Full wwPDB X-ray Structure Validation Report ⓘ

Nov 22, 2023 – 01:38 PM JST

PDB ID	:	7X5F
Title	:	Nrf2-MafG heterodimer bound with CsMBE2
Authors	:	Sengoku, T.; Shiina, M.; Suzuki, K.; Hamada, K.; Sato, K.; Uchiyama, A.; Okada, C.; Baba, S.; Ohta, T.; Motohashi, H.; Yamamoto, M.; Ogata, K.
Deposited on	:	2022-03-04
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 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:

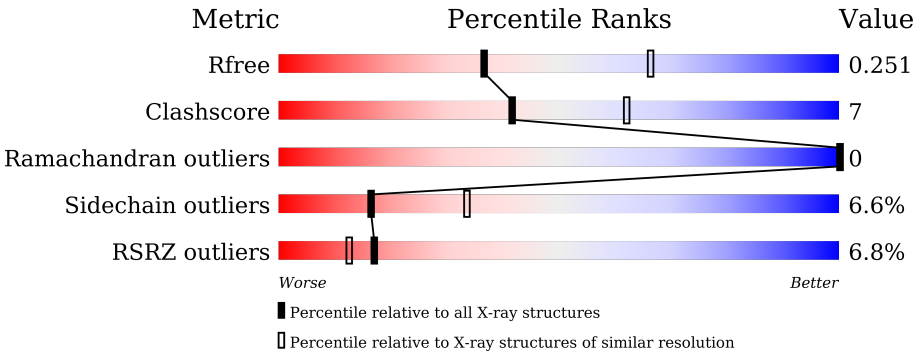
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.60 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 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	104	12% 67% 24% 5% ·
1	E	104	11% 80% 7% 5% 9%
2	B	113	78% 15% 7%
2	F	113	8% 65% 27% · 6%
3	C	16	81% 19%
3	G	16	94% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	16	81% 19%
4	H	16	81% 19%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 4555 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 Transcription factor MafG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	100	Total	C	N	O	S	0	0	0
			804	493	151	157	3			
1	E	95	Total	C	N	O	S	0	0	0
			738	450	139	146	3			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	MET	-	initiating methionine	UNP O15525
E	20	MET	-	initiating methionine	UNP O15525

- Molecule 2 is a protein called Nuclear factor erythroid 2-related factor 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	105	Total	C	N	O	S	0	0	0
			864	541	161	159	3			
2	F	106	Total	C	N	O	S	0	0	0
			842	523	160	156	3			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	448	GLY	-	expression tag	UNP Q16236
B	449	PRO	-	expression tag	UNP Q16236
B	450	HIS	-	expression tag	UNP Q16236
B	451	MET	-	expression tag	UNP Q16236
F	448	GLY	-	expression tag	UNP Q16236
F	449	PRO	-	expression tag	UNP Q16236
F	450	HIS	-	expression tag	UNP Q16236
F	451	MET	-	expression tag	UNP Q16236

- Molecule 3 is a DNA chain called Synthetic DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	16	Total	C	N	O	P	0	0	0
			325	156	57	97	15			
3	G	16	Total	C	N	O	P	0	0	0
			325	156	57	97	15			

- Molecule 4 is a DNA chain called Synthetic DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	16	Total	C	N	O	P	0	0	0
			325	155	64	91	15			
4	H	16	Total	C	N	O	P	0	0	0
			325	155	64	91	15			

- Molecule 5 is water.

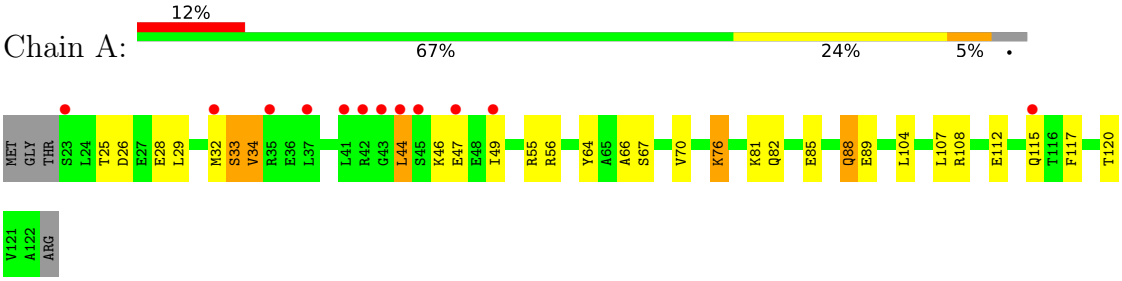
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	O	0	0
			1	1		
5	C	2	Total	O	0	0
			2	2		
5	D	1	Total	O	0	0
			1	1		
5	E	3	Total	O	0	0
			3	3		

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.

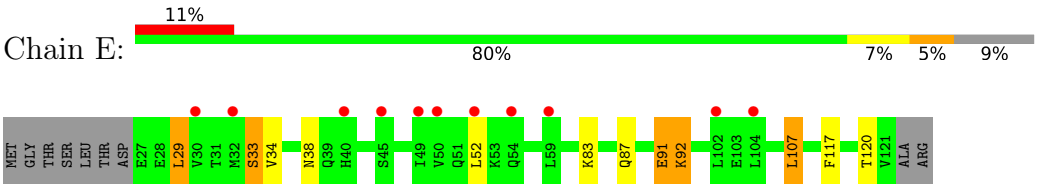
• Molecule 1: Transcription factor MafG

Chain A:



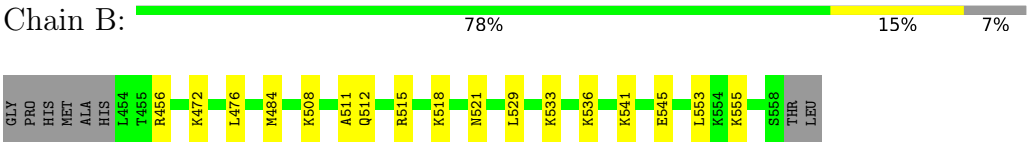
• Molecule 1: Transcription factor MafG

Chain E:



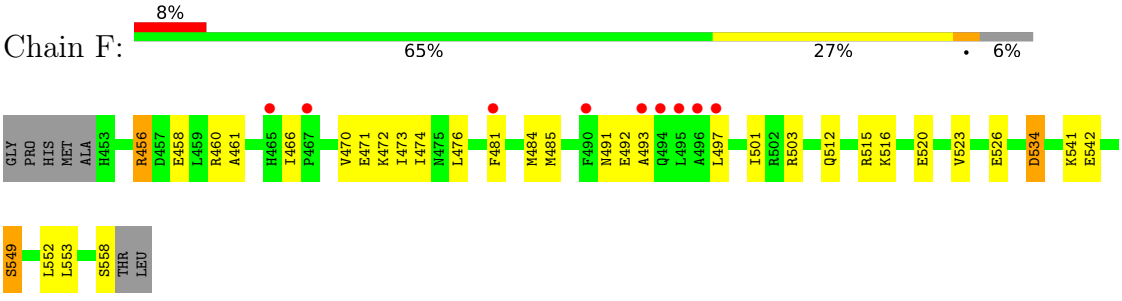
• Molecule 2: Nuclear factor erythroid 2-related factor 2

Chain B:



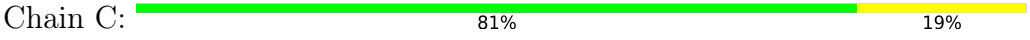
• Molecule 2: Nuclear factor erythroid 2-related factor 2

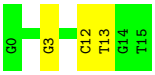
Chain F:



• Molecule 3: Synthetic DNA

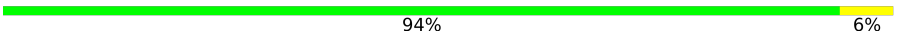
Chain C:





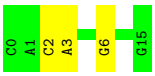
• Molecule 3: Synthetic DNA

Chain G:



• Molecule 4: Synthetic DNA

Chain D:



• Molecule 4: Synthetic DNA

Chain H:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	245.17Å 54.69Å 85.95Å 90.00° 99.18° 90.00°	Depositor
Resolution (Å)	45.27 – 2.60 45.27 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.2 (45.27-2.60) 99.2 (45.27-2.60)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.53 (at 2.61Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.226 , 0.259 0.222 , 0.251	Depositor DCC
R_{free} test set	1747 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	65.7	Xtriage
Anisotropy	0.217	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 52.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4555	wwPDB-VP
Average B, all atoms (Å^2)	89.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.83% 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	0.30	0/807	0.56	0/1077
1	E	0.30	0/741	0.54	0/991
2	B	0.32	0/872	0.52	0/1163
2	F	0.32	0/850	0.56	0/1139
3	C	0.74	0/363	0.94	0/559
3	G	0.70	0/363	0.95	0/559
4	D	0.71	0/365	0.82	0/561
4	H	0.68	0/365	0.79	0/561
All	All	0.47	0/4726	0.68	0/6610

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	804	0	831	20	0
1	E	738	0	726	8	0
2	B	864	0	906	13	0
2	F	842	0	844	21	0
3	C	325	0	183	2	0
3	G	325	0	183	1	0
4	D	325	0	180	2	0
4	H	325	0	180	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	1	0	0	0	0
5	C	2	0	0	0	0
5	D	1	0	0	0	0
5	E	3	0	0	0	0
All	All	4555	0	4033	60	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 (60) 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:112:GLU:HA	1:A:115:GLN:HG2	1.69	0.74
1:A:46:LYS:HA	1:A:49:ILE:HD12	1.74	0.70
1:A:76:LYS:NZ	2:B:521:ASN:OD1	2.26	0.68
2:F:474:ILE:HD11	2:F:501:ILE:HG22	1.79	0.64
2:F:472:LYS:O	2:F:476:LEU:HG	1.99	0.62
2:F:473:ILE:H	2:F:473:ILE:HD12	1.63	0.62
2:B:484:MET:HG3	2:F:534:ASP:OD2	2.01	0.61
1:A:47:GLU:N	1:A:47:GLU:OE1	2.35	0.60
1:A:44:LEU:HD13	1:A:49:ILE:HG13	1.85	0.58
1:E:34:VAL:O	1:E:38:ASN:ND2	2.35	0.58
1:A:82:GLN:HG3	2:B:529:LEU:HD11	1.85	0.57
1:E:107:LEU:HD23	2:F:553:LEU:HD12	1.86	0.57
1:A:107:LEU:HD23	2:B:553:LEU:HD23	1.87	0.56
4:H:13:DC:H2'	4:H:14:DA:C8	2.41	0.56
1:E:83:LYS:O	1:E:87:GLN:HG3	2.05	0.55
1:A:29:LEU:HG	1:A:56:ARG:HG3	1.88	0.55
2:B:529:LEU:O	2:B:533:LYS:HG3	2.06	0.55
1:A:117:PHE:O	1:A:120:THR:HG22	2.06	0.55
1:E:117:PHE:O	1:E:120:THR:HG22	2.08	0.54
1:A:33:SER:OG	1:A:34:VAL:N	2.40	0.53
2:B:511:ALA:O	2:B:515:ARG:HG3	2.09	0.53
1:A:28:GLU:O	1:A:32:MET:HG2	2.10	0.52
2:F:512:GLN:O	2:F:516:LYS:HG3	2.10	0.52
2:F:491:ASN:OD1	2:F:492:GLU:N	2.43	0.52
2:B:472:LYS:O	2:B:476:LEU:HB2	2.10	0.51
2:F:461:ALA:O	2:F:466:ILE:HG12	2.10	0.51
2:B:508:LYS:O	2:B:512:GLN:HG3	2.10	0.51
1:E:33:SER:OG	1:E:34:VAL:N	2.44	0.50
2:F:497:LEU:O	2:F:501:ILE:HG12	2.12	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:91:GLU:HG3	1:E:92:LYS:HE2	1.95	0.48
1:A:85:GLU:O	1:A:88:GLN:HG2	2.14	0.48
2:F:470:VAL:O	2:F:474:ILE:HG12	2.15	0.47
4:D:2:DC:H2"	4:D:3:DA:C8	2.49	0.47
2:F:473:ILE:HG21	2:F:501:ILE:HB	1.97	0.47
2:F:491:ASN:ND2	2:F:493:ALA:HB3	2.30	0.47
1:A:32:MET:O	1:A:56:ARG:NH1	2.48	0.46
2:B:476:LEU:HD12	2:B:476:LEU:HA	1.83	0.46
1:A:25:THR:OG1	1:A:26:ASP:N	2.48	0.46
1:A:104:LEU:HD11	1:A:108:ARG:NH2	2.31	0.46
2:F:456:ARG:NH1	2:F:460:ARG:HG3	2.32	0.45
2:F:520:GLU:O	2:F:523:VAL:HG22	2.16	0.45
2:B:518:LYS:NZ	4:D:6:DG:OP2	2.50	0.45
1:A:89:GLU:OE2	2:B:536:LYS:HD2	2.18	0.44
1:A:64:TYR:OH	3:C:3:DG:OP2	2.28	0.44
2:F:549:SER:O	2:F:553:LEU:HG	2.18	0.44
1:A:29:LEU:HB2	1:A:55:ARG:HH21	1.82	0.44
2:F:460:ARG:NH1	2:F:460:ARG:O	2.50	0.44
3:C:12:DC:H2'	3:C:13:DT:C6	2.53	0.44
2:F:552:LEU:HD23	2:F:552:LEU:HA	1.88	0.43
4:H:14:DA:H2"	4:H:15:DG:C8	2.54	0.43
4:H:14:DA:H2"	4:H:15:DG:H8	1.82	0.43
2:F:481:PHE:O	2:F:485:MET:HG2	2.19	0.43
1:A:66:ALA:O	1:A:70:VAL:HG23	2.19	0.43
1:A:82:GLN:CG	2:B:529:LEU:HD11	2.49	0.42
2:B:541:LYS:O	2:B:545:GLU:HG3	2.19	0.42
2:F:541:LYS:HE2	2:F:542:GLU:N	2.34	0.42
2:F:458:GLU:HA	2:F:470:VAL:HG21	2.02	0.42
2:F:515:ARG:NH2	3:G:8:DG:O6	2.54	0.40
1:E:92:LYS:HD3	1:E:92:LYS:HA	1.93	0.40
1:E:29:LEU:H	1:E:29:LEU:HD23	1.86	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	98/104 (94%)	95 (97%)	3 (3%)	0	100	100
1	E	93/104 (89%)	90 (97%)	3 (3%)	0	100	100
2	B	103/113 (91%)	100 (97%)	3 (3%)	0	100	100
2	F	104/113 (92%)	101 (97%)	3 (3%)	0	100	100
All	All	398/434 (92%)	386 (97%)	12 (3%)	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	87/92 (95%)	80 (92%)	7 (8%)	12	24
1	E	75/92 (82%)	69 (92%)	6 (8%)	12	24
2	B	96/103 (93%)	94 (98%)	2 (2%)	53	77
2	F	88/103 (85%)	80 (91%)	8 (9%)	9	18
All	All	346/390 (89%)	323 (93%)	23 (7%)	16	33

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	33	SER
1	A	34	VAL
1	A	44	LEU
1	A	67	SER
1	A	76	LYS
1	A	81	LYS
1	A	88	GLN
2	B	456	ARG
2	B	555	LYS
1	E	29	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	33	SER
1	E	52	LEU
1	E	91	GLU
1	E	92	LYS
1	E	107	LEU
2	F	456	ARG
2	F	471	GLU
2	F	484	MET
2	F	503	ARG
2	F	526	GLU
2	F	534	ASP
2	F	549	SER
2	F	558	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	88	GLN
2	B	546	ASN
1	E	82	GLN
1	E	88	GLN
2	F	512	GLN
2	F	527	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 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, 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	100/104 (96%)	0.80	12 (12%) 4 2	62, 88, 138, 152	0
1	E	95/104 (91%)	0.79	11 (11%) 4 3	52, 91, 146, 151	0
2	B	105/113 (92%)	0.35	0 100 100	49, 72, 101, 111	0
2	F	106/113 (93%)	0.79	9 (8%) 10 7	55, 104, 155, 165	0
3	C	16/16 (100%)	-0.11	0 100 100	54, 80, 100, 100	0
3	G	16/16 (100%)	0.02	0 100 100	64, 83, 109, 125	0
4	D	16/16 (100%)	-0.09	0 100 100	61, 73, 116, 124	0
4	H	16/16 (100%)	-0.02	0 100 100	63, 82, 109, 118	0
All	All	470/498 (94%)	0.58	32 (6%) 17 12	49, 85, 144, 165	0

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	49	ILE	5.5
1	A	42	ARG	4.7
2	F	496	ALA	4.3
1	E	50	VAL	4.1
1	A	44	LEU	4.0
2	F	493	ALA	3.8
1	E	45	SER	3.8
1	A	35	ARG	3.6
1	E	54	GLN	3.5
2	F	467	PRO	3.2
1	A	43	GLY	3.2
2	F	494	GLN	3.1
2	F	490	PHE	2.9
1	A	37	LEU	2.9
2	F	497	LEU	2.8
2	F	465	HIS	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	59	LEU	2.6
1	A	23	SER	2.6
1	E	32	MET	2.6
1	A	41	LEU	2.5
1	A	32	MET	2.5
1	E	52	LEU	2.4
2	F	495	LEU	2.3
1	A	45	SER	2.3
1	E	30	VAL	2.2
1	A	49	ILE	2.2
1	E	40	HIS	2.2
1	A	47	GLU	2.1
1	A	115	GLN	2.1
1	E	104	LEU	2.1
2	F	481	PHE	2.1
1	E	102	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.