Miniseries: Illustrating the Machinery of Life

Escherichia coli*

David S. Goodsell‡

From the Department of Molecular Biology, The Scripps Research Institute, La Jolla, California

Diverse biological data may be used to create illustrations of molecules in their cellular context. I describe the scientific results that support a recent textbook illustration of an *Escherichia coli* cell. The image magnifies a portion of the bacterium at one million times, showing the location and form of individual macromolecules. Results from biochemistry, electron microscopy, and X-ray crystallography were used to create the image.

Keywords: Cellular biology, molecular biology, molecular visualization, textbook, diagrams.

"A clear picture of the interior of a living cell that shows the average distribution of molecules at the proper scale, the proper concentration and with no missing parts, seems to me to be central to the understanding of the workings of life." This is how I began my 1991 article that presented several illustrations of *Escherichia coli* [1]. At the time, there was just enough information to create a convincing picture of the environment inside living cells. Today, most of the important features of this cell have been revealed, although many intriguing mysteries remain. For the new edition of "The Machinery of Life" [2], I updated these illustrations. I found that I was able to create an illustration where every molecule is specified, unlike the 1991 images where many generic proteins were included as placeholders with the expected size and oligomeric state.

Escherichia coli is arguably the most comprehensively studied organism known to science. A truly daunting amount of information is available for constructing an illustration of this cell. This includes information that spans the entire range from atoms to cells, including a full description of the genome, many studies of the proteome, ultrastructural studies of entire cells, and atomic structures of most of the major macromolecules [3]. This has prompted a number of efforts to create models of the entire cell and its processes [4–6].

When I started gathering materials for this illustration, I had hoped that I could use proteomic studies to provide an exact recipe for the cell. Unfortunately, this did not prove to be possible. Many proteomic studies are available, but the protein composition that they reveal is highly dependent on the environmental conditions of the cell. I settled on a hybrid approach. I took the concentrations of macromolecules from the same sources that I used in the 1991 article. This includes the overall value of 70% water for the cell, as well as the number of proteins, RNA, lipids, and other molecules. I also used the same values for the concentrations for the major players in protein synthesis, transport, and energy production. I then added information from two large-scale proteomic studies—the Gene-Protein Database [7] and a GeneChip approach that measures levels of mRNA [8]—to identify other proteins that are commonly present.

I created two images of *Escherichia coli* cells for the book. The first illustration, shown in Fig. 1, is an image of the entire cell. The magnification is such that molecules can barely be seen, but the illustration gives a feeling for the overall ultrastructure of the cell. The second illustration, shown in Fig. 2, enlarges a portion of the cell so that the shape and size of individual macromolecules may be more easily seen. A detailed key to Fig. 2 is included in Fig. 3.

Of course, there are many assumptions to be made when creating illustrations like this. For molecules that are present in high copy numbers, such as tRNA or ribosomes, I tried to give a representative distribution based on the concentration. However, many other molecules, such as DNA polymerase or flagellar motors, are found in much lower concentrations and would only be seen by chance in a randomly chosen view. I have chosen to include many of these in the illustration, however, to give a feeling for the diversity of function that is present in the cell.

Whenever possible, I based the shape and size of individual molecules on atomic structures of the protein from *Escherichia coli* or related bacteria. These structures were taken from the RCSB Protein Data Bank [9] and are referenced here using the four-character accession
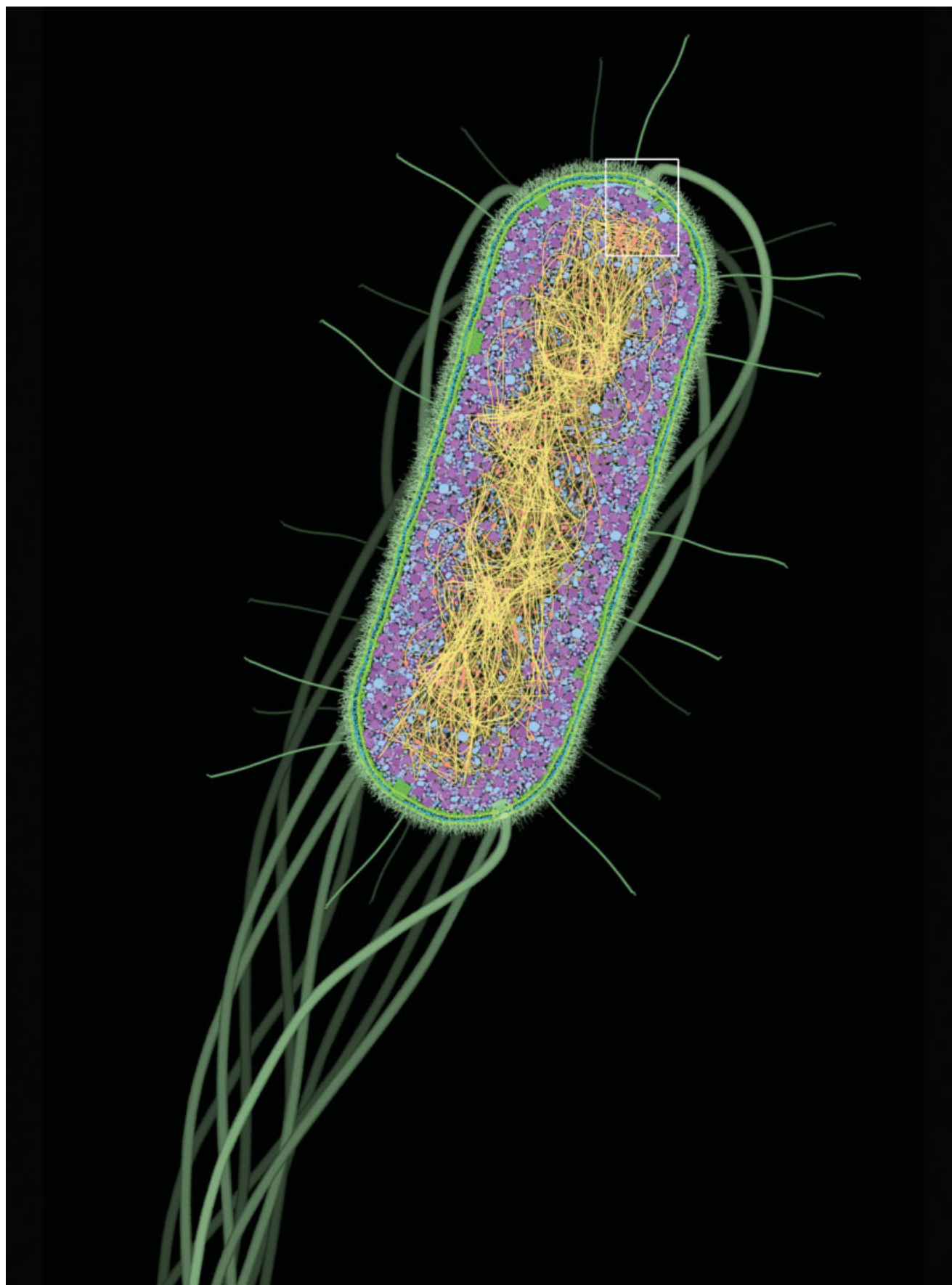
FIG. 1. **Cross section through an entire *Escherichia coli* cell.** The cell wall is in green, the cytoplasm is in magenta and blue, and the nucleoid is in yellow and orange. The boxed region is enlarged in Fig. 2. (magnification, ×70,000).
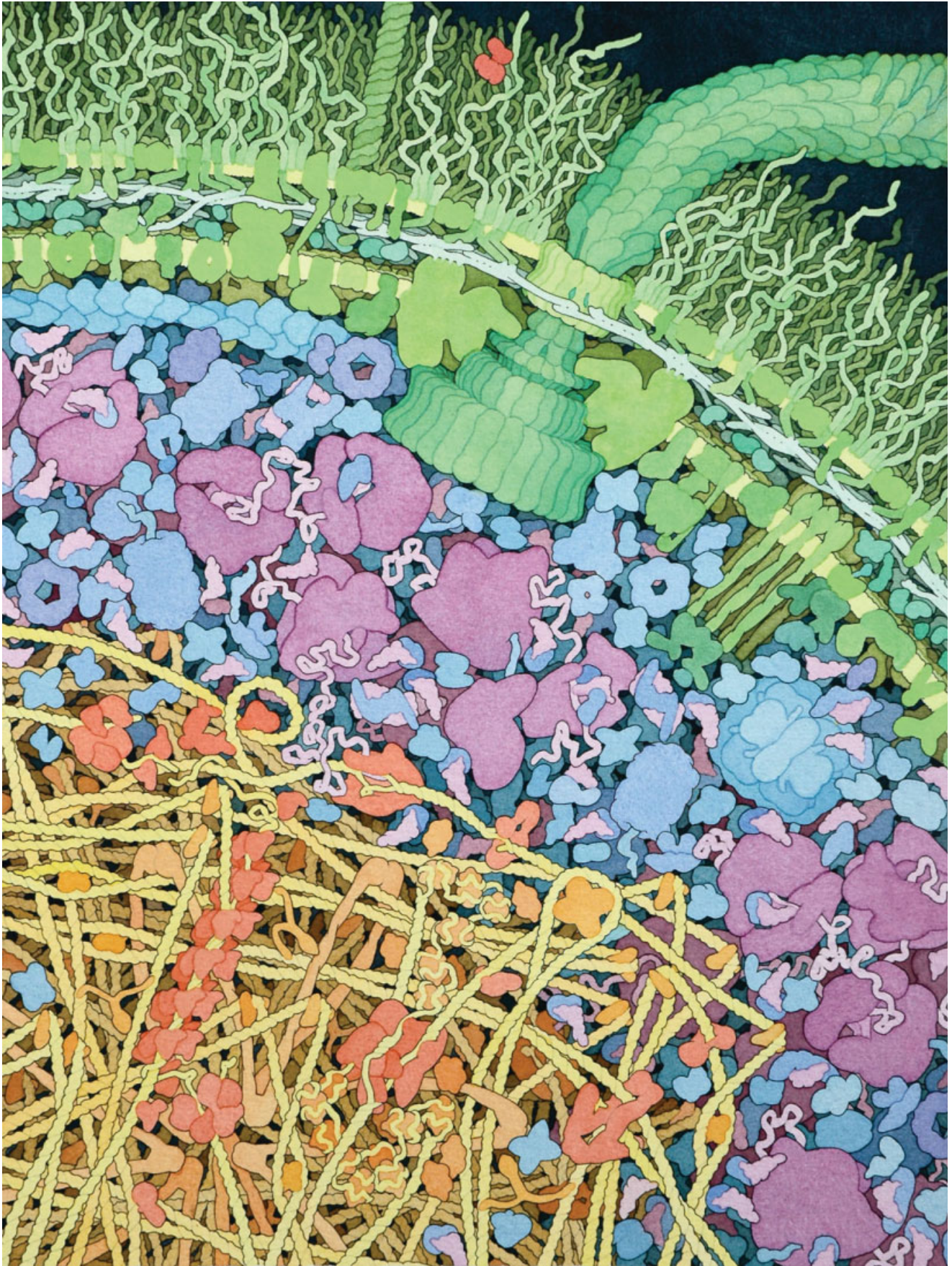
Fɪɢ. 2. **Cross section through an *Escherichia coli* cell, showing all macromolecules at a magnification of ×1,000,000.** At this magnification, individual atoms are too small to resolve (about the size of a grain of salt). The cell wall is at the top, the cytoplasm runs through the middle, and the nucleoid is at the bottom.
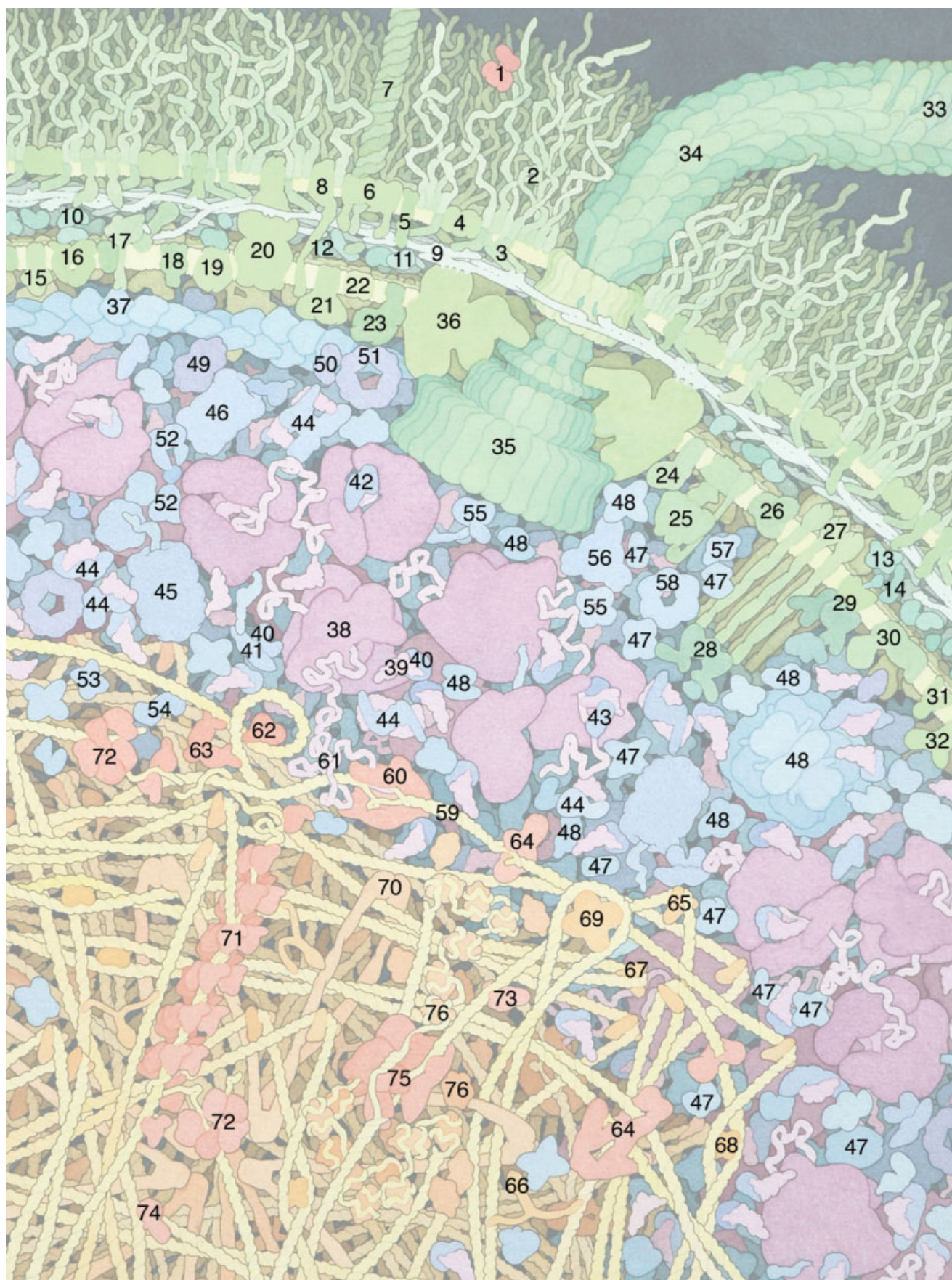
FIG. 3.

codes. I have not included the primary references to these structures, as they are easily accessible at the RCSB PDB WWW site (http://www.pdb.org).

CELL WALL

Escherichia coli cells are surrounded by a complex cell wall composed of two concentric lipid bilayers, the outer membrane and the cytoplasmic membrane with a periplasmic space in between. This cell wall plays many functional roles in protection, transport, locomotion, sensing, detoxification, and energy production.

Many molecules contribute to the structural role played by the cell wall [10, 11]. The outer membrane is filled with lipopolysaccharides—molecules that include a long polysaccharide chain connected to a complex lipid with several fatty acid tails. I based the lipid core on the structure of lipopolysaccharide bound to the outer membrane protein FhuA (1fcp). The polysaccharide chains are heterogeneous, composed of 1–40 copies of a pentasaccharide repeat, each about 15 Å long. There are about 3.5 million per cell [10].

Several proteins dominate the outer membrane. These include porin OmpF (2omf), a trimeric protein that forms a pore through the membrane. Two other plentiful proteins, OmpA and lipoprotein, interact with the peptidoglycan layer. OmpA has a large membrane spanning region (1g90) and a periplasmic domain (1r1m). Lipoprotein (1eq7), on the other hand, is tethered on the inner face of the membrane by covalently attached lipids. Estimates place 200,000 Omp proteins and 720,000 lipoproteins per cell [10]. Assuming a surface area of about 6×10^6 nm^3, this places the lipopolysaccharide about 1.3 nm apart, Omp proteins about 9.4 nm apart, and lipoprotein (trimers) about 5 nm apart. Also included is an iron transporter FhuA (1fcp) connected to the inner membrane through TonB and a fimbrial usher, described later.

The space between the membranes, the periplasmic space, is filled with peptidoglycan and a collection of small soluble proteins. The peptidoglycan appears as a dense layer in electron microscopy, 2–3 nm thick, and closer to the outer membrane than to the inner membrane. It is composed of polysaccharide chains composed of about nine disaccharide repeats (each repeat approximately 1 nm long), which are then extensively crosslinked by short peptides. The resultant network

interacts with lipoproteins and other proteins in the outer membrane, forming the primary structural support for the cell wall.

The periplasm also includes many small soluble proteins. These include periplasmic binding proteins that gather small molecules and deliver them to transporters in the cytoplasmic membrane. I have included two examples with their associated transporters: the molybdenum transporter ModA and ModBC (2onk) and the vitamin B_{12} transporter BtuF and BtuCD (2qi9). Also included are several protective enzymes, including beta-lactamase (2bls), lysozyme inhibitor (1gpq), and a monomeric copper-zinc superoxide dismutase (1eso). Two chaperone proteins are also included: heat shock protease/chaperone DegP (1ky9) and proline isomerase FkpA (1q6u).

The inner membrane is filled with proteins of many different functions. I tried to include a variety of examples, including proteins from energy production, transport, peptidoglycan synthesis, sensing, and defense, to give a cross section of the processes that are occurring. These include several transporters: magnesium transporters CorA (2bbj) and MgtE (2yvy), sodium-proton antiporter NhaA (1zcd), zinc transporter YiiP (2qfi), a calcium pump (1su4), and the two small molecule transporters mentioned earlier. A secretory channel is shown on the right side of the image, including the SecY complex in the membrane and SecA and SecB inside the cell (1rhz, 2fsg, 1ozb). The large complex that spans both membranes at the left side is a drug efflux pump AcrAB (1iwg) and TolC (1ek9).

Several enzymes of energy production are also included, such as ubiquinol oxidase (1fft), nitrate reductase NarGHI (1y4z), tricarboxylic acid cycle enzyme succinate dehydrogenase (1nek), NADH dehydrogenase (based on results from electron microscopy [12]), and ATP synthase (1c17, 1e79, 1l2p, 2a7u). A synthetic enzyme is also shown: PBP2 (penicillin binding protein 2, 2olu) is involved in the synthesis of the peptidoglycan sheath. It is associated with the structural protein MreC, discussed later.

Several of the cytoplasmic membrane proteins are involved in sensing. MscL (2oar) is one of several mechanosensitive channels that monitor the internal pressure of the cell. On the right side of the illustration, a large assembly of chemotaxis receptors are modeled after the aspartate receptor. These include array of receptors,

FIG. 3. **Key to Figure 2**. *Extracellular:* 1, Enterotoxin. *Outer membrane:* 2, lipopolysaccharide; 3, lipoprotein; 4, porin; 5, OmpA; 6, fimbrial usher; 7, pilus; 8, iron transport protein FhnA. *Periplasm:* 9, peptidoglycan; 10, periplasmic binding proteins; 11, beta-lactamase; 12, superoxide dismutase; 13, heat shock protein/chaperone DegP; 14, proline isomerase FkpA. *Inner membrane:* 15, magnesium transporter MglE; 16, vitamin B12 transporter BtuCD-F; 17, shape-determining proteins MreCD and penicillin-binding protein PBP2; 18, mechanosensory channel MscL; 19, molybdenum transporter ModBC-A; 20, drug efflux pump AcrAB and TolC; 21, magnesium transporter CorA; 22, sodium/proton antiporter NhaA; 23, nitrate reductase NarGHI; 24, succinate dehydrogenase; 25, ATP synthase; 26, ubiquinol oxidase; 27, aspartate receptor; 28, signaling proteins CheAY; 29, secretory channel SecAB; 30, NADH dehydrogenase; 31, zinc transporter YiiP; 32, calcium pump. *Flagellar motor:* 33, flagellum; 34, flagellar hook; 35, rotor; 36, motor. *Cytoplasm:* 37, cytoskeletal protein MreB; 38, ribosome; 39, transfer RNA; 40, elongation factor Tu; 41, elongation factor Ts; 42, elongation factor G; 43, initiation factors; 44, aminoacyl-tRNA synthetase; 45, chaperone GroEL; 46, proteasome HslVU; 47, glycolytic enzymes; 48, tricarboxylic acid cycle enzymes; 49, catalase; 50, Iron superoxide dismutase; 51, alkyl hydroperoxide reductase; 52, phosphoenolpyruvate: sugar phosphotransferase system; 53, nucleoside diphosphate kinase; 54, glycerol kinase; 55, acyl carrier protein system; 56, aspartate carbamoyltransferase; 57, aspartate aminotransferase; 58, glutamine synthetase. *Nucleoid:* 59, DNA; 60, RNA polymerase; 61, messenger RNA; 62, catabolite activator protein; 63, lac repressor; 64, topoisomerase; 65, HU; 66, H-NS; 67, IHF; 68, Fis; 69, Lrp; 70, condensin MukBEF; 71, RecA; 72, RecBCD; 73, DNA methyltransferase Hha1; 74, DNA glycosylase MutM; 75, DNA polymerase; 76, single strand binding protein.

modeled after two structures (1vlt and 2ch7), and the components of the Che system (1b3q), which deliver the messages inside the cell [13, 14].

FLAGELLAR MOTOR AND FIMBRIA

The flagellar motor is arguably the most impressive structure of the *Escherichia coli* cell. Typical cells have 5–10 flagella scattered at random points around the cell. The flagellum is composed of roughly 20,000 subunits, extending 5–10 μm from the cell surface [15]. It is connected to the motor with a tightly curved hook, which has roughly 130 subunits with a similar fibril structure as flagellin [16]. The motor is composed of a ring of MotA and MotB proteins surrounding a rotor composed of many different proteins. I based the structure on work from electron microscopy [17].

Fimbria is important for the attachment of cells to hosts. *Escherichia coli* can express several different types with different properties. The one shown here is a type I pilus, with a diameter of roughly 2 nm and length of 1–2 μm, composed of roughly 1,200–2,400 subunits [18]. The tip, not seen in this picture, has specialized subunits involved in adhesion. The fimbriae are extruded from the cell through an usher, shown here in the outer cell membrane [19]. The usher assembles new subunits to the end of the growing fiber, with the help of a FimC chaperone protein (1qun).

CYTOPLASMIC PROTEINS

Much of the cytoplasmic region of the bacterial cell is filled with molecules of protein synthesis. The remaining molecules are soluble enzymes performing diverse functions. As with my previous TIBS work, I included 130 glycolytic enzymes and 100 enzymes from the tricarboxylic acid cycle in each 100 nm^3 portion, which leaves about 100 enzymes of other types in this same volume. I included several of the most plentiful enzymes from proteomics studies in this collection, including aspartate carbamoyltransferase (2atc), aspartate aminotransferase (1ase), glycerol kinase (1bot), nucleoside diphosphate kinase (2hur), glutamine synthetase (2gls), alkyl hydroperoxide reductase (1n8j), catalase (1cf9), iron superoxide dismutase (1isa), and several enzymes of the phosphoenolpyruvate: sugar phosphotransferase system (2hwg, 1cm3, 1wcr) and the acyl carrier protein system (1dd8, 1i01).

There is also increasing evidence that bacterial cells have a significant cytoskeleton [20, 21]. I included the actin-like MreB protein (1jce), which is thought to form a large helical assembly inside the cytoplasmic membrane. MreC (2qf4) links this filament to the membrane and may be important for control of cell shape.

PROTEIN SYNTHESIS

Protein synthesis is one of the major tasks of the *Escherichia coli* cell, which is reflected in the composition of the cell: over a third of the molecules are involved in one way or another with production of proteins. Remarkably, atomic structures are available for most of the major molecules involved in protein synthesis. Unlike in eukaryotic cells, transcription and translation are performed in the same cellular compartment in prokaryotes, and often occur simultaneously. I included one complex at the center of the image where ribosomes have started translating an mRNA that is in the process of being transcribed. RNA polymerase was modeled after the enzyme from *Thermus thermophilus* (1iw7). Based on information in my 1991 article, there would be approximately 6 RNA polymerase molecules per 100 nm^3—I included one in foreground and one in the background of this view.

Amazingly, atomic structures are now available for ribosomes in many different states. I used two PDB files as the model for this illustration: 1yl3 and 1yl4. I used several structures for the elongation factors, including the tRNA/EF-Tu complex in 1ttt, the EF-Tu/EF-Ts complex in 1efu, and EF-G in 1dar. Structures are available for all classes of aminoacyl-tRNA synthetases—I used PDB entries 1asz, 1set, 1ffy, 1gax, 1euq, 1eiy, and 1qf6. As in my original TIBS report, I used a concentration of 30 ribosomes, 30 aminoacyl-tRNA synthetases, 340 tRNA per 100 nm^3. EF-Tu is found to be one of the most plentiful proteins in most proteomics studies [7, 8]—I included a number roughly equal to the number of tRNA molecules. I also included several chaperonin proteins. GroEL and GroES show up as being particularly plentiful in proteomics studies. I included them based on PDB entry 1aon. A HslVU proteasome is also included from PDB entry 1e94.

DNA AND DNA PACKAGING

Escherichia coli cells do not have a discrete nucleus, but they do have a loosely defined area at the center termed the nucleoid that contains most of the DNA. It is generally seen to be less densely packed with proteins than the surrounding cytoplasm, presumably due to the sieve-like exclusion of proteins by the DNA strands. A simple calculation was used to estimate the packing of DNA in the nucleoid. The *Escherichia coli* K-12 genome has 4,639,221 base pairs [22], which is about 1.5 mm in length, and a typical growing cell has 2.3 genomes per cell. The volume of the nucleoid is roughly 0.14 μm^3 [1]. To get an estimate for the packing of the DNA, we may think of the nucleoid volume as a cube with sides of about 0.52 μm, then cut the DNA into 0.52 μm lengths and stack them uniformly in the cube. From this approximation, we obtain a spacing of about 6.4 nm between DNA strands.

To calculate the number of repressors that we might expect to see in this slice, I assumed that most operons were about the size of the lac operon. It is about 5,000 base pairs long or 1,700 nm. So, if our window is a 100 nm^2, we might expect to see one repressor every 17 strands that pass through the view. I included one complex of lac repressor and catabolite activator protein in the foreground of this illustration, and we might expect to see one or two more in the background. The looped structure of the DNA is taken from the crystal structure analysis [23].

The bacterial genome is packaged by a variety of proteins that bend DNA and bridge neighboring portions of

the strand. Abundances for the DNA packaging proteins were taken from a review of the subject [24]: 50,000 histone-like protein HU, 20,000 H-NS (histone-like nucleoid structuring protein), 15,000 IHF (integration host factor), and 100,000 Fis (factor for inversion stimulation protein)—all values are the number of molecules per cell. Assuming that there are 2.3 genomes per cell, a protein with 10,000 copies per cell would be spaced about 1,000 base pairs, or 340 nm, apart. The structures of these proteins were taken from a recent review [25] and several PDB entries (3fis, 1p71, 1lr1). I also include one Lrp (leucine-responsive regulatory protein) in the view (2gqq), as well as a large star-shaped complex of condensin MukBEF [26].

I included two types of topoisomerases: a type I topoisomerase next to a transcribing RNA polymerase (1i7d) and a DNA gyrase caught in the middle of passing one DNA strand through another (1bgw, 1ei1). I also included several molecules involved in DNA repair. These include a long helical complex of RecA (2reb) and two RecBCD complexes acting on a DNA break (1w36). Also included are enzymes that correct damaged bases, such as the DNA glycosylase MutM (1kfv).

At the center of the nucleoid region, I included a replication fork. The DNA polymerase complex is based on several structures of the isolated components [27]—but the overall association of the components is still a matter of debate. The transient single-stranded regions are bound to SSB (single stranded DNA binding protein), modeled after the structure of the DNA-binding domain of the protein (1eyg).

PEDAGOGIC AND ESTHETIC CONSIDERATIONS

I wanted to capture several concepts in this illustration: that the cell is a crowded environment, that many different processes are intermingled and occur simultaneously, and that there is compartmentation even in bacterial cells. To highlight compartmentation, I chose a coloring scheme that separated each compartment. The two membranes are colored greens, with the periplasmic space in turquoise. The cytoplasm is in blues and purples, and the nucleoid is in yellows and oranges. This coloring scheme highlights the physical compartmentation—the separation of spaces by membranes and the sieve-like separation provided by the DNA—but it hides the functional localization, for instance, the fact that the cytoplasmic membrane plays multiple roles in energy production, nutrient transport, and signaling. To show this intermingled function, I created a series of "keys" for the book that isolated just the molecules involved in a particular function. These keys are similar to Fig. 3, but show the molecules of a particular function in color and the rest in gray tones. As the reader progresses through the chapter, they can compare several of these keys and explore the mixture of different functions occurring in each space.

There is also a healthy dose of artistic license applied to this type of illustration. In particular, it is always a challenge to represent fibrous and planar components in a way that is true to the actual three-dimensional geometry of the object. In this illustration, I purposefully aligned all of the DNA strands and the peptidoglycan strands approximately in the plane of the page, so that they would not be clipped. In reality, of course, we would expect a more tangled structure, with chains at all orientations. The section was chosen perpendicular to the membrane, so show the cross section through the entire cell wall. This unfortunately hides the continuous planar character of the membrane, but is necessary to reveal the different compartments.

I chose the overall view at the rounded end of the cell to give a feeling for the finite size of the cell. If I had chosen a view along one of the straight sides, there would not be any hints about the diameter of the cell. Of course, the relative size of molecules and cells is apparent in Fig. 1, which shows the whole cell, but I also wanted to capture this relationship in the enlarged cross-section. Figure 2 is a water color painting, created at twice the printed size. Figure 1 was created digitally entirely in Photoshop.

REFERENCES

[1] D. S. Goodsell (1991) Inside a living cell, *Trends Biochem. Sci.* **16**, 203–206.
[2] D. S. Goodsell (2009) The Machinery of Life, 2nd ed., Springer, New York.
[3] M. Schaechter (2001) *Escherichia coli* and *Salmonella* 2000: The view from here, *Microbiol. Mol. Biol. Rev.* **65**, 119–130.
[4] N. Nanninga (1998) Morphogenesis of *Escherichia coli*, *Microbiol. Mol. Biol. Rev.* **62**, 110–129.
[5] D. M. Morris, G. J. Jensen (2008) Toward a biomechanical understanding of whole bacterial cells, *Annu. Rev. Biochem.* **77**, 583–613.
[6] I. M. Keseler, C. Bonavides-Martinez, J. Collado-Vides, S. Gama-Castro, R. P. Gunsalus, D. A. Johnson, M. Krummenacker, L. M. Nolan, S. Paley, I. T. Paulsen, M. Peralta-Gil, A. Santos-Zavaleta, A. G. Shearer, P. D. Karp (2009) EcoCyc: A comprehensive view of *Escherichia coli* biology, *Nucleic Acids Res.* **37**, D464–470.
[7] R. A. VanBogelen, K. Z. Abshire, B. Moldover, E. R. Olson, F. C. Neidhardt (1997) *Escherichia coli* proteome analysis using the gene-protein database, *Electrophoresis* **18**, 1243–1251.
[8] R. W. Corbin, O. Paliy, F. Yang, J. Shabanowitz, M. Platt, C. E. , Jr., Lyons, K. Root, J. McAuliffe, M. I. Jordan, S. Kustu, E. Soupene, D. F. Hunt (2003) Toward a protein profile of *Escherichia coli*: Comparison to its transcription profile, *Proc. Natl. Acad. Sci. USA* **100**, 9232–9237.
[9] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne (2000) The protein data bank, *Nucleic Acids Res.* **28**, 235–242.
[10] H. Nikaido (1996) Outer Membrane. *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, F. C. Neidhardt, Ed., ASM Press, Washington, DC, pp. 29–47.
[11] H. Nikaido (2003) Molecular basis of bacterial outer membrane permeability revisited, *Microbiol. Mol. Biol. Rev.* **67**, 593–656.
[12] T. Friedrich, B. Bottcher (2004) The gross structure of the respiratory complex I: A Lego System, *Biochim. Biophys. Acta* **1608**, 1–9.
[13] G. H. Wadhams, J. P. Armitage (2004) Making sense of it all: Bacterial chemotaxis, *Nat Rev Mol Cell Biol* **5**, 1024–1037.
[14] G. L. Hazelbauer, J. J. Falke, J. S. Parkinson (2008) Bacterial chemoreceptors: High-performance signaling in networked arrays, *Trends Biochem. Sci.* **33**, 9–19.
[15] K. Yonekura, S. Maki-Yonekura, K. Namba (2003) Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy, *Nature* **424**, 643–650.
[16] F. A. Samatey, H. Matsunami, K. Imada, S. Nagashima, T. R. Shaikh, D. R. Thomas, J. Z. Chen, D. J. Derosier, A. Kitao, K. Namba (2004) Structure of the bacterial flagellar hook and implication for the molecular universal joint mechanism, *Nature* **431**, 1062–1068.
[17] G. E. Murphy, J. R. Leadbetter, G. J. Jensen (2006) In situ structure of the complete Treponema primitia flagellar motor, *Nature* **442**, 1062–1064.

[18] G. Capitani, O. Eidam, R. Glockshuber, M. G. Grutter (2006) Structural and functional insights into the assembly of type 1 pili from *Escherichia coli*, *Microbes Infect.* **8**, 2284–2290.

[19] H. Li, L. Qian, Z. Chen, D. Thibault, G. Liu, T. Liu, D. G. Thanassi (2004) The outer membrane usher forms a twin-pore secretion complex, *J. Mol. Biol.* **344**, 1397–1407.

[20] Y. L. Shih, L. Rothfield (2006) The bacterial cytoskeleton, *Microbiol. Mol. Biol. Rev.* **70**, 729–754.

[21] K. A. Michie, J. Lowe (2006) Dynamic filaments of the bacterial cytoskeleton, *Annu. Rev. Biochem.* **75**, 467–492.

[22] F. R. Blattner, G., III, Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, Y. Shao (1997) The complete genome sequence of *Escherichia coli* K-12, *Science* **277**, 1453–1474.

[23] M. Lewis, G. Chang, N. C. Horton, M. A. Kercher, H. C. Pace, M. A. Schumacher, R. G. Brennan, P. Lu (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer, *Science* **271**, 1247–1254.

[24] J. Stavans, A. Oppenheim (2006) DNA-protein interactions and bacterial chromosome architecture, *Phys. Biol.* **3**, R1–R10.

[25] M. S. Luijsterburg, M. C. Noom, G. J. Wuite, R. T. Dame (2006) The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: A molecular perspective, *J. Struct. Biol.* **156**, 262–272.

[26] R. B. Case, Y. P. Chang, S. B. Smith, J. Gore, N. R. Cozzarelli, C. Bustamante (2004) The bacterial condensin MukBEF compacts DNA into a repetitive, stable structure, *Science* **305**, 222–227.

[27] J. L. Keck, J. M. Berger (2000) DNA replication at high resolution, *Chem. Biol.* **7**, R63–R71.