Evolutionary insights into shape-shifting proteins FREE

Over millions of years a protein that now folds into two stable structures likely favored first one configuration, then the other, before settling on both.

Rachel Berkowitz

Physics Today **74** (3), 12–13 (2021);
https://doi.org/10.1063/PT.3.4693

Evolutionary insights into shape-shifting proteins

Over millions of years a protein that now folds into two stable structures likely favored first one configuration, then the other, before settling on both.

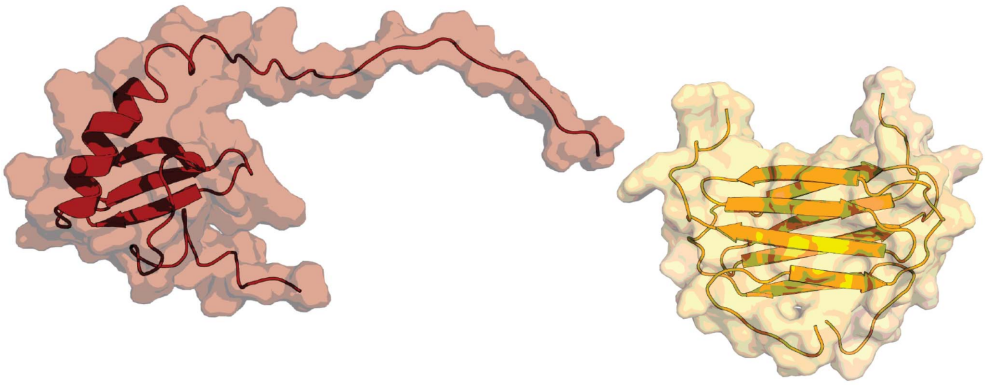

FIGURE 1. IMMUNE PROTEIN XCL1 switches between two different folded configurations that are equally stable under physiological conditions. In the chemokine fold (red), characterized by a helical structure and shared by the 50 proteins in the class, the protein binds to a receptor on a white blood cell and traffics the cell toward an infection site. In the alternative fold (gold), characterized by multiple sheet structures indicated by thick arrows, the protein directly attacks a virus or bacterium. (Courtesy of Acacia Dishman/Medical College of Wisconsin.)

According to biological paradigm, every protein has a unique stable three-dimensional structure, dictated by the specific sequence of its constituent amino acids. Through their interactions with each other, the amino acids form a well-defined, folded protein whose configuration is key to its function. Once folded, a protein may undergo conformational changes in response to environmental conditions such as temperature or pH: Subunits may rotate or hinge relative to one another to allow a channel in a cell's wall to open or close, for example. Nonetheless, the basic network of hydrogen bonds that holds the protein's shape remains unchanged.

But not all proteins follow the rules. In 2008 Brian Volkman (Medical College of Wisconsin) and colleagues reported one of the first glimpses of a human immune protein that appeared to switch seamlessly between two different folded conformations.[1] Under unchanged physiological conditions, the rule-breaking protein folded, unfolded, and refolded in equally stable structures that were held together by different hydrogen-bonded networks. How did that metamorphic behavior arise? How widespread is it? And why does it happen?

By reconstructing the likely family tree of a metamorphic protein, Volkman, graduate student Acacia Dishman, and their colleagues now identify molecular shifts that led from an ancestor that folded into a single conformation to a so-called fold-switching version that favored multiple conformations.[2] The finding suggests that metamorphic properties are not an evolutionary accident. Rather, fold switching could be a widespread adaptive trait among proteins, and the molecular sequences that encode it could offer a design strategy for engineering dual-function proteins.

Tracking a shapeshifter

The human immune system protein XCL1 was one of the first metamorphic proteins discovered. It operates mainly in the spleen and the lymph nodes and belongs to the family of 50 proteins, called chemokines, that together orchestrate the human immune response. In the chemokine fold conformation, common to all chemokines and characterized by a helical component as shown on the left in figure 1, the protein directs white blood cells to infection sites. In the alternative conformation, shown on the right, it directly kills viruses, bacteria, and fungal cells. Similar modern chemokines perform both of those functions in just one structure.

Most proteins, according to Volkman, undergo structural changes that nonetheless render them recognizable as the same protein—like an automobile whose roof opens but that still looks like a car in either state. In contrast, the two folded conformations of XCL1 don't even look at all alike. "It's like a Transformer toy, that goes from a robot back to a car," he says. In an aqueous solution, the protein switches its folding once every second, spending half its time in each conformation.

The researchers involved in XCL1's fold-switching discovery hypothesized that the protein could be an evolutionary artifact, caught in the act of changing from one version to another. As of 2020 the specific structures of six other metamorphic proteins had been examined in detail, but estimates suggest that thousands of other proteins could exhibit similar metamorphic traits.[3,4]

To investigate how a single amino acid sequence could encode two different structures and whether the shape-shifting behavior was a passing anomaly or a long-lived feature, Dishman and Volkman investigated XCL1's evolutionary history. They used software that predicted XCL1's likely ancestors based on 457 amino acid sequences from the protein's modern relatives in the chemokine family across species. The software compares amino acid sequences from different modern proteins known to share a family history and determines the sequences that are statistically most likely to have existed at dif-

ferent points going back in time.

How historical amino acid sequences folded, however, was still a mystery. (See PHYSICS TODAY, October 2008, page 20.) Dishman says, "If we reconstruct those ancestors, are they even going to look like a chemokine as we think of it today?" To find out, she produced the amino acid sequences in the lab and probed how they folded—something that cannot be predicted from first principles calculations. NMR spectroscopy revealed information about the arrangements of the amino acids in each protein after it folded into its final structure.

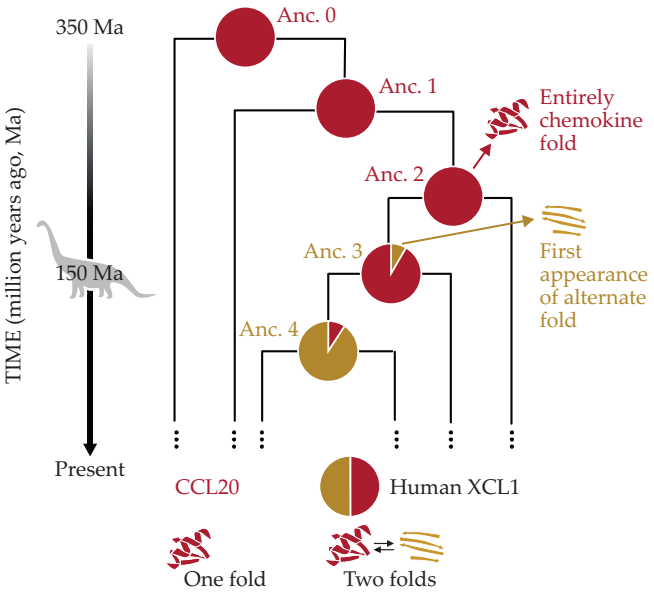

FIGURE 2. XCL1'S ANCESTORS evolved to favor an amino acid sequence that folded into the chemokine fold (red) and an alternate fold (gold). The fractional abundance of each arrangement adopted by the human immune protein shifted over time, first favoring one and then the other as represented by the pie charts, before arriving at XCL1's modern 50-50 split. Another modern protein, CCL20, retained the single fold of a 350-million-year-old Ancestor 0 (Anc. 0). Metamorphic folding may have evolved via a subset of key amino acid sequence changes that evolved from Ancestor 2 to Ancestor 3 around 150 million years ago (Ma). (Adapted from ref. 1.)

From the NMR fingerprints, the researchers found that an ancestor of XCL1, expected to have originated 350 million years ago, had just one stable structure—precisely that of a chemokine. NMR spectra of an XCL1 ancestor thought to have existed 150 million years ago contained strong peaks corresponding to the original chemokine structure, along with weak peaks indicating that 10% of molecules folded into an alternate structure. A more recent ancestor adopted mainly the novel second structure, and the modern human version, a 50-50 mixture, suggests that the dual-structure offered an evolutionary advantage.

The most exciting part, according to Dishman, is that XCL1 appears not to have evolved from one structure to another; rather, it evolved from preferring one fold to preferring a different fold and then settled on both. The fact that XCL1 folds and refolds repeatedly means that its thermodynamic stability is lower than most known proteins but not so low that it doesn't fold at all.

Family tree

After pinpointing when in the past the protein started to shift between two shapes, the researchers then deciphered the amino acid sequence that first led to XCL1's metamorphic properties. Figure 2 charts the path from the postulated oldest common ancestor shared by all chemokines, Ancestor 0, to the modern XCL1 version. In that family tree, Ancestor 2 was the last specimen to adopt just one fold, and Ancestor 3 was the first to adopt two folds. "We asked what changes in the amino acid sequence led to that shift in behavior," says Dishman.

Out of 67 amino acids that made up each of those two ancestors, 26 of them differed between the two sequences. To uncover how those sequence changes led to metamorphic behavior, the researchers analyzed the hydrogen-bond networks that formed between amino acids in each folded protein and in several other intermediate ones between Ancestors 2 and 3. The sequence changes that led from Ancestor 2 to Ancestor 3 corresponded to three structural constraints that needed to all be met for the second fold to be possible.

In one constraint, the original chemokine form of Ancestor 2 needed to incorporate new amino acids with nonpolar or sticky regions that allowed the alternate fold to arise in Ancestor 3. In another, the unfolded version needed more flexibility in some areas and more rigidity in others in order to bend into the alternate fold. Finally, the 3D chemokine fold had to be less tightly packed: The amino acids that fit together like a perfect jigsaw puzzle had to be replaced with others that meshed less well, lowering the energy barrier to structural rearrangement. If all three of those changes happened together, the protein became metamorphic; if one was missing, the metamorphism disappeared.

The ability to create artificial proteins with specific functions has opened new possibilities in drug delivery, vaccine design, functional nanomaterials, and more. Amino acid sequences that encode metamorphic proteins provide a new design strategy for researchers who want to create proteins that change their structure and function in the lab. For example, a fold-switching protein could be designed to fluoresce for use as a sensor or to serve as a moving part in a molecular machine. The sequences could also guide the search for additional metamorphic proteins in nature.

Whether metamorphic proteins are actually rare, or just rarely observed because humans have not been actively looking for them, remains to be seen. Most metamorphs have been discovered serendipitously, rather than sought out on purpose. Finding that they provide an adaptive advantage suggests that the Protein Data Bank, a global database of known proteins and their properties, could be littered with metamorphic proteins that are masquerading as monomorphs.

Rachel Berkowitz

References

1. R. L. Tuinstra et al., *Proc. Natl. Acad. Sci. USA* **105**, 5057 (2008).
2. A. F. Dishman et al., *Science* **371**, 86 (2021).
3. A. F. Dishman, B. F. Volkman, *ACS Chem. Biol.* **13**, 1438 (2018).
4. L. L. Porter, L. L. Looger, *Proc. Natl. Acad. Sci. USA* **115**, 5968 (2018).