wwPDB X-ray Structure Validation Summary Report (i)

Oct 7, 2026 – 03:10 PM JST

PDB ID	:	44LS / pdb_000044ls
Title	:	Crystal structure of the fusion protein of T-lineage protein Themis and adaptor protein Grb2
Authors	:	Ouyang, S.; Zhao, Y.
Deposited on	:	2026-08-03
Resolution	:	3.06 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:

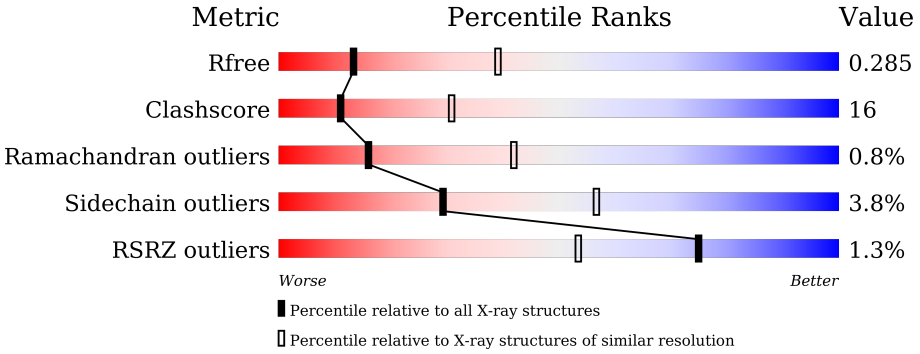
MolProbity	:	4-5-2 with Phenix2.2
Xtriage (Phenix)	:	2.2
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.51

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 3.06 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}	180053	2469 (3.10-3.02)
Clashscore	190562	2569 (3.10-3.02)
Ramachandran outliers	187476	2424 (3.10-3.02)
Sidechain outliers	187428	2423 (3.10-3.02)
RSRZ outliers	180081	2469 (3.10-3.02)

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	648	
1	B	648	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 8522 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 Protein THEMIS,Growth factor receptor-bound protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	521	4204	2702	706	777	19	0	0	0
1	A	527	4276	2754	714	788	20	0	0	0

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-23	MET	-	initiating methionine	UNP Q8N1K5
B	-22	HIS	-	expression tag	UNP Q8N1K5
B	-21	HIS	-	expression tag	UNP Q8N1K5
B	-20	HIS	-	expression tag	UNP Q8N1K5
B	-19	HIS	-	expression tag	UNP Q8N1K5
B	-18	HIS	-	expression tag	UNP Q8N1K5
B	-17	HIS	-	expression tag	UNP Q8N1K5
B	-16	SER	-	expression tag	UNP Q8N1K5
B	-15	SER	-	expression tag	UNP Q8N1K5
B	-14	GLY	-	expression tag	UNP Q8N1K5
B	-13	VAL	-	expression tag	UNP Q8N1K5
B	-12	ASP	-	expression tag	UNP Q8N1K5
B	-11	LEU	-	expression tag	UNP Q8N1K5
B	-10	GLY	-	expression tag	UNP Q8N1K5
B	-9	THR	-	expression tag	UNP Q8N1K5
B	-8	GLU	-	expression tag	UNP Q8N1K5
B	-7	ASN	-	expression tag	UNP Q8N1K5
B	-6	LEU	-	expression tag	UNP Q8N1K5
B	-5	TYR	-	expression tag	UNP Q8N1K5
B	-4	PHE	-	expression tag	UNP Q8N1K5
B	-3	GLN	-	expression tag	UNP Q8N1K5
B	-2	SER	-	expression tag	UNP Q8N1K5
B	-1	ASN	-	expression tag	UNP Q8N1K5
B	0	ALA	-	expression tag	UNP Q8N1K5
B	562	GLY	-	linker	UNP Q8N1K5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	563	GLY	-	linker	UNP Q8N1K5
B	564	SER	-	linker	UNP Q8N1K5
B	565	GLY	-	linker	UNP Q8N1K5
B	566	GLY	-	linker	UNP Q8N1K5
B	567	SER	-	linker	UNP Q8N1K5
A	-23	MET	-	initiating methionine	UNP Q8N1K5
A	-22	HIS	-	expression tag	UNP Q8N1K5
A	-21	HIS	-	expression tag	UNP Q8N1K5
A	-20	HIS	-	expression tag	UNP Q8N1K5
A	-19	HIS	-	expression tag	UNP Q8N1K5
A	-18	HIS	-	expression tag	UNP Q8N1K5
A	-17	HIS	-	expression tag	UNP Q8N1K5
A	-16	SER	-	expression tag	UNP Q8N1K5
A	-15	SER	-	expression tag	UNP Q8N1K5
A	-14	GLY	-	expression tag	UNP Q8N1K5
A	-13	VAL	-	expression tag	UNP Q8N1K5
A	-12	ASP	-	expression tag	UNP Q8N1K5
A	-11	LEU	-	expression tag	UNP Q8N1K5
A	-10	GLY	-	expression tag	UNP Q8N1K5
A	-9	THR	-	expression tag	UNP Q8N1K5
A	-8	GLU	-	expression tag	UNP Q8N1K5
A	-7	ASN	-	expression tag	UNP Q8N1K5
A	-6	LEU	-	expression tag	UNP Q8N1K5
A	-5	TYR	-	expression tag	UNP Q8N1K5
A	-4	PHE	-	expression tag	UNP Q8N1K5
A	-3	GLN	-	expression tag	UNP Q8N1K5
A	-2	SER	-	expression tag	UNP Q8N1K5
A	-1	ASN	-	expression tag	UNP Q8N1K5
A	0	ALA	-	expression tag	UNP Q8N1K5
A	562	GLY	-	linker	UNP Q8N1K5
A	563	GLY	-	linker	UNP Q8N1K5
A	564	SER	-	linker	UNP Q8N1K5
A	565	GLY	-	linker	UNP Q8N1K5
A	566	GLY	-	linker	UNP Q8N1K5
A	567	SER	-	linker	UNP Q8N1K5

- Molecule 2 is water.

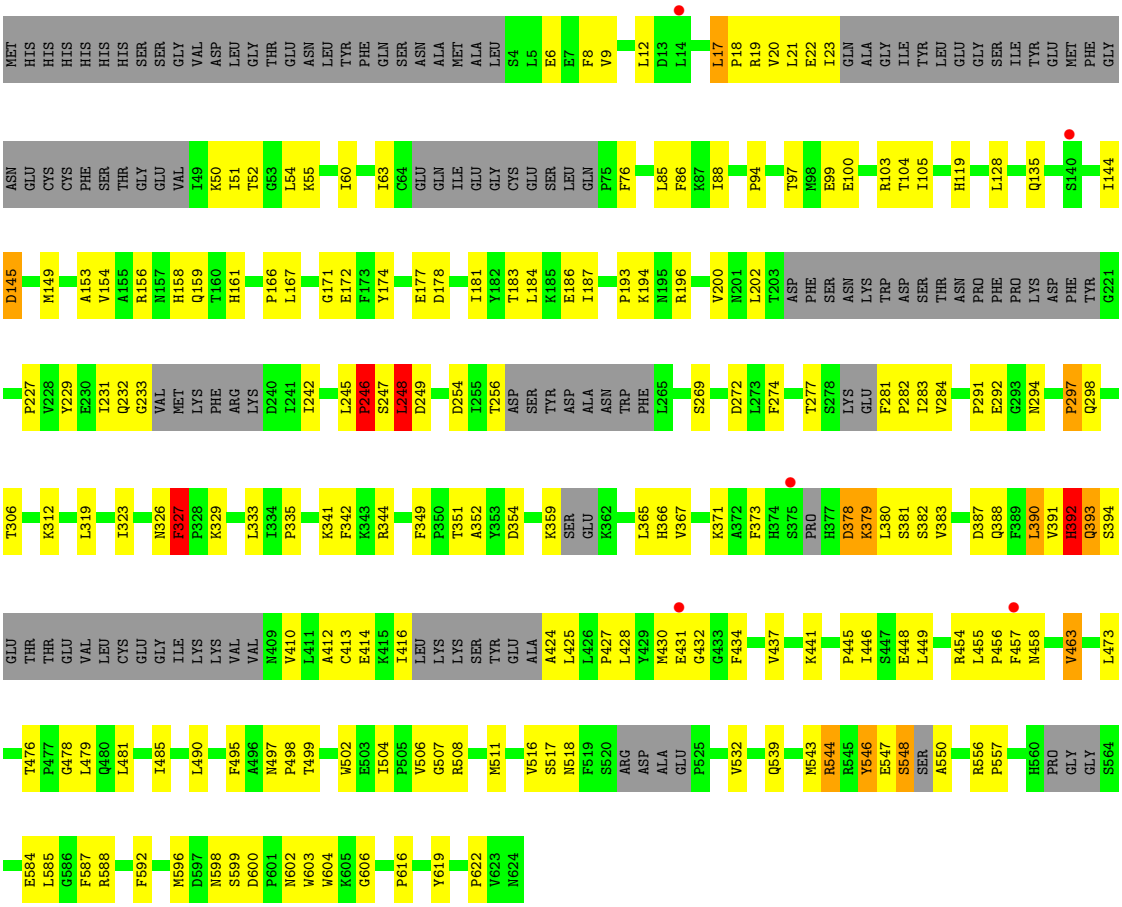
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	B	18	18	18	0	0
			Total	O		
2	A	24	24	24	0	0

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.

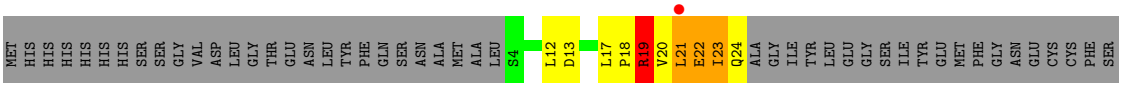
• Molecule 1: Protein THEMIS,Growth factor receptor-bound protein 2

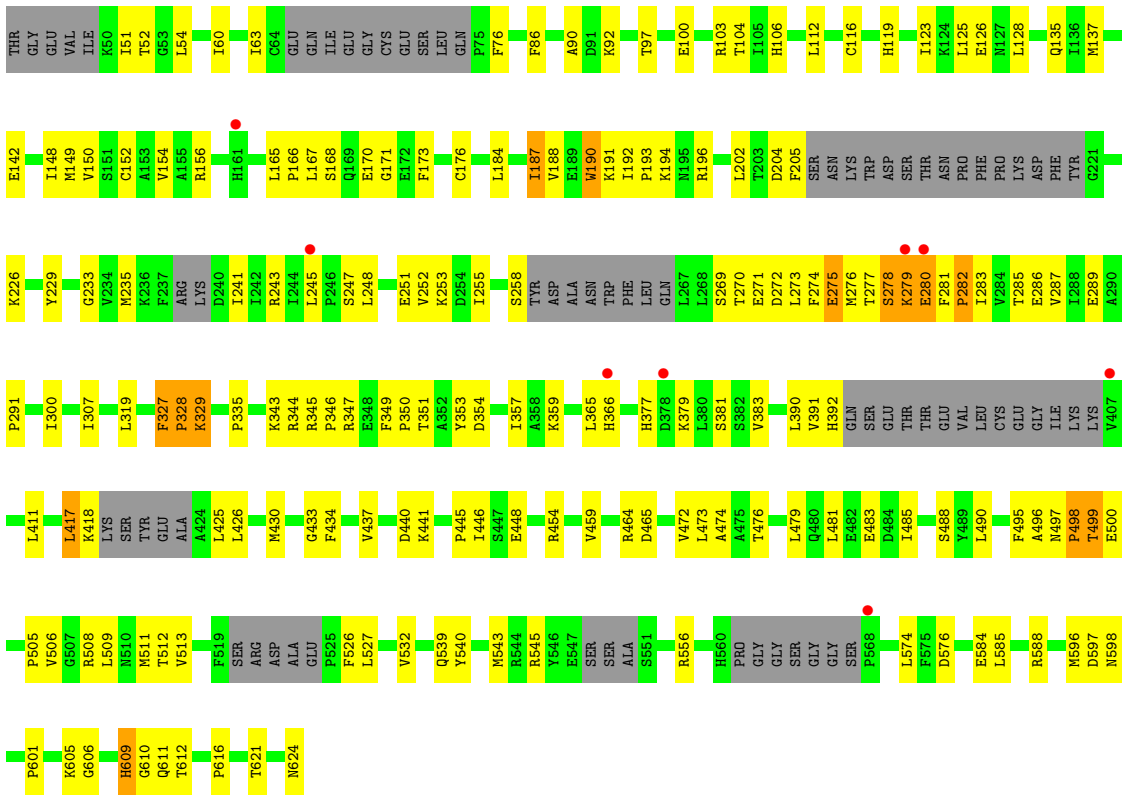
Chain B:



• Molecule 1: Protein THEMIS,Growth factor receptor-bound protein 2

Chain A:




THR GLY GLU VAL ILE K50 I51 T52 G53 L54 I60 I63 C64 GLU GLN ILE GLU GLY CYS GLU SER LEU GLN P75 F76 F86 A90 D91 K92 T97 E100 R103 T104 I105 H106 L112 C116 H119 I123 K124 L125 E126 N127 L128 Q135 I136 M137
E142 I148 M149 V150 S151 C152 A153 V154 A155 R156 H161 L165 P166 L167 S168 Q169 E170 G171 E172 F173 C176 L184 I187 V188 E189 W190 K191 I192 P193 K194 N195 R196 L202 T203 D204 F205 SER ASN LYS TRP ASP SER THR ASN PRO PHE PRO LYS ASP PHE TYR G221
K226 Y229 G233 V234 M235 K236 F237 ARG LYS D240 I241 I242 R243 I244 L245 P246 S247 L248 E251 V252 K253 D254 I255 S258 TYR ASP ALA ASN TRP PHE LEU GLN L267 L268 S269 T270 E271 D272 L273 F274 E275 M276 T277 S278 K279 E280 F281 P282 I283 V284 T285 E286 V287 I288 E289 A290
P291 I300 I307 L319 F327 P328 K329 P335 K343 R344 R345 P346 R347 E348 F349 P350 T351 A352 Y353 D354 I357 A358 K359 L365 H366 H377 D378 K379 L380 S381 S382 V383 L390 V391 H392 GLN SER GLU THR THR GLU VAL LEU CYS GLU GLY ILE LYS LYS V407
L411 L417 K418 LYS SER TYR GLU ALA A424 L425 L426 M430 G433 F434 V437 D440 K441 P445 I446 S447 E448 R454 V459 R464 D465 V472 L473 A474 A475 T476 L479 Q480 L481 E482 E483 D484 I485 S488 Y489 L490 F495 A496 N497 P498 T499 E500
P505 V506 G507 R508 L509 N510 M511 T512 V513 F519 SER ARG ASP ALA GLU P525 F526 L527 V532 Q539 Y540 M543 R544 R545 Y546 E547 SER SER ALA S551 R556 H560 PRO GLY GLY SER GLY GLY SER P568 L574 F575 D576 E584 L585 R588 M596 D597 N598
P601 K605 G606 H609 G610 Q611 T612 P616 T621 N624

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	101.80Å 136.54Å 116.52Å 90.00° 112.61° 90.00°	Depositor
Resolution (Å)	35.94 – 3.06 35.94 – 3.06	Depositor EDS
% Data completeness (in resolution range)	97.7 (35.94-3.06) 97.7 (35.94-3.06)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.59 (at 3.06Å)	Xtriage
Refinement program	PHENIX (1.18.2_3874: ???)	Depositor
R, R_{free}	0.260 , 0.286 0.260 , 0.285	Depositor DCC
R_{free} test set	1341 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å^2)	80.7	Xtriage
Anisotropy	0.635	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 71.5	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	0.94	EDS
Total number of atoms	8522	wwPDB-VP
Average B, all atoms (Å^2)	99.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 43.50 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.7472e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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	0.53	0/4373	0.85	7/5913 (0.1%)
1	B	0.51	0/4295	0.85	7/5805 (0.1%)
All	All	0.52	0/8668	0.85	14/11718 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
1	B	0	4
All	All	0	6

There are no bond length outliers.

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	499	THR	N-CA-C	-7.06	105.29	114.04
1	B	392	HIS	N-CA-C	-6.81	105.30	112.93
1	A	496	ALA	N-CA-C	6.74	125.15	110.80
1	B	498	PRO	CB-CA-C	6.67	122.56	111.56
1	B	499	THR	N-CA-C	6.42	118.85	111.02

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	19	ARG	Sidechain
1	B	281	PHE	Peptide
1	B	327	PHE	Peptide
1	B	344	ARG	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	B	544	ARG	Sidechain

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	4276	0	4287	146	0
1	B	4204	0	4205	144	0
2	A	24	0	0	7	0
2	B	18	0	0	2	0
All	All	8522	0	8492	278	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 16.

The worst 5 of 278 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:97:THR:HG23	1:A:100:GLU:H	1.36	0.91
1:A:235:MET:HA	2:A:705:HOH:O	1.76	0.84
1:B:394:SER:CB	1:B:412:ALA:H	1.93	0.80
1:B:341:LYS:HB2	1:B:463:VAL:HG13	1.65	0.78
1:A:23:ILE:HG23	1:A:252:VAL:HA	1.65	0.77

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	505/648 (78%)	456 (90%)	45 (9%)	4 (1%)	16	42
1	B	493/648 (76%)	450 (91%)	39 (8%)	4 (1%)	16	42
All	All	998/1296 (77%)	906 (91%)	84 (8%)	8 (1%)	16	42

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	328	PRO
1	A	433	GLY
1	B	478	GLY
1	A	282	PRO
1	B	282	PRO

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	486/588 (83%)	468 (96%)	18 (4%)	30	57
1	B	476/588 (81%)	457 (96%)	19 (4%)	28	56
All	All	962/1176 (82%)	925 (96%)	37 (4%)	29	57

5 of 37 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	278	SER
1	A	417	LEU
1	A	279	LYS
1	A	283	ILE
1	B	392	HIS

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

Mol	Chain	Res	Type
1	A	594	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	560	HIS
1	A	326	ASN
1	B	560	HIS
1	A	392	HIS

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 oligosaccharides 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	527/648 (81%)	0.30	9 (1%)	69	46	30, 94, 136, 168	0
1	B	521/648 (80%)	0.26	5 (0%)	79	59	62, 96, 141, 159	0
All	All	1048/1296 (80%)	0.28	14 (1%)	75	53	30, 94, 140, 168	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	280	GLU	4.2
1	A	378	ASP	3.5
1	B	140	SER	3.4
1	A	279	LYS	3.4
1	B	457	PHE	2.9

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 oligosaccharides 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.