



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 26, 2023 – 05:54 PM EDT

PDB ID : 6D7M
Title : Crystal structure of the W184R/W231R Importin alpha mutant
Authors : Pedersen, L.C.; London, R.E.; Gabel, S.A.
Deposited on : 2018-04-25
Resolution : 2.19 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 ⓘ 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 : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.35.1
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.35.1

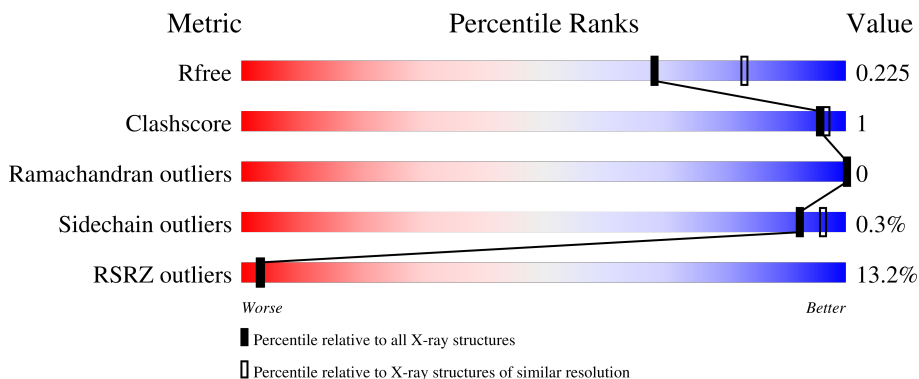
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.19 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}	130704	6864 (2.20-2.16)
Clashscore	141614	7689 (2.20-2.16)
Ramachandran outliers	138981	7564 (2.20-2.16)
Sidechain outliers	138945	7564 (2.20-2.16)
RSRZ outliers	127900	6738 (2.20-2.16)

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	B	510	

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 6428 atoms, of which 3113 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.

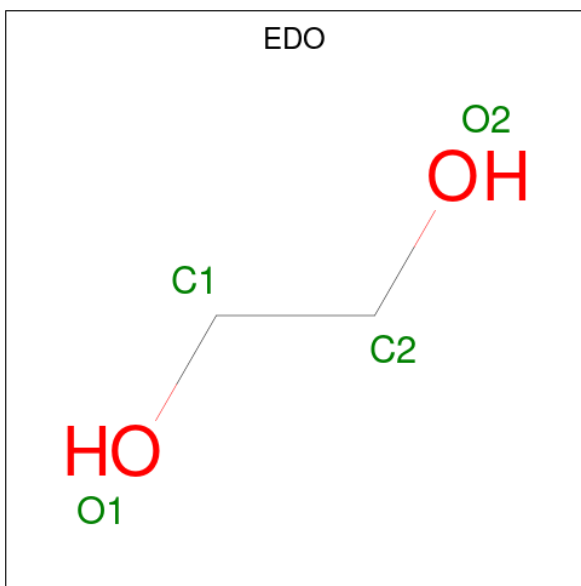
- Molecule 1 is a protein called Peroxidase,Importin subunit alpha-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	B	424	6223	1988	3093	541	591	10	0	4	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	20	MET	-	initiating methionine	PDB ?
B	184	ARG	TRP	engineered mutation	UNP P52293
B	231	ARG	TRP	engineered mutation	UNP P52293

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	H	O	0	0
			8	2	5	1		
2	B	1	Total	C	H	O	0	0
			10	2	6	2		

Continued on next page...

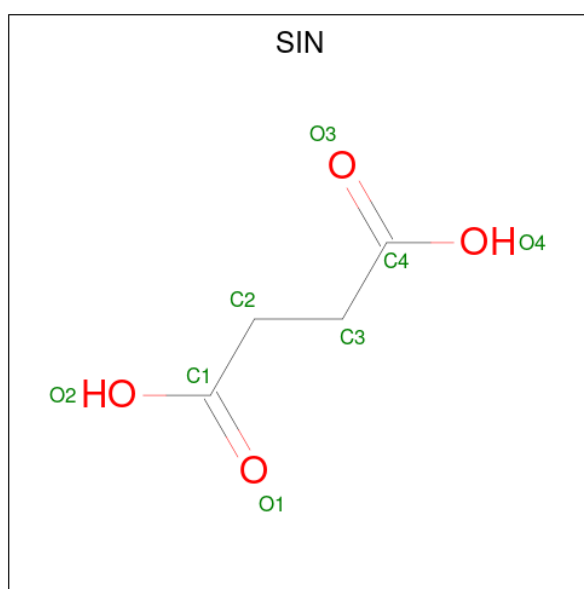
Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
2	B	1	8	2	5	1	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Cl 1 1	0	0

- Molecule 4 is SUCCINIC ACID (three-letter code: SIN) (formula: C₄H₆O₄).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
4	B	1	12	4	4	4	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	162	Total O 166 166	0	4

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	77.81Å 90.16Å 98.39Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.91 – 2.19 38.91 – 2.19	Depositor EDS
% Data completeness (in resolution range)	98.7 (38.91-2.19) 85.2 (38.91-2.19)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.09 (at 2.18Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.191 , 0.222 0.194 , 0.225	Depositor DCC
R_{free} test set	1805 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å ²)	34.4	Xtrriage
Anisotropy	0.390	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 53.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	6428	wwPDB-VP
Average B, all atoms (Å ²)	52.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.17% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: CL, EDO, SIN

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	B	0.35	0/3190	0.52	0/4363

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	B	3130	3093	3082	7	0
2	B	10	16	12	0	0
3	B	1	0	0	0	0
4	B	8	4	4	0	0
5	B	166	0	0	0	0
All	All	3315	3113	3098	7	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 1.

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:B:452:GLU:HG3	1:B:457:THR:HG21	1.83	0.61
1:B:215:VAL:O	1:B:258[A]:ARG:NH2	2.37	0.56
1:B:272:CYS:HB3	1:B:312:PRO:HB2	1.97	0.47
1:B:404:TYR:CD1	1:B:416:LEU:HD22	2.52	0.44
1:B:207:ASP:OD1	1:B:251:GLN:NE2	2.52	0.42
1:B:238:ARG:O	1:B:239:ASN:HB2	2.19	0.41
1:B:425:LEU:HG	1:B:440:ILE:HG23	2.03	0.40

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	B	426/510 (84%)	413 (97%)	13 (3%)	0	100 100

There are no Ramachandran outliers to report.

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	B	322/426 (76%)	320 (99%)	2 (1%)	86 92

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	258[A]	ARG
1	B	258[B]	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains 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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 1 is 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	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	EDO	B	602	-	3,3,3	0.52	0	2,2,2	0.13	0
4	SIN	B	605	-	7,7,7	1.00	0	8,8,8	1.49	1 (12%)
2	EDO	B	601	-	2,2,3	0.44	0	1,1,2	0.28	0
2	EDO	B	603	-	2,2,3	0.53	0	1,1,2	0.28	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
2	EDO	B	602	-	-	1/1/1/1	-
4	SIN	B	605	-	-	2/5/5/5	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	605	SIN	C2-C3-C4	-2.30	108.66	113.60

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	602	EDO	O1-C1-C2-O2
4	B	605	SIN	O2-C1-C2-C3
4	B	605	SIN	O1-C1-C2-C3

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

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	B	424/510 (83%)	0.58	56 (13%) 3 3	29, 43, 94, 120	0

All (56) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	496	PHE	7.8
1	B	454	LEU	5.7
1	B	476	LEU	5.4
1	B	455	GLY	4.8
1	B	477	GLN	4.8
1	B	495	TYR	4.7
1	B	470	LEU	4.4
1	B	457	THR	4.4
1	B	488	SER	4.3
1	B	479	HIS	4.1
1	B	471	ASP	4.1
1	B	487	ALA	4.1
1	B	431	ALA	4.1
1	B	429	LEU	4.0
1	B	73	THR	3.7
1	B	473	ILE	3.7
1	B	449	GLN	3.6
1	B	452	GLU	3.6
1	B	489	LEU	3.6
1	B	451	ALA	3.5
1	B	74	VAL	3.5
1	B	462	ILE	3.4
1	B	448	PHE	3.4
1	B	444	ILE	3.3
1	B	492	ILE	3.3
1	B	385	VAL	3.2
1	B	475	ALA	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	485	TYR	3.1
1	B	159	VAL	3.0
1	B	478	ARG	3.0
1	B	419[A]	CYS	2.9
1	B	493	GLU	2.9
1	B	381	PHE	2.8
1	B	494	LYS	2.7
1	B	383	VAL	2.7
1	B	432	LYS	2.6
1	B	382	LEU	2.6
1	B	461	SER	2.6
1	B	377	GLY	2.5
1	B	380	PRO	2.5
1	B	199	LEU	2.4
1	B	418[A]	HIS	2.4
1	B	378	LEU	2.4
1	B	441	LEU	2.4
1	B	164	ILE	2.4
1	B	384	GLY	2.3
1	B	472	LYS	2.3
1	B	486	LYS	2.3
1	B	480	GLU	2.2
1	B	481	ASN	2.2
1	B	491	LEU	2.2
1	B	445	SER	2.2
1	B	434	THR	2.2
1	B	203	HIS	2.1
1	B	379	VAL	2.1
1	B	163	ALA	2.1

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, 95th 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
2	EDO	B	603	3/4	0.78	0.20	51,61,69,69	0
4	SIN	B	605	8/8	0.82	0.28	56,61,68,68	12
2	EDO	B	602	4/4	0.87	0.17	57,68,72,76	0
2	EDO	B	601	3/4	0.94	0.10	56,67,74,74	0
3	CL	B	604	1/1	0.95	0.09	68,68,68,68	0

6.5 Other polymers [i](#)

There are no such residues in this entry.