The cell as a biomaterial

Gerald H. Pollack
Department of Bioengineering 357962, University of Washington, Seattle WA 98195
ghp@u.washington.edu

For materials scientists, the cell is evidently a biomaterial – rich with polymers, surface forces, solvent-solute interactions, liquid-crystalline structures, etc. Yet, the language of the materials scientist is as foreign to the biological world as French is to Chinese. Little of the materials scientists’ perspective has been brought to bear on the question of biological function. This review aims to begin bridging the gap between the two disciplines – to show that a materials-oriented approach has power to bring fresh insights into an otherwise impenetrably complex maze. In this approach the cell is treated as a polymer gel. If the cell is a gel, then a logical approach to the understanding of cell function is through an understanding of gel function. Great strides have been made recently in understanding the principles of polymer-gel dynamics. It has become clear that a central mechanism is the phase-transition – a major structural change prompted by a subtle change of environment. Phase-transitions are capable of doing work and such work could be responsible for much of the work of the cell. Here, we pursue this approach. We set up a polymer-gel-based foundation for cell behavior, and explore the extent to which this foundation explains how the cell achieves its everyday tasks.


Introduction
For materials scientists, incursions into the biological arena have often seemed daunting. Cell-biology textbooks densely packed with complex mechanisms pose an energy barrier difficult even for the most assiduous engineer to surmount. The language is foreign. Hardly a hint of the familiar engineering-oriented analytical approach is evident in descriptions of systems replete with the kinds of polymeric structures studied regularly by materials scientists. Engineering thought is underrepresented.

Yet, it is becoming clear that the very fabric of the cell falls within the realm of the materials scientist. The cell is a network polymers, whose interaction with solvent (water) confers a gel-like consistency. This revelation is hardly new. Even before the classic book by Frey-Wyssling a half-century ago [1], the cytoplasm’s gel-like consistency had been clear to those who have had the experience of cracking open a raw egg. The gel-like consistency is obvious. The “gel-sol” transition as a central biological mechanism is increasingly debated [2, 3], as are other consequences of the cytoplasm’s gel-like consistency [4, 5]. Such phenomena are well studied by engineers, but the fruits of their understanding have not made their way into the biological arena. Most biologists have had little exposure to the physical chemistry of gels or phase-transitions.

Perhaps it is for this reason that virtually all cell biological mechanisms build on the notion of an aqueous solution – or, more specifically on free diffusion of solutes in aqueous solution. One merely needs to peruse representative textbooks to note the many diffusional steps required in proceeding from stimulus to action. These steps invariably include; ions diffusing into and out of membrane channels; ions diffusing into and out of membrane pumps; ions diffusing through the cytoplasm; proteins diffusing toward other proteins; substrates diffusing toward enzymes; etc. A cascade of diffusional steps underlies virtually every intracellular process – notwithstanding the cytoplasm’s character as a gel, where diffusion can be slow enough to be irrelevant. This odd dichotomy between theory and evidence has grown in part because modern cell biology has been pioneered by those with limited exposure to gel function, hence limited willingness to build on a gel foundation.

What havoc has such misunderstanding wrought?

Problems with the aqueous-solution-based paradigm
Consider the consequences of assuming that the cytoplasm is an aqueous solution. To keep this solution and its solutes constrained, this liquid-like milieu is surrounded by a cell membrane, which is impervious to most solutes. But solutes need to pass into and out of the cell – to nourish the cell, to effect communication between cells, to exude waste products, etc. In order for these solutes to pass into and out of the cell, the membrane requires openable pores. Well over 100 solute-specific channels have been identified, with new ones emerging regularly.

The same goes for membrane pumps: Since ion
concentrations inside and outside the cell are rarely found to be in electrochemical equilibrium, the observed concentration gradients are thought to be maintained by active pumping by specific entities lodged within the membrane. The concepts of channels and pumps will certainly be familiar to biological readers; for others, the text by Alberts et al. [6] provides a detailed review of this foundational paradigm – along with the manner in which this paradigm accounts for basic cell function. In essence, solute partitioning between the inside and the outside of the cell is assumed to be a product of an impermeable membrane, membrane pumps, and membrane channels.

How can this foundational paradigm be tested?

If partitioning requires a continuous, impermeant barrier, violating the barrier should collapse the gradients. Metabolic processes should grind to a halt, enzymes and fuel should dissipate as they diffuse out of the cytoplasm, and the cell should be quickly brought to the edge of death.

Does this really happen?

To disrupt the membrane experimentally, scientists have concocted an array of implements not unlike swords or guns:

- **Microelectrodes.** These are plunged into cells, in order to measure electrical potentials or pass substances into the cytoplasm. The microelectrode tip may seem diminutive by conventional standards, but to the 10-μm cell, invasion by a 1-μm probe is roughly akin to a human being invaded by a fence post.
- **Electroporation** is a widely used method of effecting material transfer into a cell. By shotgunning the cell with a barrage of high-voltage pulses, the membrane becomes riddled with orifices large enough to pass genes, proteins and other macromolecules – certainly large enough to pass ions.
- The **patch-clamp** method involves the plucking of a 1-μm patch of membrane from the cell for electrophysiological investigation; the cell membrane is grossly violated.

Although such insults may sometimes cause fatal injuries, in fact, they are not necessarily consequential. Consider the microelectrode plunge. The anticipated surge of ions, proteins and metabolites might be thwarted if the hole could be kept plugged by the microelectrode shank – but this cannot always be the case. Micropipettes used to microinject calcium-sensitive dyes at multiple sites along muscle cells require repeated withdrawals and penetrations, each withdrawal leaving multiple micron-sized injuries. Yet, normal function is observed for up to several days [7]. The results of patch removal are similar. Here again, the hole in question is more than a million times the cross-section of the potassium ion. Yet, following removal of the 1-μm patch, the 10 μm isolated heart cell is commonly found to live on and continue beating [8].

Similarly innocuous is the insult of electroporation. Entry of large molecules into the cell is demonstrable even when molecules are introduced into the bath up to many hours after the end of the electrical barrage [9–13]. Hence, the pores must remain open for such long periods without resealing. Nor is there evidence of resealing in structural studies following membrane disruption in muscle and nerve cells [14, 15]. Thus, notwithstanding long-term membrane orifices of macro-molecular size, the cell does not perish.

If the examples above seem too technical, consider the common alga Caulerpa, a single cell whose length can grow to several meters. This giant cell contains stem, roots, and leaves in one cellular unit undivided by any internal walls or membranes [16]. Although battered by pounding waves and gnawed on by hungry fish, such breaches of integrity do not impair survival. In fact, deliberately cut sections of stem or leaf will grow back into entire cells. Severing of the membrane is devoid of serious consequence.

Yet another example of major insult lies within the domain of experimental genetics, where cells are routinely sectioned in order to monitor the fates of the respective fragments. When cultured epithelial cells are sectioned by a sharp micropipette, the non-nucleated fraction survives for 1–2 days, while the nucleated, centrosome-containing fraction survives indefinitely and can go on to produce progeny [17]. Sectioned muscle and nerve cells survive similarly [15, 18, 19], notwithstanding the absence of membrane resealing [14, 15]. Finally, and perhaps not surprisingly in light of all that has been said, ordinary cells in the body are continually in a “membrane-wounded” state. Cells that suffer mechanical abrasion in particular, such as skin cells, gut endothelial cells, and muscle cells are especially prone to membrane wounds – as confirmed by passive entry into the cell of large tracers that ordinarily fail to enter. Yet, such cells appear structurally and functionally normal [20, 21]. The fraction of wounded cells in different tissues is variable. In cardiac muscle cells it is ~ 20%, but the fraction rises to 60% in the presence of certain kinds of performance-enhancing drugs [22]. Thus, tears in the cell membrane occur commonly and frequently even in normal, functioning tissue, possibly due to surface abrasion.

Evidently, punching holes in the membrane does not wreak havoc with the cell even though the holes may be monumental in size relative to an ion. It appears we are stuck on the horns of a dilemma. If a continuous barrier envelops the cell and is consequential for function, one needs to explain why breaching the barrier is not more consequential than the evidence indicates. On the other hand, if we entertain the possibility that the barrier may be non-continuous, so that creating yet another opening makes little difference, we then challenge the dogma on which all mechanisms of cell biological function rest, for the continuous barrier concept has become axiomatic.

Is there an escape?

If the cytoplasm is not an aqueous solution after all, then the need for a continuous barrier (with pumps and channels) becomes less acute. If the cytoplasm is a gel, for example, the membrane could be far less consequential. This argument need not imply that the membrane is absent – only that its continuity may not be essential for function. Such an approach could go a long way toward explaining the membrane-breath anomalies described above, for gels can be sliced with relative impunity,
Major insults might, or might not, be tolerable by the gel-like cell depending on the nature of the insult and the degree of cytoplasmic damage inflicted. Death is not obligatory. A continuous barrier is not required for gel integrity – just as we have seen that a continuous barrier is not required for cell integrity. A critical feature of the cytoplasm, then, may be its gel-like consistency.

**Cells as gels**

Gels are built around a scaffold of long-chain polymers, often cross-linked to one another and invested with solvent. The cytoplasm is much the same. Cellular polymers such as proteins, polysaccharides, and nucleic acids are long chained elements, frequently cross-linked to one another to form a matrix. The matrix holds the solvent (water) – which is retained even when the cell is de-membranated. “Skinned” muscle cells, for example, retain water in the same way as gels. Very much, then, the cytoplasm resembles an ordinary gel – as textbooks properly assert.

How the gel/cell matrix holds water is a matter of some importance [23], and there are at least two mechanistic possibilities. The first is osmotic: charged surfaces attract counter-ions, which draw in water. In the second mechanism, water-molecule dipoles adsorb onto charged surfaces, and subsequently onto one another in multilayers. The first mechanism is unlikely to be the prevailing one because: (1) gels placed in a sufficiently large water bath should eventually be depleted of the counter-ions on which water retention depends; yet, the hydrated gel state is retained; and (2), cytoplasm placed under high-speed centrifugation loses ions well before it loses water [24].

The second hypothesis, that charged surfaces attract water dipoles in multilayers, is an old one [25]. The thesis is that water can build layer upon layer (Fig. 1). This view had been controversial at one time, but it has been given support by several groundbreaking observations. The first is the classical observation by Pashley and Kitchener that polished quartz surfaces placed in a humid atmosphere will adsorb films of water up to 600 molecular layers thick [26]; this implies adsorption of a substantial number of layers, one upon another. The second set of observations are those of Israelachvili and colleagues, who measured the force required to displace solvents sandwiched between closely-spaced parallel mica surfaces [27–29]. The overall behavior was largely classical. However, superimposed on the anticipated monotonic response was a series of regularly spaced peaks and valleys (Fig. 2). The spacing between peaks was always equal to the molecular diameter of the sandwiched fluid. Thus, the force oscillations appeared to arise from a layering of molecules between the surfaces.

Although the Israelachvili experiments confirm molecular layering near charged surfaces, they do not prove that the molecules are linked to one another as implied in Fig. 1. However, more recent experiments using carbon-nanotube tipped AFM probes approaching flexible monolayer surfaces in water show similar layering [30], implying that the ordering does not arise merely from packing constraints; and, the Pashley/Kitchener experiment implies that many layers are possible. Hence, the kind of layering diagrammed in Fig. 1 is strongly implied. When two charged polymeric surfaces lie in proximity of one another, the water layers can bond the surfaces much like glue (Fig. 3). This is revealed in common experience. Separating two glass slides stacked face-to-face is no problem; when the slides are wet, however, separation is formidable – sandwiched water molecules cling tenaciously to the glass surfaces and to one another, preventing separation. A similar principle holds in sand: A foot will ordinarily sink deeply into dry sand at the beach, leaving a large imprint; but in wet sand, the imprint is shallow. Water clings to the sand particles, bonding them together with enough strength to support one’s full weight.

The picture that emerges, then, is that of a cytoplasmic matrix very much resembling a gel matrix. Water molecules are retained in both cases because of their affinity for the charged (hydrophilic) surfaces and their affinity for one another. The polymer matrix and

---

**Figure 1** Organization of water molecules adjacent to charged surface.

**Figure 2** Effect of separation on force between closely spaced mica plates. Only the oscillatory part of the response is shown. After Horn and Israelachvili (1981).
adsorbed water largely make up the gel. This explains why de-membranated cells retain their integrity.

Embodied in this gel-like construct are many features that have relevance for cell function. One important one is ion partitioning. The prevailing explanation for ion gradients between extracellular and intracellular compartments lies in a balance between passive flow through channels and active transport by pumps. The gel construct invites an alternative explanation. It looks toward differences of solubility between extracellular bulk water and intracellular layered, or “structured” water – as well as differences of affinity of various ions for the cell’s charged polymeric surfaces [31]. In short, structured water excludes sodium more than it excludes potassium; and, the cells’ negatively charged polymers have higher affinity for potassium than for sodium. Hence, the cytoplasm has considerably more potassium than sodium. A fuller treatment of this fundamental biological feature is given in the recent book by the author [32].

Similarly, the gel construct provides an explanation for the cell potential. The cell is filled with negatively charged polymers. These polymers attract cations. The number of cations that can enter the cell (gel), however, is restricted by their low solubility in structured water; hence, the negative charge in the cell is not balanced by cations. The residual charge amounts to approximately 0.3 mol/kg [33]. With net negative charge, the cytoplasm will have a net negative potential, which is reflected to the cell surface by the layered water dipoles. Indeed, depending on conditions, membrane-free cells can show potentials as large as 50 mV [34]. Gels made of negatively charged polymers show comparable or larger negative potentials, while positively charged polymer gels show equivalent positive potentials (Fig. 4). Membranes, pumps, and channels evidently play no role.

Hence, the gel paradigm can go quite far in explaining the cell’s most fundamental attributes – distribution of ions, and the presence of a cell potential. These are equilibrium processes; they require no energy for maintenance.

Cell dynamics

The cell is evidently not a static structure but a machine designed to carry out a multitude of tasks. Such tasks are currently described by a broad variety of mechanisms, apparently lacking any single identifiable underlying theme [at least to this author]. For virtually every process, there appears to be another mechanism.

Whether a common underlying theme might govern the cell’s many operational tasks is a question worth asking. After all, the cell began as a simple gel, and evolved from there. As it specialized, gel structure and processes gained in intricacy. Given such lineage, the potential for a simple, common, underlying, gel-based theme should not necessarily be remote. Finding a common underlying theme has been a long-term quest in other fields. In physics, for example, protegés of Einstein continue the search for a unifying force. That nature works in a parsimonious manner, employing variations of a few simple principles to carry out multitudinous actions, is an attractive notion, which I do not believe has yet been seriously pursued in the realm of cell function, although simplicity is a guiding principle in engineering.

If the cell is a gel, then a logical approach to the question of a common underlying principle of cell function is to ask whether a common underlying principle governs gel function. Gels do “function.” They undergo transition from one state to another. The process is known as a phase-transition – much like the transition from ice to water – a small change of environment causing a huge change in structure.

Such change can generate work. Just as ice formation has sufficient power to fracture hardened concrete, gel
expansion or contraction is capable of many types of work, ranging from solute/solvent separation to force generation (Fig. 5). Common examples of useful phase-transitions are the time-release capsule, in which a gel-sol transition releases bioactive drugs, and the disposable diaper, where a condensed gel undergoes enormous hydration and expansion to capture the “load.” Such behaviors are attractive in that a large change of structure can be induced by a subtle change of environment (Fig. 6).

Like synthetic gels, the natural gel of the cell may have the capacity to undergo similarly useful transitions. The question is whether they do. This question is perhaps more aptly stated a bit differently, for the cell is not a homogeneous gel but a collection of gel-like organelles, each of which is assigned a specific task. The more relevant question, then, is whether any/all such organelles carry out their function by undergoing phase-transition.

The short answer is yes – it appears that this is the case. Pursuing so extensive a theme in a meaningful way in the short space of a review article is challenging, and for a fuller development I refer the reader to the above-mentioned book [32]. In this venue I focus on a single aspect: the relevance of phase-transitions in the production of motion.

**Gels and motion**

The classes of motion produced by phase-transitions fall largely into two categories, isotropic and linear. In isotropic gels, polymers are randomly arranged, and sometimes cross-linked. Water is held largely by its affinity to polymers (or proteins, in the case of the cell). The gel is thus well hydrated — and may in the extreme contain 99.97% water [37]. In the transitioned state, the dominant polymer-water affinity gives way to a higher polymer-polymer affinity, condensing the gel into a compact mass and expelling solvent. Thus, water moves, and polymer moves.

Linear polymers also undergo transition — from extended to shortened states. The extended state is stable because it maximizes the number of polymer-water contacts and therefore minimizes the system’s energy. Water builds layer upon layer. In the shortened state the affinity of polymer for itself exceeds the affinity of polymer for water, and the polymer folds. It may fold entirely, or it may fold regionally, along a fraction of its length. As it folds, polymer and water both move. And, if a load is placed at the end of the shortening filament, the load can move as well.

Phase-transitions are inevitably cooperative: once triggered, they go to completion. The reason lies in the transition’s razor-edge behavior: once the polymer-polymer affinity (or the polymer-water affinity) begins to prevail, its prevalence increases; hence the transition goes to completion. An example is illustrated in Fig. 7. In this example, the divalent ion, calcium, cross-links the polymer strands. Its presence thereby shifts the predominant affinity from polymer-water to polymer-polymer. Once a portion of the strand is bridged, flanking segments of the polymer are brought closer together, increasing the proclivity for additional calcium bridging. Thus, local action enhances the proclivity for action in a
neighboring segment, ensuring that the reaction proceeds to completion. In this way, transitions propagate toward completion.

Evidently, the polymer-gel phase-transition can produce different classes of motion. If the cell were to exploit this principle, it could have a simple way of producing a broad array of motions, depending on the nature and arrangement of constituent polymers. In all cases, a small shift of some environmental variable such as pH, chemical content, etc., could give rise to a cooperative, all-or-none response, which could produce massive mechanical action.

As representative examples of such action, we focus briefly on two fundamental cellular processes – secretion and contraction. The first involves an isotropic transition, the second a linear one. [Additional details on these (and other) mechanisms can be found in the above-mentioned book [32].]

**Secretion**

Secretion is the mechanism by which the cell exports chemicals. The chemicals are packed into small spherical vesicles, which lie just within the cell boundary, awaiting export (Fig. 8). According to prevailing views, the secretory vesicle is a kind of “soup” surrounded by a membrane – a miniature of the prevailing view of the cell itself. For discharge, the vesicle docks with the cell membrane as cell and vesicle membranes fuse, opening the interior of the vesicle to the extracellular space, and allowing the vesicle’s contents to escape by diffusion. Although attractive in its apparent simplicity, this mechanism does not easily reconcile with several lines of evidence.

The first is that the vesicle is by no means a clear broth containing small croutons (substances to be secreted). It is a thick matrix of tangled, almost dried noodles (negatively charged polymers), invested with the croutons. Getting these croutons leave the cell by diffusing through the entwined thicket is as implausible as envisioning a school of fish escaping from an impossibly tangled net.

A second concern is the response to solvents. De-membranated vesicle matrices can be expanded and re-condensed again and again by exposure to various solutions – but the solutions are not those expected from classical theory. When condensed matrices from mast cells or goblet cells (whose matrices hydrate to produce mucus) are exposed to low osmolarity solutions – even distilled water – they remain condensed even though the osmotic draw for water ought to be enormous [38, 40]. Something keeps the network condensed, and it appears to be multivalent cations – in some cases calcium, and in other cases the molecule to be secreted, which is commonly a multivalent cation. These multivalent cations cross-link the negatively charged matrix and keep it condensed, even in the face of extremely low osmolarity solutions.

A third issue is that discharge does not appear to be a passive event. It is often accompanied by dramatic vesicle expansion. Isolated mucin-producing secretory vesicles, for example, undergo a 600–fold volume expansion within 40 ms [38]. Vesicles of nematocysts, aquatic stinging cells, are capable of linear expansion rates of 2,000 μm/ms [39]. Such phenomenal expansion rates imply something beyond mere passive diffusion of solutes and water.

Given the above-mentioned features, it is no surprise that investigators have begun looking for mechanistic clues within the realm of the phase-transition, where expansion can be large and rapid. A feature of secretory discharge consistent with this mechanism is that discharge happens or doesn’t happen depending on a critical shift of environment – the very hallmark of the polymer-gel phase-transition. As the solvent ratio (either glycerol/water or acetone/water) is edged just past a threshold, or as the temperature edges past a threshold, goblet-cell and mast-cell matrices condense or expand abruptly – the transition thresholds in both cases lying within a window as narrow as 1% of the critical value [38]. Hence, the phase-transition’s signature criterion is

---

**Figure 8** Textbook mechanism of secretion. Chemicals are packed in vesicles, which work their way to the cell surface, poised for discharge. See text for description.

**Figure 9** Phase-transition model of secretion. The phase-transition is triggered as extracellular Na⁺ replaces the multivalent cation holding the anionic network condensed.
satisfied. The abrupt expansion and hydration would allow the relevant molecules to escape into the extracellular fluid.

How such a system might work is as follows. When the condensed matrix is exposed to the extracellular space, sodium displaces the divalent cross-linker. No longer cross-linked, the polymer can satisfy its intense thirst, imbibing water and expanding explosively, like a jack-in-the-box [38]. Meanwhile, the messenger molecules are discharged. Diffusion may play some role in release; but the principal role is played by convective forces, for multivalent ions are relatively insoluble in the layered water surrounding the charged polymers [41], and will therefore be forcefully ejected. Hence, discharge into the extracellular space occurs by explosive convection.

Muscle contraction
As is now well known, muscle sarcomeres contain three filament types: thick, thin and connecting – the latter interconnecting the ends of the thick filament with respective Z-lines. All three filaments are polymers: thin filaments consist largely of repeats of monomeric actin; thick filaments are built around multiple repeats of myosin; and connecting filaments are built of titin (also known as connectin), a huge protein containing repeating immunoglobulin-like (Ig) and other domains. Together with water, which is held with extreme tenacity [24], this array of polymers forms a gel-like lattice.

Until the mid-1950s muscle contraction was held to occur by a mechanism not much different from the phase-transition mechanism to be considered. All major research groups subscribed to this view. With the discovery of interdigitating filaments in the mid-1950s, it was tempting to dump this notion, and suppose instead that contraction arose out of pure filament sliding. This supposition led Sir Andrew Huxley and Hugh Huxley to examine independently whether filaments remained at constant length during contraction. Back-to-back papers in *Nature*, using the optical microscope [42,43] appeared to confirm this supposition. The constant filament-length paradigm took hold, and has held remarkably firm ever since – notwithstanding more than 30 subsequent reports of thick filament or A-band shortening [44,45] – a remarkable disparity of theory and evidence. The motivated reader is urged to check the cited papers and judge for him/herself.

With the emerging notion of sliding filaments, the central issue became the nature of the driving force, and the swinging cross-bridge model came to be the foremost [46]. In this model, translation is driven by oar-like elements protruding from thin filaments, propelling thin filaments to slide along the thick. This mechanism explains many known features of contraction, and has therefore become broadly accepted [47–51].

On the other hand, contradictory evidence abounds. In addition to the conflicting evidence on the constancy of filament length (above), which contradicts the pure sliding model, a serious problem is the absence of compelling evidence for cross-bridge swinging [52]. Electron-spin resonance, X-ray diffraction, and fluorescence-polarization methods have produced largely negative results, as has high-resolution electron microscopy [53]. The most positive of these results has been an angle change of 3° measured on a myosin light chain [54] – far short of the anticipated 45°. Other concerns run the gamut from instability [55], to mechanics [40], structure [56,57], and chemistry [58,59]. A glance at these reviews conveys a picture different from the one in textbooks.

An alternative approach considers the possibility that the contractile mechanism does not lie in cross-bridge rotation, but in a paradigm in which all three filaments shorten. If contiguous filaments shorten synchronously, the event is global, and may qualify as a phase-transition. We consider the three filaments one at a time.

First consider the connecting filament. Shortening of the connecting filament returns the extended, unactivated, sarcomere to its unstrained length. Conversely, applied stress lengthens the filament. Shortening may involve a sequential folding of domains along the molecule, whereas stretch includes domain unfoldings – the measured length change is stepwise [60,61]. Similarly in the intact sarcomere, passive length changes also occur in steps [62], implying that each discrete event is synchronized in parallel over many filaments.

Next, consider the thick filament. Thick filament shortening could transmit force through the thin filaments, thereby contributing to active sarcomere shortening. Evidence for thick filament shortening was mentioned above: Although rarely discussed in contemporary muscle literature, the observations of thick filament length changes are extensive: they have been carried out in more than 15 laboratories worldwide, and have employed electron and light microscopic techniques on specimens ranging from crustaceans and insects, to mammalian heart and skeletal muscle – even human muscle. Evidence to the contrary is relatively rare [63]. These extensive observations cannot be summarily dismissed merely because they are not often discussed.

Thick filament shortening could not be the sole mechanism underlying contraction. If it were, the *in vitro* motility assay in which thin filaments translate over individual myosin molecules could not work, for it contains no filaments that could shorten. On the other hand, filament shortening cannot be dismissed as irrelevant, as it is so commonly seen. Thick filament shortening could contribute directly to sarcomere

![Figure 10](image-url)
shortening. It could be mediated by and alpha-helix to random-coil transition along the myosin rod [45]. The helix-coil transition is a classical phase-transition well known to biochemists – and also to those who have put a wool sweater into a hot clothes dryer and watched it shrink.

The thin filament may also shorten. There is extensive evidence that some structural change takes place along the thin filament [57, 64, 65]; also, see below. Crystallographic evidence shows that monomers of actin can pack interchangeably in either of two configurations along the filament – a ‘‘long’’ configuration, and a shorter one [56]. The difference leads to a filament length change of 10–15%. The change in actin is worth dwelling on, for although it may be more subtle than the ones above, it may be more universal, as actin filaments are contained in all eukaryotic cells.

Isolated actin filaments show prominent undulations. Known as ‘‘reptation’’ because of its snake-like character, the constituent undulations are broadly observed: in filaments suspended in solution [66], embedded in a gel [67], and gliding on a myosin-coated surface [68]. Such undulations had been presumed to be of thermal origin, but that notion is challenged by the observation that they can be substantially intensified by exposure to myosin [66] or ATP [69]; these effects imply a specific structural change rather than thermally induced change.

In fact, structural change in actin is implied by a long history of evidence. Molecular transitions had first been noted in the 1960s and 1970s [70–72]. On exposure to myosin, actin monomers underwent a 10° rotation [73]. Conformational changes have since been confirmed in probe studies, X-ray diffraction studies, phosphorescence-anisotropy studies and fluorescence-energy transfer studies, the latter showing a myosin-triggered actin-subdomain-spacing change of 17% [74].

That such structural change propagates along the filament is shown in several experimental studies. Gelsolin is a protein that bonds to one end (the so-called “barbed” end) of the actin filament; yet the impact of binding is felt along the entire filament. Molecular orientations shift by 10°, and there is threefold decrease of the filament’s overall torsional rigidity [75]. Thus, structural change induced by point binding propagates over the entire filament. Such propagated action may account for the propagated waves seen traveling along single actin filaments – observable either by cross-correlation of point displacements [76] or by tracking fluorescence markers distributed along the filament [69, 77]. In the latter, waves of shortening can be seen propagating along the filament, much like a caterpillar.

Could such a propagating structural transition drive the thin filament to slide along the thick? A possible vehicle for such action is the inchworm mechanism (Fig. 11). By propagating along the thin filament, a shortening transition could propel the thin filament to replete past the thick filament, each propagation cycle advancing the filament incrementally toward the center of the sarcomere.

Perhaps the most critical prediction of such a mechanism is the anticipated quantal advance of the thin filament. With each propagation cycle, the filament advances by a step (Fig. 12). The advance begins as an actin monomer unbinds from a myosin bridge; it ends as the myosin bridge rebinds an actin monomomer further along the thin filament. Hence, the filament-translation step size must be an integer multiple of the actin-repeat spacing (see Fig. 12). The translation step could be 1, 2 . . . or n times the actin-repeat spacing along the thin filament.

This extraordinary prediction is confirmed. The thin filament advances in steps; and, step size is an integer multiple of the actin-monomer spacing (Figs. 13, 14). This is true both in the isolated molecular system, where myosin molecules translate along actin [78], and in the intact sarcomere, where thin filaments translate past thick filaments [62, 79]. In the latter experiments, the striated image of a single myofibril is projected onto a photodiode array. The array is scanned repeatedly, producing successive traces of intensity along the myofibril axis. Hence, single sarcomere lengths can be tracked. The sarcomere-length change is consistently stepwise (Fig. 13). Analysis of many steps showed that their size is an integer multiple of the actin-monomer spacing (Fig. 14). Agreement between this result and the model’s prediction lends support to the proposed thesis. In fact, the prediction is signature-like: Conventional mechanisms might generate a step advance during each cross-bridge stroke; and, with a fortuitously sized cross-bridge swing, the step could have the appropriate size. But it is not at all clear how integer multiples of the fundamental size might be generated in a simple way.
although they are observed regularly (Fig. 14). By contrast, the detailed quantitative observations described above are direct predictions of the reptation mechanism.

In sum, contraction of the sarcomere could well arise out of contraction of each of the three filaments – connecting, thick and thin. Connecting and thick filaments appear to shorten by local phase-transitions, each condensation shortening the respective filament by an incremental step. Because these two filaments lie in series, filament shortening leads directly to sarcomere shortening. The thin filament appears to undergo a local, propagating transition, each snake-like cycle advancing the thin filament past the thick by an increment. Repeated cycles produce large-scale translation. (A similar process may occur in the in vitro motility assay, where the myosins are firmly planted on a substrate rather than in the lattice of filaments; the filament may “snake” its way along.)

The incremental steps anticipated from these transitions are observable at various levels of organization, ranging from the single myofibrillar sarcomere [62, 79], to bundles of myofibrils [80], to segments of whole fibers [81]. Hence, the transitions are global – as the phase-transition anticipates. It is perhaps no surprise that phase-transitions arise in all three elements: This endows the system with an array of features that makes muscle the versatile and effective machine that it is.

**Conclusion**

Two examples of biological motion have been presented, each plausibly driven by phase-transitions, and each producing a different type of motion. Structures such as secretory vesicles undergo isotropic condensations and expansions, whereas filamentary bundles such as actin and myosin produce linear contraction. Linear contraction can also occur in microtubules, another of nature’s linear polymers: when cross-linked into a bundle, microtubules along one edge of the bundle are often observed to shorten [82]; this may mediate bending, as in a bimetal strip. Hence, diverse motions are possible.

Given such mechanistic versatility, it would not be surprising if the phase-transition were a generic mechanism for motion production, extending well beyond the examples considered here. Phase-transitions are simple and powerful. They can bring about large-scale motions, induced by subtle changes of environment. This results in a kind of switch-like action with huge amplification. Such features seem attractive enough to imply that if nature has chosen the phase-transition as the common denominator of cell motion – and perhaps sundry other processes – it has made a wise choice.

Another wise choice is the marriage between materials science and biology. I hope it has become apparent that cell biology is a multifaceted entity whose interfaces fit naturally with those of materials science. Opportunities abound. In this review I have barely scratched the surface; much more is contained in the book on which the review is based [32]. Indeed, with many potential intellectual bridges to span the chasm of ignorance, I foresee materials scientists eventually playing a leading role in the challenge of uncovering nature’s mysteries.

**Acknowledgment**

The consent of Ebner and Sons to reprint figures from Pollack, *Cells, Gels and the Engines of Life: A New, Unifying Approach to Cell Function* (2001), is gratefully acknowledged.

Received 21 May 2002
and accepted 2 June 2002