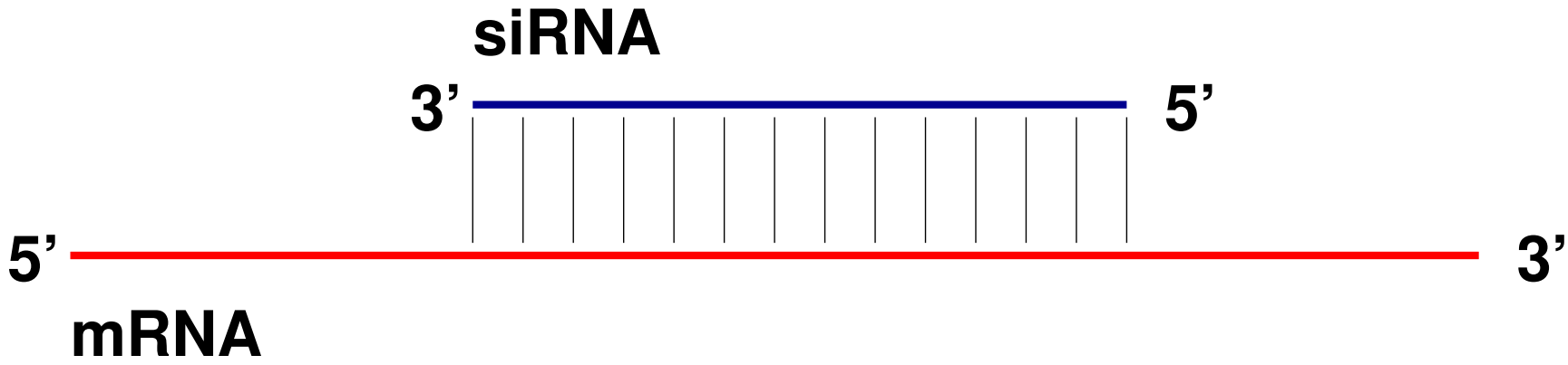
Comparison of siRNAs with high or low efficiency

Ulrike Mückstein

Institute for Theoretical Chemistry
University of Vienna

Bled, 2006

siRNA - mRNA interaction



Dataset: siRecord

- interactions between siRNA and its target mRNA
- 975 interactions with high silencing efficiency
- 551 interactions with low silencing efficiency
- subset: HEK293VECTOR
  celltyp: HEK293; transfection via vector
  - 204 interactions with high silencing efficiency
  - 76 interactions with low silencing efficiency

Information provided by RNAup

- structural context of a binding site
  - $P_u[i,j]$ region $[i,j]$ contains no secondary structure
    the unstructured regions can reside within:
    a hairpin, an interiorloop, a multiloop, the exterior loop
- location of possible binding sites
  - $P^*[i,j]$ probability of a regional interaction
- energetics of RNA-RNA interaction
  - $\Delta G = \Delta G_u + \Delta G_h$

Information provided by RNAup

- structural context of a binding site
  - $P_u[i,j]$ region $[i,j]$ contains no secondary structure
    the unstructured regions can reside within:
    a hairpin, an interiorloop, a multiloop, the exterior loop
- location of possible binding sites
  - $P^*[i,j]$ probability of a regional interaction
- energetics of RNA-RNA interaction
  - $\Delta G = \Delta G_u + \Delta G_h$

Information provided by RNAup

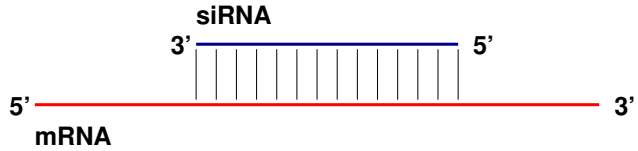
- structural context of a binding site
  - $P_u[i,j]$ region $[i,j]$ contains no secondary structure
    the unstructured regions can reside within:
    a hairpin, an interiorloop, a multiloop, the exterior loop
- location of possible binding sites
  - $P^*[i,j]$ probability of a regional interaction
- energetics of RNA-RNA interaction
  - $\Delta G = \Delta G_u + \Delta G_h$

Differences between siRNA - mRNA interactions wih high or low efficiency?

- Probability of unstructured regions
- Probability of interaction
- Energetics of interaction

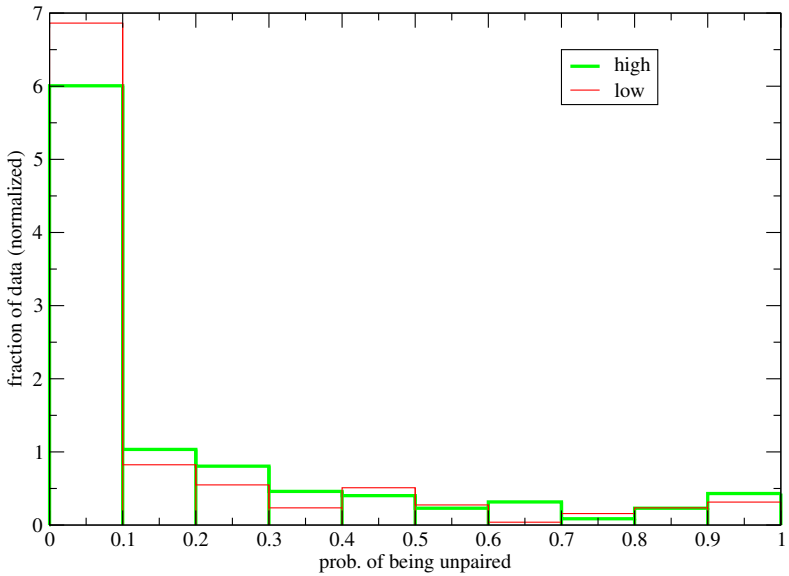
Probability of unstructured regions - target site

5’ end of siRNA, interior loop contributions - position 1,2



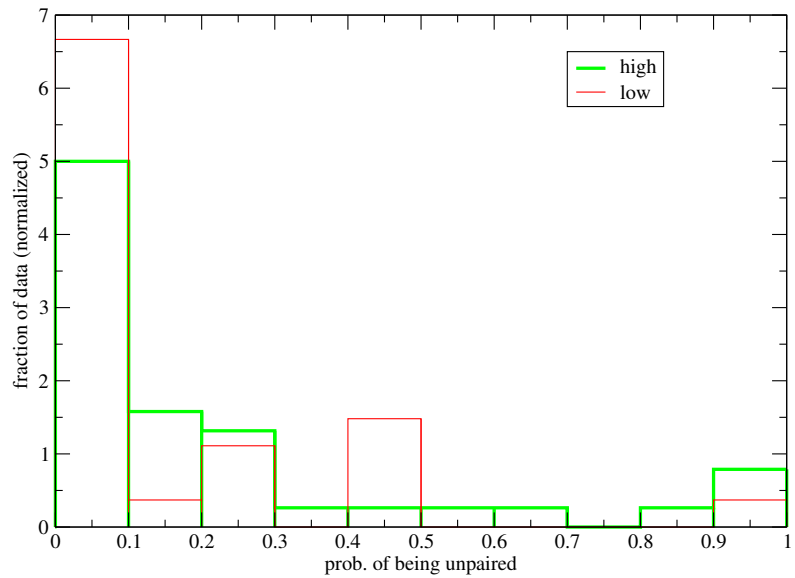
dataset: ALL

HIGH_LOW_reduced_100_1_Ulli_int_rev.dat



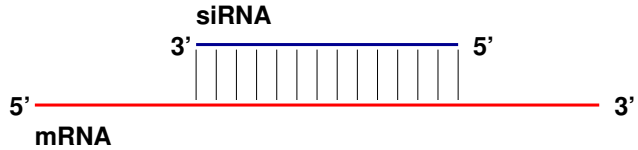
dataset: HEK293VECTOR

HIGH_LOW_reduced_100_1_Ulli_int_rev.dat



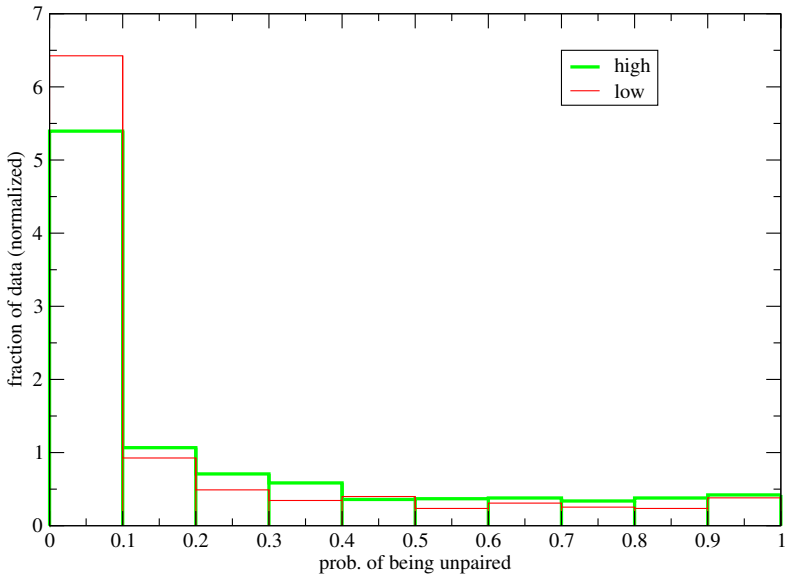
Probability of unstructured regions - off target

3’ end of siRNA, interior loop contributions - pos. [-5,-1]



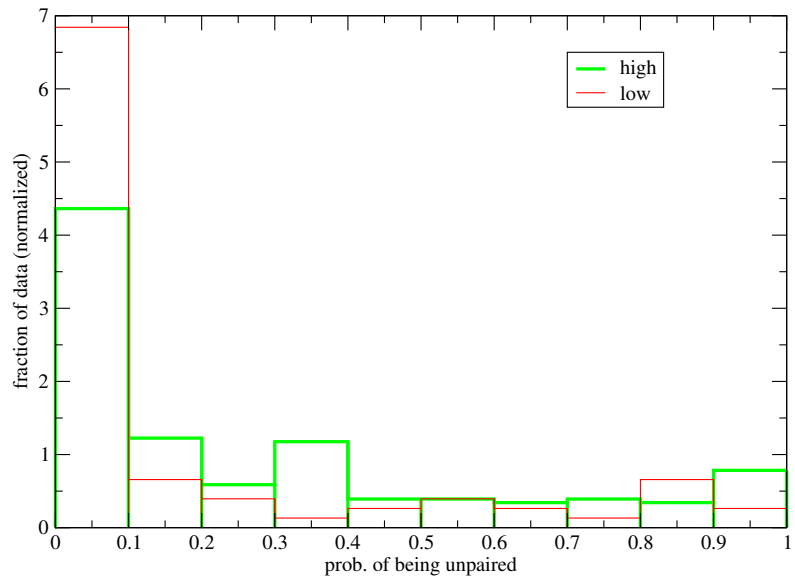
dataset: ALL

HIGH_LOW_reduced_500_198_Hakim_int_200_c.dat



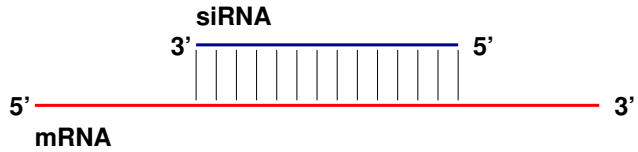
dataset: HEK293VECTOR

HIGH_LOW_reduced_500_198_Hakim_int_200_c.dat



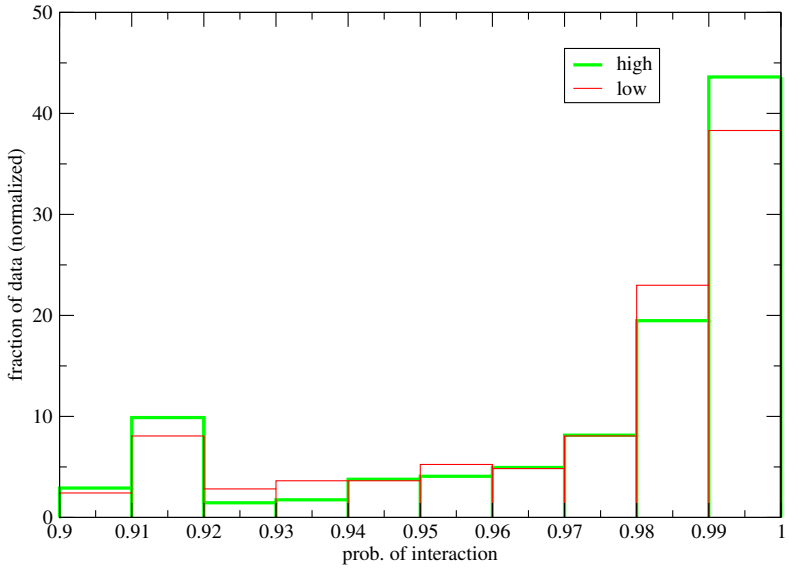
Probability of interaction

5’ end of siRNA, positions 1,2



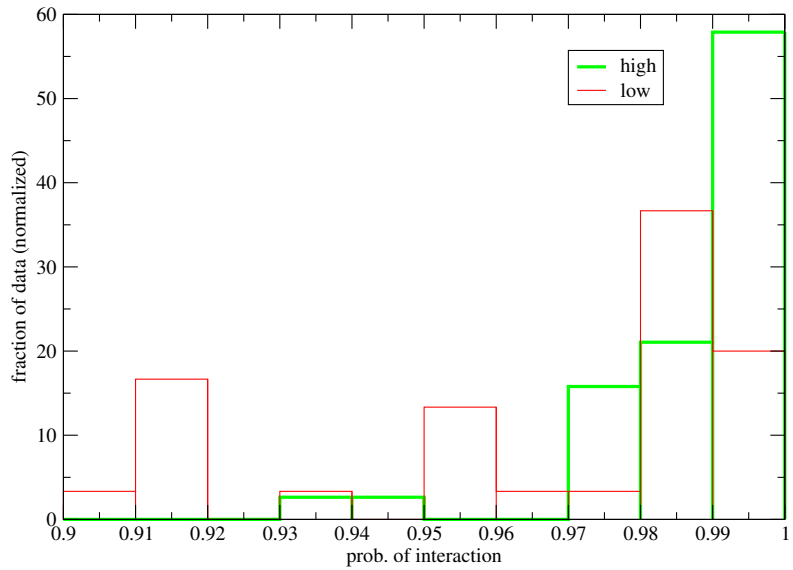
dataset: ALL

HIGH_LOW_reduced_400_1_Ulli_up_rev.dat



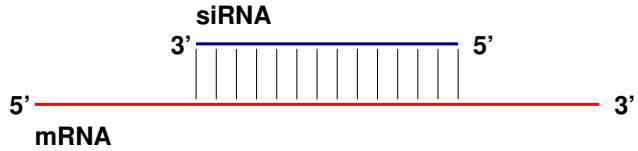
dataset: HEK293VECTOR

HIGH_LOW_reduced_400_1_Ulli_up_rev.dat



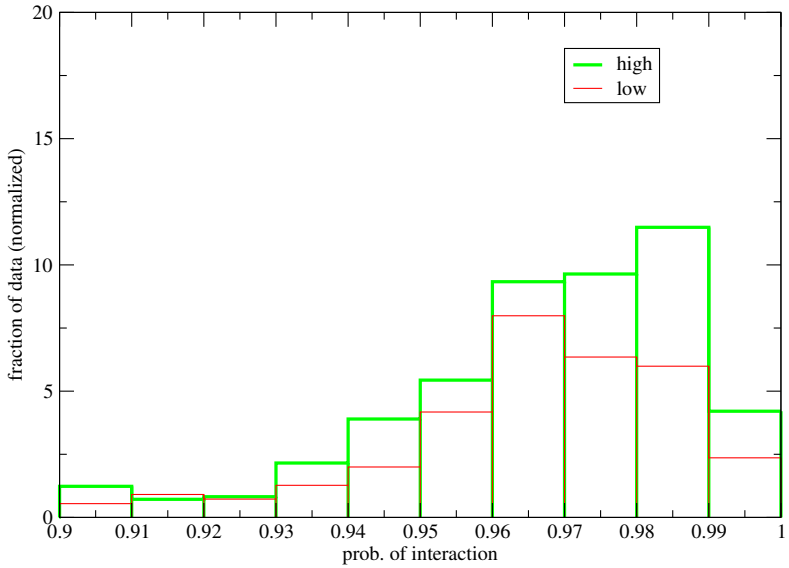
Probability of interaction

3’ end of siRNA, positions 1,2



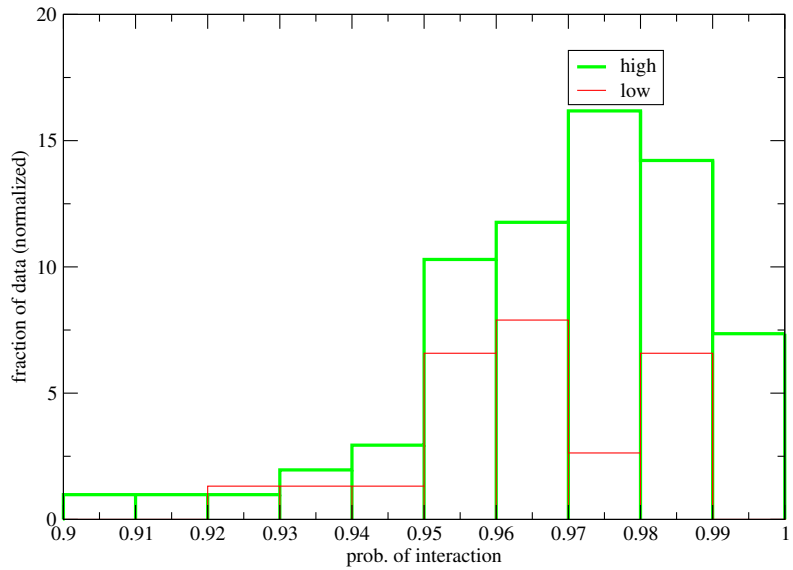
dataset: ALL

HIGH_LOW_reduced_400_1_Hakim_up.dat



dataset: HEK293VECTOR

HIGH_LOW_reduced_400_1_Hakim_up.dat



Ouestions to the audience

- differences between the structural context and the probability of interaction
- (how) can we use these differences to discriminate between siRNAs with high or low efficiency?
- support vector machines?
- statictical methods? sequence patterns?
- more data?

Thanks to

the audience
Hakim Tafer
Stephan Bernhart
Stefan Washietl