



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 9, 2024 – 12:14 PM JST

PDB ID : 8Y9K  
Title : Crystal structure of the BmCPV1 NSP9 homodimer  
Authors : Guo, H.; Wang, Y.; Ji, X.  
Deposited on : 2024-02-07  
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 : 4.02b-467  
Xtriage (Phenix) : 1.21  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.004 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.40

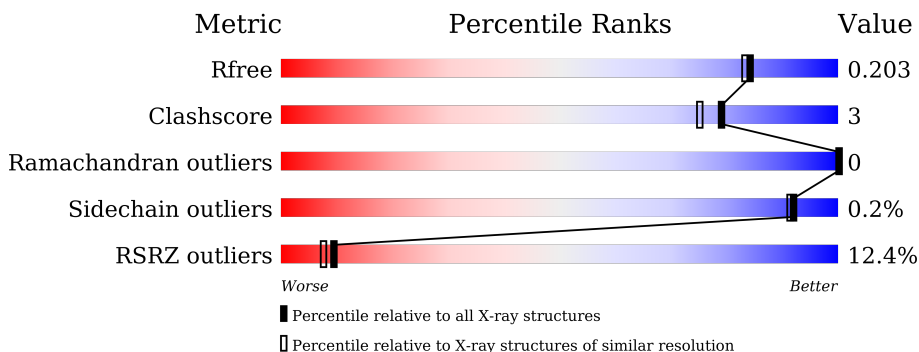
# 1 Overall quality at a glance

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	289	
1	B	289	
1	C	289	
1	D	289	

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 15796 atoms, of which 7466 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 Nonstructural protein 9.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	259	3991	1290	1983	336	369	13	0	0	0
1	B	259	3947	1280	1949	333	372	13	0	0	0
1	C	243	3698	1204	1821	315	346	12	0	0	0
1	D	236	3515	1160	1713	298	332	12	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	ALA	-	expression tag	UNP O72505
A	-1	GLY	-	expression tag	UNP O72505
A	0	PHE	-	expression tag	UNP O72505
B	-2	ALA	-	expression tag	UNP O72505
B	-1	GLY	-	expression tag	UNP O72505
B	0	PHE	-	expression tag	UNP O72505
C	-2	ALA	-	expression tag	UNP O72505
C	-1	GLY	-	expression tag	UNP O72505
C	0	PHE	-	expression tag	UNP O72505
D	-2	ALA	-	expression tag	UNP O72505
D	-1	GLY	-	expression tag	UNP O72505
D	0	PHE	-	expression tag	UNP O72505

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	218	Total	O	0	0
			218	218		
2	B	195	Total	O	0	0
			195	195		

*Continued on next page...*

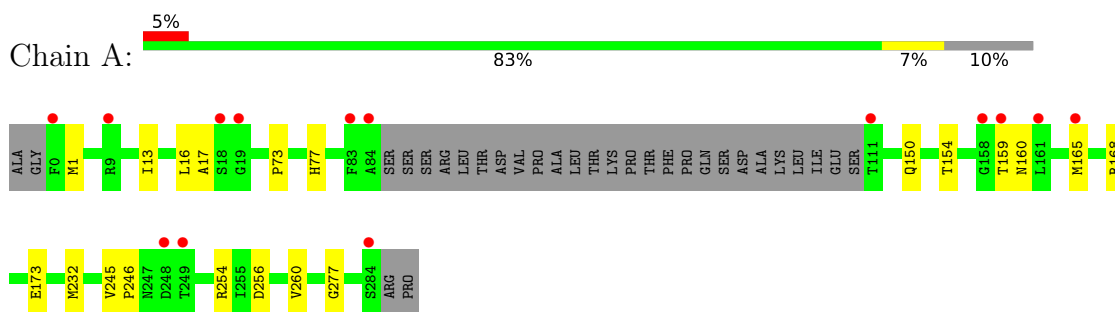
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
2	C	166	Total 166	O 166	0	0
2	D	66	Total 66	O 66	0	0

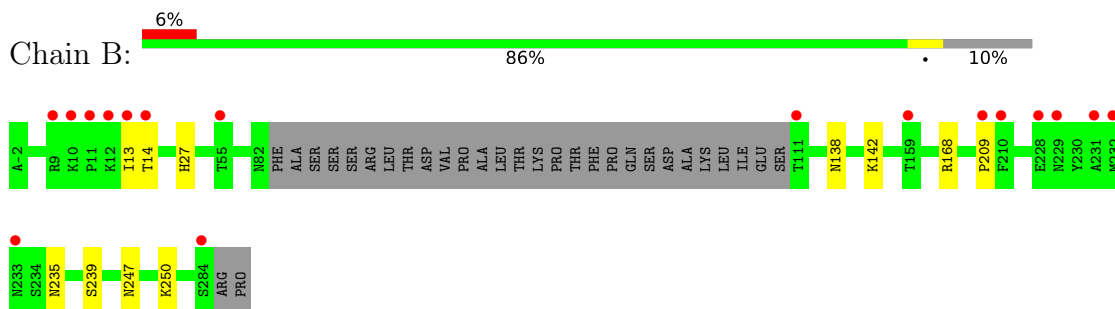
### 3 Residue-property plots [i](#)

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.

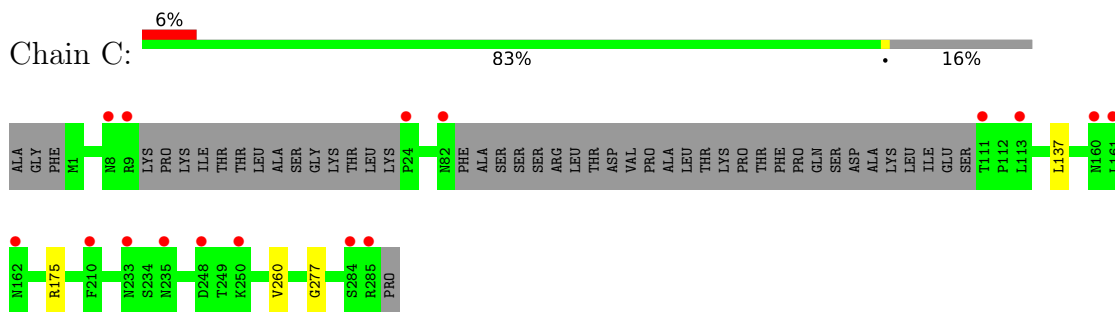
- Molecule 1: Nonstructural protein 9



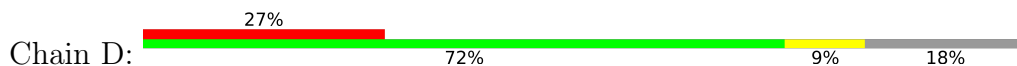
- Molecule 1: Nonstructural protein 9

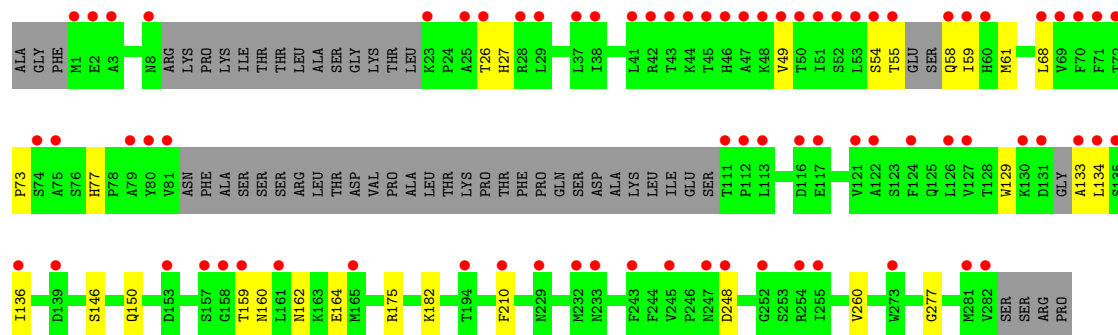


- Molecule 1: Nonstructural protein 9



- Molecule 1: Nonstructural protein 9





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.43Å 83.57Å 184.41Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.21 – 1.80 41.21 – 1.80	Depositor EDS
% Data completeness (in resolution range)	98.6 (41.21-1.80) 92.0 (41.21-1.80)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	-0.18 (at 1.79Å)	Xtrriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, $R_{free}$	0.184 , 0.204 0.186 , 0.203	Depositor DCC
$R_{free}$ test set	116398 reflections (1.71%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	28.8	Xtrriage
Anisotropy	0.221	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 43.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.024 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	15796	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.66% 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	0.36	0/2058	0.59	0/2800
1	B	0.39	0/2048	0.67	0/2791
1	C	0.37	0/1924	0.64	0/2619
1	D	0.37	0/1847	0.61	0/2518
All	All	0.37	0/7877	0.63	0/10728

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	C	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	175	ARG	Sidechain

### 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	2008	1983	1984	14	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1998	1949	1949	7	0
1	C	1877	1821	1822	2	0
1	D	1802	1713	1712	20	0
2	A	218	0	0	1	0
2	B	195	0	0	2	0
2	C	166	0	0	0	0
2	D	66	0	0	1	0
All	All	8330	7466	7467	40	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 3.

All (40) 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:133:ALA:HA	1:D:136:ILE:HD13	1.79	0.64
1:D:61:MET:HE2	1:D:68:LEU:HD11	1.83	0.60
1:D:58:GLN:C	1:D:59:ILE:HD12	2.23	0.59
1:D:129:TRP:HA	1:D:133:ALA:HB3	1.84	0.58
1:A:165:MET:O	1:A:165:MET:HE2	2.03	0.58
1:D:136:ILE:H	1:D:136:ILE:HD12	1.71	0.55
1:A:150:GLN:NE2	2:A:302:HOH:O	2.34	0.55
1:A:232:MET:HE3	1:C:137:LEU:HG	1.88	0.54
1:A:13:ILE:HD12	1:A:13:ILE:O	2.08	0.54
1:D:146:SER:O	1:D:150:GLN:HG3	2.09	0.53
1:D:26:THR:HG1	1:D:27:HIS:CE1	2.26	0.53
1:A:16:LEU:HD12	1:B:14:THR:O	2.11	0.50
1:A:254:ARG:HG2	1:A:256:ASP:OD1	2.11	0.50
1:D:49:VAL:HG21	1:D:136:ILE:HD11	1.95	0.49
1:D:61:MET:CE	1:D:68:LEU:HD11	2.43	0.49
1:D:260:VAL:O	1:D:277:GLY:HA2	2.13	0.48
1:D:159:THR:HG22	1:D:160:ASN:O	2.14	0.48
1:B:138:ASN:O	1:B:142:LYS:HG3	2.15	0.47
1:D:59:ILE:HD12	1:D:59:ILE:N	2.30	0.47
1:D:73:PRO:HB2	1:D:77:HIS:HB3	1.97	0.47
1:B:235:ASN:OD1	2:B:301:HOH:O	2.21	0.47
1:C:260:VAL:O	1:C:277:GLY:HA2	2.17	0.45
1:A:17:ALA:HA	1:B:13:ILE:HD13	1.98	0.45
1:A:165:MET:HE3	1:A:168:ARG:HD2	1.98	0.45
1:D:136:ILE:HD12	1:D:136:ILE:N	2.32	0.45
1:B:27:HIS:CE1	1:B:168:ARG:HD2	2.51	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:159:THR:HG22	1:A:160:ASN:O	2.18	0.44
1:A:260:VAL:O	1:A:277:GLY:HA2	2.18	0.43
1:A:245:VAL:HB	1:A:246:PRO:HD2	2.01	0.42
1:B:247:ASN:HB3	1:B:250:LYS:HG2	2.01	0.42
1:D:175:ARG:HG2	1:D:210:PHE:CD1	2.55	0.42
1:D:54:SER:O	1:D:55:THR:C	2.57	0.42
1:B:209:PRO:HG2	2:B:328:HOH:O	2.20	0.42
1:D:182:LYS:HE2	2:D:310:HOH:O	2.18	0.42
1:A:13:ILE:HD12	1:A:13:ILE:C	2.41	0.41
1:A:73:PRO:HB2	1:A:77:HIS:HB3	2.02	0.41
1:D:134:LEU:HD23	1:D:134:LEU:HA	1.85	0.40
1:D:248:ASP:OD1	1:D:248:ASP:N	2.52	0.40
1:D:162:ASN:OD1	1:D:164:GLU:N	2.54	0.40
1:A:1:MET:HG3	1:A:154:THR:HG21	2.02	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	255/289 (88%)	254 (100%)	1 (0%)	0	100	100
1	B	255/289 (88%)	253 (99%)	2 (1%)	0	100	100
1	C	237/289 (82%)	236 (100%)	1 (0%)	0	100	100
1	D	226/289 (78%)	223 (99%)	3 (1%)	0	100	100
All	All	973/1156 (84%)	966 (99%)	7 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	222/256 (87%)	221 (100%)	1 (0%)	86	86
1	B	220/256 (86%)	219 (100%)	1 (0%)	86	86
1	C	205/256 (80%)	205 (100%)	0	100	100
1	D	192/256 (75%)	192 (100%)	0	100	100
All	All	839/1024 (82%)	837 (100%)	2 (0%)	92	91

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	173	GLU
1	B	239	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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	259/289 (89%)	0.01	14 (5%) 32 30	25, 36, 67, 93	0
1	B	259/289 (89%)	0.27	17 (6%) 26 22	26, 42, 72, 90	0
1	C	243/289 (84%)	0.19	16 (6%) 26 22	26, 41, 65, 109	0
1	D	236/289 (81%)	1.49	77 (32%) 1 1	33, 68, 108, 142	0
All	All	997/1156 (86%)	0.47	124 (12%) 9 7	25, 44, 87, 142	0

All (124) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	53	LEU	7.8
1	B	13	ILE	7.1
1	D	51	ILE	6.5
1	D	133	ALA	5.7
1	D	55	THR	5.5
1	D	26	THR	5.3
1	D	47	ALA	5.2
1	D	54	SER	5.1
1	D	41	LEU	4.7
1	B	14	THR	4.7
1	C	161	LEU	4.5
1	D	23	LYS	4.4
1	D	81	VAL	4.3
1	C	285	ARG	4.3
1	B	12	LYS	4.3
1	A	284	SER	4.2
1	D	59	ILE	4.2
1	D	50	THR	4.2
1	C	248	ASP	4.1
1	C	24	PRO	4.1
1	D	153	ASP	4.0

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	D	58	GLN	4.0
1	C	160	ASN	3.9
1	D	48	LYS	3.8
1	D	46	HIS	3.8
1	D	44	LYS	3.7
1	C	9	ARG	3.7
1	D	37	LEU	3.7
1	D	116	ASP	3.6
1	D	281	MET	3.6
1	D	80	TYR	3.6
1	B	111	THR	3.5
1	C	233	ASN	3.4
1	C	8	ASN	3.4
1	A	161	LEU	3.4
1	B	284	SER	3.4
1	A	0	PHE	3.3
1	D	282	VAL	3.3
1	B	11	PRO	3.2
1	D	124	PHE	3.2
1	D	117	GLU	3.1
1	D	158	GLY	3.1
1	C	162	ASN	3.1
1	D	49	VAL	3.1
1	D	111	THR	3.1
1	D	69	VAL	3.0
1	D	122	ALA	3.0
1	D	157	SER	3.0
1	D	127	VAL	3.0
1	B	228	GLU	2.9
1	A	249	THR	2.9
1	D	71	PHE	2.9
1	D	42	ARG	2.9
1	D	29	LEU	2.9
1	A	165	MET	2.9
1	A	159	THR	2.8
1	D	136	ILE	2.8
1	D	130	LYS	2.8
1	D	112	PRO	2.8
1	A	84	ALA	2.8
1	D	139	ASP	2.7
1	D	159	THR	2.7
1	D	68	LEU	2.7

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	D	38	ILE	2.7
1	D	210	PHE	2.6
1	D	8	ASN	2.6
1	C	235	ASN	2.6
1	D	43	THR	2.6
1	D	3	ALA	2.6
1	D	121	VAL	2.6
1	D	1	MET	2.6
1	B	9	ARG	2.6
1	D	254	ARG	2.5
1	A	9	ARG	2.5
1	B	229	ASN	2.5
1	D	131	ASP	2.5
1	D	75	ALA	2.5
1	D	161	LEU	2.4
1	D	2	GLU	2.4
1	D	70	PHE	2.4
1	D	248	ASP	2.4
1	A	248	ASP	2.4
1	D	72	THR	2.4
1	D	255	ILE	2.4
1	D	273	TRP	2.4
1	B	209	PRO	2.4
1	D	194	THR	2.4
1	B	232	MET	2.3
1	C	250	LYS	2.3
1	D	165	MET	2.3
1	D	60	HIS	2.3
1	A	83	PHE	2.3
1	D	79	ALA	2.3
1	D	74	SER	2.3
1	D	134	LEU	2.3
1	A	19	GLY	2.2
1	B	233	ASN	2.2
1	C	82	ASN	2.2
1	D	233	ASN	2.2
1	D	247	ASN	2.2
1	B	10	LYS	2.2
1	B	210	PHE	2.2
1	B	159	THR	2.2
1	A	18	SER	2.2
1	D	113	LEU	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	111	THR	2.1
1	D	52	SER	2.1
1	D	229	ASN	2.1
1	C	210	PHE	2.1
1	D	45	THR	2.1
1	D	126	LEU	2.1
1	D	28	ARG	2.1
1	C	113	LEU	2.1
1	B	55	THR	2.1
1	A	158	GLY	2.1
1	B	231	ALA	2.0
1	D	25	ALA	2.0
1	D	243	PHE	2.0
1	D	232	MET	2.0
1	A	111	THR	2.0
1	D	252	GLY	2.0
1	D	245	VAL	2.0
1	C	284	SER	2.0
1	D	135	SER	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 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.