Structural Insights into the Stability and Flexibility of Unusual Erythroid Spectrin Repeats

Hideki Kusunoki, Ruby I. MacDonald,* and Alfonso Mondragón*
Department of Biochemistry, Molecular Biology, and Cell Biology
Northwestern University
2205 Tech Drive
Evanston, Illinois 60208

Summary

Erythroid spectrin, a major component of the cytoskeletal network of the red cell which contributes to both the stability and the elasticity of the red cell membrane, is composed of two subunits, \( \alpha \) and \( \beta \), each formed by 16–20 tandem repeats. The properties of the repeats and their relative arrangement are thought to be key determinants of spectrin flexibility. Here we report a 2.4 Å resolution crystal structure of human erythroid \( \beta \)-spectrin repeats 8 and 9. This two-repeat fragment is unusual as it exhibits low stability of folding and one of its repeats lacks two tryptophans highly conserved among spectrin repeats. Two key factors responsible for the lower stability and, possibly, its flexibility, are revealed by the structure. A third novel feature of the structure is the relative orientation of the two repeats, which increases the range of possible conformations and provides new insights into atomic models of spectrin flexibility.

Introduction

Spectrin is a ubiquitous, flexible rod-like protein which plays an important role in conferring shape and stability on metazoan cells (Bennett and Baines, 2001; De Matteis and Morrow, 2000; Lux and Palek 1995; Mohandas and Evans, 1994). The best-known isoform, erythroid spectrin, constitutes 70% of the cytoskeletal network of the red cell membrane and helps to maintain red cell shape and elasticity. The basic molecular organization of erythroid (De Matteis and Morrow, 2000) and nonerythroid (Bennett and Baines, 2001) spectrin isoforms is similar: they are comprised of two subunits, \( \alpha \) (280 kDa) (Sahr et al., 1990) and \( \beta \) (246 kDa) (Winkelmann et al., 1990), which form antiparallel heterodimers by side-to-side association (Speicher et al., 1992) and tetramers by head-to-head self-association of the heterodimers (DeSilva et al., 1992). The elongated shape of spectrin is due to its arrangement as 16–20 tandem repeating units, or spectrin repeats, each one approximately 106 residues long and with 20%–30% sequence similarity (Speicher and Marchesi, 1984). The repeats are joined by a short region consisting of five residues, which is found at the C terminus of each repeat and termed the linker region. Each repeat folds into a left-handed, antiparallel triple-helical coiled coil displaying a repetitive heptad pattern in its primary sequence that is associated with coiled-coil structures (McLachlan and Stewart, 1975). A characteristic of the spectrin repeat is the presence of tryptophan residues in the first and the third helices. These tryptophans are highly conserved, particularly in the first helix, and contribute to the conformational stability of the repeats (MacDonald et al., 1994; Pantazatos and MacDonald, 1997).

The conformation of the linker region has been the focus of models of spectrin flexibility (Bloch and Pumplin, 1992; Speicher and Marchesi, 1984). Therefore, it was of great interest that the linkers were ordered and \( \alpha \) helical in all crystal structures of a two-repeat spectrin fragment (Grum et al., 1999). This observation of an ordered linker does not preclude the presence of nonhelical linkers between some repeats, particularly in erythroid spectrin. In fact, the linker between the first complete repeat of human erythroid \( \alpha \)-spectrin and the single helix of the incomplete repeat at its N terminus is nonhelical, as shown by its NMR structure (Park et al., 2003). This region, however, may be atypical as it is involved in tetramer formation. Moreover, a disordered conformation is likely to promote proper pairing of the \( \alpha \) and \( \beta \) subunits and therefore is not an appropriate model for other repeats of spectrin. Secondary structure analyses of the linker region of aligned chicken brain and human erythroid spectrin sequences predict either a helical or nonhelical conformation for some repeats (MacDonald and Cummings, 2004), depending on the method used (Jones, 1999; King and Stemberg, 1996). Among the fragments identified with a putatively nonhelical linker is human erythroid \( \beta \)-spectrin repeats 8 and 9 (HE\( \beta \)89). Interestingly, one of these repeats, repeat 9, is also one of the few repeats without the highly conserved tryptophans in the first and the third helices. Additionally, HE\( \beta \)89 has been shown to be less stably folded than other fragments previously examined (MacDonald and Cummings, 2004; Pantazatos and MacDonald, 1997). Thus, HE\( \beta \)89 seems to be a good candidate to examine for the presence of a nonhelical linker and for investigation of the structural role of the highly conserved tryptophans in promoting folding and stability of the spectrin repeat.

Here we report the crystal structure of HE\( \beta \)89, the first structure of two repeats of human erythroid \( \beta \)-spectrin. The crystal structure clearly shows an \( \alpha \)-helical linker between the repeats. Comparison of the structure of HE\( \beta \)89 with that of repeats 16 and 17 of chicken brain \( \alpha \)-spectrin (CB\( \alpha \)1617) reveals the presence of a significant cavity in HE\( \beta \)89 due to the absence of the highly conserved tryptophans in repeat 9 and also a reduced number of canonical intrahelical hydrogen bonds. These features provide an explanation for the reduced thermal stability of HE\( \beta \)89 and support the role of the highly conserved tryptophans in folding and stability of the repeats. Equally importantly, the structure expands the range of conformational variability of a previously proposed molecular model of spectrin flexibility.

*Correspondence: rubymacd@northwestern.edu (R.I.M.), a-mondragon@northwestern.edu (A.M.)
Table 1. Data Collection and Refinement Statistics

<table>
<thead>
<tr>
<th></th>
<th>SeMet</th>
<th>Native</th>
<th>Edge</th>
<th>Peak</th>
<th>Remote</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.005</td>
<td>0.9976</td>
<td>0.9794</td>
<td>0.9794</td>
<td>0.9642</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>2.4</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td><strong>Reflections</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>95,548</td>
<td>42,477</td>
<td>42,508</td>
<td>42,442</td>
<td>42,442</td>
</tr>
<tr>
<td>Unique</td>
<td>15,165</td>
<td>13,069</td>
<td>13,083</td>
<td>13,068</td>
<td>13,068</td>
</tr>
<tr>
<td>Completeness (%)a</td>
<td>99.7  (99.8)</td>
<td>98.6 (98.4)</td>
<td>98.8 (99.0)</td>
<td>98.5 (97.9)</td>
<td>98.5 (97.9)</td>
</tr>
<tr>
<td>(I/I_{ave})a</td>
<td>7.6  (2.7)</td>
<td>12.6 (3.7)</td>
<td>12.5 (3.8)</td>
<td>12.0 (3.4)</td>
<td>12.0 (3.4)</td>
</tr>
<tr>
<td><strong>Rsym (%)a,b</strong></td>
<td>8.1  (27.9)</td>
<td>5.8 (21.0)</td>
<td>5.8 (20.5)</td>
<td>6.1 (22.7)</td>
<td>6.1 (22.7)</td>
</tr>
<tr>
<td><strong>Rmeas (%)a,c</strong></td>
<td>8.8  (30.3)</td>
<td>7.6 (27.1)</td>
<td>7.5 (26.4)</td>
<td>7.9 (29.3)</td>
<td>7.9 (29.3)</td>
</tr>
<tr>
<td><strong>Phasing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of sites</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of sites</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Figure of merit</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before density mod</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After density mod</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)a</td>
<td>45.9–2.4 (2.55–2.40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R_{bri} )a</td>
<td>22.2 (30.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R_{max} )a</td>
<td>25.9 (34.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of reflections</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working set</td>
<td>14,388</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test set</td>
<td>746</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Model composition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residues</td>
<td>211</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>182</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sulfate ions</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rms bond (Å)</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rms angle (°)</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average B factor (Å)</td>
<td>45.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ramachandran analysis (%)</strong></td>
<td>69.8</td>
<td>3.2</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Most favored</td>
<td>96.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Additional allowed</td>
<td>3.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generously allowed</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disallowed</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

aNumbers in parentheses are for the highest resolution shell.
b\(R_{sym} = \frac{\sum{I} - \langle I \rangle}{\langle I \rangle} \), where \(I\) = observed intensity and \(\langle I \rangle = \) average intensity obtained from multiple measurements.
c\(R_{meas}\) as defined by Diederichs and Karplus (Diederichs and Karplus, 1997).

d\(R_{work}\) is the R factor based on \(95\%\) of the reflections used in refinement.
e\(R_{free}\) is the R factor based on \(5\%\) of the reflections not used in refinement.

Results

Structure Determination and Overall Structure of HEβ89

The crystal structure of HEβ89 was solved by MAD phasing from a selenomethionine-substituted protein crystal and was refined to a final \(R_{work}\) of 22.2% and \(R_{free}\) of 25.9% at 2.4 Å resolution (see Experimental Procedures and Table 1). The asymmetric unit contains one molecule and the final model includes 211 residues (amino acids 1083–1273), 182 water molecules, and 2 sulfate ions. The model has excellent stereochemistry; there are no outliers and all residues occupy most favored and additional allowed regions in the Ramachandran plot (Table 1).

The overall structure of HEβ89 consists of two repeats (termed R8 and R9) connected by an extended α helix (Figure 1). The molecule has an elongated shape with a length of about 103 Å and a width of about 21 Å. The relative orientation of the two repeats is described by a translation of approximately 50 Å along the long axis of the molecule and no rotation around the same axis. Each repeat consists of three helices and two loops. We refer to the helices in the first repeat as A, B, and C, and to the helices in the second repeat as A', B', and C'. Helices C and A' form a single long helix that includes the linker region. Helices A, C, A', and C' are slightly bent, while helices B and B' are kinked at their midpoints. The three helices in each repeat show a characteristic heptad pattern and interact with each other through residues α and δ, which are generally hydrophobic, lie on the inward-facing surface of the helix, have smaller solvent accessible areas, and create a closely packed core region.

Comparison of HEβ89 and CBβ1617

The overall structure of HEβ89 is similar to that of CBβ1617 (Figure 2B). However, it is not possible to simply superpose each pair of repeats, as the relative orientation (tilt, roll, and twist angles) of the two repeats in
HEβ89 is quite different from that of the two repeats in CBα1617 (termed R16 and R17). In HEβ89, the tilt angle is about $-11.6^\circ$, the roll angle is about $3.8^\circ$, and the twist angle is about $0.2^\circ$. In the five structures of CBα1617, tilt angles range from $0.5^\circ$ to $22.4^\circ$, roll angles range from $-3.3^\circ$ to $9.5^\circ$, and twist angles range from $27.9^\circ$ to $52.5^\circ$ (Grum et al., 1999). The relative orientation of the repeats in HEβ89 was determined by the same procedure as described previously (Grum et al., 1999).

When individual repeats of HEβ89 and CBα1617 are compared, the rms differences are small (rms differences ranging from $1.07$ to $1.59$ Å for 98 to 106 Cα atoms), indicating that the individual repeats are nearly identical (Figure 2C). In R8, the BC loop and the N terminus of helix C differ significantly from those of the other repeats: the BC loop adopts a different conformation, and two extra residues are added at the N terminus of helix C (Figures 2A and 2C). A change in the phasing of the heptad pattern corresponding to a deletion of four residues and referred to as a stammer (for definition see Brown et al., 1996) that was observed in the second helices of R16 and R17 (Grum et al., 1999) is not observed in those of R8 and R9 (Figure 2A), suggesting that this stammer is not a universal feature of the spectrin repeats.

Helical Linker Region
The crystal structure directly answers the question regarding the conformation of the linker region of HEβ89, as it displays an α-helical linker without any breaks or discontinuities (Figures 1 and 3A). The α-helical linker consists of five residues and enables helix CA’ to form...
Figure 2. Comparison of the Structures of HE\(\epsilon\)89 and CB\(\epsilon\)1617
(A) Structure-based amino acid sequence alignment of individual repeats of HE\(\epsilon\)89 (SWISS-PROT P11277) and CB\(\epsilon\)1617 (SWISS-PROT P07751). Secondary structure elements were assigned according to the outputs of the program PROCHECK (Laskowski et al., 1993). To delineate the secondary structural elements, the helices in each repeat are enclosed by colored boxes. R8, R9, R16, and R17 are shown in blue, cyan, yellow, and green, respectively. The linker region of HE\(\epsilon\)89 and CB\(\epsilon\)1617 are marked by red and purple boxes, respectively. The highly conserved tryptophan residues in helix A of each spectrin repeat and the corresponding residue in R9 (Ile1183) are shown in red. Residues interacting with the tryptophan and the equivalent residue in R9 are also shown in red. The a and d positions in the heptad patterns

\[
\begin{align*}
\text{R8} & \quad 1063 \quad \text{LQPQGQLDELDFQN} & \text{N} & \text{SITQKVNSEDMPESTFPAAEQILQQHAGL} & \text{D} & \text{E} & \text{1116} \\
\text{R9} & \quad 1170 \quad \text{FGFQKDAKQAEALSHQQYTLAHLEPPDLSEAAAG} & \text{N} & \text{R16} & \quad 1771 \quad \text{LHQQFRMDDESWMDIKLVLSSEDYGRDLTGVQLRKKLRTEAELAAH} & \text{D} & \text{1223} \\
\text{R17} & \quad 1877 \quad \text{YQQFVANVEEALLNEKWLVAEDYGDA1ATQGL} & \text{N} & \text{D} & \text{1824} \\
\end{align*}
\]
the longest continuous α helix in the structure. The heptad pattern in R8 does not continue through the linker to R9, even allowing for possible stammers, skips, or stutters (Brown et al., 1996) for the following reason. The heptad pattern of helix C terminates with Leu1164 as an α residue while the pattern of helix A begins with Phe1170 as an α residue. If the heptad pattern continued from one repeat to the next, the helical linker would have to consist of six residues instead of the observed five. These features of HEβ89 are very similar to those observed in CBα1617: the existence of an α-helical linker consisting of five residues forming the longest continuous α helix, and the discontinuity of the heptad pattern (Grum et al., 1999).

In HEβ89, residues in the α-helical linker interact mainly with those in loop AB and helix B in R8 and residues in loop B’C’ in R9 (Figure 3B). Gln1166 hydrogen bonds to Asn1239 and Leu1240 in loop B’C’. Cys1167 and Leu1168 make van der Waals contacts with Met1089 in loop AB and Leu1100 in helix B, respectively. In CBα1617, Glu1873 hydrogen bonds to Asn1945 and His1947 in the loop corresponding to loop B’C’. Ser1874 hydrogen bonds to Tyr1797 in the loop corresponding to loop AB (Figure 3C) (Grum et al., 1999). The interactions observed in the linker regions of both HEβ89 and CBα1617 involve equivalent residues and are only slightly different, suggesting that they may help to maintain the helical structure of the linker.

Interactions around the Highly Conserved Tryptophan
The structures of α-spectrin repeats revealed that the highly conserved tryptophan residue (present in 14 out of 16 complete repeats in erythroid β-spectrin) in the first helix interacts with a d residue in the middle of the second helix and with a less highly conserved tryptophan residue (present in 9 out of 16 complete repeats in erythroid β-spectrin) at the a position of the third helix (Grum et al., 1999; Park et al., 2003; Pascual et al., 1997; Yan et al., 1993). In earlier studies, it was shown that both tryptophan residues promote the conformational stability of a repeat (MacDonald et al., 1994; Pantazatos and MacDonald, 1997). As R9 of HEβ89 does not possess these tryptophans, the role of these highly conserved residues could be examined by comparing the structure of R9 with that of R8 and the two repeats of CBα1617.

In R8, Trp1076 stacks on His1113, while Trp1076 interacts with Trp1150, resulting in a closely packed core region (Figure 4A). In R9, Ile1183, the equivalent of Trp1076, interacts with Asn1220. However, unlike Trp1076 and Trp1150 in R8, Ile1183 and His1256 are far...
from each other in R9 so that a large cavity is created in the core, which contains two water molecules (Figure 4B). In R16, the equivalent interaction involves Trp1784/His1821/Trp1857 (Figure 4C) and is almost the same as that of R8. In R17, this interaction involves Trp1890/His1927/Val1963. However, R17 has a valine (Val1963) instead of the conserved tryptophan in the third helix, and as a result, the core of R17 displays a much smaller cavity than in R9 and contains no water molecules (Figure 4D).

In order to stabilize their structures, heptad repeat structures generally form closely packed hydrophobic cores with residues a and d on the inward-facing surface. In contrast, the above core is formed by g, d, and a, residues, with the g residues largely exposed to the solvent. To form a closely packed core, amino acids with bulky side chains, like tryptophan, are preferred at the g position in the middle of the first helix and at the position in the middle of the third helix, respectively (Figure 2A). R9 in HE/H9252 does not contain a tryptophan residue in the first helix or in the third helix and thus possesses a large cavity in its core (Figure 4B). In addition, two water molecules are found in the large cavity of R9 (Figure 4B) and form hydrogen bonds with polar atoms on the inward-facing surface of the cavity. It has been previously shown that a cavity in a hydrophobic core can contain water molecules which hydrogen bond to polar atoms of the core (Buckle et al., 1996; Takano et al., 1997, 2003). Moreover, such water molecules may contribute to the conformational stability of a protein (Buckle et al., 1996; Takano et al., 1997, 2003). For example, in cavity-creating mutants of lysozyme, those with solvated cavities are more stable than those with empty cavities, although they are still less stable than wild-type protein (Takano et al., 1997, 2003). In addition, a cavity-filling mutation can help to stabilize a protein structure. For instance, wild-type Myb, a DNA binding protein consisting of three imperfect repeats, one of which exhibits decreased thermal stability attributed to a cavity in the hydrophobic core (Ogata et al., 1996). A valine to leucine (V103L) mutant can fill this hole and leads to higher stability (Ogata et al., 1996). Taken together, these observations suggest that buried water molecules can confer stability, although not as much as the nonpolar protein atoms forming part of the hydrophobic core. The observation of a large cavity partially filled with two water molecules in R9 suggests that the lower stability of HE/H9252 may be due, at least partially, to its more loosely packed hydrophobic core than that of other spectrin repeats. Thus, the tryptophans at the g position and at the a position are crucial for the formation of a closely packed core and a more stable protein.

Loss of Canonical Intrahelical Hydrogen Bonds in the Second Helices

In HE/H9252, both B helices are kinked at their midpoints, resulting in loss of canonical intrahelical hydrogen bonds (\( \approx 3.5 \text{ Å} \)) at this position. In helix B of R8, two canonical intrahelical hydrogen bonds are lost: the distance between the backbone nitrogen of Asp1115 and the carbonyl oxygen of Asp1119 is 4.6 Å, and the distance between the backbone nitrogen of Gly1122 and the carbonyl oxygen of Ser1116 is 5.0 Å (Figure 5A). Interestingly, the backbone nitrogens of both Asp1115 and Gly1112, as well as the carbonyl oxygens of both Asp1115 and Ser1116, form hydrogen bonds with water molecules (W1 and W2). The side chain oxygen of Asp1115 also hydrogen bonds to a water molecule (W1). Helix B’ of R9 has a similar hydrogen bonding pattern to R8 (Figure 5B), but additionally, this helix contains a proline (Pro1227) near the kink (Figure 5C). Proline lacks an amide proton and, hence, cannot form an intrahelical...
Figure 5. Stereoviews of the B Helices in All the Repeats around the Kink Region

Residues are drawn as balls-and-sticks; water molecules are shown as red spheres and are labeled from W1 to W5. Missing canonical intrahelical hydrogen bonds are indicated by black dotted lines and by green dotted lines for proline residues. Canonical intrahelical, protein-water, and C-H–O hydrogen bonds involving proline residues are indicated by orange dotted lines (all are ~3.5 Å in length).

(A) Residues from Asp1111 to Ser1116 in R8. The distance between the backbone oxygen of Asp1111 and the backbone nitrogen of Asp1115 is 4.6 Å. The distance between the backbone oxygen of Gly1112 and the backbone nitrogen of Ser1116 is 5.0 Å.

(B) Residues from Glu1218 to Lys1223 in R9. The distance between the backbone oxygen of Glu1218 and the backbone nitrogen of Asp1222 is 4.9 Å. The distance between the backbone oxygen of Asn1219 and the backbone nitrogen of Lys1223 is 5.1 Å.

(C) Residues from Lys1223 to Val1228 in R9. The distance between the backbone oxygen of Val1224 and the backbone nitrogen of Val1228 is 4.2 Å.

(D) Residues from Ala1819 to Ala1824 in R16. The distance between the backbone oxygen of Ala1820 and the backbone nitrogen of Ala1824 is 5.1 Å.

(E) Residues from Thr1925 to Arg1930 in R17. Canonical intrahelical hydrogen bonds are present between Thr1925 and Asp1929 and between Val1926 and Arg1930.
hydrogen bond between the amide and carbonyl groups, so that it may act as a helix breaker (O’Neil and DeGrado, 1990). Nevertheless, it has been reported recently that the Cα proton of a proline residue in a helix can form a C–H···O hydrogen bond so as to stabilize a helix (Chakrabarti and Chakrabarti, 1998). Most of those hydrogen bonds are formed by the proline and the carbonyl oxygen of a preceding amino acid, either three (type 3), four (type 4), or five (type 5) residues away (Chakrabarti and Chakrabarti, 1998). In R9, both a type 3 and a type 4 hydrogen bond are present in helix B and may help to stabilize it (Figure 5C). However, two canonical intrahelical hydrogen bonds are lost between Lys1223 and Pro1227 and between Val1224 and Val1228 (Figure 5C). Thus, HEβ89 loses six canonical intrahelical hydrogen bonds but forms eleven protein-water hydrogen bonds and a type 3 and a type 4 C–H···O hydrogen bond.

In CBα1617, helix B in R16 is also kinked and contains a proline residue. As a result, two canonical intrahelical hydrogen bonds are lost between Ala1819 and Pro1823 and between Ala1820 and Ala1824, and two protein-water hydrogen bonds and a type 3 and a type 4 hydrogen bond are formed (Figure 5D). The corresponding helix in R17 is not kinked, and the intrahelical hydrogen bonds are all present (Figure 5E). CBα1617 loses two canonical intrahelical hydrogen bonds and forms two protein-water hydrogen bonds and a type 3 and a type 4 C–H···O hydrogen bond.

Protein-water hydrogen bonds are observed in the B helices of R8, R9, and R16, which may contribute to stabilizing such helices. However, protein-water hydrogen bonds are likely to be less stabilizing than canonical intrahelical ones, because water molecules have high entropy associated with solvent ordering. The importance of canonical intrahelical hydrogen bonds for thermal stability of proteins has been reported. For example, the structure of wild-type Escherichia coli Fis protein contains a C–H···O hydrogen bond but lacks a canonical intrahelical hydrogen bond between residues Glu57 and Pro61 (Kostrewa et al., 1991; Yuan et al., 1991), while in the structure of a stable Fis P61A mutant, a canonical intrahelical hydrogen bond is formed (Yuan et al., 1994). As a result, the Fis P61A mutant becomes thermodynamically more stable than the wild-type protein (Yuan et al., 1994), strongly suggesting that the loss of canonical intrahelical hydrogen bonds may decrease structural stability, even though protein-water hydrogen bonds and C–H···O hydrogen bonds are formed. HEβ89 loses six canonical intrahelical hydrogen bonds, whereas CBα1617 loses only two canonical intrahelical ones. This difference indicates that a second major reason for the reduced thermal stability of HEβ89 may be the presence of fewer canonical intrahelical bonds, although some of them are replaced by other types of interactions.

Discussion

Structural Basis for the Lower Thermal Stability of HEβ89
We have solved the first crystal structure of two repeats of human erythroid β-spectrin, HEβ89, to 2.4 Å resolution. Although the linker region of HEβ89 was predicted to be nonhelical, the structure clearly shows it to be α-helical (Figures 1 and 3). In contrast, CBα1617 also possesses an α-helical linker (Figure 2B) but is thermodynamically more stable than HEβ89. Thus, the lower thermal stability of HEβ89 is not attributable to a nonhelical linker region.

Two factors appear to be important for the lower stability of HEβ89. The first factor is the presence of a large cavity with two water molecules in the hydrophobic core in R9. This cavity is presumably due to the fact that R9 does not contain the highly conserved tryptophans in its first and third helices (Figure 4B). This cavity is the first structural evidence for the importance of the tryptophan residues for the conformational stability of spectrin repeats (Figure 4). A second factor is the loss of six canonical intrahelical hydrogen bonds in the B helices of HEβ89 (Figures 5A–5C). Of the six intrahelical hydrogen bonds lost from HEβ89, four are missing from helix B of R9 (Figures 5B and 5C), strongly suggesting that R9 is the repeat that is thermodynamically less stably folded. It is noteworthy that the loss of canonical intrahelical hydrogen bonds occurs in the vicinity of the large cavity in the core, suggesting that the two factors may cooperatively decrease the thermal stability of HEβ89.

Relationship between Flexibility and Lower Stability of Spectrin Repeats
Recently, the S2 segment of myosin II has been reported to form a flexible coiled coil with poorly packed core residues (Li et al., 2003). In the crystal structure, two lysine side chains at a position in the heptad pattern adopt variable conformations in the two molecules in the asymmetric unit due to the poorly packed core. Thus, the poorly packed core regions are likely to be relevant to the flexibility of the S2 segment (Li et al., 2003). Since spectrin repeats form a triple helical coiled coil, and R9 of HEβ89 exhibits a large cavity containing two water molecules in the hydrophobic core, our observation parallels that in myosin II and suggests that R9 may contribute to spectrin flexibility through its poorly packed core region. Atomic force microscopy (AFM) and molecular dynamic simulation (MD) experiments (Altmann et al., 2002) have suggested that partial unfolding of the spectrin repeats may be an important source of flexibility. In these experiments, spectrin repeats subjected to external forces could form unfolding intermediates which were associated with kinking of the second helix, close to the proline residue in R16. (Altmann et al., 2002). In contrast, a mutation stabilizing the second helix, P1823A/G1827A, resulted in higher thermal stability than the wild-type repeat and no unfolding intermediates were observed by AFM or in MD simulations (Altmann et al., 2002). These observations are consistent with the structure of HEβ89. Destabilizing the second helix of a spectrin repeat by loss of canonical hydrogen bonds may trigger unfolding intermediates and constitute a source of flexibility.

Fluorescence polarization (Yoshino and Marchesi, 1984) and 1D 1H-NMR (Begg et al., 1994) studies on intact spectrin have suggested that certain repeats are not stably folded or are partially unfolded under physiological conditions. The midpoint temperature of un-
An Atomic Model of Spectrin Flexibility Model Based on the Structure of HEβ89

Two atomic level models of spectrin flexibility have been suggested by the crystal structures of CBα:1617 (Grum et al., 1999): a conformational rearrangement model and a bending model. The conformational rearrangement model is based on the observation that a loop between two helices becomes ordered to form part of an adjacent helix, while a neighboring helical region is melted to replace the loop. The bending model is based on the observation that the relative orientation of two repeats can differ without a change in secondary structure or major rearrangement of the linker region. It is likely that both mechanisms are important and that spectrin flexibility is achieved through a combination of both. The HEβ89 structure confirms that spectrin repeats can accommodate different relative orientations, in support of a major assumption of the bending model. In the bending model based on the structures of CBα:1617, the different tilt and roll angles between tandem repeats played a major role. The HEβ89 structure allows us to propose a new variation or refinement of the bending model based on the observation of a twist angle of almost 0°. In this refined model, differences in twist between repeatable repeats promote further variability in the arrangements of repeats so as to produce a large number of possible outcomes.

Models of tandem repeats mimicking β-spectrin were built by the repetitive superposing of two repeats (see Experimental Procedures). If the twist angle is almost 0° between all repeats, the model forms a semicircle (Figure 6A) due to the lack of twist between the repeats. On the other hand, if the twist angle is about 30° between all repeats, as in one structure of CBα:1617 (PDB code 1CUN), the model forms an extended conformation with a gentle supercoil (Figure 6B). When the model contains one or two pairs of repeats with a twist angle of almost 0°, as in HEβ89, while all other repeats adopt a conformation similar to CBα:1617, a variety of conformations result. For example, when a model contains HEβ89 at repeats 8 and 9, its trajectory begins to change (Figure 6C) from that of the extended one in Figure 6B without repeats 8 and 9. Moreover, when a model contains two HEβ89 pairs at repeats 4 and 5 and at repeats 8 and 9, as in β-spectrin, its trajectory is dramatically altered (Figure 6D). Additional combinations of twist angles would result in further unique conformations of spectrin.

A twist angle of 60° between repeats was first suggested from a model of two repeats based on the structure of a single repeat of Drosophila α-spectrin (Yan et al., 1993). Subsequently, a wide range of twist angles, from 27.9° to 52.5°, was observed in the structures of four slightly different constructs of CBα:1617 (Grum et al., 1999). The structure of HEβ89 further expands the range of twist angles to that of almost 0°. In addition, twist angles of \( \sim100° \) have been observed in the crystal structures of repeats of α-actinin (Djilovic-Carugo et al., 1999; Ylänne et al., 2001), which is a member of the spectrin superfamily. Spectrin repeats seem to be able to adopt a much wider range of twist than previously thought. If spectrin repeats can assume such a wide range of twist angles, from 0° to 100°, and the twist can vary in the same two repeats, as observed in structures...
of CBα1617 (Grum et al., 1999), this would lead to a variety of conformations for spectrin at the atomic level, in agreement with electron microscopy studies (Shotton et al., 1979).

A helical conformation of the linker region is compatible with a variety of different twist angles between spectrin repeats, leading to many different conformations. In addition, variable tilt and roll angles would add to the number of possible conformations, further increasing the repertoire of spectrin conformations. Thus, the present structure of tandem repeats of β-spectrin together with previous structural evidence of a helical linker in chicken brain α-spectrin and α-actin (Dinjovic-Carugo et al., 1999; Grum et al., 1999; Yläne et al., 2001) strongly suggest that an α helix is the preferred conformation for the linker. It is likely that spectrin uses various combinations of mechanisms to achieve its flexibility, for example, conformational rearrangements within repeats, relative bending of the repeats, and partial unfolding of the repeats. Additional biophysical and structural studies are needed to characterize more fully the range of variability of conformations of spectrin repeats and to ascertain whether other mechanisms could play a role in spectrin flexibility. For instance, the possibility of melting of the helical linker has been suggested recently (Law et al., 2003) to explain results of AFM and CD experiments on β-spectrin repeats 1–4, indicating a scenario in which the linker region is frequently helical at low temperatures, but becomes an unfolded coil at higher temperature or at the Tm (Law et al., 2003). It is not certain, however, whether the helix-to-coil transition occurs within the helices of the three-helix coiled coil or in the linker region, since direct evidence of melting of the linker was not reported. Finally, although we cannot rule out the possibility that the helical linker could melt under as yet undefined conditions or that a nonhelical linker could exist between some repeats, the possibility of a nonhelical linker between fully formed repeats appears increasingly to be less likely.

Experimental Procedures

Overexpression and Protein Purification

HEβ89 (residues 1063–1275) was overexpressed in Escherichia coli BL21(DE3) and purified by DEAE and Q-sepharose anion exchange chromatography and S-100 gel filtration chromatography as previously described (Grum et al., 1999). SeMet-HEβ89 was prepared using the methionine pathway inhibition method (Van Duyne et al., 1993) and purified in the same manner as the native HEβ88 protein. HEβ89 was more than 95% pure, which was confirmed by SDS-PAGE in 3 M urea. For crystallization, purified protein was concentrated to 15–30 mg ml⁻¹ in 10 mM HEPES (pH 7.5) using a Centriplus-10 (Amicon) and stored at 4°C prior to crystallization.

Crystallization and Structure Determination

Crystals of both native and SeMet-substituted proteins were obtained by the hanging-drop vapor-diffusion method. Crystals were grown at 4°C in a solution containing 0.1 M sodium citrate (pH 6.5), 0.3 M ammonium sulfate, and 1.2–1.3 M lithium sulfate. Crystals grew to approximately 0.06–0.08 × 0.08 × 0.8 mm³ within 1 week. For data collection, crystals were cryo protected by serial transfer into the reservoir solution supplemented with 5%, 10%, 15%, 20%, 25%, and 30% glycerol for 30–60 s/step and then instantly frozen in liquid nitrogen. Diffraction data were collected at the DuPont Northwestern Dow Collaborative Access Team (DND-CAT) beamline at the Advanced Photon Source (APS) using a MAR CCD detector, integrated with the program XDS (Kabsch, 1993), and scaled with SCALA (CCP4, 1994). The crystals belong to the tetragonal space group P4₁2₁2₁, with unit cell parameters, a = b = 122.15 Å and c = 49.5 Å with one molecule in the asymmetric unit. SeMet-HEβ89 crystals were nearly isomorphous to native crystals.

The structure of HEβ89 was solved by a MAD experiment using data from a SeMet-HEβ89 crystal collected to 2.5 Å resolution. Of the four possible selenium sites in SeMet-HEβ89, three sites were found with the program SOLVE (Terwilliger and Berendzen, 1999). Phases were calculated with the program SHARP (de la Fortelle and Bricogne, 1997) and further improved with the program SOLOMON (Abrahams and Leslie, 1996) and DM (Cowtan, 1994) as implemented in SHARP. The resulting map was sufficiently clear to build an initial model of the molecule using O (Jones et al., 1991). The model was refined using native data at 2.4 Å resolution with the program CNS (Brünger et al., 1998) and REFMAC5 (Murshudov et al., 1997). The model was finally refined with CNS and contains 211 amino acids (residues 1063–1273, one mutation, L1063E, was observed), 182 water molecules, and 2 sulfate ions with a final R_m of 22.2% and R_free of 25.9%. Repeat 8 and repeat 9 contain three α helices and two loops, respectively; helix A (1065–1085), helix B (1093–1128), helix C (1134–1169), loop AB (1086–1092), and loop BC (1129–1133) in R8 and helix A’ (1170–1192), helix B’ (1200–1236), helix C’ (1242–1271), loop A’ B’ (1193–1199), and loop B’ C’ (1237–1241) in R9 (Figure 2A). The quality of the final model was assessed with the program PROCHECK (Laskowski et al., 1993). Crystallographic statistics are summarized in Table 1.

Hypothetical Models of β-Spectrin

Hypothetical models of β-spectrin as a chain of 16 complete repeats and a partial repeat were built by the repetitive superposition of two repeats. All structural superpositions were performed with the program LSQKAB (CCP4, 1994). In Figure 6A, R9 of HEβ89 was superposed on R8 of another HEβ89, after which the superposition was repeated in turn to make 17 tandem repeats. As a result, the model forms a semicircle as the twist angle is almost 0° in all repeats. In Figure 6B, R17 of CBα1617 was superposed on R16 of another CBα1617, after which the superposition was repeated in turn to make 17 tandem repeats. As a result, the model forms an extended supercoiled conformation as the twist angle is almost 30° in all repeats. In Figure 6C, the last repeat of eight tandem repeats based on the structures of CBα1617 was superposed on R8 of HEβ89, after which R9 of HEβ89 was superposed on the first repeat of nine tandem repeats based on the structures of CBα1617. As a result, the model contains a twist angle of almost 0° at the positions of repeats 8 and 9. In Figure 6D, the last repeat of four tandem repeats based on the structures of CBα1617 was superposed on R8 of HEβ89, and then R9 of HEβ89 was superposed on the fifth repeat of the model containing a twist angle of almost 0° at the positions of repeats 8 and 9 made in Figure 6C. As a result, the model contains two twist angles of almost 0° at the positions of repeats 4 and 5 and repeats 8 and 9. In all cases, helix C of the 17th repeat was removed from the final model. All figures were generated with the programs, Molscript (Kraulis, 1991), Bobscript (Esnouf, 1999), and Raster3D (Merritt and Murphy, 1994).

Acknowledgments

We gratefully acknowledge the contributions of J. Bosman and K.A. Swanson to the initial stages of this project. We thank B. Forget for the cDNA of human β-erythroid spectrin, G. Minasov and P. Focia for their help and advice, G. Minasov, I. Radakrishnan, and J. Widom for comments on the manuscript, and the Keck Biophysics Facility of Northwestern University for instrument support. Support from the R.H. Lurie Comprehensive Cancer Center of Northwestern University to the Structural Biology Facility is acknowledged. Research was supported by an NIH grant to R.I.M. and A.M. Portions of this work were performed at the DuPont-Northwestern-Dow Collaborative Access Team (DND-CAT) Synchrotron Research Center at the Advanced Photon Source (APS). DND-CAT is supported by DuPont, Dow, and the NSF. Use of the APS is supported by the Department of Energy (DOE).


Yuan, H.S., Wang, S.S., Yang, W.Z., Finkel, S.E., and Johnson, R.C. (1994). The structure of Fis mutant Pro61Ala illustrates that the kink within the long alpha-helix is not due to the presence of the proline residue. J. Biol. Chem. 269, 28947–28954.

Accession Numbers

The coordinates have been deposited in the Protein Data Bank (accession code 1S35).