wwPDB X-ray Structure Validation Summary Report (i)

Jan 7, 2024 – 11:07 am GMT

PDB ID	:	6H8K
Title	:	Crystal structure of a variant (Q133C in PSST) of Yarrowia lipolytica complex I
Authors	:	Wirth, C.; Galemou Yoga, E.; Zickermann, V.; Hunte, C.
Deposited on	:	2018-08-02
Resolution	:	3.79 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:

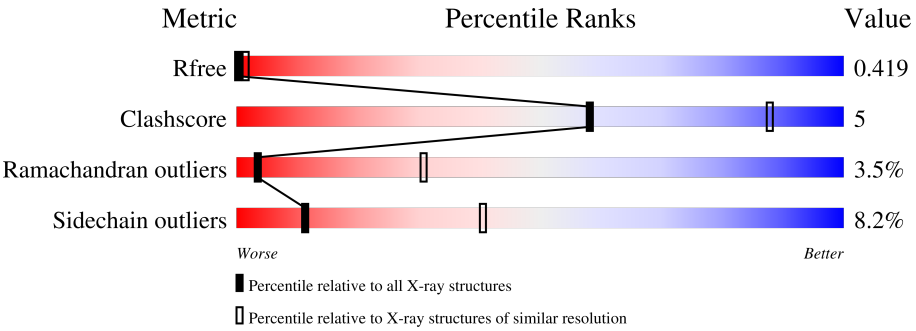
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 3.79 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.



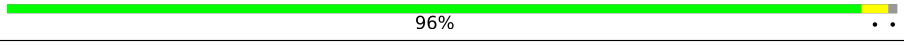
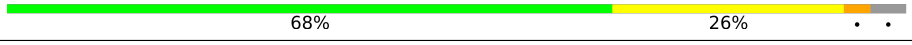
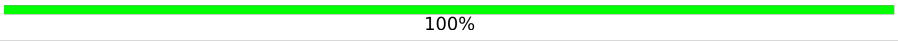
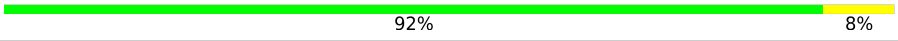
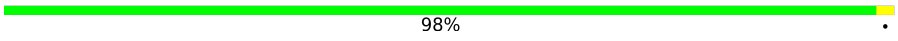
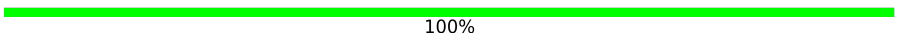
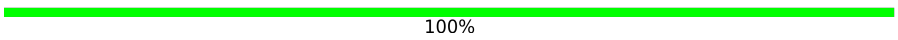
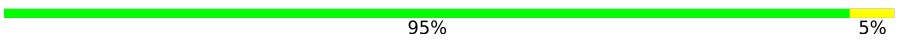
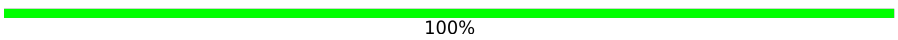
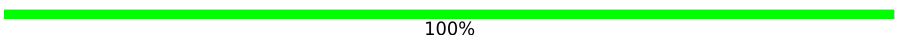
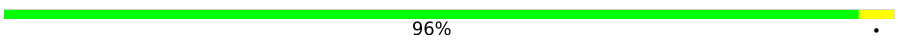
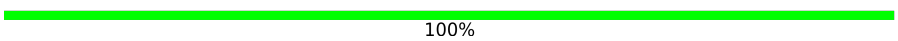
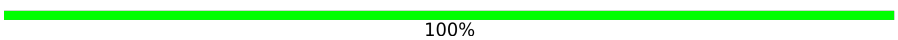
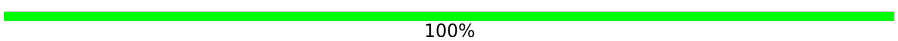
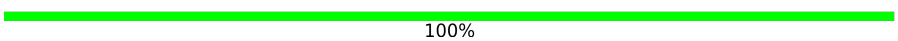
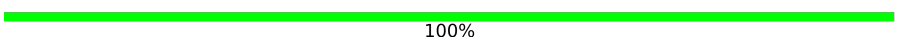
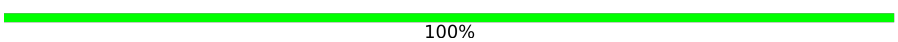
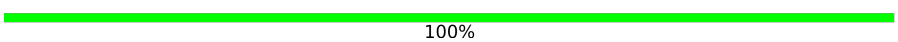
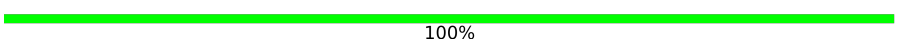
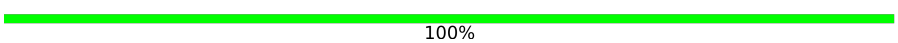
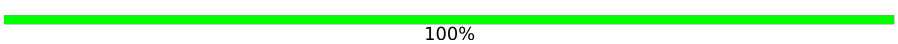
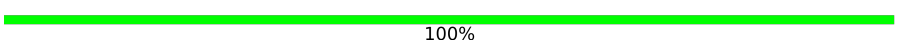
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1212 (4.00-3.60)
Clashscore	141614	1288 (4.00-3.60)
Ramachandran outliers	138981	1243 (4.00-3.60)
Sidechain outliers	138945	1237 (4.00-3.60)

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 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%

Mol	Chain	Length	Quality of chain
1	1	335	66% 27%
2	2	434	79% 19%
3	3	110	60% 17% 20%
4	4	470	83% 16%
5	5	616	86% 13%
6	6	184	46% 14% 36%
7	A	634	98%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	B	370	
9	C	383	
10	E	195	
11	G	133	
12	H	154	
13	I	140	
14	K	147	
15	L	79	
16	Z	17	
16	r	17	
17	AF	40	
17	Y	40	
18	U	10	
19	X	57	
20	W	54	
21	V	63	
22	AI	20	
22	T	20	
22	m	20	
23	AJ	19	
23	AL	19	
23	S	19	
24	R	50	
25	Q	70	
26	P	28	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
27	AC	18	100%
27	AH	18	100%
27	AM	18	100%
27	F	18	100%
27	f	18	100%
28	O	25	100%
28	l	25	100%
29	M	51	100%
30	D	30	100%
31	J	69	100%
32	N	15	100%
33	a	26	100%
33	i	26	100%
34	b	22	100%
35	AE	9	100%
35	AK	9	100%
35	AO	9	100%
35	c	9	100%
35	g	9	100%
36	AB	16	100%
36	AN	16	100%
36	d	16	100%
36	o	16	100%
37	e	13	100%
37	w	13	100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
37	z	13	100%
38	h	47	100%
39	j	48	100%
40	k	23	100%
40	s	23	100%
41	n	36	100%
41	q	36	100%
42	p	76	100%
43	t	45	100%
44	u	32	100%
45	v	11	100%
45	y	11	100%
46	AG	8	100%
46	x	8	100%
47	AA	58	100%
48	AD	39	100%

2 Entry composition (i)

There are 50 unique types of molecules in this entry. The entry contains 35631 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	325	Total	C	N	O	S	0	0	0
			2581	1758	375	441	7			

- Molecule 2 is a protein called NADH-ubiquinone oxidoreductase chain 2,NADH dehydrogenase subunit 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	2	434	Total	C	N	O	S	0	0	0
			3122	2080	478	552	12			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	3	88	Total	C	N	O	S	0	0	0
			685	475	94	113	3			

- 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	616	Total	C	N	O	S	0	0	0
			4050	2627	674	724	25			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	6	118	Total	C	N	O	S	0	0	0
			870	597	124	143	6			

- 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
			3221	1929	639	639	14			

- Molecule 8 is a protein called NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	B	367	Total	C	N	O	S	0	0	0
			1997	1204	388	396	9			

- Molecule 9 is a protein called NUCM protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	C	369	Total	C	N	O	S	0	0	0
			2882	1845	486	530	21			

- Molecule 10 is a protein called NUEM protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
10	E	195	Total	C	N	O	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 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 NADH:Ubiquinone Oxidoreductase (Com-

plex 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	143	Total	C	N	O	S	5	0	0
			1069	675	187	193	14			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	133	CYS	GLN	engineered mutation	UNP Q9UUT7

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	L	79	Total	C	N	O	S	6	0	0
			612	412	95	102	3			

- Molecule 16 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
16	Z	17	Total	C	N	O	0	0	0
			85	51	17	17			
16	r	17	Total	C	N	O	0	0	0
			85	51	17	17			

- Molecule 17 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
17	Y	40	Total	C	N	O	0	0	0
			200	120	40	40			
17	AF	40	Total	C	N	O	0	0	0
			200	120	40	40			

- Molecule 18 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
18	U	10	Total 50	C 30	N 10	O 10	0	0	0

- Molecule 19 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
19	X	57	Total 285	C 171	N 57	O 57	0	0	0

- Molecule 20 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
20	W	54	Total 270	C 162	N 54	O 54	0	0	0

- Molecule 21 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
21	V	63	Total 315	C 189	N 63	O 63	0	0	0

- Molecule 22 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
22	T	20	Total 100	C 60	N 20	O 20	0	0	0
22	m	20	Total 100	C 60	N 20	O 20	0	0	0
22	AI	20	Total 100	C 60	N 20	O 20	0	0	0

- Molecule 23 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
23	S	19	Total 95	C 57	N 19	O 19	0	0	0
23	AJ	19	Total 95	C 57	N 19	O 19	0	0	0
23	AL	19	Total 95	C 57	N 19	O 19	0	0	0

- Molecule 24 is a protein called Unknown polypeptide.

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

- Molecule 25 is a protein called Unknown polypeptide.

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

- Molecule 26 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
26	P	28	Total 140	C 84	N 28	O 28	0	0	0

- Molecule 27 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
27	F	18	Total 90	C 54	N 18	O 18	0	0	0
27	f	18	Total 90	C 54	N 18	O 18	0	0	0
27	AC	18	Total 90	C 54	N 18	O 18	0	0	0
27	AH	18	Total 90	C 54	N 18	O 18	0	0	0
27	AM	18	Total 90	C 54	N 18	O 18	0	0	0

- Molecule 28 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
28	O	25	Total 125	C 75	N 25	O 25	0	0	0
28	l	25	Total 125	C 75	N 25	O 25	0	0	0

- Molecule 29 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
29	M	51	Total 255	C 153	N 51	O 51	0	0	0

- Molecule 30 is a protein called Unknown polypeptide.

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

- Molecule 31 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
31	J	69	Total C N O 345 207 69 69	0	0	0

- Molecule 32 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
32	N	15	Total C N O 75 45 15 15	0	0	0

- Molecule 33 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
33	a	26	Total C N O 130 78 26 26	0	0	0
33	i	26	Total C N O 130 78 26 26	0	0	0

- Molecule 34 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
34	b	22	Total C N O 110 66 22 22	0	0	0

- Molecule 35 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
35	c	9	Total C N O 45 27 9 9	0	0	0
35	g	9	Total C N O 45 27 9 9	0	0	0
35	AE	9	Total C N O 45 27 9 9	0	0	0
35	AK	9	Total C N O 45 27 9 9	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
35	AO	9	Total	C	N	O	0	0	0
			45	27	9	9			

- Molecule 36 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
36	d	16	Total	C	N	O	0	0	0
			80	48	16	16			
36	o	16	Total	C	N	O	0	0	0
			80	48	16	16			
36	AB	16	Total	C	N	O	0	0	0
			80	48	16	16			
36	AN	16	Total	C	N	O	0	0	0
			80	48	16	16			

- Molecule 37 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
37	e	13	Total	C	N	O	0	0	0
			65	39	13	13			
37	w	13	Total	C	N	O	0	0	0
			65	39	13	13			
37	z	13	Total	C	N	O	0	0	0
			65	39	13	13			

- Molecule 38 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
38	h	47	Total	C	N	O	0	0	0
			235	141	47	47			

- Molecule 39 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
39	j	48	Total	C	N	O	0	0	0
			240	144	48	48			

- Molecule 40 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
40	k	23	Total	C	N	O	0	0	0
			115	69	23	23			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
40	s	23	Total 115	C 69	N 23	O 23	0	0	0

- Molecule 41 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
41	n	36	Total 180	C 108	N 36	O 36	0	0	0
41	q	36	Total 180	C 108	N 36	O 36	0	0	0

- Molecule 42 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
42	p	76	Total 380	C 228	N 76	O 76	0	0	0

- Molecule 43 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
43	t	45	Total 225	C 135	N 45	O 45	0	0	0

- Molecule 44 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
44	u	32	Total 160	C 96	N 32	O 32	0	0	0

- Molecule 45 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
45	v	11	Total 55	C 33	N 11	O 11	0	0	0
45	y	11	Total 54	C 32	N 11	O 11	0	0	0

- Molecule 46 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
46	x	8	Total C N O 40 24 8 8	0	0	0
46	AG	8	Total C N O 40 24 8 8	0	0	0

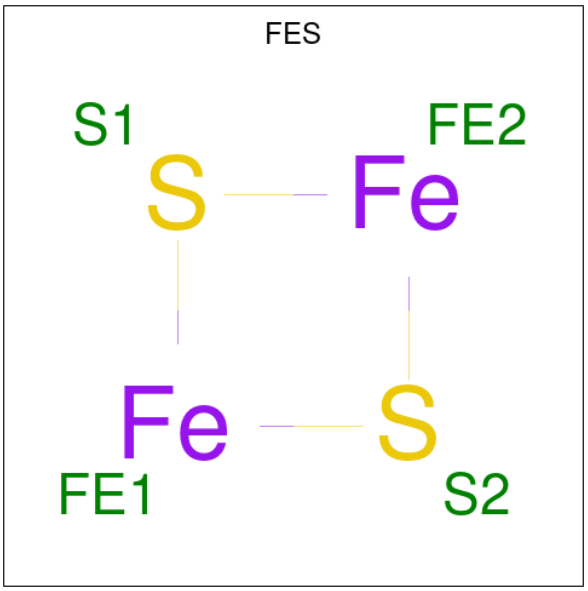
- Molecule 47 is a protein called Unknown polypeptide.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
47	AA	58	Total C N O 290 174 58 58	0	0	0

- Molecule 48 is a protein called Unknown polypeptide.

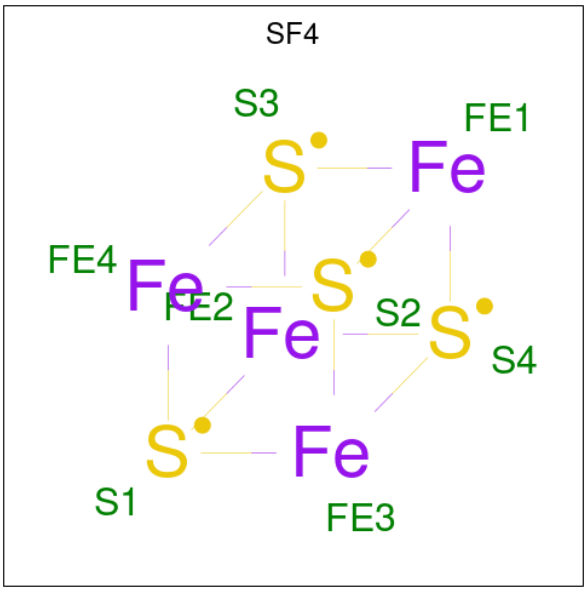
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
48	AD	39	Total C N O 195 117 39 39	0	0	0

- Molecule 49 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula: Fe_2S_2).



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

- Molecule 50 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4).



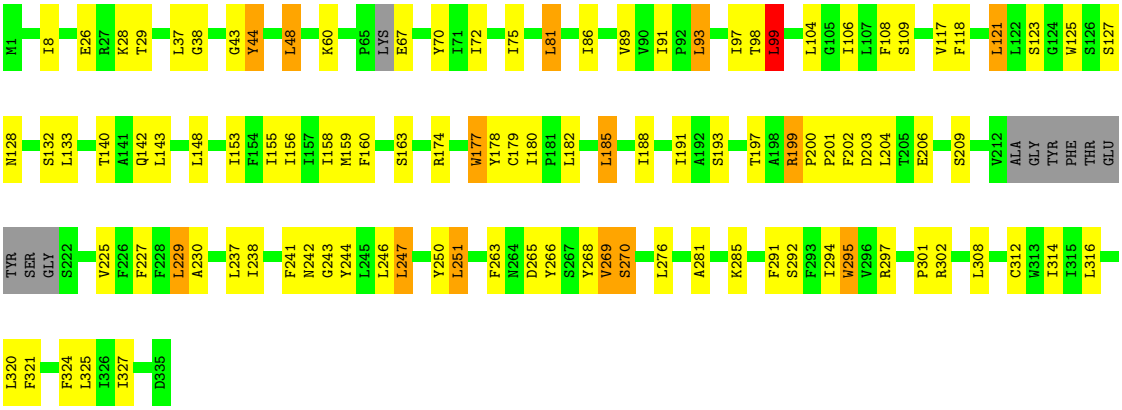
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
50	A	1	Total Fe S 8 4 4	0	0
50	A	1	Total Fe S 8 4 4	0	0
50	B	1	Total Fe S 8 4 4	0	0
50	I	1	Total Fe S 8 4 4	0	0
50	I	1	Total Fe S 8 4 4	0	0
50	K	1	Total Fe S 8 4 4	0	0

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.

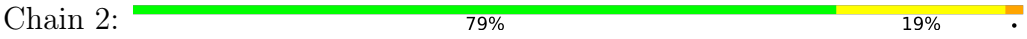
• Molecule 1: NADH-ubiquinone oxidoreductase chain 1

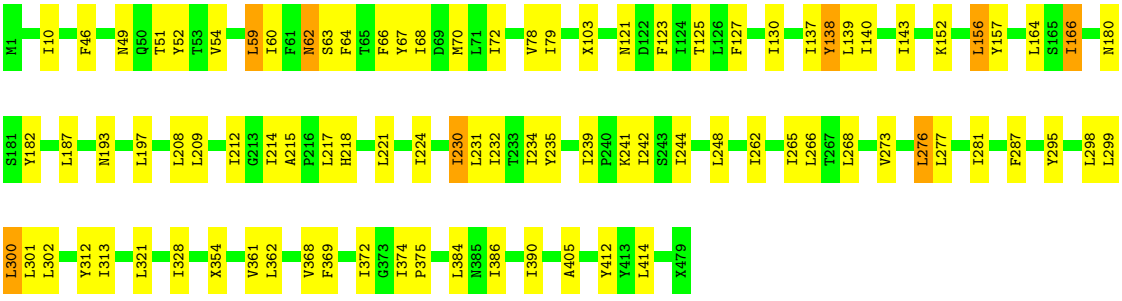
Chain 1:



• Molecule 2: NADH-ubiquinone oxidoreductase chain 2,NADH dehydrogenase subunit 2

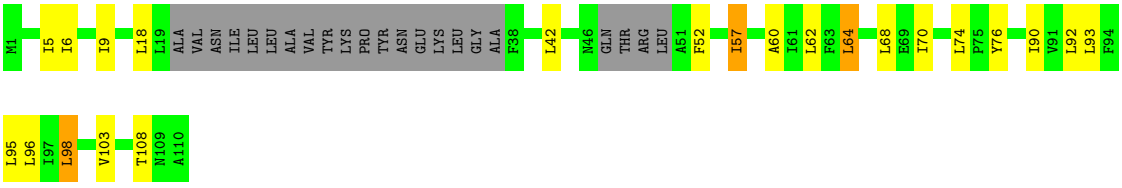
Chain 2:





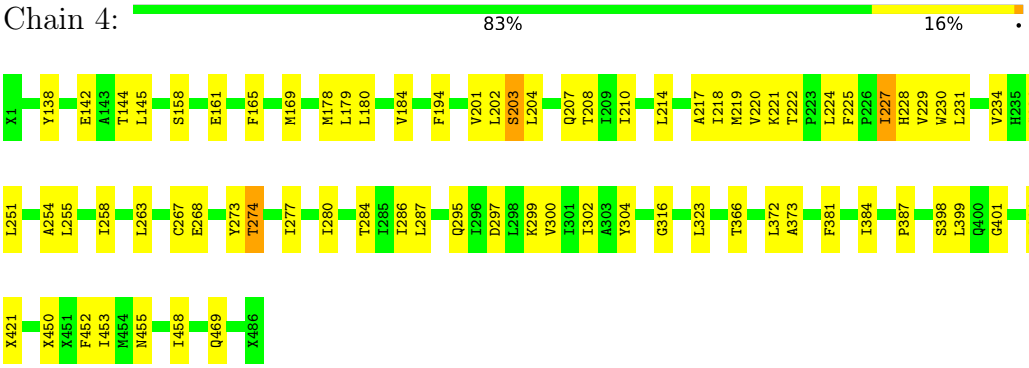
• Molecule 3: NADH-ubiquinone oxidoreductase chain 3

Chain 3:



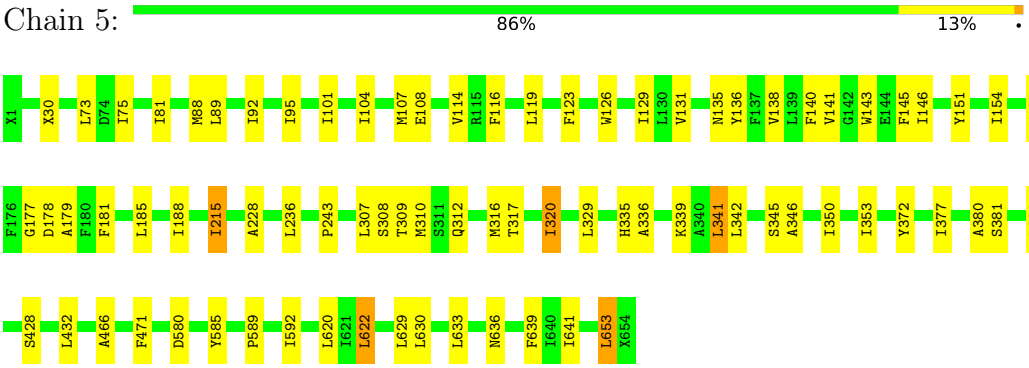
• Molecule 4: NADH-ubiquinone oxidoreductase chain 4

Chain 4:



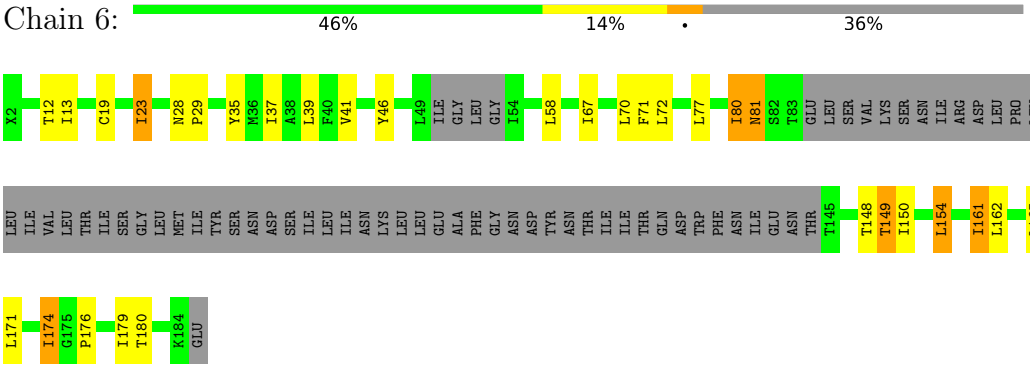
• Molecule 5: NADH-ubiquinone oxidoreductase chain 5

Chain 5:



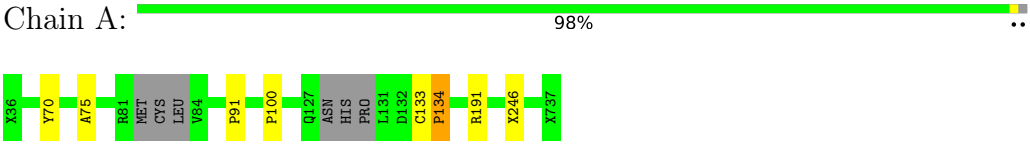
• Molecule 6: NADH-ubiquinone oxidoreductase chain 6

Chain 6:



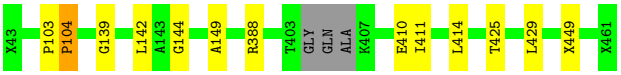
• Molecule 7: NUAM protein

Chain A:



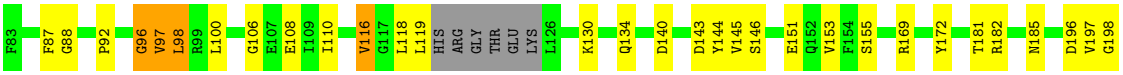
• Molecule 8: NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial

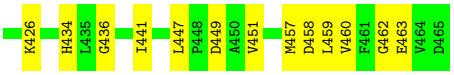
Chain B:



• Molecule 9: NUCM protein

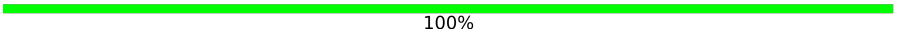
Chain C:





• Molecule 10: NUEM protein

Chain E:



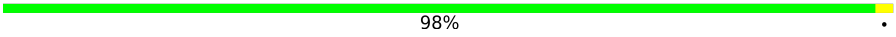
There are no outlier residues recorded for this chain.

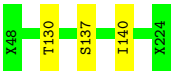
• Molecule 11: NUGM protein

Chain G:

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

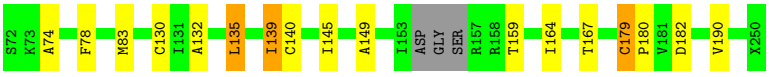
Chain H:





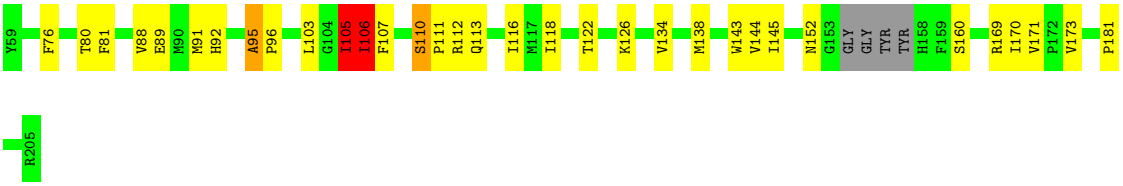
• Molecule 13: Subunit NUIM of NADH:Ubiquinone Oxidoreductase (Complex I)

Chain I:



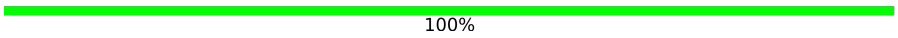
• Molecule 14: Subunit NUKM of protein NADH:Ubiquinone Oxidoreductase (Complex I)

Chain K:



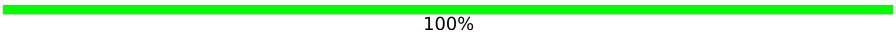
• Molecule 15: NADH-ubiquinone oxidoreductase chain 4L

• Molecule 16: Unknown polypeptide



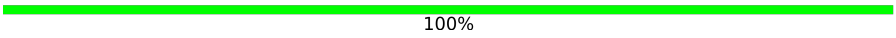
There are no outlier residues recorded for this chain.

• Molecule 16: Unknown polypeptide



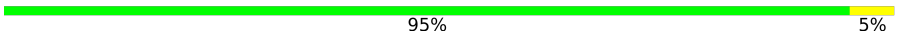
There are no outlier residues recorded for this chain.

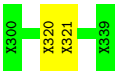
• Molecule 17: Unknown polypeptide



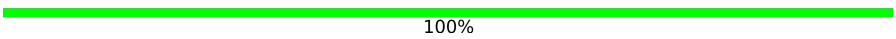
There are no outlier residues recorded for this chain.

• Molecule 17: Unknown polypeptide



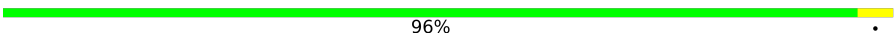


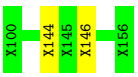
• Molecule 18: Unknown polypeptide



There are no outlier residues recorded for this chain.

• Molecule 19: Unknown polypeptide



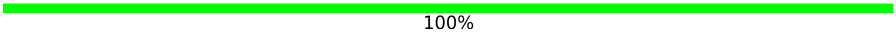


• Molecule 20: Unknown polypeptide

Chain W: 

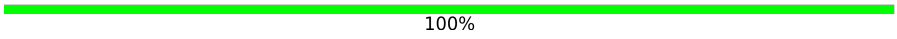
There are no outlier residues recorded for this chain.

• Molecule 21: Unknown polypeptide

Chain V: 

There are no outlier residues recorded for this chain.

• Molecule 22: Unknown polypeptide

Chain T: 

There are no outlier residues recorded for this chain.

• Molecule 22: Unknown polypeptide

Chain m: 

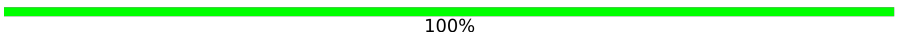
There are no outlier residues recorded for this chain.

• Molecule 22: Unknown polypeptide

Chain AI: 

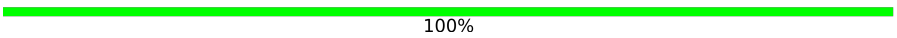
There are no outlier residues recorded for this chain.

• Molecule 23: Unknown polypeptide

Chain S: 

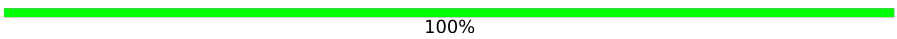
There are no outlier residues recorded for this chain.

• Molecule 23: Unknown polypeptide

Chain AJ: 

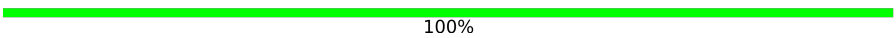
There are no outlier residues recorded for this chain.

• Molecule 23: Unknown polypeptide

Chain AL: 

There are no outlier residues recorded for this chain.

• Molecule 24: Unknown polypeptide

Chain R: 

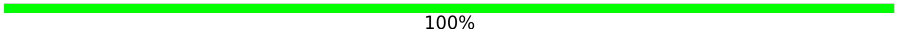
There are no outlier residues recorded for this chain.

• Molecule 25: Unknown polypeptide

Chain Q: 

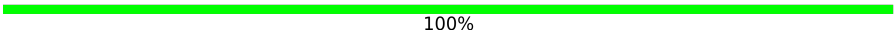
There are no outlier residues recorded for this chain.

• Molecule 26: Unknown polypeptide

Chain P: 

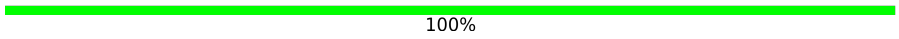
There are no outlier residues recorded for this chain.

• Molecule 27: Unknown polypeptide

Chain F: 

There are no outlier residues recorded for this chain.

• Molecule 27: Unknown polypeptide

Chain f: 

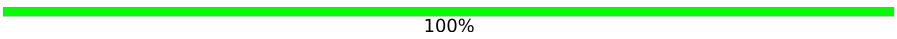
There are no outlier residues recorded for this chain.

• Molecule 27: Unknown polypeptide

Chain AC: 

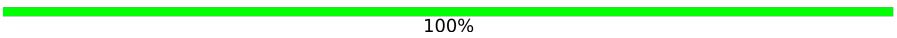
There are no outlier residues recorded for this chain.

• Molecule 27: Unknown polypeptide

Chain AH: 

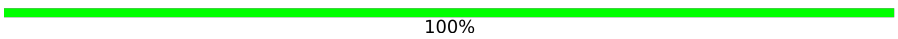
There are no outlier residues recorded for this chain.

• Molecule 27: Unknown polypeptide

Chain AM: 

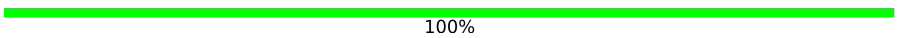
There are no outlier residues recorded for this chain.

• Molecule 28: Unknown polypeptide

Chain O: 

There are no outlier residues recorded for this chain.

• Molecule 28: Unknown polypeptide

Chain l: 

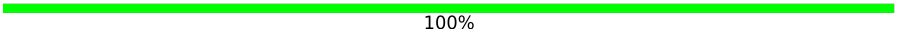
There are no outlier residues recorded for this chain.

• Molecule 29: Unknown polypeptide

Chain M: 

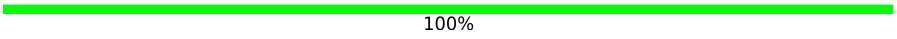
There are no outlier residues recorded for this chain.

• Molecule 30: Unknown polypeptide

Chain D: 

There are no outlier residues recorded for this chain.

• Molecule 31: Unknown polypeptide

Chain J: 

There are no outlier residues recorded for this chain.

• Molecule 32: Unknown polypeptide

Chain N: 

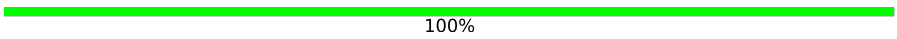
There are no outlier residues recorded for this chain.

• Molecule 33: Unknown polypeptide

Chain a: 

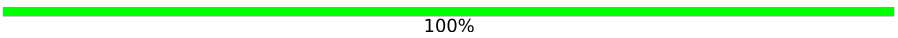
There are no outlier residues recorded for this chain.

• Molecule 33: Unknown polypeptide

Chain i: 

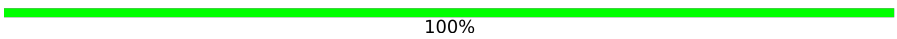
There are no outlier residues recorded for this chain.

• Molecule 34: Unknown polypeptide

Chain b: 

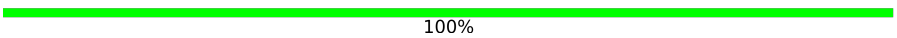
There are no outlier residues recorded for this chain.

• Molecule 35: Unknown polypeptide

Chain c: 

There are no outlier residues recorded for this chain.

• Molecule 35: Unknown polypeptide

Chain g: 

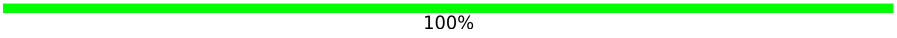
There are no outlier residues recorded for this chain.

• Molecule 35: Unknown polypeptide

Chain AE: 

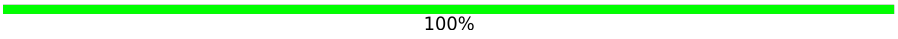
There are no outlier residues recorded for this chain.

• Molecule 35: Unknown polypeptide

Chain AK: 

There are no outlier residues recorded for this chain.

• Molecule 35: Unknown polypeptide

Chain AO: 

There are no outlier residues recorded for this chain.

• Molecule 36: Unknown polypeptide

Chain d: 

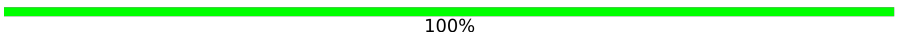
There are no outlier residues recorded for this chain.

• Molecule 36: Unknown polypeptide

Chain o: 

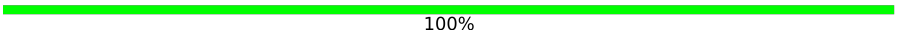
There are no outlier residues recorded for this chain.

• Molecule 36: Unknown polypeptide

Chain AB: 

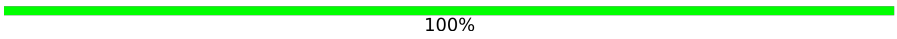
There are no outlier residues recorded for this chain.

• Molecule 36: Unknown polypeptide

Chain AN: 

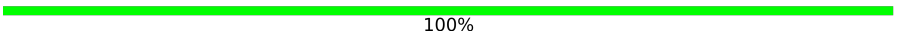
There are no outlier residues recorded for this chain.

• Molecule 37: Unknown polypeptide

Chain e: 

There are no outlier residues recorded for this chain.

• Molecule 37: Unknown polypeptide

Chain w: 

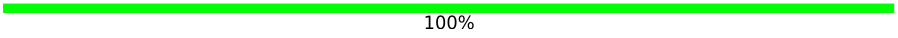
There are no outlier residues recorded for this chain.

• Molecule 37: Unknown polypeptide

Chain z: 

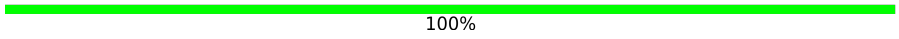
There are no outlier residues recorded for this chain.

• Molecule 38: Unknown polypeptide

Chain h: 

There are no outlier residues recorded for this chain.

• Molecule 39: Unknown polypeptide

Chain j: 

There are no outlier residues recorded for this chain.

• Molecule 40: Unknown polypeptide

Chain k: 

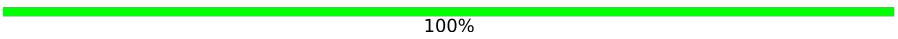
There are no outlier residues recorded for this chain.

• Molecule 40: Unknown polypeptide

Chain s: 

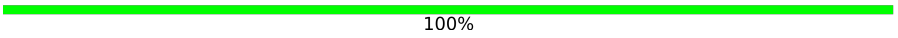
There are no outlier residues recorded for this chain.

• Molecule 41: Unknown polypeptide

Chain n: 

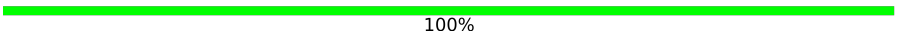
There are no outlier residues recorded for this chain.

• Molecule 41: Unknown polypeptide

Chain q: 

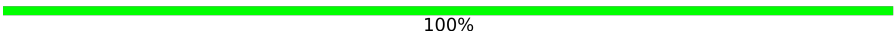
There are no outlier residues recorded for this chain.

• Molecule 42: Unknown polypeptide

Chain p: 

There are no outlier residues recorded for this chain.

• Molecule 43: Unknown polypeptide

Chain t: 

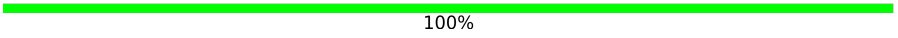
There are no outlier residues recorded for this chain.

• Molecule 44: Unknown polypeptide

Chain u: 

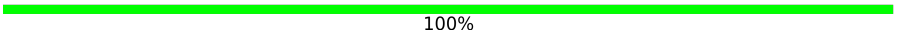
There are no outlier residues recorded for this chain.

• Molecule 45: Unknown polypeptide

Chain v: 

There are no outlier residues recorded for this chain.

• Molecule 45: Unknown polypeptide

Chain y: 

There are no outlier residues recorded for this chain.

• Molecule 46: Unknown polypeptide

Chain x: 

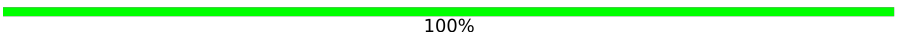
There are no outlier residues recorded for this chain.

• Molecule 46: Unknown polypeptide

Chain AG: 

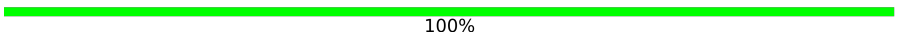
There are no outlier residues recorded for this chain.

• Molecule 47: Unknown polypeptide

Chain AA: 

There are no outlier residues recorded for this chain.

• Molecule 48: Unknown polypeptide

Chain AD: 

There are no outlier residues recorded for this chain.

4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	316.31Å 316.31Å 819.18Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.00 – 3.79 49.18 – 3.79	Depositor EDS
% Data completeness (in resolution range)	77.1 (40.00-3.79) 77.4 (49.18-3.79)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.57 (at 3.77Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.361 , 0.363 0.412 , 0.419	Depositor DCC
R_{free} test set	5835 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å^2)	142.0	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.23 , 238.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.77	EDS
Total number of atoms	35631	wwPDB-VP
Average B, all atoms (Å^2)	164.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.73% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.46	0/2648	0.70	0/3616
2	2	0.42	0/2699	0.64	0/3672
3	3	0.48	0/699	0.74	0/954
4	4	0.45	0/1991	0.66	0/2716
5	5	0.43	0/2797	0.62	0/3796
6	6	0.45	0/836	0.70	0/1141
7	A	0.43	0/547	0.71	3/740 (0.4%)
8	B	0.41	0/670	0.62	2/901 (0.2%)
9	C	0.42	0/2946	0.66	0/3989
11	G	0.39	0/581	0.55	0/787
12	H	0.39	0/172	0.63	0/210
13	I	0.45	0/614	0.69	0/836
14	K	0.44	0/1092	0.66	0/1489
15	L	0.44	0/620	0.71	0/839
All	All	0.44	0/18912	0.66	5/25686 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	A	134	PRO	N-CA-CB	6.84	111.50	103.30
8	B	104	PRO	N-CA-CB	5.92	110.40	103.30
8	B	103	PRO	N-CA-CB	5.83	110.29	103.30
7	A	91	PRO	N-CA-CB	5.73	110.18	103.30
7	A	100	PRO	N-CA-CB	5.67	110.10	103.30

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	1	2581	0	2675	53	0
2	2	3122	0	2880	48	0
3	3	685	0	719	16	0
4	4	3017	0	2290	38	0
5	5	4050	0	3162	40	0
6	6	870	0	916	25	0
7	A	3221	0	1001	2	0
8	B	1997	0	856	4	0
9	C	2882	0	2825	50	0
10	E	975	0	228	0	0
11	G	880	0	582	5	0
12	H	803	0	299	1	0
13	I	857	0	582	6	0
14	K	1069	0	1038	20	0
15	L	612	0	664	14	0
16	Z	85	0	19	0	0
16	r	85	0	20	0	0
17	AF	200	0	46	1	0
17	Y	200	0	43	0	0
18	U	50	0	13	0	0
19	X	285	0	64	1	0
20	W	270	0	59	0	0
21	V	315	0	69	0	0
22	AI	100	0	22	0	0
22	T	100	0	22	0	0
22	m	100	0	22	0	0
23	AJ	95	0	22	0	0
23	AL	95	0	21	0	0
23	S	95	0	21	0	0
24	R	250	0	53	0	0
25	Q	350	0	75	0	0
26	P	140	0	32	0	0
27	AC	90	0	22	0	0
27	AH	90	0	20	0	0
27	AM	90	0	21	0	0
27	F	90	0	20	0	0
27	f	90	0	20	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
28	O	125	0	27	0	0
28	l	125	0	27	0	0
29	M	255	0	56	0	0
30	D	150	0	32	0	0
31	J	345	0	74	0	0
32	N	75	0	17	0	0
33	a	130	0	29	0	0
33	i	130	0	31	0	0
34	b	110	0	24	0	0
35	AE	45	0	11	0	0
35	AK	45	0	11	0	0
35	AO	45	0	11	0	0
35	c	45	0	11	0	0
35	g	45	0	12	0	0
36	AB	80	0	19	0	0
36	AN	80	0	18	0	0
36	d	80	0	19	0	0
36	o	80	0	19	0	0
37	e	65	0	15	0	0
37	w	65	0	17	0	0
37	z	65	0	15	0	0
38	h	235	0	49	0	0
39	j	240	0	52	0	0
40	k	115	0	25	0	0
40	s	115	0	26	0	0
41	n	180	0	38	0	0
41	q	180	0	44	0	0
42	p	380	0	78	0	0
43	t	225	0	47	0	0
44	u	160	0	34	0	0
45	v	55	0	14	0	0
45	y	54	0	14	0	0
46	AG	40	0	10	0	0
46	x	40	0	10	0	0
47	AA	290	0	64	0	0
48	AD	195	0	43	0	0
49	A	4	0	0	0	0
49	H	4	0	0	0	0
50	A	16	0	0	0	0
50	B	8	0	0	0	0
50	I	16	0	0	0	0
50	K	8	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	35631	0	22486	279	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 279 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)
4:4:225:PHE:H	4:4:284:THR:HG22	1.39	0.86
1:1:142:GLN:HG3	1:1:308:LEU:HD23	1.62	0.82
2:2:125:THR:HA	6:6:161:ILE:HD11	1.70	0.73
5:5:126:TRP:CZ3	5:5:146:ILE:HG12	2.24	0.72
6:6:167:ILE:HA	6:6:170:LEU:HG	1.70	0.71

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	319/335 (95%)	274 (86%)	29 (9%)	16 (5%)	2 23
2	2	337/434 (78%)	309 (92%)	24 (7%)	4 (1%)	13 50
3	3	82/110 (74%)	68 (83%)	12 (15%)	2 (2%)	6 37
4	4	257/470 (55%)	228 (89%)	24 (9%)	5 (2%)	8 42
5	5	353/616 (57%)	315 (89%)	31 (9%)	7 (2%)	7 41
6	6	104/184 (56%)	89 (86%)	10 (10%)	5 (5%)	2 24
7	A	84/634 (13%)	72 (86%)	9 (11%)	3 (4%)	3 30
8	B	95/370 (26%)	90 (95%)	4 (4%)	1 (1%)	14 51
9	C	361/383 (94%)	296 (82%)	41 (11%)	24 (7%)	1 19

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
11	G	67/133 (50%)	58 (87%)	7 (10%)	2 (3%)	4 33
12	H	29/154 (19%)	27 (93%)	2 (7%)	0	100 100
13	I	82/140 (59%)	63 (77%)	13 (16%)	6 (7%)	1 16
14	K	139/147 (95%)	122 (88%)	11 (8%)	6 (4%)	2 26
15	L	77/79 (98%)	69 (90%)	5 (6%)	3 (4%)	3 28
All	All	2386/4189 (57%)	2080 (87%)	222 (9%)	84 (4%)	3 31

5 of 84 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	1	60	LYS
1	1	269	VAL
2	2	62	ASN
5	5	346	ALA
7	A	133	CYS

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	1	286/296 (97%)	252 (88%)	34 (12%)	5 26
2	2	291/308 (94%)	266 (91%)	25 (9%)	10 40
3	3	75/96 (78%)	68 (91%)	7 (9%)	9 35
4	4	215/229 (94%)	198 (92%)	17 (8%)	12 42
5	5	301/302 (100%)	277 (92%)	24 (8%)	12 42
6	6	90/157 (57%)	81 (90%)	9 (10%)	7 32
7	A	37/83 (45%)	37 (100%)	0	100 100
8	B	53/83 (64%)	50 (94%)	3 (6%)	20 52
9	C	305/324 (94%)	278 (91%)	27 (9%)	9 38
11	G	55/58 (95%)	55 (100%)	0	100 100
12	H	19/19 (100%)	18 (95%)	1 (5%)	22 54

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
13	I	55/78 (70%)	49 (89%)	6 (11%)	6 29
14	K	114/125 (91%)	107 (94%)	7 (6%)	18 50
15	L	68/69 (99%)	66 (97%)	2 (3%)	42 67
All	All	1964/2227 (88%)	1802 (92%)	162 (8%)	11 40

5 of 162 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	6	174	ILE
9	C	403	ILE
9	C	98	LEU
9	C	305	ILE
13	I	135	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
9	C	95	HIS
9	C	150	ASN
14	K	195	GLN
9	C	288	ASN
9	C	434	HIS

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)

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	$\#\|Z\| > 2$
50	SF4	B	500	8	0,12,12	-	-	-		
50	SF4	I	500	13	0,12,12	-	-	-		
49	FES	A	900	7	0,4,4	-	-	-		
50	SF4	I	501	13	0,12,12	-	-	-		
50	SF4	A	901	7	0,12,12	-	-	-		
50	SF4	A	902	7	0,12,12	-	-	-		
50	SF4	K	500	14	0,12,12	-	-	-		
49	FES	H	300	12	0,4,4	-	-	-		

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
50	SF4	B	500	8	-	-	0/6/5/5
50	SF4	I	500	13	-	-	0/6/5/5
49	FES	A	900	7	-	-	0/1/1/1
50	SF4	I	501	13	-	-	0/6/5/5
50	SF4	A	901	7	-	-	0/6/5/5
50	SF4	A	902	7	-	-	0/6/5/5
50	SF4	K	500	14	-	-	0/6/5/5
49	FES	H	300	12	-	-	0/1/1/1

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.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
7	A	27
10	E	10
8	B	10
5	5	8
12	H	5
13	I	4
2	2	4
4	4	4
11	G	2
41	q	1
25	Q	1

The worst 5 of 76 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	48.59
1	B	270:UNK	C	279:UNK	N	26.48
1	A	148:SER	C	152:UNK	N	24.96
1	E	513:UNK	C	604:UNK	N	24.71
1	I	92:GLU	C	103:UNK	N	24.13

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.