Full wwPDB X-ray Structure Validation Report ⓘ

Nov 20, 2023 – 07:24 PM JST

PDB ID	:	7DA2
Title	:	The crystal structure of the chicken FANCM-MHF complex
Authors	:	Nishino, T.; Ito, S.
Deposited on	:	2020-10-14
Resolution	:	2.79 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:

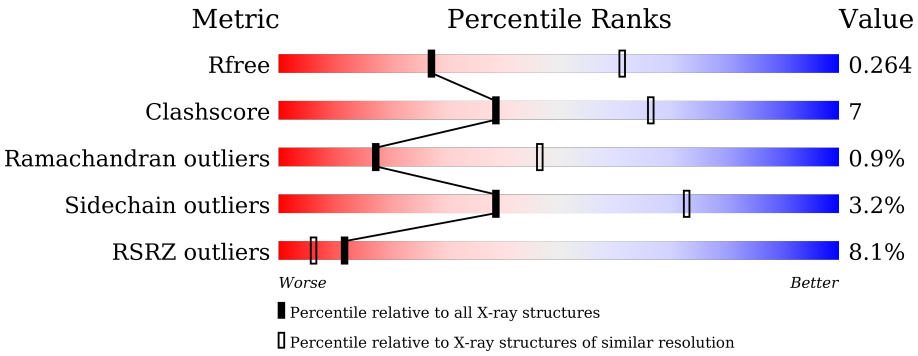
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.79 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	107	9% 69% 21% 10%
1	C	107	3% 66% 20% 14%
2	B	81	% 83% 10% 7%
2	D	81	2% 78% 14% 9%
3	E	148	15% 62% 24% · 11%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3858 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 Centromere protein S.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	96	Total	C	N	O	S	0	0	0
			758	470	141	146	1			
1	C	92	Total	C	N	O	S	0	0	0
			735	459	136	139	1			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	3	GLY	-	expression tag	UNP E1BSW7
A	4	SER	-	expression tag	UNP E1BSW7
A	29	ALA	CYS	engineered mutation	UNP E1BSW7
A	31	ALA	CYS	engineered mutation	UNP E1BSW7
A	58	ALA	CYS	engineered mutation	UNP E1BSW7
C	3	GLY	-	expression tag	UNP E1BSW7
C	4	SER	-	expression tag	UNP E1BSW7
C	29	ALA	CYS	engineered mutation	UNP E1BSW7
C	31	ALA	CYS	engineered mutation	UNP E1BSW7
C	58	ALA	CYS	engineered mutation	UNP E1BSW7

- Molecule 2 is a protein called Centromere protein X.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	75	Total	C	N	O	S	0	0	0
			604	382	114	107	1			
2	D	74	Total	C	N	O	S	0	0	0
			600	380	113	106	1			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	2	GLY	-	expression tag	UNP P0DJH7
B	3	TYR	-	expression tag	UNP P0DJH7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	2	GLY	-	expression tag	UNP P0DJH7
D	3	TYR	-	expression tag	UNP P0DJH7

- Molecule 3 is a protein called Fanconi anemia group M protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	132	Total	C	N	O	S	0	0	0
			1114	698	196	213	7			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	658	GLY	-	expression tag	UNP A0A1D5PRR9
E	659	ARG	-	expression tag	UNP A0A1D5PRR9
E	805	PRO	-	expression tag	UNP A0A1D5PRR9

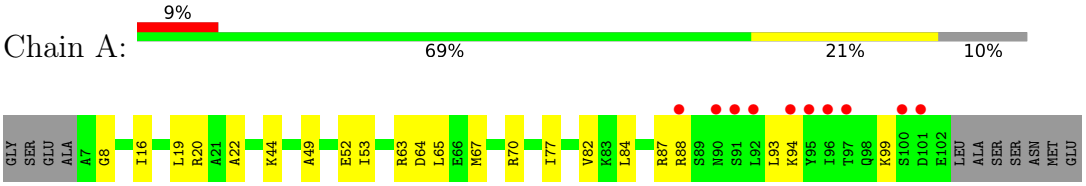
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	12	Total	O	0	0
			12	12		
4	C	7	Total	O	0	0
			7	7		
4	B	5	Total	O	0	0
			5	5		
4	D	3	Total	O	0	0
			3	3		
4	E	20	Total	O	0	0
			20	20		

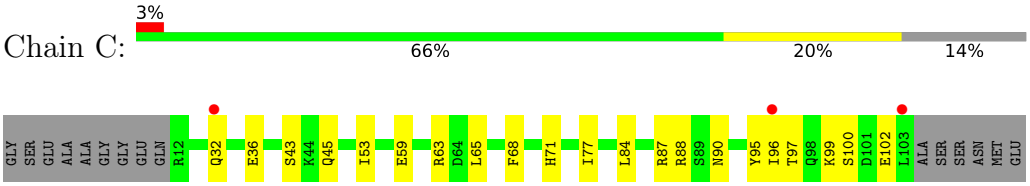
3 Residue-property plots

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.

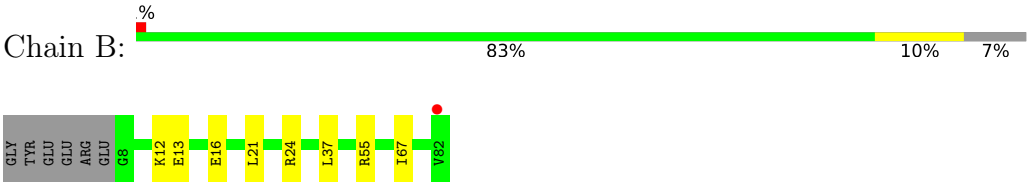
• Molecule 1: Centromere protein S



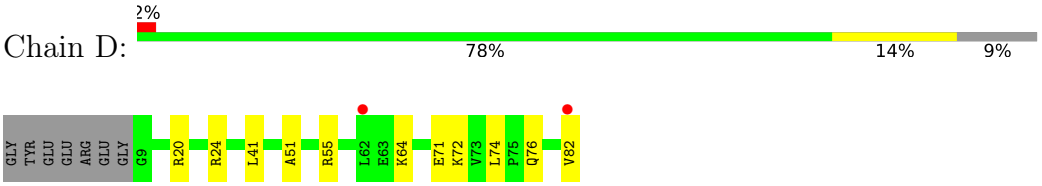
• Molecule 1: Centromere protein S



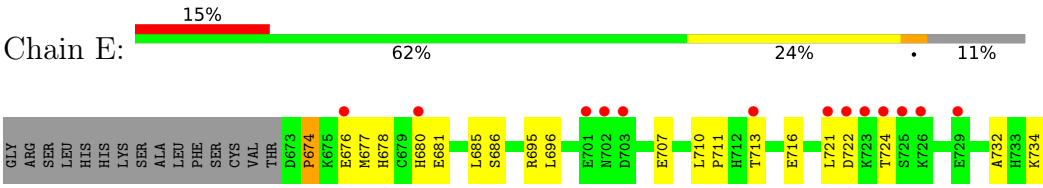
• Molecule 2: Centromere protein X

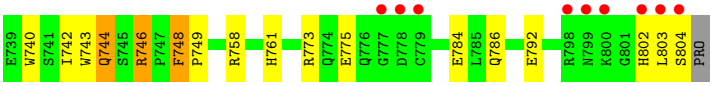


• Molecule 2: Centromere protein X



• Molecule 3: Fanconi anemia group M protein




E739
W740
S741
I742
W743
Q744
S745
R746
P747
F748
P749
R758
H761
R773
Q774
E775
Q776
G777
D778
C779
E784
L785
Q786
E792
R798
N799
K800
G801
H802
L803
S804
PRO

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	71.25Å 78.47Å 87.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.17 – 2.79 45.17 – 2.79	Depositor EDS
% Data completeness (in resolution range)	99.3 (45.17-2.79) 99.3 (45.17-2.79)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.90 (at 2.77Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.216 , 0.262 0.216 , 0.264	Depositor DCC
R_{free} test set	633 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	54.9	Xtriage
Anisotropy	0.425	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 56.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	3858	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.86% 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	0.26	0/765	0.41	0/1026
1	C	0.24	0/742	0.39	0/996
2	B	0.25	0/610	0.42	0/818
2	D	0.25	0/606	0.42	0/813
3	E	0.24	0/1145	0.40	0/1548
All	All	0.25	0/3868	0.41	0/5201

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	758	0	770	15	0
1	C	735	0	756	24	0
2	B	604	0	629	9	0
2	D	600	0	626	12	0
3	E	1114	0	1054	28	0
4	A	12	0	0	0	0
4	B	5	0	0	1	0
4	C	7	0	0	0	0
4	D	3	0	0	0	0
4	E	20	0	0	3	0
All	All	3858	0	3835	57	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 7.

All (57) 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:64:LYS:NZ	3:E:804:SER:OG	2.20	0.75
1:C:88:ARG:HD2	2:D:82:VAL:HG21	1.73	0.70
1:A:87:ARG:HH22	1:A:88:ARG:HH21	1.39	0.68
3:E:680:HIS:HB3	3:E:686:SER:HB3	1.75	0.68
1:A:67:MET:HG3	1:C:87:ARG:HB2	1.75	0.67
1:C:90:ASN:OD1	2:B:24:ARG:NH1	2.28	0.67
3:E:681:GLU:OE2	4:E:901:HOH:O	2.13	0.66
2:D:71:GLU:OE2	3:E:773:ARG:NH2	2.32	0.62
1:C:90:ASN:HB3	3:E:721:LEU:HD12	1.81	0.62
2:D:55:ARG:HD3	3:E:716:GLU:OE2	2.00	0.61
1:C:32:GLN:HE21	3:E:803:LEU:HG	1.65	0.61
2:B:55:ARG:NH2	4:B:101:HOH:O	2.33	0.60
1:A:44:LYS:NZ	3:E:696:LEU:O	2.34	0.60
1:C:63:ARG:HD3	2:D:24:ARG:HH21	1.66	0.59
1:A:20:ARG:NH1	1:A:52:GLU:OE1	2.36	0.59
3:E:746:ARG:NH1	4:E:905:HOH:O	2.36	0.58
1:C:96:ILE:HD11	2:D:41:LEU:HD13	1.85	0.58
2:D:82:VAL:O	3:E:758:ARG:NH2	2.37	0.57
1:C:96:ILE:O	1:C:100:SER:N	2.38	0.55
2:B:67:ILE:HD11	3:E:735:LEU:HD21	1.88	0.55
1:A:53:ILE:HD13	3:E:710:LEU:HD12	1.89	0.54
1:C:59:GLU:OE1	3:E:761:HIS:NE2	2.39	0.54
1:C:95:TYR:CD2	1:C:96:ILE:HG13	2.43	0.54
1:C:32:GLN:NE2	3:E:803:LEU:HG	2.24	0.53
3:E:676:GLU:HG3	3:E:677:MET:H	1.73	0.52
1:A:87:ARG:HA	1:A:93:LEU:HD13	1.93	0.51
2:B:12:LYS:HD2	2:B:12:LYS:H	1.76	0.50
3:E:740:TRP:O	3:E:744:GLN:NE2	2.44	0.50
1:C:36:GLU:OE2	3:E:802:HIS:NE2	2.45	0.50
1:A:63:ARG:NH1	1:C:88:ARG:O	2.44	0.50
1:C:87:ARG:O	2:B:24:ARG:NH2	2.45	0.50
3:E:784:GLU:OE2	4:E:902:HOH:O	2.19	0.50
1:A:19:LEU:HD12	2:B:21:LEU:HD22	1.94	0.49
1:A:16:ILE:HG12	2:B:21:LEU:HD21	1.94	0.49
1:A:65:LEU:HB3	1:A:77:ILE:HD13	1.93	0.49
1:C:63:ARG:HD3	2:D:24:ARG:NH2	2.27	0.49
1:C:65:LEU:HB3	1:C:77:ILE:HD13	1.97	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:13:GLU:O	2:B:16:GLU:HG3	2.14	0.47
3:E:674:PRO:O	3:E:678:HIS:HB2	2.15	0.47
3:E:695:ARG:HG3	3:E:696:LEU:HG	1.96	0.47
1:C:53:ILE:HD13	2:D:74:LEU:HD13	1.97	0.46
1:C:32:GLN:OE1	3:E:802:HIS:HB2	2.16	0.46
1:A:84:LEU:HD22	1:C:71:HIS:CG	2.50	0.46
3:E:707:GLU:O	3:E:732:ALA:HB1	2.15	0.46
1:A:49:ALA:O	1:A:53:ILE:HG12	2.16	0.46
1:C:68:PHE:CG	1:C:84:LEU:HD23	2.51	0.45
1:C:99:LYS:HA	1:C:102:GLU:HG2	2.00	0.44
1:C:95:TYR:O	1:C:97:THR:N	2.46	0.43
1:A:82:VAL:HG11	2:B:37:LEU:HB3	2.00	0.43
3:E:711:PRO:HB3	3:E:734:LYS:HD2	2.00	0.42
2:D:76:GLN:HG2	3:E:737:LEU:HD12	2.00	0.42
3:E:742:ILE:HG13	3:E:743:TRP:CD1	2.54	0.42
1:A:64:ASP:OD2	1:C:87:ARG:NE	2.54	0.41
2:D:51:ALA:HB1	3:E:716:GLU:HG2	2.03	0.41
1:C:43:SER:OG	1:C:45:GLN:HG3	2.21	0.40
2:D:72:LYS:HD2	3:E:738:SER:OG	2.21	0.40
1:A:22:ALA:HA	3:E:685:LEU:HD21	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	94/107 (88%)	90 (96%)	3 (3%)	1 (1%)	14 41
1	C	90/107 (84%)	88 (98%)	2 (2%)	0	100 100
2	B	73/81 (90%)	73 (100%)	0	0	100 100
2	D	72/81 (89%)	72 (100%)	0	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	E	130/148 (88%)	112 (86%)	15 (12%)	3 (2%)	6 21
All	All	459/524 (88%)	435 (95%)	20 (4%)	4 (1%)	17 46

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	E	674	PRO
3	E	749	PRO
3	E	748	PHE
1	A	8	GLY

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	79/87 (91%)	76 (96%)	3 (4%)	33 67
1	C	78/87 (90%)	78 (100%)	0	100 100
2	B	62/67 (92%)	62 (100%)	0	100 100
2	D	62/67 (92%)	61 (98%)	1 (2%)	62 88
3	E	127/141 (90%)	118 (93%)	9 (7%)	14 39
All	All	408/449 (91%)	395 (97%)	13 (3%)	39 73

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	70	ARG
1	A	94	LYS
1	A	99	LYS
2	D	20	ARG
3	E	713	THR
3	E	722	ASP
3	E	724	THR
3	E	744	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	E	746	ARG
3	E	748	PHE
3	E	775	GLU
3	E	786	GLN
3	E	792	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	17	GLN
3	E	795	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 monosaccharides 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	96/107 (89%)	0.52	10 (10%) 6 3	33, 53, 155, 176	0
1	C	92/107 (85%)	0.07	3 (3%) 46 36	32, 55, 96, 100	0
2	B	75/81 (92%)	0.00	1 (1%) 77 72	33, 49, 77, 83	0
2	D	74/81 (91%)	0.20	2 (2%) 54 44	34, 57, 86, 98	0
3	E	132/148 (89%)	0.85	22 (16%) 1 1	35, 69, 141, 165	0
All	All	469/524 (89%)	0.39	38 (8%) 12 6	32, 58, 130, 176	0

All (38) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	100	SER	10.4
3	E	802	HIS	9.6
1	A	97	THR	7.8
3	E	799	ASN	6.5
3	E	803	LEU	6.5
1	A	101	ASP	6.2
3	E	729	GLU	5.9
3	E	725	SER	5.6
1	A	92	LEU	4.6
3	E	800	LYS	4.4
3	E	723	LYS	4.3
3	E	703	ASP	4.3
3	E	777	GLY	4.3
1	A	91	SER	3.7
3	E	702	ASN	3.7
1	A	90	ASN	3.6
1	A	88	ARG	3.5
2	B	82	VAL	3.3
3	E	724	THR	3.2
1	A	96	ILE	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	32	GLN	2.8
3	E	701	GLU	2.8
3	E	713	THR	2.7
3	E	804	SER	2.7
3	E	680	HIS	2.6
3	E	726	LYS	2.5
3	E	798	ARG	2.5
2	D	62	LEU	2.5
3	E	779	CYS	2.3
1	C	103	LEU	2.3
3	E	722	ASP	2.2
3	E	676	GLU	2.2
1	C	96	ILE	2.2
1	A	95	TYR	2.2
2	D	82	VAL	2.1
3	E	778	ASP	2.1
1	A	94	LYS	2.1
3	E	721	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 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.