



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 13, 2017 – 12:15 AM EST

PDB ID : 1XFW  
Title : Crystal structure of anthrax edema factor (EF) in complex with calmodulin and 3'5' cyclic AMP (cAMP)  
Authors : Shen, Y.; Zhukovskaya, N.L.; Guo, Q.; Florian, J.; Tang, W.J.  
Deposited on : unknown  
Resolution : 3.40 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

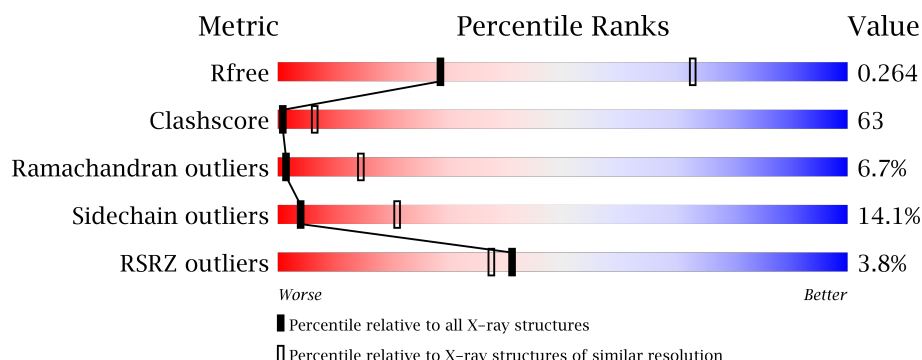
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.40 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))
$R_{free}$	100719	1679 (3.50-3.30)
Clashscore	112137	1832 (3.50-3.30)
Ramachandran outliers	110173	1789 (3.50-3.30)
Sidechain outliers	110143	1789 (3.50-3.30)
RSRZ outliers	101464	1709 (3.50-3.30)

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. 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	777	<div> <div>4%</div> <div>27% 55% 12% • 5%</div> </div>
1	B	777	<div> <div>4%</div> <div>27% 54% 12% • 5%</div> </div>
1	C	777	<div> <div>4%</div> <div>27% 54% 12% • 5%</div> </div>
1	D	777	<div> <div>4%</div> <div>28% 54% 12% • 5%</div> </div>
1	E	777	<div> <div>3%</div> <div>27% 54% 12% • 5%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	777	
2	O	149	
2	P	149	
2	Q	149	
2	R	149	
2	S	149	
2	T	149	

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
4	CMP	B	902	-	-	-	X
4	CMP	C	903	-	-	-	X
4	CMP	E	905	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 42990 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 Calmodulin-sensitive adenylate cyclase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	735	Total	C	N	O	S	0	0	0
			5992	3828	995	1163	6			
1	B	735	Total	C	N	O	S	0	0	0
			5992	3828	995	1163	6			
1	C	735	Total	C	N	O	S	0	0	0
			5992	3828	995	1163	6			
1	D	735	Total	C	N	O	S	0	0	0
			5992	3828	995	1163	6			
1	E	735	Total	C	N	O	S	0	0	0
			5992	3828	995	1163	6			
1	F	735	Total	C	N	O	S	0	0	0
			5992	3828	995	1163	6			

There are 54 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	24	MET	-	INITIATING METHIONINE	UNP P40136
A	25	HIS	-	EXPRESSION TAG	UNP P40136
A	26	HIS	-	EXPRESSION TAG	UNP P40136
A	27	HIS	-	EXPRESSION TAG	UNP P40136
A	28	HIS	-	EXPRESSION TAG	UNP P40136
A	29	HIS	-	EXPRESSION TAG	UNP P40136
A	30	HIS	-	EXPRESSION TAG	UNP P40136
A	31	ALA	-	CLONING ARTIFACT	UNP P40136
A	32	ALA	-	CLONING ARTIFACT	UNP P40136
B	24	MET	-	INITIATING METHIONINE	UNP P40136
B	25	HIS	-	EXPRESSION TAG	UNP P40136
B	26	HIS	-	EXPRESSION TAG	UNP P40136
B	27	HIS	-	EXPRESSION TAG	UNP P40136
B	28	HIS	-	EXPRESSION TAG	UNP P40136
B	29	HIS	-	EXPRESSION TAG	UNP P40136
B	30	HIS	-	EXPRESSION TAG	UNP P40136
B	31	ALA	-	CLONING ARTIFACT	UNP P40136

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	32	ALA	-	CLONING ARTIFACT	UNP P40136
C	24	MET	-	INITIATING METHIONINE	UNP P40136
C	25	HIS	-	EXPRESSION TAG	UNP P40136
C	26	HIS	-	EXPRESSION TAG	UNP P40136
C	27	HIS	-	EXPRESSION TAG	UNP P40136
C	28	HIS	-	EXPRESSION TAG	UNP P40136
C	29	HIS	-	EXPRESSION TAG	UNP P40136
C	30	HIS	-	EXPRESSION TAG	UNP P40136
C	31	ALA	-	CLONING ARTIFACT	UNP P40136
C	32	ALA	-	CLONING ARTIFACT	UNP P40136
D	24	MET	-	INITIATING METHIONINE	UNP P40136
D	25	HIS	-	EXPRESSION TAG	UNP P40136
D	26	HIS	-	EXPRESSION TAG	UNP P40136
D	27	HIS	-	EXPRESSION TAG	UNP P40136
D	28	HIS	-	EXPRESSION TAG	UNP P40136
D	29	HIS	-	EXPRESSION TAG	UNP P40136
D	30	HIS	-	EXPRESSION TAG	UNP P40136
D	31	ALA	-	CLONING ARTIFACT	UNP P40136
D	32	ALA	-	CLONING ARTIFACT	UNP P40136
E	24	MET	-	INITIATING METHIONINE	UNP P40136
E	25	HIS	-	EXPRESSION TAG	UNP P40136
E	26	HIS	-	EXPRESSION TAG	UNP P40136
E	27	HIS	-	EXPRESSION TAG	UNP P40136
E	28	HIS	-	EXPRESSION TAG	UNP P40136
E	29	HIS	-	EXPRESSION TAG	UNP P40136
E	30	HIS	-	EXPRESSION TAG	UNP P40136
E	31	ALA	-	CLONING ARTIFACT	UNP P40136
E	32	ALA	-	CLONING ARTIFACT	UNP P40136
F	24	MET	-	INITIATING METHIONINE	UNP P40136
F	25	HIS	-	EXPRESSION TAG	UNP P40136
F	26	HIS	-	EXPRESSION TAG	UNP P40136
F	27	HIS	-	EXPRESSION TAG	UNP P40136
F	28	HIS	-	EXPRESSION TAG	UNP P40136
F	29	HIS	-	EXPRESSION TAG	UNP P40136
F	30	HIS	-	EXPRESSION TAG	UNP P40136
F	31	ALA	-	CLONING ARTIFACT	UNP P40136
F	32	ALA	-	CLONING ARTIFACT	UNP P40136

- Molecule 2 is a protein called Calmodulin 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	O	146	Total	C	N	O	S	0	0	0
			1146	702	186	249	9			

*Continued on next page...*

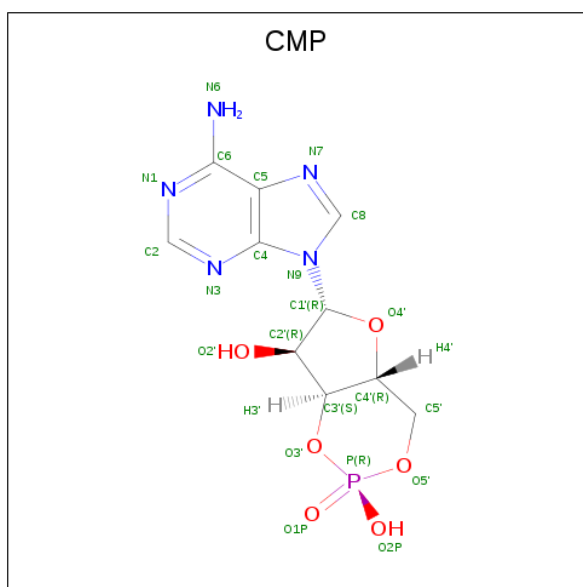
*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	P	146	Total	C	N	O	S	0	0	0
			1146	702	186	249	9			
2	Q	146	Total	C	N	O	S	0	0	0
			1146	702	186	249	9			
2	R	146	Total	C	N	O	S	0	0	0
			1146	702	186	249	9			
2	S	146	Total	C	N	O	S	0	0	0
			1146	702	186	249	9			
2	T	146	Total	C	N	O	S	0	0	0
			1146	702	186	249	9			

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	1	Total	Mg	0	0
			1	1		
3	E	1	Total	Mg	0	0
			1	1		
3	B	1	Total	Mg	0	0
			1	1		
3	C	1	Total	Mg	0	0
			1	1		
3	A	1	Total	Mg	0	0
			1	1		
3	F	1	Total	Mg	0	0
			1	1		

- Molecule 4 is ADENOSINE-3',5'-CYCLIC-MONOPHOSPHATE (three-letter code: CMP) (formula: C<sub>10</sub>H<sub>12</sub>N<sub>5</sub>O<sub>6</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
4	B	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
4	C	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
4	D	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
4	E	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
4	F	1	Total	C	N	O	P	0	0
			22	10	5	6	1		

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	P	3	Total	Ca	0	0
			3	3		
5	Q	3	Total	Ca	0	0
			3	3		
5	T	3	Total	Ca	0	0
			3	3		
5	O	3	Total	Ca	0	0
			3	3		
5	R	3	Total	Ca	0	0
			3	3		
5	S	3	Total	Ca	0	0
			3	3		

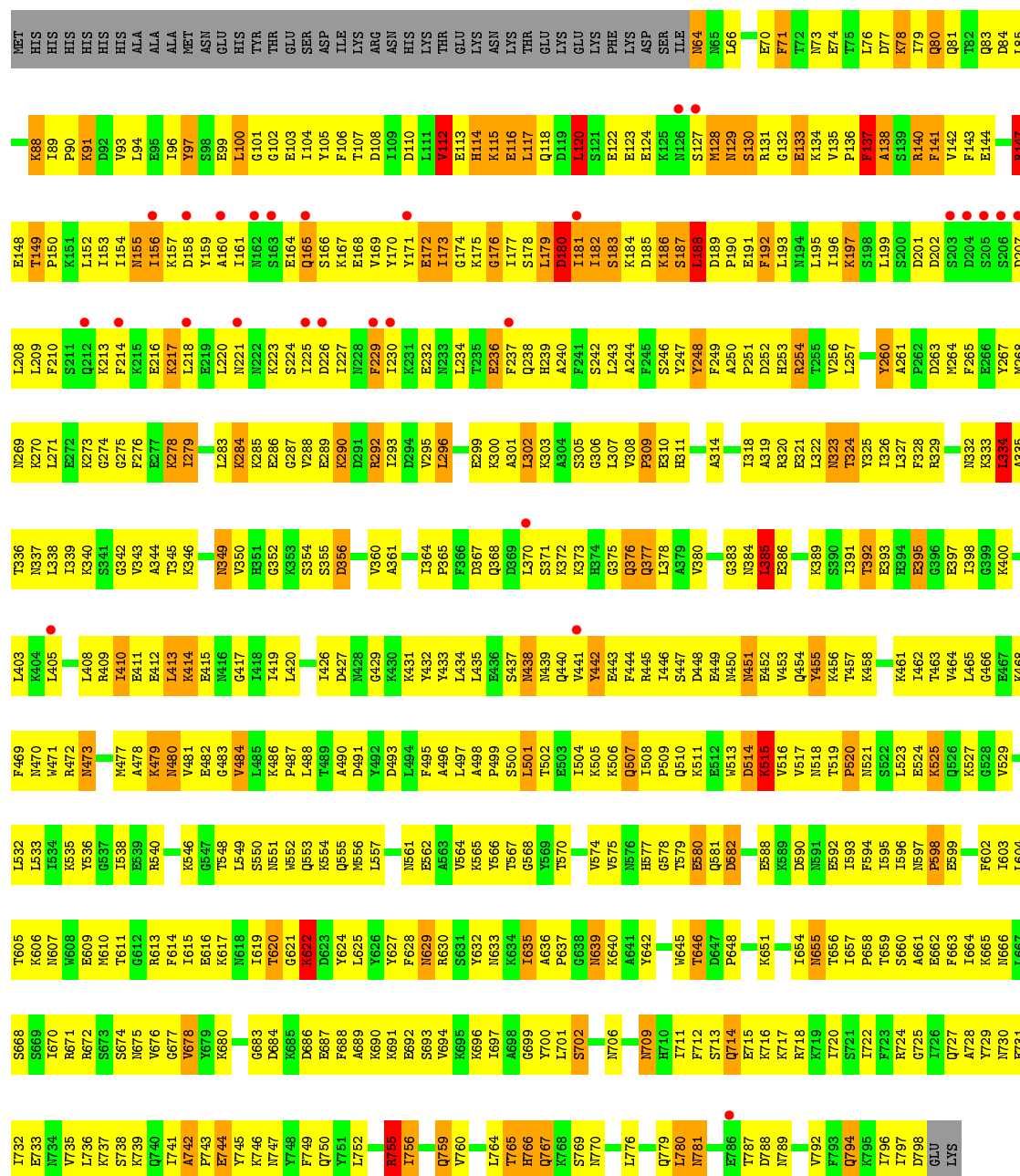
- Molecule 6 is water.

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

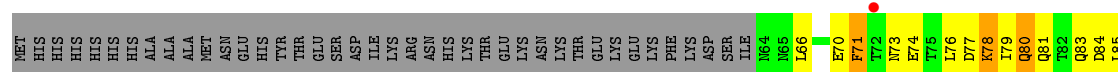


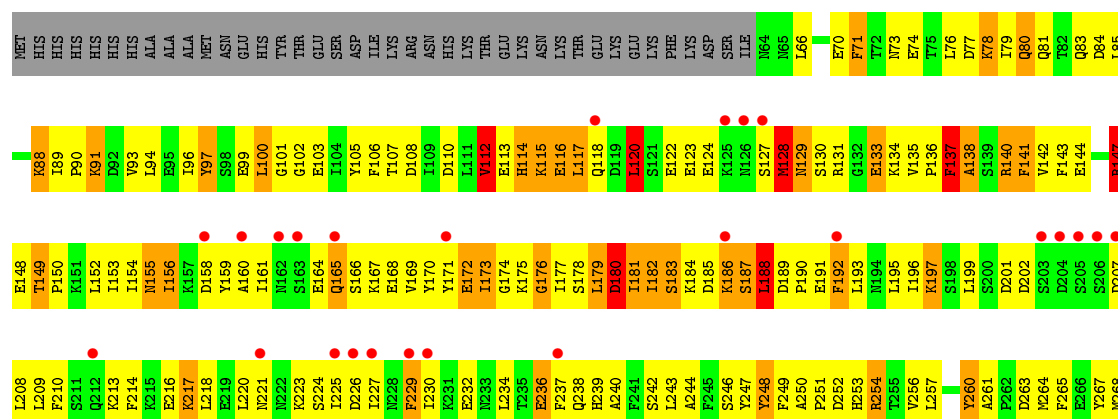
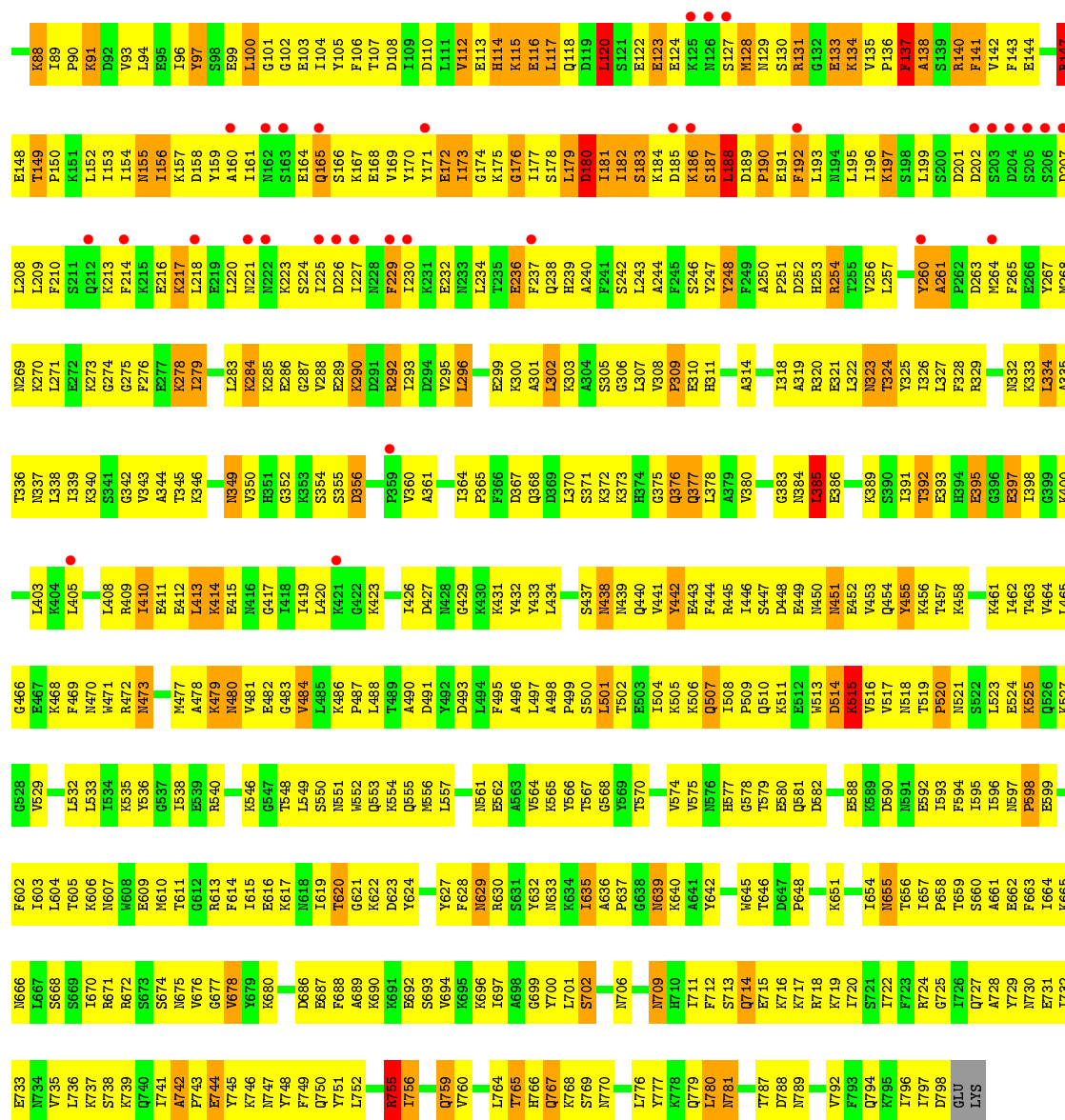


• Molecule 1: Calmodulin-sensitive adenylate cyclase

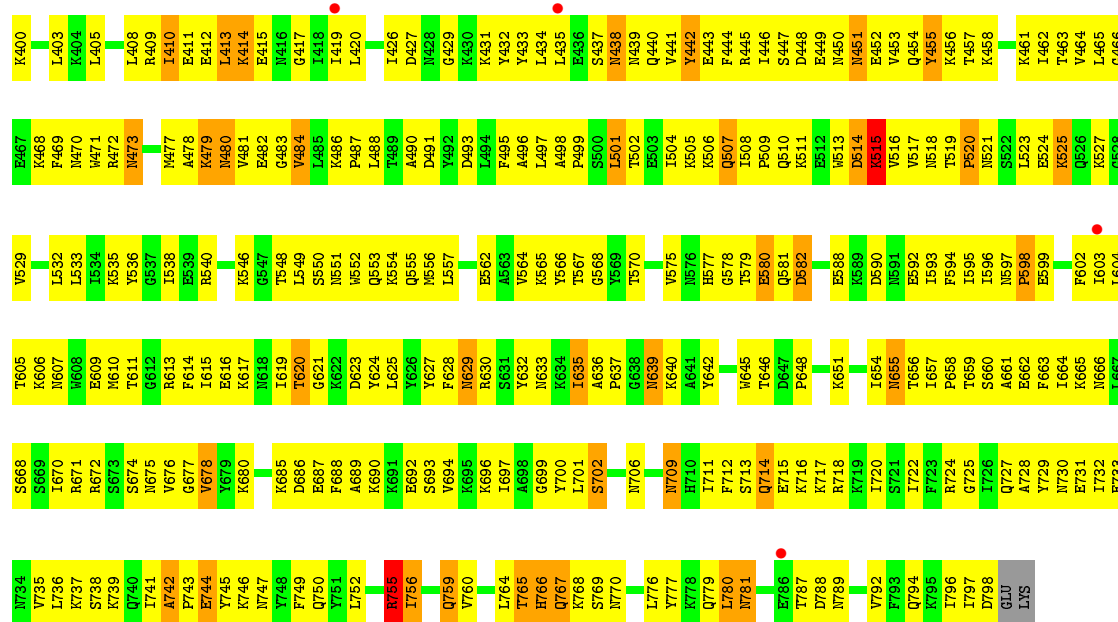


• Molecule 1: Calmodulin-sensitive adenylate cyclase

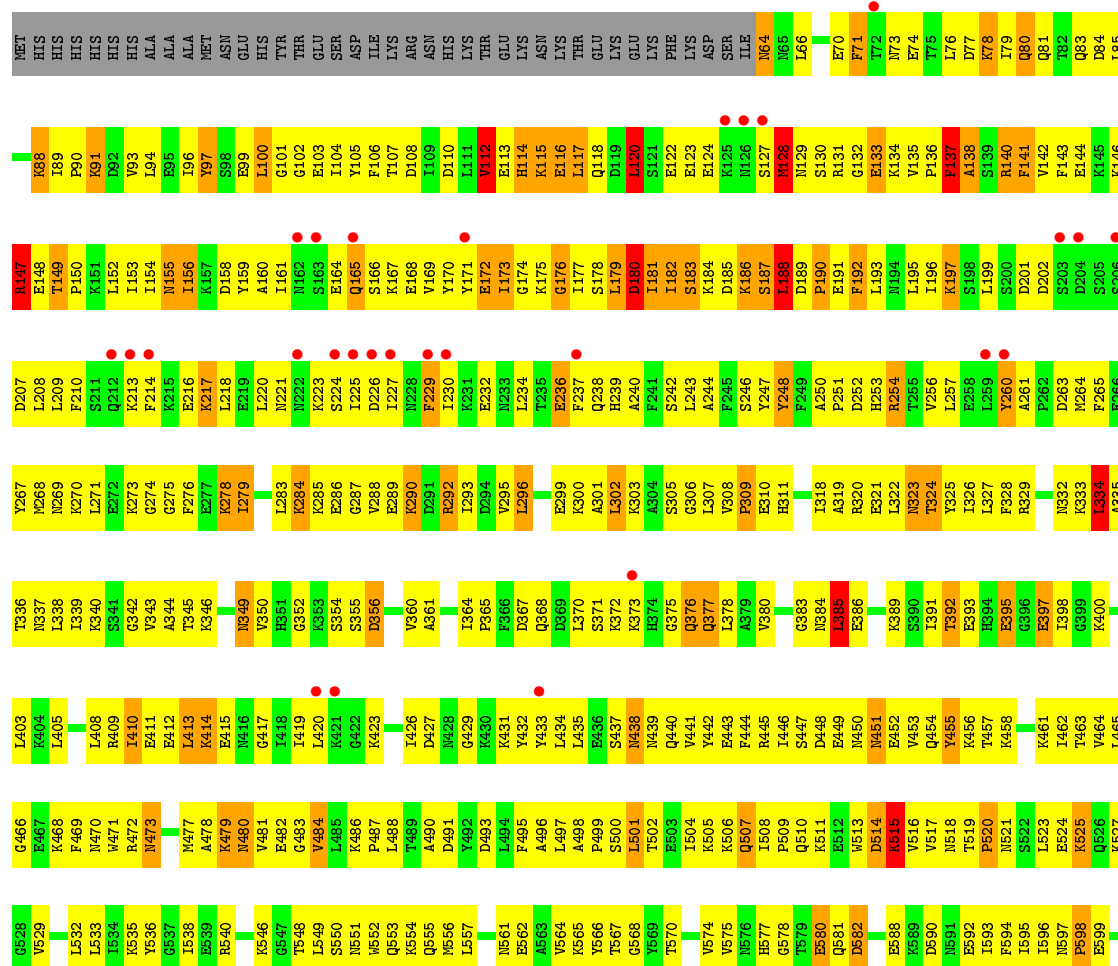


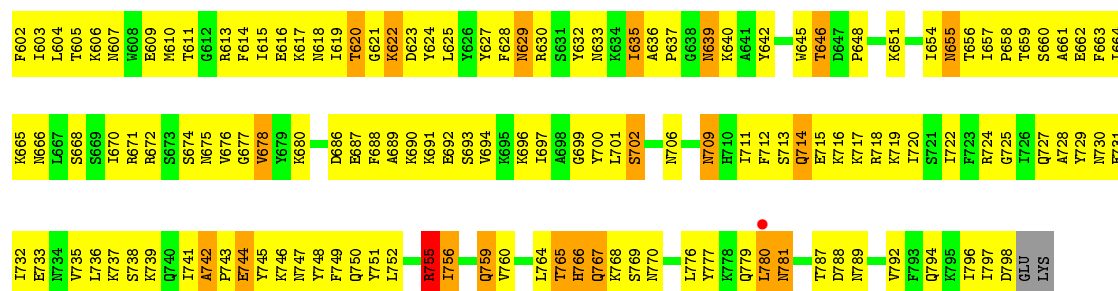




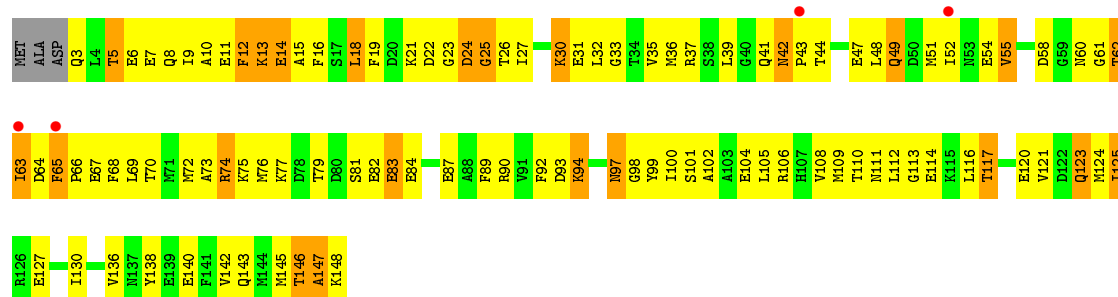


• Molecule 1: Calmodulin-sensitive adenylate cyclase





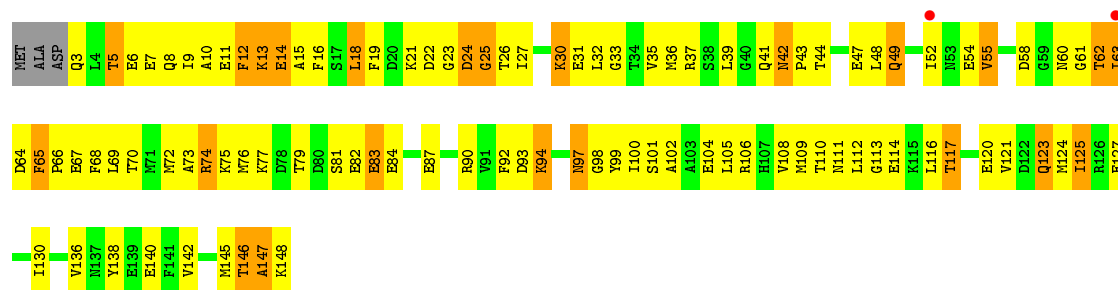
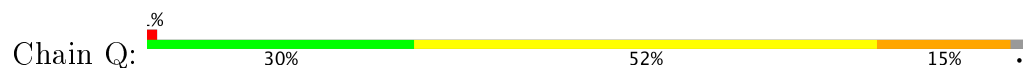
• Molecule 2: Calmodulin 2



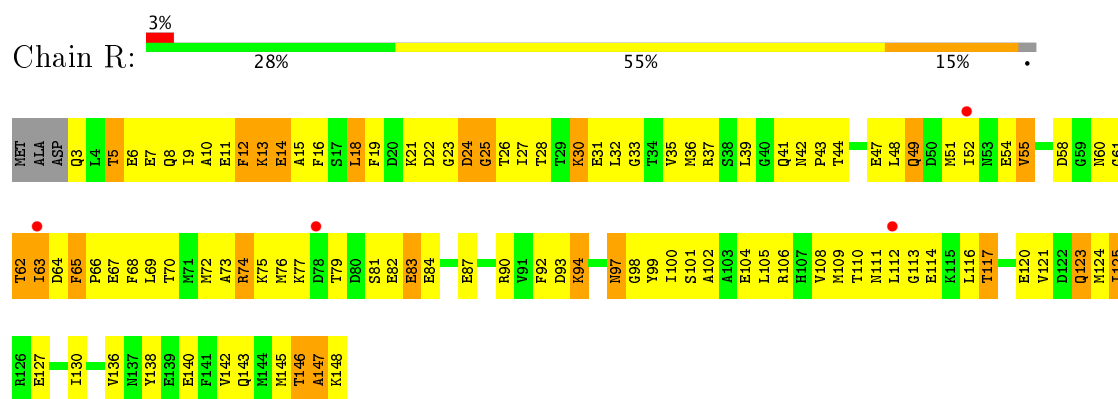
• Molecule 2: Calmodulin 2



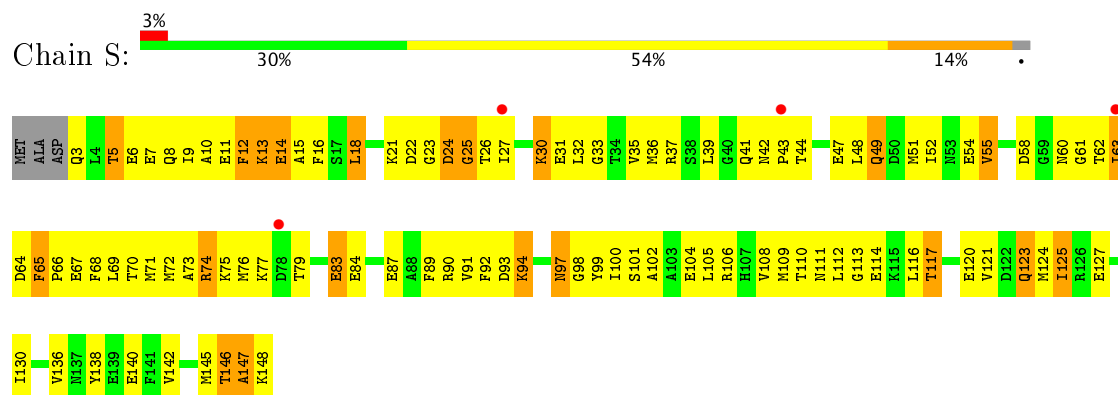
• Molecule 2: Calmodulin 2



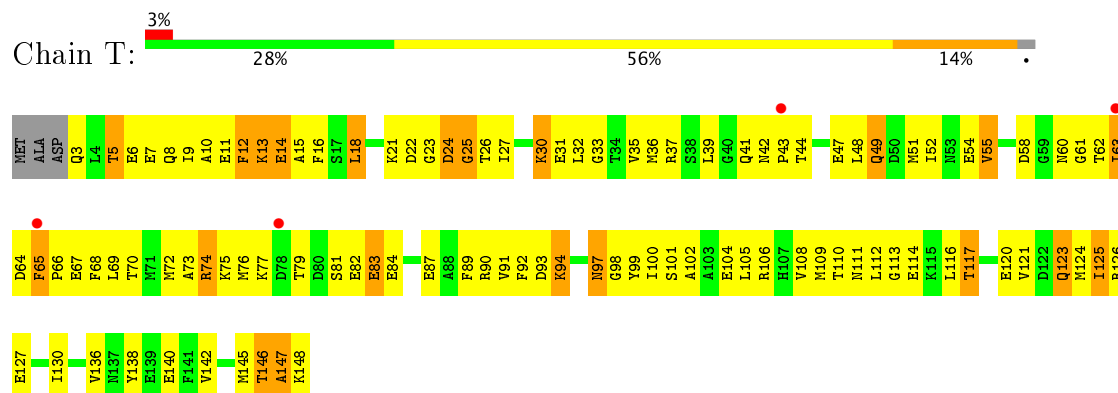
• Molecule 2: Calmodulin 2



• Molecule 2: Calmodulin 2



• Molecule 2: Calmodulin 2



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	320.50Å 185.04Å 142.45Å 90.00° 90.22° 90.00°	Depositor
Resolution (Å)	29.68 – 3.40 30.56 – 3.28	Depositor EDS
% Data completeness (in resolution range)	93.6 (29.68-3.40) 87.9 (30.56-3.28)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.21 (at 3.31Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.266 , 0.283 0.248 , 0.264	Depositor DCC
$R_{free}$ test set	5372 reflections (5.04%)	DCC
Wilson B-factor (Å <sup>2</sup> )	92.4	Xtriage
Anisotropy	0.143	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 51.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.448 for -1/2*h-3/2*k,-1/2*h+1/2*k,-l 0.449 for -1/2*h+3/2*k,1/2*h+1/2*k,-l 0.438 for 1/2*h-3/2*k,-1/2*h-1/2*k,-l 0.440 for 1/2*h+3/2*k,1/2*h-1/2*k,-l 0.440 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	42990	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.21% 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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: CA, MG, CMP

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	0.54	0/6104	0.80	8/8208 (0.1%)
1	B	0.54	0/6104	0.82	10/8208 (0.1%)
1	C	0.57	3/6104 (0.0%)	0.81	7/8208 (0.1%)
1	D	0.55	0/6104	0.82	9/8208 (0.1%)
1	E	0.54	0/6104	0.81	9/8208 (0.1%)
1	F	0.55	0/6104	0.81	8/8208 (0.1%)
2	O	0.61	1/1158 (0.1%)	0.78	0/1553
2	P	0.61	1/1158 (0.1%)	0.78	0/1553
2	Q	0.61	1/1158 (0.1%)	0.79	0/1553
2	R	0.60	1/1158 (0.1%)	0.78	0/1553
2	S	0.61	1/1158 (0.1%)	0.79	0/1553
2	T	0.61	1/1158 (0.1%)	0.78	0/1553
All	All	0.56	9/43572 (0.0%)	0.81	51/58566 (0.1%)

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	123	GLU	C-N	-6.32	1.19	1.34
1	C	261	ALA	C-N	5.87	1.45	1.34
2	S	42	ASN	N-CA	-5.57	1.35	1.46
2	T	42	ASN	N-CA	-5.55	1.35	1.46
2	P	42	ASN	N-CA	-5.54	1.35	1.46

The worst 5 of 51 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	188	LEU	N-CA-C	-7.80	89.94	111.00
1	D	188	LEU	N-CA-C	-7.80	89.95	111.00
1	A	188	LEU	N-CA-C	-7.79	89.96	111.00
1	C	188	LEU	N-CA-C	-7.79	89.97	111.00
1	E	188	LEU	N-CA-C	-7.79	89.98	111.00

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5992	0	6010	790	0
1	B	5992	0	6010	797	0
1	C	5992	0	6009	781	0
1	D	5992	0	6010	781	0
1	E	5992	0	6010	784	0
1	F	5992	0	6010	782	0
2	O	1146	0	1071	138	0
2	P	1146	0	1071	144	0
2	Q	1146	0	1071	142	0
2	R	1146	0	1071	141	0
2	S	1146	0	1071	140	0
2	T	1146	0	1071	146	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	22	0	11	4	0
4	B	22	0	11	4	0
4	C	22	0	11	4	0
4	D	22	0	11	4	0
4	E	22	0	11	4	0
4	F	22	0	11	4	0
5	O	3	0	0	0	0
5	P	3	0	0	0	0
5	Q	3	0	0	0	0
5	R	3	0	0	0	0
5	S	3	0	0	0	0
5	T	3	0	0	0	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
6	C	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	D	1	0	0	0	0
6	E	1	0	0	0	0
6	F	1	0	0	0	0
All	All	42990	0	42551	5408	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 63.

The worst 5 of 5408 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)
4:A:901:CMP:C2	4:A:901:CMP:H2	0.97	1.49
4:D:904:CMP:C2	4:D:904:CMP:H2	0.97	1.49
4:B:902:CMP:H2	4:B:902:CMP:C2	0.97	1.48
4:C:903:CMP:H2	4:C:903:CMP:C2	0.97	1.48
4:F:906:CMP:H2	4:F:906:CMP:C2	0.97	1.47

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	733/777 (94%)	512 (70%)	172 (24%)	49 (7%)	1	14
1	B	733/777 (94%)	512 (70%)	172 (24%)	49 (7%)	1	14
1	C	733/777 (94%)	510 (70%)	172 (24%)	51 (7%)	1	13
1	D	733/777 (94%)	516 (70%)	168 (23%)	49 (7%)	1	14
1	E	733/777 (94%)	512 (70%)	171 (23%)	50 (7%)	1	14
1	F	733/777 (94%)	513 (70%)	169 (23%)	51 (7%)	1	13
2	O	144/149 (97%)	100 (69%)	35 (24%)	9 (6%)	1	16

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	P	144/149 (97%)	102 (71%)	33 (23%)	9 (6%)	1	16
2	Q	144/149 (97%)	102 (71%)	33 (23%)	9 (6%)	1	16
2	R	144/149 (97%)	102 (71%)	33 (23%)	9 (6%)	1	16
2	S	144/149 (97%)	102 (71%)	33 (23%)	9 (6%)	1	16
2	T	144/149 (97%)	100 (69%)	35 (24%)	9 (6%)	1	16
All	All	5262/5556 (95%)	3683 (70%)	1226 (23%)	353 (7%)	1	14

5 of 353 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	137	PHE
1	A	183	SER
1	A	377	GLN
1	A	510	GLN
1	A	787	THR

### 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	664/705 (94%)	575 (87%)	89 (13%)	4	23
1	B	664/705 (94%)	569 (86%)	95 (14%)	4	20
1	C	664/705 (94%)	574 (86%)	90 (14%)	4	22
1	D	664/705 (94%)	572 (86%)	92 (14%)	4	21
1	E	664/705 (94%)	571 (86%)	93 (14%)	4	21
1	F	664/705 (94%)	570 (86%)	94 (14%)	4	21
2	O	123/127 (97%)	104 (85%)	19 (15%)	3	17
2	P	123/127 (97%)	104 (85%)	19 (15%)	3	17
2	Q	123/127 (97%)	104 (85%)	19 (15%)	3	17
2	R	123/127 (97%)	104 (85%)	19 (15%)	3	17
2	S	123/127 (97%)	105 (85%)	18 (15%)	3	19

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	T	123/127 (97%)	105 (85%)	18 (15%)	3	19
All	All	4722/4992 (95%)	4057 (86%)	665 (14%)	4	21

5 of 665 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	186	LYS
1	E	114	HIS
2	R	30	LYS
1	D	254	ARG
1	D	501	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 199 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	785	ASN
1	D	576	ASN
1	F	781	ASN
1	C	794	GLN
1	D	337	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 30 ligands modelled in this entry, 24 are monoatomic - leaving 6 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	CMP	A	901	-	21,25,25	1.72	4 (19%)	22,39,39	2.98	6 (27%)
4	CMP	B	902	3	21,25,25	1.73	5 (23%)	22,39,39	2.83	6 (27%)
4	CMP	C	903	3	21,25,25	1.71	5 (23%)	22,39,39	2.86	5 (22%)
4	CMP	D	904	3	21,25,25	1.70	4 (19%)	22,39,39	2.83	5 (22%)
4	CMP	E	905	3	21,25,25	1.73	6 (28%)	22,39,39	2.88	5 (22%)
4	CMP	F	906	3	21,25,25	1.76	5 (23%)	22,39,39	2.93	5 (22%)

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	CMP	A	901	-	-	0/0/31/31	0/4/4/4
4	CMP	B	902	3	-	0/0/31/31	0/4/4/4
4	CMP	C	903	3	-	0/0/31/31	0/4/4/4
4	CMP	D	904	3	-	0/0/31/31	0/4/4/4
4	CMP	E	905	3	-	0/0/31/31	0/4/4/4
4	CMP	F	906	3	-	0/0/31/31	0/4/4/4

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	903	CMP	C2'-C3'	-4.34	1.43	1.53
4	A	901	CMP	C2'-C3'	-4.25	1.43	1.53
4	D	904	CMP	C2'-C3'	-4.22	1.43	1.53
4	F	906	CMP	C2'-C3'	-4.13	1.43	1.53
4	E	905	CMP	C2'-C3'	-4.12	1.43	1.53

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	901	CMP	C4'-O4'-C1'	-10.56	98.53	109.77
4	F	906	CMP	C4'-O4'-C1'	-10.40	98.70	109.77
4	C	903	CMP	C4'-O4'-C1'	-10.14	98.98	109.77

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	E	905	CMP	C4'-O4'-C1'	-10.13	98.98	109.77
4	D	904	CMP	C4'-O4'-C1'	-10.02	99.10	109.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 24 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	901	CMP	4	0
4	B	902	CMP	4	0
4	C	903	CMP	4	0
4	D	904	CMP	4	0
4	E	905	CMP	4	0
4	F	906	CMP	4	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	123:GLU	C	124:GLU	N	1.19

## 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	735/777 (94%)	0.23	31 (4%) 37 34	18, 73, 128, 137	0
1	B	735/777 (94%)	0.23	28 (3%) 41 37	18, 73, 128, 138	0
1	C	735/777 (94%)	0.26	34 (4%) 33 30	17, 73, 128, 138	0
1	D	735/777 (94%)	0.24	30 (4%) 38 34	17, 73, 127, 137	0
1	E	735/777 (94%)	0.23	27 (3%) 42 38	17, 72, 127, 137	0
1	F	735/777 (94%)	0.22	29 (3%) 40 36	19, 73, 127, 138	0
2	O	146/149 (97%)	0.08	4 (2%) 55 51	17, 60, 119, 124	0
2	P	146/149 (97%)	0.09	3 (2%) 64 60	17, 60, 119, 124	0
2	Q	146/149 (97%)	0.08	2 (1%) 75 71	17, 60, 119, 123	0
2	R	146/149 (97%)	0.07	4 (2%) 55 51	16, 60, 119, 123	0
2	S	146/149 (97%)	0.14	4 (2%) 55 51	17, 60, 119, 123	0
2	T	146/149 (97%)	0.10	4 (2%) 55 51	16, 60, 119, 124	0
All	All	5286/5556 (95%)	0.21	200 (3%) 41 37	16, 70, 126, 138	0

The worst 5 of 200 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	162	ASN	17.4
1	D	162	ASN	9.9
1	A	163	SER	9.5
1	B	204	ASP	8.8
1	F	230	ILE	8.6

### 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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	CMP	C	903	22/22	0.90	0.51	3.42	48,63,81,92	0
4	CMP	B	902	22/22	0.88	0.42	2.78	49,63,81,89	0
4	CMP	E	905	22/22	0.89	0.39	2.41	47,63,84,91	0
4	CMP	D	904	22/22	0.90	0.33	1.87	48,62,81,86	0
4	CMP	F	906	22/22	0.86	0.39	1.86	48,63,84,93	0
4	CMP	A	901	22/22	0.89	0.36	1.83	48,65,85,93	0
5	CA	Q	806	1/1	0.97	0.19	-0.21	51,51,51,51	0
5	CA	T	812	1/1	0.98	0.18	-0.38	54,54,54,54	0
5	CA	P	804	1/1	0.99	0.18	-0.40	47,47,47,47	0
5	CA	O	802	1/1	0.96	0.18	-0.53	47,47,47,47	0
5	CA	Q	805	1/1	0.99	0.19	-0.55	29,29,29,29	0
5	CA	P	803	1/1	0.98	0.17	-0.66	30,30,30,30	0
5	CA	S	810	1/1	0.99	0.17	-0.70	51,51,51,51	0
5	CA	R	707	1/1	0.97	0.11	-0.89	69,69,69,69	0
5	CA	R	808	1/1	0.98	0.16	-0.90	47,47,47,47	0
5	CA	S	709	1/1	0.91	0.08	-1.00	78,78,78,78	0
5	CA	Q	705	1/1	0.82	0.15	-1.02	76,76,76,76	0
5	CA	O	701	1/1	0.91	0.12	-1.22	81,81,81,81	0
5	CA	P	703	1/1	0.94	0.13	-1.24	76,76,76,76	0
5	CA	T	711	1/1	0.90	0.11	-1.25	79,79,79,79	0
5	CA	T	811	1/1	0.97	0.14	-1.30	26,26,26,26	0
5	CA	O	801	1/1	0.98	0.14	-1.35	27,27,27,27	0
5	CA	R	807	1/1	0.97	0.13	-1.43	29,29,29,29	0
5	CA	S	809	1/1	0.97	0.14	-1.47	26,26,26,26	0
3	MG	A	900	1/1	0.99	0.18	-	13,13,13,13	0
3	MG	E	904	1/1	0.97	0.16	-	17,17,17,17	0
3	MG	D	903	1/1	0.98	0.14	-	14,14,14,14	0
3	MG	C	902	1/1	0.98	0.14	-	17,17,17,17	0
3	MG	B	901	1/1	0.99	0.12	-	27,27,27,27	0
3	MG	F	905	1/1	0.97	0.21	-	15,15,15,15	0

## 6.5 Other polymers

There are no such residues in this entry.