The L49F mutation in alpha erythroid spectrin induces local disorder in the tetramer association region: Fluorescence and molecular dynamics studies of free and bound alpha spectrin

Yuanli Song,¹ Nina H. Pipalia,² and L. W.-M. Fung^{1,2*}

¹Department of Chemistry, University of Illinois at Chicago, Chicago, Illinois 60607

²Department of Chemistry, Loyola University of Chicago, Chicago, Illinois 60626

Received 5 May 2009; Revised 23 June 2009; Accepted 25 June 2009

DOI: 10.1002/pro.202

Published online 10 July 2009 proteinscience.org

Abstract: The bundling of the N-terminal, partial domain helix (Helix C') of human erythroid α -spectrin (α I) with the C-terminal, partial domain helices (Helices A' and B') of erythroid β -spectrin (β I) to give a spectrin pseudo structural domain (triple helical bundle A'B'C') has long been recognized as a crucial step in forming functional spectrin tetramers in erythrocytes. We have used apparent polarity and Stern–Volmer quenching constants of Helix C' of α I bound to Helices A' and B' of β I, along with previous NMR and EPR results, to propose a model for the triple helical bundle. This model was used as the input structure for molecular dynamics simulations for both wild type (WT) and α I mutant L49F. The simulation output structures show a stable helical bundle for WT, but not for L49F. In WT, four critical interactions were identified: two hydrophobic clusters and two salt bridges. However, in L49F, the region downstream of Helix C' was unable to assume a helical conformation and one critical hydrophobic cluster was disrupted. Other molecular interactions critical to the WT helical bundle were also weakened in L49F, possibly leading to the lower tetramer levels observed in patients with this mutation-induced blood disorder.

Keywords: α -spectrin; L49F mutant; tetramers; model structure

Abbreviations: A₁B₁C₁, the first α I structural domain consisting of Helices A₁, B₁, and C₁; A'B'C', helical bundle of Helices A', B', and C'; α I, human erythroid α -spectrin; α I-N, a recombinant protein consisting of residues 1–368 of α I; α I-N Δ , single cysteine proteins of α I-N; β I, human erythroid β -spectrin; β I-C, a recombinant protein consisting of residues 1898–2083 of β I; B₁, a cysteine residue labeled with mBBR; CD, circular dichroism; Helix A₁, the first helix of the first structural domain; Helix A', residues 2009–2032 of β I-C; Helix B', residues 2044–2075 of β I-C; Helix C', residues 21–45 of α I-N; K_{sv} , Stern–Volmer quenching constant; λ_{max} , maximum emission wavelength; mBBR, monobromobimane; MD simulation, molecular dynamics simulation; PBS 7.4, 5 mM phosphate buffer with 150 mM NaCl at pH 7.4; SASA, solvent accessible surface area; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Yuanli Song and Nina H. Pipalia contributed equally to this work.

Nina H. Pipalia's current address is Department of Biochemistry, Weill Medical College of Cornell University, New York, NY 10021.

Grant sponsor: American Heart Association; Grant number: 0350617Z; Grant sponsor: National Institutes of Health; Grant number: GM68621; Grant sponsor: Searle Funds (Chicago Community Trust).

*Correspondence to: L. W.-M. Fung, Department of Chemistry, University of Illinois at Chicago, 845 W. Taylor Street, Chicago, IL 60607. E-mail: lfung@uic.edu

Introduction

The bundling of the N-terminal, partial domain helix (Helix C') of human erythroid α -spectrin (α I) with the C-terminal, partial domain helices (Helices A' and B') of erythroid β -spectrin (β I) to give a spectrin pseudo structural domain (triple helical bundle A'B'C') has long been recognized as a crucial step in forming functional spectrin tetramers.¹ Two such identical associations in two dimers (α I of dimer 1 with β I of dimer 2 and α I of dimer 2 with β I of dimer 1) lead to the formation of a tetramer. The solution structure of the first 156 residues obtained from NMR measurements shows that residues 1–52 are exposed to solvent, and residues 21–45 form a helix^{2,3} (see Fig. 4). This helix is the previously predicted partial domain Helix C'. Many residues in this helix, such as residues 24, 28, and 45, play an important role in stabilizing the triple helical bundle to form functional tetramers. Mutations at these sites often lead to blood disorders.^{4,5} However, mutations at positions outside the helical region, such as the mutation of Leu to Phe at position 49 of α I-spectrin (Spectrin Lyon),⁶ are also clinically detrimental.

It is interesting to note that residue 49 is in one of the two unstructured regions (residues 17–20 and 46–52) flanking both ends of the Helix C'. These regions undergo conformational changes, from unstructured to helical, on binding Helices A' and B', to form A'B'C'.^{7,8} We have suggested that mutations in these regions that affect side chain properties and/or the propensity of helical formation would affect spectrin tetramer formation, and may lead to blood disorders.⁸ In addition, our data suggest that the region consisting of residues 46–52, a region called the junction region since it connects Helix C' to the first structural domain in α I, is particularly important for spectrin tetramer formation.^{7,8} Thus, mutations in the junction region at position 49, from Leu to a residue with a more polar side chain, such as His, will likely lead to and has been shown to exhibit abnormal association.⁹ However, a L49F mutation preserves the hydrophobic side chain properties. In addition, the helix propensity for Phe (0.54) is higher than that of Leu (0.21).¹⁰ Furthermore, in a hydrophobic cluster analysis, V, I, L, F, Y, W, and M are considered as a set of amino acids with similar hydrophobic side chains.¹¹ Therefore, from hydrophobicity consideration, the L49F substitution should not be detrimental to tetramer formation or to red cell membrane integrity. However, L49F (Spectrin Lyon) is a clinically significant mutation that causes hereditary elliptocytosis.⁶ Structural information on the tetramer-bound state of L49F is needed to understand the molecular mechanism of this abnormal association.

Currently, no X-ray information is available for the structure of the A'B'C' bundle of either wild type (WT) or L49F, although a recent publication has indicated that a preliminary crystal structure of the WT has been obtained, but with incomplete structural

refinement.¹² We have shown that, due to the asymmetric ellipsoidal structure of the complex, high resolution solution NMR studies at 900 MHz provide only limited structural information on the complex.⁷ A simulated structure was published earlier, but revealed no specific mechanism.¹³

In this study, we have used both fluorescence methods and molecular dynamics (MD) simulations to understand the detrimental effect of the L49F mutation. Our results showed that the Phe residue prevented the conformational changes expected in the junction region (residues 46–52) on binding to β -spectrin, leading to an unstable triple helical bundle of $\alpha\beta$ -spectrin partial domains, and thus unstable tetramers.

Results

Association affinity and bound state spectra

The tetramerization affinity of erythroid spectrin, modeled by the bundling affinity of Helix C' with Helices A' and B' of model proteins in this study, is relatively low, with a K_d value of $\sim 1 \mu\text{M}$.^{14,15} When critical residues such as R28 were replaced with a labeled cysteine residue, the affinity was further reduced, making the bound state difficult to study. We systematically determined the affinity of all α I-NB₁ proteins (see Materials and Methods section for proteins used and their abbreviations) with β I-C. With the affinity information, we extracted fluorescence properties of the bound state.

Native gel electrophoresis results showed that 33 of the 43 α I-NB₁ proteins exhibited gel patterns similar to that of the parent protein, α I-N (WT), with the disappearance of the free α I-N band under the concentration conditions used,¹⁶ suggesting that their K_d values were also about $1 \mu\text{M}$. Representative gel results showed R27B₁ [Fig. 1(A), Lane 2] with two molar excess of β I-C formed complex with no detectable free R27B₁ band (Lane 3). However, for R28B₁ with two molar excess of β I-C, the free R28B₁ band (Lane 4) was still detected (Lane 5), indicating an affinity lower than $1 \mu\text{M}$. There were 10 proteins (E23B₁, I24B₁, Q25B₁, R28B₁, V31B₁, R41B₁, G46B₁, L49B₁, Y53B₁, and F58B₁) with lower affinities, and their K_d values were then determined by isothermal titration calorimetry (ITC) or fluorescence methods.

The ITC results showed that K_d values ranged from 3 to 25 μM for 8 of the 10 proteins. For example, the value for E23B₁ was 6 μM [Fig. 1(B)]. The ITC measurements for R28B₁ and L49B₁ only showed that their K_d values were greater than 25 μM , but fluorescence measurements showed 29 μM for R28B₁ [Fig. 1(C)], and about 100 μM for L49B₁.

All three methods were applied to one sample, R34B₁, and the results showed a single gel electrophoresis band (data not shown) and a K_d value of 0.6 μM from ITC [Fig. 1(B)], and 0.5 μM from fluorescence measurements [Fig. 1(C)]. Thus, all three methods provided similar affinity information, with gel

Figure 1. Association Affinity. A: A representative native polyacrylamide gel (6%) electrophoresis of the α I-NB₁ proteins ($25 \pm 3 \mu\text{M}$) without and with β I-C ($50 \mu\text{M}$). Four micrograms of protein was loaded to each lane. Lane 1, β I-C; lane 2, R27B₁; lane 3, R27B₁ with β I-C; lane 4, R28B₁; and lane 5, R28B₁ with β I-C. All samples were incubated at 4°C for 16 h before electrophoresis. For R27B₁ with β I-C sample in lane 3, the gel band for R27B₁ was not detectable, indicating a K_d value similar to that of WT, about $1 \mu\text{M}$. For R28B₁ with β I-C sample in lane 5, both 28B₁ and β I-C bands were detected, indicating weaker affinity than that of the WT. B: ITC titration curves of E23B₁ (Δ) and R34B₁ ($\circ$), titrated with β I-C in PBS 7.4 at 20°C. A K_d value of $6 \mu\text{M}$ was obtained for E23B₁, and $0.6 \mu\text{M}$ for R34B₁. The left inset is the isotherm for E23B₁ and right inset for R34B₁. C: Fluorescence intensities of R28B₁ ($\square$) and R34B₁ ($\circ$) in the presence of various amounts of β I-C. Standard deviation error bars for the data are also shown. The data points were fitted to give the curves in the plot and K_d values of $30 \mu\text{M}$ for R28B₁ and $0.6 \mu\text{M}$ for R34B₁.

electrophoresis being the simplest to operate but capable of covering only a narrow range of K_d values, and the more laborious fluorescence methods covering the broadest K_d range. In brief, with the knowledge of K_d values, we obtained free and bound α I-NB₁ fluorescence spectra for all 43 proteins (spectra not shown).

Apparent polarity

Calibration. The plot of λ_{max} and dielectric constant (ϵ) values obtained from the spectra of monobromobimane (3-(bromomethyl)-2,5,6-trimethyl-1H,7H-

pyrazolo[1,2-a]pyrazole-1,7-dione or mBBr) labeled free cysteine in mixtures of dioxane and water exhibited the expected biphasic properties [Fig. 2(A), inset]. A linear fit of the data with dielectric constant (ϵ) values above 10 gave a calibration equation: $\epsilon = (\lambda_{\text{max}} - 453)/0.35$. The calculated ϵ value for each residue in the labeled protein was referred to as local “apparent polarity” value, following the published terminology.¹⁶ Our plot was similar, but not identical, to other published calibration plots. The advantage of this approach is that it provides an instrument-independent way to compare results, with any systematic error in wavelength accuracy for a particular instrument removed by measuring an “apparent polarity scale.”^{16–23}

Free state. As shown in Figure 2(A) (open symbols), the local environment of residues 14–23 in the

Figure 2. Apparent polarity. A: Apparent polarity values, converted from λ_{max} values using a calibration curve shown in inset, of each α I-NB₁ sample, free ($\circ$) or bound to β I-C ($\bullet$). The calibration curve is a plot of λ_{max} from the spectra of cysteine labeled with mBBr ($1 \mu\text{M}$) in solutions of varying concentrations of water and dioxane to give solutions of different dielectric constants, with an excitation wavelength at 384 nm. B: The difference of apparent polarity values (bound state—free state). A negative value represents a less polar environment for the residue after association with β I-C. The residue at position 49 of Helix C' exhibited the largest change in apparent polarity on binding β I-C, whereas residue 25 was in already in an apolar environment before binding to β I-C and showed a slight increase in apparent polarity on binding β I-C.

free state were all quite polar, with an apparent polarity value of 73 at residue 14, that decreased gradually to 67 at residue 23. The segment consisting of residues 14–20 is unstructured but residues 21–23 are helical, according to NMR studies.² For the segment consisting of residues 24–45, residues 24, 26, 30–31, 33, 37, 40, and 44 were polar. They were three or four residues apart, providing a periodical variation in local solvent polarity along this segment of the protein. This periodic oscillation pattern of residues 24–45 was in good agreement with the helical conformation observed in NMR studies² and similar to that observed in side-chain mobility.³ For residues 46–51 (including residue 49), the apparent polarity values increased from ~50 at position 46 to above 60 for residues 47–50 before returning to ~55 at position 51. These values again correlated well with the unstructured conformation determined by NMR studies.² The periodic pattern returned for residues 53–58. These residues are in Helix A of the first structural domain.²

The most nonpolar local environment was at position 25, with an apparent polarity of only ~14, followed by position 28, with a value of ~36 and position 32, with a value ~32. The NMR structure indicates that these three residues shared unique structural properties, each with two hydrophobic neighbors (within 6.0 Å). Residue 25 is close to A21 (an “*i* – 4” residue) and I24 (an “*i* – 1” residue); residue 28 is close to I24 (an “*i* – 4” residue) and V31 (an “*i* + 3” residue); and residue 32 is close to V31 (an “*i* – 1” residue) and Y35 (an “*i* + 3” residue).

In general, the results of the free state were in excellent agreement with the NMR structure of the free state.

Bound state. The excellent agreement of the apparent polarity values with the free state conformation provided us a good foundation for using apparent polarity values of the bound state to obtain structural information about the bound state. The polarity values changed (either increased or decreased) at each position within the bound state, with a few residues exhibiting much larger changes than others [Fig. 2(A,B)]. Other than the large change at position 49, the changes were modest, within ± 10 [Fig. 2(B)]. The three less polar environment residues in the free state (residues 25, 28, and 34) and their immediate neighbors all exhibited somewhat larger changes, with residues 25 and 28 becoming more polar, indicating a polar contribution at these positions from Helices A' and B' of β I-C. The largest change was observed at position 49, where the apparent polarity value dropped from 64 in the free state to 35 in the bound state, suggesting a hydrophobic cluster around residue 49.

Accessibility to KI

Free state. The fluorescence quenching by iodide ions depends strongly on local conformation, hydrophobicity, and charge. Neighboring positively or negatively charged side chains may greatly affect the quenching. We found that the Stern–Volmer constant (K_{sv}) values varied periodically as a function of scanned positions, with values ranging from 5 to 30 M^{-1} [Fig. 3(A)], indicating that the side-chains of the lone Helix C' exhibited different accessibility to iodide ions, and the periodic variations were obvious for positions 21–32. For positions 33–45, since we did not have data for positions 35 and 38, it was difficult to determine whether the variations were periodic.

The residues with low K_{sv} values were at positions 21 (heptad position *a*), 24 (*d*), 25 (*e*), 27 (*g*), 28 (*a*), 31 (*d*), and 39 (*e*). Thus, the hydrophobic side (heptad positions *a*, *d*, *e*, and *g*) was less accessible to the iodide ions than the hydrophilic side (*b*, *c*, and *f*) of Helix C', as expected. Again, the accessibility results

Figure 3. KI accessibility. Stern–Volmer constants (K_{sv}) of α I-NB₁ in free (A) and bound (B) states. K_{sv} values were obtained from intensity values at λ_{max} of α I-NB₁ spectra (free or bound) with no KI (F_0) and in the presence of various concentrations of KI (F). Standard deviation error bars are also shown. Most of the residues exhibit a decrease in K_{sv} values on binding β I-C (C), with the residue 49 showing the largest decrease in KI accessibility on binding. Residues 24, 41, and 42 also show substantial decrease on binding.

were in excellent agreement with the structural information.

Bound state. For the bound state, the K_{sv} values of only eight residues, at positions 21 (a), 22 (b), 29 (b), 36 (b), 37 (c), 39 (e), 40 (f), and 47 (f), remained similar to those of the free state [Fig. 3(B,C)]. Most of these residues are at the hydrophilic positions of the heptads of Helix C', and thus were not expected to change much on binding β I-C helices. The residues with the most changes on binding were again at position 49 (a), followed by 42 (a) and then 24 (d) [Fig. 3(C)], with all of these at heptad positions a or d of Helix C' and thus expected to be affected most on binding β I-C. It is interesting to note that most of the residues at both ends of Helix C' become less accessible to KI than those in the middle.

Model for triple helical bundle A'B'C'

Since currently no experimental structural information is available for the structure at the tetramerization region including the crucial bundling of helices A', B', and C' (A'B'C') either for the WT or for mutant L49F, we used modeling, guided by fluorescence and other published data, to predict the structures. Because of heptad repeats in sequences of these three helices, the general structure of the bundle is relatively simple to predict, with the residues at positions a and d of each helix facing each other to form a hydrophobic core of the bundle. Since the sequences of the partial domain Helices A', B', and C' are homologous to those in structural domains, it is reasonable to select a known structure of a spectrin structural domain as a template to model the structure of A'B'C'. The erythroid α -spectrin first structural domain,² or the *Drosophila* 14th structural domain¹³ was used as the template for erythroid A'B'C'. With either template, the next step should be to determine whether the relative distances between the helices, helical lengths, and helix register need to be modified.

We showed that the relative distances between the helices were not critical if an MD simulation was applied. We input several structures with slightly different relative distances (not more than 5 Å) between the helices and found that Helix C' moved "back" to a position similar to its original position after the MD simulation. Thus, slightly different relative positions for helices resulted in similar MD simulation structure, suggesting that optimized interactions in the triple helical bundle were achieved after simulation. For helical lengths and helix register, we were guided by experimental results. Our previous NMR and EPR results indicated that the helical length of Helix C' in the bound state consisted of 36 residues (residues 17–52). Thus, we extended the helical length from residues 21–45 (Fig. 4, WT NMR structure) to include residues 17–52 (Fig. 4, $t = 0$). The helix register for coiled coils helical bundle was made easy due to heptad repeats in

Figure 4. MD structures of helical bundle A'B'C'. Left: NMR structure (PDB code: 1OWA) of the free state Helix C', with residues 21–45 forming a helical conformation. Top panel (WT), Left to right: Extended Helix C' (residues 17–52) in a triple helical bundle of A'B'C', with Helices A' (residues 2008–2037) and B' (residues 2004–2075) from β -spectrin. The template of the backbones of the three helices was from the first structural domain of α I-spectrin. This structure was the input file for explicit solvent MD simulations. The output structure after 400 ps simulations was similar to the input structure, except with a bend in Helix C' at position 41. Helix A' consists of residues 2009–2035) and Helix B' residues 2004–2070. Helix C' remains the same. The output structure after 2 ns simulation remained as a triple helix bundle. The helices in this bundle were more flexible and irregular than before. Bottom panel (L49F), Left to right: Corresponding triple helical bundle for L49F. The structure at $t = 0$ was made to be identical to that of WT, with only residue 49 changed to Phe. 400 ps MD simulations generated surprisingly large conformational changes. The most striking feature in this structure was that the backbone of Helix C' between 41 and 52 was no longer helical, and Helix C' was shortened to residues 19–40. Furthermore, Helix A' also shortened to residues 2013–2032. After 2 ns simulations, only segments of the partial domain helices remained as helical, and the triple helical bundle could no longer be identified.

sequences of Helices A', B', and C' with a and d residues in register to maximize the hydrophobic effect. In this model with an extended helical length for Helix C', both residues 24 and 49 were in register with residues with hydrophobic side chains in Helices A' and B' and in good agreement with the fluorescence results on both apparent polarity and K_{sv} values. Therefore, we concluded that we did not need to modify the helix register.

The structure from MD simulations after 400 ps was only slightly different from that before the simulation for the WT (see Fig. 4). The length for Helix C' remained the same (residues 17–52). An obvious bend in Helix C' was observed around residue 41. Consequently, the L49 residue moved closer to the V2044-

Figure 5. Detailed structural features of interhelix interactions. A: Hydrophobic cluster at residue 49 in WT (top) and L49F (bottom). In WT, LVLI cluster remains similar after 2 ns simulations. The distances between the residues are given in Table I. The distances from 49F(C γ) to 2044V(C γ), to 2047L(C δ), and to 2048I(C δ) are all within 6 Å. In L49F, the corresponding FVLI residues no longer form a cluster at 2 ns. The distances from 49F(C γ) to 2044V(C γ), to 2047L(C δ), and to 2048I(C δ) are larger than 6 Å. B: The hydrophobic pair 24I/2014F remains after 2 ns simulation in both WT and L49F. The salt bridge away from residue 49 (28R/2069E) remains after simulations. The salt bridge close to residue 49 (34R/2022E) remains in WT at 2 ns, but not in L49F. For display clarity, hydrogen atoms are hidden.

L2047-I2048 cluster in β I-C. The length for Helix A' shortened from 30 residues (2008–2031) to 27 residues (2009–2035), Helix B' from 32 residues (2044–2075) to 27 residues (2044–2070). It is interesting to find that Helix B' shortened to exclude residues 2071–2075. Our previous experimental studies indicate that Helix B' terminates at position 2070.^{3,24}

The structure of L49F after MD simulation of 400 ps was unexpected. In contrast to the WT, the MD simulation of L49F generated surprisingly large conformational changes. The most striking feature in this structure was that the backbone of Helix C' between 41 and 52 was no longer helical (see Fig. 4), and Helix C' was shortened to residues 19–40. Furthermore, Helix A' also shortened from 2009–2035 to 2013–2032. Helix B' remained the same as in the WT (residues 2044–2070).

Further MD simulation to 2 ns showed that the WT A'B'C' remained intact but more flexible, however, the L49F A'B'C' was no longer a triple helical bundle (see Fig. 4). The mutation of Leu to Phe at position 49 reduced or eliminated the stabilizing interactions of A'B'C'.

Interhelix interactions

The MD output structure of WT is not necessarily, and most likely is not, the atomic level, equilibrium solution structure of the partial domain helical bundle found in $\alpha\beta$ spectrin tetramers under physiological conditions. However, the structural dynamics provides information on potential interactions between helices. We first examined all possible interhelical hydrophobic clusters and charge–charge interactions (both attractive and repulsive) in the WT output structure at 400 ps and obtained a matrix for the distances of possible interactions. From these distances, we selected those within 6.0 Å. We found the following hydrophobic pairs between Helix C' and Helix B': L49 of Helix C'–V2044 of Helix B', L49–L2047, and L49–I2048

[Fig. 5(A)]. We believe that these large hydrophobic side chains (LVLI) form an important hydrophobic cluster, and, consequently, their van der Waals interactions contribute to the stabilization of A'B'C'. We also found another pair between Helix C' and Helix A': I24 and F2014 cluster [Fig. 5(B)]. In summary, one hydrophobic cluster (IF) at the top (N-terminal end) and one cluster (LVLI) at the bottom of the A'B'C' triple helical bundle contribute to triple helical bundle stability.

For interhelical salt-bridges, we found two pairs: R28 and E2069 (of Helix B'), and R34 and E2022 (of Helix A') [Fig. 5(B)]. One salt bridge pair linked Helix C' to Helix A' of β I-C, and another salt bridge pair linked Helix C' to Helix B' of β I-C.

For the L49F structure at 400 ps, despite the fact that the input structure was helical for the region consisting residues 46–52, this region appeared to be unstable in a helical conformation in the presence of Helices A' and B'. The stabilizing hydrophobic cluster in the WT structure was weakened in L49F [Fig. 5(A,B); Table I]. The atoms in the FVLI hydrophobic cluster are further apart, when compared with those in the LVLI cluster in the WT. The distance between I and F in the IF cluster also increased slightly. The distance also increased between R28/E2069 and R34/E2022. As these distances in Table I were from MD output structures, they were used only for relative comparison to show differences between WT and L49F. Presumably due to reduced interactions that stabilized the helices in the WT, A'B'C' of L49F was no longer a triple helical bundle after the 2 ns MD using these MD parameters.

Discussion

Properties of the fluorescent probe in free and bound helix C'

We have shown that the mBBR fluorescent probe attached to cysteine residues scanning Helix C'

Table I. Distances Between Residues of Helix C' Interacting with Helices A' and B'

WT				L49F			
Interacting residues	Distance (Å)			Interacting residues	Distance (Å)		
	<i>t</i> = 0	<i>t</i> = 400 ps	<i>t</i> = 2 ns		<i>t</i> = 0	<i>t</i> = 400 ps	<i>t</i> = 2 ns
L49 (Cδ)- V2044 (Cγ)	5.4	3.8	4.5	F49 (Cγ)- V2044 (Cγ)	6.8	4.0	13.3
L49 (Cδ)- L2047 (Cδ)	5.7	3.6	4.3	F49 (Cγ)- L2047 (Cδ)	6.0	4.7	9.1
L49 (Cδ)- I2048 (Cδ)	4.8	3.9	5.4	F49 (Cγ)- I2048 (Cδ)	6.7	4.8	7.4
I24 (Cδ)- F2014 (Cγ)	4.7	4.2	4.1	I24 (Cδ)- F2014 (Cγ)	4.7	4.7	4.4
R28 (Cζ)- E2069 (Cδ)	4.1	4.3	4.7	R28 (Cζ)- E2069 (Cδ)	4.1	5.2	4.7
R34 (Cζ)- E2022 (Cδ)	4.1	4.0	4.8	R34 (Cζ)- E2022 (Cδ)	4.1	8.1	8.2

Only those within 6 Å at *t* = 0 are shown. These distances were followed during MD simulations, at *t* = 400 ps and 2 ns, for both WT and L49F. As these distances were from MD output structures, they were used only for relative comparison with show differences between WT and L49F.

residues 21–45 as well as regions upstream (residues 14–20) and downstream (residues 46–59) of Helix C' exhibited properties (apparent polarity and KI accessibility) that correlated well with the known structure of free Helix C' in solution as determined by solution NMR studies.²

We then used the results of fluorescence measurements of Helix C' in the bound state to obtain structural information on Helix C' in the bound state. Fluorescence results showed that residues 24, 25, and 49, particularly residue 49, in the bound state were in an environment less polar and less accessible to KI than the other parts of the scanned region, suggesting that they were not only in the hydrophobic core of the helical bundle, but also formed hydrophobic clusters with residues in β-spectrin.

Structures of tetramerization helical bundle A'B'C'

The output structures from MD simulations depend on the force field used. A comparison of currently available force fields show different structural features for different force fields, and no single force field is available for all structures.²⁵ Although some force fields favor α-helix and others favor β-hairpin, OPLS-AA/L, the force field used in this study, has intermediate tendency,²⁵ and has been applied to coiled coil structures.²⁶ Using an identical simulation protocol, we were able to compare differences between WT and L49F.

Wild type. The WT model structure showed that residue L49 as part of Helix C' of α-spectrin was next to three hydrophobic residues (V2044, L2047, and I2048) from Helix B' of β-spectrin, forming an important four-member hydrophobic cluster at the C-terminal end of the bundle. This hydrophobic cluster was consistent with the experimental results showing the

extremely apolar environment and low KI accessibility for L49 in the bound state. Residue I24 of α-spectrin was close to residue F2014 of Helix A' of β-spectrin, forming another hydrophobic cluster. Thus, one cluster was at the N-terminal end of the helical bundle and one was at the C-terminal end of the bundle. Because of the close proximity of the atoms of these bulky side chains, we believe that their van der Waals interactions contribute to stabilizing the A'B'C' helical bundle.

Two salt bridges were also identified from the structure of A'B'C'. Residue R28 of Helix C' and residue E2069 of Helix B' formed the first salt bridge, and residue R34 of Helix C' and residue E2022 of Helix A' formed the second salt bridge.

It is interesting to note that Helix C' includes several charged side chains, with negatively charged D at position 51 and E at positions 22, 23, 26, 30, 40, 44, and 50, potentially interacting with R2016 of Helix A' and K2046, K2049, R2050, K2056, R2064, and K2070 of Helix B'. However, the distances between these charge pairs were larger than 6.0 Å in our model. For the positively charged R at positions 27, 28, 34, 39, and 41, and K at position 48 of α-spectrin, they can potentially interact with E2010, D2017, E2022, E2029, D2036, D2042, E2045, E2052, E2055, E2063, and E2069 of β-spectrin, and yet only two pairs (R28/E2069 and R34/E2022) were close.

L49F mutation. The substitution of L49 with a residue with a hydrophobic side chain to give L49F was not expected to cause a problem for Helix C' to bind to Helices A' and B' of β-spectrin. We have suggested that the replacement of residues in the region consisting of residues 17–52 with residues that exhibit lower helix propensity may reduce the binding affinity.⁸ Yet Phe has higher helix propensity than Leu.¹⁰ However, our MD output structure indicated that Phe

at position 49 prevented the conformational changes expected in the junction region (residues 46–52) on binding β -spectrin. It appeared that for Leu at position 49, the two methyl groups at C δ fit very well with the cluster of side chains of β I residues Val, Leu, and Ile. However, for Phe at position 49, the phenyl ring of the side chain did not fit well with the β I side chains and caused steric hindrance in the hydrophobic cluster. Rotation of this phenyl ring to maximize hydrophobic contacts appeared to disrupt the helical backbone. Thus, L49F remained in an unstructured conformation in the presence of β -spectrin. The length of both Helix C' and Helix A' were reduced in A'B'C' in L49F. Residues 41–52 assumed unstructured conformations in the bound state. Residue F49 could no longer be part of the four-member hydrophobic cluster. Thus, we suggest that, due to its inability to undergo a conformational change from unstructured to helical, on binding to β -spectrin, α -spectrin mutant L49F, and β -spectrin form lower levels of tetramers when compared with the WT.

It is interesting to note that for a protein molecule incapable of forming a helical conformation around the junction region (α I-L49F) in the bound state, the K_d value is about 30 μ M;¹² for a protein molecule with an unstructured junction region in the free state, but changing to a helical conformation in the bound state (α I-WT), the K_d value is about 1 μ M, and for a protein molecule with a helical junction region in the free state (α II-WT), the K_d value is about 10 nM.¹⁵ In summary, we have shown that our model structure of A'B'C' correlates well with the fluorescence and other published results, and appears to provide a molecular mechanism for the abnormal association of the clinical mutation of Spectrin Lyon. The final molecular interpretation of the effect of L49F leading to hematological disorder, of course, awaits further experimental confirmation.

Materials and Methods

Chemicals

L-cysteine and dioxane were from Sigma Chemical Company (St. Louis, MO) and mBBBr from Molecular Probes (Eugene, OR). All other reagents were similar to those used in our previous studies of spectrin.^{8,27}

Recombinant proteins of spectrin fragments

A cysteine-less recombinant protein of a spectrin α I fragment consisting of residues 1–368 (α I-N)^{3,8} was used to generate single cysteine proteins scanning positions 14–59 (α I-N Δ) (see Fig. 4), for example, glycine to cysteine replacement at position 14 (G14C). Because of low yields, we were not able to obtain sufficient P15C, Y35C, and F38C samples for this study. Forty-three α I-N Δ proteins were prepared. The recombinant protein of the first 368 residues, rather

than just the first 52 or 156 residues, of α -spectrin was used since this fragment exhibits a stability more similar to that of intact α -spectrin than the shorter fragments.²⁸ A C-terminal fragment of β I consisting of residues 1898–2083 (β I-C) was prepared as described previously.^{29,30} DNA sequencing and mass spectrometry analysis were performed at the Research Resources Center of the University of Illinois at Chicago.

All protein samples were monitored by SDS-PAGE for each preparation before and after experiments. The helical contents of α I-N Δ protein samples, measured by circular dichroism methods were about 70%, similar to that of the parent protein.²⁸

Fluorescence labeling and measurements

A five-fold molar excess of the fluorescent label mBBBr was added to each protein sample (60 μ M final concentration) in 5 mM phosphate buffer with 150 mM NaCl at pH 7.4 (PBS7.4), and the mixture was incubated at 4°C overnight.¹⁶ Excess mBBBr was removed by size-exclusion column chromatography. All mBBBr labeled α I-N Δ proteins were collectively referred to as α I-NB₁, following a published notation of using B₁ for the mBBBr labeled cysteine residue.¹⁷ Each individually labeled protein was referred to as, for example G14B₁. The cysteine-less protein was also processed in parallel to provide information on background labeling. The labeling efficiency (molar ratio of label to protein) for each labeled protein was determined according to published methods,^{21–24} with A₂₈₀ values for protein concentrations and A₃₈₄ values for label concentrations. About 0.15 label per protein was found to “label” cysteine-less parent protein; however, this corresponded to only about 0.5% of the total fluorescence intensity. Unbound mBBBr exhibits no fluorescence signal. Thus most of the background labeling was from labels not covalently attached to the protein, but trapped and that could not be removed with column chromatographic methods. After subtracting background labeling, we found about 0.9 ± 0.2 label per protein for all protein samples, in good agreement with reported values for single-cysteine proteins.¹⁷ The helicity values of the proteins generally remained similar before and after labeling.

The fluorescent spectra of each α I-NB₁ (10 μ M), with and without unlabeled β I-C (20 μ M), were obtained at 20°C with excitation at 384 nm and emission from 420 to 500 nm, using either a F-2000 (Hitachi, Japan) or a FP-6200 (JASCO, Japan) fluorescence spectrophotometer. Before spectral analysis, the background label spectrum was subtracted from all spectra.

Association affinity and bound state spectra

For each of the 43 α I-NB₁ proteins, native gel electrophoresis was used to screen for association affinities with β I-C that were different from that of the parent protein ($K_d \sim 1 \mu$ M).^{3,24} A mixture of each α I-NB₁

(10 μM) and $\beta\text{I-C}$ (20 μM) in 40 mM Tris, 20 mM sodium acetate, and 2 mM EDTA at pH 7.4 was used. Those $\alpha\text{I-NB}_1$ proteins that showed lower affinity than the parent protein were then studied by ITC methods, following our published methods.²⁴ Fluorescence methods were used for those with affinities below the ITC detection limit. Fluorescence intensities at 480 nm were measured for the labeled protein titrated with $\beta\text{I-C}$ until saturation. A one-site specific binding model was assumed, and K_d was obtained from curve fitting of the data, following published data treatment methods.⁸

With these K_d values, the bound state spectra were obtained from the spectra of the mixtures, by calculating the molar fractions of free and bound $\alpha\text{I-NB}_1$ of each mixture with its K_d value. For most $\alpha\text{I-NB}_1$ and $\beta\text{I-C}$ (1:2 molar ratio) mixtures with K_d values $\sim 1\ \mu\text{M}$, about 92% of $\alpha\text{I-NB}_1$ was bound. The bound state spectrum of a particular $\alpha\text{I-NB}_1$ protein was obtained by subtracting a proportion of the free spectrum from the spectrum of the mixture, and the resulting (bound) spectrum was used to determine the λ_{max} and intensity of the bound $\alpha\text{I-NB}_1$.

For one sample with low affinity, R28B₁ (K_d about 30 μM , see later), we also obtained spectra of a mixture of 6 μM R28B₁ and 270 μM $\beta\text{I-C}$. Under this condition, about 10% of R28B₁ remained free in the mixture, and we found that the calculated bound state was similar to those calculated from the 1:2 molar ratio mixture.

With all $\alpha\text{I-NB}_1$ and $\beta\text{I-C}$ mixture samples at 1:2 molar ratios and the same protein concentrations, constant sample conditions were maintained for all samples, and spectral comparison was straightforward. Furthermore, $\beta\text{I-C}$ samples tended to aggregate at high concentrations.

Residue polarity

Following a published method,¹⁸ a calibration curve using free L-cysteine labeled with mBBR (5:1 molar ratio with mBBR at 20 μM) in mixtures of varying amounts of water and dioxane was obtained to correlate the dielectric constant (or local solvent polarity) with λ_{max} values. This calibration curve was then used to obtain the apparent solvent polarity of each residue labeled with mBBR in free and bound state.

KI accessibility

Various amounts of freshly prepared KI solution (4M) in 200 μM $\text{Na}_2\text{S}_2\text{O}_3$ were added to $\alpha\text{I-NB}_1$ samples at 10 μM , without and with $\beta\text{I-C}$, to give different [KI] (0–0.36 mM). Samples were incubated for ~ 3 min before fluorescence measurements. Bound state spectra at different [KI] were obtained from the spectra of mixtures as indicated above. Intensity values at λ_{max} with no KI (F_0), and with different amounts of KI (F) were obtained for each $\alpha\text{I-NB}_1$, free or bound. The values of $[(F_0/F) - 1]$ were plotted against [KI],³¹ and

data points were fitted by linear regression, and the slopes provides the K_{sv} values.

Molecular modeling of helical bundle A'B'C'

We modified the simple, published model of the partial domain helices A'B'C' bundle² by extending the helical length of Helix C' from residues 21–45 to residues 17–52, since we have shown that the regions flanking Helix C' become helical on binding $\beta\text{I-C}$.^{7,8} The modification was done with the "Biopolymer" module of Insight II (v.2000), converting residues 17–20 and 46–52 from an unstructured to a coiled-coil helical conformation. This is the model for the WT. We also mutated residue 49 from Leu to Phe to give a model for L49F.

Explicit solvent MD simulations were applied to both WT and L49F, and were carried out with GRO-MACS (v.4.01)³² using the OPLS all-atom (OPLS-AA/L) force field³³ with a simple point charge model for water. The nonbonded atom interaction cutoff was set at 1.2 nm. The simulations were done for 2 ns with a time step of 2 fs, and structures at 400 ps and 2 ns were examined.

We also used other MD input model structures with the N-terminal end of Helix C' moved closer to Helices A' and B' (see Fig. 4, WT at $t = 0$) by a few Å, but no greater than 5 Å.

References

- DeSilva TM, Peng KC, Speicher KD, Speicher DW (1992) Analysis of human red cell spectrin tetramer (head-to-head) assembly using complementary univalent peptides. *Biochemistry* 31:10872–10878.
- Park S, Caffrey MS, Johnson ME, Fung LWM (2003) Solution structural studies on human erythrocyte alpha-spectrin tetramerization site. *J Biol Chem* 278: 21837–21844.
- Mehboob S, Luo BH, Fu W, Fung LWM (2005) Conformational studies of the tetramerization site of human erythroid spectrin by cysteine-scanning spin-labeling EPR methods. *Biochemistry* 44:15898–15905.
- Yawata Y (2003) Cell membrane: the red blood cell as a model. Wiley-VCH, Verlag GmbH & Co., KGaA, Weinheim, Germany.
- Park S, Johnson ME, Fung LWM (2002) Nuclear magnetic resonance studies of mutations at the tetramerization region of human alpha spectrin. *Blood* 100: 283–288.
- Morlé L, Roux AF, Alloisio N, Pothier B, Starck J, Denroy L, Morle F, Rudigoz RC, Forget BG, Delaunay J, Godet J (1990) Two elliptocytogenic alpha I/74 variants of the spectrin alpha I domain. Spectrin Culoz (GGT–GTT; alpha I 40 Gly–Val) and spectrin Lyon (CTT–TTT; alpha I 43 Leu–Phe). *J Clin Invest* 86:548–554.
- Long F, McElheny D, Jiang S, Park S, Caffrey MS, Fung LWM (2007) Conformational change of erythroid alpha-spectrin at the tetramerization site upon binding beta-spectrin. *Protein Sci* 16:2519–2530.
- Antoniu C, Lam VQ, Fung LWM (2008) Conformational changes at the tetramerization site of erythroid alpha-spectrin upon binding beta-spectrin: a spin label EPR study. *Biochemistry* 47:10765–10772.

9. Ranganathan S, Menhart N, Topouzian N, Fung LWM (2001) Laboratory method to study mutational effects on human erythrocyte spectrin tetramerization. *Am J Hematol* 67:247–251.
10. Pace CN, Scholtz JM (1998) A helix propensity scale based on experimental studies of peptides and proteins. *Biophys J* 75:422–427.
11. Silva PJ (2008) Assessing the reliability of sequence similarities detected through hydrophobic cluster analysis. *Proteins* 70:1588–1594.
12. Gaetani M, Mootien S, Harper S, Gallagher PG, Speicher DW (2008) Structural and functional effects of hereditary hemolytic anemia-associated point mutations in the alpha spectrin tetramer site. *Blood* 111:5712–5720.
13. Zhang Z, Weed SA, Gallagher PG, Morrow JS (2001) Dynamic molecular modeling of pathogenic mutations in the spectrin self-association domain. *Blood* 98:1645–1653.
14. Speicher DW, DeSilva TM, Speicher KD, Ursitti JA, Hembach P, Weglarz L (1993) Location of the human red cell spectrin tetramer binding site and detection of a related “closed” hairpin loop dimer using proteolytic footprinting. *J Biol Chem* 268:4227–4235.
15. Mehboob S, Jacob J, May M, Kotula L, Thiyagarajan P, Johnson ME, Fung LWM (2003) Structural analysis of the alpha N-terminal region of erythroid and nonerythroid spectrins by small-angle X-ray scattering. *Biochemistry* 42:14702–14710.
16. Mansoor SE, Mchourab HS, Farrens DL (1999) Determination of protein secondary structure and solvent accessibility using site-directed fluorescence studies of T4 lysozyme using the fluorescence probe monobromobimane. *Biochemistry* 38:16383–16393.
17. Mansoor SE, Mchaourab HS, Farrens DL (2002) Mapping proximity within proteins using fluorescence spectroscopy. A study of T4 lysozyme showing that tryptophan residues quench bimane fluorescence. *Biochemistry* 41:2475–2484.
18. Kosower EM, Giniger R, Radkowsky A, Hebel D, Shusterman A (1986) Bimanes 22 flexible fluorescent molecules solvent effects on the photophysical properties of syn-bimanes (1,5-diazabicyclo[3.3.0]octa-3,6-diene-2,8-diones). *J Phys Chem* 90:5552–5557.
19. Mansoor SE, Farrens DL (2004) High-throughput protein structural analysis using site-directed fluorescence labeling and the bimane derivative (2-pyridyl)dithiobimane. *Biochemistry* 43:9426–9438.
20. Sommer ME, Smith WC, Farrens DL (2005) Dynamics of arrestin-rhodopsin interactions: arrestin and retinal release are directly linked events. *J Biol Chem* 280:6861–6871.
21. Wei Z, White D, Wang J, Musse AA, Merrill AR (2007) Tilted, extended, and lying in wait: the membrane-bound topology of residues Lys-381-Ser-405 of the colicin E1 channel domain. *Biochemistry* 46:6074–6085.
22. Musse AA, Wang J, DeLeon GP, London E, Merrill AR (2006) Scanning the membrane-bound conformation of helix 1 in the colicin E1 channel domain by site-directed fluorescence labeling. *J Biol Chem* 281:885–895.
23. Sosa-Peinado A, González-Andrade M (2005) Site-directed fluorescence labeling reveals differences on the R-conformer of glucosamine 6-phosphate deaminase of *Escherichia coli* induced by active or allosteric site ligands at steady state. *Biochemistry* 44:15083–15092.
24. Mehboob S, Luo BH, Patel BM, Fung LWM (2001) $\alpha\beta$ spectrin coiled-coil association at the tetramerization site. *Biochemistry* 40:12457–12464.
25. Yoda T, Sugita Y, Okamoto Y (2004) Comparisons of force fields for proteins by generalized-ensemble simulations. *Chem Phys Lett* 386:460–467.
26. Bjelic S, Wieninger S, Jelesarov I, Karshikoff A (2008) Electrostatic contribution to the thermodynamic and kinetic stability of the homotrimeric coiled coil Lpp-56: a computational study. *Proteins* 70:810–822.
27. Li Q, Fung LWM (2009) Structural and dynamic study of the tetramerization region of non-erythroid α -spectrin: a frayed helix revealed by site-directed spin labeling electron paramagnetic resonance. *Biochemistry* 48:206–215.
28. Lusitani DM, Qtaishat N, LaBrake C, Yu RN, Davis J, Kelley MR, Fung LWM (1994) The first human alpha-spectrin structural domain begins with serine. *J Biol Chem* 269:25955–25958.
29. Luo BH, Mehboob S, Hurtuk MG, Pipalia NH, Fung LWM (2002) Important region in the beta spectrin C-terminus for spectrin tetramer formation. *Eur J Haematol* 68:73–79.
30. Lecomte MC, Nicolas G, Dhermy D, Pinder JC, Gratzer MC (1999) Properties of normal and mutant polypeptide fragments from the dimer association sites of human red cell spectrin. *Eur Biophys J* 28:208–215.
31. Lakowicz JR (1999) Principles of fluorescence spectroscopy, 2nd edition. New York: Kulwer Academic/Plenum Publishers.
32. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 4:435–447.
33. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J Phys Chem B* 105:6474–6487.