WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Oct 19, 2023 – 04:09 am BST

PDB ID	:	8PNL
Title	:	Outward-open conformation of a Major Facilitator Superfamily (MFS) transporter MHAS2168, a homologue of Rv1410 from M. tuberculosis, in complex with an alpaca nanobody
Authors	:	Remm, S.; Schoeppe, J.; Hutter, C.A.J.; Gonda, I.; Seeger, M.A.
Deposited on	:	2023-06-30
Resolution	:	2.70 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 (i) 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 (i)) were used in the production of this report:

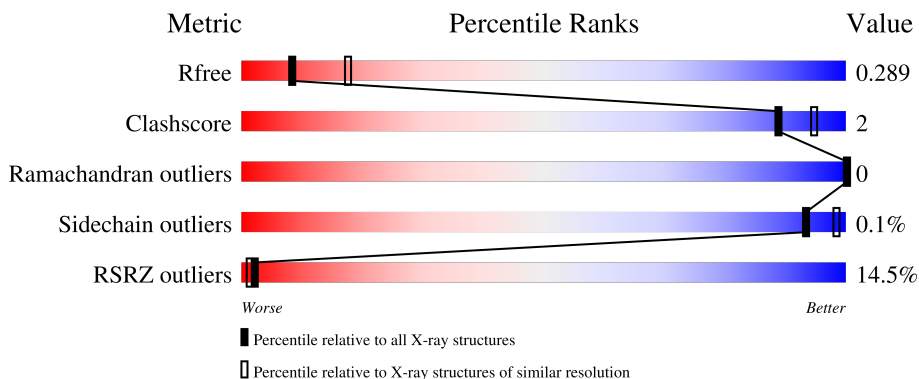|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.70 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	Percentile Ranks	Value
Rfree		0.289
Clashscore		2
Ramachandran outliers		0
Sidechain outliers		0.1%
RSRZ outliers		14.5%

Worse　　　　　*Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	543	7% / 85% · 11%
1	C	543	12% / 85% 5% 10%
2	B	121	15% / 88% · 8%
2	D	121	45% / 84% 7% 9%

WORLDWIDE PDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 18026 atoms, of which 9202 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 Putative triacylglyceride transporter.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	483	Total 7310	C 2341	H 3758	N 601	O 600	S 10	0	0	0
1	C	486	Total 7358	C 2356	H 3781	N 605	O 606	S 10	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP K5B8L6
A	1	SER	-	expression tag	UNP K5B8L6
A	536	ALA	-	expression tag	UNP K5B8L6
A	537	LEU	-	expression tag	UNP K5B8L6
A	538	GLU	-	expression tag	UNP K5B8L6
A	539	VAL	-	expression tag	UNP K5B8L6
A	540	LEU	-	expression tag	UNP K5B8L6
A	541	PHE	-	expression tag	UNP K5B8L6
A	542	GLN	-	expression tag	UNP K5B8L6
C	0	MET	-	initiating methionine	UNP K5B8L6
C	1	SER	-	expression tag	UNP K5B8L6
C	536	ALA	-	expression tag	UNP K5B8L6
C	537	LEU	-	expression tag	UNP K5B8L6
C	538	GLU	-	expression tag	UNP K5B8L6
C	539	VAL	-	expression tag	UNP K5B8L6
C	540	LEU	-	expression tag	UNP K5B8L6
C	541	PHE	-	expression tag	UNP K5B8L6
C	542	GLN	-	expression tag	UNP K5B8L6

- Molecule 2 is a protein called Nb_H2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	111	Total 1688	C 533	H 836	N 153	O 161	S 5	0	0	0

Continued on next page...

WORLDWIDE wwPDB PROTEIN DATA BANK

Continued from previous page...

| Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	D	110	1670	528	827	151	159	5	0	0	0

3 Residue-property plots ⓘ

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.
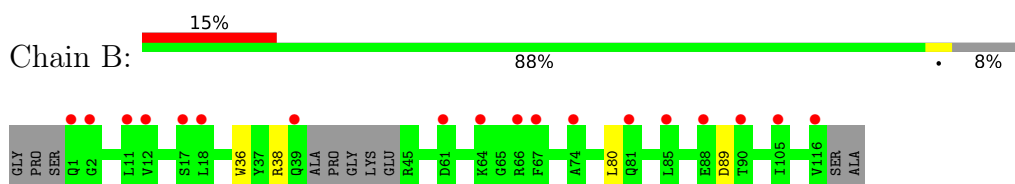
• Molecule 1: Putative triacylglyceride transporter

Chain A:
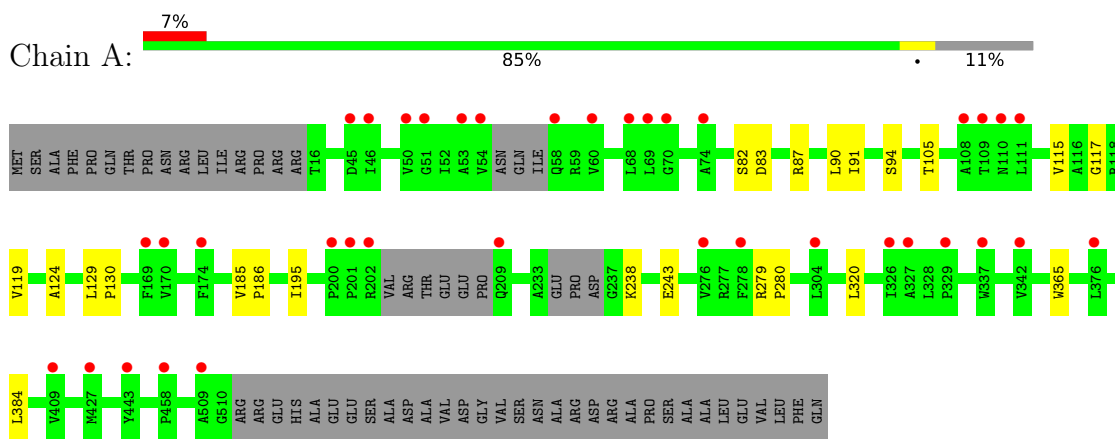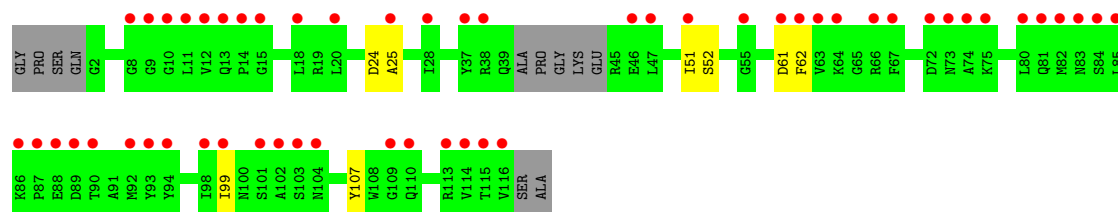
• Molecule 1: Putative triacylglyceride transporter
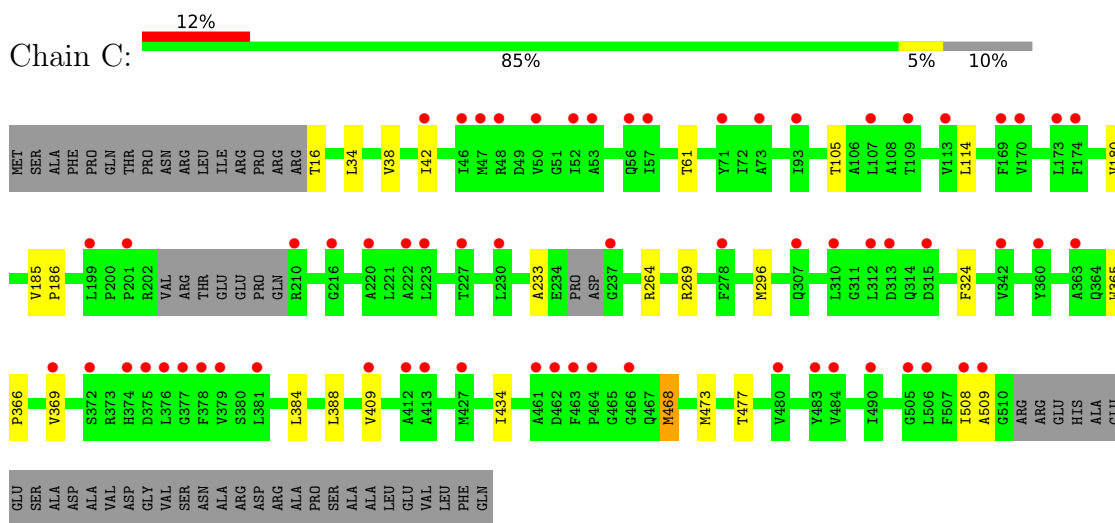
Chain C:

• Molecule 2: Nb_H2

Chain B:

• Molecule 2: Nb_H2

Chain D:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	57.78Å 160.70Å 82.96Å 90.00° 109.01° 90.00°	Depositor
Resolution (Å)	44.70 – 2.70 44.70 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.4 (44.70-2.70) 99.5 (44.70-2.70)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.35 (at 2.69Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.245 , 0.291 0.240 , 0.289	Depositor DCC
R_{free} test set	1952 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	79.6	Xtriage
Anisotropy	0.244	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 66.1	EDS
L-test for twinning[2]	$< \lvert L \rvert > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.025 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	18026	wwPDB-VP
Average B, all atoms (Å^2)	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.05% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ 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	0.26	0/3622	0.53	0/4948
1	C	0.25	0/3648	0.52	0/4985
2	B	0.27	0/867	0.63	0/1168
2	D	0.26	0/858	0.63	0/1156
All	All	0.26	0/8995	0.55	0/12257

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	3552	3758	3755	12	2
1	C	3577	3781	3779	18	2
2	B	852	836	837	2	0
2	D	843	827	826	5	0
All	All	8824	9202	9197	35	2

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:C:468:MET:SD	1:C:468:MET:N	2.59	0.75
1:A:365:TRP:CH2	1:A:384:LEU:HD12	2.40	0.57
1:C:365:TRP:CH2	1:C:384:LEU:HD12	2.39	0.57
1:C:473:MET:O	1:C:477:THR:HG23	2.06	0.56
1:C:508:ILE:HG22	1:C:509:ALA:H	1.71	0.55
1:A:91:ILE:CG2	1:A:195:ILE:HD13	2.38	0.53
1:C:477:THR:HG22	2:D:99:ILE:HG13	1.94	0.49
1:A:83:ASP:OD1	1:A:87:ARG:NH2	2.47	0.47
1:A:91:ILE:HG21	1:A:195:ILE:HD13	1.97	0.47
1:A:185:VAL:HB	1:A:186:PRO:HD3	1.96	0.47
1:C:477:THR:HG21	2:D:107:TYR:HB2	1.96	0.47
1:C:185:VAL:HB	1:C:186:PRO:HD3	1.98	0.45
1:A:129:LEU:HB3	1:A:130:PRO:HD3	1.99	0.45
1:C:296:MET:HG3	1:C:434:ILE:HG21	1.99	0.44
2:D:51:ILE:HG22	2:D:52:SER:O	2.17	0.44
1:C:38:VAL:HG11	1:C:180:VAL:HG11	1.99	0.44
1:C:409:VAL:O	1:C:409:VAL:CG1	2.66	0.44
1:A:279:ARG:HB3	1:A:280:PRO:HD3	1.98	0.44
2:B:38:ARG:NH2	2:B:89:ASP:OD1	2.51	0.43
1:A:320:LEU:HD22	1:A:384:LEU:HD21	2.00	0.43
2:D:24:ASP:OD1	2:D:25:ALA:N	2.51	0.43
1:C:42:ILE:O	1:C:42:ILE:HG22	2.19	0.43
1:C:366:PRO:O	1:C:369:VAL:HG23	2.17	0.43
2:D:61:ASP:OD1	2:D:62:PHE:N	2.52	0.43
1:A:105:THR:HG23	1:A:117:GLY:HA3	2.01	0.43
1:A:82:SER:HA	1:A:90:LEU:HD12	2.00	0.42
1:C:508:ILE:HG22	1:C:509:ALA:N	2.34	0.42
2:B:36:TRP:CD1	2:B:80:LEU:HG	2.55	0.42
1:A:115:VAL:O	1:A:119:VAL:HG23	2.21	0.41
1:C:324:PHE:HB2	1:C:388:LEU:HB3	2.03	0.41
1:C:105:THR:HG23	1:C:114:LEU:HD23	2.03	0.41
1:C:264:ARG:HA	1:C:269:ARG:NH2	2.36	0.41
1:A:94:SER:O	1:A:124:ALA:HB1	2.21	0.40
1:C:61:THR:HG21	1:C:233:ALA:HB1	2.03	0.40
1:C:34:LEU:O	1:C:38:VAL:HG23	2.21	0.40

All (2) 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:A:243:GLU:OE2	1:C:16:THR:OG1[1_556]	1.80	0.40
1:A:238:LYS:HZ1	1:C:468:MET:SD[2_555]	1.51	0.09

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	475/543 (88%)	471 (99%)	4 (1%)	0	100	100
1	C	480/543 (88%)	471 (98%)	9 (2%)	0	100	100
2	B	107/121 (88%)	105 (98%)	2 (2%)	0	100	100
2	D	106/121 (88%)	101 (95%)	5 (5%)	0	100	100
All	All	1168/1328 (88%)	1148 (98%)	20 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	351/402 (87%)	351 (100%)	0	100	100
1	C	354/402 (88%)	353 (100%)	1 (0%)	92	98
2	B	89/95 (94%)	89 (100%)	0	100	100
2	D	88/95 (93%)	88 (100%)	0	100	100
All	All	882/994 (89%)	881 (100%)	1 (0%)	93	98

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

Mol	Chain	Res	Type
1	C	468	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are

no such sidechains identified.

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, 95th 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	483/543 (88%)	0.70	37 (7%)	13	11	62, 82, 110, 134	0
1	C	486/543 (89%)	0.96	64 (13%)	3	2	66, 87, 116, 138	0
2	B	111/121 (91%)	0.91	18 (16%)	1	1	64, 93, 104, 115	0
2	D	110/121 (90%)	2.40	54 (49%)	0	0	106, 137, 151, 156	0
All	All	1190/1328 (89%)	0.98	173 (14%)	2	1	62, 87, 136, 156	0

All (173) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	10	GLY	10.5
2	D	11	LEU	10.4
2	D	114	VAL	7.6
2	B	2	GLY	6.8
1	C	509	ALA	6.8
2	D	64	LYS	6.4
2	D	116	VAL	6.3
1	C	381	LEU	6.3
2	D	47	LEU	6.2
2	B	1	GLN	6.2
1	C	379	VAL	6.0
2	D	25	ALA	6.0
2	D	74	ALA	5.7
1	C	46	ILE	5.6
2	D	90	THR	5.6
1	A	53	ALA	5.5
2	D	62	PHE	5.3
2	D	102	ALA	5.3
2	B	66	ARG	5.2
2	D	103	SER	5.1
2	B	116	VAL	5.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	115	THR	5.0
1	C	464	PRO	4.9
2	D	93	TYR	4.8
2	D	85	LEU	4.8
2	D	18	LEU	4.7
2	D	87	PRO	4.7
2	D	84	SER	4.6
1	C	109	THR	4.5
2	D	15	GLY	4.5
2	D	9	GLY	4.4
2	D	92	MET	4.3
1	C	42	ILE	4.3
2	D	12	VAL	4.2
1	C	57	ILE	4.2
1	C	50	VAL	4.2
2	B	85	LEU	4.1
2	D	113	ARG	3.9
1	C	483	TYR	3.9
1	C	376	LEU	3.8
2	D	13	GLN	3.8
1	C	52	ILE	3.7
2	D	89	ASP	3.7
1	A	202	ARG	3.7
1	C	461	ALA	3.6
2	B	64	LYS	3.6
1	A	109	THR	3.6
1	C	369	VAL	3.5
1	A	458	PRO	3.5
2	D	73	ASN	3.5
2	D	86	LYS	3.5
1	C	375	ASP	3.4
1	A	443	TYR	3.4
1	C	342	VAL	3.4
1	A	108	ALA	3.3
2	D	75	LYS	3.3
2	D	98	ILE	3.3
1	A	169	PHE	3.3
1	C	173	LEU	3.3
1	C	466	GLY	3.2
2	D	67	PHE	3.2
2	B	61	ASP	3.2
2	D	38	ARG	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	67	PHE	3.2
2	B	74	ALA	3.2
1	C	310	LEU	3.2
1	C	278	PHE	3.1
1	C	463	PHE	3.1
1	A	51	GLY	3.1
2	B	11	LEU	3.1
2	D	104	ASN	3.1
2	D	55	GLY	3.1
2	D	72	ASP	3.1
2	D	14	PRO	3.0
1	C	360	TYR	3.0
2	D	109	GLY	3.0
1	A	276	VAL	3.0
1	C	53	ALA	3.0
1	C	312	LEU	3.0
2	D	61	ASP	3.0
2	B	12	VAL	2.9
1	A	278	PHE	2.9
1	A	409	VAL	2.9
1	A	111	LEU	2.9
1	A	209	GLN	2.9
2	B	88	GLU	2.8
2	D	37	TYR	2.8
1	C	412	ALA	2.8
2	D	8	GLY	2.8
1	A	427	MET	2.8
1	C	170	VAL	2.8
1	C	223	LEU	2.7
1	A	376	LEU	2.6
1	C	508	ILE	2.6
1	C	230	LEU	2.6
2	D	66	ARG	2.6
2	B	81	GLN	2.6
1	C	377	GLY	2.6
1	A	337	TRP	2.6
2	D	110	GLN	2.6
1	C	48	ARG	2.6
1	C	210	ARG	2.6
1	C	427	MET	2.6
1	C	222	ALA	2.6
1	A	69	LEU	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	20	LEU	2.5
2	D	82	MET	2.5
2	D	63	VAL	2.5
1	C	378	PHE	2.5
2	D	83	ASN	2.5
1	C	113	VAL	2.5
1	C	409	VAL	2.5
2	B	90	THR	2.5
1	A	54	VAL	2.5
2	B	17	SER	2.5
2	D	88	GLU	2.4
1	C	413	ALA	2.4
1	A	200	PRO	2.4
1	A	50	VAL	2.4
2	B	39	GLN	2.4
1	C	462	ASP	2.4
1	C	174	PHE	2.4
2	D	94	TYR	2.4
2	D	51	ILE	2.4
1	C	237	GLY	2.4
1	A	110	ASN	2.4
1	C	216	GLY	2.4
1	C	484	VAL	2.4
1	A	326	ILE	2.3
2	D	46	GLU	2.3
1	A	304	LEU	2.3
1	A	46	ILE	2.3
1	C	506	LEU	2.3
1	A	329	PRO	2.2
1	C	201	PRO	2.2
1	A	174	PHE	2.2
1	A	201	PRO	2.2
2	D	101	SER	2.2
1	A	509	ALA	2.2
1	A	70	GLY	2.2
2	B	18	LEU	2.2
1	C	363	ALA	2.2
1	A	45	ASP	2.2
2	D	81	GLN	2.2
1	C	71	TYR	2.2
1	C	73	ALA	2.2
1	C	505	GLY	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	93	ILE	2.2
1	C	227	THR	2.2
1	A	342	VAL	2.1
1	C	372	SER	2.1
1	C	220	ALA	2.1
1	C	169	PHE	2.1
1	C	490	ILE	2.1
2	B	105	ILE	2.1
1	A	68	LEU	2.1
1	C	315	ASP	2.1
1	A	74	ALA	2.1
1	A	327	ALA	2.1
2	D	99	ILE	2.1
1	C	374	HIS	2.1
2	D	28	ILE	2.1
1	C	56	GLN	2.1
1	C	47	MET	2.1
1	A	58	GLN	2.0
1	A	60	VAL	2.0
1	C	480	VAL	2.0
1	C	307	GLN	2.0
1	C	107	LEU	2.0
1	C	199	LEU	2.0
1	C	313	ASP	2.0
2	D	80	LEU	2.0
1	A	170	VAL	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 ⓘ

There are no monosaccharides in this entry.

6.4   Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.