



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 02:57 PM UTC

PDB ID : 8ZWN / pdb_00008zwn
Title : Crystal structure of DNA-binding transcriptional activator EvgA from Escherichia coli str. K-12 substr. MG1655
Authors : Teng, Y.; Liu, R.; Luo, B.
Deposited on : 2024-06-13
Resolution : 2.90 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

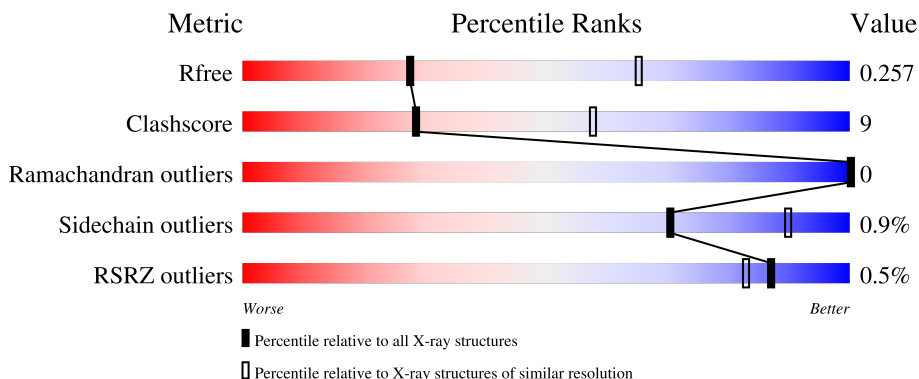
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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}	180053	2481 (2.90-2.90)
Clashscore	190562	2690 (2.90-2.90)
Ramachandran outliers	187476	2623 (2.90-2.90)
Sidechain outliers	187428	2625 (2.90-2.90)
RSRZ outliers	180081	2481 (2.90-2.90)

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	224	2% 75% 17% 8%
1	B	224	74% 17% 8%
1	C	224	74% 17% 9%
1	D	224	70% 21% 9%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 6294 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 DNA-binding transcriptional activator EvgA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	205	Total	C	N	O	S	0	0	0
			1566	989	265	303	9			
1	B	205	Total	C	N	O	S	0	0	0
			1587	1003	271	304	9			
1	C	204	Total	C	N	O	S	0	0	0
			1575	995	269	303	8			
1	D	204	Total	C	N	O	S	0	0	0
			1566	990	267	300	9			

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP P0ACZ4
A	-18	GLY	-	expression tag	UNP P0ACZ4
A	-17	SER	-	expression tag	UNP P0ACZ4
A	-16	SER	-	expression tag	UNP P0ACZ4
A	-15	HIS	-	expression tag	UNP P0ACZ4
A	-14	HIS	-	expression tag	UNP P0ACZ4
A	-13	HIS	-	expression tag	UNP P0ACZ4
A	-12	HIS	-	expression tag	UNP P0ACZ4
A	-11	HIS	-	expression tag	UNP P0ACZ4
A	-10	HIS	-	expression tag	UNP P0ACZ4
A	-9	SER	-	expression tag	UNP P0ACZ4
A	-8	SER	-	expression tag	UNP P0ACZ4
A	-7	GLY	-	expression tag	UNP P0ACZ4
A	-6	GLU	-	expression tag	UNP P0ACZ4
A	-5	ASN	-	expression tag	UNP P0ACZ4
A	-4	LEU	-	expression tag	UNP P0ACZ4
A	-3	TYR	-	expression tag	UNP P0ACZ4
A	-2	PHE	-	expression tag	UNP P0ACZ4
A	-1	GLN	-	expression tag	UNP P0ACZ4
A	0	GLY	-	expression tag	UNP P0ACZ4
B	-19	MET	-	initiating methionine	UNP P0ACZ4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-18	GLY	-	expression tag	UNP P0ACZ4
B	-17	SER	-	expression tag	UNP P0ACZ4
B	-16	SER	-	expression tag	UNP P0ACZ4
B	-15	HIS	-	expression tag	UNP P0ACZ4
B	-14	HIS	-	expression tag	UNP P0ACZ4
B	-13	HIS	-	expression tag	UNP P0ACZ4
B	-12	HIS	-	expression tag	UNP P0ACZ4
B	-11	HIS	-	expression tag	UNP P0ACZ4
B	-10	HIS	-	expression tag	UNP P0ACZ4
B	-9	SER	-	expression tag	UNP P0ACZ4
B	-8	SER	-	expression tag	UNP P0ACZ4
B	-7	GLY	-	expression tag	UNP P0ACZ4
B	-6	GLU	-	expression tag	UNP P0ACZ4
B	-5	ASN	-	expression tag	UNP P0ACZ4
B	-4	LEU	-	expression tag	UNP P0ACZ4
B	-3	TYR	-	expression tag	UNP P0ACZ4
B	-2	PHE	-	expression tag	UNP P0ACZ4
B	-1	GLN	-	expression tag	UNP P0ACZ4
B	0	GLY	-	expression tag	UNP P0ACZ4
C	-19	MET	-	initiating methionine	UNP P0ACZ4
C	-18	GLY	-	expression tag	UNP P0ACZ4
C	-17	SER	-	expression tag	UNP P0ACZ4
C	-16	SER	-	expression tag	UNP P0ACZ4
C	-15	HIS	-	expression tag	UNP P0ACZ4
C	-14	HIS	-	expression tag	UNP P0ACZ4
C	-13	HIS	-	expression tag	UNP P0ACZ4
C	-12	HIS	-	expression tag	UNP P0ACZ4
C	-11	HIS	-	expression tag	UNP P0ACZ4
C	-10	HIS	-	expression tag	UNP P0ACZ4
C	-9	SER	-	expression tag	UNP P0ACZ4
C	-8	SER	-	expression tag	UNP P0ACZ4
C	-7	GLY	-	expression tag	UNP P0ACZ4
C	-6	GLU	-	expression tag	UNP P0ACZ4
C	-5	ASN	-	expression tag	UNP P0ACZ4
C	-4	LEU	-	expression tag	UNP P0ACZ4
C	-3	TYR	-	expression tag	UNP P0ACZ4
C	-2	PHE	-	expression tag	UNP P0ACZ4
C	-1	GLN	-	expression tag	UNP P0ACZ4
C	0	GLY	-	expression tag	UNP P0ACZ4
D	-19	MET	-	initiating methionine	UNP P0ACZ4
D	-18	GLY	-	expression tag	UNP P0ACZ4
D	-17	SER	-	expression tag	UNP P0ACZ4

Continued on next page...

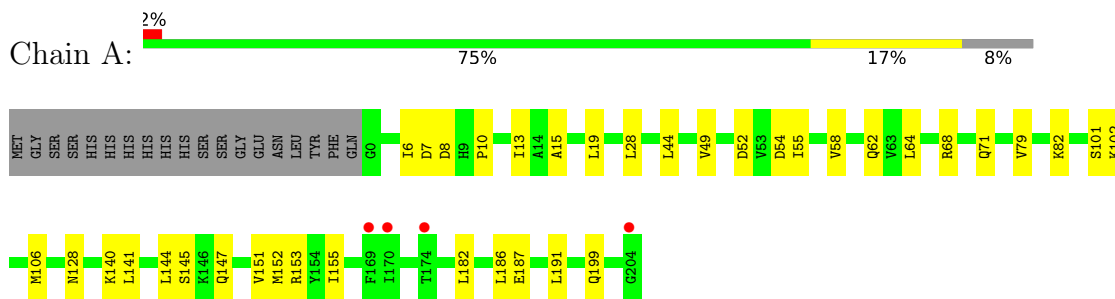
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-16	SER	-	expression tag	UNP P0ACZ4
D	-15	HIS	-	expression tag	UNP P0ACZ4
D	-14	HIS	-	expression tag	UNP P0ACZ4
D	-13	HIS	-	expression tag	UNP P0ACZ4
D	-12	HIS	-	expression tag	UNP P0ACZ4
D	-11	HIS	-	expression tag	UNP P0ACZ4
D	-10	HIS	-	expression tag	UNP P0ACZ4
D	-9	SER	-	expression tag	UNP P0ACZ4
D	-8	SER	-	expression tag	UNP P0ACZ4
D	-7	GLY	-	expression tag	UNP P0ACZ4
D	-6	GLU	-	expression tag	UNP P0ACZ4
D	-5	ASN	-	expression tag	UNP P0ACZ4
D	-4	LEU	-	expression tag	UNP P0ACZ4
D	-3	TYR	-	expression tag	UNP P0ACZ4
D	-2	PHE	-	expression tag	UNP P0ACZ4
D	-1	GLN	-	expression tag	UNP P0ACZ4
D	0	GLY	-	expression tag	UNP P0ACZ4

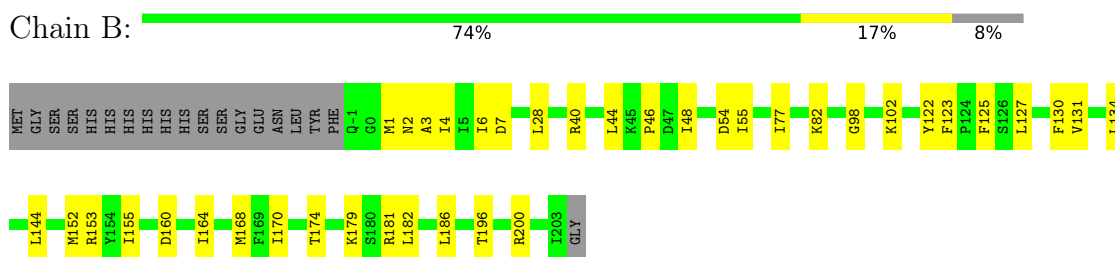
3 Residue-property plots [i](#)

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.

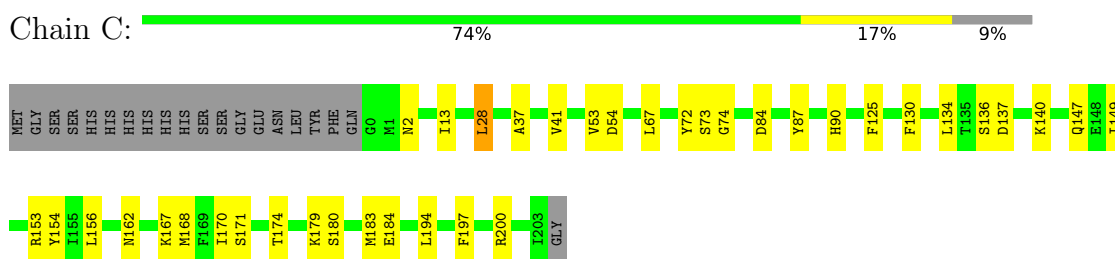
- Molecule 1: DNA-binding transcriptional activator EvgA



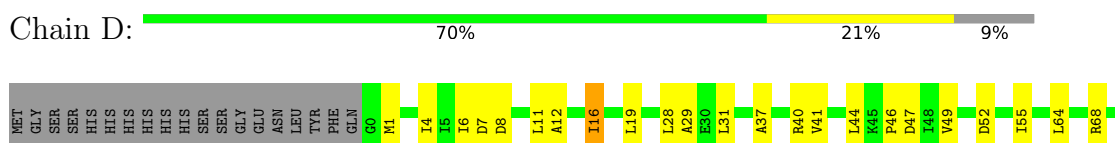
- Molecule 1: DNA-binding transcriptional activator EvgA



- Molecule 1: DNA-binding transcriptional activator EvgA



- Molecule 1: DNA-binding transcriptional activator EvgA



G74	I75	I76	I77	H90	G95	A96	N97	G105	M106	I110	A111	A112	K117	R129	S133	D142	M152	I155	G158	K159	D160	I164	K179	M183	L186	L191	M192	I203	GLY
-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	164.38Å 109.06Å 106.67Å 90.00° 125.39° 90.00°	Depositor
Resolution (Å)	46.20 – 2.90 46.20 – 2.90	Depositor EDS
% Data completeness (in resolution range)	67.7 (46.20-2.90) 67.8 (46.20-2.90)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.19 (at 2.91Å)	Xtriage
Refinement program	PHENIX v1.0	Depositor
R, R_{free}	0.199 , 0.259 0.200 , 0.257	Depositor DCC
R_{free} test set	1998 reflections (5.83%)	wwPDB-VP
Wilson B-factor (Å ²)	55.3	Xtriage
Anisotropy	0.069	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 45.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	6294	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.02% of the height of the origin peak. No significant pseudotranslation is detected.*

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](#)

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	0.15	0/1586	0.35	0/2141
1	B	0.12	0/1607	0.29	0/2165
1	C	0.12	0/1595	0.30	0/2151
1	D	0.14	0/1586	0.31	0/2140
All	All	0.13	0/6374	0.31	0/8597

There are no bond length outliers.

There are no bond angle outliers.

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	1566	0	1564	29	0
1	B	1587	0	1612	27	0
1	C	1575	0	1592	27	1
1	D	1566	0	1578	37	1
All	All	6294	0	6346	111	1

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 9.

The worst 5 of 111 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:15:ALA:HA	1:D:192:MET:HE1	1.71	0.73
1:D:68:ARG:NH1	1:D:95:GLY:O	2.22	0.72
1:D:19:LEU:HD22	1:D:106:MET:HG3	1.71	0.71
1:A:52:ASP:OD2	1:A:102:LYS:NZ	2.24	0.71
1:C:154:TYR:HE2	1:C:167:LYS:HD2	1.57	0.69

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:171:SER:OG	1:D:133:SER:OG[2_656]	2.19	0.01

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	203/224 (91%)	190 (94%)	13 (6%)	0	100	100
1	B	203/224 (91%)	198 (98%)	5 (2%)	0	100	100
1	C	202/224 (90%)	194 (96%)	8 (4%)	0	100	100
1	D	202/224 (90%)	198 (98%)	4 (2%)	0	100	100
All	All	810/896 (90%)	780 (96%)	30 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	168/192 (88%)	166 (99%)	2 (1%)	63	86
1	B	173/192 (90%)	171 (99%)	2 (1%)	63	86
1	C	171/192 (89%)	170 (99%)	1 (1%)	78	93
1	D	169/192 (88%)	168 (99%)	1 (1%)	78	93
All	All	681/768 (89%)	675 (99%)	6 (1%)	70	90

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	196	THR
1	C	28	LEU
1	D	16	ILE
1	A	145	SER
1	A	49	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	139	GLN
1	D	199	GLN
1	A	199	GLN
1	B	162	ASN
1	B	172	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 oligosaccharides 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](#)

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	205/224 (91%)	-0.01	4 (1%) 65 56	22, 50, 82, 102	0
1	B	205/224 (91%)	-0.28	0 100 100	20, 39, 64, 85	0
1	C	204/224 (91%)	0.06	0 100 100	30, 60, 79, 87	0
1	D	204/224 (91%)	-0.28	0 100 100	25, 41, 58, 69	0
All	All	818/896 (91%)	-0.13	4 (0%) 87 83	20, 47, 77, 102	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	204	GLY	3.1
1	A	169	PHE	2.8
1	A	170	ILE	2.6
1	A	174	THR	2.4

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 oligosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.