

# Full wwPDB X-ray Structure Validation Report (i)

#### Jun 12, 2024 – 07:05 PM EDT

PDB ID 3IRS

> Title CRYSTAL STRUCTURE OF UNCHARACTERIZED TIM-BARREL PRO-

> > TEIN BB4693 FROM Bordetella bronchiseptica

Authors: Patskovsky, Y.; Malashkevich, V.; Toro, R.; Foti, R.; Dickey, M.; Do, J.;

Sauder, J.M.; Raushel, F.M.; Burley, S.K.; Almo, S.C.; New York SGX Re-

search Center for Structural Genomics (NYSGXRC)

Deposited on 2009-08-24

Resolution 1.76 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 (1)) were used in the production of this report:

MolProbity 4.02b-467

> 2022.3.0, CSD as543be (2022) Mogul

Xtriage (Phenix) 1.20.1 EDS 2.36.2

Percentile statistics 20191225.v01 (using entries in the PDB archive December 25th 2019)

> Refmac 5.8.0158

CCP4 7.0.044 (Gargrove) Engh & Huber (2001)

Ideal geometry (proteins) Ideal geometry (DNA, RNA) Parkinson et al. (1996)

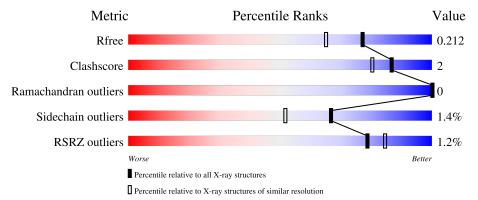
Validation Pipeline (wwPDB-VP) : 2.36.2

### 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 1.76 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	Similar resolution
Metric	$(\#  ext{Entries})$	$(\#  ext{Entries},  ext{ resolution range}( ext{Å}))$
$R_{free}$	130704	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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 >=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 <=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	291	92%	•	-
1	В	291	89%	7%	-
1	С	291	91%	6%	-



## 2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 7377 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 UNCHARACTERIZED PROTEIN BB4693.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	۸	281	Total	С	N	О	S	0	7	0
1	A	201	2253	1443	388	405	17	0	'	
1	D	281	Total	С	N	О	S	0	12	0
1	Б	201	2274	1456	392	409	17	0	12	
1	С	281	Total	С	N	О	S	0	10	0
1		201	2265	1451	388	409	17		10	

There are 33 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	expression tag	UNP Q7WEE2
A	0	SER	-	expression tag	UNP Q7WEE2
A	1	LEU	-	expression tag	UNP Q7WEE2
A	282	GLU	-	expression tag	UNP Q7WEE2
A	283	GLY	-	expression tag	UNP Q7WEE2
A	284	HIS	-	expression tag	UNP Q7WEE2
A	285	HIS	-	expression tag	UNP Q7WEE2
A	286	HIS	-	expression tag	UNP Q7WEE2
A	287	HIS	-	expression tag	UNP Q7WEE2
A	288	HIS	-	expression tag	UNP Q7WEE2
A	289	HIS	-	expression tag	UNP Q7WEE2
В	-1	MET	-	expression tag	UNP Q7WEE2
В	0	SER	-	expression tag	UNP Q7WEE2
В	1	LEU	-	expression tag	UNP Q7WEE2
В	282	GLU	-	expression tag	UNP Q7WEE2
В	283	GLY	-	expression tag	UNP Q7WEE2
В	284	HIS	-	expression tag	UNP Q7WEE2
В	285	HIS	-	expression tag	UNP Q7WEE2
В	286	HIS	-	expression tag	UNP Q7WEE2
В	287	HIS	-	expression tag	UNP Q7WEE2
В	288	HIS	-	expression tag	UNP Q7WEE2
В	289	HIS	-	expression tag	UNP Q7WEE2
С	-1	MET	-	expression tag	UNP Q7WEE2

Continued on next page...



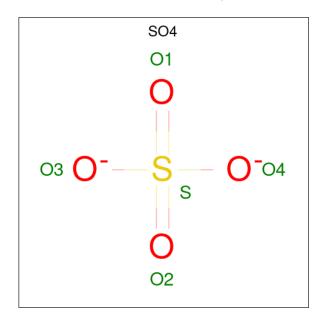
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
С	0	SER	-	expression tag	UNP Q7WEE2
С	1	LEU	-	expression tag	UNP Q7WEE2
С	282	GLU	-	expression tag	UNP Q7WEE2
С	283	GLY	-	expression tag	UNP Q7WEE2
С	284	HIS	-	expression tag	UNP Q7WEE2
С	285	HIS	-	expression tag	UNP Q7WEE2
С	286	HIS	-	expression tag	UNP Q7WEE2
С	287	HIS	-	expression tag	UNP Q7WEE2
С	288	HIS	-	expression tag	UNP Q7WEE2
С	289	HIS	-	expression tag	UNP Q7WEE2

• Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Na 1 1	0	0

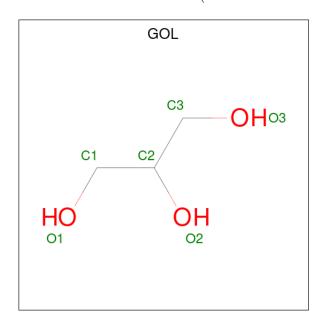
 $\bullet$  Molecule 3 is SULFATE ION (three-letter code: SO4) (formula:  $\mathrm{O_4S}).$ 



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	В	1	Total O S 5 4 1	0	0
3	С	1	Total O S 5 4 1	0	0



• Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	В	1	Total C O 6 3 3	0	0
4	С	1	Total C O 12 6 6	0	1

• Molecule 5 is water.

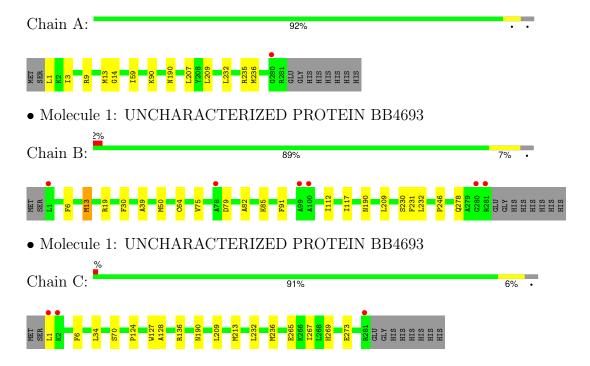
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	184	Total O 184 184	0	0
5	В	177	Total O 177 177	0	0
5	С	190	Total O 190 190	0	0



# 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: UNCHARACTERIZED PROTEIN BB4693





# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants	74.23Å 126.78Å 85.67Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $97.37^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	20.00 - 1.76	Depositor
rtesolution (A)	37.30 - 1.76	EDS
% Data completeness	98.6 (20.00-1.76)	Depositor
(in resolution range)	98.6 (37.30-1.76)	EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$< I/\sigma(I) > 1$	1.40 (at 1.76Å)	Xtriage
Refinement program	REFMAC 5.3.0034	Depositor
P. P.	0.168 , 0.214	Depositor
$R, R_{free}$	0.169 , 0.212	DCC
$R_{free}$ test set	2303 reflections $(3.00\%)$	wwPDB-VP
Wilson B-factor $(\mathring{A}^2)$	25.9	Xtriage
Anisotropy	0.195	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.33, 46.9	EDS
L-test for twinning <sup>2</sup>	$ < L >=0.49, < L^2>=0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	7377	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 5.03% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality (i)

#### 5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NA, GOL, SO4

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	$\mathbf{angles}$
IVIOI	Chain	RMSZ	# Z  > 5	RMSZ	# Z  > 5
1	A	0.57	0/2333	0.64	0/3169
1	В	0.57	0/2369	0.66	0/3219
1	С	0.56	0/2355	0.66	0/3200
All	All	0.57	0/7057	0.65	0/9588

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 (i)

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	2253	0	2244	7	0
1	В	2274	0	2276	12	0
1	С	2265	0	2253	9	0
2	A	1	0	0	0	0
3	A	5	0	0	0	0
3	В	5	0	0	0	0
3	С	5	0	0	0	0
4	В	6	0	8	0	0
4	С	12	0	16	0	0
5	A	184	0	0	1	0

Continued on next page...



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	В	177	0	0	0	0
5	С	190	0	0	1	0
All	All	7377	0	6797	28	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 2.

All (28) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic	Clash
1 1 200 1 511 11500	1 A 000[D] MDT HD0	distance (Å)	overlap (Å)
1:A:232:LEU:HD23	1:A:236[B]:MET:HE2	1.74	0.70
1:A:232:LEU:HD23	1:A:236[B]:MET:CE	2.27	0.65
1:B:50:MET:CE	1:B:91:PHE:HZ	2.17	0.58
1:B:82:ALA:HA	1:B:85:LYS:HG2	1.91	0.53
1:A:209:LEU:HD12	1:A:232:LEU:HD21	1.92	0.52
1:B:13[B]:MET:HE1	1:B:79:ASP:HA	1.91	0.51
1:C:236[A]:MET:HB3	1:C:267:ILE:HG12	1.92	0.51
1:A:90:LYS:NZ	5:A:426:HOH:O	2.45	0.49
1:B:50:MET:HE1	1:B:91:PHE:HZ	1.77	0.48
1:C:124:PRO:HA	1:C:127:TRP:CD2	2.49	0.48
1:C:209:LEU:HD12	1:C:232:LEU:HD21	1.96	0.47
1:B:209:LEU:HD12	1:B:232:LEU:HD21	1.96	0.47
1:B:19[B]:ARG:HG2	1:B:19[B]:ARG:HH21	1.79	0.46
1:B:50:MET:HE3	1:B:91:PHE:HZ	1.82	0.45
1:B:112:ILE:HG23	1:B:117:ILE:HB	1.98	0.45
1:C:34:LEU:HD11	1:C:213:MET:HB2	1.98	0.45
1:C:269[A]:HIS:NE2	1:C:273:GLU:OE1	2.50	0.44
1:C:232:LEU:HD23	1:C:236[A]:MET:CE	2.48	0.44
1:B:39:ALA:HB2	1:B:246:PRO:HA	2.01	0.43
1:C:1:LEU:HD21	1:C:265:GLU:HG2	2.00	0.43
1:B:64[A]:CYS:SG	1:B:91:PHE:CD1	3.12	0.42
1:C:70[B]:SER:OG	1:C:128:ALA:HB2	2.20	0.42
1:C:136:ARG:HG2	5:C:334:HOH:O	2.19	0.42
1:A:13:MET:HB3	1:A:14:GLY:H	1.68	0.41
1:A:207:LEU:O	1:A:235[B]:ARG:HD2	2.21	0.41
1:B:30:PHE:CD2	1:B:30:PHE:C	2.94	0.41
1:B:230[A]:SER:OG	1:B:231:PHE:N	2.54	0.40
1:A:3:ILE:HB	1:A:59:ILE:HA	2.03	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	Perce	entiles
1	A	$286/291 \ (98\%)$	278 (97%)	8 (3%)	0	100	100
1	В	291/291 (100%)	282 (97%)	9 (3%)	0	100	100
1	С	289/291 (99%)	279 (96%)	10 (4%)	0	100	100
All	All	866/873 (99%)	839 (97%)	27 (3%)	0	100	100

There are no Ramachandran outliers to report.

#### 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	240/242 (99%)	237 (99%)	3 (1%)	69	54	
1	В	245/242 (101%)	239 (98%)	6 (2%)	49	26	
1	С	243/242 (100%)	241 (99%)	2 (1%)	81	72	
All	All	728/726 (100%)	717 (98%)	11 (2%)	67	49	

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	LEU
1	A	9	ARG
1	A	190	ASN
1	В	6	PHE
1	В	13[A]	MET

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type
1	В	13[B]	MET
1	В	75	VAL
1	В	190	ASN
1	В	278	GLN
1	С	6	PHE
1	С	190	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	33	GLN
1	В	68	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 monosaccharides in this entry.

### 5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 1 is monoatomic - leaving 6 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).

Mal	Type	Chain	Res	Link	$\mathbf{B}$	ond leng	${ m gths}$	В	ond ang	gles
Mol	туре	Chain	nes .	Lilik	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	GOL	С	291[A]	-	5,5,5	0.40	0	5,5,5	0.58	0



Mol Type		Chain	n Res	es Link	B	Bond lengths			Bond angles		
MIOI	Type	Chain	nes	LIIIK	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2	
4	GOL	В	291	-	5,5,5	0.52	0	5,5,5	0.69	0	
3	SO4	A	291	-	4,4,4	0.43	0	6,6,6	0.44	0	
4	GOL	С	291[B]	-	5,5,5	0.40	0	5,5,5	0.23	0	
3	SO4	В	290	-	4,4,4	0.45	0	6,6,6	0.36	0	
3	SO4	С	290	-	4,4,4	0.32	0	6,6,6	0.44	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
4	GOL	С	291[B]	-	-	0/4/4/4	-
4	GOL	С	291[A]	-	-	3/4/4/4	-
4	GOL	В	291	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	В	291	GOL	O1-C1-C2-O2
4	В	291	GOL	O1-C1-C2-C3
4	С	291[A]	GOL	C1-C2-C3-O3
4	С	291[A]	GOL	O2-C2-C3-O3
4	С	291[A]	GOL	O1-C1-C2-C3

There are no ring outliers.

No monomer is involved in short contacts.

#### 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 (i)

#### 6.1 Protein, DNA and RNA chains (i)

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	281/291 (96%)	-0.23	1 (0%) 92 94	16, 28, 45, 91	0
1	В	281/291 (96%)	-0.05	6 (2%) 63 71	17, 28, 50, 84	0
1	С	281/291 (96%)	-0.28	3 (1%) 80 86	17, 28, 44, 88	0
All	All	843/873 (96%)	-0.19	10 (1%) 79 84	16, 28, 47, 91	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	280	GLY	5.1
1	С	281	ARG	4.0
1	В	280	GLY	4.0
1	В	1	LEU	3.6
1	В	99	ALA	3.0
1	В	281	ARG	2.8
1	В	100	ALA	2.7
1	С	2	LYS	2.6
1	В	78	ALA	2.3
1	С	1	LEU	2.2

#### 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 monosaccharides 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,  $95^{th}$  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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q < 0.9
4	GOL	В	291	6/6	0.78	0.17	33,47,49,49	0
4	GOL	С	291[A]	6/6	0.87	0.20	22,40,48,58	6
4	GOL	С	291[B]	6/6	0.87	0.20	32,42,45,55	6
2	NA	A	290	1/1	0.99	0.16	21,21,21,21	0
3	SO4	В	290	5/5	0.99	0.06	25,25,27,33	0
3	SO4	С	290	5/5	0.99	0.06	23,24,27,29	0
3	SO4	A	291	5/5	1.00	0.04	21,23,26,27	0

## 6.5 Other polymers (i)

There are no such residues in this entry.

