



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 7, 2025 – 01:22 pm BST

PDB ID : 9FKO / pdb\_00009fko  
Title : Arm segment of bacteriophage phi812  
Authors : Siborova, M.; Baska, R.; Binovsky, J.; Plevka, P.  
Deposited on : 2024-06-03  
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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity	:	4-5-2 with Phenix2.0rc1
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

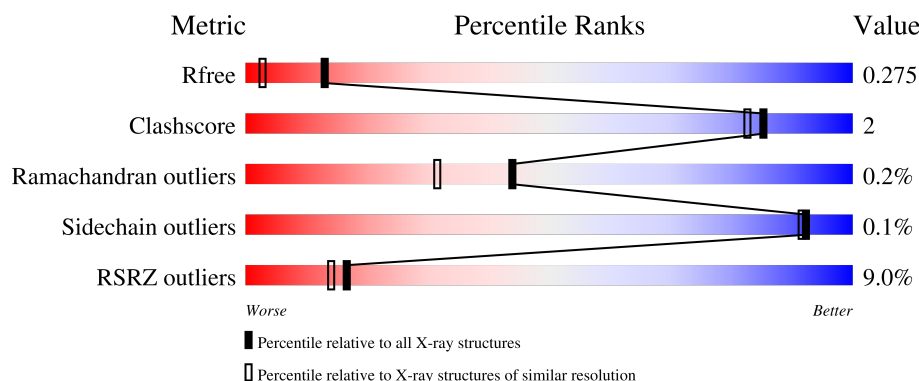
# 1 Overall quality at a glance

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))
$R_{free}$	164625	7108 (1.80-1.80)
Clashscore	180529	8162 (1.80-1.80)
Ramachandran outliers	177936	8077 (1.80-1.80)
Sidechain outliers	177891	8076 (1.80-1.80)
RSRZ outliers	164620	7108 (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 of 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	174	<div> <div>10%</div> <div>96%</div> <div>.</div> </div>
1	B	174	<div> <div>6%</div> <div>95%</div> <div>..</div> </div>
1	C	174	<div> <div>11%</div> <div>97%</div> <div>.</div> </div>
1	D	174	<div> <div>8%</div> <div>98%</div> <div>.</div> </div>
1	E	174	<div> <div>9%</div> <div>93%</div> <div>7%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	174	<div><div></div><div>10%</div><div>93%</div><div>6%</div><div></div></div>
2	G	7	<div><div></div><div>86%</div><div>14%</div><div></div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 16348 atoms, of which 8109 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 Baseplate protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	174	Total	C	H	N	O	S	0	0	0
			2714	867	1350	223	273	1			
1	B	174	Total	C	H	N	O	S	0	0	0
			2714	867	1350	223	273	1			
1	C	174	Total	C	H	N	O	S	0	0	0
			2714	867	1350	223	273	1			
1	D	174	Total	C	H	N	O	S	0	0	0
			2714	867	1350	223	273	1			
1	E	174	Total	C	H	N	O	S	0	0	0
			2714	867	1350	223	273	1			
1	F	174	Total	C	H	N	O	S	0	0	0
			2714	867	1350	223	273	1			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	174	ASP	-	expression tag	UNP A0A0U1WGQ5
A	175	PRO	-	expression tag	UNP A0A0U1WGQ5
B	174	ASP	-	expression tag	UNP A0A0U1WGQ5
B	175	PRO	-	expression tag	UNP A0A0U1WGQ5
C	174	ASP	-	expression tag	UNP A0A0U1WGQ5
C	175	PRO	-	expression tag	UNP A0A0U1WGQ5
D	174	ASP	-	expression tag	UNP A0A0U1WGQ5
D	175	PRO	-	expression tag	UNP A0A0U1WGQ5
E	174	ASP	-	expression tag	UNP A0A0U1WGQ5
E	175	PRO	-	expression tag	UNP A0A0U1WGQ5
F	174	ASP	-	expression tag	UNP A0A0U1WGQ5
F	175	PRO	-	expression tag	UNP A0A0U1WGQ5

- Molecule 2 is a protein called Unknown peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	7	Total	C	H	N	O	0	0	0
			44	21	9	7	7			

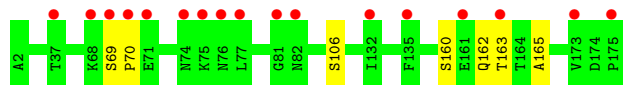
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	4	Total	O	0	0
			4	4		
3	B	5	Total	O	0	0
			5	5		
3	C	2	Total	O	0	0
			2	2		
3	D	2	Total	O	0	0
			2	2		
3	E	3	Total	O	0	0
			3	3		
3	F	4	Total	O	0	0
			4	4		

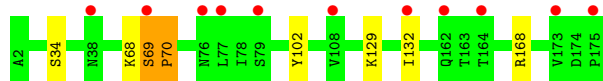
### 3 Residue-property plots [i](#)

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.

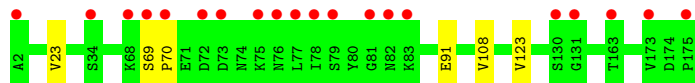
- Molecule 1: Baseplate protein



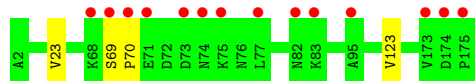
- Molecule 1: Baseplate protein



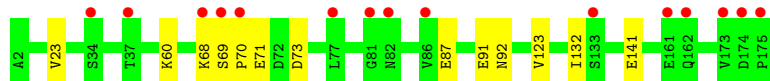
- Molecule 1: Baseplate protein



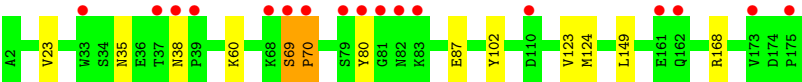
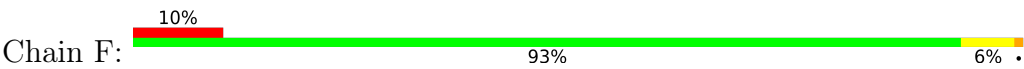
- Molecule 1: Baseplate protein



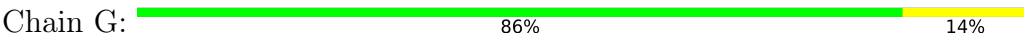
- Molecule 1: Baseplate protein



- Molecule 1: Baseplate protein



● Molecule 2: Unknown peptide



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.93Å 75.14Å 94.45Å 90.00° 109.87° 90.00°	Depositor
Resolution (Å)	46.49 – 1.80 46.49 – 1.80	Depositor EDS
% Data completeness (in resolution range)	96.1 (46.49-1.80) 96.4 (46.49-1.80)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.13 (at 1.79Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487, PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.245 , 0.275 0.245 , 0.275	Depositor DCC
$R_{free}$ test set	9506 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	38.1	Xtriage
Anisotropy	0.029	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 21.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.006 for l,k,h	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	16348	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 16.15% 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	0.49	0/1393	0.68	0/1895
1	B	0.57	0/1393	0.73	3/1895 (0.2%)
1	C	0.48	0/1393	0.61	0/1895
1	D	0.51	0/1393	0.64	0/1895
1	E	0.50	0/1393	0.71	0/1895
1	F	0.50	0/1393	0.69	0/1895
All	All	0.51	0/8358	0.68	3/11370 (0.0%)

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	E	0	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	68	LYS	CA-C-N	6.12	136.74	121.80
1	B	68	LYS	C-N-CA	6.12	136.74	121.80
1	B	70	PRO	CA-N-CD	-5.07	104.90	112.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	68	LYS	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	1364	1350	1350	6	0
1	B	1364	1350	1350	7	1
1	C	1364	1350	1350	4	0
1	D	1364	1350	1350	2	0
1	E	1364	1350	1350	8	0
1	F	1364	1350	1350	8	0
2	G	35	9	9	1	0
3	A	4	0	0	0	0
3	B	5	0	0	0	0
3	C	2	0	0	0	0
3	D	2	0	0	0	0
3	E	3	0	0	0	0
3	F	4	0	0	0	0
All	All	8239	8109	8109	35	1

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 (35) 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:69:SER:HB3	1:B:70:PRO:CD	2.15	0.77
1:F:80:TYR:OH	2:G:1:UNK:N	2.21	0.74
1:B:69:SER:CB	1:B:70:PRO:HD3	2.21	0.71
1:B:69:SER:HB3	1:B:70:PRO:HD3	1.73	0.70
1:B:69:SER:CB	1:B:70:PRO:CD	2.76	0.63
1:B:102:TYR:OH	1:B:168:ARG:HD2	2.00	0.62
1:F:69:SER:HB3	1:F:70:PRO:HD2	1.83	0.61
1:B:69:SER:OG	1:B:70:PRO:HD3	2.02	0.60
1:A:162:GLN:NE2	1:A:162:GLN:HA	2.19	0.56
1:A:162:GLN:HA	1:A:162:GLN:HE21	1.70	0.55
1:B:129:LYS:HB2	1:B:132:ILE:HD12	1.92	0.52
1:A:160:SER:OG	1:A:163:THR:HG23	2.10	0.51
1:E:60:LYS:NZ	1:E:87:GLU:OE2	2.40	0.51
1:E:69:SER:O	1:E:73:ASP:CG	2.54	0.51

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:35:ASN:HD22	1:F:38:ASN:HB3	1.76	0.51
1:F:102:TYR:OH	1:F:168:ARG:HD2	2.10	0.50
1:A:162:GLN:HE21	1:A:162:GLN:CA	2.26	0.48
1:E:69:SER:O	1:E:73:ASP:OD1	2.33	0.46
1:C:108:VAL:HG23	1:C:108:VAL:O	2.15	0.46
1:F:69:SER:CB	1:F:70:PRO:HD2	2.45	0.45
1:F:124:MET:HB3	1:F:149:LEU:HD11	1.98	0.45
1:A:106:SER:HA	1:A:165:ALA:O	2.17	0.45
1:E:69:SER:HB2	1:E:70:PRO:HD3	1.97	0.45
1:F:60:LYS:NZ	1:F:87:GLU:OE2	2.44	0.44
1:E:132:ILE:HD13	1:E:141:GLU:HB3	2.01	0.43
1:C:23:VAL:HA	1:C:123:VAL:O	2.19	0.42
1:D:69:SER:HB2	1:D:70:PRO:HD3	2.00	0.42
1:F:23:VAL:HA	1:F:123:VAL:O	2.19	0.42
1:E:70:PRO:HG2	1:E:71:GLU:OE1	2.20	0.42
1:E:23:VAL:HA	1:E:123:VAL:O	2.21	0.41
1:A:69:SER:O	1:A:70:PRO:C	2.64	0.41
1:C:69:SER:HB2	1:C:70:PRO:HD3	2.02	0.41
1:C:91:GLU:H	1:C:91:GLU:CD	2.27	0.41
1:D:23:VAL:HA	1:D:123:VAL:O	2.21	0.40
1:E:91:GLU:HG2	1:E:92:ASN:N	2.36	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:34:SER:O	1:B:69:SER:OG[2_546]	2.08	0.12

## 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	172/174 (99%)	165 (96%)	7 (4%)	0	100	100
1	B	172/174 (99%)	170 (99%)	1 (1%)	1 (1%)	22	11
1	C	172/174 (99%)	170 (99%)	2 (1%)	0	100	100
1	D	172/174 (99%)	170 (99%)	2 (1%)	0	100	100
1	E	172/174 (99%)	168 (98%)	4 (2%)	0	100	100
1	F	172/174 (99%)	169 (98%)	2 (1%)	1 (1%)	22	11
All	All	1032/1044 (99%)	1012 (98%)	18 (2%)	2 (0%)	44	31

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	69	SER
1	F	70	PRO

### 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	154/154 (100%)	154 (100%)	0	100	100
1	B	154/154 (100%)	154 (100%)	0	100	100
1	C	154/154 (100%)	154 (100%)	0	100	100
1	D	154/154 (100%)	154 (100%)	0	100	100
1	E	154/154 (100%)	154 (100%)	0	100	100
1	F	154/154 (100%)	153 (99%)	1 (1%)	84	82
All	All	924/924 (100%)	923 (100%)	1 (0%)	92	91

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	69	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14)

such sidechains are listed below:

Mol	Chain	Res	Type
1	A	45	ASN
1	A	50	GLN
1	A	162	GLN
1	B	41	GLN
1	B	155	GLN
1	C	9	HIS
1	C	50	GLN
1	D	50	GLN
1	D	162	GLN
1	E	45	ASN
1	E	50	GLN
1	F	35	ASN
1	F	82	ASN
1	F	92	ASN

### 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 oligosaccharides 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	174/174 (100%)	0.77	17 (9%) 14 12	31, 44, 73, 94	0
1	B	174/174 (100%)	0.68	11 (6%) 27 24	29, 43, 64, 80	0
1	C	174/174 (100%)	0.83	20 (11%) 11 9	32, 48, 88, 112	0
1	D	174/174 (100%)	0.68	14 (8%) 20 17	32, 43, 98, 118	0
1	E	174/174 (100%)	0.56	15 (8%) 18 15	30, 44, 65, 104	0
1	F	174/174 (100%)	0.82	17 (9%) 14 12	32, 47, 72, 101	0
2	G	0/7	-	-	-	-
All	All	1044/1051 (99%)	0.72	94 (9%) 17 14	29, 44, 72, 118	0

All (94) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	175	PRO	6.8
1	E	175	PRO	6.4
1	F	175	PRO	5.8
1	A	70	PRO	5.6
1	E	69	SER	5.6
1	C	70	PRO	5.3
1	A	76	ASN	5.3
1	D	175	PRO	5.2
1	D	70	PRO	5.0
1	B	175	PRO	5.0
1	D	173	VAL	4.8
1	A	69	SER	4.7
1	B	69	SER	4.6
1	C	77	LEU	4.5
1	F	70	PRO	4.5
1	A	77	LEU	4.4
1	B	76	ASN	4.2
1	F	69	SER	4.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	73	ASP	4.0
1	E	173	VAL	3.8
1	F	80	TYR	3.7
1	E	70	PRO	3.7
1	B	77	LEU	3.7
1	A	132	ILE	3.6
1	A	71	GLU	3.6
1	D	77	LEU	3.5
1	E	82	ASN	3.4
1	A	173	VAL	3.3
1	C	69	SER	3.3
1	A	175	PRO	3.3
1	E	174	ASP	3.3
1	F	161	GLU	3.2
1	D	68	LYS	3.2
1	C	81	GLY	3.1
1	D	174	ASP	3.1
1	A	75	LYS	3.0
1	B	162	GLN	2.9
1	F	37	THR	2.9
1	C	131	GLY	2.8
1	D	74	ASN	2.8
1	E	77	LEU	2.8
1	F	82	ASN	2.7
1	A	82	ASN	2.7
1	B	79	SER	2.7
1	A	81	GLY	2.7
1	C	75	LYS	2.7
1	B	173	VAL	2.6
1	C	79	SER	2.6
1	F	81	GLY	2.6
1	A	74	ASN	2.6
1	C	72	ASP	2.6
1	A	163	THR	2.6
1	A	68	LYS	2.5
1	F	173	VAL	2.5
1	F	33	TRP	2.5
1	C	130	SER	2.5
1	D	71	GLU	2.5
1	E	37	THR	2.5
1	C	2	ALA	2.5
1	E	162	GLN	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	78	ILE	2.5
1	B	132	ILE	2.4
1	C	163	THR	2.4
1	C	173	VAL	2.4
1	E	86	VAL	2.4
1	E	161	GLU	2.4
1	F	68	LYS	2.4
1	D	95	ALA	2.3
1	F	110	ASP	2.3
1	F	39	PRO	2.3
1	C	76	ASN	2.3
1	F	83	LYS	2.3
1	D	82	ASN	2.3
1	D	73	ASP	2.3
1	D	69	SER	2.3
1	C	83	LYS	2.2
1	E	81	GLY	2.2
1	F	38	ASN	2.2
1	A	135	PHE	2.2
1	D	83	LYS	2.2
1	E	68	LYS	2.2
1	F	162	GLN	2.2
1	A	161	GLU	2.2
1	C	34	SER	2.1
1	F	79	SER	2.1
1	D	75	LYS	2.1
1	E	133	SER	2.1
1	B	108	VAL	2.1
1	C	68	LYS	2.1
1	E	34	SER	2.0
1	A	37	THR	2.0
1	B	38	ASN	2.0
1	B	164	THR	2.0
1	C	82	ASN	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 [i](#)

There are no monosaccharides 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.