

wwPDB EM Validation Summary Report (i)

Apr 6, 2025 – 01:55 AM JST

PDB ID	:	$9L5S / pdb_00009l5s$
EMDB ID	:	EMD-62842
Title	:	Cryo-EM structure of the thermophile spliceosome (state B^*Q1)
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	:	2.90 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org* A user guide is available at https://www.wwpdb.org/validation/2017/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 2.90 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f EM} {f structures} \ (\#{f Entries})$
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 ch	ain
1	2	193	13% 5% • 81	%
2	5	116	72%	22% • •
3	6	101	59%	31% • 9%
4	А	2463	80%	19%
5	В	326	79%	21%
6	С	1011	91%	9%
7	Е	352	88%	• 12%
8	F	233	47%	53%
9	Ι	839	86%	14%

Continued on next page...



Conti	nued fron	n previous	page	
Mol	Chain	Length	Quality of chain	
10	J	687	88%	11%
11	L	768	79%	20%
12	Κ	231	100%	
13	q	480	29% 71%	
13	r	480	30% 70%	
13	s	480	29% 71%	
13	t	480	29% 71%	
14	Ν	148	98%	.
15	S	167	93%	• 6%
16	Т	496	68%	• 32%
17	М	395	62% ·	37%
18	0	408	67%	• 32%
19	R	578	58%	41%
20	W	547	17% 83%	
21	Р	260	36% 6	4%
22	Y	1416	94%	6%
23	1	698	45% •	53%
24	Z	672	97%	· ·
25	j	98	90%	10%
26	k	82	87%	• 11%
27	1	94	86%	14%
28	m	592	14% 85%	
29	0	118	73%	• 26%
30	р	211	48%	51%
31	u	114	76%	• 21%

Continued on next page...



Continued from previous page...

Mol	Chain	Length	Quality of chain						
32	V	223	36%		63%				
33	Z	678	31%	·	69%				
34	CY	510	5%		95%				
35	Ck	66		64%	36	%			
36	Cb	700	5	6%	44%				
37	8	25	24%		72%	·			
38	CV	246	28%		72%				



2 Entry composition (i)

There are 42 unique types of molecules in this entry. The entry contains 87114 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		A	toms	AltConf	Trace		
1	2	37	Total 778	C 348	N 126	0 267	Р 37	0	0

• Molecule 2 is a RNA chain called U5 snRNA.

Mol	Chain	Residues		At	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		At	toms	AltConf	Trace		
3	6	92	Total 1966	C 879	N 359	O 636	Р 92	0	0

• Molecule 4 is a protein called PRP8.

Mol	Chain	Residues		At	AltConf	Trace			
4	А	1988	Total 16398	C 10544	N 2851	O 2941	S 62	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
5	В	258	Total 1874	C 1152	N 351	O 366	${ m S}{ m 5}$	0	0

• Molecule 6 is a protein called SNU114.

Mol	Chain	Residues		Α	AltConf	Trace			
6	C	022	Total	С	Ν	Ο	\mathbf{S}	0	0
0	U	922	7301	4668	1229	1368	36	0	0



• Molecule 7 is a protein called Anaphase-promoting complex subunit 4-like WD40 domaincontaining protein.

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

• Molecule 8 is a protein called CCDC12.

Mol	Chain	Residues		Ato	\mathbf{ms}	AltConf	Trace	
8	F	110	Total 547	C 327	N 110	O 110	0	0

• Molecule 9 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues		Ator	AltConf	Trace		
9	Ι	724	Total 3596	C 2148	N 724	О 724	0	0

• Molecule 10 is a protein called Suppressor of forked domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	608	Total 4043	C 2510	N 769	0 759	$\frac{S}{5}$	0	0

• Molecule 11 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	L	611	Total 4034	C 2465	N 772	0 789	S 8	0	0

• Molecule 12 is a protein called Pre-mRNA-splicing factor SPF27.

Mol	Chain	Residues		Ato	ms	AltConf	Trace	
12	K	231	Total 1148	C 685	N 231	O 232	0	0

• Molecule 13 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues		Ato	ms	AltConf	Trace	
13	q	139	Total 691	C 413	N 139	O 139	0	0
13	t	141	Total 701	C 419	N 141	0 141	0	0





α \cdot \cdot \cdot	C		
Continued	trom	previous	page
		1	1 0

Mol	Chain	Residues		Ato	ms	AltConf	Trace	
13	r	143	Total 711	$\begin{array}{c} \mathrm{C} \\ 425 \end{array}$	N 143	0 143	0	0
13	s	140	Total 696	C 416	N 140	O 140	0	0

• Molecule 14 is a protein called Putative bud site selection protein.

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

• Molecule 15 is a protein called Peptidyl-prolyl cis-trans isomerase.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	S	157	Total 1209	C 763	N 217	O 223	S 6	0	0

• Molecule 16 is a protein called Pre-mRNA-splicing factor PRP46.

Mol	Chain	Residues		At	AltConf	Trace			
16	Т	339	Total 2653	C 1675	N 479	0 485	S 14	0	0

• Molecule 17 is a protein called Putative pre-mRNA splicing protein.

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

• Molecule 18 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	0	277	Total 2235	C 1387	N 428	0 413	${f S}{7}$	0	0

• Molecule 19 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	R	339	Total 2670	C 1659	N 508	O 496	S 7	0	0



• Molecule 20 is a protein called PRP17.

Mol	Chain	Residues		At	oms			AltConf	Trace
20	W	94	Total 738	C 456	N 141	0 137	${S \atop 4}$	0	0

• Molecule 21 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues		At	oms	AltConf	Trace		
21	Р	94	Total 772	C 483	N 150	0 138	S 1	0	0

• Molecule 22 is a protein called Pre-mRNA-splicing factor.

Mol	Chain	Residues		Ato	AltConf	Trace		
22	Y	1333	Total 6601	C 3935	N 1333	O 1333	0	0

• Molecule 23 is a protein called GPATCH1.

Mol	Chain	Residues		At	oms			AltConf	Trace
23	1	325	Total 2514	C 1582	N 443	0 484	${ m S}{ m 5}$	0	0

• Molecule 24 is a protein called RNA helicase.

Mol	Chain	Residues		At	Atoms					
24	Z	655	Total 5149	C 3270	N 878	0 983	S 18	0	0	

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

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

• Molecule 26 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	k	73	Total 577	C 369	N 101	O 105	${ m S} { m 2}$	0	0

• Molecule 27 is a protein called Sm protein F.



Mol	Chain	Residues		At	oms	AltConf	Trace		
27	1	81	Total 649	C 412	N 114	O 120	${ m S} { m 3}$	0	0

• Molecule 28 is a protein called Delta(14)-sterol reductase.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	m	86	Total 678	C 427	N 129	0 118	$\frac{S}{4}$	0	0

• Molecule 29 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues		At	oms	AltConf	Trace		
29	О	87	Total 679	C 433	N 114	0 128	${S \atop 4}$	0	0

• Molecule 30 is a protein called Sm protein B.

Mol	Chain	Residues		At	oms	AltConf	Trace		
30	р	103	Total 788	C 490	N 148	0 145	${ m S}{ m 5}$	0	0

• Molecule 31 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues		At	AltConf	Trace			
31	u	90	Total 716	C 449	N 132	0 131	${S \atop 4}$	0	0

• Molecule 32 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues		Aton	ıs	AltConf	Trace	
32	V	82	Total 478	C 287	N 99	O 92	0	0

• Molecule 33 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues		Ate	AltConf	Trace			
33	Z	211	Total 1729	C 1104	N 304	0 313	S 8	0	0

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



Mol	Chain	Residues		Ato	\mathbf{ms}	AltConf	Trace		
34	CY	25	Total 190	C 110	N 35	0 44	S 1	0	0

• Molecule 35 is a protein called GCFC2.

Mol	Chain	Residues		Aton	ıs	AltConf	Trace	
35	Ck	42	Total 205	C 121	N 42	O 42	0	0

• Molecule 36 is a protein called TFIP11.

Mol	Chain	Residues		Ator	AltConf	Trace		
36	Cb	391	Total 1938	C 1156	N 391	O 391	0	0

• Molecule 37 is a RNA chain called Unknown mRNA.

Mol	Chain	Residues		At	\mathbf{oms}	AltConf	Trace		
37	8	24	Total 468	C 228	N 72	0 144	Р 24	0	0

• Molecule 38 is a protein called Putative pre-mRNA splicing protein.

Mol	Chain	Residues		Aton	ıs	AltConf	Trace	
38	CV	68	Total 337	C 201	N 68	O 68	0	0

• Molecule 39 is N,N,7-trimethylguanosine 5'-(trihydrogen diphosphate) (CCD ID: M7M) (formula: $C_{13}H_{23}N_5O_{11}P_2$).





Mol	Chain	Residues	Atoms					AltConf
20	В	1	Total	С	Ν	Ο	Р	0
- 39	D	T	30	13	5	10	2	0

 $\bullet\,$ Molecule 40 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

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

 $\bullet \ \ \ Molecule \ 41 \ is \ GUANOSINE-5'-TRIPHOSPHATE \ (CCD \ ID: \ GTP) \ (formula: \ C_{10}H_{16}N_5O_{14}P_3).$





Mol	Chain	Residues	Atoms					AltConf
41	С	1	Total	C	N	0	Р	0
			32	10	\mathbf{b}	14	3	

• Molecule 42 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	AltConf
42	Ν	3	Total Zn 3 3	0
42	М	1	Total Zn 1 1	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.



 \bullet Molecule 1: U2 snRNA



ARA ARA ALA ALA ALA ALA ALA ALA ALA ALA
LEU ARG ARG ARG ARG ARG ARG ARG ARG ARG ARG
ALL ALL ALL ALL ALL ALL ALL ALL ALL ALL
ALA ALA FIRP SFR SFR SFR SFR SFR SFR SFR SFR SFR SFR
C CLU C CLU
ALSU LEEU ASP TILE ASP ARG ALA ASP ALA ALA ALA ALA
• Molecule 5: Pre-mRNA-splicing factor SYF2
Chain B: 79% 21%
PR0 LVS LVS LVS LVS LVS LVS LVS LVS RC RUU PR0 GLU GLU GLU GLU GLU GLU GLU CVAL CVAL CVAL CVAL CVAL CVAL CVAL CVAL
ASP ASP CILI GLIA PRO PRO V326
• Molecule 6: SNU114
Chain C: 91% 9%
Chain C: 91% 9%
Chain C: 91% 9%
Chain C: 91% 9% ####################################
Chain C: 91% 9% Main A: 88% 12%
Chain C: 91% 9% Example a state of the state o
Chain C: 91% 91% 9% Set 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
Chain C: 91% 9% A A A A A A A A A A A A A A A A A A A
Chain C: 91% 9%



ALA GLU PHE GLY THR ALA

• Molecule 9: Putative pre-mRNA splicing protein

Chain I:		86%		14%
MET PRO SER THR LEU ALA ASP ASP GLN PRO ASP	VAL ALA LEU VAL CYS CYS CYS E18 ASA ASN ASN PHE PHE	ALA SER GLN GLN ALA LEU ALA ARG GLN MET	CLN CLN CLN CLN CLN CLN CLN CLN CLN CLN	GLN GLN GLN GLN CLN GLN GLN GLN GLN GLN SER SER
SER GLY GLY GLY GLY PRO PRO CLU VAL ASP ASP ALA ASP	ALA GLN CGLN GLU GLU GLU ALA ALA ALA ALA ALA PRO FLA PLA	ALA ALA SER SER GLU CLYS GLY THR THR TLE	PALA PRO SER GLN VAL GLU VAL SER ASN PRO	ASP ALA ALA ASP ASP ASP ASP ASP ASP ASP ASP ASP AS
• Molecule 10: S	Suppressor of forked	d domain-conta	ining protein	
Chain J:		88%		11%
MET GLU SER SER ARG GLY PRO PRO ARG VAL LYS LYS	033 GLU THR THR THR TLE TLS LYS LYS PRO THR THR Q42 C	AEC MET PRO ARG GLU VAL LYS LYS LYS LYS ARG	AND ASP ASP GLU ASP CLU CLU CLU CLU ASP ASP	TYR VAL PHE PRO ASP GLU GLU GLU GLN CLN CLN SN
LEU SER ASN LEU LEU LEU ALA ALA ALA CYS GLN TRP LYS	GLU LYS LYS GLY GLY GLY GLY GLY GLY GLY	SER		
• Molecule 11: F	Putative pre-mRNA	A splicing prote	vin	
Chain L:	7	9%		20%
MET PRO VAL V4 V4 C250 ASN ASN LYS ARG LYS	SER ASP GLU GLU GLU GLU GLU ARG LYS LYS KSN KSN	ASU ASP LYS ASP ASP PRO PRO SER ALA SER CLEU GLN	ALA ALA LEU LEU ALA GLY GLY MET CIN ARC	CLU ALA CLU CLU CLN SER SER LLYS R308 S543 CLU CLU CLU
ASP ALA ALA ARG ARG ALC UAL CJY ASN STYR STYR STYR STHR	LEU ASN THR THR C361 C361 C361 C17 C17 C17 C17 C17 C17 C17 C17 C17 C1	VAL PRO LYS LYS LYS GLN VAL VAL YAL THR FRO PRO	PRON PRON LEU LEU PRO PRO LEU GLY GLY ASN	GLY GLY ALA ALA ALA PRO LEU ARG GLY GLY GLY THR PRO
LEU ARG PRO ARG ASP ASP FHE SER LEU ASN ALA	THR ASP GLY VAL SER MET GLU ALA ALA ALA CLN CLN GLN CLN	LYS GLN GLN GLN LEU ALA ALA ALA ALA PLO FRO PRO	GLU THR GLU GLU GLU CLU CLU CLU CLU GLU GLU GLU	K496 A768
• Molecule 12: F	Pre-mRNA-splicing	factor SPF27		
Chain K:		100%		
There are no out	tlier residues record	ded for this cha	in.	
• Molecule 13: F	re-mRNA-process	ing factor 19		
Chain q:	29%		71%	
M1 A139 PRO PRO ALA ALA GLV GLV ALA ALA	MET ALA VAL ASP ASP ASP ASP GLU GLU GLU GLU CLU CLU CLU	HIS VAL ASN GLU GLU GLN GLN GLN MET MET	THR ARG LYS LYS ARG ARG PRO GLN GLY TRP	ALA THR ALA ASP ASP ASP VAL ALA ALA ALA GLN VAL VAL
ALA TYR THR ASP LEU ASN VAL THR GLN SER SER	LEU ASP GLU GLU GLU ALA ALA ALA CYS GLY CSP GLY CSP	GLY LYS LYS LYS ASP TYR SER VAL VAL	L YS VAL GLU GLU LEU LEU LEU LEU CLU FLO FLO	VAL THR ALA ALA GLU THR GLU THR CLV CVAL VAL



ALA ALA CUT CUT CUT CUT CUT CUT CUT CUT CUT CUT
ALA ALA ALA ALA ALA ALA ALA ALA ALA ALA
ALA ALA ALA ALA ALA ALA ALA ALA ALA ALA
THAR IT THAR IT THAR SER SER SER SER PRO CLU VAL CLY CLY VAL CLY CLY CLY CLY CLY CLY CLY CLY CLY CL
• Molecule 13: Pre-mRNA-processing factor 19
Chain t: 29% 71%
P141 P141 ALA ALA ALA ALA ALA ALA ALA ALA ALA AL
AST AST AST AST AST AST AST AST AST AST
ALA ALA TRP VAL VAL TRP VAL TRP VAL ALA ALA ALA ALA ALA ALA ALA ALA ALA
ALA ALA ALA ALA ALA ALA ALA ALA ALA ALA
THAR THAR THAR THAR THAR THAR THAR THAR
SER TRP PRO CUU ARC ARC PRO CUV VAL LEU VAL LINS CUV CUV CUV CUV CUV CUV CUV CUV CUV CUV
• Molecule 13: Pre-mRNA-processing factor 19
Chain r: 30% 70%
A143 A143 A143 A143 A143 A144 A14 A14 A14 A14 A14 A14 A14 A14 A1
ASR VAL VAL ASP ALA ALA ALA ASP CUU CUU CUU CUU CUU CUU CUU CUU CUU CU
CLU TRP VAL VAL VAL VAL ARP ALA ALA ALA ALA ALA ALA ALA ALA ALA AL
ALA ALA ALA ALA ALA ALA ALA ALA ALA ALA
CLY CLY CLY CLY SER SER SER SER ALA ALA ALA ALA ALA ALA ALA ALA ALA AL
SER VAL ARG CLEU ARG CLEU VAL VAL CLEU VAL CLEU VAL ARG CLV VAL ARG CLV VAL CVAL CVAL CVAL CVAL CVAL CVAL C

• Molecule 13: Pre-mRNA-processing factor 19



Chain s:	29%			71%		_
M1 PRO PRO ALA ALA ALA ALA GLV GLV ALA	MET ALA VAL ASP SER SER SER LEU SER SER	GLU GLY LEU VAL GLU GLU ASN GLU	GLN GLN GLN LEU MET MET THR	ARG LYS LYS ARG PRO PRO	GLY TRP ALA ALA ASP ASP	VAL ALA ALA ALA LEU GLN GLN VAL ALA ALA
TYR THR ASP ASP LEU ASN VAL THR GLN GLN SER SER	LEU ASP LLEU GLU GLU CYS ALA ALA	VAL GLY GLY GLY LEU ASP GLY GLY LEU LEU	TYR SER VAL VAL ALA ASN TYS	VAL GLU ARG THR LEU ASP ILE	GLU PRO VAL THR ALA GLU	THR GLY THR LYS VAL VAL TLE
GLY THR ALA LYS GLY GLY TRP VAL LYS VAL LYS VAL TYR ASP	ALA GLY GLU GLU SER ALA THR PHE GLN	THR HITS ALA GLY PRO VAL THR GLY LEU	VAL VAL HIS PRO GLY GLY ARG	LEU ALA SER VAL GLY VAL ASP	LTS SER PHE VAL PHE TYR ASP LEU	THR GLY GLU ARG VAL ALA ARG
GLY TYR ALA ASP ALA ALA LEU THR THR THR THR THR THR	PHE HIS PRO ASP GLY ASN LEU PHE ALA	ALA GLY THR GLN GLN HTS HTS HTS LLE	H H H H H H H H H H H H H H H H H H H	GLU ALA GLU SER PHE PRO LEU	GLT THR PRO GLN GLN ALA ALA	FIRE GLU GLY PHE PHE PHE
ALA ALA ALA GLY CLY CLY CLY CLY SER SER SER VAL THR	ILE PHE ASP LEU LEU LEU SER SER ALA	ALA ALA ALA ALA CVAL LYS CLU CLU CLU CLU CLU	GLU VAL LEU SER SER SER SER	ASP TYR THR GLY GLY GLN LEU	ALA THR GLY GLY GLY VAL	VAL VAL GLN MET TTR TTR LYS ALA
THR LYS SER TRP SER GLU PRO VAL LEU GLY	MET PRO VAL VAL GLY TRP GLY	GLY GLU GLU ALA LYS LYS LEU VAL VAL VAL	GLU GLV GLY VAL VAL SER VAL	NTD NTD NTD NTD NTD NTD NTD		
• Molecule 14:	Putative bu	nd site selection	on protein			
Chain N:			98%			·
M1 P6 D40 N41 D148						
• Molecule 15:	Peptidyl-pr	olyl cis-trans	isomerase			
Chain S:		93	%		·	6%
M1 H78 V156 G157 GLU GLU GLU GLU GLU	ASP LYS GLN LEU VAL					
• Molecule 16:	Pre-mRNA-	-splicing facto	or PRP46			
Chain T:		68%		·	32%	_
MET ALA THR ASP VAL ASP CLY CLY PRO ASP	VAL SER LYS LYS LEU GLU ALA LEU LEU ARG	ASP ASN GLU ARG THR THR ARG SER LEU	ALA SER THR PRO ASP SER	GLY ARG LYS ARG TLE LYS LYS LEU	PRO GLY LLEU ALA SER GLU ASP	ASP ILE THR LYS SER ALA LEU
SER LEU ARG LEU HIS ALA GLU GLU GLY ASP VAL	GLN THR LEU PRO GLU ALA ALA ALA LYS	LYS LEU SER ALA ALA ALA GLY PRO PRO LYS	LYS PRO LYS PRO GLY VAL	GLU ASP GLU ALA PRO SER LYS	GLU GLU GLU HIS THR ARG LYS LYS	LLE GLV GLY FLE PRO GLN ALA LYS
PRO SER ALA GLY GLY ALA ALA SER SER ASN ASN LEU	VAL LEU ALA HIS GLY CYS PRO SER ALA	ALA GLY GLY ALA ALA ALA ALA PRO GLN	ASN GLU GLU GLN CLN LEU SFR	L158 R209 D318 D454	K488 Y496	
• Molecule 17:	Putative pr	e-mRNA spli	cing prote	in		
Chain M:		62%		·	37%	_
MET PRO PRO CLN CLYS CLYS GLN GLN ASP ASP ASP ARG	ALA GLY GLU SER THR D18 D18 A131	GLY VAL VAL GLU GLU CLU CLU F138 T139 D140	P155 TYR PHE ARG LYS GLY ARG	GLU GLU ASP SER GLU GLY ASN	VAL VAL ALA GLY SER SER SER	VAL VAL GLY GLY ASN ALA VAL VAL
			W O R L D PROTEIN DA			

GLY ALA GLY GLY GLY GLY GLY GLY GLY ALA ALA ALA ALA ALA ALA ARG ARG	ASP TASP TASP ALA ALA ALA ALA ALA ALA ALA GLY GLY GLY SER SER	ALA ALA ARG GLY PRO CLY R314 R314 R314 CLU GLU ALN ALN	ARG LYS SER ALA ALA CLV GLV GLY GLY GLY GLY
ALA ALA ALA GLY GLY ALA ALA ALA ALA ALA ALA ALA ALA ALA	GLN GLN ASP ASP ASP ASP ALA ALA ALA ALA ALA ALA ALA ALA ALA	ASP VAL GLN TYR ALA SER LEU ALA GLY ASN	
• Molecule 18: Pu	tative pre-mRNA splic	ing protein	
Chain 0:	67%	·	32%
MET ALA ALA GLU GLU CLU CLU CLU CLU CLU ALA ASP ASP ASP ALA ALA	VAL ALA GLU ASP GLU ASP SER ASN CLU CLU CLU CLU CLU CLU CLU VAL VAL VAL	ASN ASN ALU CLU CLU ALA ALA ALA ALA ALA CLU CLVS LVS LVS LVS LVS	ILE ILE ARG LYS LYS LYS LYS N210 N212 N211 N212 R213 R213 R213 R213 R213 R213 R213 R
GLU GLV GLY GLV GLU GLU GLU GLU GLN CLEU GLN LEU LEU	GLU PRO PRO GLU GLU ALA ALA ALA ALA ALA ALA GLU GLU GLU GLU	ALA GLU GLU CLYS LYS LYS LYS CLY GLY TLEU LEU SER SER SER	THR LEU ALA ALA ALA GLV GLV GLV GLV ALA ALA ALA ALA VAL ALA VAL CYS ALA ALA ALA
SER LYS PRO GLU FRO FRO GLY PRO FRO LEU VAL ALA YR	SER ASP ASP GLU GLU GLU GLU GLU		
• Molecule 19: Pr	e-mRNA-processing pro	otein 45	
Chain R:	58%	4	1%
MET A2 GLN ILE VLL ILE VLL ARG ARG ARG CLY PRO	P53 Y54 A349 A14 A14 A14 A14 A14 A14 A14 A14 SER SER SER SER	GLY TYR SER GLU SER GLU SER GLU TYR ASP GLU CLU	SER GLU ARG GLU ARG ARG ARG ARG ARG CLU CJU CJU CJU
GLU ARG LYS LYS LYS LAS ARG GLN GLN ALA ALA ALA	ARG ALA GLA VAL LEU ALA ALA ALG ALA ARG ALU ARG ARG ARG ALU SER ILE ILE	ALA LEU GLY GLY ALA ALA ALA PRO PRO SER GLY GLU SER SER	TYR ARG SER ARG LLEU PHE ARN GLN SER SER CLY THR THR THR
ASN GLU ASP ASP PRO PRO PRO FLEU PHE	GLU GLU ALA SER SER SER SER TLE TTR ARG ARG ARG ARG ARG ARG ARG ARG CUU CUU	ASP ASP GLU GLU GLU CLY GLU GLU MET GLU GLU ARG GLU ARG ARG	CLYS CLYS ASN ASN ASN ALA CLU CLU CLU CLU CLU CLY CLY CLY CLY CLY CLY CLY CLY CLY CLY
THR GLU ASP ALA CLA CLA CLA PRO CLU CLU CLU CLU CLU CLU CLU CLU CLU	GLU LYFS ASP GLU THR GLU THR ALA ALA ALA ALA ALA ALA ALA ALA ALA AL	GLU VAL GLN GLN CLY SER ASN SER ASN SER ASN SER ASN SER ASN CLY GLY GLY	TYR CLIV CLU CLU ASP CLU ASP CLU ASP ASP ASP ASP ANG ANG ANG ANG ANG ANG ANG ANG ANG ANG
ASP ASP GLU ASP ASP ASP ASP			
• Molecule 20: PF	RP17		
Chain W: 17%		83%	
MET PRO ASP PHE GLY TYR PRO PRO PRO PRO PRO VAL	ASP ASP ALA LEU LEU LEU LEU CLU SER GLU SER GLU SER AIA ALA ALA AIA ASN ASN	HIS ALA VAL VAL ARC ARC ARC ARC SER SER SER SER SER SER SER SER	SER ALA SER SER SER SER LEU LEU LEU V101 V101 ARG K69 ARG CLU GLU GLU GLU
TYR GLU VAL VAL VAL VAL VAL VAL VAL CLY GLU GLU GLU GLU	ASP GLU TYR GLU GLU GLU VAL ASP GLU GLU GLU VAL TLL GLU CLU	SER GLY THR VAL LEU LEU LYS ALA PRO PRO PRO ALA ALA ALA ALA	ARG ARG GLU ALA ALA GLU GLU GLU CLU CLU CLU CLU CLU CLU CLU CLU CLY GLY
GLU SER GLU TYR ASP TYR CI ARC GLY THR TYR HIT MET	VAL PAL ASP ASP ASP ASP ASP ASR ASN CLU ASN CLU CVAL CVAL SER THR THR	ASN TYR TYR TYR PRO LYS LYS LLYS LLYS LLYS LLYS TYR TRP TRP ARG HIS HIS	GLY GLY CLYS FRO FRO THE THE ALA CLEU CLEU CLEU CLEU CLEU CLEU CLEU CLEU
LEU GLY GLY ALA ALA ALA ALA ALA ALA ALA ALA ALA A	PHE ARP VAL TYR ARC GIV GIV GIV CIU LEU LEU LEU LEU LEU ARC SER ARC ARC GIV ARC ARC ARC GIV ARC ARC ARC GIV ARC ARC ARC ARC ARC ARC ARC ARC ARC ARC	LYS ALA THR THR ASP ER SER PHE CYS ASN ASN ASN ASN THR THR	PHE LEU SER CLY CLY CLY CLY ARP LYS LYS ARC LEU TRP ARC LEU TRP TRP TRP CLU
THR GLY GLN CYS CYS CYS CYS ASN ASN ASN ASN ILE GLY CYS	PRO HIS VAL LIE LYS PHE ASN PRO SER SER GIU ASN GLU PHE LEU	ALA ALA GLY LEU SER ASP TLS ASP CLN TTR ASP ASP ASP	ALA GLY ASP ASP ASP VAL VAL CLN GLN ASP ARG ARG ARG ARG ARG ARG CLU CLEU CLEU CLEU







• Molecule 24: RNA helicase

Chain z:	97%	• •						
MET ALA ALA ALA ALA CLY CLY CLY CLY ALA ALA ALA ALA ALA ALA ALA ALA ALA A								
• Molecule 25: Small nuclear ribonucl	eoprotein E							
Chain j:	0%	10%						
MET THR SEER ARG ALA PRO ALA ARG ARG ABG								
• Molecule 26: Small nuclear ribonucl	eoprotein G							
Chain k: 8	7% .	11%						
M1 R24 L37 L37 ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP								
\bullet Molecule 27: Sm protein F								
Chain l: 869	%	14%						
MET TARN TARN PHE VAL VAL PRO B88 ASP CTHR CTHR CTT CTT CTT CTT CTT CTT CTT CTT CTT CT								
• Molecule 28: Delta(14)-sterol reduct	tase							
Chain m: 14%	85%							
MET PRO PRO LLYS LLYS LLYS LLYS ALA ALA ALA ALA ALA ALA ALA ALA ALA AL	PHE GLY LEU LEU LEU LEU LEU LEU VAL ALA PHE PHE PHE ALA ALA ALA ALA ALA ALA ALA ALA	PRO SER LEU LEU HIS PRO LYS SER LYS LYS						
LEU ASP VAL LEU LEU CLN GLU CLY TRP PRO CLY TRP ASN ASN ASN ASN ASN ASN ASN ASN ASN ASN	ALA TYR TYR TEEU TEEU TEEU LEEU LEEU LEEU TYR ARG TYR ARG TEEU CUU CLU CLU CLU CLU	LEU ARG ASN GLY GLY GLY ARG LEU LYS TYR LYS						
LEU LEU ALSN ALA SER SER SER SER SER SER CLEU VAL LEU VAL LEU VAL THR CLEU ALA ALA ALA ALA ALA ALA ALA ALA THR THR THR THR THR THR THR THR THR THR	TIR THR PHE ALA ASP ASP ASN VAL TTR CLN TTR CLN TLEU TLEU TLEU TLEU TLEU TLEU ALA ALA ALA ALA	THR PHE VAL TYR ILE ARG SER PHE SER VAL						
LYS PRO GRV GRV GRV GRU GRV GRV GRV CRU CRU CRU CRU CRU CRU CRU CRU CRU CRU	LEU ASN PRO ARN PRO ARG THR THR TLE TLE TLE CLU TLE CLU TLE CLN TLE TLE CLN TLE CLN TLE CLN TLE CLN TLE CLN TLE TLE TLE TLE TLE TLE TL	PRO GLY LEU LEU CLY CLY TRP TRP TTRP LEU MET ASN						
CYS ALA PHE VAL ALA LYS CLY CLY CLY CLY CLY CLY CLY CLY CLY CLY	VAL LEU ALEU ALEU ALEU HIS FIR PHE PHE PHE ALA ALA ALA ALA ALA ALA ALA ALA ALA AL	GLY LEU MET LEU LEU SER PHE GLY ASP ILE VAL						
TRP VAL VAL TILE TILE TILE TILE TIRE CLN CLN CLN CLN VAL CLN TIRE TIRE TIRE TIRE CLN VAL CLN VAL CLN VAL CLN VAL CLN VAL CLN VAL VAL VAL CLN VAL	THR VAL LEUU LEUU VAL LEUU VAL VAL ARA PHE PHE PHE SER ARG ARA ARA ARA	THR ASN PRO ASN ASP PRO SER VAL LYS HIS						



LEU TYR TYR TYR TYR THR CLEU THR CLEU THR THR THR THR THR THR THR THR THR THR
MET ASP ASP ASS ASS ASS ASS ASS ASS ASS ASS
TLE PICO LEU LEU LEU PICO CTU CTU CTU CTU CTU CTU CTU CTU CTU CTU
\bullet Molecule 29: Small nuclear ribonucleoprotein Sm D1
Chain o: 73% • 26%
M1 R61 P87 LYS LYS ASN ASN ASN ASN ASN ASN ASN ASN ASN AS
• Molecule 30: Sm protein B
Chain p: 48% 51%
MET MET SER SER SER SER ARG CLIV ARG CLIV CLIV CLIV CLIV CLIV CLIV CLIV ALA ALA ALA ALA ALA ALA ALA ALA ALA AL
ALA ALA ALA ALA ALA ALA ALA ALA ALA ALA
\bullet Molecule 31: Small nuclear ribonucleoprotein Sm D3
Chain u: 76% · 21%
M1 148 062 077 077 017 017 017 017 017 017 017 017
• Molecule 32: Putative pre-mRNA splicing protein
Chain V: 36% 63%
MET BSR BSR BSR BSR BSR BSR BSR BSR BSR PFR PFR ARG ARG ARG ARG ARG ARG ARA ARA ARA AR
LYS TYR TYR GLU GLU GLU GLU GLU ARG GLU GLU GLU GLU GLU GLU GLU GLU GLU GL
PRIO PRIO PRIO PRIO PRIO PRIO PRIO PRIO
• Molecule 33: Putative pre-mRNA splicing protein
Chain Z: 31% · 69%
MET ALA SER ALA SER ALA ALA ALA ALA ALA ALA ALA CLU CLU CLU CLU CLU CLU CLU CLU CLU CLU





• Molecule 35: GCFC2

Chain Ck:

64%

36%







4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	66478	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: P5P, MG, GTP, ZN, M7M, Y5P

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		B	Bond angles		
	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5		
1	2	0.25	0/866	0.80	1/1345~(0.1%)		
2	5	0.30	1/2612~(0.0%)	0.86	5/4059~(0.1%)		
3	6	0.31	1/2200~(0.0%)	0.80	3/3425~(0.1%)		
4	А	0.27	0/16823	0.54	4/22806~(0.0%)		
5	В	0.27	0/1898	0.54	0/2558		
6	С	0.27	0/7464	0.55	0/10117		
7	Ε	0.26	0/2428	0.59	0/3295		
8	F	0.24	0/546	0.36	0/761		
9	Ι	0.23	0/3595	0.34	0/5017		
10	J	0.26	0/4103	0.47	0/5609		
11	L	0.25	0/4076	0.48	0/5552		
12	Κ	0.24	0/1147	0.32	0/1598		
13	q	0.24	0/690	0.35	0/962		
13	r	0.25	0/710	0.37	0/990		
13	\mathbf{S}	0.24	0/695	0.36	0/969		
13	t	0.24	0/700	0.36	0/976		
14	Ν	0.31	0/1227	0.53	0/1655		
15	S	0.27	0/1235	0.66	0/1671		
16	Т	0.29	0/2723	0.63	1/3701~(0.0%)		
17	М	0.26	0/2006	0.57	0/2703		
18	0	0.30	0/2289	0.59	0/3095		
19	R	0.26	0/2730	0.56	0/3689		
20	W	0.27	0/752	0.60	0/1008		
21	Р	0.25	0/787	0.54	0/1049		
22	Y	0.24	0/6599	0.38	0/9197		
23	1	0.28	0/2573	0.60	1/3473~(0.0%)		
24	Z	0.26	0/5245	0.58	1/7100~(0.0%)		
25	j	0.25	0/716	0.58	0/969		
26	k	0.26	0/584	0.72	1/787~(0.1%)		
27	1	0.25	0/661	0.61	0/898		
28	m	0.25	0/687	0.64	1/922 $(0.1%)$		
29	0	0.25	0/691	0.60	0/940		



Mal	Chain	Bond lengths		Bond angles	
IVIOI		RMSZ	# Z > 5	RMSZ	# Z > 5
30	р	0.27	0/799	0.70	1/1079~(0.1%)
31	u	0.27	0/726	0.68	1/979~(0.1%)
32	V	0.25	0/481	0.49	0/661
33	Ζ	0.26	0/1768	0.58	2/2384~(0.1%)
34	CY	0.34	0/190	0.66	0/253
38	CV	0.23	0/336	0.29	0/467
All	All	0.26	2/86358~(0.0%)	0.56	22/118719~(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
16	Т	0	1
19	R	0	1
24	Z	0	1
All	All	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	$\mathrm{Ideal}(\mathrm{\AA})$
3	6	1	G	OP3-P	-10.63	1.48	1.61
2	5	1	U	OP3-P	-10.54	1.48	1.61

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms		$Observed(^{o})$	$Ideal(^{o})$
2	5	22	С	N1-C2-O2	8.39	123.93	118.90
2	5	22	C	C2-N1-C1'	8.35	127.98	118.80
23	1	43	PRO	N-CA-CB	-7.85	93.88	103.30
31	u	77	ASP	CB-CG-OD2	7.19	124.77	118.30
3	6	23	U	N1-C2-O2	7.04	127.73	122.80

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
19	R	54	TYR	Peptide
16	Т	318	ASP	Peptide
24	Z	449	ARG	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	1984/2463~(81%)	1915 (96%)	67 (3%)	2~(0%)	48	77
5	В	256/326~(78%)	247 (96%)	9 (4%)	0	100	100
6	С	920/1011~(91%)	891 (97%)	27 (3%)	2(0%)	44	73
7	Ε	308/352~(88%)	292 (95%)	16 (5%)	0	100	100
8	F	108/233~(46%)	105 (97%)	3 (3%)	0	100	100
9	Ι	722/839~(86%)	712 (99%)	10 (1%)	0	100	100
10	J	604/687~(88%)	595 (98%)	9 (2%)	0	100	100
11	L	603/768~(78%)	593 (98%)	10 (2%)	0	100	100
12	K	229/231~(99%)	229 (100%)	0	0	100	100
13	q	137/480~(28%)	136 (99%)	1 (1%)	0	100	100
13	r	141/480~(29%)	141 (100%)	0	0	100	100
13	s	138/480~(29%)	137 (99%)	1 (1%)	0	100	100
13	t	139/480~(29%)	139 (100%)	0	0	100	100
14	Ν	146/148~(99%)	141 (97%)	4 (3%)	1 (1%)	19	49
15	S	155/167~(93%)	141 (91%)	13 (8%)	1 (1%)	22	52
16	Т	337/496~(68%)	323 (96%)	14 (4%)	0	100	100
17	М	242/395~(61%)	232 (96%)	9 (4%)	1 (0%)	30	60
18	0	275/408~(67%)	251 (91%)	24 (9%)	0	100	100
19	R	$\overline{335/578~(58\%)}$	318 (95%)	17 (5%)	0	100	100
20	W	90/547~(16%)	86 (96%)	3 (3%)	1 (1%)	12	37
21	Р	90/260~(35%)	89 (99%)	1 (1%)	0	100	100

Continued on next page...



Mol	Chain	Analysed	Favoured	Allowed	Outliers	Perce	entiles
22	Y	1329/1416~(94%)	1320 (99%)	9 (1%)	0	100	100
23	1	317/698~(45%)	300~(95%)	14 (4%)	3 (1%)	14	43
24	z	653/672~(97%)	619~(95%)	33 (5%)	1 (0%)	44	73
25	j	86/98~(88%)	85~(99%)	1 (1%)	0	100	100
26	k	71/82~(87%)	70~(99%)	1 (1%)	0	100	100
27	1	79/94~(84%)	79~(100%)	0	0	100	100
28	m	84/592~(14%)	83~(99%)	1 (1%)	0	100	100
29	О	85/118 (72%)	80 (94%)	5 (6%)	0	100	100
30	р	99/211~(47%)	94~(95%)	5 (5%)	0	100	100
31	u	88/114~(77%)	83~(94%)	3 (3%)	2(2%)	5	20
32	V	78/223~(35%)	72 (92%)	6 (8%)	0	100	100
33	Z	209/678~(31%)	200~(96%)	9 (4%)	0	100	100
34	CY	23/510~(4%)	23~(100%)	0	0	100	100
38	CV	$\overline{66/246}\ (27\%)$	66 (100%)	0	0	100	100
All	All	11226/17581~(64%)	10887 (97%)	325 (3%)	14 (0%)	50	77

Continued from previous page..

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	С	840	VAL
17	М	140	ASP
20	W	101	VAL
23	1	43	PRO
23	1	297	VAL

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	1796/2212~(81%)	1794 (100%)	2~(0%)	92 98
5	В	150/270~(56%)	150 (100%)	0	100 100

Continued on next page...



Mol	Chain	Analysed	Rotameric	Outliers Perce		ntiles
6	\mathbf{C}	809/884~(92%)	809 (100%)	0	100	100
7	Ε	254/287~(88%)	252~(99%)	2(1%)	79	93
10	J	244/592~(41%)	243 (100%)	1 (0%)	89	97
11	L	281/635~(44%)	280 (100%)	1 (0%)	89	97
14	Ν	131/131~(100%)	129~(98%)	2(2%)	60	85
15	S	126/135~(93%)	125~(99%)	1 (1%)	79	93
16	Т	286/408~(70%)	284~(99%)	2(1%)	81	94
17	М	210/293~(72%)	209 (100%)	1 (0%)	86	96
18	0	228/335~(68%)	224 (98%)	4 (2%)	54	82
19	R	279/478~(58%)	279~(100%)	0	100	100
20	W	73/459~(16%)	73 (100%)	0	100	100
21	Р	76/213~(36%)	75~(99%)	1 (1%)	65	88
23	1	266/564~(47%)	260~(98%)	6 (2%)	45	77
24	Z	559/571~(98%)	557~(100%)	2 (0%)	89	97
25	j	79/85~(93%)	79~(100%)	0	100	100
26	k	64/71~(90%)	63~(98%)	1 (2%)	58	84
27	1	72/84~(86%)	72~(100%)	0	100	100
28	m	77/497~(16%)	76~(99%)	1 (1%)	65	88
29	О	78/95~(82%)	77~(99%)	1 (1%)	65	88
30	р	84/152~(55%)	84 (100%)	0	100	100
31	u	80/94 (85%)	80 (100%)	0	100	100
32	V	22/197~(11%)	21 (96%)	1 (4%)	23	56
33	Ζ	182/575~(32%)	180 (99%)	2(1%)	70	90
34	CY	17/418 (4%)	17 (100%)	0	100	100
All	All	$6523/107\overline{35}\ (61\%)$	6492 (100%)	31 (0%)	85	96

5 of 31 residues with a non-rotameric side chain are listed below:

Mol	Chain	Res	Type
18	0	213	ARG
29	0	61	ARG
23	1	42	VAL
33	Ζ	343	ARG
24	Z	670	LYS



Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
16	Т	305	ASN
23	1	276	ASN
28	m	64	ASN
24	Z	427	GLN
4	А	1918	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	2	36/193~(18%)	11 (30%)	1 (2%)
2	5	109/116~(93%)	26~(23%)	1 (0%)
3	6	91/101~(90%)	30~(32%)	1 (1%)
37	8	0/25	-	-
All	All	236/435~(54%)	67~(28%)	3(1%)

5 of 67 RNA backbone outliers are listed below:

Mol	Chain	\mathbf{Res}	Type
1	2	13	С
1	2	14	U
1	2	15	U
1	2	16	U
1	2	17	U

All (3) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	2	13	С
2	5	70	А
3	6	83	А

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

24 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



Mal	Trune	Chain	Dec	Tinle	Bo	ond leng	$_{\rm sths}$	В	ond ang	jles
	Type	Chain	Res	Link	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
37	P5P	8	-8	37	16,23,24	0.76	0	14,33,36	0.70	0
37	Y5P	8	-4	37	14,19,20	<mark>3.69</mark>	1 (7%)	18,26,29	0.81	1 (5%)
37	Y5P	8	12	37	14,19,20	2.39	1 (7%)	18,26,29	1.05	1 (5%)
37	Y5P	8	-7	37	14,19,20	3.62	1 (7%)	18,26,29	0.81	1 (5%)
37	Y5P	8	13	37	14,19,20	2.40	1 (7%)	18,26,29	1.18	2 (11%)
37	Y5P	8	10	37	14,19,20	2.42	1 (7%)	18,26,29	0.99	1 (5%)
37	P5P	8	7	37,3	16,23,24	1.32	2 (12%)	14,33,36	1.97	2 (14%)
37	Y5P	8	6	37	14,19,20	2.39	1 (7%)	18,26,29	1.01	1 (5%)
37	P5P	8	1	37	16,23,24	1.34	2 (12%)	14,33,36	2.12	3 (21%)
37	P5P	8	-10	37	16,23,24	0.76	0	14,33,36	0.90	0
37	P5P	8	-6	37	16,23,24	1.36	2 (12%)	14,33,36	1.95	2 (14%)
37	Y5P	8	8	37	14,19,20	3.64	1 (7%)	18,26,29	0.76	1 (5%)
37	P5P	8	-9	37	16,23,24	0.75	0	14,33,36	0.87	0
37	P5P	8	3	37,3	16,23,24	0.80	0	14,33,36	0.80	0
37	P5P	8	5	37	16,23,24	1.32	2 (12%)	14,33,36	2.04	2 (14%)
37	P5P	8	0	37	16,23,24	1.37	2 (12%)	14,33,36	2.01	2 (14%)
37	Y5P	8	2	37	14,19,20	2.36	1 (7%)	18,26,29	0.97	1 (5%)
37	P5P	8	-1	37,2	16,23,24	1.34	2 (12%)	14,33,36	1.96	2 (14%)
37	Y5P	8	-5	37	14,19,20	3.71	1 (7%)	18,26,29	1.07	2 (11%)
37	Y5P	8	4	37	14,19,20	2.40	1 (7%)	18,26,29	1.08	1 (5%)
37	Y5P	8	11	37	14,19,20	2.39	1 (7%)	18,26,29	1.02	1 (5%)
37	Y5P	8	14	37	14,19,20	2.34	1 (7%)	18,26,29	1.01	1 (5%)
37	P5P	8	-3	37,2	16,23,24	0.78	0	14,33,36	0.77	0
37	P5P	8	-2	37,2	16,23,24	0.80	0	14,33,36	0.85	0

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).

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	P5P	8	-8	37	-	0/3/25/26	0/3/3/3
37	Y5P	8	-4	37	-	3/7/33/34	0/2/2/2
37	Y5P	8	12	37	-	1/7/33/34	0/2/2/2

Continued on next page...



Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
37	Y5P	8	-7	37	-	1/7/33/34	0/2/2/2
37	Y5P	8	13	37	-	4/7/33/34	0/2/2/2
37	Y5P	8	10	37	-	1/7/33/34	0/2/2/2
37	P5P	8	7	37,3	-	1/3/25/26	0/3/3/3
37	Y5P	8	6	37	-	1/7/33/34	0/2/2/2
37	P5P	8	1	37	-	3/3/25/26	0/3/3/3
37	P5P	8	-10	37	-	1/3/25/26	0/3/3/3
37	P5P	8	-6	37	-	2/3/25/26	0/3/3/3
37	Y5P	8	8	37	-	1/7/33/34	0/2/2/2
37	P5P	8	-9	37	-	0/3/25/26	0/3/3/3
37	P5P	8	3	37,3	-	0/3/25/26	0/3/3/3
37	P5P	8	5	37	-	0/3/25/26	0/3/3/3
37	P5P	8	0	37	-	2/3/25/26	0/3/3/3
37	Y5P	8	2	37	-	4/7/33/34	0/2/2/2
37	P5P	8	-1	37,2	-	2/3/25/26	0/3/3/3
37	Y5P	8	-5	37	-	6/7/33/34	0/2/2/2
37	Y5P	8	4	37	-	1/7/33/34	0/2/2/2
37	Y5P	8	11	37	-	3/7/33/34	0/2/2/2
37	Y5P	8	14	37	-	2/7/33/34	0/2/2/2
37	P5P	8	-3	37,2	-	0/3/25/26	0/3/3/3
37	P5P	8	-2	37,2	-	0/3/25/26	0/3/3/3

Continued from previous page...

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
37	8	-5	Y5P	C4-N3	-13.72	1.33	1.46
37	8	-4	Y5P	C4-N3	-13.62	1.33	1.46
37	8	8	Y5P	C4-N3	-13.49	1.33	1.46
37	8	-7	Y5P	C4-N3	-13.42	1.34	1.46
37	8	10	Y5P	C4-N3	-8.92	1.38	1.46

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
37	8	5	P5P	C6-N1-C2	6.57	125.26	115.84
37	8	-6	P5P	C6-N1-C2	6.50	125.16	115.84
37	8	7	P5P	C6-N1-C2	6.47	125.10	115.84
37	8	1	P5P	C6-N1-C2	6.43	125.05	115.84
37	8	-1	P5P	C6-N1-C2	6.36	124.94	115.84



There are no chirality outliers.

Mol	Chain	Res	Type	Atoms
37	8	-4	Y5P	C3'-C4'-C5'-O5'
37	8	1	P5P	O4'-C4'-C5'-O5'
37	8	2	Y5P	C3'-C4'-C5'-O5'
37	8	6	Y5P	O4'-C1'-N1-C2
37	8	-7	Y5P	O4'-C1'-N1-C2

5 of 39 torsion outliers are listed below:

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 7 ligands modelled in this entry, 5 are monoatomic - leaving 2 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	Chain	Dec	Tinle	В	ond leng	gths	Bond angles			
	туре	Chain	nes		Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z >2
41	GTP	С	1102	40	26,34,34	1.15	2 (7%)	32,54,54	1.50	7 (21%)
39	M7M	В	401	-	27,32,33	4.54	16 (59%)	33,49,52	1.45	5 (15%)

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
41	GTP	С	1102	40	-	5/18/38/38	0/3/3/3
39	M7M	В	401	-	-	3/17/47/48	0/3/3/3



Mol	Chain	Res	Type	Atoms	Ζ	Observed(Å)	Ideal(Å)
39	В	401	M7M	CBI-NBP	9.85	1.51	1.46
39	В	401	M7M	CBG-NBH	8.51	1.45	1.35
39	В	401	M7M	OBR-CBS	7.46	1.61	1.45
39	В	401	M7M	CBO-NBP	7.33	1.46	1.35
39	В	401	M7M	CBY-CBS	-6.99	1.35	1.53

The worst 5 of 18 bond length outliers are listed below:

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
39	В	401	M7M	CBO-CBG-NBH	5.07	109.72	106.71
41	С	1102	GTP	C5-C6-N1	3.20	119.60	113.95
41	С	1102	GTP	C8-N7-C5	3.00	108.71	102.99
41	С	1102	GTP	C2-N1-C6	-2.80	119.94	125.10
39	В	401	M7M	NBP-CBI-NBH	2.75	107.31	103.38

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
39	В	401	M7M	CBT-OBU-PBK-OBD
41	С	1102	GTP	C5'-O5'-PA-O1A
41	С	1102	GTP	C5'-O5'-PA-O3A
41	С	1102	GTP	O4'-C4'-C5'-O5'
39	В	401	M7M	CBT-OBU-PBK-OBJ

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. 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.

