



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 22, 2020 – 02:54 AM BST

PDB ID : 4WWC
Title : Crystal structure of full length YvoA in complex with palindromic operator DNA
Authors : Grau, F.C.; Fillenberg, S.B.; Muller, Y.A.
Deposited on : 2014-11-10
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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.13.1
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) : 2.13.1

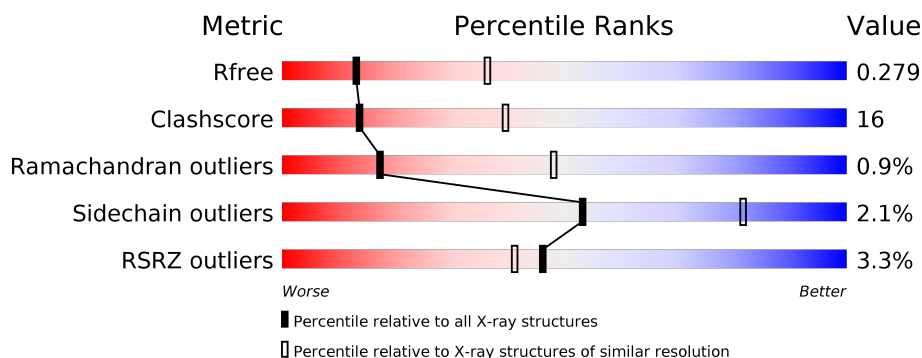
1 Overall quality at a glance

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}	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (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 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	246	<div> <div>4%</div> <div> <div></div> <div>57%</div> <div>29%</div> <div>13%</div> </div> </div>
1	B	246	<div> <div>2%</div> <div> <div></div> <div>60%</div> <div>33%</div> <div>6%</div> </div> </div>
2	E	19	<div> <div></div> <div> <div>21%</div> <div>79%</div> </div> </div>
2	F	19	<div> <div>5%</div> <div> <div></div> <div>21%</div> <div>68%</div> <div>11%</div> </div> </div>

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	EDO	E	102	-	-	-	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4350 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 repressor YvoA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	213	Total	C	N	O	S	0	1	0
			1698	1069	301	318	10			
1	B	231	Total	C	N	O	S	0	1	0
			1841	1159	321	351	10			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP O34817
A	-1	SER	-	expression tag	UNP O34817
A	0	HIS	-	expression tag	UNP O34817
B	-2	GLY	-	expression tag	UNP O34817
B	-1	SER	-	expression tag	UNP O34817
B	0	HIS	-	expression tag	UNP O34817

- Molecule 2 is a DNA chain called DNA (5'-D(P*CP*AP*GP*TP*GP*GP*TP*CP*TP*AP*GP*AP*CP*CP*AP*CP*TP*GP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	19	Total	C	N	O	P	0	0	0
			390	184	73	114	19			
2	F	19	Total	C	N	O	P	0	0	0
			391	185	73	114	19			

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).

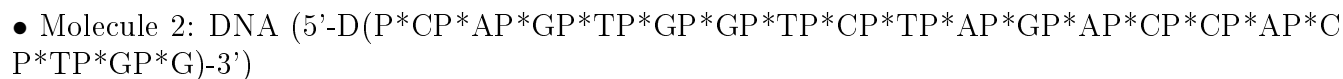
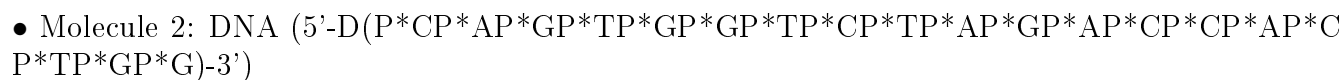
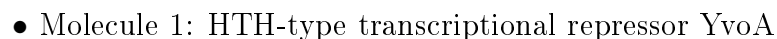


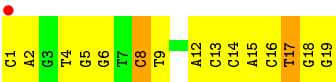
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		
3	E	1	Total	C	O	0	0
			4	2	2		
3	E	1	Total	C	O	0	0
			4	2	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	O	0	0
			1	1		
4	E	1	Total	O	0	0
			1	1		

- Molecule 1: HTH-type transcriptional repressor YvoA





4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	79.98Å 79.98Å 239.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.28 – 2.90 34.28 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.9 (34.28-2.90) 100.0 (34.28-2.90)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.23 (at 2.90Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.230 , 0.279 0.234 , 0.279	Depositor DCC
R_{free} test set	1444 reflections (8.00%)	wwPDB-VP
Wilson B-factor (Å ²)	90.1	Xtriage
Anisotropy	0.322	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 50.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	4350	wwPDB-VP
Average B, all atoms (Å ²)	85.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.95% 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](#)

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

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	0.49	0/1730	0.74	0/2327
1	B	0.50	0/1873	0.72	0/2521
2	E	1.19	0/437	1.14	1/672 (0.1%)
2	F	1.29	3/438 (0.7%)	1.16	3/674 (0.4%)
All	All	0.71	3/4478 (0.1%)	0.84	4/6194 (0.1%)

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	B	0	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	17	DT	C5-C7	-7.00	1.45	1.50
2	F	17	DT	N1-C6	-5.55	1.34	1.38
2	F	17	DT	C5-C6	-5.33	1.30	1.34

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	17	DT	C6-C5-C7	-7.19	118.59	122.90
2	E	7	DT	O5'-P-OP1	-5.76	100.52	105.70
2	F	17	DT	C4-C5-C7	5.41	122.25	119.00
2	F	8	DC	C1'-O4'-C4'	-5.24	104.86	110.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	75	SER	Peptide

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	1698	0	1720	54	0
1	B	1841	0	1840	65	0
2	E	390	0	211	18	0
2	F	391	0	214	17	0
3	A	4	0	6	0	0
3	B	16	0	24	3	0
3	E	8	0	12	0	0
4	A	1	0	0	0	0
4	E	1	0	0	0	0
All	All	4350	0	4027	133	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 16.

The worst 5 of 133 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:B:112:SER:HG	1:B:114:GLU:N	1.65	0.92
2:E:1:DC:O2	2:F:18:DG:N2	2.09	0.85
2:E:18:DG:H2'	2:E:19:DG:C8	2.19	0.78
2:E:17:DT:H2'	2:E:18:DG:C8	2.19	0.76
1:B:116:LEU:HA	1:B:119:ILE:HG12	1.69	0.74

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	205/246 (83%)	196 (96%)	7 (3%)	2 (1%)	15	45
1	B	221/246 (90%)	211 (96%)	8 (4%)	2 (1%)	17	48
All	All	426/492 (87%)	407 (96%)	15 (4%)	4 (1%)	17	48

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	122	CYS
1	B	238	TYR
1	A	86	LEU
1	B	76	LYS

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	187/215 (87%)	184 (98%)	3 (2%)	62	86
1	B	202/215 (94%)	197 (98%)	5 (2%)	47	78
All	All	389/430 (90%)	381 (98%)	8 (2%)	53	81

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	40	TYR
1	B	172	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	131	ILE
1	A	234	THR
1	B	102	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	237	HIS
1	B	237	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 ⓘ

7 ligands are modelled in this entry.

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	$\# Z > 2$
3	EDO	B	304	-	3,3,3	0.48	0	2,2,2	0.35	0
3	EDO	B	301	-	3,3,3	0.45	0	2,2,2	0.58	0
3	EDO	E	101	-	3,3,3	0.61	0	2,2,2	0.15	0
3	EDO	B	303	-	3,3,3	0.61	0	2,2,2	0.17	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	EDO	A	301	-	3,3,3	0.48	0	2,2,2	0.64	0
3	EDO	E	102	-	3,3,3	0.55	0	2,2,2	0.26	0
3	EDO	B	302	-	3,3,3	0.73	0	2,2,2	0.19	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
3	EDO	B	304	-	-	0/1/1/1	-
3	EDO	B	301	-	-	0/1/1/1	-
3	EDO	E	101	-	-	0/1/1/1	-
3	EDO	B	303	-	-	0/1/1/1	-
3	EDO	A	301	-	-	1/1/1/1	-
3	EDO	E	102	-	-	0/1/1/1	-
3	EDO	B	302	-	-	1/1/1/1	-

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
3	B	302	EDO	O1-C1-C2-O2
3	A	301	EDO	O1-C1-C2-O2

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	301	EDO	1	0
3	B	303	EDO	1	0
3	B	302	EDO	1	0

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	213/246 (86%)	0.16	9 (4%) 36 32	51, 83, 135, 147	0
1	B	231/246 (93%)	0.18	6 (2%) 56 52	51, 82, 129, 150	0
2	E	19/19 (100%)	-0.44	0 100 100	60, 77, 91, 97	0
2	F	19/19 (100%)	-0.10	1 (5%) 26 22	58, 78, 90, 92	0
All	All	482/530 (90%)	0.14	16 (3%) 46 41	51, 82, 131, 150	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	0[A]	HIS	4.6
1	A	122	CYS	3.6
1	B	175	SER	3.4
1	B	115	GLU	2.9
1	B	218	GLY	2.8

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(\AA^2)	Q<0.9
3	EDO	E	102	4/4	0.67	0.41	78,85,94,100	0
3	EDO	E	101	4/4	0.80	0.30	78,83,83,85	0
3	EDO	B	302	4/4	0.80	0.30	62,72,79,80	0
3	EDO	A	301	4/4	0.83	0.36	70,73,80,81	0
3	EDO	B	303	4/4	0.84	0.21	80,81,82,85	0
3	EDO	B	301	4/4	0.86	0.36	73,76,87,93	0
3	EDO	B	304	4/4	0.92	0.18	79,81,81,96	0

6.5 Other polymers ⓘ

There are no such residues in this entry.