



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 2, 2020 – 08:55 am BST

PDB ID : 5CX1  
Title : Nitrogenase molybdenum-iron protein beta-K400E mutant  
Authors : Owens, C.P.; Luca, M.A.; Tezcan, F.A.  
Deposited on : 2015-07-28  
Resolution : 1.75 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

<https://www.wwpdb.org/validation/2017/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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

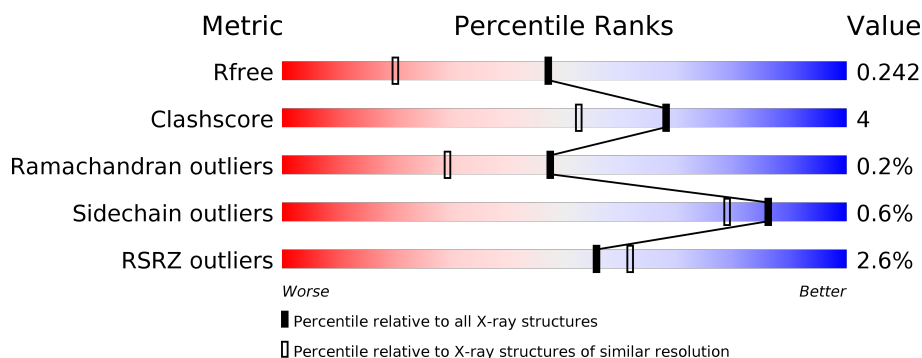
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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	3764 (1.76-1.72)
Clashscore	141614	3923 (1.76-1.72)
Ramachandran outliers	138981	3878 (1.76-1.72)
Sidechain outliers	138945	3878 (1.76-1.72)
RSRZ outliers	127900	3705 (1.76-1.72)

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 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	480	<div> <div>5%</div> <div> <div></div> <div>87%</div> <div>12%</div> <div></div> </div> </div>
1	C	480	<div> <div>4%</div> <div> <div></div> <div>89%</div> <div>10%</div> <div></div> </div> </div>
1	E	480	<div> <div>6%</div> <div> <div></div> <div>90%</div> <div>10%</div> <div></div> </div> </div>
1	G	480	<div> <div>3%</div> <div> <div></div> <div>89%</div> <div>10%</div> <div></div> </div> </div>
1	I	480	<div> <div>3%</div> <div> <div></div> <div>89%</div> <div>10%</div> <div></div> </div> </div>
1	K	480	<div> <div>4%</div> <div> <div></div> <div>89%</div> <div>10%</div> <div></div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	M	480	<div><div>%</div><div><div></div><div>88%</div><div>11%</div><div>..</div></div></div>
1	O	480	<div><div>3%</div><div><div></div><div>89%</div><div>10%</div><div>.</div></div></div>
2	B	523	<div><div>2%</div><div><div></div><div>91%</div><div>8%</div><div></div></div></div>
2	D	523	<div><div>%</div><div><div></div><div>91%</div><div>8%</div><div></div></div></div>
2	F	523	<div><div>%</div><div><div></div><div>94%</div><div>6%</div><div></div></div></div>
2	H	523	<div><div>%</div><div><div></div><div>89%</div><div>10%</div><div>.</div></div></div>
2	J	523	<div><div>2%</div><div><div></div><div>93%</div><div>6%</div><div></div></div></div>
2	L	523	<div><div>3%</div><div><div></div><div>92%</div><div>8%</div><div></div></div></div>
2	N	523	<div><div>%</div><div><div></div><div>95%</div><div>5%</div><div></div></div></div>
2	P	523	<div><div>%</div><div><div></div><div>94%</div><div>6%</div><div></div></div></div>

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 74588 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 Nitrogenase molybdenum-iron protein alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	477	Total	C	N	O	S	0	0	0
			3774	2399	642	708	25			
1	C	477	Total	C	N	O	S	0	0	0
			3782	2405	644	708	25			
1	E	477	Total	C	N	O	S	0	0	0
			3776	2402	644	705	25			
1	G	477	Total	C	N	O	S	0	0	0
			3786	2407	645	709	25			
1	I	477	Total	C	N	O	S	0	0	0
			3786	2407	645	709	25			
1	K	477	Total	C	N	O	S	0	0	0
			3782	2406	645	706	25			
1	M	477	Total	C	N	O	S	0	0	0
			3790	2410	646	709	25			
1	O	477	Total	C	N	O	S	0	0	0
			3778	2401	643	709	25			

- Molecule 2 is a protein called Nitrogenase molybdenum-iron protein beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	522	Total	C	N	O	S	0	0	0
			4170	2662	703	777	28			
2	D	522	Total	C	N	O	S	0	0	0
			4174	2665	704	777	28			
2	F	522	Total	C	N	O	S	0	0	0
			4174	2665	704	777	28			
2	H	522	Total	C	N	O	S	0	0	0
			4174	2665	704	777	28			
2	J	522	Total	C	N	O	S	0	0	0
			4172	2664	704	776	28			
2	L	522	Total	C	N	O	S	0	0	0
			4170	2662	703	777	28			

*Continued on next page...*

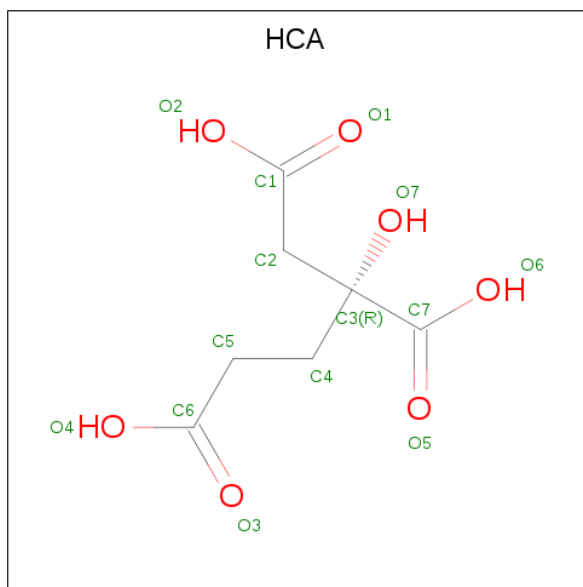
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	N	522	Total	C	N	O	S	0	0	0
			4174	2665	704	777	28			
2	P	522	Total	C	N	O	S	0	0	0
			4174	2665	704	777	28			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	400	GLU	LYS	engineered mutation	UNP P07329
D	400	GLU	LYS	engineered mutation	UNP P07329
F	400	GLU	LYS	engineered mutation	UNP P07329
H	400	GLU	LYS	engineered mutation	UNP P07329
J	400	GLU	LYS	engineered mutation	UNP P07329
L	400	GLU	LYS	engineered mutation	UNP P07329
N	400	GLU	LYS	engineered mutation	UNP P07329
P	400	GLU	LYS	engineered mutation	UNP P07329

- Molecule 3 is 3-HYDROXY-3-CARBOXY-ADIPIC ACID (three-letter code: HCA) (formula:  $C_7H_{10}O_7$ ).



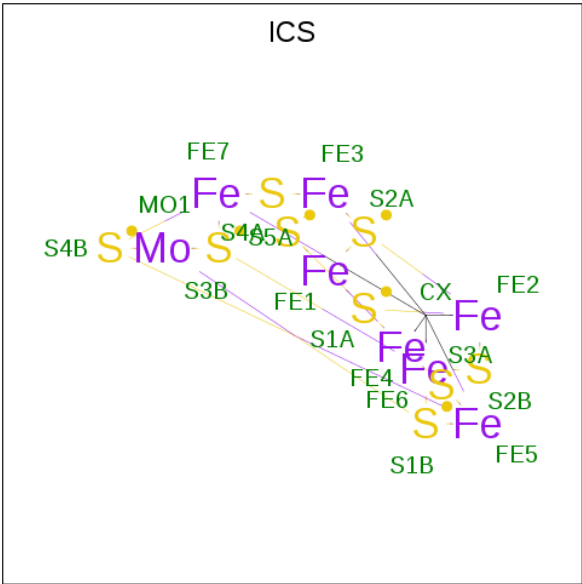
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			14	7	7		
3	C	1	Total	C	O	0	0
			14	7	7		
3	E	1	Total	C	O	0	0
			14	7	7		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	G	1	Total	C	O	0	0
			14	7	7		
3	I	1	Total	C	O	0	0
			14	7	7		
3	K	1	Total	C	O	0	0
			14	7	7		
3	M	1	Total	C	O	0	0
			14	7	7		
3	O	1	Total	C	O	0	0
			14	7	7		

- Molecule 4 is iron-sulfur-molybdenum cluster with interstitial carbon (three-letter code: ICS) (formula: CFe<sub>7</sub>MoS<sub>9</sub>).



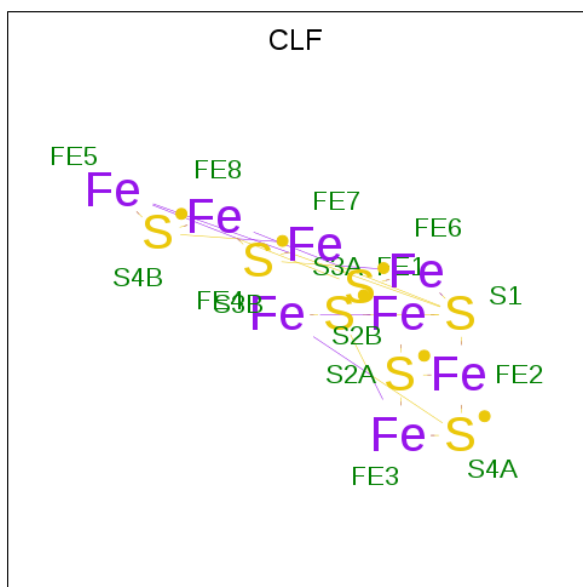
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total 18	C 1	Fe 7	Mo 1	S 9	0	0
4	C	1	Total 18	C 1	Fe 7	Mo 1	S 9	0	0
4	E	1	Total 18	C 1	Fe 7	Mo 1	S 9	0	0
4	G	1	Total 18	C 1	Fe 7	Mo 1	S 9	0	0
4	I	1	Total 18	C 1	Fe 7	Mo 1	S 9	0	0
4	K	1	Total 18	C 1	Fe 7	Mo 1	S 9	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	M	1	Total	C	Fe	Mo	S	
			18	1	7	1	9	
4	O	1	Total	C	Fe	Mo	S	
			18	1	7	1	9	

- Molecule 5 is FE(8)-S(7) CLUSTER (three-letter code: CLF) (formula:  $\text{Fe}_8\text{S}_7$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	Fe	S		
			15	8	7		
5	C	1	Total	Fe	S		
			15	8	7		
5	E	1	Total	Fe	S		
			15	8	7		
5	G	1	Total	Fe	S		
			15	8	7		
5	I	1	Total	Fe	S		
			15	8	7		
5	K	1	Total	Fe	S		
			15	8	7		
5	M	1	Total	Fe	S		
			15	8	7		
5	O	1	Total	Fe	S		
			15	8	7		

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	P	1	Total Ca 1 1	0	0
6	J	1	Total Ca 1 1	0	0
6	D	1	Total Ca 1 1	0	0
6	H	1	Total Ca 1 1	0	0
6	B	1	Total Ca 1 1	0	0
6	N	1	Total Ca 1 1	0	0
6	L	1	Total Ca 1 1	0	0
6	F	1	Total Ca 1 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	542	Total O 542 542	0	0
7	B	753	Total O 753 753	0	0
7	C	582	Total O 582 582	0	0
7	D	752	Total O 752 752	0	0
7	E	507	Total O 507 507	0	0
7	F	750	Total O 750 750	0	0
7	G	615	Total O 615 615	0	0
7	H	749	Total O 749 749	0	0
7	I	557	Total O 557 557	0	0
7	J	761	Total O 761 761	0	0
7	K	587	Total O 587 587	0	0
7	L	709	Total O 709 709	0	0

*Continued on next page...*



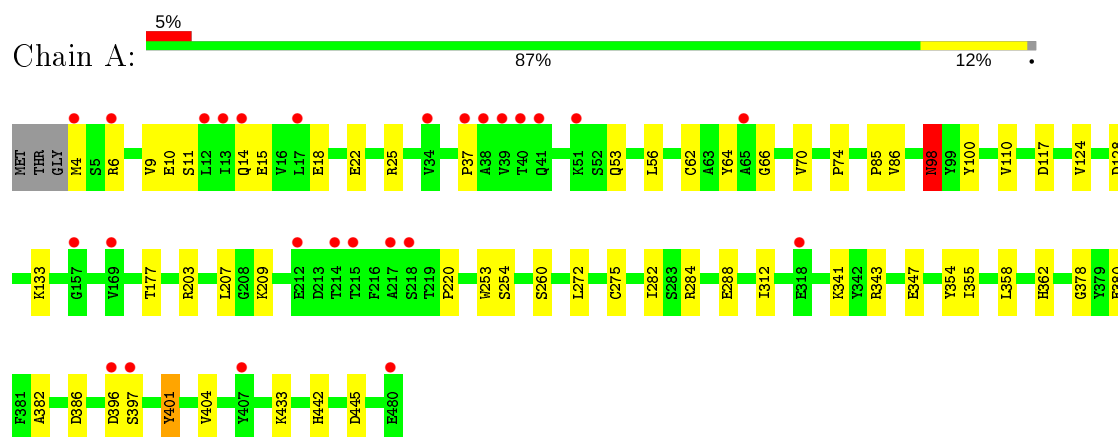
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	M	602	Total 602	O 602	0	0
7	N	797	Total 797	O 797	0	0
7	O	548	Total 548	O 548	0	0
7	P	757	Total 757	O 757	0	0

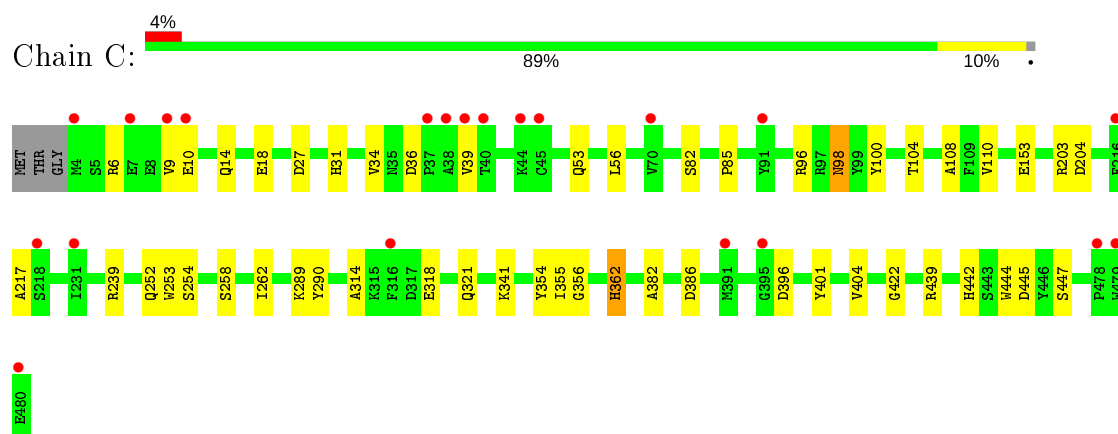
### 3 Residue-property plots [i](#)

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.

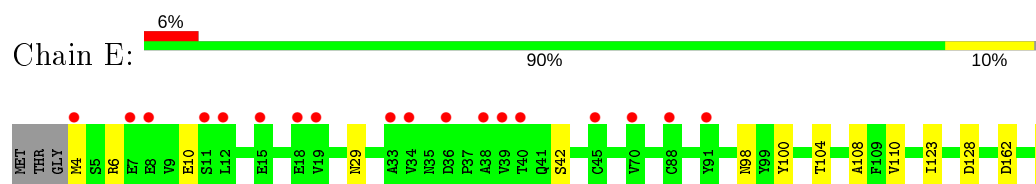
- Molecule 1: Nitrogenase molybdenum-iron protein alpha chain

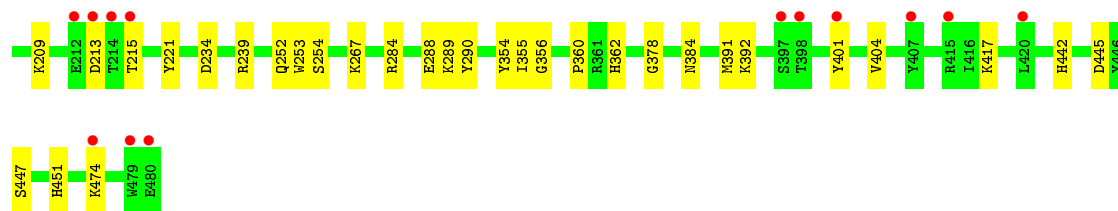


- Molecule 1: Nitrogenase molybdenum-iron protein alpha chain

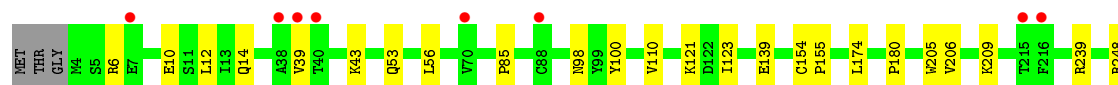
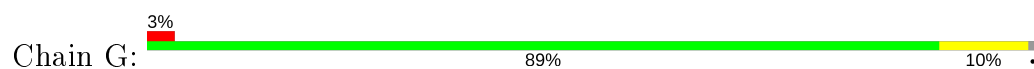


- Molecule 1: Nitrogenase molybdenum-iron protein alpha chain

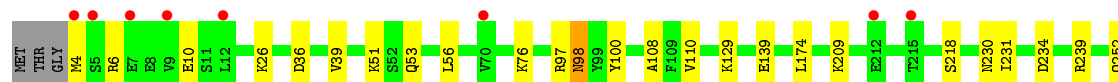
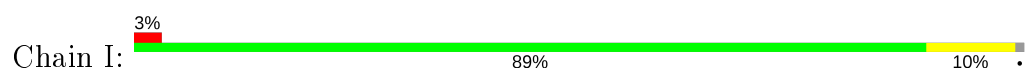




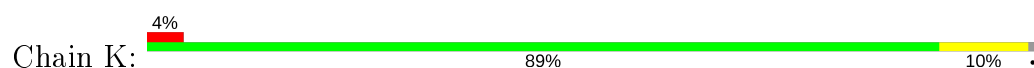
- Molecule 1: Nitrogenase molybdenum-iron protein alpha chain



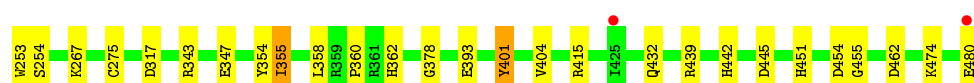
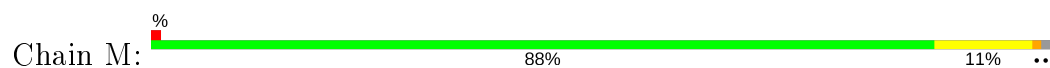
- Molecule 1: Nitrogenase molybdenum-iron protein alpha chain



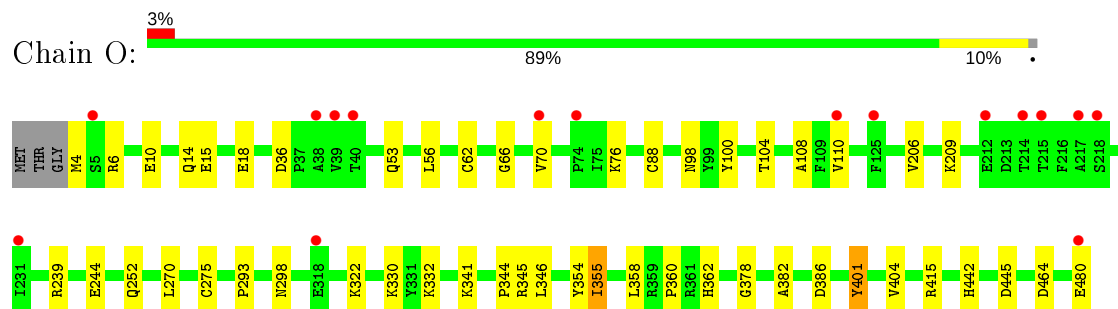
- Molecule 1: Nitrogenase molybdenum-iron protein alpha chain



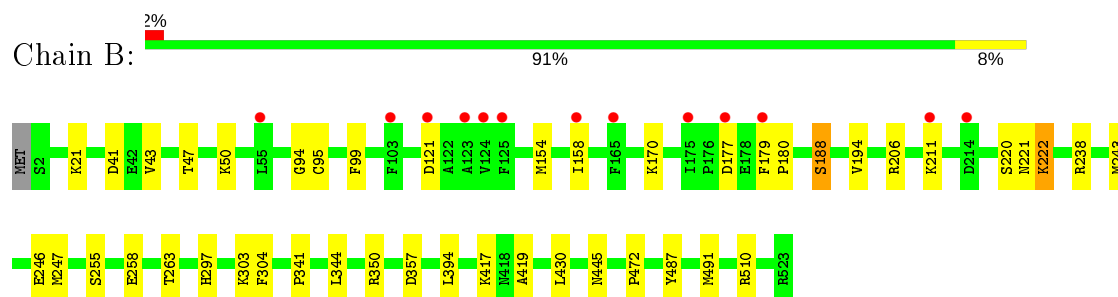
- Molecule 1: Nitrogenase molybdenum-iron protein alpha chain



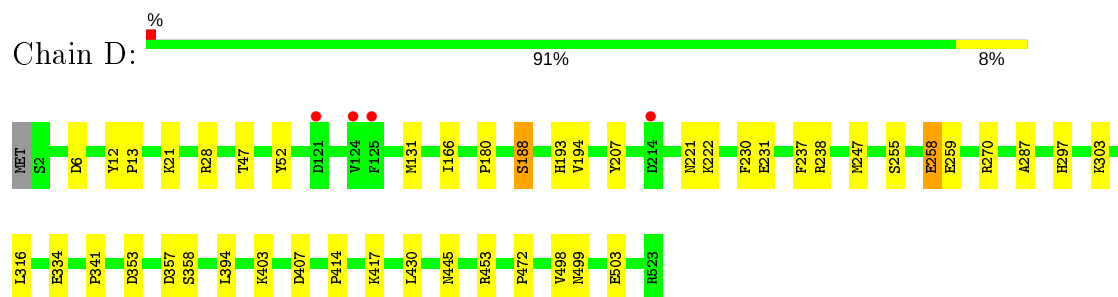
- Molecule 1: Nitrogenase molybdenum-iron protein alpha chain



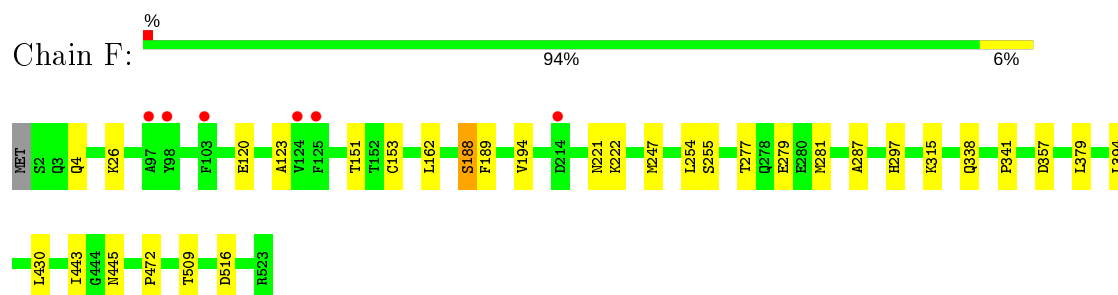
- Molecule 2: Nitrogenase molybdenum-iron protein beta chain



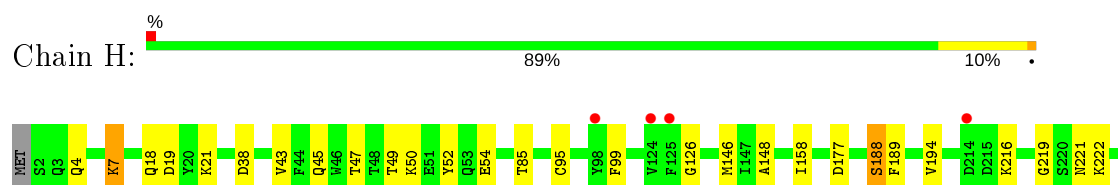
- Molecule 2: Nitrogenase molybdenum-iron protein beta chain



- Molecule 2: Nitrogenase molybdenum-iron protein beta chain



- Molecule 2: Nitrogenase molybdenum-iron protein beta chain





- Molecule 2: Nitrogenase molybdenum-iron protein beta chain



- Molecule 2: Nitrogenase molybdenum-iron protein beta chain



- Molecule 2: Nitrogenase molybdenum-iron protein beta chain



- Molecule 2: Nitrogenase molybdenum-iron protein beta chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	175.50Å 144.59Å 177.75Å 90.00° 114.27° 90.00°	Depositor
Resolution (Å)	38.68 – 1.75 39.01 – 1.75	Depositor EDS
% Data completeness (in resolution range)	95.6 (38.68-1.75) 96.4 (39.01-1.75)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.98 (at 1.75Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.198 , 0.245 0.199 , 0.242	Depositor DCC
$R_{free}$ test set	78317 reflections (9.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	13.2	Xtriage
Anisotropy	0.536	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 42.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.039 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	74588	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 34.33 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.9357e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: ICS, CLF, HCA, CA

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	0.55	0/3862	0.67	0/5212
1	C	0.55	0/3870	0.66	0/5220
1	E	0.50	0/3864	0.65	0/5212
1	G	0.54	0/3874	0.65	0/5225
1	I	0.52	0/3874	0.65	0/5225
1	K	0.54	0/3870	0.67	0/5219
1	M	0.57	0/3878	0.68	0/5229
1	O	0.53	0/3866	0.66	0/5217
2	B	0.56	1/4276 (0.0%)	0.69	3/5783 (0.1%)
2	D	0.59	1/4280 (0.0%)	0.69	3/5787 (0.1%)
2	F	0.56	2/4280 (0.0%)	0.66	3/5787 (0.1%)
2	H	0.57	2/4280 (0.0%)	0.68	4/5787 (0.1%)
2	J	0.56	2/4278 (0.0%)	0.68	3/5784 (0.1%)
2	L	0.55	2/4276 (0.0%)	0.70	5/5783 (0.1%)
2	N	0.58	1/4280 (0.0%)	0.69	3/5787 (0.1%)
2	P	0.58	2/4280 (0.0%)	0.69	3/5787 (0.1%)
All	All	0.55	13/65188 (0.0%)	0.67	27/88044 (0.0%)

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	I	0	1
1	M	0	1
All	All	0	3

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	188	SER	CA-CB	8.72	1.66	1.52
2	J	188	SER	CB-OG	8.14	1.52	1.42
2	P	188	SER	CB-OG	8.08	1.52	1.42
2	F	188	SER	CB-OG	7.64	1.52	1.42
2	F	188	SER	CA-CB	7.61	1.64	1.52

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	J	357	ASP	CB-CG-OD1	12.86	129.88	118.30
2	B	357	ASP	CB-CG-OD1	11.96	129.06	118.30
2	H	357	ASP	CB-CG-OD1	11.67	128.80	118.30
2	P	357	ASP	CB-CG-OD1	11.62	128.76	118.30
2	F	357	ASP	CB-CG-OD1	10.97	128.17	118.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	98	ASN	Sidechain
1	I	98	ASN	Sidechain
1	M	98	ASN	Sidechain

## 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	3774	0	3690	47	0
1	C	3782	0	3712	33	0
1	E	3776	0	3703	33	0
1	G	3786	0	3718	30	0
1	I	3786	0	3718	33	0
1	K	3782	0	3719	31	0
1	M	3790	0	3729	39	0
1	O	3778	0	3696	31	0
2	B	4170	0	4069	30	0
2	D	4174	0	4080	34	0
2	F	4174	0	4080	21	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	H	4174	0	4080	41	0
2	J	4172	0	4076	22	0
2	L	4170	0	4069	25	0
2	N	4174	0	4080	20	0
2	P	4174	0	4080	23	0
3	A	14	0	7	2	0
3	C	14	0	7	3	0
3	E	14	0	6	2	0
3	G	14	0	7	2	0
3	I	14	0	6	1	0
3	K	14	0	6	2	0
3	M	14	0	7	3	0
3	O	14	0	7	4	0
4	A	18	0	0	0	0
4	C	18	0	0	1	0
4	E	18	0	0	2	0
4	G	18	0	0	0	0
4	I	18	0	0	0	0
4	K	18	0	0	0	0
4	M	18	0	0	0	0
4	O	18	0	0	0	0
5	A	15	0	0	0	0
5	C	15	0	0	0	0
5	E	15	0	0	0	0
5	G	15	0	0	0	0
5	I	15	0	0	0	0
5	K	15	0	0	0	0
5	M	15	0	0	0	0
5	O	15	0	0	0	0
6	B	1	0	0	0	0
6	D	1	0	0	0	0
6	F	1	0	0	0	0
6	H	1	0	0	0	0
6	J	1	0	0	0	0
6	L	1	0	0	0	0
6	N	1	0	0	0	0
6	P	1	0	0	0	0
7	A	542	0	0	15	3
7	B	753	0	0	8	1
7	C	582	0	0	6	3
7	D	752	0	0	16	0
7	E	507	0	0	13	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	F	750	0	0	9	1
7	G	615	0	0	8	0
7	H	749	0	0	16	0
7	I	557	0	0	15	2
7	J	761	0	0	7	1
7	K	587	0	0	9	2
7	L	709	0	0	11	2
7	M	602	0	0	16	2
7	N	797	0	0	11	2
7	O	548	0	0	8	1
7	P	757	0	0	10	0
All	All	74588	0	62352	485	11

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 4.

The worst 5 of 485 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:K:286:MET:SD	7:K:728:HOH:O	2.17	1.02
4:E:502:ICS:S1A	7:E:926:HOH:O	2.20	0.98
2:H:85:THR:HG23	7:H:726:HOH:O	1.77	0.84
1:E:206:VAL:HA	1:E:209:LYS:HE2	1.60	0.82
1:O:345:ARG:HG3	7:O:601:HOH:O	1.80	0.81

The worst 5 of 11 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:I:605:HOH:O	7:L:811:HOH:O[2_748]	2.04	0.16
7:B:1319:HOH:O	7:K:828:HOH:O[2_748]	2.13	0.07
7:C:976:HOH:O	7:J:1140:HOH:O[1_554]	2.14	0.06
7:F:1442:HOH:O	7:O:1126:HOH:O[2_848]	2.14	0.06
7:A:1127:HOH:O	7:M:1160:HOH:O[1_455]	2.16	0.04

## 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	475/480 (99%)	456 (96%)	18 (4%)	1 (0%)	47	29
1	C	475/480 (99%)	457 (96%)	17 (4%)	1 (0%)	47	29
1	E	475/480 (99%)	457 (96%)	18 (4%)	0	100	100
1	G	475/480 (99%)	458 (96%)	16 (3%)	1 (0%)	47	29
1	I	475/480 (99%)	455 (96%)	19 (4%)	1 (0%)	47	29
1	K	475/480 (99%)	458 (96%)	16 (3%)	1 (0%)	47	29
1	M	475/480 (99%)	456 (96%)	18 (4%)	1 (0%)	47	29
1	O	475/480 (99%)	455 (96%)	19 (4%)	1 (0%)	47	29
2	B	520/523 (99%)	511 (98%)	8 (2%)	1 (0%)	47	29
2	D	520/523 (99%)	510 (98%)	9 (2%)	1 (0%)	47	29
2	F	520/523 (99%)	509 (98%)	10 (2%)	1 (0%)	47	29
2	H	520/523 (99%)	511 (98%)	8 (2%)	1 (0%)	47	29
2	J	520/523 (99%)	510 (98%)	9 (2%)	1 (0%)	47	29
2	L	520/523 (99%)	509 (98%)	10 (2%)	1 (0%)	47	29
2	N	520/523 (99%)	513 (99%)	6 (1%)	1 (0%)	47	29
2	P	520/523 (99%)	509 (98%)	11 (2%)	0	100	100
All	All	7960/8024 (99%)	7734 (97%)	212 (3%)	14 (0%)	47	29

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	255	SER
2	F	255	SER
2	H	255	SER
2	J	255	SER
2	L	255	SER

### 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	403/409 (98%)	398 (99%)	5 (1%)	71	56
1	C	405/409 (99%)	401 (99%)	4 (1%)	76	63
1	E	403/409 (98%)	400 (99%)	3 (1%)	84	75
1	G	406/409 (99%)	403 (99%)	3 (1%)	84	75
1	I	406/409 (99%)	402 (99%)	4 (1%)	76	63
1	K	405/409 (99%)	398 (98%)	7 (2%)	60	41
1	M	407/409 (100%)	402 (99%)	5 (1%)	71	56
1	O	404/409 (99%)	398 (98%)	6 (2%)	65	47
2	B	453/455 (100%)	452 (100%)	1 (0%)	93	90
2	D	454/455 (100%)	453 (100%)	1 (0%)	93	90
2	F	454/455 (100%)	453 (100%)	1 (0%)	93	90
2	H	454/455 (100%)	452 (100%)	2 (0%)	91	86
2	J	453/455 (100%)	452 (100%)	1 (0%)	93	90
2	L	453/455 (100%)	452 (100%)	1 (0%)	93	90
2	N	454/455 (100%)	454 (100%)	0	100	100
2	P	454/455 (100%)	454 (100%)	0	100	100
All	All	6868/6912 (99%)	6824 (99%)	44 (1%)	86	79

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	258	GLU
2	J	51	GLU
1	O	362	HIS
1	I	26	LYS
1	I	362	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	362	HIS
2	N	317	ASN
1	G	98	ASN
1	C	98	ASN
1	K	98	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 carbohydrates in this entry.

### 5.6 Ligand geometry ⓘ

Of 32 ligands modelled in this entry, 8 are monoatomic - leaving 24 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	$\# Z  > 2$
4	ICS	A	502	1	18,30,30	2.02	6 (33%)	-		
5	CLF	O	503	1,2	0,24,24	0.00	-	-		
3	HCA	G	501	-	4,13,13	1.12	1 (25%)	4,18,18	2.68	3 (75%)
3	HCA	K	501	-	4,13,13	0.65	0	4,18,18	2.36	2 (50%)
5	CLF	A	503	1,2,7	0,24,24	0.00	-	-		
4	ICS	K	502	1	18,30,30	2.49	10 (55%)	-		
4	ICS	M	502	1	18,30,30	2.69	11 (61%)	-		
4	ICS	G	502	1,7	18,30,30	2.53	8 (44%)	-		

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	HCA	C	501	-	4,13,13	0.72	0	4,18,18	2.64	3 (75%)
5	CLF	I	503	1,2	0,24,24	0.00	-	-		
5	CLF	C	503	1,2	0,24,24	0.00	-	-		
4	ICS	C	502	1	18,30,30	2.35	10 (55%)	-		
4	ICS	E	502	1	18,30,30	2.79	11 (61%)	-		
3	HCA	A	501	-	4,13,13	0.99	0	4,18,18	2.14	2 (50%)
3	HCA	E	501	-	4,13,13	0.74	0	4,18,18	2.66	2 (50%)
5	CLF	M	503	1,2	0,24,24	0.00	-	-		
3	HCA	O	501	-	4,13,13	0.82	0	4,18,18	2.41	2 (50%)
5	CLF	K	503	1,2	0,24,24	0.00	-	-		
3	HCA	M	501	-	4,13,13	0.28	0	4,18,18	2.90	4 (100%)
5	CLF	E	503	1,2	0,24,24	0.00	-	-		
4	ICS	O	502	1	18,30,30	2.74	11 (61%)	-		
5	CLF	G	503	1,2	0,24,24	0.00	-	-		
3	HCA	I	501	-	4,13,13	1.29	0	4,18,18	2.64	2 (50%)
4	ICS	I	502	1	18,30,30	2.84	10 (55%)	-		

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	HCA	M	501	-	-	1/7/17/17	-
5	CLF	E	503	1,2	-	-	0/12/10/10
5	CLF	C	503	1,2	-	-	0/12/10/10
3	HCA	C	501	-	-	2/7/17/17	-
5	CLF	O	503	1,2	-	-	0/12/10/10
3	HCA	G	501	-	-	1/7/17/17	-
3	HCA	A	501	-	-	0/7/17/17	-
5	CLF	I	503	1,2	-	-	0/12/10/10
3	HCA	K	501	-	-	0/7/17/17	-
5	CLF	G	503	1,2	-	-	0/12/10/10
3	HCA	E	501	-	-	0/7/17/17	-
5	CLF	M	503	1,2	-	-	0/12/10/10
3	HCA	O	501	-	-	0/7/17/17	-
5	CLF	K	503	1,2	-	-	0/12/10/10
3	HCA	I	501	-	-	0/7/17/17	-
5	CLF	A	503	1,2,7	-	-	0/12/10/10

The worst 5 of 78 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	I	502	ICS	S1B-FE6	-5.68	2.18	2.32
4	O	502	ICS	S1B-FE6	-5.34	2.19	2.32
4	E	502	ICS	S4B-FE7	-5.02	2.20	2.32
4	I	502	ICS	S4B-FE7	-4.76	2.20	2.32
4	G	502	ICS	S4B-FE7	-4.68	2.20	2.32

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	501	HCA	C4-C5-C6	4.29	117.89	111.39
3	E	501	HCA	C4-C5-C6	3.90	117.30	111.39
3	K	501	HCA	C4-C5-C6	3.77	117.09	111.39
3	G	501	HCA	C3-C2-C1	-3.54	109.32	114.98
3	M	501	HCA	C4-C3-C7	-3.52	105.33	111.52

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	G	501	HCA	O7-C3-C4-C5
3	M	501	HCA	O7-C3-C4-C5
3	C	501	HCA	C1-C2-C3-C4
3	C	501	HCA	C1-C2-C3-O7

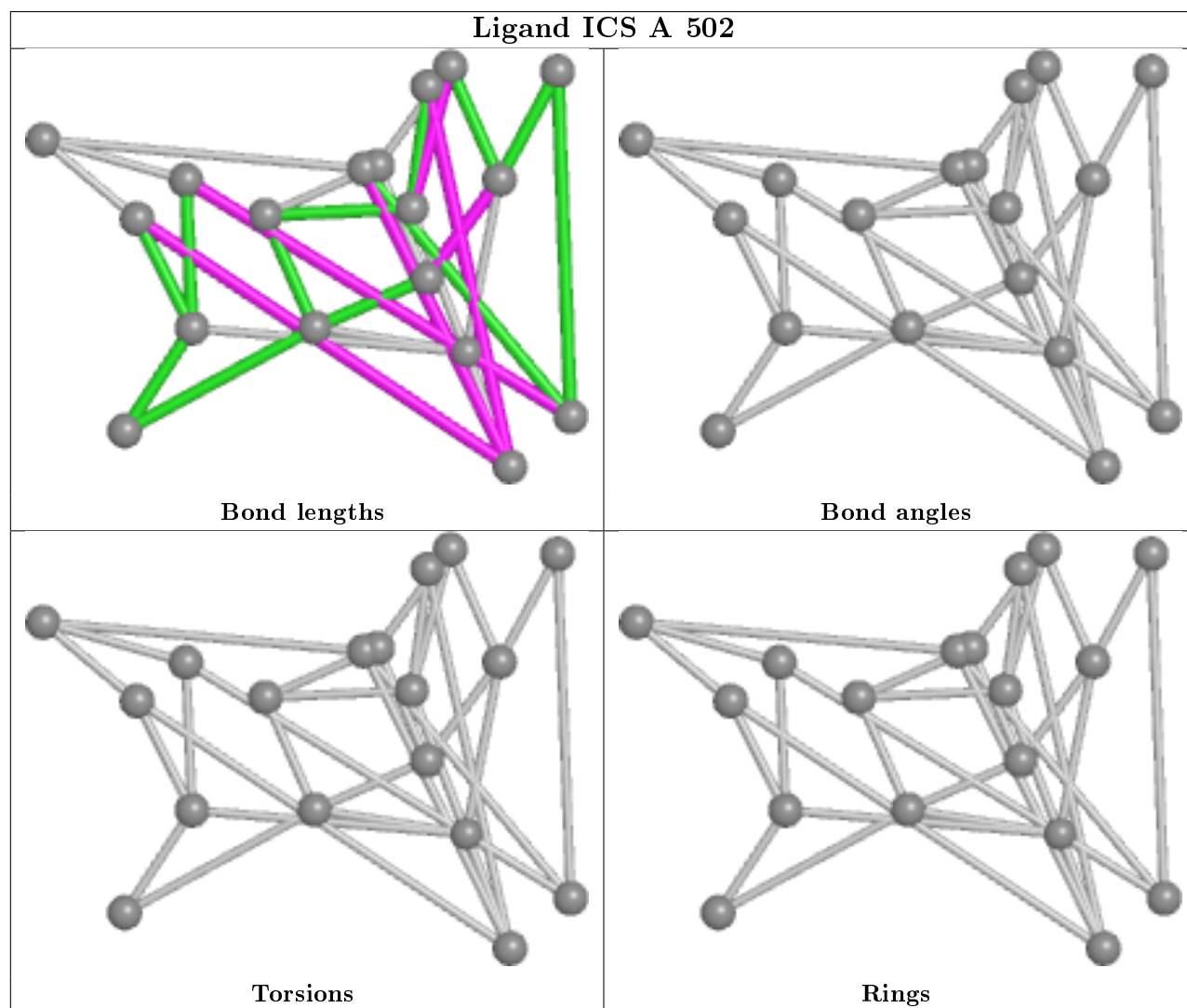
There are no ring outliers.

10 monomers are involved in 22 short contacts:

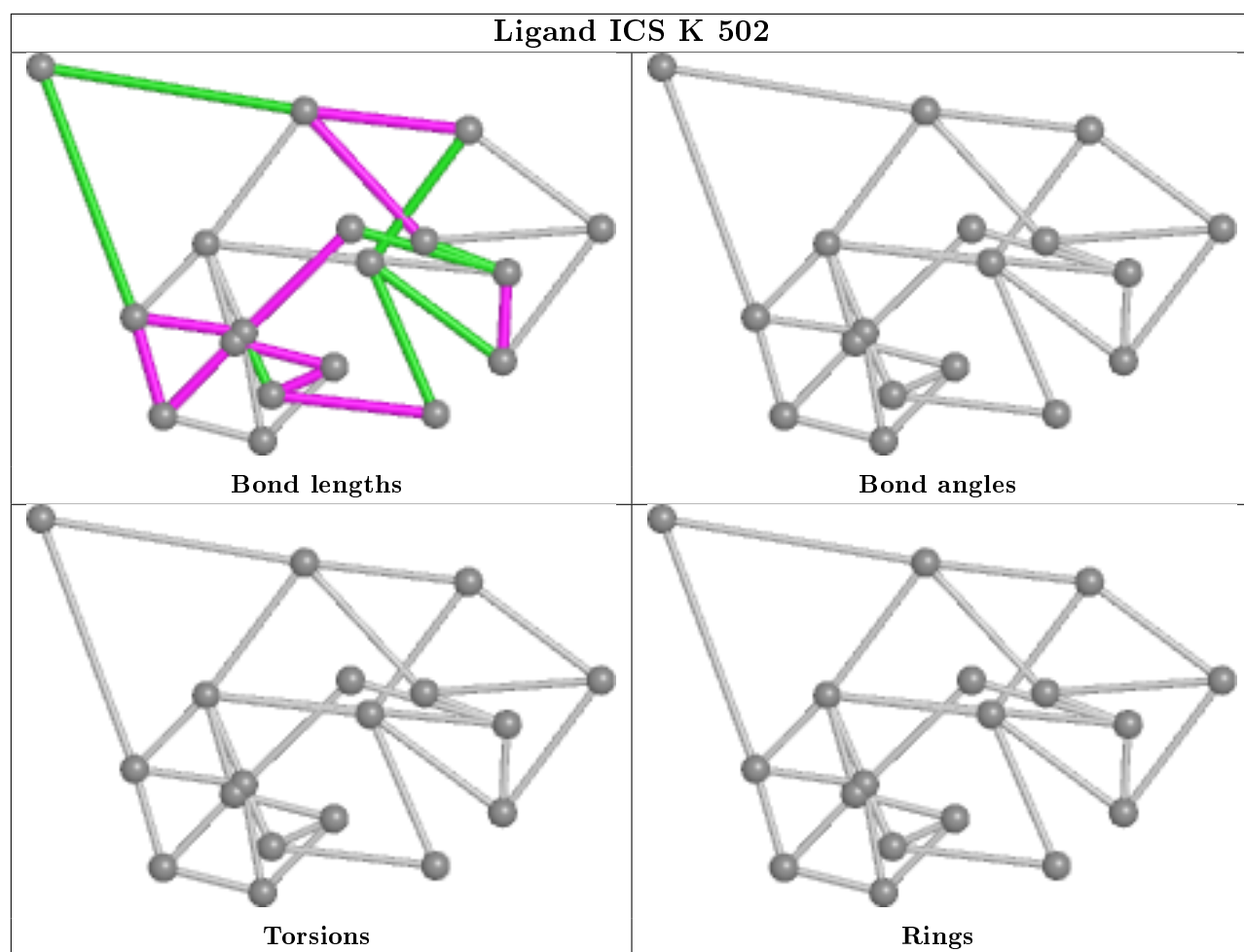
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	G	501	HCA	2	0
3	K	501	HCA	2	0
3	C	501	HCA	3	0
4	C	502	ICS	1	0
4	E	502	ICS	2	0
3	A	501	HCA	2	0
3	E	501	HCA	2	0
3	O	501	HCA	4	0
3	M	501	HCA	3	0
3	I	501	HCA	1	0

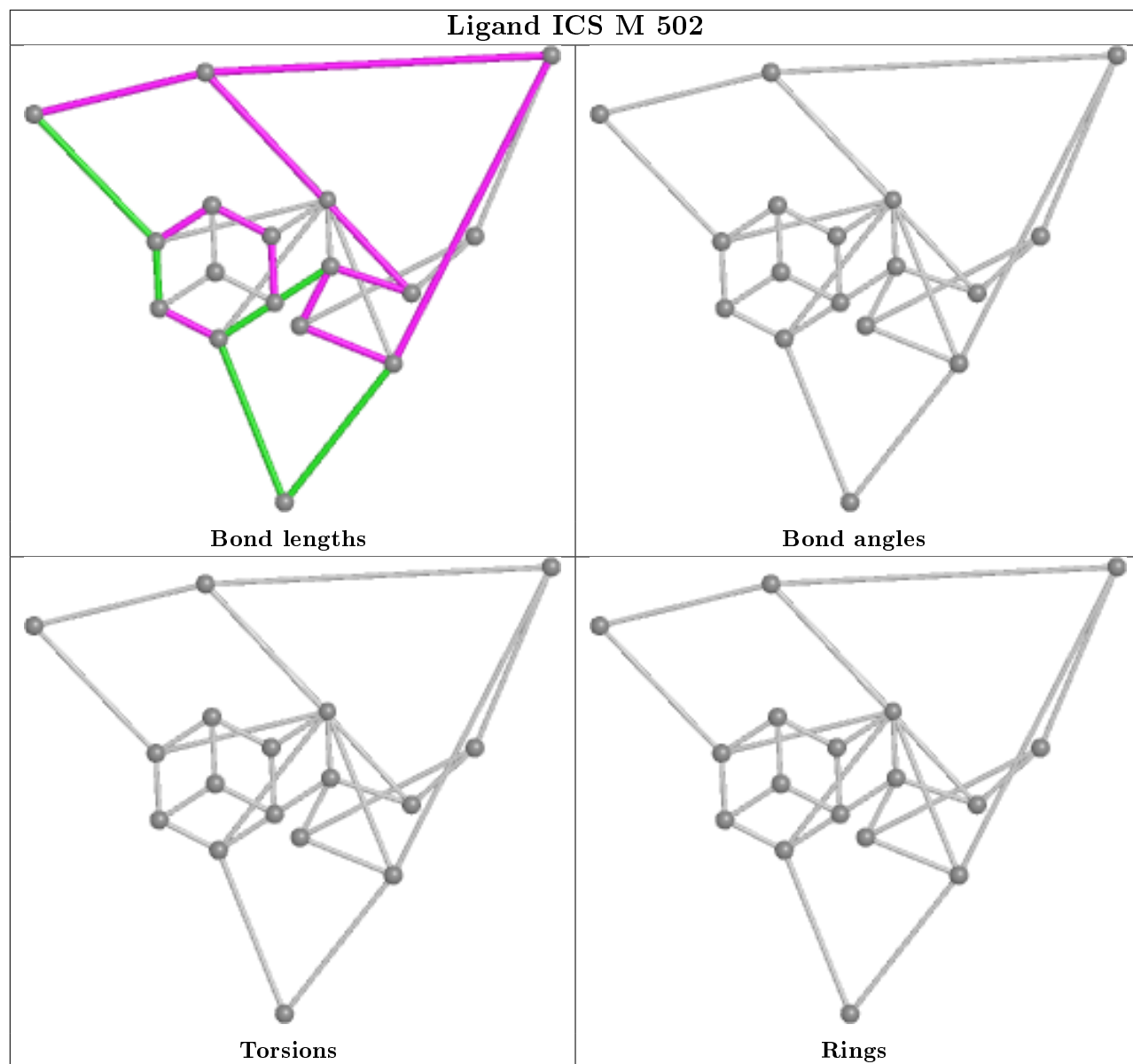
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths,

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.

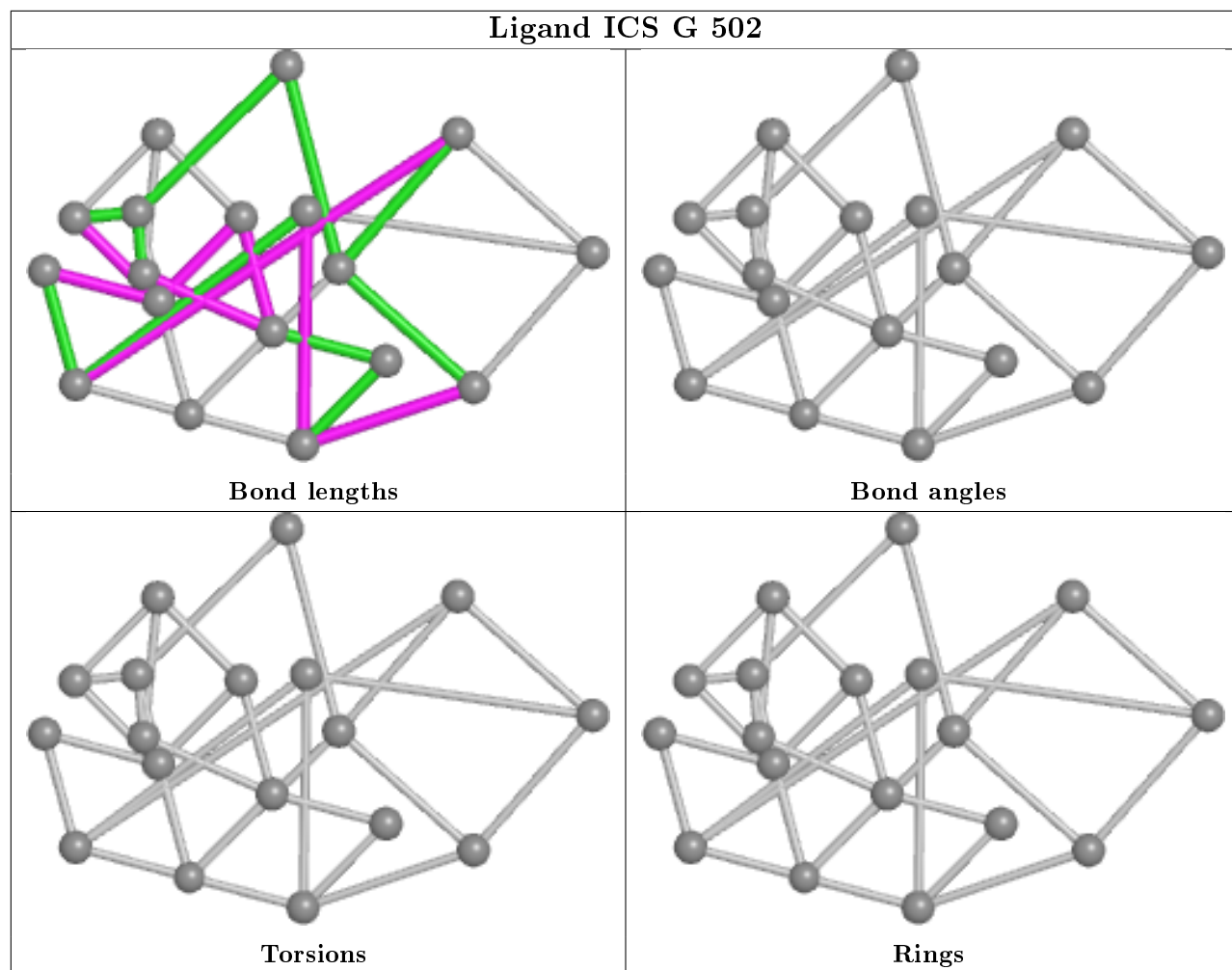


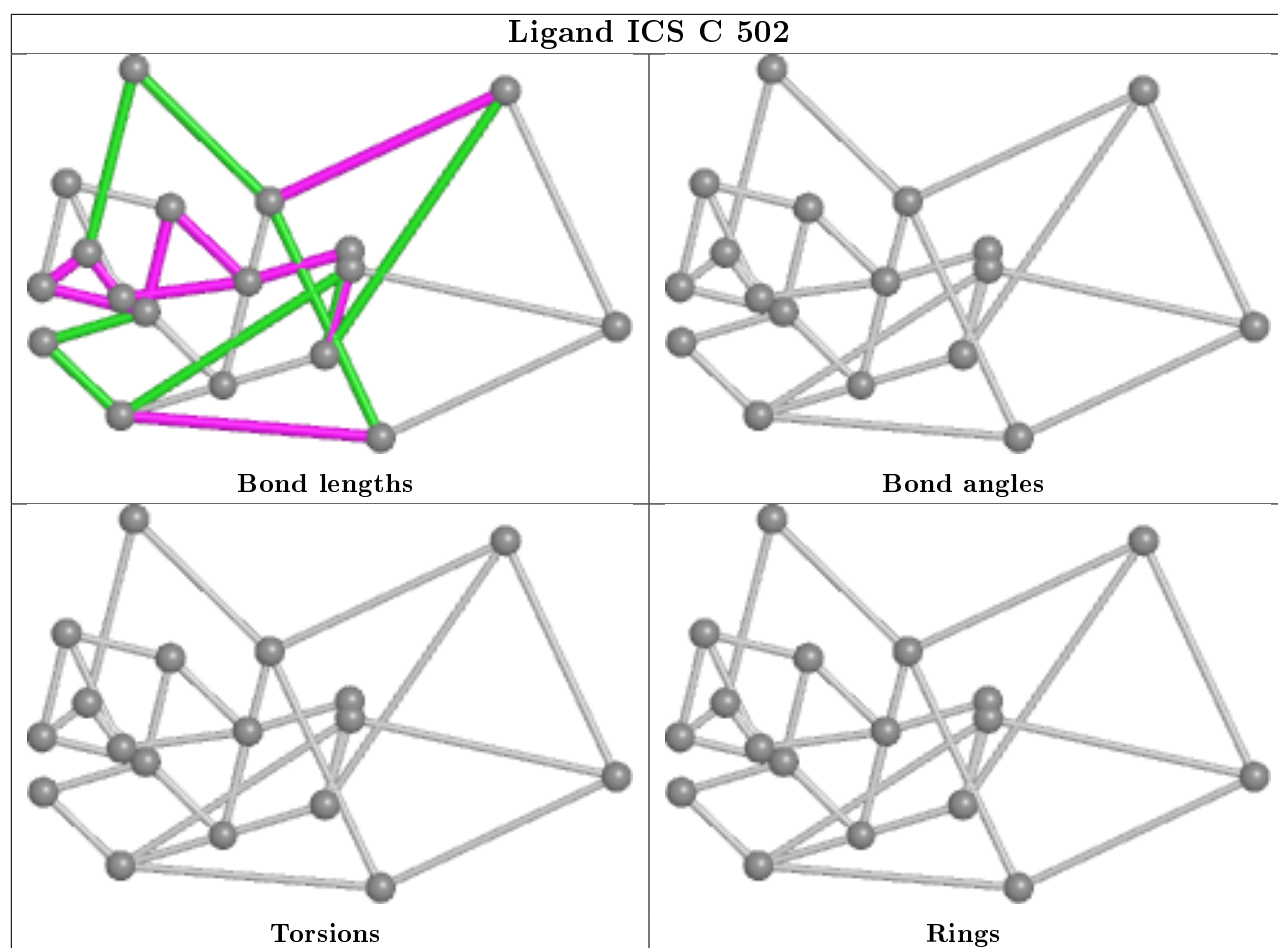


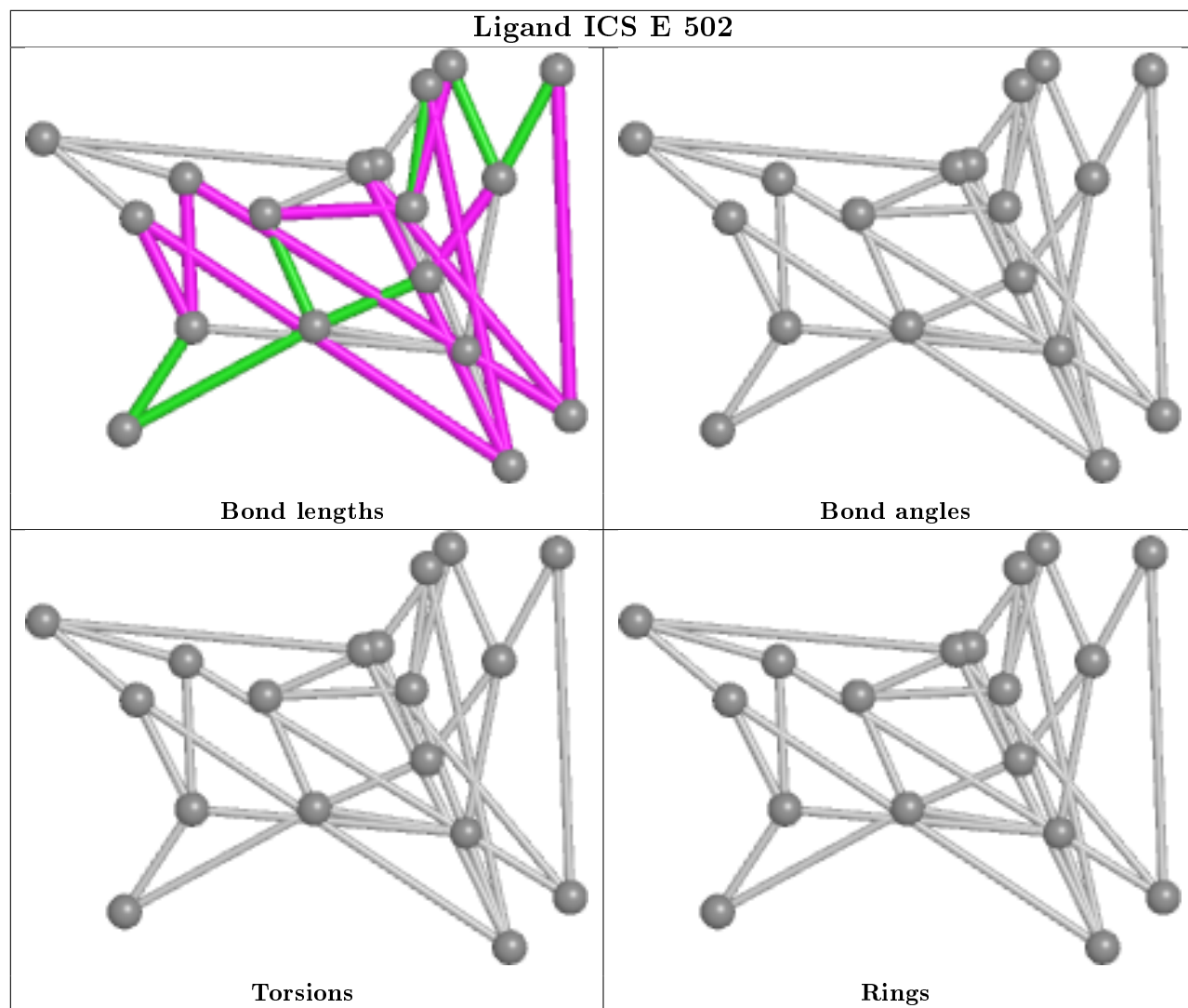




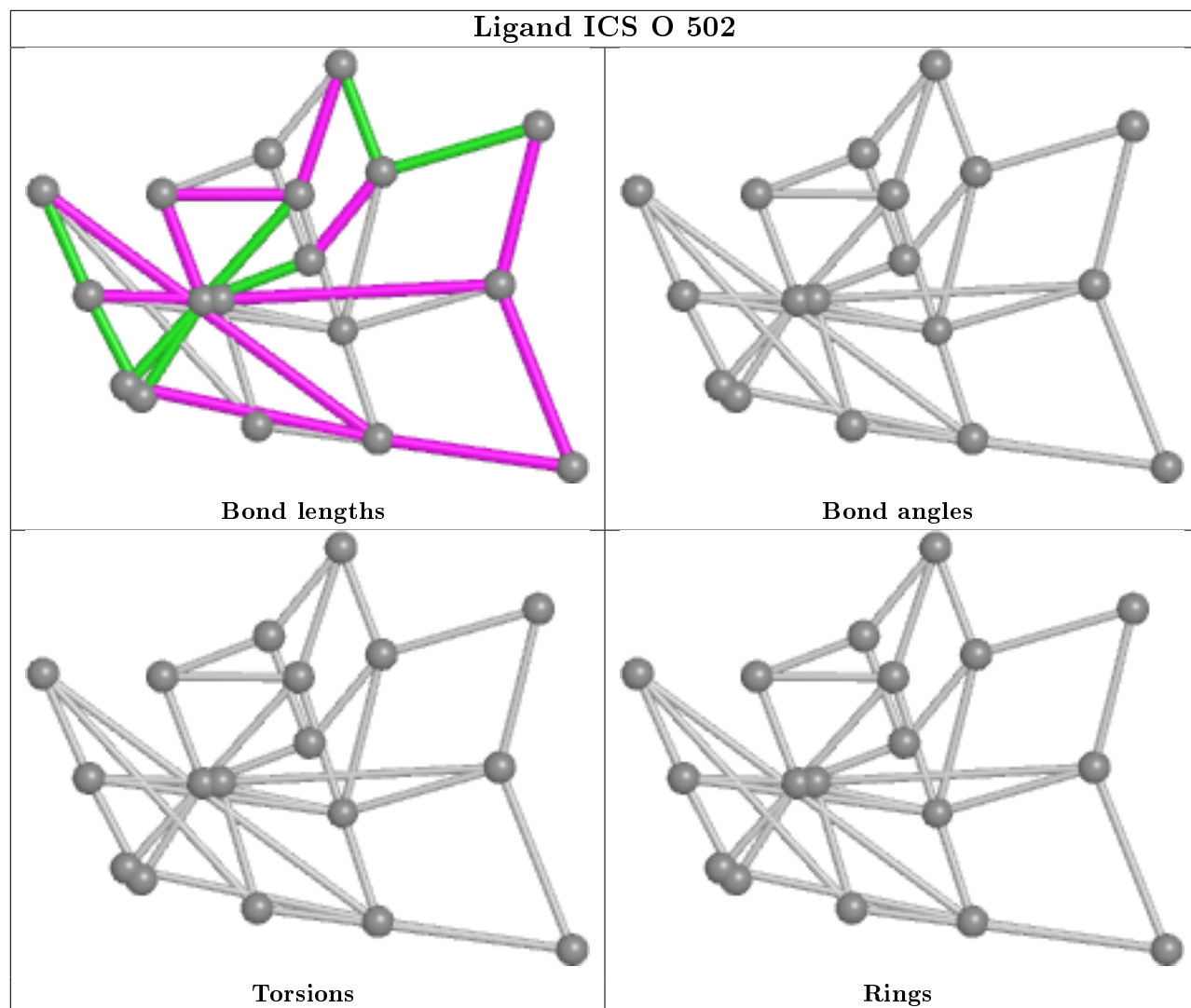
## Ligand ICS G 502

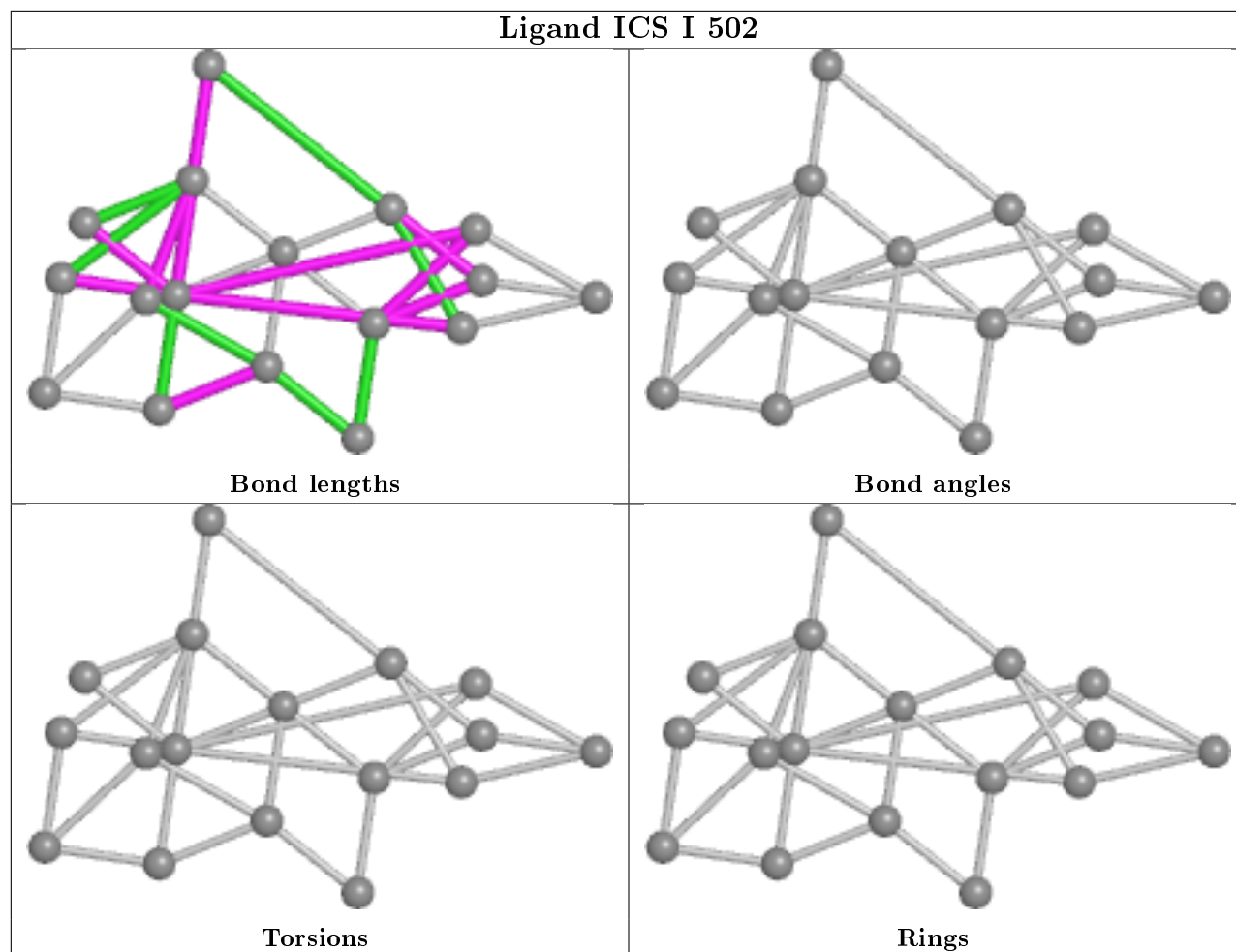






## Ligand ICS O 502





## 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 ⓘ

### 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, 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	477/480 (99%)	0.51	26 (5%)	25 30	9, 16, 31, 50	0
1	C	477/480 (99%)	0.49	21 (4%)	34 39	10, 15, 30, 44	0
1	E	477/480 (99%)	0.64	31 (6%)	18 23	11, 19, 34, 50	0
1	G	477/480 (99%)	0.29	12 (2%)	57 63	9, 14, 29, 40	0
1	I	477/480 (99%)	0.46	14 (2%)	51 57	10, 16, 27, 52	0
1	K	477/480 (99%)	0.46	19 (3%)	38 43	8, 15, 31, 43	0
1	M	477/480 (99%)	0.23	6 (1%)	77 82	8, 13, 24, 39	0
1	O	477/480 (99%)	0.39	16 (3%)	45 51	9, 15, 29, 40	0
2	B	522/523 (99%)	0.38	13 (2%)	57 63	9, 15, 26, 38	0
2	D	522/523 (99%)	0.19	4 (0%)	86 90	9, 14, 23, 37	0
2	F	522/523 (99%)	0.30	6 (1%)	80 85	8, 15, 25, 34	0
2	H	522/523 (99%)	0.23	6 (1%)	80 85	8, 15, 23, 33	0
2	J	522/523 (99%)	0.37	8 (1%)	73 80	9, 16, 25, 35	0
2	L	522/523 (99%)	0.42	15 (2%)	51 57	9, 16, 28, 42	0
2	N	522/523 (99%)	0.18	4 (0%)	86 90	7, 13, 23, 35	0
2	P	522/523 (99%)	0.17	3 (0%)	89 92	7, 13, 22, 29	0
All	All	7992/8024 (99%)	0.35	204 (2%)	56 61	7, 15, 27, 52	0

The worst 5 of 204 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	38	ALA	8.4
1	C	38	ALA	7.9
1	A	38	ALA	7.8
1	O	38	ALA	7.5
1	A	40	THR	7.1



## 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 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. 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	B-factors( $\text{\AA}^2$ )	Q<0.9
3	HCA	A	501	14/14	0.92	0.13	10,12,18,18	0
3	HCA	C	501	14/14	0.93	0.12	9,12,14,16	0
3	HCA	E	501	14/14	0.93	0.14	11,14,17,17	0
3	HCA	O	501	14/14	0.93	0.15	7,11,15,17	0
3	HCA	I	501	14/14	0.93	0.15	11,13,16,17	0
3	HCA	M	501	14/14	0.94	0.15	6,9,12,12	0
3	HCA	G	501	14/14	0.94	0.12	9,10,12,15	0
6	CA	L	601	1/1	0.95	0.13	14,14,14,14	0
3	HCA	K	501	14/14	0.95	0.14	9,12,15,15	0
6	CA	H	601	1/1	0.97	0.08	14,14,14,14	0
6	CA	N	601	1/1	0.98	0.11	11,11,11,11	0
6	CA	D	601	1/1	0.98	0.10	12,12,12,12	0
6	CA	F	601	1/1	0.98	0.11	13,13,13,13	0
4	ICS	E	502	18/18	0.98	0.07	11,14,17,18	0
5	CLF	A	503	15/15	0.98	0.06	11,12,14,15	0
4	ICS	A	502	18/18	0.99	0.06	10,12,15,15	0
5	CLF	C	503	15/15	0.99	0.05	10,11,13,14	0
4	ICS	C	502	18/18	0.99	0.06	9,12,14,14	0
5	CLF	O	503	15/15	0.99	0.07	9,11,12,12	0
4	ICS	K	502	18/18	0.99	0.07	9,12,13,13	0
4	ICS	M	502	18/18	0.99	0.08	5,10,11,12	0
5	CLF	M	503	15/15	0.99	0.07	8,9,11,13	0
6	CA	J	601	1/1	0.99	0.08	12,12,12,12	0
5	CLF	K	503	15/15	0.99	0.07	11,12,14,14	0
4	ICS	G	502	18/18	0.99	0.07	8,11,12,12	0
5	CLF	E	503	15/15	0.99	0.06	12,13,13,14	0
4	ICS	O	502	18/18	0.99	0.07	8,11,14,14	0

*Continued on next page...*

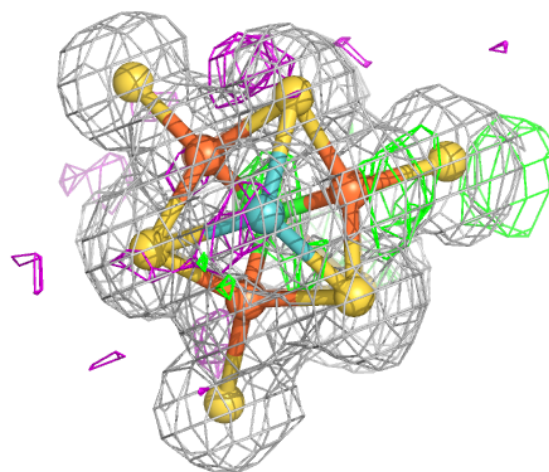
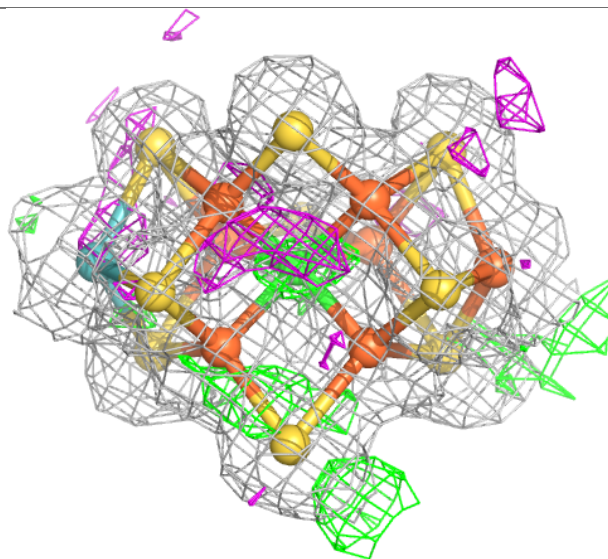
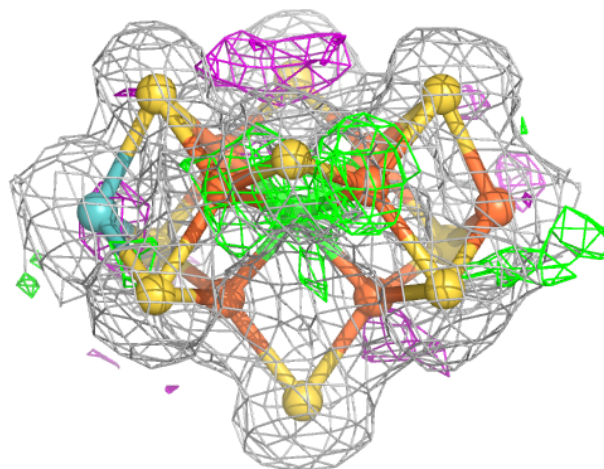
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	CA	B	601	1/1	0.99	0.10	12,12,12,12	0
6	CA	P	601	1/1	0.99	0.13	12,12,12,12	0
5	CLF	G	503	15/15	0.99	0.06	10,11,13,13	0
5	CLF	I	503	15/15	0.99	0.06	11,12,15,16	0
4	ICS	I	502	18/18	0.99	0.08	10,13,15,16	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.

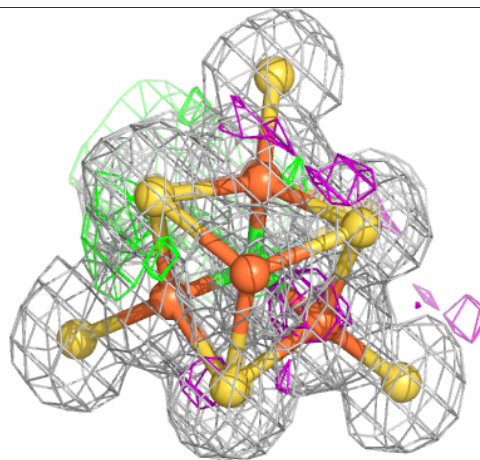
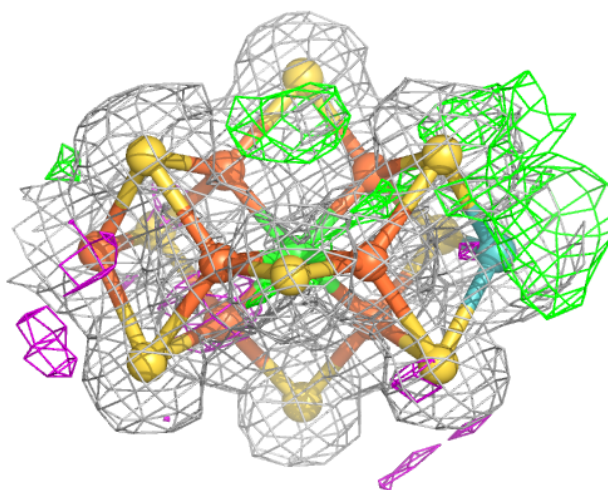
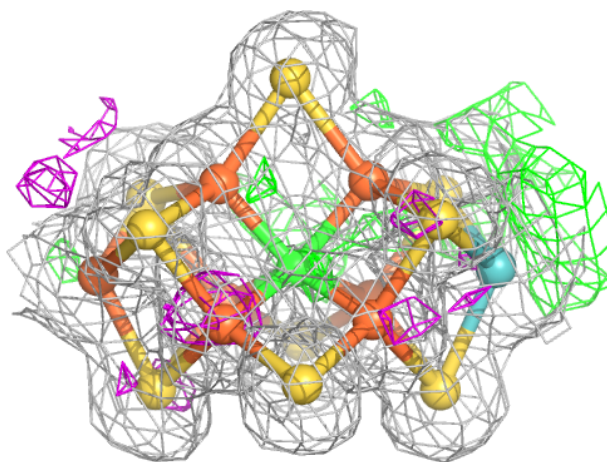
**Electron density around ICS E 502:**

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)



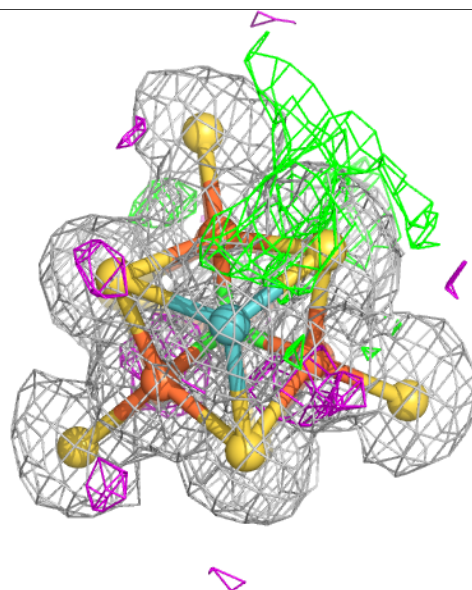
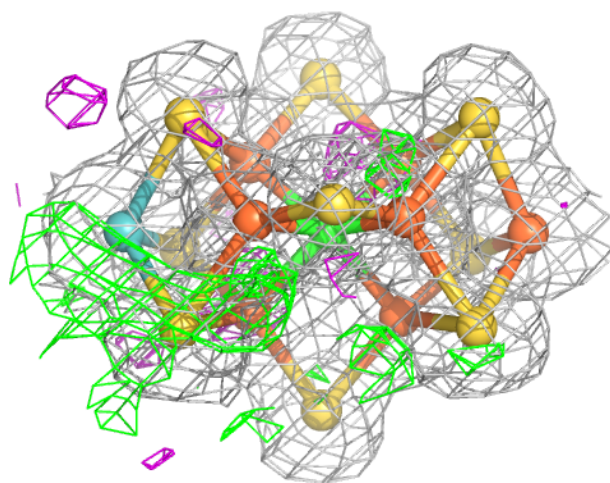
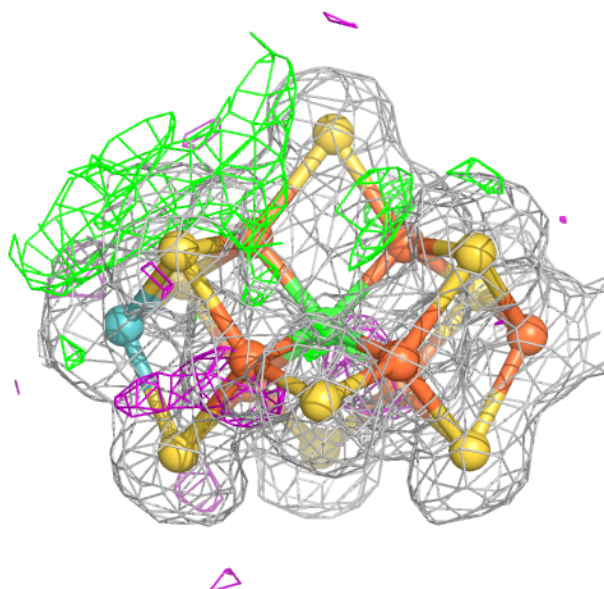
**Electron density around ICS A 502:**

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



**Electron density around ICS C 502:**

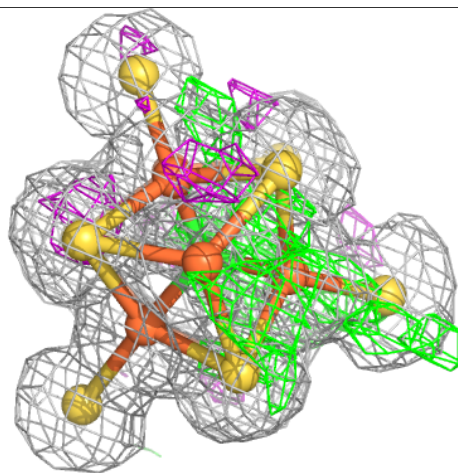
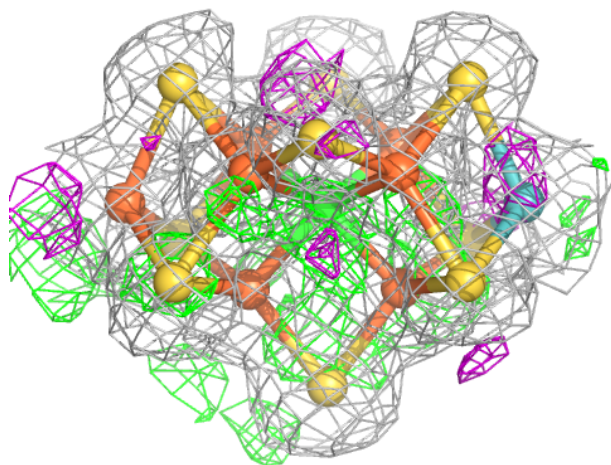
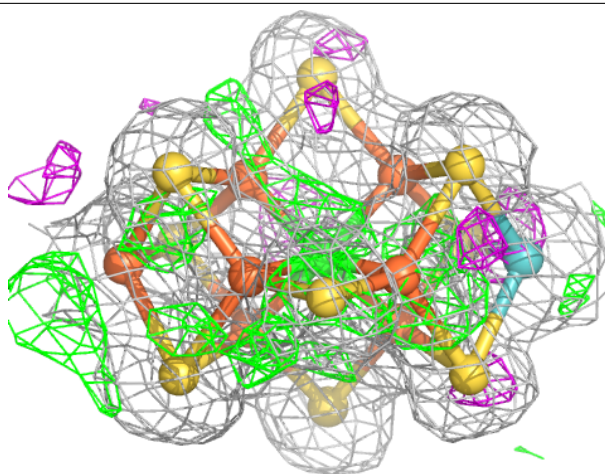
2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)





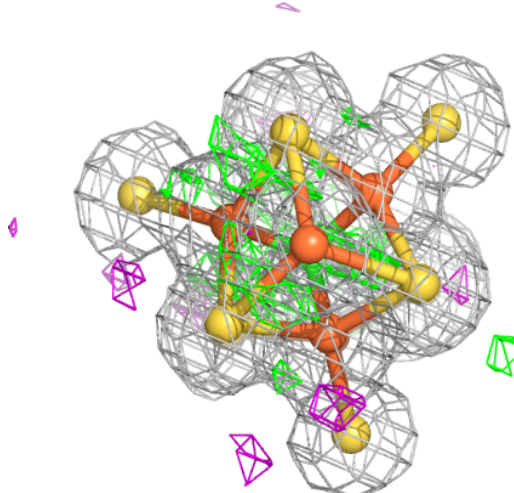
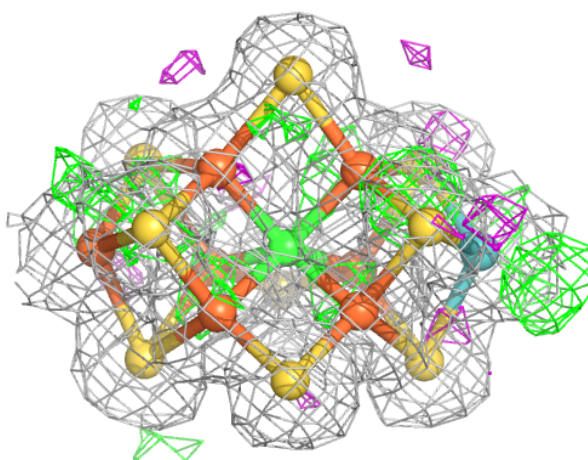
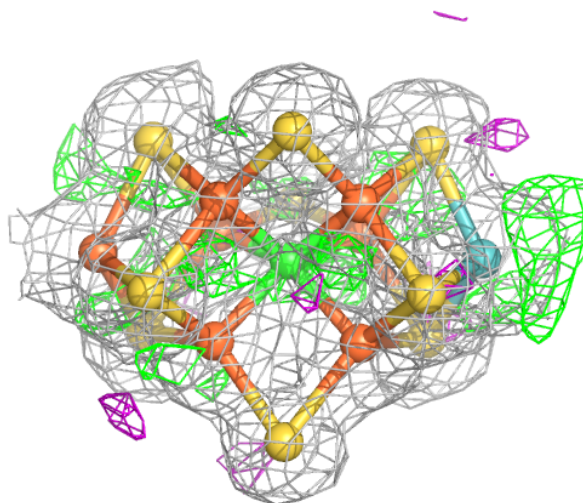
**Electron density around ICS K 502:**

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



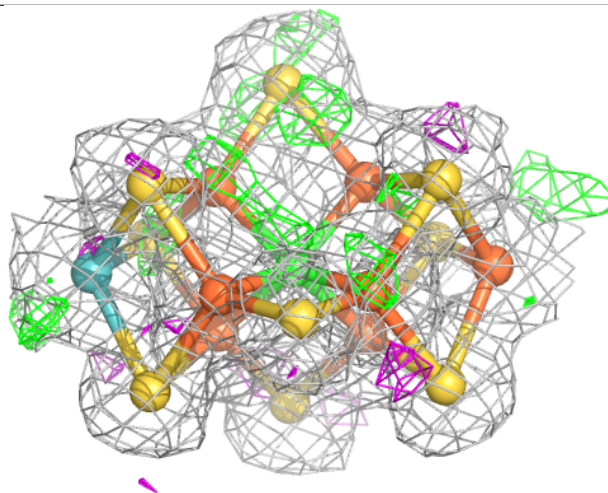
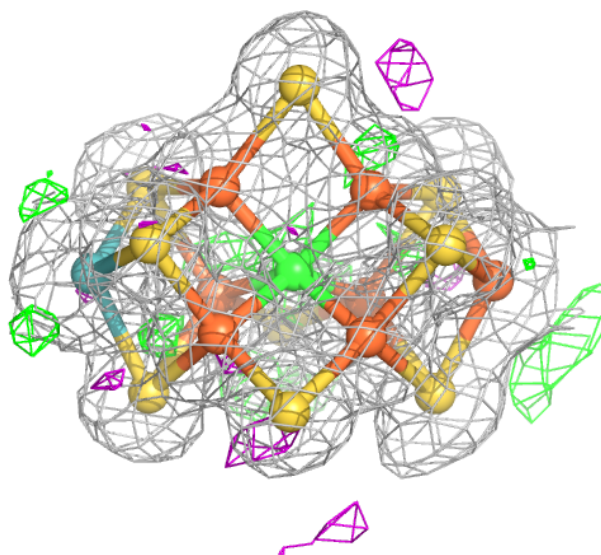
**Electron density around ICS M 502:**

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



**Electron density around ICS G 502:**

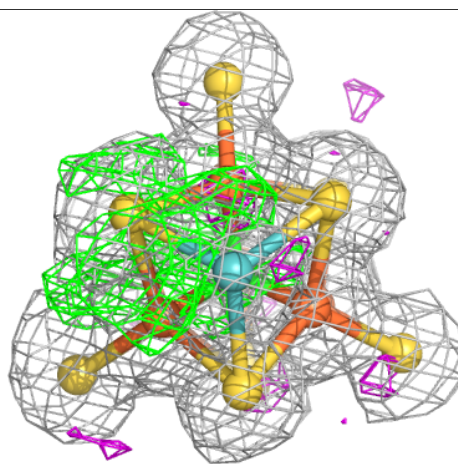
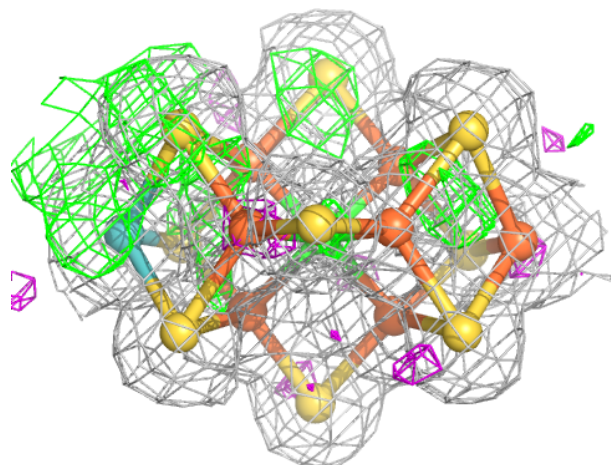
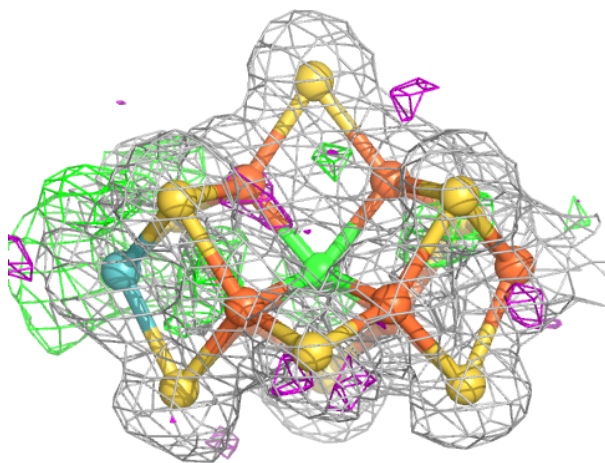
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





**Electron density around ICS O 502:**

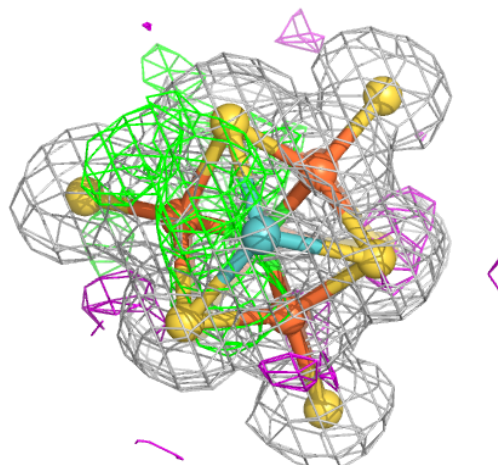
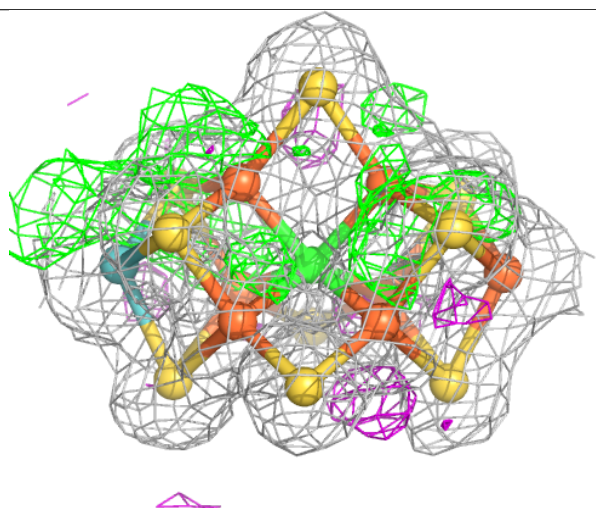
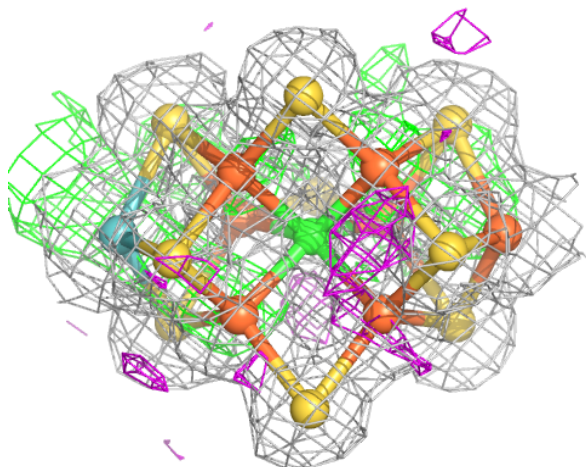
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





**Electron density around ICS I 502:**

$2mF_o-DF_c$  (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.