



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 12:27 PM GMT

PDB ID : 3VXC  
Title : Crystal Structure of Xylobiose-BxlE complex from *Streptomyces thermoviolaceus* OPC-520  
Authors : Tomoo, K.; Ishida, T.; Miyamoto, K.; Tsujibo, H.  
Deposited on : 2012-09-11  
Resolution : 1.85 Å(reported)

This is a full wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

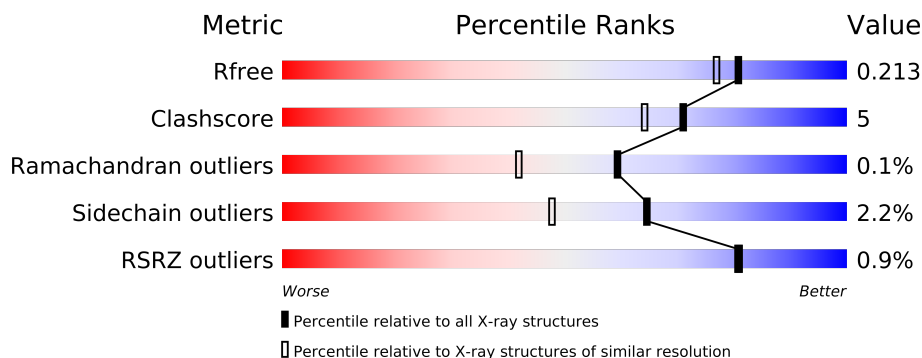
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 1.85 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))
$R_{free}$	66092	1269 (1.86-1.86)
Clashscore	79885	1470 (1.86-1.86)
Ramachandran outliers	78287	1451 (1.86-1.86)
Sidechain outliers	78261	1451 (1.86-1.86)
RSRZ outliers	66119	1269 (1.86-1.86)

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  $\geq 3$ , 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	436	
1	B	436	

## 2 Entry composition i

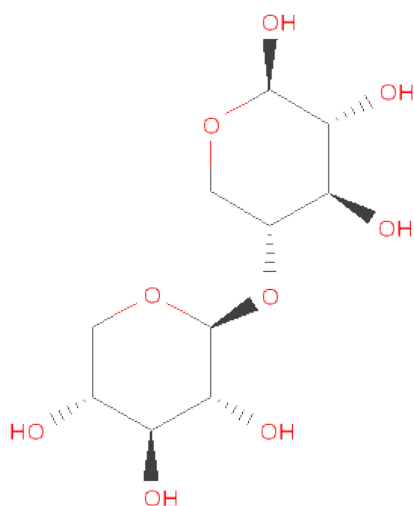
There are 3 unique types of molecules in this entry. The entry contains 6578 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Putative sugar-binding lipoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	398	Total	C	N	O	S	0	0	0
			3037	1942	506	576	13			
1	B	398	Total	C	N	O	S	0	0	0
			3037	1942	506	576	13			

- Molecule 2 is SUGAR (4-O-BETA-D-XYLOPYRANOSYL-BETA-D-XYLOPYRANOSE) (three-letter code: BXP) (formula: C<sub>10</sub>H<sub>18</sub>O<sub>9</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			19	10	9		
2	B	1	Total	C	O	0	0
			19	10	9		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	251	Total 251	O 251	0	0
3	B	215	Total 215	O 215	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	103.67Å 163.16Å 42.93Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.81 – 1.85 33.81 – 1.85	Depositor EDS
% Data completeness (in resolution range)	98.5 (33.81-1.85) 98.5 (33.81-1.85)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	12.27 (at 1.85Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.168 , 0.213 0.169 , 0.213	Depositor DCC
$R_{free}$ test set	3111 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	14.3	Xtriage
Anisotropy	0.036	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 31.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 62285 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6578	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.27% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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

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	1.13	4/3114 (0.1%)	0.90	1/4227 (0.0%)
1	B	1.02	2/3114 (0.1%)	0.86	2/4227 (0.0%)
All	All	1.07	6/6228 (0.1%)	0.88	3/8454 (0.0%)

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	353	TYR	CD2-CE2	6.15	1.48	1.39
1	B	306	CYS	CB-SG	-5.71	1.72	1.81
1	A	208	MET	CG-SD	5.56	1.95	1.81
1	A	390	GLN	CG-CD	5.49	1.63	1.51
1	A	100	PHE	CE2-CZ	5.21	1.47	1.37
1	B	161	PHE	CE1-CZ	5.10	1.47	1.37

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	334	ARG	NE-CZ-NH1	7.10	123.85	120.30
1	B	81	ASP	CB-CG-OD1	6.11	123.80	118.30
1	A	391	LYS	CD-CE-NZ	5.61	124.61	111.70

There are no chirality outliers.

There are no planarity outliers.

### 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	3037	0	2978	26	0
1	B	3037	0	2978	38	0
2	A	19	0	18	0	0
2	B	19	0	18	0	0
3	A	251	0	0	4	0
3	B	215	0	0	4	0
All	All	6578	0	5992	55	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 5.

All (55) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:95:ALA:HB3	1:B:334:ARG:NH2	1.61	1.15
1:B:95:ALA:HB3	1:B:334:ARG:HH21	1.18	1.00
1:B:95:ALA:CB	1:B:334:ARG:NH2	2.27	0.98
1:A:334:ARG:HD2	1:B:334:ARG:NH1	1.83	0.93
1:B:367:THR:HG22	3:B:751:HOH:O	1.69	0.93
1:A:334:ARG:NH2	1:B:334:ARG:HD3	1.90	0.87
1:B:101:ASN:HD22	1:B:103:GLY:H	1.23	0.86
1:A:101:ASN:HD22	1:A:103:GLY:H	1.23	0.86
1:A:334:ARG:HD2	1:B:334:ARG:HH11	1.43	0.84
1:A:122:VAL:HG23	3:A:877:HOH:O	1.79	0.82
1:A:38:THR:HG22	1:A:71:LYS:H	1.45	0.80
1:B:305:TRP:HZ2	1:B:380:GLU:HG2	1.47	0.79
1:A:390:GLN:HG2	3:A:771:HOH:O	1.83	0.78
1:B:95:ALA:CB	1:B:334:ARG:HH22	1.99	0.73
1:B:165:SER:O	1:B:169:GLU:HG3	1.91	0.71
1:A:334:ARG:NH2	1:B:334:ARG:CD	2.55	0.69
1:A:118:ASP:OD2	1:A:146:HIS:HD2	1.76	0.69
1:B:118:ASP:OD2	1:B:146:HIS:HD2	1.79	0.65
1:B:95:ALA:HB2	1:B:334:ARG:HH22	1.62	0.64
1:B:224:ASP:HB3	3:B:744:HOH:O	1.98	0.63
1:B:387:GLN:HE21	1:B:391:LYS:NZ	1.99	0.60
1:A:297:ASP:O	1:A:301:LYS:HD2	2.01	0.60
1:B:334:ARG:HG3	3:B:803:HOH:O	2.01	0.59
1:A:334:ARG:CD	1:B:334:ARG:NH1	2.63	0.59
1:B:404:ASP:OD2	1:B:406:ASP:OD2	2.22	0.58
1:A:334:ARG:CD	1:B:334:ARG:HH11	2.15	0.56

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:124:LYS:HE2	3:A:878:HOH:O	2.06	0.56
1:A:194:THR:O	1:A:278:ALA:HA	2.10	0.52
1:B:387:GLN:HG3	1:B:391:LYS:HZ2	1.75	0.51
1:A:248:GLN:O	1:A:252:GLU:HG2	2.10	0.51
1:A:391:LYS:HE2	3:A:808:HOH:O	2.10	0.51
1:B:170:HIS:CE1	1:B:191:ALA:HB1	2.47	0.50
1:B:387:GLN:HE21	1:B:391:LYS:HZ1	1.59	0.50
1:A:406:ASP:HB3	1:B:112:GLU:OE2	2.11	0.50
1:B:107:ILE:HG13	1:B:143:LEU:HD11	1.93	0.49
1:B:404:ASP:HB2	3:B:748:HOH:O	2.13	0.49
1:A:156:GLN:NE2	1:A:324:ASN:HD22	2.11	0.48
1:A:126:ASP:OD2	1:A:346:ARG:NH1	2.45	0.48
1:B:40:THR:HG22	1:B:73:LYS:HE3	1.96	0.48
1:A:334:ARG:HH22	1:B:334:ARG:HD3	1.77	0.47
1:A:118:ASP:OD2	1:A:146:HIS:CD2	2.63	0.47
1:B:118:ASP:OD2	1:B:146:HIS:CD2	2.65	0.47
1:A:335:THR:CG2	1:A:338:LYS:HG2	2.45	0.46
1:B:305:TRP:CZ2	1:B:380:GLU:HG2	2.39	0.46
1:B:338:LYS:O	1:B:342:ILE:HG12	2.19	0.43
1:B:101:ASN:ND2	1:B:103:GLY:H	2.04	0.43
1:B:156:GLN:NE2	1:B:324:ASN:HD22	2.18	0.42
1:B:50:LYS:N	1:B:50:LYS:HD3	2.34	0.42
1:B:230:ASP:OD1	1:B:232:SER:HB2	2.20	0.42
1:A:38:THR:CG2	1:A:70:ALA:HA	2.50	0.41
1:B:50:LYS:HD3	1:B:50:LYS:H	1.85	0.41
1:A:410:PRO:HB2	1:A:435:LEU:HD21	2.02	0.41
1:A:334:ARG:HH22	1:B:334:ARG:CD	2.33	0.41
1:A:38:THR:HG21	1:A:70:ALA:HA	2.02	0.40
1:B:153:ARG:HD2	1:B:327:ASN:HD22	1.86	0.40

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	396/436 (91%)	388 (98%)	8 (2%)	0	100	100
1	B	396/436 (91%)	387 (98%)	8 (2%)	1 (0%)	50	32
All	All	792/872 (91%)	775 (98%)	16 (2%)	1 (0%)	59	42

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	203	ILE

### 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	317/340 (93%)	308 (97%)	9 (3%)	56	37
1	B	317/340 (93%)	312 (98%)	5 (2%)	75	62
All	All	634/680 (93%)	620 (98%)	14 (2%)	64	48

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	38	THR
1	A	66	LYS
1	A	85	LYS
1	A	179	ASP
1	A	225	LYS
1	A	375	SER
1	A	391	LYS
1	A	395	PHE
1	A	406	ASP
1	B	50	LYS
1	B	67	LYS
1	B	171	LYS
1	B	189	LYS
1	B	395	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	101	ASN
1	A	146	HIS
1	A	156	GLN
1	A	248	GLN
1	A	327	ASN
1	A	385	GLN
1	B	69	ASN
1	B	94	ASN
1	B	101	ASN
1	B	146	HIS
1	B	156	GLN
1	B	170	HIS
1	B	327	ASN
1	B	371	ASN
1	B	385	GLN
1	B	387	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 ⓘ

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
2	BXP	A	601	-	20,20,20	1.39	2 (10%)	29,29,29	1.40	4 (13%)
2	BXP	B	501	-	20,20,20	1.49	3 (15%)	29,29,29	1.49	5 (17%)

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	BXP	A	601	-	-	0/4/38/38	0/2/2/2
2	BXP	B	501	-	-	0/4/38/38	0/2/2/2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	601	BXP	O5A-C1A	4.44	1.49	1.43
2	B	501	BXP	O5A-C5A	-3.35	1.38	1.43
2	B	501	BXP	C5A-C4A	3.24	1.57	1.51
2	B	501	BXP	O5A-C1A	2.89	1.47	1.43
2	A	601	BXP	C5A-C4A	2.62	1.56	1.51

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	601	BXP	O5B-C5B-C4B	-3.48	105.28	110.83
2	B	501	BXP	O4A-C1B-C2B	-3.38	100.03	108.12
2	B	501	BXP	C1A-C2A-C3A	-2.95	105.86	110.53
2	B	501	BXP	O5A-C1A-C2A	-2.94	105.03	109.24
2	A	601	BXP	O4A-C1B-C2B	-2.84	101.33	108.12
2	B	501	BXP	O4A-C1B-O5B	2.54	114.53	109.12
2	A	601	BXP	C1A-C2A-C3A	-2.42	106.70	110.53
2	B	501	BXP	O5B-C5B-C4B	-2.33	107.12	110.83
2	A	601	BXP	O1A-C1A-O5A	-2.14	104.03	109.88

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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, 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	398/436 (91%)	-0.24	3 (0%) 83 83	5, 12, 25, 41	0
1	B	398/436 (91%)	-0.16	4 (1%) 79 78	6, 15, 31, 49	0
All	All	796/872 (91%)	-0.20	7 (0%) 81 81	5, 14, 28, 49	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	67	LYS	3.7
1	A	67	LYS	3.4
1	A	69	ASN	2.8
1	B	171	LYS	2.8
1	A	336	GLY	2.5
1	B	64	GLU	2.3
1	B	68	ALA	2.0

### 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

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	BXP	B	501	19/19	0.12	0.63	5,6,9,10	0
2	BXP	A	601	19/19	0.10	0.06	3,5,8,11	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.