Full wwPDB X-ray Structure Validation Report (i)

Sep 8, 2025 – 05:35 pm BST

PDB ID	:	9HQE / pdb_00009hqe
Title	:	Bacteroides fragilis xenosiderophore-binding lipoprotein XusB
Authors	:	Silale, A.; van den Berg, B.
Deposited on	:	2024-12-16
Resolution	:	1.77 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

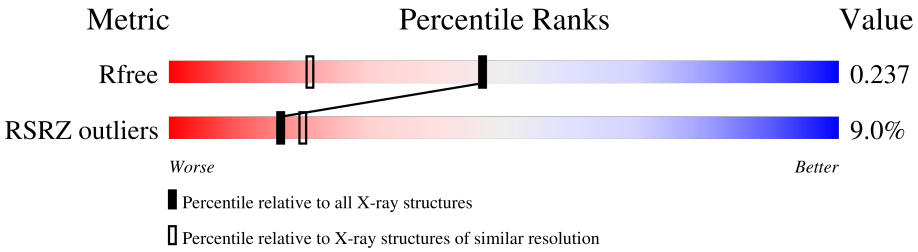
MolProbity	:	FAILED
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 1.77 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	1191 (1.78-1.78)
RSRZ outliers	164620	1191 (1.78-1.78)

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

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6081 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.

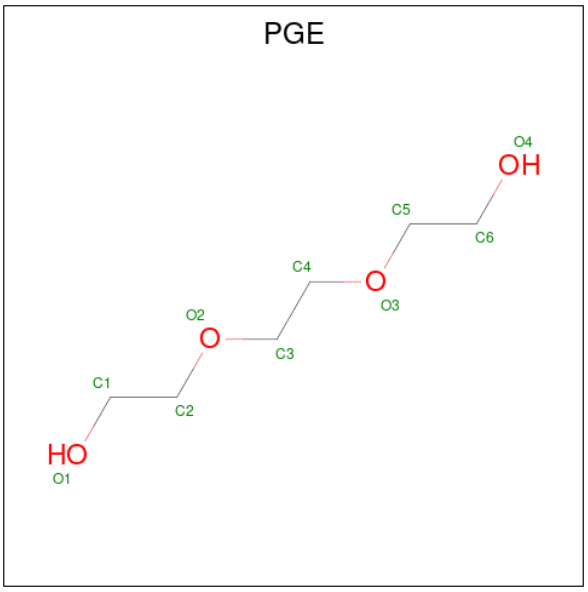
- Molecule 1 is a protein called XusB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	378	Total	C	N	O	S	0	0	0
			2908	1847	474	574	13			
1	B	365	Total	C	N	O	S	0	0	0
			2800	1775	457	555	13			

There are 16 discrepancies between the modelled and reference sequences:

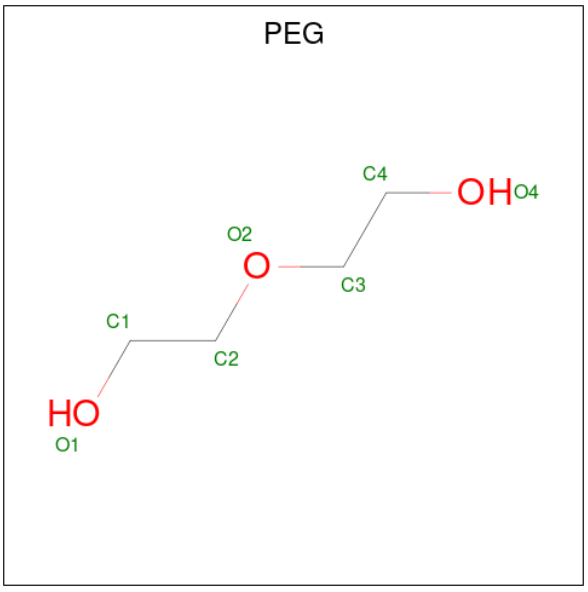
Chain	Residue	Modelled	Actual	Comment	Reference
A	407	LEU	-	expression tag	UNP Q5L7E3
A	408	GLU	-	expression tag	UNP Q5L7E3
A	409	HIS	-	expression tag	UNP Q5L7E3
A	410	HIS	-	expression tag	UNP Q5L7E3
A	411	HIS	-	expression tag	UNP Q5L7E3
A	412	HIS	-	expression tag	UNP Q5L7E3
A	413	HIS	-	expression tag	UNP Q5L7E3
A	414	HIS	-	expression tag	UNP Q5L7E3
B	407	LEU	-	expression tag	UNP Q5L7E3
B	408	GLU	-	expression tag	UNP Q5L7E3
B	409	HIS	-	expression tag	UNP Q5L7E3
B	410	HIS	-	expression tag	UNP Q5L7E3
B	411	HIS	-	expression tag	UNP Q5L7E3
B	412	HIS	-	expression tag	UNP Q5L7E3
B	413	HIS	-	expression tag	UNP Q5L7E3
B	414	HIS	-	expression tag	UNP Q5L7E3

- Molecule 2 is TRIETHYLENE GLYCOL (CCD ID: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			10	6	4		
2	B	1	Total	C	O	0	0
			10	6	4		
2	B	1	Total	C	O	0	0
			10	6	4		

- Molecule 3 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	C	O	0	0
			7	4	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C O 7 4 3	0	0

- Molecule 4 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Cl 1 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	217	Total O 217 217	0	0
5	B	111	Total O 111 111	0	0

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

3 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	50.99Å 86.92Å 154.39Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.42 – 1.77 48.42 – 1.77	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.42-1.77) 99.9 (48.42-1.77)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.06 (at 1.77Å)	Xtriage
Refinement program	PHENIX 1.21.1_5286	Depositor
R, R_{free}	0.199 , 0.237 0.199 , 0.237	Depositor DCC
R_{free} test set	3410 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	24.2	Xtriage
Anisotropy	0.317	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 41.8	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	0.96	EDS
Total number of atoms	6081	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.83% 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 6 ligands modelled in this entry, 1 is monoatomic - leaving 5 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	PGE	B	502	-	9,9,9	0.30	0	8,8,8	0.62	0
3	PEG	B	503	-	6,6,6	0.23	0	5,5,5	0.28	0
3	PEG	B	504	-	6,6,6	0.23	0	5,5,5	0.21	0
2	PGE	B	501	-	9,9,9	0.30	0	8,8,8	0.59	0
2	PGE	A	501	-	9,9,9	0.31	0	8,8,8	0.38	0

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	PGE	B	502	-	-	2/7/7/7	-
3	PEG	B	503	-	-	3/4/4/4	-
3	PEG	B	504	-	-	1/4/4/4	-
2	PGE	B	501	-	-	2/7/7/7	-
2	PGE	A	501	-	-	0/7/7/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	503	PEG	O1-C1-C2-O2
2	B	501	PGE	O2-C3-C4-O3
3	B	503	PEG	O2-C3-C4-O4
2	B	502	PGE	O1-C1-C2-O2
2	B	501	PGE	C4-C3-O2-C2
3	B	503	PEG	C1-C2-O2-C3
2	B	502	PGE	C4-C3-O2-C2
3	B	504	PEG	C1-C2-O2-C3

There are no ring outliers.

No monomer is involved in short contacts.

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	378/414 (91%)	-0.07	11 (2%) 54 60	16, 23, 39, 87	0
1	B	365/414 (88%)	0.86	56 (15%) 6 6	20, 38, 65, 97	0
All	All	743/828 (89%)	0.39	67 (9%) 17 20	16, 27, 61, 97	0

All (67) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	102	VAL	7.2
1	B	105	ALA	5.7
1	B	104	ASP	4.7
1	B	100	GLY	4.4
1	B	69	ALA	4.3
1	B	68	VAL	4.3
1	A	407	LEU	4.2
1	B	406	ASP	4.1
1	B	394	VAL	3.8
1	A	408	GLU	3.7
1	A	144	GLY	3.6
1	B	39	VAL	3.5
1	B	95	LEU	3.5
1	B	377	TYR	3.5
1	B	154	ILE	3.4
1	B	107	GLY	3.4
1	B	53	TYR	3.4
1	B	58	PRO	3.3
1	B	91	GLY	3.3
1	B	101	ILE	3.3
1	B	33	THR	3.2
1	B	116	VAL	3.2
1	B	159	ILE	3.1
1	A	115	PHE	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	143	SER	3.1
1	B	392	ALA	3.1
1	A	140	ASN	3.1
1	B	57	ALA	3.0
1	A	141	LYS	3.0
1	B	115	PHE	3.0
1	B	405	LEU	2.9
1	A	114	ASP	2.8
1	B	160	THR	2.8
1	B	373	SER	2.8
1	B	156	ASP	2.8
1	B	65	ILE	2.7
1	B	158	SER	2.7
1	B	155	SER	2.7
1	B	157	VAL	2.7
1	B	235	TRP	2.6
1	B	34	PRO	2.6
1	B	106	ASN	2.6
1	B	152	VAL	2.6
1	B	36	ALA	2.6
1	B	217	ASP	2.5
1	B	374	GLU	2.5
1	B	378	ILE	2.4
1	B	56	THR	2.4
1	B	40	LEU	2.4
1	B	54	VAL	2.4
1	B	86	THR	2.3
1	B	365	TYR	2.3
1	A	47	ASN	2.3
1	B	38	TYR	2.3
1	B	131	ASN	2.3
1	B	99	THR	2.3
1	B	103	ARG	2.2
1	B	391	GLY	2.2
1	B	64	THR	2.2
1	B	41	SER	2.2
1	A	142	GLU	2.2
1	B	375	GLY	2.2
1	B	55	VAL	2.2
1	A	145	ASP	2.1
1	B	401	GLY	2.1
1	B	153	ASN	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	60	LEU	2.0

5.2 Non-standard residues in protein, DNA, RNA chains (i)

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

5.3 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.4 Ligands (i)

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
3	PEG	B	504	7/7	0.79	0.13	36,40,47,49	0
2	PGE	B	502	10/10	0.81	0.14	39,42,45,46	0
3	PEG	B	503	7/7	0.85	0.11	37,39,43,43	0
2	PGE	B	501	10/10	0.85	0.11	32,37,39,47	0
2	PGE	A	501	10/10	0.88	0.11	28,31,37,42	0
4	CL	B	505	1/1	0.95	0.25	46,46,46,46	0

5.5 Other polymers (i)

There are no such residues in this entry.