
2	B	1045	1060	1053	10	0
3	C	104	95	89	3	0
4	A	106	0	0	1	0
4	B	55	0	0	1	0
4	C	9	0	0	0	0
All	All	2435	2312	2294	15	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

All (15) 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:B:28:GLU:HB2	2:B:78:SER:HB2	1.77	0.67
1:A:597:ASN:OD1	1:A:597:ASN:C	2.43	0.61
2:B:63:ARG:NH1	2:B:99:GLY:O	2.40	0.55
1:A:563:GLN:NE2	4:A:702:HOH:O	2.39	0.54
2:B:16:ARG:O	3:C:4:PHE:HA	2.09	0.53
2:B:73:ARG:NH2	2:B:114:GLU:OE2	2.45	0.49
2:B:95:CYS:HB3	2:B:124:ASP:HB2	1.93	0.48
2:B:90:PRO:HG3	3:C:8:TRP:CD1	2.49	0.47
2:B:70:GLY:HA3	2:B:87:GLU:OE1	2.17	0.45
1:A:565:LYS:HG3	1:A:571:TYR:CE1	2.53	0.43
2:B:93:ARG:NH2	4:B:204:HOH:O	2.52	0.42
1:A:604:MET:HE2	1:A:604:MET:HB2	1.95	0.42
2:B:46:ALA:O	2:B:47:ALA:C	2.63	0.41
3:C:4:PHE:CD1	3:C:4:PHE:C	2.98	0.41
1:A:613:MET:CG	2:B:130:ILE:HD12	2.51	0.41

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.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	142/152 (93%)	140 (99%)	2 (1%)	0	100	100
2	B	136/147 (92%)	129 (95%)	7 (5%)	0	100	100
3	C	11/15 (73%)	10 (91%)	1 (9%)	0	100	100
All	All	289/314 (92%)	279 (96%)	10 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	
1	A	130/139 (94%)	130 (100%)	0	100	100
2	B	109/115 (95%)	108 (99%)	1 (1%)	75	81
3	C	10/13 (77%)	10 (100%)	0	100	100
All	All	249/267 (93%)	248 (100%)	1 (0%)	89	92

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	32	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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 ⓘ

There are no ligands in this entry.

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, 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	143/152 (94%)	-0.59	1 (0%)	84	83	11, 28, 47, 74	1 (0%)
2	B	137/147 (93%)	0.00	1 (0%)	84	83	14, 39, 64, 78	1 (0%)
3	C	12/15 (80%)	-0.18	1 (8%)	19	17	27, 30, 67, 93	0
All	All	292/314 (92%)	-0.30	3 (1%)	79	78	11, 34, 60, 93	2 (0%)

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	483	ALA	3.4
3	C	3	HIS	2.5
2	B	11	GLY	2.4

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

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

6.3 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.