



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 1, 2026 – 06:57 AM UTC

PDB ID : 5YH0 / pdb_00005yh0
Title : The structure of DrFam20C1
Authors : Zhang, H.; Xiao, J.
Deposited on : 2017-09-27
Resolution : 3.45 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

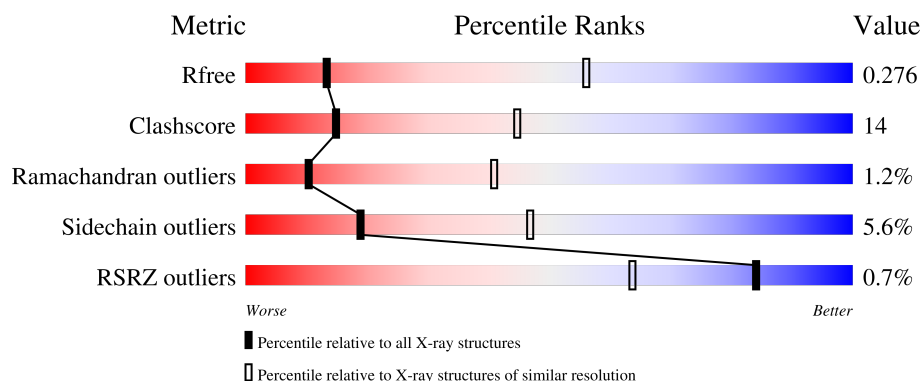
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.45 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	1070 (3.50-3.42)
Clashscore	190562	1128 (3.50-3.42)
Ramachandran outliers	187476	1101 (3.50-3.42)
Sidechain outliers	187428	1102 (3.50-3.42)
RSRZ outliers	180081	1070 (3.50-3.42)

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 $\leq 5\%$. 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	560	
1	B	560	
1	C	560	
1	D	560	
1	E	560	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	560	% 43% 23% 33%
1	G	560	47% 20% 31%
1	H	560	% 43% 23% 32%
1	I	560	44% 22% 32%
1	J	560	47% 19% 32%
1	K	560	% 45% 21% 31%
1	L	560	% 46% 20% 32%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 37919 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.

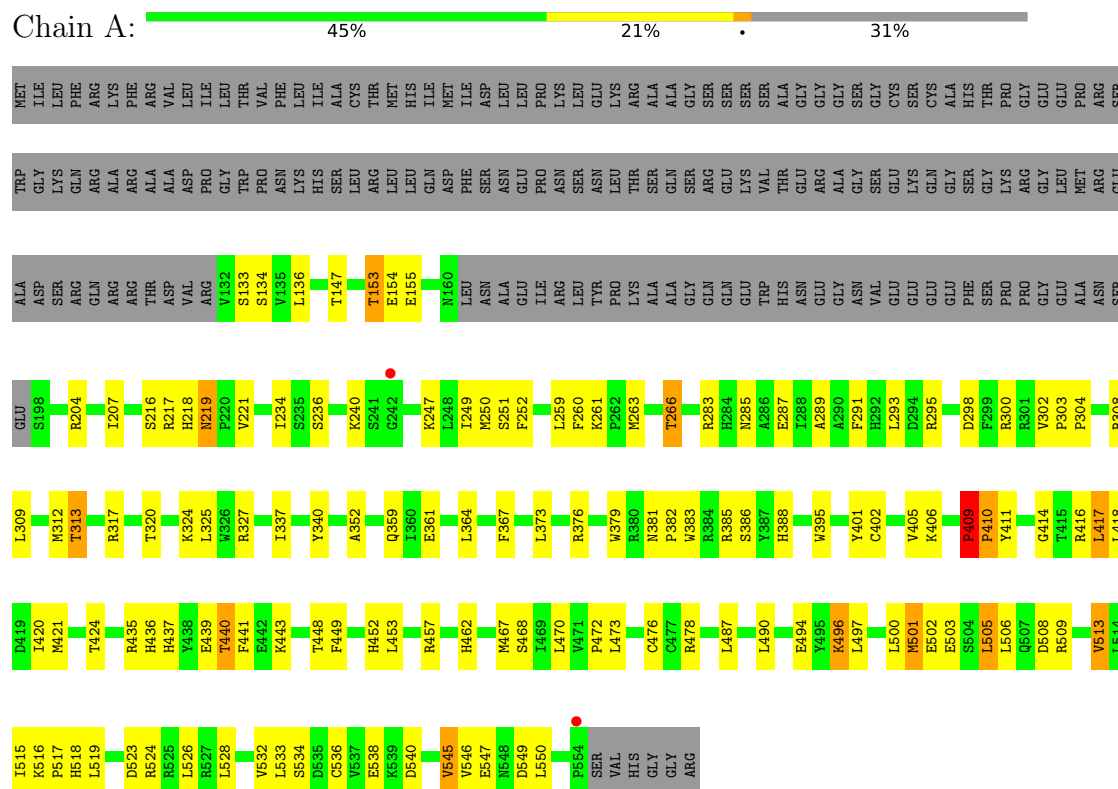
- Molecule 1 is a protein called DrFam20C1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	386	Total	C	N	O	S	0	0	0
			3181	2028	559	576	18			
1	B	380	Total	C	N	O	S	0	0	0
			3141	2005	551	567	18			
1	C	390	Total	C	N	O	S	0	0	0
			3213	2048	566	581	18			
1	D	382	Total	C	N	O	S	0	0	0
			3157	2013	554	572	18			
1	E	385	Total	C	N	O	S	0	0	0
			3173	2024	557	574	18			
1	F	377	Total	C	N	O	S	0	0	0
			3118	1991	548	561	18			
1	G	386	Total	C	N	O	S	0	0	0
			3181	2028	559	576	18			
1	H	380	Total	C	N	O	S	0	0	0
			3140	2004	551	567	18			
1	I	383	Total	C	N	O	S	0	0	0
			3162	2017	556	571	18			
1	J	378	Total	C	N	O	S	0	0	0
			3124	1994	549	563	18			
1	K	389	Total	C	N	O	S	0	0	0
			3205	2042	565	580	18			
1	L	378	Total	C	N	O	S	0	0	0
			3124	1994	549	563	18			

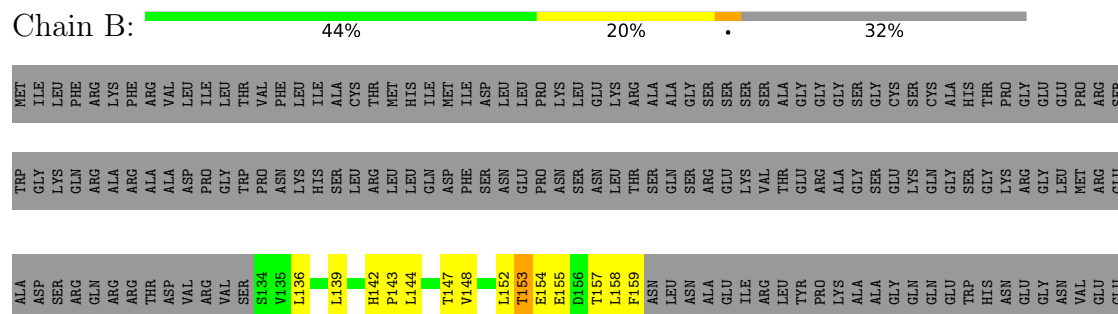
3 Residue-property plots

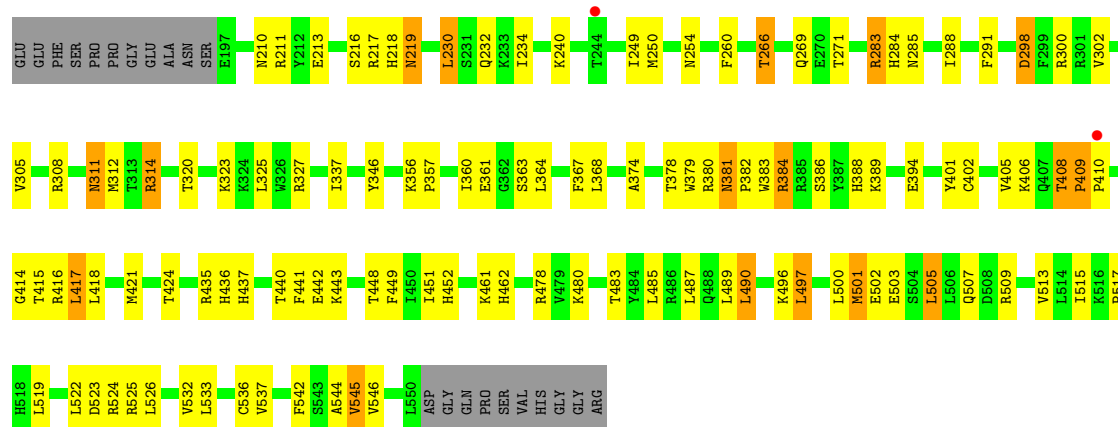
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: DrFam20C1

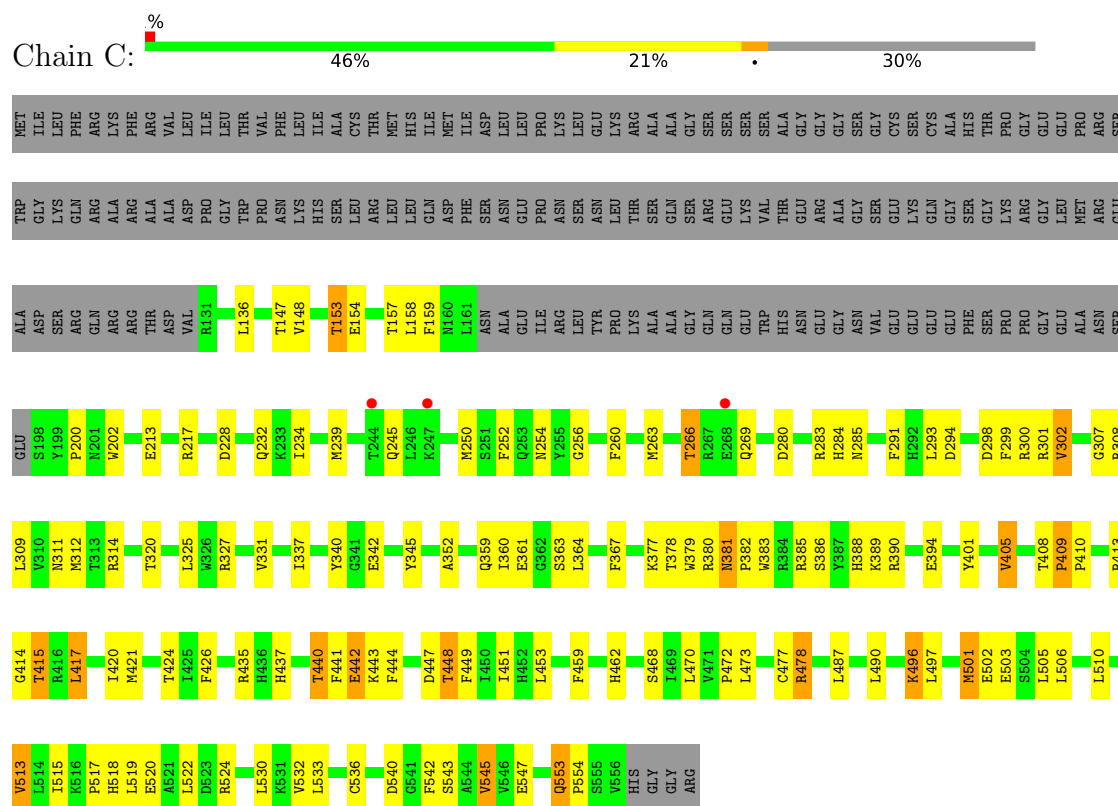


• Molecule 1: DrFam20C1

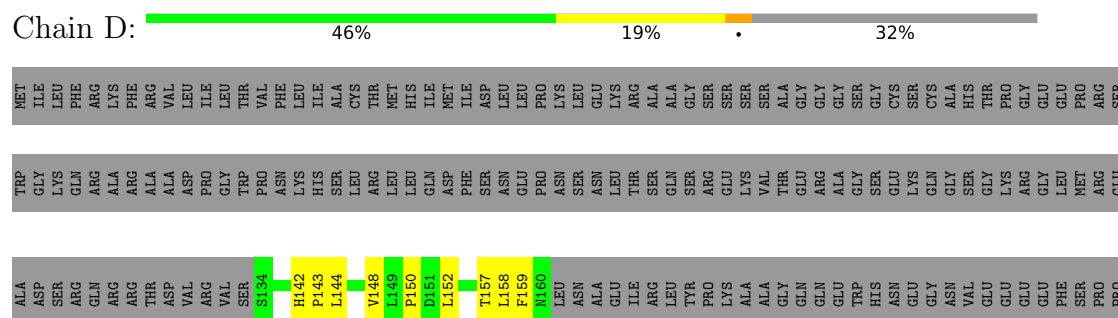


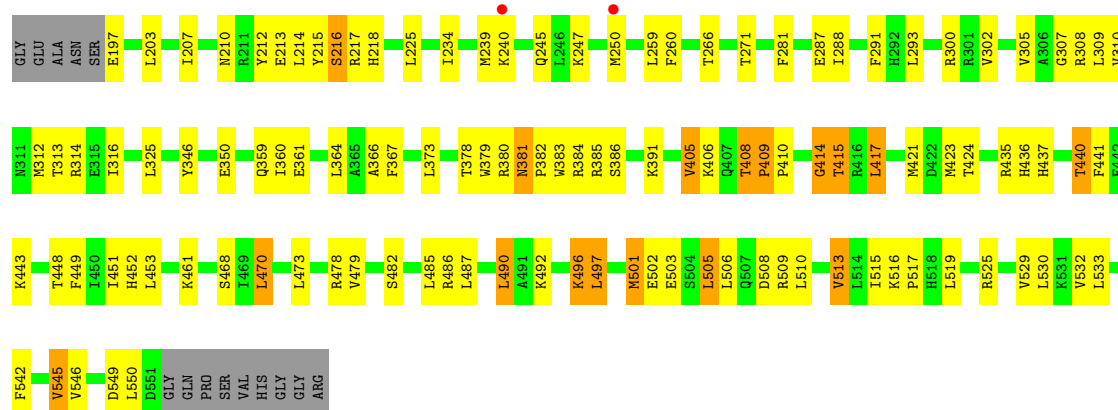


• Molecule 1: DrFam20C1



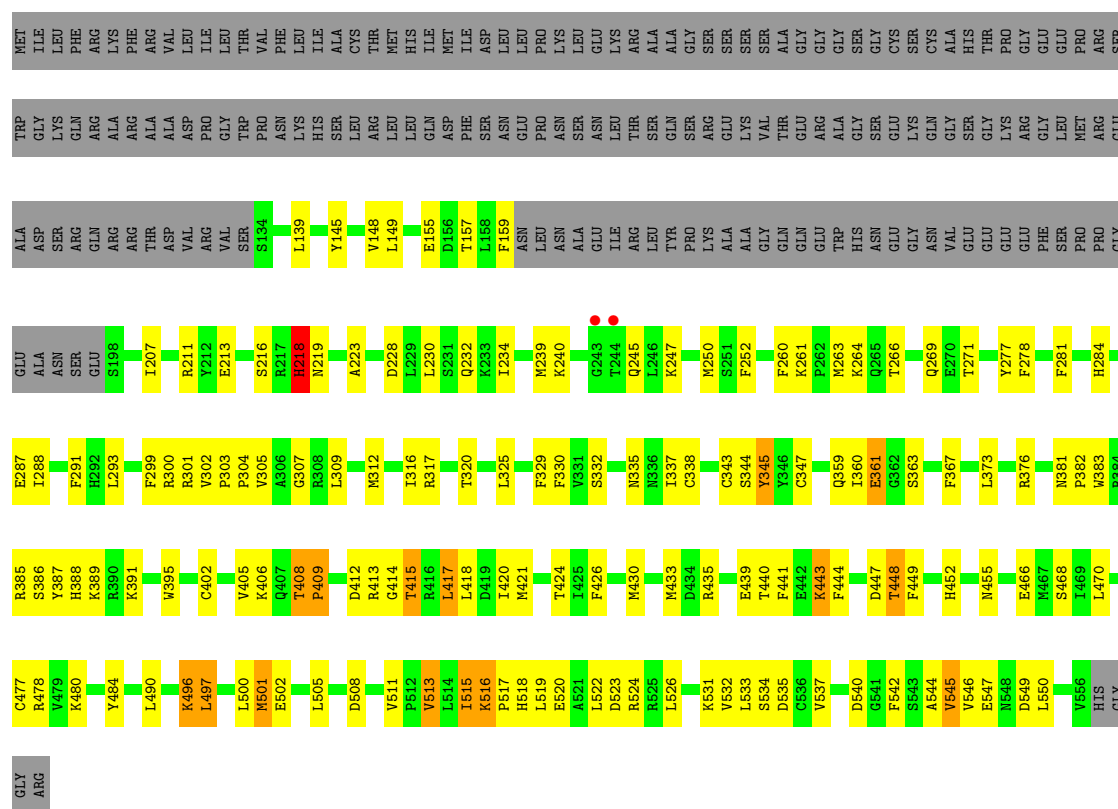
• Molecule 1: DrFam20C1





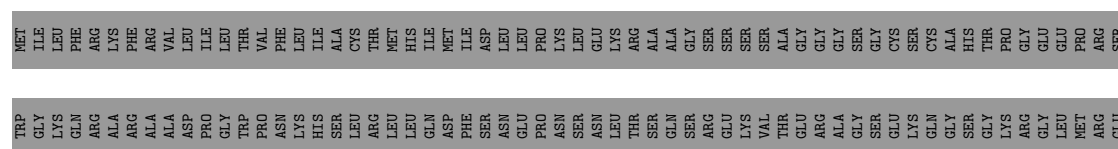
• Molecule 1: DrFam20C1

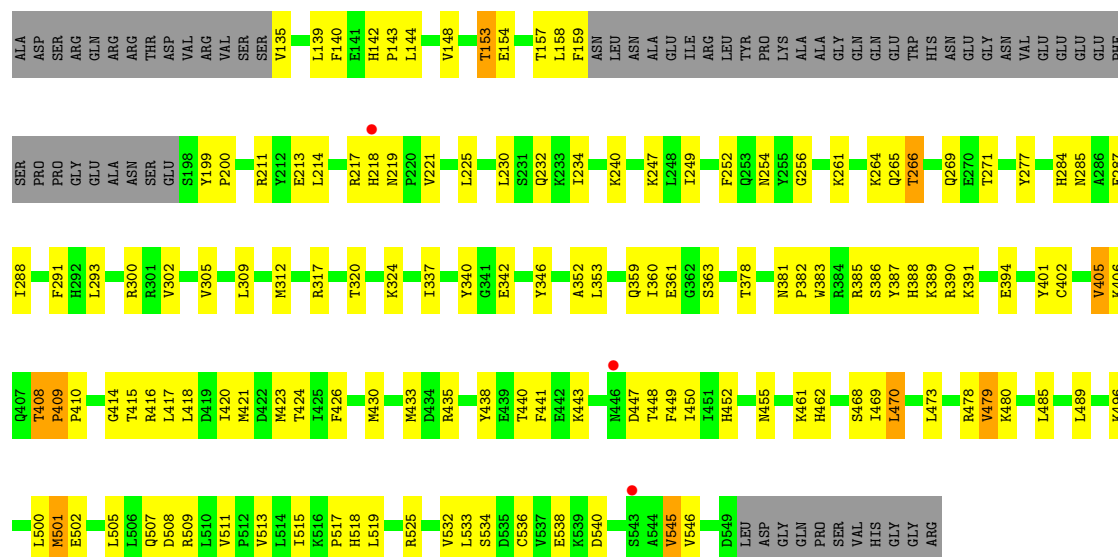
Chain E: 42% 24% 31%



• Molecule 1: DrFam20C1

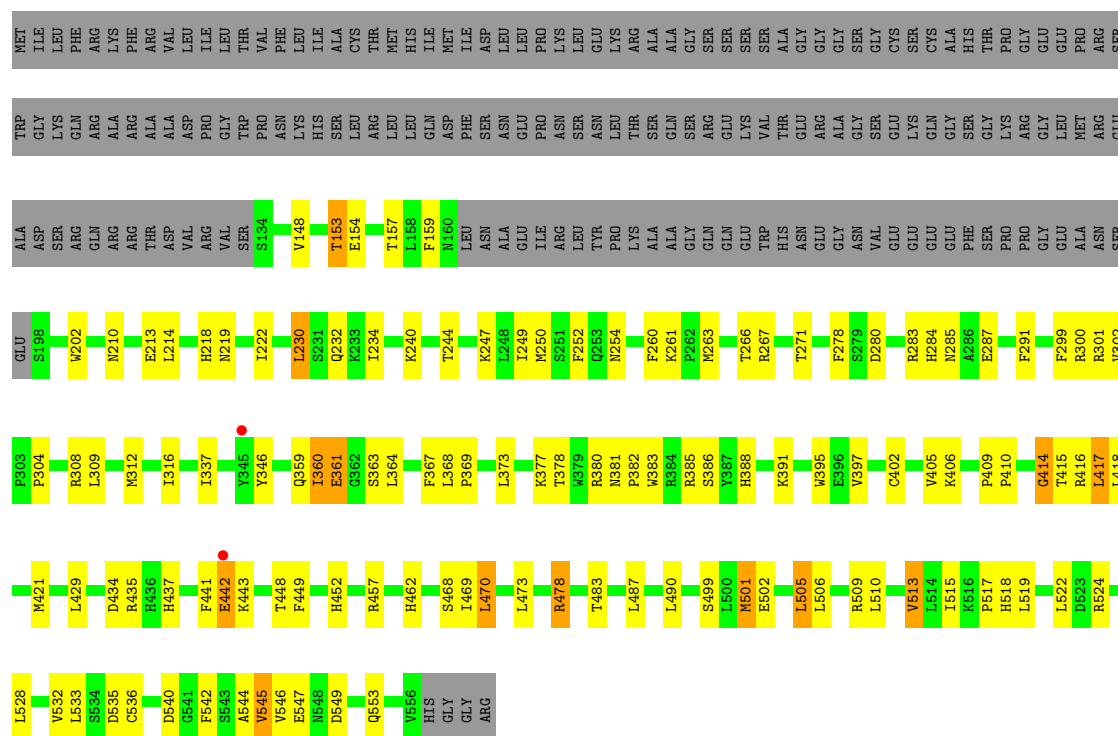
Chain F: 43% 23% 33%





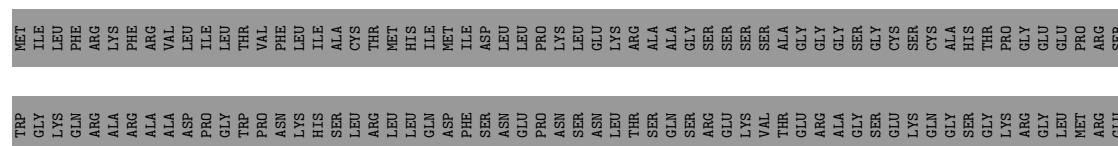
● Molecule 1: DrFam20C1

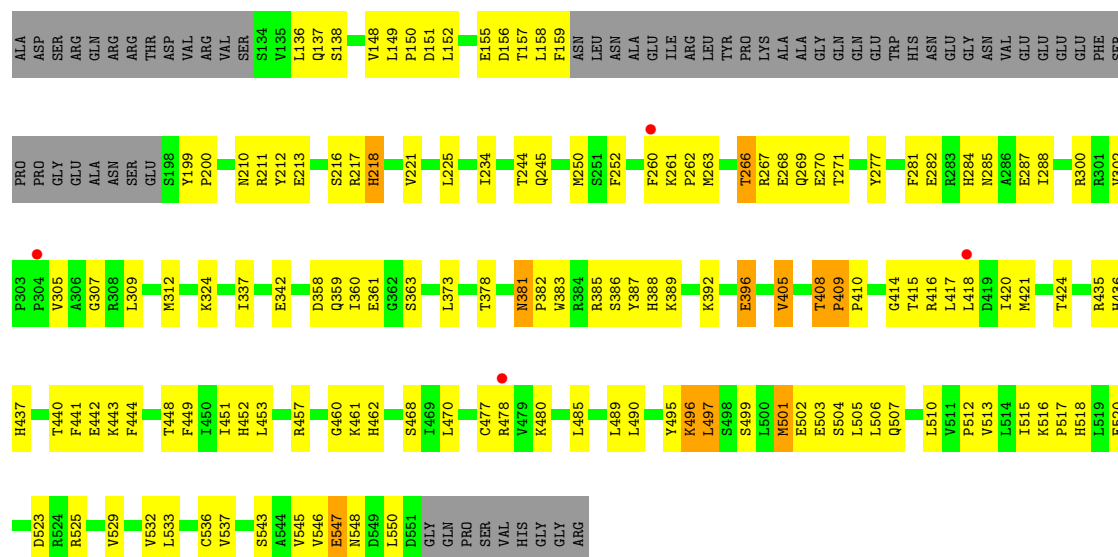
Chain G: 47% 20% 31%



● Molecule 1: DrFam20C1

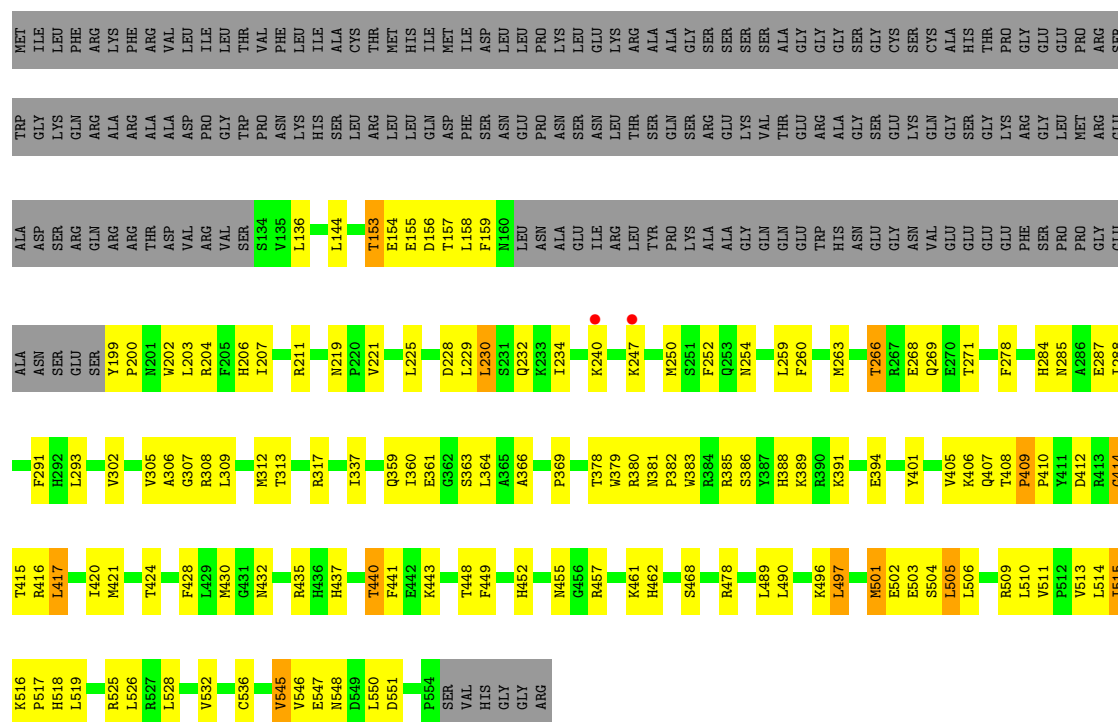
Chain H: 43% 23% 32%





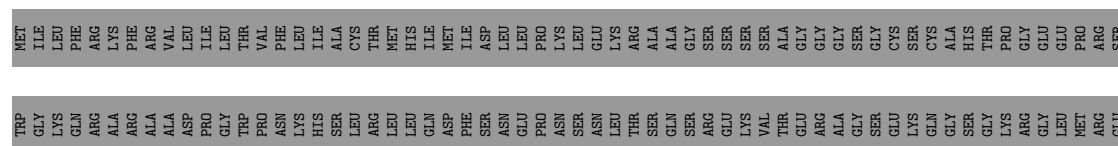
• Molecule 1: DrFam20C1

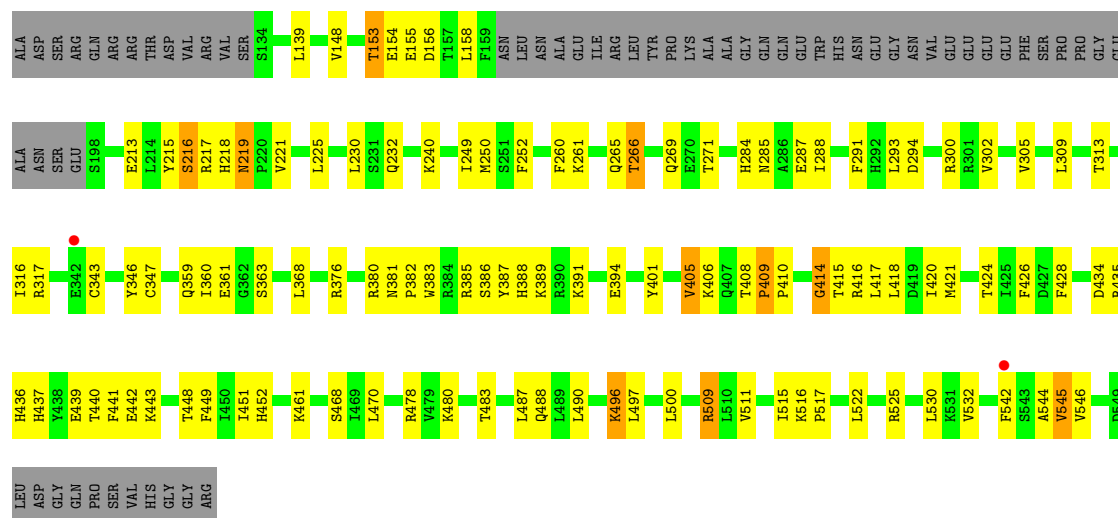
Chain I: 44% 22% 32%



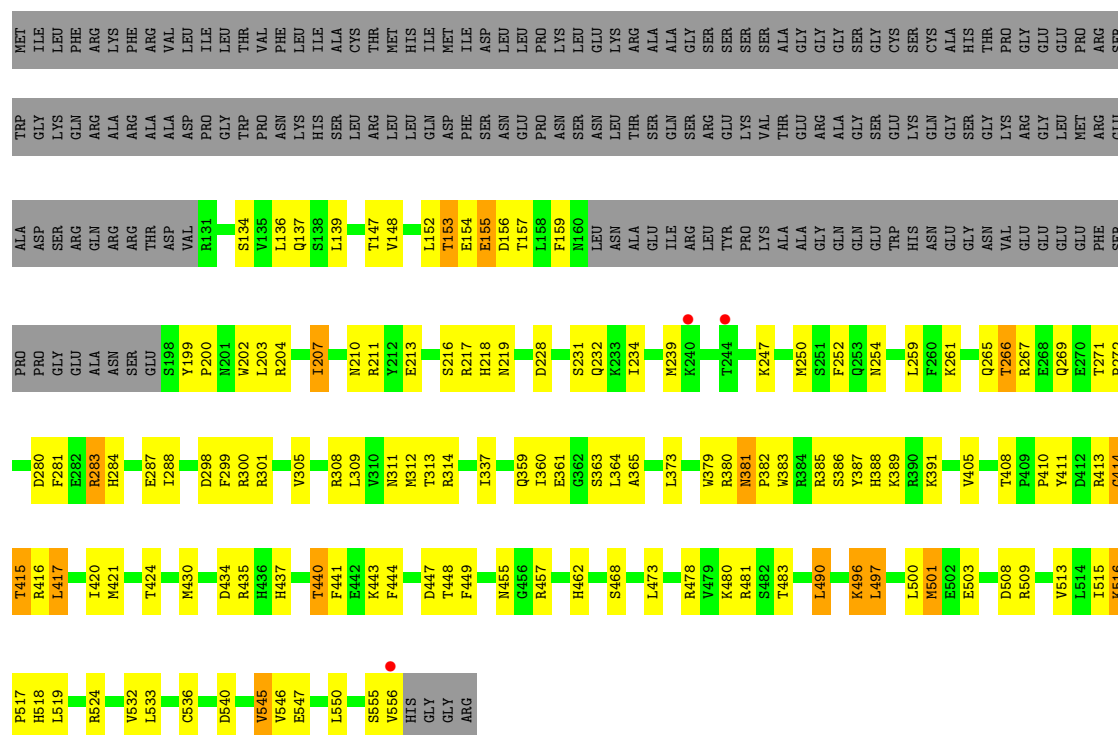
• Molecule 1: DrFam20C1

Chain J: 47% 19% 32%

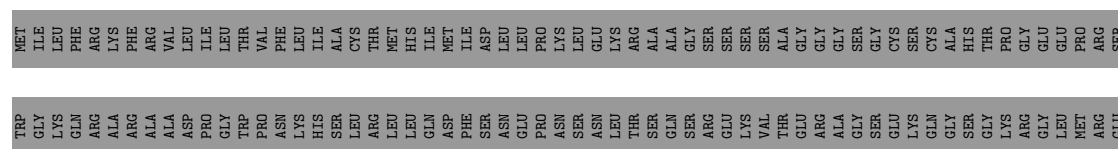


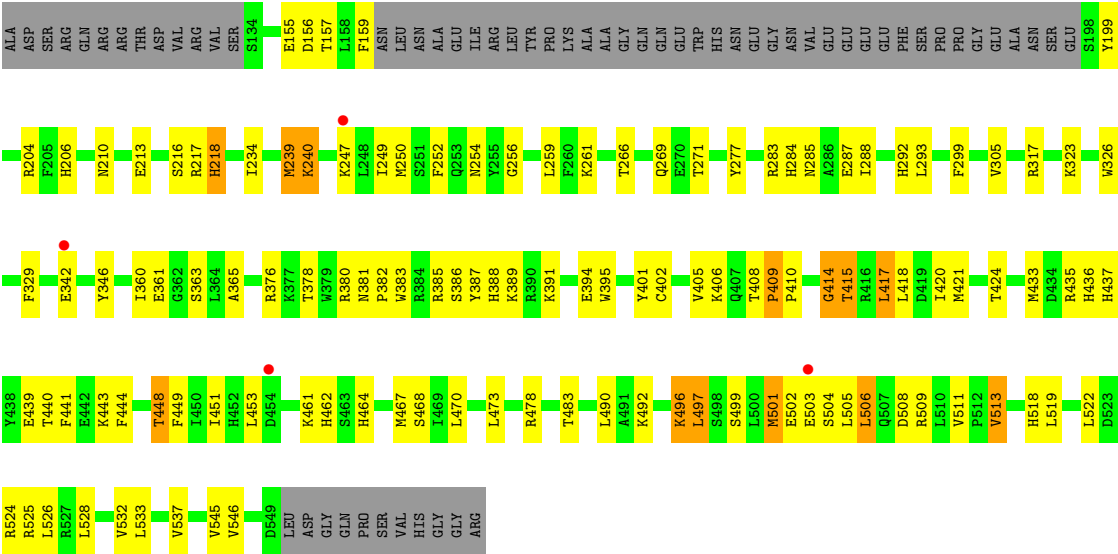


• Molecule 1: DrFam20C1



• Molecule 1: DrFam20C1





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	114.92Å 135.99Å 219.97Å 90.00° 90.50° 90.00°	Depositor
Resolution (Å)	47.84 – 3.45 47.84 – 3.45	Depositor EDS
% Data completeness (in resolution range)	99.3 (47.84-3.45) 91.7 (47.84-3.45)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.83 (at 3.48Å)	Xtriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.225 , 0.279 0.226 , 0.276	Depositor DCC
R_{free} test set	2005 reflections (2.26%)	wwPDB-VP
Wilson B-factor (Å ²)	54.3	Xtriage
Anisotropy	0.198	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 20.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.028 for h,-k,-l	Xtriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	37919	wwPDB-VP
Average B, all atoms (Å ²)	54.0	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 25.96 % 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 2.8756e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

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	0.20	0/3259	0.55	2/4403 (0.0%)
1	B	0.21	0/3218	0.57	0/4346
1	C	0.22	0/3291	0.56	0/4446
1	D	0.21	0/3234	0.56	2/4368 (0.0%)
1	E	0.21	0/3251	0.56	0/4392
1	F	0.19	0/3195	0.54	0/4315
1	G	0.19	0/3259	0.54	2/4403 (0.0%)
1	H	0.20	0/3217	0.56	0/4345
1	I	0.19	0/3240	0.54	2/4377 (0.0%)
1	J	0.19	0/3201	0.55	2/4323 (0.0%)
1	K	0.21	0/3283	0.58	2/4435 (0.0%)
1	L	0.20	0/3201	0.54	2/4323 (0.0%)
All	All	0.20	0/38849	0.55	14/52476 (0.0%)

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	1

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	414	GLY	CA-C-N	5.69	131.94	121.70
1	A	414	GLY	C-N-CA	5.69	131.94	121.70
1	J	414	GLY	CA-C-N	5.32	131.28	121.70
1	J	414	GLY	C-N-CA	5.32	131.28	121.70
1	I	414	GLY	CA-C-N	5.30	131.24	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	409	PRO	Peptide

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	3181	0	3139	88	0
1	B	3141	0	3101	86	0
1	C	3213	0	3175	91	0
1	D	3157	0	3111	78	0
1	E	3173	0	3131	101	0
1	F	3118	0	3079	92	0
1	G	3181	0	3137	86	0
1	H	3140	0	3099	92	0
1	I	3162	0	3118	98	0
1	J	3124	0	3084	76	0
1	K	3205	0	3164	84	0
1	L	3124	0	3084	81	0
All	All	37919	0	37422	1043	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 14.

The worst 5 of 1043 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:515:ILE:HG13	1:A:517:PRO:HD2	1.44	0.98
1:B:240:LYS:HG3	1:B:249:ILE:HD11	1.54	0.89
1:L:385:ARG:HH22	1:L:468:SER:HB2	1.38	0.86
1:I:291:PHE:HA	1:I:302:VAL:HG21	1.55	0.86
1:K:515:ILE:HG13	1:K:517:PRO:HD2	1.58	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	382/560 (68%)	338 (88%)	38 (10%)	6 (2%)	7	36
1	B	376/560 (67%)	334 (89%)	38 (10%)	4 (1%)	11	42
1	C	386/560 (69%)	341 (88%)	43 (11%)	2 (0%)	24	57
1	D	378/560 (68%)	334 (88%)	38 (10%)	6 (2%)	7	36
1	E	381/560 (68%)	338 (89%)	37 (10%)	6 (2%)	7	36
1	F	373/560 (67%)	331 (89%)	39 (10%)	3 (1%)	16	49
1	G	382/560 (68%)	342 (90%)	35 (9%)	5 (1%)	9	39
1	H	376/560 (67%)	330 (88%)	42 (11%)	4 (1%)	11	42
1	I	379/560 (68%)	334 (88%)	41 (11%)	4 (1%)	11	42
1	J	374/560 (67%)	328 (88%)	41 (11%)	5 (1%)	9	39
1	K	385/560 (69%)	337 (88%)	42 (11%)	6 (2%)	7	36
1	L	374/560 (67%)	331 (88%)	39 (10%)	4 (1%)	11	42
All	All	4546/6720 (68%)	4018 (88%)	473 (10%)	55 (1%)	10	41

5 of 55 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	410	PRO
1	A	496	LYS
1	D	216	SER
1	D	218	HIS
1	H	218	HIS

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	352/495 (71%)	333 (95%)	19 (5%)	20	47
1	B	347/495 (70%)	325 (94%)	22 (6%)	16	44
1	C	356/495 (72%)	334 (94%)	22 (6%)	16	44
1	D	349/495 (70%)	325 (93%)	24 (7%)	14	41
1	E	351/495 (71%)	329 (94%)	22 (6%)	16	44
1	F	344/495 (70%)	328 (95%)	16 (5%)	23	51
1	G	352/495 (71%)	330 (94%)	22 (6%)	16	44
1	H	347/495 (70%)	328 (94%)	19 (6%)	19	47
1	I	349/495 (70%)	334 (96%)	15 (4%)	26	52
1	J	345/495 (70%)	333 (96%)	12 (4%)	32	58
1	K	355/495 (72%)	332 (94%)	23 (6%)	15	43
1	L	345/495 (70%)	325 (94%)	20 (6%)	18	45
All	All	4192/5940 (71%)	3956 (94%)	236 (6%)	19	47

5 of 236 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	378	THR
1	L	285	ASN
1	G	513	VAL
1	L	271	THR
1	K	266	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	J	265	GLN
1	L	388	HIS
1	J	351	HIS
1	K	269	GLN
1	E	218	HIS

5.3.3 RNA ⓘ

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 ⓘ

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	386/560 (68%)	-0.07	2 (0%) 87 70	24, 42, 65, 90	0
1	B	380/560 (67%)	-0.06	2 (0%) 87 70	22, 44, 75, 99	0
1	C	390/560 (69%)	-0.15	3 (0%) 82 61	19, 34, 58, 82	0
1	D	382/560 (68%)	-0.01	2 (0%) 87 70	28, 51, 78, 104	0
1	E	385/560 (68%)	-0.09	2 (0%) 87 70	23, 43, 67, 90	0
1	F	377/560 (67%)	0.26	3 (0%) 82 61	34, 71, 104, 120	0
1	G	386/560 (68%)	0.01	2 (0%) 87 70	30, 50, 72, 92	0
1	H	380/560 (67%)	0.30	4 (1%) 78 54	38, 72, 107, 131	0
1	I	383/560 (68%)	0.03	2 (0%) 87 70	38, 56, 79, 100	0
1	J	378/560 (67%)	0.10	2 (0%) 87 70	32, 55, 87, 118	0
1	K	389/560 (69%)	-0.04	3 (0%) 82 61	29, 46, 68, 92	0
1	L	378/560 (67%)	0.13	4 (1%) 78 54	33, 58, 88, 114	0
All	All	4594/6720 (68%)	0.03	31 (0%) 84 64	19, 51, 88, 131	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	244	THR	3.2
1	B	244	THR	2.8
1	D	240	LYS	2.7
1	L	247	LYS	2.7
1	E	243	GLY	2.7

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 [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.