WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Nov 20, 2023 – 02:12 PM JST

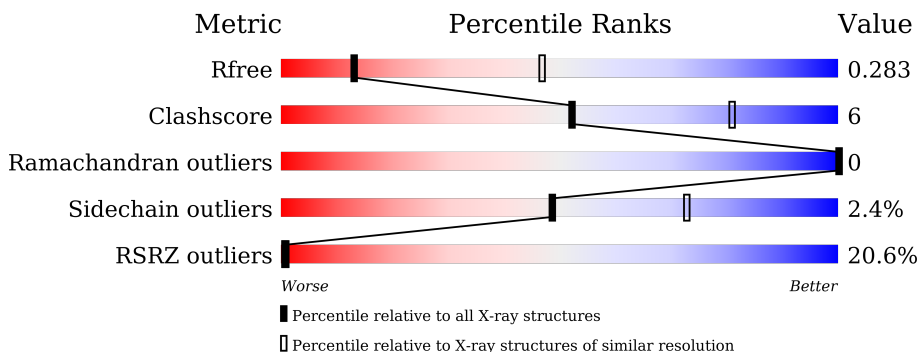|		
PDB ID	:	7CO1
Title	:	Crystal structure of SMAD2 in complex with wild-type CBP
Authors	:	Miyazono, K.; Wada, H.; Ito, T.; Tanokura, M.
Deposited on	:	2020-08-03
Resolution	:	3.30 Å(reported)

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 3.30 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		0.283
Clashscore		6
Ramachandran outliers		0
Sidechain outliers		2.4%
RSRZ outliers		20.6%

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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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	208	13% / 80% 13% 7%
1	C	208	18% / 81% 12% 7%
1	E	208	24% / 76% 16% 8%
2	B	28	11% / 61% 7% • 29%
2	D	28	7% / 61% 11% 29%
2	F	28	39% / 50% 14% 36%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 5054 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 Mothers against decapentaplegic homolog 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	194	Total	C	N	O	S	0	0	0
			1541	977	270	282	12			
1	C	193	Total	C	N	O	S	0	0	0
			1534	972	269	281	12			
1	E	192	Total	C	N	O	S	0	0	0
			1528	969	268	279	12			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	260	GLY	-	expression tag	UNP Q15796
A	261	PRO	-	expression tag	UNP Q15796
A	465	GLU	SER	engineered mutation	UNP Q15796
A	467	GLU	SER	engineered mutation	UNP Q15796
C	260	GLY	-	expression tag	UNP Q15796
C	261	PRO	-	expression tag	UNP Q15796
C	465	GLU	SER	engineered mutation	UNP Q15796
C	467	GLU	SER	engineered mutation	UNP Q15796
E	260	GLY	-	expression tag	UNP Q15796
E	261	PRO	-	expression tag	UNP Q15796
E	465	GLU	SER	engineered mutation	UNP Q15796
E	467	GLU	SER	engineered mutation	UNP Q15796

- Molecule 2 is a protein called CREB-binding protein.

| Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	20	Total	C	N	O	0	0	0
			155	93	31	31			
2	D	20	Total	C	N	O	0	0	0
			155	93	31	31			
2	F	18	Total	C	N	O	0	0	0
			141	83	29	29			

There are 15 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
B	1950	GLY	-	expression tag	UNP Q92793
B	1974	SER	-	expression tag	UNP Q92793
B	1975	ASP	-	expression tag	UNP Q92793
B	1976	GLU	-	expression tag	UNP Q92793
B	1977	ASP	-	expression tag	UNP Q92793
D	1950	GLY	-	expression tag	UNP Q92793
D	1974	SER	-	expression tag	UNP Q92793
D	1975	ASP	-	expression tag	UNP Q92793
D	1976	GLU	-	expression tag	UNP Q92793
D	1977	ASP	-	expression tag	UNP Q92793
F	1950	GLY	-	expression tag	UNP Q92793
F	1974	SER	-	expression tag	UNP Q92793
F	1975	ASP	-	expression tag	UNP Q92793
F	1976	GLU	-	expression tag	UNP Q92793
F	1977	ASP	-	expression tag	UNP Q92793

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.
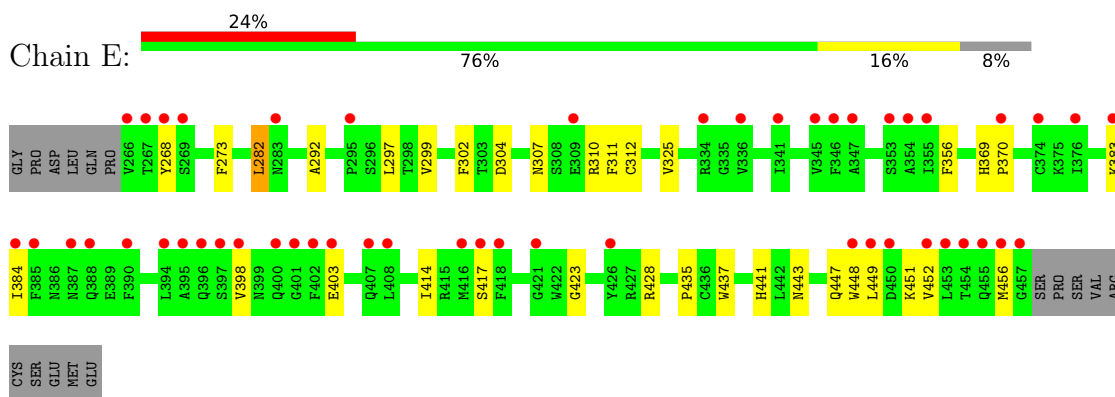
• Molecule 1: Mothers against decapentaplegic homolog 2
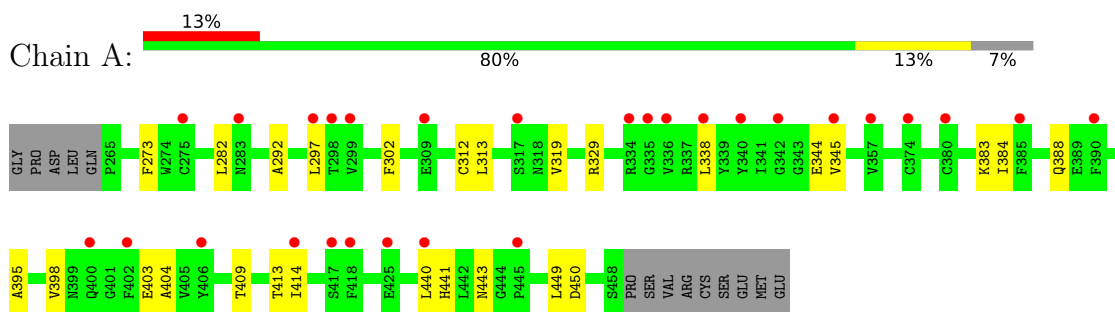
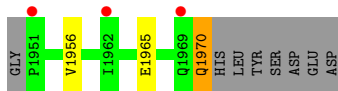• Molecule 1: Mothers against decapentaplegic homolog 2
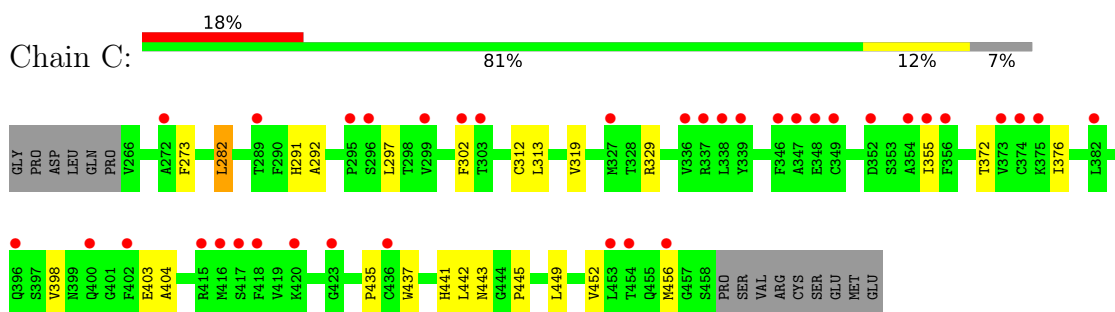
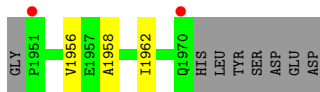• Molecule 1: Mothers against decapentaplegic homolog 2

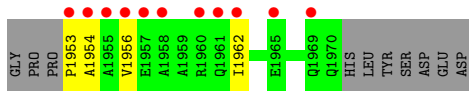• Molecule 2: CREB-binding protein

• Molecule 2: CREB-binding protein

Chain D:

• Molecule 2: CREB-binding protein

Chain F:
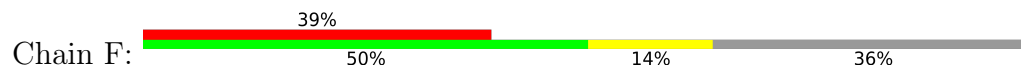
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	68.92Å 70.39Å 163.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.17 – 3.30 43.17 – 3.30	Depositor EDS
% Data completeness (in resolution range)	82.8 (43.17-3.30) 89.7 (43.17-3.30)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.86 (at 3.32Å)	Xtriage
Refinement program	PHENIX 1.12	Depositor
R, R_{free}	0.255 , 0.283 0.255 , 0.283	Depositor DCC
R_{free} test set	557 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	89.2	Xtriage
Anisotropy	0.843	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 41.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.038 for k,h,-l	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	5054	wwPDB-VP
Average B, all atoms (Å^2)	94.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.27% 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 (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	0.25	0/1582	0.45	0/2151
1	C	0.24	0/1574	0.45	0/2140
1	E	0.24	0/1568	0.45	0/2132
2	B	0.25	0/157	0.38	0/212
2	D	0.24	0/157	0.38	0/212
2	F	0.25	0/141	0.31	0/188
All	All	0.25	0/5179	0.44	0/7035

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 (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	1541	0	1489	20	0
1	C	1534	0	1481	16	0
1	E	1528	0	1476	22	0
2	B	155	0	150	2	0
2	D	155	0	150	2	0
2	F	141	0	136	4	0
All	All	5054	0	4882	55	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 6.

All (55) 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:313:LEU:O	1:A:329:ARG:NH2	2.32	0.63
1:C:398:VAL:HG21	2:D:1956:VAL:HG12	1.81	0.62
1:A:273:PHE:O	1:A:443:ASN:ND2	2.29	0.60
1:C:313:LEU:O	1:C:329:ARG:NH2	2.34	0.60
1:C:319:VAL:HG11	1:E:437:TRP:HH2	1.67	0.58
1:E:452:VAL:O	1:E:456:MET:HG2	2.06	0.56
1:A:388:GLN:NE2	2:B:1970:GLN:OE1	2.35	0.56
1:A:344:GLU:OE2	1:A:383:LYS:NZ	2.40	0.54
1:E:398:VAL:HG21	2:F:1956:VAL:HG12	1.93	0.51
1:A:450:ASP:OD2	1:E:310:ARG:NH1	2.41	0.50
1:A:403:GLU:OE1	1:A:403:GLU:N	2.36	0.50
1:C:403:GLU:OE1	1:C:403:GLU:N	2.36	0.50
1:E:292:ALA:HB2	1:E:297:LEU:HD13	1.92	0.49
1:E:403:GLU:OE1	1:E:403:GLU:N	2.35	0.48
2:F:1953:PRO:O	2:F:1956:VAL:HG22	2.13	0.48
1:C:319:VAL:HG11	1:E:437:TRP:CH2	2.47	0.48
1:A:319:VAL:HG11	1:C:437:TRP:HH2	1.78	0.48
1:C:292:ALA:HB2	1:C:297:LEU:HD13	1.97	0.47
1:C:302:PHE:O	1:C:312:CYS:HB3	2.15	0.47
1:C:273:PHE:CG	1:C:291:HIS:HB3	2.49	0.47
1:E:423:GLY:N	1:E:428:ARG:O	2.45	0.47
1:A:398:VAL:HG21	2:B:1956:VAL:HG12	1.97	0.46
1:A:292:ALA:HB2	1:A:297:LEU:HD13	1.99	0.45
1:E:302:PHE:O	1:E:312:CYS:HB3	2.16	0.45
2:D:1958:ALA:O	2:D:1962:ILE:HG12	2.15	0.45
1:E:268:TYR:CE1	1:E:451:LYS:HG2	2.52	0.45
1:A:384:ILE:CG2	1:A:414:ILE:HD13	2.47	0.45
1:A:345:VAL:O	1:A:384:ILE:N	2.49	0.45
1:A:302:PHE:O	1:A:312:CYS:HB3	2.16	0.44
1:C:452:VAL:O	1:C:456:MET:HG2	2.17	0.44
1:E:299:VAL:HG22	1:E:311:PHE:HB3	2.00	0.44
1:C:273:PHE:O	1:C:443:ASN:ND2	2.34	0.43
1:E:448:TRP:HB3	2:F:1962:ILE:HD13	2.00	0.43
1:A:450:ASP:OD2	1:E:304:ASP:HB2	2.18	0.43
1:E:325:VAL:HG22	1:E:435:PRO:HB3	2.00	0.43
1:E:447:GLN:O	1:E:451:LYS:HG3	2.18	0.43
1:A:319:VAL:HG11	1:C:437:TRP:CH2	2.53	0.43
1:C:282:LEU:HD13	1:C:435:PRO:HA	2.01	0.43
1:E:273:PHE:O	1:E:443:ASN:ND2	2.34	0.42
1:A:338:LEU:HD13	1:A:440:LEU:HD21	2.00	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:409:THR:O	1:A:413:THR:HG23	2.20	0.42
1:A:413:THR:HA	1:A:440:LEU:O	2.20	0.42
1:C:442:LEU:HB3	1:C:445:PRO:HD2	2.02	0.42
2:F:1953:PRO:HB2	2:F:1954:ALA:H	1.73	0.41
1:E:356:PHE:HB2	1:E:417:SER:HB3	2.02	0.41
1:A:384:ILE:HG22	1:A:414:ILE:HD13	2.01	0.41
1:A:403:GLU:HG2	1:A:404:ALA:H	1.85	0.41
1:A:395:ALA:O	1:A:398:VAL:HG22	2.20	0.41
1:E:282:LEU:HD12	1:E:282:LEU:HA	1.89	0.41
1:C:403:GLU:HG2	1:C:404:ALA:H	1.86	0.41
1:E:307:ASN:HB3	1:E:310:ARG:HB2	2.04	0.40
1:E:384:ILE:CG2	1:E:414:ILE:HD13	2.52	0.40
1:E:384:ILE:HG22	1:E:414:ILE:HD13	2.03	0.40
1:C:355:ILE:HG13	1:C:376:ILE:HB	2.04	0.40
1:E:369:HIS:ND1	1:E:370:PRO:HD2	2.36	0.40

There are no symmetry-related clashes.

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	192/208 (92%)	185 (96%)	7 (4%)	0	100	100
1	C	191/208 (92%)	183 (96%)	8 (4%)	0	100	100
1	E	190/208 (91%)	183 (96%)	7 (4%)	0	100	100
2	B	18/28 (64%)	18 (100%)	0	0	100	100
2	D	18/28 (64%)	18 (100%)	0	0	100	100
2	F	16/28 (57%)	16 (100%)	0	0	100	100
All	All	625/708 (88%)	603 (96%)	22 (4%)	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	168/181 (93%)	165 (98%)	3 (2%)	59	78
1	C	167/181 (92%)	163 (98%)	4 (2%)	49	73
1	E	166/181 (92%)	162 (98%)	4 (2%)	49	73
2	B	15/22 (68%)	13 (87%)	2 (13%)	4	17
2	D	15/22 (68%)	15 (100%)	0	100	100
2	F	13/22 (59%)	13 (100%)	0	100	100
All	All	544/609 (89%)	531 (98%)	13 (2%)	49	73

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

Mol	Chain	Res	Type
1	A	282	LEU
1	A	441	HIS
1	A	449	LEU
2	B	1965	GLU
2	B	1970	GLN
1	C	282	LEU
1	C	372	THR
1	C	441	HIS
1	C	449	LEU
1	E	282	LEU
1	E	383	LYS
1	E	441	HIS
1	E	449	LEU

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	194/208 (93%)	1.02	28 (14%)	2	2	73, 81, 93, 100	0
1	C	193/208 (92%)	1.18	37 (19%)	1	1	85, 95, 104, 113	0
1	E	192/208 (92%)	1.42	50 (26%)	0	0	74, 100, 126, 141	0
2	B	20/28 (71%)	0.67	3 (15%)	2	2	76, 82, 95, 97	0
2	D	20/28 (71%)	0.95	2 (10%)	7	7	92, 97, 112, 115	0
2	F	18/28 (64%)	2.43	11 (61%)	0	0	118, 127, 133, 135	0
All	All	637/708 (89%)	1.22	131 (20%)	1	1	73, 93, 121, 141	0

All (131) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	374	CYS	6.8
1	E	400	GLN	5.1
1	E	267	THR	5.0
1	C	303	THR	5.0
1	E	448	TRP	4.9
1	E	387	ASN	4.7
1	E	421	GLY	4.6
1	C	347	ALA	4.6
1	E	268	TYR	4.5
1	A	298	THR	4.4
1	E	452	VAL	4.4
1	E	397	SER	4.4
1	C	338	LEU	4.4
1	A	374	CYS	4.2
1	E	354	ALA	4.1
1	E	394	LEU	4.1
2	F	1969	GLN	4.1
2	F	1955	ALA	4.0
1	C	336	VAL	3.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	415	ARG	3.8
1	C	418	PHE	3.7
2	F	1962	ILE	3.7
1	E	457	GLY	3.6
1	E	266	VAL	3.6
1	E	402	PHE	3.6
2	F	1956	VAL	3.6
1	E	418	PHE	3.5
2	F	1958	ALA	3.5
1	E	396	GLN	3.5
1	A	402	PHE	3.4
2	D	1970	GLN	3.4
1	C	382	LEU	3.3
1	A	400	GLN	3.3
1	E	334	ARG	3.3
1	C	295	PRO	3.3
1	C	417	SER	3.2
1	A	385	PHE	3.2
1	C	352	ASP	3.2
1	A	340	TYR	3.2
1	C	337	ARG	3.2
1	C	420	LYS	3.1
1	E	355	ILE	3.1
1	E	455	GLN	3.1
1	C	355	ILE	3.1
1	A	406	TYR	3.1
1	E	401	GLY	3.1
1	E	385	PHE	3.1
1	E	403	GLU	3.1
2	D	1951	PRO	3.0
1	C	302	PHE	3.0
1	E	336	VAL	3.0
1	C	339	TYR	3.0
1	C	396	GLN	3.0
1	E	345	VAL	3.0
2	B	1951	PRO	3.0
1	E	347	ALA	2.9
1	E	346	PHE	2.9
2	F	1961	GLN	2.9
2	F	1954	ALA	2.9
1	E	398	VAL	2.9
2	F	1953	PRO	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	408	LEU	2.8
1	E	453	LEU	2.8
1	A	414	ILE	2.8
1	E	426	TYR	2.8
2	F	1957	GLU	2.7
1	E	456	MET	2.7
1	A	445	PRO	2.7
1	A	309	GLU	2.7
1	C	402	PHE	2.7
1	E	341	ILE	2.7
1	C	349	CYS	2.7
1	E	353	SER	2.6
1	E	449	LEU	2.6
1	C	348	GLU	2.6
1	C	375	LYS	2.6
1	E	416	MET	2.6
1	A	390	PHE	2.6
1	E	384	ILE	2.6
1	E	454	THR	2.6
2	F	1960	ARG	2.5
1	C	436	CYS	2.5
1	A	334	ARG	2.5
1	C	356	PHE	2.5
1	E	417	SER	2.5
1	E	395	ALA	2.5
1	A	335	GLY	2.5
1	C	327	MET	2.5
1	E	383	LYS	2.5
1	C	296	SER	2.5
1	E	376	ILE	2.5
1	A	380	CYS	2.5
1	C	453	LEU	2.5
1	C	346	PHE	2.4
1	A	283	ASN	2.4
1	C	454	THR	2.3
1	A	317	SER	2.3
1	E	295	PRO	2.3
1	E	450	ASP	2.3
1	C	299	VAL	2.3
1	A	425	GLU	2.3
1	C	289	THR	2.3
2	F	1965	GLU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	336	VAL	2.3
1	C	272	ALA	2.3
1	C	373	VAL	2.2
1	E	388	GLN	2.2
1	C	416	MET	2.2
1	E	283	ASN	2.2
1	E	407	GLN	2.2
1	A	357	VAL	2.2
1	C	400	GLN	2.2
1	E	390	PHE	2.2
1	A	342	GLY	2.2
1	A	417	SER	2.2
1	A	297	LEU	2.2
1	A	338	LEU	2.2
2	B	1962	ILE	2.2
1	E	374	CYS	2.2
1	E	309	GLU	2.2
1	C	423	GLY	2.1
1	A	345	VAL	2.1
1	A	440	LEU	2.1
1	E	269	SER	2.1
2	B	1969	GLN	2.1
1	C	354	ALA	2.1
1	A	299	VAL	2.0
1	A	418	PHE	2.0
1	A	275	CYS	2.0
1	E	370	PRO	2.0
1	C	456	MET	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 (i)

There are no such residues in this entry.