



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 01:56 am BST

PDB ID : 4Q9U
Title : Crystal structure of the Rab5, Rabex-5delta and Rabaptin-5C21 complex
Authors : Zhang, Z.; Zhang, T.; Ding, J.
Deposited on : 2014-05-01
Resolution : 4.62 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 ⓘ symbol.

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 : 2.11
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

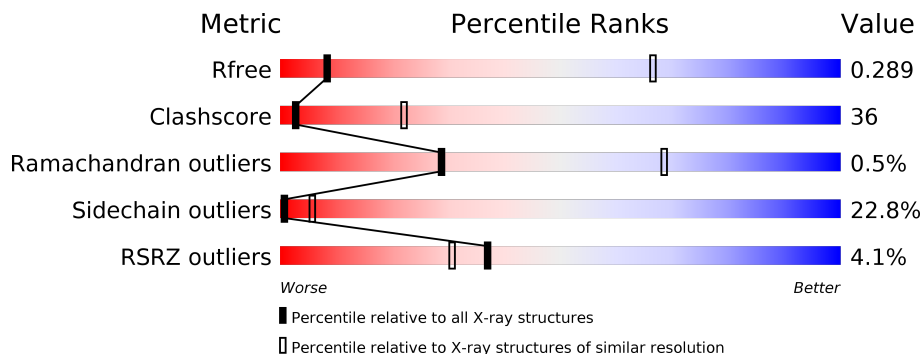
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.62 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	1062 (5.40-3.80)
Clashscore	141614	1130 (5.40-3.80)
Ramachandran outliers	138981	1074 (5.40-3.80)
Sidechain outliers	138945	1055 (5.40-3.80)
RSRZ outliers	127900	1114 (5.54-3.70)

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 $\leq 5\%$. 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	317	
1	E	317	
2	B	171	
2	F	171	
3	C	92	
3	D	92	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	92	<p>%</p> <p>32% 47% 13% 9%</p>
3	H	92	<p>24% 49% 16% 11%</p>

2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 9679 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 Rab5 GDP/GTP exchange factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	303	Total	C	N	O	S	0	0	0
			2492	1573	426	474	19			
1	E	296	Total	C	N	O	S	0	0	0
			2441	1541	419	462	19			

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	124	MET	-	EXPRESSION TAG	UNP Q9UJ41
A	125	GLY	-	EXPRESSION TAG	UNP Q9UJ41
A	126	HIS	-	EXPRESSION TAG	UNP Q9UJ41
A	127	HIS	-	EXPRESSION TAG	UNP Q9UJ41
A	128	HIS	-	EXPRESSION TAG	UNP Q9UJ41
A	129	HIS	-	EXPRESSION TAG	UNP Q9UJ41
A	130	HIS	-	EXPRESSION TAG	UNP Q9UJ41
A	131	HIS	-	EXPRESSION TAG	UNP Q9UJ41
A	?	-	LYS	DELETION	UNP Q9UJ41
A	?	-	GLN	DELETION	UNP Q9UJ41
A	?	-	GLU	DELETION	UNP Q9UJ41
A	?	-	ALA	DELETION	UNP Q9UJ41
A	?	-	GLU	DELETION	UNP Q9UJ41
A	?	-	SER	DELETION	UNP Q9UJ41
A	?	-	TRP	DELETION	UNP Q9UJ41
A	?	-	SER	DELETION	UNP Q9UJ41
A	?	-	PRO	DELETION	UNP Q9UJ41
A	?	-	ASP	DELETION	UNP Q9UJ41
A	?	-	ALA	DELETION	UNP Q9UJ41
A	?	-	CYS	DELETION	UNP Q9UJ41
A	?	-	LEU	DELETION	UNP Q9UJ41
A	?	-	GLY	DELETION	UNP Q9UJ41
A	?	-	VAL	DELETION	UNP Q9UJ41
E	124	MET	-	EXPRESSION TAG	UNP Q9UJ41
E	125	GLY	-	EXPRESSION TAG	UNP Q9UJ41

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	126	HIS	-	EXPRESSION TAG	UNP Q9UJ41
E	127	HIS	-	EXPRESSION TAG	UNP Q9UJ41
E	128	HIS	-	EXPRESSION TAG	UNP Q9UJ41
E	129	HIS	-	EXPRESSION TAG	UNP Q9UJ41
E	130	HIS	-	EXPRESSION TAG	UNP Q9UJ41
E	131	HIS	-	EXPRESSION TAG	UNP Q9UJ41
E	?	-	LYS	DELETION	UNP Q9UJ41
E	?	-	GLN	DELETION	UNP Q9UJ41
E	?	-	GLU	DELETION	UNP Q9UJ41
E	?	-	ALA	DELETION	UNP Q9UJ41
E	?	-	GLU	DELETION	UNP Q9UJ41
E	?	-	SER	DELETION	UNP Q9UJ41
E	?	-	TRP	DELETION	UNP Q9UJ41
E	?	-	SER	DELETION	UNP Q9UJ41
E	?	-	PRO	DELETION	UNP Q9UJ41
E	?	-	ASP	DELETION	UNP Q9UJ41
E	?	-	ALA	DELETION	UNP Q9UJ41
E	?	-	CYS	DELETION	UNP Q9UJ41
E	?	-	LEU	DELETION	UNP Q9UJ41
E	?	-	GLY	DELETION	UNP Q9UJ41
E	?	-	VAL	DELETION	UNP Q9UJ41

- Molecule 2 is a protein called Ras-related protein Rab-5A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	153	Total	C	N	O	S	0	0	0
			1194	759	201	229	5			
2	F	120	Total	C	N	O	S	0	0	0
			948	607	160	177	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	14	SER	-	EXPRESSION TAG	UNP P20339
F	14	SER	-	EXPRESSION TAG	UNP P20339

- Molecule 3 is a protein called Rab GTPase-binding effector protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	83	Total	C	N	O	S	0	0	0
			667	406	119	138	4			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	75	Total	C	N	O	S	0	0	0
			601	369	106	122	4			
3	G	84	Total	C	N	O	S	0	0	0
			678	412	123	139	4			
3	H	82	Total	C	N	O	S	0	0	0
			658	401	118	135	4			

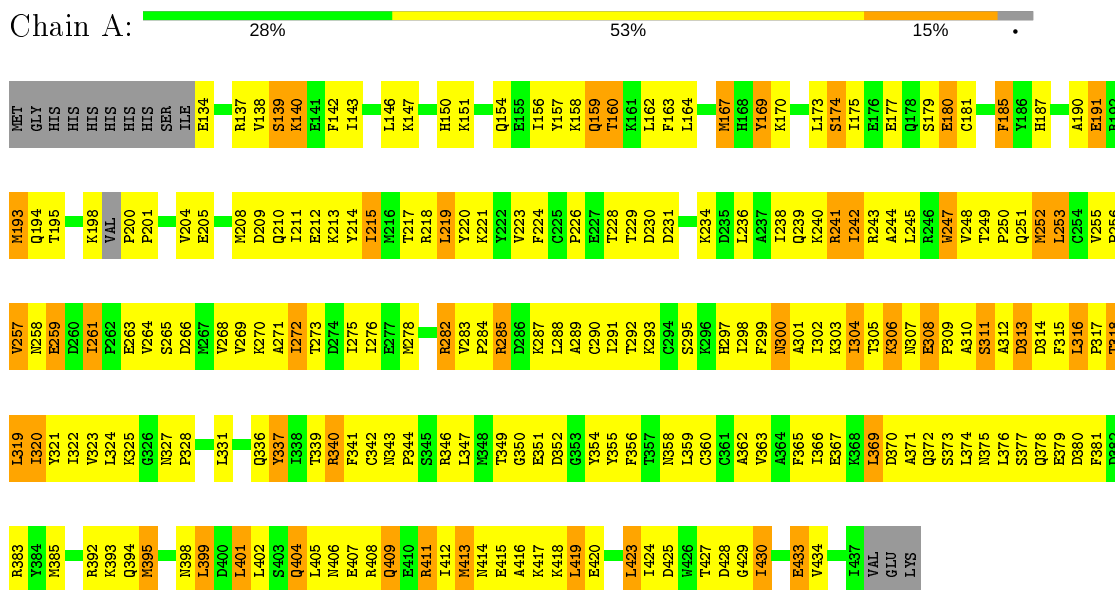
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	551	MET	-	EXPRESSION TAG	UNP Q15276
D	551	MET	-	EXPRESSION TAG	UNP Q15276
G	551	MET	-	EXPRESSION TAG	UNP Q15276
H	551	MET	-	EXPRESSION TAG	UNP Q15276

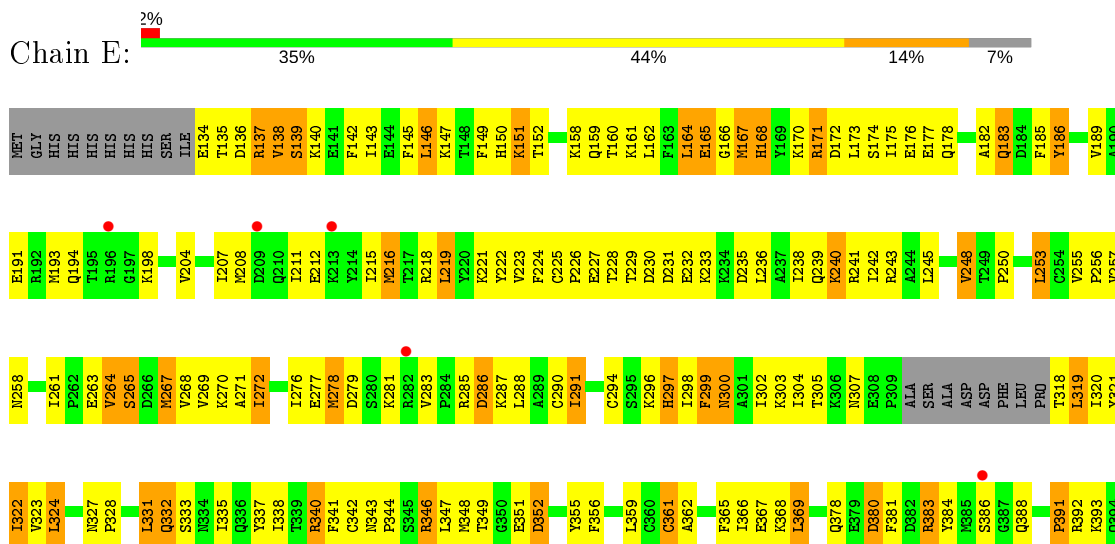
3 Residue-property plots

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: Rab5 GDP/GTP exchange factor

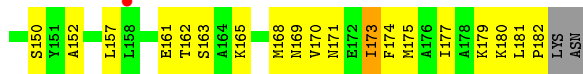
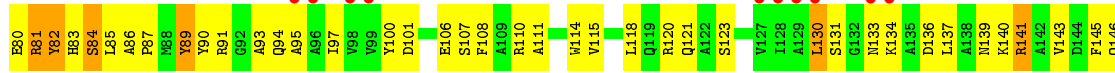
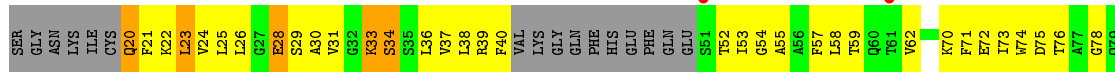


- Molecule 1: Rab5 GDP/GTP exchange factor

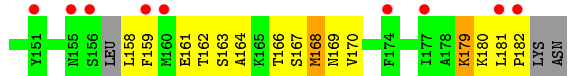
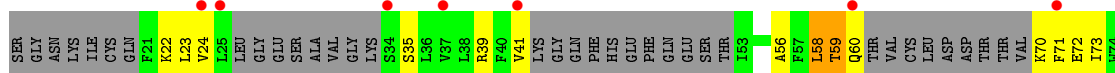




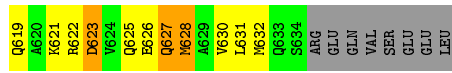
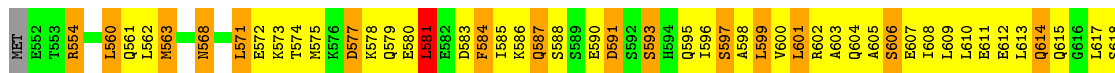
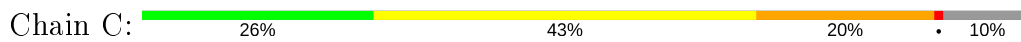
● Molecule 2: Ras-related protein Rab-5A



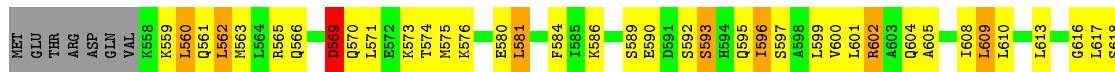
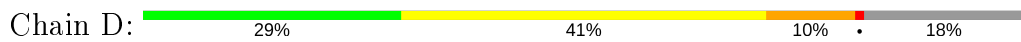
● Molecule 2: Ras-related protein Rab-5A



● Molecule 3: Rab GTPase-binding effector protein 1

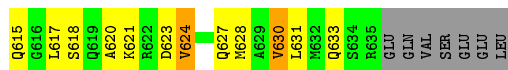


● Molecule 3: Rab GTPase-binding effector protein 1

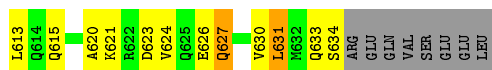
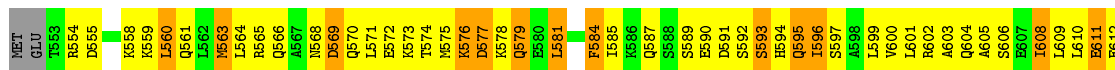




- Molecule 3: Rab GTPase-binding effector protein 1



- Molecule 3: Rab GTPase-binding effector protein 1



4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	174.82Å 174.82Å 149.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.78 – 4.62 48.49 – 4.62	Depositor EDS
% Data completeness (in resolution range)	95.9 (40.78-4.62) 95.9 (48.49-4.62)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.98 (at 4.64Å)	Xtrriage
Refinement program	PHENIX 1.8.1_1168	Depositor
R, R_{free}	0.251 , 0.343 0.258 , 0.289	Depositor DCC
R_{free} test set	621 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å ²)	172.7	Xtrriage
Anisotropy	0.423	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 208.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	9679	wwPDB-VP
Average B, all atoms (Å ²)	187.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.66% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.71	0/2535	0.92	2/3410 (0.1%)
1	E	0.60	0/2482	0.84	2/3338 (0.1%)
2	B	0.52	0/1213	0.71	0/1639
2	F	0.48	1/959 (0.1%)	0.60	0/1285
3	C	0.84	0/668	1.04	1/890 (0.1%)
3	D	0.82	0/602	1.09	1/801 (0.1%)
3	G	0.84	0/679	1.00	0/904
3	H	0.74	0/659	0.99	0/878
All	All	0.67	1/9797 (0.0%)	0.88	6/13145 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	86	ALA	CA-CB	5.09	1.63	1.52

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	581	LEU	CB-CG-CD1	-5.56	101.54	111.00
1	E	399	LEU	CB-CG-CD2	5.20	119.84	111.00
1	A	399	LEU	CB-CG-CD2	5.17	119.78	111.00
1	E	253	LEU	CA-CB-CG	5.16	127.17	115.30
1	A	308	GLU	N-CA-C	-5.14	97.11	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	2492	0	2502	226	0
1	E	2441	0	2460	185	0
2	B	1194	0	1192	81	0
2	F	948	0	937	36	0
3	C	667	0	678	61	0
3	D	601	0	618	50	0
3	G	678	0	691	60	0
3	H	658	0	672	70	0
All	All	9679	0	9750	692	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 36.

The worst 5 of 692 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:395:MET:HG3	1:A:399:LEU:HD23	1.45	0.99
1:A:250:PRO:HA	1:A:253:LEU:HG	1.49	0.92
1:A:238:ILE:HD13	1:A:331:LEU:HD23	1.53	0.89
1:E:258:ASN:H	1:E:305:THR:HG21	1.41	0.85
1:E:216:MET:HA	1:E:219:LEU:HD22	1.57	0.85

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	299/317 (94%)	279 (93%)	20 (7%)	0	100 100
1	E	292/317 (92%)	272 (93%)	18 (6%)	2 (1%)	22 62

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	149/171 (87%)	129 (87%)	17 (11%)	3 (2%)	7	40
2	F	104/171 (61%)	100 (96%)	3 (3%)	1 (1%)	15	54
3	C	81/92 (88%)	77 (95%)	4 (5%)	0	100	100
3	D	73/92 (79%)	73 (100%)	0	0	100	100
3	G	82/92 (89%)	80 (98%)	2 (2%)	0	100	100
3	H	80/92 (87%)	78 (98%)	2 (2%)	0	100	100
All	All	1160/1344 (86%)	1088 (94%)	66 (6%)	6 (0%)	29	68

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	170	VAL
2	B	30	ALA
2	B	34	SER
1	E	391	PRO
2	F	168	MET

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	281/294 (96%)	217 (77%)	64 (23%)	1	6
1	E	276/294 (94%)	217 (79%)	59 (21%)	1	6
2	B	127/143 (89%)	106 (84%)	21 (16%)	2	14
2	F	100/143 (70%)	87 (87%)	13 (13%)	4	20
3	C	76/85 (89%)	49 (64%)	27 (36%)	0	1
3	D	68/85 (80%)	48 (71%)	20 (29%)	0	2
3	G	77/85 (91%)	54 (70%)	23 (30%)	0	2
3	H	75/85 (88%)	56 (75%)	19 (25%)	0	4
All	All	1080/1214 (89%)	834 (77%)	246 (23%)	1	6

5 of 246 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	628	MET
1	E	151	LYS
3	H	560	LEU
3	D	569	ASP
3	D	609	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	133	ASN
3	H	627	GLN
1	E	388	GLN
1	A	409	GLN
1	E	300	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 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

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	303/317 (95%)	-0.37	0 100 100	45, 133, 270, 349	0
1	E	296/317 (93%)	-0.16	5 (1%) 70 61	39, 209, 344, 447	0
2	B	153/171 (89%)	0.29	13 (8%) 10 10	112, 238, 333, 441	0
2	F	120/171 (70%)	1.24	30 (25%) 0 1	137, 300, 407, 493	0
3	C	83/92 (90%)	-0.24	0 100 100	36, 120, 264, 425	0
3	D	75/92 (81%)	-0.43	0 100 100	71, 130, 262, 343	0
3	G	84/92 (91%)	-0.45	1 (1%) 79 70	54, 108, 252, 324	0
3	H	82/92 (89%)	-0.32	0 100 100	37, 116, 263, 327	0
All	All	1196/1344 (88%)	-0.07	49 (4%) 37 31	36, 178, 336, 493	0

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	130	LEU	6.1
2	B	128	ILE	4.9
2	F	98	VAL	4.8
2	F	76	THR	4.8
2	F	128	ILE	4.1

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.