WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 12:26 am BST

PDB ID	:	3V89
Title	:	The crystal structure of transferrin binding protein A (TbpA) from Neisseria meningitidis serogroup B in complex with the C-lobe of human transferrin
Authors	:	Noinaj, N.; Oke, M.; Easley, N.; Zak, O.; Aisen, P.; Buchanan, S.K.
Deposited on	:	2011-12-22
Resolution	:	3.10 Å(reported)

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

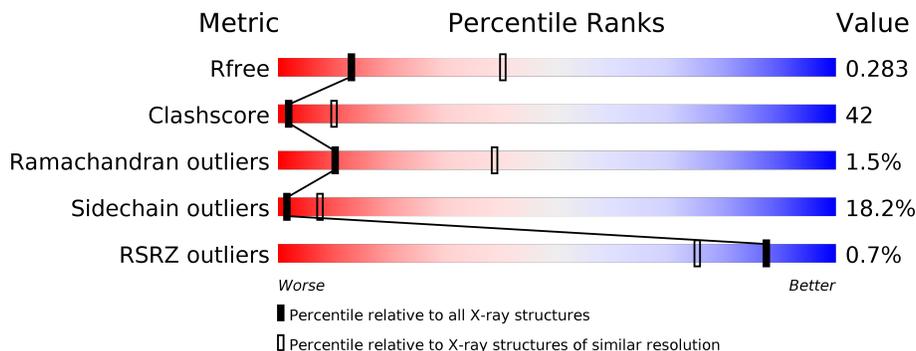|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 3.10 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		0.283
Clashscore		42
Ramachandran outliers		1.5%
Sidechain outliers		18.2%
RSRZ outliers		0.7%

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}	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 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	904	37% 45% 12% · 6%
2	B	343	47% 45% 6% ·

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 9110 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 Transferrin-binding protein A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	853	6558	4091	1182	1274	11	0	0	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	12	MET	-	EXPRESSION TAG	UNP Q9JPJ0
A	13	ASP	-	EXPRESSION TAG	UNP Q9JPJ0
A	14	ILE	-	EXPRESSION TAG	UNP Q9JPJ0
A	15	HIS	-	EXPRESSION TAG	UNP Q9JPJ0
A	16	HIS	-	EXPRESSION TAG	UNP Q9JPJ0
A	17	HIS	-	EXPRESSION TAG	UNP Q9JPJ0
A	18	HIS	-	EXPRESSION TAG	UNP Q9JPJ0
A	19	HIS	-	EXPRESSION TAG	UNP Q9JPJ0
A	20	HIS	-	EXPRESSION TAG	UNP Q9JPJ0
A	21	HIS	-	EXPRESSION TAG	UNP Q9JPJ0
A	22	HIS	-	EXPRESSION TAG	UNP Q9JPJ0
A	23	HIS	-	EXPRESSION TAG	UNP Q9JPJ0
A	24	HIS	-	EXPRESSION TAG	UNP Q9JPJ0

- Molecule 2 is a protein called Serotransferrin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	339	2552	1587	436	503	26	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	429	VAL	ILE	VARIANT	UNP P02787

3 Residue-property plots (i)

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.

• Molecule 1: Transferrin-binding protein A

K303 W304 G805 V806 N807 G808 M809 L810 T811 Y812 S813 K314 A815 K316 E817 I818 T19 E820 L821 R825 A826 L827 L828 N829 G830 N831 S832 T338 A339 R840 R841 T842 R843 P844 W845 Y846 I847 V848 Y353 Y854 T855 V856 K857 R858 H859 F860 T861 L862 R863 V366 L869 L870 N871 Y872 R873

Y874 V875 T876 W877 E878 N879 V880 R881 Q882 T883 A884 V888 N889 Q890 H891 K892 N893 V894 G895 V896 Y897 N898 R899 Y900 A901 R905 N906 Y907 L911 E912 F915

- Molecule 2: Serotransferrin

Chain B:

47%   45%   6% ·

D337 K340 P341 V342 K343 W344 L347 S348 E351 K354 C355 D356 E357 W358 S359 V360 N361 I366 E367 C368 V369 T374 E375 D376 C377 I378 A379 K380 I381 M382 N383 G384 E385 A386 D387 A388 N389 S390 L391 D392 G393 G394 F395 V396 Y397 I398 A399 G400 K401 C402 G403 L404 V407

L408 Y412 D416 N417 T421 P422 G425 Y426 F427 K434 D438 L439 T440 W441 L444 K445 G446 K447 A453 T457 I462 P463 M464 G465 L466 L467 Y468 I471 N472 H473 C474 S481 E482 G483 C484 A485 P486 G487 S488 K489 S492 S493 L494 L497 C498

N504 L505 C506 E507 P508 N509 N510 K511 E512 Y515 G516 Y517 T518 F521 R522 C523 L524 V525 E526 K527 A531 K534 H535 Q536 T537 V538 P539 Q540 N541 T542 G543 G544 K545 N546 P547 A551 K552 N53 L554 N55 E556 E560 L561 L562 T567 R568 K569 P70 E573 C577 H578

L579 A580 R581 A582 P583 N584 V587 V88 T589 R590 K593 C596 L601 R602 Q603 Q604 Q605 H606 L607 F608 G609 S610 N611 V612 T613 D614 C615 S616 G617 N618 F619 R623 S624 D628 L629 L630 F631 R632 V636 L641 N645 T646 Y647 E648 L651 G652 E653 E654 Y655 V656

K657 A658 V659 G660 N661 L662 R663 K664 C665 S666 T667 L670 L671 C674 T675 PHE ARG ARG PRO

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	58.05Å 107.59Å 130.72Å 90.00° 94.48° 90.00°	Depositor
Resolution (Å)	29.96 – 3.10 41.69 – 3.09	Depositor EDS
% Data completeness (in resolution range)	96.8 (29.96-3.10) 91.6 (41.69-3.09)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$<I/\sigma(I)>$ [1]	1.77 (at 3.06Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.2_865)	Depositor
R, R_{free}	0.224 , 0.287 0.225 , 0.283	Depositor DCC
R_{free} test set	1407 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	57.6	Xtriage
Anisotropy	0.388	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 52.8	EDS
L-test for twinning [2]	$<|L|> = 0.49$, $<L^2> = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	9110	wwPDB-VP
Average B, all atoms (Å^2)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.06% 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 (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	0.31	0/6701	0.57	0/9081
2	B	0.28	0/2605	0.50	0/3535
All	All	0.30	0/9306	0.55	0/12616

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

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	289	VAL	Peptide
1	A	743	ASP	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	6558	0	6177	605	1
2	B	2552	0	2359	148	1

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	9110	0	8536	740	1

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 42.

The worst 5 of 740 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:854:TYR:CE1	1:A:856:VAL:HG12	1.47	1.46
1:A:855:THR:CG2	1:A:861:THR:HB	1.43	1.45
1:A:52:ASP:HB3	1:A:53:ASN:CB	1.51	1.40
1:A:854:TYR:CD1	1:A:856:VAL:HG12	1.60	1.37
1:A:75:ASN:ND2	1:A:77:ARG:H	1.19	1.34

All (1) 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)
1:A:290:SER:OG	2:B:438:ASP:O[2_454]	2.02	0.18

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	847/904 (94%)	757 (89%)	72 (8%)	18 (2%)	7	30
2	B	337/343 (98%)	311 (92%)	26 (8%)	0	100	100
All	All	1184/1247 (95%)	1068 (90%)	98 (8%)	18 (2%)	10	39

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	52	ASP
1	A	55	VAL
1	A	56	THR
1	A	58	LEU
1	A	800	PRO

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	662/741 (89%)	530 (80%)	132 (20%)	1 5
2	B	271/292 (93%)	233 (86%)	38 (14%)	3 15
All	All	933/1033 (90%)	763 (82%)	170 (18%)	1 7

5 of 170 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	508	SER
1	A	625	SER
2	B	611	ASN
1	A	524	TYR
1	A	566	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	439	ASN
2	B	535	HIS
1	A	829	ASN
1	A	233	HIS
1	A	698	ASN

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 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 (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	853/904 (94%)	-0.21	8 (0%)	84	69	24, 73, 122, 160	0
2	B	339/343 (98%)	-0.34	0	100	100	31, 69, 112, 126	0
All	All	1192/1247 (95%)	-0.25	8 (0%)	87	75	24, 72, 118, 160	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	802	GLY	3.2
1	A	621	HIS	2.7
1	A	320	PHE	2.5
1	A	414	TYR	2.4
1	A	408	TYR	2.1

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 carbohydrates 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.