wwPDB X-ray Structure Validation Summary Report (i)

Sep 17, 2023 – 03:45 PM EDT

PDB ID	:	4XEI
Title	:	Orthorhombic isomorph of bovine Arp2/3 complex
Authors	:	Jurgenson, C.J.; Pollard, T.P.
Deposited on	:	2014-12-23
Resolution	:	3.87 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

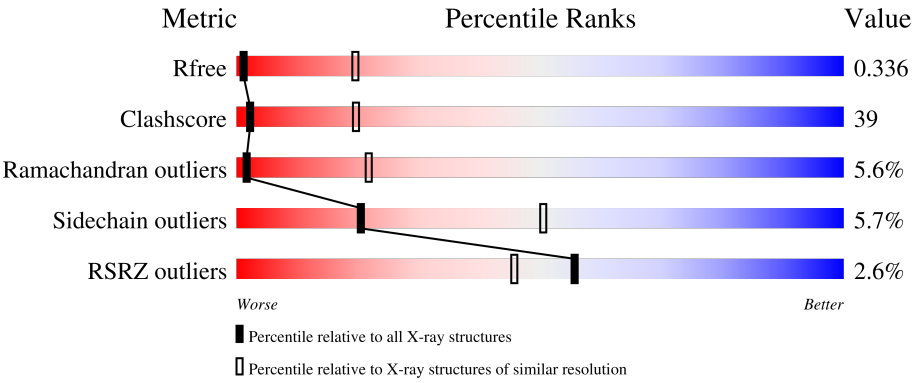
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.87 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	1026 (4.12-3.64)
Clashscore	141614	1045 (4.10-3.66)
Ramachandran outliers	138981	1008 (4.10-3.66)
Sidechain outliers	138945	1001 (4.10-3.66)
RSRZ outliers	127900	1213 (4.16-3.60)

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	418	% 51% 39% 5% ·
2	B	394	% 34% 19% · 45%
3	C	372	5% 40% 45% 7% 8%
4	D	300	% 51% 39% · 6%
5	E	178	4% 60% 32% 6% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	F	168	2% 54% 40% 5%
7	G	151	2% 54% 30% 5% 11%

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 13731 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 Actin-related protein 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	400	Total	C	N	O	S	0	0	0
			3185	2045	533	592	15			

- Molecule 2 is a protein called Actin-related protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	218	Total	C	N	O	S	0	0	0
			1759	1133	301	319	6			

- Molecule 3 is a protein called Actin-related protein 2/3 complex subunit 1B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	344	Total	C	N	O	S	0	0	0
			2661	1688	467	487	19			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	58	VAL	ILE	conflict	UNP Q58CQ2

- Molecule 4 is a protein called Actin-related protein 2/3 complex subunit 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	282	Total	C	N	O	S	0	0	0
			2272	1443	394	427	8			

- Molecule 5 is a protein called Actin-related protein 2/3 complex subunit 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	174	Total	C	N	O	S	0	0	0
			1414	908	236	261	9			

- Molecule 6 is a protein called Actin-related protein 2/3 complex subunit 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	168	Total	C	N	O	S	0	0	0
			1378	880	240	248	10			

- Molecule 7 is a protein called Actin-related protein 2/3 complex subunit 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	135	Total	C	N	O	S	0	0	0
			1014	636	171	204	3			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	17	ASP	GLY	conflict	UNP Q3SYX9
G	28	ASP	GLU	conflict	UNP Q3SYX9

- Molecule 8 is water.

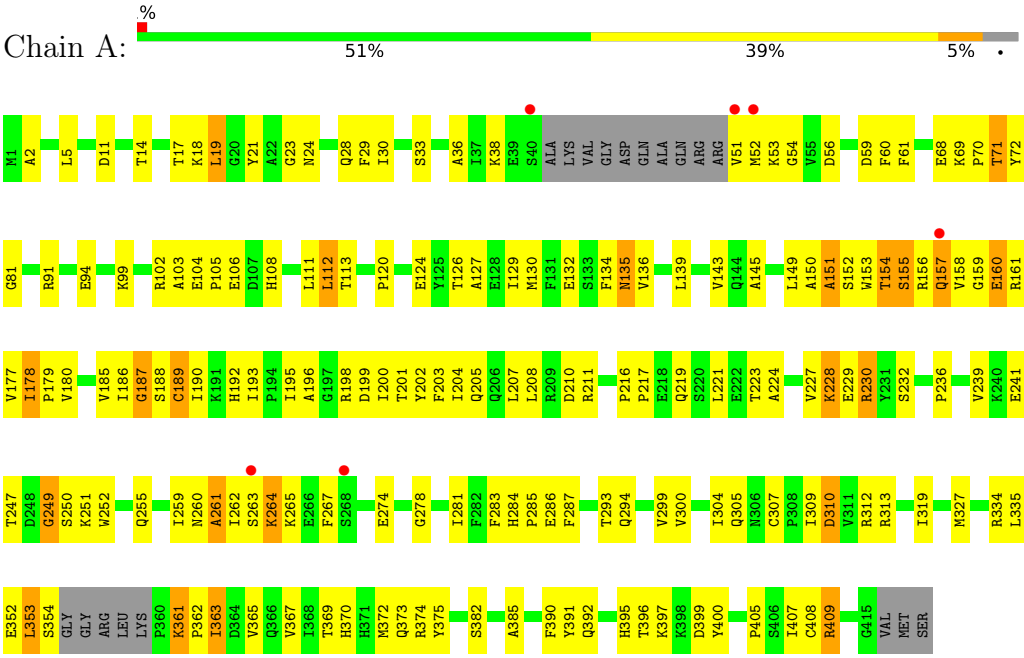
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	8	Total	O	0	0
			8	8		
8	B	5	Total	O	0	0
			5	5		
8	C	10	Total	O	0	0
			10	10		
8	D	8	Total	O	0	0
			8	8		
8	E	9	Total	O	0	0
			9	9		
8	F	3	Total	O	0	0
			3	3		
8	G	5	Total	O	0	0
			5	5		

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.

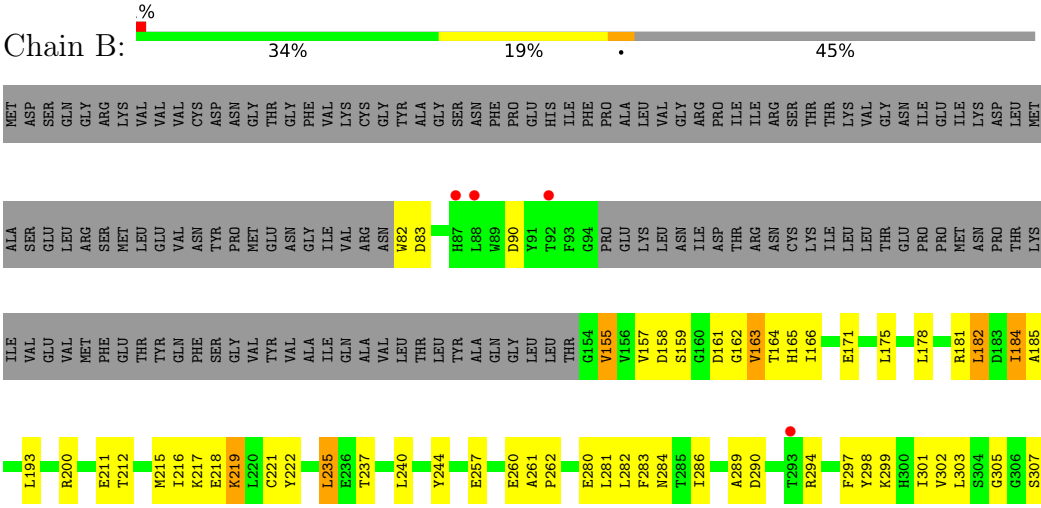
• Molecule 1: Actin-related protein 3

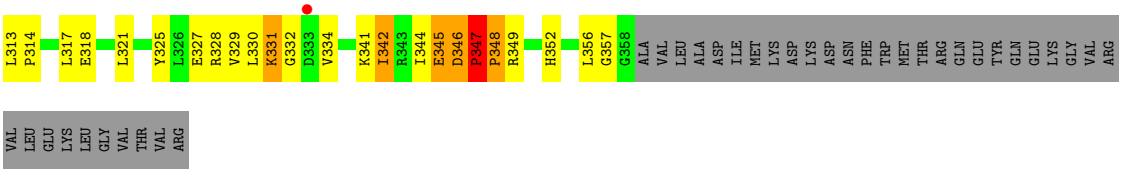
Chain A:



• Molecule 2: Actin-related protein 2

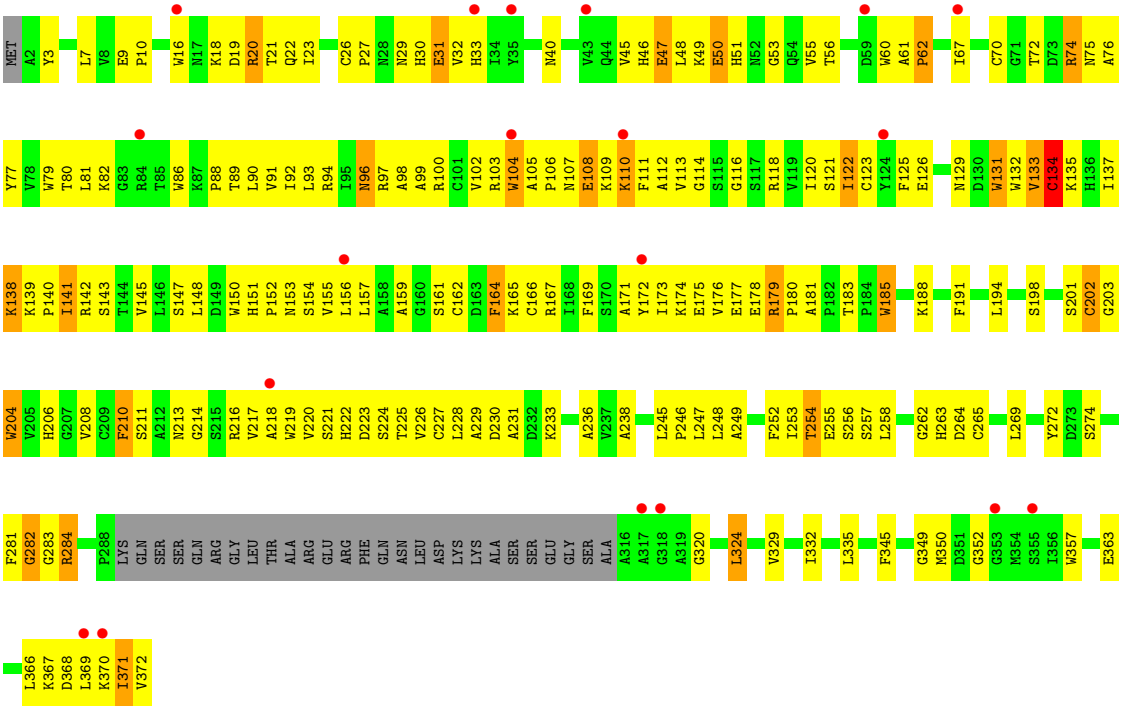
Chain B:





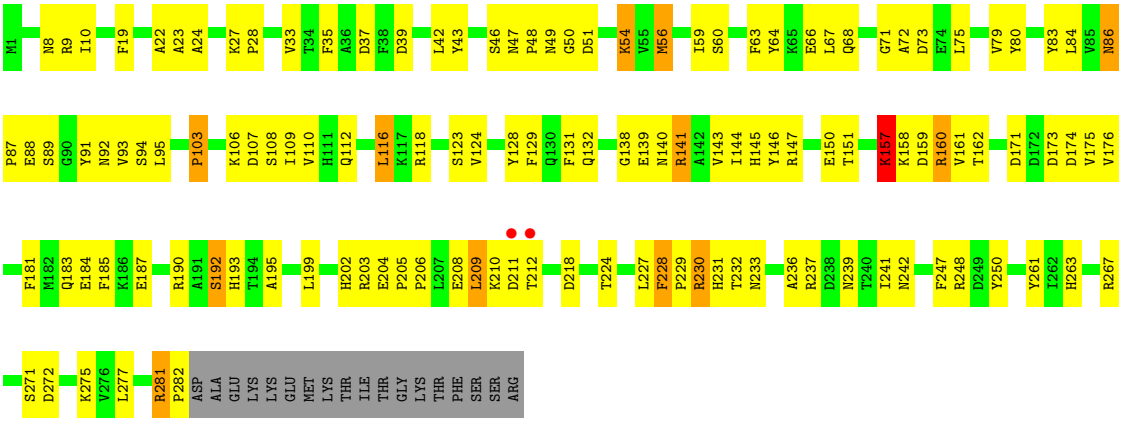
• Molecule 3: Actin-related protein 2/3 complex subunit 1B

Chain C:



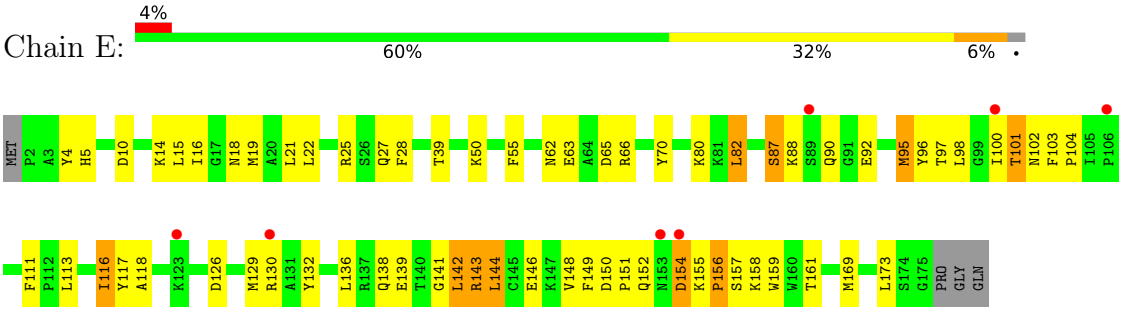
• Molecule 4: Actin-related protein 2/3 complex subunit 2

Chain D:



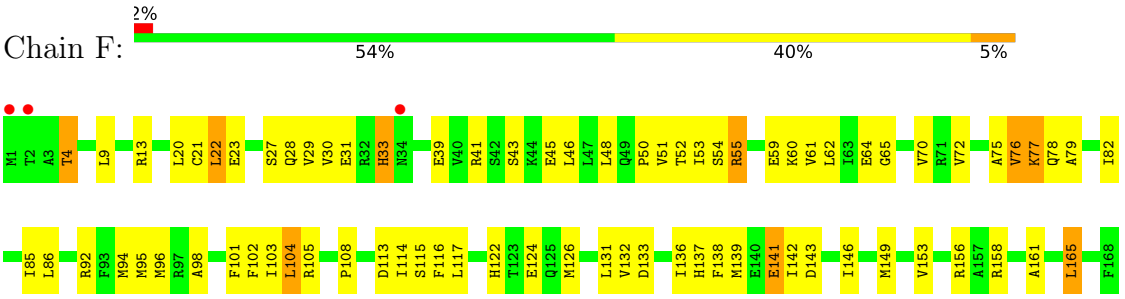
• Molecule 5: Actin-related protein 2/3 complex subunit 3

Chain E:



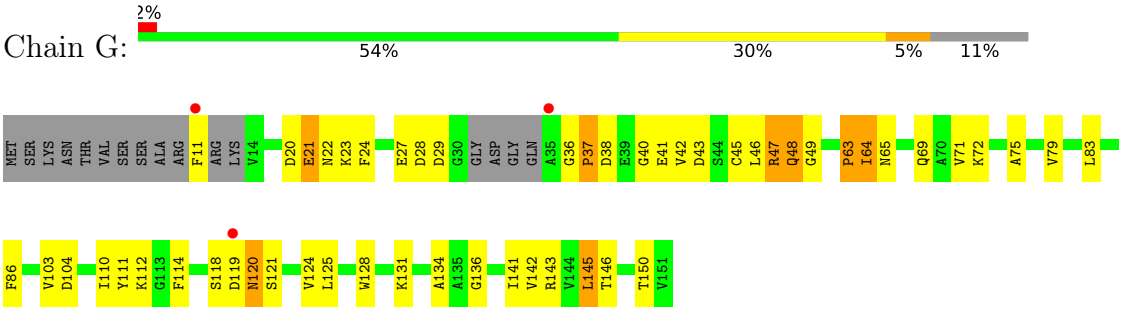
• Molecule 6: Actin-related protein 2/3 complex subunit 4

Chain F:



• Molecule 7: Actin-related protein 2/3 complex subunit 5

Chain G:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	105.57Å 156.98Å 178.09Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.41 – 3.87 45.41 – 3.87	Depositor EDS
% Data completeness (in resolution range)	95.6 (45.41-3.87) 95.6 (45.41-3.87)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.54 (at 3.88Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.279 , 0.329 0.305 , 0.336	Depositor DCC
R_{free} test set	1344 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	104.5	Xtriage
Anisotropy	0.056	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 67.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.82	EDS
Total number of atoms	13731	wwPDB-VP
Average B, all atoms (Å^2)	95.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.73% 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	0/3266	0.54	0/4432
2	B	0.30	0/1797	0.60	1/2426 (0.0%)
3	C	0.24	0/2730	0.47	0/3707
4	D	0.27	0/2321	0.51	0/3135
5	E	0.30	0/1448	0.49	0/1953
6	F	0.28	0/1400	0.49	0/1878
7	G	0.28	0/1025	0.53	0/1382
All	All	0.28	0/13987	0.52	1/18913 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	347	PRO	C-N-CD	-7.29	104.55	120.60

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	3185	0	3113	247	2
2	B	1759	0	1772	116	0
3	C	2661	0	2613	361	0
4	D	2272	0	2228	139	6
5	E	1414	0	1416	77	6

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	F	1378	0	1422	131	0
7	G	1014	0	1007	57	2
8	A	8	0	0	0	0
8	B	5	0	0	0	0
8	C	10	0	0	0	0
8	D	8	0	0	0	0
8	E	9	0	0	6	0
8	F	3	0	0	1	0
8	G	5	0	0	0	0
All	All	13731	0	13571	1071	8

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 39.

The worst 5 of 1071 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)
3:C:140:PRO:HD2	3:C:169:PHE:CZ	1.27	1.59
1:A:178:ILE:HG23	1:A:190:ILE:CD1	1.23	1.57
1:A:200:ILE:CD1	1:A:281:ILE:HD11	1.34	1.52
1:A:178:ILE:CG2	1:A:190:ILE:CD1	2.05	1.33
1:A:200:ILE:HD13	1:A:281:ILE:CD1	1.57	1.33

The worst 5 of 8 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:157:LYS:CB	5:E:27:GLN:NE2[3_555]	1.39	0.81
4:D:157:LYS:CD	5:E:27:GLN:OE1[3_555]	1.79	0.41
1:A:265:LYS:N	7:G:120:ASN:ND2[3_545]	1.83	0.37
4:D:183:GLN:NE2	5:E:151:PRO:CG[3_555]	2.02	0.18
4:D:157:LYS:CG	5:E:27:GLN:OE1[3_555]	2.08	0.12

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	394/418 (94%)	322 (82%)	47 (12%)	25 (6%)	1 20
2	B	214/394 (54%)	173 (81%)	26 (12%)	15 (7%)	1 17
3	C	340/372 (91%)	270 (79%)	49 (14%)	21 (6%)	1 20
4	D	280/300 (93%)	228 (81%)	35 (12%)	17 (6%)	1 20
5	E	172/178 (97%)	150 (87%)	16 (9%)	6 (4%)	3 30
6	F	166/168 (99%)	148 (89%)	14 (8%)	4 (2%)	6 37
7	G	130/151 (86%)	108 (83%)	15 (12%)	7 (5%)	2 23
All	All	1696/1981 (86%)	1399 (82%)	202 (12%)	95 (6%)	2 22

5 of 95 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	135	ASN
1	A	151	ALA
1	A	160	GLU
1	A	217	PRO
1	A	230	ARG

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	343/363 (94%)	331 (96%)	12 (4%)	36 62
2	B	190/345 (55%)	177 (93%)	13 (7%)	16 46
3	C	290/313 (93%)	271 (93%)	19 (7%)	16 46
4	D	246/264 (93%)	235 (96%)	11 (4%)	27 56
5	E	156/159 (98%)	142 (91%)	14 (9%)	9 36
6	F	155/155 (100%)	148 (96%)	7 (4%)	27 56

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
7	G	109/124 (88%)	100 (92%)	9 (8%)	11 40
All	All	1489/1723 (86%)	1404 (94%)	85 (6%)	20 50

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	E	25	ARG
6	F	55	ARG
5	E	62	ASN
5	E	143	ARG
6	F	165	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
3	C	96	ASN
4	D	111	HIS
4	D	132	GLN
4	D	145	HIS
5	E	18	ASN

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	400/418 (95%)	-0.06	6 (1%)	73	65	30, 83, 125, 158	0
2	B	218/394 (55%)	0.13	5 (2%)	60	51	62, 97, 143, 178	0
3	C	344/372 (92%)	0.52	19 (5%)	25	20	84, 120, 145, 165	0
4	D	282/300 (94%)	-0.13	2 (0%)	87	82	56, 82, 113, 165	0
5	E	174/178 (97%)	0.17	7 (4%)	38	31	57, 91, 119, 134	0
6	F	168/168 (100%)	-0.09	3 (1%)	68	60	55, 77, 101, 289	0
7	G	135/151 (89%)	0.15	3 (2%)	62	52	66, 102, 150, 161	0
All	All	1721/1981 (86%)	0.11	45 (2%)	56	46	30, 91, 138, 289	0

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
6	F	1	MET	6.9
1	A	51	VAL	6.1
7	G	35	ALA	5.2
2	B	87	HIS	4.8
3	C	317	ALA	4.7

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

There are no such residues in this entry.