



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 19, 2022 – 02:10 PM EDT

PDB ID : 7TUG  
Title : Crystal structure of Tapasin in complex with PaSta2-Fab  
Authors : Jiang, J.; Natarajan, K.; Taylor, D.K.; Boyd, L.F.; Margulies, D.H.  
Deposited on : 2022-02-02  
Resolution : 3.90 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 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	:	2.29
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)	:	2.29

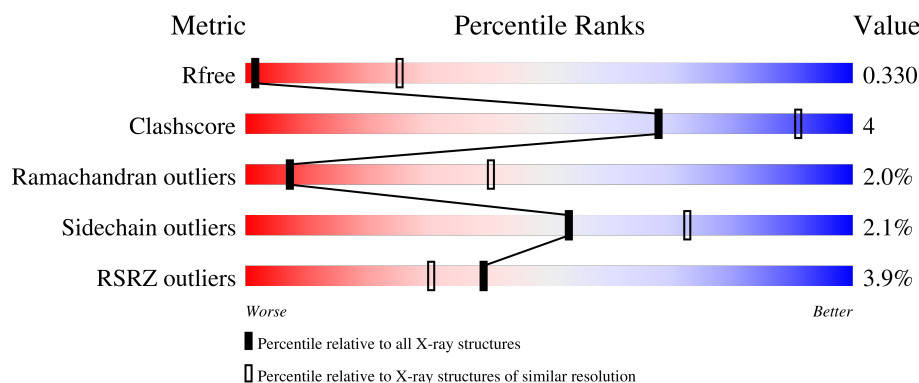
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.90 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	1002 (4.14-3.66)
Clashscore	141614	1004 (4.12-3.68)
Ramachandran outliers	138981	1021 (4.14-3.66)
Sidechain outliers	138945	1014 (4.14-3.66)
RSRZ outliers	127900	1275 (4.20-3.60)

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	D	416	
2	H	230	
3	L	220	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 5304 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 Tapasin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	D	319	Total	C	N	O	S	0	0	0
			2071	1327	360	378	6			

There are 35 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	382	GLY	-	expression tag	UNP O15533
D	383	GLY	-	expression tag	UNP O15533
D	384	LEU	-	expression tag	UNP O15533
D	385	GLU	-	expression tag	UNP O15533
D	386	VAL	-	expression tag	UNP O15533
D	387	LEU	-	expression tag	UNP O15533
D	388	PHE	-	expression tag	UNP O15533
D	389	GLN	-	expression tag	UNP O15533
D	390	GLY	-	expression tag	UNP O15533
D	391	PRO	-	expression tag	UNP O15533
D	392	GLY	-	expression tag	UNP O15533
D	393	GLY	-	expression tag	UNP O15533
D	394	GLY	-	expression tag	UNP O15533
D	395	LEU	-	expression tag	UNP O15533
D	396	ASN	-	expression tag	UNP O15533
D	397	ASP	-	expression tag	UNP O15533
D	398	ILE	-	expression tag	UNP O15533
D	399	PHE	-	expression tag	UNP O15533
D	400	GLU	-	expression tag	UNP O15533
D	401	ALA	-	expression tag	UNP O15533
D	402	GLN	-	expression tag	UNP O15533
D	403	LYS	-	expression tag	UNP O15533
D	404	ILE	-	expression tag	UNP O15533
D	405	GLU	-	expression tag	UNP O15533
D	406	TRP	-	expression tag	UNP O15533
D	407	HIS	-	expression tag	UNP O15533
D	408	GLU	-	expression tag	UNP O15533

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	409	GLY	-	expression tag	UNP O15533
D	410	GLY	-	expression tag	UNP O15533
D	411	HIS	-	expression tag	UNP O15533
D	412	HIS	-	expression tag	UNP O15533
D	413	HIS	-	expression tag	UNP O15533
D	414	HIS	-	expression tag	UNP O15533
D	415	HIS	-	expression tag	UNP O15533
D	416	HIS	-	expression tag	UNP O15533

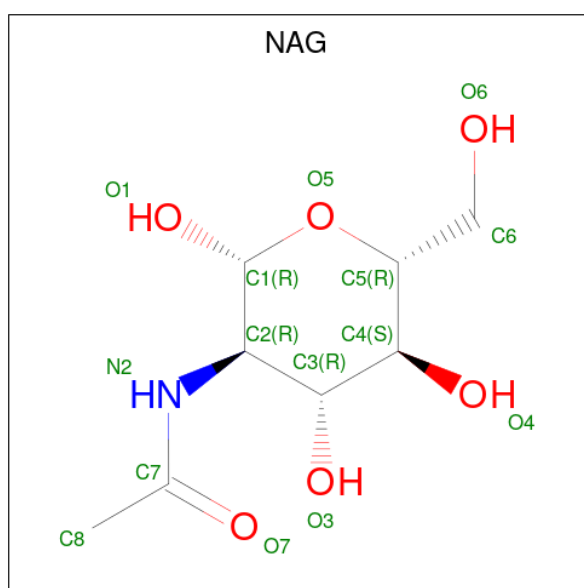
- Molecule 2 is a protein called PaSta2 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	216	Total	C	N	O	S	5	0	0
			1580	1007	257	309	7			

- Molecule 3 is a protein called PaSta2 Fab kappa light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	217	Total	C	N	O	S	3	0	0
			1639	1027	273	332	7			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

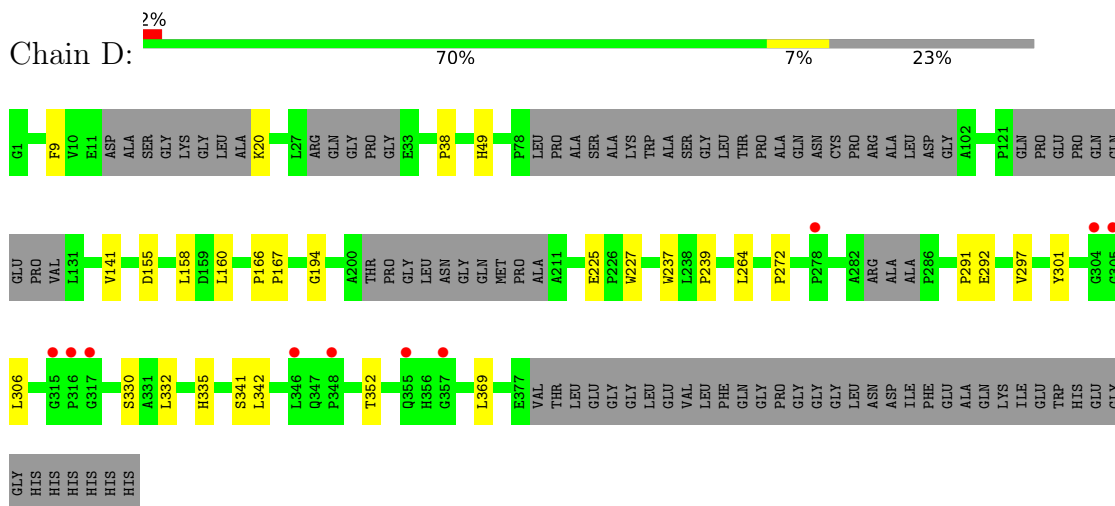


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	D	1	Total	C	N	O	0	0
			14	8	1	5		

### 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: Tapasin



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	101.09Å 168.82Å 108.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	54.33 – 3.90 67.78 – 3.90	Depositor EDS
% Data completeness (in resolution range)	93.9 (54.33-3.90) 93.9 (67.78-3.90)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.37 (at 3.89Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, $R_{free}$	0.282 , 0.332 0.282 , 0.330	Depositor DCC
$R_{free}$ test set	409 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	67.7	Xtriage
Anisotropy	0.904	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 43.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.005 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.049 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
$F_o, F_c$ correlation	0.80	EDS
Total number of atoms	5304	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.65% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG

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	D	0.25	0/2137	0.44	0/2971
2	H	0.25	0/1624	0.49	0/2232
3	L	0.24	0/1675	0.46	0/2279
All	All	0.25	0/5436	0.46	0/7482

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

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	D	2071	0	1766	11	0
2	H	1580	0	1498	20	0
3	L	1639	0	1556	8	0
4	D	14	0	13	0	0
All	All	5304	0	4833	38	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 4.

The worst 5 of 38 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:H:141:THR:HG22	2:H:186:THR:HG22	1.81	0.62
3:L:18:LYS:HA	3:L:81:ILE:O	1.99	0.62
2:H:61:ASN:OD1	2:H:62:GLU:N	2.36	0.57
1:D:335:HIS:NE2	1:D:341:SER:OG	2.30	0.57
3:L:201:GLU:HG2	3:L:212:VAL:HG22	1.87	0.54

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	D	305/416 (73%)	270 (88%)	29 (10%)	6 (2%)	7	40
2	H	214/230 (93%)	195 (91%)	13 (6%)	6 (3%)	5	34
3	L	215/220 (98%)	195 (91%)	17 (8%)	3 (1%)	11	46
All	All	734/866 (85%)	660 (90%)	59 (8%)	15 (2%)	7	40

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	38	PRO
1	D	291	PRO
2	H	150	PHE
2	H	151	PRO
2	H	152	GLU

### 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	D	169/328 (52%)	166 (98%)	3 (2%)	59	77
2	H	167/193 (86%)	162 (97%)	5 (3%)	41	64
3	L	183/197 (93%)	180 (98%)	3 (2%)	62	79
All	All	519/718 (72%)	508 (98%)	11 (2%)	53	73

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	157	THR
3	L	45	LYS
3	L	186	THR
3	L	60	ARG
2	H	145	LEU

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](#)

1 ligand is modelled in this entry.

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	$\# Z  > 2$
4	NAG	D	501	1	14,14,15	0.23	0	17,19,21	0.43	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	NAG	D	501	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

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<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	D	319/416 (76%)	0.09	10 (3%) 49 38	28, 83, 156, 204	0
2	H	216/230 (93%)	0.33	8 (3%) 41 32	27, 66, 107, 168	4 (1%)
3	L	217/220 (98%)	0.47	11 (5%) 28 23	38, 69, 119, 139	4 (1%)
All	All	752/866 (86%)	0.27	29 (3%) 39 30	27, 72, 142, 204	8 (1%)

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	316	PRO	6.5
2	H	136	THR	5.5
2	H	142	LEU	3.8
3	L	202	ALA	3.7
3	L	162	GLN	3.4

### 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 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<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
4	NAG	D	501	14/15	0.74	0.28	69,93,98,103	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.