W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 04:27 PM GMT

PDB ID	:	4F01
Title	:	Crystal structure of an artificial dimeric DnaK complex
Authors	:	Zahn, M.; Straeter, N.
Deposited on	:	2012-05-03
Resolution	:	1.40 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

http://wwpdb.org/validation/2016/XrayValidationReportHelp

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see references ⓘ) were used in the production of this report:

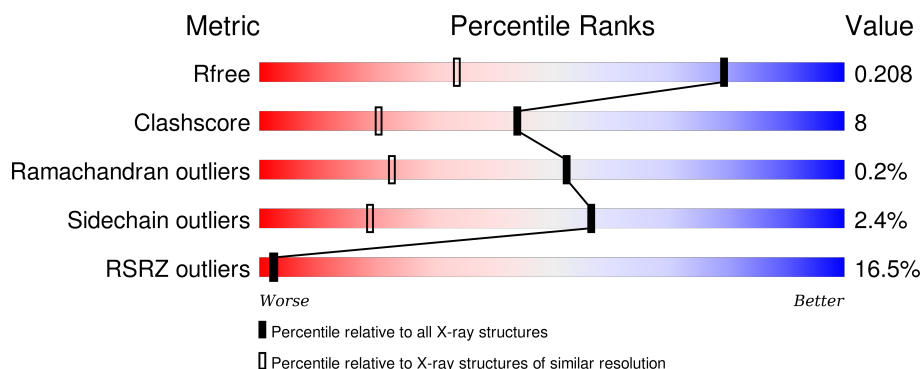|		
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

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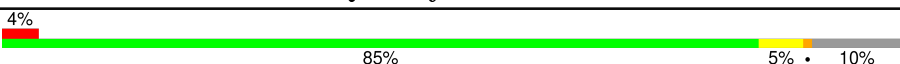*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.40 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		0.208
Clashscore		8
Ramachandran outliers		0.2%
Sidechain outliers		2.4%
RSRZ outliers		16.5%

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}	91344	1199 (1.40-1.40)
Clashscore	102246	1295 (1.40-1.40)
Ramachandran outliers	100387	1259 (1.40-1.40)
Sidechain outliers	100360	1258 (1.40-1.40)
RSRZ outliers	91569	1198 (1.40-1.40)

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. 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	241	4% / 85% / 5% · 10%
1	B	241	27% / 76% / 16% · 5%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3846 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 Chaperone protein DnaK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	218	Total	C	N	O	S	0	8	0
			1690	1041	294	346	9			
1	B	230	Total	C	N	O	S	0	3	0
			1782	1094	323	356	9			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	367	MET	-	EXPRESSION TAG	UNP P0A6Y8
A	368	GLY	-	EXPRESSION TAG	UNP P0A6Y8
A	369	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	370	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	371	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	372	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	373	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	374	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	375	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	376	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	377	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	378	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	379	SER	-	EXPRESSION TAG	UNP P0A6Y8
A	380	SER	-	EXPRESSION TAG	UNP P0A6Y8
A	381	GLY	-	EXPRESSION TAG	UNP P0A6Y8
A	382	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	383	ILE	-	EXPRESSION TAG	UNP P0A6Y8
A	384	GLU	-	EXPRESSION TAG	UNP P0A6Y8
A	385	GLY	-	EXPRESSION TAG	UNP P0A6Y8
A	386	ARG	-	EXPRESSION TAG	UNP P0A6Y8
A	387	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	388	MET	-	EXPRESSION TAG	UNP P0A6Y8
B	367	MET	-	EXPRESSION TAG	UNP P0A6Y8
B	368	GLY	-	EXPRESSION TAG	UNP P0A6Y8
B	369	HIS	-	EXPRESSION TAG	UNP P0A6Y8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	370	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	371	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	372	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	373	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	374	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	375	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	376	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	377	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	378	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	379	SER	-	EXPRESSION TAG	UNP P0A6Y8
B	380	SER	-	EXPRESSION TAG	UNP P0A6Y8
B	381	GLY	-	EXPRESSION TAG	UNP P0A6Y8
B	382	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	383	ILE	-	EXPRESSION TAG	UNP P0A6Y8
B	384	GLU	-	EXPRESSION TAG	UNP P0A6Y8
B	385	GLY	-	EXPRESSION TAG	UNP P0A6Y8
B	386	ARG	-	EXPRESSION TAG	UNP P0A6Y8
B	387	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	388	MET	-	EXPRESSION TAG	UNP P0A6Y8

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	255	Total O 255 255	0	0
2	B	119	Total O 119 119	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Chaperone protein DnaK

- Molecule 1: Chaperone protein DnaK

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	60.11Å 61.38Å 69.19Å 90.00° 103.39° 90.00°	Depositor
Resolution (Å)	25.00 – 1.40 25.14 – 1.39	Depositor EDS
% Data completeness (in resolution range)	98.5 (25.00-1.40) 98.5 (25.14-1.39)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.79 (at 1.39Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.162 , 0.213 0.160 , 0.208	Depositor DCC
R_{free} test set	1906 reflections (2.05%)	DCC
Wilson B-factor (Å^2)	15.6	Xtriage
Anisotropy	0.241	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.42 , 56.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$<\lvert L\rvert> = 0.49$, $<L^2> = 0.33$	Xtriage
Outliers	0 of 97035 reflections	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	3846	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.44% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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	0.99	2/1729 (0.1%)	1.02	6/2326 (0.3%)
1	B	1.00	4/1812 (0.2%)	0.97	5/2438 (0.2%)
All	All	0.99	6/3541 (0.2%)	0.99	11/4764 (0.2%)

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	545	SER	CB-OG	6.52	1.50	1.42
1	A	600	GLU	CD-OE1	6.05	1.32	1.25
1	B	384	GLU	CD-OE2	-5.76	1.19	1.25
1	A	540	ASP	CB-CG	5.39	1.63	1.51
1	B	520	GLU	CD-OE1	5.36	1.31	1.25
1	B	560	ASP	CB-CG	5.15	1.62	1.51

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	404[A]	MET	CG-SD-CE	-7.93	87.50	100.20
1	A	404[B]	MET	CG-SD-CE	-7.93	87.50	100.20
1	A	540	ASP	CB-CG-OD1	7.81	125.33	118.30
1	B	517	ARG	NE-CZ-NH1	7.50	124.05	120.30
1	B	391	LEU	CB-CG-CD2	6.39	121.86	111.00
1	B	517	ARG	NE-CZ-NH2	-6.17	117.21	120.30
1	B	386	ARG	NE-CZ-NH2	-5.95	117.32	120.30
1	B	507	LEU	CB-CG-CD2	-5.65	101.39	111.00
1	A	536	ARG	NE-CZ-NH1	-5.61	117.49	120.30
1	A	560	ASP	CB-CG-OD1	5.39	123.16	118.30
1	A	534	GLN	CA-CB-CG	5.12	124.67	113.40

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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	1690	0	1722	20	0
1	B	1782	0	1782	43	0
2	A	255	0	0	4	0
2	B	119	0	0	0	0
All	All	3846	0	3504	56	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 8.

All (56) 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:404[B]:MET:HA	1:A:404[B]:MET:CE	1.63	1.29
1:A:404[B]:MET:CA	1:A:404[B]:MET:HE2	1.69	1.13
1:B:483:ILE:HD12	1:B:502:LYS:HG2	1.32	1.12
1:B:560:ASP:O	1:B:563:THR:HG22	1.67	0.94
1:B:558:PRO:HD2	1:B:597:LYS:HE3	1.50	0.92
1:A:404[B]:MET:HE3	1:B:387:HIS:HB3	1.53	0.91
1:B:543:LEU:HD21	1:B:573:GLU:HG2	1.53	0.90
1:B:483:ILE:CD1	1:B:502:LYS:HG2	2.01	0.90
1:A:404[B]:MET:CE	1:B:387:HIS:HB3	2.06	0.85
1:A:404[B]:MET:HE1	1:B:387:HIS:CG	2.11	0.83
1:B:600:GLU:O	1:B:603:GLN:HG2	1.78	0.83
1:B:483:ILE:HD12	1:B:502:LYS:CG	2.07	0.83
1:A:404[B]:MET:HA	1:A:404[B]:MET:HE2	0.87	0.81
1:A:404[B]:MET:HE1	1:B:387:HIS:ND1	1.97	0.79
1:B:558:PRO:CD	1:B:597:LYS:HE3	2.16	0.75
1:A:404[B]:MET:CA	1:A:404[B]:MET:CE	2.35	0.74
1:B:546:THR:CG2	1:B:569:LEU:HD21	2.24	0.67
1:B:595:SER:O	1:B:599:MET:HG3	1.94	0.66
1:B:543:LEU:CD2	1:B:573:GLU:HG2	2.24	0.65
1:B:404[B]:MET:CE	1:B:405:GLY:H	2.09	0.64
1:B:553:ALA:HB1	1:B:601:ILE:CG2	2.29	0.63
1:B:404[B]:MET:HE2	1:B:405:GLY:H	1.65	0.61
1:A:520[A]:GLU:HG3	2:A:786:HOH:O	2.00	0.60

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:404[B]:MET:CE	1:B:387:HIS:CB	2.79	0.59
1:B:554:GLY:O	1:B:556:LYS:N	2.37	0.58
1:B:560:ASP:HA	1:B:563:THR:HB	1.85	0.57
1:B:546:THR:HG21	1:B:569:LEU:HD21	1.85	0.57
1:B:553:ALA:HB1	1:B:601:ILE:HG21	1.86	0.57
1:B:404[B]:MET:HE2	1:B:405:GLY:N	2.19	0.56
1:B:572:LEU:HD12	1:B:572:LEU:O	2.07	0.54
1:A:404[B]:MET:CE	1:B:387:HIS:CG	2.89	0.54
1:B:557:LEU:HD11	1:B:601:ILE:HG13	1.90	0.54
1:B:564:ALA:O	1:B:594:VAL:HG11	2.10	0.52
1:B:557:LEU:CD1	1:B:601:ILE:HG13	2.39	0.52
1:B:554:GLY:O	1:B:555:ASP:C	2.48	0.51
1:A:534:GLN:HG3	2:A:872:HOH:O	2.10	0.49
1:B:553:ALA:HB1	1:B:601:ILE:HG22	1.95	0.49
1:B:383:ILE:HG23	1:B:384:GLU:N	2.28	0.48
1:A:404[B]:MET:CA	1:A:404[B]:MET:HE3	2.35	0.47
1:B:548:LYS:O	1:B:552:GLU:HB3	2.15	0.47
1:A:404[B]:MET:HE1	1:B:387:HIS:CB	2.45	0.47
1:B:404[B]:MET:HE3	1:B:405:GLY:H	1.80	0.45
1:A:593:GLN:O	1:A:596:GLN:HB2	2.16	0.44
1:B:483:ILE:HD13	1:B:483:ILE:N	2.31	0.44
1:A:534:GLN:HG2	2:A:861:HOH:O	2.17	0.44
1:B:546:THR:HG21	1:B:591:LEU:HD11	2.01	0.43
1:B:404[B]:MET:CE	1:B:405:GLY:N	2.78	0.43
1:A:539:GLY:HA3	1:A:576:LEU:HD21	2.00	0.43
1:A:562:LYS:HE3	2:A:926:HOH:O	2.19	0.42
1:B:561:ASP:OD1	1:B:565:ILE:CD1	2.68	0.42
1:B:558:PRO:HB2	1:B:561:ASP:HB2	2.01	0.41
1:B:575:ALA:HB1	1:B:584:ILE:HA	2.03	0.41
1:B:546:THR:HG22	1:B:569:LEU:HD11	2.03	0.41
1:A:404[A]:MET:HB3	1:A:404[A]:MET:HE2	1.92	0.40
1:A:534:GLN:O	1:A:538:GLN:HG3	2.21	0.40
1:B:590:GLU:O	1:B:593:GLN:HB2	2.21	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	224/241 (93%)	220 (98%)	4 (2%)	0	100	100
1	B	229/241 (95%)	222 (97%)	6 (3%)	1 (0%)	39	14
All	All	453/482 (94%)	442 (98%)	10 (2%)	1 (0%)	52	22

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	555	ASP

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	188/200 (94%)	187 (100%)	1 (0%)	92	78
1	B	195/200 (98%)	187 (96%)	8 (4%)	37	6
All	All	383/400 (96%)	374 (98%)	9 (2%)	57	21

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	581	LYS
1	B	382	HIS
1	B	498	LYS
1	B	545	SER
1	B	552	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	561	ASP
1	B	563	THR
1	B	574	THR
1	B	594	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	424	GLN

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 carbohydrates 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 (i)

6.1 Protein, DNA and RNA chains (i)

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(Å²)	Q<0.9
1	A	218/241 (90%)	0.32	9 (4%)	41	39	9, 18, 43, 73	0
1	B	230/241 (95%)	1.41	65 (28%)	1	1	12, 31, 94, 144	0
All	All	448/482 (92%)	0.88	74 (16%)	2	2	9, 23, 77, 144	0

All (74) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	383	ILE	11.7
1	B	563	THR	10.1
1	B	506	GLY	8.4
1	B	559	ALA	8.3
1	B	553	ALA	7.5
1	B	561	ASP	7.4
1	B	574	THR	6.9
1	B	551	GLU	6.8
1	B	555	ASP	6.6
1	B	564	ALA	6.3
1	B	557	LEU	5.9
1	B	575	ALA	5.8
1	B	550	VAL	5.7
1	B	602	ALA	5.5
1	B	560	ASP	5.3
1	B	566	GLU	5.2
1	B	556	LYS	5.2
1	B	577	LYS	5.2
1	B	552	GLU	5.0
1	B	570	THR	4.8
1	B	562	LYS	4.8
1	B	508	ASN	4.7
1	B	583	ALA	4.6
1	B	507	LEU	4.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	568	ALA	4.6
1	B	382	HIS	4.6
1	B	505	SER	4.5
1	B	548	LYS	4.5
1	B	558	PRO	4.3
1	B	601	ILE	4.2
1	B	565	ILE	4.0
1	B	554	GLY	3.9
1	B	586	ALA	3.9
1	B	545	SER	3.9
1	A	582	ALA	3.8
1	B	464	PRO	3.7
1	B	401	ILE	3.6
1	B	589	GLN	3.5
1	B	510	ASP	3.4
1	B	567	SER	3.4
1	B	373	HIS	3.3
1	B	476	PHE	3.3
1	B	582	ALA	3.3
1	B	603	GLN	3.3
1	B	549	GLN	3.2
1	B	571	ALA	3.2
1	B	569	LEU	3.2
1	B	509	GLU	2.8
1	A	586	ALA	2.8
1	B	465	ALA	2.8
1	A	385	GLY	2.8
1	B	399	LEU	2.7
1	B	572	LEU	2.6
1	A	438	ILE	2.6
1	B	411	LEU	2.6
1	B	511	GLU	2.5
1	A	401	ILE	2.5
1	B	584	ILE	2.5
1	B	440	VAL	2.5
1	A	575	ALA	2.5
1	B	484	LEU	2.4
1	B	466	PRO	2.4
1	B	604	GLN	2.4
1	B	593	GLN	2.3
1	B	400	GLY	2.3
1	B	438	ILE	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	474	VAL	2.2
1	A	567[A]	SER	2.2
1	B	594	VAL	2.1
1	B	587	LYS	2.1
1	B	573	GLU	2.1
1	A	399	LEU	2.1
1	B	426	PHE	2.0
1	A	440	VAL	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.