Full wwPDB X-ray Structure Validation Report (i)

Nov 19, 2024 – 03:13 pm GMT

PDB ID : 8R7E

Title : MutSbeta bound to compound CHDI-00898647 in the canonical DNA-mismatch bound form

Authors : Thomsen, M.; Thieulin-Pardo, G.; Steinbacher, S.; Brace, G.; Tillack, K.; Johnson, P.; Ritzefeld, M.; Schaertl, S.; Frush, E.; Warfield, B.; Ballantyne, G.; Lee, J.-H.; Witte, D.; Finley, M.; Prasad, B.; Monteagudo, E.; Plotnikov, N.; Pacifici, R.; Maillard, M.; Wilkinson, H.; Iyer, R.; Dominguez, C.; Vogt, T.; Felsenfeld, D.; Haque, T.

Deposited on : 2023-11-24

Resolution : 2.78 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

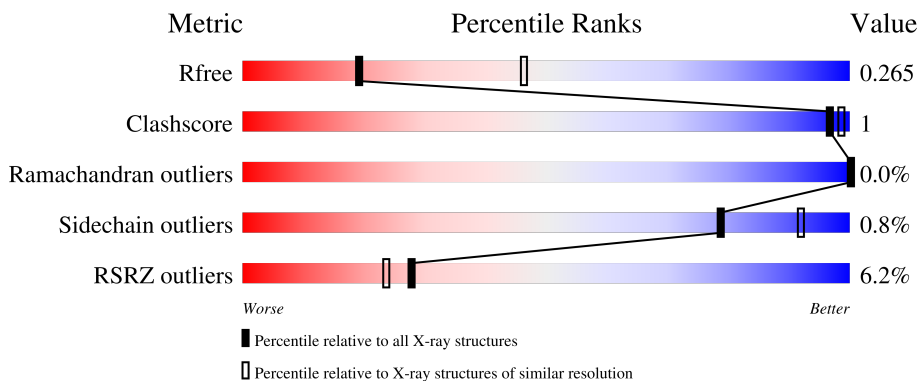|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11

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 2.78 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	Percentile Ranks	Value
Rfree		0.265
Clashscore		1
Ramachandran outliers		0.0%
Sidechain outliers		0.8%
RSRZ outliers		6.2%

Worse   *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	4924 (2.80-2.76)
Clashscore	180529	5458 (2.80-2.76)
Ramachandran outliers	177936	5386 (2.80-2.76)
Sidechain outliers	177891	5388 (2.80-2.76)
RSRZ outliers	164620	4926 (2.80-2.76)

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 <=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	930	7% 87% • 11%
1	E	930	5% 91% • 6%
2	B	918	2% 87% • 8%

Continued on next page...

Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	918	9% / 87% / 10%
3	C	24	88% / 8%
3	G	24	88% / 12%
4	D	24	83% / 12%
4	H	24	96%

2 Entry composition ⓘ

There are 10 unique types of molecules in this entry. The entry contains 28977 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 Msh2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	825	Total	C	N	O	S	373	1	0
			6570	4183	1114	1237	36			
1	E	875	Total	C	N	O	S	339	2	0
			6956	4423	1181	1316	36			

- Molecule 2 is a protein called DNA mismatch repair protein Msh3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	840	Total	C	N	O	S	177	0	0
			6721	4289	1148	1254	30			
2	F	827	Total	C	N	O	S	347	1	0
			6618	4220	1136	1232	30			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	217	GLY	-	expression tag	UNP P20585
B	218	PRO	-	expression tag	UNP P20585
F	217	GLY	-	expression tag	UNP P20585
F	218	PRO	-	expression tag	UNP P20585

- Molecule 3 is a DNA chain called DNA (5'-D(*TP*CP*TP*AP*TP*CP*TP*GP*AP*AP* GP*CP*CP*GP*AP*TP*CP*GP*AP*TP*GP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	22	Total	C	N	O	P	0	0	0
			449	215	82	131	21			
3	G	21	Total	C	N	O	P	0	0	0
			432	205	80	126	21			

- Molecule 4 is a DNA chain called DNA (5'-D(*CP*AP*TP*CP*GP*AP*TP*CP*GP*CP* AP*GP*CP*TP*TP*CP*AP*GP*AP*TP*AP*GP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	23	Total	C	N	O	P	0	0	0
			469	224	88	135	22			
4	H	23	Total	C	N	O	P	0	0	0
			469	224	88	135	22			

- Molecule 5 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: $C_{10}H_{15}N_5O_{10}P_2$).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	0
			27	10	5	10	2		
5	E	1	Total	C	N	O	P	0	0
			27	10	5	10	2		

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

EDO

O2

OH

C1

C2

HO

O1

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total 4	C 2	O 2	0	0
6	B	1	Total 4	C 2	O 2	0	0
6	B	1	Total 4	C 2	O 2	0	0
6	E	1	Total 4	C 2	O 2	0	0
6	F	1	Total 4	C 2	O 2	0	0

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total 1	Cl 1	0	0
7	B	1	Total 1	Cl 1	0	0
7	E	3	Total 3	Cl 3	0	0

- Molecule 8 is Abivertinib (three-letter code: YKQ) (formula: $C_{26}H_{26}FN_7O_2$) (labeled as "Ligand of Interest" by depositor).

YKQ

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
8	B	1	Total 36	C 26	F 1	N 7	O 2	0	0
8	F	1	Total 36	C 26	F 1	N 7	O 2	5	0

- Molecule 9 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	B	1	Total 5	O 4	S 1	0	0
9	B	1	Total 5	O 4	S 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	B	1	Total 5	O 4	S 1	0	0
9	F	1	Total 5	O 4	S 1	0	0
9	F	1	Total 5	O 4	S 1	0	0

- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	22	Total 22	O 22	0	0
10	B	27	Total 27	O 27	0	0
10	E	29	Total 29	O 29	0	0
10	F	34	Total 34	O 34	0	0
10	G	2	Total 2	O 2	0	0
10	H	3	Total 3	O 3	0	0

3 Residue-property plots ⓘ

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.
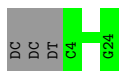
• Molecule 1: DNA mismatch repair protein Msh2
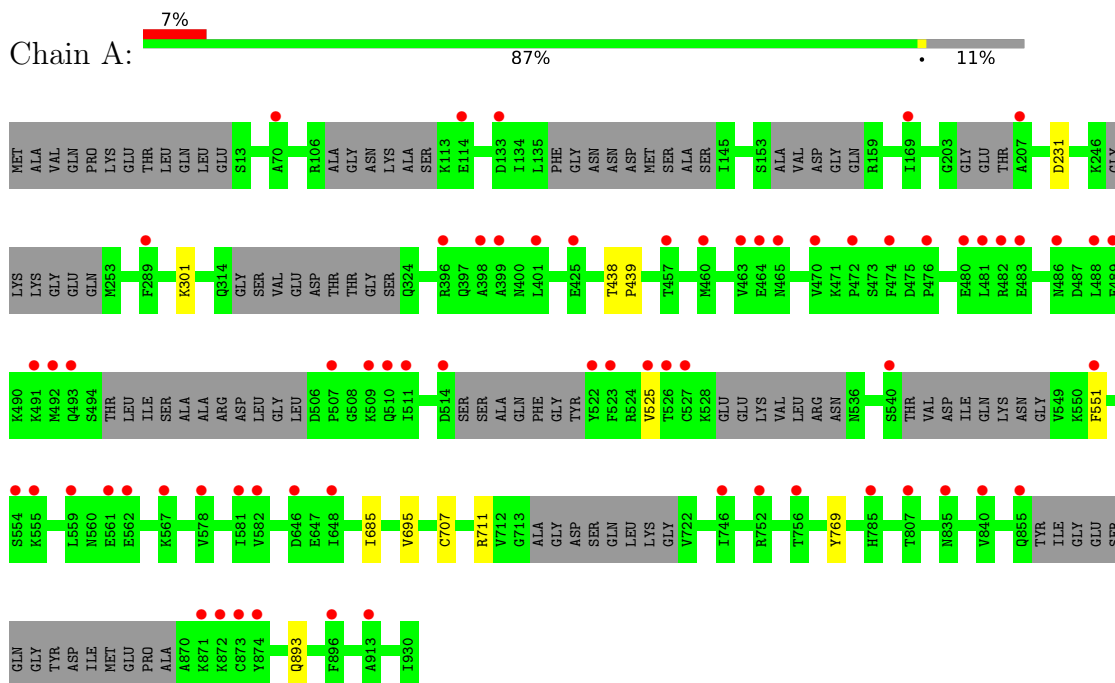
Chain A:

• Molecule 1: DNA mismatch repair protein Msh2

Chain E:

- Molecule 2: DNA mismatch repair protein Msh3

Chain B:

- Molecule 2: DNA mismatch repair protein Msh3

Chain F:

- Molecule 3: DNA (5'-D(*TP*CP*TP*AP*TP*CP*TP*GP*AP*AP*GP*CP*CP*GP*AP*TP
  *CP*GP*AP*TP*GP*G)-3')

Chain C:

- Molecule 3: DNA (5'-D(*TP*CP*TP*AP*TP*CP*TP*GP*AP*AP*GP*CP*CP*GP*AP*TP
  *CP*GP*AP*TP*GP*G)-3')
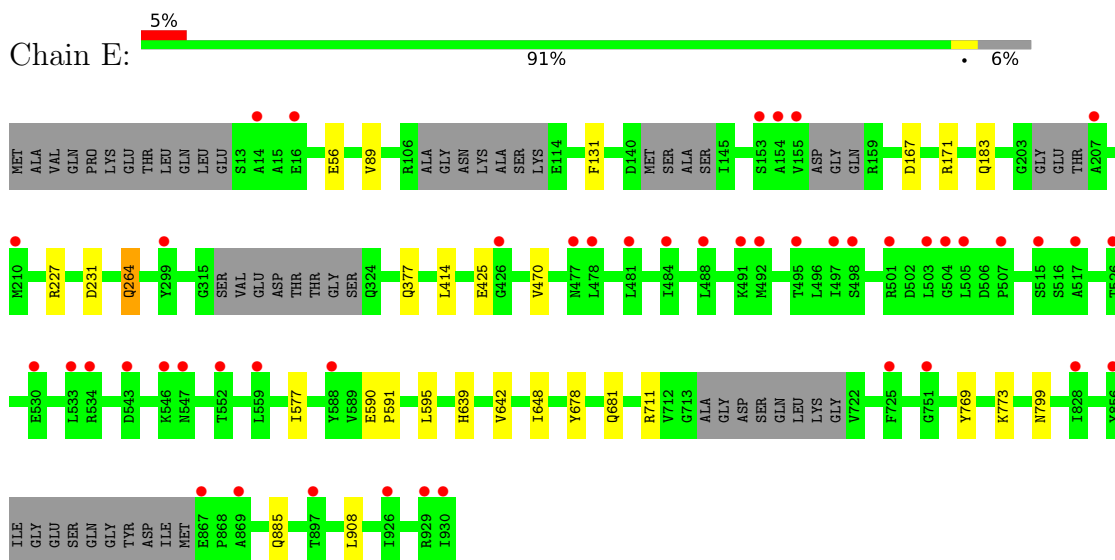
Chain G:

• Molecule 4: DNA (5'-D(*CP*AP*TP*CP*GP*AP*TP*CP*GP*CP*AP*GP*CP*TP*TP*CP *AP*GP*AP*TP*AP*GP*G)-3')

Chain D:

• Molecule 4: DNA (5'-D(*CP*AP*TP*CP*GP*AP*TP*CP*GP*CP*AP*GP*CP*TP*TP*CP *AP*GP*AP*TP*AP*GP*G)-3')

Chain H:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	100.47Å 104.10Å 122.02Å 109.33° 91.81° 109.50°	Depositor
Resolution (Å)	113.66 – 2.78 113.66 – 2.78	Depositor EDS
% Data completeness (in resolution range)	72.8 (113.66-2.78) 72.8 (113.66-2.78)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.94 (at 2.77Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.223 , 0.266 0.224 , 0.265	Depositor DCC
R_{free} test set	5400 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å^2)	62.2	Xtriage
Anisotropy	0.047	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 48.8	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	28977	wwPDB-VP
Average B, all atoms (Å^2)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.85% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: SO4, CL, ADP, EDO, YKQ

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	0.69	0/6669	0.76	0/8975
1	E	0.69	0/7066	0.77	0/9518
2	B	0.68	0/6845	0.77	0/9240
2	F	0.68	0/6737	0.77	0/9089
3	C	0.30	0/503	0.75	0/775
3	G	0.32	0/484	0.73	0/745
4	D	0.33	0/526	0.79	0/810
4	H	0.31	0/526	0.76	0/810
All	All	0.66	0/29356	0.77	0/39962

There are no bond length outliers.

There are no bond angle outliers.

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	6570	0	6637	5	0
1	E	6956	0	7010	12	0
2	B	6721	0	6816	22	0
2	F	6618	0	6720	9	0
3	C	449	0	250	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	G	432	0	237	0	0
4	D	469	0	260	2	0
4	H	469	0	260	0	0
5	A	27	0	12	0	0
5	E	27	0	12	0	0
6	A	4	0	6	0	0
6	B	8	0	12	0	0
6	E	4	0	6	0	0
6	F	4	0	6	0	0
7	A	1	0	0	0	0
7	B	1	0	0	0	0
7	E	3	0	0	0	0
8	B	36	0	0	0	0
8	F	36	0	0	0	0
9	B	15	0	0	0	0
9	F	10	0	0	0	0
10	A	22	0	0	0	0
10	B	27	0	0	0	0
10	E	29	0	0	2	0
10	F	34	0	0	0	0
10	G	2	0	0	0	0
10	H	3	0	0	0	0
All	All	28977	0	28244	48	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 1.

All (48) 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:E:183:GLN:HG3	10:E:1111:HOH:O	1.97	0.63
1:E:470:VAL:HG21	1:E:577:ILE:HD11	1.82	0.61
2:B:1079:PRO:HG2	2:B:1082:ILE:HD13	1.89	0.54
2:F:891:VAL:HG21	2:F:998:ILE:HG12	1.91	0.52
2:B:884:LEU:HB3	2:B:1004:LEU:HD22	1.92	0.51
1:A:769:TYR:CD1	2:B:1082:ILE:HD11	2.45	0.51
2:B:548:LEU:HD21	2:B:581:THR:HG23	1.95	0.49
1:A:525:VAL:HG23	1:A:551:PHE:CE1	2.47	0.48
2:B:563:VAL:HG21	2:B:866:VAL:HA	1.95	0.48
2:B:564:LEU:HA	2:B:840:VAL:HG21	1.94	0.48
1:E:89:VAL:HG11	1:E:131:PHE:CD2	2.50	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:769:TYR:CE2	1:E:773:LYS:HD2	2.50	0.47
1:A:685:ILE:HG23	1:A:695:VAL:HB	1.96	0.47
2:B:878:VAL:HG21	2:B:1057:TYR:HB2	1.97	0.46
2:F:548:LEU:HD11	2:F:581:THR:HG22	1.98	0.46
2:F:563:VAL:HG21	2:F:866:VAL:HA	1.97	0.46
1:E:264:GLN:HE21	1:E:264:GLN:HA	1.81	0.46
3:C:6:DA:C2	4:D:47:DA:C2	3.05	0.45
2:B:548:LEU:HD21	2:B:581:THR:CG2	2.47	0.45
2:B:852:GLN:HE22	2:B:856:LYS:HD3	1.82	0.45
2:B:891:VAL:HG22	2:B:1005:THR:HB	1.99	0.44
2:B:891:VAL:HG21	2:B:998:ILE:HG12	1.99	0.44
2:F:535:MET:HE3	2:F:537:ILE:HD11	2.00	0.44
2:B:607:VAL:HG22	2:B:658:ALA:HB1	2.00	0.43
1:E:769:TYR:CD2	2:F:1082:ILE:HD11	2.53	0.43
2:B:855:ARG:HD2	2:B:1004:LEU:HG	2.01	0.42
2:B:254:TYR:O	2:B:287:THR:HG23	2.18	0.42
2:B:764:ILE:HD12	2:B:764:ILE:N	2.34	0.42
1:E:414:LEU:HD21	1:E:595:LEU:HD11	2.01	0.42
1:E:678:TYR:O	1:E:681:GLN:HG2	2.20	0.42
2:F:254:TYR:O	2:F:287:THR:HG23	2.20	0.42
1:E:167:ASP:O	1:E:171:ARG:N	2.53	0.41
2:B:756:ILE:HG12	2:B:764:ILE:HD11	2.01	0.41
2:B:756:ILE:CG1	2:B:764:ILE:HD11	2.50	0.41
2:F:314:ALA:HB1	2:F:794:GLN:HG3	2.02	0.41
2:F:680:SER:N	2:F:681:PRO:CD	2.84	0.41
2:B:855:ARG:NE	2:B:968:SER:O	2.54	0.41
1:E:639:HIS:HB3	1:E:642:VAL:HB	2.02	0.41
1:A:438:THR:HB	1:A:439:PRO:HD3	2.03	0.41
2:B:680:SER:N	2:B:681:PRO:CD	2.84	0.41
2:F:426:LEU:HD13	2:F:440:ILE:HD12	2.02	0.41
4:D:39:DC:H2'	4:D:40:DT:C6	2.56	0.41
1:A:301:LYS:HB2	1:A:707:CYS:HB3	2.02	0.40
2:B:645:LEU:HD22	2:B:679:LEU:HD22	2.03	0.40
1:E:590:GLU:HB3	1:E:591:PRO:HD3	2.03	0.40
2:B:506:ILE:HG22	2:B:519:LEU:HD11	2.02	0.40
2:B:855:ARG:HD3	2:B:969:GLN:HA	2.04	0.40
1:E:227:ARG:NH2	10:E:1106:HOH:O	2.53	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	800/930 (86%)	772 (96%)	28 (4%)	0	100	100
1	E	861/930 (93%)	833 (97%)	28 (3%)	0	100	100
2	B	828/918 (90%)	798 (96%)	30 (4%)	0	100	100
2	F	812/918 (88%)	783 (96%)	28 (3%)	1 (0%)	48	76
All	All	3301/3696 (89%)	3186 (96%)	114 (4%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	322	ARG

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	725/804 (90%)	722 (100%)	3 (0%)	89	96
1	E	764/804 (95%)	754 (99%)	10 (1%)	65	86
2	B	750/818 (92%)	745 (99%)	5 (1%)	81	93
2	F	738/818 (90%)	731 (99%)	7 (1%)	75	91
All	All	2977/3244 (92%)	2952 (99%)	25 (1%)	79	92

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	231	ASP
1	A	711	ARG
1	A	893	GLN
2	B	626	CYS
2	B	633	CYS
2	B	781	HIS
2	B	938	ARG
2	B	1052	PHE
1	E	56	GLU
1	E	231	ASP
1	E	264	GLN
1	E	377	GLN
1	E	425	GLU
1	E	648	ILE
1	E	711	ARG
1	E	799	ASN
1	E	885	GLN
1	E	908	LEU
2	F	524	ASN
2	F	633	CYS
2	F	708	ASP
2	F	809	ASP
2	F	855	ARG
2	F	875	ASP
2	F	895	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	24	GLN
1	A	395	GLN
1	A	885	GLN
2	B	271	ASN
2	B	852	GLN
1	E	97	GLN
1	E	264	GLN
1	E	462	GLN
1	E	645	GLN
1	E	885	GLN
2	F	524	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 19 ligands modelled in this entry, 5 are monoatomic - leaving 14 for Mogul analysis.

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	EDO	B	1207	-	3,3,3	0.09	0	2,2,2	0.21	0
9	SO4	F	1202	-	4,4,4	0.43	0	6,6,6	0.12	0
9	SO4	B	1205	-	4,4,4	0.37	0	6,6,6	0.04	0
5	ADP	E	1001	-	24,29,29	0.64	0	29,45,45	0.82	1 (3%)
6	EDO	B	1204	-	3,3,3	0.11	0	2,2,2	0.39	0
9	SO4	B	1202	-	4,4,4	0.38	0	6,6,6	0.07	0
6	EDO	F	1204	-	3,3,3	0.05	0	2,2,2	0.25	0
9	SO4	B	1203	-	4,4,4	0.38	0	6,6,6	0.06	0
9	SO4	F	1203	-	4,4,4	0.39	0	6,6,6	0.04	0
8	YKQ	B	1201	-	38,40,40	0.70	0	48,56,56	2.63	13 (27%)
5	ADP	A	1001	-	24,29,29	0.68	0	29,45,45	0.72	1 (3%)
6	EDO	E	1004	-	3,3,3	0.11	0	2,2,2	0.38	0
6	EDO	A	1002	-	3,3,3	0.11	0	2,2,2	0.33	0
8	YKQ	F	1201	-	38,40,40	0.89	2 (5%)	48,56,56	2.61	14 (29%)

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
6	EDO	B	1207	-	-	0/1/1/1	-
5	ADP	E	1001	-	-	1/12/32/32	0/3/3/3
6	EDO	B	1204	-	-	1/1/1/1	-
6	EDO	F	1204	-	-	1/1/1/1	-
8	YKQ	B	1201	-	-	6/18/34/34	0/5/5/5
5	ADP	A	1001	-	-	3/12/32/32	0/3/3/3
6	EDO	E	1004	-	-	1/1/1/1	-
6	EDO	A	1002	-	-	1/1/1/1	-
8	YKQ	F	1201	-	-	4/18/34/34	0/5/5/5

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	F	1201	YKQ	C21-C20	-2.89	1.39	1.43
8	F	1201	YKQ	C13-N12	2.20	1.41	1.36

All (29) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	B	1201	YKQ	C20-C21-N34	-10.35	117.18	124.40
8	F	1201	YKQ	C20-C21-N34	-10.14	117.33	124.40
8	F	1201	YKQ	C31-C29-N28	9.42	120.04	113.84
8	B	1201	YKQ	C31-C29-N28	5.63	117.55	113.84
8	B	1201	YKQ	C13-N14-C15	5.23	121.22	115.28
8	B	1201	YKQ	F37-C36-C8	4.79	122.78	118.42
8	F	1201	YKQ	C13-N34-C21	4.61	123.31	115.18
8	B	1201	YKQ	C13-N34-C21	4.32	122.81	115.18
8	B	1201	YKQ	C36-C8-N5	4.30	125.58	120.47
8	F	1201	YKQ	C13-N14-C15	4.27	120.13	115.28
8	B	1201	YKQ	C9-C8-N5	-3.95	116.01	122.30
8	B	1201	YKQ	C35-C36-C8	-3.85	120.16	123.34
8	B	1201	YKQ	O22-C21-C20	3.72	122.11	115.69
8	F	1201	YKQ	C35-C36-C8	-3.69	120.29	123.34
8	F	1201	YKQ	N14-C13-N34	-3.56	120.60	126.23
8	B	1201	YKQ	C6-N5-C8	-3.29	108.48	116.27
8	B	1201	YKQ	N14-C13-N34	-3.12	121.29	126.23
8	F	1201	YKQ	C11-N12-C13	-3.01	120.50	129.23
8	F	1201	YKQ	C9-C8-N5	-2.95	117.60	122.30
8	B	1201	YKQ	C11-C35-C36	2.76	121.01	118.76
8	B	1201	YKQ	C27-N28-C29	-2.68	124.21	128.26

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	F	1201	YKQ	O30-C29-C31	-2.41	118.87	122.72
8	F	1201	YKQ	F37-C36-C8	2.36	120.56	118.42
8	F	1201	YKQ	C36-C8-N5	2.34	123.25	120.47
8	F	1201	YKQ	C23-O22-C21	2.27	122.62	118.16
8	F	1201	YKQ	C11-C35-C36	2.09	120.46	118.76
5	A	1001	ADP	C5-C6-N6	2.08	123.51	120.35
5	E	1001	ADP	C5-C6-N6	2.07	123.50	120.35
8	F	1201	YKQ	C32-C31-C29	-2.01	119.97	122.27

There are no chirality outliers.

All (18) torsion outliers are listed below:

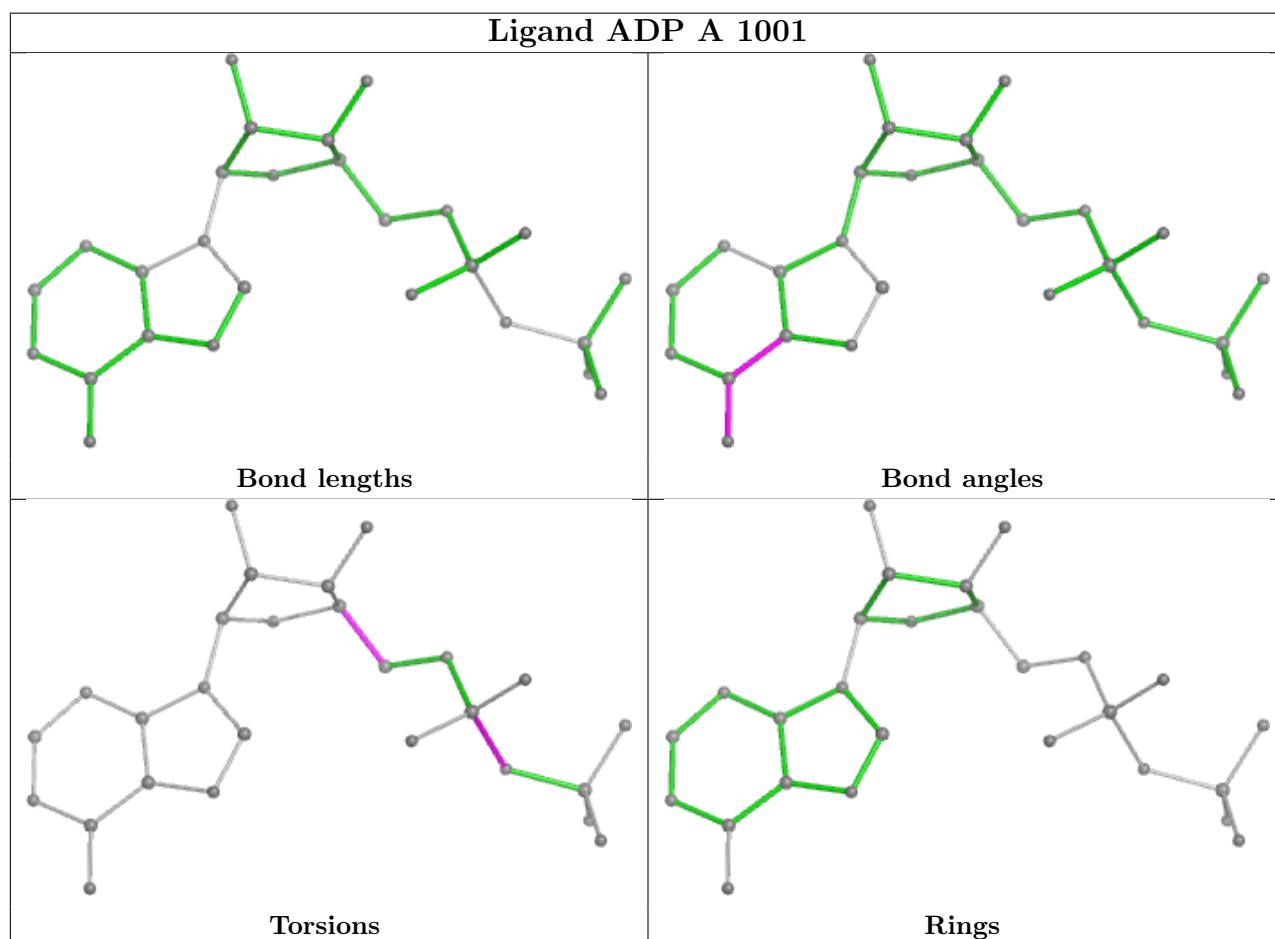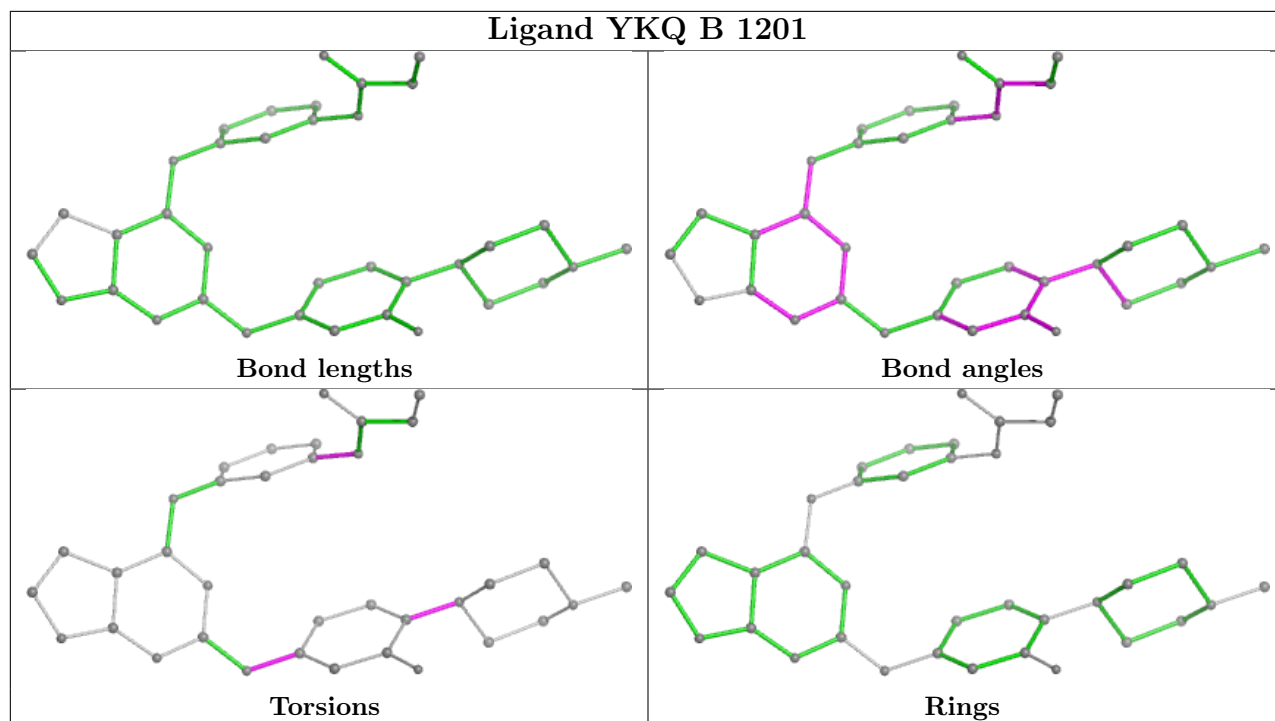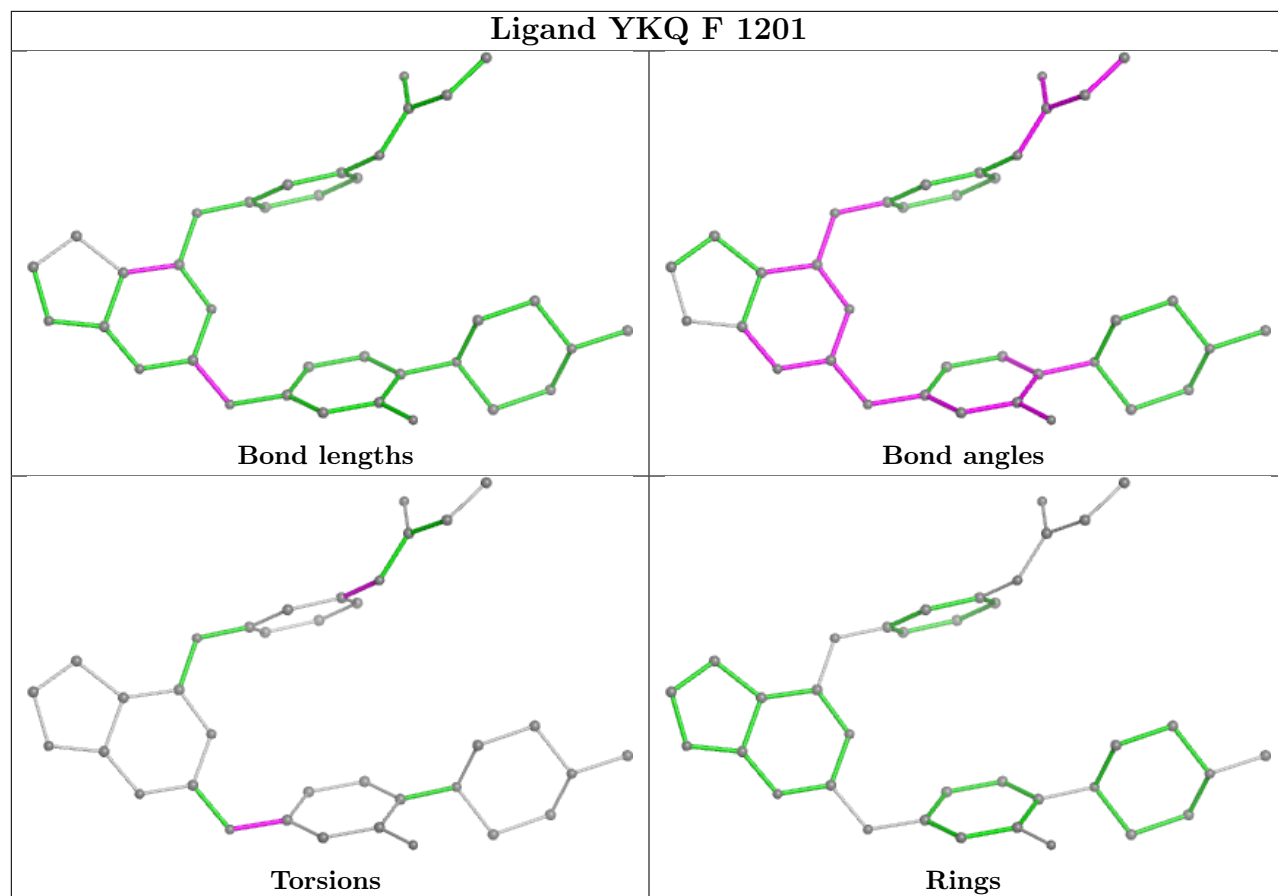| Mol	Chain	Res	Type	Atoms
8	B	1201	YKQ	C36-C8-N5-C6
5	E	1001	ADP	O4'-C4'-C5'-O5'
8	B	1201	YKQ	C26-C27-N28-C29
8	B	1201	YKQ	C33-C27-N28-C29
8	F	1201	YKQ	C33-C27-N28-C29
8	B	1201	YKQ	C35-C11-N12-C13
8	F	1201	YKQ	C26-C27-N28-C29
8	B	1201	YKQ	C10-C11-N12-C13
8	F	1201	YKQ	C35-C11-N12-C13
6	B	1204	EDO	O1-C1-C2-O2
5	A	1001	ADP	PB-O3A-PA-O1A
5	A	1001	ADP	PB-O3A-PA-O2A
8	F	1201	YKQ	C10-C11-N12-C13
6	F	1204	EDO	O1-C1-C2-O2
8	B	1201	YKQ	C9-C8-N5-C6
6	A	1002	EDO	O1-C1-C2-O2
5	A	1001	ADP	O4'-C4'-C5'-O5'
6	E	1004	EDO	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring

in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand ADP E 1001

Bond lengths

Bond angles

Torsions

Rings

Ligand YKQ B 1201

Bond lengths

Bond angles

Torsions

Rings

Ligand ADP A 1001

Bond lengths

Bond angles

Torsions

Rings

Ligand YKQ F 1201

Bond lengths

Bond angles

Torsions

Rings
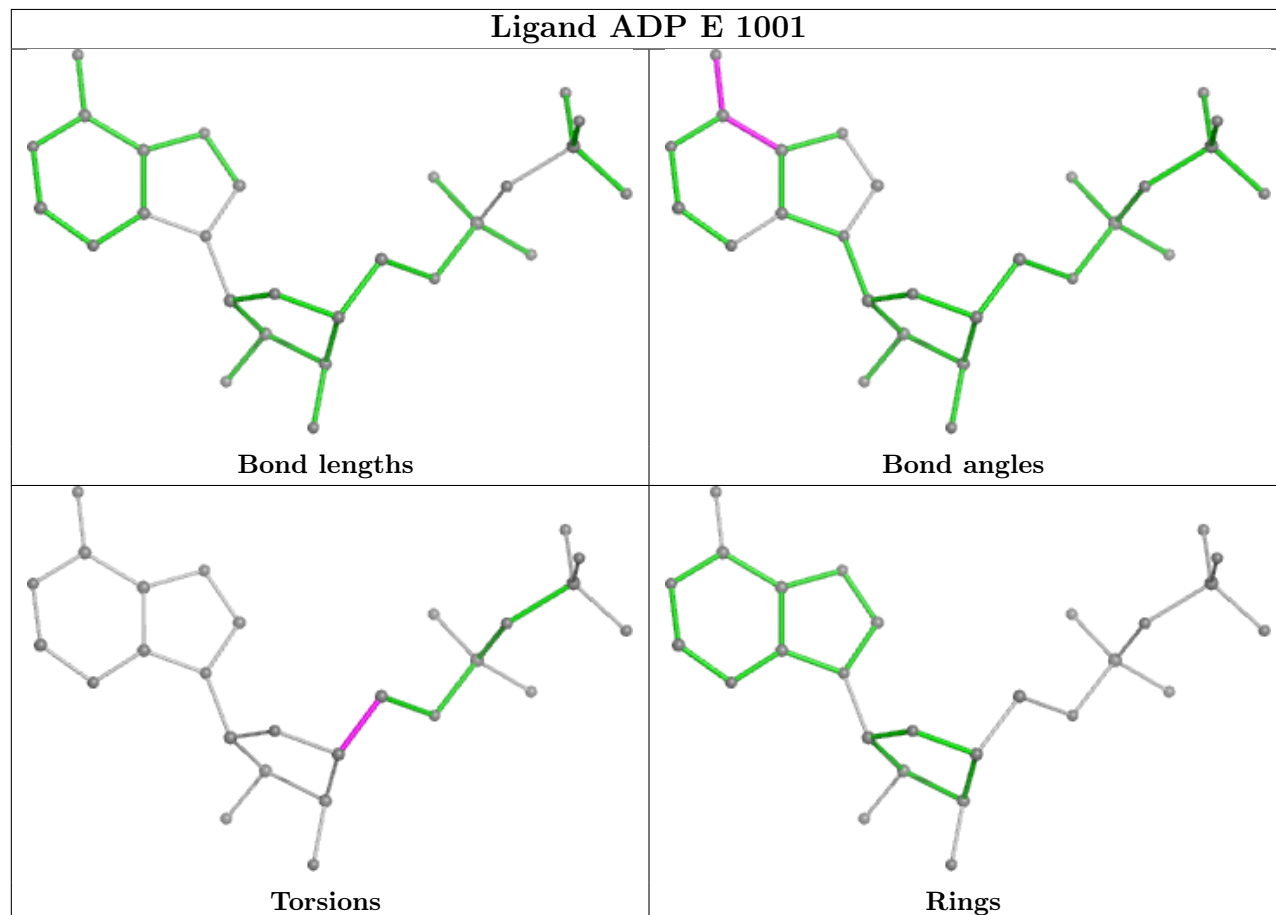
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	825/930 (88%)	0.52	67 (8%)	19	17	25, 64, 118, 168	124 (15%)
1	E	875/930 (94%)	0.37	46 (5%)	33	28	24, 59, 108, 138	123 (14%)
2	B	840/918 (91%)	0.12	20 (2%)	59	53	25, 54, 87, 139	67 (7%)
2	F	827/918 (90%)	0.63	82 (9%)	14	12	31, 64, 113, 159	132 (15%)
3	C	22/24 (91%)	0.13	0	100	100	53, 97, 141, 145	0
3	G	21/24 (87%)	-0.27	0	100	100	41, 81, 130, 145	0
4	D	23/24 (95%)	-0.09	0	100	100	47, 93, 136, 146	0
4	H	23/24 (95%)	-0.14	0	100	100	43, 70, 140, 161	0
All	All	3456/3792 (91%)	0.40	215 (6%)	28	24	24, 60, 111, 168	446 (12%)

All (215) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	490	SER	4.7
2	F	774	THR	4.7
2	F	1069	GLY	4.3
1	E	503	LEU	4.2
1	E	504	GLY	4.1
1	E	547	ASN	4.1
1	A	855	GLN	4.0
1	A	488	LEU	4.0
2	F	492	ILE	4.0
2	F	668	LEU	3.9
2	B	362	ASP	3.8
1	A	523	PHE	3.8
1	E	856	TYR	3.8
1	E	526	THR	3.8
2	F	711	LEU	3.8
1	E	517	ALA	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	514	ASP	3.7
2	F	682	VAL	3.6
2	F	655	ILE	3.6
2	F	675	ILE	3.5
1	A	509	LYS	3.5
1	A	522	TYR	3.5
2	F	1055	PHE	3.5
1	A	551	PHE	3.5
2	F	658	ALA	3.4
1	A	481	LEU	3.4
2	B	448	VAL	3.3
2	F	707	SER	3.3
2	F	615	LEU	3.3
2	B	490	SER	3.2
2	F	597	VAL	3.2
1	E	552	THR	3.1
2	F	607	VAL	3.1
2	F	851	VAL	3.1
1	E	533	LEU	3.1
2	F	828	LEU	3.1
1	A	463	VAL	3.1
2	F	605	SER	3.1
1	E	725	PHE	3.1
2	F	652	PHE	3.1
1	E	492	MET	3.1
2	F	671	VAL	3.1
1	A	581	ILE	3.0
2	F	449	GLN	3.0
1	A	873	CYS	3.0
2	B	1032	PHE	3.0
1	E	926	ILE	3.0
1	E	929	ARG	3.0
2	F	1066	ARG	3.0
1	A	396	ARG	2.9
2	F	650	SER	2.9
1	A	562	GLU	2.9
2	F	1072	VAL	2.9
1	A	465	ASN	2.9
1	A	559	LEU	2.9
1	A	582	VAL	2.9
1	A	840	VAL	2.9
2	F	657	PRO	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	594	LEU	2.9
1	A	648	ILE	2.9
1	E	828	ILE	2.9
1	A	578	VAL	2.9
2	F	646	TYR	2.9
1	A	133	ASP	2.9
2	F	678	LEU	2.9
1	A	510	GLN	2.8
1	E	507	PRO	2.8
1	A	492	MET	2.8
1	A	169	ILE	2.8
1	A	540	SER	2.8
2	F	666	ASP	2.8
1	E	501	ARG	2.8
1	A	114	GLU	2.8
2	F	598	SER	2.8
2	F	600	VAL	2.8
1	A	567	LYS	2.8
1	A	561	GLU	2.8
1	A	70	ALA	2.8
2	F	446	VAL	2.7
1	A	399	ALA	2.7
1	E	530	GLU	2.7
2	F	1014	VAL	2.7
1	A	756	THR	2.7
1	E	588	TYR	2.7
1	A	489	GLU	2.7
1	E	546	LYS	2.7
2	F	448	VAL	2.6
1	A	493	GLN	2.6
1	E	488	LEU	2.6
2	B	1033	LEU	2.6
2	F	1075	LEU	2.6
1	E	751	GLY	2.6
2	F	674	GLU	2.6
2	F	653	GLN	2.6
2	F	656	ILE	2.6
1	E	491	LYS	2.6
1	A	646	ASP	2.6
2	F	667	LEU	2.6
1	A	464	GLU	2.6
1	E	505	LEU	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	154	ALA	2.6
2	F	477	ALA	2.6
2	F	608	PHE	2.5
1	E	930	ILE	2.5
2	B	492	ILE	2.5
1	E	210	MET	2.5
2	B	532	MET	2.5
1	A	483	GLU	2.5
2	F	837	LEU	2.5
1	E	155	VAL	2.5
2	F	534	PHE	2.5
2	F	1088[A]	HIS	2.5
2	F	611	ILE	2.5
1	A	476	PRO	2.5
2	B	446	VAL	2.5
2	F	921	TYR	2.5
1	E	497	ILE	2.5
1	A	398	ALA	2.5
2	F	672	ILE	2.5
2	F	651	GLU	2.4
1	A	554	SER	2.4
2	F	877	TYR	2.4
2	F	1056	LEU	2.4
1	E	534	ARG	2.4
1	E	498	SER	2.4
2	B	344	VAL	2.4
2	F	493	VAL	2.4
1	A	874	TYR	2.4
1	E	477	ASN	2.4
2	F	529	SER	2.4
1	A	896	PHE	2.4
2	F	838	ALA	2.4
1	A	486	ASN	2.4
2	F	601	LEU	2.4
1	A	752	ARG	2.4
2	B	449	GLN	2.4
1	E	153	SER	2.4
1	A	207	ALA	2.4
1	E	869	ALA	2.4
2	F	845	ASP	2.4
2	F	602	HIS	2.3
1	A	289	PHE	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	663	ILE	2.3
1	E	207	ALA	2.3
2	F	661	SER	2.3
1	E	478	LEU	2.3
2	F	229	TYR	2.3
2	F	1119	LEU	2.3
1	A	457	THR	2.3
2	F	750	GLN	2.3
2	F	866	VAL	2.3
1	A	401	LEU	2.3
1	A	460	MET	2.3
1	A	482	ARG	2.3
2	B	1115	ASN	2.3
2	B	467	HIS	2.3
1	A	511	ILE	2.3
2	F	590	ILE	2.3
2	B	711	LEU	2.3
2	F	447	SER	2.3
2	B	249	CYS	2.2
2	B	1034	VAL	2.2
2	F	551	GLN	2.2
2	F	582	GLN	2.2
1	E	484	ILE	2.2
1	E	897	THR	2.2
2	F	524	ASN	2.2
1	A	871	LYS	2.2
1	A	872	LYS	2.2
2	B	774	THR	2.2
2	F	536	THR	2.2
2	F	664	GLN	2.2
2	F	824	ALA	2.2
1	E	481	LEU	2.2
1	E	559	LEU	2.2
1	A	507	PRO	2.2
1	A	470	VAL	2.2
1	A	474	PHE	2.2
1	A	913	ALA	2.2
2	F	759	SER	2.1
2	F	1068	TYR	2.1
1	A	491	LYS	2.1
1	E	14	ALA	2.1
1	E	867	GLU	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	599	GLU	2.1
1	E	515	SER	2.1
2	F	670	THR	2.1
1	E	299	TYR	2.1
2	F	835	PHE	2.1
1	E	426	GLY	2.1
2	B	252	CYS	2.1
1	A	425	GLU	2.1
1	A	785[A]	HIS	2.1
1	A	525	VAL	2.1
1	A	746	ILE	2.1
1	E	543	ASP	2.1
2	F	450	ASP	2.1
1	E	16	GLU	2.1
2	B	1124	GLU	2.1
2	F	683	GLU	2.1
1	A	526	THR	2.1
1	A	835	ASN	2.1
1	A	472	PRO	2.0
2	F	681	PRO	2.0
1	A	480	GLU	2.0
1	A	527	CYS	2.0
1	A	555	LYS	2.0
2	B	239	MET	2.0
2	B	776	ALA	2.0
2	F	862	GLY	2.0
1	A	807	THR	2.0
1	E	495	THR	2.0
2	F	387	PHE	2.0
2	F	659	VAL	2.0
2	F	1120	GLN	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

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
9	SO4	F	1203	5/5	0.77	0.09	127,128,130,132	0
9	SO4	B	1203	5/5	0.81	0.09	110,111,112,112	0
9	SO4	F	1202	5/5	0.84	0.14	91,92,94,100	0
8	YKQ	F	1201	36/36	0.85	0.14	70,90,105,108	5
6	EDO	A	1002	4/4	0.87	0.21	66,69,70,70	0
7	CL	E	1002	1/1	0.90	0.09	74,74,74,74	0
9	SO4	B	1205	5/5	0.90	0.12	102,104,106,106	0
6	EDO	B	1204	4/4	0.91	0.10	46,47,47,47	0
6	EDO	E	1004	4/4	0.91	0.14	63,63,63,64	0
6	EDO	B	1207	4/4	0.93	0.14	56,57,57,59	0
7	CL	E	1005	1/1	0.93	0.06	78,78,78,78	0
8	YKQ	B	1201	36/36	0.93	0.10	45,57,75,77	0
5	ADP	E	1001	27/27	0.93	0.07	52,57,61,62	0
6	EDO	F	1204	4/4	0.94	0.12	44,45,46,48	0
9	SO4	B	1202	5/5	0.94	0.11	70,72,74,74	0
5	ADP	A	1001	27/27	0.94	0.07	43,48,56,59	0
7	CL	A	1003	1/1	0.96	0.05	57,57,57,57	0
7	CL	B	1206	1/1	0.97	0.05	63,63,63,63	0
7	CL	E	1003	1/1	0.97	0.05	76,76,76,76	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around YKQ F 1201:

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
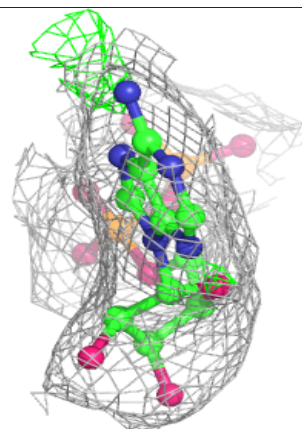mF_o-DF_c (at 3 rmsd) in purple (negative)
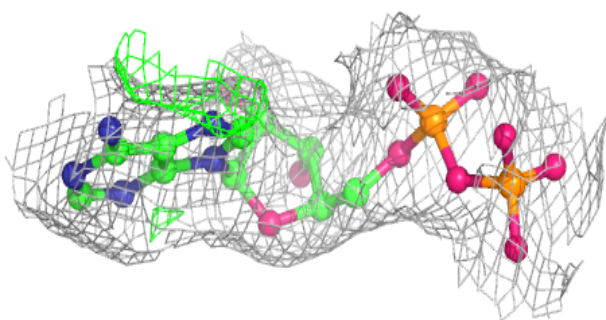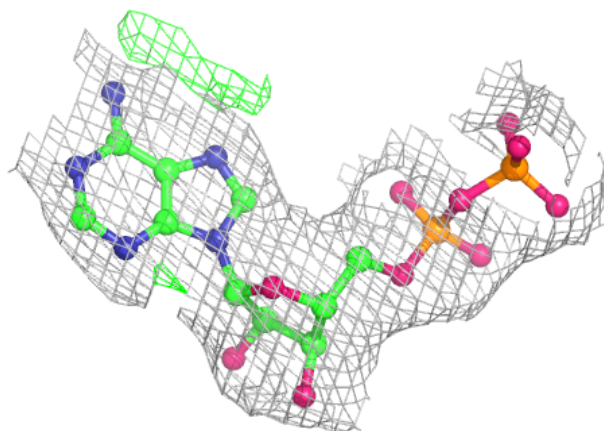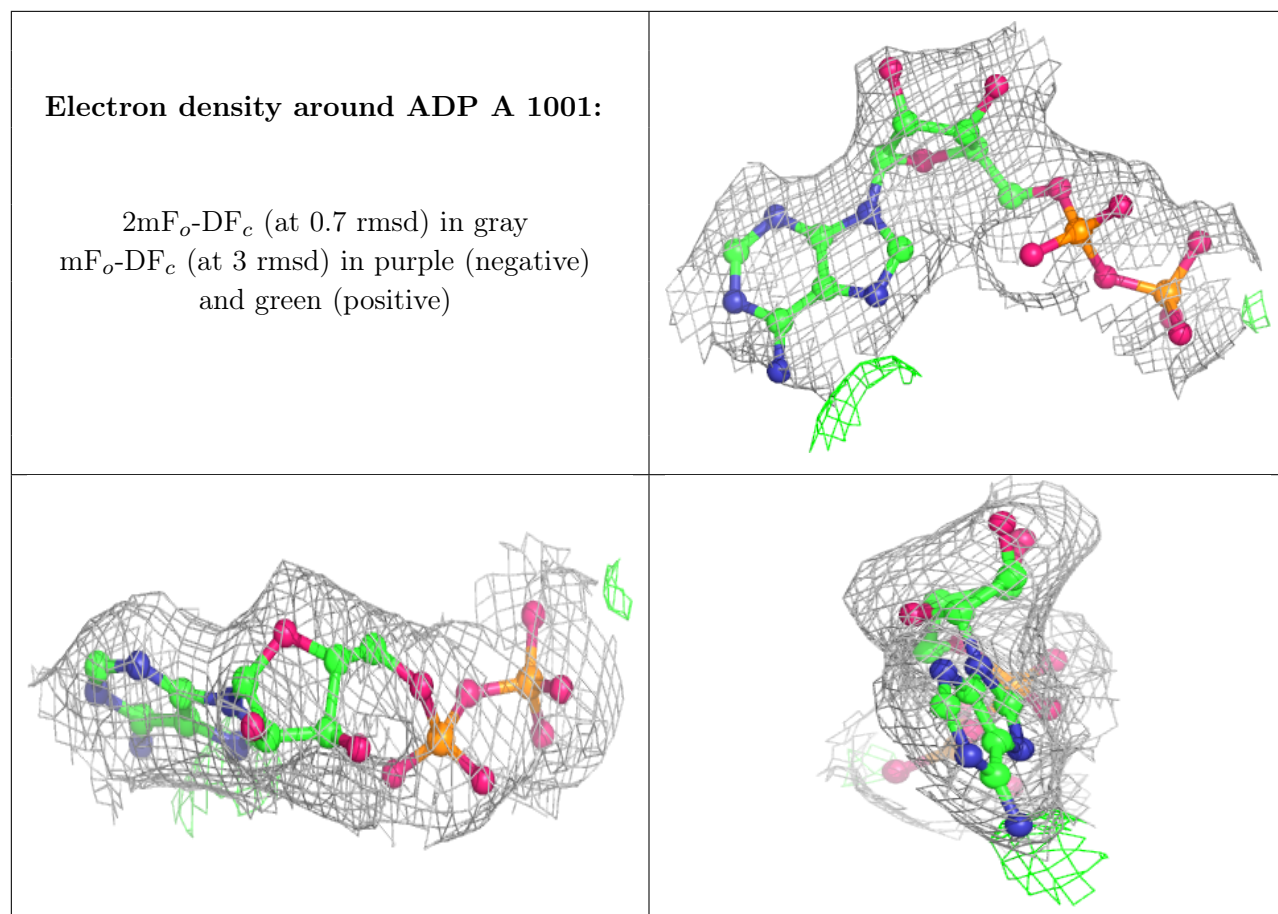and green (positive)
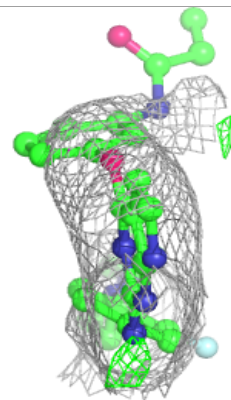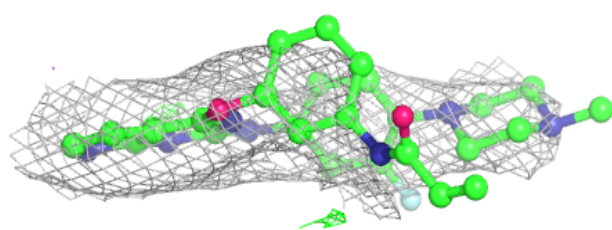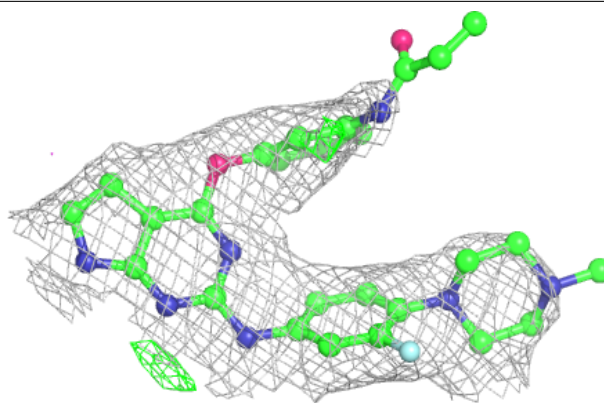
Electron density around YKQ B 1201:

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
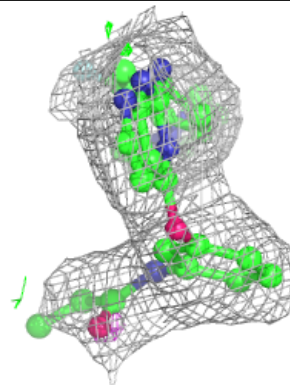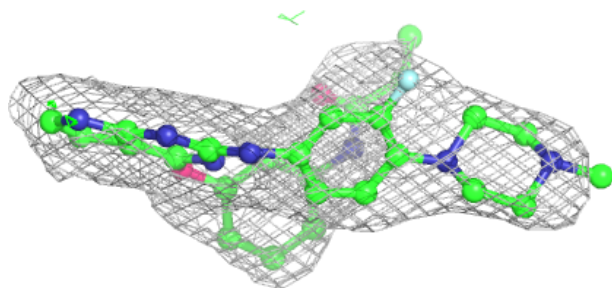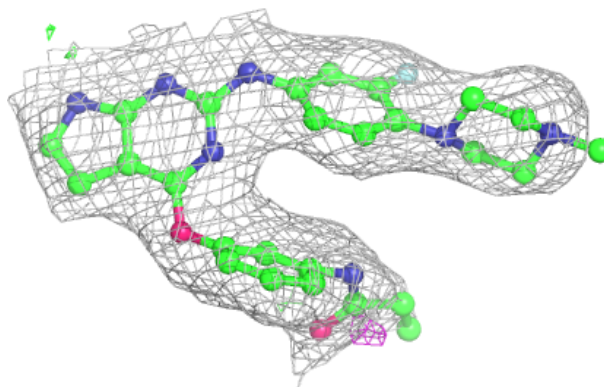
Electron density around ADP E 1001:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around ADP A 1001:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

6.5 Other polymers ⓘ

There are no such residues in this entry.