# Robert T. Batey教授讲座笔记

**A mechanistic perspective on how small molecules drive mRNA expression by riboswitches**

陈柯，黄林课题组

会议回顾

2024年7月10日，Robert T. Batey教授受中山大学孙逸仙纪念医院黄林研究员邀在Guangzhou RNA club进行题为“A mechanistic perspective on how small molecules drive mRNA expression by riboswitches”的线上学术报告。Robert T. Batey教授是著名结构生物学家，就职于美国科罗拉多大学化学与生物化学系。他长期从事于RNA调控基因表达的结构与机制研究，对相关领域做出了许多突出的贡献。值得一提的是Robert T. Batey教授在过去的二十多年时间中深耕于riboswitch领域，不仅参与解析了如SAM、purine family等多种重要的riboswitches，还深入探究如何借鉴已有的经验进一步将riboswitch运用到更加广阔的范围，如在天然riboswitch适体域的结构基础上建立文库从而筛选得到既有小分子结合位点又保留了骨架结构的重组适配体，在体内外均可发挥较好的作用。

在已有的研究中，riboswitch适体域与小分子结合形成的复合物结构已经成为该领域的重要研究模型，但对于小分子结合后具体是如何调控下游基因表达的机制探究相对较少，Robert T. Batey教授将通过此次讲座阐述小分子结合后如何介导下游mRNA的相关表达过程。

会议内容

riboswitch是mRNA中结合特定配体的调控元件，主要位于细菌mRNA 5'-untranslated region (5'-UTR)区。通常通过控制转录或翻译将小分子的结合与基因表达调节相结合。核糖开关具有两部分区域，包括适体域(aptamer domain)和表达平台(expression platform)。适体域充当特异性结合配体的受体。表达平台通过其在响应配体结合时产生的构象变化进而调控基因的转录、翻译等重要表达过程。2007年Wachter等人的研究发现并非所有核糖开关都位于原核生物的 5'-UTR部分，某些真核生物mRNA中，硫胺素焦磷酸 (TPP) 核糖开关调节3'端的剪接。

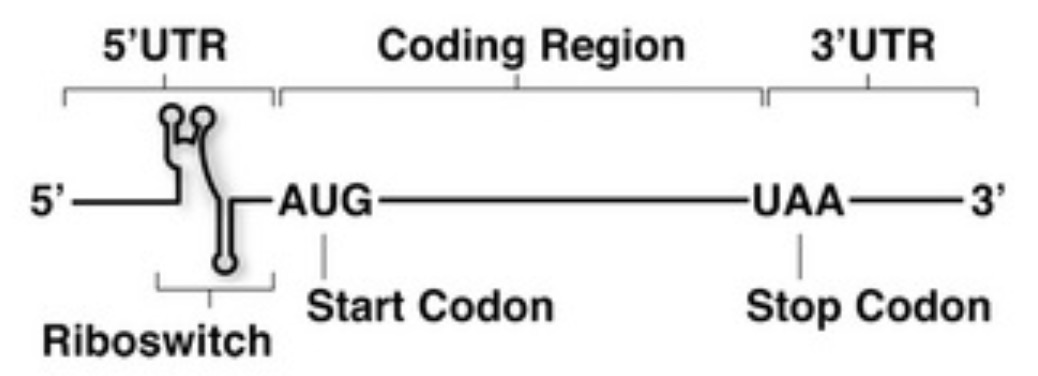


图1 mRNA的简要结构

以图2中调节转录的riboswitch为例，当配体未结合时，处于适体域和表达平台间的switching sequence（红色序列）位于抗终止子茎环中，此时转录可以正常进行。但当配体结合时，switching sequence会整合到适体域中，表达平台的构象也随之改变，折叠成终止子茎环的形态，导致转录终止。

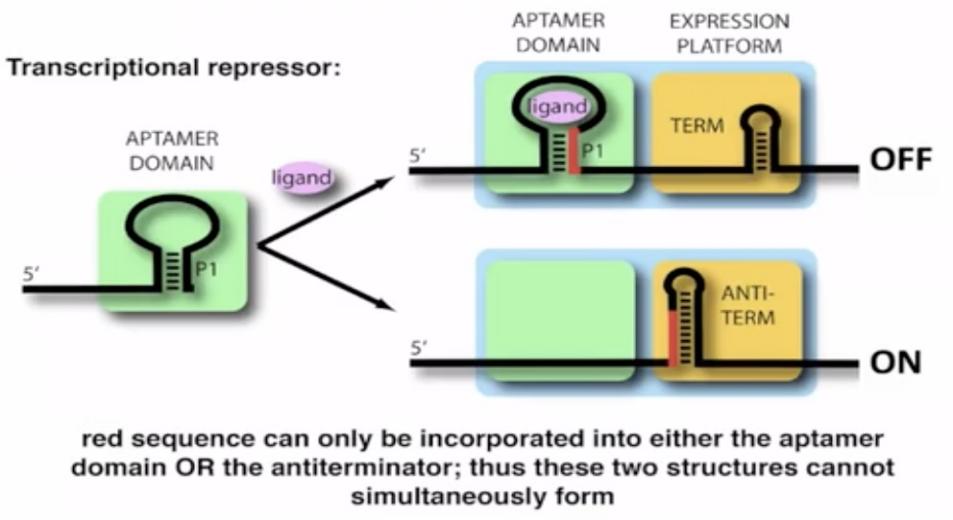


图2调节转录的riboswitch作用机制

研究发现，riboswitch广泛存在于细菌中，并且在过去的二十多年中已经发现了大于55种riboswitches，极大丰富了RNA调控基因表达领域的研究。科研人员发现在大多数细菌中往往存在着丰富的RNA调控介导机制，例如riboswitch，antisense RNA，RNA thermomter等方面。尤其是在厚壁菌门的细菌中，riboswitch的出现频率很高，说明在这一类细菌中riboswitch在调控基因表达方面具有重要作用。

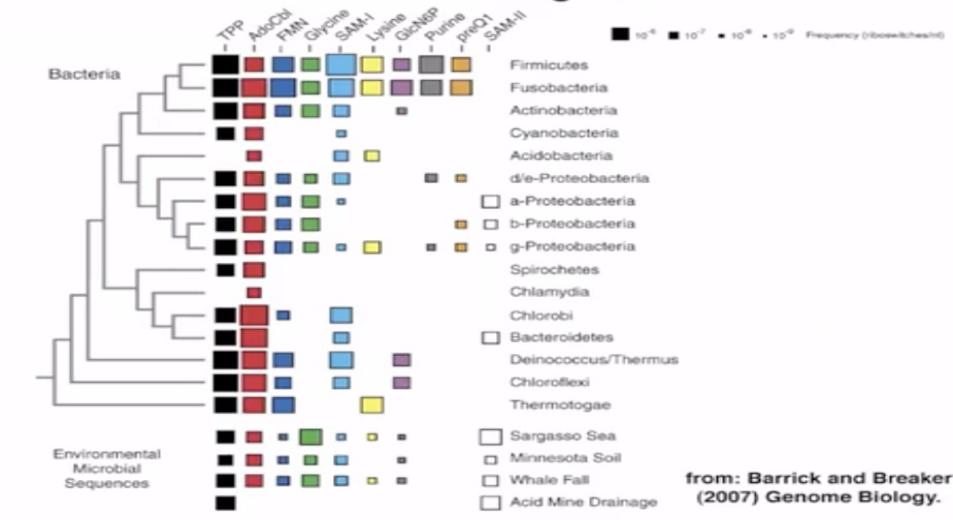


图3 riboswitch广泛分布于各类细菌中

到目前为止已发现接近60种不同类型的riboswitches，它们识别amino acids，sugars，nucleobases and nucleosides，ions等多种不同类型的配体分子从而发挥广泛的作用。尚未在哺乳动物中鉴定出天然的riboswitches，但已证实有几种合成的riboswitches也能够调节哺乳动物细胞中的基因表达，并且往往都是通过介导pre-mRNA的剪接过程，但提高其在体内的性能仍是一大挑战。在已发现的机制中，riboswitch可通过结合配体后将剪接体识别位点例如3′ SS整合入适体域，使其隔离从而导致剪接体无法识别位点，进而跳过含有过早出现终止密码子的外显子，从而让后续的翻译过程能顺利进行下去，基于这一机制，治疗型的核糖开关在今后很可能成为一大发展趋势。由腺相关病毒 (AAV) 开发的载体是在人类和动物模型中体内进行基因传递的强大工具，几种利用AAV作为载体递送的基因疗法目前已被批准用于临床。riboswitch由于其尺寸较小、非免疫原性、模块化结构以及无需蛋白质开关元件即可发挥作用的特点，是控制AAV递送的治疗性基因表达的有力方法。通过AAV将riboswitch与目的基因运送入目标区域，病人服入相关配体后，riboswitch进行调控目的基因的表达，从而达到治疗目的。

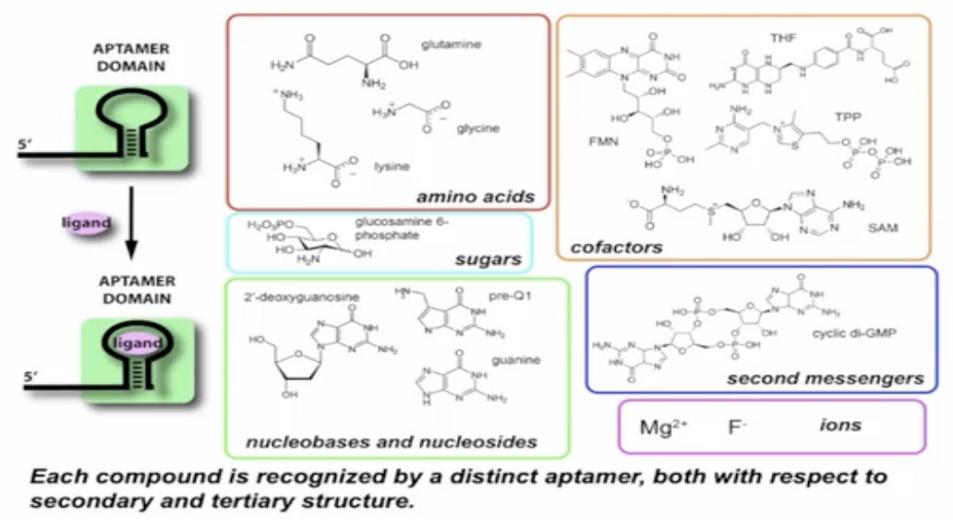


图4 riboswitch所识别的多种配体类型

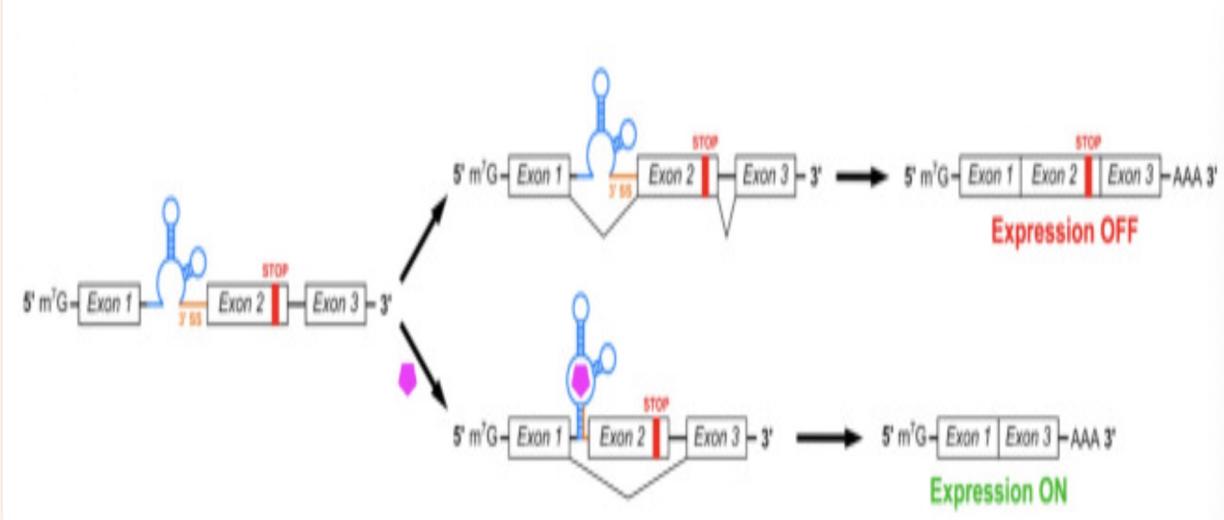


图5 riboswitch调控mRNA剪接的机制

Robert T. Batey教授教授指出，科研人员在过往的研究中大多专注于对riboswitch的适体域进行单独研究，不仅是因为其序列相对较短，并且在于配体形成复合物后构象一般都较为稳定，在后续的结晶及测量亲和力等方面都较为方便，故而对于全长的riboswitch研究较少。Robert T. Batey教授以lysine riboswitch为例向我们详细介绍了其适体域是如何结合lysine的。由图5不难发现，该适体域有典型的多螺旋结构，在解析出三级结构后研究者还发现结构中存在三个同轴堆叠现象，分别是P2与P2a。P2b与P3以及P4与P5螺旋，而且还存在环环相互作用。此外，lysine紧紧包裹在由五个stem连接处组成的结合口袋中。

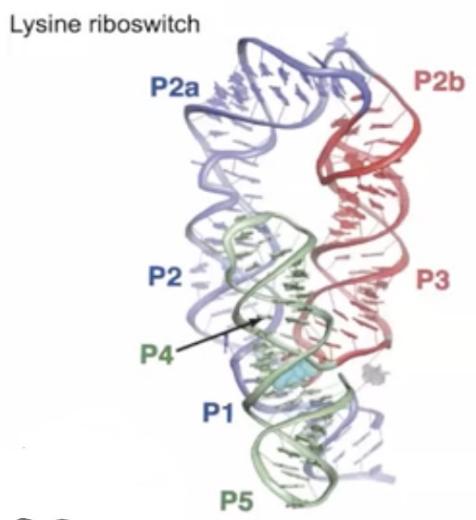
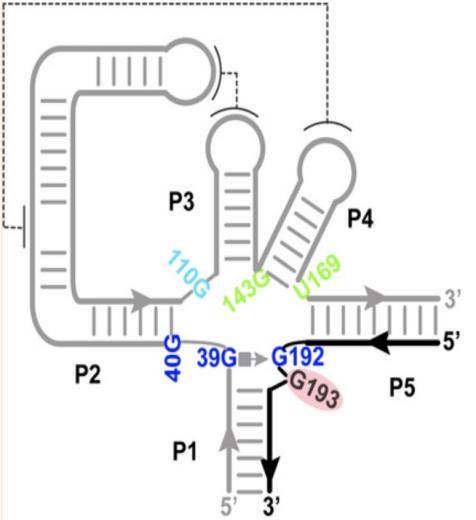


图6 lysine riboswitch适体域的二级结构与三级结构

再进一步对结合位点分析后发现，lysine riboswitch与lysine transport protein尽管在组成成分、空间结构、理化特性等方面都存在显著差异，但对lysine的识别结合位点却表现出高度的相似性。

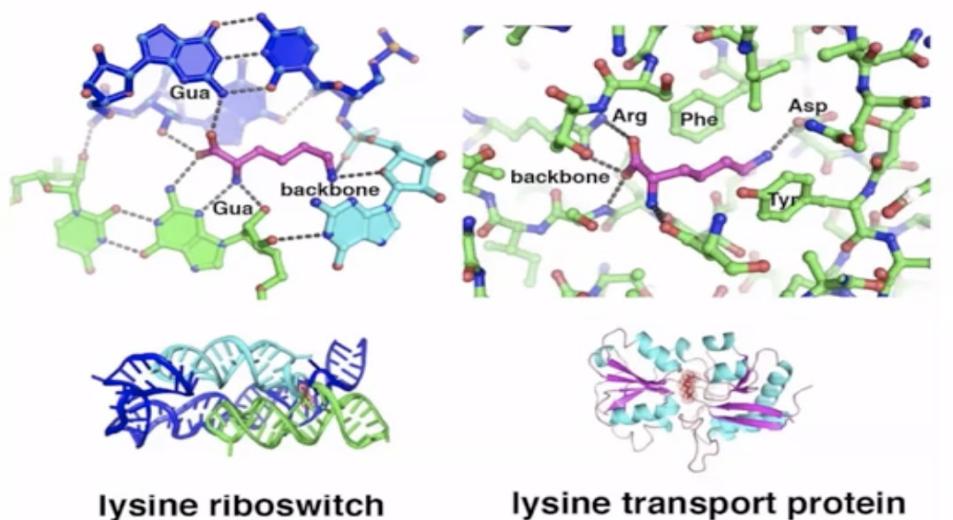


图7 lysine riboswitch与lysine transport protein与lysine的结合位点

Robert T. Batey教授教授强调目前对于适体域的研究已经较为深入，我们也应该更多关注当配体结合到适体域后，表达平台是如何影响这种变化以及后续是通过怎样的机制调节下游基因表达的。他形象地将适体域比做input domain，而表达平台则是output domain。配体起到效应分子的作用，通过其与适体域的结合，进而使表达平台开始变化，从而发挥作用。作为output domain的表达平台往往包含终止子/抗终止子序列或核糖体结合位点序列信息，起到调控转录或翻译的作用。目前认为适体域结合配体后，表达平台通过三种机制进行响应，分别是direct occlusion，interdomain docking以及strand invasion。

首先是direct occlusion，在这种机制中，Robert T. Batey教授以SAM-II riboswitch为例，适体域和表达平台往往形成一个整体不会分开成单独的两部分，当配体结合后，复合物形成的结构会对核糖体结合位点序列产生遮挡作用，使得核糖体无法与之结合，导致翻译的阻滞。

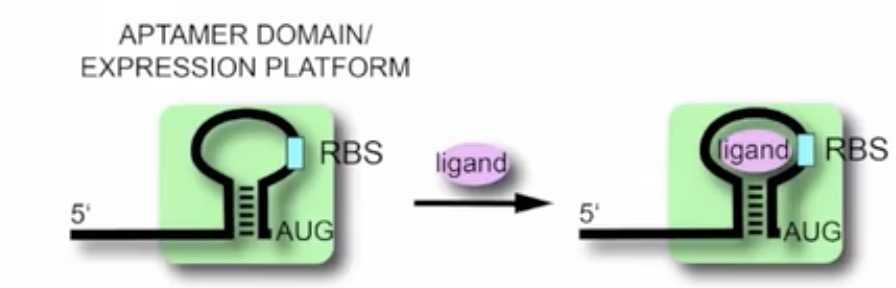


图8 direct occlusion机制的原理图



图9 SAM-II riboswitch的二级结构与三级结构

2023年1月清华大学方显杨团队发表的论文报道了THF-II riboswitch的晶体结构，研究发现当缺乏Mg2+时，THF-II riboswitch的结合口袋未折叠或部分折叠，无法与配体结合。此时P1的核糖体结合位点序列易与核糖体结合。在存在Mg2+时，结合口袋构象发生改变。P1双螺旋稳定性增强导致核糖体的结合能力减弱。Mg2+和配体同时存在条件下，riboswitch进一步发生共轴堆积变化，最终形成棒状结构。核糖体无法结合，下游翻译被抑制。有趣的是，在本次研究中发现当结合配体后，P1右臂的核糖体结合位点序列在原有三个经典配对的基础上还与左臂的序列形成了三个非经典的碱基配对，这些较强的相互作用不仅稳定了P1的螺旋结构，还使得核糖体无法结合位点进行翻译。此外，RNase H切割实验表明配体结合后减少了核糖体结合位点序列的暴露，可及性降低，说明配体的结合也导致了遮挡作用，正是第一种机制的典型体现。

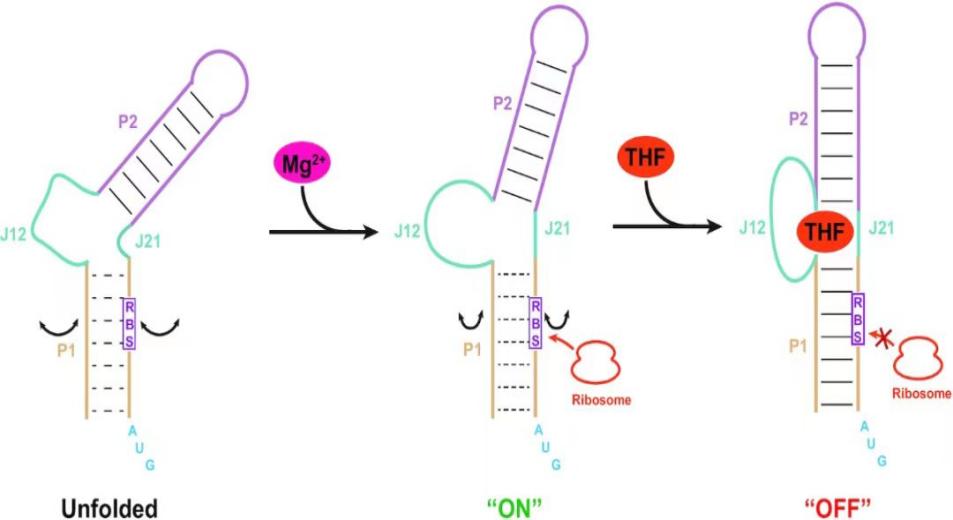


图10 THF-II riboswitch调节翻译的机制

其次是interdomain docking，在这种机制中，适体域和表达平台间是分隔开的。当配体结合到适体域后，表达平台与适体域可以发生相互作用，也会起到遮挡核糖体结合位点序列的作用，从而影响翻译过程。Robert T. Batey教授以cobalamin riboswitch为例向我们介绍了具体的机制，当cobalamin结合到适体域后，复合物形成的结构允许表达平台“停靠”并发生相互作用，含有核糖体结合位点序列的L14与适体域的L5间会发生环环相互作用形成碱基配对。同时研究人员还发现cobalamin也会提供一部分表面以稳定这种相互作用，在上述的机理下，核糖体结合位点序列不仅与适体域形成较强的相互作用且也因为空间结构的遮挡使得序列暴露程度降低，核糖体无法结合，导致翻译过程受阻。

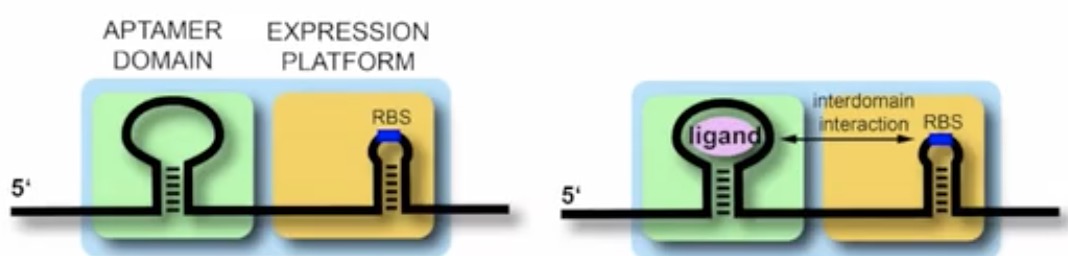


图11 interdomain docking机制的原理图

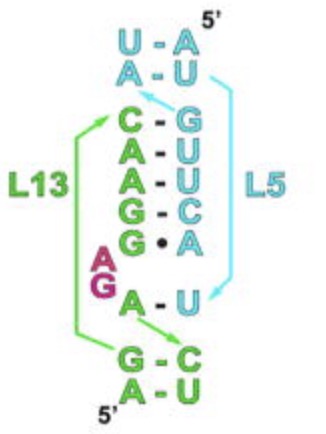
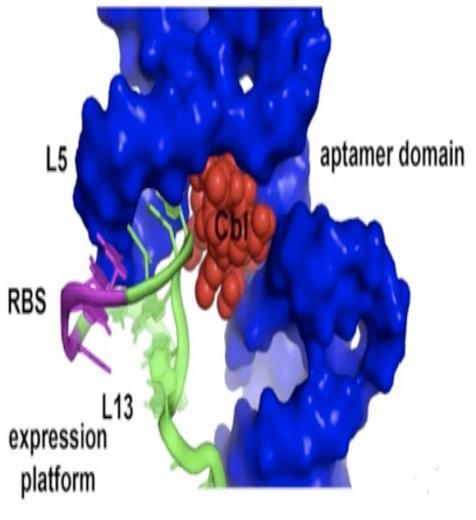


图12 cobalamin riboswitch中表达平台和适体域间的相互作用

第三种机制是strand invasion，也是最普遍存在的一种机制，大约有90%以上的riboswitch都采用这种方式调节基因表达。在这种机制中，配体结合到适体域时，switching sequence（红色序列）整合在抗终止子茎环中，此时转录可以正常进行，而当配体结合到适体域后，表达平台中的switching sequence会“入侵”到适体域中形成配对，从而稳定P1茎环结构，与此同时表达平台的构象也发生变化，形成终止子茎环结构，导致下游转录过程终止。



图13 strand invasion机制的原理图

Robert T. Batey教授强调由于转录的先后性，所以适体域比起表达平台较先形成，可以优先感知细胞内环境与代谢物浓度。若配体结合到适体域上，则switching sequence（红色序列）则会整合到适体域中，表达平台折叠成终止子构象，RNA聚合酶脱离，无法再进行转录。若不结合配体，表达平台则将switching sequence则整合到抗终止子中，下游的转录过程可以继续进行。

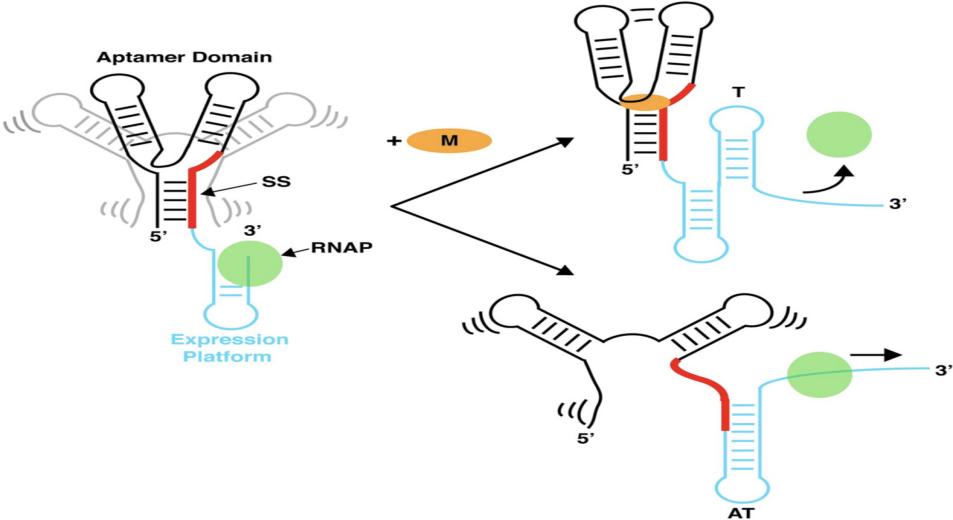


图14 riboswitch调节转录的机制示意图

随后Robert T. Batey教授简要介绍了strand invasion的原理，该过程在体内外都非常迅速高效并且不需要需要能量，在一个duplex motif中，外来的single stranded nucleic acid会识别toehold初步进行一个定位，然后通过annealing过程朝着与互补序列相反的方向进行“入侵”延伸，直到完全取代duplex motif中原有的一条链。当在riboswitch中时，需要对上述的机制做一些小修改。在配体结合后，复合物形成的三级结构会产生空间阻碍作用，阻止进一步的“入侵”作用。

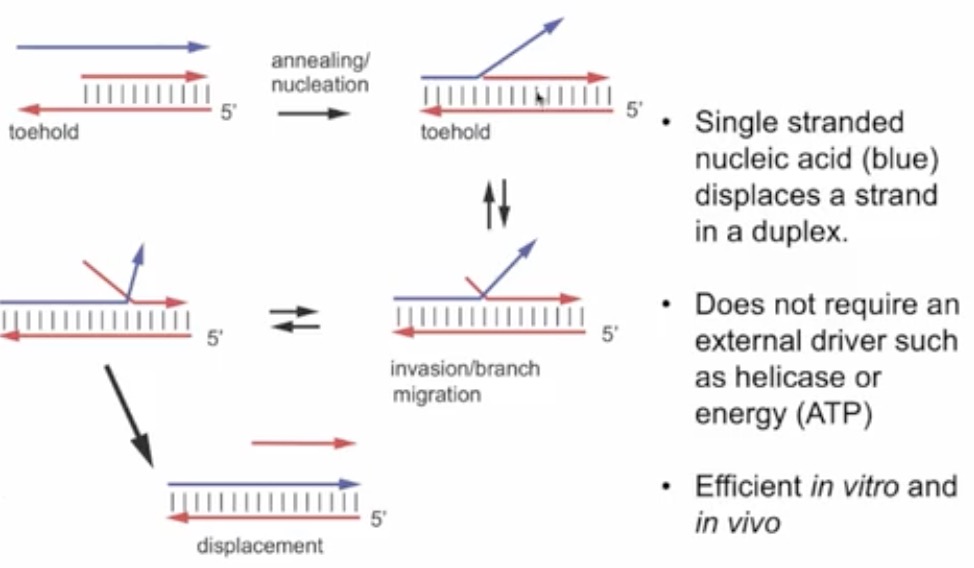


图15 strand invasion原理示意图

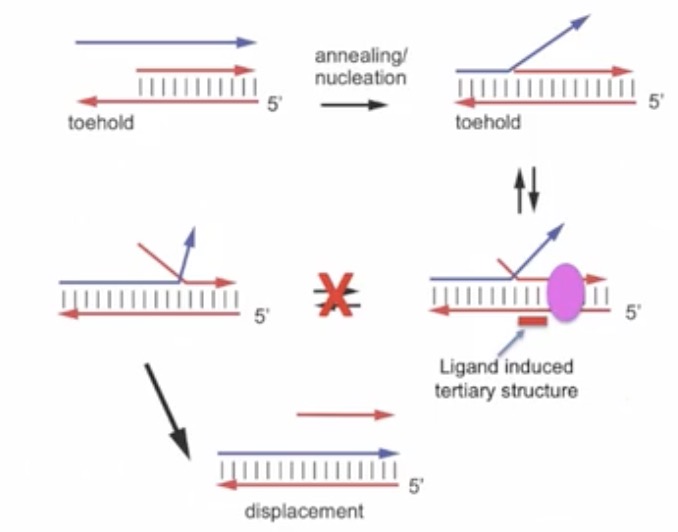


图16 riboswitch中strand invasion的不同之处

最后，Robert T. Batey教授对本次讲座的内容进行了简短的总结。他认为截至目前大多数研究还是将重心发在解析适体域与配体结合后的复合物结构，但要具体了解riboswitch的作用机制，研究人员也需要多关注配体结合适体域状态下发挥作用的表达平台是如何变化以调控介导下游基因的表达。

提问环节苗智超老师首先提出了一个非常有意思的问题，他好奇如何通过riboswitch来促进转录或翻译而非对二者产生抑制作用。对于该问题，Robert T. Batey教授介绍了曾经在cobalamin riboswitch中所做过的一个尝试性实验，他将一个linker放置在适体域和表达平台间，并不断加入核苷酸延长其长度，使适体域不能与表达平台发生相互作用，增加核糖体结合位点序列的暴露程度，使得适体域不能再阻止核糖体结合到对应的位点上。通过实验发现这种方法确实出现了提高翻译表达的结果，但将长度延长到30个核苷酸之后，由于距离太长，riboswitch已无法正常发挥作用了。苗老师总结了Robert的观点，即通过一些方法增加核糖体结合到对应位点序列的机会，从而促进翻译，他也提出如果能通过优化从而招募一些促进翻译的因子也许也会是一个解决方法。

不久前Daniel B. Haack等人通过scaffold与cryo-EM技术对TPP riboswitch结合配体前后的结构进行了解析，因为TPP riboswitch较小，所以本次研究中使用O.i. group IIC intron作为scaffold与TPP riboswitch连接起来，scaffold的高分辨率特点可以帮助我们更好地定位观察目标riboswitch的结构。

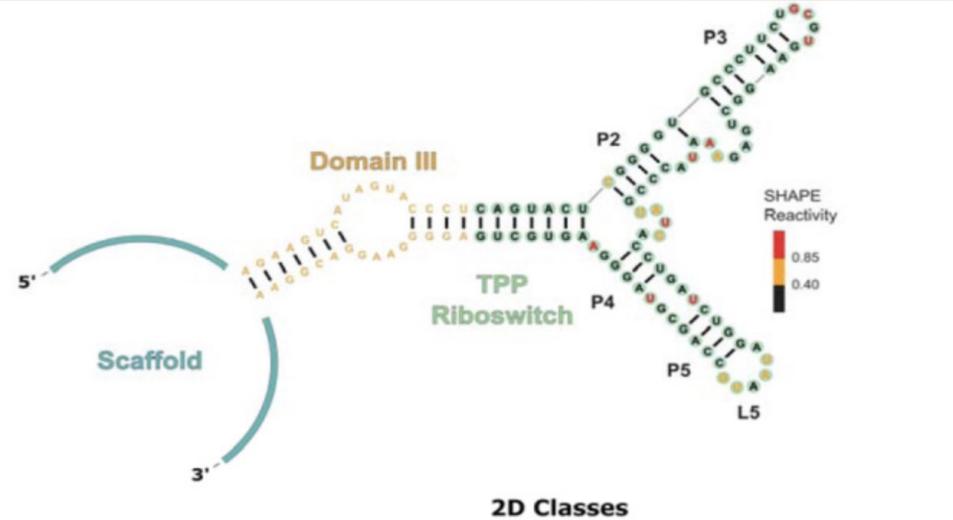


图17 O.i. group IIC intron scoffold与TPP riboswitch连接形成的复合物结构

黄林老师在最后提出我们目前是否有必要探究riboswitch在无配体结合状态下（apo state）的结构，对于这个问题Robert T. Batey教授认为无配体结合状态下的结构可能是非常复杂多变的，他并不认为目前的方法可以让我们能完整地探究到RNA折叠的多样性，但在未来如果研究人员可以改善决定分子动力学模拟精度的中的force field，也许我们能够更好地在原子层面研究生物大分子结构与动态变化。此外，诸如Alphafold这一类软件的高速发展对于未来预测RNA结构也是可以产生极大帮助的。