



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 7, 2017 – 07:56 AM EDT

PDB ID : 5NJ4  
Title : From macrocrystals to microcrystals: a strategy for membrane protein serial crystallography  
Authors : Dods, R.; Baath, P.; Branden, G.; Neutze, R.  
Deposited on : unknown  
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

<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-20029824  
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-20029824

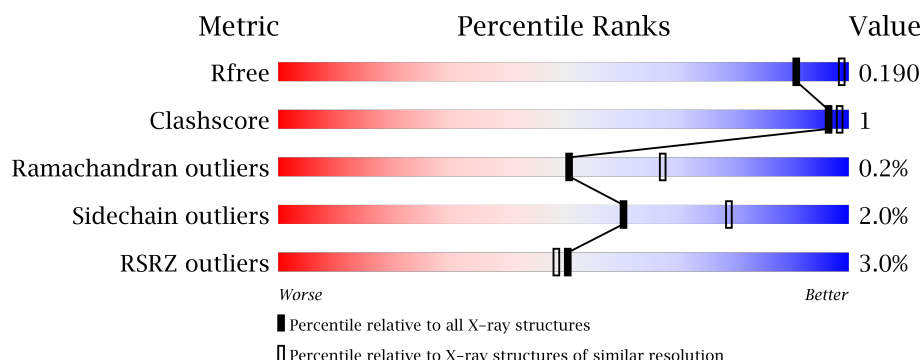
# 1 Overall quality at a glance

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))
$R_{free}$	100719	3166 (2.40-2.40)
Clashscore	112137	3674 (2.40-2.40)
Ramachandran outliers	110173	3616 (2.40-2.40)
Sidechain outliers	110143	3617 (2.40-2.40)
RSRZ outliers	101464	3195 (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 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	C	336	<div> <div>2%</div> <div>96%</div> <div>..</div> </div>
2	H	258	<div> <div>6%</div> <div>97%</div> <div>..</div> </div>
3	L	273	<div> <div>3%</div> <div>98%</div> <div>.</div> </div>
4	M	323	<div> <div>2%</div> <div>96%</div> <div>.</div> </div>

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
14	NS5	M	405	-	-	-	X
6	DGA	C	405	-	-	-	X
7	SO4	H	704	-	-	-	X
7	SO4	M	412	-	-	-	X
8	LDA	H	701	-	-	-	X
8	LDA	H	706	-	-	-	X
8	LDA	M	414	-	-	-	X
8	LDA	M	415	-	-	-	X
8	LDA	M	416	-	-	-	X
9	HTO	H	708	-	-	-	X
9	HTO	H	709	-	-	-	X
9	HTO	L	305	-	-	-	X

## 2 Entry composition

There are 15 unique types of molecules in this entry. The entry contains 10402 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 Photosynthetic reaction center cytochrome c subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	332	Total	C	N	O	S	0	0	0
			2602	1640	466	478	18			

- Molecule 2 is a protein called Reaction center protein H chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	258	Total	C	N	O	S	0	0	0
			2018	1292	344	380	2			

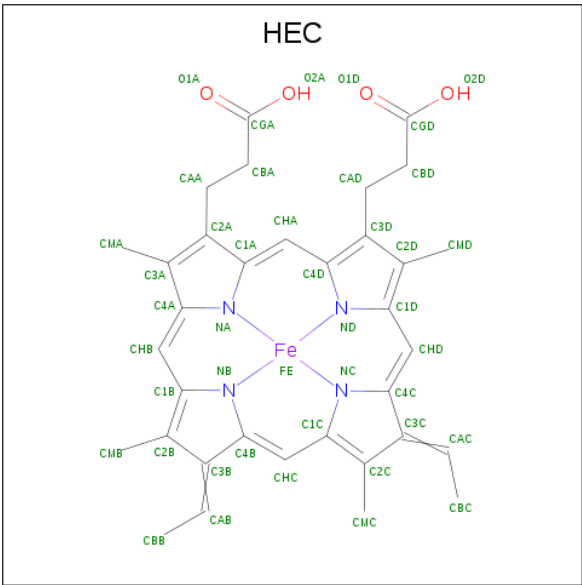
- Molecule 3 is a protein called Reaction center protein L chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	273	Total	C	N	O	S	0	1	0
			2172	1460	350	355	7			

- Molecule 4 is a protein called Reaction center protein M chain.

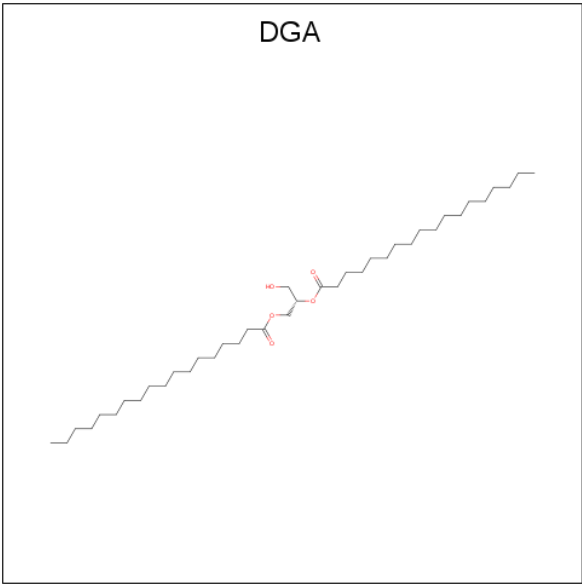
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	M	323	Total	C	N	O	S	0	0	0
			2555	1702	419	423	11			

- Molecule 5 is HEME C (three-letter code: HEC) (formula:  $C_{34}H_{34}FeN_4O_4$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	C	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
5	C	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
5	C	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
5	C	1	Total 43	C 34	Fe 1	N 4	O 4	0	0

- Molecule 6 is DIACYL GLYCEROL (three-letter code: DGA) (formula: C<sub>39</sub>H<sub>76</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	C	O	0	0
			37	33	4		

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



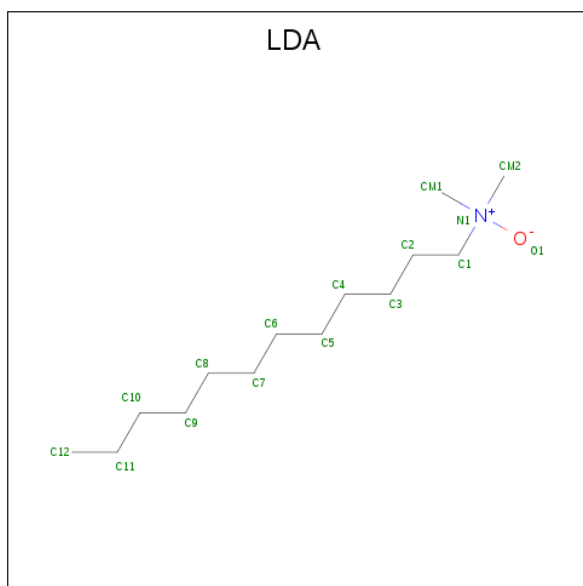
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	C	1	Total	O	S	0	0
			5	4	1		
7	C	1	Total	O	S	0	0
			5	4	1		
7	C	1	Total	O	S	0	0
			5	4	1		
7	H	1	Total	O	S	0	0
			5	4	1		
7	H	1	Total	O	S	0	0
			5	4	1		
7	H	1	Total	O	S	0	0
			5	4	1		
7	H	1	Total	O	S	0	0
			5	4	1		
7	M	1	Total	O	S	0	0
			5	4	1		
7	M	1	Total	O	S	0	0
			5	4	1		
7	M	1	Total	O	S	0	0
			5	4	1		
7	M	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

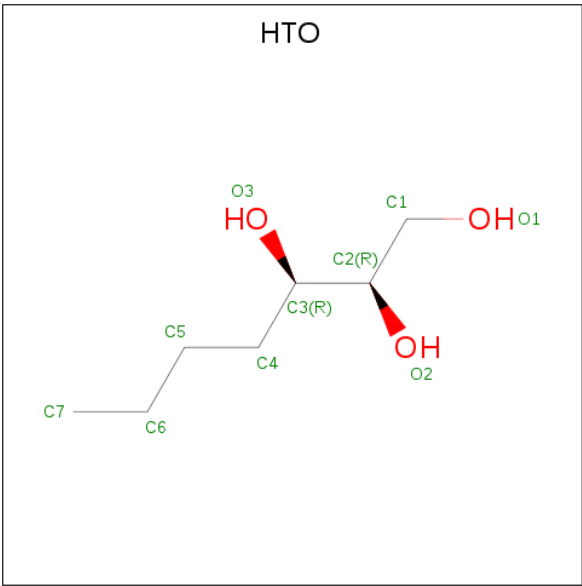
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	M	1	Total	O	S	0	0
			5	4	1		
7	M	1	Total	O	S	0	0
			5	4	1		
7	M	1	Total	O	S	0	0
			5	4	1		
7	M	1	Total	O	S	0	0
			5	4	1		

- Molecule 8 is LAURYL DIMETHYLAMINE-N-OXIDE (three-letter code: LDA) (formula:  $C_{14}H_{31}NO$ ).



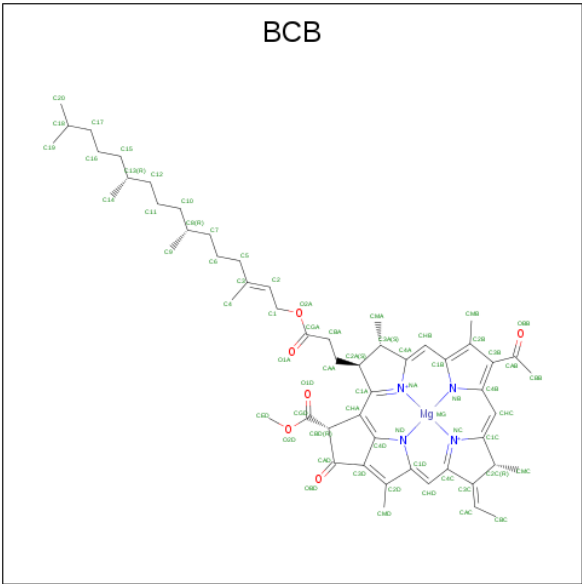
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	H	1	Total	C	N	O	0	0
			16	14	1	1		
8	H	1	Total	C	N	O	0	0
			16	14	1	1		
8	H	1	Total	C	N	O	0	0
			16	14	1	1		
8	L	1	Total	C	N	O	0	0
			16	14	1	1		
8	M	1	Total	C	N	O	0	0
			16	14	1	1		
8	M	1	Total	C	N	O	0	0
			16	14	1	1		
8	M	1	Total	C	N	O	0	0
			16	14	1	1		

- Molecule 9 is HEPTANE-1,2,3-TRIOL (three-letter code: HTO) (formula: C<sub>7</sub>H<sub>16</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	H	1	Total	C	O	0	0
			10	7	3		
9	H	1	Total	C	O	0	0
			10	7	3		
9	L	1	Total	C	O	0	0
			10	7	3		

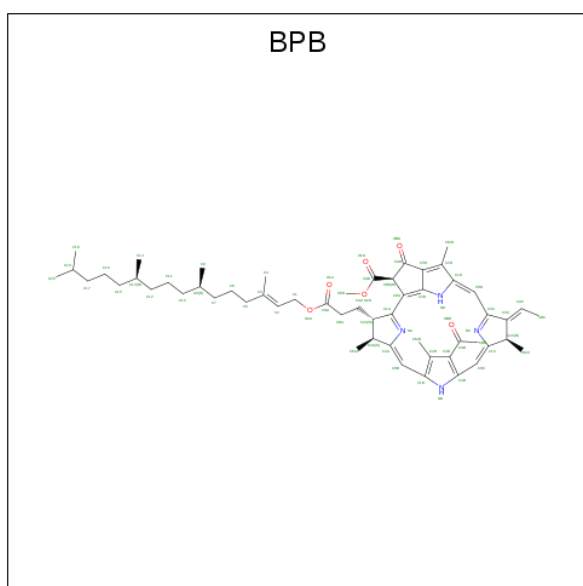
- Molecule 10 is BACTERIOCHLOROPHYLL B (three-letter code: BCB) (formula: C<sub>55</sub>H<sub>72</sub>MgN<sub>4</sub>O<sub>6</sub>).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
10	L	1	Total	C	Mg	N	O	
			66	55	1	4	6	
10	L	1	Total	C	Mg	N	O	
			66	55	1	4	6	
10	M	1	Total	C	Mg	N	O	
			66	55	1	4	6	
10	M	1	Total	C	Mg	N	O	
			66	55	1	4	6	

- Molecule 11 is BACTERIOPHEOPHYTIN B (three-letter code: BPB) (formula:  $C_{55}H_{74}N_4O_6$ ).

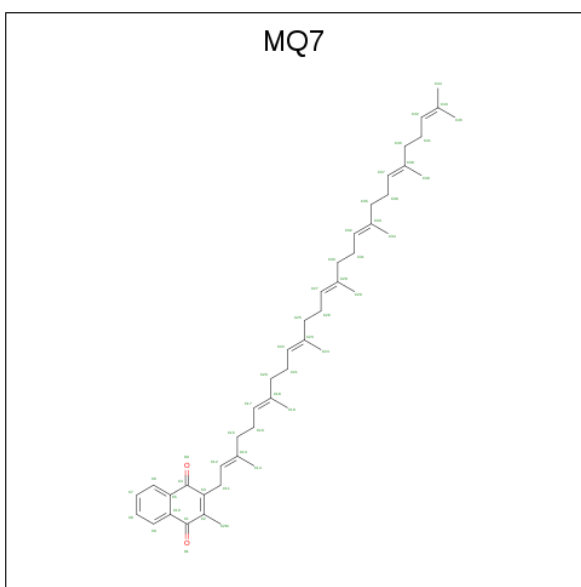


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
11	L	1	Total	C	N	O		
			65	55	4	6		
11	M	1	Total	C	N	O		
			65	55	4	6		

- Molecule 12 is FE (II) ION (three-letter code: FE2) (formula: Fe).

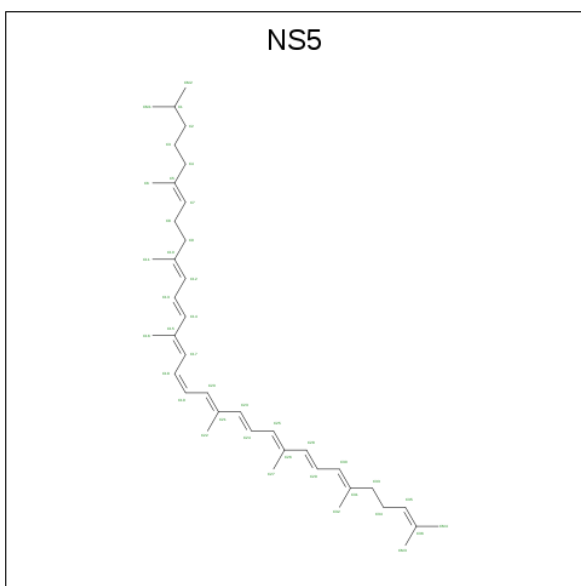
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	L	1	Total	Fe		
			1	1	0	0

- Molecule 13 is MENAQUINONE-7 (three-letter code: MQ7) (formula:  $C_{46}H_{64}O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
13	M	1	Total	C	O	0	0
			48	46	2		

- Molecule 14 is 15-cis-1,2-dihydroneurosporene (three-letter code: NS5) (formula: C<sub>40</sub>H<sub>60</sub>).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	M	1	Total	C	0	0
			40	40		

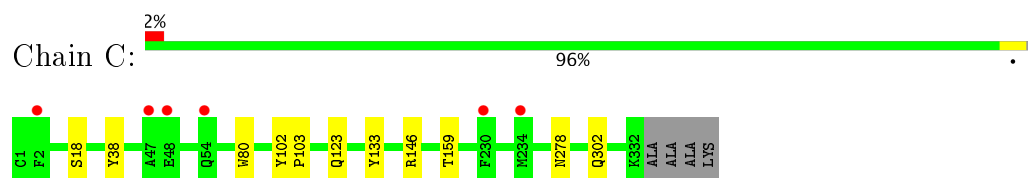
- Molecule 15 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	C	58	Total 58	O 58	0	0
15	H	23	Total 23	O 23	0	0
15	L	24	Total 24	O 24	0	0
15	M	41	Total 41	O 41	0	0

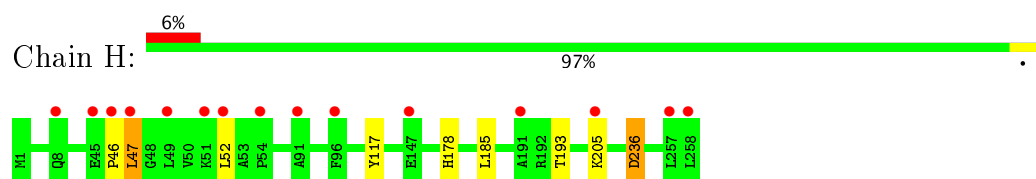
### 3 Residue-property plots [i](#)

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

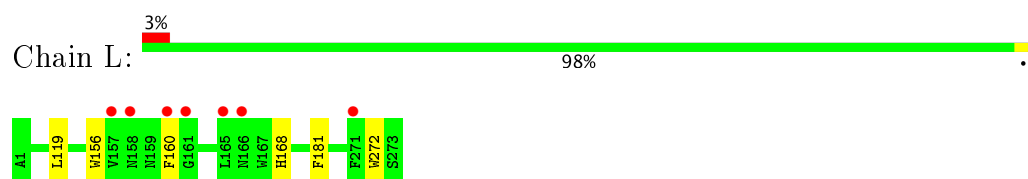
- Molecule 1: Photosynthetic reaction center cytochrome c subunit



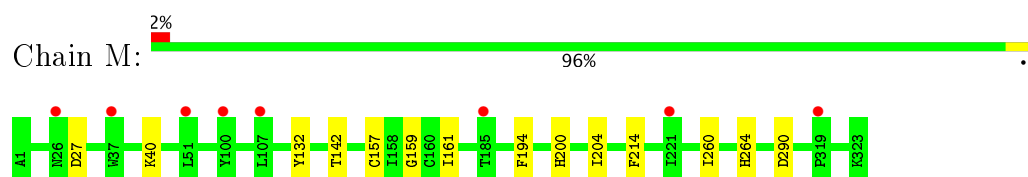
- Molecule 2: Reaction center protein H chain



- Molecule 3: Reaction center protein L chain



- Molecule 4: Reaction center protein M chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	226.50 Å   226.50 Å   113.90 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	35.70 – 2.40 35.55 – 2.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (35.70-2.40) 100.0 (35.55-2.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.55 (at 2.39 Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.161   ,   0.186 0.168   ,   0.190	Depositor DCC
$R_{free}$ test set	5711 reflections (5.21%)	DCC
Wilson B-factor (Å <sup>2</sup> )	73.2	Xtriage
Anisotropy	0.005	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 62.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	10402	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	82.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.81% 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: LDA, DGA, HTO, BCB, BPB, FE2, SO4, MQ7, HEC, FME, NS5

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	C	0.40	0/2669	0.66	0/3637
2	H	0.43	0/2055	0.66	0/2807
3	L	0.43	0/2267	0.60	0/3095
4	M	0.43	0/2659	0.59	0/3637
All	All	0.42	0/9650	0.62	0/13176

There are no bond length outliers.

There are no bond angle outliers.

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	C	2602	0	2578	6	0
2	H	2018	0	2020	1	0
3	L	2172	0	2097	4	0
4	M	2555	0	2452	5	0
5	C	172	0	120	3	0
6	C	37	0	58	0	0
7	C	15	0	0	0	0
7	H	20	0	0	0	0
7	M	40	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	H	48	0	93	0	0
8	L	16	0	31	0	0
8	M	48	0	93	0	0
9	H	20	0	32	0	0
9	L	10	0	16	0	0
10	L	132	0	144	4	0
10	M	132	0	144	1	0
11	L	65	0	74	1	0
11	M	65	0	74	3	0
12	L	1	0	0	0	0
13	M	48	0	64	0	0
14	M	40	0	60	1	0
15	C	58	0	0	1	0
15	H	23	0	0	0	0
15	L	24	0	0	0	0
15	M	41	0	0	0	0
All	All	10402	0	10150	24	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 24 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:M:159:GLY:HA3	14:M:405:NS5:H272	1.73	0.70
11:L:303:BPB:HBBB	11:L:303:BPB:HMB	1.75	0.69
11:M:404:BPB:HBBB	11:M:404:BPB:HHC	1.78	0.66
10:M:402:BCB:HBB2	10:M:402:BCB:HHC	1.82	0.62
5:C:401:HEC:HBC3	5:C:401:HEC:HMC1	1.85	0.58

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	C	330/336 (98%)	324 (98%)	6 (2%)	0	100	100
2	H	256/258 (99%)	249 (97%)	5 (2%)	2 (1%)	22	33
3	L	272/273 (100%)	265 (97%)	7 (3%)	0	100	100
4	M	321/323 (99%)	311 (97%)	10 (3%)	0	100	100
All	All	1179/1190 (99%)	1149 (98%)	28 (2%)	2 (0%)	51	67

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	46	PRO
2	H	47	LEU

### 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	C	281/282 (100%)	277 (99%)	4 (1%)	71	86
2	H	212/212 (100%)	205 (97%)	7 (3%)	43	64
3	L	219/218 (100%)	216 (99%)	3 (1%)	71	86
4	M	249/249 (100%)	244 (98%)	5 (2%)	60	79
All	All	961/961 (100%)	942 (98%)	19 (2%)	60	79

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

Mol	Chain	Res	Type
2	H	193	THR
2	H	236	ASP
4	M	40	LYS
2	H	185	LEU
4	M	194	PHE



Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	123	GLN
1	C	302	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\# Z  > 2$
2	FME	H	1	2	9,9,10	0.80	0	7,9,11	3.40	4 (57%)

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
2	FME	H	1	2	-	0/6/9/11	0/0/0/0

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	1	FME	CA-N-CN	-7.77	110.88	122.82
2	H	1	FME	O-C-CA	-2.09	120.28	125.15
2	H	1	FME	O1-CN-N	-2.07	119.42	125.20
2	H	1	FME	CE-SD-CG	2.94	110.90	100.35

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 39 ligands modelled in this entry, 1 is monoatomic - leaving 38 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	$\# Z  > 2$
5	HEC	C	401	1	28,50,50	1.63	2 (7%)	16,82,82	2.33	4 (25%)
5	HEC	C	402	1	28,50,50	1.46	2 (7%)	16,82,82	1.98	4 (25%)
5	HEC	C	403	1	28,50,50	1.55	3 (10%)	16,82,82	1.73	4 (25%)
5	HEC	C	404	1	28,50,50	1.48	4 (14%)	16,82,82	1.81	4 (25%)
6	DGA	C	405	1	36,36,43	1.20	2 (5%)	38,38,45	1.22	4 (10%)
7	SO4	C	406	-	4,4,4	0.45	0	6,6,6	0.30	0
7	SO4	C	407	-	4,4,4	0.42	0	6,6,6	0.13	0
7	SO4	C	408	-	4,4,4	0.41	0	6,6,6	0.14	0
8	LDA	H	701	-	13,15,15	2.18	1 (7%)	14,17,17	0.71	0
7	SO4	H	702	-	4,4,4	0.39	0	6,6,6	0.26	0
7	SO4	H	703	-	4,4,4	0.42	0	6,6,6	0.08	0
7	SO4	H	704	-	4,4,4	0.36	0	6,6,6	0.18	0
7	SO4	H	705	-	4,4,4	0.42	0	6,6,6	0.17	0
8	LDA	H	706	-	13,15,15	2.33	1 (7%)	14,17,17	0.86	1 (7%)
8	LDA	H	707	-	13,15,15	2.41	1 (7%)	14,17,17	0.85	0
9	HTO	H	708	-	9,9,9	0.51	0	9,10,10	0.54	0
9	HTO	H	709	-	9,9,9	0.39	0	9,10,10	0.53	0
10	BCB	L	301	-	63,74,74	3.36	23 (36%)	50,115,115	2.23	17 (34%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
10	BCB	L	302	-	63,74,74	3.36	27 (42%)	50,115,115	2.10	16 (32%)
11	BPB	L	303	-	63,70,70	2.67	15 (23%)	67,101,101	2.07	14 (20%)
8	LDA	L	304	-	13,15,15	2.39	1 (7%)	14,17,17	0.50	0
9	HTO	L	305	-	9,9,9	0.49	0	9,10,10	0.83	0
13	MQ7	M	401	-	49,49,49	1.41	2 (4%)	61,63,63	1.08	4 (6%)
10	BCB	M	402	-	63,74,74	3.36	27 (42%)	50,115,115	1.96	12 (24%)
10	BCB	M	403	-	63,74,74	3.25	24 (38%)	50,115,115	2.18	13 (26%)
11	BPB	M	404	-	63,70,70	2.84	17 (26%)	67,101,101	2.01	14 (20%)
14	NS5	M	405	-	39,39,39	1.41	1 (2%)	44,46,46	1.85	11 (25%)
7	SO4	M	406	-	4,4,4	0.30	0	6,6,6	0.25	0
7	SO4	M	407	-	4,4,4	0.36	0	6,6,6	0.06	0
7	SO4	M	408	-	4,4,4	0.28	0	6,6,6	0.47	0
7	SO4	M	409	-	4,4,4	0.47	0	6,6,6	0.18	0
7	SO4	M	410	-	4,4,4	0.40	0	6,6,6	0.17	0
7	SO4	M	411	-	4,4,4	0.44	0	6,6,6	0.16	0
7	SO4	M	412	-	4,4,4	0.43	0	6,6,6	0.18	0
7	SO4	M	413	-	4,4,4	0.42	0	6,6,6	0.07	0
8	LDA	M	414	-	13,15,15	2.41	1 (7%)	14,17,17	0.55	0
8	LDA	M	415	-	13,15,15	2.37	1 (7%)	14,17,17	0.54	0
8	LDA	M	416	-	13,15,15	2.30	1 (7%)	14,17,17	0.40	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
5	HEC	C	401	1	-	0/6/54/54	0/0/8/8
5	HEC	C	402	1	-	0/6/54/54	0/0/8/8
5	HEC	C	403	1	-	0/6/54/54	0/0/8/8
5	HEC	C	404	1	-	0/6/54/54	0/0/8/8
6	DGA	C	405	1	-	0/37/37/45	0/0/0/0
7	SO4	C	406	-	-	0/0/0/0	0/0/0/0
7	SO4	C	407	-	-	0/0/0/0	0/0/0/0
7	SO4	C	408	-	-	0/0/0/0	0/0/0/0
8	LDA	H	701	-	-	0/13/13/13	0/0/0/0
7	SO4	H	702	-	-	0/0/0/0	0/0/0/0
7	SO4	H	703	-	-	0/0/0/0	0/0/0/0
7	SO4	H	704	-	-	0/0/0/0	0/0/0/0
7	SO4	H	705	-	-	0/0/0/0	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	LDA	H	706	-	-	0/13/13/13	0/0/0/0
8	LDA	H	707	-	-	0/13/13/13	0/0/0/0
9	HTO	H	708	-	-	0/10/10/10	0/0/0/0
9	HTO	H	709	-	-	0/10/10/10	0/0/0/0
10	BCB	L	301	-	-	0/41/177/177	0/0/9/9
10	BCB	L	302	-	-	0/41/177/177	0/0/9/9
11	BPB	L	303	-	-	0/47/105/105	0/1/6/6
8	LDA	L	304	-	-	0/13/13/13	0/0/0/0
9	HTO	L	305	-	-	0/10/10/10	0/0/0/0
13	MQ7	M	401	-	-	0/41/61/61	0/2/2/2
10	BCB	M	402	-	-	0/41/177/177	0/0/9/9
10	BCB	M	403	-	-	0/41/177/177	0/0/9/9
11	BPB	M	404	-	-	0/47/105/105	0/1/6/6
14	NS5	M	405	-	-	0/43/43/43	0/0/0/0
7	SO4	M	406	-	-	0/0/0/0	0/0/0/0
7	SO4	M	407	-	-	0/0/0/0	0/0/0/0
7	SO4	M	408	-	-	0/0/0/0	0/0/0/0
7	SO4	M	409	-	-	0/0/0/0	0/0/0/0
7	SO4	M	410	-	-	0/0/0/0	0/0/0/0
7	SO4	M	411	-	-	0/0/0/0	0/0/0/0
7	SO4	M	412	-	-	0/0/0/0	0/0/0/0
7	SO4	M	413	-	-	0/0/0/0	0/0/0/0
8	LDA	M	414	-	-	0/13/13/13	0/0/0/0
8	LDA	M	415	-	-	0/13/13/13	0/0/0/0
8	LDA	M	416	-	-	0/13/13/13	0/0/0/0

The worst 5 of 156 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	L	302	BCB	C3D-C4D	-13.89	1.38	1.54
10	L	301	BCB	C3D-C4D	-13.66	1.39	1.54
10	M	402	BCB	C3D-C4D	-13.40	1.39	1.54
10	M	403	BCB	C3D-C4D	-12.43	1.40	1.54
8	M	414	LDA	O1-N1	-8.57	1.25	1.42

The worst 5 of 122 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	L	303	BPB	CBC-CAC-C3C	-7.33	110.85	127.00
11	L	303	BPB	C2C-C3C-C4C	-6.91	100.66	107.35
11	M	404	BPB	C2C-C3C-C4C	-6.48	101.08	107.35
5	C	401	HEC	CBA-CAA-C2A	-6.23	100.58	112.47

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	M	404	BPB	CBC-CAC-C3C	-6.14	113.47	127.00

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	401	HEC	1	0
5	C	402	HEC	1	0
5	C	403	HEC	1	0
5	C	404	HEC	1	0
10	L	301	BCB	2	0
10	L	302	BCB	2	0
11	L	303	BPB	1	0
10	M	402	BCB	1	0
11	M	404	BPB	3	0
14	M	405	NS5	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	C	332/336 (98%)	-0.35	6 (1%) 69 66	60, 76, 106, 138	0
2	H	257/258 (99%)	-0.20	15 (5%) 24 22	67, 89, 139, 200	0
3	L	273/273 (100%)	-0.43	7 (2%) 56 54	60, 74, 102, 120	0
4	M	323/323 (100%)	-0.18	8 (2%) 58 55	60, 74, 103, 128	0
All	All	1185/1190 (99%)	-0.29	36 (3%) 51 49	60, 77, 112, 200	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	52	LEU	4.8
4	M	37	TRP	4.4
2	H	47	LEU	4.2
2	H	49	LEU	3.9
2	H	46	PRO	3.5

### 6.2 Non-standard residues in protein, DNA, RNA chains [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(Å <sup>2</sup> )	Q<0.9
2	FME	H	1	10/11	0.98	0.10	-	74,90,102,110	0

## 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
7	SO4	H	704	5/5	0.83	0.40	26.49	90,93,97,98	5
9	HTO	L	305	10/10	0.75	0.40	26.29	102,114,122,129	0
9	HTO	H	708	10/10	0.85	0.43	15.10	102,117,128,128	0
7	SO4	M	412	5/5	0.86	0.49	9.69	143,150,158,164	0
6	DGA	C	405	37/44	0.54	0.43	5.64	106,133,171,173	0
8	LDA	M	414	16/16	0.88	0.35	5.25	115,120,158,159	0
8	LDA	H	706	16/16	0.67	0.34	4.00	88,103,162,165	0
8	LDA	M	415	16/16	0.61	0.46	3.40	100,129,160,162	0
9	HTO	H	709	10/10	0.82	0.16	3.06	108,116,128,129	0
8	LDA	H	701	16/16	0.96	0.24	2.27	72,87,99,100	0
8	LDA	M	416	16/16	0.71	0.41	2.04	119,131,154,155	0
14	NS5	M	405	40/40	0.86	0.22	2.04	70,89,128,128	0
8	LDA	H	707	16/16	0.95	0.26	1.98	95,110,135,135	0
11	BPB	L	303	65/65	0.98	0.18	1.36	60,66,76,80	0
11	BPB	M	404	65/65	0.97	0.16	1.28	64,75,157,163	0
13	MQ7	M	401	48/48	0.94	0.22	0.98	64,71,114,126	0
10	BCB	M	403	66/66	0.98	0.18	0.46	54,63,95,99	0
7	SO4	C	406	5/5	0.84	0.22	0.37	128,143,148,152	0
10	BCB	M	402	66/66	0.96	0.17	0.25	60,70,151,157	0
5	HEC	C	403	43/43	0.98	0.18	0.21	51,63,70,77	0
5	HEC	C	402	43/43	0.98	0.13	0.19	69,75,86,94	0
10	BCB	L	301	66/66	0.97	0.19	0.18	55,60,78,91	0
10	BCB	L	302	66/66	0.98	0.20	0.18	59,64,105,109	0
5	HEC	C	404	43/43	0.98	0.12	-0.10	60,67,86,105	0
5	HEC	C	401	43/43	0.98	0.11	-0.39	69,81,90,98	0
7	SO4	M	406	5/5	0.99	0.09	-0.53	88,96,107,109	0
7	SO4	H	702	5/5	0.98	0.09	-0.67	95,100,107,112	0
12	FE2	L	306	1/1	0.99	0.15	-0.92	68,68,68,68	0
8	LDA	L	304	16/16	0.81	0.45	-	133,139,169,169	0
7	SO4	M	409	5/5	0.92	0.35	-	130,130,144,150	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
7	SO4	H	703	5/5	0.96	0.39	-	163,163,172,174	0
7	SO4	C	407	5/5	0.84	0.65	-	102,107,112,113	5
7	SO4	M	408	5/5	0.99	0.08	-	77,83,86,90	0
7	SO4	H	705	5/5	0.97	0.06	-	95,96,112,113	5
7	SO4	C	408	5/5	0.80	0.47	-	100,103,108,108	5
7	SO4	M	413	5/5	0.75	0.40	-	208,211,215,229	0
7	SO4	M	411	5/5	0.83	0.32	-	140,148,160,162	0
7	SO4	M	407	5/5	0.96	0.10	-	102,105,117,122	0
7	SO4	M	410	5/5	0.84	0.25	-	89,97,101,105	5

## 6.5 Other polymers [i](#)

There are no such residues in this entry.