



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 20, 2023 – 02:09 AM EDT

PDB ID : 2HJR
Title : Crystal Structure of Cryptosporidium parvum malate dehydrogenase
Authors : Wernimont, A.K.; Dong, A.; Lew, J.; Hassani, A.; Ren, H.; Qiu, W.; Kozieradzki, I.; Weigelt, J.; Sundstrom, M.; Edwards, A.M.; Arrowsmith, C.H.; Bochkarev, A.; Hui, R.; Amani, M.; Structural Genomics Consortium (SGC)
Deposited on : 2006-06-30
Resolution : 2.20 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.35
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.35

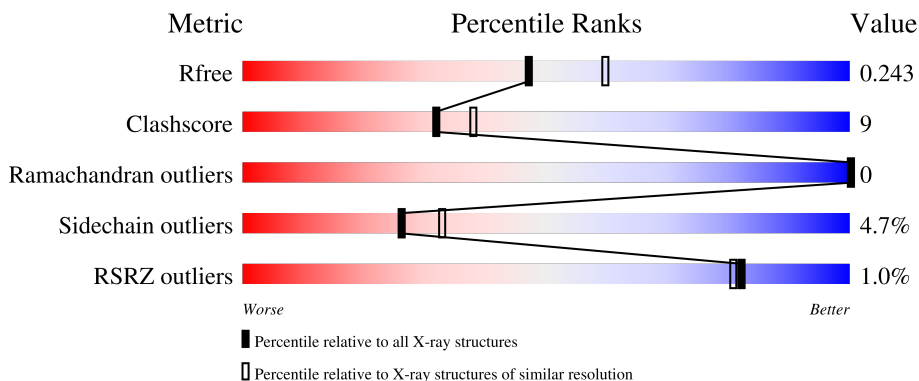
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

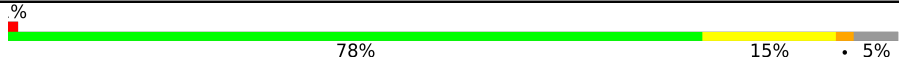


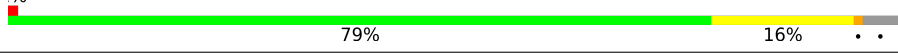
The reported resolution of this entry is 2.20 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.











Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	328	
1	B	328	
1	C	328	
1	D	328	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain	
1	E	328		%
1	F	328		%
1	G	328		%
1	H	328		%
1	I	328		%
1	J	328		2%
1	K	328		%
1	L	328		%

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
2	CIT	A	2001	-	X	X	-
2	CIT	C	2003	-	X	X	-
2	CIT	E	2008	-	-	X	-
2	CIT	F	2007	-	X	-	-
2	CIT	J	2012	-	-	X	-
2	CIT	L	2010	-	X	X	-

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 29965 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 Malate dehydrogenase.

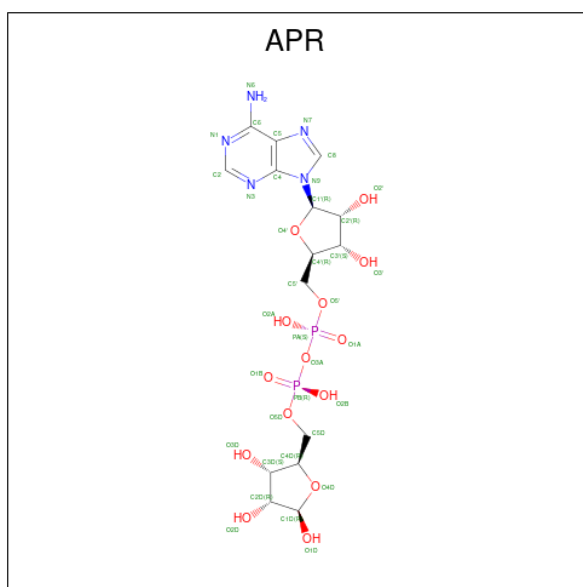
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	313	2332	1483	387	448	14	0	0	0
1	B	314	2340	1489	388	449	14	0	0	0
1	C	314	2340	1489	388	449	14	0	0	0
1	D	314	2340	1489	388	449	14	0	0	0
1	E	313	2332	1483	387	448	14	0	0	0
1	F	314	2340	1489	388	449	14	0	0	0
1	G	313	2332	1483	387	448	14	0	0	0
1	H	314	2340	1489	388	449	14	0	0	0
1	I	313	2332	1483	387	448	14	0	0	0
1	J	314	2340	1489	388	449	14	0	0	0
1	K	313	2332	1483	387	448	14	0	0	0
1	L	314	2340	1489	388	449	14	0	0	0

- Molecule 2 is CITRIC ACID (three-letter code: CIT) (formula: C₆H₈O₇).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 13 6 7	0	0
2	B	1	Total C O 13 6 7	0	0
2	C	1	Total C O 13 6 7	0	0
2	D	1	Total C O 13 6 7	0	0
2	E	1	Total C O 13 6 7	0	0
2	F	1	Total C O 13 6 7	0	0
2	G	1	Total C O 13 6 7	0	0
2	H	1	Total C O 13 6 7	0	0
2	I	1	Total C O 13 6 7	0	0
2	J	1	Total C O 13 6 7	0	0
2	K	1	Total C O 13 6 7	0	0
2	L	1	Total C O 13 6 7	0	0

- Molecule 3 is ADENOSINE-5-DIPHOSPHORIBOSE (three-letter code: APR) (formula: $C_{15}H_{23}N_5O_{14}P_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	A	1	Total 36	15	5	14	2	0	0
3	B	1	Total 36	15	5	14	2	0	0
3	C	1	Total 36	15	5	14	2	0	0
3	D	1	Total 36	15	5	14	2	0	0
3	E	1	Total 36	15	5	14	2	0	0
3	F	1	Total 36	15	5	14	2	0	0
3	G	1	Total 36	15	5	14	2	0	0
3	H	1	Total 36	15	5	14	2	0	0
3	I	1	Total 36	15	5	14	2	0	0
3	J	1	Total 36	15	5	14	2	0	0
3	K	1	Total 36	15	5	14	2	0	0
3	L	1	Total 36	15	5	14	2	0	0

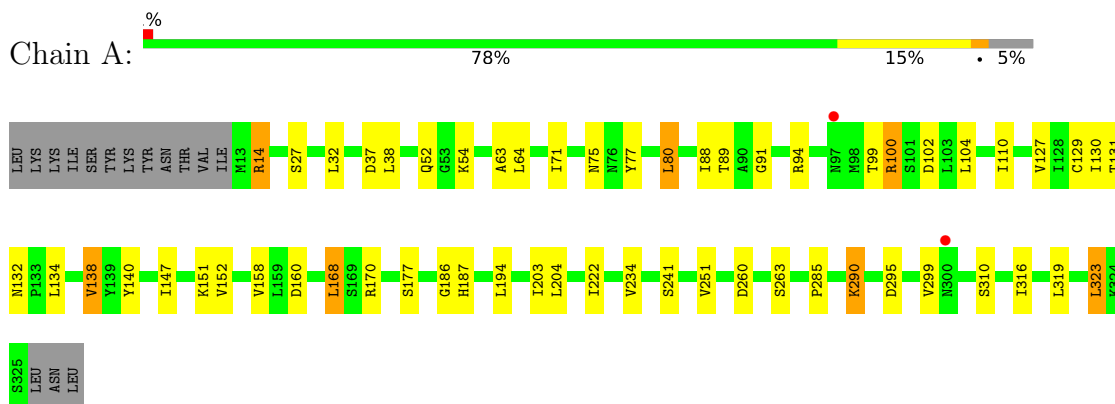
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	114	Total O 114 114	0	0
4	B	135	Total O 135 135	0	0
4	C	118	Total O 118 118	0	0
4	D	103	Total O 103 103	0	0
4	E	101	Total O 101 101	0	0
4	F	131	Total O 131 131	0	0
4	G	121	Total O 121 121	0	0
4	H	114	Total O 114 114	0	0
4	I	105	Total O 105 105	0	0
4	J	84	Total O 84 84	0	0
4	K	110	Total O 110 110	0	0
4	L	101	Total O 101 101	0	0

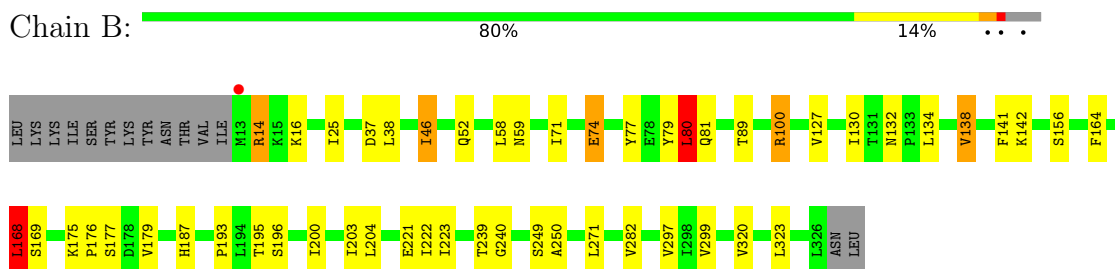
3 Residue-property plots [i](#)

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

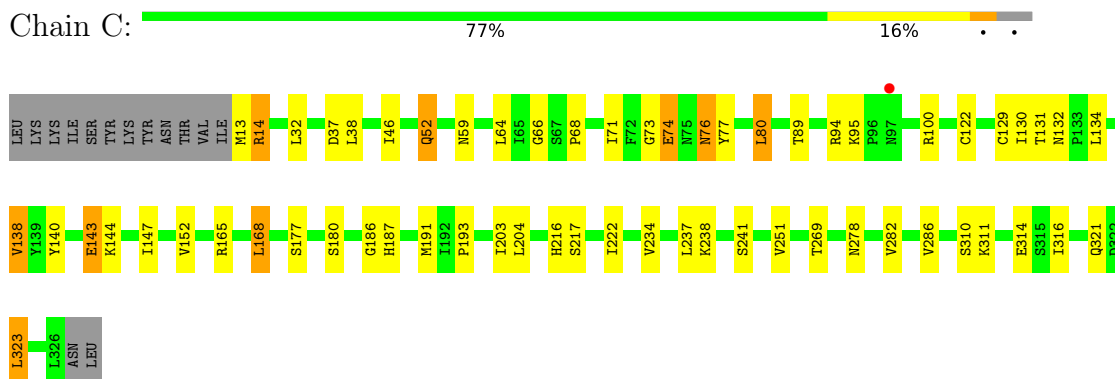
- Molecule 1: Malate dehydrogenase



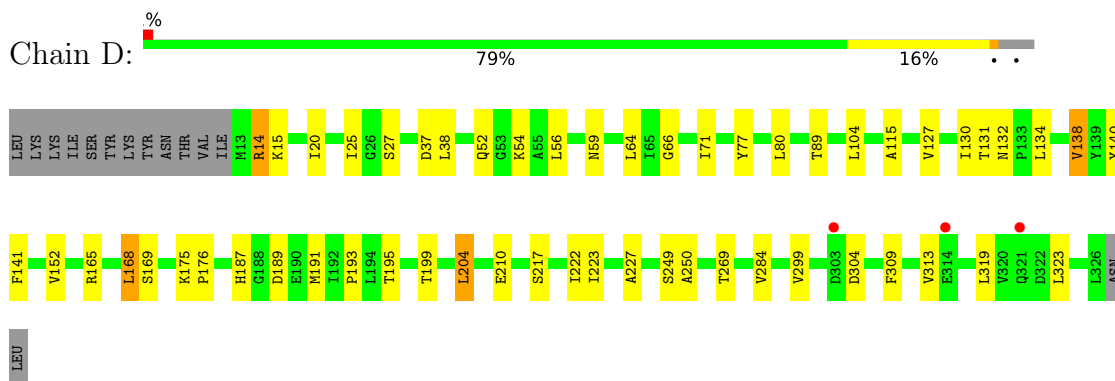
- Molecule 1: Malate dehydrogenase



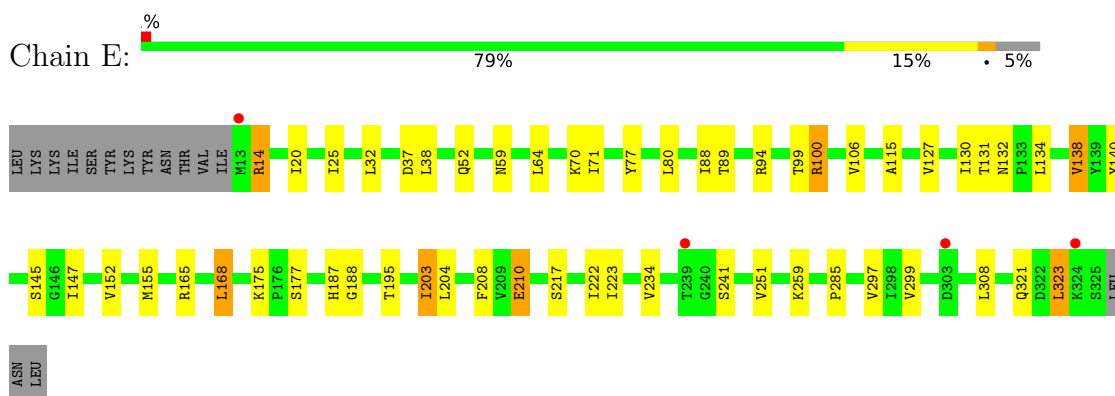
- Molecule 1: Malate dehydrogenase



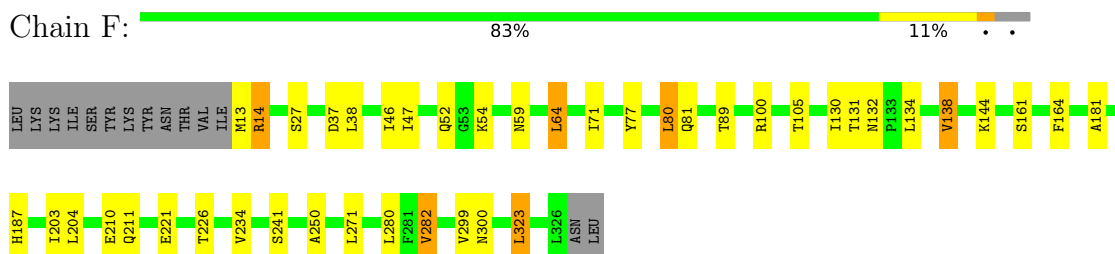
- Molecule 1: Malate dehydrogenase



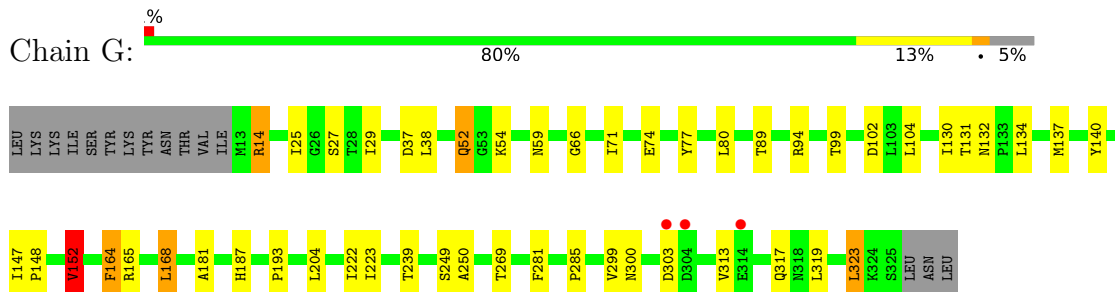
- Molecule 1: Malate dehydrogenase



- Molecule 1: Malate dehydrogenase

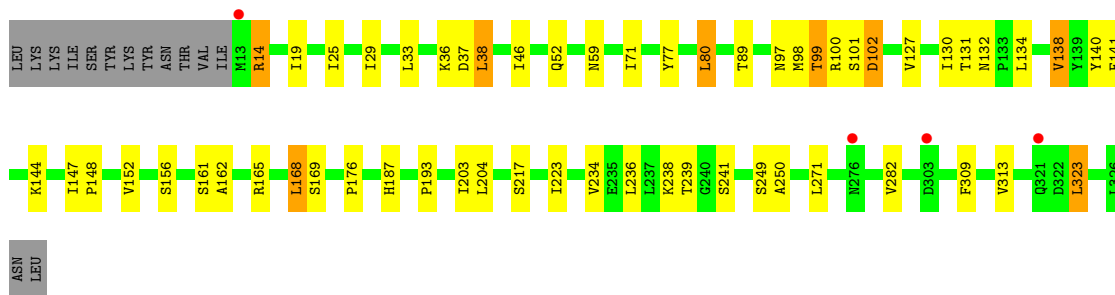


- Molecule 1: Malate dehydrogenase

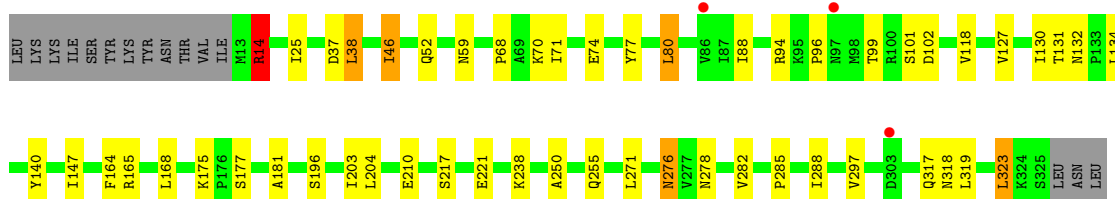
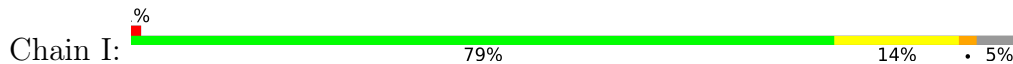


- Molecule 1: Malate dehydrogenase

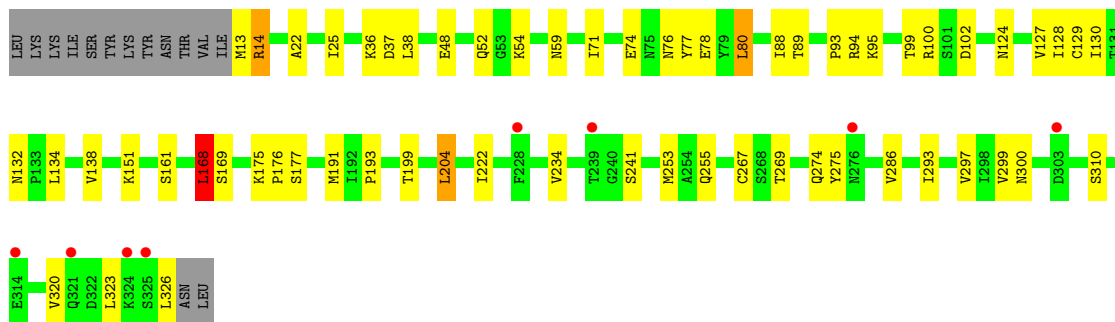
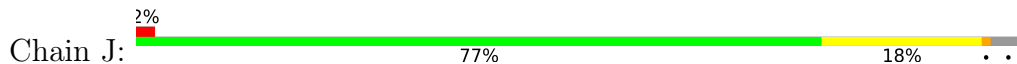




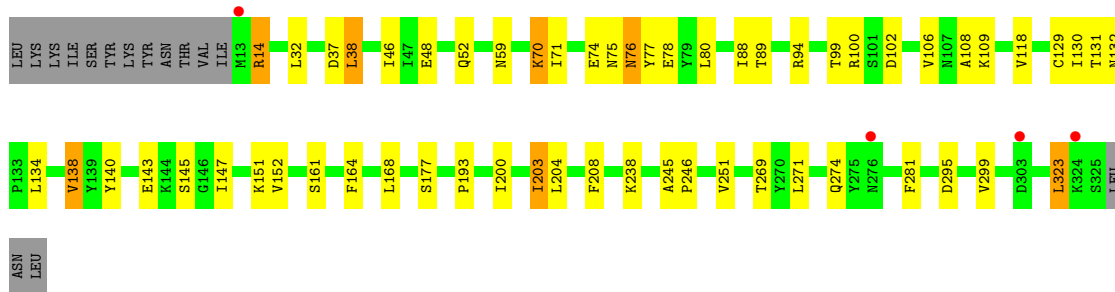
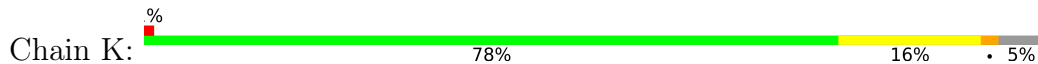
• Molecule 1: Malate dehydrogenase



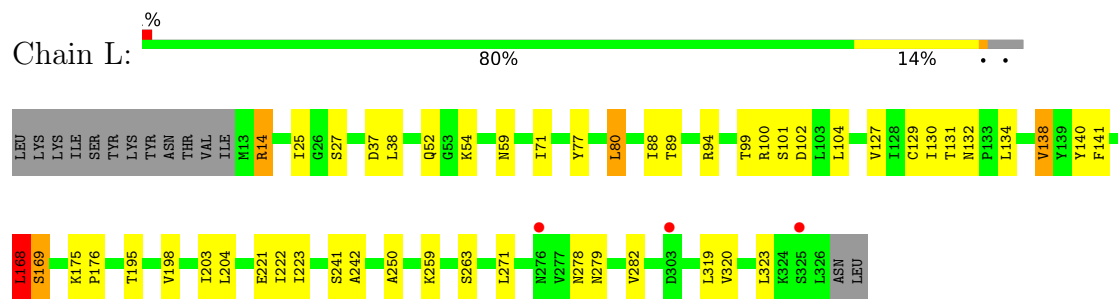
• Molecule 1: Malate dehydrogenase



• Molecule 1: Malate dehydrogenase



• Molecule 1: Malate dehydrogenase



4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α , β , γ	171.89Å 171.89Å 135.61Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.83 – 2.20 19.82 – 2.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.83-2.20) 100.0 (19.82-2.20)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.02 (at 2.21Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.198 , 0.247 0.195 , 0.243	Depositor DCC
R_{free} test set	11391 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å ²)	34.1	Xtrriage
Anisotropy	0.023	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 41.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.008 for -h,-k,l 0.017 for h,-h-k,-l 0.054 for -k,-h,-l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	29965	wwPDB-VP
Average B, all atoms (Å ²)	34.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.91% 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 i

5.1 Standard geometry i

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

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	0.57	0/2365	0.69	3/3200 (0.1%)
1	B	0.57	0/2373	0.69	5/3211 (0.2%)
1	C	0.55	0/2373	0.67	2/3211 (0.1%)
1	D	0.55	0/2373	0.68	2/3211 (0.1%)
1	E	0.57	0/2365	0.73	4/3200 (0.1%)
1	F	0.55	0/2373	0.68	4/3211 (0.1%)
1	G	0.56	0/2365	0.69	3/3200 (0.1%)
1	H	0.56	0/2373	0.69	3/3211 (0.1%)
1	I	0.54	0/2365	0.68	1/3200 (0.0%)
1	J	0.56	0/2373	0.65	3/3211 (0.1%)
1	K	0.57	0/2365	0.68	3/3200 (0.1%)
1	L	0.56	0/2373	0.67	2/3211 (0.1%)
All	All	0.56	0/28436	0.68	35/38477 (0.1%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	100	ARG	NE-CZ-NH2	-8.74	115.93	120.30
1	D	14	ARG	NE-CZ-NH2	-8.55	116.02	120.30
1	B	14	ARG	NE-CZ-NH2	-8.49	116.06	120.30
1	G	14	ARG	NE-CZ-NH2	-8.41	116.10	120.30
1	C	14	ARG	NE-CZ-NH2	-8.38	116.11	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2332	0	2399	52	0
1	B	2340	0	2410	45	0
1	C	2340	0	2410	58	0
1	D	2340	0	2410	39	0
1	E	2332	0	2399	44	0
1	F	2340	0	2410	33	0
1	G	2332	0	2399	44	0
1	H	2340	0	2410	52	0
1	I	2332	0	2399	36	0
1	J	2340	0	2410	46	0
1	K	2332	0	2399	39	0
1	L	2340	0	2410	36	0
2	A	13	0	5	11	0
2	B	13	0	5	4	0
2	C	13	0	5	12	0
2	D	13	0	5	2	0
2	E	13	0	5	7	0
2	F	13	0	5	5	0
2	G	13	0	5	0	0
2	H	13	0	5	3	0
2	I	13	0	5	1	0
2	J	13	0	5	6	0
2	K	13	0	5	1	0
2	L	13	0	5	7	0
3	A	36	0	21	3	0
3	B	36	0	21	5	0
3	C	36	0	21	4	0
3	D	36	0	21	5	0
3	E	36	0	21	4	0
3	F	36	0	21	3	0
3	G	36	0	21	7	0
3	H	36	0	21	6	0
3	I	36	0	21	4	0
3	J	36	0	21	5	0
3	K	36	0	21	3	0
3	L	36	0	21	3	0
4	A	114	0	0	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	135	0	0	11	0
4	C	118	0	0	12	0
4	D	103	0	0	7	0
4	E	101	0	0	5	0
4	F	131	0	0	4	0
4	G	121	0	0	5	0
4	H	114	0	0	11	0
4	I	105	0	0	4	0
4	J	84	0	0	5	0
4	K	110	0	0	9	0
4	L	101	0	0	3	0
All	All	29965	0	29177	517	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 9.

The worst 5 of 517 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)
3:E:3004:APR:HR'1	4:E:3072:HOH:O	1.24	1.29
1:G:52:GLN:HG2	4:G:3100:HOH:O	1.25	1.28
3:A:3001:APR:HR'1	4:A:3078:HOH:O	1.34	1.25
3:D:3003:APR:HR'1	4:D:3081:HOH:O	1.36	1.24
1:A:187:HIS:HE1	2:A:2001:CIT:C4	1.49	1.23

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	311/328 (95%)	301 (97%)	10 (3%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	312/328 (95%)	306 (98%)	6 (2%)	0	100	100
1	C	312/328 (95%)	306 (98%)	6 (2%)	0	100	100
1	D	312/328 (95%)	301 (96%)	11 (4%)	0	100	100
1	E	311/328 (95%)	304 (98%)	7 (2%)	0	100	100
1	F	312/328 (95%)	307 (98%)	5 (2%)	0	100	100
1	G	311/328 (95%)	301 (97%)	10 (3%)	0	100	100
1	H	312/328 (95%)	303 (97%)	9 (3%)	0	100	100
1	I	311/328 (95%)	308 (99%)	3 (1%)	0	100	100
1	J	312/328 (95%)	296 (95%)	16 (5%)	0	100	100
1	K	311/328 (95%)	300 (96%)	11 (4%)	0	100	100
1	L	312/328 (95%)	304 (97%)	8 (3%)	0	100	100
All	All	3739/3936 (95%)	3637 (97%)	102 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	260/275 (94%)	250 (96%)	10 (4%)	33	42
1	B	261/275 (95%)	250 (96%)	11 (4%)	30	38
1	C	261/275 (95%)	245 (94%)	16 (6%)	18	21
1	D	261/275 (95%)	252 (97%)	9 (3%)	37	47
1	E	260/275 (94%)	245 (94%)	15 (6%)	20	23
1	F	261/275 (95%)	247 (95%)	14 (5%)	22	26
1	G	260/275 (94%)	250 (96%)	10 (4%)	33	42
1	H	261/275 (95%)	250 (96%)	11 (4%)	30	38
1	I	260/275 (94%)	244 (94%)	16 (6%)	18	21
1	J	261/275 (95%)	254 (97%)	7 (3%)	44	57

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	K	260/275 (94%)	244 (94%)	16 (6%)	18	21
1	L	261/275 (95%)	250 (96%)	11 (4%)	30	38
All	All	3127/3300 (95%)	2981 (95%)	146 (5%)	26	33

5 of 146 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	J	80	LEU
1	L	221	GLU
1	K	38	LEU
1	K	200	ILE
1	E	70	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 64 such sidechains are listed below:

Mol	Chain	Res	Type
1	K	318	ASN
1	L	75	ASN
1	F	59	ASN
1	E	318	ASN
1	L	76	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

24 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	APR	G	3006	-	34,39,39	0.79	0	40,60,60	1.63	6 (15%)
3	APR	I	3008	-	34,39,39	0.96	0	40,60,60	2.87	9 (22%)
2	CIT	F	2007	-	12,12,12	2.63	2 (16%)	17,17,17	3.04	9 (52%)
3	APR	F	3005	-	34,39,39	0.91	2 (5%)	40,60,60	2.48	9 (22%)
2	CIT	A	2001	-	12,12,12	1.20	1 (8%)	17,17,17	2.23	8 (47%)
3	APR	A	3001	-	34,39,39	0.75	0	40,60,60	1.71	8 (20%)
3	APR	J	3009	-	34,39,39	0.72	0	40,60,60	1.44	4 (10%)
2	CIT	B	2002	-	12,12,12	2.45	3 (25%)	17,17,17	3.16	10 (58%)
2	CIT	H	2005	-	12,12,12	1.78	2 (16%)	17,17,17	2.50	8 (47%)
2	CIT	D	2004	-	12,12,12	1.12	0	17,17,17	2.03	6 (35%)
3	APR	C	3002	-	34,39,39	1.19	1 (2%)	40,60,60	2.67	9 (22%)
2	CIT	C	2003	-	12,12,12	1.34	2 (16%)	17,17,17	2.59	8 (47%)
3	APR	K	3010	-	34,39,39	0.84	0	40,60,60	1.68	7 (17%)
3	APR	D	3003	-	34,39,39	0.76	0	40,60,60	1.66	8 (20%)
3	APR	L	3011	-	34,39,39	0.91	0	40,60,60	2.41	7 (17%)
2	CIT	J	2012	-	12,12,12	1.52	1 (8%)	17,17,17	1.84	6 (35%)
2	CIT	I	2011	-	12,12,12	2.25	2 (16%)	17,17,17	2.70	5 (29%)
2	CIT	E	2008	-	12,12,12	2.11	5 (41%)	17,17,17	2.86	6 (35%)
2	CIT	G	2006	-	12,12,12	2.45	2 (16%)	17,17,17	3.38	7 (41%)
2	CIT	L	2010	-	12,12,12	1.84	3 (25%)	17,17,17	2.29	9 (52%)
3	APR	H	3007	-	34,39,39	0.76	0	40,60,60	2.19	7 (17%)
3	APR	E	3004	-	34,39,39	0.82	0	40,60,60	1.76	6 (15%)
3	APR	B	3012	-	34,39,39	0.78	1 (2%)	40,60,60	1.79	7 (17%)
2	CIT	K	2013	-	12,12,12	2.07	2 (16%)	17,17,17	2.54	5 (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
3	APR	G	3006	-	-	1/18/54/54	0/4/4/4
3	APR	I	3008	-	-	8/18/54/54	0/4/4/4
2	CIT	F	2007	-	-	10/16/16/16	-
3	APR	F	3005	-	-	7/18/54/54	0/4/4/4
2	CIT	A	2001	-	-	11/16/16/16	-
3	APR	A	3001	-	-	5/18/54/54	0/4/4/4
3	APR	J	3009	-	-	5/18/54/54	0/4/4/4
2	CIT	B	2002	-	-	0/16/16/16	-
2	CIT	H	2005	-	-	3/16/16/16	-
2	CIT	D	2004	-	-	4/16/16/16	-
3	APR	C	3002	-	-	4/18/54/54	0/4/4/4
2	CIT	C	2003	-	-	10/16/16/16	-
3	APR	K	3010	-	-	6/18/54/54	0/4/4/4
3	APR	D	3003	-	-	5/18/54/54	0/4/4/4
3	APR	L	3011	-	-	6/18/54/54	0/4/4/4
2	CIT	J	2012	-	-	4/16/16/16	-
2	CIT	I	2011	-	-	3/16/16/16	-
2	CIT	E	2008	-	-	4/16/16/16	-
2	CIT	G	2006	-	-	3/16/16/16	-
2	CIT	L	2010	-	-	6/16/16/16	-
3	APR	H	3007	-	-	6/18/54/54	0/4/4/4
3	APR	E	3004	-	-	5/18/54/54	0/4/4/4
3	APR	B	3012	-	-	7/18/54/54	0/4/4/4
2	CIT	K	2013	-	-	3/16/16/16	-

The worst 5 of 29 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	2007	CIT	C3-C6	-7.88	1.45	1.53
2	G	2006	CIT	C3-C6	-7.55	1.45	1.53
2	B	2002	CIT	C3-C6	-6.90	1.46	1.53
2	I	2011	CIT	C3-C6	-6.22	1.46	1.53
2	K	2013	CIT	C3-C6	-5.78	1.47	1.53

The worst 5 of 174 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	3008	APR	C1D-C2D-C3D	-12.48	86.68	102.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	3005	APR	C1D-C2D-C3D	-10.33	89.37	102.30
3	L	3011	APR	C1D-C2D-C3D	-10.00	89.78	102.30
2	F	2007	CIT	C3-C2-C1	-9.16	91.62	113.81
2	G	2006	CIT	O7-C3-C6	-9.06	96.14	108.86

There are no chirality outliers.

5 of 126 torsion outliers are listed below:

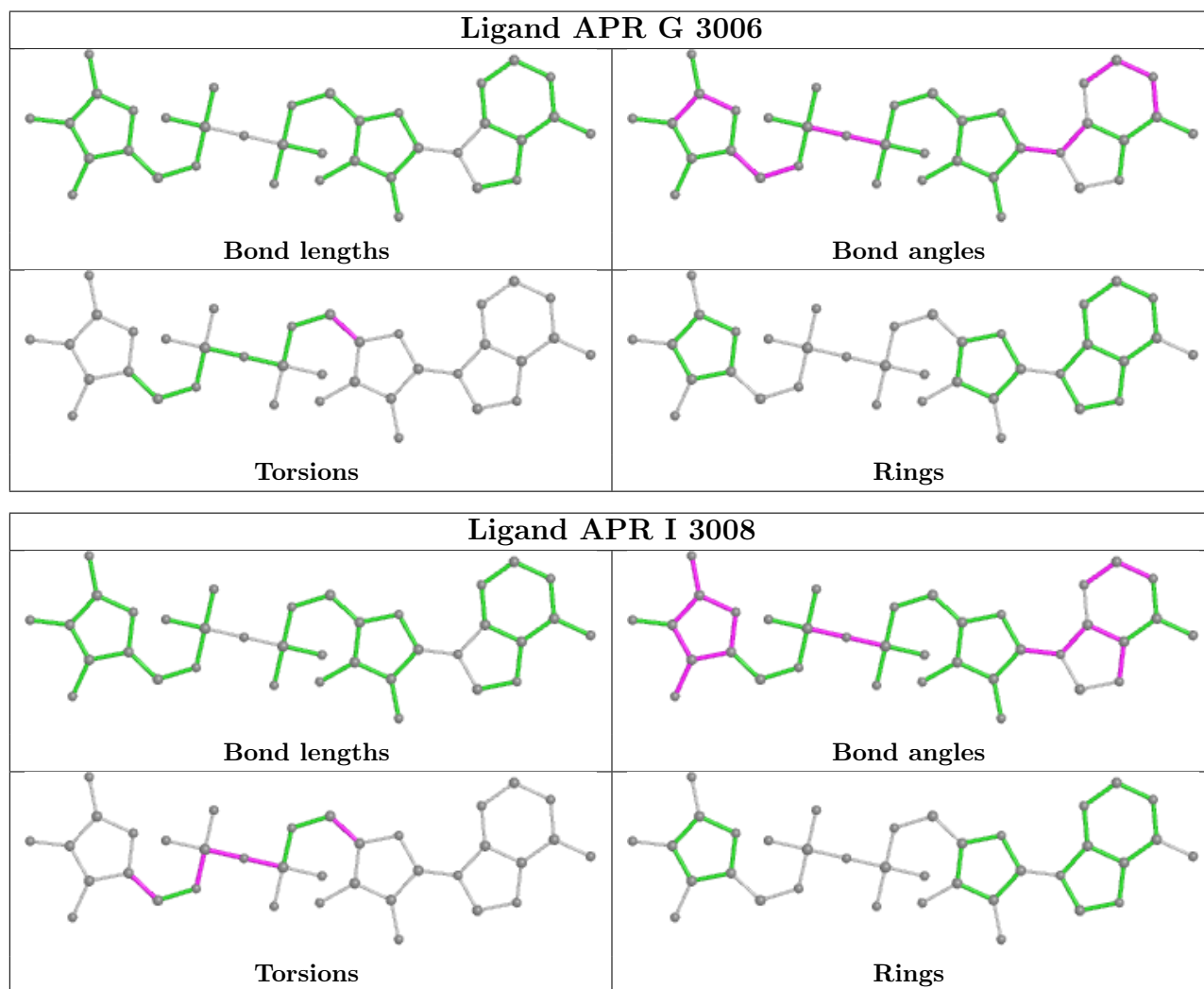
Mol	Chain	Res	Type	Atoms
2	A	2001	CIT	C2-C3-C4-C5
2	A	2001	CIT	O7-C3-C4-C5
2	A	2001	CIT	O7-C3-C6-O5
2	A	2001	CIT	O7-C3-C6-O6
2	A	2001	CIT	C4-C3-C6-O5

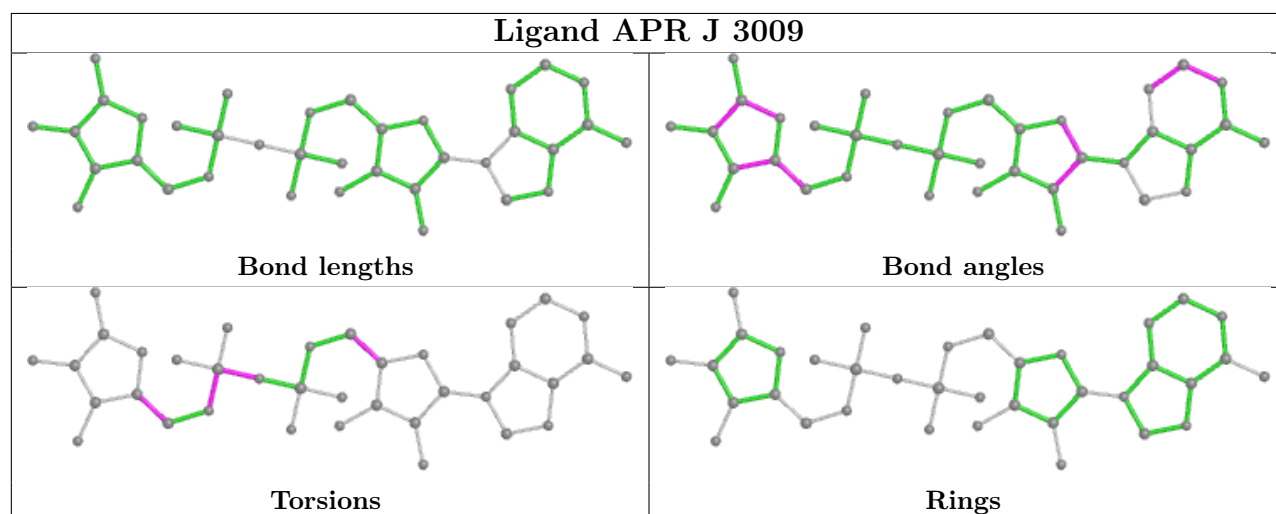
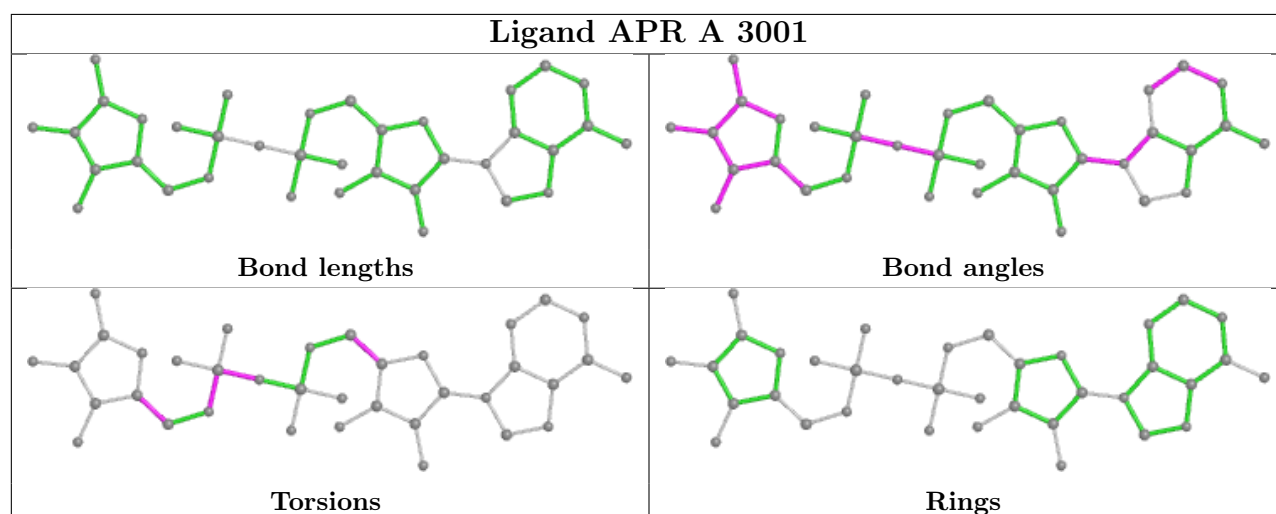
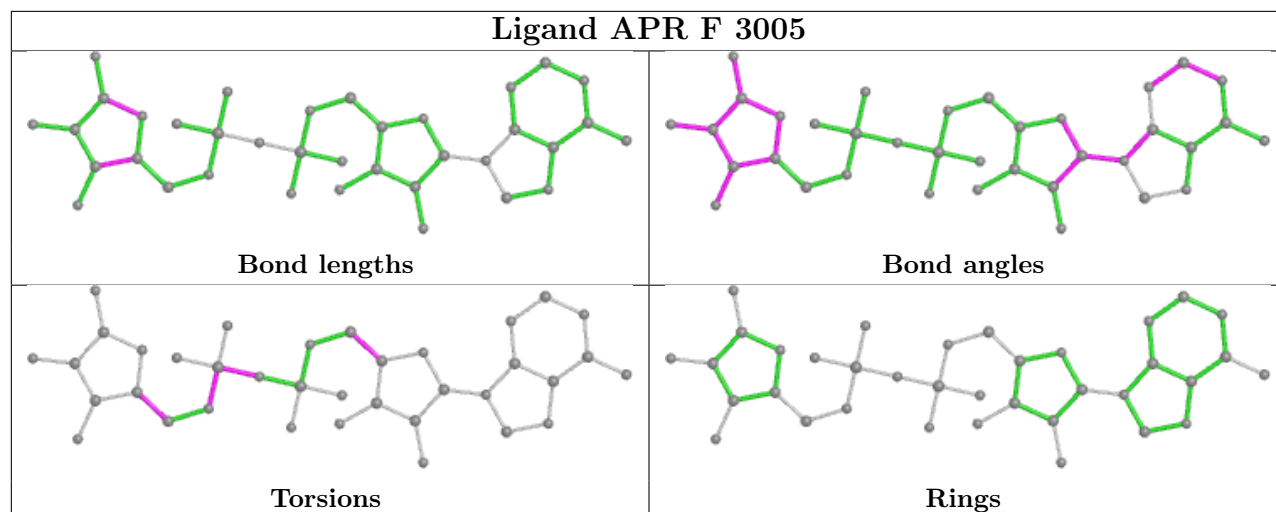
There are no ring outliers.

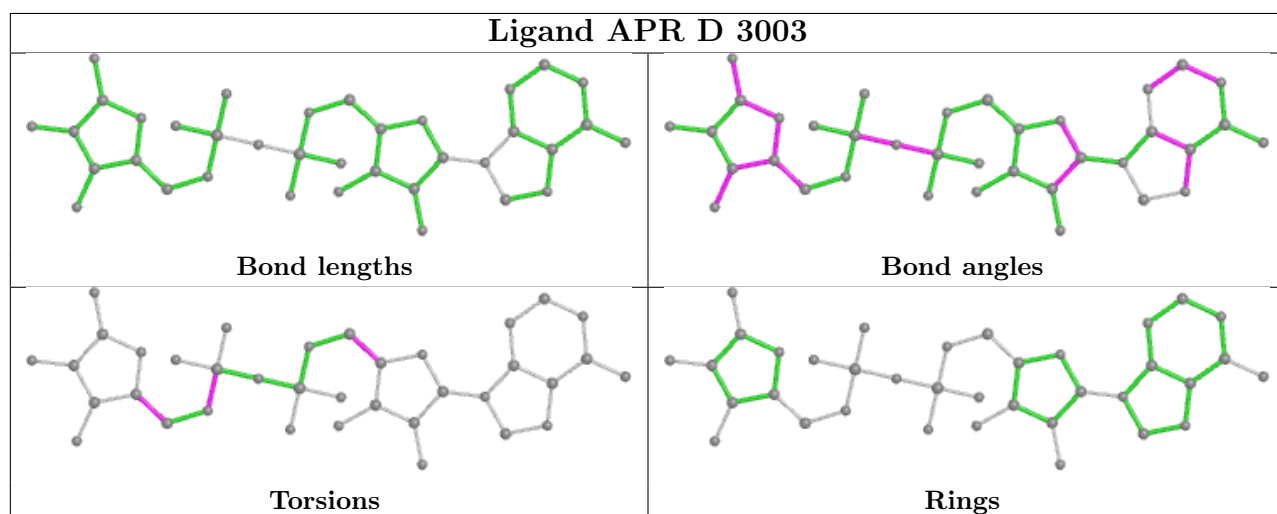
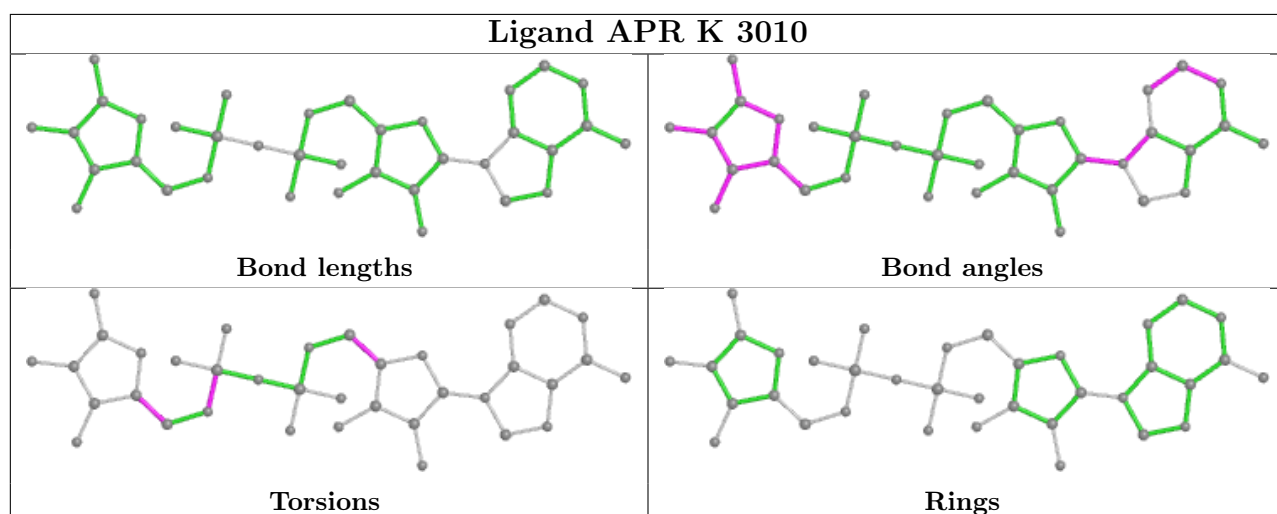
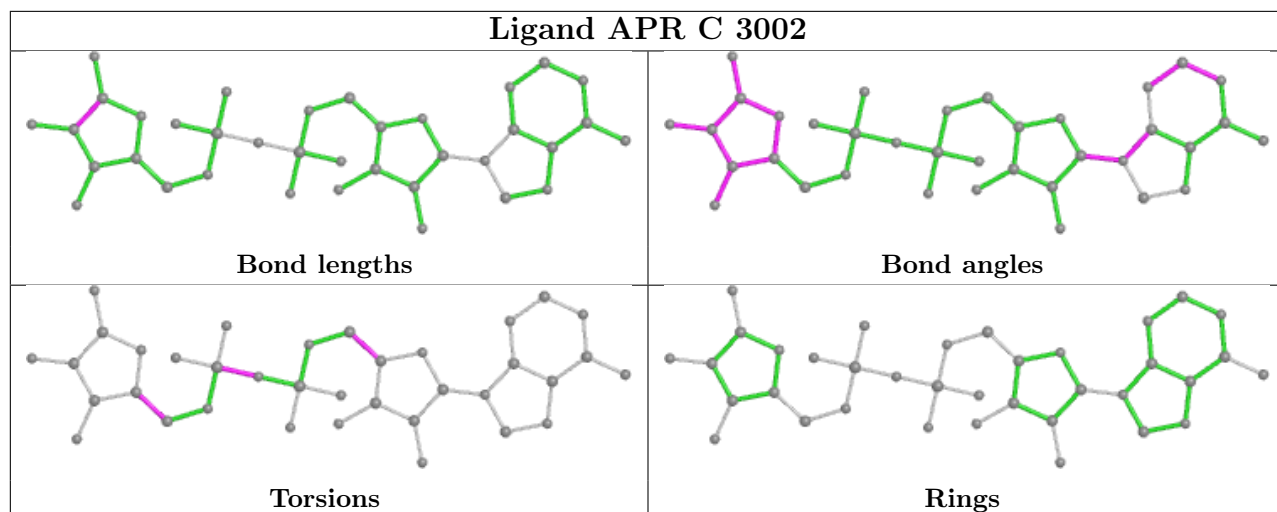
23 monomers are involved in 111 short contacts:

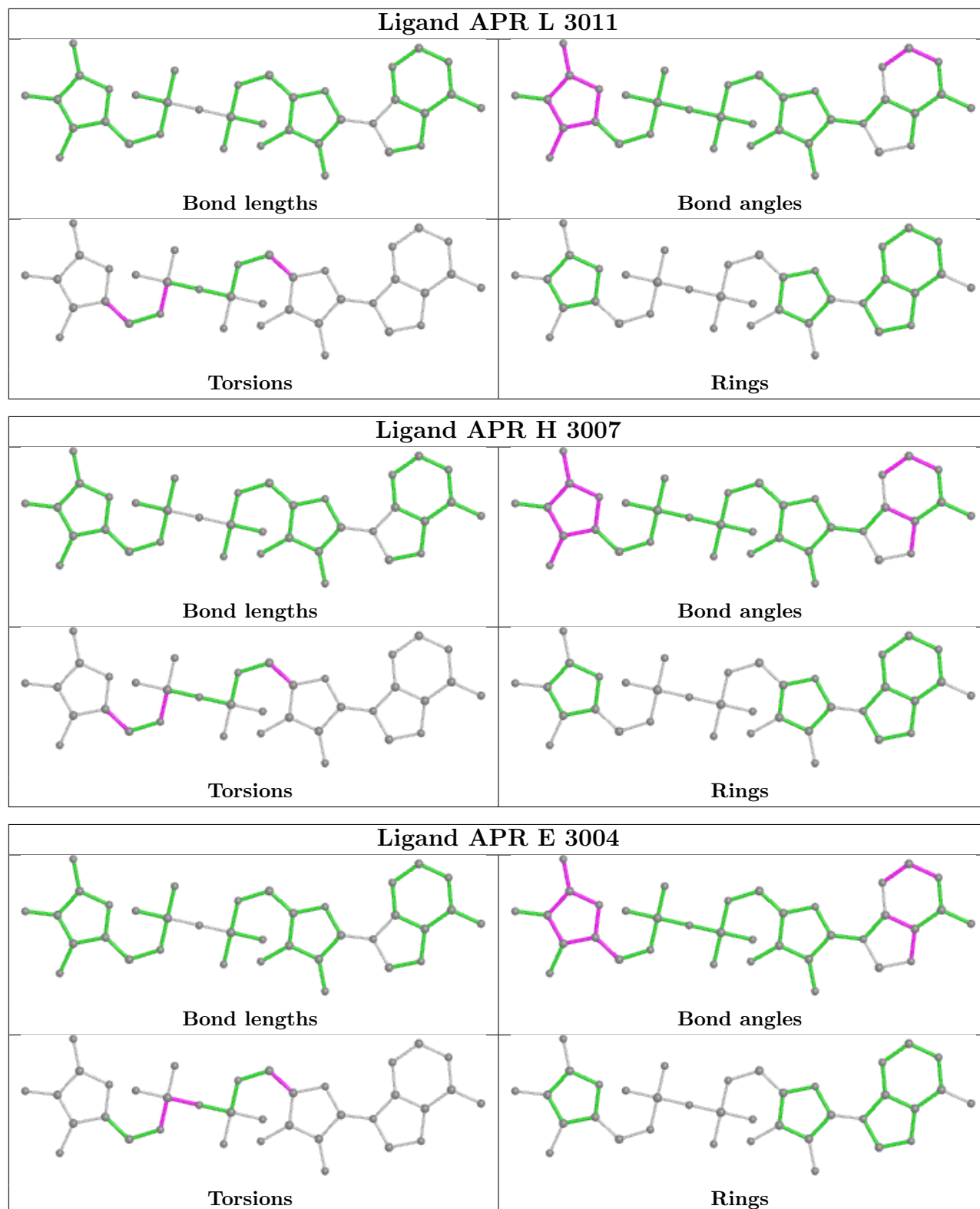
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	G	3006	APR	7	0
3	I	3008	APR	4	0
2	F	2007	CIT	5	0
3	F	3005	APR	3	0
2	A	2001	CIT	11	0
3	A	3001	APR	3	0
3	J	3009	APR	5	0
2	B	2002	CIT	4	0
2	H	2005	CIT	3	0
2	D	2004	CIT	2	0
3	C	3002	APR	4	0
2	C	2003	CIT	12	0
3	K	3010	APR	3	0
3	D	3003	APR	5	0
3	L	3011	APR	3	0
2	J	2012	CIT	6	0
2	I	2011	CIT	1	0
2	E	2008	CIT	7	0
2	L	2010	CIT	7	0
3	H	3007	APR	6	0
3	E	3004	APR	4	0
3	B	3012	APR	5	0
2	K	2013	CIT	1	0

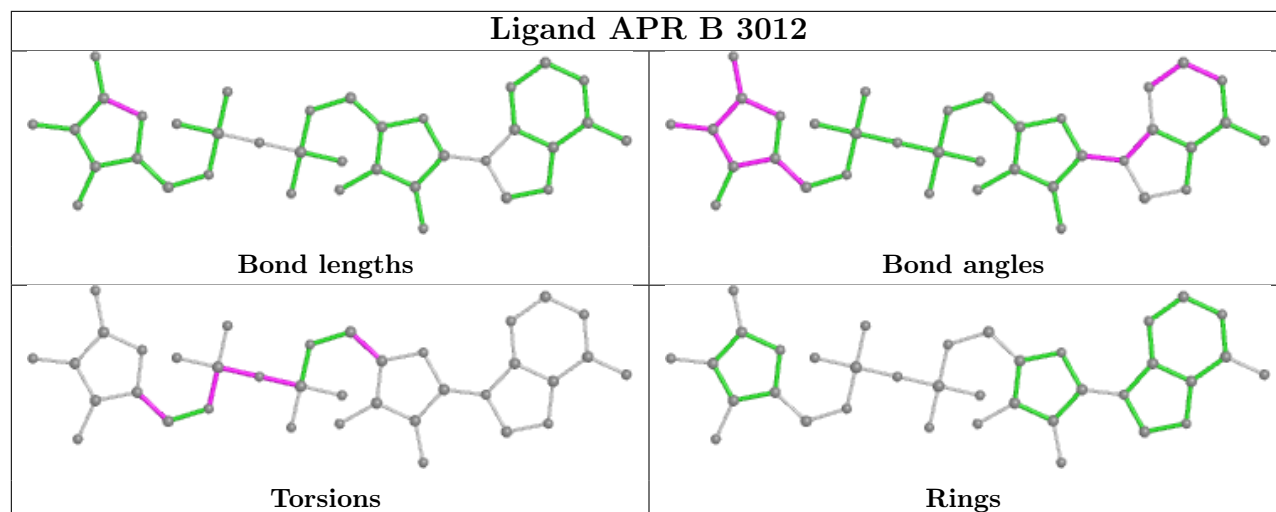
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 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.











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	313/328 (95%)	-0.30	2 (0%) 89 88	22, 32, 46, 60	0
1	B	314/328 (95%)	-0.30	1 (0%) 94 93	20, 31, 46, 58	0
1	C	314/328 (95%)	-0.32	1 (0%) 94 93	21, 32, 46, 58	0
1	D	314/328 (95%)	-0.30	3 (0%) 82 81	22, 32, 47, 59	0
1	E	313/328 (95%)	-0.31	4 (1%) 77 75	23, 32, 46, 57	0
1	F	314/328 (95%)	-0.31	0 100 100	22, 32, 46, 60	0
1	G	313/328 (95%)	-0.37	3 (0%) 82 81	22, 32, 46, 58	0
1	H	314/328 (95%)	-0.30	4 (1%) 77 75	23, 33, 47, 57	0
1	I	313/328 (95%)	-0.29	3 (0%) 82 81	22, 33, 46, 57	0
1	J	314/328 (95%)	-0.25	8 (2%) 57 55	24, 33, 48, 63	0
1	K	313/328 (95%)	-0.28	4 (1%) 77 75	23, 33, 46, 60	0
1	L	314/328 (95%)	-0.31	3 (0%) 82 81	23, 33, 47, 60	0
All	All	3763/3936 (95%)	-0.30	36 (0%) 82 81	20, 32, 47, 63	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	J	321	GLN	3.3
1	K	13	MET	3.3
1	E	303	ASP	3.2
1	H	276	ASN	3.1
1	J	239	THR	3.1

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

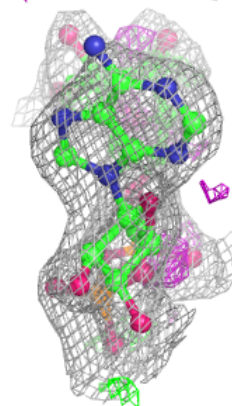
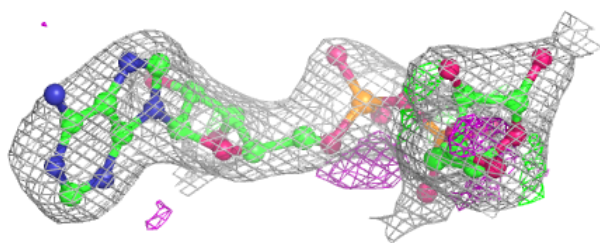
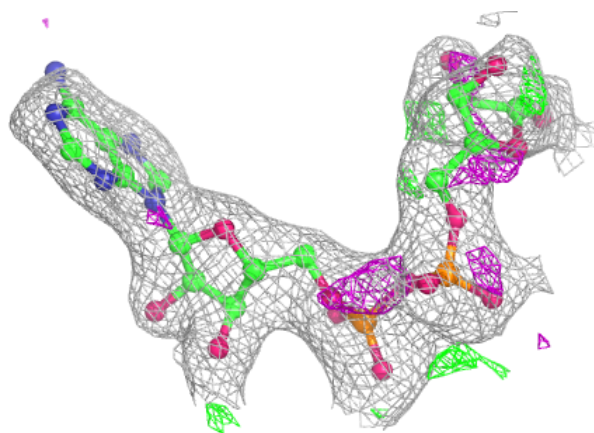
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	CIT	L	2010	13/13	0.76	0.27	34,41,48,49	0
2	CIT	J	2012	13/13	0.81	0.24	49,52,56,58	0
2	CIT	C	2003	13/13	0.89	0.15	26,31,34,34	0
2	CIT	F	2007	13/13	0.89	0.16	26,30,33,33	0
2	CIT	H	2005	13/13	0.89	0.13	30,33,36,38	0
2	CIT	A	2001	13/13	0.89	0.17	27,33,35,37	0
2	CIT	B	2002	13/13	0.89	0.15	23,27,31,32	0
3	APR	H	3007	36/36	0.89	0.17	42,47,50,51	0
2	CIT	E	2008	13/13	0.90	0.16	27,34,38,40	0
2	CIT	I	2011	13/13	0.91	0.13	27,29,32,33	0
3	APR	J	3009	36/36	0.91	0.17	42,46,51,55	0
2	CIT	K	2013	13/13	0.92	0.12	31,33,36,37	0
2	CIT	G	2006	13/13	0.92	0.12	25,29,34,34	0
3	APR	G	3006	36/36	0.93	0.12	34,36,47,53	0
3	APR	D	3003	36/36	0.93	0.15	36,42,46,49	0
3	APR	E	3004	36/36	0.93	0.15	36,41,43,48	0
3	APR	L	3011	36/36	0.93	0.13	32,40,45,48	0
3	APR	C	3002	36/36	0.94	0.11	28,32,39,40	0
3	APR	F	3005	36/36	0.94	0.11	30,35,44,46	0
3	APR	K	3010	36/36	0.94	0.12	35,40,46,52	0
3	APR	B	3012	36/36	0.94	0.13	30,40,44,46	0
3	APR	A	3001	36/36	0.95	0.11	33,36,45,51	0
2	CIT	D	2004	13/13	0.95	0.13	28,31,32,32	0
3	APR	I	3008	36/36	0.95	0.09	30,32,38,40	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.

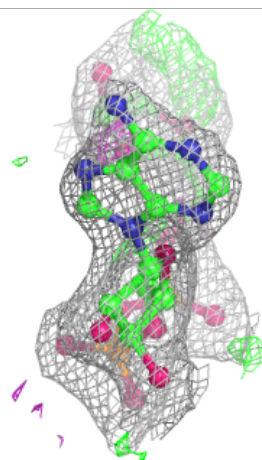
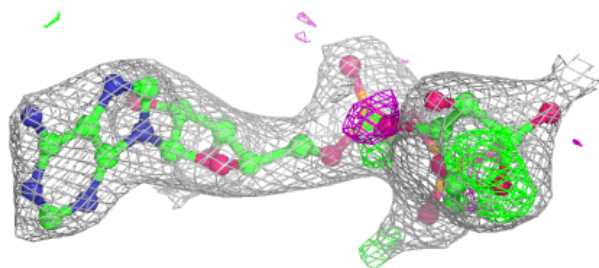
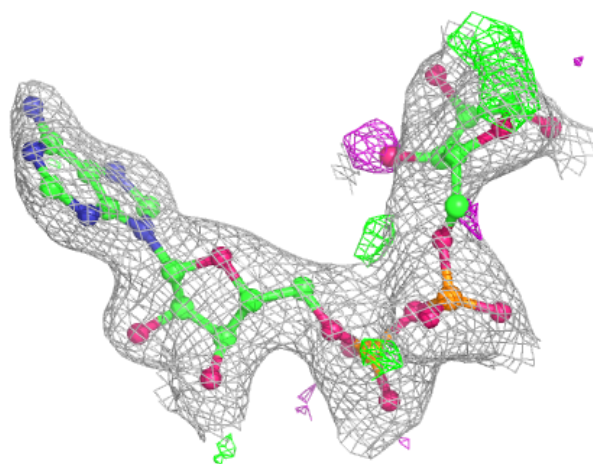
Electron density around APR H 3007:

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



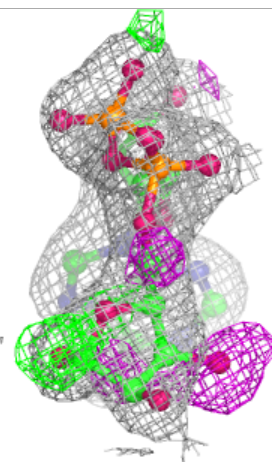
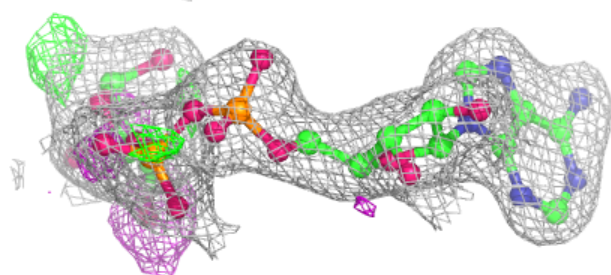
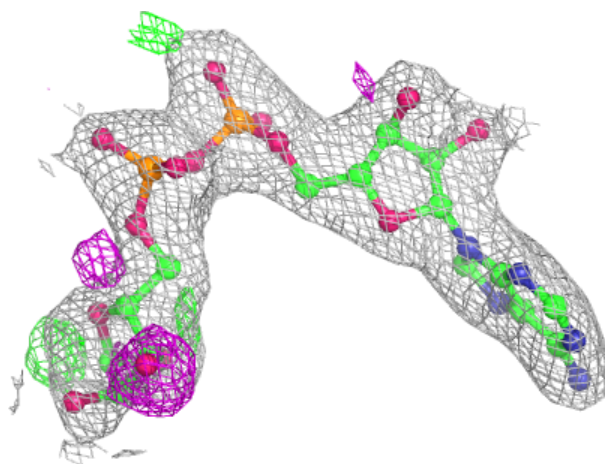
Electron density around APR J 3009:

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



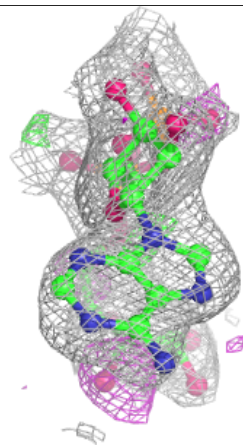
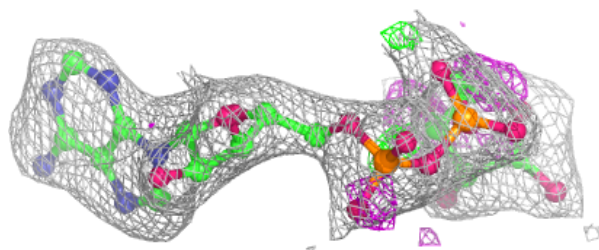
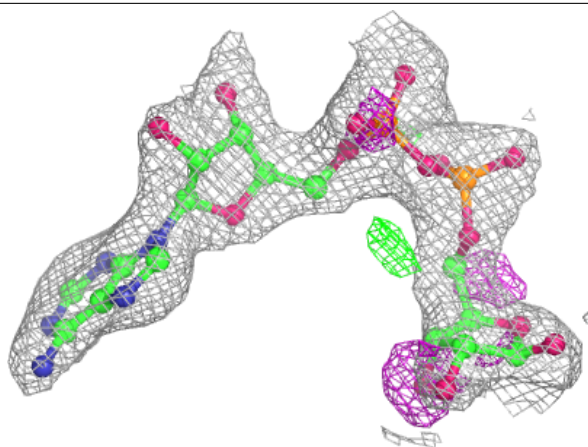
Electron density around APR G 3006:

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



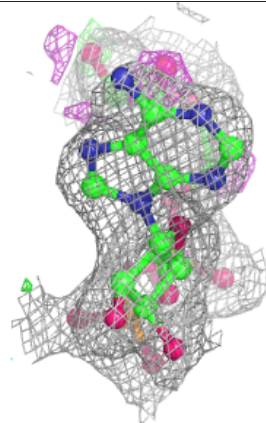
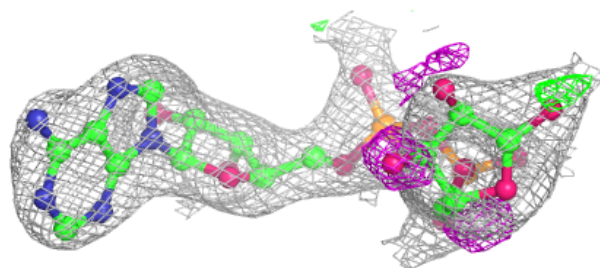
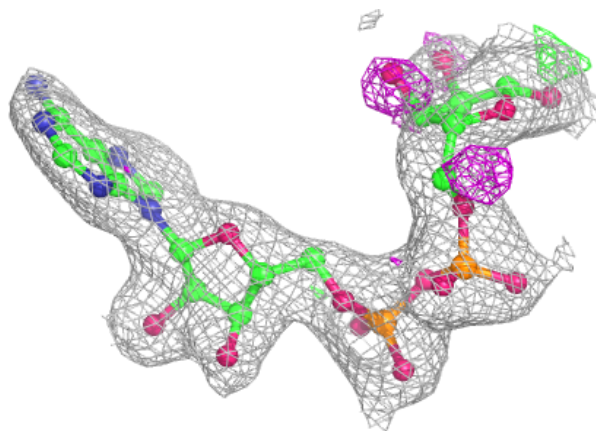
Electron density around APR D 3003:

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

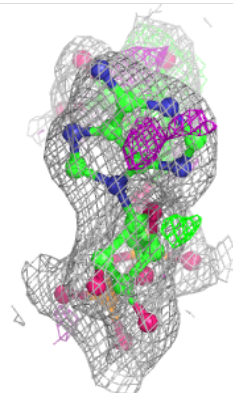
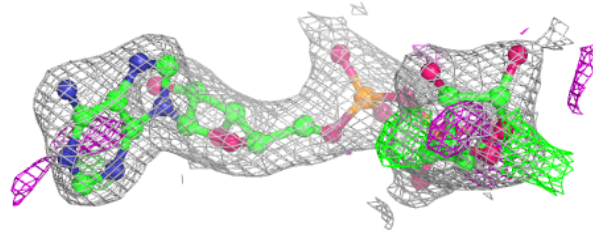
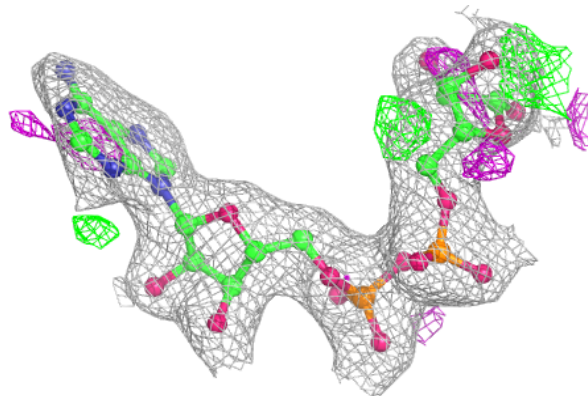


Electron density around APR E 3004:

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

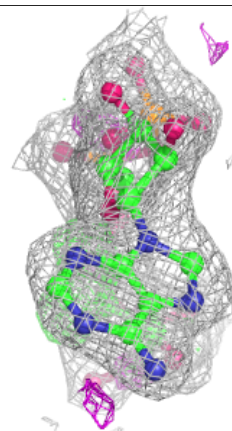
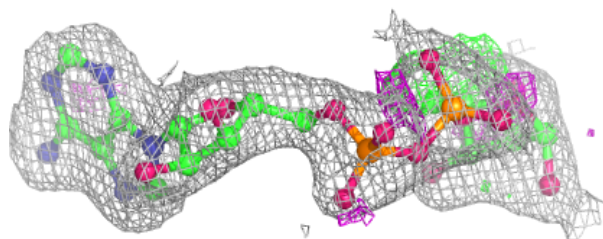
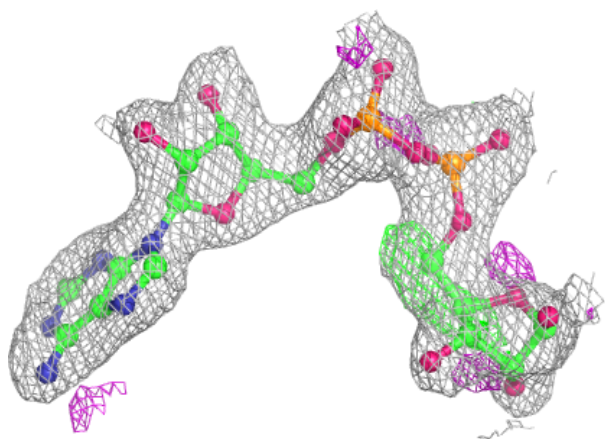
**Electron density around APR L 3011:**

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

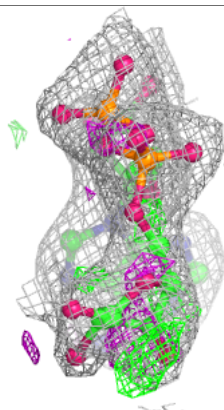
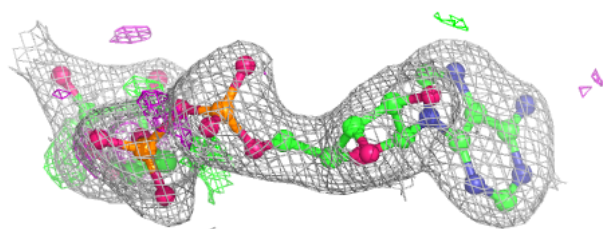
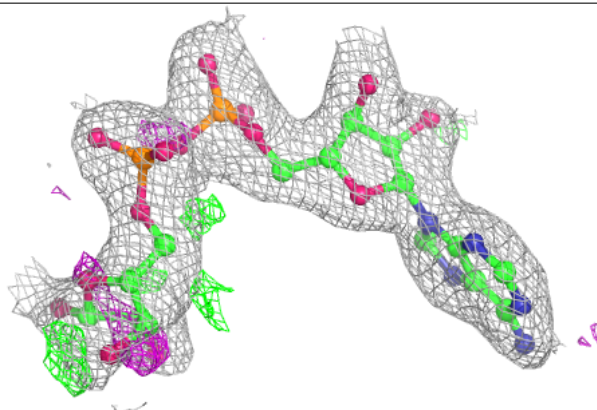


Electron density around APR C 3002:

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

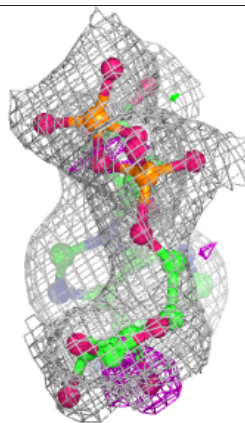
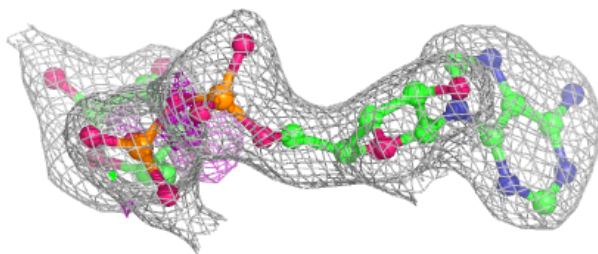
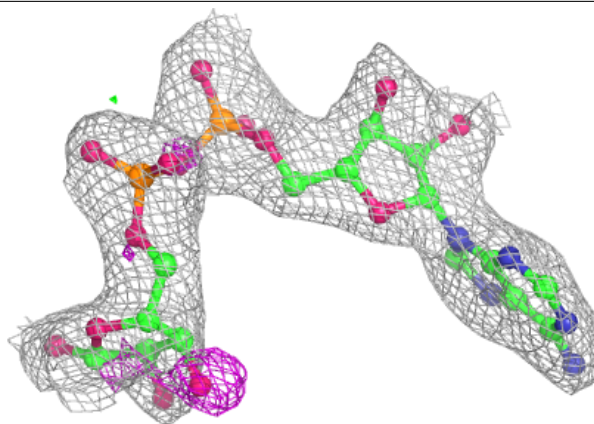
**Electron density around APR F 3005:**

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



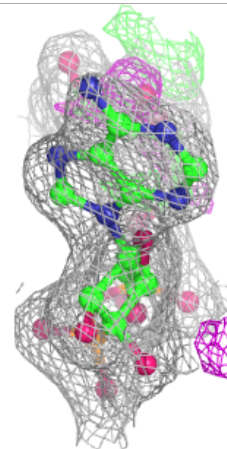
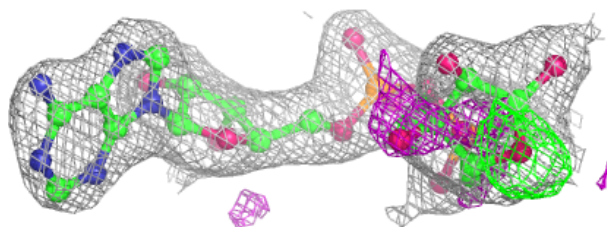
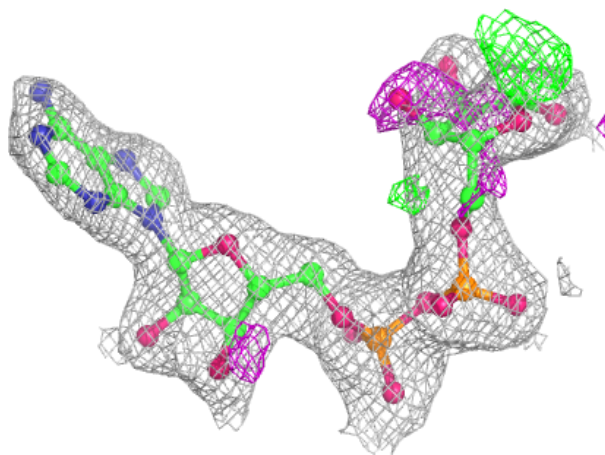
Electron density around APR K 3010:

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



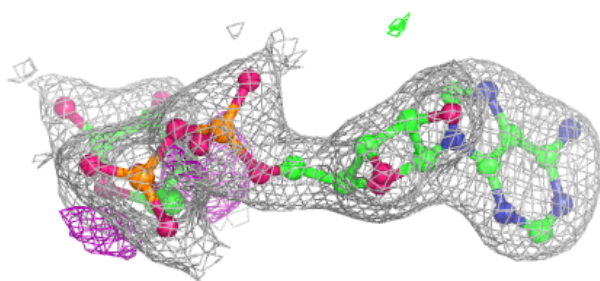
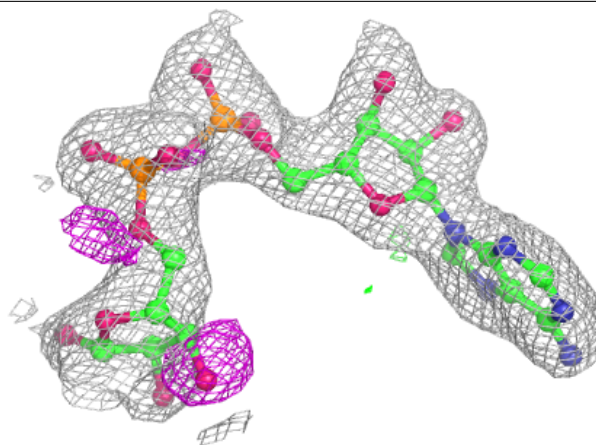
Electron density around APR B 3012:

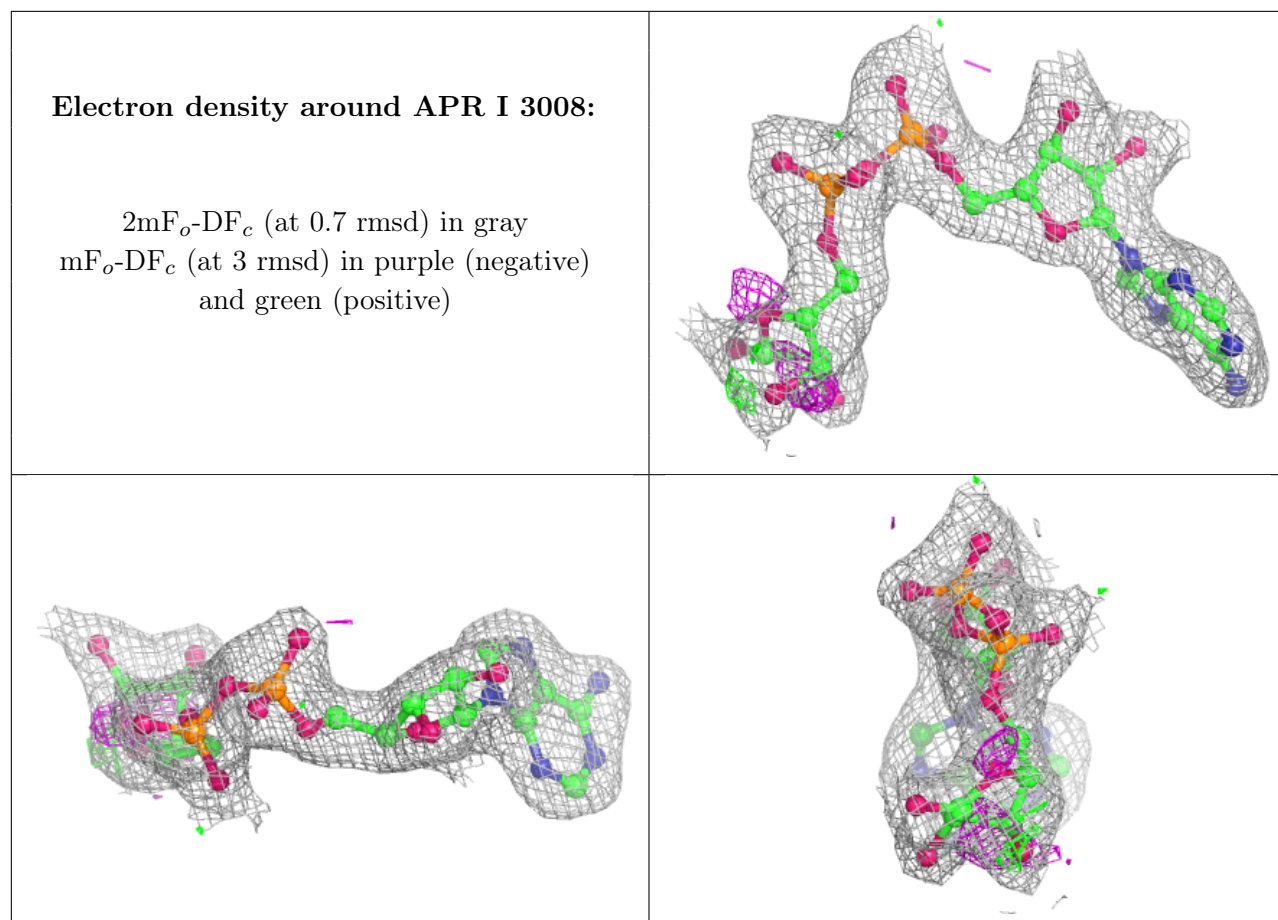
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around APR A 3001:

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





6.5 Other polymers [i](#)

There are no such residues in this entry.