wwPDB X-ray Structure Validation Summary Report (i)

Mar 6, 2024 – 09:09 pm GMT

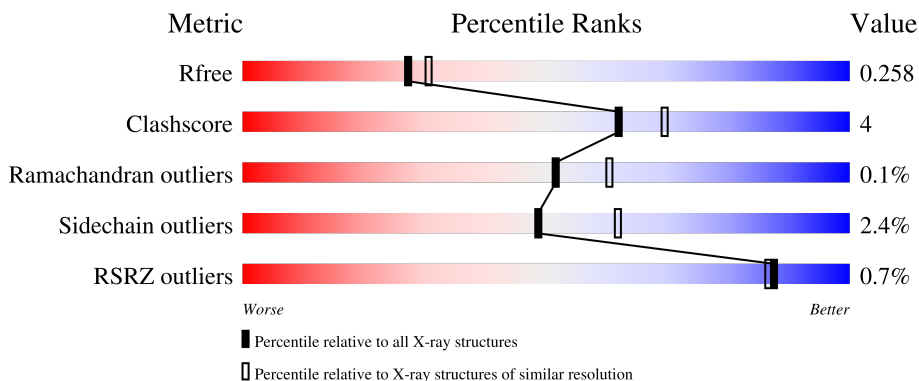|		
PDB ID	:	8Q2A
Title	:	TtX122B - A domain of unknown function from the Teredinibacter turnerae protein TERTU_2913
Authors	:	Rajagopal, B.S.; Hemsworth, G.R.
Deposited on	:	2023-08-01
Resolution	:	2.20 Å(reported)

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 2.20 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		0.258
Clashscore		4
Ramachandran outliers		0.1%
Sidechain outliers		2.4%
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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	245	86%　　　　　　　9%　　·
1	B	245	%　84%　　　　　　10%　· ·
1	C	245	%　81%　　　　　12%　· 6%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 10348 atoms, of which 4853 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 Putative lipoprotein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	234	Total	C	H	N	O	S	56	0	0
			3393	1114	1614	310	347	8			
1	B	234	Total	C	H	N	O	S	55	0	0
			3408	1117	1626	313	344	8			
1	C	231	Total	C	H	N	O	S	56	0	0
			3381	1107	1613	309	344	8			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	55	GLN	LYS	conflict	UNP C5BNC6
A	125	ILE	VAL	conflict	UNP C5BNC6
A	238	LEU	-	expression tag	UNP C5BNC6
A	239	GLU	-	expression tag	UNP C5BNC6
A	240	HIS	-	expression tag	UNP C5BNC6
A	241	HIS	-	expression tag	UNP C5BNC6
A	242	HIS	-	expression tag	UNP C5BNC6
A	243	HIS	-	expression tag	UNP C5BNC6
A	244	HIS	-	expression tag	UNP C5BNC6
A	245	HIS	-	expression tag	UNP C5BNC6
B	55	GLN	LYS	conflict	UNP C5BNC6
B	125	ILE	VAL	conflict	UNP C5BNC6
B	238	LEU	-	expression tag	UNP C5BNC6
B	239	GLU	-	expression tag	UNP C5BNC6
B	240	HIS	-	expression tag	UNP C5BNC6
B	241	HIS	-	expression tag	UNP C5BNC6
B	242	HIS	-	expression tag	UNP C5BNC6
B	243	HIS	-	expression tag	UNP C5BNC6
B	244	HIS	-	expression tag	UNP C5BNC6
B	245	HIS	-	expression tag	UNP C5BNC6
C	55	GLN	LYS	conflict	UNP C5BNC6
C	125	ILE	VAL	conflict	UNP C5BNC6
C	238	LEU	-	expression tag	UNP C5BNC6

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	239	GLU	-	expression tag	UNP C5BNC6
C	240	HIS	-	expression tag	UNP C5BNC6
C	241	HIS	-	expression tag	UNP C5BNC6
C	242	HIS	-	expression tag	UNP C5BNC6
C	243	HIS	-	expression tag	UNP C5BNC6
C	244	HIS	-	expression tag	UNP C5BNC6
C	245	HIS	-	expression tag	UNP C5BNC6

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Ca 1	0	0
2	B	1	Total 1	Ca 1	0	0
2	C	1	Total 1	Ca 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	59	Total 59	O 59	0	0
3	B	58	Total 58	O 58	0	0
3	C	46	Total 46	O 46	0	0

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.
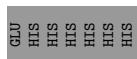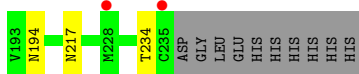
- Molecule 1: Putative lipoprotein

Chain A:
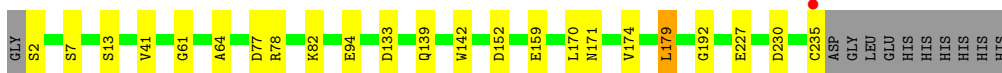
- Molecule 1: Putative lipoprotein

Chain B:

- Molecule 1: Putative lipoprotein

Chain C:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	69.14Å 75.32Å 116.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.69 – 2.20 46.69 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (46.69-2.20) 99.0 (46.69-2.20)	Depositor EDS
R_{merge}	0.36	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.72 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.8.0419	Depositor
R, R_{free}	0.208 , 0.259 0.214 , 0.258	Depositor DCC
R_{free} test set	1575 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	10.9	Xtriage
Anisotropy	0.800	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 26.5	EDS
L-test for twinning[2]	$< |L| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	10348	wwPDB-VP
Average B, all atoms (Å^2)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.13% 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)

Bond lengths and bond angles in the following residue types are not validated in this section: CA

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	0.88	5/1830 (0.3%)	1.06	4/2497 (0.2%)
1	B	0.85	5/1833 (0.3%)	1.07	5/2500 (0.2%)
1	C	0.77	1/1818 (0.1%)	1.08	2/2478 (0.1%)
All	All	0.84	11/5481 (0.2%)	1.07	11/7475 (0.1%)

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	159	GLU	CD-OE2	10.75	1.37	1.25
1	B	115	GLU	CD-OE1	8.68	1.35	1.25
1	A	227	GLU	CD-OE2	7.25	1.33	1.25
1	A	133	ASP	CG-OD1	7.11	1.41	1.25
1	B	100	GLU	CD-OE2	6.39	1.32	1.25

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	78	ARG	NE-CZ-NH2	-6.80	116.90	120.30
1	A	152	ASP	CB-CG-OD2	-6.31	112.62	118.30
1	B	143	TYR	OH-CZ-CE2	-6.05	103.75	120.10
1	A	78	ARG	NE-CZ-NH1	5.58	123.09	120.30
1	B	143	TYR	CB-CG-CD2	-5.49	117.71	121.00

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	1779	1614	1587	8	0
1	B	1782	1626	1599	17	0
1	C	1768	1613	1591	19	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	A	59	0	0	1	0
3	B	58	0	0	2	0
3	C	46	0	0	3	0
All	All	5495	4853	4777	44	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 4.

The worst 5 of 44 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:C:13:SER:HB2	1:C:45:VAL:HG21	1.77	0.65
1:C:165:LEU:HD23	1:C:167:TYR:O	1.97	0.65
1:C:135:ILE:HD11	1:C:142:TRP:HZ2	1.64	0.62
1:B:192:GLY:HA3	3:B:405:HOH:O	2.02	0.59
1:C:135:ILE:HD11	1:C:142:TRP:CZ2	2.38	0.58

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.

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	232/245 (95%)	218 (94%)	13 (6%)	1 (0%)	34	37
1	B	232/245 (95%)	217 (94%)	15 (6%)	0	100	100
1	C	227/245 (93%)	216 (95%)	11 (5%)	0	100	100
All	All	691/735 (94%)	651 (94%)	39 (6%)	1 (0%)	51	60

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	192	GLY

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	179/197 (91%)	175 (98%)	4 (2%)	52	65
1	B	179/197 (91%)	177 (99%)	2 (1%)	73	85
1	C	181/197 (92%)	174 (96%)	7 (4%)	32	41
All	All	539/591 (91%)	526 (98%)	13 (2%)	49	62

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	13	SER
1	C	17	GLN
1	C	165	LEU
1	C	74	GLN
1	C	103	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	216	ASN
1	B	217	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	74	GLN
1	C	194	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 ⓘ

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	234/245 (95%)	0.16	1 (0%)	92	91	5, 14, 30, 44	0
1	B	234/245 (95%)	0.12	2 (0%)	84	83	5, 14, 28, 47	0
1	C	231/245 (94%)	0.27	2 (0%)	84	83	5, 16, 32, 58	0
All	All	699/735 (95%)	0.18	5 (0%)	87	86	5, 15, 31, 58	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	235	CYS	3.1
1	B	164	ASP	3.0
1	B	3	GLY	2.3
1	C	228	MET	2.1
1	A	235	CYS	2.0

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 ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å²)	Q<0.9
2	CA	B	301	1/1	0.97	0.06	12,12,12,12	0
2	CA	C	301	1/1	0.97	0.07	12,12,12,12	0
2	CA	A	301	1/1	0.99	0.05	7,7,7,7	0

6.5 Other polymers ⓘ

There are no such residues in this entry.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK