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)

Mar 10, 2024 – 06:02 AM EDT

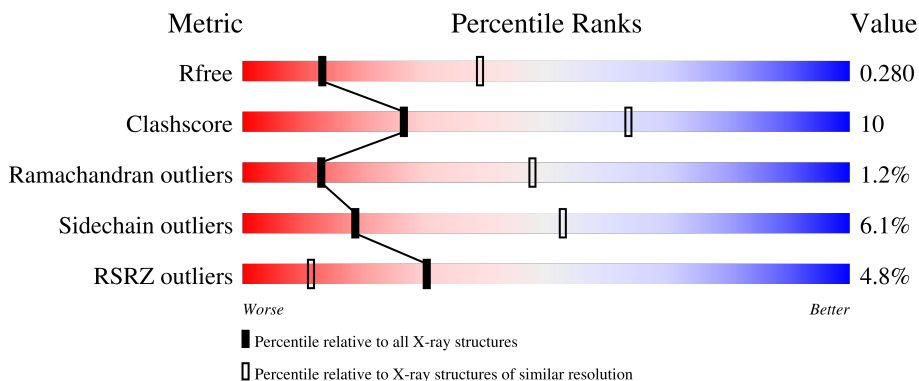|		
PDB ID	:	4G1U
Title	:	X-ray structure of the bacterial heme transporter HmuUV from Yersinia pestis
Authors	:	Woo, J.-S.; Goetz, B.A.; Zeltina, A.; Locher, K.P.
Deposited on	:	2012-07-11
Resolution	:	3.01 Å(reported)

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 3.01 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		0.280
Clashscore		10
Ramachandran outliers		1.2%
Sidechain outliers		6.1%
RSRZ outliers		4.8%

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}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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% 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	357	5% 56% / 25% / • / 17%
1	B	357	5% 58% / 24% / • / 17%
2	C	266	% 72% / 24% / • •
2	D	266	6% 73% / 21% / • •

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 8499 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 Hemin transport system permease protein hmuU.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	297	Total	C	N	O	S	0	0	0
			2209	1451	380	365	13			
1	B	297	Total	C	N	O	S	0	0	0
			2209	1451	380	365	13			

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-22	MET	-	expression tag	UNP Q56992
A	-21	GLY	-	expression tag	UNP Q56992
A	-20	HIS	-	expression tag	UNP Q56992
A	-19	HIS	-	expression tag	UNP Q56992
A	-18	HIS	-	expression tag	UNP Q56992
A	-17	HIS	-	expression tag	UNP Q56992
A	-16	HIS	-	expression tag	UNP Q56992
A	-15	HIS	-	expression tag	UNP Q56992
A	-14	HIS	-	expression tag	UNP Q56992
A	-13	HIS	-	expression tag	UNP Q56992
A	-12	HIS	-	expression tag	UNP Q56992
A	-11	HIS	-	expression tag	UNP Q56992
A	-10	SER	-	expression tag	UNP Q56992
A	-9	SER	-	expression tag	UNP Q56992
A	-8	GLY	-	expression tag	UNP Q56992
A	-7	GLU	-	expression tag	UNP Q56992
A	-6	ASN	-	expression tag	UNP Q56992
A	-5	LEU	-	expression tag	UNP Q56992
A	-4	TYR	-	expression tag	UNP Q56992
A	-3	PHE	-	expression tag	UNP Q56992
A	-2	GLN	-	expression tag	UNP Q56992
A	-1	GLY	-	expression tag	UNP Q56992
A	0	HIS	-	expression tag	UNP Q56992
B	-22	MET	-	expression tag	UNP Q56992
B	-21	GLY	-	expression tag	UNP Q56992

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-20	HIS	-	expression tag	UNP Q56992
B	-19	HIS	-	expression tag	UNP Q56992
B	-18	HIS	-	expression tag	UNP Q56992
B	-17	HIS	-	expression tag	UNP Q56992
B	-16	HIS	-	expression tag	UNP Q56992
B	-15	HIS	-	expression tag	UNP Q56992
B	-14	HIS	-	expression tag	UNP Q56992
B	-13	HIS	-	expression tag	UNP Q56992
B	-12	HIS	-	expression tag	UNP Q56992
B	-11	HIS	-	expression tag	UNP Q56992
B	-10	SER	-	expression tag	UNP Q56992
B	-9	SER	-	expression tag	UNP Q56992
B	-8	GLY	-	expression tag	UNP Q56992
B	-7	GLU	-	expression tag	UNP Q56992
B	-6	ASN	-	expression tag	UNP Q56992
B	-5	LEU	-	expression tag	UNP Q56992
B	-4	TYR	-	expression tag	UNP Q56992
B	-3	PHE	-	expression tag	UNP Q56992
B	-2	GLN	-	expression tag	UNP Q56992
B	-1	GLY	-	expression tag	UNP Q56992
B	0	HIS	-	expression tag	UNP Q56992

- Molecule 2 is a protein called Hemin import ATP-binding protein HmuV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	262	Total 2052	C 1294	N 374	O 374	S 10	0	0	0
2	D	257	Total 2019	C 1272	N 369	O 368	S 10	0	0	0

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).

PO4

O3

O$^-$

O2 O$^-$ P O O1

P

O$^-$

O4
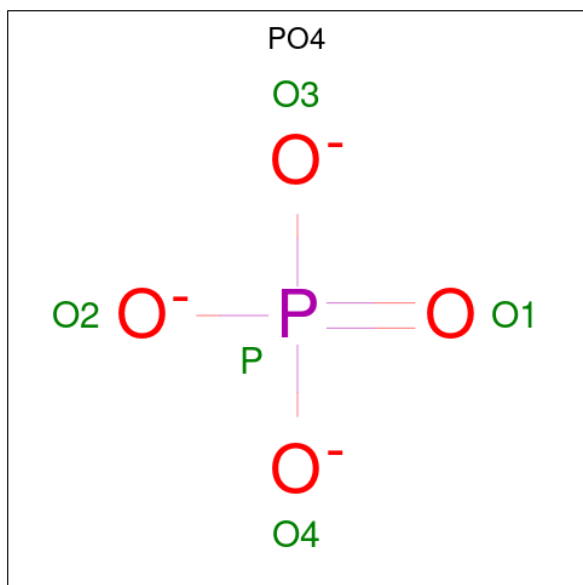
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	O	P	0	0
			5	4	1		
3	D	1	Total	O	P	0	0
			5	4	1		

WORLDWIDE

PDB

PROTEIN DATA BANK

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 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: Hemin transport system permease protein hmuU

Chain A:

• Molecule 1: Hemin transport system permease protein hmuU

Chain B:

• Molecule 2: Hemin import ATP-binding protein HmuV

%

Chain C:

72%   24%

● Molecule 2: Hemin import ATP-binding protein HmuV

6%

Chain D:

73%   21%
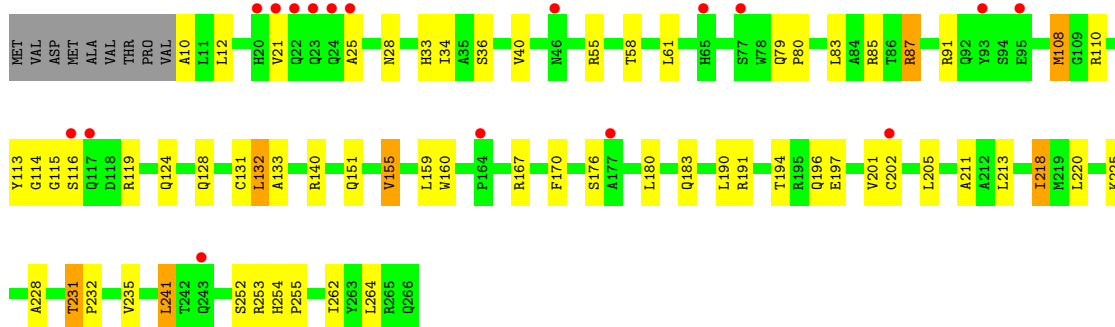
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	100.87Å 101.23Å 105.21Å 90.00° 91.30° 90.00°	Depositor
Resolution (Å)	29.77 – 3.01 29.77 – 3.01	Depositor EDS
% Data completeness (in resolution range)	88.4 (29.77-3.01) 88.4 (29.77-3.01)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.37 (at 3.00Å)	Xtriage
Refinement program	PHENIX 1.7.2_869	Depositor
R, R_{free}	0.263 , 0.295 0.256 , 0.280	Depositor DCC
R_{free} test set	1971 reflections (5.28%)	wwPDB-VP
Wilson B-factor (Å^2)	72.6	Xtriage
Anisotropy	0.009	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 42.2	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.000 for -h,l,k 0.004 for -h,-l,-k 0.000 for -l,k,h 0.000 for -k,-h,-l 0.000 for k,h,-l 0.000 for k,l,h 0.000 for l,h,k 0.000 for l,-h,-k 0.000 for -k,-l,h 0.000 for h,-k,-l 0.016 for l,-k,h	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	8499	wwPDB-VP
Average B, all atoms (Å^2)	106.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.08% 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PO4

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	0.35	0/2247	0.51	0/3061
1	B	0.33	0/2247	0.52	0/3061
2	C	0.32	0/2095	0.48	0/2854
2	D	0.27	0/2061	0.46	0/2805
All	All	0.32	0/8650	0.49	0/11781

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 ⓘ

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	2209	0	2380	61	0
1	B	2209	0	2380	59	0
2	C	2052	0	2060	39	0
2	D	2019	0	2023	36	0
3	C	5	0	0	1	0
3	D	5	0	0	0	0
All	All	8499	0	8843	179	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 179 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)
2:C:167:ARG:NH2	2:C:197:GLU:OE1	2.11	0.83
2:D:167:ARG:NH2	2:D:197:GLU:OE1	2.12	0.81
1:A:76:MET:HE1	1:A:247:ILE:HD11	1.65	0.77
2:C:176:SER:HB2	2:D:176:SER:HB2	1.69	0.75
1:A:86:ASP:OD1	1:A:88:GLY:N	2.23	0.72

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	291/357 (82%)	271 (93%)	17 (6%)	3 (1%)	15	53
1	B	291/357 (82%)	272 (94%)	17 (6%)	2 (1%)	22	60
2	C	260/266 (98%)	239 (92%)	17 (6%)	4 (2%)	10	42
2	D	255/266 (96%)	233 (91%)	18 (7%)	4 (2%)	9	40
All	All	1097/1246 (88%)	1015 (92%)	69 (6%)	13 (1%)	13	48

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	304	PRO
1	B	304	PRO
2	C	114	GLY
2	D	114	GLY
1	B	146	ALA

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	231/283 (82%)	218 (94%)	13 (6%)	21	56
1	B	231/283 (82%)	219 (95%)	12 (5%)	23	59
2	C	219/223 (98%)	204 (93%)	15 (7%)	16	48
2	D	215/223 (96%)	200 (93%)	15 (7%)	15	47
All	All	896/1012 (88%)	841 (94%)	55 (6%)	18	53

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	83	LEU
2	C	155	VAL
2	D	241	LEU
2	D	155	VAL
2	C	87	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	6	GLN
1	B	6	GLN
2	C	128	GLN
2	D	128	GLN

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.

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

2 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	PO4	D	301	-	4,4,4	0.94	0	6,6,6	0.53	0
3	PO4	C	301	-	4,4,4	0.55	0	6,6,6	0.69	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	301	PO4	1	0

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 ⓘ

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	297/357 (83%)	0.04	17 (5%)	23	8	52, 102, 163, 207	0
1	B	297/357 (83%)	0.04	17 (5%)	23	8	74, 110, 176, 215	0
2	C	262/266 (98%)	-0.31	2 (0%)	86	65	67, 79, 124, 163	0
2	D	257/266 (96%)	0.13	17 (6%)	18	5	64, 107, 171, 208	0
All	All	1113/1246 (89%)	-0.02	53 (4%)	30	11	52, 99, 168, 215	0

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	164	PRO	6.1
2	D	116	SER	6.0
2	D	65	HIS	5.9
1	B	326	LEU	5.1
1	A	142	HIS	4.7

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 ⓘ

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(Å^2)	Q<0.9
3	PO4	C	301	5/5	0.74	0.35	56,61,80,101	0
3	PO4	D	301	5/5	0.87	0.12	133,135,165,185	0

6.5 Other polymers ⓘ

There are no such residues in this entry.