



---

# Overhang set report

---



## Example

There are 5 overhangs in this set. The restriction enzyme used for this set is **Esp3I**.

Overhangs in this set: GAAA, CACC, GACT, ATGG, CCAG.

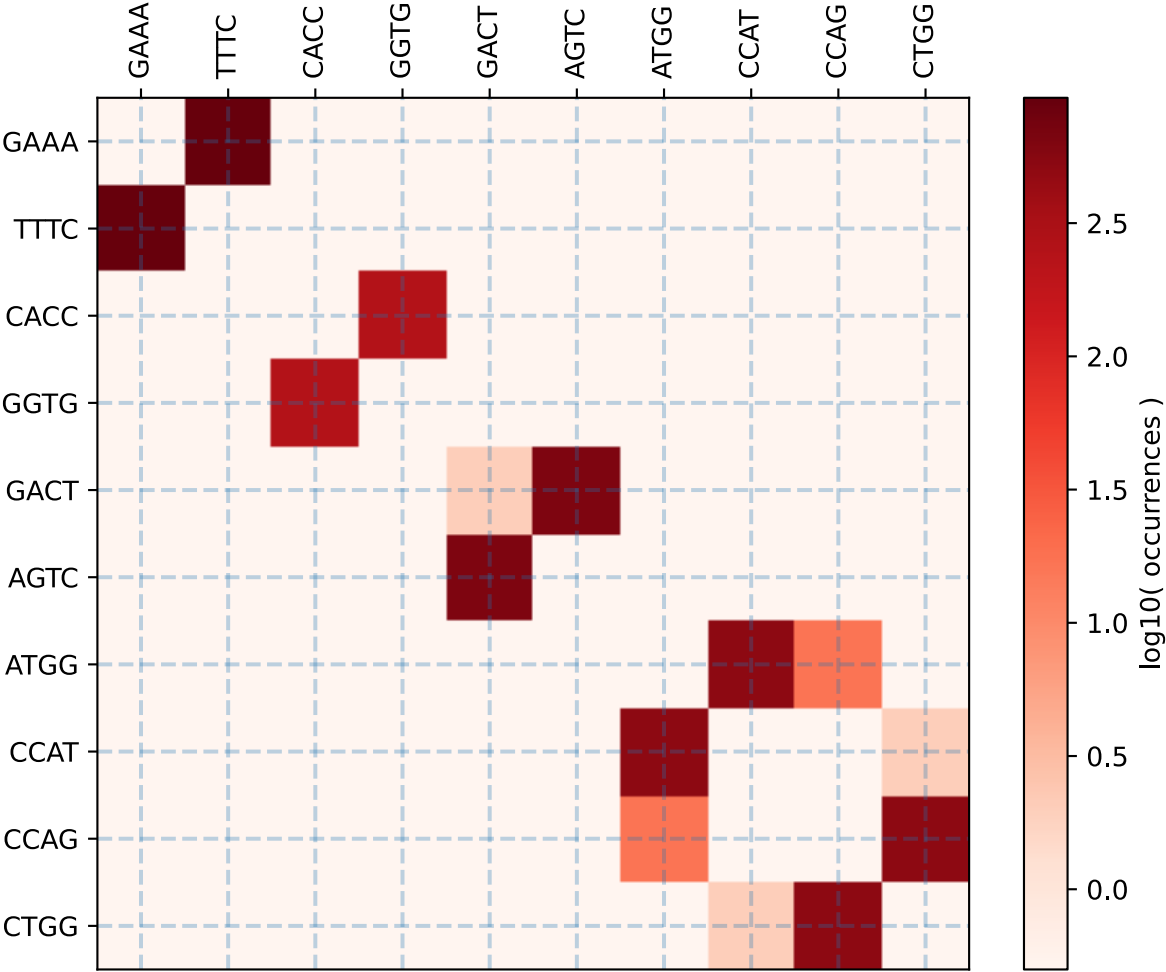
**Warning!** Weakly annealing overhang(s): CACC/GGTG

**Warning!** The following may have self-misannealing issues: AGTC/GACT

**Warning!** The following pairs may have misannealing issues: ATGG/CCAT ~ CCAG/CTGG

These overhang pairs (including reverse complements) have fewer than 2 differences: ATGG/CCAT ~ CCAG/CTGG ;

Please see the Appendix on the last page for an explanation of details.





# GAAA

# TTTC

GC content: 25 %.

Has 3 identical bases in a row. However, this has not shown to be very important.

Can form the following amino acids in 6 translation frames:

E[NITRSMK]

[LT\*RSEMKQPAWVG]K

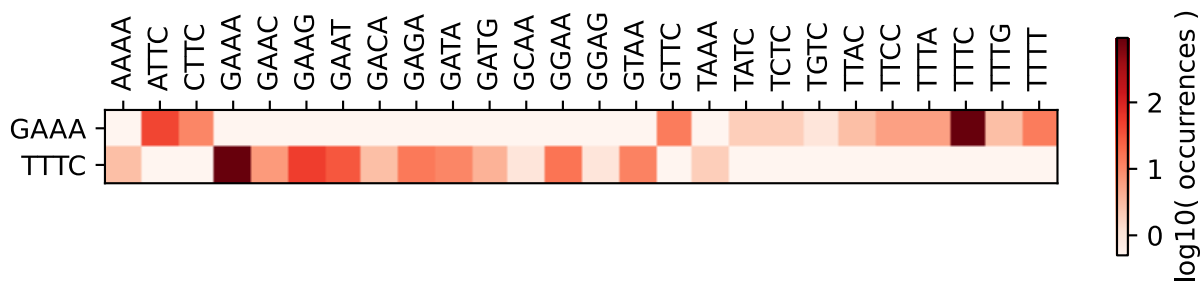
[RG\*][NK]

F[LHRPQ]

[NIFLHSTRCDPAYGV]F

[FILV][S]

Misannealing overhangs:





# CACC

# GGTG

GC content: **75 %**.

Can form the following amino acids in 6 translation frames:

H[LHRPQ]

[NIFLHSTRCDPAYGV]T

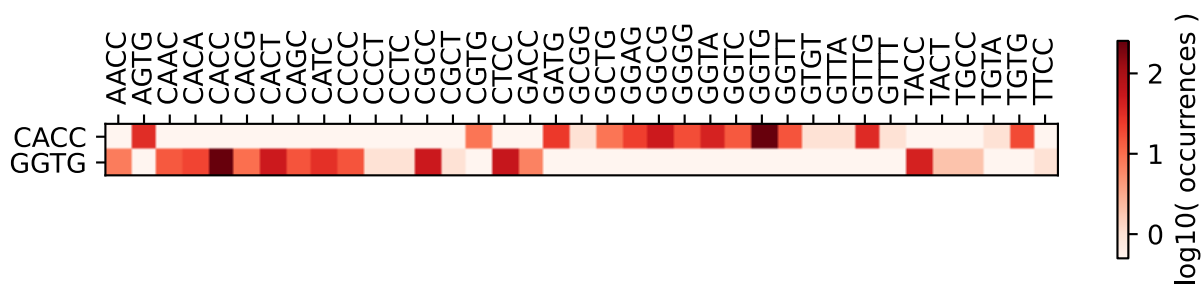
[TPAS][P]

G[EDAGV]

[LT\*RSEMKQPAWVG]V

[RGW][\*CW]

Misannealing overhangs:





# AGTC

# GACT

GC content: **50 %**.

Can form the following amino acids in 6 translation frames:

S[LHRPQ]

[ILT\*RSEPKQAGV]V

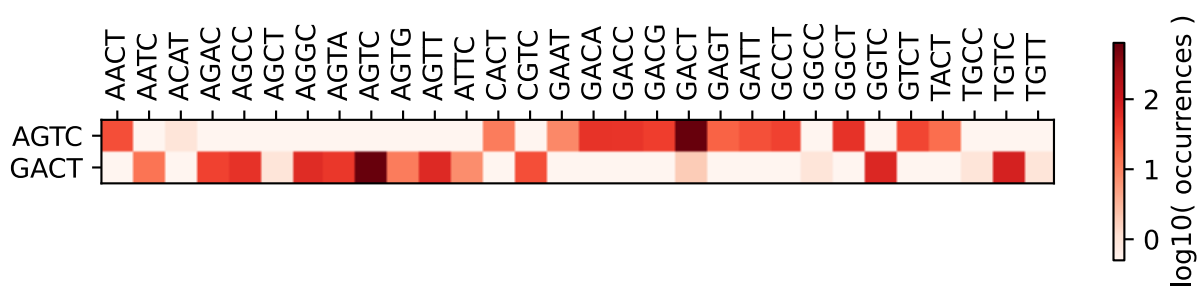
[\*KEQ][S]

D[FL\*SCYW]

[LT\*RSEMKQPAWVG]T

[RG\*][L]

Misannealing overhangs:





# ATGG

# CCAT

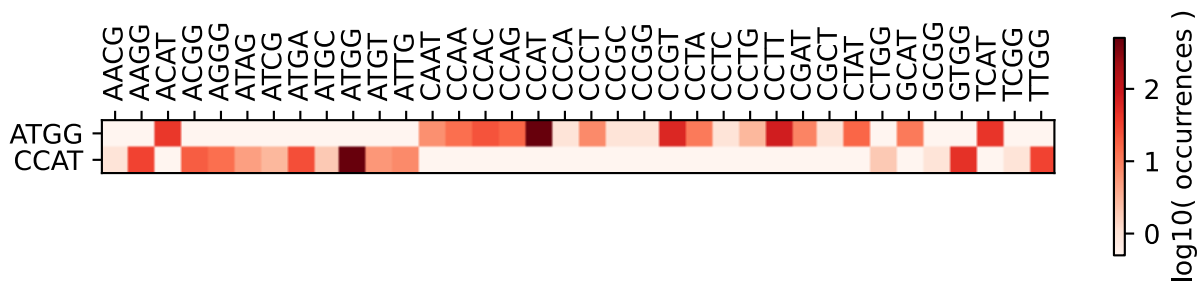
GC content: **50 %**.

The overhang contains the start codon ATG.

Can form the following amino acids in 6 translation frames:

M[EDAGV]  
[ILT\*RSEPKQAGV]W  
[NDYH][G]  
P[FL\*SCYW]  
[NIFLHSTRCDPAYGV]H  
[TPAS][IM]

Misannealing overhangs:





# CCAG

# CTGG

GC content: **75 %**.

Can form the following amino acids in 6 translation frames:

P[EDAGV]

[NIFLHSTRCDPAYGV]Q

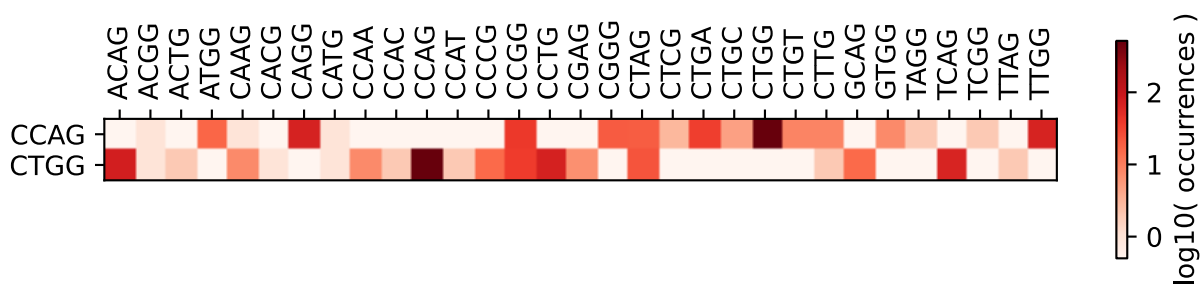
[TPAS][RS]

L[EDAGV]

[NIFLHSTRCDPAYGV]W

[TPAS][G]

Misannealing overhangs:



# Appendix

The report consists of 3 sections: results, overhangs, appendix.

## Result page(s)

The first page describes the overhang set. The result is also summarised with a symbol:

- ☑ : good overhang set
- ⚠ : warning; there are ways to significantly improve the set
- ☒ : error: the set cannot be used for DNA assembly

## Overhang pages

Each overhang is also analysed separately. The result is summarised with a symbol:

- ☑ : good overhang
- ☒ : unusable palindromic sequence

**Overhangs** are unpaired nucleotides at the end of a double-stranded linear DNA molecule. These overhangs create "sticky" (non-blunt) DNA ends. Overhangs can be on either strand; 5' or 3' overhangs. During DNA assembly, these overhangs are created by type IIS restriction enzymes, and ligated with another DNA with a complementary overhang, by ligases. The type of the restriction enzyme and the ligase influences the misannealing rate, which is displayed in a summary plot. For more details, see the [Tatapov](#) package and Pryor et al. ([PLoS ONE \(2020\) 15\(9\): e0238592](#)).

After the assembly, these overhangs remain in the sequence as fusion sites ("scars"). If this is in a coding sequence (CDS), then addition of nucleotides can ensure that the two joining parts remain in the same translation frame. By carefully choosing which nucleotides we add, we can select suitable amino acids, or start/stop codons. For each overhang, a list of codons is also displayed that shows translation options. The [GeneDom](#) package can be used for automating the addition of these nucleotides, the overhangs and the enzyme sites.

### Overhang sets

Use the [GoldenHinges](#) Python package to generate a set of mutually compatible overhangs that can be used for DNA assembly.