Full wwPDB X-ray Structure Validation Report ⓘ

Oct 15, 2023 – 02:39 PM EDT

PDB ID	:	7SR5
Title	:	Single chain trimer HLA-A*02:01 (Y108C, A163C) with Wilms tumor protein peptide RMFPNAPYL
Authors	:	Finton, K.A.K.; Rupert, P.B.
Deposited on	:	2021-11-07
Resolution	:	2.35 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*
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:

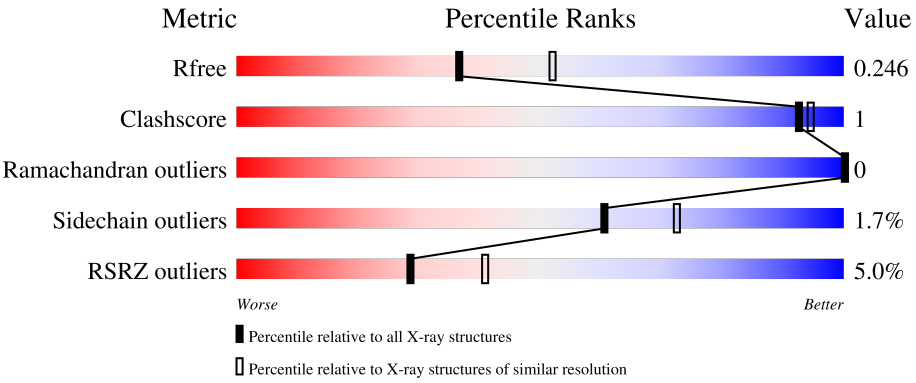
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.35 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	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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 >=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 <=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	424	2% 85% 11%
1	C	424	6% 78% 17%
2	B	116	5% 94%
2	D	116	4% 91% 6%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7504 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 Wilms tumor protein peptide,Beta-2-microglobulin,MHC class I antigen chimera.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	379	Total	C	N	O	S	0	0	0
			2919	1840	518	547	14			
1	C	350	Total	C	N	O	S	0	0	0
			2706	1706	485	503	12			

There are 86 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	10	GLY	-	linker	UNP P19544
A	11	GLY	-	linker	UNP P19544
A	12	GLY	-	linker	UNP P19544
A	13	GLY	-	linker	UNP P19544
A	14	SER	-	linker	UNP P19544
A	15	GLY	-	linker	UNP P19544
A	16	GLY	-	linker	UNP P19544
A	17	GLY	-	linker	UNP P19544
A	18	GLY	-	linker	UNP P19544
A	19	SER	-	linker	UNP P19544
A	20	GLY	-	linker	UNP P19544
A	21	GLY	-	linker	UNP P19544
A	22	GLY	-	linker	UNP P19544
A	23	GLY	-	linker	UNP P19544
A	24	SER	-	linker	UNP P19544
A	124	GLY	-	linker	UNP P16213
A	125	GLY	-	linker	UNP P16213
A	126	GLY	-	linker	UNP P16213
A	127	GLY	-	linker	UNP P16213
A	128	SER	-	linker	UNP P16213
A	129	GLY	-	linker	UNP P16213
A	130	GLY	-	linker	UNP P16213
A	131	GLY	-	linker	UNP P16213
A	132	GLY	-	linker	UNP P16213

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	133	SER	-	linker	UNP P16213
A	134	GLY	-	linker	UNP P16213
A	135	GLY	-	linker	UNP P16213
A	136	GLY	-	linker	UNP P16213
A	137	GLY	-	linker	UNP P16213
A	138	SER	-	linker	UNP P16213
A	139	GLY	-	linker	UNP P16213
A	140	GLY	-	linker	UNP P16213
A	141	GLY	-	linker	UNP P16213
A	142	GLY	-	linker	UNP P16213
A	143	SER	-	linker	UNP P16213
A	227	CYS	TYR	engineered mutation	UNP A0A678ZGP6
A	282	CYS	ALA	engineered mutation	UNP A0A678ZGP6
A	419	HIS	-	expression tag	UNP A0A678ZGP6
A	420	HIS	-	expression tag	UNP A0A678ZGP6
A	421	HIS	-	expression tag	UNP A0A678ZGP6
A	422	HIS	-	expression tag	UNP A0A678ZGP6
A	423	HIS	-	expression tag	UNP A0A678ZGP6
A	424	HIS	-	expression tag	UNP A0A678ZGP6
C	10	GLY	-	linker	UNP P19544
C	11	GLY	-	linker	UNP P19544
C	12	GLY	-	linker	UNP P19544
C	13	GLY	-	linker	UNP P19544
C	14	SER	-	linker	UNP P19544
C	15	GLY	-	linker	UNP P19544
C	16	GLY	-	linker	UNP P19544
C	17	GLY	-	linker	UNP P19544
C	18	GLY	-	linker	UNP P19544
C	19	SER	-	linker	UNP P19544
C	20	GLY	-	linker	UNP P19544
C	21	GLY	-	linker	UNP P19544
C	22	GLY	-	linker	UNP P19544
C	23	GLY	-	linker	UNP P19544
C	24	SER	-	linker	UNP P19544
C	124	GLY	-	linker	UNP P16213
C	125	GLY	-	linker	UNP P16213
C	126	GLY	-	linker	UNP P16213
C	127	GLY	-	linker	UNP P16213
C	128	SER	-	linker	UNP P16213
C	129	GLY	-	linker	UNP P16213
C	130	GLY	-	linker	UNP P16213
C	131	GLY	-	linker	UNP P16213

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	132	GLY	-	linker	UNP P16213
C	133	SER	-	linker	UNP P16213
C	134	GLY	-	linker	UNP P16213
C	135	GLY	-	linker	UNP P16213
C	136	GLY	-	linker	UNP P16213
C	137	GLY	-	linker	UNP P16213
C	138	SER	-	linker	UNP P16213
C	139	GLY	-	linker	UNP P16213
C	140	GLY	-	linker	UNP P16213
C	141	GLY	-	linker	UNP P16213
C	142	GLY	-	linker	UNP P16213
C	143	SER	-	linker	UNP P16213
C	227	CYS	TYR	engineered mutation	UNP A0A678ZGP6
C	282	CYS	ALA	engineered mutation	UNP A0A678ZGP6
C	419	HIS	-	expression tag	UNP A0A678ZGP6
C	420	HIS	-	expression tag	UNP A0A678ZGP6
C	421	HIS	-	expression tag	UNP A0A678ZGP6
C	422	HIS	-	expression tag	UNP A0A678ZGP6
C	423	HIS	-	expression tag	UNP A0A678ZGP6
C	424	HIS	-	expression tag	UNP A0A678ZGP6

- Molecule 2 is a protein called VHH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	113	Total	C	N	O	S	0	0	0
			829	512	142	171	4			
2	D	112	Total	C	N	O	S	0	0	0
			825	509	142	170	4			

- Molecule 3 is water.

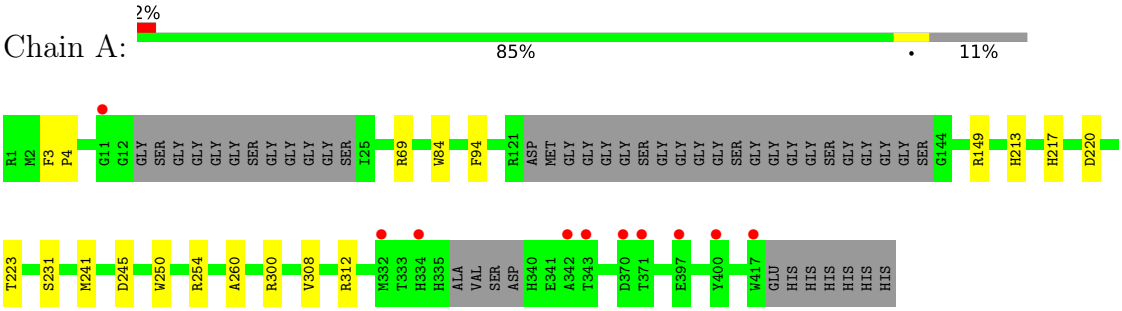
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	88	Total	O	0	0
			88	88		
3	B	44	Total	O	0	0
			44	44		
3	C	53	Total	O	0	0
			53	53		
3	D	40	Total	O	0	0
			40	40		

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.

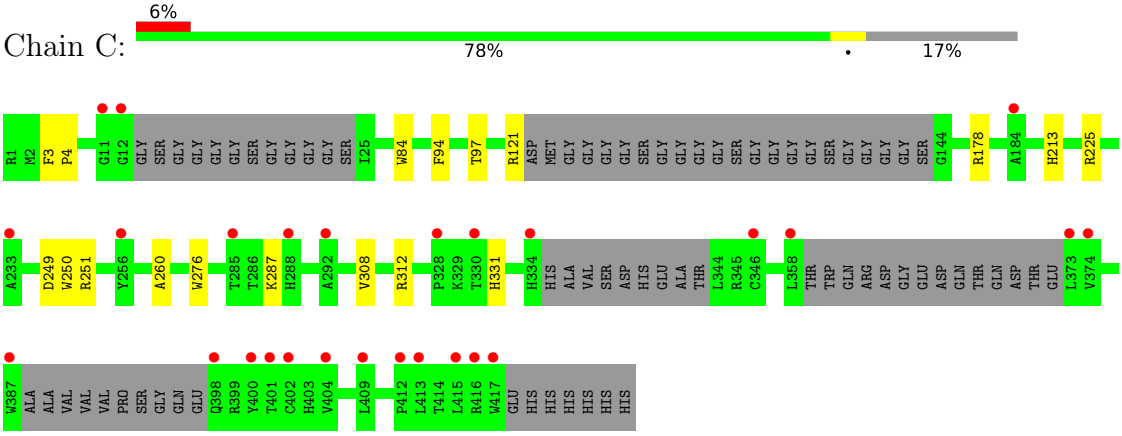
• Molecule 1: Wilms tumor protein peptide,Beta-2-microglobulin,MHC class I antigen chimera

Chain A:



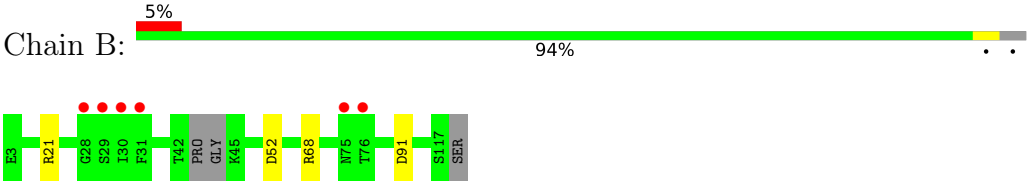
• Molecule 1: Wilms tumor protein peptide,Beta-2-microglobulin,MHC class I antigen chimera

Chain C:



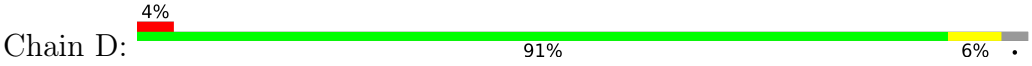
• Molecule 2: VHH

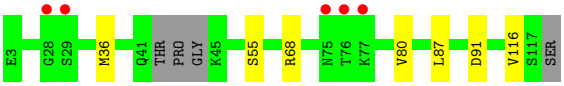
Chain B:



• Molecule 2: VHH

Chain D:




E3
G28
S29
M36
Q41
THR
PRO
GLY
K45
S55
R68
N75
T76
K77
V80
L87
D91
V116
S117
SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	117.48Å 117.48Å 261.27Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.74 – 2.35 48.74 – 2.35	Depositor EDS
% Data completeness (in resolution range)	99.4 (48.74-2.35) 99.4 (48.74-2.35)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.00 (at 2.34Å)	Xtriage
Refinement program	PHENIX 1.19.1_4122	Depositor
R, R_{free}	0.229 , 0.248 0.226 , 0.246	Depositor DCC
R_{free} test set	3825 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	41.0	Xtriage
Anisotropy	0.092	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 33.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	7504	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 35.82 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.4979e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.24	0/3003	0.49	0/4094
1	C	0.24	0/2781	0.49	0/3785
2	B	0.25	0/842	0.52	0/1141
2	D	0.24	0/838	0.51	0/1135
All	All	0.24	0/7464	0.50	0/10155

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	2919	0	2594	8	0
1	C	2706	0	2416	7	0
2	B	829	0	762	3	0
2	D	825	0	759	3	0
3	A	88	0	0	1	0
3	B	44	0	0	1	0
3	C	53	0	0	0	0
3	D	40	0	0	0	0
All	All	7504	0	6531	20	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 1.

All (20) 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)
2:D:68:ARG:NH2	2:D:91:ASP:OD2	2.28	0.64
1:A:213:HIS:NE2	3:A:501:HOH:O	2.31	0.60
2:B:68:ARG:NH2	2:B:91:ASP:OD2	2.33	0.57
1:A:3:PHE:CD1	1:A:4:PRO:HD2	2.43	0.53
2:B:21:ARG:NH2	3:B:201:HOH:O	2.42	0.53
1:C:3:PHE:CD1	1:C:4:PRO:HD2	2.46	0.51
1:A:69:ARG:NH2	2:B:52:ASP:OD1	2.44	0.50
1:C:249:ASP:OD2	1:C:251:ARG:NE	2.47	0.48
1:C:276:TRP:HB2	1:C:287:LYS:HG3	1.99	0.45
1:A:308:VAL:O	1:A:312:ARG:HG3	2.17	0.45
1:C:84:TRP:CE2	1:C:260:ALA:HB2	2.52	0.45
1:C:308:VAL:O	1:C:312:ARG:HG3	2.18	0.44
1:A:250:TRP:HB3	1:A:312:ARG:HD3	1.99	0.44
2:D:87:LEU:HB3	2:D:116:VAL:HG21	2.01	0.42
1:A:149:ARG:NH2	1:A:245:ASP:OD1	2.53	0.42
1:C:97:THR:O	1:C:121:ARG:NH1	2.45	0.41
2:D:36:MET:HB2	2:D:80:VAL:HG11	2.03	0.41
1:A:84:TRP:CE2	1:A:260:ALA:HB2	2.56	0.41
1:C:250:TRP:HB3	1:C:312:ARG:HD3	2.01	0.41
1:A:217:HIS:HA	1:A:220:ASP:HB2	2.03	0.41

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	371/424 (88%)	363 (98%)	8 (2%)	0	100	100
1	C	338/424 (80%)	333 (98%)	5 (2%)	0	100	100
2	B	109/116 (94%)	107 (98%)	2 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	108/116 (93%)	106 (98%)	2 (2%)	0	100	100
All	All	926/1080 (86%)	909 (98%)	17 (2%)	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	282/347 (81%)	276 (98%)	6 (2%)	53	65
1	C	263/347 (76%)	258 (98%)	5 (2%)	57	68
2	B	86/97 (89%)	86 (100%)	0	100	100
2	D	86/97 (89%)	85 (99%)	1 (1%)	71	82
All	All	717/888 (81%)	705 (98%)	12 (2%)	60	72

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	94	PHE
1	A	223	THR
1	A	231	SER
1	A	241	MET
1	A	254	ARG
1	A	300	ARG
1	C	94	PHE
1	C	178	ARG
1	C	213	HIS
1	C	225	ARG
1	C	331	HIS
2	D	55	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	41	ASN
1	A	284	GLN
1	C	331	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

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

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^{th} 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^2)	Q<0.9
1	A	379/424 (89%)	0.25	10 (2%)	56	65	28, 47, 89, 110	0
1	C	350/424 (82%)	0.25	27 (7%)	13	20	30, 50, 103, 156	0
2	B	113/116 (97%)	0.07	6 (5%)	26	38	28, 36, 78, 97	0
2	D	112/116 (96%)	0.05	5 (4%)	33	46	26, 37, 82, 97	0
All	All	954/1080 (88%)	0.20	48 (5%)	28	41	26, 45, 92, 156	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	417	TRP	6.6
1	C	400	TYR	5.2
2	B	30	ILE	5.0
1	C	412	PRO	4.7
2	D	75	ASN	4.6
1	A	342	ALA	3.9
1	A	11	GLY	3.8
2	B	75	ASN	3.7
1	A	400	TYR	3.6
1	C	328	PRO	3.3
1	A	370	ASP	3.3
1	C	233	ALA	3.2
1	A	371	THR	3.2
1	A	332	MET	3.2
1	C	11	GLY	3.0
1	C	402	CYS	3.0
1	A	334	HIS	3.0
2	D	28	GLY	2.9
1	C	373	LEU	2.9
1	A	397	GLU	2.9
2	B	76	THR	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	413	LEU	2.7
1	C	330	THR	2.6
1	A	417	TRP	2.6
2	D	76	THR	2.5
2	B	28	GLY	2.5
1	C	416	ARG	2.5
1	C	334	HIS	2.5
1	C	358	LEU	2.4
1	C	404	VAL	2.4
1	C	346	CYS	2.4
1	C	12	GLY	2.4
1	C	409	LEU	2.3
1	C	292	ALA	2.3
1	A	343	THR	2.3
1	C	288	HIS	2.3
2	B	31	PHE	2.3
2	B	29	SER	2.3
1	C	387	TRP	2.2
1	C	285	THR	2.2
2	D	29	SER	2.2
1	C	398	GLN	2.1
1	C	374	VAL	2.1
1	C	184	ALA	2.1
1	C	401	THR	2.1
1	C	256	TYR	2.1
1	C	415	LEU	2.0
2	D	77	LYS	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.