

Additional File 6: Consensus diagrams of all RNA motifs supplementary for: Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea and their metagenomes

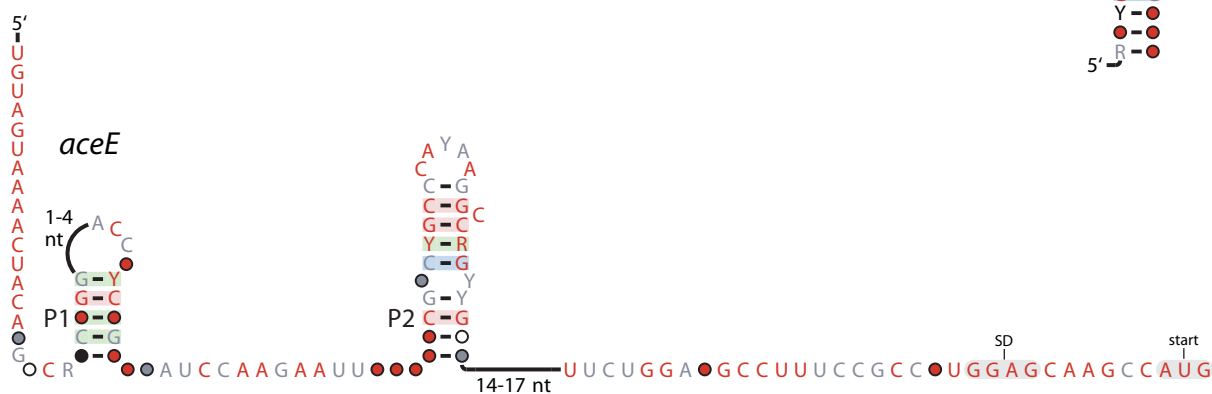
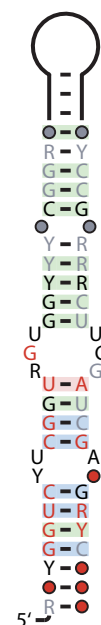
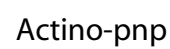
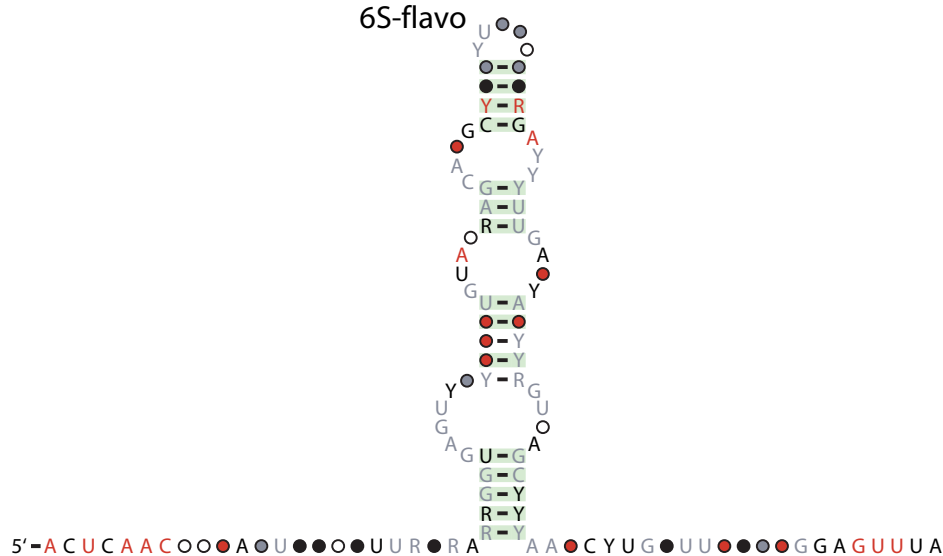
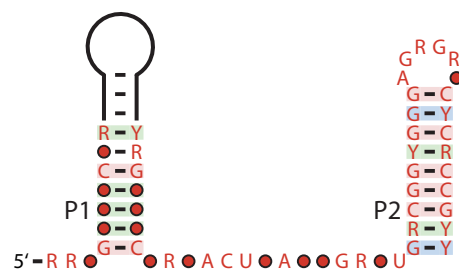
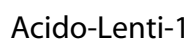
Zasha Weinberg, Joy X. Wang, Jarrod Bogue, Jingying Yang,
Keith Corbino, Ryan H. Moy, Ronald R. Breaker

February 27, 2010

Contents

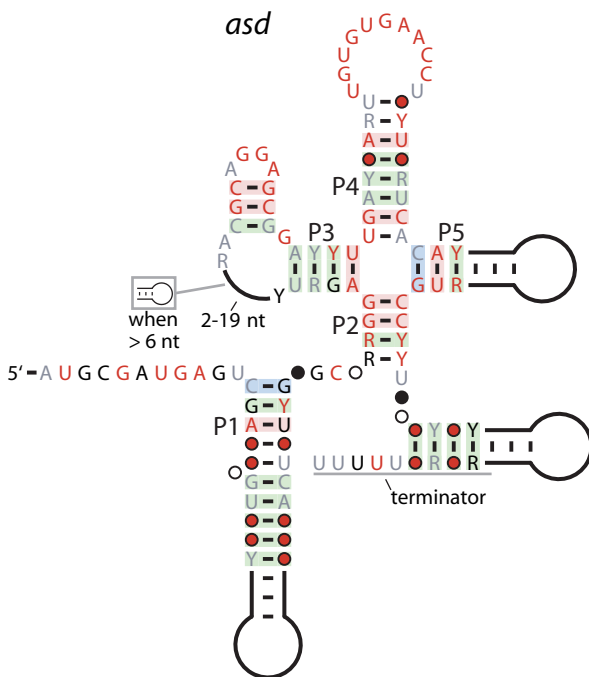
6S-Flavo , <i>aceE</i> , Acido-1 , Acido-Lenti-1 , Actino-pnp	3
AdoCbl-variant , <i>asd</i> , <i>atoC</i> , Bacillaceae-1 , <i>Bacillus</i> -plasmid	4
Bacteroidales-1 , <i>Bacteroides-1</i> , <i>Bacteroides-2</i> , c4 antisense RNA , c4 antisense RNA target a1b1	5
Chlorobi-1 , Chlorobi-RRM , Chloroflexi-1 , Clostridiales-1 , <i>Collinsella-1</i>	6
<i>crcB</i> , Cyano-1 , Cyano-2 , <i>Desulfotalea-1</i> , Downstream-peptide	7
Dictyoglomi-1 , <i>epsC</i> , <i>fixA</i> , Flavo-1	8
<i>flpD</i> , <i>flg</i> - Rhizobiales , <i>gabT</i> , Gamma-cis-1 , <i>glnA</i> , GUCCY-hairpin	9
Gut-1 , <i>gyrA</i> , <i>hopC</i> , <i>icd</i> , JUMPstart , L17 downstream element	10
<i>lactis</i> -plasmid, Lacto-int , Lacto-rpoB , Lacto-usp , leu/phe leader , Lnt	11
<i>manA</i> , <i>Methylobacterium-1</i> , <i>metK</i> - Rhodobacter , Moco-II , <i>msiK</i>	12
Ocean-V , Ocean-VI , <i>pan</i> , Pedo-repair	13
<i>pfl</i> , <i>psaA</i> , <i>pheA</i> , PhotoRC-I , PhotoRC-II	14
<i>Polynucleobacter-1</i> , <i>potC</i> , Pseudomon-1	15
<i>psbNH</i> , Pseudomon-2 , Pseudomon-groES , <i>Pseudomon-Rho</i> , <i>Pyrobac-1</i> , <i>Pyrobac-HINT</i>	16
<i>radC</i> , Rhizobiales-1 , Rhizobiales-2 , <i>Rhodopirellula-1</i> , <i>rmf</i>	17
<i>rne-II</i> , SAM-Chlorobi , SAM-I-nil	18
SAM/SAH , <i>sanguinis</i> -hairpin, <i>sbcD</i> , ScRE , Soil-1 , Solibacter-1	19
STAXI , <i>sucA-II</i> , <i>sucC</i>	20

Termite-<i>flg</i>, Termite-<i>leu</i>, <i>traJ</i>-II, TwoAYGGAY	21
<i>wcaG</i>, Whalefall-1, <i>yjdF</i>, <i>ykkC</i>-III	22

6S-Flavo, *aceE*, Acido-1, Acido-Lenti-1, Actino-pnp

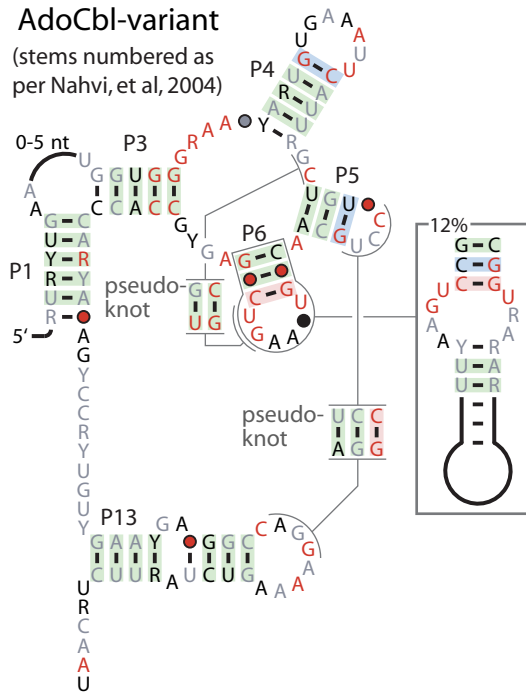
A portion of this figure was adapted from the supplementary data of a previous publication [21].

AdoCbl-variant, *asd*, *atoC*, Bacillaceae-1, *Bacillus*-plasmid

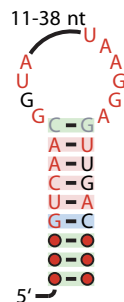


AdoCbl-variant

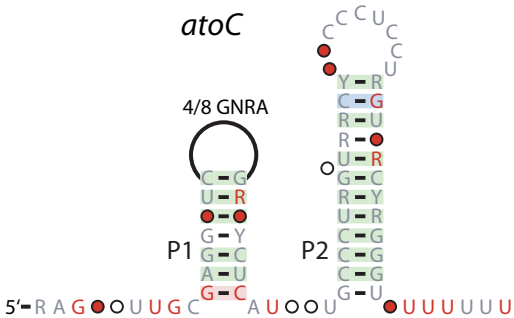
(stems numbered as per Nahvi, et al, 2004)



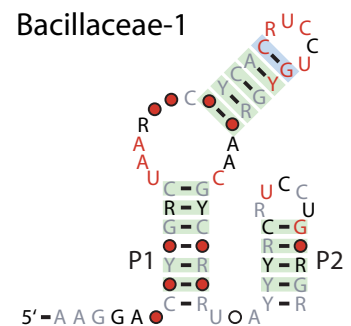
Bacillus-plasmid



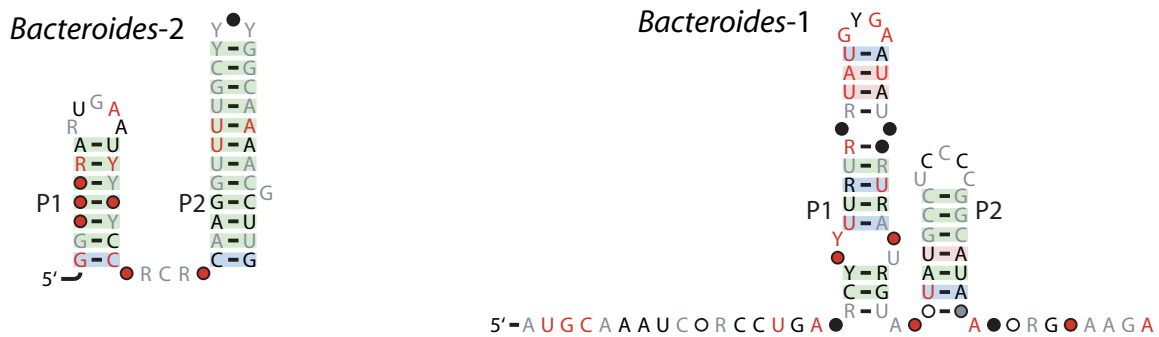
atoC



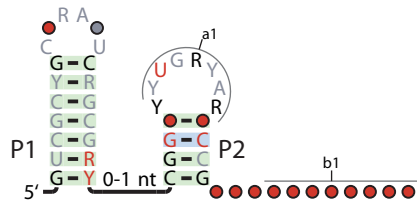
Bacillaceae-1



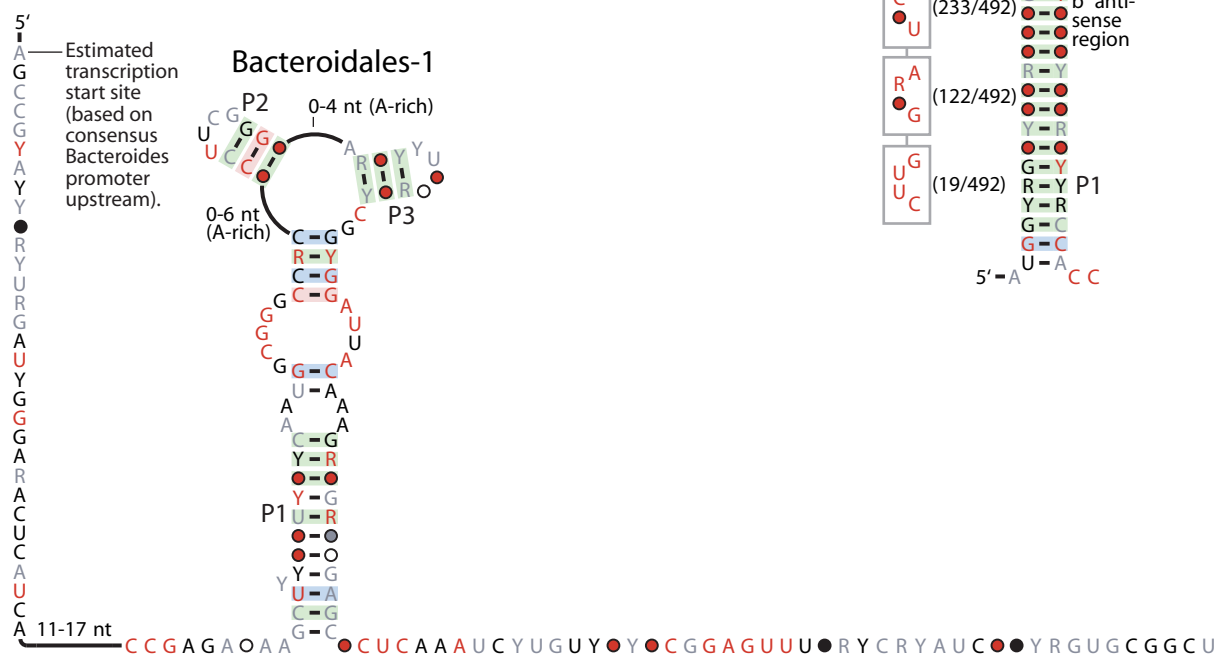
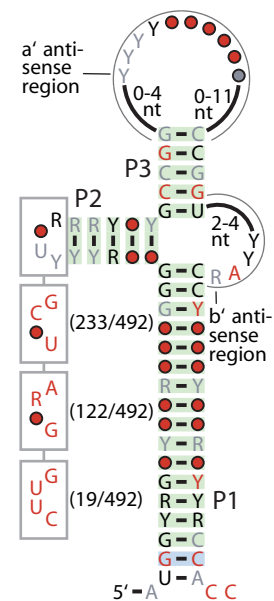
Bacteroidales-1, *Bacteroides*-1, *Bacteroides*-2, c4 antisense RNA, c4 antisense RNA target a1b1



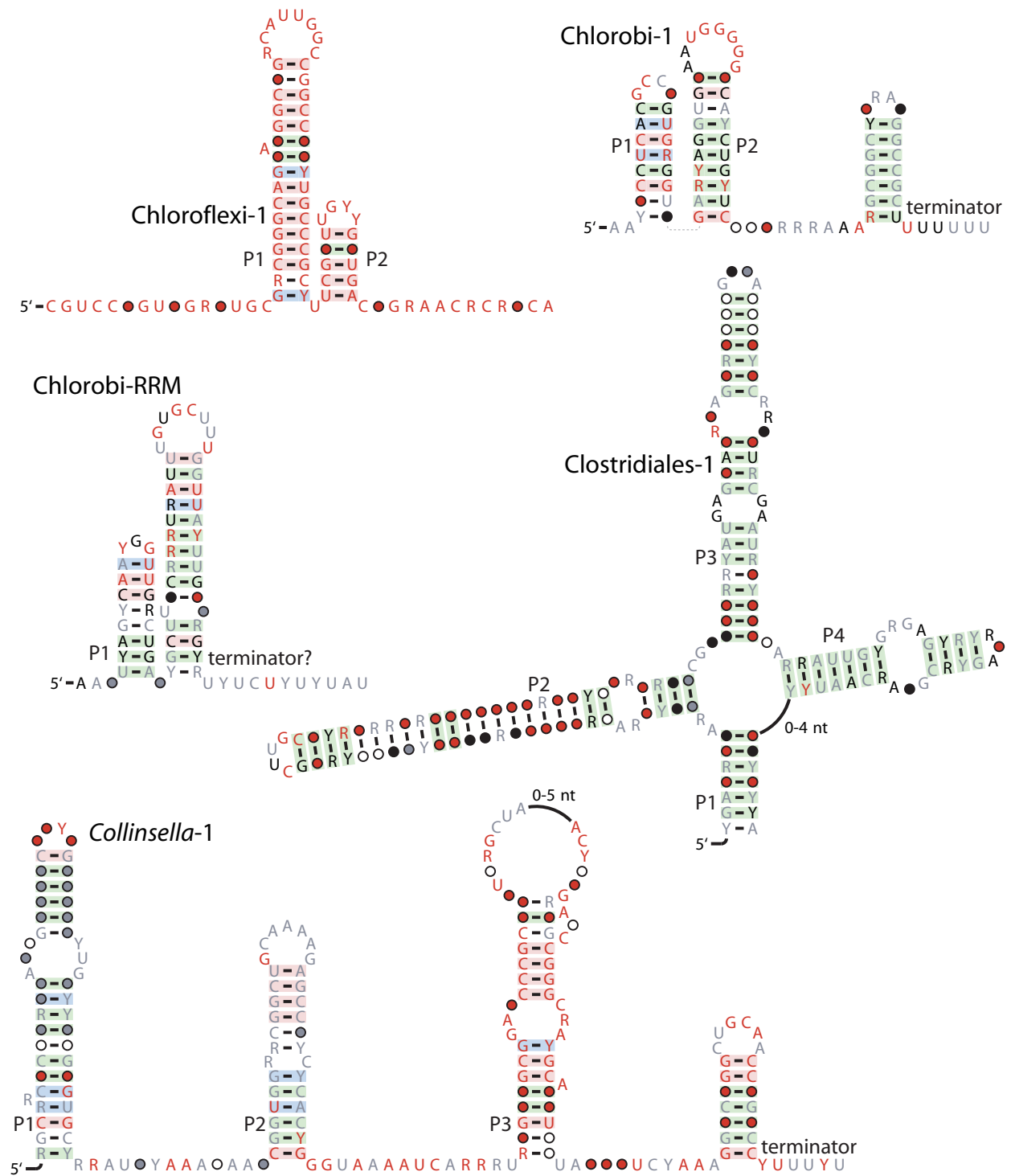
c4 antisense RNA a1b1



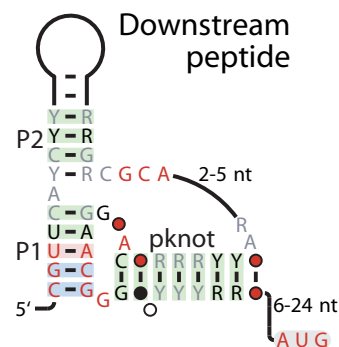
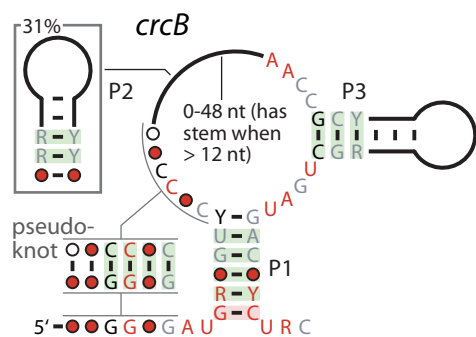
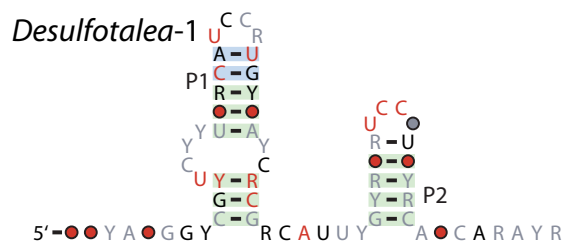
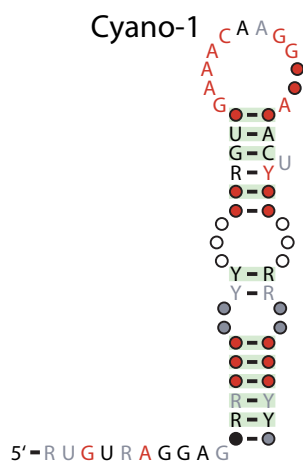
c4 antisense RNA



Chlorobi-1, Chlorobi-RRM, Chloroflexi-1, Clostridiales-1, *Collinsella*-1



crcB, Cyano-1, Cyano-2, *Desulfotalea*-1, Downstream-peptide

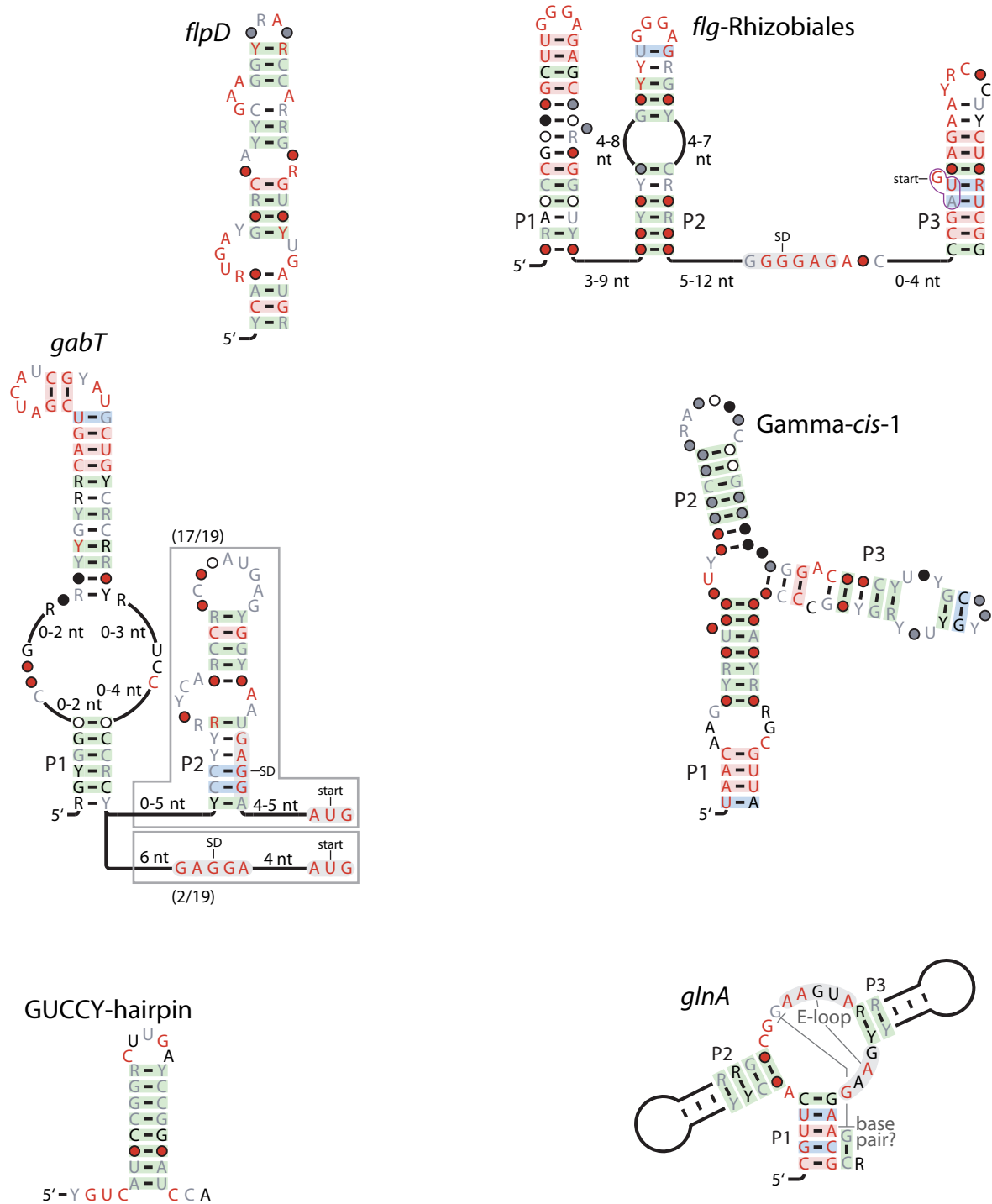


The figure displays four RNA secondary structure models:

- Flavo-1**: A long, linear RNA molecule with multiple stem-loops and internal loops. The sequence starts with 5'-Y Y and ends with G A. Key regions include P1, E-loop, and several other stems.
- Dictyoglomi-1**: A complex RNA molecule with multiple stem-loops and internal loops. The sequence starts with R A A and ends with C G. Key regions include P2, E-loop, and several other stems.
- epsC**: A small RNA molecule with a single stem-loop structure. The sequence starts with 5'-G G U R A and ends with G R. Key regions include P1, P2, P3, and P4.
- fixA**: A small RNA molecule with a single stem-loop structure. The sequence starts with 5'-CGGUCAAGAUACGRGCGGUU and ends with R R R R G. Key regions include P1 and E-loop.

8

flpD, *flg*-Rhizobiales, *gabT*, Gamma-*cis*-1, *glnA*, GUCCY-hairpin

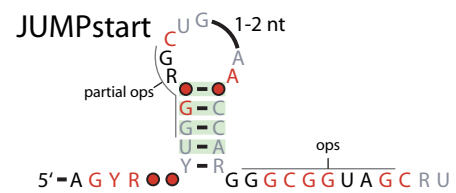
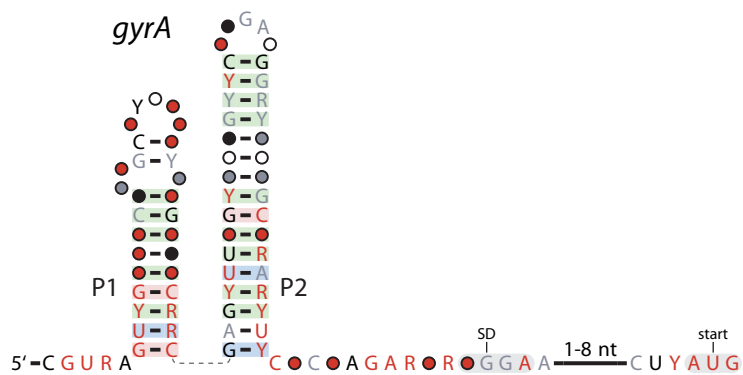
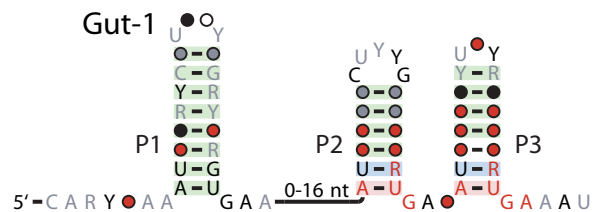
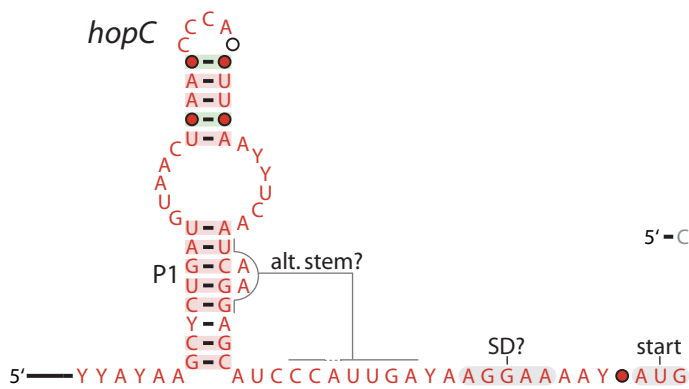
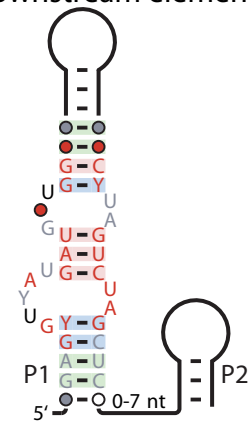


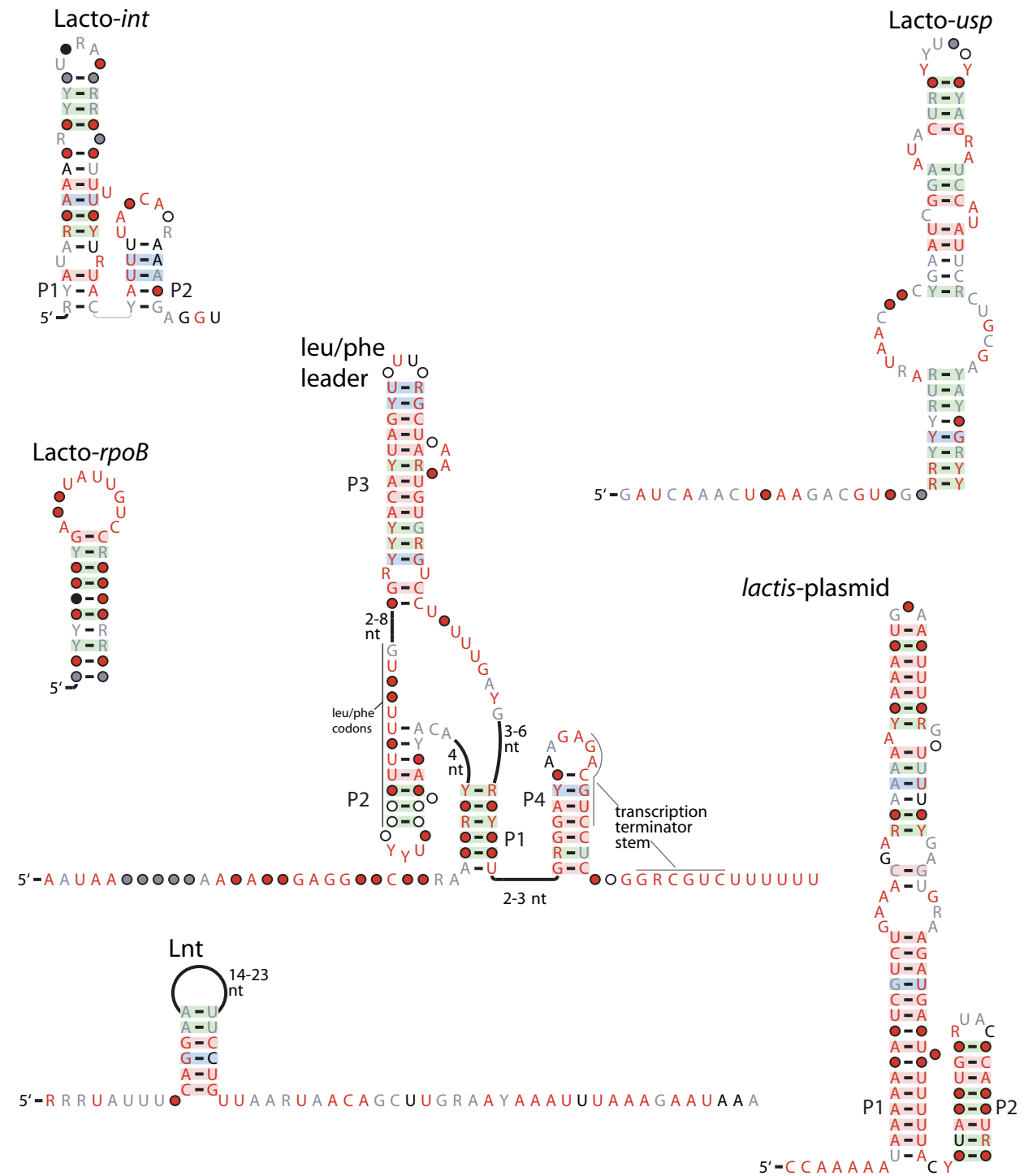
A portion of this figure was adapted from the supplementary data of a previous publication [21].

Gut-1, *gyrA*, *hopC*, *icd*, JUMPstart, L17 downstream element

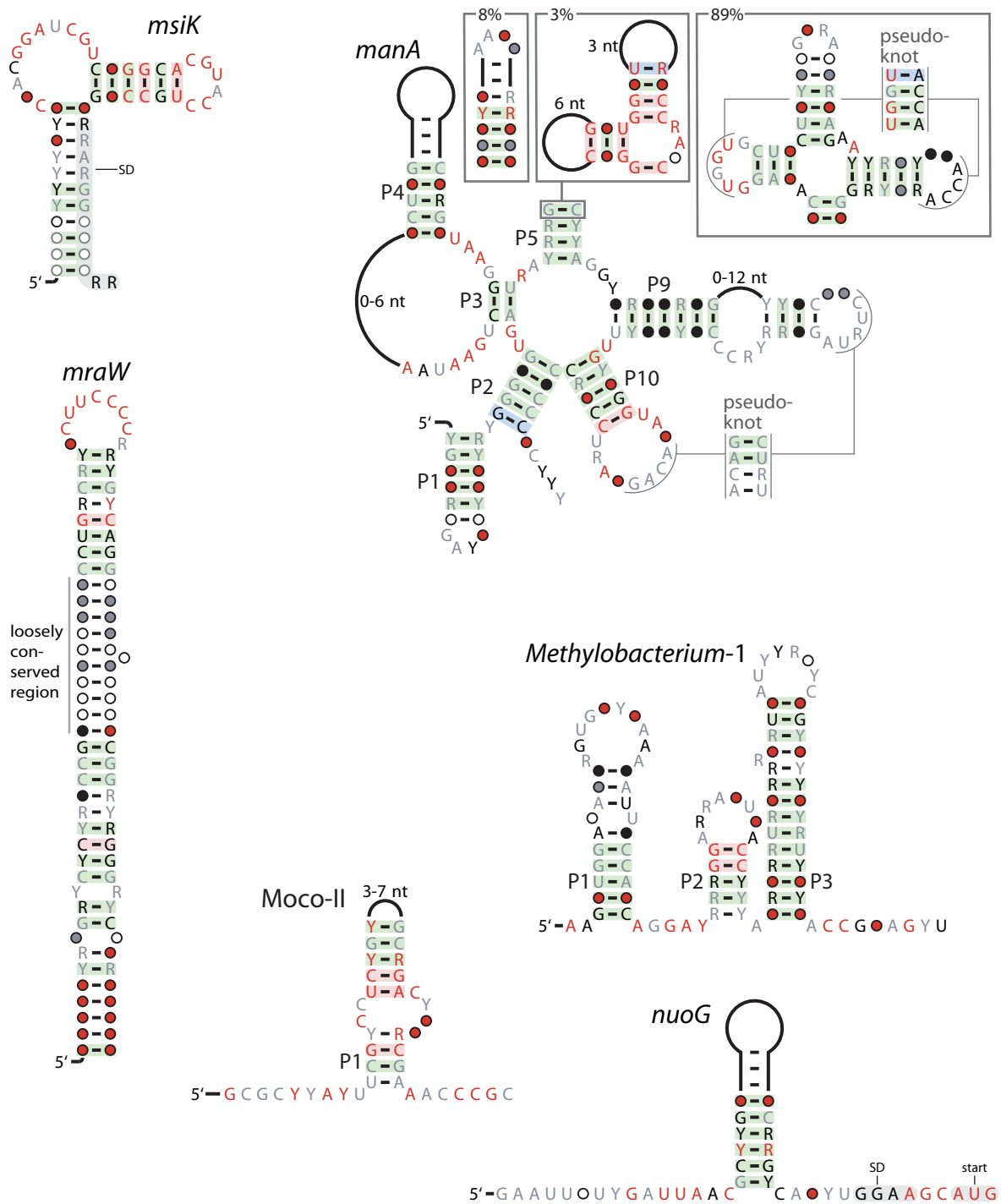


L17 downstream element



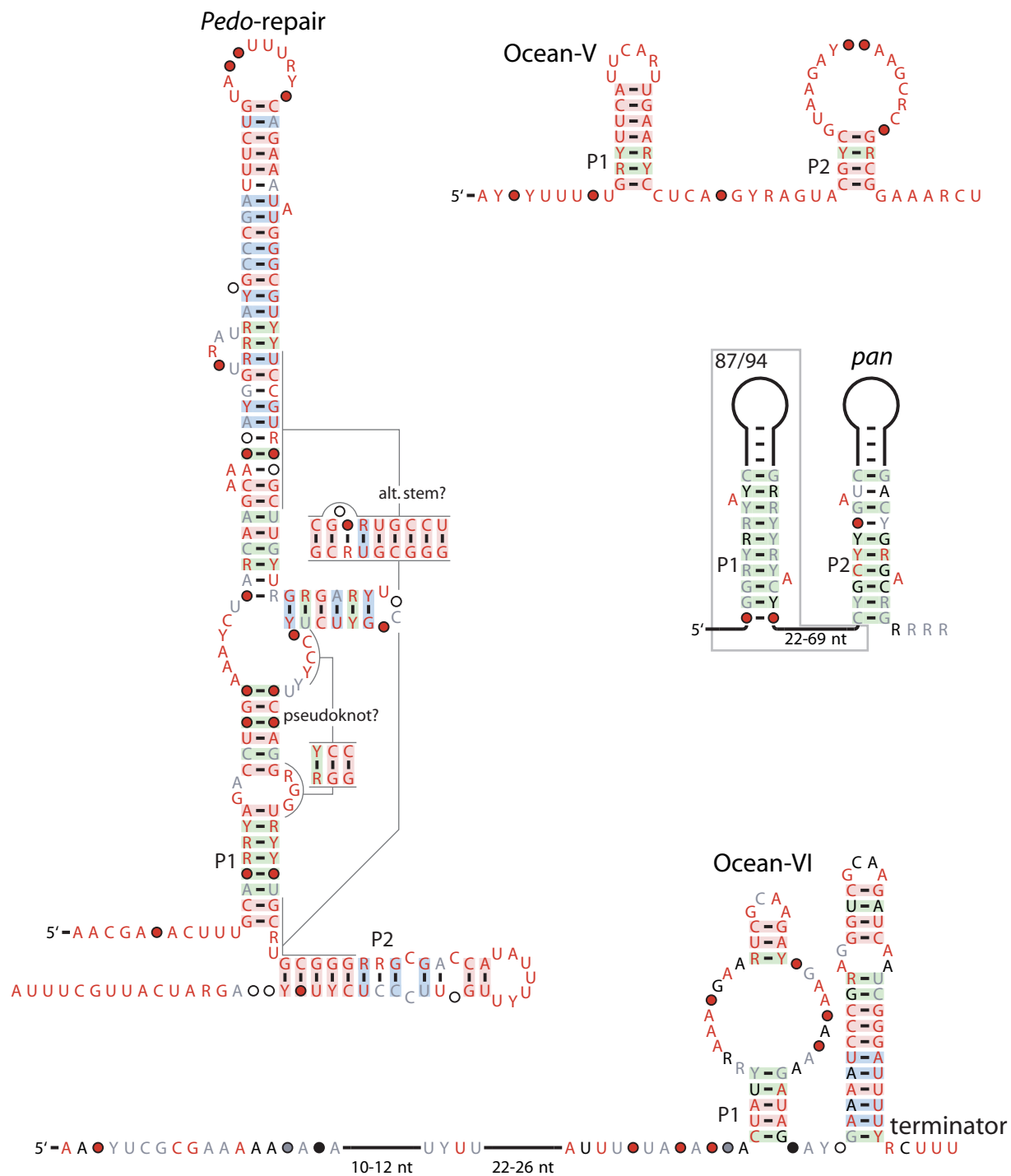
lactis-plasmid, Lacto-*int*, Lacto-*rpoB*, Lacto-*usp*, leu/phe leader, Lnt

manA, *Methylobacterium-1*, *metK*-Rhodobacter, *Moco-II*, *msiK*

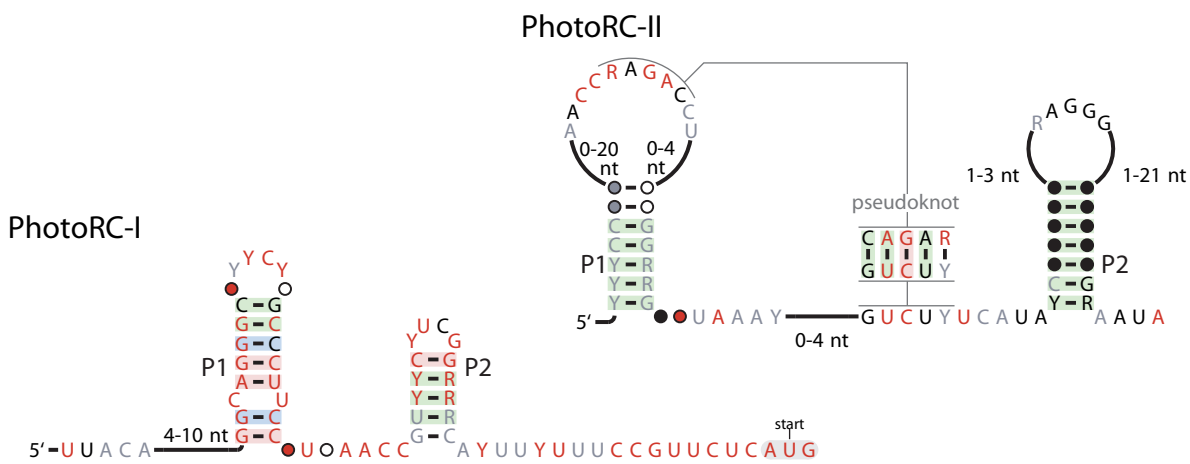
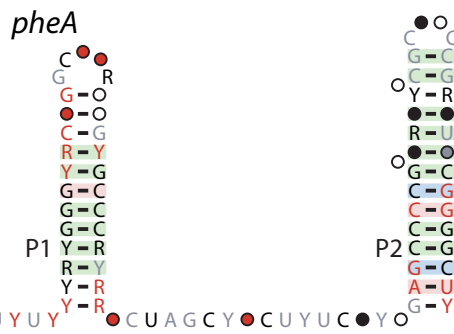
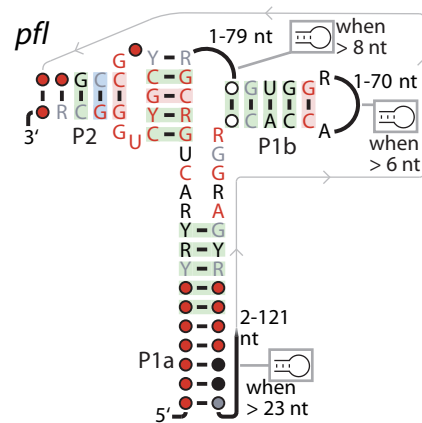
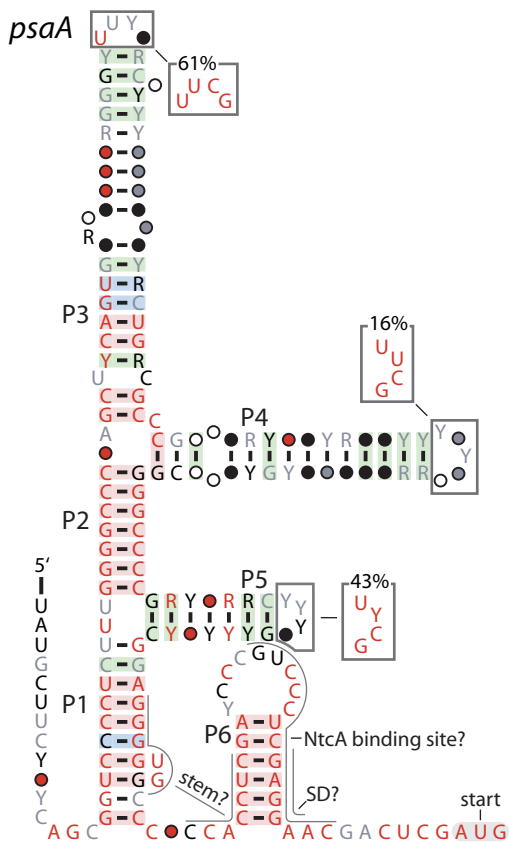


A portion of this figure was adapted from the supplementary data of a previous publication [21].

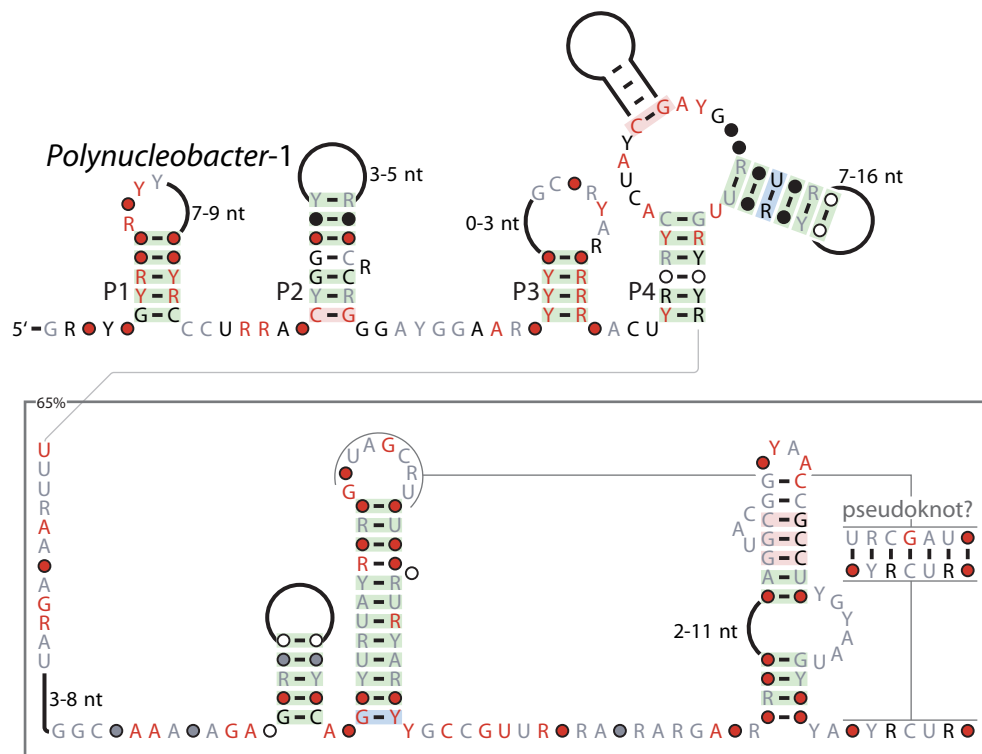
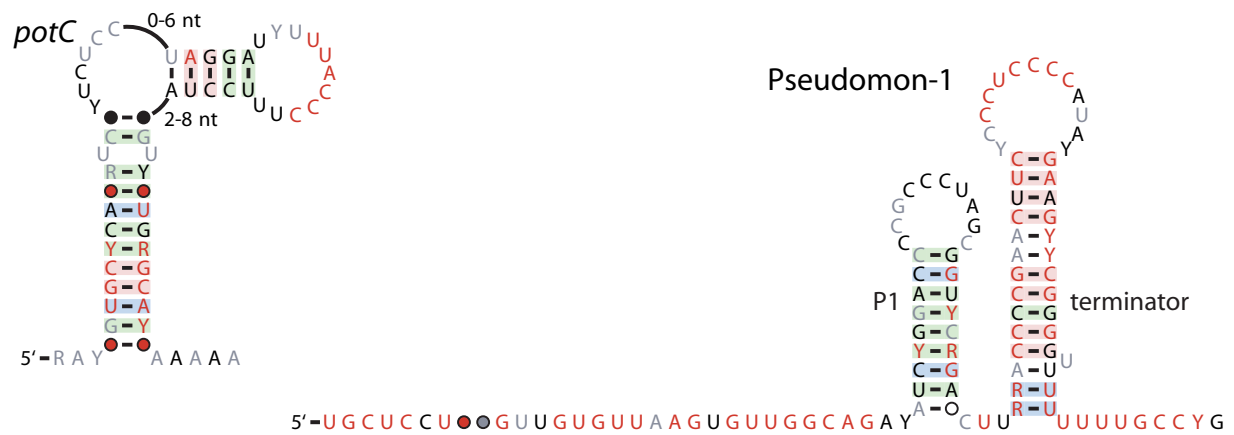
Ocean-V, Ocean-VI, *pan*, *Pedo*-repair



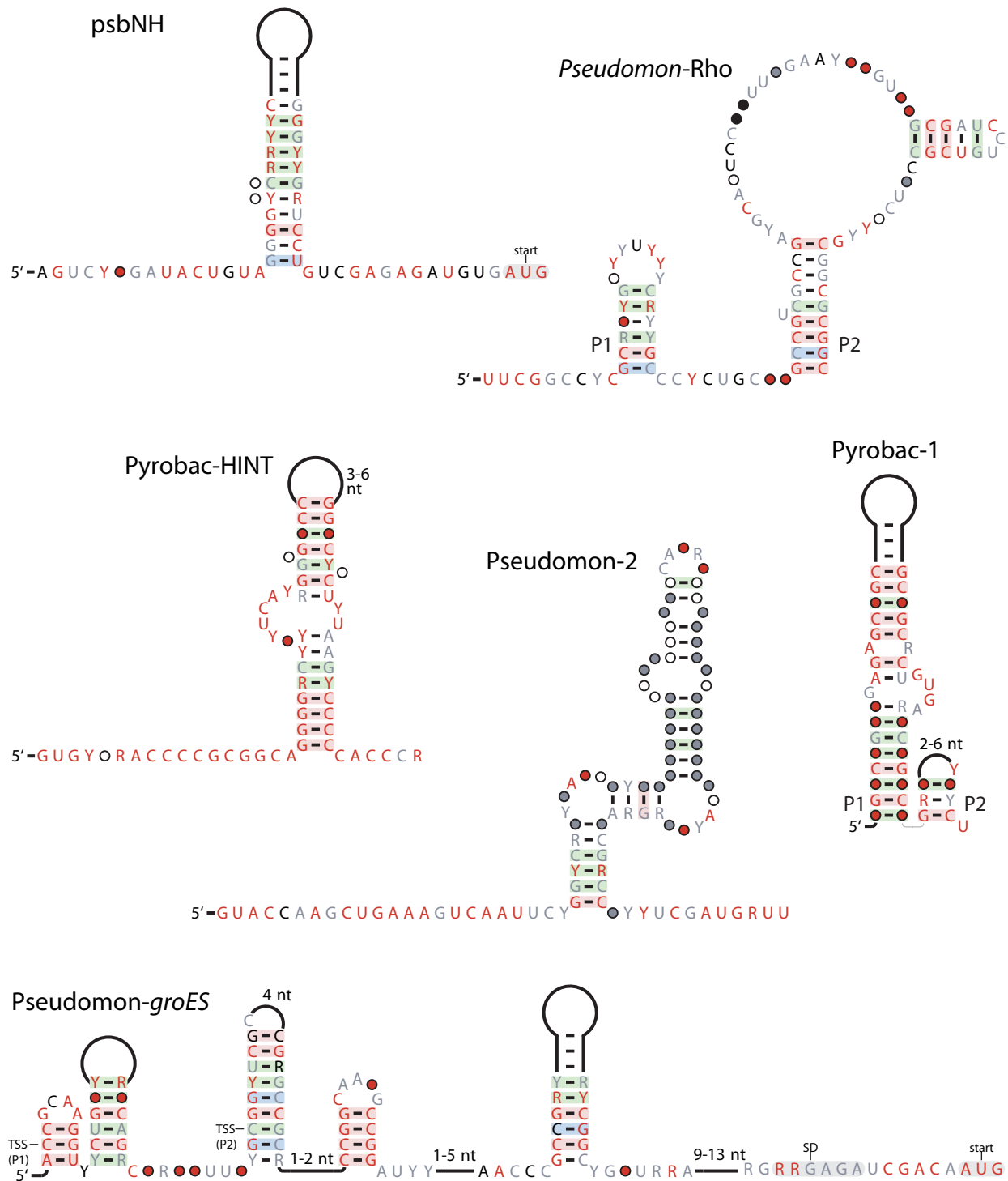
pfl, *psaA*, *pheA*, **PhotoRC-I**, **PhotoRC-II**



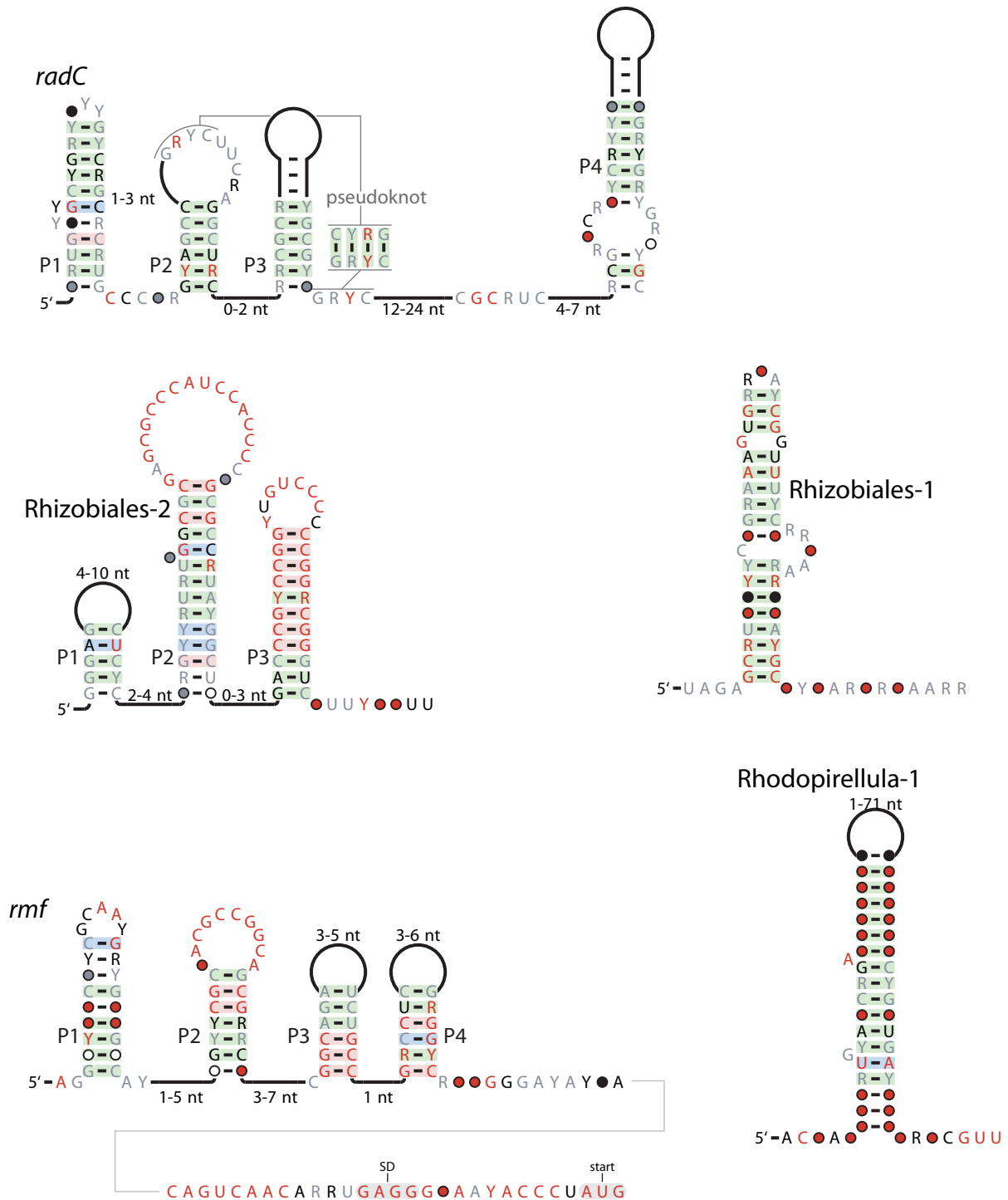
Polynucleobacter-1, *potC*, *Pseudomon-1*



psbNH, **Pseudomon-2**, *Pseudomon-groES*, *Pseudomon-Rho*, *Pyrobac-1*,
Pyrobac-HINT



radC, Rhizobiales-1, Rhizobiales-2, Rhodopirellula-1, *rmf*

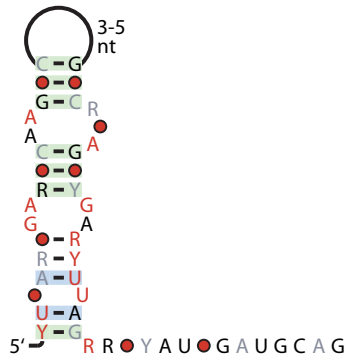


[illegible]

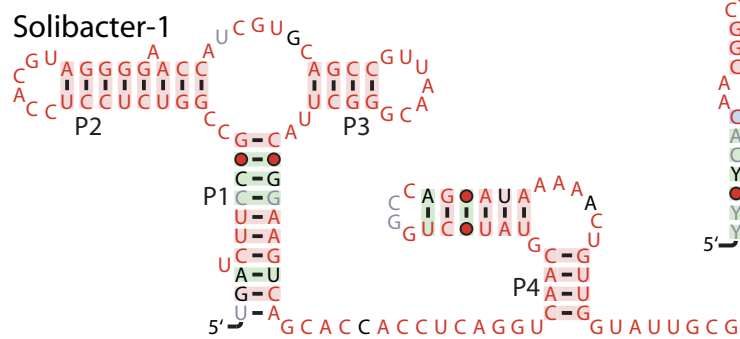
SAM/SAH, *sanguinis*-hairpin, *sbcD*, ScRE, Soil-1, Solibacter-1



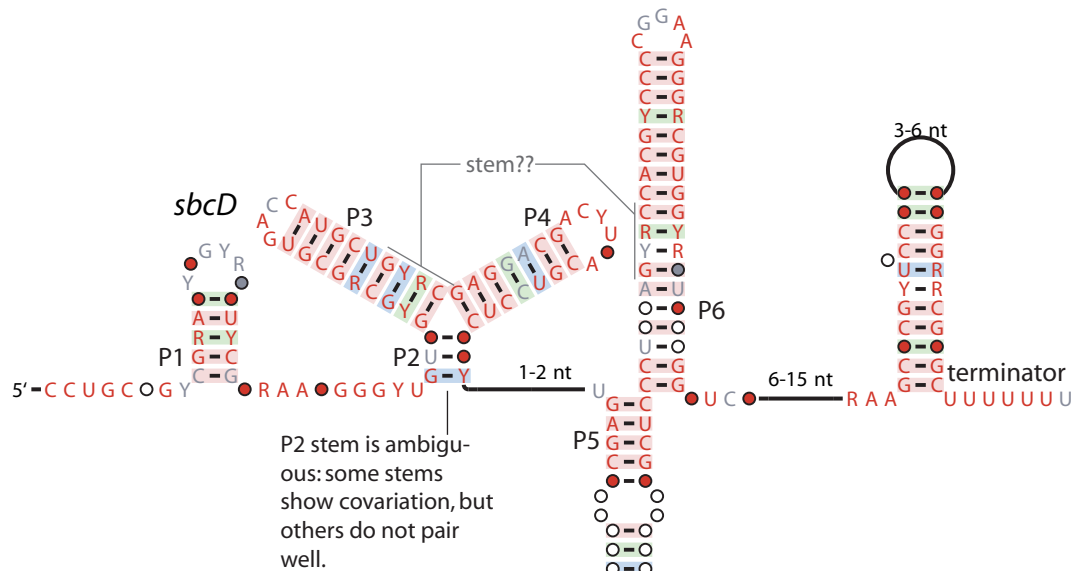
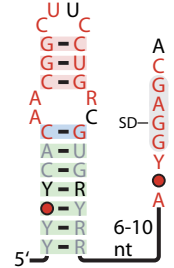
sanguinis-hairpin



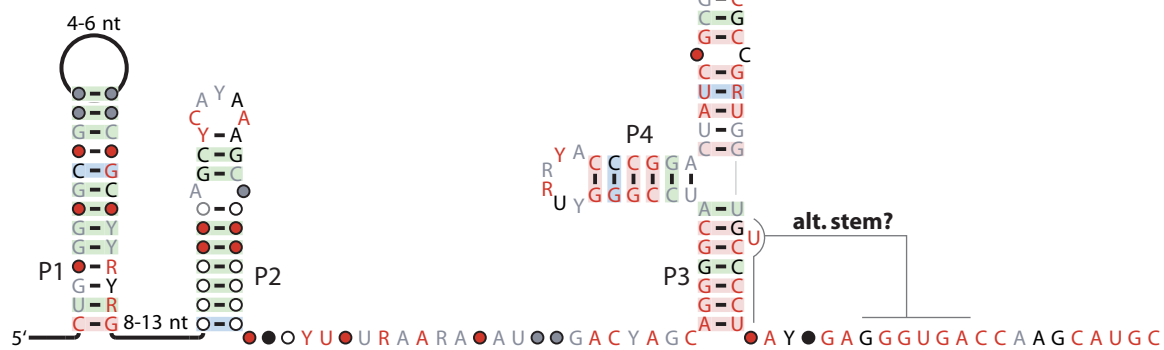
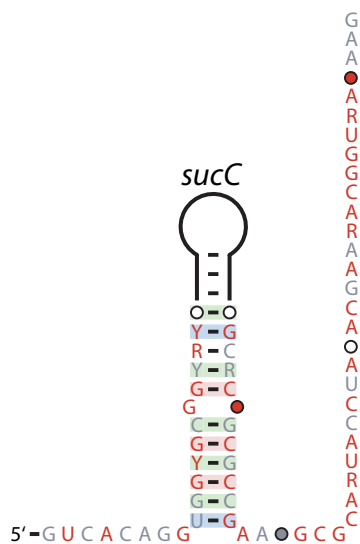
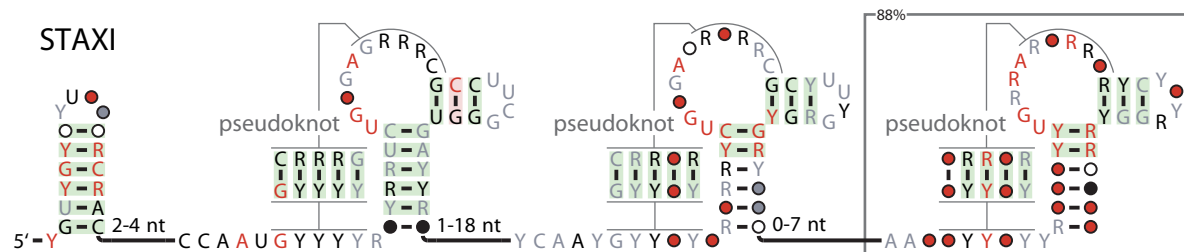
Solibacter-1



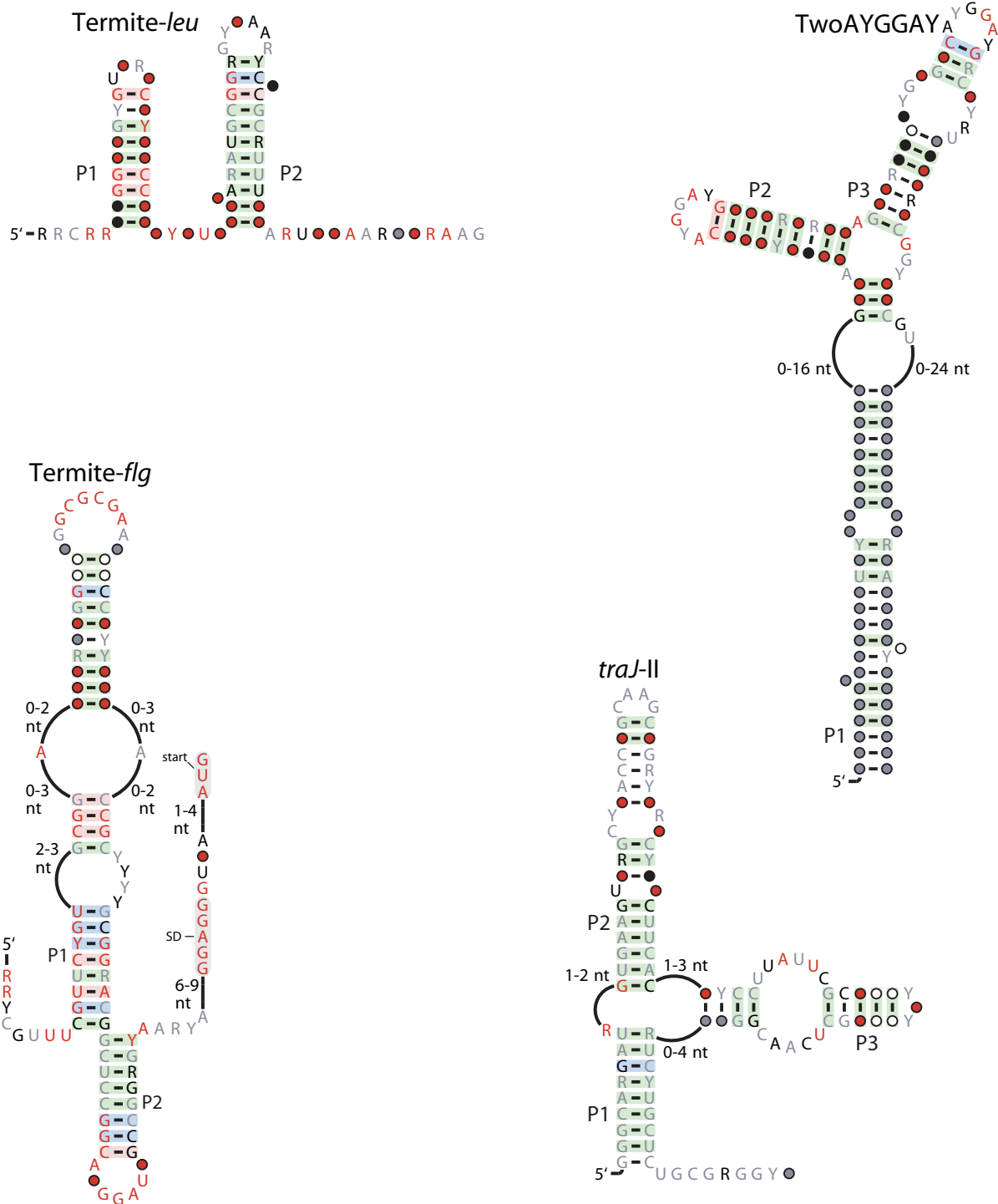
SAM/SAH



STAXI, *sucA*-II, *sucC*

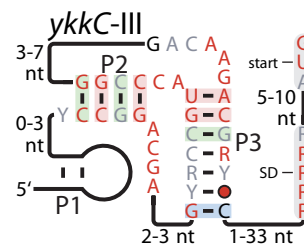
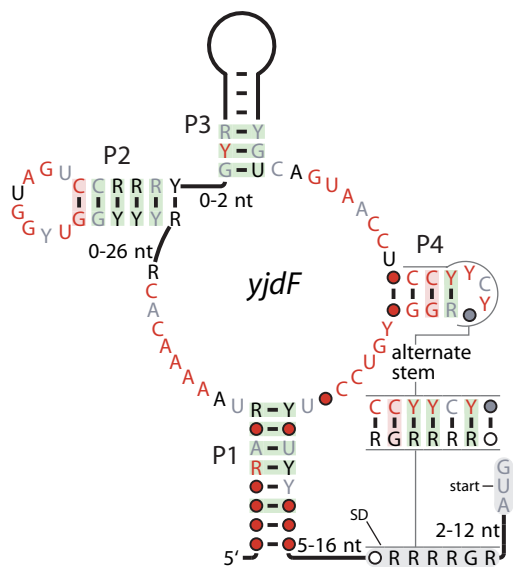
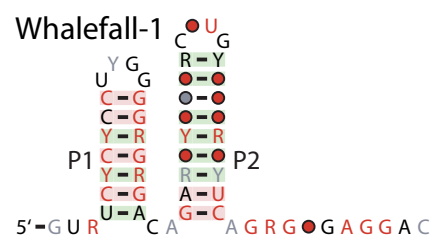
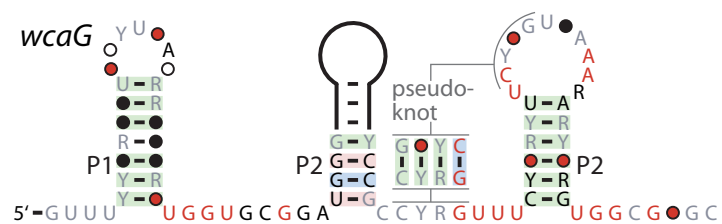


Termite-*flg*, Termite-*leu*, *traJ*-II, TwoAYGGAY



A portion of this figure was adapted from the supplementary data of a previous publication [21].

wcaG, Whalefall-1, *yjdF*, *ykkC-III*



A portion of this figure was adapted from the supplementary data of a previous publication [21].

References

References are listed in the manuscript.