



# Full wwPDB X-ray Structure Validation Report

Feb 26, 2014 – 06:03 PM GMT

PDB ID : 2HSI  
Title : Crystal structure of putative peptidase M23 from pseudomonas aeruginosa,  
New York Structural Genomics Consortium  
Authors : Patskovsky, Y.; Almo, S.C.; Burley, S.K.; New York SGX Research Center for  
Structural Genomics (NYSGXRC)  
Deposited on : 2006-07-21  
Resolution : 1.90 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>

---

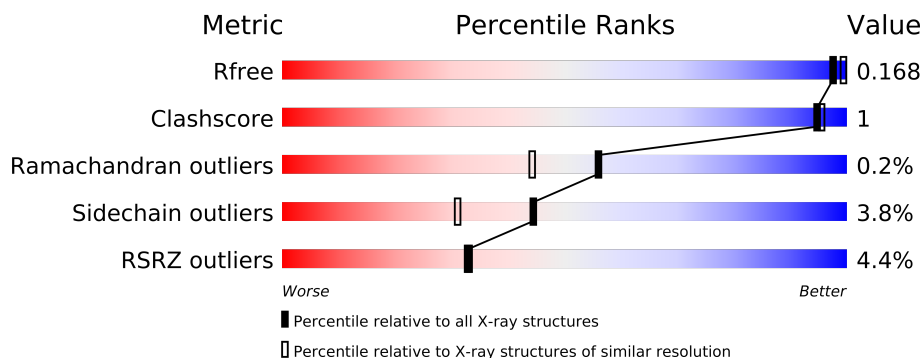
The following versions of software and data (see [references](#)) were used in the production of this report:

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 1.90 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	3684 (1.90-1.90)
Clashscore	79885	4465 (1.90-1.90)
Ramachandran outliers	78287	4413 (1.90-1.90)
Sidechain outliers	78261	4414 (1.90-1.90)
RSRZ outliers	66119	3686 (1.90-1.90)

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	282	
1	B	282	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3971 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 peptidase M23.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	227	Total	C	N	O	S	0	4	0
			1746	1122	309	310	5			
1	B	228	Total	C	N	O	S	0	2	0
			1738	1118	309	306	5			

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Zn	0	0
			1	1		
2	A	1	Total	Zn	0	0
			1	1		

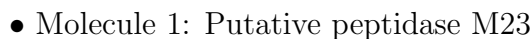
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	249	Total	O	0	0
			249	249		
3	B	236	Total	O	0	0
			236	236		

i

- Molecule 1: Putative peptidase M23

Age Group	Percentage
18-29	90%
30-49	85%
50-64	75%
65 and older	65%



Age Group	Percentage of Respondents
18-29	90%
30-49	86%
50-64	78%
65+	66%

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	108.31Å 108.31Å 83.18Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 1.90 27.08 – 1.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (20.00-1.90) 99.9 (27.08-1.90)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.54 (at 1.91Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.204 , 0.245 0.150 , 0.168	Depositor DCC
$R_{free}$ test set	1360 reflections (3.22%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.2	Xtriage
Anisotropy	0.181	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 35.3	EDS
Estimated twinning fraction	0.254 for h,-h-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.41$ , $\langle L^2 \rangle = 0.24$	Xtriage
Outliers	0 of 43675 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3971	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.01% 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: ZN

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	0.46	0/1803	0.64	0/2442
1	B	0.44	0/1789	0.63	1/2424 (0.0%)
All	All	0.45	0/3592	0.64	1/4866 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	126	ASP	CB-CG-OD2	5.85	123.56	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	165	PRO	Peptide

### 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	1746	0	1771	7	0
1	B	1738	0	1766	3	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	249	0	0	1	0
3	B	236	0	0	0	0
All	All	3971	0	3537	10	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 1.

All (10) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:150:SER:OG	1:A:151:ASN:OD1	2.31	0.48
1:B:123:LEU:HB3	1:B:124:PRO:HD2	1.97	0.47
1:A:165:PRO:O	1:A:180:HIS:ND1	2.48	0.47
1:A:225:MET:HB3	1:A:263:ASN:HB2	1.97	0.46
1:B:128:LYS:HE2	1:B:128:LYS:HB2	1.42	0.44
1:A:124:PRO:HA	1:A:125:GLU:HA	1.66	0.43
1:B:195:ALA:HA	1:B:196:PRO:HD3	1.90	0.43
1:A:85:GLN:HG3	1:A:102:VAL:HG23	2.01	0.41
1:A:128:LYS:HE2	1:A:128:LYS:HB2	1.92	0.41
1:A:90:ARG:NH2	3:A:518:HOH:O	2.52	0.41

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	225/282 (80%)	220 (98%)	5 (2%)	0	100	100
1	B	224/282 (79%)	217 (97%)	6 (3%)	1 (0%)	43	29

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	449/564 (80%)	437 (97%)	11 (2%)	1 (0%)	56 44

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	166	PHE

### 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	186/229 (81%)	180 (97%)	6 (3%)	51 39
1	B	184/229 (80%)	176 (96%)	8 (4%)	40 26
All	All	370/458 (81%)	356 (96%)	14 (4%)	44 31

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

Mol	Chain	Res	Type
1	A	166	PHE
1	A	225	MET
1	A	231	LYS
1	A	248	LYS
1	A	256	THR
1	A	270	ARG
1	B	90	ARG
1	B	128	LYS
1	B	166	PHE
1	B	225	MET
1	B	231	LYS
1	B	248	LYS
1	B	256	THR
1	B	270	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.



### 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 ⓘ

Of 2 ligands modelled in this entry, 2 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.

## 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	227/282 (80%)	0.05	10 (4%) 33 33	15, 23, 56, 94	0
1	B	228/282 (80%)	-0.08	10 (4%) 33 33	15, 24, 56, 70	0
All	All	455/564 (80%)	-0.01	20 (4%) 33 33	15, 24, 56, 94	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	179	PRO	5.2
1	A	256	THR	5.0
1	A	281	GLN	4.9
1	A	125	GLU	4.5
1	B	26	PHE	4.1
1	B	167	GLY	3.6
1	B	254	ARG	3.5
1	B	29	ARG	3.5
1	B	256	THR	3.5
1	A	26	PHE	3.3
1	A	124	PRO	3.2
1	A	29	ARG	3.2
1	B	280	PHE	3.2
1	A	280	PHE	3.1
1	A	165	PRO	3.1
1	A	127	LEU	3.0
1	B	165	PRO	2.5
1	B	124	PRO	2.4
1	B	27	ILE	2.3
1	B	259	HIS	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(Å <sup>2</sup> )	Q<0.9
2	ZN	B	283	1/1	0.04	-2.10	46,46,46,46	0
2	ZN	A	283	1/1	0.03	-2.15	50,50,50,50	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.