o Basepaired columns are not independent.

o Many RNAs exhibit compensatory mutations to maintain basepairing.

o Maximum *extra* info: 2 bit per basepair

$\log_2 (16/4) = 2$ bits

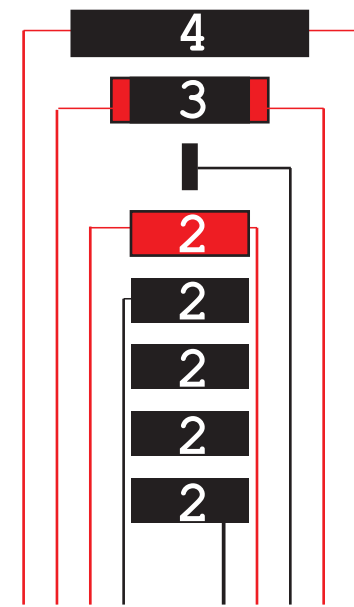o this occurs in red columns, that have 4 of possible 16 bps: GC, CG, AU, UA

sequence + structure profile: 17 bits
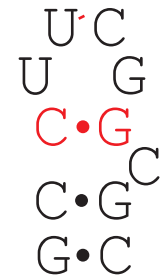
```
        4
      [ 3 ]
        ▮
      [ 2 ]
      [ 2 ]
      [ 2 ]
      [ 2 ]
      [ 2 ]
```

yeast
```
   U C
  U   G
  G•C
  U•A  U
  G•C
```

human
```
   U C
  U   G
  C•G
  C•G  C
  G•C
```

fly
```
   U C
  U   G
  A•U
  C•G  A
  G•C
```

```
struct  <<<<---->->>
yeast   GUGUUCGCUAC
human   GCCUUCGGCGC
fly     GCAUUCGUAGC
fish    GUUUUCGAUAC
orc     GUCUUCGGAAC
frog    GCGUUCGCGGC
worm    GCUUUCGAGGC
mouse   GUAUUCGUCAC
        21_2222__12
```

sequence profile: 14 bits