



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 20, 2018 – 10:43 pm GMT

PDB ID : 1FP4  
Title : CRYSTAL STRUCTURE OF THE ALPHA-H195Q MUTANT OF NITRO-GENASE  
Authors : Sorlie, M.; Christiansen, J.; Lemon, B.J.; Peters, J.W.; Dean, D.R.; Hales, B.J.  
Deposited on : 2000-08-30  
Resolution : 2.50 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30686

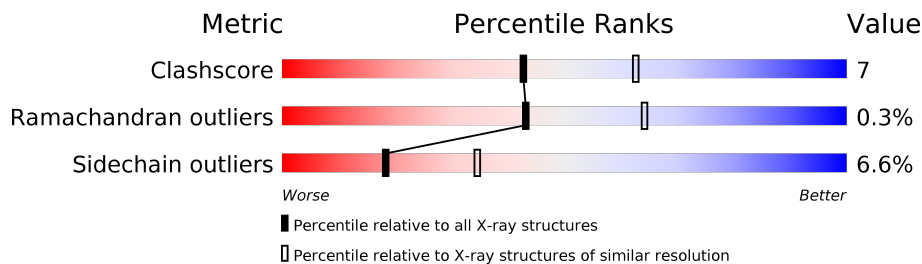
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	122078	4826 (2.50-2.50)
Ramachandran outliers	120005	4734 (2.50-2.50)
Sidechain outliers	119972	4736 (2.50-2.50)

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	

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
4	CFM	A	496	-	-	X	-
4	CFM	C	497	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 16295 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	467	Total	C	N	O	S	0	0	0
			3708	2360	629	695	24			
1	C	468	Total	C	N	O	S	0	0	0
			3712	2363	630	695	24			

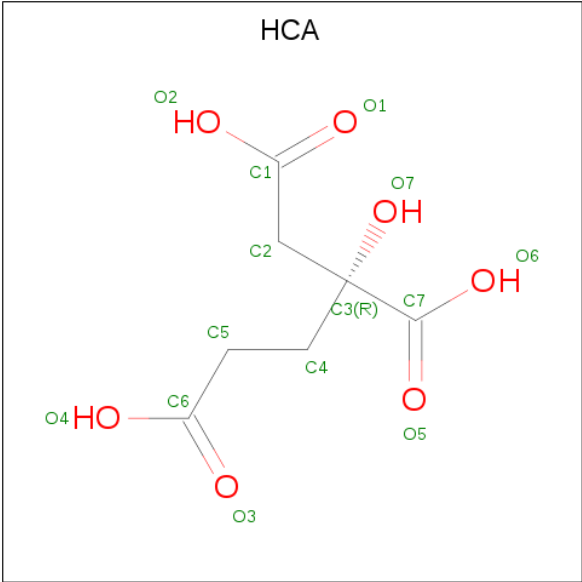
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	195	GLN	HIS	ENGINEERED	UNP P07328
C	195	GLN	HIS	ENGINEERED	UNP P07328

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

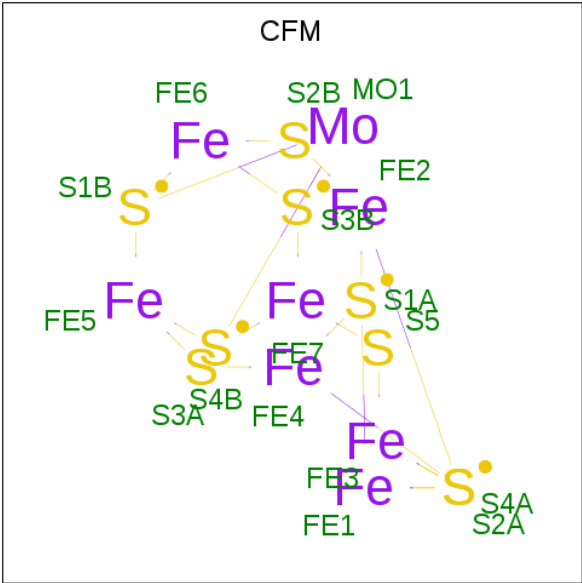
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	522	Total	C	N	O	S	0	0	0
			4174	2666	705	775	28			
2	D	522	Total	C	N	O	S	0	0	0
			4174	2666	705	775	28			

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



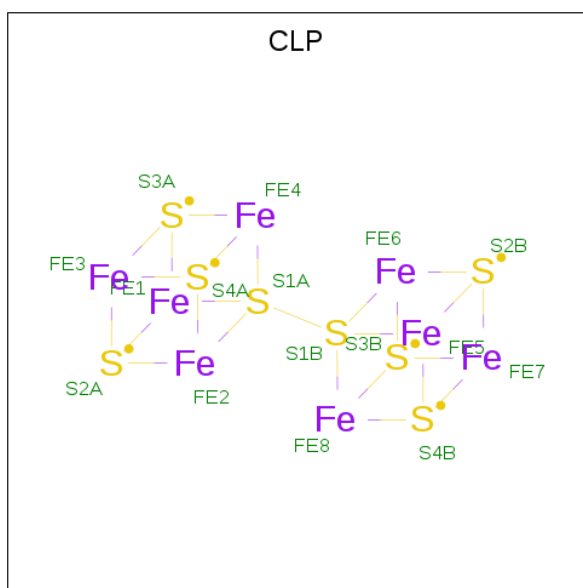
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			14	7	7		
3	C	1	Total	C	O	0	0
			14	7	7		

- Molecule 4 is 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	0	0
			17	7	1	9		
4	C	1	Total	Fe	Mo	S	0	0
			17	7	1	9		

- Molecule 5 is FE-S CLUSTER (three-letter code: CLP) (formula:  $\text{Fe}_8\text{S}_8$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	Fe	S	0	0
			15	8	7		
5	C	1	Total	Fe	S	0	0
			15	8	7		

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 7 is water.

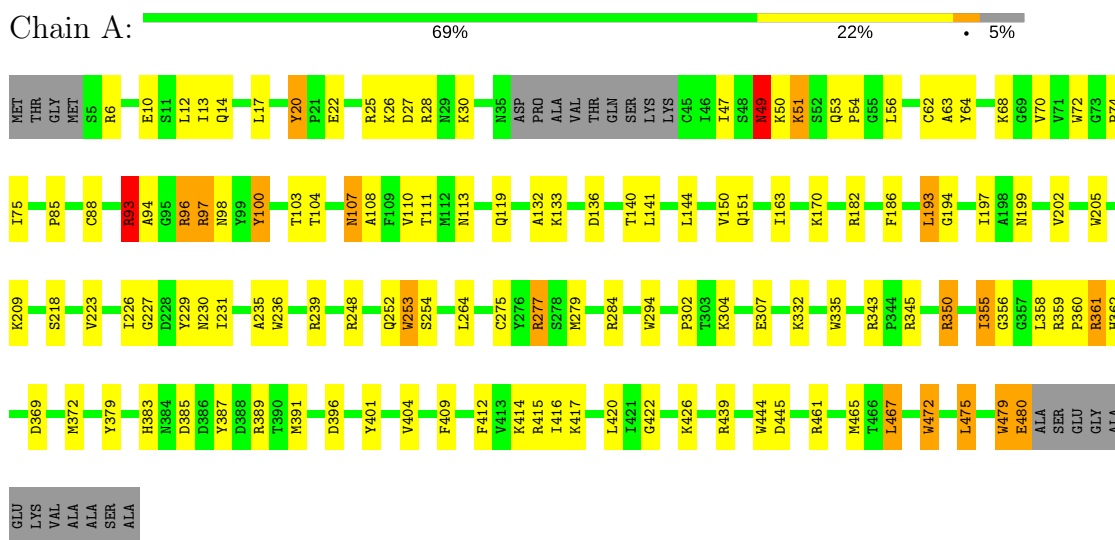
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	102	Total	O	0	0
			102	102		
7	B	114	Total	O	0	0
			114	114		
7	C	74	Total	O	0	0
			74	74		
7	D	143	Total	O	0	0
			143	143		

### 3 Residue-property plots

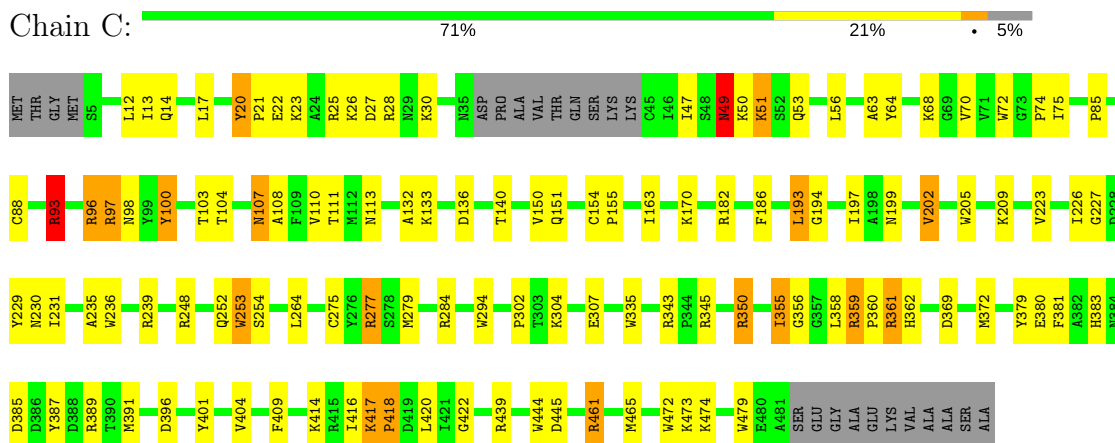
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are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

Note EDS was not executed.

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

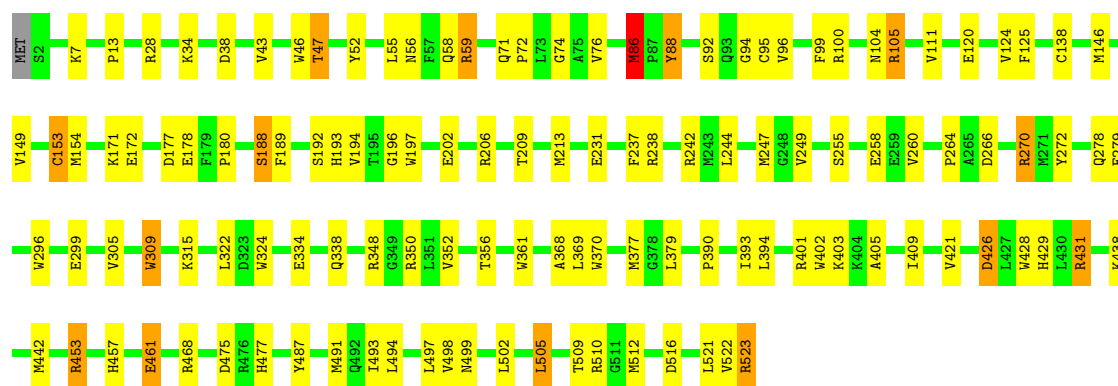


#### • Molecule 1: NITROGENASE MOLYBDENUM-IRON PROTEIN ALPHA CHAIN



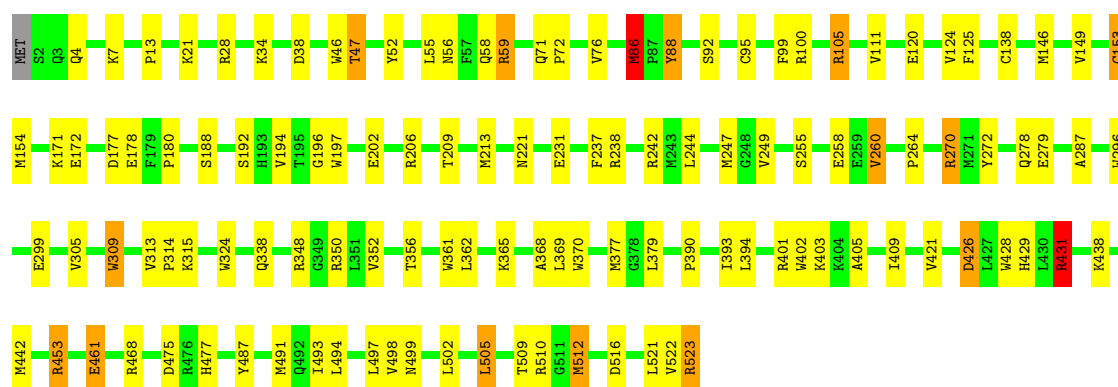
#### • Molecule 2: NITROGENASE MOLYBDENUM-IRON PROTEIN BETA CHAIN





• Molecule 2: NITROGENASE MOLYBDENUM-IRON PROTEIN BETA CHAIN

Chain D: 77% 20%



## 4 Data and refinement statistics

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, $\alpha$ , $\beta$ , $\gamma$	107.20Å 130.20Å 80.40Å 90.00° 111.20° 90.00°	Depositor
Resolution (Å)	20.00 – 2.50	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-2.50)	Depositor
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.184 , 0.240	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	16295	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	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, CLP, CFM

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	0.74	0/3793	1.41	49/5114 (1.0%)
1	C	0.85	1/3797 (0.0%)	1.44	56/5120 (1.1%)
2	B	0.72	0/4280	1.31	55/5786 (1.0%)
2	D	0.72	0/4280	1.30	56/5786 (1.0%)
All	All	0.76	1/16150 (0.0%)	1.36	216/21806 (1.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	418	PRO	N-CD	-26.97	1.10	1.47

The worst 5 of 216 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	49	ASN	CB-CA-C	-14.28	81.84	110.40
1	C	49	ASN	CB-CA-C	-13.93	82.54	110.40
1	C	418	PRO	CA-N-CD	13.66	130.82	111.70
1	C	418	PRO	N-CA-CB	-10.43	90.78	103.30
1	C	50	LYS	N-CA-CB	-10.30	92.05	110.60

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	3708	0	3638	70	0
1	C	3712	0	3640	61	0
2	B	4174	0	4087	67	0
2	D	4174	0	4087	59	0
3	A	14	0	6	0	0
3	C	14	0	6	0	0
4	A	17	0	0	5	0
4	C	17	0	0	6	0
5	A	15	0	0	3	0
5	C	15	0	0	2	0
6	B	1	0	0	0	0
6	D	1	0	0	0	0
7	A	102	0	0	7	0
7	B	114	0	0	2	0
7	C	74	0	0	2	0
7	D	143	0	0	2	0
All	All	16295	0	15464	232	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 7.

The worst 5 of 232 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:146:MET:HG3	2:B:180:PRO:HB2	1.62	0.82
2:D:390:PRO:HB2	2:D:393:ILE:HD11	1.63	0.81
2:B:390:PRO:HB2	2:B:393:ILE:HD11	1.64	0.80
1:C:350:ARG:NH1	1:C:416:ILE:O	2.17	0.77
2:D:146:MET:HG3	2:D:180:PRO:HB2	1.66	0.77

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	463/492 (94%)	433 (94%)	28 (6%)	2 (0%)	36	57
1	C	464/492 (94%)	437 (94%)	26 (6%)	1 (0%)	49	71
2	B	520/523 (99%)	508 (98%)	11 (2%)	1 (0%)	49	71
2	D	520/523 (99%)	507 (98%)	12 (2%)	1 (0%)	49	71
All	All	1967/2030 (97%)	1885 (96%)	77 (4%)	5 (0%)	43	64

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	6	ARG
2	B	255	SER
2	D	255	SER
1	A	355	ILE
1	C	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	397/415 (96%)	367 (92%)	30 (8%)	14	28
1	C	396/415 (95%)	369 (93%)	27 (7%)	17	33
2	B	454/455 (100%)	426 (94%)	28 (6%)	20	38
2	D	454/455 (100%)	427 (94%)	27 (6%)	21	40
All	All	1701/1740 (98%)	1589 (93%)	112 (7%)	18	34

5 of 112 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	338	GLN
1	C	93	ARG
2	D	315	LYS
2	B	350	ARG
2	B	505	LEU

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

Mol	Chain	Res	Type
2	B	518	ASN
1	C	230	ASN
2	D	513	GLN
2	B	519	HIS
1	C	252	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 carbohydrates 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	0.41	0	3,18,18	0.69	0
4	CFM	A	496	1	0,24,24	0.00	-	0,45,45	0.00	-
5	CLP	A	498	1,2	0,21,25	0.00	-	0,39,54	0.00	-
3	HCA	C	495	-	4,13,13	0.56	0	3,18,18	0.85	0
4	CFM	C	497	1	0,24,24	0.00	-	0,45,45	0.00	-
5	CLP	C	499	1,2	0,21,25	0.00	-	0,39,54	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
3	HCA	A	494	-	-	0/7/17/17	0/0/0/0
4	CFM	A	496	1	-	0/0/84/84	0/0/8/8
5	CLP	A	498	1,2	-	0/0/72/117	0/9/9/10
3	HCA	C	495	-	-	0/7/17/17	0/0/0/0
4	CFM	C	497	1	-	0/0/84/84	0/0/8/8
5	CLP	C	499	1,2	-	0/0/72/117	0/9/9/10

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	496	CFM	5	0
5	A	498	CLP	3	0
4	C	497	CFM	6	0
5	C	499	CLP	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

### 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.