



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 01:46 am BST

PDB ID : 2XRO
Title : Crystal structure of TtgV in complex with its DNA operator
Authors : Lu, D.; Fillet, S.; Meng, C.; Alguel, Y.; Kloppsteck, P.; Bergeron, J.; Krell, T.; Gallegos, M.-T.; Ramos, J.; Zhang, X.
Deposited on : 2010-09-17
Resolution : 3.40 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 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

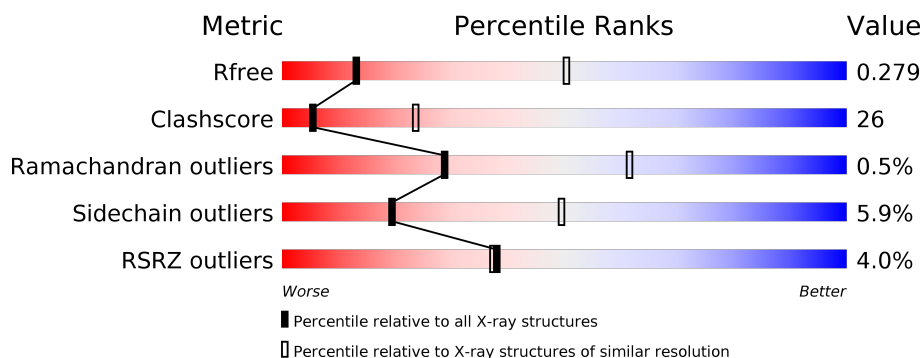
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)

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 $\leq 5\%$. 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	241	<div> <div>3%</div> <div> <div></div> <div>55%</div> <div>42%</div> <div>..</div> </div> </div>
1	B	241	<div> <div>5%</div> <div> <div></div> <div>61%</div> <div>33%</div> <div>..</div> </div> </div>
1	E	241	<div> <div>3%</div> <div> <div></div> <div>54%</div> <div>43%</div> <div>..</div> </div> </div>
1	F	241	<div> <div>5%</div> <div> <div></div> <div>59%</div> <div>36%</div> <div>5%</div> </div> </div>
2	X	42	<div> <div>7%</div> <div> <div></div> <div>38%</div> <div>55%</div> </div> </div>
3	Y	43	<div> <div>•</div> <div> <div></div> <div>23%</div> <div>74%</div> </div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8840 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 HTH-TYPE TRANSCRIPTIONAL REGULATOR TTGV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	239	Total	C	N	O	S	0	0	0
			1768	1123	309	334	2			
1	B	240	Total	C	N	O	S	0	0	0
			1780	1130	313	335	2			
1	E	239	Total	C	N	O	S	0	0	0
			1774	1126	312	334	2			
1	F	240	Total	C	N	O	S	0	0	0
			1780	1130	313	335	2			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	13	SER	-	expression tag	UNP Q93PU6
A	109	SER	CYS	engineered mutation	UNP Q93PU6
A	205	SER	CYS	engineered mutation	UNP Q93PU6
B	13	SER	-	expression tag	UNP Q93PU6
B	109	SER	CYS	engineered mutation	UNP Q93PU6
B	205	SER	CYS	engineered mutation	UNP Q93PU6
E	13	SER	-	expression tag	UNP Q93PU6
E	109	SER	CYS	engineered mutation	UNP Q93PU6
E	205	SER	CYS	engineered mutation	UNP Q93PU6
F	13	SER	-	expression tag	UNP Q93PU6
F	109	SER	CYS	engineered mutation	UNP Q93PU6
F	205	SER	CYS	engineered mutation	UNP Q93PU6

- Molecule 2 is a DNA chain called TTGV OPERATOR DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	X	42	Total	C	N	O	P	0	0	0
			850	408	153	248	41			

- Molecule 3 is a DNA chain called TTGV OPERATOR DNA.

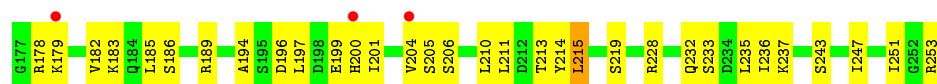
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	Y	43	Total	C	N	O	P	0	0	0
			886	424	161	259	42			

- Molecule 4 is OSMIUM ION (three-letter code: OS) (formula: Os).

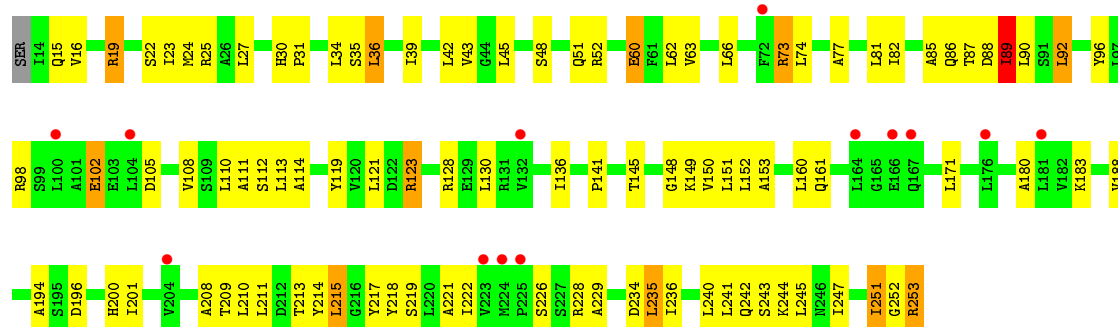
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Os	0	0
			1	1		
4	A	1	Total	Os	0	0
			1	1		

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 ($\text{RSRZ} > 2$). 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.

SR	SR
S91	SER
L92	ILE
V93	Q15
K94	V16
P95	
Y96	
	R19
	A20
	A21
	S22
	I23
	Y24
	R25
	A26
	L27
	L34
	S35
	K36
	I39
	A40
	Q41
	L42
	V43
	G44
	L45
	P46
	R47
	S48
	T49
	V50
	Q51
	G52
	I53
	E60
	E64
	A65
	L66
	G71
	F72
	R73
	L74
	G75
	P76
	G79
	Q80
	L81
	I82
	N83
	Q84
	A85
	Q86
	T87
	D88
	I89
	L90
	T97
	L98
	L99
	L100
	L101
	L102
	L103
	L104
	L105
	L106
	L107
	L108
	L109
	L110
	L111
	L112
	L113
	D116
	K117
	I118
	R123
	I124
	V125
	S126
	E127
	R128
	R131
	F134
	P135
	I136
	G137
	L138
	N139
	V140
	P141
	L146
	A147
	G148
	K149
	V150
	L151
	L152
	L155
	P156
	D157
	E158
	T159
	L160
	L168
	P169
	V170
	L171
	T175
	L176



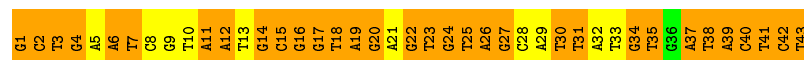
• Molecule 1: HTH-TYPE TRANSCRIPTIONAL REGULATOR TTGV



• Molecule 2: TTGV OPERATOR DNA



• Molecule 3: TTGV OPERATOR DNA



4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α , β , γ	89.58Å 89.58Å 416.75Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	62.23 – 3.40 77.58 – 3.40	Depositor EDS
% Data completeness (in resolution range)	92.6 (62.23-3.40) 96.5 (77.58-3.40)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.71 (at 3.41Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.214 , 0.282 0.207 , 0.279	Depositor DCC
R_{free} test set	1203 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å ²)	116.6	Xtriage
Anisotropy	0.371	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 171.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.447 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	8840	wwPDB-VP
Average B, all atoms (Å ²)	152.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.20% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical 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 ⓘ

5.1 Standard geometry ⓘ

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

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	0.26	0/1789	0.47	0/2429
1	B	0.25	0/1801	0.51	0/2444
1	E	0.25	0/1795	0.48	0/2436
1	F	0.25	0/1801	0.51	0/2444
2	X	1.27	2/952 (0.2%)	2.24	62/1465 (4.2%)
3	Y	1.25	2/994 (0.2%)	2.29	77/1535 (5.0%)
All	All	0.62	4/9132 (0.0%)	1.18	139/12753 (1.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	X	25	DA	C3'-O3'	6.37	1.52	1.44
3	Y	15	DC	N1-C6	-5.67	1.33	1.37
3	Y	15	DC	C3'-O3'	-5.55	1.36	1.44
2	X	14	DT	C1'-N1	-5.02	1.40	1.47

The worst 5 of 139 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	X	21	DC	C4'-C3'-C2'	-10.42	93.72	103.10
3	Y	10	DT	N3-C4-O4	10.12	125.97	119.90
3	Y	18	DT	O4'-C1'-N1	9.81	114.87	108.00
3	Y	35	DT	C1'-O4'-C4'	-9.79	100.31	110.10
2	X	24	DT	N3-C4-O4	9.57	125.64	119.90

There are no chirality outliers.

There are no planarity outliers.

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	1768	0	1852	103	0
1	B	1780	0	1867	80	0
1	E	1774	0	1863	106	0
1	F	1780	0	1867	83	0
2	X	850	0	475	64	0
3	Y	886	0	489	65	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
All	All	8840	0	8413	449	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 26.

The worst 5 of 449 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:F:89:ILE:H	1:F:89:ILE:HD13	1.28	0.97
1:B:89:ILE:H	1:B:89:ILE:HD13	1.31	0.94
2:X:1:DG:H5''	2:X:1:DG:H8	1.32	0.92
1:A:135:PRO:HD2	1:A:138:ILE:HD12	1.56	0.87
1:E:135:PRO:HD2	1:E:138:ILE:HD12	1.56	0.87

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	237/241 (98%)	196 (83%)	40 (17%)	1 (0%)	34	67
1	B	238/241 (99%)	206 (87%)	30 (13%)	2 (1%)	19	51
1	E	237/241 (98%)	197 (83%)	39 (16%)	1 (0%)	34	67
1	F	238/241 (99%)	207 (87%)	30 (13%)	1 (0%)	34	67
All	All	950/964 (98%)	806 (85%)	139 (15%)	5 (0%)	29	61

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	89	ILE
1	F	89	ILE
1	B	251	ILE
1	A	138	ILE
1	E	138	ILE

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	187/193 (97%)	181 (97%)	6 (3%)	39	67
1	B	188/193 (97%)	172 (92%)	16 (8%)	10	35
1	E	188/193 (97%)	182 (97%)	6 (3%)	39	67
1	F	188/193 (97%)	172 (92%)	16 (8%)	10	35
All	All	751/772 (97%)	707 (94%)	44 (6%)	19	49

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	235	LEU
1	E	86	GLN
1	F	234	ASP
1	B	251	ILE
1	E	51	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	86	GLN
1	E	80	GLN
1	F	184	GLN
1	B	184	GLN
1	B	232	GLN

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](#)

Of 2 ligands modelled in this entry, 2 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 [i](#)

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 [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	239/241 (99%)	0.19	8 (3%) 46 45	63, 154, 257, 392	0
1	B	240/241 (99%)	0.18	13 (5%) 25 26	77, 138, 233, 292	0
1	E	239/241 (99%)	0.25	8 (3%) 46 45	68, 151, 245, 403	0
1	F	240/241 (99%)	0.18	13 (5%) 25 26	79, 136, 228, 303	0
2	X	42/42 (100%)	-0.38	0 100 100	102, 162, 183, 196	0
3	Y	43/43 (100%)	-0.38	0 100 100	108, 153, 190, 227	0
All	All	1043/1049 (99%)	0.15	42 (4%) 38 37	63, 146, 241, 403	0

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	176	LEU	5.0
1	B	224	MET	3.9
1	F	224	MET	3.9
1	A	110	LEU	3.6
1	F	176	LEU	3.6

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](#)

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(\AA^2)	Q<0.9
4	OS	A	1001	1/1	0.89	0.55	370,370,370,370	0
4	OS	B	1001	1/1	0.93	0.57	373,373,373,373	0

6.5 Other polymers [i](#)

There are no such residues in this entry.