



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 2, 2025 – 04:06 pm BST

PDB ID : 9HJO / pdb\_00009hjo  
Title : FANCM-FAAP24-dsDNA complex  
Authors : Coulthard-Graf, R.J.; Briggs, D.C.; McDonald, N.Q.; Deans, A.  
Deposited on : 2024-11-29  
Resolution : 2.40 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

<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	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

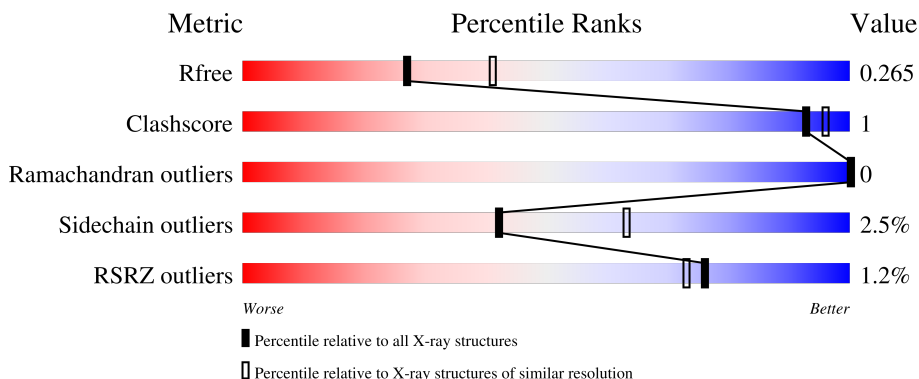
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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.



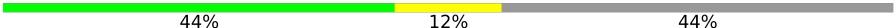


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	164625	4642 (2.40-2.40)
Clashscore	180529	5218 (2.40-2.40)
Ramachandran outliers	177936	5158 (2.40-2.40)
Sidechain outliers	177891	5159 (2.40-2.40)
RSRZ outliers	164620	4642 (2.40-2.40)

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	A	234	<div> <div>81%</div> <div>8%</div> <div>11%</div> </div>
1	C	234	<div> <div>83%</div> <div>6%</div> <div>11%</div> </div>
2	B	215	<div> <div>2%</div> <div>86%</div> <div>5%</div> <div>9%</div> </div>
2	D	215	<div> <div>85%</div> <div>5%</div> <div>9%</div> </div>
3	F	25	<div> <div>36%</div> <div>12%</div> <div>52%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	H	25	 44% 12% 44%
4	G	25	 32% 64%
4	I	25	 20% 8% 72%

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 14364 atoms, of which 6933 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 Fanconi anemia group M protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	208	Total	C	H	N	O	S	0	2	0
			3344	1057	1656	292	321	18			
1	C	209	Total	C	H	N	O	S	0	4	0
			3401	1076	1689	296	322	18			

- Molecule 2 is a protein called Fanconi anemia core complex-associated protein 24.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	196	Total	C	H	N	O	S	0	1	0
			3117	983	1582	267	278	7			
2	D	195	Total	C	H	N	O	S	0	0	0
			3077	972	1558	263	277	7			

- Molecule 3 is a DNA chain called DNA (25-MER).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	F	12	Total	C	H	N	O	P	0	0	0
			371	115	128	44	72	12			
3	H	14	Total	C	H	N	O	P	0	0	0
			434	134	150	52	84	14			

- Molecule 4 is a DNA chain called DNA (25-MER).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	G	9	Total	C	H	N	O	P	0	0	0
			282	88	95	35	55	9			
4	I	7	Total	C	H	N	O	P	0	0	0
			220	68	75	28	42	7			

- Molecule 5 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	2	Total 2	Mg 2	0	0
5	D	2	Total 2	Mg 2	0	0

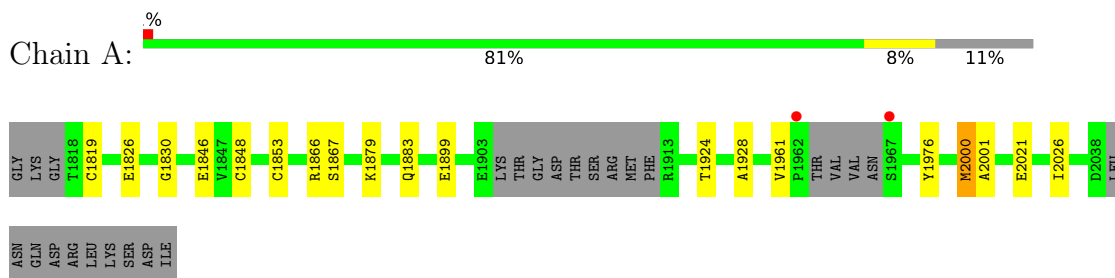
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	42	Total 42	O 42	0	0
6	B	15	Total 15	O 15	0	0
6	C	47	Total 47	O 47	0	0
6	D	9	Total 9	O 9	0	0
6	H	1	Total 1	O 1	0	0

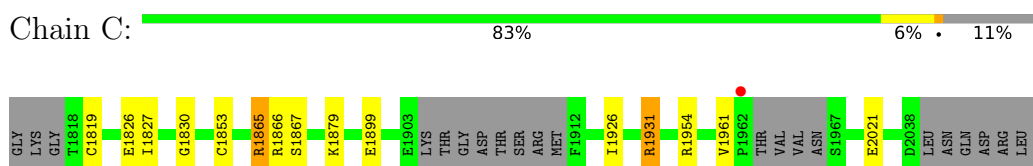
### 3 Residue-property plots [i](#)

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.

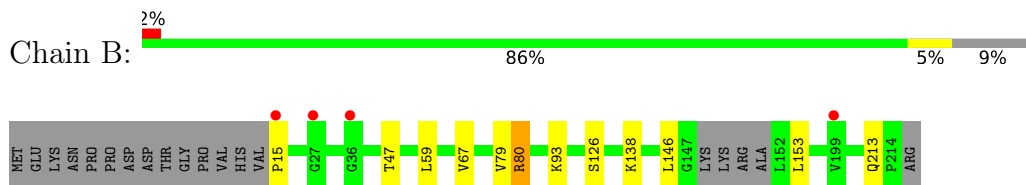
- Molecule 1: Fanconi anemia group M protein



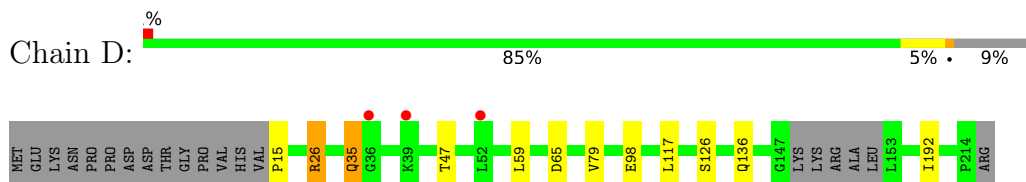
- Molecule 1: Fanconi anemia group M protein



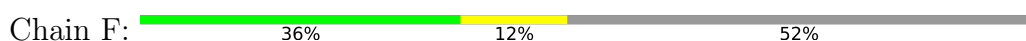
- Molecule 2: Fanconi anemia core complex-associated protein 24



- Molecule 2: Fanconi anemia core complex-associated protein 24

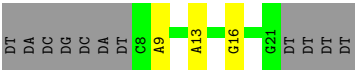


- Molecule 3: DNA (25-MER)

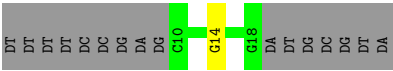




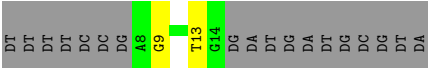
• Molecule 3: DNA (25-MER)



• Molecule 4: DNA (25-MER)



• Molecule 4: DNA (25-MER)



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.01Å 69.97Å 86.76Å 91.63° 89.93° 115.95°	Depositor
Resolution (Å)	86.72 – 2.40 86.72 – 2.40	Depositor EDS
% Data completeness (in resolution range)	96.1 (86.72-2.40) 95.6 (86.72-2.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.64 (at 2.40Å)	Xtriage
Refinement program	REFMAC 5.8.0430 (refmacat 0.4.88)	Depositor
R, $R_{free}$	0.214 , 0.265 0.215 , 0.265	Depositor DCC
$R_{free}$ test set	1979 reflections (4.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	49.7	Xtriage
Anisotropy	0.378	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 53.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.459 for h,-h-k,-l 0.011 for -h,-k,l 0.009 for -h,h+k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	14364	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.66% 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: CME, MG

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	0.64	0/1672	1.22	4/2243 (0.2%)
1	C	0.66	0/1703	1.23	2/2284 (0.1%)
2	B	0.61	0/1562	1.14	1/2111 (0.0%)
2	D	0.61	0/1543	1.17	2/2086 (0.1%)
3	F	0.46	0/271	1.25	4/415 (1.0%)
3	H	0.45	0/317	1.12	3/486 (0.6%)
4	G	0.47	0/209	1.22	1/321 (0.3%)
4	I	0.50	0/162	1.18	2/248 (0.8%)
All	All	0.61	0/7439	1.19	19/10194 (0.2%)

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	A	0	1
1	C	0	5
2	B	0	2
2	D	0	1
All	All	0	9

There are no bond length outliers.

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	I	9	DG	O3'-P-O5'	-7.38	92.93	104.00
2	D	98	GLU	CB-CG-CD	7.21	124.86	112.60
3	F	16	DG	O3'-P-O5'	-6.78	93.83	104.00
4	I	13	DT	O3'-P-O5'	-6.76	93.86	104.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	14	DG	O3'-P-O5'	-6.75	93.88	104.00

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1866	ARG	Sidechain
2	B	80[A]	ARG	Sidechain
2	B	80[B]	ARG	Sidechain
1	C	1865	ARG	Sidechain
1	C	1866	ARG	Sidechain

## 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	1688	1656	1702	8	0
1	C	1712	1689	1735	9	0
2	B	1535	1582	1608	2	0
2	D	1519	1558	1584	4	0
3	F	243	128	135	0	0
3	H	284	150	157	0	0
4	G	187	95	102	0	0
4	I	145	75	79	0	0
5	B	2	0	0	0	0
5	D	2	0	0	0	0
6	A	42	0	0	0	0
6	B	15	0	0	0	0
6	C	47	0	0	0	0
6	D	9	0	0	0	0
6	H	1	0	0	0	0
All	All	7431	6933	7102	20	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 1.

The worst 5 of 20 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1976:TYR:HE1	1:A:2000:MET:HG2	1.64	0.60
1:C:1931[B]:ARG:NH1	1:C:1931[B]:ARG:HG3	2.19	0.58
1:A:2000:MET:HG3	1:A:2001:ALA:N	2.19	0.56
1:C:1926:ILE:HD11	2:D:117:LEU:HD21	1.93	0.50
2:D:192:ILE:HD12	2:D:192:ILE:H	1.78	0.49

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	200/234 (86%)	194 (97%)	6 (3%)	0	100	100
1	C	203/234 (87%)	198 (98%)	5 (2%)	0	100	100
2	B	193/215 (90%)	189 (98%)	4 (2%)	0	100	100
2	D	191/215 (89%)	187 (98%)	4 (2%)	0	100	100
All	All	787/898 (88%)	768 (98%)	19 (2%)	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	A	190/211 (90%)	187 (98%)	3 (2%)	58	76
1	C	193/211 (92%)	192 (100%)	1 (0%)	86	94

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	172/188 (92%)	163 (95%)	9 (5%)	19	34
2	D	170/188 (90%)	164 (96%)	6 (4%)	31	51
All	All	725/798 (91%)	706 (97%)	19 (3%)	42	62

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	26	ARG
2	D	126	SER
2	D	136	GLN
2	D	47	THR
2	B	126	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	1873	ASN
1	C	1992	HIS
2	D	186	GLN
1	C	2006	GLN
2	B	136	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

8 non-standard protein/DNA/RNA residues 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
1	CME	C	1991	1	8,9,10	0.44	0	5,9,11	0.43	0
1	CME	C	1819	1	8,9,10	0.48	0	5,9,11	0.75	0
1	CME	A	1819	1	8,9,10	0.44	0	5,9,11	0.95	0
1	CME	A	1848	1	8,9,10	0.46	0	5,9,11	0.51	0
1	CME	C	1848	1	8,9,10	0.48	0	5,9,11	0.24	0
1	CME	A	1853	1	8,9,10	0.31	0	5,9,11	1.20	1 (20%)
1	CME	C	1853	1	8,9,10	0.31	0	5,9,11	1.09	1 (20%)
1	CME	A	1991	1	8,9,10	0.40	0	5,9,11	0.49	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
1	CME	C	1991	1	-	2/5/8/10	-
1	CME	C	1819	1	-	1/5/8/10	-
1	CME	A	1819	1	-	1/5/8/10	-
1	CME	A	1848	1	-	2/5/8/10	-
1	CME	C	1848	1	-	2/5/8/10	-
1	CME	A	1853	1	-	1/5/8/10	-
1	CME	C	1853	1	-	1/5/8/10	-
1	CME	A	1991	1	-	2/5/8/10	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1853	CME	CZ-CE-SD	-2.26	105.52	113.37
1	C	1853	CME	CZ-CE-SD	-2.04	106.30	113.37

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	1819	CME	SD-CE-CZ-OH
1	A	1848	CME	SD-CE-CZ-OH
1	A	1991	CME	CE-SD-SG-CB
1	C	1819	CME	SD-CE-CZ-OH
1	C	1848	CME	SD-CE-CZ-OH

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	C	1819	CME	1	0
1	A	1819	CME	1	0
1	A	1848	CME	1	0

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	204/234 (87%)	-0.70	2 (0%) 79 76	24, 46, 93, 133	2 (0%)
1	C	205/234 (87%)	-0.89	1 (0%) 87 85	20, 46, 97, 130	4 (1%)
2	B	196/215 (91%)	-0.27	4 (2%) 64 61	35, 76, 129, 150	1 (0%)
2	D	195/215 (90%)	-0.29	3 (1%) 71 68	33, 78, 129, 163	0
3	F	12/25 (48%)	-0.98	0 100 100	62, 98, 134, 141	0
3	H	14/25 (56%)	-0.79	0 100 100	65, 107, 138, 142	0
4	G	9/25 (36%)	-0.40	0 100 100	93, 101, 145, 153	0
4	I	7/25 (28%)	-0.72	0 100 100	97, 99, 138, 147	0
All	All	842/998 (84%)	-0.55	10 (1%) 76 73	20, 61, 126, 163	7 (0%)

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1962	PRO	3.2
1	A	1967	SER	2.9
2	D	52	LEU	2.9
2	D	36	GLY	2.7
2	B	15	PRO	2.5

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

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
1	CME	C	1819	10/11	0.98	0.08	35,52,95,95	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
1	CME	A	1848	10/11	0.99	0.05	40,42,67,74	0
1	CME	A	1853	10/11	0.99	0.04	36,44,61,65	0
1	CME	A	1819	10/11	0.99	0.07	36,54,101,101	0
1	CME	C	1848	10/11	0.99	0.06	37,42,66,76	0
1	CME	C	1853	10/11	0.99	0.06	39,46,65,74	0
1	CME	C	1991	10/11	0.99	0.04	36,43,71,72	0
1	CME	A	1991	10/11	1.00	0.04	39,47,79,87	0

### 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( $\text{\AA}^2$ )	Q<0.9
5	MG	B	301	1/1	0.98	0.04	52,52,52,52	0
5	MG	B	302	1/1	0.98	0.07	73,73,73,73	0
5	MG	D	301	1/1	0.98	0.07	69,69,69,69	0
5	MG	D	302	1/1	0.98	0.12	50,50,50,50	0

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

There are no such residues in this entry.