wwPDB X-ray Structure Validation Summary Report (i)

Feb 14, 2024 – 11:59 AM EST

PDB ID	:	3MIN
Title	:	NITROGENASE MOFE PROTEIN FROM AZOTOBACTER VINELANDII, OXIDIZED STATE
Authors	:	Peters, J.W.; Stowell, M.H.B.; Soltis, S.M.; Day, M.W.; Kim, J.; Rees, D.C.
Deposited on	:	1996-12-20
Resolution	:	2.03 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

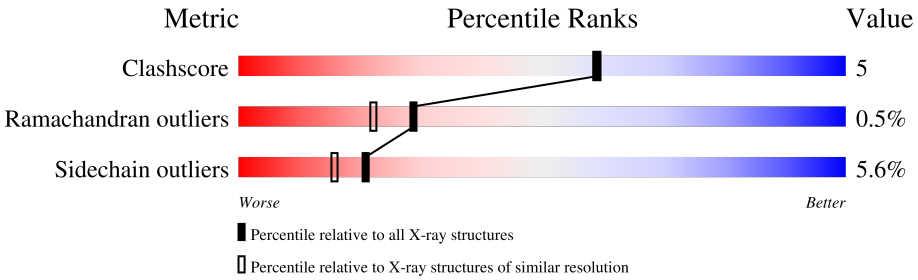
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.03 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	141614	11643 (2.04-2.00)
Ramachandran outliers	138981	11493 (2.04-2.00)
Sidechain outliers	138945	11492 (2.04-2.00)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	491	74% 16% • 5%
1	C	491	76% 15% • 5%
2	B	522	82% 15% •
2	D	522	83% 16% •

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 16490 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.

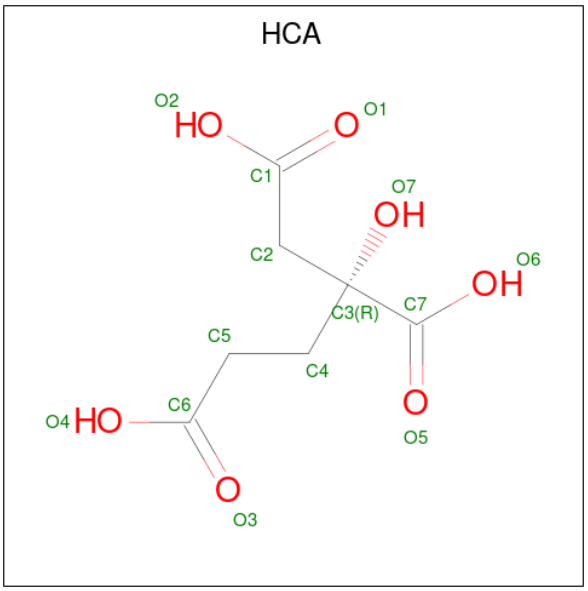
- Molecule 1 is a protein called NITROGENASE MOLYBDENUM IRON PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	467	Total	C	N	O	S	0	0	0
			3709	2361	630	694	24			
1	C	468	Total	C	N	O	S	0	0	0
			3713	2364	631	694	24			

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

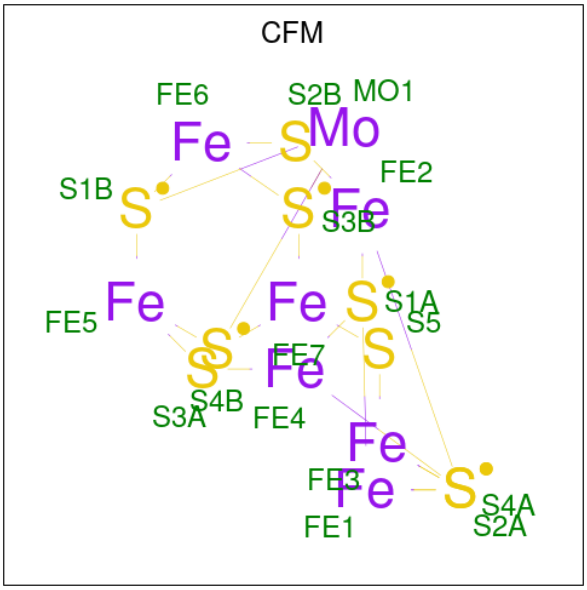
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	522	Total	C	N	O	S	0	0	0
			4174	2666	705	775	28			
2	D	522	Total	C	N	O	S	0	0	0
			4173	2666	705	774	28			

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



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

- Molecule 4 is FE-MO-S CLUSTER (three-letter code: CFM) (formula: Fe_7MoS_9).

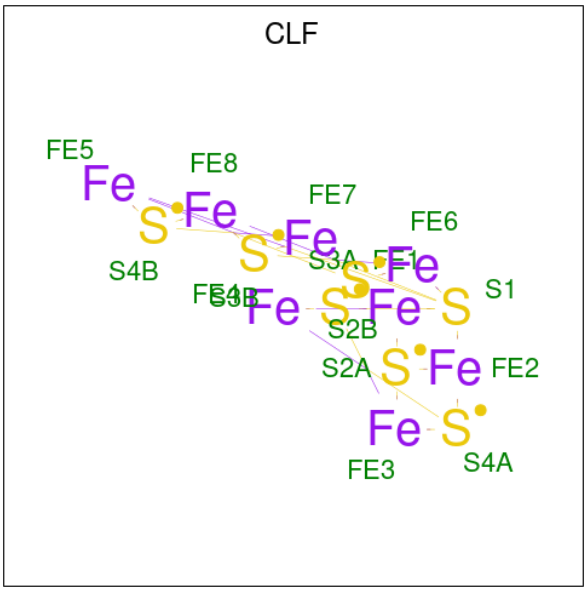


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Fe Mo S 17 7 1 9	0	0
4	C	1	Total Fe Mo S 17 7 1 9	0	0

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



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 water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	134	Total O 134 134	0	0
7	B	185	Total O 185 185	0	0
7	C	129	Total O 129 129	0	0
7	D	179	Total O 179 179	0	0

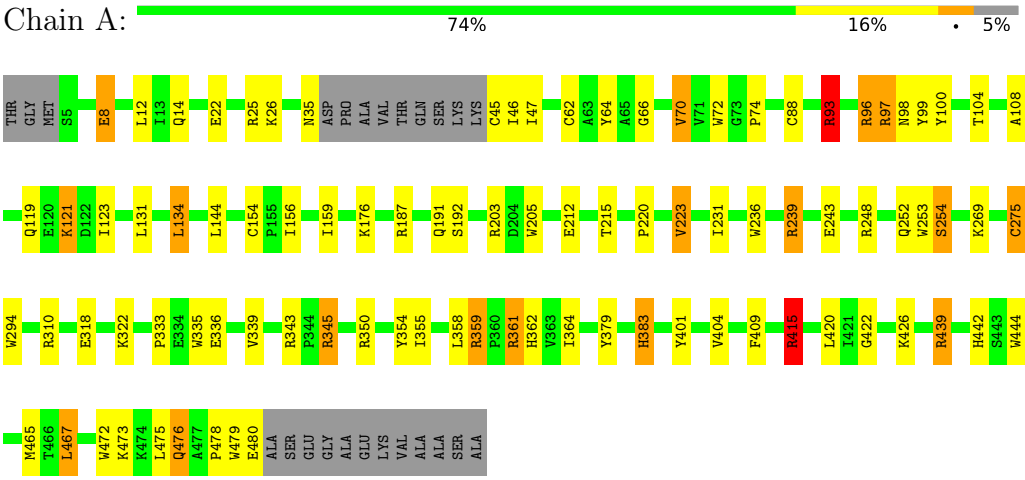
3 Residue-property plots (i)

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. 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. 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.

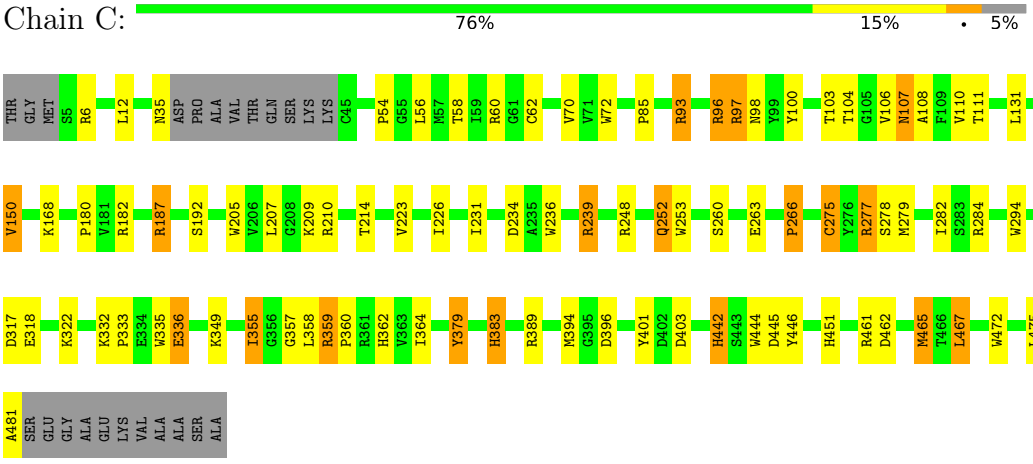
• Molecule 1: NITROGENASE MOLYBDENUM IRON PROTEIN

Chain A:



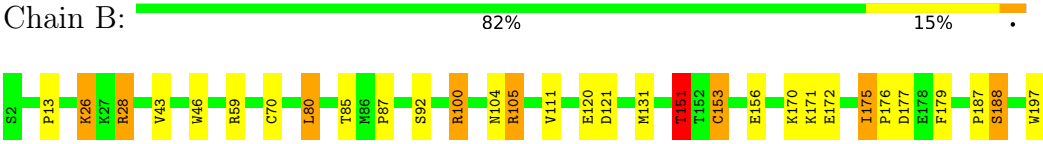
• Molecule 1: NITROGENASE MOLYBDENUM IRON PROTEIN

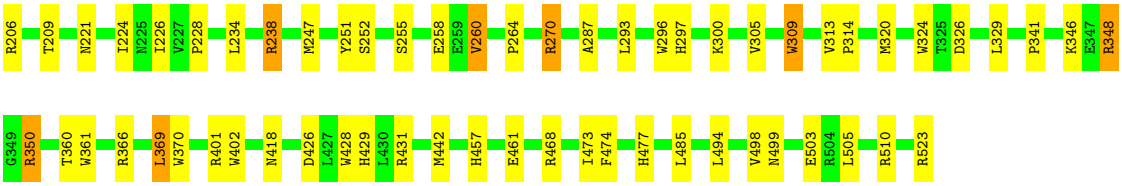
Chain C:



• Molecule 2: NITROGENASE MOLYBDENUM IRON PROTEIN

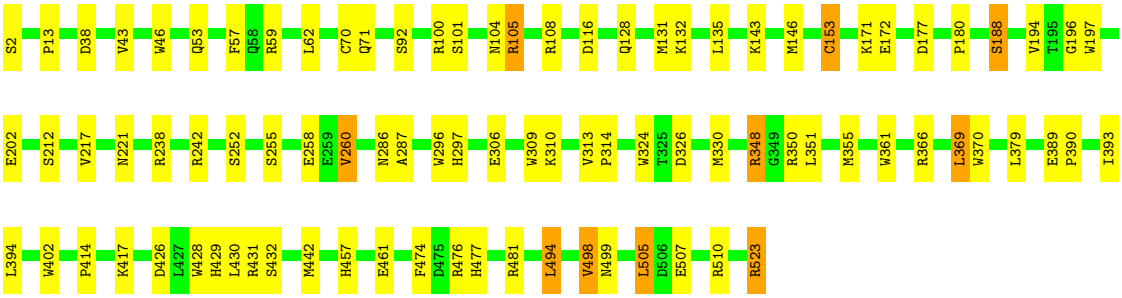
Chain B:





• Molecule 2: NITROGENASE MOLYBDENUM IRON PROTEIN

Chain D:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	108.00Å 131.30Å 81.00Å 90.00° 110.70° 90.00°	Depositor
Resolution (Å)	30.00 – 2.03	Depositor
% Data completeness (in resolution range)	89.6 (30.00-2.03)	Depositor
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR 3.98	Depositor
R, R_{free}	0.206 , 0.264	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	16490	wwPDB-VP
Average B, all atoms (Å^2)	29.0	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: CA, HCA, CFM, CLF

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	0.77	3/3795 (0.1%)	1.45	66/5117 (1.3%)
1	C	0.75	3/3799 (0.1%)	1.39	53/5123 (1.0%)
2	B	0.77	2/4280 (0.0%)	1.34	63/5786 (1.1%)
2	D	0.75	2/4279 (0.0%)	1.34	58/5785 (1.0%)
All	All	0.76	10/16153 (0.1%)	1.38	240/21811 (1.1%)

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	1
1	C	0	1
All	All	0	2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	153	CYS	CB-SG	-14.54	1.57	1.82
2	D	153	CYS	CB-SG	-10.80	1.63	1.82
1	A	275	CYS	CB-SG	10.34	1.99	1.82
1	A	88	CYS	CB-SG	-9.51	1.66	1.82
1	C	275	CYS	CB-SG	8.26	1.96	1.82

The worst 5 of 240 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	345	ARG	NE-CZ-NH1	12.69	126.64	120.30
1	A	359	ARG	NE-CZ-NH1	12.57	126.59	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	361	ARG	NE-CZ-NH1	12.51	126.55	120.30
2	D	523	ARG	NE-CZ-NH2	-11.37	114.62	120.30
1	A	96	ARG	NE-CZ-NH2	-10.90	114.85	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	446	TYR	Sidechain
1	C	446	TYR	Sidechain

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	3709	0	3637	43	0
1	C	3713	0	3639	42	0
2	B	4174	0	4088	41	0
2	D	4173	0	4088	36	0
3	A	14	0	6	0	0
3	C	14	0	6	0	0
4	A	17	0	0	1	0
4	C	17	0	0	2	0
5	B	1	0	0	0	0
5	D	1	0	0	0	0
6	B	15	0	0	0	0
6	D	15	0	0	1	0
7	A	134	0	0	9	0
7	B	185	0	0	1	0
7	C	129	0	0	2	0
7	D	179	0	0	4	0
All	All	16490	0	15464	148	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 5.

The worst 5 of 148 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)
2:B:499:ASN:HD21	2:D:477:HIS:H	1.19	0.86
1:A:93:ARG:HG3	1:A:113:ASN:HB2	1.60	0.84
2:B:209:THR:HG21	2:B:309:TRP:HE1	1.45	0.81
2:D:131:MET:HE2	2:D:135:LEU:HD11	1.67	0.77
2:B:209:THR:HG21	2:B:309:TRP:NE1	2.02	0.75

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	463/491 (94%)	443 (96%)	17 (4%)	3 (1%)	25	18
1	C	464/491 (94%)	432 (93%)	28 (6%)	4 (1%)	17	10
2	B	520/522 (100%)	496 (95%)	23 (4%)	1 (0%)	47	43
2	D	520/522 (100%)	501 (96%)	18 (4%)	1 (0%)	47	43
All	All	1967/2026 (97%)	1872 (95%)	86 (4%)	9 (0%)	29	22

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	318	GLU
1	C	6	ARG
1	A	254	SER
2	B	255	SER
2	D	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	397/414 (96%)	367 (92%)	30 (8%)	13 8
1	C	396/414 (96%)	371 (94%)	25 (6%)	18 12
2	B	454/454 (100%)	430 (95%)	24 (5%)	22 17
2	D	454/454 (100%)	438 (96%)	16 (4%)	36 34
All	All	1701/1736 (98%)	1606 (94%)	95 (6%)	21 16

5 of 95 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	106	VAL
1	C	362	HIS
1	C	131	LEU
1	C	214	THR
1	C	401	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	252	GLN
2	D	499	ASN
1	C	383	HIS
2	D	518	ASN
2	D	168	ASN

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 8 ligands modelled in this entry, 2 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	$\#\|Z\| > 2$
3	HCA	A	494	4	13,13,13	2.86	3 (23%)	14,18,18	2.96	6 (42%)
3	HCA	C	494	4	13,13,13	2.99	4 (30%)	14,18,18	2.83	7 (50%)
4	CFM	C	496	1,3	0,24,24	-	-	-		
6	CLF	D	525	1,2	0,24,24	-	-	-		
6	CLF	B	525	1,2	0,24,24	-	-	-		
4	CFM	A	496	1,3	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
6	CLF	D	525	1,2	-	-	0/12/10/10
6	CLF	B	525	1,2	-	-	0/12/10/10
3	HCA	A	494	4	-	6/17/17/17	-
3	HCA	C	494	4	-	4/17/17/17	-

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	494	HCA	C3-C7	-8.26	1.44	1.53
3	A	494	HCA	C3-C7	-7.92	1.45	1.53
3	A	494	HCA	O5-C7	5.50	1.39	1.22
3	C	494	HCA	O5-C7	5.15	1.38	1.22
3	C	494	HCA	C4-C3	3.04	1.59	1.54

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	494	HCA	O5-C7-C3	-6.31	113.32	122.25
3	C	494	HCA	O5-C7-C3	-6.24	113.42	122.25
3	A	494	HCA	O6-C7-C3	5.61	122.79	113.05
3	C	494	HCA	O6-C7-C3	5.49	122.58	113.05
3	A	494	HCA	C3-C2-C1	3.42	122.09	113.81

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	494	HCA	C2-C3-C4-C5
3	A	494	HCA	C7-C3-C4-C5
3	A	494	HCA	O7-C3-C4-C5
3	C	494	HCA	C2-C3-C4-C5
3	A	494	HCA	C3-C4-C5-C6

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	496	CFM	2	0
6	D	525	CLF	1	0
4	A	496	CFM	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 is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.