



wwPDB X-ray Structure Validation Summary Report

Jan 7, 2024 – 04:46 pm GMT

PDB ID : 6ERG
Title : Complex of XLF and heterodimer Ku bound to DNA
Authors : Nemoz, C.; Legrand, P.; Ropars, V.; Charbonnier, J.B.
Deposited on : 2017-10-18
Resolution : 2.90 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 : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
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) : 2.36

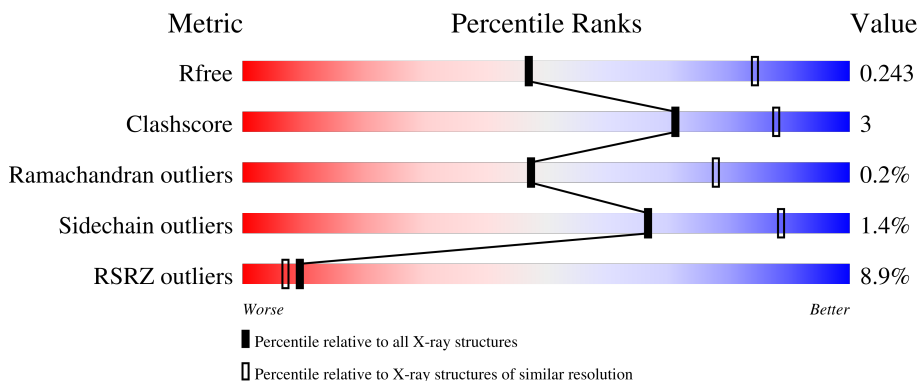
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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.



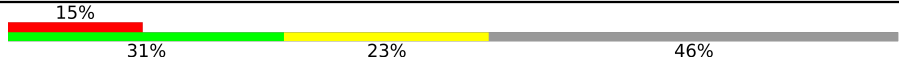




Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	544	
1	D	544	
2	B	572	
2	E	572	
3	C	13	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	13	
4	H	21	
4	K	21	
5	M	34	
5	R	34	

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 18997 atoms, of which 8 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 X-ray repair cross-complementing protein 6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	501	4048	2589	686	755	18	0	0	0
1	D	503	4066	2601	688	759	18	0	0	0

- Molecule 2 is a protein called X-ray repair cross-complementing protein 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	531	4278	2736	731	787	24	0	0	0
2	E	536	4315	2755	737	799	24	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-16	MET	-	initiating methionine	UNP P13010
B	-15	HIS	-	expression tag	UNP P13010
B	-14	HIS	-	expression tag	UNP P13010
B	-13	HIS	-	expression tag	UNP P13010
B	-12	HIS	-	expression tag	UNP P13010
B	-11	HIS	-	expression tag	UNP P13010
B	-10	HIS	-	expression tag	UNP P13010
B	-9	HIS	-	expression tag	UNP P13010
B	-8	HIS	-	expression tag	UNP P13010
B	-7	HIS	-	expression tag	UNP P13010
B	-6	HIS	-	expression tag	UNP P13010
B	-5	GLU	-	expression tag	UNP P13010
B	-4	ASN	-	expression tag	UNP P13010
B	-3	LEU	-	expression tag	UNP P13010
B	-2	TYR	-	expression tag	UNP P13010
B	-1	PHE	-	expression tag	UNP P13010

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	GLN	-	expression tag	UNP P13010
B	1	GLY	-	expression tag	UNP P13010
E	-16	MET	-	initiating methionine	UNP P13010
E	-15	HIS	-	expression tag	UNP P13010
E	-14	HIS	-	expression tag	UNP P13010
E	-13	HIS	-	expression tag	UNP P13010
E	-12	HIS	-	expression tag	UNP P13010
E	-11	HIS	-	expression tag	UNP P13010
E	-10	HIS	-	expression tag	UNP P13010
E	-9	HIS	-	expression tag	UNP P13010
E	-8	HIS	-	expression tag	UNP P13010
E	-7	HIS	-	expression tag	UNP P13010
E	-6	HIS	-	expression tag	UNP P13010
E	-5	GLU	-	expression tag	UNP P13010
E	-4	ASN	-	expression tag	UNP P13010
E	-3	LEU	-	expression tag	UNP P13010
E	-2	TYR	-	expression tag	UNP P13010
E	-1	PHE	-	expression tag	UNP P13010
E	0	GLN	-	expression tag	UNP P13010
E	1	GLY	-	expression tag	UNP P13010

- Molecule 3 is a protein called Non-homologous end-joining factor 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
3	C	7	Total	C	H	N	O	0	0	0
			61	37	4	11	9			
3	F	7	Total	C	H	N	O	0	0	0
			61	37	4	11	9			

- Molecule 4 is a DNA chain called DNA (21-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
4	H	21	Total	C	N	O	P	0	0	0
			431	208	71	132	20			
4	K	21	Total	C	N	O	P	0	0	0
			431	208	71	132	20			

- Molecule 5 is a DNA chain called DNA (34-MER).

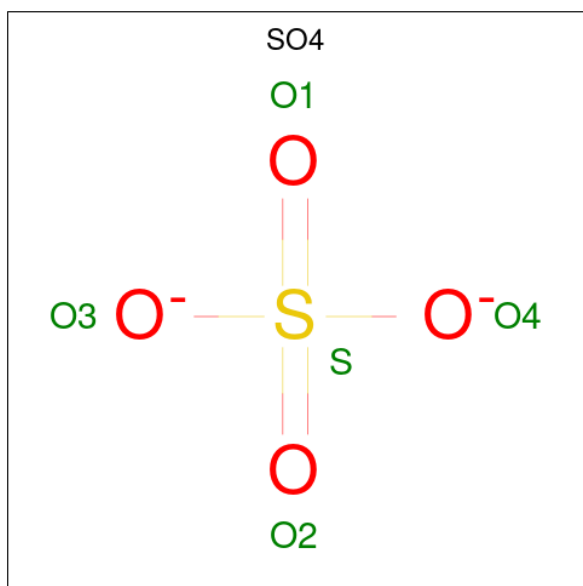
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
5	M	32	Total	C	N	O	P	0	0	0
			645	308	124	182	31			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
5	R	31	626	299	121	176	30	0	0	0

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
6	B	1	5	4	1	5	0
6	D	1	5	4	1	5	0
6	E	1	5	4	1	5	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	2	2	2	0	0
7	B	5	5	5	0	0
7	D	2	2	2	0	0
7	E	5	5	5	0	0
7	H	1	1	1	0	0

Continued on next page...

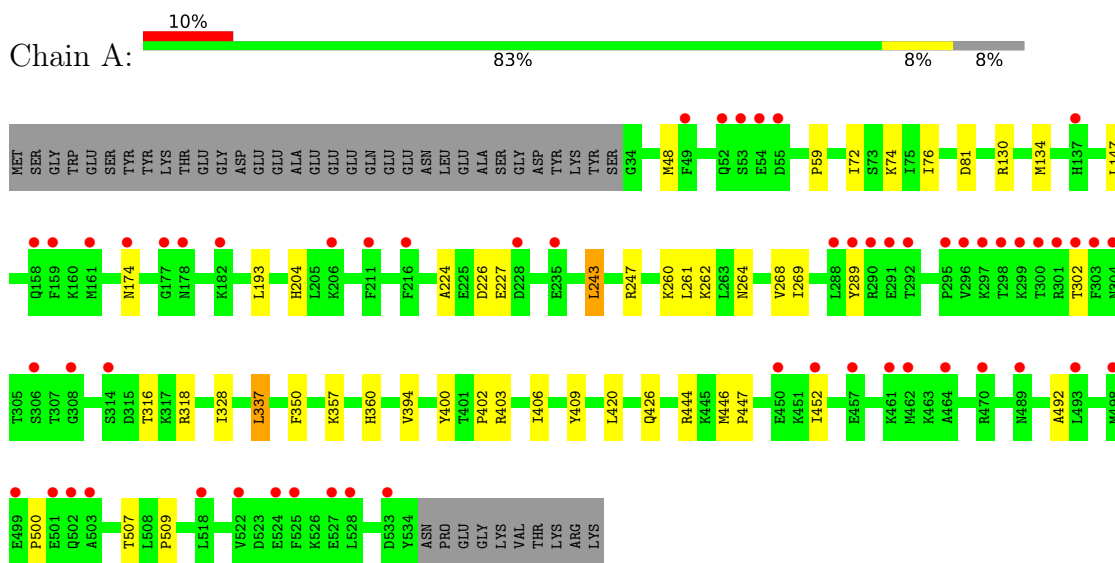
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	K	2	Total O 2 2	0	0
7	M	3	Total O 3 3	0	0

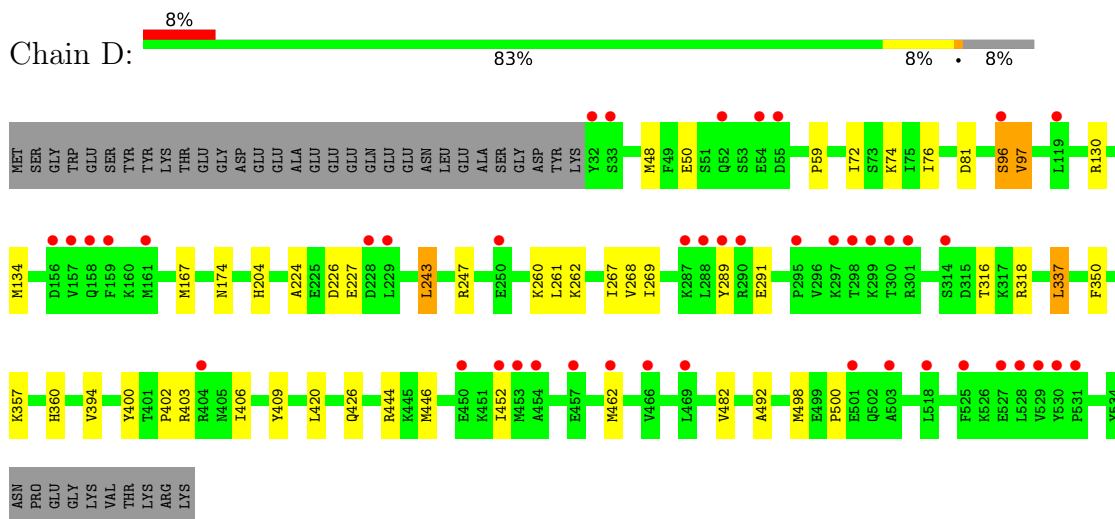
3 Residue-property plots [i](#)

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.

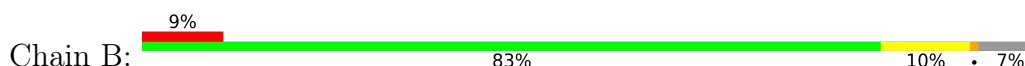
- Molecule 1: X-ray repair cross-complementing protein 6

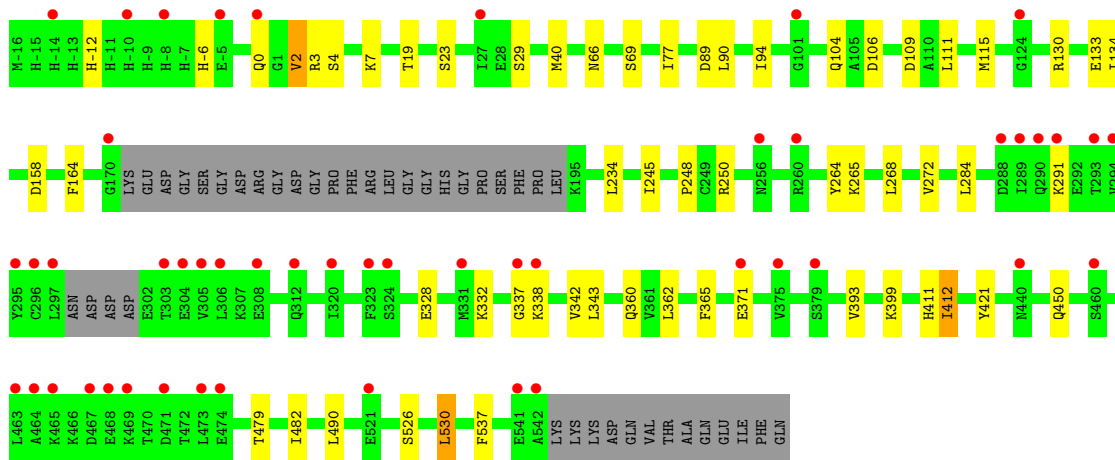


- Molecule 1: X-ray repair cross-complementing protein 6

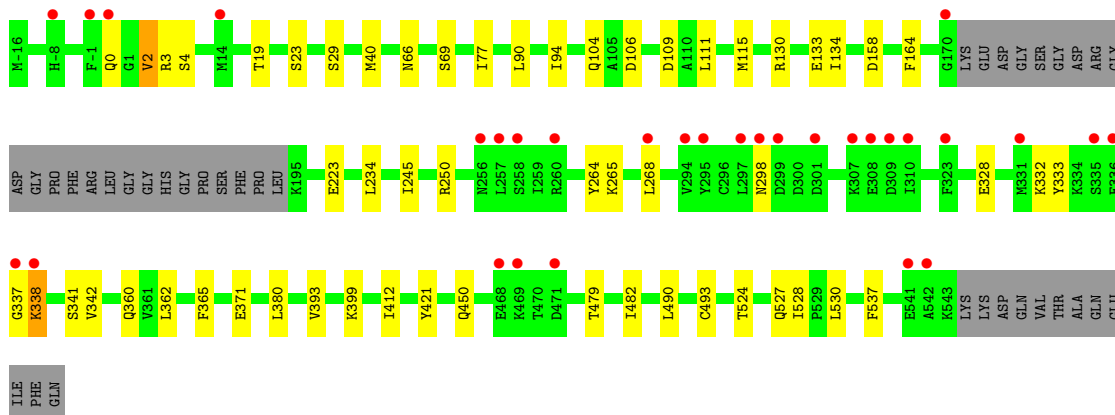
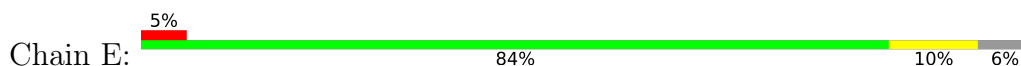


- Molecule 2: X-ray repair cross-complementing protein 5

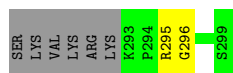




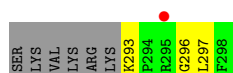
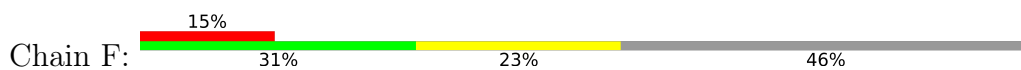
• Molecule 2: X-ray repair cross-complementing protein 5



• Molecule 3: Non-homologous end-joining factor 1



• Molecule 3: Non-homologous end-joining factor 1



• Molecule 4: DNA (21-MER)

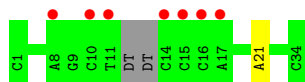
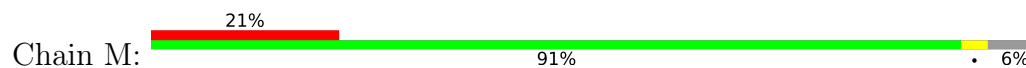




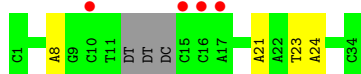
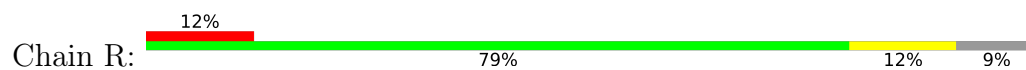
- Molecule 4: DNA (21-MER)



- Molecule 5: DNA (34-MER)



- Molecule 5: DNA (34-MER)



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	111.71Å 114.26Å 127.17Å 90.00° 93.14° 90.00°	Depositor
Resolution (Å)	48.76 – 2.90 47.62 – 2.90	Depositor EDS
% Data completeness (in resolution range)	72.9 (48.76-2.90) 72.9 (47.62-2.90)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.26 (at 2.91Å)	Xtrriage
Refinement program	BUSTER 2.10.3	Depositor
R, R_{free}	0.218 , 0.244 0.228 , 0.243	Depositor DCC
R_{free} test set	2605 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	80.8	Xtrriage
Anisotropy	0.049	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 51.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	18997	wwPDB-VP
Average B, all atoms (Å ²)	100.0	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 28.03 % 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 1.9798e-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 [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	0.34	0/4127	0.57	0/5560
1	D	0.35	0/4146	0.58	1/5586 (0.0%)
2	B	0.33	0/4370	0.60	2/5893 (0.0%)
2	E	0.34	0/4408	0.59	0/5947
3	C	0.44	0/58	0.61	0/74
3	F	0.42	0/58	0.57	0/74
4	H	0.99	0/481	0.99	0/743
4	K	0.97	0/481	0.98	0/743
5	M	0.94	0/723	0.92	0/1108
5	R	0.98	0/702	0.95	1/1076 (0.1%)
All	All	0.47	0/19554	0.65	4/26804 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	411	HIS	C-N-CA	7.34	140.06	121.70
5	R	8	DA	O4'-C1'-N9	6.71	112.69	108.00
2	B	530	LEU	CA-CB-CG	5.80	128.64	115.30
1	D	96	SER	C-N-CA	5.15	134.56	121.70

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	4048	0	4125	32	0
1	D	4066	0	4139	33	0
2	B	4278	0	4302	42	0
2	E	4315	0	4323	38	0
3	C	57	4	60	1	0
3	F	57	4	60	3	0
4	H	431	0	243	4	0
4	K	431	0	243	4	0
5	M	645	0	359	1	0
5	R	626	0	348	2	0
6	B	5	0	0	0	0
6	D	5	0	0	0	0
6	E	5	0	0	0	0
7	A	2	0	0	0	0
7	B	5	0	0	0	0
7	D	2	0	0	0	0
7	E	5	0	0	0	0
7	H	1	0	0	0	0
7	K	2	0	0	0	0
7	M	3	0	0	0	0
All	All	18989	8	18202	129	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 3.

The worst 5 of 129 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:D:446:MET:HE2	2:E:365:PHE:HE2	1.48	0.78
2:B:111:LEU:HD13	2:B:134:ILE:HD11	1.71	0.73
2:E:111:LEU:HD13	2:E:134:ILE:HD11	1.71	0.72
1:A:264:ASN:HB3	2:B:530:LEU:HD21	1.72	0.72
2:E:479:THR:HA	2:E:482:ILE:HD12	1.72	0.71

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	499/544 (92%)	490 (98%)	9 (2%)	0	100	100
1	D	501/544 (92%)	489 (98%)	11 (2%)	1 (0%)	47	78
2	B	525/572 (92%)	513 (98%)	11 (2%)	1 (0%)	47	78
2	E	532/572 (93%)	518 (97%)	14 (3%)	0	100	100
3	C	5/13 (38%)	4 (80%)	0	1 (20%)	0	0
3	F	5/13 (38%)	4 (80%)	0	1 (20%)	0	0
All	All	2067/2258 (92%)	2018 (98%)	45 (2%)	4 (0%)	47	78

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	97	VAL
3	C	296	GLY
3	F	296	GLY
2	B	412	ILE

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	455/492 (92%)	448 (98%)	7 (2%)	65	87
1	D	457/492 (93%)	448 (98%)	9 (2%)	55	82
2	B	480/513 (94%)	475 (99%)	5 (1%)	76	92

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	E	484/513 (94%)	478 (99%)	6 (1%)	71	91
3	C	6/12 (50%)	6 (100%)	0	100	100
3	F	6/12 (50%)	6 (100%)	0	100	100
All	All	1888/2034 (93%)	1861 (99%)	27 (1%)	67	89

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	174	ASN
1	D	243	LEU
2	E	338	LYS
1	D	226	ASP
1	D	337	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	0	GLN
1	D	174	ASN
2	E	0	GLN
1	D	360	HIS
1	A	360	HIS

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](#)

3 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	$\# Z > 2$
6	SO4	B	601	-	4,4,4	0.14	0	6,6,6	0.05	0
6	SO4	E	601	-	4,4,4	0.14	0	6,6,6	0.05	0
6	SO4	D	601	-	4,4,4	0.14	0	6,6,6	0.05	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	501/544 (92%)	0.78	57 (11%) 5 3	55, 103, 179, 211	0
1	D	503/544 (92%)	0.49	44 (8%) 10 7	35, 86, 160, 192	0
2	B	531/572 (92%)	0.55	49 (9%) 9 6	48, 90, 155, 210	0
2	E	536/572 (93%)	0.40	31 (5%) 23 19	35, 80, 145, 198	0
3	C	7/13 (53%)	0.11	0 100 100	69, 72, 97, 108	0
3	F	7/13 (53%)	1.03	2 (28%) 0 0	69, 75, 113, 113	0
4	H	21/21 (100%)	-0.15	0 100 100	88, 105, 174, 194	0
4	K	21/21 (100%)	0.06	0 100 100	82, 105, 160, 165	0
5	M	32/34 (94%)	0.95	7 (21%) 0 0	69, 126, 210, 218	0
5	R	31/34 (91%)	0.84	4 (12%) 3 2	67, 135, 180, 201	0
All	All	2190/2368 (92%)	0.55	194 (8%) 9 7	35, 91, 166, 218	0

The worst 5 of 194 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	159	PHE	7.3
2	B	468	GLU	7.3
1	A	300	THR	7.0
1	A	289	TYR	6.9
5	M	15	DC	6.2

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](#)

LIGAND-RSR INFOmissingINFO

6.5 Other polymers [i](#)

There are no such residues in this entry.