wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 27, 2014 – 03:37 PM GMT

PDB ID	:	1VH0
Title	:	Crystal structure of a hypothetical protein
Authors	:	Structural GenomiX
Deposited on	:	2003-11-03
Resolution	:	2.31 Å(reported)

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 2.31 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
Rfree		0.275
Clashscore		14
Ramachandran outliers		0.1%
Sidechain outliers		5.1%
RSRZ outliers		13.5%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
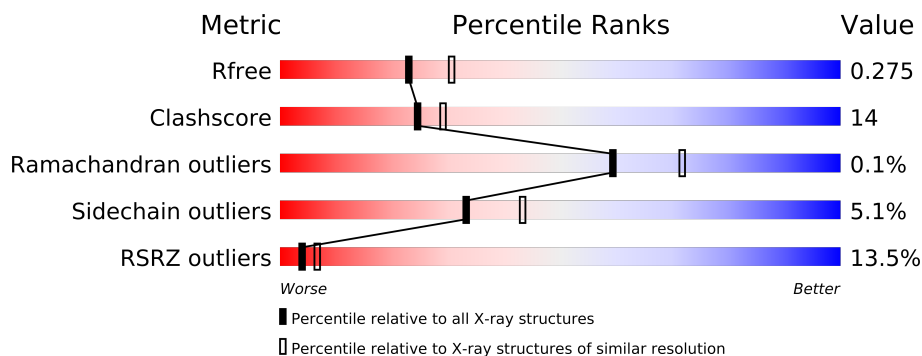
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	3293 (2.34-2.30)
Clashscore	79885	4097 (2.34-2.30)
Ramachandran outliers	78287	4055 (2.34-2.30)
Sidechain outliers	78261	4054 (2.34-2.30)
RSRZ outliers	66119	3294 (2.34-2.30)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	161	
1	B	161	
1	C	161	
1	D	161	
1	E	161	
1	F	161	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 7758 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Hypothetical UPF0247 protein SAV0024/SA0023.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	157	Total	C	N	O	S	0	0	0
			1260	804	215	234	7			
1	B	157	Total	C	N	O	S	0	0	0
			1260	804	215	234	7			
1	C	157	Total	C	N	O	S	0	0	0
			1258	804	215	232	7			
1	D	148	Total	C	N	O	S	0	0	0
			1187	761	204	216	6			
1	E	157	Total	C	N	O	S	0	0	0
			1248	797	214	230	7			
1	F	154	Total	C	N	O	S	0	0	0
			1224	784	209	225	6			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	cloning artifact	UNP P0A0N8
A	0	SER	-	cloning artifact	UNP P0A0N8
A	1	LEU	-	cloning artifact	UNP P0A0N8
B	-1	MET	-	cloning artifact	UNP P0A0N8
B	0	SER	-	cloning artifact	UNP P0A0N8
B	1	LEU	-	cloning artifact	UNP P0A0N8
C	-1	MET	-	cloning artifact	UNP P0A0N8
C	0	SER	-	cloning artifact	UNP P0A0N8
C	1	LEU	-	cloning artifact	UNP P0A0N8
D	-1	MET	-	cloning artifact	UNP P0A0N8
D	0	SER	-	cloning artifact	UNP P0A0N8
D	1	LEU	-	cloning artifact	UNP P0A0N8
E	-1	MET	-	cloning artifact	UNP P0A0N8
E	0	SER	-	cloning artifact	UNP P0A0N8
E	1	LEU	-	cloning artifact	UNP P0A0N8
F	-1	MET	-	cloning artifact	UNP P0A0N8
F	0	SER	-	cloning artifact	UNP P0A0N8

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

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	1	LEU	-	cloning artifact	UNP P0A0N8

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	76	Total 76	O 76	0	0
2	B	43	Total 43	O 43	0	0
2	C	45	Total 45	O 45	0	0
2	D	47	Total 47	O 47	0	0
2	E	49	Total 49	O 49	0	0
2	F	61	Total 61	O 61	0	0

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 errors 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.
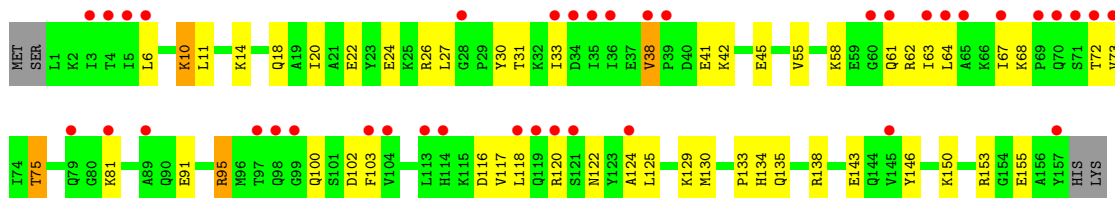
- Molecule 1: Hypothetical UPF0247 protein SAV0024/SA0023
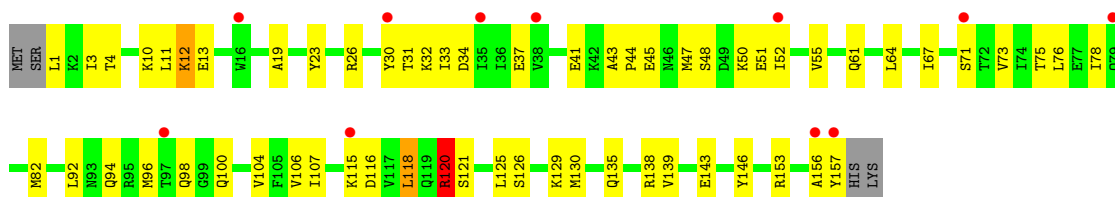
Chain A:
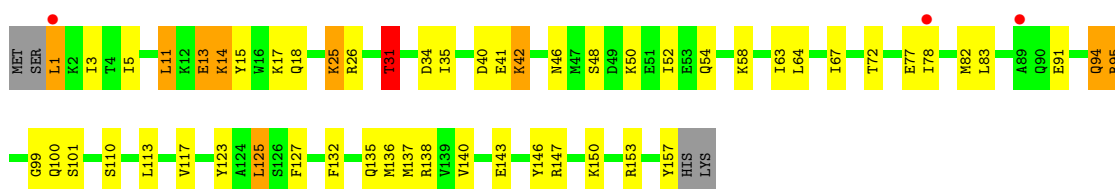
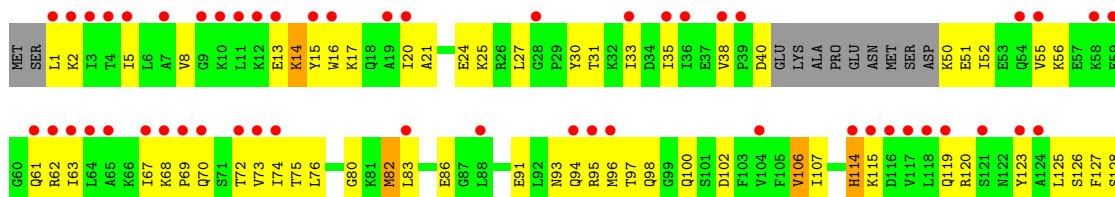- Molecule 1: Hypothetical UPF0247 protein SAV0024/SA0023

Chain B:

- Molecule 1: Hypothetical UPF0247 protein SAV0024/SA0023

Chain C:

- Molecule 1: Hypothetical UPF0247 protein SAV0024/SA0023

Chain D:

H134 Q135 V139 I142 Y146 M152 Y157 HIS LYS

• Molecule 1: Hypothetical UPF0247 protein SAV0024/SA0023

Chain E:

MET SER L1 K2 I3 T4 I5 L6 K10 L11 K12 E13 K14 I20 A21 E22 Y23 E24 K25 R26 L27 G28 T31 K32 I33 I36 E37 E41 K42 A43 P44 D49 K50 Q54 V55 K56 E57 K58 E59 G60 Q61 R62 I63 L64 A65 K66 I67 I74 M82 Q90

E91 L92 N93 Q94 R95 M96 T97 Q98 S101 F105 G109 G112 R120 S126 F127 S128 K129 M130 M136 M137 R138 L141 I142 V145 Y146 F149 R153 G154 E155 A156 Y157 HIS LYS

• Molecule 1: Hypothetical UPF0247 protein SAV0024/SA0023

Chain F:

MET SER L1 L11 K12 E13 K14 Q18 A19 I20 E24 K25 R26 I33 P39 D40 E41 K42 A43 P44 GLU ASN MET S48 E53 Q54 V55 K58 E59 R62 T75 K81 M82 L83 G87 Q90 R95 S101 V106 N111 K115 D116

Y123 A124 L125 S126 F127 M130 P133 H134 Q135 M136 Y157 HIS LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	48.53Å 74.50Å 81.53Å 91.06° 107.14° 90.14°	Depositor
Resolution (Å)	19.92 – 2.31 19.94 – 2.31	Depositor EDS
% Data completeness (in resolution range)	(Not available) (19.92-2.31) 96.9 (19.94-2.31)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	8.52 (at 2.30Å)	Xtriage
Refinement program	REFMAC 4	Depositor
R, R_{free}	0.260 , 0.318 0.250 , 0.275	Depositor DCC
R_{free} test set	2359 reflections (5.31%)	DCC
Wilson B-factor (Å^2)	30.2	Xtriage
Anisotropy	0.554	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 45.5	EDS
Estimated twinning fraction	0.269 for h,-k,-h-l 0.012 for -h,k,-l 0.006 for -h,-k,h+l	Xtriage
L-test for twinning	$<\|L\|> = 0.50, <L^2> = 0.34$	Xtriage
Outliers	0 of 46732 reflections	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	7758	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.62% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

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	0.92	0/1279	1.33	7/1714 (0.4%)
1	B	0.90	0/1279	1.43	10/1714 (0.6%)
1	C	0.86	0/1277	1.25	2/1711 (0.1%)
1	D	0.85	0/1204	1.21	1/1612 (0.1%)
1	E	0.84	0/1267	1.26	5/1699 (0.3%)
1	F	0.86	0/1242	1.27	8/1665 (0.5%)
All	All	0.87	0/7548	1.30	33/10115 (0.3%)

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	4
1	B	0	1
1	C	0	2
1	D	0	3
1	E	0	2
All	All	0	12

There are no bond length outliers.

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	96	MET	CA-CB-CG	11.89	133.51	113.30
1	B	153	ARG	CD-NE-CZ	11.59	139.83	123.60
1	A	95	ARG	NE-CZ-NH1	-10.03	115.28	120.30
1	B	95	ARG	NE-CZ-NH1	-9.85	115.38	120.30
1	F	95	ARG	NE-CZ-NH2	-8.28	116.16	120.30

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	132	PHE	Mainchain
1	A	146	TYR	Mainchain
1	A	64	LEU	Mainchain
1	A	94	GLN	Mainchain
1	B	102	ASP	Mainchain

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1260	0	1303	32	0
1	B	1260	0	1303	44	0
1	C	1258	0	1303	50	0
1	D	1187	0	1229	44	0
1	E	1248	0	1281	38	0
1	F	1224	0	1262	18	0
2	A	76	0	0	4	0
2	B	43	0	0	4	0
2	C	45	0	0	3	0
2	D	47	0	0	7	0
2	E	49	0	0	2	0
2	F	61	0	0	0	0
All	All	7758	0	7681	212	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 14.

The worst 5 of 212 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:3:ILE:HG12	1:A:31:THR:HG21	1.54	0.86
1:B:91:GLU:OE2	1:B:95:ARG:NH1	2.09	0.84
1:C:3:ILE:HG12	1:C:31:THR:HG21	1.59	0.84
1:C:104:VAL:HG22	2:C:183:HOH:O	1.76	0.84
1:B:146:TYR:CZ	1:B:150:LYS:HD2	2.15	0.82

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	155/161 (96%)	153 (99%)	2 (1%)	0	100	100
1	B	155/161 (96%)	148 (96%)	7 (4%)	0	100	100
1	C	155/161 (96%)	149 (96%)	6 (4%)	0	100	100
1	D	144/161 (89%)	137 (95%)	6 (4%)	1 (1%)	30	34
1	E	155/161 (96%)	150 (97%)	5 (3%)	0	100	100
1	F	150/161 (93%)	148 (99%)	2 (1%)	0	100	100
All	All	914/966 (95%)	885 (97%)	28 (3%)	1 (0%)	59	72

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	80	GLY

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	137/142 (96%)	125 (91%)	12 (9%)	14	16
1	B	137/142 (96%)	132 (96%)	5 (4%)	47	62
1	C	136/142 (96%)	129 (95%)	7 (5%)	33	43
1	D	127/142 (89%)	118 (93%)	9 (7%)	21	25
1	E	133/142 (94%)	130 (98%)	3 (2%)	63	80
1	F	131/142 (92%)	126 (96%)	5 (4%)	44	59

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	801/852 (94%)	760 (95%)	41 (5%)	33 43

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	12	LYS
1	C	125	LEU
1	F	53	GLU
1	C	78	ILE
1	C	94	GLN

Some 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	61	GLN
1	D	98	GLN
1	F	90	GLN
1	C	111	ASN
1	E	98	GLN

5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

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

There are no chain breaks in this entry.

6   Fit of model and data ⓘ

6.1   Protein, DNA and RNA chains ⓘ

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	157/161 (97%)	0.49	3 (1%)	64	73	28, 41, 55, 62	0
1	B	157/161 (97%)	1.34	39 (24%)	1	2	30, 47, 60, 64	0
1	C	157/161 (97%)	0.95	11 (7%)	16	24	31, 51, 69, 72	0
1	D	148/161 (91%)	1.68	53 (35%)	1	1	31, 50, 65, 75	0
1	E	157/161 (97%)	1.05	17 (10%)	6	10	29, 52, 69, 82	0
1	F	154/161 (95%)	0.54	3 (1%)	64	73	30, 43, 63, 68	0
All	All	930/966 (96%)	1.00	126 (13%)	4	6	28, 48, 65, 82	0

The worst 5 of 126 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	60	GLY	10.1
1	D	65	ALA	7.0
1	D	55	VAL	6.4
1	E	153	ARG	6.4
1	D	117	VAL	6.2

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 carbohydrates in this entry.

6.4   Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.