wwPDB X-ray Structure Validation Summary Report (i)

Oct 17, 2021 – 06:12 AM EDT

PDB ID	:	1MR5
Title	:	Orthorhombic form of Trypanosoma cruzi trans-sialidase
Authors	:	Buschiazzo, A.; Amaya, M.F.; Cremona, M.L.; Frasch, A.C.; Alzari, P.M.
Deposited on	:	2002-09-18
Resolution	:	2.25 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 (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

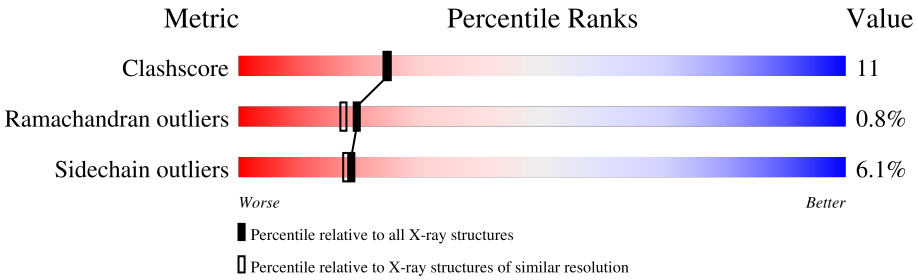
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.25 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	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)

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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	648	73% 19%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4974 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
			Total	C	N	O	S			
1	A	621	4842	3063	847	916	16	0	2	0

There are 25 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	MET	-	expression tag	UNP Q26966
A	-12	GLY	-	expression tag	UNP Q26966
A	-11	GLY	-	expression tag	UNP Q26966
A	-10	SER	-	expression tag	UNP Q26966
A	-9	HIS	-	expression tag	UNP Q26966
A	-8	HIS	-	expression tag	UNP Q26966
A	-7	HIS	-	expression tag	UNP Q26966
A	-6	HIS	-	expression tag	UNP Q26966
A	-5	HIS	-	expression tag	UNP Q26966
A	-4	HIS	-	expression tag	UNP Q26966
A	-3	GLY	-	expression tag	UNP Q26966
A	-2	MET	-	expression tag	UNP Q26966
A	-1	ALA	-	expression tag	UNP Q26966
A	0	SER	-	expression tag	UNP Q26966
A	58	PHE	ASN	engineered mutation	UNP Q26966
A	262	THR	SER	SEE REMARK 999	UNP Q26966
A	476	HIS	ARG	SEE REMARK 999	UNP Q26966
A	484	LEU	VAL	SEE REMARK 999	UNP Q26966
A	495	LYS	SER	engineered mutation	UNP Q26966
A	496	GLY	VAL	engineered mutation	UNP Q26966
A	520	LYS	GLU	engineered mutation	UNP Q26966
A	558	VAL	GLU	SEE REMARK 999	UNP Q26966
A	593	GLY	ASP	engineered mutation	UNP Q26966
A	597	ASP	ILE	engineered mutation	UNP Q26966
A	599	ARG	HIS	engineered mutation	UNP Q26966

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	132	Total	O	0	0
			132	132		

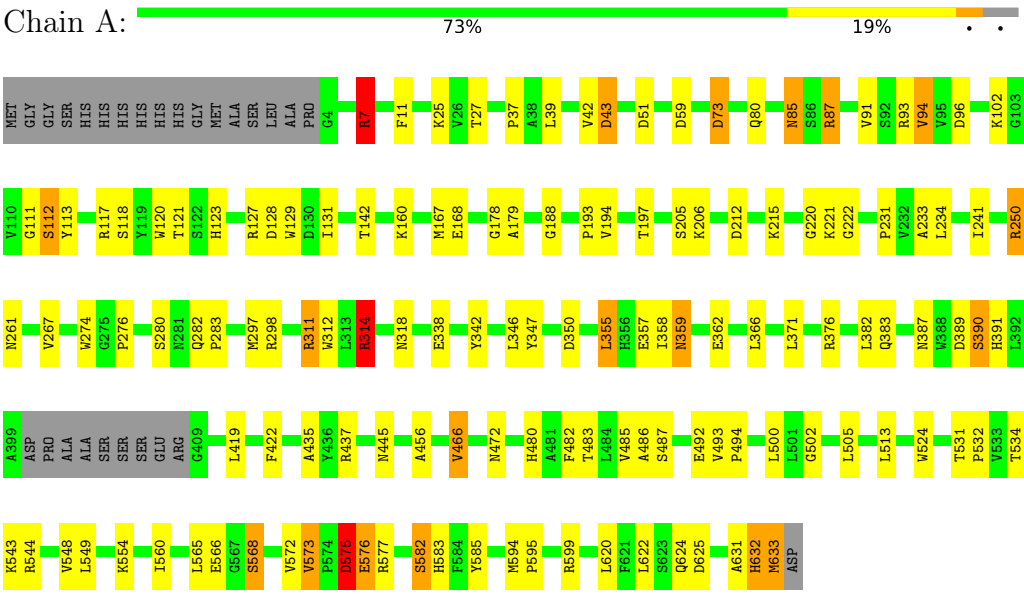
3 Residue-property plots (i)

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. 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. 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

Chain A:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	87.03Å 88.02Å 89.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	27.84 – 2.25	Depositor
% Data completeness (in resolution range)	99.5 (27.84-2.25)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
Refinement program	REFMAC 5.0	Depositor
R, R_{free}	0.218 , 0.274	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4974	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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	0.99	0/4954	1.09	20/6722 (0.3%)

There are no bond length outliers.

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	96	ASP	CB-CG-OD2	9.93	127.24	118.30
1	A	258	ASP	CB-CG-OD2	8.83	126.25	118.30
1	A	212	ASP	CB-CG-OD2	8.76	126.18	118.30
1	A	250	ARG	NE-CZ-NH2	-8.44	116.08	120.30
1	A	43	ASP	CB-CG-OD2	8.43	125.89	118.30

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	A	4842	0	4764	105	0
2	A	132	0	0	5	0
All	All	4974	0	4764	105	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 105 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:505:LEU:HD11	1:A:513:LEU:HB2	1.46	0.97
1:A:383:GLN:HE21	1:A:387:ASN:ND2	1.71	0.89
1:A:312:TRP:O	1:A:314:ARG:HD2	1.80	0.81
1:A:312:TRP:O	1:A:314:ARG:CD	2.34	0.76
1:A:359:ASN:HD21	1:A:362:GLU:H	1.35	0.75

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	619/648 (96%)	586 (95%)	28 (4%)	5 (1%)	19 17

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	632	HIS
1	A	575	ASP
1	A	576	GLU
1	A	398	PRO
1	A	466	VAL

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	527/547 (96%)	495 (94%)	32 (6%)	18 18

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	573	VAL
1	A	577	ARG
1	A	250	ARG
1	A	215	LYS
1	A	582	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	604	ASN
1	A	632	HIS
1	A	359	ASN
1	A	383	GLN
1	A	472	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

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

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.