W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 17, 2020 – 03:22 am BST

PDB ID	:	3VOX
Title	:	X-ray Crystal Structure of Wild Type HrtR in the Apo Form
Authors	:	Sawai, H.; Sugimoto, H.; Shiro, Y.; Aono, S.
Deposited on	:	2012-02-23
Resolution	:	3.10 Å(reported)

The following versions of software and data (see references (1)) 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 3.10 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.
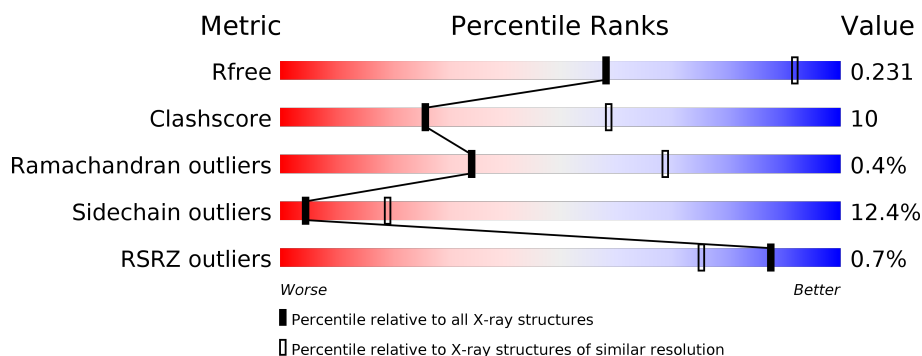
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 <=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	189	61% 26% 8% · ·
1	B	189	66% 23% · · 6%
1	C	189	62% 28% · 7%
1	D	189	63% 26% · 6%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 6136 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 regulator.

| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	181	Total	C	N	O	S	0	1	0
			1563	1015	250	296	2			
1	B	177	Total	C	N	O	S	0	1	0
			1530	996	242	290	2			
1	C	175	Total	C	N	O	S	0	0	0
			1510	983	237	288	2			
1	D	177	Total	C	N	O	S	0	1	0
			1533	997	243	291	2			

3 Residue-property plots ⓘ

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 (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.
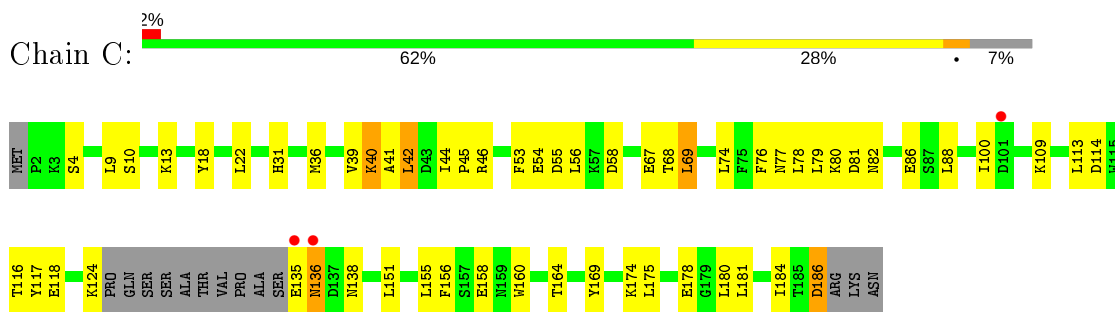
- Molecule 1: Transcriptional regulator
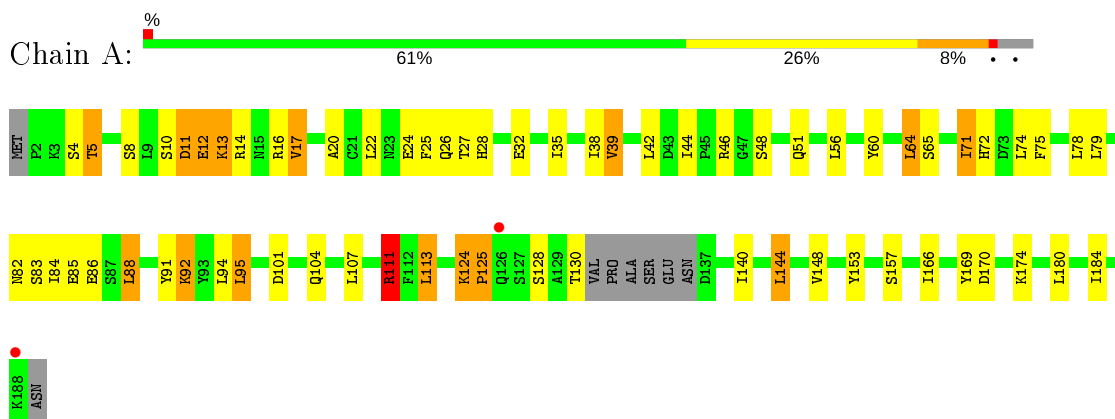
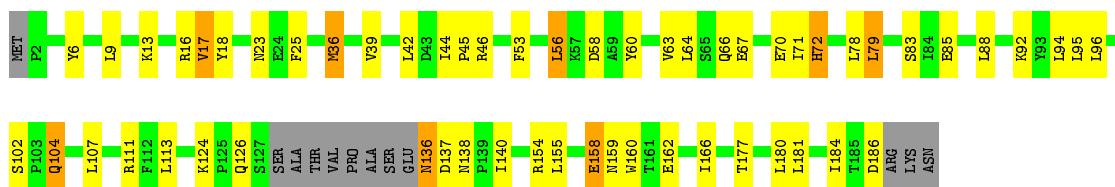

- Molecule 1: Transcriptional regulator
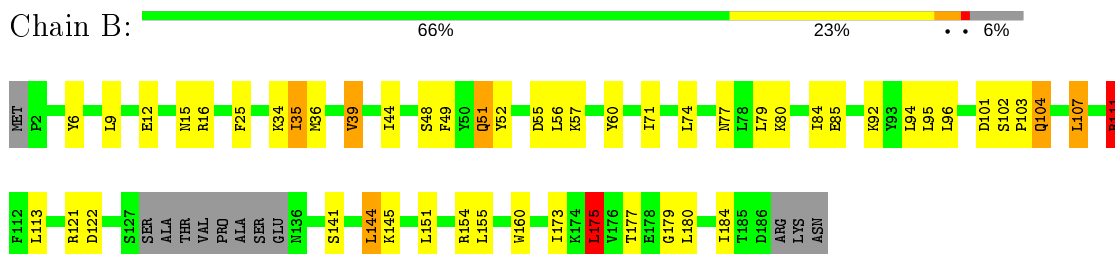
- Molecule 1: Transcriptional regulator

- Molecule 1: Transcriptional regulator

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	90.84Å　　190.18Å　　152.84Å 90.00°　　　90.00°　　　90.00°	Depositor
Resolution (Å)	29.90　–　3.10 29.90　–　3.07	Depositor EDS
% Data completeness (in resolution range)	99.8 (29.90-3.10) 98.6 (29.90-3.07)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	4.38 (at 3.06Å)	Xtriage
Refinement program	PHENIX 1.7.3_928	Depositor
R, R_{free}	0.184　,　0.241 0.174　,　0.231	Depositor DCC
R_{free} test set	1266 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	79.6	Xtriage
Anisotropy	0.397	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 62.9	EDS
L-test for twinning[2]	$< |L| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6136	wwPDB-VP
Average B, all atoms (Å^2)	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.76% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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 ⓘ

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	0.50	0/1604	0.67	2/2168 (0.1%)
1	B	0.52	0/1574	0.71	2/2129 (0.1%)
1	C	0.42	0/1549	0.59	0/2094
1	D	0.45	0/1574	0.59	0/2129
All	All	0.48	0/6301	0.64	4/8520 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	111	ARG	NE-CZ-NH1	7.64	124.12	120.30
1	B	111	ARG	NE-CZ-NH1	7.63	124.12	120.30
1	A	111	ARG	NE-CZ-NH2	-5.29	117.65	120.30
1	B	175	LEU	CA-CB-CG	5.26	127.41	115.30

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	1563	0	1511	40	0
1	B	1530	0	1475	29	0
1	C	1510	0	1454	26	0
1	D	1533	0	1474	33	0
All	All	6136	0	5914	126	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 10.

The worst 5 of 126 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:39:VAL:HG11	1:D:46:ARG:HA	1.61	0.83
1:A:111:ARG:HG2	1:A:111:ARG:HH11	1.47	0.77
1:A:140:ILE:HG23	1:B:180:LEU:HD11	1.72	0.71
1:D:44:ILE:HG12	1:D:45:PRO:HD2	1.72	0.70
1:C:136:ASN:HB3	1:C:138:ASN:H	1.57	0.70

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	178/189 (94%)	166 (93%)	10 (6%)	2 (1%)	14	46
1	B	174/189 (92%)	162 (93%)	12 (7%)	0	100	100
1	C	171/189 (90%)	166 (97%)	5 (3%)	0	100	100
1	D	174/189 (92%)	166 (95%)	7 (4%)	1 (1%)	25	59
All	All	697/756 (92%)	660 (95%)	34 (5%)	3 (0%)	34	69

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	71	ILE
1	A	125	PRO
1	A	11	ASP

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	175/181 (97%)	150 (86%)	25 (14%)	3	14
1	B	172/181 (95%)	153 (89%)	19 (11%)	6	25
1	C	169/181 (93%)	147 (87%)	22 (13%)	4	18
1	D	172/181 (95%)	153 (89%)	19 (11%)	6	25
All	All	688/724 (95%)	603 (88%)	85 (12%)	4	19

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	107	LEU
1	C	42	LEU
1	D	111	ARG
1	B	111	ARG
1	B	184	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	72	HIS
1	A	138	ASN
1	D	126	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)

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(Å²)	Q<0.9
1	A	181/189 (95%)	-0.30	2 (1%)	80	64	59, 94, 143, 185	0
1	B	177/189 (93%)	-0.54	0	100	100	52, 76, 128, 171	0
1	C	175/189 (92%)	-0.24	3 (1%)	70	49	72, 108, 154, 190	0
1	D	177/189 (93%)	-0.36	0	100	100	68, 96, 140, 166	0
All	All	710/756 (93%)	-0.36	5 (0%)	87	75	52, 94, 143, 190	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	135	GLU	4.1
1	A	126	GLN	3.2
1	C	101	ASP	2.4
1	A	188	LYS	2.3
1	C	136	ASN	2.1

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.