Full wwPDB X-ray Structure Validation Report (i)

Dec 8, 2025 – 12:40 PM EST

PDB ID	:	9MHK / pdb_00009mhk
Title	:	The structure of Zcp triple mutant H29A, H31A, H142A
Authors	:	Bera, A.K.; Liyayi, I.K.; Criss, A.K.; Noinaj, N.
Deposited on	:	2024-12-12
Resolution	:	2.11 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 2.11 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		0.196
Clashscore		3
Ramachandran outliers		2.7%
Sidechain outliers		1.8%
RSRZ outliers		8.8%

Worse   *Better*

▊ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
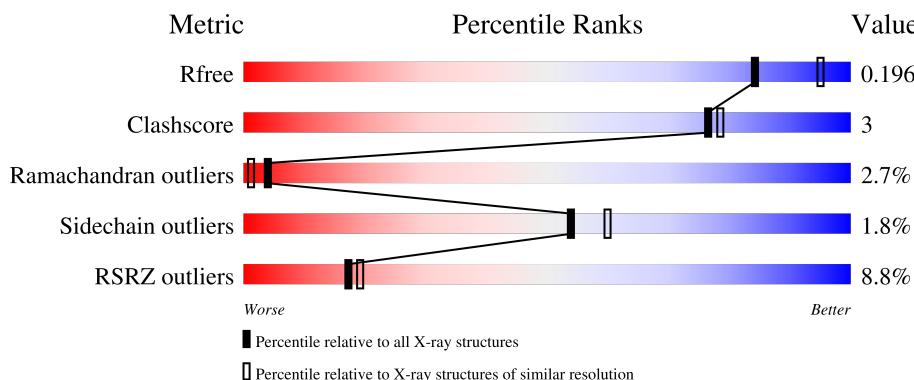
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	7689 (2.14-2.10)
Clashscore	180529	8431 (2.14-2.10)
Ramachandran outliers	177936	8366 (2.14-2.10)
Sidechain outliers	177891	8367 (2.14-2.10)
RSRZ outliers	164620	7689 (2.14-2.10)

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	274	7% / 81% / 8% / 11%
1	B	274	9% / 81% / 8% / 11%
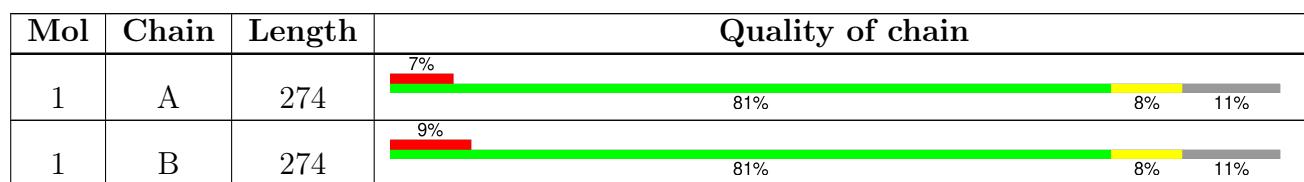
2 Entity composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 4096 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 Phosphoribosylglycinamide formyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	245	Total 1903	C 1198	N 331	O 369	S 5	0	0	0
1	B	245	Total 1882	C 1187	N 325	O 365	S 5	0	0	0

There are 58 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	SER	-	expression tag	UNP Q5F7W6
A	-4	TYR	-	expression tag	UNP Q5F7W6
A	-3	TYR	-	expression tag	UNP Q5F7W6
A	-2	HIS	-	expression tag	UNP Q5F7W6
A	-1	HIS	-	expression tag	UNP Q5F7W6
A	0	HIS	-	expression tag	UNP Q5F7W6
A	1	HIS	-	expression tag	UNP Q5F7W6
A	2	HIS	-	expression tag	UNP Q5F7W6
A	3	HIS	-	expression tag	UNP Q5F7W6
A	4	ASP	-	expression tag	UNP Q5F7W6
A	5	TYR	-	expression tag	UNP Q5F7W6
A	6	ASP	-	expression tag	UNP Q5F7W6
A	7	ILE	-	expression tag	UNP Q5F7W6
A	8	PRO	-	expression tag	UNP Q5F7W6
A	9	THR	-	expression tag	UNP Q5F7W6
A	10	THR	-	expression tag	UNP Q5F7W6
A	11	GLU	-	expression tag	UNP Q5F7W6
A	12	ASN	-	expression tag	UNP Q5F7W6
A	13	LEU	-	expression tag	UNP Q5F7W6
A	14	TYR	-	expression tag	UNP Q5F7W6
A	15	PHE	-	expression tag	UNP Q5F7W6
A	16	GLN	-	expression tag	UNP Q5F7W6
A	17	GLY	-	expression tag	UNP Q5F7W6
A	18	ALA	-	expression tag	UNP Q5F7W6
A	19	MET	-	expression tag	UNP Q5F7W6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	VAL	-	expression tag	UNP Q5F7W6
A	29	ALA	HIS	engineered mutation	UNP Q5F7W6
A	31	ALA	HIS	engineered mutation	UNP Q5F7W6
A	142	ALA	HIS	engineered mutation	UNP Q5F7W6
B	-5	SER	-	expression tag	UNP Q5F7W6
B	-4	TYR	-	expression tag	UNP Q5F7W6
B	-3	TYR	-	expression tag	UNP Q5F7W6
B	-2	HIS	-	expression tag	UNP Q5F7W6
B	-1	HIS	-	expression tag	UNP Q5F7W6
B	0	HIS	-	expression tag	UNP Q5F7W6
B	1	HIS	-	expression tag	UNP Q5F7W6
B	2	HIS	-	expression tag	UNP Q5F7W6
B	3	HIS	-	expression tag	UNP Q5F7W6
B	4	ASP	-	expression tag	UNP Q5F7W6
B	5	TYR	-	expression tag	UNP Q5F7W6
B	6	ASP	-	expression tag	UNP Q5F7W6
B	7	ILE	-	expression tag	UNP Q5F7W6
B	8	PRO	-	expression tag	UNP Q5F7W6
B	9	THR	-	expression tag	UNP Q5F7W6
B	10	THR	-	expression tag	UNP Q5F7W6
B	11	GLU	-	expression tag	UNP Q5F7W6
B	12	ASN	-	expression tag	UNP Q5F7W6
B	13	LEU	-	expression tag	UNP Q5F7W6
B	14	TYR	-	expression tag	UNP Q5F7W6
B	15	PHE	-	expression tag	UNP Q5F7W6
B	16	GLN	-	expression tag	UNP Q5F7W6
B	17	GLY	-	expression tag	UNP Q5F7W6
B	18	ALA	-	expression tag	UNP Q5F7W6
B	19	MET	-	expression tag	UNP Q5F7W6
B	20	VAL	-	expression tag	UNP Q5F7W6
B	29	ALA	HIS	engineered mutation	UNP Q5F7W6
B	31	ALA	HIS	engineered mutation	UNP Q5F7W6
B	142	ALA	HIS	engineered mutation	UNP Q5F7W6

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	186	Total 186	O 186	0	0
2	B	125	Total 125	O 125	0	0

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: Phosphoribosylglycinamide formyltransferase

- Molecule 1: Phosphoribosylglycinamide formyltransferase
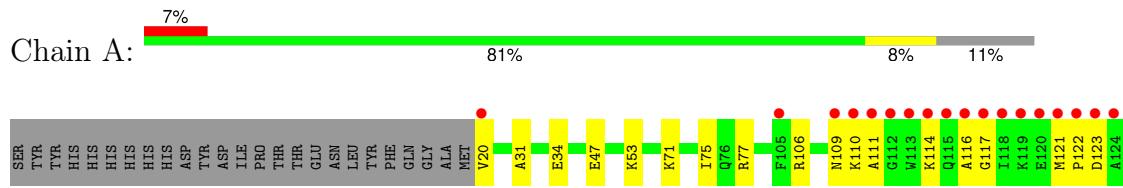
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	55.48Å 91.42Å 126.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.43 – 2.11 47.43 – 2.11	Depositor EDS
% Data completeness (in resolution range)	98.3 (47.43-2.11) 98.2 (47.43-2.11)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.03 (at 2.10Å)	Xtriage
Refinement program	PHENIX (1.18rc1_3777: ???)	Depositor
R, R_{free}	0.168 , 0.194 0.169 , 0.196	Depositor DCC
R_{free} test set	2000 reflections (5.29%)	wwPDB-VP
Wilson B-factor (Å^2)	32.6	Xtriage
Anisotropy	0.185	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 50.0	EDS
L-test for twinning[2]	$< |L| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4096	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.65% 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 ⓘ

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	0.24	0/1944	0.52	0/2634
1	B	0.21	0/1922	0.51	0/2607
All	All	0.22	0/3866	0.52	0/5241

There are no bond length outliers.

There are no bond angle outliers.

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	1903	0	1830	11	0
1	B	1882	0	1793	13	0
2	A	186	0	0	4	0
2	B	125	0	0	2	0
All	All	4096	0	3623	21	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 (21) 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:230:ARG:NH1	1:B:230:ARG:HH12	1.75	0.85

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:230:ARG:HH12	1:B:230:ARG:HH12	1.21	0.84
1:B:205:LYS:O	1:B:207:HIS:N	2.20	0.75
1:A:230:ARG:HH12	1:B:230:ARG:NH1	1.85	0.74
1:A:47:GLU:OE1	2:A:301:HOH:O	2.12	0.68
1:A:243:ASP:OD1	1:A:252:LYS:NZ	2.26	0.62
1:A:174:ARG:NH2	2:A:304:HOH:O	2.36	0.57
1:B:205:LYS:C	1:B:207:HIS:H	2.11	0.57
1:B:77:ARG:CZ	1:B:85:ARG:HH21	2.19	0.56
1:A:31:ALA:O	1:A:34:GLU:HG2	2.10	0.51
1:B:75:ILE:HG22	1:B:77:ARG:HG2	1.93	0.51
1:A:75:ILE:HG22	1:A:77:ARG:HG2	1.93	0.51
1:B:121:MET:O	1:B:123:ASP:N	2.45	0.50
1:A:71:LYS:HE2	2:A:401:HOH:O	2.12	0.50
1:B:101:TYR:HB3	1:B:132:MET:HB2	1.95	0.48
1:B:31:ALA:O	1:B:34:GLU:HG2	2.14	0.47
1:A:106:ARG:HG2	2:A:316:HOH:O	2.16	0.45
1:A:53:LYS:HA	1:A:53:LYS:HD3	1.73	0.44
1:B:247:GLN:O	2:B:302:HOH:O	2.21	0.41
1:B:22:ARG:NE	2:B:301:HOH:O	2.04	0.40
1:B:121:MET:O	1:B:122:PRO:C	2.64	0.40

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	
1	A	243/274 (89%)	226 (93%)	10 (4%)	7 (3%)	3	1
1	B	243/274 (89%)	230 (95%)	7 (3%)	6 (2%)	4	1
All	All	486/548 (89%)	456 (94%)	17 (4%)	13 (3%)	4	1

All (13) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	122	PRO
1	B	122	PRO
1	B	206	THR
1	B	112	GLY
1	B	203	ARG
1	A	111	ALA
1	A	114	LYS
1	A	116	ALA
1	A	117	GLY
1	A	121	MET
1	A	123	ASP
1	B	114	LYS
1	B	118	ILE

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	198/233 (85%)	194 (98%)	4 (2%)	50	56
1	B	192/233 (82%)	189 (98%)	3 (2%)	58	65
All	All	390/466 (84%)	383 (98%)	7 (2%)	54	60

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	20	VAL
1	A	109	ASN
1	A	110	LYS
1	A	143	GLU
1	B	20	VAL
1	B	143	GLU
1	B	208	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	102	GLN
1	A	157	ASN
1	A	212	GLN
1	A	231	GLN
1	B	21	HIS
1	B	102	GLN
1	B	136	ASN
1	B	212	GLN
1	B	231	GLN

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	245/274 (89%)	-0.14	19 (7%)	20	23	20, 34, 122, 150	0
1	B	245/274 (89%)	0.15	24 (9%)	14	16	22, 40, 111, 134	0
All	All	490/548 (89%)	0.00	43 (8%)	17	19	20, 38, 119, 150	0

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	116	ALA	8.1
1	A	116	ALA	6.9
1	B	113	TRP	6.4
1	A	118	ILE	5.8
1	A	117	GLY	5.6
1	B	114	LYS	4.9
1	A	111	ALA	4.8
1	A	113	TRP	4.8
1	B	206	THR	4.6
1	B	203	ARG	4.4
1	B	111	ALA	4.4
1	B	118	ILE	4.1
1	B	115	GLN	4.1
1	B	20	VAL	3.9
1	B	204	SER	3.9
1	B	122	PRO	3.8
1	B	117	GLY	3.5
1	B	120	GLU	3.5
1	B	207	HIS	3.4
1	A	115	GLN	3.4
1	A	120	GLU	3.3
1	A	20	VAL	3.3
1	B	121	MET	3.3
1	A	121	MET	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	119	LYS	3.1
1	A	122	PRO	3.0
1	B	205	LYS	3.0
1	B	208	LYS	2.8
1	B	124	ALA	2.6
1	A	109	ASN	2.5
1	A	124	ALA	2.5
1	B	112	GLY	2.4
1	B	105	PHE	2.4
1	B	123	ASP	2.4
1	A	114	LYS	2.2
1	A	110	LYS	2.2
1	B	109	ASN	2.2
1	A	105	PHE	2.2
1	A	112	GLY	2.2
1	A	123	ASP	2.1
1	A	125	SER	2.1
1	A	119	LYS	2.1
1	B	110	LYS	2.1

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.