

Apr 6, 2025 – 02:08 AM JST

PDB ID	:	$9L5R / pdb_{00009l5r}$
EMDB ID	:	EMD-62841
Title	:	Cryo-EM structure of the thermophile spliceosome (state ILS)
Authors	:	Li, Y.; Fischer, P.; Wang, M.; Yuan, R.; Meng, W.; Luehrmann, R.; Lau, B.;
		Hurt, E.; Cheng, J.
Deposited on	:	2024-12-23
Resolution	:	2.80 Å(reported)
	-	

This is a Full wwPDB EM Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org* A user guide is available at https://www.wwpdb.org/validation/2017/EMValidationReportHelp with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

EMDB validation analysis	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7(2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.42

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $ELECTRON\ MICROSCOPY$

The reported resolution of this entry is 2.80 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f EM\ structures}\ (\#{f Entries})$		
Ramachandran outliers	207382	16835		
Sidechain outliers	206894	16415		
RNA backbone	6643	2191		

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain								
1	2	193	45% 9% ·	45%							
2	5	116	75%	2	0% • •						
3	6	101	60%	36%	•••						
4	А	2463	81%	•	19%						
5	В	326	78%	•	21%						
6	С	1011	91%		9%						
7	D	325	29% • 71%								
8	Е	352	87%		• 12%						
9	F	233	46% •	53%							



Conti	nued fron	ı previous	page	
Mol	Chain	Length	Quality of chain	
10	Ι	839	87%	12%
11	J	687	88%	11%
12	L	768	82%	• 17%
13	К	231	99%	
14	q	480	29% 71%	
14	r	480	29% 70%	
14	s	480	29% 71%	
14	t	480	29% 71%	
15	N	148	97%	••
16	S	167	93%	• 6%
17	Т	496	64%	34%
18	U	757	45% ·	54%
19	М	395	62%	37%
20	0	408	65%	32%
21	R	578	58% •	41%
22	W	547	72%	27%
23	Р	260	43%	6%
24	Y	1416	95%	5%
25	a	98	88%	12%
25	j	98	90%	10%
26	b	94	76%	24%
26	1	94	85%	• 14%
27	с	592	14% 86%	
27	m	592	14% 85%	
28	d	118	69%	31%



Mol	Chain	Length	Quality of chai	n	
28	0	118	72%	•	26%
29	е	211	37%	63%	
29	р	211	48%	51%	
30	f	114	74%		26%
30	u	114	79%		21%
31	g	82	88%		12%
31	k	82	88%		• 11%
32	h	242	50%	50%	
33	i	201	51%	49%	
34	CY	510	18% 82%		
35	Cb	975	52%	48%	
36	Cc	764	93%		7%
37	7	101	• 21% • 74	1%	
38	Ci	153	55%	45%	
39	Н	22	100%		



2 Entry composition (i)

There are 44 unique types of molecules in this entry. The entry contains 104144 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	2	106	Total 2224	C 994	N 357	O 767	Р 106	0	0

• Molecule 2 is a RNA chain called U5 snRNA.

Mol	Chain	Residues		A	AltConf	Trace			
2	5	111	Total 2343	C 1048	N 398	0 786	Р 111	0	0

• Molecule 3 is a RNA chain called U6 snRNA.

Mol	Chain	Residues		A	toms	AltConf	Trace		
3	6	98	Total 2090	C 935	N 379	O 678	Р 98	0	0

• Molecule 4 is a protein called PRP8.

Mol	Chain	Residues		At	AltConf	Trace			
4	А	2006	Total 16540	C 10628	N 2874	O 2976	S 62	0	0

• Molecule 5 is a protein called Pre-mRNA-splicing factor SYF2.

Mol	Chain	Residues		Ate	AltConf	Trace			
5	В	258	Total 2071	C 1273	N 388	0 405	$\frac{S}{5}$	0	0

• Molecule 6 is a protein called SNU114.

Mol	Chain	Residues		Α	AltConf	Trace			
6	С	922	Total 7201	C	N 1000	0	S	0	0
			7301	4668	1229	1368	36		



• Molecule 7 is a protein called SDE2.

Mol	Chain	Residues		At	oms			AltConf	Trace
7	D	95	Total 786	C 477	N 150	0 154	${f S}{5}$	0	0

• Molecule 8 is a protein called Anaphase-promoting complex subunit 4-like WD40 domaincontaining protein.

Mol	Chain	Residues		At	oms			AltConf	Trace
8	Е	310	Total 2379	C 1493	N 414	0 462	S 10	0	0

• Molecule 9 is a protein called CCDC12.

Mol	Chain	Residues		At	oms			AltConf	Trace
9	F	110	Total 879	С 544	N 166	O 167	${S \over 2}$	0	0

• Molecule 10 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues		Α	toms			AltConf	Trace
10	Ι	738	Total 6135	C 3924	N 1071	0 1112	S 28	0	0

• Molecule 11 is a protein called Suppressor of forked domain-containing protein.

Mol	Chain	Residues		At		AltConf	Trace		
11	J	608	Total 5176	C 3297	N 937	O 929	S 13	0	0

• Molecule 12 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues		At	AltConf	Trace			
12	L	637	Total 5052	C 3111	N 952	0 974	S 15	0	0

• Molecule 13 is a protein called Pre-mRNA-splicing factor SPF27.

Mol	Chain	Residues		Ate			AltConf	Trace	
13	K	231	Total 1797	C 1122	N 317	0 354	${S \atop 4}$	0	0

• Molecule 14 is a protein called Pre-mRNA-processing factor 19.



Mol	Chain	Residues		At	oms			AltConf	Trace
14	a	130	Total	С	Ν	0	S	0	0
14	q	159	1110	699	195	214	2	0	0
14	+	1.4.1	Total	С	Ν	0	S	0	0
14	U	141	1122	707	197	216	2	0	0
14	r	1/12	Total	С	Ν	0	S	0	0
14	1	140	1132	713	199	218	2	0	0
14	G	140	Total	С	Ν	0	S	0	0
14	a	140	1115	702	196	215	2		0

• Molecule 15 is a protein called Putative bud site selection protein.

Mol	Chain	Residues		\mathbf{A}	toms			AltConf	Trace
15	Ν	148	Total 1200	C 755	N 213	O 220	S 12	0	0

• Molecule 16 is a protein called Peptidyl-prolyl cis-trans isomerase.

Mol	Chain	Residues		At	oms		AltConf	Trace	
16	S	157	Total 1209	C 763	N 217	0 223	S 6	0	0

• Molecule 17 is a protein called Pre-mRNA-splicing factor PRP46.

Mol	Chain	Residues		At	oms			AltConf	Trace
17	Т	328	Total 2565	C 1621	N 461	O 469	S 14	0	0

• Molecule 18 is a protein called Cell cycle control protein.

Mol	Chain	Residues		At	AltConf	Trace			
18	U	348	Total 2821	C 1770	N 516	0 524	S 11	0	0

• Molecule 19 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues	Atoms			AltConf	Trace		
19	М	248	Total 1964	C 1238	N 355	0 354	S 17	0	0

• Molecule 20 is a protein called Putative pre-mRNA splicing protein.



Mol	Chain	Residues		Atoms			AltConf	Trace	
20	0	276	Total 2224	C 1381	N 424	0 412	S 7	0	0

• Molecule 21 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues		A	toms	5			AltConf	Trace
21	R	340	Total 2686	C 1664	N 509	O 503	Р 2	S 8	0	0

• Molecule 22 is a protein called PRP17.

Mol	Chain	Residues		Ate	oms			AltConf	Trace
22	W	397	Total 2224	C 1336	N 444	0 440	$\frac{S}{4}$	0	0

• Molecule 23 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues		At	oms			AltConf	Trace
23	Р	114	Total 925	C 577	N 182	0 165	S 1	0	0

• Molecule 24 is a protein called Pre-mRNA-splicing factor.

Mol	Chain	Residues	Atoms				AltConf	Trace	
24	Y	1344	Total 10819	C 6892	N 1900	O 2002	$\begin{array}{c} \mathrm{S} \\ \mathrm{25} \end{array}$	0	0

• Molecule 25 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms	AltConf	Trace
25	a	86	Total C N O 426 253 86 87	0	0
25	j	88	Total C N O S 704 456 121 126 1	0	0

• Molecule 26 is a protein called Sm protein F.

Mol	Chain	Residues	Atoms	AltConf	Trace
26	b	71	Total C N O 351 209 71 71	0	0
26	1	81	Total C N O S 649 412 114 120 3	0	0



• Molecule 27 is a protein called Delta(14)-sterol reductase.

Mol	Chain	Residues	Atoms	AltConf	Trace
27	с	84	Total C N O 416 247 84 85	0	0
27	m	86	Total C N O S 678 427 129 118 4	0	0

• Molecule 28 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms	AltConf	Trace
28	d	81	Total C N O 401 239 81 81	0	0
28	О	87	Total C N O S 679 433 114 128 4	0	0

• Molecule 29 is a protein called Sm protein B.

Mol	Chain	Residues	Atoms	AltConf	Trace
29	е	79	Total C N O 388 230 79 79	0	0
29	р	103	Total C N O S 788 490 148 145 5	0	0

• Molecule 30 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms	AltConf	Trace
30	f	84	Total C N O 414 246 84 84	0	0
30	u	90	Total C N O S 716 449 132 131 4	0	0

• Molecule 31 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms	AltConf	Trace
31	g	72	Total C N O 354 210 72 72	0	0
31	k	73	Total C N O S 577 369 101 105 2	0	0

• Molecule 32 is a protein called U2 small nuclear ribonucleoprotein A'.



Mol	Chain	Residues		Ato	\mathbf{ms}		AltConf	Trace
32	h	122	Total 605	C 361	N 122	0 122	0	0

• Molecule 33 is a protein called U2 small nuclear ribonucleoprotein B'-like protein.

Mol	Chain	Residues		Ato	ms		AltConf	Trace
33	i	102	Total 504	C 300	N 102	O 102	0	0

• Molecule 34 is a protein called Nineteen complex-related protein 2-domain-containing protein.

Mol	Chain	Residues		Aton	ns		AltConf	Trace
34	CY	91	Total 444	C 262	N 91	0 91	0	0

• Molecule 35 is a protein called G-patch domain-containing protein.

Mol	Chain	Residues		Ator	ns		AltConf	Trace
35	Cb	505	Total 2496	C 1486	N 505	O 505	0	0

• Molecule 36 is a protein called RNA helicase.

Mol	Chain	Residues		Ator	ns		AltConf	Trace
36	Cc	714	Total 3541	C 2112	N 714	O 715	0	0

• Molecule 37 is a RNA chain called Unknown mRNA.

Mol	Chain	Residues		At	oms			AltConf	Trace
37	7	26	Total 494	C 241	N 68	O 159	Р 26	0	0

• Molecule 38 is a protein called Putative cyclophilin protein.

Mol	Chain	Residues		Aton	ıs		AltConf	Trace
38	Ci	84	Total 415	C 247	N 84	0 84	0	0

• Molecule 39 is a protein called Unknown protein.



Mol	Chain	Residues		Ator	\mathbf{ns}		AltConf	Trace
39	Н	22	Total 110	C 66	N 22	O 22	0	0

• Molecule 40 is N,N,7-trimethylguanosine 5'-(trihydrogen diphosphate) (CCD ID: M7M) (formula: $C_{13}H_{23}N_5O_{11}P_2$).



Mol	Chain	Residues		Ate	oms			AltConf
40	Р	1	Total	С	Ν	Ο	Р	0
40	D	1	30	13	5	10	2	0

• Molecule 41 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	AltConf
41	С	1	Total Mg 1 1	0

• Molecule 42 is GUANOSINE-5'-TRIPHOSPHATE (CCD ID: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).





Mol	Chain	Residues		Ate	oms			AltConf
42	С	1	Total	С	Ν	0	Р	0
			32	10	5	14	3	

• Molecule 43 is INOSITOL HEXAKISPHOSPHATE (CCD ID: IHP) (formula: $C_6H_{18}O_{24}P_6$).



Mol	Chain	Residues	Α	tor	ns		AltConf
/3	T	1	Total	С	Ο	Р	0
40	5	I	36	6	24	6	0

 $\bullet\,$ Molecule 44 is ZINC ION (CCD ID: ZN) (formula: Zn).



Mol	Chain	Residues	Atoms	AltConf
44	Ν	3	Total Zn 3 3	0
44	U	1	Total Zn 1 1	0
44	М	2	Total Zn 2 2	0



3 Residue-property plots (i)

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



• Molecule 1: U2 snRNA



L1728 L1728 L18989 L1963 L1963 R2035 R205 R205 R205 R205 R205 R205 R205 R20	GLN GLN ALA ALA ALA CLU CLU GLN GLN GLN GLN GLN	LEU THR ALA ALA ALA THR LYS CLY GLN ASN VAL	GLY GLU GLU GLU ILE ILE THE THR THR SER SER SER	PHE GLU GLN CHLN THR PHE ALA SER LYS
THR GLU TRP ARG ARG ALA ALA ALA ALA SER SER ALEU	ARG THR ALA ALA ASN MET MET ASN PRO PRO ASN	ASF ASP ASP ASP THR TTR TTR MET PRO LYS LYS LSN	LEU LYS LYS PHE TLE TLE ALA ALA ALA ALA ALA ARG	VAL GLN VAL ALA GLY PHE LEU TYR GLY
CYS SER PRO ASP ASP ASP CLN VAL LYS GLU TLE ARC	CYS TLE WAL WAL VAL PRO GLY GLY ARG GLY ARG SER VAL	LEU PRO CLN CLN CLN CLN CLN CLU CLU CLU CLU CLU	GLY LEU GLU PRO FRO LEU CLY LEU TLE HIS THR	ALA GLY ASN CLU CLU LEU PRO PRO MET SER
PR0 ALA ASP ASP ASP THR THR HIS LVS LVS LEU VAL ALA	HIS PRO PRO SER TRP LYS ASN ASN ASN ASN THR THR THR THR VAL	PALA THR PRO GLY SER VAL SER LEU SER LEU SER ALA ALA	LEU THR PRO LEU LEU LEU LYS TYR CLY VAL CLU	ASN LYS ASP PRO ASN VAL ASP ASN PRO
GLN GLY PHE THR THR THR MET GLV GLV GLV ARG GLN LEU	LEU LEU SER ASP ASP LYS PHE LYS PHE LEU VAL PRO PRO	LYS LYS LYS ASN ASN TYR ASN PHE GLY SER SER SER	SER GLY ILE GLU LYS PRO TYR HIS VAL	ASP ASP PRO LEU PRO PRO PRO PRO SER
GLU GLU HIS ARG PRO TLE TLE THR SER PHE SER PHE CLU	LEU GLU GLU ALF TRP VAL ASP ASP ASP ASP ASP ASN ALA			
• Molecule 5: Pre	e-mRNA-splicing fac	tor SYF2		
Chain B:	78%		• 21%	
MET PRO PRO PRO LYS LYS LYS LYS LYS CLY GLU SER GLN GLN GLN	PRO PRO PRO PRO PRO PRO PRO PRO CLN CLN CLN CLN CLN CLN CLN CLN CLN CLN	GLU GLU GLU GLU PRO THR ASP SER SER FU GLU	VAL GLU SER ARG SER ASP ASP LYS VAL VAL	GLN SER PRO THR LEU ALA ASP PRO
SER ASP ASP THR SER GLU GLU S69 411 11 0265 0265	32 0			
• Molecule 6: SN	U114			
• Molecule 6: SN Chain C:	U114	91%		9%
• Molecule 6: SN Chain C:	U114 U114 Multiple Mu	ALM TIM ALA ALA ASP ASP ASP ASP ASP ASP PHE	GLY ARS PRIS PRIS TILE GLU GLU SER SER GLU SER GLU	GUU SER ASP ASP GLY ALA ALA ALA
Molecule 6: SN Chain C:	U1114 U1114 U1114 U1114 U114 U14 U14 U15 U15 U15 U15 U15 U15 U15 U15 U15 U15	RRU BER NS27 NS27 NS27 NS27 ASP NS53 ASP ASP ASP ASP ASP ASP ASP ASP PHE PHE	P1011 CLY PHE PHE TLE CLU GLU ALA SER SER CLU	GUU SER ASP ASP GLY ASN ALA ALA
 Molecule 6: SN Chain C: Chain C:	U114 U114 H M M M M M M M M M M M M M M M M M M M	SER SER 100 110 110 110 110 110 110 11	P1011 GLY ASN ASN ASN ALL CLU GLU SSR SSR GLU	GUU SER ASP ASP ALA ALA ALA ALA
 Molecule 6: SN Chain C: Chain S: Chain S: Molecule 7: SD Chain D: 		91% 91% 91% 91% 91% 91% 91% 91%	PIOI1 GLY ASN ASN ASN ALL GLU GLU SER SER GLU SER	010 SER ASP ASP ALA ALA ALA
 Molecule 6: SN Chain C: Chain C: Chain S: Chain D: Chain D: 	U114 H 10 H H H H 4 V H H H 4 V H H 10 H H H H 4 V H H H 4 V H H 10 H H H H H 4 V H H H 4 V H H 10 H H H H H 4 V H H H 4 V H H 10 H H H H H 4 V H H H H 4 V H H 10 H H H H H 4 V H H H H 4 V H H 10 H H H H H 4 V H H H H 4 V H H 10 H H H H H 4 V H H H H 1 V H H H H 1 V H E 2 29% • H 10 H H H H 1 V H H H 1 V H H H 1 V H H 10 H H H H 1 V H H H H 1 V H H H H 1 V H H 10 H H H H H 1 V H H H H H 1 V H H 10 H H H H H 1 V H H H H H H H H H H H H H	91% 91% 91% 91% 91% 91% 91% 91% 91% 91%	ASP P1011 CLY SER ARG LEU PRO PRO THR CLU TRE PHE TLEU TRE PRO CLU CLU CLU CLU CLU CLU CLU CLU CLU CLU	THR GUU GUU GUU ARC LEU LEU ARC ASP LEU ARS ASP LEU ASP ASP LEU ASP HIS LEU ASN THR ASN ALA ALA ALA ALA ALA ALA ALA ALA ALA AL
 Molecule 6: SN Chain C: Chain C: Chain 2: Chain 2: Chain D: Chain D: Chain D: Chain D: Chain D: Chain D: 	Image:	ANAL ASP FRO CLU RRA ASP SER ASP SER ASP SER ALA ASP ALA ASP ASP LEU VAL ASS ASP ASP ASP LEU VAL ASS ASP ASP ASP LEU ARC ASP ASP ASP LEU ASP	ARG ASP Piol1 CLY LEU SER PIE PHE ALA ARG PL PL VAL LEU FRO PRO PL VAL LEU FRO PL PL ULU THR CLU CLU CLU LEU THR CLU CLU CLU LEU THR CLU CLU CLU LEU TYR CLU CLU CLU CLY ALA CLU CLU CLU LYS PRO SER CLU CLU	PHETERGUUGLYSERARGGLYARGARGGLNLEUARGCLNLEUASPCLNLEUARGTHRARGTHRALATHRALALEUALALEUALALEUALALEU
 Molecule 6: SN Chain C: Chain C: Chain C: Chain C: Chain C: Chain D: <	Way W	91% 51% 51% 51% 51% 51% 51% 51% 5	ARC ARC ASP P1011 CLY THR LED SER P40.1 CLY VLL ALA ARG P41. P41. SER LEU TRG CLU CLU SER LEU TRR CLU CLU SER LEU TRR CLU CLU SER LU KLA ALA CLU SER LU CLU CLU CLU GLU TTR CLU GLU CLU	ALA GLY SER PHE TRR CLU SER CLU GLY ARC CLU SER LEU ASP CLU SER LEU ASP CLU SER LEU ASP CLY ARC ASP CLY ARC CLU LEU CLU LEU ALA THR ALA ARA ALA THR ALA ALA ALA ALA ALA ALA ALA ALA ALA AL

• Molecule 8: Anaphase-promoting complex subunit 4-like WD40 domain-containing protein



Chain E:		87%		• 12%	
MET SER GLU GLU LYS ARG ARG ARA ASP ASP ASP PRO GLU GLU CLV	LEU VAL VAL LYS LYS ARG GLN ASN VAL GLY SER SER	ARG ALA LEU THR ARG GLY ASP SER ALA	GLY ALA LEU ILEU THR T43 D92 D92 N95	R176 N229 R259 T352	
• Molecule 9: CCD0	C12				
Chain F:	46%	·	53%		
MET SER SER SER SER SER SER SER ALA ALA ALA ALA ALA ALA ALA ALA ALA	ALA ARG ARG LEU LEU LEU ALA SER EEU LYS SER ARG	LYS GLN GLU ASP ASP HTS PRO PRO SER	GLN ASP GLN GLN GLU GLU SER ALLA GLU GLU	SER LYS CLN GLU GLU ASP SER SER	GLN ALA S58 A98
R120 LYS LYS ALA ALA ALA ALA ALA ALA ALA ALA ALA AL	GLY ALA ALA ALA ASP GLY GLY GLY GLY GLY	ASN ASN GLY GLV GLU GLU VAL LEU LEU	GLU GLY GLY GLU TLE CLU GLY MET MET	ARG GLU GLU GLU GLU GLU GLU GLU	ALA ARG ARG GLU LYS
GLU ALA GLU GLU GLU PHE GLU THR ALA ALA					
• Molecule 10: Puta	tive pre-mRNA	splicing prot	ein		
Chain I:		87%		12%	
MET PRO SER LEU LEU LEU ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP	VAL VAL CYS CYS CYS C410 C416 C416 C416 C416 C416	P558 W727 MET LEU LYS GLN GLN	NTS CTN CTN CTN CTN CTN CTN CTN CTN CTN CTN	LEU GLN GLN GLN ASN ASN	SER SER GLY ALA GLY
ASP PRO GLU VAL ASP ALA ALA ALA GLU GLU GLU GLU GLU	ARG ALA ALA ALA ALA ALA ALA SER SER GLU	ASSN 1LE LYS GLY THR TLE ALA PRO SER THR	GLN SER VAL GLU VAL SER ASN PRO ALA ALA	ASP LEU ASP ASP MET ASP GLU	
• Molecule 11: Supp	pressor of forked	domain-cont	aining protein		
Chain J:		88%		11%	
MET GLU GLU SER SER ARG CLY PRO PRO PRO PRO M12 CLI CLU CLU	THR TILE TILE TILE TILE PRO PRO PRO PRO PRO PRO PRO PRO PRO PRO	USZ MET PRO ARG GLU VAL LYS LYS LYS LYS LYS	ARG ILLE ASP GLU ASP THR THR THR GLU CLU VAL	ASP TYR VAL PHE PRO ASP GLU	GLN GLN THR LYS ASN
LEU SER ASN LEU LEU LEU ALA ALA ALA CYS GLN CYS GLU CYS	617 617 617 617 617 617 617 617 617 617	N. C.			
• Molecule 12: Puta	tive pre-mRNA	splicing prot	ein		
Chain L:	8	2%	·	17%	
MET PRU VAL V4 V4 V210 C200 C300 C300 C300 C300 C300 C300 C30	GLU GLU GLU GLU ALA ALA ALA CLV CLVS LVS LVS ARG LVS	ASN LYS LYS ASP ASP C282 S343 GLU ASN ASN	ALA THR ARC ARC GLY LEU VAL ASN ASN SET STHR STHR	LEU ASN THR G361 9390 A405	GLY THR THR GLY PHE
GLU SER VAL VAL PRO LYS CLYS CLYS CLYS CLYS CLYS CLYS CLYS	LEU LEU ALA PRO LEU LEU GLY PRO GLY CLY	GLY ALA PRO LEU ARG PRO GLY GLY THR	PRO LEU ARG PRO PRO ARG ARG ARG ARG CEU CEU	ASN ALA ALA ASP GLY VAL SER MET	GLU HIS ALA LEU ARG
GLN GLN LEU LYS LYS GLN GLN ALA ALA ALA ALA ALA PRO PRO PRO CUTS CHI	THR CLU CLU CLU CLU CLU LEU CLU CLU CLU CLU CLU CLU	L505 L539 P540 R541 L668 L686	A7 68		

• Molecule 13: Pre-mRNA-splicing factor SPF27



Chain K:	99% .	
M1 K36 K80 N231		
• Molecule 14:	Pre-mRNA-processing factor 19	
Chain q:	29% 71%	
M1 K25 A139 A139 A13 A13 A13 A13 A13 A14 A14 A14 A14 A14 A14 A14 A14 A14 A14	GLY MET ALA ALA ALA ALA ALA ASP SER ASP CLU CLU CLU CLU CLU ASP ASP ASP ASP ASP ASP ASP ASP ASP CLU ASP ASP CLU ASP CLU ASP CLU ASP CLU CLU CLU CLU CLU CLU CLU CLU CLU CLU	NHD
GLN VAL TYR TYR THR ASP LEU VAL THR GLN	ALA ALA SER SER SER SER SER SER ASP CVS CVS CVS CVS CVS CVS CVS CVS CVS CVS	1
VAL VAL VAL ILE GLY ALA ALA LYS GLY VAL VAL	VAL VAL ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP	MINU.
VAL ALA ARG GLY TYR ALA ALA ALA LEU THR THR	THR THR ALM ALM ALM ALM ALM ALM ALM ALM ALM ALM	145
PHE TRP PHE ALA ALA ALA ALA GLY CYS GLY THR SER	SER VAL THR THR PHE ASP ASP ASP ALA ALA ALA ALA ALA ALA ALA ALA ALA AL	1 144
THR LYS ALA ALA THR LYS SER SER SER GLU PRO VAL	ARG LEEU MET PRO VAL VAL VAL VAL CYS GLY GLY GLY CLYS GLU VAL VAL VAL VAL VAL VAL VAL VAL VAL VAL	
• Molecule 14:	Pre-mRNA-processing factor 19	
Chain t:	29% 71%	
M1 P141 ALA ALA ALA GLU GLU ALD MET	ALA ALA ASP ASP ASP ASP ASP SER SER SER SER SER SER SER SER SER SER	1 1 1
THR ASP LEU ASN VAL THR GLN SER SER SER LEU	ASP ASP ASN ASN ASN ASN ASN ASN ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP	110
THR ALA LYS LYS GLY TRP VAL LYS VAL LYS VAL TYR ASP	0LY ARG CARG CARG CARG CARG CARG CARG CARG	145
TYR ALA ASP ALA ALA ALA ALA LEU THR THR CYS ALA ALA	HIS PRO ASP ASP ASP ASP ASN ALA ALA ALA ALA ALA ALA ALA ALA CLV CLV CLV CLV CLV CLV CLV CLV CLU CLV CLU CLV CLV CLV CLV CLV CLV CLV CLV CLV CLV	ULU U
ALA THR GLY GLY GLY GLY THR SER SER SER VAL THR THR	APHE APHE ARG LEU ARG CLY SER ALA ALA ALA ALA ALA ALA ALA ALA ALA AL	11114
LYS SER TRP SER GLU PRO PRO VAL ARG LEU GLY MET	PRO VAL VAL VAL VAL TRP GIT ALA ALA ALA ALA ALA CIT VAL VAL VAL VAL VAL VAL VAL VAL VAL VAL	
• Molecule 14:	Pre-mRNA-processing factor 19	
• Molecule 14: Chain r:	Pre-mRNA-processing factor 19 29% 70%	



ALA TYR THR ASP LEU ASN VAL	THR GLN ALA SER	SER LEU ASP LEU	GLU GLU GLU	CYS ALA ALA	VAL GLY GLY	LEU ASP	T EIL CTYS CLY	ASP ILE	TYR SER	VAL VAL ATA	ASN LYS	VAL GLU	THR LEU	ASP ILE	GLU	PRO VAL	THR ALA	GLU	TRP THR	GLY THR	LYS	VAL
ILE GLY GLY ALA ALA LYS GLY TRP	VAL LYS VAL TYR	ASP ALA GLY ARG	GLU SER ALA	THR PHE GLN	THR HIS ALA	GLY PRO	VAL THR	GLY LEU ALA	VAL HIS	PRO GLY CL V	ARG	LEU ALA	VAL GLY	VAL ASP	LYS SER	PHE VAL	PHE TYR	ASP	GLU THR	GLU	ARG	ALA
ARG GLY ALA ALA ALA ALA	LEU THR THR CYS	ALA PHE HIS PRO	ASP GLY ASN	LEU PHE ALA	ALA GLY THR	GLN THR	GLY GLY	LEU VAL	PHE HIS	THR THR	GLU	GLN	SER PHF.	PRO LEU	GLY THR	PRO ILE	GLN ALA	LEU ALA	PHE SER	GLU ASN	GLY	TRP
PHE ALA ALA THR GLY LYS GLY	THR SER SER VAL	THR ILE PHE ASP	LEU ARG LYS	SER GLY ALA	ALA ALA ALA	VAL LYS	GLU	THR GLY	GLU VAL	LEU SER	SER	ASP TYR	GLY GLN	TYR LEU	ALA THR	GLY GLY	GLY THR	GLY VAL	THR VAL	GLN MET	TYR	LYS
ALA THR LYS SER TRP SER GLU	PRO VAL ARG LEU	GLY MET PRO VAL	VAL GLY VAL	LYS TRP GLY	GLY GLU ALA	LYS ARG	LEU VAL	VAL VAL SER	ARG GLU	GLY VAL VAT	SER VAL	GLY GLY	LYS GLU	GLU								
• Molecul	le 14:	Pre-r	nRN	A-p	roce	essii	ng i	fact	or	19												
Chain s:		29%									71	%										
M1 K30 L96 A140	PRO ALA ALA GLN	ASN GLY GLU ALA	MET ALA VAL	ASP SER GLU	SER LEU SER	GLU GLU	LEU VAL	HIS VAL	ASN GLU	VAL GLN CT N	GLN	MET LYS	ARG LYS	LYS ARG	PRO ILE	PR0 GLN	GLY TRP	ALA THR	ALA ASP	ASP VAL	ALA	LEU
GLN GLN VAL ALA TYR THR ASP	LEU ASN VAL THR	GLN ALA SER SER	LEU ASP LEU	GLU ASN GLU	CYS ALA ALA	VAL GLY	GLY LEU	GLY LYS	LEU ASP	ILE TYR SFD	VAL VAL	ALA ASN	VAL GLU	ARG THR	LEU ASP	GLY	GLU PRO	VAL THR	ALA THR	GLU TRP	THR	THR
LYS VAL VAL ILE GLY ALA	LYS GLY TRP VAL	LYS VAL TYR ASP	ALA GLY ARG	GLU SER ALA	THR PHE GLN	THR	ALA GLY	VAL VAL THR	GLY LEU	ALA VAL HTS	PRO GLY	GLY ARG	LEU ALA	SER VAL	GLY VAL	ASP LYS	SER PHE	VAL PHE	TYR ASP	GLU	THR	GLU
ARG VAL ALA ARG GLY TYR ALA	ASP ALA ALA LEU	THR THR CYS ALA	PHE HIS PRO	ASP GLY ASN	LEU PHE ALA	ALA GLY	GLN THR	GLY HIS	ILE LEU	VAL PHE HTS	THR	LEU	GLN ALA	GLU SER	PHE PRO	GLY	THR PRO	GLN	ALA LEU	ALA PHE	SER	ASN
GLY PHE TRP PHE ALA ALA THR	GLY CLY GLY THR	SER SER VAL THR	ILE PHE ASP	LEU ARG LYS	SER GLY ALA	ALA	ALA VAL	GLU LEU	GLN THR	GLV GLU VAT	LEU SER	ILE SER	ASP TYR	THR GLY	GLN TYR	LEU	THR GLY	GLY GLY	THR GLY	VAL THR	VAL	MET
TYR THR LYS ALA THR LYS SER	TRP SER GLU PRO	VAL ARG LEU GLY	MET PRO VAL	VAL GLY VAL	LYS TRP GLY	GL Y	ALA LYS	LEU VAL	VAL VAL	SER ARG CI II	GLY VAL	VAL SER	UAL LEU GLY	LYS LYS	CLU GLU							
• Molecul	le 15:	Puta	tive	bud	site	e sel	lect	ion	pr	otei	'n											
Chain N:								979	%										••			
M1 P6 A7 C141	C144 D148																					
• Molecul	le 16:	Pepti	idyl-j	prol	yl ci	is-tı	ran	s is	om	eras	se											
Chain S:							ç	93%										• 6	5%			
M1 F51 G157 GLU GLU	GLU GLU ASP ASP	LYS GLN LEU VAL																				

• Molecule 17: Pre-mRNA-splicing factor PRP46

Chain T: 64% • 34%



MAG
MAG
LEU
CU
MAG
<

SER VAL

• Molecule 19: Putative pre-mRNA splicing protein



• Molecule 20: Putative pre-mRNA splicing protein



Chain 0:	65%	• 32%
MET ALA GLU GLU GLU GLU FRC PRO ASP ASP ASP	VAL VAL ALA ALA ALA ALA ALA ASP GLU GLU ASN ASN ASN ASN ASN ASN ASN ASN ASN ASN	VAL VAL ALA ALA ALA CLUS CLUS CLUS LUS LUS LUS ARG ARG ARG ARG ARG CLUS ARG ARG ARG CLUS ARG ARG ARG ARG ARG ARG ARG ARG ARG ARG
R208 V209 N212 R213 R213 R294 R295 G326 ARG	GLU GLY GLY GLY GLU GLU GLU GLU MET LEU GLU GLU GLU GLU GLU GLU GLU GLU GLU GL	ALA GLU TLE ALA ALA ALA PRO GLU CLU CLS CLN GLU CLS CLN GLN GLN GLY CLS CLA CLA CLA ALA ALA ALA ALA ALA ALA ALA
GLN GLY GLY ALA LYS LYS ALA ALA ALA THR THR THR CYS ASP	SER LYS LYS GUU GUU PRO GUY SER ASP GUU GUU GUU GUU	
• Molecule 21:	Pre-mRNA-processing protein	45
Chain R:	58%	• 41%
M1 T42 GLN ILE ILE ILE ILE ARG ARG ARG ARG ARG ARG	PRO PRO PS2 PS2 PS2 PS2 N177 N177 N177 S234 S234 A14 S242 A14 A14 A14 A14 A14 A14 A14 S242 A14 A14 A14 A14 A14 A14 A14 A14 A14 A14	ARG SER SER TYR SER TYR SER SER SER SER SER SER ARG ARG ARG ARG ARG ARG CLU SER ARG CLU SER ARG CLU CLU CLU CLU CLU CLU CLU CLU CLU CLU
LYS LEU LEU CLU GLU GLU GLU LYS LEU ARG CLN	SER ARG ARG GLY ALA ALA ALA ARG ALU ARG ALA ARG ARG ARG ARG ARG ARG ARG ARG ARG AR	LVS LVS ALA ALA ALA ALA ALA ALA ALA ALA ALA ASP ASP ASP ASP ASP ASP ASP ASP ASP AS
PHE GLY GLY THR THR THR ALA GLU GLU ASP ASN PRO	ASP LYS PRO PHE LEU PHE ALA ALA ALA ALA ALA ALA ALA TYR TYR TYR TYR ASN ARG PRO ALU ARG ARG ALU ARG ARG ALU ALA ALA ALA ALA ALA ALA ALA ALA ALA	TYR GLU ASP ASP GLU GLU GLU GLU GLU GLU ARG GLU ARG GLY GLU ARG GLY GLU GLY GLU GLY GLU GLY GLU ARG GLY GLY GLU ARG ARD ARG ARD ARP ARP ARP ARP ARP ARP ARP ARP ARP ARP
GLY THR PHE LYS GLY GLY GLV ALA ALA GLN PRO	ARG GLV GLV VAL VAL VAL CVAL CVAL ALA ALA ALA ALA ASP PRO ASP ASP CVAL VAL VAL CVAL	SER CULV CULN CULN CULN CULN CULY SER ASN CULY CULN CULN CULN CULN CULN CULN CULN CULN
LYS ARG SER ARG VAL ASP GLU ASP ASP	ASP	
• Molecule 22:	PRP17	
Chain W:	72%	27%
MET PRO PRO ASP ASP GLV GLV TYR PRO PRO ASN	PRO VAL ASP ASP ALA ALA LEU LEU LEU LEU CLU CLU CLU ALA ALA ALA ALA ALA ALA ALA ALA ALA A	P34 P52 ARC ARC ARC SER SER SER SER ARC ARC M108 M108 M108 ARG ARG
GLU GLU GLU GLU VAL VAL VAL VAL VAL LYS CLY GLU	611 612 617 617 617 617 611 718 718 718 718 718 718 718 718 718 7	THR VAL LYS LYS LYS ALA ALA ALA ALA ALA ALA ALA ALA ALA AL
LEU GLY GLU GLU SER GLU TYR LEU GLY ARG GLY	THR MET MET MET MIS MIS PRO CLEU MIS MIS MIS MIS MIS MIS MIS MIS MIS MIS	ITR REC REC REC REC REC REC REC REC REC RE
• Molecule 23:	Putative pre-mRNA splicing p	protein
Chain P:	43%	56%
MET T2 P71 P70 P10 CLY CLY SER SER SER	SER ASN ASN ASN ASN ASN ASN PRO PRO ALA ALA ALA SER SER SER SER SER SER SER CUY CUY SER SER SER SER SER SER SER SER SER SER	PRU SER SER ASP ASP GLY GLY GLY GLY ASP PRO ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP
ALA GLU THR ARG ASP ASP ASP ASP ASP ASP	SER SER SER SER SER SER ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP	ASP ASP ALA ALA ALA CLU CLU CLU CLU CLU ARG ALA ALA ALA ALA



GLU ARG LYS ALA ALA GLU GLU ALA ARG ARG	GLU GLU GLU GLU GLU GLU ASN ASN ASN ASN ASP PHE ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP	VAL R217 R217 R238 R260		
• Molecule 24	: Pre-mRNA-splicing fa	ctor		
Chain Y:		95%		5%
MET PRO PRO ALA LVS ARG LVS ILYS SER SER	ALA ASP SER SER ASO ALA ALA ALA ALA ALA ALA ALA ALA ALA AL	K1396 K1396 K1367 GLY GLU GLU VAL VAL VAL VAL VAL	MET PRO LYS GLU VAL VAL SER ALA ALA ALA ALA ALA	GLU GLY TYR LEU PRO LEU ASP
GLU GLU GLU GLU GLU GLU GLU GLU GLU GLU	GLY ARG LYS CLY CLY CLY CLY CLY CLU CLU CLU CLU SER SER			
• Molecule 25	: Small nuclear ribonuc	leoprotein E		
Chain a:	{	88%	12%	_
MET THR SER SER ARG GLY PRO GLY GLY ARG	VARG L13 A98 A98			
• Molecule 25	: Small nuclear ribonuc	leoprotein E		
Chain j:		90%	10%)
MET THR SER ARG GLY ALA ALA ALA ALA ARG	A 98			
• Molecule 26	: Sm protein F			
Chain b:	76%		24%	_
MET ASN THR GLY PHE VAL PRO VAL VAL	A 4 A 4 GLY CLYS CLYS CLN MET AER AER AER ASP CLU GLU GLU GLU			
• Molecule 26	: Sm protein F			
Chain l:	85'	%	• 14%	-
MET ASN THR GLY PHE VAL VAL VAL R2 <u>5</u>	ASP ASP LYS MET GLU GLU			
• Molecule 27	: Delta(14)-sterol reduc	tase		
Chain c: 14	%	86%		-
MET ALA PRO LYS LYS LYS GLN ALA VAL	SER STRG THR GLN HIS TTR GLU PHE PHE PHE PHE PHE ALA ALA ALA ALA	PHEN PHEN GLY LEU LEU LEU VAL LEU VAL ALA PHE ALA ALA ALA CYS	ASN ASP ILLE THR GLY CYS PRO PRO PRO SER SER LEU	LEU HIS PRO LYS SER LYS LYS LYS
LEU ASP VAL LEU LEU LYS GLN GLN GLN TRP	LYNU LYNU GLY CLY CLY CLE CLY CLE CLY CLE CLY CLE CLY CLE CLY CLE CLY CLE CLY CLE CLY CLE CLY CLE CLY CLE CLY CLE CLY CLE CLY CLY CLY CLY CLY CLY CLY CLY CLY CLY	ALEO TYR TYR TEU TRP LEU TRP LEU TRP LEU TRP LEU TRP LEU TRP FRO	GLY ASP VAL VAL VAL CLU GLU GLU GLU CLU CLU CLU ASN	GLY GLY ARG LEU LYS TYR LYS





• Molecule 28: Small nuclear ribonucleoprotein Sm D1

Chain d:

69%

31%



	-	р.	р.	A		20		20	N	20	A	5	20	5	A	P.	5	5	3	7	5	7	5	7	5	7	7	5	7	7	5	7	>	0	A	5	7	3
- 1	ΩΩ.	n i	03	ف	~	5-	8	5-	03	5	<u> </u>	è.	5-	<u> </u>	<u> </u>	03	æ	<u> </u>	e ce	<u> </u>	è.	<u> </u>	è.	<u> </u>	è.	<u> </u>	<u> </u>	è.	<u> </u>	<u> </u>	è.	10	1	8	<u> </u>	æ	ы,	2
	Ĥ.	₹.	A.	A	E	<u></u>	E	1	A.	<u> </u>	A	A	<u></u>	25	A	A	A	25	A	25	A	25	A.	25	A.	25	25	A.	25	25	A	23	25	A	A	A	25 .	A.

 \bullet Molecule 28: Small nuclear ribonucleoprotein Sm D1

Chain o:	72%		• 26%	
MI K9 R44 LYS ASIN LYS ALA ALA CLYS GUU	ALA ASP ARG GLY GLY ARG GLY ARG CLY ARC CLY ARC CLY ARC	GLY GLY ARG GLY GLY ARG ARG GLY ARG GLY		
• Molecule 29: Sm	protein B			
Chain e:	37%	63%	6	
MET SER SER SER SER SER GLN GLN GJ GJ CJN G G G S THR THR	ARG MET GL/M GL/M L/YS PRO PRO ASN SER THR THR THR THR THR THR THR	ILE GLN GLN GLU GLV GLY ALA ALA ALA ALA ALA ALA ALA SER	LIHK LIHK ALA ALA GLY CLY CLY CLY CLY ALA ALA ALA ALA ALA ALA ALA	ALA ALA ALA PRO TLE
SER LEU ALA ALA ALA PRO GLY VAL VAL VAL	PRO PRO PRO PRO PRO GLY PRO ALA ALA ALA ALA CVO CVO CVO CVO CVO CVO CVO CVO CVO CVO	PHE PRO GLY GLY GLY PRO PRO FRO FRO	PKU ALA ALA PRO PRO PRO ALA GLY PRO GLY PRO GLY	GLY PHE PRO PRO GLY
PHE PRO PRO GLY GLY GLY PRO PRO PRO PRO PRO PRO PRO PRO PRO	PRO PRO ARG ARG			
• Molecule 29: Sm	protein B			
Chain p:	48%		51%	
MET SER SER SER SER SER SER ALN ALN MET ALA MET ALA V14	N85 P96 ALA ALA ALA ALA ALA ALA CLY CLY CLY THR	THR GLY GLY ALA ALA ALA TLE SER SER SER SER	ALA ALA PRO ALA ALA PRO CLY CLY CLY CLY CLY CLY CLY CLY CLY CLY	PHE PRO GLY PHE PRO
ALA ALA ALA ALA PRO PRO PRO GLY GLY FRO GLY ALA	PRO PRO GLY PHE PHC PHC PRO PRO PRO PRO PRO CLY GLY	PHE PRO GLY ALA PRO PRO PRO PLY PHE PRO PRO	PAKU GLY GLY ALA PRO PRO PRO PHE PRO PRO PRO	ARG
• Molecule 30: Sm	all nuclear ribonucl	eoprotein Sm D3		
Chain f:	74%		26%	
MET THR SER SER SER SER IS ASN VAL ARG GLY VAL VAL	GLY LEU ALA ARG GLY GLY ARG ARG ARG ARG ARG ARG	SER GLY ARG GLY ARG		
• Molecule 30: Sma	all nuclear ribonucl	eoprotein Sm D3		
Chain u:	79%		21%	•
M1 V90 ARG GLY ARG ARG ALA LEU ALA ARG ARG ARG	ALA THR VAL SER ARG ALA ALA ALA SER GLY GLY ARG GLY ARG			
• Molecule 31: Sma	all nuclear ribonucl	eoprotein G		
Chain g:	88	3%	12%	



MET ALA ABSP ABSP ABSP ABSC ABSP ABSC ABSC ABSC ABSC ABSC ABSC ABSC ABSC
• Molecule 31: Small nuclear ribonucleoprotein G
Chain k: 88% • 11%
MI 237 237 ASP ASP ASP ASP ASP ASP ASP ASP ARG CLY
• Molecule 32: U2 small nuclear ribonucleoprotein A'
Chain h: 50% 50%
M M V 122 V 122 V 122 V 122 V 122 V 122 V 122 V 122 V 124 V
LYNS LYNS GLUU GLUU GLUU GLUU GLUU GLUU GLUU GLU
ASP ASP
\bullet Molecule 33: U2 small nuclear ribonucleoprotein B'-like protein
Chain 1: 51% 49%
Chain 1: 51% 49%
Chain 1: 51% 49%
Chain 1: 51% 49% U
Chain 1: 51% 49% ####################################
Chain 1: 51% 49% b b b b b b b b b b
Chain 1: 51% 49% Image: Sign of the set
Chain I: 51%
Chain I: 51% 49% W 51% 51% 49% W 51%
Chain 1: 51% 49% # 66 # 4 5 7 # 4 # 4 # 5 7 # 5 # 4 # 5 # 5 # 4 # 4 # 5 # 5 # 4 # 4



ASP GLY

• Molecule 35: G-patch domain-containing protein

Chain	Cb:					52	2%							-	-	-	-	48	%	-	-	-		-		
MET ALA ARG LEU ASP	ASP LEU GLN	PHE ASP PRO	ALA ARG	LFU ASN	SER	VAL ASP AI A	TYR SER	SER SER	ASP SER	ASP GLY	ASP SER	ASP SER	ASP ASP	TYR VAL	ILE PRO	SER ARG	ASN PRO	HIS ASP	ASP GLU	PHE AT A	ASP	ASN	PRO	LYS ARG	ARG ARG	THR GLY
ARG ASP PRO LYS GLU	SER ALA ALA	LEU GLY ILE	PHE ALA	ASP SER	ASP GLU	ASP GLU GL V	PRO ALA	ARG ARG	TRP LYS	LYS LYS	PRO LEU	ARG ASN	ARG GLY	VAL ASN	PHE VAL	SER GLY	GLN GLN	GLN	GLU	GLU	SER	GLU	SER ASP	GLU GLU	GLU SER	ASN ASP
ASP GLU ARG LYS GLN	GLN GLU ALA	VAL ASP ASN	GLU TYR	ASP ASP ASP	ASP PRO	GLY GLY VAI	PRO ALA	ALA LEU	THR GLU	ASP ASP	VAL ASP	GLU ASP	GLU ASP	ASP GLU	MET GLY	GLY VAL	GLY LEU	GLY PHE	LYS PRO	SER	THR	ALA	THR SER	GLN PRO	GLY GLY	GLY LEU
GLY TRP THR LEU PRO	THR GLN GLN	GLU THR THR	THR LYS LYS	PRO PRO	PRO PHE	GLN ARG PHF	VAL LYS	SER LYS	ILE ASP	PRO SER	ASN PRO	GLY	PRO GLY	PHE THR	PRO THR	SER ALA	ARG ALA	PR0 THR	LEU	ALA	ASN	GLU	PRO THR	PRO PRO	R237	<mark>A299</mark> VAL
LYS GLU THR LYS	GLN GLU ARG	GLU GLU	LYS ARG	ALA ALA ARG	LEU ARG	GLU GLU	VAL ILE	ASP SER	SER GLU	GLU GLU	ARG LYS	LYS LYS	ALA ALA	LYS LYS	LYS LYS	GLY LEU	ALA GLY	LYS ILE	GLY SER	GLY	GLY	GLY	ALA SER	THR PRO	ARG PRO	GLN
PRO LYS TYR LEU THR	MET GLU GLU	ILE LYS LYS	VAL ALA	GLY	HIS ILE	PRO GLU AT A	PHE THR	PRO ILE	LEU ASP	MET THR	GLY PRO	GLU D	LEU	LEU THR	SER THR	SER GLY	LEU MET	THR PRO	THR GLY	GLY		DEBU DEBU	LYS PRO	PROGLU	ARG GLY	PRO ALA
LEU ARG THR GLY ARG	GLU GLU ARG	ARG ARG THR	CLU GLU	ARG GLU	ARG GLU	ARG GLU ARC	GLU GLU	GLU LYS	GLU ARG	ALA THR	VAL GLY	LYS PRO	PR.0 GLU	GLU ILE	THR PHE	ARG GLN	VAL LEU	GLU GLU	TRP CYS	GLN GLN	SIH	LEU	GLN	MET PRO	GLU ARG	ASN LYS
ILE HIS ALA GLU GLY	PRO LEU TYR	ARG ILE SER	GLY PRO	GLY LYS	GLY GLY	VAL SER VAT	TYR PHE	GLY GLY	ASN SER	VAL PHE	VAL LEU	MET LYS	ASP ARG	LYS PRO	GLU GLU	PHE ARG	ARG ASP	ASP GLU	ALA GLU	TRP	ILE	LEU	ASP MET	VAL TYR		
• Mole	ecule	36:	RN	IA	heli	cas	е																			
Chain	Cc:										939	%											7	7%	I	
MET ALA ASP ILE THR	LYS GLY THR	LYS ARG LEU	SER GLY	ALA GLU	GL Y SER	GLN ASP SFR	LYS ARG	VAL LYS	THR ASN	GL Y ASP	LYS MET	ASP ALA	LYS ASN	PRO TYR	LEU ALA	HIS	GLU GLU	LYS PRO	GLU ALA	ASP DHF	GLU GLU	Top	A7 64			
• Mole	ecule	37:	Un	kno	own	ml	RN.	А																		
Chain	7:		2	1%		•		-	-	-	-	-	-	7	4%	-	-	-	-	-	-	-				
G1 U2 A3 Y5P4 A5	Y5P6 A7 Y5P8	N Y5P10 Y5P11	Y5P12 Y5P13 VED14	N N	NN	NNN	N N	N	N	NN	N	NN	N	NN	N	NN	N	N	N	N	N N	NN	N N	N N	N N	NN
N N N N	N N	N N	NN	N N N	NN	NNN	N N	N	N	N <mark>Y5P89</mark>	Y5P90 A91	Y5P92 A93	Y5P94 Y5P95	Y5P96 A97	Y5P98 Y5P99	Y5P100 Y5P101										
• Mole	ecule	38:	Pu	tati	ive	cyc	lop	hili	in j	pro	tei	n														
Chain	Ci:					5	5%								_			45	5%			_	_			

W I D E DB

LEU GLU ASP LEU ASP VAL ALA ALA ALA ALA CLY PRO CLY CTU

• Molecule 39: Unknown protein

Chain H:

100%

There are no outlier residues recorded for this chain.



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	77668	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE; Relion	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{\AA}^2)$	50	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV $(4k \ge 4k)$	Depositor



5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: Y5P, SEP, P5P, ZN, M7M, IHP, MG, GTP

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).

Mal	Chain	Bo	ond lengths	Bond angles		
	Unam	RMSZ	# Z > 5	RMSZ	# Z > 5	
1	2	0.43	0/2471	0.82	2/3832~(0.1%)	
2	5	0.73	1/2612~(0.0%)	0.90	3/4059~(0.1%)	
3	6	0.86	1/2338~(0.0%)	0.96	3/3639~(0.1%)	
4	А	0.46	1/16969~(0.0%)	0.62	11/23006~(0.0%)	
5	В	0.41	0/2099	0.63	1/2806~(0.0%)	
6	С	0.40	0/7464	0.59	1/10117~(0.0%)	
7	D	0.33	0/793	0.66	0/1051	
8	Ε	0.35	0/2428	0.66	0/3295	
9	F	0.31	0/891	0.66	0/1201	
10	Ι	0.32	0/6282	0.59	1/8491~(0.0%)	
11	J	0.42	0/5301	0.60	1/7149~(0.0%)	
12	L	0.35	0/5120	0.63	2/6882~(0.0%)	
13	Κ	0.29	0/1833	0.57	0/2493	
14	q	0.27	0/1128	0.61	0/1533	
14	r	0.27	0/1151	0.58	0/1566	
14	\mathbf{S}	0.26	0/1133	0.58	1/1540~(0.1%)	
14	\mathbf{t}	0.25	0/1141	0.56	0/1552	
15	Ν	0.52	0/1227	0.66	1/1655~(0.1%)	
16	S	0.36	0/1235	0.66	0/1671	
17	Т	0.56	0/2635	0.76	3/3582~(0.1%)	
18	U	0.34	0/2883	0.62	1/3895~(0.0%)	
19	М	0.39	0/2006	0.67	1/2703~(0.0%)	
20	0	0.39	0/2278	0.61	0/3081	
21	R	0.44	0/2724	0.63	0/3675	
22	W	0.32	0/2237	0.57	1/3068~(0.0%)	
23	Р	0.46	0/945	0.66	0/1264	
24	Y	0.25	0/11057	0.49	0/14995	
25	a	0.24	0/425	0.43	0/589	
25	j	0.26	0/716	0.60	0/969	
26	b	0.24	0/350	0.47	0/486	
26	1	0.26	0/661	0.63	0/898	
27	с	0.24	0/415	0.45	0/575	



Mal	Chain	Bo	ond lengths	E	Bond angles
INIOI	Unam	RMSZ	# Z > 5	RMSZ	# Z > 5
27	m	0.27	0/687	0.61	0/922
28	d	0.24	0/400	0.44	0/556
28	0	0.29	0/691	0.61	0/940
29	е	0.25	0/386	0.47	0/533
29	р	0.29	0/799	0.72	0/1079
30	f	0.24	0/413	0.48	0/573
30	u	0.30	0/726	0.72	0/979
31	g	0.24	0/353	0.47	0/489
31	k	0.28	0/584	0.76	1/787~(0.1%)
32	h	0.24	0/604	0.44	0/841
33	i	0.24	0/503	0.40	0/699
34	CY	0.24	0/443	0.34	0/612
35	Cb	0.24	0/2494	0.35	0/3471
36	Cc	0.24	0/3540	0.39	0/4935
37	7	0.73	0/71	1.48	1/106~(0.9%)
38	Ci	0.25	0/414	0.42	0/575
All	All	0.40	3/106056~(0.0%)	0.62	35/145415~(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
4	А	0	1
8	Е	0	1
10	Ι	0	1
12	L	0	2
16	S	0	1
All	All	0	6

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	6	1	G	OP3-P	-10.80	1.48	1.61
2	5	1	U	OP3-P	-10.61	1.48	1.61
4	А	1111	VAL	CB-CG1	-6.05	1.40	1.52

All (35) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms		$Observed(^{o})$	$Ideal(^{o})$
4	А	725	LEU	CB-CG-CD1	-8.57	96.42	111.00



Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
22	W	108	MET	CA-CB-CG	8.11	127.08	113.30
10	Ι	424	ILE	CG1-CB-CG2	-7.69	94.49	111.40
3	6	74	G	C4-C5-N7	7.67	113.87	110.80
15	Ν	144	CYS	CB-CA-C	-7.66	95.08	110.40
31	k	37	LEU	CA-CB-CG	7.52	132.59	115.30
17	Т	319	LEU	CA-CB-CG	7.26	132.00	115.30
4	А	725	LEU	CA-CB-CG	7.21	131.89	115.30
2	5	27	U	N3-C2-O2	-7.16	117.19	122.20
4	А	458	ASP	CB-CG-OD1	6.82	124.44	118.30
11	J	304	ASP	CB-CG-OD1	6.68	124.31	118.30
5	В	265	ASP	CB-CG-OD1	6.67	124.31	118.30
4	А	1728	LEU	CA-CB-CG	6.58	130.43	115.30
2	5	46	G	O4'-C1'-N9	6.36	113.29	108.20
4	А	1898	LEU	CA-CB-CG	6.32	129.82	115.30
3	6	74	G	C6-C5-N7	-6.12	126.73	130.40
18	U	683	LEU	CA-CB-CG	6.06	129.23	115.30
19	М	86	ILE	CG1-CB-CG2	-5.84	98.56	111.40
4	А	1363	ASP	CB-CG-OD2	-5.81	113.07	118.30
12	L	668	LEU	CA-CB-CG	5.75	128.53	115.30
2	5	27	U	N1-C2-O2	5.70	126.79	122.80
1	2	13	С	P-O3'-C3'	5.69	126.53	119.70
4	А	775	LEU	CA-CB-CG	5.62	128.22	115.30
6	С	852	ASP	CB-CG-OD1	5.57	123.31	118.30
4	А	414	ASP	CB-CG-OD1	5.53	123.27	118.30
14	s	96	LEU	CA-CB-CG	5.39	127.69	115.30
4	А	1963	LEU	CA-CB-CG	5.34	127.58	115.30
17	Т	286	LEU	CA-CB-CG	5.24	127.35	115.30
4	А	1111	VAL	CG1-CB-CG2	-5.19	102.60	110.90
3	6	74	G	N9-C4-C5	-5.14	103.34	105.40
12	L	686	LEU	CA-CB-CG	5.13	127.11	115.30
4	А	1535	LEU	CB-CG-CD1	5.09	119.65	111.00
1	2	16	U	C5-C6-N1	5.07	125.23	122.70
17	Т	286	LEU	CB-CG-CD1	-5.07	102.39	111.00
37	7	97	A	C4-N9-C1'	5.02	135.33	126.30

There are no chirality outliers.

All (6) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	А	1437	ARG	Peptide
8	Е	92	ASP	Peptide
10	Ι	416	GLY	Peptide



Continued from previous page...

Mol	Chain	Res	Type	Group
12	L	505	LEU	Peptide
12	L	539	LEU	Peptide
16	S	51	PHE	Peptide

5.2 Too-close contacts (i)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

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	Perce	ntiles
4	А	2004/2463~(81%)	1911 (95%)	90 (4%)	3~(0%)	48	77
5	В	256/326~(78%)	249~(97%)	6 (2%)	1 (0%)	30	61
6	С	920/1011~(91%)	889~(97%)	29 (3%)	2(0%)	44	73
7	D	91/325~(28%)	90~(99%)	1 (1%)	0	100	100
8	Ε	308/352~(88%)	285~(92%)	23 (8%)	0	100	100
9	F	108/233~(46%)	104 (96%)	3(3%)	1 (1%)	14	42
10	Ι	736/839~(88%)	705~(96%)	29 (4%)	2(0%)	37	67
11	J	604/687~(88%)	592~(98%)	12 (2%)	0	100	100
12	L	629/768~(82%)	602~(96%)	26 (4%)	1 (0%)	44	73
13	Κ	229/231~(99%)	224 (98%)	5 (2%)	0	100	100
14	q	137/480~(28%)	135~(98%)	2 (2%)	0	100	100
14	r	141/480~(29%)	140 (99%)	1 (1%)	0	100	100
14	s	138/480~(29%)	137~(99%)	1 (1%)	0	100	100
14	t	139/480~(29%)	136 (98%)	3 (2%)	0	100	100
15	N	146/148~(99%)	136 (93%)	8 (6%)	2 (1%)	9	30
16	S	155/167~(93%)	148 (96%)	7 (4%)	0	100	100



\mathbf{Mol}	Chain	Analysed	Favoured	Allowed	Outliers	Perce	\mathbf{ntiles}
17	Т	326/496~(66%)	311~(95%)	14~(4%)	1 (0%)	37	67
18	U	346/757~(46%)	321~(93%)	23~(7%)	2(1%)	22	51
19	М	242/395~(61%)	225~(93%)	17~(7%)	0	100	100
20	0	274/408~(67%)	261~(95%)	13~(5%)	0	100	100
21	R	334/578~(58%)	312~(93%)	21~(6%)	1 (0%)	37	67
22	W	391/547~(72%)	376~(96%)	14~(4%)	1 (0%)	37	67
23	Р	110/260~(42%)	102~(93%)	8~(7%)	0	100	100
24	Y	1342/1416~(95%)	1315~(98%)	27~(2%)	0	100	100
25	a	84/98~(86%)	83~(99%)	1 (1%)	0	100	100
25	j	86/98~(88%)	84 (98%)	2(2%)	0	100	100
26	b	69/94~(73%)	69 (100%)	0	0	100	100
26	1	79/94~(84%)	77~(98%)	2(2%)	0	100	100
27	с	82/592~(14%)	80~(98%)	2(2%)	0	100	100
27	m	84/592~(14%)	80~(95%)	4(5%)	0	100	100
28	d	79/118~(67%)	76~(96%)	3~(4%)	0	100	100
28	О	85/118~(72%)	78~(92%)	7 (8%)	0	100	100
29	е	75/211~(36%)	73~(97%)	2(3%)	0	100	100
29	р	99/211~(47%)	92~(93%)	7~(7%)	0	100	100
30	f	82/114~(72%)	79~(96%)	3~(4%)	0	100	100
30	u	88/114~(77%)	83 (94%)	5~(6%)	0	100	100
31	g	70/82~(85%)	69~(99%)	1 (1%)	0	100	100
31	k	71/82~(87%)	68~(96%)	3~(4%)	0	100	100
32	h	120/242~(50%)	117~(98%)	3~(2%)	0	100	100
33	i	100/201~(50%)	100 (100%)	0	0	100	100
34	CY	89/510~(18%)	88~(99%)	1 (1%)	0	100	100
35	Cb	501/975~(51%)	495 (99%)	6 (1%)	0	100	100
36	Cc	$712/\overline{764}~(93\%)$	694 (98%)	18 (2%)	0	100	100
38	Ci	82/153~(54%)	82 (100%)	0	0	100	100
All	All	$12\overline{843/19790}~(65\%)$	12373~(96%)	453 (4%)	17~(0%)	50	77

All (17) Ramachandran outliers are listed below:



Mol	Chain	Res	Type
4	А	494	LYS
4	А	1219	ILE
5	В	111	GLN
6	С	840	VAL
9	F	98	ALA
15	N	7	ALA
18	U	400	VAL
21	R	177	ASN
4	А	595	GLN
12	L	390	GLN
6	С	846	TRP
10	Ι	558	TYR
22	W	101	VAL
18	U	399	ILE
15	Ν	6	PRO
10	Ι	416	GLY
17	Т	489	PRO

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent side chain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

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
4	А	1813/2212~(82%)	1804 (100%)	9~(0%)	86 95
5	В	208/270~(77%)	207~(100%)	1 (0%)	86 95
6	С	809/884~(92%)	807 (100%)	2~(0%)	92 97
7	D	80/276~(29%)	78~(98%)	2(2%)	42 75
8	Ε	254/287~(88%)	250~(98%)	4 (2%)	58 85
9	F	92/179~(51%)	91~(99%)	1 (1%)	70 90
10	Ι	646/729~(89%)	644 (100%)	2~(0%)	91 97
11	J	525/592~(89%)	525~(100%)	0	100 100
12	L	520/635~(82%)	517~(99%)	3~(1%)	84 95
13	Κ	186/186~(100%)	184 (99%)	2(1%)	70 90
14	q	122/385~(32%)	121 (99%)	1 (1%)	79 93



Mol	Chain	Analysed	Rotameric	Outliers	Perce	ntiles
14	r	123/385~(32%)	121 (98%)	2 (2%)	58	85
14	\mathbf{S}	122/385~(32%)	121 (99%)	1 (1%)	79	93
14	t	123/385~(32%)	123 (100%)	0	100	100
15	Ν	131/131~(100%)	129 (98%)	2 (2%)	60	86
16	S	126/135~(93%)	126 (100%)	0	100	100
17	Т	276/408~(68%)	269~(98%)	7 (2%)	42	75
18	U	294/649~(45%)	291 (99%)	3 (1%)	73	91
19	М	210/293~(72%)	207~(99%)	3 (1%)	62	87
20	0	227/335~(68%)	216 (95%)	11 (5%)	21	53
21	R	278/476~(58%)	277 (100%)	1 (0%)	89	96
22	W	73/459~(16%)	73 (100%)	0	100	100
23	Р	91/213~(43%)	90~(99%)	1 (1%)	70	90
24	Y	1173/1231~(95%)	1169 (100%)	4 (0%)	91	97
25	j	79/85~(93%)	79~(100%)	0	100	100
26	1	72/84~(86%)	71 (99%)	1 (1%)	62	87
27	m	77/497~(16%)	76~(99%)	1 (1%)	65	88
28	О	78/95~(82%)	76~(97%)	2(3%)	41	75
29	р	84/152~(55%)	83~(99%)	1 (1%)	67	89
30	u	80/94~(85%)	80 (100%)	0	100	100
31	k	64/71~(90%)	64 (100%)	0	100	100
All	All	$90\overline{36}/13198~(68\%)$	8969 (99%)	67 (1%)	80	94

All (67) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	А	351	ASN
4	А	420	ASN
4	А	589	LYS
4	А	593	SER
4	А	902	ASN
4	А	974	ARG
4	А	1141	ASN
4	A	1268	ARG
4	А	2035	ARG
5	В	298	ASN



Mol	Chain	Res	Type
6	С	527	ASN
6	С	853	ARG
7	D	207	LYS
7	D	218	ARG
8	Е	95	ASN
8	Е	176	ARG
8	Е	229	ASN
8	Е	259	ARG
9	F	120	ARG
10	Ι	472	ARG
10	Ι	727	LYS
12	L	209	ILE
12	L	210	LYS
12	L	541	ARG
13	K	36	LYS
13	K	80	LYS
14	q	25	LYS
14	r	25	LYS
14	r	59	ARG
14	S	30	LYS
15	N	141	CYS
15	Ν	144	CYS
17	Т	178	LYS
17	Т	305	ASN
17	Т	488	LYS
17	Т	491	LEU
17	Т	492	VAL
17	Т	493	ARG
17	Т	495	LYS
18	U	468	ARG
18	U	520	ARG
18	U	736	ARG
19	М	144	ARG
19	М	149	ARG
19	М	290	LYS
20	0	85	ASP
20	0	86	ARG
20	0	88	ASP
20	0	89	LYS
20	0	91	LEU
20	0	208	ARG
20	0	209	VAL



Mol	Chain	Res	Type
20	0	212	ASN
20	0	213	ARG
20	0	294	LYS
20	0	295	ARG
21	R	132	LEU
23	Р	238	ARG
24	Y	376	ARG
24	Y	395	ARG
24	Y	396	LYS
24	Y	1307	ARG
26	1	25	ARG
27	m	86	ARG
28	0	9	LYS
28	0	44	ARG
29	р	85	ASN

Sometimes side chains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such side chains are listed below:

Mol	Chain	Res	Type
4	А	834	ASN
4	А	1141	ASN
4	А	1525	GLN
4	А	1737	GLN
14	q	110	HIS
17	Т	220	GLN
17	Т	305	ASN
20	0	212	ASN
24	Y	683	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	2	101/193~(52%)	18 (17%)	2(1%)
2	5	109/116~(93%)	21 (19%)	1 (0%)
3	6	97/101~(96%)	36~(37%)	2(2%)
37	7	2/101~(1%)	2(100%)	0
All	All	309/511~(60%)	77 (24%)	5 (1%)

All (77) RNA backbone outliers are listed below:



Mol	Chain	\mathbf{Res}	Type
1	2	4	U
1	2	14	U
1	2	16	U
1	2	17	U
1	2	18	G
1	2	19	G
1	2	23	А
1	2	24	G
1	2	28	А
1	2	39	С
1	2	40	U
1	2	41	G
1	2	100	U
1	2	111	G
1	2	120	С
1	2	137	U
1	2	142	G
1	2	145	U
2	5	6	G
2	5	8	А
2	5	20	А
2	5	21	С
2	5	22	С
2	5	23	G
2	5	24	А
2	5	27	U
2	5	33	С
2	5	50	G
2	5	71	А
2	5	72	U
2	5	75	U
2	5	77	С
2	5	87	U
2	5	88	С
2	5	89	U
2	5	91	U
2	5	93	G
2	5	94	А
2	5	106	U
3	6	6	U
3	6	7	С
3	6	9	G
3	6	15	U
L			



Mol	Chain	Res	Type
3	6	18	G
3	6	19	U
3	6	20	С
3	6	21	A
3	6	22	А
3	6	24	U
3	6	28	А
3	6	32	А
3	6	38	А
3	6	39	G
3	6	42	G
3	6	45	U
3	6	47	G
3	6	49	A
3	6	61	С
3	6	62	U
3	6	63	A
3	6	68	U
3	6	70	A
3	6	74	G
3	6	81	А
3	6	82	А
3	6	84	G
3	6	85	А
3	6	88	С
3	6	89	G
3	6	90	С
3	6	91	U
3	6	95	А
3	6	96	А
3	6	97	U
3	6	98	U
37	7	2	U
37	7	97	A

Continued from previous page...

All (5) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	2	13	С
1	2	23	А
2	5	70	А
3	6	15	U



Continued from previous page...

\mathbf{Mol}	Chain	\mathbf{Res}	Type
3	6	38	А

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

25 non-standard protein/DNA/RNA residues 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).

Mal	Turne	Chain	Dec	PosLinkBond lengthsBond angles			Bond lengths		gles	
IVIOI	Type	Ullalli	nes		Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
37	Y5P	7	14	37	14,19,20	2.35	1 (7%)	18,26,29	1.02	1 (5%)
37	Y5P	7	4	37	14,19,20	2.42	1 (7%)	18,26,29	1.11	1 (5%)
37	Y5P	7	13	37	14,19,20	2.41	1 (7%)	18,26,29	1.00	1 (5%)
37	P5P	7	3	37,3	16,23,24	0.80	0	14,33,36	0.75	0
37	Y5P	7	95	37	14,19,20	2.34	1 (7%)	18,26,29	0.99	1 (5%)
37	Y5P	7	8	37	14,19,20	3.67	1 (7%)	18,26,29	0.81	1 (5%)
21	SEP	R	242	21	8,9,10	0.68	0	8,12,14	1.20	1 (12%)
37	Y5P	7	101	37	14,19,20	2.30	1 (7%)	18,26,29	1.03	1 (5%)
37	Y5P	7	99	37	14,19,20	3.84	1 (7%)	18,26,29	0.79	0
37	P5P	7	93	37,1	16,23,24	0.79	0	14,33,36	0.76	0
37	Y5P	7	98	37	14,19,20	2.43	1 (7%)	18,26,29	0.98	1 (5%)
37	Y5P	7	90	37	14,19,20	<mark>3.73</mark>	1 (7%)	18,26,29	0.78	1 (5%)
37	Y5P	7	6	37	14,19,20	2.37	1 (7%)	18,26,29	0.94	1 (5%)
37	P5P	7	7	37,3	16,23,24	1.28	2 (12%)	14,33,36	1.96	2 (14%)
37	Y5P	7	96	37	14,19,20	2.48	1 (7%)	18,26,29	1.02	1(5%)
37	Y5P	7	94	37	14,19,20	3.78	1 (7%)	18,26,29	0.77	1(5%)
37	Y5P	7	92	37	14,19,20	2.35	1 (7%)	18,26,29	0.96	1(5%)
37	Y5P	7	100	37	14,19,20	3.75	1 (7%)	18,26,29	0.79	1(5%)
21	SEP	R	234	21	8,9,10	1.37	1 (12%)	8,12,14	0.95	0
37	P5P	7	91	37,1	16,23,24	0.77	0	14,33,36	0.73	0
37	P5P	7	5	37	16,23,24	1.33	2 (12%)	14,33,36	1.95	2(14%)
37	Y5P	7	12	37	14,19,20	2.43	1 (7%)	18,26,29	1.01	1 (5%)



Mol Type Cl	Turne	Chain	Dec	Link	Bond lengths			Bond angles		
	Unain	nes		Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2	
37	Y5P	7	11	37	14,19,20	2.39	1 (7%)	18,26,29	0.99	1 (5%)
37	Y5P	7	89	37	14,19,20	2.36	1 (7%)	18,26,29	0.96	1 (5%)
37	Y5P	7	10	37	14,19,20	2.41	1 (7%)	18,26,29	0.99	1 (5%)

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
37	Y5P	7	14	37	-	2/7/33/34	0/2/2/2
37	Y5P	7	4	37	-	1/7/33/34	0/2/2/2
37	Y5P	7	13	37	-	1/7/33/34	0/2/2/2
37	P5P	7	3	37,3	-	2/3/25/26	0/3/3/3
37	Y5P	7	95	37	-	4/7/33/34	0/2/2/2
37	Y5P	7	8	37	-	1/7/33/34	0/2/2/2
21	SEP	R	242	21	-	2/5/8/10	-
37	Y5P	7	101	37	-	1/7/33/34	0/2/2/2
37	Y5P	7	99	37	-	1/7/33/34	0/2/2/2
37	P5P	7	93	37,1	-	1/3/25/26	0/3/3/3
37	Y5P	7	98	37	-	1/7/33/34	0/2/2/2
37	Y5P	7	90	37	-	2/7/33/34	0/2/2/2
37	Y5P	7	6	37	-	1/7/33/34	0/2/2/2
37	P5P	7	7	37,3	-	2/3/25/26	0/3/3/3
37	Y5P	7	96	37	-	1/7/33/34	0/2/2/2
37	Y5P	7	94	37	-	1/7/33/34	0/2/2/2
37	Y5P	7	92	37	-	1/7/33/34	0/2/2/2
37	Y5P	7	100	37	-	1/7/33/34	0/2/2/2
21	SEP	R	234	21	-	1/5/8/10	-
37	P5P	7	91	37,1	-	0/3/25/26	0/3/3/3
37	P5P	7	5	37	-	0/3/25/26	0/3/3/3
37	Y5P	7	12	37	-	4/7/33/34	0/2/2/2
37	Y5P	7	11	37	-	4/7/33/34	0/2/2/2
37	Y5P	7	89	37	-	3/7/33/34	0/2/2/2
37	Y5P	7	10	37	-	3/7/33/34	0/2/2/2

All (23) bond length outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
37	7	99	Y5P	C4-N3	-14.20	1.33	1.46
37	7	94	Y5P	C4-N3	-13.97	1.33	1.46
37	7	100	Y5P	C4-N3	-13.85	1.33	1.46
37	7	90	Y5P	C4-N3	-13.72	1.33	1.46
37	7	8	Y5P	C4-N3	-13.53	1.33	1.46
37	7	96	Y5P	C4-N3	-9.10	1.37	1.46
37	7	12	Y5P	C4-N3	-8.99	1.38	1.46
37	7	98	Y5P	C4-N3	-8.93	1.38	1.46
37	7	13	Y5P	C4-N3	-8.91	1.38	1.46
37	7	10	Y5P	C4-N3	-8.89	1.38	1.46
37	7	4	Y5P	C4-N3	-8.88	1.38	1.46
37	7	11	Y5P	C4-N3	-8.86	1.38	1.46
37	7	14	Y5P	C4-N3	-8.72	1.38	1.46
37	7	6	Y5P	C4-N3	-8.68	1.38	1.46
37	7	89	Y5P	C4-N3	-8.65	1.38	1.46
37	7	92	Y5P	C4-N3	-8.56	1.38	1.46
37	7	95	Y5P	C4-N3	-8.55	1.38	1.46
37	7	101	Y5P	C4-N3	-8.48	1.38	1.46
37	7	5	P5P	C6-N1	3.80	1.39	1.32
37	7	7	P5P	C6-N1	3.72	1.39	1.32
21	R	234	SEP	P-01P	3.04	1.60	1.50
37	7	5	P5P	C8-N7	-2.43	1.30	1.34
37	7	7	P5P	C8-N7	-2.32	1.30	1.34

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms		$Observed(^{o})$	$Ideal(^{o})$
37	7	7	P5P	C6-N1-C2	6.43	125.06	115.84
37	7	5	P5P	C6-N1-C2	6.36	124.94	115.84
37	7	101	Y5P	N1-C2-N3	-3.67	114.56	125.33
37	7	10	Y5P	N1-C2-N3	-3.60	114.76	125.33
37	7	11	Y5P	N1-C2-N3	-3.59	114.79	125.33
37	7	12	Y5P	N1-C2-N3	-3.59	114.80	125.33
37	7	4	Y5P	N1-C2-N3	-3.59	114.81	125.33
37	7	95	Y5P	N1-C2-N3	-3.59	114.81	125.33
37	7	13	Y5P	N1-C2-N3	-3.58	114.84	125.33
37	7	14	Y5P	N1-C2-N3	-3.58	114.84	125.33
37	7	98	Y5P	N1-C2-N3	-3.51	115.03	125.33
37	7	89	Y5P	N1-C2-N3	-3.50	115.07	125.33
37	7	92	Y5P	N1-C2-N3	-3.50	115.07	125.33
37	7	6	Y5P	N1-C2-N3	-3.43	115.26	125.33
37	7	96	Y5P	N1-C2-N3	-3.39	115.40	125.33
37	7	7	P5P	N1-C2-N3	-2.95	123.88	127.65



Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
37	7	5	P5P	N1-C2-N3	-2.84	124.02	127.65
21	R	242	SEP	OG-CB-CA	2.74	110.81	108.14
37	7	8	Y5P	N1-C2-N3	-2.20	118.88	125.33
37	7	100	Y5P	N1-C2-N3	-2.05	119.31	125.33
37	7	90	Y5P	N1-C2-N3	-2.03	119.37	125.33
37	7	94	Y5P	N1-C2-N3	-2.02	119.40	125.33

There are no chirality outliers.

Mol	Chain	Res	Type	Atoms
21	R	234	SEP	N-CA-CB-OG
37	7	6	Y5P	O4'-C1'-N1-C2
37	7	11	Y5P	O4'-C1'-N1-C2
37	7	89	Y5P	O4'-C1'-N1-C2
37	7	94	Y5P	O4'-C1'-N1-C2
37	7	96	Y5P	O4'-C1'-N1-C2
37	7	98	Y5P	O4'-C1'-N1-C2
37	7	99	Y5P	O4'-C1'-N1-C2
37	7	100	Y5P	O4'-C1'-N1-C2
37	7	4	Y5P	O4'-C1'-N1-C2
37	7	8	Y5P	O4'-C1'-N1-C2
37	7	10	Y5P	O4'-C1'-N1-C2
37	7	14	Y5P	O4'-C1'-N1-C2
37	7	90	Y5P	O4'-C1'-N1-C2
37	7	92	Y5P	O4'-C1'-N1-C2
37	7	95	Y5P	O4'-C1'-N1-C2
37	7	89	Y5P	C4'-C5'-O5'-P
37	7	3	P5P	C3'-C4'-C5'-O5'
37	7	11	Y5P	O4'-C4'-C5'-O5'
37	7	11	Y5P	C3'-C4'-C5'-O5'
37	7	12	Y5P	C2'-C1'-N1-C2
37	7	12	Y5P	C2'-C1'-N1-C6
37	7	3	P5P	O4'-C4'-C5'-O5'
37	7	95	Y5P	C3'-C4'-C5'-O5'
37	7	12	Y5P	O4'-C1'-N1-C6
37	7	93	P5P	C4'-C5'-O5'-P
37	7	12	Y5P	O4'-C1'-N1-C2
37	7	10	Y5P	C4'-C5'-O5'-P
37	7	7	P5P	C3'-C4'-C5'-O5'
37	7	13	Y5P	O4'-C1'-N1-C2
37	7	11	Y5P	C4'-C5'-O5'-P

All (41) torsion outliers are listed below:



Mol	Chain	Res	Type	Atoms							
37	7	95	Y5P	C4'-C5'-O5'-P							
37	7	95	Y5P	O4'-C4'-C5'-O5'							
21	R	242	SEP	N-CA-CB-OG							
21	R	242	SEP	CB-OG-P-O1P							
37	7	14	Y5P	C4'-C5'-O5'-P							
37	7	10	Y5P	O4'-C4'-C5'-O5'							
37	7	101	Y5P	O4'-C1'-N1-C2							
37	7	89	Y5P	O4'-C4'-C5'-O5'							
37	7	7	P5P	O4'-C4'-C5'-O5'							
37	7	90	Y5P	C4'-C5'-O5'-P							

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 10 ligands modelled in this entry, 7 are monoatomic - leaving 3 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with |Z| > 2 is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol Type	Turne	Chain	Dec		Bond lengths			Bond angles		
	Ullalli	nes	LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z > 2	
43	IHP	J	1001	-	36,36,36	1.47	6 (16%)	54,60,60	0.98	3 (5%)
42	GTP	С	1102	41	26,34,34	1.38	3 (11%)	32,54,54	1.54	6 (18%)
40	M7M	В	401	-	27,32,33	4.11	15 (55%)	33,49,52	1.31	4 (12%)

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
43	IHP	J	1001	-	-	9/30/54/54	0/1/1/1
42	GTP	С	1102	41	-	6/18/38/38	0/3/3/3
40	M7M	В	401	-	-	2/17/47/48	0/3/3/3

All (24) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(A)	$\operatorname{Ideal}(\operatorname{\AA})$
40	В	401	M7M	CBI-NBP	8.02	1.50	1.46
40	В	401	M7M	CBG-NBH	7.81	1.44	1.35
40	В	401	M7M	CBY-CBS	-7.10	1.34	1.53
40	В	401	M7M	OBR-CBS	6.91	1.60	1.45
40	В	401	M7M	CBO-NBP	6.25	1.44	1.35
40	В	401	M7M	CBM-NBE	5.34	1.45	1.32
40	В	401	M7M	OBR-CBQ	-5.25	1.29	1.42
40	В	401	M7M	CBO-NBN	5.21	1.46	1.37
40	В	401	M7M	CBM-NBV	5.20	1.45	1.35
40	В	401	M7M	CBM-NBN	4.48	1.47	1.36
42	С	1102	GTP	C5-C6	-4.43	1.38	1.47
43	J	1001	IHP	P4-014	3.57	1.66	1.59
40	В	401	M7M	OBB-CBF	-3.30	1.18	1.23
40	В	401	M7M	OCB-CBY	3.23	1.50	1.43
40	В	401	M7M	OCA-CBX	-3.16	1.35	1.43
43	J	1001	IHP	P3-O13	3.15	1.65	1.59
43	J	1001	IHP	P5-O15	3.11	1.65	1.59
40	В	401	M7M	CBF-NBE	3.10	1.43	1.38
40	В	401	M7M	CBG-CBO	3.06	1.43	1.37
43	J	1001	IHP	P1-011	2.98	1.64	1.59
43	J	1001	IHP	P2-O12	2.92	1.64	1.59
43	J	1001	IHP	P6-O16	2.91	1.64	1.59
42	С	1102	GTP	C5-C4	-2.08	1.37	1.43
42	С	1102	GTP	O4'-C4'	-2.07	1.40	1.45

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
40	В	401	M7M	CBO-CBG-NBH	4.35	109.30	106.71
43	J	1001	IHP	C5-C4-C3	-3.73	102.25	110.41
42	С	1102	GTP	PB-O3B-PG	-3.69	120.15	132.83
42	С	1102	GTP	C5-C6-N1	3.38	119.91	113.95
42	С	1102	GTP	C8-N7-C5	3.07	108.85	102.99
43	J	1001	IHP	O15-C5-C4	3.06	115.91	108.69
42	С	1102	GTP	C2-N1-C6	-3.04	119.50	125.10



Mol	Chain	Res	Type	Atoms	Ζ	$Observed(^{o})$	$Ideal(^{o})$
42	С	1102	GTP	O2G-PG-O3B	2.38	112.60	104.64
40	В	401	M7M	NBP-CBI-NBH	2.33	106.71	103.38
40	В	401	M7M	CBY-CBX-CBQ	2.24	105.68	101.43
42	С	1102	GTP	C3'-C2'-C1'	2.11	104.16	100.98
40	В	401	M7M	CBG-CBO-NBN	-2.11	120.28	124.00
43	J	1001	IHP	C6-C1-C2	2.07	114.94	110.41

There are no chirality outliers.

Mol	Chain	Res	Type	Atoms
42	С	1102	GTP	C5'-O5'-PA-O1A
42	С	1102	GTP	C5'-O5'-PA-O2A
43	J	1001	IHP	C4-C5-O15-P5
43	J	1001	IHP	C3-O13-P3-O23
43	J	1001	IHP	C5-O15-P5-O35
42	С	1102	GTP	PA-O3A-PB-O1B
40	В	401	M7M	OBR-CBS-CBT-OBU
43	J	1001	IHP	C6-C5-O15-P5
43	J	1001	IHP	C4-O14-P4-O44
42	С	1102	GTP	PB-O3A-PA-O2A
43	J	1001	IHP	C1-O11-P1-O21
42	С	1102	GTP	C5'-O5'-PA-O3A
43	J	1001	IHP	C1-O11-P1-O31
43	J	1001	IHP	C1-O11-P1-O41
43	J	1001	IHP	C3-O13-P3-O43
42	С	1102	GTP	PB-O3A-PA-O1A
40	В	401	M7M	CBT-OBU-PBK-OBL

All (17) torsion outliers are listed below:

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier.



Ligand IHP J 1001 Bond lengths Bond angles Torsions Rings

The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.







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.

