



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 05:24 pm BST

PDB ID : 5I5D  
Title : Salmonella global domain 245  
Authors : Dong, C.; Dong, H.  
Deposited on : 2016-02-15  
Resolution : 1.64 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

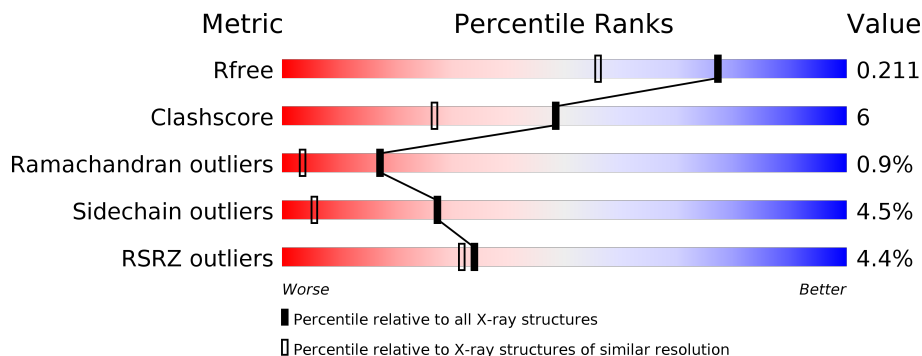
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.64 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))
$R_{free}$	130704	3122 (1.66-1.62)
Clashscore	141614	3268 (1.66-1.62)
Ramachandran outliers	138981	3215 (1.66-1.62)
Sidechain outliers	138945	3215 (1.66-1.62)
RSRZ outliers	127900	3079 (1.66-1.62)

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 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  $\leq 5\%$ . 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	586	 2% 50% 6% 42%
1	B	586	 3% 50% 6% 42%
1	C	586	 3% 48% 7% 42%
1	D	586	 2% 51% 7% 42%

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11987 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 Inner membrane protein YejM.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	339	2681	1690	468	515	8	0	0	0
1	B	340	2686	1693	469	516	8	0	0	0
1	C	338	2675	1687	467	513	8	0	0	0
1	D	342	2702	1703	471	520	8	0	0	0

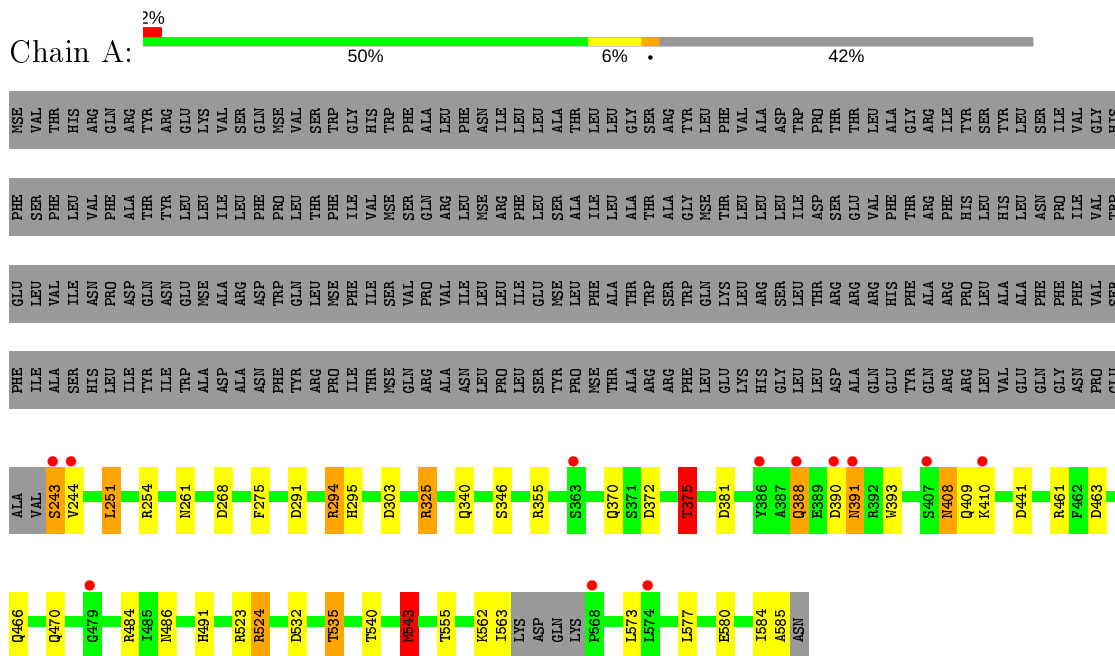
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	332	332	332	0	0
2	B	316	316	316	0	0
2	C	247	247	247	0	0
2	D	348	348	348	0	0

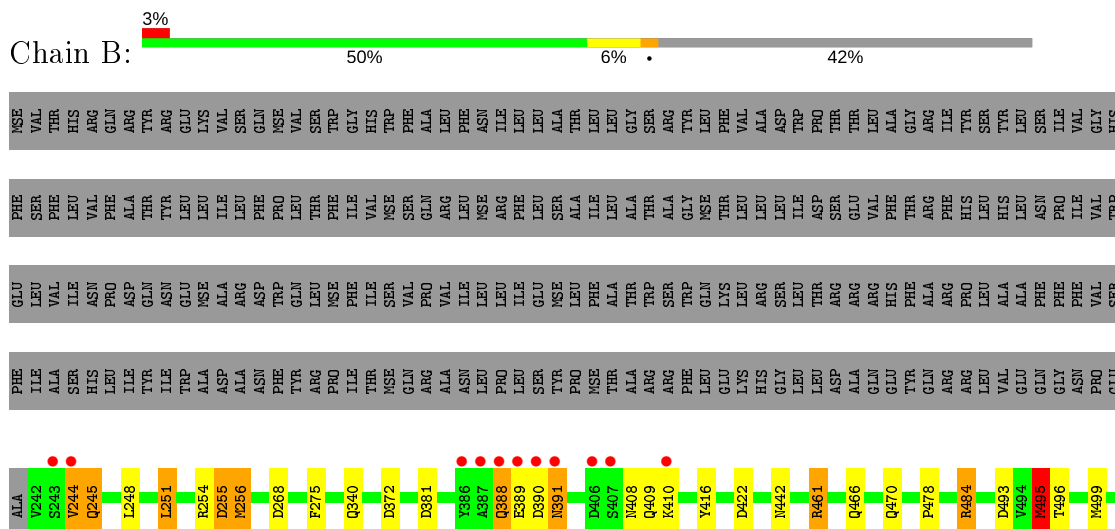
### 3 Residue-property plots

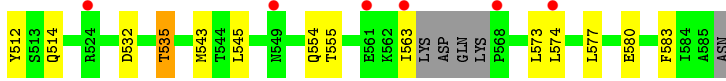
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Inner membrane protein YejM

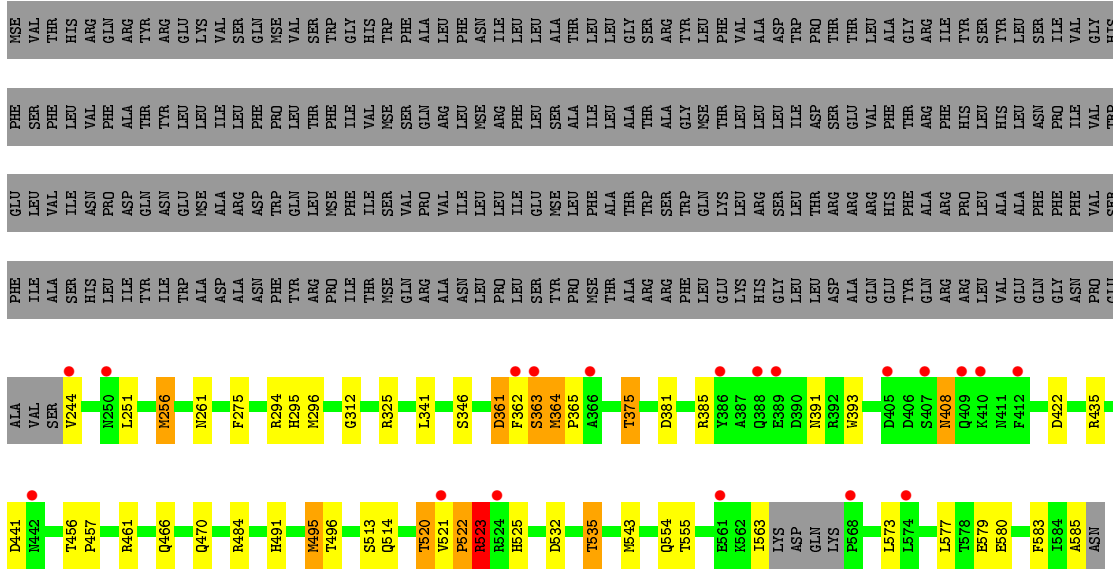


- Molecule 1: Inner membrane protein YejM

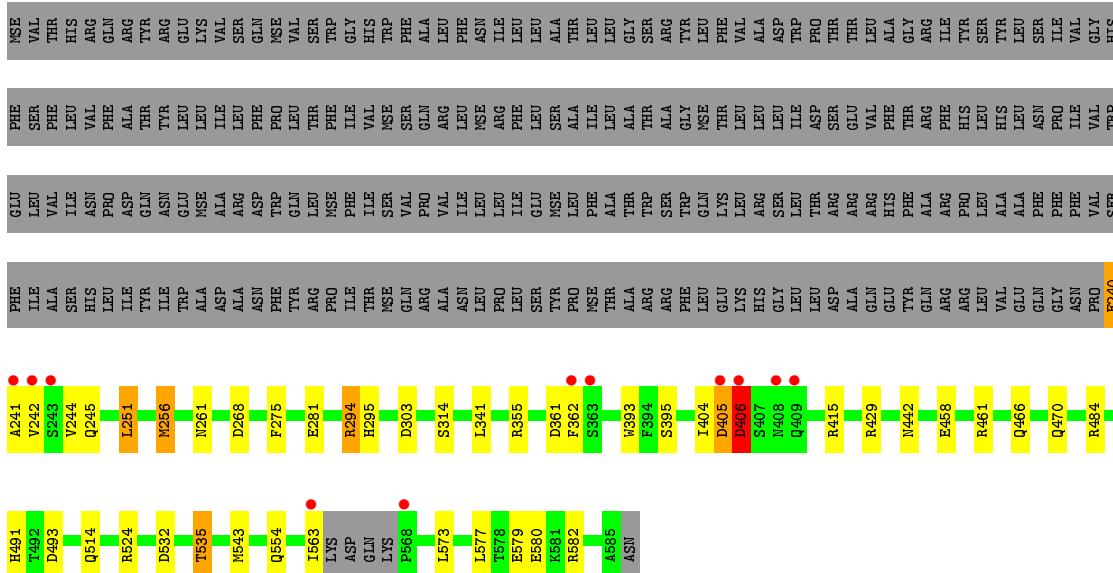




- Molecule 1: Inner membrane protein YejM



- Molecule 1: Inner membrane protein YejM



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.60Å 196.09Å 70.27Å 90.00° 95.79° 90.00°	Depositor
Resolution (Å)	98.04 – 1.64 98.04 – 1.64	Depositor EDS
% Data completeness (in resolution range)	85.8 (98.04-1.64) 85.8 (98.04-1.64)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.50 (at 1.64Å)	Xtrriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.192 , 0.222 0.184 , 0.211	Depositor DCC
$R_{free}$ test set	7305 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.5	Xtrriage
Anisotropy	0.176	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 44.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	11987	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 71.57 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.5653e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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	0.89	2/2736 (0.1%)	1.06	21/3714 (0.6%)
1	B	0.91	2/2741 (0.1%)	1.04	12/3721 (0.3%)
1	C	0.78	1/2730 (0.0%)	1.02	12/3706 (0.3%)
1	D	0.91	3/2757 (0.1%)	1.15	23/3743 (0.6%)
All	All	0.87	8/10964 (0.1%)	1.07	68/14884 (0.5%)

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	A	0	1
1	B	0	1
1	C	0	2
1	D	0	3
All	All	0	7

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	405	ASP	C-N	-10.01	1.11	1.34
1	D	404	ILE	C-N	-8.02	1.15	1.34
1	A	243	SER	CA-CB	-8.01	1.41	1.52
1	D	395	SER	CB-OG	-7.70	1.32	1.42
1	B	256	MSE	C-N	6.52	1.44	1.33

The worst 5 of 68 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	405	ASP	O-C-N	-20.26	90.28	122.70
1	D	405	ASP	CA-C-N	14.62	149.37	117.20
1	A	325	ARG	NE-CZ-NH1	12.48	126.54	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	405	ASP	C-N-CA	12.34	152.54	121.70
1	B	495	MSE	CG-SE-CE	10.47	121.93	98.90

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	390	ASP	Peptide
1	B	255	ASP	Mainchain
1	C	522	PRO	Peptide
1	C	523	ARG	Peptide
1	D	240	GLU	Peptide

## 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	2681	0	2584	26	0
1	B	2686	0	2585	40	0
1	C	2675	0	2579	43	0
1	D	2702	0	2602	17	0
2	A	332	0	0	6	0
2	B	316	0	0	3	0
2	C	247	0	0	10	0
2	D	348	0	0	11	0
All	All	11987	0	10350	126	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 6.

The worst 5 of 126 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:B:255:ASP:O	1:B:256:MSE:N	1.69	1.23
1:B:255:ASP:C	1:B:256:MSE:CA	2.09	1.21

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:495:MSE:SE	2:C:814:HOH:O	2.10	1.17
1:B:255:ASP:CA	1:B:256:MSE:N	2.09	1.14
1:B:495:MSE:SE	2:B:851:HOH:O	2.15	1.12

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	335/586 (57%)	321 (96%)	11 (3%)	3 (1%)	17	3
1	B	336/586 (57%)	319 (95%)	14 (4%)	3 (1%)	17	3
1	C	334/586 (57%)	319 (96%)	12 (4%)	3 (1%)	17	3
1	D	338/586 (58%)	323 (96%)	12 (4%)	3 (1%)	17	3
All	All	1343/2344 (57%)	1282 (96%)	49 (4%)	12 (1%)	17	3

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	391	ASN
1	A	409	GLN
1	A	562	LYS
1	B	244	VAL
1	B	391	ASN

### 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	291/493 (59%)	279 (96%)	12 (4%)	30	7
1	B	291/493 (59%)	279 (96%)	12 (4%)	30	7
1	C	290/493 (59%)	275 (95%)	15 (5%)	23	4
1	D	293/493 (59%)	279 (95%)	14 (5%)	25	5
All	All	1165/1972 (59%)	1112 (96%)	53 (4%)	27	6

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	580	GLU
1	C	363	SER
1	D	573	LEU
1	C	251	LEU
1	C	275	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	554	GLN
1	C	340	GLN
1	D	514	GLN
1	C	261	ASN
1	C	388	GLN

### 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](#)

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	D	2
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	404:ILE	C	405:ASP	N	1.15
1	D	405:ASP	C	406:ASP	N	1.11
1	B	255:ASP	C	256:MSE	N	0.80

## 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	331/586 (56%)	-0.04	12 (3%) 42 40	12, 21, 46, 71	0
1	B	332/586 (56%)	0.15	17 (5%) 28 25	11, 22, 51, 83	0
1	C	330/586 (56%)	0.23	19 (5%) 23 21	15, 27, 56, 84	0
1	D	334/586 (56%)	-0.01	11 (3%) 46 44	12, 21, 45, 83	0
All	All	1327/2344 (56%)	0.08	59 (4%) 34 32	11, 23, 50, 84	0

The worst 5 of 59 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	407	SER	12.2
1	D	242	VAL	10.4
1	B	386	TYR	8.3
1	C	405	ASP	7.6
1	D	563	ILE	7.3

### 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

There are no such residues in this entry.