



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 14, 2020 – 03:10 pm GMT

PDB ID : 6TXK  
Title : Crystal structure of thermotoga maritima E65K Ferritin  
Deposited on : 2020-01-14  
Resolution : 2.36 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

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 : 1.8.0 (224370), CSD as540be (2019)  
Xtriage (Phenix) : 1.13  
EDS : 2.7  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
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) : 2.7

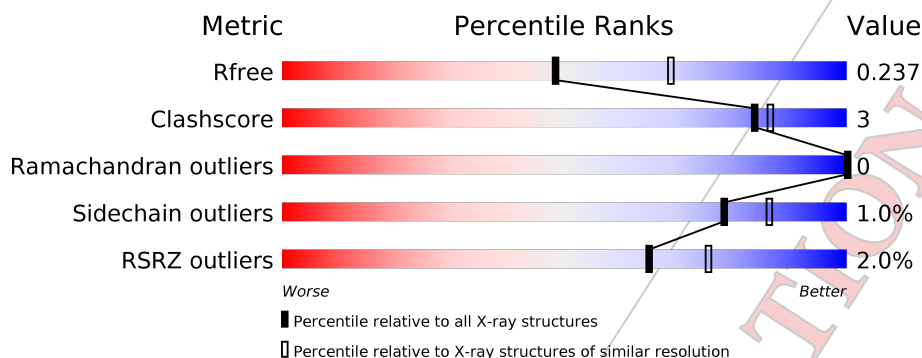
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.36 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))
$R_{free}$	111664	1015 (2.36-2.36)
Clashscore	122126	1081 (2.36-2.36)
Ramachandran outliers	120053	1066 (2.36-2.36)
Sidechain outliers	120020	1067 (2.36-2.36)
RSRZ outliers	108989	1002 (2.36-2.36)

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	164	<div> <div>2%</div> <div> <div></div> <div>91%</div> <div>9%</div> </div> </div>
1	B	164	<div> <div>2%</div> <div> <div></div> <div>98%</div> <div></div> </div> </div>
1	C	164	<div> <div>%</div> <div> <div></div> <div>90%</div> <div>10%</div> </div> </div>
1	D	164	<div> <div>%</div> <div> <div></div> <div>93%</div> <div>7%</div> </div> </div>
1	E	164	<div> <div>%</div> <div> <div></div> <div>95%</div> <div>5%</div> </div> </div>
1	F	164	<div> <div>%</div> <div> <div></div> <div>94%</div> <div>5%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	164	
1	H	164	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	SO4	D	206	-	-	-	X

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 11850 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 Ferritin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	164	Total	C	N	O	S	0	3	0
			1385	888	226	265	6			
1	B	164	Total	C	N	O	S	0	6	0
			1405	901	230	268	6			
1	C	164	Total	C	N	O	S	0	2	0
			1379	883	226	264	6			
1	D	164	Total	C	N	O	S	0	2	0
			1379	884	226	263	6			
1	E	164	Total	C	N	O	S	0	3	0
			1387	889	229	263	6			
1	F	164	Total	C	N	O	S	0	2	0
			1379	883	225	265	6			
1	G	164	Total	C	N	O	S	0	1	0
			1373	880	225	262	6			
1	H	164	Total	C	N	O	S	0	5	0
			1401	897	232	266	6			

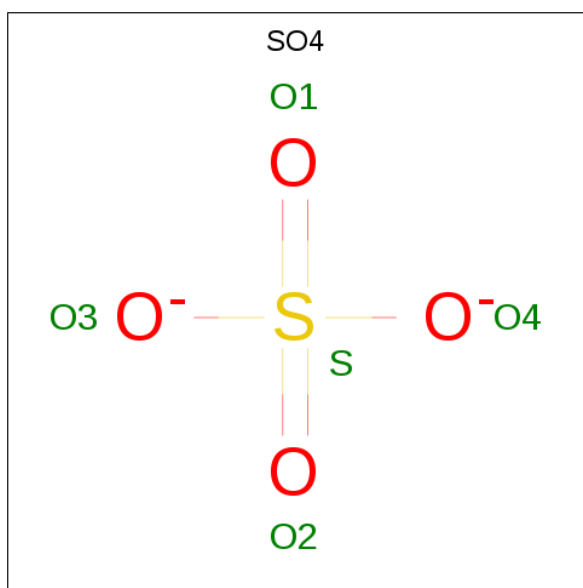
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	65	LYS	GLU	engineered mutation	UNP Q9X0L2
B	65	LYS	GLU	engineered mutation	UNP Q9X0L2
C	65	LYS	GLU	engineered mutation	UNP Q9X0L2
D	65	LYS	GLU	engineered mutation	UNP Q9X0L2
E	65	LYS	GLU	engineered mutation	UNP Q9X0L2
F	65	LYS	GLU	engineered mutation	UNP Q9X0L2
G	65	LYS	GLU	engineered mutation	UNP Q9X0L2
H	65	LYS	GLU	engineered mutation	UNP Q9X0L2

- Molecule 2 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	G	1	Total Fe 1 1	0	0
2	D	1	Total Fe 1 1	0	0
2	E	1	Total Fe 1 1	0	0
2	H	1	Total Fe 1 1	0	0
2	B	1	Total Fe 1 1	0	0
2	C	1	Total Fe 1 1	0	0
2	A	1	Total Fe 1 1	0	0
2	F	1	Total Fe 1 1	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	E	1	Total	O	S	0	0
			5	4	1		
3	E	1	Total	O	S	0	0
			5	4	1		
3	F	1	Total	O	S	0	0
			5	4	1		
3	F	1	Total	O	S	0	0
			5	4	1		
3	G	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	G	1	Total	O	S	0	0
			5	4	1		
3	G	1	Total	O	S	0	0
			5	4	1		
3	G	1	Total	O	S	0	0
			5	4	1		
3	H	1	Total	O	S	0	0
			5	4	1		
3	H	1	Total	O	S	0	0
			5	4	1		
3	H	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



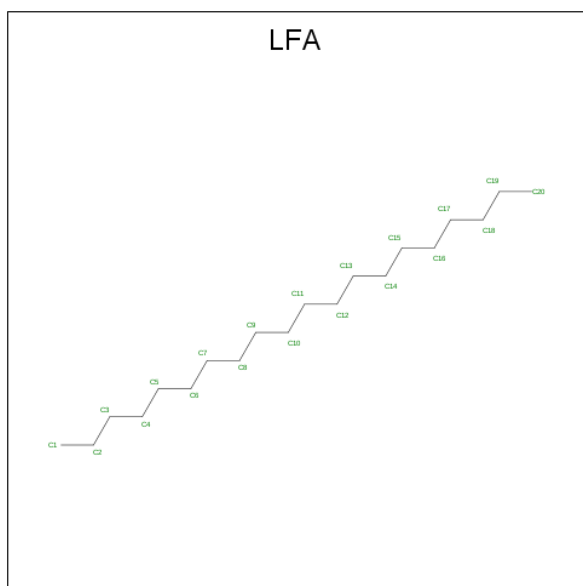
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total	C	O	0	0
			6	3	3		
4	C	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		
4	E	1	Total	C	O	0	0
			6	3	3		
4	F	1	Total	C	O	0	0
			6	3	3		
4	G	1	Total	C	O	0	0
			6	3	3		
4	H	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is EICOSANE (three-letter code: LFA) (formula:  $C_{20}H_{42}$ ).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	C	0	0
			20	20		
5	C	1	Total	C	0	0
			20	20		
5	E	1	Total	C	0	0
			20	20		

Continued on next page...



*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	H	1	Total C 20 20	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	67	Total O 67 67	0	0
6	B	71	Total O 71 71	0	0
6	C	53	Total O 53 53	0	0
6	D	64	Total O 64 64	0	0
6	E	54	Total O 54 54	0	0
6	F	54	Total O 54 54	0	0
6	G	38	Total O 38 38	0	0
6	H	41	Total O 41 41	0	0

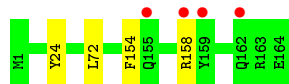
### 3 Residue-property plots [i](#)

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.

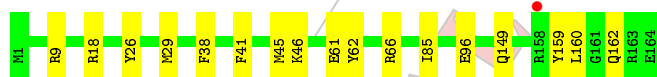
- Molecule 1: Ferritin



- Molecule 1: Ferritin



- Molecule 1: Ferritin



- Molecule 1: Ferritin

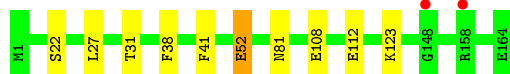


- Molecule 1: Ferritin




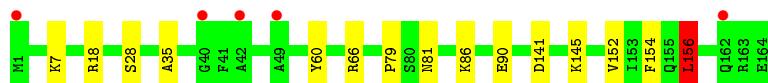
- Molecule 1: Ferritin

Chain F:  94% 5%




• Molecule 1: Ferritin

Chain G:  91% 9% 3%



• Molecule 1: Ferritin

Chain H:  90% 10% 5%



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	175.81Å 175.81Å 354.14Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.44 – 2.36 48.25 – 2.36	Depositor EDS
% Data completeness (in resolution range)	99.7 (20.44-2.36) 99.8 (48.25-2.36)	Depositor EDS
$R_{merge}$	0.23	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.19 (at 2.37Å)	Xtriage
Refinement program	PHENIX 1.15.1_3469	Depositor
R, $R_{free}$	0.201 , 0.236 0.201 , 0.237	Depositor DCC
$R_{free}$ test set	2100 reflections (2.43%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.2	Xtriage
Anisotropy	0.281	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 43.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	11850	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.69% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup> Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality (i)

### 5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, LFA, FE, SO4

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	0.25	0/1423	0.38	0/1910
1	B	0.25	0/1449	0.38	0/1944
1	C	0.26	0/1414	0.41	0/1899
1	D	0.27	0/1414	0.40	0/1898
1	E	0.25	0/1425	0.42	1/1912 (0.1%)
1	F	0.24	0/1414	0.36	0/1899
1	G	0.31	0/1405	0.44	1/1886 (0.1%)
1	H	0.34	0/1445	0.47	1/1939 (0.1%)
All	All	0.27	0/11389	0.41	3/15287 (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	H	0	2

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	87	ASP	CB-CG-OD1	6.01	123.71	118.30
1	H	150	MET	CB-CG-SD	5.79	129.76	112.40
1	G	156	LEU	CA-CB-CG	5.35	127.59	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	H	149	GLN	Peptide
1	H	163	ARG	Peptide

## 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1385	0	1353	9	0
1	B	1405	0	1379	2	0
1	C	1379	0	1342	12	0
1	D	1379	0	1347	7	0
1	E	1387	0	1360	5	0
1	F	1379	0	1340	7	0
1	G	1373	0	1339	12	0
1	H	1401	0	1374	11	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
3	A	25	0	0	1	0
3	B	30	0	0	0	0
3	C	15	0	0	0	0
3	D	30	0	0	0	0
3	E	10	0	0	0	0
3	F	10	0	0	0	0
3	G	20	0	0	0	0
3	H	20	0	0	1	0
4	A	6	0	8	0	0
4	B	18	0	24	0	0
4	C	12	0	16	0	0
4	D	12	0	16	0	0
4	E	6	0	8	0	0
4	F	6	0	8	0	0
4	G	6	0	8	1	0
4	H	6	0	8	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	20	0	42	0	0
5	C	20	0	42	1	0
5	E	20	0	42	0	0
5	H	20	0	42	0	0
6	A	67	0	0	1	0
6	B	71	0	0	0	0
6	C	53	0	0	2	0
6	D	64	0	0	2	0
6	E	54	0	0	1	0
6	F	54	0	0	1	0
6	G	38	0	0	2	0
6	H	41	0	0	0	0
All	All	11850	0	11098	59	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 3.

All (59) 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:H:5:SER:HB2	4:H:203:GOL:H12	1.69	0.72
1:G:141:ASP:O	1:G:145:LYS:HG3	1.92	0.69
1:D:69:ARG:NH2	6:D:301:HOH:O	2.22	0.69
1:H:159:TYR:O	1:H:162:GLN:HG3	1.97	0.64
1:F:108:GLU:O	1:F:112:GLU:HG3	1.99	0.63
1:E:50[A]:GLN:NE2	6:E:301:HOH:O	2.32	0.62
1:A:155:GLN:NE2	6:A:301:HOH:O	2.32	0.61
1:C:159:TYR:O	1:C:162:GLN:HB2	2.04	0.58
1:B:158:ARG:HG2	1:B:158:ARG:HH11	1.68	0.57
1:G:18:ARG:NH2	6:G:301:HOH:O	2.39	0.55
1:C:149:GLN:NE2	6:C:304:HOH:O	2.38	0.53
1:G:152:VAL:O	1:G:156:LEU:HD23	2.10	0.52
1:C:46:LYS:HB3	5:C:204:LFA:H191	1.92	0.51
1:H:159:TYR:HA	1:H:162:GLN:HG3	1.91	0.51
1:G:28:SER:OG	1:G:79:PRO:O	2.24	0.51
1:G:7:LYS:NZ	6:G:304:HOH:O	2.43	0.51
1:D:18:ARG:NH2	6:D:306:HOH:O	2.44	0.51
1:C:38:PHE:CZ	1:C:85:ILE:HD11	2.47	0.50
1:E:107:LEU:HD11	1:E:119:VAL:HG13	1.93	0.50
1:A:28:SER:HB2	1:B:72:LEU:HD12	1.94	0.49
1:D:69:ARG:HG2	1:D:70:VAL:N	2.27	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:112:GLU:OE1	6:F:301:HOH:O	2.19	0.49
1:D:123:LYS:NZ	1:D:127:ASP:OD2	2.42	0.48
1:A:29:MET:HG2	1:A:82:TRP:CH2	2.47	0.48
1:C:29:MET:HB2	1:C:45:MET:HE2	1.95	0.48
1:A:107:LEU:HD11	1:A:119:VAL:HG13	1.95	0.48
1:F:22:SER:HB2	1:F:52:GLU:HG2	1.95	0.48
1:G:86:LYS:O	1:G:90:GLU:HG3	2.14	0.47
1:A:80:SER:N	3:A:204:SO4:O3	2.38	0.47
1:A:26:TYR:OH	1:A:96:GLU:OE2	2.26	0.47
1:H:27:LEU:O	1:H:31:THR:HG23	2.15	0.46
1:G:18:ARG:HH11	4:G:202:GOL:H2	1.80	0.46
1:D:66:ARG:HA	1:D:66:ARG:HD2	1.83	0.46
1:F:38:PHE:HB3	1:F:41:PHE:HD2	1.81	0.46
1:C:38:PHE:CE2	1:C:85:ILE:HD11	2.50	0.46
1:A:66:ARG:HD2	1:A:66:ARG:HA	1.88	0.45
1:E:27:LEU:O	1:E:31:THR:HG23	2.16	0.45
1:E:69[B]:ARG:NH2	1:F:81:ASN:OD1	2.50	0.45
1:C:26:TYR:OH	1:C:96:GLU:OE2	2.26	0.44
1:A:27:LEU:O	1:A:31:THR:HG23	2.17	0.44
1:D:41:PHE:CD1	1:D:160:LEU:HD12	2.53	0.44
1:H:103:ILE:HD13	1:H:125:PHE:HB3	2.00	0.44
1:G:28:SER:HB2	1:H:72:LEU:HD12	1.99	0.44
1:G:81:ASN:OD1	1:H:69[A]:ARG:NH2	2.52	0.43
1:F:123:LYS:HD3	1:F:123:LYS:HA	1.86	0.42
1:E:66:ARG:HA	1:E:66:ARG:HD2	1.88	0.42
1:F:27:LEU:O	1:F:31:THR:HG23	2.19	0.42
1:G:60:TYR:HE2	1:H:31:THR:HG22	1.84	0.42
1:G:66:ARG:HA	1:G:66:ARG:HD2	1.86	0.42
1:C:66:ARG:HD2	1:C:66:ARG:HA	1.81	0.42
1:H:114:LYS:HE2	1:H:114:LYS:HB3	1.79	0.41
1:D:141:ASP:OD1	1:D:145:LYS:NZ	2.48	0.41
1:H:80:SER:N	3:H:205:SO4:O2	2.47	0.41
1:C:18:ARG:NH2	6:C:307:HOH:O	2.52	0.41
1:A:15:GLN:HB2	1:A:106:ILE:HD11	2.03	0.41
1:C:41:PHE:CD1	1:C:160:LEU:HD12	2.56	0.41
1:C:41:PHE:CE2	1:C:85:ILE:HD13	2.56	0.41
1:G:35:ALA:HB2	1:H:69[B]:ARG:HG3	2.03	0.41
1:C:61:GLU:HG3	1:C:62:TYR:N	2.35	0.40

There are no symmetry-related clashes.



## 5.3 Torsion angles

### 5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	165/164 (101%)	165 (100%)	0	0	100	100
1	B	168/164 (102%)	167 (99%)	1 (1%)	0	100	100
1	C	164/164 (100%)	163 (99%)	1 (1%)	0	100	100
1	D	164/164 (100%)	162 (99%)	2 (1%)	0	100	100
1	E	165/164 (101%)	165 (100%)	0	0	100	100
1	F	164/164 (100%)	164 (100%)	0	0	100	100
1	G	163/164 (99%)	162 (99%)	1 (1%)	0	100	100
1	H	167/164 (102%)	166 (99%)	1 (1%)	0	100	100
All	All	1320/1312 (101%)	1314 (100%)	6 (0%)	0	100	100

There are no Ramachandran outliers to report.

### 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	147/144 (102%)	147 (100%)	0	100	100
1	B	150/144 (104%)	148 (99%)	2 (1%)	71	81
1	C	146/144 (101%)	145 (99%)	1 (1%)	85	91
1	D	146/144 (101%)	145 (99%)	1 (1%)	85	91
1	E	147/144 (102%)	145 (99%)	2 (1%)	69	80
1	F	146/144 (101%)	145 (99%)	1 (1%)	85	91

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	145/144 (101%)	143 (99%)	2 (1%)	69	80
1	H	149/144 (104%)	147 (99%)	2 (1%)	71	81
All	All	1176/1152 (102%)	1165 (99%)	11 (1%)	78	88

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	24	TYR
1	B	154	PHE
1	C	9	ARG
1	D	9	ARG
1	E	1	MET
1	E	87	ASP
1	F	52	GLU
1	G	154	PHE
1	G	156	LEU
1	H	111	SER
1	H	123	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	55	HIS
1	B	162	GLN
1	D	83	ASN
1	E	149	GLN
1	E	155	GLN
1	G	83	ASN
1	G	135	GLN
1	G	162	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 56 ligands modelled in this entry, 8 are monoatomic - leaving 48 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	$\# Z  > 2$
3	SO4	A	202	-	4,4,4	0.15	0	6,6,6	0.06	0
4	GOL	A	203	-	5,5,5	0.90	0	5,5,5	1.00	0
3	SO4	A	204	-	4,4,4	0.17	0	6,6,6	0.05	0
3	SO4	A	205	-	4,4,4	0.16	0	6,6,6	0.05	0
3	SO4	A	206	-	4,4,4	0.15	0	6,6,6	0.07	0
3	SO4	A	207	-	4,4,4	0.17	0	6,6,6	0.05	0
3	SO4	B	202	-	4,4,4	0.16	0	6,6,6	0.05	0
3	SO4	B	203	-	4,4,4	0.16	0	6,6,6	0.06	0
4	GOL	B	204	-	5,5,5	0.90	0	5,5,5	0.98	0
3	SO4	B	205	-	4,4,4	0.17	0	6,6,6	0.09	0
4	GOL	B	206	-	5,5,5	0.90	0	5,5,5	0.98	0
4	GOL	B	207	-	5,5,5	0.88	0	5,5,5	1.00	0
5	LFA	B	208	-	19,19,19	0.08	0	18,18,18	0.20	0
3	SO4	B	209	-	4,4,4	0.17	0	6,6,6	0.06	0
3	SO4	B	210	-	4,4,4	0.16	0	6,6,6	0.05	0
3	SO4	B	211	-	4,4,4	0.16	0	6,6,6	0.06	0
4	GOL	C	202	-	5,5,5	0.90	0	5,5,5	0.99	0
4	GOL	C	203	-	5,5,5	0.93	0	5,5,5	0.99	0
5	LFA	C	204	-	19,19,19	0.08	0	18,18,18	0.14	0
3	SO4	C	205	-	4,4,4	0.16	0	6,6,6	0.05	0
3	SO4	C	206	-	4,4,4	0.16	0	6,6,6	0.05	0
3	SO4	C	207	-	4,4,4	0.17	0	6,6,6	0.06	0
3	SO4	D	202	-	4,4,4	0.16	0	6,6,6	0.06	0
4	GOL	D	203	-	5,5,5	0.96	0	5,5,5	0.86	0
4	GOL	D	204	-	5,5,5	0.93	0	5,5,5	0.98	0
3	SO4	D	205	-	4,4,4	0.15	0	6,6,6	0.05	0
3	SO4	D	206	-	4,4,4	0.15	0	6,6,6	0.06	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	SO4	D	207	-	4,4,4	0.17	0	6,6,6	0.04	0
3	SO4	D	208	-	4,4,4	0.16	0	6,6,6	0.06	0
3	SO4	D	209	-	4,4,4	0.16	0	6,6,6	0.06	0
3	SO4	E	202	-	4,4,4	0.17	0	6,6,6	0.05	0
4	GOL	E	203	-	5,5,5	0.92	0	5,5,5	1.00	0
5	LFA	E	204	-	19,19,19	0.08	0	18,18,18	0.10	0
3	SO4	E	205	-	4,4,4	0.16	0	6,6,6	0.05	0
3	SO4	F	202	-	4,4,4	0.16	0	6,6,6	0.05	0
4	GOL	F	203	-	5,5,5	0.92	0	5,5,5	0.97	0
3	SO4	F	204	-	4,4,4	0.16	0	6,6,6	0.06	0
4	GOL	G	202	-	5,5,5	0.88	0	5,5,5	1.01	0
3	SO4	G	203	-	4,4,4	0.16	0	6,6,6	0.04	0
3	SO4	G	204	-	4,4,4	0.17	0	6,6,6	0.05	0
3	SO4	G	205	-	4,4,4	0.17	0	6,6,6	0.04	0
3	SO4	G	206	-	4,4,4	0.16	0	6,6,6	0.05	0
3	SO4	H	202	-	4,4,4	0.15	0	6,6,6	0.06	0
4	GOL	H	203	-	5,5,5	0.92	0	5,5,5	0.99	0
5	LFA	H	204	-	19,19,19	0.08	0	18,18,18	0.09	0
3	SO4	H	205	-	4,4,4	0.16	0	6,6,6	0.05	0
3	SO4	H	206	-	4,4,4	0.16	0	6,6,6	0.05	0
3	SO4	H	207	-	4,4,4	0.14	0	6,6,6	0.07	0

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
4	GOL	A	203	-	-	4/4/4/4	-
4	GOL	B	204	-	-	0/4/4/4	-
4	GOL	B	206	-	-	2/4/4/4	-
4	GOL	B	207	-	-	0/4/4/4	-
5	LFA	B	208	-	-	2/17/17/17	-
4	GOL	C	202	-	-	4/4/4/4	-
4	GOL	C	203	-	-	2/4/4/4	-
5	LFA	C	204	-	-	1/17/17/17	-
4	GOL	D	203	-	-	2/4/4/4	-
4	GOL	D	204	-	-	0/4/4/4	-
4	GOL	E	203	-	-	2/4/4/4	-
5	LFA	E	204	-	-	1/17/17/17	-
4	GOL	F	203	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	G	202	-	-	2/4/4/4	-
4	GOL	H	203	-	-	4/4/4/4	-
5	LFA	H	204	-	-	2/17/17/17	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (30) torsion outliers are listed below:

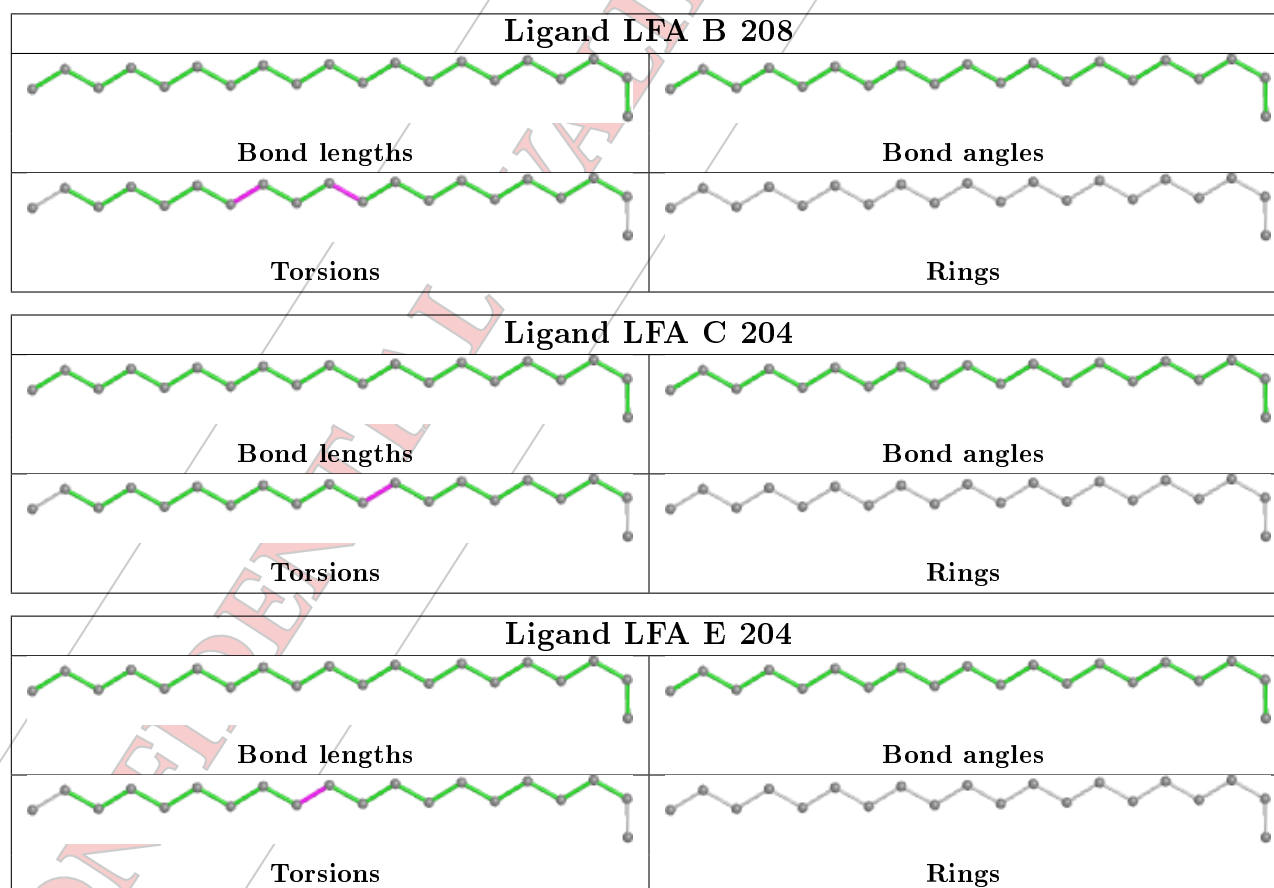
Mol	Chain	Res	Type	Atoms
4	H	203	GOL	C1-C2-C3-O3
4	G	202	GOL	O1-C1-C2-C3
4	A	203	GOL	C1-C2-C3-O3
4	C	203	GOL	C1-C2-C3-O3
4	D	203	GOL	C1-C2-C3-O3
4	C	202	GOL	C1-C2-C3-O3
4	C	203	GOL	O2-C2-C3-O3
4	C	202	GOL	O2-C2-C3-O3
4	A	203	GOL	O1-C1-C2-C3
4	B	206	GOL	C1-C2-C3-O3
4	F	203	GOL	C1-C2-C3-O3
5	H	204	LFA	C13-C14-C15-C16
5	E	204	LFA	C10-C11-C12-C13
4	G	202	GOL	O1-C1-C2-O2
4	D	203	GOL	O2-C2-C3-O3
5	C	204	LFA	C11-C10-C9-C8
5	H	204	LFA	C10-C11-C12-C13
5	B	208	LFA	C12-C13-C14-C15
5	B	208	LFA	C9-C10-C11-C12
4	H	203	GOL	O2-C2-C3-O3
4	A	203	GOL	O2-C2-C3-O3
4	F	203	GOL	O2-C2-C3-O3
4	C	202	GOL	O1-C1-C2-O2
4	H	203	GOL	O1-C1-C2-O2
4	A	203	GOL	O1-C1-C2-O2
4	B	206	GOL	O2-C2-C3-O3
4	E	203	GOL	O2-C2-C3-O3
4	H	203	GOL	O1-C1-C2-C3
4	E	203	GOL	C1-C2-C3-O3
4	C	202	GOL	O1-C1-C2-C3

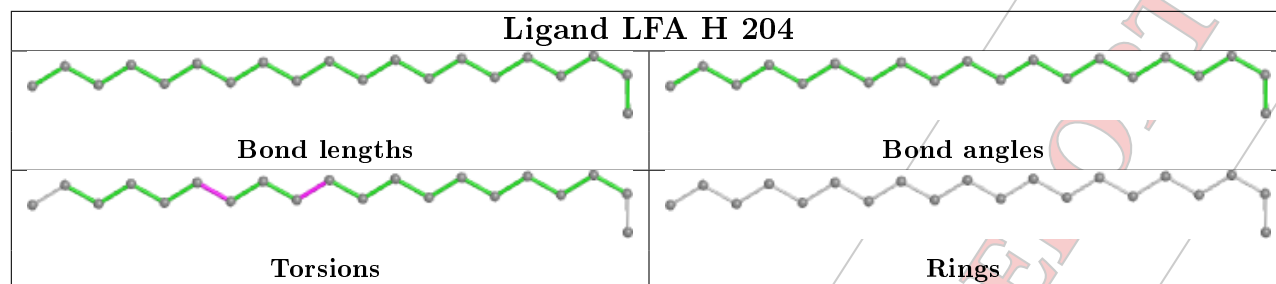
There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	204	LFA	1	0
4	G	202	GOL	1	0
4	H	203	GOL	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

### 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	164/164 (100%)	0.14	4 (2%) 59 69	33, 43, 59, 103	0
1	B	164/164 (100%)	0.15	4 (2%) 59 69	34, 44, 69, 90	0
1	C	164/164 (100%)	0.14	1 (0%) 89 94	38, 48, 71, 93	0
1	D	164/164 (100%)	0.07	1 (0%) 89 94	37, 46, 70, 88	0
1	E	164/164 (100%)	0.11	1 (0%) 89 94	35, 44, 62, 89	0
1	F	164/164 (100%)	0.14	2 (1%) 79 86	35, 46, 71, 85	0
1	G	164/164 (100%)	0.28	5 (3%) 50 62	41, 52, 73, 101	0
1	H	164/164 (100%)	0.39	8 (4%) 29 43	42, 54, 84, 115	0
All	All	1312/1312 (100%)	0.18	26 (1%) 65 75	33, 47, 73, 115	0

All (26) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	162	GLN	4.5
1	G	1	MET	4.3
1	H	150	MET	4.1
1	H	164	GLU	3.4
1	H	151	SER	3.3
1	B	158	ARG	3.3
1	G	162	GLN	3.2
1	F	158	ARG	3.1
1	B	162	GLN	3.1
1	G	40	GLY	3.1
1	H	61[A]	GLU	3.0
1	A	1	MET	2.9
1	A	98	PHE	2.8
1	F	148	GLY	2.6
1	C	158	ARG	2.6
1	G	42	ALA	2.5

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	H	158	ARG	2.4
1	H	6	GLU	2.4
1	A	155	GLN	2.3
1	A	6	GLU	2.3
1	B	159	TYR	2.2
1	G	49	ALA	2.2
1	H	159	TYR	2.1
1	D	83	ASN	2.0
1	E	164	GLU	2.0
1	B	155	GLN	2.0

## 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(Å <sup>2</sup> )	Q<0.9
3	SO4	D	206	5/5	0.56	0.46	75,84,88,89	5
3	SO4	D	209	5/5	0.58	0.24	69,70,76,80	5
3	SO4	A	206	5/5	0.66	0.39	85,88,93,93	5
3	SO4	H	202	5/5	0.68	0.24	107,108,111,116	0
4	GOL	B	206	6/6	0.69	0.24	65,70,74,76	0
4	GOL	D	203	6/6	0.74	0.23	68,74,85,89	0
4	GOL	F	203	6/6	0.75	0.19	71,72,75,81	0
3	SO4	A	202	5/5	0.76	0.31	85,90,92,93	5
3	SO4	E	202	5/5	0.76	0.21	97,100,102,106	0
4	GOL	H	203	6/6	0.79	0.31	78,80,82,84	0
2	FE	A	201	1/1	0.80	0.30	94,94,94,94	1
3	SO4	B	210	5/5	0.80	0.86	82,82,82,82	5
4	GOL	B	207	6/6	0.81	0.18	68,75,76,79	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	SO4	C	207	5/5	0.81	0.38	76,77,78,79	5
3	SO4	G	203	5/5	0.82	0.28	104,107,112,112	5
4	GOL	G	202	6/6	0.83	0.20	76,81,86,89	0
3	SO4	G	204	5/5	0.83	0.31	74,75,77,77	5
3	SO4	C	205	5/5	0.83	0.24	100,100,106,106	5
3	SO4	A	207	5/5	0.83	0.22	97,99,101,106	5
3	SO4	B	203	5/5	0.84	0.17	91,92,96,98	0
3	SO4	B	211	5/5	0.84	0.18	142,142,143,144	0
3	SO4	H	207	5/5	0.84	0.14	69,75,77,77	5
3	SO4	E	205	5/5	0.84	0.25	91,93,96,98	5
3	SO4	G	205	5/5	0.85	0.24	71,80,86,89	5
3	SO4	A	205	5/5	0.85	0.31	61,67,74,74	5
3	SO4	D	202	5/5	0.85	0.22	109,109,110,111	0
3	SO4	H	206	5/5	0.86	0.19	122,123,124,124	5
4	GOL	A	203	6/6	0.86	0.15	65,74,75,82	0
2	FE	F	201	1/1	0.86	0.10	49,49,49,49	1
2	FE	G	201	1/1	0.87	0.14	58,58,58,58	1
5	LFA	C	204	20/20	0.87	0.27	46,53,66,67	0
4	GOL	C	202	6/6	0.87	0.23	63,75,81,82	0
3	SO4	H	205	5/5	0.88	0.22	86,86,87,88	5
5	LFA	H	204	20/20	0.88	0.28	49,62,86,87	0
2	FE	B	201	1/1	0.89	0.12	54,54,54,54	1
3	SO4	D	207	5/5	0.89	0.17	76,81,82,84	5
3	SO4	F	204	5/5	0.89	0.33	65,70,74,75	5
5	LFA	E	204	20/20	0.89	0.34	49,55,77,79	0
3	SO4	G	206	5/5	0.90	0.14	109,110,111,113	0
3	SO4	F	202	5/5	0.91	0.17	90,95,97,98	0
3	SO4	D	208	5/5	0.91	0.20	80,82,83,83	5
3	SO4	B	202	5/5	0.91	0.20	59,68,69,69	5
3	SO4	B	209	5/5	0.91	0.27	77,78,80,80	5
2	FE	C	201	1/1	0.92	0.17	41,41,41,41	1
4	GOL	C	203	6/6	0.92	0.19	58,65,69,69	0
3	SO4	D	205	5/5	0.92	0.21	74,79,82,82	5
4	GOL	D	204	6/6	0.93	0.16	44,55,57,60	0
2	FE	D	201	1/1	0.94	0.08	51,51,51,51	1
3	SO4	C	206	5/5	0.94	0.20	75,76,79,80	5
3	SO4	A	204	5/5	0.94	0.18	67,67,69,72	5
5	LFA	B	208	20/20	0.94	0.26	38,46,73,77	0
4	GOL	E	203	6/6	0.94	0.14	63,67,68,68	0
3	SO4	B	205	5/5	0.96	0.18	53,55,58,59	5
4	GOL	B	204	6/6	0.96	0.14	54,55,59,60	0
2	FE	E	201	1/1	0.96	0.14	57,57,57,57	1

Continued on next page...

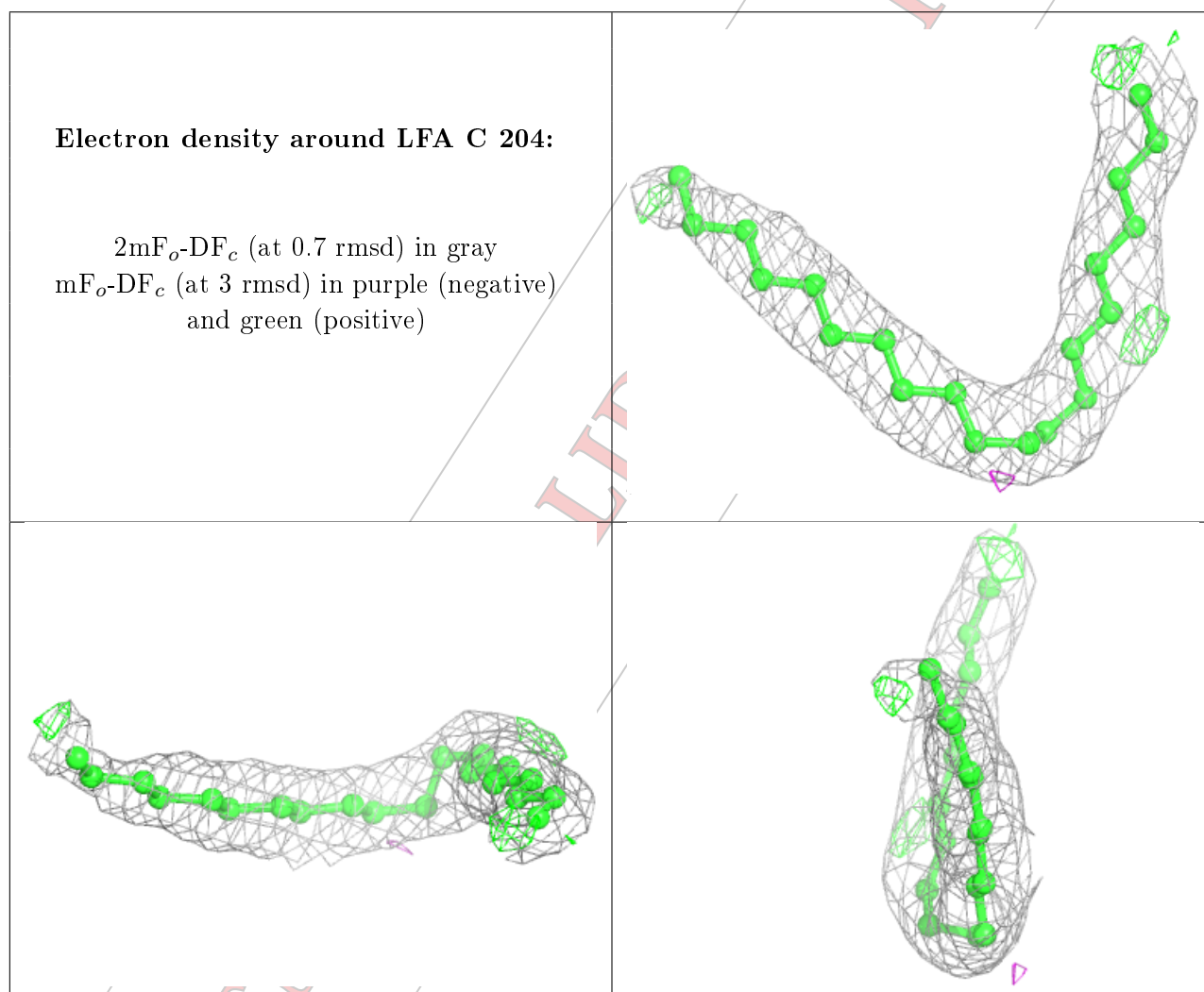
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	FE	H	201	1/1	0.99	0.13	62,62,62,62	1

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.

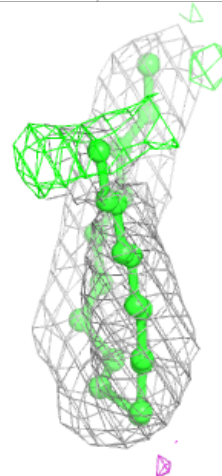
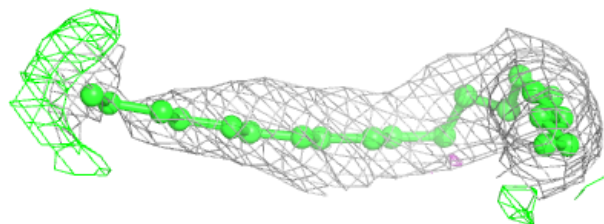
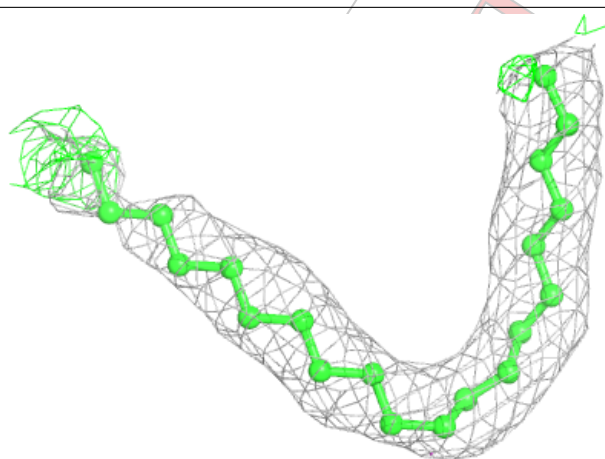
**Electron density around LFA C 204:**

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



**Electron density around LFA H 204:**

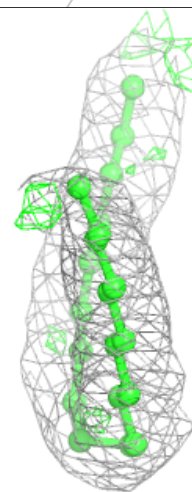
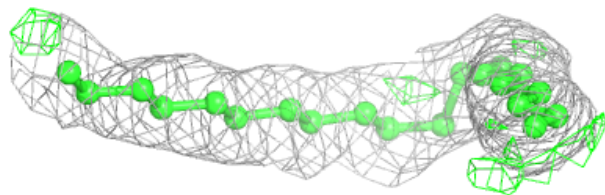
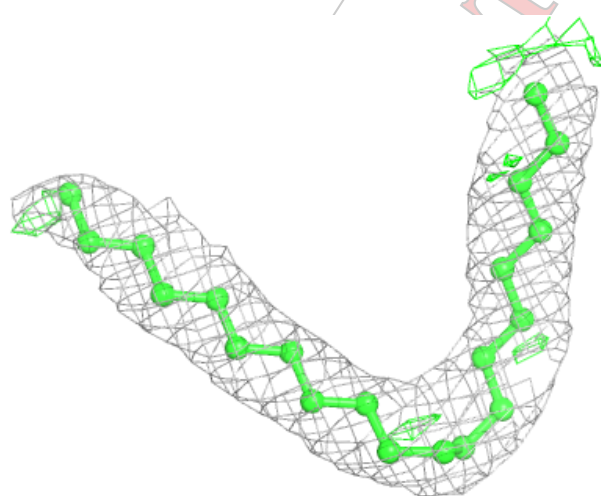
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



CONFIDENTIAL

**Electron density around LFA E 204:**

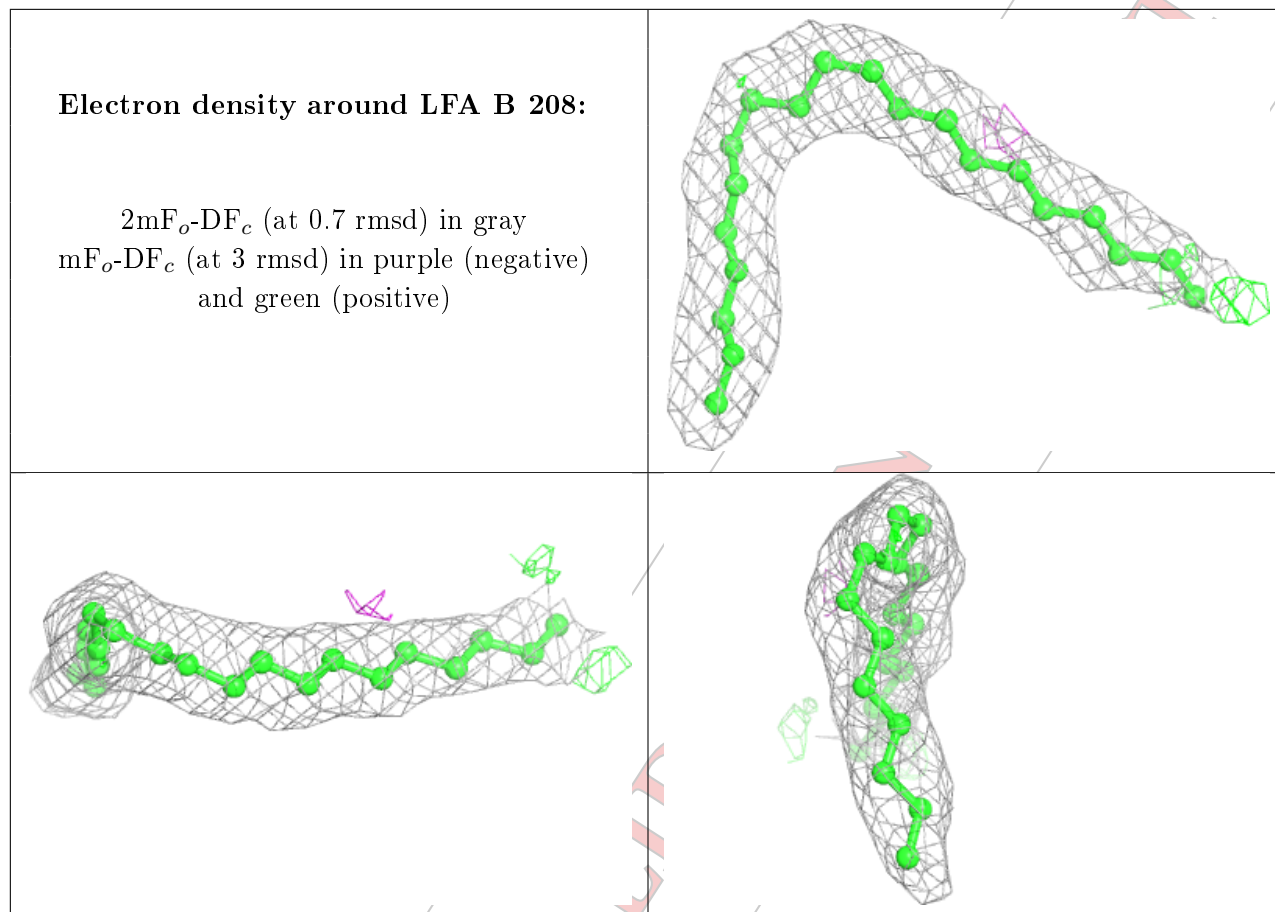
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



CONFIDENTIAL

**Electron density around LFA B 208:**

$2mF_o - DF_c$  (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.

CONFIDENTIAL