wwPDB X-ray Structure Validation Summary Report (i)

Dec 10, 2022 – 11:25 PM EST

PDB ID	:	1KAF
Title	:	DNA Binding Domain Of The Phage T4 Transcription Factor MotA (AA105-211)
Authors	:	Li, N.; Sickmier, E.A.; Zhang, R.; Joachimiak, A.; White, S.W.
Deposited on	:	2001-11-01
Resolution	:	1.60 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 (i) 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 (i)) were used in the production of this report:

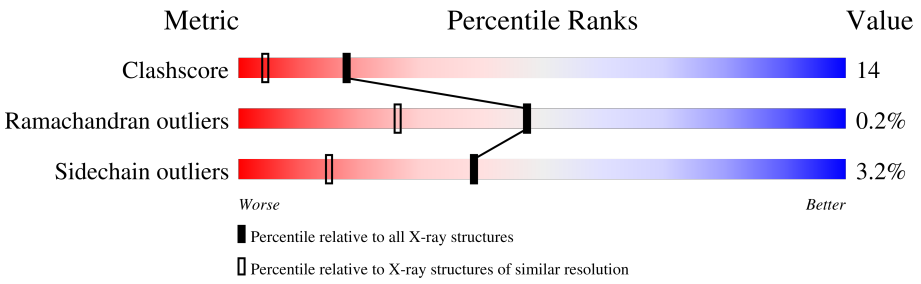
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 1.60 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))
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	108	86% 12% •
1	B	108	73% 25% •
1	C	108	78% 20% •
1	D	108	81% 13% • •
1	E	108	69% 23% • • 5%
1	F	108	66% 27% • •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5785 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 Transcription regulatory protein MOTA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	108	874	550	150	166	8	0	0	0
1	B	108	874	550	150	166	8	0	0	0
1	C	108	874	550	150	166	8	0	0	0
1	D	104	842	530	146	159	7	0	0	0
1	E	103	836	527	145	157	7	0	0	0
1	F	104	842	530	146	159	7	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	104	MET	-	initiating methionine	UNP P22915
B	104	MET	-	initiating methionine	UNP P22915
C	104	MET	-	initiating methionine	UNP P22915
D	104	MET	-	initiating methionine	UNP P22915
E	104	MET	-	initiating methionine	UNP P22915
F	104	MET	-	initiating methionine	UNP P22915

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	100	100	100	0	0
2	B	119	119	119	0	0
2	C	142	142	142	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	D	105	105	105	0	0
			Total	O		
2	E	93	93	93	0	0
			Total	O		
2	F	84	84	84	0	0

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. 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. 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.

Note EDS was not executed.

• Molecule 1: Transcription regulatory protein MOTA

Chain A:

• Molecule 1: Transcription regulatory protein MOTA

Chain B:

• Molecule 1: Transcription regulatory protein MOTA

Chain C:

• Molecule 1: Transcription regulatory protein MOTA

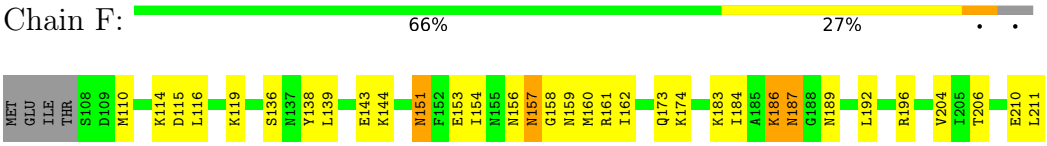
Chain D:

• Molecule 1: Transcription regulatory protein MOTA

Chain E:

• Molecule 1: Transcription regulatory protein MOTA

Chain F:


66%
27%
MET
GLU
ILE
THR
S108
D109
M110
K114
D115
L116
K119
S136
N137
Y138
L139
E143
K144
N151
F152
E153
I154
N155
N156
N157
G158
N159
M160
R161
I162
Q173
K174
K183
I184
A185
K186
N187
G188
N189
L192
R196
V204
I205
T206
E210
L211

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	31.28Å 54.73Å 91.44Å 94.77° 91.10° 95.61°	Depositor
Resolution (Å)	30.00 – 1.60	Depositor
% Data completeness (in resolution range)	(Not available) (30.00-1.60)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.226 , 0.258	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5785	wwPDB-VP
Average B, all atoms ($Å^2$)	12.0	wwPDB-VP

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	0.33	0/883	0.63	2/1177 (0.2%)
1	B	0.32	0/883	0.60	0/1177
1	C	0.31	0/883	0.64	0/1177
1	D	0.31	0/851	0.60	0/1134
1	E	0.31	0/845	0.62	0/1126
1	F	0.30	0/851	0.59	0/1134
All	All	0.31	0/5196	0.61	2/6925 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	160	MET	N-CA-C	-5.04	97.40	111.00
1	A	133	ILE	N-CA-C	-5.03	97.42	111.00

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	874	0	891	13	0
1	B	874	0	891	28	0
1	C	874	0	891	27	0
1	D	842	0	858	16	0
1	E	836	0	853	30	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	842	0	858	31	0
2	A	100	0	0	1	0
2	B	119	0	0	2	0
2	C	142	0	0	1	0
2	D	105	0	0	1	0
2	E	93	0	0	2	0
2	F	84	0	0	1	0
All	All	5785	0	5242	141	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 14.

The worst 5 of 141 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:D:186:LYS:HE3	1:D:186:LYS:H	1.27	0.96
1:E:195:LYS:HD3	1:E:195:LYS:H	1.29	0.95
1:C:186:LYS:HE3	1:C:186:LYS:H	1.28	0.94
1:D:186:LYS:H	1:D:186:LYS:CE	1.85	0.89
1:C:104:MET:HE2	1:C:133:ILE:HG22	1.56	0.86

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	106/108 (98%)	104 (98%)	2 (2%)	0	100	100
1	B	106/108 (98%)	105 (99%)	1 (1%)	0	100	100
1	C	106/108 (98%)	104 (98%)	2 (2%)	0	100	100
1	D	102/108 (94%)	101 (99%)	1 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	E	101/108 (94%)	98 (97%)	2 (2%)	1 (1%)	15 3
1	F	102/108 (94%)	101 (99%)	1 (1%)	0	100 100
All	All	623/648 (96%)	613 (98%)	9 (1%)	1 (0%)	47 26

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	112	GLU

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	97/97 (100%)	95 (98%)	2 (2%)	53 29
1	B	97/97 (100%)	95 (98%)	2 (2%)	53 29
1	C	97/97 (100%)	95 (98%)	2 (2%)	53 29
1	D	93/97 (96%)	89 (96%)	4 (4%)	29 9
1	E	92/97 (95%)	89 (97%)	3 (3%)	38 14
1	F	93/97 (96%)	88 (95%)	5 (5%)	22 5
All	All	569/582 (98%)	551 (97%)	18 (3%)	39 15

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	157	ASN
1	F	187	ASN
1	F	186	LYS
1	D	186	LYS
1	F	151	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	189	ASN
1	F	137	ASN
1	E	189	ASN
1	F	151	ASN
1	B	189	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 monosaccharides 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)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.