



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 04:50 PM UTC

PDB ID : 3VJJ / pdb_00003vjj
Title : Crystal Structure Analysis of the P9-1
Authors : Akita, F.; Higashiura, A.; Suzuki, M.; Tsukihara, T.; Nakagawa, A.; Omura, T.
Deposited on : 2011-10-24
Resolution : 3.00 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

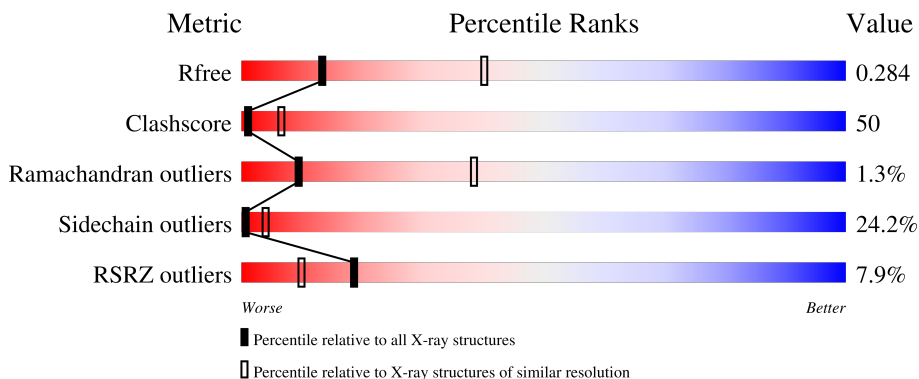
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	2672 (3.00-3.00)
Clashscore	190562	2977 (3.00-3.00)
Ramachandran outliers	187476	2877 (3.00-3.00)
Sidechain outliers	187428	2880 (3.00-3.00)
RSRZ outliers	180081	2671 (3.00-3.00)

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	368	5% 24% 38% 13% 25%
1	B	368	7% 23% 37% 13% 27%

2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 4475 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 P9-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	277	Total	C	N	O	S	0	0	0
			2265	1459	372	427	7			
1	B	270	Total	C	N	O	S	0	0	0
			2210	1428	364	411	7			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP Q913E4
A	2	SER	-	expression tag	UNP Q913E4
A	3	TYR	-	expression tag	UNP Q913E4
A	4	TYR	-	expression tag	UNP Q913E4
A	5	HIS	-	expression tag	UNP Q913E4
A	6	HIS	-	expression tag	UNP Q913E4
A	7	HIS	-	expression tag	UNP Q913E4
A	8	HIS	-	expression tag	UNP Q913E4
A	9	HIS	-	expression tag	UNP Q913E4
A	10	HIS	-	expression tag	UNP Q913E4
A	11	LEU	-	expression tag	UNP Q913E4
A	12	GLY	-	expression tag	UNP Q913E4
A	13	SER	-	expression tag	UNP Q913E4
A	14	THR	-	expression tag	UNP Q913E4
A	15	SER	-	expression tag	UNP Q913E4
A	16	LEU	-	expression tag	UNP Q913E4
A	17	TYR	-	expression tag	UNP Q913E4
A	18	LYS	-	expression tag	UNP Q913E4
A	19	LYS	-	expression tag	UNP Q913E4
A	20	ALA	-	expression tag	UNP Q913E4
A	21	GLY	-	expression tag	UNP Q913E4
A	346	LYS	ARG	conflict	UNP Q913E4
B	1	MET	-	expression tag	UNP Q913E4
B	2	SER	-	expression tag	UNP Q913E4
B	3	TYR	-	expression tag	UNP Q913E4

Continued on next page...

Continued from previous page...

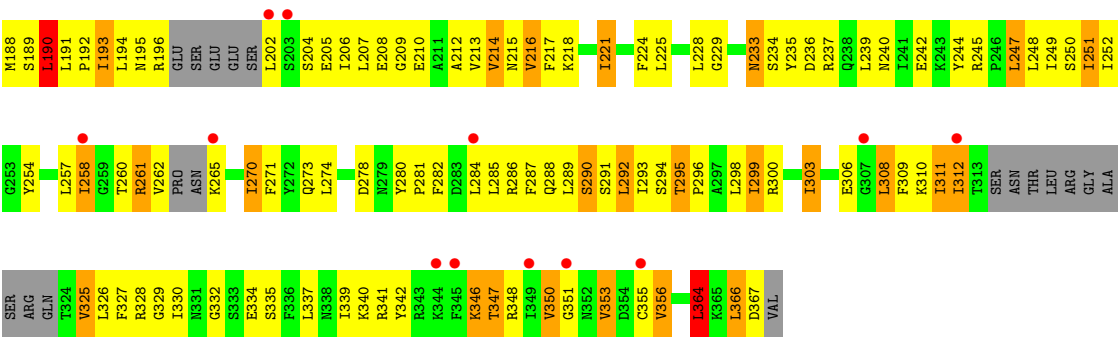
Chain	Residue	Modelled	Actual	Comment	Reference
B	4	TYR	-	expression tag	UNP Q913E4
B	5	HIS	-	expression tag	UNP Q913E4
B	6	HIS	-	expression tag	UNP Q913E4
B	7	HIS	-	expression tag	UNP Q913E4
B	8	HIS	-	expression tag	UNP Q913E4
B	9	HIS	-	expression tag	UNP Q913E4
B	10	HIS	-	expression tag	UNP Q913E4
B	11	LEU	-	expression tag	UNP Q913E4
B	12	GLY	-	expression tag	UNP Q913E4
B	13	SER	-	expression tag	UNP Q913E4
B	14	THR	-	expression tag	UNP Q913E4
B	15	SER	-	expression tag	UNP Q913E4
B	16	LEU	-	expression tag	UNP Q913E4
B	17	TYR	-	expression tag	UNP Q913E4
B	18	LYS	-	expression tag	UNP Q913E4
B	19	LYS	-	expression tag	UNP Q913E4
B	20	ALA	-	expression tag	UNP Q913E4
B	21	GLY	-	expression tag	UNP Q913E4
B	346	LYS	ARG	conflict	UNP Q913E4

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.

Chain A:

Chain B:

Amino Acid	Frequency (%)
Met	7%
Ser	23%
Tyr	23%
K64	23%
His	23%
I66	23%
P67	23%
I68	23%
Ser	23%
L69	23%
D70	23%
Lys	23%
Gly	23%
Ser	23%
Thr	23%
Ser	23%
Leu	23%
L76	23%
L77	23%
N78	23%
Lys	23%
Lys	23%
Ala	23%
Gly	23%
Met	23%
Ala	23%
Asp	23%
Gln	23%
E26	37%
R27	37%
R28	37%
T29	37%
F30	37%
G31	37%
S32	37%
Y33	37%
K34	37%
I35	37%
E36	37%
E37	37%
L38	37%
T39	37%
L40	37%
K41	37%
M42	37%
Asp	13%
Gln	13%
Pro	13%
Asn	13%
Arg	13%
Asn	13%
Thr	13%
Asn	13%
Ser	13%
Ser	13%
Ser	13%
Thr	13%
Thr	13%
Ser	13%
Thr	13%
Ser	13%
Gln	13%
Gln	13%
Pro	13%
Asn	13%
Arg	13%
Asn	13%
Arg	13%



4 Data and refinement statistics

Property	Value	Source
Space group	P 4 21 2	Depositor
Cell constants a, b, c, α , β , γ	127.28Å 127.28Å 143.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.18 – 3.00 42.18 – 3.00	Depositor EDS
% Data completeness (in resolution range)	94.6 (42.18-3.00) 93.6 (42.18-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.33 (at 3.01Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7_650)	Depositor
R, R_{free}	0.243 , 0.293 0.238 , 0.284	Depositor DCC
R_{free} test set	974 reflections (4.25%)	wwPDB-VP
Wilson B-factor (Å ²)	94.8	Xtriage
Anisotropy	0.580	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 99.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	4475	wwPDB-VP
Average B, all atoms (Å ²)	116.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.08% 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

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	0.86	2/2306 (0.1%)	1.18	13/3106 (0.4%)
1	B	0.66	0/2247	1.15	10/3021 (0.3%)
All	All	0.77	2/4553 (0.0%)	1.16	23/6127 (0.4%)

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	1
All	All	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	41	LYS	C-N	24.72	1.75	1.33
1	A	40	ILE	C-N	-5.85	1.23	1.33

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	41	LYS	N-CA-C	-11.55	99.72	114.04
1	A	41	LYS	CA-C-N	8.42	138.89	122.73
1	A	41	LYS	C-N-CA	8.42	138.89	122.73
1	B	261	ARG	NE-CZ-NH2	8.10	126.49	119.20
1	B	181	GLU	N-CA-C	7.67	121.24	112.57

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	182	GLU	Peptide
1	A	201	SER	Peptide
1	B	182	GLU	Peptide

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	2265	0	2296	236	0
1	B	2210	0	2251	222	2
All	All	4475	0	4547	450	2

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 50.

The worst 5 of 450 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:41:LYS:C	1:A:63:THR:N	1.75	1.44
1:A:201:SER:H	1:A:202:LEU:HD22	1.16	1.05
1:A:97:SER:HB3	1:A:99:LEU:H	1.22	1.03
1:B:97:SER:HB3	1:B:99:LEU:H	1.19	1.02
1:A:353:VAL:O	1:A:356:VAL:HG12	1.62	0.98

All (2) 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:204:SER:OG	1:B:215:ASN:OD1[8_555]	2.01	0.19
1:B:105:ASP:OD1	1:B:208:GLU:OE2[8_555]	2.16	0.04

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	271/368 (74%)	243 (90%)	23 (8%)	5 (2%)	6	31
1	B	258/368 (70%)	238 (92%)	18 (7%)	2 (1%)	16	50
All	All	529/736 (72%)	481 (91%)	41 (8%)	7 (1%)	9	38

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	262	VAL
1	A	349	ILE
1	B	96	TYR
1	A	69	LEU
1	A	96	TYR

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	258/342 (75%)	196 (76%)	62 (24%)	1	4
1	B	251/342 (73%)	190 (76%)	61 (24%)	1	4
All	All	509/684 (74%)	386 (76%)	123 (24%)	1	4

5 of 123 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	325	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	299	ILE
1	B	91	TYR
1	B	298	LEU
1	B	350	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	233	ASN
1	B	101	ASN
1	B	233	ASN
1	B	168	ASN
1	A	168	ASN

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, 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	277/368 (75%)	0.43	17 (6%) 27 14	65, 113, 177, 227	0
1	B	270/368 (73%)	0.50	26 (9%) 13 7	61, 109, 167, 223	0
All	All	547/736 (74%)	0.46	43 (7%) 18 10	61, 111, 174, 227	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	202	LEU	5.1
1	A	360	ASP	4.8
1	B	203	SER	4.5
1	B	126	LEU	3.9
1	A	69	LEU	3.8

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.