Full wwPDB X-ray Structure Validation Report ⓘ

May 17, 2020 – 11:42 am BST

PDB ID	:	5LM5
Title	:	Structure of C-terminal domain from S. cerevisiae Pat1 decapping activator bound to Dcp2 HLM2 peptide (region 435-451)
Authors	:	Charenton, C.; Gaudon-Plesse, C.; Fourati, Z.; Taverniti, V.; Back, R.; Kolesnikova, O.; Seraphin, B.; Graille, M.
Deposited on	:	2016-07-29
Resolution	:	2.60 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 following versions of software and data (see references ①) were used in the production of this report:

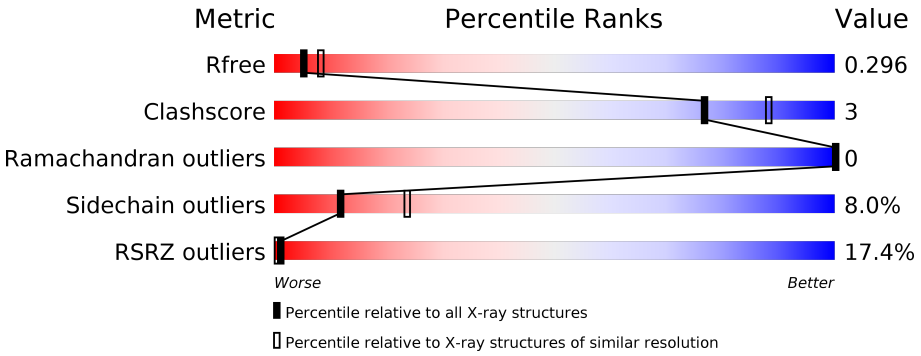
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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 on 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	369	11% 76% 9% • 15%
1	B	369	20% 72% 12% • 15%
2	C	14	7% 43% 43% 14%
2	D	14	14% 71% 14% 14%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5339 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 DNA topoisomerase 2-associated protein PAT1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	315	Total	C	N	O	S	0	0	0
			2574	1664	419	483	8			
1	B	315	Total	C	N	O	S	0	0	0
			2578	1669	420	481	8			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	428	MET	-	initiating methionine	UNP P25644
A	429	HIS	-	expression tag	UNP P25644
A	430	HIS	-	expression tag	UNP P25644
A	431	HIS	-	expression tag	UNP P25644
A	432	HIS	-	expression tag	UNP P25644
A	433	HIS	-	expression tag	UNP P25644
A	434	HIS	-	expression tag	UNP P25644
A	706	ALA	GLN	engineered mutation	UNP P25644
A	713	ALA	LEU	engineered mutation	UNP P25644
B	428	MET	-	initiating methionine	UNP P25644
B	429	HIS	-	expression tag	UNP P25644
B	430	HIS	-	expression tag	UNP P25644
B	431	HIS	-	expression tag	UNP P25644
B	432	HIS	-	expression tag	UNP P25644
B	433	HIS	-	expression tag	UNP P25644
B	434	HIS	-	expression tag	UNP P25644
B	706	ALA	GLN	engineered mutation	UNP P25644
B	713	ALA	LEU	engineered mutation	UNP P25644

- Molecule 2 is a protein called mRNA decapping protein 2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	14	Total	C	N	O	0	0	0
			101	62	17	22			

Continued on next page...

Continued from previous page...

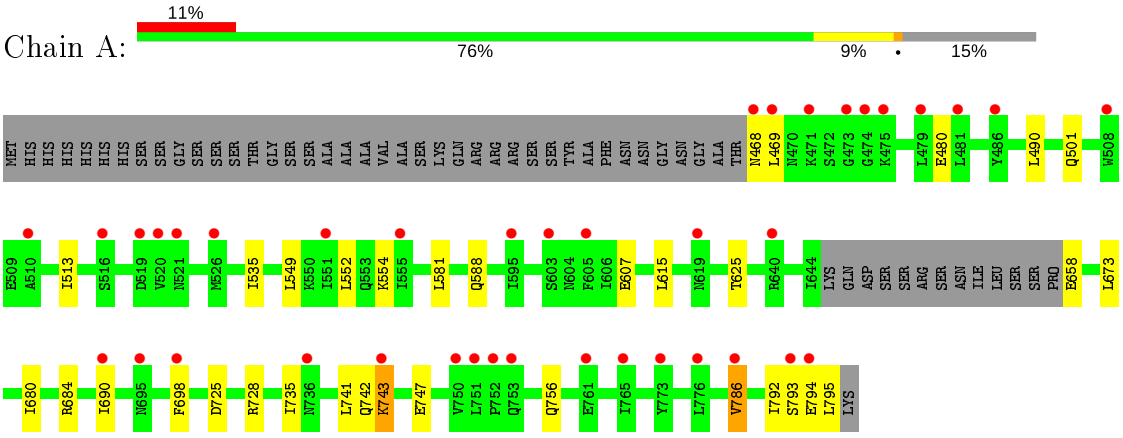
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	D	12	Total	C	N	O	0	0	0
			86	53	14	19			

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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.

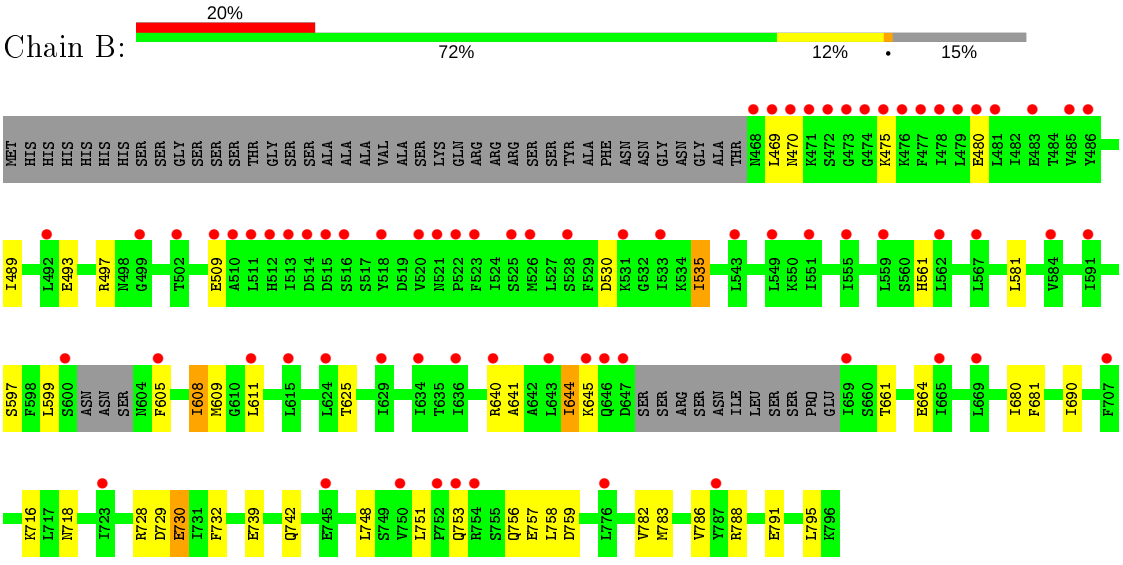
- Molecule 1: DNA topoisomerase 2-associated protein PAT1

Chain A:



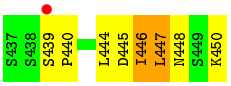
- Molecule 1: DNA topoisomerase 2-associated protein PAT1

Chain B:



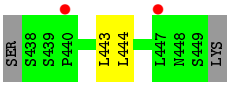
- Molecule 2: mRNA decapping protein 2

Chain C:



- Molecule 2: mRNA decapping protein 2

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	98.28Å 115.96Å 122.01Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.14 – 2.60 49.14 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.2 (49.14-2.60) 99.2 (49.14-2.60)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.06 (at 2.61Å)	Xtriage
Refinement program	BUSTER-TNT 2.10.2	Depositor
R, R_{free}	0.256 , 0.268 0.280 , 0.296	Depositor DCC
R_{free} test set	1126 reflections (5.20%)	wwPDB-VP
Wilson B-factor (Å^2)	64.6	Xtriage
Anisotropy	0.735	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 46.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.000 for -h,-l,-k	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5339	wwPDB-VP
Average B, all atoms (Å^2)	108.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.58% of the height of the origin peak. No significant pseudotranslation is detected.*

[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	0.46	0/2614	0.63	0/3525
1	B	0.46	0/2617	0.66	0/3525
2	C	0.61	0/101	0.87	0/135
2	D	0.54	0/86	0.81	0/116
All	All	0.46	0/5418	0.65	0/7301

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	2574	0	2657	14	0
1	B	2578	0	2671	19	0
2	C	101	0	104	5	0
2	D	86	0	86	1	0
All	All	5339	0	5518	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 3.

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:B:609:MET:HE3	1:B:641:ALA:HB1	1.58	0.85
1:B:625:THR:HB	1:B:680:ILE:HG22	1.75	0.68
1:A:725:ASP:HB2	2:C:446:ILE:HD11	1.75	0.68
1:B:599:LEU:HB3	1:B:640:ARG:HG2	1.81	0.63
1:B:489:ILE:HG21	1:B:535:ILE:HD11	1.83	0.60
1:B:605:PHE:CD1	1:B:644:ILE:HG23	2.36	0.60
1:B:605:PHE:HB2	1:B:644:ILE:HD12	1.84	0.59
1:B:661:THR:O	1:B:664:GLU:HG2	2.03	0.59
1:A:625:THR:HB	1:A:680:ILE:HG22	1.84	0.58
1:A:794:GLU:HG2	2:C:440:PRO:HB2	1.86	0.56
1:B:609:MET:CE	1:B:641:ALA:HB1	2.32	0.55
1:B:728:ARG:HG3	1:B:729:ASP:N	2.23	0.53
1:A:698:PHE:CZ	1:A:741:LEU:HD21	2.44	0.53
1:A:786:VAL:CG1	1:A:795:LEU:HD13	2.39	0.53
1:B:608:ILE:HG12	1:B:641:ALA:HB2	1.93	0.51
1:B:581:LEU:HD21	1:B:690:ILE:HD12	1.93	0.50
1:B:728:ARG:O	1:B:732:PHE:HD1	1.96	0.48
1:A:581:LEU:HD21	1:A:690:ILE:HD12	1.95	0.48
1:A:786:VAL:HG13	1:A:795:LEU:HD13	1.97	0.46
1:A:549:LEU:HD21	1:A:607:GLU:HG2	1.98	0.46
1:A:728:ARG:HG3	2:C:447:LEU:HD22	1.97	0.46
1:A:735:ILE:HD11	1:A:792:ILE:HD11	1.98	0.45
1:A:490:LEU:HD13	1:B:795:LEU:HD21	1.97	0.45
1:B:605:PHE:CD1	1:B:644:ILE:CG2	2.99	0.44
1:A:793:SER:HA	2:C:444:LEU:HD11	2.00	0.44
1:B:681:PHE:CD2	1:B:730:GLU:HG2	2.53	0.43
2:C:439:SER:N	2:C:440:PRO:HD2	2.34	0.42
1:B:493:GLU:O	1:B:497:ARG:HD3	2.20	0.42
1:A:743:LYS:HE3	1:A:747:GLU:HG3	2.01	0.41
1:A:625:THR:HG21	1:A:673:LEU:HD21	2.03	0.40
1:B:748:LEU:HA	1:B:751:LEU:HD12	2.04	0.40
1:B:681:PHE:CG	1:B:730:GLU:HG2	2.56	0.40
1:B:783:MET:HB3	2:D:443:LEU:HD13	2.03	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	311/369 (84%)	309 (99%)	2 (1%)	0	100	100
1	B	309/369 (84%)	306 (99%)	3 (1%)	0	100	100
2	C	12/14 (86%)	8 (67%)	4 (33%)	0	100	100
2	D	10/14 (71%)	8 (80%)	2 (20%)	0	100	100
All	All	642/766 (84%)	631 (98%)	11 (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	294/338 (87%)	278 (95%)	16 (5%)	22	44
1	B	294/338 (87%)	267 (91%)	27 (9%)	9	17
2	C	13/13 (100%)	8 (62%)	5 (38%)	0	0
2	D	11/13 (85%)	10 (91%)	1 (9%)	9	18
All	All	612/702 (87%)	563 (92%)	49 (8%)	12	24

All (49) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	468	ASN
1	A	469	LEU
1	A	480	GLU
1	A	501	GLN
1	A	513	ILE
1	A	535	ILE
1	A	552	LEU
1	A	554	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	588	GLN
1	A	615	LEU
1	A	658	GLU
1	A	684	ARG
1	A	742	GLN
1	A	743	LYS
1	A	756	GLN
1	A	786	VAL
1	B	469	LEU
1	B	470	ASN
1	B	475	LYS
1	B	480	GLU
1	B	509	GLU
1	B	530	ASP
1	B	535	ILE
1	B	561	HIS
1	B	597	SER
1	B	608	ILE
1	B	611	LEU
1	B	644	ILE
1	B	645	LYS
1	B	716	LYS
1	B	718	ASN
1	B	730	GLU
1	B	739	GLU
1	B	742	GLN
1	B	753	GLN
1	B	756	GLN
1	B	757	GLU
1	B	758	LEU
1	B	759	ASP
1	B	782	VAL
1	B	786	VAL
1	B	788	ARG
1	B	791	GLU
2	C	445	ASP
2	C	446	ILE
2	C	447	LEU
2	C	448	ASN
2	C	450	LYS
2	D	444	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such

sidechains are listed below:

Mol	Chain	Res	Type
1	B	561	HIS
2	D	442	GLN

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 carbohydrates 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	315/369 (85%)	0.78	39 (12%) 4 2	71, 98, 140, 169	0
1	B	315/369 (85%)	1.16	72 (22%) 0 0	73, 114, 155, 178	0
2	C	14/14 (100%)	0.86	1 (7%) 16 11	74, 90, 118, 119	0
2	D	12/14 (85%)	0.76	2 (16%) 1 1	86, 96, 112, 113	0
All	All	656/766 (85%)	0.96	114 (17%) 1 0	71, 105, 149, 178	0

All (114) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	479	LEU	7.9
1	A	520	VAL	7.6
1	B	472	SER	7.2
1	B	752	PRO	6.3
1	B	514	ASP	5.7
1	B	659	ILE	5.5
1	B	476	LYS	5.5
1	B	510	ALA	5.1
1	B	473	GLY	5.1
1	A	510	ALA	5.1
1	A	521	ASN	5.1
1	B	551	ILE	4.8
1	A	473	GLY	4.8
1	A	471	LYS	4.6
1	B	502	THR	4.6
1	A	793	SER	4.6
1	A	469	LEU	4.6
1	B	615	LEU	4.6
1	B	520	VAL	4.5
1	B	526	MET	4.5
1	B	647	ASP	4.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	469	LEU	4.4
1	B	555	ILE	4.4
1	B	543	LEU	4.4
1	B	474	GLY	4.2
1	B	481	LEU	4.2
2	D	447	LEU	4.2
1	B	754	ARG	4.1
1	A	752	PRO	3.9
1	B	750	VAL	3.9
1	B	753	GLN	3.9
1	B	516	SER	3.9
1	B	521	ASN	3.9
1	B	475	LYS	3.8
1	B	486	TYR	3.8
1	B	549	LEU	3.7
1	A	786	VAL	3.7
1	A	481	LEU	3.6
1	B	522	PRO	3.6
1	B	645	LYS	3.6
1	B	600	SER	3.6
1	B	513	ILE	3.5
1	A	555	ILE	3.5
1	B	562	LEU	3.5
1	B	525	SER	3.4
1	B	471	LYS	3.4
1	A	690	ILE	3.4
1	B	515	ASP	3.3
1	B	723	ILE	3.3
1	A	773	TYR	3.3
1	A	750	VAL	3.3
1	B	643	LEU	3.3
1	A	479	LEU	3.3
1	A	640	ARG	3.2
1	B	470	ASN	3.1
1	A	475	LYS	3.1
1	A	619	ASN	3.0
1	B	531	LYS	3.0
1	A	751	LEU	2.9
1	B	499	GLY	2.9
1	B	528	SER	2.9
1	B	518	TYR	2.9
2	D	440	PRO	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	787	TYR	2.8
1	B	646	GLN	2.8
1	A	765	ILE	2.8
1	B	605	PHE	2.8
1	A	474	GLY	2.7
1	B	624	LEU	2.7
1	B	634	ILE	2.7
1	B	492	LEU	2.6
1	B	611	LEU	2.6
1	A	519	ASP	2.6
1	B	567	LEU	2.6
1	B	523	PHE	2.6
1	A	526	MET	2.6
1	A	605	PHE	2.5
1	A	698	PHE	2.5
1	A	743	LYS	2.5
1	A	603	SER	2.5
1	B	636	ILE	2.5
1	B	776	LEU	2.5
1	A	761	GLU	2.5
1	B	480	GLU	2.5
1	B	512	HIS	2.4
1	A	595	ILE	2.4
1	B	511	LEU	2.4
1	B	591	ILE	2.4
1	B	745	GLU	2.4
1	A	753	GLN	2.4
1	B	483	GLU	2.4
1	B	640	ARG	2.3
1	A	551	ILE	2.3
1	B	478	ILE	2.3
1	B	584	VAL	2.3
1	B	485	VAL	2.3
1	B	509	GLU	2.3
1	B	665	ILE	2.3
1	A	776	LEU	2.2
1	A	736	ASN	2.2
1	B	533	ILE	2.2
1	B	669	LEU	2.2
1	B	468	ASN	2.2
1	B	477	PHE	2.2
1	A	468	ASN	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	486	TYR	2.1
1	A	516	SER	2.1
1	B	629	ILE	2.1
2	C	439	SER	2.1
1	A	794	GLU	2.1
1	B	707	PHE	2.1
1	A	695	ASN	2.0
1	A	508	TRP	2.0
1	B	559	LEU	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 carbohydrates 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.