



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 02:28 am GMT

PDB ID : 2GRX  
Title : Crystal structure of TonB in complex with FhuA, E. coli outer membrane receptor for ferrichrome  
Authors : Pawelek, P.D.; Allaire, M.; Coulton, J.W.  
Deposited on : 2006-04-25  
Resolution : 3.30 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	:	trunk28620
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)	:	recalc28949

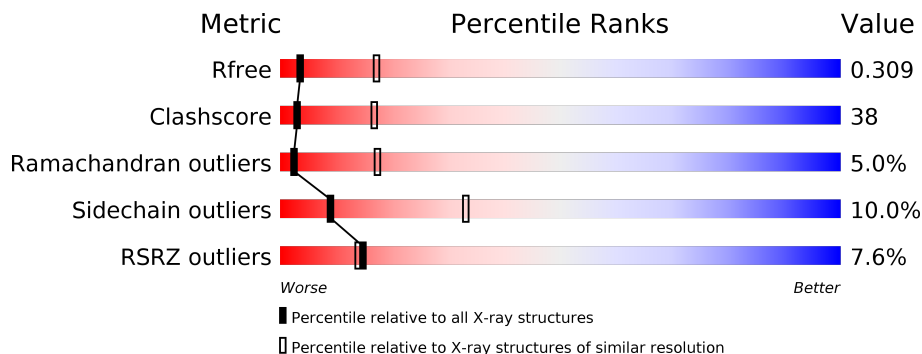
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.30 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	1034 (3.36-3.24)
Clashscore	112137	1100 (3.36-3.24)
Ramachandran outliers	110173	1081 (3.36-3.24)
Sidechain outliers	110143	1080 (3.36-3.24)
RSRZ outliers	101464	1039 (3.36-3.24)

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	725	
1	B	725	
2	C	229	
2	D	229	

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GCN	A	920	-	-	X	-
3	GCN	B	920	-	-	X	-
3	GCN	B	921	-	-	X	-
3	KDO	B	971	-	-	X	-
5	FTT	A	900	-	-	X	X
5	FTT	A	901	-	-	X	X
5	FTT	A	902	-	-	-	X
5	FTT	B	900	-	-	X	X
5	FTT	B	901	-	-	X	-
5	FTT	B	902	-	-	X	X
5	FTT	B	903	-	-	X	-
6	DPO	A	910	-	-	X	-
6	DPO	B	910	-	-	X	-
7	DAO	A	930	-	-	-	X
7	DAO	B	930	-	-	X	-
8	MYR	B	940	-	-	-	X

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 12533 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 Ferrichrome-iron receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	702	Total	C	N	O	S	0	0	0
			5466	3444	926	1082	14			
1	B	691	Total	C	N	O	S	0	0	0
			5394	3398	914	1068	14			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	406	SER	-	SEE REMARK 999	UNP P06971
A	407	SER	-	SEE REMARK 999	UNP P06971
A	408	HIS	-	SEE REMARK 999	UNP P06971
A	409	HIS	-	SEE REMARK 999	UNP P06971
A	410	HIS	-	SEE REMARK 999	UNP P06971
A	411	HIS	-	SEE REMARK 999	UNP P06971
A	412	HIS	-	SEE REMARK 999	UNP P06971
A	413	HIS	-	SEE REMARK 999	UNP P06971
A	414	GLY	-	SEE REMARK 999	UNP P06971
A	415	SER	-	SEE REMARK 999	UNP P06971
A	416	SER	-	SEE REMARK 999	UNP P06971
B	406	SER	-	SEE REMARK 999	UNP P06971
B	407	SER	-	SEE REMARK 999	UNP P06971
B	408	HIS	-	SEE REMARK 999	UNP P06971
B	409	HIS	-	SEE REMARK 999	UNP P06971
B	410	HIS	-	SEE REMARK 999	UNP P06971
B	411	HIS	-	SEE REMARK 999	UNP P06971
B	412	HIS	-	SEE REMARK 999	UNP P06971
B	413	HIS	-	SEE REMARK 999	UNP P06971
B	414	GLY	-	SEE REMARK 999	UNP P06971
B	415	SER	-	SEE REMARK 999	UNP P06971
B	416	SER	-	SEE REMARK 999	UNP P06971

- Molecule 2 is a protein called Protein tonB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	78	Total 624	C 393	N 120	O 109	S 2	0	0	0
2	D	78	Total 624	C 393	N 120	O 109	S 2	0	0	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	11	GLY	-	SEE REMARK 999	UNP P94739
C	12	SER	-	SEE REMARK 999	UNP P94739
C	13	SER	-	SEE REMARK 999	UNP P94739
C	14	HIS	-	SEE REMARK 999	UNP P94739
C	15	HIS	-	SEE REMARK 999	UNP P94739
C	16	HIS	-	SEE REMARK 999	UNP P94739
C	17	HIS	-	SEE REMARK 999	UNP P94739
C	18	HIS	-	SEE REMARK 999	UNP P94739
C	19	HIS	-	SEE REMARK 999	UNP P94739
C	20	SER	-	SEE REMARK 999	UNP P94739
C	21	SER	-	SEE REMARK 999	UNP P94739
C	22	GLY	-	SEE REMARK 999	UNP P94739
C	23	LEU	-	SEE REMARK 999	UNP P94739
C	24	VAL	-	SEE REMARK 999	UNP P94739
C	25	PRO	-	SEE REMARK 999	UNP P94739
C	26	ARG	-	SEE REMARK 999	UNP P94739
C	27	GLY	-	SEE REMARK 999	UNP P94739
C	28	SER	-	SEE REMARK 999	UNP P94739
C	29	HIS	-	SEE REMARK 999	UNP P94739
C	30	MET	-	SEE REMARK 999	UNP P94739
C	31	SER	-	SEE REMARK 999	UNP P94739
C	32	VAL	-	SEE REMARK 999	UNP P94739
D	11	GLY	-	SEE REMARK 999	UNP P94739
D	12	SER	-	SEE REMARK 999	UNP P94739
D	13	SER	-	SEE REMARK 999	UNP P94739
D	14	HIS	-	SEE REMARK 999	UNP P94739
D	15	HIS	-	SEE REMARK 999	UNP P94739
D	16	HIS	-	SEE REMARK 999	UNP P94739
D	17	HIS	-	SEE REMARK 999	UNP P94739
D	18	HIS	-	SEE REMARK 999	UNP P94739
D	19	HIS	-	SEE REMARK 999	UNP P94739
D	20	SER	-	SEE REMARK 999	UNP P94739
D	21	SER	-	SEE REMARK 999	UNP P94739
D	22	GLY	-	SEE REMARK 999	UNP P94739
D	23	LEU	-	SEE REMARK 999	UNP P94739

*Continued on next page...*

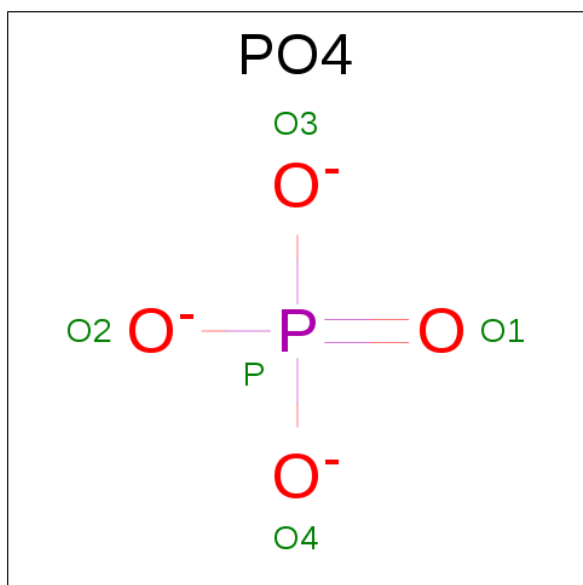
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	24	VAL	-	SEE REMARK 999	UNP P94739
D	25	PRO	-	SEE REMARK 999	UNP P94739
D	26	ARG	-	SEE REMARK 999	UNP P94739
D	27	GLY	-	SEE REMARK 999	UNP P94739
D	28	SER	-	SEE REMARK 999	UNP P94739
D	29	HIS	-	SEE REMARK 999	UNP P94739
D	30	MET	-	SEE REMARK 999	UNP P94739
D	31	SER	-	SEE REMARK 999	UNP P94739
D	32	VAL	-	SEE REMARK 999	UNP P94739

- Molecule 3 is a polymer of unknown type called SUGAR (5-MER).

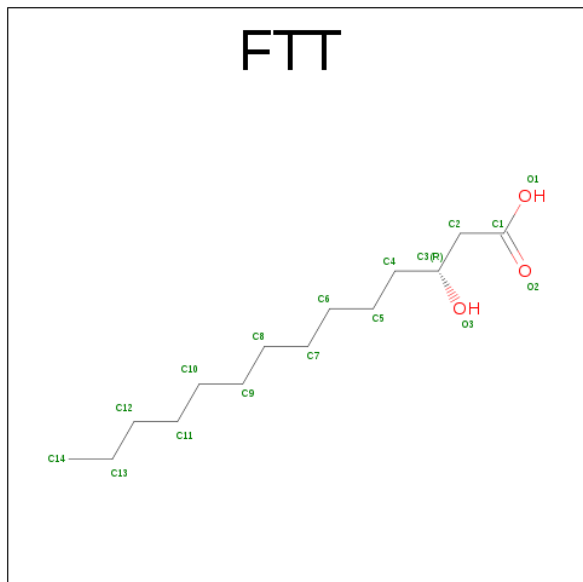
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	5	Total	C	N	O	0	0
			63	35	2	26		
3	B	5	Total	C	N	O	0	0
			63	35	2	26		

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



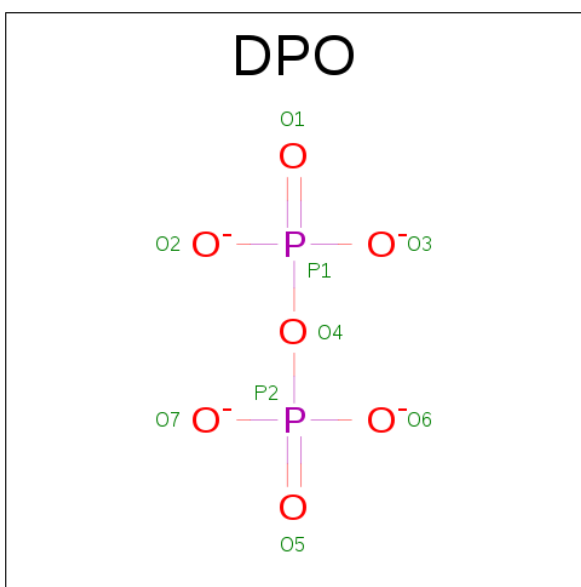
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	P	0	0
			4	3	1		
4	B	1	Total	O	P	0	0
			4	3	1		

- Molecule 5 is 3-HYDROXY-TETRADECANOIC ACID (three-letter code: FTT) (formula:  $C_{14}H_{28}O_3$ ).



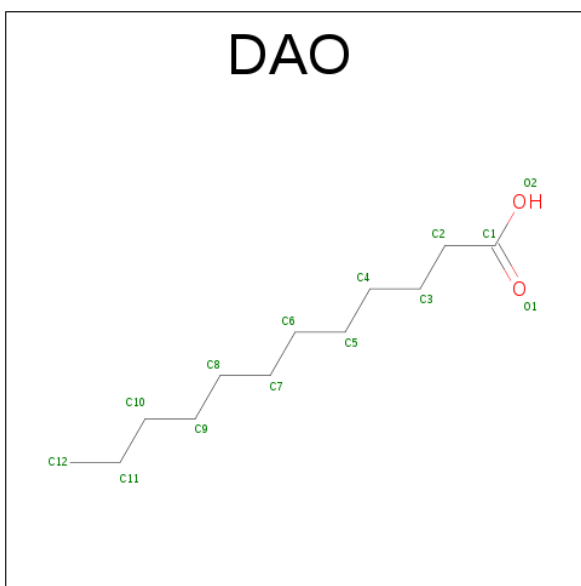
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			16	14	2		
5	A	1	Total	C	O	0	0
			16	14	2		
5	A	1	Total	C	O	0	0
			16	14	2		
5	A	1	Total	C	O	0	0
			16	14	2		
5	B	1	Total	C	O	0	0
			16	14	2		
5	B	1	Total	C	O	0	0
			16	14	2		
5	B	1	Total	C	O	0	0
			16	14	2		

- Molecule 6 is DIPHOSPHATE (three-letter code: DPO) (formula:  $O_7P_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	O	P	0	0
			8	6	2		
6	B	1	Total	O	P	0	0
			8	6	2		

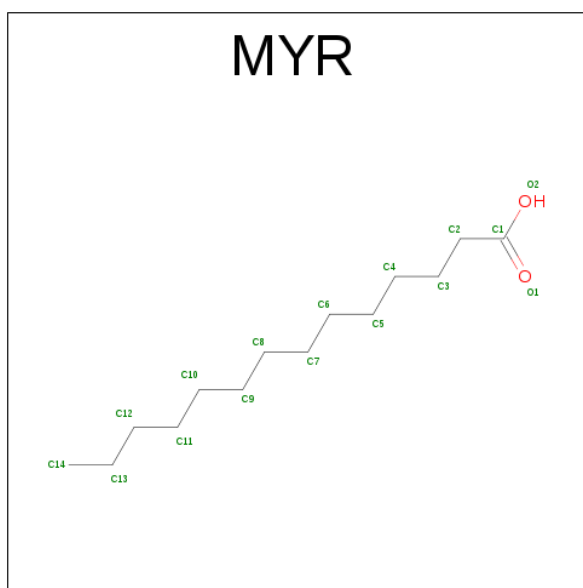
- Molecule 7 is LAURIC ACID (three-letter code: DAO) (formula:  $C_{12}H_{24}O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			13	12	1		
7	B	1	Total	C	O	0	0
			13	12	1		

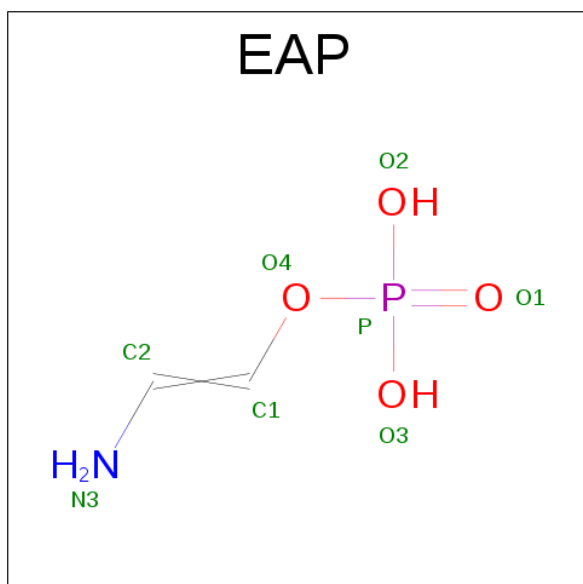


- Molecule 8 is MYRISTIC ACID (three-letter code: MYR) (formula:  $C_{14}H_{28}O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	B	1	Total	C	O	0	0
			15	14	1		

- Molecule 9 is 2-AMINO-VINYL-PHOSPHATE (three-letter code: EAP) (formula:  $C_2H_6NO_4P$ ).



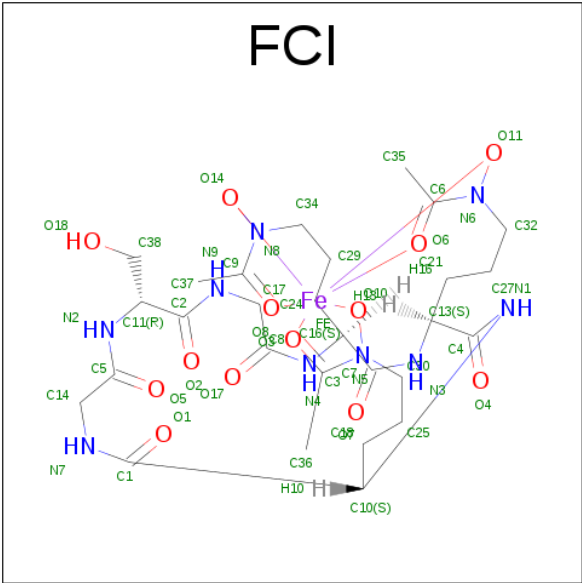
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
9	A	1	Total	C	N	O	P	0	0
			7	2	1	3	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
9	B	1	7	2	1	3	1	0	0

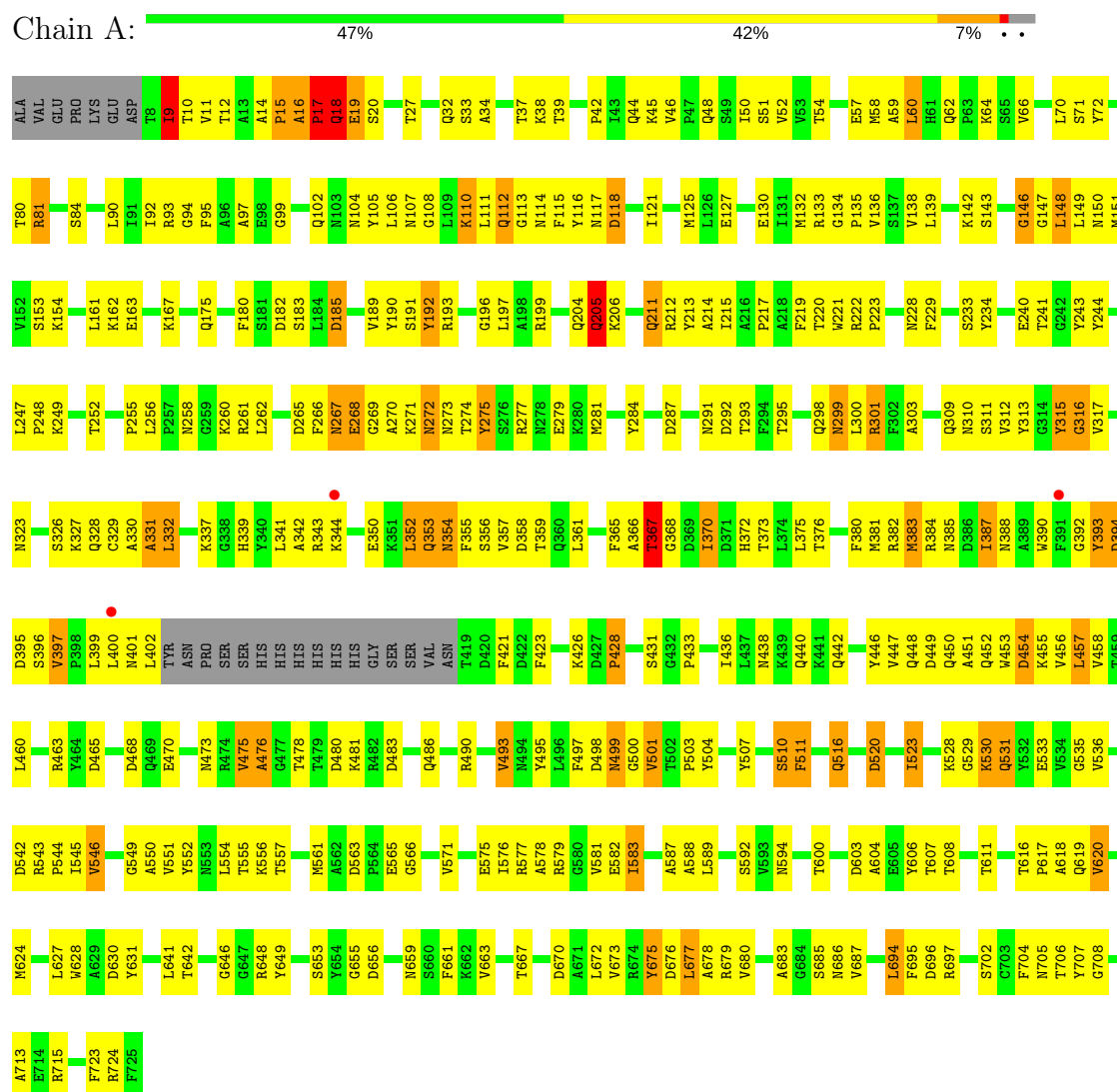
- Molecule 10 is FERRICROCIN-IRON (three-letter code: FCI) (formula: C<sub>28</sub>H<sub>44</sub>FeN<sub>9</sub>O<sub>13</sub>).



### 3 Residue-property plots

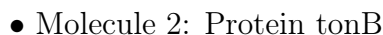
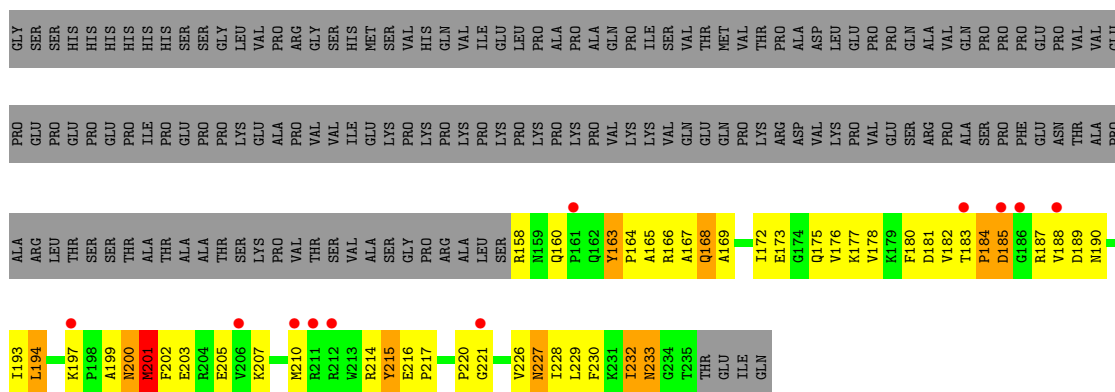
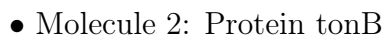
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

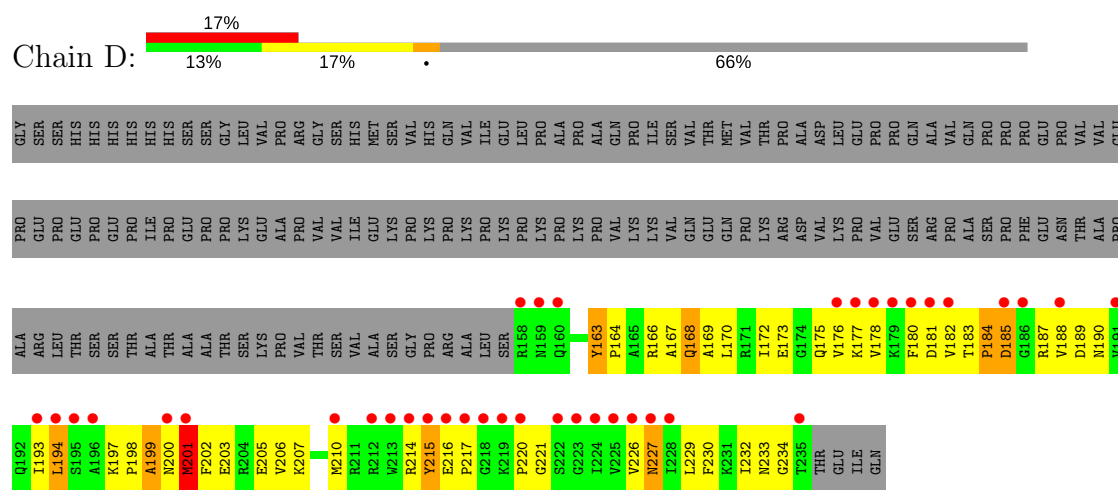
#### • Molecule 1: Ferrichrome-iron receptor



#### • Molecule 1: Ferrichrome-iron receptor







## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	234.32Å 91.84Å 138.51Å 90.00° 118.86° 90.00°	Depositor
Resolution (Å)	50.00 – 3.30 45.94 – 3.30	Depositor EDS
% Data completeness (in resolution range)	99.2 (50.00-3.30) 99.2 (45.94-3.30)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.33 (at 3.32Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.284 , 0.329 0.265 , 0.309	Depositor DCC
$R_{free}$ test set	2325 reflections (6.39%)	DCC
Wilson B-factor (Å <sup>2</sup> )	93.6	Xtriage
Anisotropy	0.301	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 68.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	12533	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	98.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.62% 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: GMH, FTT, PO4, MYR, DAO, GCN, KDO, FCI, DPO, EAP

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	0.50	0/5599	0.69	1/7609 (0.0%)
1	B	0.40	0/5525	0.66	0/7504
2	C	0.35	0/636	0.56	0/857
2	D	0.31	0/636	0.54	0/857
All	All	0.44	0/12396	0.66	1/16827 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	17	PRO	N-CA-C	5.33	125.96	112.10

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	5466	0	5193	358	0
1	B	5394	0	5120	344	0
2	C	624	0	640	53	0
2	D	624	0	640	48	0
3	A	63	0	50	18	0
3	B	63	0	51	24	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	4	0	0	0	0
4	B	4	0	0	0	0
5	A	64	0	107	30	0
5	B	64	0	107	100	0
6	A	8	0	0	8	0
6	B	8	0	0	8	0
7	A	13	0	23	5	0
7	B	13	0	23	7	0
8	B	15	0	27	4	0
9	A	7	0	4	2	0
9	B	7	0	4	3	0
10	A	46	0	31	2	0
10	B	46	0	31	1	0
All	All	12533	0	12051	922	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 38.

The worst 5 of 922 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)
5:B:902:FTT:H92	5:B:903:FTT:C8	1.28	1.55
3:B:971:KDO:C1	3:B:971:KDO:O1B	1.70	1.39
5:B:902:FTT:C9	5:B:903:FTT:H82	1.52	1.39
5:B:902:FTT:C13	5:B:903:FTT:H131	1.55	1.34
5:B:901:FTT:C14	5:B:903:FTT:H112	1.58	1.33

There are no symmetry-related clashes.

## 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	A	698/725 (96%)	590 (84%)	77 (11%)	31 (4%)	3	20
1	B	687/725 (95%)	582 (85%)	79 (12%)	26 (4%)	4	24
2	C	76/229 (33%)	55 (72%)	11 (14%)	10 (13%)	0	2
2	D	76/229 (33%)	53 (70%)	13 (17%)	10 (13%)	0	2
All	All	1537/1908 (81%)	1280 (83%)	180 (12%)	77 (5%)	2	17

5 of 77 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	16	ALA
1	A	18	GLN
1	A	316	GLY
1	A	331	ALA
1	A	367	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	579/601 (96%)	518 (90%)	61 (10%)	8	31
1	B	572/601 (95%)	516 (90%)	56 (10%)	9	34
2	C	66/200 (33%)	60 (91%)	6 (9%)	11	38
2	D	66/200 (33%)	61 (92%)	5 (8%)	15	47
All	All	1283/1602 (80%)	1155 (90%)	128 (10%)	9	33

5 of 128 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	620	VAL
1	B	192	TYR
2	C	168	GLN
1	A	679	ARG
1	B	80	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 73 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	692	ASN
1	B	205	GLN
2	C	208	ASN
1	B	107	ASN
1	B	267	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

10 carbohydrates 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	$\# Z  > 2$
3	GCN	A	920	3,5	10,10,11	0.80	0	9,13,15	0.72	0
3	GCN	A	921	3,5,4	10,10,11	1.01	1 (10%)	9,13,15	1.09	1 (11%)
3	GMH	A	960	9,3	13,13,14	0.80	0	16,18,20	0.68	0
3	KDO	A	970	3	12,15,16	0.74	0	13,21,24	0.73	0
3	KDO	A	971	3	12,15,16	0.45	0	13,21,24	0.72	0
3	GCN	B	920	3,5	10,10,11	0.80	0	9,13,15	0.73	0
3	GCN	B	921	3,5,4	10,10,11	1.02	1 (10%)	9,13,15	1.09	1 (11%)
3	GMH	B	960	9,3	13,13,14	0.80	0	16,18,20	0.69	0
3	KDO	B	970	3	12,15,16	0.74	0	13,21,24	0.72	0
3	KDO	B	971	3	12,15,16	0.45	0	13,21,24	0.73	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
3	GCN	A	920	3,5	-	0/2/15/18	0/1/1/1
3	GCN	A	921	3,5,4	-	0/2/15/18	0/1/1/1
3	GMH	A	960	9,3	-	0/6/23/26	0/1/1/1
3	KDO	A	970	3	-	0/6/26/30	0/1/1/1
3	KDO	A	971	3	-	0/6/26/30	0/1/1/1
3	GCN	B	920	3,5	-	0/2/15/18	0/1/1/1
3	GCN	B	921	3,5,4	-	0/2/15/18	0/1/1/1
3	GMH	B	960	9,3	-	0/6/23/26	0/1/1/1
3	KDO	B	970	3	-	0/6/26/30	0/1/1/1
3	KDO	B	971	3	-	0/6/26/30	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	921	GCN	C1-C2	2.61	1.55	1.52
3	B	921	GCN	C1-C2	2.64	1.55	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	921	GCN	C1-O5-C5	2.50	115.61	112.17
3	A	921	GCN	C1-O5-C5	2.53	115.65	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 42 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	920	GCN	7	0
3	A	921	GCN	5	0
3	A	960	GMH	4	0
3	A	970	KDO	3	0
3	A	971	KDO	3	0
3	B	920	GCN	7	0
3	B	921	GCN	7	0
3	B	960	GMH	4	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	970	KDO	3	0
3	B	971	KDO	7	0

## 5.6 Ligand geometry

19 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	$\# Z  > 2$
10	FCI	A	1050	-	42,51,56	1.49	4 (9%)	51,78,87	1.13	3 (5%)
5	FTT	A	900	3	15,15,16	0.39	0	15,15,17	0.65	0
5	FTT	A	901	-	15,15,16	0.31	0	15,15,17	0.76	0
5	FTT	A	902	3,7	15,15,16	0.31	0	15,15,17	0.77	0
5	FTT	A	903	-	15,15,16	0.45	0	15,15,17	0.51	0
6	DPO	A	910	-	5,7,8	3.30	2 (40%)	4,10,13	0.32	0
7	DAO	A	930	5	12,12,13	1.03	1 (8%)	11,11,13	0.80	1 (9%)
4	PO4	A	950	3	0,3,4	0.00	-	0,3,6	0.00	-
9	EAP	A	980	3	3,6,7	3.51	2 (66%)	0,6,9	0.00	-
10	FCI	B	1050	-	42,51,56	1.49	4 (9%)	51,78,87	1.13	3 (5%)
5	FTT	B	900	3	15,15,16	4.07	1 (6%)	15,15,17	2.22	2 (13%)
5	FTT	B	901	-	15,15,16	0.31	0	15,15,17	0.77	0
5	FTT	B	902	3,7	15,15,16	0.32	0	15,15,17	0.77	0
5	FTT	B	903	8	15,15,16	0.45	0	15,15,17	0.51	0
6	DPO	B	910	-	5,7,8	3.31	2 (40%)	4,10,13	0.32	0
7	DAO	B	930	5	12,12,13	1.03	1 (8%)	11,11,13	0.80	1 (9%)
8	MYR	B	940	5	14,14,15	0.97	1 (7%)	13,13,15	0.81	1 (7%)
4	PO4	B	950	3	0,3,4	0.00	-	0,3,6	0.00	-
9	EAP	B	980	3	3,6,7	3.50	2 (66%)	0,6,9	0.00	-

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
10	FCI	A	1050	-	-	0/57/102/116	0/0/6/6
5	FTT	A	900	3	-	0/14/14/15	0/0/0/0
5	FTT	A	901	-	-	0/14/14/15	0/0/0/0
5	FTT	A	902	3,7	-	0/14/14/15	0/0/0/0
5	FTT	A	903	-	-	0/14/14/15	0/0/0/0
6	DPO	A	910	-	-	0/2/5/6	0/0/0/0
7	DAO	A	930	5	-	0/10/10/11	0/0/0/0
4	PO4	A	950	3	-	0/0/0/0	0/0/0/0
9	EAP	A	980	3	-	0/0/4/5	0/0/0/0
10	FCI	B	1050	-	-	0/57/102/116	0/0/6/6
5	FTT	B	900	3	-	0/14/14/15	0/0/0/0
5	FTT	B	901	-	-	0/14/14/15	0/0/0/0
5	FTT	B	902	3,7	-	0/14/14/15	0/0/0/0
5	FTT	B	903	8	-	0/14/14/15	0/0/0/0
6	DPO	B	910	-	-	0/2/5/6	0/0/0/0
7	DAO	B	930	5	-	0/10/10/11	0/0/0/0
8	MYR	B	940	5	-	0/12/12/13	0/0/0/0
4	PO4	B	950	3	-	0/0/0/0	0/0/0/0
9	EAP	B	980	3	-	0/0/4/5	0/0/0/0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	910	DPO	P2-O4	-5.66	1.50	1.60
6	A	910	DPO	P2-O4	-5.65	1.50	1.60
9	A	980	EAP	P-O4	-4.47	1.49	1.63
9	B	980	EAP	P-O4	-4.46	1.49	1.63
6	B	910	DPO	P1-O4	-4.28	1.50	1.63

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	A	1050	FCI	O17-C8-N4	-2.11	119.33	122.97
10	B	1050	FCI	O17-C8-N4	-2.09	119.36	122.97
5	B	900	FTT	C10-C9-C8	2.01	124.82	114.45
8	B	940	MYR	O2-C1-C2	2.06	125.23	111.65
10	B	1050	FCI	C34-N8-C9	2.06	130.56	125.65

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

17 monomers are involved in 138 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
10	A	1050	FCI	2	0
5	A	900	FTT	12	0
5	A	901	FTT	9	0
5	A	902	FTT	6	0
5	A	903	FTT	3	0
6	A	910	DPO	8	0
7	A	930	DAO	5	0
9	A	980	EAP	2	0
10	B	1050	FCI	1	0
5	B	900	FTT	17	0
5	B	901	FTT	39	0
5	B	902	FTT	39	0
5	B	903	FTT	64	0
6	B	910	DPO	8	0
7	B	930	DAO	7	0
8	B	940	MYR	4	0
9	B	980	EAP	3	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	A	702/725 (96%)	-0.16	3 (0%) 92 92	24, 61, 110, 177	0
1	B	691/725 (95%)	0.45	66 (9%) 9 8	44, 115, 186, 210	0
2	C	78/229 (34%)	0.76	11 (14%) 3 3	43, 112, 179, 209	0
2	D	78/229 (34%)	2.38	38 (48%) 0 0	102, 187, 210, 210	0
All	All	1549/1908 (81%)	0.29	118 (7%) 15 14	24, 88, 186, 210	0

The worst 5 of 118 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	178	VAL	7.9
1	B	255	PRO	7.8
2	D	225	VAL	7.6
2	D	159	ASN	7.3
1	B	250	GLU	6.9

### 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](#)

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( $\text{\AA}^2$ )	Q<0.9
3	KDO	A	970	15/16	0.92	0.17	-1.34	65,72,98,98	0
3	KDO	B	970	15/16	0.93	0.21	-1.96	65,72,98,98	0
3	KDO	B	971	15/16	0.77	0.24	-	73,80,98,98	0
3	GMH	A	960	13/14	0.88	0.14	-	62,67,79,81	0
3	GMH	B	960	13/14	0.88	0.19	-	62,67,79,81	0
3	GCN	A	921	10/11	0.91	0.14	-	49,60,64,74	0
3	GCN	B	921	10/11	0.89	0.08	-	49,60,64,74	0
3	GCN	A	920	10/11	0.88	0.23	-	62,74,79,87	0
3	GCN	B	920	10/11	0.94	0.11	-	62,74,79,87	0
3	KDO	A	971	15/16	0.88	0.20	-	73,80,98,98	0

## 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. 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( $\text{\AA}^2$ )	Q<0.9
8	MYR	B	940	15/16	0.72	1.04	12.73	69,86,94,95	0
7	DAO	A	930	13/14	0.79	0.60	6.18	69,75,82,86	0
5	FTT	A	900	16/17	0.80	0.46	5.56	76,87,96,102	0
5	FTT	A	902	16/17	0.88	0.47	5.01	63,70,78,79	0
5	FTT	B	900	16/17	0.83	0.41	4.65	84,95,98,102	0
5	FTT	A	901	16/17	0.79	0.43	3.84	63,70,78,79	0
5	FTT	B	902	16/17	0.84	0.57	3.46	63,70,78,79	0
10	FCI	A	1050	46/51	0.87	0.28	-0.45	44,57,70,81	0
10	FCI	B	1050	46/51	0.81	0.25	-1.25	44,57,70,81	0
4	PO4	B	950	4/5	0.95	0.13	-	71,74,76,81	0
5	FTT	B	901	16/17	0.63	0.73	-	63,70,78,79	0
5	FTT	B	903	16/17	0.42	0.88	-	75,88,92,97	0
4	PO4	A	950	4/5	0.94	0.10	-	71,74,76,81	0
9	EAP	A	980	7/8	0.89	0.14	-	81,87,93,95	0
6	DPO	B	910	8/9	0.69	0.22	-	62,80,93,95	0
9	EAP	B	980	7/8	0.89	0.16	-	81,87,93,95	0
6	DPO	A	910	8/9	0.70	0.22	-	62,80,93,95	0
7	DAO	B	930	13/14	0.83	0.40	-	69,75,82,86	0
5	FTT	A	903	16/17	0.74	0.47	-	75,88,92,97	0



## 6.5 Other polymers [i](#)

There are no such residues in this entry.