



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 7, 2018 – 01:22 pm GMT

PDB ID : 4WZ7  
Title : Crystal structure of mitochondrial NADH:ubiquinone oxidoreductase from Yarrowia lipolytica.  
Authors : Wirth, C.; Zickermann, V.; Brandt, U.; Hunte, C.  
Deposited on : 2014-11-18  
Resolution : 3.60 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) : 1.13  
EDS : trunk30967  
Percentile statistics : (not set)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

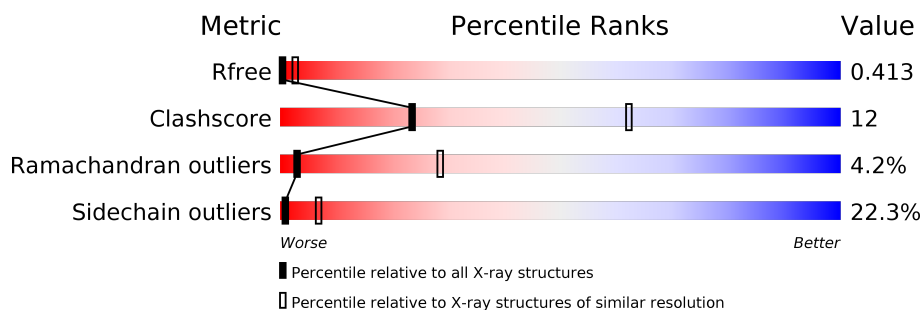
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.60 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.





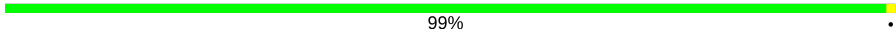

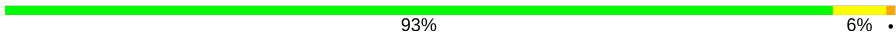


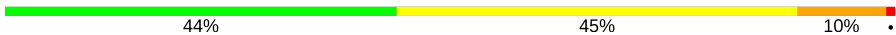


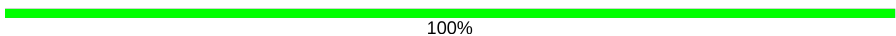
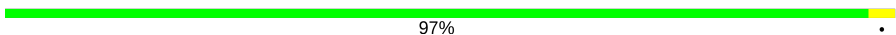



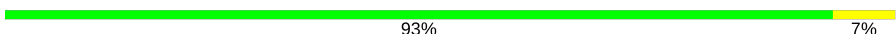
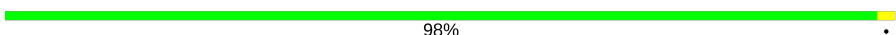
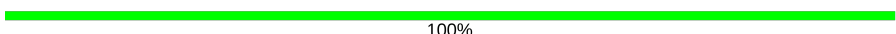

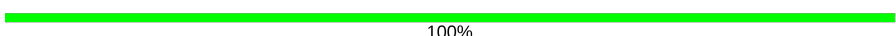

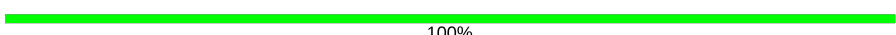

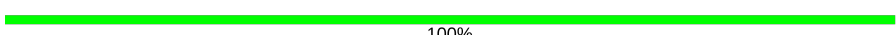

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	111664	1053 (3.70-3.50)
Clashscore	122126	1141 (3.70-3.50)
Ramachandran outliers	120053	1102 (3.70-3.50)
Sidechain outliers	120020	1102 (3.70-3.50)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	1	327	
2	2	438	
3	3	89	
4	4	470	
5	5	619	
6	6	185	
7	A	628	


























*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
8	B	370	
9	C	444	
10	E	195	
11	G	133	
12	H	154	
13	I	137	
14	K	183	
15	L	89	
16	D	57	
16	Z	57	
17	F	54	
18	J	63	
19	M	29	
20	N	50	
21	O	70	
22	P	46	
23	Q	51	
24	R	30	
25	S	69	
26	AH	15	
26	T	15	
27	U	26	
28	V	22	
29	AB	9	
29	AY	9	

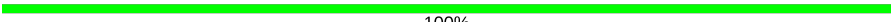
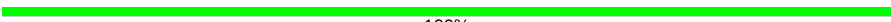



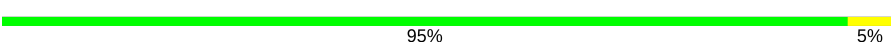

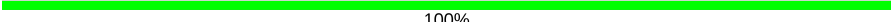


*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
29	BE	9	 100%
29	W	9	 78% 22%
30	AJ	16	 100%
30	AV	16	 100%
30	BH	16	 88% 13%
30	X	16	 100%
31	AR	13	 100%
31	AT	13	 100%
31	Y	13	 100%
32	AA	18	 100%
32	AW	18	 100%
32	BB	18	 100%
32	BG	18	 100%
33	AC	47	 85% 15%
33	AD	47	 100%
34	AE	48	 92% 8%
35	AF	35	 100%
36	AG	25	 100%
37	AI	36	 100%
37	AL	36	 100%
38	AK	76	 100%
38	AN	76	 97% .
39	AM	17	 100%
40	AO	32	 100%
41	AP	11	 100%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
41	AS	11	 100%
42	AQ	8	 100%
42	BA	8	 100%
43	AU	58	 100%
44	AX	39	 100%
45	AZ	40	 95% 5%
46	BC	20	 90% 10%
47	BD	19	 100%
47	BF	19	 100%
48	BI	905	 99%

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
50	SF4	K	500	-	-	X	-

## 2 Entry composition [i](#)

There are 50 unique types of molecules in this entry. The entry contains 35169 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 NADH-ubiquinone oxidoreductase chain 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	1	318	Total	C	N	O	S	0	0	0
			2177	1431	352	389	5			

- Molecule 2 is a protein called NADH dehydrogenase subunit 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	2	438	Total	C	N	O	S	0	0	0
			3142	2092	482	556	12			

- Molecule 3 is a protein called NADH-ubiquinone oxidoreductase chain 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	3	89	Total	C	N	O	S	0	0	0
			641	444	91	104	2			

- Molecule 4 is a protein called NADH-ubiquinone oxidoreductase chain 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	4	470	Total	C	N	O	S	0	0	0
			3017	1952	507	546	12			

- Molecule 5 is a protein called NADH-ubiquinone oxidoreductase chain 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	5	619	Total	C	N	O	S	0	0	0
			4065	2636	677	727	25			

- Molecule 6 is a protein called NADH-ubiquinone oxidoreductase chain 6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	6	149	Total	C	N	O	S	0	0	0
			1078	735	156	180	7			

- Molecule 7 is a protein called NUAM protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	A	628	Total	C	N	O	S	0	0	0
			3226	1933	639	639	15			

- Molecule 8 is a protein called NUBM protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	B	370	Total	C	N	O	S	0	0	0
			2013	1214	391	399	9			

- Molecule 9 is a protein called NUCM protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	C	366	Total	C	N	O	S	0	0	0
			2647	1675	455	496	21			

- Molecule 10 is a protein called 39-kDa subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	E	195	Total	C	N	O	S	0	0	0
			975	585	195	195				

- Molecule 11 is a protein called NUGM protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	G	133	Total	C	N	O	S	0	0	0
			880	558	154	164	4			

- Molecule 12 is a protein called Subunit NUHM of protein NADH:Ubiquinone Oxidoreductase (Complex I).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	H	154	Total	C	N	O	S	0	0	0
			803	476	156	164	7			

- Molecule 13 is a protein called Subunit NUIM of protein NADH:Ubiquinone Oxidoreductase (Complex I).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	I	137	Total	C	N	O	S	0	0	0
			857	533	145	169	10			

- Molecule 14 is a protein called Subunit NUKM of protein NADH:Ubiquinone Oxidoreductase (Complex I).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	K	147	Total	C	N	O	S	0	0	0
			1023	642	182	187	12			

- Molecule 15 is a protein called NADH-ubiquinone oxidoreductase chain 4L.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	L	89	Total	C	N	O	S	0	0	0
			660	437	108	112	3			

- Molecule 16 is a protein called unknown subunits 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
16	Z	57	Total	C	N	O	0	0	0
			285	171	57	57			
16	D	57	Total	C	N	O	0	0	0
			285	171	57	57			

- Molecule 17 is a protein called unknown subunits 2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
17	F	54	Total	C	N	O	0	0	0
			270	162	54	54			

- Molecule 18 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
18	J	63	Total	C	N	O	0	0	0
			315	189	63	63			

- Molecule 19 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
19	M	29	Total	C	N	O	0	0	0
			145	87	29	29			

- Molecule 20 is a protein called unknown subunits.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
20	N	50	Total	C	N	O	0	0	0
			250	150	50	50			

- Molecule 21 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
21	O	70	Total	C	N	O	0	0	0
			350	210	70	70			

- Molecule 22 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
22	P	46	Total	C	N	O	0	0	0
			230	138	46	46			

- Molecule 23 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
23	Q	51	Total	C	N	O	0	0	0
			255	153	51	51			

- Molecule 24 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
24	R	30	Total	C	N	O	0	0	0
			150	90	30	30			

- Molecule 25 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
25	S	69	Total	C	N	O	0	0	0
			345	207	69	69			

- Molecule 26 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
26	T	15	Total	C	N	O	0	0	0
			75	45	15	15			
26	AH	15	Total	C	N	O	0	0	0
			75	45	15	15			

- Molecule 27 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
27	U	26	Total	C	N	O	0	0	0
			130	78	26	26			

- Molecule 28 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
28	V	22	Total	C	N	O	0	0	0
			110	66	22	22			

- Molecule 29 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
29	W	9	Total	C	N	O	0	0	0
			45	27	9	9			
29	AB	9	Total	C	N	O	0	0	0
			45	27	9	9			
29	AY	9	Total	C	N	O	0	0	0
			45	27	9	9			
29	BE	9	Total	C	N	O	0	0	0
			45	27	9	9			

- Molecule 30 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
30	X	16	Total	C	N	O	0	0	0
			80	48	16	16			
30	AJ	16	Total	C	N	O	0	0	0
			80	48	16	16			
30	AV	16	Total	C	N	O	0	0	0
			80	48	16	16			
30	BH	16	Total	C	N	O	0	0	0
			80	48	16	16			

- Molecule 31 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
31	Y	13	Total	C	N	O	0	0	0
			65	39	13	13			
31	AR	13	Total	C	N	O	0	0	0
			65	39	13	13			
31	AT	13	Total	C	N	O	0	0	0
			65	39	13	13			

- Molecule 32 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
32	AA	18	Total	C	N	O	0	0	0
			90	54	18	18			
32	AW	18	Total	C	N	O	0	0	0
			90	54	18	18			
32	BB	18	Total	C	N	O	0	0	0
			90	54	18	18			
32	BG	18	Total	C	N	O	0	0	0
			90	54	18	18			

- Molecule 33 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
33	AC	47	Total	C	N	O	0	0	0
			235	141	47	47			
33	AD	47	Total	C	N	O	0	0	0
			235	141	47	47			

- Molecule 34 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
34	AE	48	Total	C	N	O	0	0	0
			240	144	48	48			

- Molecule 35 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
35	AF	35	Total	C	N	O	0	0	0
			175	105	35	35			

- Molecule 36 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
36	AG	25	Total	C	N	O	0	0	0
			125	75	25	25			

- Molecule 37 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
37	AI	36	Total	C	N	O	0	0	0
			180	108	36	36			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
37	AL	36	Total	C	N	O	0	0	0
			180	108	36	36			

- Molecule 38 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
38	AK	76	Total	C	N	O	0	0	0
			380	228	76	76			
38	AN	76	Total	C	N	O	0	0	0
			380	228	76	76			

- Molecule 39 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
39	AM	17	Total	C	N	O	0	0	0
			85	51	17	17			

- Molecule 40 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
40	AO	32	Total	C	N	O	0	0	0
			160	96	32	32			

- Molecule 41 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
41	AP	11	Total	C	N	O	0	0	0
			55	33	11	11			
41	AS	11	Total	C	N	O	0	0	0
			54	32	11	11			

- Molecule 42 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
42	AQ	8	Total	C	N	O	0	0	0
			40	24	8	8			
42	BA	8	Total	C	N	O	0	0	0
			40	24	8	8			

- Molecule 43 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
43	AU	58	Total	C	N	O	0	0	0
			290	174	58	58			

- Molecule 44 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
44	AX	39	Total	C	N	O	0	0	0
			195	117	39	39			

- Molecule 45 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
45	AZ	40	Total	C	N	O	0	0	0
			200	120	40	40			

- Molecule 46 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
46	BC	20	Total	C	N	O	0	0	0
			100	60	20	20			

- Molecule 47 is a protein called unknown subunits.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
47	BD	19	Total	C	N	O	0	0	0
			95	57	19	19			
47	BF	19	Total	C	N	O	0	0	0
			95	57	19	19			

- Molecule 48 is a protein called unknown subunits.

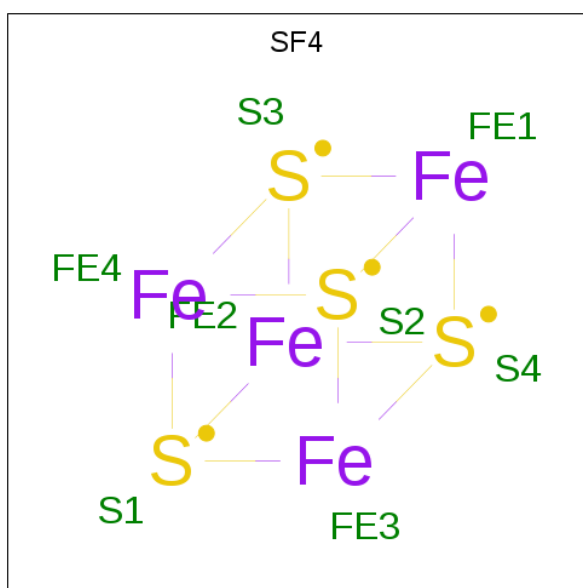
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
48	BI	9	Total	C	N	O	0	0	0
			45	27	9	9			

- Molecule 49 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula: Fe<sub>2</sub>S<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
49	A	1	Total	Fe	S	0	0
			4	2	2		
49	H	1	Total	Fe	S	0	0
			4	2	2		

- Molecule 50 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula:  $\text{Fe}_4\text{S}_4$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
50	A	1	Total	Fe	S	0	0
			8	4	4		
50	A	1	Total	Fe	S	0	0
			8	4	4		

*Continued on next page...*

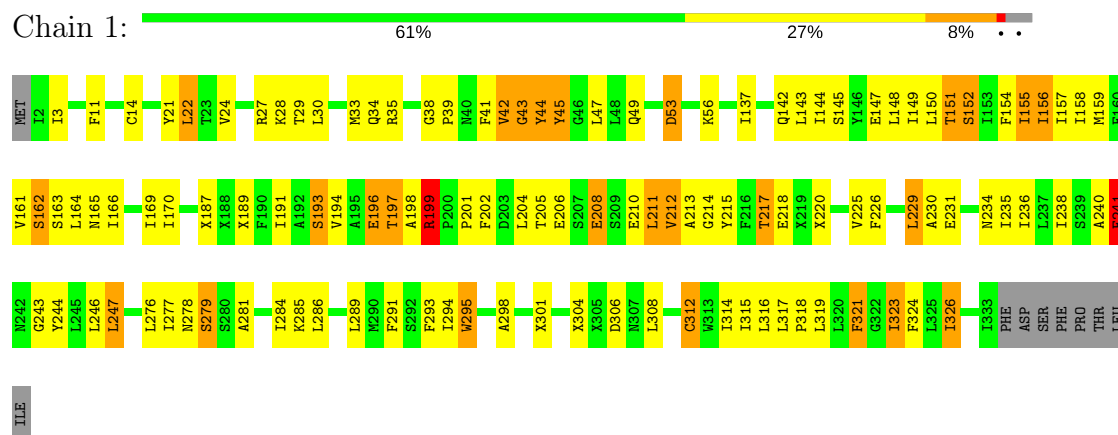
*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
50	B	1	Total 8	Fe 4	S 4	0	0
50	I	1	Total 8	Fe 4	S 4	0	0
50	I	1	Total 8	Fe 4	S 4	0	0
50	K	1	Total 8	Fe 4	S 4	0	0

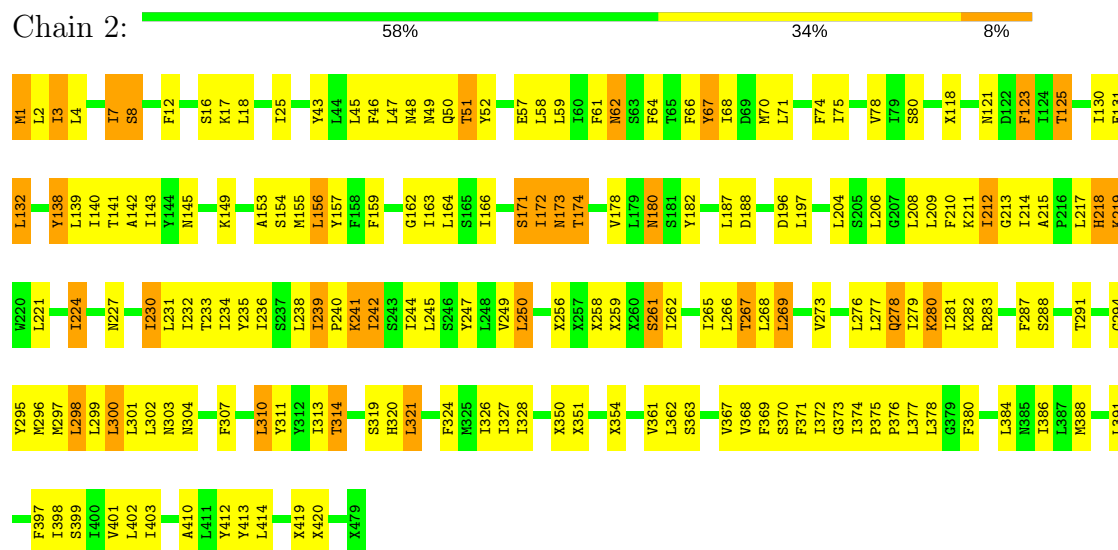
### 3 Residue-property plots

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. 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.

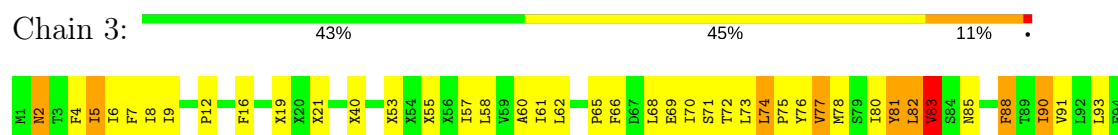
#### • Molecule 1: NADH-ubiquinone oxidoreductase chain 1



#### • Molecule 2: NADH dehydrogenase subunit 2

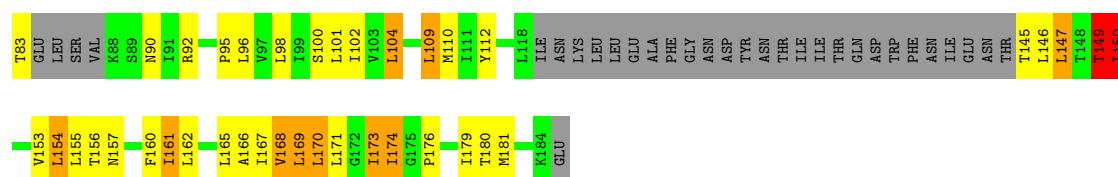


#### • Molecule 3: NADH-ubiquinone oxidoreductase chain 3



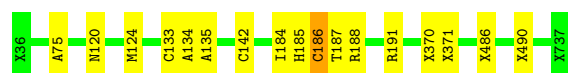






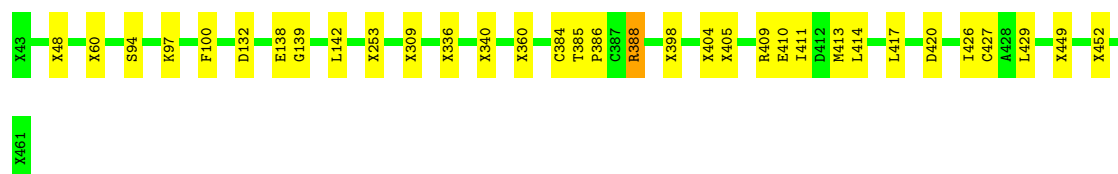
- Molecule 7: NUAM protein

Chain A: 97%



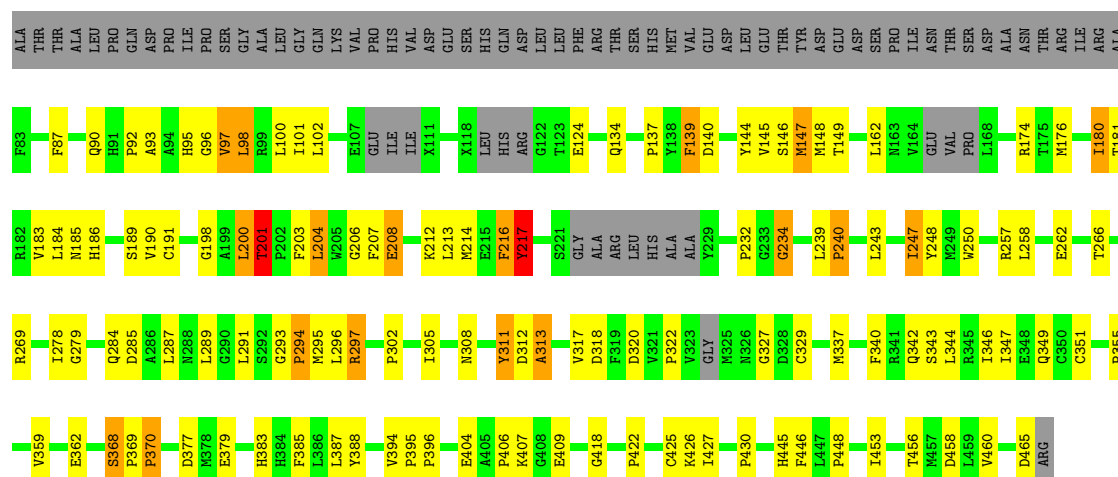
- Molecule 8: NUBM protein

Chain B: 91% 9%



- Molecule 9: NUCM protein

Chain C: 54% 24% 18%




- Molecule 10: 39-kDa subunit

Chain E: 99%



- Molecule 11: NUGM protein

Chain G:  80% 20%



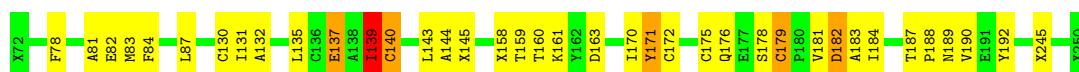
- Molecule 12: Subunit NUHM of protein NADH:Ubiquinone Oxidoreductase (Complex I)

Chain H:  93% 6%



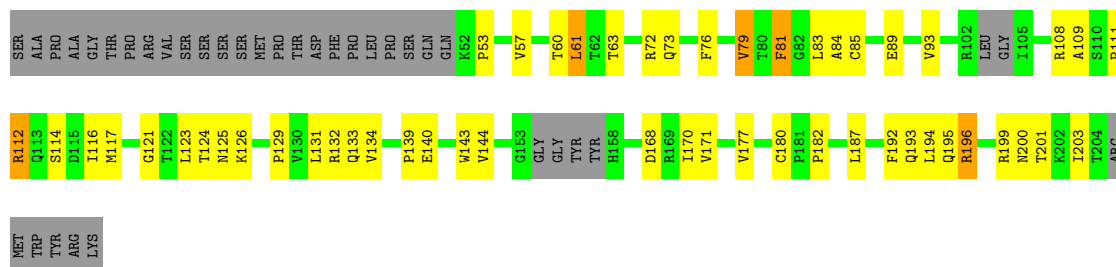
- Molecule 13: Subunit NUIM of protein NADH:Ubiquinone Oxidoreductase (Complex I)

Chain I:  72% 23%



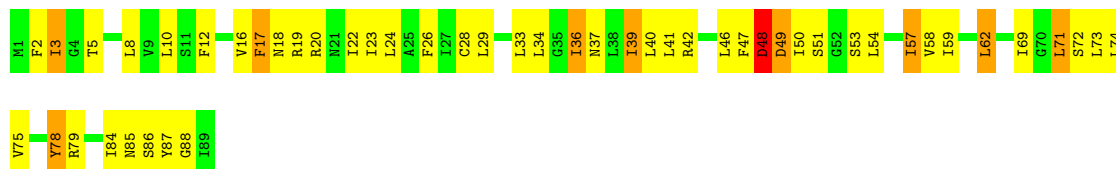
- Molecule 14: Subunit NUKM of protein NADH:Ubiquinone Oxidoreductase (Complex I)

Chain K:  52% 26% 20%




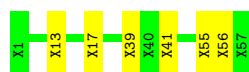
- Molecule 15: NADH-ubiquinone oxidoreductase chain 4L

Chain L:  44% 45% 10%




- Molecule 16: unknown subunits 1

Chain Z:  89% 11%



- Molecule 16: unknown subunits 1

Chain D:  84% 16%



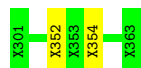
- Molecule 17: unknown subunits 2

Chain F:  100%

There are no outlier residues recorded for this chain.

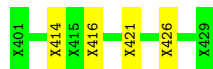
- Molecule 18: unknown subunits

Chain J:  97%



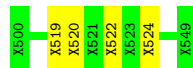
- Molecule 19: unknown subunits

Chain M:  86%



- Molecule 20: unknown subunits

Chain N:  92%



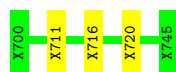
- Molecule 21: unknown subunits

Chain O:  94%



- Molecule 22: unknown subunits

Chain P:  93%



- Molecule 23: unknown subunits

Chain Q:  98%



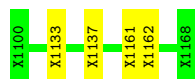
- Molecule 24: unknown subunits

Chain R:  100%

There are no outlier residues recorded for this chain.

- Molecule 25: unknown subunits

Chain S:  94% 6%



- Molecule 26: unknown subunits

Chain T:  100%

There are no outlier residues recorded for this chain.

- Molecule 26: unknown subunits

Chain AH:  100%


There are no outlier residues recorded for this chain.

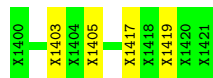
- Molecule 27: unknown subunits

Chain U:  100%


There are no outlier residues recorded for this chain.

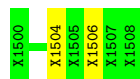
- Molecule 28: unknown subunits

Chain V:  82% 18%



- Molecule 29: unknown subunits

Chain W:  78% 22%



- Molecule 29: unknown subunits

Chain AB:  100%

There are no outlier residues recorded for this chain.

- Molecule 29: unknown subunits

Chain AY:  100%

There are no outlier residues recorded for this chain.

- Molecule 29: unknown subunits

Chain BE:  100%

There are no outlier residues recorded for this chain.

- Molecule 30: unknown subunits

Chain X:  100%

There are no outlier residues recorded for this chain.

- Molecule 30: unknown subunits

Chain AJ:  100%


There are no outlier residues recorded for this chain.

- Molecule 30: unknown subunits

Chain AV:  100%

There are no outlier residues recorded for this chain.

- Molecule 30: unknown subunits

Chain BH:  88% 13%



- Molecule 31: unknown subunits

Chain Y:  100%

There are no outlier residues recorded for this chain.

- Molecule 31: unknown subunits

Chain AR:  100%

There are no outlier residues recorded for this chain.

- Molecule 31: unknown subunits

Chain AT:  100%

There are no outlier residues recorded for this chain.

- Molecule 32: unknown subunits

Chain AA:  100%

There are no outlier residues recorded for this chain.

- Molecule 32: unknown subunits

Chain AW:  100%

There are no outlier residues recorded for this chain.

- Molecule 32: unknown subunits

Chain BB:  100%


There are no outlier residues recorded for this chain.

- Molecule 32: unknown subunits

Chain BG:  100%

There are no outlier residues recorded for this chain.

- Molecule 33: unknown subunits

Chain AC:  85% 15%



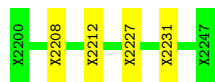
- Molecule 33: unknown subunits

Chain AD:  100%

There are no outlier residues recorded for this chain.

- Molecule 34: unknown subunits

Chain AE:  92% 8%



- Molecule 35: unknown subunits

Chain AF:  100%

There are no outlier residues recorded for this chain.

- Molecule 36: unknown subunits

Chain AG:  100%

There are no outlier residues recorded for this chain.

- Molecule 37: unknown subunits

Chain AI:  100%

There are no outlier residues recorded for this chain.

- Molecule 37: unknown subunits

Chain AL:  100%

There are no outlier residues recorded for this chain.

- Molecule 38: unknown subunits

Chain AK:  100%

There are no outlier residues recorded for this chain.

- Molecule 38: unknown subunits

Chain AN:  97%



- Molecule 39: unknown subunits

Chain AM:  100%

There are no outlier residues recorded for this chain.

- Molecule 40: unknown subunits

Chain AO:  100%

There are no outlier residues recorded for this chain.

- Molecule 41: unknown subunits

Chain AP:  100%

There are no outlier residues recorded for this chain.

- Molecule 41: unknown subunits

Chain AS:  100%

There are no outlier residues recorded for this chain.

- Molecule 42: unknown subunits

Chain AQ:  100%

There are no outlier residues recorded for this chain.

- Molecule 42: unknown subunits



Chain BA:  100%

There are no outlier residues recorded for this chain.

- Molecule 43: unknown subunits

Chain AU:  100%

There are no outlier residues recorded for this chain.

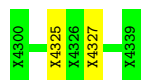
- Molecule 44: unknown subunits

Chain AX:  100%

There are no outlier residues recorded for this chain.

- Molecule 45: unknown subunits

Chain AZ:  95% 5%



- Molecule 46: unknown subunits

Chain BC:  90% 10%



- Molecule 47: unknown subunits

Chain BD:  100%

There are no outlier residues recorded for this chain.

- Molecule 47: unknown subunits

Chain BF:  100%

There are no outlier residues recorded for this chain.

- Molecule 48: unknown subunits

Chain BI:  99%



[illegible]

## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	317.74Å 317.74Å 818.97Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	25.00 – 3.60 50.00 – 3.60	Depositor EDS
% Data completeness (in resolution range)	84.8 (25.00-3.60) 85.0 (50.00-3.60)	Depositor EDS
$R_{merge}$	0.52	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.01 (at 3.57Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, $R_{free}$	0.316 , 0.341 0.379 , 0.413	Depositor DCC
$R_{free}$ test set	2191 reflections (1.41%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	106.8	Xtriage
Anisotropy	0.020	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 999.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.41$ , $\langle L^2 \rangle = 0.24$	Xtriage
Estimated twinning fraction	0.039 for $-1/3^*h+1/3^*k+1/3^*l,-k,8/3^*h+4/3^*k+1/3^*l$ 0.097 for $-2/3^*h-1/3^*k-1/3^*l,-1/3^*h-2/3^*k+1/3^*l,-4/3^*h+4/3^*k+1/3^*l$ 0.067 for $-h,1/3^*h-1/3^*k-1/3^*l,-4/3^*h-8/3^*k+1/3^*l$	Xtriage
$F_o, F_c$ correlation	0.77	EDS
Total number of atoms	35169	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	109.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.62% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SF4, FES

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	1	0.49	0/1638	0.88	3/2224 (0.1%)
2	2	0.56	0/2613	0.89	1/3550 (0.0%)
3	3	0.52	0/533	0.94	1/728 (0.1%)
4	4	0.54	0/1920	0.86	0/2615
5	5	0.51	0/2813	0.82	0/3820
6	6	0.54	0/1090	0.87	0/1491
7	A	0.48	0/412	0.79	0/531
8	B	0.46	0/527	0.73	0/701
9	C	0.46	0/2653	0.82	7/3592 (0.2%)
11	G	0.39	0/566	0.67	0/766
12	H	0.44	0/178	0.77	0/220
13	I	0.41	0/499	0.79	0/675
14	K	0.47	0/1042	0.80	2/1424 (0.1%)
15	L	0.58	0/666	0.91	0/902
All	All	0.50	0/17150	0.84	14/23239 (0.1%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	3	82	LEU	C-N-CA	7.67	140.88	121.70
9	C	240	PRO	N-CA-CB	6.37	110.94	103.30
9	C	395	PRO	N-CA-CB	5.93	110.42	103.30
9	C	396	PRO	N-CA-CB	5.80	110.26	103.30
1	1	43	GLY	C-N-CA	5.68	135.90	121.70

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	1	2177	0	1809	76	0
2	2	3142	0	2868	136	0
3	3	641	0	608	45	0
4	4	3017	0	2276	107	0
5	5	4065	0	3176	91	0
6	6	1078	0	1125	55	0
7	A	3226	0	1029	10	0
8	B	2013	0	839	22	0
9	C	2647	0	2375	71	0
10	E	975	0	229	1	0
11	G	880	0	581	21	0
12	H	803	0	302	7	0
13	I	857	0	563	31	0
14	K	1023	0	899	24	0
15	L	660	0	701	34	0
16	D	285	0	64	5	0
16	Z	285	0	63	3	0
17	F	270	0	60	0	0
18	J	315	0	71	1	0
19	M	145	0	36	2	0
20	N	250	0	53	2	0
21	O	350	0	75	2	0
22	P	230	0	52	2	0
23	Q	255	0	55	2	0
24	R	150	0	32	0	0
25	S	345	0	77	2	0
26	AH	75	0	17	0	0
26	T	75	0	17	0	0
27	U	130	0	29	0	0
28	V	110	0	24	2	0
29	AB	45	0	14	0	0
29	AY	45	0	11	0	0
29	BE	45	0	12	0	0
29	W	45	0	11	1	0
30	AJ	80	0	19	0	0
30	AV	80	0	19	0	0
30	BH	80	0	18	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
30	X	80	0	19	0	0
31	AR	65	0	16	0	0
31	AT	65	0	15	0	0
31	Y	65	0	16	0	0
32	AA	90	0	20	0	0
32	AW	90	0	24	0	0
32	BB	90	0	21	0	0
32	BG	90	0	22	0	0
33	AC	235	0	49	4	0
33	AD	235	0	52	0	0
34	AE	240	0	54	2	0
35	AF	175	0	37	0	0
36	AG	125	0	29	0	0
37	AI	180	0	39	0	0
37	AL	180	0	44	0	0
38	AK	380	0	79	0	0
38	AN	380	0	84	1	0
39	AM	85	0	20	0	0
40	AO	160	0	34	0	0
41	AP	55	0	13	0	0
41	AS	54	0	13	0	0
42	AQ	40	0	10	0	0
42	BA	40	0	10	0	0
43	AU	290	0	64	0	0
44	AX	195	0	44	0	0
45	AZ	200	0	48	1	0
46	BC	100	0	22	1	0
47	BD	95	0	22	0	0
47	BF	95	0	21	0	0
48	BI	45	0	11	0	0
49	A	4	0	0	0	0
49	H	4	0	0	1	0
50	A	16	0	0	1	0
50	B	8	0	0	0	0
50	I	16	0	0	2	0
50	K	8	0	0	2	0
All	All	35169	0	21161	682	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 12.

The worst 5 of 682 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)
5:5:585:TYR:O	5:5:589:PRO:HD2	1.36	1.25
3:3:98:LEU:HD11	6:6:170:LEU:HA	1.27	1.15
13:I:175:CYS:HB2	13:I:184:ILE:HD13	1.17	1.09
12:H:217:UNK:CB	12:H:220:UNK:O	2.01	1.08
6:6:28:ASN:HB3	6:6:29:PRO:HD3	1.34	1.06

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	1	202/327 (62%)	166 (82%)	30 (15%)	6 (3%)	5	37
2	2	320/438 (73%)	278 (87%)	31 (10%)	11 (3%)	4	35
3	3	64/89 (72%)	54 (84%)	7 (11%)	3 (5%)	2	26
4	4	243/470 (52%)	212 (87%)	24 (10%)	7 (3%)	5	38
5	5	356/619 (58%)	295 (83%)	46 (13%)	15 (4%)	3	28
6	6	141/185 (76%)	118 (84%)	14 (10%)	9 (6%)	1	19
7	A	63/628 (10%)	55 (87%)	5 (8%)	3 (5%)	2	25
8	B	72/370 (20%)	67 (93%)	5 (7%)	0	100	100
9	C	348/444 (78%)	303 (87%)	27 (8%)	18 (5%)	2	23
11	G	65/133 (49%)	55 (85%)	8 (12%)	2 (3%)	4	37
12	H	30/154 (20%)	28 (93%)	1 (3%)	1 (3%)	4	35
13	I	64/137 (47%)	47 (73%)	11 (17%)	6 (9%)	1	10
14	K	141/183 (77%)	126 (89%)	12 (8%)	3 (2%)	8	45
15	L	87/89 (98%)	72 (83%)	6 (7%)	9 (10%)	0	9
All	All	2196/4266 (52%)	1876 (85%)	227 (10%)	93 (4%)	3	28

5 of 93 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	1	42	VAL
1	1	44	TYR
1	1	45	TYR
1	1	199	ARG
1	1	212	VAL

### 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	1	175/186 (94%)	129 (74%)	46 (26%)	0	4
2	2	291/291 (100%)	220 (76%)	71 (24%)	1	5
3	3	60/60 (100%)	42 (70%)	18 (30%)	0	3
4	4	215/215 (100%)	163 (76%)	52 (24%)	1	5
5	5	301/305 (99%)	225 (75%)	76 (25%)	0	5
6	6	110/167 (66%)	73 (66%)	37 (34%)	0	2
7	A	39/40 (98%)	36 (92%)	3 (8%)	14	49
8	B	53/56 (95%)	50 (94%)	3 (6%)	23	61
9	C	238/371 (64%)	203 (85%)	35 (15%)	3	22
11	G	55/55 (100%)	53 (96%)	2 (4%)	38	72
12	H	19/19 (100%)	18 (95%)	1 (5%)	25	63
13	I	55/55 (100%)	44 (80%)	11 (20%)	1	9
14	K	91/157 (58%)	74 (81%)	17 (19%)	1	11
15	L	69/77 (90%)	46 (67%)	23 (33%)	0	2
All	All	1771/2054 (86%)	1376 (78%)	395 (22%)	1	7

5 of 395 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	4	405	ARG
5	5	234	ASN
14	K	140	GLU

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
4	4	461	THR
5	5	137	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
5	5	162	GLN
7	A	185	HIS
14	K	113	GLN
5	5	113	GLN
14	K	133	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](#)

8 ligands are modelled in this entry.

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	$\# Z  > 2$
49	FES	A	900	7	0,4,4	0.00	-	0,4,4	0.00	-
50	SF4	A	901	7	0,12,12	0.00	-	0,24,24	0.00	-

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
50	SF4	A	902	7	0,12,12	0.00	-	0,24,24	0.00	-
50	SF4	B	500	8	0,12,12	0.00	-	0,24,24	0.00	-
49	FES	H	300	12	0,4,4	0.00	-	0,4,4	0.00	-
50	SF4	I	500	13	0,12,12	0.00	-	0,24,24	0.00	-
50	SF4	I	501	13	0,12,12	0.00	-	0,24,24	0.00	-
50	SF4	K	500	14	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
49	FES	A	900	7	-	0/0/4/4	0/1/1/1
50	SF4	A	901	7	-	0/0/48/48	0/6/5/5
50	SF4	A	902	7	-	0/0/48/48	0/6/5/5
50	SF4	B	500	8	-	0/0/48/48	0/6/5/5
49	FES	H	300	12	-	0/0/4/4	0/1/1/1
50	SF4	I	500	13	-	0/0/48/48	0/6/5/5
50	SF4	I	501	13	-	0/0/48/48	0/6/5/5
50	SF4	K	500	14	-	0/0/48/48	0/6/5/5

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.

5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
50	A	902	SF4	1	0
49	H	300	FES	1	0
50	I	500	SF4	1	0
50	I	501	SF4	1	0
50	K	500	SF4	2	0

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
7	A	30
8	B	11
10	E	10
5	5	8
12	H	5
13	I	5
4	4	4
2	2	4
1	1	3
11	G	2
3	3	2
22	P	1
16	Z	1
21	O	1
37	AL	1

The worst 5 of 88 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	E	422:UNK	C	499:UNK	N	48.62
1	B	270:UNK	C	279:UNK	N	26.71
1	E	513:UNK	C	604:UNK	N	26.57
1	A	148:UNK	C	152:UNK	N	24.97
1	I	92:UNK	C	103:UNK	N	24.67

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.