1

The Nature of Chance

1.1 *Silk, Strength, and Statistics*

Spider silk is an amazing material. Pound for pound it is four times as strong as steel and can absorb three times as much energy as the KEVLAR from which bullet-proof vests are made (Gosline et al. 1986). Better yet, silk doesn't require a blast furnace or a chemical factory for its production; it begins as a viscous liquid produced by small glands in the abdomen of a spider, and is tempered into threads as the spider uses its legs to pull the secretion through small spigots. Silk threads, each a tenth the diameter of a human hair, are woven into webs that allow spiders to catch prey as large as hummingbirds. Incredible stuff!

The strength and resilience of spiders' silks have been known since antiquity. Indeed, anyone who has walked face first into a spiderweb while ambling down a woodland path has firsthand experience in the matter. Furthermore, the basic chemistry of silk has been known since early in this century: it is a protein, formed from the same amino acids that make our skin and muscles. But how does a biological material, produced at room temperature, get to be as strong as steel and as energy-absorbing as KEVLAR? Therein lies a mystery that has taken years to solve.

The first clues came in the 1950s with the application of X-ray crystallography to biological problems. The information provided by this technique, used so successfully by James Watson and Francis Crick to deduce the structure of DNA, allowed physical chemists to search for an orderly arrangement of the amino acids in silks—and order they found. Spider silk is what is known as a crystalline polymer. As with any protein, the amino-acid building blocks of silk are bound together in long chains. But in silks, portions of these chains are held in strict alignment—frozen parallel to each other to form crystals—and these long, thin crystals are themselves aligned with the axis of the thread. The arrangement is reminiscent of other biological crystalline polymers (such as cellulose), and in fact can account for silk's great strength.

It cannot, however, account for the KEVLAR-like ability of spider silk to absorb energy before it breaks. Energy absorption has two requirements: strength (how

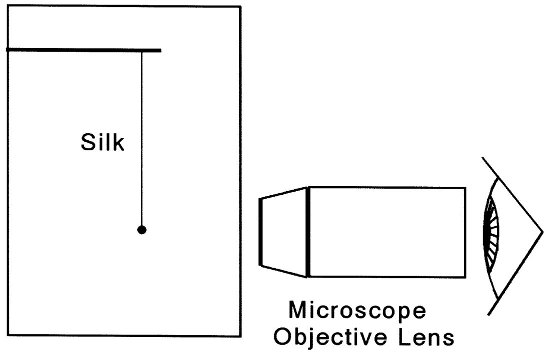

FIG. 1.1 A schematic diagram of the apparatus used to measure the change in length of a piece of spider silk. The vial in which the silk is suspended is itself immersed in a temperature-controlled water bath, and the microscope is used to track the location of the silk's free end

much force the material can resist) and extensibility (how far the material can stretch). Many crystalline polymers have the requisite strength; cellulose, for example, is almost as strong as silk. But the strength of crystalline polymers is usually gained at the loss of extensibility. Like a chain aligned with its load, a protein polymer in a crystal can extend only by stretching its links, and these do not have much give. Cellulose fibers typically can extend by only about 5% before they break. In contrast, spider silk can extend by as much as 30% (Denny 1980). As a result of this difference in extensibility, spider silk can absorb ten times more energy than cellulose. Again, how does silk do it?

This mystery went unsolved for 30 years. As powerful a tool as X-ray crystallography is, it only allows one to "see" the ordered (aligned) parts of a molecule, and the ordered parts of silk (the crystals) clearly don't allow for the requisite extension. If silk is to be extensible, some portion of its molecular structure must be sufficiently disordered to allow for rearrangement without stretching the bonds between amino acids in the protein chain. But how does one explore the structure of these amorphous molecules?

The answer arrived one dank and dreary night in Vancouver, British Columbia, as John Gosline performed an elegant, if somewhat bizarre, experiment (Gosline et al. 1984). He glued a short length of spider silk to a tiny glass rod, and, like a sinker on a fishing line, glued an even smaller piece of glass to the thread's loose end. Pole, line, and sinker were then placed in a small vial of water so that the spider silk was vertical, held taut by the weight at its end (fig. 1.1). The vial was stoppered and placed in a second container of water in front of a microscope. By circulating water through this second bath, Gosline could control the temperature of the silk without otherwise disturbing it, and by watching the weight through the microscope he could keep track of the thread's length. The stage was set.

Slowly Gosline raised the temperature. If silk behaved like a "normal" material, its length would increase as the temperature rose. Heat a piece of steel, for

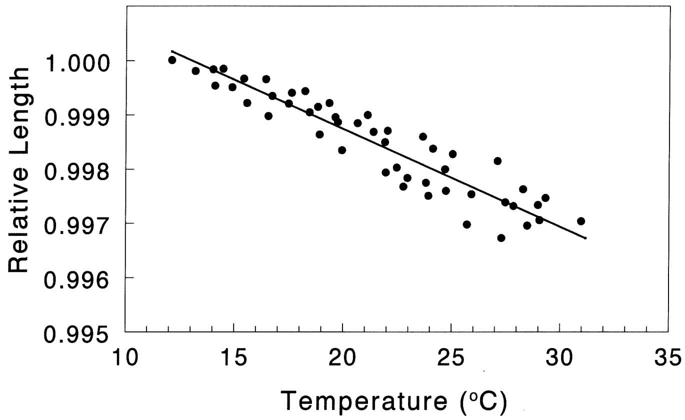

FIG. 1.2 The length of a piece of spider silk decreases slightly as temperature increases, an effect explained by the random rearrangement of the amorphous portions of the material. (Data from Gosline et al. 1984)

instance, and it will expand. In contrast, if the molecules in the amorphous portions of the silk are sufficiently free to rearrange, they should behave differently. In that case, as the temperature rose, the amorphous protein chains would be increasingly rattled by thermal agitation and should become increasingly contorted. As with a piece of string, the more contorted the molecules, the closer together their ends should be. In other words, if the amorphous proteins in silk were free to move around, the silk should get shorter as the temperature was raised. This strange effect could be predicted on the basis of statistics and thermodynamics (that's how Gosline knew to try the experiment) and had already been observed in man-made rubbery materials. Was spider silk built like steel or like rubber?

As the temperature slowly drifted up—10°C, 12°C, and higher—the silk slowly shortened (fig. 1.2). The amorphous portions of silk are a rubber! From this simple experiment, we now know a great deal about how the noncrystalline portions of the silk molecules are arranged (fig. 1.3), and we can indeed account for silk's great extensibility. Knowing the basis for both silk's strength and extensibility, we can in turn explain its amazing capacity to absorb energy—knowledge that can potentially help us to design man-made materials that are lighter, stronger, and tougher.

There are two morals to this story. First, it is not the orderly part of spiders' silk that makes it special. Instead, it is the molecular disorder, the random, ever-shifting, stochastic arrangement of amorphous protein chains that gives the material its unique properties. Second, it was a knowledge of probability and

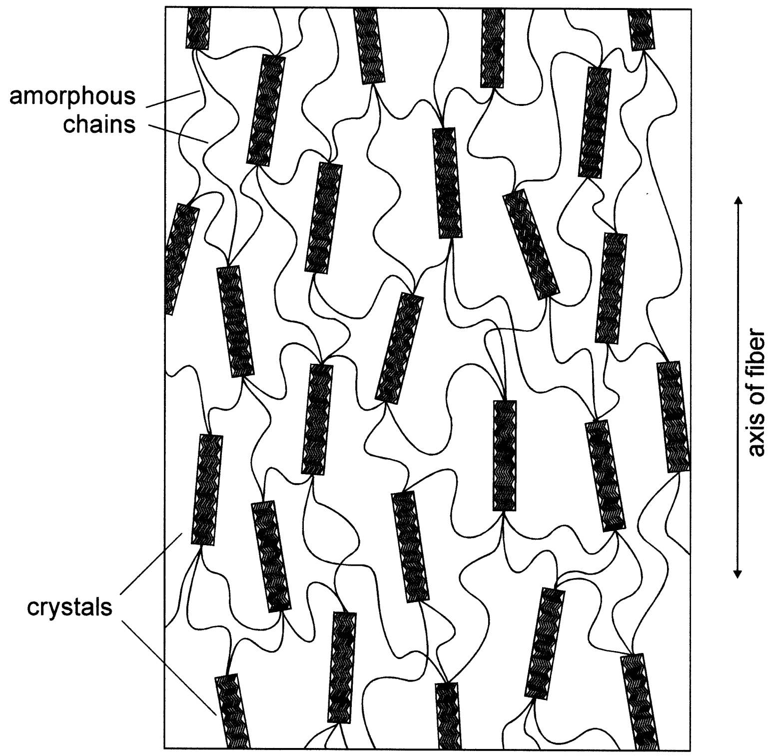

FIG. 1.3 The molecular architecture of spider silk. Crystals of ordered proteins are aligned with the silk fiber's axis, providing the material with great strength. Randomly arranged (amorphous) protein chains connect the crystals in an extensible network.

statistics that allowed Gosline to predict the consequences of this disorder and thereby perform the critical experiment. As we will see, the theory of probability can be an invaluable tool.

Probability theory was originally devised to predict the outcome in games of chance, but its utility has been extended far beyond games. Life itself is a chancy proposition, a fact apparent in our daily lives. Some days you are lucky—every stoplight turns green as you approach and you breeze in to work. Other days, just by chance, you are stopped by every light. The probability of rain coinciding with weddings, picnics, and parades is a standard worry. On a more profound level, many of the defining moments of our lives (when we are born, whom we marry, when we die) have elements of chance associated with them. However, as we have seen with spider silk, the role of chance in biology extends far beyond the random events that shape human existence. Chance is everywhere, and its role in life is the subject of this book.

1.2 *What Is Certain?*

As an instructive example, imagine yourself sitting with a friend beside a mountain stream, the afternoon sun shining through the trees overhead, the water babbling as it flows by. What can you say with absolute certainty about the scene in front of you? Well, yes, the light will get predictably dimmer as the afternoon progresses toward sunset, but if you look at one spot on the river bank you notice that there is substantial short-term variation in light intensity as well. As sunlight propagates through the foliage on its way to the ground, the random motion of leaves modulates the rays, and the intensity of light on the bank varies unpredictably both in space and in time. Yes, a leaf falling off a tree will accelerate downward due to the steady pull of gravity, but even if you knew exactly where the leaf started its fall, you would be hard-pressed to predict exactly where it would end up. Turbulent gusts of wind and the leaf's own tumbling will affect its trajectory. A close look into the stream reveals a pair of trout spawning, doing their instinctive best to reproduce. But even with the elaborate rituals and preparations of spawning, and even if all the eggs are properly fertilized, there is chance involved. Which of the parents' genes are incorporated into each gamete is a matter of chance, and which of the millions of sperm actually fertilize the hundreds of eggs is impossible to predict with precision.

Even the act of talking to a friend is fraught with chance when done next to a mountain stream. The babbling sound of the brook is pleasing because it is so unpredictable, but this lack of predictability can make communication difficult. Somehow your ears and your brain must extract from this background noise the information in speech.

So, chance in life is unavoidable. Given this fact, how should a biologist react? In many disciplines, the traditional reaction is to view the random variations of life as a necessary evil that can be exorcised (or at least tamed) through the application of clever ideas and (as a last resort) inferential statistics. Even then we are taught in our statistics classes to abhor unexplained variation. In a well-designed experiment, the less *chance* involved in the outcome, the better!

There is an alternative, however: the approach taken by Gosline in his experiment on spider silk. If chance is a given in life, why not use it to our advantage? In other words, if we know that a system will behave in a random fashion in the short term and at small scale (as with the random thermal motions of protein chains in silk), we can use this information to make accurate predictions as to how the system will behave in the long run and on a larger scale. Therein lies the thread of our tale. The diffusion of molecules, the drift of genes in a population, the longevity of phytoplankton, all include a large element of chance, and we will see why. How soft can a sound be before no animal can detect it? How fast must a mouse move in the moonlight before no owl can see it? We will be able to make predictions.

Before we embark on this exploration, we need to discuss briefly the nature of variation and which of nature's variations will (and will not) be included here.

1.3 *Determinism versus Chance*

One of Sir Isaac Newton's grand legacies is the idea that much about how the universe works can be precisely known. For example, if we know the exact mass of the moon and Earth and their current speed relative to each other, Newtonian mechanics and the law of gravitation should be able to tell us exactly where the moon is relative to Earth at any future time. As a practical matter, this is very close to being true. We can, for instance, predict solar and lunar eclipses with reasonable accuracy centuries in advance. Processes such as the moon's orbital mechanics are said to be *deterministic*, implying that, given sufficient knowledge of the initial state of a system, its future is determined exactly.

In fact, good examples of real-world deterministic processes are difficult to find. As our example of an afternoon spent observing a mountain stream is meant to convey, many of the processes that seem simple when described in the abstract (the variation in light intensity with the position of the sun, the downward acceleration of an object falling from a height) are exceedingly complex in reality. Details (rustling leaves and atmospheric turbulence) inevitably intrude, bringing with them an element of unpredictability. In some cases, the amount of variability associated with a process is sufficiently small that we are willing to view the system as being deterministic, and accept as fact predictions regarding its behavior. The physics of a pendulum clock, for instance, is so straightforward that we are content to use these machines as an accurate means of measuring time. In biology, few systems are so reliable, and deterministic behavior can be viewed at best as a polite fiction. As you might expect from the title of this book, deterministic processes will have no place here.

If a system or process is not deterministic, it is by definition *stochastic*. Even if we know *exactly* the state of a stochastic system at one time, we can never predict *exactly* what its state will be in the future. Some element of chance is involved. Unlike pregnancy and perfection, stochasticity can manifest itself to a variable degree. Many stochastic processes are approximately predictable with just a minor overlay of random behavior. The light intensity at our mountain stream is an example. Yes, there are minor random short-term fluctuations, but if we were to take 5-minute averages of the light level at the forest floor, they would closely follow predictions based on knowing the elevation of the sun. In other cases, the predictability of a system is negligible, and chance alone governs its behavior. The movement of molecules in a room-temperature gas is a good example. Both types of systems will be included in our exploration.

As a practical matter, the dividing line between "deterministic" and "stochastic" is open to interpretation. For example, it is common practice (both in sports and introductory texts on probability theory) to accept the flip of a coin as a chance proposition, a stochastic process. But if you know enough about the height above the ground at which the coin is flipped, the angular velocity initially imparted to the coin, and the effects of air resistance, it should be possible to decide in advance whether the coin will land heads up. Indeed, much of what we accept as stochastic may well be deterministic given sufficient understanding of the mechanism involved. In this respect, the line between "deterministic" and "stochastic" is often drawn as a matter of convenience. If the precise predictions that are possible in theory are too difficult to carry out in practice, we shift the line a bit and think of the process as being stochastic.

This is not to imply that *all* processes are deterministic, however. As far as physicists have been able to divine, there are aspects of nature, encountered at very small scales of time and space, that are unpredictable *even in theory*. For example, there are limits to the precision with which you can know both the velocity and the location of an object (this is a rough statement of Heisenberg's uncertainty principle). In other words, if you could know exactly where an electron is at some point in time, you couldn't know what its velocity is. Conversely, if you know exactly what its velocity is, you can't know its position. In either case, you can't predict exactly where the electron will be even a short time in the future. This is the strange realm of quantum mechanics, where chance reigns and human intuition is of little use. In this text we make scant use of the principles of quantum mechanics (a single, brief mention of the unpredictability of light emission in chapter 8). We introduce the subject here only to note that there is indeed a dividing line between "deterministic" and "stochastic," even if it is fuzzy.

1.4 *Chaos*

In recent years, this dividing line has become even fuzzier. Beginning in the late 1970s, a wide variety of physical systems that should behave deterministically were found in fact to behave unpredictably. These systems are said to exhibit *deterministic chaos*, or just *chaos* for short. But if they are deterministic, how can they be unpredictable? This apparent conflict is solved by the fact that chaotic systems are exquisitely sensitive to the state in which they are started.

Consider a "normal" deterministic system, something like the flight of a baseball through still air. In this case, if we know the initial speed of the ball (20 m s^{-1}), its initial location (home plate in Fenway Park, Boston), and its direction of motion (45° to the horizontal), we can predict where the ball will land (just beyond second base). If we make a small error in the measurement of any of these initial conditions (say, the ball is moving at 20.001 rather than 20.000 m s^{-1}), the error in predicting the ball's landing is concomitantly small.

If the motion of a baseball were chaotic, however, its flight would be quite different. Every time a chaotic ball were launched at exactly 20 m s^{-1} and an angle of exactly 45° from the center of home plate, it would land in the same spot near second base; the system is deterministic. If, however, a chaotic ball were launched with even a slight error in any of these initial conditions, its eventual landing spot would be drastically different. A shift from 20.000 to 20.001 m s^{-1} might cause it to land in the town square of, say, Emporia, Kansas. Granted, every time the ball is launched at *exactly* 20.001 m s^{-1}, it ends up in the same place in Kansas, so the system is still deterministic, but it is *extremely* sensitive to the initial conditions.[1]

This hypothetical example is intended only to provide an intuitive "feel" for the character of a chaotic system; it is sorely lacking as a definition. Unfortunately, a compact, formal definition of chaos is not easy to come by, and a lengthy explanation would be out of place here. If you want to pursue this intriguing field further, we suggest that you consult Moon (1992).

The strong sensitivity to initial conditions described above can make real-world chaotic systems appear to be stochastic. Shifts in initial conditions that are too small to be measured can cause the behavior of the system to fluctuate drastically, and these fluctuations can reasonably be assigned to "chance." The primary disadvantage of treating a chaotic system as if it were stochastic is a loss of insight. Once a process is stamped with the title "random," it is easy to stop looking for a mechanistic cause for its behavior. At present, however, there are few alternatives for understanding many chaotic systems, and in this text processes that actually may be chaotic will be treated as if they were stochastic.

[1] It is interesting to note that the motion of the planets, which has long been cited as the classical example of deterministic mechanics, is in fact chaotic. Because each planet is subject to a gravitational pull from all other planets, there is the possibility that at some time in the future the alignment of the solar system may be such that one of the planets could be thrown substantially off its present orbit, and this potential makes it virtually impossible to predict accurately where the planets will be at a given date in the future. See Peterson (1993) for an enlightening and readable discussion of chaos in planetary physics.

1.5 *A Road Map*

Our exploration of chance in biology begins in chapter 2 with a brief review of the theory of probability, culminating in a discussion of Bayes' formula and the difficulty of testing the general population for a rare disease. In chapters 3 and 4 we move from the examination of single events to the probabilities associated with large numbers of events. To this end, we introduce the notion of a Bernoulli trial and the binomial, geometric, and normal distributions.

With these basic tools in hand, we then examine the multifaceted role of random walks in biology (chapters 5 and 6). We explore the mechanics of molecular diffusion and use them to predict the maximum size of plankton in both water and air. An analogy between molecular diffusion and genetic drift allows us to predict the time expected before an allele becomes fixed or lost in a population. The nature of three-dimensional random walks is examined, leading to an explanation of why your arteries are elastic, how horses hold their heads up, and how scallops swim. It is here that we find out exactly why heated spider silk shrinks.

We then shift gears in chapter 7 and expand on our knowledge of probability distributions by exploring the statistics of extremes. We show why it is possible (but unlikely) that you will have your eardrums burst at a cocktail party, and how to predict the size of waves crashing on a rocky shore. It is here, too, that we predict the absolute limits to human longevity, the likelihood of the next 0.400 hitter in baseball, and why jet engines only occasionally flame out.

And finally, in chapter 8 we explore how our ability to see and hear the world around us is affected by the inevitable thermal and quantum noise of the environment. The trade-off between spatial and temporal resolution in sight is explained, accounting for why it is so hard to catch a ball at dusk. We predict the lower limit to size in nerve cells, and show how random noise can actually allow nerves to respond to signals that would otherwise be undetectable.