Homogeneous Nucleation in Sickle Hemoglobin: Stochastic Measurements with a Parallel Method

Zhiqi Cao and Frank A. Ferrone
Department of Physics and Atmospheric Science, Drexel University, Philadelphia, Pennsylvania 19104 USA

ABSTRACT The homogeneous nucleation rate for sickle hemoglobin polymerization has been measured for concentrations from 3.9 to 4.9 mM and temperatures from 13°C to 35°C by observing the stochastic fluctuations of the time to complete 10% of the reaction after photolysis of the carboxy derivative. To allow efficient data collection, a mesh was used to divide the photolysis beam into an array of smaller beams, which allowed parallel observation of about 100 different regions. Nucleation rates measured here are consistent with more restricted previouosly published data and, when combined with directly measured monomer addition rates, are consistent with previous analysis of progress curves. By describing these rates with equilibrium nucleation theory, the concentration of nuclei and hence their stability can be ascertained. Consequently, the chemical potential by which a monomer is attached to the polymer is determined. This attachment energy ranges from −6.6 to −8.0 kcal/mol between 15°C and 35°C. The enthalpic part of that chemical potential is found to be equal to the enthalpy determined by solubility measurements, as expected from thermodynamic considerations. The entropic portion of the contact chemical potential contributes from −21.4 to −8.7 kcal/mol. The vibrational chemical potential of monomers in the polymer ranges from −25.7 to −27.4 kcal/mol over the same temperatures.

INTRODUCTION

Assembly in most biological systems requires organizing templates to ensure the correct location and directionality of growth. For such a strategy to work it is necessary to impede assembly without the template, and this is accomplished by imposition of kinetic barriers to such assembly. In the absence of the template, the assembly (e.g., actin, tubulin) must undertake a series of unfavorable steps before "downhill" growth is allowed. In sickle cell disease, a similar process occurs in the formation of sickle hemoglobin fibers. Energy is required for molecules of sickle hemoglobin, which are competent to form long fibers, to assemble into small aggregates. Thermal fluctuations provide the needed energy, and thus transient species of various sizes form spontaneously. At some critical size (which is labeled the nucleus) each additional molecule lowers the energy of the aggregate. The formation of nuclei is thus characterized by molecular fluctuations, and the time to form a nucleus is inherently stochastic.

Unlike assemblies that have been perfected by evolution for functional activity, sickle hemoglobin assembly involves numerous weak contacts and requires high concentrations to aggregate. The primary structure of sickle hemoglobin differs from that of hemoglobin A, which does not assemble, by the replacement of two negatively charged glutamic acids on the molecular surface by hydrophobic valines. This mutation leads to the reversible polymerization of solid-core 14-stranded fibers. Regions that make contact within the fiber also appear on the fiber surface (Mirchev and Ferrone, manuscript submitted for publication). The existence of potential contact regions on the surface of the fiber allows nucleation to occur there as well, thereby gaining added stability from the available surface contact regions. This gives rise to a pathway known as heterogeneous nucleation, and accounts for the exponential proliferation of polymers once the initial polymer nucleus is formed (Ferrone et al., 1985a). The polymers that form in this manner create a spatially extensive array called a polymer domain. Because the domain begins from a single homogeneous nucleus, any fluctuations in the formation time for that nucleus are reflected in the observed domain growth. This amplification permits the observation of fluctuations at a molecular scale by simple light-scattering intensity measurements. The distribution of characteristic polymerization times generated by such fluctuations is very useful because it can be used to determine the homogeneous nucleation rate (Hofrichter, 1986; Szabo, 1988).

To create a testable mechanism, the foregoing concept of double nucleation was embodied in a pair of differential equations (Ferrone et al., 1980). The parameters needed for that description of the double nucleation mechanism were related to a single kinetic parameter, viz. the monomer addition rate, plus simple thermodynamic constructs in equilibrium nucleation theory (Ferrone et al., 1985a). The mathematical model for double nucleation then contained five adjustable parameters (Ferrone et al., 1985a). These were determined by fitting the model predictions to an extensive set of light-scattering progress curves. The plausibility of the parameters coupled with the quality of the fit of the data suggested that the thermodynamic and kinetic descriptions of the model were accurate. The most novel part of the conceptual model was validated by direct observation of the postulated heterogeneous nucleation process

Received for publication 28 May 1996 and in final form 23 October 1996.
Address reprint requests to Dr. Frank A. Ferrone, Department of Physics, Drexel University, 32 & Chestnut St., Philadelphia, PA 19104. Tel.: 215-895-2778; Fax: 215-895-5934; E-mail: frank ferrone@coasmail.drexel.edu.

0006-3495/97/01/343/10 $2.00

by differential interference contrast (DIC) microscopy (Samuel et al., 1990). However, the quantitative description of the model was challenged when the initial use of stochastic methods revealed a homogeneous nucleation rate 100-fold larger than the double nucleation model's predictions (Hofrichter, 1986). The obvious interest in repeating and expanding upon the first stochastic measurements was offset by the intrinsic difficulty of the method. This formed the direct impetus for the development of this technique.

The need to rectify this discrepancy transcends the issue of finding a viable quantitative description of sickle hemoglobin polymerization. The description of the nucleation process employed in the model was derived from elementary constructs: the addition of a molecule to a small aggregate, the concentration of which must be determined to set the rate (Ferrone et al., 1985a). Thus the mechanism must be able to describe the stability of small aggregates accurately. Monomers in solution possess rotational and translational entropy. The reduction of that entropy in assembly is partially offset by entropy regained by center-of-mass vibrational motion of monomers within the polymer, and the remaining cost is borne by the energy of bonding molecules together in the polymer. It is important to discover whether any significant features of this assembly have been omitted in such a description, or whether this elementary description is sufficient. Conversely, once the nucleation process is accurately described, it provides a unique window into the thermodynamics of assembly. Although a monomer in the nucleus or in the infinite polymer possesses both contact energy and vibrational entropy, the contributions to the total free energy of each of these different effects depend differently on aggregate size. Hence it becomes possible to distinguish between contact free energy and free energy arising from polymer vibrational entropy when both nucleus and polymer are thermodynamically characterized. Such information is vital for a molecular description of any assembly process and provides the ultimate motivation for the measurements of nucleation undertaken here.

To characterize the fluctuations that lead to nucleation requires collection of a distribution of characteristic times, which is taken here as the time to reach one-tenth of the reaction. Because the nucleation reaction is so highly concentration dependent, its characteristic tenth time can vary from milliseconds to kiloseconds (Ferrone et al., 1985b). When a distribution of long times must be collected, the duration of the experiment can become prohibitive, and at the very least can create stringent demands for sample stability. To circumvent this difficulty, we have developed a method for performing such measurements in parallel rather than in series (Cao and Ferrone, 1996). Because the reaction is initiated by laser photolysis of the CO derivative, we have broken the laser beam into an array of smaller beams, and view all of the photolyzed spots with a detector with spatial resolution, viz. a CCD camera. Light scattering in the forward direction can then be used to follow the formation of polymers.

In this paper we describe the implementation of such a parallel stochastic method. We find it possible to measure nucleation rates as a function of concentration and temperature. The data fundamentally agree with the homogeneous nucleation theory previously proposed once experimentally measured elongation rates are used. With such agreement, the chemical potential of the monomers aggregated in polymers can be examined to shed new light on the aggregation process.

MATERIALS AND METHODS

Sickle hemoglobin was prepared by standard techniques (Cho and Ferrone, 1991) and stored in liquid nitrogen as the oxy derivative until use. Water-saturated CO was added in a glove box, which was flushed extensively with the gas. Sodium dithionite was added to give a final concentration of 60 mM. Samples (about 2 μl) were prepared between glass coverslips in the CO box and sealed with Kerr sticky wax. Typical thickness of the samples was 4–5 μm, as determined by optical absorption of the sample. Concentrations were determined by measurement of the COHb Soret spectrum. To permit high-precision determination of concentration, a series of control experiments were used to calibrate the amount of material lost by adherence to the micropipettes used (0.10 μl is lost in this fashion). Concentration accuracy is validated in the control experiments described below.

The apparatus is shown in Fig. 1. The sample is placed on a thermoelectric stage (Cambion) with a controller of our own design. The sample is enclosed in a plastic bag into which the objective is inserted and through which nitrogen gently circulated to allow measurements at temperatures below the dew point. Temperatures were calibrated by placing a thermistor at the stage position. The sample is observed in a horizontal microspectrophotometer. Long working distance, strain-free objectives (Leitz $\times$10) were used as condenser and objective. A projective 2.5$\times$ eyepiece (Olympus) projects an image of the sample onto a CCD detector (Thomson TH7883 in a cooled enclosure and camera system produced by Photometrix, Tuscon, AZ). The beam is split to a Newvicon detector and monitor as well. For absorption measurements, the light from a xenon high-pressure 150-W arc lamp is passed through an Oriel f 3.5 Monochromator and is imaged on the condenser in a Koehler configuration. A beam splitter sends a portion of the monochromator light to a photomultiplier tube (Hammamatsu R453). The photolysis beam and the source for light scattering is a Spectra Physics 164-06 argon ion laser, emitting at 488 nm. When shutter S1 is open, this beam is reflected into the primary optical train by a dichroic beam splitter M3, which passes the arc lamp light in the region <450 nm and reflects the 488-nm laser line. At the position typically occupied by the field diaphragm is placed a steel mesh (Buckbee-Meers), the image of which shown in Fig. 1, inset *A*. This mesh is imaged at the sample plane as well, where the spot size of the holes is 6.7 μm radius. Beyond the eyepiece, these spots are brought to a common focal spot, at which point a glass coverslip with a small painted dot is inserted. This dot blocks the direct beam and allows rays scattering at larger angles from all spots to pass to the detector. Thus the detector forms an image of only the scattered light at the different spots photolyzed by the laser beam. To reduce the overall intensity striking the CCD, polarizers are inserted between the objective and eyepiece.

For measurement of the COHb absorption, the shutter is closed, and the beam stop and polarizers are removed. Because it is important to measure the desaturation effected by the photolysis beam, the absorption spectrum was measured while the laser impinged on the sample. This experiment is difficult in that precise optical density as high as 1.00 must be recorded while rejecting a ~300-mW laser beam head on. This rejection was accomplished by inserting polarizers between the objective and eyepiece rotated fully to reject the polarized laser. The beam stop was replaced by a 488-nm optical notch filter (Kaiser Optical Systems), which passes all wavelengths except a narrow band around 488 nm. In addition to the notch

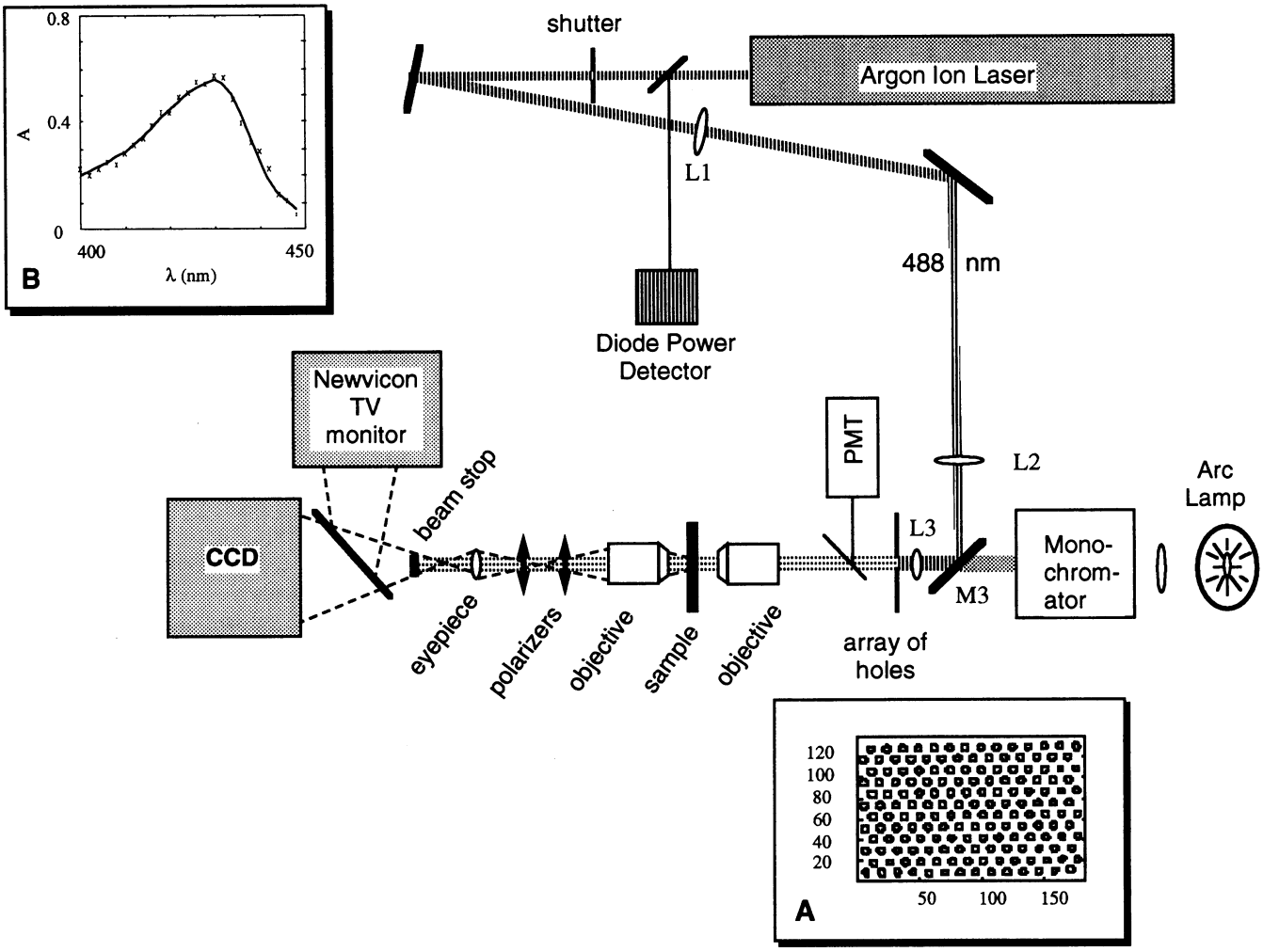

FIGURE 1 A schematic of the apparatus. The sample is observed by a 10× Leitz LWD strain-free objective, and its image is projected onto a CCD camera and a TV monitor. For light scattering measurements, the beam of an argon ion laser is sent through an array of holes placed where a field diaphragm is conventionally located. These holes are imaged at the sample. To reject the direct laser light, polarizers are placed after the collecting objective. In addition, a beam stop is placed at a position after the eyepiece, which is conjugate to the back aperture of the collecting objective. Thus light scattered at a greater angle than that of the laser passes the stop and is imaged at the CCD. M3 is a partially reflecting mirror that passes light in the Soret band region (400–450 nm) while reflecting the laser line at 488 nm. For absorbance measurements the beam stop may be removed. In some absorbance measurements, a pair of filters is also placed before the CCD. For absorbance, the light is provided by a 150-W xenon lamp and passed through a monochromator. The reference photomultiplier (PMT) is used for measurement of absorbance spectra. A MacII fx computer controls the CCD data transfer, as well as the shutter for the laser, and the stepping of the monochromator. (*Inset A*) Image of the array of holes taken with the CCD camera. (*Inset B*) Absorbance spectrum taken with the laser on. Note that the sample has been converted to deoxyHb, and that an absorbance spectrum of good quality can be recorded, despite a laser beam with intensity that is orders of magnitude greater.

filter, a small horizontal stripe was added at the beam stop position. This masked a portion of the beam, but allowed the extent of the beam and overall optical spectrum to be determined. The mesh was removed for this experiment. The spectrum measured with the laser on is shown in Fig. 1, inset *B*. As can be seen, it is fit very well by the standard deoxyHb spectrum, verifying the deoxygenation of the sample. This agrees with expectations based on the previous experiments of Ferrone et al. (1985b). Power density at the sample is 7.1 kW/cm^2, well in excess of the 3.7 kW/cm^2 typically needed (Ferrone et al., 1985b). With the mesh in place, a check of desaturation at a single wavelength (430 nm) was employed, and this was done for each experiment. In that case, a narrow-band filter was employed. Only spots with complete desaturation were used in the analysis of the experimental results.

When the spots were imaged on the CCD, the image area of 512 × 512 pixels was broken into an array of 33 × 14 "superpixels." Each superpixel is a binned array of 17 × 28 pixels. The selection of the binning numbers is totally dependent on the sizes and arrangement of the holes on the mesh. Each optically imaged spot was thus a single superpixel, which alternated with empty superpixels. The empty elements were sorted and discarded. In a kinetic experiment, 200 to 500 frames were collected. To collect additional kinetic data on the same sample, the sample was displaced slightly by a micrometer, so that a new area was photolyzed, thereby relieving any concern that polymer melting or CO recombination might be incomplete at the start of the next kinetic experiment.

The experiment was controlled by a Macintosh MacII fx. This also generated the timing signals, as well as controlling the opening of the

shutter and the motion of the monochromator for spectral scanning experiments.

RESULTS

Data were collected at samples of concentrations from 3.92 to 4.89 mM and at temperatures ranging from 13 to 35°C. (Not all temperatures were used for all concentrations.) Typical data are shown in Fig. 2.

The first 10% of each progress curve was employed to analyze exponential growth, and was fit to the function $Y(t) = A \exp(Bt) - Y_o$. The exponential growth rate B is a particularly sensitive parameter and thus serves as a valuable control for comparison against the extensive data set of Ferrone et al. (1985b). Fig. 3 shows the concentration dependence of B at four different temperatures. As can be seen, the agreement is excellent. The experimental values are listed in Table 1.

Distributions of the tenth time (time to reach 1/10 of the maximum signal) were collected for each concentration and temperature. Fig. 4 shows a series of such distributions. As the exponential growth rate slowed, the width of the distribution increased, as expected. This can be demonstrated by calculating a dimensionless reduced width, defined as the ratio of the width of the distribution at half-height divided by the average value of the tenth time. Keeping the concentration fixed at 3.9 mM, the reduced width increased from 0.47 at 35°C to 1.05 at 18°C. Fixing the temperature at 20°C, the reduced width increased from 0.29 at 4.45 mM to 0.89 at 3.92 mM.

Szabo (1988) has developed an analytical formulation to describe the distribution of tenth times in such an experiment. The homogeneous nucleation rate is denoted by ζ, and B is the rate of exponential growth of the concentration of polymerized monomers. n is a parameter that describes the point at which observation of the domain is made, and which is equal to $\exp(Bt_o)$, where t_o is the tenth time measured in a volume with a large number of nucleations and no stochastic variation. Given these definitions, $T(t)$, the distribution of tenth times t, is given by

$$T(t) = \frac{Bn^{\zeta/B}}{\Gamma(\zeta/B)}(1 - e^{-Bt})^n e^{-\zeta t}, \quad (1)$$

in which Γ is a gamma function. The initial part of the distribution is kept small by the $1 - e^{-Bt}$ term. Once $t > B^{-1}$, the distribution becomes a decaying exponential whose decay constant is ζ, the rate of homogeneous nucleation.

The curves in Fig. 4 show the fit of the Eq. 1 to the distributions. The determination of B is obtained independently of the distribution by observing the exponential growth of the initial part of the progress curve. Two parameters are varied in the fits of Eq. 1 to the data: the homogeneous nucleation rate ζ and the threshold parameter n. It is evident that the fit is excellent.

Once the nucleation rate ζ is known, the rate constant for nucleation, f_o, measured in mM/s, can be determined by dividing ζ by the photolyzed volume V and Avogadro's number N_o. Thus

$$f_o = \zeta/N_o V. \quad (2)$$

Nucleation volume is determined from the product of the photolytic areas, directly measured, times the thickness of the sample, determined by sample absorbance. Volumes are shown in Table 1. Nucleation rates determined for all of the data are shown in Table 1 as well.

Fig. 5 shows the concentration dependence of the nucleation rate at different temperatures. The error bars are dominated by the variation in photolyzed volume. The observed concentration dependence includes a linear dependence on initial concentration. However, it is evident that a degree of curvature can also be accommodated.

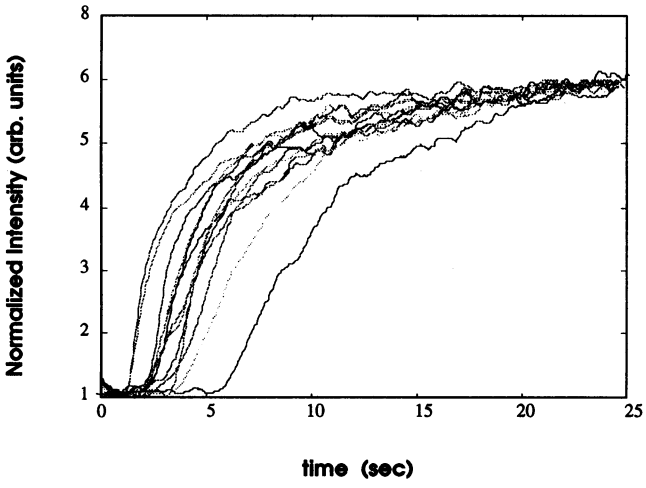

FIGURE 2 Typical scattering data. The initial concentration was 3.92 mM, and the temperature was 35°C. Scattering intensity is shown as a function of time after photolysis has begun. Initial data points have all been normalized to unity. The scattering at the initial time is a background effect that serves to allow different laser intensities at the different spots to be normalized.

DISCUSSION

The double nucleation model first and foremost provides a conceptual framework to explain the observed phenomena of sickle hemoglobin polymerization. The observations of the unique features of this mechanism, such as stochastic nucleation, domain formation, and the direct observation of the heterogeneous process, serve to validate this framework. The quantitative formulation of this or any model is useful as a means of summarizing a large body of data concisely. Beyond that, however, a quantitative model provides insights into the fundamental interactions from which it is derived.

To reap these benefits it is necessary that the model accurately describe data under a variety of conditions. As described in the Introduction, the initial impetus for this work was the 100-fold discrepancy between the predicted homogeneous nucleation rates and the rates first measured

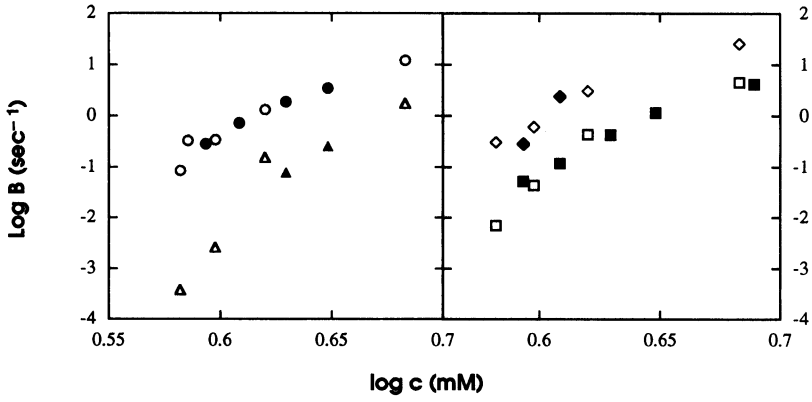

FIGURE 3 Log B versus log c for four different temperatures. B is the exponent in the exponential growth rate characteristic of the first 10–15% of polymerization; c is the initial concentration. Data obtained in this study (●, ▲, ◆, ■) is compared with the data of Ferrone et al. (1985b) (○, △, ◇, □) for 16°C (▲, △), 20°C (■, □), 25°C (●, ○), and 30°C (◆, ◇). The close agreement of these data is a sensitive control experiment for the comparison of the two data sets.

by stochastics (Hofrichter, 1986). Those rates are shown in Fig. 5 as the open circles, and as can be seen, they agree well with the larger data set shown here. In Fig. 5 the dashed line shows the prediction of the double nucleation model at 25°C, compared with the data collected here. It is apparent that the curves are simply shifted on the log scale and thus differ by a constant factor. Remarkably, the concentration dependence is correctly described, whereas the magnitude of the predicted rate is discrepant. We were thus led to reexamine the predictions of the double nucleation model.

The homogeneous nucleation rate is the product of an association rate and the activity of nuclei, divided by the activity coefficient of the activated complex. In this case, the activated complex can be considered as the aggregate one monomer larger than the nucleus. Thus if f_o is the nucleation rate,

$$f_o = k_+ \gamma c_o \gamma_{i*} c_{i*} / \gamma_{i*+1}. \quad (3)$$

TABLE 1 Exponential growth rate B and nucleation rate f_o for various conditions

c (mM)	T (°C)	V (μm)^3	log B (s^{-1})	log f_o (mM/s)
3.92	25	9.5	−0.539	−8.75
	30		0.023	−7.66
	35		0.240	−7.43
4.06	25	9.5	−0.137	−7.71
	30		0.382	−7.57
4.26	20	4.0	−0.361	−8.35
	25		0.192	−6.97
4.45	13.5	7.6	−0.577	−8.45
	16		−0.584	−8.27
	18.5		−0.180	−7.47
	20		0.061	−7.40
4.89	13.5	7.4	0.000	−6.87
	16		0.020	−6.97
	18.5		0.502	−6.24
	20		0.623	−6.14

Here the monomer addition rate constant is k_+, γ_o is the monomer activity coefficient at the initial monomer concentration c_o, i^* is the nucleus size, and c_{i^*} is the concentration of aggregates (nuclei) of size i^*. γ_{i^*} is the activity coefficient for the nucleus, and γ_{i^*+1} is the activity coefficient for the "activated complex," defined as one monomer larger than the nucleus. The activity coefficients depend both on the aggregate size i^* and initial monomer concentration c_o. The concentration of nuclei is assumed to be in

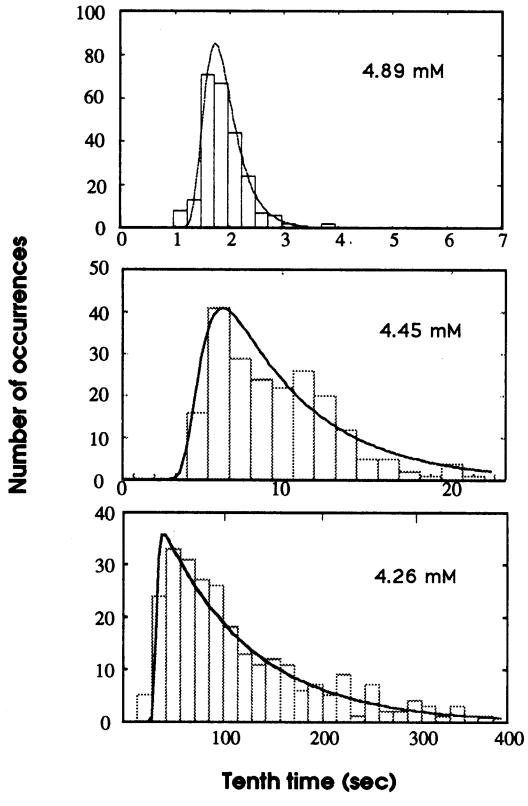

FIGURE 4 Distributions of tenth times for various concentrations as shown, at 20°C. Note the different time axes. The distributions have been fit by Eq. 1 using the parameters shown in Table 1. The homogeneous nucleation rate is determined by the exponential decay of the tail of the distribution.

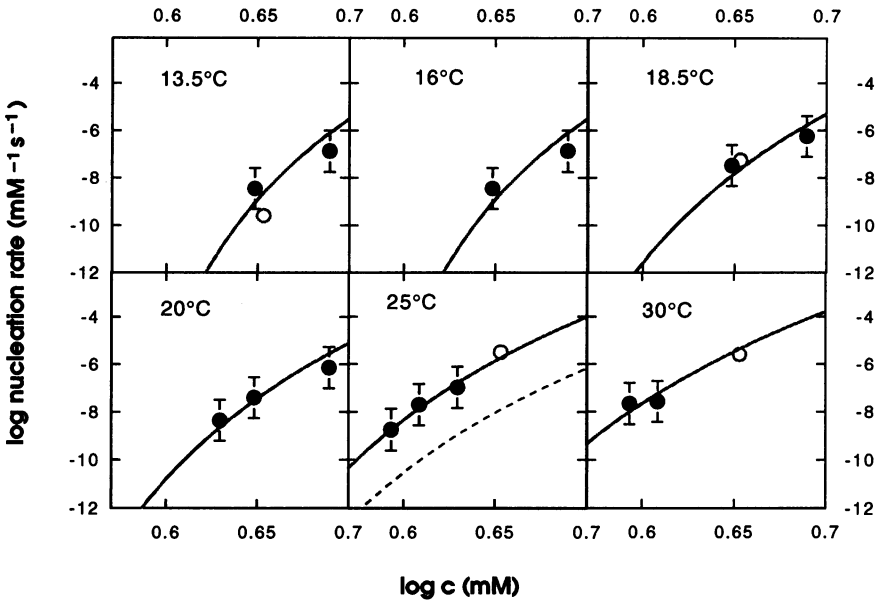

FIGURE 5 Nucleation rates f_o as a function of concentration for different temperatures as shown. Rates are obtained by fitting distributions similar to those shown in Fig. 4. The solid lines are the best fit obtained by varying a single parameter, the chemical potential of the monomer contacts, μ_{PC}. The chemical potentials so obtained are shown in Fig. 6. A single data point collected at 35°C is not shown. The open circles are the results obtained by Hofrichter (1986), using a serial technique, and have not been included in determining the best-fit line. The data of Hofrichter were taken at 14.3°C, 18.9°C, and 29.2°C, and the point at 25°C was interpolated in the original paper. It is a mark of the success of the nucleation theory employed that fits to one set of data extrapolate correctly to the other data set. The error bars in the data are primarily due to variations of the volume V in Eq. 2 due to variation of thickness across the sample and variation in imaged spot size.

equilibrium with the monomers, and so the activity of nuclei $(\gamma_i{\cdot}c_i{\cdot})$ can be formally written as $K_i{\cdot}(\gamma_o c_o)^{i*}$, where $K_i{\cdot}$ is the equilibrium constant for forming a nucleus of size i^*. In this equation, only k_+ is independent of the concentration of monomers. The concentration of nuclei, $c_i{\cdot}$, can be described with a single unknown parameter, the chemical potential of the contacts of a monomer inside a polymer, μ_{PC}, which represents the interaction energy between monomers in the polymer. Thus there are only two unknown, adjustable parameters in the description of f_o: the monomer addition rate constant k_+ and the chemical potential of contacts between monomers within a polymer, μ_{PC}.

These two parameters, k_+ and μ_{PC}, were determined originally as part of a set of five parameters that were fit to the light-scattering data on which the double nucleation model was based (Ferrone et al., 1985a). Subsequent to the formulation of the double nucleation mechanism, the rate of polymer elongation was measured by use of differential interference contrast microscopy to give a direct and completely independent determination of k_+ (Samuel et al., 1990). This value differed from that determined in the fitting. However, because of interdependence of the fit parameters, it was not immediately clear whether the direct measurement could be incorporated into the fit of the original light-scattering data. As shown in the Appendix, with four varied parameters and k_+ set to the observed value, a very reasonable fit to the data can indeed be obtained. After fitting to the original 25°C scattering data for progress curves, the new value for μ_{PC} is -7.5 kcal/mol compared to the old value of -8.6 kcal/mol. It is intrinsically a negative quantity because it represents the bonding of molecules into a polymer. With these new parameters, the prediction of the homogeneous nucleation rate f_o changes as well.

The prediction of f_o given these changes is drawn as the solid line through the 25°C data in Fig. 5. As can be seen, the agreement is excellent. This means that nucleation rates measured by stochastic methods agree with those determined by parameterizing the time course of light scattering, and argues that the quantitative formulation of homogeneous nucleation is sound. In turn, this implies that fitting the stochastic data will give an accurate determination of the polymer contact energy for other temperatures. By extrapolating the measured values of k_+ (using the observed activation energy of 13.9 kcal/mol) to obtain monomer addition rates at all temperatures used in this study, it is possible to fit the log f_o versus log c curves to determine $\mu_{PC}(T)$, which is the only variable parameter for determination of the homogeneous nucleation rate. The solid curves in Fig. 5 show the various fits obtained by changing the value of μ_{PC}. As can be seen, the fits are all quite good, especially given that only a single parameter is varied from one to another. Moreover, the theory successfully connects data collected by this technique and that collected by the prior serial stochastic method (Hofrichter, 1986). Because those data were not used in the fits, the agreement between the theory fit to this data and that of Hofrichter at other concentrations must be considered a success for the model.

As shown in Fig. 6, the polymer contact energy μ_{PC} varies with temperature, becoming more negative (i.e., showing greater stability) at higher temperatures. This is somewhat similar to the behavior of the free energy of association calculated from the solubility as $RT \ln \gamma_s c_s$. It is useful to distinguish between the two free energies. This is done in the framework of a two-component, two-phase model of assembly, which treats the process as equivalent to crystallization (Ferrone et al., 1985a). To perform that analysis, the chemical potential of a monomer in solution is set equal to that of a monomer in the aggregate, with the latter treated like a crystal phase. In solution the chemical potential is the sum of the translational and rotational free energy, μ_{TR}, and the solubility term $RT \ln \gamma_s c_s$. The aggregate contains two terms: energy from the "bonds" or contacts,

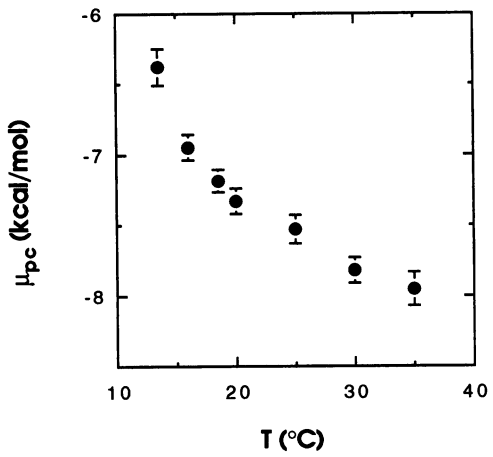

FIGURE 6 Chemical potential of the contact energy of monomers within a polymer for different temperatures. μ_{PC} is obtained by fitting nucleation rate equations to the rates shown in Fig. 5. Error bars include the uncertainty propagated from the extrapolation of k_+ to temperatures not measured in the original data of Samuel et al. (1990).

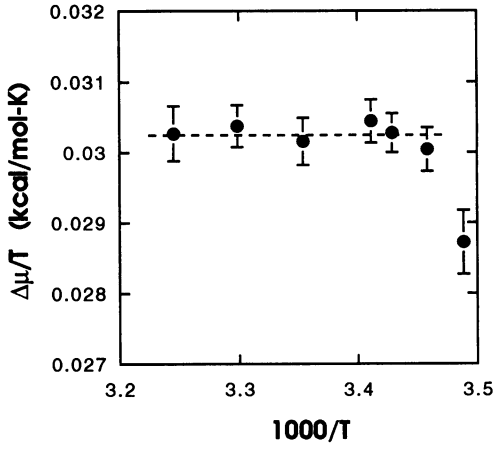

FIGURE 7 A plot of the difference in chemical potential between polymer contact energy and energy of association, divided by temperature, as a function of reciprocal temperature. The difference in chemical potential $\Delta\mu$ is defined as $\mu_{PC} - RT \ln \gamma_s c_s$. The slope of such a plot gives the difference in the enthalpy between contacts and association energy. The enthalpy difference is consistent with zero, with the exception of the lowest temperature point. The enthalpy difference is expected to be zero, because all of the enthalpic contributions computed from the solubility should arise from the contact energy, with the exception of a small motional term that has the same value in the polymer and monomer phases. Whereas $RT \ln \gamma_s c_s$ is a true equilibrium measurement, μ_{PC} is determined from kinetic measurements using equilibrium nucleation theory. The fact that there is no difference between the enthalpies calculated in these different ways argues that the kinetics are well described by equilibrium nucleation theory.

μ_{PC}, and the chemical potential from the vibrations of monomers within the aggregate, μ_{PV}. Thus

$$\mu_{TR} + RT \ln \gamma_s c_s = \mu_{PC} + \mu_{PV}. \quad (4)$$

μ_{PV} and μ_{TR} both arise from motional freedom, in solution or in the polymer.

The enthalpic part of μ_{PC} should be reflected in $RT \ln \gamma_s c_s$. The enthalpy per mole of monomer in solution or incorporated into the polymer is the same for the μ_{TR} or μ_{PV} terms, having a value of $6RT$. The enthalpy per mole of monomer associated is defined as $\Delta h = \Delta H/N$ and may be determined from the temperature dependence of the free energy essentially as the van't Hoff enthalpy, viz. (Hill, 1986),

$$\Delta h = \Delta H/N = \partial(\mu/T)/\partial(1/T). \quad (5)$$

To explore the similarity between Δh_{PC} (the enthalpy of the polymer contacts) and Δh_s (the enthalpy of the assembly calculated from the solubility), the difference

$$\Delta\mu = RT \ln \gamma_s c_s - \mu_{PC} \quad (6)$$

was computed and plotted as $\Delta\mu/T$ versus $1/T$. μ_{PC} was taken from this data set, and c_s was determined from the data of Ross et al. (1977). The slope of a plot of $\Delta\mu/T$ versus $1/T$ is the difference in enthalpies $\Delta h_{PC} - \Delta h_s$. Fig. 7 shows this plot; with the exception of the lowest temperature point, the plot shows a flat line whose zero slope indicates no enthalpy difference. (The lowest temperature point is also the one point at which these data and the data of Hofrichter (1986) differ, as seen in Fig. 5. Using the previous value at the point of discrepancy would remove this one discrepant point.) Thus there is no difference between the enthalpic contributions of the polymer contacts and the enthalpy calculated from the solubility.

Although this conclusion does not provide new quantitative information, because the enthalpy from solubility, Δh_s, is already known, it is an important confirmation that this analysis of kinetics by equilibrium nucleation theory can provide thermodynamic information. The chemical potential μ_{PC} is the result of the fit of the nucleation rates to the theory described, and the solubility is determined by entirely different means. The agreement of the enthalpy determined by entirely different methods argues that μ_{PC} has been correctly determined.

The chemical potential μ_{PC} describes the "glue" that holds the polymer together. The accurate measurement of this fundamental quantity is one of the principal quantitative findings of this work. Because this has been separated from other terms, it can be compared with standard values for the expected interactions. A considerable fraction of the contact energy is expected to be the hydrophobic interaction that favors inserting one of the β6 val into the pocket formed by β73, β85, and β88. To estimate the hydrophobic interaction one would like to know the buried surface area. The β6 val is mostly solvent exposed and is expected to account for ~1.5 kcal/mol by calibrations with model compounds (Nozaki and Tanford, 1971). Only parts of the β73, β85, and β88 groups are exposed. However, given the complementarity of the receptor region to the β6val surface, the receptor region must contribute an approximately equivalent buried surface and hence hydrophobic free energy. Although the second β6 is not grouped with the same receptor region, it is reasonable to assume that that it too

buries a hydrophobic region, also equal in area to the β6 exposed area. Therefore, the total free energy due to hydrophobic interactions must be about 4 × 1.5 kcal/mol = 6.0 kcal/mol. This would account for ~80% of the contact energy.

The free energy of association is customarily determined by the solubility as $RT \ln \gamma_s c_s$, an energy that is significantly smaller than the contact energy μ_{PC} described above. This results from the interplay between motional freedom in solution that is lost upon association and only partially regained by vibrational motions of the molecules within the polymer or aggregate. The recovery of lost solution entropy by motion within a protein aggregate was suggested as long ago as 1963 by Steinberg and Scheraga, and the general concept was employed in droplet condensation theories well before that (cf. Abraham, 1974). Although a number of approaches have been adopted to estimate this quantity (Horton and Lewis, 1992; Tidor and Karplus, 1994), the present method is one of the very few ways in which experimental data are directly brought to bear on this important issue. The fundamental reason that the present work can experimentally differentiate between these effects is that aggregates of vastly different size are being characterized. The concentration of small nuclei is determined because the nucleation rate depends on their concentration, whereas the concentration of monomers in infinite polymers is determined from equilibrium quantities such as the solubility. The vibrational modes depend on the number of molecules in the aggregate, whereas the contact energy depends on the number of bonds in the aggregate. Small aggregates differ from large ones in the number of bonds per molecule, and thus the energy terms μ_{PC} and μ_{PV} scale in different and distinguishable ways.

Rotational and translational chemical potentials account for some 36 kcal/mol, which is forsaken upon assembly. Seventy-five percent of this is recovered in vibrational energy, as shown in Fig. 8. The linear trend of the chemical potential is striking. (The lowest temperature data point is an exception, and the enthalpic contributions also differed from the rest of the data for this point, as noted above.) This linear behavior, however, is not unexpected. In any solid in which the vibrational modes are all excited, the general form of the chemical potential will be $-RT(\ln RT$ + constants) (Hill, 1986). Given the small range of absolute temperature explored here, the foregoing simplifies to an equation linear in T. A simple model of such a solid is the Einstein lattice model, in which all molecules in a solid are assumed to oscillate in the same potential given by the average potential with all of the other molecules fixed. The Einstein lattice predictions, adjusted to match the data at one point, are also drawn in the figure. However, other models for solids (e.g., the Debye model) will necessarily adopt the same behavior over the range employed. Although discrimination between different models is not possible with the temperature range employed here, the temperature dependence of the vibrational free energy argues for the correctness of the overall formulation, because the different temperature data are in no way constrained by the models or the analysis to show such linearity in the temperature.

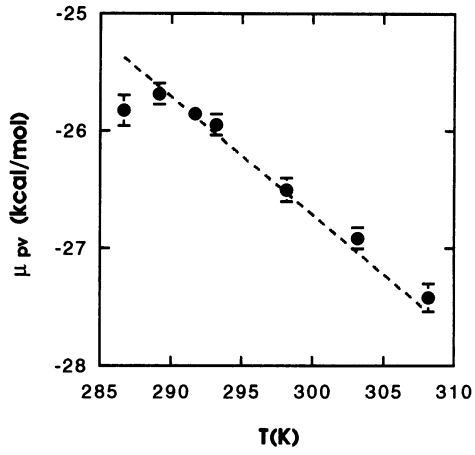

FIGURE 8 Vibrational chemical potential μ_{PV} for monomers within the polymer, as a function of temperature. This chemical potential is determined by using Eq. 4. The vibrational chemical potential represents the redemption of some of the 36 kcal/mol of rotational and translational free energy lost upon association. With the exception of the lowest temperature point, the temperature dependence of the curve agrees with expectations of an Einstein model, or any model in which all modes of vibration are excited at the temperatures of observation.

Because the enthalpic contributions to the contact chemical potential are known, it is possible to deduce the entropic contributions to that chemical potential. The entropic energy, $-T\Delta S$ is shown in Fig. 9 as a function of temperature. This arises from the entropy at the contact region itself, rather than the overall entropy discussed above. Because the enthalpy is positive, the entropic contribution to the chem-

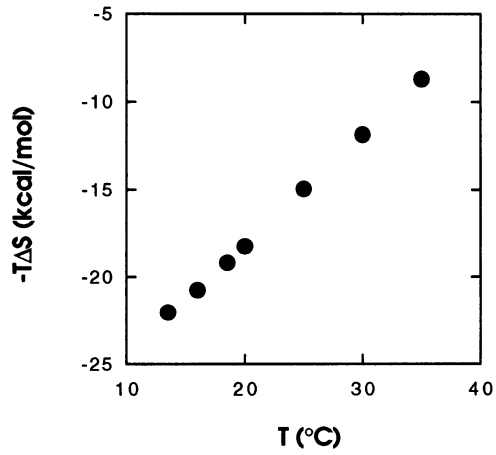

FIGURE 9 Entropic contribution to the free energy of the intrapolymer contacts as a function of temperature. The enthalpy was calculated from the solubility. As shown above, the enthalpy for overall association and the enthalpic part of the chemical potential μ_{PC} are the same. From μ_{PC} the enthalpy was subtracted to give the free energy due to entropic contributions, $-T\Delta s$. This allowed the entropy Δs to be determined.

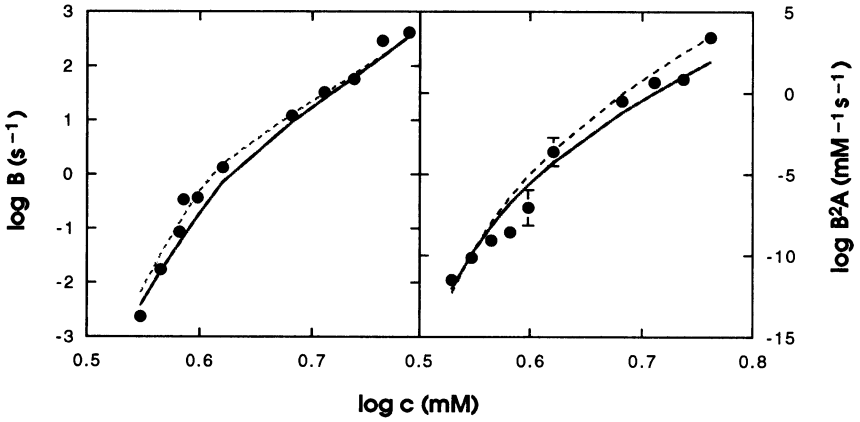

FIGURE 10 Fit of B and B^2A. The data are those of Ferrone et al. (1985a,b), and the dashed line shows the original fit to those data. When the monomer addition rate constant is fixed, B^2A depends on one parameter, μ_{PC}, and B on three parameters in addition to μ_{PC}. Thus the fit shown by the solid lines was obtained by variation of μ_{PC} to fit B^2A, and then by variation of the remaining three parameters to optimize the fit to B. The original fit was carried out by simultaneous fitting of all five parameters to both data sets.

ical potential μ_{PC} is greater than μ_{PC} itself, being in the range of -11 kcal/mol. A remarkable series of cancellations is thus revealed. The modest overall association energy of $\sim$1.5 kcal/mol arises from a contact energy of -7.5 kcal/mol plus entropic recovery of -26.5 kcal/mol, competing with a loss of solution freedom amounting to -35.5 kcal/mol. The -7.5 kcal/mol of contact free energy in turn decomposes into $+7.55$ kcal of enthalpy and -15.0 kcal/mol from the entropic contributions.

Finally, we note that the method described here could be used for any system that is capable of photolytic initiation, even if the process is not reversible, in contrast to the procedures employed in serial stochastic measurements. Given the increasing availability of caged compounds with photochemical triggers, such a method could probe other reactions that involve a millisecond-scale molecular process as a trigger for a readily observable chain of events.

APPENDIX

The polymerization of sickle hemoglobin has been characterized by observing light that is scattered as polymers form. During the initial 10–20% of the reaction, the light scattering is assumed to be proportional to the concentration of monomers polymerized, denoted as Δ. The differential equations describing the double nucleation model can be linearized, and the concentration of polymerized monomers can be shown to have the form

$$\Delta = A[\cosh Bt - 1], \qquad \text{(A1)}$$

in which A and B are related to the molecular parameters of the description (Bishop and Ferrone, 1984). As Bt becomes large, Eq. A1 is approximately given by

$$\Delta = A/2 \exp(Bt). \qquad \text{(A2)}$$

Although molecular parameters of both homogeneous and heterogeneous nucleation contribute to both A and B, the product B^2A is independent of the details of the description of the heterogeneous process as long as the rate of that process is proportional to Δ. Therefore, the parameterization of progress curves involved fitting Eq. A1 to the light scattering data, and then fitting a detailed model to the concentration and temperature dependence of B and B^2A (Bishop and Ferrone, 1984; Ferrone et al., 1985a,b). Five parameters were used to describe the concentration-dependent values of B and B^2A at a given temperature. B^2A is controlled by only two of these parameters: the monomer addition rate constant k_+ and the contact energy μ_{PC}. B is dependent on these plus three additional parameters that describe the magnitude and size dependence of the energy needed to attach a heterogeneous nucleus to the polymer.

Initially a simultaneous fit was carried out in which all five parameters were varied to provide the best overall fit (Ferrone et al., 1985a). Because k_+ has been determined by DIC microscopy (Samuel et al., 1990), the fit can be limited to four parameters, with only one of the parameters variable for B^2A. Therefore, we undertook a refitting by first finding the best value of μ_{PC} to fit the B^2A data, and then varying the three remaining parameters to fit the data of log B versus log c. Fig. 10 shows the result of this fitting procedure. The fit remains acceptable and illustrates the ability of a nonlinear fit to trade errors in one region of a curve for those in another.

We thank Dr. Qun Dou for his assistance in developing the software for the video interface employed in these studies, and Dr. Michael Cho for assistance in preparation of the samples.

REFERENCES

Abraham, F. F. 1974. Homogeneous Nucleation Theory. New York: Academic Press.

Bishop, M. F., and F. A. Ferrone. 1984. Kinetics of nucleation controlled polymerization: a perturbation treatment for use with a secondary pathway. *Biophys. J.* 46:631–644.

Cao, Z., and F. A. Ferrone. 1996. A 50th order reaction predicted and observed for sickle hemoglobin nucleation. *J. Mol. Biol.* 256:219–222.

Cho, M. R., and F. A. Ferrone. 1991. Monomer diffusion and polymer alignment in domains of sickle hemoglobin. *Biophys. J.* 63:205–214.

Ferrone, F. A., J. Hofrichter, and W. A. Eaton. 1985a. Kinetics of sickle hemoglobin polymerization. II. A double nucleation mechanism. *J. Mol. Biol.* 183:611–631.

Ferrone, F. A., J. Hofrichter, and W. A. Eaton. 1985b. Kinetics of sickle hemoglobin polymerization. I. Studies using temperature-jump and laser photolysis techniques. *J. Mol. Biol.* 183:591–610.

Ferrone, F. A., J. Hofrichter, H. Sunshine, and W. A. Eaton. 1980. Kinetic studies on photolysis-induced gelation of sickle cell hemoglobin suggest a new mechanism. *Biophys. J.* 32:361–377.

Hill, T. L. 1986. An Introduction to Statistical Thermodynamics. Dover Publications, New York.

Hofrichter, J. 1986. Kinetics of sickle hemoglobin polymerization. III. Nucleation rates determined from stochastic fluctuations in polymerization progress curves. *J. Mol. Biol.* 189:553–571.

Horton, N., and M. Lewis. 1992. Calculation of the free energy of association for protein complexes. *Protein Sci.* 1:169–181.

Nozaki, Y., and C. Tanford. 1971. The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. *J. Biol. Chem.* 246:2211–2217.

Ross, P. D., J. Hofrichter, and W. A. Eaton. 1977. Thermodynamics of gelation of sickle cell deoxyhemoglobin. *J. Mol. Biol.* 115:111–134.

Samuel, R. E., E. D. Salmon, and R. W. Briehl. 1990. Nucleation and growth of fibres and gel formation in sickle cell haemoglobin. *Nature.* 345:833–835.

Steinberg, I. Z., and H. A. Scheraga. 1963. Entropy changes accompanying association reactions of proteins. *J. Biol. Chem.* 238:172–181.

Szabo, A. 1988. Fluctuations in the polymerization of sickle hemoglobin: a simple analytical model. *J. Mol. Biol.* 199:539–542.

Tidor, B., and M. Karplus. 1994. The contribution of vibrational entropy to molecular association. *J. Mol. Biol.* 238:404–414.