



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 06:24 AM GMT

PDB ID : 2WZA  
Title : TWO INTRAMOLECULAR ISOPEPTIDE BONDS ARE IDENTIFIED IN  
THE CRYSTAL STRUCTURE OF THE STREPTOCOCCUS GORDONII  
SSPB C-TERMINAL DOMAIN  
Authors : Forsgren, N.; Lamont, R.J.; Persson, K.  
Deposited on : 2009-11-26  
Resolution : 2.08 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  
<http://wwpdb.org/validation/2016/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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

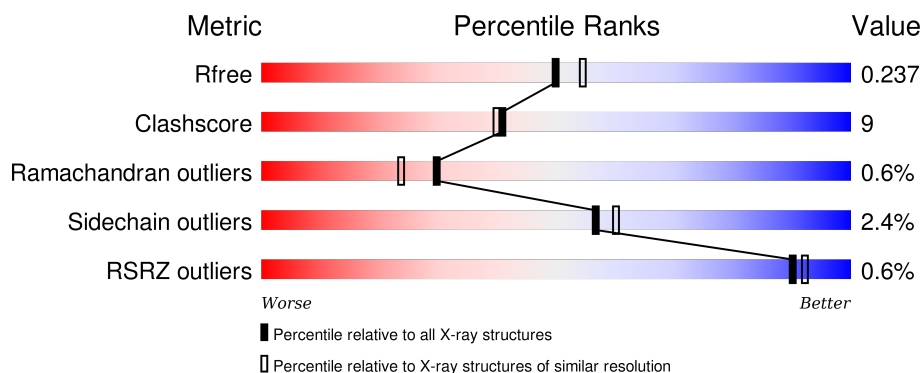
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*


The reported resolution of this entry is 2.08 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))
$R_{free}$	91344	4546 (2.10-2.06)
Clashscore	102246	5101 (2.10-2.06)
Ramachandran outliers	100387	5048 (2.10-2.06)
Sidechain outliers	100360	5049 (2.10-2.06)
RSRZ outliers	91569	4556 (2.10-2.06)

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	374	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2808 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 AGGLUTININ RECEPTOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	338	Total	C	N	O	S	0	2	0
			2667	1684	428	552	3			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1040	MET	-	EXPRESSION TAG	UNP P16952
A	1041	LYS	-	EXPRESSION TAG	UNP P16952
A	1042	PRO	-	EXPRESSION TAG	UNP P16952
A	1043	MET	-	EXPRESSION TAG	UNP P16952
A	1044	SER	-	EXPRESSION TAG	UNP P16952
A	1045	ASP	-	EXPRESSION TAG	UNP P16952
A	1046	TYR	-	EXPRESSION TAG	UNP P16952
A	1047	ASP	-	EXPRESSION TAG	UNP P16952
A	1048	ILE	-	EXPRESSION TAG	UNP P16952
A	1049	PRO	-	EXPRESSION TAG	UNP P16952
A	1050	THR	-	EXPRESSION TAG	UNP P16952
A	1051	THR	-	EXPRESSION TAG	UNP P16952
A	1052	GLU	-	EXPRESSION TAG	UNP P16952
A	1053	ASN	-	EXPRESSION TAG	UNP P16952
A	1054	LEU	-	EXPRESSION TAG	UNP P16952
A	1055	TYR	-	EXPRESSION TAG	UNP P16952
A	1056	PHE	-	EXPRESSION TAG	UNP P16952
A	1057	GLN	-	EXPRESSION TAG	UNP P16952
A	1058	GLY	-	EXPRESSION TAG	UNP P16952
A	1059	ALA	-	EXPRESSION TAG	UNP P16952
A	1060	MET	-	EXPRESSION TAG	UNP P16952

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	3	Total 3	Ca 3	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	138	Total 138	O 138	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.75Å 80.15Å 87.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.58 – 2.08 42.58 – 2.08	Depositor EDS
% Data completeness (in resolution range)	98.2 (42.58-2.08) 99.6 (42.58-2.08)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.81 (at 2.08Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.192 , 0.255 0.185 , 0.237	Depositor DCC
$R_{free}$ test set	2024 reflections (5.12%)	DCC
Wilson B-factor (Å <sup>2</sup> )	21.0	Xtriage
Anisotropy	0.601	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 50.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 21096 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	2808	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.05% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: CA

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	0.63	0/2721	0.68	0/3692

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	2667	0	2594	45	0
2	A	3	0	0	0	0
3	A	138	0	0	2	0
All	All	2808	0	2594	45	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 9.

All (45) 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:1259:LYS:NZ	1:A:1393:ASN:CG	1.86	1.28

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1082:LYS:NZ	1:A:1232:ASN:CG	1.91	1.22
1:A:1259:LYS:HZ1	1:A:1393:ASN:CG	1.51	1.03
1:A:1082:LYS:HZ1	1:A:1232:ASN:CG	1.60	1.00
1:A:1082:LYS:HZ3	1:A:1232:ASN:CG	1.56	0.98
1:A:1259:LYS:HZ3	1:A:1393:ASN:CG	1.56	0.96
1:A:1254:LYS:HE3	1:A:1256:THR:HG21	1.63	0.80
1:A:1229:LYS:HZ3	1:A:1251:ASN:HB2	1.48	0.78
1:A:1410:ARG:NH1	1:A:1410:ARG:HB2	2.00	0.76
1:A:1150:ASP:OD1	1:A:1152:LYS:HG3	1.89	0.71
1:A:1229:LYS:NZ	1:A:1251:ASN:HB2	2.05	0.71
1:A:1141:VAL:O	1:A:1143:PRO:HD3	1.93	0.69
1:A:1225:LYS:O	1:A:1254:LYS:HG3	1.96	0.65
1:A:1255:ILE:HG23	1:A:1290:GLY:HA3	1.79	0.65
1:A:1198:GLU:O	1:A:1202:LYS:HG3	2.00	0.61
1:A:1410:ARG:HB2	1:A:1410:ARG:HH11	1.66	0.61
1:A:1150:ASP:OD2	1:A:1154:ASN:HB2	2.03	0.57
1:A:1354:ASN:HD22	1:A:1354:ASN:N	2.06	0.53
1:A:1152:LYS:HD2	1:A:1154:ASN:HD22	1.73	0.53
1:A:1410:ARG:CB	1:A:1410:ARG:HH11	2.23	0.51
1:A:1145:LEU:HB3	1:A:1216:PRO:HG2	1.93	0.51
1:A:1254:LYS:HE3	1:A:1256:THR:CG2	2.37	0.50
1:A:1259:LYS:HE2	1:A:1393:ASN:HB3	1.94	0.50
1:A:1294:GLN:HG2	1:A:1374:PRO:HD3	1.92	0.50
1:A:1214:THR:HG22	1:A:1216:PRO:HD3	1.93	0.49
1:A:1159:VAL:HA	1:A:1193:SER:O	2.12	0.49
1:A:1263:VAL:HG13	1:A:1272:LEU:HD11	1.95	0.49
1:A:1115:LYS:HB2	1:A:1208:GLY:HA3	1.96	0.48
1:A:1099:ALA:HB1	1:A:1221:SER:HA	1.99	0.44
1:A:1082:LYS:HE2	1:A:1232:ASN:HB3	1.99	0.44
1:A:1150:ASP:CG	1:A:1152:LYS:HG3	2.37	0.43
1:A:1222:GLU:HG2	1:A:1223:PHE:N	2.33	0.43
1:A:1177:LEU:HD11	1:A:1237:ILE:HG21	2.01	0.43
1:A:1263:VAL:HG22	3:A:2067:HOH:O	2.19	0.43
1:A:1162:GLN:HG2	1:A:1164:TYR:CZ	2.54	0.43
1:A:1199:GLU:O	1:A:1203:GLN:HB2	2.19	0.43
1:A:1385:ARG:HG2	1:A:1413:THR:OG1	2.19	0.42
1:A:1155:GLN:HG3	3:A:2030:HOH:O	2.20	0.42
1:A:1306:VAL:HG22	1:A:1358:THR:HG23	2.02	0.41
1:A:1294:GLN:CG	1:A:1374:PRO:HD3	2.49	0.41
1:A:1152:LYS:HD2	1:A:1154:ASN:ND2	2.36	0.41
1:A:1107:LEU:HB2	1:A:1213:ILE:HB	2.02	0.41

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1266:ASP:HA	1:A:1267:PRO:HD3	1.85	0.41
1:A:1354:ASN:N	1:A:1354:ASN:ND2	2.69	0.40
1:A:1229:LYS:HZ3	1:A:1251:ASN:CB	2.28	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	338/374 (90%)	331 (98%)	5 (2%)	2 (1%)	30 24

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1269	SER
1	A	1142	ARG

### 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	297/328 (90%)	290 (98%)	7 (2%)	57 60

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1118	LYS
1	A	1144	ASP
1	A	1152	LYS
1	A	1203	GLN
1	A	1226	THR
1	A	1354	ASN
1	A	1395	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1154	ASN
1	A	1190	GLN
1	A	1354	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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 [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	338/374 (90%)	-0.05	2 (0%) 90 92	10, 22, 44, 60	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1118	LYS	2.8
1	A	1125	GLN	2.1

### 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](#)

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	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	CA	A	2415	1/1	0.99	0.13	0.31	19,19,19,19	0
2	CA	A	2416	1/1	0.99	0.10	-0.52	16,16,16,16	0
2	CA	A	2414	1/1	0.98	0.08	-1.55	21,21,21,21	0

## 6.5 Other polymers

There are no such residues in this entry.