



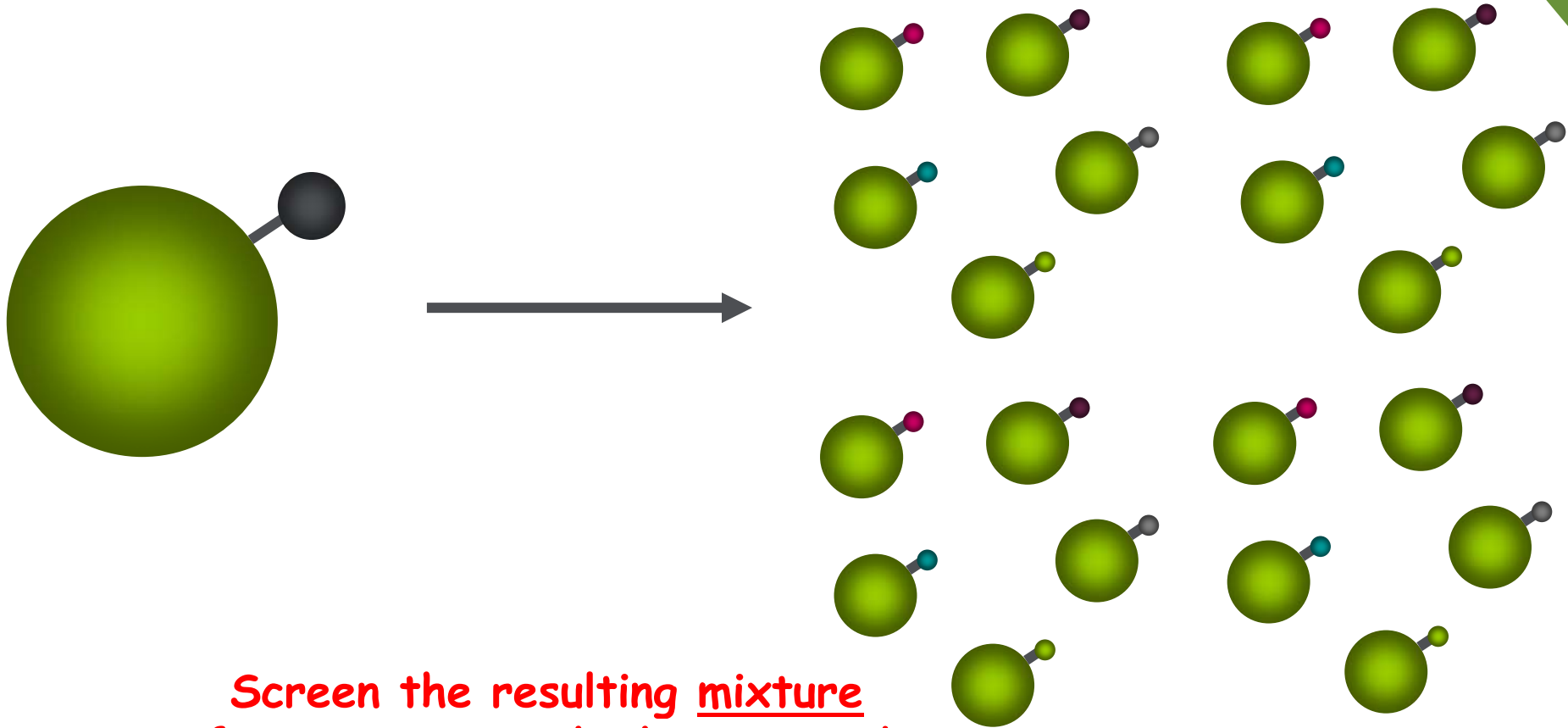
Aston University
Birmingham

Protein engineering: MAX randomisation

Dr. Anna V. Hine



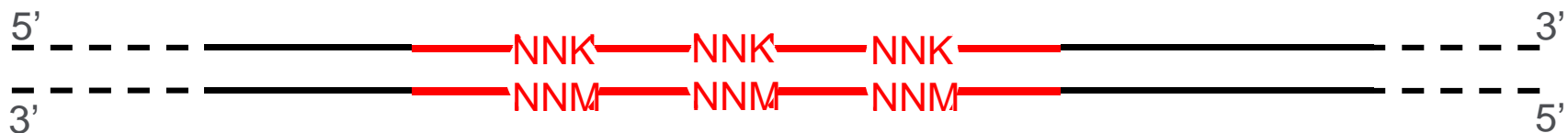
Alternative approach: make a protein library



Screen the resulting mixture
for a protein with the required
activity

Saturation mutagenesis

- ▶ Also known as codon randomisation
- ▶ Used when you KNOW which amino acids are key to protein activity (or at least you can make a good, educated guess)
- ▶ These codon(s) are replaced with randomised codons via cassette mutagenesis eg.



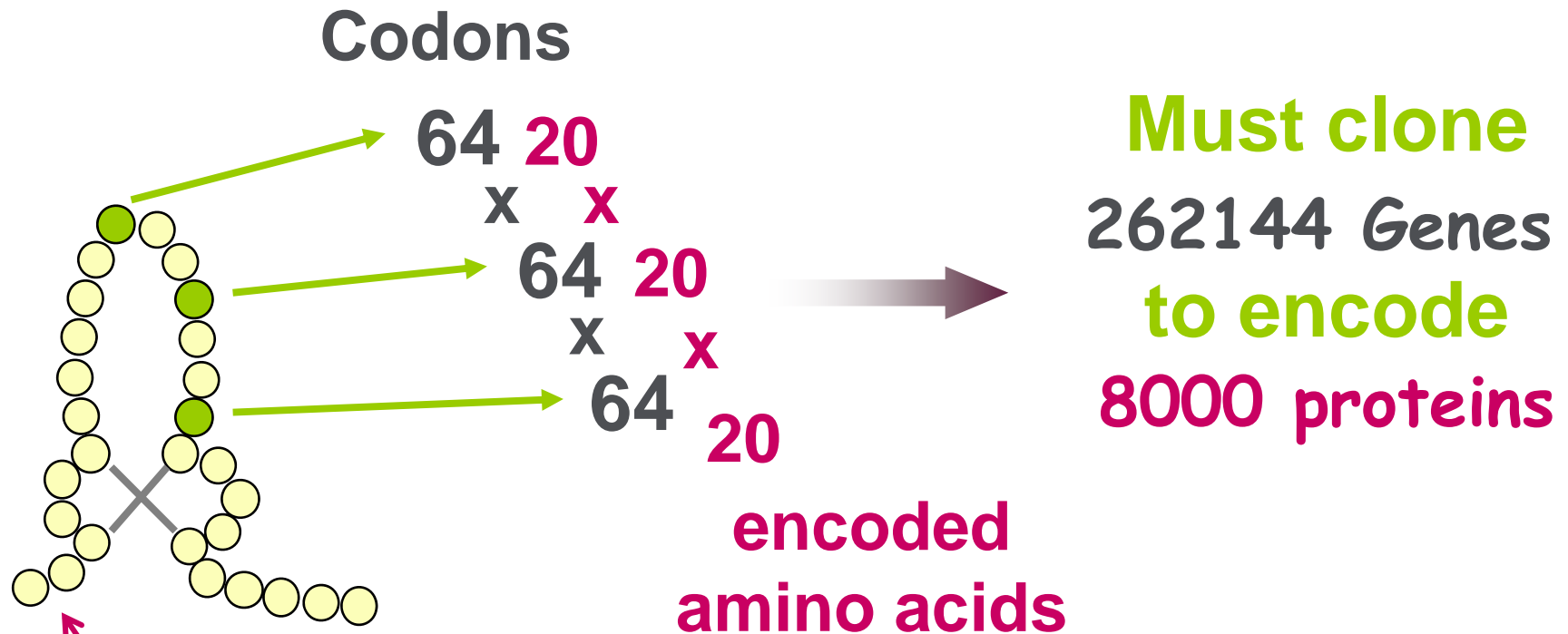
Where N = A/C/G/T
K = G/T,
M = A/C

Designer randomisation – the genetic code

	T	C	A	G
T	TTT Phe (F) TTC " TTA Leu (L) TTG "	TCT Ser (S) TCC " TCA " TCG "	TAT Tyr (Y) TAC TAA Ter TAG Ter	TGT Cys (C) TGC TGA Ter TGG Trp (W)
C	CTT Leu (L) CTC " CTA " CTG "	CCT Pro (P) CCC " CCA " CCG "	CAT His (H) CAC " CAA Gln (Q) CAG "	CGT Arg (R) CGC " CGA " CGG "
A	ATT Ile (I) ATC " ATA " ATG Met (M)	ACT Thr (T) ACC " ACA " ACG "	AAT Asn (N) AAC " AAA Lys (K) AAG "	AGT Ser (S) AGC " AGA Arg (R) AGG "
G	GTT Val (V) GTC " GTA " GTG "	GCT Ala (A) GCC " GCA " GCG "	GAT Asp (D) GAC " GAA Glu (E) GAG "	GGT Gly (G) GGC " GGA " GGG "

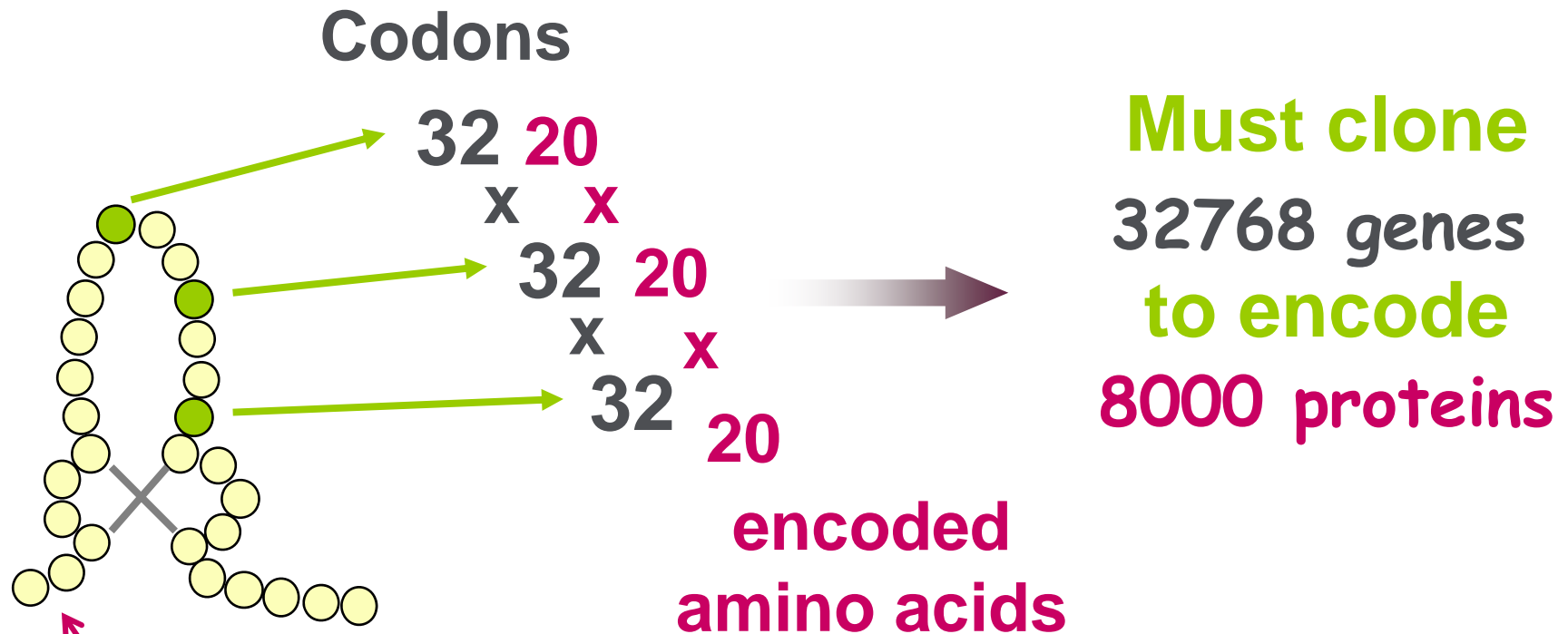
MAX randomisation lets you "pick out" individual codons

Codon randomisation with NNN



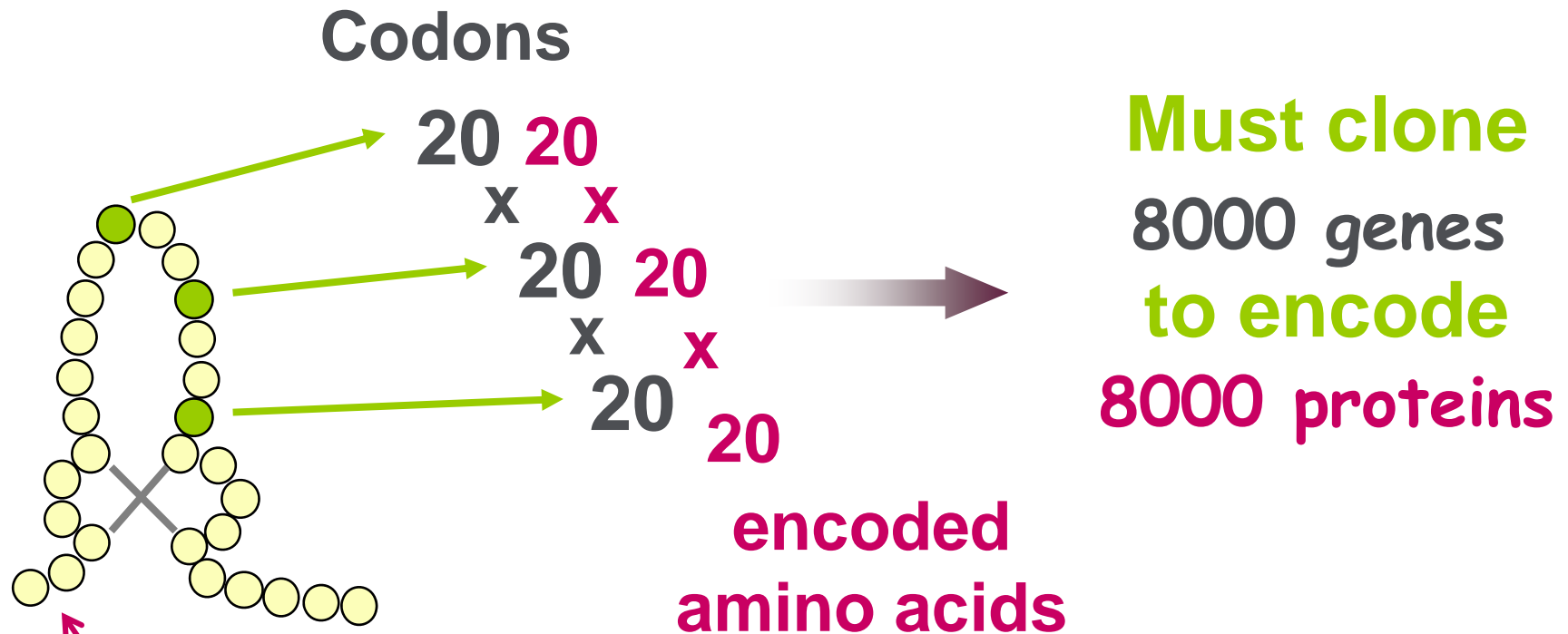
eg. zinc finger protein
(3 aa's key for activity)

Codon randomisation with NNK



eg. zinc finger protein
(3 aa's key for activity)

Codon randomisation with "MAX"



eg. zinc finger protein
(3 aa's key for activity)

The issue of sequence space.....

- ▶ The biggest library that anyone can generate has 10^{10} clones
- ▶ Ideally, you should have enough “sequence space” to have 3 copies of each clone
- ▶ In terms of saturation mutagenesis:

Codon	6 codons	clones	Maximum no of codons within sequence space
NNN	64^6	6.9×10^{10}	≤ 6
NNK	32^6	1.1×10^9	≤ 7
MAX	20^6	64×10^6	8

The screening issue

- ▶ Protein libraries are screened by biopanning
- ▶ Biopanning relies on the law of mass action
- ▶ Which requires even representation within libraries
- ▶ Even representation is not possible with conventional libraries, which contain vast differences in the concentrations of individual proteins
- ▶ MAX randomisation has the power to encode proteins with even representation

Summary

- ▶ MAX randomisation can deliver:
- ▶ Small libraries
- ▶ With bespoke components
- ▶ And no encoded bias