Full wwPDB EM Validation Report (i)

Mar 14, 2026 – 01:07 AM UTC

PDB ID	:	9MU2 / pdb_00009mu2
EMDB ID	:	EMD-48617
Title	:	SaPI1 neck structure with DNA, tail completion protein, and tape measure protein
Authors	:	Kizziah, J.L.; Dokland, T.
Deposited on	:	2025-01-13
Resolution	:	3.54 Å(reported)

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

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

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

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

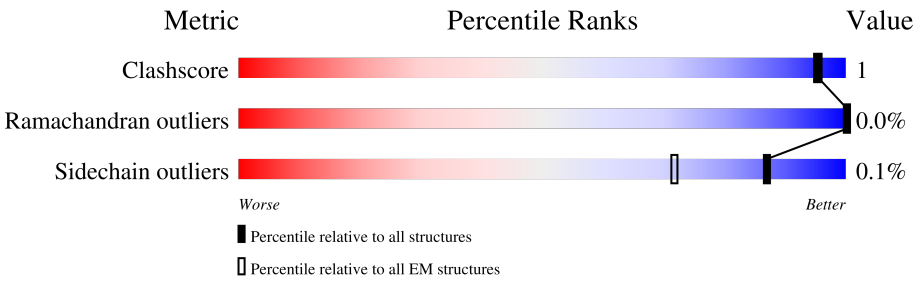
EMDB validation analysis	:	0.0.1.dev132
MolProbity	:	4-5-2 with Phenix2.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
EM percentile statistics	:	202505.v01 (Using data in the EMDB archive up until May 2025)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

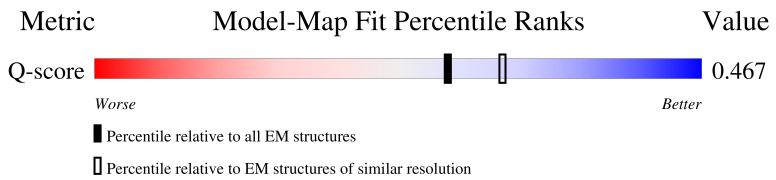
1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.54 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)	Similar EM resolution (#Entries, resolution range(Å))
Clashscore	229148	23984	-
Ramachandran outliers	224038	23583	-
Sidechain outliers	223484	23102	-
Q-score	-	25397	12891 (3.04 - 4.04)

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% The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	110	87% 13%
1	B	110	85% 13%
1	G	110	85% 13%
1	H	110	85% 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	S	110	83% 5% 13%
1	T	110	86% 13%
1	Y	110	85% 13%
1	Z	110	85% 13%
1	e	110	85% 13%
1	f	110	84% 13%
1	k	110	84% 13%
1	l	110	85% 13%
2	C	100	98%
2	I	100	97%
2	U	100	100%
2	a	100	100%
2	g	100	98%
2	m	100	99%
3	D	127	94% 6%
3	J	127	95% 5%
3	V	127	97%
3	b	127	94% 6%
3	h	127	96%
3	n	127	97%
4	E	193	5% 82% 16%
4	F	193	6% 94%
4	K	193	78% 6% 16%
4	L	193	6% 94%
4	W	193	80% 16%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	X	193	6% 94%
4	c	193	6% 81% 16%
4	d	193	6% 94%
4	i	193	7% 82% 16%
4	j	193	6% 94%
4	o	193	7% 80% 16%
4	p	193	6% 94%
5	M	1154	97%
5	N	1154	97%
5	O	1154	97%
6	P	115	7% 93% 5%
7	Q	43864	100%
8	R	43864	100%

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 31396 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 protein called Head-tail connector protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	96	780	489	128	159	4	0	0
1	B	96	780	489	128	159	4	0	0
1	G	96	780	489	128	159	4	0	0
1	H	96	780	489	128	159	4	0	0
1	S	96	780	489	128	159	4	0	0
1	T	96	780	489	128	159	4	0	0
1	Y	96	780	489	128	159	4	0	0
1	Z	96	780	489	128	159	4	0	0
1	e	96	780	489	128	159	4	0	0
1	f	96	780	489	128	159	4	0	0
1	k	96	780	489	128	159	4	0	0
1	l	96	780	489	128	159	4	0	0

- Molecule 2 is a protein called adaptor.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	C	100	832	527	145	158	2	0	0
2	I	100	832	527	145	158	2	0	0
2	U	100	832	527	145	158	2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	a	100	832	527	145	158	2	0	0
			Total	C	N	O	S		
2	g	100	832	527	145	158	2	0	0
			Total	C	N	O	S		
2	m	100	832	527	145	158	2	0	0

- Molecule 3 is a protein called DUF3168 domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	D	127	1037	656	163	214	4	0	0
			Total	C	N	O	S		
3	J	127	1037	656	163	214	4	0	0
			Total	C	N	O	S		
3	V	127	1037	656	163	214	4	0	0
			Total	C	N	O	S		
3	b	127	1037	656	163	214	4	0	0
			Total	C	N	O	S		
3	h	127	1037	656	163	214	4	0	0
			Total	C	N	O	S		
3	n	127	1037	656	163	214	4	0	0

- Molecule 4 is a protein called Major tail protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	E	162	1288	809	214	260	5	0	0
			Total	C	N	O			
4	F	12	105	66	18	21		0	0
			Total	C	N	O	S		
4	K	162	1288	809	214	260	5	0	0
			Total	C	N	O			
4	L	12	105	66	18	21		0	0
			Total	C	N	O	S		
4	W	162	1288	809	214	260	5	0	0
			Total	C	N	O			
4	X	12	105	66	18	21		0	0
			Total	C	N	O	S		
4	c	162	1288	809	214	260	5	0	0
			Total	C	N	O			
4	d	12	105	66	18	21		0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	i	162	1288	809	214	260	5	0	0
			Total	C	N	O			
4	j	12	105	66	18	21		0	0
			Total	C	N	O	S		
4	o	162	1288	809	214	260	5	0	0
			Total	C	N	O			
4	p	12	105	66	18	21		0	0

- Molecule 5 is a protein called Tape measure protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	M	38	306	193	56	56	1	0	0
			Total	C	N	O	S		
5	N	37	299	189	55	54	1	0	0
			Total	C	N	O	S		
5	O	37	299	189	55	54	1	0	0

- Molecule 6 is a protein called HK97 gp10 family phage protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	P	109	904	576	151	173	4	0	0

- Molecule 7 is a DNA chain called DNA, 5'-3'.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
7	Q	16	324	159	45	104	16	0	0

- Molecule 8 is a DNA chain called DNA, 3'-5'.

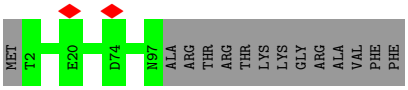
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
8	R	16	332	159	69	88	16	0	0

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 and atom inclusion in map density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). 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.

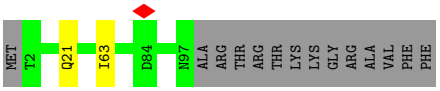
• Molecule 1: Head-tail connector protein

Chain A:



• Molecule 1: Head-tail connector protein

Chain B:



• Molecule 1: Head-tail connector protein

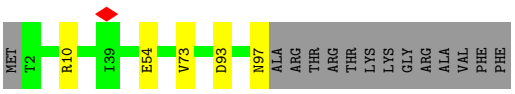
Chain G:

• Molecule 1: Head-tail connector protein

Chain H:

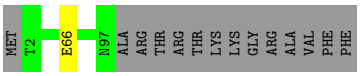
• Molecule 1: Head-tail connector protein

Chain S:



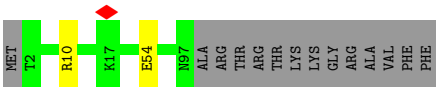
• Molecule 1: Head-tail connector protein

Chain T: 86% 13%



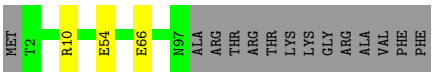
• Molecule 1: Head-tail connector protein

Chain Y: 85% 13%



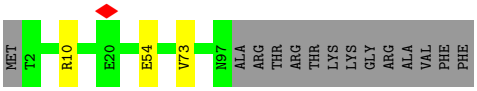
• Molecule 1: Head-tail connector protein

Chain Z: 85% 13%



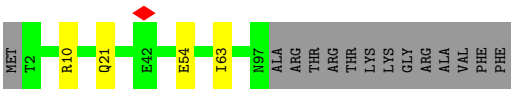
• Molecule 1: Head-tail connector protein

Chain e: 85% 13%



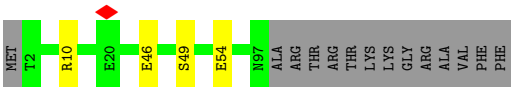
• Molecule 1: Head-tail connector protein

Chain f: 84% 13%



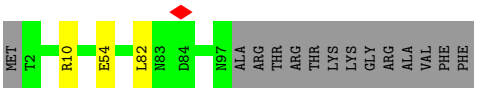
• Molecule 1: Head-tail connector protein

Chain k: 84% 13%



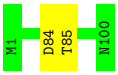
• Molecule 1: Head-tail connector protein

Chain l: 85% 13%



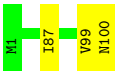
• Molecule 2: adaptor

Chain C: 98%



• Molecule 2: adaptor

Chain I: 97%



• Molecule 2: adaptor

Chain U: 100%

There are no outlier residues recorded for this chain.

• Molecule 2: adaptor

Chain a: 100%

There are no outlier residues recorded for this chain.

• Molecule 2: adaptor

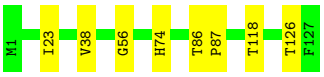
Chain g: 98%

• Molecule 2: adaptor

Chain m: 99%

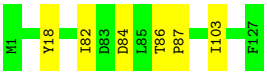
• Molecule 3: DUF3168 domain-containing protein

Chain D: 94% 6%



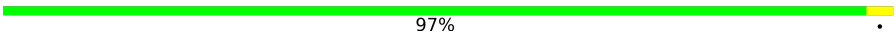
• Molecule 3: DUF3168 domain-containing protein

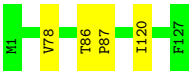
Chain J: 95% 5%



• Molecule 3: DUF3168 domain-containing protein

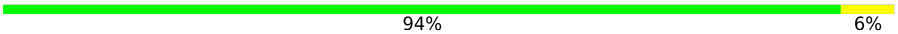
Chain V:





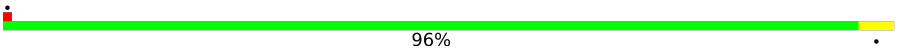
• Molecule 3: DUF3168 domain-containing protein

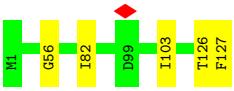
Chain b:



• Molecule 3: DUF3168 domain-containing protein

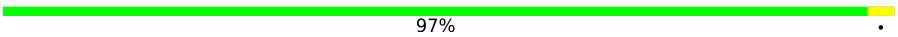
Chain h:





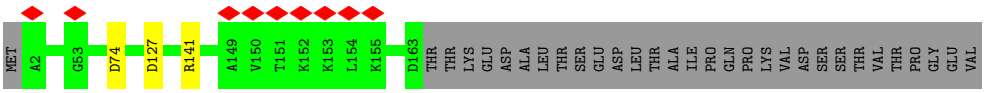
• Molecule 3: DUF3168 domain-containing protein

Chain n:



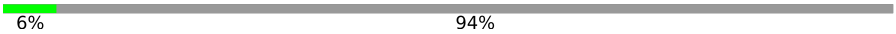
• Molecule 4: Major tail protein

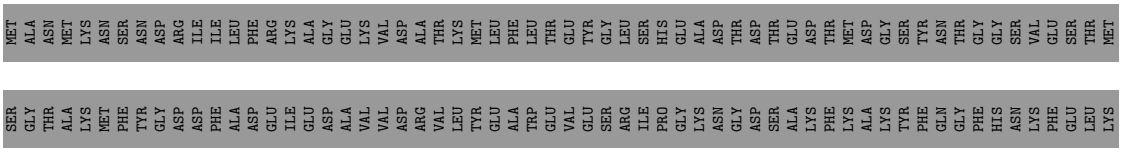
Chain E:

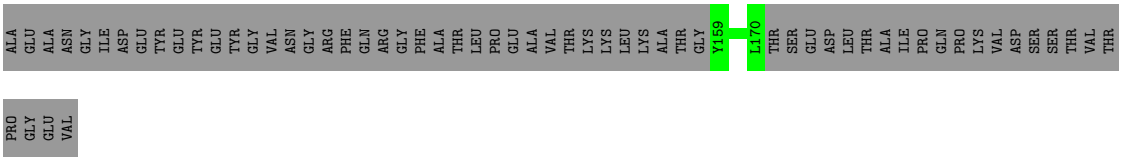


• Molecule 4: Major tail protein

Chain F:

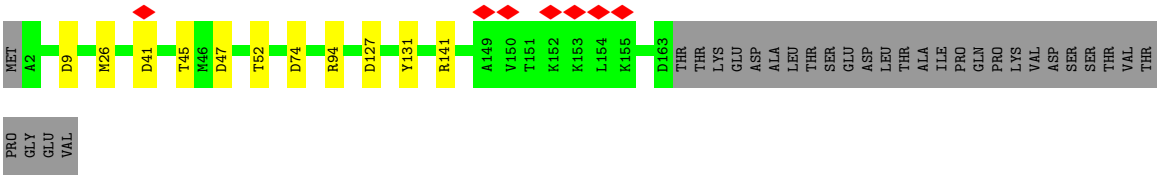






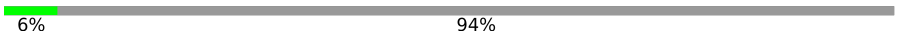
• Molecule 4: Major tail protein

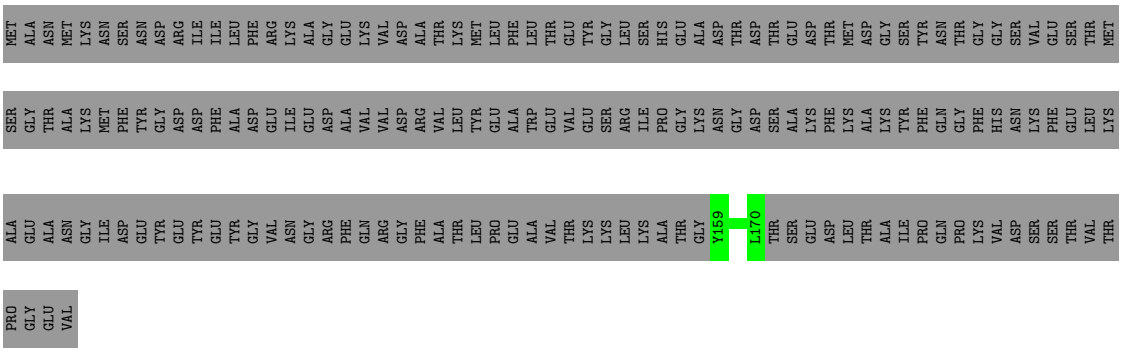
Chain K:



• Molecule 4: Major tail protein

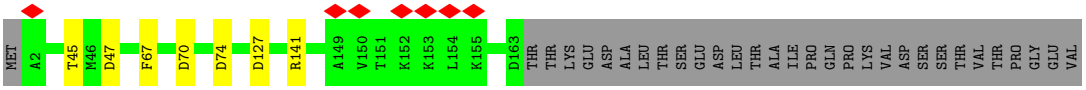
Chain L:





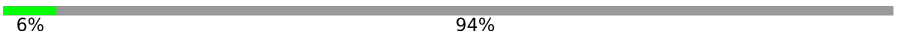
• Molecule 4: Major tail protein

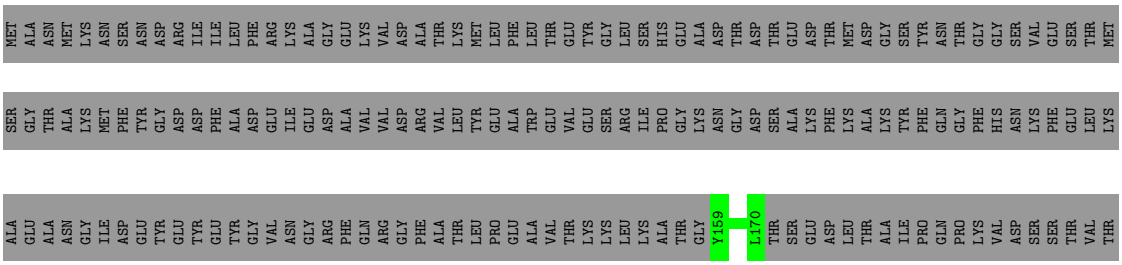
Chain W:



• Molecule 4: Major tail protein

Chain X:





PRO GLY GLU VAL

• Molecule 4: Major tail protein

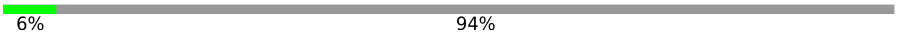
Chain c:

MET A2 D9 D39 T40 D41 D44 N51 T52 G53 D74 R94 D127 R141 A149 V150 T151 K152 K153 L154 K155 D163 THR THR LYS GLU ASP ALA LEU THR SER GLU ASP LEU THR ALA ILE PRO GLN PRO LYS VAL ASP SER SER THR VAL THR PRO GLY

GLU VAL

• Molecule 4: Major tail protein

Chain d:



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

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

ALA GLU ALA ASN GLY ILE ASP GLU TYR GLU TYR GLU TYR GLY VAL ASN GLY ARG PHE GLN ARG GLY PHE ALA THR LEU PRO GLU ALA VAL THR LYS LYS LEU LYS ALA THR GLY Y159 L170 THR SER GLU ASP LEU THR ALA ILE PRO GLN PRO LYS VAL ASP SER SER THR VAL THR

PRO GLY GLU VAL

• Molecule 4: Major tail protein

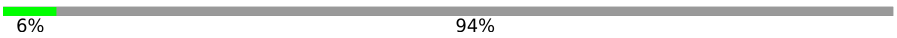
Chain i:

MET A2 D39 D44 D47 G48 S49 Y50 N51 T52 G53 D74 I126 D127 R141 A149 V150 T151 K152 K153 L154 K155 D163 THR THR LYS GLU ASP ALA LEU THR SER GLU ASP LEU THR ALA ILE PRO GLN PRO LYS VAL ASP SER SER THR VAL THR PRO GLY GLU

VAL

• Molecule 4: Major tail protein

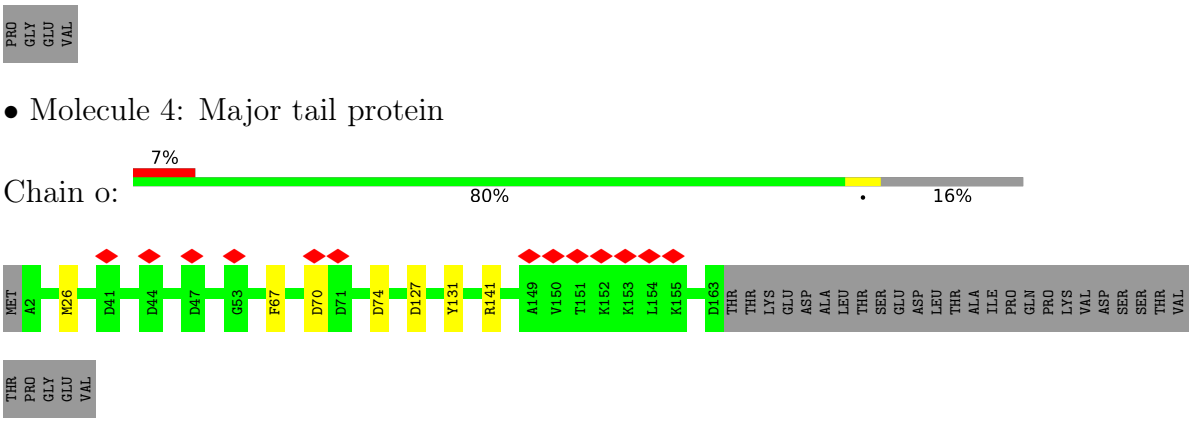
Chain j:

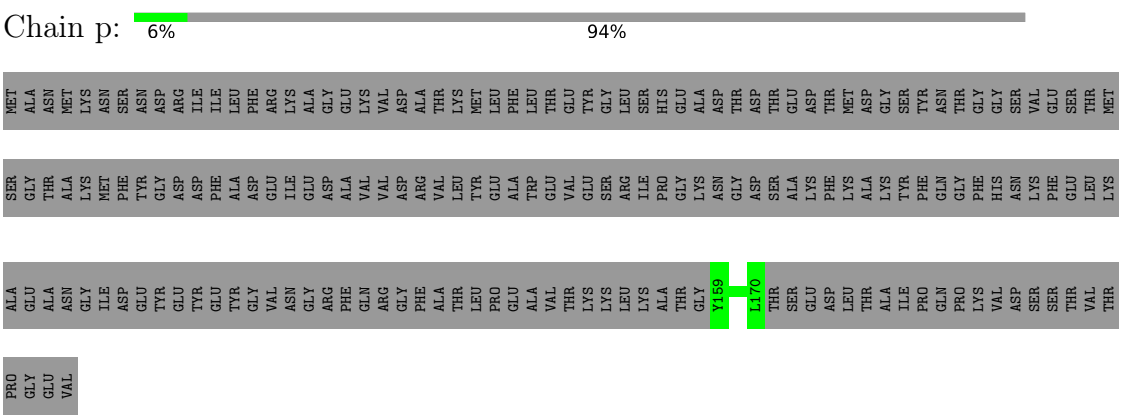


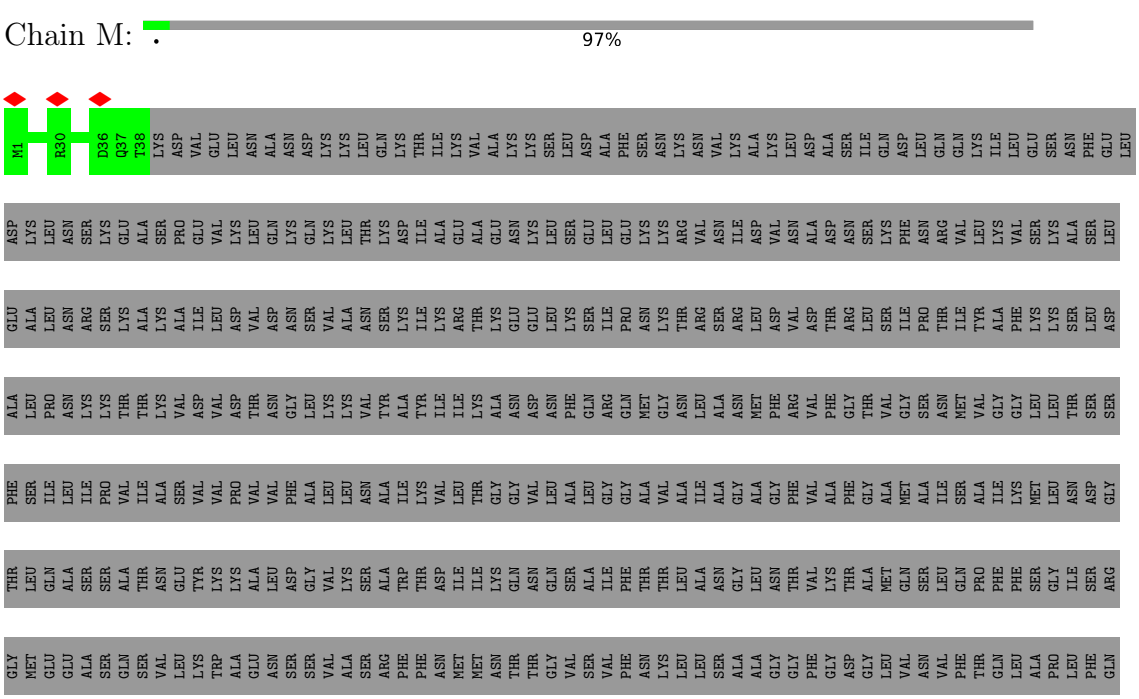
MET ALA ASN MET LYS ASN SER ASN ASP ARG ILE ILE LEU PHE ARG LYS ALA GLY GLU LYS VAL ASP ALA THR LYS MET LEU PHE LEU THR GLU TYR GLY LEU SER HIS GLU ALA ASP THR ASP THR GLU ASP THR MET ASP GLY SER TYR ASN THR GLY GLY SER VAL GLU SER THR MET

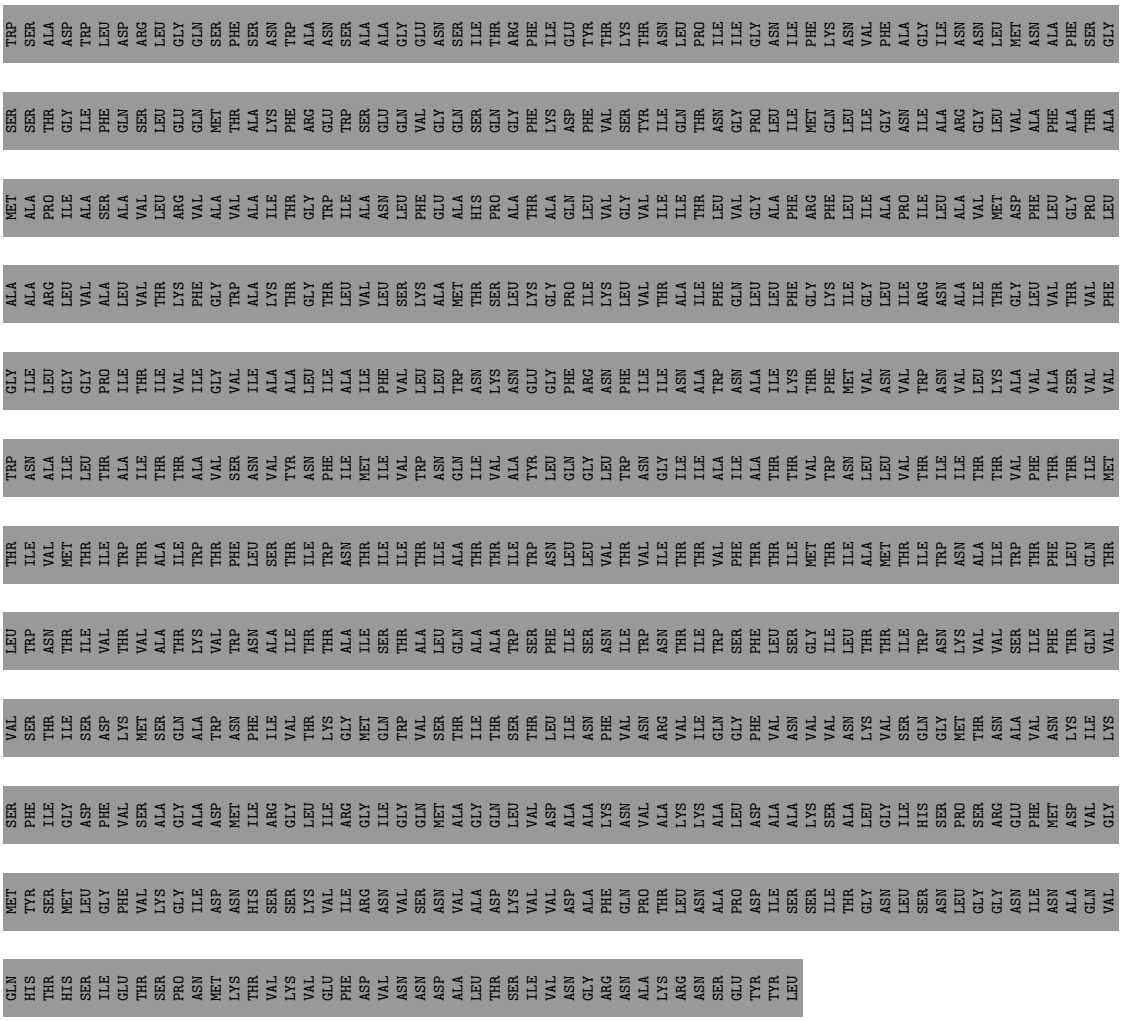
SER GLY THR ALA LYS MET PHE TYR GLY ASP ASP PHE ALA ASP GLU ILE GLU ASP ALA VAL VAL ASP ARG VAL LEU TYR GLU ALA TRP GLU VAL GLU SER ARG ILE PRO GLY LYS ASN GLY ASP SER ALA LYS PHE LYS ALA LYS TYR PHE GLN GLY PHE HIS ASN LYS PHE GLU LEU LYS

ALA GLU ALA ASN GLY ILE ASP GLU TYR GLU TYR GLU TYR GLY VAL ASN GLY ARG PHE GLN ARG GLY PHE ALA THR LEU PRO GLU ALA VAL THR LYS LYS LEU LYS ALA THR GLY Y159 L170 THR SER GLU ASP LEU THR ALA ILE PRO GLN PRO LYS VAL ASP SER SER THR VAL THR


PRO GLY GLU VAL
• Molecule 4: Major tail protein
Chain o:
7%
80%
16%
MET A2 M26 D41 D44 D47 G53 F67 D70 D71 D74 D127 Y131 R141 A149 V150 T151 K152 K153 L154 K155 D163 THR THR LYS GLU ASP ALA LEU THR SER GLU ASP LEU THR ALA ILE PRO GLN PRO LYS VAL ASP SER SER THR VAL
THR PRO GLY GLU VAL

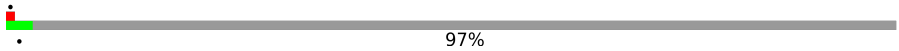

• Molecule 4: Major tail protein
Chain p:
6%
94%
MET ALA ASN MET LYS ASN SER ASN ASP ARG ILE ILE LEU PHE ARG LYS ALA GLY GLU LYS VAL ASP ALA THR LYS MET LEU PHE LEU THR GLU TYR GLY LEU SER HIS GLU ALA ASP THR ASP THR GLU ASP THR MET ASP GLY SER TYR ASN THR GLY GLY SER VAL GLU SER THR MET
SER GLY THR ALA LYS MET PHE TYR GLY ASP ASP PHE ALA ASP GLU ILE GLU ASP ALA VAL VAL ASP ARG VAL LEU TYR GLU ALA TRP GLU VAL GLU SER ARG ILE PRO GLY LYS ASN GLY ASP SER ALA LYS PHE LYS ALA LYS TYR PHE GLN GLY PHE HIS ASN LYS PHE GLU LEU LYS
ALA GLU ALA ASN GLY ILE ASP GLU TYR GLU TYR GLU TYR GLY VAL ASN GLY ARG PHE GLN ARG GLY PHE ALA THR LEU PRO GLU ALA VAL THR LYS LYS LEU LYS ALA THR GLY Y159 L170 THR SER GLU ASP LEU THR ALA ILE PRO GLN PRO LYS VAL ASP SER SER THR VAL THR
PRO GLY GLU VAL

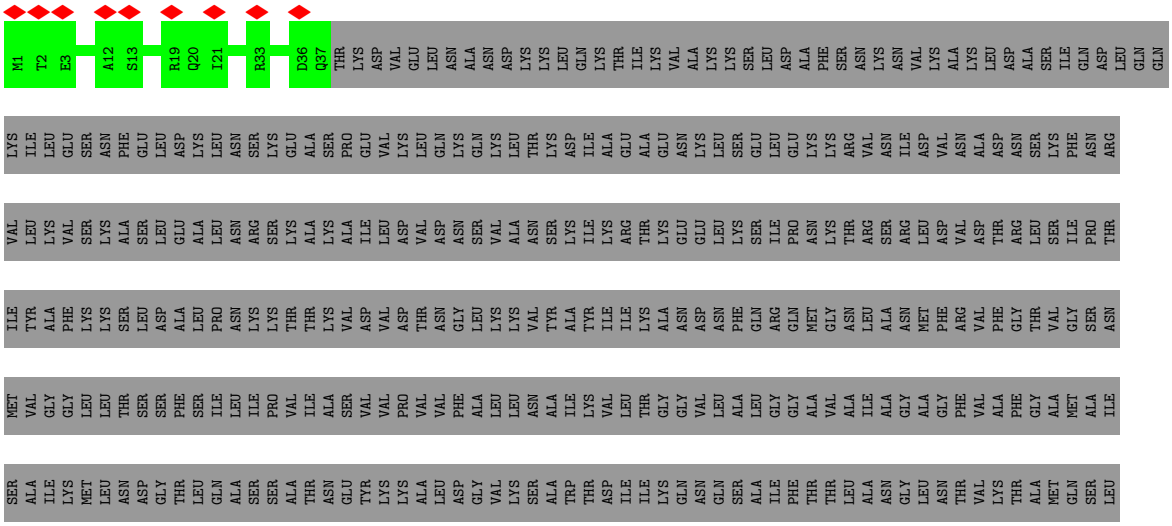

• Molecule 5: Tape measure protein
Chain M:
97%
M1 R30 D36 Q37 T38 LYS ASP VAL GLU LEU ASN ALA ASN ASP LYS LYS LEU GLN LYS THR ILE LYS VAL ALA LYS LYS SER LEU ASP ALA PHE SER ASN LYS ASN VAL LYS ALA LYS LEU ASP ALA SER ILE GLN ASP LEU GLN GLN LYS ILE LEU GLU SER ASN PHE GLU LEU
ASP LYS LEU ASN SER LYS GLU ALA SER PRO GLU VAL LYS LEU GLN LYS GLN LYS LEU THR LYS ASP ILE ALA GLU ALA GLU ASN LYS LEU SER GLU LEU GLU LYS LYS ARG VAL ASN ILE ASP VAL ASN ALA ASP ASN SER LYS PHE ASN ARG VAL LEU LYS VAL SER LYS ALA SER LEU
GLU ALA LEU ASN ARG SER LYS ALA LYS ALA ILE LEU ASP VAL ASP ASN SER VAL ALA ASN SER LYS ILE ILE LYS ARG THR LYS GLU GLU LEU LYS SER ILE PRO ASN LYS THR ARG SER ARG LEU ASP VAL ASP THR ARG LEU SER ILE PRO THR ILE TYR ALA PHE LYS LYS LEU LEU ASP
ALA LEU PRO ASN LYS LYS THR THR LYS VAL ASP VAL ASP THR ASN GLY LEU LYS LYS VAL TYR ALA TYR ILE ILE LYS ALA ASN ASP ASN VAL ASN PHE GLN ARG GLN MET GLY ASN LEU ALA ASN MET PHE ARG VAL PHE GLY THR VAL GLY SER ASN MET VAL GLY GLY LEU LEU THR SER SER
PHE SER ILE LEU ILE PRO VAL ILE ALA SER VAL VAL PRO VAL VAL PHE ALA LEU LEU ASN ALA ILE LYS VAL LEU THR GLY GLY VAL LEU ALA LEU GLY GLY ALA VAL ALA ALA ILE ALA GLY ALA GLY PHE VAL ALA PHE GLY ALA MET ALA ILE SER ALA ILE LYS MET LEU ASN ASP GLY
THR LEU GLN ALA SER SER ALA THR ASN GLU TYR LYS LYS ALA LEU ASP GLY VAL LYS SER ALA TRP THR THR ASP ILE ILE ILE LYS GLN ASN GLN SER ALA ILE PHE THR THR LEU ALA ASN GLY LEU ASN THR VAL LYS THR ALA MET GLN SER LEU GLN PRO PHE PHE SER GLY ILE SER ARG
GLY MET GLU GLU ALA SER GLN SER VAL LEU LYS TRP ALA GLU ASN SER SER VAL ALA SER ARG PHE PHE ASN MET MET MET ASN THR THR GLY VAL SER VAL PHE ASN LYS LEU LEU SER ALA ALA GLY GLY PHE GLY ASP GLY LEU VAL ASN VAL PHE THR GLN LEU ALA PRO LEU PHE GLN



- Molecule 5: Tape measure protein

Chain N:





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

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

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

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

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

ALA ILE THR GLY LEU VAL THR VAL PHE GLY ILE LEU GLY PRO ILE THR ILE VAL ILE GLY VAL ILE ALA ALA LEU ILE ALA ILE PHE VAL LEU LEU TRP ASN LYS ASN GLU GLY PHE ARG ASN PHE ILE ILE ASN ALA TRP ASN ALA ILE LYS THR PHE MET VAL ASN VAL TRP ASN

VAL LEU LYS ALA VAL ALA SER VAL VAL TRP ASN ALA ILE LEU THR ALA ILE THR THR ALA VAL SER ASN VAL TYR ASN PHE ILE MET ILE VAL TRP ASN GLN ILE VAL ALA TYR LEU GLN GLY LEU TRP ASN GLY ILE ILE ALA ILE ALA THR THR VAL TRP ASN LEU LEU VAL THR ILE

ILE THR THR VAL PHE THR THR ILE MET THR ILE VAL MET THR ILE TRP THR ALA ILE TRP THR PHE LEU SER THR ILE TRP ASN THR ILE ILE THR ILE ALA THR THR ILE TRP ASN LEU LEU VAL THR VAL ILE THR THR VAL PHE THR THR ILE MET THR ILE ALA MET THR ILE TRP

ASN ALA ILE TRP THR PHE LEU GLN THR LEU TRP ASN THR ILE VAL THR VAL ALA THR LYS VAL TRP ASN ALA ILE THR THR ALA ILE SER THR ALA LEU GLN ALA ALA TRP SER PHE ILE SER ASN ILE TRP ASN THR ILE TRP SER PHE LEU SER GLY ILE LEU THR THR ILE TRP ASN

LYS VAL VAL SER ILE PHE THR GLN VAL VAL SER THR ILE SER ASP LYS MET SER GLN ALA TRP ASN PHE ILE VAL THR LYS GLY MET GLN TRP VAL SER THR ILE THR SER THR LEU ILE ASN PHE VAL ASN ARG VAL ILE GLN GLY PHE VAL ASN VAL VAL ASN LYS VAL SER GLN GLY

MET THR ASN ALA VAL ASN LYS ILE LYS SER PHE ILE GLY ASP PHE VAL SER ALA GLY ALA ASP MET ILE ARG GLY LEU ILE ARG GLY ILE GLY GLN MET ALA GLY GLN LEU VAL ASP ALA ALA LYS ASN VAL ALA LYS LYS ALA LEU ASP ALA ALA LYS SER ALA LEU GLY ILE HIS SER

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

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

- Molecule 5: Tape measure protein

Chain O: 97%

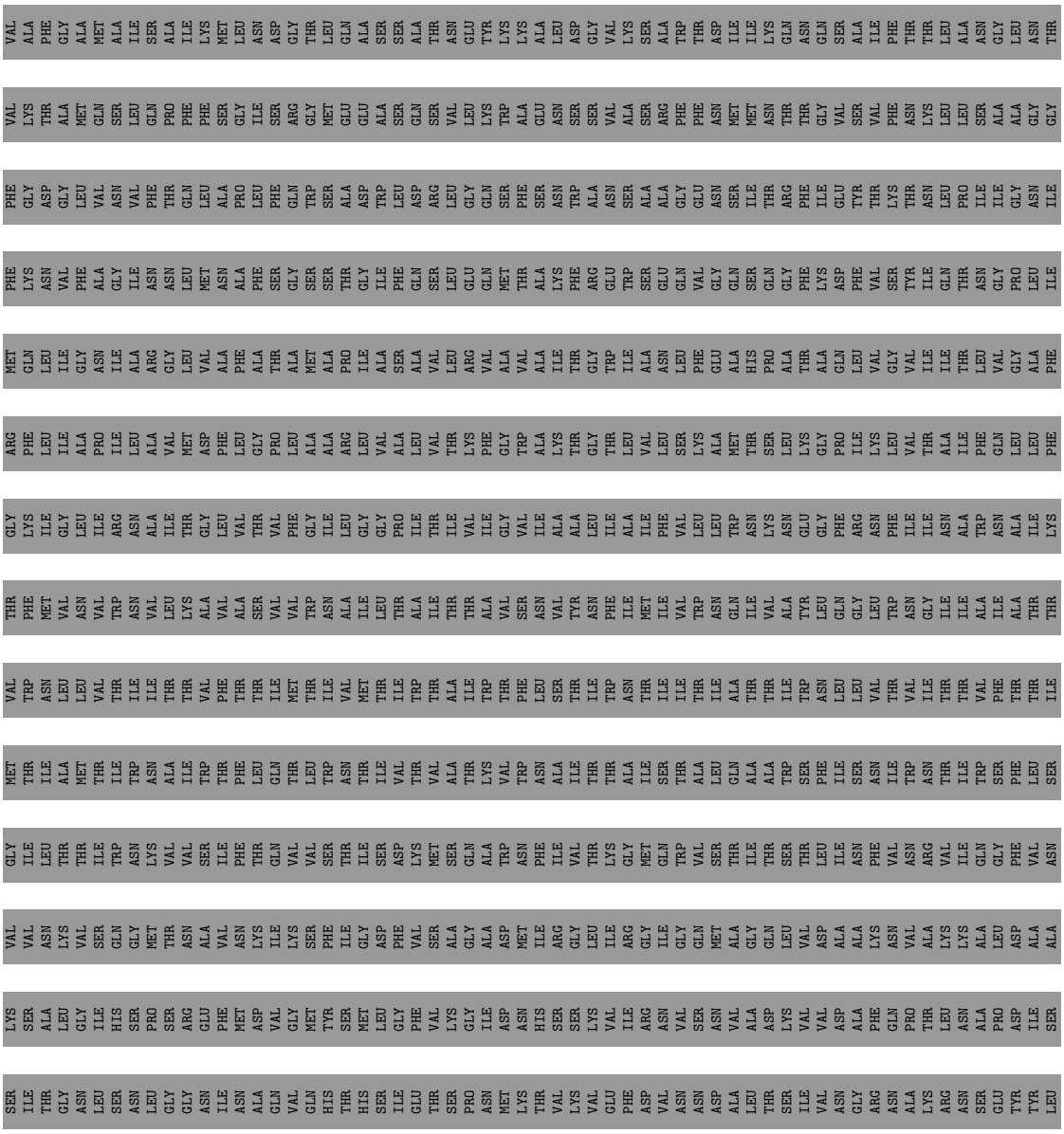
M1 T2 K5 A8 E11 A12 S13 V14 A15 R19 D22 R30 F31 K32 R33 Q37 THR LYS ASP VAL GLU LEU ASN ALA ASN ASP LYS LYS LEU GLN LYS THR ILE LYS VAL ALA LYS LYS SER LEU ASP ALA PHE SER ASN LYS ASN VAL LYS ALA LYS LEU ASP

ALA SER ILE GLN ASP LEU GLN GLN LYS ILE LEU GLU SER ASN PHE GLU LEU ASP LYS LEU ASN SER LYS GLU ALA SER PRO GLU VAL LYS LEU GLN LYS GLN LYS LEU THR LYS ASP ILE ALA GLU ALA GLU ASN LYS LEU SER GLU LEU GLU LYS LYS ARG VAL ASN ILE ASP VAL ASN

ALA ASP ASN SER LYS PHE ASN ARG VAL LEU LYS VAL SER LYS ALA SER LEU GLU ALA LEU ASN ARG SER LYS ALA LYS ALA ILE LEU ASP VAL ASP ASN SER VAL ALA ASN SER LYS ILE LYS ARG THR LYS GLU LEU LYS SER ILE PRO ASN LYS THR ARG SER ARG LEU ASP VAL

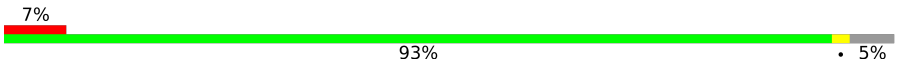
ASP THR ARG LEU SER ILE PRO THR ILE TYR ALA PHE LYS LYS SER LEU ASP ALA LEU PRO ASN LYS LYS THR THR LYS VAL ASP VAL ASP THR ASN GLY LEU LYS LYS VAL TYR ALA TYR ILE ILE LYS ALA ASN ASP ASN PHE GLN ARG GLN MET GLY ASN LEU ALA ASN MET PHE ARG

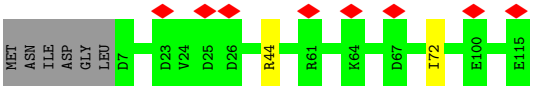
VAL PHE GLY THR VAL GLY SER ASN MET VAL GLY GLY LEU LEU THR SER SER PHE SER ILE LEU ILE PRO VAL ILE ALA SER VAL VAL PRO VAL VAL PHE ALA LEU LEU ASN ALA ILE LYS VAL LEU THR GLY GLY VAL LEU ALA LEU GLY GLY ALA VAL ALA ILE ALA GLY ALA GLY PHE



- Molecule 6: HK97 gp10 family phage protein

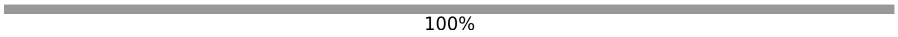
Chain P:

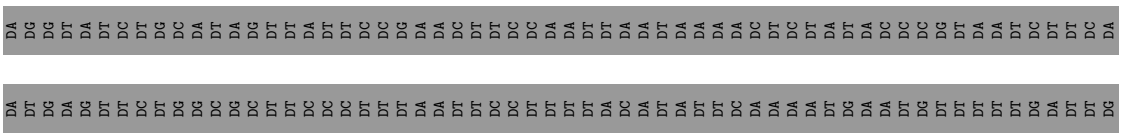




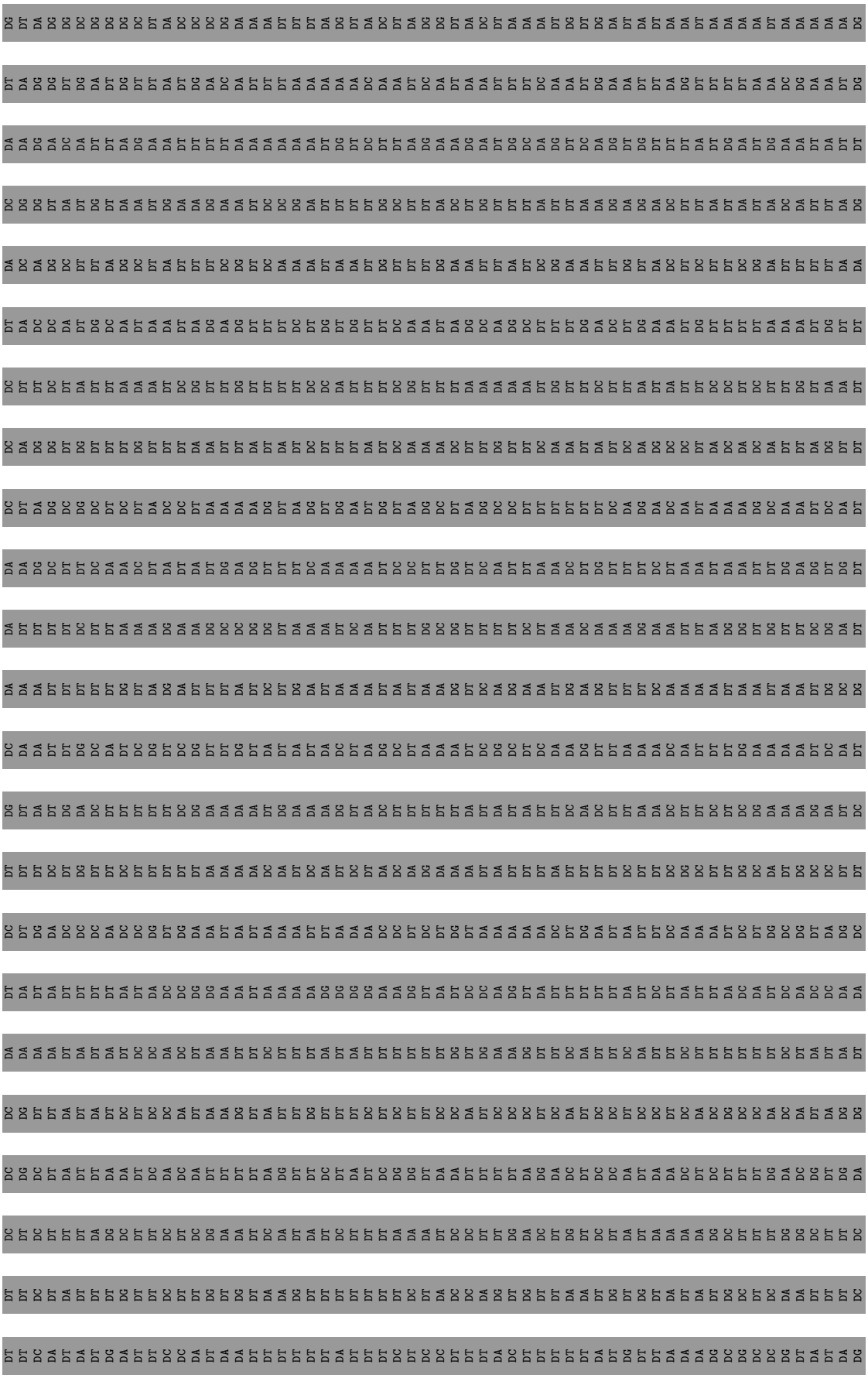
- Molecule 7: DNA, 5'-3'

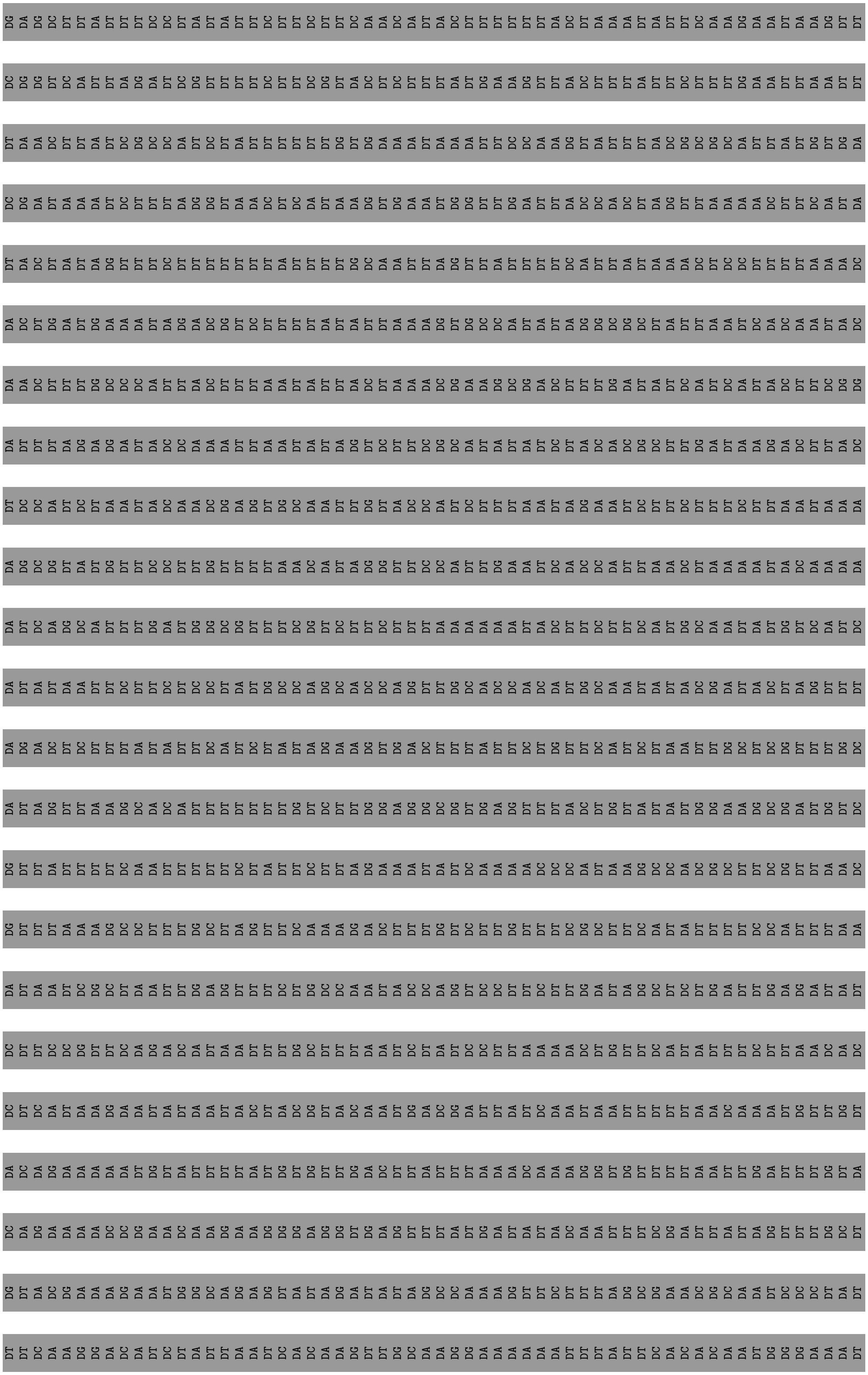
Chain Q:

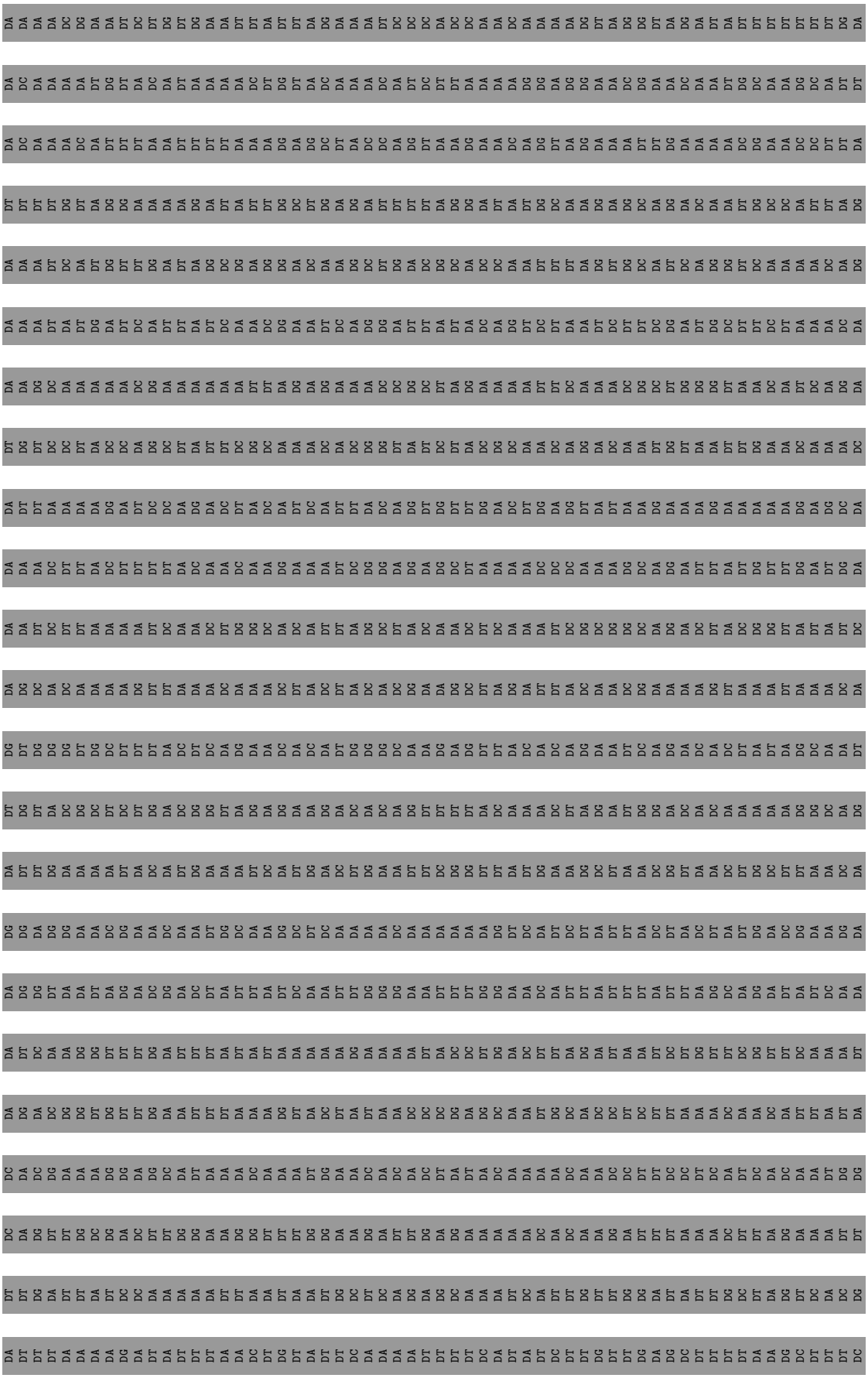


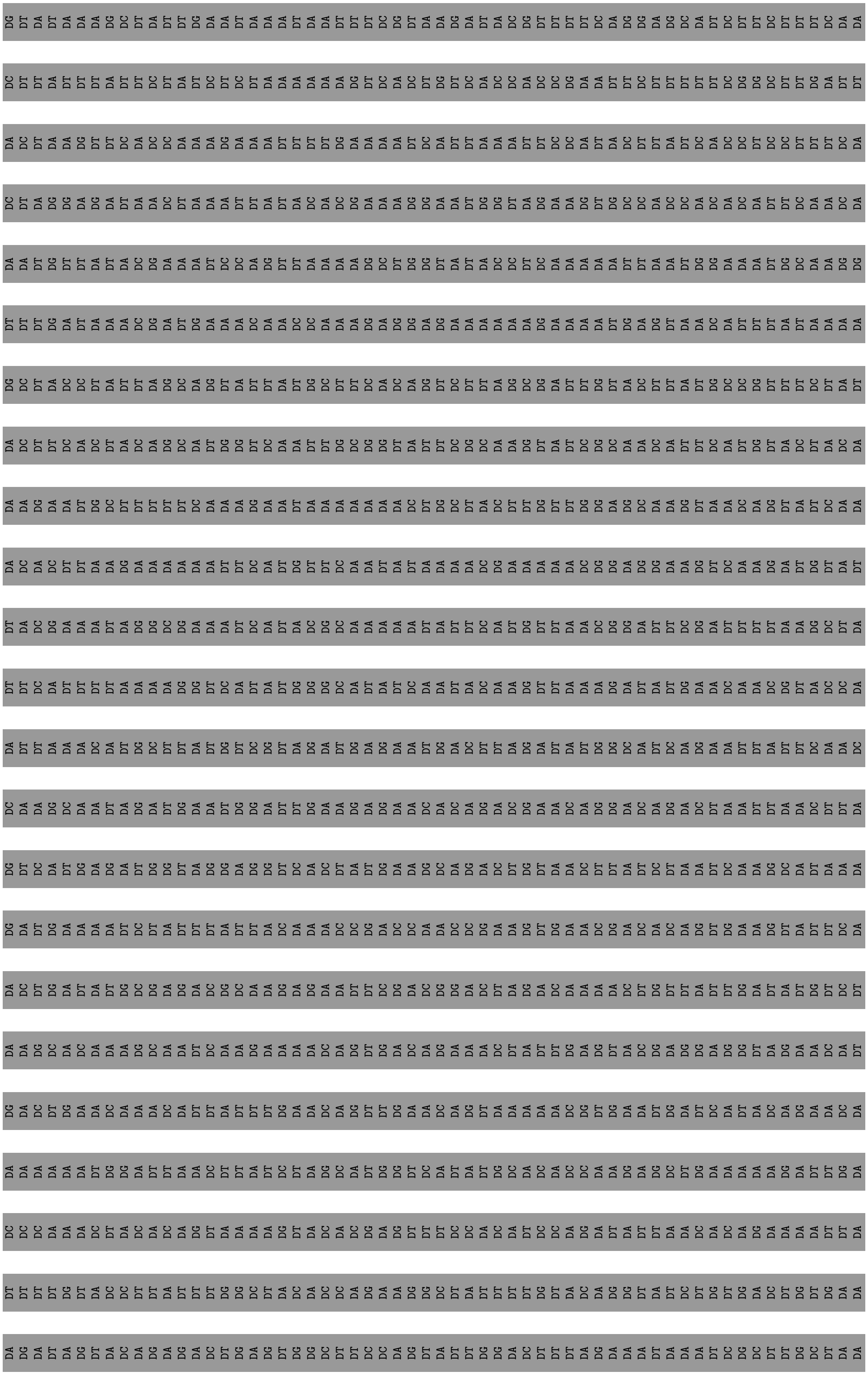


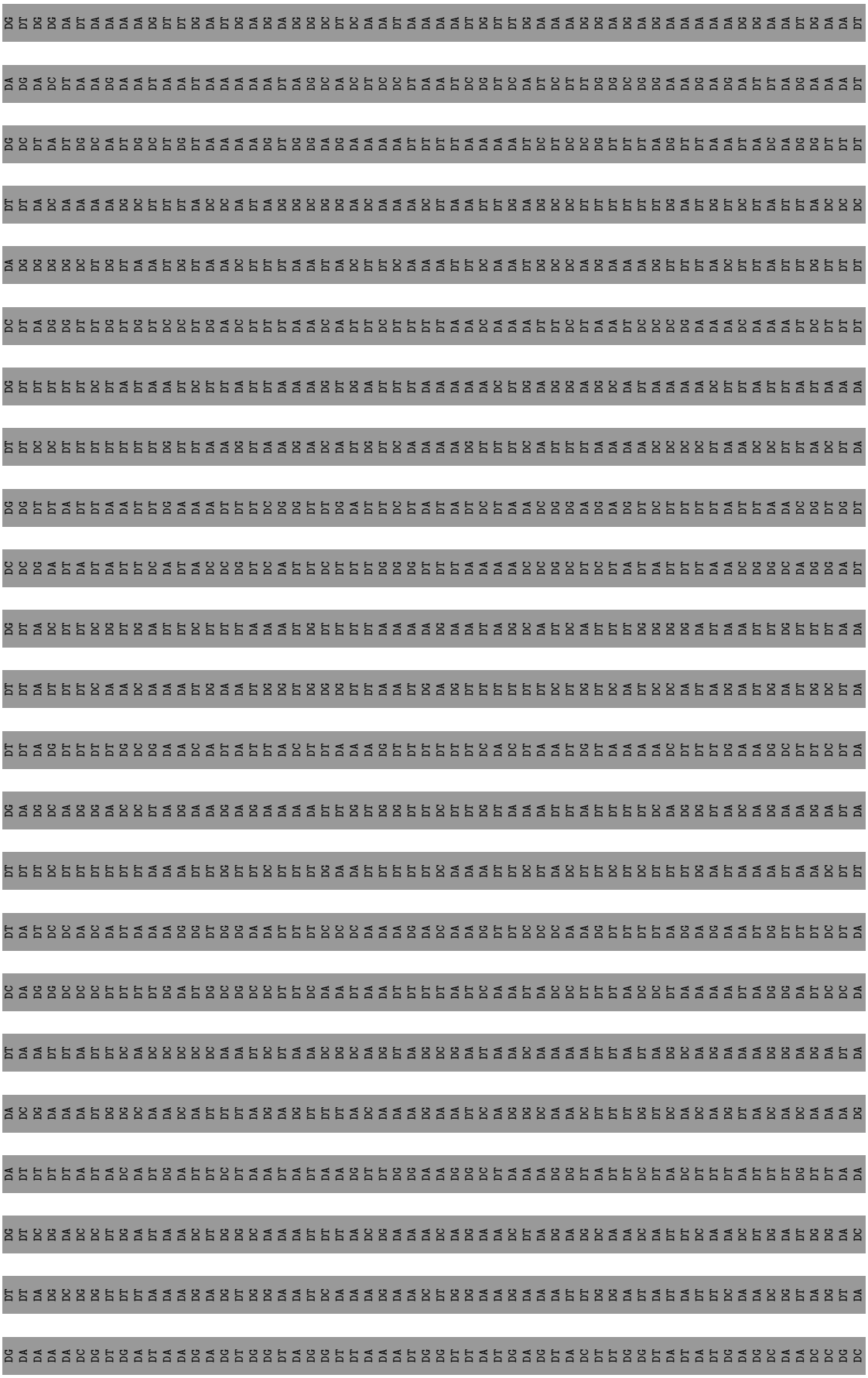
A1
T16

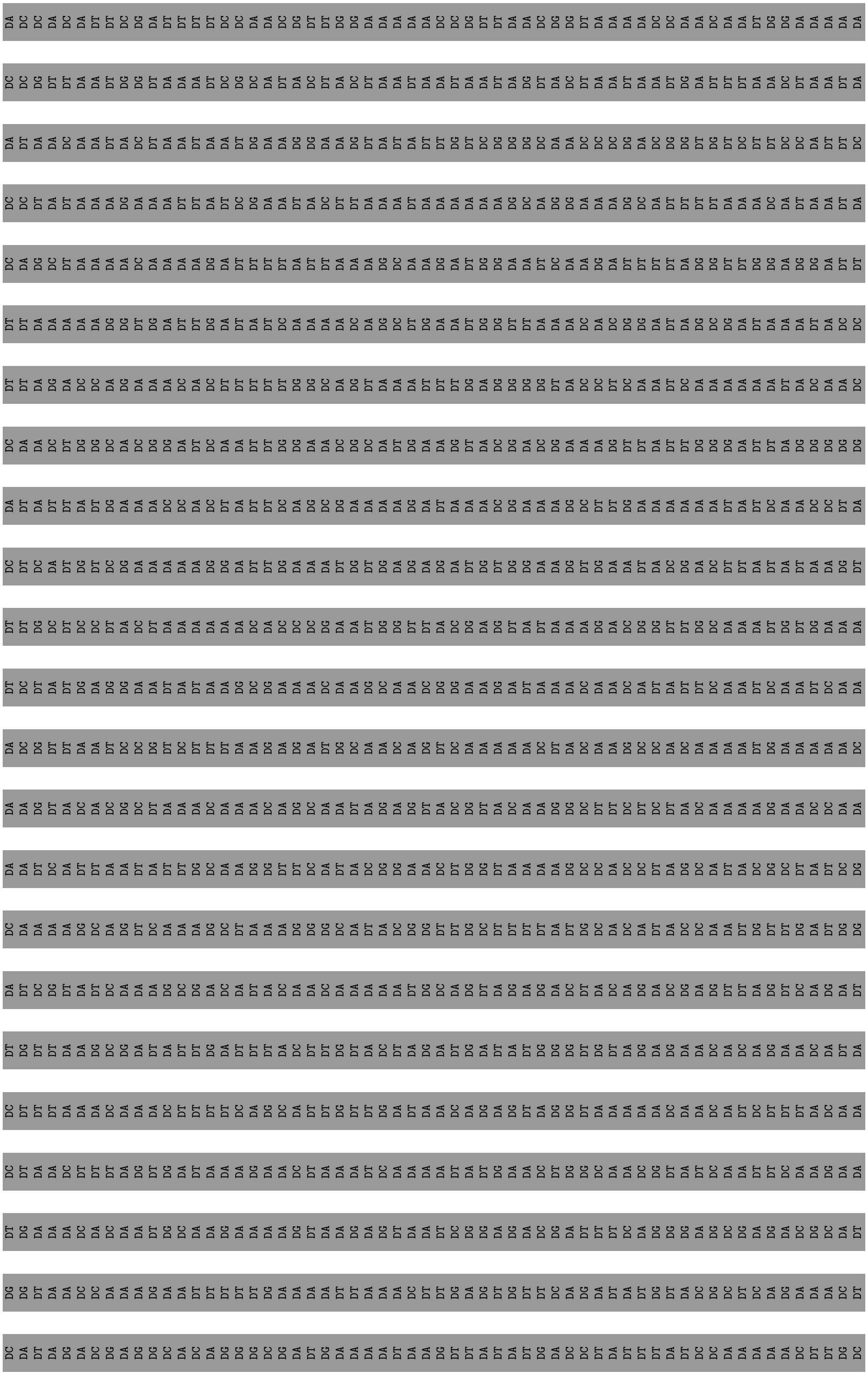

DG DT DA DG DG DC DG DG DG DC DT DA DC DC DC DG DA DA DA DT DT DT DA DG DT DA DC DT DA DG DG DT DA DC DT DA DA DT DG DT DG DA DT DA DA DA DT DA DA DA DT DA DA DA DT DG
DT DA DG DG DT DG DA DT DG DT DT DA DT DG DA DC DA DT DT DT DA DA DA DA DA DC DA DA DT DT DA DA DT DA DA DT DT DA DT DA DT DT DA DT DT DT DT DT DG DA DA DT DT DA DC DG DA DA DT DG
DA DA DG DA DC DA DT DT DA DG DA DA DT DT DT DT DA DA DA DA DA DA DA DT DT DG DT DC DT DT DA DG DA DA DG DA DT DG DC DA DA DG DT DT DG DT DT DT DA DT DT DG DA DA DT DA DT DT
DC DG DG DT DA DT DG DT DA DA DT DG DA DA DG DA DA DA DT DC DC DG DA DT DT DT DT DT DT DT DA DC DT DG DT DT DT DA DA DT DT DA DC DA DT DT DT DT DT DG DA DA DT DT DA DC DA DG
DA DC DA DG DC DT DT DA DG DC DT DA DT DT DT DC DG DT DC DA DA DA DT DT DA DA DT DT DG DT DT DT DG DA DA DT DT DT DA DT DC DG DA DA DT DT DG DT DA DC DT DC DT DT DT DC DG DA DT DT DT DA
DT DA DC DC DA DT DG DC DA DT DA DA DT DA DG DA DG DT DT DT DC DT DG DT DG DT DT DC DA DA DT DA DG DC DA DG DC DT DT DT DG DA DA DT DG DT DT DT DT DA DA DA DT DG DT DT
DC DT DT DC DT DT DA DT DT DA DA DA DT DC DG DT DT DG DT DT DT DT DC DC DA DT DT DT DC DG DT DT DT DA DA DA DA DT DT DG DT DT DA DT DT DC DC DC DT DC DT DT DG DA DA DT
DC DA DG DG DT DT DG DT DT DG DT DT DT DG DT DT DA DT DA DT DT DT DT DC DT DA DA DA DC DT DT DG DT DT DC DA DA DT DT DA DC DC DA DT DT DA DT DC DG DC DC DT DA DC DC DA DT DT DA DG DT DT
DC DT DA DG DC DC DG DC DT DC DT DA DC DC DT DA DA DA DA DG DT DT DA DT DG DT DG DT DT DA DG DC DT DA DG DC DC DT DT DT DT DT DC DA DA DG DC DA DA DT DC DA DA DG DA DA DT
DA DA DG DC DT DT DC DA DA DC DT DA DT DT DA DT DG DA DG DT DT DT DC DA DA DA DT DC DC DT DT DG DT DC DT DA DA DT DA DA DG DT DT DC DT DA DA DG DT DT DT DG DA DT DG DT DT
DA DT DT DT DT DC DT DT DA DA DG DA DG DC DC DG DG DT DT DA DA DT DT DC DA DT DC DA DA DC DT DT DC DA DA DT DT DC DC DT DT DG DT DT DT DC DT DT DG DA DA DT DC DC DG DT DT
DA DA DA DT DT DT DT DT DA DG DA DA DT DT DT DA DT DC DT DT DG DA DA DG DA DA DT DT DG DA DA DT DT DG DT DT DT DC DA DA DA DT DA DA DT DA DA DT DG DC DC DT DC DG DG
DC DA DA DT DT DG DC DA DT DC DG DT DC DG DT DT DG DT DA DT DA DA DC DT DA DA DA DT DC DG DC DT DA DA DG DC DT DT DA DA DA DT DC DA DA DT DT DG DC DC DA DA DT DC DC DA DT
DG DT DA DT DG DA DC DT DT DT DT DT DC DG DA DA DA DA DT DG DA DT DA DC DT DT DT DT DT DA DT DA DT DT DC DT DA DA DC DT DT DC DT DC DG DA DA DA DG DA DT DT DC
DT DT DT DC DT DG DT DT DC DT DT DT DT DT DA DA DA DA DC DA DT DC DT DA DC DA DG DA DA DA DT DA DT DT DT DT DC DT DT DC DC DG DC DA DT DG DC DC DT DT DT
DC DT DG DA DC DC DC DA DC DC DG DT DG DA DA DA DT DA DA DA DT DT DA DA DA DC DC DT DT DC DT DC DG DA DT DA DT DT DC DT DA DA DT DC DT DG DC DG DT DA DG
DT DA DT DA DT DT DT DT DA DT DC DC DA DA DA DA DG DG DG DG DA DA DG DT DA DT DT DT DA DT DA DT DT DC DT DT DT DT DT DC DA DA DT DT DT DC DA DC DC DC DA DA
DA DA DA DT DT DA DC DC DA DC DC DT DT DA DT DC DC DC DT DT DC DT DC DT DT DC DC DT DT DC DT DC DC DC DT DT DC DT DT DC DC DT DT DC DT DC DC DC DT DT DC DT DT DC DC DT DT
DC DT DT DC DA DT DT DT DT DT DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA
DC DG DT DT DA DA DT DC DT DC DC DA DT DA DA DG DT DA DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT
DC DG DC DT DA DT DT DA DA DT DC DA DC DA DT DT DT DT DA DA DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT
DC DT DC DT DT DT DA DT DC DT DC DC DA DT DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC DC
DT DT DC DT DA DT DT DA DT DT DG DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT
DT DT DC DA DT DT DT DT DT DC DC DA DT DA DA DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT

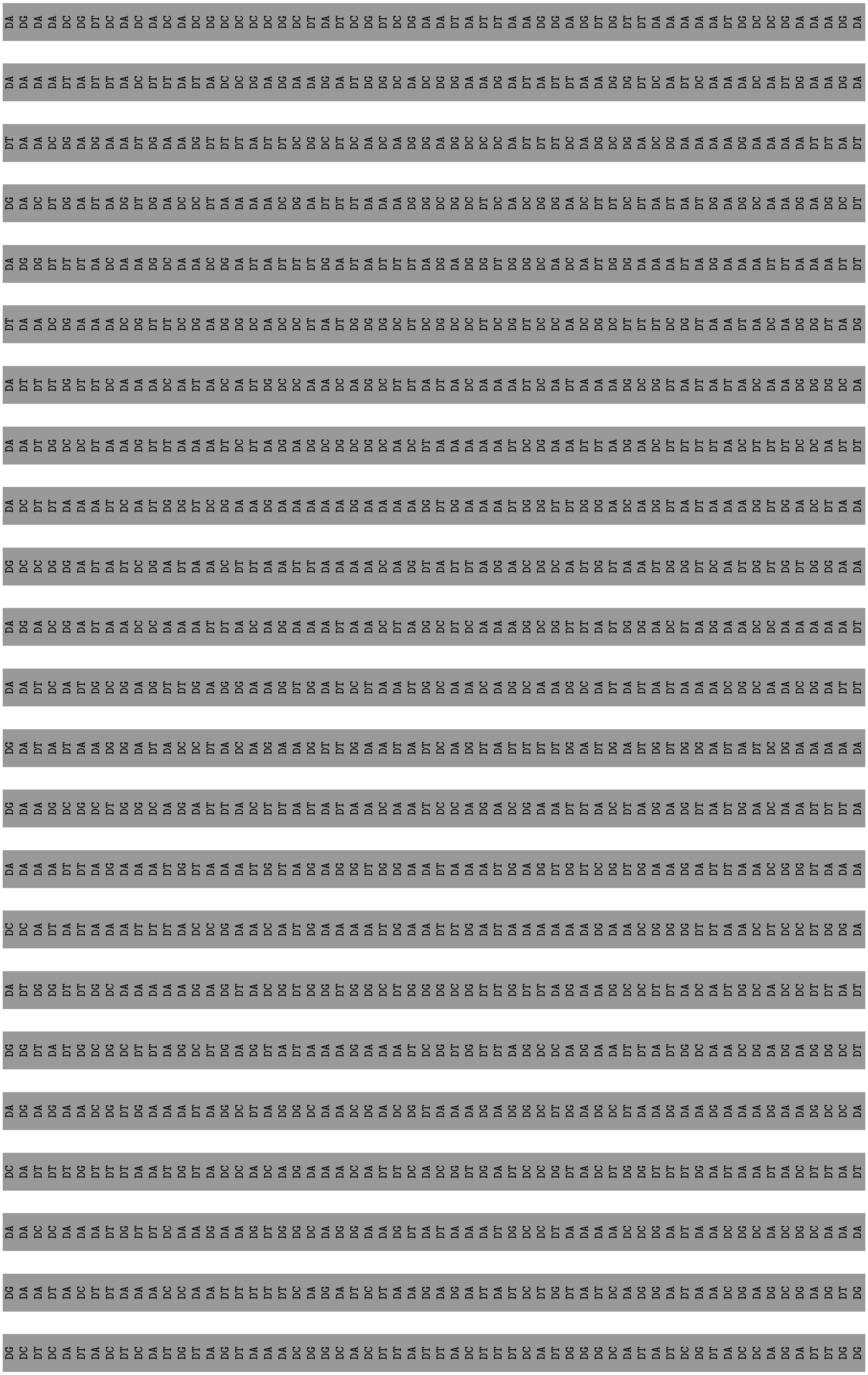












DA DT DA DT DG DA DG DG DA DG DG DA DA DT DA DG DG DA DA DA DA DT DG DA DC DT DA DA DC DA DC DA DT DT DA DC DA DA DG DT DA DA DA DC DT DA DT DT DA DT DC DA DA DA DA DA DA DT DG

DC DT DA DG DA DA DT DG DC DC DC DG DA DA DC DG DA DA DA DT DC DA DT DA DA DG DA DC DG DG DA DT DG DC DA DG DG DT DT DA DT DG DA DC DA DT DA DT DT DC DT DC DA DG DC DT DG DA DA DA

DC DT DG DT DC DG DT DA DC DT DC DG DA DA DC DC DA DC DA DA DG DA DA DA DA DA DG DC DA DG DT DG DA DT DC DA DA DA DA DC DA DG DA DT DG DT DA DG DC DT DG DT DG DA DG DT DA DT DA DC

DC DA DG DA DG DG DG DC DT DA DT DG DT DC DG DG DA DC DT DA DT DT DA DA DC DT DA DG DT DC DG DT DA DG DT DG DG DT DG DT DA DA DG DT DA DG DT DA DA DA DA DC DA DC DA DT DT DT DA DG

DT DG DA DT DT DG DA DA DA DC DA DG DG DC DA DA DG DA DT DA DG DA DC DG DC DG DG DG DA DT DA DT DC DA DT DG DG DC DA DA DT DT DT DA DG DG DG DA DT DT DA DA DT DA DT DC DA DA DG DA

DA DT DG DA DC DC DA DT DG DA DA DG DA DT DG DA DC DA DA DA DA DT DG DC DA DA DA DC DT DA DT DC DT DT DT DT DT DA DA DG DA DA DA DT DA DT DT DG DA DT DA DA DC DG DA DA DA DA DA DA

DT DT DT DT DC DG DA DA DA DA DA DG DA DA DC DG DT DC DA DT DT DT DA DT DA DT DA DA DG DC DT DA DG DG DT DA DG DT DT DA DC DC DG DT DA DT DC DG DA DA DA DA DA DG DG DA DG DA DA DC

DG DT DA DT DA DG DC DA DC DA DG DT DT DA DG DT DT DA DT DT DG DT DA DC DC DT DA DT DA DT DG DG DA DC DA DC DC DT DG DA DA DC DT DA DA DA DG DC DA DA DG DT DG DG DA DG DG DA DA DT

DT DC DG DA DA DA DG DT DG DT DT DT DC DA DG DA DA DC DG DT DG DG DA DG DA DA DA DA DG DG DC DT DT DC DG DG DA DA DG DT DA DG DC DG DG DA DG DT DG DT DA DA DA DG DA DC DA DT DA

DT DT DA DG DA DT DA DG DA DG DT DC DA DA DG DG DA DG DG DT DT DT DT DG DG DT DG DA DA DG DT DG DA DC DG DC DA DA DT DA DC DT DT DA DG DT DC DA DC DA DA DC DA DT DT DT DA DA DG

DA DT DT DC DA DA DC DA DG DG DA DC DG DT DA DA DG DC DA DT DA DC DA DC DA DC DA DT DA DA DC DT DA DA DA DA DC DT DA DA DG DA DG DC DA DA DT DC DA DA DA DG DG DT DT DT DA DC DA DG

DT DT DG DT DT DG DA DG DG DC DA DG DA DG DA DG DT DA DA DA DG DA DA DG DA DA DG DC DG DA DA DA DG DA DG DA DA DG DT DA DC DA DA DG DG DC DG DC DA DA DG DT DT DA DA DA DA DG DA DG

DA DT DG DC DA DG DT DT DA DT DT DA DA DA DG DT DG DG DG DT DC DA DG DT DT DG DT DA DT DG DA DA DA DA DT DA DT DA DA DG DG DG DA DG DT DG DT DG DG DG DA DA DA DT DG DA DC DG DG DA

DG DG DT DT DA DG DA DA DT DT DA DA DA DA DC DT DA DT DT DT DC DA DG DA DT DA DG DA DG DT DT DT DA DT DT DA DC DA DC DA DA DC DA DA DC DA DG DA DT DC DT DA DG DC DT DT DC DT DG DG

DT DG DA DT DT DA DT DA DT DT DA DA DA DC DT DT DG DT DT DA DT DG DA DA DG DT DT DA DG DG DG DA DT DT DG DA DG DT DA DT DT DT DT DC DT DT DC DC DG DG DT DC DA DA DA DG DA DT DG DT

DG DT DT DC DA DA DC DA DA DT DG DA DA DG DT DA DT DG DG DG DT DT DA DA DA DA DG DA DG DA DT DG DA DG DA DT DT DG DA DA DT DC DA DT DT DT DA DC DA DT DT DT DA DT DT DG DA DG DG DA

DG DG DC DA DG DA DC DG DA DT DG DA DT DT DA DA DC DA DT DA DC DC DT DA DA DA DA DT DG DA DA DA DT DT DC DT DC DG DA DA DA DA DA DG DT DA DC DA DC DT DG DA DA DA DT DA DA DT DC DA

DA DA DA DA DA DT DA DT DA DA DA DA DA DT DA DA DA DA DC DA DC DC DT DG DA DA DG DA DA DA DA DG DG DC DT DA DA DG DA DT DT DG DA DA DG DA DT DG DA DT DT DT DT DA DT DT DA DA DA DG

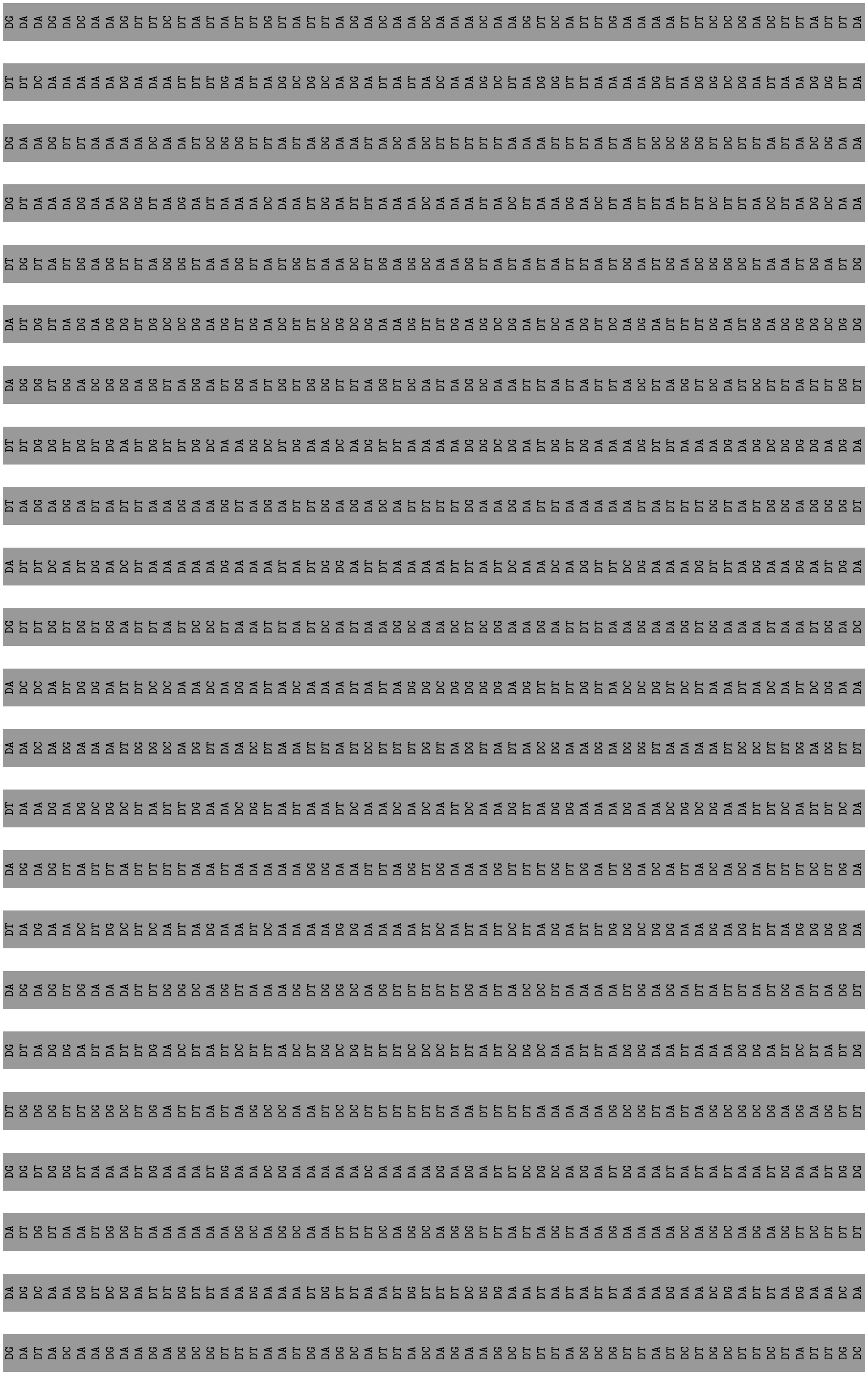
DA DA DA DT DT DA DA DA DG DA DT DA DA DA DG DA DC DA DG DT DG DA DA DT DT DT DT DA DC DA DG DT DC DC DT DA DC DG DA DT DG DG DC DT DA DA DT DA DT DG DA DA DT DG DA DA DT DA DT DG

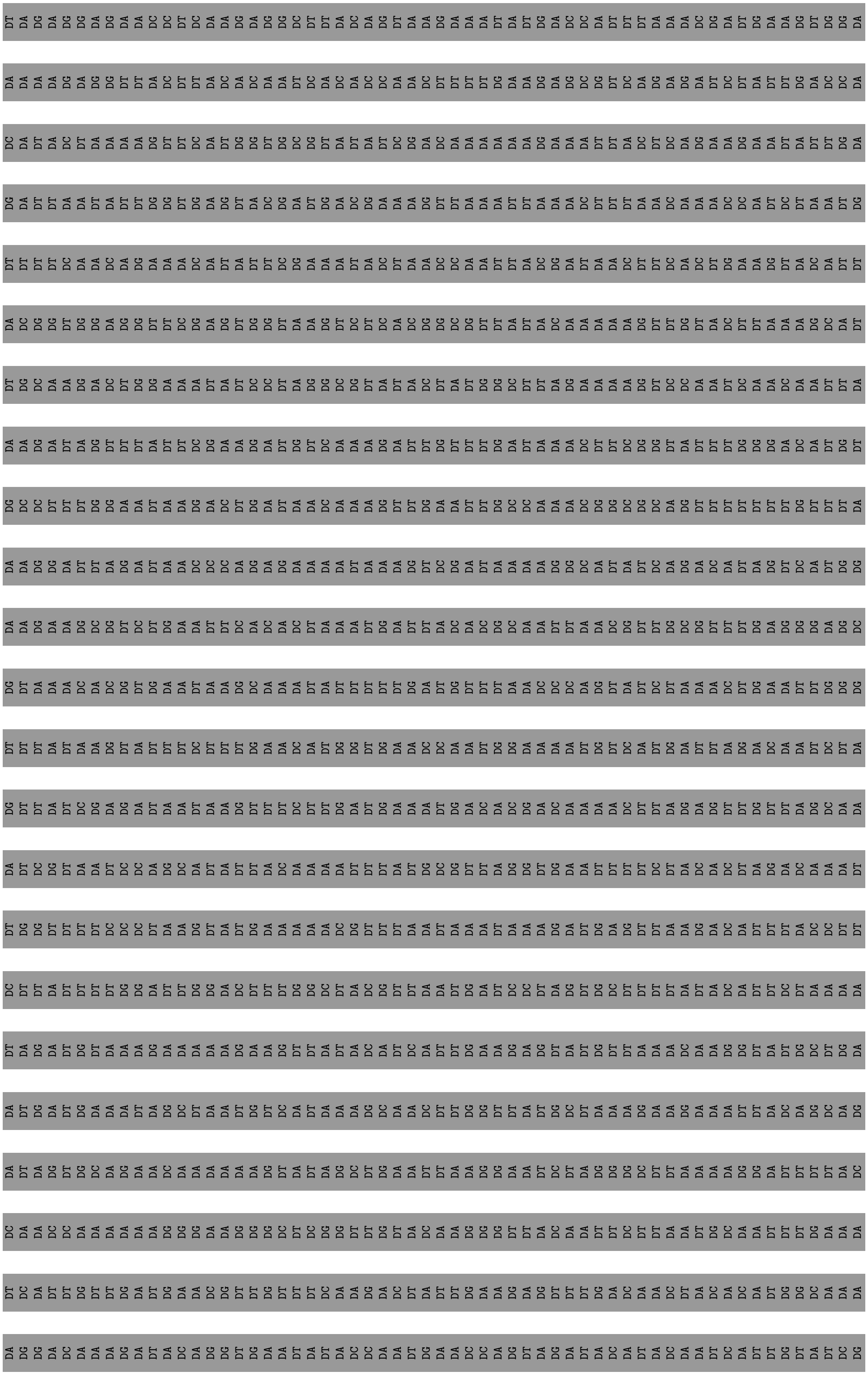
DA DA DT DT DA DA DG DG DG DC DT DA DT DG DT DT DA DA DG DA DA DT DG DA DT DG DC DC DT DA DG DT DT DT DA DA DT DT DG DA DT DA DC DT DG DG DA DG DA DT DG DA DC DA DA DT DG DA DT DG

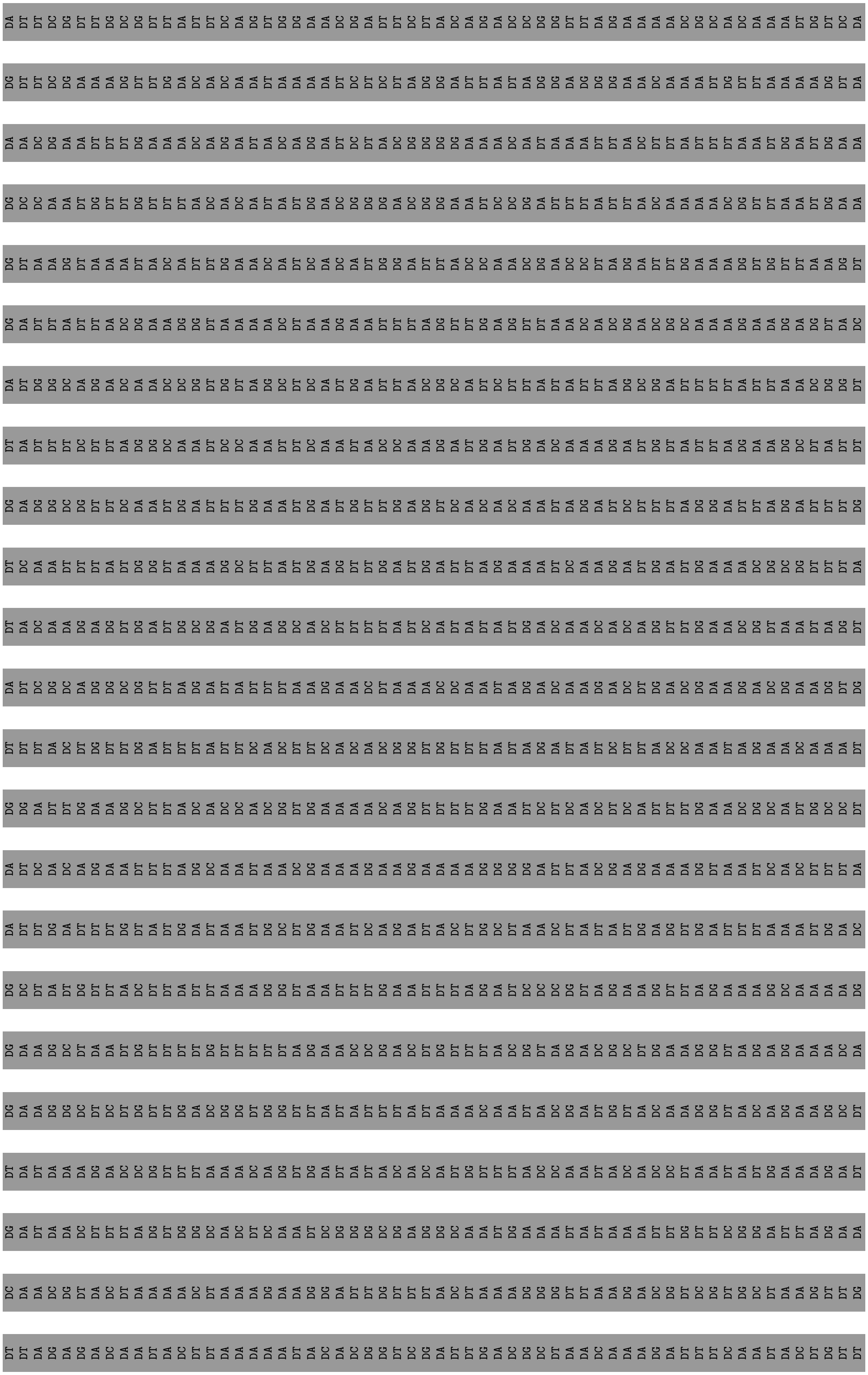
DA DT DT DA DA DA DA DA DA DC DT DT DA DA DA DA DA DT DA DT DG DG DA DT DT DG DG DT DT DC DG DA DT DA DT DC DT DT DT DA DT DT DG DC DT DG DG DA DA DT DA DC DT DG DC DG DA DT DT DA

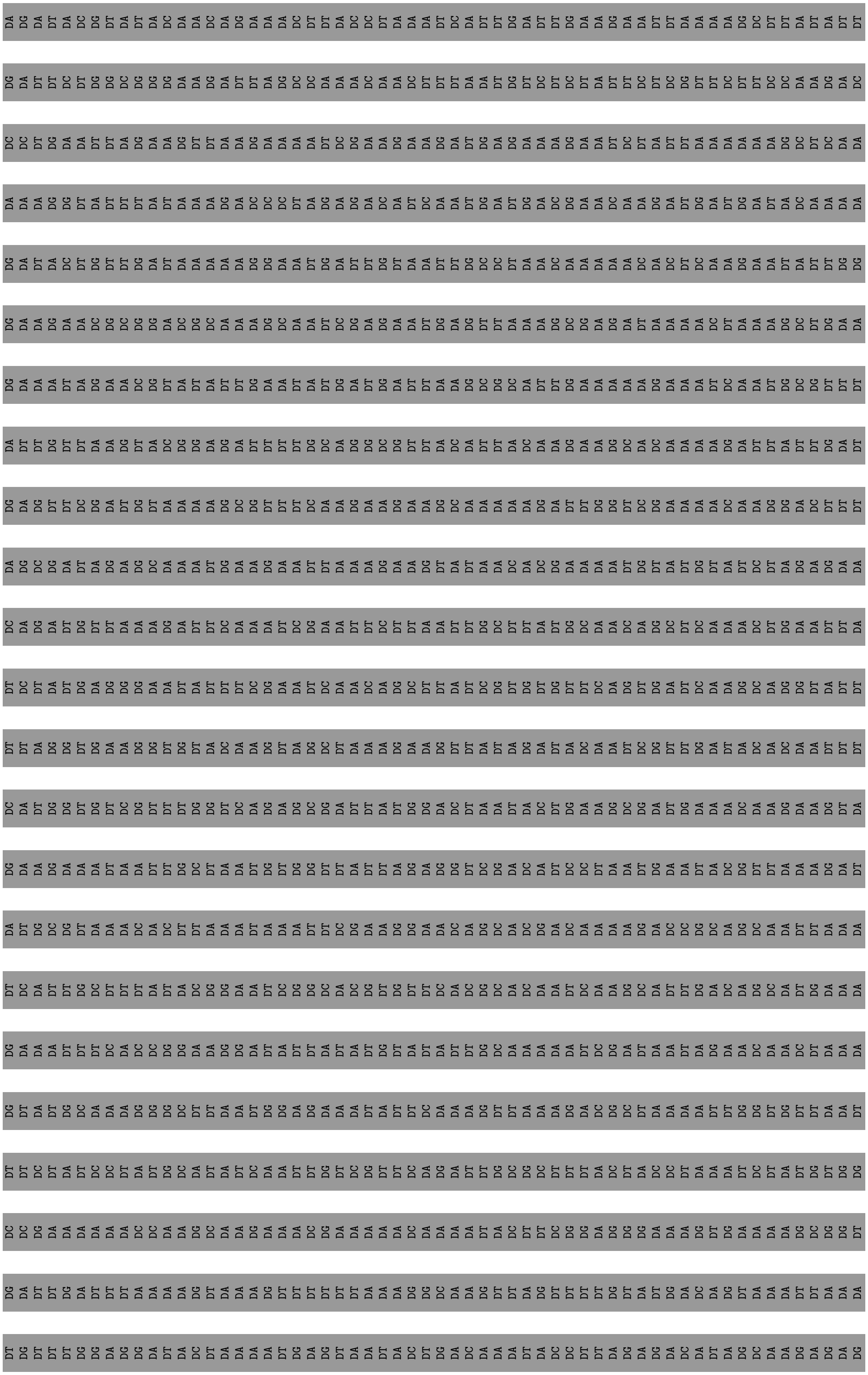
DT DT DC DG DG DC DG DT DA DA DT DC DG DC DA DC DT DG DA DT DG DC DT DT DG DT DT DG DT DC DA DT DA DT DC DG DC DC DT DA DT DC DT DA DT DA DC DA DG DT DG DG DC DT DA DG DT DT DA DC

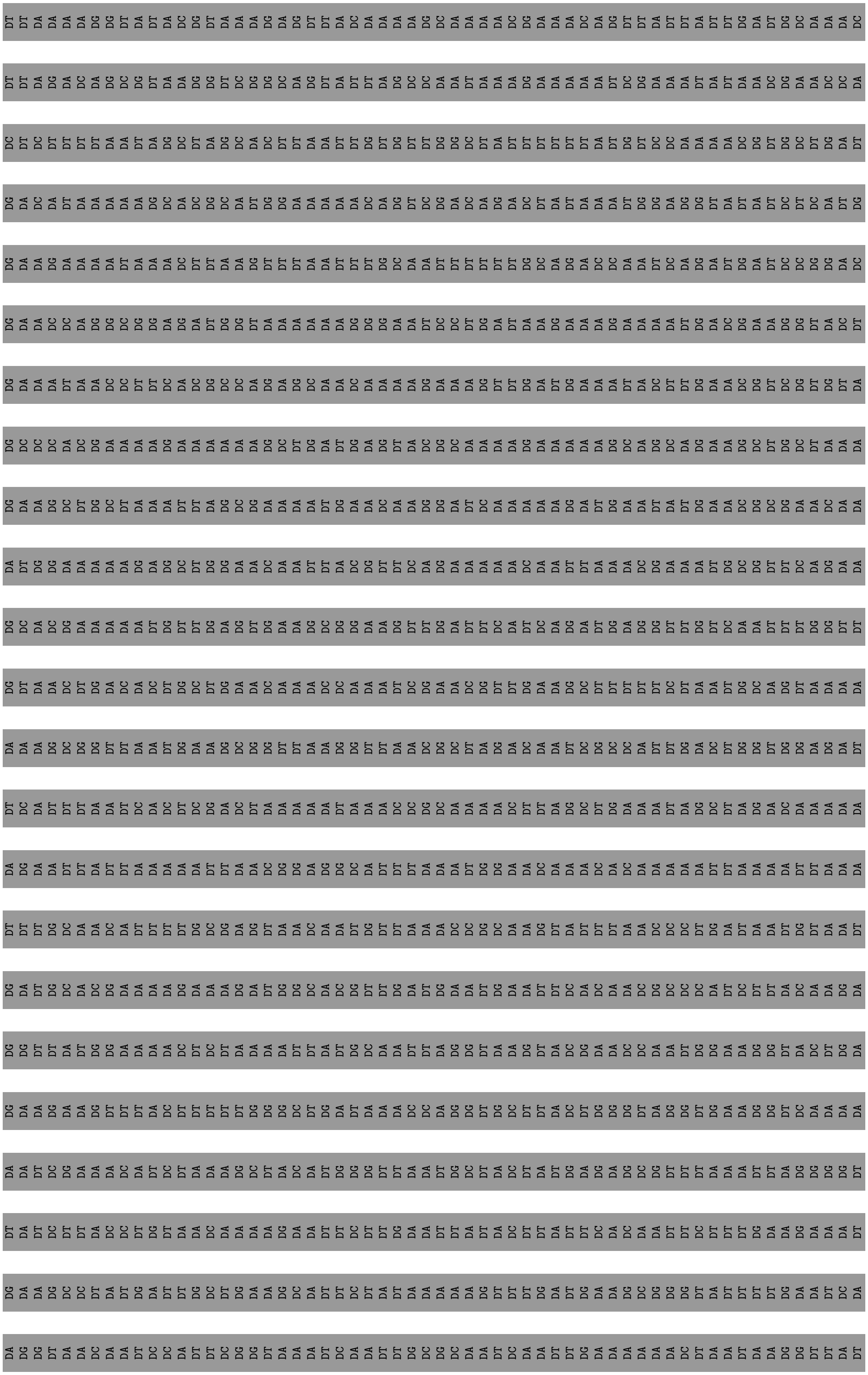
DC DA DA DA DA DC DA DA DA DG DA DA DG DT DA DC DA DT DC DA DA DG DG DG DA DC DA DA DT DT DA DC DA DG DA DT DA DA DA DT DA DT DA DA DC DA DA DG DA DG DA DC DA DA DG DA DT DA DA DA

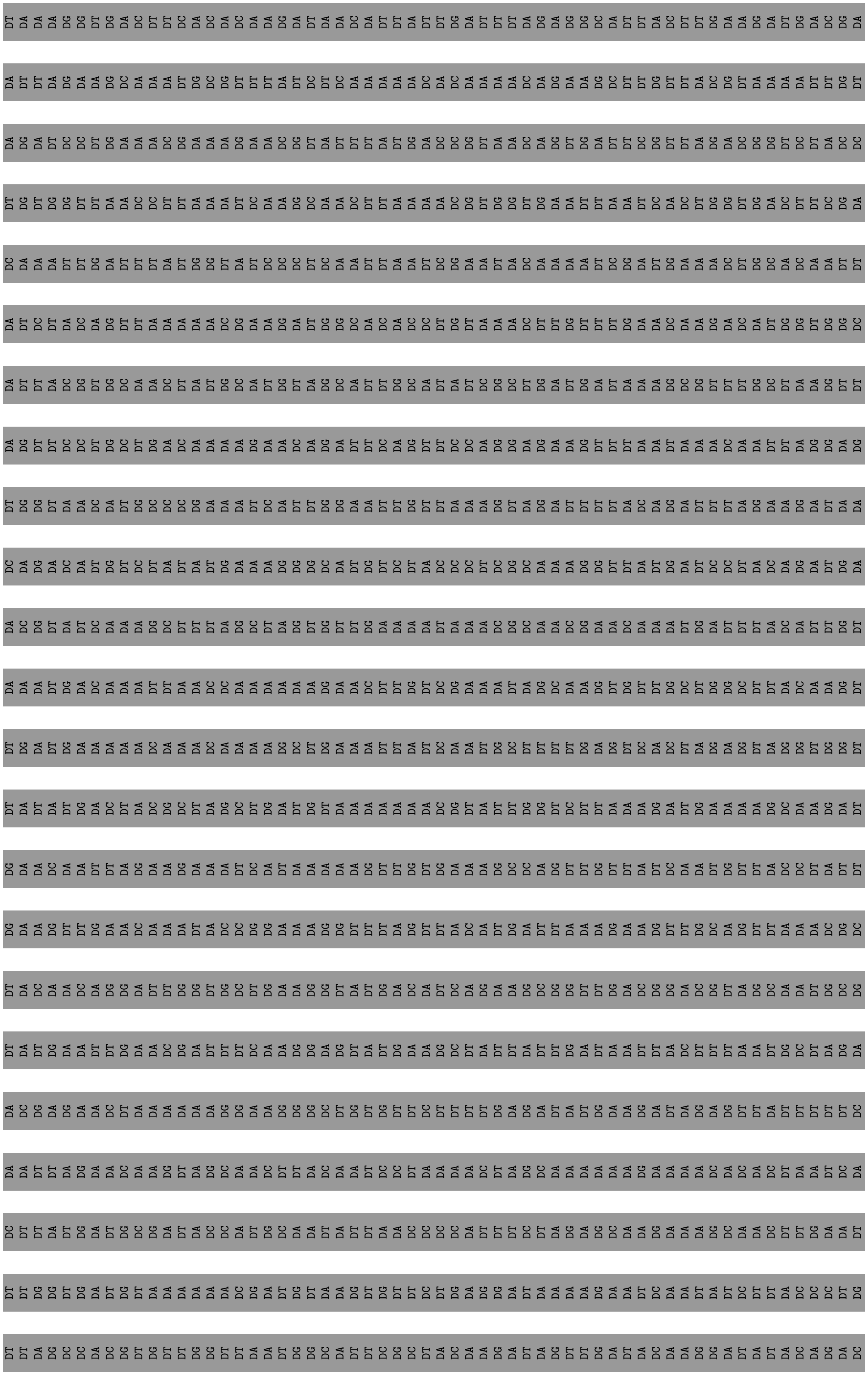


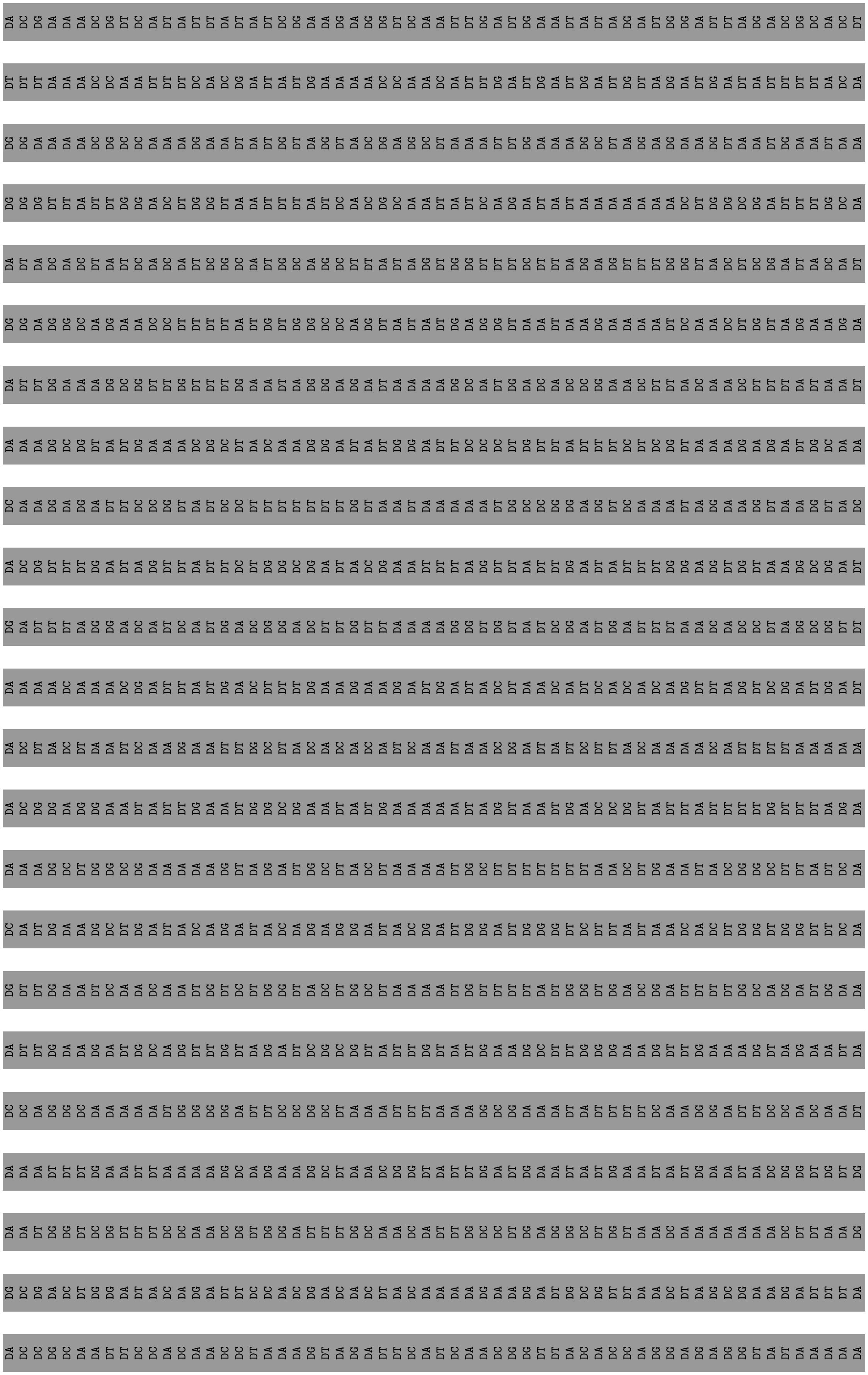


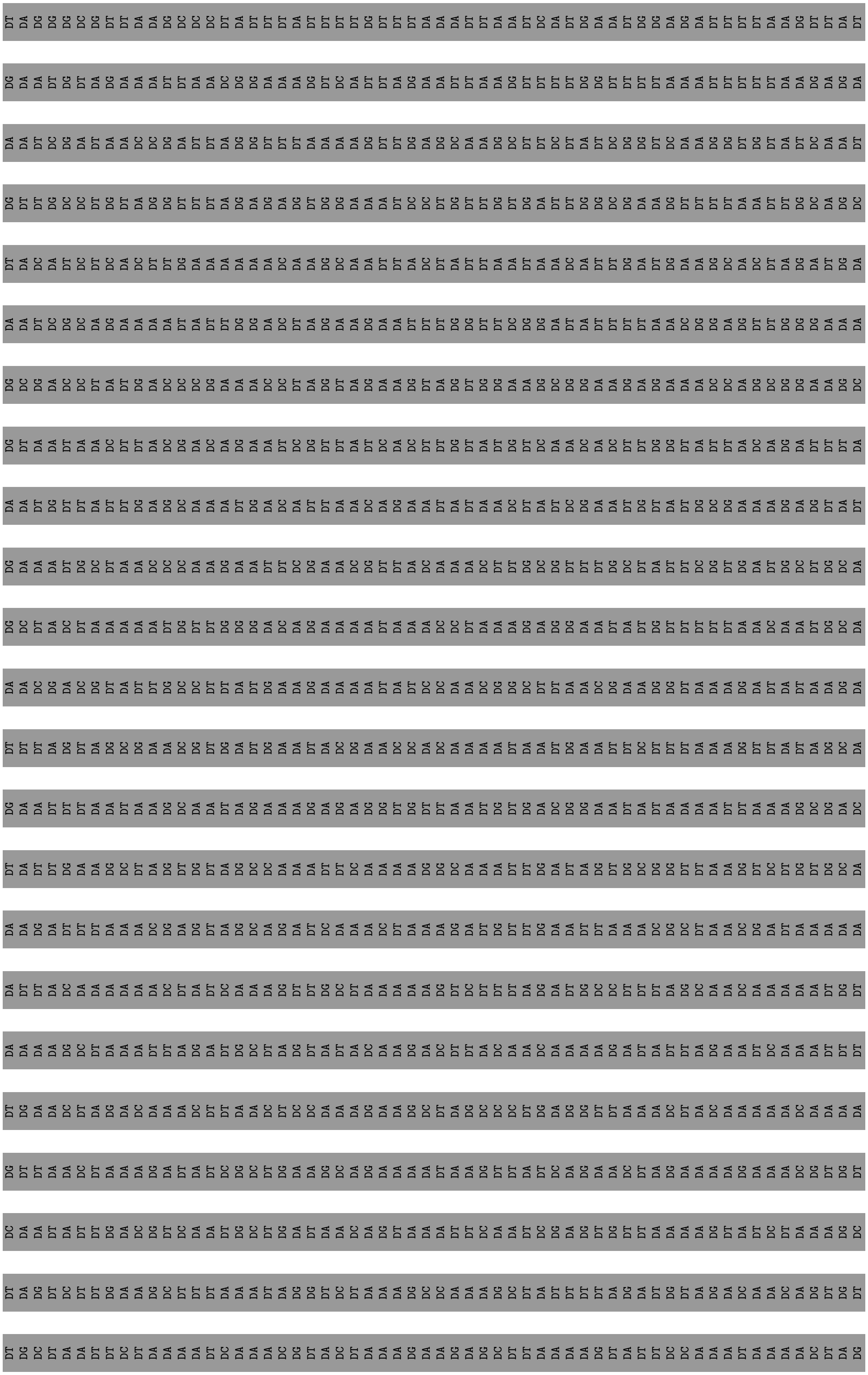


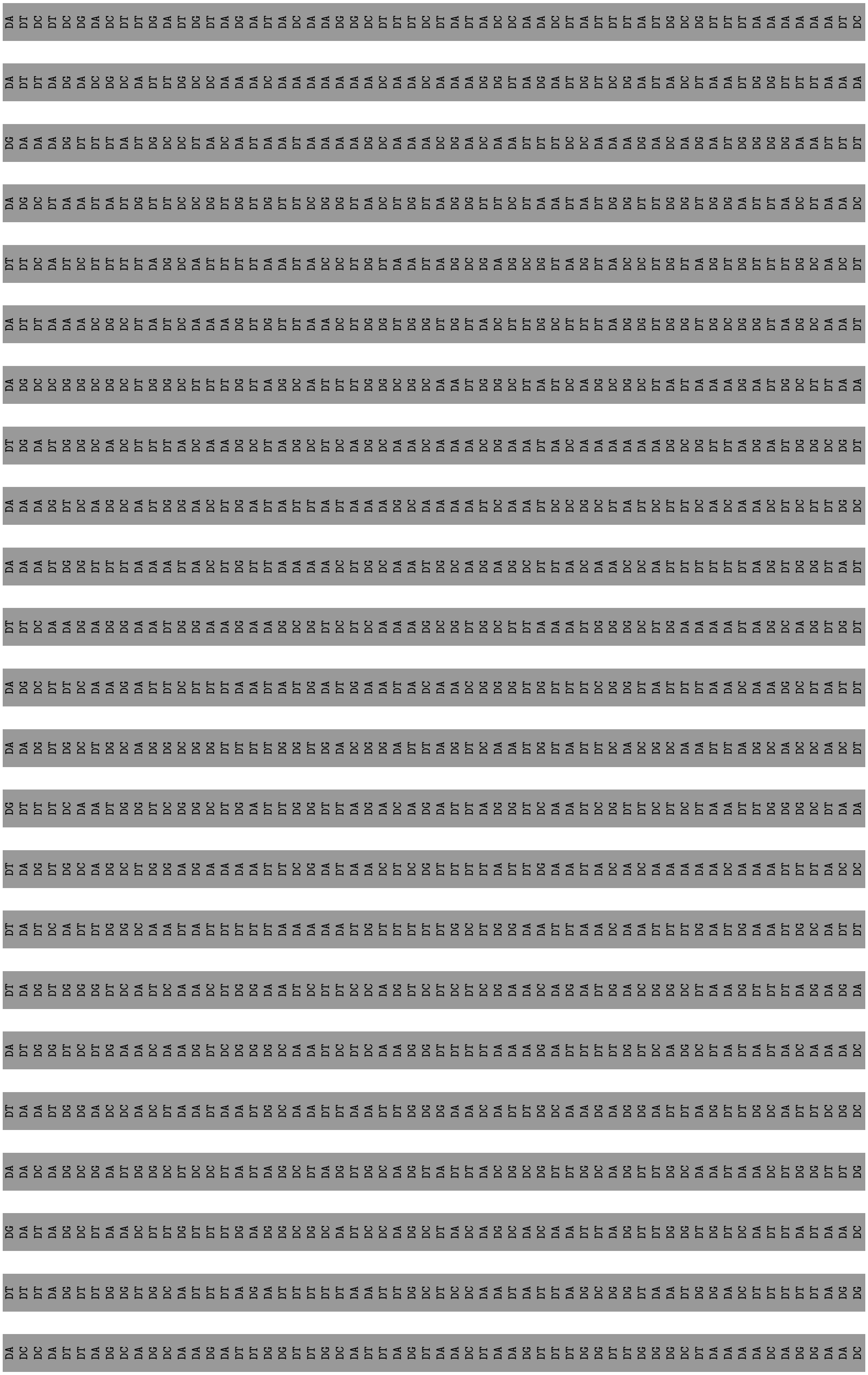


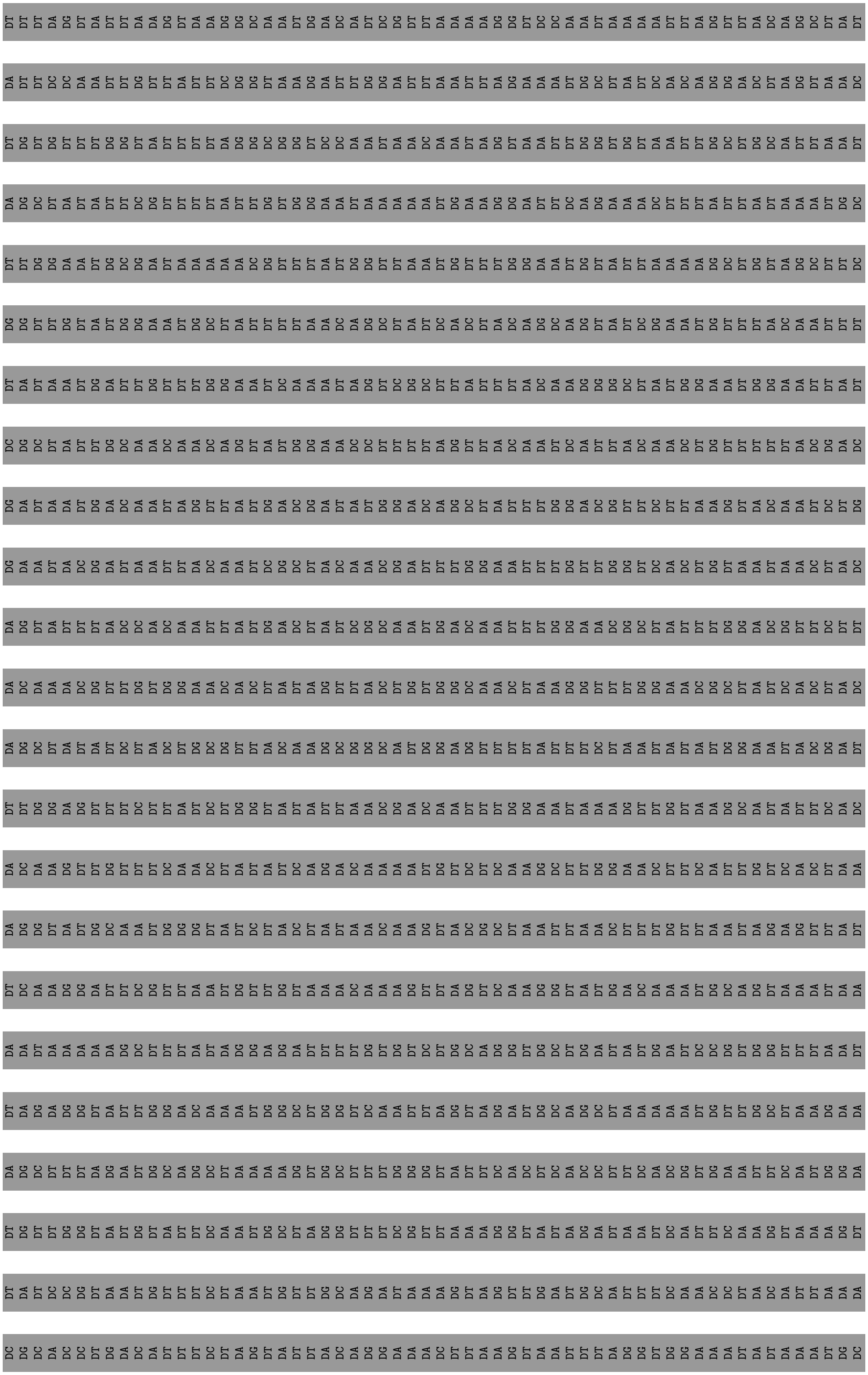












DG DC DA DA DG DT DA DC DA DA DC DA DC DA DC DA DC DA DT DT DC DT DA DT DT DG DA DA DA DC DA DT DC DA DC DC DG DA DA DC DA DT DG DA DA DA DA DC DT DG DT DT DA DA DA DG DT DT DG DA

DA DT DT DC DG DA DT DG DT DC DA DA DT DA DA DC DG DA DT DG DC DG DC DT DT DA DC DT DA DG DT DA DT DT DG DT DT DA DA DC DG DG DC DA DG DA DA DA DT DG DC DT DA DA DA DC DG DC DA DA

DT DT DC DT DG DA DG DT DA DT DT DA DC DT DT DA DT DA DA DA DG DG DA DG DG DT DT DA DC DA DA DA DT DG DG DA DC DA DT DA DG DA DA DT DT DA DA DC DA DA DA DA DA DA DG DA DT DG DG

DT DA DC DT DG DT DA DA DT DC DA DA DA DT DT DA DA DG DT DG DA DA DT DA DC DG DG DG DT DT DT DA DT DC DG DT DT DA DA DC DG DA DT DA DT DA DG DT DA DA DT DT DG DA DT DA DG DC DA DT

DG DC DA DA DA DT DC DA DA DC DA DC DA DA DA DG DT DA DT DC DA DA DG DA DC DA DA DA DG DA DA DA DA DT DA DT DG DA DA DC DG DG DT DC DG DT DA DT DA DT DT DA DA DT DG DG DG DG DA DG

DC DA DA DT DT DA DT DA DT DC DA DG DT DA DG DA DG DA DT DA DT DA DG DT DT DG DT DT DC DC DT DT DG DT DT DT DT DG DT DA DA DA DG DT DT DA DA DA DA DA DT DC DG DT DT DC DA DG DA

DC DA DT DT DG DC DT DT DA DT DA DT DG DC DG DA DG DA DT DA DT DG DT DT DG DT DA DT DT DC DG DT DT DA DA DC DG DA DC DA DG DA DC DA DT DA DG DA DA DC DC DT DA DT DG DT DA DT DT DT

DG DC DG DA DG DA DA DA DT DC DA DG DA DA DG DA DA DA DA DG DA DA DG DA DG DT DT DG DA DA DT DT DA DC DA DG DG DT DT DT DA DC DT DC DA DA DC DC DA DA DC DT DT DC DT DG DA DT DG DA

DT DT DA DC DG DT DG DA DA DA DT DT DA DG DA DT DA DA DA DA DA DC DA DA DC DT DT DC DC DC DG DG DA DT DT DA DC DG DA DA DT DA DT DT DC DA DA DG DA DC DA DC DG DA DT DC DA DA DC DA

DA DA DT DT DT DA DT DG DT DA DA DA DT DG DG DT DA DA DA DC DA DG DT DA DT DA DA DA DG DT DT DA DT DT DT DT DT DA DA DC DG DG DA DG DT DT DA DT DA DA DA DC DC DC DT DA DA DA DC DA

DA DA DA DA DG DG DT DA DA DT DA DA DA DG DT DT DT DC DT DT DT DT DG DA DA DC DT DA DA DA DA DT DT DC DG DA DA DA DC DT DA DC DA DG DA DA DT DT DA DC DC DA DT DA DC DG DG DT DG DA

DA DA DG DT DA DT DT DG DG DA DA DC DA DA DG DC DC DT DA DG DA DG DT DT DA DG DA DA DG DA DA DA DA DC DA DA DA DA DG DG DT DT DG DG DA DT DT DG DT DG DG DT DC DG DT DT DT DG DA

DT DT DT DT DA DA DT DA DT DT DG DA DT DT DG DG DC DA DT DG DC DA DG DG DC DG DG DA DG DA DC DA DA DA DA DG DA DA DA DG DT DA DT DA DC DA DT DT DT DG DA DA DA DA DT DT DT DG DA DG

DC DA DA DA DG DG DT DA DC DA DG DT DT DT DA DT DT DA DT DC DA DC DG DG DT DA DG DT DG DC DT DC DC DT DA DA DC DG DA DC DC DA DA DT DT DC DA DA DC DA DT DG DT DA DT DA DA DA DA DA

DG DA DT DA DA DC DA DA DT DT DA DT DT DT DT DA DG DG DC DG DA DA DG DA DT DA DC DA DG DA DA DT DC DG DT DT DT DG DT DA DT DG DG DA DA DT DT DT DA DA DC DG DC DA DT DG DC DT DG DA

DA DA DT DA DA DT DG DA DA DA DA DT DC DG DA DA DG DG DG DA DT DC DA DA DA DC DT DA DA DA DA DG DC DT DG DG DA DG DA DC DA DG DA DA DT DT DG DT DT DT DA DT DG DA DT DA DG DC DT DT

DC DC DG DA DG DT DT DT DA DT DA DA DA DA DA DC DG DG DT DG DT DT DG DA DA DA DT DA DA DG DT DA DC DC DG DA DA DA DC DG DA DA DT DA DT DA DT DC DC DC DA DA DC DC DA DA DA DA DT DT

DT DA DA DA DT DA DC DG DG DA DG DC DT DA DA DT DA DA DA DT DT DT DG DA DG DT DT DT DA DA DT DC DA DA DA DC DG DG DT DA DC DA DA DA DA DA DG DT DT DC DA DG DT DT DT DG DA DT DT DT

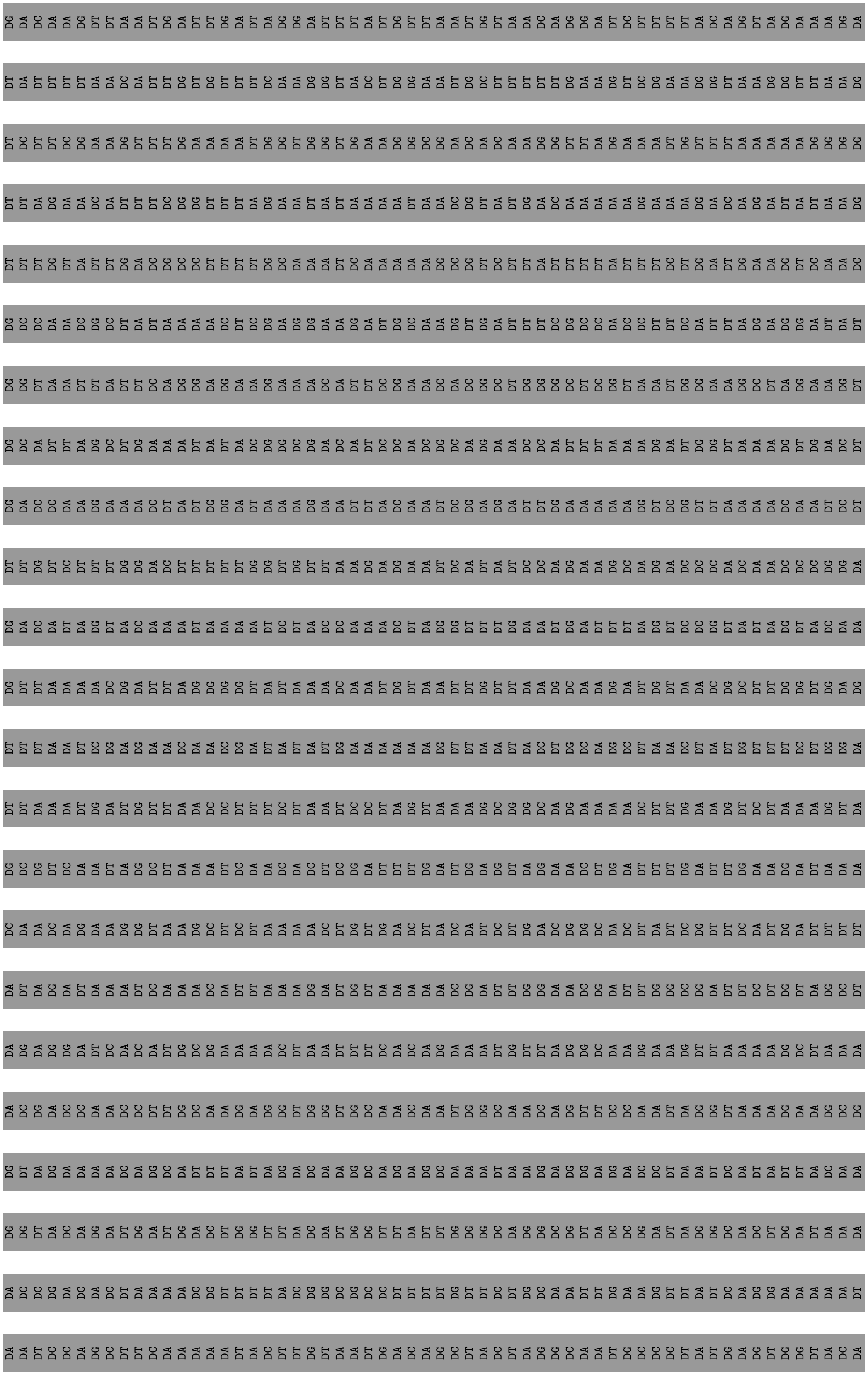
DG DA DA DA DT DT DT DT DA DT DT DA DT DA DA DG DT DA DG DG DT DG DT DC DA DG DA DA DT DG DA DC DA DA DT DA DA DC DT DA DT DT DA DA DA DC DC DA DC DC DT DA DA DA DG DG DT DA DA DT

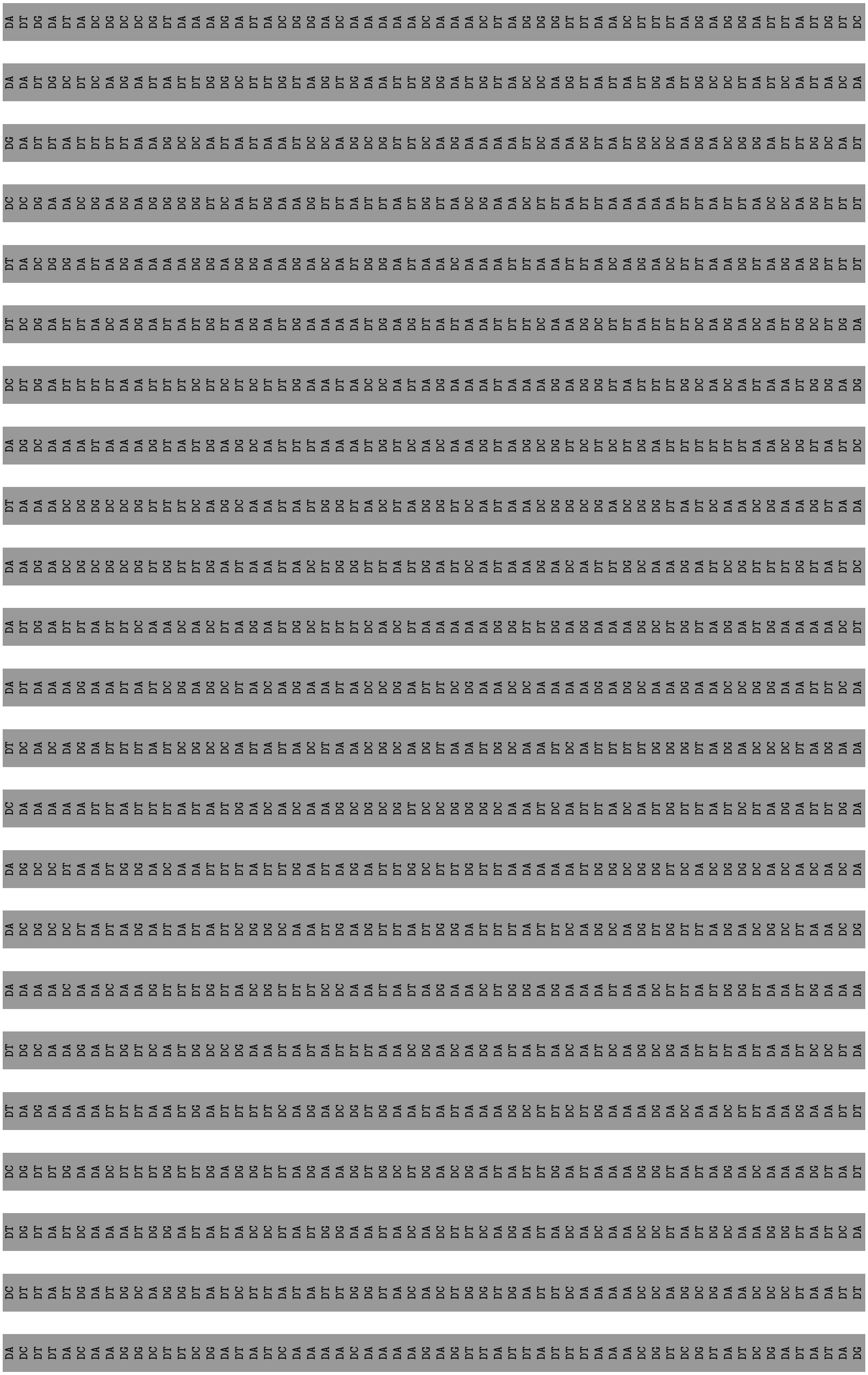
DG DG DC DG DC DA DC DC DT DG DT DA DC DC DA DG DT DA DG DA DA DA DC DA DA DC DT DT DT DA DG DT DA DA DA DA DA DA DG DT DT DA DA DT DG DC DT DG DA DC DG DG DT DG DT DA DT DT DA

DA DC DT DT DT DT DG DA DT DA DT DT DC DT DA DG DA DA DA DA DT DA DA DA DT DA DT DA DC DT DT DA DT DG DA DA DG DT DT DA DT DT DA DA DC DG DC DT DA DT DA DG DG DG DA DA DA DA DG DA

DT DG DG DA DT DT DG DT DT DA DG DT DC DA DT DG DT DC DG DA DA DG DG DT DG DA DA DA DA DC DG DA DC DA DA DG DA DA DA DG DA DA DT DA DT DG DT DA DA DT DA DA DC DT DG DT DC DA DT DT

DG DA DT DA DG DG DA DA DA DT DC DA DG DA DA DG DG DC DG DA DC DA DG DA DC DA DA DC DT DG DG DT DT DG DA DA DT DG DT DA DC DT DG DC DT DA DG DA DG DA DG DA DT DT DC DC DT DA DT DA





DG DC DG DG DT DG DT DG DA DA DT DA DA DC DA DA DC DT DT DT DA DA DA DG DG DA DG DA DT DT DT DC DC DA DA DG DA DG DG DC DT DG DA DA DG DG DT DC DT DA DG DA DT DA DT DG DT DA DT DT

DA DC DG DA DT DC DT DA DG DA DA DA DC DA DG DG DA DC DG DT DA DA DA DG DC DA DC DT DT DT DT DA DA DT DC DG DG DG DG DT DA DA DC DT DA DT DT DG DG DA DC DC DT DG DG DT DA DA DC DA

DA DC DA DG DA DC DA DT DC DA DC DT DC DA DA DT DT DT DA DT DT DC DT DA DT DC DG DG DC DC DA DA DA DG DA DG DG DT DG DT DT DA DA DC DC DA DA DT DT DC DT DT DG DA DA DA DA DA DC DA

DT DC DG DC DG DC DC DT DC DA DA DG DT DA DT DC DA DA DT DG DA DC DT DG DA DT DT DC DA DG DG DC DG DG DA DC DG DT DG DT DT DA DA DA DC DC DG DT DT DA DC DC DA DA DT DA DC DA DG DA

DA DC DC DC DA DG DC DA DT DA DT DC DT DA DA DG DT DG DA DT DA DT DT DA DC DG DG DA DA DG DT DT DG DG DT DC DA DT DT DA DC DT DA DT DA DT DC DT DA DT DA DC DG DC DA DA DG DA DC DA

DC DA DC DA DA DA DA DT DG DC DG DT DT DA DG DA DT DT DT DC DC DC DG DT DT DA DC DC DG DA DA DA DG DC DG DT DT DT DA DG DA DG DA DT DG DC DA DG DG DT DT DG DG DT DT DC DT DT DT DG

DA DT DG DT DA DC DT DG DC DC DT DG DG DA DC DA DC DT DA DT DA DA DT DG DG DT DG DC DT DC DT DA DA DG DA DC DA DA DG DT DA DC DT DT DA DC DC DA DG DA DA DA DC DA DG DC DA DC DA DG

DG DT DA DG DA DA DA DT DA DT DG DC DT DT DA DA DA DT DT DC DG DA DA DC DG DT DG DT DC DA DT DT DG DA DC DA DT DT DT DT DC DA DA DT DA DA DG DA DA DA DA DA DC DA DA DC DG DG DA DG

DC DA DT DG DG DA DA DT DT DT DC DT DG DC DC DC DG DC DA DA DA DA DC DG DC DC DG DG DT DT DA DT DT DG DG DG DA DA DC DA DT DA DT DC DC DC DT DA DA DG DA DG DT DA DT DT DA DC DA DA

DA DA DT DT DA DT DC DA DG DA DT DT DT DA DA DA DA DA DT DC DG DT DT DG DG DT DT DT DA DG DA DT DT DT DC DT DA DT DA DT DC DA DC DT DA DC DT DG DA DA DG DA DA DT DC DA DA DA DC DC

DG DA DT DT DT DA DC DT DG DA DT DT DT DC DC DC DT DA DA DA DG DA DC DT DT DT DA DA DA DG DG DT DA DT DT DG DC DA DG DG DT DT DG DG DA DT DA DT DT DA DG DA DA DG DT DA DA DA DA DT

DC DG DA DA DT DA DC DA DC DC DA DG DG DT DA DA DT DA DC DA DA DC DA DC DA DA DG DT DA DT DT DA DA DG DA DC DG DT DA DA DT DA DA DC DT DT DC DC DC DG DT DC DT DG DC DA DC DA DT DC

DA DA DT DT DT DT DT DA DG DT DT DA DG DA DA DA DC DT DT DT DG DG DT DA DC DT DG DG DT DG DG DC DG DT DT DG DG DT DA DA DA DT DG DG DA DG DT DT DT DA DT DT DC DG DA DA DG DG DA DA

DA DG DG DT DG DG DT DT DG DA DA DT DA DA DT DG DG DT DA DG DT DA DG DA DT DA DA DT DT DT DT DT DC DG DA DA DA DG DA DT DG DA DT DA DA DC DT DT DA DA DT DC DG DA DG DT DT DA DC DA

DA DA DC DA DA DC DA DT DC DA DC DA DA DT DA DT DA DA DT DC DC DG DG DT DT DA DT DT DG DA DC DA DC DA DA DA DC DA DT DC DA DG DT DT DT DC DT DA DT DG DA DA DT DC DA DG DA DT DA DG

DA DG DG DA DA DC DT DG DG DT DG DT DT DT DT DA DA DA DT DT DT DT DG DC DA DG DT DA DA DC DT DA DA DG DA DA DT DA DA DC DA DG DA DC DC DG DT DT DA DT DC DT DA DT DA DA DG DT DT DC

DT DG DA DA DC DA DT DG DT DT DA DA DA DA DC DA DT DT DT DA DT DC DG DT DG DT DT DA DA DA DA DA DC DC DG DA DT DG DA DT DT DA DT DA DA DC DG DT DA DG DA DT DA DG DA DG DG DC DG DC

DT DT DA DT DA DT DT DT DC DA DG DA DC DG DA DA DT DT DA DA DC DG DA DT DA DG DT DA DG DA DC DG DC DA DA DT DT DA DA DT DG DG DG DC DG DT DT DT DG DC DA DG DT DA DT DG DT DG DA DT

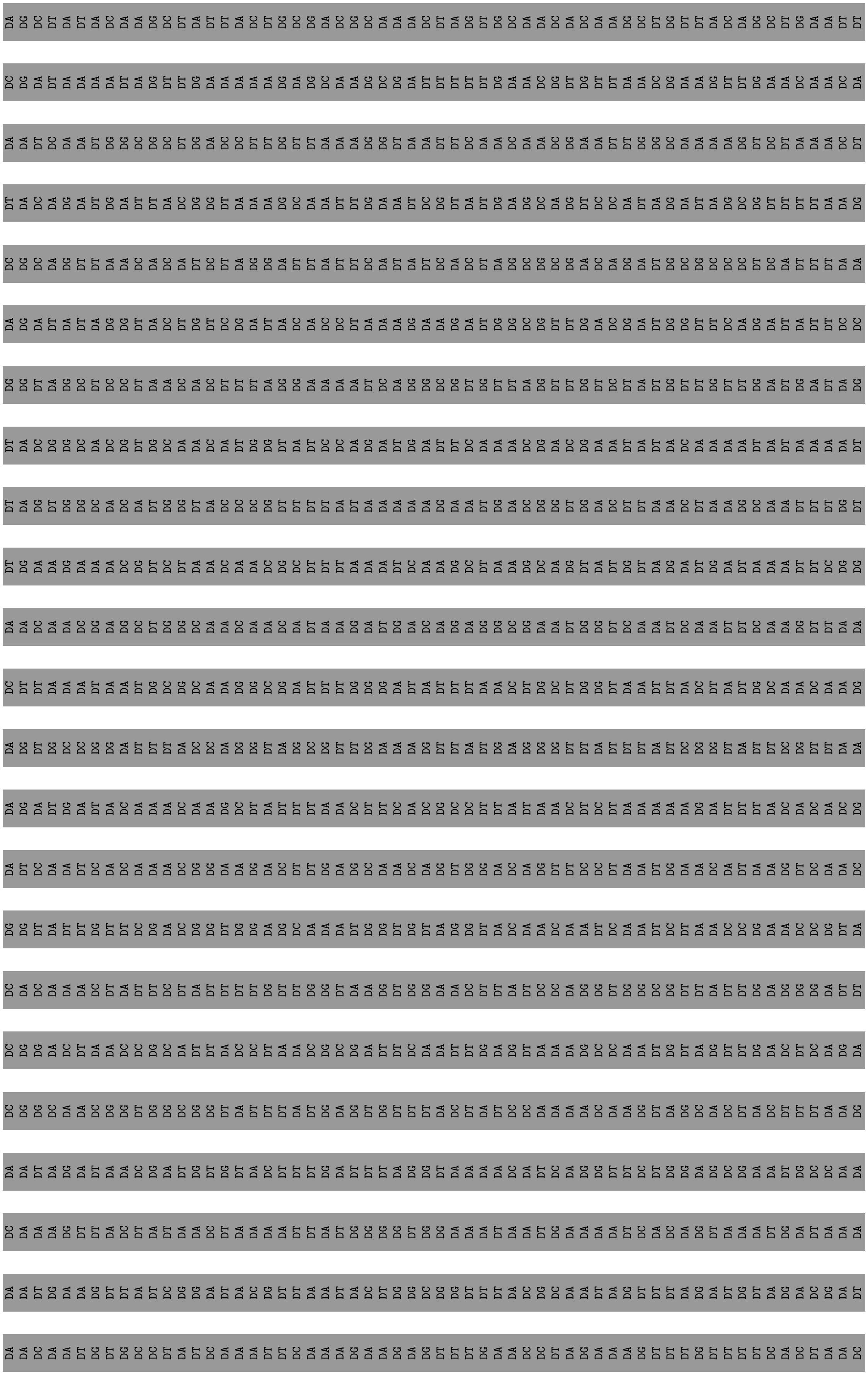
DA DC DC DG DA DA DT DG DA DA DT DT DT DT DT DA DA DA DA DC DA DT DT DC DA DG DG DC DA DA DG DG DT DG DC DA DT DG DC DT DC DA DG DG DC DA DT DT DC DT DT DT DA DC DA DC DA DA DA DA

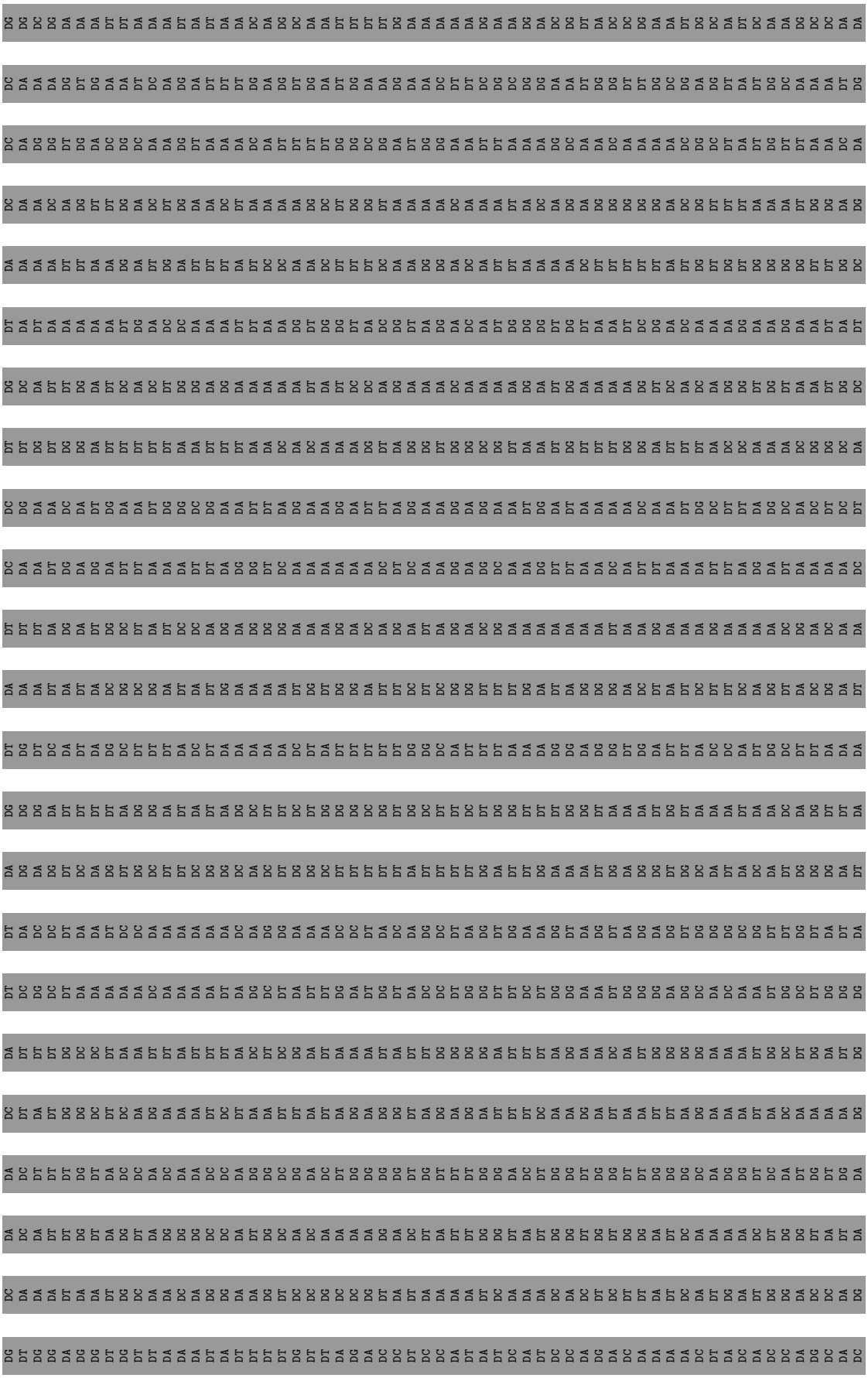
DC DG DG DG DA DG DT DG DA DT DA DA DT DG DT DT DG DT DT DG DT DT DG DA DA DC DG DT DC DA DA DT DT DT DA DG DC DT DT DC DA DA DT DA DT DT DG DA DA DA DA DT DG DA DT DT DT DA DG DT

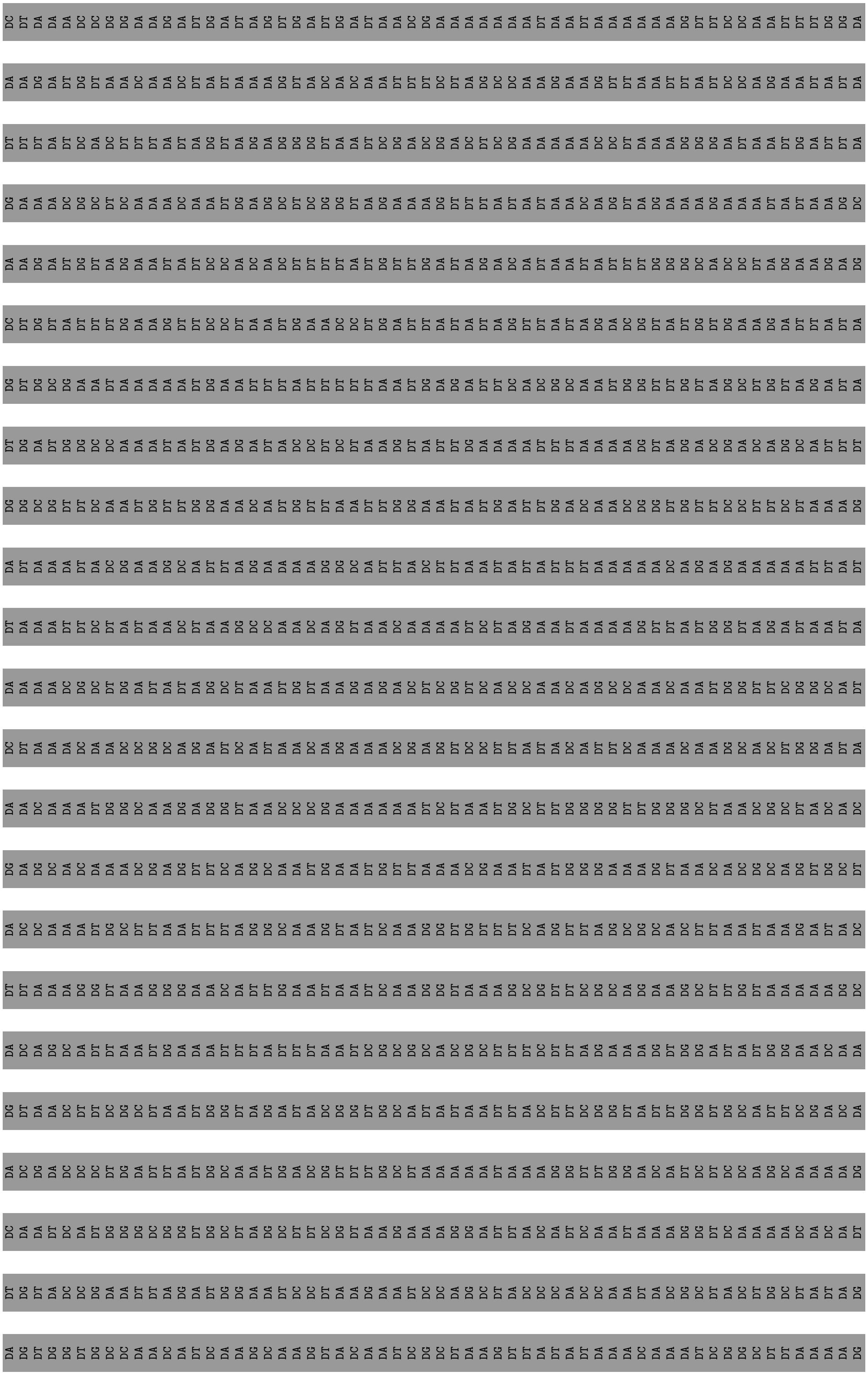
DT DA DG DT DG DG DG DT DT DT DG DA DT DG DG DT DA DT DA DA DC DA DA DA DG DC DT DT DG DT DT DT DA DT DA DT DC DA DA DA DT DC DT DA DT DT DC DA DA DG DA DT DA DC DT DA DT DC DG DA

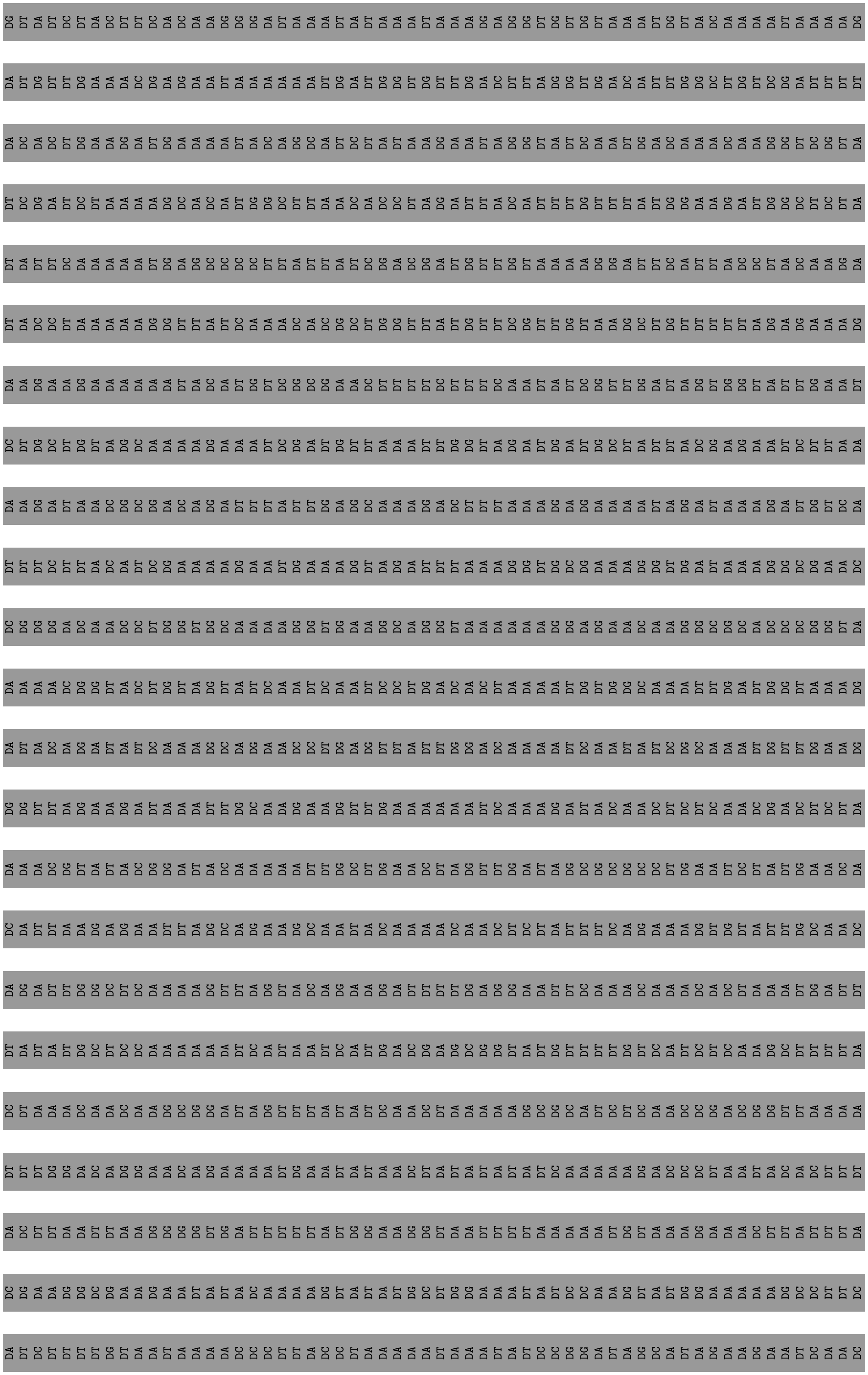
DA DG DC DT DG DT DC DG DG DT DA DA DA DG DA DC DT DT DT DA DA DC DC DA DA DT DT DA DA DA DG DC DA DA DA DA DT DA DT DG DG DC DT DG DA DT DA DC DA DC DA DA DA DC DG DT DT DA DA DT

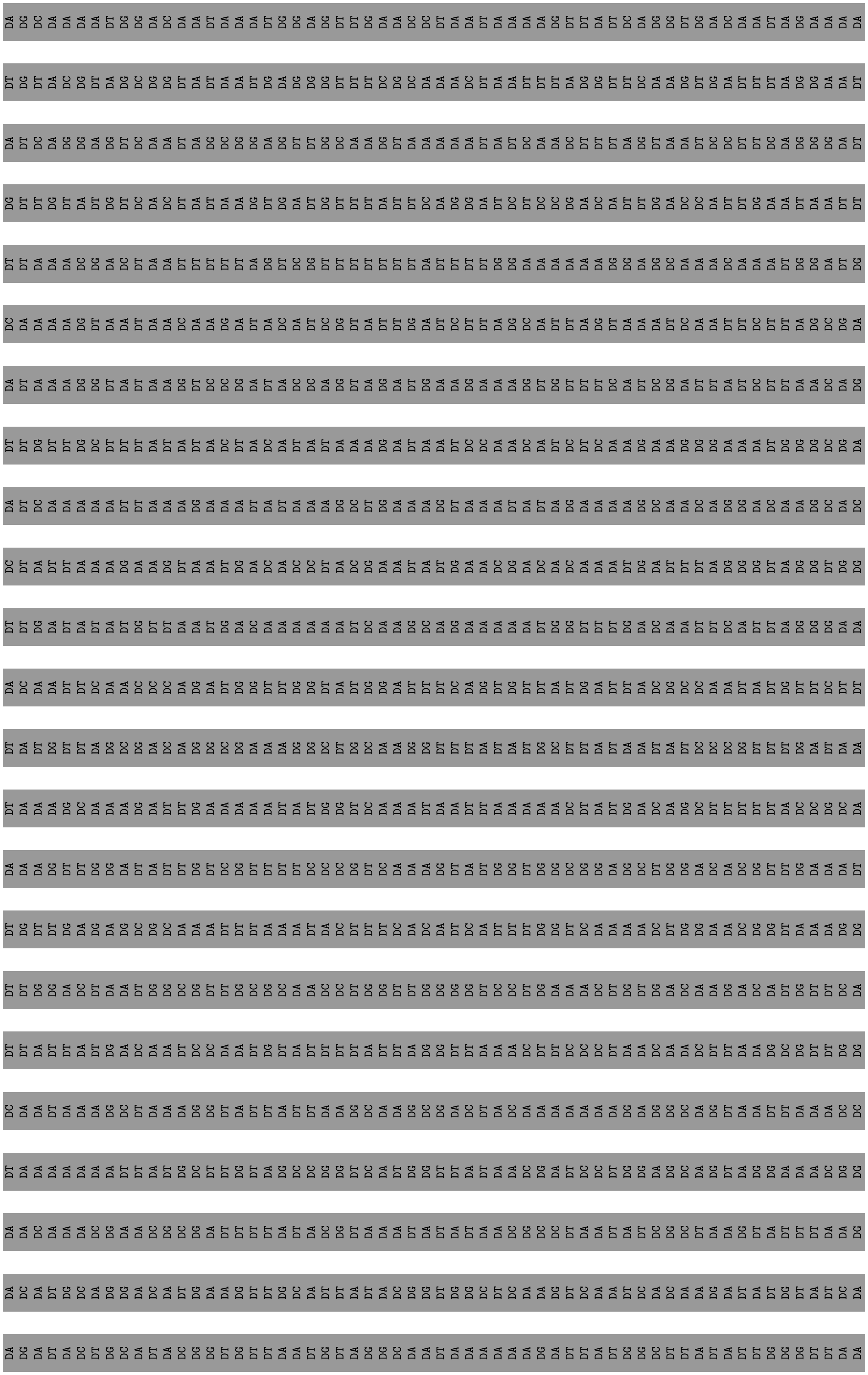
DA DG DC DA DA DA DA DG DT DG DA DA DT DG DA DT DA DG DT DG DC DG DA DC DA DA DA DA DG DG DC DA DT DT DC DA DA DC DA DA DA DT DC DG DA DA DA DT DC DA DA DG DC DA DA DA DA DC DG DA

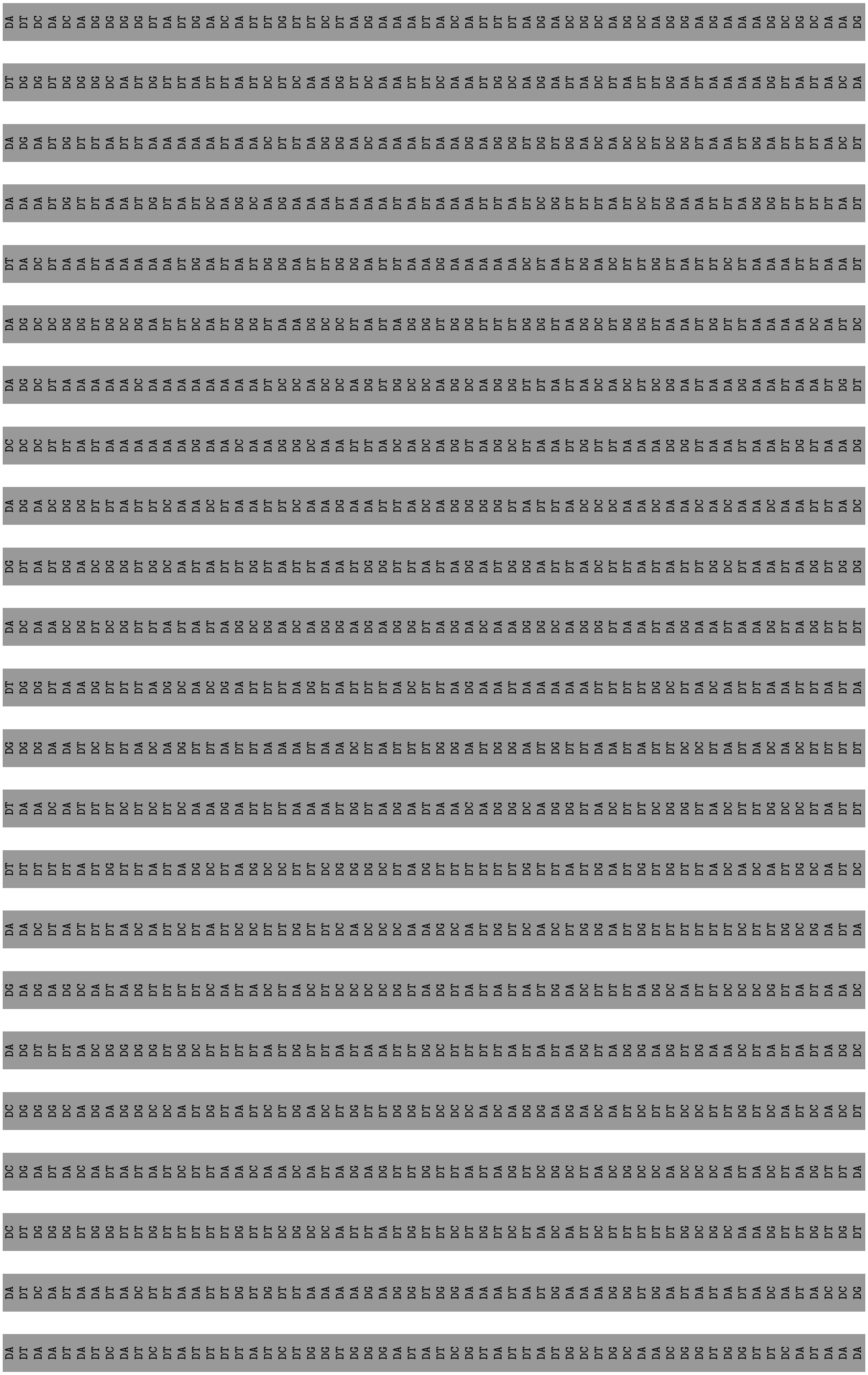


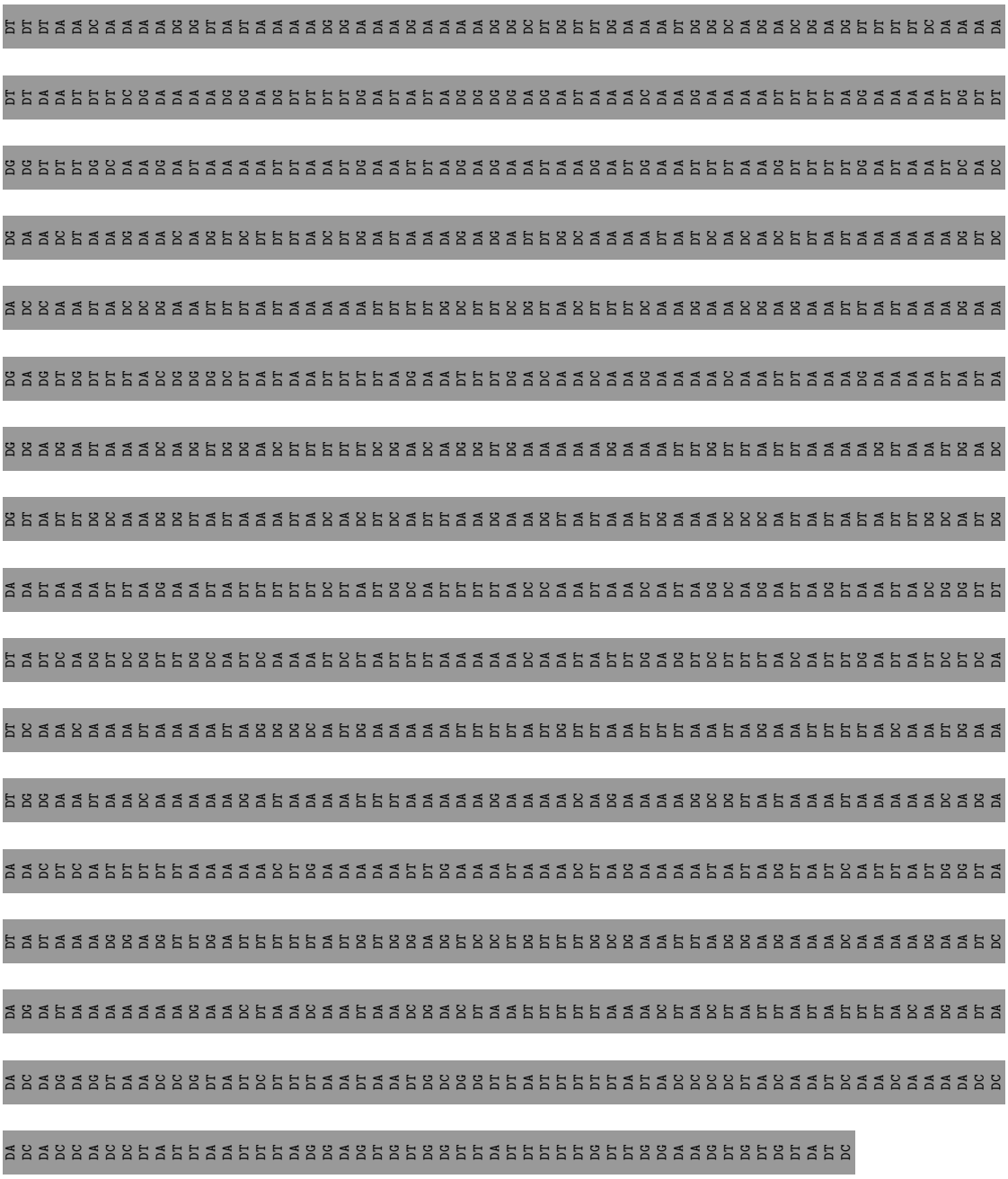






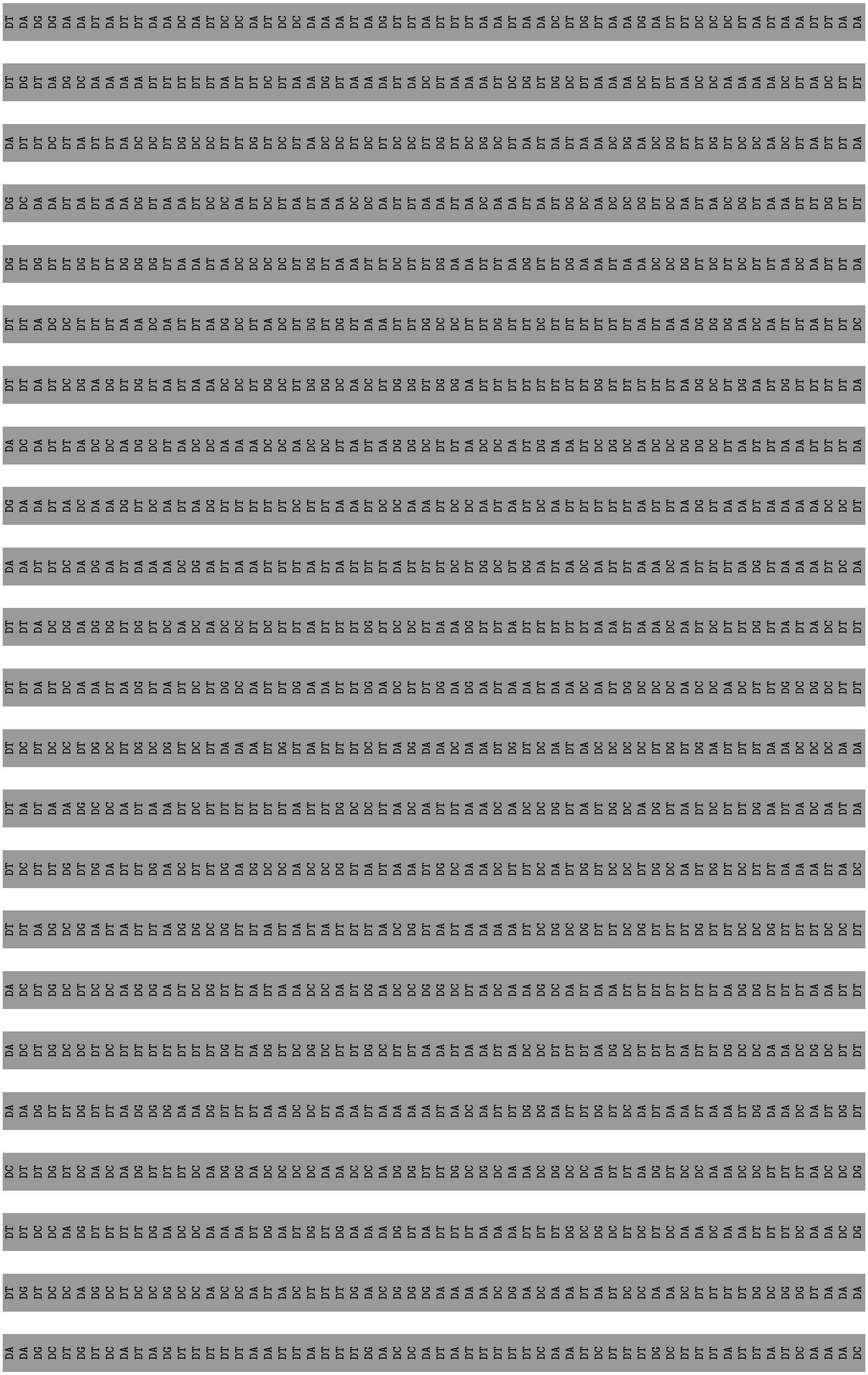


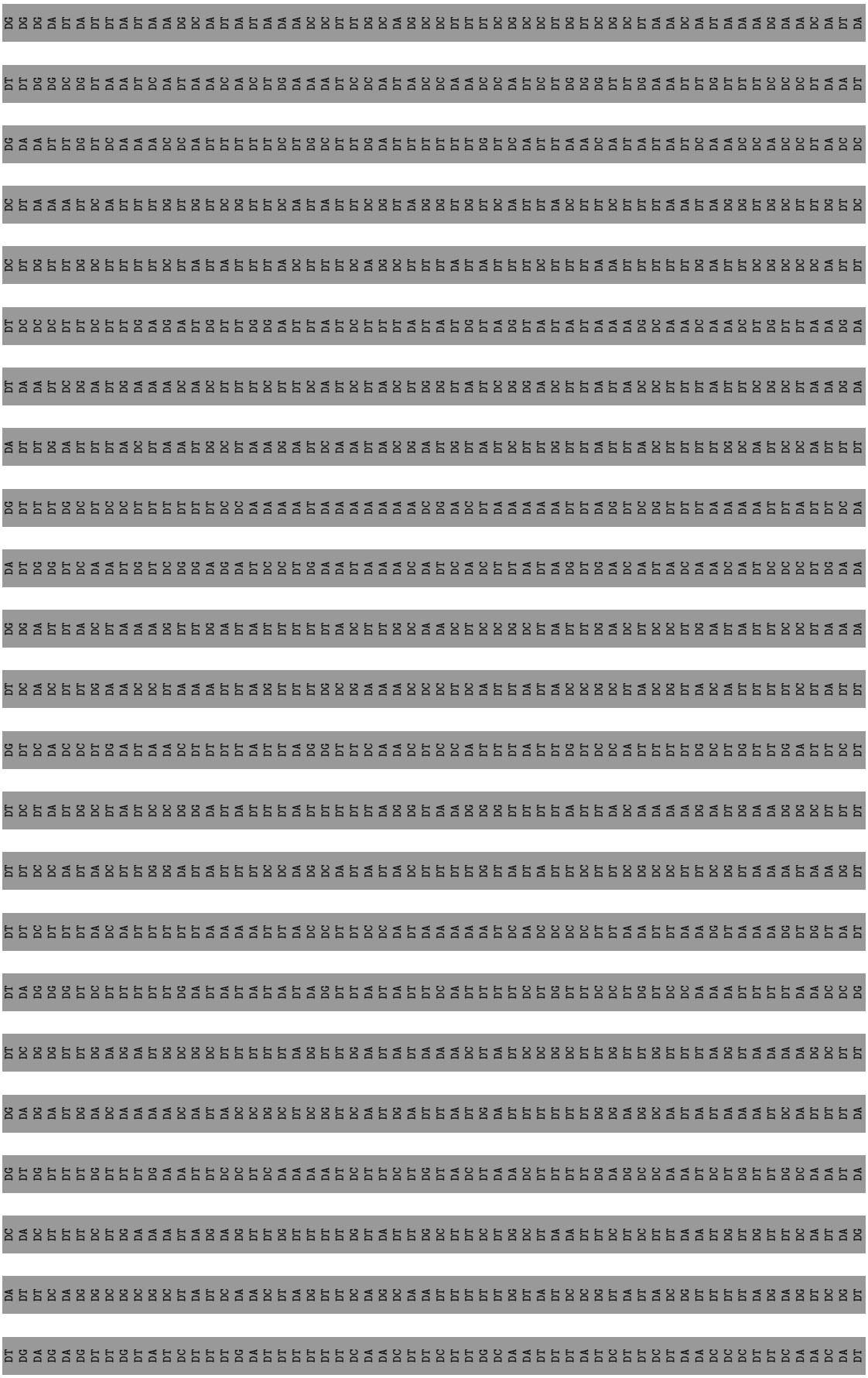

DA DT DC DA DC DA DG DG DG DT DA DT DG DA DC DA DT DT DG DT DT DC DT DA DG DA DA DA DT DA DC DA DT DT DT DA DG DA DC DG DC DA DG DC DA DG DG DA DG DA DA DA DG DC DG DC DA DG
DT DG DG DT DG DG DC DA DT DG DT DT DA DT DA DT DA DT DC DT DC DA DA DG DT DC DA DA DT DT DC DA DA DT DG DC DA DG DA DT DA DC DT DA DT DT DG DA DT DA DA DA DG DT DA DT DA DC DA
DA DG DA DT DG DT DT DA DT DT DA DA DA DA DT DA DA DC DT DT DA DG DG DA DC DA DA DA DT DA DA DG DA DG DG DT DG DT DG DA DC DA DC DC DT DC DG DT DA DA DT DG DA DT DT DT DA DC DT
DA DA DA DT DG DT DT DA DA DT DG DT DA DT DC DA DG DC DA DG DA DA DA DT DA DA DA DT DA DT DA DA DA DT DT DA DT DC DG DT DT DT DA DT DC DT DG DA DA DT DT DA DG DG DT DT DT DT DA DT
DT DA DC DT DA DA DT DA DA DA DA DT DG DA DT DA DT DG DG DA DT DT DG DG DA DT DT DA DA DG DA DA DA DA DC DT DA DT DG DA DC DT DT DG DT DA DT DT DC DT DA DA DA DT DT DA DA DT
DA DG DC DC DG DG DT DG DC DG DA DT DT DC DA DT DG DG DT DA DA DG DC DC DT DA DT DA DG DG DT DG DG DT DT DT DG DG DT DA DG DC DT DG DG DT DA DA DT DG DT DT DA DA DA DC DA DT DC
DA DG DC DT DA DA DA DA DC DA DA DA DA DA DA DT DC DC DA DC DC DA DG DT DG DC DC DA DG DC DA DG DG DT DT DA DT DA DC DA DC DT DC DG DA DT DA DA DG DA DA DT DA DA DT DG DT
DC DC DC DT DT DA DT DA DA DA DA DG DA DC DA DG DG DC DA DA DT DT DA DC DA DC DA DG DT DA DG DC DT DA DA DT DG DT DT DA DA DG DG DT DA DT DA DA DT DG DT DA DG
DA DG DA DC DG DG DT DT DA DT DC DA DA DC DT DA DA DT DT DC DA DA DG DA DA DT DT DA DC DA DG DG DG DG DT DA DT DT DA DC DC DC DA DA DC DA DA DC DA DC DA DA DC DA DT DT DA DC
DG DT DA DT DG DA DC DG DG DT DG DC DA DT DA DT DT DG DT DA DT DT DA DA DT DG DG DT DT DA DT DA DG DA DT DG DG DA DT DT DA DC DT DT DA DT DA DT DT DG DC DT DA DA DT DA DG DT DG DG
DA DC DA DA DC DG DT DC DG DT DT DA DT DA DT DA DG DC DG DA DC DA DG DG DA DG DA DG DG DT DA DG DA DC DA DA DG DG DC DA DG DG DT DA DA DT DA DG DA DA DT DA DA DG DT DA DG DT DT DT
DT DG DG DT DA DA DG DT DT DT DA DG DC DA DC DG DA DT DT DT DA DG DT DA DT DT DT DA DC DT DT DA DG DA DA DT DA DA DA DA DT DT DT DG DC DT DA DC DA DT DT DA DA DT DT DA DT DA
DG DG DG DA DA DT DC DT DT DA DC DA DG DT DT DA DT DT DA DA DT DA DA DC DT DA DT DT DT DG DG DA DT DG DG DA DT DG DT DT DA DA DT DA DT DT DC DC DT DA DT DA DC DA DC DT DT DT
DT DA DA DC DA DT DT DT DC DT DC DT DC DA DA DG DA DT DT DT DA DA DA DT DG DT DA DG DA DT DA DC DA DG DG DC DA DG DG DT DA DC DT DT DC DG DG DT DA DC DT DT DG DC DC DT DA DT DT
DT DT DT DT DT DA DT DG DT DT DA DT DA DG DC DT DA DG DC DC DT DT DC DG DG DC DT DA DG DT DT DT DT DT DT DG DT DT DA DT DG DA DT DG DT DG DT DT DA DC DA DC DA DT DG DC DA DT DC
DA DA DC DT DA DT DT DT DA DC DA DT DC DT DA DT DC DC DT DT DG DT DT DC DA DC DC DC DA DA DG DC DA DT DG DT DC DA DC DT DG DG DA DT DG DT DT DT DT DT DC DT DT DG DC DG DA DT DA
DG DA DG DA DG DC DA DT DA DG DT DT DT DT DC DA DT DA DC DT DA DC DT DC DC DC DG DT DA DG DT DA DT DA DT DA DT DG DA DC DT DT DT DA DG DC DA DT DT DC DC DC DG DT DA DT DA DC
DA DG DT DT DT DA DC DG DG DG DT DG DC DT DT DT DT DA DT DG DT DT DA DT DA DA DT DT DG DC DT DT DT DT DA DT DA DT DA DG DT DA DG DG DA DG DT DG DA DC DT DA DT DA DT DA DG DC
DC DG DG DG DC DA DG DA DG DC DC DA DT DG DT DA DT DC DT DG DA DC DT DG DT DT DG DG DT DC DC DC DA DC DA DG DG DA DG DA DC DA DT DC DT DT DC DC DT DT DG DT DC DA DT DC DA DC DT
DC DG DA DT DA DC DA DT DA DT DA DT DC DT DT DA DA DC DA DA DC DA DT DA DG DA DG DT DT DG DT DT DA DT DA DG DT DC DG DC DT DA DC DG DC DC DA DC DC DC DA DT DA DC DT DA DG DT DA
DC DT DG DG DG DT DG DG DT DT DG DT DT DT DT DG DG DT DC DG DC DC DA DT DT DA DT DG DT DT DC DT DG DT DC DT DA DC DA DT DC DT DT DT DT DG DC DG DC DA DA DG DT DT DG DT DG DT
DA DT DC DA DT DA DA DT DA DC DT DT DA DA DT DT DG DT DG DT DT DA DA DA DG DA DG DG DT DG DG DA DA DA DT DA DT DG DA DA DA DG DG DT DG DA DT DA DT DA DT DA DC DA DT DA DC DC DG
DA DT DA DA DT DA DT DT DC DA DT DC DT DA DT DT DT DT DA DT DC DT DG DG DT DG DG DA DT DA DT DC DG DT DA DT DT DA DT DG DC DT DG DC DA DA DC DG DG DT DG DT DT DC DA DT DA DA DA

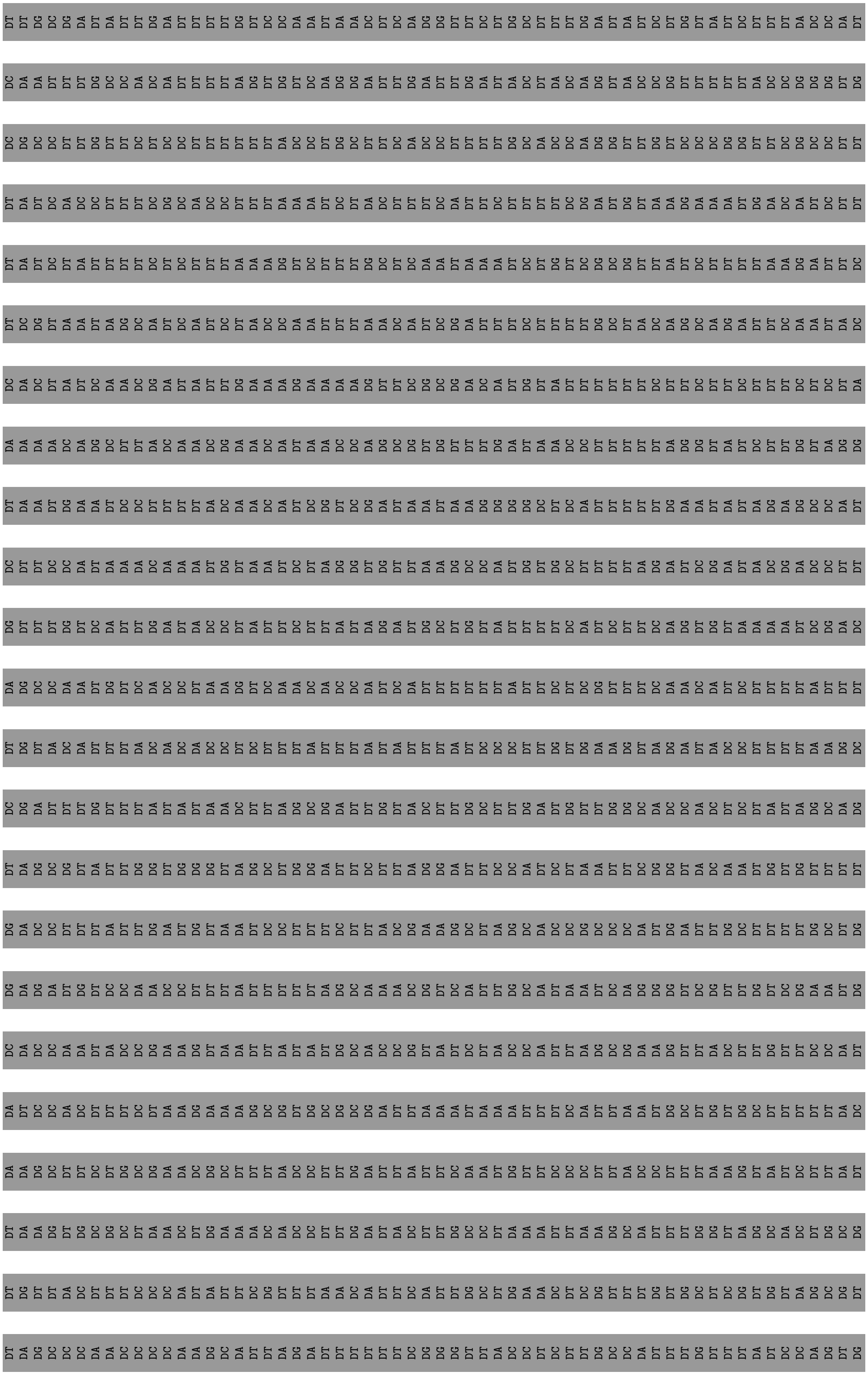


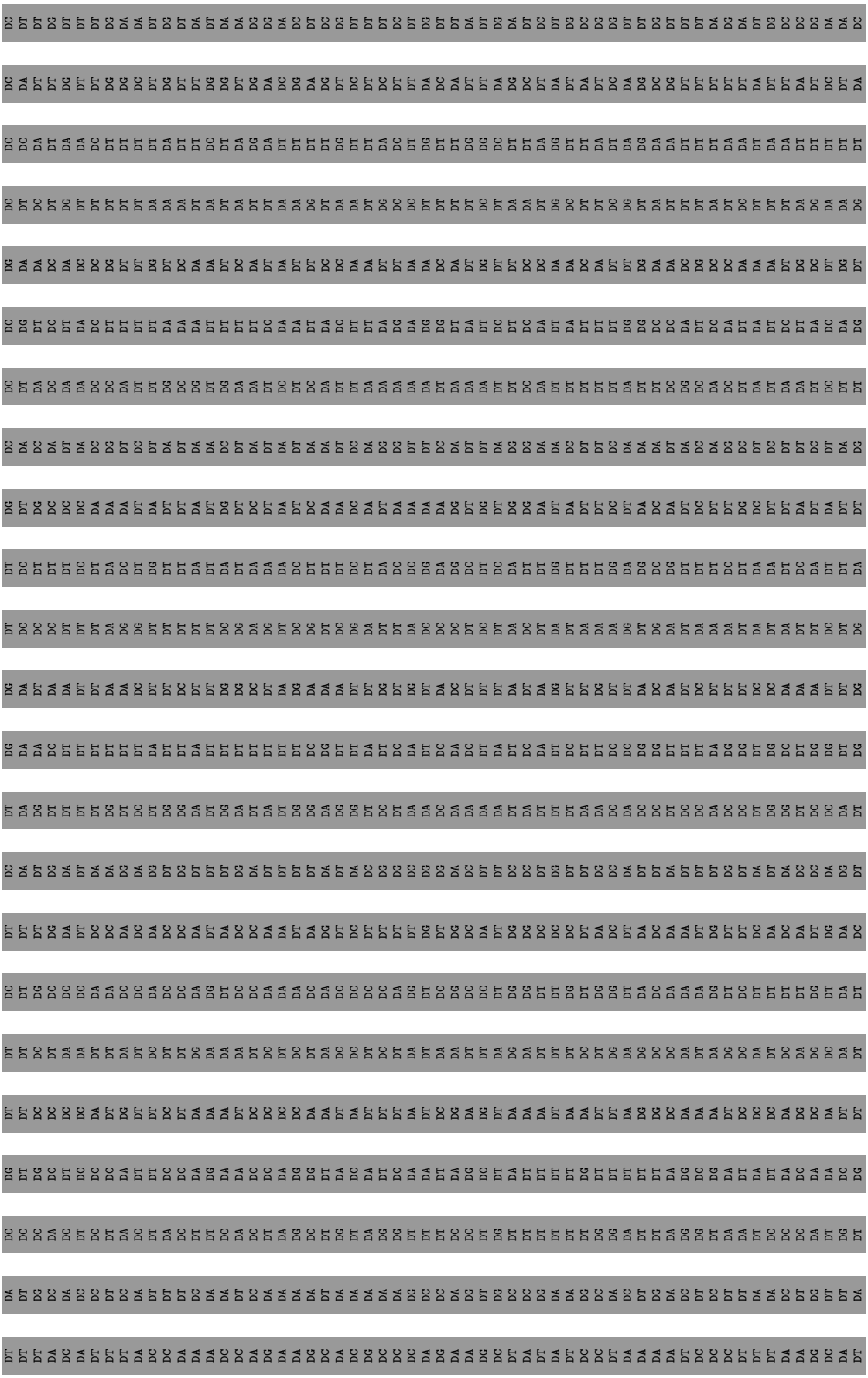
- Molecule 8: DNA, 3'-5'

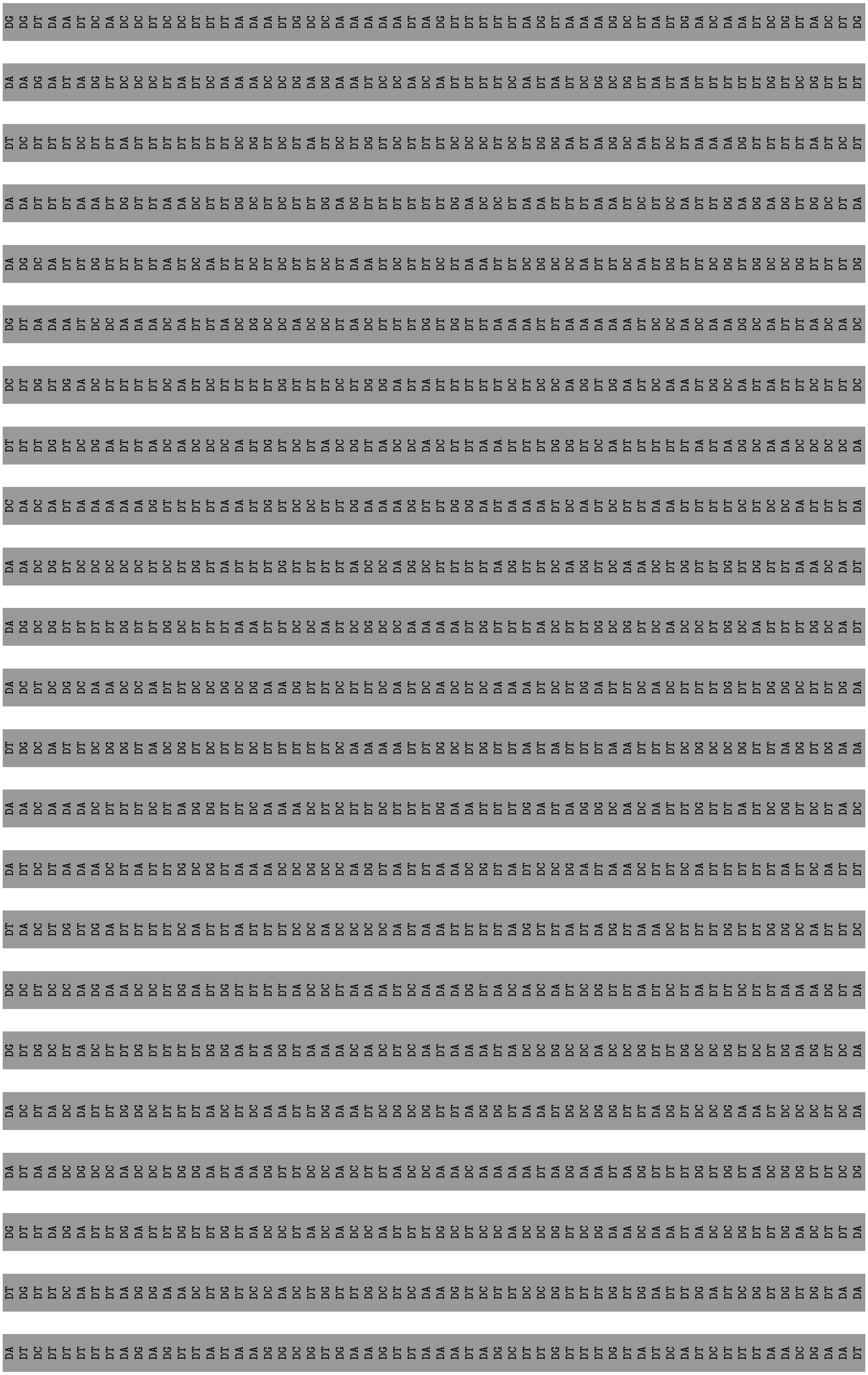
Chain R: 100%








DC DT DT DG DT DT DT DG DA DA DT DG DT DA DT DA DA DG DG DA DC DT DC DG DT DT DT DC DT DG DT DT DA DT DG DA DT DC DT DG DC DG DG DT DT DG DT DT DT DA DG DA DT DC DC DG DA DA DC
DC DA DT DT DG DT DT DG DG DC DT DG DT DT DG DG DT DG DA DC DT DG DA DG DT DC DT DT DC DT DT DA DC DA DT DT DA DT DA DG DC DT DA DT DC DA DG DC DG DT DT DT DT DA DT DT DA DT DC DT DA
DC DC DA DT DA DA DC DT DT DT DT DA DT DT DC DT DA DG DA DT DT DT DC DT DA DT DT DG DT DT DA DC DT DG DG DC DT DA DG DA DA DT DT DT DA DA DT DT DA DT DT DT DT DA DA DT DT DT DT
DC DT DC DT DG DT DT DT DT DT DA DA DT DA DA DT DT DA DT DT DA DA DG DT DT DA DC DC DT DT DT DC DT DA DA DT DT DG DC DT DA DT DT DT DT DA DT DT DA DT DG DA DA DG DA DA DG
DG DA DA DC DA DC DC DG DT DT DG DT DC DA DA DT DC DA DT DA DT DT DC DC DA DA DT DT DA DA DC DA DT DG DT DT DC DC DA DA DC DA DT DT DG DA DA DC DG DC DC DA DA DT DG DC DT DG DT
DC DG DT DC DT DA DC DT DT DT DT DA DA DT DT DT DT DC DA DA DT DT DA DA DG DG DT DA DT DC DT DC DA DT DA DT DT DT DT DG DG DC DC DA DT DC DC DA DT DC DT DA DC DA DG
DC DT DA DC DA DA DC DC DA DT DT DT DG DC DG DT DA DG DA DA DT DC DC DA DA DA DA DA DT DA DA DT DT DA DT DT DC DC DA DT DT DT DT DT DG DC DC DA DT DT DA DA DT DC DT DT
DC DA DC DA DT DA DC DG DT DC DT DA DT DA DA DT DA DT DA DT DT DC DA DA DG DG DA DA DC DT DT DC DA DA DG DC DT DT DC DA DG DC DT DT DA DA DT DC DT DT DA DG
DG DT DG DC DC DC DA DA DA DT DA DT DT DA DT DG DT DC DT DA DA DC DA DT DA DA DA DA DG DT DG DT DA DT DT DT DC DT DA DC DA DT DC DT DT DG DC DT DT DA DT DT
DT DC DT DT DT DC DT DA DC DT DG DT DT DA DT DA DA DT DA DC DT DT DT DC DT DA DC DC DG DA DG DC DT DC DC DA DT DT DG DT DT DG DA DG DC DG DT DT DT DC DT DA DA DT DC DA DT DT DA
DT DC DC DC DT DT DT DA DG DG DT DT DT DT DT DC DG DA DG DT DT DC DT DT DA DC DC DC DT DC DA DT DT DG DT DT DT DG DA DG DT DG DT DA DA DT DA DA DT DT DC DT DT DG
DG DA DT DA DA DT DT DA DA DC DT DT DC DT DA DA DG DG DC DT DA DA DT DG DT DG DT DA DC DT DT DT DA DT DT DG DT DT DA DC DA DT DC DT DT DG DC DT DT DA DA DA DT DT DG
DG DA DA DC DT DT DT DT DT DA DT DT DT DT DT DT DT DC DG DT DT DA DT DC DA DA DT DC DA DT DC DT DT DC DC DG DG DT DT DT DA DG DG DT DG DC DT DG DG DT DG
DT DA DG DT DT DT DT DC DT DG DG DG DA DT DG DA DT DA DT DT DT DG DG DG DA DA DG DG DT DC DT DT DA DC DA DA DA DA DT DA DT DT DT DC DC DA DA DC DC DT DC DA DC DC DT DG DG DT DA DT
DC DA DT DG DA DT DA DA DG DA DG DT DG DT DT DT DT DG DA DT DT DT DA DT DA DC DC DG DG DG DA DC DT DT DC DC DT DG DT DT DG DC DA DT DT DA DT DT DT DT DG DT DA DA DC DC DA DG DT
DT DT DT DG DA DT DC DC DA DC DC DC DA DT DA DC DC DA DA DT DG DT DC DT DT DT DT DG DT DG DC DA DT DT DC DC DA DC DC DC DC DT DA DC DA DA DT DG DG DT DT DT DA DG DG DT DA DC
DC DT DG DC DC DC DA DA DC DC DC DC DA DG DT DC DC DA DA DA DC DA DC DC DC DC DG DT DC DG DC DC DT DT DG DG DT DG DT DG DT DA DA DA DG DT DC DT DT DT DT DA DT DT DA DC DT DC
DT DT DC DT DA DA DT DT DA DT DT DC DT DT DC DT DC DT DC DT DC DC DA DA DT DA DA DG DA DT DT DT DT DG DG DA DT DT DG DA DA DT DC DC DA DA DG DC DC DA DA DT DC DC DA DA DT
DT DT DC DC DC DC DA DT DG DT DT DC DT DA DA DA DA DT DC DC DA DA DA DT DA DA DT DT DA DT DT DA DT DT DT DT DT DA DG DG DC DC DA DA DA DT DC DC DC DC DA DG DC DA DT DT DT
DG DT DG DC DT DT DC DC DC DA DT DT DC DC DA DG DA DA DC DC DA DC DA DG DG DT DA DC DA DA DA DT DA DG DC DT DT DT DT DT DT DT DA DG DC DG DA DA DT DA DA DC DC DA DG DC DG
DC DC DC DA DC DT DC DT DA DC DT DA DC DT DT DG DC DG DT DC DC DC DC DT DG DG DT DT DT DT DT DG DG DT DG DT DG DT DG DG DG DT DG DG DT DG DG DT DG DG DT DG DG DT DG DG DT
DA DT DG DC DA DC DC DT DC DA DT DT DC DT DA DA DA DA DT DC DC DA DA DA DA DA DA DG DC DC DA DA DG DC DC DA DA DG DC DC DA DA DG DC DC DA DA DG DC DC DA DA DG DC DC DA
DT DT DT DA DC DA DT DT DT DA DC DC DC DA DA DA DG DA DA DG DG DC DC DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA DA



DA DC DC DG DA DT DA DA DA DT DA DA DC DC DC DT DC DA DT DA DA DC DT DT DT DC DA DA DC DG DC DT DA DC DC DT DG DG DT DA DA DA DT DC DC DG DG DC DA DC DT DC DT DT DG DT DT DG DC DA

DT DA DG DT DA DA DT DT DA DC DC DA DG DC DA DG DT DT DA DA DA DT DA DT DC DC DC DA DA DA DT DC DG DC DC DT DT DG DC DG DC DA DT DT DA DT DT DT DA DA DG DT DT DA DA DC DT DT DG DA

DA DT DT DG DA DT DT DG DA DC DC DA DT DT DC DG DC DC DT DC DT DG DT DC DA DT DC DT DT DA DT DG DT DT DG DT DT DG DC DC DA DG DC DT DC DG DT DT DG DT DT DC DC DG DA DA DT DT DT DA

DT DC DA DT DC DT DA DC DA DT DA DC DT DG DC DT DT DA DG DC DT DT DG DA DT DT DT DA DA DA DG DC DG DT DT DG DT DT DA DG DA DC DG DT DT DT DC DT DT DC DA DA DC DA DA DA DT DT DG DC

DT DT DA DG DT DT DA DA DG DT DC DA DC DC DG DT DC DA DT DT DC DT DT DT DT DT DA DT DA DA DA DA DC DG DG DG DT DA DC DC DA DT DG DT DG DC DC DA DC DT DA DA DT DT DT DT DA DT DA DT

DT DT DT DG DT DA DT DA DT DT DC DG DT DC DG DT DT DT DG DA DA DT DC DA DT DC DT DG DG DA DT DA DC DC DA DT DG DT DT DG DC DA DC DG DT DG DC DC DG DT DA DC DT DA DT DC DA DT DC DA

DA DC DA DA DC DA DT DA DG DA DC DA DA DC DT DA DA DC DA DC DG DC DC DT DG DA DT DT DT DT DC DC DT DA DA DA DG DT DG DT DT DA DG DG DA DG DC DT DA DC DC DG DG DA DA DT DA DT DC DT

DG DA DA DC DC DA DT DC DG DT DC DA DA DC DG DC DC DA DT DC DT DT DC DT DT DT DA DG DG DT DG DT DA DT DC DG DA DC DA DG DT DA DC DC DT DA DT DA DT DC DT DT DT DA DA DA DT DG DA DG

DG DG DC DG DC DA DT DC DT DG DT DC DG DC DG DC DT DA DG DT DG DA DT DA DT DG DA DA DT DA DA DT DC DC DT DA DG DA DT DG DT DG DT DT DA DA DC DT DG DC DG DC DT DT DA DA DA DA DC DG

DC DT DA DT DC DT DA DT DG DG DA DC DT DG DC DT DC DA DT DA DC DG DA DT DT DC DA DA DT DT DG DC DT DT DT DA DC DC DG DT DA DA DT DC DA DT DC DT DG DT DA DA DG DT DT DT DA DG DA DC

DT DT DT DT DG DC DC DA DA DT DT DC DG DT DT DG DT DT DG DA DA DT DT DA DC DC DT DT DT DA DA DC DA DA DG DG DT DC DA DG DC DG DC DC DA DT DT DG DA DT DT DT DG DT DT DG DT DT DC DA

DA DC DT DT DC DG DT DT DA DA DC DA DC DG DT DT DC DA DA DA DA DA DT DC DG DC DT DT DG DC DT DC DT DT DT DT DT DC DA DA DC DT DA DT DT DT DT DA DT DC DG DA DA DT DT DC DA DG DC DT

DG DT DA DA DC DA DG DC DT DT DG DT DG DT DT DG DC DA DC DT DA DG DT DT DT DG DC DG DT DC DG DC DA DG DT DA DA DT DA DG DC DT DT DG DT DA DT DA DG DC DT DT DC DG DT DT DT DT DG DC

DT DT DG DA DT DT DT DC DG DA DT DT DT DG DT DT DG DA DA DT DG DC DC DT DT DT DT DG DT DC DG DC DA DC DT DA DT DC DA DT DT DC DA DC DT DT DT DT DG DC DT DA DT DT DA DA DC DG DT DT

DT DG DT DG DT DA DT DC DA DG DC DC DA DT DA DT DT DT DT DG DC DT DT DT DA DA DT DT DG DG DT DT DA DA DA DG DT DC DT DT DT DA DC DC DG DA DC DA DG DC DT DT DC DG DA DT DA DG DT DA

DT DC DT DT DG DA DA DT DA DG DA DT DT DT DG DA DT DA DT DA DA DA DC DA DA DG DC DT DT DT DG DT DT DA DT DA DC DC DA DT DC DA DA DA DC DC DC DA DC DT DA DA DC DT DA DA DA DT DC DA

DT DT DT DT DC DA DA DT DA DT DT DG DA DA DG DC DT DA DA DA DT DT DG DA DC DG DT DT DC DA DA DC DA DA DC DA DA DC DA DT DT DA DT DC DA DC DT DC DC DC DG DT DT DT DT DG DT DG DT DA

DA DA DG DA DA DT DG DC DC DT DG DA DG DC DA DT DG DC DA DC DC DT DT DG DC DC DT DG DA DA DT DG DT DT DT DT DA DA DA DA DT DT DC DA DT DT DC DG DG DT DA DT DC DA DC DA DT DA DC

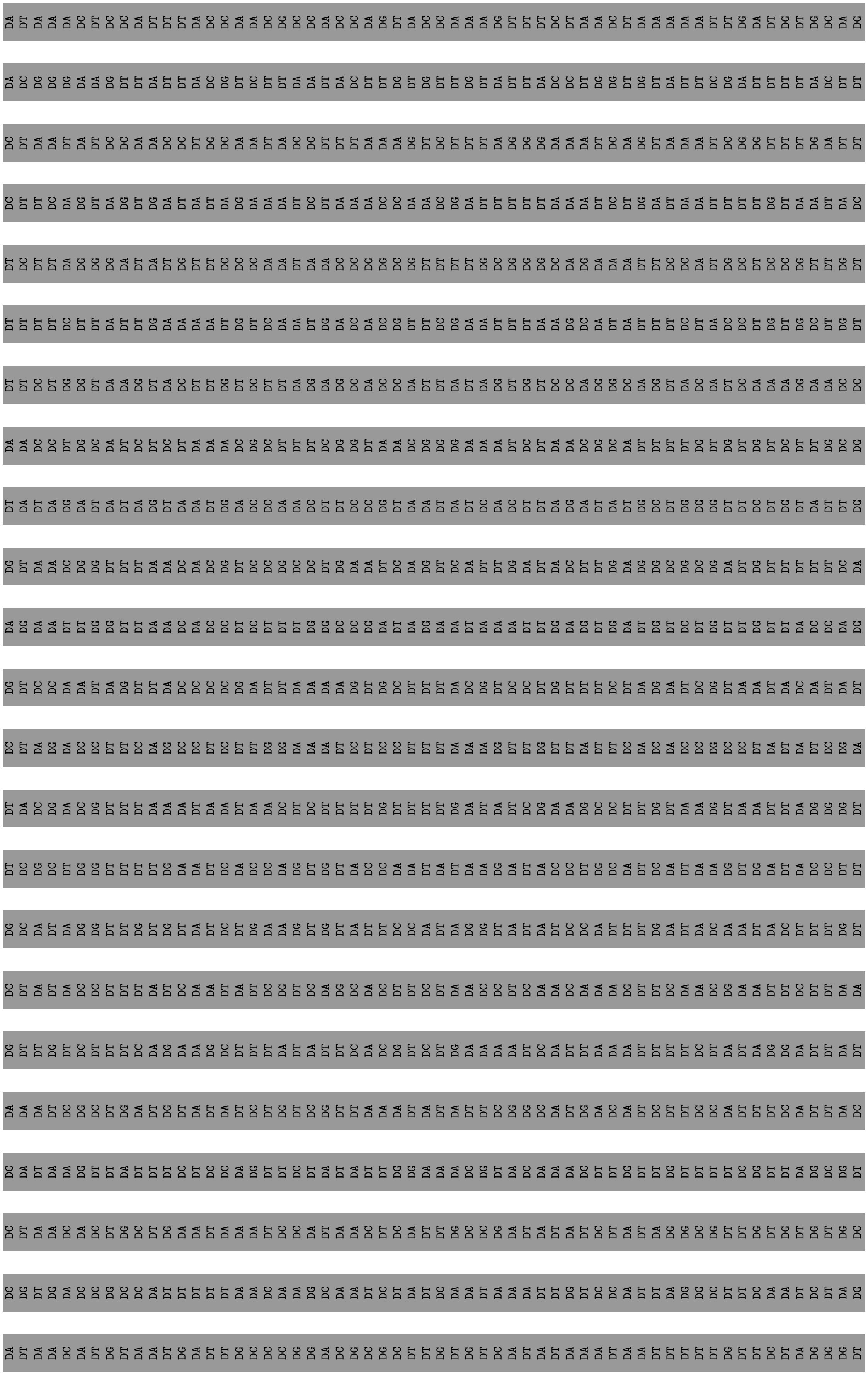
DT DG DC DA DA DA DC DG DC DC DC DA DT DT DA DA DT DT DG DC DG DT DC DT DA DC DT DA DT DC DG DT DT DA DA DT DT DC DG DT DC DT DG DA DA DA DT DA DT DA DA DG DC DG DC DC DT DC DT DA

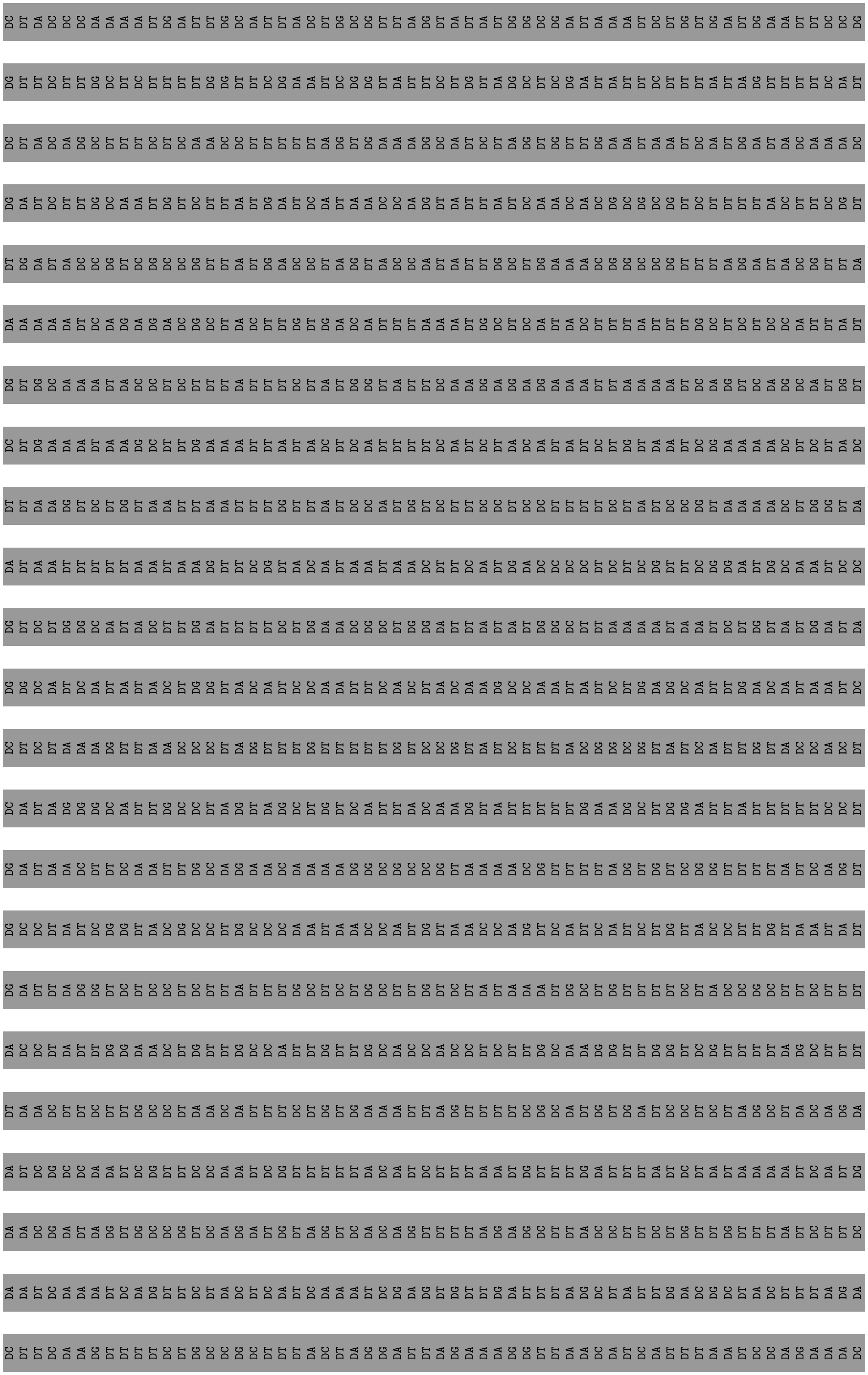
DT DC DT DA DC DG DT DT DA DT DA DA DT DC DA DT DC DG DG DT DT DT DT DT DA DA DC DA DC DG DA DT DA DA DA DT DG DT DT DT DT DA DA DC DA DT DG DT DT DC DA DG DA DA DC DT DT DA DT DA

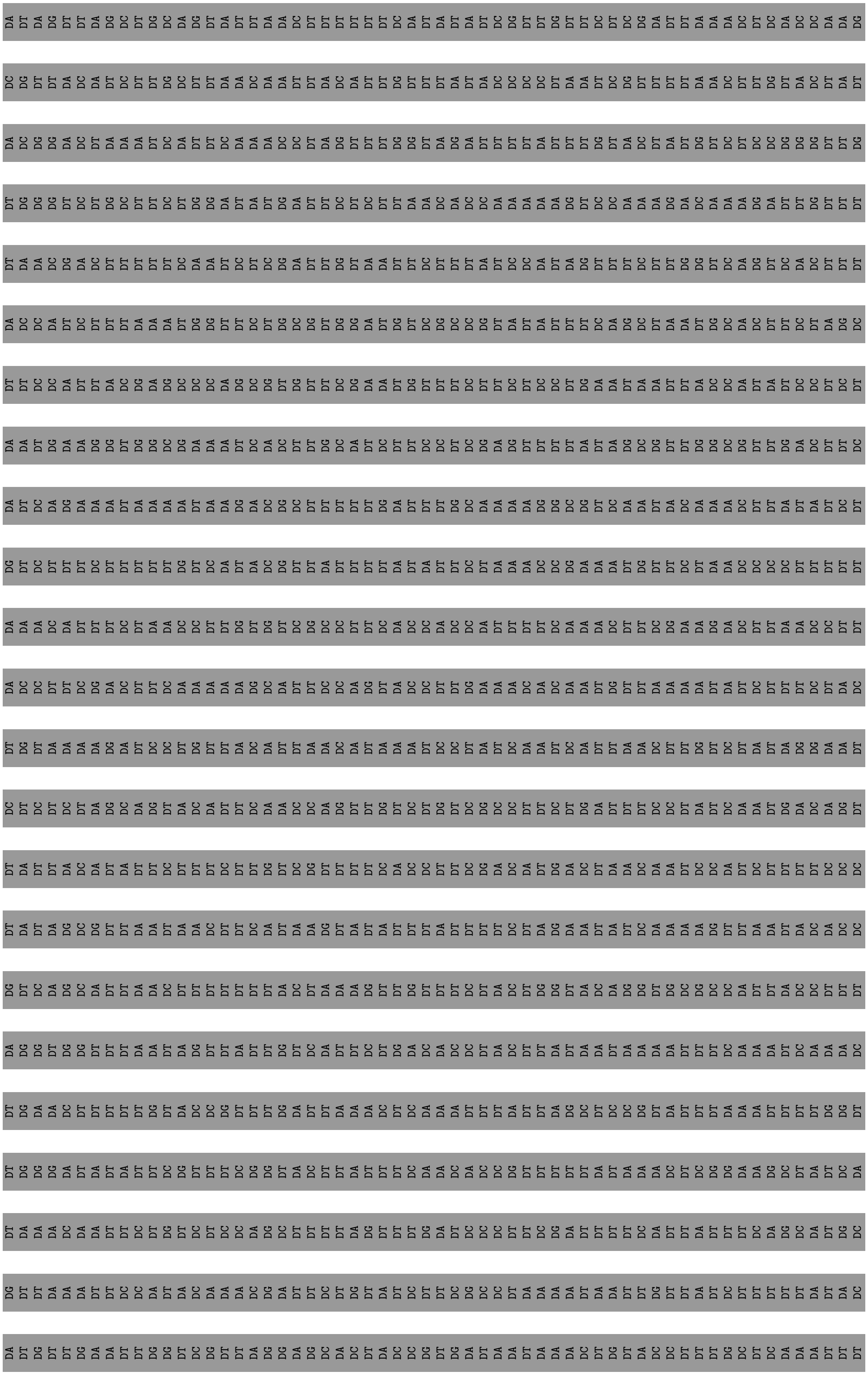
DG DA DT DA DA DC DG DG DT DC DT DG DT DT DA DT DT DC DT DT DA DG DT DT DA DC DT DG DC DA DA DA DA DT DT DT DA DA DA DA DC DA DC DC DA DG DT DT DC DC DT DC DT DA DT DC DT DG DA DT

DT DC DA DT DA DG DA DA DA DC DT DG DA DT DG DT DT DT DG DT DG DT DC DA DA DT DA DA DC DC DG DG DA DT DT DA DT DA DT DT DG DT DG DA DT DG DT DT DG DT DT DT DG DT DA DA DC DT DC DG

DA DT DT DA DA DG DT DT DA DT DC DA DT DC DT DT DT DC DG DA DA DA DA DA DT DT DA DT DC DT DA DC DT DA DC DC DA DT DT DA DT DT DC DA DA DC DC DA DC DC DT DT DT DC DC DT DT DC DG DA







DT DC DA DA DA DT DG DT DA DT DA DC DT DT DT DC DT DT DT DT DG DT DC DT DC DC DG DC DC DT DG DC DA DT DG DC DC DA DA DT DC DA DA DT DA DT DT DA DA DA DA DT DC DA DA DA DC DG DA DC

DC DA DC DA DA DT DC DC DA DA DC DC DT DT DT DT DT DG DT DT DT DT DC DT DT DC DT DA DA DC DT DC DT DA DG DG DC DT DT DG DT DT DC DC DA DA DT DA DC DT DT DT DC DA DC DC DG DT DA DT

DG DG DT DA DA DT DT DC DT DG DT DA DG DT DT DT DC DG DA DA DT DT DT DT DA DG DT DT DC DA DA DA DA DG DA DA DA DC DT DT DT DA DT DT DA DC DC DT DT DT DT DT DG DT DT DT DA DG DG DG

DT DT DT DA DT DA DA DC DT DC DC DG DT DT DA DA DA DA DA DT DA DA DC DT DT DT DA DT DA DC DT DG DT DT DT DA DC DC DA DT DT DT DA DC DA DT DA DA DA DT DT DT DG DT DT DG DA DT DC DG

DT DG DT DC DT DT DG DA DA DT DA DT DT DC DG DT DA DA DT DC DC DG DG DG DA DA DG DT DT DG DT DT DT DT DT DA DT DC DT DA DA DT DT DT DC DA DC DG DT DA DA DT DC DA DT DC DA DG DA DA

DG DT DT DG DG DT DT DG DA DG DT DA DA DA DC DC DT DG DT DA DA DT DT DC DA DA DC DT DC DT DT DC DT DT DT DT DC DT DT DC DT DG DA DT DT DT DC DT DC DG DC DA DA DA DT DA DC DA DT DA

DG DG DT DT DC DT DA DT DG DT DC DT DG DT DC DG DT DT DA DA DC DG DA DA DT DA DC DA DA DC DA DT DA DT DC DT DC DG DC DA DT DA DT DA DA DG DC DA DA DT DG DT DC DT DG DA DA DC DG DA

DT DT DT DT DT DA DA DC DT DT DT DA DC DA DA DA DA DA DC DA DA DG DG DA DA DC DA DA DC DT DA DT DA DT DC DT DC DT DA DC DT DG DA DT DA DT DA DA DT DT DG DC DT DC DC DC DC DA DT DT

DA DA DT DA DT DA DC DG DA DC DC DG DT DT DC DA DT DA DT DT DT DT DC DT DT DT DG DT DC DT DT DG DA DT DA DC DT DT DT DG DT DG DT DT DG DA DT DT DT DG DC DA DT DG DC DT DA DT DC DA

DA DT DT DA DC DT DA DT DA DT DC DG DT DT DA DA DC DG DA DT DA DA DA DC DC DC DG DT DA DT DT DC DA DC DT DT DA DA DT DT DT DG DA DT DT DA DC DA DG DT DA DC DC DA DT DC DT DT DT DT

DT DT DT DG DT DT DA DA DT DT DC DT DA DT DG DT DC DC DA DT DT DT DG DT DA DA DC DC DT DC DC DT DT DT DA DT DA DA DG DT DA DA DT DA DC DT DC DA DG DA DA DT DT DG DC DG DT DT DT DA

DG DC DA DT DT DT DC DT DG DC DC DG DT DT DA DA DC DA DA DT DA DC DT DA DG DT DA DA DG DC DG DC DA DT DC DG DT DT DA DT DT DG DA DC DA DT DC DG DA DA DT DT DC DA DA DC DT DT DT DA

DA DC DA DG DT DT DT DT DC DA DT DG DT DT DC DG DG DT DG DA DT DG DT DT DT DC DA DA DT DA DG DA DA DT DG DT DG DT DG DT DG DT DT DG DT DA DC DT DT DG DC DG DC DA DT DT DT DA DT DA

DT DT DT DC DC DA DC DC DT DA DA DA DT DT DA DC DT DT DA DA DG DT DT DT DC DC DT DG DT DA DA DT DA DC DT DA DG DA DA DA DT DG DT DC DA DG DG DT DG DC DG DT DT DT DA DA DT DG DT DA

DG DG DT DT DG DA DA DA DT DG DC DA DT DC DA DA DC DT DA DC DT DT DT DA DT DC DT DG DC DA DA DC DA DT DT DA DG DA DA DA DC DA DT DT DA DC DG DG DA DT DA DA DC DT DT DT DA DC DT DT

DG DA DA DT DG DA DT DT DA DT DC DT DA DT DA DC DC DT DT DT DA DA DC DG DA DA DA DC DC DT DA DG DC DA DT DT DG DA DA DT DA DC DA DT DA DC DC DA DA DC DA DT DC DC DA DT DG DA DA DT

DT DC DA DC DG DT DG DA DA DG DG DT DG DA DG DT DG DA DA DT DA DC DC DC DA DA DA DG DC DA DC DT DT DT DT DA DG DC DT DG DC DA DT DC DT DA DA DA DG DC DT DT DT DC DT DT DA DG DC DA

DA DC DA DT DT DT DT DT DA DG DC DT DG DC DA DT DC DT DA DC DT DA DA DT DT DG DA DC DC DA DG DC DC DA DT DT DT DG DT DC DC DA DA DT DA DC DC DT DC DT DA DA DT DT DA DA DA DC DC DA

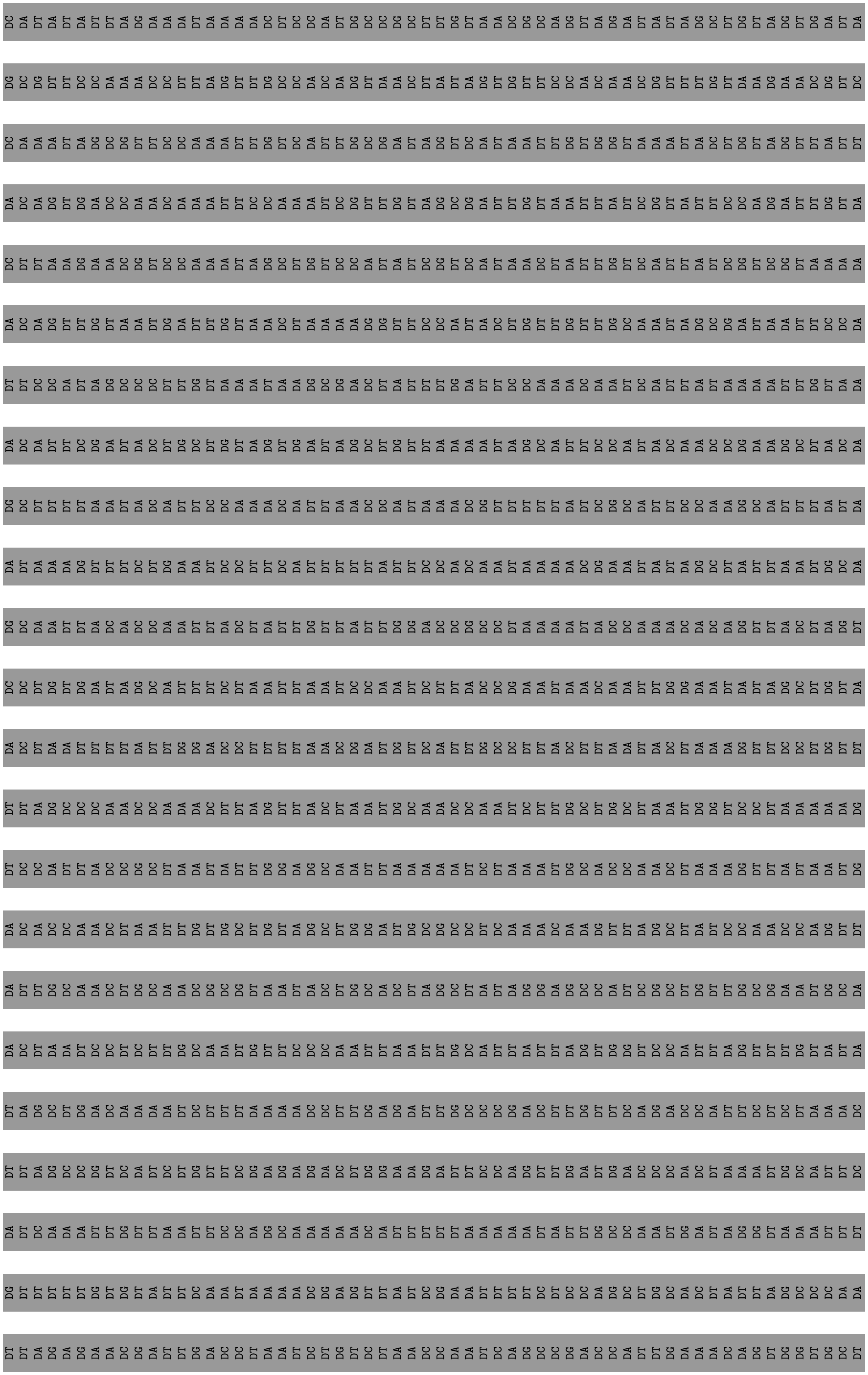
DC DG DG DA DT DC DA DT DA DT DC DA DG DC DA DC DC DT DG DC DA DG DA DC DA DC DA DA DA DA DT DC DT DC DC DT DA DT DA DA DA DG DC DT DT DT DT DT DA DT DT DT DT DA DT DT DT DA DC DT

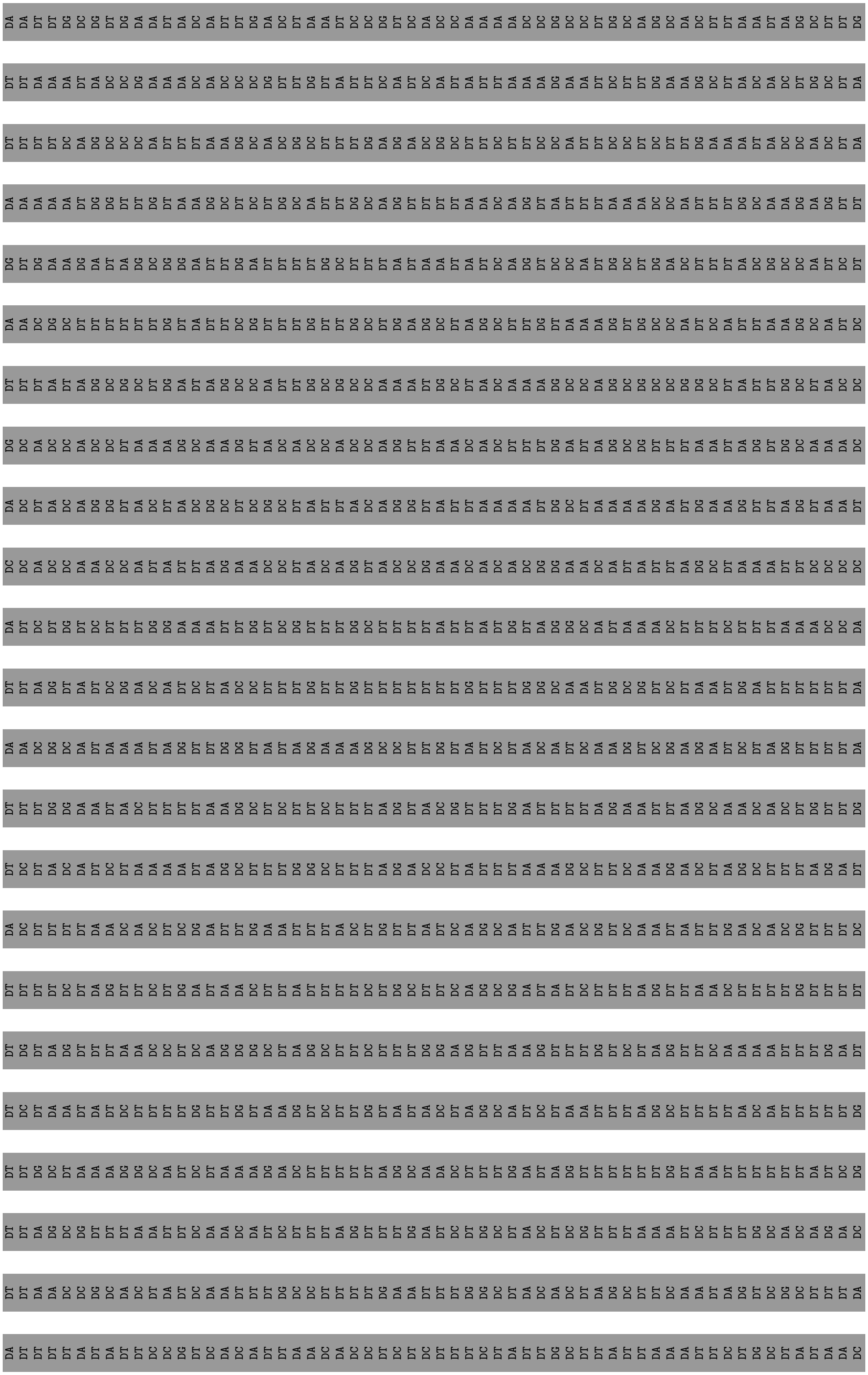
DG DC DA DT DT DT DG DT DC DA DT DA DC DC DT DT DG DA DC DT DA DA DC DT DT DT DG DT DT DT DA DC DA DA DC DA DT DT DA DA DC DG DA DA DT DC DC DT DT DG DA DA DT DA DA DC DT DC DT DA

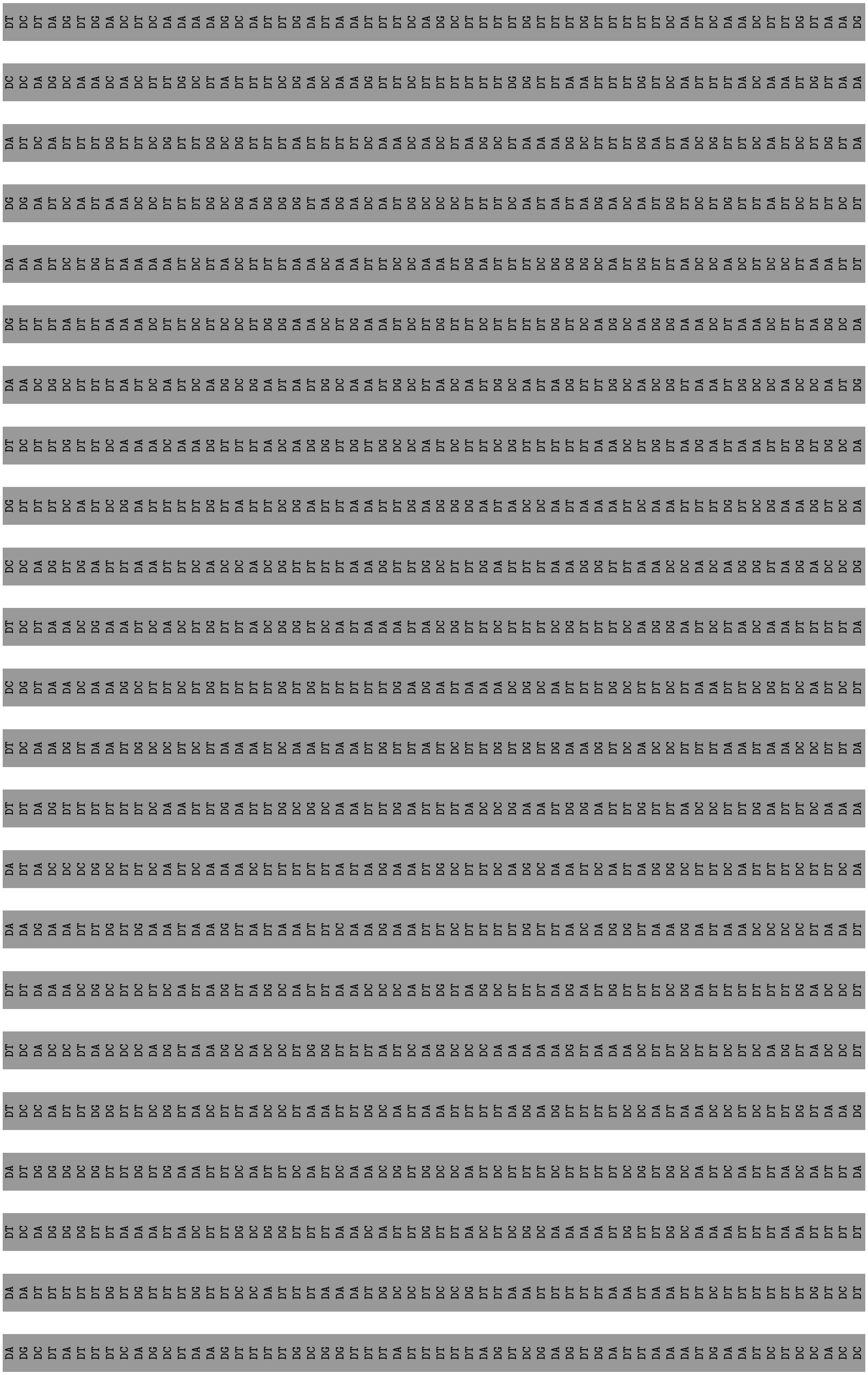
DT DT DA DA DC DA DA DA DG DT DT DA DA DT DT DA DG DC DG DT DA DC DT DT DG DT DT DA DT DA DG DT DA DG DA DT DA DC DC DC DA DT DT DG DC DA DT DA DC DC DT DT DT DA DG DT DG DA DC DA

DA DT DG DA DA DG DT DT DC DC DA DA DG DC DT DT DG DA DG DA DC DA DT DT DT DT DG DT DC DT DG DA DT DA DT DA DG DT DT DG DA DA DA DC DA DA DC DT DT DG DT DG DT DG DA DA DT DA DT DG

DC DT DT DA DC DA DA DC DT DT DT DA DT DT DC DC DA DA DA DT DT DG DT DC DG DT DT DA DA DT DA DT DA DC DC DA DG DA DT DA DA DG DA DA DA DC DT DC DC DA DA DA DT DC DG DT DA DT DT DC







DA DG DT DC DA DA DT DG DG DC DG DA DT DT DG DT DC DT DA DG DC DG DT DT DA DA DC DC DT DT DA DA DC DC DG DC DT DT DC DA DT DT DA DA DC DC DG DC DT DT DT DT DT DT DT DA DC DT DG DC

DA DT DT DA DG DA DA DA DA DA DG DC DT DT DC DA DA DC DG DT DT DC DG DA DT DT DT DG DG DT DT DT DG DT DT DC DA DG DC DA DG DT DG DT DC DA DG DT DT DA DC DA DA DC DC DA DA DA DT DT

DG DA DC DA DA DC DC DT DC DA DT DC DT DG DA DT DG DA DA DT DC DA DA DC DT DT DC DC DG DC DT DT DC DA DC DT DC DA DA DC DA DT DT DT DT DT DC DG DT DG DC DT DT DC DT DG DA DA DC DG

DC DA DT DT DT DC DG DT DT DT DA DA DT DT DG DT DT DT DT DT DC DT DG DA DA DC DG DT DA DA DT DT DG DT DT DC DC DA DG DC DT DC DT DT DT DT DT DC DC DA DT DT DT DG DT DT DC DG DC DG

DT DT DC DA DT DA DT DT DC DA DT DC DT DT DT DT DT DG DA DT DC DC DT DT DG DT DT DC DA DT DT DT DT DC DG DC DT DA DA DT DT DT DA DG DC DA DG DC DT DT DC DT DT DT DA DG DC DA DG DC

DT DT DC DT DG DC DT DG DC DT DT DT DT DT DC DT DT DT DT DG DC DG DT DA DC DT DC DA DT DC DA DG DC DT DT DT DT DT DC DT DT DT DT DC DG DT DG DG DG DC DT DA DC DA DC DG DA DC DG

DT DT DC DA DA DG DT DA DT DT DT DC DA DT DC DA DA DC DT DT DT DC DT DT DT DG DT DT DG DC DT DC DT DG DG DC DG DT DG DA DA DG DG DT DT DA DT DT DT DC DA DG DT DA DC DC DT DT DC

DG DT DC DA DT DT DT DT DC DT DT DT DC DT DT DA DT DC DA DG DG DA DT DT DC DC DC DT DT DT DT DT DT DA DC DC DA DT DC DT DC DC DG DC DC DT DG DG DT DT DC DG DT DC DC DG DG DA DT DC

DA DT DC DT DG DA DT DT DG DG DT DC DT DG DC DA DA DA DA DA DA DT DT DG DC DA DA DA DT DT DA DA DA DC DT DT DA DA DG DT DT DT DA DT DT DT DT DC DT DT DC DC DA DT DG DA DG DA DT DA

DT DA DC DC DT DC DC DA DT DT DT DA DT DA DG DT DC DT DG DT DC DG DA DC DT DG DT DT DT DT DT DC DC DA DT DG DC DG DT DG DC DT DT DT DT DT DA DT DG DT DC DA DT DC DA DG DC DA DC DG

DT DT DT DT DG DG DA DC DA DT DA DA DA DA DA DA DT DA DG DC DC DA DA DC DA DC DA DA DT DT DA DA DG DT DG DC DT DA DG DC DT DA DT DT DA DA DA DA DG DA DG DT DG DG DT DT DC DG DT DT

DA DT DA DT DT DT DC DG DA DT DT DT DT DT DC DT DT DT DA DT DT DG DG DC DT DA DA DT DA DC DT DG DC DC DG DA DC DC DT DT DA DC DG DC DT DG DT DC DT DA DA DG DT DT DT DG DC DA DT DC

DA DA DT DA DA DT DA DA DC DT DG DT DT DT DC DG DT DT DT DT DG DC DT DT DT DT DG DT DA DA DC DT DC DT DT DT DA DC DG DT DA DT DA DC DC DT DT DT DT DA DA DC DT DC DT DC DT DT DG DC

DT DA DT DG DT DC DT DC DT DA DA DG DG DT DA DT DT DT DG DT DC DA DG DT DA DT DT DA DC DT DC DA DT DT DT DT DA DG DT DA DT DC DC DT DC DC DA DA DA DC DA DT DT DT DA DA DT DT DT DA

DC DT DG DT DC DA DT DA DC DA DA DA DA DC DT DA DA DC DT DT DG DC DC DT DT DT DA DA DA DA DA DC DT DT DT DA DC DT DT DT DT DA DA DA DT DC DA DA DT DC DA DC DC DG DC DT DT DT DT

DC DA DC DT DT DT DC DC DC DT DC DC DG DA DA DG DT DA DT DT DT DT DG DT DT DT DT DT DC DG DT DT DT DC DT DT DG DC DT DT DG DG DT DT DT DT DT DT DC DG DG DC DC DA DC DA DT DA DG DA

DT DT DT DA DG DG DT DA DG DT DA DA DA DG DC DG DC DA DA DT DC DT DG DA DA DC DG DA DC DA DA DT DT DG DA DT DA DT DG DC DA DT DA DG DG DA DT DA DG DA DA DA DT DT DA DA DC DA DC DC

DA DA DT DT DT DT DA DG DC DG DT DC DT DT DT DA DA DC DT DT DT DG DA DA DT DA DT DT DT DC DT DC DC DA DT DT DA DA DG DC DC DC DT DT DT DG DC DA DT DA DC DT DT DT DA DG DT DT DG DT

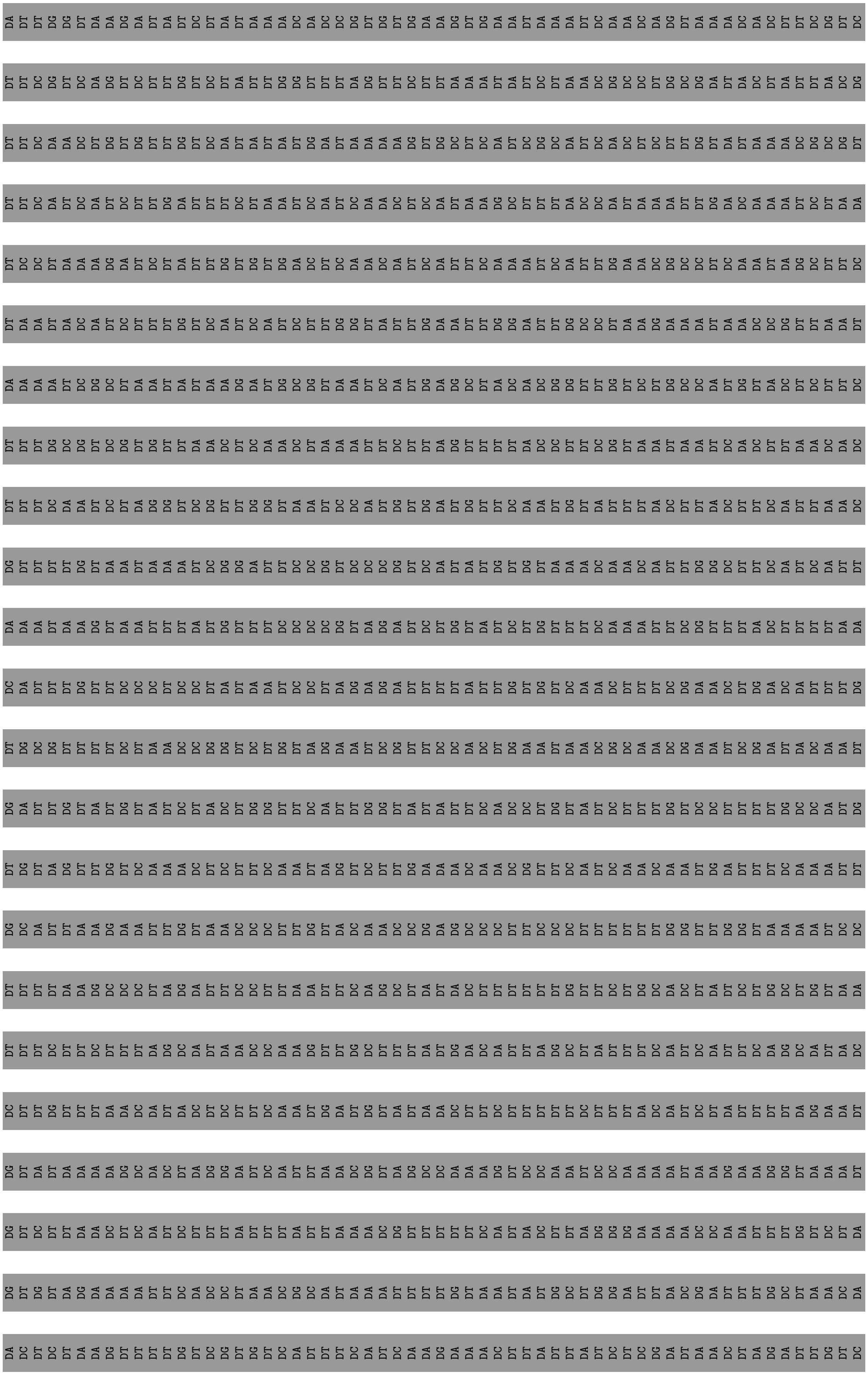
DT DC DT DA DT DT DA DT DC DG DA DT DT DT DT DT DG DC DA DA DT DA DT DA DC DA DT DA DT DA DA DT DA DT DC DC DT DT DC DC DG DG DT DG DA DA DA DT DT DT DC DT DT DT DC DA DT DG DC DT

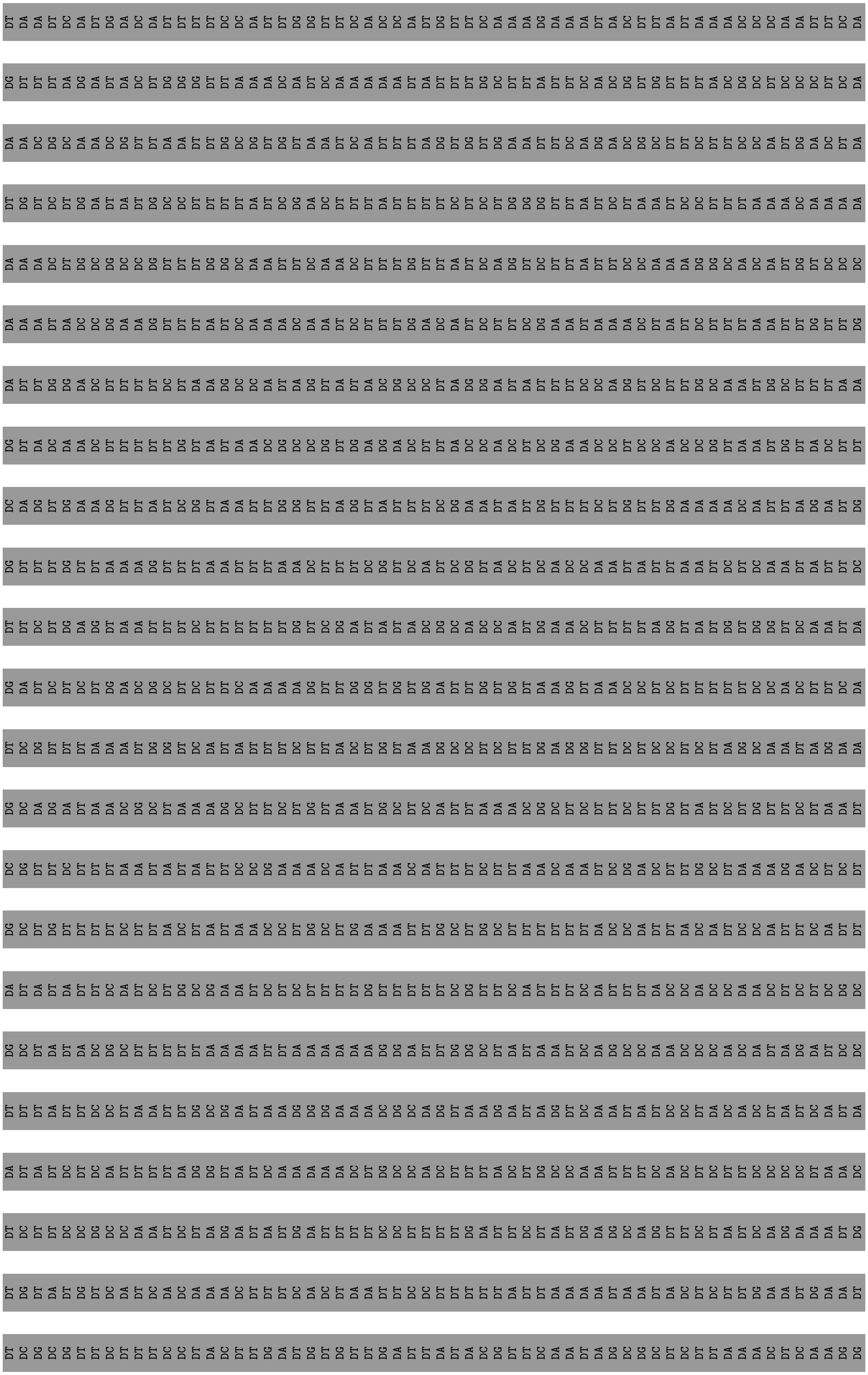
DG DT DC DA DA DT DG DC DT DT DG DA DT DT DG DT DG DC DG DT DG DA DA DC DA DC DG DT DG DC DC DG DA DT DT DC DC DG DT DA DT DA DA DA DG DC DA DA DT DG DA DT DT DT DA DA DT DT DG DC

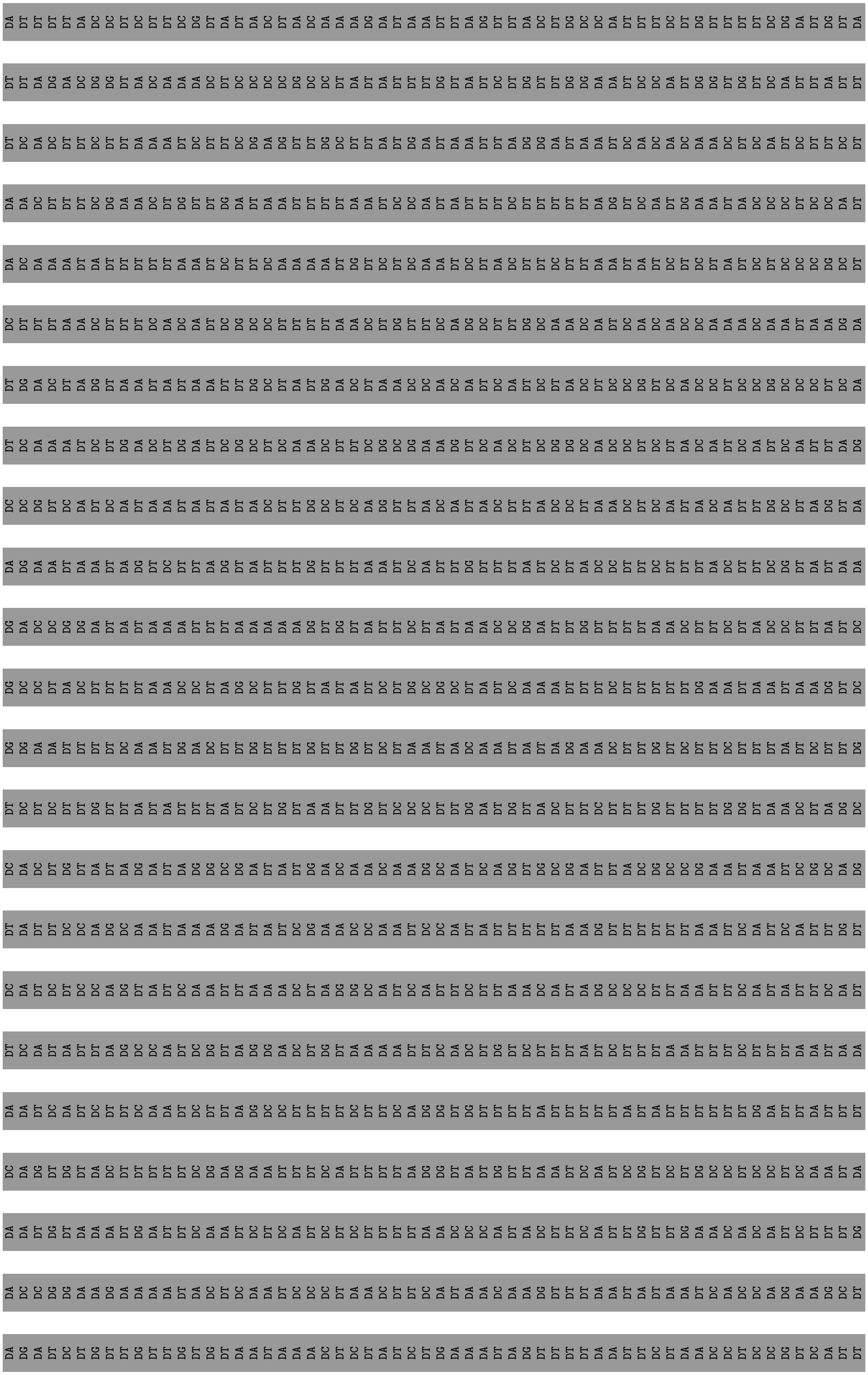
DT DG DC DG DG DT DC DT DT DT DT DG DT DC DG DT DG DC DT DG DT DT DC DC DT DT DC DG DA DA DT DT DT DA DT DT DT DA DA DG DT DG DT DT DT DA DC DG DC DA DT DA DT DC DT DT DT DA DA DC

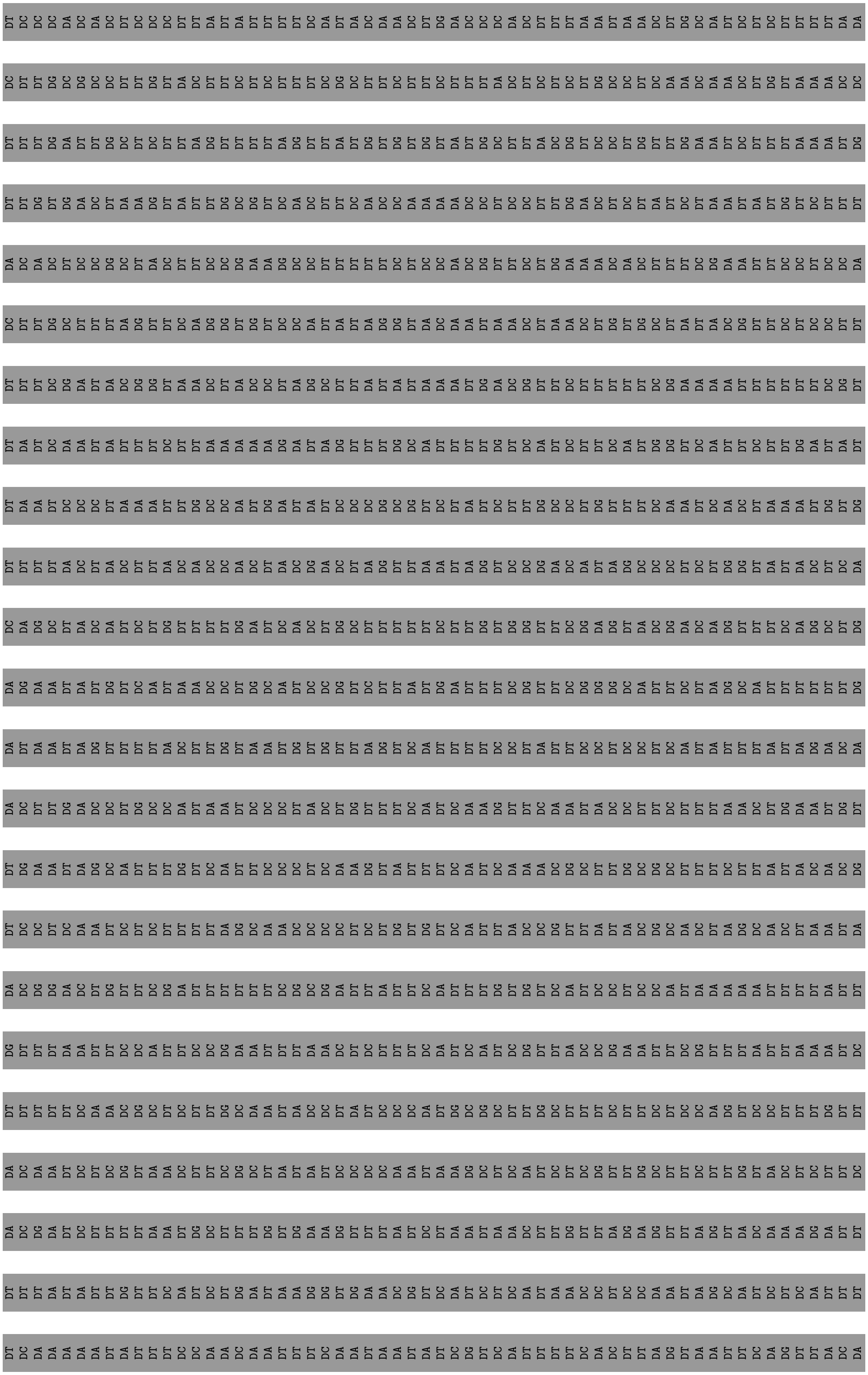
DG DT DA DT DT DC DA DT DT DA DG DG DA DT DG DT DC DG DA DC DC DT DC DT DA DA DT DA DA DC DC DA DC DA DT DT DA DG DC DA DA DT DT DA DT DT DT DC DT DT DC DT DA DC DT DT DC DT DT DG

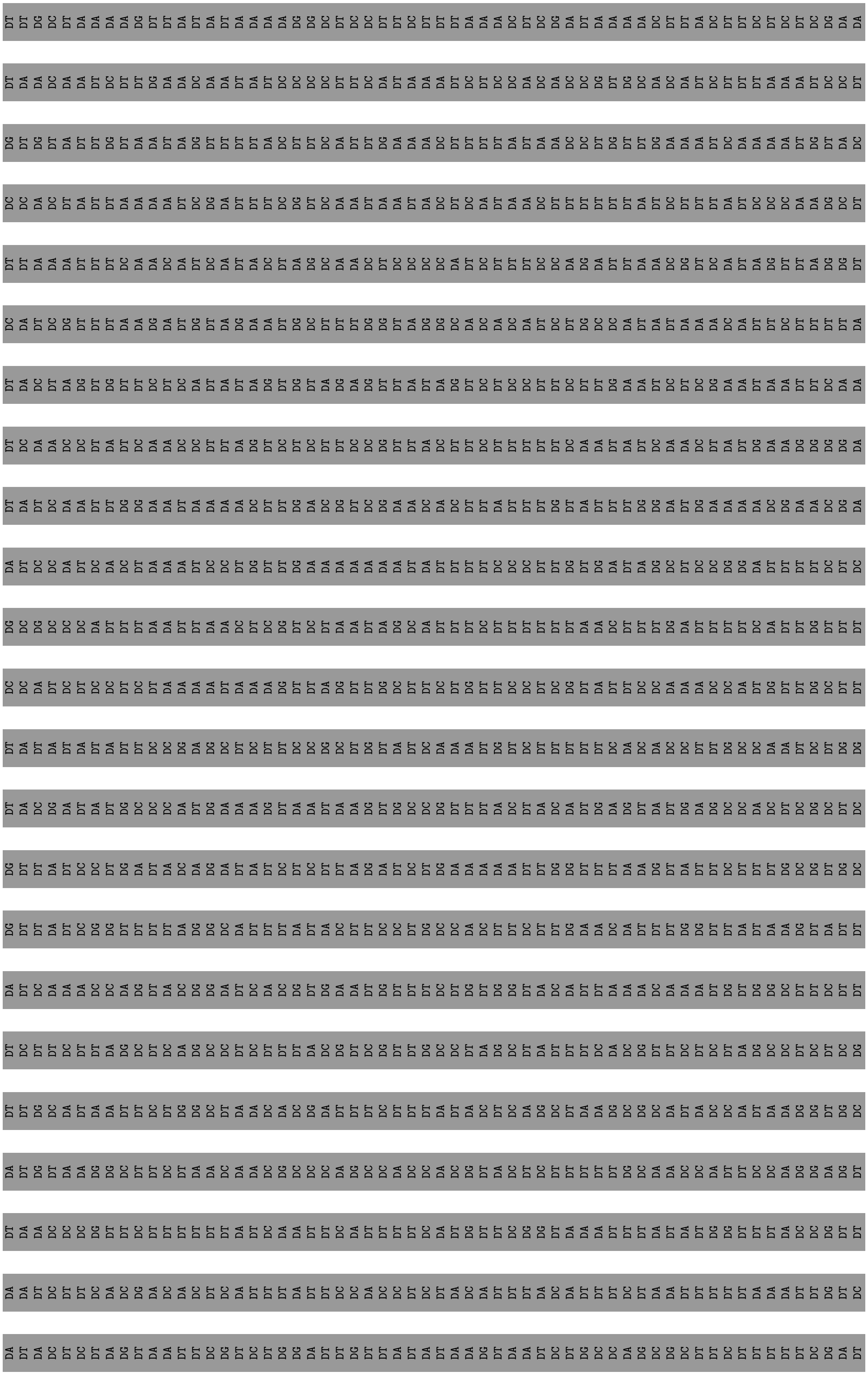
DT DT DT DC DA DT DC DG DC DT DT DC DA DG DT DA DT DT DA DG DT DC DC DA DT DA DA DT DC DG DC DT DC DT DG DA DC DC DA DA DA DC DG DA DC DA DC DC DA DT DG DA DA DA DT DT DG DT DG DT



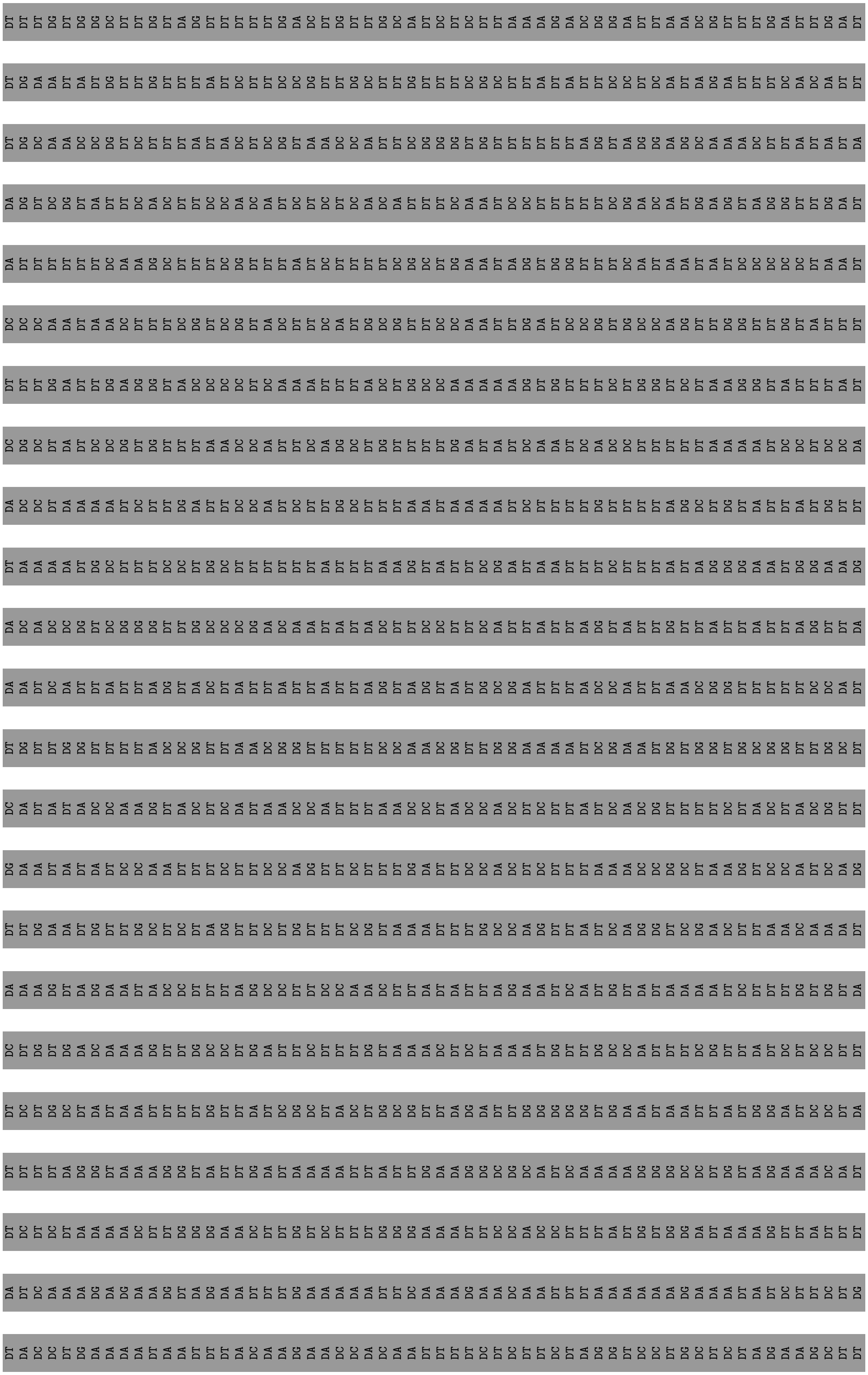


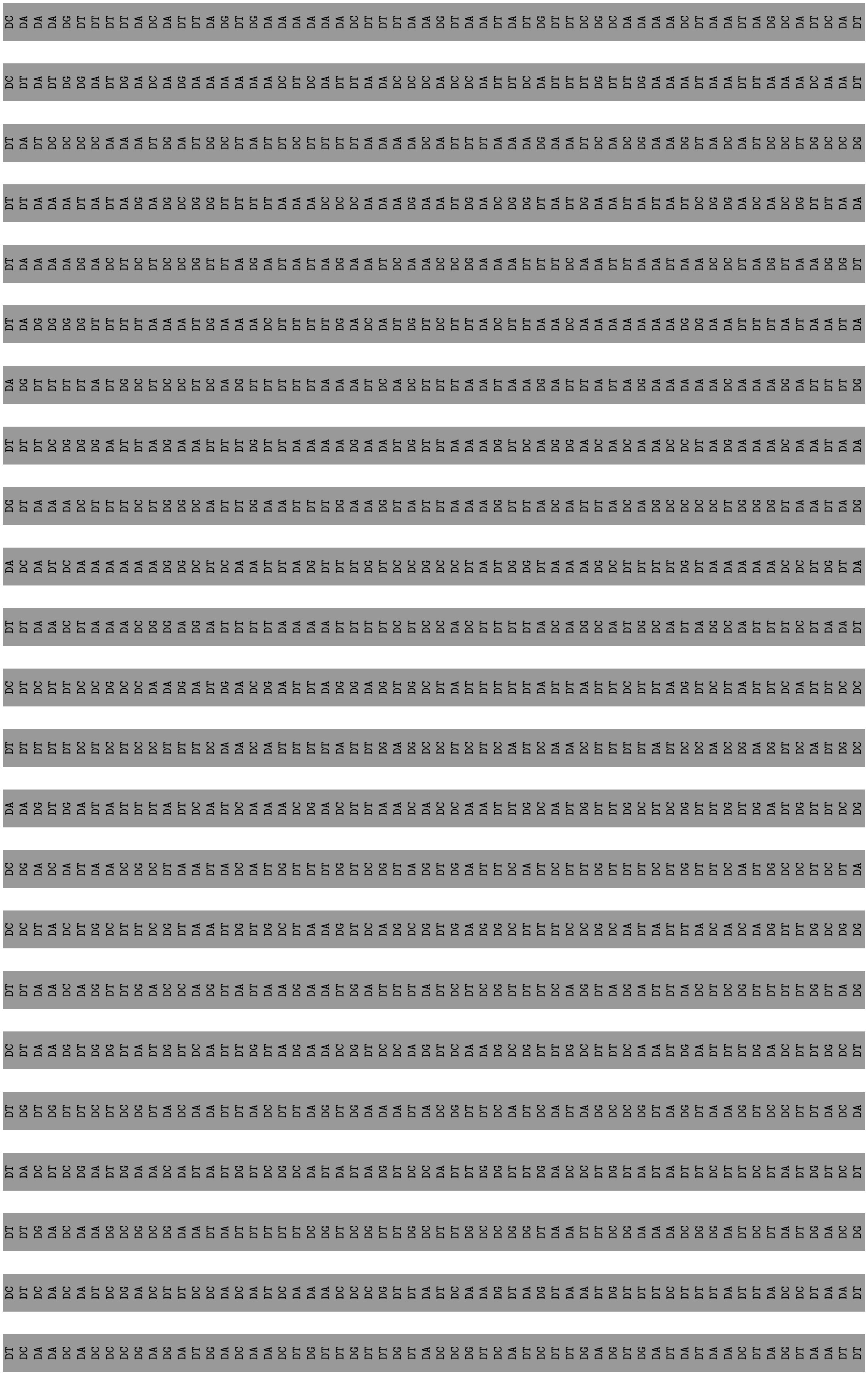


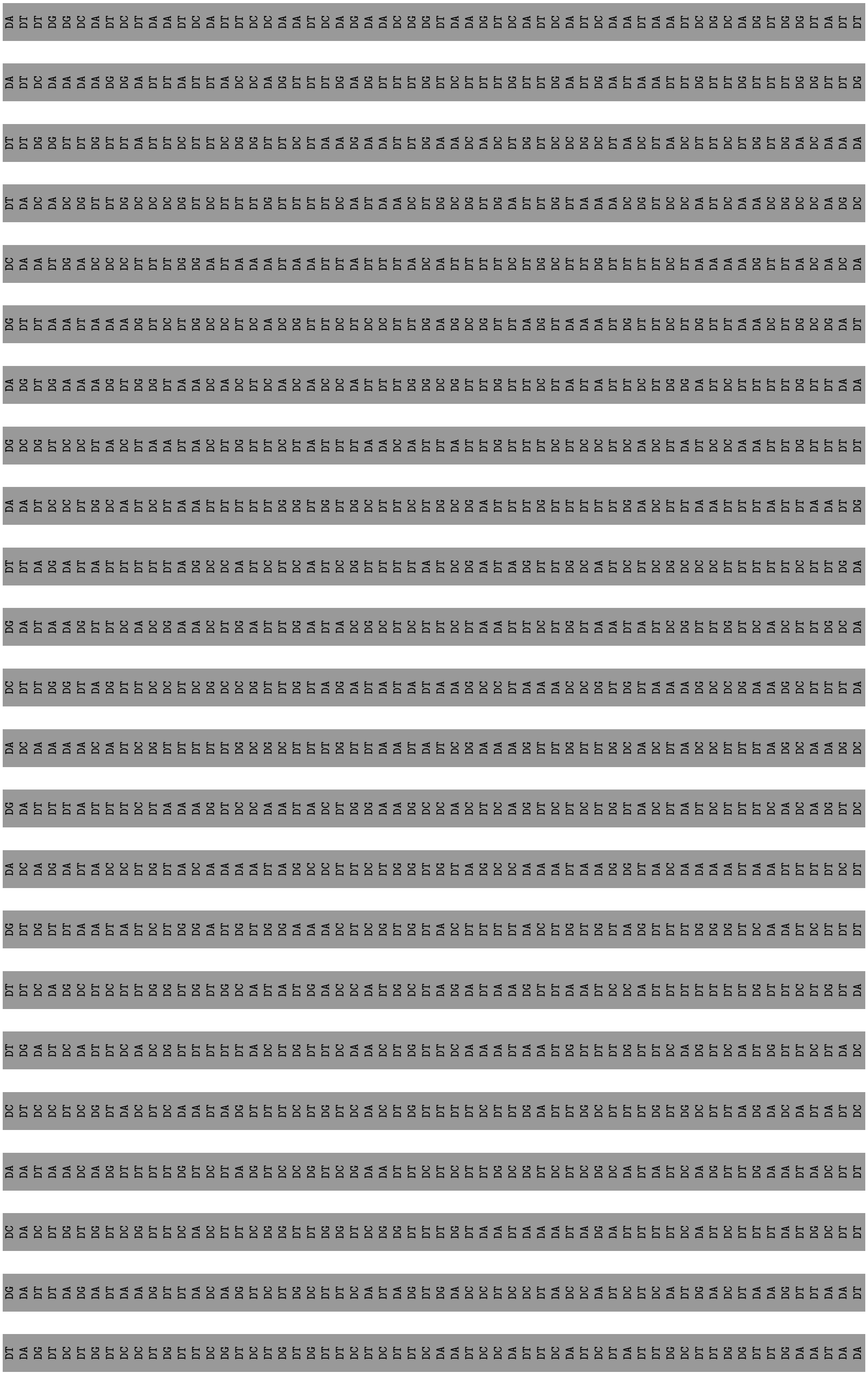




WORLDWIDE
PDB
PROTEIN DATA BANK







DT DT DC DT DG DA DT DG DC DC DA DT DA DT DC DT DA DA DG DT DC DA DT DT DC DT DC DA DT DC DT DA DC DG DA DC DA DT DA DA DG DC DA DT DG DT DT DT DA DA DT DT DG DG DT DA DC DG DT DT

DG DT DT DC DA DT DA DT DC DT DT DA DA DC DT DT DG DT DA DT DT DG DC DC DC DT DA DA DT DA DC DC DT DT DT DA DA DA DT DG DA DA DT DT DG DC DT DA DA

DA DT DC DG DA DA DT DC DC DG DT DT DA DA DC DA DT DG DA DA DT DT DT DT DG DC DG DT DA DT DG DA DT DT DC DG DC DC DT DA DT DT DT DC DG DT DA DA DT DA DC DA DT DC DT DT

DG DA DC DT DT DC DC DT DC DC DG DT DT DT DT DC DC DG DT DT DT DT DT DG DA DA DC DA DT DG DA DA DT DT DT DT DC DT DT DT DA DA DG DT DG DT DT DT DA DA DT DT DG DA DT DA DC DT DG

DT DT DA DC DT DT DG DC DT DC DC DA DA DC DA DA DG DT DA DG DC DA DG DT DT DT DT DT DA DT DT DC DT DT DT DG DA DA DA DA DG DC DA DT DT DC DT DT DT DG DT DA DT DA DC DA

DT DG DA DA DT DG DT DT DG DC DG DA DT DA DC DT DT DG DC DG DA DA DG DT DA DC DG DC DA DA DT DG DA DC DC DA DT DG DC DT DG DT DA DG DT DG DA DA DG DT DA DT DG DA DA DC DC

DG DC DA DT DA DA DG DT DA DC DA DA DT DC DG DA DA DG DA DC DT DG DT DG DA DA DG DC DA DT DA DA DC DT DG DC DT DG DT DA DG DG DT DA DT DT DT DA DA DA

DT DG DT DT DA DC DT DC DA DT DT DT DT DT DT DT DT DT DC DT DC DC DT DC DT DT DG DG DT DT DG DT DT DC DC DA DT DC DG DT DT DA DA DC DC DT DT DG DC DA DT DT DT

DT DC DC DA DT DT DA DA DT DT DT DT DT DG DA DG DG DT DA DT DA DC DC DC DA DG DC DT DT DT DA DA DC DT DG DG DA DT DT DT DT DC DG DT DA DA DG DA DA DT DG

DT DG DT DG DG DT DG DG DC DA DC DT DT DC DT DA DC DC DA DT DT DT DC DG DT DT DG DT DA DA DT DT DT DC DT DC DC DC DT DA DG DT DG DA DA DG DG DG DA DG

DG DT DG DA DT DA DA DG DT DA DT DG DG DA DA DT DT DT DA DA DT DC DC DA DA DA DT DT DT DA DG DA DA DC DT DT DA DG DT DT DA DA DT DC DC DC

DG DA DA DA DA DA DG DA DT DT DC DG DG DT DG DG DT DG DA DC DA DT DT DT DT DT DA DG DA DA DT DA DG DA DA DT DA DA DA DG DT DT DG DA DA DA DG DA DA

DG DA DT DG DC DT DC DC DT DG DA DA DA DC DG DT DA DT DC DT DT DA DC DG DA DA DT DA DT DT DT DA DT DC DT DG DC DT DT DT DA DT DA DA DG DC DT DT DA

DA DA DA DG DC DT DC DA DA DC DA DA DG DA DT DA DT DG DA DA DT DT DT DG DA DA DT DA DC DA DG DA DT DT DC DT DT DT DA DA DA DT DC DG DT DG DA DA DT DG DA

DC DA DA DT DA DT DC DC DA DA DC DA DT DT DG DA DT DT DT DT DG DA DG DC DA DT DT DA DT DT DA DA DT DT DT DG DG DA DT DA DA DT DC DA DA DA DT DT DT DC DT DA

DG DT DT DT DA DA DA DT DT DC DT DG DT DG DT DT DT DT DC DT DC DA DT DT DC DC DC DA DA DT DT DC DC DA DG DT DC DG DC DA DA DT DC DT DT DC DC DA DT DT DG DA

DT DG DA DG DG DG DA DG DT DT DG DT DT DG DT DG DT DG DT DT DC DA DT DT DT DG DA DA DT DT DT DT DT DT DA DA DT DT DG DT DT DG DT DA DA DT DT DG DT DT DA

DG DT DT DT DA DG DG DT DT DG DC DA DT DT DG DC DT DC DC DT DT DT DG DA DA DT DG DA DA DT DT DT DT DT DG DT DC DC DT DA DA DT DC DG DT DA DA DT DG DT DG

DT DG DA DG DT DA DA DT DT DC DT DT DT DG DG DA DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT

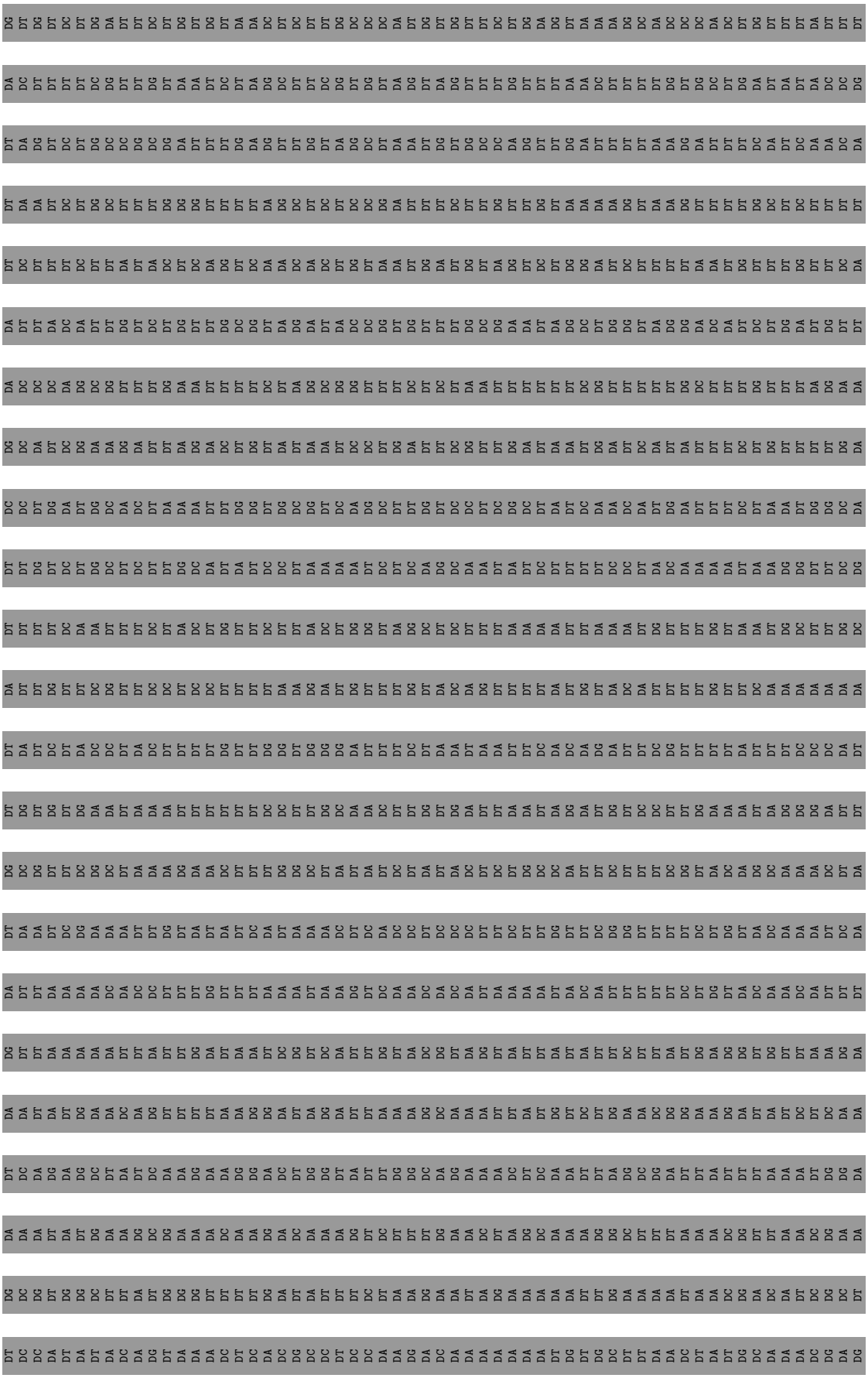
DA DA DC DA DG DA DT DT DC DT DC DA DA DG DG DT DA DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT DT

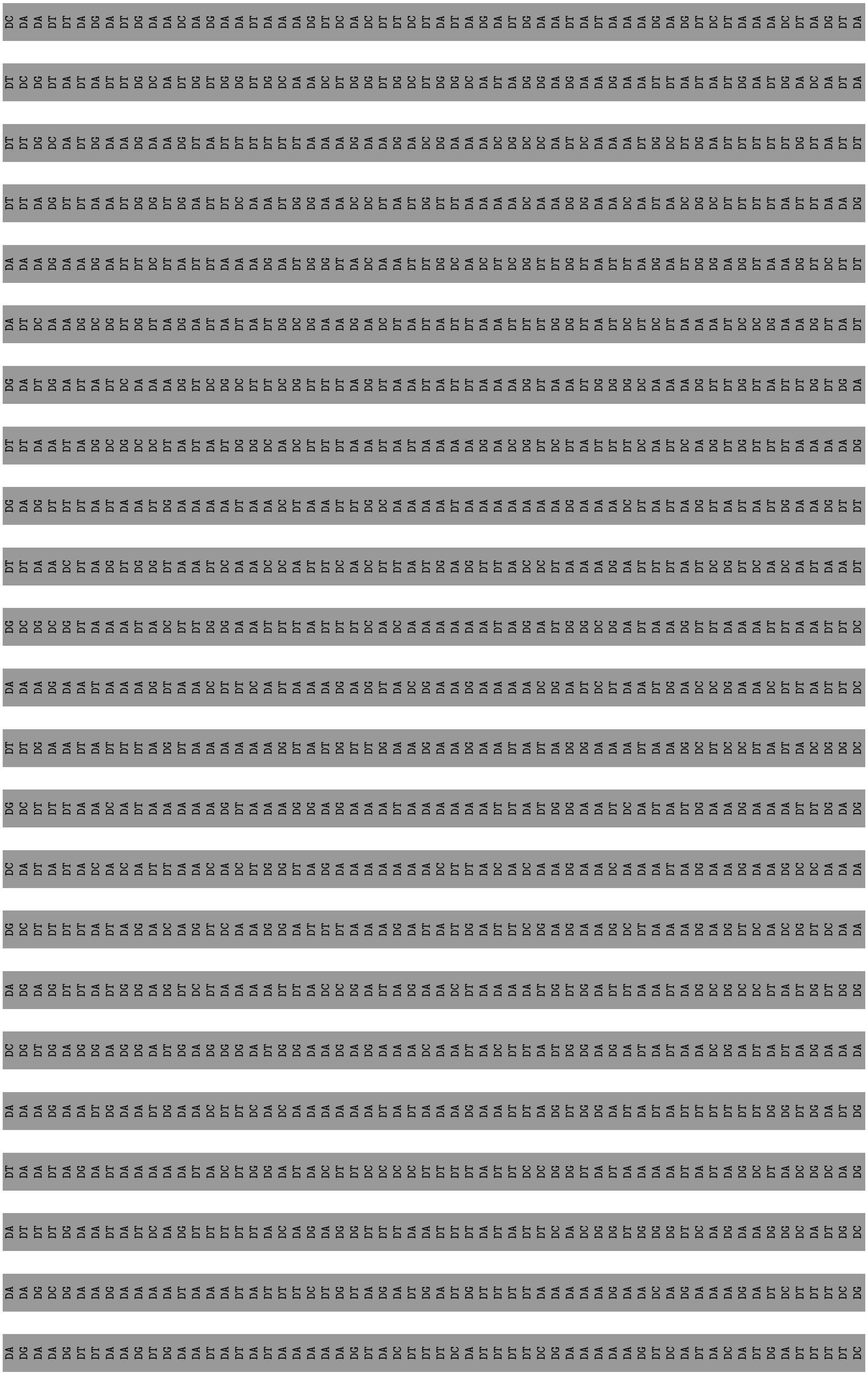
DG DC DT DA DA DT DA DA DA DT DA DA DT DG DT DT DC DC DC DA DA DT DT DT DG DA DT DA DT DA DT DT DT DC DT DC DC DC DT DT DT DC DG DT DT DC DC DA DC DC DA

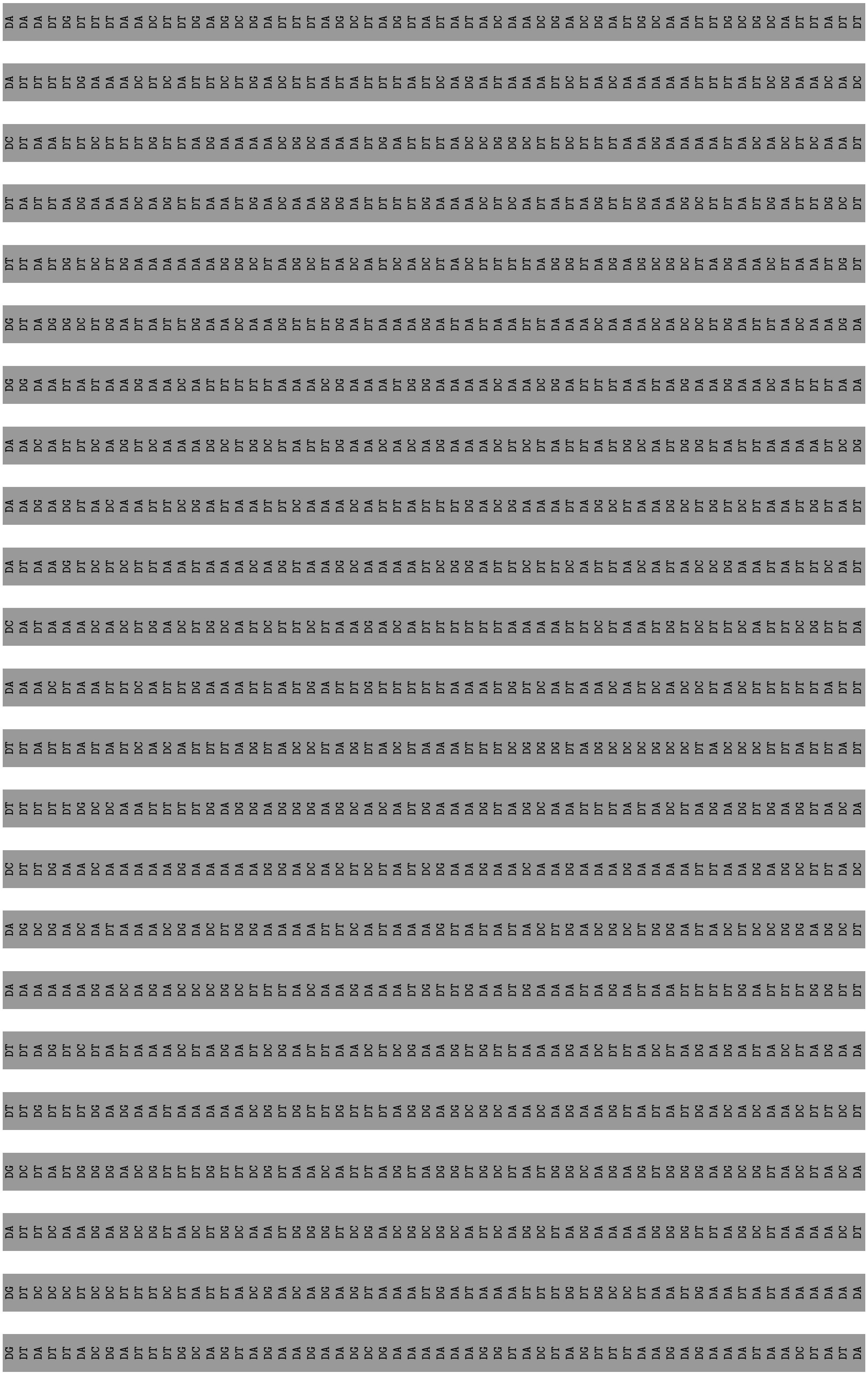
DT DA DG DT DA DA DA DA DT DA DA DG DA DT DG DC DT DT DT DT DG DG DG DC DT DC DG DG DT DT DT DC DC DC DT DT DT DC DC DC DT DT DT DC DC DC DT DT DC DA

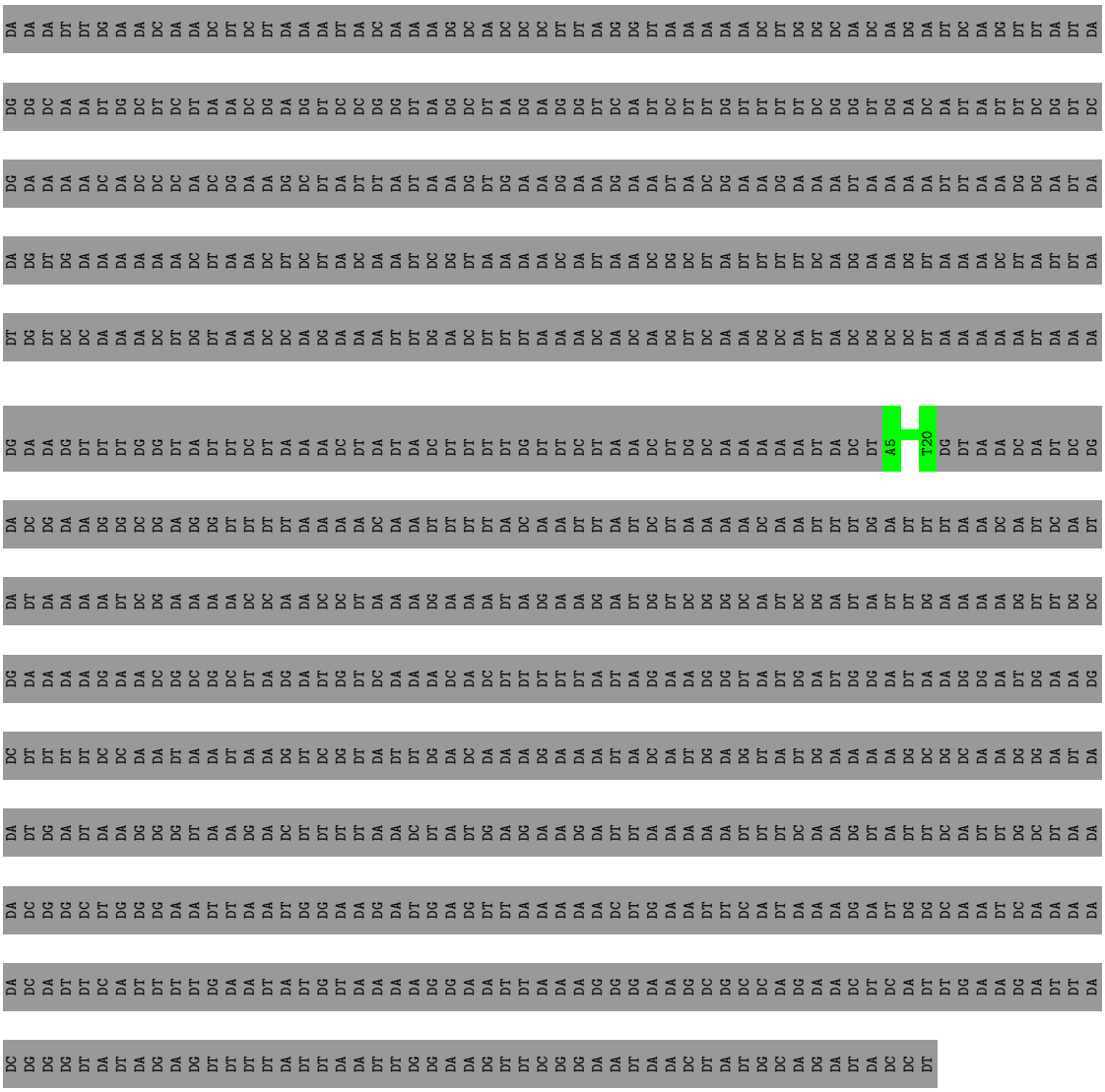
DG DT DT DT DA DC DG DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT

DT DG DT DG DT DC DC DA DT DC DT DT DA DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT DG DT DT








A5
T20

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	35724	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	35.26	Depositor
Minimum defocus (nm)	700	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	2.602	Depositor
Minimum map value	-1.352	Depositor
Average map value	-0.013	Depositor
Map value standard deviation	0.139	Depositor
Recommended contour level	0.5	Depositor
Map size (Å)	319.2, 319.2, 319.2	wwPDB
Map dimensions	240, 240, 240	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.33, 1.33, 1.33	Depositor

5 Model quality (i)

5.1 Standard geometry (i)

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	0.31	0/789	0.57	0/1058
1	B	0.32	0/789	0.60	0/1058
1	G	0.31	0/789	0.58	0/1058
1	H	0.31	0/789	0.59	0/1058
1	S	0.31	0/789	0.59	0/1058
1	T	0.31	0/789	0.57	0/1058
1	Y	0.31	0/789	0.59	0/1058
1	Z	0.31	0/789	0.59	0/1058
1	e	0.31	0/789	0.58	0/1058
1	f	0.32	0/789	0.58	0/1058
1	k	0.31	0/789	0.60	0/1058
1	l	0.32	0/789	0.59	0/1058
2	C	0.34	0/847	0.55	0/1145
2	I	0.34	0/847	0.56	0/1145
2	U	0.33	0/847	0.56	0/1145
2	a	0.34	0/847	0.55	0/1145
2	g	0.35	0/847	0.62	0/1145
2	m	0.33	0/847	0.55	0/1145
3	D	0.35	0/1061	0.55	0/1442
3	J	0.34	0/1061	0.52	0/1442
3	V	0.34	0/1061	0.54	0/1442
3	b	0.35	0/1061	0.64	0/1442
3	h	0.32	0/1061	0.61	0/1442
3	n	0.34	0/1061	0.62	0/1442
4	E	0.30	0/1314	0.51	0/1764
4	F	0.32	0/107	0.34	0/143
4	K	0.30	0/1314	0.52	0/1764
4	L	0.32	0/107	0.35	0/143
4	W	0.30	0/1314	0.52	0/1764
4	X	0.32	0/107	0.36	0/143
4	c	0.30	0/1314	0.55	0/1764
4	d	0.31	0/107	0.33	0/143
4	i	0.30	0/1314	0.55	0/1764
4	j	0.34	0/107	0.36	0/143

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
4	o	0.29	0/1314	0.54	0/1764
4	p	0.33	0/107	0.36	0/143
5	M	0.33	0/308	0.61	0/409
5	N	0.30	0/301	0.55	0/399
5	O	0.35	0/301	0.71	0/399
6	P	0.24	0/923	0.52	0/1243
7	Q	0.24	0/359	0.67	0/552
8	R	0.22	0/375	0.59	0/576
All	All	0.32	0/32009	0.57	0/43238

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
5	O	0	1
6	P	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	O	30	ARG	Sidechain
6	P	44	ARG	Sidechain

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	780	0	770	0	0
1	B	780	0	770	1	0
1	G	780	0	770	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	H	780	0	770	1	0
1	S	780	0	770	3	0
1	T	780	0	770	1	0
1	Y	780	0	770	1	0
1	Z	780	0	770	2	0
1	e	780	0	770	2	0
1	f	780	0	770	2	0
1	k	780	0	770	2	0
1	l	780	0	770	1	0
2	C	832	0	841	1	0
2	I	832	0	841	1	0
2	U	832	0	841	0	0
2	a	832	0	841	0	0
2	g	832	0	841	0	0
2	m	832	0	841	0	0
3	D	1037	0	984	4	0
3	J	1037	0	984	3	0
3	V	1037	0	984	2	0
3	b	1037	0	984	4	0
3	h	1037	0	984	3	0
3	n	1037	0	984	2	0
4	E	1288	0	1210	3	0
4	F	105	0	94	0	0
4	K	1288	0	1210	7	0
4	L	105	0	94	0	0
4	W	1288	0	1210	5	0
4	X	105	0	94	0	0
4	c	1288	0	1210	4	0
4	d	105	0	94	0	0
4	i	1288	0	1210	4	0
4	j	105	0	94	0	0
4	o	1288	0	1210	5	0
4	p	105	0	94	0	0
5	M	306	0	329	0	0
5	N	299	0	322	0	0
5	O	299	0	322	0	0
6	P	904	0	879	1	0
7	Q	324	0	188	1	0
8	R	332	0	180	0	0
All	All	31396	0	30234	61	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

All (61) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:W:74:ASP:OD1	4:c:141:ARG:NH2	2.17	0.77
4:c:74:ASP:OD1	4:i:141:ARG:NH2	2.20	0.75
4:i:74:ASP:OD1	4:o:141:ARG:NH2	2.31	0.64
4:E:141:ARG:NH2	4:o:74:ASP:OD1	2.36	0.58
4:E:74:ASP:OD1	4:K:141:ARG:NH2	2.38	0.57
6:P:72:ILE:HD12	6:P:72:ILE:N	2.23	0.54
4:K:74:ASP:OD1	4:W:141:ARG:NH2	2.42	0.53
4:c:127:ASP:OD1	4:c:127:ASP:O	2.28	0.51
3:h:82:ILE:HD11	3:h:103:ILE:HD11	1.91	0.51
4:i:127:ASP:O	4:i:127:ASP:OD1	2.29	0.50
3:b:74:HIS:HE1	3:b:118:THR:OG1	1.95	0.50
1:k:10:ARG:NH1	1:k:54:GLU:OE2	2.45	0.49
1:G:10:ARG:NH1	1:G:54:GLU:OE2	2.47	0.48
1:Y:10:ARG:NH1	1:Y:54:GLU:OE2	2.46	0.48
4:o:127:ASP:OD1	4:o:127:ASP:O	2.31	0.47
1:Z:10:ARG:NH1	1:Z:54:GLU:OE1	2.48	0.47
3:J:82:ILE:HD11	3:J:103:ILE:HD11	1.97	0.47
4:W:127:ASP:O	4:W:127:ASP:OD1	2.33	0.47
3:b:74:HIS:CE1	3:b:118:THR:OG1	2.69	0.46
1:f:10:ARG:NH1	1:f:54:GLU:OE1	2.48	0.46
3:J:18:TYR:OH	3:J:84:ASP:OD2	2.25	0.46
4:K:127:ASP:O	4:K:127:ASP:OD1	2.34	0.45
1:e:10:ARG:NH1	1:e:54:GLU:OE2	2.49	0.45
4:K:26:MET:HE2	4:K:131:TYR:CE1	2.52	0.45
4:E:127:ASP:OD1	4:E:127:ASP:O	2.33	0.44
7:Q:1:DA:P	7:Q:1:DA:O4'	2.75	0.44
1:S:73:VAL:O	1:S:73:VAL:HG13	2.16	0.44
1:S:93:ASP:OD1	1:S:97:ASN:ND2	2.44	0.44
3:b:56:GLY:O	3:b:126:THR:HG22	2.18	0.44
1:l:10:ARG:NH1	1:l:54:GLU:OE1	2.51	0.44
1:f:21:GLN:NE2	1:f:63:ILE:HD11	2.33	0.43
3:h:127:PHE:CZ	4:i:126:ILE:HD11	2.53	0.43
3:V:86:THR:HB	3:V:87:PRO:HA	2.00	0.43
3:n:23:ILE:O	3:n:38:VAL:HA	2.18	0.42
3:n:74:HIS:HE1	3:n:118:THR:OG1	2.02	0.42
4:K:45:THR:HG22	4:K:47:ASP:H	1.85	0.42
1:B:21:GLN:NE2	1:B:63:ILE:HD11	2.34	0.42
1:e:73:VAL:HG13	1:e:73:VAL:O	2.20	0.42
3:J:86:THR:HB	3:J:87:PRO:HA	2.02	0.42
4:K:41:ASP:O	4:K:52:THR:HG22	2.20	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:W:67:PHE:CZ	4:W:70:ASP:HB2	2.55	0.42
3:b:23:ILE:O	3:b:38:VAL:HA	2.20	0.42
4:W:45:THR:HG22	4:W:47:ASP:H	1.84	0.42
3:D:86:THR:HB	3:D:87:PRO:HA	2.01	0.41
1:T:66:GLU:OE1	1:T:66:GLU:N	2.52	0.41
1:G:73:VAL:O	1:G:73:VAL:HG13	2.20	0.41
3:D:23:ILE:O	3:D:38:VAL:HA	2.19	0.41
1:k:46:GLU:OE2	1:k:49:SER:OG	2.36	0.41
4:o:26:MET:HE2	4:o:131:TYR:CE1	2.56	0.41
1:Z:66:GLU:OE1	1:Z:66:GLU:N	2.51	0.41
3:V:78:VAL:HG13	3:V:120:ILE:HD11	2.02	0.41
3:h:56:GLY:O	3:h:126:THR:HG22	2.21	0.41
2:C:84:ASP:OD1	2:C:85:THR:N	2.54	0.41
3:D:56:GLY:O	3:D:126:THR:HG22	2.20	0.41
3:D:74:HIS:HE1	3:D:118:THR:OG1	2.04	0.41
1:S:10:ARG:NH1	1:S:54:GLU:OE2	2.54	0.41
4:o:67:PHE:CZ	4:o:70:ASP:HB2	2.56	0.41
1:H:21:GLN:NE2	1:H:63:ILE:HD11	2.35	0.41
4:K:9:ASP:OD1	4:K:94:ARG:NH1	2.52	0.40
4:c:9:ASP:OD1	4:c:94:ARG:NH1	2.55	0.40
2:I:99:VAL:O	2:I:100:ASN:HB2	2.22	0.40

There are no symmetry-related clashes.

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	
1	A	94/110 (86%)	90 (96%)	4 (4%)	0	100	100
1	B	94/110 (86%)	92 (98%)	2 (2%)	0	100	100
1	G	94/110 (86%)	91 (97%)	3 (3%)	0	100	100
1	H	94/110 (86%)	91 (97%)	3 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	S	94/110 (86%)	91 (97%)	3 (3%)	0	100	100
1	T	94/110 (86%)	91 (97%)	3 (3%)	0	100	100
1	Y	94/110 (86%)	91 (97%)	3 (3%)	0	100	100
1	Z	94/110 (86%)	91 (97%)	3 (3%)	0	100	100
1	e	94/110 (86%)	92 (98%)	2 (2%)	0	100	100
1	f	94/110 (86%)	92 (98%)	2 (2%)	0	100	100
1	k	94/110 (86%)	90 (96%)	4 (4%)	0	100	100
1	l	94/110 (86%)	92 (98%)	2 (2%)	0	100	100
2	C	98/100 (98%)	93 (95%)	5 (5%)	0	100	100
2	I	98/100 (98%)	94 (96%)	4 (4%)	0	100	100
2	U	98/100 (98%)	92 (94%)	6 (6%)	0	100	100
2	a	98/100 (98%)	92 (94%)	6 (6%)	0	100	100
2	g	98/100 (98%)	93 (95%)	4 (4%)	1 (1%)	12	45
2	m	98/100 (98%)	93 (95%)	5 (5%)	0	100	100
3	D	125/127 (98%)	124 (99%)	1 (1%)	0	100	100
3	J	125/127 (98%)	123 (98%)	2 (2%)	0	100	100
3	V	125/127 (98%)	124 (99%)	1 (1%)	0	100	100
3	b	125/127 (98%)	123 (98%)	2 (2%)	0	100	100
3	h	125/127 (98%)	122 (98%)	3 (2%)	0	100	100
3	n	125/127 (98%)	124 (99%)	1 (1%)	0	100	100
4	E	160/193 (83%)	154 (96%)	6 (4%)	0	100	100
4	F	10/193 (5%)	10 (100%)	0	0	100	100
4	K	160/193 (83%)	154 (96%)	6 (4%)	0	100	100
4	L	10/193 (5%)	9 (90%)	1 (10%)	0	100	100
4	W	160/193 (83%)	155 (97%)	5 (3%)	0	100	100
4	X	10/193 (5%)	10 (100%)	0	0	100	100
4	c	160/193 (83%)	154 (96%)	6 (4%)	0	100	100
4	d	10/193 (5%)	10 (100%)	0	0	100	100
4	i	160/193 (83%)	154 (96%)	6 (4%)	0	100	100
4	j	10/193 (5%)	10 (100%)	0	0	100	100
4	o	160/193 (83%)	154 (96%)	6 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	p	10/193 (5%)	10 (100%)	0	0	100	100
5	M	36/1154 (3%)	35 (97%)	1 (3%)	0	100	100
5	N	35/1154 (3%)	35 (100%)	0	0	100	100
5	O	35/1154 (3%)	34 (97%)	1 (3%)	0	100	100
6	P	107/115 (93%)	106 (99%)	1 (1%)	0	100	100
All	All	3699/8575 (43%)	3585 (97%)	113 (3%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	g	25	GLU

5.3.2 Protein sidechains (i)

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	
1	A	86/97 (89%)	86 (100%)	0	100	100
1	B	86/97 (89%)	86 (100%)	0	100	100
1	G	86/97 (89%)	86 (100%)	0	100	100
1	H	86/97 (89%)	86 (100%)	0	100	100
1	S	86/97 (89%)	86 (100%)	0	100	100
1	T	86/97 (89%)	86 (100%)	0	100	100
1	Y	86/97 (89%)	86 (100%)	0	100	100
1	Z	86/97 (89%)	86 (100%)	0	100	100
1	e	86/97 (89%)	86 (100%)	0	100	100
1	f	86/97 (89%)	86 (100%)	0	100	100
1	k	86/97 (89%)	86 (100%)	0	100	100
1	l	86/97 (89%)	85 (99%)	1 (1%)	63	73
2	C	95/95 (100%)	95 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	I	95/95 (100%)	94 (99%)	1 (1%)	65	74
2	U	95/95 (100%)	95 (100%)	0	100	100
2	a	95/95 (100%)	95 (100%)	0	100	100
2	g	95/95 (100%)	94 (99%)	1 (1%)	65	74
2	m	95/95 (100%)	94 (99%)	1 (1%)	65	74
3	D	121/121 (100%)	121 (100%)	0	100	100
3	J	121/121 (100%)	121 (100%)	0	100	100
3	V	121/121 (100%)	121 (100%)	0	100	100
3	b	121/121 (100%)	120 (99%)	1 (1%)	73	77
3	h	121/121 (100%)	121 (100%)	0	100	100
3	n	121/121 (100%)	121 (100%)	0	100	100
4	E	132/160 (82%)	132 (100%)	0	100	100
4	F	11/160 (7%)	11 (100%)	0	100	100
4	K	132/160 (82%)	132 (100%)	0	100	100
4	L	11/160 (7%)	11 (100%)	0	100	100
4	W	132/160 (82%)	132 (100%)	0	100	100
4	X	11/160 (7%)	11 (100%)	0	100	100
4	c	132/160 (82%)	132 (100%)	0	100	100
4	d	11/160 (7%)	11 (100%)	0	100	100
4	i	132/160 (82%)	132 (100%)	0	100	100
4	j	11/160 (7%)	11 (100%)	0	100	100
4	o	132/160 (82%)	132 (100%)	0	100	100
4	p	11/160 (7%)	11 (100%)	0	100	100
5	M	33/965 (3%)	33 (100%)	0	100	100
5	N	32/965 (3%)	32 (100%)	0	100	100
5	O	32/965 (3%)	32 (100%)	0	100	100
6	P	96/101 (95%)	96 (100%)	0	100	100
All	All	3379/7376 (46%)	3374 (100%)	5 (0%)	87	87

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	I	87	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	b	62	ILE
2	g	87	ILE
1	l	82	LEU
2	m	87	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (44) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	21	GLN
1	B	94	ASN
2	C	72	HIS
2	C	100	ASN
3	D	74	HIS
3	D	102	ASN
3	D	113	GLN
4	E	136	ASN
1	H	21	GLN
1	H	94	ASN
2	I	72	HIS
2	I	100	ASN
3	J	45	ASN
3	J	74	HIS
3	J	113	GLN
4	K	136	ASN
6	P	69	GLN
1	T	21	GLN
2	U	90	HIS
3	V	74	HIS
4	W	136	ASN
1	Z	21	GLN
1	Z	94	ASN
2	a	90	HIS
2	a	100	ASN
3	b	74	HIS
3	b	113	GLN
4	c	136	ASN
1	f	21	GLN
2	g	90	HIS
2	g	100	ASN
3	h	74	HIS
3	h	105	GLN
3	h	113	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	i	136	ASN
4	j	162	HIS
1	l	21	GLN
2	m	90	HIS
2	m	100	ASN
3	n	74	HIS
3	n	102	ASN
3	n	105	GLN
3	n	113	GLN
4	o	136	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

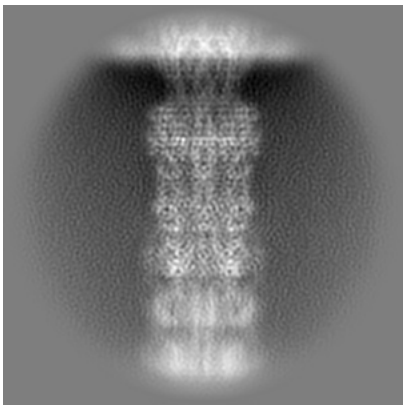
6 Map visualisation (i)

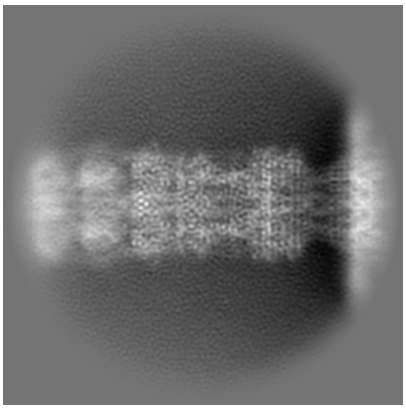
This section contains visualisations of the EMDB entry EMD-48617. These allow visual inspection of the internal detail of the map and identification of artifacts.

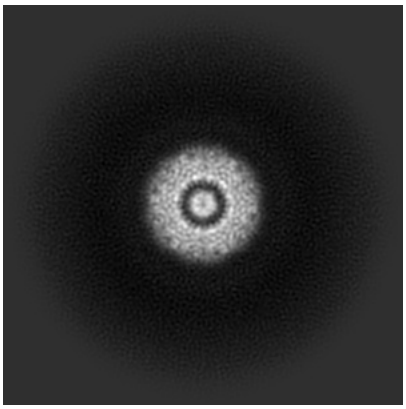
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections (i)

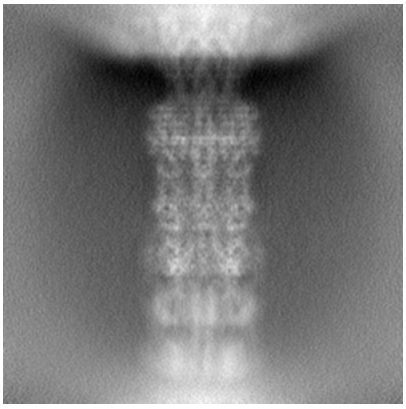
6.1.1 Primary map

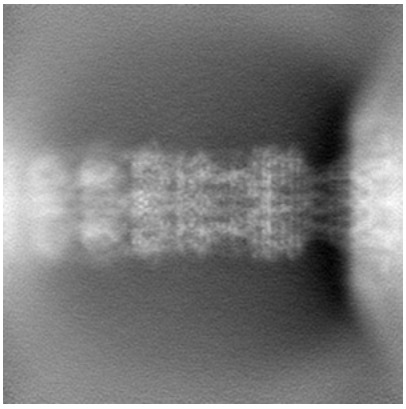

X

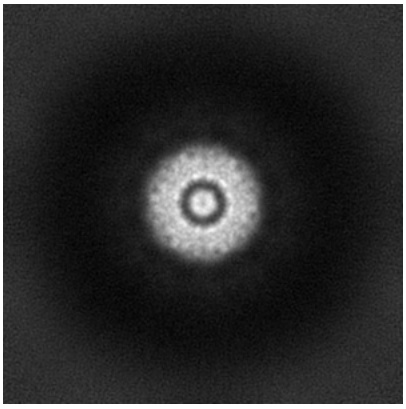

Y


Z

6.1.2 Raw map


X

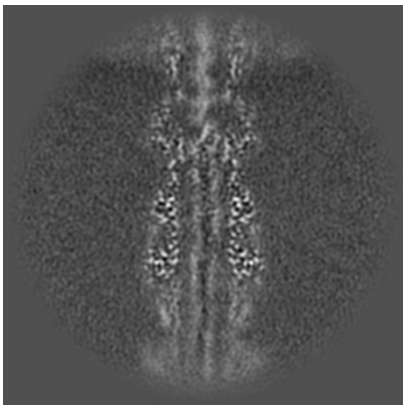

Y

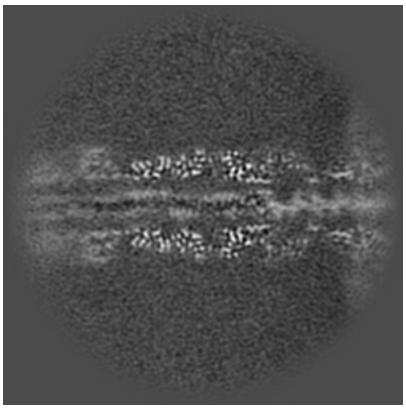

Z

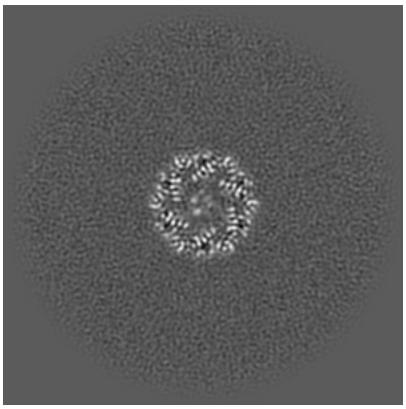
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

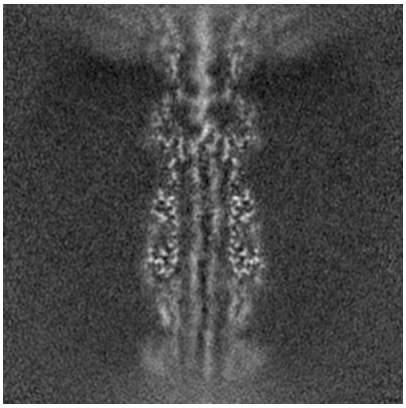
6.2.1 Primary map

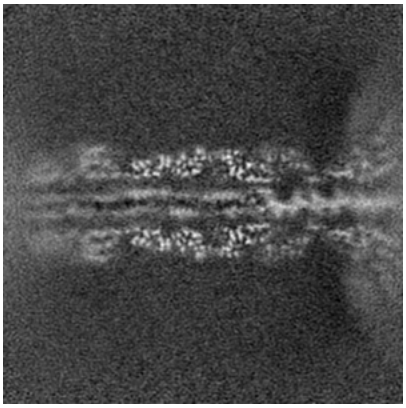

X Index: 120

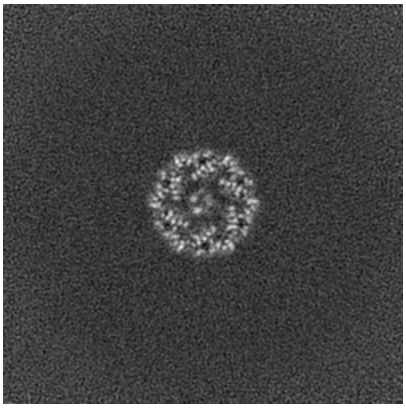

Y Index: 120


Z Index: 120

6.2.2 Raw map


X Index: 120

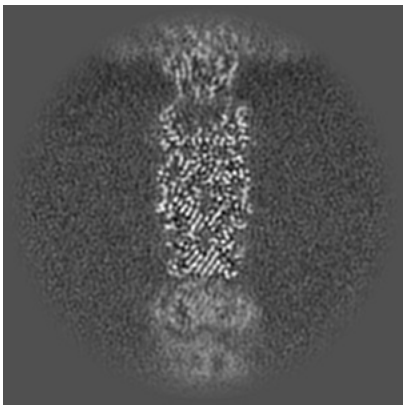

Y Index: 120

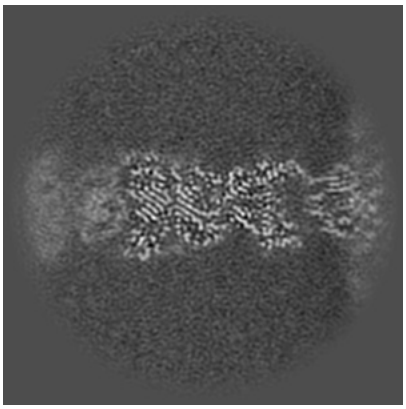

Z Index: 120

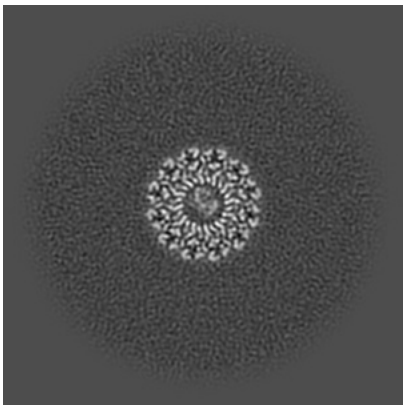
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

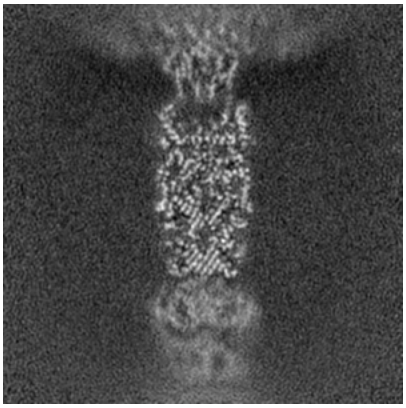
6.3.1 Primary map

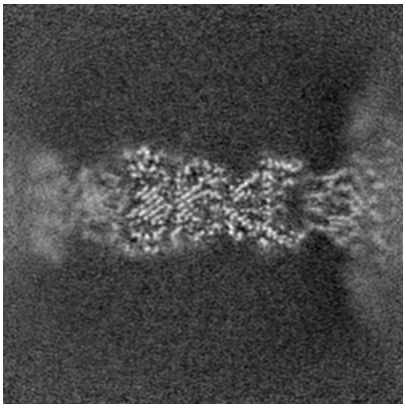

X Index: 136

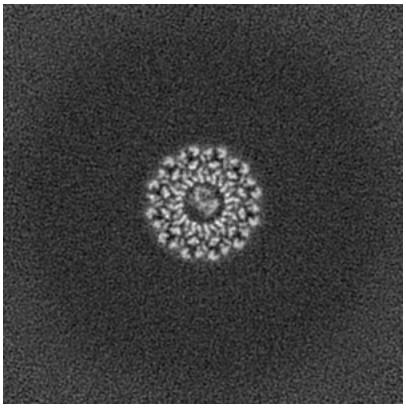

Y Index: 136


Z Index: 159

6.3.2 Raw map


X Index: 136

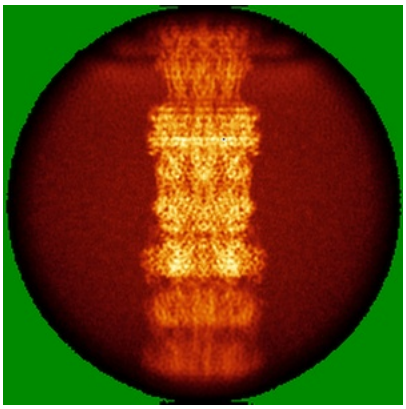

Y Index: 104

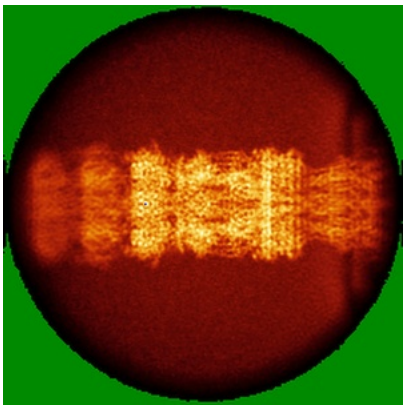

Z Index: 159

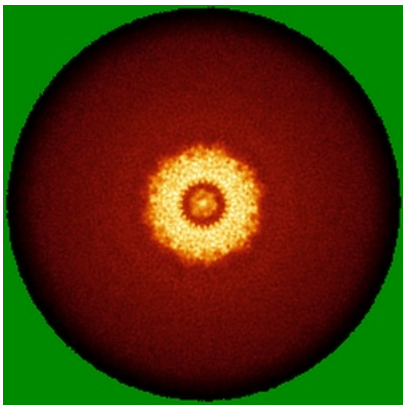
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) (i)

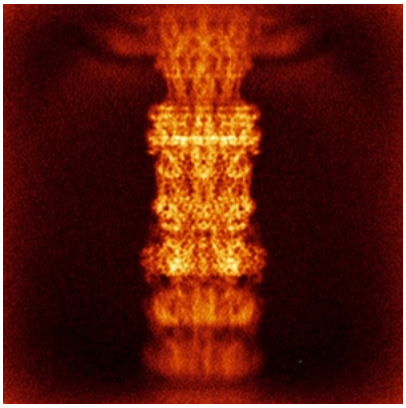
6.4.1 Primary map

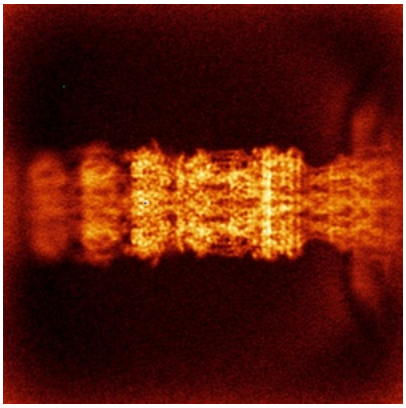

X

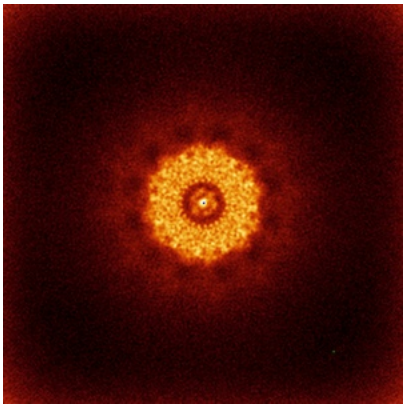

Y


Z

6.4.2 Raw map


X

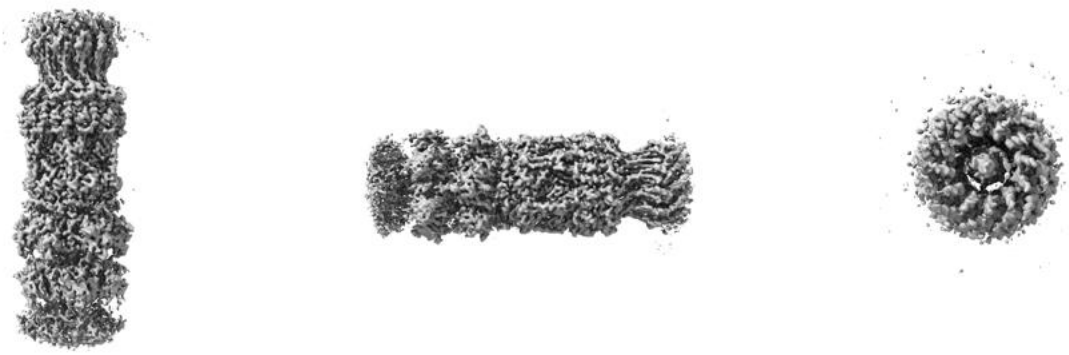

Y


Z

The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views (i)

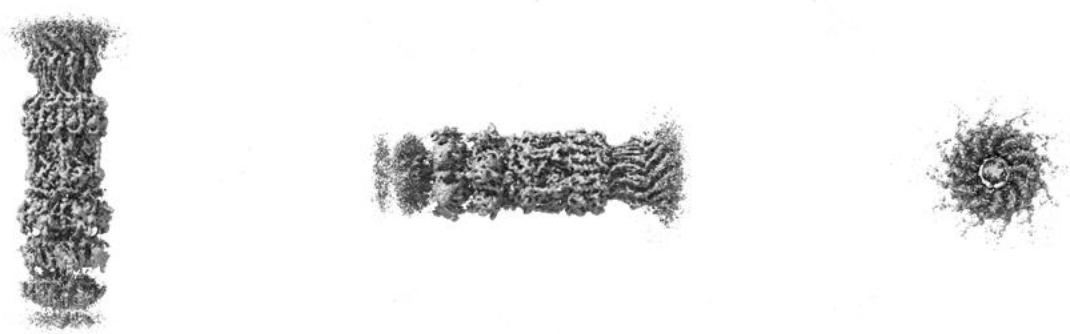
6.5.1 Primary map



X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.5. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.5.2 Raw map



X Y Z

These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

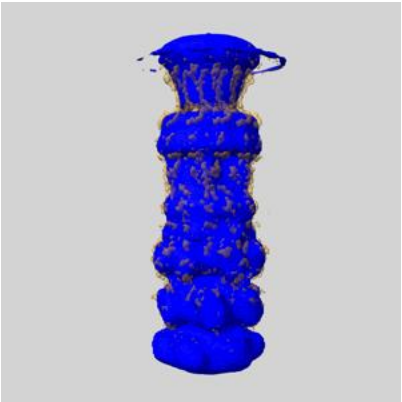
6.6 Mask visualisation (i)

This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency

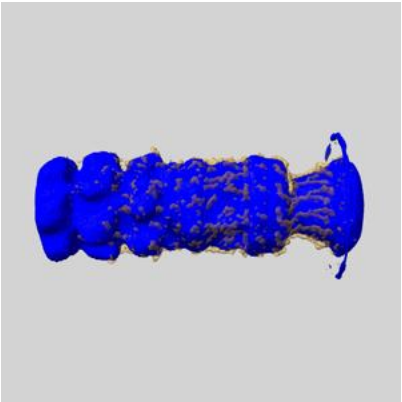
A mask typically either:

- Encompasses the whole structure
- Separates out a domain, a functional unit, a monomer or an area of interest from a larger structure

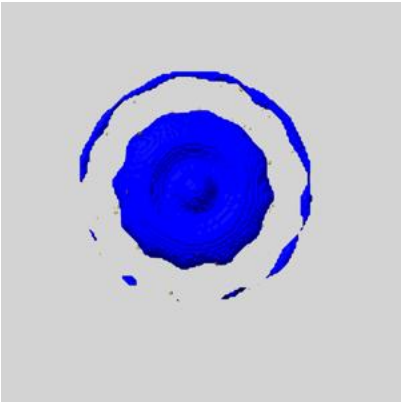
6.6.1 emd_48617_msk_1.map (i)



X



Y

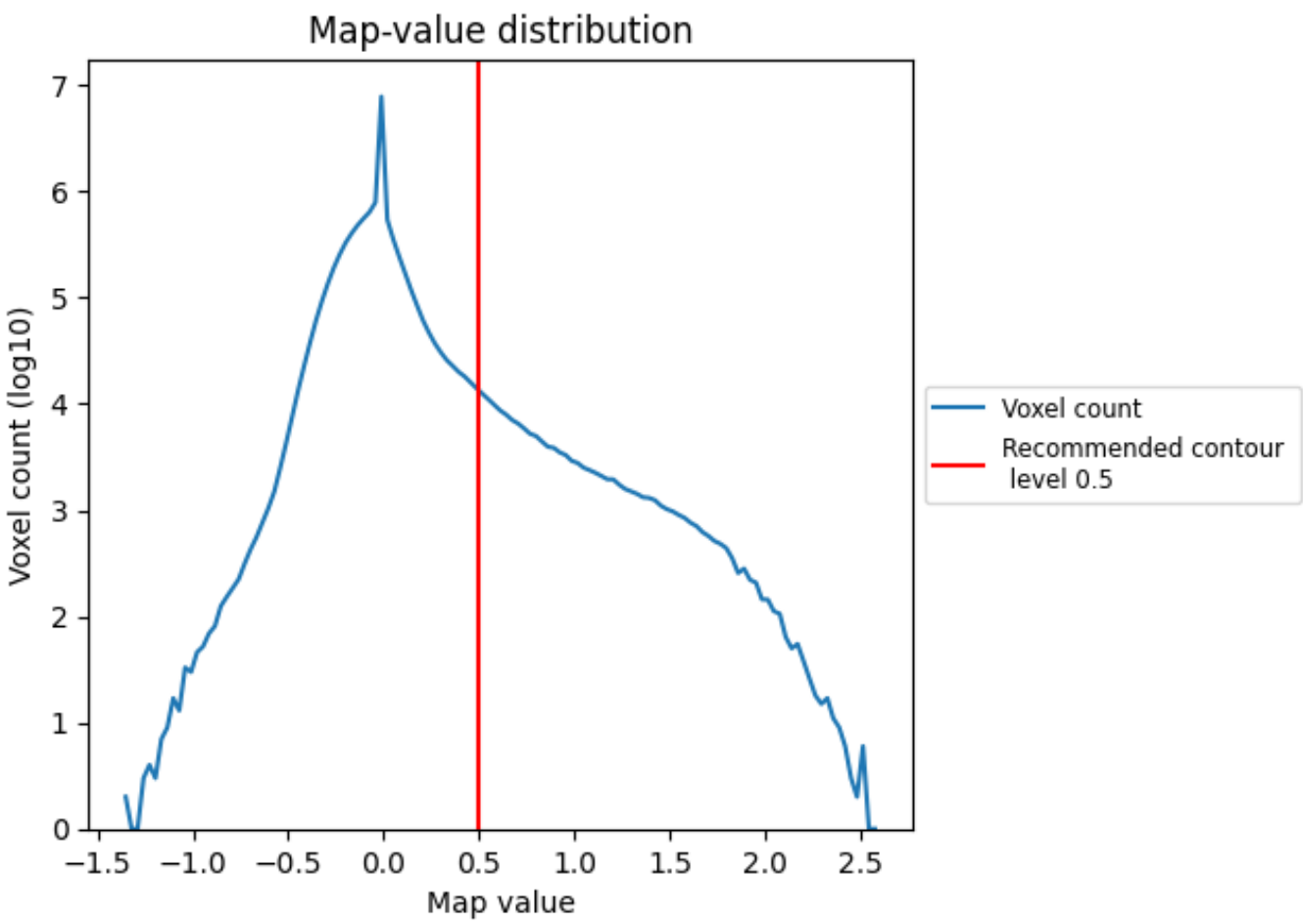


Z

7 Map analysis (i)

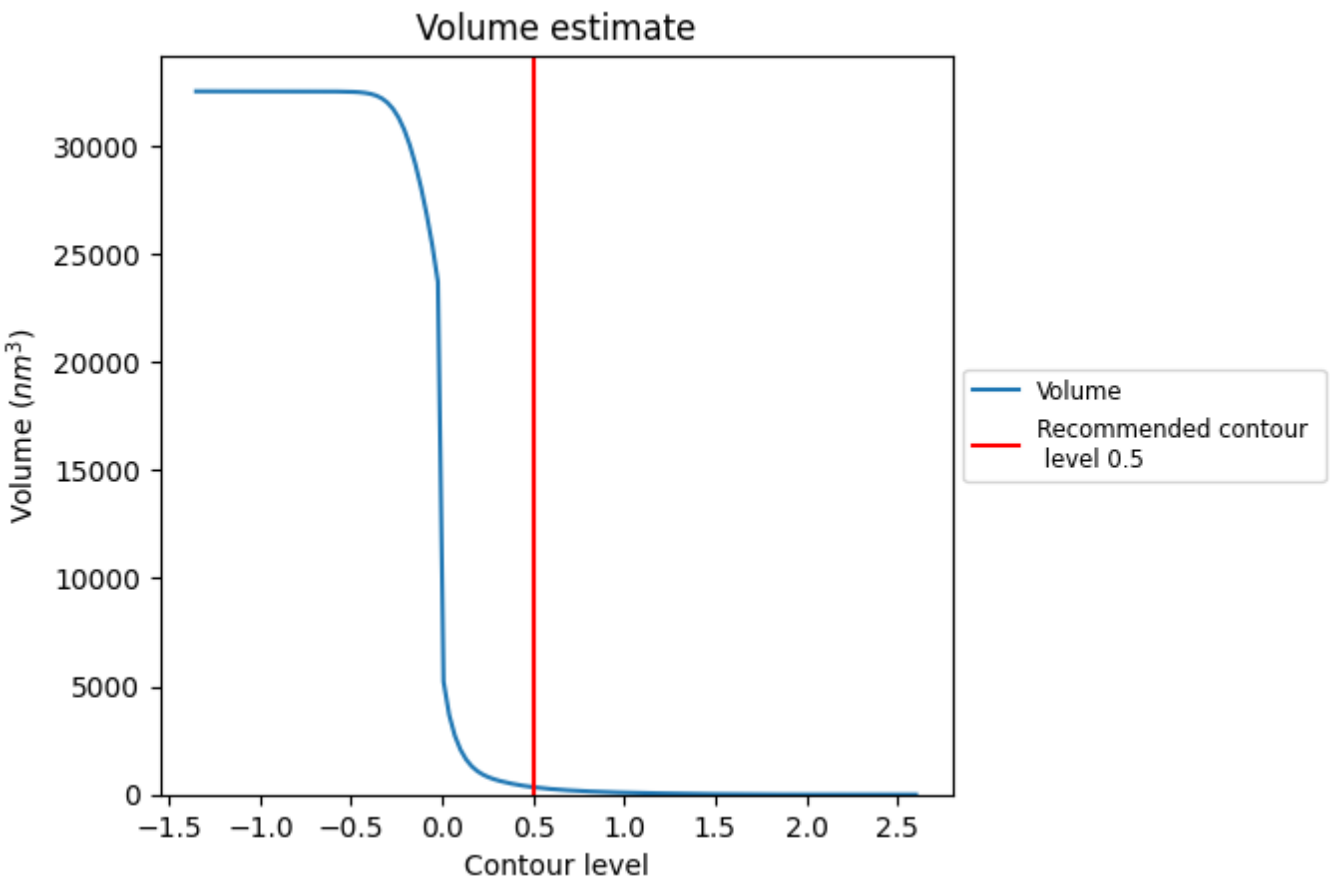
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution (i)



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

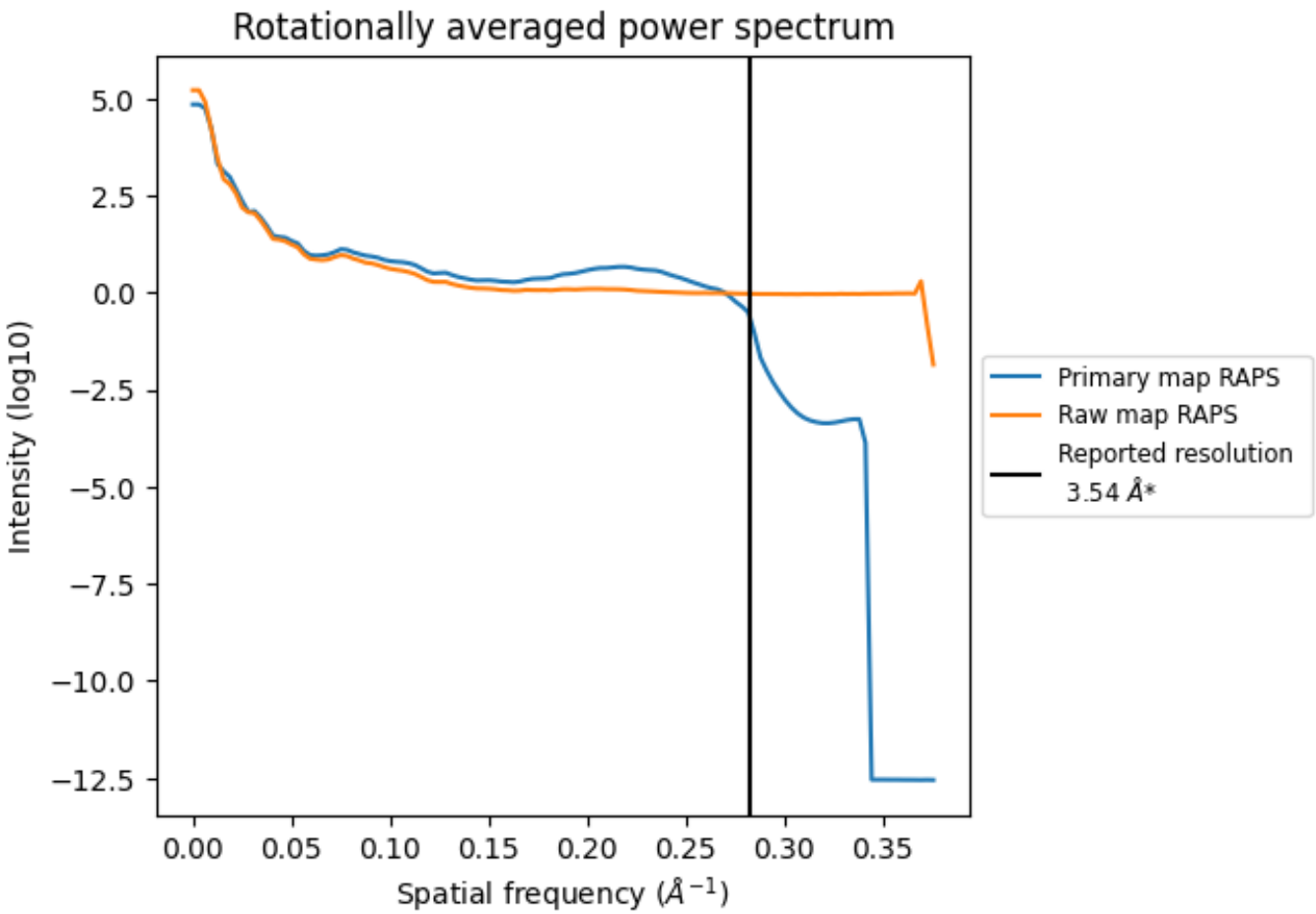
7.2 Volume estimate (i)



The volume at the recommended contour level is 344 nm^3; this corresponds to an approximate mass of 310 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

7.3 Rotationally averaged power spectrum (i)

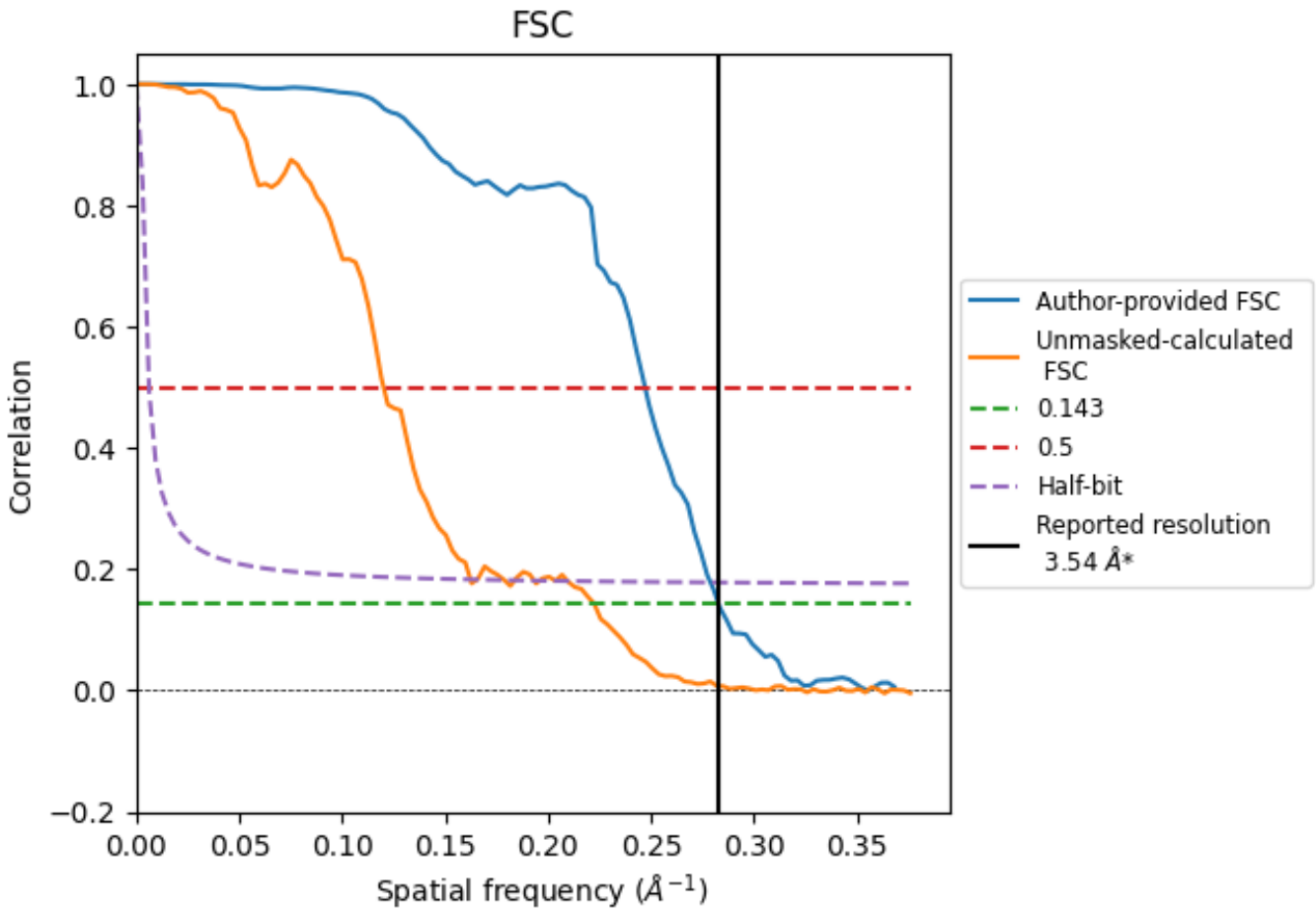


*Reported resolution corresponds to spatial frequency of 0.282 Å^{-1}

8 Fourier-Shell correlation (i)

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC (i)



*Reported resolution corresponds to spatial frequency of 0.282 Å^{-1}

8.2 Resolution estimates (i)

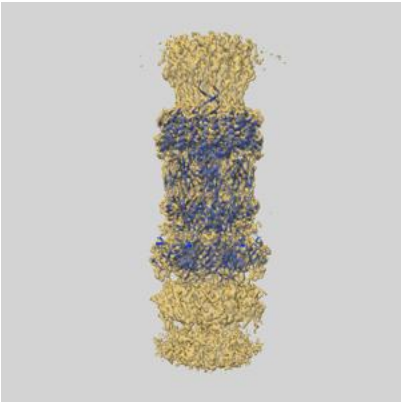
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.54	-	-
Author-provided FSC curve	3.54	4.05	3.59
Unmasked-calculated*	4.50	8.32	6.16

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 4.50 differs from the reported value 3.54 by more than 10 %

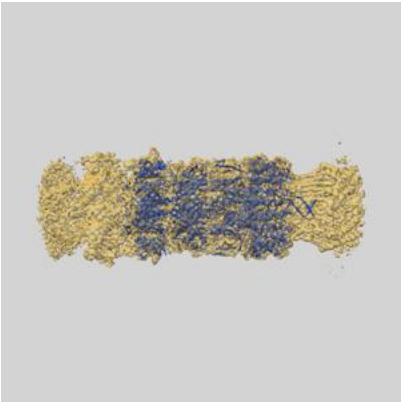
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-48617 and PDB model 9MU2. Per-residue inclusion information can be found in section 3 on page 8.

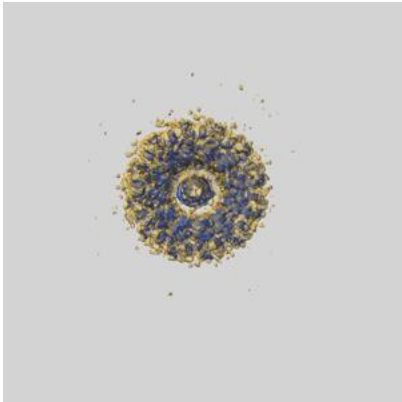
9.1 Map-model overlay (i)



X



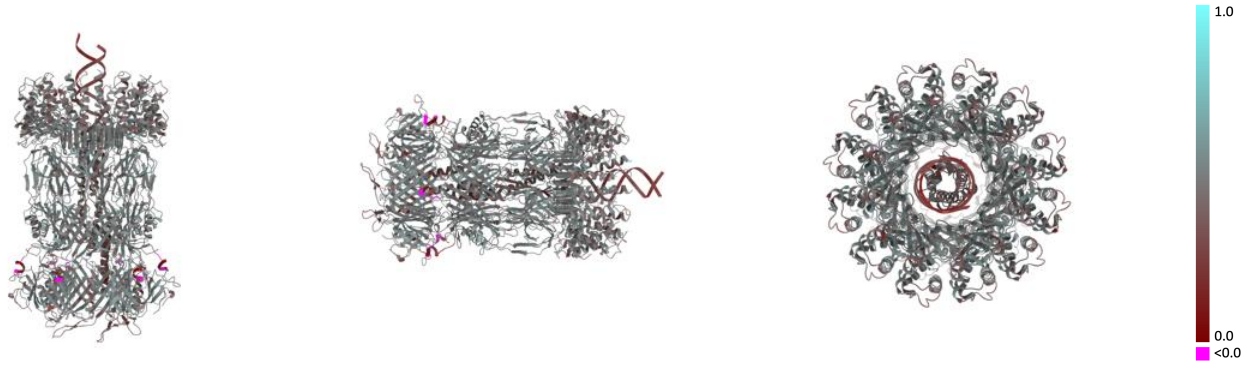
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.5 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

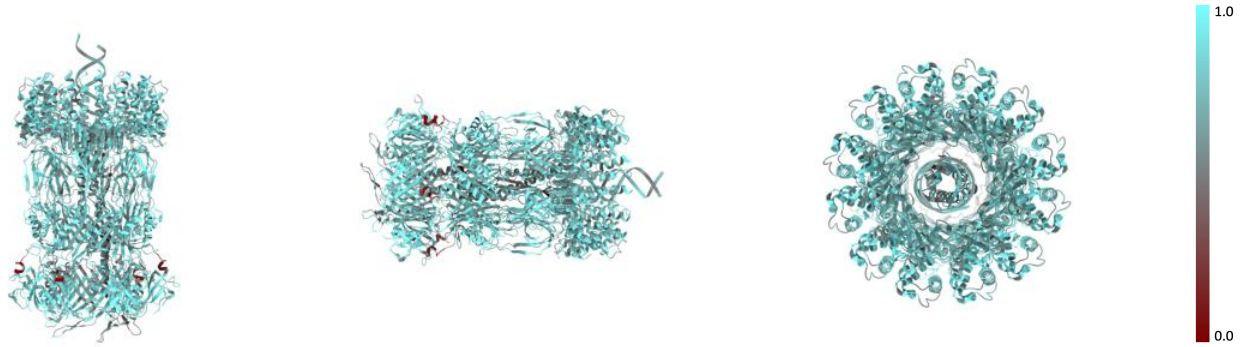
9.2 Q-score mapped to coordinate model (i)



X Y Z

The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

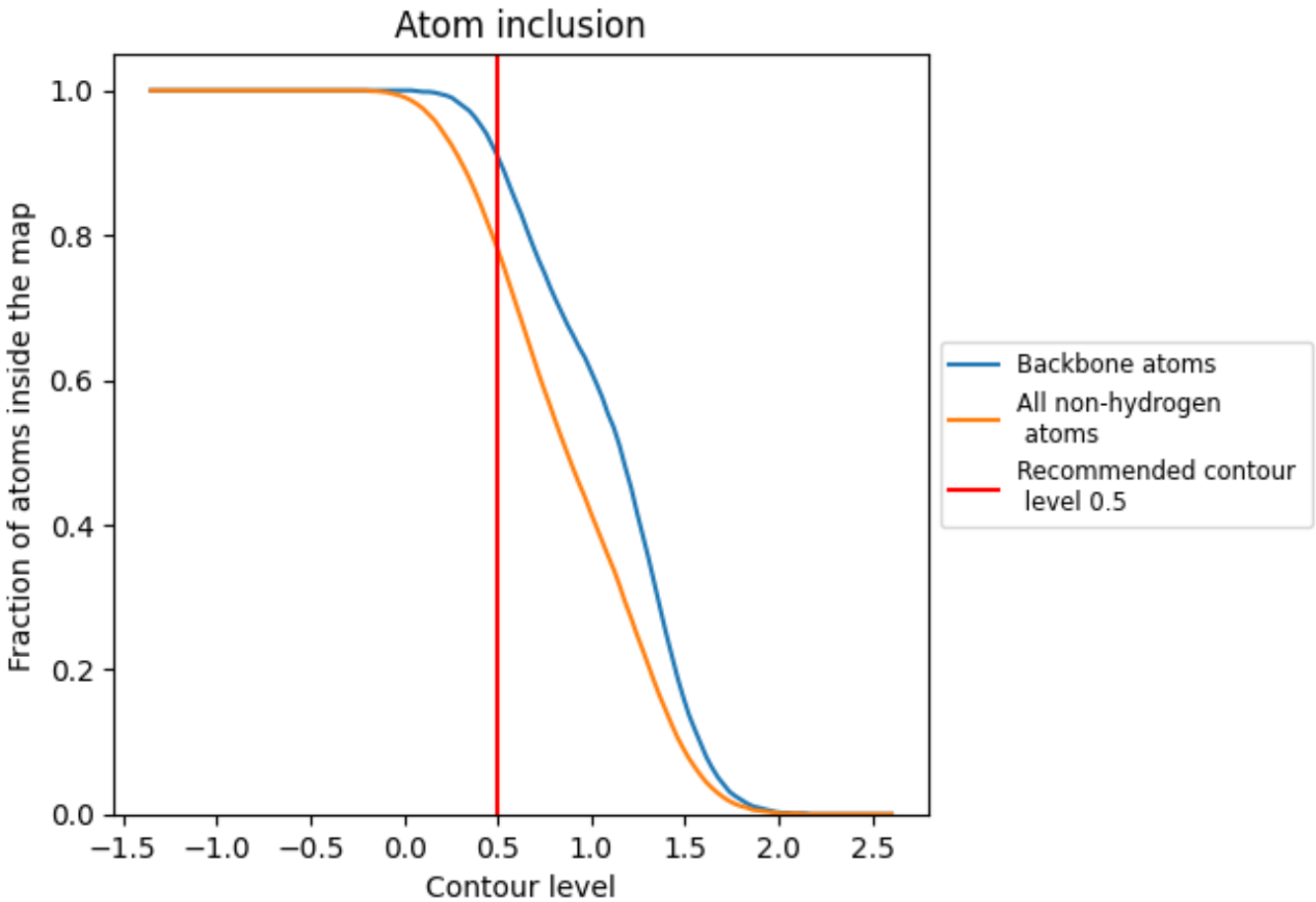
9.3 Atom inclusion mapped to coordinate model (i)



X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.5).

9.4 Atom inclusion (i)



At the recommended contour level, 91% of all backbone atoms, 78% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (0.5) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7810	0.4670
A	0.7720	0.4660
B	0.7780	0.4680
C	0.8210	0.5010
D	0.8240	0.4930
E	0.7630	0.4570
F	0.8330	0.4480
G	0.7690	0.4620
H	0.7740	0.4560
I	0.8090	0.4920
J	0.8350	0.4990
K	0.7740	0.4630
L	0.8530	0.4780
M	0.5920	0.3690
N	0.5310	0.3550
O	0.5000	0.3560
P	0.6650	0.4120
Q	0.6980	0.2890
R	0.6750	0.2600
S	0.7720	0.4630
T	0.7870	0.4620
U	0.8270	0.4990
V	0.8280	0.5020
W	0.7640	0.4580
X	0.8240	0.4380
Y	0.7670	0.4670
Z	0.7890	0.4840
a	0.8350	0.5050
b	0.8400	0.4960
c	0.7740	0.4670
d	0.8040	0.4580
e	0.7790	0.4630
f	0.7800	0.4710
g	0.8300	0.5050
h	0.8090	0.4960

Continued on next page...

Continued from previous page...

Chain	Atom inclusion	Q-score
i	0.7590	0.4550
j	0.8240	0.4480
k	0.7620	0.4600
l	0.8010	0.4760
m	0.8420	0.5030
n	0.8230	0.5010
o	0.7500	0.4540
p	0.8240	0.4500