Coarse-grained, foldable, physical model of the polypeptide chain

Promita Chakraborty and Ronald N. Zuckermann[1]

The Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

Edited* by Ken A. Dill, Stony Brook University, Stony Brook, NY, and approved June 25, 2013 (received for review March 29, 2013)

Although nonflexible, scaled molecular models like Pauling–Corey's and its descendants have made significant contributions in structural biology research and pedagogy, recent technical advances in 3D printing and electronics make it possible to go one step further in designing physical models of biomacromolecules: to make them conformationally dynamic. We report here the design, construction, and validation of a flexible, scaled, physical model of the polypeptide chain, which accurately reproduces the bond rotational degrees of freedom in the peptide backbone. The coarse-grained backbone model consists of repeating amide and α-carbon units, connected by mechanical bonds (corresponding to φ and ψ) that include realistic barriers to rotation that closely approximate those found at the molecular scale. Longer-range hydrogen-bonding interactions are also incorporated, allowing the chain to readily fold into stable secondary structures. The model is easily constructed with readily obtainable parts and promises to be a tremendous educational aid to the intuitive understanding of chain folding as the basis for macromolecular structure. Furthermore, this physical model can serve as the basis for linking tangible biomacromolecular models directly to the vast array of existing computational tools to provide an enhanced and interactive human–computer interface.

protein folding | self-assembly | biomimetic modular robotics | rotational energy barrier | conformational isomerism

Understanding protein folding pathways, predicting protein structure and the de novo designing of functional proteins have been a long-standing grand challenge in computational and structural biology. Although tremendous advances are being made (1), folded protein structures are very difficult to visualize in our mind due to their complexity and shear size. State-of-the-art computer visualization techniques are well developed and provide an array of powerful interactive tools for exploring the 3D structures of biomacromolecules (2–4). While increasingly complex molecules can be visualized with computers, the mode of user interaction has been mostly limited to the mouse and keyboard. Although the use of haptic devices is on the rise, there are only a few low-cost, specialized input devices particularly designed for interaction with biomacromolecules. Augmented-reality (AR) and immersive environments enhance user interaction experiences during the handling of existing visualization tools (2, 5, 6), but the physical models used in these environments are not flexible or precise enough to represent the conformational dynamism of polypeptides by themselves. There is a strong need for scaled, realistically foldable, but inexpensive, physical models to go hand-in-hand with the AR and other computer interfaces, while concomitantly taking better advantage of current computational capabilities.

Physical molecular models of organic small molecules and biomacromolecules with atomic representations have been around for many decades (7–17). Although early pioneers like Pauling and Corey's scaled physical model of the polypeptide chain (11) helped to elucidate the molecular packing details of protein secondary structures (15, 17), these were nonflexible models and did not capture the inherent dynamism now known to exist in protein structures. Although there are a variety of physical molecular models commercially available (9, 10, 13, 14, 16), none have captured the true conformational degrees of freedom of the polymer chain, as the macromolecules have too many atoms, complex short-range and long-range conformational constraints, and specific folding behaviors. There are construction kits for α-helices, β-sheets, and nucleic acids (14) that capture the scale and complexity of these molecules, but they are not flexible. Some models focus on chemical structures with multiple types of bonding (9) but are not made to scale. Other models spontaneously self-assemble into 3D molecules aided by internal magnetic fasteners (16) but are too simplistic to represent the folding of the polypeptide chain. Entire molecules of folded protein structures have also been generated with 3D printing, but these models do not explicitly represent the backbone and thus cannot be folded or unfolded.

Although many of the existing models are accurate with respect to physical dimensions (atomic radii, bond lengths, and bond angles), none is able to freely sample the bond rotational degrees of freedom that are needed to represent the motion of the polypeptide chain. The most dynamic representation reported is Olson's articulated model that has been used to make flexible polypeptides by chaining the constituents through an elastic string, with the elasticity representing the pull between atoms (18). These models are flexible and can be folded into protein secondary and tertiary structures, but do not have the space-filling impact or a realistic representation of the dihedral angle rotational barrier that are so central to protein backbone behavior.

Despite these limitations, physical models have been rising in popularity (7, 8), as they play a critical role, both as educational tools and as aids to chemistry and biochemistry researchers to gain insight into protein-folding mechanisms. Physical models engage visuospatial thinking of biomolecules much more effectively than textbook images and computer screens can, via a process termed "tactile visualization" (19). Moreover, experiments with gaming interfaces like FoldIt have demonstrated that humans have superior 3D pattern matching skills than any existing software for solving challenging scientific problems (20, 21). We believe that this intuition and skill can be even more channelized while playing with physical, foldable models and may result in unexpected and surprising discoveries, as in the case of FoldIt.

With this vision in mind, we report here the design and fabrication of a tangible, coarse-grained, dimensionally accurate, physical molecular model of the polypeptide chain, which has the necessary degrees of freedom and bond rotational barriers to accurately emulate the backbone folding dynamics of the polypeptide chain. Our approach was to break down the component amino acids into constituent coarse-grained components linked

Author contributions: P.C. and R.N.Z. designed research; P.C. and R.N.Z. performed research; R.N.Z. contributed new reagents/analytic tools; P.C. and R.N.Z. analyzed data; and P.C. and R.N.Z. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

[1]To whom correspondence should be addressed. E-mail: rnzuckermann@lbl.gov.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1305741110/-/DCSupplemental.

by rotatable bonds. The flexibility of the backbone chain in our model has made it possible to readily build all of the common protein secondary structure elements. This model is a necessary first step toward a sophisticated computer input device that can manipulate and intuitively interact with computer visualization tools. It should ultimately be possible for these models to provide interactive dihedral angle information to computers while transitioning between various conformations. This would enable real-time feedback about the chain's conformational energy and direct comparison with known protein structures (structural homology searching). The model will serve as a first step in implementing an intuitive, computationally augmented physical model that can help people instinctively understand and hypothesize new details about the science of protein-folding pathways.

Methods: Design and Testing of the Model

In real polypeptide chains, the bond rotations along the backbone are restricted, such that only certain bonds can rotate while others remain relatively rigid. Each amino acid monomer contains two backbone rotational degrees of freedom, the φ and ψ dihedral angles (Fig. 1). When we consider the rigid and flexible backbone elements, the chain can be dissected into two repeating units: a set of four atoms confined to a rigid plane (forming the amide), and the α-carbon atom, where two amides are connected via φ and ψ bonds. Thus, the polypeptide backbone can be represented as an alternating copolymer of the amide unit and the α-carbon unit. In our model, referred to as a Peppytide, we emulate this basic structure of the polypeptide chain backbone by linking two types of units together: the amide units and the α-carbon units (C_α), connected alternately at φ and ψ bonds (Fig. 2 *C* and *D*). Thus, the simplest assembly that contains both φ and ψ bonds is an α-carbon linked to two amides. This forms an amide-C_α-amide arrangement in the model that we refer to as an amino acid diamide (Fig. 2*A*). We have also developed a third unit, a methyl-group unit (Fig. 2*E*) representing alanine, as a generic side-chain residue. Alanine is the smallest amino acid side chain, where the methyl group can approximate the impact of side-chain substitution and chirality on the general dynamics of a small peptide chain. Polyalanine has also been known to form α-helices and β-sheets (22, 23).

Polypeptide backbone conformations are dominated by both short-range interactions about φ and ψ, and longer-range intrachain hydrogen bonding interactions, as well as the interactions of the side chains. The Peppytide model embodies both the short-range and long-range interactions of the backbone in addition to the steric hindrances of atoms that are within spatial proximity.

The factors most important for steric hindrances are the shapes and sizes of the constituent parts. By close analysis of protein crystal structures, the shapes of amide units (*trans*) and α-carbon units (corresponding to L-amino acids) were designed (Fig. 2*A*). The most widely accepted values of interatomic distances have been used for the atomic-scale dimensions of the units (Fig. 2*B*) (24). All parts were drawn to scale with a scale factor of 1 Å = 0.368″ in a computer-aided design (CAD) software (Fig. 2 *C*–*E*). The φ and ψ bonds, which are the linkages between the amide and α-carbon units, were implemented with freely (360°) rotating nut-and-screw arrangements. Rotational barriers were also included to reproduce the dihedral angle preferences observed in protein structures (see below). As the constituent atoms of each of the units need to be within their covalent bonding distances, the bonding atoms were cut along specific planes, as had been previously done with the Corey-Pauling-Koltun (CPK) and other models (11).

Theoretically, the atom size for the elements in the backbone chain, namely the model radii (R_M), in any model should be equal to their Van der Waals radius (R_{VDW}). However, in a dynamic physical model, the R_M needs to be a fraction of R_{VDW} for the chain to move freely and avoid getting interlocked with itself. This was examined by checking for steric clashes in a CAD software using a 3D drawing of the alanine diamide molecule assembled using Peppytide model units. R_M was varied from 0.6 R_{VDW} to 0.8 R_{VDW} (Fig. 3*C* and *SI Appendix*, Fig. S1), and it was found that $R_M = 0.7\ R_{VDW}$ is the largest size possible for representing hard spheres while maintaining access to the entire conformational landscape accessible by polypeptides.

The backbone dihedral angles in polypeptide chains do not rotate freely. There are barriers to rotation about both φ and ψ that limit the conformational flexibility of the chain, which is a result of the local bonding geometry, steric and electronic effects (25). We therefore introduced a conformational bias into each rotatable bond in the backbone of the model. The favored dihedral angles (φ, ψ) in polypeptides are well known from experimental data and are typically illustrated in a Ramachandran plot (25). The densest regions of the plot, that is, the most favorable regions, are low-energy positions of dihedrals of the polypeptide backbone (Fig. 3*D*). These preferred regions mostly correspond to the α-helix (left- and right-handed) and the β-sheet conformations—the secondary structures universally found in proteins.

To represent these barriers within the physical model, it was necessary to decouple φ and ψ from each other and to study their behavior separately. The information from the Ramachandran plot was decoupled to get independent values of φ and ψ over the full range of rotation (−180°, 180°) (Fig. 3 *A* and *B*). We used data from ∼78,000 known protein structures in the Protein Data Bank (PDB) because these data are a direct manifestation of the favored angles adopted in proteins. For comparison, we also generated a Ramachandran energy map with OPLS 95 (optimized potentials for liquid simulations) force fields in the Maestro framework for alanine dipeptide, where the energy minima mostly clusters around φ = −70° and −140° (Fig. 3*E*). However, these calculated energies are an indirect measurement of the same effect as they reflect only very local interactions. As we wanted to incorporate the effects of both short-range and long-range interactions in the physical model, we used the data from the PDB histogram instead of the minima in the energy maps to design the rotational barriers.

The decoupled φ and ψ distributions giving preferred angles for φ and ψ (Fig. 3 *A* and *B*) were calculated from ∼59 million (φ, ψ) values obtained from 77,873 protein structure files from the PDB (crystallography and NMR structures only). The four peaks obtained (two for φ; two for ψ) correspond to the darkest regions of the Ramachandran plot for α-helix and β-sheet conformations, and denote the corresponding minimum energy configurations. This analysis shows that the φ-peaks are 56° apart (at −62° and −118°) and ψ-peaks are 180° apart (at −42°, 138°). There is a third φ-peak at ∼61° that corresponds to left-handed helices, which is not represented in this version of the model. The peaks in the φ and ψ distributions were each fit to Gaussian distribution (Fig. 3*A*) to facilitate their approximation in the physical model.

To introduce these dihedral angle preferences, or rotational barriers into the physical model, we used a customized circular magnet array for each φ and ψ bond. Magnet arrays can be quite intricate and can produce a wide variety of mechanical interactions (26, 27). Magnets are attractive choice in this application because they are noncontact, frictionless, cheap, passive (need no power to operate), exhibit strong coupling behavior, and can generate Gaussian barriers.

In Peppytides, we reproduced the conformational biasing due to rotational barriers by two separate arrangements of magnets for φ and ψ, respectively, that work in unison to form a physical rotation barrier (Fig. 4). By arranging small, powerful neodymium magnets across the rotational interfaces (Fig. 4, *Left*), certain bond rotation angles (or angle ranges) are preferred by the model during the 360° rotation of the φ or ψ bonds. Thus, we are able to embody, with reasonable precision, the natural torsional angle biases for the entire landscape of the Ramachandran plot in the model. Based on the distribution functions of φ and ψ, the magnets are positioned 56° and 180° apart, respectively (Fig. 4*A* for φ coupled faces, and Fig. 4*B* for ψ coupled faces). In our design, the most stable conformations of both φ and ψ

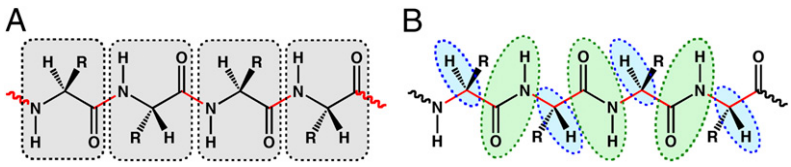

Fig. 1. Representations of the repeating units of the polypeptide chain. (*A*) A repeating homopolymer of amino acid units connected by amide bonds (red), (*B*) an alternating copolymer of rigid amide (green) and α-carbon (blue) units connected by the rotatable bonds φ and ψ (red).

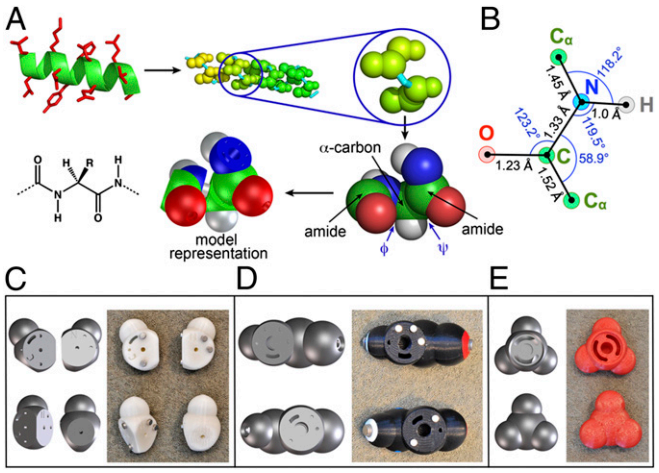

Fig. 2. Shapes and dimensions of the repeating units in the polypeptide chain model. (*A*) Equivalence of the shape of amides and α-carbons in a leucine zipper motif (from pid:2ZTA), and that of the Peppytide model; C, green; H, gray; N, blue; O, red. (*B*) Bond dimensions and angles of atoms drawn to scale, comprising amides and α-carbons used for the model (24). (*C*–*E*) CAD drawings and finished parts: (*C*) α-carbon unit corresponding to L-amino acids: ψ-face (*Upper Left*); φ-face (*Upper Right*); side-chain face (*Lower Right*), (*D*) amide unit (trans configuration): φ-face (*Upper*); ψ-face (*Lower*). (*E*) Methyl-group unit (for alanine side chain).

are stabilized by two pairs of magnets in each face providing added strength. For ψ, the magnets on each of the coupled faces were positioned at −42° and 138° (180° apart). For φ, the arrangements were slightly more complicated. To stabilize the magnets, we needed to place three magnets 56° apart on amide (N) face, and two magnets 56° apart on the α-carbon face in precise locations. This coupled arrangement of φ resulted in two primary energy minima (peaks 2 and 3), and two weaker satellite minima (peaks 1 and 4) (Fig. 4, *Upper Right*). The actual macroscopic barriers to rotation in the physical model due to the magnetic arrays were experimentally determined (for methods, see *SI Appendix*, section S5). To quantify the energy spent, the coupled faces of φ (and separately ψ) were slowly rotated over multiple cycles with a DC motor and the current drawn during rotation was measured as a function of rotation angle. Under the conditions used, the current drawn by DC motor is directly proportional to the shaft torque, which is proportional to the output energy (28). The current data were processed to extract the corresponding energy barrier (*SI Appendix*, Fig. S13). We found that the two primary energy minima for φ align well with the φ-distribution peaks at −62° and −118° (Fig. 4, *Upper Right*, red curve). However, the φ energy curve also has two additional weaker satellite minima at around −7° and −173° that broaden the energy well as compared to the natural system (blue curve). However, the five-magnet design allows for two sets of magnets to overlap, providing an energy barrier on the same order as the ψ bond. For ψ, the two energy minima values match well with that of the PDB distribution at −42° and 138°. Importantly, with these measured energy profiles, we can generate an equivalent Ramachandran plot that overlaps remarkably well with the natural system (comparing Fig. 3 *D* and *F*).

The representation of hydrogen bonding is another important feature reproduced in the Peppytide model. The long-range hydrogen bond interactions of the polypeptide chain are key to formation of the secondary and tertiary structure. The model reproduces the hydrogen bond donor (NH group) and acceptor (C=O group) behavior of distal amides by using a pair of rod magnets (Fig. 5*B*). Magnets are a reasonable approximation of the hydrogen bond interaction because in reality the NH and C=O groups attract each other but not themselves, similar to the north and south poles of a magnet. Importantly, this feature allows the model to reproduce long-range interactions between monomers that are separated in sequence space, yet are in close contact in 3D space, thus enabling and stabilizing secondary structure. Previously, Olson's articulated polypeptide chain model has demonstrated the use of magnets as hydrogen bonds (18). Magnets are only an approximate means to represent the H-bond interactions, as the force–distance relationship in magnets is different from that of H-bond strengths. Although both have exponential decay rates, the H-bond strength (29) exhibits a different decay rate than magnet-array fields (30). However, their decay curves follow roughly similar trends, making the magnets a practical choice for representing H-bonds approximately. Another advantage is that they are passive components requiring no power to operate.

All of the three units of the model were designed to be hollow to make them as light as possible to minimize the impact of gravity. The parts were created using a 3D printer using acrylonitrile butadiene styrene (ABS) plastic. The 3D-printable stereo lithography (STL) files for these parts are provided to enable anyone to readily produce the Peppytide model themselves (*SI Appendix*, section S3; Datasets S1–S3). Undersized pilot holes are designed in each part to guide drilling precision bores for the bond pieces and magnet installation. The parts were subsequently assembled in a chain using cheap screws and spacers (for details, see *SI Appendix*, section S4). The Peppytide model is ∼93,000,000 times magnified (Fig. 5*C*), 1.3×10^{23} times heavier than its biological counterpart, and can be manually folded into secondary and tertiary structures.

Results and Discussion

We tested the model by folding it manually into a variety of secondary structures, including a right-handed 3_{10} helix, a π-helix, an α-helix, and parallel and antiparallel β-sheets, and compared the model structures with those of the analogous crystal structures (Fig. 6, Movie S1, and *SI Appendix*, section S6). The α-helix, measured over ∼3.5 turns (measured from the α-carbon of residue 1 to the α-carbon of residue 13) in the Peppytide, is 6.78″ ± 0.16″ (equivalent to 18.43 ± 0.45 Å), which is in excellent agreement with an α-helix of same length measured at 18.4 Å in the protein structure (pid:2ZTA chain B) (Fig. 6*A*). The β-sheets made with the Peppytide model have a natural curvature as found in protein β-sheets. The parallel β-sheet measured in the Peppytide over five amides in each of the two strands (measured from the nitrogen of amide 1 to the nitrogen of amide 5 in the same strand) is 4.85″ ± 0.04″ (equivalent to 13.20 ± 0.12 Å). This agrees well with the parallel β-sheet of equivalent length in protein structure as 13.4 and 12.9 Å on the two strands (pid:202J, chain A).

We made type I and type II β-turns with the model, and compared them, respectively, with the type I β-hairpin turn found in ubiquitin (pid:1AAR, turn-seq:TLTG) (31), and the type II turn found as a subpart of the β-barrel in factor H binding protein (pid:3KVD, turn-seq:GSDD) (32) (Fig. 6*C*). Protein β-turns often contain a glycine at the R_2 or R_3 position (33). To facilitate folding into the various turn conformations, the side-chain methyl group in Peppytides can be removed to create a glycine residue. β-turn types I, I′, II, and II′ were constructed with a glycine version of the model (*SI Appendix*, Fig. S15) based on existing turn angle values (34). Type I′ and type II′ are more common in β-hairpins found in nature (35–37). All of the side chains faced outward, so it was not a problem to form the folds in the model. Interestingly, some of the turns once formed had a tendency to shift their conformations to attain greater stability. For example, the β-turn type I and II models showed a propensity to revert to their more stable counterparts, the type II′ and type I′ turns, respectively. The turns could be folded in the model even though the values of φ and ψ for the turns are different from those enforced by the magnet arrays. This was because the turns were stabilized by a combination of the conformational constraints imposed by both steric interactions and the hydrogen-bonding magnet interactions.

All secondary structures made with the model are very stable to handling due to the combined stabilizing effects of the H-bond magnets and the rotational barrier magnet arrays. The strength of the H-bond magnets (pull force, 2.49 lb) was chosen to overcome the effect of the model weight and hence the influence of gravity, while still forming stable H-bonds (*SI Appendix*, section S2). The H-bond magnets have been designed to touch and form the CO···HN bond based on the standard O–N distance in polypeptides. For two distal amides forming a hydrogen bond, the O–N distance is typically 3.00 ± 0.12 Å, from α-helix crystal structure (pid:2ZTA). In the Peppytide model, the O–N distance is 1.17″ ± 0.04″ (equivalent to 3.18 ± 0.11 Å).

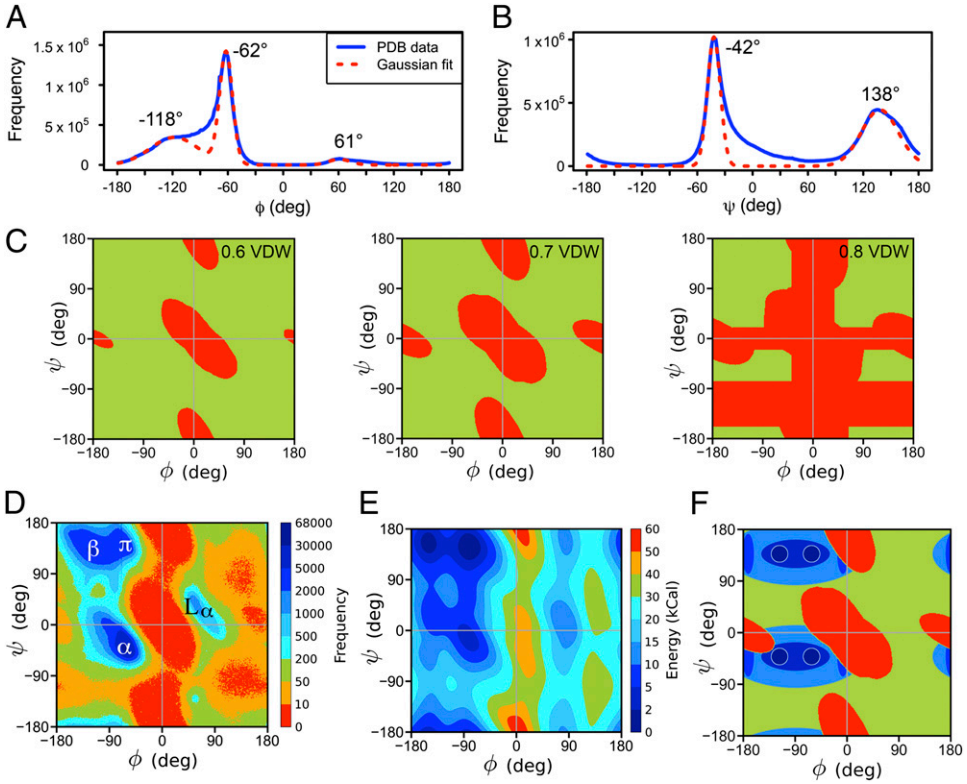

Fig. 3. Distribution of φ and ψ bond angles in the polypeptide chain, and in the Peppytide model. (*A* and *B*) Distribution of φ and ψ bond angles from 77,873 protein structures (~59 million points) from the PDB: (*A*) φ has two peaks (energy minima positions) at −62° and −118°. (*B*) ψ has two peaks (energy minima) at −42° and 138°. (*C*) Effects of steric hindrance on the conformational flexibility in the Peppytide model (measured at 5° intervals) for atom radii of 0.6, 0.7, and 0.8 R_{VDW}, respectively (from *Left* to *Right*). (*D*) Ramachandran histogram plot representation of the PDB data from *A* and *B*. (*E*) Ramachandran energy plot from the quantum mechanical OPLS force field calculations computed at 1° intervals using alanine dipeptide molecule in the Maestro platform (in kilocalories per mole). (*F*) Measured energy landscape of the Peppytide model resulting from the magnet array rotational barriers of φ and ψ bonds. The darkest blue circles are the energy minima dictated by the magnet arrays. The red regions in *C* and *F* represent sterically inaccessible conformations due to occlusion between parts.

Because of the combined stabilization of each local dihedral angle and the longer-range hydrogen bonds, Peppytides can form very stable secondary structures quite easily. The β-strands in the model can be formed with very slight human intervention with just light shaking of the chains, when the model attains the minimum energy positions, and can be easily converted into parallel or antiparallel β-sheets, or β-turns (Fig. 6 *B* and *C*). The α-helix can be easily formed with the help of a template (that can double as a stand) that facilitates the "nucleation" of the helical fold (Fig. 6*A*). Once formed, the helix is very stable to external stress/strain. To test the unfolding of the α-helix manually, we applied mechanical pull directly along the helical axis. Pulling in the direction of the helix axis does not unfold the helix easily. However, with the application of a slight unwinding torque (along the helical axis) on both sides, unfolding starts readily at the termini and gradually proceeds inward (*SI Appendix*, Fig. S16). This unfolding mechanism has been studied both experimentally (38) and computationally (39), as α-helices have been known to undergo transformation into β-sheet under mechanical pressure, and is especially relevant to amylogenic diseases. We

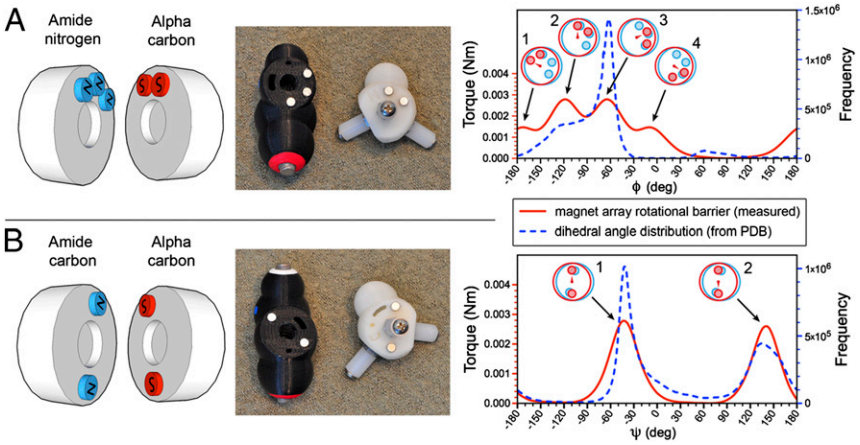

Fig. 4. Enforcing rotational barriers on the φ and ψ bonds using circular magnet arrays; N, north pole; S, south pole. (*A*) (*Left*) The coupled faces for φ, simulating the rotational constraints on φ; peaks are 56° apart. (*Center*) φ-faces of the amide and α-carbon units. (*B*) (*Left*) The coupled faces for ψ, simulating the rotational constraints on ψ; peaks are 180° apart. (*Center*) ψ-faces of the amide and α-carbon units. (*A* and *B*) (*Right*) The measured energy landscape for φ (*Upper*) and ψ (*Lower*) in the model corresponding to magnet arrangements (red), overlaid with the distribution of φ and ψ from protein structures (blue) for performance comparison; (*Insets*) the coupling of interface magnets that lead to the respective peaks in the model, blue (φ-face), red (ψ-face).

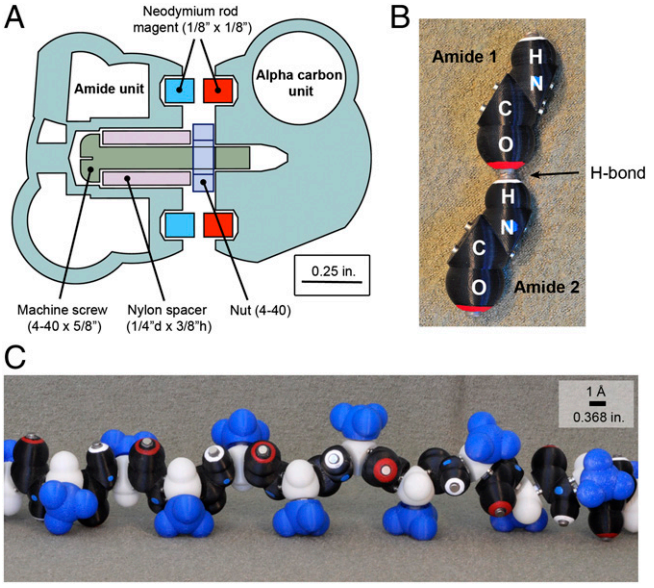

Fig. 5. Peppytide model assembly. (*A*) Bore dimensions and assembly plan for the amide unit and α-carbon unit joint (cross-section view drawn to scale); the same scheme was used for the C_α-CH_3 joint. (*B*) Representation of a hydrogen bond between distal amides. (*C*) A Peppytide homopolymer (polyalanine); black part, amide unit; white part, C_α unit; blue part, methyl group unit; red ring, oxygen; white ring, hydrogen; blue dot, nitrogen.

have also successfully folded longer Peppytide chains into several known protein conformations, including a minimal ββα motif (28-mer, pid:1FSD) (40), and fish osteocalcin (chain A; 45-mer pid:1VZM) (Fig. 6*D* and *SI Appendix*, section S6 and Fig. S14). It demonstrates that, as expected, a generic Peppytide chain can readily adopt a variety of specific folds. They can be used to make extremely complex structures, the folding of which is highly instructive.

Self-Folding and Biomimetic Modular Robotics

Although the current Peppytide model is a good tool for studying and teaching polypeptide chain folding, it also illustrates a fundamental architectural principle ubiquitous in biology: that a linear chain of modular units can be configured into a fantastic variety of 3D shapes. There is growing interest in translating this concept to the macroscopic scale to create reconfigurable objects from a universal set of modular units. The intersection of microelectronics, pervasive computing, and growing interests in biolocomotion have paved a path for the emerging field of biomimetic robotics with multimodular units working distributively to accomplish a single task. Advances are being made to fold a generic linear chain or a flat sheet into almost any 3D shape, to ultimately provide "programmable matter" (41–48). Engineers have created complex, dynamic multiunit systems that operate electronically and can interface with one another and/or a computer. These robots enable dynamic conformational information to be sent to a computer base station or to each other in real time. With the advent of miniaturized actuating technologies, this has broader impacts for future computational models for studying molecules, especially folding pathways and protein receptors, if one molecule could communicate with another wirelessly and convey its structure. Moteins, a 1D string of simple modular (polygonal or polyhedral) robots, have been shown to programmatically fold and self-assemble into 3D shapes (42). Posey is a physical construction kit that captures the shape of the assembled objects and virtually represents that in the host computer (45, 46). PolyBot, a modular robot, has been used to emulate a variety of gaits (e.g., snake-like horizontal sinusoidal motion, caterpillar-like vertical climbing motion, etc.) by propagating a wave signal through the modules (47). Inspired by paper origami, programmable folding has been used to direct the folding of a 2D sheet into various 3D shapes (43). In close analogy to protein folding, these examples have a fundamentally flexible, almost universal ability to form any given arbitrary 3D structure from a standard set of building blocks. In this respect, the Peppytide model reported here represents an important step to bridge the gap between structural biology and macroscopic design—an architectural bridge across great length scales to directly adopt nanoscale, macromolecular structural design principles to human-scale objects.

Conclusions

Despite previous efforts to build interactive physical models of biomacromolecules, there still lacks a mechanically faithful reproduction of the polypeptide chain that captures the mechanical flexibility, degrees of freedom, short- and long-range (nonbonding) interactions, all of which are essential features of the molecular system. The Peppytide model developed here reproduces several critical aspects of the natural system that impact chain dynamics including the following: (*i*) dimensional accuracy of bond lengths and bond angles, (*ii*) a faithful representation of the short-range rotational barrier imposed on all of the backbone dihedral angles, and (*iii*) long-range stabilization resulting from intrabackbone hydrogen bonding. The model is foldable into stable secondary structures of proteins with considerable accuracy, and is an excellent tool with which to intuitively understand the process of biopolymer chain folding and unfolding of tertiary structures. Because folding of linear polymer chains is a fundamental architectural concept ubiquitous in biology, tools like the Peppytide

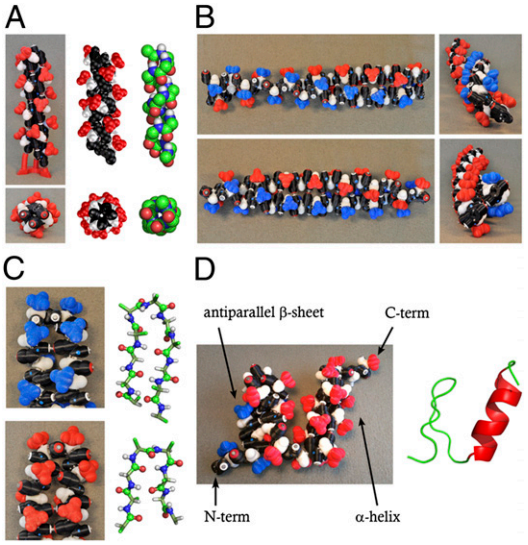

Fig. 6. Secondary and tertiary structures formed from the Peppytide model. (*A*) Comparison of a 13-mer polyalanine α-helix (R_M = 0.7 R_{VDW}): (*Left*) Peppytide physical model (alanine side chains in red); (*Center*) Peppytide in CAD representation with theoretically ideal values of $\varphi = -62°$, $\psi = -42°$; (*Right*) α-helix from crystal structure (leu-zipper pid:2ZTA, residues 16–28) (*Upper*: front view; *Lower*: top view). (*B*) Two strands of polyalanine Peppytide model folded into β-sheet conformations (with blue alanine side chains in one strand, and red in the other): (*Upper*) antiparallel, (*Lower*) parallel; the views to the *Right* show the natural curvature of the sheets. (*C*) β-turns: (*Upper*) type I in Peppytide compared with a turn in pid:1AAR, residues 4–14. (*Lower*) Type II in Peppytide compared with a turn in pid:3KVD, residues 221–228. (*D*) De novo ββα motif (pid:1FSD), a 28-mer; blue side chains indicate N-term; (*Right*) protein ribbon structure, green indicates loop and β-sheet, and red indicates α-helix.

model promise to play an important role to teach and conceive the concepts of protein folding.

The simple design of the model makes it ideal for further elaboration. An obvious next step is to expand to the full set of amino acid side chains, so that a complete protein tertiary structure can be folded. Potential further improvements to this model would be the use of softer materials, which would allow the model to sample more of conformational space while using R_{VDW} closer to 1, and to perhaps elaborate on them to include representations of electrostatic or hydrophobic forces. Sensor, actuator, or microprocessor control could also be incorporated to create a more realistic, user-friendly input/interaction device for computational tools. One application might be to make Peppytide "display" the folding pathway as a function of time, given the ability to self-fold through actuators. It should also be possible to create a model with assignable and distinct bond angles for each backbone dihedral angle, which would bias them to fold into a predetermined structure. Getting multimaterial 3D printed models to generate flexible structures or to self-fold is also an exciting possibility (49, 50).

Although polypeptides are a compelling first target for this type of model, the work is not limited to this class of compounds. Similar models of other macromolecular systems, including polynucleotides, peptidomimetics (e.g., polypeptoids and β-peptides), as well as synthetic polymers (e.g., Kevlar, conducting polymers, polystyrene, polyethyleneoxide, etc.), could be made to inform the emerging field of protein-mimetic nanostructures.

ACKNOWLEDGMENTS. We thank Prof. Alexey Onufriev (Virginia Tech), Dr. Babak Sanii, Helen Tran, and Michael Connolly (Lawrence Berkeley National Laboratory) for helpful discussions. We also thank Prof. Joseph DeRisi (University of California, San Francisco) for helpful comments as well as initial support with 3D printing. This work was supported by the Defense Threat Reduction Agency under Grant IACRO B1144571. In addition, we thank the Molecular Graphics and Computation Facility (supported by National Science Foundation Grant CHE-0840505) at the University of California, Berkeley, for use of the facilities in energy plot computations. This work was performed at the Molecular Foundry, Lawrence Berkeley National Laboratory, and was partially supported by the Office of Science, Office of Basic Energy Sciences, Scientific User Facilities Division, of the Department of Energy under Contract DE-AC02-05CH11231.

1. Dill KA, MacCallum JL (2012) The protein-folding problem, 50 years on. *Science* 338 (6110):1042–1046.
2. O'Donoghue SI, et al. (2010) Visualization of macromolecular structures. *Nat Methods* 7(s):S42–S55.
3. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graph* 14(1), 33–38, 27–28.
4. Pettersen EF, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.
5. Stone JE, Vandivort KL, Schulten K (2011) Immersive out-of-core visualization of large-size and long-timescale molecular dynamics trajectories. *Lect Notes Comput Sci* 6939:1–12.
6. Gillet A, Sanner M, Stoffler D, Olson A (2005) Tangible interfaces for structural molecular biology. *Structure* 13(3):483–491.
7. Center for BioMolecular Modeling, Milwaukee School of Engineering. Available at http://cbm.msoe.edu/. Accessed July 17, 2013.
8. Scripps Physical Model Service. Available at http://models.scripps.edu/. Accessed July 17, 2013.
9. Anderson GR (2006) Chemical modeling apparatus. US Patent 2006/0099877 A1.
10. Buist PH, Raffler AA (1991) Dynamic molecular model. US Patent 5,030,103.
11. Corey RB, Pauling L (1953) Molecular models of amino acids, peptides, and proteins. *Rev Sci Instrum* 24(8):621–627.
12. Fletterick RJ, Matela R (1982) Color-coded (alpha)-carbon models of proteins. *Biopolymers* 21(5):999–1003.
13. Fullerton LW, Roberts M, Richards JL (2010) Coded linear magnet arrays in two dimensions. US Patent 7,750,781.
14. Herman TM, et al. (2004) Molecular models. US Patent 6,793,497 B2.
15. Pauling L, Corey RB, Branson HR (1951) The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. *Proc Natl Acad Sci USA* 37(4): 205–211.
16. Roth E, Nickel A-ML, Herman TM (2008) Molecular models. US Patent 7,465,169 B2.
17. Eisenberg D (2003) The discovery of the α-helix and β-sheet, the principal structural features of proteins. *Proc Natl Acad Sci USA* 100(20):11207–11210.
18. Olson AJ, *Tangible Models (Molecular Graphics Lab)*. Available at http://mgl.scripps.edu/projects/projects/tangible_models/articulatedmodels. Accessed July 17, 2013.
19. Herman T, et al. (2006) Tactile teaching: Exploring protein structure/function using physical models. *Biochem Mol Biol Educ* 34(4):247–254.
20. Khatib F, et al. (2011) Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nat Struct Mol Biol* 18(10):1175–1177.
21. Eiben CB, et al. (2012) Increased Diels-Alderase activity through backbone remodeling guided by Foldit players. *Nat Biotechnol* 30(2):190–192.
22. Heitmann B, Job GE, Kennedy RJ, Walker SM, Kemp DS (2005) Water-solubilized, capstabilized, helical polyalanines: Calibration standards for NMR and CD analyses. *J Am Chem Soc* 127(6):1690–1704.
23. Rathore O, Sogah DY (2001) Self-assembly of beta-sheets into nanostructures by poly (alanine) segments incorporated in multiblock copolymers inspired by spider silk. *J Am Chem Soc* 123(22):5231–5239.
24. Creighton TE (1984) *Proteins: Structures and Molecular Properties* (Freeman, New York).
25. Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol* 7(1):95–99.
26. Fullerton LW, Roberts MD (2010) System and method for producing a spatial force. US Patent 7,760,058.
27. Fullerton LW, Roberts MD (2010) System and method for alignment of objects. US Patent 7,800,472.
28. Kleppner D, Kolenkow R (1973) *An Introduction to Mechanics* (McGraw-Hill, Boston).
29. Espinosa E, Molins E, Lecomte C (1998) Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. *Chem Phys Lett* 285(3-4): 170–173.
30. Agashe JS, Arnold DP (2009) A study of scaling and geometry effects on the forces between cuboidal and cylindrical magnets using analytical force solutions. *J Phys D Appl Phys* 42(9):099801.
31. Rotondi KS, Gierasch LM (2006) Natural polypeptide scaffolds: β-sheets, β-turns, and β-hairpins. *Biopolymers* 84(1):13–22.
32. Cendron L, Veggi D, Girardi E, Zanotti G (2011) Structure of the uncomplexed *Neisseria meningitidis* factor H-binding protein fHbp (rLP2086). *Acta Crystallogr Sect F Struct Biol Cryst Commun* 67(Pt 5):531–535.
33. Stanger HE, Gellman SH (1998) Rules for antiparallel β-sheet design: D-Pro-Gly is superior to L-Asn-Gly for β-hairpin nucleation. *J Am Chem Soc* 120(17):4236–4237.
34. Raghava GPS, Kaur H, A server for β-turn types prediction. Available at www.imtech.res.in/raghava/betaturns/turn.html. Accessed July 17, 2013.
35. Yang AS, Honig B (1995) Free energy determinants of secondary structure formation: II. Antiparallel β-sheets. *J Mol Biol* 252(3):366–376.
36. Chou KC, Blinn JR (1997) Classification and prediction of beta-turn types. *J Protein Chem* 16(6):575–595.
37. Chou KC (2000) Prediction of tight turns and their types in proteins. *Anal Biochem* 286 (1):1–16.
38. Miserez A, Wasko SS, Carpenter CF, Waite JH (2009) Non-entropic and reversible long-range deformation of an encapsulating bioelastomer. *Nat Mater* 8(11):910–916.
39. Qin Z, Buehler MJ (2010) Molecular dynamics simulation of the α-helix to β-sheet transition in coiled protein filaments: Evidence for a critical filament length scale. *Phys Rev Lett* 104(19):198304.
40. Dahiyat BI, Mayo SL (1997) De novo protein design: Fully automated sequence selection. *Science* 278(5335):82–87.
41. Boncheva M, et al. (2005) Magnetic self-assembly of three-dimensional surfaces from planar sheets. *Proc Natl Acad Sci USA* 102(11):3924–3929.
42. Cheung KC, et al. (2011) Programmable assembly with universally foldable strings (Moteins). *IEEE Trans Robot* 27(4):718–729.
43. Hawkes E, et al. (2010) Programmable matter by folding. *Proc Natl Acad Sci USA* 107 (28):12441–12445.
44. Raffle HS, Parkes AJ, Ishii H (2004) Topobo: A constructive assembly system with kinetic memory. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04)* (ACM, New York), pp 647–654. Available at http://dl.acm.org/citation.cfm?id=985774. Accessed July 17, 2013.
45. Weller MP, Do EY-L, Gross MD (2008) Posey: Instrumenting a poseable hub and strut construction toy. *Proceedings of the Second International Conference on Tangible and Embedded Interaction (TEI '08)* (ACM, New York), pp 39–46. Available at http://dl.acm.org/citation.cfm?id=1347402. Accessed July 17, 2013.
46. Weller MP, Gross MD, Do EY-L (2009) Tangible Sketching in 3D with Posey. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)* (ACM, New York), pp 3193–3198. Available at http://dl.acm.org/citation.cfm?doid=1520340.1520455. Accessed July 17, 2013.
47. Yim M, Duff DG, Roufas KD (2002) Walk on the wild side [modular robot motion]. *IEEE Robot Autom Mag* 9(4):49–53.
48. Zykov V, et al. (2007) Evolved and designed self-reproducing modular robotics. *IEEE Trans Robot* 23(2):308–319.
49. Oxman N (2011) Variable property rapid prototyping. *Virtual Phys Prototyp* 6(1):3–31.
50. Tibbits S (2013) 4D printing: Multi-material shape change. *Archit Des J*, in press.