wwPDB X-ray Structure Validation Summary Report (i)

Feb 14, 2024 – 08:10 AM EST

PDB ID	:	3MKQ
Title	:	Crystal structure of yeast alpha/betaprime-COP subcomplex of the COPI vesicular coat
Authors	:	Lee, C.; Goldberg, J.
Deposited on	:	2010-04-15
Resolution	:	2.50 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 (1)) were used in the production of this report:

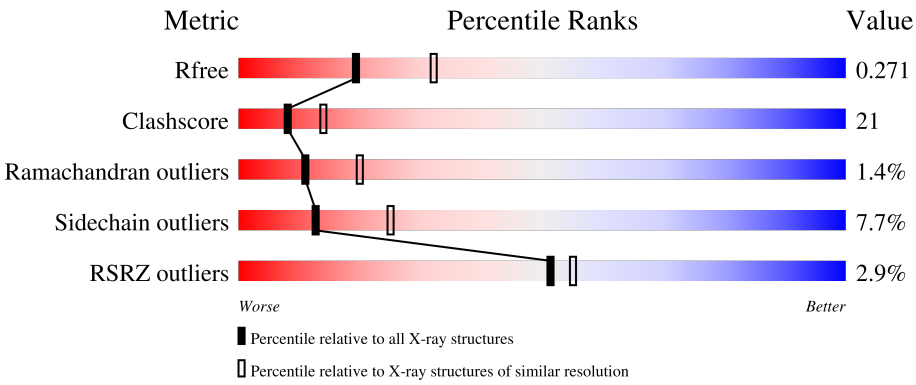
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	814	2%, 64%, 31%, 5%
1	C	814	3%, 60%, 35%, 5%
1	E	814	2%, 59%, 35%, 5%
2	B	177	2%, 68%, 27%, 5%
2	D	177	5%, 63%, 30%, 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	177	6% 62% 34%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 23730 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 Coatomer beta'-subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	811	6438	4110	1063	1252	13	0	0	0
1	C	811	6437	4109	1063	1252	13	0	0	0
1	E	810	6430	4107	1061	1248	14	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	701	ILE	VAL	variant	UNP A6ZU46
C	701	ILE	VAL	variant	UNP A6ZU46
E	701	ILE	VAL	variant	UNP A6ZU46

- Molecule 2 is a protein called Coatomer subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	177	1390	877	230	279	4	0	0	0
2	D	177	1390	877	230	279	4	0	0	0
2	F	175	1375	869	228	274	4	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	664	ALA	VAL	variant	UNP P53622
B	676	ILE	SER	variant	UNP P53622
D	664	ALA	VAL	variant	UNP P53622
D	676	ILE	SER	variant	UNP P53622
F	664	ALA	VAL	variant	UNP P53622

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	676	ILE	SER	variant	UNP P53622

- Molecule 3 is water.

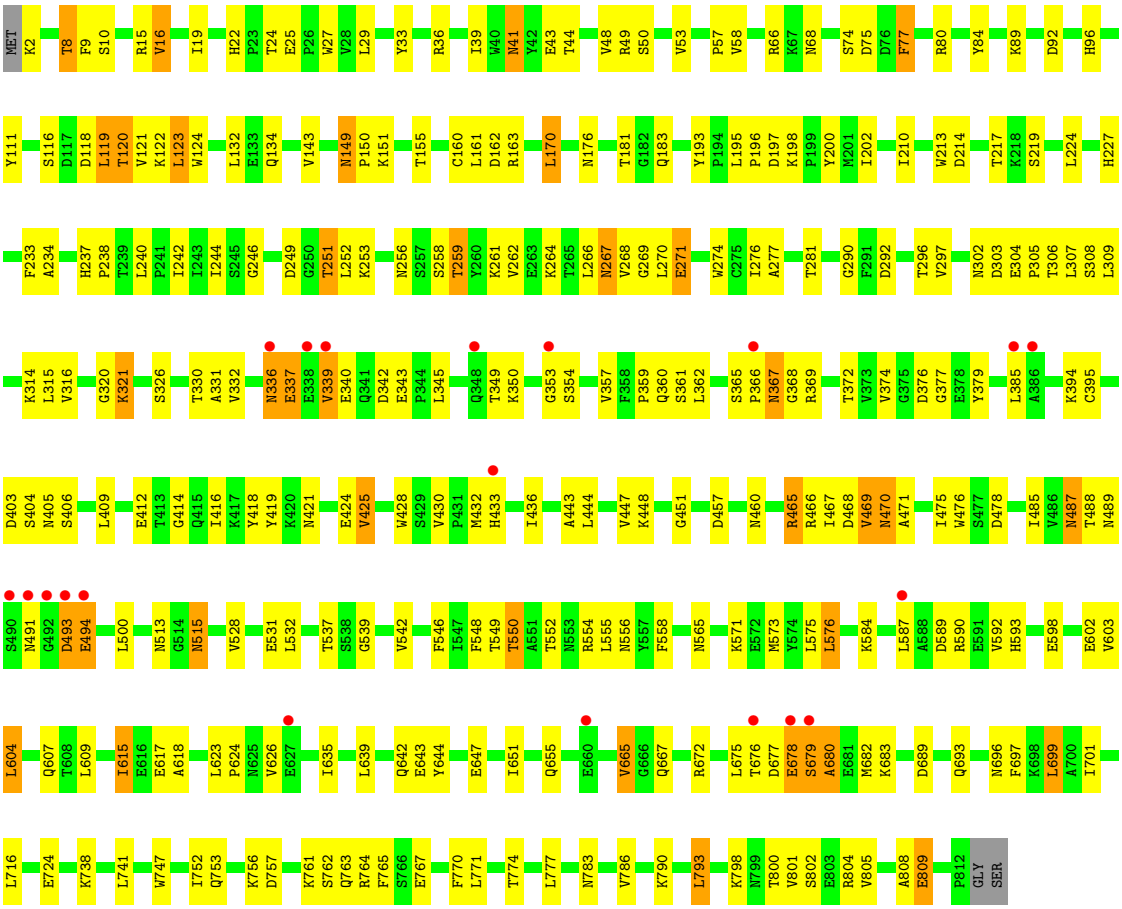
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	100	Total O 100 100	0	0
3	B	17	Total O 17 17	0	0
3	C	76	Total O 76 76	0	0
3	D	5	Total O 5 5	0	0
3	E	66	Total O 66 66	0	0
3	F	6	Total O 6 6	0	0

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.

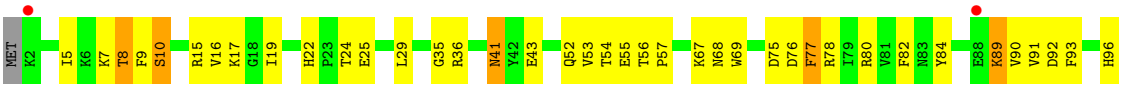
- Molecule 1: Coatomer beta'-subunit

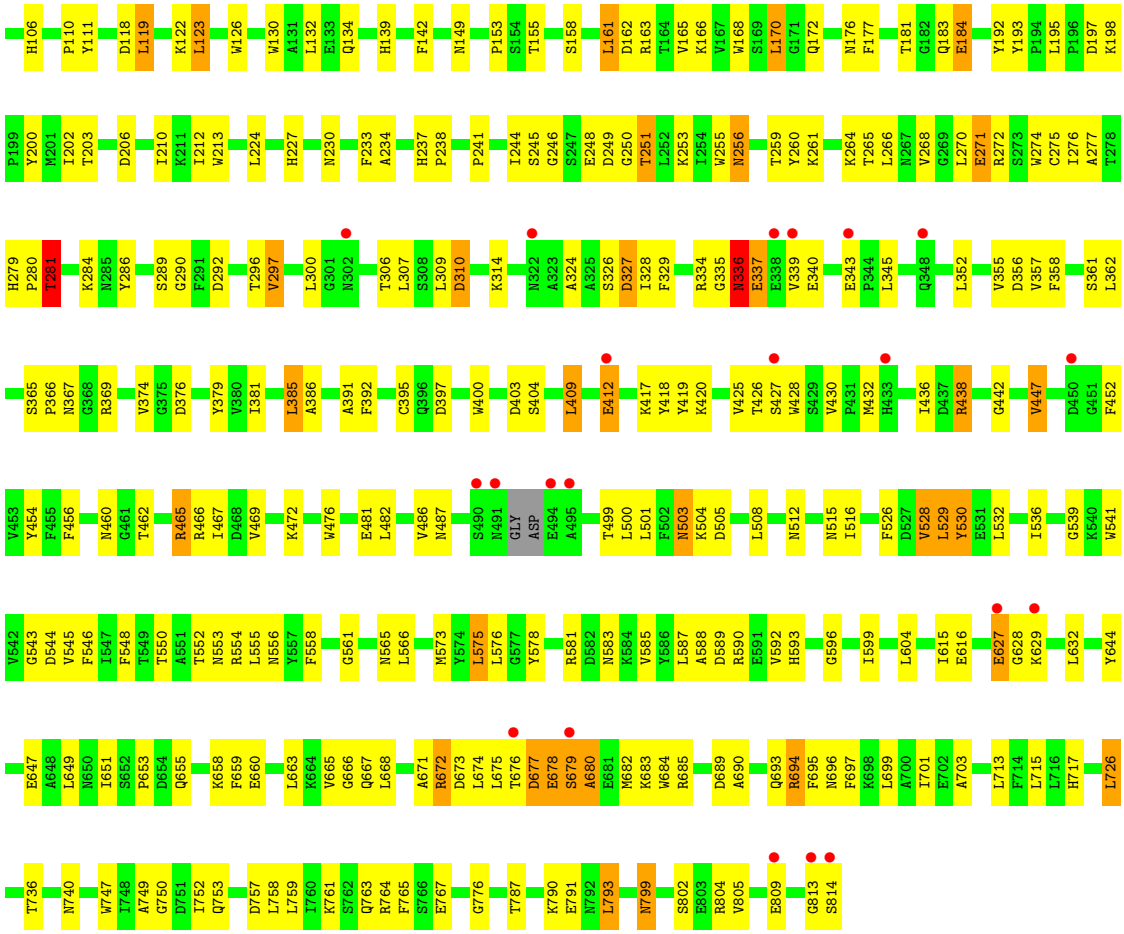
Chain A: 2% 64% 31% 5%



- Molecule 1: Coatomer beta'-subunit

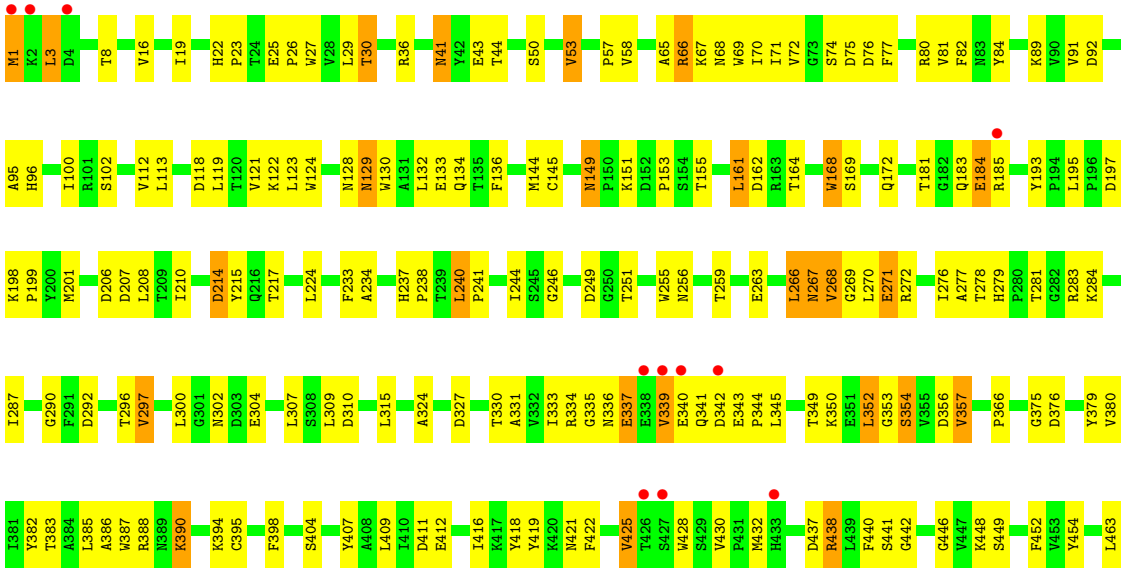
Chain C: 3% 60% 35% 5%

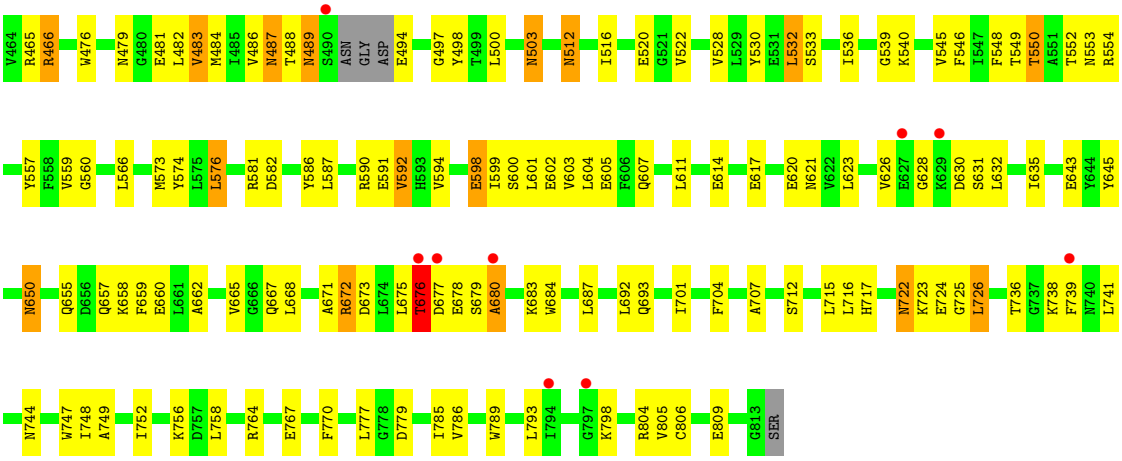




- Molecule 1: Coatomer beta'-subunit

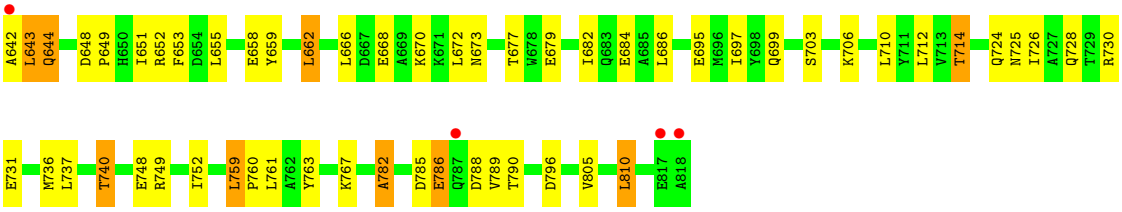
Chain E:





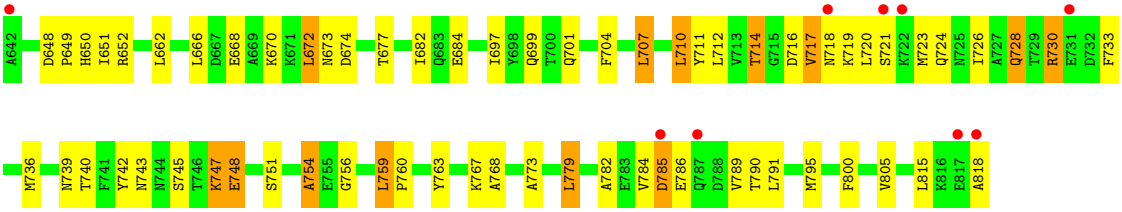
- Molecule 2: Coatomer subunit alpha

Chain B:



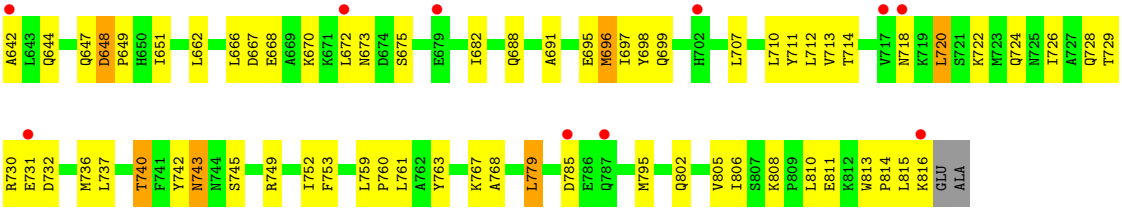
- Molecule 2: Coatomer subunit alpha

Chain D:



- Molecule 2: Coatomer subunit alpha

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	152.64Å 152.64Å 294.54Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	24.98 – 2.50 24.98 – 2.50	Depositor EDS
% Data completeness (in resolution range)	90.9 (24.98-2.50) 91.0 (24.98-2.50)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	2.67 (at 2.50Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.221 , 0.272 0.219 , 0.271	Depositor DCC
R_{free} test set	6381 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å^2)	36.3	Xtriage
Anisotropy	0.693	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 57.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.012 for -h,-k,l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	23730	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.16% 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	0.41	0/6591	0.68	0/8960
1	C	0.39	0/6589	0.66	0/8954
1	E	0.39	0/6582	0.65	0/8945
2	B	0.37	0/1415	0.57	1/1915 (0.1%)
2	D	0.35	0/1415	0.56	0/1915
2	F	0.36	0/1400	0.55	1/1896 (0.1%)
All	All	0.39	0/23992	0.64	2/32585 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	782	ALA	N-CA-C	-5.43	96.33	111.00
2	F	647	GLN	N-CA-C	5.21	125.06	111.00

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	6438	0	6225	276	0
1	C	6437	0	6225	288	0
1	E	6430	0	6226	273	0
2	B	1390	0	1349	49	0
2	D	1390	0	1349	62	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	1375	0	1338	68	0
3	A	100	0	0	3	0
3	B	17	0	0	0	0
3	C	76	0	0	5	0
3	D	5	0	0	0	0
3	E	66	0	0	2	0
3	F	6	0	0	0	0
All	All	23730	0	22712	985	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 21.

The worst 5 of 985 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:550:THR:HG22	1:A:554:ARG:H	1.11	1.12
1:C:334:ARG:HA	1:C:581:ARG:HH22	1.06	1.10
1:A:615:ILE:HD12	1:A:615:ILE:H	1.22	1.04
1:A:111:TYR:HB3	1:A:123:LEU:HD11	1.41	1.03
2:B:810:LEU:HD23	2:B:810:LEU:H	1.23	1.03

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	809/814 (99%)	755 (93%)	43 (5%)	11 (1%)	11 20
1	C	807/814 (99%)	741 (92%)	55 (7%)	11 (1%)	11 20
1	E	806/814 (99%)	749 (93%)	43 (5%)	14 (2%)	9 16
2	B	175/177 (99%)	168 (96%)	7 (4%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	D	175/177 (99%)	153 (87%)	17 (10%)	5 (3%)	4 6
2	F	173/177 (98%)	166 (96%)	7 (4%)	0	100 100
All	All	2945/2973 (99%)	2732 (93%)	172 (6%)	41 (1%)	11 20

5 of 41 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	339	VAL
1	A	493	ASP
1	A	494	GLU
1	C	336	ASN
1	C	679	SER

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	698/700 (100%)	647 (93%)	51 (7%)	14 27
1	C	698/700 (100%)	643 (92%)	55 (8%)	12 24
1	E	697/700 (100%)	641 (92%)	56 (8%)	12 23
2	B	148/148 (100%)	136 (92%)	12 (8%)	11 23
2	D	148/148 (100%)	138 (93%)	10 (7%)	16 30
2	F	147/148 (99%)	136 (92%)	11 (8%)	13 26
All	All	2536/2544 (100%)	2341 (92%)	195 (8%)	13 25

5 of 195 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	759	LEU
1	E	267	ASN
2	D	707	LEU
1	E	30	THR
1	E	357	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 101 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	644	GLN
1	E	149	ASN
2	F	739	ASN
2	D	688	GLN
2	D	744	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 (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(Å^2)	Q<0.9
1	A	811/814 (99%)	-0.00	20 (2%)	57	61	17, 39, 62, 84	0
1	C	811/814 (99%)	0.08	23 (2%)	53	56	20, 43, 66, 89	0
1	E	810/814 (99%)	0.07	20 (2%)	57	61	19, 42, 67, 89	0
2	B	177/177 (100%)	0.12	4 (2%)	60	63	29, 44, 68, 74	0
2	D	177/177 (100%)	0.57	9 (5%)	28	29	32, 55, 77, 83	0
2	F	175/177 (98%)	0.34	10 (5%)	23	25	30, 53, 73, 80	0
All	All	2961/2973 (99%)	0.10	86 (2%)	51	55	17, 43, 68, 89	0

The worst 5 of 86 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	818	ALA	9.0
1	C	814	SER	7.3
2	B	818	ALA	6.8
1	C	679	SER	6.0
1	A	676	THR	5.6

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.