



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 28, 2017 – 04:31 PM EST

PDB ID : 5WBW  
Title : Yeast Hsp104 fragment 1-360  
Authors : Lee, S.  
Deposited on : 2017-06-29  
Resolution : 2.60 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  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030736  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : rb-20030736

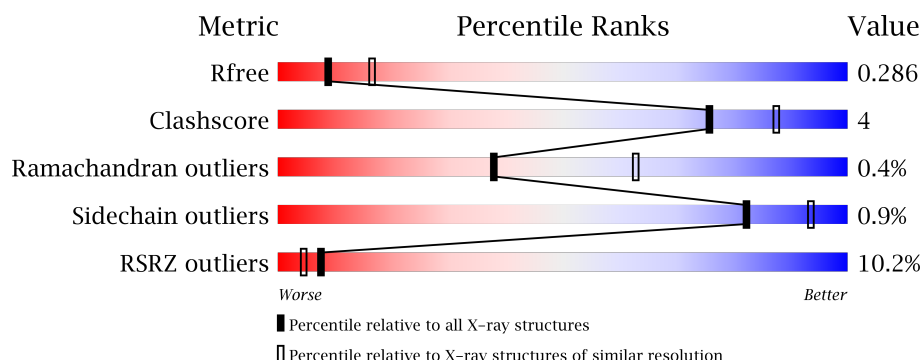
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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))
$R_{free}$	100719	2542 (2.60-2.60)
Clashscore	112137	2895 (2.60-2.60)
Ramachandran outliers	110173	2848 (2.60-2.60)
Sidechain outliers	110143	2848 (2.60-2.60)
RSRZ outliers	101464	2550 (2.60-2.60)

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	353	
1	B	353	
1	D	353	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7886 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 Heat shock protein 104.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	338	Total	C	N	O	S	0	0	0
			2621	1662	457	498	4			
1	B	351	Total	C	N	O	S	0	0	0
			2636	1673	452	508	3			
1	D	345	Total	C	N	O	S	0	0	0
			2605	1656	448	497	4			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	356	GLU	GLN	conflict	UNP P31539
B	356	GLU	GLN	conflict	UNP P31539
D	356	GLU	GLN	conflict	UNP P31539

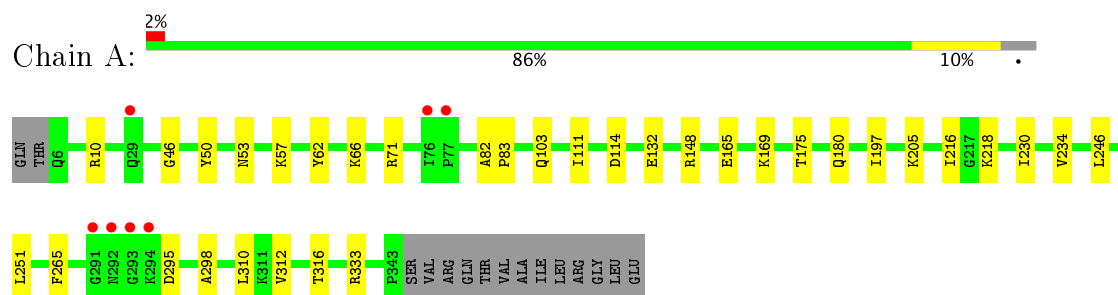
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	15	Total	O	0	0
			15	15		
2	B	6	Total	O	0	0
			6	6		
2	D	3	Total	O	0	0
			3	3		

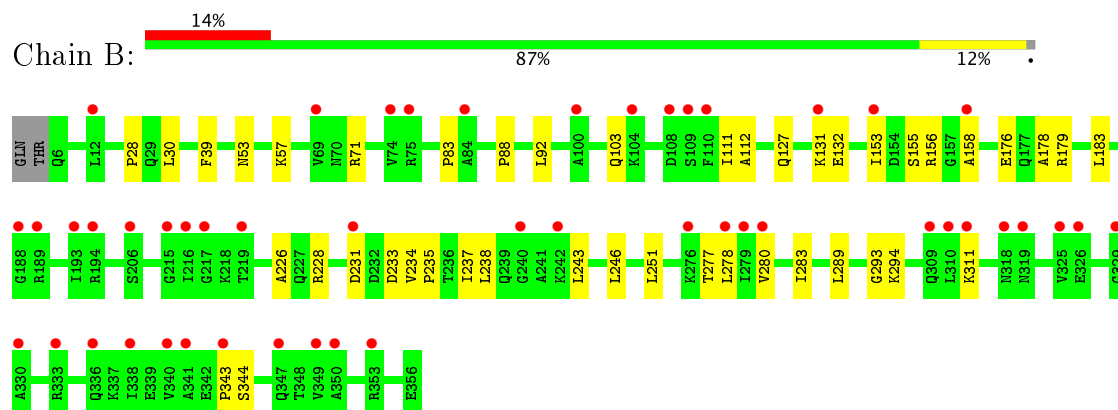
### 3 Residue-property plots [i](#)

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

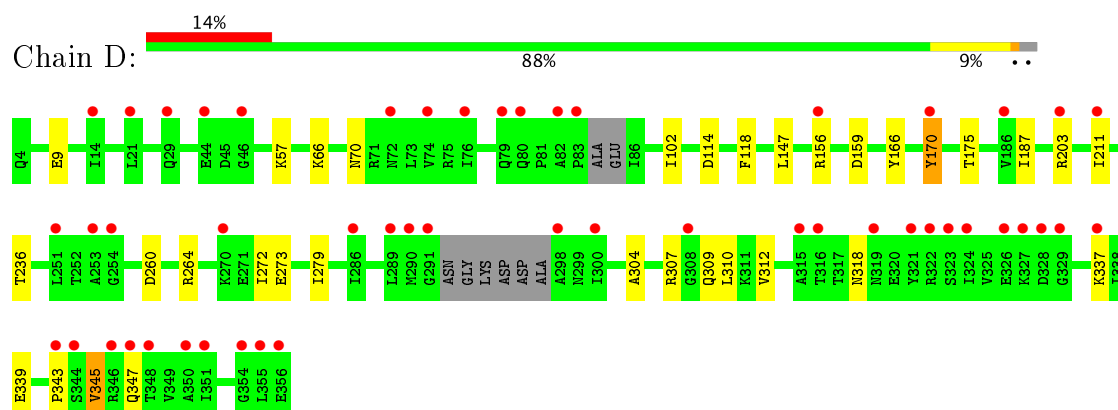
- Molecule 1: Heat shock protein 104



- Molecule 1: Heat shock protein 104



- Molecule 1: Heat shock protein 104



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.56 Å 75.82 Å 235.65 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.04 – 2.60 46.04 – 2.60	Depositor EDS
% Data completeness (in resolution range)	97.8 (46.04-2.60) 97.8 (46.04-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.20 (at 2.61 Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.237 , 0.286 0.237 , 0.286	Depositor DCC
$R_{free}$ test set	2586 reflections (7.03%)	DCC
Wilson B-factor (Å <sup>2</sup> )	64.5	Xtriage
Anisotropy	0.033	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 57.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	7886	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.62% 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, 0.333 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	0.21	0/2661	0.36	0/3600
1	B	0.21	0/2676	0.38	0/3631
1	D	0.20	0/2641	0.37	0/3581
All	All	0.21	0/7978	0.37	0/10812

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 [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	2621	0	2642	19	0
1	B	2636	0	2585	23	0
1	D	2605	0	2585	16	0
2	A	15	0	0	0	0
2	B	6	0	0	1	0
2	D	3	0	0	0	0
All	All	7886	0	7812	56	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 4.

All (56) 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:112:ALA:HB2	1:B:153:ILE:HG21	1.78	0.66
1:B:243:LEU:HG	1:B:280:VAL:HB	1.82	0.62
1:B:155:SER:HB2	1:B:158:ALA:HB2	1.84	0.59
1:B:176:GLU:OE2	1:B:179:ARG:NH2	2.35	0.59
1:D:272:ILE:HG23	1:D:279:ILE:HB	1.85	0.58
1:A:218:LYS:NZ	1:A:316:THR:O	2.37	0.58
1:B:156:ARG:NH1	2:B:401:HOH:O	2.38	0.57
1:D:57:LYS:HE2	1:D:147:LEU:HD21	1.86	0.57
1:A:10:ARG:NH1	1:A:165:GLU:OE1	2.37	0.57
1:B:178:ALA:HB2	1:B:183:LEU:HD12	1.87	0.57
1:A:50:TYR:OH	1:A:114:ASP:OD1	2.16	0.56
1:B:283:ILE:HD13	1:B:289:LEU:HD11	1.88	0.56
1:D:66:LYS:O	1:D:70:ASN:ND2	2.33	0.56
1:B:88:PRO:HB3	1:B:92:LEU:HD23	1.87	0.55
1:D:273:GLU:O	1:D:309:GLN:NE2	2.38	0.55
1:D:9:GLU:HG3	1:D:156:ARG:HB3	1.87	0.55
1:D:304:ALA:HA	1:D:307:ARG:HE	1.73	0.54
1:B:226:ALA:HB2	1:B:243:LEU:HD13	1.90	0.54
1:B:237:ILE:HG23	1:B:238:LEU:HG	1.91	0.53
1:B:228:ARG:NH2	1:B:233:ASP:O	2.40	0.53
1:A:53:ASN:O	1:A:57:LYS:HB2	2.09	0.52
1:A:46:GLY:O	1:D:203:ARG:NH1	2.44	0.51
1:A:103:GLN:HG2	1:A:111:ILE:HG12	1.94	0.49
1:B:277:THR:OG1	1:B:278:LEU:N	2.45	0.49
1:A:175:THR:HB	1:A:230:ILE:HD11	1.95	0.48
1:A:62:TYR:CE2	1:A:66:LYS:HD2	2.48	0.48
1:B:28:PRO:HA	1:B:83:PRO:HG2	1.95	0.47
1:B:53:ASN:O	1:B:57:LYS:HB2	2.14	0.47
1:B:228:ARG:HA	1:B:231:ASP:HB3	1.97	0.47
1:B:127:GLN:O	1:B:131:LYS:HG2	2.14	0.46
1:A:71:ARG:NH2	1:A:132:GLU:OE2	2.35	0.46
1:A:197:ILE:HD13	1:A:234:VAL:HG12	1.97	0.46
1:B:71:ARG:NH2	1:B:132:GLU:OE2	2.44	0.45
1:A:205:LYS:HE2	1:A:333:ARG:HA	1.97	0.45
1:B:280:VAL:HG22	1:B:311:LYS:HE2	1.98	0.45
1:B:293:GLY:HA2	1:B:294:LYS:HA	1.64	0.45
1:D:260:ASP:O	1:D:264:ARG:HG3	2.17	0.44
1:D:159:ASP:OD1	1:D:159:ASP:N	2.47	0.44
1:A:82:ALA:HA	1:A:83:PRO:HD3	1.81	0.43
1:B:234:VAL:HG22	1:B:235:PRO:HD2	2.00	0.43
1:D:211:ILE:HG22	1:D:339:GLU:HA	1.99	0.43
1:B:30:LEU:HB3	1:B:88:PRO:HA	2.01	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:216:ILE:HD12	1:A:216:ILE:HA	1.90	0.42
1:A:295:ASP:HA	1:A:298:ALA:HB3	2.02	0.42
1:A:148:ARG:HA	1:A:148:ARG:HD3	1.81	0.42
1:D:310:LEU:HG	1:D:312:VAL:HG23	2.01	0.42
1:B:103:GLN:HG2	1:B:111:ILE:HG12	2.01	0.41
1:D:211:ILE:HG12	1:D:337:LYS:HD2	2.02	0.41
1:A:246:LEU:HD11	1:A:251:LEU:HD11	2.02	0.41
1:D:102:ILE:HD13	1:D:118:PHE:HB3	2.03	0.41
1:A:180:GLN:HG3	1:D:236:THR:OG1	2.20	0.41
1:D:166:TYR:HD1	1:D:170:TYR:HE2	1.69	0.41
1:D:187:ILE:O	1:D:345:VAL:HG21	2.21	0.41
1:B:246:LEU:HD11	1:B:251:LEU:HD11	2.03	0.40
1:A:165:GLU:HG2	1:A:169:LYS:HE2	2.03	0.40
1:A:310:LEU:HG	1:A:312:VAL:HG23	2.03	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	336/353 (95%)	331 (98%)	5 (2%)	0	100	100
1	B	349/353 (99%)	331 (95%)	16 (5%)	2 (1%)	28	53
1	D	339/353 (96%)	326 (96%)	11 (3%)	2 (1%)	28	53
All	All	1024/1059 (97%)	988 (96%)	32 (3%)	4 (0%)	38	63

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	343	PRO
1	D	318	ASN
1	D	343	PRO

Continued on next page...



*Continued from previous page...*

Mol	Chain	Res	Type
1	B	344	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	274/299 (92%)	273 (100%)	1 (0%)	93	98
1	B	265/299 (89%)	264 (100%)	1 (0%)	93	98
1	D	266/299 (89%)	261 (98%)	5 (2%)	62	84
All	All	805/897 (90%)	798 (99%)	7 (1%)	82	93

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

Mol	Chain	Res	Type
1	A	265	PHE
1	B	39	PHE
1	D	114	ASP
1	D	170	TYR
1	D	175	THR
1	D	345	VAL
1	D	347	GLN

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 molecules 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 [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

### 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	338/353 (95%)	-0.08	7 (2%) 64 58	35, 61, 125, 207	0
1	B	351/353 (99%)	0.69	48 (13%) 3 2	43, 117, 184, 250	0
1	D	345/353 (97%)	0.72	50 (14%) 3 1	41, 101, 202, 268	0
All	All	1034/1059 (97%)	0.45	105 (10%) 7 4	35, 92, 181, 268	0

All (105) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	351	ILE	12.0
1	B	325	VAL	11.4
1	D	291	GLY	9.4
1	B	347	GLN	7.8
1	D	326	GLU	7.1
1	A	292	ASN	6.7
1	D	344	SER	6.6
1	D	290	MET	5.8
1	B	189	ARG	5.4
1	A	294	LYS	5.4
1	D	327	LYS	5.4
1	D	355	LEU	5.3
1	A	291	GLY	5.3
1	D	308	GLY	5.1
1	B	350	ALA	5.0
1	D	324	ILE	4.8
1	D	82	ALA	4.7
1	D	348	THR	4.5
1	D	346	ARG	4.3
1	D	356	GLU	4.3
1	B	109	SER	4.2
1	B	206	SER	4.2
1	D	44	GLU	4.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	293	GLY	4.0
1	D	254	GLY	4.0
1	B	75	ARG	3.9
1	D	156	ARG	3.9
1	D	170	TYR	3.8
1	D	316	THR	3.8
1	D	354	GLY	3.7
1	B	100	ALA	3.7
1	B	215	GLY	3.7
1	B	329	GLY	3.6
1	D	76	ILE	3.6
1	B	309	GLN	3.3
1	B	341	ALA	3.3
1	B	353	ARG	3.3
1	D	350	ALA	3.3
1	B	110	PHE	3.3
1	B	330	ALA	3.2
1	B	188	GLY	3.2
1	D	298	ALA	3.2
1	B	12	LEU	3.1
1	B	333	ARG	3.1
1	B	276	LYS	3.1
1	D	322	ARG	3.0
1	D	72	ASN	3.0
1	A	77	PRO	2.9
1	D	289	LEU	2.9
1	B	217	GLY	2.9
1	D	211	ILE	2.8
1	B	338	ILE	2.8
1	D	80	GLN	2.8
1	B	343	PRO	2.8
1	D	323	SER	2.8
1	B	340	VAL	2.7
1	B	311	LYS	2.7
1	B	74	VAL	2.7
1	B	318	ASN	2.7
1	B	242	LYS	2.6
1	B	279	ILE	2.6
1	B	310	LEU	2.6
1	B	240	GLY	2.5
1	D	329	GLY	2.5
1	D	74	VAL	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	315	ALA	2.5
1	D	29	GLN	2.4
1	B	231	ASP	2.4
1	B	131	LYS	2.4
1	D	253	ALA	2.4
1	D	286	ILE	2.4
1	B	319	ASN	2.4
1	D	270	LYS	2.4
1	D	300	ILE	2.3
1	D	347	GLN	2.3
1	D	83	PRO	2.3
1	B	278	LEU	2.3
1	B	216	ILE	2.3
1	D	21	LEU	2.3
1	D	321	TYR	2.3
1	D	337	LYS	2.3
1	A	29	GLN	2.3
1	B	219	THR	2.2
1	B	153	ILE	2.2
1	B	193	ILE	2.2
1	B	349	VAL	2.2
1	D	328	ASP	2.2
1	D	14	ILE	2.2
1	D	203	ARG	2.2
1	B	104	LYS	2.2
1	B	336	GLN	2.1
1	D	251	LEU	2.1
1	B	194	ARG	2.1
1	A	76	ILE	2.1
1	D	186	VAL	2.1
1	B	84	ALA	2.1
1	B	280	VAL	2.1
1	D	79	GLN	2.1
1	B	158	ALA	2.1
1	B	69	VAL	2.0
1	B	326	GLU	2.0
1	D	343	PRO	2.0
1	D	319	ASN	2.0
1	B	108	ASP	2.0
1	D	46	GLY	2.0

## 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.