



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 30, 2023 – 09:08 am GMT

PDB ID : 8BGB  
Title : Structure of the citrate-bound extracytoplasmic PAS domain of histidine kinase CitA from *Geobacillus thermodenitrificans*  
Authors : Becker, S.  
Deposited on : 2022-10-27  
Resolution : 1.70 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 types of validation reports are described at

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

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

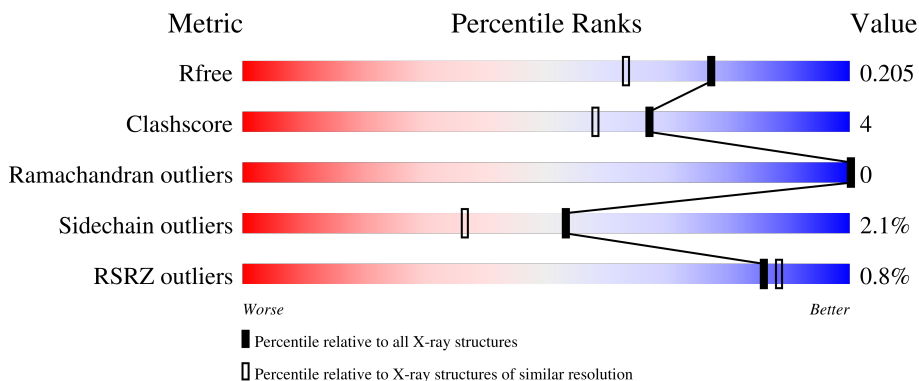
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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  $\geq 3$ , 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	AAA	131	 92% 7% ..
1	BBB	131	 88% 8% .. 2%
1	CCC	131	 87% 10% ..
1	DDD	131	 83% 12% ..
1	EEE	131	 92% 7% .. 2%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	FFF	131	 89% 6% • •
1	GGG	131	 2% 91% 7% •
1	HHH	131	 86% 11% •

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8715 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 Histidine kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	130	1010	628	186	194	2	0	4	0
1	BBB	126	953	595	175	181	2	0	2	1
1	CCC	129	1014	633	182	196	3	0	7	1
1	DDD	127	969	606	178	182	3	0	3	1
1	EEE	130	978	612	176	188	2	0	1	1
1	FFF	127	964	603	174	184	3	0	3	0
1	GGG	128	991	618	180	191	2	0	5	0
1	HHH	128	988	617	178	190	3	0	5	0

There are 24 discrepancies between the modelled and reference sequences:

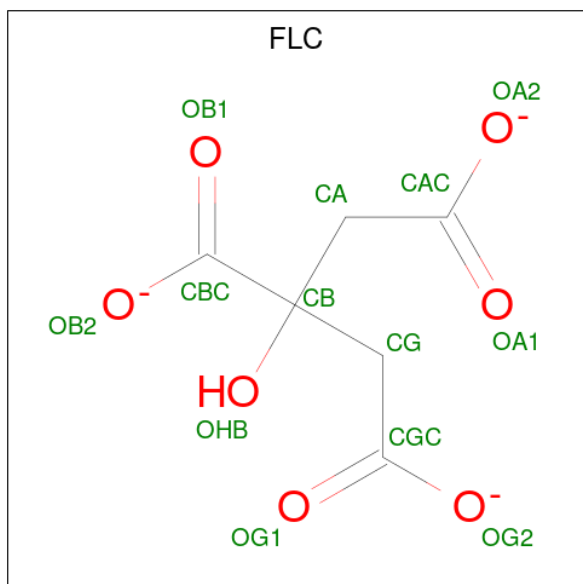
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	31	GLY	-	expression tag	UNP A4IPE6
AAA	32	SER	-	expression tag	UNP A4IPE6
AAA	53	SER	PRO	variant	UNP A4IPE6
BBB	31	GLY	-	expression tag	UNP A4IPE6
BBB	32	SER	-	expression tag	UNP A4IPE6
BBB	53	SER	PRO	variant	UNP A4IPE6
CCC	31	GLY	-	expression tag	UNP A4IPE6
CCC	32	SER	-	expression tag	UNP A4IPE6
CCC	53	SER	PRO	variant	UNP A4IPE6
DDD	31	GLY	-	expression tag	UNP A4IPE6
DDD	32	SER	-	expression tag	UNP A4IPE6
DDD	53	SER	PRO	variant	UNP A4IPE6
EEE	31	GLY	-	expression tag	UNP A4IPE6

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
EEE	32	SER	-	expression tag	UNP A4IPE6
EEE	53	SER	PRO	variant	UNP A4IPE6
FFF	31	GLY	-	expression tag	UNP A4IPE6
FFF	32	SER	-	expression tag	UNP A4IPE6
FFF	53	SER	PRO	variant	UNP A4IPE6
GGG	31	GLY	-	expression tag	UNP A4IPE6
GGG	32	SER	-	expression tag	UNP A4IPE6
GGG	53	SER	PRO	variant	UNP A4IPE6
HHH	31	GLY	-	expression tag	UNP A4IPE6
HHH	32	SER	-	expression tag	UNP A4IPE6
HHH	53	SER	PRO	variant	UNP A4IPE6

- Molecule 2 is CITRATE ANION (three-letter code: FLC) (formula: C<sub>6</sub>H<sub>5</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	1	Total C O 13 6 7	0	0
2	BBB	1	Total C O 13 6 7	0	0
2	CCC	1	Total C O 13 6 7	0	0
2	DDD	1	Total C O 13 6 7	0	0
2	EEE	1	Total C O 13 6 7	0	0
2	FFF	1	Total C O 13 6 7	0	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	GGG	1	Total	C	O	0	0
			13	6	7		
2	HHH	1	Total	C	O	0	0
			13	6	7		

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	GGG	1	Total	Na	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	118	Total	O	0	2
			120	120		
4	BBB	77	Total	O	0	3
			80	80		
4	CCC	112	Total	O	0	4
			116	116		
4	DDD	90	Total	O	0	0
			90	90		
4	EEE	67	Total	O	0	0
			67	67		
4	FFF	83	Total	O	0	3
			86	86		
4	GGG	83	Total	O	0	3
			86	86		
4	HHH	96	Total	O	0	2
			98	98		

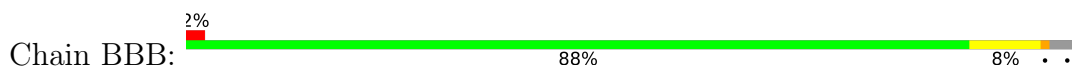
### 3 Residue-property plots [i](#)

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

- Molecule 1: Histidine kinase



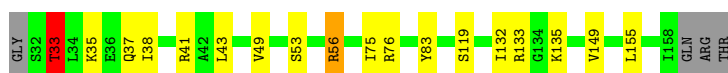
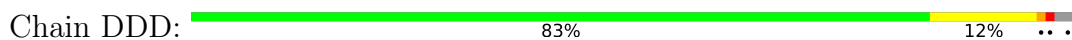
- Molecule 1: Histidine kinase



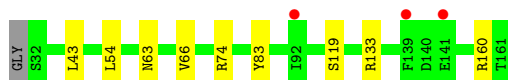
- Molecule 1: Histidine kinase



- Molecule 1: Histidine kinase



- Molecule 1: Histidine kinase

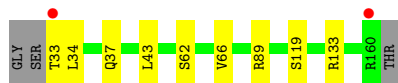


- Molecule 1: Histidine kinase

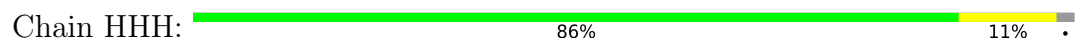




- Molecule 1: Histidine kinase



- Molecule 1: Histidine kinase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	64.21Å 101.14Å 77.88Å 90.00° 105.63° 90.00°	Depositor
Resolution (Å)	48.79 – 1.70 48.74 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.79-1.70) 99.9 (48.74-1.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.71 (at 1.70Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.167 , 0.196 0.178 , 0.205	Depositor DCC
$R_{free}$ test set	5202 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.3	Xtrriage
Anisotropy	0.107	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 46.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8715	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.82% 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: NA, FLC

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	AAA	0.77	0/1021	0.88	2/1376 (0.1%)
1	BBB	0.74	0/963	0.84	0/1296
1	CCC	0.82	2/1025 (0.2%)	0.86	2/1378 (0.1%)
1	DDD	0.72	0/980	0.85	1/1319 (0.1%)
1	EEE	0.71	0/989	0.77	0/1331
1	FFF	0.76	1/975 (0.1%)	0.80	0/1314
1	GGG	0.75	0/1001	0.82	0/1349
1	HHH	0.78	1/999 (0.1%)	0.78	0/1346
All	All	0.76	4/7953 (0.1%)	0.83	5/10709 (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	AAA	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	CCC	100	GLU	CD-OE2	6.36	1.32	1.25
1	FFF	82	GLU	CD-OE1	-5.64	1.19	1.25
1	CCC	100	GLU	CD-OE1	5.42	1.31	1.25
1	HHH	100	GLU	CD-OE2	5.32	1.31	1.25

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	AAA	67	ARG	NE-CZ-NH2	-6.34	117.13	120.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	CCC	52	THR	CA-CB-OG1	-6.17	96.04	109.00
1	AAA	67	ARG	NE-CZ-NH1	5.47	123.04	120.30
1	CCC	52	THR	OG1-CB-CG2	5.40	122.43	110.00
1	DDD	33	THR	CA-C-O	-5.19	109.21	120.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	AAA	141	GLU	Mainchain

## 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	AAA	1010	0	1028	7	0
1	BBB	953	0	972	6	0
1	CCC	1014	0	1025	10	0
1	DDD	969	0	989	13	1
1	EEE	978	0	1003	8	0
1	FFF	964	0	967	6	1
1	GGG	991	0	984	8	0
1	HHH	988	0	999	12	0
2	AAA	13	0	5	1	0
2	BBB	13	0	5	1	0
2	CCC	13	0	5	1	0
2	DDD	13	0	5	1	0
2	EEE	13	0	5	1	0
2	FFF	13	0	5	0	0
2	GGG	13	0	5	0	0
2	HHH	13	0	5	1	0
3	GGG	1	0	0	0	0
4	AAA	120	0	0	1	0
4	BBB	80	0	0	1	1
4	CCC	116	0	0	2	0
4	DDD	90	0	0	2	0
4	EEE	67	0	0	2	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	FFF	86	0	0	1	1
4	GGG	86	0	0	1	0
4	HHH	98	0	0	0	0
All	All	8715	0	8007	62	2

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.

All (62) 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:HHH:49:VAL:O	1:HHH:52[B]:THR:HG22	1.74	0.88
1:CCC:57[B]:GLU:OE2	4:CCC:301:HOH:O	1.96	0.83
1:FFF:63:ASN:HD22	1:GGG:37:GLN:HE22	1.29	0.81
1:CCC:76:ARG:NH2	1:CCC:77[B]:GLN:OE1	2.17	0.76
1:DDD:33:THR:HG23	1:DDD:37:GLN:HE21	1.52	0.74
1:DDD:41[A]:ARG:NH1	4:DDD:302:HOH:O	2.24	0.69
1:FFF:63:ASN:HD22	1:GGG:37:GLN:NE2	1.91	0.69
1:HHH:52[B]:THR:CG2	1:HHH:55:VAL:HG23	2.25	0.67
1:CCC:76:ARG:NH1	4:CCC:302:HOH:O	2.27	0.66
1:DDD:41[A]:ARG:NH2	4:DDD:301:HOH:O	2.05	0.66
1:AAA:100:GLU:OE1	1:AAA:104:LYS:NZ	2.28	0.66
1:DDD:76:ARG:HH22	1:EEE:160:ARG:HD2	1.62	0.63
1:DDD:132:ILE:HG13	1:DDD:155:LEU:HD21	1.81	0.63
1:AAA:61:ASP:OD2	1:AAA:67:ARG:HD3	1.99	0.63
1:FFF:35:LYS:HD3	1:FFF:123[A]:GLU:OE2	1.99	0.62
1:CCC:48:THR:O	1:CCC:52:THR:HG23	2.02	0.60
1:EEE:66:VAL:HG13	1:HHH:38:ILE:HD11	1.84	0.58
1:AAA:61:ASP:OD2	1:AAA:67:ARG:CD	2.53	0.57
1:BBB:41:ARG:NH1	4:BBB:302:HOH:O	2.30	0.54
1:EEE:54:LEU:C	1:EEE:54:LEU:HD23	2.29	0.54
1:HHH:52[B]:THR:HG21	1:HHH:55:VAL:HG23	1.89	0.53
1:DDD:76:ARG:NH2	1:EEE:160:ARG:HD2	2.24	0.52
1:AAA:54[B]:LEU:C	1:AAA:54[B]:LEU:HD23	2.32	0.51
1:AAA:159[A]:GLN:HG2	4:AAA:321:HOH:O	2.11	0.51
1:BBB:100:GLU:OE2	1:BBB:104:LYS:NZ	2.42	0.51
1:EEE:63:ASN:O	4:EEE:301:HOH:O	2.18	0.50
1:CCC:38:ILE:HD11	1:GGG:66:VAL:HG13	1.95	0.48
1:GGG:89:ARG:CZ	1:GGG:89:ARG:CD	2.92	0.48
1:EEE:74:ARG:NH2	4:EEE:304:HOH:O	2.47	0.47
1:FFF:49:VAL:HG22	1:FFF:75:ILE:HD13	1.96	0.47

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:HHH:83:TYR:CZ	2:HHH:201:FLC:HA2	2.51	0.46
1:DDD:35:LYS:HG2	1:DDD:132:ILE:HD11	1.97	0.46
1:AAA:83:TYR:CZ	2:AAA:201:FLC:HG1	2.51	0.46
1:HHH:49:VAL:HG22	1:HHH:75:ILE:HD13	1.98	0.46
1:EEE:43:LEU:HD11	1:EEE:119:SER:HB3	1.98	0.46
1:FFF:41:ARG:NH1	4:FFF:302:HOH:O	2.35	0.45
1:DDD:49:VAL:HG22	1:DDD:75:ILE:HD13	1.98	0.45
1:CCC:53[B]:SER:O	1:CCC:57[B]:GLU:HG3	2.17	0.45
1:DDD:83:TYR:CZ	2:DDD:201:FLC:HA2	2.53	0.44
1:AAA:66:VAL:HG22	1:DDD:38:ILE:HD11	1.98	0.44
1:GGG:33:THR:HG23	1:GGG:34:LEU:N	2.33	0.44
1:HHH:52[B]:THR:CG2	1:HHH:55:VAL:CG2	2.95	0.43
1:GGG:43:LEU:HD11	1:GGG:119:SER:HB3	2.00	0.43
1:BBB:56:ARG:HD3	1:BBB:56:ARG:C	2.39	0.43
1:FFF:56:ARG:HD3	1:FFF:56:ARG:C	2.40	0.42
1:HHH:52[B]:THR:HG21	1:HHH:71:PHE:CZ	2.55	0.42
1:CCC:56:ARG:C	1:CCC:56:ARG:HD3	2.40	0.42
1:EEE:83:TYR:CZ	2:EEE:201:FLC:HA2	2.54	0.42
1:CCC:49:VAL:HG22	1:CCC:75:ILE:HD13	2.02	0.42
1:DDD:56:ARG:HD3	1:DDD:56:ARG:C	2.40	0.42
1:GGG:33:THR:HG22	4:GGG:371:HOH:O	2.19	0.42
1:HHH:139:PHE:HA	1:HHH:144[B]:SER:O	2.20	0.41
1:BBB:83:TYR:CZ	2:BBB:201:FLC:HG1	2.54	0.41
1:BBB:43:LEU:HD11	1:BBB:119:SER:HB3	2.02	0.41
1:DDD:43:LEU:HD11	1:DDD:119:SER:HB3	2.02	0.41
1:HHH:40[B]:MET:HE3	1:HHH:40[B]:MET:HB3	1.93	0.41
1:BBB:89:ARG:HG2	1:BBB:90:GLN:NE2	2.36	0.41
1:CCC:83:TYR:CZ	2:CCC:201:FLC:HA2	2.56	0.41
1:HHH:43:LEU:HD11	1:HHH:119:SER:HB3	2.03	0.41
1:HHH:52[B]:THR:HG23	1:HHH:55:VAL:H	1.86	0.41
1:CCC:159:GLN:O	1:GGG:62:SER:HB2	2.21	0.40
1:DDD:135:LYS:HA	1:DDD:149:VAL:O	2.22	0.40

All (2) 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)
1:DDD:53:SER:OG	1:FFF:57[B]:GLU:OE2[1_455]	1.81	0.39
4:BBB:370:HOH:O	4:FFF:383:HOH:O[1_454]	2.10	0.10

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	AAA	132/131 (101%)	131 (99%)	1 (1%)	0	100	100
1	BBB	124/131 (95%)	123 (99%)	1 (1%)	0	100	100
1	CCC	134/131 (102%)	133 (99%)	1 (1%)	0	100	100
1	DDD	128/131 (98%)	127 (99%)	1 (1%)	0	100	100
1	EEE	129/131 (98%)	128 (99%)	1 (1%)	0	100	100
1	FFF	128/131 (98%)	128 (100%)	0	0	100	100
1	GGG	131/131 (100%)	130 (99%)	1 (1%)	0	100	100
1	HHH	131/131 (100%)	131 (100%)	0	0	100	100
All	All	1037/1048 (99%)	1031 (99%)	6 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	AAA	108/106 (102%)	107 (99%)	1 (1%)	78	70
1	BBB	101/106 (95%)	98 (97%)	3 (3%)	41	22
1	CCC	107/106 (101%)	104 (97%)	3 (3%)	43	25
1	DDD	101/106 (95%)	98 (97%)	3 (3%)	41	22
1	EEE	104/106 (98%)	103 (99%)	1 (1%)	76	67
1	FFF	99/106 (93%)	96 (97%)	3 (3%)	41	22

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	GGG	102/106 (96%)	101 (99%)	1 (1%)	76	67
1	HHH	104/106 (98%)	102 (98%)	2 (2%)	57	41
All	All	826/848 (97%)	809 (98%)	17 (2%)	53	36

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	133	ARG
1	BBB	56	ARG
1	BBB	105	SER
1	BBB	133	ARG
1	CCC	33	THR
1	CCC	56	ARG
1	CCC	133	ARG
1	DDD	33	THR
1	DDD	56	ARG
1	DDD	133	ARG
1	EEE	133	ARG
1	FFF	41	ARG
1	FFF	56	ARG
1	FFF	133	ARG
1	GGG	133	ARG
1	HHH	56	ARG
1	HHH	133	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

Of 9 ligands modelled in this entry, 1 is monoatomic - leaving 8 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
2	FLC	GGG	201	3	12,12,12	1.21	1 (8%)	17,17,17	1.27	3 (17%)
2	FLC	CCC	201	-	12,12,12	1.49	2 (16%)	17,17,17	1.33	2 (11%)
2	FLC	HHH	201	-	12,12,12	1.08	1 (8%)	17,17,17	1.29	2 (11%)
2	FLC	AAA	201	-	12,12,12	1.16	1 (8%)	17,17,17	1.13	2 (11%)
2	FLC	FFF	201	-	12,12,12	1.09	1 (8%)	17,17,17	1.12	1 (5%)
2	FLC	EEE	201	-	12,12,12	1.24	1 (8%)	17,17,17	1.40	3 (17%)
2	FLC	DDD	201	-	12,12,12	1.16	1 (8%)	17,17,17	1.31	2 (11%)
2	FLC	BBB	201	-	12,12,12	1.34	1 (8%)	17,17,17	1.50	3 (17%)

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	FLC	GGG	201	3	-	0/16/16/16	-
2	FLC	CCC	201	-	-	0/16/16/16	-
2	FLC	HHH	201	-	-	0/16/16/16	-
2	FLC	AAA	201	-	-	2/16/16/16	-
2	FLC	FFF	201	-	-	2/16/16/16	-
2	FLC	EEE	201	-	-	0/16/16/16	-
2	FLC	DDD	201	-	-	2/16/16/16	-
2	FLC	BBB	201	-	-	2/16/16/16	-

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	CCC	201	FLC	CB-CBC	3.08	1.56	1.53

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	CCC	201	FLC	OB1-CBC	2.82	1.31	1.22
2	BBB	201	FLC	CB-CBC	2.78	1.56	1.53
2	AAA	201	FLC	OA1-CAC	2.43	1.30	1.22
2	GGG	201	FLC	OB1-CBC	2.43	1.30	1.22
2	DDD	201	FLC	OB1-CBC	2.21	1.29	1.22
2	FFF	201	FLC	OA1-CAC	2.16	1.29	1.22
2	EEE	201	FLC	OA1-CAC	2.12	1.29	1.22
2	HHH	201	FLC	OG1-CGC	2.00	1.28	1.22

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	BBB	201	FLC	OB1-CBC-CB	-3.79	116.89	122.25
2	DDD	201	FLC	OB1-CBC-CB	-3.76	116.93	122.25
2	EEE	201	FLC	OB1-CBC-CB	-3.64	117.10	122.25
2	CCC	201	FLC	OB1-CBC-CB	-3.52	117.27	122.25
2	AAA	201	FLC	OB1-CBC-CB	-3.31	117.57	122.25
2	BBB	201	FLC	OB2-CBC-CB	2.96	118.19	113.05
2	CCC	201	FLC	OB2-CBC-CB	2.89	118.08	113.05
2	HHH	201	FLC	OB1-CBC-CB	-2.85	118.21	122.25
2	GGG	201	FLC	OB1-CBC-CB	-2.65	118.50	122.25
2	FFF	201	FLC	OB1-CBC-CB	-2.64	118.52	122.25
2	HHH	201	FLC	OB2-CBC-CB	2.35	117.12	113.05
2	DDD	201	FLC	OB2-CBC-CB	2.30	117.04	113.05
2	EEE	201	FLC	OB2-CBC-CB	2.29	117.03	113.05
2	BBB	201	FLC	OA1-CAC-CA	-2.26	116.35	122.94
2	EEE	201	FLC	OG1-CGC-CG	-2.21	116.48	122.94
2	GGG	201	FLC	OB2-CBC-CB	2.19	116.84	113.05
2	GGG	201	FLC	OHB-CB-CBC	2.03	111.72	108.86
2	AAA	201	FLC	OB2-CBC-CB	2.03	116.57	113.05

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	BBB	201	FLC	CG-CB-CBC-OB1
2	BBB	201	FLC	CG-CB-CBC-OB2
2	FFF	201	FLC	CB-CA-CAC-OA2
2	AAA	201	FLC	CB-CA-CAC-OA2
2	AAA	201	FLC	CB-CA-CAC-OA1
2	DDD	201	FLC	CB-CG-CGC-OG1
2	DDD	201	FLC	CB-CG-CGC-OG2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	FFF	201	FLC	CB-CA-CAC-OA1

There are no ring outliers.

6 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	CCC	201	FLC	1	0
2	HHH	201	FLC	1	0
2	AAA	201	FLC	1	0
2	EEE	201	FLC	1	0
2	DDD	201	FLC	1	0
2	BBB	201	FLC	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	AAA	130/131 (99%)	-0.30	1 (0%) 86   88	14, 20, 33, 47	0
1	BBB	126/131 (96%)	-0.14	2 (1%) 72   76	16, 27, 44, 51	0
1	CCC	129/131 (98%)	-0.34	0 100   100	14, 19, 34, 45	0
1	DDD	127/131 (96%)	-0.36	0 100   100	17, 23, 37, 55	0
1	EEE	130/131 (99%)	0.03	3 (2%) 60   65	18, 29, 52, 80	0
1	FFF	127/131 (96%)	-0.29	0 100   100	17, 24, 40, 58	0
1	GGG	128/131 (97%)	-0.25	2 (1%) 72   76	17, 23, 38, 56	0
1	HHH	128/131 (97%)	-0.28	0 100   100	15, 21, 38, 54	0
All	All	1025/1048 (97%)	-0.24	8 (0%) 86   88	14, 23, 41, 80	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	161	THR	4.0
1	EEE	141	GLU	3.8
1	BBB	144	SER	2.7
1	BBB	62	SER	2.7
1	GGG	160	ARG	2.3
1	GGG	33	THR	2.2
1	EEE	139	PHE	2.0
1	EEE	92	ILE	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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 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
2	FLC	DDD	201	13/13	0.94	0.09	19,21,23,23	0
2	FLC	BBB	201	13/13	0.95	0.09	19,21,24,24	0
2	FLC	EEE	201	13/13	0.95	0.08	21,23,27,28	0
2	FLC	AAA	201	13/13	0.96	0.08	16,17,20,21	0
2	FLC	FFF	201	13/13	0.96	0.07	18,20,22,22	0
2	FLC	CCC	201	13/13	0.97	0.07	15,17,19,19	0
2	FLC	GGG	201	13/13	0.97	0.07	19,20,23,23	0
2	FLC	HHH	201	13/13	0.97	0.06	17,18,21,22	0
3	NA	GGG	202	1/1	0.99	0.08	33,33,33,33	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.