Supplementary figures for

Dissecting the target specificity of RNase H recruiting oligonucleotides using massively parallel reporter analysis of short RNA motifs.

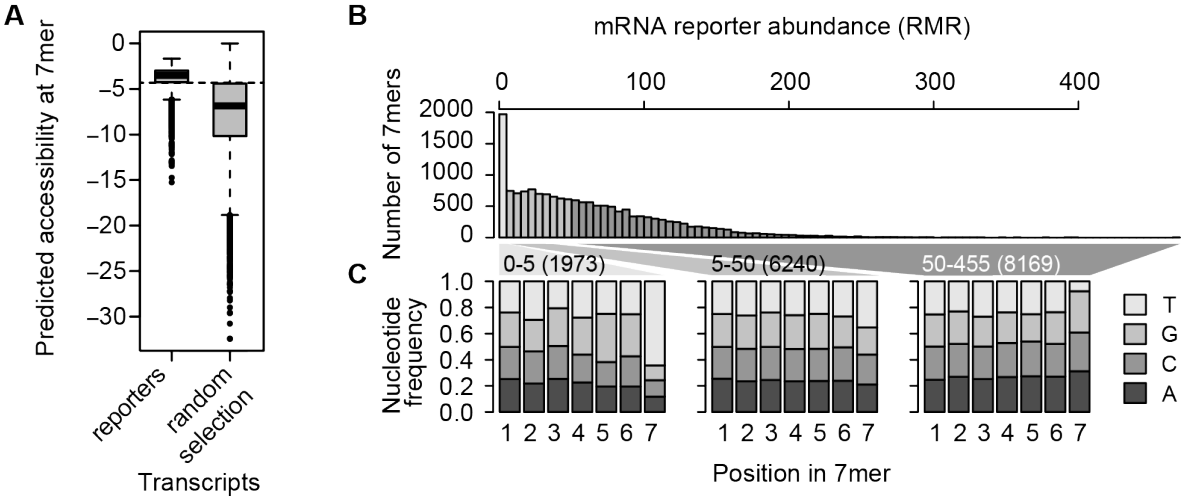
Jakob Lewin Rukov[1], Peter H Hagedorn[2], Isabel Bro Høy[1], Yanping Feng[1,3], Morten Lindow[2], Jeppe Vinther[1,*]

[1]Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200 Copenhagen N, Denmark

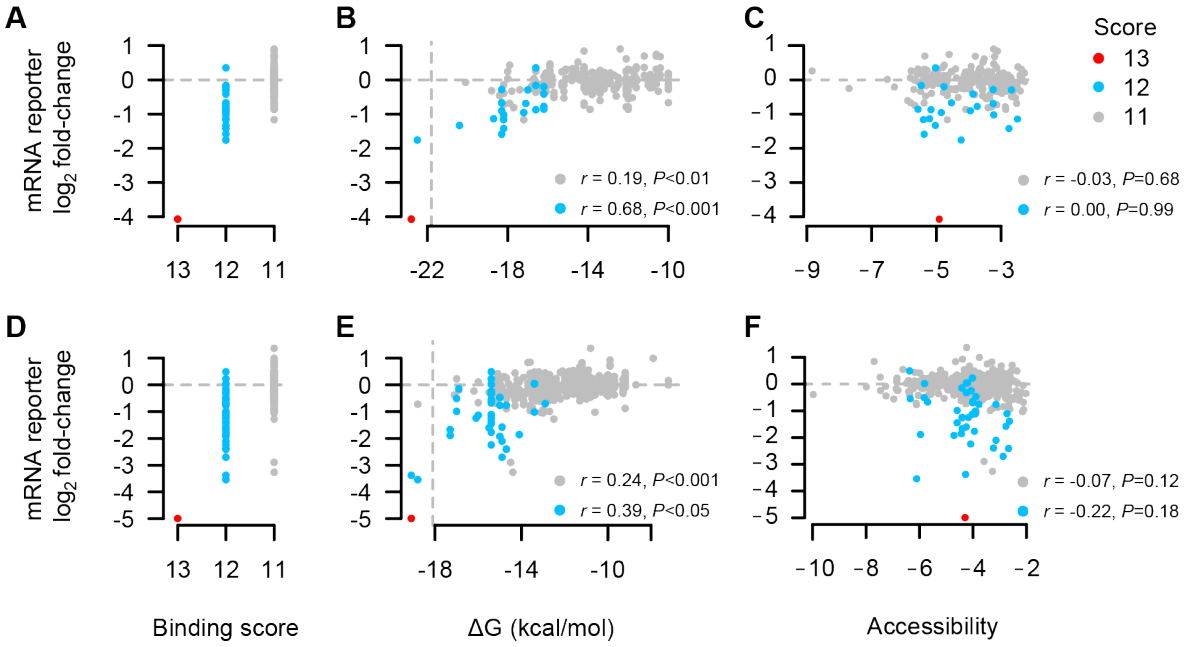
[2]Roche Pharmaceutical Discovery and Early Development, Roche Innovation Center Copenhagen A/S, Fremtidsvej 3, DK-2970 Hørsholm, Denmark

[3]College of Animal Sciences and Technology, Huazhong Agricultural University, Wuhan, Hubei Province, 430070, P.R.China

* To whom correspondence should be addressed. Tel: +4535321264; Fax: +4535322128; Email: jvinther@bio.ku.dk



Supplementary Figure 1. Predicted accessibility, distribution and position-specific nucleotide frequencies for the 7mers. (**A**) Distribution of predicted accessibilities of a 13nt long oligonucleotide at the position of the 7mer degenerate subsequence in each of the 16382 different reporter mRNAs. For comparison, random regions of the same size as the reporter mRNA was chosen among 16382 different randomly selected mRNAs, and the accessibly calculated (see supplementary methods) for the central 13 nts corresponding to the same position as in the reporter mRNA. (**B**) Distribution of 7mer abundances (RMR) (**C**) The nucleotide frequencies at each of the seven positions in the 7mer, for low- (0 to 5 RMR), medium- (5 to 50 RMR), and high-abundance (50 to 455 RMR) 7mers.



Supplementary Figure 2. Association between changes in mRNA reporter abundance and binding affinity or predicted accessibility. For oligonucleotide A, correlation between mRNA reporter fold-change and (**A**) binding score, (**B**) free energy of binding, or (**C**) predicted accessibility. Similarly for oligonucleotide B, correlation between mRNA reporter fold-change and (**D**) binding score, (**E**) free energy of binding, or (**F**) predicted accessibility. The correlation is quantified by the Pearson's correlation coefficient, *r*, its significance *P*.

Supplementary Methods

Accessibility was calculated in a sliding window approach across the mRNA reporter (window size 80nt), where base pairing within the window was restricted to a maximum distance of 40nt, as implemented in RNAplfold algorithm (Tafer et al., 2008). The accessibility of the 13nt region centered on the 7nt degenerate subsequence is the log-base-2 summed up probability of being single-stranded across all windows.

Supplementary reference

Tafer H, Ameres SL, Obernosterer G, Gebeshuber CA, Schroeder R, Martinez J, Hofacker IL (2008) The impact of target site accessibility on the design of effective siRNAs. *Nat Biotechnol* **26**(5):578-83