



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 08:09 am GMT

PDB ID : 5AKC  
Title : MutS in complex with the N-terminal domain of MutL - crystal form 2  
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 : 6.60 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](mailto: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

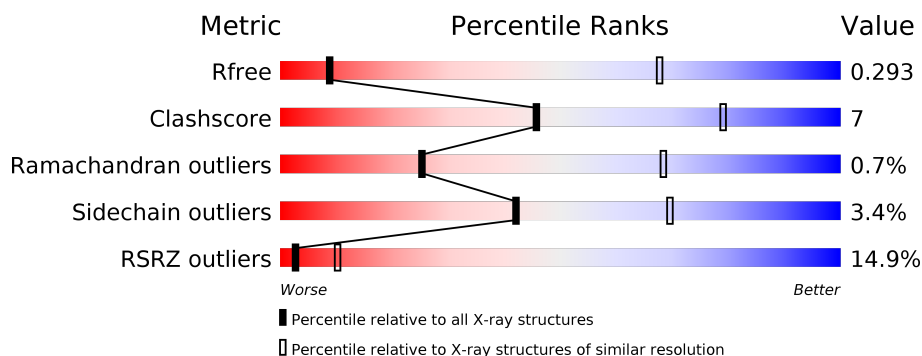
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 6.60 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	1099 (9.50-3.70)
Clashscore	112137	1031 (9.00-3.80)
Ramachandran outliers	110173	1000 (9.00-3.76)
Sidechain outliers	110143	1097 (9.50-3.70)
RSRZ outliers	101464	1000 (9.00-3.72)

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  $\geq 3$ , 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>10%</div> <div> <div></div> <div>74%</div> <div>8%</div> <div>•</div> <div>17%</div> </div> </div>
1	B	800	<div> <div>8%</div> <div> <div></div> <div>72%</div> <div>10%</div> <div>•</div> <div>17%</div> </div> </div>
1	E	800	<div> <div>11%</div> <div> <div></div> <div>72%</div> <div>10%</div> <div>•</div> <div>17%</div> </div> </div>
1	F	800	<div> <div>9%</div> <div> <div></div> <div>71%</div> <div>11%</div> <div>•</div> <div>17%</div> </div> </div>
1	I	800	<div> <div>14%</div> <div> <div></div> <div>71%</div> <div>12%</div> <div>•</div> <div>17%</div> </div> </div>
1	J	800	<div> <div>20%</div> <div> <div></div> <div>71%</div> <div>11%</div> <div>•</div> <div>17%</div> </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	E	1801	-	-	X	-

## 2 Entry composition [i](#)

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

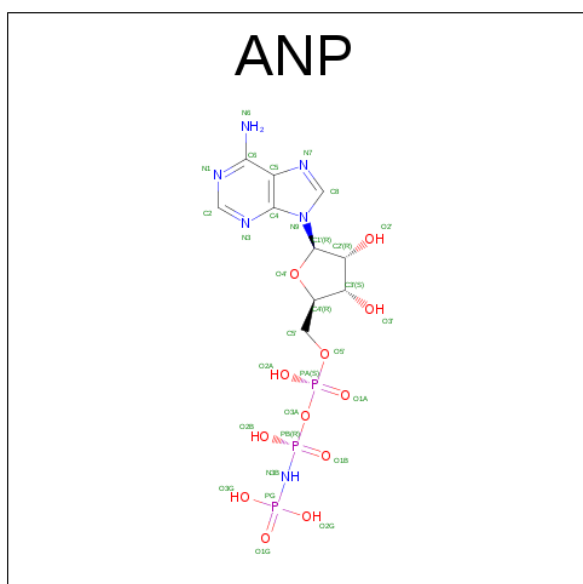
*Continued on next page...*



*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:  $C_{10}H_{17}N_6O_{12}P_3$ ).

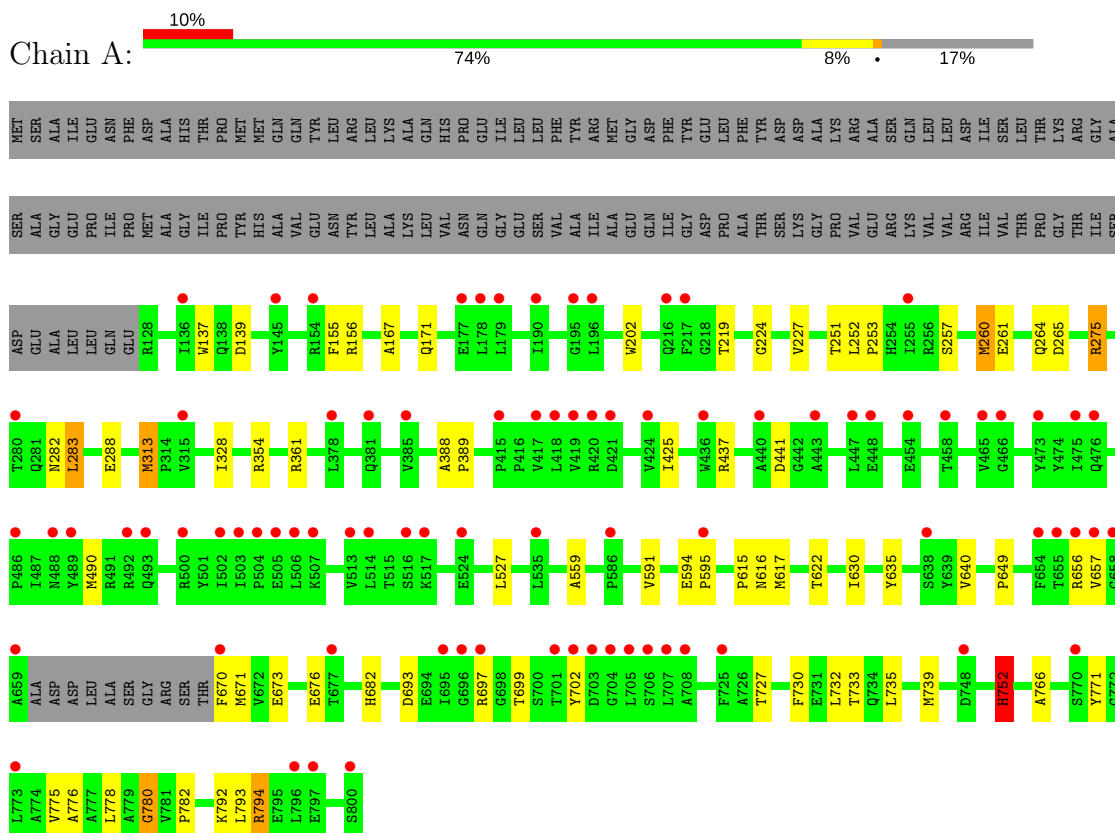


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
3	B	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
3	E	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
3	F	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
3	I	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
3	J	1	Total	C	N	O	P	0	0
			31	10	6	12	3		

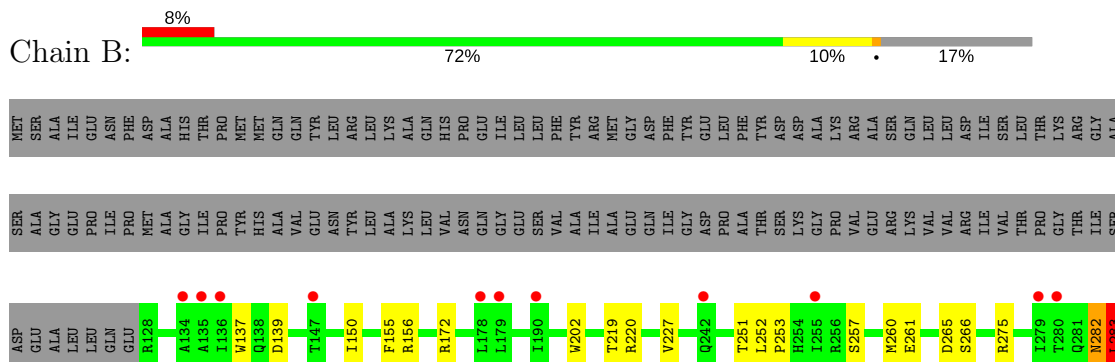
### 3 Residue-property plots [i](#)

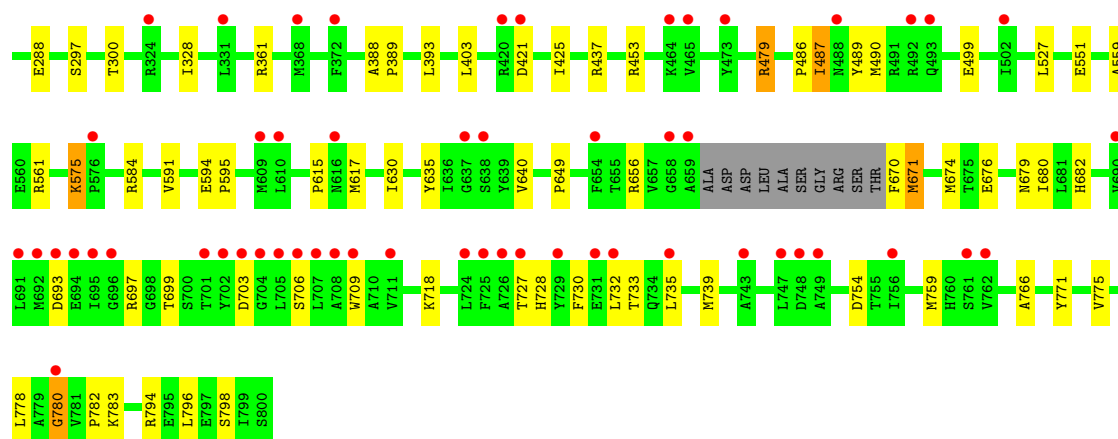
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

#### • Molecule 1: DNA MISMATCH REPAIR PROTEIN MUTS

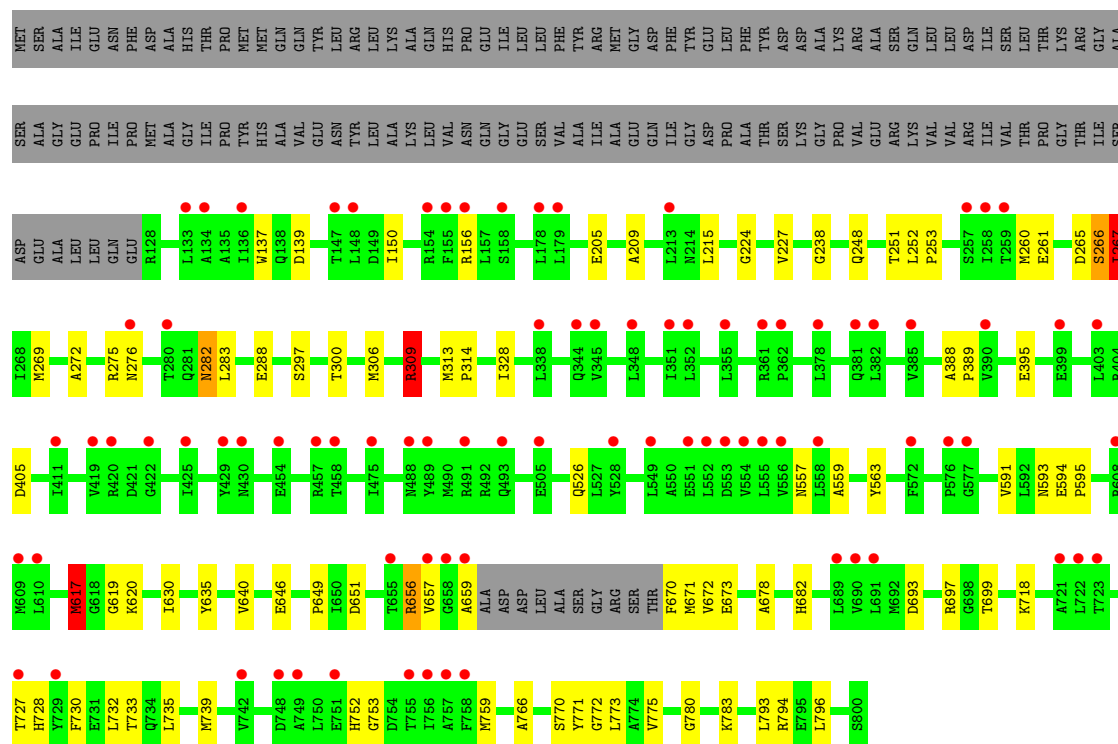
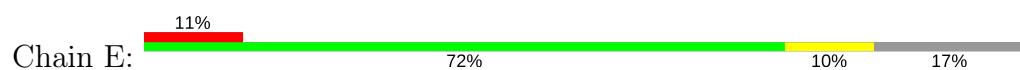


#### • Molecule 1: DNA MISMATCH REPAIR PROTEIN MUTS

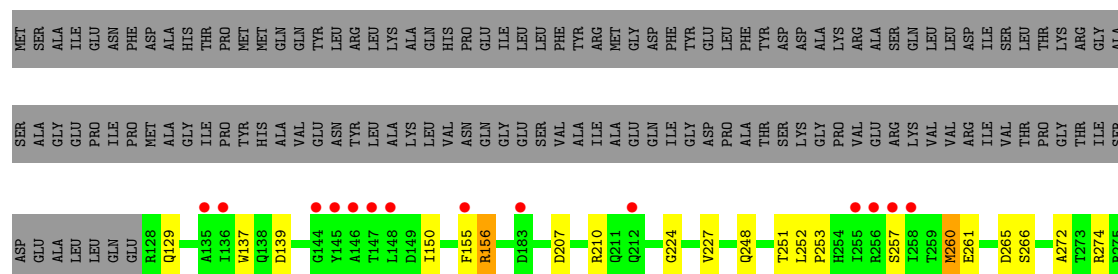




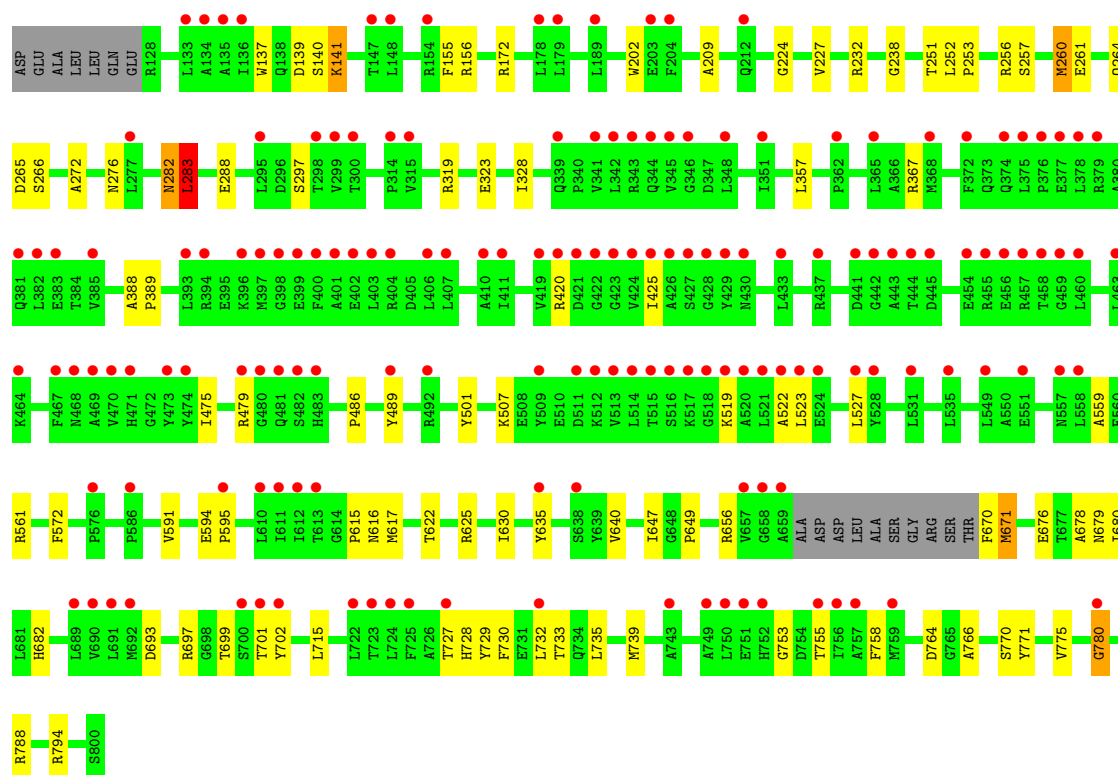
• Molecule 1: DNA MISMATCH REPAIR PROTEIN MUTS



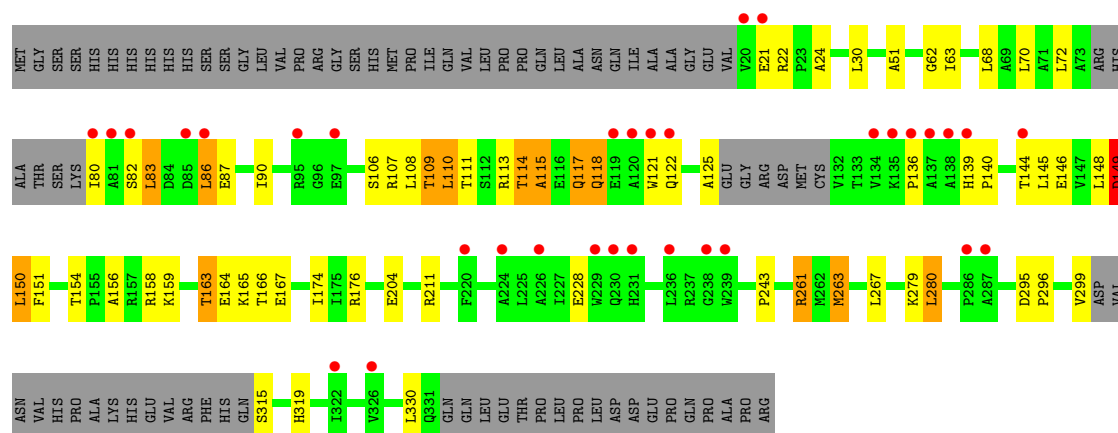
• Molecule 1: DNA MISMATCH REPAIR PROTEIN MUTS



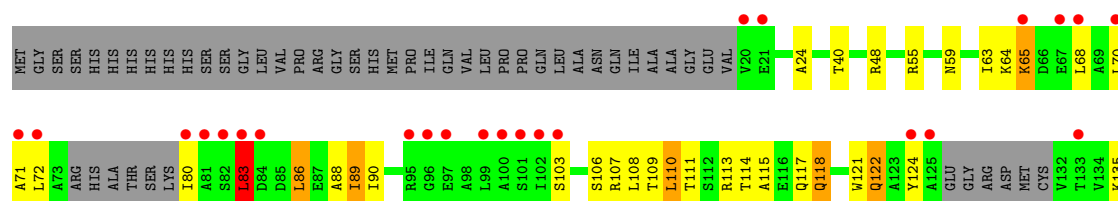


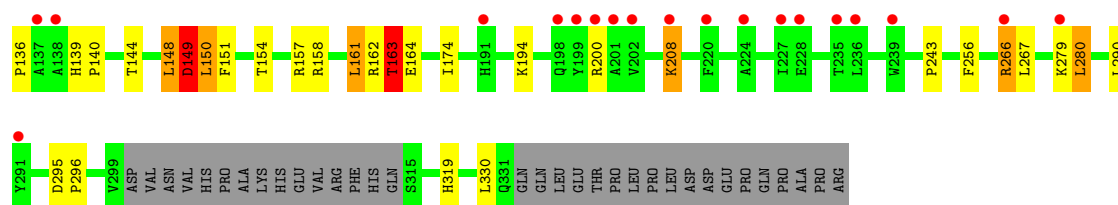


• Molecule 2: DNA MISMATCH REPAIR PROTEIN MUTL

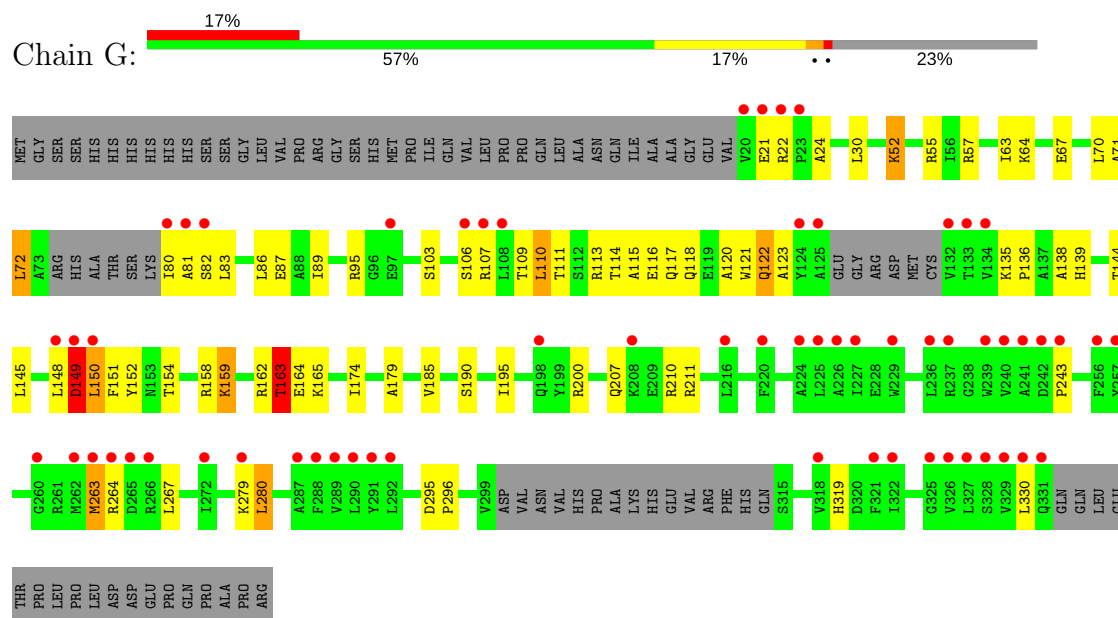


• Molecule 2: DNA MISMATCH REPAIR PROTEIN MUTL

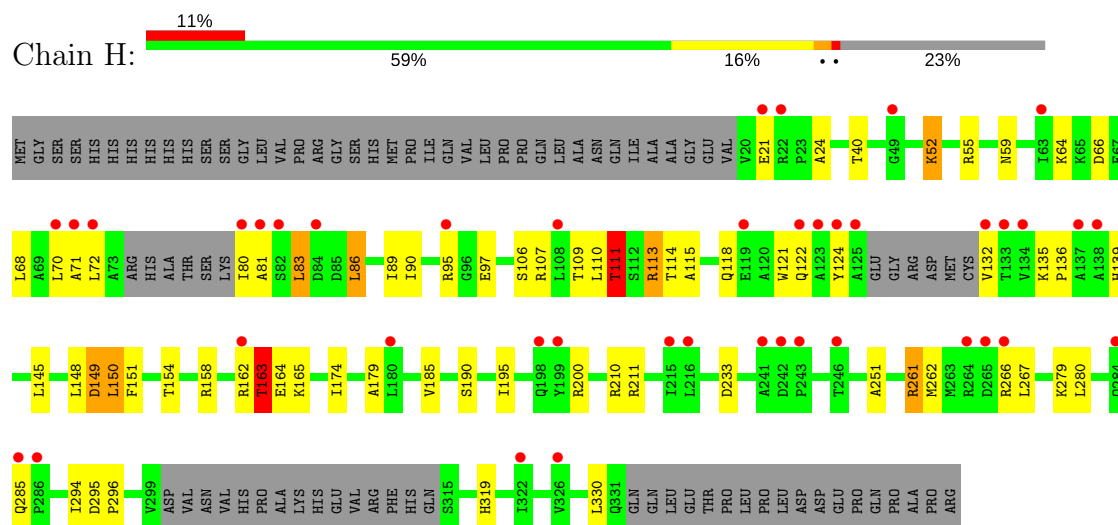




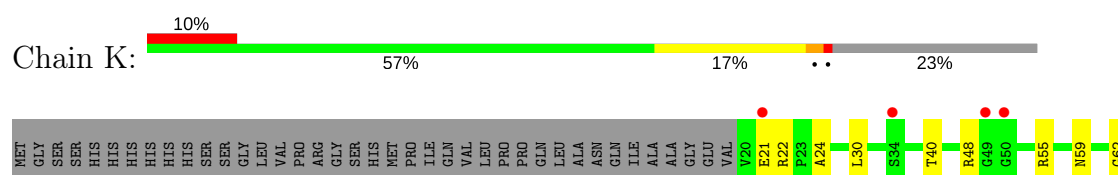
### • Molecule 2: DNA MISMATCH REPAIR PROTEIN MUTL



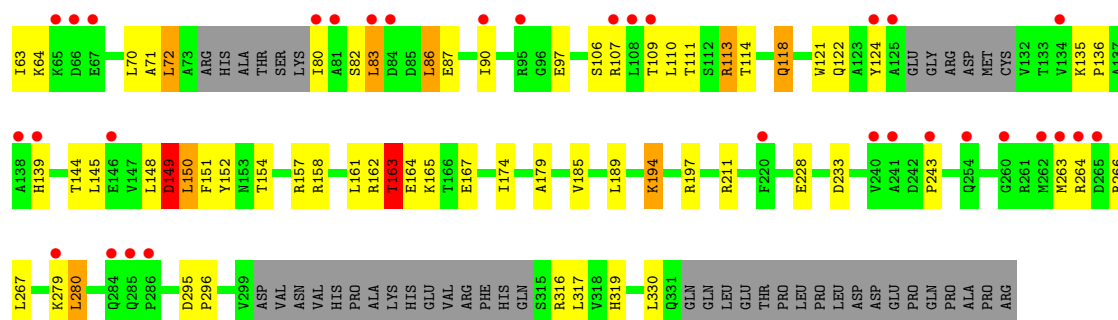
### • Molecule 2: DNA MISMATCH REPAIR PROTEIN MUTL



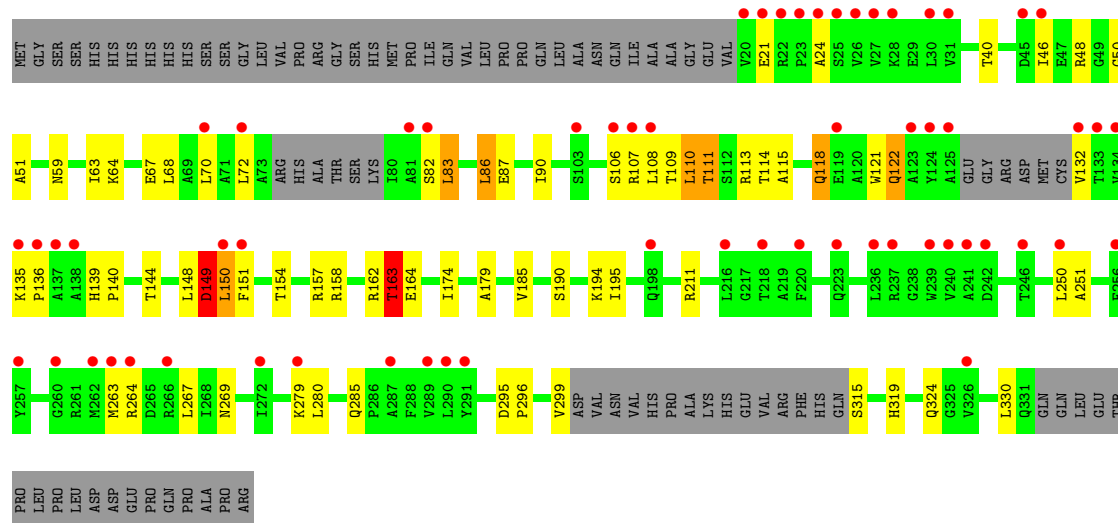
### • Molecule 2: DNA MISMATCH REPAIR PROTEIN MUTL







• Molecule 2: DNA MISMATCH REPAIR PROTEIN MUTL



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	380.63Å 126.46Å 243.35Å 90.00° 91.41° 90.00°	Depositor
Resolution (Å)	243.28 – 6.60 49.94 – 6.60	Depositor EDS
% Data completeness (in resolution range)	96.0 (243.28-6.60) 89.8 (49.94-6.60)	Depositor EDS
$R_{merge}$	0.21	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.75 (at 6.68Å)	Xtriage
Refinement program	REFMAC 5.8.0107	Depositor
R, $R_{free}$	0.259 , 0.291 0.266 , 0.293	Depositor DCC
$R_{free}$ test set	1088 reflections (5.39%)	DCC
Wilson B-factor (Å <sup>2</sup> )	170.9	Xtriage
Anisotropy	0.989	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 383.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.39$ , $\langle L^2 \rangle = 0.21$	Xtriage
Estimated twinning fraction	0.095 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.84	EDS
Total number of atoms	45054	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	253.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.63% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	0.64	1/5311 (0.0%)	0.77	8/7186 (0.1%)
1	B	0.63	0/5311	0.81	18/7186 (0.3%)
1	E	0.67	1/5311 (0.0%)	0.78	6/7186 (0.1%)
1	F	0.67	2/5311 (0.0%)	0.84	17/7186 (0.2%)
1	I	0.58	0/5311	0.74	5/7186 (0.1%)
1	J	0.59	0/5311	0.78	15/7186 (0.2%)
2	C	0.68	1/2288 (0.0%)	0.87	8/3096 (0.3%)
2	D	0.77	1/2288 (0.0%)	0.90	10/3096 (0.3%)
2	G	0.65	0/2288	0.83	6/3096 (0.2%)
2	H	0.73	2/2288 (0.1%)	0.87	4/3096 (0.1%)
2	K	0.67	0/2288	0.84	5/3096 (0.2%)
2	L	0.62	1/2288 (0.0%)	0.81	2/3096 (0.1%)
All	All	0.65	9/45594 (0.0%)	0.81	104/61692 (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	A	0	1
1	B	0	1
1	I	0	2
2	C	0	1
2	H	0	1
2	K	0	2
2	L	0	1
All	All	0	9

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	208	LYS	CE-NZ	14.66	1.85	1.49
1	F	671	MET	CG-SD	13.22	2.15	1.81
1	E	718	LYS	CD-CE	8.36	1.72	1.51
2	C	228	GLU	CG-CD	7.86	1.63	1.51
1	F	129	GLN	CG-CD	5.82	1.64	1.51

The worst 5 of 104 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	283	LEU	CB-CG-CD2	-13.89	87.39	111.00
1	B	283	LEU	CB-CG-CD2	-11.85	90.85	111.00
1	J	283	LEU	CB-CG-CD2	-10.76	92.70	111.00
1	I	671	MET	CG-SD-CE	9.48	115.37	100.20
1	F	394	ARG	NE-CZ-NH1	9.13	124.87	120.30

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	752	HIS	Peptide
1	B	487	ILE	Peptide
2	C	86	LEU	Peptide
2	H	81	ALA	Peptide
1	I	129	GLN	Peptide

## 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	51	0
1	B	5226	0	5283	57	0
1	E	5226	0	5283	79	0
1	F	5226	0	5283	74	0
1	I	5226	0	5283	84	0
1	J	5226	0	5283	87	0
2	C	2252	0	2272	69	0
2	D	2252	0	2272	51	0
2	G	2252	0	2271	59	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	H	2252	0	2272	58	0
2	K	2252	0	2272	66	0
2	L	2252	0	2272	49	0
3	A	31	0	13	2	0
3	B	31	0	13	1	0
3	E	31	0	13	9	0
3	F	31	0	13	4	0
3	I	31	0	13	3	0
3	J	31	0	13	5	0
All	All	45054	0	45407	655	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 7.

The worst 5 of 655 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:D:208:LYS:NZ	2:D:208:LYS:CE	1.85	1.38
1:F:671:MET:SD	1:F:671:MET:CG	2.15	1.34
2:C:110:LEU:CD2	2:C:145:LEU:HA	1.64	1.26
2:C:110:LEU:HD21	2:C:145:LEU:CA	1.76	1.14
2:G:83:LEU:HA	2:G:87:GLU:HA	1.18	1.13

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	659/800 (82%)	626 (95%)	31 (5%)	2 (0%)	44 81
1	B	659/800 (82%)	627 (95%)	30 (5%)	2 (0%)	44 81

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	659/800 (82%)	629 (95%)	27 (4%)	3 (0%)	32	74
1	F	659/800 (82%)	630 (96%)	26 (4%)	3 (0%)	32	74
1	I	659/800 (82%)	627 (95%)	31 (5%)	1 (0%)	51	84
1	J	659/800 (82%)	631 (96%)	25 (4%)	3 (0%)	32	74
2	C	277/369 (75%)	255 (92%)	18 (6%)	4 (1%)	13	54
2	D	277/369 (75%)	252 (91%)	20 (7%)	5 (2%)	10	49
2	G	277/369 (75%)	250 (90%)	22 (8%)	5 (2%)	10	49
2	H	277/369 (75%)	247 (89%)	25 (9%)	5 (2%)	10	49
2	K	277/369 (75%)	253 (91%)	21 (8%)	3 (1%)	17	60
2	L	277/369 (75%)	255 (92%)	18 (6%)	4 (1%)	13	54
All	All	5616/7014 (80%)	5282 (94%)	294 (5%)	40 (1%)	25	68

5 of 40 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	115	ALA
2	D	89	ILE
2	D	115	ALA
2	D	163	THR
2	G	115	ALA

### 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%)	538 (98%)	12 (2%)	57	79
1	B	550/662 (83%)	534 (97%)	16 (3%)	48	73
1	E	550/662 (83%)	535 (97%)	15 (3%)	50	74
1	F	550/662 (83%)	535 (97%)	15 (3%)	50	74
1	I	550/662 (83%)	536 (98%)	14 (2%)	53	77
1	J	550/662 (83%)	539 (98%)	11 (2%)	60	82

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	C	235/308 (76%)	222 (94%)	13 (6%)	25	58
2	D	235/308 (76%)	224 (95%)	11 (5%)	30	62
2	G	235/308 (76%)	226 (96%)	9 (4%)	38	67
2	H	235/308 (76%)	221 (94%)	14 (6%)	22	55
2	K	235/308 (76%)	223 (95%)	12 (5%)	28	60
2	L	235/308 (76%)	219 (93%)	16 (7%)	18	51
All	All	4710/5820 (81%)	4552 (97%)	158 (3%)	42	69

5 of 158 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	288	GLU
2	G	159	LYS
2	L	111	THR
1	F	361	ARG
1	F	738	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
2	G	118	GLN
2	H	118	GLN
2	K	324	GLN
1	F	248	GLN
2	L	118	GLN

### 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	2.11	10 (34%)	28,52,52	2.93	13 (46%)
3	ANP	B	1801	-	29,33,33	2.34	11 (37%)	28,52,52	2.11	7 (25%)
3	ANP	E	1801	-	29,33,33	1.81	7 (24%)	28,52,52	2.18	7 (25%)
3	ANP	F	1801	-	29,33,33	2.08	8 (27%)	28,52,52	2.11	9 (32%)
3	ANP	I	1801	1	29,33,33	2.08	8 (27%)	28,52,52	2.33	8 (28%)
3	ANP	J	1801	-	29,33,33	1.78	7 (24%)	28,52,52	2.33	9 (32%)

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	1	-	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	E	1801	ANP	PG-O2G	-3.47	1.47	1.56
3	J	1801	ANP	PG-O2G	-3.41	1.47	1.56
3	I	1801	ANP	PB-O2B	-3.24	1.47	1.56
3	F	1801	ANP	PG-O3G	-3.18	1.48	1.56
3	E	1801	ANP	PG-O3G	-3.09	1.48	1.56

The worst 5 of 53 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1801	ANP	O1G-PG-N3B	-8.53	99.03	111.79
3	E	1801	ANP	N3-C2-N1	-6.88	122.87	128.86
3	J	1801	ANP	N3-C2-N1	-6.26	123.40	128.86
3	I	1801	ANP	N3-C2-N1	-6.11	123.54	128.86
3	F	1801	ANP	N3-C2-N1	-5.88	123.74	128.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 24 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1801	ANP	2	0
3	B	1801	ANP	1	0
3	E	1801	ANP	9	0
3	F	1801	ANP	4	0
3	I	1801	ANP	3	0
3	J	1801	ANP	5	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	663/800 (82%)	0.79	83 (12%) 4 12	177, 220, 306, 373	0
1	B	663/800 (82%)	0.75	66 (9%) 8 15	196, 245, 311, 362	0
1	E	663/800 (82%)	0.80	84 (12%) 4 12	191, 224, 348, 398	0
1	F	663/800 (82%)	0.73	68 (10%) 7 14	181, 220, 296, 350	0
1	I	663/800 (82%)	1.12	111 (16%) 2 9	198, 244, 528, 643	0
1	J	663/800 (82%)	1.25	158 (23%) 1 7	213, 266, 510, 602	0
2	C	285/369 (77%)	0.81	33 (11%) 5 13	200, 237, 282, 323	0
2	D	285/369 (77%)	0.89	43 (15%) 3 10	193, 234, 273, 296	0
2	G	285/369 (77%)	1.08	61 (21%) 1 7	217, 239, 319, 347	0
2	H	285/369 (77%)	0.93	41 (14%) 3 10	210, 227, 268, 294	0
2	K	285/369 (77%)	0.90	36 (12%) 4 12	197, 233, 295, 329	0
2	L	285/369 (77%)	1.05	61 (21%) 1 7	233, 264, 336, 375	0
All	All	5688/7014 (81%)	0.92	845 (14%) 3 10	177, 237, 355, 643	0

The worst 5 of 845 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	421	ASP	13.9
1	I	422	GLY	13.1
1	I	749	ALA	8.3
1	J	515	THR	8.3
1	I	420	ARG	7.8

### 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
3	ANP	A	1801	31/31	0.86	0.32	-0.87	198,206,212,215	0
3	ANP	I	1801	31/31	0.84	0.38	-0.91	198,200,203,204	0
3	ANP	F	1801	31/31	0.88	0.33	-0.99	195,202,206,209	0
3	ANP	B	1801	31/31	0.78	0.33	-1.04	216,223,227,230	0
3	ANP	J	1801	31/31	0.91	0.23	-1.11	218,226,233,235	0
3	ANP	E	1801	31/31	0.93	0.28	-1.13	198,201,206,207	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.