

Compendium of overhangs

There are 136 overhangs in this report. The restriction enzyme used in this compendium is Esp3I.

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





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:

K[EAVGD]

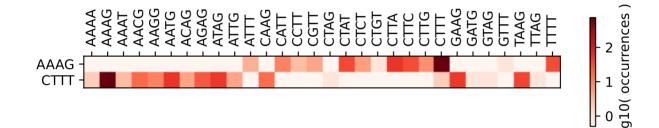
[IESAV*KQGPLRT]K

[*EKQ][RS]

L[SFW*CLY]

[DINSFAVRCGPLYHT]F

[APST][FL]







Extreme GC content: 0 %.

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:

K[SFW*CLY]

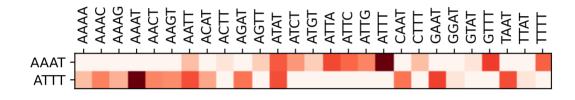
[IESAV*KQGPLRT]N

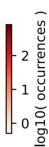
[*EKQ][MI]

I[SFW*CLY]

[IESAV*KQGPLRT]F

[NDYH][FL]







GGTT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

N[PLRHQ]

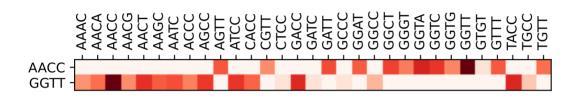
[IESAV*KQGPLRT]T

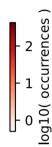
[*EKQ][P]

G[SFW*CLY]

[GESWAV*KQMPLRT]V

[WGR][FL]







CGTT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

N[EAVGD]

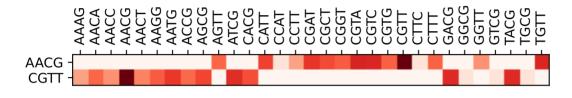
[IESAV*KQGPLRT]T

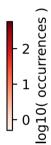
[*EKQ][R]

R[SFW*CLY]

[DINSFAVRCGPLYHT]V

[APST][FL]







GCTT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

K[PLRHQ]

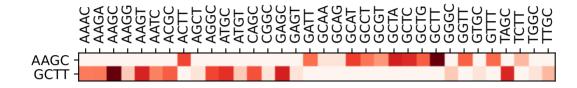
[IESAV*KQGPLRT]S

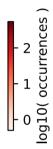
[*EKQ][A]

A[SFW*CLY]

[GESWAV*KQMPLRT]L

[GRCS][FL]







CATT

GC content: 25 %.

The overhang contains the start codon ATG.

Can form the following amino acids in 6 translation frames:

N[EAVGD]

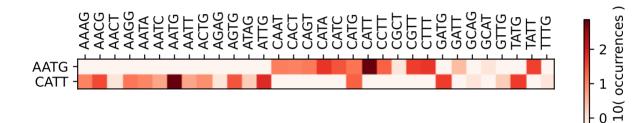
[IESAV*KQGPLRT]M

[*EKQ][W*C]

 $\mathsf{H}[\mathsf{SFW}^*\mathsf{CLY}]$

[DINSFAVRCGPLYHT]I

[APST][FL]







Extreme GC content: 0 %.

The overhang is palindromic, cannot be used for DNA assembly.

Can form the following amino acids in 6 translation frames:

N[SFW*CLY]

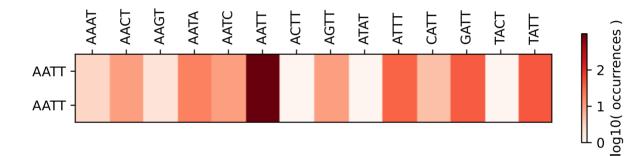
[IESAV*KQGPLRT]I

[*EKQ][FL]

N[SFW*CLY]

[IESAV*KQGPLRT]I

[*EKQ][FL]





TTGT

GC content: 25 %.

Can form the following amino acids in 6 translation frames:

T[INSKMRT]

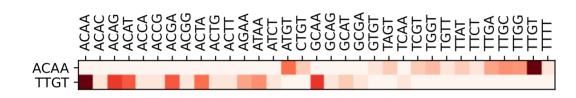
[IESAV*KQGPLRT]Q

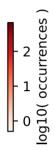
[NDYH][NK]

L[SFW*CLY]

[DINSFAVRCGPLYHT]C

[FVIL][V]







ATGT

GC content: 25 %.

The overhang contains the start codon ATG.

Can form the following amino acids in 6 translation frames:

T[SFW*CLY]

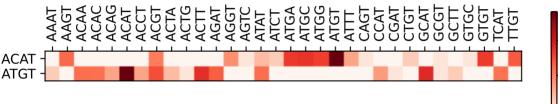
[IESAV*KQGPLRT]H

[NDYH][MI]

M[SFW*CLY]

[IESAV*KQGPLRT]C

[NDYH][V]





TGGT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

T[INSKMRT]

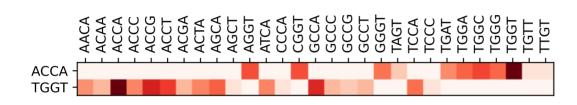
[IESAV*KQGPLRT]P

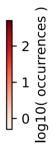
[NDYH][HQ]

W[SFW*CLY]

[DINSFAVRCGPLYHT]G

[MVL][V]







GGGT

GC content: 75 %.

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

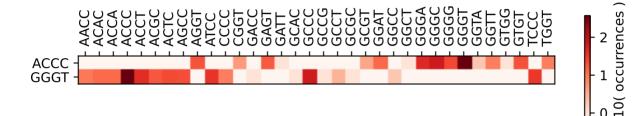
Can form the following amino acids in 6 translation frames:

T[PLRHQ]
[IESAV*KQGPLRT]P
[NDYH][P]

G[SFW*CLY]

[GESWAV*KQMPLRT]G

[WGR][V]





TCGT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

T[INSKMRT]

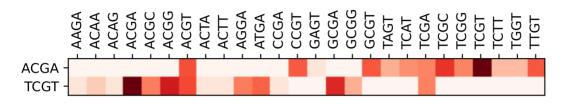
[IESAV*KQGPLRT]R

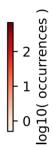
[NDYH][ED]

S[SFW*CLY]

[DINSFAVRCGPLYHT]R

[FVIL][V]







GCGT

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

T[PLRHQ]

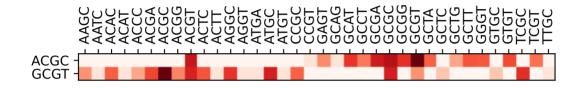
[IESAV*KQGPLRT]R

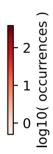
[NDYH][A]

A[SFW*CLY]

[GESWAV*KQMPLRT]R

[GRCS][V]







ACGT

ACGT

GC content: 50 %.

The overhang is palindromic, cannot be used for DNA assembly.

Can form the following amino acids in 6 translation frames:

T[SFW*CLY]

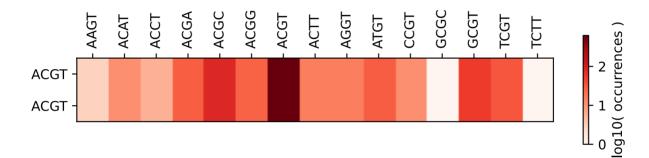
[IESAV*KQGPLRT]R

[NDYH][V]

T[SFW*CLY]

[IESAV*KQGPLRT]R

[NDYH][V]





TAGT

GC content: 25 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

T[INSKMRT]

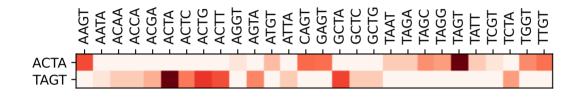
[IESAV*KQGPLRT]L

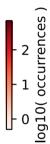
[NDYH][Y*]

*[SFW*CLY]

[DINSFAVRCGPLYHT]S

[VIL][V]







GAGT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

T[PLRHQ]

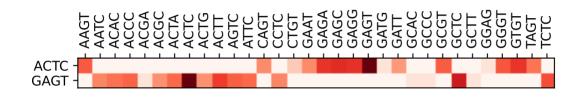
[IESAV*KQGPLRT]L

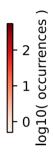
[NDYH][S]

E[SFW*CLY]

[GESWAV*KQMPLRT]S

[G*R][V]







CAGT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

T[EAVGD]

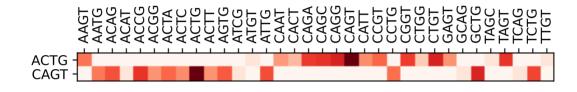
[IESAV*KQGPLRT]L

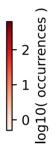
[NDYH][W*C]

Q[SFW*CLY]

[DINSFAVRCGPLYHT]S

[APST][V]







ACTT

GC content: 25 %.

Can form the following amino acids in 6 translation frames:

K[SFW*CLY]

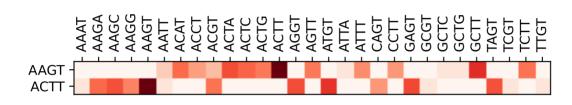
[IESAV*KQGPLRT]S

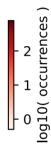
[*EKQ][V]

T[SFW*CLY]

[IESAV*KQGPLRT]L

[NDYH][FL]







TTCT

GC content: 25 %.

Can form the following amino acids in 6 translation frames:

R[INSKMRT]

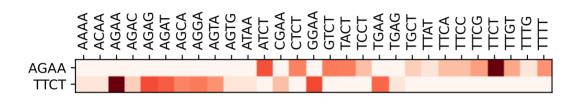
[IESAV*KQGPLRT]E

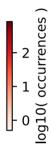
[*EKQ][NK]

F[SFW*CLY]

[DINSFAVRCGPLYHT]S

[FVIL][L]







TGCT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

S[INSKMRT]

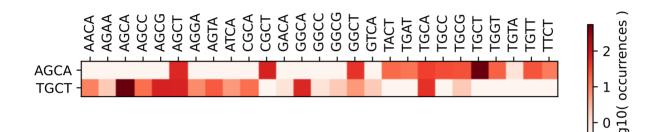
[IESAV*KQGPLRT]A

[*EKQ][HQ]

C[SFW*CLY]

[DINSFAVRCGPLYHT]A

[MVL][L]





GGCT

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

S[PLRHQ]

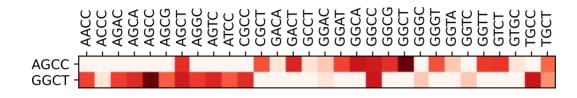
[IESAV*KQGPLRT]A

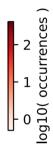
[*EKQ][P]

G[SFW*CLY]

[GESWAV*KQMPLRT]A

[WGR][L]







CGCT

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

S[EAVGD]

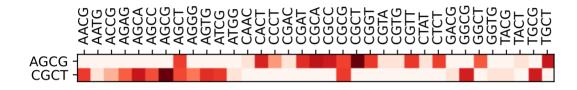
[IESAV*KQGPLRT]A

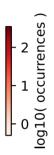
[*EKQ][R]

R[SFW*CLY]

[DINSFAVRCGPLYHT]A

[APST][L]







AGCT

GC content: 50 %.

The overhang is palindromic, cannot be used for DNA assembly.

Can form the following amino acids in 6 translation frames:

 $\mathsf{S}[\mathsf{SFW}^*\mathsf{CLY}]$

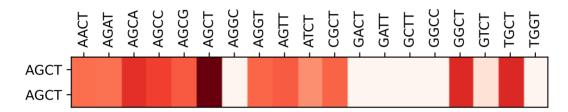
[IESAV*KQGPLRT]A

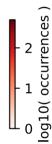
[*EKQ][L]

S[SFW*CLY]

[IESAV*KQGPLRT]A

[*EKQ][L]







GCCT

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

R[PLRHQ]

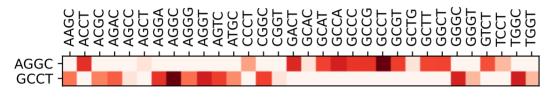
[IESAV*KQGPLRT]G

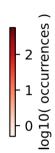
[*EKQ][A]

A[SFW*CLY]

[GESWAV*KQMPLRT]P

[GRCS][L]







AGGT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

T[SFW*CLY]

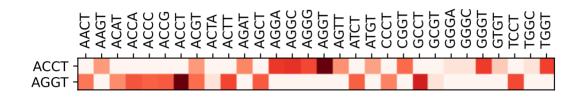
[IESAV*KQGPLRT]P

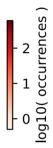
[NDYH][L]

R[SFW*CLY]

[IESAV*KQGPLRT]G

[*EKQ][V]







TACT

GC content: 25 %.

Can form the following amino acids in 6 translation frames:

S[INSKMRT]

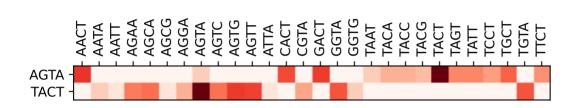
[IESAV*KQGPLRT]V

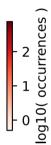
[*EKQ][Y*]

Y[SFW*CLY]

[DINSFAVRCGPLYHT]T

[VIL][L]







AGTT

GC content: 25 %.

Can form the following amino acids in 6 translation frames:

N[SFW*CLY]

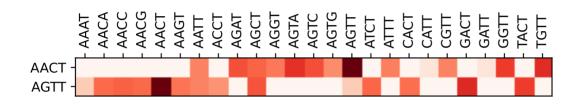
[IESAV*KQGPLRT]T

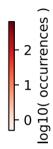
[*EKQ][L]

S[SFW*CLY]

[IESAV*KQGPLRT]V

[*EKQ][FL]







GTAT

GC content: 25 %.

Can form the following amino acids in 6 translation frames:

I[PLRHQ]

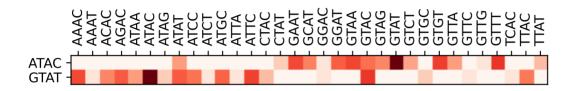
[IESAV*KQGPLRT]Y

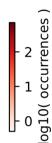
[NDYH][T]

V[SFW*CLY]

[GESWAV*KQMPLRT]Y

[GRCS][MI]







ATAT

Extreme GC content: 0 %.

The overhang is palindromic, cannot be used for DNA assembly.

Can form the following amino acids in 6 translation frames:

I[SFW*CLY]

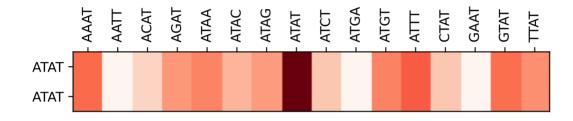
[IESAV*KQGPLRT]Y

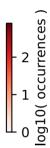
[NDYH][MI]

I[SFW*CLY]

[IESAV*KQGPLRT]Y

[NDYH][MI]







TGAT

GC content: 25 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

I[INSKMRT]

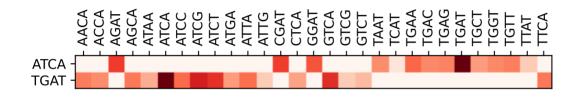
[IESAV*KQGPLRT]S

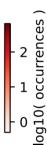
[NDYH][HQ]

*[SFW*CLY]

[DINSFAVRCGPLYHT]D

[MVL][MI]







GGAT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

I[PLRHQ]

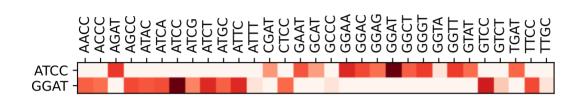
[IESAV*KQGPLRT]S

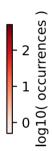
[NDYH][P]

G[SFW*CLY]

[GESWAV*KQMPLRT]D

[WGR][MI]







ATCT

GC content: 25 %.

Can form the following amino acids in 6 translation frames:

R[SFW*CLY]

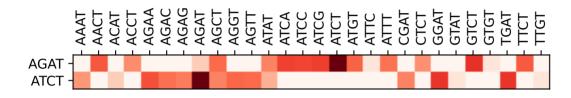
[IESAV*KQGPLRT]D

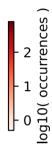
[*EKQ][MI]

I[SFW*CLY]

[IESAV*KQGPLRT]S

[NDYH][L]







CCAT

GC content: 50 %.

The overhang contains the start codon ATG.

Can form the following amino acids in 6 translation frames:

M[EAVGD]

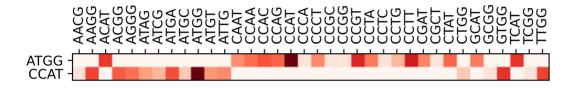
[IESAV*KQGPLRT]W

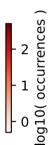
[NDYH][G]

P[SFW*CLY]

[DINSFAVRCGPLYHT]H

[APST][MI]







GAAT

GC content: 25 %.

Can form the following amino acids in 6 translation frames:

I[PLRHQ]

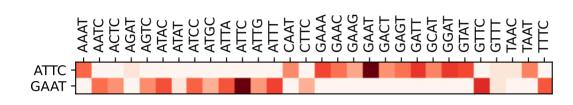
[IESAV*KQGPLRT]F

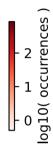
[NDYH][S]

E[SFW*CLY]

[GESWAV*KQMPLRT]N

[G*R][MI]







CAAT

GC content: 25 %.

Can form the following amino acids in 6 translation frames:

I[EAVGD]

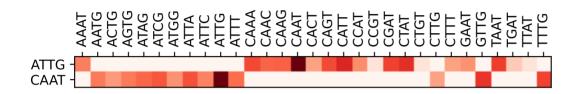
[IESAV*KQGPLRT]L

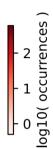
[NDYH][W*C]

Q[SFW*CLY]

[DINSFAVRCGPLYHT]N

[APST][MI]







TTTG

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:

Q[INSKMRT]

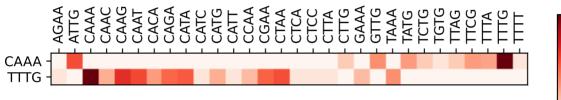
[DINSFAVRCGPLYHT]K

[APST][NK]

F[EAVGD]

[DINSFAVRCGPLYHT]L

[FVIL][W*C]





GTTG

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

Q[PLRHQ]

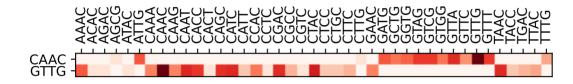
[DINSFAVRCGPLYHT]N

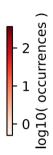
[APST][T]

V[EAVGD]

[GESWAV*KQMPLRT]L

[GRCS][W*C]







TGTG

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

H[INSKMRT]

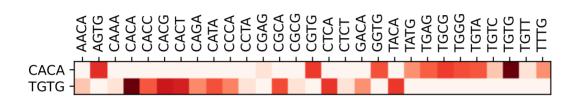
[DINSFAVRCGPLYHT]T

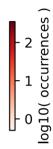
[APST][HQ]

C[EAVGD]

[DINSFAVRCGPLYHT]V

[MVL][W*C]







GGTG

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

H[PLRHQ]

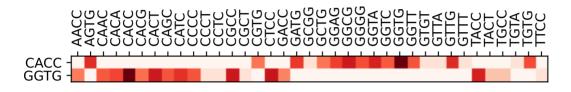
[DINSFAVRCGPLYHT]T

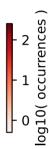
[APST][P]

G[EAVGD]

[GESWAV*KQMPLRT]V

[WGR][W*C]







CGTG

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

H[EAVGD]

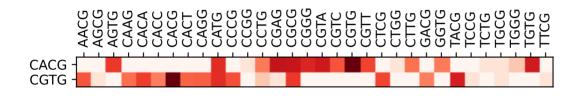
[DINSFAVRCGPLYHT]T

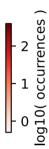
[APST][R]

R[EAVGD]

[DINSFAVRCGPLYHT]V

[APST][W*C]







CACT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

S[EAVGD]

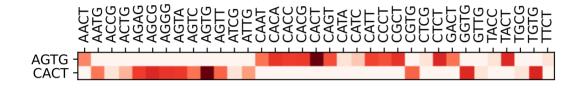
[IESAV*KQGPLRT]V

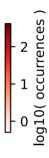
[*EKQ][W*C]

H[SFW*CLY]

[DINSFAVRCGPLYHT]T

[APST][L]







TCTG

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

Q[INSKMRT]

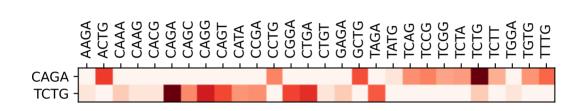
[DINSFAVRCGPLYHT]R

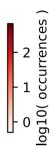
[APST][ED]

S[EAVGD]

[DINSFAVRCGPLYHT]L

[FVIL][W*C]







CATG

GC content: 50 %.

The overhang is palindromic, cannot be used for DNA assembly.

The overhang contains the start codon ATG.

Can form the following amino acids in 6 translation frames:

H[EAVGD]

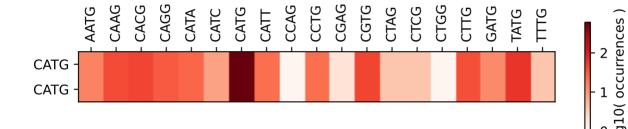
[DINSFAVRCGPLYHT]M

[APST][W*C]

H[EAVGD]

[DINSFAVRCGPLYHT]M

[APST][W*C]





TTGG

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

P[INSKMRT]

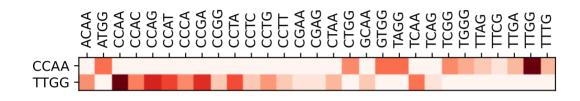
[DINSFAVRCGPLYHT]Q

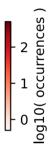
[APST][NK]

L[EAVGD]

[DINSFAVRCGPLYHT]W

[FVIL][G]







TGGG

GC content: 75 %.

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:

P[INSKMRT]

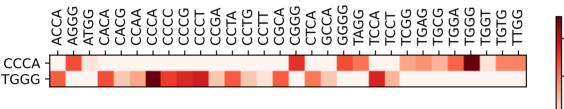
[DINSFAVRCGPLYHT]P

[APST][HQ]

W[EAVGD]

[DINSFAVRCGPLYHT]G

[MVL][G]







CCCT

GC content: 75 %.

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:

R[EAVGD]

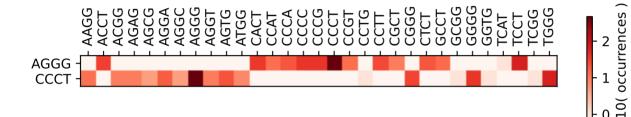
[IESAV*KQGPLRT]G

[*EKQ][G]

P[SFW*CLY]

[DINSFAVRCGPLYHT]P

[APST][L]





CCGG

CCGG

Extreme GC content: 100 %.

The overhang is palindromic, cannot be used for DNA assembly.

Can form the following amino acids in 6 translation frames:

P[EAVGD]

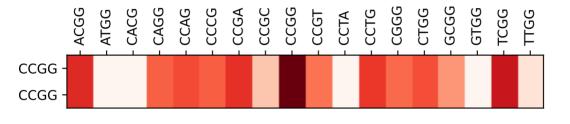
[DINSFAVRCGPLYHT]R

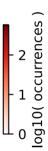
[APST][G]

P[EAVGD]

[DINSFAVRCGPLYHT]R

[APST][G]







CCGT

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

T[EAVGD]

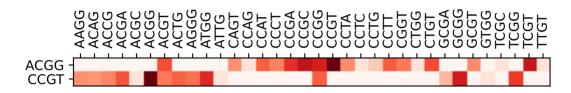
[IESAV*KQGPLRT]R

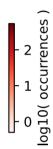
[NDYH][G]

P[SFW*CLY]

[DINSFAVRCGPLYHT]R

[APST][V]







TAGG

GC content: 50 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

P[INSKMRT]

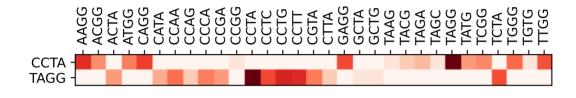
[DINSFAVRCGPLYHT]L

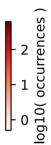
[APST][Y*]

*[EAVGD]

[DINSFAVRCGPLYHT]R

[VIL][G]







GAGG

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

P[PLRHQ]

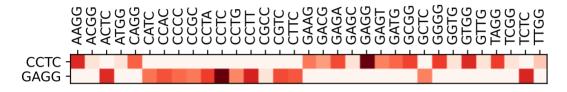
[DINSFAVRCGPLYHT]L

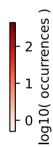
[APST][S]

E[EAVGD]

[GESWAV*KQMPLRT]R

[G*R][G]







CCTG

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

Q[EAVGD]

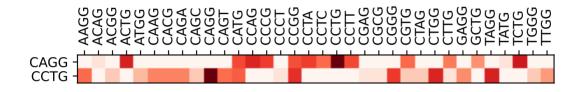
[DINSFAVRCGPLYHT]R

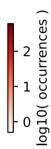
[APST][G]

P[EAVGD]

[DINSFAVRCGPLYHT]L

[APST][W*C]







CCTT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

K[EAVGD]

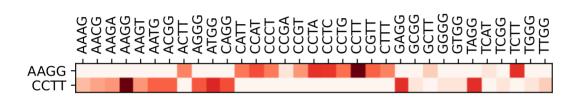
[IESAV*KQGPLRT]R

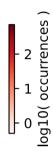
[*EKQ][G]

P[SFW*CLY]

[DINSFAVRCGPLYHT]L

[APST][FL]







CTCG

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

R[EAVGD]

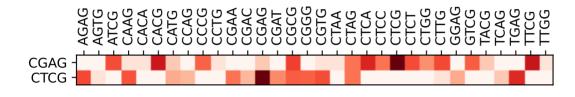
[DINSFAVRCGPLYHT]E

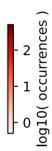
[APST][RS]

L[EAVGD]

[DINSFAVRCGPLYHT]S

[APST][R]







CGAT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

I[EAVGD]

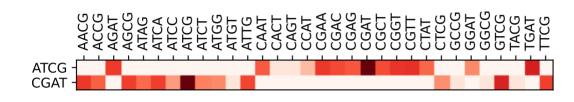
[IESAV*KQGPLRT]S

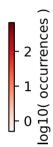
[NDYH][R]

R[SFW*CLY]

[DINSFAVRCGPLYHT]D

[APST][MI]





CGCG

CGCG

Extreme GC content: 100 %.

The overhang is palindromic, cannot be used for DNA assembly.

Can form the following amino acids in 6 translation frames:

R[EAVGD]

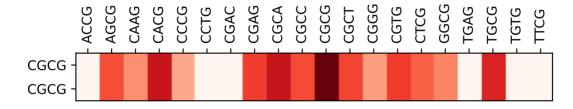
[DINSFAVRCGPLYHT]A

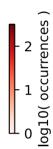
[APST][R]

R[EAVGD]

[DINSFAVRCGPLYHT]A

[APST][R]







CGGG

Extreme GC content: 100 %.

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:

P[EAVGD]

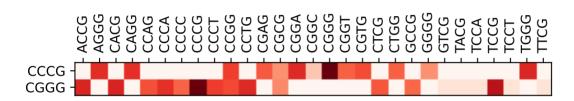
[DINSFAVRCGPLYHT]P

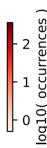
[APST][R]

R[EAVGD]

[DINSFAVRCGPLYHT]G

[APST][G]







CGGT

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

T[EAVGD]

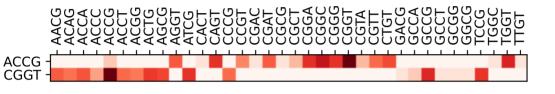
[IESAV*KQGPLRT]P

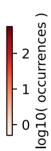
[NDYH][R]

R[SFW*CLY]

[DINSFAVRCGPLYHT]G

[APST][V]







GACG

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

R[PLRHQ]

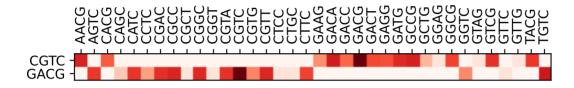
[DINSFAVRCGPLYHT]V

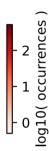
[APST][S]

D[EAVGD]

[GESWAV*KQMPLRT]T

[G*R][R]







TTAG

GC content: 25 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

L[INSKMRT]

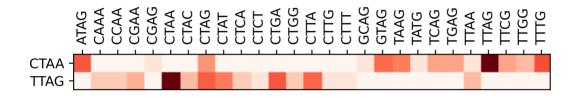
[DINSFAVRCGPLYHT]*

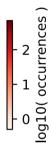
[APST][NK]

L[EAVGD]

[DINSFAVRCGPLYHT]*

[FVIL][RS]







CTAG

GC content: 50 %.

The overhang is palindromic, cannot be used for DNA assembly.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

L[EAVGD]

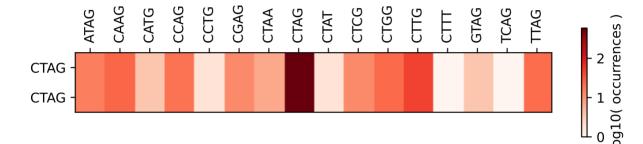
[DINSFAVRCGPLYHT]*

[APST][RS]

L[EAVGD]

[DINSFAVRCGPLYHT]*

[APST][RS]





CTAT

GC content: 25 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

I[EAVGD]

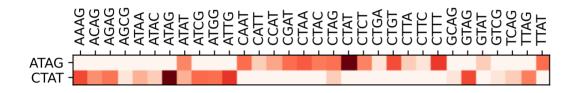
[IESAV*KQGPLRT]*

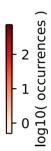
[NDYH][RS]

L[SFW*CLY]

[DINSFAVRCGPLYHT]Y

[APST][MI]







GGAG

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

L[PLRHQ]

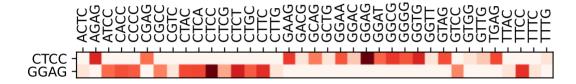
[DINSFAVRCGPLYHT]S

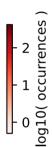
[APST][P]

G[EAVGD]

[GESWAV*KQMPLRT]E

[WGR][RS]







CTCT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

R[EAVGD]

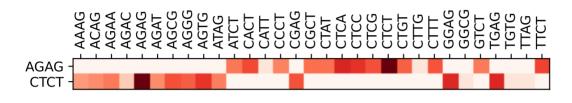
[IESAV*KQGPLRT]E

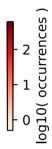
[*EKQ][RS]

L[SFW*CLY]

[DINSFAVRCGPLYHT]S

[APST][L]







CTGG

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

P[EAVGD]

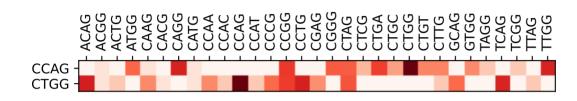
[DINSFAVRCGPLYHT]Q

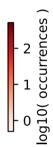
[APST][RS]

L[EAVGD]

[DINSFAVRCGPLYHT]W

[APST][G]







CTGT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

T[EAVGD]

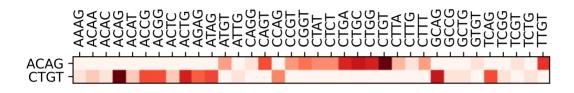
[IESAV*KQGPLRT]Q

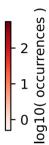
[NDYH][RS]

L[SFW*CLY]

[DINSFAVRCGPLYHT]C

[APST][V]







TAAG

GC content: 25 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

L[INSKMRT]

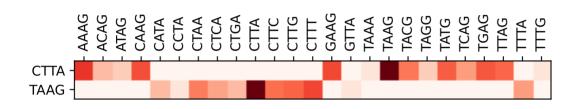
[DINSFAVRCGPLYHT]L

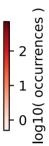
[APST][Y*]

*[EAVGD]

[DINSFAVRCGPLYHT]K

[VIL][RS]







CTTG

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

Q[EAVGD]

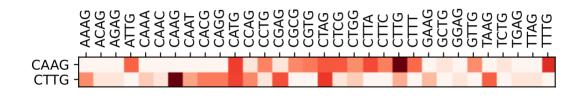
[DINSFAVRCGPLYHT]K

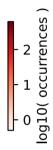
[APST][RS]

L[EAVGD]

[DINSFAVRCGPLYHT]L

[APST][W*C]







GTTC

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

E[PLRHQ]

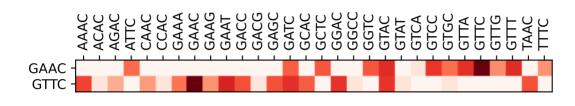
[GESWAV*KQMPLRT]N

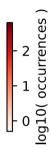
[G*R][T]

V[PLRHQ]

[GESWAV*KQMPLRT]F

[GRCS][S]







GAAG

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

L[PLRHQ]

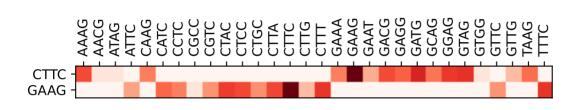
[DINSFAVRCGPLYHT]F

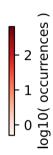
[APST][S]

E[EAVGD]

[GESWAV*KQMPLRT]K

[G*R][RS]







GACT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

S[PLRHQ]

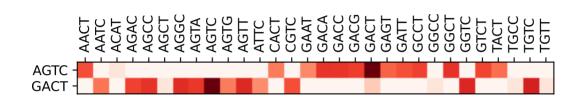
[IESAV*KQGPLRT]V

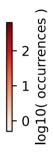
[*EKQ][S]

D[SFW*CLY]

[GESWAV*KQMPLRT]T

[G*R][L]







GC content: 50 %.

Can form the following amino acids in 6 translation frames:

E[INSKMRT]

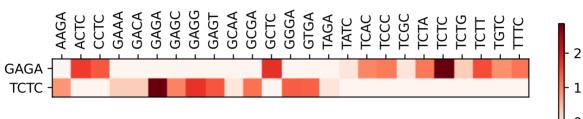
[GESWAV*KQMPLRT]R

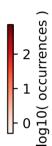
[G*R][ED]

S[PLRHQ]

[DINSFAVRCGPLYHT]L

[FVIL][S]







GCTC

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

E[PLRHQ]

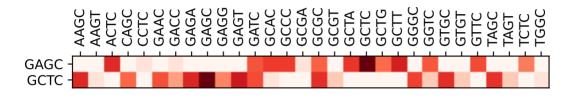
[GESWAV*KQMPLRT]S

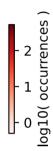
[G*R][A]

A[PLRHQ]

[GESWAV*KQMPLRT]L

[GRCS][S]







GATC

GC content: 50 %.

The overhang is palindromic, cannot be used for DNA assembly.

Can form the following amino acids in 6 translation frames:

D[PLRHQ]

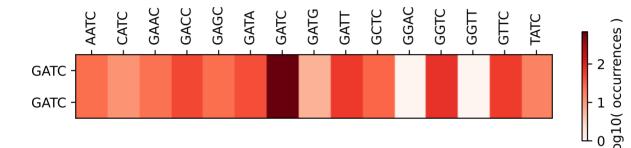
[GESWAV*KQMPLRT]I

[G*R][S]

D[PLRHQ]

[GESWAV*KQMPLRT]I

[G*R][S]





GATG

GC content: 50 %.

The overhang contains the start codon ATG.

Can form the following amino acids in 6 translation frames:

H[PLRHQ]

[DINSFAVRCGPLYHT]I

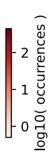
[APST][S]

D[EAVGD]

[GESWAV*KQMPLRT]M

 $[G^*R][W^*C]$







GATT

GC content: 25 %.

Can form the following amino acids in 6 translation frames:

N[PLRHQ]

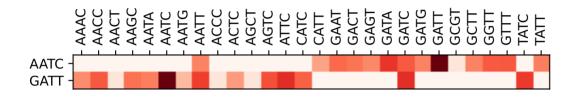
[IESAV*KQGPLRT]I

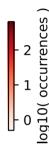
[*EKQ][S]

D[SFW*CLY]

[GESWAV*KQMPLRT]I

[G*R][FL]







GCAG

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

L[PLRHQ]

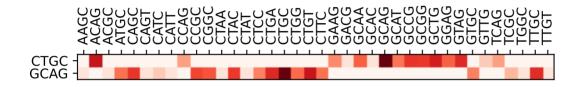
[DINSFAVRCGPLYHT]C

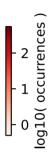
[APST][A]

A[EAVGD]

[GESWAV*KQMPLRT]Q

[GRCS][RS]







GCAT

GC content: 50 %.

The overhang contains the start codon ATG.

Can form the following amino acids in 6 translation frames:

M[PLRHQ]

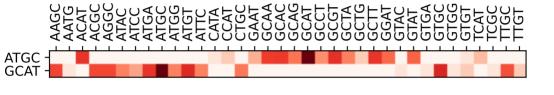
[IESAV*KQGPLRT]C

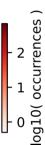
[NDYH][A]

A[SFW*CLY]

[GESWAV*KQMPLRT]H

[GRCS][MI]







TGGC

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

A[INSKMRT]

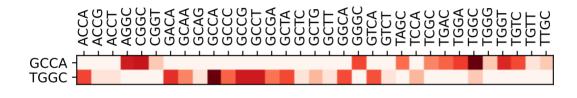
[GESWAV*KQMPLRT]P

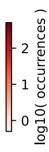
[GRCS][HQ]

W[PLRHQ]

[DINSFAVRCGPLYHT]G

[MVL][A]







CGGC

GCCG

Extreme GC content: 100 %.

Can form the following amino acids in 6 translation frames:

R[PLRHQ]

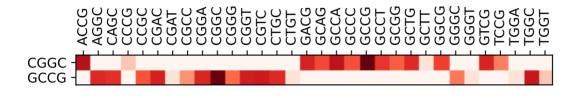
[DINSFAVRCGPLYHT]G

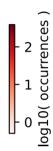
[APST][A]

A[EAVGD]

[GESWAV*KQMPLRT]P

[GRCS][R]







GCGC

GCGC

Extreme GC content: 100 %.

The overhang is palindromic, cannot be used for DNA assembly.

Can form the following amino acids in 6 translation frames:

A[PLRHQ]

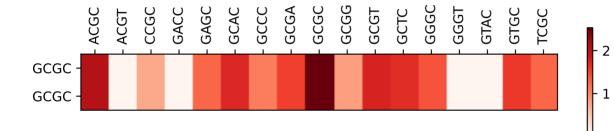
[GESWAV*KQMPLRT]R

[GRCS][A]

A[PLRHQ]

[GESWAV*KQMPLRT]R

[GRCS][A]





GCGG

Extreme GC content: 100 %.

Can form the following amino acids in 6 translation frames:

P[PLRHQ]

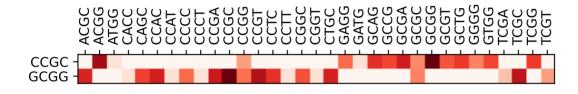
[DINSFAVRCGPLYHT]R

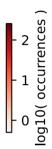
[APST][A]

A[EAVGD]

[GESWAV*KQMPLRT]R

[GRCS][G]







TAGC

GC content: 50 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

A[INSKMRT]

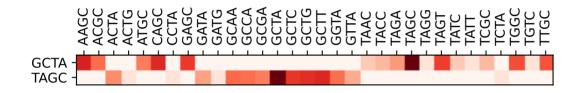
[GESWAV*KQMPLRT]L

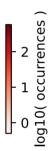
[GRCS][Y*]

*[PLRHQ]

[DINSFAVRCGPLYHT]S

[VIL][A]







GCTG

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

Q[PLRHQ]

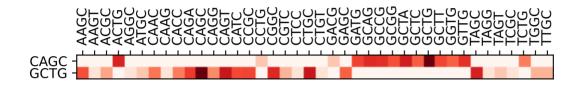
[DINSFAVRCGPLYHT]S

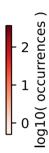
[APST][A]

A[EAVGD]

[GESWAV*KQMPLRT]L

[GRCS][W*C]







GTCC

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

G[PLRHQ]

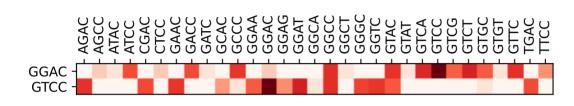
[GESWAV*KQMPLRT]D

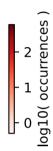
[WGR][T]

V[PLRHQ]

[GESWAV*KQMPLRT]S

[GRCS][P]







TGCC

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

G[INSKMRT]

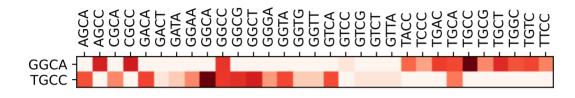
[GESWAV*KQMPLRT]A

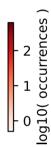
[WGR][HQ]

C[PLRHQ]

[DINSFAVRCGPLYHT]A

[MVL][P]







GGCC

Extreme GC content: 100 %.

The overhang is palindromic, cannot be used for DNA assembly.

Can form the following amino acids in 6 translation frames:

G[PLRHQ]

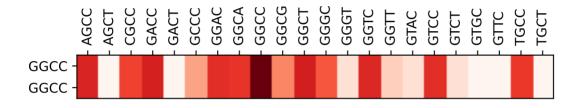
[GESWAV*KQMPLRT]A

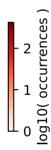
[WGR][P]

G[PLRHQ]

[GESWAV*KQMPLRT]A

[WGR][P]







GGCG

Extreme GC content: 100 %.

Can form the following amino acids in 6 translation frames:

R[PLRHQ]

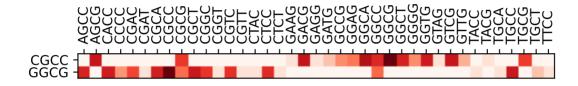
[DINSFAVRCGPLYHT]A

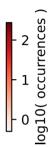
[APST][P]

G[EAVGD]

[GESWAV*KQMPLRT]A

[WGR][R]







TCCC

GC content: 75 %.

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:

G[INSKMRT]

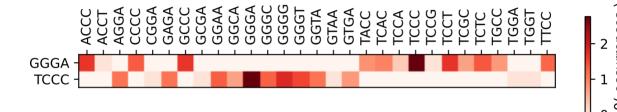
[GESWAV*KQMPLRT]G

[WGR][ED]

S[PLRHQ]

[DINSFAVRCGPLYHT]P

[FVIL][P]





GGGC

Extreme GC content: 100 %.

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:

A[PLRHQ]

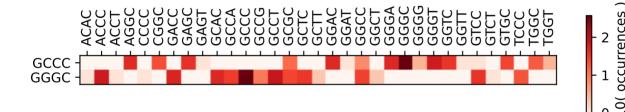
[GESWAV*KQMPLRT]P

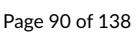
[GRCS][P]

G[PLRHQ]

[GESWAV*KQMPLRT]G

[WGR][A]







GGGG

Extreme GC content: 100 %.

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:

P[PLRHQ]

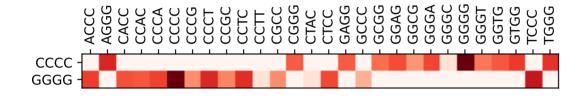
[DINSFAVRCGPLYHT]P

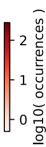
[APST][P]

G[EAVGD]

[GESWAV*KQMPLRT]G

[WGR][G]







TACC

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

G[INSKMRT]

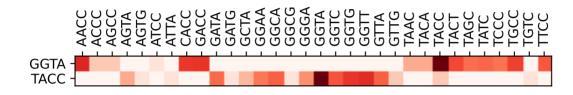
[GESWAV*KQMPLRT]V

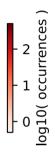
 $[WGR][Y^*]$

Y[PLRHQ]

[DINSFAVRCGPLYHT]T

[VIL][P]







GGTC

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

D[PLRHQ]

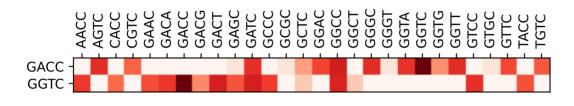
[GESWAV*KQMPLRT]T

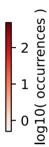
[G*R][P]

G[PLRHQ]

 $[\mathsf{GESWAV}^*\mathsf{KQMPLRT}]\mathsf{V}$

[WGR][S]







GTAC

GC content: 50 %.

The overhang is palindromic, cannot be used for DNA assembly.

Can form the following amino acids in 6 translation frames:

V[PLRHQ]

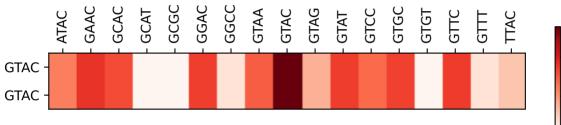
[GESWAV*KQMPLRT]Y

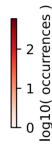
[GRCS][T]

V[PLRHQ]

[GESWAV*KQMPLRT]Y

[GRCS][T]







GTAG

GC content: 50 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

L[PLRHQ]

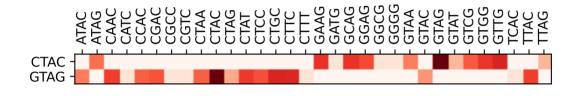
[DINSFAVRCGPLYHT]Y

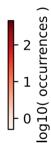
[APST][T]

V[EAVGD]

[GESWAV*KQMPLRT]*

[GRCS][RS]







GTCG

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

R[PLRHQ]

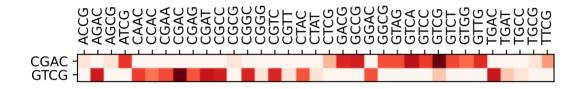
[DINSFAVRCGPLYHT]D

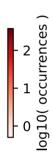
[APST][T]

V[EAVGD]

[GESWAV*KQMPLRT]S

[GRCS][R]







GTCT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

R[PLRHQ]

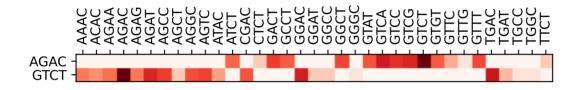
[IESAV*KQGPLRT]D

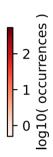
[*EKQ][T]

V[SFW*CLY]

[GESWAV*KQMPLRT]S

[GRCS][L]







TCAC

GC content: 50 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

V[INSKMRT]

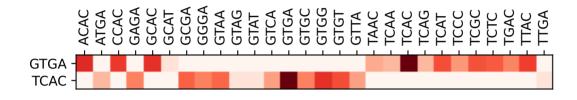
[GESWAV*KQMPLRT]*

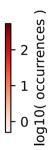
[GRCS][ED]

S[PLRHQ]

[DINSFAVRCGPLYHT]H

[FVIL][T]







GTGC

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

A[PLRHQ]

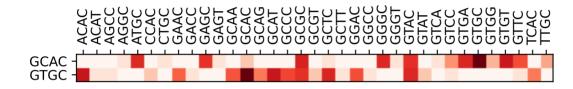
[GESWAV*KQMPLRT]H

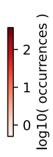
[GRCS][T]

V[PLRHQ]

[GESWAV*KQMPLRT]C

[GRCS][A]







GTGG

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

P[PLRHQ]

[DINSFAVRCGPLYHT]H

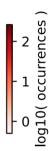
[APST][T]

V[EAVGD]

[GESWAV*KQMPLRT]W

[GRCS][G]







GTGT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

T[PLRHQ]

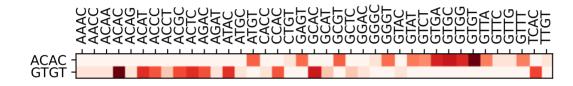
[IESAV*KQGPLRT]H

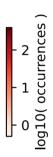
[NDYH][T]

V[SFW*CLY]

[GESWAV*KQMPLRT]C

[GRCS][V]







GTTT

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:

K[PLRHQ]

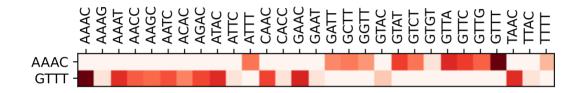
[IESAV*KQGPLRT]N

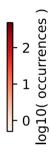
[*EKQ][T]

V[SFW*CLY]

[GESWAV*KQMPLRT]F

[GRCS][FL]







TAAC

GC content: 25 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

V[INSKMRT]

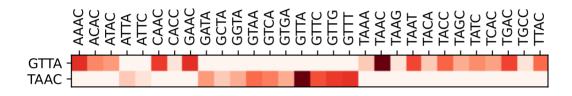
[GESWAV*KQMPLRT]L

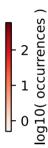
[GRCS][Y*]

*[PLRHQ]

[DINSFAVRCGPLYHT]N

[VIL][T]







TAAT

Extreme GC content: 0 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

I[INSKMRT]

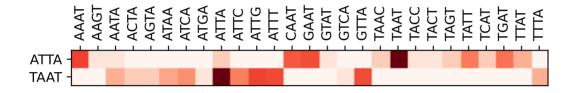
[IESAV*KQGPLRT]L

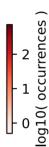
[NDYH][Y*]

*[SFW*CLY]

[DINSFAVRCGPLYHT]N

[VIL][MI]







TACG

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

R[INSKMRT]

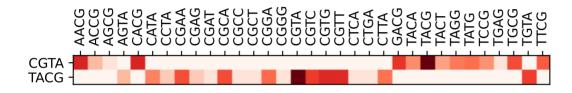
[DINSFAVRCGPLYHT]V

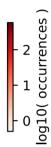
 $[APST][Y^*]$

Y[EAVGD]

[DINSFAVRCGPLYHT]T

[VIL][R]







TCTA

GC content: 25 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

*[INSKMRT]

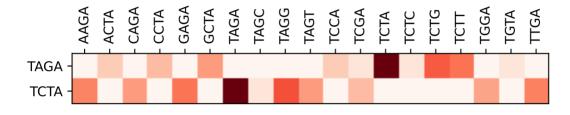
[DINSFAVRCGPLYHT]R

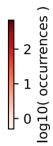
[VIL][ED]

S[INSKMRT]

[DINSFAVRCGPLYHT]L

[FVIL][Y*]







TATA

Extreme GC content: 0 %.

The overhang is palindromic, cannot be used for DNA assembly.

Can form the following amino acids in 6 translation frames:

Y[INSKMRT]

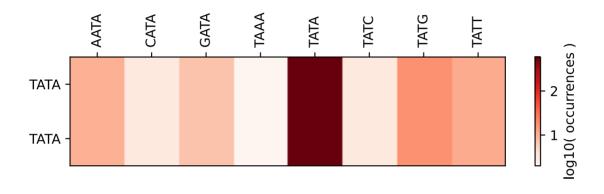
[DINSFAVRCGPLYHT]I

 $[VIL][Y^*]$

Y[INSKMRT]

[DINSFAVRCGPLYHT]I

 $[VIL][Y^*]$





TATC

GC content: 25 %.

Can form the following amino acids in 6 translation frames:

D[INSKMRT]

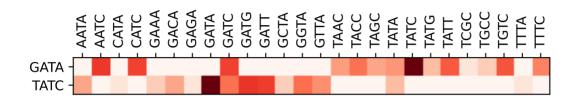
[GESWAV*KQMPLRT]I

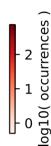
 $[G^*R][Y^*]$

Y[PLRHQ]

[DINSFAVRCGPLYHT]I

[VIL][S]







TATG

GC content: 25 %.

The overhang contains the start codon ATG.

Can form the following amino acids in 6 translation frames:

H[INSKMRT]

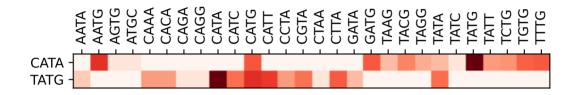
[DINSFAVRCGPLYHT]I

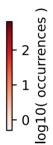
[APST][Y*]

Y[EAVGD]

[DINSFAVRCGPLYHT]M

[VIL][W*C]







TATT

Extreme GC content: 0 %.

Can form the following amino acids in 6 translation frames:

N[INSKMRT]

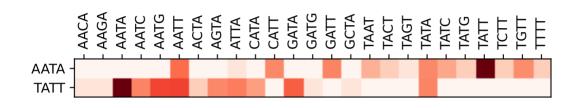
[IESAV*KQGPLRT]I

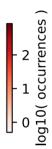
[*EKQ][Y*]

Y[SFW*CLY]

[DINSFAVRCGPLYHT]I

[VIL][FL]







TCAG

GC content: 50 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

L[INSKMRT]

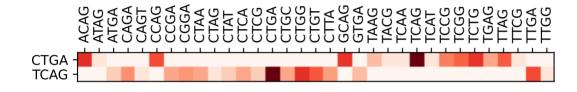
[DINSFAVRCGPLYHT]*

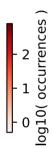
[APST][ED]

S[EAVGD]

[DINSFAVRCGPLYHT]Q

[FVIL][RS]







TCAT

GC content: 25 %.

The overhang contains the start codon ATG.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

M[INSKMRT]

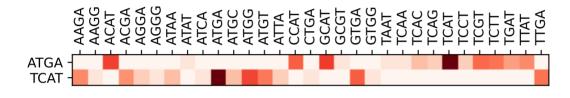
[IESAV*KQGPLRT]*

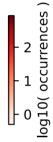
[NDYH][ED]

S[SFW*CLY]

[DINSFAVRCGPLYHT]H

[FVIL][MI]







TCCG

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

R[INSKMRT]

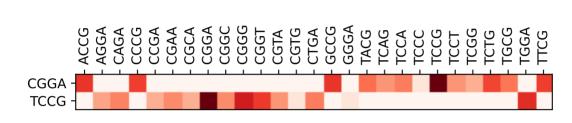
[DINSFAVRCGPLYHT]G

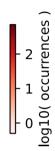
[APST][ED]

S[EAVGD]

[DINSFAVRCGPLYHT]P

[FVIL][R]







TCCT

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

R[INSKMRT]

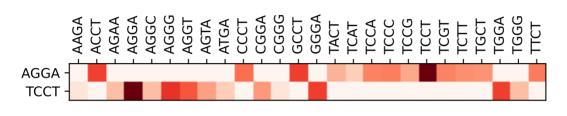
[IESAV*KQGPLRT]G

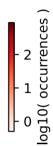
[*EKQ][ED]

S[SFW*CLY]

[DINSFAVRCGPLYHT]P

[FVIL][L]







TCGA

GC content: 50 %.

The overhang is palindromic, cannot be used for DNA assembly.

Can form the following amino acids in 6 translation frames:

S[INSKMRT]

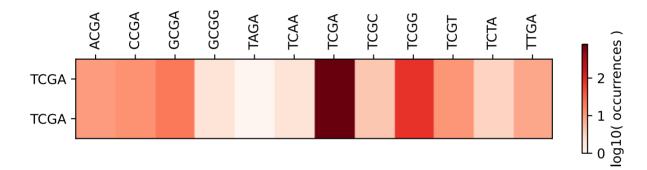
[DINSFAVRCGPLYHT]R

[FVIL][ED]

S[INSKMRT]

[DINSFAVRCGPLYHT]R

[FVIL][ED]





TCGC

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

A[INSKMRT]

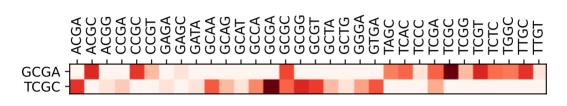
[GESWAV*KQMPLRT]R

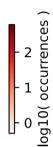
[GRCS][ED]

S[PLRHQ]

[DINSFAVRCGPLYHT]R

[FVIL][A]







TCGG

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

P[INSKMRT]

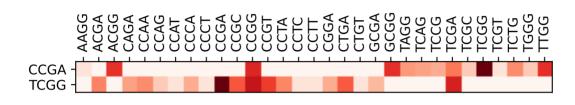
[DINSFAVRCGPLYHT]R

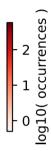
[APST][ED]

S[EAVGD]

[DINSFAVRCGPLYHT]R

[FVIL][G]







TCTT

GC content: 25 %.

Can form the following amino acids in 6 translation frames:

K[INSKMRT]

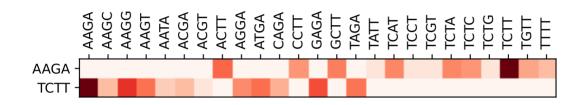
[IESAV*KQGPLRT]R

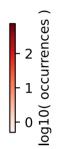
[*EKQ][ED]

S[SFW*CLY]

[DINSFAVRCGPLYHT]L

[FVIL][FL]







TGAC

GC content: 50 %.

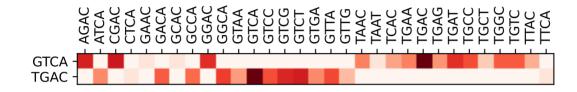
The overhang contains a stop codon (TAA, TAG or TGA).

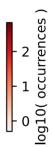
Can form the following amino acids in 6 translation frames:

V[INSKMRT] [GESWAV*KQMPLRT]S [GRCS][HQ] *[PLRHQ]

[DINSFAVRCGPLYHT]D

[MVL][T]







TGAG

GC content: 50 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

L[INSKMRT]

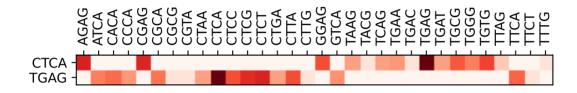
[DINSFAVRCGPLYHT]S

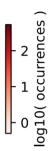
[APST][HQ]

*[EAVGD]

[DINSFAVRCGPLYHT]E

[MVL][RS]







TGCA

GC content: 50 %.

The overhang is palindromic, cannot be used for DNA assembly.

Can form the following amino acids in 6 translation frames:

C[INSKMRT]

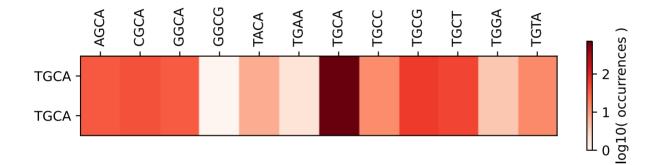
[DINSFAVRCGPLYHT]A

[MVL][HQ]

C[INSKMRT]

[DINSFAVRCGPLYHT]A

[MVL][HQ]





TGCG

GC content: 75 %.

Can form the following amino acids in 6 translation frames:

R[INSKMRT]

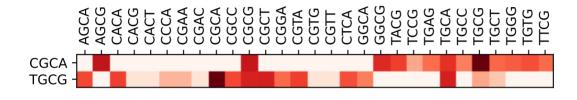
[DINSFAVRCGPLYHT]A

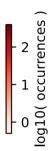
[APST][HQ]

C[EAVGD]

[DINSFAVRCGPLYHT]A

[MVL][R]







TGGA

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

S[INSKMRT]

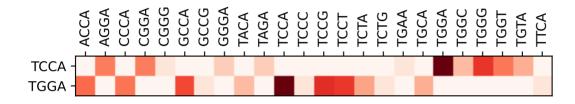
[DINSFAVRCGPLYHT]P

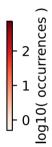
[FVIL][HQ]

W[INSKMRT]

[DINSFAVRCGPLYHT]G

[MVL][ED]







TGTA

GC content: 25 %.

Can form the following amino acids in 6 translation frames:

Y[INSKMRT]

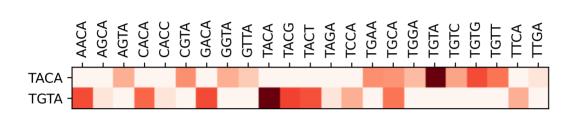
[DINSFAVRCGPLYHT]T

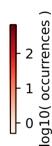
[VIL][HQ]

C[INSKMRT]

[DINSFAVRCGPLYHT]V

[MVL][Y*]







TGTC

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

D[INSKMRT]

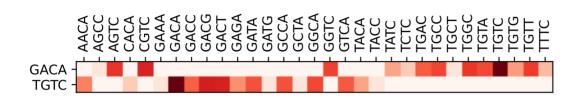
[GESWAV*KQMPLRT]T

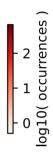
[G*R][HQ]

C[PLRHQ]

[DINSFAVRCGPLYHT]V

[MVL][S]







TGTT

GC content: 25 %.

Can form the following amino acids in 6 translation frames:

N[INSKMRT]

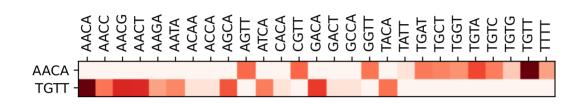
[IESAV*KQGPLRT]T

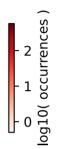
[*EKQ][HQ]

C[SFW*CLY]

[DINSFAVRCGPLYHT]V

[MVL][FL]







TTAA

Extreme GC content: 0 %.

The overhang is palindromic, cannot be used for DNA assembly.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

L[INSKMRT]

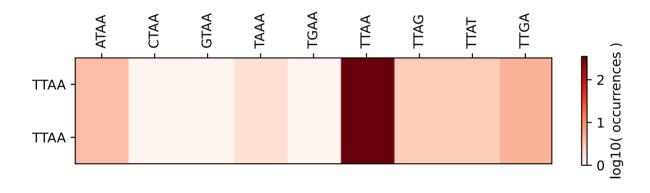
[DINSFAVRCGPLYHT]*

[FVIL][NK]

L[INSKMRT]

[DINSFAVRCGPLYHT]*

[FVIL][NK]





TTAC

GC content: 25 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

V[INSKMRT]

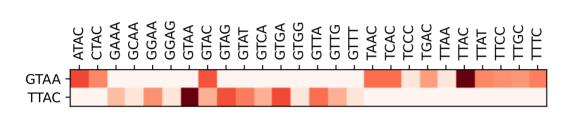
[GESWAV*KQMPLRT]*

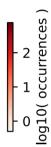
[GRCS][NK]

L[PLRHQ]

[DINSFAVRCGPLYHT]Y

[FVIL][T]







TTAT

Extreme GC content: 0 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

I[INSKMRT]

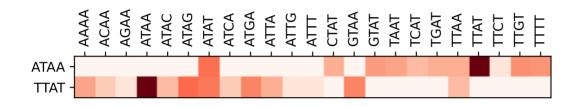
[IESAV*KQGPLRT]*

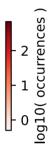
[NDYH][NK]

L[SFW*CLY]

[DINSFAVRCGPLYHT]Y

[FVIL][MI]







TTCA

GC content: 25 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

*[INSKMRT]

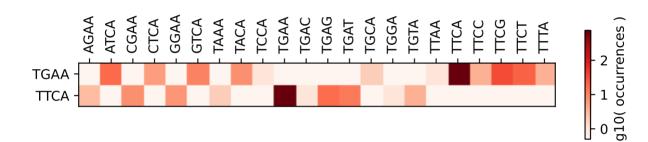
[DINSFAVRCGPLYHT]E

[MVL][NK]

F[INSKMRT]

[DINSFAVRCGPLYHT]S

[FVIL][HQ]





TTCC

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

G[INSKMRT]

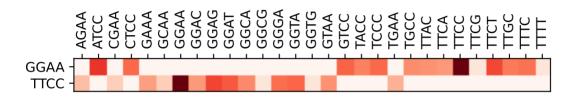
[GESWAV*KQMPLRT]E

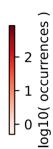
[WGR][NK]

F[PLRHQ]

[DINSFAVRCGPLYHT]S

[FVIL][P]







TTCG

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

R[INSKMRT]

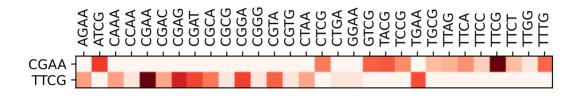
[DINSFAVRCGPLYHT]E

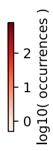
[APST][NK]

F[EAVGD]

[DINSFAVRCGPLYHT]S

[FVIL][R]







TTGA

GC content: 25 %.

The overhang contains a stop codon (TAA, TAG or TGA).

Can form the following amino acids in 6 translation frames:

S[INSKMRT]

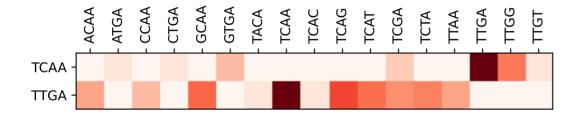
[DINSFAVRCGPLYHT]Q

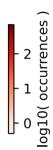
[FVIL][NK]

L[INSKMRT]

[DINSFAVRCGPLYHT]*

[FVIL][ED]







TTGC

GC content: 50 %.

Can form the following amino acids in 6 translation frames:

A[INSKMRT]

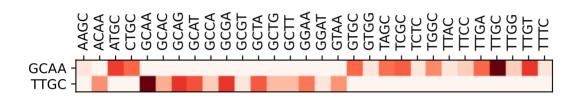
[GESWAV*KQMPLRT]Q

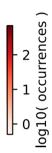
[GRCS][NK]

L[PLRHQ]

[DINSFAVRCGPLYHT]C

[FVIL][A]









Extreme GC content: 0 %.

The overhang contains a stop codon (TAA, TAG or TGA).

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:

*[INSKMRT]

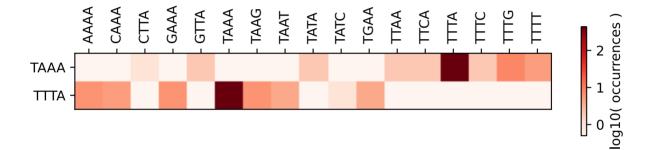
[DINSFAVRCGPLYHT]K

[VIL][NK]

F[INSKMRT]

[DINSFAVRCGPLYHT]L

[FVIL][Y*]







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[INSKMRT]

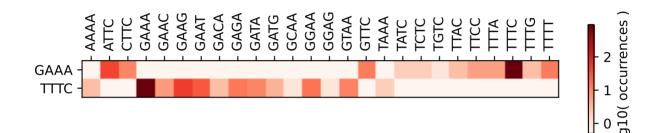
[GESWAV*KQMPLRT]K

[G*R][NK]

F[PLRHQ]

[DINSFAVRCGPLYHT]F

[FVIL][S]







Extreme GC content: 0 %.

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:

K[INSKMRT]

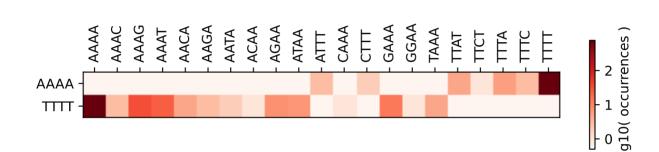
[IESAV*KQGPLRT]K

[*EKQ][NK]

F[SFW*CLY]

[DINSFAVRCGPLYHT]F

[FVIL][FL]



Appendix

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

Summary page(s)

The first page summarises the compendium.

Overhang pages

Each overhang is 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.