

Full wwPDB X-ray Structure Validation Report (i)

Nov 3, 2024 – 06:19 AM EST

PDB ID	:	3W8J
Title	:	Crystal structure of P5 a0 in a complex with Prx4 c-term
Authors	:	Inaba, K.; Suzuki, M.; Kojima, R.
Deposited on		
Resolution	:	2.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (i) 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 (1)) were used in the production of this report:

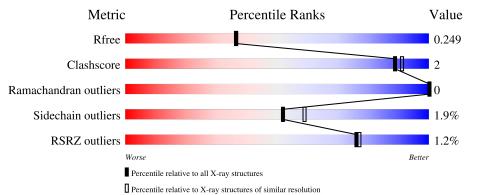
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY \, DIFFRACTION$

The reported resolution of this entry is 2.10 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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R_{free}	164625	6234 (2.10-2.10)
Clashscore	180529	6893 (2.10-2.10)
Ramachandran outliers	177936	6839 (2.10-2.10)
Sidechain outliers	177891	6840 (2.10-2.10)
RSRZ outliers	164620	6234 (2.10-2.10)

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	А	142	2%		82%		•	15%
1	В	142			81%			14%
2	С	20	20%	10%		70%		
2	D	20	25%	5%		70%		



2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 2053 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Δ	121	Total	С	Ν	0	S	0	0	0
	A	121	941	595	166	179	1	0		
1	В	122	Total	С	Ν	0	S	0	0	0
	I B	122	945	597	167	180	1	0	0	0

• Molecule 1 is a protein called Protein disulfide-isomerase A6.

Chain	Residue	Modelled	Actual	Comment	Reference
А	-1	MET	-	expression tag	UNP Q15084
А	0	GLY	-	expression tag	UNP Q15084
А	1	SER	-	expression tag	UNP Q15084
А	2	SER	-	expression tag	UNP Q15084
А	3	HIS	-	expression tag	UNP Q15084
А	4	HIS	-	expression tag	UNP Q15084
А	5	HIS	-	expression tag	UNP Q15084
А	6	HIS	-	expression tag	UNP Q15084
А	7	HIS	-	expression tag	UNP Q15084
А	8	HIS	-	expression tag	UNP Q15084
А	9	SER	-	expression tag	UNP Q15084
А	10	SER	-	expression tag	UNP Q15084
А	11	GLY	-	expression tag	UNP Q15084
А	12	LEU	-	expression tag	UNP Q15084
А	13	VAL	-	expression tag	UNP Q15084
А	14	PRO	-	expression tag	UNP Q15084
А	15	ARG	-	expression tag	UNP Q15084
А	16	GLY	-	expression tag	UNP Q15084
А	17	SER	-	expression tag	UNP Q15084
А	18	HIS	-	expression tag	UNP Q15084
А	19	MET	-	expression tag	UNP Q15084
А	58	ALA	CYS	engineered mutation	UNP Q15084
В	-1	MET	-	expression tag	UNP Q15084
В	0	GLY	-	expression tag	UNP Q15084
В	1	SER	-	expression tag	UNP Q15084

There are 44 discrepancies between the modelled and reference sequences:

Continued on next page...



Chain	Residue	Modelled	Actual	Comment	Reference
В	2	SER	-	expression tag	UNP Q15084
В	3	HIS	-	expression tag	UNP Q15084
В	4	HIS	-	expression tag	UNP Q15084
В	5	HIS	-	expression tag	UNP Q15084
В	6	HIS	-	expression tag	UNP Q15084
В	7	HIS	-	expression tag	UNP Q15084
В	8	HIS	-	expression tag	UNP Q15084
В	9	SER	-	expression tag	UNP Q15084
В	10	SER	-	expression tag	UNP Q15084
В	11	GLY	-	expression tag	UNP Q15084
В	12	LEU	-	expression tag	UNP Q15084
В	13	VAL	-	expression tag	UNP Q15084
В	14	PRO	-	expression tag	UNP Q15084
В	15	ARG	-	expression tag	UNP Q15084
В	16	GLY	-	expression tag	UNP Q15084
В	17	SER	-	expression tag	UNP Q15084
В	18	HIS	-	expression tag	UNP Q15084
В	19	MET	-	expression tag	UNP Q15084
В	58	ALA	CYS	engineered mutation	UNP Q15084

Continued from previous page...

• Molecule 2 is a protein called C-terminal peptide from Peroxiredoxin-4.

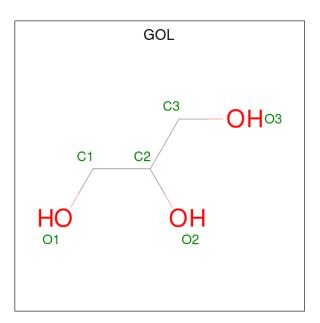
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
2	С	6	Total				S	0	0	0
		Ű	38	23	6	8	1	Ŭ		
2	Л	6	Total	С	Ν	Ο	\mathbf{S}	0	0	0
2	2 D	0	38	23	6	8	1	0	0	0

• Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	2	Total K 2 2	0	0
3	В	2	Total K 2 2	0	0

• Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	В	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 6 & 3 & 3 \end{array}$	0	0
4	В	1	$\begin{array}{ccc} \text{Total} \text{C} \text{O} \\ 6 3 3 \end{array}$	0	0
4	В	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 6 & 3 & 3 \end{array}$	0	0
4	В	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 6 & 3 & 3 \end{array}$	0	0

• Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	А	18	Total O 18 18	0	0
5	В	41	$\begin{array}{cc} \text{Total} & \text{O} \\ 42 & 42 \end{array}$	0	1
5	С	1	Total O 1 1	0	0
5	D	2	Total O 2 2	0	0



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.

- Chain A: 82% 15% MET GLY SER HIS HHIS HHIS HHIS HHIS HHIS SER HHIS SER HHIS SER HHIS SER HHIS SER HHIS SER HHIS MET HHIS MHI • Molecule 1: Protein disulfide-isomerase A6 Chain B: 81% 14% MET GLY SERR SERR HIS SER HIS HIS HIS HIS SER HIS SER CLY VAL U LEU VAL LEU VAL I LEU RAG CA RAGONINA • Molecule 2: C-terminal peptide from Peroxiredoxin-4 Chain C: 20% 10% 70% TRP LYS PRO GLY SER GLU THR THR ILE ILE ILE PRO PRO • Molecule 2: C-terminal peptide from Peroxiredoxin-4 Chain D: 25% 5% 70% GLY SER GLU GLU THR ILE ILE PRO ASP
- Molecule 1: Protein disulfide-isomerase A6



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants	39.05Å 53.38Å 133.35Å	Depositor
a, b, c, α , β , γ	90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.67 - 2.10	Depositor
Resolution (A)	41.67 - 2.10	EDS
% Data completeness	82.7 (41.67-2.10)	Depositor
(in resolution range)	82.7 (41.67-2.10)	EDS
R _{merge}	(Not available)	Depositor
R _{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	6.50 (at 2.10Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
D D	0.182 , 0.245	Depositor
R, R_{free}	0.189 , 0.249	DCC
R_{free} test set	709 reflections (5.05%)	wwPDB-VP
Wilson B-factor $(Å^2)$	29.5	Xtriage
Anisotropy	0.766	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.35, 50.3	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	0.95	EDS
Total number of atoms	2053	wwPDB-VP
Average B, all atoms $(Å^2)$	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 7.14% of the height of the origin peak. No significant pseudotranslation is detected.

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.



¹Intensities estimated from amplitudes.

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: GOL, K

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	Mol Chain		lengths	Bond angles	
	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5
1	А	0.34	0/962	0.49	0/1304
1	В	0.38	0/966	0.51	0/1309
2	С	0.54	0/38	0.68	0/51
2	D	0.38	0/38	0.71	0/51
All	All	0.36	0/2004	0.51	0/2715

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 (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	А	941	0	919	3	0
1	В	945	0	921	3	0
2	С	38	0	36	1	0
2	D	38	0	36	1	0
3	А	2	0	0	0	0
3	В	2	0	0	0	0
4	В	24	0	32	0	0
5	А	18	0	0	1	0
5	В	42	0	0	0	0

Continued on next page...



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 2.

All (7) 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:60:ARG:NH1	5:A:315:HOH:O	2.27	0.66
1:B:102:LYS:NZ	1:B:113:ASP:OD2	2.25	0.56
2:C:0:GLU:HG3	2:C:1:VAL:HG23	1.93	0.51
1:A:69:ALA:HA	1:A:78:VAL:HG21	1.97	0.45
1:A:46:TRP:HB2	1:A:104:PHE:HB2	1.99	0.44
1:B:118:ARG:NH1	2:D:0:GLU:HB2	2.32	0.44
1:B:85:LYS:HE3	1:B:86:HIS:NE2	2.34	0.42

There are no symmetry-related clashes.

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	А	119/142~(84%)	114 (96%)	5(4%)	0	100 100
1	В	120/142~(84%)	116~(97%)	4(3%)	0	100 100
2	С	4/20~(20%)	4 (100%)	0	0	100 100
2	D	4/20~(20%)	4 (100%)	0	0	100 100
All	All	247/324~(76%)	238~(96%)	9~(4%)	0	100 100

There are no Ramachandran outliers to report.



Mol Chain Non-H H(model) H(added) Clashes Symm-Clashes 5 $\overline{\mathbf{C}}$ 0 0 0 0 1 5 D 2 0 0 0 0 All All 20530 1944 7 0

Continued from previous page...

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 side chain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	А	99/117~(85%)	99~(100%)	0	100 100
1	В	99/117~(85%)	96~(97%)	3(3%)	36 40
2	С	4/16~(25%)	4 (100%)	0	100 100
2	D	4/16~(25%)	3~(75%)	1 (25%)	0 0
All	All	206/266~(77%)	202~(98%)	4 (2%)	52 59

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	В	73	LYS
1	В	113	ASP
1	В	134	LEU
2	D	0	GLU

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

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)

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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| > 2 is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Turne	Chain	Dec	Link	B	ond leng	gths	B	ond ang	gles
	Type	Chain	nes		Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z >2
4	GOL	В	203	-	5,5,5	0.30	0	$5,\!5,\!5$	0.30	0
4	GOL	В	204	-	$5,\!5,\!5$	0.38	0	$5,\!5,\!5$	0.26	0
4	GOL	В	202	-	$5,\!5,\!5$	0.37	0	$5,\!5,\!5$	0.36	0
4	GOL	В	201	-	$5,\!5,\!5$	0.35	0	$5,\!5,\!5$	0.40	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	В	203	-	-	2/4/4/4	-
4	GOL	В	204	-	-	4/4/4/4	-
4	GOL	В	202	-	-	3/4/4/4	-
4	GOL	В	201	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	В	204	GOL	O1-C1-C2-C3
4	В	201	GOL	O1-C1-C2-C3
4	В	202	GOL	O1-C1-C2-C3
4	В	204	GOL	C1-C2-C3-O3
4	В	202	GOL	O1-C1-C2-O2
4	В	201	GOL	O1-C1-C2-O2
4	В	204	GOL	O1-C1-C2-O2

Continued on next page...



Mol	Chain		Type	Atoms
4	В	203	GOL	O1-C1-C2-O2
4	В	204	GOL	O2-C2-C3-O3
4	В	202	GOL	C1-C2-C3-O3
4	В	203	GOL	O1-C1-C2-C3

Continued from previous page...

There are no ring outliers.

No monomer is involved in short contacts.

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^{th} 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	$OWAB(Å^2)$	$\mathbf{Q} {<} 0.9$
1	А	121/142~(85%)	0.18	3 (2%) 58 60	26, 44, 75, 88	0
1	В	122/142~(85%)	-0.34	0 100 100	21, 33, 53, 75	0
2	С	6/20~(30%)	0.71	0 100 100	40, 44, 53, 57	0
2	D	6/20~(30%)	0.29	0 100 100	36, 40, 50, 53	0
All	All	255/324 (78%)	-0.06	3 (1%) 76 77	21, 38, 70, 88	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	140	GLY	2.8
1	А	36	ASN	2.4
1	А	32	PRO	2.3

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 monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.



Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q < 0.9
4	GOL	В	204	6/6	0.68	0.15	72,73,76,77	0
4	GOL	В	202	6/6	0.72	0.16	$61,\!68,\!73,\!75$	0
4	GOL	В	201	6/6	0.74	0.16	44,59,65,68	0
3	Κ	В	206	1/1	0.79	0.13	84,84,84,84	0
4	GOL	В	203	6/6	0.82	0.16	56,60,61,61	0
3	Κ	А	202	1/1	0.86	0.14	78,78,78,78	0
3	Κ	А	201	1/1	0.89	0.26	83,83,83,83	0
3	Κ	В	205	1/1	0.94	0.28	69,69,69,69	0

6.5 Other polymers (i)

There are no such residues in this entry.

