Supporting Information

Estimation of drag-tag lengths for direct quantitative analysis of multiple miRNAs

Leonid T. Cherney and Sergey N. Krylov

1. Inapplicability of ELFSE to short probe-miRNA hybrids

The theory of ELFSE was developed for significantly long oligonucleotides that presumably behave as a semiflexible random coil.[1-6] The latter is true if the polymer contour length $L = Nb$ is much greater than the Kuhn length b_K of the polymer. Here, N is the number of monomers and b is the monomer length. The b_K value is often related to persistence length p ($p = b_K/2$ for semiflexible polymers described by the worm-like chain model), which is a measure of the polymer's stiffness.[7,8] We approximately have $b = 0.34$ nm and $b_K = 100$ nm for double stranded DNA (dsDNA).[5] The single stranded DNA (ssDNA) is much more flexible and is described by values $b = 0.43$ nm and $b_K = 6$ nm.[5] In particular, calculations of the mobility of ssDNA with an attached drag tag[5] is based on the theory of polyampholyte electrophoresis that itself was developed for long polymer chains adopting a Gaussian conformation.[9] Since miRNAs contain a small number of nucleotides (~ 18-26) we have to understand if the ELFSE theory can be applied to migration of DNA-miRNA duplexes with attached drag tags. Structural studies of DNA-RNA hybrids show that they form hybrid helixes.[10] Their conformation is intermediate between A- and B-forms of dsDNA but more resembles the A-form.[11-14] The persistence length of dsDNA can be estimated as 45-50 nm[15], which is in agreement with the Kuhn length ~ 100 nm.[5] Double stranded RNA (dsRNA) is stiffer than dsDNA and the persistence length of dsRNA helix is larger (65-80 nm) than that of dsDNA.[14,16] One could expect that **a** DNA-RNA hybrid has intermediate stiffness and its persistence length lies in a range of 50-75 nm. There are only a few experimental studies on the persistence length of DNA-RNA hybrids. They indicate that persistence lengths **can be** as low as 20 nm in some cases.[17,18] However, even such surprisingly low values are still higher than the contour length $L_{hyb} < 9$ nm of short hybrids containing 18-26 nucleotides that we deal with in this work. As a result, the assumption made in ELFSE that long oligonucleotides behave as semiflexible random coils, is not applicable to short DNA-RNA complexes. Rather, these short hybrids seem to behave like ridged rods.

2. Flexibility of short peptide drag tags

The Kuhn length of short peptides, $b_{K,tag}$, is of the order of 1 nm magnitude and the persistence length is just ~ 0.5 nm.[19-22] On the other hand, contour lengths of peptides used for example in **Table 1** (see main text) are 1.8, 3.6, 5.4, and 7.2 nm for drag tags containing 5, 10, 15, and 20 residues, respectively. These values are several times higher than the persistence length of peptides. Thus, the latter should take a more compact conformation that can be properly described by the gyration radius $R_{G,tag}$. In the case of long polymers with the number of

monomers $N_{tag} > N^*_{tag}$ calculations of $R_{G,tag}$ require taking into account the excluded volume. The critical value N^*_{tag} is determined by relation[1,5]

$$N^*_{tag} = \frac{b^3_{K,tag}}{b_{tag} d^2_{tag}} \quad \text{(S1)}$$

where d_{tag} is the diameter of the peptide, $b_{tag} = 0.36$ nm is the crystallographic length per residue. This value is slightly smaller than the dynamic value of $b_{tag} = 0.4$ nm that accounts for the internal degrees of freedom of the peptide.[23] For peptides formed by glycine, alanine, and threonine we have $d_{tag} < 0.24$ nm and, therefore, $N_{tag}^* > 48$ that is larger than the number of residues in the drag tags used in this study. Thus we can neglect the effect of excluded volume on interactions in peptides. In this case, the gyration radius is determined by the Kratky-Porod equation[7,8]

$$R^2_{G,tag} = \frac{b_{K,tag} L_{tag}}{6}\left[1 - 3\left(\frac{b_{K,tag}}{2L_{tag}}\right) + 6\left(\frac{b_{K,tag}}{2L_{tag}}\right)^2 - 6\left(\frac{b_{K,tag}}{2L_{tag}}\right)^3\left(1 - \exp\left(-\frac{2L_{tag}}{b_{K,tag}}\right)\right)\right] \quad \text{(S2)}$$

where $L_{tag} = N_{tag}\, b_{tag}$ is the contour length of the peptide. For example, the gyration radii of drag tags containing 5, 10, 15, and 20 residues are 0.39, 0.64, 0.83, and 0.99 nm, respectively. These values are significantly smaller than the length of the ssDNA-miRNA hybrids and we can assume that the drag tag forms a worm-like chain located at the end of the ssDNA-miRNA hybrid.

The hydrodynamic radius of the drag tag, $R_{H,tag}$, can be related to the gyration radius of polymers by equation (10) (hereafter equation numbers without S refer to the main text). On the other hand, studies of unfolded proteins result in the following dependence for the hydrodynamic radius (in nm):[24]

$$R_{H,tag} = (0.22 \pm 0.11) N_{tag}^{0.57 \pm 0.02} \quad \text{(S3)}$$

that was obtained by fitting experimental data. Here, N_{tag} is the number of residues in the peptide chain. Expression (S3) gives approximately twice higher values for $R_{H,tag}$ (0.55, 0.82, 1.03, and 1.21 nm for drag tags containing 5, 10, 15, and 20 residues, respectively) than those found from (10). However, the order of magnitude of $R_{H,tag}$ remains the same.

3. Mobility of unreacted probes and probes bound to SSB

For ssDNA b_K = 6 nm[5] and is comparable to the length of ssDNA section (8 – 11 nm) in the probe. Thus, relation (3) can be used to estimate the mobility of ssDNA without a drag tag. In this case, a value of d_{hyb} present in (3) should be replaced with the diameter of ssDNA, d_{ssDNA}. The latter can be estimated as $d_{hyb}/2$[25,26] or can be considered as an adjustable parameter. The probe mobility can then be evaluated using the second relation (12) if μ_{hyb}, L_{hyb}, and d_{hyb} are replaced with μ_{ssDNA}, L_{ssDNA} and d_{ssDNA}, respectively. After such modification relation (12) will depend on the numbers of monomers in ssDNA, N_{DNA}, and in the tag, N_{tag}, through the ratio $R_{G,tag}(N_{tag})/N_{DNA}$ (since $L_{ssDNA} \sim N_{DNA}$), where function $R_{G,tag}(N_{tag})$ is determined by the Kratky-Porod equation (S2). The averaging procedure employed in ELFSE leads to a different dependence of the probe mobility on the ratio N_{tag}/N_{DNA}:

$$\mu_{ssDNA+tag} = \mu_{ssDNA}\left(1+\alpha\frac{N_{tag}}{N_{DNA}}\right)^{-1}, \quad \alpha = \frac{b_{tag}b_{K,tag}}{b_{ssDNA}b_{K,ssDNA}} \tag{S4}$$

Here, b_{ssDNA} and $b_{K,ssDNA}$ are the monomer size and the Kuhn length of ssDNA, respectively. Though the second relation (12) (modified for probes, i.e. for ssDNA with a tag) gives different dependencies of $\mu_{ssDNA+tag}/\mu_{ssDNA}$ on N_{DNA} and N_{tag} than the first relation (S4) does, these relations lead to close values in our ranges of N_{DNA} and N_{tag}. For example, at N_{DNA} = 20 and N_{tag} = 20 we obtain $\mu_{ssDNA+tag}/\mu_{ssDNA}$ = 0.72 and 0.87, respectively, from relations (12) (modified for probes) and (S4). Since $d_{ssDNA} < d_{hyb}$, we always have $\mu_{ssDNA} > \mu_{hyb}$. The mobility of the probe (i.e. $\mu_{ssDNA+tag}$) is also larger than that of the tagged hybrid.

Finally, let us consider the mobility of the probe bound to SSB, $\mu_{probe+SSB}$. SSB is a globular protein consisting of 177 residues and having the characteristic size ~ 10 nm. ssDNA bound to a SSB surface can be assumed to have a length <10 nm and a diameter ~1 nm. As a result, we can consider a complex of ssDNA-SSB as a globular object with approximately the same diameter ($d_{comp} \sim$ 10 nm), as the native protein. In this case, its mobility, $\mu_{ssDNA+SSB}$ can be estimated by relations similar to equations (3)

$$\mu_{ssDNA+SSB} = \frac{\sigma_{ssDNA+SSB}\lambda_D}{\eta} = \frac{Q_{ssDNA+SSB}\lambda_D}{\pi\eta d_{SSB}^2} \quad \left(\sigma_{ssDNA+SSB} = \frac{Q_{ssDNA+SSB}}{\pi d_{SSB}^2}\right) \tag{S5}$$

Where $\sigma_{ssDNA+SSB}$ is the surface density of the electric charge in the diffuse part of the double layer around the complex of SSB and ssDNA, $Q_{ssDNA+SSB}$ is the total charge of this complex, and we neglected the Stern layer around the complex. Given the relatively small number of residues in the drag tag we can assume that $\mu_{probe+SSB} \approx \mu_{ssDNA+SSB}$. For much longer drag tags a relation for $\mu_{probe+SSB}$ can be derived similarly to relation (12) for $\mu_{hyb+tag}$. Indeed, an expression (S5) can be derived from the balance of electric and hydrodynamic forces, $F_{E,ssDNA+SSB}$ and $F_{H,ssDNA+SSB}$,

acting upon the ssDNA-SSB complex if we assume the following expressions for them (similarly to (5)):

$$F_{\mathrm{E,ssDNA+SSB}} = Q_{\mathrm{ssDNA+SSB}} E, \quad F_{\mathrm{H,ssDNA+SSB}} = \frac{\pi \eta d_{\mathrm{SSB}}^2}{\lambda_{\mathrm{D}}} u \tag{S6}$$

Then the mobility $\mu_{\mathrm{probe+SSB}}$ of the probe bound to SSB and having the long drag tag can be found from the balance of all forces acting upon such complex:

$$F_{\mathrm{E,ssDNA+SSB}} = F_{\mathrm{H,ssDNA+SSB}} + F_{\mathrm{H,tag}} \tag{S7}$$

Substituting expressions (S6) into equation (S7) and taking into account relation (9) we obtain

$$\mu_{\mathrm{probe+SSB}} = \frac{\mu_{\mathrm{ssDNA+SSB}}}{1 + \dfrac{6\lambda_{\mathrm{D}} R_{\mathrm{H,tag}}}{d_{\mathrm{SSB}}^2}} \tag{S8}$$

Here, the hydrodynamic radius of the drag tag, $R_{\mathrm{H,tag}}$, is determined by relations (10) and (S2). For short drag tags with $6\lambda_{\mathrm{D}} R_{\mathrm{H,tag}} << d^2_{\mathrm{SSB}}$ expression (S8) is reduced to $\mu_{\mathrm{probe+SSB}} \approx \mu_{\mathrm{ssDNA+SSB}}$.

References

1. Viovy, J.-L. *Rev. Modern Phys.* **2000**, *72*, 813-872.
2. Mayer, P.; Slater, G. W.; Drouin, G. *Anal. Chem.* **1994**, *66*, 1777-1780.
3. Hubert, S. J.; Slater, G. W. *Electrophoresis* **1995**, *16*, 2137-2142.
4. Desruisseaux, C.; Long, D.; Drouin, G.; Slater, G. W. *Macromolecules* **2001,** *34,* 44-52.
5. Meagher, R. J.; Won, J. - I.; McCormick, L., C.; Nedelcu, S.; Bertrand, M. M.; Bertram, J. L.; Drouin, G.; Barron, A., E.; Slater, G. W. *Electrophoresis* **2005**, *26*, 331-350.
6. Meagher, R. J.; McCormick, L., C.; Haynes, R. D.; Won, J. - I.; Lin, J. S.; Slater, G. W.; Barron, A. E. *Electrophoresis* **2006**, *27*, 1702-1712.
7. Teraoka, I. *Polymer Solutions* John Wiley & Sons, New York **2002**, 360 pp.
8. Strobl, G. R. *The physics of polymers: concepts for understanding their structures and behavior*, Springer, Berlin, **2007**, 518 pp.
9. Long, D.; Dobrynin, A. V.; Rubinstein, M.; Ajdari, A. *J. Chem. Phys.* **1998**, *108*, 1234-1244.
10. Rich, A. *J. Biol. Chem.* **2006**, *281*, 7693-7696.

11. Milman, G.; Langridge, R.; Chamberlin, M. J.; *Proc. Natl. Acad. Sci. U.S.A* **1967**, *57*, 1804-1810.
12. Lesnik, E. A.; Freier, S. M. *Biochemistry* **1995**, *34*, 10807-10815.
13. Fedoroff, O. Yu.; Salazar, M.; Reid, B. R. *J. Mol. Biol.* **1993**, *233*, 509-523.
14. Noy, A.; Pérez, A.; Márquez, M.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2005**, *127*, 4910-4920.
15. Cheatham, T. E., III; Kollman, P. A. *J. Mol. Biol.* **1966**, *259*, 434-444.
16. Kebbekus, P.; Draper, D. E.; Hagermad, P. *Biochemistry* **1995**, *34*, 4354-4357.
17. Wen, J.-D.; Manosas, M.; Li, P. T. X.; Smith, S. B.; Bustamante, C.; Ritort, F.; Tinoco, I., Jr. *Biophys. J.* **2007**, *92*, 2996-3009.
18. Collin, D.; Ritort, F., Jarzynski, C.; Smith, S. B.; Tinoco, I., Jr.; Bustamante, C. *Nature* **2005**, *437*, 231-234.
19. Hanke, F.; Serr, A.; Kreuzer, H. J.; Netz, R. R. *EPLA*, **2010**, *92*, 53001_p1-53001_p6.
20. Zhang, Y.; and Skolnick, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 1029-1034.
21. Rawat, N.; and Biswas, P. *J. Chem. Phys*. 2009, 131, 165104_1-165104_9.
22. Gräter, F.; Heider, P.; Zangi, R.; Berne, B. J. *J. Am. Chem. Soc*. **2008**, *130*, 11578–11579.
23. Ainavarapu, S. R. K.; Brujic, J.; Huang, H. H.; Wiita, A. P.; Lu, H.; Li, L.; Walther, K. A.; Carrion-Vazquez, M.; Li, H.; Fernandez, J. M. *Biophys. J.* **2007**, *92*, 225–233.
24. Wilkins, D. K.; Grimshaw, S. B.; Receveur, V.; Dobson, C. M.; Jones, J. A.; and Smith, L. J. *Biochemistry* **1999,** *38,* 16424-16431.
25. Simmel, F. C.; Dittmer, W. U. *Small* **2005**, *1*, 284-299.
26. Krishnan, Y.; Simmel, F. C. *Angew. Chem. Int. Ed.* **2011**, *50*, 3124–3156.