

Observed versus predicted structure of fluorescent self-quenching reporter molecules (SQRM): Caveats with respect to the use of “stem–loop” oligonucleotides as probes for mRNA folding

VIKRAM PATTANAYAK,^{1,3} LIDA K. GIFFORD,^{2,3} PONZY LU,² and ALAN M. GEWIRTZ¹

¹Department of Medicine, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA

²Department of Chemistry, School of Arts and Sciences, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

ABSTRACT

We developed self-quenching reporter molecules (SQRM), oligodeoxynucleotides with fluorophore and quencher moieties at the 5' and 3' ends respectively, to probe mRNAs for single-stranded, hybridization accessible sequences. SQRM and their homologous antecedents, Molecular Beacons (MB), are designed with the assumption that they adopt a stem–loop structure thought critical for regulating their reporter function. Recently, we observed that stem–loop structures are not required for SQRM function, and on this basis proposed that these reporter molecules be classified according to whether they were stemmed (Type I) or not (Type II). This finding further stimulated us to investigate whether Type I SQRM, and by extension MBs, actually adopt a stem–loop configuration under physiologic conditions. Accordingly, we synthesized Type I and Type II SQRM and studied the thermodynamic characteristics of each by fluorescence melting analysis. The results of these studies suggested that the majority of stem–loop Type I SQRM are unstructured at 37°C, while some of the stemless Type II SQRM are, surprisingly, structured. These results were not predicted by the mfold computer program. Type I and II SQRM were then employed to “map” the mRNA secondary structure of a gene encoding a tyrosine kinase receptor, *c-kit*. Neither experimentally determined melting temperatures nor mfold-“predicted” thermodynamic parameters were useful in predicting the fluorescence signal-to-noise ratios obtained for SQRM incubated with *c-kit* mRNA. We conclude that stem–loop reporter molecules are in fact unlikely to adopt their presumed structures at 37°C, and this design consideration may be dispensed when their use is contemplated under physiologic conditions.

Keywords: fluorescent probes; oligodeoxynucleotides; structure; RNA silencing

INTRODUCTION

Antisense nucleic acids in the form of oligodeoxynucleotides, ribozymes, DNAzymes, siRNA, shRNA, and microRNA can be used to silence the expression of specific genes by eliciting the destruction or blocking the use of their messenger RNA (mRNA) transcripts (for review, see Scherer and Rossi 2003). This approach to manipulating gene expression is now widely used to understand gene function, and it is hoped that gene silencing will make equally important contributions to the treatment of dis-

eases in which overexpression of known genes is thought to play a role (for review, see Mahato et al. 2005). In the course of applying various gene silencing methodologies in living cells, it has become well known that not all sequences within an mRNA are equally accessible for hybridization. In hopes of making silencing more predictable we developed self-quenching reporter molecules (SQRM), oligodeoxynucleotides with fluorophore and quencher moieties at the 5' and 3' ends respectively, to help identify hybridization accessible sequence within a given mRNA (Gifford et al. 2005b). As a result, we have hypothesized that these molecules might enable a rational approach to target site selection, and in so doing facilitate the design of more effective DNA, and RNA, antisense oligonucleotide probes and drugs (Gifford et al. 2005a,b)

The design of Type I SQRM and their antecedents, the fluorescent reporter molecules known as molecular beacons (MB) (Tyagi and Kramer 1996), is based on the solution

³These authors contributed equally to this work.

Reprint requests to: Alan M. Gewirtz, Department of Medicine, 421 Curie Boulevard, Room 713, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA; e-mail: gewirtz@mail.med.upenn.edu; fax: (215) 573-2078.

Article and publication are at <http://www.majournal.org/cgi/doi/10.1261/rna.890408>.

structures they are assumed to adopt in vivo. These structures are best characterized as simple stem-loops whose formation is driven by Watson-Crick base-pairing of complementary ends (Fig. 1). SQRMs are distinguished from MB by the fact that the entire SQRm sequence, including the stem region, is completely complementary to the target mRNA sequence while only the loop of MB need be complementary. The stem was originally thought important for holding the 5' fluorophore in close proximity to the quencher molecule linked to the 3' end. In this configuration, the fluorophore is subject to fluorescence resonance energy transfer (FRET) or collisional quenching (Kuhn et al. 2002) and, as a result, cannot fluoresce. Previous studies have shown that linear probes can be used to signal accessible sites (Parkhurst and Parkhurst 1995a,b) and that complementary ends are not required for the functionality of these molecules (Gifford et al. 2005a).

The finding that a stem is not required for reporter function of these molecules caused us to question our assumption that Type I SQRm, and by extension MBs, actually adopt a stem-loop configuration. To examine this question, we synthesized Type I and Type II SQRm which were designed to target the *c-kit* protooncogene. KIT is a tyrosine kinase receptor that is involved in hematopoietic cell maintenance and proliferation. Mutations in the *c-kit* gene have been associated with many different types of malignancies, including leukemia and gastrointestinal stromal cell tumors (for review, see Ashman 1999; Sattler and Salgia 2004), and overexpression of *c-kit* is thought to play a role in some cancers, making it an attractive therapeutic target. We studied the thermodynamic characteristics of each SQRm employed in these studies by fluorescence melting analysis. Our results suggest that the majority of stem-loop DNA SQRms are in fact completely or partially unstructured at 37°C, and that this design consideration

could likely be dispensed with when these molecules are being employed as reporters in physiologic systems.

RESULTS

Mapping the 1.2-kb *c-kit* mRNA fragment

Several methods were employed to design probes for mapping the *c-kit* RNA with Type II SQRms. The first sequence targeted, 420 ← 446, was chosen based on predictions by the mfold computer program (Table 1; Zuker 2003). The entire sequence of 1.2-kb *c-kit* mRNA was entered into mfold as a coarse selection method for a part of the sequence that is most likely to be single stranded. The program generated 22 possible secondary structures that were determined based on free energy calculations for an unalterable buffer solution of 1 M NaCl at 37°C. Mfold yielded a single-stranded count containing the number of structures in which each particular base in the *c-kit* fragment sequence could be found in a loop region. This single-stranded count was converted into a single-stranded propensity, which is the percentage of the 22 theoretical structures containing a particular base in a loop region. The single-stranded propensity data were sorted to identify sites at least 18 bases long in which at least 13 out of the 18 bases had a single-stranded propensity of more than 50%. Thirteen sites were identified, and the longest sequence with the highest single-stranded propensity was chosen for further study.

Three additional SQRms were designed to avoid the targeting of a potential stem-loop structure in the RNA with a Type I SQRm (Table 1; Gifford et al. 2005a). The first site in the RNA, 802 → 832, is shifted 5 bases in the 5' direction from 807 → 837; the second, 812 → 843, is shifted 5 bases in the 3' direction from 802 → 837. The

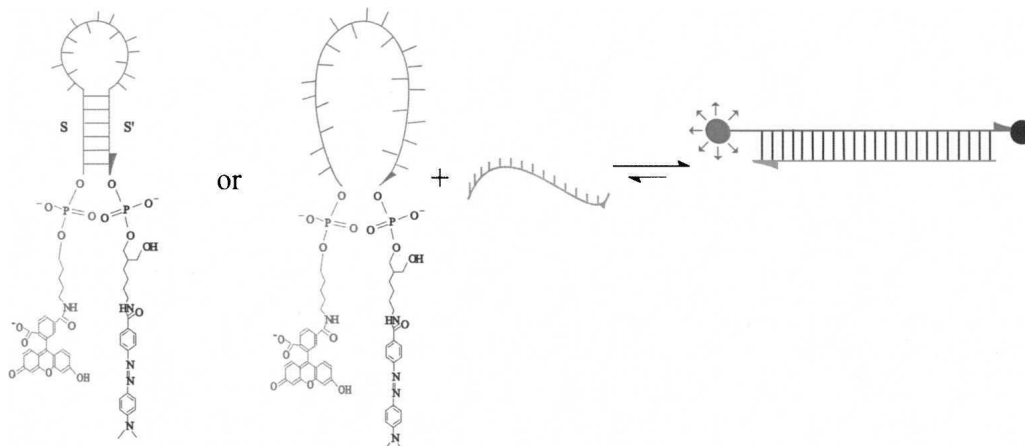


FIGURE 1. Type I (left) or Type II (right) SQRMs recognize target RNA and hybridize fully.

TABLE 1. Type I and II *c-kit*-targeted SQRM sequences

Name	Stem-loop length	SQRM sequence (5' → 3')
Type I SQRM 39 ← 64	4/18	F- TCTGGACGCGAAGCAGTAGGAGCAGA -D
Type I SQRM 137 ← 163	4/19	F- TAATCTCGTCGCCACGCGACTATTA -D
Type I SQRM 287 ← 304	4/10	F- CGTGTGTTGGTGCACG -D
Type I SQRM 351 ← 381	5/21	F- TTCCCATACAAGGAGCGGTCAACAAGGAAA -D
Type II SQRM 420 ← 446	0/27	F-TTGAGGAATAATTGGTCACTTCTGGG-D
Type I SQRM 580 ← 607	4/20	F- TCAGGATGAATTTTCCGACAGCACTGA -D
Type I SQRM 663 ← 683	5/11	F- TCTTCCCCTCCCTAAGAAGA -D
Type II SQRM 802 ← 832	0/31	F-TGATAGTCAACGTTGCCTGACGTTTATAATT-D
Type I SQRM 807 ← 837	5/21	F- TGAACTGATAGTCAACGTTGCCTGACGTTCA -D
Type II SQRM 812 ← 842	0/31	F-CTCGCTGAACTGATAGTCAACGTTGCCTGAC-D
Type II SQRM 807 ← 832 (loop)	0/21	F-TGATAGTCAACGTTGCCTGAC-D
Type I SQRM 924 ← 946	5/13	F- AGATATTAATGAATCCTTTATCT -D
Type I SQRM 962 ← 994	5/23	F- CTACATTTCTCCATCGTTTACAAACTGTAG -D
Type I SQRM 1093 ← 1116	5/14	F- TCTGATATTACTTTCATTCTCAGA -D

Stem regions are in boldface. F, Fluorescein; D, DABCYL.

third site is targeted using SQRM 807 loop, a 21-base SQRM that actually hybridizes from bases 812 to 832 and contains only the “loop” sequence between the complementary regions of site 807 → 837.

Determination of SQRM fluorescent melting temperatures

A characteristic thermal denaturation curve for Type I SQRM 807 ← 832 is shown in Figure 2A. There is a linear dependence of fluorescence on temperature for a probe labeled with fluorescein in the absence of a quencher. The fluorescence of each SQRM was corrected with a linear correction factor derived from the high temperature fluorescence values for each SQRM. The data in Figure 2A show that the fluorescence of a SQRM in the absence of target is low at low temperatures, when the SQRM is closed, and high at higher temperatures, when the SQRM is open. A van't Hoff plot for this SQRM is shown in Figure 2B. The slope of the fitted curve corresponds to $-\Delta H_{2 \rightarrow 3}^{\circ}$ and the intercept corresponds to $\Delta S_{2 \rightarrow 3}^{\circ}$. Table 2 is a summary of the melting temperatures and thermodynamic parameters of *c-kit* targeted SQRM in the absence of ODN target.

SQRM structure predictions

The Type I SQRM used in this experiment possess sequences capable of forming stem-loop structures. In contrast to molecular beacons, and shared-stem molecular beacons, which are postulated to have the same ability to form stem-loops, the SQRM stem sequences are fully complementary to the target. Since SQRM sequences directly follow from the sequence of the target gene, desired features such as high GC content in the stem are difficult to design. To predict whether stem-loop structures are likely

to form, the oligonucleotide sequences were entered into mfold. Three different mfold predictions for an oligonucleotide corresponding to Type I SQRM 807 ← 837 are shown in Figure 3. The stem-loop hairpin structure (Fig. 3A) that the SQRM was intended to form is an unstable structure. This structure, which had to be modeled by forcing the 5 nucleotides (nt) on each end to base pair and prohibiting any of the “loop” nucleotides from base-pairing, has a melting temperature of 36.1°C and a stability of 0.1 kcal/mol at the mapping temperature. If this were the only possible structure, more than half of the SQRM in solution would be unstructured, while the rest would exist as hairpins. If the mfold program is asked to calculate a structure that contains a 5-base paired stem, it predicts a fold with 9 base pairs (bp) and a bulge (Fig. 3B). The most favorable structure predicted is not a stem-loop. Instead, it is an asymmetrical hairpin (Fig. 3C). The most favorable structure has a melting temperature of 58.6°C and a stability of 2.4 kcal/mol at 37°C, indicating that it is very stable at the mapping temperature. The predominately double-stranded stemmed structure shown in Figure 3 has a melting temperature of 38.6°C and a stability of 0.3 kcal/mol at the mapping temperature. Even if this were the most favorable SQRM structure, its ~37°C melting temperature means that a large equilibrium fraction of this SQRM would have no structure at all. It should be stated that the contribution of fluorophore-quencher interaction is ignored by mfold.

Out of the 10 Type I SQRM identified using Access-Search, only 3, SQRM 287 ← 304, SQRM 351 ← 381, and SQRM 1093 ← 1116, are predicted by mfold to actually form stem-loop hairpins. Since a large number of SQRM are not predicted to fold into their intended stem-loop structures, the melting temperatures of all of the SQRM were experimentally determined. The SQRM melting data (Tables 3 and 4) suggest that Type I SQRM do not

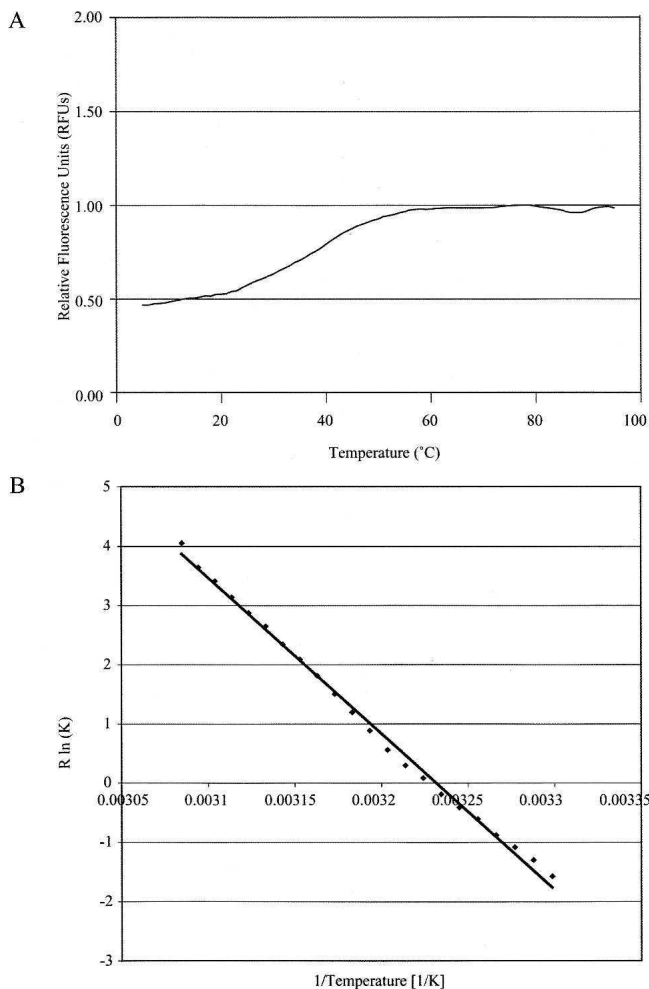


FIGURE 2. (A) Thermal denaturation profiles for SQRM 807 loop after correction using fluorescence of a fluorescein-conjugated oligonucleotide without DABCYL. The melting temperature for the internal structure of SQRM 807 loop was determined to be 37.5°C. (B) van't Hoff plot for SQRM 807 loop and linear curve fit used to determine thermodynamic parameters.

necessarily form hairpins at 37°C. Four of 10 Type I SQRMs were over 75% unstructured at the 37°C mapping temperature, while only 1 was more than 75% structured. The data regarding Type II SQRM structures were similar. For example, Type II SQRM 802 ← 832 had a 31.7°C melting temperature and was mostly unstructured at 37°C, while two Type II SQRMs 812 ← 842, and the 807 loop (807 ← 832), each with melting temperatures of ~37.5°C, were half structured, and half unstructured at 37°C.

c-kit 3-kb mapping experiment

SQRMs were incubated with full-length *c-kit* mRNA and the results were compared to the previous experiment using the 1.2-kb *c-kit* RNA fragment as the target (Gifford et al. 2005b). All SQRMs were incubated with *c-kit* RNA, an oligonucleotide (DNA) complementary to the SQRM

(positive control) and a scrambled oligodeoxynucleotide (SCR; negative control). Multiple SQRMs were screened simultaneously on a fluorescence plate reader. The signal-to-noise ratios of the SQRMs in the presence of *c-kit* 1.2 and 3 kb are shown in Figure 4. The signal is the fluorescence intensity of each SQRM with RNA, and the noise is the background fluorescence of each SQRM (SQRM alone). Target sites with a ratio close to 1 are thought not to be accessible to SQRM binding *in vitro*, while those with an arbitrarily chosen signal-to-noise ratio of 5.0 or greater are more accessible to the SQRMs (Gifford et al. 2005a,b).

The fluorescence intensity of SQRM with RNA signal-to-noise ratios varied according to target site. Fluorescence intensity also depended on the length of the RNA screened. The data in Figure 4 indicate that Type II SQRM 420 ← 446 and Type I SQRMs 663 ← 683, 807 ← 837, and 962 ← 994 are able to bind to the *c-kit* 1.2-kb fragment. In contrast to this, Type II SQRM 802 ← 832 and Type I SQRMs 663 ← 683, 807 ← 837, 962 ← 994, and 1093 ← 1116 are able to bind with *in vitro*-transcribed *c-kit* 3-kb RNA. Type II SQRM 420 ← 446 does not hybridize with the full-length RNA, while Type II SQRM 802 ← 832 and Type I SQRM 1093 ← 1116 exhibit signal-to-noise ratios high enough to indicate binding to *c-kit* 3 kb. This indicates some sequence accessibility changes around 420 → 446, 802 → 832, and 1093 → 1116 due to the length of the target RNA. These data show that there is little difference in the ability of Type I and II SQRMs to hybridize to their RNA target sites.

DISCUSSION

Our group has been interested in developing strategies for identifying hybridization accessible regions within mRNA molecules. The ultimate goal of these studies is to elucidate a rational approach for designing antisense oligonucleotides. While many strategies for accomplishing this goal have been proposed (Sczakiel and Far 2002), no universally effective approach has been identified. During the course of these studies we began to realize that it was highly likely our “stem-loop” probes were not stem-loops at all and our “stemless” probes were likely structured. In this way, the structure of a SQRM seemed as unpredictable based on sequence alone as the structure of the mRNA molecules we were trying to target. To investigate this hypothesis, the melting temperatures, entropies, and enthalpies of folding of SQRM and MBs were determined. In accord with the hypothesis that prompted these studies, we found little correlation between the experimentally deduced conformation of the probes and the structures they were assumed to have. Further, experimental data and mfold structural predictions also failed to correlate (Tables 2–4). In fact, only one of eight Type I SQRMs mfold predicted to be more than 70% structured at 37°C turned out to be more

TABLE 2. Thermodynamic parameters of SQRM used in the *c-kit* 3 kb mapping experiments

Name	T_m (SQRM) (°C)	$\Delta H_{2 \rightarrow 3}^{\circ}$ (kcal/mol)	$\Delta S_{2 \rightarrow 3}^{\circ}$ [cal/(mol K)]
Type I SQRM 39 (4/18)	45.9 ± 1.9	26.9 ± 5.2	84 ± 17
Type I SQRM 137 (4/19)	36.4 ± 1.0	13.1 ± 1.0	42 ± 3
Type I SQRM 287 (4/10)	39.4 ± 2.0	28.8 ± 1.8	92 ± 6
Type I SQRM 351 (5/21)	<23.9	N/A	N/A
Type II SQRM 420 (0/27)	27.5 ± 0.2	25.1 ± 0.5	84 ± 2
Type I SQRM 580 (4/20)	37.9 ± 1.5	28.1 ± 3.4	90 ± 11
Type I SQRM 663 (5/11)	41.0 ± 0.7	25.7 ± 1.6	82 ± 5
Type II SQRM 802 (0/31)	31.7 ± 0.5	23.4 ± 1.8	77 ± 6
Type I SQRM 807 (5/21)	37.8 ± 0.8	14.6 ± 0.4	47 ± 2
Type II SQRM 812 (0/31)	37.5 ± 0.6	12.4 ± 1.0	40 ± 3
Type II SQRM 807 loop (0/21)	37.5 ± 0.5	27.8 ± 1.9	89 ± 6
Type I SQRM 924 (5/13)	21.2 ± 0.3	30.7 ± 0.8	104 ± 3
Type I SQRM 962 (5/23)	23.5 ± 0.4	37.3 ± 2.3	126 ± 8
Type I SQRM 1093 (5/14)	27.1 ± 0.5	30.4 ± 0.4	101 ± 1

All parameters were obtained using van't Hoff plots as described in the Materials and Methods section. T_m refers to the melting temperature of a SQRM, $\Delta H_{2 \rightarrow 3}^{\circ}$ refers to the change in enthalpy upon unfolding of a SQRM, $\Delta S_{2 \rightarrow 3}^{\circ}$ refers to the change in entropy upon unfolding of a SQRM. Numbers in parentheses next to SQRM name correspond to stem and loop nucleotide length, respectively.

than 70% structured in the melting experiments were performed. Retrospective analysis of previously reported studies revealed similar results. For example, melting curves were obtained for a total of 12 MBs with reverse complementary ends (Bonnet et al. 1999; Tsourkas et al. 2002a,b, 2003). Table 5 shows both the experimental and predicted melting temperatures of the MB employed in these studies. While mfold predicted all but MB 18-5 from Tsourkas et al. (2003) to fold into a stem-loop configuration, the melting temperature predictions differed by at least 10°C for half of the 12 beacons and at least 7°C for 10 of the 12. In some cases, these differences were significant because they implied a change in predicted structure at 37°C. Mfold predicts the Tsourkas et al. (2003) MBs 17-4, 19-4, and 20-4 would be predominately unstructured, since their melting temperatures are below 37°C. Experimentally, their melting temperatures are all above 37°C, demonstrating that the opposite is true; they are predominately structured.

As well as demonstrating that Type I SQRM and MBs are highly unlikely to solely adopt stem-loop configurations, at least under physiologic conditions, our studies are also of interest because they point out a lack of correlation between experimentally observed and computationally derived structural predictions. In addition, and of greater consequence, while the experimentally determined thermodynamic parameters

might indicate to what extent SQRM would be expected to have structure at the assay temperature (Tables 3 and 4), we found no apparent correlation between the predicted amount of structure and signal-to-noise ratio of SQRM fluorescence when probe and target were allowed to interact (Fig. 4). Type I SQRM 39 ← 64, for example, was over 75% structured at 37°C and had a signal-to-noise ratio of ~2:1 when incubated with *c-kit* mRNA. The six SQRM with the best signal-to-noise ratios in the *c-kit* 1.2-kb and *c-kit* 3-kb mapping experiments, Type I SQRM 663 ← 683, 807 ← 837, 962 ← 994, and 1093 ← 1116 and Type II SQRM 420 ← 446 and 802 ← 832, range from predominately structured (Type I SQRM 663 ← 683 is 72.9% structured) to predominately unstructured (Type I SQRM 962 ← 994 is 93.9% unstructured).

In our earlier work, we hypothesized that the accessibility of RNA to SQRM hybridization would change as a function of RNA target length, assuming interactions of various domains within the RNA might affect local folding (Tinoco and Bustamante 1999). Therefore, two different length *c-kit* mRNAs were synthesized in vitro, a 1.2-kb fragment and the full-length RNA (3 kb). We found that SQRM signal-to-noise ratios were similar when associating with either the 1.2- or 3-kb RNAs, with a few exceptions (Fig. 4). When hybridizing with the 1.2-kb fragment, Type II SQRM 420 ← 446 had a signal-to-noise ratio twice that when hybridized to the full-length mRNA. Conversely, Type I SQRM 351 ← 381 and Type II SQRM 802 ← 832 exhibited

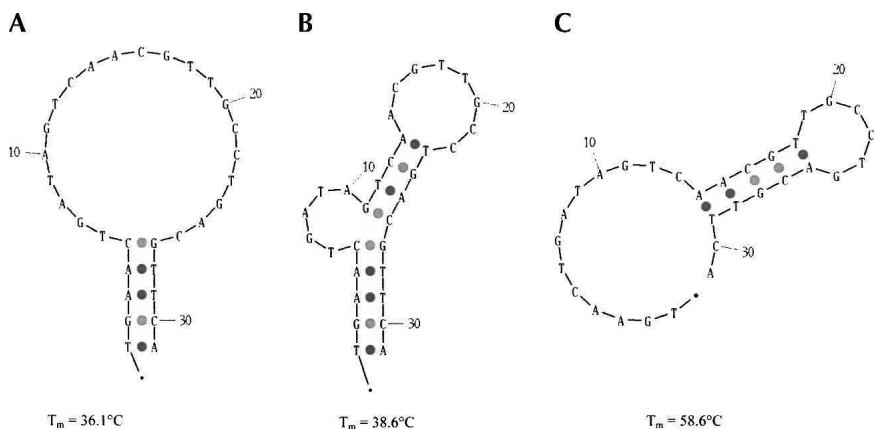


FIGURE 3. Folds and melting temperatures of SQRM 807 predicted using mfold. (A) SQRM 807 constrained to form traditional stem-loop structure. (B) SQRM 807 constrained to form a 5-nt base-paired stem. (C) Folding of SQRM 807 with no constraints.

TABLE 3. Melting temperature results and mfold predictions for Type I SQRM

Name	T_m (SQRM) experimental (°C)	T_m (SQRM) mfold (°C)	T_m (SQRM stem-loop) mfold (°C)	% Unstructured at 310 K experimental (%)	% Unstructured at 310 K mfold (%)
<i>c-kit</i> 39 (4/18)	45.9 ± 1.9	61.7	31.5	23.5 ± 1.2	6.9
<i>c-kit</i> 137 (4/19)	36.4 ± 1.0	59.4	5.3	51.0 ± 1.7	3.7
<i>c-kit</i> 287 (4/10)	39.4 ± 2.0	49.1	49.1	41.5 ± 6.6	12.5
<i>c-kit</i> 351 (5/21)	<23.9	38.0	38.0	84.9 ± 6.1	46.0
<i>c-kit</i> 580 (4/20)	37.9 ± 1.5	42.2	30.1	47.3 ± 5.1	27.4
<i>c-kit</i> 663 (5/11)	41.0 ± 0.7	56.1	42.9	37.1 ± 1.6	4.3
<i>c-kit</i> 807 (5/21)	37.8 ± 0.8	58.6	36.1	48.5 ± 1.4	2.0
<i>c-kit</i> 924 (5/13)	21.2 ± 0.3	26.2	28.1	93.5 ± 0.7	78.9
<i>c-kit</i> 962 (5/23)	23.5 ± 0.4	45.2	28.3	93.9 ± 1.5	27.8
<i>c-kit</i> 1093 (5/14)	27.1 ± 0.5	42.7	42.7	83.5 ± 1.4	27.9

higher signal-to-noise ratios when incubated with 3-kb *c-kit* mRNA. These results indicate that local changes in folding can be affected by the presence of additional RNA sequence, possibly due to tertiary structure interactions.

Previously, it was shown that mfold was unable to predict the folding of long RNA molecules (e.g., Sohail et al. 1999). In summary, the discrepancies elucidated by these studies indicate that the use of mfold to predict the structure of short DNA oligonucleotides, such as SQRM and MBs, is likely to be misleading, and that even if structure is experimentally determined, it appears to have little effect on the signal-to-noise ratio generated by a given probe. There are a variety of explanations for this finding. First, mfold cannot account for the fluorescein–DABCYL interaction, which increases the melting temperature of a hybrid (Marras et al. 2002). While the quenching efficiency was found to be enhanced by this effect, it is not known the extent to which the fluorescein–DABCYL interaction may promote or stabilize SQRM stem formation. In addition, the use of mfold to identify potentially single-stranded sites is not an exact method. Mfold does not allow for the exact experimental solvent conditions to be used in the thermodynamic parameter calculations with RNA. In vivo, other concerns arise that cannot be predicted by mfold, such as molecular crowding (Miyoshi et al. 2004) and the presence of helix-dissociating proteins (Weston and Sommerville 2006). These same molecular crowding effects and helix-dissociating proteins would influence SQRM in vivo behavior. In previous work, we have shown that the presence of a

stem-loop structure is unnecessary, at least for our purposes (Gifford et al. 2005a). This work suggests there is little structural difference between Type I and Type II SQRM and therefore the additional steps of folding SQRM in mfold to ensure the formation of a stem-loop structure and determining the melting characteristics of each SQRM are likely unnecessary as well.

MATERIALS AND METHODS

SQRM design

Several different methods were used to select the SQRM sequences employed to map the two different lengths of *c-kit* mRNA (1.2 and 3 kb, respectively) (Gifford et al. 2005a,b). Of note, all of the SQRM sequences chosen targeted sequence common to both RNA lengths. The first method, a computer program written in C++, Access-Search, was used to search 1249 bases of the *c-kit* coding RNA sequence (bases –21 to 1228; *c-kit* 1.2 kb) for target sequences with 4 or 5 complementary bases separated by 10–23 bases; when the full complement is synthesized this would form the stem-loop structure of the Type I SQRM (Gifford et al. 2005b). A total of 81 sequences were identified, and a set of 10 sites distributed along the 1.2-kb fragment were randomly chosen for mapping experiments (Table 1).

Oligodeoxynucleotides (ODN) employed as positive hybridization controls had sequences identical to the selected target sequences. Negative controls for hybridization reactions had scrambled (SCR) sequences designed by making purine or pyrimidine transitions (A with G, C with T, G with A, T with C) in the ODN target sequence.

TABLE 4. Melting temperature results and mfold predictions for Type II SQRM

Name	T_m (SQRM) experimental (°C)	T_m (SQRM) mfold (°C)	% unstructured at 310 K experimental (%)	% unstructured at 310 K mfold (%)
<i>c-kit</i> 420 (0/27)	27.5 ± 0.2	25.0	78.3 ± 0.8	76.0
<i>c-kit</i> 802 (0/31)	31.7 ± 0.5	58.6	65.8 ± 0.2	2.0
<i>c-kit</i> 812 (0/31)	37.5 ± 0.6	39.1	49.3 ± 0.9	41.8
<i>c-kit</i> 807 loop (0/21)	37.5 ± 0.5	39.1	48.3 ± 1.8	41.8

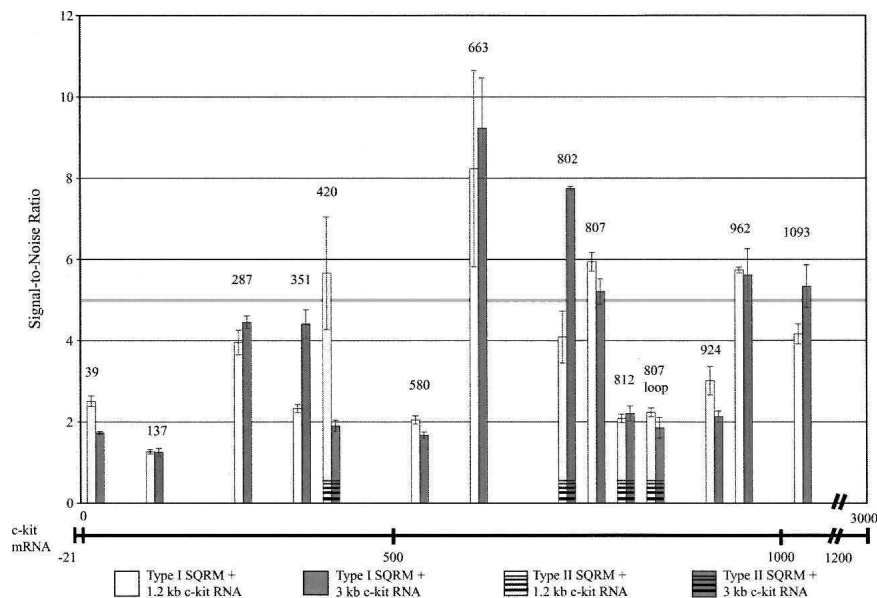


FIGURE 4. Comparison of signal-to-noise ratios for SQRM binding *c-kit* 1.2-kb RNA and *c-kit* 3-kb RNA. The light gray line represents the 5:1 signal-to-noise ratio threshold for accessibility. The white bars show data for SQRM hybridized to the *c-kit* 1.2-kb fragment; gray bars show SQRM plus *c-kit* 3-kb fragment data. The target sites are shown on the horizontal axis and are spaced proportionally to their location in the *in vitro* transcribed RNA transcript. All assays were performed in triplicate.

SQRM synthesis

All nucleic acid molecules used in these studies were obtained from the University of Pennsylvania Cancer Center Nucleic Acid Facility. SQRM were synthesized as described previously, using fluorescein and 3'-C7-4-(4'dimethylaminophenylazo)benzoic acid (DABCYL) as the fluorophore-quencher pair (Gifford et al. 2005b). Each SQRM has five phosphorothioate nucleotides at each end flanking a central region containing phosphodiester nucleotides, as well as fluorescein and DABCYL moieties conjugated to the 5' and 3' ends, respectively. SQRM were numbered according to the following convention: A SQRM's number designator corresponds to the 5' base of mRNA to which the 3' base of the SQRM complements, the adenine of the initiation codon is base (+1), and the base 5' of the initiation A is designated (-1), no zero base. SQRM sequences are listed in Table 1. Arrows indicate 5' to 3' direction of the nucleotide sequence.

In vitro transcription of RNA

cDNA corresponding to *c-kit* mRNA bases -21 to 1228 or -21 to 3031 was cloned into a pcDNA3 vector (Promega). The resulting plasmids were then utilized in the RiboMAX Large Scale RNA Production System-T7 (Promega) to transcribe 1.2- and 3-kb fragments of *c-kit* mRNA. The *in vitro*-transcribed RNA transcripts were purified using the RNeasy Mini Kit (Qiagen).

Melting temperature determination

Thermal denaturation profiles for the SQRM used to map *c-kit* 3 kb were obtained by monitoring the fluorescence of SQRM (100 nM) incubated in SQRM Buffer using an iCycler iQ (Bio-Rad), 490-nm excitation filter and 525-nm emission filter. Fifty micro-

liters of SQRM solution were loaded onto 96-well 0.2-mL thin-wall PCR plates (Bio-Rad). The solutions were incubated at 95°C for 10 min, and the temperature was subsequently decreased in steps of 1°C. Each step lasted 2.5 min, and the fluorescence of each SQRM solution was measured at the end of each step. To ensure the absence of hysteresis, each solution was then incubated again at 95°C for 3 min, followed by 4°C for 10 min, and the temperature was subsequently increased in steps of 1°C. Each step lasted 2.5 min, and the fluorescence of each SQRM solution was measured at the end of each step.

To correct for the temperature dependence of fluorescence in SQRM buffer, a linear correction factor was obtained for each SQRM by fitting the high temperature fluorescence intensity data (where the SQRM are thought to be unstructured) with straight lines. Each SQRM thermal denaturation curve was corrected using this factor and normalized by dividing by the maximal corrected fluorescence (Bonnet et al. 1999).

Thermodynamic determinations

As described by Bonnet et al. (1999) SQRM alone in solution can be approximated by a system with two states: closed and open. When equilibrium is reached, the system can be described by the following equation:

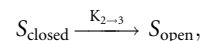


TABLE 5. Melting temperature comparison between other investigators' experiments and mfold predictions

Name	T_m (SQRM) experimental (°C)	T_m (SQRM) mfold (°C)
PNAS 96 6171 ^a	54	55.2
NAR 30 5168 ^b 2'-deoxy-MB	43	45.3
NAR 31 1319 ^c		
MB 17-4	44.4	26.2
MB 19-4	41.0	26.2
MB 20-4	39.9	29.2
MB 17-5	58.2	43.5
MB 18-5	56.7	49.6
MB 19-5	57.4	45.2
FamMB 19-5	53.9	45.2
MB 17-6	67.0	56.6
MB 18-6	65.5	55.8
MB 19-7	67.0	59.7

^aBonnet et al. (1999).

^bTsourkas et al. (2002a).

^cTsourkas et al. (2003).

where S_{closed} is the closed SQRM, S_{open} is the open SQRM, and $K_{2 \rightarrow 3}$ describes the equilibrium constant for the transition from a structured to an unstructured SQRM. The Bonnet et al. (1999) notation is used here for consistency.

The fluorescence of the solutions of SQRM without ODN targets can be written as the sum of the fluorescence intensity of closed and open SQRM:

$$F = \beta \frac{[S_{\text{closed}}]}{S_o} + \gamma \frac{[S_{\text{open}}]}{S_o},$$

where β is the characteristic fluorescence intensity of SQRM in the closed state, γ is the characteristic fluorescence intensity of SQRM in the open state, and S_o is the total concentration of SQRM:

$$S_o = [S_{\text{closed}}] + [S_{\text{open}}].$$

The equilibrium constant for the transition from the closed state to the open state can be written as

$$K_{2 \rightarrow 3}^o = \left(\frac{F - \beta}{\gamma - F} \right),$$

as derived by Bonnet et al. (1999). The above equation can be substituted into the van't Hoff equation and plotted as

$$R \ln \left(\frac{F - \beta}{\gamma - F} \right) = - \Delta H_{2 \rightarrow 3}^o \frac{1}{T} + \Delta S_{2 \rightarrow 3}^o,$$

where $\Delta H_{2 \rightarrow 3}^o$ is the change in enthalpy of the transition and $\Delta S_{2 \rightarrow 3}^o$ is the change in entropy of the transition, $R = 1.99$ cal/mol K is the universal gas constant, and T is the temperature of the solution in Kelvin, since $\Delta G^o = -RT \ln K(T) = \Delta H^o - T\Delta S^o$ where ΔG^o is free energy.

The melting temperature, T_m , for each transition from closed to open SQRM can be determined by substituting $\Delta G^o = 0$ to get

$$T_m = \frac{\Delta H_{2 \rightarrow 3}^o}{\Delta S_{2 \rightarrow 3}^o}.$$

The fraction of SQRM that do not have structure and fluoresce at a given temperature can be determined empirically using (Tsourkas et al. 2003)

$$\% \text{unstructured} = \frac{K_{2 \rightarrow 3}}{1 + K_{2 \rightarrow 3}}.$$

Computational predictions of thermodynamic parameters

The sequence of each SQRM was given to mfold (Zuker 2003) with a theoretical buffer solution (100 mM Na⁺, 2 mM Mg²⁺ at 37°C), similar to the SQRM buffer used in previous experiments (Gifford et al. 2005a,b). The program output consisted of a set of lowest free energy secondary structures as well as predicted enthalpies and entropies of folding and melting temperatures (<http://www.bioinfo.rpi.edu/applications/mfold>).

Solution mapping experiments

SQRMs (100 nM) were incubated with in vitro transcribed *c-kit* mRNA (1 μM), ODN target (1 μM), or SCR target (1 μM) in SQRM Buffer (100 mM Tris, 2 mM MgCl₂ at pH 7.5). Fluorescence was monitored after reactions were incubated for 30 min at 37°C in a Packard FluoroCount fluorometer using an excitation filter of 490 nm and an emission filter of 525 nm. The signal-to-noise ratios (SQRM incubated with target divided by the SQRM alone) were computed as described previously (Gifford et al. 2005b), using the median of ratios method and implemented with the bootstrap algorithm (Brody et al. 2002).

ACKNOWLEDGMENTS

This study was supported in part by grants from the NIH (PO1-CA72765 and RO1CA101859) (A.M.G.), the Doris Duke Charitable Foundation (A.M.G.), and the Pennsylvania Department of Health. The Department specifically disclaims responsibility for any analyses, interpretations, or conclusions. A.M.G. is a Distinguished Clinical Scientist of the Doris Duke Charitable Foundation. V.P. was a member of the Roy and Diana Vagelos Scholars Program. We thank Adam Peritz for helpful discussions and David Jordan for computational assistance. All of the SQRM and ODNs were synthesized by X. Zhang and A.E. Peritz of the University of Pennsylvania Cancer Center Nucleic Acid Facility.

Received October 26, 2007; accepted December 27, 2007.

REFERENCES

- Ashman, L.K. 1999. The biology of stem cell factor and its receptor C-kit. *Int. J. Biochem. Cell Biol.* **31**: 1037–1051.
- Bonnet, G., Tyagi, S., Libchaber, A., and Kramer, F.R. 1999. Thermodynamic basis of the enhanced specificity of structured DNA probes. *Proc. Natl. Acad. Sci.* **96**: 6171–6176.
- Brody, J.P., Williams, B.A., Wold, B.J., and Quake, S.R. 2002. Significance and statistical errors in the analysis of DNA microarray data. *Proc. Natl. Acad. Sci.* **99**: 12975–12978.
- Gifford, L.K., Jordan, D., Pattanayak, V., Vernovsky, K., Do, B.T., Gewirtz, A.M., and Lu, P. 2005a. Stemless self-quenching reporter molecules identify target sequences in mRNA. *Anal. Biochem.* **347**: 77–88.
- Gifford, L.K., Opalinska, J.B., Jordan, D., Pattanayak, V., Greenham, P., Kalota, A., Robbins, M., Vernovsky, K., Rodriguez, L.C., Do, B.T., et al. 2005b. Identification of antisense nucleic acid hybridization sites in mRNA molecules with self-quenching fluorescent reporter molecules. *Nucleic Acids Res.* **33**: e28. doi: 10.1093/nar/gni024.
- Kuhn, H., Demidov, V.V., Coull, J.M., Fiandaca, M.J., Gildea, B.D., and Frank-Kamenetskii, M.D. 2002. Hybridization of DNA and PNA molecular beacons to single-stranded and double-stranded DNA targets. *J. Am. Chem. Soc.* **124**: 1097–1103.
- Mahato, R.I., Cheng, K., and Guntaka, R.V. 2005. Modulation of gene expression by antisense and antigene oligodeoxynucleotides and small interfering RNA. *Expert Opin Drug Deliv* **2**: 3–28.
- Marras, S.A.E., Kramer, F.R., and Tyagi, S. 2002. Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes. *Nucleic Acids Res.* **30**: e122. doi: 10.1093/nar/gnf121.
- Miyoshi, D., Matsumura, S., Nakano, S., and Sugimoto, N. 2004. Duplex dissociation of telomere DNAs induced by molecular crowding. *J. Am. Chem. Soc.* **126**: 165–169.
- Parkhurst, K.M. and Parkhurst, L.J. 1995a. Kinetic studies by fluorescence resonance energy transfer employing a double-labeled

- oligonucleotide: Hybridization to the oligonucleotide complement and to a single-stranded DNA. *Biochemistry* **34**: 285–292.
- Parkhurst, K.M. and Parkhurst, L.J. 1995b. Donor–acceptor distance distributions in a double-labeled fluorescent oligonucleotide both as a single strand and in duplexes. *Biochemistry* **34**: 293–300.
- Sattler, M. and Salgia, R. 2004. Targeting c-Kit mutations: Basic science to novel therapies. *Leuk. Res.* **28**: S11–S20.
- Scherer, L.J. and Rossi, J.J. 2003. Approaches for the sequence-specific knockdown of mRNA. *Nat. Biotechnol.* **21**: 1457–1465.
- Sczakiel, G. and Far, R.K. 2002. The role of target accessibility for antisense inhibition. *Curr. Opin. Mol. Ther.* **4**: 149–153.
- Sohail, M., Akhtar, S., and Southern, E.M. 1999. The folding of large RNAs studied by hybridization to arrays of complementary oligonucleotides. *RNA* **5**: 646–655.
- Tinoco Jr., I. and Bustamante, C. 1999. How RNA folds. *J. Mol. Biol.* **293**: 271–281.
- Tsourkas, A., Behlke, M.A., and Bao, G. 2002a. Hybridization of 2'-O-methyl and 2'-deoxy molecular beacons to RNA and DNA targets. *Nucleic Acids Res.* **30**: 5168–5174. doi: 10.1093/nar/gkf635.
- Tsourkas, A., Behlke, M.A., and Bao, G. 2002b. Structure–function relationships of shared-stem and conventional molecular beacons. *Nucleic Acids Res.* **30**: 4208–4215. doi: 10.1093/nar/gkf536.
- Tsourkas, A., Behlke, M.A., Rose, S.D., and Bao, G. 2003. Hybridization kinetics and thermodynamics of molecular beacons. *Nucleic Acids Res.* **31**: 1319–1330. doi: 10.1093/nar/gkg212.
- Tyagi, S. and Kramer, F.R. 1996. Molecular beacons: Probes that fluoresce upon hybridization. *Nat. Biotechnol.* **14**: 303–308.
- Weston, A. and Sommerville, J. 2006. Xp54 and related (DDX6-like) RNA helicases: Roles in messenger RNP assembly, translation regulation and RNA degradation. *Nucleic Acids Res.* **34**: 3082–3094. doi: 10.1093/nar/gkl409.
- Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**: 3406–3415. doi: 10.1093/nar/gkg595.