



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 08:09 am GMT

PDB ID : 5AKD
Title : MutS in complex with the N-terminal domain of MutL - crystal form 3
Authors : Groothuizen, F.S.; Winkler, I.; Cristovao, M.; Fish, A.; Winterwerp, H.H.K.; Reumer, A.; Marx, A.D.; Hermans, N.; Nicholls, R.A.; Murshudov, G.N.; Lebbink, J.H.G.; Friedhoff, P.; Sixma, T.K.
Deposited on : 2015-03-03
Resolution : 7.60 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

<http://wwpdb.org/validation/2016/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	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

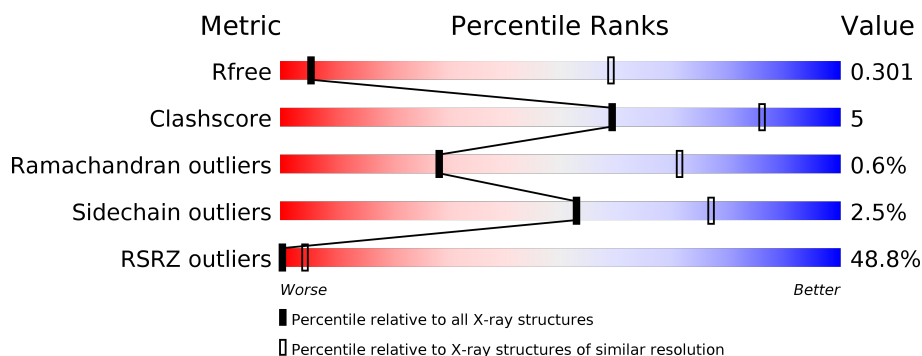
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 7.60 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}	100719	1100 (10.00-3.70)
Clashscore	112137	1035 (10.00-3.80)
Ramachandran outliers	110173	1003 (10.00-3.76)
Sidechain outliers	110143	1099 (11.50-3.70)
RSRZ outliers	101464	1003 (10.00-3.72)

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. 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	800	<div> <div>39%</div> <div>72%10%17%</div> </div>
1	B	800	<div> <div>36%</div> <div>73%9%17%</div> </div>
1	E	800	<div> <div>43%</div> <div>72%10%17%</div> </div>
1	F	800	<div> <div>35%</div> <div>73%10%17%</div> </div>
1	I	800	<div> <div>41%</div> <div>75%7%17%</div> </div>
1	J	800	<div> <div>45%</div> <div>74%9%17%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	C	369	
2	D	369	
2	G	369	
2	H	369	
2	K	369	
2	L	369	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	ANP	A	1801	-	-	-	X
3	ANP	E	1801	-	-	-	X
3	ANP	I	1801	-	-	-	X
3	ANP	J	1801	-	-	-	X

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 45054 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 DNA MISMATCH REPAIR PROTEIN MUTS.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	663	Total	C	N	O	S	0	0	0
			5226	3284	938	984	20			
1	B	663	Total	C	N	O	S	0	0	0
			5226	3284	938	984	20			
1	E	663	Total	C	N	O	S	0	0	0
			5226	3284	938	984	20			
1	F	663	Total	C	N	O	S	0	0	0
			5226	3284	938	984	20			
1	I	663	Total	C	N	O	S	0	0	0
			5226	3284	938	984	20			
1	J	663	Total	C	N	O	S	0	0	0
			5226	3284	938	984	20			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	93	ALA	CYS	ENGINEERED MUTATION	UNP P23909
A	235	SER	CYS	ENGINEERED MUTATION	UNP P23909
A	239	ALA	CYS	ENGINEERED MUTATION	UNP P23909
A	246	CYS	ASP	ENGINEERED MUTATION	UNP P23909
A	297	SER	CYS	ENGINEERED MUTATION	UNP P23909
A	569	SER	CYS	ENGINEERED MUTATION	UNP P23909
A	711	VAL	CYS	ENGINEERED MUTATION	UNP P23909
B	93	ALA	CYS	ENGINEERED MUTATION	UNP P23909
B	235	SER	CYS	ENGINEERED MUTATION	UNP P23909
B	239	ALA	CYS	ENGINEERED MUTATION	UNP P23909
B	246	CYS	ASP	ENGINEERED MUTATION	UNP P23909
B	297	SER	CYS	ENGINEERED MUTATION	UNP P23909
B	569	SER	CYS	ENGINEERED MUTATION	UNP P23909
B	711	VAL	CYS	ENGINEERED MUTATION	UNP P23909
E	93	ALA	CYS	ENGINEERED MUTATION	UNP P23909
E	235	SER	CYS	ENGINEERED MUTATION	UNP P23909
E	239	ALA	CYS	ENGINEERED MUTATION	UNP P23909

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	246	CYS	ASP	ENGINEERED MUTATION	UNP P23909
E	297	SER	CYS	ENGINEERED MUTATION	UNP P23909
E	569	SER	CYS	ENGINEERED MUTATION	UNP P23909
E	711	VAL	CYS	ENGINEERED MUTATION	UNP P23909
F	93	ALA	CYS	ENGINEERED MUTATION	UNP P23909
F	235	SER	CYS	ENGINEERED MUTATION	UNP P23909
F	239	ALA	CYS	ENGINEERED MUTATION	UNP P23909
F	246	CYS	ASP	ENGINEERED MUTATION	UNP P23909
F	297	SER	CYS	ENGINEERED MUTATION	UNP P23909
F	569	SER	CYS	ENGINEERED MUTATION	UNP P23909
F	711	VAL	CYS	ENGINEERED MUTATION	UNP P23909
I	93	ALA	CYS	ENGINEERED MUTATION	UNP P23909
I	235	SER	CYS	ENGINEERED MUTATION	UNP P23909
I	239	ALA	CYS	ENGINEERED MUTATION	UNP P23909
I	246	CYS	ASP	ENGINEERED MUTATION	UNP P23909
I	297	SER	CYS	ENGINEERED MUTATION	UNP P23909
I	569	SER	CYS	ENGINEERED MUTATION	UNP P23909
I	711	VAL	CYS	ENGINEERED MUTATION	UNP P23909
J	93	ALA	CYS	ENGINEERED MUTATION	UNP P23909
J	235	SER	CYS	ENGINEERED MUTATION	UNP P23909
J	239	ALA	CYS	ENGINEERED MUTATION	UNP P23909
J	246	CYS	ASP	ENGINEERED MUTATION	UNP P23909
J	297	SER	CYS	ENGINEERED MUTATION	UNP P23909
J	569	SER	CYS	ENGINEERED MUTATION	UNP P23909
J	711	VAL	CYS	ENGINEERED MUTATION	UNP P23909

- Molecule 2 is a protein called DNA MISMATCH REPAIR PROTEIN MUTL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	285	Total	C	N	O	S	0	0	0
			2252	1423	409	418	2			
2	D	285	Total	C	N	O	S	0	0	0
			2252	1423	409	418	2			
2	G	285	Total	C	N	O	S	0	0	0
			2252	1423	409	418	2			
2	H	285	Total	C	N	O	S	0	0	0
			2252	1423	409	418	2			
2	K	285	Total	C	N	O	S	0	0	0
			2252	1423	409	418	2			
2	L	285	Total	C	N	O	S	0	0	0
			2252	1423	409	418	2			

There are 150 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-19	MET	-	EXPRESSION TAG	UNP P23367
C	-18	GLY	-	EXPRESSION TAG	UNP P23367
C	-17	SER	-	EXPRESSION TAG	UNP P23367
C	-16	SER	-	EXPRESSION TAG	UNP P23367
C	-15	HIS	-	EXPRESSION TAG	UNP P23367
C	-14	HIS	-	EXPRESSION TAG	UNP P23367
C	-13	HIS	-	EXPRESSION TAG	UNP P23367
C	-12	HIS	-	EXPRESSION TAG	UNP P23367
C	-11	HIS	-	EXPRESSION TAG	UNP P23367
C	-10	HIS	-	EXPRESSION TAG	UNP P23367
C	-9	SER	-	EXPRESSION TAG	UNP P23367
C	-8	SER	-	EXPRESSION TAG	UNP P23367
C	-7	GLY	-	EXPRESSION TAG	UNP P23367
C	-6	LEU	-	EXPRESSION TAG	UNP P23367
C	-5	VAL	-	EXPRESSION TAG	UNP P23367
C	-4	PRO	-	EXPRESSION TAG	UNP P23367
C	-3	ARG	-	EXPRESSION TAG	UNP P23367
C	-2	GLY	-	EXPRESSION TAG	UNP P23367
C	-1	SER	-	EXPRESSION TAG	UNP P23367
C	0	HIS	-	EXPRESSION TAG	UNP P23367
C	61	SER	CYS	ENGINEERED MUTATION	UNP P23367
C	131	CYS	ASN	ENGINEERED MUTATION	UNP P23367
C	216	LEU	CYS	ENGINEERED MUTATION	UNP P23367
C	256	PHE	CYS	ENGINEERED MUTATION	UNP P23367
C	276	TYR	CYS	ENGINEERED MUTATION	UNP P23367
D	-19	MET	-	EXPRESSION TAG	UNP P23367
D	-18	GLY	-	EXPRESSION TAG	UNP P23367
D	-17	SER	-	EXPRESSION TAG	UNP P23367
D	-16	SER	-	EXPRESSION TAG	UNP P23367
D	-15	HIS	-	EXPRESSION TAG	UNP P23367
D	-14	HIS	-	EXPRESSION TAG	UNP P23367
D	-13	HIS	-	EXPRESSION TAG	UNP P23367
D	-12	HIS	-	EXPRESSION TAG	UNP P23367
D	-11	HIS	-	EXPRESSION TAG	UNP P23367
D	-10	HIS	-	EXPRESSION TAG	UNP P23367
D	-9	SER	-	EXPRESSION TAG	UNP P23367
D	-8	SER	-	EXPRESSION TAG	UNP P23367
D	-7	GLY	-	EXPRESSION TAG	UNP P23367
D	-6	LEU	-	EXPRESSION TAG	UNP P23367
D	-5	VAL	-	EXPRESSION TAG	UNP P23367
D	-4	PRO	-	EXPRESSION TAG	UNP P23367
D	-3	ARG	-	EXPRESSION TAG	UNP P23367
D	-2	GLY	-	EXPRESSION TAG	UNP P23367

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-1	SER	-	EXPRESSION TAG	UNP P23367
D	0	HIS	-	EXPRESSION TAG	UNP P23367
D	61	SER	CYS	ENGINEERED MUTATION	UNP P23367
D	131	CYS	ASN	ENGINEERED MUTATION	UNP P23367
D	216	LEU	CYS	ENGINEERED MUTATION	UNP P23367
D	256	PHE	CYS	ENGINEERED MUTATION	UNP P23367
D	276	TYR	CYS	ENGINEERED MUTATION	UNP P23367
G	-19	MET	-	EXPRESSION TAG	UNP P23367
G	-18	GLY	-	EXPRESSION TAG	UNP P23367
G	-17	SER	-	EXPRESSION TAG	UNP P23367
G	-16	SER	-	EXPRESSION TAG	UNP P23367
G	-15	HIS	-	EXPRESSION TAG	UNP P23367
G	-14	HIS	-	EXPRESSION TAG	UNP P23367
G	-13	HIS	-	EXPRESSION TAG	UNP P23367
G	-12	HIS	-	EXPRESSION TAG	UNP P23367
G	-11	HIS	-	EXPRESSION TAG	UNP P23367
G	-10	HIS	-	EXPRESSION TAG	UNP P23367
G	-9	SER	-	EXPRESSION TAG	UNP P23367
G	-8	SER	-	EXPRESSION TAG	UNP P23367
G	-7	GLY	-	EXPRESSION TAG	UNP P23367
G	-6	LEU	-	EXPRESSION TAG	UNP P23367
G	-5	VAL	-	EXPRESSION TAG	UNP P23367
G	-4	PRO	-	EXPRESSION TAG	UNP P23367
G	-3	ARG	-	EXPRESSION TAG	UNP P23367
G	-2	GLY	-	EXPRESSION TAG	UNP P23367
G	-1	SER	-	EXPRESSION TAG	UNP P23367
G	0	HIS	-	EXPRESSION TAG	UNP P23367
G	61	SER	CYS	ENGINEERED MUTATION	UNP P23367
G	131	CYS	ASN	ENGINEERED MUTATION	UNP P23367
G	216	LEU	CYS	ENGINEERED MUTATION	UNP P23367
G	256	PHE	CYS	ENGINEERED MUTATION	UNP P23367
G	276	TYR	CYS	ENGINEERED MUTATION	UNP P23367
H	-19	MET	-	EXPRESSION TAG	UNP P23367
H	-18	GLY	-	EXPRESSION TAG	UNP P23367
H	-17	SER	-	EXPRESSION TAG	UNP P23367
H	-16	SER	-	EXPRESSION TAG	UNP P23367
H	-15	HIS	-	EXPRESSION TAG	UNP P23367
H	-14	HIS	-	EXPRESSION TAG	UNP P23367
H	-13	HIS	-	EXPRESSION TAG	UNP P23367
H	-12	HIS	-	EXPRESSION TAG	UNP P23367
H	-11	HIS	-	EXPRESSION TAG	UNP P23367
H	-10	HIS	-	EXPRESSION TAG	UNP P23367

Continued on next page...

Continued from previous page...

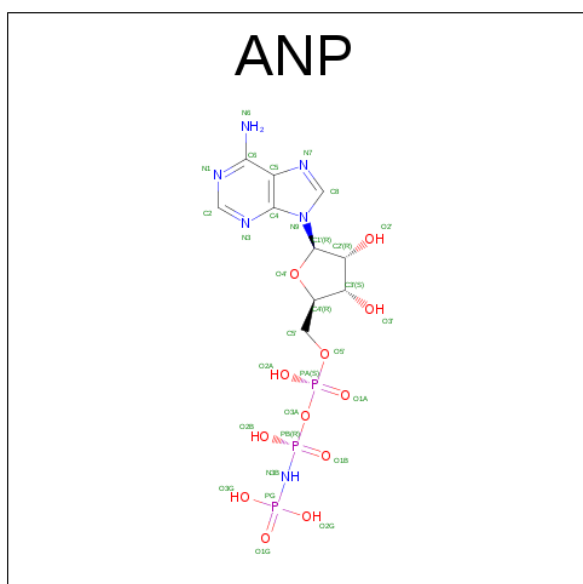
Chain	Residue	Modelled	Actual	Comment	Reference
H	-9	SER	-	EXPRESSION TAG	UNP P23367
H	-8	SER	-	EXPRESSION TAG	UNP P23367
H	-7	GLY	-	EXPRESSION TAG	UNP P23367
H	-6	LEU	-	EXPRESSION TAG	UNP P23367
H	-5	VAL	-	EXPRESSION TAG	UNP P23367
H	-4	PRO	-	EXPRESSION TAG	UNP P23367
H	-3	ARG	-	EXPRESSION TAG	UNP P23367
H	-2	GLY	-	EXPRESSION TAG	UNP P23367
H	-1	SER	-	EXPRESSION TAG	UNP P23367
H	0	HIS	-	EXPRESSION TAG	UNP P23367
H	61	SER	CYS	ENGINEERED MUTATION	UNP P23367
H	131	CYS	ASN	ENGINEERED MUTATION	UNP P23367
H	216	LEU	CYS	ENGINEERED MUTATION	UNP P23367
H	256	PHE	CYS	ENGINEERED MUTATION	UNP P23367
H	276	TYR	CYS	ENGINEERED MUTATION	UNP P23367
K	-19	MET	-	EXPRESSION TAG	UNP P23367
K	-18	GLY	-	EXPRESSION TAG	UNP P23367
K	-17	SER	-	EXPRESSION TAG	UNP P23367
K	-16	SER	-	EXPRESSION TAG	UNP P23367
K	-15	HIS	-	EXPRESSION TAG	UNP P23367
K	-14	HIS	-	EXPRESSION TAG	UNP P23367
K	-13	HIS	-	EXPRESSION TAG	UNP P23367
K	-12	HIS	-	EXPRESSION TAG	UNP P23367
K	-11	HIS	-	EXPRESSION TAG	UNP P23367
K	-10	HIS	-	EXPRESSION TAG	UNP P23367
K	-9	SER	-	EXPRESSION TAG	UNP P23367
K	-8	SER	-	EXPRESSION TAG	UNP P23367
K	-7	GLY	-	EXPRESSION TAG	UNP P23367
K	-6	LEU	-	EXPRESSION TAG	UNP P23367
K	-5	VAL	-	EXPRESSION TAG	UNP P23367
K	-4	PRO	-	EXPRESSION TAG	UNP P23367
K	-3	ARG	-	EXPRESSION TAG	UNP P23367
K	-2	GLY	-	EXPRESSION TAG	UNP P23367
K	-1	SER	-	EXPRESSION TAG	UNP P23367
K	0	HIS	-	EXPRESSION TAG	UNP P23367
K	61	SER	CYS	ENGINEERED MUTATION	UNP P23367
K	131	CYS	ASN	ENGINEERED MUTATION	UNP P23367
K	216	LEU	CYS	ENGINEERED MUTATION	UNP P23367
K	256	PHE	CYS	ENGINEERED MUTATION	UNP P23367
K	276	TYR	CYS	ENGINEERED MUTATION	UNP P23367
L	-19	MET	-	EXPRESSION TAG	UNP P23367
L	-18	GLY	-	EXPRESSION TAG	UNP P23367

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	-17	SER	-	EXPRESSION TAG	UNP P23367
L	-16	SER	-	EXPRESSION TAG	UNP P23367
L	-15	HIS	-	EXPRESSION TAG	UNP P23367
L	-14	HIS	-	EXPRESSION TAG	UNP P23367
L	-13	HIS	-	EXPRESSION TAG	UNP P23367
L	-12	HIS	-	EXPRESSION TAG	UNP P23367
L	-11	HIS	-	EXPRESSION TAG	UNP P23367
L	-10	HIS	-	EXPRESSION TAG	UNP P23367
L	-9	SER	-	EXPRESSION TAG	UNP P23367
L	-8	SER	-	EXPRESSION TAG	UNP P23367
L	-7	GLY	-	EXPRESSION TAG	UNP P23367
L	-6	LEU	-	EXPRESSION TAG	UNP P23367
L	-5	VAL	-	EXPRESSION TAG	UNP P23367
L	-4	PRO	-	EXPRESSION TAG	UNP P23367
L	-3	ARG	-	EXPRESSION TAG	UNP P23367
L	-2	GLY	-	EXPRESSION TAG	UNP P23367
L	-1	SER	-	EXPRESSION TAG	UNP P23367
L	0	HIS	-	EXPRESSION TAG	UNP P23367
L	61	SER	CYS	ENGINEERED MUTATION	UNP P23367
L	131	CYS	ASN	ENGINEERED MUTATION	UNP P23367
L	216	LEU	CYS	ENGINEERED MUTATION	UNP P23367
L	256	PHE	CYS	ENGINEERED MUTATION	UNP P23367
L	276	TYR	CYS	ENGINEERED MUTATION	UNP P23367

- Molecule 3 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: $C_{10}H_{17}N_6O_{12}P_3$).

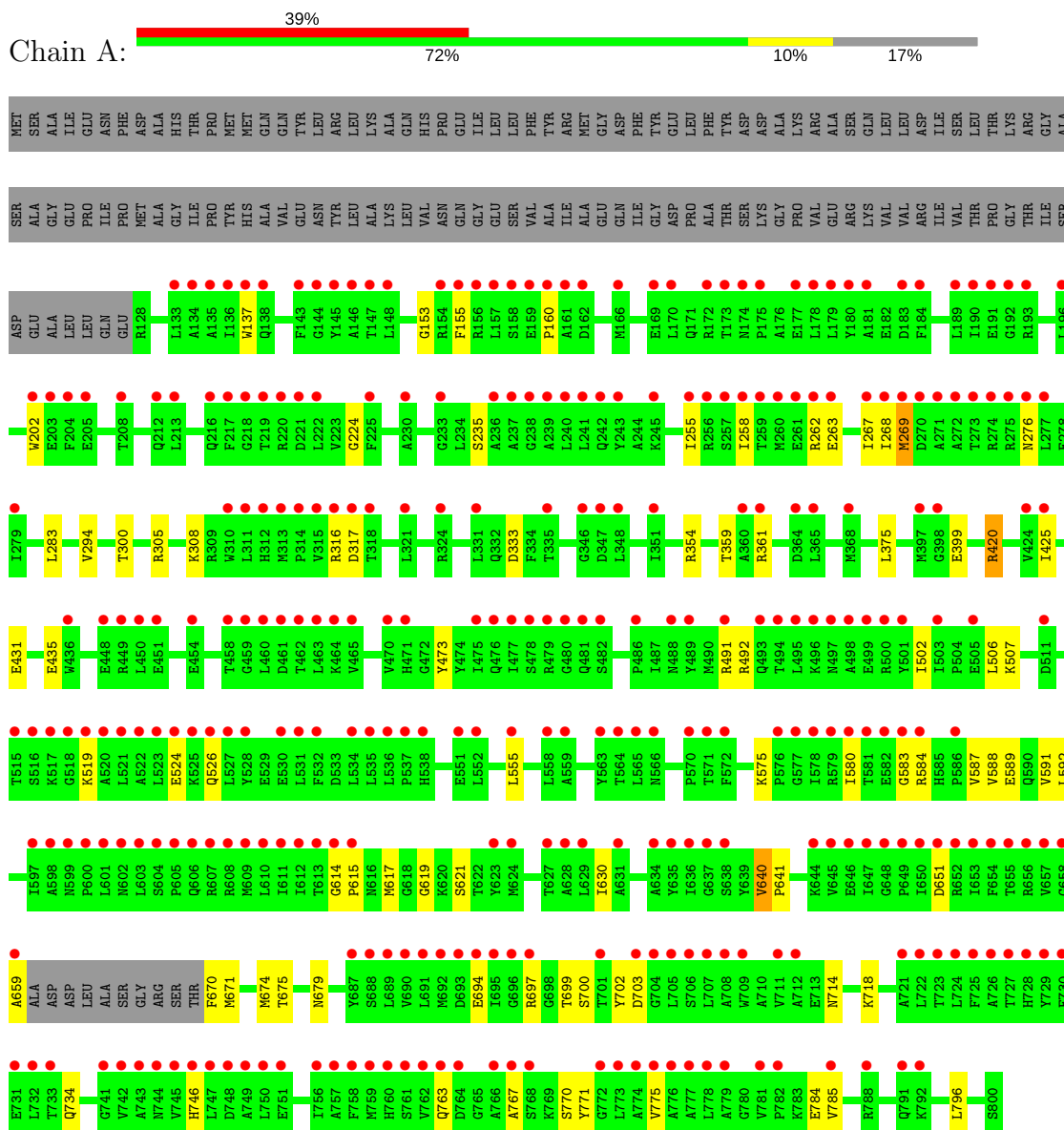


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
3	B	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
3	E	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
3	F	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
3	I	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
3	J	1	Total	C	N	O	P	0	0
			31	10	6	12	3		

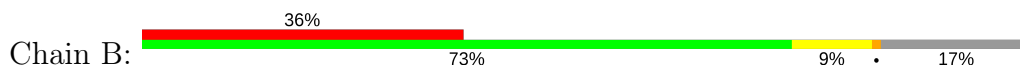
3 Residue-property plots

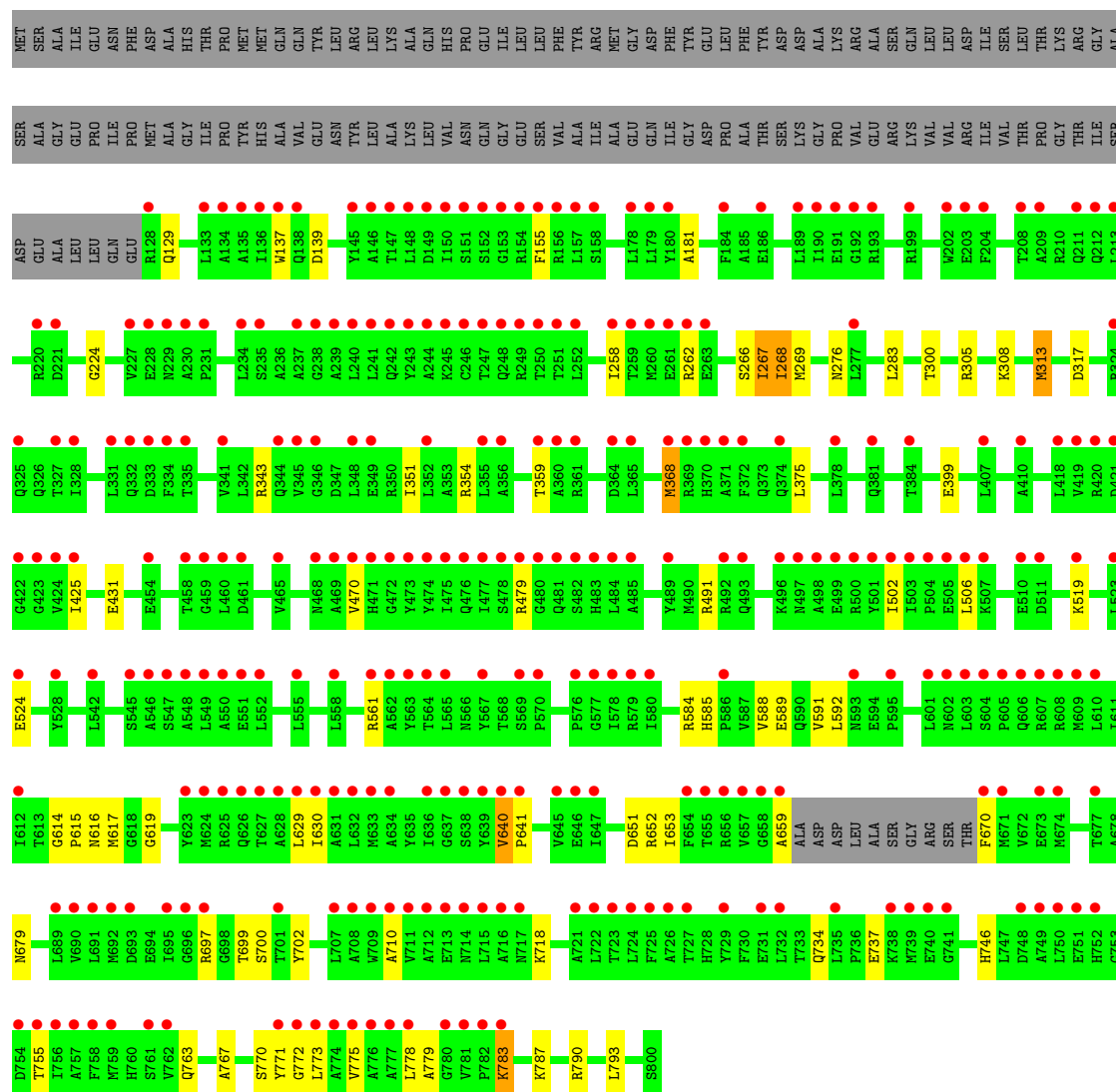
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: DNA MISMATCH REPAIR PROTEIN MUTS

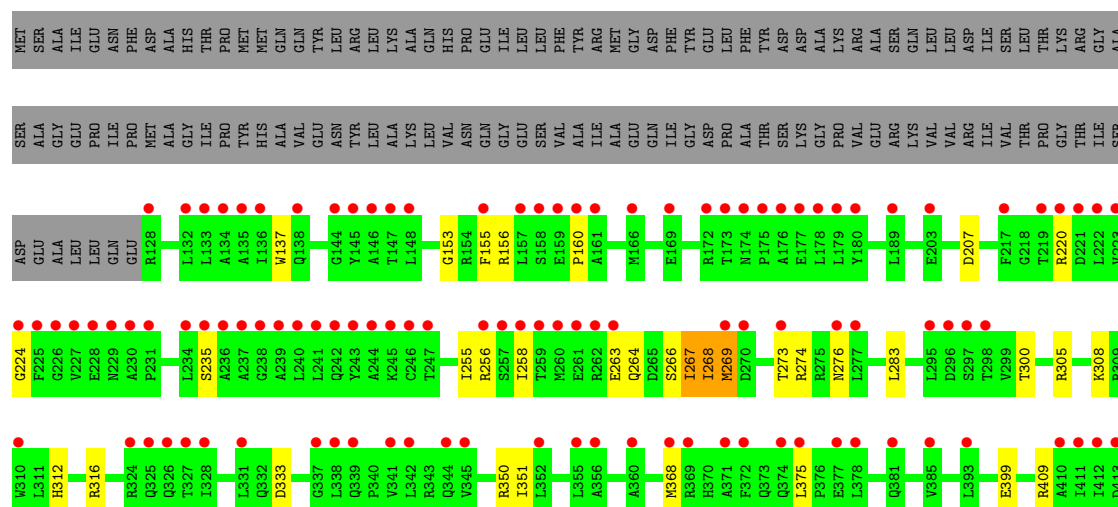


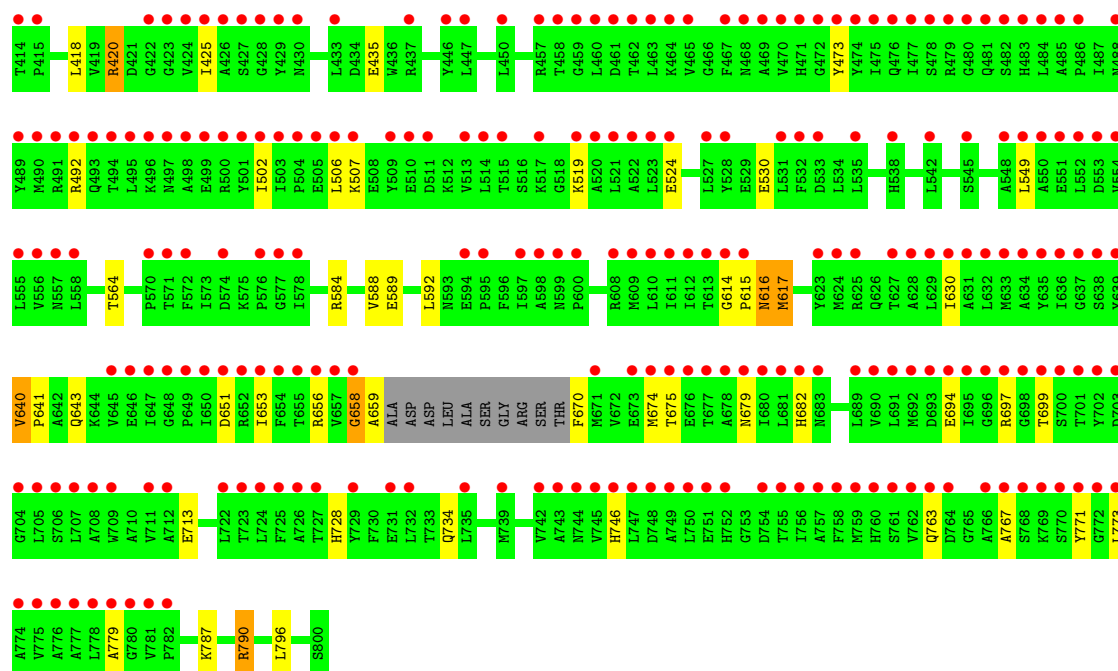
• Molecule 1: DNA MISMATCH REPAIR PROTEIN MUTS





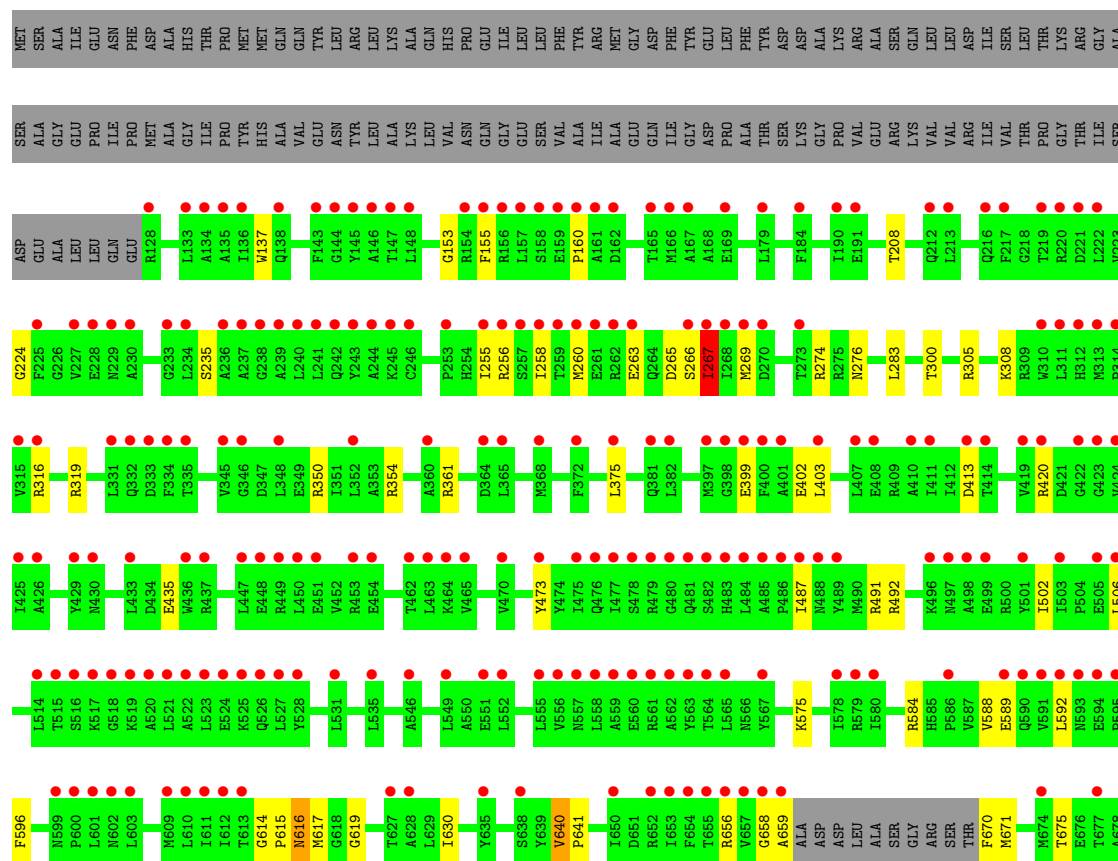
• Molecule 1: DNA MISMATCH REPAIR PROTEIN MUTS

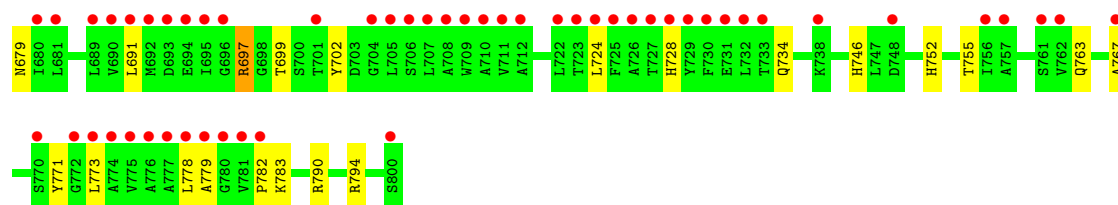




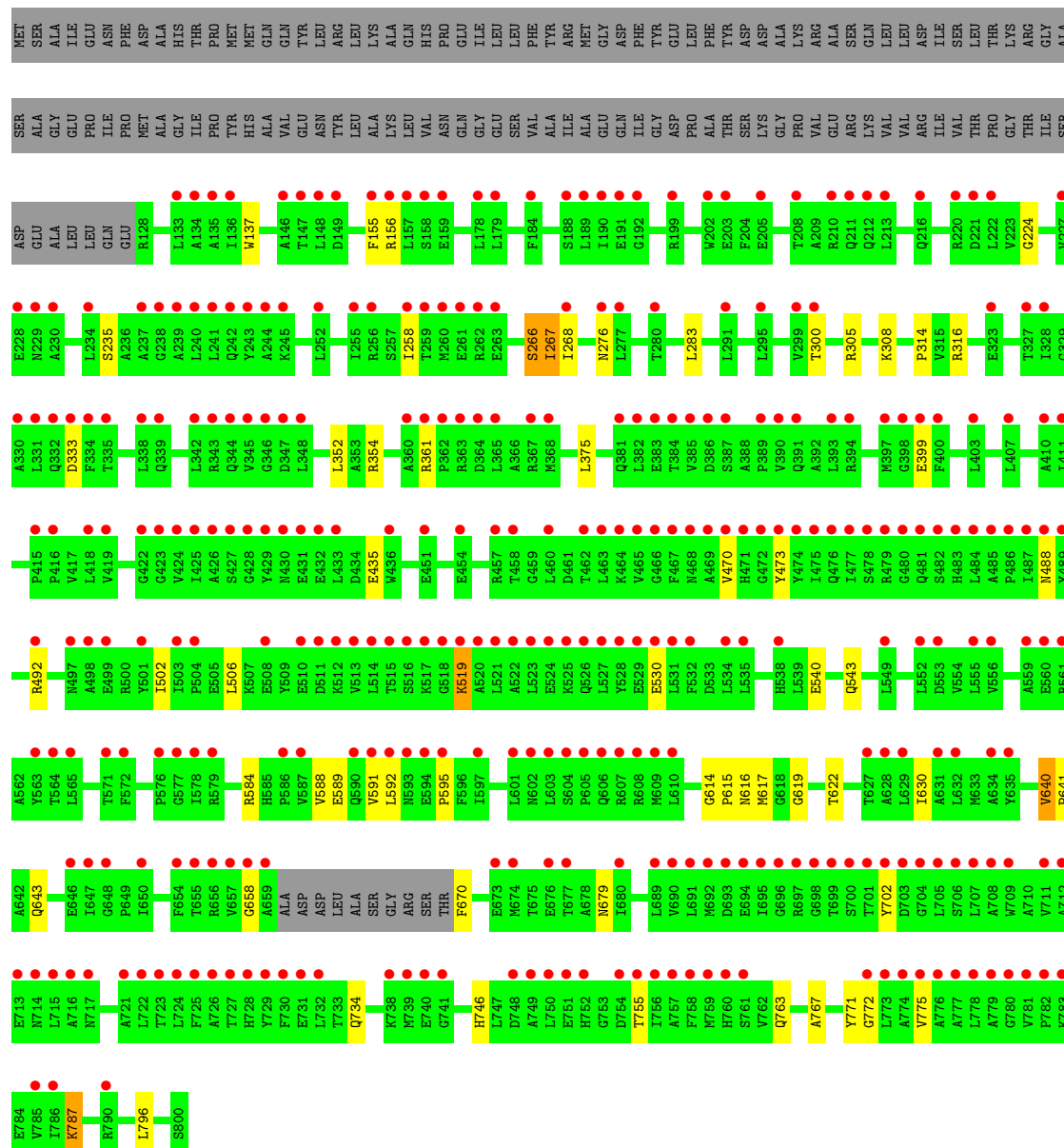
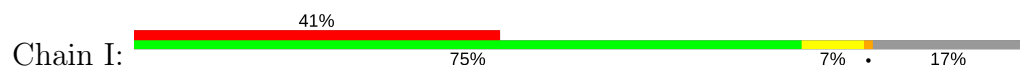
● Molecule 1: DNA MISMATCH REPAIR PROTEIN MUTS

Chain F:

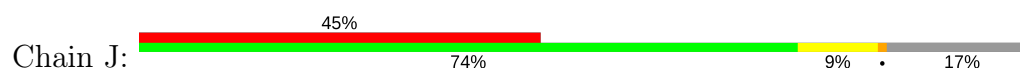


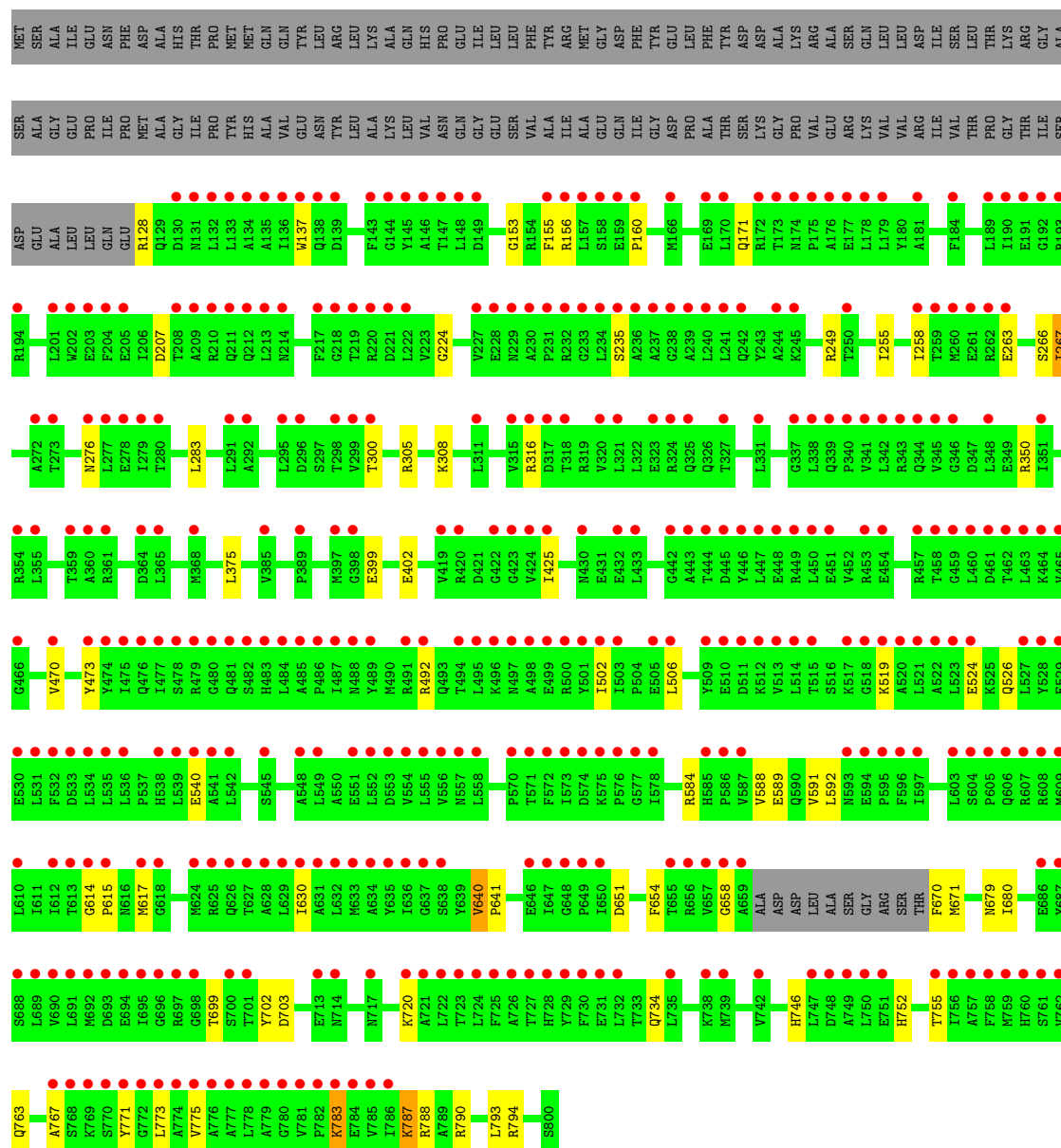


• Molecule 1: DNA MISMATCH REPAIR PROTEIN MUTS

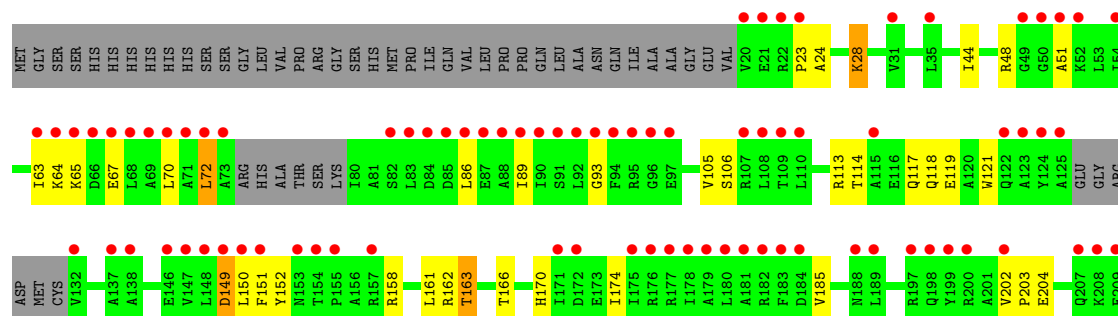


• Molecule 1: DNA MISMATCH REPAIR PROTEIN MUTS

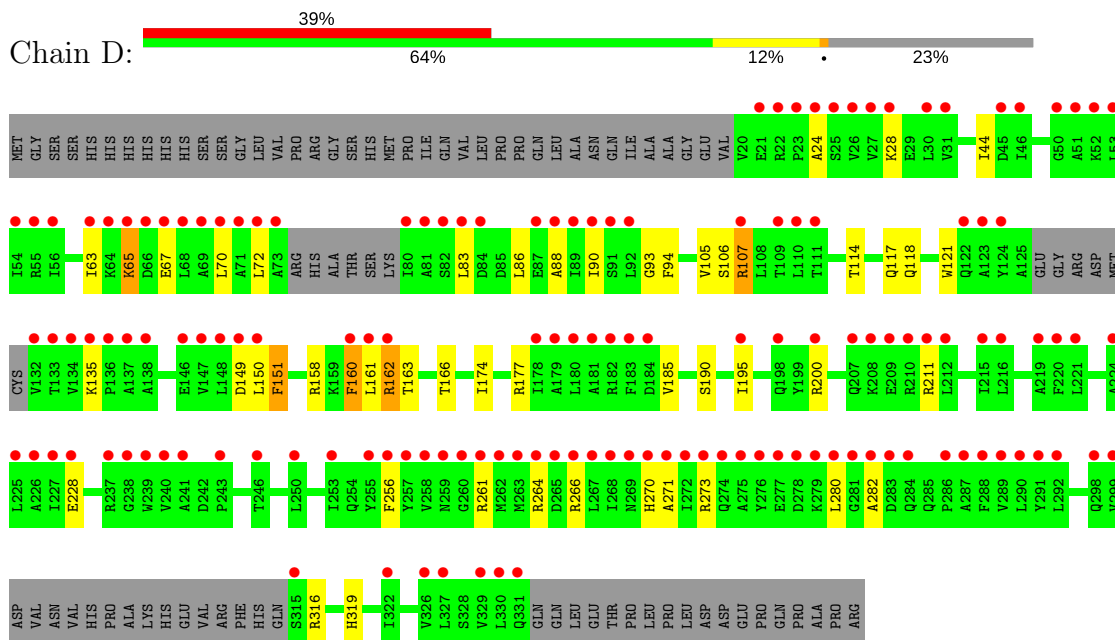




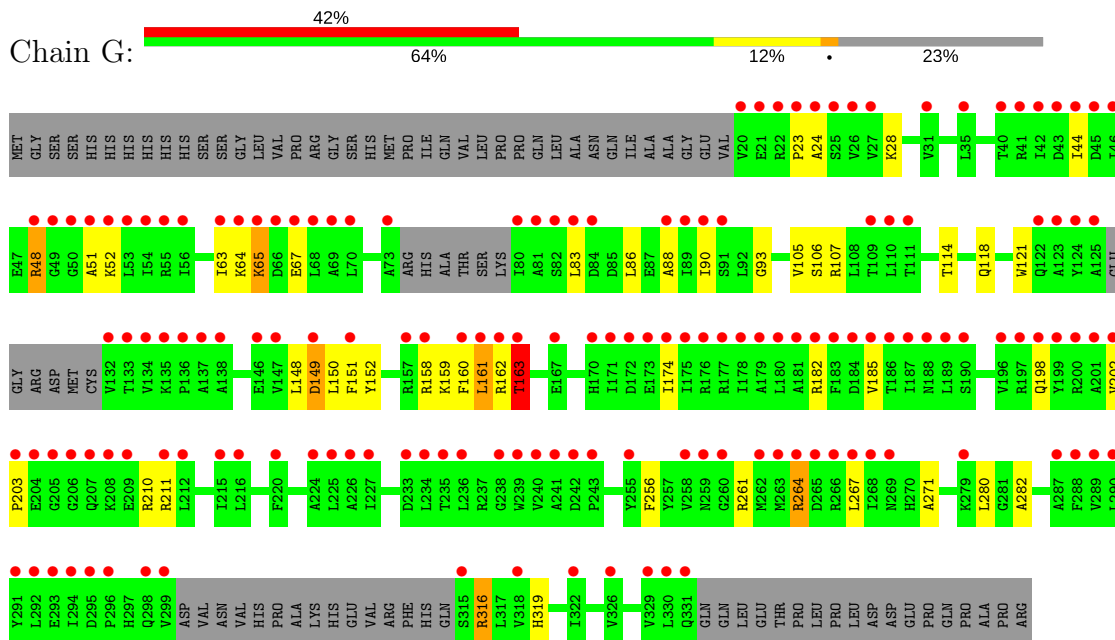
● Molecule 2: DNA MISMATCH REPAIR PROTEIN MUTL



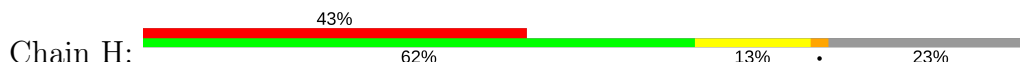
- Molecule 2: DNA MISMATCH REPAIR PROTEIN MUTL

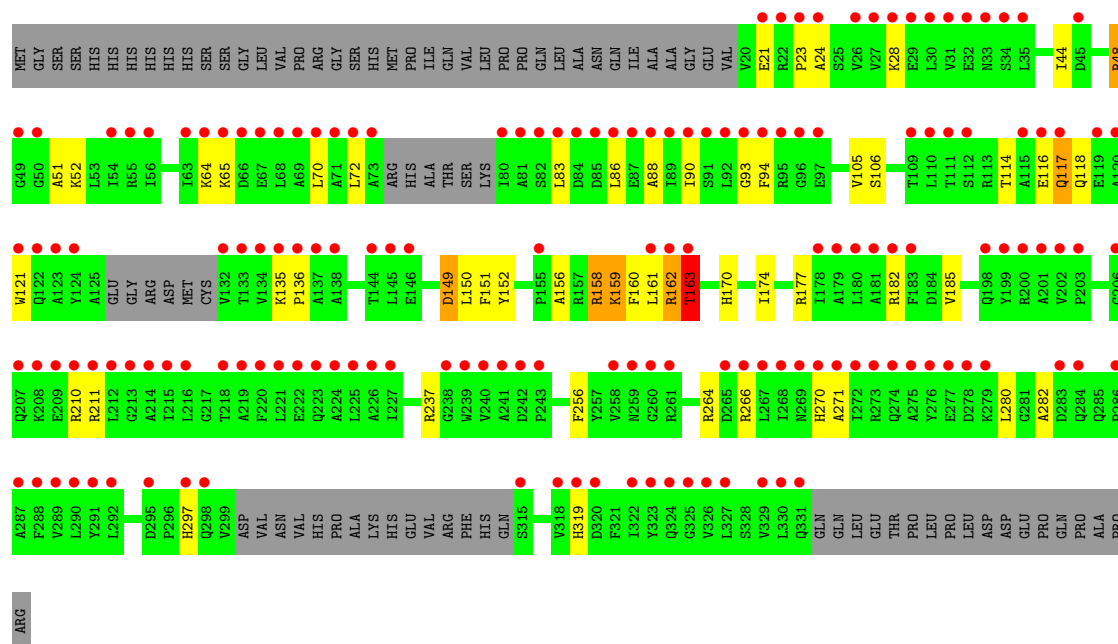


- Molecule 2: DNA MISMATCH REPAIR PROTEIN MUTL

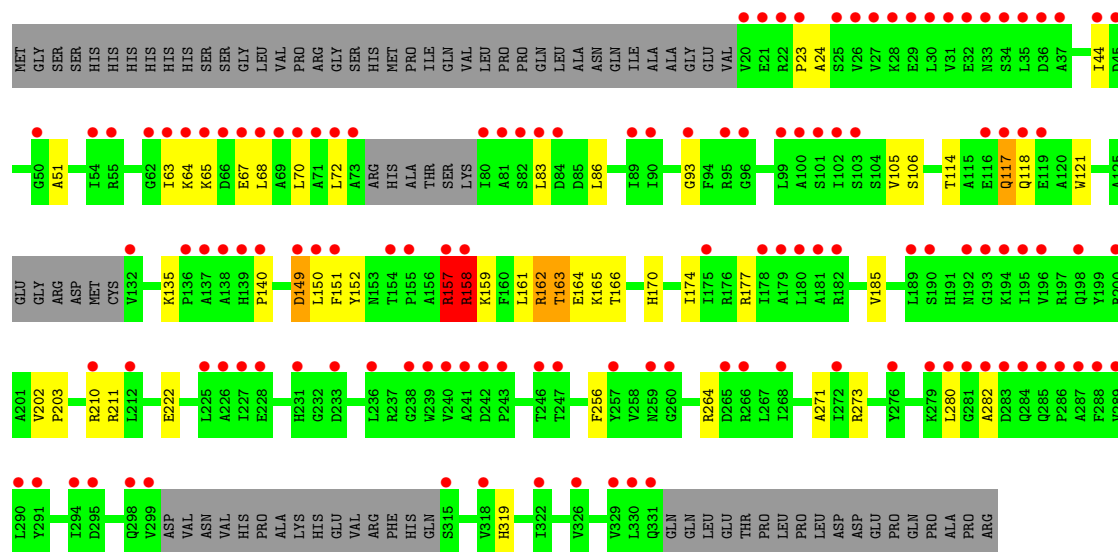


- Molecule 2: DNA MISMATCH REPAIR PROTEIN MUTL

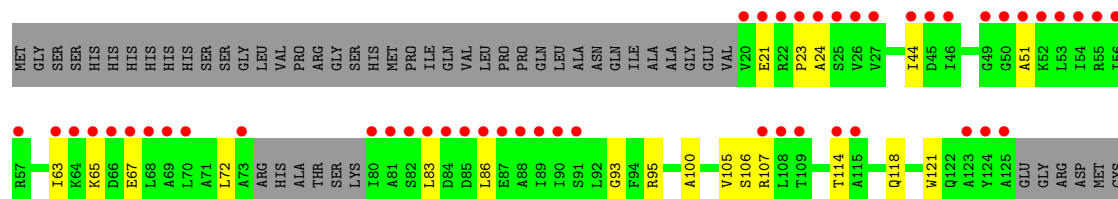
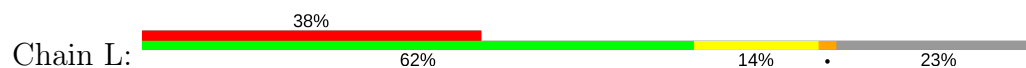


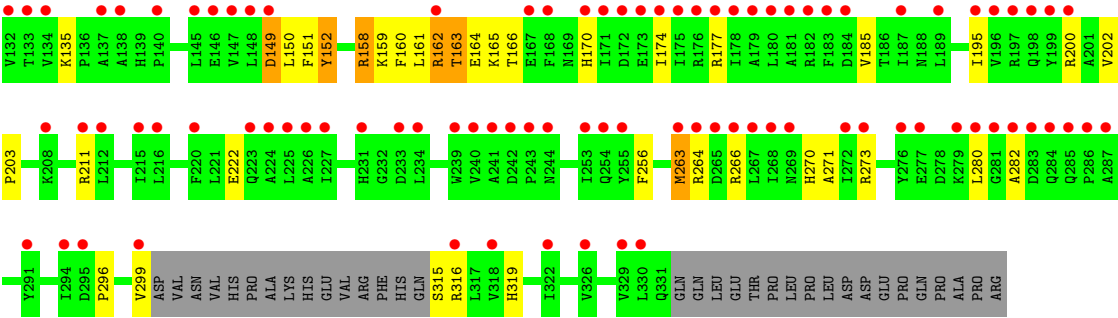


• Molecule 2: DNA MISMATCH REPAIR PROTEIN MUTL



• Molecule 2: DNA MISMATCH REPAIR PROTEIN MUTL





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	193.02Å 109.76Å 275.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	275.84 – 7.60 49.30 – 7.60	Depositor EDS
% Data completeness (in resolution range)	80.1 (275.84-7.60) 80.5 (49.30-7.60)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.14 (at 7.37Å)	Xtriage
Refinement program	REFMAC 5.8.0107	Depositor
R, R_{free}	0.264 , 0.306 0.260 , 0.301	Depositor DCC
R_{free} test set	585 reflections (5.24%)	DCC
Wilson B-factor (Å ²)	441.7	Xtriage
Anisotropy	0.279	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 500.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.37$, $\langle L^2 \rangle = 0.19$	Xtriage
Estimated twinning fraction	0.090 for h,-k,-l	Xtriage
F_o, F_c correlation	0.79	EDS
Total number of atoms	45054	wwPDB-VP
Average B, all atoms (Å ²)	231.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.81% of the height of the origin peak. No significant pseudotranslation is detected.*

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

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

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	0.66	1/5311 (0.0%)	0.80	9/7186 (0.1%)
1	B	0.66	0/5311	0.80	10/7186 (0.1%)
1	E	0.65	3/5311 (0.1%)	0.84	16/7186 (0.2%)
1	F	0.64	0/5311	0.77	13/7186 (0.2%)
1	I	0.62	0/5311	0.74	2/7186 (0.0%)
1	J	0.70	3/5311 (0.1%)	0.86	11/7186 (0.2%)
2	C	0.72	1/2288 (0.0%)	0.79	2/3096 (0.1%)
2	D	0.82	2/2288 (0.1%)	0.85	6/3096 (0.2%)
2	G	0.71	0/2288	0.81	4/3096 (0.1%)
2	H	0.71	1/2288 (0.0%)	0.84	8/3096 (0.3%)
2	K	0.90	3/2288 (0.1%)	1.03	8/3096 (0.3%)
2	L	0.71	0/2288	0.82	5/3096 (0.2%)
All	All	0.69	14/45594 (0.0%)	0.82	94/61692 (0.2%)

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	B	0	2
2	K	0	1
All	All	0	3

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	K	157	ARG	CD-NE	17.42	1.76	1.46
2	K	157	ARG	NE-CZ	17.13	1.55	1.33
1	J	787	LYS	CE-NZ	15.77	1.88	1.49
1	J	249	ARG	NE-CZ	13.73	1.50	1.33

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	228	GLU	CD-OE1	11.54	1.38	1.25

The worst 5 of 94 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	K	157	ARG	NE-CZ-NH1	31.56	136.08	120.30
1	J	249	ARG	NE-CZ-NH1	29.02	134.81	120.30
1	J	249	ARG	NE-CZ-NH2	-20.07	110.27	120.30
1	E	156	ARG	NE-CZ-NH2	16.87	128.74	120.30
1	E	420	ARG	NE-CZ-NH1	13.33	126.97	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	129	GLN	Peptide
1	B	268	ILE	Peptide
2	K	157	ARG	Sidechain

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	5226	0	5283	64	0
1	B	5226	0	5283	74	1
1	E	5226	0	5283	86	0
1	F	5226	0	5283	61	1
1	I	5226	0	5283	46	1
1	J	5226	0	5283	42	0
2	C	2252	0	2272	29	0
2	D	2252	0	2272	37	0
2	G	2252	0	2272	38	0
2	H	2252	0	2272	41	0
2	K	2252	0	2272	29	1
2	L	2252	0	2272	39	0
3	A	31	0	13	5	0
3	B	31	0	13	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	31	0	13	2	0
3	F	31	0	13	3	0
3	I	31	0	13	6	0
3	J	31	0	13	2	0
All	All	45054	0	45408	488	2

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 5.

The worst 5 of 488 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:C:28:LYS:NZ	2:C:28:LYS:CE	1.71	1.50
2:K:157:ARG:NE	2:K:157:ARG:CD	1.76	1.46
1:J:787:LYS:CE	1:J:787:LYS:NZ	1.88	1.35
1:E:269:MET:SD	1:E:653:ILE:HB	1.69	1.31
2:D:105:VAL:O	2:D:150:LEU:CD1	1.90	1.20

All (2) 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:B:737:GLU:OE2	1:F:491:ARG:NH1[2_544]	1.87	0.33
1:I:488:ASN:ND2	2:K:117:GLN:OE1[1_545]	1.91	0.29

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	659/800 (82%)	636 (96%)	23 (4%)	0	100	100
1	B	659/800 (82%)	638 (97%)	21 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	659/800 (82%)	632 (96%)	24 (4%)	3 (0%)	32	74
1	F	659/800 (82%)	634 (96%)	23 (4%)	2 (0%)	44	81
1	I	659/800 (82%)	635 (96%)	22 (3%)	2 (0%)	44	81
1	J	659/800 (82%)	639 (97%)	18 (3%)	2 (0%)	44	81
2	C	277/369 (75%)	258 (93%)	15 (5%)	4 (1%)	13	54
2	D	277/369 (75%)	260 (94%)	13 (5%)	4 (1%)	13	54
2	G	277/369 (75%)	261 (94%)	13 (5%)	3 (1%)	17	60
2	H	277/369 (75%)	260 (94%)	13 (5%)	4 (1%)	13	54
2	K	277/369 (75%)	262 (95%)	11 (4%)	4 (1%)	13	54
2	L	277/369 (75%)	261 (94%)	12 (4%)	4 (1%)	13	54
All	All	5616/7014 (80%)	5376 (96%)	208 (4%)	32 (1%)	28	71

5 of 32 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	163	THR
2	D	151	PHE
1	F	267	ILE
2	G	163	THR
2	H	163	THR

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	550/662 (83%)	539 (98%)	11 (2%)	60	82
1	B	550/662 (83%)	541 (98%)	9 (2%)	68	85
1	E	550/662 (83%)	539 (98%)	11 (2%)	60	82
1	F	550/662 (83%)	541 (98%)	9 (2%)	68	85
1	I	550/662 (83%)	538 (98%)	12 (2%)	57	79
1	J	550/662 (83%)	541 (98%)	9 (2%)	68	85

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	C	235/308 (76%)	228 (97%)	7 (3%)	46	72
2	D	235/308 (76%)	225 (96%)	10 (4%)	33	64
2	G	235/308 (76%)	227 (97%)	8 (3%)	42	69
2	H	235/308 (76%)	223 (95%)	12 (5%)	28	60
2	K	235/308 (76%)	227 (97%)	8 (3%)	42	69
2	L	235/308 (76%)	224 (95%)	11 (5%)	30	62
All	All	4710/5820 (81%)	4593 (98%)	117 (2%)	53	77

5 of 117 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	616	ASN
2	H	21	GLU
2	L	72	LEU
1	F	697	ARG
2	G	48	ARG

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	C	319	HIS
1	F	212	GLN
1	I	585	HIS
1	E	214	ASN
1	E	585	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

6 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$
3	ANP	A	1801	-	29,33,33	2.03	8 (27%)	28,52,52	2.31	7 (25%)
3	ANP	B	1801	-	29,33,33	1.94	7 (24%)	28,52,52	2.24	8 (28%)
3	ANP	E	1801	-	29,33,33	2.02	10 (34%)	28,52,52	2.22	8 (28%)
3	ANP	F	1801	-	29,33,33	1.84	9 (31%)	28,52,52	2.23	8 (28%)
3	ANP	I	1801	-	29,33,33	1.94	9 (31%)	28,52,52	2.05	7 (25%)
3	ANP	J	1801	-	29,33,33	1.94	8 (27%)	28,52,52	2.09	6 (21%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	ANP	A	1801	-	-	0/13/38/38	0/3/3/3
3	ANP	B	1801	-	-	0/13/38/38	0/3/3/3
3	ANP	E	1801	-	-	0/13/38/38	0/3/3/3
3	ANP	F	1801	-	-	0/13/38/38	0/3/3/3
3	ANP	I	1801	-	-	0/13/38/38	0/3/3/3
3	ANP	J	1801	-	-	0/13/38/38	0/3/3/3

The worst 5 of 51 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1801	ANP	PG-O3G	-3.51	1.47	1.56
3	I	1801	ANP	PG-O2G	-3.42	1.47	1.56
3	E	1801	ANP	PG-O2G	-3.22	1.47	1.56
3	I	1801	ANP	PG-O3G	-3.09	1.48	1.56
3	B	1801	ANP	PB-O2B	-3.02	1.48	1.56

The worst 5 of 44 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1801	ANP	N3-C2-N1	-7.60	122.24	128.86
3	B	1801	ANP	O1G-PG-N3B	-6.73	101.73	111.79
3	F	1801	ANP	O1G-PG-N3B	-6.55	101.99	111.79
3	J	1801	ANP	N3-C2-N1	-6.03	123.61	128.86
3	E	1801	ANP	N3-C2-N1	-5.80	123.81	128.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 22 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1801	ANP	5	0
3	B	1801	ANP	4	0
3	E	1801	ANP	2	0
3	F	1801	ANP	3	0
3	I	1801	ANP	6	0
3	J	1801	ANP	2	0

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	663/800 (82%)	2.23	315 (47%) 0 4	120, 187, 361, 469	0
1	B	663/800 (82%)	2.06	289 (43%) 0 5	125, 203, 272, 321	0
1	E	663/800 (82%)	2.67	343 (51%) 0 4	170, 262, 510, 665	0
1	F	663/800 (82%)	2.01	276 (41%) 0 5	127, 197, 351, 463	0
1	I	663/800 (82%)	2.38	325 (49%) 0 4	174, 224, 402, 487	0
1	J	663/800 (82%)	2.73	363 (54%) 0 4	172, 231, 378, 460	0
2	C	285/369 (77%)	2.04	136 (47%) 0 4	178, 213, 254, 293	0
2	D	285/369 (77%)	2.37	144 (50%) 0 4	170, 201, 273, 306	0
2	G	285/369 (77%)	2.38	155 (54%) 0 4	204, 224, 251, 265	0
2	H	285/369 (77%)	2.60	159 (55%) 0 4	154, 218, 301, 389	0
2	K	285/369 (77%)	2.12	130 (45%) 0 5	185, 220, 264, 283	0
2	L	285/369 (77%)	2.31	139 (48%) 0 4	210, 231, 253, 269	0
All	All	5688/7014 (81%)	2.33	2774 (48%) 0 4	120, 220, 361, 665	0

The worst 5 of 2774 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	J	749	ALA	21.6
1	E	757	ALA	19.1
1	J	147	THR	16.8
1	E	134	ALA	14.9
1	A	135	ALA	14.7

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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
3	ANP	I	1801	31/31	0.70	0.41	-0.52	177,180,186,187	0
3	ANP	E	1801	31/31	0.70	0.51	-0.72	204,216,227,232	0
3	ANP	A	1801	31/31	0.81	0.51	-0.74	156,164,170,171	0
3	ANP	J	1801	31/31	0.64	0.49	-0.83	193,199,212,213	0
3	ANP	B	1801	31/31	0.72	0.31	-0.87	165,174,179,181	0
3	ANP	F	1801	31/31	0.85	0.30	-1.06	154,162,169,174	0

6.5 Other polymers [i](#)

There are no such residues in this entry.