Palindromic Sequences

Restriction enzymes cut double-stranded DNA* at specific locations based the pattern of bases found at those locations. These enzymes predictably cut both strands because the sequences they recognize are palindromic. The recognition sequences are short strings of identical bases on both DNA strands.

Palindromic sequences are similar to language palindromes but follow distinct rules. By following these rules, any string of bases can be made into a palindromic sequence

Repeat
New Sequence

```
AAATTT
TTTAAA
```

Palindromic sequences are a short run of bases (typically 4 to 8 bases in length), followed by their complementary bases in reverse order. For example, the recognition sequence for BamHI is GGATCC.

Note the first three bases GGA are followed by the complement of those three bases in reverse order: TCC. The complement to the whole six base strand is CCTAGG, read backwards (as it would be when reading from 5' to 3' on the complementary strand) is GGATCC, an exact match for the original strand.

This pattern makes it possible to reconstruct a palindromic sequence from one-half of one strand. For example, the six-base recognition sequence (e.g. TAGCTA) can be reconstructed from just knowing the first three bases on one strand:

- Starting with the original sequence - TAG
- Calculate the reverse complement of the sequence - ATC
- Reverse the order of the reverse complement (CTA) and add it to the end of the forward strand - TAGCTA
- Calculate the reverse complement of the whole forward strand to finish the reverse strand: ATCGAT

Final result:
TAGCTA
ATCGAT
Having short stretches of DNA that read the same on both strands of double-stranded DNA allow restriction enzymes to cut both strands in the same place.

But, the sequence of one strand defines the other, so what ,does non palindromic mean when we speak of DNA?

Take any random stretch of DNA such as AGTCCGATCCGT
find its reverse complement (ie, write the bases that pair in each position): TCAGGCTAGGCA
Because the strands are antiparallel, this sequence is 3' to 5' if read left to right. Therefore, to write the sequence of the complementary strand "as usual" ie 5' to 3', we
flip it to the proper 5'-3' orientation for the complementary strand: ACGGATCGCCACT
The sequence you see is not the same, so it is not a palindrome. ACGGATCGCCACT ≠ AGTCCGATCCGT

The next page is an extract from a restriction enzyme catalogue, giving a partial list of enzymes whose recognition sequence is NOT palindromic, according to the definition above. In general, these enzymes do not cut within the recognition sequence.

tags:

Enzymes with Nonpalindromic Sequences

All recognition sequences are written 5' to 3' using the single letter code nomenclature. Numbers in parentheses indicate the point of cleavage.

Example, GGTCTC(1/5) indicates cleavage at:
5'...GGTCTCN/...3'
3'...CCAGAGNNNN/...5'

Enzyme	Recognition Sequence
AciI	CCGC(-3/-1)
AcuI	CTGAAG(16/14)
AlwI	GGATC(4/5)
BaeI	(10/15)ACNNNNGTAYC(12/7)
BbsI §	GAAGAC(2/6)
BbsI-HF®	GAAGAC(2/6)
BbvCI	CCTCAGC(-5/-2)
BbvI	GCAGC(8/12)
BccI	CCATC(4/5)
BceAI	ACGGC(12/14)
BcgI	(10/12)CGANNNNNNTGC(12/10)
BciVI	GTATCC(6/5)
BcoDI	GTCTC(1/5)
BfuAI	ACCTGC(4/8)
BmgBI	CACGTC(-3/-3)
BmrI	ACTGGG(5/4)
BpmI	CTGGAG(16/14)
BpuEI	CTTGAG(16/14)
Bpu10I	CCTNAGC(-5/-2)
BsaI-HF®v2	GGTCTC(1/5)
BsaXI	(9/12)ACNNNNNCTCC(10/7)
BseRI	GAGGAG(10/8)
BseYI	CCCAGC(-5/-1)