



# Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 04:45 pm BST

PDB ID : 2F41  
Title : Crystal structure of FapR- a global regulator of fatty acid biosynthesis in *B. subtilis*  
Authors : Buschiazzo, A.; Guerin, M.E.; Alzari, P.M.  
Deposited on : 2005-11-22  
Resolution : 2.50 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  
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

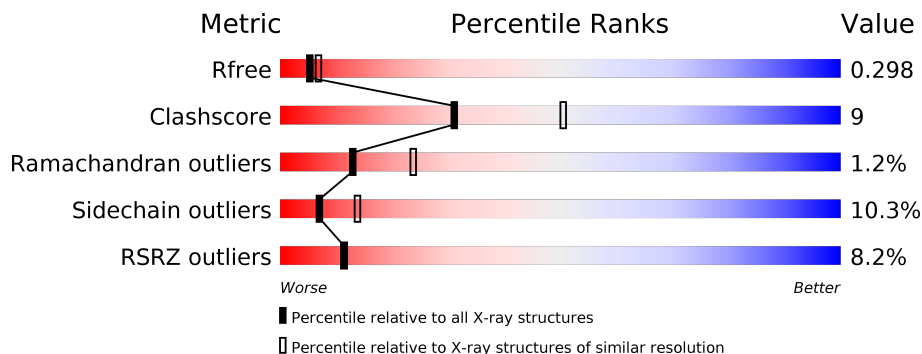
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	121	<div style="display: flex; align-items: center;"> <div style="width: 4%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 72%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 14%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: grey; margin-right: 5px;"></div> </div> <p style="text-align: center;">4%      72%      14%      •      11%</p>
1	B	121	<div style="display: flex; align-items: center;"> <div style="width: 4%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 69%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 17%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: grey; margin-right: 5px;"></div> </div> <p style="text-align: center;">4%      69%      17%      •      10%</p>
1	C	121	<div style="display: flex; align-items: center;"> <div style="width: 16%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 66%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 16%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: grey; margin-right: 5px;"></div> </div> <p style="text-align: center;">16%      66%      16%      •      15%</p>
1	D	121	<div style="display: flex; align-items: center;"> <div style="width: 5%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 71%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 16%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: grey; margin-right: 5px;"></div> </div> <p style="text-align: center;">5%      71%      16%      ••      11%</p>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2968 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 Transcription factor fapR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	108	740	472	130	137	1	0	0	0
1	B	109	772	489	133	149	1	0	0	0
1	C	103	703	448	129	125	1	0	0	0
1	D	108	745	476	135	133	1	0	0	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	O	0	0
			2	2		
2	B	2	Total	O	0	0
			2	2		
2	C	2	Total	O	0	0
			2	2		
2	D	2	Total	O	0	0
			2	2		

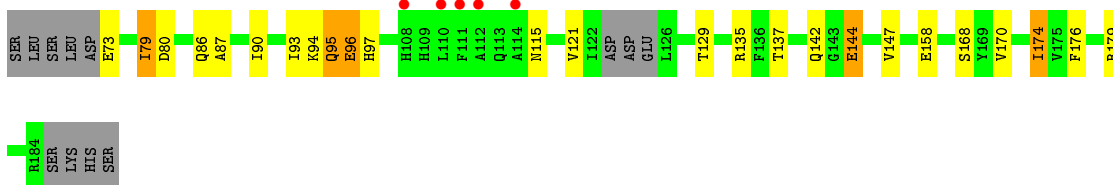
### 3 Residue-property plots [i](#)

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.

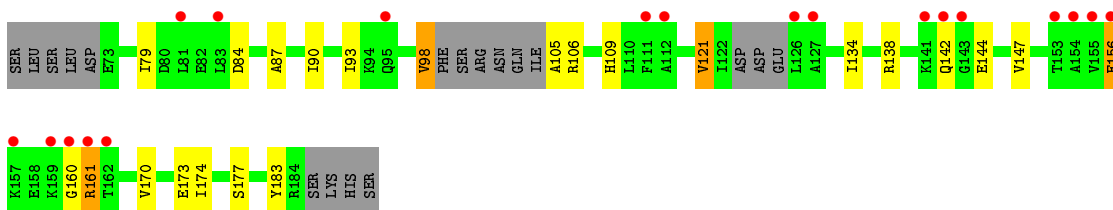
- Molecule 1: Transcription factor fapR



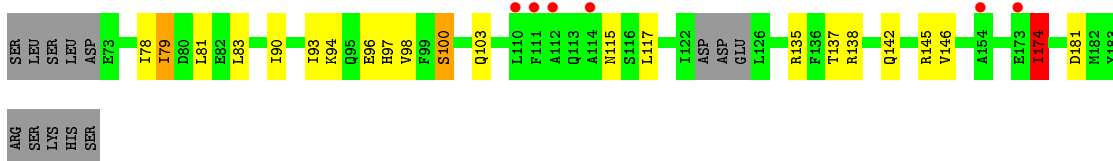
- Molecule 1: Transcription factor fapR



- Molecule 1: Transcription factor fapR



- Molecule 1: Transcription factor fapR



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	39.24Å 84.32Å 155.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.50 29.32 – 2.50	Depositor EDS
% Data completeness (in resolution range)	79.8 (30.00-2.50) 79.8 (29.32-2.50)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.14 (at 2.51Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.219 , 0.274 0.254 , 0.298	Depositor DCC
$R_{free}$ test set	790 reflections (5.33%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	56.5	Xtrriage
Anisotropy	0.097	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 63.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	2968	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	62.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.76% 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	0.94	0/747	0.86	2/1020 (0.2%)
1	B	1.04	1/780 (0.1%)	0.90	0/1062
1	C	0.86	0/708	0.87	0/963
1	D	1.02	1/752 (0.1%)	0.90	1/1024 (0.1%)
All	All	0.97	2/2987 (0.1%)	0.88	3/4069 (0.1%)

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

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	138	ARG	CA-CB	5.39	1.65	1.53
1	B	144	GLU	CB-CG	5.16	1.61	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	174	ILE	CB-CA-C	-5.45	100.71	111.60
1	A	93	ILE	CG1-CB-CG2	-5.10	100.17	111.40
1	A	90	ILE	CB-CA-C	-5.00	101.60	111.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	156	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	740	0	683	11	0
1	B	772	0	714	15	0
1	C	703	0	660	12	0
1	D	745	0	701	16	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
All	All	2968	0	2758	52	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 9.

All (52) 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:D:93:ILE:HG22	1:D:142:GLN:HA	1.66	0.77
1:A:93:ILE:HG22	1:A:142:GLN:HA	1.66	0.76
1:D:103:GLN:NE2	1:D:142:GLN:HE22	1.83	0.76
1:B:94:LYS:H	1:B:97:HIS:HD2	1.33	0.76
1:C:98:VAL:HG13	1:C:105:ALA:N	2.03	0.74
1:C:79:ILE:HG12	1:C:90:ILE:HG12	1.71	0.73
1:D:79:ILE:HG12	1:D:90:ILE:HG12	1.72	0.71
1:B:93:ILE:HG22	1:B:142:GLN:HA	1.77	0.67
1:C:90:ILE:HD12	1:C:147:VAL:HG22	1.78	0.65
1:A:79:ILE:HG12	1:A:90:ILE:HG12	1.78	0.65
1:D:94:LYS:H	1:D:97:HIS:HD2	1.44	0.64
1:D:103:GLN:NE2	1:D:142:GLN:NE2	2.46	0.64
1:B:137:THR:HG21	1:B:174:ILE:HD13	1.81	0.63
1:C:156:GLU:HB2	1:C:161:ARG:HB3	1.81	0.62
1:B:144:GLU:OE2	1:B:170:VAL:HG11	2.01	0.61
1:A:98:VAL:HG21	1:A:142:GLN:HB2	1.82	0.61
1:D:94:LYS:O	1:D:98:VAL:HG23	2.03	0.59
1:D:98:VAL:HG11	1:D:103:GLN:HA	1.86	0.58
1:C:93:ILE:HG22	1:C:142:GLN:HA	1.88	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:106:ARG:HH12	1:B:73:GLU:HA	1.73	0.53
1:B:94:LYS:H	1:B:97:HIS:CD2	2.22	0.53
1:D:137:THR:HG21	1:D:174:ILE:HD13	1.91	0.52
1:D:98:VAL:CG1	1:D:103:GLN:HA	2.39	0.52
1:A:102:ASN:HB3	1:A:104:ILE:HD11	1.93	0.51
1:B:90:ILE:CD1	1:B:147:VAL:HG22	2.42	0.50
1:B:115:ASN:HD21	1:B:129:THR:HG1	1.61	0.49
1:A:155:VAL:HG22	1:A:162:THR:HG23	1.96	0.47
1:C:160:GLY:O	1:C:183:TYR:HA	2.15	0.47
1:D:81:LEU:HD11	1:D:83:LEU:HD23	1.98	0.46
1:C:98:VAL:HG21	1:C:142:GLN:HB2	1.98	0.46
1:D:78:ILE:HD13	1:D:117:LEU:CD1	2.46	0.45
1:D:79:ILE:CG1	1:D:90:ILE:HG12	2.44	0.45
1:B:168:SER:HB2	1:B:176:PHE:HB3	1.98	0.45
1:D:93:ILE:HD13	1:D:146:VAL:HG23	1.98	0.44
1:C:79:ILE:CG1	1:C:90:ILE:HG12	2.45	0.44
1:A:99:PHE:HB3	1:A:106:ARG:HE	1.82	0.44
1:D:98:VAL:CG1	1:D:103:GLN:C	2.86	0.43
1:A:88:ILE:HD13	1:A:88:ILE:HG21	1.80	0.43
1:A:94:LYS:H	1:A:97:HIS:HD2	1.67	0.42
1:A:90:ILE:CD1	1:A:147:VAL:HG22	2.49	0.42
1:C:87:ALA:CB	1:C:121:VAL:HG13	2.49	0.42
1:B:96:GLU:H	1:B:96:GLU:HG3	1.56	0.42
1:B:176:PHE:CD2	1:B:176:PHE:C	2.93	0.42
1:B:80:ASP:O	1:B:87:ALA:HA	2.20	0.42
1:D:94:LYS:H	1:D:97:HIS:CD2	2.30	0.41
1:C:106:ARG:HB2	1:C:109:HIS:CD2	2.55	0.41
1:B:95:GLN:HE21	1:B:95:GLN:HA	1.85	0.41
1:B:79:ILE:HG12	1:B:90:ILE:HG12	2.03	0.41
1:C:144:GLU:OE1	1:C:170:VAL:HG11	2.21	0.41
1:A:131:SER:HB3	1:B:135:ARG:HD3	2.01	0.41
1:C:93:ILE:HD12	1:C:93:ILE:N	2.36	0.40
1:D:90:ILE:HG23	1:D:145:ARG:HG3	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	104/121 (86%)	99 (95%)	4 (4%)	1 (1%)	15	28
1	B	105/121 (87%)	99 (94%)	5 (5%)	1 (1%)	15	28
1	C	97/121 (80%)	93 (96%)	3 (3%)	1 (1%)	15	28
1	D	104/121 (86%)	98 (94%)	4 (4%)	2 (2%)	8	13
All	All	410/484 (85%)	389 (95%)	16 (4%)	5 (1%)	13	24

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	158	GLU
1	C	84	ASP
1	D	100	SER
1	A	158	GLU
1	D	96	GLU

### 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	65/103 (63%)	60 (92%)	5 (8%)	13	25
1	B	72/103 (70%)	65 (90%)	7 (10%)	8	16
1	C	60/103 (58%)	51 (85%)	9 (15%)	3	5
1	D	65/103 (63%)	59 (91%)	6 (9%)	9	18
All	All	262/412 (64%)	235 (90%)	27 (10%)	7	14

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	79	ILE
1	A	86	GLN
1	A	121	VAL
1	A	162	THR
1	A	174	ILE
1	B	79	ILE
1	B	86	GLN
1	B	95	GLN
1	B	96	GLU
1	B	121	VAL
1	B	174	ILE
1	B	179	ARG
1	C	98	VAL
1	C	121	VAL
1	C	134	ILE
1	C	138	ARG
1	C	156	GLU
1	C	161	ARG
1	C	173	GLU
1	C	174	ILE
1	C	177	SER
1	D	79	ILE
1	D	100	SER
1	D	115	ASN
1	D	135	ARG
1	D	174	ILE
1	D	181	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (20) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	97	HIS
1	A	113	GLN
1	A	142	GLN
1	A	167	ASN
1	B	95	GLN
1	B	97	HIS
1	B	113	GLN
1	B	115	ASN
1	B	142	GLN
1	C	97	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	109	HIS
1	C	113	GLN
1	C	115	ASN
1	C	142	GLN
1	C	167	ASN
1	D	97	HIS
1	D	103	GLN
1	D	113	GLN
1	D	115	ASN
1	D	142	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

### 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	108/121 (89%)	0.40	5 (4%) 32 34	51, 63, 72, 83	0
1	B	109/121 (90%)	0.44	5 (4%) 32 34	50, 64, 73, 83	0
1	C	103/121 (85%)	0.89	19 (18%) 1 1	51, 62, 72, 83	0
1	D	108/121 (89%)	0.37	6 (5%) 24 25	50, 63, 72, 83	0
All	All	428/484 (88%)	0.52	35 (8%) 11 11	50, 63, 73, 83	0

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	159	LYS	4.7
1	A	156	GLU	4.7
1	C	95	GLN	4.3
1	C	161	ARG	4.1
1	C	143	GLY	3.9
1	C	155	VAL	3.8
1	D	111	PHE	3.7
1	C	162	THR	3.6
1	A	111	PHE	3.3
1	C	157	LYS	3.3
1	C	160	GLY	3.2
1	C	126	LEU	2.9
1	C	153	THR	2.9
1	B	111	PHE	2.9
1	C	142	GLN	2.9
1	B	112	ALA	2.7
1	D	112	ALA	2.7
1	C	127	ALA	2.7
1	C	154	ALA	2.7
1	C	156	GLU	2.6
1	D	114	ALA	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	81	LEU	2.6
1	B	114	ALA	2.5
1	D	154	ALA	2.4
1	A	155	VAL	2.4
1	A	157	LYS	2.3
1	B	110	LEU	2.3
1	C	111	PHE	2.3
1	B	108	HIS	2.2
1	C	83	LEU	2.2
1	C	112	ALA	2.1
1	D	110	LEU	2.1
1	D	173	GLU	2.1
1	C	141	LYS	2.1
1	A	96	GLU	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.