Heterogeneous Nucleation in Sickle Hemoglobin: Experimental Validation of a Structural Mechanism

Maria A. Rotter,* Suzanna Kwong,[†] Robin W. Briehl,[†] and Frank A. Ferrone*

*Department of Physics, Drexel University, Philadelphia, Pennsylvania; and [†]Department of Physiology and Biophysics, and Department of Medicine, Albert Einstein College of Medicine, Bronx, New York

ABSTRACT   Sickle hemoglobin polymerizes by two types of nucleation: homogeneous nucleation of aggregates in solution, and heterogeneous nucleation on preexisting polymers. It has been proposed that the same contact that is made in the interior of the polymer between the mutant site $\beta 6$ and its receptor pocket on an adjacent molecule is the primary contact site for the heterogeneous nucleus. We have constructed cross-linked hybrid molecules in which one β-subunit is from HbA with Glu at $\beta 6$, and the other is from HbS with a Val at $\beta 6$. We measured solubility (using sedimentation) and polymerization kinetics (using laser photolysis) on cross-linked hybrids, and cross-linked HbS as controls. We find ~4000 times less heterogeneous nucleation in the cross-linked AS molecules than in cross-linked HbS, in strong confirmation of the proposal. In addition, changes in stability of the nucleus support a further proposal that more than one $\beta 6$ contact is involved in the homogeneous nucleus.

INTRODUCTION

Sickle hemoglobin is a genetic mutation of the oxygen-carrying tetramer, hemoglobin A (HbA). HbA consists of two α- and two β-chains, and in HbS the sixth position of each β-chain has a hydrophobic Val in place of the charged Glu. The structural consequence of this mutation is that when HbS loses oxygen, it can form long, rigid, 14-stranded polymers. As shown in Fig. 1, the polymers are composed of seven double-strands, and the molecules in each double-strand lie in half-staggered registry with one another (1). Within the diagonal contact region, one of the $\beta 6$ mutation sites forms a contact with a receptor region on a diagonally adjacent molecule. This contact region is known as a lateral contact and the less specific contact region directly along the double-strand axis is known as the axial contact region.

The polymerization of deoxygenated sickle hemoglobin is an unusual reaction, displaying both an extraordinary concentration dependence (equivalent to a reaction of 30–50th order) and exponential growth over the first 10–20% of the reaction (2). These features directly contribute to the pathophysiology of the disease, which fundamentally arises from the difficulty of deforming polymer-containing erythrocytes to permit their transit through the narrow capillaries in the circulation. It is in these capillaries where oxygen exchange primarily (though not exclusively) occurs and where polymerization will be most catastrophic. Because of the sharpness of the exponential timecourse, very few polymers form for a period of time (called a delay time) sufficient for most cells to escape the narrow vessels in which they would become trapped once enough hemoglobin has polymerized (3). The high concentration-dependence, however, gives the reaction a strong sensitivity to conditions, and thus small changes can significantly displace the delay time so that cells no longer escape. In short, this extraordinary reaction is an integral aspect of the disease.

This reaction is the consequence of a double nucleation mechanism, as illustrated in Fig. 2 (4,5). In this mechanism, the overall reaction is initially bottle-necked by the need to form a nucleus. The nucleus is assumed to have no special structure, but is simply a piece of the polymer. It rate-limits the reaction because it is the species of lowest concentration, and thus by definition, the least stable species along the reaction pathway. Once this nucleus forms, the reaction is subsequently downhill, and each nucleus forms a very long polymer. The surface of such polymers can also act to catalyze further nucleation, called heterogeneous nucleation. Nuclei that form on the surface of a polymer act like homogeneous nuclei in being able to generate polymers of their own. These polymers are indistinguishable from the original polymers and therefore can, themselves, assist more heterogeneous nucleation and more polymerization, so that the exponential growth of polymerized hemoglobin is readily explained. The concentration-dependence arises from the need to create nuclei by the spontaneous coalescence of many monomers, and is further enhanced by significant crowding (or solution nonideality) (6). Double nucleation was proposed in 1980 as a way to account for a variety of physico-chemical kinetic experiments (4), and directly observed in 1990 (5).

No specific molecular model for heterogeneous nucleation was put forth until 1996, when Mirchev and Ferrone (7) proposed that the same contact partners seen in the double-strand, i.e., the $\beta 6$ Val donor and the corresponding acceptor regions, were the critical elements of the heterogeneous process. This proposal was based on the observation that four of the 14 strands had donor or acceptor regions exterior to the polymer, and that docking of the donor and acceptor
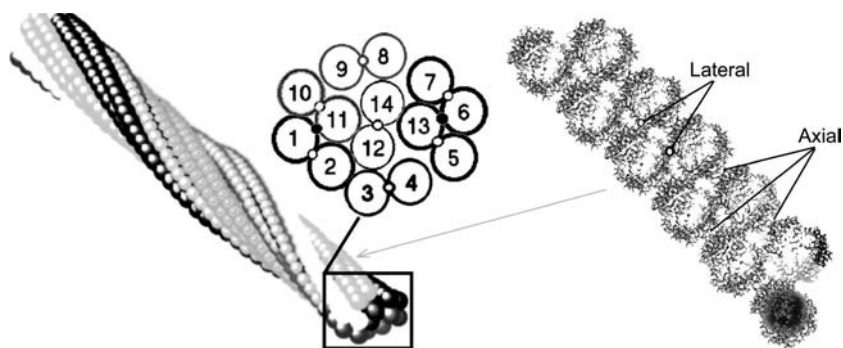
FIGURE 1 The structure of sickle fibers. On the left is the sickle hemoglobin polymer. Seven double-strands are wrapped together, thereby forming a fiber of varying width. A helical section is shown, in which lateral contacts between fibers are indicated by small dots. The $\beta6$ sites are located at the lateral contacts. The white dots are the lateral contacts known from the double-strands; the black dots are proposed based on model building and thermodynamic measurements (21). The linear double-strands determined by crystallography are shown on the right. (22) In the picture shown, a spherical core has been cut from each Hb molecule to highlight the contact regions. On one end molecule, the excised core is shown as a solid sphere. The $\beta6$ Val is in the lateral contacts, again highlighted by open dots.

appeared possible. For a hemoglobin molecule on the polymer surface, both $\beta6$ Val groups would be engaged in contacts: one within the polymer and another as a nucleation attachment site. This suggested a straightforward way to test this proposal: hybrid molecules composed of one β_A and one β_S subunit should frustrate heterogeneous nucleation. Because Hb tetramers can exchange subunits, this experiment requires one more ingredient. It has been shown that it is possible to cross-link sickle hemoglobin and yet permit its polymerization while thwarting subunit exchange (8,9). If subunit exchange is thus prevented so that no molecules are formed having both $\beta6$ Val groups, the experiment should directly show the suppression of heterogeneous nucleation. If heterogeneous nucleation still occurs, another site must be involved. This idea is the core concept of this work.

The foregoing analysis is somewhat oversimplified, however. If some hemoglobin hybrids assume the ''wrong'' position in the polymer, so that their $\beta6$ Glu site is internal and the $\beta6$val site external, heterogeneous nucleation could still proceed. Although this is unlikely because of the energetic penalty associated with trying to place a charged group into a hydrophobic region, it is not impossible. With so many heterogeneous nuclei forming in the usual case, it remains completely plausible that, although suppressed, the amount of heterogeneous nucleation remains readily observable. Moreover, nucleation in general may be suppressed in these hybrids. All nuclei will be less stable if only for the reason that hybrids no longer have two equivalent, favorable

ways to dock in the polymer. These considerations make this experiment, so simple in concept, one in which a detailed quantitative analysis is essential.

Consequently, the best way to understand this experiment is by recourse to the model developed for nucleation in sickle hemoglobin (10). This in no way biases the outcome, since the model has never made specific assumptions about the nature of the contact sites that generate heterogeneous nucleation. Thus, the comparison can be executed with confidence, since both nucleation processes have been modeled very successfully and can be measured directly. In this article, then, we have analyzed the nucleation of cross-linked HbAS hybrids. To provide the framework we recapitulate a number of results of the theory. In the analysis, we find that a strong suppression of heterogeneous nucleation has occurred, and thus we find very strong evidence for the proposal of Mirchev and Ferrone (7) that the same contact site is involved in heterogeneous nucleation as in the polymer. Moreover, homogeneous nucleation is also affected by the use of hybrids, providing strong evidence that both $\beta6$ sites are present for at least some molecules within the nucleus.

METHODS

Sickle hemoglobin and hemoglobin A were prepared by column chromatography using standard procedures. Cross-linked HbAS was prepared as follows. Oxyhemoglobin A and S were mixed in 0.2 M bis tris buffer,

Homogeneous Nucleation
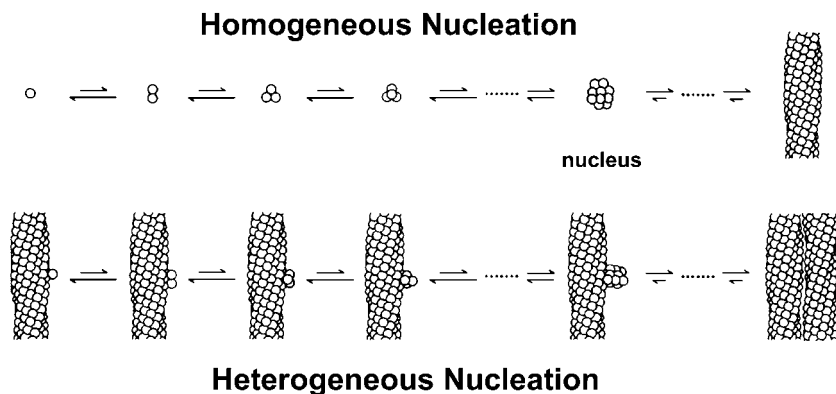

nucleus

Heterogeneous Nucleation

FIGURE 2 Double-nucleation model of Ferrone et al. (10). In homogeneous nucleation, a nucleus forms from solution, whereas in heterogeneous nucleation, a nucleus forms on the surface of another polymer. Nuclei are always unfavorable, so that the equilibrium arrows point more strongly away from nuclei than toward them.

pH 7.2, and the mixture then converted to deoxyhemoglobin by a stream of humidified nitrogen for 1–2 h at 25°C in the presence of inositol hexaphosphate in 10-fold molar excess over hemoglobin. Cross-linking was performed using Bis(3,5-dibromosalicyl) fumarate (DBBF) according to the procedure of Chatterjee et al. (11) A solution of DBBF at a ratio of 1.5 molar excess over Hb was added under anaerobic conditions. The cross-linking reaction was carried out for 2 h at 37°C. At the end of the reaction, glycine was added to a final concentration of 1 M to consume any remaining amount of the reagent. The reaction mixture was then saturated with CO and passed through Sephadex G-25 (Amersham Biosciences, Piscataway, NJ) to remove DBBF and IHP. The mixture was then purified by gel filtration on Sephadex G-100 (superfine) in 1 M MgCl$_2$ + 0.1% Tris, pH 7.0 to separate the un-cross-linked dimers from the cross-linked tetramer fractions (12). Finally, the different cross-linked fractions were separated by use of ion exchange chromatography on a High Q (Biorad, Hercules, CA) column in 25 mM Tris/bis tris pH 8.0. Final purity was checked by polyacrylamide gel electrophoresis. Samples were exchanged into 0.15 M phosphate buffer (pH 7.35) on PD-10 columns, and then concentrated by Centricon or Microcon concentrators (Millipore, Billerica, MA).

Solubility measurements were made using centrifugation as described elsewhere (13) after deoxygenation by 50 mM dithionite.

Sample concentrations for kinetic measurements were determined by diluting the concentrated hemoglobin using micropipettes with 1% accuracy and measuring the Soret absorption spectrum of the dilute sample in a 1-cm-pathlength cuvette. Concentration determinations were repeated three or more times to further improve precision. Kinetic studies were performed on thin slides of COHb which had 50 mM sodium dithionite added to reduce any methemoglobin and scavenged oxygen.

Experiments on the kinetics of polymer formation were performed using laser photolysis of COHb, with parallel data collection (14). Photolysis was achieved by introducing the 488-nm line of a CW-argon ion laser into the optical train of a horizontal microspectrophotometer. Detection was by CCD. The laser beam was focused on the sample by a Leitz 10× LWD strain-free objective (Wetzlar, Germany), and collected by an equivalent objective. Photolysis was confirmed as complete by observing the change in optical density. Photolyzed volume was determined by measuring the size of the photolysis spots and deducing sample thickness from absorption and previously determined concentration. Polymerization was observed by detection of the scattered photolysis light. The direct beam was blocked by a spot at a point conjugate to the back aperture of the objective and scattering proceeded around the spot. To collect data on many different regions simultaneously, the laser was passed through a mesh that was subsequently imaged at the position of the field diaphragm. This created an array of photolyzed spots at the sample of which 60–200 were observed and analyzed. To repeat the experiment the sample was moved by micrometer stage so that there was no question of the full recovery of the photolyzed region. This allowed rapid collection of large numbers of progress curves in a short time, which insured sample stability. Temperature was regulated by a thermoelectric stage.

Homogeneous nucleation rates were obtained by analysis of the distribution of the initiation of light scattering for the different curves, using Szabo's equation (15), i.e.,

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

where Γ is a gamma function. The parameter n is related to the threshold number of monomers required for detection, and is $2\theta (c_o - c_s)N_o VB/J$ (F.A. Ferrone, unpublished), in which N_o is Avogadro's number, and J is the net rate of polymer elongation. The values c_o and c_s are initial concentration and solubility of the hemoglobin solution, respectively. The value θ is a threshold parameter, which is taken as 1/10 for measurements of 10th-time. The value n lies in the range 10^3–10^6. $T(t)$ is small for small t, and once $t > 1/B$, the distribution becomes a decaying exponential, whose decay constant is ζ, the rate of homogeneous nucleation. The rate constant for homogeneous

nucleation, f_o, is related to ζ by including Avogadro's number and the volume in the relationship $\zeta = f_o N_o V$.

The curves were analyzed to yield an exponential growth rate, denoted B. The constant B is dominated by heterogeneous nucleation, although small terms in the expression for B arise from the concentration of homogeneous nucleation as well. Specifically,

$$B_2 = J_o(g_o - df_o/dc), \qquad (2)$$

in which J_o is the net rate of polymer growth, f_o the rate of homogeneous nucleation, and g_o the rate of heterogeneous nucleation (10).

THEORY

Equilibrium

The solubility can be expressed by equating the chemical potential of monomers in solution with that of monomers incorporated into polymers, as

$$RT \ln \gamma_S c_S = -RT \ln 2 + \mu_{PC} + \mu_{PV} - \mu_{RT}, \qquad (3)$$

in which μ_{PC} and μ_{PV} are, respectively, the chemical potential for contacts and vibration of the center of mass of a molecule in the polymer, and μ_{RT} is the chemical potential of rotations and translations of a molecule in solution. R is the gas constant, and T is the absolute temperature. The ln 2 term arises from the fact that there are two $\beta6$ Val that can dock in a receptor site. This could be considered as an entropic contribution to stability in the polymer. Equivalently, one could consider the equilibrium that of $\beta6$ Val with receptors, in which case each HbS molecule having two $\beta6$ sites makes the net concentration of $\beta6$ Val twice that of HbS. For the hybrids, Eq. 3 becomes

$$kT \ln \gamma_s^h c_s^h = \mu_{PC}^h + \mu_{PV}^h - \mu_{RT} + kT \ln f_p^+, \qquad (4)$$

where the superscript h designates quantities for the hybrid molecule. Hybrid molecules have only a single $\beta6$ Val and so have no ln 2. The value f_p^+ is the fraction of molecules in the polymer that are correctly docked, and from mixture data, we expect $f_p^+ \approx 1$, i.e., the number of incorrect dockings is small. Then, with f_p^- being the fraction incorrectly docked,

$$kT \ln \gamma_s^h c_s^h = \mu_{PC}^h + \mu_{PV}^h - \mu_{RT} + kT \ln(1 - f_p^-)$$
$$= \mu_{PC}^h + \mu_{PV}^h - \mu_{RT} - kT f_p^-. \qquad (5)$$

Kinetic equations

There are two principal equations that describe polymer formation. The concentration of monomers incorporated into polymers is denoted Δ (defined as $c - c_o$) and is given by

$$\frac{d\Delta}{dt} = Jc_p, \qquad (6)$$

where c_p is the concentration of polymers, and J is the elongation rate,

$$J = k_+ \gamma c - k_- = k_+ (\gamma c - \gamma_s c_s). \quad (7)$$

The concentration of polymers changes because of polymer creation by homogeneous nucleation, the rate of which is denoted by f or by heterogeneous nucleation, with a rate given by the product $g\Delta$. Since heterogeneous nucleation requires the existence of polymers, the rate is proportional to the mass of polymers, Δ. Hence,

$$\frac{dc_p}{dt} = f + g\Delta. \quad (8)$$

These nonlinear equations have been solved by linearization and expansion for the initial growth phase, and thus the quantities of interest are subscripted by zero to designate the initial value of a quantity that will change during the reaction.

Homogeneous nucleation

The homogeneous nucleation rate f_o 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

$$f_o = k_+ \frac{\gamma_o c_o \gamma_{i^*} c_{i^*}}{\gamma^{\ddagger}}, \quad (9)$$

in which $\gamma^{\ddagger}$ is the activity coefficient of the activated complex (Hill (16); i.e., Eq. 2 and Eq. A2.1 of Ferrone et al. (10)). The value k_+ is the monomer addition rate, taken as size-independent. The value c_{i^*} can be related to the contact energy, μ_{PC}, and the chemical potential from vibrational entropy μ_{PV}. 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, i.e., $\delta(i)$. The total contact energy for size i is given by

$$\mu_{iC} \equiv i \delta(i) \mu_{PC} \approx (i + \delta_1 \ln i + \delta_2) \mu_{PC}, \quad (10)$$

in which δ_1 and δ_2 are determined from fitting the above functional form of $\delta(i)$ to the contacts determined from close-packed spheres (17). The total chemical potential from vibrations is $(i - 1) \mu_{PV}$.

For concise notation the parameter ξ is introduced, which contains μ_{PC} and other constants specified by the geometry of the nucleation process. The value ξ is defined by

$$\xi = -(4 + \delta_1 \mu_{PC}/RT). \quad (11)$$

In fitting data, μ_{PC} is determined by varying ξ while the geometrical parameters are unchanged. The homogeneous nucleation rate f_o is given by

$$f_o = q k_+ \frac{\gamma_o c_o \gamma_s c_s}{\gamma^{\ddagger}} \left[\frac{\ln S}{\xi}\right]^{\xi} e^{1.12\xi}, \quad (12)$$

where q contains only geometrically determined constants (i.e., Ivanova et al. (18)), and S is the activity supersaturation at the initial concentration c_o, defined as

$$S = \gamma_o c_o/\gamma_s c_s, \quad (13)$$

in which c_s is the solubility, and γ_s is the activity coefficient at solubility. The value c_o is the concentration of deoxyhemoglobin S at the initiation of the polymerization. The value γ_o is the activity coefficient of the total hemoglobin concentration at initiation. The nucleus size, i^*, is only used in the activity coefficient for the activated complex. The expression for the nucleus size is

$$i^* = \xi/\ln S. \quad (14)$$

Crowding effects appear in the activity coefficient for the monomer, γ, which depends on the total concentration of hemoglobin and in $\gamma^{\ddagger}$, the activity coefficient for the activated complex, an aggregate of size $i^* + 1$. This depends on the nucleus size (i^*) and the concentration of hemoglobin c, but since i^* is determined (Eq. 14) without added parameters or variables, $\gamma^{\ddagger}$ is fully specified.

Heterogeneous nucleation

The rate of heterogeneous nucleation $g_o\Delta$ is proportional to the concentration of monomers already present in polymers. Analogous to Eq. 9,

$$g_o\Delta = \frac{k_+ \gamma c \gamma'_{j^*} c'_{j^*}}{\gamma'_{j^*+1}}. \quad (15)$$

Here the primes indicate the concentration and activity coefficients of attached aggregates of size j^*. The activity coefficient of the attached aggregate 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. Since the heterogeneous nucleus consists of an aggregate attached to a polymer, the activated complex is no longer a spherical object. The heterogeneous nucleus, an attached aggregate, is written with a prime to distinguish it from a free aggregate like the homogeneous nucleus, which would merely differ in size.

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

$$\gamma'_{j^*} c'_{j^*} = K'_{j^*} \gamma_{j^*} c_{j^*} \gamma_p \phi\Delta, \quad (16)$$

in which primes indicate attached aggregates, and unprimed symbols indicate solution aggregates. The value K'_{j^*} is the equilibrium constant for the attachment process, and Δ is the concentration of monomers in polymers. The value ϕ was originally taken as simply designating the fraction of polymerized monomers that can accept an aggregate (however, see below), and γ_p is the activity coefficient of the polymer with no aggregate attached. Then Eq. 16 becomes

$$g_o\Delta = \frac{k_+ \gamma c K'_{j^*} \gamma_{j^*} c_{j^*} \gamma_p \phi\Delta}{\gamma'_{j^*+1}} = k_+ \gamma c \gamma_{j^*} c_{j^*} K'_{j^*} \phi\Gamma\Delta, \quad (17)$$

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. The value Γ was once taken to be unity (10), an assumption that has been shown to be untenable (18), and an approximate expression has been derived (19), i.e.,

$$\ln \Gamma = -\frac{3j*2/3(c_{pp}v)^{-2/3}vc}{1-vc} - \frac{j*c}{c_{pp}}\left[\frac{1+vc+(vc)^2}{(1-vc)^3}\right], \quad (18)$$

where c_{pp} is the concentration of hemoglobin in the polymer phase, and v is the specific volume of the monomer. The size of the heterogeneous nucleus $j*$ is computed in a thermodynamic treatment similar to that used for the homogeneous nucleus, but with the addition of an energy of attachment. The additional terms that appear in the calculation of the attachment can be traced to $K'_{j*}\phi$ in Eq. 17. In energetic terms the added stability of the heterogeneous nucleus is

$$-RT \ln K'_{j*}\phi = -RT \ln \phi + \mu_{CC}\sigma_1 j* + \mu_{CC}\sigma_2 \ln j*, \quad (19)$$

where μ_{CC} is the chemical potential per unit contact area between polymers and thus is the energetic term causing a heterogeneous nucleus to stick to the polymer. In contrast, μ_{PC} is the energetic term describing the contact energy of a monomer within the polymer. The surface area in contact has linear and logarithmic contributions with coefficients σ_1 and σ_2. This is similar to the expansion of the contact energy within the homogeneous nucleus in constant, linear, and log terms (recall Eq. 10). Originally ϕ was 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.

From the above definitions, analogous to Eq. 12,

$$g_o = k_+ \gamma_o c_o \phi \Gamma \left[\frac{\ln S + \xi_1}{\xi_2}\right]^{\xi_2} e^{\xi_2}, \quad (20a)$$

where

$$\xi_1 = -\sigma_1 \mu_{CC}/RT, \quad (20b)$$

and

$$\xi_2 = \xi + 4 + \sigma_2 \mu_{CC}/RT. \quad (20c)$$

S is the activity supersaturation as above. The value ξ is also the same variable used in Eq. 11. In this notation, $j*$ can be written as

$$j* = \frac{\xi_2}{\ln S + \xi_1}. \quad (21)$$

The heterogeneous nucleus size $j*$ does not depend on ϕ.

RESULTS

Solubility measurements were performed on cross-linked HbS and AS, as well as HbS for reference. Three separate preparations were used, and solubility measured two or more times for each. HbS solubility was measured at 25°C to be 2.91 ± 0.08 mM (tetramer). With the cross-link, the solubility decreased to 2.75 ± 0.03 mM (i.e., ~0.1 kcal/mol more stable). The cross-linked hybrid, on the other hand, had significantly higher solubility, i.e., 4.27 ± 0.06 mM (~2.5 kcal/mol less stable). All errors represent replication uncertainty. From these solubility measurements, overall stability ΔG can be determined as $RT \ln \gamma_s c_s$, as listed in Table 1.

When kinetics were measured, all hemoglobins showed the same qualitative features, i.e., exponential growth of the initial progress curves, and stochastic distributions of the tenth times. Those distributions were used to deduce homogeneous nucleation rates. For the parameter B as well as homogeneous nucleation rates, all hemoglobins showed high concentration dependences. Fig. 3 shows homogeneous nucleation rates for cross-linked SS and cross-linked AS hemoglobins in a and b, respectively, as a function of initial concentration. (In passing, we note that HbSxl rates are similar to the HbS rates.) The difference in homogeneous nucleation rates between Fig. 3 a for HbSxl and Fig. 3 b of HbASxl is substantial. Fig. 4 shows the parameter B, dominated by heterogeneous nucleation (Eq. 2) as a function of initial concentration for HbS, cross-linked HbSS, and HbAS. Again, there is a significant change when one $\beta 6$ is Glu rather than Val.

The curves through the data show the theoretical fits. For each sample, its respective solubility determined by sedimentation was used as a fixed parameter in the fit. Homogeneous nucleation was fit by varying ξ, which in turn is the result of varying the average attachment energy, μ_{PC}. We found that $\xi = 14.6$ and 11.0 for HbSxl and HbASxl, respectively. At 25° this translates into $\mu_{PC} = -8.5$ kcal and -6.9 kcal/mol, as shown in Table 1.

Overall stability ΔG is distinct from contact energy. The overall stability is dictated by the combination of contact energy μ_{PC} and a free energy due to motion of the monomers around their equilibrium position in the polymer. This motional term arises from vibrational entropy and is dominated by the number of normal modes that are excited, and that in turn is proportional to the number of molecules in an aggregate. It is not proportional to the number of contacts and is thus experimentally distinct from the contact energy. For example, in HbS, μ_{PC} is -7.6 kcal/mol, and μ_{PV} is much larger, being -26.3 kcal/mol.

TABLE 1 Analysis of homogeneous nucleation energies in kcal/mol

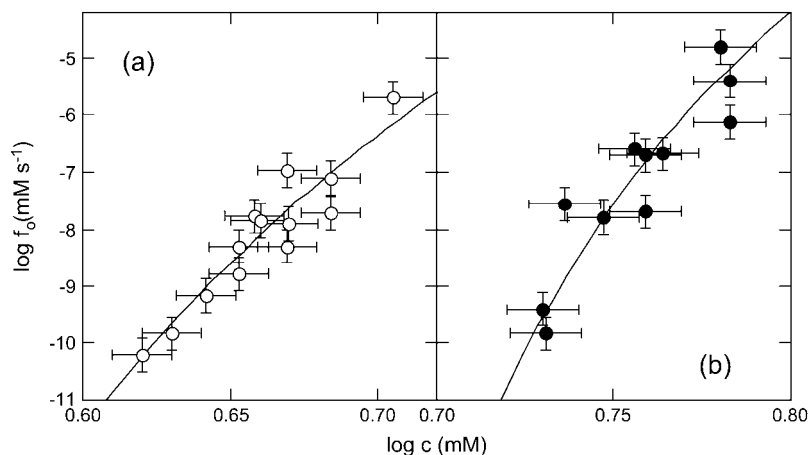|	HbSxl	HbASxl
μ_{PC}	-8.5 ± 0.1	-6.9 ± 0.1
$RT \ln \gamma_s c_s$	1.48 ± 0.01	2.52 ± 0.01
Orientation	0	0.4
μ_{PV}	-25.5	-26.5
$i*$ in range studied	5–9	4–7

FIGURE 3 Homogeneous nucleation rates, as a function of initial concentration. (*a*) Data for HbSxl. (*b*) Data for HbASxl. Note the different concentration ranges. Homogeneous nucleation is substantially slowed by the presence of one $\beta 6$ Glu, as seen in *b*. Solid lines show fits of the double nucleation model to the data. In this case, there is only one varied parameter, μ_{PC}, the polymer contact energy.

HbASxl is less stable than HbSxl by ~ 1 kcal/mol in ΔG. In part, this is because an AS hybrid has one strongly preferred orientation within the polymer. This preference is made quantitative by an entropic term of size $RT \ln 2 = 0.4$ kcal/mol. Compared to HbSxl, the contact energy in HbASxl contributes even less stability—by 1.6 kcal/mol. Remarkably, the vibrational chemical potential, despite its large size, varies by no more than 1 kcal/mol, a mere 4%.

The changes in homogeneous nucleation rates are profoundly affected by change in solubility, which in turn affects supersaturation S. For example, at $c = 5.25$ (log $c = 0.72$), supersaturation for HbSxl is 21.3 versus only 3.8 for HbASxl. From Eq. 12, and the fact that ξ is in the range 12–14, it is clear that the change in ln S alone is enough to account for greater than five orders of magnitude in the homogeneous nucleation rate.

Next we turn to heterogeneous nucleation. A heterogeneous nucleus is thermodynamically equivalent to constructing an aggregate in solution and then attaching it to a polymer, and thus all the effects contributing to the stability of homogeneous nuclei must also contribute to heterogeneous nuclei as well. In addition, there is the energy due to the contact of the aggregate with another polymer. That

energy has four parts: it has a simple statistical part because of the relative frequency of the contact sites, which is the focus of the study here. Once a contact is made, the size-dependence of the contact energy is taken to given rise to three parts, i.e., constant, linear, and nonlinear size-dependent terms.

When only ϕ is allowed to vary, the results of the fit to B are shown in Fig. 4 as the solid line. In a complete treatment, the quantities $\mu_{CC}\sigma_1$ and $\mu_{CC}\sigma_2$ should also be varied since these terms in the model alter the dependence on size. However, the data set is not sufficiently large to obtain convergence with all three parameters adjustable. Therefore, we held either $\mu_{CC}\sigma_1$ or $\mu_{CC}\sigma_2$ fixed and varied the other member of the pair. The results are shown in Table 2 and drawn as the small dashed and dot-dashed lines in Fig. 4. The fits are all quite similar, and in all cases, $RT \ln \phi$ drops significantly, which is the central result of this work. The long dashed line shows the expectation that there was no change in ϕ, but simply a change in solubility and homogeneous nucleation, as seen in Fig. 3.

Although again the change in solubility exerts a substantial influence here, there is an offsetting influence due to the change in heterogeneous nucleus size. As seen in Table 2,
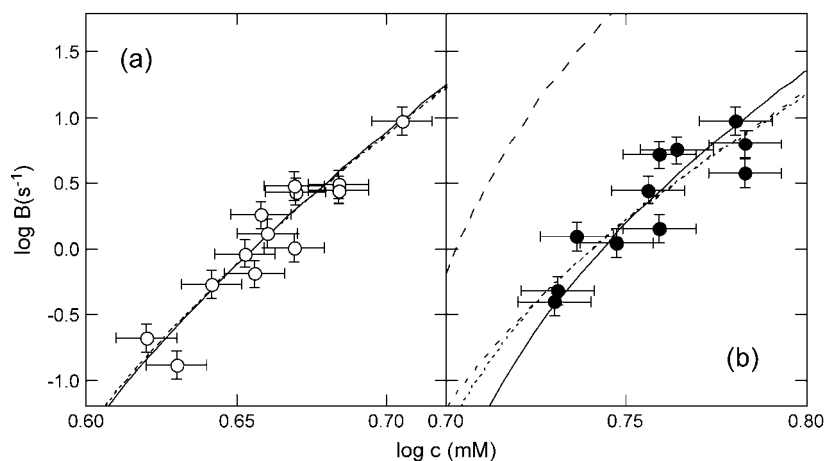

FIGURE 4 Exponential growth parameter B as a function of initial concentration. As shown in Eq. 2, B is dominated by the rate of heterogeneous nucleation, g_o. Data is shown for the same samples observed in Fig. 3, i.e., (*a*) HbSxl and (*b*) HbASxl. For the cross-linked samples, the solid curve was obtained by only varying ϕ, whereas the small dashed and dot-dashed curves have varied either $\mu_{CC}\sigma_2$ or $\mu_{CC}\sigma_1$ in addition to ϕ. The long dashed curve in *c* shows what would be expected for HbASxl if there had been no change in ϕ.

TABLE 2 Analysis of heterogeneous nucleation energies in kcal/mol

	HbSxl	HbASxl
$RT \ln \phi$	5.0 ± 0.5*	0.3 ± 0.5*
	4.7 ± 1.4	-1.2 ± 1.0
	4.8 ± 0.8	-0.7 ± 0.9
$\mu_{CC}\sigma_1$	0.17 ± 0.01	0.17 ± 0.01
	0.17 ± 0.01	0.17 ± 0.01
	0.16 ± 0.07	-0.01 ± 0.16
$\mu_{CC}\sigma_2$	-6.5 ± 1.0	-6.5 ± 1.0
	-6.7 ± 1.1	-7.2 ± 1.1
	-6.5 ± 1.0	-6.5 ± 1.0
j* in range studied	2–3	1

*In the first row, only $\ln \phi$ is varied; in the second row, $\mu_{CC}\sigma_2$ is also varied; and in the third row, $\mu_{CC}\sigma_1$ is varied but $\mu_{CC}\sigma_2$ is held fixed.

the nucleus size drops from between 2 and 3, for HbSxl, to 1 for the cross-linked AS molecules. This has a substantial effect on the activity coefficient ratio, Γ, as seen in Eq. 18. This effect of a factor of 2–3 occurs in the exponent, and thus has more impact than the small change in homogeneous nucleus (range of 5–9 for HbSxl, and 4–7 for HbASxl). The change in heterogeneous nucleus size causes the activity coefficient ratio to change in a way that partially offsets the direct changes in $\ln S$ in Eq. 20a. As a result, the change in Fig. 4 between HbSxl and HbASxl, due to everything except changes in ϕ, is much less than would have naively expected from the large change evidenced in Fig. 3. The changes excluding alteration of ϕ are those seen when the data of Fig. 4 *a* are compared with the dashed line of Fig. 4 *b*.

From the results as seen in Fig. 4 when ϕ is varied, heterogeneous nuclei are more than 4000 times less frequent in HbASxl than seen in HbSxl. This large decrease in the fraction of available sites for heterogeneous nucleation in a cross-linked polymer provides the strongest evidence supporting the Mirchev-Ferrone proposal for the origin of heterogeneous nucleation.

DISCUSSION

The presence of the cross-link affects solubility, homogeneous nucleation, and heterogeneous nucleation. It is precisely this interrelationship that makes it necessary to analyze the data in a detailed fashion so as to distinguish the molecular origin of the various results. We first examine heterogeneous nucleation, and then homogeneous nucleation.

Heterogeneous nucleation

One reason that any heterogeneous nucleation is observed may be that some hybrids have assumed a more energetically costly reversed position. (Another reason, of course, could be that there is more than one pathway.) The free energy difference of ~5 kcal (= $\Delta RT \ln \phi$), therefore, is interpreted

as the penalty for reversing a hybrid molecule. This corresponds well with the difference in hydrophobic energy of ~4 kcal/mol between removing a Val and Glu from solvent (20). This would be the energetic cost, for example, of reversing a hybrid molecule on the surface of the polymer, so as to put Val on the polymer exterior and place the Glu in the hydrophobic pocket receptor pocket of an adjacent molecule inside the polymer. This agreement is also quite good, given the rough nature of the approximation of hydrophobic energy.

Homogeneous nucleation

Until recently, it was widely believed that only one $\beta6$ group per molecule participated in contacts within the polymer (2). If this were the case for nucleation, it would be hard to understand the difference in μ_{PC} between HbSxl and HbASxl. A model in which only one $\beta6$ group is involved in the intermolecular contacts would predict the only change would be that of statistics, i.e., $RT \ln 2 = 0.4$ kcal/mol, far less than the 1.6 kcal/mol observed.

It has recently been proposed that, for some monomers, both $\beta6$ groups have interactions within the polymer (21). Those monomers with additional contacts are denoted as tetrads. If nucleation began with the tetrad, this would give a qualitative explanation for the change in the contact energy per molecule, μ_{PC}, on going from HbSxl to HbASxl. Since, in the concentration range studied here, the homogeneous nucleus sizes range between 4 and 7, we shall consider a nucleus size of 5.5 for illustration. If nucleation were to proceed through the tetrad geometry, two added contacts could have been made in HbSxl, which would be absent in the hybrid. Using the standard value of 4 kcal/mol for changing between Val and Glu gives a prediction of 2 contacts $\times$ 4 kcal/mol contact/5.5 molecules in the nucleus = 1.5 kcal/mol. Using the 5 kcal/mol, as deduced in the heterogeneous nucleation analysis, gives 1.8 kcal/mol for the equivalent calculation. These bracket the measured 1.6 kcal/mol change in contact chemical potential (per monomer basis). This then provides a quantitative support for the notion that the tetrad geometry provides a detailed path for nucleation.

As a final step in the consideration of homogeneous nucleation, it is useful to ask why this large change due to the tetrad is not apparent in overall stability ΔG. If the second $\beta6$ is customarily buried in a hydrophobic pocket, then replacing it with a Glu could entail as much as 4–5 kcal/mol as a penalty. But this 4–5 kcal/mol is at two contacts, whereas the energy deduced from solubility is mathematically distributed across 14 molecules (since differences are not distinguished). Thus, one expects a penalty of 4–5 kcal/mol $\times$ 2/14. This gives a cost of 0.6–0.7 kcal/mol per average molecule in the polymer. In addition, a statistical penalty occurs of an additional 0.4 kcal/mol. This penalty is exacted for all molecules not having the extra contact—that is, the ones where only one contact matters. This adds 0.4 kcal/mol

$\times 10/14 = 0.29$ kcal/mol. The result is 0.89–0.99, which is not very close to the observed 1.0 kcal/mol observed difference in stability ΔG.

Significance and implications

The data presented here provide strong evidence that the molecular contacts involved in heterogeneous nucleation are also a subset of the same molecular contacts as seen in the polymer interior. This provides a degree of simplicity in understanding the thermodynamics that underlie this somewhat complex assembly process. In addition, it provides a degree of closure to the concept of heterogeneous nucleation, first proposed in 1979. As this and other work (19) has shown, there will be more than one amino acid in contact for the heterogeneous nucleus, and the additional participants need to be identified, perhaps by model building and searching the molecular surface, before the understanding of heterogeneous nucleation can be considered complete. This, however, is an important step in that direction.

At the same time, the results demonstrate that the interior of the polymer is more complex than once thought, even though the same amino acid partners are involved, and in some cases the details will need to be incorporated into present models for a more complete description. In both cases, the connection between molecular and kinetic models has been strengthened, and this can be expected to provide further insights into the detailed process of polymerization and the search for appropriate therapies for sickle cell disease.

REFERENCES

1. Wishner, B. C., K. B. Ward, E. E. Lattman, and W. E. Love. 1975. Crystal structure of sickle-cell deoxyhemoglobin at 5 Å resolution. *J. Mol. Biol.* 98:179–194.

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

3. Mozzarelli, A., J. Hofrichter, and W. A. Eaton. 1987. Delay time of hemoglobin S polymerization prevents most cells from sickling. *Science.* 237:500–506.

4. 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.

5. 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.

6. Ferrone, F. A., and M. A. Rotter. 2004. Crowding and the polymerization of sickle hemoglobin. *J. Mol. Recognit.* 17:497–504.

7. 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.

8. Benesch, R. E., R. Edalji, R. Benesch, and S. Kwong. 1980. Solubilization of hemoglobin S by other hemoglobins. *Proc. Natl. Acad. Sci. USA.* 77:5130–5134.

9. Bookchin, R. M., T. Balazs, R. L. Nagel, and I. Tellez. 1977. Polymerisation of haemoglobin SA hybrid tetramers. *Nature.* 269:526–527.

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

11. Chatterjee, R., E. V. Welty, R. Y. Walder, S. L. Pruitt, P. H. Rogers, A. Arnone, and J. A. Walder. 1986. Isolation and characterization of a new hemoglobin derivative cross-linked between the Å chains (Lysine 99α1-Lysine 99α2). *J. Biol. Chem.* 261:9929–9937.

12. Benesch, R., and S. Kwong. 1991. Hemoglobin tetramers stabilized by a single intramolecular cross-link. *J. Protein Chem.* 10:503–510.

13. Hofrichter, J., P. D. Ross, and W. A. Eaton. 1976. Supersaturation in sickle cell hemoglobin solutions. *Proc. Natl. Acad. Sci. USA.* 73:3035–3039.

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

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

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

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

18. 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.

19. Ferrone, F. A., M. Ivanova, and R. Jasuja. 2002. Heterogeneous nucleation and crowding in sickle hemoglobin: an analytic approach. *Biophys. J.* 82:399–406.

20. Creighton, T. E. 1993. Proteins: Structures and Molecular Properties. W.H. Freeman and Company, New York.

21. Roufberg, A., and F. A. Ferrone. 2000. A model for the sickle hemoglobin fiber using both mutation sites. *Protein Sci.* 9:1031–1034.

22. 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.