wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 07:34 PM GMT

PDB ID	:	1G20
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	:	2.20 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
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

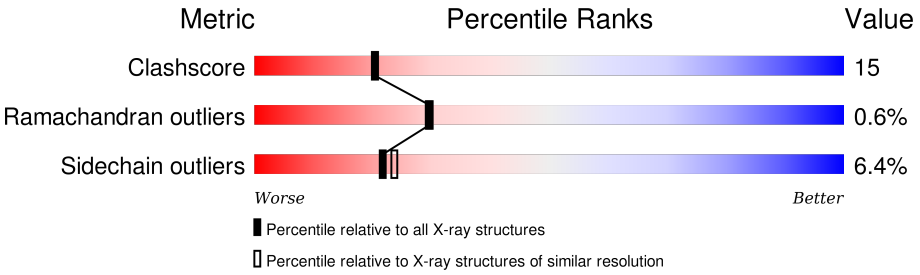
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)

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 >=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 <=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	73% 21%
1	C	492	71% 23%
2	B	523	81% 17%
2	D	523	78% 20%
3	E	289	58% 27% 10%
3	F	289	49% 38% 10%
3	G	289	51% 35% 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	H	289	51% 35% 9%

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	CFM	A	1496	-	-	X	-
7	CFM	C	3496	-	-	X	-

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 24903 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
			4174	2666	705	775	28			
2	D	522	Total	C	N	O	S	0	0	0
			4174	2666	705	775	28			

- Molecule 3 is a protein called NITROGENASE IRON PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	259	Total	C	N	O	S	3	0	0
			1953	1227	330	376	20			
3	F	260	Total	C	N	O	S	0	0	0
			1952	1226	327	379	20			
3	G	257	Total	C	N	O	S	5	0	0
			1938	1219	327	372	20			
3	H	262	Total	C	N	O	S	3	0	0
			1984	1244	337	383	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...

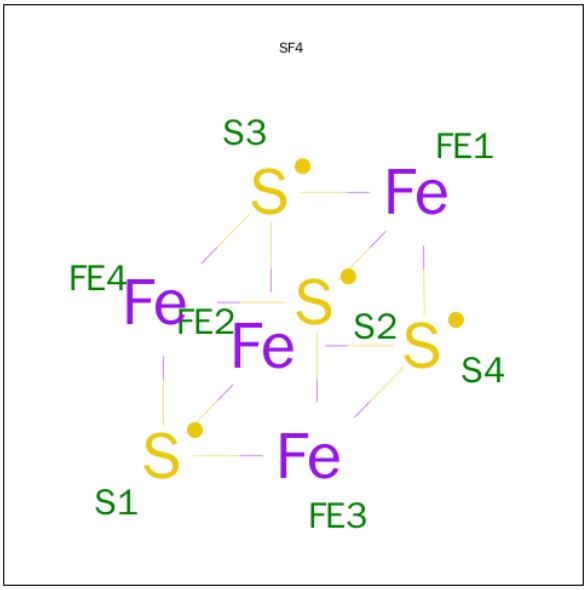
Continued from previous page...

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

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

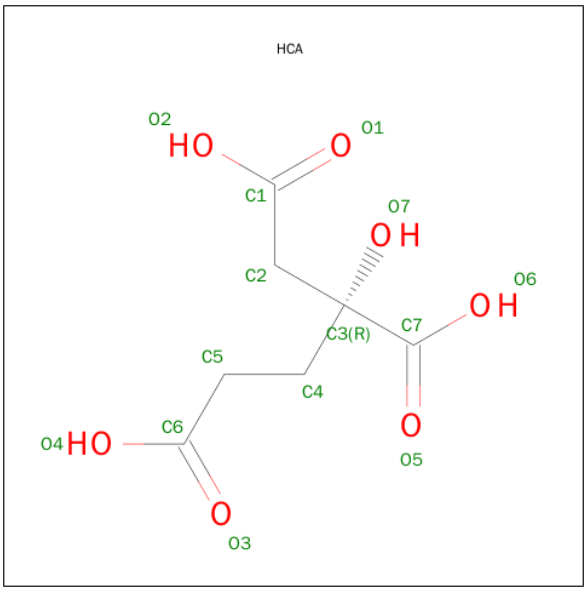
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Ca 1 1	0	0
4	D	1	Total Ca 1 1	0	0

- Molecule 5 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4).



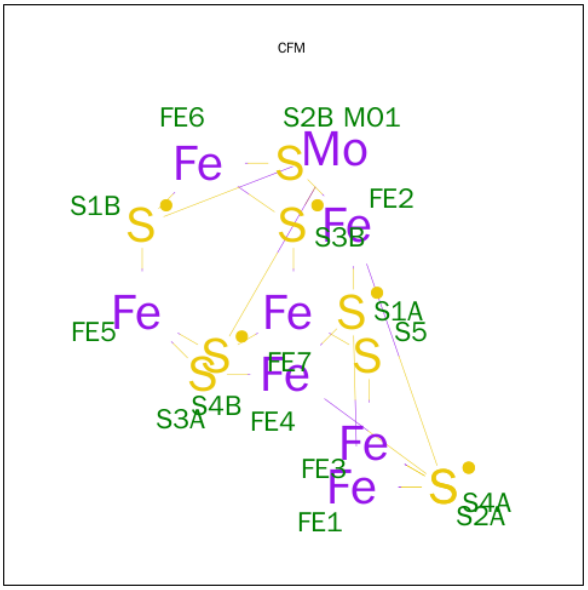
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	E	1	Total Fe S 8 4 4	0	0
5	G	1	Total Fe S 8 4 4	0	0

- Molecule 6 is 3-HYDROXY-3-CARBOXY-ADIPIC ACID (three-letter code: HCA) (formula: $C_7H_{10}O_7$).



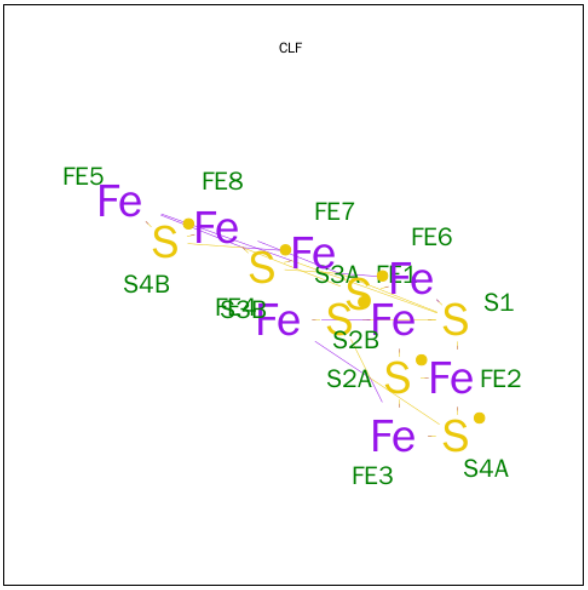
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 14 7 7	0	0
6	C	1	Total C O 14 7 7	0	0

- Molecule 7 is FE-MO-S CLUSTER (three-letter code: CFM) (formula: Fe_7MoS_9).



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

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



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

- Molecule 9 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	211	Total O 211 211	0	0
9	B	309	Total O 309 309	0	0
9	C	104	Total O 104 104	0	0
9	D	235	Total O 235 235	0	0
9	E	56	Total O 56 56	0	0
9	F	62	Total O 62 62	0	0
9	G	37	Total O 37 37	0	0
9	H	52	Total O 52 52	0	0

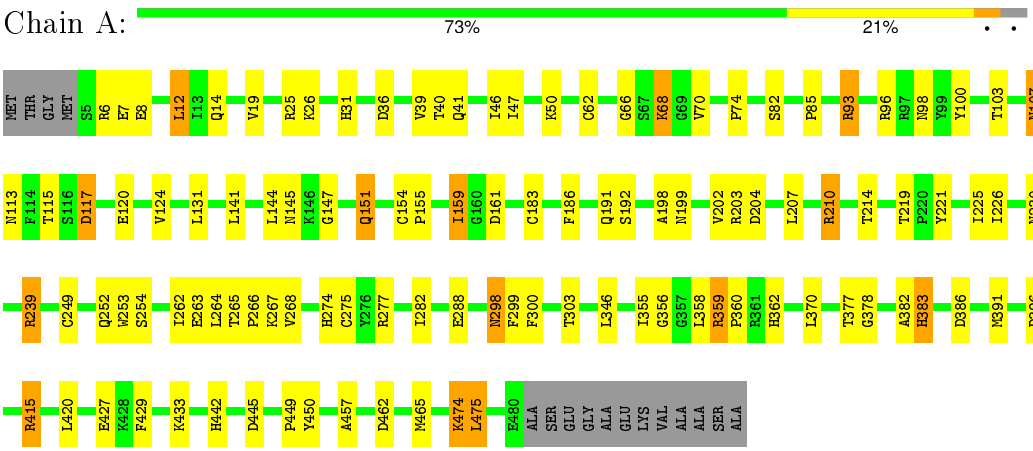
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

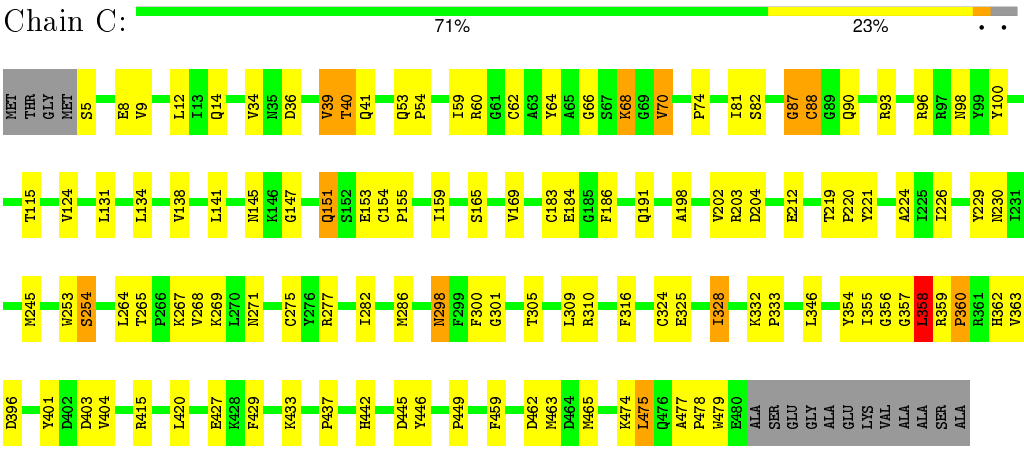
• Molecule 1: NITROGENASE MOLYBDENUM-IRON PROTEIN ALPHA CHAIN

Chain A:



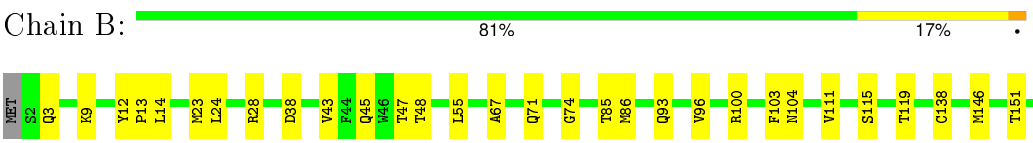
• Molecule 1: NITROGENASE MOLYBDENUM-IRON PROTEIN ALPHA CHAIN

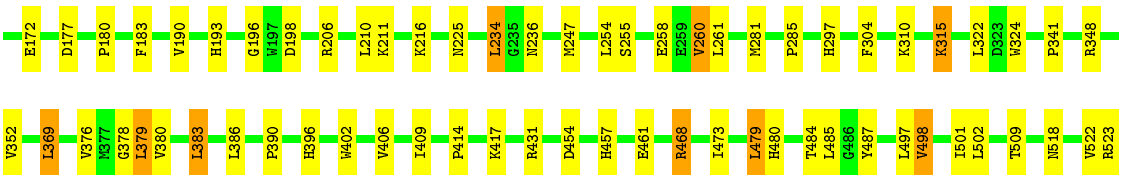
Chain C:



• Molecule 2: NITROGENASE MOLYBDENUM-IRON PROTEIN BETA CHAIN

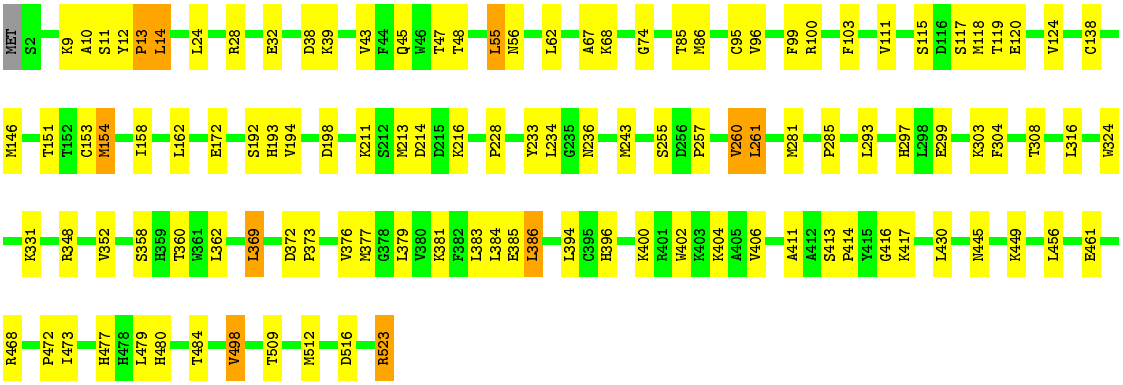
Chain B:





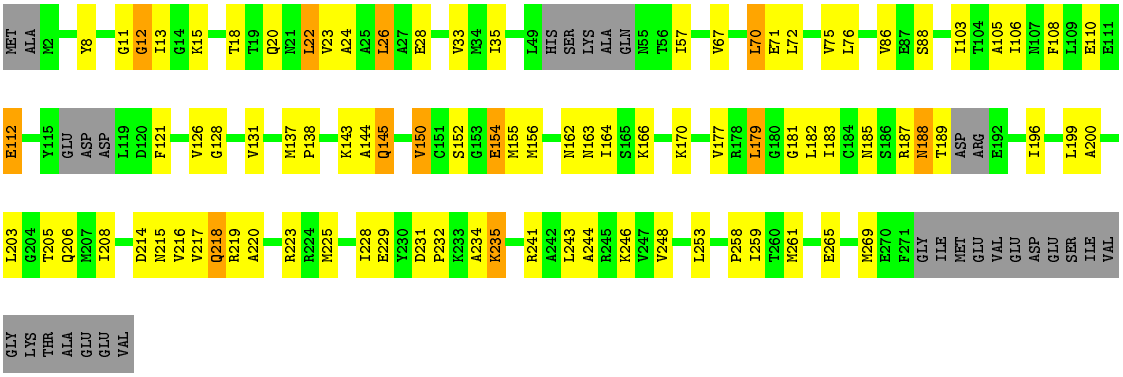
• Molecule 2: NITROGENASE MOLYBDENUM-IRON PROTEIN BETA CHAIN

Chain D:



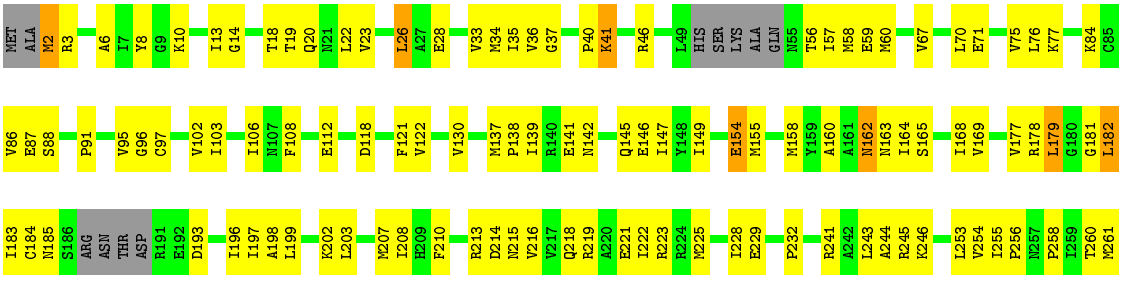
• Molecule 3: NITROGENASE IRON PROTEIN

Chain E:



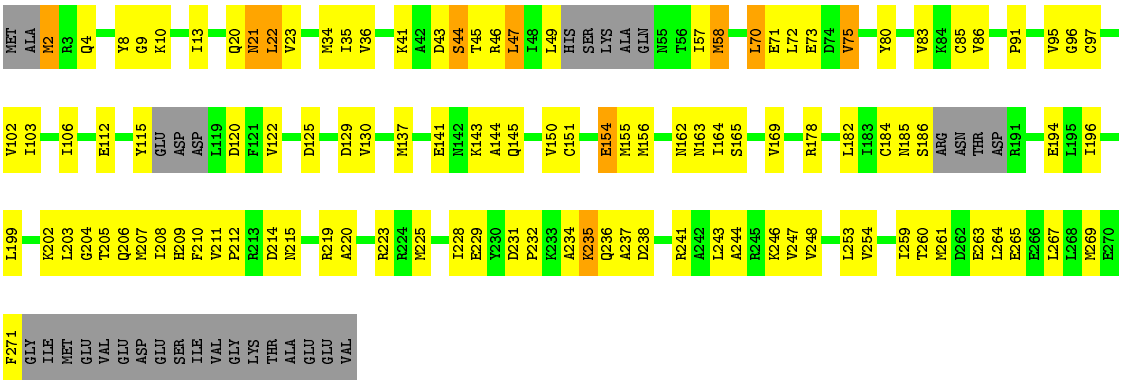
• Molecule 3: NITROGENASE IRON PROTEIN

Chain F:



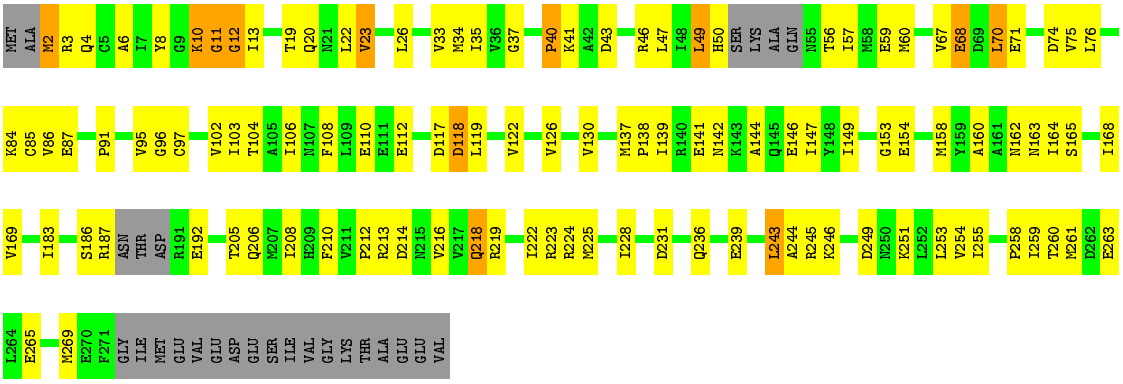
• Molecule 3: NITROGENASE IRON PROTEIN

Chain G:



• Molecule 3: NITROGENASE IRON PROTEIN

Chain H:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	264.24Å 111.46Å 121.59Å 90.00° 97.40° 90.00°	Depositor
Resolution (Å)	20.00 – 2.20	Depositor
% Data completeness (in resolution range)	89.1 (20.00-2.20)	Depositor
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.218 , 0.255	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	24903	wwPDB-VP
Average B, all atoms (Å^2)	37.0	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: SF4, CFM, CLF, CA, HCA

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	0.47	0/3864	0.87	4/5212 (0.1%)
1	C	0.49	0/3864	0.86	5/5212 (0.1%)
2	B	0.49	0/4280	0.82	3/5786 (0.1%)
2	D	0.44	0/4280	0.80	2/5786 (0.0%)
3	E	0.29	0/1973	0.63	0/2654
3	F	0.28	0/1973	0.62	0/2655
3	G	0.31	0/1958	0.60	1/2633 (0.0%)
3	H	0.32	0/2006	0.63	1/2699 (0.0%)
All	All	0.42	0/24198	0.77	16/32637 (0.0%)

There are no bond length outliers.

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	210	ARG	NE-CZ-NH2	-8.68	115.96	120.30
1	A	210	ARG	NE-CZ-NH1	7.90	124.25	120.30
1	C	420	LEU	CA-CB-CG	6.50	130.26	115.30
1	A	420	LEU	CA-CB-CG	6.30	129.78	115.30
1	C	88	CYS	CA-CB-SG	-6.02	103.17	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	3708	86	0
1	C	3776	0	3708	98	0
2	B	4174	0	4087	59	0
2	D	4174	0	4087	76	0
3	E	1953	0	1971	72	0
3	F	1952	0	1958	98	0
3	G	1938	0	1958	115	0
3	H	1984	0	1993	143	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
5	E	8	0	0	0	0
5	G	8	0	0	0	0
6	A	14	0	6	3	0
6	C	14	0	6	2	0
7	A	17	0	0	5	0
7	C	17	0	0	6	0
8	B	15	0	0	1	0
8	D	15	0	0	1	0
9	A	211	0	0	6	0
9	B	309	0	0	6	0
9	C	104	0	0	1	0
9	D	235	0	0	6	0
9	E	56	0	0	2	0
9	F	62	0	0	0	0
9	G	37	0	0	0	0
9	H	52	0	0	3	0
All	All	24903	0	23482	692	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 15.

The worst 5 of 692 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:G:231:ASP:OD2	3:G:234:ALA:HB2	1.31	1.25
1:C:253:TRP:CZ3	1:C:282:ILE:HD13	1.78	1.16
1:A:275:CYS:HA	1:A:358:LEU:HD22	1.28	1.14
3:F:23:VAL:HG21	3:F:35:ILE:HD11	1.32	1.04
3:G:23:VAL:HG21	3:G:35:ILE:HD11	1.40	1.01

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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%)	443 (94%)	28 (6%)	3 (1%)	30	29
1	C	474/492 (96%)	441 (93%)	29 (6%)	4 (1%)	24	22
2	B	520/523 (99%)	502 (96%)	17 (3%)	1 (0%)	52	59
2	D	520/523 (99%)	507 (98%)	11 (2%)	2 (0%)	39	42
3	E	251/289 (87%)	234 (93%)	16 (6%)	1 (0%)	39	42
3	F	254/289 (88%)	240 (94%)	12 (5%)	2 (1%)	24	22
3	G	249/289 (86%)	231 (93%)	16 (6%)	2 (1%)	24	22
3	H	256/289 (89%)	244 (95%)	10 (4%)	2 (1%)	24	22
All	All	2998/3186 (94%)	2842 (95%)	139 (5%)	17 (1%)	30	29

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	40	THR
3	F	118	ASP
1	A	40	THR
1	C	396	ASP
3	G	44	SER

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%)	375 (93%)	30 (7%)	17	17

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	C	405/415 (98%)	379 (94%)	26 (6%)	22 24
2	B	454/455 (100%)	429 (94%)	25 (6%)	27 30
2	D	454/455 (100%)	427 (94%)	27 (6%)	24 27
3	E	206/233 (88%)	186 (90%)	20 (10%)	10 9
3	F	205/233 (88%)	191 (93%)	14 (7%)	20 21
3	G	204/233 (88%)	194 (95%)	10 (5%)	31 36
3	H	209/233 (90%)	198 (95%)	11 (5%)	28 32
All	All	2542/2672 (95%)	2379 (94%)	163 (6%)	22 24

5 of 163 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	360	PRO
2	D	214	ASP
3	G	154	GLU
1	C	401	TYR
2	D	13	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 62 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	3	GLN
2	D	518	ASN
3	H	4	GLN
2	D	163	ASN
3	E	142	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 (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 10 ligands modelled in this entry, 2 are monoatomic - leaving 8 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	HCA	A	1494	-	4,13,13	1.48	2 (50%)	3,18,18	6.44	2 (66%)
7	CFM	A	1496	1	0,24,24	0.00	-	0,45,45	0.00	-
8	CLF	B	1498	1,2	0,24,24	0.00	-	0,57,57	0.00	-
6	HCA	C	3494	-	4,13,13	0.90	0	3,18,18	6.50	2 (66%)
7	CFM	C	3496	1	0,24,24	0.00	-	0,45,45	0.00	-
8	CLF	D	3498	1,2	0,24,24	0.00	-	0,57,57	0.00	-
5	SF4	E	5290	3	0,12,12	0.00	-	0,24,24	0.00	-
5	SF4	G	7290	3	0,12,12	0.00	-	0,24,24	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
6	HCA	A	1494	-	-	0/7/17/17	0/0/0/0
7	CFM	A	1496	1	-	0/0/84/84	0/0/8/8
8	CLF	B	1498	1,2	-	0/0/132/132	0/12/10/10
6	HCA	C	3494	-	-	0/7/17/17	0/0/0/0
7	CFM	C	3496	1	-	0/0/84/84	0/0/8/8
8	CLF	D	3498	1,2	-	0/0/132/132	0/12/10/10
5	SF4	E	5290	3	-	0/0/48/48	0/6/5/5
5	SF4	G	7290	3	-	0/0/48/48	0/6/5/5

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	1494	HCA	O7-C3	2.00	1.46	1.43
6	A	1494	HCA	C4-C3	2.17	1.56	1.53

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	1494	HCA	C4-C3-C2	5.34	129.24	111.33
6	C	3494	HCA	C4-C3-C2	5.73	130.55	111.33
6	C	3494	HCA	C3-C2-C1	9.69	130.45	114.96
6	A	1494	HCA	C3-C2-C1	9.78	130.60	114.96

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	1494	HCA	3	0
7	A	1496	CFM	5	0
8	B	1498	CLF	1	0
6	C	3494	HCA	2	0
7	C	3496	CFM	6	0
8	D	3498	CLF	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

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.