



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 11, 2022 – 12:03 PM JST

PDB ID : 7VPW  
Title : Crystal structure of Transportin-1 in complex with BAP1 PY-NLS (residues 706-724)  
Authors : Yang, T.J.; Hsu, S.T.D.  
Deposited on : 2021-10-18  
Resolution : 3.76 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](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.27
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)	:	2.27

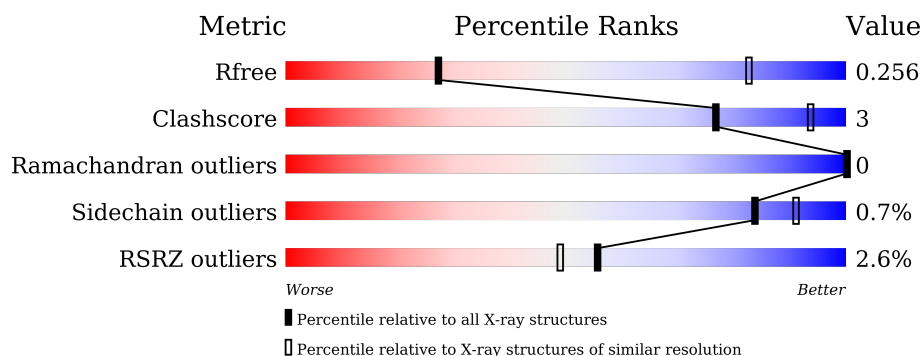
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.76 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	1039 (3.94-3.58)
Clashscore	141614	1051 (3.92-3.60)
Ramachandran outliers	138981	1015 (3.92-3.60)
Sidechain outliers	138945	1011 (3.92-3.60)
RSRZ outliers	127900	1050 (3.96-3.56)

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  $\geq 3$ , 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	19	 26% 11% 63%
2	A	870	 3% 88% 8% .

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6705 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 BRCA1-associated protein 1 (BAP1).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	B	7	Total	C	N	O	0	0	0
			67	41	17	9			

- Molecule 2 is a protein called Transportin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	834	Total	C	N	O	S	0	0	0
			6638	4257	1107	1223	51			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP Q92973
A	-2	GLY	-	expression tag	UNP Q92973
A	-1	SER	-	expression tag	UNP Q92973
A	361	GLY	-	linker	UNP Q92973
A	362	GLY	-	linker	UNP Q92973
A	363	SER	-	linker	UNP Q92973
A	364	GLY	-	linker	UNP Q92973
A	365	GLY	-	linker	UNP Q92973
A	366	SER	-	linker	UNP Q92973
A	367	GLY	-	linker	UNP Q92973

### 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.

- Molecule 1: BRCA1-associated protein 1 (BAP1)

Chain B: 

ILE  
GLY  
ARG  
LEU  
HIS  
LYS  
GLN  
ARG  
LYS  
PRO  
ASP  
ARG  
R718  
R722  
P723  
Y724

- Molecule 2: Transportin-1

Chain A: 

GLY  
GLY  
SER  
LYS  
MET  
GLU  
TYR  
TRP  
K6  
E21  
V33  
E38  
M41  
Q42  
Y43  
L57  
E60  
G69  
K92  
C95  
L96  
L113  
W126  
L130  
P131  
F149  
G150  
A151  
I155  
D166  
L239  
H253  
I257  
R264  
A276  
F279  
L297

S311  
I315  
K319  
GLY  
ASP  
VAL  
GLU  
GLU  
ASP  
GLU  
THR  
ILE  
PRO  
ASP  
SER  
GLU  
GLN  
ASP  
ILE  
ARG  
GLY  
GLY  
SER  
GLY  
SER  
GLY  
ASP  
THR  
ILE  
S371  
R376  
Y390  
L395  
L399  
E415  
L419  
G422  
I440  
I444  
L453  
R464  
L486  
R495

I522  
D580  
K581  
R611  
C612  
L615  
D631  
Q632  
L647  
L648  
S649  
G650  
L651  
A652  
T670  
Q674  
G675  
M676  
R684  
D693  
F709  
H710  
P711  
T715  
N716  
L717  
S736  
M743  
I758  
H770  
R777  
C798  
I817  
M820  
I821  
V828  
F835

A840  
F860  
Q863  
V864  
N868  
F872  
R883  
L884  
A885  
A886  
F887  
Y888  
Q889  
V890

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.90Å 129.69Å 170.21Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	63.86 – 3.76 103.16 – 3.76	Depositor EDS
% Data completeness (in resolution range)	98.3 (63.86-3.76) 93.3 (103.16-3.76)	Depositor EDS
$R_{merge}$	0.21	Depositor
$R_{sym}$	0.23	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.17 (at 3.78Å)	Xtriage
Refinement program	PHENIX 1.19.1_4122	Depositor
R, $R_{free}$	0.229 , 0.258 0.230 , 0.256	Depositor DCC
$R_{free}$ test set	1607 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	124.4	Xtriage
Anisotropy	0.825	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 94.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	6705	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	145.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.60% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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](#)

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	0.28	0/68	0.85	0/88
2	A	0.29	0/6779	0.53	0/9205
All	All	0.29	0/6847	0.54	0/9293

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	67	0	72	0	0
2	A	6638	0	6710	41	0
All	All	6705	0	6782	41	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.

The worst 5 of 41 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)
2:A:38:GLU:HA	2:A:41:ASN:ND2	1.99	0.77
2:A:57:LEU:HD22	2:A:60:GLU:HB2	1.75	0.68

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:649:SER:OG	2:A:693:ASP:OD2	2.10	0.64
2:A:864:VAL:HG22	2:A:868:ASN:HB2	1.84	0.58
2:A:670:THR:O	2:A:674:GLN:HG3	2.04	0.58

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	5/19 (26%)	3 (60%)	2 (40%)	0	100	100
2	A	830/870 (95%)	800 (96%)	30 (4%)	0	100	100
All	All	835/889 (94%)	803 (96%)	32 (4%)	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	7/18 (39%)	5 (71%)	2 (29%)	0	2
2	A	749/777 (96%)	746 (100%)	3 (0%)	91	95
All	All	756/795 (95%)	751 (99%)	5 (1%)	84	91

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	718	ARG
1	B	722	ARG
2	A	264	ARG
2	A	319	LYS
2	A	887	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	716	ASN
2	A	744	GLN

### 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](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	B	7/19 (36%)	-0.14	0 100 100	116, 138, 164, 174	0
2	A	834/870 (95%)	0.42	22 (2%) 56 49	92, 142, 188, 215	0
All	All	841/889 (94%)	0.41	22 (2%) 56 49	92, 142, 188, 215	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	495	ARG	4.4
2	A	717	LEU	3.7
2	A	758	ILE	3.0
2	A	777	ARG	3.0
2	A	835	PHE	3.0

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.