



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 14, 2020 – 08:26 AM BST

PDB ID : 6TMS
Title : Crystal structure of a de novo designed hexameric helical-bundle protein
Authors : Xu, C.; Pei, X.Y.; Luisi, B.F.; Baker, D.
Deposited on : 2019-12-05
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
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.14.4.dev1
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) : 2.14.4.dev1

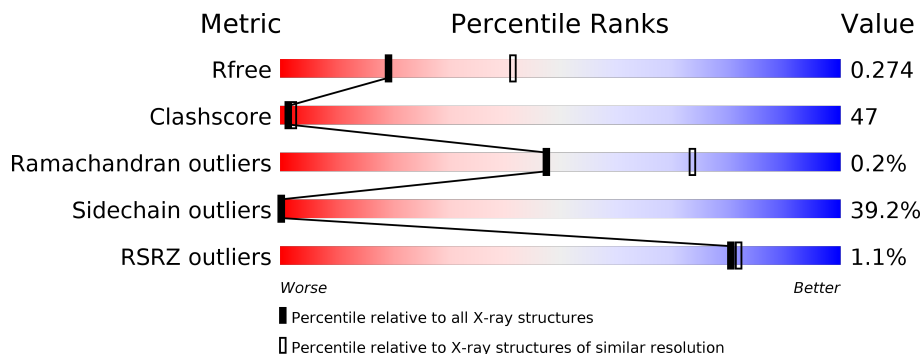
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 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 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 $\leq 5\%$. 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	69	
1	B	69	
1	C	69	
1	D	69	
1	E	69	
1	F	69	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	69	
1	I	69	
1	J	69	
1	K	69	
2	G	69	
2	L	69	
3	Q	16	
3	R	16	

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 7764 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 a novel designed pore protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	69	553	348	91	113	1	0	0	0
1	B	69	553	348	91	113	1	0	0	0
1	C	69	553	348	91	113	1	0	0	0
1	E	69	553	348	91	113	1	0	0	0
1	F	69	553	348	91	113	1	0	0	0
1	D	69	553	348	91	113	1	0	0	0
1	H	69	553	348	91	113	1	0	0	0
1	I	69	553	348	91	113	1	0	0	0
1	J	69	553	348	91	113	1	0	0	0
1	K	69	549	346	91	111	1	0	0	0

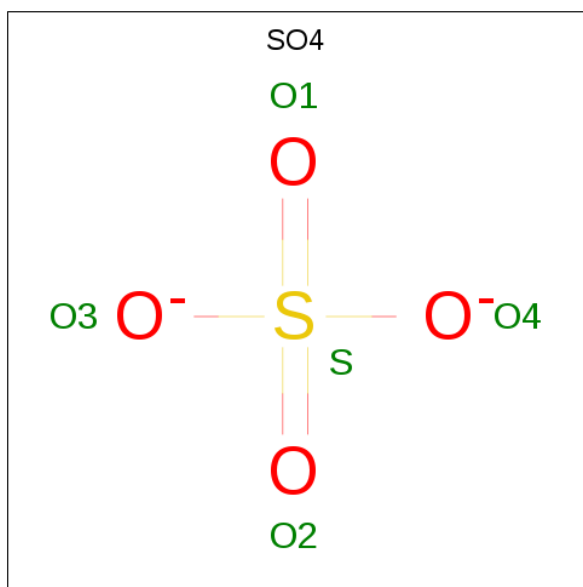
- Molecule 2 is a protein called a novel designed pore protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	G	69	551	348	89	113	1	0	0	0
2	L	69	551	348	89	113	1	0	0	0

- Molecule 3 is a protein called affinity purification tag.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	Q	3	Total	C	N	O	0	0	0
			14	8	3	3			
3	R	4	Total	C	N	O	0	0	0
			20	12	4	4			

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	O	S	0	0
			5	4	1		
4	F	1	Total	O	S	0	0
			5	4	1		
4	H	1	Total	O	S	0	0
			5	4	1		
4	I	1	Total	O	S	0	0
			5	4	1		
4	J	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	98	Total	O	0	0
			98	98		
5	B	100	Total	O	0	0
			100	100		
5	C	106	Total	O	0	0
			106	106		

Continued on next page...

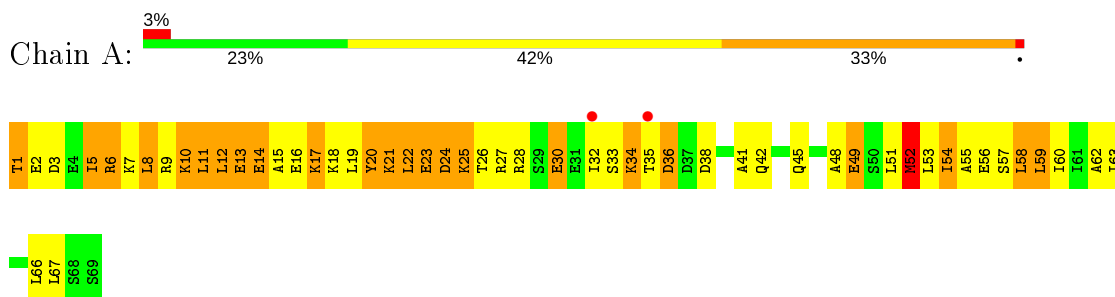
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	E	91	Total 91	O 91	0	0
5	F	102	Total 102	O 102	0	0
5	G	108	Total 108	O 108	0	0
5	D	82	Total 82	O 82	0	0
5	H	79	Total 79	O 79	0	0
5	I	77	Total 77	O 77	0	0
5	J	69	Total 69	O 69	0	0
5	K	75	Total 75	O 75	0	0
5	L	77	Total 77	O 77	0	0
5	Q	3	Total 3	O 3	0	0
5	R	10	Total 10	O 10	0	0

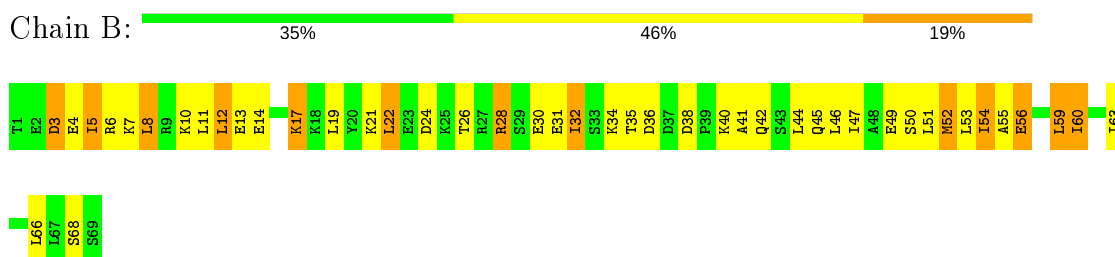
3 Residue-property plots [i](#)

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.

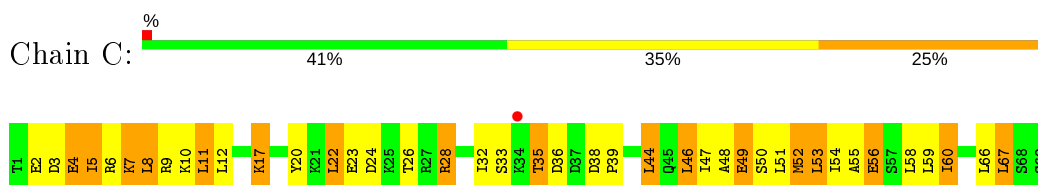
- Molecule 1: a novel designed pore protein



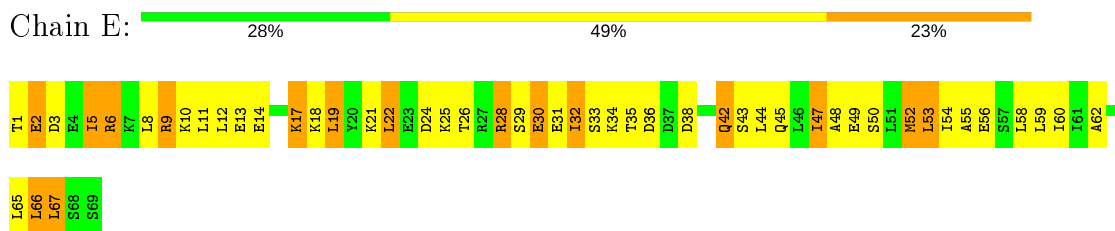
- Molecule 1: a novel designed pore protein



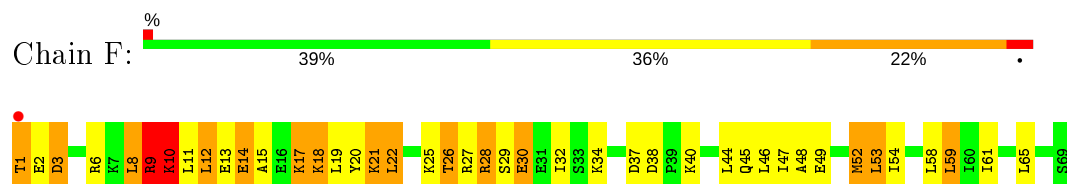
- Molecule 1: a novel designed pore protein



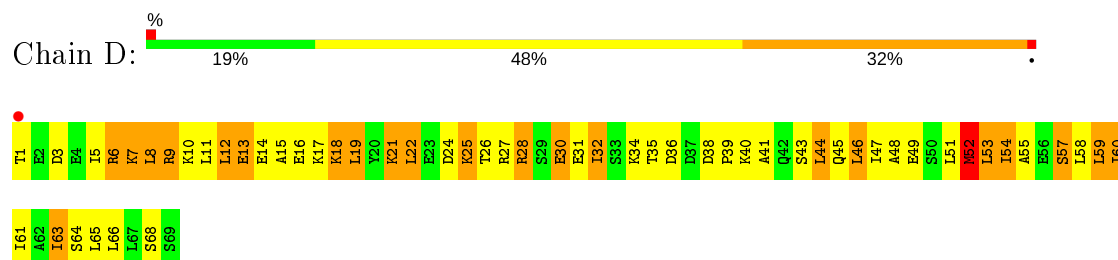
- Molecule 1: a novel designed pore protein



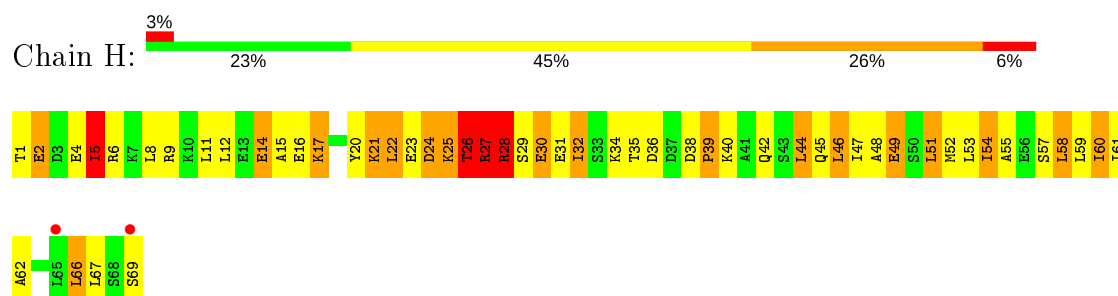
- Molecule 1: a novel designed pore protein



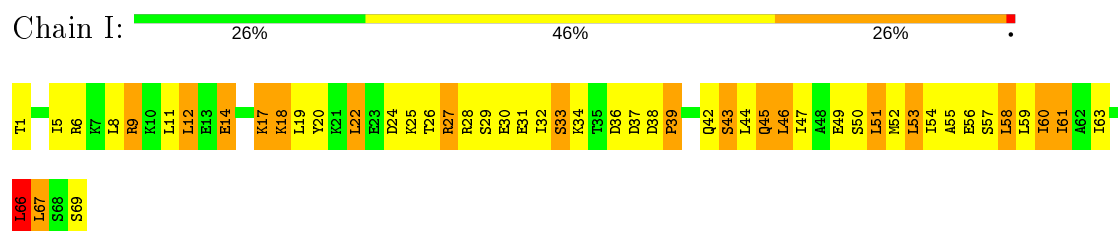
- Molecule 1: a novel designed pore protein



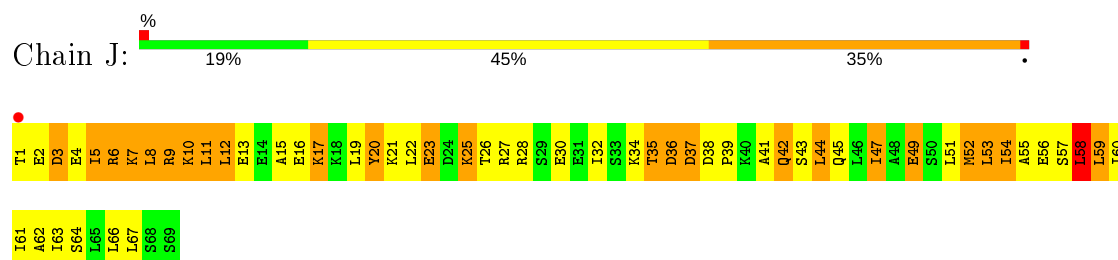
- Molecule 1: a novel designed pore protein



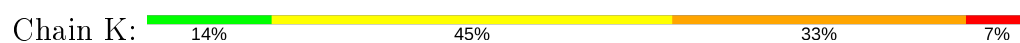
- Molecule 1: a novel designed pore protein

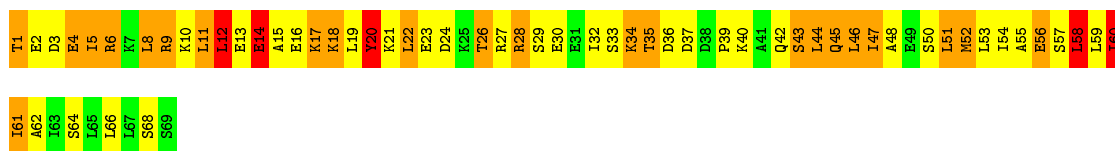


- Molecule 1: a novel designed pore protein



- Molecule 1: a novel designed pore protein

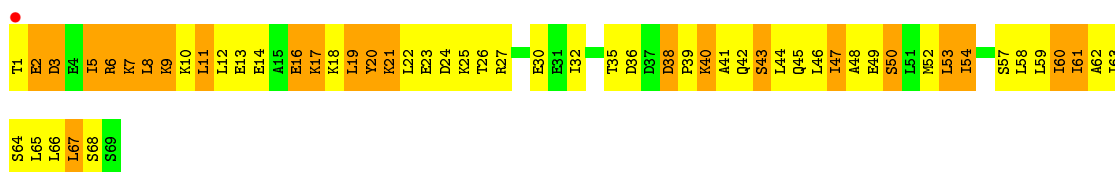
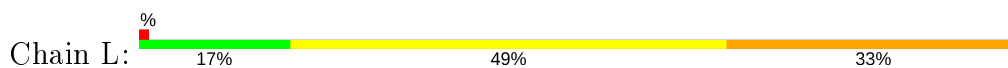




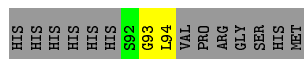
- Molecule 2: a novel designed pore protein



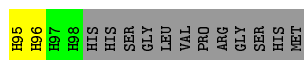
- Molecule 2: a novel designed pore protein



- Molecule 3: affinity purification tag



- Molecule 3: affinity purification tag



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	45.88Å 54.33Å 79.33Å 89.80° 90.08° 89.77°	Depositor
Resolution (Å)	15.99 – 2.70 54.33 – 2.50	Depositor EDS
% Data completeness (in resolution range)	95.7 (15.99-2.70) 98.0 (54.33-2.50)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.63 (at 2.51Å)	Xtrriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.288 , 0.299 0.250 , 0.274	Depositor DCC
R_{free} test set	1348 reflections (5.15%)	wwPDB-VP
Wilson B-factor (Å ²)	36.4	Xtrriage
Anisotropy	0.456	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.24 , 83.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.419 for h,-k,-l 0.400 for -h,k,-l 0.410 for -h,-k,l	Xtrriage
Reported twinning fraction	0.500 for h,-k,-l	Depositor
Outliers	0 of 26197 reflections	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	7764	wwPDB-VP
Average B, all atoms (Å ²)	47.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 40.43 % 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 2.7386e-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 i

5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	1.52	7/554 (1.3%)	1.57	15/740 (2.0%)
1	B	0.90	3/554 (0.5%)	1.19	3/740 (0.4%)
1	C	1.37	10/554 (1.8%)	1.33	6/740 (0.8%)
1	D	0.73	0/554	1.25	3/740 (0.4%)
1	E	0.76	1/554 (0.2%)	1.22	4/740 (0.5%)
1	F	0.91	2/554 (0.4%)	1.43	7/740 (0.9%)
1	H	1.08	3/554 (0.5%)	1.61	11/740 (1.5%)
1	I	1.07	5/554 (0.9%)	1.45	7/740 (0.9%)
1	J	0.75	1/554 (0.2%)	1.28	2/740 (0.3%)
1	K	1.04	2/550 (0.4%)	1.59	13/735 (1.8%)
2	G	0.79	1/552 (0.2%)	1.27	3/737 (0.4%)
2	L	0.81	0/552	1.33	0/737
3	Q	0.79	0/13	0.77	0/16
3	R	0.50	0/19	0.59	0/25
All	All	1.01	35/6672 (0.5%)	1.38	74/8910 (0.8%)

The worst 5 of 35 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	23	GLU	CD-OE1	-16.92	1.07	1.25
1	A	23	GLU	CD-OE2	-15.55	1.08	1.25
1	C	23	GLU	CD-OE1	-13.72	1.10	1.25
1	C	23	GLU	CD-OE2	-13.62	1.10	1.25
1	A	20	TYR	C-O	-11.79	1.00	1.23

The worst 5 of 74 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	28	ARG	NE-CZ-NH1	-11.54	114.53	120.30
1	H	28	ARG	CG-CD-NE	-10.08	90.64	111.80

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	I	66	LEU	CA-CB-CG	-9.39	93.70	115.30
1	I	38	ASP	CB-CG-OD2	8.96	126.37	118.30
1	F	28	ARG	NE-CZ-NH2	8.01	124.31	120.30

There are no chirality outliers.

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	553	0	596	71	0
1	B	553	0	596	48	0
1	C	553	0	596	37	0
1	D	553	0	596	104	1
1	E	553	0	596	42	1
1	F	553	0	596	59	0
1	H	553	0	596	81	2
1	I	553	0	596	56	1
1	J	553	0	596	90	1
1	K	549	0	592	94	0
2	G	551	0	596	54	0
2	L	551	0	596	106	0
3	Q	14	0	6	10	0
3	R	20	0	7	3	0
4	B	5	0	0	0	0
4	F	5	0	0	0	0
4	H	5	0	0	0	0
4	I	5	0	0	0	0
4	J	5	0	0	0	0
5	A	98	0	0	6	1
5	B	100	0	0	4	0
5	C	106	0	0	5	0
5	D	82	0	0	11	0
5	E	91	0	0	7	0
5	F	102	0	0	10	0
5	G	108	0	0	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	H	79	0	0	13	0
5	I	77	0	0	6	0
5	J	69	0	0	4	1
5	K	75	0	0	10	0
5	L	77	0	0	7	0
5	Q	3	0	0	0	0
5	R	10	0	0	1	0
All	All	7764	0	7161	652	4

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 47.

The worst 5 of 652 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:L:18:LYS:HD3	5:L:109:HOH:O	1.35	1.26
2:G:28:ARG:NH1	5:G:101:HOH:O	1.72	1.22
2:G:31:GLU:OE2	1:H:24:ASP:OD2	1.61	1.18
2:L:18:LYS:CD	5:L:109:HOH:O	1.88	1.10
1:H:31:GLU:OE1	5:H:201:HOH:O	1.66	1.09

All (4) 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:E:28:ARG:NH1	1:D:14:GLU:OE2[1_655]	1.82	0.38
1:H:1:THR:OG1	1:J:36:ASP:OD1[1_545]	2.10	0.10
1:H:1:THR:N	1:I:34:LYS:NZ[1_545]	2.14	0.06
5:A:122:HOH:O	5:J:246:HOH:O[1_556]	2.18	0.02

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	67/69 (97%)	65 (97%)	2 (3%)	0	100	100
1	B	67/69 (97%)	66 (98%)	1 (2%)	0	100	100
1	C	67/69 (97%)	64 (96%)	3 (4%)	0	100	100
1	D	67/69 (97%)	65 (97%)	2 (3%)	0	100	100
1	E	67/69 (97%)	65 (97%)	2 (3%)	0	100	100
1	F	67/69 (97%)	62 (92%)	5 (8%)	0	100	100
1	H	67/69 (97%)	64 (96%)	2 (3%)	1 (2%)	10	26
1	I	67/69 (97%)	63 (94%)	3 (4%)	1 (2%)	10	26
1	J	67/69 (97%)	64 (96%)	3 (4%)	0	100	100
1	K	67/69 (97%)	65 (97%)	2 (3%)	0	100	100
2	G	67/69 (97%)	66 (98%)	1 (2%)	0	100	100
2	L	67/69 (97%)	66 (98%)	1 (2%)	0	100	100
3	Q	1/16 (6%)	1 (100%)	0	0	100	100
3	R	2/16 (12%)	2 (100%)	0	0	100	100
All	All	807/860 (94%)	778 (96%)	27 (3%)	2 (0%)	47	73

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	39	PRO
1	I	39	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	64/64 (100%)	44 (69%)	20 (31%)	0	0
1	B	64/64 (100%)	39 (61%)	25 (39%)	0	0
1	C	64/64 (100%)	43 (67%)	21 (33%)	0	0
1	D	64/64 (100%)	37 (58%)	27 (42%)	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	64/64 (100%)	38 (59%)	26 (41%)	0	0
1	F	64/64 (100%)	45 (70%)	19 (30%)	0	1
1	H	64/64 (100%)	37 (58%)	27 (42%)	0	0
1	I	64/64 (100%)	38 (59%)	26 (41%)	0	0
1	J	64/64 (100%)	36 (56%)	28 (44%)	0	0
1	K	63/64 (98%)	31 (49%)	32 (51%)	0	0
2	G	64/64 (100%)	47 (73%)	17 (27%)	0	1
2	L	64/64 (100%)	31 (48%)	33 (52%)	0	0
All	All	767/768 (100%)	466 (61%)	301 (39%)	0	0

5 of 301 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	17	LYS
1	H	35	THR
2	L	16	GLU
1	D	22	LEU
1	D	59	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	42	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](#)

5 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	SO4	J	101	-	4,4,4	0.21	0	6,6,6	0.12	0
4	SO4	I	101	-	4,4,4	0.24	0	6,6,6	0.19	0
4	SO4	F	101	-	4,4,4	0.20	0	6,6,6	0.25	0
4	SO4	H	101	-	4,4,4	0.29	0	6,6,6	0.47	0
4	SO4	B	101	-	4,4,4	0.28	0	6,6,6	0.26	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	69/69 (100%)	-0.61	2 (2%) 51 52	18, 34, 64, 75	0
1	B	69/69 (100%)	-0.68	0 100 100	19, 41, 60, 65	0
1	C	69/69 (100%)	-0.42	1 (1%) 75 77	18, 40, 65, 71	0
1	D	69/69 (100%)	-0.36	1 (1%) 75 77	35, 49, 64, 83	0
1	E	69/69 (100%)	-0.52	0 100 100	18, 40, 62, 67	0
1	F	69/69 (100%)	-0.39	1 (1%) 75 77	18, 40, 70, 85	0
1	H	69/69 (100%)	-0.15	2 (2%) 51 52	33, 52, 74, 81	0
1	I	69/69 (100%)	-0.18	0 100 100	32, 52, 66, 68	0
1	J	69/69 (100%)	-0.26	1 (1%) 75 77	33, 48, 68, 79	0
1	K	69/69 (100%)	-0.27	0 100 100	28, 51, 71, 82	0
2	G	69/69 (100%)	-0.61	0 100 100	20, 37, 61, 67	0
2	L	69/69 (100%)	0.02	1 (1%) 75 77	28, 55, 70, 84	0
3	Q	3/16 (18%)	0.38	0 100 100	64, 64, 66, 69	0
3	R	4/16 (25%)	-0.09	0 100 100	71, 79, 84, 85	0
All	All	835/860 (97%)	-0.37	9 (1%) 80 82	18, 45, 69, 85	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	L	1	THR	5.3
1	F	1	THR	3.3
1	J	1	THR	3.3
1	A	35	THR	3.0
1	H	69	SER	2.4

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	SO4	I	101	5/5	0.95	0.17	46,46,53,86	0
4	SO4	J	101	5/5	0.97	0.13	43,43,43,43	0
4	SO4	F	101	5/5	0.98	0.10	41,41,46,51	0
4	SO4	H	101	5/5	0.98	0.13	29,29,29,29	0
4	SO4	B	101	5/5	0.98	0.24	32,32,32,32	0

6.5 Other polymers [i](#)

There are no such residues in this entry.