wwPDB X-ray Structure Validation Summary Report (i)

Sep 8, 2025 – 05:47 pm BST

PDB ID	:	9HQ1 / pdb_00009hq1
Title	:	XusB lipoprotein bound to ferric salmochelin
Authors	:	Silale, A.; Soo, Y.L.; van den Berg, B.
Deposited on	:	2024-12-16
Resolution	:	1.80 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

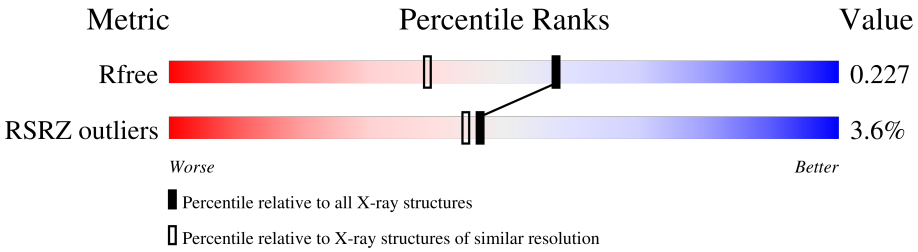
MolProbity	:	FAILED
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.45.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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}	164625	7108 (1.80-1.80)
RSRZ outliers	164620	7108 (1.80-1.80)

MolProbity failed to run properly - the sequence quality summary graphics cannot be shown.

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 7471 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.

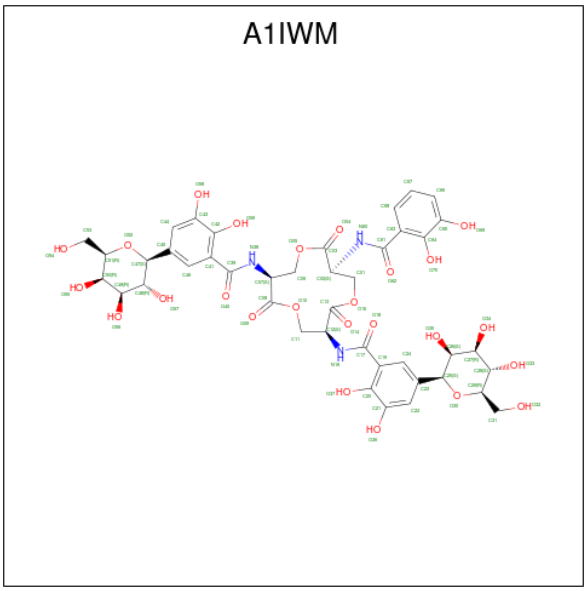
- Molecule 1 is a protein called DUF4374 domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	426	Total	C	N	O	S	0	1	0
			3307	2090	547	660	10			
1	B	426	Total	C	N	O	S	0	0	0
			3298	2084	545	659	10			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	465	LEU	-	expression tag	UNP Q8A622
A	466	GLU	-	expression tag	UNP Q8A622
A	467	HIS	-	expression tag	UNP Q8A622
A	468	HIS	-	expression tag	UNP Q8A622
A	469	HIS	-	expression tag	UNP Q8A622
A	470	HIS	-	expression tag	UNP Q8A622
A	471	HIS	-	expression tag	UNP Q8A622
A	472	HIS	-	expression tag	UNP Q8A622
B	465	LEU	-	expression tag	UNP Q8A622
B	466	GLU	-	expression tag	UNP Q8A622
B	467	HIS	-	expression tag	UNP Q8A622
B	468	HIS	-	expression tag	UNP Q8A622
B	469	HIS	-	expression tag	UNP Q8A622
B	470	HIS	-	expression tag	UNP Q8A622
B	471	HIS	-	expression tag	UNP Q8A622
B	472	HIS	-	expression tag	UNP Q8A622

- Molecule 2 is Salmochelin S4 (CCD ID: A1IWM) (formula: $C_{42}H_{47}N_3O_{25}$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C N O 70 42 3 25	0	0

- Molecule 3 is FE (III) ION (CCD ID: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Fe 1 1	0	0

- Molecule 4 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Ca 2 2	0	0
4	B	4	Total Ca 4 4	0	0

- Molecule 5 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	2	Total Cl 2 2	0	0
5	B	2	Total Cl 2 2	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	384	Total O 384 384	0	0
6	B	401	Total O 401 401	0	0

MolProbity failed to run properly - this section is therefore empty.

3 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	45.44Å 108.72Å 93.83Å 90.00° 96.66° 90.00°	Depositor
Resolution (Å)	54.36 – 1.80 54.36 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (54.36-1.80) 99.7 (54.36-1.80)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.66 (at 1.80Å)	Xtriage
Refinement program	PHENIX 1.21.1_5286	Depositor
R, R_{free}	0.193 , 0.229 0.193 , 0.227	Depositor DCC
R_{free} test set	4149 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	23.3	Xtriage
Anisotropy	0.213	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 32.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7471	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.73% 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.

4 Model quality

4.1 Standard geometry

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles

4.3.1 Protein backbone

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA

MolProbity failed to run properly - this section is therefore empty.

4.4 Non-standard residues in protein, DNA, RNA chains

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

4.5 Carbohydrates

There are no oligosaccharides in this entry.

4.6 Ligand geometry

Of 12 ligands modelled in this entry, 11 are monoatomic - leaving 1 for Mogul analysis.

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	$\#\|Z\| > 2$
2	A1IWM	A	501	3	75,75,75	2.07	22 (29%)	110,110,110	1.92	32 (29%)

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
2	A1IWM	A	501	3	-	2/63/103/103	0/5/6/6

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	A1IWM	C26-C25	5.31	1.65	1.53
2	A	501	A1IWM	C17-N16	5.20	1.45	1.34
2	A	501	A1IWM	O05-C03	4.31	1.42	1.33
2	A	501	A1IWM	O52-C47	-4.19	1.38	1.44
2	A	501	A1IWM	C61-N60	4.08	1.43	1.34

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	501	A1IWM	C23-C25-C26	6.22	127.81	112.62
2	A	501	A1IWM	O52-C47-C45	5.74	117.96	108.00
2	A	501	A1IWM	O30-C29-C31	5.00	118.86	106.44
2	A	501	A1IWM	O30-C25-C23	-4.18	100.74	108.00
2	A	501	A1IWM	O10-C08-O09	-4.14	116.28	124.13

There are no chirality outliers.

All (2) torsion outliers are listed below:

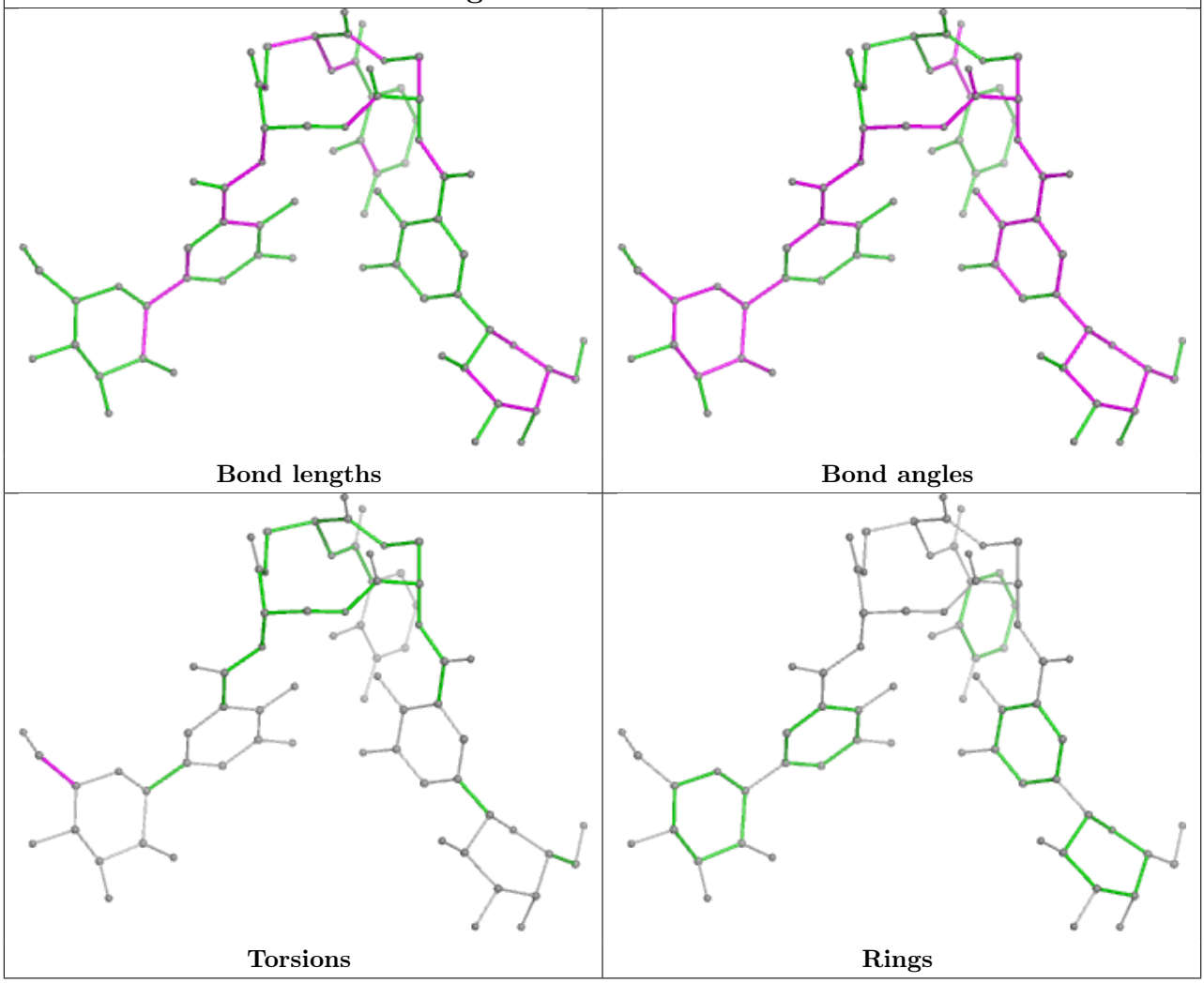
Mol	Chain	Res	Type	Atoms
2	A	501	A1IWM	C28-C29-C31-O32
2	A	501	A1IWM	O30-C29-C31-O32

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand A1IWM A 501



4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data

5.1 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, 95^{th} 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	426/472 (90%)	-0.06	10 (2%)	61	59	11, 24, 40, 53	1 (0%)
1	B	426/472 (90%)	-0.04	21 (4%)	36	33	16, 23, 44, 63	0
All	All	852/944 (90%)	-0.05	31 (3%)	46	44	11, 24, 42, 63	1 (0%)

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	418	VAL	5.9
1	B	417	GLY	4.0
1	B	415	ASN	3.6
1	B	411	TYR	3.2
1	A	388	ASN	3.2

5.2 Non-standard residues in protein, DNA, RNA chains

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

5.3 Carbohydrates

There are no oligosaccharides in this entry.

5.4 Ligands

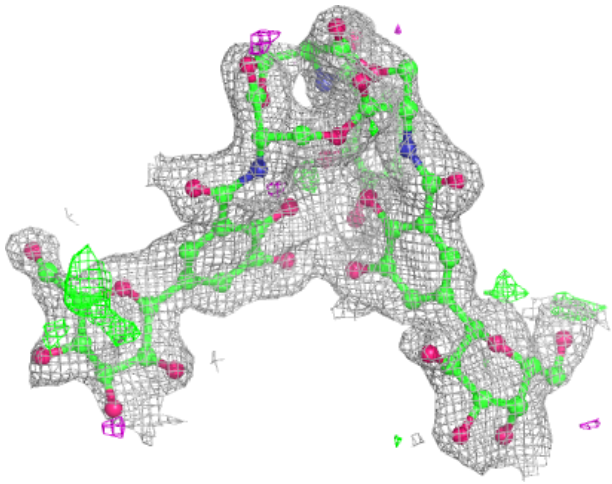
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

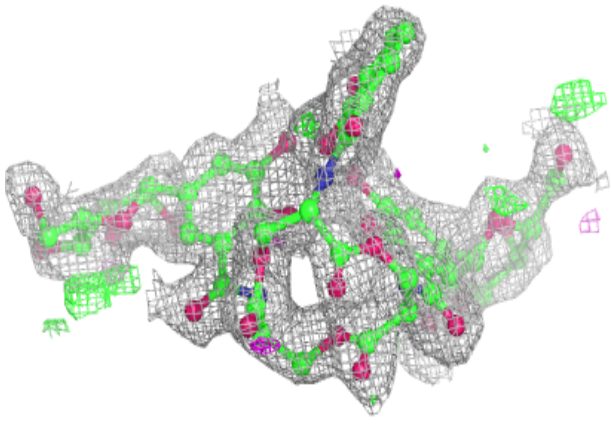
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	A1IWM	A	501	70/70	0.86	0.11	26,35,40,42	70
5	CL	A	505	1/1	0.93	0.10	40,40,40,40	0
4	CA	B	502	1/1	0.95	0.11	42,42,42,42	0
4	CA	B	504	1/1	0.96	0.11	49,49,49,49	0
4	CA	A	504	1/1	0.97	0.12	43,43,43,43	0
5	CL	B	506	1/1	0.97	0.07	47,47,47,47	0
4	CA	B	503	1/1	0.98	0.16	41,41,41,41	0
4	CA	B	501	1/1	0.98	0.07	27,27,27,27	0
4	CA	A	503	1/1	0.99	0.03	24,24,24,24	0
5	CL	A	506	1/1	0.99	0.03	25,25,25,25	0
5	CL	B	505	1/1	0.99	0.04	26,26,26,26	0
3	FE	A	502	1/1	0.99	0.03	27,27,27,27	0

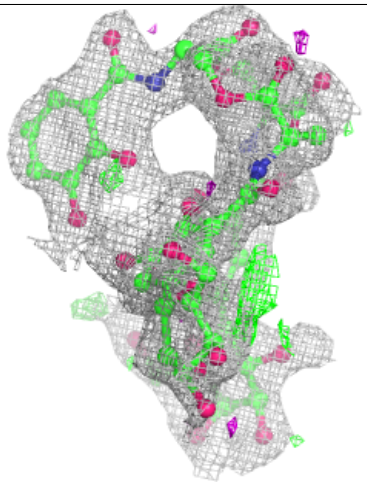
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around A1IWM A 501:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)







5.5 Other polymers (i)

There are no such residues in this entry.