Full wwPDB X-ray Structure Validation Report (i)

May 14, 2020 – 03:55 pm BST

PDB ID	:	5ZOJ
Title	:	Crystal structure of human SMAD2-MAN1 complex
Authors	:	Miyazono, K.; Ohno, Y.; Ito, T.; Tanokura, M.
Deposited on	:	2018-04-13
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 (i) symbol.

The following versions of software and data (see references (i)) 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 ⓘ

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.
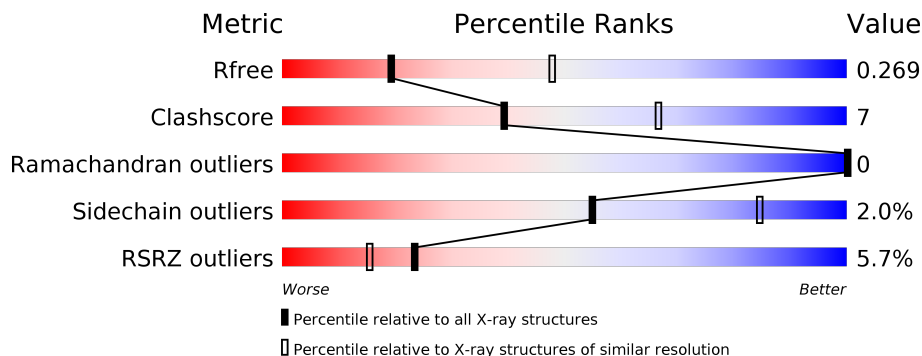
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 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	200	79% 15% 7%
1	B	200	76% 16% · 8%
1	C	200	77% 16% · 7%
2	D	132	78% 17% · ·
2	E	132	65% 23% 12%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6383 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	187	Total 1492	C 946	N 263	O 271	S 12	0	0	0
1	B	185	Total 1476	C 936	N 261	O 267	S 12	0	0	0
1	C	186	Total 1483	C 941	N 262	O 268	S 12	0	0	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	259	GLY	-	expression tag	UNP Q15796
A	260	PRO	-	expression tag	UNP Q15796
A	261	GLY	-	expression tag	UNP Q15796
B	259	GLY	-	expression tag	UNP Q15796
B	260	PRO	-	expression tag	UNP Q15796
B	261	GLY	-	expression tag	UNP Q15796
C	259	GLY	-	expression tag	UNP Q15796
C	260	PRO	-	expression tag	UNP Q15796
C	261	GLY	-	expression tag	UNP Q15796

- Molecule 2 is a protein called Inner nuclear membrane protein Man1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	127	Total 1020	C 649	N 184	O 181	S 6	0	0	0
2	E	116	Total 912	C 581	N 159	O 166	S 6	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	759	GLY	-	expression tag	UNP Q9Y2U8

Continued on next page...

Continued from previous page...

| Chain	Residue	Modelled	Actual	Comment	Reference
D	760	PRO	-	expression tag	UNP Q9Y2U8
D	761	GLY	-	expression tag	UNP Q9Y2U8
E	759	GLY	-	expression tag	UNP Q9Y2U8
E	760	PRO	-	expression tag	UNP Q9Y2U8
E	761	GLY	-	expression tag	UNP Q9Y2U8

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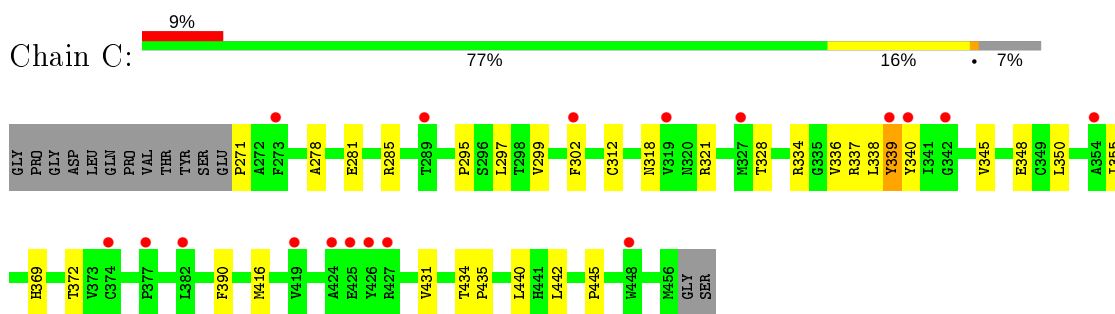
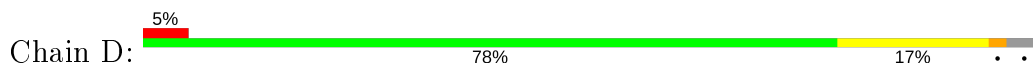• Molecule 1: Mothers against decapentaplegic homolog 2
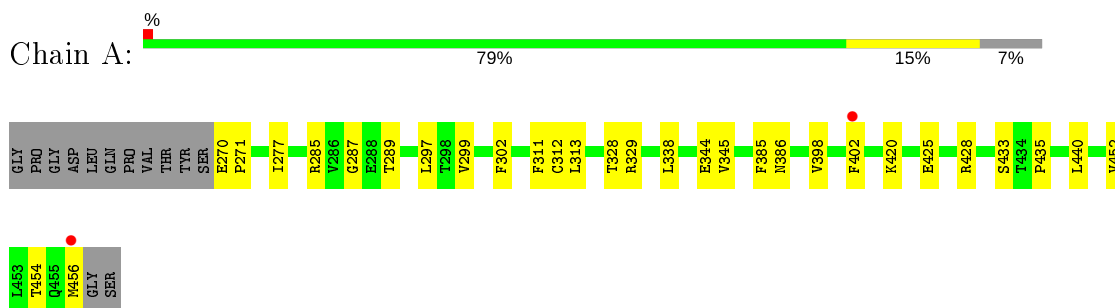
• Molecule 1: Mothers against decapentaplegic homolog 2
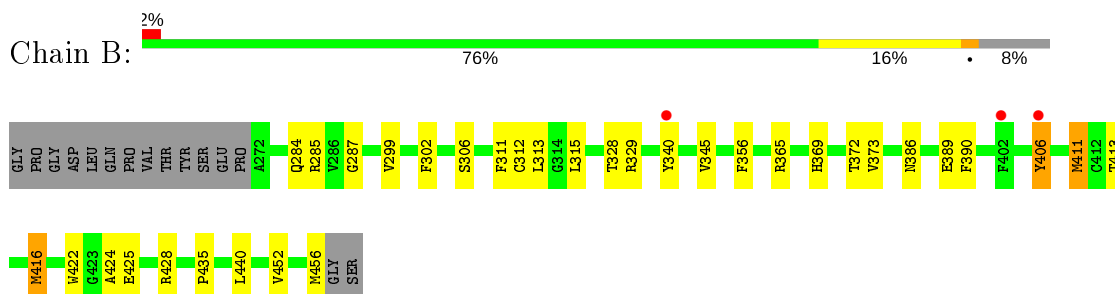
• Molecule 1: Mothers against decapentaplegic homolog 2

• Molecule 2: Inner nuclear membrane protein Man1

GLY PRO GLY S762 H771 L772 D773 R774 R775 P784 C785 L786 K787 I788 R789 E808 E812 N817 I823 A824 V825 R830 C833 V834 Y835 V836 K837 P841 E842 V863 L866 R867 D869 R870 Y871 A878 P884 S888 ASN LYS

• Molecule 2: Inner nuclear membrane protein Man1

Chain E:

13%

65% 23% 12%

GLY PRO GLY S762 W765 D773 R774 R775 P779 L782 T783 P784 C785 L786 K787 I788 R789 F792 D793 P794 I798 Q801 W802 H803 K813 H822 I823 A824 V825 D826 K827 N828 S829 R830 V834 K837 C838 L839 S840 P841 E842 Y843 A844 G845 K846 K849 V863

K864 Y865 L866 ARG LEU ASP ARG TYR HIS HIS ARG PHE PRO GLN A878 L879 T880 S881 N882 T883 P884 L885 S888 ASN LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	176.81Å　　176.81Å　　73.85Å 90.00°　　　90.00°　　　120.00°	Depositor
Resolution (Å)	42.47　–　2.79 44.20　–　2.79	Depositor EDS
% Data completeness (in resolution range)	91.3 (42.47-2.79) 97.0 (44.20-2.79)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.11 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.12rc2_2821: ???	Depositor
R, R_{free}	0.242　,　0.269 0.242　,　0.269	Depositor DCC
R_{free} test set	1526 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å^2)	65.8	Xtriage
Anisotropy	0.439	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 51.3	EDS
L-test for twinning[2]	$< |L| > = 0.51, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.017 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	6383	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.96% 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	0/1531	0.44	0/2081
1	B	0.26	0/1514	0.47	0/2057
1	C	0.25	0/1522	0.47	0/2068
2	D	0.24	0/1048	0.44	0/1419
2	E	0.25	0/934	0.42	0/1264
All	All	0.25	0/6549	0.45	0/8889

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	1492	0	1443	21	0
1	B	1476	0	1430	20	0
1	C	1483	0	1438	20	0
2	D	1020	0	1009	17	0
2	E	912	0	907	22	0
All	All	6383	0	6227	93	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 (93) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Full wwPDB X-ray Structure Validation Report

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:784:PRO:HD3	2:E:841:PRO:HG3	1.68	0.74
1:A:338:LEU:HD13	1:A:440:LEU:HD21	1.72	0.71
2:D:823:ILE:HG22	2:D:836:VAL:HG22	1.75	0.68
1:A:425:GLU:HG2	2:D:771:HIS:HB3	1.77	0.66
2:D:869:ASP:OD1	2:D:869:ASP:N	2.26	0.65
2:E:789:ARG:HH12	2:E:864:LYS:HB2	1.62	0.64
2:E:840:SER:HB3	2:E:843:TYR:HD1	1.62	0.64
2:D:787:LYS:NZ	2:D:833:CYS:SG	2.72	0.63
1:B:369:HIS:HB3	1:B:372:THR:HG23	1.81	0.62
2:E:774:ARG:HA	2:E:774:ARG:HH11	1.63	0.62
1:C:338:LEU:HD13	1:C:440:LEU:HD21	1.83	0.60
2:D:775:ARG:HB2	2:D:775:ARG:HH11	1.66	0.60
1:A:285:ARG:NH2	1:A:289:THR:OG1	2.39	0.55
1:C:278:ALA:HB1	1:C:285:ARG:NH1	2.22	0.55
2:E:825:VAL:HG12	2:E:834:VAL:HG13	1.89	0.54
1:C:336:VAL:HG12	1:C:355:ILE:HG21	1.89	0.54
1:A:328:THR:HG21	1:A:435:PRO:HG2	1.90	0.54
1:B:328:THR:HG21	1:B:435:PRO:HG2	1.89	0.54
1:B:373:VAL:HG11	1:B:422:TRP:O	2.08	0.54
1:C:281:GLU:OE1	1:C:318:ASN:ND2	2.37	0.53
2:D:775:ARG:H	2:D:775:ARG:HD3	1.75	0.52
1:B:365:ARG:NH1	1:B:389:GLU:OE1	2.44	0.51
2:E:773:ASP:OD1	2:E:775:ARG:HG2	2.10	0.51
2:E:845:GLY:O	2:E:849:LYS:HG3	2.11	0.51
1:A:285:ARG:NH1	1:A:287:GLY:O	2.41	0.50
1:A:452:VAL:O	1:A:456:MET:HG3	2.11	0.50
1:C:328:THR:HG21	1:C:435:PRO:HG2	1.94	0.50
1:C:431:VAL:O	1:C:434:THR:HG22	2.11	0.50
2:E:774:ARG:NH2	2:E:779:PRO:HG3	2.27	0.50
2:E:842:GLU:OE1	2:E:842:GLU:N	2.44	0.49
1:C:369:HIS:HB3	1:C:372:THR:HG23	1.94	0.49
2:D:867:ARG:NH1	2:D:869:ASP:OD2	2.46	0.49
2:E:803:HIS:HB2	2:E:825:VAL:HG21	1.94	0.48
2:D:808:GLU:O	2:D:812:GLU:HG3	2.13	0.48
1:B:285:ARG:NH2	1:B:287:GLY:O	2.46	0.48
2:D:825:VAL:HG12	2:D:834:VAL:HG13	1.94	0.48
1:B:428:ARG:NH1	1:B:435:PRO:HD3	2.28	0.48
1:C:442:LEU:HB3	1:C:445:PRO:HD2	1.96	0.47
2:E:823:ILE:HD12	2:E:885:LEU:HD12	1.95	0.47
2:D:837:LYS:HD2	2:D:878:ALA:HB1	1.97	0.47
1:B:425:GLU:CD	2:E:775:ARG:HH22	2.17	0.47
2:D:785:CYS:HB2	2:D:866:LEU:O	2.15	0.47

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:295:PRO:HB2	1:C:339:TYR:HE1	1.80	0.47
2:E:840:SER:HB3	2:E:843:TYR:CD1	2.47	0.47
2:E:846:LYS:HD2	2:E:846:LYS:HA	1.80	0.47
1:A:313:LEU:O	1:A:329:ARG:NH2	2.43	0.46
1:C:299:VAL:HB	1:C:336:VAL:HG22	1.97	0.46
1:B:369:HIS:HB2	2:E:775:ARG:HB2	1.96	0.46
1:C:281:GLU:OE2	1:C:321:ARG:NH1	2.47	0.46
1:B:299:VAL:HG22	1:B:311:PHE:HB3	1.98	0.46
2:E:774:ARG:CZ	2:E:779:PRO:HG3	2.45	0.46
1:A:420:LYS:HD2	1:A:428:ARG:NH1	2.30	0.46
1:C:339:TYR:HD1	1:C:340:TYR:H	1.64	0.46
1:A:270:GLU:HB2	1:A:271:PRO:HD3	1.98	0.45
1:C:337:ARG:HB2	1:C:348:GLU:HB3	1.99	0.45
2:D:788:ILE:HG12	2:D:863:VAL:HG22	1.99	0.45
1:C:302:PHE:O	1:C:312:CYS:HB3	2.16	0.45
1:A:425:GLU:OE2	2:D:775:ARG:NH2	2.38	0.45
1:C:337:ARG:HG2	1:C:350:LEU:HD11	1.97	0.45
1:B:340:TYR:HD1	1:B:345:VAL:HG22	1.82	0.45
1:B:313:LEU:O	1:B:329:ARG:NH2	2.46	0.44
2:E:783:THR:HG21	2:E:866:LEU:O	2.16	0.44
1:A:299:VAL:HG22	1:A:311:PHE:HB3	1.99	0.44
2:D:775:ARG:NH1	2:D:775:ARG:HB2	2.30	0.44
1:A:297:LEU:N	1:A:338:LEU:O	2.51	0.44
1:A:398:VAL:HG23	1:A:456:MET:SD	2.58	0.43
1:A:302:PHE:O	1:A:312:CYS:HB3	2.18	0.43
1:A:344:GLU:OE2	1:A:386:ASN:ND2	2.50	0.43
1:A:402:PHE:HA	1:A:456:MET:HE1	1.99	0.43
2:E:765:TRP:CD2	2:E:813:LYS:HD3	2.54	0.43
1:B:424:ALA:HB3	2:E:775:ARG:NH2	2.34	0.43
1:B:413:THR:HA	1:B:440:LEU:O	2.19	0.42
1:B:406:TYR:O	1:B:406:TYR:HD1	2.03	0.42
1:A:454:THR:HG22	1:C:302:PHE:CD2	2.55	0.42
1:B:452:VAL:O	1:B:456:MET:HG2	2.20	0.42
2:E:792:PHE:HE2	2:E:794:PRO:HG3	1.84	0.42
1:B:386:ASN:OD1	1:B:389:GLU:HG3	2.20	0.42
1:A:428:ARG:HD3	1:A:433:SER:O	2.20	0.41
1:B:306:SER:HG	1:C:271:PRO:N	2.18	0.41
1:C:340:TYR:HD1	1:C:345:VAL:HG22	1.84	0.41
1:A:277:ILE:HG12	1:A:440:LEU:HG	2.03	0.41
2:E:787:LYS:O	2:E:863:VAL:HA	2.20	0.41
1:B:302:PHE:O	1:B:312:CYS:HB3	2.20	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:830:ARG:O	2:E:830:ARG:HD3	2.21	0.41
1:A:345:VAL:CG1	1:A:385:PHE:HB3	2.51	0.41
1:A:420:LYS:HD2	1:A:428:ARG:HH12	1.85	0.41
1:B:390:PHE:HD1	1:B:411:MET:HG2	1.85	0.41
1:C:297:LEU:HD22	1:C:338:LEU:HD12	2.03	0.41
1:B:356:PHE:O	1:B:416:MET:HA	2.21	0.40
1:C:390:PHE:CE2	1:C:445:PRO:HB3	2.57	0.40
2:D:842:GLU:OE1	2:D:842:GLU:N	2.54	0.40
2:D:773:ASP:OD1	2:D:775:ARG:NH1	2.54	0.40
2:D:784:PRO:HD3	2:D:841:PRO:HG3	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	185/200 (92%)	180 (97%)	5 (3%)	0	100	100
1	B	183/200 (92%)	178 (97%)	5 (3%)	0	100	100
1	C	184/200 (92%)	180 (98%)	4 (2%)	0	100	100
2	D	125/132 (95%)	119 (95%)	6 (5%)	0	100	100
2	E	112/132 (85%)	108 (96%)	4 (4%)	0	100	100
All	All	789/864 (91%)	765 (97%)	24 (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	162/172 (94%)	162 (100%)	0	100	100
1	B	160/172 (93%)	155 (97%)	5 (3%)	40	74
1	C	161/172 (94%)	158 (98%)	3 (2%)	57	85
2	D	112/115 (97%)	108 (96%)	4 (4%)	35	69
2	E	101/115 (88%)	99 (98%)	2 (2%)	55	84
All	All	696/746 (93%)	682 (98%)	14 (2%)	55	84

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	284	GLN
1	B	315	LEU
1	B	406	TYR
1	B	411	MET
1	B	416	MET
1	C	334	ARG
1	C	339	TYR
1	C	416	MET
2	D	775	ARG
2	D	789	ARG
2	D	868	LEU
2	D	869	ASP
2	E	801	GLN
2	E	883	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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	187/200 (93%)	0.20	2 (1%)	80	75	36, 59, 104, 134	0
1	B	185/200 (92%)	0.19	3 (1%)	72	66	41, 60, 87, 103	0
1	C	186/200 (93%)	0.71	18 (9%)	7	4	53, 97, 125, 161	0
2	D	127/132 (96%)	0.41	6 (4%)	31	22	51, 75, 113, 137	0
2	E	116/132 (87%)	0.90	17 (14%)	2	1	50, 96, 132, 162	0
All	All	801/864 (92%)	0.45	46 (5%)	23	15	36, 73, 118, 162	0

All (46) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	424	ALA	6.2
1	C	354	ALA	5.4
2	E	782	LEU	5.0
2	E	866	LEU	4.8
2	E	823	ILE	4.6
2	E	789	ARG	3.6
2	D	817	ASN	3.5
2	E	822	HIS	3.4
2	E	843	TYR	3.4
1	C	426	TYR	3.4
2	E	837	LYS	3.4
2	E	785	CYS	3.3
1	C	419	VAL	3.2
1	A	402	PHE	3.2
1	C	340	TYR	3.2
1	C	427	ARG	3.1
2	E	839	LEU	3.0
2	E	798	ILE	3.0
2	E	825	VAL	3.0
1	C	425	GLU	3.0

Continued on next page...

Full wwPDB X-ray Structure Validation Report

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	273	PHE	2.9
2	E	828	ASN	2.9
2	E	882	ASN	2.8
1	C	302	PHE	2.8
1	B	406	TYR	2.7
2	E	784	PRO	2.7
1	C	377	PRO	2.6
1	C	448	TRP	2.5
2	E	880	THR	2.5
1	C	342	GLY	2.4
2	D	871	TYR	2.4
1	A	456	MET	2.4
1	B	340	TYR	2.3
2	D	823	ILE	2.3
2	E	826	ASP	2.3
1	C	339	TYR	2.3
2	D	884	PRO	2.3
1	C	374	CYS	2.3
1	C	327	MET	2.3
2	D	774	ARG	2.2
1	C	289	THR	2.1
2	E	885	LEU	2.1
1	C	319	VAL	2.0
2	D	830	ARG	2.0
1	B	402	PHE	2.0
1	C	382	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.