



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 28, 2024 – 06:05 pm GMT

PDB ID : 9GCO
Title : Xenorhabdus bovienii Thioredoxin complex with Rhs toxin TreX and immunity protein TriX
Authors : Jurenas, D.; Terradot, L.
Deposited on : 2024-08-02
Resolution : 2.00 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 ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

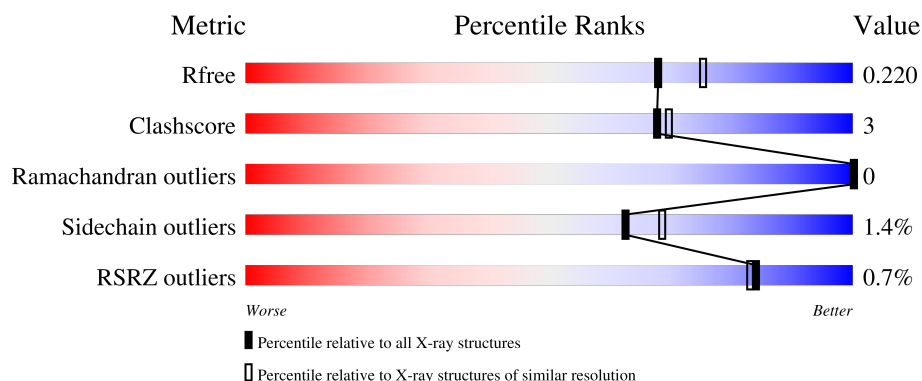
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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))
R_{free}	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (2.00-2.00)

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 $\leq 5\%$. 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.

Mol	Chain	Length	Quality of chain
1	A	115	<div> <div>%</div> <div> <div></div> <div>82%</div> <div>10%</div> <div>6%</div> </div> </div>
1	D	115	<div> <div>%</div> <div> <div></div> <div>82%</div> <div>12%</div> <div>6%</div> </div> </div>
2	B	140	<div> <div>%</div> <div> <div></div> <div>91%</div> <div>7%</div> </div> </div>
2	E	140	<div> <div></div> <div> <div>89%</div> <div>6%</div> </div> </div>
3	C	142	<div> <div>%</div> <div> <div></div> <div>85%</div> <div>8%</div> <div>6%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	142	 A horizontal bar chart showing the quality of chain F. The bar is 92% green and 6% yellow. The text '92%' is centered below the green portion, and '6%' is centered below the yellow portion.

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	PEG	E	201	-	-	X	-

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 11981 atoms, of which 5826 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 Thioredoxin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	108	Total	C	H	N	O	S	0	0	0
			1685	536	850	140	155	4			
1	D	108	Total	C	H	N	O	S	0	0	0
			1685	536	850	140	155	4			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP D3V6L5
A	2	ALA	-	expression tag	UNP D3V6L5
A	3	SER	-	expression tag	UNP D3V6L5
D	1	GLY	-	expression tag	UNP D3V6L5
D	2	ALA	-	expression tag	UNP D3V6L5
D	3	SER	-	expression tag	UNP D3V6L5

- Molecule 2 is a protein called Toxin TreX.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	130	Total	C	H	N	O	S	0	0	0
			1994	641	981	171	197	4			
2	E	131	Total	C	H	N	O	S	0	0	0
			2016	647	994	173	198	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MET	-	initiating methionine	UNP D3UXR3
E	1	MET	-	initiating methionine	UNP D3UXR3

- Molecule 3 is a protein called Immunity Protein TriX.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	134	Total	C	H	N	O	S	0	0	0
			2139	687	1048	185	211	8			
3	F	139	Total	C	H	N	O	S	0	0	0
			2224	717	1083	200	216	8			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	135	VAL	-	expression tag	UNP D3UXR2
C	136	ASP	-	expression tag	UNP D3UXR2
C	137	HIS	-	expression tag	UNP D3UXR2
C	138	HIS	-	expression tag	UNP D3UXR2
C	139	HIS	-	expression tag	UNP D3UXR2
C	140	HIS	-	expression tag	UNP D3UXR2
C	141	HIS	-	expression tag	UNP D3UXR2
C	142	HIS	-	expression tag	UNP D3UXR2
F	135	VAL	-	expression tag	UNP D3UXR2
F	136	ASP	-	expression tag	UNP D3UXR2
F	137	HIS	-	expression tag	UNP D3UXR2
F	138	HIS	-	expression tag	UNP D3UXR2
F	139	HIS	-	expression tag	UNP D3UXR2
F	140	HIS	-	expression tag	UNP D3UXR2
F	141	HIS	-	expression tag	UNP D3UXR2
F	142	HIS	-	expression tag	UNP D3UXR2

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C₄H₁₀O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	H	O	0	0
			17	4	10	3		
5	E	1	Total	C	H	O	0	0
			17	4	10	3		

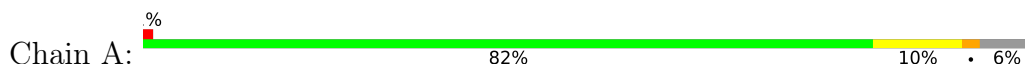
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	29	Total	O	0	0
			29	29		
6	B	24	Total	O	0	0
			24	24		
6	C	21	Total	O	0	0
			21	21		
6	D	28	Total	O	0	0
			28	28		
6	E	42	Total	O	0	0
			42	42		
6	F	53	Total	O	0	0
			53	53		

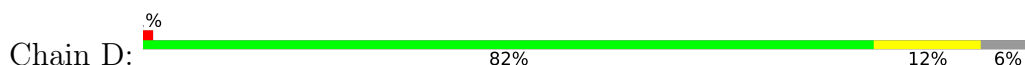
3 Residue-property plots [i](#)

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 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.

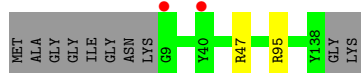
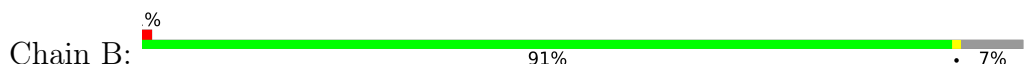
- Molecule 1: Thioredoxin



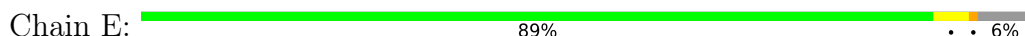
- Molecule 1: Thioredoxin



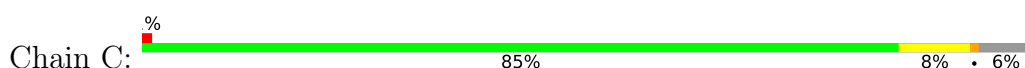
- Molecule 2: Toxin TreX



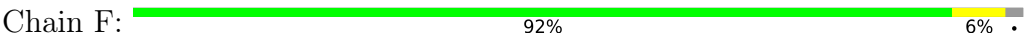
- Molecule 2: Toxin TreX



- Molecule 3: Immunity Protein TriX



- Molecule 3: Immunity Protein TriX



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	81.95Å 82.35Å 121.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.92 – 2.00 48.92 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.92-2.00) 99.8 (48.92-2.00)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.41 (at 2.00Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.188 , 0.223 0.188 , 0.220	Depositor DCC
R_{free} test set	54478 reflections (3.29%)	wwPDB-VP
Wilson B-factor (Å ²)	34.2	Xtriage
Anisotropy	0.500	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.41 , 40.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	11981	wwPDB-VP
Average B, all atoms (Å ²)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 62.65 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.0912e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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, PEG

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	0.57	0/851	0.65	0/1150
1	D	0.53	0/851	0.67	0/1150
2	B	0.52	0/1037	0.70	0/1403
2	E	0.63	0/1046	0.75	0/1414
3	C	0.50	0/1114	0.64	0/1503
3	F	0.54	0/1169	0.65	0/1578
All	All	0.55	0/6068	0.68	0/8198

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	1
1	D	0	1
All	All	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	81	ARG	Sidechain
1	D	81	ARG	Sidechain

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	835	850	850	15	0
1	D	835	850	850	8	0
2	B	1013	981	981	3	0
2	E	1022	994	994	6	0
3	C	1091	1048	1048	9	0
3	F	1141	1083	1083	4	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	2	0	0	0	0
4	D	1	0	0	0	0
4	F	2	0	0	0	0
5	B	7	10	10	0	0
5	E	7	10	10	4	0
6	A	29	0	0	0	0
6	B	24	0	0	0	0
6	C	21	0	0	0	0
6	D	28	0	0	0	0
6	E	42	0	0	1	0
6	F	53	0	0	1	0
All	All	6155	5826	5826	41	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 3.

The worst 5 of 41 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)
2:B:95:ARG:HD3	3:C:135:VAL:HG22	1.71	0.72
1:A:88:LEU:HD21	1:A:115:ILE:HD12	1.72	0.72
1:A:88:LEU:HD21	1:A:115:ILE:CD1	2.21	0.70
1:D:88:LEU:HD21	1:D:115:ILE:CD1	2.23	0.69
1:A:88:LEU:HD11	1:A:115:ILE:HD11	1.76	0.68

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	106/115 (92%)	104 (98%)	2 (2%)	0	100	100
1	D	106/115 (92%)	104 (98%)	2 (2%)	0	100	100
2	B	128/140 (91%)	124 (97%)	4 (3%)	0	100	100
2	E	129/140 (92%)	124 (96%)	5 (4%)	0	100	100
3	C	132/142 (93%)	130 (98%)	2 (2%)	0	100	100
3	F	137/142 (96%)	134 (98%)	3 (2%)	0	100	100
All	All	738/794 (93%)	720 (98%)	18 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	89/94 (95%)	88 (99%)	1 (1%)	70	76
1	D	89/94 (95%)	88 (99%)	1 (1%)	70	76
2	B	105/110 (96%)	105 (100%)	0	100	100
2	E	106/110 (96%)	103 (97%)	3 (3%)	38	40
3	C	123/131 (94%)	121 (98%)	2 (2%)	58	64
3	F	128/131 (98%)	126 (98%)	2 (2%)	58	64
All	All	640/670 (96%)	631 (99%)	9 (1%)	62	68

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	19	GLU
3	F	23	HIS
1	D	60	LYS
2	E	95	ARG
2	E	111	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	71	HIS

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 7 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
5	PEG	E	201	-	6,6,6	0.15	0	5,5,5	0.27	0
5	PEG	B	201	-	6,6,6	0.18	0	5,5,5	0.24	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	PEG	E	201	-	-	1/4/4/4	-
5	PEG	B	201	-	-	1/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	E	201	PEG	O2-C3-C4-O4
5	B	201	PEG	O2-C3-C4-O4

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	201	PEG	4	0

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	108/115 (93%)	0.02	1 (0%) 81 80	38, 48, 72, 89	0
1	D	108/115 (93%)	-0.10	1 (0%) 81 80	34, 44, 61, 91	0
2	B	130/140 (92%)	-0.11	2 (1%) 71 70	37, 45, 65, 104	0
2	E	131/140 (93%)	-0.22	0 100 100	32, 40, 57, 68	0
3	C	134/142 (94%)	0.06	1 (0%) 84 83	36, 53, 84, 95	0
3	F	139/142 (97%)	-0.13	0 100 100	32, 43, 69, 77	0
All	All	750/794 (94%)	-0.08	5 (0%) 84 83	32, 45, 71, 104	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	115	ILE	3.5
2	B	9	GLY	3.2
1	D	115	ILE	3.2
2	B	40	TYR	2.4
3	C	15	GLY	2.0

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 monosaccharides in this entry.

6.4 Ligands [i](#)

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(Å ²)	Q<0.9
5	PEG	E	201	7/7	0.91	0.09	40,50,60,61	0
5	PEG	B	201	7/7	0.92	0.08	41,50,61,66	0
4	CA	F	202	1/1	0.93	0.07	41,41,41,41	0
4	CA	F	201	1/1	0.96	0.05	45,45,45,45	0
4	CA	A	201	1/1	0.96	0.06	48,48,48,48	0
4	CA	C	201	1/1	0.96	0.05	57,57,57,57	0
4	CA	C	202	1/1	0.96	0.08	51,51,51,51	0
4	CA	D	201	1/1	0.97	0.04	43,43,43,43	0
4	CA	B	202	1/1	0.98	0.05	39,39,39,39	1

6.5 Other polymers [i](#)

There are no such residues in this entry.