Supporting Information

Mahmoudabadi et al. 10.1073/pnas.1701670114

I. Introduction: Viral Rewiring of the Host Metabolism

Viruses rely entirely on their host as an energy source. Instead of passively exploiting the host's metabolic energy, some viruses appear to augment it (2). A particularly compelling example is demonstrated by phages that infect cyanobacteria. Cyanophages carry genes for photosystem II, high-light inducible protein, transaldolase, and ribonucleotide reductase, which are all transcribed during an infection (3). Given the unprecedented presence of photosynthetic genes in viral genomes and the active expression of these genes during an infection, it is proposed that cyanophages carry these genes to increase the host energy supply and deoxynucleotide production for their own replication (3). An analogous finding is the presence of sulfite reductase genes in genomes of phages that infect deep-sea bacteria, which use sulfur as their energy source (36). Here, too, these phages are hypothesized to add to the host's metabolic output. It was also shown that large DNA algal viruses encode deoxynucleotide synthesis enzymes. For example, PBCV-1 encodes 13 nucleotide metabolism enzymes and EsV-1 encodes an ATPase and both subunits of ribonucleotide reductase (34). The most recent study on this topic identified more than 200 virus-encoded auxiliary metabolic genes such as those used in nitrogen and sulfur metabolism in marine viral metagenomes (6).

Adenoviruses have been shown to reprogram the host's glutamine metabolism by up-regulating glutamine transporters and glutamine catabolism enzymes. Glutamine is a critical amino acid used in the synthesis and import of other amino acids. Interestingly, this viral rewiring of glutamine metabolism is shown to boost the concentration of certain amino acids as well as increase glutamine reductive carboxylation. Together, these effects are required for optimal viral production not only during an adenovirus infection but also during herpes and influenza viral infections (4). In addition to virus-infected cells, which have high demand for molecular building blocks and energy, cancer cells, immune cells, and other proliferating cells similarly rewire glutamine metabolism for energy production and biosynthesis (44).

Moreover, several examples from the early years of virology have shown that viruses such as Rous sarcoma virus and feline leukemia virus increase their hosts' glycolytic rate upon infection (35). Similarly, the vaccinia virus was shown to up-regulate mitochondrial genes involved in the electron transport chain (ETC), thereby increasing the ATP production within its host (5).

II. Energetic Cost Definitions and Assumptions

Viral synthesis requires the expenditure of the host's ATP-equivalent molecules as well as the usurpation of the host's monomeric building blocks such as nucleotides, amino acids, and in the case of some viruses, lipids. To synthesize these monomeric building blocks and generate ATP-equivalent molecules, heterotrophic cells, such as the hosts of T4 or influenza, rely on reduced carbon sources from their environment. In calculating the energetic cost of a viral infection and its impact on the host energy supply, it is critical that we state several assumptions about the host growth conditions. First, we assume that the host is growing aerobically at 37 °C, with glucose as the sole carbon source. Second, we assume that sources of nitrogen, sulfur, phosphorus, and other trace elements are in excess, which is typical of culture conditions in the laboratory, and from which burst size measurements are commonly obtained. Third, we assume that the growth media contain only inorganic sources of nitrogen, requiring cells to synthesize amino acids rather than salvaging them from a growth medium supplemented with peptides (although this assumption should be modified in the case of a mammalian host cell that cannot synthesize all amino acids).

As the sole carbon source in the growth media, glucose will serve both as a source of energy and biomass. As glucose molecules enter the cell, some are fully metabolized through glycolysis and the tricarboxylic acid (TCA) cycle, and result in the formation of carbon dioxide, water, and ATPs (Fig. S1). Other glucose molecules, however, are not fully metabolized. Instead, they serve as precursor metabolites for the synthesis of building blocks. In fact, all building blocks, coenzymes, and prosthetic groups are synthesized from just 12 precursor metabolites, which serve as links between fueling and biosynthesis reactions (21). The diversion of these precursor metabolites away from energy-producing pathways toward the synthesis of building blocks (Fig. S1) results in an "opportunity cost." We define the opportunity cost of a precursor metabolite as the number of ATP molecules that could have been generated had the precursor metabolite not been diverted away from energy-producing pathways (Fig. S1, step 1). Fig. S2*A* shows the placement of precursor metabolites along metabolic pathways, and summarizes the net energetic gains and losses of converting one precursor metabolite to another. Fig. S2*B* builds upon Fig. S2*A* and the detailed estimates provided in SI sections III and IV to delineate the opportunity costs of precursor metabolites in heterotrophic bacteria and eukaryotes.

In the synthesis of building blocks from precursor metabolites, there is an additional source of opportunity cost, namely, the oxidation of electron carrier molecules (Fig. S1, steps 2). Had these electron carriers been preserved for energy production rather than being oxidized during biosynthesis pathways, 2 ATP molecules could have been generated from each NAD(P)H molecule using the bacterial ETC (ref. 21, pp. 158–159) and 2.5 ATP molecules from the eukaryotic ETC (45). The combination of the opportunity cost of precursor metabolites and the opportunity cost of building-block synthesis from precursor metabolites will be referred to as the opportunity cost of a building block, or E_O (Fig. S1, steps 1 and 2). To simplify, we will generally refer to the opportunity cost of a building block in our calculations.

In addition to an opportunity cost, a monomeric building block incurs a direct expenditure of ATP-equivalent molecules. To convert a precursor metabolite into a monomeric building block such as an amino acid, nucleotide, or lipid, the cell must pay a synthesis cost (Fig. S1, step 3). The cell must also pay in ATP-equivalent molecules to polymerize or assemble building blocks into genomes, proteins, and membranes (Fig. S1, step 4).

III. Opportunity Costs of Precursor Metabolites for Heterotrophic Bacteria

In this section, we will use the units of P, without any subscripts, to refer to ATP (and ATP-equivalent) hydrolysis events (using the subscripts without having first derived the opportunity cost of a metabolite is meaningless). Once we have derived the opportunity cost of a precursor metabolite here in this section, we will accompany it with the symbol P_O in later figures and tables to clearly mark these costs as opportunity costs. We will estimate the opportunity cost of precursor metabolites, $C_{opportunity}$, by the following:

$$C_{opportunity} = E_{glu} - E_{partial}, \quad \textbf{[S1]}$$

where E_{glu} represents the net energetic gain from the complete metabolism of a glucose molecule into water and carbon dioxide,

and $E_{partial}$ is the net energetic gain from the partial metabolism of a glucose molecule into a precursor metabolite. Under aerobic respiration, E_{glu} is ≈ 26 P in *E. coli* (24). In the event that there is a net energetic cost from the conversion of glucose into a precursor metabolite, $E_{partial}$ will be a negative value.

In the synthesis of lipids, dihydroxyacetone-phosphate (dhap) is used as a precursor. This precursor is generated during glycolysis (Fig. S2*A*). Each glucose molecule results in the production of two dihydroxyacetone-phosphate molecules, with this process having a net energetic cost of 2 P ($E_{partial} = -2$) (Fig. S2*A*). The opportunity cost of two dihydroxyacetone-phosphate molecules is 2 P greater than that of glucose, or 28 P (Eq. **S1**). The opportunity cost of one dihydroxyacetone-phosphate molecule is therefore ≈ 14 P (Fig. S2*B*).

To simplify the opportunity cost estimates further, we could use a shortcut. Rather than estimating the opportunity cost of each precursor metabolite by calculating the $E_{partial}$ from glucose as the starting point, we could obtain the opportunity cost of metabolite *j*, $C^{j}_{opportunity}$, from the opportunity cost of metabolite *i*, $C^{i}_{opportunity}$, by the following:

$$C^{j}_{opportunity} = C^{i}_{opportunity} - E^{i\rightarrow j}_{partial}, \qquad \textbf{[S2]}$$

where $E^{i\rightarrow j}_{partial}$ represents the net energetic gain from the conversion of metabolite *i* to metabolite *j*.

For example, during glycolysis each molecule of dihydroxyacetone-phosphate is converted to a molecule of 3-phosphoglycerate (3pg), resulting in the production of 1 NADH molecule and 1 ATP. In *E. coli*, each NADH molecule results in the production of ≈ 2 ATP molecules under aerobic conditions (ref. 21, pp. 158–159); therefore, the net energetic gain from this conversion is ≈ 3 P. Hence, the opportunity cost of each 3-phosphoglycerate molecule would be ≈ 11 P, which is 3 P less than the opportunity cost of a dihydroxyacetone-phosphate molecule (Eq. **S2**) (Fig. S2*B*).

If not used as a precursor, 3-phosphoglycerate is converted to phosphoenolpyruvate (pep) in glycolysis (Fig. S2*A*). In this process, however, there is zero energy expenditure or gain. Hence, the opportunity cost of a phosphoenolpyruvate molecule is the same as that of a 3-phosphoglycerate molecule, which is ≈ 11 P (Fig. S2*B*). Both of these precursors come before pyruvate (pyr) in glycolysis. In converting a phosphoenolpyruvate molecule into a pyruvate molecule, there is a net energy gain of 1 P (Fig. S2*A*). The opportunity cost of a pyruvate molecule is therefore ≈ 10 P (1 P less than a phosphoenolpyruvate opportunity cost) (Fig. S2*B*). Pyruvate can be converted to oxaloacetate (oaa) with the expenditure of 1 ATP. The opportunity cost of oxaloacetate is ≈ 11 P (10 P + 1 P) (Fig. S2*B*).

Pyruvate is further converted to acetyl-CoA (acCoA), and in the process one molecule of NADH is generated, which is equivalent to 2 P (Fig. S2*A*). The opportunity cost of acetyl-CoA is therefore ≈ 8 P (10 P − 2 P) (Fig. S2*B*). One molecule of acetyl-CoA and one molecule of oxaloacetate are then eventually converted to α-ketoglutarate (αkg) (Fig. S2*A*). The sum of the opportunity costs of acetyl-CoA (8 P) and oxaloacetate (11 P) is 19 P, and because 1 molecule of NADH is generated in their conversion to α-ketoglutarate, the opportunity cost of α-ketoglutarate is ≈ 17 P (or 19 P – 2 P) (Fig. S2*B*). Similarly, α-ketoglutarate is eventually converted to oxaloacetate, and 2 NADH, 1 GTP, and 1 $FADH_2$ molecules are generated (Fig. S2*A*). This is a net gain of ≈ 6 (assuming 1 P from each $FADH_2$), reducing the opportunity cost of oxaloacetate to ≈ 11 P (17 P – 6 P). Note that this is consistent with the opportunity cost of oxaloacetate derived from the anaplerotic pathway described earlier (the conversion of pyruvate to oxaloacetate via the pyruvate decarboxylase enzyme).

Glucose can also be converted to ribose-5-phosphate (r5p) in the pentose phosphate pathway, and in the process 2 NADPH molecules are generated, which is equivalent to 4 P (Fig. S2*A*). We subtract 4 P from the possible 26 P that glucose would be converted to under respiratory conditions, and we arrive at 22 P as the opportunity cost of ribose-5-phosphate (Eq. **S1**) (Fig. S2*B*). The same calculation can be used for erythrose-4-phosphate (e4p), resulting in 22 P as its opportunity cost (Fig. S2*B*).

IV. Opportunity Costs of Precursor Metabolites for Heterotrophic Eukaryotes

To estimate the opportunity costs of precursor metabolites in heterotrophic eukaryotes, we can carry out very similar calculations to those performed for heterotrophic bacteria (Fig. 1 *B* and *C*). Similar to SI section III, we will refrain from using subscripts for P in this section. For eukaryotes, E_{glu}, or the total energetic gain from the complete metabolism of a glucose molecule into carbon dioxide and water is higher. This is because each NAD(P)H and $FADH_2$ molecule results in a higher number of ATPs within the mitochondrial ETC compared with the bacterial ETC. Specifically, each NAD(P)H molecule is equivalent to ≈ 2.5 P and each $FADH_2$ molecule corresponds to ≈ 1.5 P (ref. 45 and ref. 46, pp. 517–518), resulting in 30–32 P per glucose molecule. Note that the theoretical yield of 38 P per glucose molecule has been shown to be an overestimate due to the outdated assumptions that each NAD(P)H molecule is equivalent to 3 P and that each $FADH_2$ molecule generates 2 P (45). We will therefore use 32 P as E_{glu}.

Each glucose molecule results in the production of two dihydroxyacetone-phosphate molecules, with this process having a net energetic cost of 2 P. The opportunity cost of two dihydroxyacetone-phosphate molecules is 2 P greater than that of glucose, or 34 P (Eq. **S1**). The opportunity cost of one dihydroxyacetone-phosphate molecule is therefore 17 P (Fig. S2*B*). As described earlier, in the conversion of dihydroxyacetone-phosphate molecule into a 3-phosphoglycerate molecule, 1 NADH and 1 ATP molecules are produced. This is equivalent to a net energetic gain of ≈ 3.5 P. The opportunity cost of a 3-phosphoglycerate molecule is therefore ≈ 13.5 P (17 P – 3.5 P) (Eq. **S2**) (Fig. S2*B*).

The opportunity cost of a phosphoenolpyruvate is the same as that of a 3-phosphoglycerate, which is ≈ 13.5 P. In converting phosphoenolpyruvate into pyruvate, there is a net energy gain of 1 P (1 ATP molecule is formed). The opportunity cost of pyruvate is therefore ≈ 12.5 P (Fig. S2*B*). Pyruvate can be converted to oxaloacetate with the expenditure of 1 ATP. Hence, the opportunity cost of oxaloacetate will be ≈ 13.5 P (Fig. S2*B*).

Pyruvate is converted to acetyl-CoA, and in the process one molecule of NADH is generated. As a result, the opportunity cost of acetyl-CoA is ≈ 10 P (12.5 P – 2.5 P). One molecule of acetyl-CoA and one molecule of oxaloacetate are converted to α-ketoglutarate in the TCA cycle. The sum of the opportunity costs of acetyl-CoA (10 P) and oxaloacetate (13.5 P) is 23.5 P, and because 1 molecule of NADH is generated in their conversion to α-ketoglutarate, the opportunity cost of α-ketoglutarate is ≈ 21 P (23.5 P – 2.5 P) (Fig. S2*B*). α-Ketoglutarate is eventually converted to oxaloacetate, and 2 NADH, 1 GTP, and 1 $FADH_2$ molecules are generated. This is a net gain of ≈ 7.5 P (assuming 1.5 P from each $FADH_2$), reducing the opportunity cost of oxaloacetate to ≈ 13.5 P (21 P – 7.5 P). Note that this is again consistent with the opportunity cost of oxaloacetate derived from the anaplerotic pathway.

Glucose can also be converted to ribose-5-phosphate in the pentose phosphate pathway, and in the process 2 NADPH molecules are generated, which is equivalent to 5 P. The opportunity cost of ribose-5-phosphate is thus ≈ 27 P (Fig. S2*B*). The same calculation can be used for erythrose-4-phosphate, resulting in ≈ 27 P as its opportunity cost (Fig. S2*B*).

V. Viral Entry Cost

The process of viral entry varies extensively across different viral groups. Although many animal viruses enter their host cell through clathrin-mediated endocytosis or fusion with the cell membrane

(47), most phages inject their genetic material with the capsid remaining outside of the cell. In both cases, however, entry is mediated by the interaction between viral entry proteins and host receptors (47).

For T4, it is the interactions between a minimum of three long tail fibers and cellular receptors that initiates a cascade of conformational changes (48) (Fig. 1, step 1). After this preliminary interaction, the base plate is subsequently brought closer to the cell membrane, allowing the short tail fibers to interact with their host receptors. The tail sheath contracts, resulting in the tail tube puncturing the outer cell membrane (Fig. 1, step 2). Then, conformational changes in gp5 phage protein activate its lysozyme domains, resulting in the digestion of the peptidoglycan layer (Fig. 1, step 2). Effectively, the tail tube passes through the periplasm. At this point, the phage DNA is passed through the inner membrane via the tail tube and is then exposed to the intracellular environment (Fig. 1, step 2). In general, viral entry proceeds through protein conformational changes and does not rely on ATP expenditure (49). In the case of the T5 phage, this point has been explicitly demonstrated (50).

The influenza virus is composed of a capsid that is enveloped by a lipid membrane (Fig. 2). Inside the capsid reside the ribonucleoprotein complexes, which are composed of the segmented viral genome encapsidated by proteins. The viral membrane is decorated with hemagglutinin (HA) proteins, which bind to the host sialic acid receptors, thereby initiating clathrin-mediated endocytosis (Fig. 2, step 1). During endocytosis, a self-assembled protein cage composed of clathrin triskelions forms around the inward budding vesicle. Once the clathrin cage has formed, dynamins perform the last stage of endocytosis (Fig. 2, step 2).

Dynamin is a mechanochemical GTPase that self-assembles into multimeric spirals at the necks of clathrin-coated endocytic vesicles to catalyze membrane fission. The dynamin helix wrapped around the neck of an endocytic vesicle forms a protein–lipid tube with an inner diameter of 20 nm. In the presence of GTP, the dynamin helix undergoes structural changes that result in the reduction of the inner diameter (51). Because the dynamin helix is composed of at least two turns (37), each turn is composed of 13 dynamins, and each dynamin consumes 1 GTP, the energy requirement for vesicle fission can be approximated as 30–100 P_D (Fig. 2, step 2).

Once the vesicle is released from the cell membrane, the clathrin cage has to be disassembled. This process requires the expenditure of 3 ATPs per triskelion (52). For a clathrin cage composed of 36 triskelions, this is equivalent to about 100 P_D (Fig. 2, step 3). As the endosome matures, the endosomal lumen becomes more acidic. The influenza virus has exploited this feature for the uncoating of its lipid membrane as well as the disassembly of its capsid. The endosomal pH drop activates the viral transmembrane proton channel, M2. The influx of H^+ ions from the endosome through M2 leads to the dissociation of the viral capsid proteins. The acidic environment also triggers conformational changes that expose the viral HA2 fusion peptide, which subsequently fuses the viral and endosomal membranes (53). These two events together enable the release of the viral ribonucleoprotein complexes into the host's cytoplasm (54). To summarize, the cost of entry for influenza per virion, E_{Entry}, falls within the range of 10^2 to 10^3 P_D. It is important to note that, in our estimates, we will generally not include the cost of host protein production, unless a protein is exclusively produced for viral synthesis. For example, the cost of producing the clathrin cage and dynamin proteins are not included due to the fact that these proteins are recycled and produced for the host's own functions.

T4 efficiently exposes its genomic material to the host cytoplasm where it can be readily transcribed and translated. The influenza virus, however, due to the much larger volume and extensive compartmentalization of its eukaryotic host, faces additional obstacles in the way of reaching sites of transcription and translation. We will describe these obstacles in the next section.

VI. Intracellular Transport Cost

Replication and transcription of the influenza virus occur inside the nucleus. Like other cargo destined toward the nucleus from the plasma membrane, the endosome carrying the influenza virus is transported via the action of the dynein motor proteins along microtubule tracks (Fig. 2, step 4). Unlike the kinesin motor protein, dynein is one that takes variable step sizes along the microtubule. As a result of having a hexameric ring of AAA+ ATPases (similar to Vps4, described later in SI section XI), dynein has multiple ATP binding sites. For the purposes of our estimates, we assume a step size of 8 nm and the expenditure of 1 ATP per step (55). If we assume that the nucleus resides roughly in the center of a cell $\approx$ 10 μm in diameter (18), this will require $\approx 10^3$ dynein steps in carrying the endosome. As a result, the cost of transport for the viral genome from the vicinity of the cell membrane to the nucleus is $\approx 10^3$ P_D (Fig. 2, step 4).

The ribonucleoproteins are imported to and exported from the nucleus via nuclear localization signals (Fig. 2, step 5). The nuclear import of influenza ribonucleoproteins has been experimentally demonstrated to be an energy-consuming process (56). To get access to the nucleus, these ribonucleoproteins have to pass through the nuclear pore complex. They do so by binding to importins using nuclear localization signals. Each cargo imported into the nucleus will require the hydrolysis of one GTP. Whether ribonucleoproteins travel in and out of the nucleus separately or together is not definitely known, but there is recent evidence suggesting that the eight ribonucleoproteins travel as one cargo (57). Hence, we estimate that the import of the viral genome will cost 1 P_D per virus.

Once inside the nucleus, the influenza genome is transcribed and replicated, the costs of which will be discussed in later sections. The production of the full ribonucleoprotein complex is a convoluted process, requiring the transcription of the viral genome (Fig. 2, step 6), the nuclear export of the resulting transcripts (Fig. 2, step 5), as well as the translation of viral mRNA transcripts (Fig. 2, steps 7). Once the viral proteins are generated, some proteins travel back into the nucleus to encapsidate the viral −ssRNA genome and form the next generation of ribonucleoprotein complexes (PB1, PB2, NP, and PA proteins) (Fig. 2, steps 5 and 8). Other viral proteins are transported to the cell membrane, and together with ribonucleoprotein complexes, which are destined to the same site, bud off.

Similar to the nuclear entrance, cargo exiting the nucleus must also pay a price. The nuclear export cost through the CRM1 export pathway, which is the one used by influenza, incurs 1 P_D per cargo, similar to the import pathway (57). Considering that on average 6,000 influenza genomes have to leave the nucleus to eventually give rise to 6,000 influenza virions, we estimate the export of ribonucleoproteins from the nucleus will cost $\approx 6 \times 10^3$ P_D. The mechanism and energetics of viral mRNA export from the nucleus are not well understood (58), although there is a growing body of evidence implicating the involvement of the nuclear RNA export factor 1. We suspect the cost of mRNA nuclear export is on par with the cost of ribonucleoprotein nuclear export and estimate the cost of nuclear import and export to be $\approx 10^4$ P_D.

Upon exiting the nucleus, ribonucleoproteins have to be transported toward the cell membrane where they can coassemble with other viral proteins and bud off. Their path to the cytoplasm starts at the microtubule organizing center and proceeds via recycling endosomes bound to the RAB-11 GTPase (57). This endocytic transport is powered by kinesin motor proteins trekking along microtubules with a step size of 8 nm and the hydrolysis of 1 ATP per step (59) (Fig. 2, step 9). Assuming the eight ribonucleoproteins travel as one cargo—consistent with the reported typical size of an endosome (200 nm in diameter) (60)—a 5-μm

transport of 6,000 ribonucleoproteins from the nucleus to the cytoplasm costs $\approx 3 \times 10^6$ P_D. Under a similar assumption (i.e., each endosome carries the proteins required to build a single virus), the cost of protein transport to the apical cell surface is similar to the cost of ribonucleoprotein transport. Considering the sum of the costs from nuclear import and export as well as travel along the microtubules, the cost of intracellular transport for an influenza infection, $E_{Transit/i}$, can be approximated as 6×10^6 P_D.

The cost of GTPases mediating endosome docking and fusion is likely negligible compared with the heavy cost of motor proteins mobilizing the endosomes. This is because, during endosome trafficking, there are far more steps taken by motor proteins than there are endosome fusion or docking events. However, each fusion or docking event incurs a similar cost to a step taken by a motor protein. More detailed estimates of endosomal trafficking are exceedingly difficult due to the unavailability of studies that shed light on (*i*) the average number of proteins carried per endosome, (*ii*) the average number of GTPases needed per endosome, and (*iii*) the average length that an endosome travels within the cell, among other topics.

VII. Genome Replication Cost

The Direct Cost. Genome replication is a complex phenomenon requiring many different steps, such as the unwinding of the parent helix, RNA primer synthesis, Okazaki fragment ligation, and proofreading. Although there are many facets of genome replication in cells that require energy expenditure, the direct cost of replication lies primarily in the direct synthesis cost of nucleotides from precursor metabolites as well as the polymerization of individual nucleotides. The energetic cost of genome replication for a virus with a dsDNA genome can be approximated as follows:

$$E_{REP(dsDNA)/v} \approx 2L_g(e_d + e_p). \quad \text{[S3]}$$

Here, L_g corresponds to the genome length and is multiplied by 2 to account for T4's double-stranded genome. The cost of each DNA nucleotide can be stated as the sum of e_d, which represents the average direct cost of DNA synthesis from precursor metabolites, and e_p, which denotes the cost of chain elongation per base [equivalent to 2 P_D (20)] (Fig. 1, step 5). The energetic cost of replicating a −ssRNA genome is similarly as follows:

$$E_{REP(-ssRNA)/v} \approx 2L_g(e_r + e_p), \quad \text{[S4]}$$

where e_r represents the average direct cost of RNA synthesis from precursor metabolites. The factor of 2 stems from the fact that a −ssRNA has to first be converted to +ssRNA before a second copy of −ssRNA can be synthesized (Fig. 2, step 8). e_d in the context of bacterial metabolism is equivalent to 11 P_D (Fig. S3*A* and Dataset S1). The average synthesis cost of an RNA base from a precursor metabolite in the context of eukaryotic metabolism is 10 P_D (Fig. S3*B* and Dataset S1). With these values in hand, we estimate the direct cost of T4 ($L_g \approx 2 \times 10^5$) and influenza ($L_g \approx 1 \times 10^4$) genome replication to be $\approx 4 \times 10^6$ P_D and $\approx 3 \times 10^5$ P_D, respectively.

The Opportunity Cost. The opportunity cost of T4 phage replication can be estimated as $2L_g e_{od}$, where e_{od} corresponds to the average opportunity cost of a DNA base synthesized in bacteria and is ≈ 34 P_O (Fig. S3*A* and Dataset S1). Note that e_{od} represents the sum of the average opportunity cost of precursor metabolites required for DNA synthesis (33 P_O) and the average opportunity cost of DNA synthesis from precursor metabolites (1 P_O). Under this estimate, the opportunity cost of genome replication for a T4 phage is $\approx 1 \times 10^7$ P_O. Similarly, the opportunity cost of genome replication for an influenza virus can be estimated as $2L_g e_{or}$, where e_{or} refers to the average opportunity cost of an RNA nucleotide synthesized in a eukaryotic organism, which is ≈ 39 P_O (Fig. S3*B* and Dataset S1). Hence, the opportunity cost of influenza genome replication is $\approx 8 \times 10^5$ P_O.

The Total Cost. The total cost of T4 phage replication is the sum of opportunity and direct costs: $2L_g(e_d + e_p + e_{od})$. Under this estimate, the cost of genome replication for a T4 phage is $\approx 2 \times 10^7$ P_T. The same logic follows for influenza, where the total cost of its genome replication can be estimated as $2L_g(e_r + e_p + e_{or})$. Hence, the total cost of influenza genome replication is $\approx 1 \times 10^6$ P_T.

The Replication Cost Component of an Infection. Furthermore, the replication cost per infection is the cost of replication for a virus multiplied by its burst size, B, resulting in the following:

$$E_{REP/i} \approx BE_{REP/v}. \quad \text{[S5]}$$

The direct, opportunity, and total costs of genome replication during a T4 infection with an average burst size of 200 are $\approx 9 \times 10^8$ P_D, $\approx 2 \times 10^9$ P_O, and $\approx 3 \times 10^9$ P_T (Fig. 1, step 5). The direct, opportunity, and total costs of genome replication during an influenza infection with an average burst size of 6,000 are $\approx 2 \times 10^9$ P_D, $\approx 5 \times 10^9$ P_O, and $\approx 6 \times 10^9$ P_T (Fig. 2, step 8).

VIII. Transcriptional Cost

As in the case of replication, transcription involves various energy-consuming processes, such as transcriptional activation, initiation and termination, as well as proofreading and mRNA splicing. However, the dominant cost in transcription is similar to that in replication, namely, the investment in the synthesis of the nucleotides themselves (20). See Fig. 1 (step 3) and Fig. 2 (step 6).

The Direct Cost. The direct cost of transcription is approximately the following:

$$E_{TX/v} \approx (e_{er} + e_r) \sum_{i=1}^{m} L_{ri} N_{ri}, \quad \text{[S6]}$$

where L_{ri} and N_{ri} represent the length and the copy number of each viral mRNA transcript, respectively. The average direct cost of synthesizing an RNA base from precursor metabolites is denoted as e_r, which is 10 P_D (Fig. S3 and Dataset S1). The direct cost of polymerizing an RNA base is symbolized by e_{er}, which is equal to 2 P_D (20). Because the mRNA copy number for each viral gene is largely an unknown parameter, we approximated the viral genome as one long gene. This allowed us to eliminate the L_r and N_r dependence, and replace them with constants, l_r and n_r, such that

$$E_{TX/v} \approx (e_{er} + e_r) l_r n_r. \quad \text{[S7]}$$

The constant l_r corresponds to the length of the mRNA transcript, and thus is equal to the length of the genome. The average copy number of this transcript, n_r, can be approximated by the observed ratio of 1 mRNA transcript per 1,000 resulting proteins both in prokaryotes (18) as well as in mammalian cells (61). The average protein copy number of a virus, n_p, can be related to its average transcript number according to $n_r \approx n_p/1{,}000$.

To obtain n_p, we have used available data on viral protein copy numbers. For a T4 phage, with a total estimated protein count of 5,000 representing 35 structural and 3 lysis genes (Table S1), the average protein copy number per virus is ≈ 130. Influenza's total protein count is also $\approx 5{,}000$, representing the products of nine proteins (Table S2). This results in an estimate for the average protein copy number for an influenza virus of ≈ 550. With

these numbers in hand, we estimate n_r to be ≈ 0.13 for a single T4 phage and 0.55 for an influenza virus. We note that, generally, mRNAs are relatively short-lived and each individual mRNA on average produces between 10 and 100 proteins. However, the pool of such proteins is then used to synthesize all of the virus particles that make up a given burst.

Although it is sufficient to consider only the synthesis and polymerization costs of nucleotides in the direct cost estimates for genome replication, the transcriptional cost should additionally encompass the cost of mRNA repolymerization (20). This is because mRNA transcripts have a short life span and must be regenerated throughout the duration of the infection. This cost is perhaps more prominent in bacteria where the lifetime of a transcript is about 3 min (18), whereas the lifetime of a mammalian cellular transcript is $\approx$ 10 h (18), and therefore comparable to the lifetime of the influenza infection itself. The first step in calculating the mRNA repolymerization cost is to multiply the lifetime of the infection, t, by the mRNA degradation rate, δ_r.

The second step is to take into account the average transcript copy number and its length to determine the number of RNA bases that have to be repolymerized during the course of an infection. Effectively, the assumption is that the RNA nucleotides are only being repolymerized, not resynthesized. The repolymerization cost of transcription can be stated as the number of nucleotides to be repolymerized, $l_r n_r \delta_r t$, multiplied by e_{er}. The direct cost of transcription per virus can then be revised as such:

$$E_{TX/v} \approx l_r n_r (e_r + e_{er}\delta_r t). \quad \text{[S8]}$$

The lifetime of a T4 infection is $\approx$ 30 min (62), and the lifetime of an influenza infection is roughly 12 h (14). With these parameters in hand, the direct cost of transcription (per virus) for T4 and influenza are $\approx 7 \times 10^5$ P_D and $\approx 7 \times 10^4$ P_D, respectively (Table 1).

The Opportunity Cost. The opportunity cost of transcription can be obtained by $l_r n_r e_{or}$, where e_{or} represents the opportunity cost of an RNA nucleotide. e_{or} is $\approx$ 31 P_O in bacteria and 39 P_O in eukaryotes (Dataset S1 and Fig. S3). Note that, in this expression, we do not account for repolymerization events as RNA nucleotides are assumed to be recycled rather than resynthesized. The opportunity costs of transcription for a single T4 phage and an influenza virus are thus $\approx 7 \times 10^5$ P_O and $\approx 2 \times 10^5$ P_O, respectively.

The Total Cost. The total cost of transcription can be obtained by $l_r n_r (e_{or} + e_r + e_{er}\delta_r t)$, which represents the sum of opportunity and direct costs of transcription. According to this estimate, the total costs of transcription for a single T4 phage and an influenza virus are $\approx 1 \times 10^6$ P_T and $\approx 3 \times 10^5$ P_T, respectively (Table 1).

The Transcriptional Cost of an Infection. The transcriptional cost of an infection is the transcriptional cost of a virus multiplied by its burst size, namely, the following:

$$E_{TX/i} \approx BE_{TX/v}. \quad \text{[S9]}$$

For T4 with an average burst size of 200 and for influenza with an average burst size of 6,000, the direct cost of transcription at the level of an infection is $\approx 1 \times 10^8$ P_D (Fig. 1, step 3) and $\approx 4 \times 10^8$ P_D, respectively (Fig. 2, step 6). The opportunity cost of transcription for these two infections are $\approx 1 \times 10^8$ P_O (T4) and $\approx 1 \times 10^9$ P_O (influenza).

Their total costs are $\approx 3 \times 10^8$ P_T (T4, Fig. 1, step 3) and $\approx 2 \times 10^9$ P_T (influenza, Fig. 2, step 6) (Table 1).

IX. Translational Cost

There are important biosynthetic costs associated with proteins just as there are with nucleic acids. Here, we attempt to capture the most significant costs in the protein synthesis pathway while making some simplifying assumptions that neglect substantially smaller cost components such as the costs of translational initiation and termination and posttranslational modifications (20). The estimate for translational cost follows the same rationale as the cost calculation for transcription.

The Direct Cost. The direct cost of translating the viral proteome can be estimated as follows:

$$E_{TL/v} \approx (e_a + e_{ea}) \sum_{j=1}^{k} L_{pj} N_{pj}. \quad \text{[S10]}$$

Here, we have multiplied the total number of amino acids by the per-amino acid costs of synthesis from precursor metabolites, e_a, and polymerization, e_{ea}. The arrays L_p and N_p hold values for the length of each protein and its copy number per virus, respectively (Tables S1 and S2). We show that $\sum_{j=1}^{k} L_{pj} N_{pj}$ for the influenza virus and the T4 phage are about 1.7 and 1.2 million amino acids, respectively (Tables S1 and S2). In both bacteria and eukaryotes, e_a, is on average equal to 2 P_D (Fig. S3 and Dataset S1) and e_{ea} is 4 P_D (25). Due to the relatively slow protein degradation rates for bacteria (0.4/h) and human cells (0.08/h) (20, 26) compared with infection lifetimes, we have neglected costs associated with this process. Using this information, the direct cost of translation for a T4 phage and an influenza virus are $\approx 7 \times 10^6$ P_D and $\approx 1 \times 10^7$ P_D, respectively (Table 1). This finding stems mainly from the fact that both viruses are based on $\approx 10^6$ aa.

The Opportunity Cost. The opportunity cost of viral translation is approximately $e_{oa} \sum_{j=1}^{k} L_{pj} N_{pj}$, where e_{oa} denotes the average opportunity cost of an amino acid and corresponds to 25 P_O in bacteria and 30 P_O in eukaryotes (Fig. S3 and Dataset S1). The opportunity cost of viral translation for a T4 phage and an influenza virus are therefore $\approx 3 \times 10^7$ P_O and $\approx 5 \times 10^7$ P_O, respectively.

The Total Cost. The total cost of viral translation is the sum of direct and opportunity costs of translation. The total cost of viral translation for a T4 phage and an influenza virus are therefore $\approx 4 \times 10^7$ P_T and $\approx 6 \times 10^7$ P_T, respectively (Table 1).

The Translational Cost Component of an Infection. The translational cost of an infection is simply the cost of translation per virion multiplied by its burst size, namely, the following:

$$E_{TL/i} \approx BE_{TL/v}. \quad \text{[S11]}$$

The direct, opportunity, and total translational costs for a T4 infection with a burst size of 200 are $\approx 1 \times 10^9$ P_D, 6×10^9 P_O, and 8×10^9 P_T (Fig. 1, step 4). For an influenza infection with a burst size of 6,000, these costs are $\approx 6 \times 10^{10}$ P_D, 3×10^{11} P_O, and 4×10^{11} P_T (Fig. 2, step 7) (Table 1).

Protein Folding and Quality Control. Just as in any other biological process, protein folding is subject to errors. To correct for such errors and prevent the aggregation of misfolded proteins, cells from all three domains of life have evolved elaborate mechanisms for the detection of misfolded proteins. Through various ATP-dependent (e.g., Hsp90, Hsp70, and Hsp60) and ATP-independent processes, a triage is carried out in which some proteins are refolded and others are degraded (63–65). From an energetic standpoint, protein quality control mechanisms are likely to cost substantially less than the cost of translation. As shown above, a protein with an

average length of 300 aa (26) will have a direct translational cost of 1,800 P_D and a total cost of 9,300 P_T in bacteria and 10,800 P_T in eukaryotes, respectively. On the other hand, the energetic cost of various protein quality control pathways can range from a few ATPs (66) to a few hundred ATPs per protein (67), which is required for protein degradation. Thus, it is likely that protein quality control will be a fractional cost compared with the translational cost of a protein. A similar conclusion was drawn in the context of cellular protein cost (20). Because we were unable to ascertain the fraction of viral proteins that may be degraded, and because different proteins require different quality control pathways (68), any more detailed estimates are difficult to make. Future experimental studies would be needed to determine any substantial costs incurred by protein quality control or translation at large that may be missing from our estimates.

Another very interesting experimental avenue would be to explore the consequences of the quality control mechanisms as they are being partially recruited toward maintaining the viral proteome. Particularly in the context of a host cell that survives the infection, how the cell responds to the additional burden from viral protein production and maintenance would be a fascinating topic of study. Constructed from roughly 5,000 proteins (Table S2), and with an average burst size of about 6,000, the influenza infection will produce about 3×10^7 viral proteins. Considering that a human cell will harbor more than 10^9 proteins (69), we would expect the extra load from viral proteins on the quality control machinery to be minimal. For a T4 phage, composed of roughly 5,000 proteins (Table S1) and with a burst size of roughly 200, the total number of viral proteins during an infection would be $\approx 10^6$. This is comparable to the number of bacterial proteins, which is estimated to be 10^6 per cubic micrometer (70), or the approximate volume of an *E. coli* cell. Hence, in the case of a T4 infection, the viral quality control pathways are likely to more heavily affected than in the case of an influenza infection.

X. Assembly and Genome Packaging Cost

Upon translation, the influenza viral proteins and ribonucleoproteins travel toward the cell membrane via Rab-bound endosomes that are carried by kinesins on microtubules. Interactions of the matrix protein M1 with ribonucleoproteins and the viral transmembrane proteins, namely HA, NA, and M2, result in the assembly of the influenza virus (71). Although the kinetics of the assembly steps remains to be delineated, influenza virus assembly and genome packaging are not regarded as energy-consuming processes. In general, virus assembly is described as an energetically favorable process, typically driven by the burial of hydrophobic surfaces (72, 73), and therefore independent of host energy expenditure. As an example, the assembly of hepatitis B virus is shown to occur spontaneously through weak protein–protein interactions (74).

Although the assembly of the T4 capsid is spontaneous, the packaging of the genome inside the capsid is not (Fig. 1, step 7). The cost of genome packaging for a T4 phage is as follows:

$$E_{Pack/v} = e_p L_g, \qquad \textbf{[S12]}$$

where the cost to package a base pair, e_p, is 2 P_D (75). For the 169-kb genome of T4, this cost is as follows: $\approx 3 \times 10^5$ P_D.

The packaging cost of a T4 infection is simply the cost of packaging for a single T4 phage multiplied by the T4 burst size:

$$E_{Pack/i} = BE_{Pack/v}. \qquad \textbf{[S13]}$$

For T4, with a burst size of 200, the contribution of packaging to the total cost of infection is $\approx 7 \times 10^7$ P_D (Fig. 1, step 7) (Table 1).

XI. Viral Exit Cost

Viruses use two primary exit strategies. Generally, enveloped viruses such as influenza and HIV bud off from the host membrane. Phages, on the other hand, generally lyse their host cells. For T4, the cost of exit is primarily the production cost of proteins that together break down the cell wall. We have already included the cost of lysis proteins in our translational cost estimates. The lysis proteins include holin, endolysin, and spanin proteins. The holins create holes in the host inner membrane, enabling the endolysins to reach the peptidoglycan layer. The spanins fuse the inner and the outer membrane as a requirement for lysis of Gram-negative bacteria (Fig. 1, step 9, only holins and endolysins are shown). Considering that T4 holin, endolysin, and spanin proteins are 218, 164, and 216 aa in length, respectively, and each has a copy number of about 4,000 per infection (or ≈ 20 per virus considering a burst size of 200) (76), the contribution of lysis proteins to the translational cost of an infection is negligible.

In the case of influenza, the exocytosis of virions is energy-consuming. However, the exact mechanism remains a mystery, with influenza M2 protein so far serving as the most likely agent for mediating exocytosis (77). Three separate costs exist: (*i*) the cost to locally bend the membrane outward, (*ii*) the cost to scisse the budding virion from the cell membrane, and (*iii*) the cost of the cellular membrane that becomes part of the viral membrane (Fig. 2, step 10). The cost to bend the membrane into the shape of a sphere of any size, is equivalent to 25 P_D (25). One way to estimate the cost of scission is to assume that it incurs a comparable cost to the HIV scission process, which, similar to several other enveloped viruses, is mediated through the ESCRT-III (endosomal sorting complexes required for transport) assembly. ESCRT-III complex self-assembles into filaments around the neck of a budding vesicle (similar to dynamin), and its disassembly requires the expenditure of 6 P_D via the Vps4 ATPase (78). The sum of these two costs results in the use of 31 P_D as one influenza virus leaves the cell. Because 6,000 virions exit the cell on average, exocytosis costs $\approx 2 \times 10^5$ P_D.

An alternative, order-of-magnitude estimate could be made with the assumption that the cost of membrane scission during endocytosis equals the cost of membrane scission during exocytosis. In estimating the influenza entry cost, we showed that the cost of membrane scission during influenza endocytosis via the dynamin polymer is ≈ 30 P_D. Together with the cost of membrane bending, ≈ 25 P_D, this exit estimate is slightly higher (55 P_D per virus) than the previous (31 P_D per virus). Under this estimate, the cost of exocytosis for all 6,000 influenza virions amounts to $\approx 3 \times 10^5$ P_D.

The primary cost of viral exit for influenza, however, is the cost of lipids that are taken from the host cell to form the viral membrane. The cost of lipids per virus can be estimated by the number of lipid molecules needed per virion multiplied by the cost of a lipid molecule, e_l:

$$E_{Exit/v} = \frac{8\pi r^2}{s} e_l. \qquad \textbf{[S14]}$$

In Eq. **S14**, the numerator represents twice the viral surface area (accounting for the bilayer) and the denominator, s, denotes the surface area of a lipid head group, which is ≈ 0.5 nm^2 (26, 79). With an average radius of 50 nm, the influenza virus is composed of $\approx 1 \times 10^5$ lipid molecules. The direct, opportunity, and total costs of a lipid molecule in a eukaryotic cell are 18 P_D, 264 P_O, and 282 P_T, respectively (Dataset S1 and Fig. S3*B*). As a result, the direct, opportunity, and total costs of lipids per influenza virus are 2×10^6 P_D, 3×10^7 P_O, and 4×10^7 P_T, respectively. The exit costs of an influenza infection can be derived by the following:

$$E_{Exit/i} = BE_{Exit/v}, \quad \textbf{[S15]}$$

and are $\approx 1 \times 10^{10}$ P_D, 2×10^{11} P_O, and 2×10^{11} P_T for an infection with a burst size of 6,000 (Table 1).

XII. Estimating the Total Host Energy Budget from Growth Experiments in Chemostats

The total basal and growth metabolic requirements of various organisms have been shown to correlate with the cellular volume (20). We have used Eqs. **S16**–**S18** presented by Lynch and Marinov (20) to estimate the energetic budget of hosts considered in this study. Basal metabolic requirement of a cell scales with the cell volume according to the following:

$$E_M \approx 0.39V^{0.88}, \quad \textbf{[S16]}$$

where V, represents the cell volume in units of cubic micrometers, and E_M is in the units of 10^9 ATP per cell per hour. The growth metabolic requirement of a cell similarly scales with the cell volume according to the following:

$$E_G \approx 27V^{0.97}, \quad \textbf{[S17]}$$

where E_G is in the units of 10^9 ATP per cell. The total energy requirement of a cell is simply the sum of the basal and growth energy requirements,

$$E_T = E_G + tE_M, \quad \textbf{[S18]}$$

where t is the cell division time in the units of hours.

For a bacterial cell with a volume of 1 μm^3, the maintenance metabolic cost is $\approx 10^8$ P_T in the duration of a T4 phage infection that lasts about 30 min. A mammalian cell, on the other hand, with a characteristic volume of 2,000 μm^3, has a basal metabolic cost of $\approx 10^{12}$ P_T over the course of a 12-h influenza infection. The total energetic cost of a cell should also encompass the cost of cellular growth. Hence, the total energetic cost of a bacterium and a mammalian cell with the dimensions highlighted above are $\approx 3 \times 10^{10}$ P_T and $\approx 5 \times 10^{13}$ P_T, respectively, during the course of their viral infection. The correlation between cellular volume and metabolic capacity is supported by the observation that larger *E. coli* cells produce higher T4 burst sizes (80).

XIII. The Heat Production and Power Consumption of a Viral Infection

In our estimates for heat production and power consumption of a viral infection, we will not consider the total cost of an infection as it contains the opportunity costs; by definition, these opportunity costs do not represent direct use of ATP (and ATP-equivalent) molecules; rather, they represent the ATP (and ATP-equivalent) molecules that could have been generated in the absence of a viral infection (SI section II). For these estimates, we will rely on the direct costs.

To estimate the amount of heat that is generated due to a viral infection, we have to consider the inefficiency of aerobic metabolism. The burning of glucose results in the production of $\approx$ 2,800 kJ/mol of heat (81). The same reaction takes place inside our cells, with the difference being that cells are capable of harnessing a fraction of the free energy that would otherwise be liberated as heat. When glucose is aerobically metabolized into water and carbon dioxide, a fraction of the free energy is used to convert ADP into ATP, whereas the remaining free energy is dissipated as heat. By assuming physiological conditions, the free-energy change of ATP hydrolysis can be approximated as –50 kJ/mol (16). In bacterial metabolism, 26 ATPs are generated from each glucose molecule; hence the free energy captured by the conversion of ADP into ATP is approximately –1,300 kJ/mol of glucose. As a result, in this simple estimate, we consider that about 50% of the energy from the aerobic metabolism of glucose is dissipated as heat: $(1 - (1{,}300\ \text{kJ/mol}/2{,}800\ \text{kJ/mol})) \times 100\%$. For eukaryotic cells, with 32 ATPs generated per glucose molecule, about 40% of the energy stored in glucose is dissipated as heat.

The T4 infection has a direct cost of 3×10^9 P_D (Table 1). Because each glucose molecule results in 26 ATPs in the bacterial metabolism, T4 infection's direct cost would require the complete metabolism of 10^8 glucose molecules. The influenza infection's direct cost is about an order of magnitude higher than that of T4 (Table 1) and is equivalent to the aerobic metabolism of $\approx 10^9$ glucose molecules. Based on the number of glucose molecules required to cover the direct costs of each infection, the free energy stored in glucose (–2,800 kJ/mol), and the percentage of the energy released as heat during the aerobic metabolism of glucose ($\approx$ 40–50%), we can conclude that the heat generated during T4 and influenza infections are $\approx$ 0.2 and 2 nJ, respectively.

In half an hour, the T4 infection results in the hydrolysis of ATP-equivalent molecules at an average rate of 2×10^6 P_D per second. In half a day, an influenza infection also has an average ATP-hydrolysis rate of 2×10^6 P_D per second. Put in terms of the more familiar units of watts (by assuming –50 kJ/mol of free energy change per P_D), the power of both viral infections is on the order of 200 fW.

XIV. Generalizing Viral Energetics for dsDNA Phages

Fig. 5*A* shows how we can generalize the estimates presented here by thinking of dsDNA phages as approximately spherical objects with an outer layer of thickness, t. In this model, the inner volume of a virus containing the viral genome is given by $4\pi r^3/3$, which can be used to estimate the viral genome length (28). The total cost of genome replication for a dsDNA genome can be obtained from Eq. **S3** (Fig. 4). However, instead of using the viral genome length directly, we can divide the capsid inner volume by the volume of a base pair, v_d (≈ 1 nm^3) (26):

$$E_{REP(dsDNA)/v} \approx \frac{4\pi r^3}{3v_d}\left(e_d + e_p + e_{od}\right). \quad \textbf{[S19]}$$

Because for many dsDNA phages only about one-half of the capsid is filled with the viral genome (28), the cost of a DNA base was not multiplied by a factor of 2 (even though the genome is double stranded) as the two multipliers cancel each other out. The direct cost of replication can be obtained similarly by the exclusion of the opportunity cost of a DNA base or e_{od} (in bacteria) from Eq. **S19**. Moreover, the translational cost of a virus can be obtained from a modification of $(e_a + e_{ea} + e_{oa})\sum_{j=1}^{k} L_{pj}N_{pj}$, derived previously (SI section IX). The total cost of viral translation,

$$E_{TL/v} \approx \frac{4\pi\left(R^3 - r^3\right)}{3v_a}\left(e_a + e_{ea} + e_{oa}\right), \quad \textbf{[S20]}$$

can be obtained from multiplying the total number of amino acids by the total cost of an amino acid. The number of amino acids is estimated by dividing the outer capsid volume (denoted by the shaded blue region in Fig. 5*A*), $4\pi(R^3 - r^3)/3$, by the volume of an amino acid, v_a, which can be approximate as 0.1 nm^3 (82). This expression can be further simplified by replacing the outer radius, R, with $r + t$ (Fig. 4):

$$E_{TL/v} \approx \frac{4\pi\left((r+t)^3 - r^3\right)}{3v_a}\left(e_a + e_{ea} + e_{oa}\right) \approx \frac{4\pi r^2 t}{v_a}\left(e_a + e_{ea} + e_{oa}\right). \quad \textbf{[S21]}$$

The direct translation cost of a virus can be similarly obtained from Eq. **S21** by excluding the average opportunity cost of an

amino acid, e_{oa}. The critical radius, r_{crit}, at which translation and replication will have equal cost can be obtained by setting Eqs. **S19** and **S21** equal and solving for r (Fig. 4). Because capsid shell thickness is relatively conserved across icosahedral viruses studied to date, it can be treated as a constant equal to 3 nm (83). The critical radius for total cost estimates, $r_{crit-Tot}$, is 59 nm. For the direct cost estimates, the critical radius, $r_{crit-Dir}$, is 42 nm (Fig. 4).

XV. Discussion

In addition to the energetics of viral synthesis, another burst-limiting metric to consider is a volumetric one, namely the fraction of the host volume occupied by viruses during an infection. Taking influenza and T4 as our representative viruses, it is clear that they both occupy a relatively small percentage of the total host volume (84). A T4 infection takes up less than 5% of its host's total volume. An influenza infection takes up even less space (≈0.2% of its host volume). These estimates suggest that (*i*) for T4, the energy requirement is more likely a burst-limiting factor than the volumetric requirement, and (*ii*) for influenza, neither energetic nor volumetric factors seem to be limiting the burst size.

We have already considered possible causes for the inefficiencies of an influenza infection, which have experimentally been shown to result in only 1 infectious virus out of every 10 produced (40). Accounting for this inefficiency boosts the total cost of an influenza infection to ≈ 10% of its host's. A second consideration that could explain the relatively low energy uptake of an influenza infection is the growth state of its host cell. Our current estimate assumes that the host cell is under maximal growth conditions. When the host cell is not dividing, however, its energy supply could be as low as 10^{12} P_T (estimated for a 12-h infection period; SI section VIII). In considering both the 10% infection efficiency and assuming a slow-growing host cell, the influenza infection's total cost could also be a significant fraction of its host's total cost.

A final consideration is that implicit in our original question is the assumption that all viruses must conform to producing the maximum burst size allowed by their host energetic supply. This assumption, although it appears compatible with our findings for phages, may simply not be true for viruses such as influenza. It may be that a eukaryotic virus within a multicellular setting has an entirely different growth strategy than a prokaryotic virus infecting a single-celled organism. The influenza virus burst size is not only under selection pressure within its host cell but also within the multicellular organism that serves as its secondary host.

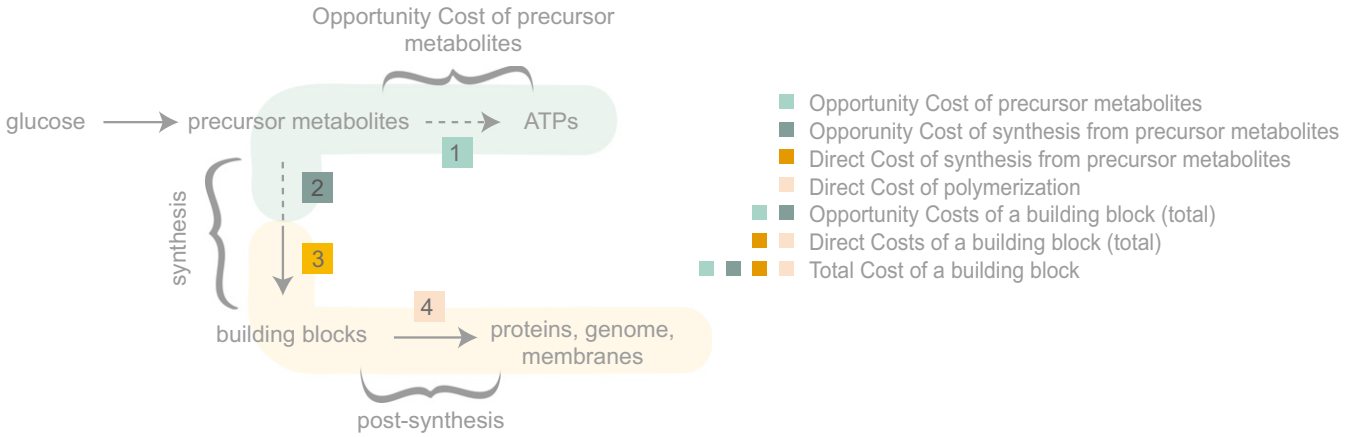

Fig. S1. Defining the opportunity and the direct costs of monomeric building blocks. During aerobic conditions, glucose is converted to precursor metabolites and eventually results in the production of 26 ATPs (in bacteria) when fully metabolized into water and carbon dioxide. If a precursor metabolite is instead channeled toward the synthesis of building blocks (step 2 and 3), the number of ATPs that could have been generated from its full metabolism minus the number of ATPs that have already been generated from its partial metabolism serves as its opportunity cost (step 1). The synthesis of building blocks from precursor metabolites incurs both opportunity (step 2) and direct (step 3) costs. The creation of macromolecular structures such as proteins from building blocks invokes postsynthesis costs such as the cost of polymerization (step 4). Direct costs are denoted in shades of orange, and opportunity costs are denoted in shades of green. The direct costs of a building block can be attained by the sum of steps 3 and 4, whereas the opportunity costs of a building block is the sum of steps 1 and 2. The total cost of a building block is the sum of steps 1–4.

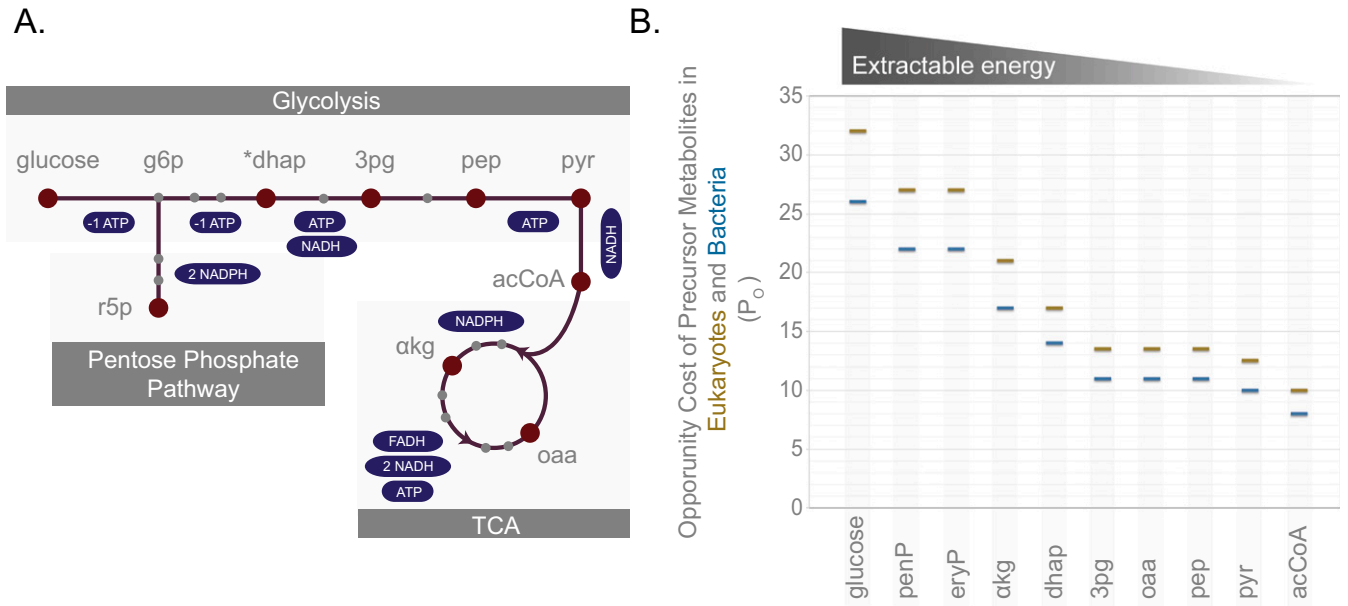

Fig. S2. The metabolic pathways and opportunity cost associated with each precursor metabolites. (*A*) Metabolic pathways and precursor metabolites involved in the synthesis of building blocks. Metabolites that do not serve as precursors are denoted as gray circles. The energetic gain or loss from the conversion of one precursor metabolite to another is depicted (the energetic loss is denoted by a negative sign). The asterisk signifies that, from dhap onward, two molecules of each precursor metabolite are generated; however, the energetic gains or losses are reported on a per-molecule basis. (*B*) The opportunity cost of each precursor metabolite in the context of bacterial and eukaryotic aerobic metabolism (using glucose as the sole carbon source) in units of P_O. See SI sections II–IV for a full description of these opportunity cost estimates.

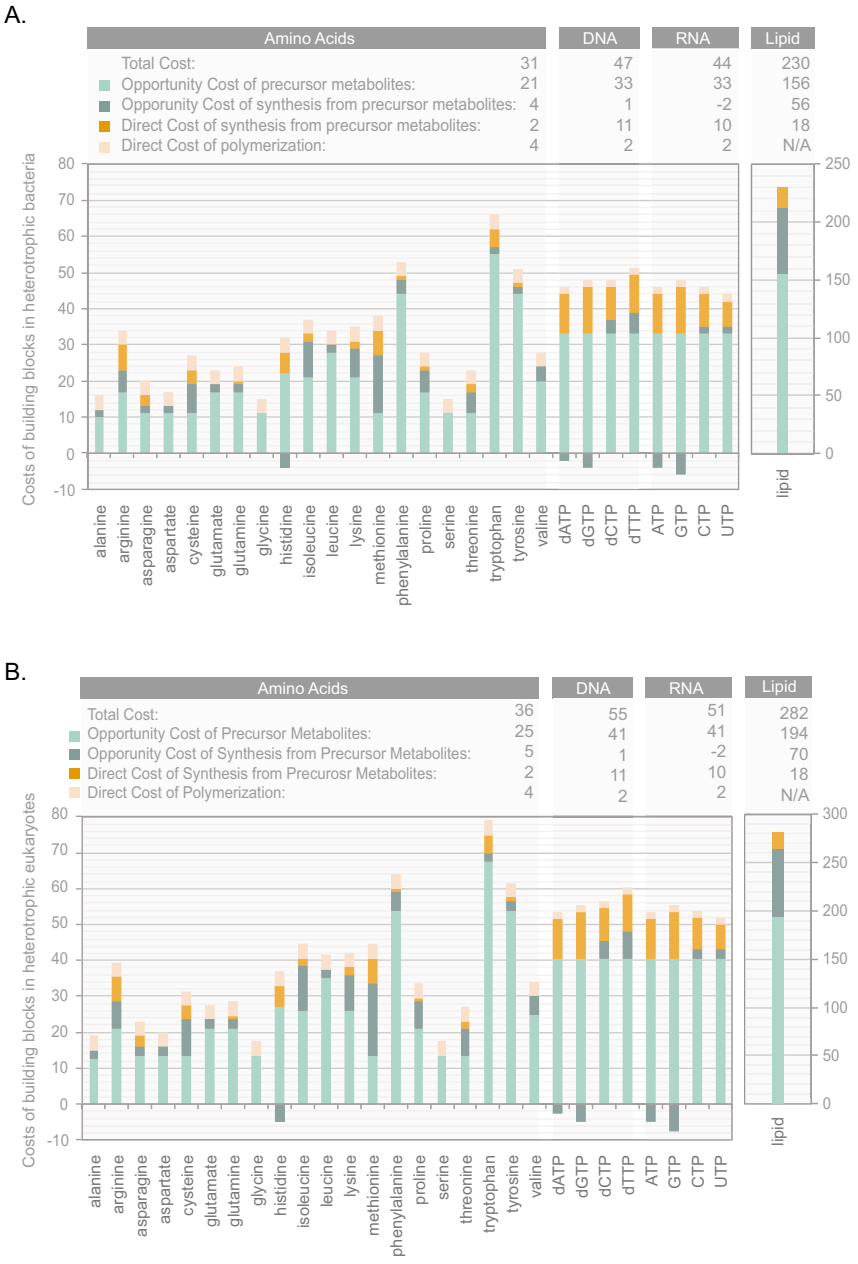

Fig. S3. The breakdown of direct and opportunity costs associated with amino acids, DNA, RNA, and lipids in the context of (*A*) bacterial and (*B*) eukaryotic metabolism. Average cost values are reported in the table above the chart. See Dataset S1 for a detailed derivation of these costs.

Table S1. T4 bacteriophage structural proteins and their average copy numbers per virion

T4 bacteriophage			
Protein	Protein length (L_{pj})	Average protein copy number (N_{pj})	Total no. of amino acids in protein j ($L_{pj}N_{pj}$)
23*	521	930	484,530
20*	524	12	6,288
24*	427	55	23,485
soc*	80	840	67,200
hoc*	376	160	60,160
22*	269	576	154,944
21*	212	72	15,264
IPIII*	194	370	71,780
IPI*	95	360	34,200
IPII*	100	360	36,000
alt*	682	40	27,280
68*	141	240	33,840
67*	80	341	27,280
3†	176	6	1,056
53†	196	6	1,176
5†	575	3	1,725
6†	660	12	7,920
7†	1,032	6	6,192
8†	334	12	4,008
9†	288	18	5,184
10†	602	18	10,836
11†	219	18	3,942
12†	527	18	9,486
15†	272	6	1,632
18†	659	144	94,896
19†	163	144	23,472
25†	132	6	792
26†	208	Assumed 1	208
27†	391	3	1,173
28†	177	Assumed 1	177
29†	590	6	3,540
48†	364	6	2,184
54†	320	6	1,920
td†	286	3	858
frd†	193	6	1,158
holin‡	218	20	4,360
endolysin‡	164	20	3,280
spanin‡	216	20	4,320
Totals		$\sum_{j=1}^{k} N_{pj}$ = 4,805 proteins	$\sum_{j=1}^{k} L_{pj}N_{pj}$ = 1,225,786 aa

The number of amino acids comprising each virion is calculated by the product of the average protein copy number and the length of the corresponding protein.

*These genes together compose the phage head.

†These genes are those that make up the tail tube, the tail sheath, and the base plate (table modified from ref. 85).

‡These genes are those that are involved in lysis (76).

Table S2. Influenza A virus proteins and their average copy numbers per virion

Influenza A virus				
RNA segment lengths (no. of nucleotides)	Protein product	Protein length (L_{pj})	Average protein copy number (N_{pj})	Total no. of amino acids in protein j ($L_{pj}N_{pj}$)
1 (2,341)	Polymerase PB2	759	45	34,155
2 (2,341)	Polymerase PB1	757	45	34,065
3 (2,233)	Polymerase PA	716	45	32,220
4 (1,778)	Hemagglutinin	566	500	283,000
5 (1,565)	Nucleoprotein	498	1,000	498,000
6 (1,413)	Neuraminidase	454	100	45,400
7 (1,027)	Matrix protein M1	252	3,000	756,000
	Matrix protein M2	97	40	3,880
8 (890)	NS1	230	0	0
	NS2	121	165	19,965
Totals			$\sum_{j=1}^{k} N_{pj}$ = 4,940 proteins	$\sum_{j=1}^{k} L_{pj}N_{pj}$ = 1,706,685 aa

The number of amino acids comprising each virion is calculated by the product of the average protein copy number and the length of the corresponding protein (table modified from ref. 86).

Dataset S1. The detailed breakdown of opportunity and direct costs of building blocks across heterotrophic bacterial and eukaryotic metabolisms (using glucose as the sole carbon source). All references are provided as notes in the Excel sheets

Dataset S1

Dataset S2. A list of viruses and their associated costs used to estimate replication to translation cost ratios shown in Fig. 4

Dataset S2

Dataset S3. A list of direct and total fractional cost estimates, E_g, for genetic elements of lengths 1, 10, 100, 1,000, and 10,000 bp across 30 dsDNA viruses (Fig. 5). Genetic elements are assumed to have no functional benefit to the virus and to be nontranscribed. Viruses A, B, and C correspond to hypothetical viruses

Dataset S3