WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 13, 2024 – 07:58 pm BST

PDB ID	:	4BP8
Title	:	Oligopeptidase B from Trypanosoma brucei - open form
Authors	:	Canning, P.; Rea, D.; Morty, R.; Fulop, V.
Deposited on	:	2013-05-23
Resolution	:	2.40 Å(reported)

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 2.40 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   Percentile Ranks   Value

Rfree — 0.244
Clashscore — 10
Ramachandran outliers — 1.1%
Sidechain outliers — 7.6%
RSRZ outliers — 0.8%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
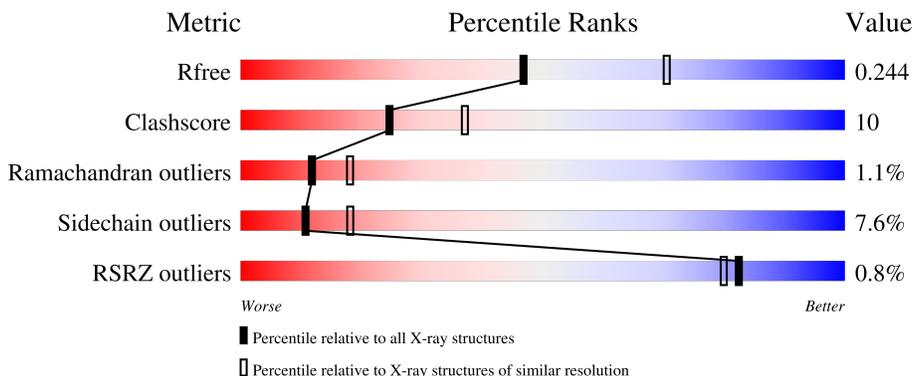
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	4642 (2.40-2.40)
Clashscore	180529	5218 (2.40-2.40)
Ramachandran outliers	177936	5158 (2.40-2.40)
Sidechain outliers	177891	5159 (2.40-2.40)
RSRZ outliers	164620	4642 (2.40-2.40)

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% 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	715	80% 15% · ·
1	B	715	79% 17% · ·

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 12080 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 OLIGOPEPTIDASE B.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	713	Total	C	N	O	S	Se	0	0	1
			5651	3577	979	1064	14	17			
1	B	713	Total	C	N	O	S	Se	0	0	1
			5651	3577	979	1064	14	17			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	397	Total	O	0	0
			397	397		
2	B	381	Total	O	0	0
			381	381		

3 Residue-property plots ⓘ

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: OLIGOPEPTIDASE B

● Molecule 1: OLIGOPEPTIDASE B

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	124.14Å 124.14Å 249.14Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	65.73 – 2.40 65.73 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.3 (65.73-2.40) 99.3 (65.73-2.40)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.80 (at 2.40Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.184 , 0.239 0.193 , 0.244	Depositor DCC
R_{free} test set	3521 reflections (4.05%)	wwPDB-VP
Wilson B-factor (Å^2)	46.0	Xtriage
Anisotropy	0.463	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 36.7	EDS
L-test for twinning [2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.021 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12080	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.24% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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 ⓘ

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	0.76	2/5772 (0.0%)	0.90	16/7805 (0.2%)
1	B	0.75	1/5772 (0.0%)	0.86	15/7805 (0.2%)
All	All	0.76	3/11544 (0.0%)	0.88	31/15610 (0.2%)

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	2
1	B	0	3
All	All	0	5

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	209	ASP	CB-CG	7.69	1.67	1.51
1	A	151	CYS	CB-SG	-6.77	1.70	1.82
1	A	209	ASP	CB-CG	5.54	1.63	1.51

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	576	ARG	NE-CZ-NH1	-12.12	114.24	120.30
1	A	354	ARG	NE-CZ-NH2	-10.82	114.89	120.30
1	A	156	PRO	C-N-CD	-10.74	96.96	120.60
1	B	49	ARG	NE-CZ-NH2	-9.89	115.36	120.30
1	B	463	LEU	CA-CB-CG	9.67	137.54	115.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	156	PRO	Peptide
1	A	482	TYR	Peptide
1	B	156	PRO	Peptide
1	B	179	VAL	Peptide
1	B	482	TYR	Peptide

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	5651	0	5536	99	0
1	B	5651	0	5536	118	0
2	A	397	0	0	19	0
2	B	381	0	0	21	0
All	All	12080	0	11072	216	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 10.

The worst 5 of 216 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:576:ARG:HH11	1:A:576:ARG:CG	1.47	1.26
1:B:576:ARG:HH11	1:B:576:ARG:HG2	1.03	1.19
1:A:180:ARG:N	1:A:180:ARG:HD2	1.48	1.12
1:A:614:PRO:HB3	1:A:623:MSE:CE	1.78	1.12
1:A:576:ARG:NH1	1:A:576:ARG:HG2	1.46	1.11

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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	711/715 (99%)	674 (95%)	29 (4%)	8 (1%)	12	18
1	B	711/715 (99%)	673 (95%)	30 (4%)	8 (1%)	12	18
All	All	1422/1430 (99%)	1347 (95%)	59 (4%)	16 (1%)	12	18

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	157	PRO
1	A	687	SER
1	B	157	PRO
1	A	121	ALA
1	A	285	ARG

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	616/601 (102%)	567 (92%)	49 (8%)	10	16
1	B	616/601 (102%)	571 (93%)	45 (7%)	11	20
All	All	1232/1202 (102%)	1138 (92%)	94 (8%)	11	18

5 of 94 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	139	LYS
1	B	351	LEU
1	B	162	VAL
1	B	263	SER
1	B	393	CYS

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	B	370	ASN
1	B	556	GLN
1	A	643	GLN
1	A	699	GLN
1	B	224	GLN

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 oligosaccharides 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 ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	696/715 (97%)	-0.56	4 (0%)	85	83	30, 45, 68, 90	0
1	B	696/715 (97%)	-0.49	7 (1%)	79	76	29, 46, 67, 92	0
All	All	1392/1430 (97%)	-0.53	11 (0%)	82	80	29, 45, 68, 92	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	209	ASP	4.0
1	A	3	THR	3.4
1	A	368	PRO	2.9
1	B	715	LYS	2.9
1	B	371	LEU	2.7

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK