



# Full wwPDB X-ray Structure Validation Report ⓘ

Apr 26, 2026 – 05:46 AM EDT

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](mailto: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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

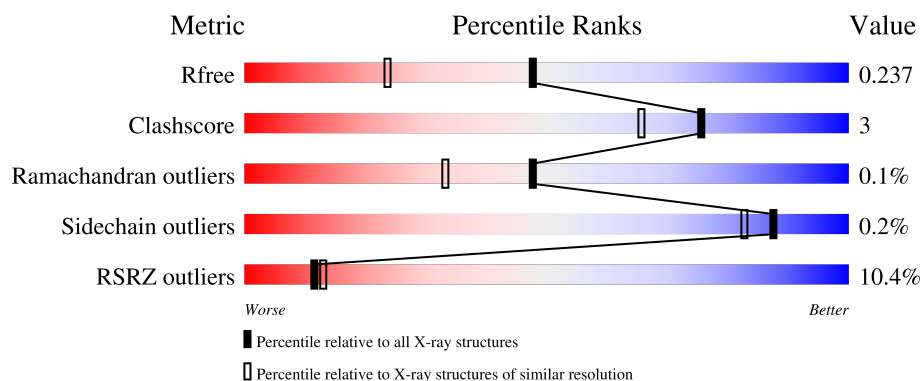
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 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))
$R_{free}$	180053	1365 (1.78-1.78)
Clashscore	190562	1395 (1.78-1.78)
Ramachandran outliers	187476	1382 (1.78-1.78)
Sidechain outliers	187428	1382 (1.78-1.78)
RSRZ outliers	180081	1365 (1.78-1.78)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	414	<div> <div>3%</div> <div>87%</div> <div>9%</div> </div>
1	B	414	<div> <div>16%</div> <div>80%</div> <div>8%</div> <div>12%</div> </div>

## 2 Entry composition

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.

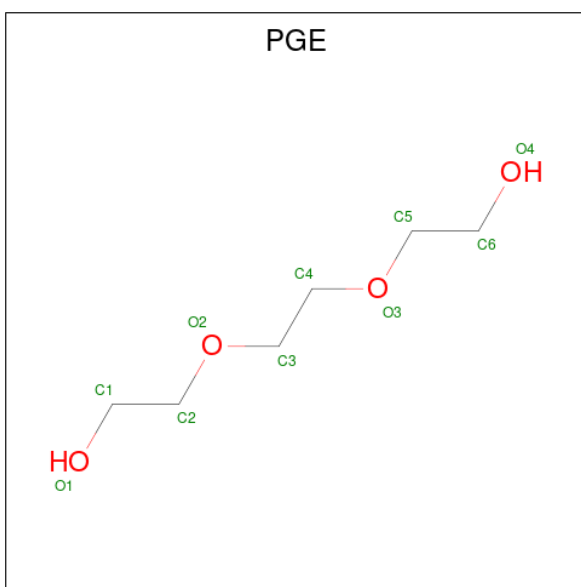
- Molecule 1 is a protein called XusB.

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:

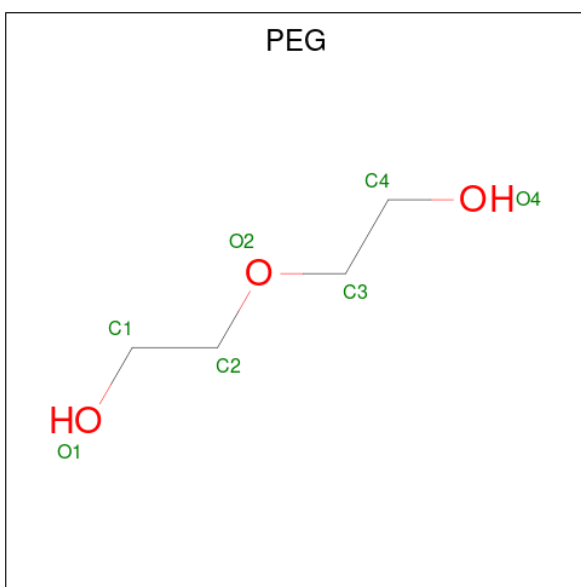
Chain	Residue	Modelled	Actual	Comment	Reference
A	407	LEU	-	expression tag	UNP Q5L7E3
A	408	GLU	-	expression tag	UNP Q5L7E3
A	409	HIS	-	expression tag	UNP Q5L7E3
A	410	HIS	-	expression tag	UNP Q5L7E3
A	411	HIS	-	expression tag	UNP Q5L7E3
A	412	HIS	-	expression tag	UNP Q5L7E3
A	413	HIS	-	expression tag	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<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).



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	0	0
			7	4	3		

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

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

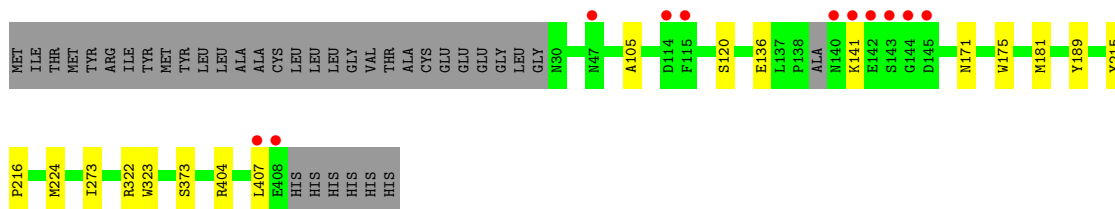
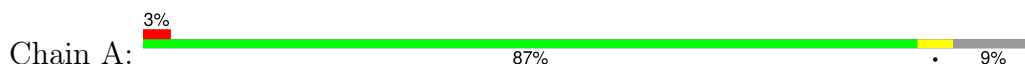
- Molecule 5 is water.

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

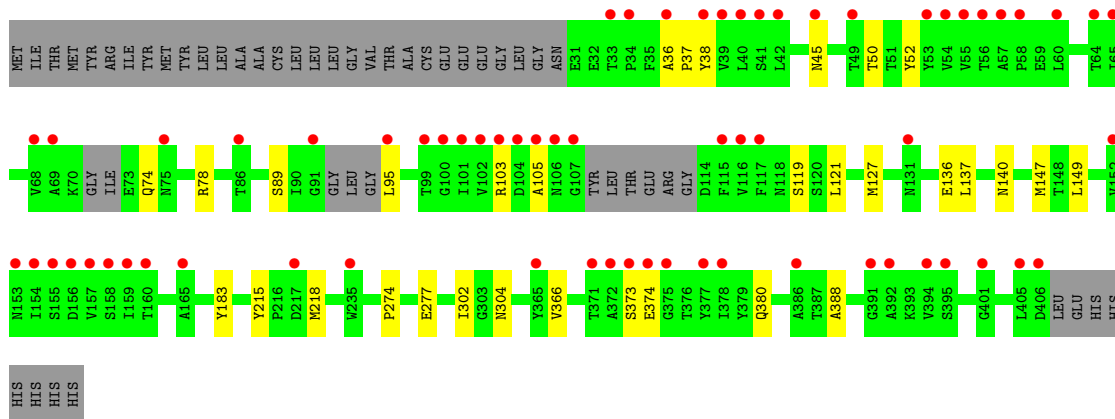
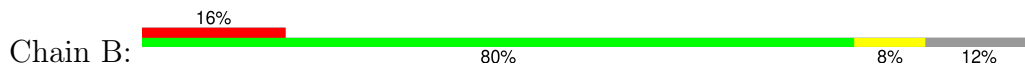
### 3 Residue-property plots [i](#)

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: XusB



#### • Molecule 1: XusB



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	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 <sup>2</sup> )	24.2	Xtriage
Anisotropy	0.317	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 41.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	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 <sup>2</sup> )	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.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: PGE, PEG, CL

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	0.30	0/2975	0.51	0/4044
1	B	0.23	0/2862	0.45	0/3889
All	All	0.27	0/5837	0.48	0/7933

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	2908	0	2792	10	0
1	B	2800	0	2682	25	0
2	A	10	0	14	0	0
2	B	20	0	28	0	0
3	B	14	0	20	2	0
4	B	1	0	0	0	0
5	A	217	0	0	0	0
5	B	111	0	0	0	0
All	All	6081	0	5536	34	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 3.

All (34) 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:404:ARG:HH12	1:B:304:ASN:HB3	1.55	0.71
1:B:149:LEU:HD13	1:B:218:MET:HE2	1.75	0.69
1:B:36:ALA:O	1:B:103:ARG:NH1	2.31	0.63
1:B:140:ASN:HD22	3:B:503:PEG:H31	1.65	0.61
1:B:37:PRO:HA	1:B:103:ARG:HH22	1.70	0.55
1:B:37:PRO:HA	1:B:103:ARG:NH2	2.22	0.55
1:A:404:ARG:NH1	1:B:304:ASN:HB3	2.21	0.54
1:B:127:MET:HE1	1:B:215:TYR:CE1	2.43	0.53
1:B:274:PRO:HB2	1:B:277:GLU:HG3	1.91	0.52
1:A:322:ARG:HG2	1:A:323:TRP:CD1	2.44	0.52
1:B:45:ASN:ND2	1:B:50:THR:OG1	2.44	0.51
1:A:105:ALA:HA	1:A:407:LEU:HD11	1.92	0.50
1:B:137:LEU:HD13	1:B:147:MET:HB3	1.93	0.49
1:B:119:SER:N	1:B:136:GLU:OE2	2.40	0.47
1:B:127:MET:HE3	1:B:183:TYR:CE1	2.49	0.47
1:B:52:TYR:O	1:B:74:GLN:N	2.35	0.46
1:B:121:LEU:HD12	1:B:121:LEU:HA	1.76	0.46
1:B:373:SER:O	1:B:374:GLU:HB2	2.16	0.45
1:A:224:MET:HE1	1:A:273:ILE:HG12	1.97	0.45
1:B:37:PRO:HG2	1:B:38:TYR:CE2	2.51	0.45
1:B:78:ARG:HD2	1:B:89:SER:OG	2.17	0.44
1:B:78:ARG:HH21	1:B:95:LEU:HD13	1.84	0.43
1:B:302:ILE:HG21	1:B:366:VAL:HG21	2.00	0.43
3:B:504:PEG:H12	3:B:504:PEG:H32	1.75	0.43
1:A:141:LYS:HG3	1:A:175:TRP:CE3	2.53	0.42
1:A:141:LYS:HE2	1:A:171:ASN:HD22	1.84	0.42
1:A:215:TYR:CG	1:A:216:PRO:HA	2.55	0.42
1:B:127:MET:HE1	1:B:215:TYR:CD1	2.55	0.41
1:B:127:MET:HE3	1:B:183:TYR:CZ	2.55	0.41
1:B:380:GLN:O	1:B:388:ALA:HA	2.20	0.40
1:A:120:SER:O	1:A:136:GLU:HG3	2.21	0.40
1:B:149:LEU:HB2	1:B:218:MET:CE	2.51	0.40
1:B:37:PRO:HG2	1:B:38:TYR:CD2	2.56	0.40
1:A:181:MET:HA	1:A:189:TYR:O	2.21	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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	374/414 (90%)	361 (96%)	13 (4%)	0	100	100
1	B	357/414 (86%)	340 (95%)	16 (4%)	1 (0%)	36	22
All	All	731/828 (88%)	701 (96%)	29 (4%)	1 (0%)	48	33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	105	ALA

### 5.3.2 Protein sidechains [i](#)

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	314/344 (91%)	313 (100%)	1 (0%)	86	80
1	B	303/344 (88%)	303 (100%)	0	100	100
All	All	617/688 (90%)	616 (100%)	1 (0%)	87	83

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	373	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	130	GLN
1	A	172	GLN
1	A	219	GLN
1	A	263	GLN
1	A	336	ASN
1	B	45	ASN
1	B	106	ASN
1	B	130	GLN
1	B	131	ASN
1	B	140	ASN
1	B	162	GLN
1	B	172	GLN
1	B	219	GLN
1	B	304	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

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	A	501	-	9,9,9	0.32	0	8,8,8	0.38	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	PEG	B	504	-	6,6,6	0.24	0	5,5,5	0.21	0
2	PGE	B	501	-	9,9,9	0.31	0	8,8,8	0.59	0
2	PGE	B	502	-	9,9,9	0.31	0	8,8,8	0.62	0
3	PEG	B	503	-	6,6,6	0.24	0	5,5,5	0.27	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	A	501	-	-	0/7/7/7	-
3	PEG	B	504	-	-	1/4/4/4	-
2	PGE	B	501	-	-	2/7/7/7	-
2	PGE	B	502	-	-	2/7/7/7	-
3	PEG	B	503	-	-	3/4/4/4	-

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	501	PGE	C4-C3-O2-C2
3	B	503	PEG	C1-C2-O2-C3
2	B	502	PGE	O1-C1-C2-O2
2	B	502	PGE	C4-C3-O2-C2
3	B	504	PEG	C1-C2-O2-C3

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	504	PEG	1	0
3	B	503	PEG	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	378/414 (91%)	0.01	11 (2%) 53 61	16, 23, 39, 87	0
1	B	365/414 (88%)	0.94	66 (18%) 3 3	20, 38, 65, 97	0
All	All	743/828 (89%)	0.47	77 (10%) 11 13	16, 27, 61, 97	0

All (77) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	102	VAL	7.5
1	B	105	ALA	6.0
1	B	104	ASP	4.8
1	B	100	GLY	4.5
1	B	69	ALA	4.5
1	B	68	VAL	4.4
1	A	407	LEU	4.3
1	B	406	ASP	4.2
1	B	394	VAL	4.0
1	A	408	GLU	3.8
1	A	144	GLY	3.8
1	B	39	VAL	3.7
1	B	95	LEU	3.6
1	B	107	GLY	3.6
1	B	377	TYR	3.6
1	B	154	ILE	3.6
1	B	53	TYR	3.5
1	B	91	GLY	3.5
1	B	58	PRO	3.5
1	B	101	ILE	3.4
1	B	33	THR	3.4
1	B	116	VAL	3.4
1	B	159	ILE	3.3
1	B	392	ALA	3.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	115	PHE	3.2
1	A	143	SER	3.2
1	A	140	ASN	3.2
1	B	57	ALA	3.2
1	B	115	PHE	3.1
1	A	141	LYS	3.0
1	B	405	LEU	3.0
1	B	160	THR	2.9
1	A	114	ASP	2.9
1	B	373	SER	2.9
1	B	65	ILE	2.9
1	B	156	ASP	2.9
1	B	157	VAL	2.8
1	B	158	SER	2.8
1	B	155	SER	2.8
1	B	235	TRP	2.7
1	B	34	PRO	2.7
1	B	152	VAL	2.7
1	B	106	ASN	2.7
1	B	36	ALA	2.7
1	B	217	ASP	2.6
1	B	374	GLU	2.6
1	B	378	ILE	2.6
1	B	56	THR	2.5
1	B	40	LEU	2.5
1	B	54	VAL	2.5
1	B	86	THR	2.5
1	B	365	TYR	2.4
1	A	47	ASN	2.4
1	B	38	TYR	2.4
1	B	391	GLY	2.4
1	B	131	ASN	2.4
1	B	99	THR	2.4
1	B	64	THR	2.3
1	B	41	SER	2.3
1	B	375	GLY	2.3
1	B	103	ARG	2.3
1	A	142	GLU	2.3
1	B	55	VAL	2.3
1	B	401	GLY	2.3
1	A	145	ASP	2.2
1	B	153	ASN	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	60	LEU	2.1
1	B	386	ALA	2.1
1	B	371	THR	2.1
1	B	372	ALA	2.1
1	B	75	ASN	2.1
1	B	395	SER	2.1
1	B	49	THR	2.1
1	B	42	LEU	2.0
1	B	45	ASN	2.0
1	B	117	PHE	2.0
1	B	165	ALA	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 6.4 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<sup>th</sup> 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 <sup>2</sup> )	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

## 6.5 Other polymers [i](#)

There are no such residues in this entry.