Michael Wang

Unevolved De Novo *Proteins Have Innate Tendencies to Bind Transition Metals*

1
00:00:00,240 --> 00:00:10,990
[Music]

2
00:00:18,229 --> 00:00:13,759
hello everyone it's the last talk we're

3
00:00:20,540 --> 00:00:18,239
almost done I'll try to keep it fun I'm

4
00:00:22,809 --> 00:00:20,550
going to tell you about de novo proteins

5
00:00:24,769 --> 00:00:22,819
and how they bind transition metals and

6
00:00:27,139 --> 00:00:24,779
the way I'm going to do that is I'm

7
00:00:29,839 --> 00:00:27,149
going to talk to you about what makes

8
00:00:34,040 --> 00:00:29,849
all of us come together our common

9
00:00:36,050 --> 00:00:34,050
ancestry how proteins run can tell us a

10
00:00:39,140 --> 00:00:36,060
little bit about how we might be

11
00:00:41,140 --> 00:00:39,150
different get into the actual bulk of

12
00:00:43,220 --> 00:00:41,150
the science and conclude with some

13
00:00:45,980 --> 00:00:43,230

perspective for the future

14
00:00:49,090 --> 00:00:45,990
so I know we just met but I'd like you

15
00:00:52,250 --> 00:00:49,100
to meet my family these are my ancestors

16
00:00:55,790 --> 00:00:52,260
from a couple hundred well maybe a

17
00:00:58,640 --> 00:00:55,800
hundred years ago and what makes them my

18
00:01:01,160 --> 00:00:58,650
ancestors is that we share DNA

19
00:01:03,740 --> 00:01:01,170
and we share DNA that means we share a

20
00:01:05,359 --> 00:01:03,750
gym type and because the DNA makes us

21
00:01:07,489 --> 00:01:05,369
who we are we share a phenotype so you

22
00:01:12,590 --> 00:01:07,499
look into my eyes and their eyes see the

23
00:01:14,690 --> 00:01:12,600
same neuroses not only do I have

24
00:01:17,300 --> 00:01:14,700
ancestors but we all have ancestors and

25
00:01:20,139 --> 00:01:17,310
our ancestors as humans were related to

26
00:01:23,419 --> 00:01:20,149
other primates ancestors so we are all

27
00:01:26,510 --> 00:01:23,429
sort of connected and if we go a little

28
00:01:28,849 --> 00:01:26,520
further Charles Darwin says not only are

29
00:01:30,889 --> 00:01:28,859
all of us connected evolutionarily to

30
00:01:33,050 --> 00:01:30,899
other primates but you know all

31
00:01:35,090 --> 00:01:33,060
eukaryotes are connected to our ka are

32
00:01:37,989 --> 00:01:35,100
connected to bacteria because there was

33
00:01:40,429 --> 00:01:37,999
this one last Universal common ancestor

34
00:01:42,739 --> 00:01:40,439
we're all connected so that's common

35
00:01:45,940 --> 00:01:42,749
ancestry and where this might become a

36
00:01:49,789 --> 00:01:45,950
problem is when we try to look at the

37
00:01:52,639 --> 00:01:49,799
tips of the tree branches and understand

38
00:01:55,249 --> 00:01:52,649

how life works because all those three

39
00:01:58,849 --> 00:01:55,259
branches trace back to a common stem all

40
00:02:01,730 --> 00:01:58,859
right and we can imagine that if the

41
00:02:03,109 --> 00:02:01,740
Luca looked different then we would all

42
00:02:04,519 --> 00:02:03,119
look different all the bent branches of

43
00:02:06,559 --> 00:02:04,529
the tree would look different so

44
00:02:08,780 --> 00:02:06,569
anything that we conclude by looking at

45
00:02:11,780 --> 00:02:08,790
the recipe has the fingerprints of this

46
00:02:13,780 --> 00:02:11,790
common ancestry in it and that's not an

47
00:02:16,059 --> 00:02:13,790
easy thing to solve because

48
00:02:18,280 --> 00:02:16,069
in considering astrobiology we want to

49
00:02:20,229 --> 00:02:18,290
consider all possibilities of life and

50
00:02:23,170 --> 00:02:20,239
that's sort of hampering our

51
00:02:25,179 --> 00:02:23,180
understanding of all possibilities so

52
00:02:26,979 --> 00:02:25,189
how would we go about getting over a

53
00:02:29,589 --> 00:02:26,989
common ancestry well we've consider

54
00:02:31,180 --> 00:02:29,599
alternative ancestry and you know if we

55
00:02:33,490 --> 00:02:31,190
found another plant with life that

56
00:02:36,490 --> 00:02:33,500
that'd be it we'd be done but we haven't

57
00:02:39,490 --> 00:02:36,500
done that yet so the proposition is we

58
00:02:41,500 --> 00:02:39,500
can use a de novo genotype to find a de

59
00:02:44,830 --> 00:02:41,510
novo to create the de novo phenotype

60
00:02:46,500 --> 00:02:44,840
with new properties or maybe the same

61
00:02:48,490 --> 00:02:46,510
properties maybe life is sort of

62
00:02:50,250 --> 00:02:48,500
constrained in some ways either way

63
00:02:52,509 --> 00:02:50,260

would be informative

64
00:02:54,430 --> 00:02:52,519
that's where protein design comes in we

65
00:02:56,979 --> 00:02:54,440
want to create a de novo phenotype from

66
00:02:59,379 --> 00:02:56,989
a Genova genotype de novo meaning from

67
00:03:02,710 --> 00:02:59,389
scratch something completely new not

68
00:03:03,970 --> 00:03:02,720
related to what came before and when

69
00:03:07,360 --> 00:03:03,980
considering protein design we could

70
00:03:10,059 --> 00:03:07,370
consider a sequence of all amino acids

71
00:03:11,470 --> 00:03:10,069
you know every possibility and what

72
00:03:13,330 --> 00:03:11,480
you'd end up with is a bunch of

73
00:03:16,089 --> 00:03:13,340
insoluble aggregates and it's hard to

74
00:03:18,729 --> 00:03:16,099
study those so we're going to bias the

75
00:03:20,920 --> 00:03:18,739
system by asking first that it form

76
00:03:22,990 --> 00:03:20,930
secondary and tertiary structures so

77
00:03:24,280 --> 00:03:23,000
that it's sort of tractable to work with

78
00:03:26,619 --> 00:03:24,290
and the way we're going to accomplish

79
00:03:28,750 --> 00:03:26,629
that is with binary patterning so binary

80
00:03:30,129 --> 00:03:28,760
patterning is we break all our amino

81
00:03:31,539 --> 00:03:30,139
acids into those that are polar and

82
00:03:34,119 --> 00:03:31,549
those that are non poor those that are

83
00:03:36,129 --> 00:03:34,129
nonpolar want to come together in an

84
00:03:38,439 --> 00:03:36,139
aqueous environment so I've coated the

85
00:03:40,809 --> 00:03:38,449
nonpolar ones here and the hydrophobic

86
00:03:43,150 --> 00:03:40,819
effect drives all those how those

87
00:03:45,369 --> 00:03:43,160
hydrophobic residues into the core of a

88
00:03:46,839 --> 00:03:45,379

protein so if this binary pattern we

89
00:03:50,319 --> 00:03:46,849
haven't looked at the specific residues

90
00:03:52,300 --> 00:03:50,329
we can get this constant secondary and

91
00:03:54,699 --> 00:03:52,310
tertiary structure and because we've

92
00:03:56,920 --> 00:03:54,709
broken them down into you know generic

93
00:03:58,720 --> 00:03:56,930
properties of polarity we get enormous

94
00:04:02,050 --> 00:03:58,730
diversity because this could be you know

95
00:04:03,990 --> 00:04:02,060
any nonpolar amino acid so this could be

96
00:04:05,979 --> 00:04:04,000
any polar amino acid enormous diversity

97
00:04:07,420 --> 00:04:05,989
when we actually make it we're not going

98
00:04:08,649 --> 00:04:07,430
to achieve that enormous diversity it's

99
00:04:10,420 --> 00:04:08,659
something like more than all the

100
00:04:13,149 --> 00:04:10,430
molecules in the universe but a

101
00:04:14,110 --> 00:04:13,159
tractable number is a million may be

102
00:04:16,180 --> 00:04:14,120
nice to look at a million different

103
00:04:18,310 --> 00:04:16,190
genes that aren't affected by common

104
00:04:20,979 --> 00:04:18,320
ancestry so that's what we do we express

105
00:04:26,000 --> 00:04:20,989
it in e.coli and they have no

106
00:04:30,710 --> 00:04:28,130
small limitation because we broke it

107
00:04:33,890 --> 00:04:30,720
down into polar and nonpolar and we want

108
00:04:37,580 --> 00:04:33,900
to do this in genetically driven way eco

109
00:04:39,380 --> 00:04:37,590
lies make our proteins for us we limited

110
00:04:41,660 --> 00:04:39,390
our selection of amino acids to those

111
00:04:44,240 --> 00:04:41,670
that were kind of easy so if the middle

112
00:04:48,350 --> 00:04:44,250
position of the three nuclear bases that

113
00:04:52,460 --> 00:04:48,360

code for an amino acid is T it'll be

114
00:04:56,690 --> 00:04:52,470
nonpolar and if it's a it will be polar

115
00:04:59,450 --> 00:04:56,700
so NTN that triplet gives us any of five

116
00:05:01,400 --> 00:04:59,460
nonpolar amino acids and then and gives

117
00:05:02,750 --> 00:05:01,410
us any of six polar amino acids that's

118
00:05:07,220 --> 00:05:02,760
sort of an easy way to build this

119
00:05:10,130 --> 00:05:07,230
diversity so we build a library there's

120
00:05:11,600 --> 00:05:10,140
a million of them what can they do well

121
00:05:14,570 --> 00:05:11,610
I'm not going to be telling you about

122
00:05:17,210 --> 00:05:14,580
this my labs done this but I think we

123
00:05:18,440 --> 00:05:17,220
can replace essential genes I'll say

124
00:05:20,570 --> 00:05:18,450
that again we can replace the central

125
00:05:22,700 --> 00:05:20,580
genes we can take a binary pattern

126
00:05:25,730 --> 00:05:22,710
design protein with no common ancestry

127
00:05:29,870 --> 00:05:25,740
put it into an e coli that's going to

128
00:05:31,760 --> 00:05:29,880
die and it lives that's kind of weird

129
00:05:33,380 --> 00:05:31,770
we can also gene expression levels

130
00:05:36,200 --> 00:05:33,390
there's some interaction with RNA there

131
00:05:37,730 --> 00:05:36,210
and we can evolve this sort of early

132
00:05:43,910 --> 00:05:37,740
functional protein into something that's

133
00:05:47,060 --> 00:05:43,920
more functional and better this sort of

134
00:05:49,310 --> 00:05:47,070
system is based on amino acids alone and

135
00:05:50,960 --> 00:05:49,320
that's not the scope of all protein

136
00:05:52,070 --> 00:05:50,970
function that we find in nature and

137
00:05:54,620 --> 00:05:52,080
that's not the scope of all protein

138
00:05:56,780 --> 00:05:54,630

function that's interesting so the idea

139
00:05:58,550 --> 00:05:56,790
that I had was we're gonna add metals

140
00:05:59,990 --> 00:05:58,560
maybe it'll bind the metals maybe

141
00:06:01,880 --> 00:06:00,000
they'll give it an additional function

142
00:06:04,250 --> 00:06:01,890
and a side benefit of this is if we

143
00:06:06,680 --> 00:06:04,260
started with ten to the six a million

144
00:06:08,150 --> 00:06:06,690
and we added two metals we went you know

145
00:06:11,960 --> 00:06:08,160
two times ten to the six we're getting

146
00:06:14,360 --> 00:06:11,970
additional diversity and the reason

147
00:06:16,940 --> 00:06:14,370
those are interesting is because they

148
00:06:19,430 --> 00:06:16,950
perform a lot of key functions and life

149
00:06:21,920 --> 00:06:19,440
as we know it so some examples are

150
00:06:24,050 --> 00:06:21,930
oxidation and reduction small molecule

151
00:06:26,660 --> 00:06:24,060
binding sort of biologically important

152
00:06:28,400 --> 00:06:26,670
functions and it's not just biologically

153
00:06:31,250 --> 00:06:28,410
important functions that we find in

154
00:06:33,020 --> 00:06:31,260
modern organisms if we trace back that

155
00:06:35,810 --> 00:06:33,030
Luca the last Universal common ancestor

156
00:06:37,340 --> 00:06:35,820
that we all share iron was essential for

157
00:06:38,900 --> 00:06:37,350
some of the functions that it was doing

158
00:06:41,390 --> 00:06:38,910
so if we were

159
00:06:43,430 --> 00:06:41,400
operating medals and asking random or

160
00:06:45,410 --> 00:06:43,440
sort of semi random sequences to do

161
00:06:46,460 --> 00:06:45,420
those functions then those are important

162
00:06:48,920 --> 00:06:46,470
functions that we'd like to see

163
00:06:53,930 --> 00:06:48,930

recapitulated without that common

164
00:06:56,990 --> 00:06:53,940
ancestry bias so there's steps to get to

165
00:06:59,270 --> 00:06:57,000
that big answer and the big question we

166
00:07:02,770 --> 00:06:59,280
have to answer is can we take this

167
00:07:05,210 --> 00:07:02,780
binary pattern sequence add medals and

168
00:07:06,530 --> 00:07:05,220
then hopefully get function afterwards

169
00:07:09,740 --> 00:07:06,540
right but there's a prerequisite step

170
00:07:10,850 --> 00:07:09,750
that they stick to the medals so to get

171
00:07:13,970 --> 00:07:10,860
at that question we're just going to

172
00:07:16,060 --> 00:07:13,980
pick 52 of them express them and see if

173
00:07:18,440 --> 00:07:16,070
they bind some representative medals and

174
00:07:20,810 --> 00:07:18,450
if they do then we'd like to know how

175
00:07:22,550 --> 00:07:20,820
that works sort of characterization I'll

176
00:07:24,320 --> 00:07:22,560
mention at this point that this is all

177
00:07:26,210 --> 00:07:24,330
published it's under the same title as

178
00:07:29,630 --> 00:07:26,220
this talk so if you're interested you

179
00:07:32,720 --> 00:07:29,640
can read further but this is we'll start

180
00:07:35,990 --> 00:07:32,730
out with the qualitative screen of 52

181
00:07:39,440 --> 00:07:36,000
and again what we're looking for is does

182
00:07:40,850 --> 00:07:39,450
the protein stick to a metal very simple

183
00:07:43,340 --> 00:07:40,860
questions so we're going to immobilize

184
00:07:44,750 --> 00:07:43,350
the metal we're gonna add the protein

185
00:07:47,000 --> 00:07:44,760
and if they stick we should be able to

186
00:07:49,820 --> 00:07:47,010
detect it so this is the important part

187
00:07:52,460 --> 00:07:49,830
of a protein gel so if you see the

188
00:07:54,920 --> 00:07:52,470

protein there that means we were able to

189
00:07:57,500 --> 00:07:54,930
detected in some part of the sample so

190
00:07:59,420 --> 00:07:57,510
we had it hopefully it sticks and this

191
00:08:00,890 --> 00:07:59,430
is how much protein got added so there

192
00:08:04,310 --> 00:08:00,900
was a bunch of protein added the thicker

193
00:08:06,170 --> 00:08:04,320
the band the more protein we can wash

194
00:08:08,240 --> 00:08:06,180
away all the stuff that didn't stick you

195
00:08:11,560 --> 00:08:08,250
know all the endogenous proteins those

196
00:08:14,270 --> 00:08:11,570
sorts of things and it didn't come off

197
00:08:16,760 --> 00:08:14,280
didn't come off that's good because when

198
00:08:19,250 --> 00:08:16,770
we eventually wash it off I mean we rate

199
00:08:21,500 --> 00:08:19,260
it by the percent retained we can see

200
00:08:23,750 --> 00:08:21,510
that this protein that I added here in

201
00:08:26,360 --> 00:08:23,760
this example stuck very well to the

202
00:08:29,090 --> 00:08:26,370
beads so we washed off all the non

203
00:08:31,400 --> 00:08:29,100
binding stuff and this was specifically

204
00:08:34,969 --> 00:08:31,410
balanced there was some interaction of

205
00:08:36,800 --> 00:08:34,979
our protein with a metal and that's

206
00:08:39,770 --> 00:08:36,810
promising for looking for later metal

207
00:08:42,589 --> 00:08:39,780
dependent functionality so we did this

208
00:08:45,020 --> 00:08:42,599
with 52 different things and to our

209
00:08:48,290 --> 00:08:45,030
surprise a lot of them stuck in fact

210
00:08:48,800 --> 00:08:48,300
almost all of them stuck we wanted to do

211
00:08:51,680 --> 00:08:48,810
that

212
00:08:52,600 --> 00:08:51,690
sticking quantitatively though so it

213
00:08:55,150 --> 00:08:52,610

wasn't just you know

214
00:08:56,980 --> 00:08:55,160
it stick to a bead yes or no wanted that

215
00:08:59,740 --> 00:08:56,990
to be quantitative so we did this with

216
00:09:00,579 --> 00:08:59,750
equilibrium dialysis so these hexa our

217
00:09:02,740 --> 00:09:00,589
Pentagon's

218
00:09:05,590 --> 00:09:02,750
are the protein and there's some free

219
00:09:08,139 --> 00:09:05,600
metal floating in solution and there's a

220
00:09:11,319 --> 00:09:08,149
membrane in between the two that the

221
00:09:15,280 --> 00:09:11,329
metal can transfer through but the

222
00:09:18,340 --> 00:09:15,290
protein can't so we can take some from

223
00:09:19,840 --> 00:09:18,350
this side measure the free metal and add

224
00:09:21,550 --> 00:09:19,850
more metal and that will diffuse through

225
00:09:24,100 --> 00:09:21,560
and bind to the protein and if we keep

226
00:09:25,240 --> 00:09:24,110
iterating that over and over again and

227
00:09:28,870 --> 00:09:25,250
measuring the free metal we can

228
00:09:30,790 --> 00:09:28,880
interpolate how much was you know we

229
00:09:33,069 --> 00:09:30,800
could pull out because it wound up stuck

230
00:09:36,040 --> 00:09:33,079
to the protein so we can do this

231
00:09:37,569 --> 00:09:36,050
quantitatively and I'll just sort of

232
00:09:41,620 --> 00:09:37,579
summarize the results from all these

233
00:09:43,389 --> 00:09:41,630
studies so what we saw was many possible

234
00:09:44,860 --> 00:09:43,399
binding sites it wasn't just one metal

235
00:09:47,980 --> 00:09:44,870
per protein like you might find in

236
00:09:49,840 --> 00:09:47,990
evolved systems we saw many possible

237
00:09:51,819 --> 00:09:49,850
binding sites and many of them ended up

238
00:09:54,250 --> 00:09:51,829

being used and that comes down to a

239
00:09:55,660 --> 00:09:54,260
binary patterning approach there's two

240
00:09:57,790 --> 00:09:55,670
residues in the binary patterning

241
00:10:01,750 --> 00:09:57,800
approach that can bind metals they're

242
00:10:04,389 --> 00:10:01,760
histidine and carboxylic acid containing

243
00:10:07,660 --> 00:10:04,399
residues both of which combine those

244
00:10:08,860 --> 00:10:07,670
with different affinities and there is a

245
00:10:10,680 --> 00:10:08,870
bunch of ways that this could work so

246
00:10:13,210 --> 00:10:10,690
they could have it on one helix sort of

247
00:10:14,680 --> 00:10:13,220
over here and one he looks over here and

248
00:10:17,259 --> 00:10:14,690
between the two there's a binding site

249
00:10:19,509 --> 00:10:17,269
it could all be shared on long helix it

250
00:10:20,889 --> 00:10:19,519
could be on the turns there's a lot of

251
00:10:22,389 --> 00:10:20,899
different places this could bind and we

252
00:10:25,930 --> 00:10:22,399
ended up seeing a lot of different

253
00:10:27,550 --> 00:10:25,940
binding sites the screening results I

254
00:10:30,519 --> 00:10:27,560
said a bunch of them in fact almost all

255
00:10:31,990 --> 00:10:30,529
of them stuck so here's that on sort of

256
00:10:33,310 --> 00:10:32,000
a general look so what we want to look

257
00:10:36,759 --> 00:10:33,320
at is over here the protein of interest

258
00:10:38,740 --> 00:10:36,769
it's around ten kilotons so this is a

259
00:10:41,170 --> 00:10:38,750
protein it's called us at twenty four we

260
00:10:44,230 --> 00:10:41,180
add this much protein we wash away

261
00:10:47,019 --> 00:10:44,240
everything that doesn't bind and then we

262
00:10:49,509 --> 00:10:47,029
have see what bound to zinc and so

263
00:10:51,790 --> 00:10:49,519

there's something that bounces neck we

264
00:10:53,350 --> 00:10:51,800
cut that histidine the number of

265
00:10:56,170 --> 00:10:53,360
histidines honest at twenty four and a

266
00:10:57,579 --> 00:10:56,180
half and a lot of them stop binding to

267
00:10:59,769 --> 00:10:57,589
the metal so the hissings

268
00:11:02,800 --> 00:10:59,779
were as we expected important for the

269
00:11:03,460 --> 00:11:02,810
binding and if we get rid of all the

270
00:11:04,110 --> 00:11:03,470
histidines

271
00:11:08,370 --> 00:11:04,120
then not

272
00:11:10,530 --> 00:11:08,380
sticks um I said there was 52 so this is

273
00:11:12,510 --> 00:11:10,540
the 11th one sort of showing that it's

274
00:11:15,720 --> 00:11:12,520
representative many of them bind the

275
00:11:18,269 --> 00:11:15,730
20th one also sticks and an interesting

276
00:11:20,460 --> 00:11:18,279
comparison is these are all the protein

277
00:11:21,840 --> 00:11:20,470
of interest but you can take this bead

278
00:11:24,180 --> 00:11:21,850
experiment and look at the whole

279
00:11:27,840 --> 00:11:24,190
proteome that you can see on a gel of

280
00:11:30,810 --> 00:11:27,850
e.coli and so the whole all the coli

281
00:11:32,910 --> 00:11:30,820
proteins that we add are there you know

282
00:11:35,160 --> 00:11:32,920
there's thicker band's thinner bands but

283
00:11:37,110 --> 00:11:35,170
it's a lot of proteins if you do the

284
00:11:40,380 --> 00:11:37,120
same see what sticks to metal and see

285
00:11:41,670 --> 00:11:40,390
what comes off not much sticks to metal

286
00:11:44,790 --> 00:11:41,680
so there's something different between

287
00:11:46,650 --> 00:11:44,800
our proteins where it's abundant and

288
00:11:50,130 --> 00:11:46,660

evolved proteins where this sort of

289
00:11:52,260 --> 00:11:50,140
sticking isn't we did the equilibrium

290
00:11:55,019 --> 00:11:52,270
dialysis we quantify the affinity it was

291
00:11:57,950 --> 00:11:55,029
between 100 and animal or 3 micro molar

292
00:12:00,690 --> 00:11:57,960
if that means anything to you

293
00:12:02,010 --> 00:12:00,700
stoichiometry I said there's multiple

294
00:12:03,720 --> 00:12:02,020
possible binding sites in there are

295
00:12:06,420 --> 00:12:03,730
multiple actual binding sites so it's 1

296
00:12:08,480 --> 00:12:06,430
to 4 binding sites per protein so this

297
00:12:11,340 --> 00:12:08,490
one will bind you to equivalents of

298
00:12:14,610 --> 00:12:11,350
cobalt that will bind four equivalents

299
00:12:17,070 --> 00:12:14,620
of zinc and more interesting is it's

300
00:12:19,710 --> 00:12:17,080
specific so a single protein might bind

301
00:12:21,810 --> 00:12:19,720
one equivalent of cobalt but more

302
00:12:25,079 --> 00:12:21,820
equivalents of zinc suggesting that some

303
00:12:28,530 --> 00:12:25,089
of these unselected undesigned binding

304
00:12:30,329 --> 00:12:28,540
sites are being are preferentially

305
00:12:32,070 --> 00:12:30,339
binding one metal and not another so

306
00:12:34,949 --> 00:12:32,080
there's some specificity going on here

307
00:12:36,870 --> 00:12:34,959
and in summary because we're building

308
00:12:38,820 --> 00:12:36,880
towards you know the function metal

309
00:12:41,449 --> 00:12:38,830
binding is ubiquitous it's weak it

310
00:12:45,620 --> 00:12:41,459
occurs in multiple locations in our

311
00:12:47,790 --> 00:12:45,630
hands in this in this library and

312
00:12:51,990 --> 00:12:47,800
because we don't see that same

313
00:12:53,820 --> 00:12:52,000

stickiness in modern proteins and might

314
00:12:55,980 --> 00:12:53,830
suggest that they evolved away from this

315
00:12:59,250 --> 00:12:55,990
sort of permissive binding but that's

316
00:13:02,550 --> 00:12:59,260
sort of speculative hard to prove so if

317
00:13:04,140 --> 00:13:02,560
natural systems great proteins that have

318
00:13:08,150 --> 00:13:04,150
functions that make cats and dogs

319
00:13:11,010 --> 00:13:08,160
the alternative seems less specific

320
00:13:12,810 --> 00:13:11,020
bindings everywhere and hard to say what

321
00:13:15,360 --> 00:13:12,820
it'll do and that's sort of what I'd

322
00:13:16,160 --> 00:13:15,370
like to segue into this is sort of

323
00:13:18,500 --> 00:13:16,170
tentative

324
00:13:20,510 --> 00:13:18,510
publish stuff but if there's an

325
00:13:23,840 --> 00:13:20,520
important function in life it might be

326
00:13:26,030 --> 00:13:23,850
the use of ATP right because ATP is the

327
00:13:28,730 --> 00:13:26,040
currency of the cell and we've been

328
00:13:32,360 --> 00:13:28,740
hearing RNAi talks ATP hydrolysis is

329
00:13:36,920 --> 00:13:32,370
important reaction and metalloproteins

330
00:13:38,930 --> 00:13:36,930
can be used to do this and tentatively

331
00:13:40,010 --> 00:13:38,940
very tentatively one of the

332
00:13:43,550 --> 00:13:40,020
metalloproteins

333
00:13:46,280 --> 00:13:43,560
seems to hydrolyze ATP and do it better

334
00:13:50,000 --> 00:13:46,290
in the presence of magnesium so

335
00:13:53,930 --> 00:13:50,010
tentatively this might be a result of

336
00:13:55,970 --> 00:13:53,940
the middle button future directions so

337
00:13:58,340 --> 00:13:55,980
can we create useful functional

338
00:14:00,470 --> 00:13:58,350

metalloproteins without ancestry can we

339
00:14:03,220 --> 00:14:00,480
evolve them to be comparable to the

340
00:14:05,330 --> 00:14:03,230
things we see today can we see how

341
00:14:06,620 --> 00:14:05,340
they're different from the things we see

342
00:14:08,300 --> 00:14:06,630
today because if they're different but

343
00:14:10,220 --> 00:14:08,310
to let's say the same function that

344
00:14:12,400 --> 00:14:10,230
would highlight this difference between

345
00:14:14,450 --> 00:14:12,410
common ancestry and alternative ancestry

346
00:14:16,760 --> 00:14:14,460
and sort of a question I'd like to open

347
00:14:18,590 --> 00:14:16,770
up to everyone here is there's a lot of

348
00:14:20,030 --> 00:14:18,600
reactions that I don't realize are

349
00:14:21,970 --> 00:14:20,040
important but I'm sure everyone here is

350
00:14:24,230 --> 00:14:21,980
more knowledgeable and so if you have

351
00:14:27,680 --> 00:14:24,240
you know the reaction that you're

352
00:14:33,650 --> 00:14:27,690
feeling patriotic about let me know and

353
00:14:36,020 --> 00:14:33,660
I'll convert to your country can't end

354
00:14:38,900 --> 00:14:36,030
without acknowledging the work the

355
00:14:44,950 --> 00:14:38,910
fellow members of the heck lab and thank

356
00:15:00,920 --> 00:14:57,970
awesome questions for Mike hey Great Auk

357
00:15:03,650 --> 00:15:00,930
two things what are the lengths of the

358
00:15:05,390 --> 00:15:03,660
proteins that you were using a hundred

359
00:15:08,840 --> 00:15:05,400
and two amino acids okay so that's not

360
00:15:12,140 --> 00:15:08,850
very long in terms of the proteins that

361
00:15:16,160 --> 00:15:12,150
we have in ourselves today yeah it's

362
00:15:18,670 --> 00:15:16,170
sort of it's sort of the the minimum to

363
00:15:21,050 --> 00:15:18,680

be a protein and not sort of a peptide

364
00:15:24,320 --> 00:15:21,060
there's debate about Erica's insulins of

365
00:15:26,690 --> 00:15:24,330
protein but yeah they're small and the

366
00:15:30,080 --> 00:15:26,700
other thing is I was wondering what is

367
00:15:33,380 --> 00:15:30,090
your hypothesis on why the proteins that

368
00:15:36,170 --> 00:15:33,390
are natural to eco like what they didn't

369
00:15:38,150 --> 00:15:36,180
bind to the metals like I have my theory

370
00:15:41,090 --> 00:15:38,160
but I'm wondering what you think

371
00:15:43,580 --> 00:15:41,100
I'd like to hear your theory but my

372
00:15:50,300 --> 00:15:43,590
hypothesis is sort of the reverse

373
00:15:51,920 --> 00:15:50,310
pressure metal concentrations inside of

374
00:15:54,950 --> 00:15:51,930
the cell are tightly regulated right you

375
00:15:56,660 --> 00:15:54,960
want only the chemistry that you want to

376
00:15:59,720 --> 00:15:56,670
have happen to happen and if there is

377
00:16:04,370 --> 00:15:59,730
some metal binding site that is on the

378
00:16:06,680 --> 00:16:04,380
surface a it might mediate a protein

379
00:16:08,180 --> 00:16:06,690
protein interaction and so two things

380
00:16:11,420 --> 00:16:08,190
with that surface binding I'll stick

381
00:16:14,240 --> 00:16:11,430
together and be it might just mess with

382
00:16:17,920 --> 00:16:14,250
your whole regulation system but we can

383
00:16:25,900 --> 00:16:22,640
hi great oh so we heard that the sizes

384
00:16:30,830 --> 00:16:25,910
are around a hundred and that you get

385
00:16:33,830 --> 00:16:30,840
multiple binding sites per protein so on

386
00:16:38,060 --> 00:16:33,840
one is attainable for you to make them

387
00:16:43,250 --> 00:16:38,070
smaller so that you could get one

388
00:16:46,760 --> 00:16:43,260

binding site per protein it is doable to

389
00:16:48,470 --> 00:16:46,770
make them smaller they become more

390
00:16:50,900 --> 00:16:48,480
disordered the smaller the hydrophobic

391
00:16:54,920 --> 00:16:50,910
chorus and so getting structures on

392
00:16:55,550 --> 00:16:54,930
smaller versions is not something that

393
00:17:06,310 --> 00:16:55,560
we've been

394
00:17:12,220 --> 00:17:09,100
um just to clarify something you started

395
00:17:15,780 --> 00:17:12,230
from talking about how protein design is

396
00:17:19,870 --> 00:17:15,790
at such a point that it can even replace

397
00:17:22,450 --> 00:17:19,880
critical functional proteins and things

398
00:17:24,430 --> 00:17:22,460
like e.coli mmm at the same time about

399
00:17:27,220 --> 00:17:24,440
metalloproteins you know you ended with

400
00:17:29,110 --> 00:17:27,230
but saying that perhaps we can make

401
00:17:31,750 --> 00:17:29,120
something that would have this this

402
00:17:36,310 --> 00:17:31,760
critical function and so is it the case

403
00:17:38,290 --> 00:17:36,320
that it is harder to make Nutella

404
00:17:40,240 --> 00:17:38,300
proteins that would be functional and

405
00:17:41,350 --> 00:17:40,250
the the things that you had in mind at

406
00:17:43,450 --> 00:17:41,360
the beginning just didn't have these

407
00:17:47,220 --> 00:17:43,460
metal cofactors or what's the difference

408
00:17:49,330 --> 00:17:47,230
I think because we haven't specified

409
00:17:51,970 --> 00:17:49,340
from the start that these are designed

410
00:17:54,040 --> 00:17:51,980
to bind a certain metal that in the cell

411
00:17:56,740 --> 00:17:54,050
it's sort of a toss-up if Li well unless

412
00:18:00,820 --> 00:17:56,750
we go that further design step inside

413
00:18:02,920 --> 00:18:00,830

the cell whereas what we did design for

414
00:18:06,850 --> 00:18:02,930
which was just structure was enough to

415
00:18:08,350 --> 00:18:06,860
do certain other functions so improving

416
00:18:11,710 --> 00:18:08,360
the metal binding will make it easier to

417
00:18:15,990 --> 00:18:11,720
then do the next step of metal

418
00:18:21,010 --> 00:18:16,000
dependence in vivo function thank you

419
00:18:22,660 --> 00:18:21,020
great uh any last burning questions all