



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 13, 2017 – 11:44 pm GMT

PDB ID : 3VJJ
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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

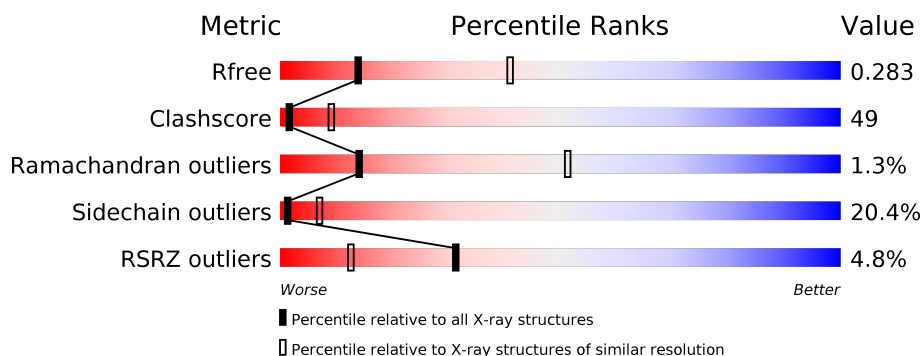
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

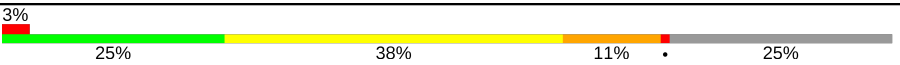

The reported resolution of this entry is 3.00 Å.

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}	100719	1692 (3.00-3.00)
Clashscore	112137	2037 (3.00-3.00)
Ramachandran outliers	110173	1973 (3.00-3.00)
Sidechain outliers	110143	1976 (3.00-3.00)
RSRZ outliers	101464	1716 (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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	
1	B	368	

2 Entry composition

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 and DNA 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 ($\text{RSRZ} > 2$). 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:

3% 25% 38% 11% 25%

MET LEU TYR ASP ARG GLN ALA ASN ILE SER VAL LYS LEU D70 HIS LEU D71 GLY SER THR SER SER E75 LEU L77 L76 W78 Y79 L80 ALA GLY D82 MET MET ALA ASP Q25 E26 R27 R28 T29 F30 G31 S32 Y33 K34 I35 E36 E37 L38 T39 I40 K41 ASN ASP GLN PRO PRO ASN ARG ASN THR SER SER THR GLN THR THR GLU THR ASN ASP

S189 L190 L191 P192 I193 I194 N195 R196 E197 S198 S199 E200 S201 L202 E206 L207 E208 T209 F210 F211 F212 Q213 Q214 V215 V216 F217 K218 M219 Y220 Y221 F224 L225 L228 G229 N232 N233 S234 D236 R237 Q238 L239 E242 R245 L246 L247 L248 L249 S250 S251 V252 V253 V254 V255

L257 L258 R261 V262 P263 R264 K265 L270 F271 F272 Q273 Q274 L274 D278 R279 V280 P281 F282 D283 L284 L285 L286 F287 Q288 L289 S290 S291 L292 L293 S294 T295 L298 L299 I303 E306 G307 L308 F309 L310 L311 L312 T313 S313 SER ASN THR THR ARG LEU ARG GLY ALA SER SER ARG GLN T324 V254

Chain B:

Amino Acid	Percentage
MET	4%
SER	25%
TYR	39%
GLY	10%
ASP	27%



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 0.237 , 0.283	Depositor DCC
R_{free} test set	973 reflections (4.25%)	DCC
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 , 89.3	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	0.70	1/2306 (0.0%)	0.87	7/3106 (0.2%)
1	B	0.56	0/2247	0.84	3/3021 (0.1%)
All	All	0.63	1/4553 (0.0%)	0.86	10/6127 (0.2%)

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 (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	41	LYS	C-N	18.06	1.75	1.34

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	261	ARG	NE-CZ-NH1	13.60	127.10	120.30
1	B	261	ARG	NE-CZ-NH2	12.38	126.49	120.30
1	A	261	ARG	NE-CZ-NH2	-12.34	114.13	120.30
1	B	261	ARG	NE-CZ-NH1	-11.62	114.49	120.30
1	A	41	LYS	C-N-CA	6.88	138.89	121.70

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	231	0
1	B	2210	0	2251	215	2
All	All	4475	0	4547	438	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 49.

The worst 5 of 438 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	1.38
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%)	10	43
1	B	258/368 (70%)	238 (92%)	18 (7%)	2 (1%)	22	64
All	All	529/736 (72%)	481 (91%)	41 (8%)	7 (1%)	14	51

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%)	204 (79%)	54 (21%)	1	6
1	B	251/342 (73%)	201 (80%)	50 (20%)	1	8
All	All	509/684 (74%)	405 (80%)	104 (20%)	1	7

5 of 104 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	311	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	75	GLU
1	B	311	ILE
1	A	325	VAL
1	A	347	THR

Some 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	233	ASN
1	B	101	ASN
1	A	101	ASN
1	B	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 carbohydrates 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%)	0.12	11 (3%) 39 16	65, 113, 177, 227	0
1	B	270/368 (73%)	0.19	15 (5%) 25 10	61, 109, 167, 223	0
All	All	547/736 (74%)	0.15	26 (4%) 31 12	61, 111, 174, 227	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	349	ILE	5.1
1	B	345	PHE	4.9
1	B	265	LYS	4.5
1	B	202	LEU	4.5
1	B	63	THR	4.4

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 carbohydrates 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.