Full wwPDB X-ray Structure Validation Report (i)

Dec 2, 2024 – 11:39 AM EST

PDB ID	:	2FX0
Title	:	Crystal Structure of HlyIIR, a Hemolysin II transcriptional Regulator
Authors	:	Kovalevskiy, O.V.; Lebedev, A.A.; Solonin, A.S.; Antson, A.A.
Deposited on	:	2006-02-03
Resolution	:	2.40 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

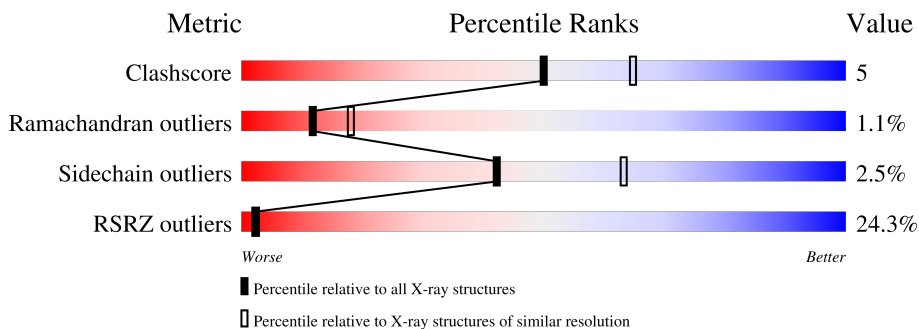|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 2.40 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	Percentile Ranks	Value
Clashscore		5
Ramachandran outliers		1.1%
Sidechain outliers		2.5%
RSRZ outliers		24.3%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	180529	5218 (2.40-2.40)
Ramachandran outliers	177936	5158 (2.40-2.40)
Sidechain outliers	177891	5159 (2.40-2.40)
RSRZ outliers	164620	4642 (2.40-2.40)

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%

Mol	Chain	Length	Quality of chain
1	A	201	21% 75% 14% 11%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 1590 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 hemolysin II regulatory protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	179	Total 1476	C 962	N 240	O 272	Se 2	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	modified residue	GB 31322346
A	9	MSE	MET	modified residue	GB 31322346
A	43	MSE	MET	modified residue	GB 31322346
A	183	MSE	MET	modified residue	GB 31322346

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	114	Total 114	O 114	0	0

3 Residue-property plots ⓘ

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.

• Molecule 1: hemolysin II regulatory protein
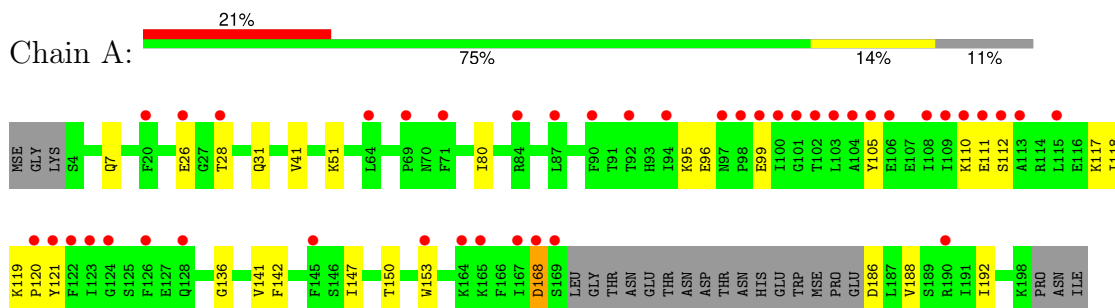
4　Data and refinement statistics　ⓘ

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α, β, γ	124.21Å　124.21Å　79.67Å 90.00°　　90.00°　　120.00°	Depositor
Resolution (Å)	24.68　–　2.40 24.68　–　2.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (24.68-2.40) 99.2 (24.68-2.40)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.33 (at 2.41Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.221　,　　　0.276 0.223　,　(Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	44.4	Xtriage
Anisotropy	0.235	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 69.5	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	1590	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.76% 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	0.65	2/1506 (0.1%)	0.56	0/2019

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	A	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	99	GLU	CD-OE1	13.35	1.40	1.25
1	A	99	GLU	CB-CG	6.42	1.64	1.52

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	110	LYS	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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1476	0	1483	15	1
2	A	114	0	0	2	0
All	All	1590	0	1483	15	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 5.

All (15) 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:80:ILE:HD11	1:A:141:VAL:HG21	1.73	0.71
1:A:26:GLU:CD	1:A:26:GLU:H	2.09	0.55
1:A:119:LYS:C	1:A:121:TYR:H	2.13	0.52
1:A:119:LYS:O	1:A:121:TYR:N	2.45	0.50
1:A:95:LYS:HG3	1:A:96:GLU:N	2.26	0.48
1:A:7:GLN:NE2	2:A:306:HOH:O	2.46	0.48
1:A:31:GLN:HG3	1:A:41:VAL:HG21	1.97	0.46
1:A:28:THR:O	1:A:51:LYS:HE3	2.15	0.46
1:A:111:GLU:HB3	1:A:112:SER:H	1.52	0.45
1:A:136:GLY:HA3	1:A:142:PHE:CD2	2.52	0.44
1:A:31:GLN:HA	1:A:41:VAL:HG22	2.00	0.44
1:A:188:VAL:O	1:A:192:ILE:HG12	2.19	0.42
1:A:147:ILE:HG22	2:A:301:HOH:O	2.20	0.41
1:A:150:THR:HA	1:A:153:TRP:CD1	2.56	0.41
1:A:119:LYS:C	1:A:121:TYR:N	2.75	0.40

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:A:105:TYR:CB	1:A:105:TYR:CZ[10_666]	2.18	0.02

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	175/201 (87%)	164 (94%)	9 (5%)	2 (1%)	12	18

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	168	ASP
1	A	120	PRO

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	159/175 (91%)	155 (98%)	4 (2%)	42	63

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	117	LYS
1	A	118	ILE
1	A	168	ASP
1	A	186	ASP

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

Mol	Chain	Res	Type
1	A	7	GLN
1	A	53	ASN

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

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 ⓘ

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 ⓘ

Warning: The R factor obtained from EDS is 0.2955, which does not match the depositor's R factor of 0.22085. Please interpret the results in this section carefully.

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	177/201 (88%)	1.49	43 (24%)	2	2	26, 51, 67, 72	7 (3%)

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	99	GLU	7.6
1	A	110	LYS	6.2
1	A	102	THR	5.3
1	A	100	ILE	5.1
1	A	64	LEU	4.7
1	A	101	GLY	4.7
1	A	165	LYS	4.2
1	A	113	ALA	4.2
1	A	164	LYS	4.1
1	A	121	TYR	3.9
1	A	168	ASP	3.8
1	A	105	TYR	3.7
1	A	71	PHE	3.5
1	A	112	SER	3.5
1	A	94	ILE	3.5
1	A	111	GLU	3.3
1	A	122	PHE	3.2
1	A	104	ALA	3.2
1	A	26	GLU	3.1
1	A	92	THR	3.0
1	A	115	LEU	2.8
1	A	124	GLY	2.7
1	A	169	SER	2.7
1	A	145	PHE	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	153	TRP	2.6
1	A	120	PRO	2.5
1	A	108	ILE	2.5
1	A	167	ILE	2.5
1	A	28	THR	2.5
1	A	103	LEU	2.4
1	A	106	GLU	2.4
1	A	90	PHE	2.4
1	A	69	PRO	2.3
1	A	97	ASN	2.3
1	A	190	ARG	2.3
1	A	128	GLN	2.3
1	A	109	ILE	2.3
1	A	98	PRO	2.2
1	A	20	PHE	2.1
1	A	126	PHE	2.1
1	A	84	ARG	2.1
1	A	87	LEU	2.1
1	A	123	ILE	2.1

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.