



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 23, 2023 – 04:53 AM EDT

PDB ID : 3CVL
Title : Structure of Peroxisomal Targeting Signal 1 (PTS1) binding domain of Trypanosoma brucei Peroxin 5 (TbPEX5) complexed to T. brucei Phosphofruktokinase (PFK) PTS1 peptide
Authors : Sampathkumar, P.; Roach, C.; Michels, P.A.M.; Hol, W.G.J.
Deposited on : 2008-04-18
Resolution : 2.15 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

<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
Xtriage (Phenix) : 1.13
EDS : 2.35
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

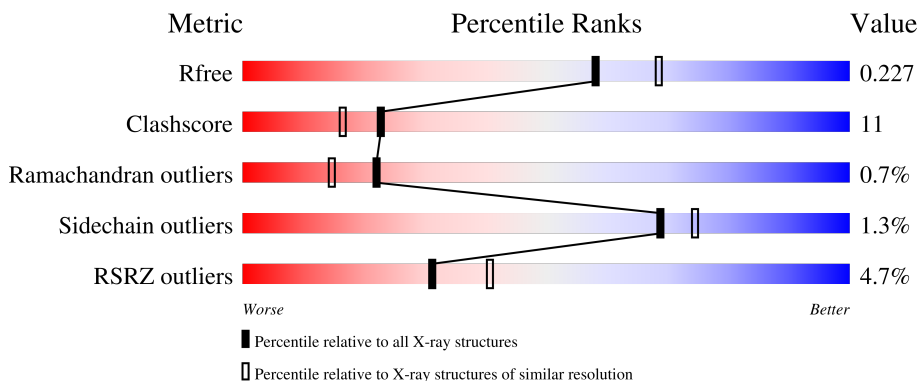
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.15 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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	A	327	 4% 70% 18% 10%
2	B	7	 14% 57% 14% 29%

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 2399 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 Peroxisome targeting signal 1 receptor PEX5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	294	2292	1440	397	441	14	0	6	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	329	GLY	-	expression tag	UNP Q9U7C3
A	330	HIS	-	expression tag	UNP Q9U7C3
A	331	MET	-	expression tag	UNP Q9U7C3
A	378	ALA	LYS	engineered mutation	UNP Q9U7C3
A	379	ALA	GLU	engineered mutation	UNP Q9U7C3

- Molecule 2 is a protein called T. brucei PFK PTS1 peptide Ac-HEELAKL.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	5	40	26	6	8	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	65	Total	O	0	0
			65	65		
3	B	2	Total	O	0	0
			2	2		

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	44.39Å 66.36Å 49.68Å 90.00° 105.13° 90.00°	Depositor
Resolution (Å)	26.17 – 2.15 26.17 – 2.15	Depositor EDS
% Data completeness (in resolution range)	97.4 (26.17-2.15) 97.4 (26.17-2.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.12 (at 2.15Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.212 , 0.278 0.221 , 0.227	Depositor DCC
R_{free} test set	743 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	27.1	Xtrriage
Anisotropy	0.329	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 48.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	2399	wwPDB-VP
Average B, all atoms (Å ²)	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.43% 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](#)

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	0.74	2/2358 (0.1%)	0.72	2/3217 (0.1%)
2	B	0.58	0/39	0.71	0/49
All	All	0.74	2/2397 (0.1%)	0.72	2/3266 (0.1%)

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	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	345	PRO	C-N	-9.42	1.12	1.34
1	A	346	TYR	C-N	-9.24	1.12	1.34

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	337	ASP	CB-CG-OD1	6.54	124.19	118.30
1	A	346	TYR	O-C-N	-6.31	112.61	122.70

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	363	ALA	Peptide
1	A	364	ASN	Peptide
1	A	602	THR	Peptide

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	A	2292	0	2158	52	0
2	B	40	0	45	1	0
3	A	65	0	0	3	0
3	B	2	0	0	0	0
All	All	2399	0	2203	52	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 11.

The worst 5 of 52 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:453:LEU:HD11	1:A:495:MET:HE3	1.46	0.95
1:A:453:LEU:HD11	1:A:495:MET:CE	1.99	0.92
1:A:343[B]:ASN:ND2	1:A:602:THR:HG23	1.89	0.86
1:A:595:ILE:HG22	1:A:641:VAL:HG11	1.67	0.75
1:A:604:PRO:HD2	3:A:746:HOH:O	1.90	0.71

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	A	294/327 (90%)	285 (97%)	7 (2%)	2 (1%)	22 15
2	B	3/7 (43%)	3 (100%)	0	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	297/334 (89%)	288 (97%)	7 (2%)	2 (1%)	22 15

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	364	ASN
1	A	600	GLY

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	230/266 (86%)	226 (98%)	4 (2%)	60 65
2	B	4/6 (67%)	4 (100%)	0	100 100
All	All	234/272 (86%)	230 (98%)	4 (2%)	69 65

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

Mol	Chain	Res	Type
1	A	535[A]	GLN
1	A	535[B]	GLN
1	A	567	TYR
1	A	598	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	449	GLN
1	A	551	GLN
1	A	639	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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	345:PRO	C	346:TYR	N	1.12
1	A	346:TYR	C	347:MET	N	1.12

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	294/327 (89%)	0.36	13 (4%) 34 43	19, 26, 41, 56	0
2	B	5/7 (71%)	1.23	1 (20%) 1 1	29, 32, 43, 45	0
All	All	299/334 (89%)	0.37	14 (4%) 31 41	19, 26, 41, 56	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	348	TYR	4.4
1	A	343[A]	ASN	4.3
2	B	4	LEU	3.8
1	A	601	GLY	3.7
1	A	602	THR	3.7

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.