Heterogeneous Nucleation and Crowding in Sickle Hemoglobin: An Analytic Approach

Frank A. Ferrone, Maria Ivanova, and Ravi Jasuja
Department of Physics, Drexel University, Philadelphia, Pennsylvania 19104 USA

ABSTRACT Sickle hemoglobin nucleation occurs in solution as a homogeneous process or on existing polymers in a heterogeneous process. We have developed an analytic formulation to describe the solution crowding and large nonideality that affects the heterogeneous nucleation of sickle hemoglobin by using convex particle theory. The formulation successfully fits the concentration and temperature dependence of the heterogeneous nucleation process over 14 orders of magnitude. Unlike previous approaches, however, the new formulation can also accurately describe the effects of adding nonpolymerizing agents to the solution. Without additional adjustable parameters, the model now describes the data of M. Ivanova, R. Jasuja, S. Kwong, R. W. Briehl, and F. A. Ferrone, (*Biophys. J.* 2000, 79:1016–1022), in which up to 50% of the sickle hemoglobin is substituted by cross-linked hemoglobin A, which does not polymerize, and which substitution causes the rates to decrease by 10^5. The success of this approach provides insight into the polymerization process: from the size-dependence of the contact energy deduced here, it also appears that various contacts of unknown origin are energetically significant in the heterogeneous nucleation process.

INTRODUCTION

The nucleation of sickle hemoglobin is a remarkable process involving, in the same reaction, a very strong concentration dependence (up to 50th power) and an exponential time dependence. Reconciliation of these two disparate characteristics was made possible by the postulation of a novel, double nucleation mechanism in which it was proposed that the surface of a polymer could catalyze the nucleation of additional polymers (Ferrone et al., 1980; Ferrone et al., 1985b) (Fig. 1). The formation of polymers thus could proceed either by nucleation in bulk solution (homogeneous nucleation) or by polymer nucleation on other polymers (heterogeneous nucleation). Originally deduced on the basic of kinetic measurements, the heterogeneous nucleation process was subsequently observed directly in differential interference contrast microscopy (Samuel et al., 1990). Recently, a structural rationale for the process has been advanced using the same primary contact partners found in the polymer itself., viz. β6 val seating into a hydrophobic pocket around β88 Leu (Mirchev and Ferrone, 1997).

The assembly of sickle hemoglobin requires high hemoglobin concentrations (>160 mg/ml) in vivo and in phosphate buffers of low molarity (Eaton and Hofrichter, 1990). Consequently, the description of the assembly process in thermodynamically sound terms must account for substantial nonideality in the solutions. The mutual crowding of monomers has been well studied, and accurate monomer activity coefficients are available, having been determined in various colligative experiments (Ross and Minton, 1977; Ross et al., 1977, 1978; Prouty et al., 1985). Among other uses, the activity coefficients have successfully rationalized the copolymerization of sickle hemoglobin in mixtures with nonpolymerizing agents (Minton, 1977; Eaton and Hofrichter, 1990). Describing the formation of homogeneous nuclei is not as simple as describing solubility, but nonetheless can be modeled by an approach known as scaled particle theory (SPT), in which activity coefficients can be determined for objects of similar geometry (e.g., spheres) (Ferrone et al., 1985a). The adequacy of SPT modeling of homogeneous nucleation was recently tested, and it was found to be highly successful (Ivanova et al., 2000). Heterogeneous nucleation poses a still more difficult challenge because the nucleus forms on a surface, so scaled particle approaches are inad-

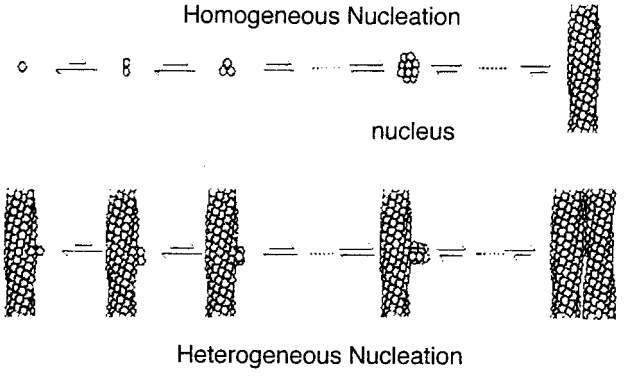

FIGURE 1 The double nucleation mechanism. The first polymer forms by homogeneous nucleation, followed by elongation. Subsequent polymers may either form by additional homogeneous nucleation events or by nucleation onto the surface of polymers that have already formed. Once formed by either pathway, the polymers are the same. Nuclei are assumed to have no special structure, but maintain the same geometry as monomers within the polymer.

Received for publication 23 January 2001 and in final form 28 August 2001.

Address reprint requests to Frank A. Ferrone, Drexel University, Dept. of Physics, 32 & Chestnut St., Philadelphia, PA 19104. Tel.: 215-895-2778; Fax: 215-895-5934; E-mail: fferrone@drexel.edu.

0006-3495/02/01/399/08 $2.00

equate. Although lattice approaches that model the aggregates as formed from cubes, for example, are numerically workable, they do not provide the convenience of an analytic form (cf. for example, Madden and Herzfeld, 1993). Having an analytic description is especially useful because the equations must be incorporated into a kinetic model and then fit to data. The original approach to the problem of describing the activity coefficients of the heterogeneous nucleus was to assume that the activity coefficient of the polymer equaled the activity coefficient of the polymer plus nucleus (or at least were proportional in a concentration-independent ratio) (Ferrone et al., 1985a). Recent tests, however, have shown that such a fortuitous cancellation did not occur, and, unlike the case for homogeneous nucleation, the quantitative description of the effect of crowding on heterogeneous nucleation rates was quite inaccurate (Ivanova et al., 2000).

Such a deficiency has various implications. The models used to describe the kinetics use the energetics of attachment of nuclei. Without accuracy in the model, it is impossible to know if the energetics deduced for the heterogeneous nucleation have meaning. This is especially important because contacts contribute to the stiffness of the gel, which in turn leads to the pathophysiology. Likewise, a systematic study of the contact energies is frustrated until they can be unambiguously identified. It should also be noted that, not only does sickle hemoglobin assemble in a solution crowded by deoxyHbS molecules, but it is also crowded by the presence of molecules that do not polymerize, such as oxy-HbS or HbA.

In this paper, we report a new approach based on a variation of scaled particle theory that we label as convex particle theory (CPT) (Boublik, 1974). Unlike SPT, CPT can be used when the geometry of the particles is different, as long as the particles themselves are convex. We find that the formulation gives rise to a useful approximation that permits a simple analytic expression to be given for the relevant activity coefficient ratio, and requires no new undetermined parameters. With this new formulation, we have refit the extant data for the exponential growth process, which depends heavily on the rate of homogeneous nucleation. The new formulation provides a fit that is as good, or better than, that originally proposed by Ferrone et al. (1985a). Once the data is fit to obtain a standard set of parameters for temperatures between 15 and 35°C, the parameters and theory are used to predict the results expected from addition of crowding molecules, such as cross-linked HbA. (The presence of the cross-link prohibits the HbA from reproportionating subunits with HbS.) The model provides very good predictive power for the crowding of the solution, arguing that the approach taken is sound, and the approximations reasonable. From this vantage point, it becomes possible to use the heterogeneous nucleation parameters as a tool to understand the molecular energetics of this unique and critical process.

THEORY

Homogeneous nucleation

We begin with the description of homogeneous nucleation because heterogeneous nucleation involves the attachment of an aggregate to a surface of the polymer. The homogeneous nucleation rate f_0 is the result of monomer addition to a spontaneously formed nucleus. If nuclei of size i^* have concentration denoted by c_{i*} and possess activity coefficient γ_{i*} then the nucleation rate is given by

$$f_0 = k_+ \frac{\gamma_0 c_0 \gamma_{i*} c_{i*}}{\gamma^{\ddagger}}, \tag{1}$$

in which $\gamma^{\ddagger}$ is the activity coefficient of the activated complex (Hill, 1986; cf. Eqs. 2 and A2.1 of Ferrone et al., 1985a). k_+ is the monomer addition rate constant, taken as size independent. By straightforward statistical thermodynamics, c_{i*} can be related to μ_{PC}, the chemical potential that holds the molecules within the polymer, because the nucleus is approximated as being in equilibrium with the monomers. In the computation to determine the nucleus, the contributions of the contact energy, μ_{PC}, and the chemical potential from vibrational entropy μ_{PV}, must be considered. For an aggregate of size i, the chemical potential of the aggregate depends on the fraction of contact sites in the infinite polymer that have been made, viz., $\delta(i)$. The total contact energy for size i is given by $i\delta(i)\mu_{PC}$. A useful approximation is to set $\delta(i) = \delta_0 i + \delta_1 \ln i + \delta_2$ where the various constants are determined from fitting this function to the contacts determined from studying close-packed spheres (Ginnel, 1961). The total chemical potential from vibrations, in contrast, is taken to be just $(i - 1)\mu_{PV}$. (The prefactor is not simply i because the aggregate retains center-of-mass freedom, which removes 6 degrees of freedom, the equivalent of one less member in the aggregate.) The critical ingredient here is that the net vibrational chemical potential rises linearly with the addition of monomers to the aggregate. Because total contact energy and total vibrational chemical potential depend on size i differently, they are distinguishable in the fitting process.

For the sake of concise notation, the parameter ξ is used, which contains the unknown μ_{PC} and other constants specified by the geometry of the nucleation process and not varied in any fits. ξ is defined by

$$\xi = -(4 + \delta_1 \mu_{PC}/RT). \tag{2}$$

Then the homogeneous nucleation rate f_0 can be shown to be

$$f_0 = q k_+ \frac{\gamma_0 c_0 \gamma_s c_s}{\gamma^{\ddagger}} \left[\frac{\ln S}{\xi}\right]^{\xi} e^{1.12\xi}, \tag{3}$$

where q is a combination of other geometrically determined constants (cf. Ivanova et al., 2000), and S is the activity supersaturation at the initial concentration c_0, defined as

$$S = \gamma_0 c_0 / \gamma_s c_s, \tag{4}$$

in which c_s is the solubility, and γ_s is the activity coefficient at solubility. c_0 is the concentration of deoxyhemoglobin S at the initiation of the polymerization. γ_0 is the activity coefficient of the total hemoglobin concentration at initiation. The nucleus size, i^*, does not appear explicitly in this equation, but it is related in a very simple way to the parameters used, viz.,

$$i^* = \xi / \ln S. \tag{5}$$

The effects of crowding appear in two ways. First, they appear in the activity coefficient for the monomer, γ, which only depends on the total concentration of hemoglobin. This activity coefficient is known for different concentrations from both osmotic pressure and sedimentation experiments (Ross and Minton, 1977), and has been extensively used in the study of the thermodynamics of hemoglobin S (Eaton and Hofrichter, 1990).

Nonideality also affects $\gamma^\ddagger$, the activity coefficient for the activated complex, an aggregate of size $i^* + 1$. Taking both nucleus and monomer as roughly spherical allows the use of SPT to determine $\gamma^\ddagger$ (Minton, 1981). This depends on the nucleus size (i^*) and the concentration of hemoglobin c, but, given such input data, $\gamma^\ddagger$ is fully specified.

Heterogeneous nucleation

The rate of heterogeneous nucleation is proportional to the concentration of monomers already present in polymers, and may be written as $g_0\Delta$. Δ is the concentration of monomers incorporated into polymers. Analogous to Eq. 1, we may write

$$g_0\Delta = \frac{k_+ \gamma c \gamma'_{j^*} c'_{j^*}}{\gamma'_{j^*+1}}. \tag{6}$$

Here the primes indicate the concentration and activity coefficients of attached aggregates of size j^*. The activity coefficient of the attached aggregate necessarily includes the volume excluded by the polymer to which the aggregate is attached. The activity coefficient in the denominator is that of the activated complex, an attached aggregate of size $j^* + 1$, which again includes the polymer in the calculation. Although Eq. 6 is similar in form to Eq. 1, there are important differences. Because the heterogeneous nucleus consists of an aggregate attached to a polymer, the activated complex is no longer is a spherical object. The heterogeneous nucleus is also an attached aggregate, and is written with a prime to distinguish it from an aggregate like the homogeneous nucleus, but simply different in size.

The activity of an attached aggregate of size j^* to polymers is given by its equilibrium with the solution aggregates and the polmer sites to which it attaches, i.e.,

$$\gamma'_{j^*} c'_{j^*} = K'_{j^*} \gamma_{j^*} c_{j^*} \gamma_p \phi \Delta, \tag{7}$$

where primes indicate attached aggregates, and unprimed symbols indicate solution aggregates. K'_{j^*} is the equilibrium constant for the attachment process, and Δ is the concentration of monomers in polymers. ϕ was originally taken as the fraction of such polymerized monomers that can accept an aggregate (however, cf. below), and γ_p is the activity coefficient of the polymer with no aggregate attached. Then, Eq. 6 becomes

$$g_0\Delta = \frac{k_+ \gamma c K'_{j^*} \gamma_{j^*} c_{j^*} \gamma_p \phi \Delta}{\gamma'_{j^*+1}} = k_+ \gamma c \gamma_{j^*} c_{j^*} K'_{j^*} \phi \Gamma \Delta, \tag{8}$$

in which Γ is here defined as γ_p/γ'_{j^*+1} i.e., the activity coefficient for a polymer with no aggregate attached divided by the activity coefficient of a polymer with aggregate size $j^* + 1$ attached. Γ was previously taken to be unity, i.e., the activity coefficients were assumed to be equal (Ferrone et al., 1985a). This assumption has been conclusively shown to be inadequate (Ivanova, et al. 2000).

Although SPT cannot be used here for Γ, another approach to the calculation of the activity coefficients is that of Boublik (1974), which has been summarized by Minton (1998). This approach allows particles of dissimilar shapes to be compared (e.g., rods and spheres) as long as the particles are convex, and we shall refer to it here as CPT. Using Minton's notation, the nonideality for convex particles depends upon a length function H_n, a surface area S_n and a volume V_n. The subscript n simply denotes which class of particle is described, monomers, or some specific aggregate. Because we will approximate the solution as having monomers and aggregates with size j^* attached to a polymer, there are only two classes in this description. The length function H_n is the Kihara supporting function, equal to one-half the projection of the particle onto a single directional axis, averaged over all orientations of the particle relative to its axis. For example, for a sphere, H_n is simply the radius r; for a right-circular cylinder of length L and radius r, it is $(\pi r + L)/4$. Then the activity coefficient for the nth particle class is given by

$$\ln \gamma_n = -\ln(1 - \langle\langle V\rangle\rangle) + \frac{H_n\langle\langle S\rangle\rangle + S_n\langle\langle H\rangle\rangle + V_n\langle\langle 1\rangle\rangle}{1 - \langle\langle V\rangle\rangle} + \frac{H_n^2\langle\langle S\rangle\rangle^2 + 2V_n\langle\langle H\rangle\rangle\langle\langle S\rangle\rangle}{2(1 - \langle\langle V\rangle\rangle)^2} + \frac{V_n\langle\langle H^2\rangle\rangle\langle\langle S\rangle\rangle^2}{3(1 - \langle\langle V\rangle\rangle)^3}, \tag{9}$$

where the double bracket indicates a weighted average, viz., $\langle\langle X\rangle\rangle = \Sigma\ \rho_k X_k$, and $\langle\langle 1\rangle\rangle = \Sigma\ \rho_k$ and ρ_k is the number density of the kth species or class in the solution. The

quantity of interest here is $\ln \Gamma = \ln \gamma_p - \ln \gamma'_{j^*+1}$, the difference between logs of the activity coefficient of a polymer and of a polymer with a nucleus attached. Observing in Eq. 9 that the denominators are the same, (because all double bracketed terms are the same) the computation of $\ln \Gamma$ will involve determination of the differences in Kihara function, surface area and volume of a polymer with a nucleus minus a polymer. Determining the volume difference (ΔV) is straightforward, and is just the volume of the nucleus. The surface area difference (ΔV) is approximately the surface area of the attached nucleus, with perhaps a correction for the attachment area. This can be found given the volume. The projected, averaged length is essentially the same for the polymer with and without attached nucleus, and, moreover, the projected distance is averaged over orientations, suggesting that the Kihara function difference $\Delta H \sim 0$. Thus, we have

$$\ln \Gamma = \frac{\Delta S\langle\langle H\rangle\rangle + \Delta V\langle\langle 1\rangle\rangle}{1 - \langle\langle V\rangle\rangle} + \frac{2\Delta V\langle\langle H\rangle\rangle\langle\langle S\rangle\rangle}{2(1 - \langle\langle V\rangle\rangle)^2} + \frac{\Delta V\langle\langle H^2\rangle\rangle\langle\langle S\rangle\rangle^2}{3(1 - \langle\langle V\rangle\rangle)^3}. \quad (10)$$

Finally, because the number concentration of monomers is far greater than the concentrations of any other species, the sums over index k reduce to the contributions of the monomers. Let v be the volume of a monomer, and ρ' be the relative density, i.e., $\rho' = c/c_{pp}$, in which c is the concentration of pure monomers and c_{pp} is the concentration of hemoglobin in the polymer phase, taken here as 69 g/dl (Eaton and Hofrichter, 1990). Then ΔV, the volume of the nucleus of size j^*, is given by $\Delta V = j^* v \rho'$. From this expression and the assumption that the nucleus is spherical, we can derive an effective nuclear radius, and then determine the surface area ΔS. We then arrive at a simple expression for $\ln \Gamma$, viz.,

$$\ln \Gamma = -\frac{3j^{*2/3}(c_{pp}v)^{-2/3}vc}{1 - vc} - \frac{j^*c}{c_{pp}}\left[\frac{1 + vc + (vc)^2}{(1 - vc)^3}\right]. \quad (11)$$

This expression contains no new parameters, but uses known parameters in a new combination to account for nonideality of the heterogeneous nucleation process.

In the calculation of Γ, the size of the heterogeneous nucleus j^* is required. This is computed in a thermodynamic treatment like that used to compute the homogeneous nucleus. The differences in describing homogeneous and heterogeneous nuclei arise in that the latter must include an energy of attachment, but otherwise involves the same energetic considerations involved in computing homogeneous nucleation. The additional terms that appear in the calculation of the attachment are due to $K'_{j^*}\ \phi$ in Eq. 8. In energetic terms, the added stability of the heterogeneous nucleus is

$$-RT \ln K'_{j^*}\phi = -RT \ln \phi + \mu_{CC}\sigma_1 j^* + \mu_{CC}\sigma_2 \ln j^*, \quad (12)$$

where the terms have the following simple definitions: μ_{CC} is the chemical potential per unit contact area between polymers. The surface area in contact is broken up into linear and logarithmic contributions with coefficients σ_1 and σ_2. Equivalently, this represents the expansion of the energy in constant, linear, and log terms as was done for i^*. These are the unknown parameters, to be determined by fitting the double nucleation description to the data. It is possible that the best fit combination will yield parameters that display a maximum surface area in contact. In this case, for all nucleus sizes greater than the maximum, j^* is set to the maximum. The nucleus is determined by finding the turning point of the free energy versus size.

The parameter ϕ was originally taken to specify the fraction of surface molecules available. Physically, it is impossible from kinetic measurements to distinguish between a small number of sites (small ϕ) and a larger number of sites with weaker attachment energies.

Using the above definitions and solving for the nucleation turning point, we may then write, analogous to Eq. 3,

$$K'_{j^*}\gamma_{j^*}c_{j^*} = \left[\frac{\ln(S + \xi_1)}{\xi_2}\right]^{\xi_2} e^{\xi_2} \quad (13)$$

where

$$\xi_1 = -\frac{\sigma_1\mu_{CC}}{RT} \quad \text{and} \quad \xi_2 = \xi + 4 + \frac{\sigma_2\mu_{CC}}{RT}.$$

S is still the activity supersaturation as described above. ξ is also the same variable used in Eq. 2 derived from considering homogeneous nucleation. With this condensed notation, the heterogeneous nucleus j^* can be written as

$$j^* = \frac{\xi_2}{\ln S + \xi_1}. \quad (14)$$

In the previous derivation, we note one final assumption. Whereas the homogeneous nucleus retained its center of mass rotations and translations (that gave rise to the $i - 1$ rather than i in the vibrational chemical potential), the heterogeneous nucleus center-of-mass motion is intrinsically vibrational. To limit the number of free parameters, the frequency of that motion is taken as the same as the frequency of the monomers within the polymer, so that there is simply an additional chemical potential in the amount μ_{PV}. Any deviation of the center-of-mass vibrational frequency from the frequency of the average center-of-mass motion will appear with RT ln ϕ, from which it cannot be empirically split.

In fitting the experimental data, the three parameters that are operationally varied become ϕ, ξ_1, and ξ_2, which uniquely determine ϕ, $\mu_{CC}\sigma_1$, and $\mu_{CC}\sigma_2$. Once these are known, the equilibrium constant (through Eq. 12) is specified as well as the nucleus size j^*, which is used in the activity coefficient Γ.

DATA ANALYSIS

The equations that describe the double nucleation model in terms of its constituent rate and equilibrium constants can be linearized to describe the initial phase of polymer formation. When this is done, the concentration of monomers in polymers, denoted Δ, grows as

$$\Delta = A[\cosh Bt - 1] \sim A/2e^{Bt} \tag{15}$$

(Bishop and Ferrone, 1984), where the last approximation is true for large values of Bt. The exponential growth rate B is related to the heterogeneous nucleation rate g_0 by

$$B^2 = J_0[g_0 - df_0/dc], \tag{16}$$

where f_0 is the homogeneous nucleation rate, and J_0 is the net polymer growth rate. The subscripts zero indicate that these are evaluated at the initial concentration (c_0). Hence, the predictions of the heterogeneous nucleation model will be reflected in the rate B. The data for rates B have been collected at a variety of temperatures and concentrations (Ferrone et al., 1985b; Cao and Ferrone, 1997; Ivanova et al., 2000). B values cover a range of 10^3–10^{-4}, reflecting a range in heterogeneous nucleation rate g of $\sim 10^{14}$. To unify the data collected at different temperatures, we constructed an explicit temperature-dependent function for $\mu_{CC}\sigma_1$, $\mu_{CC}\sigma_2$, and ln ϕ. In this, it was assumed that the enthalpy had a linear temperature-dependent term in addition to a constant value. That is, each energetic term had a temperature dependence described by $\Delta H_{app}(T) = \Delta H + \Delta C_p T$ where ΔH and ΔC_p are temperature independent constants. Thus,

$$\mu_{CC}\sigma_1(T) = \mu_{CC}\sigma_1(T_0) + \Delta H_1[1 - T/T_0] + \Delta C_{p1} T \ln(T_0/T), \tag{17a}$$

$$\mu_{CC}\sigma_2(T) = \mu_{CC}\sigma_2(T_0) + \Delta H_2[1 - T/T_0] + \Delta C_{p2} T \ln(T_0/T), \tag{17b}$$

$$\ln \phi(T) = \ln \phi(T_0) + (\Delta H_\phi/R)[1/T - 1/T_0] + (\Delta C_{p\phi}/R)\ln(T_0/T). \tag{17c}$$

Here, ΔH is the temperature-independent enthalpy, and ΔC_p the heat capacity for the different contact energy terms as designated by their subscripts. T_0 is an arbitrary reference temperature, and R is the gas constant. Values of k_+ and ξ are taken from the measured temperature dependence of homogeneous nucleation (Cao and Ferrone, 1997). B values for four different temperatures were thus simultaneously fit to obtain the best nine parameters in Eq. 17 that described the concentration and temperature dependence of the B.

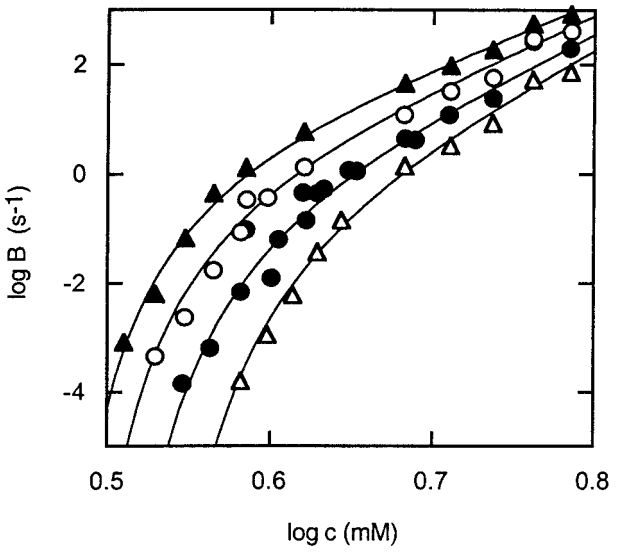

FIGURE 2 Exponential growth rate B as a function of the log of initial concentration in mM for four different temperatures: 15°C (*open triangles*), 20°C (*filled circles*), 25°C (*open circles*), and 35°C (*filled triangles*). Curves are the result of fitting the revised double nucleation model to the data, using the parameters shown in Table 1. Although in principle three parameters are required for each curve, the thermal analysis of Eq. 17 links parameters, so that 9 rather than 12 parameters have been varied in total. Because the heterogeneous nucleation rate varies as B^2, the heterogeneous nucleation rates cover ~14 orders of magnitude in this plot.

Figure 2 shows the measured rates for four different temperatures (15, 20, 25, and 35°C) for a range of concentrations, along with theoretical curves fit to the data. The data is a compilation of that published by Ferrone et al. (1985b), Cao and Ferrone (1997), and Ivanova et al. (2000). It is evident that this description provides an excellent representation of the data shown. The 51 measurements (over four temperatures) are described by variation of nine parameters. In general, the fit to each curve would have required variation of three parameters; by using Eq. 17, we have been able to reduce the number. The parameters obtained in the fits are shown in Table 1.

TABLE 1 Parameters describing heterogeneous nucleation

$\mu_{CC}\sigma_1$ (20°C)	=	0.068 kcal/mol
$\mu_{CC}\sigma_2$ (20°C)	=	−5.68 kcal/mol
ln ϕ (20°C)	=	−7.97
ΔH_1	=	−88.8 kcal/mol
ΔH_2	=	1059. kcal/mol
ΔH_ϕ	=	586.2 kcal/mol
ΔC_{p1}	=	0.284 kcal/mol-K
ΔC_{p2}	=	−3.49 kcal/mol-K
$\Delta C_{p\phi}$	=	1.98 kcal/mol-K

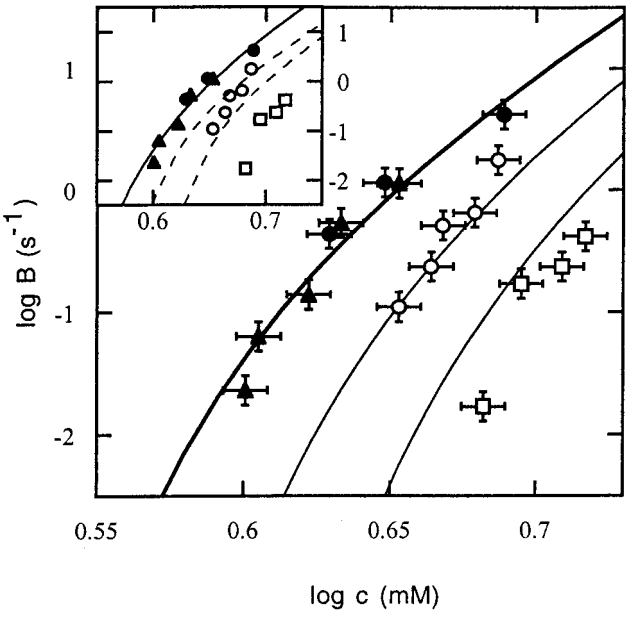

FIGURE 3 *B* as a function of log of initial concentration for different fractional substitutions, at 20°C. Filled symbols are collected for pure HbS, open circles are collected for 30% of the HbS replaced by cross-linked HbA, and open squares are collected for 50% of the HbS replaced by cross-linked HbA (Ivanova et al., 2000). Curves are drawn as predictions based on the fits shown in Fig. 2 (parameters of Table 1), with no further adjustment of parameters to match the theory to the data of 30% or 50% HbA. The purpose of the cross-link is to prevent reproportionation of the HbA into hybrids with the HbS. The inset shows the failure of the previous theory (Ferrone et al., 1985, Ivanova et al, 2000) to fit the data shown.

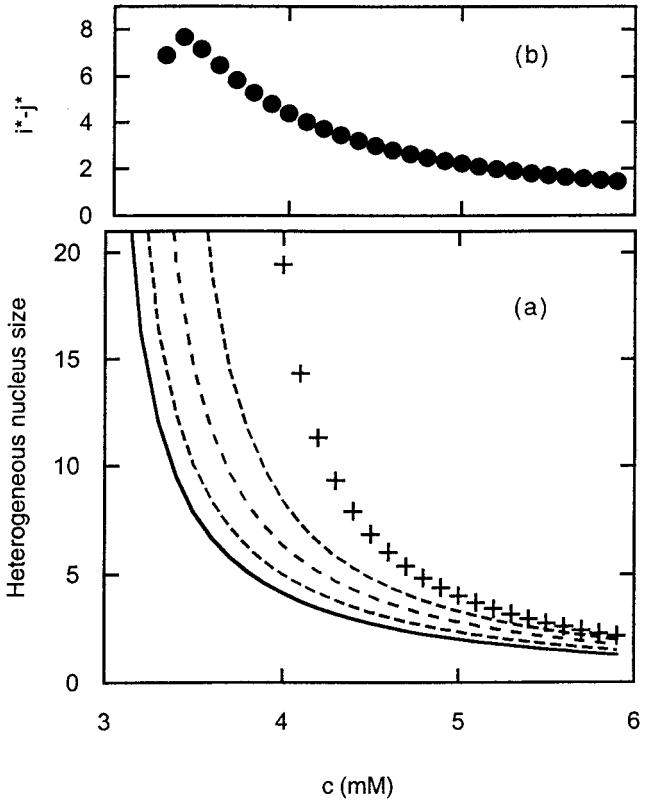

FIGURE 4 (*a*) Heterogeneous nucleus size (j^*) as a function of initial concentrations. Four temperatures are shown (15, 20, 25, and 35°C) as solid lines, with the smallest nuclei at the highest temperature. The + signs show j^* at 20°C with 50% of the HbS replaced by cross-linked HbA. (*b*) Difference in the homogeneous nucleus minus the heterogeneous nucleus ($i^* - j^*$). As expected, the heterogeneous nucleus is generally a few monomers smaller than the homogeneous nucleus.

The critical test of this model is to see whether the nonideality data of Ivanova et al. (2000) can be described once the parameters of the model have been determined in the absence of added crowding agents. The nonideality description developed above is thus used for solutions in which there is a 30% and 50% crowding by nonpolymerizing cross-linked HbA, with the results shown in Fig. 3. Although the predictions are not perfect, they are both quite close. Because the values of *B* vary as the square root of the heterogeneous nucleation rate, the variation in *B* seen here is indicative of a change in heterogeneous nucleation rate *g* of the order of 10^4. In the inset to Fig. 3, the lack of agreement from the previous model is shown for comparison.

DISCUSSION

The results obtained here show that the double nucleation mechanism correctly describes polymerization in crowded solutions, once the proper approach is taken to nonideality. Because heterogeneous nucleation is the least well understood part of the entire model, previous lack of success in fitting cross-linked mixture data might have also signaled some fundamental mistake in the description of the kinetic equations that governed the model. In turn, the present success is an important milestone in permitting the model to address such important issues as the polymerization in the presence of oxygen, or of HbF, which significantly ameliorates the disease. These topics will be dealt with in subsequent investigations. The new parameterization gives insights into the heterogeneous nucleation process. The size of the heterogeneous nucleus varies at 25°C from ~10 to 1.5 as the initial concentration is varied from 3.5 to 6 mM. These results and the values for other temperatures are shown in Fig. 4. A noninteger nucleus merely means that the turning point in the free energy versus size is between two integral aggregate sizes. Large nuclei are the exception rather than the rule, and, under physiological conditions of temperature and concentration, the nucleus ranges between ~1 and 5. The temperature dependence is weak but observable. When 50% of the hemoglobin is replaced by nonpolymerizing species (such as cross-linked HbA) the nucleus size rises, as shown by the crosses in Fig. 4. The heterogeneous nucleus, as expected, is generally smaller than the homogeneous nucleus because of the added stability con-

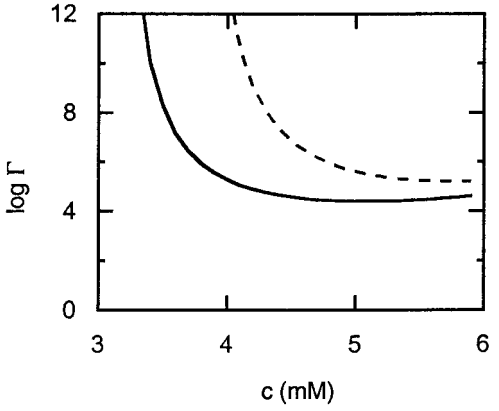

FIGURE 5 Activity coefficient ratio Γ as a function of initial concentration. The solid line shows the activity coefficient ratio for 20°C; the dashed line shows the same ratio but with 50% of the hemoglobin replaced by non-polymerizing HbA. Note the region of constancy in the activity coefficient, which is consistent with the prior assumptions in which Γ was taken as a constant.

ferred by the contact with the parent polymer. In Fig. 4 *b*, the difference between the homogeneous nucleus (i^*) and heterogeneous nucleus (j^*) is shown. Somewhat subtle vibrational effects cause the difference to decrease at high nucleus sizes. In general, though, the homogeneous nucleus is larger by 2 to as many as eight molecules.

Γ, the log of the activity ratio, is shown in Fig. 5 as a function of initial concentration. As can be seen, it is almost constant between 4 and 6 mM, which is not altogether surprising because the prior approximations of constant Γ provided a reasonable fit of the data. However, when 50% of the sickle hemoglobin is replaced by nonpolymerizing hemoglobin, the value of Γ rises, and the region of relative constancy diminishes. This is what accounts for the failure of the prior approach to fit the data containing nonpolymerizing species.

Perhaps the most interesting feature to arise from the revised analysis of polymerization is the strength of the interpolymer interactions and what it implies about the heterogeneous process. With the description used here, the contact energy as a function of size is given by Eq. 12 and is shown in Fig. 6 for four temperatures. As can be seen, there is very little net temperature effect on the attachment energy. The value of the energy for an aggregate of the size of a monomer might at first seem contradictory, being almost zero, whereas two monomers attach with ~ -4.5 kcal/mol. There are geometrical terms to consider, however. For example, if 4 of the 14 molecules had attachment sites, then the near zero of the first monomer is just a cancellation of a small intrinsic affinity by the rarity of a site for attachment to happen. In such a case, the actual energy of attachment at the site would become RT ln $4/14 = -0.75$ kcal/mol. As was mentioned above, the ln ϕ term was initially thought to be solely a geometrical term, but, on further consideration, must include any size-independent contribution to the stabilization energy.

In Fig. 5, the energy of the contacts of the monomer within the polymer, μ_{PC}, is shown for comparison. Within the polymer, one can distinguish three regions: lateral contacts that zig-zag up the constituent double strands and primarily involve the mutation site, axial contacts that occur between molecules displaced along the axis of the polymer, and inter-strand contacts that occur only between the double strands. Although the analysis of homogeneous nucleation gives a total contact energy $\mu_{PC} = -7.5$ kcal/mol (Cao and Ferrone, 1997) within the polymer, there has been no further breakdown of the total among the possible contact categories. The inter-double strand contacts vary considerably depending on the pairs considered, and, in general, model building has shown that these are not extensive (Watowich et al., 1989). Axial contacts are also found to be not very specific (Harrington et al., 1997), and it also may be that these are stretched in polymerization (Watowich et al., 1993). The strongest of the categories appears to be the lateral contacts, and it is these, after all, that are required to promote polymerization. Each monomer has two lateral contact regions, one donor and one acceptor, so the polymer attachment energy might be expected to be no greater than $-7.5/2 = -3.75$ kcal/mol. This would apply to each monomer in the aggregate. Although close, it is less than that found in the attachment energy. In addition, there is the issue of reconciling the size of the first monomer to attach to a polymer. If one were to require that the energy of the first monomer attachment be greater than the polymer-attachment energy of the second monomer, the number of heterogeneous sites drops to ~ 1 in 1000 monomers in a polymer.

The implication is that the heterogeneous site might arise from some type of polymer defect rather than a regular feature of the polymer geometry. Such a defect, in turn, might have a strong enough energy to bind the first or second monomer, but one which could not be sustained in the polymer due to structural incompatability.

It has been proposed that heterogeneous nucleation occurs because the β6 val-β receptor contact region that stabilizes the double strand also occurs on four external sites of the polymer (Mirchev and Ferrone, 1997). From that proposal, the geometrical part of ϕ must be $4/14$. Because this proposal invokes the same contact between polymers as found within them, the energetics governing polymer stability should be simply related to the contact energy attaching nucleus to polymer. The results of the present analysis do not appear to be consistent with the demands of such a model, most notably, in requiring a dimer to have almost four times greater energy of attachment to the polymer than the monomer does.

A new feature to emerge from the analysis is that there must be additional contact sites that are also energetically significant beyond the first contact site. As can be seen in

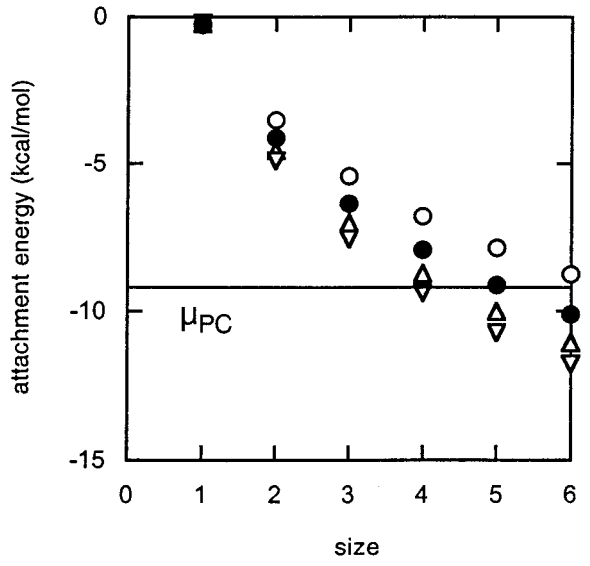

FIGURE 6 Free energy of attachment of an aggregate to its parent polymer as a function of size for four temperatures as computed using Eq. 12 and the parameters of Table 1. *Open circles*, 15°C; *filled circles*, 20°C; *triangles up*, 25°C; and *triangles down*, 35°C. The horizontal line shows the value of μ_{PC}, the contact energy of a monomer within the polymer. No individual step would be expected to have a greater contact energy than μ_{PC}. The infrequency of contact sites at which such attachment occurs will have the effect of offsetting the zero line.

Fig. 6, the stabilization continues to increase as the number of molecules in the aggregate increases. The ability of the model to generate energetic parameters provides a valuable tool to explore the nature of the heterogeneous process.

We especially thank Dr. Alan Minton for bringing the work by Boublik to our attention. We also thank Dr. Robin Briehl for critically reading the manuscript, and acknowledge the support of National Institutes of Health grants R01HL57549 and PO1HL58512.

REFERENCES

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.

Boublik, T. 1974. Statistical thermodynamics of convex molecule fluids. *Mol. Phys.* 27:1415–1427.

Cao, Z., and F. A. Ferrone. 1997. Homogeneous nucleation in sickle hemoglobin. Stochastic measurements with a parallel method. *Biophys. J.* 72:343–372.

Eaton, W. A., and J. Hofrichter. 1990. Sickle cell hemoglobin polymerization. *Adv. Protein Chem.* 40:63–280.

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.

Ginnel, R. 1961. Geometric basis of phase change. *J. Chem. Phys.* 34: 992–998.

Harrington, D. L., K. Adachi, and W. E. Royer, Jr. 1997. The high resolution crystal structure of deoxyhemoglobin S. *J. Mol. Biol.* 272: 398–407.

Hill, T. L. 1986. An Introduction to Statistical Thermodynamics. Dover Publications, New York. 194–198.

Ivanova, M., R. Jasuja, S. Kwong, R. W. Briehl, and F. A. Ferrone. 2000. Nonideality and the nucleation of sickle hemoglobin. *Biophys. J.* 79: 1016–1022.

Madden, T. L., and J. Herzfeld. 1993. Crowding-induced organization of cytoskeletal elements. I. Spontaneous demixing of cytosolic proteins and model filaments to form filament bundles. *Biophys. J.* 65:1147–1154.

Minton, A. P. 1977. Non-ideality and the thermodynamics of sickle-cell hemoglobin gelation. *J. Mol. Biol.* 100:519–542.

Minton, A. P. 1981. Excluded volume as a determinant of macromolecular structure and reactivity. *Biopolymers.* 20:2093–2120.

Minton, A. P. 1998. Molecular crowding: Analysis of effects of high concentrations of inert cosolutes on biochemical equilibria and rates in terms of volume exclusion. *In* Methods in Enzymology. G. K. A. a. M. L. Johnson, editor. Academic Press, San Diego, CA. 127–149.

Mirchev, R., and F. A. Ferrone. 1997. The structural origin of heterogeneous nucleation and polymer cross linking in sickle hemoglobin. *J. Mol. Biol.* 265:475–479.

Prouty, M. S., A. N. Schechter, and V. A. Parsegian. 1985. Chemical potential measurements of deoxyhemoglobin S polymerization. *J. Mol. Biol.* 184:517–528.

Ross, P. D., R. W. Briehl, and A. P. Minton. 1978. Temperature dependence of nonideality in concentrated solutions of hemoglobin. *Biopolymers.* 17:2285–2288.

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

Ross, P. D., and A. P. Minton. 1977. Analysis of non-ideal behavior in concentrated hemoglobin solutions. *J. Mol. Biol.* 112:437–452.

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.

Watowich, S. J., L. J. Gross, and R. Josephs. 1993. Analysis of the intermolecular contacts within sickle hemoglobin fibers: effect of site-specific substitutions, fiber pitch and double-strand disorder. *J. Struct. Biol.* 111:161–179.

Watowich, S. J., L. J. Gross, and R. J. Josephs. 1989. Intramolecular contacts within sickle hemoglobin fibers. *J. Mol. Biol.* 209:821–828.