



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 30, 2021 – 02:51 PM EDT

PDB ID : 3MM0
Title : Crystal structure of chimeric avidin
Authors : Livnah, O.; Eisenberg-Domovich, Y.; Maatta, J.A.E.; Kulomaa, M.S.; Hytonen, V.P.; Nordlund, H.R.
Deposited on : 2010-04-19
Resolution : 2.70 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
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)	:	2.23.2

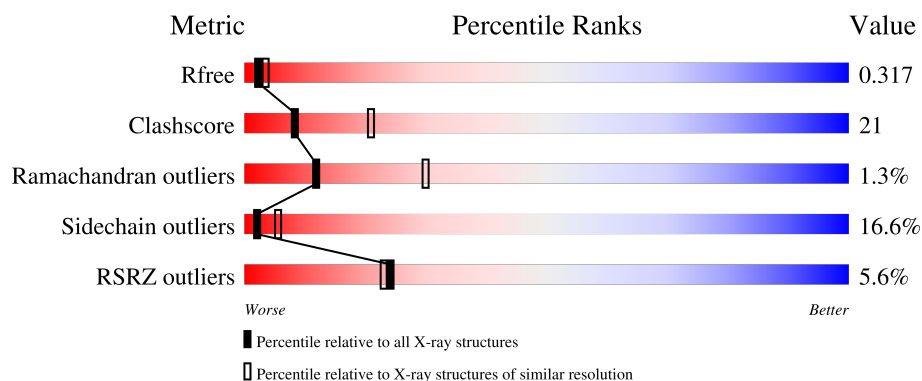
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	129	
1	B	129	
1	C	129	
1	D	129	
1	E	129	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	129	<div><div>5%</div><div><div></div><div>39%</div><div>46%</div><div>6%</div><div>9%</div></div></div>
1	G	129	<div><div>7%</div><div><div></div><div>49%</div><div>33%</div><div>9%</div><div>9%</div></div></div>
1	H	129	<div><div>8%</div><div><div></div><div>50%</div><div>32%</div><div>9%</div><div>10%</div></div></div>
1	I	129	<div><div>%</div><div><div></div><div>46%</div><div>40%</div><div>5%</div><div>9%</div></div></div>
1	K	129	<div><div>2%</div><div><div></div><div>67%</div><div>19%</div><div>6%</div><div>8%</div></div></div>
1	M	129	<div><div>%</div><div><div></div><div>47%</div><div>40%</div><div>••</div><div>10%</div></div></div>
1	N	129	<div><div>22%</div><div><div></div><div>62%</div><div>24%</div><div>5%</div><div>9%</div></div></div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11014 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 Avidin, Avidin-related protein 4/5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	115	Total	C	N	O	S	0	0	0
			892	567	153	168	4			
1	B	115	Total	C	N	O	S	0	0	0
			892	567	153	168	4			
1	C	118	Total	C	N	O	S	0	0	0
			920	583	161	172	4			
1	D	119	Total	C	N	O	S	0	0	0
			928	587	162	175	4			
1	E	116	Total	C	N	O	S	0	0	0
			900	571	155	170	4			
1	F	118	Total	C	N	O	S	0	0	0
			920	583	160	173	4			
1	G	117	Total	C	N	O	S	0	0	0
			912	579	158	171	4			
1	H	116	Total	C	N	O	S	0	0	0
			901	573	155	169	4			
1	I	118	Total	C	N	O	S	0	0	0
			920	583	160	173	4			
1	K	119	Total	C	N	O	S	0	0	0
			928	587	162	175	4			
1	M	116	Total	C	N	O	S	0	0	0
			901	573	155	169	4			
1	N	117	Total	C	N	O	S	0	0	0
			909	577	157	171	4			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLN	-	expression tag	UNP P02701
A	-1	THR	-	expression tag	UNP P02701
A	0	VAL	-	expression tag	UNP P02701
A	117	TYR	ILE	engineered mutation	UNP P02701
B	-2	GLN	-	expression tag	UNP P02701

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	THR	-	expression tag	UNP P02701
B	0	VAL	-	expression tag	UNP P02701
B	117	TYR	ILE	engineered mutation	UNP P02701
C	-2	GLN	-	expression tag	UNP P02701
C	-1	THR	-	expression tag	UNP P02701
C	0	VAL	-	expression tag	UNP P02701
C	117	TYR	ILE	engineered mutation	UNP P02701
D	-2	GLN	-	expression tag	UNP P02701
D	-1	THR	-	expression tag	UNP P02701
D	0	VAL	-	expression tag	UNP P02701
D	117	TYR	ILE	engineered mutation	UNP P02701
E	-2	GLN	-	expression tag	UNP P02701
E	-1	THR	-	expression tag	UNP P02701
E	0	VAL	-	expression tag	UNP P02701
E	117	TYR	ILE	engineered mutation	UNP P02701
F	-2	GLN	-	expression tag	UNP P02701
F	-1	THR	-	expression tag	UNP P02701
F	0	VAL	-	expression tag	UNP P02701
F	117	TYR	ILE	engineered mutation	UNP P02701
G	-2	GLN	-	expression tag	UNP P02701
G	-1	THR	-	expression tag	UNP P02701
G	0	VAL	-	expression tag	UNP P02701
G	117	TYR	ILE	engineered mutation	UNP P02701
H	-2	GLN	-	expression tag	UNP P02701
H	-1	THR	-	expression tag	UNP P02701
H	0	VAL	-	expression tag	UNP P02701
H	117	TYR	ILE	engineered mutation	UNP P02701
I	-2	GLN	-	expression tag	UNP P02701
I	-1	THR	-	expression tag	UNP P02701
I	0	VAL	-	expression tag	UNP P02701
I	117	TYR	ILE	engineered mutation	UNP P02701
K	-2	GLN	-	expression tag	UNP P02701
K	-1	THR	-	expression tag	UNP P02701
K	0	VAL	-	expression tag	UNP P02701
K	117	TYR	ILE	engineered mutation	UNP P02701
M	-2	GLN	-	expression tag	UNP P02701
M	-1	THR	-	expression tag	UNP P02701
M	0	VAL	-	expression tag	UNP P02701
M	117	TYR	ILE	engineered mutation	UNP P02701
N	-2	GLN	-	expression tag	UNP P02701
N	-1	THR	-	expression tag	UNP P02701
N	0	VAL	-	expression tag	UNP P02701

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
N	117	TYR	ILE	engineered mutation	UNP P02701

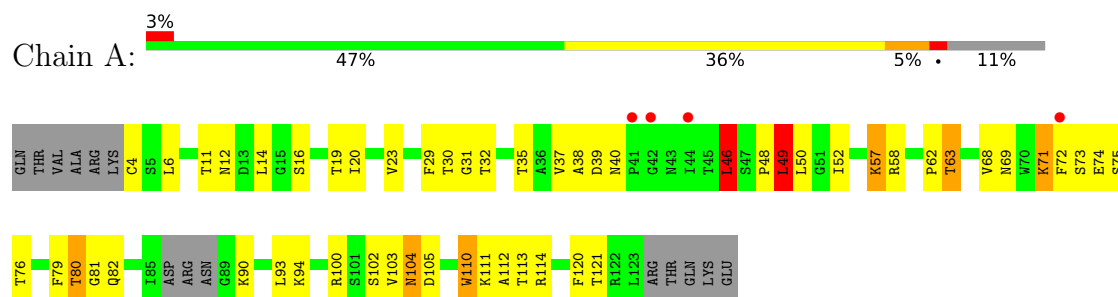
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	8	Total O 8 8	0	0
2	B	6	Total O 6 6	0	0
2	C	9	Total O 9 9	0	0
2	D	7	Total O 7 7	0	0
2	E	8	Total O 8 8	0	0
2	F	9	Total O 9 9	0	0
2	G	10	Total O 10 10	0	0
2	H	4	Total O 4 4	0	0
2	I	14	Total O 14 14	0	0
2	K	10	Total O 10 10	0	0
2	M	5	Total O 5 5	0	0
2	N	1	Total O 1 1	0	0

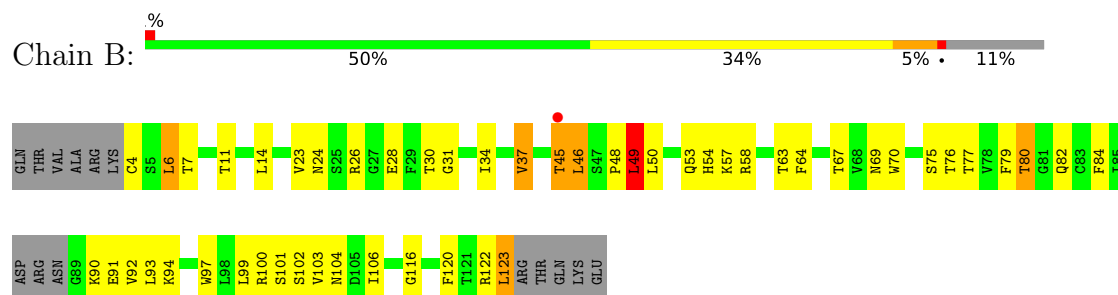
3 Residue-property plots [i](#)

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.

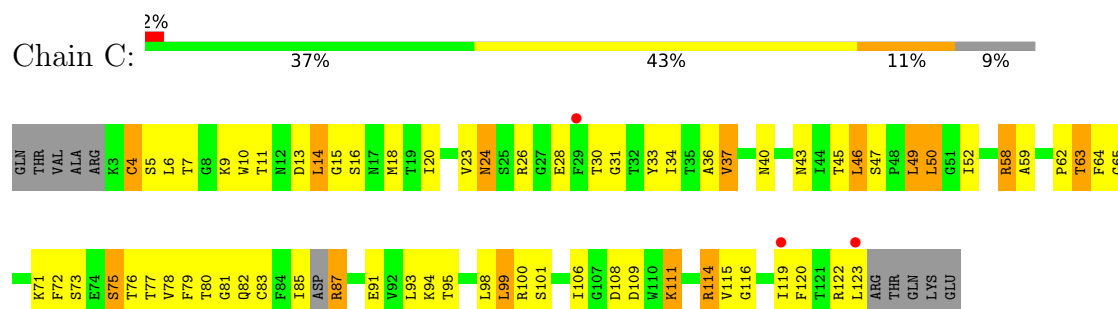
- Molecule 1: Avidin, Avidin-related protein 4/5



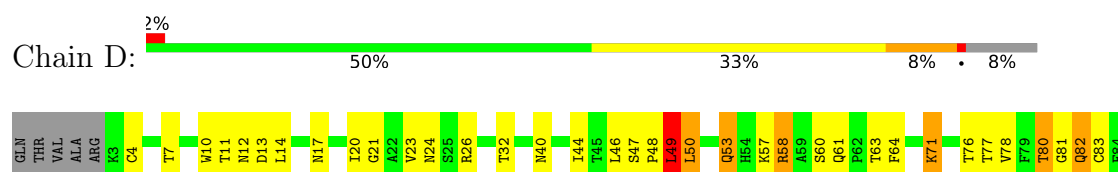
- Molecule 1: Avidin, Avidin-related protein 4/5



- Molecule 1: Avidin, Avidin-related protein 4/5

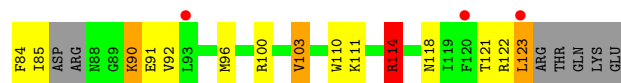


- Molecule 1: Avidin, Avidin-related protein 4/5

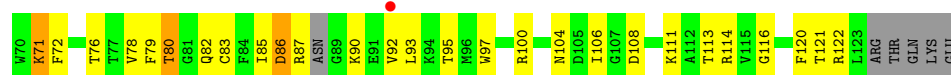
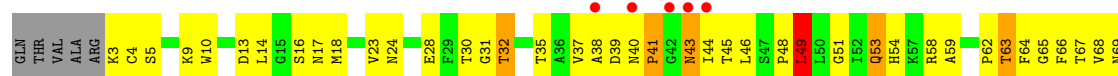




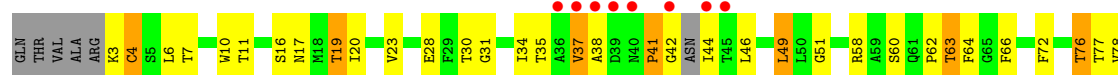
- Molecule 1: Avidin, Avidin-related protein 4/5



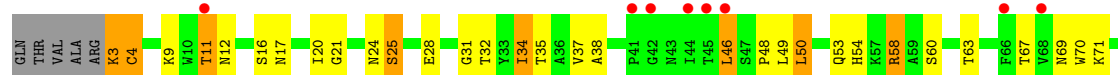
- Molecule 1: Avidin, Avidin-related protein 4/5



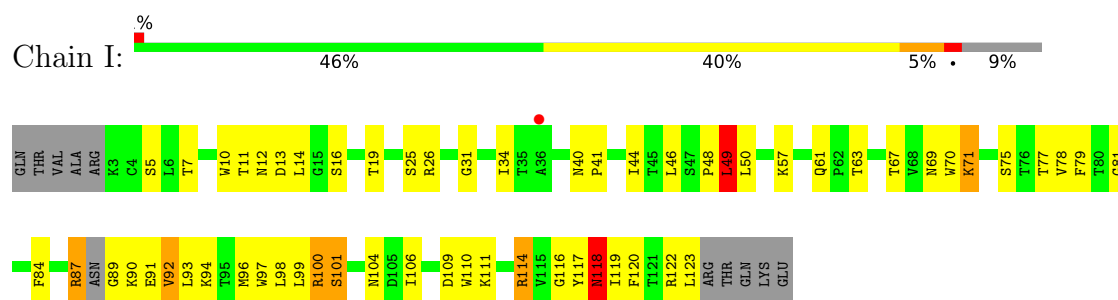
- Molecule 1: Avidin, Avidin-related protein 4/5



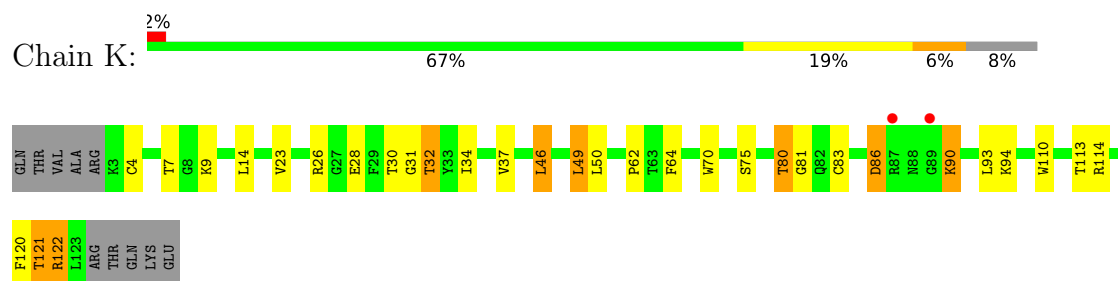
- Molecule 1: Avidin, Avidin-related protein 4/5



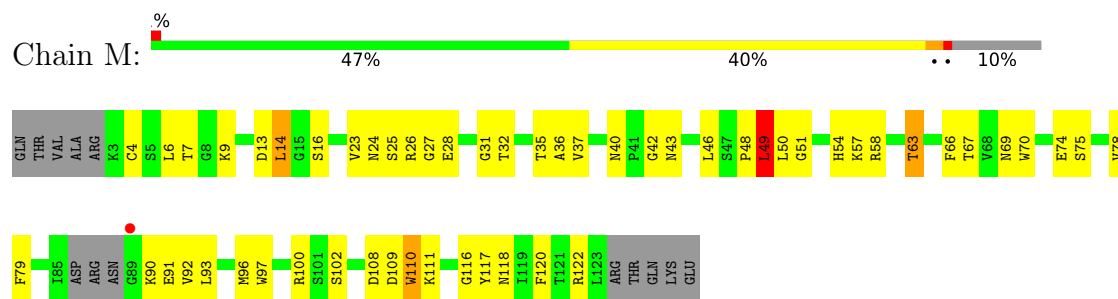
- Molecule 1: Avidin, Avidin-related protein 4/5



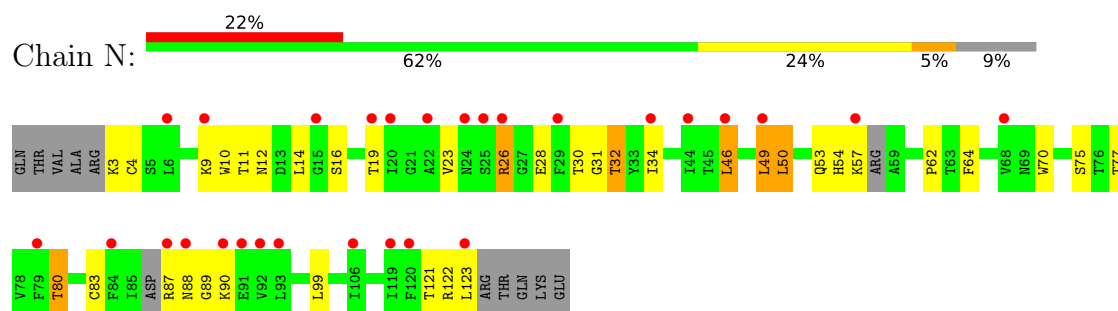
- Molecule 1: Avidin, Avidin-related protein 4/5



- Molecule 1: Avidin, Avidin-related protein 4/5



- Molecule 1: Avidin, Avidin-related protein 4/5



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	158.66Å 124.79Å 88.89Å 90.00° 114.45° 90.00°	Depositor
Resolution (Å)	50.00 – 2.70 49.41 – 2.69	Depositor EDS
% Data completeness (in resolution range)	98.0 (50.00-2.70) 97.5 (49.41-2.69)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.08 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.222 , 0.309 0.230 , 0.317	Depositor DCC
R_{free} test set	2149 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	52.3	Xtriage
Anisotropy	0.615	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 72.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	11014	wwPDB-VP
Average B, all atoms (Å ²)	80.0	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 42.18 % 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 2.1174e-04. 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	0.78	0/911	0.91	2/1235 (0.2%)
1	B	0.78	0/911	0.84	1/1235 (0.1%)
1	C	0.68	0/939	0.86	1/1271 (0.1%)
1	D	0.68	0/948	0.82	1/1285 (0.1%)
1	E	0.66	0/919	0.87	4/1246 (0.3%)
1	F	0.83	0/939	0.90	1/1271 (0.1%)
1	G	0.80	0/930	0.92	0/1257
1	H	0.74	0/920	0.88	3/1246 (0.2%)
1	I	0.80	0/939	0.94	2/1271 (0.2%)
1	K	0.68	0/948	0.86	2/1285 (0.2%)
1	M	0.75	0/920	0.90	1/1246 (0.1%)
1	N	0.55	0/927	0.80	3/1254 (0.2%)
All	All	0.73	0/11151	0.88	21/15102 (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	G	0	1

There are no bond length outliers.

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	N	50	LEU	CA-CB-CG	11.32	141.34	115.30
1	K	50	LEU	CA-CB-CG	9.96	138.21	115.30
1	E	49	LEU	CA-CB-CG	8.29	134.37	115.30
1	C	49	LEU	CA-CB-CG	8.10	133.93	115.30
1	I	49	LEU	CA-CB-CG	8.08	133.87	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	G	16	SER	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	892	0	876	48	0
1	B	892	0	876	61	0
1	C	920	0	908	71	0
1	D	928	0	913	49	0
1	E	900	0	882	31	0
1	F	920	0	906	65	0
1	G	912	0	899	58	0
1	H	901	0	889	38	0
1	I	920	0	906	44	0
1	K	928	0	913	17	0
1	M	901	0	889	32	0
1	N	909	0	894	27	0
2	A	8	0	0	1	0
2	B	6	0	0	0	0
2	C	9	0	0	1	0
2	D	7	0	0	0	0
2	E	8	0	0	3	0
2	F	9	0	0	1	0
2	G	10	0	0	2	0
2	H	4	0	0	0	0
2	I	14	0	0	4	0
2	K	10	0	0	1	0
2	M	5	0	0	2	0
2	N	1	0	0	0	0
All	All	11014	0	10751	454	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 21.

The worst 5 of 454 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:41:PRO:HB2	1:F:44:ILE:HG13	1.18	1.12
1:N:11:THR:HB	1:N:123:LEU:HD11	1.36	1.05
1:F:82:GLN:HE21	1:G:100:ARG:NH2	1.57	1.00
1:F:38:ALA:HB3	1:F:41:PRO:HG3	1.43	1.00
1:H:58:ARG:HH11	1:H:58:ARG:HB3	1.25	0.99

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	111/129 (86%)	102 (92%)	8 (7%)	1 (1%)	17	40
1	B	111/129 (86%)	105 (95%)	5 (4%)	1 (1%)	17	40
1	C	114/129 (88%)	108 (95%)	4 (4%)	2 (2%)	8	21
1	D	117/129 (91%)	105 (90%)	9 (8%)	3 (3%)	5	13
1	E	112/129 (87%)	105 (94%)	7 (6%)	0	100	100
1	F	114/129 (88%)	108 (95%)	4 (4%)	2 (2%)	8	21
1	G	111/129 (86%)	101 (91%)	8 (7%)	2 (2%)	8	21
1	H	112/129 (87%)	110 (98%)	2 (2%)	0	100	100
1	I	114/129 (88%)	105 (92%)	7 (6%)	2 (2%)	8	21
1	K	117/129 (91%)	110 (94%)	7 (6%)	0	100	100
1	M	112/129 (87%)	101 (90%)	7 (6%)	4 (4%)	3	7
1	N	111/129 (86%)	106 (96%)	4 (4%)	1 (1%)	17	40
All	All	1356/1548 (88%)	1266 (93%)	72 (5%)	18 (1%)	12	30

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	13	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	F	41	PRO
1	I	106	ILE
1	F	13	ASP
1	G	41	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	98/111 (88%)	82 (84%)	16 (16%)	2	6
1	B	98/111 (88%)	84 (86%)	14 (14%)	3	8
1	C	101/111 (91%)	84 (83%)	17 (17%)	2	5
1	D	102/111 (92%)	82 (80%)	20 (20%)	1	3
1	E	99/111 (89%)	78 (79%)	21 (21%)	1	3
1	F	101/111 (91%)	82 (81%)	19 (19%)	1	4
1	G	100/111 (90%)	84 (84%)	16 (16%)	2	6
1	H	99/111 (89%)	79 (80%)	20 (20%)	1	3
1	I	101/111 (91%)	86 (85%)	15 (15%)	3	7
1	K	102/111 (92%)	88 (86%)	14 (14%)	3	8
1	M	99/111 (89%)	87 (88%)	12 (12%)	5	11
1	N	100/111 (90%)	85 (85%)	15 (15%)	3	7
All	All	1200/1332 (90%)	1001 (83%)	199 (17%)	2	5

5 of 199 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	80	THR
1	I	5	SER
1	G	104	ASN
1	H	46	LEU
1	I	78	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
1	K	118	ASN
1	N	118	ASN
1	M	43	ASN
1	M	118	ASN
1	D	82	GLN

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 monosaccharides 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	115/129 (89%)	0.01	4 (3%) 44 44	45, 73, 95, 100	0
1	B	115/129 (89%)	-0.02	1 (0%) 84 85	40, 71, 95, 107	0
1	C	118/129 (91%)	0.10	3 (2%) 57 59	48, 87, 115, 137	0
1	D	119/129 (92%)	0.04	3 (2%) 57 59	47, 82, 118, 141	0
1	E	116/129 (89%)	0.33	10 (8%) 10 8	49, 81, 116, 134	0
1	F	118/129 (91%)	0.12	6 (5%) 28 26	43, 71, 99, 111	0
1	G	117/129 (90%)	0.07	9 (7%) 13 11	44, 70, 98, 115	0
1	H	116/129 (89%)	0.48	10 (8%) 10 8	53, 86, 106, 115	0
1	I	118/129 (91%)	0.04	1 (0%) 86 87	45, 74, 95, 112	0
1	K	119/129 (92%)	0.08	2 (1%) 70 72	43, 71, 96, 112	0
1	M	116/129 (89%)	-0.05	1 (0%) 84 85	51, 72, 86, 102	0
1	N	117/129 (90%)	1.37	28 (23%) 0 0	67, 141, 204, 230	0
All	All	1404/1548 (90%)	0.21	78 (5%) 24 23	40, 77, 130, 230	0

The worst 5 of 78 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	N	87	ARG	8.9
1	N	88	ASN	6.4
1	N	123	LEU	6.0
1	H	68	VAL	5.6
1	H	44	ILE	5.5

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 monosaccharides 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.