The machinery and principles of vesicle transport in the cell

JAMES E. ROTHMAN

A defining feature of the eukaryotic cell is its compartmentalized cytoplasm of functionally specialized membrane-bound organelles. This internal organization creates a network of complementary reaction vessels that enable biochemical processes to occur without interference, as well as a matrix of distinct surfaces on which needed signaling modules can coalesce on demand. The molecular mechanisms by which compartments are generated and maintained are thus fundamental to biology and medicine. Discovering them posed daunting challenges two decades ago when Randy Schekman and I independently initiated the genetic and biochemical approaches. To appreciate the context, I will set the stage from my perspective and background.

By 1970 the already-classic work of George Palade had made it evident that secreted proteins are carried from the endoplasmic reticulum (ER) to the cell surface in specialized containers, or transport vesicles, that bud from one membrane and fuse with the next, transiting the Golgi stack *en route*. We now know that such intracellular protein transport is a universal process in all eukaryotes. Many kinds of vesicles traverse the cell, laden with many kinds of cargo for delivery. The result is a choreographed program of secretory, biosynthetic and endocytic protein traffic that serves the cell's internal physiologic needs, propagates its internal organization and allows it to communicate with the outside world and to receive nutrients and signals from it.

All vesicle transfer processes can be thought of as having two basic steps: budding (when the vesicle pinches off from a 'donor' membrane) and fusion (when the membrane of the vesicle merges with the 'acceptor' membrane of the intended target). The membrane fusion process has special importance for both intracellular and extracellular physiology. Fusion of vesicles within the cell must be done with exquisite specificity to prevent one organelle from taking on another's functional properties. Fusion with the cell surface (plasma) membrane (exocytosis) results in the release of the vesicle's contents, almost always consisting of highly active substances, and therefore must be exquisitely regulated. Exocytosis is used by almost every cell and tissue in the body. The dizzying array of signaling molecules secreted by exocytosis affords a veritable tour of physiology and, frequently, related diseases: neurotransmitters and their ion channel receptors, endocrine hormones like insulin, transporters for glucose and other nutrients, systemic mediators such as histamine and adrenaline, growth factors, and many others.

Setting the stage

I took up the problem in the late 1970s with a broad background outside the field that proved helpful in itself and that left me sufficiently naive to be uninhibited. I was immersed in theoretical physics from my first exposure to it as a teenager; physics taught me how to rigorously analyze the components of a problem by first imagining the form a solution would take. This can be a useful approach when engulfed in the fog that envelopes the uncharted waters of biology. A last-minute and nearly instantaneous conversion to biology occurred during my junior year in college (at Yale) as the result of an introductory lecture by Fred Richards. Happenstance then led me to Don Engelman, and so my imprinting in experimental science was in the biophysics of membranes. Yale allowed me to drop all formal course work (and yet still graduate; they probably do not permit that anymore!) to pursue full-time research, for which I will always be grateful. That year, I learned from Engelman how to dissect each morsel of data to get the most from it, and I became a scientist. Next, I entered Harvard Medical School (in 1971) with the idea of learning biology broadly (rather than practicing). I succeeded in the former, leaving the MD program more or less after the basic sciences (but with enough clinical exposure to know that leaving would be a good thing for the patients).

It was as a medical student in histology that I first learned about secretion. What an astonishing process—How could cells make vesicles from membranes? How could each vesicle know where to go? How could it fuse? It was particularly astonishing because at the time it was not even possible to begin to imagine the form a molecular solution might take. This captured my imagination, but not enough was known to productively take up the problem then. My PhD thesis with Eugene Kennedy established how the lipid bilayer is formed by asymmetric biosynthesis. Kennedy, a brilliant mind and an original thinker, taught me how to formulate a problem in biochemical terms. My year or so as a postdoctoral fellow with Harvey Lodish (at the Massachusetts Institute of Technology) was short but formative. Lodish taught me how to work with viruses and complex cell-free systems (translocation across membranes coupled to protein synthesis) and (frankly) how a large laboratory can be run boldly and energetically.

My first professorial appointment (1978) was in Stanford's biochemistry department, where I fell under the spell of the great enzymologist Arthur Kornberg (I must admit I was already under his spell before I arrived, which is why I joined the department). From Kornberg I learned two critical things at a critical time: how to formulate the strategy for a successful biochemical dissection of a complex system; and a deep faith that no matter how complex the problem, biochemistry would (eventually) succeed and would indeed provide the only sure route to the underlying molecular mechanisms. Kornberg's preaching on this subject was convincing because it stood on very solid ground. After all, cell-free reconstitution had been the central experimental approach of all biochemistry since its founding with the discovery of alcoholic fermentation in yeast extracts by the Buchner brothers at the end of the nineteenth century. And by the late 1970s, the core principles of ATP synthesis, DNA replication, RNA transcription, protein synthesis and even the genetic code were all relatively fresh 'trophies' of the reconstitution approach, which effectively strips away the subtleties of physiologic regulation to show the robust core machinery beneath.

And now it was time for the three-dimensional organization within cells to yield its secrets. Right? However, a brief sobering look at the complexities evident in electron microscope images of intracellular membranes suffices to give one pause. What made reconstituting intracellular transport seem especially
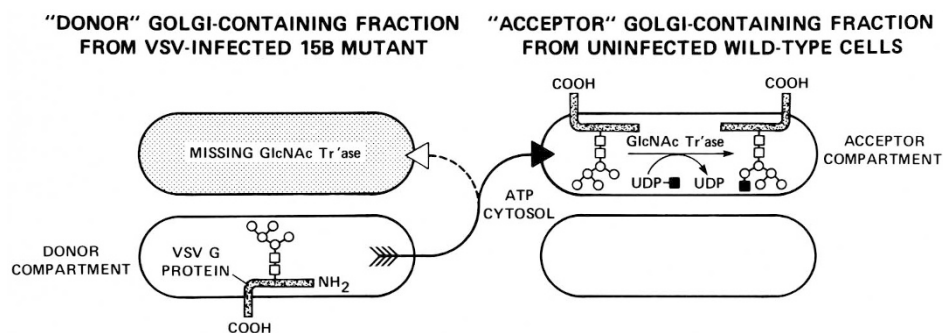
COMMENTARY

Fig. 1 Scheme for reconstitution of transport. (Reprinted from ref. 5, with permission from Elsevier Science.)

daunting was that membrane compartments engaging in transport (like the ER and the Golgi) are frequently in intimate proximity in the living cell. As a result, it was almost universally assumed that transport could never take place in homogenates, because physical proximity was believed to be the very basis on which vesicles found their target. Very distinguished cell biologists cautioned me not to bother because it would never work (the good news is that you do not have much competition that way). Their strong prejudice, which I believe was deeply rooted in the origin of cell biology as a branch of anatomy, no doubt contributed to the skepticism with which our results were greeted in the early years until the protein machinery we had isolated[1] began to match up with genes required for transport in the living cell[2].

Reconstitution of transport

Our initial goal was to detect transport of a protein between membrane-bound compartments in a cell-free extract. As it was not yet possible to express cloned genes in animal cells, we studied the transport of a membrane glycoprotein (G protein) that is copiously expressed during infection by vesicular stomatitis virus (VSV), then a very popular system that I had learned in Harvey Lodish's laboratory. The processing of G protein's oligosaccharide chains during passage through the Golgi also provided a necessary biochemical handle to follow potential transport in homogenates. One of the sugars added to VSV G protein in the Golgi is *N*-acetylglucosamine (GlcNAc).

After many trials and tribulations ('Protein Transport', Great Experiments, www.ergito.com), my first postdoctoral fellow, Erik Fries, and I found a successful formula[3]. We mixed two homogenates: One contained G protein (VSV-infected) and lacked an essential Golgi GlcNAc transferase because of mutation of the host cell (the donor); the other lacked G protein (uninfected) but contained the GlcNAc transferase (the acceptor, made from wild-type cells). If vesicles carrying G protein were to bud off from ER membranes or Golgi membranes in the donor homogenate, and then fuse with the Golgi from the acceptor homogenate, the radioactively tagged GlcNAc could be incorporated into G protein. Transfer of G protein within the donor homogenate would not produce positive results in this assay, as its Golgi lacks the needed GlcNAc transferase (Fig. 1). This design was very powerful, as transport would have to occur between organelles derived from entirely separate cells, establishing the fundamental principle that transport depends on the intrinsic chemical specificity of the membranes involved, rather than on the proximity of compartments in the cell.

Soon we learned that the donor membrane, like the acceptor membrane, is the Golgi, and as a result we had reconstituted transfer within the Golgi stack[4]. With another early, brave postdoc, William Balch, we were therefore able to simplify the assay by pre-isolating the Golgi membranes from each homogenate, and we made a number of other essential technological improvements that allowed enzymology to begin[5]. We also found that the Golgi stacks from each homogenate remained separate and unaffected in size during the cell-free reaction, and that the glycosylated G protein resides exclusively in the acceptor population. This effectively ruled out the possibility of direct fusion of the Golgi stacks under the conditions used, and indicated a vesicle intermediate[6]. Electron microscopy demonstrated copious quantities of exactly such vesicles, about 70 nm in diameter, budding from the Golgi cisternae as a result of the cell-free transport incubation[7].

Principle of vesicle budding

At this stage I began what is still an active collaboration with the great electron microscopist and cell biologist, Lelio Orci, at the University of Geneva School of Medicine. This made it possible to use electron microscope immunocytochemistry to more precisely delineate the nature and composition of the intermediates in transport. Together with a graduate student, Benjamin Glick, we confirmed that the budding transport vesicles contained VSV G protein in their membranes and found that they had a coat on their cytoplasmic surface distinct from the then-known clathrin coat[8].

The essential step in establishing the budding mechanism stemmed from my finding (during a brief sabbatical in Balch's lab, then at Yale) that transport was inhibited by a non-hydrolyzable analog of GTP, GTPγS. We then found that GTPγS blocks uncoating[9], so that the 'COP-coated' transport vesicles (now called COPI) massively accumulate, which enabled their purification by Vivek Malhotra and Tito Serafini[10]. From the vesicles came two central findings: the seven subunit 'coatomer' that assembles to constitute the coat[11]; and the discovery that the GTPase ADP ribosylation factor (ARF) is present along with coatomer in stoichiometric amounts, explaining the previously 'mysterious' effect of GTPγS. The latter also indicated the budding mechanism[12]: GTP-bound ARF recruits the coatomer to the Golgi (triggering coat assembly and vesicle budding), and releases it back to the cytosol after it hydrolyzes the GTP (uncoating).

By 1993, the validity of this simple and intuitive mechanism was confirmed using pure proteins[13]. ARF is charged with GTP at the Golgi surface, 'switching on' budding by recruiting coatomer from the cytosol. Coatomer, now locally concentrated and oriented on the membrane surface, self-assembles by polymerization into the coat, including ARF [GTP]. The growing coat acts as a mechanical device to sculpt the applied membrane into the shape of a vesicle whose size is determined by the inner diameter of the coat. The coat now forms an exoskeleton that must be shed to enable the enclosed vesicle to fuse, which occurs when ARF hydrolyzes GTP.

The same principle extends to clathrin-coated vesicles[14] and to COPII-coated vesicles budding from ER, as shown by Orci and Schekman[15]. The particular ARF GTPase family member used and the species of coatomer varies, allowing diversity in physiologic regulation (by GTP exchange/hydrolysis) and in cargo selection (by binding subunits of the coat).

**The SNARE hypothesis and the principle of membrane fusion**
This story begins in 1987 with the finding by Benjamin Glick that cell-free transport is blocked by low concentrations of the sulfhydryl alkylating reagent *N*-ethylmaleimide (NEM). Felix Wieland (on a sabbatical) and a fellow, Mark Block, then purified the *N*-ethylmaleimide-sensitive factor (NSF) based on its ability to restore transport[1]. NSF is a cytosolic protein that binds to membranes by means of the soluble NSF attachment protein (SNAP) (ref.16). With Lelio Orci and Vivek Malhotra we learned that NSF functions in the fusion process: The coated vesicles bud and lose their coats in the absence of NSF, but fail to fuse[17]. The identification of NSF as the animal equivalent of Schekman's SEC18 yeast gene was pivotal because it cemented the physiologic relevance of the mechanistic results from the cell-free system and foreshadowed the universality of the fusion mechanism[2].

But the most important insight came from the identification of the membrane receptors for SNAP, called SNAREs, which were purified by Thomas Söllner using affinity chromatography with recombinant NSF and SNAP proteins[18]. For mainly technical reasons, we used brain as the source material, which proved a most propitious choice. This enabled us to deduce how neurotransmitters are released at synapses while also affording profound insight into both the principle of membrane fusion and the basis of its inherent chemical specificity, the existence of which had been indicated by our first reconstitution experiments.

The purification yielded an unexpected result: a stoichiometric (1:1:1) complex of three membrane proteins, syntaxin, SNAP-25 and VAMP/synaptobrevin (Great Experiments, www.ergito.com). All three of these SNARE proteins were known to be present in synapses. Although previously cloned, sequenced and localized (in the laboratories of Pietro de Camilli, Reinhard Jahn, Richard Scheller, Thomas Südhof and Michael Wilson), the actual function of these proteins at the synapse was then essentially unknown.

The association of the SNAREs with the general fusion machinery NSF and SNAP indicated their involvement in membrane fusion, and the existence of a complex among them immediately indicated a simple mechanism to initiate membrane fusion[18]. VAMP resides mainly in the synaptic vesicle membrane facing into the cytosol, whereas syntaxin and SNAP-25 reside mainly in the plasma membrane, with which the synaptic vesicle fuses. A complex of all three integral membrane proteins would bring these two membranes together and position their lipid bilayers within the molecular contact range for fusion.

The SNARE hypothesis generalized this model for fusion by proposing that every vesicle in the cell has its own v-SNARE (like VAMP), and that every target membrane has its own t-SNARE (like the syntaxin–SNAP-25 dimer), and that fusion could only occur when the cognate v-SNAREs and t-SNAREs engage. The SNARE hypothesis triggered intensive research that identified and explored in detail the properties of many more SNAREs in yeast, plants and animals, including hu-
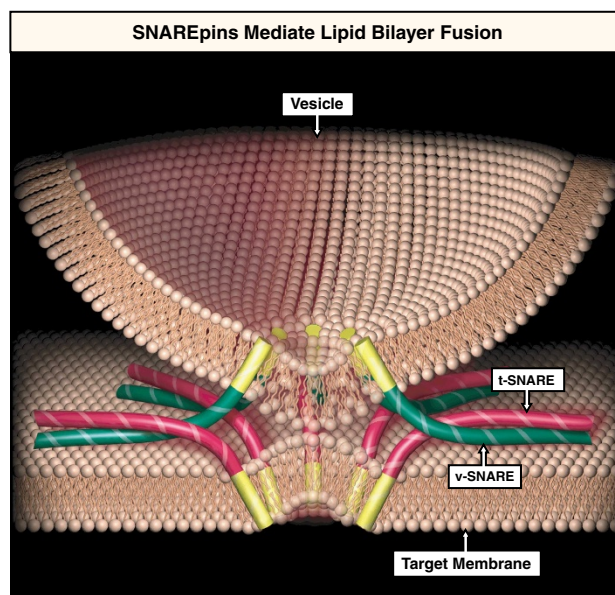
Fig. 2 An early concept of the role of the SNAREpin in mediating fusion. The t-SNARE is now known to consist of three helices and the v-SNARE of one helix. (Reproduced from ergito.com.)

mans, which, as we predicted, localize to and function at compartments engaging in fusion. In all cases, the SNARE complex structurally consists of a bundle of four α-helices, one from the v-SNARE and three from the t-SNARE (one helix from each subunit[19]).

The essential function of NSF and SNAP is to enable ongoing fusion by recycling SNAREs after fusion, using energy derived from ATP hydrolysis by NSF to disrupt the extraordinarily stable SNARE complex. Spontaneous dissociation of SNAREs is extraordinarily slow (requiring a billion years, according to recent estimates!), so were it not for NSF and SNAP, SNAREs could be used only once. This vital though indirect function of NSF and SNAP in bilayer fusion focused our attention on the simplest remaining possibility, that the SNARE complex itself, bridging two bilayers as a 'SNAREpin' (Fig. 2), is the core principle of membrane fusion.

This was tested by Thomas Weber by reconstitution of the recombinant neuronal v- and t-SNAREs into separate liposomes, which then fused spontaneously with high efficiency, albeit very slowly[20]. The t-SNARE is now known to fold into an auto-inhibited conformation that blocks access to the v-SNARE, slowing fusion; when the auto-inhibitory domain is removed, fusion occurs much more rapidly. (The extraordinarily rapid release of neurotransmitter from pre-docked vesicles after calcium entry results from v- and t-SNAREs that are already mostly 'zipped-up' between vesicle and plasma membrane. As a result, only a millisecond or so is needed to complete this process.) The generality of the biophysical principle underlying SNARE-dependent fusion is evident from the fact that enveloped viruses use 'hairpins' analogous to SNAREpins to enter cells.

The SNARE hypothesis predicts the pattern of membrane flow in the cell based on the intrinsic specificity of the cell's complement of SNARE proteins. All 11 of the potential v-SNAREs encoded in the yeast genome were tested in all possible combinations for their capacity to fuse (as isolated proteins) with the t-SNAREs representing *cis* and *trans* Golgi, plasma

membrane, early and late endosomes, and vacuoles (lysosomes). Only 10 of a total of 275 of these and other tetrameric combinations are fusogenic, and all but 1 correspond to a known transport pathway in the cell[21,22]. For example, the v-SNARE from ER only fuses with the t-SNARE of the entry face of the Golgi stack, and no other t-SNARE. This result predicts membrane flow from ER to Golgi, the first step in the secretory pathway, and so on for other cognate v- and t-SNAREs. The sole outlier predicted that the ER can fuse directly with the plasma membrane, a pathway that was not known then but has now been found[23].

Reflections and perspectives

In retrospect, 'the stars aligned' to make this work happen the way it did. I came from outside the field of cell biology and therefore had no preconceived ideas of what could not be done. The existing background made it clear that reconstitution of transport would be of crucial importance. Enough was known to successfully formulate the reconstitution experiment, and I was fortunate to be in ideal settings for the type of work involved. The climate in the Stanford biochemistry department encouraged persistent risk-taking with the long-range view in mind, and never forced me to doubt that we would succeed, eventually. This allowed the early reconstitution phase to develop and prosper. Princeton provided a productive, if brief, interlude. Molecular mechanisms were established at Sloan-Kettering in New York City, where Paul Marks provided an environment ideal for intensive fundamental research without distraction, and an opportunity to recruit and work along with wonderful membrane biochemists (including at one time Ulrich Hartl, and now Jon Goldberg, Thomas Söllner and Martin Wiedmann). And at every stage, I have been privileged to work with many very talented students and postdoctoral fellows, only some of whom I have been able to mention here.

As a result, a problem that seemed unapproachable a quarter century ago, understanding how membranes flow within the cell, has now been solved in its main features. Still, more remains to be learned. How do regulatory networks harness the now-familiar engine of core transport machinery to flexibly adjust the pace and direction of membrane flow according to need to permit an integrated response? There are no shortage of regulatory proteins or structures available to accomplish this task, either by adjusting the rate of transport or by adding layers of specificity: rab GTPases, GAPs and nucleotide exchange factors, tethering proteins, ion channels, calcium sensors, phosphoinositide kinases and binding proteins, among others, as well as lipid 'rafts' to help organize these. Looking ahead, we can expect existing genetic and biochemical approaches, combined with rapidly evolving imaging methods and high-throughput functional tests of whole genomes, to demonstrate the versatile means by which the principles of vesicle transport are used in the service of the physiology of the cell, the organ and the organism, and how imbalances can result in disease.

1. Block, M., Glick, B., Wilcox, C., Wieland, F. & Rothman, J.E. Purification of an N-ethylmaleimide-sensitive protein catalyzing vesicular transport. *Proc. Natl. Acad. Sci. USA* **85**, 7852–7856 (1988).
2. Wilson, D.W. *et al.* A fusion protein required for vesicle-mediated transport in both mammalian cells and yeast. *Nature* **339**, 355–359 (1989).
3. Fries, E. & Rothman, J.E. Transport of vesicular stomatitis virus glycoprotein in a cell-free extract. *Proc. Natl. Acad. Sci. USA* **77**, 3870–3874 (1980).
4. Fries, E. & Rothman, J.E. Transient activity of Golgi-like membranes as donors of vesicular stomatitis viral glycoprotein *in vitro. J. Cell Biol.* **90**, 697–704 (1981).
5. Balch, W., Dunphy, W., Braell, W. & Rothman, J.E. Reconstitution of the transport of protein between successive compartments of the Golgi measured by the coupled incorporation of N-acetylglucosamine. *Cell* **39**, 405–416 (1984).
6. Braell, W., Balch, W., Dobbertin, D. & Rothman, J.E. The glycoprotein that is transported between successive compartments of the Golgi in a cell-free system resides in stacks of cisternae. *Cell* **39**, 511–524 (1984).
7. Balch, W., Glick, B. & Rothman, J.E. Sequential intermediates in the pathway of intercompartmental transport in a cell-free system. *Cell* **39**, 525–536 (1984).
8. Orci, L., Glick, B.S. & Rothman, J.E. A new type of coated vesicular carrier that appears not to contain clathrin: Its possible role in protein transport within the Golgi stack. *Cell* **46**, 171–184 (1986).
9. Melançon, P. *et al.* Involvement of GTP-binding "G" proteins in transport through the Golgi stack. *Cell* **51**, 1053–1062 (1987).
10. Malhotra, V., Serafini, T., Orci, L., Shepherd, J.C. & Rothman, J.E. Purification of a novel class of coated vesicles mediating biosynthetic protein transport through the Golgi stack. *Cell* **58**, 329–336 (1989).
11. Waters, M., Serafini, T. & Rothman, J.E. 'Coatomer': A cytosolic protein complex containing subunits of non-clathrin-coated Golgi transport vesicles. *Nature* **349**, 248–251 (1991).
12. Serafini, T. *et al.* ADP-ribosylation factor is a subunit of the coat of Golgi-derived COP-coated vesicles: a novel role for a GTP-binding protein. *Cell* **67**, 239–253 (1991).
13. Orci, L., Palmer, D.J., Amherdt, M. & Rothman J.E. Coated vesicle assembly in the Golgi requires only coatomer and ARF proteins from the cytosol. *Nature* **364**, 732–734 (1993).
14. Stamnes, M.A. & Rothman, J.E. The binding of AP-1 clathrin adaptor particles to Golgi membranes requires ADP-ribosylation factor, a small GTP-binding protein. *Cell* **73**, 999–1005 (1993).
15. Barlowe, C. *et al.* COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell* **77**, 895–907 (1994).
16. Clary, D.O., Griff, I.C. & Rothman, J.E. SNAPs, a family of NSF attachment proteins involved in intracellular membrane fusion in animals and yeast. *Cell* **61**, 709–721 (1990).
17. Malhotra, V., Orci, L., Glick, B., Block, M. & Rothman, J.E. Role of an N-ethylmaleimide-sensitive transport component in promoting fusion of transport vesicles with cisternae of the Golgi stack. *Cell* **54**, 221–227 (1988).
18. Söllner T.H. *et al.* SNAP receptors implicated in vesicle targeting and fusion. *Nature* **362**, 318–324 (1993).
19. Sutton, R.B., Fasshauer, D., Jahn, R. & Brunger, A.T. Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature* **395**, 347–353 (1998).
20. Weber, T. *et al.* SNAREpins: minimal machinery for membrane fusion. *Cell* **92**, 759–772 (1998).
21. McNew, J. *et al.* Compartmental specificity of cellular membrane fusion encoded in SNARE proteins. *Nature* **407**, 153–159 (2000).
22. Parlati, F. *et al.* Distinct SNARE complexes mediating membrane fusion in Golgi transport based on combinatorial specificity. *Proc. Natl. Acad. Sci. USA* **99**, 5424–5429 (2002).
23. Gagnon, R. *et al.* Endoplasmic reticulum-mediated phagocytosis is a mechanism of entry into macrophages. *Cell* **110**, 119–131 (2002).

James E. Rothman

Cellular Biophysics and Biochemistry Program

Sloan-Kettering Institute

Memorial Sloan-Kettering Cancer Center

New York, New York, USA