



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:32 PM GMT

PDB ID : 3ZRW  
Title : The structure of the dimeric Hamp-Dhp fusion A291V mutant  
Authors : Zeth, K.; Hulko, M.; Ferris, H.U.; Martin, J.; Lupas, A.N.  
Deposited on : 2011-06-20  
Resolution : 2.25 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  
<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

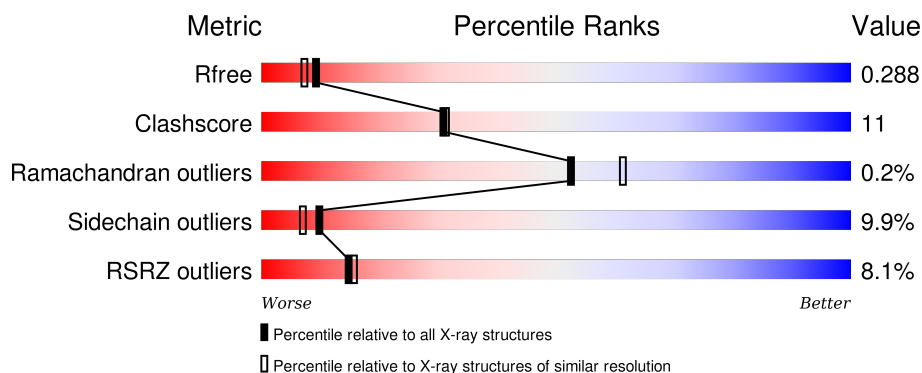
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.25 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	1640 (2.28-2.24)
Clashscore	102246	1095 (2.26-2.26)
Ramachandran outliers	100387	1063 (2.26-2.26)
Sidechain outliers	100360	1063 (2.26-2.26)
RSRZ outliers	91569	1647 (2.28-2.24)

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	116	<div> <div>12%</div> <div>70%</div> <div>19%</div> <div>6%</div> <div>• •</div> </div>
1	C	116	<div> <div>4%</div> <div>67%</div> <div>26%</div> <div>• •</div> </div>
1	D	116	<div> <div>3%</div> <div>76%</div> <div>16%</div> <div>• 5%</div> </div>
2	B	116	<div> <div>11%</div> <div>72%</div> <div>17%</div> <div>5% 5%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3714 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 AF1503 PROTEIN, OSMOLARITY SENSOR PROTEIN ENVZ.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	111	Total	C	N	O	S	5	1	0
			874	541	152	176	5			
1	C	111	Total	C	N	O	S	0	0	0
			877	541	158	174	4			
1	D	110	Total	C	N	O	S	6	1	0
			871	539	154	174	4			

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	291	VAL	ALA	ENGINEERED MUTATION	UNP O28769
A	274	GLY	-	EXPRESSION TAG	UNP O28769
A	275	SER	-	EXPRESSION TAG	UNP O28769
A	276	HIS	-	EXPRESSION TAG	UNP O28769
A	277	MET	-	EXPRESSION TAG	UNP O28769
C	291	VAL	ALA	ENGINEERED MUTATION	UNP O28769
C	274	GLY	-	EXPRESSION TAG	UNP O28769
C	275	SER	-	EXPRESSION TAG	UNP O28769
C	276	HIS	-	EXPRESSION TAG	UNP O28769
C	277	MET	-	EXPRESSION TAG	UNP O28769
D	291	VAL	ALA	ENGINEERED MUTATION	UNP O28769
D	274	GLY	-	EXPRESSION TAG	UNP O28769
D	275	SER	-	EXPRESSION TAG	UNP O28769
D	276	HIS	-	EXPRESSION TAG	UNP O28769
D	277	MET	-	EXPRESSION TAG	UNP O28769

- Molecule 2 is a protein called AF1503 PROTEIN, OSMOLARITY SENSOR PROTEIN ENVZ.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	110	Total	C	N	O	S	0	1	0
			866	530	156	176	4			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	291	VAL	ALA	ENGINEERED MUTATION	UNP O28769
B	329	GLN	LEU	ENGINEERED MUTATION	UNP O28769
B	274	GLY	-	EXPRESSION TAG	UNP O28769
B	275	SER	-	EXPRESSION TAG	UNP O28769
B	276	HIS	-	EXPRESSION TAG	UNP O28769
B	277	MET	-	EXPRESSION TAG	UNP O28769

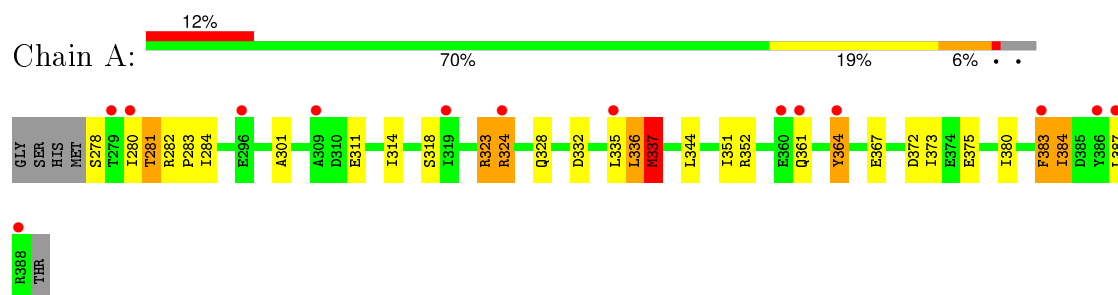
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	26	Total O 26 26	0	0
3	B	32	Total O 32 32	0	0
3	C	93	Total O 93 93	0	0
3	D	75	Total O 75 75	0	0

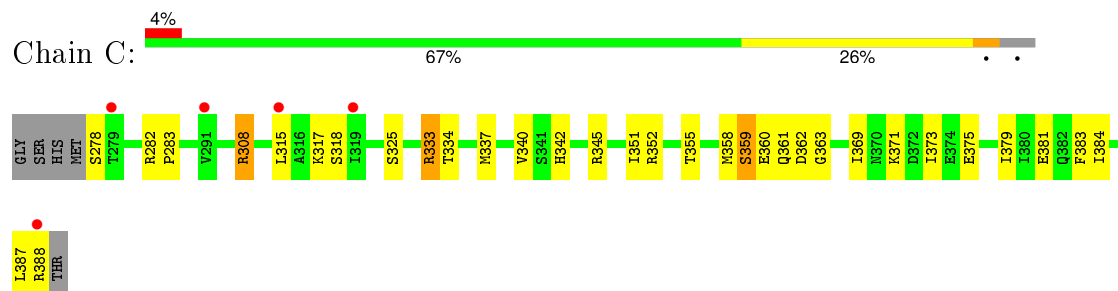
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

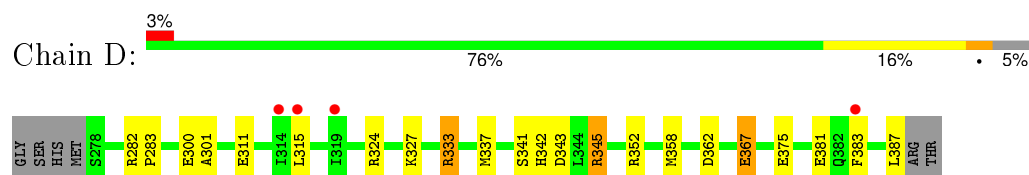
- Molecule 1: AF1503 PROTEIN, OSMOLARITY SENSOR PROTEIN ENVZ



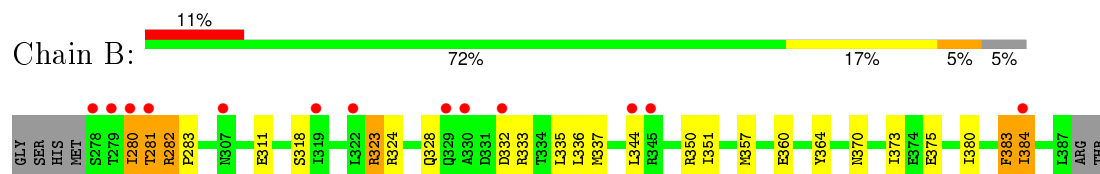
- Molecule 1: AF1503 PROTEIN, OSMOLARITY SENSOR PROTEIN ENVZ



- Molecule 1: AF1503 PROTEIN, OSMOLARITY SENSOR PROTEIN ENVZ



- Molecule 2: AF1503 PROTEIN, OSMOLARITY SENSOR PROTEIN ENVZ



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.86Å 46.67Å 157.46Å 90.00° 94.31° 90.00°	Depositor
Resolution (Å)	25.00 – 2.25 19.79 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.8 (25.00-2.25) 99.9 (19.79-2.20)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.26 (at 2.19Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.214 , 0.269 0.241 , 0.288	Depositor DCC
$R_{free}$ test set	1264 reflections (5.40%)	DCC
Wilson B-factor (Å <sup>2</sup> )	29.1	Xtriage
Anisotropy	0.221	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 34.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 26379 reflections (0.004%)	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3714	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.94% 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 [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	1.01	6/884 (0.7%)	1.08	4/1193 (0.3%)
1	C	1.14	1/884 (0.1%)	1.10	4/1192 (0.3%)
1	D	1.24	6/881 (0.7%)	1.12	8/1189 (0.7%)
2	B	0.93	4/875 (0.5%)	1.00	4/1180 (0.3%)
All	All	1.09	17/3524 (0.5%)	1.08	20/4754 (0.4%)

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	C	0	1

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	367[A]	GLU	N-CA	-9.40	1.27	1.46
1	D	367[B]	GLU	N-CA	-9.40	1.27	1.46
1	A	337[A]	MET	N-CA	9.34	1.65	1.46
1	A	337[B]	MET	N-CA	9.34	1.65	1.46
2	B	364	TYR	CE1-CZ	-7.84	1.28	1.38

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	333	ARG	NE-CZ-NH1	9.84	125.22	120.30
1	C	333	ARG	NE-CZ-NH1	9.14	124.87	120.30
1	A	332	ASP	CB-CG-OD1	-7.78	111.30	118.30
2	B	332	ASP	CB-CG-OD1	-7.67	111.40	118.30
1	A	332	ASP	CB-CG-OD2	7.52	125.07	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	387	LEU	Peptide

## 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	874	0	882	28	0
1	C	877	0	891	23	0
1	D	871	0	881	17	0
2	B	866	0	869	24	0
3	A	26	0	0	3	0
3	B	32	0	0	1	0
3	C	93	0	0	3	0
3	D	75	0	0	3	0
All	All	3714	0	3523	76	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 11.

The worst 5 of 76 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:383:PHE:CZ	2:B:383:PHE:HZ	1.65	1.14
2:B:323:ARG:HH11	2:B:323:ARG:HG2	1.09	1.11
2:B:351:ILE:HG22	2:B:373:ILE:HD11	1.16	1.10
2:B:351:ILE:CG2	2:B:373:ILE:HD11	1.81	1.09
1:A:323:ARG:HH11	1:A:323:ARG:HG2	1.20	1.05

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	110/116 (95%)	108 (98%)	2 (2%)	0	100	100
1	C	109/116 (94%)	107 (98%)	1 (1%)	1 (1%)	21	18
1	D	109/116 (94%)	107 (98%)	2 (2%)	0	100	100
2	B	109/116 (94%)	108 (99%)	1 (1%)	0	100	100
All	All	437/464 (94%)	430 (98%)	6 (1%)	1 (0%)	52	61

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	359	SER

### 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	97/102 (95%)	81 (84%)	16 (16%)	3	1
1	C	97/102 (95%)	87 (90%)	10 (10%)	9	6
1	D	96/102 (94%)	95 (99%)	1 (1%)	82	89
2	B	96/102 (94%)	84 (88%)	12 (12%)	6	3
All	All	386/408 (95%)	347 (90%)	39 (10%)	10	7

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	281	THR
2	B	328	GLN
1	C	361	GLN
2	B	318	SER
2	B	323	ARG

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	328	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](#)

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	111/116 (95%)	0.86	14 (12%) 5 5	23, 55, 87, 108	0
1	C	111/116 (95%)	0.26	5 (4%) 37 41	20, 29, 48, 67	0
1	D	110/116 (94%)	0.32	4 (3%) 46 50	21, 29, 52, 59	0
2	B	110/116 (94%)	0.91	13 (11%) 6 6	21, 59, 88, 129	0
All	All	442/464 (95%)	0.59	36 (8%) 15 16	20, 36, 79, 129	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	280	ILE	7.0
2	B	278	SER	6.2
2	B	279	THR	5.7
1	A	309	ALA	5.4
1	A	279	THR	5.2

### 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 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.