



wwPDB X-ray Structure Validation Summary Report

Mar 31, 2014 – 06:13 PM BST

PDB ID : 4F77
Title : The 8S snRNP Assembly Intermediate
Authors : Grimm, C.; Pelz, J.P.; Schindelin, H.; Diederichs, K.; Kuper, J.; Kisker, C.
Deposited on : 2012-05-15
Resolution : 3.10 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

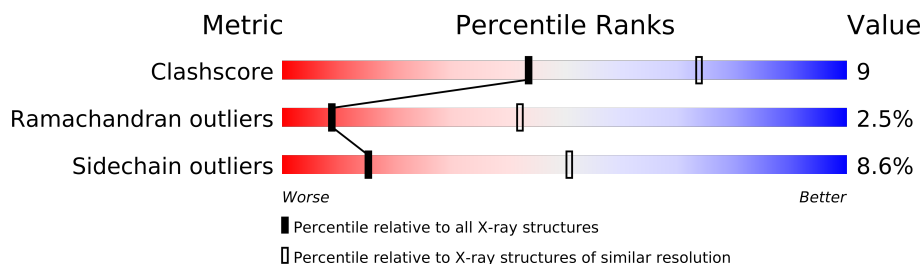
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	NOT EXECUTED
Percentile statistics	:	21963
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable23004

1 Overall quality at a glance

The reported resolution of this entry is 3.10 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	79885	1078 (3.16-3.04)
Ramachandran outliers	78287	1044 (3.16-3.04)
Sidechain outliers	78261	1044 (3.16-3.04)









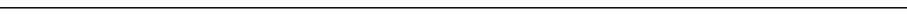













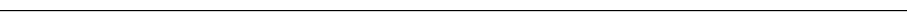

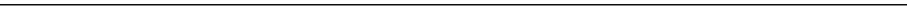











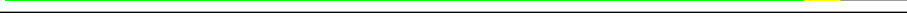


The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	119	
1	I	119	
1	Q	119	
1	Y	119	
1	g	119	
1	o	119	
1	w	119	
2	B	118	
2	J	118	
2	R	118	
2	Z	118	
2	h	118	
2	p	118	
2	x	118	
3	C	92	
3	K	92	
3	S	92	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	a	92	
3	i	92	
3	q	92	
3	y	92	
4	D	86	
4	L	86	
4	T	86	
4	b	86	
4	j	86	
4	r	86	
4	z	86	
5	1	124	
5	E	124	
5	M	124	
5	U	124	
5	c	124	
5	k	124	
5	s	124	
6	2	247	
6	F	247	
6	N	247	
6	V	247	
6	d	247	
6	l	247	
6	t	247	
7	3	186	
7	G	186	
7	O	186	
7	W	186	
7	e	186	
7	m	186	
7	u	186	
8	4	76	
8	H	76	
8	P	76	
8	X	76	
8	f	76	
8	n	76	
8	v	76	

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 42689 atoms, of which 0 are hydrogen and 0 are deuterium.

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 D1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	82	Total	C	N	O	S	0	0	0
			648	413	113	119	3			
1	A	82	Total	C	N	O	S	0	0	0
			648	413	113	119	3			
1	Q	82	Total	C	N	O	S	0	0	0
			648	413	113	119	3			
1	Y	82	Total	C	N	O	S	0	0	0
			648	413	113	119	3			
1	g	82	Total	C	N	O	S	0	0	0
			648	413	113	119	3			
1	o	82	Total	C	N	O	S	0	0	0
			648	413	113	119	3			
1	w	82	Total	C	N	O	S	0	0	0
			648	413	113	119	3			

- Molecule 2 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	J	100	Total	C	N	O	S	0	0	0
			807	505	146	150	6			
2	B	100	Total	C	N	O	S	0	0	0
			807	505	146	150	6			
2	R	100	Total	C	N	O	S	0	0	0
			807	505	146	150	6			
2	Z	100	Total	C	N	O	S	0	0	0
			807	505	146	150	6			
2	h	100	Total	C	N	O	S	0	0	0
			807	505	146	150	6			
2	p	100	Total	C	N	O	S	0	0	0
			807	505	146	150	6			
2	x	100	Total	C	N	O	S	0	0	0
			807	505	146	150	6			

- Molecule 3 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	K	77	Total	C	N	O	S	0	0	0
			638	405	113	115	5			
3	C	77	Total	C	N	O	S	0	0	0
			638	405	113	115	5			
3	S	77	Total	C	N	O	S	0	0	0
			638	405	113	115	5			
3	a	77	Total	C	N	O	S	0	0	0
			638	405	113	115	5			
3	i	77	Total	C	N	O	S	0	0	0
			638	405	113	115	5			
3	q	77	Total	C	N	O	S	0	0	0
			638	405	113	115	5			
3	y	77	Total	C	N	O	S	0	0	0
			638	405	113	115	5			

- Molecule 4 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	71	Total	C	N	O	S	0	0	0
			556	358	92	101	5			
4	D	71	Total	C	N	O	S	0	0	0
			556	358	92	101	5			
4	T	71	Total	C	N	O	S	0	0	0
			556	358	92	101	5			
4	b	71	Total	C	N	O	S	0	0	0
			556	358	92	101	5			
4	j	71	Total	C	N	O	S	0	0	0
			556	358	92	101	5			
4	r	71	Total	C	N	O	S	0	0	0
			556	358	92	101	5			
4	z	71	Total	C	N	O	S	0	0	0
			556	358	92	101	5			

- Molecule 5 is a protein called LD23602p.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	M	17	Total	C	N	O	0	0	0
			133	85	19	29			
5	E	17	Total	C	N	O	0	0	0
			133	85	19	29			
5	U	17	Total	C	N	O	0	0	0
			133	85	19	29			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	c	17	Total	C	N	O	0	0	0
			133	85	19	29			
5	k	17	Total	C	N	O	0	0	0
			133	85	19	29			
5	s	17	Total	C	N	O	0	0	0
			133	85	19	29			
5	1	17	Total	C	N	O	0	0	0
			133	85	19	29			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	6991	GLY	-	EXPRESSION TAG	UNP Q9VV74
M	6992	ALA	-	EXPRESSION TAG	UNP Q9VV74
E	6991	GLY	-	EXPRESSION TAG	UNP Q9VV74
E	6992	ALA	-	EXPRESSION TAG	UNP Q9VV74
U	6991	GLY	-	EXPRESSION TAG	UNP Q9VV74
U	6992	ALA	-	EXPRESSION TAG	UNP Q9VV74
c	6991	GLY	-	EXPRESSION TAG	UNP Q9VV74
c	6992	ALA	-	EXPRESSION TAG	UNP Q9VV74
k	6991	GLY	-	EXPRESSION TAG	UNP Q9VV74
k	6992	ALA	-	EXPRESSION TAG	UNP Q9VV74
s	6991	GLY	-	EXPRESSION TAG	UNP Q9VV74
s	6992	ALA	-	EXPRESSION TAG	UNP Q9VV74
1	6991	GLY	-	EXPRESSION TAG	UNP Q9VV74
1	6992	ALA	-	EXPRESSION TAG	UNP Q9VV74

- Molecule 6 is a protein called CG10419.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	N	216	Total	C	N	O	S	0	0	0
			1787	1138	309	331	9			
6	F	216	Total	C	N	O	S	0	0	0
			1787	1138	309	331	9			
6	V	216	Total	C	N	O	S	0	0	0
			1787	1138	309	331	9			
6	d	216	Total	C	N	O	S	0	0	0
			1787	1138	309	331	9			
6	l	216	Total	C	N	O	S	0	0	0
			1787	1138	309	331	9			
6	t	216	Total	C	N	O	S	0	0	0
			1787	1138	309	331	9			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	2	216	Total	C	N	O	S	0	0	0
			1787	1138	309	331	9			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
N	7969	GLY	-	EXPRESSION TAG	UNP Q9VVDX0
N	7970	ALA	-	EXPRESSION TAG	UNP Q9VVDX0
F	7969	GLY	-	EXPRESSION TAG	UNP Q9VVDX0
F	7970	ALA	-	EXPRESSION TAG	UNP Q9VVDX0
V	7969	GLY	-	EXPRESSION TAG	UNP Q9VVDX0
V	7970	ALA	-	EXPRESSION TAG	UNP Q9VVDX0
d	7969	GLY	-	EXPRESSION TAG	UNP Q9VVDX0
d	7970	ALA	-	EXPRESSION TAG	UNP Q9VVDX0
l	7969	GLY	-	EXPRESSION TAG	UNP Q9VVDX0
l	7970	ALA	-	EXPRESSION TAG	UNP Q9VVDX0
t	7969	GLY	-	EXPRESSION TAG	UNP Q9VVDX0
t	7970	ALA	-	EXPRESSION TAG	UNP Q9VVDX0
2	7969	GLY	-	EXPRESSION TAG	UNP Q9VVDX0
2	7970	ALA	-	EXPRESSION TAG	UNP Q9VVDX0

- Molecule 7 is a protein called Ichn.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	O	127	Total	C	N	O	S	0	0	0
			984	626	161	188	9			
7	G	127	Total	C	N	O	S	0	0	0
			984	626	161	188	9			
7	W	127	Total	C	N	O	S	0	0	0
			984	626	161	188	9			
7	e	127	Total	C	N	O	S	0	0	0
			984	626	161	188	9			
7	m	127	Total	C	N	O	S	0	0	0
			984	626	161	188	9			
7	u	127	Total	C	N	O	S	0	0	0
			984	626	161	188	9			
7	3	127	Total	C	N	O	S	0	0	0
			984	626	161	188	9			

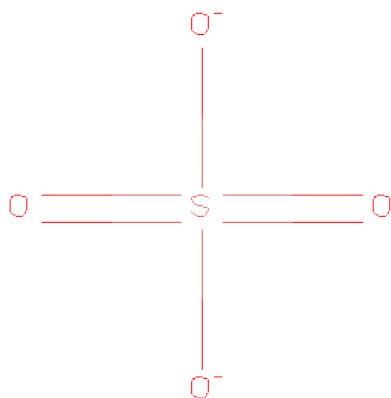
There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
O	6180	HIS	-	EXPRESSION TAG	UNP Q9U3W1
O	6181	HIS	-	EXPRESSION TAG	UNP Q9U3W1
O	6182	HIS	-	EXPRESSION TAG	UNP Q9U3W1
O	6183	HIS	-	EXPRESSION TAG	UNP Q9U3W1
O	6184	HIS	-	EXPRESSION TAG	UNP Q9U3W1
O	6185	HIS	-	EXPRESSION TAG	UNP Q9U3W1
G	6180	HIS	-	EXPRESSION TAG	UNP Q9U3W1
G	6181	HIS	-	EXPRESSION TAG	UNP Q9U3W1
G	6182	HIS	-	EXPRESSION TAG	UNP Q9U3W1
G	6183	HIS	-	EXPRESSION TAG	UNP Q9U3W1
G	6184	HIS	-	EXPRESSION TAG	UNP Q9U3W1
G	6185	HIS	-	EXPRESSION TAG	UNP Q9U3W1
W	6180	HIS	-	EXPRESSION TAG	UNP Q9U3W1
W	6181	HIS	-	EXPRESSION TAG	UNP Q9U3W1
W	6182	HIS	-	EXPRESSION TAG	UNP Q9U3W1
W	6183	HIS	-	EXPRESSION TAG	UNP Q9U3W1
W	6184	HIS	-	EXPRESSION TAG	UNP Q9U3W1
W	6185	HIS	-	EXPRESSION TAG	UNP Q9U3W1
e	6180	HIS	-	EXPRESSION TAG	UNP Q9U3W1
e	6181	HIS	-	EXPRESSION TAG	UNP Q9U3W1
e	6182	HIS	-	EXPRESSION TAG	UNP Q9U3W1
e	6183	HIS	-	EXPRESSION TAG	UNP Q9U3W1
e	6184	HIS	-	EXPRESSION TAG	UNP Q9U3W1
e	6185	HIS	-	EXPRESSION TAG	UNP Q9U3W1
m	6180	HIS	-	EXPRESSION TAG	UNP Q9U3W1
m	6181	HIS	-	EXPRESSION TAG	UNP Q9U3W1
m	6182	HIS	-	EXPRESSION TAG	UNP Q9U3W1
m	6183	HIS	-	EXPRESSION TAG	UNP Q9U3W1
m	6184	HIS	-	EXPRESSION TAG	UNP Q9U3W1
m	6185	HIS	-	EXPRESSION TAG	UNP Q9U3W1
u	6180	HIS	-	EXPRESSION TAG	UNP Q9U3W1
u	6181	HIS	-	EXPRESSION TAG	UNP Q9U3W1
u	6182	HIS	-	EXPRESSION TAG	UNP Q9U3W1
u	6183	HIS	-	EXPRESSION TAG	UNP Q9U3W1
u	6184	HIS	-	EXPRESSION TAG	UNP Q9U3W1
u	6185	HIS	-	EXPRESSION TAG	UNP Q9U3W1
3	6180	HIS	-	EXPRESSION TAG	UNP Q9U3W1
3	6181	HIS	-	EXPRESSION TAG	UNP Q9U3W1
3	6182	HIS	-	EXPRESSION TAG	UNP Q9U3W1
3	6183	HIS	-	EXPRESSION TAG	UNP Q9U3W1
3	6184	HIS	-	EXPRESSION TAG	UNP Q9U3W1
3	6185	HIS	-	EXPRESSION TAG	UNP Q9U3W1

- Molecule 8 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	P	70	Total	C	N	O	S	0	0	0
			544	344	96	98	6			
8	H	70	Total	C	N	O	S	0	0	0
			544	344	96	98	6			
8	X	70	Total	C	N	O	S	0	0	0
			544	344	96	98	6			
8	f	70	Total	C	N	O	S	0	0	0
			544	344	96	98	6			
8	n	70	Total	C	N	O	S	0	0	0
			544	344	96	98	6			
8	v	70	Total	C	N	O	S	0	0	0
			544	344	96	98	6			
8	4	70	Total	C	N	O	S	0	0	0
			544	344	96	98	6			

- Molecule 9 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



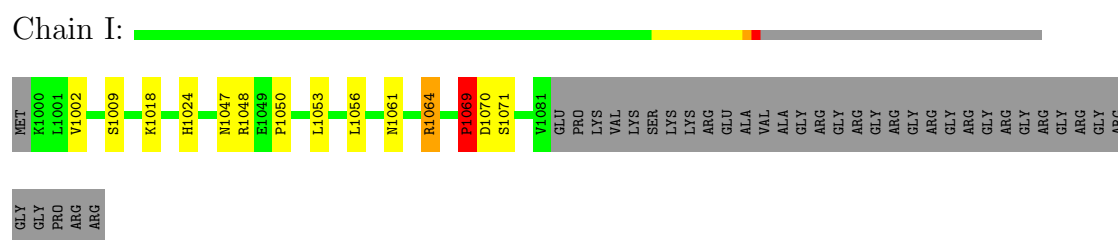
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	t	1	Total	O	S	0	0
			5	4	1		
9	2	1	Total	O	S	0	0
			5	4	1		

3 Residue-property plots

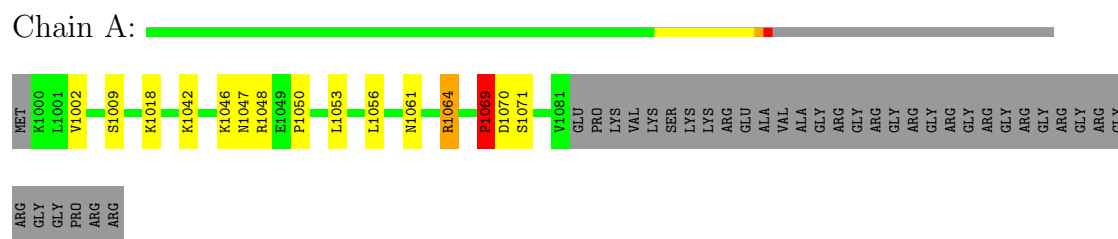
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according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

Note EDS was not executed.

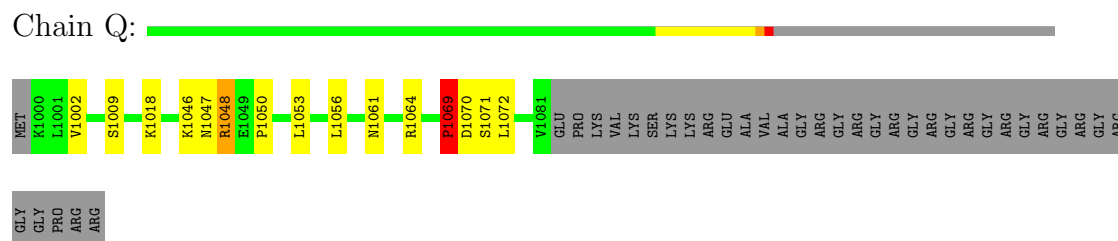
- Molecule 1: Small nuclear ribonucleoprotein Sm D1



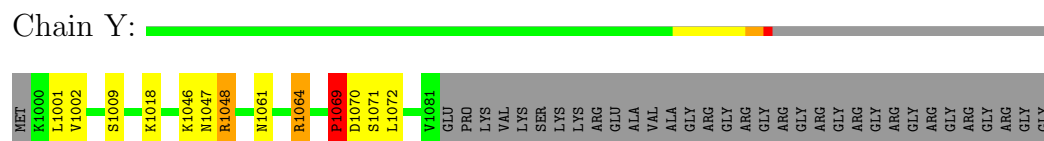
- Molecule 1: Small nuclear ribonucleoprotein Sm D1



- Molecule 1: Small nuclear ribonucleoprotein Sm D1

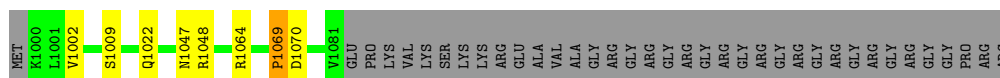


- Molecule 1: Small nuclear ribonucleoprotein Sm D1



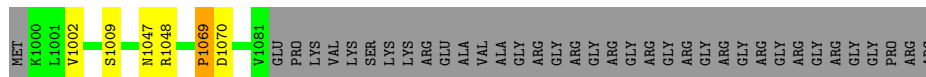
- Molecule 1: Small nuclear ribonucleoprotein Sm D1





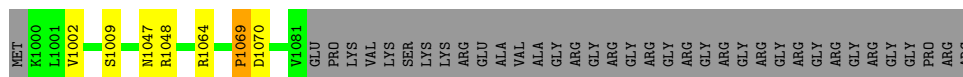
- Molecule 1: Small nuclear ribonucleoprotein Sm D1

Chain 0:



- Molecule 1: Small nuclear ribonucleoprotein Sm D1

Chain w:



- Molecule 2: Small nuclear ribonucleoprotein Sm D2

Chain J:



- Molecule 2: Small nuclear ribonucleoprotein Sm D2

Chain B:



- Molecule 2: Small nuclear ribonucleoprotein Sm D2

Chain R:



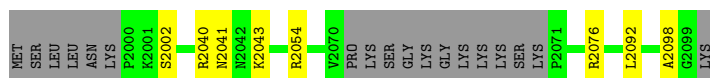
- Molecule 2: Small nuclear ribonucleoprotein Sm D2

Chain Z:



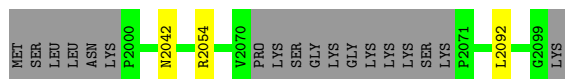
- Molecule 2: Small nuclear ribonucleoprotein Sm D2

Chain h:



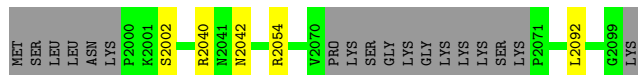
- Molecule 2: Small nuclear ribonucleoprotein Sm D2

Chain p:



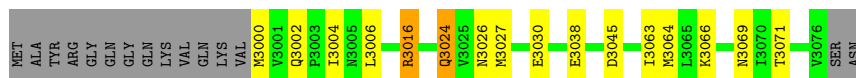
- Molecule 2: Small nuclear ribonucleoprotein Sm D2

Chain x:



- Molecule 3: Small nuclear ribonucleoprotein E

Chain K:



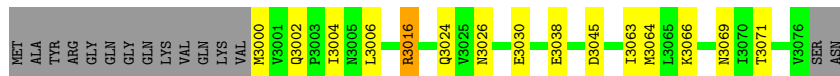
- Molecule 3: Small nuclear ribonucleoprotein E

Chain C:



- Molecule 3: Small nuclear ribonucleoprotein E

Chain S:



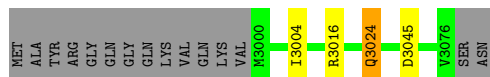
- Molecule 3: Small nuclear ribonucleoprotein E

Chain a:



- Molecule 3: Small nuclear ribonucleoprotein E

Chain i:



- Molecule 3: Small nuclear ribonucleoprotein E

Chain q: 

MET	ALA	TYR	ARG	GLY	GLN	GLY	GLN	LYS	VAL	GLN	LYS	VAL	K3000	I3004	R3016	Q3024	Y3025	N3026	D3045	V3076	SER	ASN
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-------	-----	-----

- Molecule 3: Small nuclear ribonucleoprotein E

Chain y: 

MET	ALA	TYR	ARG	GLY	GLN	GLY	GLN	LYS	VAL	GLN	LYS	VAL	K3000	R3016	Q3024	V3025	N3026	D3045	V3076	SER	ASN
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-----	-----

- Molecule 4: Small nuclear ribonucleoprotein F

Chain L: 

MET	SER	LEU	PRO	LEU	N4000	P4001	K4002	P4003	F4004	K4016	L4017	K4018	W4019	G4020	M4021	V4030	D4031	D4046	L4053	V4056	R4059	C4060	N4061	N4062	V4063	L4064	R4067	E4070	GLU	GLU	GLU	GLU	ASP	GLY	GLY	MET	ARG	GLU
-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Molecule 4: Small nuclear ribonucleoprotein F

Chain D: 

MET	SER	LEU	PRO	LEU	N4000	P4001	K4002	P4003	F4004	K4016	L4017	K4018	W4019	G4020	M4021	V4030	D4031	D4046	L4053	V4056	R4059	C4060	N4061	N4062	V4063	L4064	R4067	E4070	GLU	GLU	GLU	GLU	ASP	GLY	GLY	MET	ARG	GLU
-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Molecule 4: Small nuclear ribonucleoprotein F

Chain T: 

MET	SER	LEU	PRO	LEU	N4000	P4001	K4002	P4003	F4004	L4005	K4016	L4017	K4018	W4019	G4020	M4021	V4030	D4031	D4046	L4053	V4056	R4059	C4060	N4061	N4062	V4063	L4064	R4067	E4070	GLU	GLU	GLU	GLU	ASP	GLY	GLY	MET	ARG	GLU
-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Molecule 4: Small nuclear ribonucleoprotein F

Chain b: 

MET	SER	LEU	PRO	LEU	N4000	P4001	K4002	M4021	D4046	L4053	E4070	GLU	GLU	GLU	GLU	ASP	GLY	GLY	MET	ARG	GLU
-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Molecule 4: Small nuclear ribonucleoprotein F

Chain j: 

MET	SER	LEU	PRO	LEU	N4000	P4001	K4002	M4021	D4046	L4053	E4070	GLU	GLU	GLU	GLU	ASP	GLY	GLY	MET	ARG	GLU
-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Molecule 4: Small nuclear ribonucleoprotein F

Chain r: 

MET	SER	LEU	PRO	LEU	N4000	P4001	K4002	M4021	D4046	L4053	E4070	GLU	GLU	GLU	GLU	ASP	GLY	GLY	MET	ARG	GLU
-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Molecule 4: Small nuclear ribonucleoprotein F

Chain z:

MET	SER	LEU	PRO	LEU	N4000	M4021	D4046	L4053	E4070	GLU	GLU	GLU	GLU	ASP	GLY	MET	ARG	GLU
-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Molecule 5: LD23602p

Chain M:

GLY	ALA	MET	SER	PRO	ASP	GLU	THR	ASN	LYS	A7000	V7001	W7002	S7005	L7006	L7007	T7010	S7014	V7015	G7016	LEU	ALA	ARG	GLU	GLU	ALA	ALA	VAL	ALA	ALA	ARG	ASN	LEU	GLU	ASP	LYS	THR	ASN	LYS	CYS	VAL	LEU	ARG	GLU	GLU	ASN	ALA	ALA	TYR	GLY	LEU	ASN	GLN	GLU	VAL	ALA	GLY	THR	ASP	LEU	PRO	GLY	THR	ALA	SER	TRP
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

THR	SER	PRO	GLU	VAL	PHE	SER	THR	ASN	LYS	VAL	GLY	ASP	TYR	ALA	ARG	THR	THR	VAL	ASP	GLY	VAL	GLY	GLU	TYR	ASP	VAL	VAL	VAL	VAL	ASN	ILE	ASN	GLU	ASP	LYS	THR	ASN	LYS	CYS	THR	VAL	LEU	ARG	GLU	GLU	ASN	ALA	ALA	TYR	GLY	LEU	ASN	GLN	GLU	VAL	LEU	VAL	GLY	THR	ASP	LEU	PRO	GLY	THR	ALA	SER	TRP
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Molecule 5: LD23602p

Chain E:

GLY	ALA	MET	SER	PRO	ASP	GLU	THR	ASN	LYS	A7000	V7001	W7002	D7003	S7005	L7007	G7016	LEU	ALA	ARG	GLU	GLU	ALA	ALA	LEU	ALA	ALA	ARG	ASN	ILE	ASN	GLU	ASP	LYS	THR	ASN	LYS	CYS	THR	VAL	LEU	ARG	GLU	GLU	ASN	ALA	ALA	TYR	GLY	LEU	ASN	GLN	GLU	VAL	LEU	VAL	GLY	THR	ASP	LEU	PRO	GLY	THR	ALA	SER	TRP
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

GLU	PRO	VAL	SER	PHE	LYS	VAL	GLY	ASP	TYR	ALA	ALA	ARG	ALA	THR	VAL	ASP	GLY	VAL	ASP	TYR	GLU	GLY	ALA	VAL	VAL	VAL	ILE	ASN	GLU	ALA	GLY	LYS	GLY	THR	ASN	CYS	THR	VAL	LEU	ARG	GLU	GLU	ASN	ALA	ALA	TYR	GLY	LEU	ASN	GLN	GLU	VAL	LEU	VAL	GLY	THR	ASP	LEU	PRO	GLY	THR	ALA	SER	TRP
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Molecule 5: LD23602p

Chain U:

GLY	ALA	MET	SER	PRO	ASP	GLU	THR	ASN	LYS	A7000	V7001	W7002	D7003	S7005	L7007	V7008	G7016	LEU	ALA	ARG	GLU	GLU	ALA	ALA	LEU	ALA	ALA	ARG	ASN	ILE	ASN	GLU	ASP	LYS	THR	ASN	LYS	CYS	THR	VAL	LEU	ARG	GLU	GLU	ASN	ALA	ALA	TYR	GLY	LEU	ASN	GLN	GLU	VAL	LEU	VAL	GLY	THR	ASP	LEU	PRO	GLY	THR	ALA	SER	TRP
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

PRO	GLU	PRO	VAL	SER	PHE	LYS	VAL	GLY	ASP	TYR	ALA	ARG	ALA	THR	VAL	ASP	GLY	VAL	ASP	TYR	GLU	GLY	ALA	VAL	VAL	VAL	ILE	ASN	GLU	ALA	GLY	LYS	GLY	THR	ASN	CYS	THR	VAL	LEU	ARG	GLU	GLU	ASN	ALA	ALA	TYR	GLY	LEU	ASN	GLN	GLU	VAL	LEU	VAL	GLY	THR	ASP	LEU	PRO	GLY	THR	ALA	SER	TRP
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Molecule 5: LD23602p

Chain c:

GLY	ALA	MET	SER	PRO	ASP	GLU	THR	ASN	LYS	A7000	V7001	W7002	L7006	L7007	G7016	LEU	ALA	ARG	GLU	GLU	ALA	ALA	LEU	ALA	ALA	ARG	ARG	ASN	ILE	ASN	GLU	ASP	LYS	THR	ASN	LYS	CYS	THR	VAL	LEU	ARG	GLU	GLU	ASN	ALA	ALA	TYR	GLY	LEU	ASN	GLN	GLU	VAL	LEU	VAL	GLY	THR	ASP	LEU	PRO	GLY	THR	ALA	SER	TRP
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

VAL	SER	PHE	LYS	VAL	ASP	GLY	THR	ASN	LYS	ALA	ARG	ALA	THR	VAL	VAL	ASP	GLY	VAL	ASP	TYR	GLU	GLY	ALA	VAL	VAL	VAL	ILE	ASN	GLU	ALA	GLY	LYS	THR	ASN	LYS	CYS	THR	VAL	LEU	ARG	GLU	GLU	ASN	ALA	ALA	TYR	GLY	LEU	ASN	GLN	GLU	VAL	LEU	VAL	GLY	THR	ASP	LEU	PRO	GLY	THR	ALA	SER	TRP
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Molecule 5: LD23602p

Chain k:

GLY	ALA	MET	SER	PRO	ASP	GLU	THR	ASN	LYS	A7000	V7001	W7002	L7006	G7016	LEU	ALA	ARG	GLU	GLU	ALA	ALA	LEU	ALA	ALA	ARG	ARG	ASN	ILE	ASN	GLU	ASP	LYS	THR	ASN	LYS	CYS	THR	VAL	LEU	ARG	GLU	GLU	ASN	ALA	ALA	TYR	GLY	LEU	ASN	GLN	GLU	VAL	LEU	VAL	GLY	THR	ASP	LEU	PRO	GLY	THR	ALA	SER	TRP
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

SER PHE LYS VAL SER GLY ASP TYR ALA ARG ALA THR TYR VAL ASP GLY VAL ASP TYR GLY ALA VAL VAL SER ILE ASN GLU GLY LYS SER THR CYS VAL VAL LEU ARG LEU TYR LEU GLY TYR ASN GLU GLN GLU VAL LEU LEU VAL ASP LEU LEU PRO SER TRP

• Molecule 5: LD23602p

Chain s:

GLY ALA MET SER ASP GLY THR ALA A7000 V7001 W7002 L7006 L7007 G7016 LEU ALA ARG ALA ALA VAL VAL ALA ARG ARG LEU ASP ASP LYS THR ASN LYS VAL ARG LEU TYR GLY ASN ALA ALA ALA GLN GLU VAL LEU LEU VAL ILE SER LEU LEU PRO PRO

VAL SER PHE LYS VAL SER GLY ASP TYR ALA ARG ALA THR TYR VAL ASP GLY VAL ASP TYR GLY ALA VAL VAL SER ILE ASN GLU GLY LYS SER THR CYS VAL VAL LEU ARG LEU TYR GLY ASN ALA ALA GLN GLU VAL LEU LEU VAL ILE SER LEU LEU PRO SER TRP

• Molecule 5: LD23602p

Chain 1:

GLY ALA MET SER ASP GLY THR ALA A7000 V7001 W7002 L7006 L7007 S7014 V7015 G7016 LEU ALA ARG ALA ALA VAL VAL ALA ARG ARG LEU ASP ASP LYS THR ASN LYS VAL ARG LEU TYR GLY ASN ALA ALA ALA GLN GLU VAL LEU LEU VAL ILE SER LEU LEU PRO PRO

GLU PRO SER PHE LYS VAL VAL GLY TYR ALA ARG ALA THR TYR VAL ASP GLY VAL ASP TYR GLY ALA VAL VAL SER ILE ASN GLU GLY LYS SER THR CYS VAL VAL LEU ARG LEU TYR GLY ASN ALA ALA GLN GLU VAL LEU LEU VAL ILE SER LEU LEU PRO SER TRP

• Molecule 6: CG10419

Chain N:

GLY ALA MET GLN HIS PRO ASP GLN THR PHE GLN LEU GLN ALA LEU LEU ILE CYS GLU PRO ASP TYR VAL ASP TYR GLY ALA VAL VAL SER ILE ASN GLU GLY LYS SER THR CYS VAL VAL LEU ARG LEU TYR GLY ASN ALA ALA GLN GLU VAL LEU LEU VAL ILE SER LEU LEU PRO SER TRP

E8044 L8045 P8046 P8047 T8054 W8057 Q8066 L8073 L8079 N8083 L8086 E8089 T8093 S8094 D8095 Q8096 E8097 K8098 R8099 W8099 L8100 P8104 N8105 Q8106 G8103 E8104 E8105 Y8106 L8109 L8113 L8117 L8124 F8127 W8131 L8132 Q8133 P8135 N8136 T8137 W8138 L8139 L8140 L8141 L8142 W8146 M8147

A8148 R8149 W8150 T8154 H8159 L8160 P8161 F8167 L8170 R8171 A8174 T8176 C8177 L8178 L8180 R8181 L8184 R8185 E8186 D8187 E8188 R8189 Q8190 R8191 A8192 Y8195 N8196 L8197 L8198 L8199 T8200 Y8201 V8203 Q8204 V8205 F8206 Q8208 N8209 D8210 F8211 K8212 T8213 W8214 L8215

• Molecule 6: CG10419

Chain F:

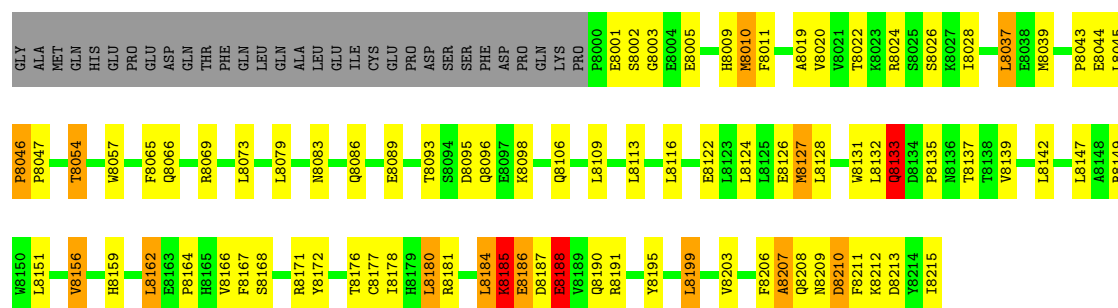
GLY ALA MET GLN HIS PRO ASP GLN THR PHE GLN LEU GLN ALA LEU LEU ILE CYS GLU PRO ASP TYR VAL ASP TYR GLY ALA VAL VAL SER ILE ASN GLU GLY LYS SER THR CYS VAL VAL LEU ARG LEU TYR GLY ASN ALA ALA GLN GLU VAL LEU LEU VAL ILE SER LEU LEU PRO SER TRP

L8045 P8046 P8047 T8054 W8057 Q8071 V8072 L8073 V8074 L8075 L8076 R8077 N8083 L8086 E8089 T8093 S8094 D8095 Q8096 E8097 K8098 R8099 W8099 L8100 P8104 N8105 Q8106 G8103 E8104 E8105 Y8106 L8109 L8113 L8117 L8118 N8119 D8120 L8121 L8122 L8123 L8124 L8125 E8126 N8127 W8131 L8132 Q8133 P8135 N8136 T8137

T8138 W8139 L8142 W8145 R8149 W8150 L8151 Y8152 A8153 T8154 L8155 W8156 L8159 R8160 P8161 L8162 F8167 L8170 T8176 L8180 R8181 L8184 R8185 E8186 D8187 E8188 R8189 Q8190 R8191 A8192 Y8195 N8196 L8197 L8198 L8199 T8200 Y8201 V8203 Q8204 V8205 F8206 Q8208 N8209 D8210 F8211 K8212 T8213 W8214 L8215

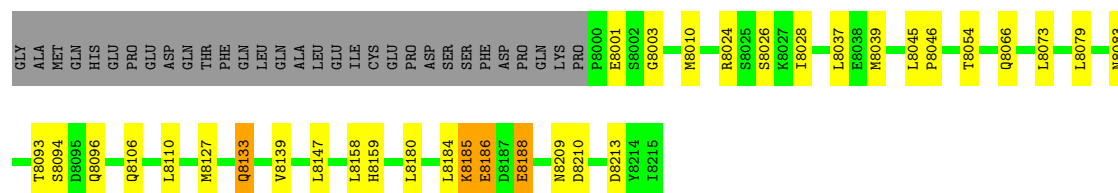
• Molecule 6: CG10419

Chain V:



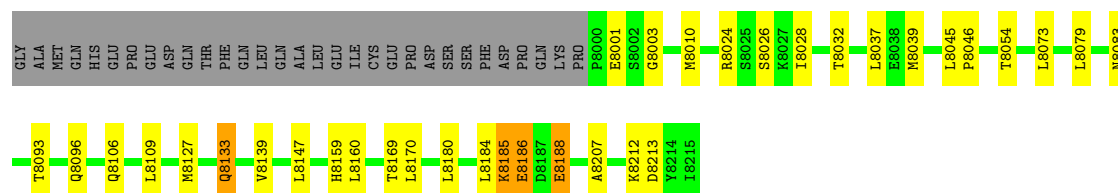
• Molecule 6: CG10419

Chain d:



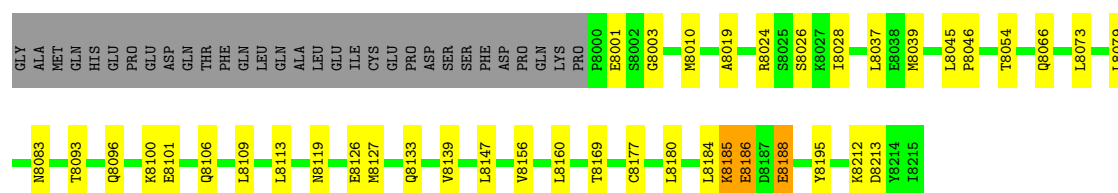
• Molecule 6: CG10419

Chain l:



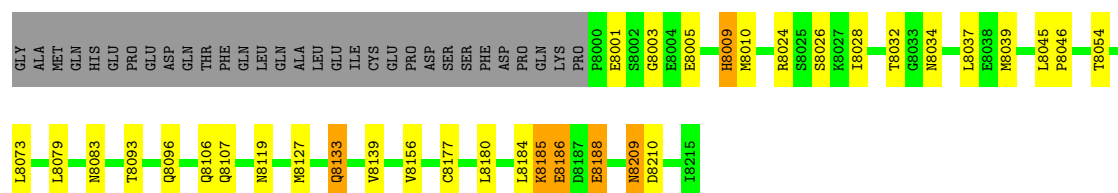
• Molecule 6: CG10419

Chain t:



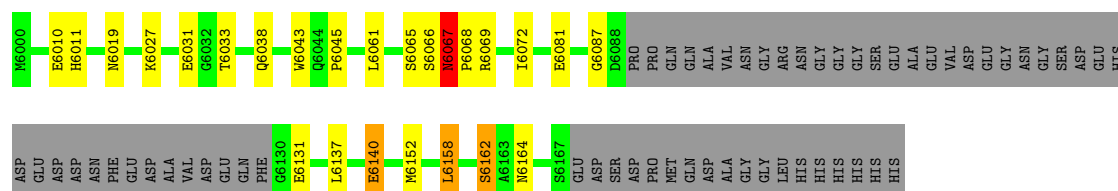
• Molecule 6: CG10419

Chain 2:



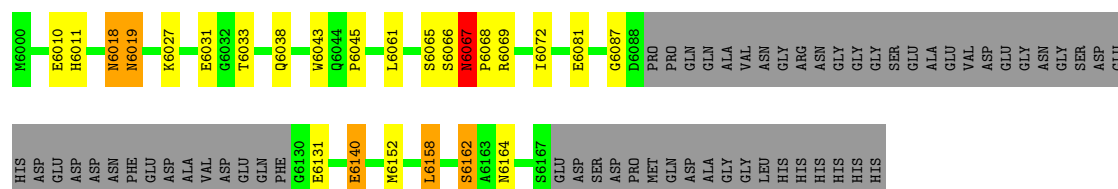
• Molecule 7: Ichn

Chain O:



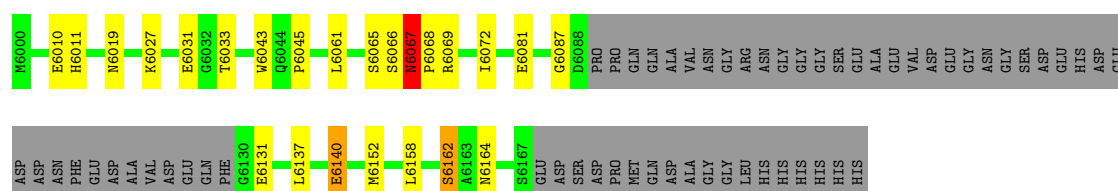
- Molecule 7: Ichn

Chain G:



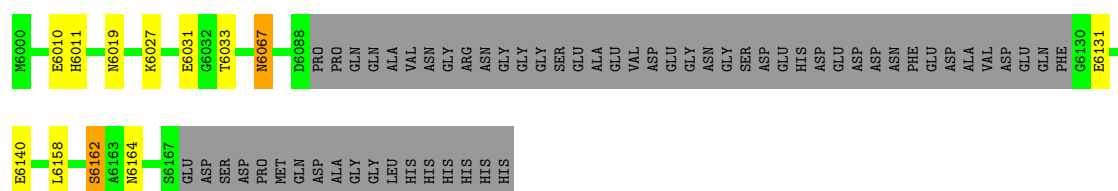
- Molecule 7: Ichn

Chain W:



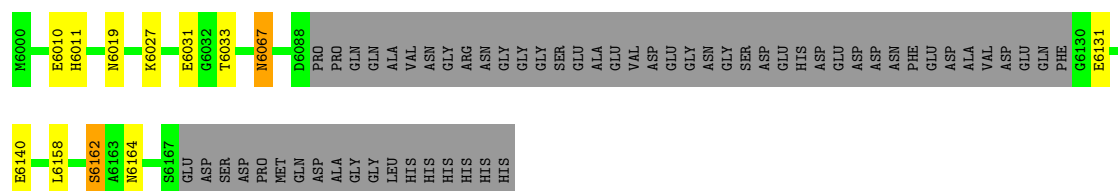
- Molecule 7: Ichn

Chain e:



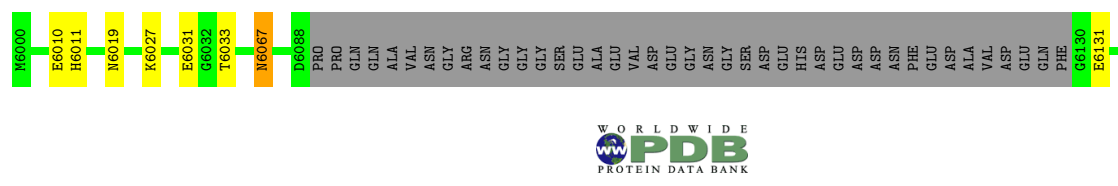
- Molecule 7: Ichn

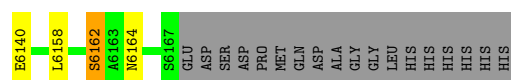
Chain m:



- Molecule 7: Ichn

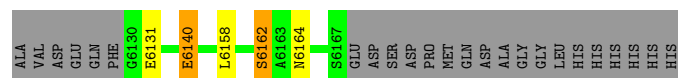
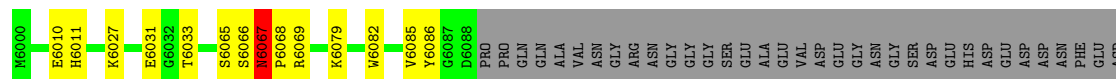
Chain u:





- Molecule 7: Ichn

Chain 3:



- Molecule 8: Small nuclear ribonucleoprotein G

Chain P:



- Molecule 8: Small nuclear ribonucleoprotein G

Chain H:



- Molecule 8: Small nuclear ribonucleoprotein G

Chain X:



- Molecule 8: Small nuclear ribonucleoprotein G

Chain f:



- Molecule 8: Small nuclear ribonucleoprotein G

Chain n:



- Molecule 8: Small nuclear ribonucleoprotein G

Chain v:



- Molecule 8: Small nuclear ribonucleoprotein G

Chain 4: 

MET	SER	LYS	ALA	HE000	HE008	LE026	RE068	HE069	ARG	VAL
-----	-----	-----	-----	-------	-------	-------	-------	-------	-----	-----

4 Data and refinement statistics

EDS was not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	150.93Å 356.81Å 230.75Å 90.00° 97.31° 90.00°	Depositor
Resolution (Å)	59.47 – 3.10	Depositor
% Data completeness (in resolution range)	64.6 (59.47-3.10)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.04	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.37 (at 3.13Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, R_{free}	0.232 , 0.256	Depositor
Wilson B-factor (Å ²)	49.1	Xtriage
Anisotropy	0.038	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.39$, $\langle L^2 \rangle = 0.22$	Xtriage
Outliers	39 of 281573 reflections (0.014%)	Xtriage
Total number of atoms	42689	wwPDB-VP
Average B, all atoms (Å ²)	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 77.30 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 8.7504e-07. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.51	0/656	0.75	1/888 (0.1%)
1	I	0.51	0/656	0.74	1/888 (0.1%)
1	Q	0.49	0/656	0.76	1/888 (0.1%)
1	Y	0.49	0/656	0.75	1/888 (0.1%)
1	g	0.49	0/656	0.74	1/888 (0.1%)
1	o	0.51	0/656	0.75	1/888 (0.1%)
1	w	0.46	0/656	0.74	1/888 (0.1%)
2	B	0.52	0/817	0.70	0/1096
2	J	0.50	0/817	0.70	0/1096
2	R	0.48	0/817	0.69	0/1096
2	Z	0.47	0/817	0.68	0/1096
2	h	0.47	0/817	0.68	0/1096
2	p	0.47	0/817	0.67	0/1096
2	x	0.48	0/817	0.69	0/1096
3	C	0.43	0/646	0.75	1/867 (0.1%)
3	K	0.46	0/646	0.76	1/867 (0.1%)
3	S	0.44	0/646	0.75	1/867 (0.1%)
3	a	0.40	0/646	0.74	1/867 (0.1%)
3	i	0.41	0/646	0.74	1/867 (0.1%)
3	q	0.43	0/646	0.74	1/867 (0.1%)
3	y	0.42	0/646	0.74	1/867 (0.1%)
4	D	0.39	0/567	0.65	0/765
4	L	0.39	0/567	0.65	0/765
4	T	0.40	0/567	0.64	0/765
4	b	0.40	0/567	0.64	0/765
4	j	0.39	0/567	0.64	0/765
4	r	0.39	0/567	0.67	0/765
4	z	0.38	0/567	0.64	0/765
5	l	0.65	0/135	0.90	0/184
5	E	0.57	0/135	0.88	0/184
5	M	0.63	0/135	0.89	0/184
5	U	0.56	0/135	0.83	0/184

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
5	c	0.64	0/135	0.94	0/184
5	k	0.64	0/135	0.91	0/184
5	s	0.55	0/135	0.87	0/184
6	2	0.56	0/1830	0.88	2/2489 (0.1%)
6	F	0.58	0/1830	0.89	2/2489 (0.1%)
6	N	0.60	0/1830	0.91	2/2489 (0.1%)
6	V	0.56	0/1830	0.87	2/2489 (0.1%)
6	d	0.53	0/1830	0.86	2/2489 (0.1%)
6	l	0.58	0/1830	0.89	2/2489 (0.1%)
6	t	0.60	0/1830	0.89	2/2489 (0.1%)
7	3	0.46	0/1008	0.74	0/1373
7	G	0.57	0/1008	0.76	0/1373
7	O	0.55	0/1008	0.75	0/1373
7	W	0.53	0/1008	0.75	0/1373
7	e	0.47	0/1008	0.72	0/1373
7	m	0.57	0/1008	0.75	0/1373
7	u	0.55	0/1008	0.75	0/1373
8	4	0.40	0/551	0.72	0/737
8	H	0.42	0/551	0.73	0/737
8	P	0.42	0/551	0.73	0/737
8	X	0.42	0/551	0.73	0/737
8	f	0.39	0/551	0.72	0/737
8	n	0.42	0/551	0.73	0/737
8	v	0.41	0/551	0.73	0/737
All	All	0.51	0/43470	0.78	28/58793 (0.0%)

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
7	3	0	1
7	G	0	1
7	O	0	1
7	W	0	1
7	e	0	1
7	m	0	1
7	u	0	1
All	All	0	7

There are no bond length outliers.

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	N	8185	LYS	C-N-CA	7.12	139.50	121.70
6	F	8185	LYS	C-N-CA	7.10	139.44	121.70
6	2	8185	LYS	C-N-CA	7.08	139.40	121.70
6	l	8185	LYS	C-N-CA	7.01	139.21	121.70
6	t	8185	LYS	C-N-CA	7.01	139.22	121.70

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
7	G	6067	ASN	Mainchain
7	O	6067	ASN	Mainchain
7	W	6067	ASN	Mainchain
7	e	6067	ASN	Mainchain
7	m	6067	ASN	Mainchain

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	648	0	690	6	0
1	I	648	0	690	10	0
1	Q	648	0	690	10	0
1	Y	648	0	690	7	0
1	g	648	0	690	0	0
1	o	648	0	690	0	0
1	w	648	0	690	0	0
2	B	807	0	833	13	0
2	J	807	0	833	18	0
2	R	807	0	833	19	0
2	Z	807	0	833	9	0
2	h	807	0	833	0	0
2	p	807	0	833	0	0
2	x	807	0	833	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	638	0	657	13	0
3	K	638	0	657	15	0
3	S	638	0	657	13	0
3	a	638	0	657	0	0
3	i	638	0	657	0	0
3	q	638	0	657	0	0
3	y	638	0	657	0	0
4	D	556	0	561	16	0
4	L	556	0	561	17	0
4	T	556	0	561	17	0
4	b	556	0	561	0	0
4	j	556	0	561	0	0
4	r	556	0	561	0	0
4	z	556	0	561	0	0
5	1	133	0	0	0	0
5	E	133	0	123	2	0
5	M	133	0	123	2	0
5	U	133	0	123	2	0
5	c	133	0	123	0	0
5	k	133	0	123	0	0
5	s	133	0	123	0	0
6	2	1787	0	0	1	0
6	F	1787	0	1779	50	0
6	N	1787	0	1779	57	0
6	V	1787	0	1779	54	0
6	d	1787	0	1779	0	0
6	l	1787	0	1779	0	0
6	t	1787	0	1779	0	0
7	3	984	0	0	6	0
7	G	984	0	943	21	0
7	O	984	0	943	19	0
7	W	984	0	943	19	0
7	e	984	0	943	0	0
7	m	984	0	943	0	0
7	u	984	0	943	0	0
8	4	544	0	0	1	0
8	H	544	0	563	7	0
8	P	544	0	563	7	0
8	X	544	0	563	8	0
8	f	544	0	563	0	0
8	n	544	0	563	0	0
8	v	544	0	563	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	2	5	0	0	0	0
9	t	5	0	0	0	0
All	All	42689	0	39635	370	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 9.

The worst 5 of 370 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:R:2043:LYS:HG2	2:R:2067:TRP:HB3	1.49	0.94
7:O:6067:ASN:HB2	7:O:6068:PRO:HD3	1.49	0.92
2:R:2092:LEU:HD23	4:T:4056:VAL:HG22	1.51	0.91
2:B:2092:LEU:HD23	4:D:4056:VAL:HG22	1.51	0.91
2:J:2092:LEU:HD23	4:L:4056:VAL:HG22	1.52	0.90

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	80/119 (67%)	76 (95%)	2 (2%)	2 (2%)	9	42
1	I	80/119 (67%)	76 (95%)	2 (2%)	2 (2%)	9	42
1	Q	80/119 (67%)	76 (95%)	2 (2%)	2 (2%)	9	42
1	Y	80/119 (67%)	76 (95%)	2 (2%)	2 (2%)	9	42
1	g	80/119 (67%)	76 (95%)	2 (2%)	2 (2%)	9	42
1	o	80/119 (67%)	77 (96%)	1 (1%)	2 (2%)	9	42
1	w	80/119 (67%)	77 (96%)	1 (1%)	2 (2%)	9	42
2	B	96/118 (81%)	92 (96%)	4 (4%)	0	100	100
2	J	96/118 (81%)	90 (94%)	6 (6%)	0	100	100
2	R	96/118 (81%)	91 (95%)	5 (5%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	Z	96/118 (81%)	92 (96%)	4 (4%)	0	100	100
2	h	96/118 (81%)	90 (94%)	3 (3%)	3 (3%)	7	36
2	p	96/118 (81%)	91 (95%)	5 (5%)	0	100	100
2	x	96/118 (81%)	92 (96%)	3 (3%)	1 (1%)	22	68
3	C	75/92 (82%)	73 (97%)	2 (3%)	0	100	100
3	K	75/92 (82%)	73 (97%)	2 (3%)	0	100	100
3	S	75/92 (82%)	73 (97%)	2 (3%)	0	100	100
3	a	75/92 (82%)	73 (97%)	2 (3%)	0	100	100
3	i	75/92 (82%)	73 (97%)	2 (3%)	0	100	100
3	q	75/92 (82%)	73 (97%)	2 (3%)	0	100	100
3	y	75/92 (82%)	73 (97%)	2 (3%)	0	100	100
4	D	69/86 (80%)	66 (96%)	3 (4%)	0	100	100
4	L	69/86 (80%)	66 (96%)	3 (4%)	0	100	100
4	T	69/86 (80%)	67 (97%)	2 (3%)	0	100	100
4	b	69/86 (80%)	66 (96%)	3 (4%)	0	100	100
4	j	69/86 (80%)	67 (97%)	2 (3%)	0	100	100
4	r	69/86 (80%)	66 (96%)	3 (4%)	0	100	100
4	z	69/86 (80%)	66 (96%)	3 (4%)	0	100	100
5	1	15/124 (12%)	11 (73%)	3 (20%)	1 (7%)	2	14
5	E	15/124 (12%)	11 (73%)	3 (20%)	1 (7%)	2	14
5	M	15/124 (12%)	11 (73%)	3 (20%)	1 (7%)	2	14
5	U	15/124 (12%)	11 (73%)	3 (20%)	1 (7%)	2	14
5	c	15/124 (12%)	9 (60%)	5 (33%)	1 (7%)	2	14
5	k	15/124 (12%)	10 (67%)	4 (27%)	1 (7%)	2	14
5	s	15/124 (12%)	11 (73%)	3 (20%)	1 (7%)	2	14
6	2	214/247 (87%)	187 (87%)	17 (8%)	10 (5%)	4	23
6	F	214/247 (87%)	182 (85%)	22 (10%)	10 (5%)	4	23
6	N	214/247 (87%)	174 (81%)	31 (14%)	9 (4%)	4	27
6	V	214/247 (87%)	182 (85%)	21 (10%)	11 (5%)	3	22
6	d	214/247 (87%)	183 (86%)	20 (9%)	11 (5%)	3	22
6	l	214/247 (87%)	183 (86%)	21 (10%)	10 (5%)	4	23

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
6	t	214/247 (87%)	182 (85%)	21 (10%)	11 (5%)	3	22
7	3	123/186 (66%)	114 (93%)	5 (4%)	4 (3%)	6	33
7	G	123/186 (66%)	114 (93%)	5 (4%)	4 (3%)	6	33
7	O	123/186 (66%)	114 (93%)	5 (4%)	4 (3%)	6	33
7	W	123/186 (66%)	114 (93%)	5 (4%)	4 (3%)	6	33
7	e	123/186 (66%)	114 (93%)	5 (4%)	4 (3%)	6	33
7	m	123/186 (66%)	114 (93%)	5 (4%)	4 (3%)	6	33
7	u	123/186 (66%)	114 (93%)	5 (4%)	4 (3%)	6	33
8	4	68/76 (90%)	64 (94%)	3 (4%)	1 (2%)	15	57
8	H	68/76 (90%)	64 (94%)	3 (4%)	1 (2%)	15	57
8	P	68/76 (90%)	64 (94%)	3 (4%)	1 (2%)	15	57
8	X	68/76 (90%)	64 (94%)	3 (4%)	1 (2%)	15	57
8	f	68/76 (90%)	64 (94%)	3 (4%)	1 (2%)	15	57
8	n	68/76 (90%)	64 (94%)	3 (4%)	1 (2%)	15	57
8	v	68/76 (90%)	64 (94%)	3 (4%)	1 (2%)	15	57
All	All	5180/7336 (71%)	4740 (92%)	308 (6%)	132 (2%)	9	42

5 of 132 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	M	7006	LEU
6	N	8028	ILE
6	N	8185	LYS
6	N	8186	GLU
6	N	8212	LYS

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	77/101 (76%)	70 (91%)	7 (9%)	14	45
1	I	77/101 (76%)	72 (94%)	5 (6%)	24	65

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	Q	77/101 (76%)	72 (94%)	5 (6%)	24	65
1	Y	77/101 (76%)	71 (92%)	6 (8%)	18	57
1	g	77/101 (76%)	71 (92%)	6 (8%)	18	57
1	o	77/101 (76%)	73 (95%)	4 (5%)	32	73
1	w	77/101 (76%)	72 (94%)	5 (6%)	24	65
2	B	94/110 (86%)	90 (96%)	4 (4%)	40	80
2	J	94/110 (86%)	92 (98%)	2 (2%)	66	92
2	R	94/110 (86%)	91 (97%)	3 (3%)	51	87
2	Z	94/110 (86%)	90 (96%)	4 (4%)	40	80
2	h	94/110 (86%)	89 (95%)	5 (5%)	32	72
2	p	94/110 (86%)	91 (97%)	3 (3%)	51	87
2	x	94/110 (86%)	90 (96%)	4 (4%)	40	80
3	C	72/84 (86%)	67 (93%)	5 (7%)	22	62
3	K	72/84 (86%)	68 (94%)	4 (6%)	30	70
3	S	72/84 (86%)	68 (94%)	4 (6%)	30	70
3	a	72/84 (86%)	68 (94%)	4 (6%)	30	70
3	i	72/84 (86%)	68 (94%)	4 (6%)	30	70
3	q	72/84 (86%)	68 (94%)	4 (6%)	30	70
3	y	72/84 (86%)	69 (96%)	3 (4%)	40	81
4	D	60/74 (81%)	56 (93%)	4 (7%)	23	63
4	L	60/74 (81%)	57 (95%)	3 (5%)	34	75
4	T	60/74 (81%)	56 (93%)	4 (7%)	23	63
4	b	60/74 (81%)	56 (93%)	4 (7%)	23	63
4	j	60/74 (81%)	56 (93%)	4 (7%)	23	63
4	r	60/74 (81%)	56 (93%)	4 (7%)	23	63
4	z	60/74 (81%)	57 (95%)	3 (5%)	34	75
5	l	15/97 (16%)	11 (73%)	4 (27%)	1	2
5	E	15/97 (16%)	12 (80%)	3 (20%)	2	8
5	M	15/97 (16%)	12 (80%)	3 (20%)	2	8
5	U	15/97 (16%)	13 (87%)	2 (13%)	6	22
5	c	15/97 (16%)	12 (80%)	3 (20%)	2	8

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	k	15/97 (16%)	13 (87%)	2 (13%)	6	22
5	s	15/97 (16%)	12 (80%)	3 (20%)	2	8
6	2	203/231 (88%)	176 (87%)	27 (13%)	6	22
6	F	203/231 (88%)	170 (84%)	33 (16%)	3	12
6	N	203/231 (88%)	171 (84%)	32 (16%)	4	14
6	V	203/231 (88%)	177 (87%)	26 (13%)	6	24
6	d	203/231 (88%)	178 (88%)	25 (12%)	7	26
6	l	203/231 (88%)	176 (87%)	27 (13%)	6	22
6	t	203/231 (88%)	172 (85%)	31 (15%)	4	15
7	3	108/159 (68%)	98 (91%)	10 (9%)	13	44
7	G	108/159 (68%)	98 (91%)	10 (9%)	13	44
7	O	108/159 (68%)	99 (92%)	9 (8%)	16	53
7	W	108/159 (68%)	99 (92%)	9 (8%)	16	53
7	e	108/159 (68%)	99 (92%)	9 (8%)	16	53
7	m	108/159 (68%)	99 (92%)	9 (8%)	16	53
7	u	108/159 (68%)	99 (92%)	9 (8%)	16	53
8	4	61/66 (92%)	59 (97%)	2 (3%)	50	86
8	H	61/66 (92%)	59 (97%)	2 (3%)	50	86
8	P	61/66 (92%)	59 (97%)	2 (3%)	50	86
8	X	61/66 (92%)	59 (97%)	2 (3%)	50	86
8	f	61/66 (92%)	59 (97%)	2 (3%)	50	86
8	n	61/66 (92%)	59 (97%)	2 (3%)	50	86
8	v	61/66 (92%)	59 (97%)	2 (3%)	50	86
All	All	4830/6454 (75%)	4413 (91%)	417 (9%)	15	51

5 of 417 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	Y	1064	ARG
7	e	6033	THR
6	2	8184	LEU
3	a	3016	ARG
6	d	8073	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 49 such sidechains are listed below:

Mol	Chain	Res	Type
2	Z	2031	ASN
7	e	6018	ASN
2	x	2031	ASN
3	a	3024	GLN
1	g	1047	ASN

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

2 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	$\# Z > 2$
9	SO4	2	8301	-	4,4,4	0.25	0	6,6,6	0.38	0
9	SO4	t	8301	-	4,4,4	0.58	0	6,6,6	0.37	0

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
9	SO4	2	8301	-	-	0/0/0/0	0/0/0/0
9	SO4	t	8301	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.