



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 02:40 am GMT

PDB ID : 3DRF  
Title : Lactococcal OppA complexed with an endogenous peptide in the closed conformation  
Authors : Berntsson, R.P.-A.; Doeven, M.K.; Duurkens, R.H.; Sengupta, D.; Marrink, S.-J.; Thunnissen, A.-M.; Poolman, B.; Slotboom, D.-J.  
Deposited on : 2008-07-11  
Resolution : 1.30 Å(reported)

This is a wwPDB X-ray Structure 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

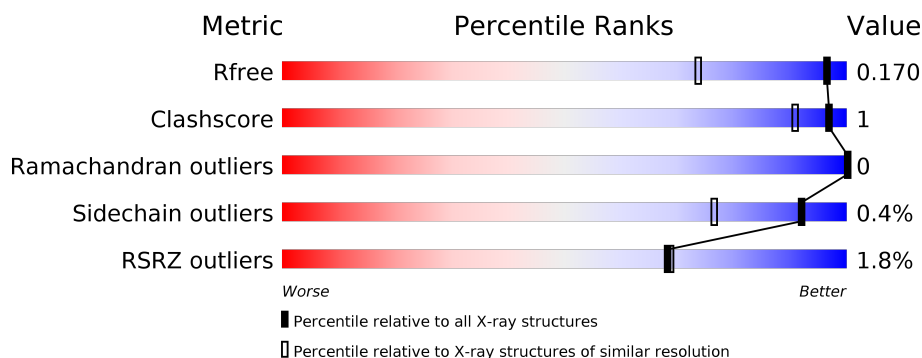
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.30 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))
$R_{free}$	111664	1286 (1.32-1.28)
Clashscore	122126	1332 (1.32-1.28)
Ramachandran outliers	120053	1282 (1.32-1.28)
Sidechain outliers	120020	1282 (1.32-1.28)
RSRZ outliers	108989	1250 (1.32-1.28)

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. 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	590	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 25%, orange 25%, yellow 25%, green 25%, grey 25%);"></div> <div style="display: flex; justify-content: space-between; padding: 0 5px;"> <span>25%</span> <span>87%</span> <span>9%</span> </div> </div>
2	B	8	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 25%, orange 25%, yellow 25%, green 25%, grey 25%);"></div> <div style="display: flex; justify-content: space-between; padding: 0 5px;"> <span>25%</span> <span>100%</span> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5516 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 Oligopeptide-binding protein oppA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	539	Total	C	N	O	S	0	20	0
			4349	2749	715	872	13			

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	EXPRESSION TAG	UNP A2RJ53
A	579	GLY	-	EXPRESSION TAG	UNP A2RJ53
A	580	SER	-	EXPRESSION TAG	UNP A2RJ53
A	581	ILE	-	EXPRESSION TAG	UNP A2RJ53
A	582	GLU	-	EXPRESSION TAG	UNP A2RJ53
A	583	GLY	-	EXPRESSION TAG	UNP A2RJ53
A	584	ARG	-	EXPRESSION TAG	UNP A2RJ53
A	585	HIS	-	EXPRESSION TAG	UNP A2RJ53
A	586	HIS	-	EXPRESSION TAG	UNP A2RJ53
A	587	HIS	-	EXPRESSION TAG	UNP A2RJ53
A	588	HIS	-	EXPRESSION TAG	UNP A2RJ53
A	589	HIS	-	EXPRESSION TAG	UNP A2RJ53
A	590	HIS	-	EXPRESSION TAG	UNP A2RJ53

- Molecule 2 is a protein called endogenous peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	8	Total	C	N	O	0	0	0
			48	27	9	12			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1111	Total	O	0	0
			1111	1111		

*Continued on next page...*

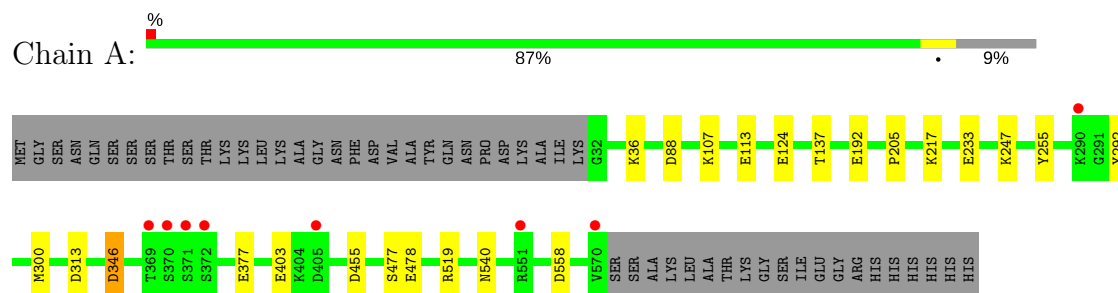
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	8	Total	O	0	0
			8	8		

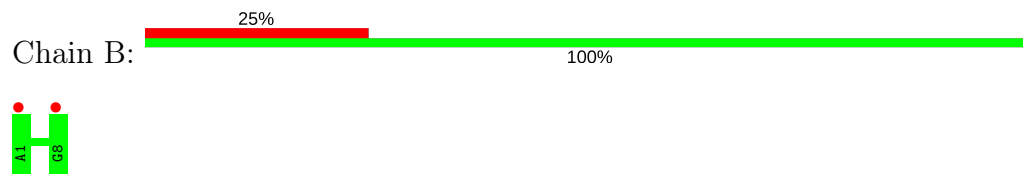
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: Oligopeptide-binding protein oppA



- Molecule 2: endogenous peptide



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	59.11Å 74.65Å 115.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	62.75 – 1.30 30.44 – 1.30	Depositor EDS
% Data completeness (in resolution range)	99.6 (62.75-1.30) 99.7 (30.44-1.30)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.67 (at 1.30Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.131 , 0.161 0.142 , 0.170	Depositor DCC
$R_{free}$ test set	6300 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	8.5	Xtriage
Anisotropy	0.039	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 48.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	5516	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	12.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.86% 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 [i](#)

### 5.1 Standard geometry [i](#)

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	1.31	8/4446 (0.2%)	0.86	7/6021 (0.1%)
2	B	1.55	0/47	0.94	0/62
All	All	1.31	8/4493 (0.2%)	0.87	7/6083 (0.1%)

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	403	GLU	CD-OE2	-7.27	1.17	1.25
1	A	205	PRO	N-CD	-6.88	1.38	1.47
1	A	377	GLU	CD-OE1	-6.36	1.18	1.25
1	A	478	GLU	CD-OE2	-6.23	1.18	1.25
1	A	124	GLU	CD-OE1	-6.04	1.19	1.25

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	205	PRO	CA-N-CD	-7.19	101.44	111.50
1	A	455	ASP	CB-CG-OD2	-6.29	112.64	118.30
1	A	519	ARG	NE-CZ-NH2	-5.68	117.46	120.30
1	A	558	ASP	CB-CG-OD1	5.59	123.33	118.30
1	A	346	ASP	CB-CG-OD2	-5.43	113.41	118.30

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	4349	0	4208	9	0
2	B	48	0	47	0	0
3	A	1111	0	0	6	1
3	B	8	0	0	0	0
All	All	5516	0	4255	9	1

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 1.

The worst 5 of 9 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:247:LYS:HE2	3:A:1377:HOH:O	1.59	1.02
1:A:36:LYS:NZ	3:A:1328:HOH:O	1.88	1.02
1:A:217:LYS:HE3	3:A:1086:HOH:O	1.77	0.81
1:A:107:LYS:NZ	3:A:980:HOH:O	2.22	0.73
1:A:247:LYS:CE	3:A:1377:HOH:O	2.29	0.67

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:965:HOH:O	3:A:1640:HOH:O[1_455]	2.17	0.03

## 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	556/590 (94%)	547 (98%)	9 (2%)	0	100	100
2	B	6/8 (75%)	6 (100%)	0	0	100	100
All	All	562/598 (94%)	553 (98%)	9 (2%)	0	100	100



There are no Ramachandran outliers to report.

### 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	474/496 (96%)	472 (100%)	2 (0%)	92	76
2	B	5/5 (100%)	5 (100%)	0	100	100
All	All	479/501 (96%)	477 (100%)	2 (0%)	92	76

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	137	THR
1	A	255	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	379	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules 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 [i](#)

There are no ligands in this entry.

## 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<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	539/590 (91%)	0.18	8 (1%) 73 74	5, 9, 17, 25	0
2	B	8/8 (100%)	1.36	2 (25%) 0 0	10, 11, 18, 21	1 (12%)
All	All	547/598 (91%)	0.20	10 (1%) 68 69	5, 9, 17, 25	1 (0%)

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	1	ALA	5.5
1	A	371	SER	5.2
1	A	290	LYS	4.1
1	A	570[A]	VAL	4.1
1	A	405	ASP	3.9

### 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 carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.