wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2024 – 05:41 AM EDT

PDB ID	:	1TYH
Title	:	Crystal Structure of Transcriptional Activator tenA from Bacillus subtilis
Authors	:	Eswaramoorthy, S.; Swaminathan, S.; Burley, S.K.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on	:	2004-07-07
Resolution	:	2.54 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:

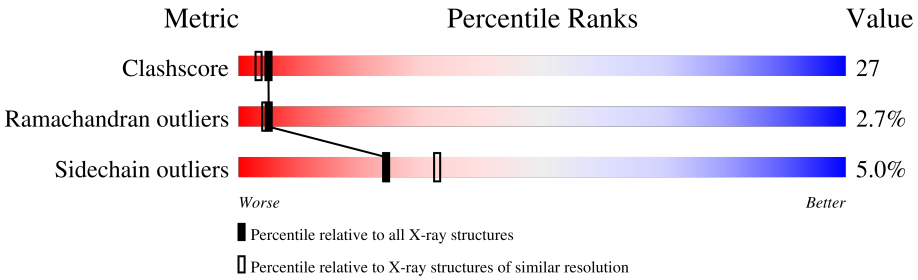
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.54 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	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)

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	248	63% 24% • 11%
1	B	248	62% 24% • 11%
1	D	248	35% 48% 6% 11%
1	E	248	40% 44% 6% 11%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 7252 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 Transcriptional activator tenA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	221	Total	C	N	O	S	Se	0	0	0
			1783	1154	287	334	3	5			
1	B	221	Total	C	N	O	S	Se	0	0	0
			1783	1154	287	334	3	5			
1	D	221	Total	C	N	O	S	Se	0	0	0
			1791	1158	287	338	3	5			
1	E	221	Total	C	N	O	S	Se	0	0	0
			1783	1154	287	334	3	5			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MSE	-	cloning artifact	UNP P25052
A	0	SER	-	cloning artifact	UNP P25052
A	1	LEU	-	cloning artifact	UNP P25052
A	71	MSE	MET	modified residue	UNP P25052
A	83	MSE	MET	modified residue	UNP P25052
A	116	MSE	MET	modified residue	UNP P25052
A	194	MSE	MET	modified residue	UNP P25052
A	211	MSE	MET	modified residue	UNP P25052
A	237	GLU	-	expression tag	UNP P25052
A	238	GLY	-	expression tag	UNP P25052
A	239	GLY	-	expression tag	UNP P25052
A	240	SER	-	expression tag	UNP P25052
A	241	HIS	-	expression tag	UNP P25052
A	242	HIS	-	expression tag	UNP P25052
A	243	HIS	-	expression tag	UNP P25052
A	244	HIS	-	expression tag	UNP P25052
A	245	HIS	-	expression tag	UNP P25052
A	246	HIS	-	expression tag	UNP P25052
B	-1	MSE	-	cloning artifact	UNP P25052
B	0	SER	-	cloning artifact	UNP P25052
B	1	LEU	-	cloning artifact	UNP P25052

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	71	MSE	MET	modified residue	UNP P25052
B	83	MSE	MET	modified residue	UNP P25052
B	116	MSE	MET	modified residue	UNP P25052
B	194	MSE	MET	modified residue	UNP P25052
B	211	MSE	MET	modified residue	UNP P25052
B	237	GLU	-	expression tag	UNP P25052
B	238	GLY	-	expression tag	UNP P25052
B	239	GLY	-	expression tag	UNP P25052
B	240	SER	-	expression tag	UNP P25052
B	241	HIS	-	expression tag	UNP P25052
B	242	HIS	-	expression tag	UNP P25052
B	243	HIS	-	expression tag	UNP P25052
B	244	HIS	-	expression tag	UNP P25052
B	245	HIS	-	expression tag	UNP P25052
B	246	HIS	-	expression tag	UNP P25052
D	-1	MSE	-	cloning artifact	UNP P25052
D	0	SER	-	cloning artifact	UNP P25052
D	1	LEU	-	cloning artifact	UNP P25052
D	71	MSE	MET	modified residue	UNP P25052
D	83	MSE	MET	modified residue	UNP P25052
D	116	MSE	MET	modified residue	UNP P25052
D	194	MSE	MET	modified residue	UNP P25052
D	211	MSE	MET	modified residue	UNP P25052
D	237	GLU	-	expression tag	UNP P25052
D	238	GLY	-	expression tag	UNP P25052
D	239	GLY	-	expression tag	UNP P25052
D	240	SER	-	expression tag	UNP P25052
D	241	HIS	-	expression tag	UNP P25052
D	242	HIS	-	expression tag	UNP P25052
D	243	HIS	-	expression tag	UNP P25052
D	244	HIS	-	expression tag	UNP P25052
D	245	HIS	-	expression tag	UNP P25052
D	246	HIS	-	expression tag	UNP P25052
E	-1	MSE	-	cloning artifact	UNP P25052
E	0	SER	-	cloning artifact	UNP P25052
E	1	LEU	-	cloning artifact	UNP P25052
E	71	MSE	MET	modified residue	UNP P25052
E	83	MSE	MET	modified residue	UNP P25052
E	116	MSE	MET	modified residue	UNP P25052
E	194	MSE	MET	modified residue	UNP P25052
E	211	MSE	MET	modified residue	UNP P25052
E	237	GLU	-	expression tag	UNP P25052

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	238	GLY	-	expression tag	UNP P25052
E	239	GLY	-	expression tag	UNP P25052
E	240	SER	-	expression tag	UNP P25052
E	241	HIS	-	expression tag	UNP P25052
E	242	HIS	-	expression tag	UNP P25052
E	243	HIS	-	expression tag	UNP P25052
E	244	HIS	-	expression tag	UNP P25052
E	245	HIS	-	expression tag	UNP P25052
E	246	HIS	-	expression tag	UNP P25052

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	36	Total O 36 36	0	0
2	B	34	Total O 34 34	0	0
2	D	17	Total O 17 17	0	0
2	E	25	Total O 25 25	0	0

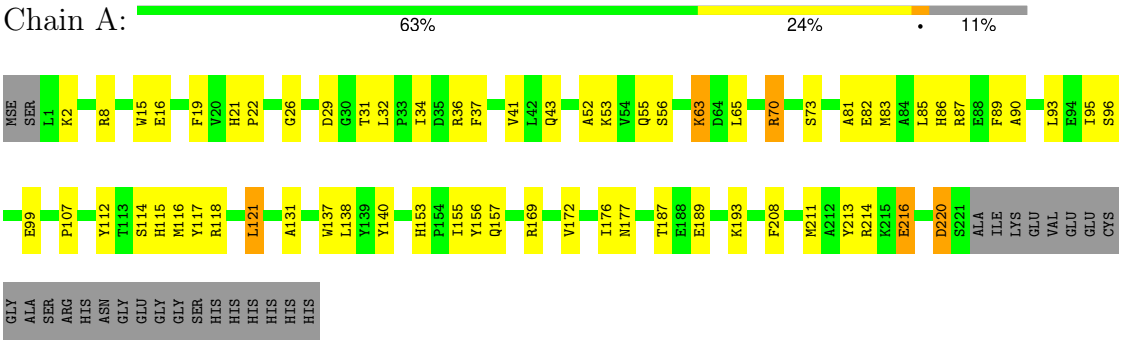
3 Residue-property plots (i)

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.

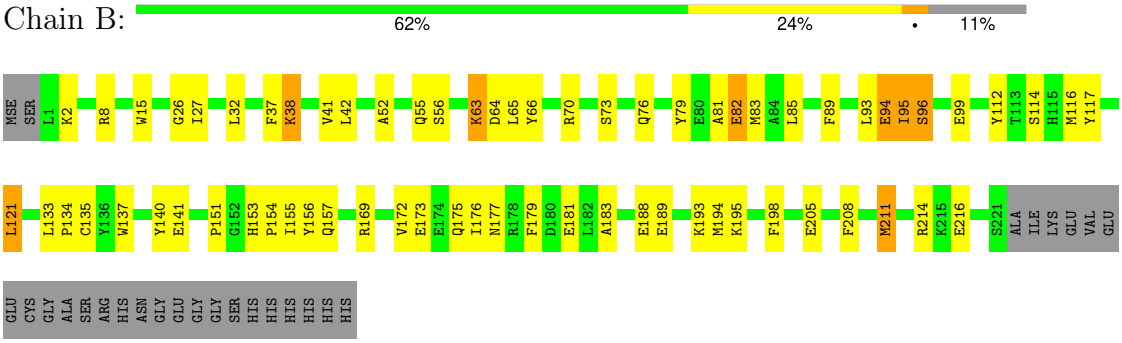
• Molecule 1: Transcriptional activator tenA

Chain A:



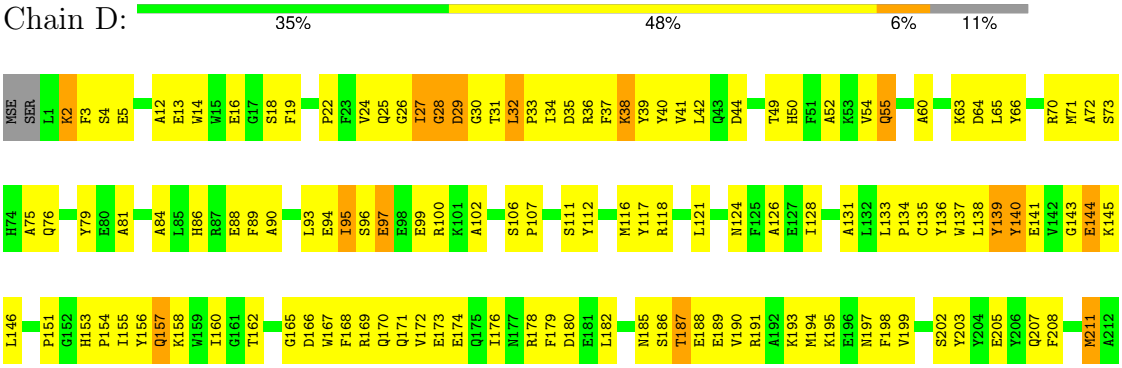
• Molecule 1: Transcriptional activator tenA

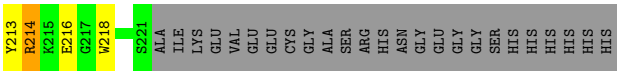
Chain B:



• Molecule 1: Transcriptional activator tenA

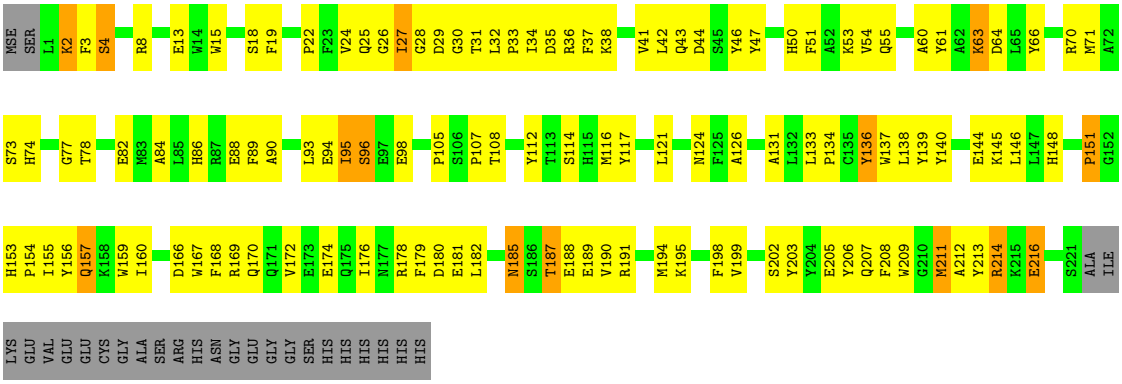
Chain D:





• Molecule 1: Transcriptional activator tenA

Chain E:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α, β, γ	58.43Å 58.43Å 297.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.54	Depositor
% Data completeness (in resolution range)	96.9 (50.00-2.54)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.244 , 0.304	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7252	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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	0.44	0/1836	0.60	1/2483 (0.0%)
1	B	0.44	0/1836	0.60	1/2483 (0.0%)
1	D	0.46	0/1844	0.62	0/2493
1	E	0.45	0/1836	0.62	0/2483
All	All	0.45	0/7352	0.61	2/9942 (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	B	214	ARG	N-CA-C	-5.92	95.01	111.00
1	A	214	ARG	N-CA-C	-5.90	95.06	111.00

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	1783	0	1625	50	0
1	B	1783	0	1625	56	0
1	D	1791	0	1633	144	0
1	E	1783	0	1625	136	0
2	A	36	0	0	8	0
2	B	34	0	0	7	0
2	D	17	0	0	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	E	25	0	0	7	0
All	All	7252	0	6508	374	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 27.

The worst 5 of 374 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:99:GLU:HG3	2:A:268:HOH:O	1.46	1.11
1:A:220:ASP:HB2	2:A:272:HOH:O	1.71	0.90
1:E:151:PRO:HG2	1:E:157:GLN:OE1	1.70	0.90
1:D:155:ILE:HG23	1:D:156:TYR:CD1	2.09	0.88
1:A:86:HIS:CB	2:A:271:HOH:O	2.22	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	219/248 (88%)	208 (95%)	9 (4%)	2 (1%)	17 24
1	B	219/248 (88%)	206 (94%)	8 (4%)	5 (2%)	6 6
1	D	219/248 (88%)	180 (82%)	30 (14%)	9 (4%)	3 1
1	E	219/248 (88%)	181 (83%)	30 (14%)	8 (4%)	3 2
All	All	876/992 (88%)	775 (88%)	77 (9%)	24 (3%)	5 4

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	30	GLY

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	95	ILE
1	D	187	THR
1	E	95	ILE
1	E	96	SER

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	175/201 (87%)	170 (97%)	5 (3%)	42 57
1	B	175/201 (87%)	169 (97%)	6 (3%)	37 50
1	D	177/201 (88%)	161 (91%)	16 (9%)	9 11
1	E	175/201 (87%)	167 (95%)	8 (5%)	27 36
All	All	702/804 (87%)	667 (95%)	35 (5%)	24 33

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	27	ILE
1	E	63	LYS
1	E	185	ASN
1	D	27	ILE
1	D	2	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	148	HIS
1	D	171	GLN
1	E	207	GLN
1	E	55	GLN
1	E	148	HIS

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

There are no ligands in this entry.

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

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.