



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 07:33 PM GMT

PDB ID : 1G21  
Title : MGATP-BOUND AND NUCLEOTIDE-FREE STRUCTURES OF A NITROGENASE PROTEIN COMPLEX BETWEEN LEU127DEL-FE PROTEIN AND THE MOFE PROTEIN  
Authors : Chiu, H.-J.; Peters, J.W.; Lanzilotta, W.N.; Ryle, M.J.; Seefeldt, L.C.; Howard, J.B.; Rees, D.C.  
Deposited on : 2000-10-16  
Resolution : 3.00 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

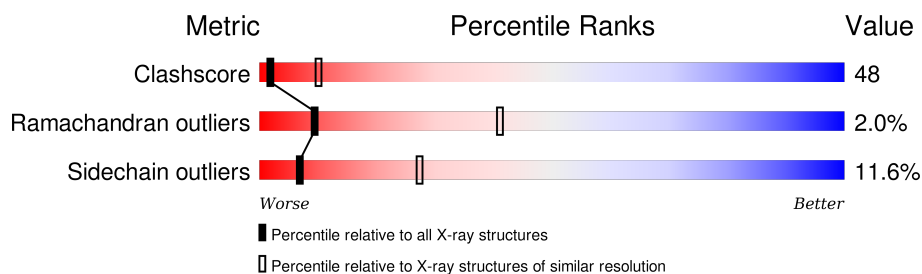
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)

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	492	
1	C	492	
2	B	523	
2	D	523	
3	E	289	
3	F	289	
3	G	289	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	H	289	

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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	ATP	G	7292	-	-	X	-
9	CFM	A	1496	-	-	X	-
9	CFM	C	3496	-	-	X	-

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 24237 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	476	Total	C	N	O	S	0	0	0
			3776	2402	642	708	24			
1	C	476	Total	C	N	O	S	0	0	0
			3776	2402	642	708	24			

- 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	0	0
			4170	2663	704	775	28			
2	D	522	Total	C	N	O	S	0	0	0
			4170	2663	704	775	28			

- Molecule 3 is a protein called NITROGENASE IRON PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	268	Total	C	N	O	S	109	0	0
			2029	1265	347	397	20			
3	F	267	Total	C	N	O	S	71	0	0
			2020	1260	346	394	20			
3	G	268	Total	C	N	O	S	130	0	0
			2029	1265	347	397	20			
3	H	268	Total	C	N	O	S	116	0	0
			2029	1265	347	397	20			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	?	-	LEU	DELETION	UNP P00459
F	?	-	LEU	DELETION	UNP P00459
G	?	-	LEU	DELETION	UNP P00459

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	?	-	LEU	DELETION	UNP P00459

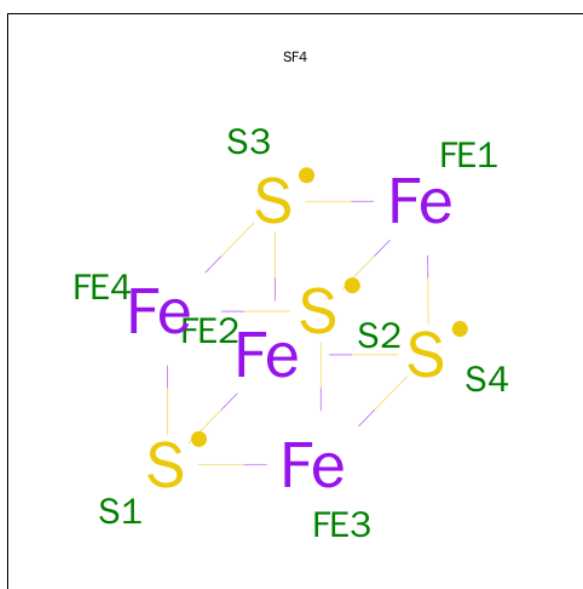
- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	H	1	Total Mg 1 1	0	0
4	G	1	Total Mg 1 1	0	0
4	F	1	Total Mg 1 1	0	0
4	E	1	Total Mg 1 1	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

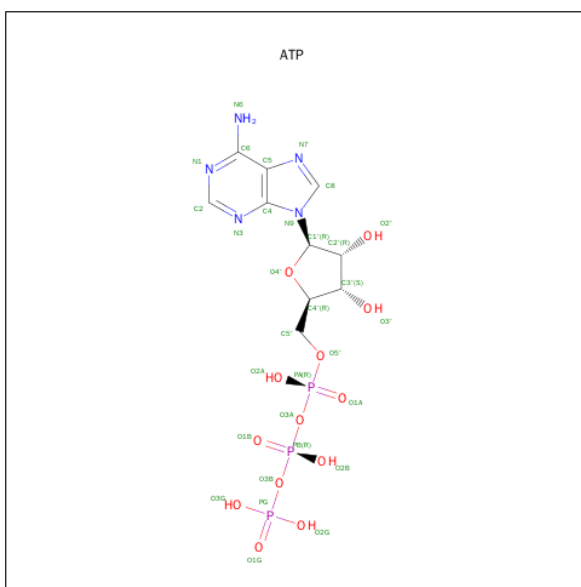
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total Ca 1 1	0	0
5	D	1	Total Ca 1 1	0	0

- Molecule 6 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe<sub>4</sub>S<sub>4</sub>).



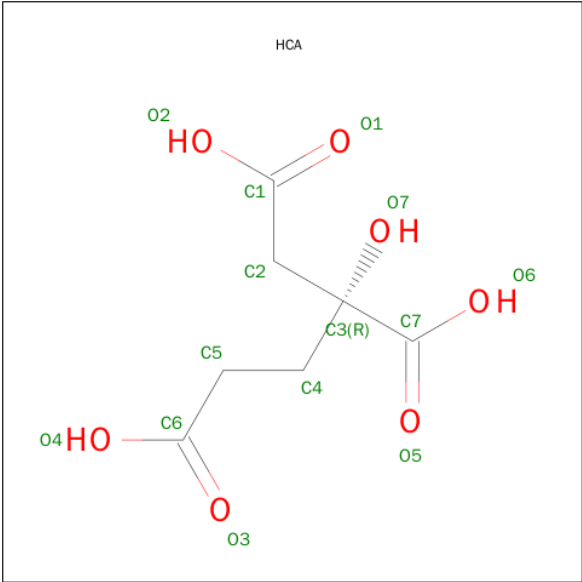
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	F	1	Total 8	Fe 4	S 4	0	0
6	G	1	Total 8	Fe 4	S 4	0	0

- Molecule 7 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula:  $\text{C}_{10}\text{H}_{16}\text{N}_5\text{O}_{13}\text{P}_3$ ).



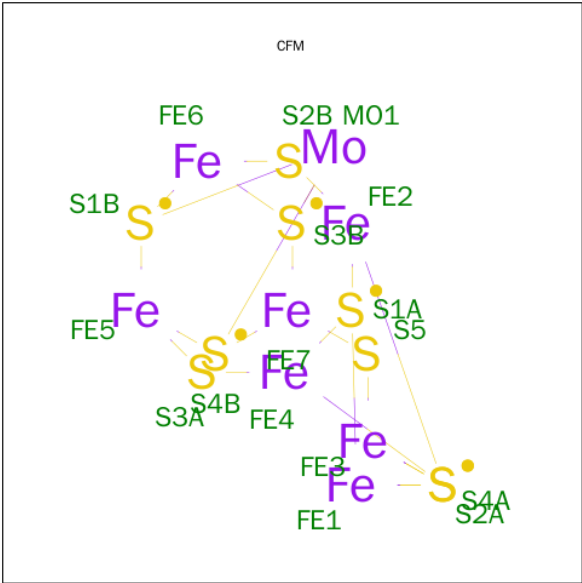
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	E	1	Total 31	C 10	N 5	O 13	P 3	0	0
7	F	1	Total 31	C 10	N 5	O 13	P 3	0	0
7	G	1	Total 31	C 10	N 5	O 13	P 3	0	0
7	H	1	Total 31	C 10	N 5	O 13	P 3	0	0

- Molecule 8 is 3-HYDROXY-3-CARBOXY-ADIPIC ACID (three-letter code: HCA) (formula:  $C_7H_{10}O_7$ ).



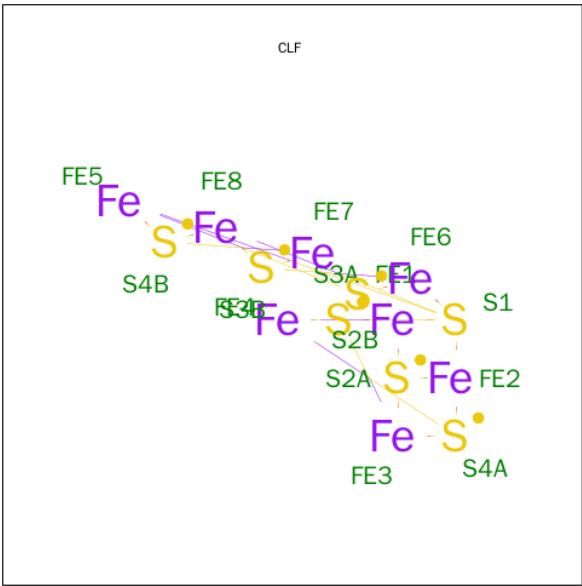
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			14	7	7		
8	C	1	Total	C	O	0	0
			14	7	7		

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



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

- Molecule 10 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
10	A	1	Total	Fe	S	0	0
			15	8	7		
10	D	1	Total	Fe	S	0	0
			15	8	7		

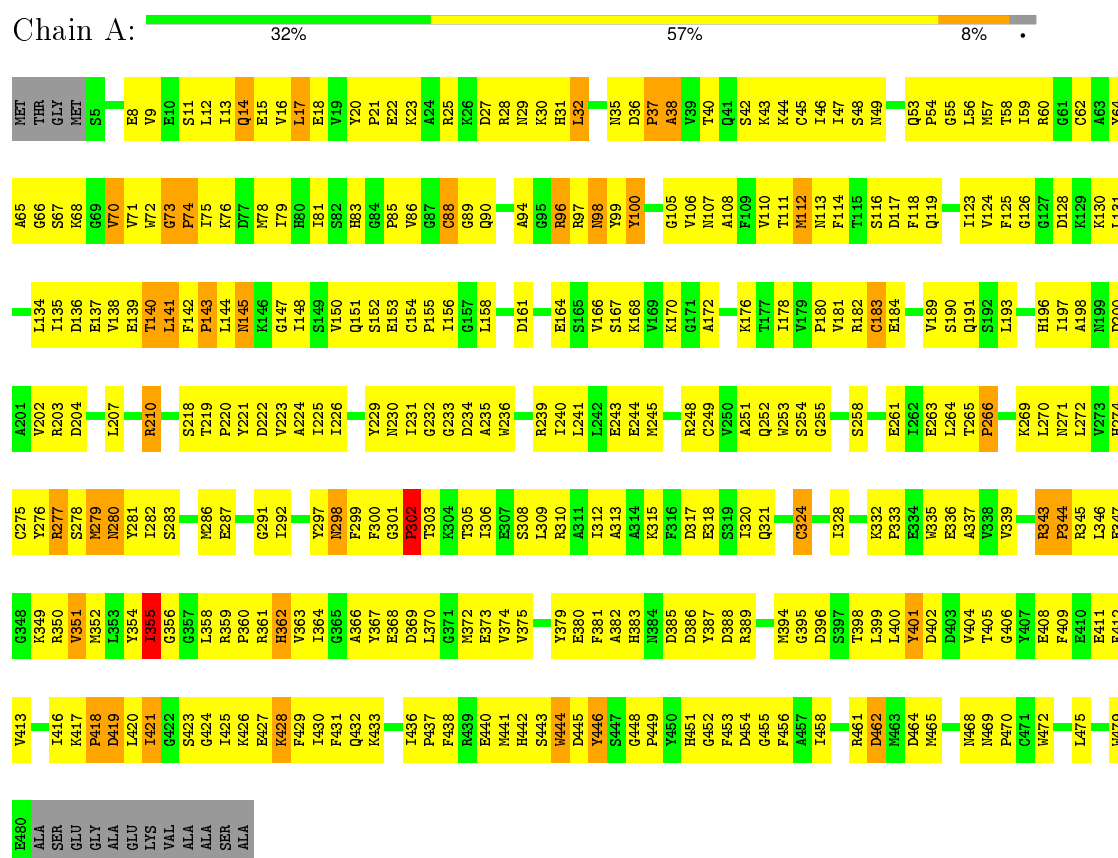


### 3 Residue-property plots

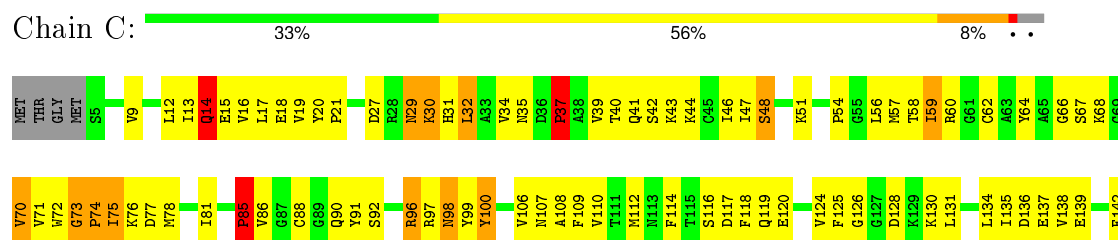
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according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

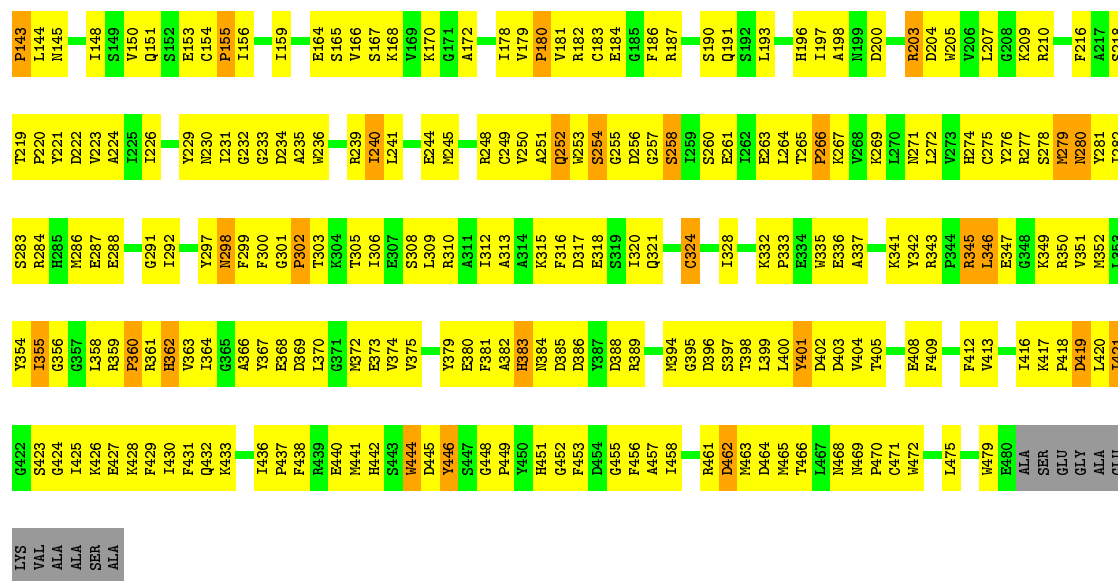
Note EDS was not executed.

#### • Molecule 1: NITROGENASE MOLYBDENUM-IRON PROTEIN ALPHA CHAIN

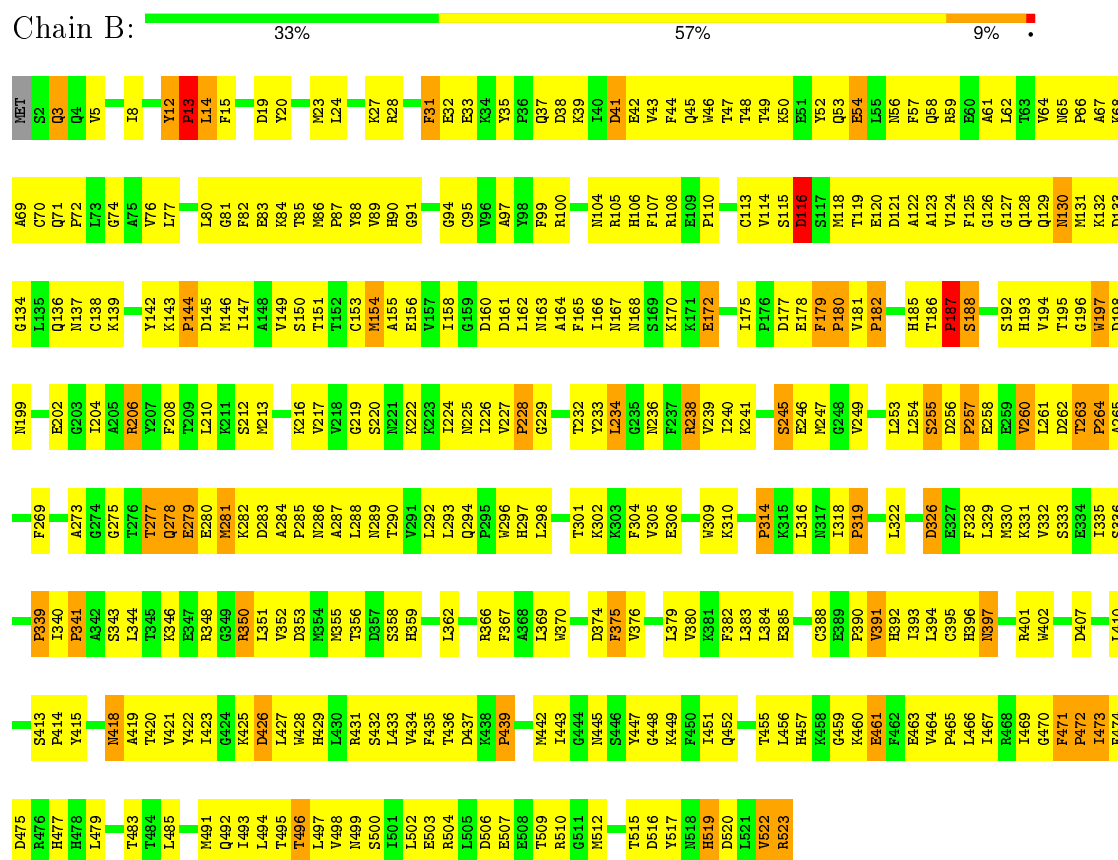


#### • Molecule 1: NITROGENASE MOLYBDENUM-IRON PROTEIN ALPHA CHAIN

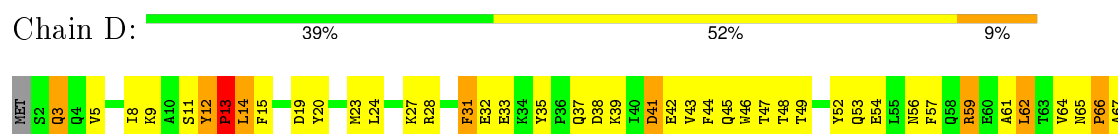


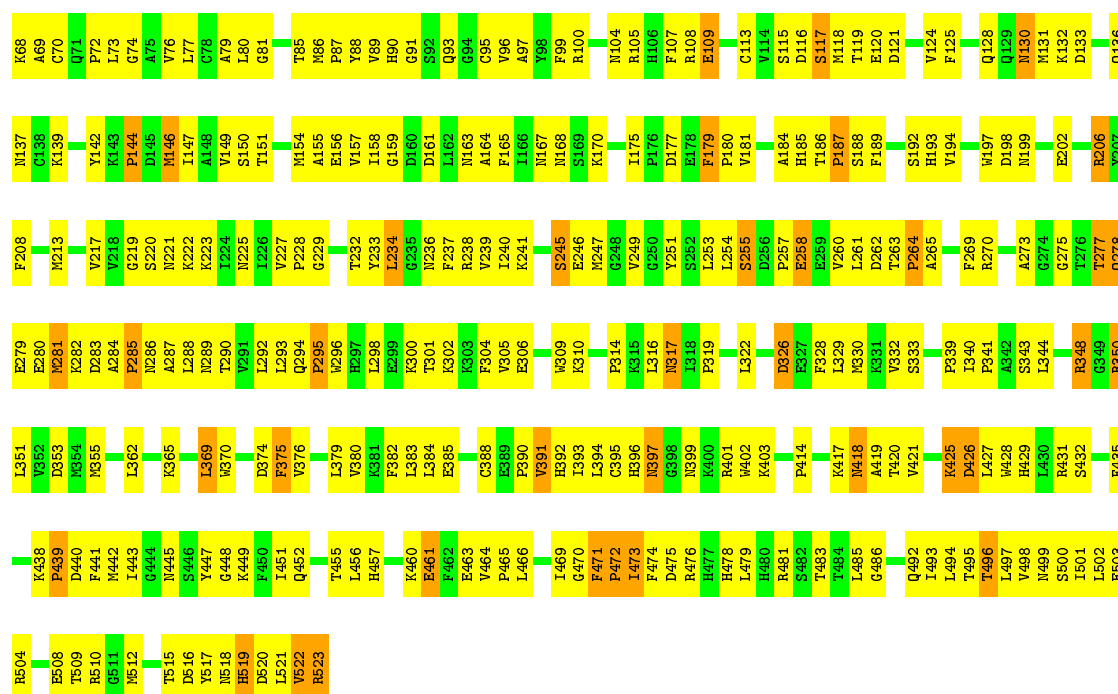


### • Molecule 2: NITROGENASE MOLYBDENUM-IRON PROTEIN BETA CHAIN



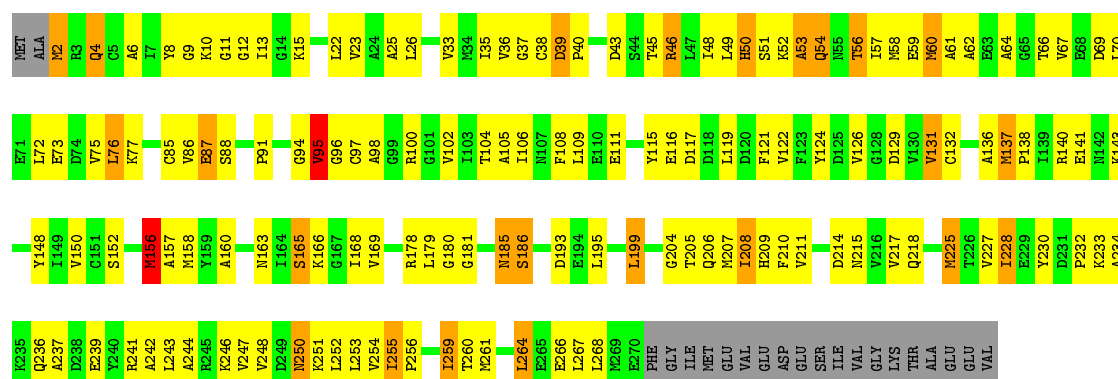
### • Molecule 2: NITROGENASE MOLYBDENUM-IRON PROTEIN BETA CHAIN





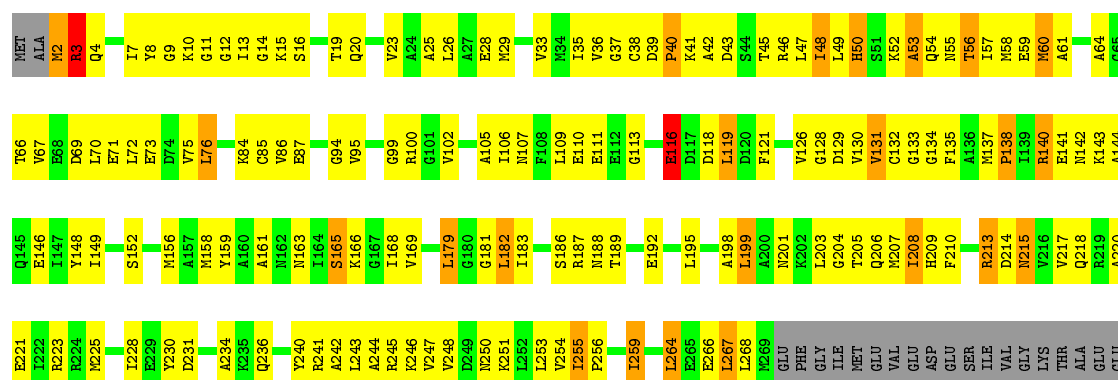
### • Molecule 3: NITROGENASE IRON PROTEIN

Chain E: 42% 42% 8% • 7%



### • Molecule 3: NITROGENASE IRON PROTEIN

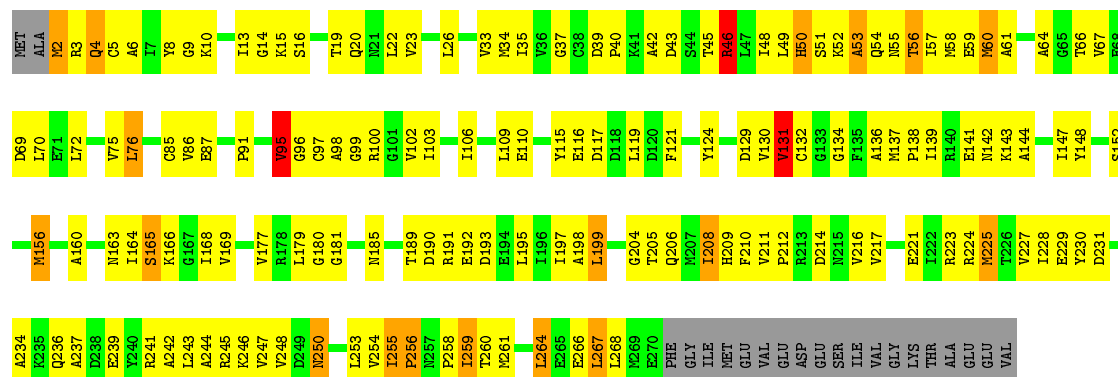
Chain F: 37% 47% 8% • 8%



VAL

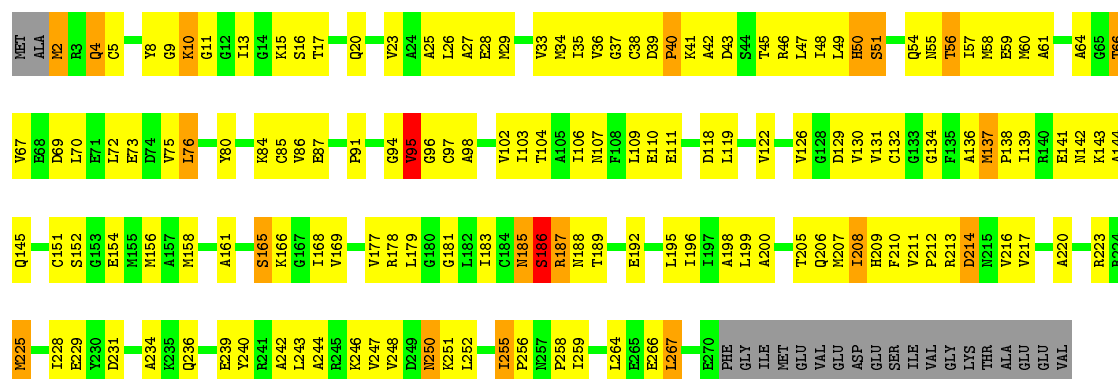
- Molecule 3: NITROGENASE IRON PROTEIN

Chain G: 40% 46% 6% • 7%



- Molecule 3: NITROGENASE IRON PROTEIN

Chain H: 39% 46% 7% • 7%



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	110.50Å 121.50Å 264.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.00	Depositor
% Data completeness (in resolution range)	71.6 (20.00-3.00)	Depositor
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.238 , 0.285	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	24237	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	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: MG, CFM, CLF, HCA, ATP, CA, SF4

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	0/3864	0.73	2/5212 (0.0%)
1	C	0.48	0/3864	0.73	2/5212 (0.0%)
2	B	0.54	0/4276	0.82	7/5782 (0.1%)
2	D	0.50	1/4276 (0.0%)	0.88	15/5782 (0.3%)
3	E	0.51	0/2052	0.80	2/2764 (0.1%)
3	F	0.55	1/2043 (0.0%)	0.84	4/2752 (0.1%)
3	G	0.47	0/2052	0.75	0/2764
3	H	0.54	1/2052 (0.0%)	0.81	1/2764 (0.0%)
All	All	0.51	3/24479 (0.0%)	0.80	33/33032 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	H	40	PRO	N-CD	5.37	1.55	1.47
3	F	40	PRO	N-CD	5.31	1.55	1.47
2	D	109	GLU	C-N	-5.08	1.24	1.34

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	12	TYR	C-N-CD	-21.77	72.71	120.60
2	B	12	TYR	C-N-CD	-20.40	75.71	120.60
2	B	12	TYR	C-N-CA	13.32	177.95	122.00
2	D	12	TYR	C-N-CA	12.49	174.45	122.00
1	A	88	CYS	CA-CB-SG	-10.83	94.51	114.00

There are no chirality outliers.

There are no planarity outliers.

## 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	3776	0	3709	452	0
1	C	3776	0	3709	490	0
2	B	4170	0	4076	441	0
2	D	4170	0	4076	461	0
3	E	2029	0	2040	150	0
3	F	2020	0	2034	166	0
3	G	2029	0	2041	158	0
3	H	2029	0	2041	153	0
4	E	1	0	0	0	0
4	F	1	0	0	0	0
4	G	1	0	0	0	0
4	H	1	0	0	0	0
5	B	1	0	0	0	0
5	D	1	0	0	0	0
6	F	8	0	0	0	0
6	G	8	0	0	0	0
7	E	31	0	12	4	0
7	F	31	0	12	6	0
7	G	31	0	12	11	0
7	H	31	0	12	6	0
8	A	14	0	6	1	0
8	C	14	0	6	2	0
9	A	17	0	0	4	0
9	C	17	0	0	7	0
10	A	15	0	0	1	0
10	D	15	0	0	3	0
All	All	24237	0	23786	2236	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 48.

The worst 5 of 2236 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:F:39:ASP:OD1	3:F:40:PRO:HD2	1.28	1.29

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:91:GLY:HA2	2:D:118:MET:CE	1.62	1.28
1:C:124:VAL:HG21	3:H:58:MET:CE	1.70	1.21
1:A:355:ILE:HG23	1:A:356:GLY:H	0.99	1.15
3:F:39:ASP:OD1	3:F:40:PRO:CD	1.95	1.13

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	474/492 (96%)	420 (89%)	47 (10%)	7 (2%)	13	50
1	C	474/492 (96%)	415 (88%)	52 (11%)	7 (2%)	13	50
2	B	520/523 (99%)	470 (90%)	43 (8%)	7 (1%)	15	53
2	D	520/523 (99%)	450 (86%)	60 (12%)	10 (2%)	10	43
3	E	266/289 (92%)	232 (87%)	29 (11%)	5 (2%)	10	43
3	F	265/289 (92%)	232 (88%)	25 (9%)	8 (3%)	5	29
3	G	266/289 (92%)	225 (85%)	33 (12%)	8 (3%)	5	29
3	H	266/289 (92%)	234 (88%)	24 (9%)	8 (3%)	5	29
All	All	3051/3186 (96%)	2678 (88%)	313 (10%)	60 (2%)	9	41

5 of 60 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	13	PRO
2	B	391	VAL
2	D	13	PRO
2	D	439	PRO
3	E	53	ALA



### 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	405/415 (98%)	362 (89%)	43 (11%)	8	31
1	C	405/415 (98%)	357 (88%)	48 (12%)	6	26
2	B	453/455 (100%)	398 (88%)	55 (12%)	6	25
2	D	453/455 (100%)	411 (91%)	42 (9%)	11	39
3	E	216/233 (93%)	186 (86%)	30 (14%)	4	19
3	F	215/233 (92%)	190 (88%)	25 (12%)	7	27
3	G	216/233 (93%)	188 (87%)	28 (13%)	5	22
3	H	216/233 (93%)	189 (88%)	27 (12%)	6	24
All	All	2579/2672 (96%)	2281 (88%)	298 (12%)	7	27

5 of 298 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	346	LEU
2	D	281	MET
3	H	17	THR
1	C	398	THR
2	D	41	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 81 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	362	HIS
2	D	130	ASN
3	H	54	GLN
1	C	384	ASN
2	D	18	GLN

### 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 18 ligands modelled in this entry, 6 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$
8	HCA	A	1494	9	4,13,13	0.61	0	3,18,18	1.42	1 (33%)
9	CFM	A	1496	1,8	0,24,24	0.00	-	0,45,45	0.00	-
10	CLF	A	1498	1,2	0,24,24	0.00	-	0,57,57	0.00	-
8	HCA	C	3494	9	4,13,13	0.58	0	3,18,18	0.67	0
9	CFM	C	3496	1,8	0,24,24	0.00	-	0,45,45	0.00	-
10	CLF	D	3498	1,2	0,24,24	0.00	-	0,57,57	0.00	-
7	ATP	E	5292	4	24,33,33	1.21	3 (12%)	31,52,52	1.16	2 (6%)
6	SF4	F	5290	3	0,12,12	0.00	-	0,24,24	0.00	-
7	ATP	F	6292	4	24,33,33	1.13	0	31,52,52	1.24	3 (9%)
6	SF4	G	7290	3	0,12,12	0.00	-	0,24,24	0.00	-
7	ATP	G	7292	4	24,33,33	1.83	5 (20%)	31,52,52	1.25	4 (12%)
7	ATP	H	8292	4	24,33,33	1.22	3 (12%)	31,52,52	1.26	4 (12%)

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
8	HCA	A	1494	9	-	0/7/17/17	0/0/0/0
9	CFM	A	1496	1,8	-	0/0/84/84	0/0/8/8

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
10	CLF	A	1498	1,2	-	0/0/132/132	0/12/10/10
8	HCA	C	3494	9	-	0/7/17/17	0/0/0/0
9	CFM	C	3496	1,8	-	0/0/84/84	0/0/8/8
10	CLF	D	3498	1,2	-	0/0/132/132	0/12/10/10
7	ATP	E	5292	4	-	0/18/38/38	0/3/3/3
6	SF4	F	5290	3	-	0/0/48/48	0/6/5/5
7	ATP	F	6292	4	-	0/18/38/38	0/3/3/3
6	SF4	G	7290	3	-	0/0/48/48	0/6/5/5
7	ATP	G	7292	4	-	0/18/38/38	0/3/3/3
7	ATP	H	8292	4	-	0/18/38/38	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	H	8292	ATP	C8-N7	-2.41	1.30	1.34
7	H	8292	ATP	PB-O2B	-2.28	1.45	1.54
7	G	7292	ATP	PG-O2G	-2.17	1.46	1.54
7	G	7292	ATP	PB-O2B	-2.12	1.45	1.54
7	E	5292	ATP	PG-O2G	-2.08	1.47	1.54

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	G	7292	ATP	O3'-C3'-C4'	-3.20	101.46	111.05
7	F	6292	ATP	O3'-C3'-C4'	-2.65	103.11	111.05
7	G	7292	ATP	N3-C2-N1	-2.37	127.08	128.89
7	H	8292	ATP	O3'-C3'-C4'	-2.15	104.60	111.05
8	A	1494	HCA	C3-C2-C1	2.03	118.20	114.96

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 45 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	A	1494	HCA	1	0
9	A	1496	CFM	4	0
10	A	1498	CLF	1	0
8	C	3494	HCA	2	0
9	C	3496	CFM	7	0
10	D	3498	CLF	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	E	5292	ATP	4	0
7	F	6292	ATP	6	0
7	G	7292	ATP	11	0
7	H	8292	ATP	6	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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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

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

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

### 6.3 Carbohydrates ⓘ

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

### 6.4 Ligands ⓘ

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

### 6.5 Other polymers ⓘ

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