W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Aug 27, 2023 – 09:34 AM EDT

PDB ID	:	3I6W
Title	:	Structure and Activation Mechanism of the CHK2 DNA-Damage Checkpoint Kinase
Authors	:	Pavletich, N.P.
Deposited on	:	2009-07-07
Resolution	:	3.25 Å(reported)

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

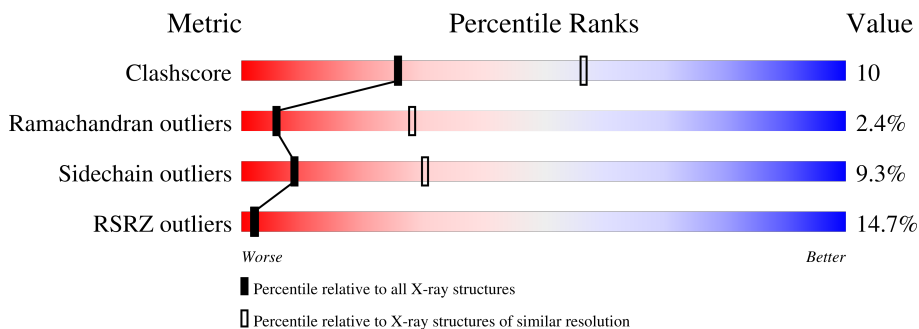|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1   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 3.25 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.
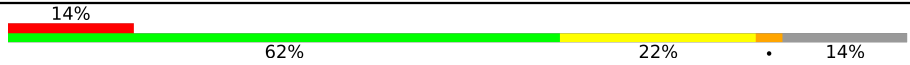
Metric	Percentile Ranks	Value
Clashscore		10
Ramachandran outliers		2.4%
Sidechain outliers		9.3%
RSRZ outliers		14.7%

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))
Clashscore	141614	1251 (3.30-3.22)
Ramachandran outliers	138981	1229 (3.30-3.22)
Sidechain outliers	138945	1228 (3.30-3.22)
RSRZ outliers	127900	1154 (3.30-3.22)

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	443	8% / 60% 19% • 17%
1	B	443	9% / 63% 21% • 14%
1	C	443	6% / 58% 21% • 17%
1	D	443	13% / 63% 21% • 14%
1	E	443	17% / 60% 20% • 17%
1	F	443	21% / 65% 18% • 14%
1	G	443	12% / 59% 20% • 17%

WORLDWIDE PDB PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	443	

14%

62%   22%   •   14%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 24560 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 Serine/threonine-protein kinase Chk2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	369	Total	C	N	O	S	0	0	0
			3022	1952	504	553	13			
1	B	382	Total	C	N	O	S	0	0	0
			3120	2012	523	571	14			
1	C	368	Total	C	N	O	S	0	0	0
			3018	1950	503	552	13			
1	D	382	Total	C	N	O	S	0	0	0
			3120	2012	523	571	14			
1	E	368	Total	C	N	O	S	0	0	0
			3018	1950	503	552	13			
1	F	382	Total	C	N	O	S	0	0	0
			3120	2012	523	571	14			
1	G	369	Total	C	N	O	S	0	0	0
			3022	1952	504	553	13			
1	H	382	Total	C	N	O	S	0	0	0
			3120	2012	523	571	14			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	249	ARG	LYS	engineered mutation	UNP O96017
B	249	ARG	LYS	engineered mutation	UNP O96017
C	249	ARG	LYS	engineered mutation	UNP O96017
D	249	ARG	LYS	engineered mutation	UNP O96017
E	249	ARG	LYS	engineered mutation	UNP O96017
F	249	ARG	LYS	engineered mutation	UNP O96017
G	249	ARG	LYS	engineered mutation	UNP O96017
H	249	ARG	LYS	engineered mutation	UNP O96017

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: Serine/threonine-protein kinase Chk2

- Molecule 1: Serine/threonine-protein kinase Chk2

L460 D461 L462 V463 K464 K465 P471 K472 F475 T476 T477 E478 R482 L486 Q487 D488 E489 D490 M491 K492 D497 E501 GLU ASN GLU SER THR ALA LEU PRO GLN VAL LEU

• Molecule 1: Serine/threonine-protein kinase Chk2

Chain C:

6%

58% 21% 17%

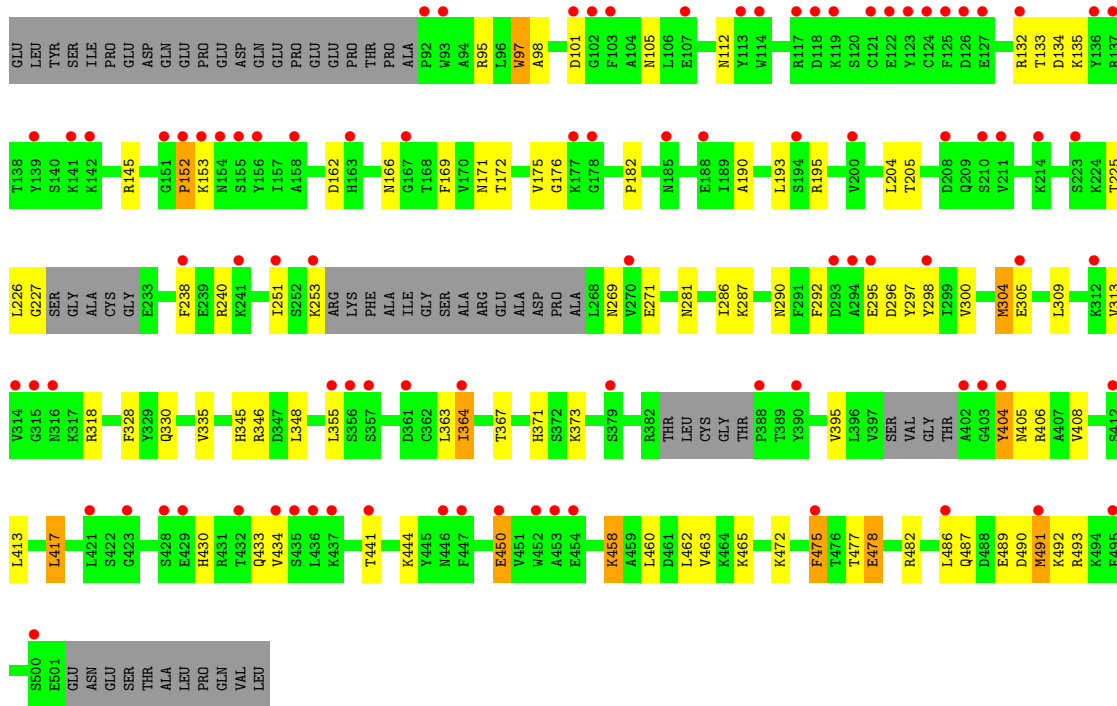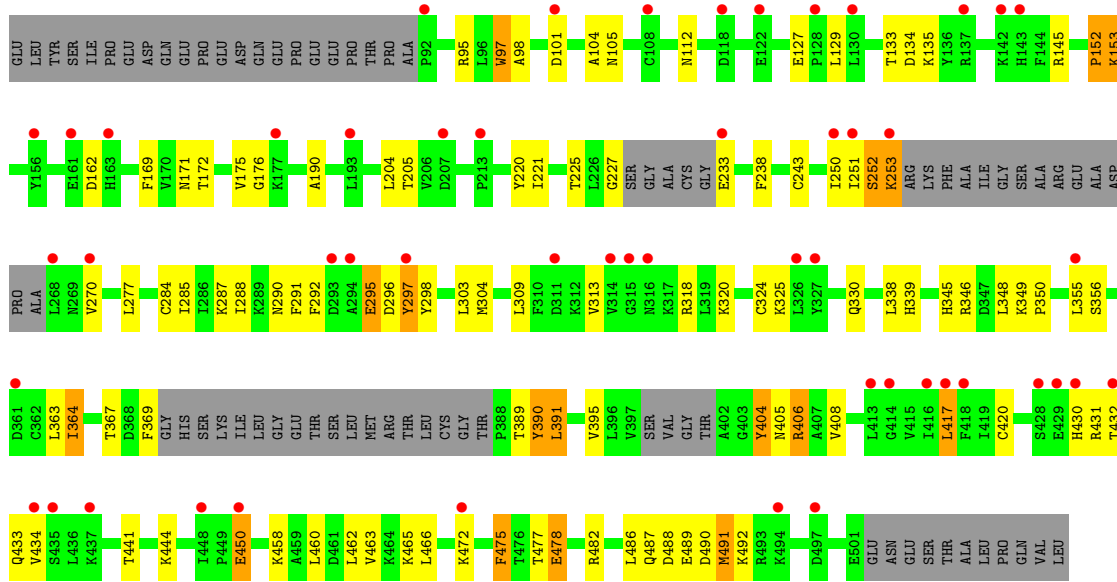GLU LEU TYR SER ILE PRO GLU ASP GLN GLU PRO GLU ASP GLN GLU PRO GLU GLU PRO THR PRO ALA P92 R95 L96 W97 A98 N105 N112 E127 P128 L129 T133 D134 K135 T138 H143 F144 R145 R148 P152 K153 D162 H163 S164 G165 N166 F169 V170

N171 T172 V175 G176 P182 A190 R195 L204 T205 V206 D207 K214 A215 S223 L226 GLY SER GLY ALA CYS GLY E233 A237 K241 V246 I250 I251 S252 K253 ARG LYS PHE ILE GLY SER ALA ARG GLU ALA ASP PRO ALA L268 N269 V270 E271 T272

E273 I274 E275 I276 L277 H282 P283 C284 I285 I286 K287 I288 K289 N290 F291 F292 D293 A294 E295 D296 Y297 Y298 L303 M304 L309 V313 N316 K317 R318 Y329 Q330 A334 V335 Q336 Y337 L338 H345 R346 D347 L348 K349 P350 L355 S356 L363 I364 K365 I366 T367 D368

F369 GLY HIS SER LYS ILE LEU LEU GLY GLU THR SER LEU MET ARG THR LEU CYS GLY THR P388 T389 Y390 L391 V395 L396 V397 SER VAL GLY THR A402 G403 Y404 N405 R406 A407 V408 L413 I416 L417 E429 H430 R431 T432 Q433 V434 S435 L436 T441 K444 E450 E457

K458 A459 L460 D461 L462 V463 K464 K465 K472 F475 T476 T477 E478 R482 L486 Q487 D488 E489 D490 M491 K492 Q496 E501 GLU ASN GLU SER THR ALA LEU PRO GLN VAL LEU

• Molecule 1: Serine/threonine-protein kinase Chk2

Chain D:

13%

63% 21% 14%

GLU LEU TYR SER ILE PRO GLU ASP GLN GLU PRO GLU ASP GLN GLU PRO GLU GLU PRO THR PRO ALA P92 R95 L96 W97 A98 D101 G102 F103 A104 N105 N112 Y113 W114 R117 D118 K119 S120 C121 D126 L129 L130 R131 R132 T133 D134 K135 Y136 K142 R145

V150 G151 P152 K153 N154 S155 Y156 D162 H163 F169 V170 N171 T172 V175 G176 K177 N184 N185 M186 S187 E188 I189 A190 L193 F202 D203 L204 T205 Q209 S210 V211 R217 M222 T225 L226 G227 SER GLY ALA CYS GLY E233 V234 F238 E239 K245 I251

S252 K253 ARG LYS PHE ALA ILE GLY SER ALA ARG GLU ALA ASP PRO ALA L268 N269 V270 E271 L277 L280 N281 C284 I285 I286 K287 I288 F291 F292 E295 D296 Y297 Y298 M304 L309 V313 V314 G315 N316 K317 R318 C324 Q330 A334 V335 I343 I344

H345 R346 D347 L348 K349 P350 L355 S356 S357 C362 L363 I364 K365 I366 T367 H371 S372 K373 R382 THR LEU CYS GLY THR P388 A392 P393 E394 V395 L396 V397 SER VAL GLY THR A402 G403 Y404 N405 R406 A407 V408 L413 I416 L417 C420 S428 E429 H430 R431 T432

Q433 V434 S435 L436 T441 K444 Y445 N446 E450 E454 E457 K458 A459 L460 D461 L462 V463 K464 K465 K472 F475 T476 T477 E478 R482 L486 Q487 D488 E489 D490 M491 K492 F495 E501 GLU ASN GLU SER THR ALA LEU PRO GLN VAL LEU

• Molecule 1: Serine/threonine-protein kinase Chk2

Chain E:

17%

60% 20% 17%

• Molecule 1: Serine/threonine-protein kinase Chk2

Chain F:

• Molecule 1: Serine/threonine-protein kinase Chk2

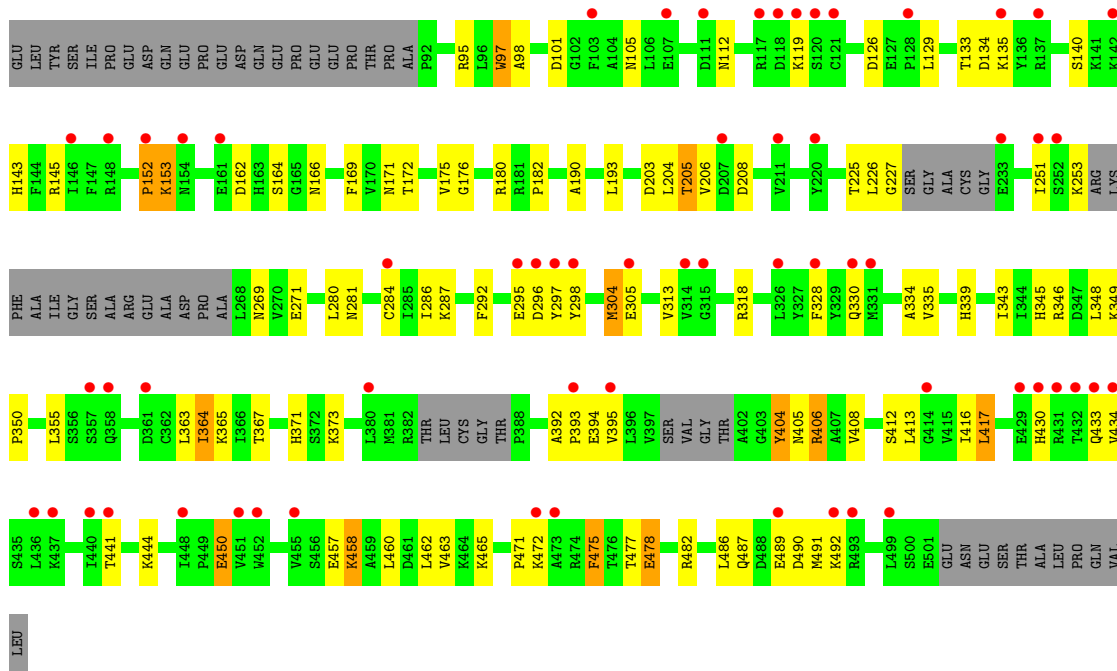Chain G:

• Molecule 1: Serine/threonine-protein kinase Chk2

Chain H:

14%
62% 22% · 14%

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	76.20Å 114.70Å 123.00Å 84.10° 81.20° 80.70°	Depositor
Resolution (Å)	30.00 – 3.25 29.86 – 3.25	Depositor EDS
% Data completeness (in resolution range)	90.8 (30.00-3.25) 90.9 (29.86-3.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	1.59 (at 3.24Å)	Xtriage
Refinement program	REFMAC 5.3.0036	Depositor
R, R_{free}	0.251 , 0.287 0.240 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	119.5	Xtriage
Anisotropy	0.185	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 172.8	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	24560	wwPDB-VP
Average B, all atoms (Å^2)	137.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.71% 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	0.48	0/3087	0.59	0/4163
1	B	0.49	0/3186	0.60	1/4294 (0.0%)
1	C	0.55	0/3083	0.63	1/4158 (0.0%)
1	D	0.46	0/3186	0.60	1/4294 (0.0%)
1	E	0.44	0/3083	0.58	0/4158
1	F	0.41	0/3186	0.57	1/4294 (0.0%)
1	G	0.43	0/3087	0.58	1/4163 (0.0%)
1	H	0.40	0/3186	0.57	0/4294
All	All	0.46	0/25084	0.59	5/33818 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	309	LEU	CA-CB-CG	5.17	127.20	115.30
1	G	309	LEU	CA-CB-CG	5.12	127.07	115.30
1	C	309	LEU	CA-CB-CG	5.09	127.01	115.30
1	F	309	LEU	CA-CB-CG	5.05	126.92	115.30
1	D	309	LEU	CA-CB-CG	5.04	126.89	115.30

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3022	0	3026	80	6
1	B	3120	0	3130	79	0
1	C	3018	0	3023	74	5
1	D	3120	0	3130	71	0
1	E	3018	0	3023	81	6
1	F	3120	0	3130	69	8
1	G	3022	0	3026	69	5
1	H	3120	0	3130	84	4
All	All	24560	0	24618	516	20

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 10.

The worst 5 of 516 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:129:LEU:HD23	1:D:193:LEU:CD2	1.47	1.44
1:A:129:LEU:CD2	1:D:193:LEU:HD21	1.52	1.40
1:F:132:ARG:CD	1:H:129:LEU:HD13	1.74	1.16
1:B:166:ASN:HD22	1:C:129:LEU:HB2	1.15	1.09
1:F:132:ARG:HD2	1:H:129:LEU:CD1	1.84	1.07

The worst 5 of 20 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)
1:G:320:LYS:CD	1:H:457:GLU:OE2[1_455]	0.80	1.40
1:C:138:THR:CG2	1:C:496:GLN:NE2[1_655]	1.64	0.56
1:E:137:ARG:CD	1:F:450:GLU:CG[1_455]	1.64	0.56
1:G:320:LYS:CE	1:H:457:GLU:OE2[1_455]	1.71	0.49
1:A:195:ARG:O	1:F:493:ARG:NH1[1_456]	1.73	0.47

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	359/443 (81%)	326 (91%)	23 (6%)	10 (3%)	5	25
1	B	372/443 (84%)	337 (91%)	28 (8%)	7 (2%)	8	34
1	C	358/443 (81%)	326 (91%)	21 (6%)	11 (3%)	4	24
1	D	372/443 (84%)	334 (90%)	31 (8%)	7 (2%)	8	34
1	E	358/443 (81%)	327 (91%)	21 (6%)	10 (3%)	5	25
1	F	372/443 (84%)	337 (91%)	28 (8%)	7 (2%)	8	34
1	G	359/443 (81%)	328 (91%)	21 (6%)	10 (3%)	5	25
1	H	372/443 (84%)	335 (90%)	30 (8%)	7 (2%)	8	34
All	All	2922/3544 (82%)	2650 (91%)	203 (7%)	69 (2%)	6	28

5 of 69 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	134	ASP
1	A	390	TYR
1	A	433	GLN
1	B	134	ASP
1	B	433	GLN

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	333/395 (84%)	303 (91%)	30 (9%)	9	32
1	B	344/395 (87%)	311 (90%)	33 (10%)	8	29
1	C	333/395 (84%)	302 (91%)	31 (9%)	9	30
1	D	344/395 (87%)	312 (91%)	32 (9%)	9	30
1	E	333/395 (84%)	303 (91%)	30 (9%)	9	32
1	F	344/395 (87%)	312 (91%)	32 (9%)	9	30
1	G	333/395 (84%)	302 (91%)	31 (9%)	9	30

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	H	344/395 (87%)	312 (91%)	32 (9%)	9	30
All	All	2708/3160 (86%)	2457 (91%)	251 (9%)	9	30

5 of 251 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	458	LYS
1	H	153	LYS
1	E	444	LYS
1	H	97	TRP
1	H	417	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 54 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	105	ASN
1	F	112	ASN
1	H	112	ASN
1	E	112	ASN
1	E	358	GLN

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	369/443 (83%)	0.65	35 (9%)	8	9	118, 136, 157, 175	0
1	B	382/443 (86%)	0.58	40 (10%)	6	6	115, 136, 154, 182	0
1	C	368/443 (83%)	0.56	28 (7%)	13	12	119, 134, 158, 178	0
1	D	382/443 (86%)	0.72	56 (14%)	2	2	118, 136, 154, 182	0
1	E	368/443 (83%)	1.15	77 (20%)	1	1	118, 136, 157, 180	0
1	F	382/443 (86%)	1.07	91 (23%)	0	0	118, 137, 153, 182	0
1	G	369/443 (83%)	0.96	51 (13%)	2	2	115, 137, 156, 178	0
1	H	382/443 (86%)	0.93	62 (16%)	1	2	119, 136, 153, 181	0
All	All	3002/3544 (84%)	0.82	440 (14%)	2	2	115, 136, 156, 182	0

The worst 5 of 440 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	152	PRO	10.2
1	A	434	VAL	9.6
1	G	434	VAL	8.7
1	E	432	THR	8.5
1	E	241	LYS	7.0

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 ⓘ

There are no such residues in this entry.