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

wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 16, 2023 – 08:50 PM EDT

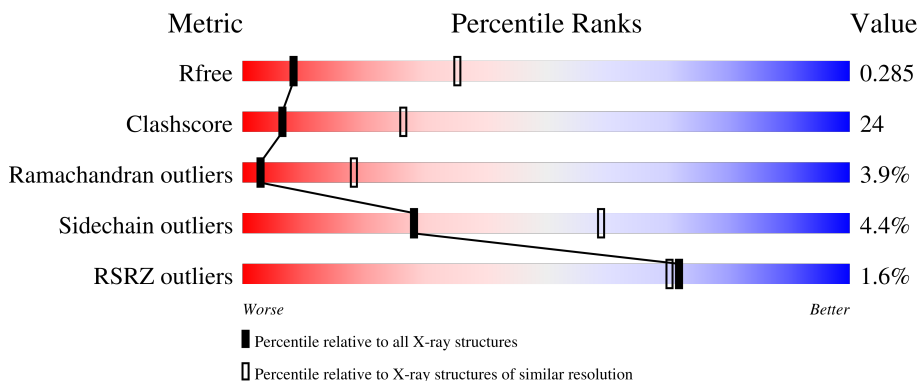|		
PDB ID	:	2EU1
Title	:	Crystal structure of the chaperonin GroEL-E461K
Authors	:	Cabo-Bilbao, A.; Spinelli, S.; Sot, B.; Agirre, J.; Mechaly, A.E.; Muga, A.; Guerin, D.M.A.
Deposited on	:	2005-10-28
Resolution	:	3.29 Å(reported)

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 3.29 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.
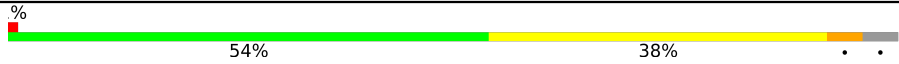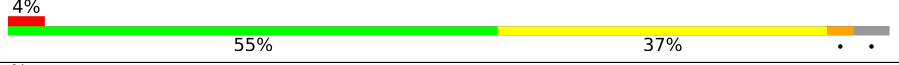
Metric	Percentile Ranks	Value
Rfree		0.285
Clashscore		24
Ramachandran outliers		3.9%
Sidechain outliers		4.4%
RSRZ outliers		1.6%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
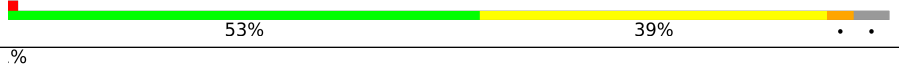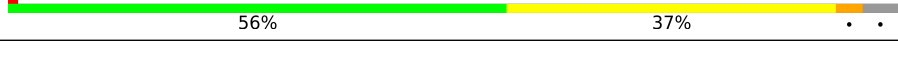
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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	548	2% 53% 38% · ·
1	B	548	% 56% 36% · ·
1	C	548	% 57% 35% · ·
1	D	548	3% 53% 39% · ·
1	E	548	% 53% 39% · ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	548	%　54%　38%
1	G	548	%　54%　38%
1	H	548	4%　55%　37%
1	I	548	%　57%　37%
1	J	548	2%　57%　35%
1	K	548	%　55%　38%
1	L	548	%　53%　39%
1	M	548	%　53%　39%
1	N	548	%　56%　37%

2 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 53970 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 GROEL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	524	Total 3855	C 2398	N 666	O 771	S 20	0	0	0
1	B	524	Total 3855	C 2398	N 666	O 771	S 20	0	0	0
1	C	524	Total 3855	C 2398	N 666	O 771	S 20	0	0	0
1	D	524	Total 3855	C 2398	N 666	O 771	S 20	0	0	0
1	E	524	Total 3855	C 2398	N 666	O 771	S 20	0	0	0
1	F	524	Total 3855	C 2398	N 666	O 771	S 20	0	0	0
1	G	524	Total 3855	C 2398	N 666	O 771	S 20	0	0	0
1	H	524	Total 3855	C 2398	N 666	O 771	S 20	0	0	0
1	I	524	Total 3855	C 2398	N 666	O 771	S 20	0	0	0
1	J	524	Total 3855	C 2398	N 666	O 771	S 20	0	0	0
1	K	524	Total 3855	C 2398	N 666	O 771	S 20	0	0	0
1	L	524	Total 3855	C 2398	N 666	O 771	S 20	0	0	0
1	M	524	Total 3855	C 2398	N 666	O 771	S 20	0	0	0
1	N	524	Total 3855	C 2398	N 666	O 771	S 20	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	461	LYS	GLU	engineered mutation	UNP P0A6F5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	461	LYS	GLU	engineered mutation	UNP P0A6F5
C	461	LYS	GLU	engineered mutation	UNP P0A6F5
D	461	LYS	GLU	engineered mutation	UNP P0A6F5
E	461	LYS	GLU	engineered mutation	UNP P0A6F5
F	461	LYS	GLU	engineered mutation	UNP P0A6F5
G	461	LYS	GLU	engineered mutation	UNP P0A6F5
H	461	LYS	GLU	engineered mutation	UNP P0A6F5
I	461	LYS	GLU	engineered mutation	UNP P0A6F5
J	461	LYS	GLU	engineered mutation	UNP P0A6F5
K	461	LYS	GLU	engineered mutation	UNP P0A6F5
L	461	LYS	GLU	engineered mutation	UNP P0A6F5
M	461	LYS	GLU	engineered mutation	UNP P0A6F5
N	461	LYS	GLU	engineered mutation	UNP P0A6F5

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: GROEL

Chain A:

• Molecule 1: GROEL

Chain B:

• Molecule 1: GROEL

Chain C:

• Molecule 1: GROEL

Chain D:

• Molecule 1: GROEL

Chain E:

• Molecule 1: GROEL

Chain F:
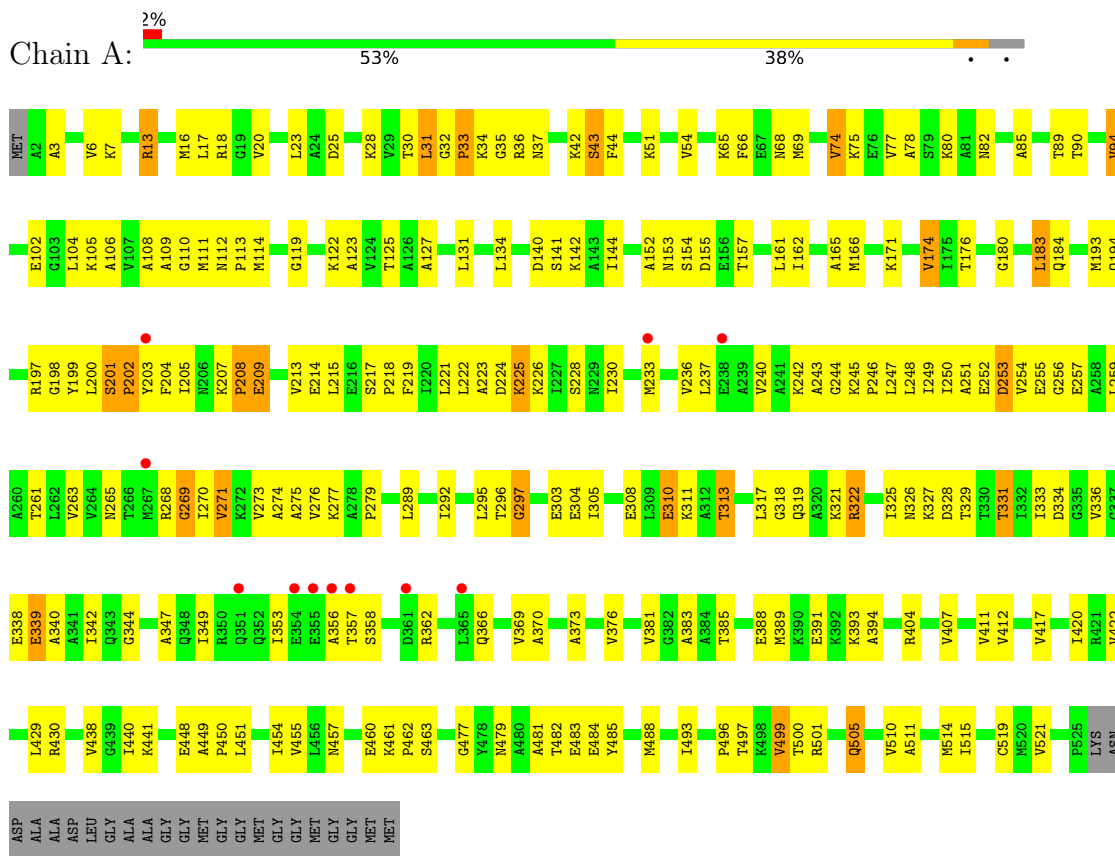
● Molecule 1: GROEL

● Molecule 1: GROEL

Chain H:

• Molecule 1: GROEL

Chain I:

• Molecule 1: GROEL

Chain J:

• Molecule 1: GROEL

Chain K:

• Molecule 1: GROEL

Chain L:

• Molecule 1: GROEL

Chain M:

• Molecule 1: GROEL

Chain M:

• Molecule 1: GROEL

Chain N:
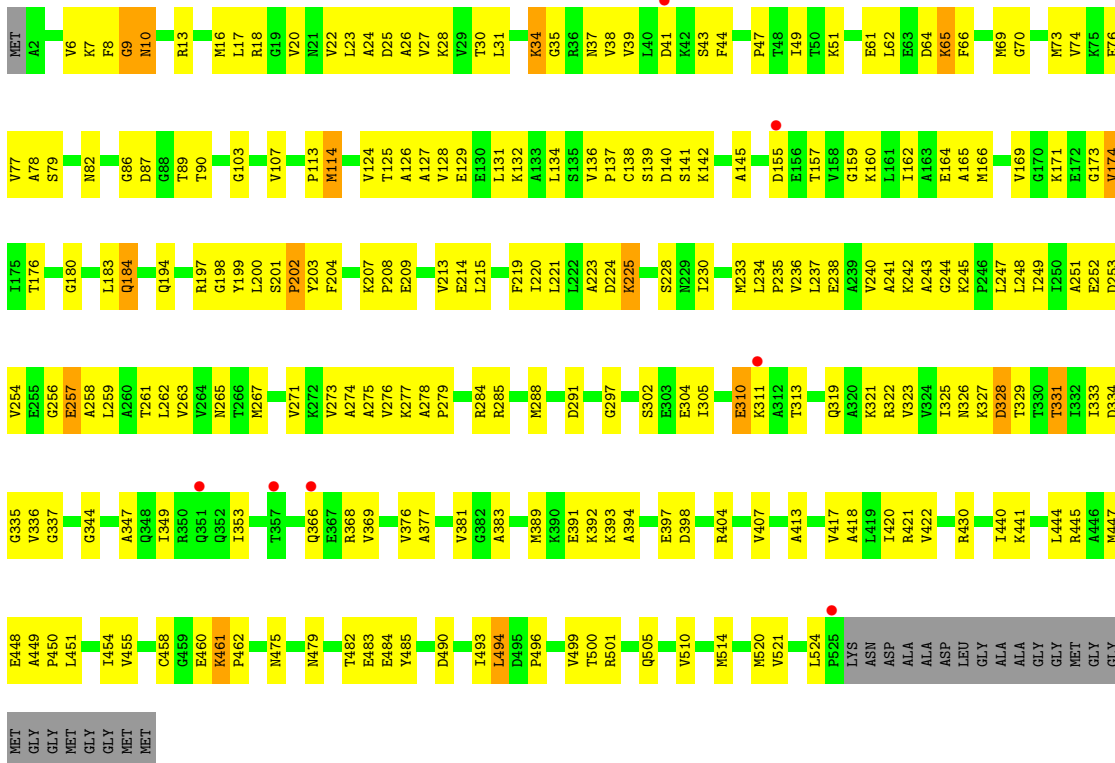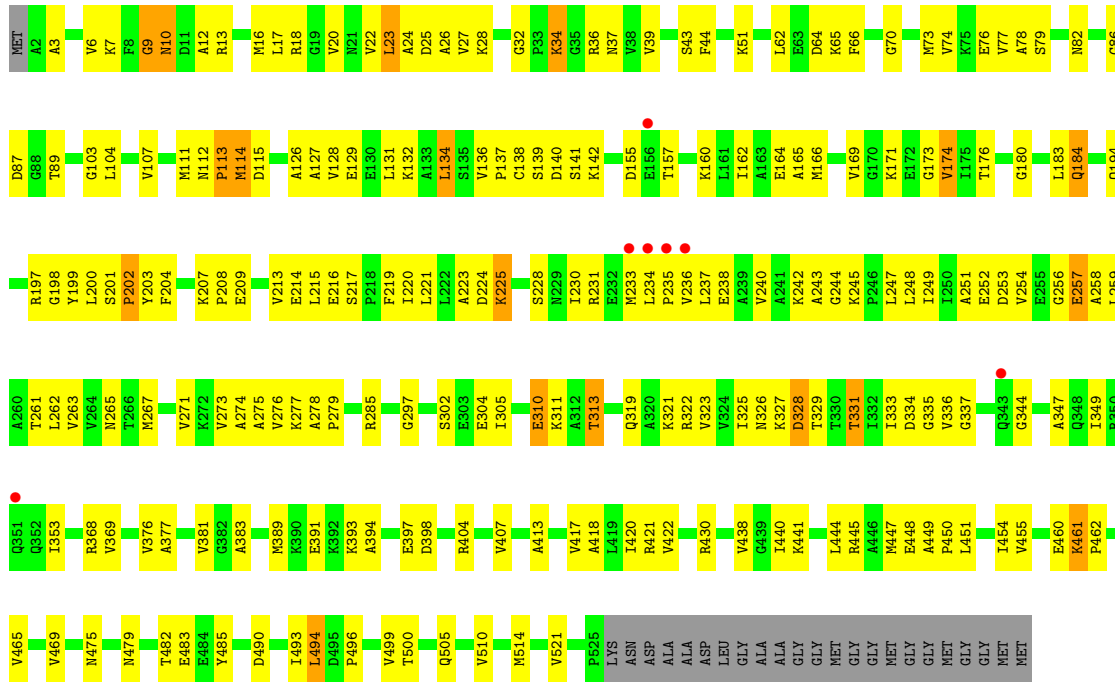
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	267.69Å 290.64Å 247.01Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 3.29 39.75 – 3.29	Depositor EDS
% Data completeness (in resolution range)	(Not available) (15.00-3.29) 85.4 (39.75-3.29)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.70 (at 3.32Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.276 , 0.296 0.266 , 0.285	Depositor DCC
R_{free} test set	6133 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	70.6	Xtriage
Anisotropy	0.555	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 56.8	EDS
L-test for twinning[2]	$< |L| > = 0.43, < L^2 > = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	53970	wwPDB-VP
Average B, all atoms (Å^2)	57.0	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 44.51 % 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 1.5101e-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 ⓘ

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	0.56	1/3883 (0.0%)	0.78	3/5242 (0.1%)
1	B	0.59	0/3883	0.66	2/5242 (0.0%)
1	C	0.55	0/3883	0.68	2/5242 (0.0%)
1	D	0.54	0/3883	0.69	2/5242 (0.0%)
1	E	0.53	1/3883 (0.0%)	0.77	3/5242 (0.1%)
1	F	0.54	0/3883	0.68	2/5242 (0.0%)
1	G	0.54	0/3883	0.67	2/5242 (0.0%)
1	H	0.54	0/3883	0.79	3/5242 (0.1%)
1	I	0.57	1/3883 (0.0%)	0.68	2/5242 (0.0%)
1	J	0.53	0/3883	0.82	6/5242 (0.1%)
1	K	0.52	0/3883	0.68	2/5242 (0.0%)
1	L	0.52	0/3883	0.79	3/5242 (0.1%)
1	M	0.54	0/3883	0.83	5/5242 (0.1%)
1	N	0.52	0/3883	0.78	5/5242 (0.1%)
All	All	0.54	3/54362 (0.0%)	0.74	42/73388 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	519	CYS	CB-SG	-5.20	1.73	1.81
1	E	519	CYS	CB-SG	-5.05	1.73	1.81
1	I	519	CYS	CB-SG	-5.05	1.73	1.81

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	M	284	ARG	NE-CZ-NH1	-24.36	108.12	120.30
1	A	13	ARG	NE-CZ-NH2	-23.32	108.64	120.30
1	H	13	ARG	NE-CZ-NH2	-22.86	108.87	120.30
1	M	284	ARG	NE-CZ-NH2	22.53	131.56	120.30
1	L	13	ARG	NE-CZ-NH2	-22.29	109.15	120.30

There are no chirality outliers.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3855	0	3983	183	0
1	B	3855	0	3983	168	0
1	C	3855	0	3983	170	0
1	D	3855	0	3983	195	0
1	E	3855	0	3983	198	0
1	F	3855	0	3983	200	1
1	G	3855	0	3983	175	0
1	H	3855	0	3983	205	0
1	I	3855	0	3983	184	0
1	J	3855	0	3983	181	0
1	K	3855	0	3983	208	0
1	L	3855	0	3983	227	0
1	M	3855	0	3983	215	1
1	N	3855	0	3983	206	0
All	All	53970	0	55762	2596	1

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 24.

The worst 5 of 2596 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:F:326:ASN:HD22	1:F:329:THR:HB	1.03	1.12
1:A:326:ASN:HD22	1:A:329:THR:HB	1.05	1.11
1:C:326:ASN:HD22	1:C:329:THR:HB	1.07	1.11
1:E:326:ASN:HD22	1:E:329:THR:HB	1.08	1.09
1:D:326:ASN:HD22	1:D:329:THR:HB	1.04	1.08

All (1) 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:F:484:GLU:OE1	1:M:484:GLU:OE1[8_556]	2.13	0.07

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	522/548 (95%)	449 (86%)	54 (10%)	19 (4%)	3	20
1	B	522/548 (95%)	450 (86%)	51 (10%)	21 (4%)	3	18
1	C	522/548 (95%)	455 (87%)	49 (9%)	18 (3%)	3	22
1	D	522/548 (95%)	448 (86%)	54 (10%)	20 (4%)	3	19
1	E	522/548 (95%)	453 (87%)	50 (10%)	19 (4%)	3	20
1	F	522/548 (95%)	449 (86%)	52 (10%)	21 (4%)	3	18
1	G	522/548 (95%)	450 (86%)	56 (11%)	16 (3%)	4	23
1	H	522/548 (95%)	434 (83%)	69 (13%)	19 (4%)	3	20
1	I	522/548 (95%)	438 (84%)	67 (13%)	17 (3%)	4	22
1	J	522/548 (95%)	441 (84%)	58 (11%)	23 (4%)	2	16
1	K	522/548 (95%)	441 (84%)	60 (12%)	21 (4%)	3	18
1	L	522/548 (95%)	434 (83%)	63 (12%)	25 (5%)	2	14
1	M	522/548 (95%)	438 (84%)	62 (12%)	22 (4%)	3	17
1	N	522/548 (95%)	440 (84%)	61 (12%)	21 (4%)	3	18
All	All	7308/7672 (95%)	6220 (85%)	806 (11%)	282 (4%)	3	18

5 of 282 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	33	PRO
1	A	43	SER
1	A	152	ALA
1	B	32	GLY

Continued on next page...

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

Continued from previous page...

Mol	Chain	Res	Type
1	B	33	PRO

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	404/415 (97%)	381 (94%)	23 (6%)	20	51
1	B	404/415 (97%)	383 (95%)	21 (5%)	23	54
1	C	404/415 (97%)	383 (95%)	21 (5%)	23	54
1	D	404/415 (97%)	383 (95%)	21 (5%)	23	54
1	E	404/415 (97%)	383 (95%)	21 (5%)	23	54
1	F	404/415 (97%)	381 (94%)	23 (6%)	20	51
1	G	404/415 (97%)	384 (95%)	20 (5%)	24	55
1	H	404/415 (97%)	388 (96%)	16 (4%)	31	61
1	I	404/415 (97%)	391 (97%)	13 (3%)	39	67
1	J	404/415 (97%)	391 (97%)	13 (3%)	39	67
1	K	404/415 (97%)	389 (96%)	15 (4%)	34	63
1	L	404/415 (97%)	389 (96%)	15 (4%)	34	63
1	M	404/415 (97%)	391 (97%)	13 (3%)	39	67
1	N	404/415 (97%)	391 (97%)	13 (3%)	39	67
All	All	5656/5810 (97%)	5408 (96%)	248 (4%)	28	59

5 of 248 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	125	THR
1	L	505	GLN
1	G	331	THR
1	L	490	ASP
1	N	65	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 92 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	319	GLN
1	K	72	GLN
1	H	475	ASN
1	J	37	ASN
1	K	475	ASN

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	524/548 (95%)	0.01	11 (2%)	63	62	5, 52, 107, 117	0
1	B	524/548 (95%)	0.04	5 (0%)	82	82	3, 50, 106, 118	0
1	C	524/548 (95%)	-0.05	3 (0%)	89	90	3, 52, 106, 118	0
1	D	524/548 (95%)	0.11	16 (3%)	49	48	6, 52, 107, 118	0
1	E	524/548 (95%)	0.02	7 (1%)	77	77	8, 53, 107, 117	0
1	F	524/548 (95%)	-0.01	7 (1%)	77	77	7, 52, 107, 118	0
1	G	524/548 (95%)	-0.00	8 (1%)	73	72	6, 53, 107, 118	0
1	H	524/548 (95%)	0.17	21 (4%)	38	36	4, 56, 114, 120	0
1	I	524/548 (95%)	0.05	5 (0%)	82	82	8, 55, 113, 120	0
1	J	524/548 (95%)	-0.02	9 (1%)	70	68	4, 56, 113, 120	0
1	K	524/548 (95%)	-0.03	4 (0%)	86	86	9, 56, 112, 120	0
1	L	524/548 (95%)	-0.01	6 (1%)	80	81	10, 56, 112, 120	0
1	M	524/548 (95%)	0.03	7 (1%)	77	77	10, 56, 113, 120	0
1	N	524/548 (95%)	0.04	7 (1%)	77	77	9, 56, 113, 120	0
All	All	7336/7672 (95%)	0.02	116 (1%)	72	70	3, 54, 110, 120	0

The worst 5 of 116 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	357	THR	4.7
1	A	361	ASP	4.4
1	A	238	GLU	4.3
1	I	268	ARG	4.1
1	J	203	TYR	3.7

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 ⓘ

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.