



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 10:09 pm GMT

PDB ID : 1MS3
Title : Monoclinic form of Trypanosoma cruzi trans-sialidase
Authors : Buschiazzo, A.; Amaya, M.F.; Cremona, M.L.; Frasch, A.C.; Alzari, P.M.
Deposited on : 2002-09-19
Resolution : 1.65 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

<http://wwpdb.org/validation/2016/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) : NOT EXECUTED
EDS : NOT EXECUTED
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc28949

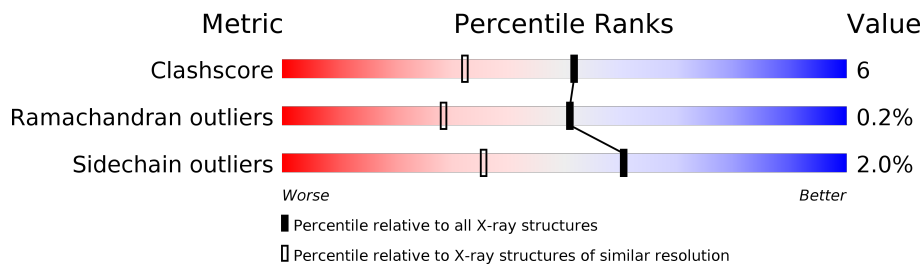
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.65 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))
Clashscore	112137	1468 (1.66-1.66)
Ramachandran outliers	110173	1438 (1.66-1.66)
Sidechain outliers	110143	1438 (1.66-1.66)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	648	
1	B	648	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 10960 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 trans-sialidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	623	Total	C	N	O	S	0	10	0
			4914	3111	856	933	14			
1	B	623	Total	C	N	O	S	0	7	0
			4894	3099	853	928	14			

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	MET	-	EXPRESSION TAG	UNP Q26964
A	-12	GLY	-	EXPRESSION TAG	UNP Q26964
A	-11	GLY	-	EXPRESSION TAG	UNP Q26964
A	-10	SER	-	EXPRESSION TAG	UNP Q26964
A	-9	HIS	-	EXPRESSION TAG	UNP Q26964
A	-8	HIS	-	EXPRESSION TAG	UNP Q26964
A	-7	HIS	-	EXPRESSION TAG	UNP Q26964
A	-6	HIS	-	EXPRESSION TAG	UNP Q26964
A	-5	HIS	-	EXPRESSION TAG	UNP Q26964
A	-4	HIS	-	EXPRESSION TAG	UNP Q26964
A	-3	GLY	-	EXPRESSION TAG	UNP Q26964
A	-2	MET	-	EXPRESSION TAG	UNP Q26964
A	-1	ALA	-	EXPRESSION TAG	UNP Q26964
A	0	SER	-	EXPRESSION TAG	UNP Q26964
A	58	PHE	ASN	ENGINEERED	UNP Q26964
A	262	THR	SER	SEE REMARK 999	UNP Q26964
A	476	HIS	ARG	SEE REMARK 999	UNP Q26964
A	484	LEU	VAL	SEE REMARK 999	UNP Q26964
A	495	LYS	SER	ENGINEERED	UNP Q26964
A	496	GLY	VAL	ENGINEERED	UNP Q26964
A	520	LYS	GLU	ENGINEERED	UNP Q26964
A	558	VAL	GLU	SEE REMARK 999	UNP Q26964
A	593	GLY	ASP	ENGINEERED	UNP Q26964
A	597	ASP	ILE	ENGINEERED	UNP Q26964
A	599	ARG	HIS	ENGINEERED	UNP Q26964

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-13	MET	-	EXPRESSION TAG	UNP Q26964
B	-12	GLY	-	EXPRESSION TAG	UNP Q26964
B	-11	GLY	-	EXPRESSION TAG	UNP Q26964
B	-10	SER	-	EXPRESSION TAG	UNP Q26964
B	-9	HIS	-	EXPRESSION TAG	UNP Q26964
B	-8	HIS	-	EXPRESSION TAG	UNP Q26964
B	-7	HIS	-	EXPRESSION TAG	UNP Q26964
B	-6	HIS	-	EXPRESSION TAG	UNP Q26964
B	-5	HIS	-	EXPRESSION TAG	UNP Q26964
B	-4	HIS	-	EXPRESSION TAG	UNP Q26964
B	-3	GLY	-	EXPRESSION TAG	UNP Q26964
B	-2	MET	-	EXPRESSION TAG	UNP Q26964
B	-1	ALA	-	EXPRESSION TAG	UNP Q26964
B	0	SER	-	EXPRESSION TAG	UNP Q26964
B	58	PHE	ASN	ENGINEERED	UNP Q26964
B	262	THR	SER	SEE REMARK 999	UNP Q26964
B	476	HIS	ARG	SEE REMARK 999	UNP Q26964
B	484	LEU	VAL	SEE REMARK 999	UNP Q26964
B	495	LYS	SER	ENGINEERED	UNP Q26964
B	496	GLY	VAL	ENGINEERED	UNP Q26964
B	520	LYS	GLU	ENGINEERED	UNP Q26964
B	558	VAL	GLU	SEE REMARK 999	UNP Q26964
B	593	GLY	ASP	ENGINEERED	UNP Q26964
B	597	ASP	ILE	ENGINEERED	UNP Q26964
B	599	ARG	HIS	ENGINEERED	UNP Q26964

- Molecule 2 is water.

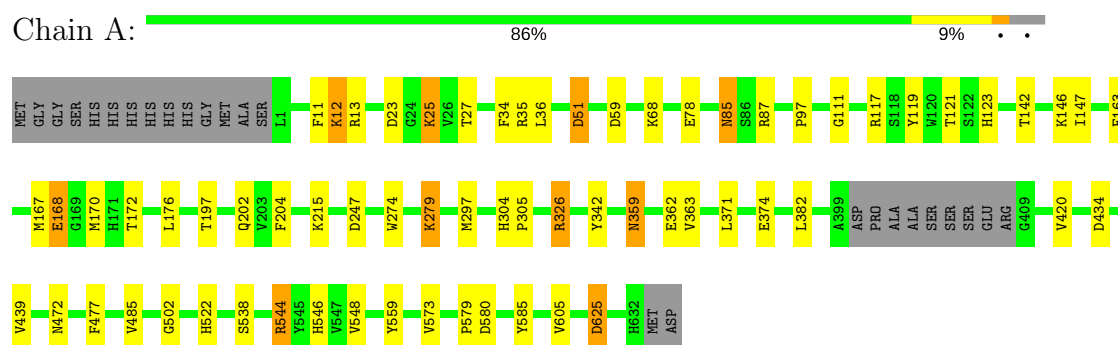
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	595	Total O 595 595	0	0
2	B	557	Total O 557 557	0	0

3 Residue-property plots

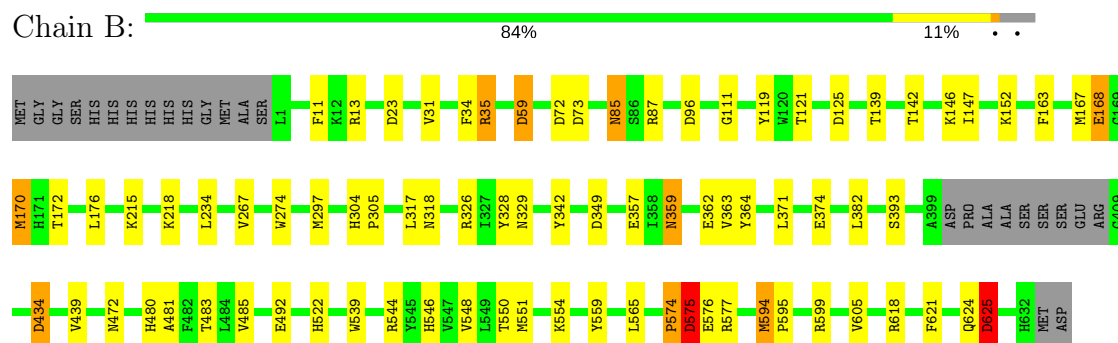
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

Note EDS was not executed.

- Molecule 1: trans-sialidase



- Molecule 1: trans-sialidase



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	63.22Å 128.85Å 86.62Å 90.00° 90.99° 90.00°	Depositor
Resolution (Å)	14.97 – 1.65	Depositor
% Data completeness (in resolution range)	94.8 (14.97-1.65)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
Refinement program	REFMAC 5	Depositor
R, R_{free}	0.163 , 0.192	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	10960	wwPDB-VP
Average B, all atoms (Å ²)	20.0	wwPDB-VP

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	0.61	0/5030	0.88	7/6831 (0.1%)
1	B	0.59	0/5010	0.92	15/6803 (0.2%)
All	All	0.60	0/10040	0.90	22/13634 (0.2%)

There are no bond length outliers.

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	35	ARG	NE-CZ-NH2	-13.86	113.37	120.30
1	B	35	ARG	NE-CZ-NH1	11.10	125.85	120.30
1	B	575	ASP	CB-CG-OD2	10.39	127.65	118.30
1	A	247	ASP	CB-CG-OD2	7.36	124.92	118.30
1	A	625	ASP	CB-CG-OD2	6.85	124.46	118.30

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	A	4914	0	4824	58	0
1	B	4894	0	4805	73	0
2	A	595	0	0	5	2
2	B	557	0	0	10	2
All	All	10960	0	9629	125	2

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 6.

The worst 5 of 125 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:544:ARG:NH2	1:A:625:ASP:OD1	1.60	1.35
1:B:35:ARG:HD2	1:B:357:GLU:OE1	1.23	1.24
1:B:594:MET:HA	1:B:594:MET:HE2	1.23	1.09
1:B:35:ARG:CD	1:B:357:GLU:OE1	2.03	1.06
1:B:167:MET:HB2	1:B:170:MET:HE3	1.57	0.85

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:1180:HOH:O	2:B:885:HOH:O[1_455]	1.93	0.27
2:A:986:HOH:O	2:B:1076:HOH:O[1_454]	2.18	0.02

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	629/648 (97%)	605 (96%)	23 (4%)	1 (0%)	51	28
1	B	626/648 (97%)	599 (96%)	25 (4%)	2 (0%)	44	23
All	All	1255/1296 (97%)	1204 (96%)	48 (4%)	3 (0%)	51	28

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	168	GLU
1	B	575	ASP
1	B	574	PRO

5.3.2 Protein sidechains ⓘ

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	536/547 (98%)	525 (98%)	11 (2%)	59	32
1	B	533/547 (97%)	522 (98%)	11 (2%)	59	32
All	All	1069/1094 (98%)	1047 (98%)	22 (2%)	60	32

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	359	ASN
1	B	146	LYS
1	B	594	MET
1	A	544	ARG
1	B	85	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	604	ASN
1	B	85	ASN
1	B	603	ASN
1	A	603	ASN
1	B	604	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates 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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.