



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 13, 2024 – 06:36 AM EST

PDB ID : 4JEP  
Title : Crystal structure of Toxoplasma gondii nucleoside triphosphate diphosphohydrolase 1 (NTPDase1)  
Authors : Krug, U.; Totzauer, R.; Strater, N.  
Deposited on : 2013-02-27  
Resolution : 3.10 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)	:	1.20.1
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)	:	2.39

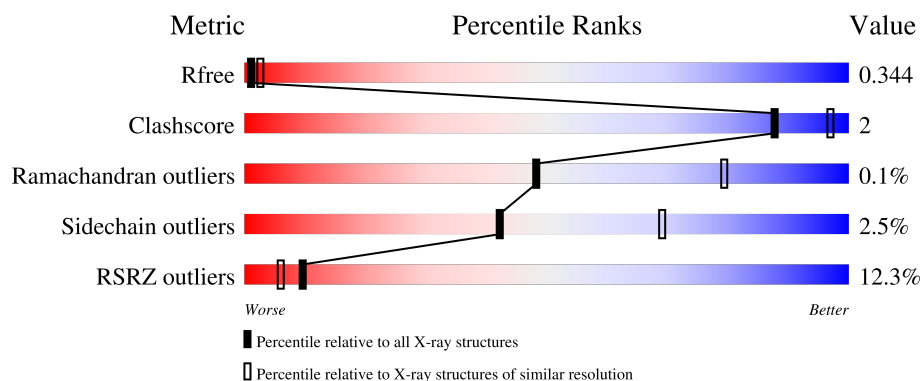
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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	1351 (3.10-3.10)
Clashscore	180529	1454 (3.10-3.10)
Ramachandran outliers	177936	1391 (3.10-3.10)
Sidechain outliers	177891	1391 (3.10-3.10)
RSRZ outliers	164620	1351 (3.10-3.10)

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	611	<div> <div>10%</div> <div>86%</div> <div>8%</div> <div>6%</div> </div>
1	B	611	<div> <div>13%</div> <div>83%</div> <div>7%</div> <div>10%</div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 8772 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 Nucleoside-triphosphatase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	574	Total	C	N	O	S	0	0	0
			4471	2805	789	853	24			
1	B	552	Total	C	N	O	S	0	0	0
			4301	2703	756	816	26			

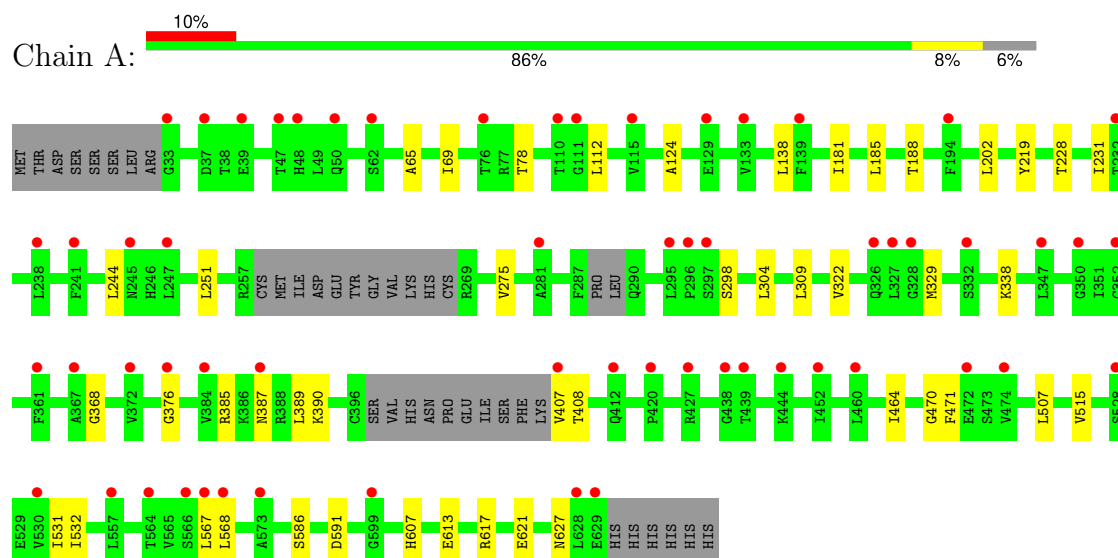
There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	25	MET	-	expression tag	UNP Q27895
A	629	GLU	-	expression tag	UNP Q27895
A	630	HIS	-	expression tag	UNP Q27895
A	631	HIS	-	expression tag	UNP Q27895
A	632	HIS	-	expression tag	UNP Q27895
A	633	HIS	-	expression tag	UNP Q27895
A	634	HIS	-	expression tag	UNP Q27895
A	635	HIS	-	expression tag	UNP Q27895
B	25	MET	-	expression tag	UNP Q27895
B	629	GLU	-	expression tag	UNP Q27895
B	630	HIS	-	expression tag	UNP Q27895
B	631	HIS	-	expression tag	UNP Q27895
B	632	HIS	-	expression tag	UNP Q27895
B	633	HIS	-	expression tag	UNP Q27895
B	634	HIS	-	expression tag	UNP Q27895
B	635	HIS	-	expression tag	UNP Q27895

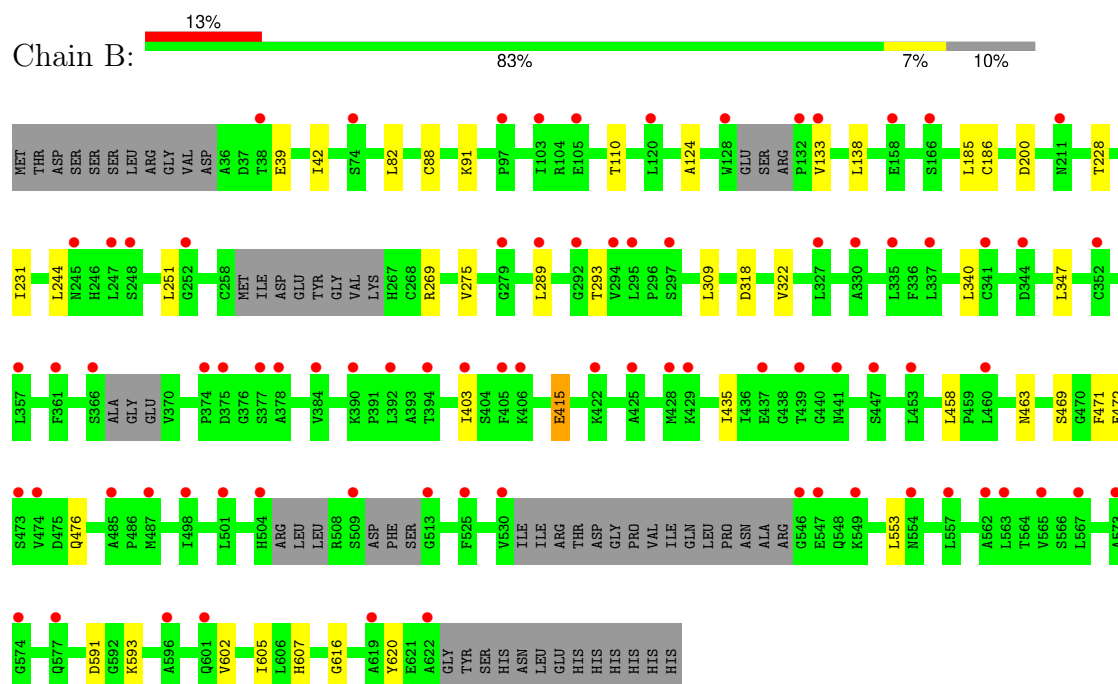
### 3 Residue-property plots [i](#)

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: Nucleoside-triphosphatase 2



#### • Molecule 1: Nucleoside-triphosphatase 2



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	72.74Å 161.52Å 236.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.31 – 3.10 36.31 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.4 (36.31-3.10) 99.6 (36.31-3.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.20 (at 3.12Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, $R_{free}$	0.270 , 0.305 0.303 , 0.344	Depositor DCC
$R_{free}$ test set	1306 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	65.1	Xtriage
Anisotropy	0.228	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 62.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.81	EDS
Total number of atoms	8772	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	71.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.13% 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	0.36	0/4554	0.55	0/6161
1	B	0.37	0/4381	0.54	0/5922
All	All	0.37	0/8935	0.55	0/12083

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

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	4471	0	4420	19	0
1	B	4301	0	4249	16	0
All	All	8772	0	8669	33	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 (33) 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:338:LYS:HB2	1:A:389:LEU:HD11	1.86	0.56
1:B:616:GLY:HA2	1:B:620:TYR:HD1	1.73	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:185:LEU:HB3	1:B:228:THR:HG23	1.89	0.54
1:A:69:ILE:HG12	1:A:78:THR:HG22	1.89	0.53
1:B:289:LEU:HD11	1:B:293:THR:HG21	1.90	0.53
1:A:185:LEU:HB3	1:A:228:THR:HG23	1.91	0.52
1:A:368:GLY:HA2	1:A:387:ASN:HA	1.92	0.52
1:B:322:VAL:HG11	1:B:471:PHE:HB2	1.91	0.51
1:A:124:ALA:HB2	1:A:138:LEU:HD21	1.93	0.50
1:A:407:VAL:HG22	1:B:476:GLN:HE21	1.76	0.49
1:B:244:LEU:HD22	1:B:275:VAL:HB	1.96	0.47
1:B:415:GLU:HG2	1:B:458:LEU:HD13	1.96	0.47
1:A:188:THR:HG22	1:A:231:ILE:HD11	1.95	0.46
1:B:124:ALA:HB2	1:B:138:LEU:HD21	1.97	0.46
1:B:269:ARG:HH21	1:B:318:ASP:HB2	1.81	0.46
1:B:309:LEU:HA	1:B:607:HIS:HD2	1.81	0.45
1:B:39:GLU:HA	1:B:42:ILE:HD12	1.98	0.45
1:A:244:LEU:HD22	1:A:275:VAL:HB	1.98	0.44
1:A:112:LEU:HD22	1:A:202:LEU:HD21	1.98	0.44
1:A:464:ILE:HG12	1:A:470:GLY:HA2	2.00	0.44
1:A:376:GLY:HA3	1:A:532:ILE:HG12	2.00	0.43
1:B:602:VAL:HA	1:B:605:ILE:HD12	1.99	0.43
1:A:181:ILE:HD12	1:A:219:TYR:HB3	2.01	0.43
1:B:322:VAL:HG13	1:B:469:SER:HB2	2.00	0.43
1:A:322:VAL:HG11	1:A:471:PHE:HB2	2.00	0.43
1:A:385:ARG:HA	1:A:390:LYS:HD2	2.00	0.43
1:B:88:CYS:HB2	1:B:91:LYS:HB2	2.00	0.43
1:A:617:ARG:HA	1:A:621:GLU:HB2	2.00	0.42
1:A:298:SER:HB3	1:B:463:ASN:HD22	1.85	0.41
1:A:304:LEU:HD23	1:A:309:LEU:HD13	2.02	0.41
1:A:515:VAL:HG22	1:A:568:LEU:HB3	2.03	0.41
1:B:186:CYS:HB3	1:B:231:ILE:HG12	2.03	0.40
1:A:65:ALA:HB3	1:A:181:ILE:HG12	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	566/611 (93%)	545 (96%)	20 (4%)	1 (0%)	44	74
1	B	538/611 (88%)	516 (96%)	22 (4%)	0	100	100
All	All	1104/1222 (90%)	1061 (96%)	42 (4%)	1 (0%)	48	79

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	329	MET

### 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	491/527 (93%)	481 (98%)	10 (2%)	50	74
1	B	475/527 (90%)	461 (97%)	14 (3%)	37	65
All	All	966/1054 (92%)	942 (98%)	24 (2%)	42	69

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	251	LEU
1	A	408	THR
1	A	507	LEU
1	A	531	ILE
1	A	567	LEU
1	A	586	SER
1	A	591	ASP
1	A	607	HIS
1	A	613	GLU
1	A	627	ASN
1	B	82	LEU
1	B	110	THR

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	B	133	VAL
1	B	200	ASP
1	B	251	LEU
1	B	340	LEU
1	B	347	LEU
1	B	403	ILE
1	B	415	GLU
1	B	435	ILE
1	B	472	GLU
1	B	553	LEU
1	B	591	ASP
1	B	593	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	101	GLN
1	B	463	ASN
1	B	476	GLN
1	B	607	HIS
1	B	608	HIS

### 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	574/611 (93%)	0.95	59 (10%) 13 8	32, 61, 99, 128	0
1	B	552/611 (90%)	1.11	80 (14%) 7 4	38, 78, 112, 173	0
All	All	1126/1222 (92%)	1.03	139 (12%) 9 6	32, 69, 107, 173	0

All (139) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	341	CYS	6.0
1	A	39	GLU	4.6
1	B	567	LEU	4.6
1	A	566	SER	4.5
1	A	372	VAL	4.0
1	A	111	GLY	4.0
1	A	472	GLU	3.9
1	B	366	SER	3.9
1	B	97	PRO	3.7
1	A	48	HIS	3.6
1	A	110	THR	3.6
1	B	357	LEU	3.5
1	B	132	PRO	3.5
1	B	337	LEU	3.5
1	B	252	GLY	3.5
1	A	247	LEU	3.4
1	B	390	LYS	3.4
1	B	297	SER	3.4
1	B	384	VAL	3.3
1	B	557	LEU	3.3
1	B	405	PHE	3.3
1	B	394	THR	3.2
1	A	439	THR	3.1
1	B	361	PHE	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	375	ASP	3.0
1	B	120	LEU	3.0
1	B	103	ILE	3.0
1	B	378	ALA	3.0
1	A	139	PHE	3.0
1	A	528	SER	3.0
1	A	564	THR	2.9
1	B	428	MET	2.9
1	A	568	LEU	2.9
1	B	453	LEU	2.9
1	B	525	PHE	2.9
1	B	549	LYS	2.8
1	B	158	GLU	2.8
1	B	335	LEU	2.8
1	B	530	VAL	2.8
1	A	37	ASP	2.8
1	A	573	ALA	2.8
1	B	248	SER	2.7
1	B	406	LYS	2.7
1	B	403	ILE	2.7
1	A	438	GLY	2.7
1	B	289	LEU	2.7
1	A	557	LEU	2.7
1	B	509	SER	2.6
1	B	562	ALA	2.6
1	B	563	LEU	2.6
1	B	513	GLY	2.6
1	A	295	LEU	2.6
1	B	374	PRO	2.6
1	B	554	ASN	2.6
1	A	33	GLY	2.5
1	B	547	GLU	2.5
1	A	444	LYS	2.5
1	B	377	SER	2.5
1	B	327	LEU	2.5
1	B	622	ALA	2.5
1	B	565	VAL	2.5
1	A	245	ASN	2.5
1	A	326	GLN	2.5
1	A	530	VAL	2.5
1	B	504	HIS	2.5
1	A	628	LEU	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	439	THR	2.5
1	A	384	VAL	2.5
1	B	392	LEU	2.4
1	A	412	GLN	2.4
1	A	452	ILE	2.4
1	A	361	PHE	2.4
1	B	425	ALA	2.4
1	A	567	LEU	2.4
1	A	129	GLU	2.4
1	B	546	GLY	2.4
1	B	573	ALA	2.4
1	B	577	GLN	2.4
1	B	601	GLN	2.4
1	A	629	GLU	2.3
1	B	38	THR	2.3
1	A	347	LEU	2.3
1	B	498	ILE	2.3
1	A	241	PHE	2.3
1	B	473	SER	2.3
1	B	245	ASN	2.3
1	B	344	ASP	2.3
1	A	352	CYS	2.3
1	A	407	VAL	2.3
1	B	474	VAL	2.3
1	A	387	ASN	2.3
1	A	327	LEU	2.3
1	B	352	CYS	2.3
1	B	501	LEU	2.3
1	A	47	THR	2.3
1	A	474	VAL	2.3
1	A	460	LEU	2.3
1	B	447	SER	2.2
1	A	115	VAL	2.2
1	A	281	ALA	2.2
1	B	292	GLY	2.2
1	A	50	GLN	2.2
1	B	487	MET	2.2
1	A	599	GLY	2.2
1	B	437	GLU	2.2
1	B	211	ASN	2.2
1	A	76	THR	2.2
1	B	247	LEU	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	376	GLY	2.1
1	B	105	GLU	2.1
1	B	133	VAL	2.1
1	B	294	VAL	2.1
1	B	279	GLY	2.1
1	B	460	LEU	2.1
1	A	133	VAL	2.1
1	A	328	GLY	2.1
1	A	194	PHE	2.1
1	A	238	LEU	2.1
1	B	485	ALA	2.1
1	B	574	GLY	2.1
1	A	420	PRO	2.1
1	B	128	TRP	2.1
1	B	330	ALA	2.1
1	A	427	ARG	2.1
1	B	422	LYS	2.1
1	A	62	SER	2.1
1	B	74	SER	2.1
1	B	166	SER	2.1
1	A	296	PRO	2.1
1	A	367	ALA	2.0
1	B	619	ALA	2.0
1	A	332	SER	2.0
1	B	441	ASN	2.0
1	B	596	ALA	2.0
1	A	232	THR	2.0
1	A	297	SER	2.0
1	B	295	LEU	2.0
1	B	429	LYS	2.0
1	A	350	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 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.