



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 13, 2020 – 10:00 AM BST

PDB ID : 1USI  
Title : L-leucine-binding protein with phenylalanine bound  
Authors : Magnusson, U.; Salopek-Sondi, B.; Luck, L.A.; Mowbray, S.L.  
Deposited on : 2003-11-24  
Resolution : 1.80 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

<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 : 2.14.4.dev1  
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) : 2.14.4.dev1

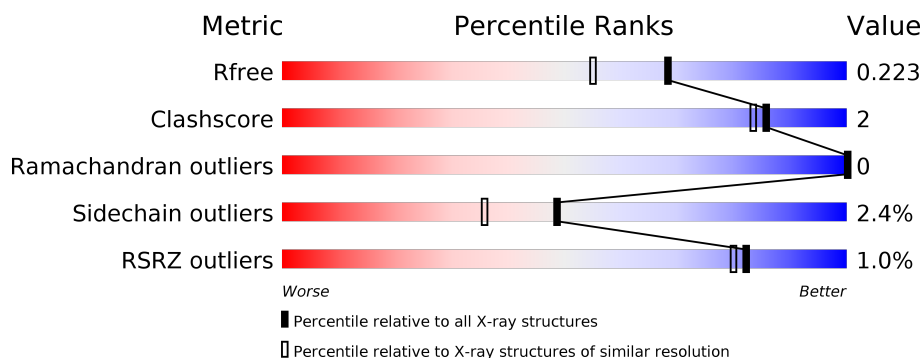
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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))
$R_{free}$	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	346	<div> <div style="width: 100%; height: 10px; background-color: red; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">%</span> </div> <div style="width: 88%; height: 10px; background-color: green; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">88%</span> </div> <div style="width: 11%; height: 10px; background-color: yellow; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">11%</span> </div> </div>
1	C	346	<div> <div style="width: 100%; height: 10px; background-color: red; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">%</span> </div> <div style="width: 90%; height: 10px; background-color: green; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">90%</span> </div> <div style="width: 9%; height: 10px; background-color: yellow; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">9%</span> </div> </div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 5515 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.

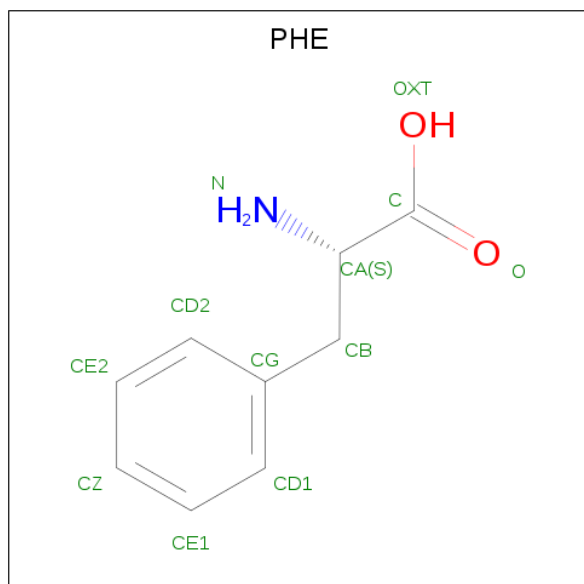
- Molecule 1 is a protein called LEUCINE-SPECIFIC BINDING PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	346	Total	C	N	O	S	0	0	1
			2596	1630	445	510	11			
1	C	346	Total	C	N	O	S	0	0	1
			2596	1630	445	510	11			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	344	LYS	ALA	conflict	UNP P04816
C	344	LYS	ALA	conflict	UNP P04816

- Molecule 2 is PHENYLALANINE (three-letter code: PHE) (formula:  $C_9H_{11}NO_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			12	9	1	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	C	1	Total	C	N	O	0	0
			12	9	1	2		

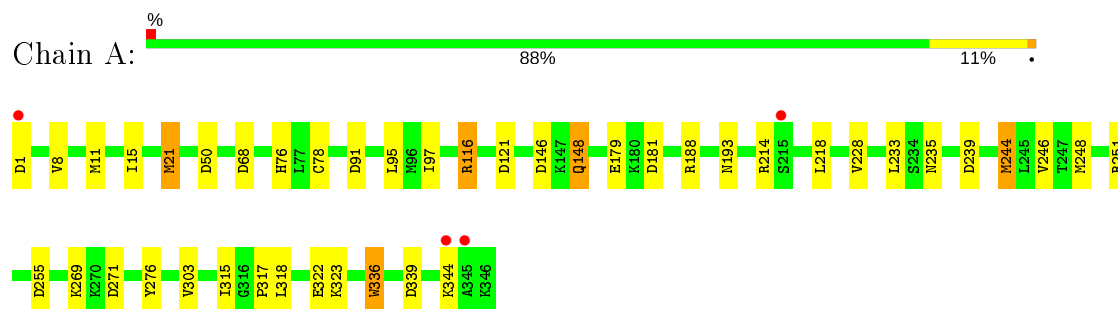
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	151	Total	O	0	0
			151	151		
3	C	148	Total	O	0	0
			148	148		

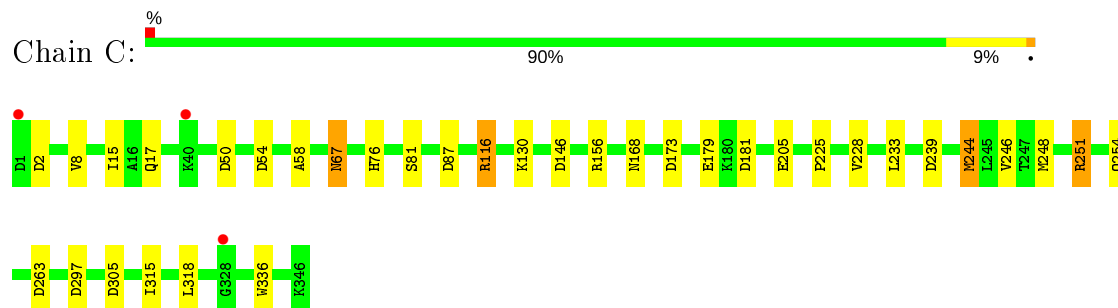
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: LEUCINE-SPECIFIC BINDING PROTEIN



#### • Molecule 1: LEUCINE-SPECIFIC BINDING PROTEIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.24Å 78.14Å 70.18Å 90.00° 102.82° 90.00°	Depositor
Resolution (Å)	35.00 – 1.80 31.34 – 1.80	Depositor EDS
% Data completeness (in resolution range)	96.2 (35.00-1.80) 96.3 (31.34-1.80)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.91 (at 1.80Å)	Xtriage
Refinement program	REFMAC 5.1.19	Depositor
R, $R_{free}$	0.189 , 0.222 0.194 , 0.223	Depositor DCC
$R_{free}$ test set	3242 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.3	Xtriage
Anisotropy	0.209	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 41.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.019 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5515	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.56% 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

### 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	1.08	6/2644 (0.2%)	1.06	18/3578 (0.5%)
1	C	1.05	3/2644 (0.1%)	1.05	14/3578 (0.4%)
All	All	1.06	9/5288 (0.2%)	1.05	32/7156 (0.4%)

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	116	ARG	CB-CG	-8.62	1.29	1.52
1	A	21	MET	SD-CE	-7.28	1.37	1.77
1	A	322	GLU	CG-CD	6.74	1.62	1.51
1	A	244	MET	CB-CG	6.20	1.71	1.51
1	A	322	GLU	CD-OE1	5.92	1.32	1.25
1	C	244	MET	CB-CG	5.89	1.70	1.51
1	C	205	GLU	CD-OE1	5.72	1.31	1.25
1	A	323	LYS	CD-CE	5.35	1.64	1.51
1	A	303	VAL	CB-CG2	-5.30	1.41	1.52

All (32) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	271	ASP	CB-CG-OD2	8.77	126.20	118.30
1	C	2	ASP	CB-CG-OD2	8.55	126.00	118.30
1	C	297	ASP	CB-CG-OD2	8.49	125.94	118.30
1	C	50	ASP	CB-CG-OD2	8.09	125.58	118.30
1	C	263	ASP	CB-CG-OD2	8.08	125.57	118.30
1	A	188	ARG	NE-CZ-NH2	-7.76	116.42	120.30
1	C	87	ASP	CB-CG-OD2	7.68	125.21	118.30
1	A	251	ARG	NE-CZ-NH2	-7.57	116.52	120.30
1	C	54	ASP	CB-CG-OD2	7.37	124.93	118.30
1	C	146	ASP	CB-CG-OD2	7.36	124.93	118.30
1	C	156	ARG	NE-CZ-NH2	-7.02	116.79	120.30
1	C	239	ASP	CB-CG-OD2	6.99	124.59	118.30
1	A	1	ASP	CB-CG-OD2	6.62	124.26	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	255	ASP	CB-CG-OD2	6.55	124.20	118.30
1	A	68	ASP	CB-CG-OD2	6.11	123.80	118.30
1	A	248	MET	CG-SD-CE	5.96	109.73	100.20
1	A	181	ASP	CB-CG-OD2	5.79	123.51	118.30
1	A	248	MET	CB-CG-SD	-5.74	95.19	112.40
1	A	239	ASP	CB-CG-OD1	5.71	123.44	118.30
1	A	244	MET	CB-CG-SD	-5.71	95.28	112.40
1	A	91	ASP	CB-CG-OD2	5.66	123.39	118.30
1	A	146	ASP	CB-CG-OD2	5.62	123.36	118.30
1	C	181	ASP	CB-CG-OD2	5.57	123.31	118.30
1	C	305	ASP	CB-CG-OD2	5.43	123.18	118.30
1	A	214	ARG	NE-CZ-NH2	-5.38	117.61	120.30
1	A	218	LEU	CA-CB-CG	5.38	127.67	115.30
1	C	130	LYS	CD-CE-NZ	-5.33	99.45	111.70
1	A	116	ARG	NE-CZ-NH2	-5.09	117.75	120.30
1	A	121	ASP	CB-CG-OD2	5.05	122.84	118.30
1	A	339	ASP	CB-CG-OD1	5.05	122.84	118.30
1	C	173	ASP	CB-CG-OD1	5.03	122.83	118.30
1	C	248	MET	CB-CG-SD	-5.01	97.37	112.40

There are no chirality outliers.

There are no planarity outliers.

## 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	2596	0	2555	14	0
1	C	2596	0	2555	10	0
2	A	12	0	8	0	0
2	C	12	0	8	0	0
3	A	151	0	0	1	0
3	C	148	0	0	0	0
All	All	5515	0	5126	23	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 2.



All (23) 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:233:LEU:HD21	1:A:244:MET:SD	2.38	0.64
1:A:21:MET:HE2	1:A:276:TYR:HB2	1.84	0.59
1:C:233:LEU:HD21	1:C:244:MET:SD	2.47	0.54
1:A:8:VAL:HG11	1:A:76:HIS:CE1	2.43	0.52
1:A:11:MET:CE	1:A:50:ASP:HB2	2.39	0.52
1:A:15:ILE:HD11	1:A:78:CYS:SG	2.50	0.51
1:A:315:ILE:HG13	1:A:318:LEU:HD21	1.92	0.51
1:A:228:VAL:O	1:A:233:LEU:HD22	2.12	0.48
1:A:244:MET:CE	1:A:246:VAL:HB	2.44	0.47
1:A:95:LEU:HD21	1:A:97:ILE:HG13	1.98	0.46
1:C:8:VAL:HG11	1:C:76:HIS:CE1	2.52	0.45
1:A:148:GLN:HG3	3:A:2042:HOH:O	2.17	0.44
1:A:148:GLN:HE21	1:A:148:GLN:CA	2.29	0.44
1:C:251:ARG:HD3	1:C:254:GLN:HE22	1.81	0.44
1:C:15:ILE:HG22	1:C:15:ILE:O	2.18	0.43
1:C:246:VAL:HG11	1:C:336:TRP:CE3	2.54	0.43
1:C:225:PRO:O	1:C:228:VAL:HG22	2.19	0.43
1:A:246:VAL:HG11	1:A:336:TRP:CZ3	2.53	0.43
1:C:315:ILE:HG13	1:C:318:LEU:HD21	2.01	0.42
1:A:235:ASN:O	1:C:17:GLN:HG3	2.19	0.42
1:A:148:GLN:HA	1:A:148:GLN:HE21	1.85	0.42
1:C:67:ASN:HD22	1:C:67:ASN:HA	1.64	0.40
1:C:58:ALA:HB2	1:C:81:SER:HB2	2.04	0.40

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	344/346 (99%)	339 (98%)	5 (2%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	344/346 (99%)	341 (99%)	3 (1%)	0	100	100
All	All	688/692 (99%)	680 (99%)	8 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	267/268 (100%)	259 (97%)	8 (3%)	41	27
1	C	267/268 (100%)	262 (98%)	5 (2%)	57	46
All	All	534/536 (100%)	521 (98%)	13 (2%)	49	36

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	116	ARG
1	A	148	GLN
1	A	179	GLU
1	A	193	ASN
1	A	269	LYS
1	A	317	PRO
1	A	336	TRP
1	A	344	LYS
1	C	67	ASN
1	C	116	ARG
1	C	168	ASN
1	C	179	GLU
1	C	251	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (15) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	17	GLN
1	A	83	GLN
1	A	148	GLN
1	A	149	GLN
1	A	159	GLN
1	A	168	ASN
1	A	235	ASN
1	A	254	GLN
1	C	67	ASN
1	C	83	GLN
1	C	139	GLN
1	C	149	GLN
1	C	168	ASN
1	C	235	ASN
1	C	254	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	PHE	C	1346	-	9,12,12	1.32	1 (11%)	10,15,15	1.08	1 (10%)
2	PHE	A	1346	-	9,12,12	0.74	0	10,15,15	0.88	1 (10%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PHE	C	1346	-	-	2/4/8/8	0/1/1/1
2	PHE	A	1346	-	-	1/4/8/8	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1346	PHE	CE2-CD2	2.94	1.45	1.38

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1346	PHE	CB-CA-C	-2.55	106.57	110.69
2	C	1346	PHE	CB-CA-C	-2.53	106.59	110.69

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	1346	PHE	CA-CB-CG-CD1
2	C	1346	PHE	CA-CB-CG-CD2
2	A	1346	PHE	CA-CB-CG-CD1

There are no ring outliers.

No monomer is involved in short contacts.

## 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 [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	346/346 (100%)	-0.20	4 (1%) 79 76	16, 23, 33, 44	0
1	C	346/346 (100%)	-0.24	3 (0%) 84 82	15, 23, 36, 45	0
All	All	692/692 (100%)	-0.22	7 (1%) 82 80	15, 23, 35, 45	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	344	LYS	4.0
1	A	1	ASP	3.4
1	C	328	GLY	3.4
1	A	215	SER	2.8
1	C	1	ASP	2.6
1	C	40	LYS	2.3
1	A	345	ALA	2.3

### 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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

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. 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	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	PHE	C	1346	12/12	0.94	0.18	20,22,29,29	0
2	PHE	A	1346	12/12	0.97	0.14	15,19,26,26	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.