怎么测序抗生素？

目前为止，抗生素对治疗细菌感染是不能替代的。从抗生素的发现到现在，抗生素拯救了众多生命。但是当今的细菌在发展抗生素抗性——换句话说，抗生素的治疗效用在降低。因此，很多的科学家们在努力研究抗生素，让他们培养新的抗生素。但是想研究抗生素的人常遇到一些特别难的障碍。

一般而言，细胞的蛋白质（多肽）是按照细胞的DNA序列制造的。先转录DNA制造RNA，然后用这些转录的代码制造多肽。但是这些抗生素算是一个例外。他们是“非核糖体肽”，不会依赖于DNA序列被制造的。因此，生物信息学家不能依赖于DNA序列预测抗生素的氨基酸序列。此外，目前分析多肽的方式都专注在线性多肽，但是大部分的抗生素是环肽。因此，包括环肽序列的数据特别少，不能依赖于氨基酸数据来测序环肽。

因为环肽很独特，科学家们需要找到一个从头测序环肽的方式。最理想测序环肽的方式是使用质谱分析去做。简单来说，质谱分析会破坏这些环肽，分成几个片段，然后称片的重量。质谱分析比较便宜，也很快。但是用质谱分析会遇到两个问题。其一，质谱分析的数据不完美，会包括很多的假数据。其二，质谱数据只能包括环肽所有片段的一小部分（可能三分之一的片段）。

以下是一个例子。Tyrocidine B1就是一个抗生素。它的序列很简单，是以下的测序：

Val\* – Lys – Leu – Phe – Pro – Trp – Phe – Asn – Gln – Tyr

这个是一个环肽，所以Val不一定是第一个氨基酸。它的缩写序列是“VKLFPWFNQY”。

“VKL”的片段重量是V+K+L的重量，99.06+113.08+128.09=340.23,所以质谱分析的结果可能会包括340.23的数字。TyrocidineB1的总体重量等于1322左右。 使用质谱分析的方式制造理想TyrocidineB1的所有片段有89个。但是只有30个左右会出现，而且大概70跟TyrocidineB1没有关系的重量也会出现。生物信息学家们还没有培养分别真假数据的算法，也还没有培养用这些数据来测序环肽的序列。如果可以的话，科学家们对抗生素的理解会有明显的帮助，可以开始考虑怎么重新制造抗生素。

“VKLFPWFNQY”。