



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 21, 2016 – 02:04 AM EDT

PDB ID : 4WZJ  
Title : Spliceosomal U4 snRNP core domain  
Authors : Leung, A.K.W.; Nagai, K.; Li, J.  
Deposited on : 2014-11-19  
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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027939

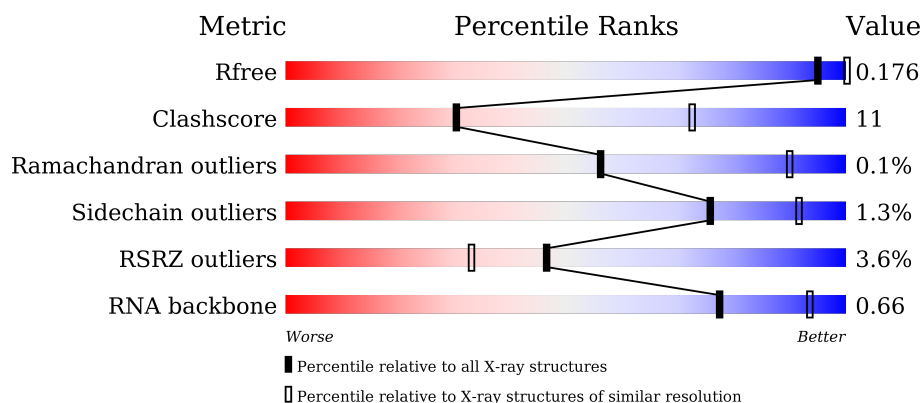
# 1 Overall quality at a glance

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))
$R_{free}$	91344	1408 (3.80-3.40)
Clashscore	102246	1010 (3.74-3.46)
Ramachandran outliers	100387	1007 (3.76-3.44)
Sidechain outliers	100360	1007 (3.76-3.44)
RSRZ outliers	91569	1003 (3.78-3.42)
RNA backbone	2183	1058 (4.40-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	125	<div> <div>3%</div> <div> <div></div> <div>54%</div> <div>12%</div> <div>34%</div> </div> </div>
1	AA	125	<div> <div>%</div> <div> <div></div> <div>55%</div> <div>11%</div> <div>34%</div> </div> </div>
1	AAA	125	<div> <div>4%</div> <div> <div></div> <div>54%</div> <div>11%</div> <div>34%</div> </div> </div>
1	AAAA	125	<div> <div>3%</div> <div> <div></div> <div>66%</div> <div>34%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	125	
1	HH	125	
1	HHH	125	
1	HHHH	125	
1	O	125	
1	OO	125	
1	OOO	125	
1	OOOO	125	
2	B	95	
2	BB	95	
2	BBB	95	
2	BBBB	95	
2	I	95	
2	II	95	
2	III	95	
2	III	95	
2	P	95	
2	PP	95	
2	PPP	95	
2	PPPP	95	
3	C	118	
3	CC	118	
3	CCC	118	
3	CCCC	118	
3	J	118	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	JJ	118	
3	JJJ	118	
3	JJJJ	118	
3	Q	118	
3	QQ	118	
3	QQQ	118	
3	QQQQ	118	
4	D	118	
4	DD	118	
4	DDD	118	
4	DDDD	118	
4	K	118	
4	KK	118	
4	KKK	118	
4	KKKK	118	
4	R	118	
4	RR	118	
4	RRR	118	
4	RRRR	118	
5	F	86	
5	FF	86	
5	FFF	86	
5	FFFF	86	
5	M	86	
5	MM	86	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	MMM	86	
5	MMMM	86	
5	T	86	
5	TT	86	
5	TTT	86	
5	TTTT	86	
6	E	92	
6	EE	92	
6	EEE	92	
6	EEEE	92	
6	L	92	
6	LL	92	
6	LLL	92	
6	LLLL	92	
6	S	92	
6	SS	92	
6	SSS	92	
6	SSSS	92	
7	G	76	
7	GG	76	
7	GGG	76	
7	GGGG	76	
7	N	76	
7	NN	76	
7	NNN	76	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
7	NNNN	76	
7	U	76	
7	UU	76	
7	UUU	76	
7	UUUU	76	
8	V	68	
8	VV	68	
8	VVV	68	
8	VVVV	68	
8	X	68	
8	XX	68	
8	XXX	68	
8	XXXX	68	
8	Y	68	
8	YY	68	
8	YYY	68	
8	YYYY	68	

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 71485 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 Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	83	Total	C	N	O	S	0	0	0
			652	409	115	122	6			
1	H	84	Total	C	N	O	S	0	0	0
			658	412	116	124	6			
1	O	83	Total	C	N	O	S	0	0	0
			652	409	115	122	6			
1	AA	83	Total	C	N	O	S	0	0	0
			652	409	115	122	6			
1	HH	84	Total	C	N	O	S	0	0	0
			653	409	115	123	6			
1	OO	81	Total	C	N	O	S	0	0	0
			637	400	112	119	6			
1	AAA	83	Total	C	N	O	S	0	0	0
			648	406	114	122	6			
1	HHH	84	Total	C	N	O	S	0	0	0
			658	412	116	124	6			
1	OOO	82	Total	C	N	O	S	0	0	0
			646	406	114	120	6			
1	AAAA	82	Total	C	N	O	S	0	0	0
			646	406	114	120	6			
1	HHHH	82	Total	C	N	O	S	0	0	0
			643	403	113	121	6			
1	OOOO	82	Total	C	N	O	S	0	0	0
			646	406	114	120	6			

- Molecule 2 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	86	Total	C	N	O	S	0	0	0
			690	434	126	123	7			
2	I	72	Total	C	N	O	S	0	0	0
			574	364	103	100	7			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	P	71	Total	C	N	O	S	0	0	0
			569	361	102	99	7			
2	BB	71	Total	C	N	O	S	0	0	0
			565	358	101	99	7			
2	II	75	Total	C	N	O	S	0	0	0
			585	370	105	103	7			
2	PP	71	Total	C	N	O	S	0	0	0
			569	361	102	99	7			
2	BBB	71	Total	C	N	O	S	0	0	0
			565	358	101	99	7			
2	III	71	Total	C	N	O	S	0	0	0
			565	358	101	99	7			
2	PPP	71	Total	C	N	O	S	0	0	0
			559	355	98	99	7			
2	BBBB	74	Total	C	N	O	S	0	0	0
			592	375	106	104	7			
2	IIII	75	Total	C	N	O	S	0	0	0
			602	381	109	105	7			
2	PPPP	75	Total	C	N	O	S	0	0	0
			598	381	107	103	7			

- Molecule 3 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	82	Total	C	N	O	S	0	0	0
			649	413	113	119	4			
3	J	82	Total	C	N	O	S	0	0	0
			649	413	113	119	4			
3	Q	82	Total	C	N	O	S	0	0	0
			649	413	113	119	4			
3	CC	82	Total	C	N	O	S	0	0	0
			649	413	113	119	4			
3	JJ	82	Total	C	N	O	S	0	0	0
			649	413	113	119	4			
3	QQ	82	Total	C	N	O	S	0	0	0
			649	413	113	119	4			
3	CCC	82	Total	C	N	O	S	0	0	0
			649	413	113	119	4			
3	JJJ	82	Total	C	N	O	S	0	0	0
			649	413	113	119	4			
3	QQQ	82	Total	C	N	O	S	0	0	0
			649	413	113	119	4			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	CCCC	82	Total	C	N	O	S	0	0	0
			649	413	113	119	4			
3	JJJJ	82	Total	C	N	O	S	0	0	0
			649	413	113	119	4			
3	QQQQ	82	Total	C	N	O	S	0	0	0
			649	413	113	119	4			

- Molecule 4 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	97	Total	C	N	O	S	0	0	0
			776	488	143	140	5			
4	K	100	Total	C	N	O	S	0	0	0
			796	499	149	143	5			
4	R	98	Total	C	N	O	S	0	0	0
			787	494	147	141	5			
4	DD	98	Total	C	N	O	S	0	0	0
			783	491	146	141	5			
4	KK	104	Total	C	N	O	S	0	0	0
			838	526	155	152	5			
4	RR	98	Total	C	N	O	S	0	0	0
			787	494	147	141	5			
4	DDD	98	Total	C	N	O	S	0	0	0
			787	494	147	141	5			
4	KKK	99	Total	C	N	O	S	0	0	0
			786	494	146	141	5			
4	RRR	98	Total	C	N	O	S	0	0	0
			787	494	147	141	5			
4	DDDD	104	Total	C	N	O	S	0	0	0
			838	526	155	152	5			
4	KKKK	104	Total	C	N	O	S	0	0	0
			838	526	155	152	5			
4	RRRR	98	Total	C	N	O	S	0	0	0
			787	494	147	141	5			

- Molecule 5 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	74	Total	C	N	O	S	0	0	0
			576	373	95	103	5			
5	M	79	Total	C	N	O	S	0	0	0
			609	392	100	112	5			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	T	76	Total	C	N	O	S	0	0	0
			594	383	97	109	5			
5	FF	75	Total	C	N	O	S	0	0	0
			585	378	96	106	5			
5	MM	80	Total	C	N	O	S	0	0	0
			621	399	101	115	6			
5	TT	78	Total	C	N	O	S	0	0	0
			596	385	99	107	5			
5	FFF	74	Total	C	N	O	S	0	0	0
			576	373	95	103	5			
5	MMM	76	Total	C	N	O	S	0	0	0
			590	381	97	107	5			
5	TTT	76	Total	C	N	O	S	0	0	0
			590	381	97	107	5			
5	FFFF	74	Total	C	N	O	S	0	0	0
			576	373	95	103	5			
5	MMMM	78	Total	C	N	O	S	0	0	0
			604	389	99	111	5			
5	TTTT	77	Total	C	N	O	S	0	0	0
			599	386	98	110	5			

- Molecule 6 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	E	79	Total	C	N	O	S	0	0	0
			652	412	116	119	5			
6	L	79	Total	C	N	O	S	0	0	0
			652	412	116	119	5			
6	S	79	Total	C	N	O	S	0	0	0
			652	412	116	119	5			
6	EE	79	Total	C	N	O	S	0	0	0
			652	412	116	119	5			
6	LL	79	Total	C	N	O	S	0	0	0
			652	412	116	119	5			
6	SS	79	Total	C	N	O	S	0	0	0
			652	412	116	119	5			
6	EEE	79	Total	C	N	O	S	0	0	0
			652	412	116	119	5			
6	LLL	79	Total	C	N	O	S	0	0	0
			652	412	116	119	5			
6	SSS	79	Total	C	N	O	S	0	0	0
			652	412	116	119	5			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	EEEE	79	Total	C	N	O	S	0	0	0
			652	412	116	119	5			
6	LLLL	77	Total	C	N	O	S	0	0	0
			638	405	113	115	5			
6	SSSS	79	Total	C	N	O	S	0	0	0
			652	412	116	119	5			

- Molecule 7 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	74	Total	C	N	O	S	0	0	0
			577	364	104	103	6			
7	N	74	Total	C	N	O	S	0	0	0
			577	364	104	103	6			
7	U	74	Total	C	N	O	S	0	0	0
			577	364	104	103	6			
7	GG	74	Total	C	N	O	S	0	0	0
			577	364	104	103	6			
7	NN	74	Total	C	N	O	S	0	0	0
			577	364	104	103	6			
7	UU	74	Total	C	N	O	S	0	0	0
			571	361	101	103	6			
7	GGG	74	Total	C	N	O	S	0	0	0
			577	364	104	103	6			
7	NNN	74	Total	C	N	O	S	0	0	0
			577	364	104	103	6			
7	UUU	74	Total	C	N	O	S	0	0	0
			577	364	104	103	6			
7	GGGG	74	Total	C	N	O	S	0	0	0
			577	364	104	103	6			
7	NNNN	74	Total	C	N	O	S	0	0	0
			577	364	104	103	6			
7	UUUU	74	Total	C	N	O	S	0	0	0
			577	364	104	103	6			

- Molecule 8 is a RNA chain called U4 small nuclear RNA variant: Native sequence 85-145, of which nucleotides 97-104 are replaced with GAAA tetraloop and nucleotides 134-137 are replaced with GAAA tetraloop receptor..

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	V	68	Total	C	N	O	P	0	0	0
			1453	650	263	473	67			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	X	68	Total 1453	C 650	N 263	O 473	P 67	0	0	0
8	Y	68	Total 1453	C 650	N 263	O 473	P 67	0	0	0
8	VV	68	Total 1453	C 650	N 263	O 473	P 67	0	0	0
8	XX	68	Total 1453	C 650	N 263	O 473	P 67	0	0	0
8	YY	68	Total 1453	C 650	N 263	O 473	P 67	0	0	0
8	VVV	68	Total 1453	C 650	N 263	O 473	P 67	0	0	0
8	XXX	68	Total 1453	C 650	N 263	O 473	P 67	0	0	0
8	YYY	68	Total 1453	C 650	N 263	O 473	P 67	0	0	0
8	VVVV	68	Total 1453	C 650	N 263	O 473	P 67	0	0	0
8	XXXX	68	Total 1453	C 650	N 263	O 473	P 67	0	0	0
8	YYYY	68	Total 1453	C 650	N 263	O 473	P 67	0	0	0

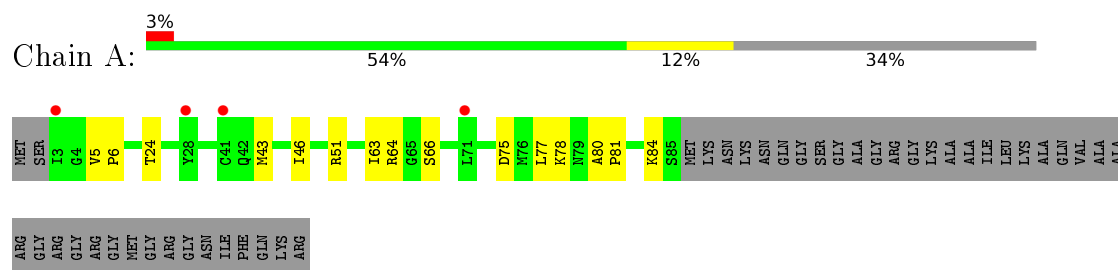
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	R	1	Total 1	O 1	0	0
9	DD	1	Total 1	O 1	0	0
9	RR	1	Total 1	O 1	0	0

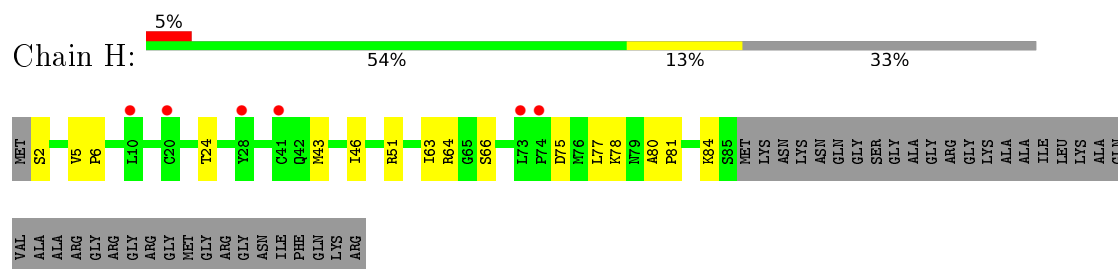
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

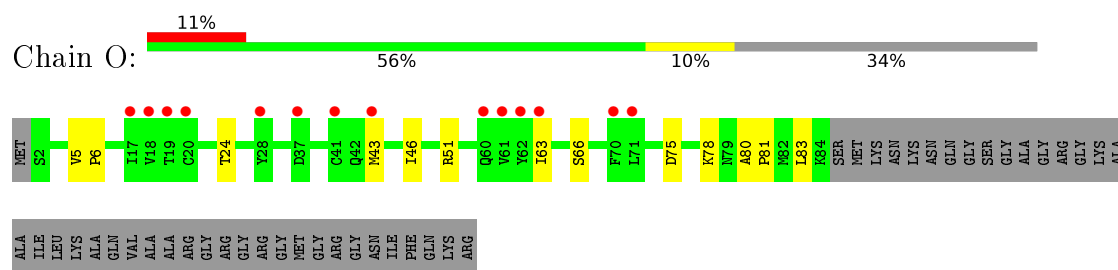
- Molecule 1: Small nuclear ribonucleoprotein Sm D3



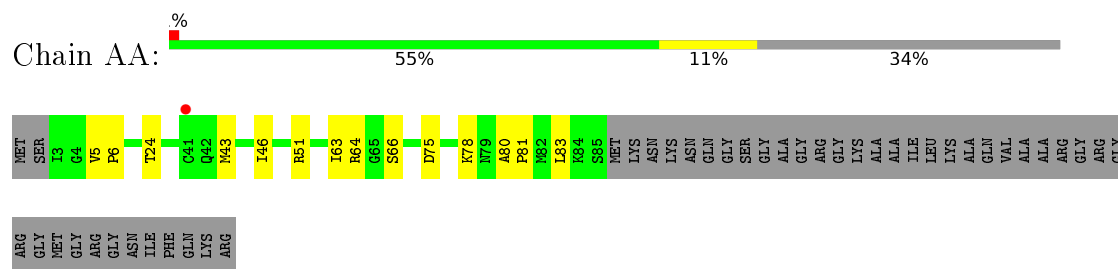
- Molecule 1: Small nuclear ribonucleoprotein Sm D3



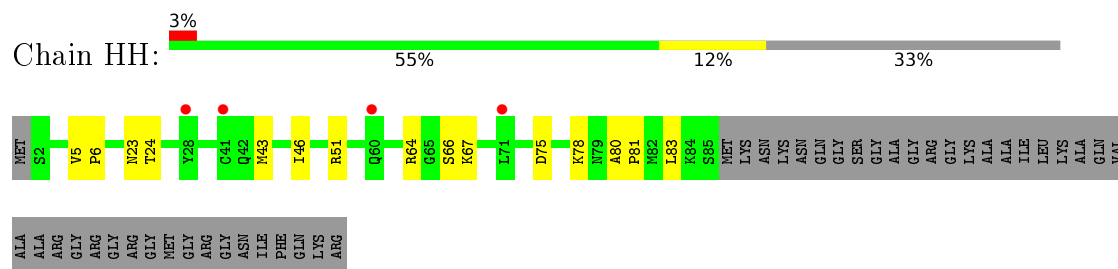
- Molecule 1: Small nuclear ribonucleoprotein Sm D3



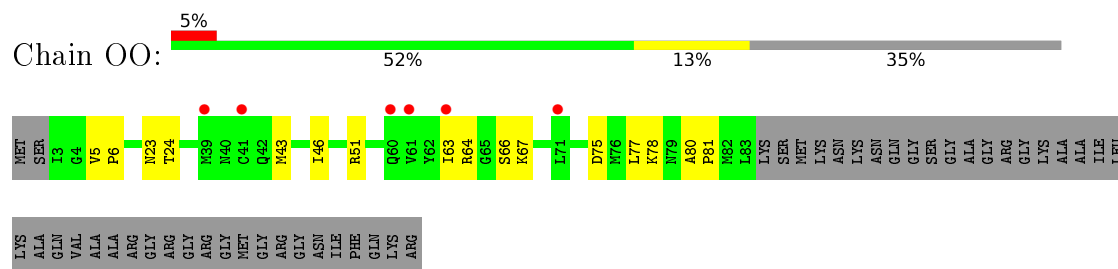
- Molecule 1: Small nuclear ribonucleoprotein Sm D3



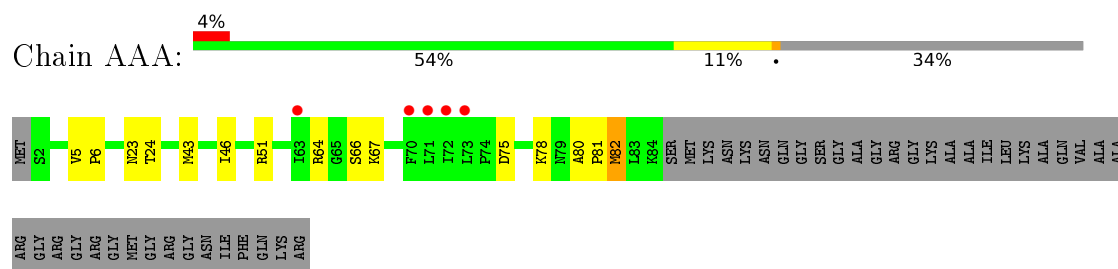
- Molecule 1: Small nuclear ribonucleoprotein Sm D3



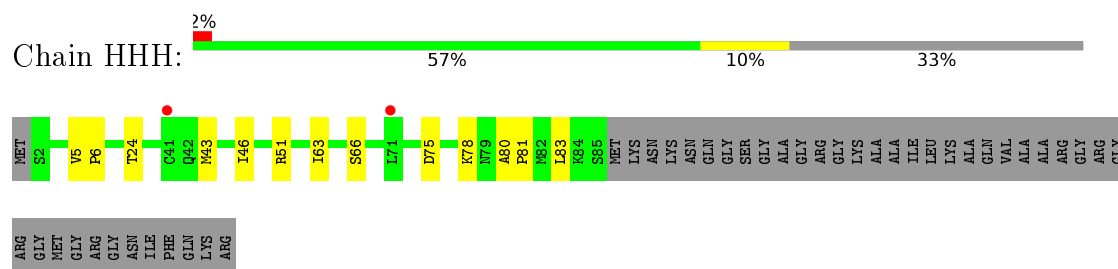
- Molecule 1: Small nuclear ribonucleoprotein Sm D3



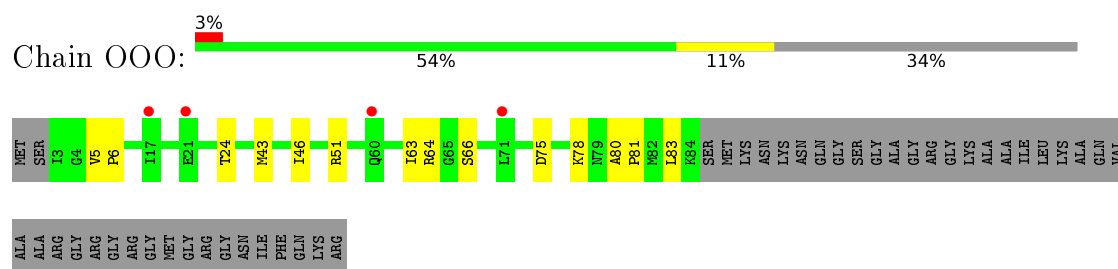
- Molecule 1: Small nuclear ribonucleoprotein Sm D3



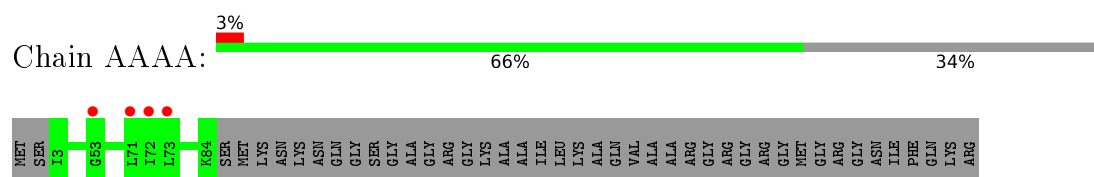
- Molecule 1: Small nuclear ribonucleoprotein Sm D3



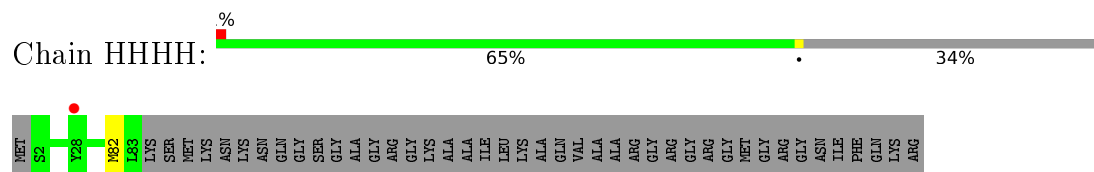
- Molecule 1: Small nuclear ribonucleoprotein Sm D3



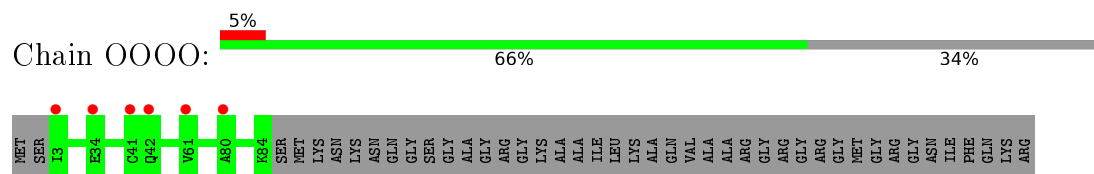
- Molecule 1: Small nuclear ribonucleoprotein Sm D3



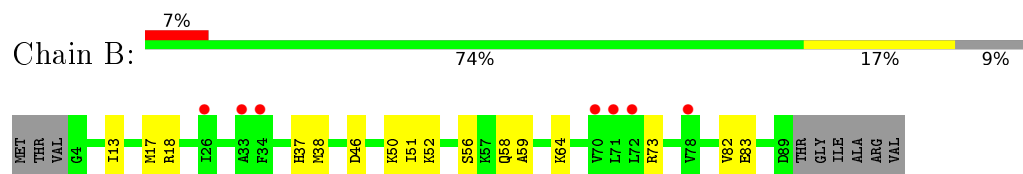
- Molecule 1: Small nuclear ribonucleoprotein Sm D3



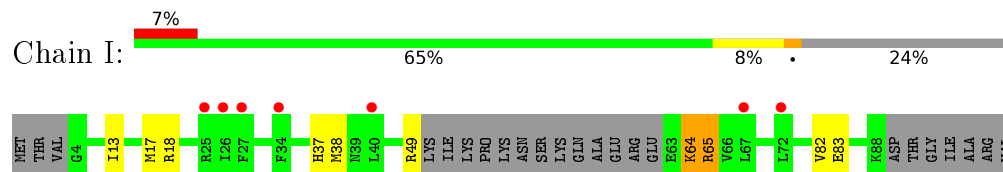
- Molecule 1: Small nuclear ribonucleoprotein Sm D3



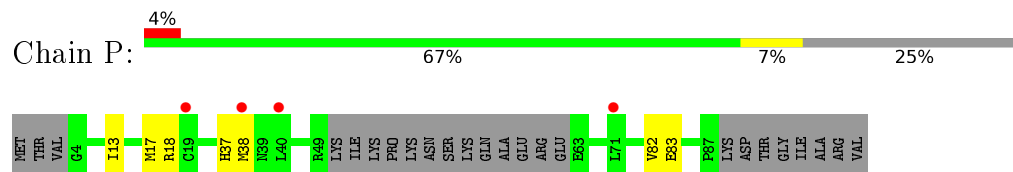
- Molecule 2: Small nuclear ribonucleoprotein-associated proteins B and B'



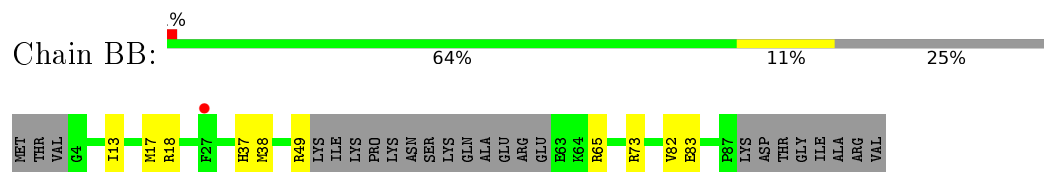
- Molecule 2: Small nuclear ribonucleoprotein-associated proteins B and B'



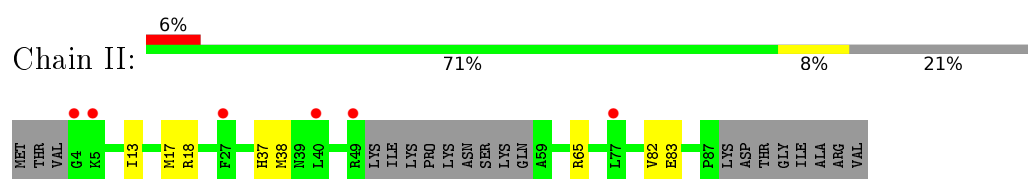
- Molecule 2: Small nuclear ribonucleoprotein-associated proteins B and B'



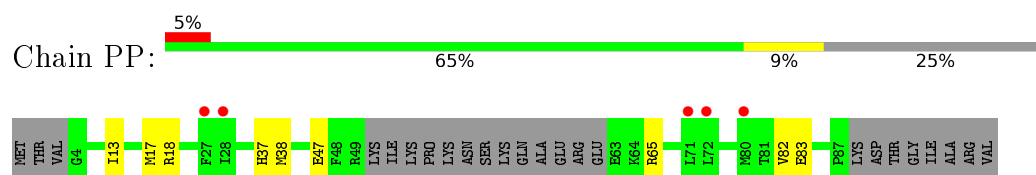
- Molecule 2: Small nuclear ribonucleoprotein-associated proteins B and B'



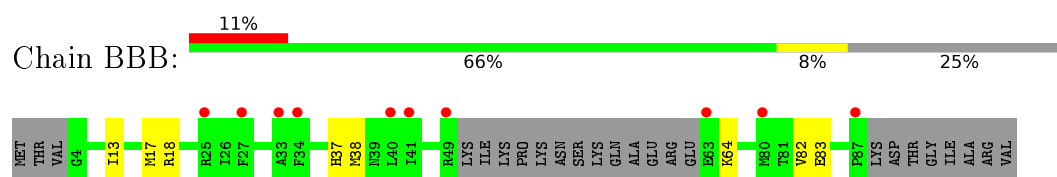
- Molecule 2: Small nuclear ribonucleoprotein-associated proteins B and B'



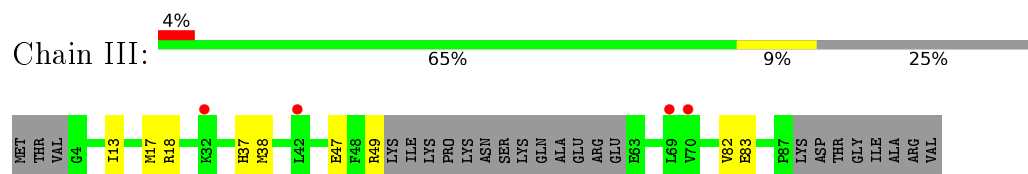
- Molecule 2: Small nuclear ribonucleoprotein-associated proteins B and B'



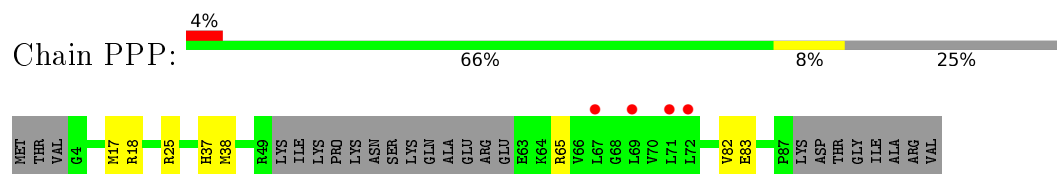
- Molecule 2: Small nuclear ribonucleoprotein-associated proteins B and B'



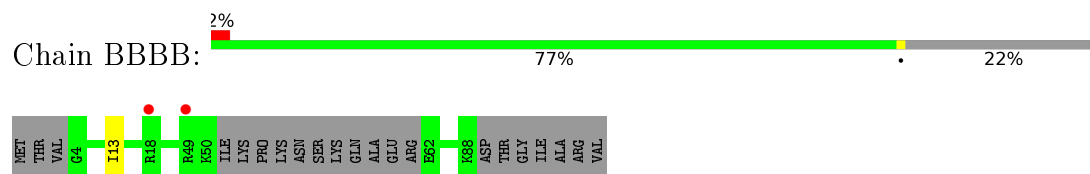
- Molecule 2: Small nuclear ribonucleoprotein-associated proteins B and B'



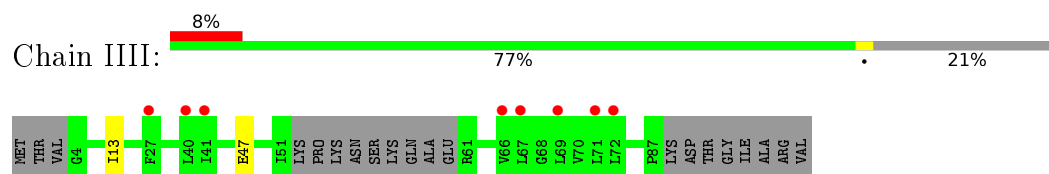
- Molecule 2: Small nuclear ribonucleoprotein-associated proteins B and B'



- Molecule 2: Small nuclear ribonucleoprotein-associated proteins B and B'



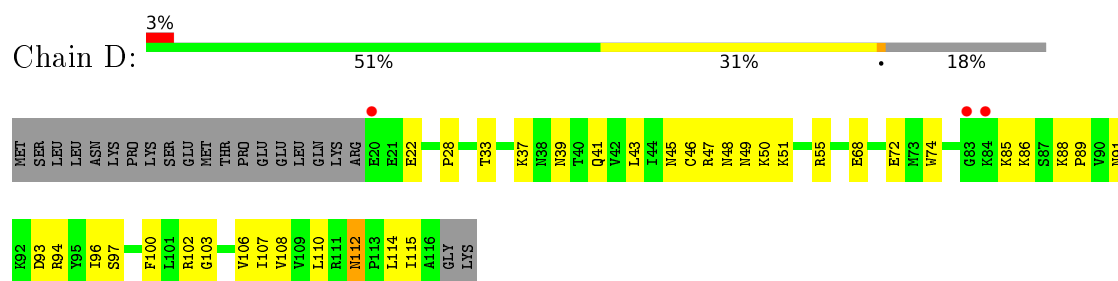
- Molecule 2: Small nuclear ribonucleoprotein-associated proteins B and B'



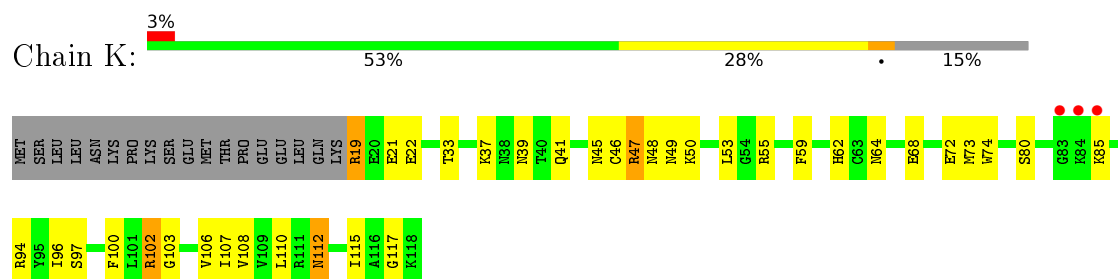
- Molecule 2: Small nuclear ribonucleoprotein-associated proteins B and B'



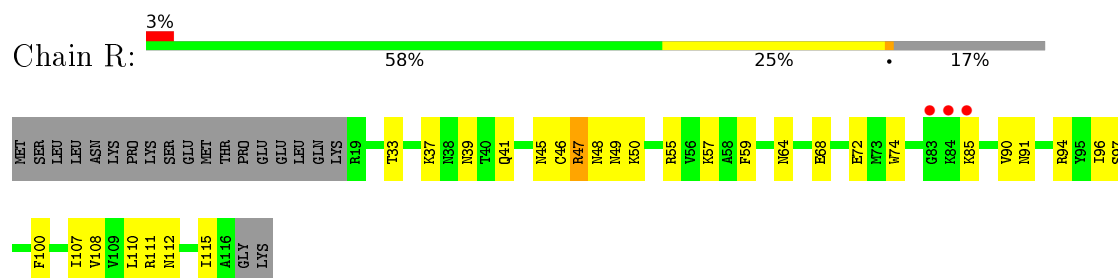




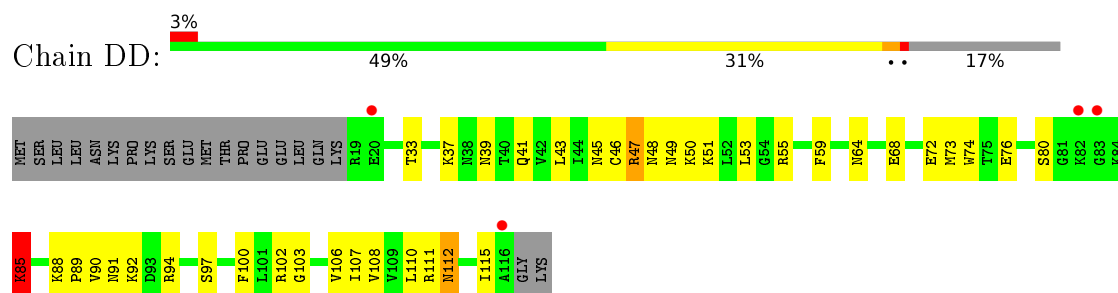
- Molecule 4: Small nuclear ribonucleoprotein Sm D2



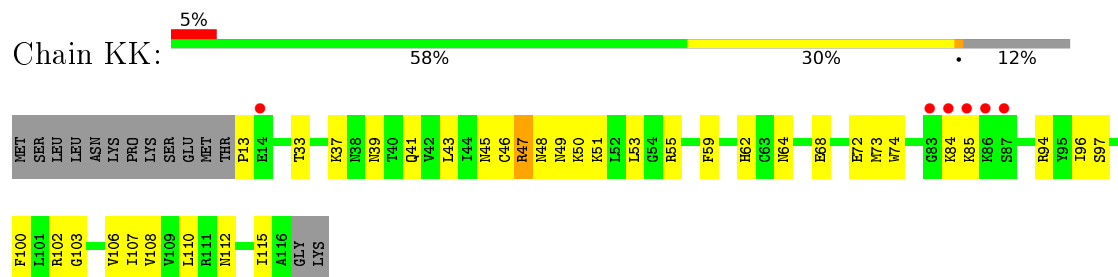
- Molecule 4: Small nuclear ribonucleoprotein Sm D2



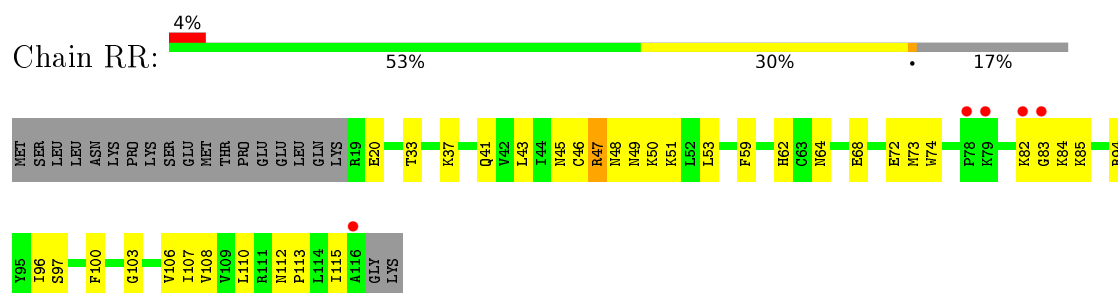
- Molecule 4: Small nuclear ribonucleoprotein Sm D2



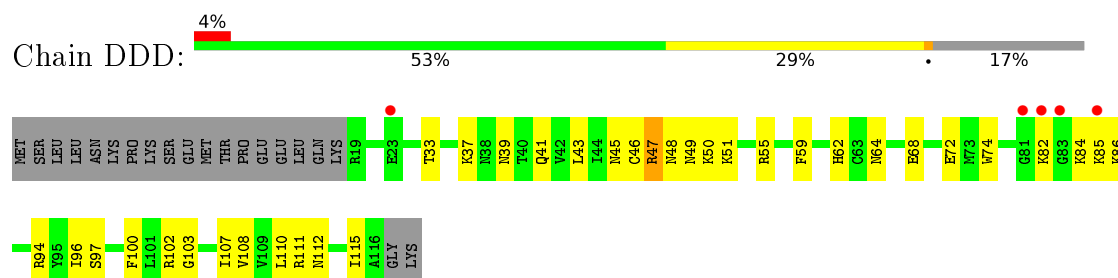
- Molecule 4: Small nuclear ribonucleoprotein Sm D2



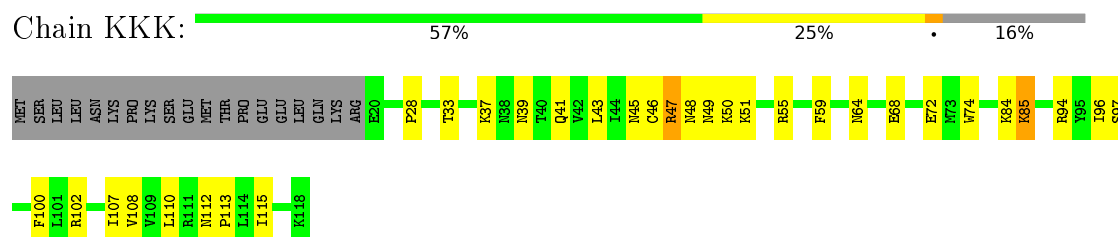
- Molecule 4: Small nuclear ribonucleoprotein Sm D2



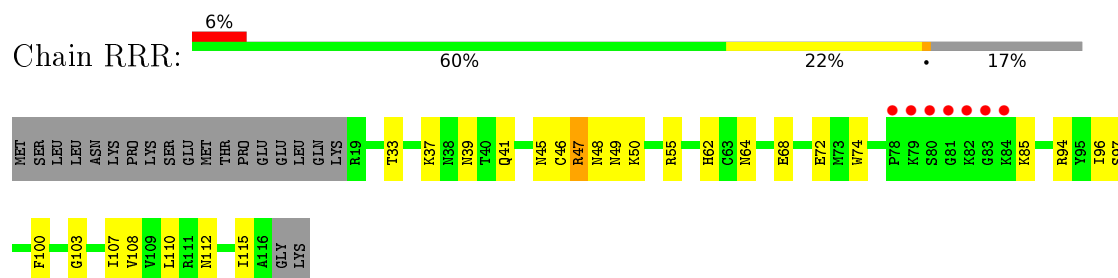
- Molecule 4: Small nuclear ribonucleoprotein Sm D2



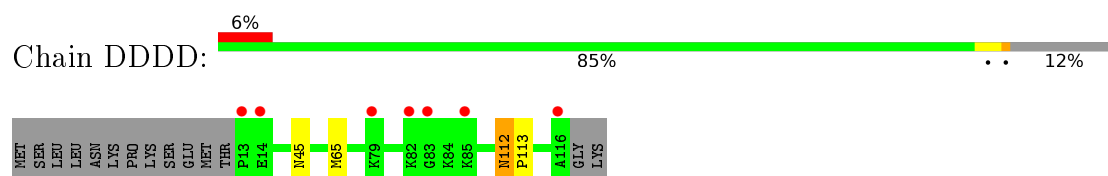
- Molecule 4: Small nuclear ribonucleoprotein Sm D2



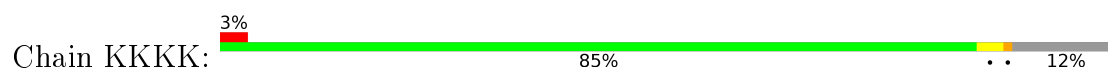
- Molecule 4: Small nuclear ribonucleoprotein Sm D2

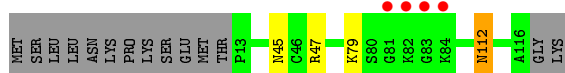


- Molecule 4: Small nuclear ribonucleoprotein Sm D2

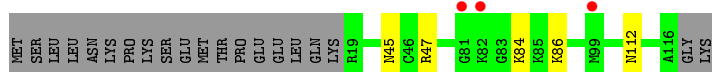
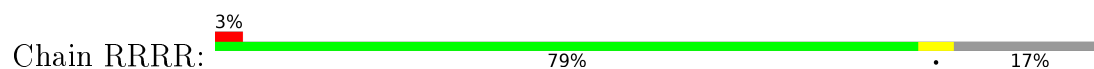


- Molecule 4: Small nuclear ribonucleoprotein Sm D2





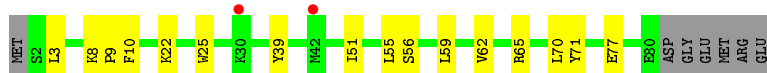
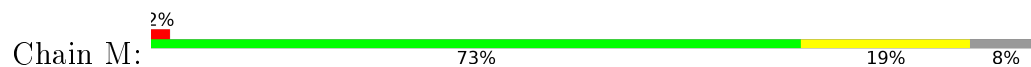
- Molecule 4: Small nuclear ribonucleoprotein Sm D2



- Molecule 5: Small nuclear ribonucleoprotein F



- Molecule 5: Small nuclear ribonucleoprotein F



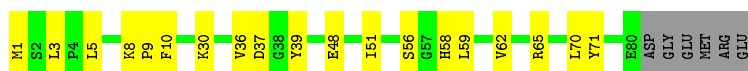
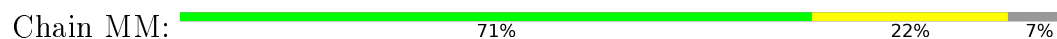
- Molecule 5: Small nuclear ribonucleoprotein F



- Molecule 5: Small nuclear ribonucleoprotein F

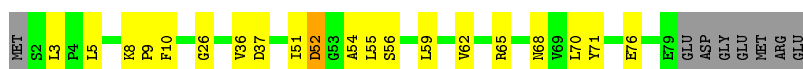


- Molecule 5: Small nuclear ribonucleoprotein F



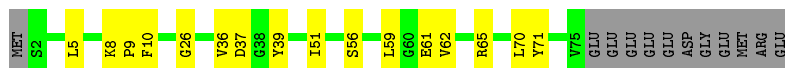
- Molecule 5: Small nuclear ribonucleoprotein F





- Molecule 5: Small nuclear ribonucleoprotein F

Chain FFF: 67% 19% 14%



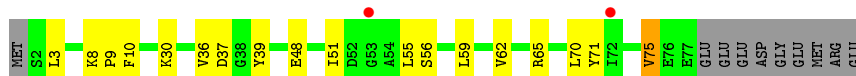
- Molecule 5: Small nuclear ribonucleoprotein F

Chain MMM: 72% 16% 12%



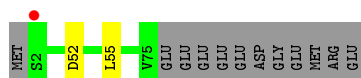
- Molecule 5: Small nuclear ribonucleoprotein F

Chain TTT: 2% 67% 20% 12%



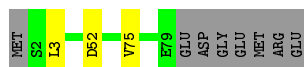
- Molecule 5: Small nuclear ribonucleoprotein F

Chain FFFF: 84% 14% 2%



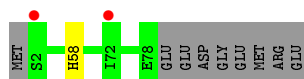
- Molecule 5: Small nuclear ribonucleoprotein F

Chain MMMM: 87% 9% 4%



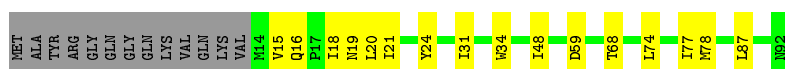
- Molecule 5: Small nuclear ribonucleoprotein F

Chain TTTT: 2% 88% 10%



- Molecule 6: Small nuclear ribonucleoprotein E

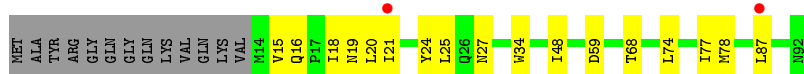
Chain E: 68% 17% 14%



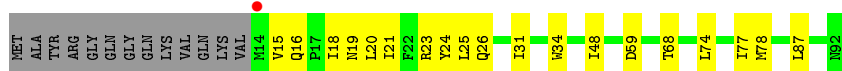
- Molecule 6: Small nuclear ribonucleoprotein E



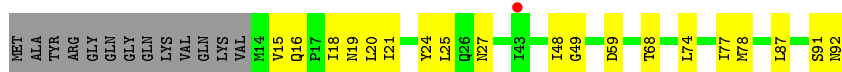
- Molecule 6: Small nuclear ribonucleoprotein E



- Molecule 6: Small nuclear ribonucleoprotein E



- Molecule 6: Small nuclear ribonucleoprotein E



- Molecule 6: Small nuclear ribonucleoprotein E



- Molecule 6: Small nuclear ribonucleoprotein E

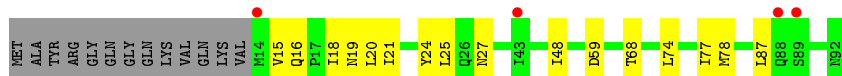
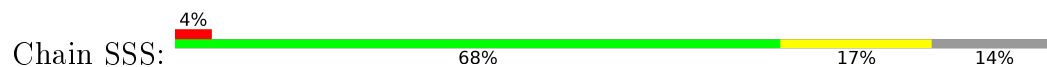


- Molecule 6: Small nuclear ribonucleoprotein E

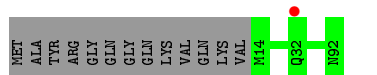
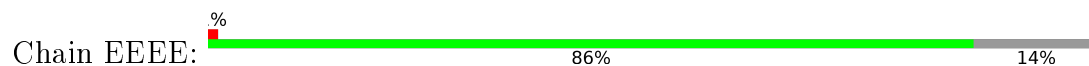




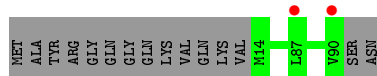
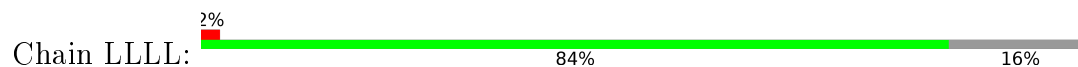
- Molecule 6: Small nuclear ribonucleoprotein E



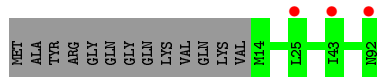
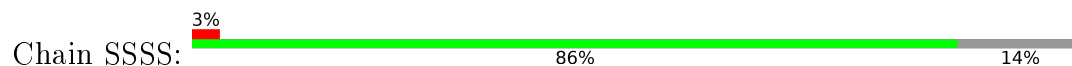
- Molecule 6: Small nuclear ribonucleoprotein E



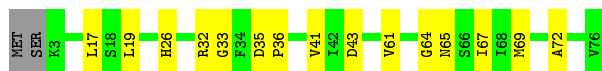
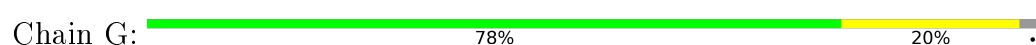
- Molecule 6: Small nuclear ribonucleoprotein E



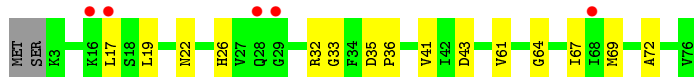
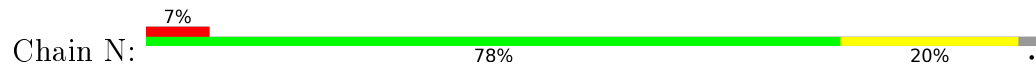
- Molecule 6: Small nuclear ribonucleoprotein E



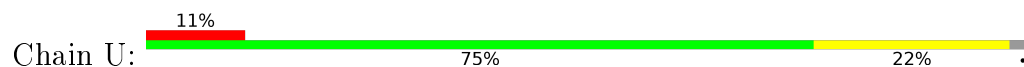
- Molecule 7: Small nuclear ribonucleoprotein G



- Molecule 7: Small nuclear ribonucleoprotein G

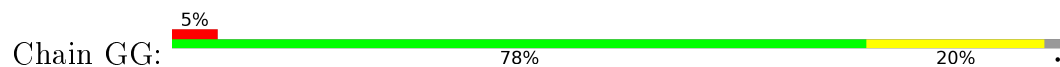


- Molecule 7: Small nuclear ribonucleoprotein G

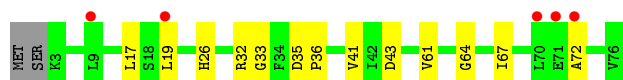
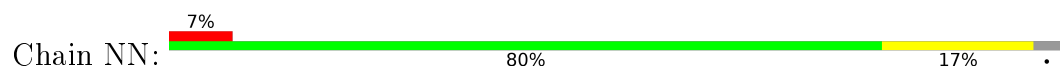




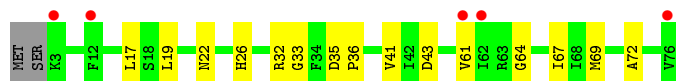
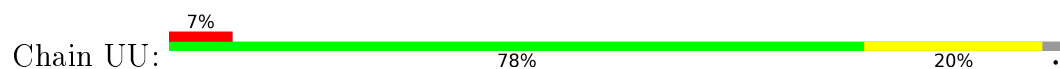
- Molecule 7: Small nuclear ribonucleoprotein G



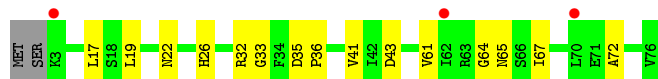
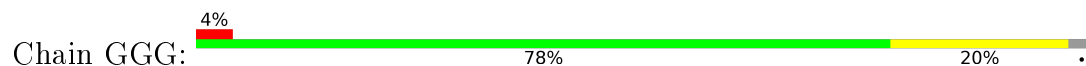
- Molecule 7: Small nuclear ribonucleoprotein G



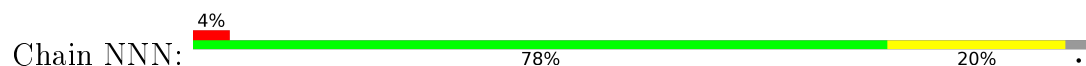
- Molecule 7: Small nuclear ribonucleoprotein G



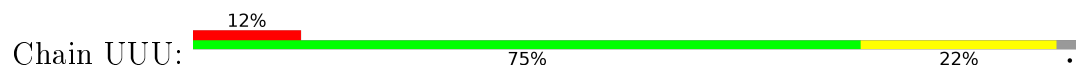
- Molecule 7: Small nuclear ribonucleoprotein G



- Molecule 7: Small nuclear ribonucleoprotein G

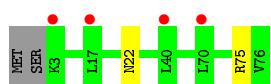


- Molecule 7: Small nuclear ribonucleoprotein G

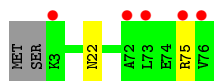


- Molecule 7: Small nuclear ribonucleoprotein G





- Molecule 7: Small nuclear ribonucleoprotein G



- Molecule 7: Small nuclear ribonucleoprotein G



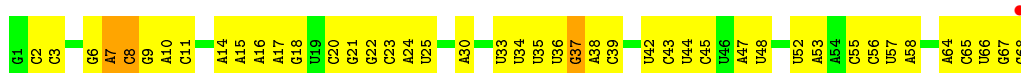
- Molecule 8: U4 small nuclear RNA variant: Native sequence 85-145, of which nucleotides 97-104 are replaced with GAAA tetraloop and nucleotides 134-137 are replaced with GAAA tetraloop receptor.



- Molecule 8: U4 small nuclear RNA variant: Native sequence 85-145, of which nucleotides 97-104 are replaced with GAAA tetraloop and nucleotides 134-137 are replaced with GAAA tetraloop receptor.



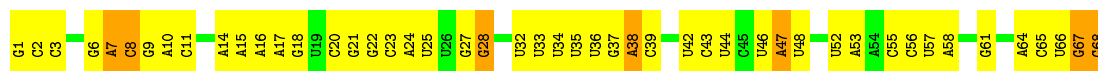
- Molecule 8: U4 small nuclear RNA variant: Native sequence 85-145, of which nucleotides 97-104 are replaced with GAAA tetraloop and nucleotides 134-137 are replaced with GAAA tetraloop receptor.



- Molecule 8: U4 small nuclear RNA variant: Native sequence 85-145, of which nucleotides 97-104 are replaced with GAAA tetraloop and nucleotides 134-137 are replaced with GAAA tetraloop receptor.

receptor.

Chain VV: 



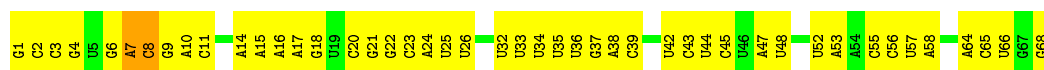
- Molecule 8: U4 small nuclear RNA variant: Native sequence 85-145, of which nucleotides 97-104 are replaced with GAAA tetraloop and nucleotides 134-137 are replaced with GAAA tetraloop receptor.

Chain XX: 



- Molecule 8: U4 small nuclear RNA variant: Native sequence 85-145, of which nucleotides 97-104 are replaced with GAAA tetraloop and nucleotides 134-137 are replaced with GAAA tetraloop receptor.

Chain YY: 



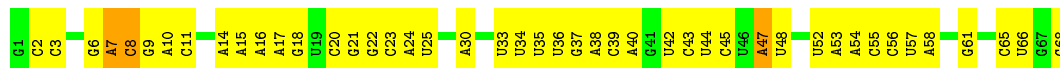
- Molecule 8: U4 small nuclear RNA variant: Native sequence 85-145, of which nucleotides 97-104 are replaced with GAAA tetraloop and nucleotides 134-137 are replaced with GAAA tetraloop receptor.

Chain VVV: 



- Molecule 8: U4 small nuclear RNA variant: Native sequence 85-145, of which nucleotides 97-104 are replaced with GAAA tetraloop and nucleotides 134-137 are replaced with GAAA tetraloop receptor.

Chain XXX: 



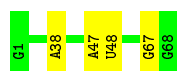
- Molecule 8: U4 small nuclear RNA variant: Native sequence 85-145, of which nucleotides 97-104 are replaced with GAAA tetraloop and nucleotides 134-137 are replaced with GAAA tetraloop receptor.

Chain YYY: 



- Molecule 8: U4 small nuclear RNA variant: Native sequence 85-145, of which nucleotides 97-104 are replaced with GAAA tetraloop and nucleotides 134-137 are replaced with GAAA tetraloop receptor.

Chain VVVV: 94% 6%



- Molecule 8: U4 small nuclear RNA variant: Native sequence 85-145, of which nucleotides 97-104 are replaced with GAAA tetraloop and nucleotides 134-137 are replaced with GAAA tetraloop receptor.

Chain XXXX: 90% 9%



- Molecule 8: U4 small nuclear RNA variant: Native sequence 85-145, of which nucleotides 97-104 are replaced with GAAA tetraloop and nucleotides 134-137 are replaced with GAAA tetraloop receptor.

Chain YYYY: 93% 7%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	248.01Å 248.01Å 251.94Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	66.15 – 3.60 66.15 – 3.48	Depositor EDS
% Data completeness (in resolution range)	83.1 (66.15-3.60) 75.9 (66.15-3.48)	Depositor EDS
$R_{merge}$	0.21	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.20 (at 3.49Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, $R_{free}$	0.177 , 0.224 0.179 , 0.176	Depositor DCC
$R_{free}$ test set	8363 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	84.3	Xtriage
Anisotropy	0.018	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 35.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.32$ , $\langle L^2 \rangle = 0.15$	Xtriage
Estimated twinning fraction	0.309 for -h,-k,l 0.306 for h,-h-k,-l 0.306 for -k,-h,-l	Xtriage
Reported twinning fraction	0.218 for H, K, L 0.282 for -K, -H, -L 0.283 for K, H, -L 0.216 for -H, -K, L	Depositor
Outliers	0 of 169321 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	71485	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	128.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.59% 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 ⓘ

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.46	0/660	0.67	0/889
1	AA	0.45	0/660	0.69	0/889
1	AAA	0.46	0/656	0.71	1/885 (0.1%)
1	AAAA	0.49	0/654	0.69	0/881
1	H	0.50	0/666	0.71	0/897
1	HH	0.49	0/661	0.68	0/892
1	HHH	0.49	0/666	0.67	0/897
1	HHHH	0.45	0/651	0.71	1/878 (0.1%)
1	O	0.45	0/660	0.68	0/889
1	OO	0.44	0/645	0.66	0/870
1	OOO	0.46	0/654	0.67	0/881
1	OOOO	0.44	0/654	0.69	0/881
2	B	0.53	0/700	0.82	0/933
2	BB	0.56	0/573	0.81	1/765 (0.1%)
2	BBB	0.49	0/573	0.73	0/765
2	BBBB	0.50	0/600	0.84	0/799
2	I	0.56	0/582	0.83	2/776 (0.3%)
2	II	0.49	0/593	0.77	0/793
2	III	0.61	2/573 (0.3%)	0.76	1/765 (0.1%)
2	IIII	0.61	0/610	0.84	1/813 (0.1%)
2	P	0.46	0/577	0.74	0/769
2	PP	0.49	0/577	0.69	0/769
2	PPP	0.49	0/567	0.92	3/758 (0.4%)
2	PPPP	0.46	0/607	0.73	0/810
3	C	0.56	0/657	0.77	0/888
3	CC	0.55	0/657	0.78	0/888
3	CCC	0.56	0/657	0.76	0/888
3	CCCC	0.54	0/657	0.76	0/888
3	J	0.54	0/657	0.76	0/888
3	JJ	0.53	0/657	0.76	0/888
3	JJJ	0.53	0/657	0.75	0/888
3	JJJJ	0.54	0/657	0.78	0/888
3	Q	0.52	0/657	0.74	0/888
3	QQ	0.53	0/657	0.76	0/888

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
3	QQQ	0.56	0/657	0.77	0/888
3	QQQQ	0.51	0/657	0.75	0/888
4	D	0.69	0/786	0.86	0/1053
4	DD	0.66	0/793	0.88	1/1063 (0.1%)
4	DDD	0.62	0/797	0.86	1/1067 (0.1%)
4	DDDD	0.62	0/849	0.88	1/1136 (0.1%)
4	K	0.69	0/806	0.88	4/1079 (0.4%)
4	KK	0.65	0/849	0.84	1/1136 (0.1%)
4	KKK	0.63	0/796	0.86	2/1064 (0.2%)
4	KKKK	0.76	0/849	0.93	2/1136 (0.2%)
4	R	0.63	0/797	0.85	1/1067 (0.1%)
4	RR	0.68	1/797 (0.1%)	0.89	2/1067 (0.2%)
4	RRR	0.63	0/797	0.84	1/1067 (0.1%)
4	RRRR	0.66	0/797	0.89	2/1067 (0.2%)
5	F	0.81	0/588	0.83	0/795
5	FF	0.72	0/597	0.80	0/807
5	FFF	0.69	0/588	0.82	0/795
5	FFFF	0.70	0/588	0.85	2/795 (0.3%)
5	M	0.73	0/621	0.81	0/840
5	MM	0.74	0/633	0.83	0/855
5	MMM	0.68	0/602	0.80	0/814
5	MMMM	0.76	0/616	0.84	0/833
5	T	0.72	0/606	0.80	0/819
5	TT	0.71	0/608	0.82	0/823
5	TTT	0.74	0/602	0.77	0/814
5	TTTT	0.69	0/611	0.79	0/826
6	E	0.61	0/660	0.83	0/886
6	EE	0.60	0/660	0.86	0/886
6	EEE	0.63	0/660	0.86	0/886
6	EEEE	0.62	0/660	0.86	0/886
6	L	0.59	0/660	0.86	1/886 (0.1%)
6	LL	0.57	0/660	0.82	0/886
6	LLL	0.58	0/660	0.82	0/886
6	LLLL	0.66	0/646	0.85	0/867
6	S	0.60	0/660	0.83	0/886
6	SS	0.62	0/660	0.84	0/886
6	SSS	0.57	0/660	0.83	0/886
6	SSSS	0.59	0/660	0.80	0/886
7	G	0.52	0/584	0.78	0/779
7	GG	0.54	0/584	0.78	0/779
7	GGG	0.49	0/584	0.75	0/779
7	GGGG	0.50	0/584	0.79	1/779 (0.1%)
7	N	0.50	0/584	0.77	0/779

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
7	NN	0.52	0/584	0.79	0/779
7	NNN	0.49	0/584	0.83	2/779 (0.3%)
7	NNNN	0.55	0/584	0.82	2/779 (0.3%)
7	U	0.49	0/584	0.79	2/779 (0.3%)
7	UU	0.48	0/578	0.74	0/772
7	UUU	0.53	0/584	0.81	1/779 (0.1%)
7	UUUU	0.49	0/584	0.76	0/779
8	V	0.59	1/1626 (0.1%)	0.80	3/2534 (0.1%)
8	VV	0.53	0/1626	0.79	7/2534 (0.3%)
8	VVV	0.46	0/1626	0.80	5/2534 (0.2%)
8	VVVV	0.53	0/1626	0.76	4/2534 (0.2%)
8	X	0.58	1/1626 (0.1%)	0.83	6/2534 (0.2%)
8	XX	0.46	0/1626	0.81	8/2534 (0.3%)
8	XXX	0.42	0/1626	0.76	4/2534 (0.2%)
8	XXXX	0.51	0/1626	0.99	12/2534 (0.5%)
8	Y	0.46	0/1626	0.78	6/2534 (0.2%)
8	YY	0.43	0/1626	0.75	4/2534 (0.2%)
8	YYY	0.41	0/1626	0.75	4/2534 (0.2%)
8	YYYY	0.48	0/1626	0.77	4/2534 (0.2%)
All	All	0.56	5/74296 (0.0%)	0.80	106/103980 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	PPPP	0	1
4	D	0	1
4	DD	0	1
4	DDD	0	1
4	DDDD	0	1
4	K	0	1
4	KK	0	1
4	KKK	0	1
4	KKKK	0	1
4	R	0	1
4	RR	0	1
4	RRR	0	1
4	RRRR	0	1
5	M	0	1
5	MM	0	2

*Continued on next page...*

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
5	MMM	0	1
5	MMMM	0	1
5	TT	0	2
5	TTT	0	1
8	V	0	1
8	VVVV	0	1
All	All	0	23

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	RR	20	GLU	CB-CG	6.88	1.65	1.52
8	X	8	C	O3'-P	-6.72	1.53	1.61
2	III	47	GLU	CD-OE1	5.94	1.32	1.25
8	V	7	A	O3'-P	-5.39	1.54	1.61
2	III	47	GLU	CD-OE2	-5.10	1.20	1.25

The worst 5 of 106 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	XXXX	48	U	O5'-P-OP2	-23.61	82.36	110.70
2	PPP	25	ARG	NE-CZ-NH2	-11.55	114.52	120.30
8	XXXX	18	G	O5'-P-OP2	-11.18	95.64	105.70
2	PPP	25	ARG	NE-CZ-NH1	10.24	125.42	120.30
8	XXX	47	A	C2'-C3'-O3'	-9.57	88.44	109.50

There are no chirality outliers.

5 of 23 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	D	112	ASN	Peptide
4	DD	112	ASN	Peptide
4	K	112	ASN	Peptide
5	M	3	LEU	Peptide
4	R	112	ASN	Peptide

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within

the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	652	0	670	10	0
1	AA	652	0	670	9	0
1	AAA	648	0	1324	10	0
1	AAAA	646	0	0	0	0
1	H	658	0	675	10	0
1	HH	653	0	661	9	0
1	HHH	658	0	1332	8	0
1	HHHH	643	0	0	0	0
1	O	652	0	670	8	0
1	OO	637	0	652	10	0
1	OOO	646	0	1330	9	0
1	OOOO	646	0	0	0	0
2	B	690	0	712	13	0
2	BB	565	0	575	6	0
2	BBB	565	0	1151	4	0
2	BBBB	592	0	0	0	0
2	I	574	0	588	8	0
2	II	585	0	586	5	0
2	III	565	0	1151	5	0
2	IIII	602	0	0	0	0
2	P	569	0	586	4	0
2	PP	569	0	586	5	0
2	PPP	559	0	1139	4	0
2	PPPP	598	0	0	0	0
3	C	649	0	693	11	0
3	CC	649	0	693	9	0
3	CCC	649	0	1386	9	0
3	CCCC	649	0	0	0	0
3	J	649	0	693	9	0
3	JJ	649	0	693	7	0
3	JJJ	649	0	1386	9	0
3	JJJJ	649	0	0	0	0
3	Q	649	0	693	9	0
3	QQ	649	0	693	11	0
3	QQQ	649	0	1386	7	0
3	QQQQ	649	0	0	0	0
4	D	776	0	819	38	0
4	DD	783	0	821	44	0
4	DDD	787	0	1665	38	0
4	DDDD	838	0	0	0	1
4	K	796	0	837	32	0
4	KK	838	0	884	33	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	KKK	786	0	1651	27	1
4	KKKK	838	0	0	0	0
4	R	787	0	832	31	0
4	RR	787	0	832	31	0
4	RRR	787	0	1664	23	0
4	RRRR	787	0	0	0	0
5	F	576	0	589	21	0
5	FF	585	0	595	15	0
5	FFF	576	0	1178	18	0
5	FFFF	576	0	0	0	0
5	M	609	0	607	16	0
5	MM	621	0	623	15	1
5	MMM	590	0	1198	17	0
5	MMMM	604	0	0	0	0
5	T	594	0	601	25	0
5	TT	596	0	597	32	0
5	TTT	590	0	1194	15	0
5	TTTT	599	0	0	0	1
6	E	652	0	668	17	0
6	EE	652	0	668	20	0
6	EEE	652	0	1336	19	0
6	EEEE	652	0	0	0	0
6	L	652	0	668	19	0
6	LL	652	0	668	23	0
6	LLL	652	0	1325	19	0
6	LLLL	638	0	0	0	0
6	S	652	0	668	18	0
6	SS	652	0	668	25	0
6	SSS	652	0	1336	17	0
6	SSSS	652	0	0	0	0
7	G	577	0	603	12	0
7	GG	577	0	603	13	0
7	GGG	577	0	1206	13	0
7	GGGG	577	0	0	0	0
7	N	577	0	603	12	0
7	NN	577	0	603	11	0
7	NNN	577	0	1206	13	0
7	NNNN	577	0	0	0	0
7	U	577	0	603	12	0
7	UU	571	0	592	13	0
7	UUU	577	0	1206	12	0
7	UUUU	577	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	V	1453	0	733	48	0
8	VV	1453	0	733	45	0
8	VVV	1453	0	1466	51	0
8	VVVV	1453	0	0	0	0
8	X	1453	0	733	53	0
8	XX	1453	0	733	43	0
8	XXX	1453	0	1466	50	0
8	XXXX	1453	0	0	0	0
8	Y	1453	0	733	43	0
8	YY	1453	0	733	46	0
8	YYY	1453	0	1466	40	0
8	YYYY	1453	0	0	0	0
9	DD	1	0	0	0	0
9	R	1	0	0	0	0
9	RR	1	0	0	0	0
All	All	71485	0	64587	1106	2

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 11.

The worst 5 of 1106 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)
8:V:37:G:H2'	8:V:38:A:H5''	1.37	1.07
8:Y:37:G:H2'	8:Y:38:A:H5''	1.36	1.06
8:XX:37:G:H2'	8:XX:38:A:H5''	1.37	1.06
8:YYY:37:G:H2'	8:YYY:38:A:H5''	1.38	1.05
8:VVV:37:G:H2'	8:VVV:38:A:H5''	1.43	1.02

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:MM:58:HIS:NE2	4:KKK:113:PRO:O[2_665]	2.06	0.14
4:DDDD:113:PRO:O	5:TTTT:58:HIS:NE2[3_565]	2.07	0.13

## 5.3 Torsion angles

### 5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	81/125 (65%)	80 (99%)	1 (1%)	0	100	100
1	AA	81/125 (65%)	78 (96%)	3 (4%)	0	100	100
1	AAA	81/125 (65%)	80 (99%)	1 (1%)	0	100	100
1	AAAA	80/125 (64%)	78 (98%)	2 (2%)	0	100	100
1	H	82/125 (66%)	81 (99%)	1 (1%)	0	100	100
1	HH	82/125 (66%)	78 (95%)	4 (5%)	0	100	100
1	HHH	82/125 (66%)	81 (99%)	1 (1%)	0	100	100
1	HHHH	80/125 (64%)	78 (98%)	2 (2%)	0	100	100
1	O	81/125 (65%)	80 (99%)	1 (1%)	0	100	100
1	OO	79/125 (63%)	78 (99%)	1 (1%)	0	100	100
1	OOO	80/125 (64%)	78 (98%)	2 (2%)	0	100	100
1	OOOO	80/125 (64%)	79 (99%)	1 (1%)	0	100	100
2	B	84/95 (88%)	82 (98%)	2 (2%)	0	100	100
2	BB	67/95 (70%)	66 (98%)	1 (2%)	0	100	100
2	BBB	67/95 (70%)	66 (98%)	1 (2%)	0	100	100
2	BBBB	70/95 (74%)	69 (99%)	1 (1%)	0	100	100
2	I	68/95 (72%)	66 (97%)	2 (3%)	0	100	100
2	II	71/95 (75%)	70 (99%)	1 (1%)	0	100	100
2	III	67/95 (70%)	65 (97%)	2 (3%)	0	100	100
2	IIII	71/95 (75%)	70 (99%)	1 (1%)	0	100	100
2	P	67/95 (70%)	65 (97%)	2 (3%)	0	100	100
2	PP	67/95 (70%)	65 (97%)	2 (3%)	0	100	100
2	PPP	67/95 (70%)	66 (98%)	1 (2%)	0	100	100
2	PPPP	71/95 (75%)	69 (97%)	2 (3%)	0	100	100
3	C	80/118 (68%)	77 (96%)	3 (4%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	CC	80/118 (68%)	77 (96%)	3 (4%)	0	100	100
3	CCC	80/118 (68%)	77 (96%)	3 (4%)	0	100	100
3	CCCC	80/118 (68%)	78 (98%)	2 (2%)	0	100	100
3	J	80/118 (68%)	77 (96%)	3 (4%)	0	100	100
3	JJ	80/118 (68%)	77 (96%)	3 (4%)	0	100	100
3	JJJ	80/118 (68%)	77 (96%)	3 (4%)	0	100	100
3	JJJJ	80/118 (68%)	77 (96%)	3 (4%)	0	100	100
3	Q	80/118 (68%)	77 (96%)	3 (4%)	0	100	100
3	QQ	80/118 (68%)	78 (98%)	2 (2%)	0	100	100
3	QQQ	80/118 (68%)	78 (98%)	2 (2%)	0	100	100
3	QQQQ	80/118 (68%)	78 (98%)	2 (2%)	0	100	100
4	D	95/118 (80%)	91 (96%)	4 (4%)	0	100	100
4	DD	96/118 (81%)	92 (96%)	3 (3%)	1 (1%)	19	66
4	DDD	96/118 (81%)	92 (96%)	4 (4%)	0	100	100
4	DDDD	102/118 (86%)	98 (96%)	4 (4%)	0	100	100
4	K	98/118 (83%)	93 (95%)	4 (4%)	1 (1%)	19	66
4	KK	102/118 (86%)	98 (96%)	4 (4%)	0	100	100
4	KKK	97/118 (82%)	93 (96%)	4 (4%)	0	100	100
4	KKKK	102/118 (86%)	98 (96%)	4 (4%)	0	100	100
4	R	96/118 (81%)	92 (96%)	4 (4%)	0	100	100
4	RR	96/118 (81%)	90 (94%)	6 (6%)	0	100	100
4	RRR	96/118 (81%)	92 (96%)	4 (4%)	0	100	100
4	RRRR	96/118 (81%)	92 (96%)	4 (4%)	0	100	100
5	F	72/86 (84%)	68 (94%)	4 (6%)	0	100	100
5	FF	73/86 (85%)	69 (94%)	4 (6%)	0	100	100
5	FFF	72/86 (84%)	69 (96%)	3 (4%)	0	100	100
5	FFFF	72/86 (84%)	69 (96%)	3 (4%)	0	100	100
5	M	77/86 (90%)	73 (95%)	4 (5%)	0	100	100
5	MM	78/86 (91%)	74 (95%)	4 (5%)	0	100	100
5	MMM	74/86 (86%)	69 (93%)	5 (7%)	0	100	100
5	MMMM	76/86 (88%)	69 (91%)	6 (8%)	1 (1%)	15	60

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
5	T	74/86 (86%)	70 (95%)	4 (5%)	0	100	100
5	TT	76/86 (88%)	73 (96%)	3 (4%)	0	100	100
5	TTT	74/86 (86%)	70 (95%)	3 (4%)	1 (1%)	14	59
5	TTTT	75/86 (87%)	71 (95%)	4 (5%)	0	100	100
6	E	77/92 (84%)	76 (99%)	1 (1%)	0	100	100
6	EE	77/92 (84%)	75 (97%)	2 (3%)	0	100	100
6	EEE	77/92 (84%)	75 (97%)	2 (3%)	0	100	100
6	EEEE	77/92 (84%)	75 (97%)	2 (3%)	0	100	100
6	L	77/92 (84%)	74 (96%)	3 (4%)	0	100	100
6	LL	77/92 (84%)	75 (97%)	2 (3%)	0	100	100
6	LLL	77/92 (84%)	75 (97%)	2 (3%)	0	100	100
6	LLLL	75/92 (82%)	74 (99%)	1 (1%)	0	100	100
6	S	77/92 (84%)	75 (97%)	2 (3%)	0	100	100
6	SS	77/92 (84%)	75 (97%)	2 (3%)	0	100	100
6	SSS	77/92 (84%)	75 (97%)	2 (3%)	0	100	100
6	SSSS	77/92 (84%)	75 (97%)	2 (3%)	0	100	100
7	G	72/76 (95%)	70 (97%)	2 (3%)	0	100	100
7	GG	72/76 (95%)	70 (97%)	2 (3%)	0	100	100
7	GGG	72/76 (95%)	70 (97%)	2 (3%)	0	100	100
7	GGGG	72/76 (95%)	70 (97%)	2 (3%)	0	100	100
7	N	72/76 (95%)	70 (97%)	2 (3%)	0	100	100
7	NN	72/76 (95%)	70 (97%)	2 (3%)	0	100	100
7	NNN	72/76 (95%)	70 (97%)	2 (3%)	0	100	100
7	NNNN	72/76 (95%)	70 (97%)	2 (3%)	0	100	100
7	U	72/76 (95%)	70 (97%)	2 (3%)	0	100	100
7	UU	72/76 (95%)	70 (97%)	2 (3%)	0	100	100
7	UUU	72/76 (95%)	70 (97%)	2 (3%)	0	100	100
7	UUUU	72/76 (95%)	70 (97%)	2 (3%)	0	100	100
All	All	6617/8520 (78%)	6400 (97%)	213 (3%)	4 (0%)	56	90

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	DD	85	LYS
5	TTT	75	VAL
4	K	117	GLY
5	MMMM	75	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	A	73/100 (73%)	73 (100%)	0	100	100
1	AA	73/100 (73%)	73 (100%)	0	100	100
1	AAA	72/100 (72%)	72 (100%)	0	100	100
1	AAAA	72/100 (72%)	72 (100%)	0	100	100
1	H	74/100 (74%)	73 (99%)	1 (1%)	74	91
1	HH	72/100 (72%)	72 (100%)	0	100	100
1	HHH	74/100 (74%)	74 (100%)	0	100	100
1	HHHH	72/100 (72%)	72 (100%)	0	100	100
1	O	73/100 (73%)	73 (100%)	0	100	100
1	OO	71/100 (71%)	71 (100%)	0	100	100
1	OOO	72/100 (72%)	72 (100%)	0	100	100
1	OOOO	72/100 (72%)	72 (100%)	0	100	100
2	B	77/85 (91%)	75 (97%)	2 (3%)	54	83
2	BB	63/85 (74%)	62 (98%)	1 (2%)	70	90
2	BBB	63/85 (74%)	61 (97%)	2 (3%)	46	81
2	BBBB	66/85 (78%)	65 (98%)	1 (2%)	72	90
2	I	64/85 (75%)	62 (97%)	2 (3%)	47	81
2	II	63/85 (74%)	62 (98%)	1 (2%)	70	90
2	III	63/85 (74%)	62 (98%)	1 (2%)	70	90
2	IIII	67/85 (79%)	66 (98%)	1 (2%)	72	90
2	P	64/85 (75%)	63 (98%)	1 (2%)	70	90

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	PP	64/85 (75%)	63 (98%)	1 (2%)	70	90
2	PPP	62/85 (73%)	62 (100%)	0	100	100
2	PPPP	67/85 (79%)	65 (97%)	2 (3%)	48	82
3	C	77/100 (77%)	75 (97%)	2 (3%)	54	83
3	CC	77/100 (77%)	75 (97%)	2 (3%)	54	83
3	CCC	77/100 (77%)	75 (97%)	2 (3%)	54	83
3	CCCC	77/100 (77%)	75 (97%)	2 (3%)	54	83
3	J	77/100 (77%)	75 (97%)	2 (3%)	54	83
3	JJ	77/100 (77%)	75 (97%)	2 (3%)	54	83
3	JJJ	77/100 (77%)	75 (97%)	2 (3%)	54	83
3	JJJJ	77/100 (77%)	75 (97%)	2 (3%)	54	83
3	Q	77/100 (77%)	75 (97%)	2 (3%)	54	83
3	QQ	77/100 (77%)	75 (97%)	2 (3%)	54	83
3	QQQ	77/100 (77%)	75 (97%)	2 (3%)	54	83
3	QQQQ	77/100 (77%)	75 (97%)	2 (3%)	54	83
4	D	90/110 (82%)	88 (98%)	2 (2%)	60	86
4	DD	90/110 (82%)	85 (94%)	5 (6%)	26	68
4	DDD	91/110 (83%)	90 (99%)	1 (1%)	80	92
4	DDDD	97/110 (88%)	95 (98%)	2 (2%)	61	87
4	K	91/110 (83%)	89 (98%)	2 (2%)	60	86
4	KK	97/110 (88%)	96 (99%)	1 (1%)	82	93
4	KKK	90/110 (82%)	89 (99%)	1 (1%)	80	92
4	KKKK	97/110 (88%)	95 (98%)	2 (2%)	61	87
4	R	91/110 (83%)	90 (99%)	1 (1%)	80	92
4	RR	91/110 (83%)	90 (99%)	1 (1%)	80	92
4	RRR	91/110 (83%)	90 (99%)	1 (1%)	80	92
4	RRRR	91/110 (83%)	89 (98%)	2 (2%)	60	86
5	F	63/74 (85%)	62 (98%)	1 (2%)	70	90
5	FF	64/74 (86%)	64 (100%)	0	100	100
5	FFF	63/74 (85%)	63 (100%)	0	100	100
5	FFFF	63/74 (85%)	62 (98%)	1 (2%)	70	90

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	M	65/74 (88%)	64 (98%)	1 (2%)	72	90
5	MM	67/74 (90%)	67 (100%)	0	100	100
5	MMM	64/74 (86%)	64 (100%)	0	100	100
5	MMMM	65/74 (88%)	64 (98%)	1 (2%)	72	90
5	T	65/74 (88%)	65 (100%)	0	100	100
5	TT	63/74 (85%)	62 (98%)	1 (2%)	70	90
5	TTT	64/74 (86%)	63 (98%)	1 (2%)	70	90
5	TTTT	65/74 (88%)	65 (100%)	0	100	100
6	E	74/84 (88%)	74 (100%)	0	100	100
6	EE	74/84 (88%)	74 (100%)	0	100	100
6	EEE	74/84 (88%)	74 (100%)	0	100	100
6	EEEE	74/84 (88%)	74 (100%)	0	100	100
6	L	74/84 (88%)	74 (100%)	0	100	100
6	LL	74/84 (88%)	74 (100%)	0	100	100
6	LLL	74/84 (88%)	74 (100%)	0	100	100
6	LLLL	72/84 (86%)	72 (100%)	0	100	100
6	S	74/84 (88%)	74 (100%)	0	100	100
6	SS	74/84 (88%)	74 (100%)	0	100	100
6	SSS	74/84 (88%)	74 (100%)	0	100	100
6	SSSS	74/84 (88%)	74 (100%)	0	100	100
7	G	64/66 (97%)	63 (98%)	1 (2%)	70	90
7	GG	64/66 (97%)	63 (98%)	1 (2%)	70	90
7	GGG	64/66 (97%)	62 (97%)	2 (3%)	47	81
7	GGGG	64/66 (97%)	63 (98%)	1 (2%)	70	90
7	N	64/66 (97%)	63 (98%)	1 (2%)	70	90
7	NN	64/66 (97%)	64 (100%)	0	100	100
7	NNN	64/66 (97%)	64 (100%)	0	100	100
7	NNNN	64/66 (97%)	63 (98%)	1 (2%)	70	90
7	U	64/66 (97%)	62 (97%)	2 (3%)	47	81
7	UU	63/66 (96%)	62 (98%)	1 (2%)	70	90
7	UUU	64/66 (97%)	62 (97%)	2 (3%)	47	81

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
7	UUUU	64/66 (97%)	63 (98%)	1 (2%)	70 90
All	All	6108/7428 (82%)	6028 (99%)	80 (1%)	76 91

5 of 80 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	PP	13	ILE
3	CCC	40	LEU
2	PPPP	11	GLN
3	QQ	40	LEU
5	TT	52	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 47 such sidechains are listed below:

Mol	Chain	Res	Type
4	R	91	ASN
3	CC	64	ASN
6	SS	27	ASN
1	AA	60	GLN
4	DD	69	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
8	V	67/68 (98%)	0	0
8	VV	67/68 (98%)	2 (2%)	1 (1%)
8	VVV	67/68 (98%)	1 (1%)	0
8	VVVV	67/68 (98%)	0	0
8	X	67/68 (98%)	0	0
8	XX	67/68 (98%)	3 (4%)	2 (2%)
8	XXX	67/68 (98%)	1 (1%)	1 (1%)
8	XXXX	67/68 (98%)	2 (2%)	0
8	Y	67/68 (98%)	2 (2%)	0
8	YY	67/68 (98%)	1 (1%)	0
8	YYY	67/68 (98%)	2 (2%)	0
8	YYYY	67/68 (98%)	1 (1%)	0
All	All	804/816 (98%)	15 (1%)	4 (0%)

5 of 15 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
8	Y	8	C
8	Y	68	G
8	VV	8	C
8	VV	68	G
8	XX	8	C

All (4) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
8	VV	67	G
8	XX	47	A
8	XX	67	G
8	XXX	47	A

## 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](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	83/125 (66%)	0.26	4 (4%) 34 24	58, 143, 193, 230	0
1	AA	83/125 (66%)	0.09	1 (1%) 81 69	82, 130, 175, 210	0
1	AAA	83/125 (66%)	0.06	5 (6%) 25 17	63, 130, 180, 217	0
1	AAAA	82/125 (65%)	0.18	4 (4%) 33 23	56, 136, 187, 201	0
1	H	84/125 (67%)	0.32	6 (7%) 19 12	58, 144, 210, 248	0
1	HH	84/125 (67%)	0.17	4 (4%) 34 24	84, 139, 195, 231	0
1	HHH	84/125 (67%)	0.21	2 (2%) 62 47	70, 143, 195, 230	0
1	HHHH	82/125 (65%)	0.10	1 (1%) 81 69	71, 142, 192, 240	0
1	O	83/125 (66%)	0.69	14 (16%) 2 2	92, 164, 222, 306	0
1	OO	81/125 (64%)	0.25	6 (7%) 17 12	64, 143, 192, 213	0
1	OOO	82/125 (65%)	0.28	4 (4%) 33 23	74, 146, 192, 214	0
1	OOOO	82/125 (65%)	0.36	6 (7%) 18 12	50, 152, 202, 231	0
2	B	86/95 (90%)	0.31	7 (8%) 15 10	72, 139, 198, 223	0
2	BB	71/95 (74%)	0.25	1 (1%) 78 65	63, 125, 171, 191	0
2	BBB	71/95 (74%)	0.62	10 (14%) 4 3	78, 136, 190, 273	1 (1%)
2	BBBB	74/95 (77%)	0.21	2 (2%) 58 43	57, 128, 164, 201	0
2	I	72/95 (75%)	0.34	7 (9%) 10 7	79, 131, 178, 240	0
2	II	75/95 (78%)	0.32	6 (8%) 15 10	68, 126, 202, 302	0
2	III	71/95 (74%)	0.28	4 (5%) 28 19	60, 129, 183, 202	0
2	IIII	75/95 (78%)	0.46	8 (10%) 8 6	89, 154, 220, 250	0
2	P	71/95 (74%)	0.29	4 (5%) 28 19	82, 146, 184, 219	0
2	PP	71/95 (74%)	0.39	5 (7%) 19 13	84, 138, 195, 207	0
2	PPP	71/95 (74%)	0.16	4 (5%) 28 19	79, 137, 191, 204	0
2	PPPP	75/95 (78%)	0.54	9 (12%) 6 5	63, 135, 191, 222	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
3	C	82/118 (69%)	-0.02	0 100 100	70, 104, 135, 172	0
3	CC	82/118 (69%)	-0.05	1 (1%) 81 69	57, 91, 139, 162	0
3	CCC	82/118 (69%)	0.05	1 (1%) 81 69	43, 96, 151, 213	0
3	CCCC	82/118 (69%)	0.09	2 (2%) 62 47	56, 99, 159, 214	0
3	J	82/118 (69%)	-0.00	0 100 100	51, 102, 146, 168	0
3	JJ	82/118 (69%)	0.03	2 (2%) 62 47	56, 104, 168, 212	0
3	JJJ	82/118 (69%)	-0.07	1 (1%) 81 69	57, 101, 144, 178	0
3	JJJJ	82/118 (69%)	0.05	0 100 100	49, 91, 150, 191	0
3	Q	82/118 (69%)	0.13	1 (1%) 81 69	57, 108, 151, 275	0
3	QQ	82/118 (69%)	-0.02	0 100 100	46, 98, 152, 208	0
3	QQQ	82/118 (69%)	0.03	0 100 100	46, 103, 157, 174	0
3	QQQQ	82/118 (69%)	0.03	1 (1%) 81 69	64, 103, 178, 220	0
4	D	97/118 (82%)	0.00	3 (3%) 52 38	48, 93, 168, 229	0
4	DD	98/118 (83%)	0.13	4 (4%) 41 29	57, 106, 187, 244	0
4	DDD	98/118 (83%)	0.13	5 (5%) 32 22	34, 92, 211, 237	0
4	DDDD	104/118 (88%)	0.18	7 (6%) 21 13	36, 96, 184, 219	0
4	K	100/118 (84%)	0.02	3 (3%) 54 38	29, 99, 180, 200	0
4	KK	104/118 (88%)	0.09	6 (5%) 26 18	45, 106, 191, 249	0
4	KKK	99/118 (83%)	-0.04	0 100 100	59, 110, 173, 200	0
4	KKKK	104/118 (88%)	0.06	4 (3%) 44 32	47, 100, 198, 262	0
4	R	98/118 (83%)	0.04	3 (3%) 52 38	48, 103, 175, 212	0
4	RR	98/118 (83%)	0.06	5 (5%) 32 22	49, 100, 207, 227	0
4	RRR	98/118 (83%)	0.41	7 (7%) 19 12	26, 95, 263, 292	0
4	RRRR	98/118 (83%)	0.02	3 (3%) 52 38	36, 96, 189, 246	0
5	F	74/86 (86%)	-0.01	0 100 100	25, 74, 117, 136	0
5	FF	75/86 (87%)	-0.04	0 100 100	25, 88, 149, 188	0
5	FFF	74/86 (86%)	0.11	0 100 100	47, 89, 139, 166	0
5	FFFF	74/86 (86%)	0.10	1 (1%) 78 65	42, 88, 131, 165	0
5	M	79/86 (91%)	0.05	2 (2%) 61 46	45, 90, 153, 171	0
5	MM	80/86 (93%)	-0.10	0 100 100	39, 89, 139, 167	0
5	MMM	76/86 (88%)	0.13	0 100 100	45, 97, 143, 167	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
5	MMMM	78/86 (90%)	-0.07	0 100 100	37, 85, 136, 165	0
5	T	76/86 (88%)	0.10	2 (2%) 59 44	49, 94, 153, 219	0
5	TT	78/86 (90%)	-0.08	0 100 100	45, 89, 138, 174	0
5	TTT	76/86 (88%)	0.13	2 (2%) 59 44	46, 91, 142, 213	0
5	TTTT	77/86 (89%)	0.05	2 (2%) 59 44	44, 92, 156, 203	0
6	E	79/92 (85%)	-0.06	0 100 100	46, 100, 142, 181	0
6	EE	79/92 (85%)	0.06	1 (1%) 79 66	53, 102, 159, 210	0
6	EEE	79/92 (85%)	0.10	2 (2%) 61 46	51, 100, 158, 195	0
6	EEEE	79/92 (85%)	0.04	1 (1%) 79 66	54, 96, 149, 168	0
6	L	79/92 (85%)	0.14	3 (3%) 44 32	57, 100, 157, 221	0
6	LL	79/92 (85%)	0.04	1 (1%) 79 66	62, 103, 154, 173	0
6	LLL	79/92 (85%)	0.27	4 (5%) 32 22	57, 114, 193, 234	0
6	LLLL	77/92 (83%)	-0.03	2 (2%) 59 44	42, 91, 157, 190	0
6	S	79/92 (85%)	0.09	2 (2%) 61 46	47, 112, 180, 205	0
6	SS	79/92 (85%)	-0.01	0 100 100	52, 98, 157, 202	0
6	SSS	79/92 (85%)	0.09	4 (5%) 32 22	58, 109, 179, 203	0
6	SSSS	79/92 (85%)	0.17	3 (3%) 44 32	51, 107, 180, 227	0
7	G	74/76 (97%)	0.24	0 100 100	58, 130, 187, 200	0
7	GG	74/76 (97%)	0.12	4 (5%) 29 20	45, 120, 173, 204	0
7	GGG	74/76 (97%)	0.17	3 (4%) 41 29	64, 133, 182, 222	0
7	GGGG	74/76 (97%)	0.16	4 (5%) 29 20	57, 129, 183, 233	0
7	N	74/76 (97%)	0.31	5 (6%) 20 13	76, 128, 186, 212	0
7	NN	74/76 (97%)	0.24	5 (6%) 20 13	70, 120, 172, 243	0
7	NNN	74/76 (97%)	0.31	3 (4%) 41 29	75, 142, 187, 214	0
7	NNNN	74/76 (97%)	0.31	5 (6%) 20 13	63, 111, 186, 233	0
7	U	74/76 (97%)	0.38	8 (10%) 8 6	86, 132, 212, 239	0
7	UU	74/76 (97%)	0.31	5 (6%) 20 13	61, 133, 181, 226	0
7	UUU	74/76 (97%)	0.49	9 (12%) 5 5	50, 141, 213, 254	0
7	UUUU	74/76 (97%)	0.33	5 (6%) 20 13	74, 144, 188, 206	0
8	V	68/68 (100%)	-0.63	1 (1%) 76 64	67, 171, 209, 272	0
8	VV	68/68 (100%)	-0.60	0 100 100	63, 148, 219, 272	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
8	VVV	68/68 (100%)	-0.63	1 (1%) 76 64	75, 163, 213, 248	0
8	VVVV	68/68 (100%)	-0.60	0 100 100	67, 168, 222, 263	0
8	X	68/68 (100%)	-0.60	0 100 100	81, 164, 218, 272	0
8	XX	68/68 (100%)	-0.60	0 100 100	74, 161, 218, 267	0
8	XXX	68/68 (100%)	-0.66	0 100 100	85, 158, 210, 221	0
8	XXXX	68/68 (100%)	-0.55	0 100 100	69, 158, 220, 249	0
8	Y	68/68 (100%)	-0.62	1 (1%) 76 64	79, 173, 228, 245	0
8	YY	68/68 (100%)	-0.61	0 100 100	81, 155, 229, 291	0
8	YYY	68/68 (100%)	-0.57	1 (1%) 76 64	78, 164, 247, 310	0
8	YYYY	68/68 (100%)	-0.62	0 100 100	83, 164, 226, 341	0
All	All	7623/9336 (81%)	0.07	275 (3%) 46 33	25, 118, 195, 341	1 (0%)

The worst 5 of 275 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	Q	1	MET	8.9
4	RRR	83	GLY	7.8
4	RRR	82	LYS	7.8
2	II	4	GLY	7.4
4	KKKK	83	GLY	7.4

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.