Articles

Rearrangement of Partially Ordered Stacked Conformations Contributes to the Rugged Energy Landscape of a Small RNA Hairpin

Aaron L. Stancik and Eric B. Brauns*

Department of Chemistry, University of Idaho, Moscow, Idaho 83844-2343

Received June 23, 2008; Revised Manuscript Received August 7, 2008

ABSTRACT: We have studied the fast relaxation kinetics of a small RNA hairpin tetraloop using time-resolved infrared spectroscopy. A laser-induced temperature jump initiated the relaxation by rapidly perturbing the thermal equilibrium of the sample. We probed the relaxation kinetics at two different wavenumbers, 1574 and 1669 cm^{-1}. The latter is due to the C6═O6 carbonyl stretch of the base guanine and is a direct measure of guanine base pairing. The former is assigned to a ring vibration of guanine and tracks structure by sensing base stacking interactions. Overall, the kinetics at 1574 cm^{-1} are faster than those observed at 1669 cm^{-1}. When relaxation occurs at the melting temperature, the kinetics at both wavenumbers are biexponential. When relaxation occurs at a temperature that is higher than the melting temperature, the data at 1669 cm^{-1} are still biexponential while only a single fast phase is resolved in the data at 1574 cm^{-1}. The fast phases are in the range of microseconds, while the slower phases are in the range of tens of microseconds. At both wavenumbers, a portion of the relaxation is not resolved, indicating the existence of a very fast, sub-100 ns phase. Our results provide additional evidence that small, fast folding hairpin loops are characterized by a rugged energy landscape. Furthermore, our data suggest that single-strand stacking interactions and stacking interactions in the loop contribute significantly to the ruggedness of the energy landscape. This work also demonstrates the utility of time-resolved infrared spectroscopy in studying RNA folding.

RNA folds in a hierarchical manner; small, independently stable secondary structural motifs fold first and are subsequently assembled into larger, three-dimensional architectures (*1*−*3*). Numerous types of structural motifs exist, but the hairpin loop is the most ubiquitous and is found in all RNA molecules (*4, 5*). Understanding the formation of the hairpin loop is therefore an integral component of the overall RNA folding problem. A hairpin loop forms when a region of RNA folds back on itself, forming a stem comprised of base pairs and a loop comprised of unpaired bases. Many loop sizes are found in nature, ranging from diloops (two unpaired bases in the loop) to tetradecaloops and larger. However, the tetraloop is by far the most common. Furthermore, tetraloops with the sequence UNCG (where N is any nucleotide) are the most common of all the tetraloops and occur at a statistically higher rate that other sequences in nature (*4, 5*). The impetus for studying the folding kinetics of RNA tetraloops is clear.

The underlying folding dynamics of RNA tetraloops (and hairpins in general) are not nearly as straightforward as their apparent structural simplicity might suggest. As newer techniques are employed, increasing detail is uncovered and the true complexity is revealed. Recent evidence (experimental, theoretical, and computational) supports the notion that the energy landscape of even small hairpin loops is quite rugged and that simple two-state kinetics cannot accurately describe folding. For example, Ansari has developed a "configurational diffusion" model to describe the folding mechanism of small nucleic acid hairpins (*6*). The recognition that hairpin formation was much slower than the rates of end-to-end collisions stressed the necessity of this model (*7, 8*). Similarly, work by the Bevilacqua and Gruebele groups using mutants of well-characterized tetraloops provides strong evidence that supports a rugged energy landscape (*9*−*12*). An underlying theme in their work is that the location of the probe determines the observed folding kinetics. Essentially, populations of intermediate states are exposed by monitoring the kinetics from different locations within the molecule. A recent communication by Thirumalai and coworkers explicitly states that multiple probes are required to explore the rugged energy landscape of RNA hairpins (*13*).

Folding complexity is attributed to misfolded traps and alternative populations, both of which contribute to the ruggedness of the energy landscape. Recent theoretical work treats the non-native states as structures with incorrect or partially folded stems. However, it is possible that the non-native states could also be due to single-stranded conformations having alternative base stacking orientations. Accounting for such states from a theoretical standpoint poses a

* To whom correspondence should be addressed. Phone: (208) 885-4349. Fax: (208) 885-6173. E-mail: ebrauns@uidaho.edu.

10.1021/bi801170c CCC: $40.75
Published on Web 09/23/2008

challenge and has not yet been explicitly addressed. Since these theoretical models are often used to interpret experimental data, the contribution of single-stranded intermediates has been neglected. It could also be said that the omission of single-stranded intermediates from theoretical models is due to the lack of experimental evidence of their existence. We believe that the work presented here provides this evidence.

To explore the ruggedness of the folding landscape, two things are needed: (1) a method for initiating folding and unfolding on a sufficiently short time scale so that mechanistic nuances can be accessed and (2) a detection method that provides some type of structurally specific information. The first requirement can be satisfied by using a laser-induced temperature jump (*T*-jump)[1] to initiate unfolding on nanosecond time scales (picosecond laser-induced *T*-jumps are also possible). This is the method used in the experimental work cited above and is also the method that we employ in the work reported here. The second requirement is more difficult to fulfill. One way is to insert non-native probes at various locations within the molecule. This method provides localized kinetic information but is cumbersome (involving novel synthetic strategies) and may perturb the native structure to some small degree. An alternative solution is to exploit some intrinsic spectroscopic feature that is native to the molecule. For example, it is well-known that UV−visible absorption in the 260−280 nm range is a reliable indicator of nucleic acid structure. However, probing specific structural locations within the molecule is not possible because this absorption band is broad and featureless. As a result, poorly populated (but not insignificant) intermediates or misfolded states may not be identifiable.

Our method mitigates some of these issues by probing the relaxation kinetics (subsequent to a laser-induced *T*-jump) with infrared (IR) spectroscopy. IR spectroscopic transitions arise from vibrations of specific molecular groups within a molecule. For example, C═O, C═C, and C═N vibrations tend to have strong, distinguishable absorptions in the 1500−1700 cm^{-1} range (*14*). These groups are found in all nucleotide bases. When RNA undergoes a structural change, the local environment surrounding these groups is altered and spectral changes manifest. This is quite similar to the kind of hypo- or hyperchromic changes observed in the UV spectral region. However, there are several structural features that are present that can be monitored so that localized kinetic information can be obtained. In this way, we can exploit the intrinsic properties of the molecule and monitor structurally specific kinetics. In earlier work, we demonstrated the utility of this technique by studying the folding of a larger tRNA molecule (*15*). From this study, we identified two distinct unfolded populations and postulated that tRNA folds via two parallel pathways each comprised of multiple steps. To the best of our knowledge, this was the first demonstration of time-resolved IR spectroscopy of RNA following a laser-induced *T*-jump.

In the work reported here, we study the folding kinetics of the UCCG variant of the UNCG class of tetraloops. In addition to highlighting the utility or our technique, these results provide further support for the idea that the folding landscape of a small hairpin loop is quite rugged and that stacking interactions contribute significantly to the ruggedness.

EXPERIMENTAL PROCEDURES

Sample Preparation. The RNA oligonucleotide sequence used in this work is 5′-gcUCCGgc-3′. The bases written in uppercase are the unpaired bases found in the loop, and those written in lowercase comprise the stem. The HPLC-purified oligonucleotide was purchased from Integrated DNA Technologies. Samples were prepared by dialysis against 100 mM phosphate buffer and then lyophilization against D_2O three times to remove labile protons. The final solution contained RNA at a concentration of ~15 mM (nucleotide) and a buffer pD of 7.2. This concentration is below the level needed to promote duplex formation (*16, 17*) and has been verified by our own dilution experiments.

The sample was placed in a custom-built sample cell that was used for the *T*-jump experiments and FTIR spectroscopy. Reliable comparisons between the two experimental components are facilitated by using the same cells for both. The cells are comprised of two CaF_2 windows separated by a Teflon spacer that defines the cell path length. The spacer also divides the cell into two compartments: one for the sample (RNA and buffer) and one for the reference (buffer alone). In this way, sample and reference measurements are obtained by simply translating the cell from side to side. For the experiments described here, the cell path length was 53 μm. The relative path lengths of the sample and reference compartments were determined by measuring interference fringes and were found to be within 0.3 μm of each other. The CaF_2 windows are then placed in a copper housing. The windows are secured by fastening a copper faceplate to the housing. The cell is then mounted on a larger copper block that is coupled to a circulating water bath for temperature control. The cell temperature is maintained within ±0.1 °C.

FTIR Spectroscopy. Equilibrium FTIR spectra as a function of temperature were recorded on a MIR 8025 FTIR spectrometer (Newport). The copper block with an affixed cell was mounted to a computer-controlled translation stage and placed in a custom-built sample compartment. The stage allowed sample and reference spectra to be recorded without compromising the sample compartment purge. Temperature-dependent spectra were recorded in 1 °C increments from 20 to 90 °C. At each temperature, 256 co-added scans were recorded for the sample and reference. Rather than recording 256 scans for the reference followed by 256 sample scans, we translated the cell back and forth so that sample and reference spectra were recorded alternately. This was done to minimize any long-term baseline drift.

Because of the spectral collection method just described, baseline deviations were essentially eliminated, and no baseline adjustment of the final absorbance spectra was necessary. For presentation, the spectra were smoothed using a second-order, 11-point Savitzky−Golay algorithm. Absorbance versus temperature was plotted for two different wavenumbers, 1574 and 1669 cm^{-1}, to produce melting curves. The melting curves were fit to a thermodynamic model (described in the text). The melt curves shown in Figure 2 were obtained from the raw data (not the smoothed data) and were normalized using parameters obtained from the fit.

[1] Abbreviations: FTIR, Fourier transform infrared; *T*-jump, temperature jump; MCT, mercury cadmium telluride.

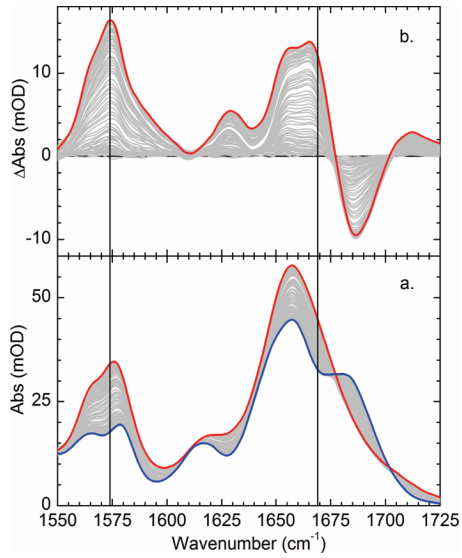

FIGURE 1: (a) Equilibrium FTIR spectra and (b) corresponding difference spectra. In panel a, the blue curve is the spectrum at the lowest temperature (20 °C) and the red curve is the spectrum at the highest temperature (90 °C). The difference spectra were calculated by subtracting the lowest-temperature spectrum from the spectra at all other temperatures. The solid vertical lines are at 1574 and 1669 cm^{-1}, the probe wavenumbers.

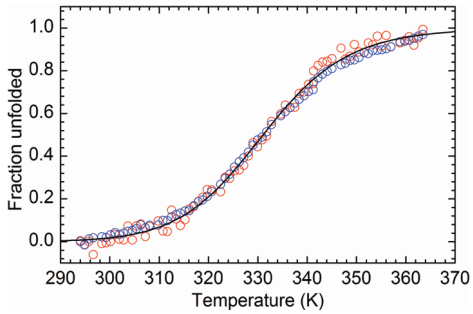

FIGURE 2: Normalized melting curves for the data recorded at 1574 cm^{-1} (blue circles) and 1669 cm^{-1} (red circles). The solid line through the data points is a fit to the thermodynamic model described in the text.

Time-Resolved IR Spectroscopy. The sample is probed by a CW diode laser tuned to either 1574 or 1669 cm^{-1}. Spectral resolution is provided by the narrow line width of the diode laser (<0.5 cm^{-1}). The diode laser is focused onto the sample by a gold-coated off axis paraboloidal mirror (90° off axis; parent focal length of 25.4 mm). The focused beam diameter is less than 100 μm. The transmitted probe radiation is collected by a second off axis paraboloidal mirror and then focused onto a 20 MHz MCT detector. The *T*-jump pulse is generated by Raman shifting the fundamental output of a Q-switched Nd:YAG laser in a 1 m long Raman cell filled with H_2 gas at 500 psi. The laser operates at a repetition rate of 10 Hz and has a pulse temporal width of 10 ns. The first Stokes line is at a wavelength of 1.9 μm and serves as the *T*-jump pulse. A Pellin-Broca prism is used to separate the 1.9 μm radiation from the Stokes, anti-Stokes, and residual Rayleigh radiation. The *T*-jump pulse overlaps with the probe laser at the sample. A long focal length lens (f = 250 mm) is used to control the diameter of the *T*-jump pulse. The pump pulse is adjusted to be roughly 5 times the diameter of the probe laser at the point of overlap. This ensures uniform heating of the probe volume and minimizes the effects of drift.

D_2O has a near-IR absorption band centered around 1.9 μm. The D_2O solvent transmits 75−80% of the *T*-jump pulse in our cells (path length of 53 μm). The energy from the fraction that is absorbed generates the temperature jump. The specific heat capacity of D_2O is 4.22 J K^{-1} g^{-1}. Assuming 20% absorption and a sample volume of 3.9 × 10^{-5} mL, one can calculate that ~9 mJ of 1.9 μm radiation is required to generate a 10 °C temperature jump. Since the thermal equilibration of the aqueous solvent occurs on a picosecond time scale, the temperature jump is instantaneous when induced by a 10 ns laser pulse. Hence, the time resolution of the experiment is limited by the laser pulse width and/or the detector rise time (~20 ns combined).

Prior to the collection of time-resolved data, the transmission of the probe wavenumber through the reference as a function of temperature is measured. Using these data, the transient absorption through the reference is used as an internal thermometer to measure the magnitude of the *T*-jump.

The arrival of the *T*-jump pulse defines time zero and triggers data collection. The transient absorption profile subsequent to the *T*-jump is detected by the MCT detector described above. The transient waveform from 0 to 1.4 ms is digitized at 5 ns intervals by a 14-bit high-speed digitizer (100 MHz bandwidth, 200 MS/s). Each waveform is averaged over 3000 laser shots. The time-dependent absorption is calculated for each sample/reference pair. Finally, the time-resolved absorbance of the sample is obtained by subtracting the reference absorption from the sample absorption.

Digitizing the waveforms at 5 ns increments over nearly 6 decades yields more than 260000 data points. This has two drawbacks. One is that it is simply inconvenient; the excessive number of data points is cumbersome and impedes data analysis and display. The second drawback is more practical. Digitizing a long time span in small intervals leads to unnecessary point density at longer times. For example, at 25 ns, the next point would be at 30 ns. However, at 500 μs, the next point would be at 500.005 μs. Aside from being impractical, subsequent data analysis is heavily weighted toward longer times and could lead to erroneous results. We circumvent this by interpolating the transient absorption profiles (i.e., the time-resolved data after the reference has been subtracted from the sample) using a logarithmically spaced time axis. Using this method, we reduce the number of data points to 3004. The interpolated data exactly overlay the original data.

Since the time resolution of our experimental setup is instrument-limited to ~20 ns, we fit all of our kinetic data from 100 ns. This provides ample separation between instrumental artifacts and meaningful data.

RESULTS

Equilibrium FTIR Spectroscopy. Equilibrium FTIR spectra recorded at different temperatures are shown in Figure 1. The sensitivity of IR spectroscopy to RNA structure is immediately evident from these data. In addition to hypo- and hyperchromic changes, there are band shifts, measureable changes in bandwidths, and (with more involved analyses) changes in band shape. Furthermore, most of these changes are linked to specific structural moieties within the hairpin so that structurally specific information can be extracted (*14*).

While the spectral features are somewhat broad and overlap to some degree, structural assignments can still be made. The spectral changes at 1574 and 1669 cm^{-1} are the focus of the work reported here. These are assigned in the following paragraph.

In folded RNA, the C6═O6 stretch of guanine occurs near 1685 cm^{-1}. Upon melting, this transition shifts to 1669 cm^{-1} and increases in intensity. The two low-frequency peaks around 1564 and 1577 cm^{-1} are a doublet due to ring vibrations of guanine. Both red shift and increase in intensity as the RNA unfolds. These changes manifest in the difference spectra as a strong feature at 1574 cm^{-1}. When the RNA unfolds, base pairing as well as base stacking interactions is disrupted. The proximity of the involved group to the environmental perturbation determines its sensitivity. Consequently, the spectral changes at 1574 and 1669 cm^{-1} are weighted contributions from both pairing and stacking interactions. For example, the guanine C6═O6 group is directly involved in a Watson−Crick base pair with cytosine. Therefore, changes at 1669 cm^{-1} are primarily due to base pairing with a much smaller contribution from base stacking. In contrast, the changes at 1574 cm^{-1} (due to guanine ring modes) are more sensitive to base stacking and, to a lesser extent, base pairing.

Equilibrium Melting. Normalized melt curves for each probe wavenumber are shown in Figure 2. In the equilibrium limit, both curves show two-state behavior and were analyzed accordingly. The absorbance is proportional to the fraction of unfolded RNA hairpins, and the fraction of unfolded hairpins is related to the equilibrium constant according to

$$f_u = \frac{K_{eq}}{1 + K_{eq}} \quad (1)$$

where f_u is the fraction of unfolded hairpins and K_{eq} is the equilibrium constant. Furthermore, the equilibrium constant is related to the Gibbs free energy of unfolding, ΔG, by

$$K_{eq} = \exp(-\Delta G/RT) \quad (2)$$

Finally, the enthalpy of unfolding (ΔH), the melting temperature (T_m), and the change in heat capacity (ΔC_p) are introduced into the model by noting that the temperature dependence of the Gibbs free energy is given by the Gibbs−Helmholtz relationship

$$\Delta G(T) = \Delta H(T)\left(1 - \frac{T}{T_m}\right) + \Delta C_p\left(T - T_m - T \ln \frac{T}{T_m}\right) \quad (3)$$

However, we held ΔC_p constant at zero to reduce the number of parameters in the model. While a bit of accuracy is sacrificed by doing this, more general conclusions can still be drawn, and the reliability of the fits are increased substantially. By combining the previous relationships, we fit the raw melt curves to the following model

$$A(T) = a + (b - a)\frac{e^{-\Delta G(T)/RT}}{1 + e^{-\Delta G(T)/RT}} \quad (4)$$

In this expression, a and b are constants that give the lower and upper absorption limits, respectively. After the fits were obtained, the data were converted to fraction unfolded versus temperature so that the two sets of data could be overlapped and compared directly. These results are shown in Figure 2. Within the error limit, the melting behavior for both is identical. Both show a melting transition at 331 K (58 °C). Both also have an unfolding enthalpy on the order of 100 kJ mol^{-1}. Although the data obtained from the equilibrium melting are useful, it is unremarkable without further analyses and/or experiments.

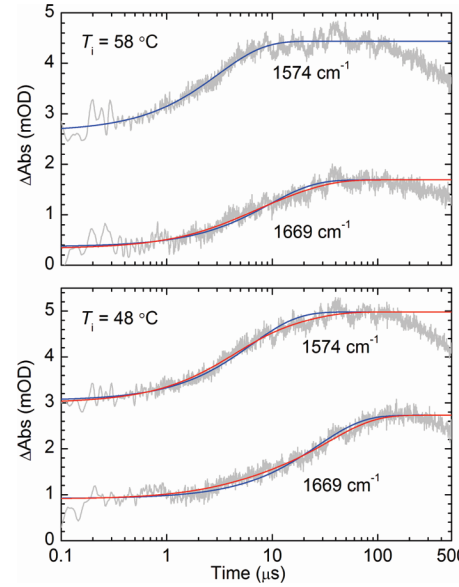

FIGURE 3: Temperature jump relaxation kinetics. The bottom plot shows the data starting from an initial temperature of 48 °C, and the top plot shows the data starting from an initial temperature of 58 °C. The gray lines are the raw data, the blue lines single-exponential fits, and the red lines double-exponential fits. Note that the time axis is logarithmic. Also note that the data have not been scaled or offset in any way; the y axes values represent the raw data.

Time-Resolved Spectroscopy. Despite the similarities in the equilibrium data, the kinetic data reveal distinct differences. The kinetic data are shown in Figure 3. Two sets of kinetic traces were recorded at each wavenumber. One trace starts from an initial temperature of 49 °C (bottom panels) and jumps 9 °C to the melting transition. The second trace at each wavenumber starts from the transition temperature (58 °C) and jumps 9 °C to a final temperature of 67 °C. Figure 2 shows that at a starting temperature of 49 °C, ~24% of the RNA population is unfolded while at the melting transition, the fraction unfolded is 50%. Likewise, 73% is unfolded at a final temperature of 67 °C.

Before discussing the quantitative aspects of the kinetics, we will describe the more general features of the time-resolved data. A particularly prominent feature in all of the kinetic traces is the deviation from the zero baseline at early times. This is most pronounced in the data recorded at 1574 cm^{-1}. At the earliest resolvable times (down to 10 ns, data not shown), the transient absorption does not extrapolate to zero. Instead, it has a finite value between 2 and 3 mOD. There must be a very fast process that is virtually complete in less than 10−100 ns. Although not nearly as distinct, this is also seen in the data recorded at 1669 cm^{-1}.

The time at which the transient absorption reaches a maximum is another qualitative feature that is worthy of discussion. When relaxation occurs at the melting temperature (i.e., initial temperature of 49 °C followed by a 9 °C T-jump), the data at 1574 cm^{-1} reach a maximum by ~78 μs. By comparison, however, the corresponding relaxation recorded at 1669 cm^{-1} does not reach a maximum until

Table 1

ν (cm^{-1})	T_i (°C)	a_1	τ_1 (μs)	a_2	τ_2 (μs)
1574	48	0.26	3.6 ± 0.2	0.14	15.6 ± 1.5
	58	0.40	3.1 ± 0.04	–	–
1669	48	0.10	3.8 ± 0.3	0.56	34.7 ± 0.4
	58	0.31	3.7 ± 0.7	0.49	14.9 ± 1.6

nearly ~180 μs. Further insight can be gained by comparing the transient absorption maxima with the equilibrium counterparts for the same temperature interval. For both wavenumbers, these are equal, indicating that the corresponding dynamics are able to fully relax to their equilibrium conformations within 78 μs (1574 cm^{-1}) and 180 μs (1669 cm^{-1}).

Similarly, when relaxation occurs at a temperature above the melting temperature (i.e., initial temperature of 58 °C and jumping to 67 °C), the transient absorption maxima are reached in 28 and 85 μs for the data recorded at 1574 and 1669 cm^{-1}, respectively. In both cases, they are faster than their lower-temperature counterparts (as would be expected). Once again, the maxima for the data recorded at 1669 cm^{-1} take roughly 2−3 times longer to reach than the 1574 cm^{-1} data. Comparison to the equilibrium data is not as revealing as in the previous case. At 1574 cm^{-1}, the transient absorption maximum and the equilibrium absorption maximum over the same temperature interval are equal. The same conclusion can be drawn as before; the corresponding dynamics are able to fully relax to their equilibrium conformation within 28 μs. However, this is not the case for the data at 1669 cm^{-1}. The transient absorption maximum is ~1 mOD less than its equilibrium maximum. However, we note that these data have the lowest signal-to-noise ratio of all the time-resolved data. Because of this and the relatively small discrepancy (1 mOD), we are reserving further analysis until more detailed experiments can be performed.

With the exception of the high temperature data recorded at 1574 cm^{-1}, all traces are best fit to a double exponential (single exponentials are shown in the figures for comparison) with the following form

$$\Delta A(t) = a_1 \exp(-t/\tau_1) + a_2 \exp(-t/\tau_2) + y_0 \qquad (5)$$

where τ_1 and τ_2 are the relaxation constants, a_1 and a_2 are their relative amplitudes, and y_0 is the maximum transient absorbance. The results are listed in Table 1. In the table, the amplitude of each exponential has been converted to a fractional amplitude for easier comparison. This was done by assuming that at time zero, $\Delta A(t)$ should equal zero. In accordance with this, y_0 represents the "full scale" and can be used to normalize the amplitudes. In addition, by assuming that the transients should extrapolate to zero at time zero, we can quantify the magnitude of the "missing" kinetics mentioned earlier.

Perhaps the most noteworthy result is that the kinetics recorded at 1574 and 1669 cm^{-1} are substantially different. As would be expected from the previous qualitative discussion, the kinetics recorded at 1574 cm^{-1} are faster (overall) than those at 1669 cm^{-1}. At the melting temperature, both sets of data show biexponential behavior. However, there are marked differences. The resolvable kinetics recorded at 1574 cm^{-1} are dominated by the faster of the two relaxation components (3.6 μs vs 15.6 μs). In contrast, the transient data recorded at 1669 cm^{-1} are dominated by the slower of the two relaxation components (34.7 μs vs 3.8 μs). The unresolvable fast component accounts for approximately ~60% of the relaxation recorded at 1574 cm^{-1}, while it accounts for roughly ~22% of the relaxation recorded at 1669 cm^{-1}.

At the higher relaxation temperature, the 1574 cm^{-1} data are fit best by a single exponential. The slower component seen at the lower initial temperature vanishes, and the resolvable relaxation is dominated by the faster component (3.1 μs). Interestingly, the fraction of the missing relaxation is still 60%. At the same higher relaxation temperature, the data recorded at 1669 cm^{-1} are still biexponential and the majority of the resolvable relaxation is still dominated by the slower component. However, it is nearly a factor of 2 faster than at the lower temperature (14.9 μs compared to 34.7 μs). The fastest of the resolvable phases is only slightly faster at 3.7 μs. However, its amplitude has increased from 10 to 31%.

DISCUSSION

We probed the relaxation kinetics at two wavenumbers, 1574 and 1669 cm^{-1}. The former is due to a ring vibration of guanine, and the latter is due to the C6═O6 stretch of guanine. The C6═O6 stretch of guanine is sensitive to its hydrogen-bonded state (i.e., base pairing). The free guanine has an absorption band around ~1685 cm^{-1}. When base pairing occurs, electron density is withdrawn from the carbonyl and the vibrational frequency red shifts and increases in intensity. However, the ring vibration around 1574 cm^{-1} is not as straightforward to assign. There is little doubt that changes in the absorption around 1574 cm^{-1} are sensitive to RNA structure. However, the exact nature of the changes is not as clear. The electron density changes that accompany the C6═O6 carbonyl are "felt" in the ring, thereby altering the vibrational characteristics of the ring mode. However, base stacking interactions would also affect this mode. While there is no precedent for this specific assignment (to the best of our knowledge), we are confident that base stacking interactions can be inferred from changes in this vibrational mode.

Recent theoretical treatments may provide molecular insight into our results (*18*−*21*). Via analysis of small hairpin sequences, it becomes feasible to enumerate possible intermediate structures. The statistical weights of the intermediates can be calculated by determining their energies from published data. A kinetic model can be constructed from these results by calculating the rates of interconversions between each structure (or cluster of related structures). This approach has been used quite successfully by Chen and coworkers with excellent results. Not only does their work predict folding intermediates, it also predicts the presence of alternative folding routes (kinetic heterogeneity). Chen's results are extremely significant and provide a much needed link between experimental data and molecular details. Unfortunately, his work cannot be used directly to model our data. However, indirect application of the model can still provide molecular insight.

According to Chen's model, base stacking interactions stabilize RNA structure; isolated base pairs do not. Hence, for a structure to be considered statistically relevant, it must

contain at least one base stack (i.e., two adjacent base pairs). The absence of stabilizing stacking interactions renders other structures insignificant. This has two benefits. For one, the so-called Turner rules can be used to calculate the energies of the potential intermediates (*22*). Second, via elimination of structures that lack base stacks, the configurational space is dramatically reduced and subsequent calculations become more tractable. By design, the model highlights stem interactions. Any intermediate ensembles or alternate pathways that are identified are due to structural diversity in the stem.

Our model system is quite small and contains only a single base stack in the stem of the native structure. Folding intermediates in our system simply cannot be due to misfolded or partially folded stems. Enumeration of "possible" intermediates for our system shows that alternative base pairings (with or without stacks) would result in structures that are energetically highly unfavorable. Other than partially folded native hairpins (either of the stem native base pairs formed), any other structure would have either non-native base pairs, small loops (e.g., di- or triloops), or a combination of the two. Despite the lack of stem diversity, we still see kinetic intermediates and evidence of kinetic heterogeneity. From this, we conclude that loop interactions and single-strand stacking interactions must also play a key role in the folding mechanism.

Indeed, favorable stacking interactions between the bases within the loop contribute to the stability of the native structure. This has been known for some time and is included in some models (*23, 24*). However, these models are still largely dominated by nearest neighbor interactions, and the loop contribution remains somewhat phenomenological; it improves the overall model accuracy but provides little additional molecular insight. It is quite feasible, however, that these intraloop and single-strand stacking interactions are strong enough to produce local energy minima as well as possible routes for alternate folding populations.

We explain our data by treating stacking and base pairing as separate degrees of freedom. Earlier work has shown that base stacking occurs in the 0.1−1 μs range (*25, 26*). Accordingly, the unresolvable fast phases in our data are attributed to this. However, the resolvable fast phases occur on the microsecond time scale. In addition, these show only a small temperature dependence, indicating that the corresponding process is opposed by a small activation energy and that the barrier is predominantly entropic. This is consistent with stacking rearrangement previously observed by Porschke (*27*). For a hairpin to form, the ends must come together, which requires that the chain bend somewhat sharply. For this to occur, the adjacent bases must be oriented to permit favorable stacking interactions. This partial conformational ordering explains the unfavorable entropy and results in microsecond relaxation dynamics. In this regard, the partially stacked intermediate is "on-pathway" since it connects the native and unfolded states. However, this does not imply that only native intramolecular interactions stabilize the intermediate. The data presented here are insufficient to make such an assignment.

In contrast, the slow phases behave differently. At 1669 cm^{-1}, this phase becomes almost twice as fast as the temperature is increased, corresponding to an activated process. On the basis of this observation and our equilibrium thermodynamic data, we believe that this is consistent with stem formation. In addition, its amplitude decreases slightly as the temperature is increased. Assuming that the fast phase corresponds to stacking rearrangement, at higher temperatures the chain is more mobile and more configurations become available. Thus, the entropic barrier is lowered. In so doing, the chain configuration is closer to its equilibrium state when the ends come together. Thereby, the amount of structural relaxation that must occur for the stem to form is smaller. Hence, its amplitude is decreased. The same argument is used to explain the disappearance of the slow phase recorded at 1574 cm^{-1} with an increase in temperature. However, since stacking precedes base pairing, the ring vibration "equilibrates" faster than the carbonyl vibration so the effects of stem formation at the higher temperature are not felt as strongly at 1574 cm^{-1}.

In the previous work of Gruebele and Bevilacqua, single-exponential relaxation kinetics are observed near the melting temperature (*9*). A fast phase appears below the melting temperature, and a slow phase appears above it. Our data are different. We see biexponential kinetics at the melting temperature for both wavenumbers. However, above the melting temperature, the kinetics at 1574 cm^{-1} are single-exponential. We do not have kinetic data below the melting temperature for comparison, but we can make comparisons with the relaxations at and above the melting temperature.

They see a main kinetic phase that is due to the transition from the native ensemble to the unfolded ensemble. Near the melting temperature, only these two states are populated enough to be observed experimentally and the relaxation is single-exponential. They argue that at higher temperatures, an ensemble of off pathway, stem sensitive structures is populated. These serve as misfolded traps that are separated from the unfolded state by a large barrier. Consequently, a slow kinetic phase is observed. Below the melting temperature, an ensemble comprised of hairpins with partially folded (native) stems is populated. These are on-pathway intermediates, and the rate of interconversion between them and the native state is fast. This results in biexponential kinetics with a fast phase in addition to the main phase.

Rather than assume that our results are in opposition, we believe that there are marked similarities and that the differences provide further evidence that the energy landscape of small RNA hairpins is extremely rugged. Recognizing that the same process viewed using different experiments can produce different results accounts for the differences. Each type of experiment is sensitive to slightly different phenomena. Previous work of Dyer and co-workers is a good example (*28*). *T*-Jump relaxation kinetics of apomyoglobin probed using tryptophan fluorescence revealed monoexponential kinetics, while the same system under the same conditions, studied using IR, revealed biexponential kinetics. The differences are not in conflict. Rather, they highlight the sensitivity of different types of measurements. In addition, these differences can be interpreted as further evidence of a relaxation process governed by a rugged energy landscape. In this regard, different types of measurements are complementary, and a variety of experiments are needed to fully characterize the process.

In terms of parallels, the main kinetic phase observed by Gruebele and our "slow" phase are likely to be analogous. Both are activated, and the rates are similar. We both assign

this phase as the interconversion between the native and unfolded states. However, we further assign this to the formation of stem base pairs. They attribute their fast phase to the interconversion between two closely related ensembles, whereas we assign our fast phase to stacking rearrangement, motion that precedes stem formation. These are not mutually exclusive observations.

CONCLUSION

Using a fast folding, relatively "simple" hairpin structure, our results show that the most fundamental aspect of RNA folding (hairpin formation) is a complex process. Multiple kinetic phases are observed that indicate the presence of local energy minima between the unfolded and native states. This supports the notion that the energy landscape is rugged. Furthermore, we monitor folding using two different (but intrinsic) spectral markers. This is akin to probing a multiply labeled RNA molecule. Our data show that the observed kinetics depend upon the probe wavenumber. Since each probe wavenumber is assigned to a structural aspect of the sample, we can discern alternative populations. Probe-dependent kinetics suggest that kinetic heterogeneity is also a characteristic of RNA hairpin formation. In fact, we observe substantially different kinetics at the two wavenumbers that were probed. Overall, the kinetics recorded at 1574 cm^{-1} are roughly twice as fast as those recorded at 1669 cm^{-1}.

We interpret our findings as evidence that misfolding in the stem is insufficient to explain the folding complexity. Our results show that loop interactions and/or single-strand stacking fluctuations also contribute to the folding complexity. The earliest stages of hairpin folding involve base stacking. Before the chain can collapse permitting loop closure, the stacks rearrange, leading to partial conformational ordering. With the bases properly oriented relative to their neighbors, the chain can collapse and zipping of the stem base pairs can occur.

REFERENCES

1. Brion, P., and Westhof, E. (1997) Hierarchy and dynamics of RNA folding. *Annu. Rev. Biophys. Biomol. Struct. 26*, 113–137.
2. Kwok, L. W., Shcherbakova, I., Lamb, J. S., Park, H. Y., Andresen, K., Smith, H., Brenowitz, M., and Pollack, L. (2006) Concordant exploration of the kinetics of RNA folding from global and local perspectives. *J. Mol. Biol. 355*, 282–293.
3. SantaLucia, J., Jr., and Turner, D. H. (1997) Measuring the thermodynamics of RNA secondary structure formation. *Biopolymers 44*, 309–319.
4. Varani, G. (1995) Exceptionally stable nucleic acid hairpins. *Annu. Rev. Biophys. Biomol. Struct. 24*, 379–404.
5. Woese, C. R., Winker, S., and Gutell, R. R. (1990) Architecture of ribosomal RNA: Constraints on the sequence of "tetra-loops". *Proc. Natl. Acad. Sci. U.S.A. 87*, 8467–8471.
6. Ansari, A., Kuznetsov, S. V., and Shen, Y. (2001) Configurational diffusion down a folding funnel describes the dynamics of DNA hairpins. *Proc. Natl. Acad. Sci. U.S.A. 98*, 7771–7776.
7. Shen, Y., Kuznetsov, S. V., and Ansari, A. (2001) Loop dependence of the dynamics of DNA hairpins. *J. Phys. Chem. B 105*, 12202–12211.
8. Wang, X., and Nau, W. M. (2004) Kinetics of end-to-end collision in short single-stranded nucleic acids. *J. Am. Chem. Soc. 126*, 808–813.
9. Ma, H., Proctor, D. J., Kierzek, E., Kierzek, R., Bevilacqua, P. C., and Gruebele, M. (2006) Exploring the energy landscape of a small RNA hairpin. *J. Am. Chem. Soc. 128*, 1523–1530.
10. Moody, E. M., Feerrar, J. C., and Bevilacqua, P. C. (2004) Evidence that folding of an RNA tetraloop hairpin is less cooperative than its DNA counterpart. *Biochemistry 43*, 7992–7998.
11. Proctor, D. J., Ma, H., Kierzek, E., Kierzek, R., Gruebele, M., and Bevilacqua, P. C. (2004) Folding thermodynamics and kinetics of YNMG RNA hairpins: Specific incorporation of 8-bromoguanosine leads to stabilization by enhancement of the folding rate. *Biochemistry 43*, 14004–14014.
12. Siegfried, N. A., Metzger, S. L., and Bevilacqua, P. C. (2007) Folding Cooperativity in RNA and DNA Is Dependent on Position in the Helix. *Biochemistry 46*, 172–181.
13. Hyeon, C., and Thirumalai, D. (2008) Multiple Probes are Required to Explore and Control the Rugged Energy Landscape of RNA Hairpins. *J. Am. Chem. Soc. 130*, 1538–1539.
14. Banyay, M., Sarkar, M., and Gräslund, A. (2003) A library of IR bands of nucleic acids in solution. *Biophys. Chem. 104*, 477–488.
15. Brauns, E. B., and Dyer, R. B. (2005) Time-resolved infrared spectroscopy of RNA folding. *Biophys. J. 89*, 3523–3530.
16. Abdelkafi, M., Leulliot, N., Baurmruk, V., Bednárová, L., Turpin, P. Y., Namane, A., Gouyette, C., Huynh-Dinh, T., and Ghomi, M. (1998) Structural Features of the UCCG and UGCG Tetraloops in Very Short Hairpins as Evidenced by Optical Spectroscopy. *Biochemistry 37*, 7878–7884.
17. Baurmruk, V., Gouyette, C., Huynh-Dinh, T., Sun, J., and Ghomi, M. (2001) Comparison Between CUUG and UUCG Tetraloops: Thermodynamic Stability and Structural Features Analyzed by UV Absorption and Vibrational Spectroscopy. *Nucleic Acids Res. 29*, 4089–4096.
18. Zhang, W., and Chen, S. J. (2002) RNA hairpin-folding kinetics. *Proc. Natl. Acad. Sci. U.S.A. 99*, 1931–1936.
19. Zhang, W., and Chen, S. J. (2006) Exploring the complex folding kinetics of RNA hairpins: I. General folding kinetics analysis. *Biophys. J. 90*, 765–777.
20. Zhang, W., and Chen, S. J. (2006) Exploring the complex folding kinetics of RNA hairpins: II. Effect of sequence, length, and misfolded states. *Biophys. J. 90*, 778–787.
21. Zhang, W., and Chen, S. J. (2003) Master equation approach to finding the rate-limiting steps in biopolymer folding. *J. Chem. Phys. 118*, 3413–3420.
22. Serra, M. J., and Turner, D. H. (1989) Predicting Thermodynamic Properties of RNA. *Methods Enzymol. 259*, 242–261.
23. Ivanov, V., Zeng, Y., and Zocchi, G. (2004) Statistical Mechanics of Base Stacking and Pairing in DNA Melting. *Phys. Rev. E 70*, 051907.
24. Kuznetsov, S. V., Shen, Y., Benight, A. S., and Ansari, A. (2001) A Seminflexible Polymer Model Applied to Loop Formation in DNA Hairpins. *Biophys. J. 81*, 2864–2875.
25. Dewey, T. G., and Turner, D. H. (1979) Laser Temperature-Jump Study of Stacking in Adenylic Acid Polymers. *Biochemistry 18*, 5757–5762.
26. Porschke, D. (1978) Molecular States in Single-Stranded Adenylate Chains by Relaxation Analysis. *Biopolymers 17*, 315–323.
27. Menger, M., Eckstein, F., and Porschke, D. (2000) Dynamics of the RNA hairpin GNRA tetraloop. *Biochemistry 39*, 4500–4507.
28. Gulotta, M., Gilmanshin, R., Buscher, T. C., Callender, R. H., and Dyer, R. B. (2001) Core Formation in Apomyoglobin: Probing the Upper Reaches of the Folding Energy Landscape. *Biochemistry 40*, 5137–5143.

BI801170C