Full wwPDB X-ray Structure Validation Report ⓘ

Oct 13, 2024 – 03:45 am BST

PDB ID	:	7Z5G
Title	:	human apo MATCAP
Authors	:	Bak, J.; Adamoupolos, A.; Heidebrecht, T.; Perrakis, A.
Deposited on	:	2022-03-09
Resolution	:	2.11 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:

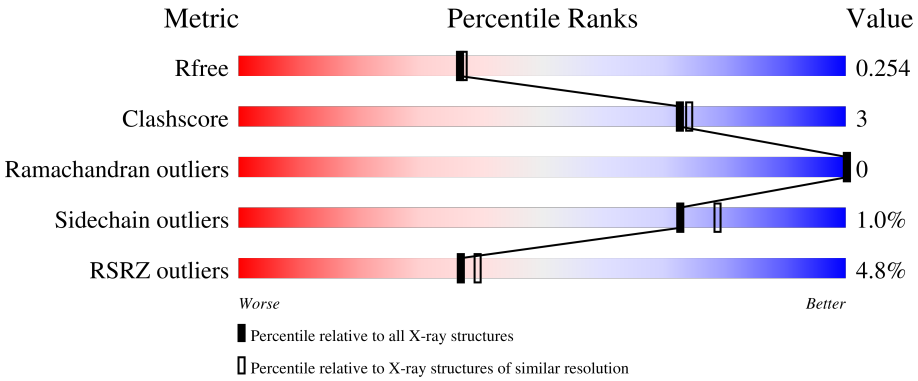
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.11 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	7689 (2.14-2.10)
Clashscore	180529	8431 (2.14-2.10)
Ramachandran outliers	177936	8366 (2.14-2.10)
Sidechain outliers	177891	8367 (2.14-2.10)
RSRZ outliers	164620	7689 (2.14-2.10)

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 >=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	335	93% 6%
1	B	335	12% 92% 6%
1	C	335	4% 90% 8%
1	D	335	95%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 21770 atoms, of which 10714 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 Uncharacterized protein KIAA0895-like.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	332	Total	C	H	N	O	S	77	0	0
			5408	1722	2686	495	491	14			
1	B	331	Total	C	H	N	O	S	77	0	0
			5395	1718	2682	494	487	14			
1	C	330	Total	C	H	N	O	S	77	0	0
			5376	1712	2671	493	486	14			
1	D	331	Total	C	H	N	O	S	77	0	0
			5389	1716	2675	494	490	14			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	82	Total	O	0	0
			82	82		
2	B	17	Total	O	0	0
			17	17		
2	C	32	Total	O	0	0
			32	32		
2	D	71	Total	O	0	0
			71	71		

3 Residue-property plots (i)

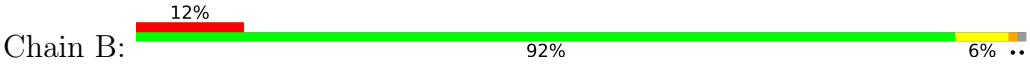
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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.

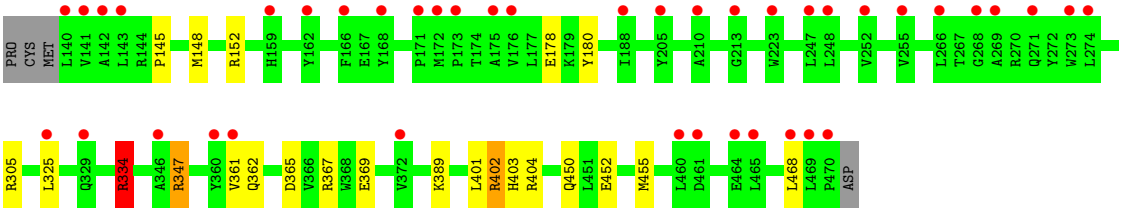
• Molecule 1: Uncharacterized protein KIAA0895-like

Chain A:

• Molecule 1: Uncharacterized protein KIAA0895-like

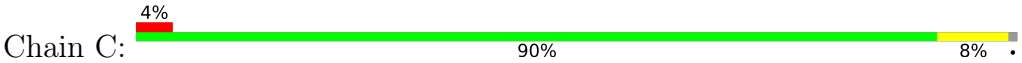
Chain B:

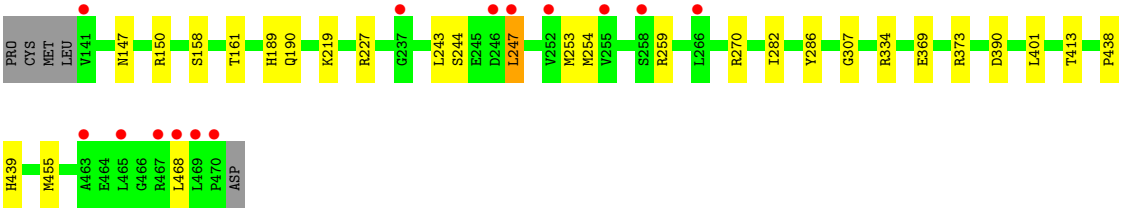




• Molecule 1: Uncharacterized protein KIAA0895-like

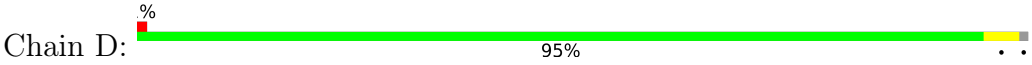
Chain C:

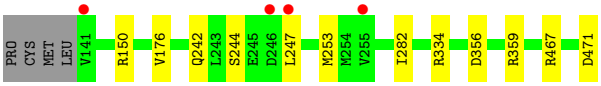




• Molecule 1: Uncharacterized protein KIAA0895-like

Chain D:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	55.85Å 88.03Å 165.61Å 90.00° 90.68° 90.00°	Depositor
Resolution (Å)	47.20 – 2.11 47.20 – 2.11	Depositor EDS
% Data completeness (in resolution range)	98.7 (47.20-2.11) 98.7 (47.20-2.11)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.97 (at 2.12Å)	Xtriage
Refinement program	REFMAC 5.8.0267, REFMAC 5.8.0267	Depositor
R, R_{free}	0.217 , 0.251 0.218 , 0.254	Depositor DCC
R_{free} test set	4629 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	36.4	Xtriage
Anisotropy	0.378	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.42 , 49.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.020 for h,-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	21770	wwPDB-VP
Average B, all atoms (Å^2)	58.0	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 24.39 % 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 3.8679e-03. 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	0.64	0/2787	0.81	1/3771 (0.0%)
1	B	0.55	0/2778	0.78	3/3760 (0.1%)
1	C	0.58	0/2770	0.80	1/3749 (0.0%)
1	D	0.63	0/2779	0.80	1/3760 (0.0%)
All	All	0.60	0/11114	0.80	6/15040 (0.0%)

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	441	MET	CG-SD-CE	-8.09	87.26	100.20
1	D	334	ARG	NE-CZ-NH2	-6.09	117.26	120.30
1	B	347	ARG	NE-CZ-NH2	-5.72	117.44	120.30
1	C	270	ARG	NE-CZ-NH2	-5.33	117.63	120.30
1	B	468	LEU	CA-CB-CG	5.26	127.41	115.30
1	B	334	ARG	NE-CZ-NH1	5.08	122.84	120.30

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	2722	2686	2672	19	0
1	B	2713	2682	2668	17	0
1	C	2705	2671	2657	18	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	2714	2675	2661	14	0
2	A	82	0	0	2	0
2	B	17	0	0	0	0
2	C	32	0	0	3	0
2	D	71	0	0	2	0
All	All	11056	10714	10658	66	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 (66) 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:C:521:HOH:O	1:D:176:VAL:HG23	1.71	0.89
1:C:373:ARG:NH2	1:C:390:ASP:OD2	2.06	0.89
1:B:325:LEU:O	1:B:450:GLN:OE1	1.91	0.88
1:A:441:MET:SD	2:A:561:HOH:O	2.34	0.85
1:B:361:VAL:O	1:B:367:ARG:HD2	1.78	0.83
1:C:307:GLY:HA2	2:C:521:HOH:O	1.80	0.81
1:C:147:ASN:HD22	1:C:150:ARG:HH21	1.31	0.77
1:D:244:SER:OG	1:D:247:LEU:CD1	2.46	0.63
1:C:244:SER:OG	1:C:247:LEU:HD23	2.00	0.62
1:A:409:PHE:CB	1:A:441:MET:HE1	2.31	0.61
1:B:402:ARG:HD2	1:B:403:HIS:CE1	2.36	0.60
1:D:467:ARG:NH1	2:D:501:HOH:O	2.17	0.59
1:A:152:ARG:NH2	1:A:365:ASP:OD1	2.36	0.57
1:B:362:GLN:O	1:B:367:ARG:NH1	2.34	0.57
1:B:401:LEU:HB3	1:B:455:MET:HE2	1.87	0.56
1:B:152:ARG:NH2	1:B:365:ASP:OD1	2.38	0.56
1:A:459:ARG:HH11	1:A:459:ARG:HG3	1.71	0.54
1:A:294:ARG:O	1:A:431:VAL:HG11	2.08	0.53
1:B:145:PRO:HD2	1:B:148:MET:HE2	1.90	0.53
1:A:410:PRO:HD3	1:A:441:MET:HE1	1.90	0.52
1:C:219:LYS:HG3	1:C:243:LEU:HD23	1.91	0.52
1:C:334:ARG:NH1	1:C:369:GLU:OE2	2.32	0.52
1:A:343:HIS:CD2	1:A:344:ARG:HH11	2.28	0.51
1:D:247:LEU:HD12	1:D:247:LEU:N	2.25	0.51
1:B:334:ARG:NH1	1:B:369:GLU:OE1	2.42	0.51
1:D:253:MET:HE1	1:D:282:ILE:HD13	1.92	0.51
1:C:253:MET:CE	1:C:282:ILE:HD13	2.41	0.50
1:B:404:ARG:HD3	1:B:455:MET:CE	2.42	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:244:SER:CB	1:D:247:LEU:HD11	2.42	0.49
1:B:145:PRO:HD2	1:B:148:MET:CE	2.41	0.49
1:C:219:LYS:HG3	1:C:243:LEU:CD2	2.43	0.49
1:A:334:ARG:NH1	1:A:369:GLU:OE2	2.45	0.49
1:D:356:ASP:OD1	1:D:359:ARG:NH2	2.46	0.48
1:C:373:ARG:HH22	1:C:390:ASP:CG	2.16	0.48
1:D:253:MET:CE	1:D:282:ILE:HD13	2.43	0.48
1:C:158:SER:HG	1:C:161:THR:HG1	1.60	0.47
1:A:409:PHE:HB2	1:A:441:MET:HE1	1.97	0.47
1:D:150:ARG:NH1	2:D:503:HOH:O	2.48	0.47
1:C:401:LEU:HD13	1:C:455:MET:HG2	1.97	0.46
1:A:180:TYR:CZ	1:A:389:LYS:HD3	2.51	0.46
1:D:244:SER:OG	1:D:247:LEU:HD11	2.15	0.46
1:C:190:GLN:HE22	1:C:468:LEU:HD13	1.80	0.46
1:B:404:ARG:HD3	1:B:455:MET:HE3	1.97	0.46
1:A:171:PRO:O	1:B:305:ARG:NH1	2.42	0.45
1:C:413:THR:HG21	1:C:438:PRO:HD3	1.98	0.45
1:A:459:ARG:NH1	2:A:504:HOH:O	2.49	0.45
1:D:244:SER:CB	1:D:247:LEU:CD1	2.94	0.45
1:B:404:ARG:NH1	1:B:455:MET:HE1	2.33	0.44
1:C:439:HIS:HB2	2:C:504:HOH:O	2.17	0.44
1:A:382:SER:OG	1:A:383:LEU:CD1	2.66	0.44
1:A:459:ARG:HG3	1:A:459:ARG:NH1	2.31	0.44
1:B:401:LEU:HB3	1:B:455:MET:CE	2.47	0.43
1:A:247:LEU:HD23	1:A:247:LEU:HA	1.87	0.43
1:C:253:MET:CE	1:C:286:TYR:HB2	2.49	0.43
1:D:244:SER:HB3	1:D:247:LEU:CD1	2.49	0.43
1:C:190:GLN:NE2	1:C:468:LEU:HD13	2.34	0.43
1:C:253:MET:HE1	1:C:282:ILE:HD13	2.01	0.43
1:B:452:GLU:HA	1:B:455:MET:HG2	2.00	0.43
1:D:244:SER:HB3	1:D:247:LEU:HD11	2.00	0.43
1:A:343:HIS:HD2	1:A:344:ARG:HH11	1.65	0.42
1:B:180:TYR:CZ	1:B:389:LYS:HD3	2.54	0.42
1:D:247:LEU:CD1	1:D:247:LEU:N	2.83	0.42
1:A:410:PRO:CD	1:A:441:MET:HE1	2.50	0.42
1:B:347:ARG:HB3	1:C:189:HIS:CE1	2.55	0.42
1:A:410:PRO:HD3	1:A:441:MET:CE	2.50	0.42
1:A:413:THR:HG21	1:A:438:PRO:HD3	2.01	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	330/335 (98%)	323 (98%)	7 (2%)	0	100	100
1	B	329/335 (98%)	321 (98%)	8 (2%)	0	100	100
1	C	328/335 (98%)	321 (98%)	7 (2%)	0	100	100
1	D	329/335 (98%)	322 (98%)	7 (2%)	0	100	100
All	All	1316/1340 (98%)	1287 (98%)	29 (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	288/291 (99%)	286 (99%)	2 (1%)	81	86
1	B	287/291 (99%)	284 (99%)	3 (1%)	73	79
1	C	286/291 (98%)	282 (99%)	4 (1%)	62	69
1	D	287/291 (99%)	285 (99%)	2 (1%)	81	86
All	All	1148/1164 (99%)	1137 (99%)	11 (1%)	73	79

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	245	GLU
1	A	383	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	178	GLU
1	B	334	ARG
1	B	402	ARG
1	C	227	ARG
1	C	247	LEU
1	C	254	MET
1	C	259	ARG
1	D	242	GLN
1	D	471	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	343	HIS
1	B	163	ASN
1	B	439	HIS
1	B	442	GLN
1	B	450	GLN
1	C	189	HIS
1	C	439	HIS
1	D	439	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^{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	332/335 (99%)	-0.02	5 (1%) 71 73	26, 48, 79, 113	0
1	B	331/335 (98%)	0.93	41 (12%) 9 10	43, 72, 104, 129	0
1	C	330/335 (98%)	0.21	14 (4%) 41 44	36, 56, 92, 127	0
1	D	331/335 (98%)	-0.20	4 (1%) 76 77	26, 44, 76, 107	0
All	All	1324/1340 (98%)	0.23	64 (4%) 36 39	26, 55, 94, 129	0

All (64) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	255	VAL	4.9
1	A	266	LEU	4.4
1	B	470	PRO	4.3
1	C	470	PRO	4.1
1	C	468	LEU	3.8
1	D	255	VAL	3.8
1	C	465	LEU	3.7
1	A	140	LEU	3.6
1	B	252	VAL	3.4
1	B	140	LEU	3.4
1	C	463	ALA	3.3
1	B	273	TRP	3.2
1	B	176	VAL	3.2
1	D	141	VAL	3.2
1	B	205	TYR	3.2
1	B	274	LEU	3.2
1	B	210	ALA	3.2
1	B	469	LEU	3.2
1	C	141	VAL	3.1
1	A	255	VAL	3.0
1	B	468	LEU	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	171	PRO	3.0
1	B	247	LEU	3.0
1	B	255	VAL	2.9
1	B	461	ASP	2.9
1	B	141	VAL	2.8
1	B	266	LEU	2.8
1	B	465	LEU	2.8
1	B	188	ILE	2.8
1	B	460	LEU	2.7
1	B	361	VAL	2.7
1	B	175	ALA	2.7
1	D	247	LEU	2.6
1	B	248	LEU	2.6
1	B	172	MET	2.6
1	B	223	TRP	2.5
1	B	268	GLY	2.5
1	C	247	LEU	2.5
1	B	213	GLY	2.5
1	B	173	PRO	2.3
1	B	372	VAL	2.3
1	B	346	ALA	2.3
1	C	252	VAL	2.3
1	B	166	PHE	2.3
1	B	360	TYR	2.3
1	B	325	LEU	2.3
1	B	271	GLN	2.3
1	B	159	HIS	2.2
1	C	467	ARG	2.2
1	A	247	LEU	2.2
1	D	246	ASP	2.1
1	B	142	ALA	2.1
1	B	269	ALA	2.1
1	B	162	TYR	2.1
1	C	237	GLY	2.1
1	B	329	GLN	2.1
1	B	143	LEU	2.1
1	C	266	LEU	2.1
1	C	469	LEU	2.1
1	A	142	ALA	2.1
1	B	168	TYR	2.1
1	C	258	SER	2.0
1	B	464	GLU	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	246	ASP	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.