



# Full wwPDB X-ray Structure Validation Report ⓘ

Apr 21, 2016 – 08:35 AM EDT

PDB ID : 5BVG  
Title : Selenium incorporated nitrogenase MoFe-protein (Av1-Se2B) from *A. vinelandii*  
Authors : Spatzal, T.; Perez, K.A.; Howard, J.B.; Rees, D.C.  
Deposited on : 2015-06-05  
Resolution : 1.60 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027257  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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-20027257

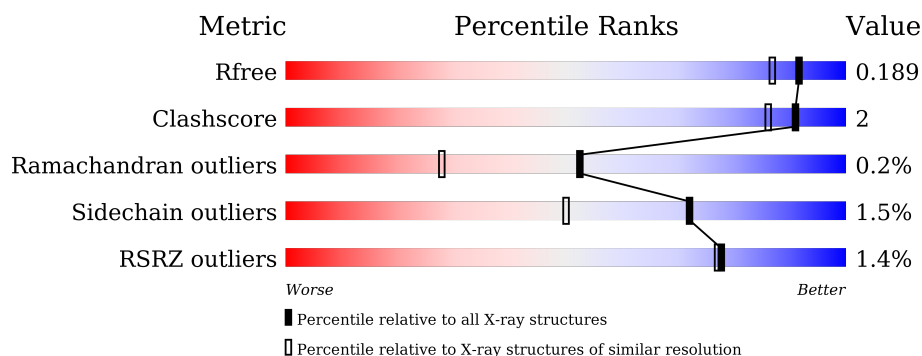
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.60 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))
$R_{free}$	91344	2475 (1.60-1.60)
Clashscore	102246	2732 (1.60-1.60)
Ramachandran outliers	100387	2654 (1.60-1.60)
Sidechain outliers	100360	2653 (1.60-1.60)
RSRZ outliers	91569	2479 (1.60-1.60)

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	A	492	<div> <div>2%</div> <div>90%</div> <div>6%</div> <div>.</div> </div>
1	C	492	<div> <div>%</div> <div>91%</div> <div>5%</div> <div>.</div> </div>
2	B	523	<div> <div>2%</div> <div>95%</div> <div>5%</div> <div>.</div> </div>
2	D	523	<div> <div>%</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
5	IMD	C	505	-	-	X	X
5	IMD	D	602	-	-	-	X
9	SE	D	604[A]	-	-	X	-

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 17833 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	9	0
			3863	2451	662	722	28			
1	C	477	Total	C	N	O	S	0	12	0
			3880	2460	670	722	28			

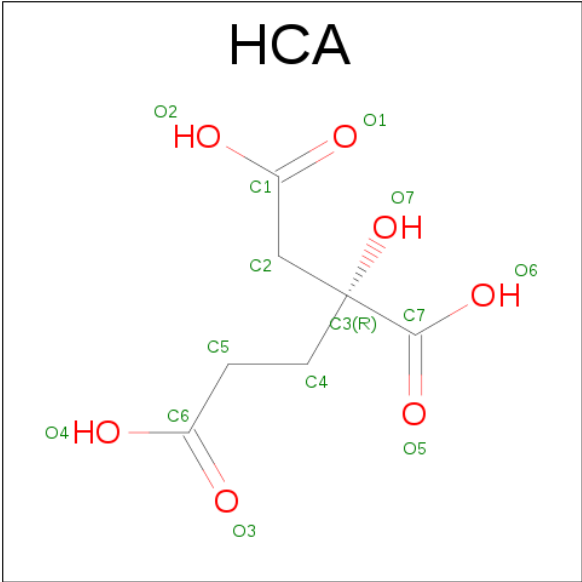
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	440	GLN	GLU	conflict	UNP P07328
C	440	GLN	GLU	conflict	UNP P07328

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

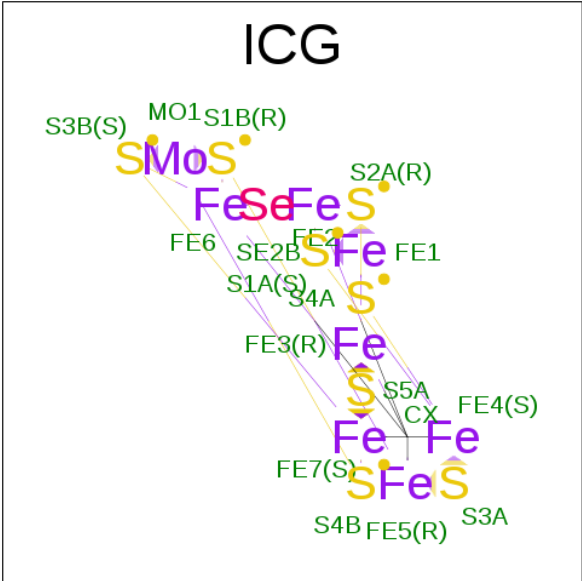
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	522	Total	C	N	O	S	0	12	0
			4257	2715	719	793	30			
2	D	522	Total	C	N	O	S	0	10	0
			4255	2712	718	795	30			

- Molecule 3 is 3-HYDROXY-3-CARBOXY-ADIPIC ACID (three-letter code: HCA) (formula: C<sub>7</sub>H<sub>10</sub>O<sub>7</sub>).



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

- Molecule 4 is iron-sulfur-molybdenum cluster with interstitial carbon with selenium incorporated (three-letter code: ICG) (formula: CFe<sub>7</sub>MoS<sub>8</sub>Se).



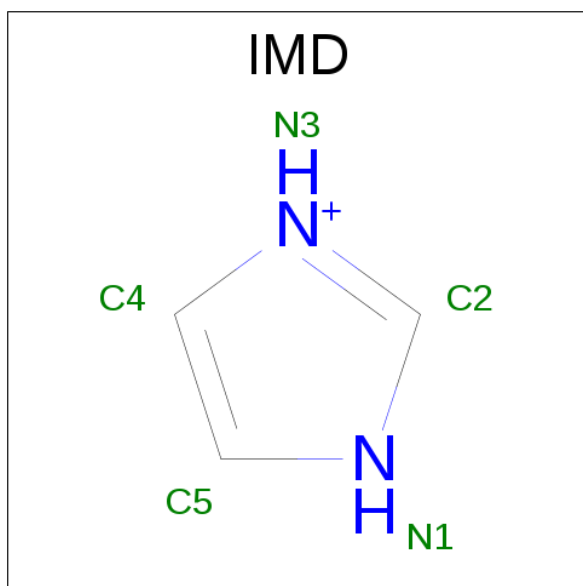
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
4	A	1	Total	C	Fe	Mo	S	Se	0	0
			18	1	7	1	8	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	C	1	Total	C	Fe	Mo	S	Se	0	0
			18	1	7	1	8	1		

- Molecule 5 is IMIDAZOLE (three-letter code: IMD) (formula:  $C_3H_5N_2$ ).

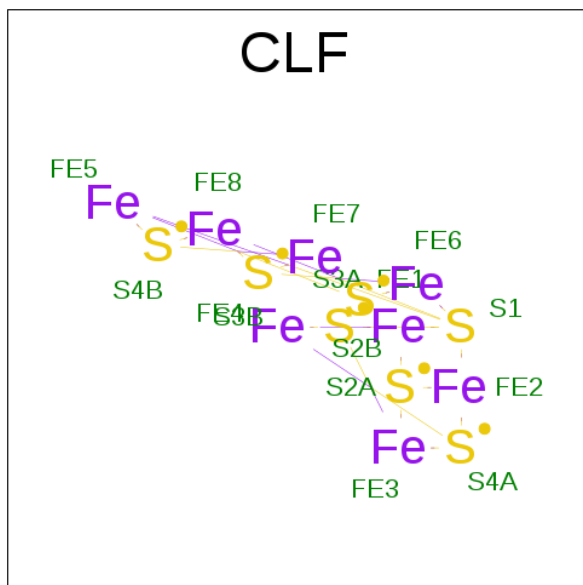


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	N	0	0
			5	3	2		
5	B	1	Total	C	N	0	0
			5	3	2		
5	B	1	Total	C	N	0	0
			5	3	2		
5	B	1	Total	C	N	0	0
			5	3	2		
5	C	1	Total	C	N	0	0
			5	3	2		
5	C	1	Total	C	N	0	0
			5	3	2		
5	D	1	Total	C	N	0	0
			5	3	2		
5	D	1	Total	C	N	0	0
			5	3	2		
5	D	1	Total	C	N	0	0
			5	3	2		

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula:  $Cl$ ).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Cl 1 1	0	0
6	C	1	Total Cl 1 1	0	0

- Molecule 7 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
7	A	1	Total Fe S 15 8 7	0	0
7	C	1	Total Fe S 15 8 7	0	0

- Molecule 8 is FE (II) ION (three-letter code: FE2) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	B	2	Total Fe 2 2	0	0

- Molecule 9 is SELENIUM ATOM (three-letter code: SE) (formula: Se).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	B	1	Total Se 1 1	0	1
9	D	1	Total Se 1 1	0	1

- Molecule 10 is water.

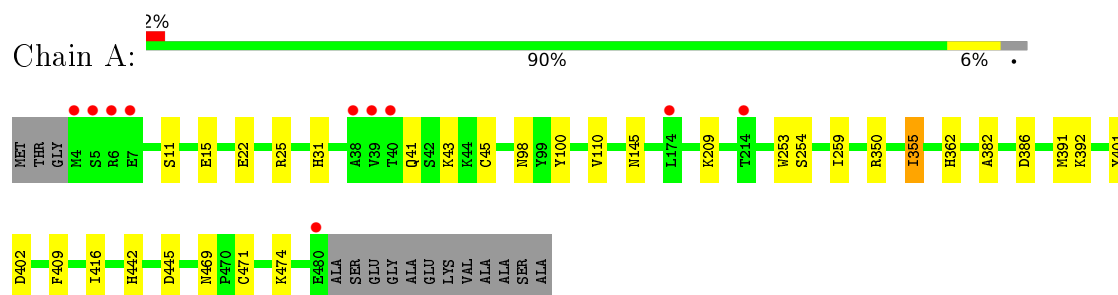
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	296	Total 296	O 296	0	0
10	B	393	Total 393	O 393	0	1
10	C	309	Total 309	O 309	0	0
10	D	435	Total 435	O 435	0	1



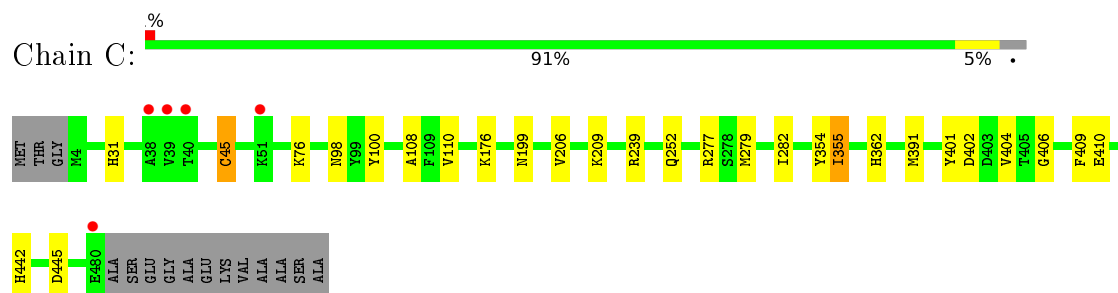
### 3 Residue-property plots

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.

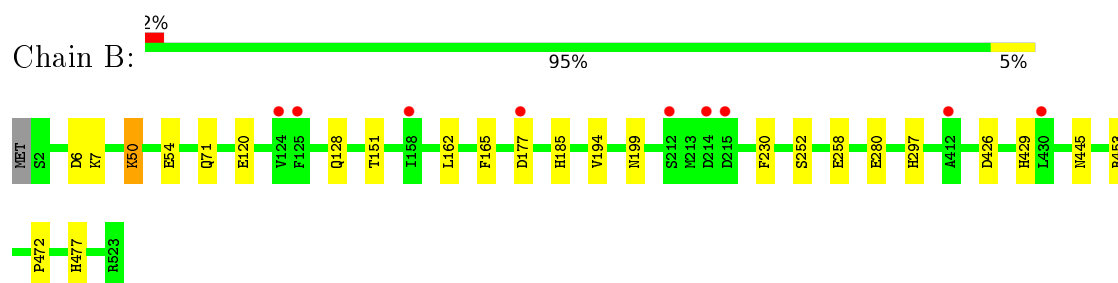
- Molecule 1: Nitrogenase molybdenum-iron protein alpha chain



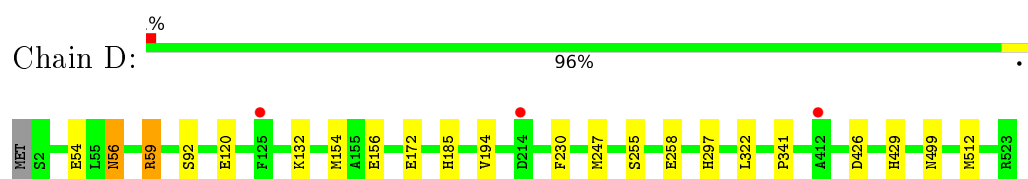
- Molecule 1: Nitrogenase molybdenum-iron protein alpha chain



- 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	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.58Å 130.82Å 107.13Å 90.00° 108.85° 90.00°	Depositor
Resolution (Å)	101.38 – 1.60 37.49 – 1.60	Depositor EDS
% Data completeness (in resolution range)	96.6 (101.38-1.60) 96.6 (37.49-1.60)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.23 (at 1.60Å)	Xtriage
Refinement program	REFMAC 5.8.0123	Depositor
R, $R_{free}$	0.152 , 0.179 0.164 , 0.189	Depositor DCC
$R_{free}$ test set	12691 reflections (5.21%)	DCC
Wilson B-factor (Å <sup>2</sup> )	10.2	Xtriage
Anisotropy	0.529	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 40.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.025 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	17833	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.46% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: IMD, CL, CLF, ICG, HCA, FE2, SE

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	0.43	0/3951	0.68	0/5325
1	C	0.44	0/3968	0.69	1/5346 (0.0%)
2	B	0.42	0/4363	0.66	0/5896
2	D	0.45	0/4360	0.67	1/5891 (0.0%)
All	All	0.43	0/16642	0.68	2/22458 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	59	ARG	NE-CZ-NH1	5.30	122.95	120.30
1	C	277	ARG	NE-CZ-NH1	5.29	122.94	120.30

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	A	3863	0	3794	17	0
1	C	3880	0	3797	16	0
2	B	4257	0	4141	17	0
2	D	4255	0	4134	17	0
3	A	14	0	6	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	14	0	6	1	0
4	A	18	0	0	0	0
4	C	18	0	0	0	0
5	A	5	0	5	0	0
5	B	15	0	15	2	0
5	C	10	0	10	6	0
5	D	15	0	15	2	0
6	A	1	0	0	1	0
6	C	1	0	0	1	0
7	A	15	0	0	0	0
7	C	15	0	0	0	0
8	B	2	0	0	0	0
9	B	1	0	0	1	0
9	D	1	0	0	2	0
10	A	296	0	0	1	0
10	B	393	0	0	2	0
10	C	309	0	0	5	0
10	D	435	0	0	4	0
All	All	17833	0	15923	72	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 2.

All (72) 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)
6:A:504:CL:CL	9:B:605[A]:SE:SE	2.53	1.14
6:C:504:CL:CL	9:D:604[A]:SE:SE	2.66	1.01
1:A:45[B]:CYS:SG	1:A:391:MET:CE	2.51	0.99
1:A:11[B]:SER:O	1:A:15[B]:GLU:HG3	1.67	0.94
1:A:45[B]:CYS:SG	1:A:391:MET:HE1	2.11	0.89
2:B:477:HIS:H	2:D:499:ASN:HD21	1.23	0.86
1:A:45[B]:CYS:SG	1:A:391:MET:HE2	2.18	0.84
1:C:45[A]:CYS:SG	1:C:391:MET:CE	2.72	0.76
1:C:45[A]:CYS:SG	1:C:391:MET:HE2	2.26	0.76
1:A:22:GLU:OE2	1:A:25:ARG:NH1	2.21	0.73
5:C:505:IMD:H5	10:C:848:HOH:O	1.89	0.71
1:C:206:VAL:HA	1:C:209[B]:LYS:HE2	1.74	0.69
5:D:601:IMD:H4	10:D:723:HOH:O	1.91	0.69
2:D:426:ASP:H	2:D:429:HIS:HD2	1.42	0.68
1:C:45[A]:CYS:SG	1:C:391:MET:HE1	2.38	0.63

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:280:GLU:OE1	5:B:601:IMD:H5	1.99	0.63
2:D:92:SER:OG	2:D:154[B]:MET:HG2	1.99	0.62
1:C:410:GLU:HB2	5:C:505:IMD:C2	2.31	0.61
2:B:426:ASP:H	2:B:429:HIS:HD2	1.48	0.61
5:C:505:IMD:C5	10:C:864:HOH:O	2.48	0.60
1:A:31:HIS:HE1	10:A:603:HOH:O	1.85	0.59
2:B:71:GLN:HE22	2:B:199:ASN:HD22	1.52	0.58
2:B:50:LYS:HE2	2:B:54:GLU:HG3	1.85	0.58
2:D:230:PHE:H	2:D:297:HIS:HE1	1.52	0.56
1:C:31:HIS:HD2	1:C:402:ASP:OD2	1.87	0.56
2:B:297:HIS:HD2	10:B:777:HOH:O	1.89	0.56
1:C:31:HIS:HE1	10:C:602:HOH:O	1.88	0.55
2:D:132:LYS:NZ	2:D:172[B]:GLU:OE2	2.25	0.55
1:A:31:HIS:HD2	1:A:402:ASP:OD2	1.90	0.55
2:D:297:HIS:HD2	10:D:737:HOH:O	1.89	0.55
2:D:230:PHE:H	2:D:297:HIS:CE1	2.25	0.55
1:C:239:ARG:HH11	1:C:252:GLN:HE21	1.56	0.53
2:D:185:HIS:HE1	10:D:830:HOH:O	1.91	0.52
1:C:199:ASN:HD21	1:C:279[A]:MET:HA	1.75	0.52
2:B:230:PHE:H	2:B:297:HIS:HE1	1.58	0.51
2:D:194:VAL:HB	2:D:297:HIS:CG	2.46	0.50
2:B:194:VAL:HB	2:B:297:HIS:CG	2.47	0.49
9:D:604[A]:SE:SE	10:D:896:HOH:O	2.81	0.49
1:A:442:HIS:CG	3:A:501:HCA:H52	2.48	0.48
1:A:350[B]:ARG:NH2	1:A:416:ILE:O	2.46	0.48
2:B:230:PHE:H	2:B:297:HIS:CE1	2.30	0.48
1:A:469:ASN:HD22	1:A:471:CYS:H	1.61	0.48
2:D:120:GLU:OE2	5:D:603:IMD:H4	2.14	0.47
2:B:426:ASP:H	2:B:429:HIS:CD2	2.32	0.47
1:C:100:TYR:CE1	1:C:110:VAL:HB	2.49	0.47
1:C:354:TYR:CZ	1:C:404:VAL:HG12	2.51	0.45
2:B:453[B]:ARG:HD2	2:D:512:MET:HE2	1.98	0.45
1:A:442:HIS:HB3	3:A:501:HCA:O6	2.16	0.45
1:C:442:HIS:HB3	3:C:501:HCA:O5	2.18	0.44
1:A:100:TYR:CE1	1:A:110:VAL:HB	2.52	0.44
2:D:426:ASP:H	2:D:429:HIS:CD2	2.29	0.44
2:B:185:HIS:HE1	10:B:853:HOH:O	2.00	0.43
2:B:120:GLU:OE2	5:B:603:IMD:H2	2.18	0.43
2:D:156:GLU:OE2	2:D:185:HIS:HD2	2.00	0.43
1:A:382:ALA:HB1	1:A:386:ASP:HB2	2.00	0.43
1:C:406:GLY:CA	5:C:505:IMD:H5	2.48	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:56:ASN:HD22	2:D:59:ARG:HE	1.68	0.42
2:D:247:MET:HG3	2:D:341:PRO:HD3	2.01	0.42
1:A:209:LYS:NZ	1:A:259:ILE:HD11	2.35	0.42
2:B:151:THR:HG23	2:B:162:LEU:HD11	2.00	0.42
1:A:474:LYS:HB3	2:D:322:LEU:HD21	2.01	0.42
5:C:505:IMD:C4	10:C:864:HOH:O	2.67	0.42
1:C:76:LYS:O	1:C:108:ALA:HA	2.20	0.41
1:C:199:ASN:HD22	1:C:282:ILE:H	1.67	0.41
2:B:445:ASN:HB2	2:B:472:PRO:O	2.20	0.41
2:B:128:GLN:HE22	2:B:165:PHE:HA	1.86	0.41
2:D:499:ASN:HA	2:D:499:ASN:HD22	1.68	0.41
2:B:50:LYS:CE	2:B:54:GLU:HG3	2.50	0.41
5:C:503:IMD:H2	10:C:830:HOH:O	2.21	0.41
1:A:41:GLN:HE21	1:A:43:LYS:H	1.68	0.40
1:C:199:ASN:HD21	1:C:279[B]:MET:HA	1.85	0.40
1:A:253:TRP:HA	1:A:254:SER:HA	1.85	0.40

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	484/492 (98%)	465 (96%)	18 (4%)	1 (0%)	52	28
1	C	487/492 (99%)	470 (96%)	16 (3%)	1 (0%)	52	28
2	B	532/523 (102%)	526 (99%)	6 (1%)	0	100	100
2	D	530/523 (101%)	523 (99%)	6 (1%)	1 (0%)	52	28
All	All	2033/2030 (100%)	1984 (98%)	46 (2%)	3 (0%)	52	31

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	255	SER
1	C	355	ILE
1	A	355	ILE

### 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	416/415 (100%)	408 (98%)	8 (2%)	65	39
1	C	415/415 (100%)	406 (98%)	9 (2%)	60	31
2	B	462/455 (102%)	454 (98%)	8 (2%)	68	44
2	D	461/455 (101%)	458 (99%)	3 (1%)	88	78
All	All	1754/1740 (101%)	1726 (98%)	28 (2%)	72	47

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	98	ASN
1	A	145	ASN
1	A	355	ILE
1	A	362	HIS
1	A	392	LYS
1	A	401	TYR
1	A	409	PHE
1	A	445	ASP
2	B	6[A]	ASP
2	B	6[B]	ASP
2	B	7	LYS
2	B	50	LYS
2	B	177	ASP
2	B	252[A]	SER
2	B	252[B]	SER
2	B	258	GLU
1	C	45[A]	CYS
1	C	45[B]	CYS
1	C	98	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	176	LYS
1	C	355	ILE
1	C	362	HIS
1	C	401	TYR
1	C	409	PHE
1	C	445	ASP
2	D	54	GLU
2	D	56	ASN
2	D	258	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (36) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	31	HIS
1	A	41	GLN
1	A	145	ASN
1	A	384	ASN
1	A	468	ASN
1	A	469	ASN
2	B	71	GLN
2	B	128	GLN
2	B	130	ASN
2	B	168	ASN
2	B	185	HIS
2	B	225	ASN
2	B	294	GLN
2	B	297	HIS
2	B	363	HIS
2	B	418	ASN
2	B	429	HIS
1	C	31	HIS
1	C	119	GLN
1	C	199	ASN
1	C	252	GLN
1	C	271	ASN
1	C	384	ASN
1	C	468	ASN
2	D	56	ASN
2	D	130	ASN
2	D	185	HIS
2	D	225	ASN
2	D	286	ASN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
2	D	294	GLN
2	D	297	HIS
2	D	363	HIS
2	D	418	ASN
2	D	429	HIS
2	D	499	ASN
2	D	518	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 21 ligands modelled in this entry, 6 are monoatomic - leaving 15 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	$\# Z  > 2$
3	HCA	A	501	-	4,13,13	0.52	0	3,18,18	0.93	0
4	ICG	A	502	1	6,30,30	2.03	2 (33%)	0,78,78	0.00	-
5	IMD	A	503	-	3,5,5	0.32	0	4,5,5	0.57	0
7	CLF	A	505	1,2	0,24,24	0.00	-	0,57,57	0.00	-
5	IMD	B	601	-	3,5,5	0.48	0	4,5,5	0.80	0
5	IMD	B	602	-	3,5,5	0.36	0	4,5,5	0.64	0
5	IMD	B	603	-	3,5,5	0.45	0	4,5,5	0.46	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	HCA	C	501	-	4,13,13	0.68	0	3,18,18	1.53	0
4	ICG	C	502	1	6,30,30	1.93	1 (16%)	0,78,78	0.00	-
5	IMD	C	503	-	3,5,5	0.48	0	4,5,5	0.58	0
5	IMD	C	505	-	3,5,5	0.58	0	4,5,5	0.97	0
7	CLF	C	506	1,2	0,24,24	0.00	-	0,57,57	0.00	-
5	IMD	D	601	-	3,5,5	0.52	0	4,5,5	0.56	0
5	IMD	D	602	-	3,5,5	0.43	0	4,5,5	0.61	0
5	IMD	D	603	-	3,5,5	0.41	0	4,5,5	0.39	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
3	HCA	A	501	-	-	0/7/17/17	0/0/0/0
4	ICG	A	502	1	-	0/0/204/204	0/0/13/13
5	IMD	A	503	-	-	0/0/0/0	0/1/1/1
7	CLF	A	505	1,2	-	0/0/132/132	0/12/10/10
5	IMD	B	601	-	-	0/0/0/0	0/1/1/1
5	IMD	B	602	-	-	0/0/0/0	0/1/1/1
5	IMD	B	603	-	-	0/0/0/0	0/1/1/1
3	HCA	C	501	-	-	0/7/17/17	0/0/0/0
4	ICG	C	502	1	-	0/0/204/204	0/0/13/13
5	IMD	C	503	-	-	0/0/0/0	0/1/1/1
5	IMD	C	505	-	-	0/0/0/0	0/1/1/1
7	CLF	C	506	1,2	-	0/0/132/132	0/12/10/10
5	IMD	D	601	-	-	0/0/0/0	0/1/1/1
5	IMD	D	602	-	-	0/0/0/0	0/1/1/1
5	IMD	D	603	-	-	0/0/0/0	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	502	ICG	S1B-FE6	-4.20	2.20	2.25
4	A	502	ICG	S1B-FE6	-4.07	2.20	2.25
4	A	502	ICG	S5A-FE7	-2.09	2.20	2.24

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	501	HCA	2	0
5	B	601	IMD	1	0
5	B	603	IMD	1	0
3	C	501	HCA	1	0
5	C	503	IMD	1	0
5	C	505	IMD	5	0
5	D	601	IMD	1	0
5	D	603	IMD	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 ⓘ

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	477/492 (96%)	0.01	10 (2%) 67 65	8, 13, 22, 38	11 (2%)
1	C	477/492 (96%)	-0.07	5 (1%) 84 84	7, 12, 21, 26	10 (2%)
2	B	522/523 (99%)	-0.12	9 (1%) 73 71	8, 14, 22, 29	11 (2%)
2	D	522/523 (99%)	-0.22	3 (0%) 90 90	7, 11, 18, 26	7 (1%)
All	All	1998/2030 (98%)	-0.10	27 (1%) 78 77	7, 13, 21, 38	39 (1%)

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	39	VAL	6.7
2	D	125	PHE	6.1
1	A	38	ALA	4.6
1	C	38	ALA	4.6
2	B	125	PHE	4.5
1	C	39	VAL	4.0
1	A	4	MET	3.6
2	B	124	VAL	3.4
1	A	40	THR	3.4
1	A	5	SER	3.0
2	B	214	ASP	3.0
1	A	7	GLU	2.7
2	B	430	LEU	2.7
1	A	174	LEU	2.6
2	D	214	ASP	2.6
2	D	412	ALA	2.4
2	B	212	SER	2.4
1	C	40[A]	THR	2.4
1	C	51[A]	LYS	2.4
2	B	158[A]	ILE	2.3
1	C	480	GLU	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	412	ALA	2.1
1	A	480	GLU	2.1
1	A	214	THR	2.0
1	A	6	ARG	2.0
2	B	215	ASP	2.0
2	B	177	ASP	2.0

## 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 carbohydrates 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. 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
5	IMD	C	505	5/5	0.72	0.39	14.97	21,22,23,24	0
5	IMD	D	602	5/5	0.90	0.14	12.80	23,24,24,25	0
5	IMD	B	601	5/5	0.95	0.10	1.59	20,21,21,22	0
5	IMD	D	601	5/5	0.96	0.10	1.03	13,15,15,16	0
5	IMD	B	602	5/5	0.96	0.10	0.88	16,16,17,18	0
6	CL	A	504	1/1	1.00	0.11	0.57	13,13,13,13	1
5	IMD	B	603	5/5	0.94	0.13	-0.03	21,21,23,24	0
5	IMD	D	603	5/5	0.94	0.12	-0.06	20,22,23,23	0
6	CL	C	504	1/1	1.00	0.10	-0.55	13,13,13,13	1
3	HCA	A	501	14/14	0.97	0.09	-0.70	9,9,10,10	0
7	CLF	C	506	15/15	0.99	0.07	-1.26	7,8,9,10	0
7	CLF	A	505	15/15	0.99	0.06	-1.27	10,11,12,14	0
3	HCA	C	501	14/14	0.98	0.09	-1.65	7,8,9,9	0
8	FE2	B	606	1/1	1.00	0.06	-2.03	10,10,10,10	1
4	ICG	C	502	18/18	0.99	0.07	-2.16	8,8,9,10	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
4	ICG	A	502	18/18	0.99	0.07	-2.32	8,9,10,11	0
9	SE	D	604[A]	1/1	1.00	0.02	-2.46	8,8,8,8	1
9	SE	B	605[A]	1/1	1.00	0.04	-2.72	12,12,12,12	1
8	FE2	B	604	1/1	1.00	0.05	-3.10	10,10,10,10	1
5	IMD	C	503	5/5	0.91	0.14	-	17,19,21,22	0
5	IMD	A	503	5/5	0.96	0.09	-	21,22,23,23	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.