



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 6, 2023 – 03:06 am GMT

PDB ID : 1H1L  
Title : NITROGENASE MO-FE PROTEIN FROM KLEBSIELLA PNEUMONIAE,  
NIFV MUTANT  
Authors : Mayer, S.M.; Gormal, C.A.; Smith, B.E.; Lawson, D.M.  
Deposited on : 2002-07-18  
Resolution : 1.90 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 : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

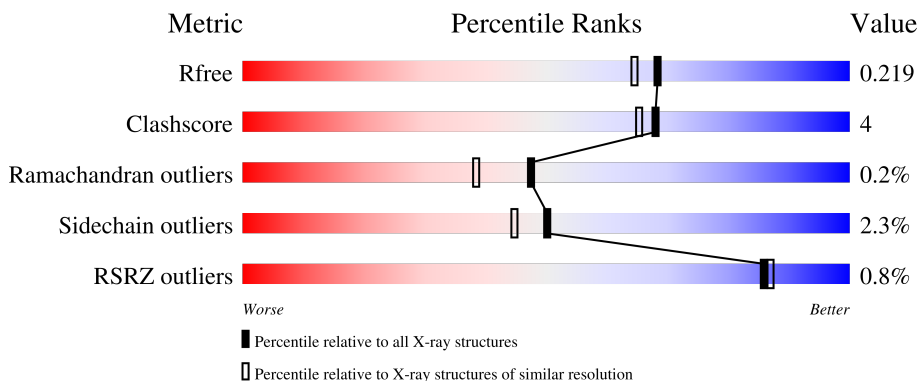
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	481	 80% 18% ..
1	C	481	 75% 22% ..
2	B	519	 83% 16% .
2	D	519	 82% 16% .

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 17025 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
			Total	C	N	O	S			
1	A	478	3664	2331	627	682	24	0	0	0
1	C	478	3674	2335	629	686	24	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	85	VAL	ALA	conflict	UNP P00466
A	94	GLY	GLU	conflict	UNP P00466
C	85	VAL	ALA	conflict	UNP P00466
C	94	GLY	GLU	conflict	UNP P00466

- Molecule 2 is a protein called NITROGENASE MOLYBDENUM IRON PROTEIN BETA CHAIN.

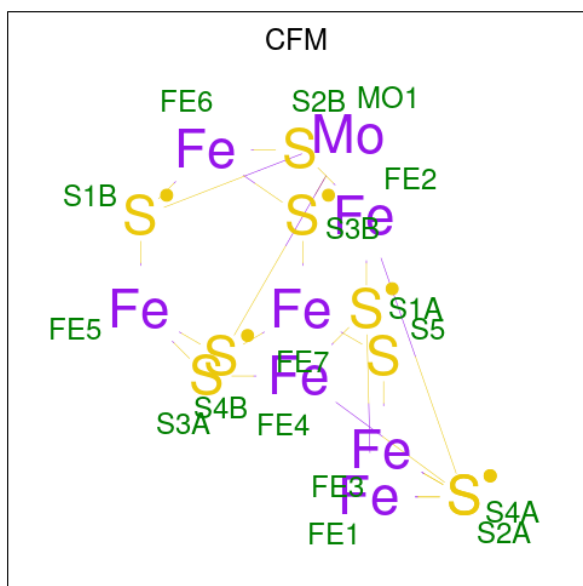
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	519	4026	2567	679	752	28	0	0	0
2	D	519	4028	2568	678	754	28	0	0	0

- Molecule 3 is CITRIC ACID (three-letter code: CIT) (formula: C<sub>6</sub>H<sub>8</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 13 6 7	0	0
3	C	1	Total C O 13 6 7	0	0

- Molecule 4 is FE-MO-S CLUSTER (three-letter code: CFM) (formula: Fe<sub>7</sub>MoS<sub>9</sub>).

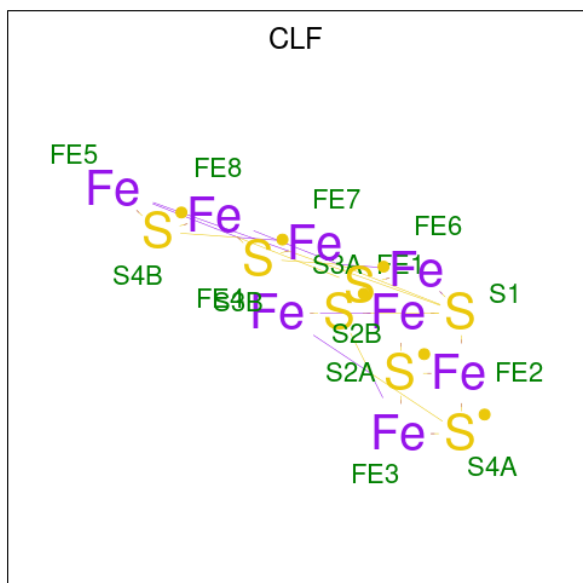


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Fe Mo S 17 7 1 9	0	0
4	C	1	Total Fe Mo S 17 7 1 9	0	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 6 is FE(8)-S(7) CLUSTER (three-letter code: CLF) (formula: Fe<sub>8</sub>S<sub>7</sub>).



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

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	1	Total Cl 1 1	0	0
7	D	1	Total Cl 1 1	0	0

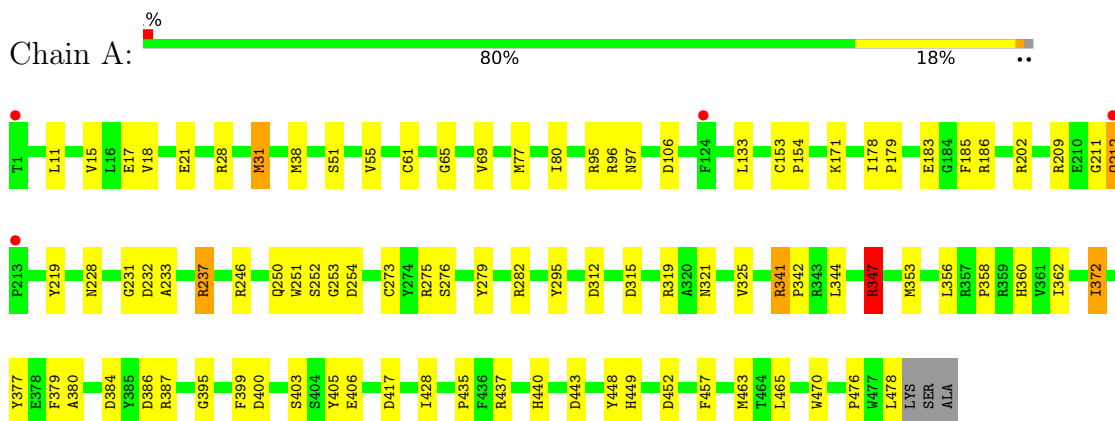
- Molecule 8 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
8	A	309	Total 309	O 309	0	0
8	B	474	Total 474	O 474	0	0
8	C	283	Total 283	O 283	0	0
8	D	471	Total 471	O 471	0	0

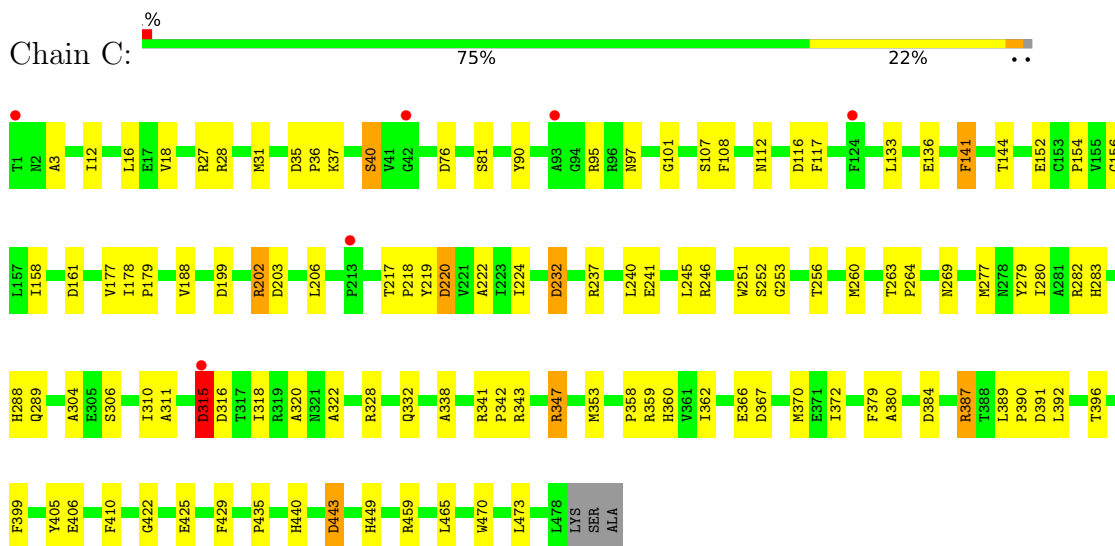
### 3 Residue-property plots

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.

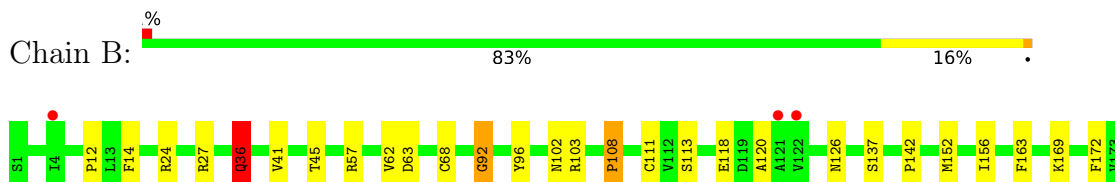
- Molecule 1: NITROGENASE MOLYBDENUM IRON PROTEIN ALPHA CHAIN

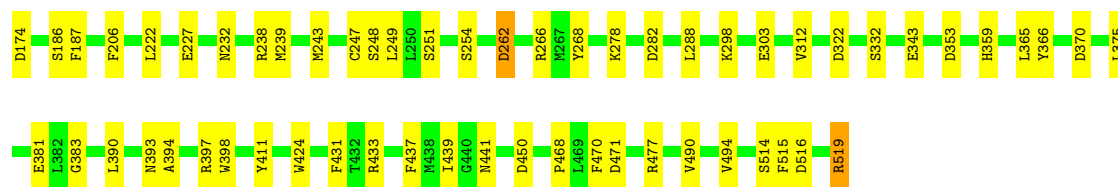


- Molecule 1: NITROGENASE MOLYBDENUM IRON PROTEIN ALPHA CHAIN

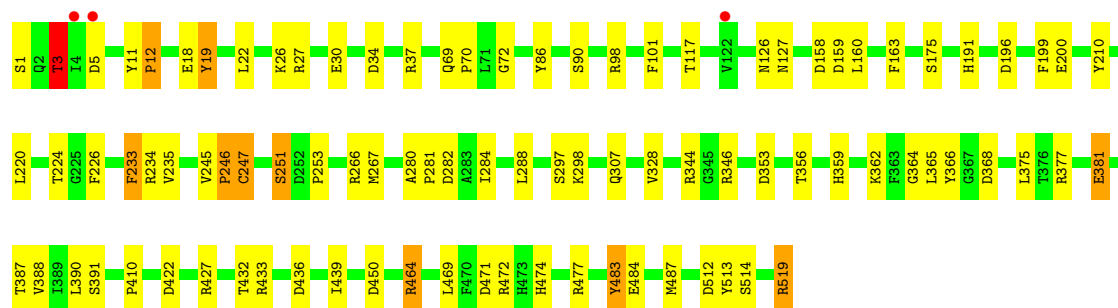
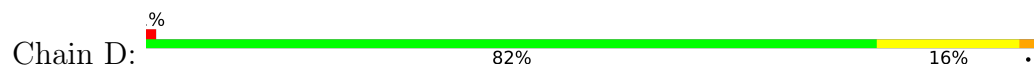


- Molecule 2: NITROGENASE MOLYBDENUM IRON PROTEIN BETA CHAIN





• Molecule 2: NITROGENASE MOLYBDENUM IRON PROTEIN BETA CHAIN





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	204.06Å 75.06Å 164.26Å 90.00° 123.95° 90.00°	Depositor
Resolution (Å)	40.00 – 1.90 39.36 – 1.90	Depositor EDS
% Data completeness (in resolution range)	94.2 (40.00-1.90) 94.3 (39.36-1.90)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.50 (at 1.89Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.176 , 0.236 0.164 , 0.219	Depositor DCC
$R_{free}$ test set	7646 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.3	Xtrriage
Anisotropy	0.255	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 57.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	17025	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.23% 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: CIT, CFM, MG, CLF, CL

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	0.68	0/3751	1.53	48/5090 (0.9%)
1	C	0.65	0/3761	1.53	45/5102 (0.9%)
2	B	0.70	0/4124	1.48	53/5604 (0.9%)
2	D	0.74	2/4126 (0.0%)	1.49	56/5606 (1.0%)
All	All	0.69	2/15762 (0.0%)	1.51	202/21402 (0.9%)

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	12
1	C	0	22
2	B	0	13
2	D	0	14
All	All	0	61

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	307	GLN	CD-NE2	6.91	1.50	1.32
2	D	307	GLN	CG-CD	-5.23	1.39	1.51

The worst 5 of 202 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	328	ARG	NE-CZ-NH1	15.59	128.10	120.30
2	D	433	ARG	NE-CZ-NH1	14.45	127.53	120.30
1	A	96	ARG	NE-CZ-NH1	-14.41	113.09	120.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	282	ASP	CB-CG-OD1	13.86	130.78	118.30
2	D	344	ARG	NE-CZ-NH2	-12.54	114.03	120.30

There are no chirality outliers.

5 of 61 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	17	GLU	Mainchain
1	A	171	LYS	Mainchain
1	A	179	PRO	Mainchain
1	A	211	GLY	Mainchain
1	A	31	MET	Mainchain

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3664	0	3503	30	0
1	C	3674	0	3518	45	0
2	B	4026	0	3872	19	0
2	D	4028	0	3882	23	0
3	A	13	0	4	3	0
3	C	13	0	4	2	0
4	A	17	0	0	0	0
4	C	17	0	0	0	0
5	A	2	0	0	0	0
5	B	2	0	0	0	0
6	B	15	0	0	0	0
6	D	15	0	0	0	0
7	B	1	0	0	0	0
7	D	1	0	0	0	0
8	A	309	0	0	4	0
8	B	474	0	0	2	0
8	C	283	0	0	3	0
8	D	471	0	0	3	0
All	All	17025	0	14783	114	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 4.

The worst 5 of 114 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)
3:A:1479:CIT:O4	8:A:2309:HOH:O	1.71	1.07
1:C:237:ARG:HD3	2:D:22:LEU:HD21	1.50	0.93
1:C:217:THR:HG22	1:C:219:TYR:H	1.29	0.92
1:A:428:ILE:HG13	2:B:108:PRO:HB3	1.69	0.74
2:D:3:THR:HG22	2:D:5:ASP:H	1.52	0.73

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	476/481 (99%)	455 (96%)	21 (4%)	0	100	100
1	C	476/481 (99%)	457 (96%)	19 (4%)	0	100	100
2	B	517/519 (100%)	503 (97%)	13 (2%)	1 (0%)	47	38
2	D	517/519 (100%)	503 (97%)	12 (2%)	2 (0%)	34	24
All	All	1986/2000 (99%)	1918 (97%)	65 (3%)	3 (0%)	47	38

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	251	SER
2	B	251	SER
2	D	246	PRO

### 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	372/399 (93%)	364 (98%)	8 (2%)	52	47
1	C	375/399 (94%)	365 (97%)	10 (3%)	44	38
2	B	422/442 (96%)	413 (98%)	9 (2%)	53	48
2	D	425/442 (96%)	416 (98%)	9 (2%)	53	48
All	All	1594/1682 (95%)	1558 (98%)	36 (2%)	50	45

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	3	THR
2	D	514	SER
2	D	12	PRO
2	D	235	VAL
2	B	152	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	91	GLN
2	D	359	HIS
2	D	434	GLN
1	C	112	ASN
1	C	118	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 6 are monoatomic - leaving 6 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
6	CLF	B	1520	1,2	0,24,24	-	-	-		
4	CFM	A	1480	1	0,24,24	-	-	-		
4	CFM	C	1480	1	0,24,24	-	-	-		
3	CIT	A	1479	-	12,12,12	1.49	1 (8%)	17,17,17	1.95	3 (17%)
3	CIT	C	1479	-	12,12,12	1.49	1 (8%)	17,17,17	2.33	7 (41%)
6	CLF	D	1521	1,2	0,24,24	-	-	-		

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	CIT	A	1479	-	-	2/16/16/16	-
3	CIT	C	1479	-	-	5/16/16/16	-
6	CLF	D	1521	1,2	-	-	0/12/10/10
6	CLF	B	1520	1,2	-	-	0/12/10/10

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1479	CIT	C3-C6	-3.77	1.49	1.53
3	C	1479	CIT	C3-C6	-3.01	1.50	1.53

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	1479	CIT	O5-C6-C3	-4.54	115.82	122.25
3	C	1479	CIT	C3-C4-C5	4.48	124.66	113.81
3	A	1479	CIT	O5-C6-C3	-4.19	116.32	122.25
3	A	1479	CIT	O4-C5-O3	-4.07	113.15	123.30
3	A	1479	CIT	O6-C6-C3	3.32	118.81	113.05

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	1479	CIT	C1-C2-C3-C4
3	C	1479	CIT	C1-C2-C3-C6
3	C	1479	CIT	C1-C2-C3-O7
3	C	1479	CIT	O2-C1-C2-C3
3	A	1479	CIT	C1-C2-C3-C6

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1479	CIT	3	0
3	C	1479	CIT	2	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	478/481 (99%)	-0.20	4 (0%) 86 87	14, 23, 35, 44	0
1	C	478/481 (99%)	-0.09	6 (1%) 77 79	15, 26, 39, 50	0
2	B	519/519 (100%)	-0.32	3 (0%) 89 90	13, 20, 34, 46	0
2	D	519/519 (100%)	-0.39	3 (0%) 89 90	12, 20, 34, 44	0
All	All	1994/2000 (99%)	-0.25	16 (0%) 86 87	12, 22, 35, 50	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	122	VAL	4.4
2	D	4	ILE	4.0
2	D	5	ASP	3.6
1	C	315	ASP	3.5
1	A	1	THR	3.1

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

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

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q < 0.9’ lists the number of atoms with occupancy less than 0.9.



Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	MG	A	3001	1/1	0.82	0.10	28,28,28,28	0
5	MG	A	3003	1/1	0.86	0.20	37,37,37,37	0
3	CIT	C	1479	13/13	0.88	0.51	14,16,21,22	13
3	CIT	A	1479	13/13	0.93	0.53	12,16,17,18	13
5	MG	B	1522	1/1	0.98	0.12	23,23,23,23	0
7	CL	B	1521	1/1	0.98	0.04	26,26,26,26	0
4	CFM	A	1480	17/17	0.99	0.11	12,15,17,17	0
5	MG	B	1523	1/1	0.99	0.06	21,21,21,21	0
6	CLF	B	1520	15/15	0.99	0.08	14,16,19,22	0
6	CLF	D	1521	15/15	0.99	0.10	14,16,18,20	0
4	CFM	C	1480	17/17	0.99	0.08	13,18,19,20	0
7	CL	D	1520	1/1	1.00	0.07	23,23,23,23	0

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

There are no such residues in this entry.