



# Full wwPDB X-ray Structure Validation Report

Feb 27, 2014 – 01:01 AM GMT

PDB ID : 4A3V  
Title : yeast regulatory particle proteasome assembly chaperone Hsm3 in complex with Rpt1 C-terminal fragment  
Authors : Richet, N.; Barrault, M.B.; Godart, C.; Murciano, B.; Le Tallec, B.; Rousseau, E.; Ledu, M.H.; Charbonnier, J.B.; Legrand, P.; Guerois, R.; Peyroche, A.; Ochsenbein, F.  
Deposited on : 2011-10-04  
Resolution : 3.80 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>

---

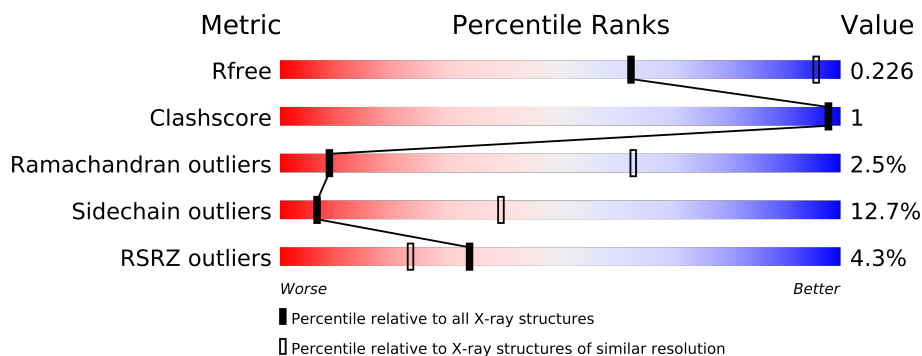
The following versions of software and data (see [references](#)) were used in the production of this report:

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 3.80 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	1162 (4.20-3.40)
Clashscore	79885	1100 (4.10-3.50)
Ramachandran outliers	78287	1050 (4.10-3.50)
Sidechain outliers	78261	1042 (4.10-3.50)
RSRZ outliers	66119	1163 (4.20-3.40)

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	496	
1	C	496	
2	B	95	
2	D	95	
3	E	10	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8722 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 DNA MISMATCH REPAIR PROTEIN HSM3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	455	Total	C	N	O	S	0	0	0
			3716	2396	594	713	13			
1	C	458	Total	C	N	O	S	0	0	0
			3735	2409	596	717	13			

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	EXPRESSION TAG	UNP P38348
A	-3	ALA	-	EXPRESSION TAG	UNP P38348
A	-2	MET	-	EXPRESSION TAG	UNP P38348
A	-1	ALA	-	EXPRESSION TAG	UNP P38348
A	0	ASP	-	EXPRESSION TAG	UNP P38348
A	1	PRO	-	EXPRESSION TAG	UNP P38348
A	481	SER	-	EXPRESSION TAG	UNP P38348
A	482	GLY	-	EXPRESSION TAG	UNP P38348
A	483	SER	-	EXPRESSION TAG	UNP P38348
A	484	GLY	-	EXPRESSION TAG	UNP P38348
A	485	GLY	-	EXPRESSION TAG	UNP P38348
A	486	LEU	-	EXPRESSION TAG	UNP P38348
A	487	GLU	-	EXPRESSION TAG	UNP P38348
A	488	VAL	-	EXPRESSION TAG	UNP P38348
A	489	LEU	-	EXPRESSION TAG	UNP P38348
A	490	PHE	-	EXPRESSION TAG	UNP P38348
A	491	GLN	-	EXPRESSION TAG	UNP P38348
C	-4	GLY	-	EXPRESSION TAG	UNP P38348
C	-3	ALA	-	EXPRESSION TAG	UNP P38348
C	-2	MET	-	EXPRESSION TAG	UNP P38348
C	-1	ALA	-	EXPRESSION TAG	UNP P38348
C	0	ASP	-	EXPRESSION TAG	UNP P38348
C	1	PRO	-	EXPRESSION TAG	UNP P38348
C	481	SER	-	EXPRESSION TAG	UNP P38348
C	482	GLY	-	EXPRESSION TAG	UNP P38348

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	483	SER	-	EXPRESSION TAG	UNP P38348
C	484	GLY	-	EXPRESSION TAG	UNP P38348
C	485	GLY	-	EXPRESSION TAG	UNP P38348
C	486	LEU	-	EXPRESSION TAG	UNP P38348
C	487	GLU	-	EXPRESSION TAG	UNP P38348
C	488	VAL	-	EXPRESSION TAG	UNP P38348
C	489	LEU	-	EXPRESSION TAG	UNP P38348
C	490	PHE	-	EXPRESSION TAG	UNP P38348
C	491	GLN	-	EXPRESSION TAG	UNP P38348

- Molecule 2 is a protein called 26S PROTEASE REGULATORY SUBUNIT 7 HOMOLOG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	79	Total	C	N	O	S	0	0	0
			618	386	117	111	4			
2	D	77	Total	C	N	O	S	0	0	0
			603	377	114	108	4			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	373	GLY	-	EXPRESSION TAG	UNP P33299
B	374	PRO	-	EXPRESSION TAG	UNP P33299
B	375	GLY	-	EXPRESSION TAG	UNP P33299
B	376	GLY	-	EXPRESSION TAG	UNP P33299
B	377	SER	-	EXPRESSION TAG	UNP P33299
D	373	GLY	-	EXPRESSION TAG	UNP P33299
D	374	PRO	-	EXPRESSION TAG	UNP P33299
D	375	GLY	-	EXPRESSION TAG	UNP P33299
D	376	GLY	-	EXPRESSION TAG	UNP P33299
D	377	SER	-	EXPRESSION TAG	UNP P33299

- Molecule 3 is a protein called LINKER.

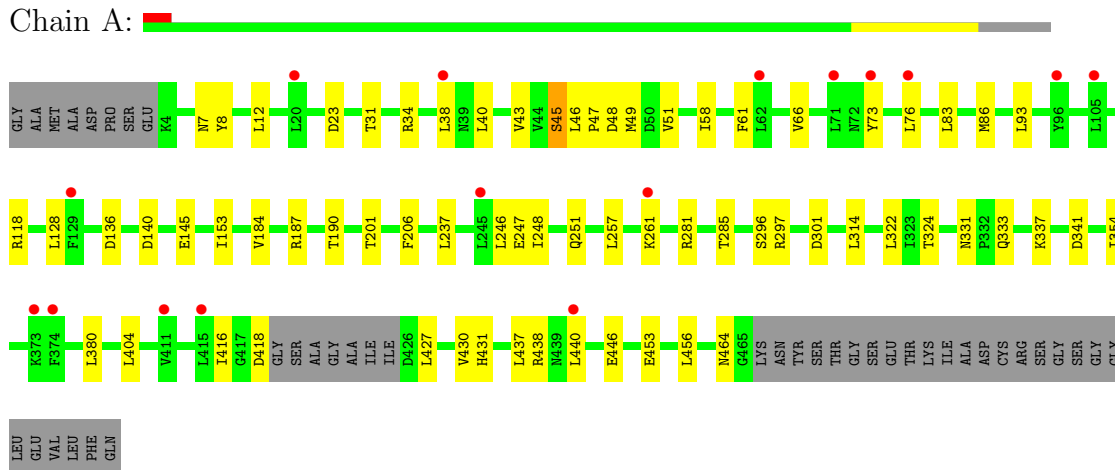
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	E	10	Total	C	N	O	0	0	0
			50	30	10	10			

### 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 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.

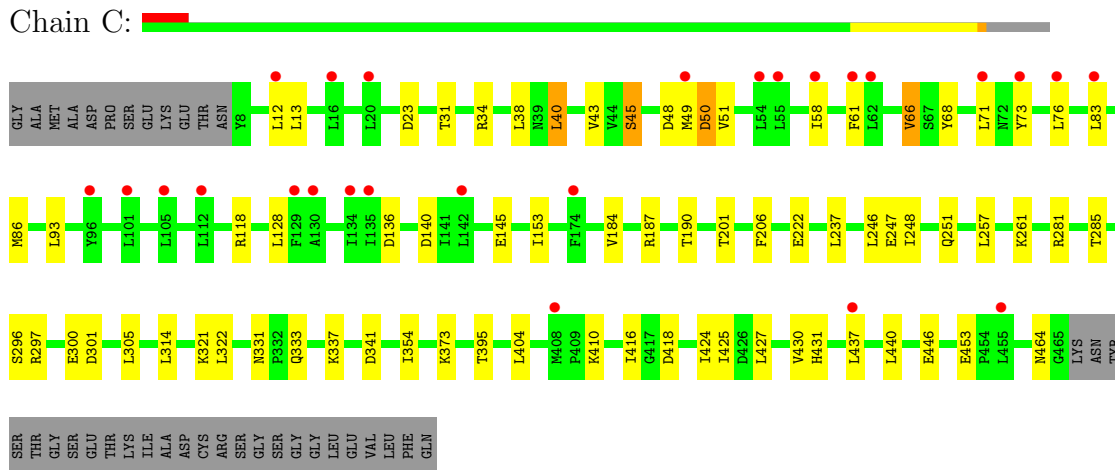
#### • Molecule 1: DNA MISMATCH REPAIR PROTEIN HSM3

Chain A:



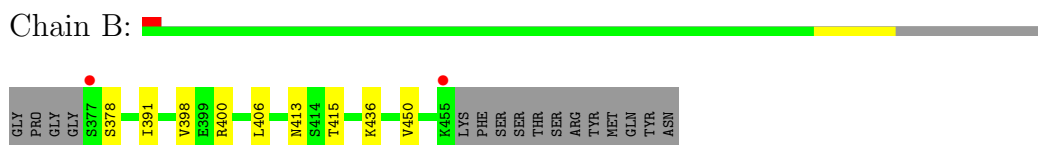
#### • Molecule 1: DNA MISMATCH REPAIR PROTEIN HSM3

Chain C:



#### • Molecule 2: 26S PROTEASE REGULATORY SUBUNIT 7 HOMOLOG

Chain B:



- Molecule 2: 26S PROTEASE REGULATORY SUBUNIT 7 HOMOLOG

Chain D:



- Molecule 3: LINKER

Chain E:



There are no outlier residues recorded for this chain.

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	186.66Å 186.66Å 373.77Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	58.35 – 3.80 58.35 – 3.80	Depositor EDS
% Data completeness (in resolution range)	(Not available) (58.35-3.80) 100.0 (58.35-3.80)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.46 (at 3.77Å)	Xtriage
Refinement program	BUSTER 2.8.0	Depositor
R, $R_{free}$	0.190 , 0.208 0.210 , 0.226	Depositor DCC
$R_{free}$ test set	1933 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	179.6	Xtriage
Anisotropy	0.237	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 236.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 38668 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8722	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	225.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.53% 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

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	0.51	0/3782	0.75	1/5125 (0.0%)
1	C	0.50	0/3802	0.75	0/5152
2	B	0.52	0/626	0.75	0/837
2	D	0.48	0/611	0.71	0/818
All	All	0.51	0/8821	0.75	1/11932 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	7	ASN	C-N-CA	5.34	135.06	121.70

There are no chirality outliers.

There are no planarity outliers.

### 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	3716	0	15	2	0
1	C	3735	0	278	4	0
2	B	618	0	0	0	0
2	D	603	0	0	1	0
3	E	50	0	0	0	0
All	All	8722	0	293	6	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 (6) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:68:TYR:HB2	1:C:71:LEU:HB3	1.91	0.52
1:A:296:SER:O	1:A:331:ASN:CB	2.66	0.43
1:C:296:SER:O	1:C:331:ASN:CB	2.66	0.43
1:A:456:LEU:CG	2:D:428:MET:CE	2.98	0.41
1:C:13:LEU:HD21	1:C:40:LEU:HD11	2.01	0.41
1:C:300:GLU:OE1	1:C:305:LEU:N	2.54	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	451/496 (91%)	413 (92%)	25 (6%)	13 (3%)	7	60
1	C	456/496 (92%)	418 (92%)	26 (6%)	12 (3%)	8	62
2	B	77/95 (81%)	68 (88%)	8 (10%)	1 (1%)	18	76
2	D	75/95 (79%)	67 (89%)	7 (9%)	1 (1%)	18	76
All	All	1059/1182 (90%)	966 (91%)	66 (6%)	27 (2%)	9	63

All (27) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	8	TYR
1	A	45	SER
1	A	48	ASP
1	A	51	VAL
1	A	66	VAL
1	A	301	ASP
1	A	354	ILE
2	B	400	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	45	SER
1	C	48	ASP
1	C	51	VAL
1	C	66	VAL
1	C	301	ASP
1	C	354	ILE
2	D	400	ARG
1	A	43	VAL
1	A	47	PRO
1	C	43	VAL
1	A	261	LYS
1	C	261	LYS
1	C	425	ILE
1	A	46	LEU
1	A	190	THR
1	C	73	TYR
1	C	190	THR
1	A	73	TYR
1	C	50	ASP

### 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	422/455 (93%)	370 (88%)	52 (12%)	7	41
1	C	424/455 (93%)	367 (87%)	57 (13%)	6	37
2	B	66/80 (82%)	58 (88%)	8 (12%)	7	42
2	D	64/80 (80%)	57 (89%)	7 (11%)	9	48
All	All	976/1070 (91%)	852 (87%)	124 (13%)	6	39

All (124) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	LEU
1	A	23	ASP
1	A	31	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	34	ARG
1	A	38	LEU
1	A	40	LEU
1	A	45	SER
1	A	49	MET
1	A	58	ILE
1	A	61	PHE
1	A	76	LEU
1	A	83	LEU
1	A	86	MET
1	A	93	LEU
1	A	118	ARG
1	A	128	LEU
1	A	136	ASP
1	A	140	ASP
1	A	145	GLU
1	A	153	ILE
1	A	184	VAL
1	A	187	ARG
1	A	201	THR
1	A	206	PHE
1	A	237	LEU
1	A	246	LEU
1	A	247	GLU
1	A	248	ILE
1	A	251	GLN
1	A	257	LEU
1	A	281	ARG
1	A	285	THR
1	A	297	ARG
1	A	314	LEU
1	A	322	LEU
1	A	324	THR
1	A	333	GLN
1	A	337	LYS
1	A	341	ASP
1	A	380	LEU
1	A	404	LEU
1	A	416	ILE
1	A	418	ASP
1	A	427	LEU
1	A	430	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	431	HIS
1	A	437	LEU
1	A	438	ARG
1	A	440	LEU
1	A	446	GLU
1	A	453	GLU
1	A	464	ASN
2	B	378	SER
2	B	391	ILE
2	B	398	VAL
2	B	406	LEU
2	B	413	ASN
2	B	415	THR
2	B	436	LYS
2	B	450	VAL
1	C	12	LEU
1	C	23	ASP
1	C	31	THR
1	C	34	ARG
1	C	38	LEU
1	C	40	LEU
1	C	45	SER
1	C	49	MET
1	C	50	ASP
1	C	58	ILE
1	C	61	PHE
1	C	66	VAL
1	C	76	LEU
1	C	83	LEU
1	C	86	MET
1	C	93	LEU
1	C	118	ARG
1	C	128	LEU
1	C	136	ASP
1	C	140	ASP
1	C	145	GLU
1	C	153	ILE
1	C	184	VAL
1	C	187	ARG
1	C	201	THR
1	C	206	PHE
1	C	222	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	237	LEU
1	C	246	LEU
1	C	247	GLU
1	C	248	ILE
1	C	251	GLN
1	C	257	LEU
1	C	281	ARG
1	C	285	THR
1	C	297	ARG
1	C	314	LEU
1	C	321	LYS
1	C	322	LEU
1	C	333	GLN
1	C	337	LYS
1	C	341	ASP
1	C	373	LYS
1	C	395	THR
1	C	404	LEU
1	C	410	LYS
1	C	416	ILE
1	C	418	ASP
1	C	424	ILE
1	C	427	LEU
1	C	430	VAL
1	C	431	HIS
1	C	437	LEU
1	C	440	LEU
1	C	446	GLU
1	C	453	GLU
1	C	464	ASN
2	D	391	ILE
2	D	398	VAL
2	D	400	ARG
2	D	406	LEU
2	D	413	ASN
2	D	415	THR
2	D	450	VAL

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 ⓘ

There are no ligands in this entry.

### 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	455/496 (91%)	0.51	16 (3%) 42 28	142, 211, 286, 300	0
1	C	458/496 (92%)	0.60	26 (5%) 23 16	146, 235, 300, 300	0
2	B	79/95 (83%)	0.49	2 (2%) 54 36	149, 198, 283, 290	0
2	D	77/95 (81%)	0.40	2 (2%) 53 35	163, 219, 285, 293	0
3	E	0/10	-	-	-	-
All	All	1069/1192 (89%)	0.54	46 (4%) 34 23	142, 219, 298, 300	0

All (46) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	62	LEU	4.1
1	C	20	LEU	3.9
1	C	12	LEU	3.2
1	C	96	TYR	3.2
1	C	129	PHE	3.1
1	C	105	LEU	2.8
1	C	54	LEU	2.7
2	D	454	TYR	2.6
1	A	71	LEU	2.6
1	A	373	LYS	2.6
2	D	378	SER	2.6
1	C	76	LEU	2.5
1	A	261	LYS	2.5
1	A	96	TYR	2.5
1	C	134	ILE	2.5
1	C	455	LEU	2.5
1	C	55	LEU	2.5
1	A	73	TYR	2.4
1	C	83	LEU	2.4
1	C	16	LEU	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	101	LEU	2.4
1	C	142	LEU	2.4
1	A	62	LEU	2.3
1	C	71	LEU	2.3
1	A	105	LEU	2.2
1	A	129	PHE	2.2
1	C	73	TYR	2.2
1	A	20	LEU	2.2
1	C	58	ILE	2.2
1	C	112	LEU	2.1
2	B	455	LYS	2.1
1	C	437	LEU	2.1
1	A	411	VAL	2.1
1	C	408	MET	2.1
2	B	377	SER	2.1
1	A	38	LEU	2.1
1	C	130	ALA	2.1
1	A	415	LEU	2.1
1	C	61	PHE	2.1
1	A	245	LEU	2.1
1	C	135	ILE	2.1
1	A	374	PHE	2.1
1	A	76	LEU	2.0
1	C	174	PHE	2.0
1	A	440	LEU	2.0
1	C	49	MET	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 ⓘ

There are no ligands in this entry.



## 6.5 Other polymers ⓘ

There are no such residues in this entry.