wwPDB X-ray Structure Validation Summary Report (i)

May 18, 2020 – 07:02 am BST

PDB ID	:	4HKB
Title	:	CH67 Fab (unbound) from the CH65-67 Lineage
Authors	:	Schmidt, A.G.; Harrison, S.C.
Deposited on	:	2012-10-15
Resolution	:	3.60 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 following versions of software and data (see references (1)) were used in the production of this report:

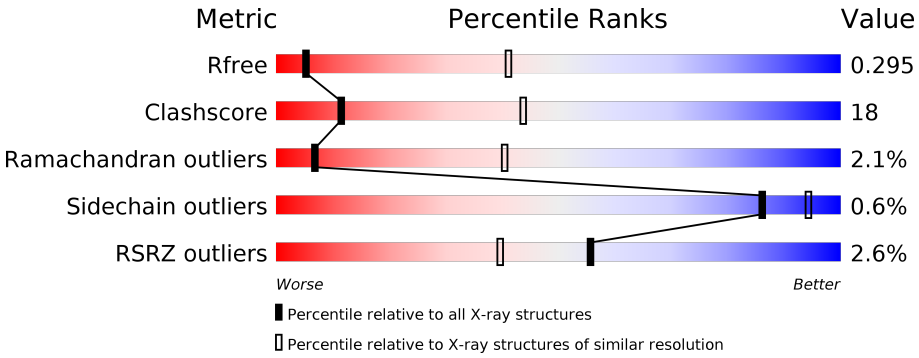
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.60 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.



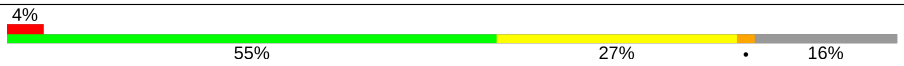
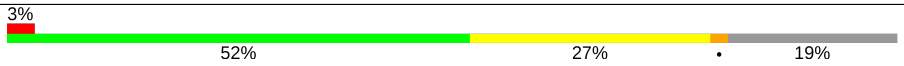
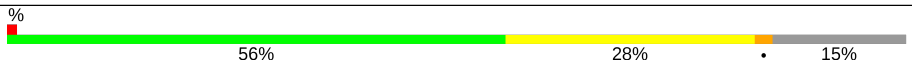
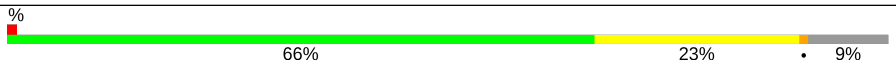
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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 <=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	236	65% 27% 7%
1	C	236	5% 69% 22% 8%
1	E	236	5% 65% 26% 7%
1	G	236	3% 69% 22% 8%
1	I	236	68% 24% 7%
1	J	236	64% 27% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	213	
2	D	213	
2	F	213	
2	H	213	
2	K	213	
2	N	213	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 17971 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 CH67 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	J	218	Total	C	N	O	S	0	0	0
			1665	1056	281	321	7			
1	A	220	Total	C	N	O	S	0	0	0
			1675	1061	283	324	7			
1	C	218	Total	C	N	O	S	0	0	0
			1662	1053	281	321	7			
1	E	220	Total	C	N	O	S	0	0	0
			1675	1061	283	324	7			
1	G	218	Total	C	N	O	S	0	0	0
			1664	1055	281	321	7			
1	I	220	Total	C	N	O	S	0	0	0
			1675	1061	283	324	7			

- Molecule 2 is a protein called CH67 light chain.

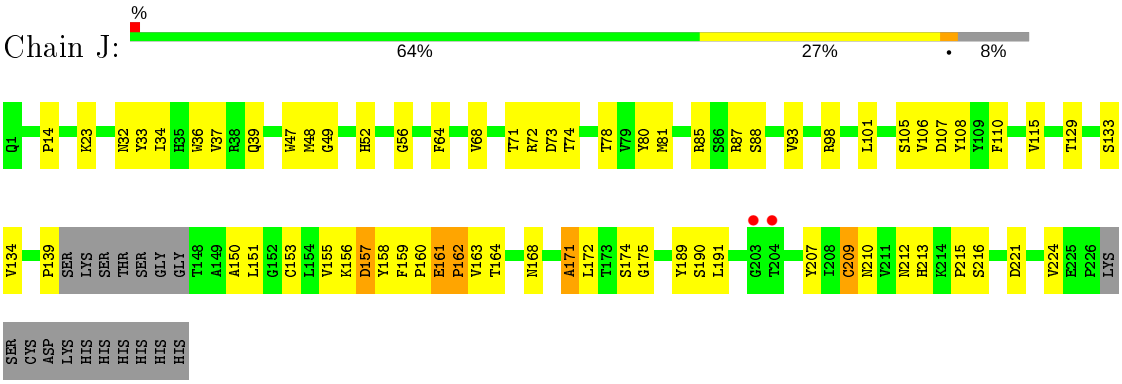
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	N	193	Total	C	N	O	S	0	0	0
			1433	894	241	294	4			
2	B	164	Total	C	N	O	S	0	0	0
			1216	755	205	252	4			
2	D	179	Total	C	N	O	S	0	0	0
			1341	834	228	275	4			
2	F	173	Total	C	N	O	S	0	0	0
			1303	809	222	268	4			
2	H	174	Total	C	N	O	S	0	0	0
			1306	813	221	268	4			
2	K	182	Total	C	N	O	S	0	0	0
			1356	845	229	278	4			

3 Residue-property plots (i)

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.

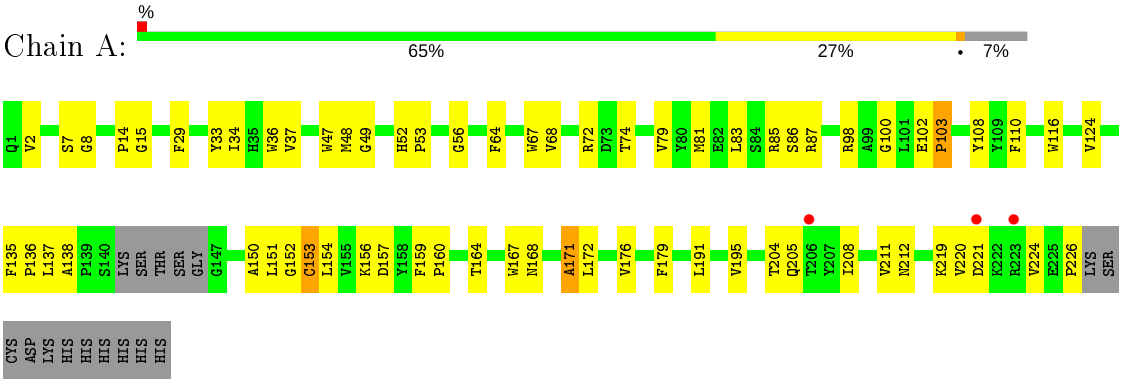
- Molecule 1: CH67 heavy chain

Chain J:



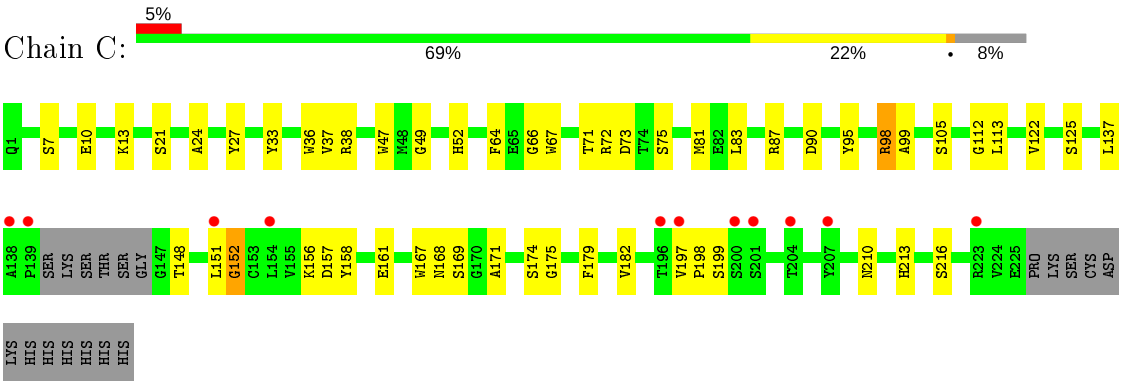
- Molecule 1: CH67 heavy chain

Chain A:



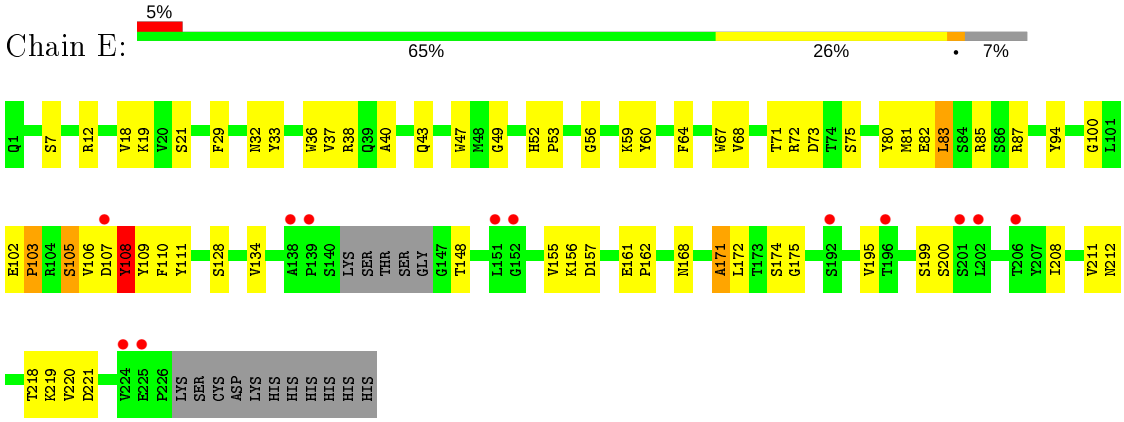
- Molecule 1: CH67 heavy chain

Chain C:



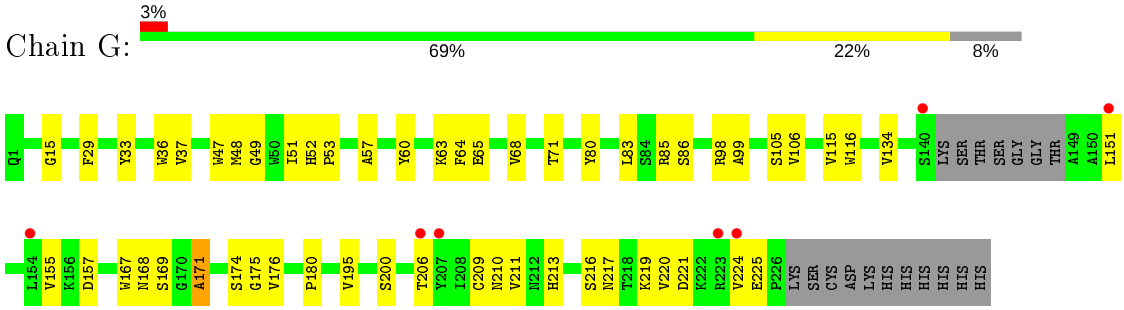
- Molecule 1: CH67 heavy chain

Chain E:



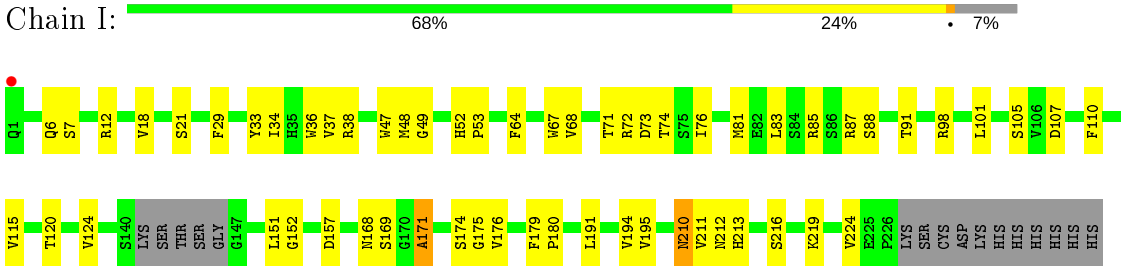
- Molecule 1: CH67 heavy chain

Chain G:



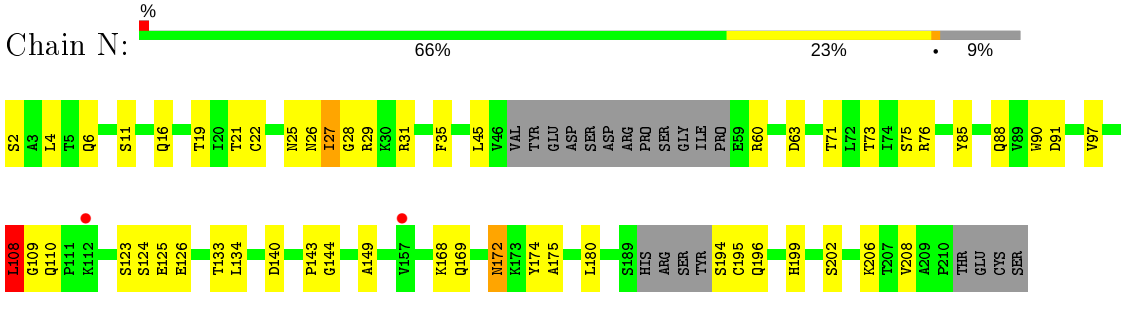
- Molecule 1: CH67 heavy chain

Chain I:



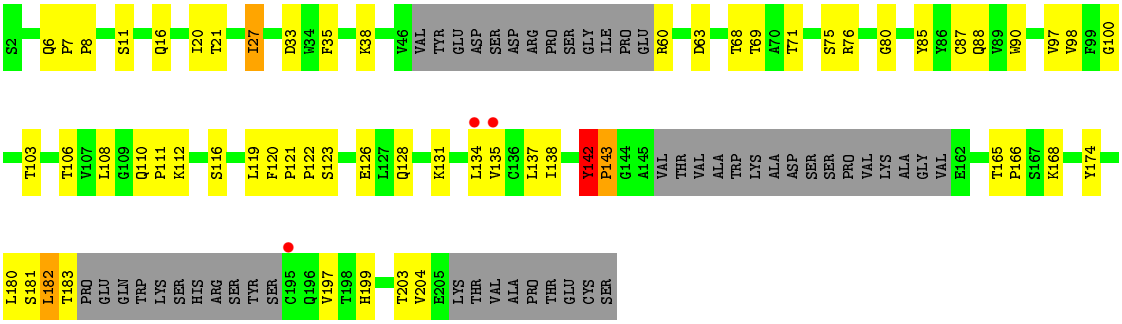
- Molecule 2: CH67 light chain

Chain N:

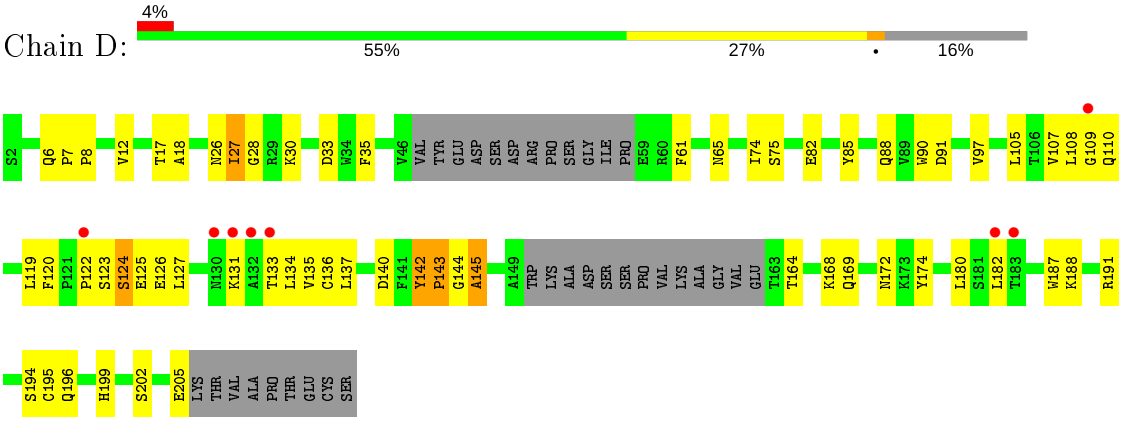


- Molecule 2: CH67 light chain

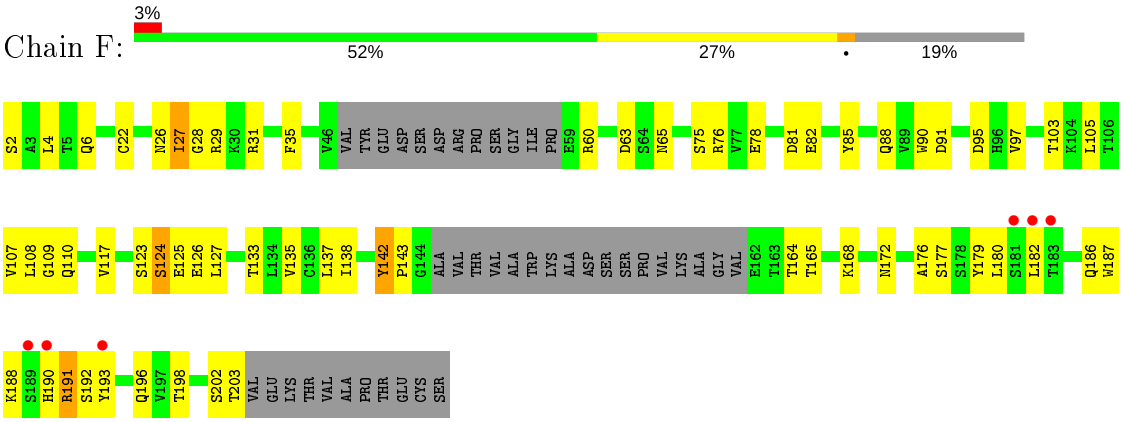
Chain B:



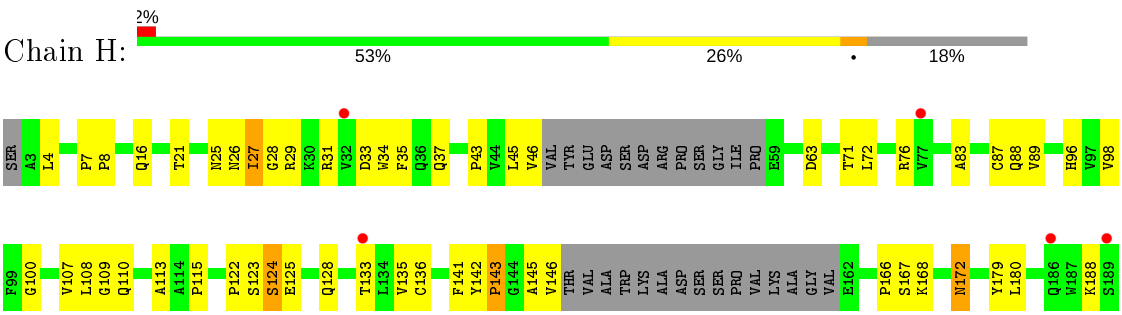
- Molecule 2: CH67 light chain

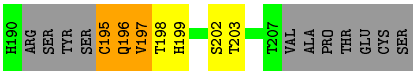


- Molecule 2: CH67 light chain



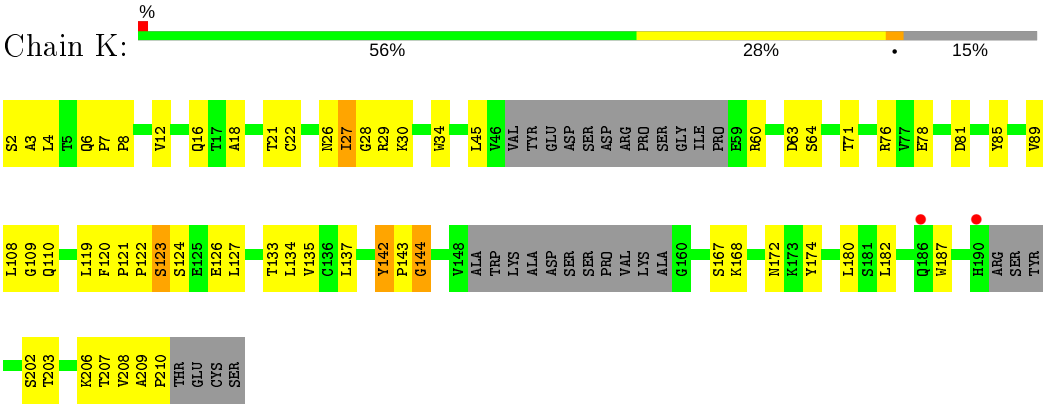
- Molecule 2: CH67 light chain





- Molecule 2: CH67 light chain

Chain K:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	113.25Å 123.27Å 228.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.70 – 3.60 47.94 – 3.55	Depositor EDS
% Data completeness (in resolution range)	94.9 (41.70-3.60) 91.9 (47.94-3.55)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.26 (at 3.57Å)	Xtriage
Refinement program	PHENIX 1.7.3_928	Depositor
R, R_{free}	0.258 , 0.306 0.257 , 0.295	Depositor DCC
R_{free} test set	2005 reflections (5.20%)	wwPDB-VP
Wilson B-factor (Å^2)	89.7	Xtriage
Anisotropy	0.213	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.22 , 56.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	17971	wwPDB-VP
Average B, all atoms (Å^2)	122.0	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 33.56 % 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 7.8460e-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	0.35	1/1720 (0.1%)	0.69	2/2348 (0.1%)
1	C	0.32	0/1706	0.68	2/2328 (0.1%)
1	E	0.33	1/1720 (0.1%)	0.62	2/2348 (0.1%)
1	G	0.31	0/1709	0.59	0/2333
1	I	0.30	0/1720	0.62	2/2348 (0.1%)
1	J	0.28	0/1710	0.58	1/2335 (0.0%)
2	B	0.35	1/1239 (0.1%)	0.69	1/1689 (0.1%)
2	D	0.33	1/1370 (0.1%)	0.65	0/1870
2	F	0.32	0/1332	0.63	0/1816
2	H	0.26	0/1333	0.57	0/1817
2	K	0.31	0/1384	0.62	2/1889 (0.1%)
2	N	0.31	0/1464	0.59	0/2000
All	All	0.32	4/18407 (0.0%)	0.63	12/25121 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	103	PRO	N-CD	5.34	1.55	1.47
2	D	122	PRO	N-CD	5.23	1.55	1.47
2	B	143	PRO	N-CD	5.20	1.55	1.47
1	E	103	PRO	N-CD	5.19	1.55	1.47

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	J	34	ILE	CG1-CB-CG2	-5.99	98.22	111.40
1	A	191	LEU	CA-CB-CG	5.92	128.92	115.30
2	K	142	TYR	C-N-CD	-5.69	108.08	120.60
1	E	102	GLU	C-N-CD	5.68	140.32	128.40
2	K	144	GLY	N-CA-C	5.67	127.27	113.10

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	1675	0	1623	58	0
1	C	1662	0	1611	47	0
1	E	1675	0	1623	61	0
1	G	1664	0	1613	41	0
1	I	1675	0	1623	53	0
1	J	1665	0	1617	70	0
2	B	1216	0	1167	83	0
2	D	1341	0	1288	66	0
2	F	1303	0	1242	48	0
2	H	1306	0	1255	60	0
2	K	1356	0	1307	52	0
2	N	1433	0	1387	40	0
All	All	17971	0	17356	617	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 18.

The worst 5 of 617 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)
2:B:182:LEU:CD1	2:B:183:THR:H	1.19	1.52
2:B:182:LEU:HD12	2:B:183:THR:N	1.12	1.43
1:J:160:PRO:C	1:J:162:PRO:HD2	1.47	1.33
2:D:142:TYR:CD2	2:D:143:PRO:HA	1.61	1.33
2:B:180:LEU:CD1	2:B:181:SER:O	1.82	1.28

There are no symmetry-related clashes.

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	216/236 (92%)	199 (92%)	14 (6%)	3 (1%)	11	48
1	C	214/236 (91%)	195 (91%)	16 (8%)	3 (1%)	11	48
1	E	216/236 (92%)	194 (90%)	19 (9%)	3 (1%)	11	48
1	G	214/236 (91%)	196 (92%)	16 (8%)	2 (1%)	17	57
1	I	216/236 (92%)	199 (92%)	15 (7%)	2 (1%)	17	57
1	J	214/236 (91%)	193 (90%)	17 (8%)	4 (2%)	8	42
2	B	156/213 (73%)	139 (89%)	13 (8%)	4 (3%)	5	35
2	D	173/213 (81%)	149 (86%)	17 (10%)	7 (4%)	3	26
2	F	167/213 (78%)	143 (86%)	18 (11%)	6 (4%)	3	29
2	H	166/213 (78%)	145 (87%)	14 (8%)	7 (4%)	3	25
2	K	174/213 (82%)	154 (88%)	16 (9%)	4 (2%)	6	38
2	N	187/213 (88%)	163 (87%)	20 (11%)	4 (2%)	7	40
All	All	2313/2694 (86%)	2069 (90%)	195 (8%)	49 (2%)	7	40

5 of 49 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	J	161	GLU
1	A	153	CYS
2	B	142	TYR
2	B	182	LEU
2	D	191	ARG

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	185/200 (92%)	185 (100%)	0	100 100
1	C	183/200 (92%)	183 (100%)	0	100 100
1	E	185/200 (92%)	183 (99%)	2 (1%)	73 88
1	G	184/200 (92%)	184 (100%)	0	100 100
1	I	185/200 (92%)	184 (100%)	1 (0%)	88 95
1	J	184/200 (92%)	183 (100%)	1 (0%)	88 95
2	B	137/180 (76%)	137 (100%)	0	100 100
2	D	151/180 (84%)	149 (99%)	2 (1%)	69 86
2	F	147/180 (82%)	147 (100%)	0	100 100
2	H	147/180 (82%)	143 (97%)	4 (3%)	44 73
2	K	153/180 (85%)	153 (100%)	0	100 100
2	N	161/180 (89%)	158 (98%)	3 (2%)	57 80
All	All	2002/2280 (88%)	1989 (99%)	13 (1%)	86 94

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	N	108	LEU
2	N	172	ASN
2	H	29	ARG
2	N	25	ASN
2	H	25	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
2	N	171	ASN
2	B	6	GLN
2	H	169	GLN
2	N	196	GLN
2	N	199	HIS

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 carbohydrates 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, 95^{th} 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	220/236 (93%)	-0.24	3 (1%)	75	61	48, 92, 208, 272	0
1	C	218/236 (92%)	-0.10	11 (5%)	28	18	42, 102, 217, 358	0
1	E	220/236 (93%)	-0.04	12 (5%)	25	15	25, 110, 239, 350	0
1	G	218/236 (92%)	-0.13	7 (3%)	47	32	43, 101, 217, 323	0
1	I	220/236 (93%)	-0.40	1 (0%)	91	83	46, 87, 165, 256	0
1	J	218/236 (92%)	-0.32	2 (0%)	84	73	46, 93, 178, 261	0
2	B	164/213 (76%)	-0.18	3 (1%)	68	53	61, 130, 216, 258	0
2	D	179/213 (84%)	-0.04	8 (4%)	33	21	53, 136, 219, 309	0
2	F	173/213 (81%)	-0.17	6 (3%)	44	29	48, 129, 241, 351	0
2	H	174/213 (81%)	-0.01	5 (2%)	51	35	66, 145, 263, 365	0
2	K	182/213 (85%)	-0.22	2 (1%)	80	68	62, 119, 183, 266	0
2	N	193/213 (90%)	-0.25	2 (1%)	82	70	60, 126, 187, 296	0
All	All	2379/2694 (88%)	-0.18	62 (2%)	56	40	25, 113, 220, 365	0

The worst 5 of 62 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	206	THR	7.3
2	D	183	THR	5.0
2	F	183	THR	4.9
2	D	131	LYS	4.9
2	D	132	ALA	4.7

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 carbohydrates 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.