
Amanda Nourse, Manuela Trabi, Norelle L. Daly, and David J. Craik‡
From the Institute for Molecular Bioscience, Australian Research Council Special Research Centre for Functional and Applied Genomics, University of Queensland, Brisbane, Qld 4072, Australia

The recently discovered cyclotides kalata B1 and kalata B2 are miniproteins containing a head-to-tail cyclized backbone and a cystine knot motif, in which disulfide bonds and the connecting backbone segments form a ring that is penetrated by the third disulfide bond. This arrangement renders the cyclotides extremely stable against thermal and enzymatic decay, making them a possible template onto which functionalities can be grafted. We have compared the hydrodynamic properties of two prototypic cyclotides, kalata B1 and kalata B2, using analytical ultracentrifugation techniques. Direct evidence for oligomerization of kalata B2 was shown by sedimentation velocity experiments in which a method for determining size distribution of polydisperse molecules in solution was employed. The shape of the oligomers appears to be spherical. Both sedimentation velocity and equilibrium experiments indicate that in phosphate buffer kalata B1 exists mainly as a monomer, even at millimolar concentrations. In contrast, at 1.8 mM, kalata B2 exists as an equilibrium mixture of monomer (30%), tetramer (42%), octamer (25%), and possibly a small proportion of higher oligomers. The results from the sedimentation equilibrium experiments show that this self-association is concentration dependent and reversible. We link our findings to the three-dimensional structures of both cyclotides, and propose two putative interaction interfaces on opposite sides of the kalata B2 molecule, one involving a hydrophobic interaction with the Phe residue, and the second involving a charge-charge interaction with the Asp residue. An understanding of the factors affecting solution aggregation is of vital importance for future pharmaceutical application of these molecules.

The cyclotides (1) are a recently discovered family of circular proteins that have been isolated from various Rubiaceae and Violaceae plants. They are one of several groups of naturally occurring circular proteins discovered in a wide range of micro- and higher organisms over recent years (2, 3). To date, around 50 different cyclotides are known, ranging in size from 28 to 37 amino acids. In addition to their head-to-tail cyclized peptide backbones, they all share a high sequence similarity as well as the striking feature of a cystine knot: two disulfide bonds and their connecting backbone segments form a ring that is penetrated by the third disulfide bond. This cyclic cystine knot (CCK) motif (4), i.e. the combination of circular peptide backbone and knotted disulfide arrangement, renders the cyclotides extremely stable, both against thermal degradation and enzymatic digestion. The stability the cyclotides may serve as useful protein templates for biopharmaceutical applications (12). The aim is to graft pharmacologically relevant peptide epitopes onto the stable CCK framework. Since knowledge of protein oligomerization and self-association behavior under various solution conditions will be vital in the evaluation of its suitability as a biopharmaceutical, we studied the sedimentation properties of two prototypic cyclotides, kalata B1 and kalata B2, with the purpose of shedding light on their properties in solution. These molecules were the first cyclotides to be discovered, are the most abundant peptides in the native plant O. affinis, and have dissimilar solution behaviors.

In this study we employed analytical ultracentrifugation techniques, where absorbance (for low concentrations) and interference optics (for high concentrations) were used to access a broad range of protein concentrations for the characterization of these cyclotides and their oligomers. Analytical ultracentrifugation provides important information about the solution behavior of biomolecules, including the overall size and shape of the molecules and their propensity to self-associate. We utilized sedimentation velocity experiments to resolve the nature of the aggregation behavior employing a newly applied method for determining the continuous size distribution of polydisperse macromolecules in solution (13, 14). We also characterized the hydrodynamic properties of these cyclotides, including the sedimentation coefficient (s), the translational diffusion coefficient...
Aggregation Behavior of Kalata B1 and Kalata B2

Fig. 1. Continuous size distribution analysis of sedimentation velocity data of kalata B1. Panel A, fringe displacements are plotted versus position from the center of rotation at time intervals of 2 min. A total of 180 scans were used in the analysis. The cell bottom value is between 7.13 and 7.15 cm (not stretched). For clarity every 4th scan is shown. The sedimentation velocity profiles were fitted to a continuous sedimentation distribution model according to Equation 1. Panels B and C are the bitmap and overlay representation of the residuals plotted as sedimentation distribution model according to Equation 1.

EXPERIMENTAL PROCEDURES

Sample Preparation—The limiting solubility of a freshly prepared solution of kalata B2 in buffer is ~6.5 mg/ml. This solution was stable over a 2-week period during which the analytical ultracentrifugation experiments were carried out at 20 °C. However, on storage at low temperature over time the solution becomes colloidal, with some precipitation occurring. By contrast, kalata B1 solutions appear to be less susceptible to precipitation. All sedimentation experiments were done with freshly prepared solutions.

Analytical Ultracentrifugation—Sedimentation equilibrium and velocity experiments were carried out using an Optima XL-A XL-I analytical ultracentrifuge (Beckman Coulter, Fullerton, CA), equipped with both absorbance and interference optical detection systems, using a Beckman An-60 Ti rotor with cells containing sapphire windows, and titanium double-sector centerpiece (pathlength 1.2 cm) (Nanalytics GmbH, Dallgow, Germany). The value of the radial distance of the bottom of the cell is between 7.13 and 7.15 cm (not stretched). Kalata B1 and kalata B2 were isolated from the aerial parts of O. affinis as described previously (1) and purified by repeated reverse phase HPLC. Prior to centrifugation, the peptides were either dissolved in or dialyzed into 100 mM sodium phosphate buffer, pH 7.4. The molar extinction coefficient at 280 nm (5875 M⁻¹ cm⁻¹), partial specific volumes (0.702 liters/g), and molecular masses for both kalata B1 and B2 (2,892 and 2,955 Da, respectively) were calculated based on its amino acid composition using the program Sednterp (16). In experiments using interference optics, 3.33 fringes were taken as equal to 1 mg/ml protein.

Sedimentation Velocity—For sedimentation velocity experiments, samples (340 μl, concentrations between 0.5 and 6.5 mg/ml) and reference solutions (350 μl) were loaded into cells. The rotor temperature was equilibrated at 20 °C in the vacuum chamber for 1–2 h prior to the start of the run. Experiments were conducted at 20 °C and rotor speeds of 60,000 rpm. Interference scans were collected at time intervals of 1 min. Data were analyzed with the software Sedfit Version 8.52, and errors reported are from replicate experiments.

Continuous Size Distribution Analysis—The continuous size distribution analysis of the sedimentation velocity data of the kalata proteins was performed with Sedfit (Version 8.52). In this analysis a differential sedimentation coefficient distribution \( c(s) \) that deconvolutes diffusion effects, based on the direct boundary modeling with distributions of Lamm equation solutions (13) is determined. A sedimentation coefficient distribution \( c(s) \) can be defined as in Equation 1,

\[
\alpha(x) = c(s) = sDs(x,Ds)dx + \epsilon
\]  

(1)

with the measured absorbance or interference profiles \( \alpha(x) \) denoting the observed sedimentation data, \( c(s) \) the concentration of species with sedimentation coefficients between \( s \) and \( s + ds \), \( \epsilon(x,Ds) \) the solution of the Lamm equation described above (17) and \( \epsilon \) the noise components. The Lamm equation was solved by finite element methods on a static or moving frame of reference (18, 19). For each species, the diffusion coefficient \( D(s) \) was estimated as a function of the sedimentation \( s \) based on the known partial specific volume and on an estimated anhydrous frictional ratio \( f/f_o \) (14). The \( c(s) \) distribution was converted into a molar mass distribution \( c(M) \) (13, 14).

All size distributions were solved on a radial grid of 1000 radius values between the meniscus and bottom, a confidence level of \( p = 0.68 \), and a resolution \( N \) of 300 sedimentation coefficients between 0.1 and 10 \( s \) (or molar masses between 500 and 100,000 Da).

Rapid Monomer-Tetramer-Octamer Self-association Model—The sedimentation velocity analysis was performed with Sedfit (Version 8.52) using a model for rapid reversible monomer-tetramer-octamer self-association by calculating finite element solutions of the Lamm equation according to methods described by Claverie (18), combined with a weight-average sedimentation coefficient and gradient average diffusion coefficients (20) and a two-step propagation scheme as described by Schuck (19). In modeling the experimental data the association constants and species sedimentation coefficients were treated as floating parameters.

Sedimentation Equilibrium—For the sedimentation equilibrium experiments the cells were filled with 130–150 μl of sample at loading concentrations between 1.0–6.5 mg/ml and reference (140–160 μl) solutions. Sedimentation equilibrium was attained at 24 h at a rotor temperature of 20 °C, and at rotor speeds of 40,000, 45,000, 50,000, and 60,000 rpm respectively. Interferometric patterns were recorded with only water in the cell at appropriate speeds and used for correcting for radial-dependent fluctuation in Rayleigh response across the cell (21).

Data analysis was performed by global analysis of 9 datasets obtained at different loading concentrations and rotor speeds using XL-A XL-I Data Analysis Software Version 4.0 (Beckman Instruments, Inc. Beckman Coulter, Fullerton, CA) for fitting the self-association models.
Aggregation Behavior of Kalata B1 and Kalata B2

Best-fit values and estimates of the continuous size distribution analysis of kalata B1 and kalata B2

<table>
<thead>
<tr>
<th>Sample</th>
<th>Concentration (mg/ml)</th>
<th>$\beta f_0^a$</th>
<th>$s_{20}^b$ (Svedberg)</th>
<th>$M^c$</th>
<th>$D_{20}^d$ $\times 10^{-6}$ cm²/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kalata B1</td>
<td>1.2</td>
<td>1.15</td>
<td>0.73</td>
<td>3146</td>
<td>1.94</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>1.15</td>
<td>0.72</td>
<td>3087</td>
<td>1.92</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>1.15</td>
<td>0.75</td>
<td>3252</td>
<td>1.92</td>
</tr>
<tr>
<td></td>
<td>6.5</td>
<td>1.15</td>
<td>0.75</td>
<td>3252</td>
<td>1.92</td>
</tr>
<tr>
<td>Kalata B2</td>
<td>0.5</td>
<td>1.15</td>
<td>0.70</td>
<td>2905</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>1.15</td>
<td>0.73</td>
<td>3149</td>
<td>1.94</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>1.15</td>
<td>0.75</td>
<td>3273</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>1.15</td>
<td>0.70</td>
<td>3054</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td>6.5</td>
<td>1.18</td>
<td>0.67</td>
<td>2857</td>
<td>1.97</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>1.15</td>
<td>0.72</td>
<td>3087</td>
<td>1.94</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>1.15</td>
<td>0.73</td>
<td>3149</td>
<td>1.94</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>1.15</td>
<td>0.75</td>
<td>3273</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>1.15</td>
<td>0.70</td>
<td>3054</td>
<td>1.91</td>
</tr>
</tbody>
</table>

*a* Best-fit weight-average anhydrous frictional ratio ($f_0$).

*b* Sedimentation coefficient taken from the ordinate maximum of each peak in the best-fit $c(s)$ distribution (Fig. 3).

*c* Molar mass values taken from the ordinate maximum of each peak in the best-fit $c(M)$ distribution (13, 14).

*d* Diffusion coefficient calculated using the best-fit values for sedimentation coefficient, frictional ratio and molar mass.

Estimates of the weight-average molar masses ($M_w$) were obtained from the least-squares fits that were based on the Boltzmann distributions of ideal species in the centrifugal field as shown in Equation 2,

$$S(r) = S(r_0)\exp[\sigma(r^2 - r_0^2)] - E$$

where $S(r)$ is the experimentally observed concentration signal in absorbance or fringes at radius $r$, $S(r_0)$ the concentration signal at the reference radius $r_0$, $\sigma = M(1 - y/F)^{1/2}$ (where $M$ is the weight-average molar mass of the monomer, $\sigma$ the partial specific volume of the solute, $p$ the solvent density, $y$ the rotor angular velocity, $T$ the temperature in Kelvin) and $E$ the baseline offset, and models found in the data analysis software involving multiple species (22–24).

RESULTS

In previous NMR spectroscopic studies we had noted that a well determined NMR structure of kalata B1 could be obtained in aqueous solution at millimolar concentrations, while kalata B2 required 20–30% of an organic co-solvent to achieve a similar quality structure. One possible explanation would be an increased propensity of kalata B2 to aggregate, but it was not possible to determine this from the NMR data. Analytical ultracentrifugation techniques were therefore employed to compare the hydrodynamic properties and resolve the oligomeric structures of kalata B1 and kalata B2. This was achieved by fitting the sedimentation velocity data to a continuous size distribution (sedimentation coefficient and molar mass), non-interacting discrete species and rapid monomer-tetramer-octamer self-association models. In addition, sedimentation equilibrium data over a wide range of solution conditions were fitted to non-interacting discrete species model, as well as a number of self-association models.

Sedimentation Velocity—The sedimentation velocity experiments for kalata B1 and kalata B2 were conducted at a concentration of 5.0 mg/ml. The data, obtained at time intervals of 1 min (Fig. 1), show an excellent fit to a continuous size distribution model (Equation 1). The best-fit sedimentation profiles shown in panel A are of very high quality (r.m.s.d. = 0.018, which represents a fit with a relative error of approx. 0.1%), as reflected in the small, randomly distributed residuals and virtually no systematic deviation visible in the residuals bitmap (Fig. 1, B and C). The resulting continuous size distribution profiles, $c(s)$ (sedimentation coefficient) (Fig. 1D), show the presence of a single monomeric species in solution with a sedimentation coefficient ($s_{20}$) of 0.75 S, a diffusion coefficient ($D_{20}$) of $1.92 \times 10^{-6}$ cm²/s, and a molar mass ($M$) of 3252 Da. The weight-average frictional ratio ($f_0$) was optimized by least-squares regression and converged to a best-fit value of...
1.15. Similar values were also obtained from experiments at other initial protein concentrations (1.25 and 2.5 mg/ml) as summarized in Table I. The presence of up to 90% monomer was evident, even at millimolar concentrations. The molar mass determined by this analysis is close to the known molecular mass of kalata B1 (2892 Da), based on its amino acid composition, and verified by mass spectrometry.

For comparison, the sedimentation profiles of kalata B2 at concentrations of 0.5, 2.0, 5.0, and 6.5 mg/ml were fitted to a continuous size distribution model. A single best-fit weight-average frictional ratio for all species at different concentrations was extracted from the data by virtue of the criterion of the quality of the fit. The results of the analysis are also summarized in Table I. At a loading protein concentration of 5 mg/ml, Fig. 2 reveals the presence of three major species, with sedimentation coefficients of ~0.70 (30% of the total), 1.91 (42%), and 2.60 S (25%). These $s_{20}$ values correspond to molar masses of 3,054, 13,810, and 22,108 Da. It thus appears that the solution is a mixture of monomers, with tetramers and octamers. The oligomers appear to be very stable, as confirmed by similar size distribution measurements made 2 weeks apart.

The mass of the 13,810 Da species (4.67 times the monomer molar mass) could be attributed to either a tetramer or a pentamer. Additional analysis of the sedimentation velocity
and equilibrium data suggests that this species is most likely a tetramer and that the 22,108 Da species is an octamer (see below).

Fig. 5 shows a comparison of the continuous size distribution profiles, c(s) in panels A–C of kalata B2 at initial protein concentrations of 0.5, 2.0, and 6.5 mg/ml. At low concentration kalata B2 (0.5 mg/ml, panel A) is present mainly as a single monomeric species in solution with an s_{20} of 0.70 S, a diffusion coefficient (D_{20}) of 2.0 \times 10^{-6} cm^2/s, and an apparent molar mass of 2,895 Da. The apparent molar mass value is consistent with that deduced from its amino acid composition (2,955 Da), and experimentally verified by mass spectrometry. When the concentration is increased to 2 mg/ml (panel B), a species of a molecular mass that corresponds to a tetramer (12,269 Da), corresponding to ~24% of the total, also appears in solution. At 6.5 mg/ml (Fig. 3, panel C) a third aggregate, in addition to the monomer (~19%) and the putative tetramer (~47%), corresponding to an approximate molar mass of an octamer of 23,891 Da (27%) becomes apparent.

To determine whether a rapid reversible equilibrium exists between the kalata B2 species, the sedimentation velocity data at an initial protein concentration of 5.0 mg/ml were fitted to a monomer-tetramer-octamer rapid self-association model (Sedfit Version 8.5) as shown in Fig. 4. For this fit, the respective values obtained for s_{20} from fits to the continuous size distribution model (Table I), and association constants from fits of the sedimentation equilibrium data to a monomer-tetramer-octamer model (see below), were used as starting values. For the monomer, tetramer, and octamer, respectively, the calculated best-fit sedimentation profiles converge to s_{20} values of 0.7, 1.6, and 2.6 S, with association constant values of K_{1,4} = 1.4 \times 10^{10} M^{-1} (0.008 fringe units) and K_{1,8} = 1.0 \times 10^{14} M^{-7} (1.0 \times 10^{-8} fringe units). The r.m.s.d. of the global fit is 0.04. Panels B and C are the bitmap and overlay representation of the residuals plotted as a function of radial position.

The oligomeric states of kalata B2, the Stokes radius and dimensions of equivalent hydrodynamic ellipsoids (oblate and prolate) of each oligomer were estimated from the experimental sedimentation coefficients (Table I) and a theoretical estimate for hydration of 0.30 g/g as predicted from the amino acid composition (16, 25).

The oligomeric states of kalata B2 at initial protein concentrations of 0.5, 2.0, and 6.5 mg/ml were fitted to a monomer-tetramer-octamer self-association model (solid lines), which converged to s_{20} values of 0.70, 1.6, and 2.6 S for the monomer, tetramer, and octamer, respectively. The association constants were determined as K_{1,4} = 1.4 \times 10^{10} M^{-1} (0.008 fringe units) and K_{1,8} = 1.0 \times 10^{14} M^{-7} (1.0 \times 10^{-8} fringe units). The r.m.s.d. of the global fit is 0.04. Panels B and C are the bitmap and overlay representation of the residuals plotted as a function of radial position.

**Table II**

<table>
<thead>
<tr>
<th>Oligomer</th>
<th>M</th>
<th>s_{20}</th>
<th>f_{D0}</th>
<th>Stokes radius^a</th>
<th>Oblate ellipsoid^b</th>
<th>Prolate ellipsoid^c</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Da</td>
<td>S</td>
<td>nm</td>
<td></td>
<td>nm \times nm</td>
<td>nm \times nm</td>
</tr>
<tr>
<td>Monomer</td>
<td>2,955</td>
<td>0.70</td>
<td>1.14</td>
<td>1.06</td>
<td>1.40</td>
<td>2.49</td>
</tr>
<tr>
<td>Tetramer</td>
<td>11,820</td>
<td>1.65</td>
<td>1.21</td>
<td>1.80</td>
<td>2.51</td>
<td>5.55</td>
</tr>
<tr>
<td>Octamer</td>
<td>23,891</td>
<td>2.60</td>
<td>1.22</td>
<td>2.28</td>
<td>2.63</td>
<td>8.22</td>
</tr>
</tbody>
</table>

^a Calculated from the theoretical molar mass and sedimentation coefficient.

^b Axial ratios for oblate or prolate ellipsoid calculated by the v method (16) using the program SEDNTERP and employing theoretical molar masses with their respective sedimentation coefficients (Table I) and a theoretical estimate for hydration of 0.30 g/g as predicted from the amino acid composition (16, 25).

^c Overall dimensions, calculated with the assumption of a hydration of 0.3 g/g.
monomeric species, acquired at a loading concentrations of 1.25, 2.5, and 5.0 mg/ml and rotorspeeds of 40,000, 50,000, and 60,000 rpm. The data fit well to a single monomeric species (Equation 1) with an r.m.s.d. of 0.029. Panel B, residuals for the nonlinear least squares best-fits at all concentrations and all rotor speeds are plotted versus the distance from the axis of rotation. Panel C, fringe displacement at equilibrium is plotted versus the distance from the axis of rotation for kalata B1. Samples at initial concentrations of 1.0 (circle), 3.0 (square), and 6.5 mg/ml (triangle) in 100 mM sodium phosphate, pH 7.4, were centrifuged at 20 °C for at least 24 h at 40,000, 45,000, and 60,000 rpm, respectively. For clarity only data at a rotor speed of 40,000 (all offset by 2 fringes and -0.1 cm) and 50,000 rpm are shown. The solid lines represent the global nonlinear least squares best-fits to the data at all concentrations and all rotor speeds according to a monomer-tetramer-octamer model. Panel D, residuals for the nonlinear least squares best-fits at all concentrations and all rotor speeds are plotted versus the distance from the axis of rotation. The r.m.s.d. was 0.033.

Fig. 5. Sedimentation equilibrium of kalata B1 and kalata B2. Panel A, fringe displacements at equilibrium are plotted versus the distance from the axis of rotation for kalata B1. Samples at initial concentrations of 1.25 (circle), 2.5 (square), and 5 mg/ml (triangle) in 100 mM sodium phosphate, pH 7.4, were centrifuged at 20 °C for at least 24 h at 40,000, 50,000, and 60,000 rpm, respectively. For clarity only data at rotor speeds of 40,000 (all offset by 2 fringes and -0.1 cm) and 60,000 rpm are shown. The solid line represents the global nonlinear least squares best-fit to the data to mainly a single monomeric species (Equation 1) with an r.m.s.d of 0.029. Panel B, residuals for the nonlinear least squares best-fits at all concentrations and all rotor speeds are plotted versus the distance from the axis of rotation. Panel C, fringe displacement at equilibrium is plotted versus the distance from the axis of rotation for kalata B2. Samples at initial concentrations of 1.0 (circle), 3.0 (square), and 6.5 mg/ml (triangle) in 100 mM sodium phosphate, pH 7.4, were centrifuged at 20 °C for at least 24 h at 40,000, 45,000, and 60,000 rpm, respectively. For clarity only data at a rotor speed of 40,000 (all offset by 2 fringes and -0.1 cm) and 50,000 rpm are shown. The solid lines represent the global nonlinear least squares best-fits to the data at all concentrations and all rotor speeds according to a monomer-tetramer-octamer model. Panel D, residuals for the nonlinear least squares best-fits at all concentrations and all rotor speeds are plotted versus the distance from the axis of rotation. The r.m.s.d. was 0.033.

Surface-exposed Residues—In order to explain the difference in the solution behavior of kalata B1 and kalata B2, their three-dimensional structures, obtained from solution NMR experiments, were compared. The amino acid sequences of kalata B1 (26) and kalata B2 (1) differ in only five residues, so particular attention was given to the properties and the extent of surface exposure of those residues. The locations of the five amino acid “substitutions” (V6F, T16S, S18T, V21I, and N25D)
of kalata B2 compared with kalata B1, along with the overall amino acid sequence, are displayed in Fig. 6. The overall three-dimensional structures of the two peptides are almost identical, as might be expected, because both contain a cystine knot motif (27) that strongly influences the core structure. We therefore turned to an analysis of the surface properties of the molecules.

Fig. 7A shows the surfaces of kalata B1 and B2 with residues color-coded into categories of positive, negative, polar or hydrophobic. The cysteine residues are colored separately from the other polar residues for convenience. This makes it clear that most of the cystine residues are buried in the core of the molecule, with only Cys210 and Cys17 surface-exposed. This exposure occurs only on one face of the molecule and forms part of a polar or hydrophilic region. Surface-exposed hydrophobic residues dominate the opposite face of the molecule, presumably forced to the surface because there is no room for them in the protein core because of the cystine knot. Differences in the extent of this hydrophobic patch between the two molecules are only revealed by a comparison of the average solvent accessible surface areas of kalata B1 and kalata B2. This comparison, given in Å² and a percentage of the total surface area, and sorted by amino acid properties is summarized in Table III.

There is an almost 115 Å² increase in the hydrophobic solvent accessible surface area of kalata B2 compared with kalata B1, which accounts for most of the increase in the overall solvent exposed area of 125 Å², the other being an increase of 10 Å² in the hydrophilic area. The increase in hydrophobic area is largely due to the substitution of V6 in kalata B1 for a Phe in kalata B2, which leads to an increase in the solvent exposed area by an average 62 Å². This is one of two hydrophobic-hydrophobic substitutions, the other being the Val24 for Ile on the same face of the protein. Phe6 and Ile21 form part of an extended hydrophobic patch, which also includes Gly7, Gly8, Trp16, and Pro20, covering 28% of the overall surface of kalata B2.

The other noticeable difference between the surfaces of the two peptides is in the charge distribution, as illustrated in Fig. 7B, which shows that the oppositely charged residues Arg24 and Asp25 are in close proximity in kalata B2, creating an exposed bipolar patch on one end of the molecule that does not exist in kalata B1. This bipolar region is also almost completely solvent-exposed. We therefore suggest that this solvent-exposed bipolar patch, involving Asp25, in combination with the hydrophobic patch, involving Phe6, might be involved in the oligomerization of kalata B2 in solution.

Molecular modeling studies were carried out to assist in predicting the way in which the bipolar region and hydrophobic patch might be involved in self-association. Fig. 8 illustrates the potential self-interaction interfaces that might lead to tetramer formation. Fig. 8A focuses on the potential charge-charge interactions between the bipolar patches of different monomers. Fig. 8B shows a possible relative orientation of the two putative self-interaction surfaces of each monomer in the tetramer. The key interacting regions that distinguish kalata

---

**Fig. 6. Ribbon representation of the overall fold of kalata B1 and B2.** The Cα atoms of amino acids common to both peptides are shown as balls labeled with one-letter code (residues other than cysteine) or numbered according to their position in the peptide sequence (cysteine residues) and colored according to the properties of the amino acids (hydrophobic, gray; hydrophilic, light green; acidic, red; basic, blue; cysteine, yellow). For residues that differ in kalata B1 and kalata B2, the amino acid side chains are shown as ball and stick (in blue for kalata B1, in green for kalata B2) and labeled with their one-letter codes and residue numbers in the corresponding color.

**Fig. 7. Surface representations of kalata B1 (right) and kalata B2 (left).** A, solvent accessible surface colored according to the properties of the residues (hydrophobic, gray; hydrophilic, green; acidic, red; basic, blue; cysteine, yellow). For orientation, the position of various residues is indicated by the one-letter code and residue number. The models in the second row are rotated by 180° around the x-axis. B, electrostatic surface potentials with negatively charged areas given in red and positively charged areas in blue. The illustration in the lower left corner shows the bipolar patch in kalata B2 created by residues Arg24 and Asp25.

**Table III**

<table>
<thead>
<tr>
<th>Kalata B1</th>
<th>Kalata B2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute</td>
<td>Percent</td>
</tr>
<tr>
<td>Hydrophobic</td>
<td>1131.8</td>
</tr>
<tr>
<td>Acidic</td>
<td>34.7</td>
</tr>
<tr>
<td>Basic</td>
<td>86.6</td>
</tr>
<tr>
<td>Cysteine</td>
<td>132.7</td>
</tr>
<tr>
<td>Total</td>
<td>2025.6</td>
</tr>
</tbody>
</table>
We found that the self-assembly of kalata B2 into multimers takes place in a reversible concentration-dependent manner, i.e. at very low concentrations only the monomeric species is present, at intermediate concentrations most likely tetramers appear, and at high concentrations octamers and possibly higher order oligomers emerge. The analysis consistently reveals the presence of three major species at high concentration with sedimentation coefficients of −0.7, 1.91, and 2.6 S (Table I). Furthermore, the peak of the monomer does not shift significantly (Fig. 3), which shows that there is no significant concentration dependence of the sedimentation coefficient, an indication of ideal sedimentation.

Although the number and relative concentration of individual kalata B2 oligomers can be ascertained by using the continuous size distribution method, it was difficult to be certain about the oligomeric identity of the 1.91 S peak, i.e. either a tetramer or a pentamer. Possible errors can be introduced when transforming the c(s) distribution into molar mass distribution c(M) (14). In addition, the minor systematic error in the fit indicates that either the oligomers are in reversible equilibrium with each other, possibly on the time scale of the sedimentation process, or that the oligomers differ slightly in their shape, in which case the c(s) assumption of a single weight-average frictional ratio is not strictly fulfilled. There is some indication that this 1.91 S peak represents a tetramer if the analysis of additional sedimentation velocity and equilibrium data is taken into account. For instance the velocity data fits well to a rapid monomer-tetramer-octamer model of self-association (Fig. 4). This model of self-association is also supported by the analysis of equilibrium data (Fig. 5, B and C) that fit well to a monomer in equilibrium with tetramer and octamer, but not with a pentamer. This strongly suggests that a monomer-tetramer-octamer self-association pattern is mainly the manner in which kalata B2 self-associates.

The sedimentation data enabled an estimation of the shape and asymmetry of each oligomer. The overall dimensions of the monomer were calculated as 2.40 × 1.70 nm (equivalent an oblate ellipsoid) or 2.63 × 1.89 nm (equivalent to a prolate ellipsoid), with both corresponding very well to the dimensions measured from its NMR structure (2.5 × 1.9 × 1.5 nm). The Stokes radius of 1.06 nm, with the axial ratio values of the monomer being <2, suggests that kalata B2 is spatially compact and has a roughly spherical shape. Thus, the structural information on kalata B2 supports our hydrodynamic findings. The finding that an increase in oligomeric size causes only a slight increase in the axial ratios (Table II), would suggest that the monomers are arranged mainly in a symmetrical fashion that is compatible with a spherical kalata B2 oligomer. The data rule out a linear oligomerization pattern that would result in asymmetric multimeric structures in solution.

To try to explain the difference in solution behavior, the properties of the surface exposed residues of the two cyclotides were examined, with particular attention given to differences in the amino acid sequence. There are two surface areas on these proteins with distinguishing properties and features that might explain the difference in their solution behavior. One is tides has not previously been undertaken. Since these proteins possess a CCK scaffold that has been proposed as a potential template onto which desirable properties can be grafted, our results show that in the design of such modified templates, changes in aggregation behavior should also be taken into account. Furthermore, aggregation in vivo is likely to be enhanced, since these peptides are intended to perform their function in biological environments that are generally characterized by high total concentrations of macromolecules, referred to as “crowded” (29, 30).

We resolved the oligomeric forms of these cyclotides in aqueous buffer by the application of sedimentation equilibrium and sedimentation velocity experiments, employing a new method for determining the continuous size distribution of polydisperse macromolecules (14); that has been shown to be the best means to detect and quantitatively characterize self-association of macromolecules in solution (14, 28). A detailed comparison of the hydrodynamic properties and solution behavior of cyclo-
the bipolar area that includes the N25D substitution, and the other, on the opposite side of the molecule, is the hydrophobic surface patch that includes the V6F substitution (Fig. 7). The bipolar patch created by the N25D substitution, Arg24-Asp25, might theoretically facilitate intermolecular ionic self-interaction of kalata B2 (as depicted in Fig. 8A) but not kalata B1, since the latter lacks this second negatively charged amino acid. To support the notion of charge-charge interaction between different kalata B2 molecules, there are many examples where salt bridges exist between two associating molecules in buffer solution. One such instance is the affinity of an antibody for lysozyme via a Guu-Arg salt bridge, confirmed in analytical ultracentrifugation studies (31).

With respect to the V6F substitution, a recent analysis has shown that Phe has a high probability of being found in many types of protein-protein interfaces, whereas Val does not (15). It is thus possible to hypothesize that the V6F substitution in kalata B2 facilitates self-association. The two surface regions discussed are substantially solvent-exposed, which makes them good candidates for interaction areas. Moreover, a model of the tetramer (Fig. SB), showing a possible relative orientation of the two putative self-interaction surfaces, illustrates that the Phe6-Phe6 interaction in the hydrophobic patch interface is possible while maintaining the bipolar patch interactions. It is noted that in this model the other hydrophobic substitution, V21I, is also present in this hydrophobic patch interface and that the Ile lines up favorably for an Ile21-Ile21 interaction. Furthermore, Ile is preferred (only slightly), as shown in the V6F model, over Val in self-interaction complexes (15). It is thus possible that, in addition to Phe6, Ile21 may also contribute to the self-aggregation of kalata B2. As illustrated in Fig. 8B, several other hydrophobic residues that are conserved in both kalata B1 and B2 line the putative interaction surface. We therefore propose a hypothetical model of self-association that involves a Phe6-Phe6, and possibly an Ile21-Ile21, as well as both kalata B1 and B2 line the putative interaction surface.

In conclusion, we have shown that kalata B2 assembles into a monomer-tetramer-octamer self-association pattern, while kalata B1 exists as a monomer in solution, even at relatively high concentration. Furthermore, we have determined that the kalata B2 oligomers are most likely globular rather than elongated. Based on the properties of the surface exposed residues, and residues highly likely to be found in self-interaction interfaces (15), we have proposed two putative self-association surface areas on opposite sides of the circular protein kalata B2. These two substantially solvent exposed areas, where the V6F and the N25D changes are found, represent the most distinguishing features of kalata B2 relative to kalata B1. Being on opposite sides of the protein, Phe6-Phe6 and Arg24-Asp25 interactions between different kalata B2 molecules to form a globular tetramer are sterically feasible.

Acknowledgments—We thank P. Schuck for valuable comments on the analysis of the analytical ultracentrifugation data, K. Rosengren for critical reading of the manuscript, and Rekha Bharathi for purification of kalata B2. The Institute for Molecular Bioscience is a Special Research Centre for the ARC.

REFERENCES
5. Gran, L. (1973) Lloydia 36, 174–178
Amanda Nourse, Manuela Trabi, Norelle L. Daly and David J. Craik

doi: 10.1074/jbc.M306826200 originally published online October 15, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M306826200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 29 references, 6 of which can be accessed free at http://www.jbc.org/content/279/1/562.full.html#ref-list-1