
AB
GRAD
CON23

1
00:00:12,530 --> 00:00:10,450
[Music]

2
00:00:13,850 --> 00:00:12,540
all right hello everyone how's it going

3
00:00:16,010 --> 00:00:13,860
my name's Sean I'm from the University

4
00:00:17,390 --> 00:00:16,020
of Maryland Baltimore County and today

5
00:00:19,189 --> 00:00:17,400
we're going to be looking at xenoamino

6
00:00:21,529 --> 00:00:19,199
acids a look into biochemistry as we

7
00:00:22,849 --> 00:00:21,539
don't know it so I'm going to fly

8
00:00:24,410 --> 00:00:22,859
through a lot of this because it's only

9
00:00:26,210 --> 00:00:24,420
a six minute talk but please come for

10
00:00:28,189 --> 00:00:26,220
feel free to come find me at my poster

11
00:00:30,830 --> 00:00:28,199
outside afterwards so my project

12
00:00:33,530 --> 00:00:30,840
attempts to merge Theory with experiment

13
00:00:35,389 --> 00:00:33,540

in order to test if a alternative amino

14
00:00:37,370 --> 00:00:35,399
acid set is capable of building protein

15
00:00:38,870 --> 00:00:37,380
structure now to do this we need to

16
00:00:41,510 --> 00:00:38,880
select a candidate alphabet from

17
00:00:43,850 --> 00:00:41,520
plausible Alternatives and from there

18
00:00:46,510 --> 00:00:43,860
use our collaborators over at the

19
00:00:49,670 --> 00:00:46,520
Charles University in Prague to build

20
00:00:51,170 --> 00:00:49,680
oligopeptides now I'm going to take a

21
00:00:52,850 --> 00:00:51,180
step back here and go over some very

22
00:00:55,850 --> 00:00:52,860
quick background

23
00:00:57,529 --> 00:00:55,860
um so all life on Earth since Luca has

24
00:00:59,750 --> 00:00:57,539
used the same set of 20 Alpha amino

25
00:01:01,369 --> 00:00:59,760
acids to construct metabolism now

26
00:01:03,110 --> 00:01:01,379
multiple disciplines agree that there

27
00:01:05,750 --> 00:01:03,120
were far more than 20 available for

28
00:01:08,149 --> 00:01:05,760
early evolution in the origins of life

29
00:01:09,170 --> 00:01:08,159
and over the last decade or so there's

30
00:01:11,330 --> 00:01:09,180
been

31
00:01:13,969 --> 00:01:11,340
um theoretical work that identifies a

32
00:01:16,429 --> 00:01:13,979
simple but statistical profile that

33
00:01:18,890 --> 00:01:16,439
distinguishes life's amino acids from

34
00:01:21,469 --> 00:01:18,900
Alternatives and that is that the 20

35
00:01:24,289 --> 00:01:21,479
amino acids used by life exhibit a

36
00:01:26,510 --> 00:01:24,299
non-random coverage of size and

37
00:01:28,670 --> 00:01:26,520
hydrophobicity when we Define coverage

38
00:01:31,310 --> 00:01:28,680

as the range of values in a set and with

39
00:01:33,710 --> 00:01:31,320
how and how they how evenly they

40
00:01:35,450 --> 00:01:33,720
distribute across that range

41
00:01:36,890 --> 00:01:35,460
so what do I mean by that exactly I want

42
00:01:39,530 --> 00:01:36,900
you to consider a set of five amino

43
00:01:42,590 --> 00:01:39,540
acids and pick any physical chemical

44
00:01:45,530 --> 00:01:42,600
descriptor this could be size this could

45
00:01:47,690 --> 00:01:45,540
be volume I'm sorry uh log P this could

46
00:01:49,910 --> 00:01:47,700
be charge and you're going to measure

47
00:01:52,670 --> 00:01:49,920
them rank order them and plot them on

48
00:01:55,490 --> 00:01:52,680
that axis from here what we can do is we

49
00:01:58,310 --> 00:01:55,500
can get the intervals between each amino

50
00:02:00,230 --> 00:01:58,320
acid and this ties into our our idea of

51
00:02:02,030 --> 00:02:00,240
coverage which breaks down into two

52
00:02:04,670 --> 00:02:02,040
components which is range and evenness

53
00:02:07,190 --> 00:02:04,680
so range is simply just the sum of these

54
00:02:09,410 --> 00:02:07,200
intervals between each amino acid and

55
00:02:11,930 --> 00:02:09,420
evenness is the sample variance of these

56
00:02:14,510 --> 00:02:11,940
intervals now this ties back into what I

57
00:02:18,050 --> 00:02:14,520
said the coded 20 amino acids are

58
00:02:20,890 --> 00:02:18,060
statistically non-random in one in about

59
00:02:23,869 --> 00:02:20,900
two and a half million alternative

60
00:02:27,670 --> 00:02:23,879
alphabets would have the same or better

61
00:02:31,250 --> 00:02:27,680
coverage in size and hydrophobicity

62
00:02:34,010 --> 00:02:31,260
now why exactly are we looking at range

63
00:02:36,470 --> 00:02:34,020

and evenness well range allows for a

64
00:02:39,650 --> 00:02:36,480
broader diversity of structures and

65
00:02:41,270 --> 00:02:39,660
functions and evenness allows for the

66
00:02:42,949 --> 00:02:41,280
best approximation of any desired

67
00:02:45,770 --> 00:02:42,959
physical chemistry so I like to think of

68
00:02:48,290 --> 00:02:45,780
it as trying to recreate a black and

69
00:02:51,110 --> 00:02:48,300
white painting using a defined set of

70
00:02:53,869 --> 00:02:51,120
tiles and you need to pick which colors

71
00:02:56,570 --> 00:02:53,879
you want to use to try and get that

72
00:02:58,550 --> 00:02:56,580
painting as close as possible now range

73
00:03:01,130 --> 00:02:58,560
is effectively the difference between

74
00:03:03,410 --> 00:03:01,140
having black and white tiles versus just

75
00:03:04,910 --> 00:03:03,420
different Shades of Gray and evenness is

76
00:03:07,130 --> 00:03:04,920
among the black and white tiles having

77
00:03:08,270 --> 00:03:07,140
an even distribution going from black to

78
00:03:11,390 --> 00:03:08,280
white

79
00:03:13,550 --> 00:03:11,400
now why exactly are we looking at log p

80
00:03:15,710 --> 00:03:13,560
and volume log p is just a measure of

81
00:03:18,350 --> 00:03:15,720
hydrophobicity and hydrophobic collapse

82
00:03:19,850 --> 00:03:18,360
is crucial for protein folding and for

83
00:03:22,790 --> 00:03:19,860
volume volume determines the physical

84
00:03:24,229 --> 00:03:22,800
space that allows for protein folding

85
00:03:26,210 --> 00:03:24,239
so now that we have this idea of

86
00:03:29,030 --> 00:03:26,220
coverage in our mind I would like to

87
00:03:30,110 --> 00:03:29,040
very quickly go over our workflow I'm

88
00:03:33,350 --> 00:03:30,120

going to skip essentially everything

89
00:03:34,850 --> 00:03:33,360
except that top right corner where all

90
00:03:36,770 --> 00:03:34,860
we're doing is we have a heuristic

91
00:03:38,869 --> 00:03:36,780
search protocol that searches a library

92
00:03:41,270 --> 00:03:38,879
of purchasable amino acids

93
00:03:43,490 --> 00:03:41,280
and from there we are in a constant

94
00:03:45,410 --> 00:03:43,500
feedback loop with our empiricist

95
00:03:48,530 --> 00:03:45,420
collaborators in attempting to

96
00:03:51,110 --> 00:03:48,540
eventually get a candidate alphabet

97
00:03:52,610 --> 00:03:51,120
now I'm sure you are all dying to see

98
00:03:54,110 --> 00:03:52,620
what one of these alphabets look like

99
00:03:56,330 --> 00:03:54,120
right

100
00:04:00,649 --> 00:03:56,340
so here we go uh this is our current

101
00:04:02,809 --> 00:04:00,659
alphabet now big emphasis on current

102
00:04:04,309 --> 00:04:02,819
um because of that iterative feedback

103
00:04:05,630 --> 00:04:04,319
loop this is definitely subject to

104
00:04:08,330 --> 00:04:05,640
change and probably will change before

105
00:04:11,509 --> 00:04:08,340
we start constructing oligopeptides from

106
00:04:14,990 --> 00:04:11,519
a set but here is our first set um this

107
00:04:18,170 --> 00:04:15,000
is simply just one example of around 10

108
00:04:20,270 --> 00:04:18,180
to the 14 possible High coverage sets

109
00:04:23,090 --> 00:04:20,280
that we could pick from so there are a

110
00:04:27,170 --> 00:04:23,100
lot of variation that we can change and

111
00:04:29,629 --> 00:04:27,180
play with with the empiricists now again

112
00:04:31,430 --> 00:04:29,639
the goal here is to synthesize these

113
00:04:33,249 --> 00:04:31,440

like a peptides from a completely

114
00:04:36,230 --> 00:04:33,259
non-canonical set

115
00:04:38,090 --> 00:04:36,240
and that's hopefully what we'll be doing

116
00:04:41,870 --> 00:04:38,100
in a few weeks

117
00:04:45,310 --> 00:04:41,880
now I would like to very quickly end on

118
00:04:47,930 --> 00:04:45,320
a deeper note of who cares

119
00:04:51,409 --> 00:04:47,940
why would an advanced science to design

120
00:04:53,450 --> 00:04:51,419
an alternative amino acid set and I

121
00:04:56,510 --> 00:04:53,460
would like to kind of bring that

122
00:04:58,790 --> 00:04:56,520
question to this idea of is there such a

123
00:05:02,150 --> 00:04:58,800
thing as a good or bad amino acid

124
00:05:03,590 --> 00:05:02,160
alphabet now before the empiricists and

125
00:05:05,150 --> 00:05:03,600
the crowd jump out of their chairs and

126
00:05:07,550 --> 00:05:05,160
strangle me

127
00:05:10,129 --> 00:05:07,560
um yes the empiricists are saying of

128
00:05:12,290 --> 00:05:10,139
course there are bad amino acid sets the

129
00:05:13,909 --> 00:05:12,300
20 are very good at what they do but

130
00:05:15,890 --> 00:05:13,919
theorists are pushing back a little bit

131
00:05:19,310 --> 00:05:15,900
and saying well you know Evolution can

132
00:05:21,650 --> 00:05:19,320
find any needle in any Haystack so I

133
00:05:23,749 --> 00:05:21,660
would like to believe that the the truth

134
00:05:25,790 --> 00:05:23,759
lies somewhere in between but what we're

135
00:05:29,090 --> 00:05:25,800
really trying to do here is begin to

136
00:05:31,670 --> 00:05:29,100
characterize good versus bad using an

137
00:05:34,129 --> 00:05:31,680
adaptation of 1980s behavioral thinking

138
00:05:37,790 --> 00:05:34,139

so optimality theory for life's amino

139
00:05:39,890 --> 00:05:37,800
acids now I'd like to conclude just by

140
00:05:43,670 --> 00:05:39,900
thanking our sponsors or I'm sorry our

141
00:05:46,129 --> 00:05:43,680
sponsors are funders

142
00:05:48,529 --> 00:05:46,139
and our collaborators over at the

143
00:05:50,330 --> 00:05:48,539
Charles University in Prague and if you

144
00:05:51,950 --> 00:05:50,340
would like to talk to me about this in

145
00:05:54,350 --> 00:05:51,960
much much greater depth please find me

146
00:05:56,400 --> 00:05:54,360
at poster 24 which is outside thank you

147
00:06:06,290 --> 00:05:56,410
all so much I really appreciate it

148
00:06:06,300 --> 00:06:12,710
[Music]