



Full wwPDB EM Validation Report ⓘ

Nov 5, 2024 – 10:22 AM JST

PDB ID : 8H6L
EMDB ID : EMD-34508
Title : Cryo-EM structure of human exon-defined spliceosome in the early B state.
Authors : Zhang, W.; Zhan, X.; Zhang, X.; Bai, R.; Lei, J.; Yan, C.; Shi, Y.
Deposited on : 2022-10-18
Resolution : 2.60 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 ⓘ 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 ⓘ](#)) 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) : 2.39

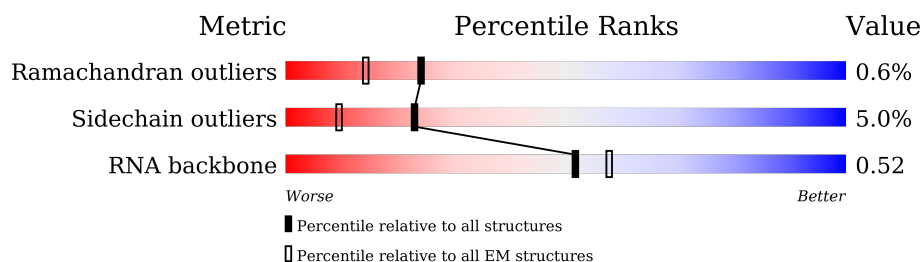
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 2.60 Å.

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



Metric	Whole archive (#Entries)	EM structures (#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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	144	
2	5A	117	
3	5B	2335	
4	5C	972	
5	5D	2136	
6	5E	357	
7	2a	231	
7	4a	231	
7	5a	231	






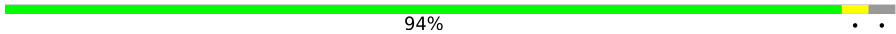
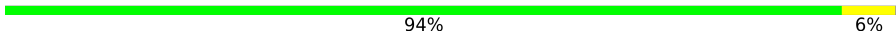

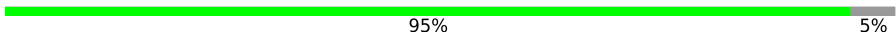



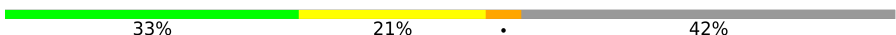







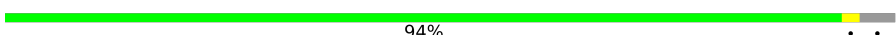




Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	2b	119	
8	4b	119	
8	5b	119	
9	2c	118	
9	4c	118	
9	5c	118	
10	2d	86	
10	4d	86	
10	5d	86	
11	2e	92	
11	4e	92	
11	5e	92	
12	2f	76	
12	4f	76	
12	5f	76	
13	2g	126	
13	4g	126	
13	5g	126	
14	6A	107	
15	6a	95	
16	6b	102	
17	6c	139	
18	6d	91	
19	6e	80	
20	6f	103	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
21	6g	96	
22	4A	145	
23	4B	683	
24	4C	522	
25	4D	499	
26	4E	128	
27	4F	142	
28	4G	941	
29	4H	177	
30	4I	376	
31	4J	800	
32	4Z	513	
33	2A	188	
34	2B	255	
35	2C	225	
36	2D	793	
37	2E	464	
38	2F	501	
39	2G	1304	
40	2H	895	
41	2I	1217	
42	2J	424	
43	2K	125	
44	2L	110	
45	2M	86	

2 Entry composition

There are 49 unique types of molecules in this entry. The entry contains 94667 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 pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	57	Total	C	N	O	P	0	0
			1187	531	183	416	57		

- Molecule 2 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	5A	115	Total	C	N	O	P	0	0
			2420	1084	403	818	115		

- Molecule 3 is a protein called Pre-mRNA-processing-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	5B	2253	Total	C	N	O	S	0	0
			18642	11992	3250	3319	81		

- Molecule 4 is a protein called 116 kDa U5 small nuclear ribonucleoprotein component.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	5C	818	Total	C	N	O	S	0	0
			6436	4114	1085	1205	32		

- Molecule 5 is a protein called U5 small nuclear ribonucleoprotein 200 kDa helicase.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	5D	1696	Total	C	N	O	S	0	0
			13633	8715	2329	2519	70		

- Molecule 6 is a protein called U5 small nuclear ribonucleoprotein 40 kDa protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
6	5E	299	Total	C	N	O	0	0
			1196	598	299	299		

- Molecule 7 is a protein called Isoform SM-B of Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms				AltConf	Trace
7	5a	84	Total	C	N	O	0	0
			336	168	84	84		
7	4a	64	Total	C	N	O	0	0
			256	128	64	64		
7	2a	86	Total	C	N	O	0	0
			344	172	86	86		

- Molecule 8 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms				AltConf	Trace
8	5b	82	Total	C	N	O	0	0
			328	164	82	82		
8	4b	82	Total	C	N	O	0	0
			334	170	82	82		
8	2b	82	Total	C	N	O	0	0
			328	164	82	82		

- Molecule 9 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms				AltConf	Trace
9	5c	97	Total	C	N	O	0	0
			388	194	97	97		
9	4c	74	Total	C	N	O	0	0
			300	152	74	74		
9	2c	85	Total	C	N	O	0	0
			340	170	85	85		

- Molecule 10 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms				AltConf	Trace
10	5d	74	Total	C	N	O	0	0
			296	148	74	74		
10	4d	71	Total	C	N	O	0	0
			292	150	71	71		
10	2d	74	Total	C	N	O	0	0
			296	148	74	74		

- Molecule 11 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms				AltConf	Trace
11	5e	79	Total	C	N	O	0	0
			316	158	79	79		
11	4e	78	Total	C	N	O	0	0
			314	158	78	78		
11	2e	79	Total	C	N	O	0	0
			316	158	79	79		

- Molecule 12 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms				AltConf	Trace
12	5f	72	Total	C	N	O	0	0
			288	144	72	72		
12	4f	73	Total	C	N	O	0	0
			298	152	73	73		
12	2f	68	Total	C	N	O	0	0
			272	136	68	68		

- Molecule 13 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms				AltConf	Trace
13	5g	76	Total	C	N	O	0	0
			304	152	76	76		
13	4g	71	Total	C	N	O	0	0
			288	146	71	71		
13	2g	80	Total	C	N	O	0	0
			320	160	80	80		

- Molecule 14 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	6A	59	Total	C	N	O	P	0	0
			1251	558	230	404	59		

- Molecule 15 is a protein called U6 snRNA-associated Sm-like protein LSm2.

Mol	Chain	Residues	Atoms				AltConf	Trace
15	6a	90	Total	C	N	O	0	0
			360	180	90	90		

- Molecule 16 is a protein called U6 snRNA-associated Sm-like protein LSm3.

Mol	Chain	Residues	Atoms				AltConf	Trace
16	6b	74	Total	C	N	O	0	0
			296	148	74	74		

- Molecule 17 is a protein called U6 snRNA-associated Sm-like protein LSm4.

Mol	Chain	Residues	Atoms				AltConf	Trace
17	6c	74	Total	C	N	O	0	0
			296	148	74	74		

- Molecule 18 is a protein called U6 snRNA-associated Sm-like protein LSm5.

Mol	Chain	Residues	Atoms				AltConf	Trace
18	6d	72	Total	C	N	O	0	0
			288	144	72	72		

- Molecule 19 is a protein called U6 snRNA-associated Sm-like protein LSm6.

Mol	Chain	Residues	Atoms				AltConf	Trace
19	6e	70	Total	C	N	O	0	0
			280	140	70	70		

- Molecule 20 is a protein called U6 snRNA-associated Sm-like protein LSm7.

Mol	Chain	Residues	Atoms				AltConf	Trace
20	6f	65	Total	C	N	O	0	0
			260	130	65	65		

- Molecule 21 is a protein called U6 snRNA-associated Sm-like protein LSm8.

Mol	Chain	Residues	Atoms				AltConf	Trace
21	6g	61	Total	C	N	O	0	0
			244	122	61	61		

- Molecule 22 is a RNA chain called U4 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	4A	129	Total	C	N	O	P	0	0
			2744	1225	472	917	130		

- Molecule 23 is a protein called U4/U6 small nuclear ribonucleoprotein Prp3.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	4B	256	Total	C	N	O	S	0	0
			2076	1316	385	367	8		

- Molecule 24 is a protein called U4/U6 small nuclear ribonucleoprotein Prp4.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	4C	426	Total	C	N	O	S	0	0
			3370	2118	612	620	20		

- Molecule 25 is a protein called U4/U6 small nuclear ribonucleoprotein Prp31.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	4D	376	Total	C	N	O	S	0	0
			2874	1788	524	550	12		

- Molecule 26 is a protein called NHP2-like protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	4E	124	Total	C	N	O	S	0	0
			962	608	171	178	5		

- Molecule 27 is a protein called Thioredoxin-like protein 4A.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	4F	141	Total	C	N	O	S	0	0
			1169	751	194	214	10		

- Molecule 28 is a protein called Pre-mRNA-processing factor 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	4G	801	Total	C	N	O	S	0	0
			5504	3419	1043	1026	16		

- Molecule 29 is a protein called Peptidyl-prolyl cis-trans isomerase H.

Mol	Chain	Residues	Atoms				AltConf	Trace
29	4H	169	Total	C	N	O	0	0
			844	506	169	169		

- Molecule 30 is a protein called WW domain-binding protein 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	4I	75	Total	C	N	O	S	0	0
			494	304	96	91	3		

- Molecule 31 is a protein called U4/U6.U5 tri-snRNP-associated protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	4J	153	Total	C	N	O	S	0	0
			1153	715	206	230	2		

- Molecule 32 is a protein called WD40 repeat-containing protein SMU1.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	4Z	420	Total	C	N	O		0	0
			2093	1253	420	420			

- Molecule 33 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	2A	109	Total	C	N	O	P	0	0
			2311	1032	396	774	109		

- Molecule 34 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	2B	162	Total	C	N	O		0	0
			648	324	162	162			

- Molecule 35 is a protein called U2 small nuclear ribonucleoprotein B''.

Mol	Chain	Residues	Atoms					AltConf	Trace
35	2C	94	Total	C	N	O		0	0
			376	188	94	94			

- Molecule 36 is a protein called Splicing factor 3A subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
36	2D	236	Total	C	N	O	S	0	0
			1380	793	285	299	3		

- Molecule 37 is a protein called Splicing factor 3A subunit 2.

Mol	Chain	Residues	Atoms				AltConf	Trace
37	2E	94	Total	C	N	O	0	0
			376	188	94	94		

- Molecule 38 is a protein called Splicing factor 3A subunit 3.

Mol	Chain	Residues	Atoms				AltConf	Trace
38	2F	423	Total	C	N	O	0	0
			1693	847	423	423		

- Molecule 39 is a protein called Splicing factor 3B subunit 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
39	2G	1048	Total	C	N	O	0	0
			4192	2096	1048	1048		

- Molecule 40 is a protein called Splicing factor 3B subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
40	2H	213	Total	C	N	O	S	0	0
			959	510	220	226	3		

- Molecule 41 is a protein called Splicing factor 3B subunit 3.

Mol	Chain	Residues	Atoms				AltConf	Trace
41	2I	1168	Total	C	N	O	0	0
			4672	2336	1168	1168		

- Molecule 42 is a protein called Splicing factor 3B subunit 4.

Mol	Chain	Residues	Atoms				AltConf	Trace
42	2J	78	Total	C	N	O	0	0
			312	156	78	78		

- Molecule 43 is a protein called Splicing factor 3B subunit 6.

Mol	Chain	Residues	Atoms				AltConf	Trace
43	2K	108	Total	C	N	O	0	0
			432	216	108	108		

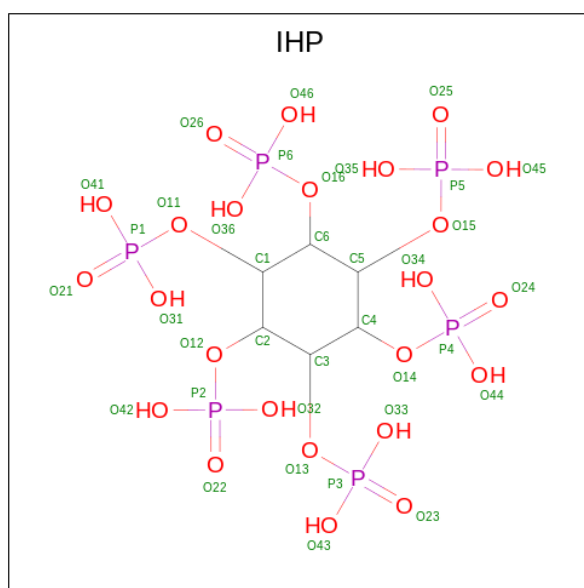
- Molecule 44 is a protein called PHD finger-like domain-containing protein 5A.

Mol	Chain	Residues	Atoms				AltConf	Trace
44	2L	89	Total	C	N	O	0	0
			356	178	89	89		

- Molecule 45 is a protein called Splicing factor 3B subunit 5.

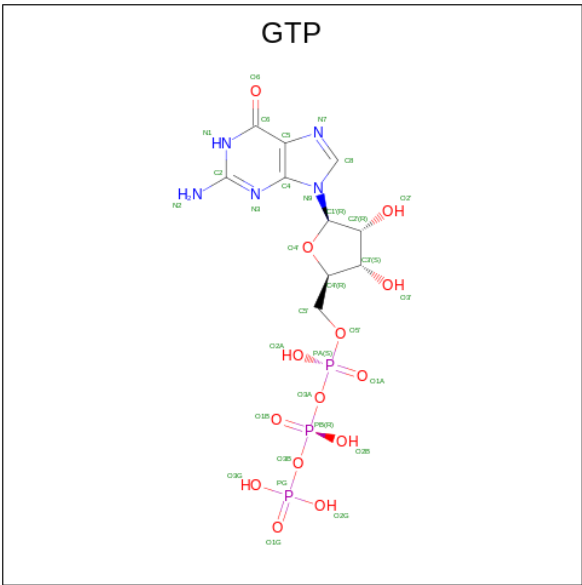
Mol	Chain	Residues	Atoms				AltConf	Trace
45	2M	66	Total	C	N	O	0	0
			264	132	66	66		

- Molecule 46 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula: $C_6H_{18}O_{24}P_6$).



Mol	Chain	Residues	Atoms				AltConf
46	5B	1	Total	C	O	P	0
			36	6	24	6	

- Molecule 47 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).



Mol	Chain	Residues	Atoms					AltConf
47	5C	1	Total	C	N	O	P	0
			32	10	5	14	3	

- Molecule 48 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
48	5C	1	Total	Mg	0
			1	1	

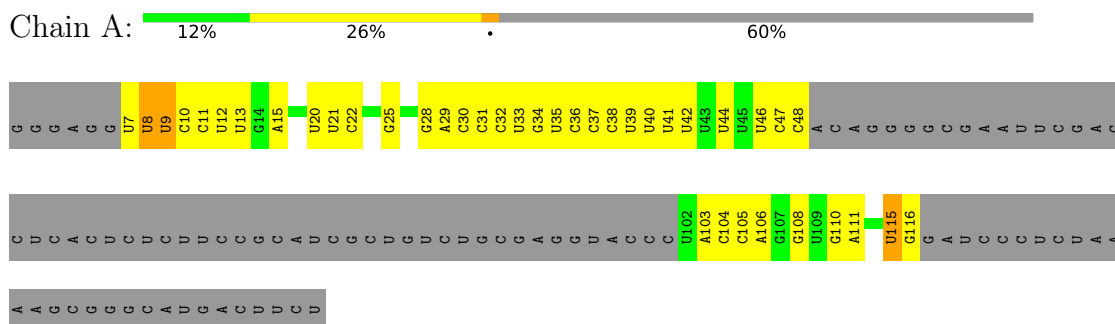
- Molecule 49 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
49	4I	1	Total	Zn	0
			1	1	

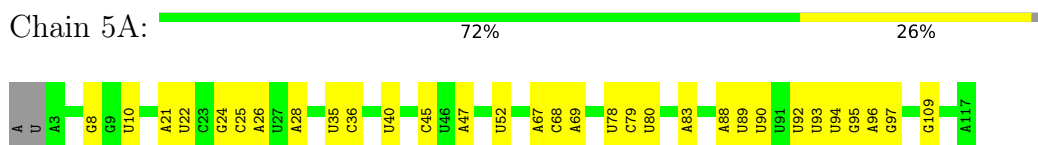
3 Residue-property plots [i](#)

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.

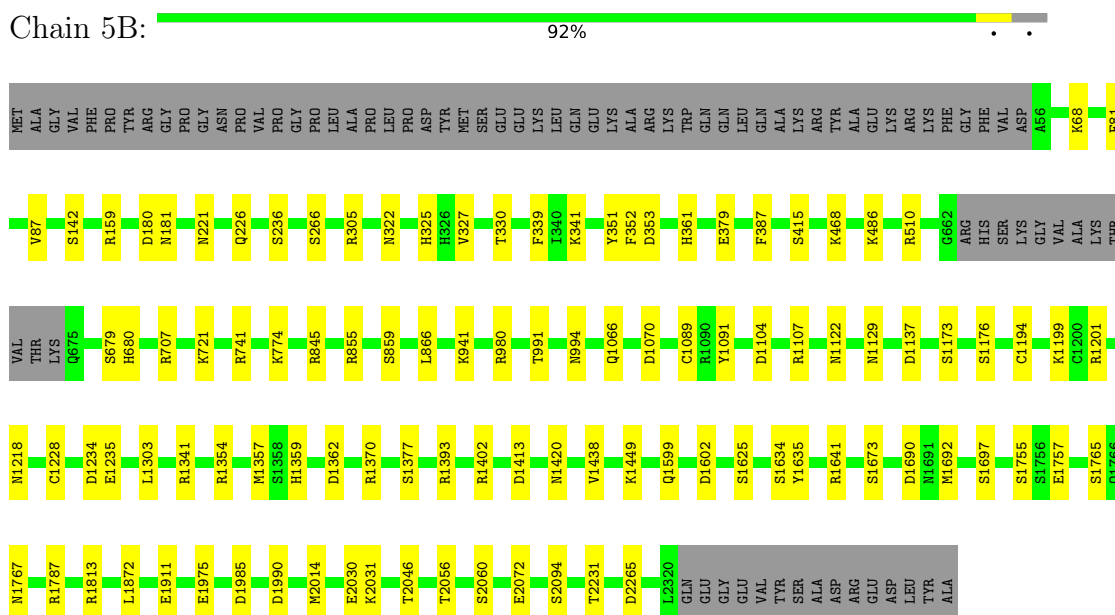
• Molecule 1: pre-mRNA



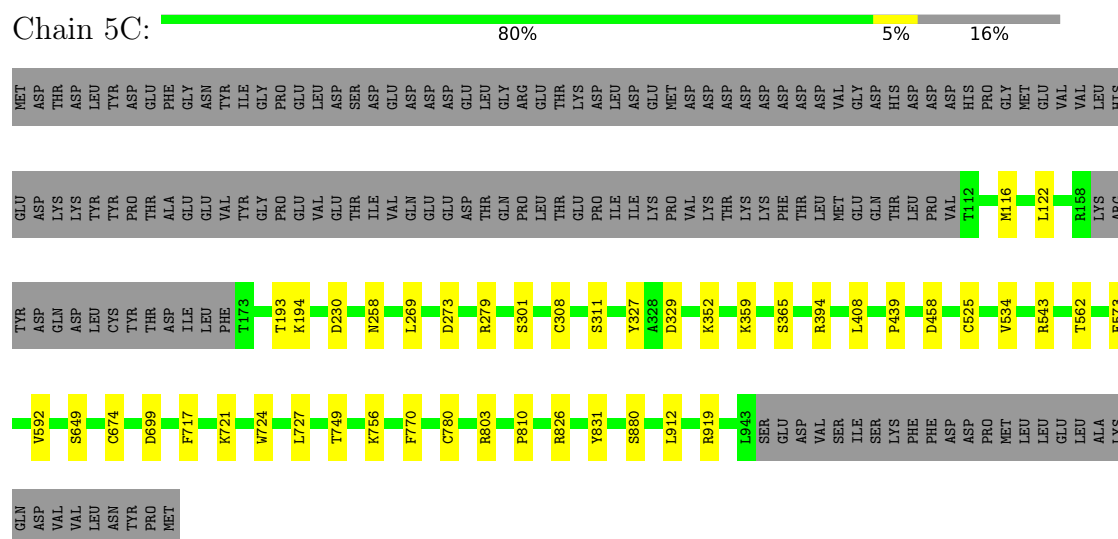
• Molecule 2: U5 snRNA



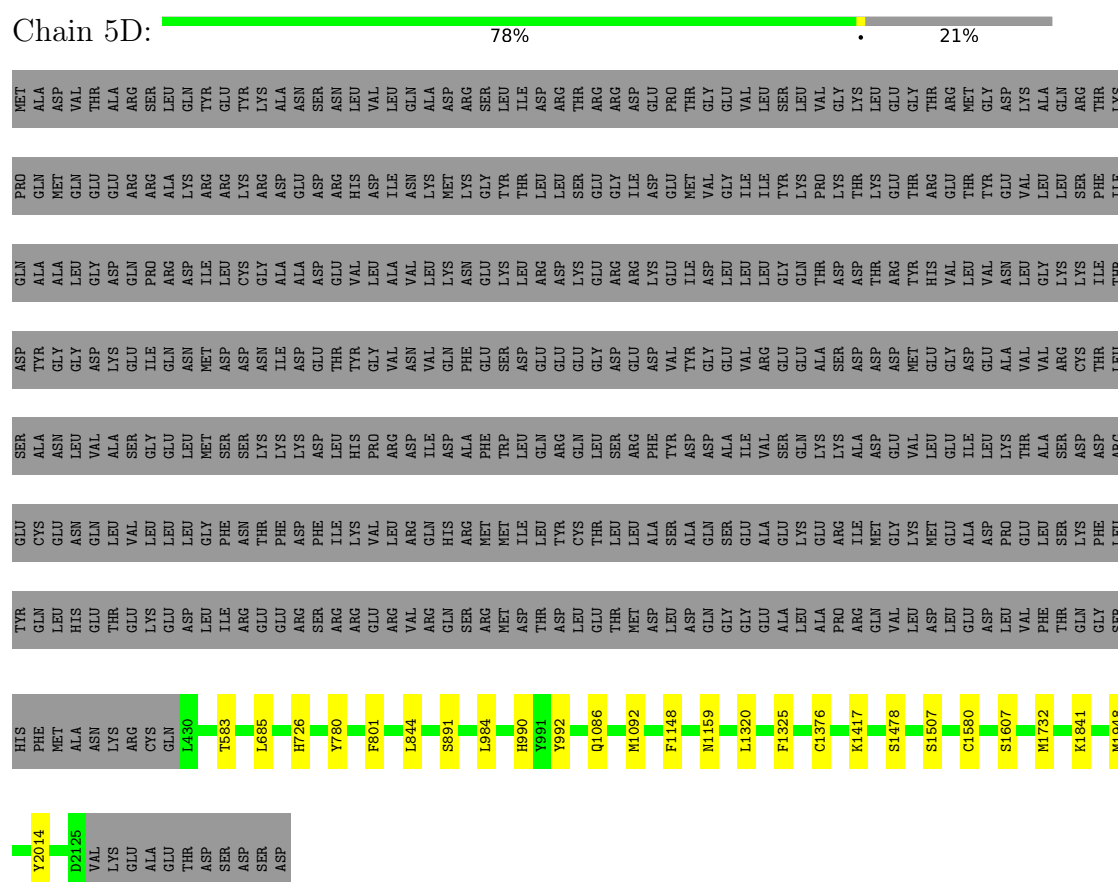
• Molecule 3: Pre-mRNA-processing-splicing factor 8



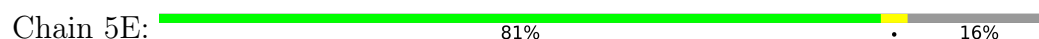
- Molecule 4: 116 kDa U5 small nuclear ribonucleoprotein component



- Molecule 5: U5 small nuclear ribonucleoprotein 200 kDa helicase

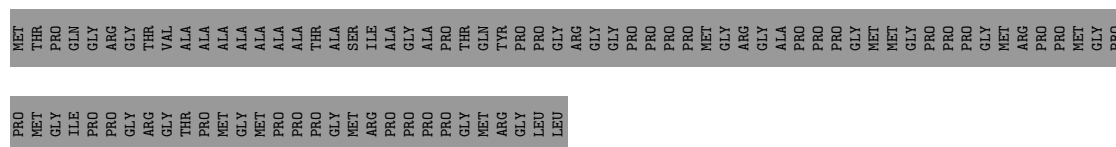


- Molecule 6: U5 small nuclear ribonucleoprotein 40 kDa protein

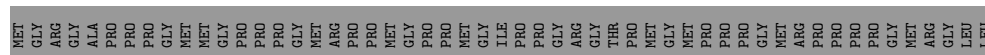




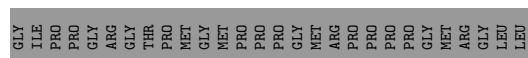
- Chain 5a:  36% 64%



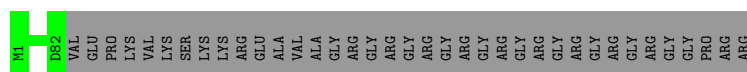
- Chain 4a:  28% 72%



- Chain 2a: 37% 63%



- Chain 5b:  69% 31%



- 
- WORLDWIDE
PDB
PROTEIN DATA BANK

MET	K2	V63
GLU		
PRO		
LYS		
VAL		
LYS		
SER		
LYS		
LYS		
ARG		
GLU		
ALA		
VAL		
ALA		
GLY		
ARG		
GLY		
ARG		
GLY		
ARG		
GLY		
ARG		
GLY		
ARG		
GLY		
ARG		
GLY		
ARG		
GLY		
ARG		
GLY		
PRO		
ARG		

- Chain 2b:  69% 31%

M1	D82	VAL	GLU	PRO	LYS	VAL	LYS	SER	LYS	LYS	ARG	GLU	ALA	VAL	ALA	GLY	ARG	GLY	GLY	ARG	ARG	GLY	ARG	GLY	GLY	ARG	GLY	ARG	GLY	ARG	GLY	ARG	GLY	ARG	GLY	GLY	PRO	ARG	ARG
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Chain 5c: 81% . 18%

MET SER LEU LEU ASN LYS PRO LYS SER GLU MET THR PRO GLU GLU LEU GLN LYS ARG E20 N112 A116 GLY LYS

- Chain 4c: 63% 37%

MET	SER	LEU	LEU	ASN	ASN	LYS	PRO	PRO	SER	GLU	MET	THR	PRO	GLU	GLU	LEU	GLU	LYS	ARG	GLU	GLU	GLU	GLU	PHE	ASN	T26	T75	GLU	VAL	PRO	SER	SER	GLY	LYS	LYS	LYS	LYS	SER	PRO	PRO	VAL	R91	L114	L1E	ALA	GLY	GLY
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	-----	-----	-----	-----

- Chain 2c:  71% : 28%

MET	SER	LEU	LEU	ASN	LYS	PRO	LYS	SER	GLU	MET	THR	PRO	GLU	GLU	LEU	GLN	LYS	ARG	E20	E76	VAL	PRO	LYS	SER	GLY	LYS	GLY	LYS	LYS	SER	LYS	P69	N112	A116	GLY	LYS
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	-----	-----

- Chain 5d: 86% 14%

MET	S2	V75	GLU	GLU	GLU	GLU	GLU	ASP	GLY	GLU	MET	ARG	GLU
-----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Chain 4d: 83% 17%

MET
SER
LEU
P4
G74
VAL
GLU
GLU
GLU
GLU
GLU
ASP
GLY
GLU
MET
ARG
GLU

- Molecule 10: Small nuclear ribonucleoprotein F

MET	S2	V75	GLU	GLU	GLU	GLU	GLU	ASP	GLY	GLU	MET	ARG	GLU
-----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- | |
|-----|
| MET |
| ALA |
| TYR |
| ARG |
| GLY |
| GLN |
| GLY |
| GLN |
| LYS |
| VAL |
| GLN |
| LYS |
| VAL |
| M14 |
| N92 |

- MET
 ALA
 TYR
 ARG
 GLY
 GLN
 GLY
 GLN
 LYS
 VAL
 GLN
 LYS
 VAL
 MET
 V15
 N92

- | |
|-----|
| MET |
| ALA |
| TYR |
| ARG |
| GLY |
| GLN |
| GLY |
| GLN |
| LYS |
| VAL |
| GLN |
| LYS |
| VAL |
| M14 |
| N92 |

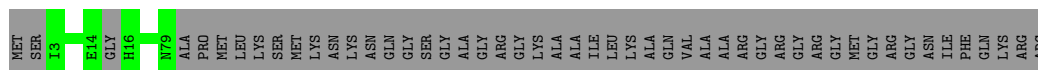
- MET
SER
LYS
ALA
H5
V76

- MET
SER
LYS
- A4
- V76

- MET**
SER
LYS
A4
- M48**
ALA
THR
SER
GLY
GLN
Q54
- V76**

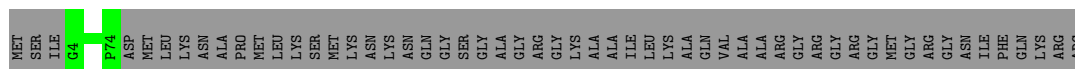
- 
- WORLD WIDE
PDB
PROTEIN DATA BANK

Chain 5g:  60% 40%



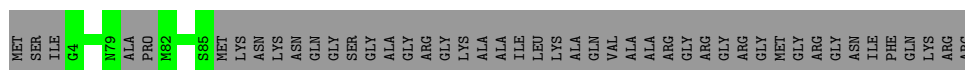
- Molecule 13: Small nuclear ribonucleoprotein Sm D3

Chain 4g:  56% 44%



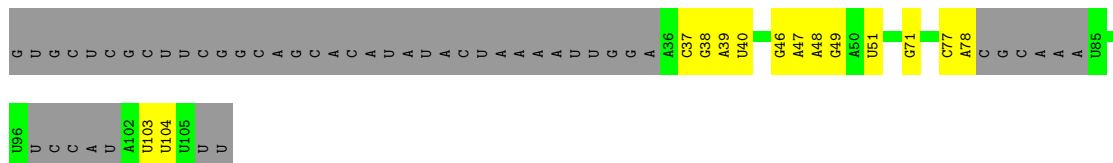
- Molecule 13: Small nuclear ribonucleoprotein Sm D3

Chain 2g:  63% 37%



- Molecule 14: U6 snRNA

Chain 6A:  42% 13% 45%



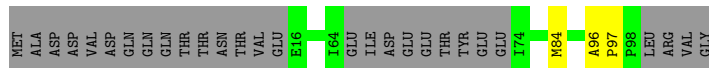
- Molecule 15: U6 snRNA-associated Sm-like protein LSm2

Chain 6a:  91% 5%



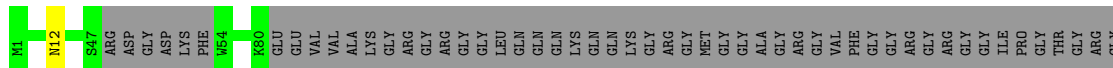
- Molecule 16: U6 snRNA-associated Sm-like protein LSm3

Chain 6b:  70% 27%




- Molecule 17: U6 snRNA-associated Sm-like protein LSm4

Chain 6c:  53% 47%




GLN
PRO
GLU
LYS
LYS
PRO
GLY
ARG
GLN
ALA
GLY
LYS
GLN

- Molecule 18: U6 snRNA-associated Sm-like protein LSm5

Chain 6d:  78% 21%

MET
ALA
ALA
ASN
ALA
THR
THR
ASN
PRO
S10
I59
THR
PRO
GLU
GLY
R64
D70
G85
GLY
GLY
PRO
GLU
VAL

- Molecule 19: U6 snRNA-associated Sm-like protein LSm6

Chain 6e:  85% 12%

MET
SER
LEU
ARG
LYS
GLN
T7
V52
Q55
G76
LYS
ARG
ARG
MET

- Molecule 20: U6 snRNA-associated Sm-like protein LSm7

Chain 6f:  63% 37%

MET
ALA
ASP
LYS
GLU
LYS
LYS
S11
R55
ARG
ASP
PRO
GLY
ASP
GLN
TYR
LYS
LEU
THR
GLU
ASP
T68
Q87
ASP
GLY
MET
GLU
ALA
ILE
ARG
PRO
ASN
PRO
PHE
ILE
GLN
GLN
GLN
ASP
ALA

- Molecule 21: U6 snRNA-associated Sm-like protein LSm8

Chain 6g:  63% 36%

MET
THR
SER
A4
I34
E43
ARG
VAL
PHE
SER
SER
SER
GLN
GLY
VAL
E53
I73
ASP
GLU
GLU
THR
GLU
ASP
SER
SER
ALA
LEU
LEU
LEU
GLY
ASN
ILE
ARG
ALA
GLU
PRO
PRO
LEU
ASN
SER
VAL
ALA
HIS

- Molecule 22: U4 snRNA

Chain 4A:  63% 26% 11%

H7H0
A1
G2
A17
G18
U19
A25
G26
C37
U38
A39
U40
G45
U53
A54
U63
G
A
A
A
A
C
U70
U71
H72
U73
C74
C75
C76
A
A
U
U
A
C
C
C83
C84
G85
G90
C99
A100
A103
G109
U114
G115
A118
A119
U120

U121
U124
G125
A126
C127
G133
U
A
C
G
G138
G144

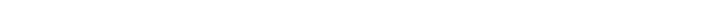
- Molecule 23: U4/U6 small nuclear ribonucleoprotein Prp3

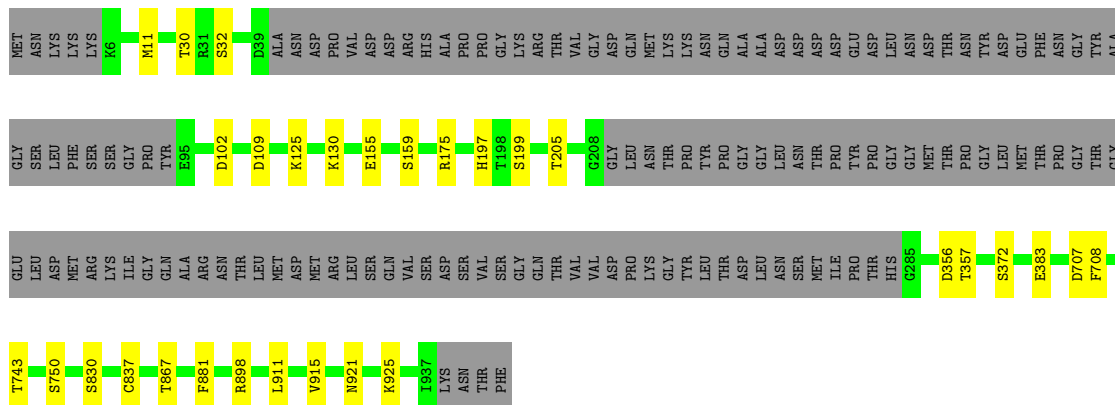
Chain 4B:  35% 63%

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

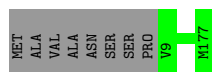
ARG
PHE
VAL
ASP
LYS
LEU
PHE
GLU
ALA
VAL
GLU
GLY
GLY
ARG
SER
SER
HIS
SER
LYS
SER
SER
VAL
SER
LEU
PHE
PHE
GLY
ASN
CYS
VAL
VAL
GLY
LYS
MET
ILE
SER
LYS
GLU
SER
SER
SER
VAL
VAL
LYS
HIS
ASP
HIS
LEU
ARG
LYS
PRO
PHE
PHE
ASP
ASP
GLU
SER
THR
VAL
GLU

MET
S2
Y3
V25
I63
L94
S132
S137
T138
R141
Y142

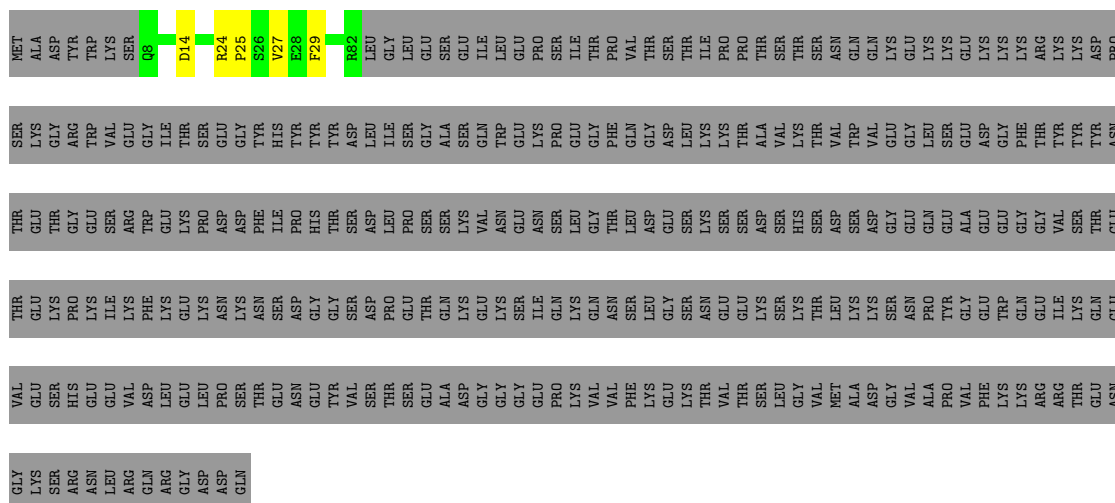
- Chain 4G:  82% 15%



- Chain 4H:  95% 5%



- Chain 4I:  19% . 80%



- WORLDWIDE
PDB
PROTEIN DATA BANK

THR	LYS	ALA	ASN	GLU	PRO	LEU	ARG	VAL	GLU	THR	GLY	MET
ILE	GLU	PRO	ILE	PRO	PRO	LEU	GLU	GLU	GLU	THR	GLU	GLY
VAL	ALA	ASN	GLY	PRO	GLY	PRO	LEU	GLU	GLU	THR	ASN	GLY
LEU	PHE	LYS	TRP	SER	SER	GLY	GLU	PHE	GLY	LYS	GLY	SER
SER	ARG	SER	THR	VAL	GLN	ASP	ILE	GLY	GLN	ARG	GLY	LYS
GLY	GLN	LEU	THR	LYS	GLN	GLY	ARG	GLY	ILE	GLY	LEU	GLY
SER	LEU	PRO	VAL	VAL	VAL	GLN	ARG	GLN	ARG	ARG	ARG	HIS
GLY	SER	SER	ASN	ALA	LEU	THR	ALA	ARG	ALA	ALA	ARG	ARG
LYS	HIS	ALA	LEU	ILE	GLU	GLN	GLN	GLY	LYS	LYS	GLY	GLY
SER	ARG	VAL	ASP	VAL	GLU	ASP	LEU	GLY	GLN	ALA	ALA	GLY
ARG	THR	TYR	GLU	PHE	ASP	GLY	ARG	ASP	ASP	GLY	GLY	LYS
MET	HIS	CYS	GLU	ASN	GLU	ASP	ARG	LEU	ARG	LEU	ALA	GLY
ALA	GLY	ILE	LYS	ALA	ALA	PHE	GLY	GLN	GLY	LEU	ALA	GLY
ASN	LYS	GLU	GLN	THR	GLU	GLY	GLN	D264	LYS	PRO	GLU	ALA
ASN	GLY	ASP	GLN	LEU	LEU	SER	GLN	S265	GLY	PRO	ARG	GLY
THR	GLY	THR	GLN	GLU	GLU	ARG	GLY	GLY	GLY	LEU	SER	GLY
THR	THR	MET	ASP	PHE	LEU	LEU	LEU	E284	SER	VAL	THR	THR
LYS	LYS	ILE	SER	ARG	GLN	ARG	THR	V290	SER	ASN	THR	ALA
MET	LYS	ASP	ALA	THR	LYS	GLY	VAL	E305	GLY	ILE	HIS	ALA
THR	THR	LYS	SER	THR	LEU	GLY	GLY	GLY	GLY	LYS	ARG	GLY
ARG	ARG	TYR	THR	GLU	LYS	ARG	PRO	R310	ARG	GLY	GLY	GLY
MET	MET	SER	THR	PRO	GLY	VAL	ALA	PRO	ASP	GLY	SER	ALA
LYS	LYS	ARG	LEU	THR	ARG	GLY	GLU	TYR	T146	ALA	GLN	ALA
LEU	LEU	GLU	ASP	LEU	LEU	GLY	GLU	LEU	E149	GLU	GLU	THR
LEU	LEU	GLU	GLU	VAL	GLN	VAL	TYR	PRO	GLY	PRO	GLN	GLN
ASP	ASP	TYR	GLU	GLY	GLN	GLU	LEU	TYR	L179	PRO	SER	PRO
GLU	GLU	ARG	PRO	ALA	LEU	GLU	THR	ALA	GLU	ARG	GLU	GLU
GLY	GLY	ILE	ILE	ASN	GLN	LYS	PRO	GLU	ARG	ARG	ARG	ARG
ALA	ALA	PHE	VAL	ASN	GLN	GLY	GLY	ASP	K188	THR	VAL	ARG
LEU	LEU	THR	GLN	GLU	LEU	GLU	GLU	GLU	THR	LEU	LYS	GLU
LYS	LYS	ASP	ARG	GLY	GLN	VAL	VAL	VAL	GLY	GLY	GLY	HIS
MET	MET	PHE	LEU	GLN	LEU	PRO	THR	ASP	E192	ARG	GLU	LYS
SER	SER	GLU	ALA	ARG	ARG	GLN	PHE	ASP	GLY	LYS	LYS	LYS
SER	SER	GLU	ALA	GLU	ASP	PRO	LYS	LEU	K214	ARG	ARG	HIS
SER	SER	LYS	ALA	SER	SER	LEU	LYS	ALA	ASP	ASP	ASP	LYS
SER	SER	ASP	LEU	MET	GLY	PRO	THR	GLN	LEU	ASP	ASP	HIS
ASP	ASP	GLY	LEU	ASP	GLU	LYS	LYS	GLN	ALA	ALA	GLY	ARG
THR	THR	TYR	LEU	PHE	LYS	ASP	ARG	LYS	GLU	TYR	TYR	SER
PRO	PRO	LYS	CYS	GLU	VAL	THR	VAL	PRO	LYS	GLU	GLU	GLY
LEU	LEU	PRO	GLN	ARG	VAL	THR	VAL	ARG	ARG	ALA	ALA	GLY
GLY	GLY	ASP	ASN	ASP	GLU	ARG	LYS	S332	LYS	ALA	ALA	GLY
THR	THR	VAL	LYS	ILE	VAL	VAL	LYS	E342	LEU	LEU	ALA	GLY
ALA	ALA	ILE	LEU	ARG	LYS	ASN	ILE	GLY	GLY	LEU	SER	GLY
LEU	LEU	GLU	LEU	SER	LYS	MET	ARG	GLY	GLY	LEU	SER	SER
LEU	LEU	TYR	GLU	ASN	LEU	ASP	LYS	GLY	GLY	GLY	GLY	GLY
LEU	LEU	TYR	THR	THR	GLU	ILE	GLU	P346	MET	MET	THR	GLY
GLN												

[illegible]

A A C3 C14 U U U U U G G C C U U A A G G C C U A A30 G31 U37 C40 C45 U46 U47 A48 A51 U65 A A C G G C C C C U U C C U A U A U A C C C G G G G C C A A A U U U U



- Chain 2B:  63% . 36%

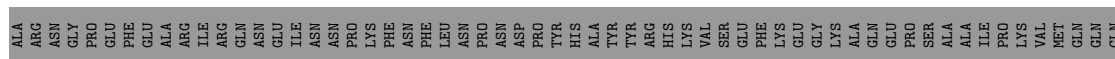


- Chain 2C:  42% 58%



ILE
THR
PRO
SER
HIS
ALA
MET
LYS
ILE
THR
TYR
ALA
LYS
LYS

- Chain 2D:  28% 70%



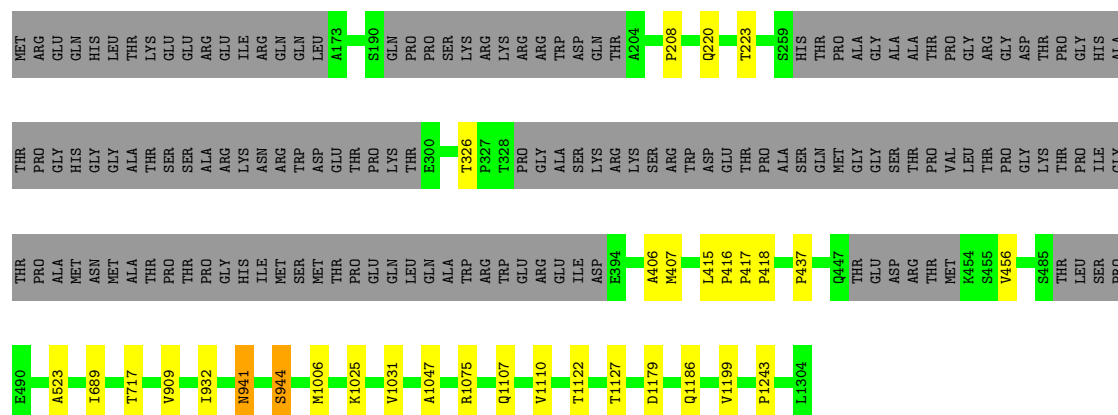
GLN	GLN	THR	GLN	GLN	GLN	LEU	PRO	PRO	GLN	LYS	VAL	VAL	GLN	ALA	GLN	VAL	ILE	ILE	GLN	PRO	PRO	LYS	GLU	PRO	PRO	PRO	GLU	PHE	PHE	ILE	ALA	ASP	PRO	PRO	PRO	SER	ILE	SL160	P221	P222	K223	K228	LEU	LYS	LYS	GLU	ALA	ALA	GLU	GLU	ASN	PRO	ARG	GLU	VAL	L240	V380
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------

V285	ASP	PHE	GLN	PRO	ASN	GLU	GLN	GLY	M294	P298	T299	T300	P301	X302	GLU	LEU	GLY	ALA	ARG	ARG	TYR	LYS	PHE	GLY	GLY	SER	SER	GLU	GLU	VAL	GLU	GLU	GLU	MET	GLU	VAL	GLU	ASP	ASP	LVS	GLN	GLN	GLU	GLU	LVS	ALA	GLU	GLU	PRO	PRO	SER	SER	CYT
------	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

LEU ASP GLN ASP THR GLN VAL GLN ASP MET ASP GLU GLY SER ASP ASP ASP ASP GLU GLU GLY GLY GLN LYS VAL PRO PRO PRO PRO GLU THR PRO MET PRO PRO PRO PRO PRO PRO GLN VAL VAL VAL ARG LYS ASP ASP TYR ASP PRO PRO LYS ALA ALA LYS PRO PRO PRO PRO PRO PRO

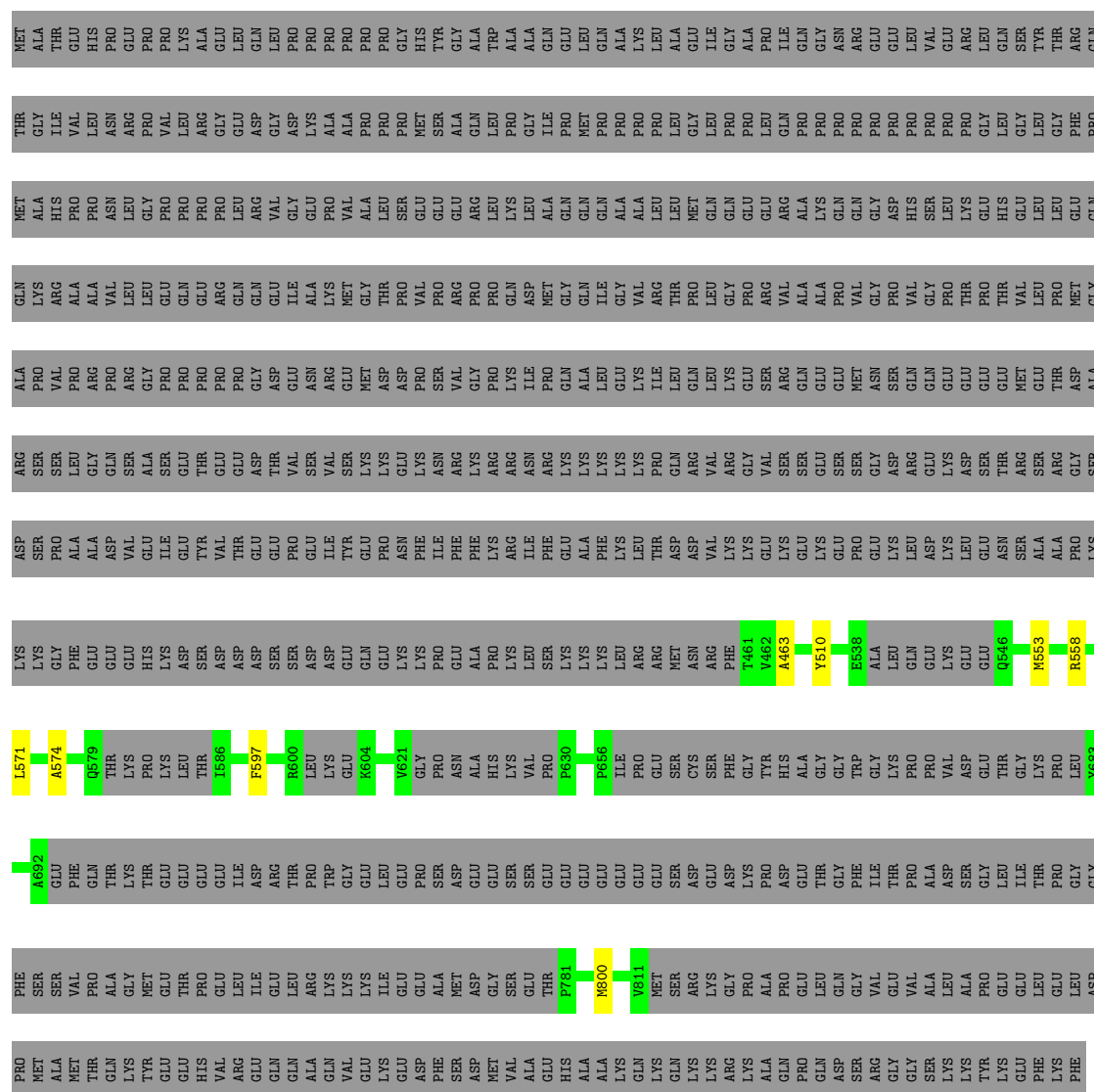
ALA	PRO	ASP	GLU	TYR	L411	S423	Q426	R442	T482	G490	GLU	GLU	GLU	GLN	LYS	PRO	GLU	GLU	R499	S506	G507	S508	T521	L522	H530	LYS	ALA	LYS	GLY	LEU	VAL	PRO	GLU	ASP	ASP	THR	LYS	LYS	GLU	LYS	ILE	GLY	PRO	PRO	LYS	PRO	PRO	GLN	LYS	GLN
-----	-----	-----	-----	-----	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

[illegible]



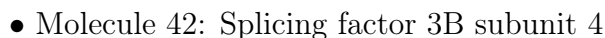
• Molecule 40: Splicing factor 3B subunit 2

Chain 2H: 23% . 76%

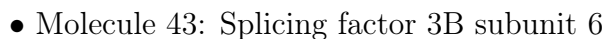


• Molecule 41: Splicing factor 3B subunit 3

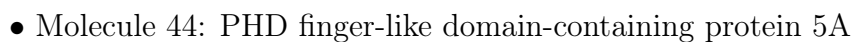
94%



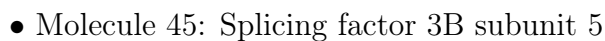
Response	Percentage
Satisfied	82%
Not Satisfied	18%



83% 14%



81% 19%



73% 23%



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	716083	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN 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	GATAN K3 (6k x 4k)	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	0.65	3/1317 (0.2%)	0.90	1/2042 (0.0%)
2	5A	0.29	0/2698	0.82	0/4195
3	5B	0.28	1/19157 (0.0%)	0.51	2/26004 (0.0%)
4	5C	0.27	1/6580 (0.0%)	0.56	3/8938 (0.0%)
5	5D	0.26	0/13923	0.49	1/18868 (0.0%)
6	5E	0.67	0/1195	0.71	0/1492
7	2a	0.50	0/343	0.69	0/427
7	4a	0.22	0/254	0.48	0/314
7	5a	0.50	0/335	0.68	0/417
8	2b	0.56	0/327	0.68	0/407
8	4b	0.22	0/333	0.48	0/416
8	5b	0.57	0/327	0.67	0/407
9	2c	0.70	0/338	0.73	0/419
9	4c	0.23	0/298	0.48	0/370
9	5c	0.69	0/387	0.72	0/482
10	2d	0.77	0/295	0.76	0/367
10	4d	0.24	0/291	0.49	0/363
10	5d	0.77	0/295	0.76	0/367
11	2e	0.64	0/315	0.75	0/392
11	4e	0.22	0/313	0.49	0/390
11	5e	0.65	0/315	0.74	0/392
12	2f	0.55	0/270	0.63	0/334
12	4f	0.24	0/297	0.51	0/371
12	5f	0.54	0/287	0.61	0/357
13	2g	0.47	0/318	0.56	0/394
13	4g	0.23	0/287	0.49	0/358
13	5g	0.46	0/302	0.56	0/374
14	6A	0.30	0/1398	0.81	0/2172
15	6a	0.43	0/359	0.67	0/447
16	6b	0.46	0/294	0.75	0/364
17	6c	0.34	0/294	0.61	0/364
18	6d	0.43	0/286	0.59	0/354

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
19	6e	0.43	0/279	0.72	0/347
20	6f	0.38	0/258	0.61	0/319
21	6g	0.41	0/242	0.64	0/299
22	4A	0.31	0/3025	0.77	1/4702 (0.0%)
23	4B	0.25	0/2114	0.50	0/2836
24	4C	0.25	0/3452	0.53	0/4675
25	4D	0.25	0/2912	0.50	0/3924
26	4E	0.25	0/974	0.47	0/1316
27	4F	0.28	0/1198	0.50	0/1620
28	4G	0.24	0/5592	0.48	1/7615 (0.0%)
29	4H	0.24	0/853	0.45	0/1188
30	4I	0.28	0/502	0.62	2/683 (0.3%)
31	4J	0.25	0/1158	0.52	0/1553
32	4Z	0.24	0/2101	0.45	0/2928
33	2A	0.86	11/2576 (0.4%)	1.43	55/4003 (1.4%)
34	2B	0.63	0/647	1.42	0/807
35	2C	0.61	0/375	1.20	0/467
36	2D	0.23	0/1388	0.48	0/1813
37	2E	0.22	0/373	0.58	1/461 (0.2%)
38	2F	0.25	0/1688	0.47	0/2102
39	2G	1.04	4/4184 (0.1%)	0.83	2/5216 (0.0%)
40	2H	0.65	0/957	0.67	0/1209
41	2I	0.85	0/4664	0.76	0/5816
42	2J	0.62	0/311	0.64	0/387
43	2K	0.79	0/431	0.79	0/537
44	2L	0.74	0/355	0.68	0/442
45	2M	1.01	0/263	0.77	0/327
All	All	0.45	20/96900 (0.0%)	0.65	69/130950 (0.1%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
3	5B	0	1
9	2c	0	1
9	5c	0	1
25	4D	0	1
38	2F	0	1
39	2G	0	11
40	2H	0	3

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
41	2I	0	11
43	2K	0	1
45	2M	0	1
All	All	0	32

All (20) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
39	2G	407	MET	N-CA	12.36	1.71	1.46
39	2G	406	ALA	C-N	7.94	1.52	1.34
33	2A	142	C	C1'-N1	7.32	1.59	1.48
39	2G	1243	PRO	N-CA	-7.11	1.35	1.47
33	2A	182	U	C1'-N1	6.94	1.59	1.48
33	2A	150	U	C1'-N1	6.74	1.58	1.48
33	2A	151	C	C1'-N1	6.53	1.58	1.48
33	2A	97	G	C1'-N9	-6.41	1.37	1.46
33	2A	141	C	C1'-N1	6.38	1.58	1.48
33	2A	184	C	C1'-N1	6.35	1.58	1.48
33	2A	148	C	C1'-N1	6.33	1.58	1.48
39	2G	944	SER	N-CA	-5.72	1.34	1.46
33	2A	65	U	C1'-N1	5.54	1.57	1.48
3	5B	1228	CYS	CB-SG	5.51	1.91	1.82
33	2A	48	A	C1'-N9	-5.48	1.39	1.46
4	5C	810	PRO	CG-CD	-5.20	1.33	1.50
1	A	9	U	C1'-N1	5.14	1.56	1.48
1	A	8	U	C1'-N1	5.13	1.56	1.48
1	A	7	U	C1'-N1	5.13	1.56	1.48
33	2A	110	A	C1'-N9	-5.08	1.39	1.46

All (69) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	5C	810	PRO	CA-N-CD	-14.23	91.58	111.50
33	2A	167	U	C5-C4-O4	11.61	132.87	125.90
39	2G	406	ALA	C-N-CA	10.28	147.39	121.70
33	2A	164	C	N1-C2-O2	-10.12	112.83	118.90
3	5B	1194	CYS	CA-CB-SG	9.56	131.21	114.00
33	2A	162	U	N3-C2-O2	-8.97	115.92	122.20
33	2A	164	C	C5'-C4'-O4'	-8.21	99.25	109.10
33	2A	169	C	P-O3'-C3'	8.20	129.54	119.70
33	2A	166	G	O4'-C1'-N9	8.03	114.62	108.20
33	2A	167	U	N3-C4-O4	-7.89	113.88	119.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
33	2A	164	C	P-O3'-C3'	7.54	128.75	119.70
33	2A	167	U	N1-C2-O2	7.41	127.98	122.80
33	2A	164	C	N3-C2-O2	7.38	127.07	121.90
4	5C	810	PRO	N-CD-CG	-7.37	92.14	103.20
33	2A	113	G	OP2-P-O3'	7.26	121.17	105.20
33	2A	149	A	OP2-P-O3'	7.26	121.16	105.20
33	2A	141	C	OP2-P-O3'	7.24	121.13	105.20
33	2A	114	A	OP2-P-O3'	7.24	121.12	105.20
33	2A	183	G	OP2-P-O3'	7.22	121.08	105.20
33	2A	181	G	OP2-P-O3'	7.21	121.07	105.20
33	2A	182	U	OP2-P-O3'	7.21	121.07	105.20
33	2A	180	G	OP2-P-O3'	7.21	121.06	105.20
33	2A	150	U	OP2-P-O3'	7.20	121.05	105.20
33	2A	148	C	OP2-P-O3'	7.18	121.00	105.20
33	2A	168	A	P-O5'-C5'	-7.14	109.47	120.90
30	4I	25	PRO	N-CD-CG	-6.97	92.74	103.20
33	2A	167	U	N3-C2-O2	-6.96	117.33	122.20
33	2A	180	G	O3'-P-O5'	-6.83	91.02	104.00
33	2A	149	A	O3'-P-O5'	-6.83	91.03	104.00
39	2G	406	ALA	CA-C-O	-6.82	105.77	120.10
33	2A	155	C	P-O3'-C3'	6.82	127.88	119.70
33	2A	182	U	O3'-P-O5'	-6.81	91.06	104.00
4	5C	308	CYS	CA-CB-SG	6.79	126.22	114.00
33	2A	183	G	O3'-P-O5'	-6.79	91.11	104.00
33	2A	148	C	O3'-P-O5'	-6.78	91.12	104.00
33	2A	141	C	O3'-P-O5'	-6.77	91.14	104.00
33	2A	150	U	O3'-P-O5'	-6.75	91.17	104.00
33	2A	113	G	O3'-P-O5'	-6.75	91.19	104.00
33	2A	181	G	O3'-P-O5'	-6.74	91.19	104.00
33	2A	114	A	O3'-P-O5'	-6.73	91.21	104.00
30	4I	25	PRO	CA-N-CD	-6.53	102.36	111.50
33	2A	165	A	O4'-C1'-N9	-6.19	103.25	108.20
33	2A	166	G	N9-C4-C5	6.14	107.86	105.40
33	2A	166	G	C8-N9-C4	-6.11	103.96	106.40
1	A	115	U	P-O3'-C3'	6.11	127.03	119.70
33	2A	162	U	N1-C2-O2	6.07	127.05	122.80
33	2A	166	G	N3-C4-C5	-6.01	125.60	128.60
33	2A	168	A	C5'-C4'-C3'	-5.93	106.51	116.00
33	2A	172	C	P-O3'-C3'	5.80	126.66	119.70
33	2A	156	U	P-O3'-C3'	-5.77	112.78	119.70
37	2E	146	MET	C-N-CA	5.75	146.17	122.00
5	5D	583	THR	C-N-CA	5.75	136.07	121.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
33	2A	167	U	O3'-P-O5'	-5.74	93.09	104.00
33	2A	164	C	C5-C4-N4	-5.72	116.20	120.20
28	4G	707	ASP	C-N-CA	5.66	135.84	121.70
33	2A	157	G	O4'-C1'-N9	-5.50	103.80	108.20
33	2A	166	G	C6-N1-C2	-5.46	121.83	125.10
33	2A	106	G	O5'-P-OP1	5.41	117.19	110.70
3	5B	1194	CYS	N-CA-CB	5.28	120.11	110.60
33	2A	156	U	OP2-P-O3'	5.28	116.82	105.20
33	2A	160	A	P-O5'-C5'	-5.27	112.46	120.90
33	2A	170	C	O4'-C1'-C2'	-5.23	100.57	105.80
22	4A	70	U	C2-N1-C1'	5.20	123.94	117.70
33	2A	164	C	C6-N1-C2	5.17	122.37	120.30
33	2A	157	G	P-O5'-C5'	-5.15	112.66	120.90
33	2A	170	C	N3-C4-C5	-5.14	119.84	121.90
33	2A	156	U	C4'-C3'-C2'	5.08	107.69	102.60
33	2A	162	U	C2-N3-C4	-5.03	123.98	127.00
33	2A	176	G	N9-C4-C5	5.01	107.41	105.40

There are no chirality outliers.

All (32) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
38	2F	443	THR	Peptide
39	2G	1025	LYS	Peptide
39	2G	1122	THR	Peptide
39	2G	1127	THR	Peptide
39	2G	1179	ASP	Peptide
39	2G	1199	VAL	Peptide
39	2G	220	GLN	Peptide
39	2G	415	LEU	Mainchain,Peptide
39	2G	689	ILE	Peptide
39	2G	941	ASN	Peptide
39	2G	944	SER	Peptide
40	2H	553	MET	Peptide
40	2H	558	ARG	Peptide
40	2H	571	LEU	Peptide
41	2I	261	PHE	Peptide
41	2I	366	ASP	Peptide
41	2I	468	ASP	Peptide
41	2I	530	ASP	Peptide
41	2I	534	ASN	Peptide
41	2I	552	ARG	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
41	2I	670	GLN	Peptide
41	2I	678	VAL	Peptide
41	2I	74	THR	Peptide
41	2I	980	LYS	Peptide
41	2I	986	ILE	Peptide
43	2K	29	LYS	Peptide
45	2M	74	GLN	Peptide
9	2c	112	ASN	Peptide
25	4D	358	ARG	Sidechain
3	5B	941	LYS	Peptide
9	5c	112	ASN	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	Percentiles	
3	5B	2249/2335 (96%)	2145 (95%)	104 (5%)	0	100	100
4	5C	814/972 (84%)	745 (92%)	68 (8%)	1 (0%)	48	71
5	5D	1694/2136 (79%)	1618 (96%)	75 (4%)	1 (0%)	48	71
6	5E	297/357 (83%)	272 (92%)	16 (5%)	9 (3%)	3	6
7	2a	84/231 (36%)	82 (98%)	2 (2%)	0	100	100
7	4a	60/231 (26%)	57 (95%)	3 (5%)	0	100	100
7	5a	82/231 (36%)	80 (98%)	2 (2%)	0	100	100
8	2b	80/119 (67%)	77 (96%)	3 (4%)	0	100	100
8	4b	80/119 (67%)	76 (95%)	4 (5%)	0	100	100
8	5b	80/119 (67%)	77 (96%)	3 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	2c	81/118 (69%)	78 (96%)	3 (4%)	0	100	100
9	4c	70/118 (59%)	68 (97%)	2 (3%)	0	100	100
9	5c	95/118 (80%)	91 (96%)	4 (4%)	0	100	100
10	2d	72/86 (84%)	68 (94%)	4 (6%)	0	100	100
10	4d	69/86 (80%)	67 (97%)	2 (3%)	0	100	100
10	5d	72/86 (84%)	69 (96%)	3 (4%)	0	100	100
11	2e	77/92 (84%)	76 (99%)	1 (1%)	0	100	100
11	4e	76/92 (83%)	70 (92%)	6 (8%)	0	100	100
11	5e	77/92 (84%)	76 (99%)	1 (1%)	0	100	100
12	2f	64/76 (84%)	62 (97%)	2 (3%)	0	100	100
12	4f	71/76 (93%)	67 (94%)	4 (6%)	0	100	100
12	5f	70/76 (92%)	68 (97%)	2 (3%)	0	100	100
13	2g	76/126 (60%)	75 (99%)	1 (1%)	0	100	100
13	4g	69/126 (55%)	69 (100%)	0	0	100	100
13	5g	72/126 (57%)	70 (97%)	2 (3%)	0	100	100
15	6a	88/95 (93%)	77 (88%)	7 (8%)	4 (4%)	2	2
16	6b	70/102 (69%)	64 (91%)	3 (4%)	3 (4%)	2	3
17	6c	70/139 (50%)	63 (90%)	6 (9%)	1 (1%)	9	19
18	6d	68/91 (75%)	63 (93%)	4 (6%)	1 (2%)	8	18
19	6e	68/80 (85%)	64 (94%)	2 (3%)	2 (3%)	3	6
20	6f	61/103 (59%)	56 (92%)	5 (8%)	0	100	100
21	6g	57/96 (59%)	52 (91%)	4 (7%)	1 (2%)	7	14
23	4B	248/683 (36%)	229 (92%)	19 (8%)	0	100	100
24	4C	422/522 (81%)	388 (92%)	33 (8%)	1 (0%)	44	66
25	4D	372/499 (74%)	354 (95%)	18 (5%)	0	100	100
26	4E	122/128 (95%)	112 (92%)	10 (8%)	0	100	100
27	4F	139/142 (98%)	134 (96%)	5 (4%)	0	100	100
28	4G	795/941 (84%)	745 (94%)	50 (6%)	0	100	100
29	4H	167/177 (94%)	156 (93%)	11 (7%)	0	100	100
30	4I	73/376 (19%)	71 (97%)	2 (3%)	0	100	100
31	4J	143/800 (18%)	136 (95%)	7 (5%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
32	4Z	414/513 (81%)	401 (97%)	12 (3%)	1 (0%)	44	66
34	2B	160/255 (63%)	146 (91%)	12 (8%)	2 (1%)	10	21
35	2C	92/225 (41%)	90 (98%)	2 (2%)	0	100	100
36	2D	226/793 (28%)	208 (92%)	12 (5%)	6 (3%)	4	7
37	2E	88/464 (19%)	63 (72%)	16 (18%)	9 (10%)	0	0
38	2F	413/501 (82%)	367 (89%)	41 (10%)	5 (1%)	11	24
39	2G	1032/1304 (79%)	844 (82%)	166 (16%)	22 (2%)	5	11
40	2H	199/895 (22%)	179 (90%)	16 (8%)	4 (2%)	6	12
41	2I	1152/1217 (95%)	1053 (91%)	89 (8%)	10 (1%)	14	31
42	2J	76/424 (18%)	75 (99%)	1 (1%)	0	100	100
43	2K	106/125 (85%)	85 (80%)	18 (17%)	3 (3%)	4	7
44	2L	87/110 (79%)	74 (85%)	13 (15%)	0	100	100
45	2M	64/86 (74%)	55 (86%)	7 (11%)	2 (3%)	3	5
All	All	13703/20230 (68%)	12707 (93%)	908 (7%)	88 (1%)	24	43

All (88) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	5D	1086	GLN
6	5E	193	THR
15	6a	55	LEU
16	6b	84	MET
18	6d	70	ASP
19	6e	52	VAL
19	6e	55	GLN
24	4C	459	PRO
32	4Z	383	CYS
36	2D	301	PRO
37	2E	139	PRO
37	2E	141	ILE
37	2E	146	MET
37	2E	162	PRO
37	2E	165	ARG
37	2E	218	PRO
38	2F	284	ARG
39	2G	208	PRO
39	2G	416	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
39	2G	418	PRO
39	2G	456	VAL
39	2G	717	THR
39	2G	941	ASN
39	2G	1107	GLN
41	2I	405	SER
41	2I	919	SER
43	2K	99	GLN
43	2K	105	LYS
15	6a	74	ALA
16	6b	97	PRO
17	6c	12	ASN
34	2B	160	LYS
36	2D	223	LYS
36	2D	280	VAL
38	2F	277	THR
39	2G	113	ALA
39	2G	1110	VAL
40	2H	597	PHE
41	2I	917	PRO
6	5E	60	MET
6	5E	88	ARG
6	5E	256	ASP
38	2F	177	ARG
38	2F	393	PRO
40	2H	510	TYR
6	5E	162	ARG
16	6b	96	ALA
34	2B	32	PRO
36	2D	300	THR
39	2G	112	ILE
39	2G	437	PRO
39	2G	523	ALA
39	2G	909	VAL
39	2G	1006	MET
40	2H	463	ALA
40	2H	574	ALA
41	2I	529	ALA
41	2I	578	THR
43	2K	75	ASP
6	5E	159	PRO
15	6a	73	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
37	2E	147	PRO
37	2E	217	PRO
39	2G	1047	ALA
39	2G	1075	ARG
39	2G	1186	GLN
41	2I	95	SER
41	2I	229	GLU
6	5E	270	LYS
21	6g	34	ILE
37	2E	220	PRO
39	2G	326	THR
39	2G	932	ILE
41	2I	918	ARG
41	2I	1138	HIS
45	2M	56	ALA
6	5E	149	GLY
39	2G	417	PRO
36	2D	221	PRO
39	2G	223	THR
4	5C	439	PRO
38	2F	229	TRP
41	2I	1204	VAL
6	5E	324	PRO
36	2D	298	PRO
39	2G	1031	VAL
45	2M	64	VAL
15	6a	52	PRO

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all 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	
3	5B	2034/2108 (96%)	1932 (95%)	102 (5%)	20	43
4	5C	718/866 (83%)	676 (94%)	42 (6%)	17	36
5	5D	1517/1908 (80%)	1493 (98%)	24 (2%)	58	79

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
23	4B	225/599 (38%)	205 (91%)	20 (9%)	8	17
24	4C	362/442 (82%)	330 (91%)	32 (9%)	8	17
25	4D	299/424 (70%)	278 (93%)	21 (7%)	12	27
26	4E	108/111 (97%)	104 (96%)	4 (4%)	29	55
27	4F	129/130 (99%)	121 (94%)	8 (6%)	15	33
28	4G	417/792 (53%)	388 (93%)	29 (7%)	12	27
29	4H	10/148 (7%)	10 (100%)	0	100	100
30	4I	32/333 (10%)	28 (88%)	4 (12%)	3	7
31	4J	113/681 (17%)	102 (90%)	11 (10%)	6	14
32	4Z	11/450 (2%)	11 (100%)	0	100	100
36	2D	95/709 (13%)	87 (92%)	8 (8%)	9	19
40	2H	26/776 (3%)	25 (96%)	1 (4%)	28	54
All	All	6096/10477 (58%)	5790 (95%)	306 (5%)	23	43

All (306) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	5B	68	LYS
3	5B	81	PHE
3	5B	87	VAL
3	5B	142	SER
3	5B	159	ARG
3	5B	180	ASP
3	5B	181	ASN
3	5B	221	ASN
3	5B	226	GLN
3	5B	236	SER
3	5B	266	SER
3	5B	305	ARG
3	5B	322	ASN
3	5B	325	HIS
3	5B	327	VAL
3	5B	330	THR
3	5B	339	PHE
3	5B	341	LYS
3	5B	351	TYR
3	5B	352	PHE
3	5B	353	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	5B	361	HIS
3	5B	379	GLU
3	5B	387	PHE
3	5B	415	SER
3	5B	468	LYS
3	5B	486	LYS
3	5B	510	ARG
3	5B	679	SER
3	5B	680	HIS
3	5B	707	ARG
3	5B	721	LYS
3	5B	741	ARG
3	5B	774	LYS
3	5B	845	ARG
3	5B	855	ARG
3	5B	859	SER
3	5B	866	LEU
3	5B	980	ARG
3	5B	991	THR
3	5B	994	ASN
3	5B	1066	GLN
3	5B	1070	ASP
3	5B	1089	CYS
3	5B	1091	TYR
3	5B	1104	ASP
3	5B	1107	ARG
3	5B	1122	ASN
3	5B	1129	ASN
3	5B	1137	ASP
3	5B	1173	SER
3	5B	1176	SER
3	5B	1199	LYS
3	5B	1201	ARG
3	5B	1218	ASN
3	5B	1234	ASP
3	5B	1235	GLU
3	5B	1303	LEU
3	5B	1341	ARG
3	5B	1354	ARG
3	5B	1357	MET
3	5B	1359	HIS
3	5B	1362	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	5B	1370	ARG
3	5B	1377	SER
3	5B	1393	ARG
3	5B	1402	ARG
3	5B	1413	ASP
3	5B	1420	ASN
3	5B	1438	VAL
3	5B	1449	LYS
3	5B	1599	GLN
3	5B	1602	ASP
3	5B	1625	SER
3	5B	1634	SER
3	5B	1635	TYR
3	5B	1641	ARG
3	5B	1673	SER
3	5B	1690	ASP
3	5B	1692	MET
3	5B	1697	SER
3	5B	1755	SER
3	5B	1757	GLU
3	5B	1765	SER
3	5B	1767	ASN
3	5B	1787	ARG
3	5B	1813	ARG
3	5B	1872	LEU
3	5B	1911	GLU
3	5B	1975	GLU
3	5B	1985	ASP
3	5B	1990	ASP
3	5B	2014	MET
3	5B	2030	GLU
3	5B	2031	LYS
3	5B	2046	THR
3	5B	2056	THR
3	5B	2060	SER
3	5B	2072	GLU
3	5B	2094	SER
3	5B	2231	THR
3	5B	2265	ASP
4	5C	116	MET
4	5C	122	LEU
4	5C	193	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	5C	194	LYS
4	5C	230	ASP
4	5C	258	ASN
4	5C	269	LEU
4	5C	273	ASP
4	5C	279	ARG
4	5C	301	SER
4	5C	311	SER
4	5C	327	TYR
4	5C	329	ASP
4	5C	352	LYS
4	5C	359	LYS
4	5C	365	SER
4	5C	394	ARG
4	5C	408	LEU
4	5C	458	ASP
4	5C	525	CYS
4	5C	534	VAL
4	5C	543	ARG
4	5C	562	THR
4	5C	573	GLU
4	5C	592	VAL
4	5C	649	SER
4	5C	674	CYS
4	5C	699	ASP
4	5C	717	PHE
4	5C	721	LYS
4	5C	724	TRP
4	5C	727	LEU
4	5C	749	THR
4	5C	756	LYS
4	5C	770	PHE
4	5C	780	CYS
4	5C	803	ARG
4	5C	826	ARG
4	5C	831	TYR
4	5C	880	SER
4	5C	912	LEU
4	5C	919	ARG
5	5D	685	LEU
5	5D	726	HIS
5	5D	780	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	5D	801	PHE
5	5D	844	LEU
5	5D	891	SER
5	5D	984	LEU
5	5D	990	HIS
5	5D	992	TYR
5	5D	1092	MET
5	5D	1148	PHE
5	5D	1159	ASN
5	5D	1320	LEU
5	5D	1325	PHE
5	5D	1376	CYS
5	5D	1417	LYS
5	5D	1478	SER
5	5D	1507	SER
5	5D	1580	CYS
5	5D	1607	SER
5	5D	1732	MET
5	5D	1841	LYS
5	5D	1948	MET
5	5D	2014	TYR
23	4B	424	LEU
23	4B	434	VAL
23	4B	459	GLU
23	4B	518	ARG
23	4B	526	LYS
23	4B	531	LYS
23	4B	538	SER
23	4B	548	VAL
23	4B	561	GLU
23	4B	577	LYS
23	4B	580	ASN
23	4B	595	LYS
23	4B	605	ASP
23	4B	639	LYS
23	4B	640	ASP
23	4B	642	SER
23	4B	653	THR
23	4B	658	ARG
23	4B	659	GLU
23	4B	663	LYS
24	4C	90	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
24	4C	92	ARG
24	4C	105	ASP
24	4C	106	SER
24	4C	143	ASP
24	4C	147	LYS
24	4C	150	LYS
24	4C	166	TRP
24	4C	168	HIS
24	4C	199	HIS
24	4C	200	LYS
24	4C	211	MET
24	4C	223	ASN
24	4C	231	ASP
24	4C	262	ASP
24	4C	265	LEU
24	4C	288	SER
24	4C	292	LYS
24	4C	304	SER
24	4C	334	SER
24	4C	355	GLN
24	4C	356	GLU
24	4C	370	ASP
24	4C	378	SER
24	4C	395	ARG
24	4C	399	CYS
24	4C	403	LEU
24	4C	407	LEU
24	4C	451	LEU
24	4C	456	LYS
24	4C	460	ILE
24	4C	483	SER
25	4D	116	ASP
25	4D	117	LYS
25	4D	119	SER
25	4D	120	LYS
25	4D	122	PHE
25	4D	137	ARG
25	4D	144	ASN
25	4D	150	LYS
25	4D	163	THR
25	4D	168	SER
25	4D	177	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
25	4D	214	PHE
25	4D	260	SER
25	4D	311	SER
25	4D	362	MET
25	4D	391	ASP
25	4D	395	SER
25	4D	402	SER
25	4D	421	SER
25	4D	426	ARG
25	4D	431	GLN
26	4E	51	SER
26	4E	81	VAL
26	4E	85	SER
26	4E	122	SER
27	4F	3	TYR
27	4F	25	VAL
27	4F	63	ILE
27	4F	94	LEU
27	4F	132	SER
27	4F	137	SER
27	4F	138	THR
27	4F	141	ARG
28	4G	11	MET
28	4G	30	THR
28	4G	32	SER
28	4G	102	ASP
28	4G	109	ASP
28	4G	125	LYS
28	4G	130	LYS
28	4G	155	GLU
28	4G	159	SER
28	4G	175	ARG
28	4G	197	HIS
28	4G	199	SER
28	4G	205	THR
28	4G	356	ASP
28	4G	357	THR
28	4G	372	SER
28	4G	383	GLU
28	4G	708	PHE
28	4G	743	THR
28	4G	750	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
28	4G	830	SER
28	4G	837	CYS
28	4G	867	THR
28	4G	881	PHE
28	4G	898	ARG
28	4G	911	LEU
28	4G	915	VAL
28	4G	921	ASN
28	4G	925	LYS
30	4I	14	ASP
30	4I	24	ARG
30	4I	27	VAL
30	4I	29	PHE
31	4J	149	GLU
31	4J	179	LEU
31	4J	188	LYS
31	4J	264	ASP
31	4J	265	SER
31	4J	284	GLU
31	4J	290	VAL
31	4J	305	GLU
31	4J	332	SER
31	4J	347	HIS
31	4J	350	ARG
36	2D	423	SER
36	2D	426	GLN
36	2D	442	ARG
36	2D	482	THR
36	2D	506	SER
36	2D	508	SER
36	2D	521	THR
36	2D	522	LEU
40	2H	800	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (31) such sidechains are listed below:

Mol	Chain	Res	Type
3	5B	73	HIS
3	5B	105	ASN
3	5B	545	HIS
3	5B	680	HIS
3	5B	1487	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	5B	1563	HIS
3	5B	1791	HIS
3	5B	2166	HIS
4	5C	154	HIS
4	5C	208	HIS
4	5C	245	HIS
4	5C	502	HIS
4	5C	627	HIS
4	5C	642	HIS
5	5D	785	HIS
5	5D	911	GLN
5	5D	1515	HIS
23	4B	480	ASN
23	4B	511	HIS
23	4B	515	ASN
24	4C	282	HIS
24	4C	322	HIS
24	4C	364	HIS
24	4C	421	HIS
25	4D	270	HIS
26	4E	17	HIS
27	4F	89	HIS
28	4G	741	HIS
28	4G	908	HIS
31	4J	261	HIS
36	2D	505	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	55/144 (38%)	35 (63%)	10 (18%)
14	6A	55/107 (51%)	14 (25%)	2 (3%)
2	5A	114/117 (97%)	30 (26%)	5 (4%)
22	4A	124/145 (85%)	35 (28%)	4 (3%)
33	2A	105/188 (55%)	22 (20%)	3 (2%)
All	All	453/701 (64%)	136 (30%)	24 (5%)

All (136) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	8	U

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	9	U
1	A	10	C
1	A	11	C
1	A	12	U
1	A	13	U
1	A	15	A
1	A	20	U
1	A	21	U
1	A	22	C
1	A	25	G
1	A	29	A
1	A	30	C
1	A	31	C
1	A	32	C
1	A	33	U
1	A	34	G
1	A	35	U
1	A	36	C
1	A	37	C
1	A	39	U
1	A	41	U
1	A	42	U
1	A	44	U
1	A	46	U
1	A	47	C
1	A	48	C
1	A	103	A
1	A	104	C
1	A	105	C
1	A	106	A
1	A	108	G
1	A	110	G
1	A	111	A
1	A	116	G
2	5A	8	G
2	5A	10	U
2	5A	21	A
2	5A	22	U
2	5A	24	G
2	5A	25	C
2	5A	26	A
2	5A	28	A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	5A	35	U
2	5A	36	C
2	5A	40	U
2	5A	45	C
2	5A	47	A
2	5A	52	U
2	5A	68	C
2	5A	69	A
2	5A	78	U
2	5A	79	C
2	5A	80	U
2	5A	83	A
2	5A	88	A
2	5A	89	U
2	5A	90	U
2	5A	92	U
2	5A	93	U
2	5A	94	U
2	5A	95	G
2	5A	96	A
2	5A	97	G
2	5A	109	G
14	6A	37	C
14	6A	38	G
14	6A	39	A
14	6A	40	U
14	6A	46	G
14	6A	47	A
14	6A	48	A
14	6A	49	G
14	6A	51	U
14	6A	71	G
14	6A	77	C
14	6A	78	A
14	6A	103	U
14	6A	104	U
22	4A	2	G
22	4A	17	A
22	4A	18	G
22	4A	19	U
22	4A	25	A
22	4A	26	G

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
22	4A	37	C
22	4A	38	U
22	4A	39	A
22	4A	40	U
22	4A	45	G
22	4A	53	U
22	4A	54	A
22	4A	71	U
22	4A	73	U
22	4A	74	C
22	4A	75	C
22	4A	76	C
22	4A	84	C
22	4A	85	G
22	4A	90	G
22	4A	100	A
22	4A	103	A
22	4A	109	G
22	4A	114	U
22	4A	115	G
22	4A	118	A
22	4A	119	A
22	4A	120	U
22	4A	121	U
22	4A	124	U
22	4A	125	G
22	4A	126	A
22	4A	127	C
22	4A	144	G
33	2A	31	G
33	2A	37	U
33	2A	40	C
33	2A	45	C
33	2A	47	U
33	2A	51	A
33	2A	65	U
33	2A	112	G
33	2A	143	A
33	2A	147	G
33	2A	152	G
33	2A	153	A
33	2A	154	C

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
33	2A	156	U
33	2A	157	G
33	2A	164	C
33	2A	165	A
33	2A	168	A
33	2A	169	C
33	2A	177	A
33	2A	178	A
33	2A	179	C

All (24) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	9	U
1	A	21	U
1	A	28	G
1	A	33	U
1	A	35	U
1	A	36	C
1	A	38	C
1	A	40	U
1	A	41	U
1	A	115	U
2	5A	67	A
2	5A	78	U
2	5A	79	C
2	5A	94	U
2	5A	96	A
14	6A	37	C
14	6A	77	C
22	4A	18	G
22	4A	38	U
22	4A	99	C
22	4A	114	U
33	2A	156	U
33	2A	164	C
33	2A	168	A

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

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

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 2 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	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
47	GTP	5C	1500	48	26,34,34	1.13	2 (7%)	32,54,54	1.53	7 (21%)
46	IHP	5B	3000	-	36,36,36	0.73	0	54,60,60	1.07	2 (3%)

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
47	GTP	5C	1500	48	-	6/18/38/38	0/3/3/3
46	IHP	5B	3000	-	-	3/30/54/54	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
47	5C	1500	GTP	C5-C6	-4.03	1.39	1.47
47	5C	1500	GTP	C2-N3	2.08	1.38	1.33

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
47	5C	1500	GTP	PB-O3B-PG	-3.28	121.57	132.83
47	5C	1500	GTP	C5-C6-N1	3.26	119.72	113.95
47	5C	1500	GTP	PA-O3A-PB	-3.15	122.01	132.83
47	5C	1500	GTP	C8-N7-C5	3.01	108.72	102.99

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
47	5C	1500	GTP	C2-N1-C6	-2.88	119.80	125.10
47	5C	1500	GTP	C3'-C2'-C1'	2.51	104.75	100.98
47	5C	1500	GTP	O6-C6-C5	-2.23	120.01	124.37
46	5B	3000	IHP	C6-C5-C4	2.08	114.96	110.41
46	5B	3000	IHP	C5-C4-C3	2.02	114.82	110.41

There are no chirality outliers.

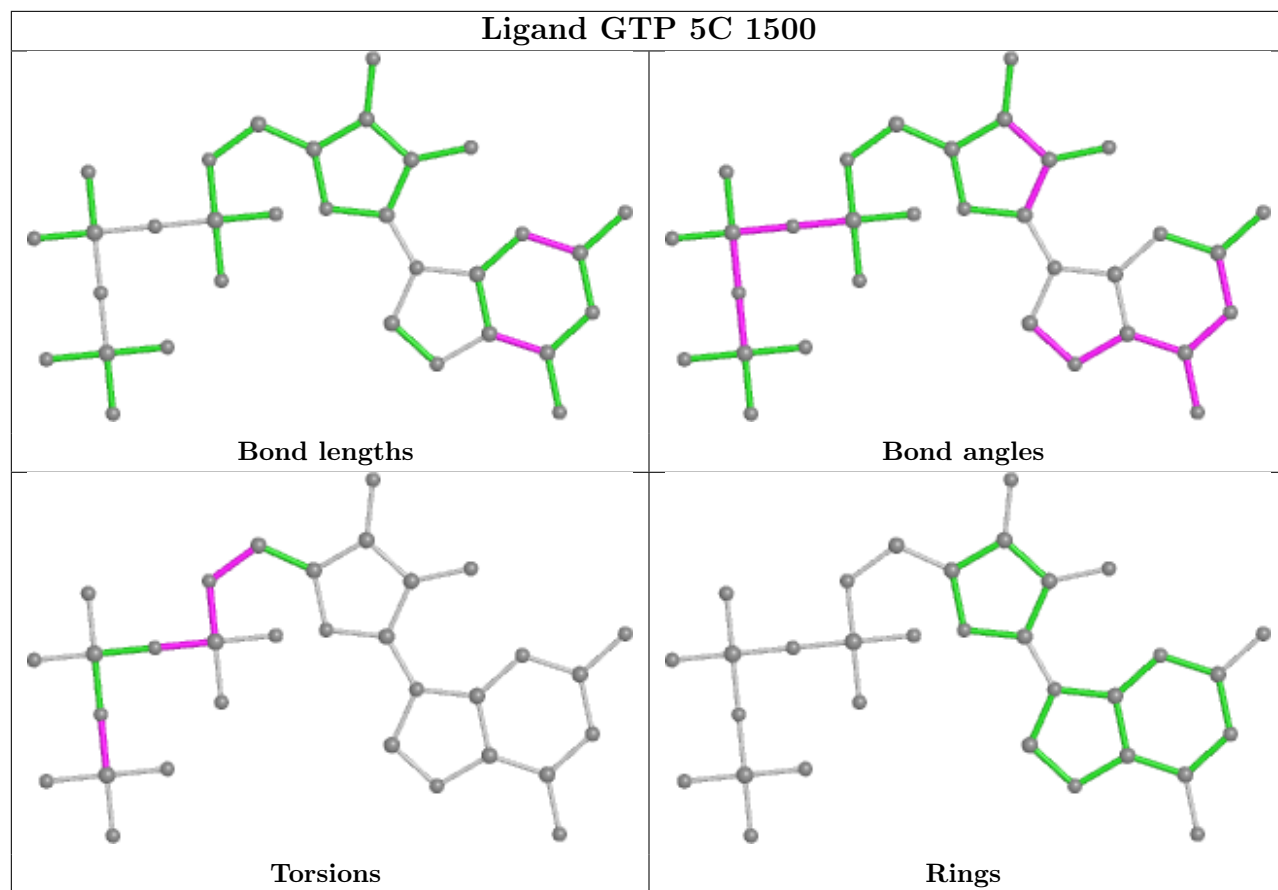
All (9) torsion outliers are listed below:

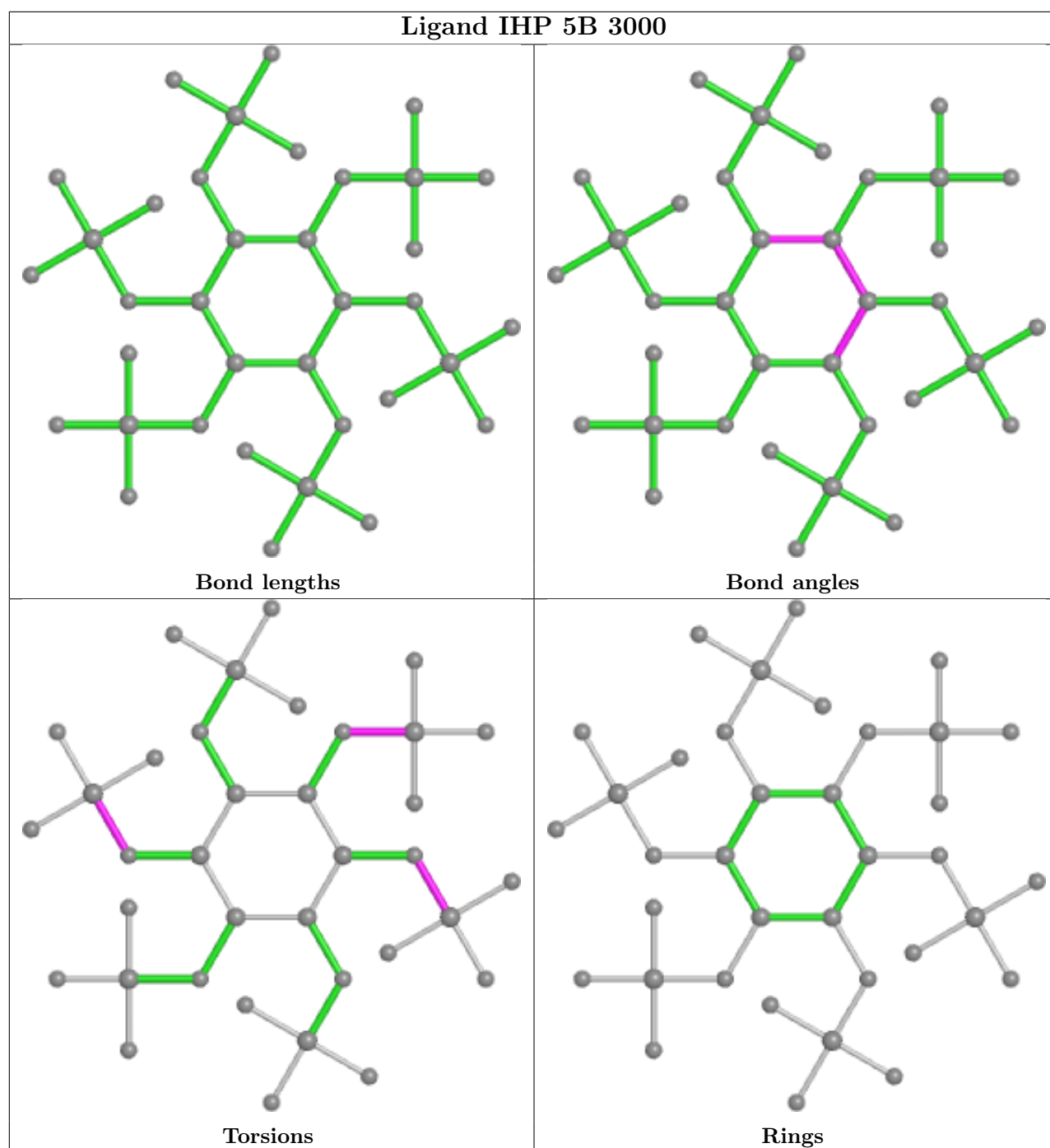
Mol	Chain	Res	Type	Atoms
46	5B	3000	IHP	C2-O12-P2-O42
46	5B	3000	IHP	C4-O14-P4-O44
47	5C	1500	GTP	PB-O3B-PG-O3G
47	5C	1500	GTP	C5'-O5'-PA-O1A
47	5C	1500	GTP	PB-O3A-PA-O5'
47	5C	1500	GTP	C5'-O5'-PA-O3A
47	5C	1500	GTP	C5'-O5'-PA-O2A
46	5B	3000	IHP	C5-O15-P5-O35
47	5C	1500	GTP	C4'-C5'-O5'-PA

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 than 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.