



# wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 09:44 PM GMT

PDB ID : 3B8O  
Title : Structure of WzzE- Bacterial Polysaccharide Co-polymerase  
Authors : Tocilj, A.; Matte, A.; Cygler, M.  
Deposited on : 2007-11-01  
Resolution : 2.40 Å(reported)

This is a wwPDB validation summary 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>

---

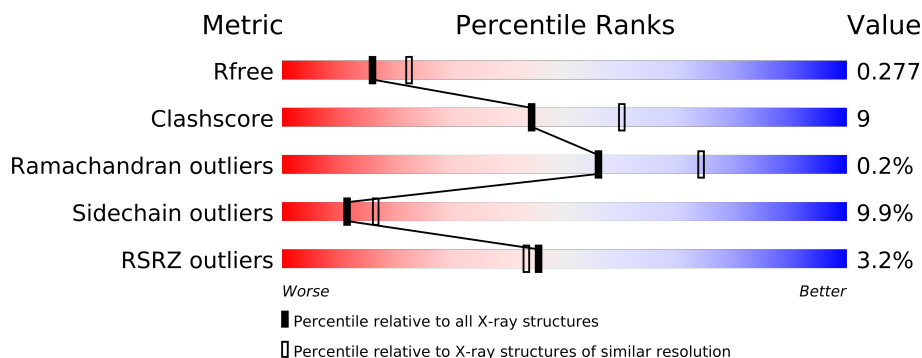
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.40 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	2207 (2.40-2.40)
Clashscore	79885	2789 (2.40-2.40)
Ramachandran outliers	78287	2736 (2.40-2.40)
Sidechain outliers	78261	2737 (2.40-2.40)
RSRZ outliers	66119	2210 (2.40-2.40)

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	265	
1	B	265	
1	C	265	
1	D	265	
1	E	265	
1	F	265	
1	G	265	
1	H	265	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 14771 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 Lipopolysaccharide biosynthesis protein wzzE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	221	Total	C	N	O	Se	0	0	0
			1797	1124	321	344	8			
1	B	221	Total	C	N	O	Se	0	0	0
			1797	1124	321	344	8			
1	C	221	Total	C	N	O	Se	0	0	0
			1797	1124	321	344	8			
1	D	221	Total	C	N	O	Se	0	0	0
			1797	1124	321	344	8			
1	E	221	Total	C	N	O	Se	0	0	0
			1797	1124	321	344	8			
1	F	221	Total	C	N	O	Se	0	0	0
			1797	1124	321	344	8			
1	G	221	Total	C	N	O	Se	0	0	0
			1797	1124	321	344	8			
1	H	221	Total	C	N	O	Se	0	0	0
			1797	1124	321	344	8			

- Molecule 2 is water.

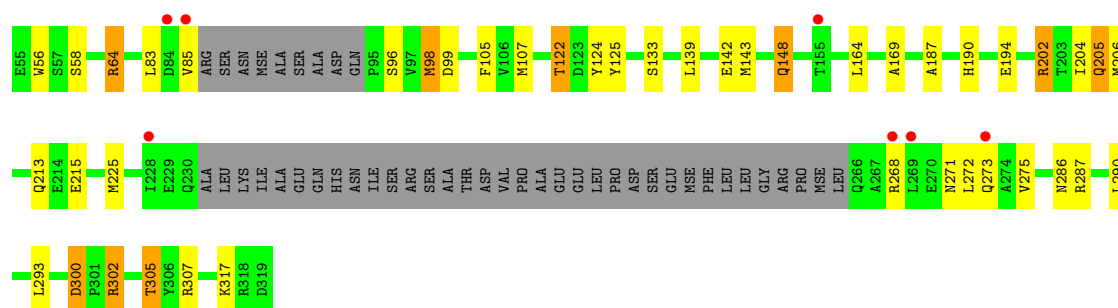
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	46	Total	O	0	0
			46	46		
2	B	48	Total	O	0	0
			48	48		
2	C	52	Total	O	0	0
			52	52		
2	D	48	Total	O	0	0
			48	48		
2	E	48	Total	O	0	0
			48	48		
2	F	58	Total	O	0	0
			58	58		

*Continued on next page...*

*Continued from previous page...*

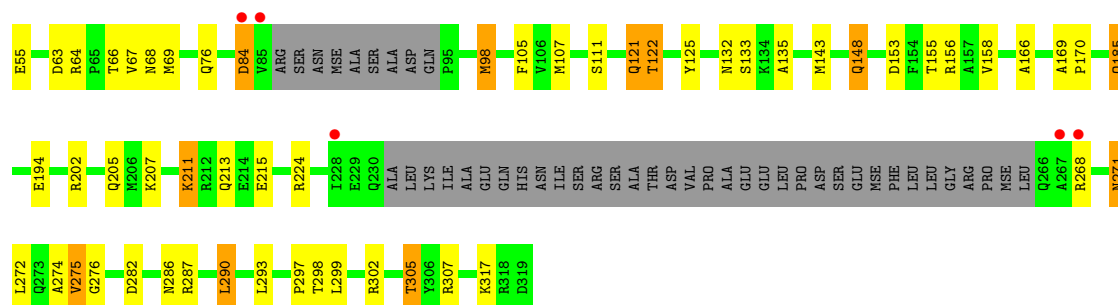
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	45	Total	O	0	0
			45	45		
2	H	50	Total	O	0	0
			50	50		





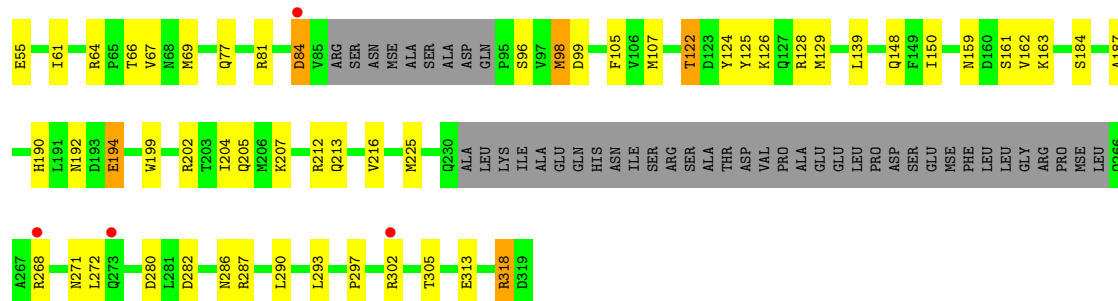
- Molecule 1: Lipopolysaccharide biosynthesis protein wzzE

Chain E:



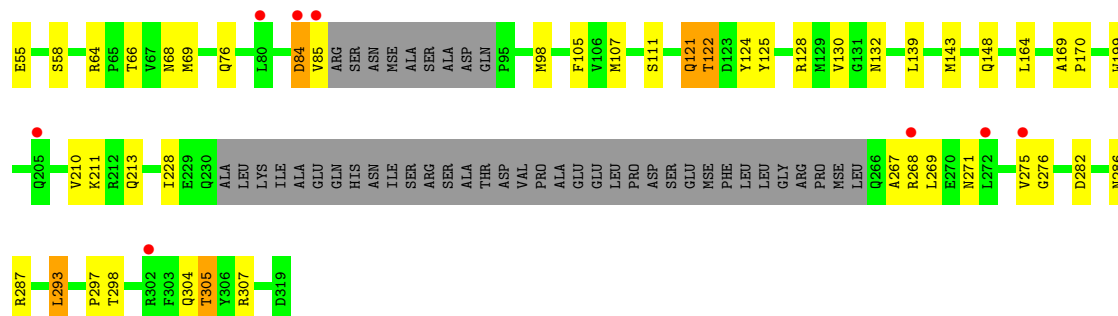
- Molecule 1: Lipopolysaccharide biosynthesis protein wzzE

Chain F:



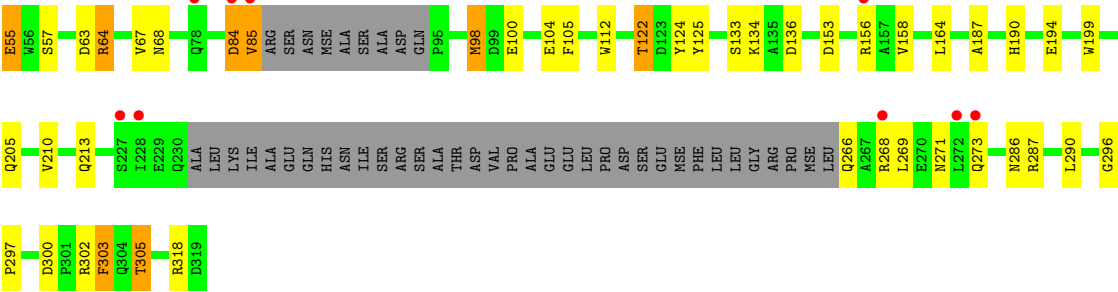
- Molecule 1: Lipopolysaccharide biosynthesis protein wzzE

Chain G:



- Molecule 1: Lipopolysaccharide biosynthesis protein wzzE

Chain H: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.44Å 108.30Å 122.57Å 90.00° 102.15° 90.00°	Depositor
Resolution (Å)	47.89 – 2.40 46.23 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.1 (47.89-2.40) 99.1 (46.23-2.40)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.78 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.223 , 0.266 0.236 , 0.277	Depositor DCC
$R_{free}$ test set	5290 reflections (5.23%)	DCC
Wilson B-factor (Å <sup>2</sup> )	37.2	Xtriage
Anisotropy	0.109	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 41.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	8 of 106363 reflections (0.008%)	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	14771	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.65% 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 ⓘ

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	0.69	0/1822	0.73	1/2451 (0.0%)
1	B	0.71	0/1822	0.77	1/2451 (0.0%)
1	C	0.73	1/1822 (0.1%)	0.77	0/2451
1	D	0.71	0/1822	0.76	1/2451 (0.0%)
1	E	0.79	0/1822	0.81	1/2451 (0.0%)
1	F	0.76	0/1822	0.75	0/2451
1	G	0.74	0/1822	0.78	2/2451 (0.1%)
1	H	0.67	0/1822	0.71	0/2451
All	All	0.73	1/14576 (0.0%)	0.76	6/19608 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	266	GLN	CD-NE2	5.16	1.45	1.32

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	113	ASP	CB-CG-OD1	5.48	123.23	118.30
1	G	121	GLN	CA-CB-CG	5.45	125.39	113.40
1	B	83	LEU	CA-CB-CG	5.39	127.70	115.30
1	D	307	ARG	NE-CZ-NH2	-5.11	117.75	120.30
1	G	293	LEU	CA-CB-CG	5.07	126.97	115.30

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	1797	0	1748	42	0
1	B	1797	0	1748	29	0
1	C	1797	0	1748	30	0
1	D	1797	0	1748	27	0
1	E	1797	0	1748	43	0
1	F	1797	0	1748	38	0
1	G	1797	0	1748	26	0
1	H	1797	0	1748	35	0
2	A	46	0	0	2	0
2	B	48	0	0	1	0
2	C	52	0	0	4	0
2	D	48	0	0	0	0
2	E	48	0	0	3	0
2	F	58	0	0	1	0
2	G	45	0	0	1	0
2	H	50	0	0	1	0
All	All	14771	0	13984	245	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 9.

The worst 5 of 245 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:105:PHE:HB2	1:A:305:THR:HG23	1.26	1.13
1:C:98:MSE:HG3	2:C:351:HOH:O	1.51	1.10
1:C:105:PHE:HB2	1:C:305:THR:HG23	1.29	1.09
1:A:65:PRO:HA	1:A:69:MSE:CE	1.82	1.09
1:E:105:PHE:HB2	1:E:305:THR:HG23	1.39	1.04

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	215/265 (81%)	208 (97%)	7 (3%)	0	100	100
1	B	215/265 (81%)	212 (99%)	3 (1%)	0	100	100
1	C	215/265 (81%)	212 (99%)	3 (1%)	0	100	100
1	D	215/265 (81%)	210 (98%)	5 (2%)	0	100	100
1	E	215/265 (81%)	210 (98%)	4 (2%)	1 (0%)	38	53
1	F	215/265 (81%)	209 (97%)	5 (2%)	1 (0%)	38	53
1	G	215/265 (81%)	211 (98%)	4 (2%)	0	100	100
1	H	215/265 (81%)	212 (99%)	2 (1%)	1 (0%)	38	53
All	All	1720/2120 (81%)	1684 (98%)	33 (2%)	3 (0%)	56	74

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	275	VAL
1	H	303	PHE
1	F	318	ARG

### 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	190/216 (88%)	169 (89%)	21 (11%)	9	12
1	B	190/216 (88%)	174 (92%)	16 (8%)	16	23
1	C	190/216 (88%)	172 (90%)	18 (10%)	12	18
1	D	190/216 (88%)	170 (90%)	20 (10%)	10	14
1	E	190/216 (88%)	163 (86%)	27 (14%)	5	5
1	F	190/216 (88%)	175 (92%)	15 (8%)	18	26
1	G	190/216 (88%)	172 (90%)	18 (10%)	12	18
1	H	190/216 (88%)	174 (92%)	16 (8%)	16	23
All	All	1520/1728 (88%)	1369 (90%)	151 (10%)	11	16

5 of 151 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	287	ARG
1	E	155	THR
1	H	85	VAL
1	D	293	LEU
1	E	68	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 81 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	213	GLN
1	E	192	ASN
1	H	82	ASN
1	D	294	ASN
1	E	132	ASN

### 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 ⓘ

There are no ligands in this entry.

### 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	221/265 (83%)	0.13	6 (2%) 52 49	21, 30, 38, 44	0
1	B	221/265 (83%)	-0.02	10 (4%) 32 30	21, 30, 39, 44	0
1	C	221/265 (83%)	0.18	7 (3%) 45 43	22, 30, 38, 44	0
1	D	221/265 (83%)	0.11	7 (3%) 45 43	22, 30, 39, 44	0
1	E	221/265 (83%)	0.13	5 (2%) 57 55	22, 30, 39, 44	0
1	F	221/265 (83%)	-0.00	4 (1%) 65 63	21, 30, 38, 44	0
1	G	221/265 (83%)	0.21	8 (3%) 41 39	21, 30, 38, 44	0
1	H	221/265 (83%)	0.13	9 (4%) 35 33	21, 30, 38, 44	0
All	All	1768/2120 (83%)	0.11	56 (3%) 45 43	21, 30, 39, 44	0

The worst 5 of 56 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	228	ILE	6.2
1	H	84	ASP	4.9
1	C	154	PHE	4.2
1	F	268	ARG	4.1
1	D	228	ILE	4.1

### 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 ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.