wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 23, 2023 – 07:30 PM EDT

PDB ID	:	1B48
Title	:	CRYSTAL STRUCTURE OF MGSTA4-4 IN COMPLEX WITH GSH CONJUGATE OF 4-HYDROXYNONENAL IN ONE SUBUNIT AND GSH IN THE OTHER: EVIDENCE OF SIGNALING ACROSS DIMER INTERFACE IN MGSTA4-4
Authors	:	Xiao, B.; Zimniak, P.; Ji, X.
Deposited on	:	1999-01-06
Resolution	:	2.60 Å(reported)

The following versions of software and data (see references ①) were used in the production of this report:

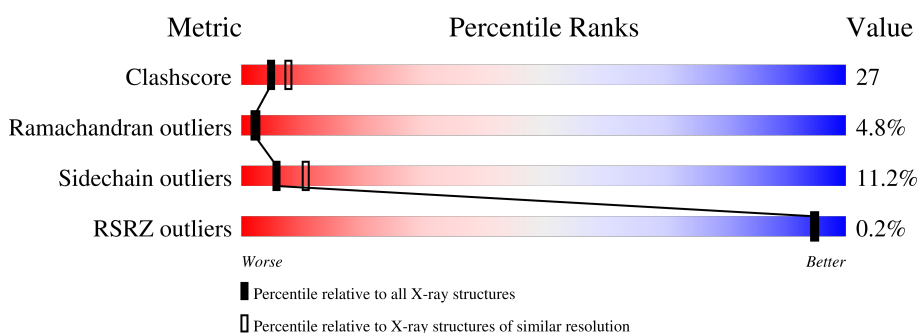|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 2.60 Å.

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

Metric	Percentile Ranks	Value
Clashscore		27
Ramachandran outliers		4.8%
Sidechain outliers		11.2%
RSRZ outliers		0.2%

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))
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	221	43% 51% 5%
1	B	221	46% 45% 9%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3677 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 PROTEIN (GLUTATHIONE S-TRANSFERASE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	221	Total	C	N	O	S	0	0	0
			1793	1166	295	323	9			
1	B	221	Total	C	N	O	S	0	0	0
			1793	1166	295	323	9			

- Molecule 2 is 4-S-GLUTATHIONYL-5-PENTYL-TETRAHYDRO-FURAN-2-OL (three-letter code: HAG) (formula: $C_{19}H_{33}N_3O_8S$).
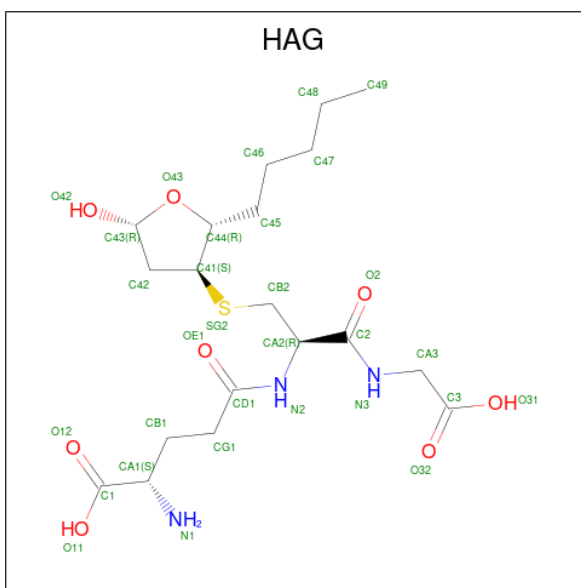
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			31	19	3	8	1		

- Molecule 3 is GLUTATHIONE (three-letter code: GSH) (formula: $C_{10}H_{17}N_3O_6S$).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	B	1	Total	C	N	O	S	0	0
			20	10	3	6	1		

• Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	18	Total	O	0	0
			18	18		
4	B	22	Total	O	0	0
			22	22		

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: PROTEIN (GLUTATHIONE S-TRANSFERASE)

Chain A:

43%   51%   5%

• Molecule 1: PROTEIN (GLUTATHIONE S-TRANSFERASE)

Chain B:

46%   45%   9%
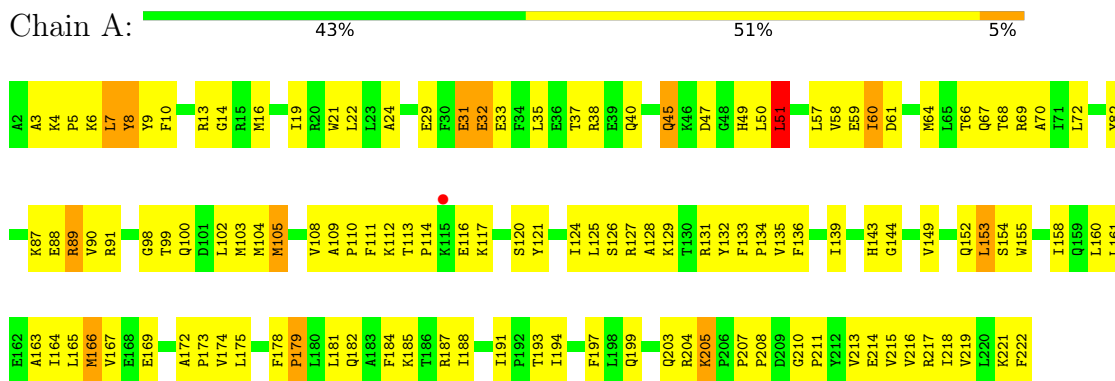
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	114.45Å 96.14Å 50.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.60 19.68 – 2.60	Depositor EDS
% Data completeness (in resolution range)	87.4 (20.00-2.60) 87.5 (19.68-2.60)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.88 (at 2.59Å)	Xtriage
Refinement program	X-PLOR 3.843	Depositor
R, R_{free}	0.244 , 0.317 0.233 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	45.3	Xtriage
Anisotropy	0.798	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.22 , 64.1	EDS
L-test for twinning[2]	$< |L| > = 0.45, < L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	3677	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.23% 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 (i)

5.1   Standard geometry (i)

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

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	0.45	0/1833	0.64	0/2471
1	B	0.43	0/1833	0.63	0/2471
All	All	0.44	0/3666	0.64	0/4942

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 (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1793	0	1834	103	0
1	B	1793	0	1834	105	0
2	A	31	0	31	2	0
3	B	20	0	16	2	0
4	A	18	0	0	2	0
4	B	22	0	0	3	0
All	All	3677	0	3715	202	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 27.

The worst 5 of 202 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:A:217:ARG:HA	1:A:222:PHE:HB2	1.61	0.82
1:B:163:ALA:O	1:B:167:VAL:HG23	1.83	0.78
1:B:149:VAL:HG13	1:B:150:GLY:H	1.48	0.78
1:A:214:GLU:O	1:A:218:ILE:HG12	1.86	0.76
1:B:112:LYS:HB3	1:B:116:GLU:HB3	1.68	0.76

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	219/221 (99%)	174 (80%)	34 (16%)	11 (5%)	2	2
1	B	219/221 (99%)	182 (83%)	27 (12%)	10 (5%)	2	3
All	All	438/442 (99%)	356 (81%)	61 (14%)	21 (5%)	2	2

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	68	THR
1	A	112	LYS
1	B	13	ARG
1	B	68	THR
1	B	80	ASN

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	192/192 (100%)	174 (91%)	18 (9%)	8	17
1	B	192/192 (100%)	167 (87%)	25 (13%)	4	7
All	All	384/384 (100%)	341 (89%)	43 (11%)	6	10

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	100	GLN
1	B	140	LEU
1	B	101	ASP
1	B	117	LYS
1	B	155	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	100	GLN
1	B	151	ASN
1	B	159	GLN
1	B	11	ASN
1	B	49	HIS

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5   Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

2 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	$\#|Z| > 2$
2	HAG	A	223	-	30,31,31	2.07	9 (30%)	29,40,40	2.80	14 (48%)
3	GSH	B	223	-	18,19,19	1.90	6 (33%)	23,24,24	3.28	11 (47%)

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
2	HAG	A	223	-	-	10/32/44/44	0/1/1/1
3	GSH	B	223	-	-	8/24/24/24	-

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	223	HAG	O12-C1	6.33	1.41	1.22
2	A	223	HAG	O42-C43	3.85	1.49	1.39
2	A	223	HAG	O32-C3	3.79	1.34	1.22
3	B	223	GSH	CB1-CA1	3.50	1.60	1.53
2	A	223	HAG	CB2-SG2	-3.38	1.78	1.82

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	223	GSH	CA2-CB2-SG2	-9.32	103.72	114.19
2	A	223	HAG	O11-C1-O12	-6.97	108.26	124.09
3	B	223	GSH	O12-C1-O11	-6.61	109.09	124.09
3	B	223	GSH	O2-C2-N3	6.31	136.53	122.99
2	A	223	HAG	O2-C2-N3	6.29	136.50	122.99

There are no chirality outliers.
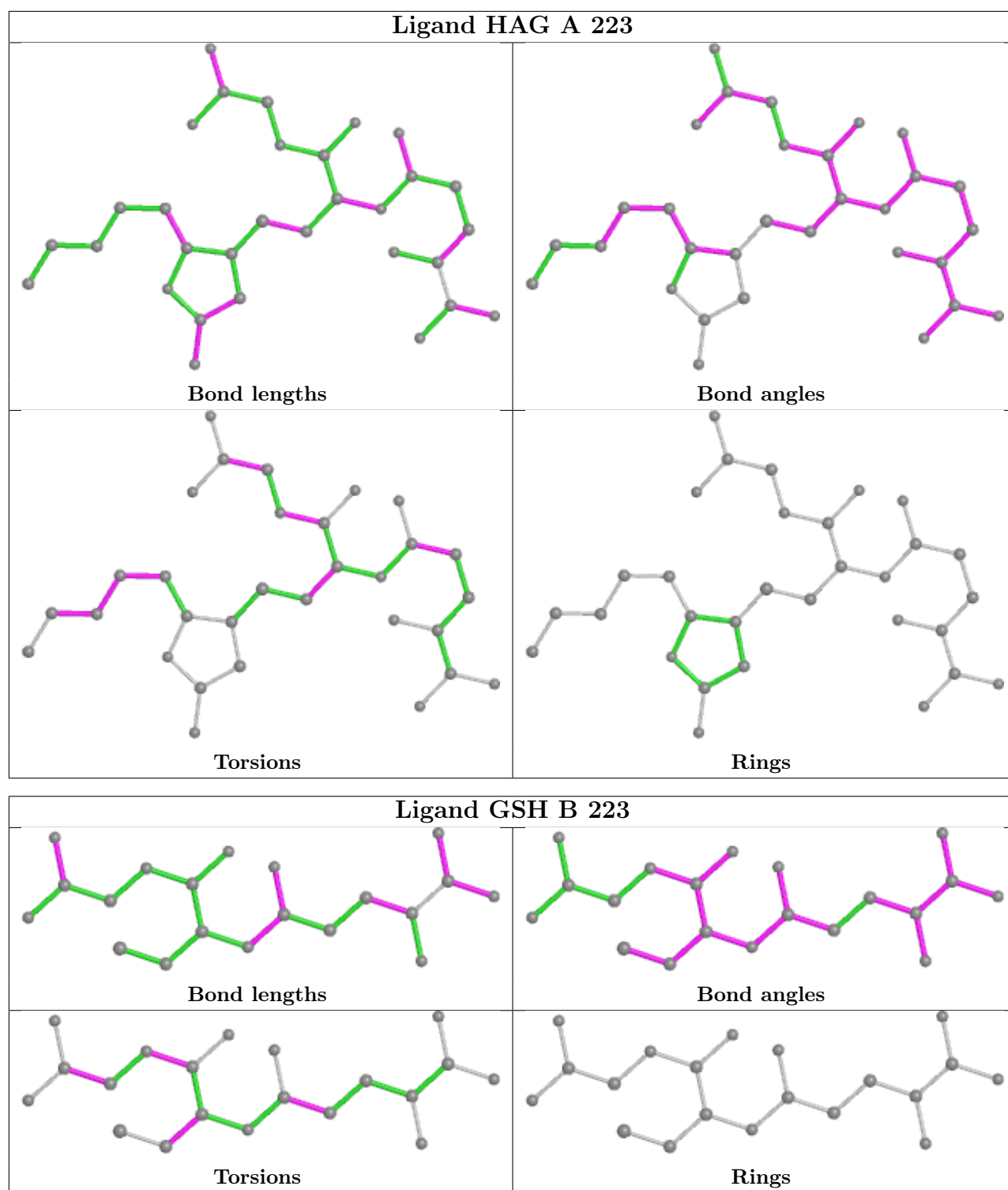
5 of 18 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	223	HAG	CA2-C2-N3-CA3
2	A	223	HAG	O2-C2-N3-CA3
2	A	223	HAG	N2-CA2-CB2-SG2
2	A	223	HAG	C44-C45-C46-C47
3	B	223	GSH	N2-CA2-CB2-SG2

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	223	HAG	2	0
3	B	223	GSH	2	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 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 HAG A 223

Bond lengths

Bond angles

Torsions

Rings

Ligand GSH B 223

Bond lengths

Bond angles

Torsions

Rings

5.7 Other polymers ⓘ

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, 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(Å²)	Q<0.9
1	A	221/221 (100%)	-0.42	1 (0%)	91	89	15, 51, 73, 83	0
1	B	221/221 (100%)	-0.51	0	100	100	14, 48, 69, 76	0
All	All	442/442 (100%)	-0.47	1 (0%)	95	95	14, 50, 72, 83	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	115	LYS	2.6

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

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

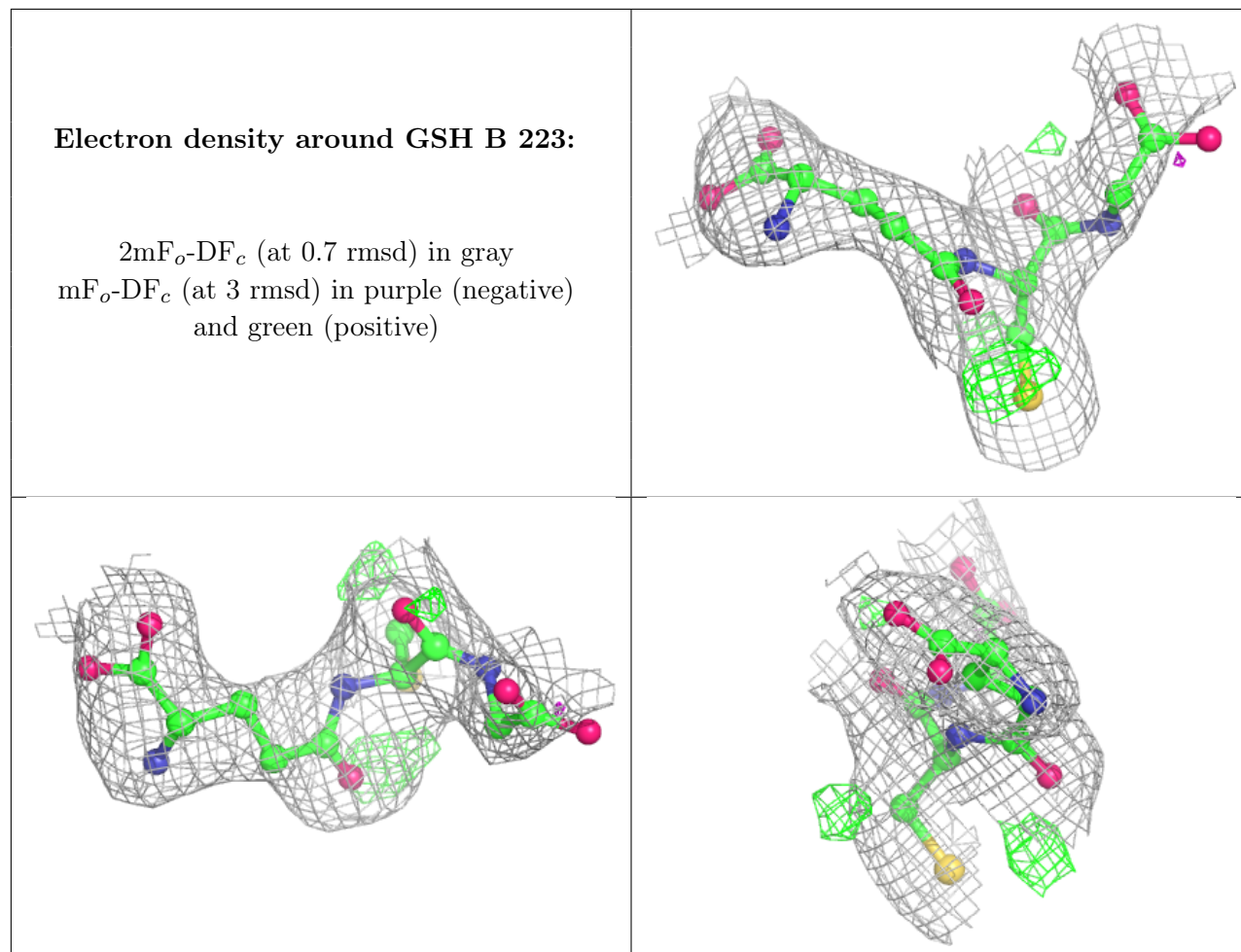## 6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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.
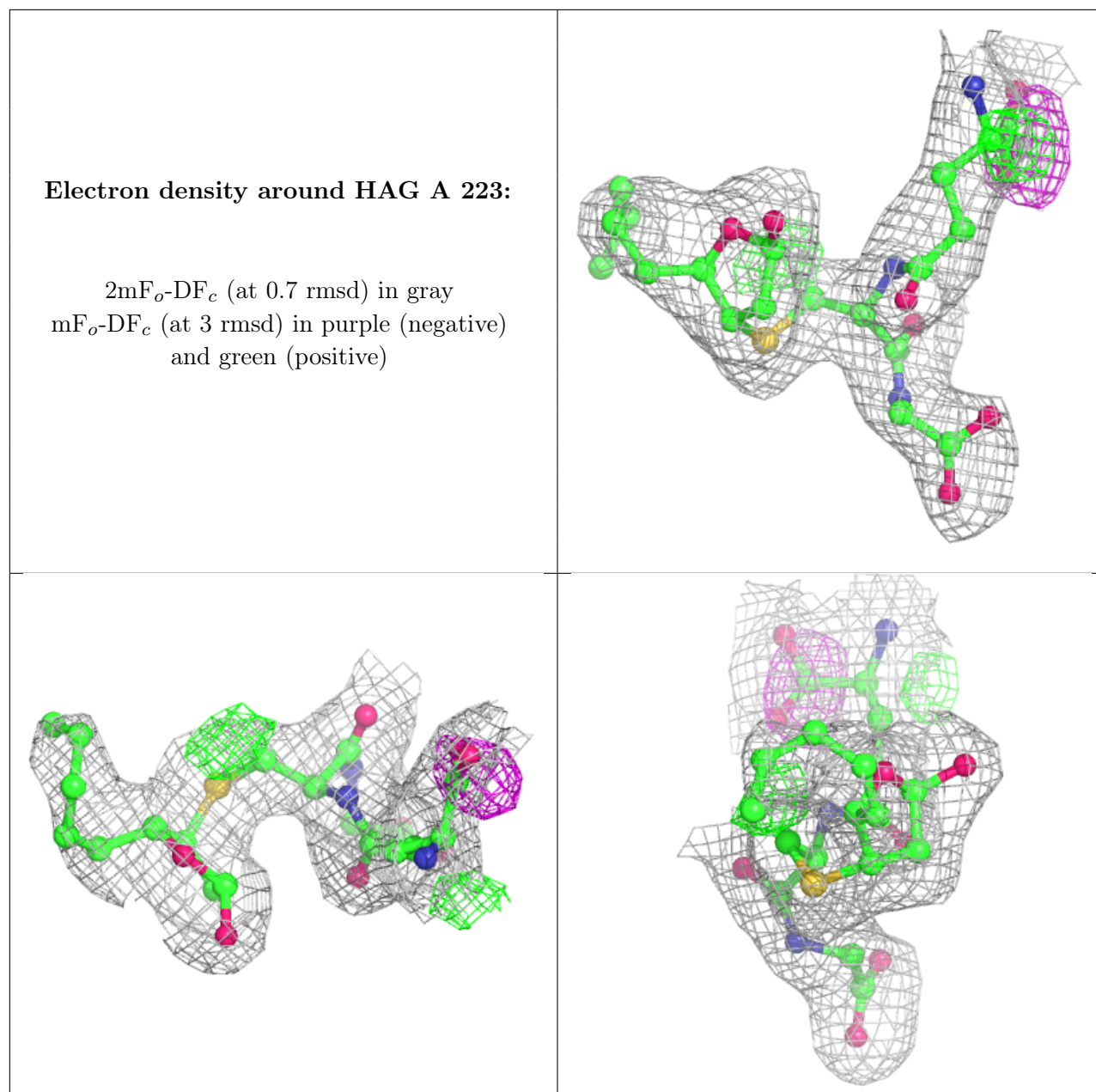
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å²)	Q<0.9
3	GSH	B	223	20/20	0.87	0.15	39,50,74,75	0
2	HAG	A	223	31/31	0.91	0.16	26,49,60,61	0

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 GSH B 223:

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

Electron density around HAG A 223:

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

6.5 Other polymers (i)

There are no such residues in this entry.