



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 08:50 PM GMT

PDB ID : 1M1N  
Title : Nitrogenase MoFe protein from Azotobacter vinelandii  
Authors : Einsle, O.; Tezcan, F.A.; Andrade, S.L.A.; Schmid, B.; Yoshida, M.; Howard, J.B.; Rees, D.C.  
Deposited on : 2002-06-19  
Resolution : 1.16 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

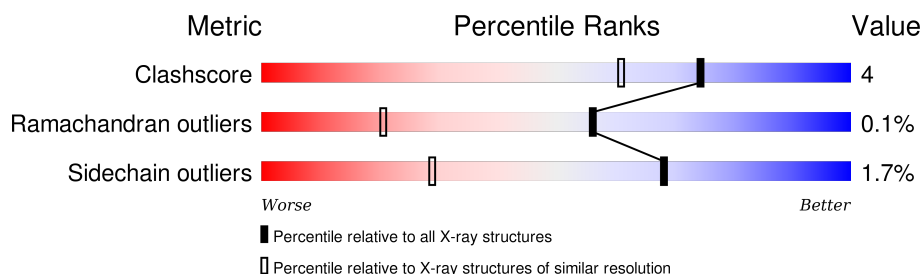
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.16 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))
Clashscore	102246	1109 (1.20-1.12)
Ramachandran outliers	100387	1058 (1.20-1.12)
Sidechain outliers	100360	1058 (1.20-1.12)


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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	491	
1	C	491	
1	E	491	
1	G	491	
2	B	522	
2	D	522	
2	F	522	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	522	 87%12%•

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 37384 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 Nitrogenase molybdenum-iron protein alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	477	Total	C	N	O	S	0	6	0
			3810	2423	647	713	27			
1	C	477	Total	C	N	O	S	0	6	0
			3814	2424	652	710	28			
1	E	477	Total	C	N	O	S	0	7	0
			3815	2426	647	715	27			
1	G	477	Total	C	N	O	S	0	7	0
			3819	2428	653	710	28			

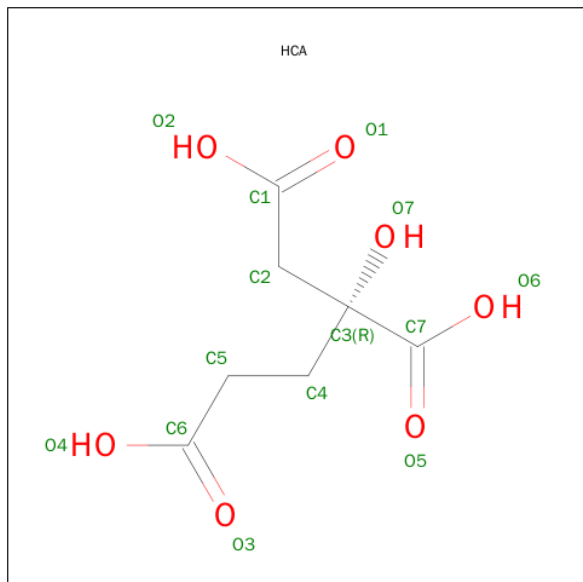
- Molecule 2 is a protein called Nitrogenase molybdenum-iron protein beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	522	Total	C	N	O	S	0	18	0
			4236	2710	707	784	35			
2	D	522	Total	C	N	O	S	0	15	0
			4225	2701	707	782	35			
2	F	522	Total	C	N	O	S	0	17	0
			4234	2709	707	783	35			
2	H	522	Total	C	N	O	S	0	13	0
			4218	2696	707	780	35			

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

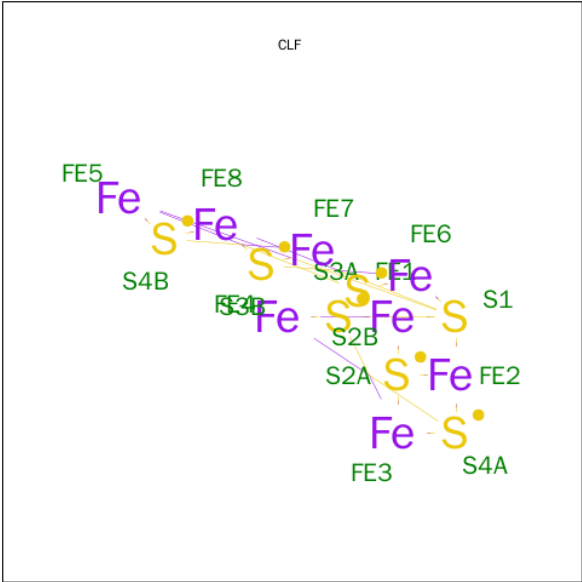
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	H	1	Total	Ca	0	0
			1	1		
3	B	1	Total	Ca	0	0
			1	1		
3	D	1	Total	Ca	0	0
			1	1		
3	F	1	Total	Ca	0	0
			1	1		

- Molecule 4 is 3-HYDROXY-3-CARBOXY-ADIPIC ACID (three-letter code: HCA) (formula:  $C_7H_{10}O_7$ ).



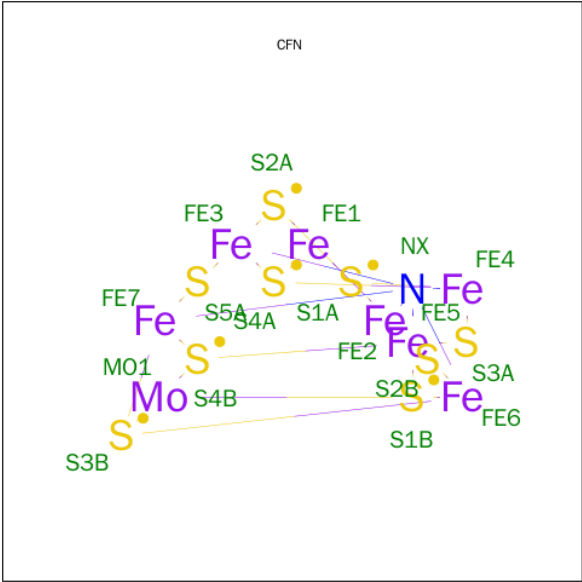
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			14	7	7		
4	C	1	Total	C	O	0	0
			14	7	7		
4	E	1	Total	C	O	0	0
			14	7	7		
4	G	1	Total	C	O	0	0
			14	7	7		

- Molecule 5 is FE(8)-S(7) CLUSTER (three-letter code: CLF) (formula:  $Fe_8S_7$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	Fe	S	0	0
			15	8	7		
5	D	1	Total	Fe	S	0	0
			15	8	7		
5	F	1	Total	Fe	S	0	0
			15	8	7		
5	H	1	Total	Fe	S	0	0
			15	8	7		

- Molecule 6 is FE(7)-MO-S(9)-N CLUSTER (three-letter code: CFN) (formula: Fe<sub>7</sub>MoNS<sub>9</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total 18	Fe 7	Mo 1	N 1	S 9	0	0
6	C	1	Total 18	Fe 7	Mo 1	N 1	S 9	0	0
6	E	1	Total 18	Fe 7	Mo 1	N 1	S 9	0	0
6	G	1	Total 18	Fe 7	Mo 1	N 1	S 9	0	0

- Molecule 7 is water.

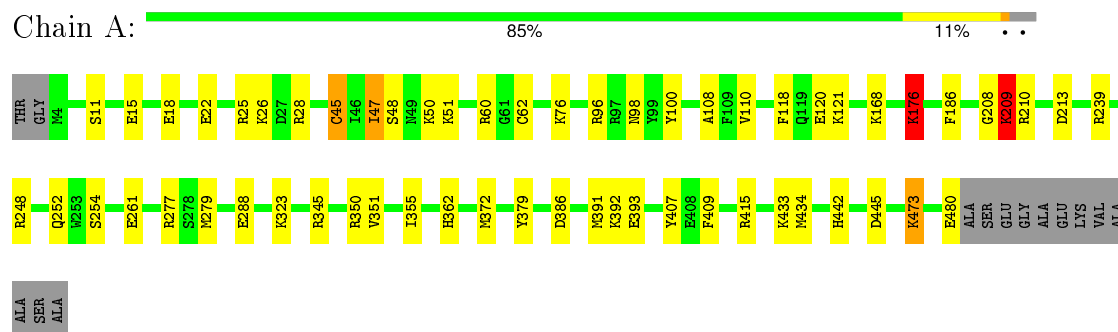
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	553	Total 553	O 553	0	0
7	B	723	Total 723	O 723	0	0
7	C	570	Total 570	O 570	0	0
7	D	704	Total 704	O 704	0	0
7	E	536	Total 536	O 536	0	0
7	F	708	Total 708	O 708	0	0
7	G	533	Total 533	O 533	0	0
7	H	694	Total 694	O 694	0	0

### 3 Residue-property plots

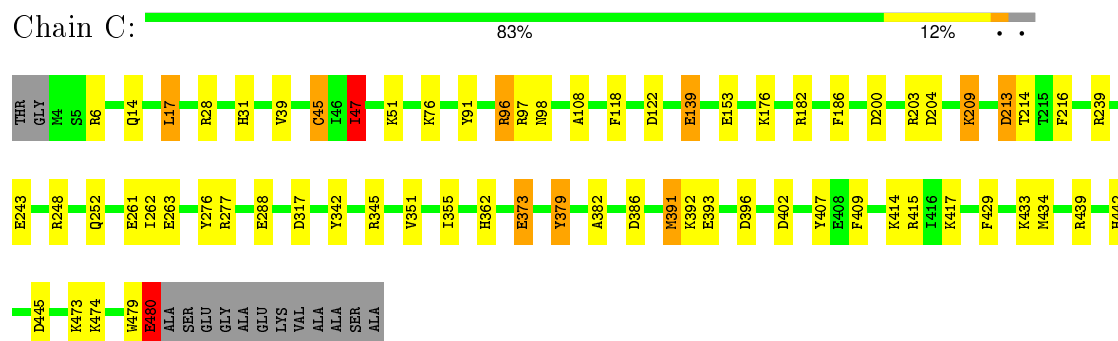
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according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

Note EDS was not executed.

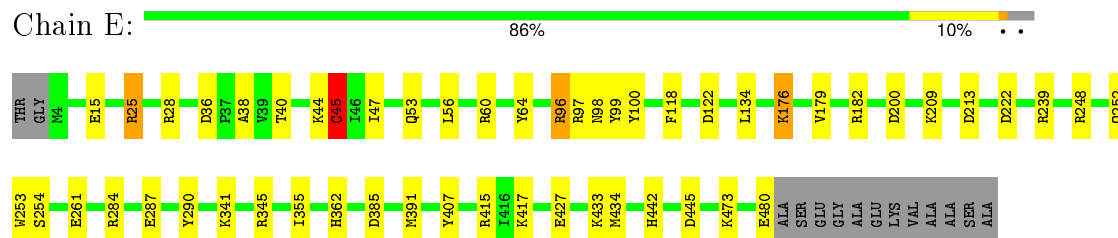
- Molecule 1: Nitrogenase molybdenum-iron protein alpha chain



- Molecule 1: Nitrogenase molybdenum-iron protein alpha chain




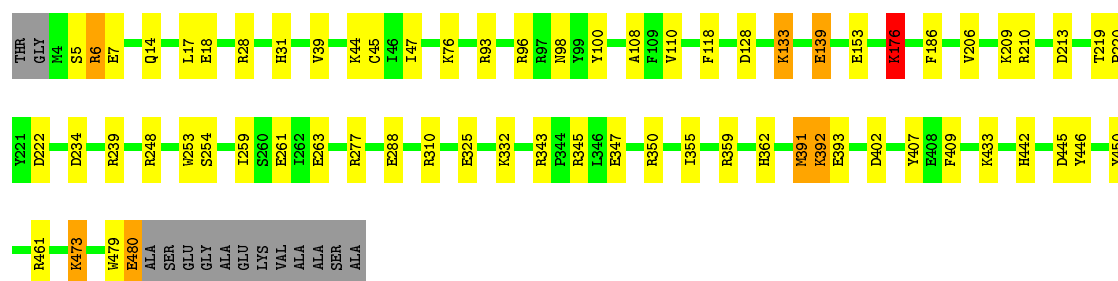
- Molecule 1: Nitrogenase molybdenum-iron protein alpha chain




- Molecule 1: Nitrogenase molybdenum-iron protein alpha chain

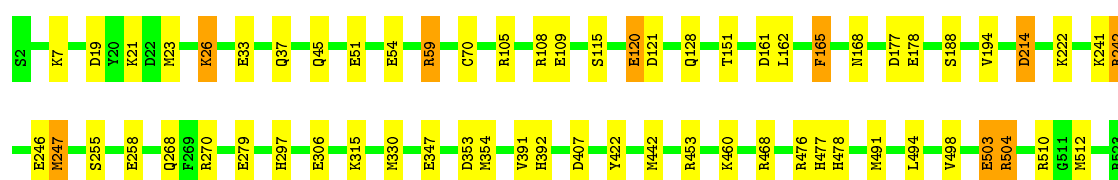


Chain G:  83% 12% ..



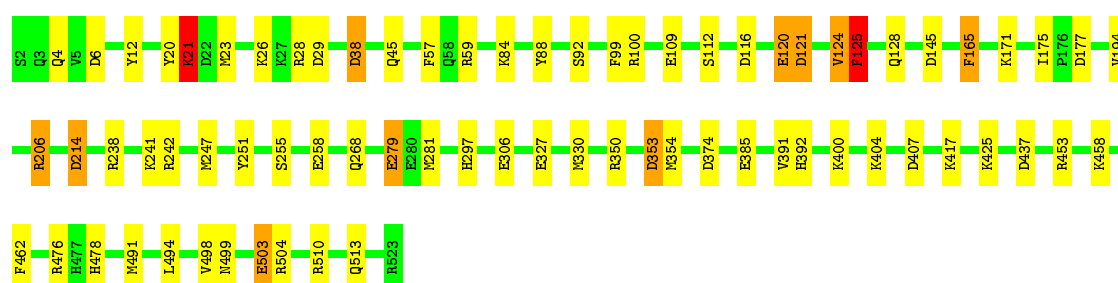
- Molecule 2: Nitrogenase molybdenum-iron protein beta chain

Chain B:  88% 11% .




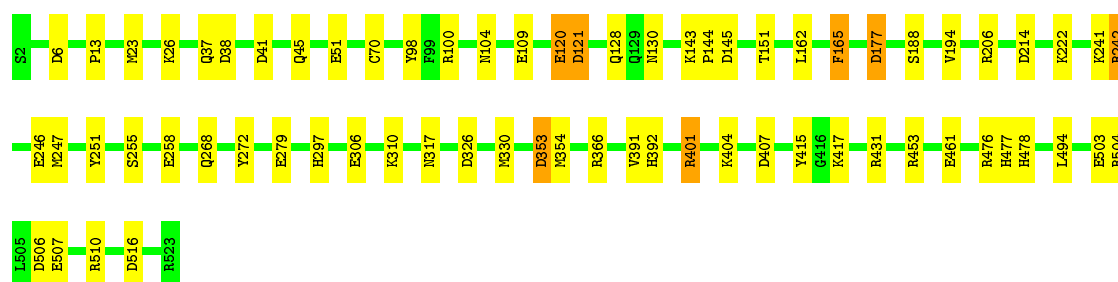
- Molecule 2: Nitrogenase molybdenum-iron protein beta chain

Chain D:  86% 12% .




- Molecule 2: Nitrogenase molybdenum-iron protein beta chain

Chain F:  87% 12% .



- Molecule 2: Nitrogenase molybdenum-iron protein beta chain

Chain H:  87% 12% .



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	108.31Å 131.63Å 159.16Å 90.00° 108.37° 90.00°	Depositor
Resolution (Å)	50.00 – 1.16	Depositor
% Data completeness (in resolution range)	95.6 (50.00-1.16)	Depositor
$R_{merge}$	0.09	Depositor
$R_{sym}$	0.08	Depositor
Refinement program	REFMAC 5	Depositor
R, $R_{free}$	0.123 , 0.149	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	37384	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	wwPDB-VP

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: CFN, CLF, HCA, CA

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	1.27	21/3922 (0.5%)	1.20	25/5286 (0.5%)
1	C	1.32	28/3926 (0.7%)	1.26	34/5291 (0.6%)
1	E	1.25	15/3931 (0.4%)	1.19	27/5298 (0.5%)
1	G	1.23	18/3935 (0.5%)	1.16	27/5302 (0.5%)
2	B	1.29	29/4414 (0.7%)	1.19	25/5960 (0.4%)
2	D	1.29	29/4391 (0.7%)	1.23	42/5930 (0.7%)
2	F	1.21	22/4408 (0.5%)	1.12	21/5952 (0.4%)
2	H	1.24	25/4376 (0.6%)	1.18	30/5909 (0.5%)
All	All	1.26	187/33303 (0.6%)	1.19	231/44928 (0.5%)

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	2
1	C	0	2
1	E	0	1
1	G	0	2
2	B	0	3
2	H	0	1
All	All	0	11

The worst 5 of 187 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	391	MET	CG-SD	-11.51	1.51	1.81
1	C	407	TYR	CE1-CZ	-10.23	1.25	1.38
2	D	503	GLU	CD-OE1	10.23	1.36	1.25
1	G	391	MET	CG-SD	-10.04	1.55	1.81

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	92	SER	CB-OG	-9.96	1.29	1.42

The worst 5 of 231 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	247	MET	CG-SD-CE	14.08	122.72	100.20
2	H	38	ASP	CB-CG-OD2	13.25	130.23	118.30
2	B	491	MET	CG-SD-CE	-12.97	79.44	100.20
2	B	510	ARG	NE-CZ-NH1	12.90	126.75	120.30
2	D	491	MET	CG-SD-CE	-12.35	80.44	100.20

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	379	TYR	Sidechain
1	A	96	ARG	Sidechain
2	B	108	ARG	Sidechain
2	B	315	LYS	Mainchain
2	B	59	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	3810	0	3748	30	0
1	C	3814	0	3755	40	2
1	E	3815	0	3750	15	0
1	G	3819	0	3764	32	0
2	B	4236	0	4156	31	0
2	D	4225	0	4141	46	0
2	F	4234	0	4155	38	0
2	H	4218	0	4136	43	0
3	B	1	0	0	1	0
3	D	1	0	0	1	0
3	F	1	0	0	0	0
3	H	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	14	0	6	2	0
4	C	14	0	6	2	0
4	E	14	0	6	1	0
4	G	14	0	6	2	0
5	B	15	0	0	0	0
5	D	15	0	0	0	0
5	F	15	0	0	0	0
5	H	15	0	0	0	0
6	A	18	0	0	0	0
6	C	18	0	0	0	0
6	E	18	0	0	0	0
6	G	18	0	0	0	0
7	A	553	0	0	12	5
7	B	723	0	0	15	2
7	C	570	0	0	10	12
7	D	704	0	0	24	6
7	E	536	0	0	5	0
7	F	708	0	0	14	4
7	G	533	0	0	7	4
7	H	694	0	0	17	5
All	All	37384	0	31629	258	20

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.

The worst 5 of 258 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:C:209:LYS:NZ	1:C:209:LYS:CE	1.69	1.56
1:C:391:MET:CG	1:C:391:MET:CB	1.76	1.55
2:F:247[A]:MET:CE	2:F:247[A]:MET:SD	2.01	1.48
2:B:247[A]:MET:SD	2:B:247[A]:MET:CE	2.10	1.38
2:D:124:VAL:HG13	7:D:8044:HOH:O	1.35	1.20

The worst 5 of 20 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)
7:G:3329:HOH:O	7:H:4423:HOH:O[2_756]	0.55	1.65
7:C:8019:HOH:O	7:D:8052:HOH:O[2_545]	0.56	1.64

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:A:6891:HOH:O	7:B:8141:HOH:O[2_655]	0.71	1.49
7:C:7897:HOH:O	7:F:4626:HOH:O[1_454]	0.84	1.36
7:C:7978:HOH:O	7:H:4816:HOH:O[1_454]	0.98	1.22

## 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	481/491 (98%)	463 (96%)	18 (4%)	0	100	100
1	C	481/491 (98%)	461 (96%)	20 (4%)	0	100	100
1	E	482/491 (98%)	462 (96%)	20 (4%)	0	100	100
1	G	482/491 (98%)	462 (96%)	20 (4%)	0	100	100
2	B	538/522 (103%)	529 (98%)	8 (2%)	1 (0%)	52	20
2	D	535/522 (102%)	525 (98%)	9 (2%)	1 (0%)	52	20
2	F	537/522 (103%)	527 (98%)	9 (2%)	1 (0%)	52	20
2	H	533/522 (102%)	524 (98%)	8 (2%)	1 (0%)	52	20
All	All	4069/4052 (100%)	3953 (97%)	112 (3%)	4 (0%)	56	21

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	255	SER
2	D	255	SER
2	F	255	SER
2	H	255	SER

### 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	413/414 (100%)	402 (97%)	11 (3%)	52	11
1	C	413/414 (100%)	404 (98%)	9 (2%)	60	18
1	E	414/414 (100%)	406 (98%)	8 (2%)	65	24
1	G	414/414 (100%)	401 (97%)	13 (3%)	47	9
2	B	472/454 (104%)	469 (99%)	3 (1%)	90	69
2	D	469/454 (103%)	462 (98%)	7 (2%)	72	34
2	F	471/454 (104%)	466 (99%)	5 (1%)	80	47
2	H	467/454 (103%)	463 (99%)	4 (1%)	84	55
All	All	3533/3472 (102%)	3473 (98%)	60 (2%)	68	29

5 of 60 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	124	VAL
1	E	176	LYS
1	G	480	GLU
1	E	40	THR
1	E	362	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 46 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	168	ASN
1	E	271	ASN
2	H	130	ASN
2	D	499	ASN
1	E	29	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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 16 ligands modelled in this entry, 4 are monoatomic - leaving 12 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	HCA	A	6494	-	4,13,13	0.94	0	3,18,18	1.42	1 (33%)
6	CFN	A	6496	1	6,30,30	1.34	1 (16%)	0,78,78	0.00	-
5	CLF	B	6498	1,2	0,24,24	0.00	-	0,57,57	0.00	-
4	HCA	C	7494	-	4,13,13	1.31	0	3,18,18	1.85	1 (33%)
6	CFN	C	7496	1	6,30,30	0.77	0	0,78,78	0.00	-
5	CLF	D	7498	1,2	0,24,24	0.00	-	0,57,57	0.00	-
4	HCA	E	8494	-	4,13,13	1.05	0	3,18,18	1.13	0
6	CFN	E	8496	1	6,30,30	0.79	0	0,78,78	0.00	-
5	CLF	F	8498	1,2	0,24,24	0.00	-	0,57,57	0.00	-
4	HCA	G	9494	-	4,13,13	1.42	1 (25%)	3,18,18	1.28	1 (33%)
6	CFN	G	9496	1	6,30,30	1.17	1 (16%)	0,78,78	0.00	-
5	CLF	H	9498	1,2	0,24,24	0.00	-	0,57,57	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
4	HCA	A	6494	-	-	0/7/17/17	0/0/0/0
6	CFN	A	6496	1	-	0/0/204/204	0/0/13/13

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	CLF	B	6498	1,2	-	0/0/132/132	0/12/10/10
4	HCA	C	7494	-	-	0/7/17/17	0/0/0/0
6	CFN	C	7496	1	-	0/0/204/204	0/0/13/13
5	CLF	D	7498	1,2	-	0/0/132/132	0/12/10/10
4	HCA	E	8494	-	-	0/7/17/17	0/0/0/0
6	CFN	E	8496	1	-	0/0/204/204	0/0/13/13
5	CLF	F	8498	1,2	-	0/0/132/132	0/12/10/10
4	HCA	G	9494	-	-	0/7/17/17	0/0/0/0
6	CFN	G	9496	1	-	0/0/204/204	0/0/13/13
5	CLF	H	9498	1,2	-	0/0/132/132	0/12/10/10

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	6496	CFN	S2B-FE6	-2.62	2.17	2.24
6	G	9496	CFN	S2B-FE6	-2.44	2.17	2.24
4	G	9494	HCA	O7-C3	2.23	1.46	1.43

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	7494	HCA	C3-C2-C1	-2.50	110.97	114.96
4	A	6494	HCA	C4-C5-C6	2.03	120.49	112.10
4	G	9494	HCA	C4-C5-C6	2.11	120.84	112.10

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	6494	HCA	2	0
4	C	7494	HCA	2	0
4	E	8494	HCA	1	0
4	G	9494	HCA	2	0

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

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

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

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.