Cell Geometry and Membrane Protein Crowding Constrain Growth Rate, Overflow Metabolism, Respiration, and Maintenance Energy.

Ross P. Carlson[1*], Tomáš Gedeon[2], Mauricio Garcia Benitez[3], William R. Harcombe[4], Radhakrishnan Mahadevan[3], Ashley E. Beck[5],

1. Department of Chemical and Biological Engineering, Center for Biofilm Engineering, Montana State University, Bozeman, MT USA
2. Department of Mathematical Sciences, Montana State University, Bozeman, MT USA
3. Department of Chemical Engineering and Applied Chemistry, University of Toronto, Canada
4. Department of Ecology, Evolution, and Behavior, University of Minnesota, St. Paul, MN USA
5. Department of Biological and Environmental Sciences, Carroll College, Helena, MT USA

[*]Corresponding author: R.P.C., rossc@montana.edu

Keywords: systems biology, cell biology, metabolism, cell geometry, surface area to volume ratio

Abstract

A metabolic theory is presented for predicting maximum growth rate, overflow metabolism, respiration efficiency, and maintenance energy flux based on the intersection of cell geometry, membrane protein crowding, and metabolism. The presented membrane-centric theory employs biophysical properties and metabolic systems analysis to successfully predict phenotypic properties of *Escherichia coli* K-12 strains MG1655 and NCM3722. The strains are genetically similar but differ in surface area to volume (SA:V) ratios by up to 30%, maximum growth rates on glucose media by 40%, and overflow-inducing growth rates by 80%. The predictions were tested against experimental evidence including phenomics data, membrane proteomics data, and MG1655 SA:V mutant growth rates. The predictions were remarkably consistent with experimental data and provided a membrane-centric explanation for maximum growth rate, maintenance energy generation, respiration chain efficiency (P/O number), and optimal biomass yield of a strain. These analyses did not consider cytosolic macromolecular crowding, highlighting the distinct properties of the presented theory and gaps in current cell biology literature. Cell geometry and membrane protein crowding are significant biophysical constraints and consideration of both provide a more complete theoretical framework for improved understanding and control of cell biology.

INTRODUCTION

Bacterial cell geometry, including cell shape and size, is highly regulated and constrains the surface area available for acquiring nutrients and the volume available for synthesizing proteins (1-4). Cell geometry also impacts molecular crowding by fixing the two- and three-dimensional spaces available for macromolecules (5-9). Three-dimensional cytosolic molecular crowding has been analyzed both explicitly and implicitly using computational biology to predict catabolite repression and phenotype shifts including overflow metabolism or the production of fermentation products in the presence of electron acceptors (5-7, 10, 11). A quantitative and predictive understanding of the intersection of cellular geometry, two-dimensional membrane protein crowding, and phenotype, however, is not well studied (12-16).

Cell membranes provide a barrier against environmental stressors, retain macromolecules, provide a platform for the selective transport of molecules, and play a crucial role in energy metabolism including respiration (17, 18) (Fig. 1A). Membranes support major metabolite fluxes for macromolecule synthesis and maintenance energy generation. Intracellular reactions for macromolecular synthesis are well characterized while requirements for maintenance energy are largely a black box. Maintenance energy is cellular energy consumed for functions other than the direct production and polymerization of biomass components (19). Meeting maintenance energy requirements accounts for a significant fraction (30 to nearly 100%) of substrate fluxes (20-23). Maintenance energy calculations have not previously considered the impact of biophysical constraints like finite membrane surface area and membrane protein crowding.

Cellular surface area to volume ratios (SA:V) and membrane macromolecular crowding are the two foci of the presented work. Systems biology studies have applied enzyme metrics, such as total protein mass to approximate crowding on a two-dimensional membrane surface, but these approaches generally miss the significant difference between a two- and three-dimensional crowding constraint (10, 18, 24) (Fig. 1B, 1C). The current study makes a distinction between two- and three-dimensional protein crowding on the two-dimensional membrane surface by accounting for biophysical properties of embedded and associated enzymes (10, 24). Enzyme requirements for surface area are not readily predictable from the enzyme mass or volume (Fig.

1B, 1C, supplemental S1), highlighting the need to study surface properties separately from intracellular crowding. Additionally, the lipid bilayer is a major building block of the membrane and has a finite capacity to host embedded and adsorbed proteins due to a combination of steric effects and the loss of membrane integrity at high protein loading (17, 25-28).

Escherichia coli is a facultative anaerobe and a convenient host for studying tradeoffs between cell geometry, membrane protein crowding, and phenotype. Two well studied and genetically similar *E. coli* K-12 strains, MG1655 and NCM3722, have distinct phenotypes including different maximum growth rates on minimal media, different acetate overflow metabolisms, and different cellular geometries (supplemental S1, S2). These cells display exquisite control over their geometry including cell length and width, and we propose that this contributes to the phenotypic differences. The SA:V ratios influence the balance of nutrient and energy fluxes between surface area- and volume-associated processes (1-3). While theoretical studies have addressed the role of surface area on metabolism (15, 16), no study has developed a quantitative and predictive molecular level theory that accounts for strain-specific differences in SA:V ratios, growth rate dependent changes in SA:V ratios, and growth rate dependent changes in membrane protein crowding (29). The presented systems biology construct demonstrates the remarkable theoretical impact of cell geometry and membrane protein crowding on foundational cellular processes including maximum growth rate, onset of overflow metabolism, respiration efficiency, and ultimately the cellular capacity for energy conservation and maintenance energy generation. These predictions are all accomplished without consideration of cytosolic macromolecular crowding nor absolute cellular volume, highlighting the incomplete nature of current cell biology theory. We propose that membrane-associated constraints are complementary to previously described cytosolic-based constraints.

RESULTS

E. coli K-12 strains MG1655 and NCM3722 have different SA:V ratios and phenotypes

We hypothesize that the biophysical properties of the bacterial inner membrane constrain phenotype and developed a systems biology methodology to test this hypothesis. The methodology was developed using *E. coli* strain MG1655 data, and the predictions were then tested using data from *E. coli* strain NCM3722. These two strains are genetically similar but have

distinct phenotypes including different maximum growth rates on minimal salts medium, different acetate overflow metabolisms, and different surface area to volume (SA:V) ratios (Fig. 2A, 2B, 2C). Si *et al*. measured MG1655 and NCM3722 cellular dimensions as a function of growth rate (2). These foundational, but technically challenging, measurements can differ substantially between research groups (2, 30, 31). Having a cell size data set for both strains collected from a single research group builds confidence in the comparisons. SA:V ratios decreased with increasing growth rate because the cell length and width increased (Fig. 2C) (1, 2). NCM3722 cells are smaller than MG1655 at low growth rates, leading to higher SA:V ratios (Fig. 2B, 2C). The different SA:V ratios are proposed to play a role in strain phenotypes. NCM3722 has a specific growth rate on glucose salts media that is ~40% faster than MG1655 (0.97 ± 0.06 h^{-1} vs. 0.69 ± 0.02 h^{-1}, respectively; Fig 2A, supplementary data S1, S2). MG1655 displays an acetate overflow metabolism at growth rates $\geq 0.4 \pm 0.1$ h^{-1} while NCM3722 displays acetate overflow at growth rates $\geq 0.75 \pm 0.05$ h^{-1}, which are faster than μ_{max} for MG1655 (Fig. 2A, supplementary data S1). Additionally, MG1655 has a cellular volume that is approximately two times larger than NCM3722 at a growth rate of 0.65 h^{-1}, whereas MG1655 has a SA:V ratio ~30% smaller than NCM3722 at this growth rate (Fig. 2B, supplementary data S2). We present evidence that the cell dimensions, and the membrane capacity to host proteins, correlate with the phenotypic properties of the strains, suggesting these biophysical properties are important for analyses of cell biology.

Membrane proteome is dynamic with respect to growth rate and environment

The *E. coli* membrane proteome is dynamic as a function of growth rate and environment (32-36) (Fig. 3A, 3B, 3C). The membrane hosting of central metabolism proteins (protein crowding) per cell volume increased with growth rate (Fig. 3A) (33-36) (supplementary data S3). In fact, the protein hosting increased faster than the decrease in SA:V (Fig. 2B), resulting in a net gain of membrane catalytic potential with growth rate. Note both sets of data are plotted on a volume basis to facilitate comparisons. The areal densities of central metabolism proteins for *E. coli* K-12 were determined using three proteomics data sets (33-36) and cell geometry (2). The presented analyses pooled data for *E. coli* K-12 strains MG1655 and BW25113 which are closely related and have similar growth rates, cell geometries, and phenotypes (30). For simplicity they are collectively referred to here as MG1655. Copy number of membrane-associated proteins was

converted into an occupied surface area per cell volume using enzyme properties and cell geometry (2, 33-36) (supplementary data S3). The presented trends are for glucose minimal salts media. Faster growth rates were reported in media containing multiple substrates, but the studies did not include detailed substrate uptake data, precluding its inclusion in the current analysis.

E. coli MG1655 synthesized increasing numbers of glucose transporter complexes (PtsG) per volume with growth rate (Fig. 3B). However, *E. coli* cultures grown in the absence of glucose expressed much lower numbers of PtsG complexes. The proteomics data are from chemostat cultures suggesting the cells have a finite surface area and prioritize environmentally relevant membrane proteomes at high growth rates (Fig. 3B). Meanwhile, *E. coli* grown on glucose at low dilution rates synthesized increasing levels of alternative, high affinity substrate transporters for arabinose, glycerol, and xylose as the dilution rate decreased (Fig. 3C). These substrates were not present in the media and the phenotypes likely represented a hedge strategy preparing the energy limited cells for alternative substrates, should they become available (32, 33) (supplementary data S3). These unused proteins were only synthesized at low dilution rates.

We propose that finite membrane surface area and membrane protein crowding preclude the simultaneous and constitutive expression of all types of membrane proteins at all growth rates. The dynamic nature of the membrane proteome is hypothesized to reflect the substantial constraint it places on *E. coli* phenotypes. The remainder of the analysis presents theory and data to support this proposal.

Systems biology of cellular SA:V ratios and membrane protein crowding

The foci of this study are the inner membrane surface area, membrane protein hosting capacity, the areal requirements of membrane associated proteins, and the collective role of these biophysical parameters on phenotype. The specific Membrane Surface Area (*sMSA*, units: nm^2 (g cdw)$^{-1}$) required for membrane-associated enzyme fluxes was quantified using enzyme properties including flux per g cell dry weight (cdw), surface area requirements (nm^2 enzyme^{-1}), and k_{cat} (molecules (s enzyme)$^{-1}$). Total *sMSA* required for a phenotype was tabulated by summing the contribution of all active membrane-associated enzymes and by accounting for the

growth rate dependent SA:V ratio. Mathematical development is presented in Box 1 while parameters are reviewed in supplementary data S1 and S4.

The *sMSA* constraint was evaluated using a metabolic model of *E. coli* (37, 38) and flux balance analysis (FBA) (39). The *sMSA* constraint was implemented by incorporating a membrane surface area requirement for each membrane-associated reaction, termed MRE_i (membrane real estate) (Box 1) (15). Flux through membrane-associated enzymes was constrained by a finite *sMSA* pool which changed in magnitude based on growth rate dependent SA:V ratios and protein crowding (quantified as the fraction of the Membrane Surface Area occupied by central metabolism proteins, f_{MSA}) (supplementary data S5, Box 1). The *sMSA* balance was enforced concurrently but orthogonally to standard FBA mass balances (mmol (g cdw h)$^{-1}$). The *in silico* theory was termed specific Membrane Surface Area-constrained Flux Balance Analysis (*sMSAc-FBA*) (39). This novel formulation contrasts with previous studies where the weights on the membrane-associated fluxes were the molecular masses of the protein, which correlates with enzyme volume (10, 18, 24, 40), as opposed to the surface area occupied as in the presented work (Fig. 1C). The presented analyses considered neither cytosolic protein investment nor cytosolic macromolecular crowding.

E. coli MG1655 growth rate and biomass yield are constrained by membrane surface area, protein crowding, and maintenance energy

Cell growth requires fluxes of nutrients across the membrane for biomass synthesis and maintenance energy generation. The theoretical impact of finite *sMSA* on these fluxes was quantified for *E. coli* strain MG1655 growing on glucose salts medium (Fig. 4A). Briefly, a matrix of biomass synthesis rates (q_x) and maintenance energy fluxes (q_{ATP}, aggregate of growth- and nongrowth-associated maintenance energy (GAM, NGAM, respectively)) was defined and applied as FBA flux constraints. For each parameter combination comprised of a unique q_x and q_{ATP}, an *sMSA* optimized phenotype was identified by minimizing glucose uptake (f_{MSA} = 0.07, Fig. 4A, Box 1).

Metabolic processes competing for a finite pool of *sMSA* resulted in rate-yield tradeoffs (16, 41). Phenotypes with high biomass yields (g biomass (g glucose)$^{-1}$) required larger investments of *sMSA* for high efficiency electron transfer chain (ETC) components, quantified by the P/O

number (Fig. 4B). The P/O number is the ratio of ATP produced via respiration per two electrons transferred to O_2. *E. coli* has the genetic potential to operate ETC configurations with P/O numbers from ~0 to 2 (42-44). ETC efficiency is not influenced by substrate-level phosphorylation such as ATP synthesis associated with acetate overflow. Phenotypes with low growth rates and small maintenance energy fluxes can support high P/O numbers because of the modest *sMSA* requirements for transporters and ATP synthase (upper left Fig. 4B). Phenotypes with high growth rates require larger investments of *sMSA* for glucose transporters (PtsG) (Fig. 4C, supplementary data S6). These tradeoffs generate gradients in other phenotypic properties including the potential for maintenance energy fluxes (q_{ATP}) (Fig. 4A), use of parallel ETC NADH dehydrogenase enzymes Nuo (Fig. 4D) or Ndh II (Fig. 4E), and acetate overflow metabolism (q_{acetate}) (Fig. 4F). Secretion of acetate generates ATP via substrate-level phosphorylation and reduces the need for *sMSA*-intensive ETC components and ATP synthase complexes (15, 16). Presented q_{ATP} accounts for the capacity to generate the sum of growth- and nongrowth-associated maintenance energy (GAM and NGAM, respectively) but did not include ATP required to polymerize monomers into macromolecules for biomass synthesis (supplementary data S5).

μ_{MAX} is a Pareto tradeoff between biomass production and maintenance energy generation

Finite *sMSA* may necessitate phenotypic tradeoffs providing a new lens with which to interpret maximum growth rate. Consensus extracellular fluxes were assembled for *E. coli* strain MG1655 growing on glucose salts media (Fig. 5A, supplementary data S1). Data span growth rates from 0.1 h^{-1} to μ_{max} (6, 35, 40, 45, 46). The consensus data were interpreted using *sMSA* theory by applying the fluxes as constraints to *sMSAc-FBA* analyses and quantifying maximum possible maintenance energy generation (q_{ATP}) as a function of growth rate, cell geometry, and f_{MSA} (0.02-0.09) (Fig 5B, 5C, 5D). Predicted maintenance energy was therefore a sum of growth and nongrowth-associated maintenance energy. q_{ATP} was calculated 100 times for each scenario by randomly and independently perturbing each experimental consensus flux within an assumed error interval (±7% of flux) and using these perturbed fluxes as the *sMSAc-FBA* reaction constraints. Data points in Fig. 5C – 5E show the mean from 100 simulations; error bars represent the standard deviation.

Positive slopes between growth rate and maintenance energy flux (Fig. 5C) indicate that both rates can theoretically increase simultaneously without penalty as the membrane can accommodate a balanced increase of substrate transporters, ETC complexes, and ATP synthase complexes. A slope of zero indicates that the growth rate can increase with no reduction in q_{ATP}; the membrane hosts more substrate transporters while holding the other processes constant. A negative slope indicates that the growth rate can increase only with a concurrent decrease in q_{ATP}; the increased substrate transporter demand requires a decrease in ETC components and ATP synthase complexes. Negative slopes define a Pareto front between simultaneous increases in growth rate and q_{ATP} (Fig. 5C, 5D) (47).

Membrane protein crowding (f_{MSA}) has the potential to constrain the energetic efficiency of respiration by limiting the ability of the cell to extract useful energy from substrate. For example, there was ~4-fold increase in predicted MG1655 maintenance energy potential (10 ± 3 vs. 41 ± 4 mmol ATP (g cdw h)$^{-1}$) for the same consensus fluxes (μ = 0.5 h^{-1}, $q_{glucose}$ = 7.2 ± 0.5 mmol (g cdw h)$^{-1}$, $q_{acetate}$ = 1.7 ± 0.1 mmol (g cdw h)$^{-1}$) when the enforced protein crowding increased from 4% to 9% (f_{MSA} = 0.04, 0.09) (Fig. 5D, supplementary data S6). The difference in q_{ATP} potential is based on the availability of *sMSA* required to use higher efficiency ETC components and assemble more ATP synthase complexes. We hypothesize that *E. coli* operates at a maximum membrane protein capacity as this would significantly increase the possible maintenance energy generation, thereby improving fitness.

Experimental *E. coli* MG1655 cultures have a μ_{max} of 0.69 ± 0.02 h^{-1} on glucose salts media (Fig. 2A, blue shaded area, Fig. 5D) which coincides with the start of a negative slope between growth rate and q_{ATP} (f_{MSA} = 0.07). The predicted f_{MSA} = 0.07 is lower than values estimated by Szenk *et al.* (15). The current study adjusted enzyme k_{cat} values to a culture temperature of 37°C, resulting in some faster k_{cat} values and therefore smaller f_{MSA} requirements (48). All reactions in the *in silico* model were balanced for atoms and electrons (supplementary material S5). Therefore, all simulations necessarily balanced consensus fluxes with appropriate CO_2 (necessary for carbon balance) and O_2 (necessary for electron balance) fluxes as no other electron acceptors were available in the simulations.

Submaximal P/O numbers optimize energy conserving potential of finite membrane surface area

E. coli has a branched ETC that can convey electrons using multiple enzymatic routes to acceptors including O_2 (42-44). These enzymes can conserve varying amounts of substrate chemical energy by converting it into chemiosmotic energy stored as proton motive force. As mentioned previously, *E. coli* ETC components can operate at P/O numbers between ~0 and 2 (42-44). P/O number can significantly influence maintenance energy calculation (49). *E. coli* cultures typically operate at submaximal P/O numbers in the range of 1 to 1.5 even at O_2 sufficiency (reviewed in supplementary data S4). Few theories have been proposed to explain this phenotype (50, 51). We hypothesize *sMSA* theory is relevant to the experimentally observed submaximal P/O numbers.

Data from four proteomics studies (33-36) and enzyme k_{cat} numbers were used to calculate experimental P/O numbers as a function of growth rate (Fig. 5E, supplementary data S6, S7). Theoretical growth rate dependent P/O numbers were calculated using *sMSAc-FBA* using experimental consensus fluxes, SA:V ratios, and growth rate dependent membrane protein crowding (Fig. 5E, supplementary data S6, S7). *In silico* P/O numbers at μ_{max} were comparable to the experimental values (1.04 ± 0.01 and 1.01-1.07, respectively). We propose the submaximal P/O numbers represent an optimized use of the finite *sMSA* that enables optimal substrate fluxes for growth and cellular energy generation and now present additional evidence for this conclusion. Predicted P/O numbers increase rapidly as growth rate decreases. We propose slow-growing cells are not substantially constrained by *sMSA* as the slow growth rate does not require as many glucose transporters, ETC components, or ATP synthase complexes. Slow growing cells have sufficient *sMSA* to permit the expression of alternative substrate transporters such as those shown in Fig. 3C.

Membrane surface area and membrane protein crowding predict onset of overflow metabolism.

The intersection of metabolism, membrane protein crowding, and cellular geometry predicts the onset of acetate overflow based on optimization of finite *sMSA*. The presented analysis is designed to further test the *sMSA* theory by highlighting its potential significance as a mediator

of major phenotypic strategies. *sMSAc-FBA* was used to calculate growth rate dependent P/O numbers for strain MG1655 using the experimental consensus fluxes, SA:V, and growth rate dependent membrane protein crowding (Fig. 5A, Box 1, supplementary data S7).

The theoretical P/O number vs. growth rate curve was fit empirically (analyses of different empirical fittings are presented in supplementary data S7). The inflection point was determined from the second derivative with respect to growth rate. The inflection point identified the growth rate where *sMSA* became sufficiently limiting to warrant the overflow metabolism. The inflection point at $\mu = 0.5\ h^{-1}$ correlates well with the experimental values for overflow onset, $0.4 \pm 0.1\ h^{-1}$ (Fig. 2A, supplementary data S1). At growth rates less than the inflection point, the analysis suggests cells have sufficient *sMSA* to operate substrate transporters, ETC apparatus, and ATP synthase without requiring the overflow tradeoff (Fig. 5A, 5E). The P/O number inflection also coincides with the maximum experimental biomass yields, providing a basis for defining an optimal biomass phenotype (Fig. 5A, 5E, supplementary data S1).

Modulating E. coli MG1655 SA:V alters maximum growth rate in predictable manner

Si *et al*. (2) examined the effect of MG1655 SA:V on batch growth rate. A titratable MreB knockdown system was used to modulate *E. coli* geometry by changing the cell diameter and cell length and therefore changing SA:V ratios (1-3, 52). The batch growth rate on defined glucose medium supplemented with amino acids was measured as a function of titrated SA:V. Cells with higher SA:V ratio grew faster than cells with lower SA:V ratios, supporting the *sMSA* theory presented here (Fig. 5F). Unfortunately, no amino acid uptake data was reported so a quantitative analysis of the fluxes, SA:V, and f_{MSA} was not possible. Interestingly, when the same data set was plotted as total cell volume vs. growth rate, a negative relationship was observed (Fig. 5F) (1, 3, 53). The same study also analyzed a strain with a titratable FtsZ system. The strain did not achieve a large range of SA:V ratios and the specific growth rates during batch growth were largely unchanged (supplementary data S2)

Testing sMSA predictions using data from E. coli strain NCM3722

sMSA theory, developed using *E. coli* strain MG1655 data, predicted several phenotypic properties based on SA:V ratios and membrane protein crowding (f_{MSA}). We propose if these

phenotypic properties can be predicted for another *E. coli* strain, such as strain NCM3722, then the theory holds meritorious promise. Synthesizing the theoretical framework, we predict the following behaviors will hold for *E. coli* strain NCM3722: **prediction 1,** *sMSAc-FBA* using an f_{MSA} of 0.07 will predict μ_{max} at the Pareto front between biomass synthesis and maintenance energy generation; **prediction 2,** *sMSAc-FBA* will only predict an accurate μ_{max} when biologically relevant SA:V ratios are applied to the analyses; **prediction 3,** *sMSAc-FBA* will quantitatively predict P/O ratios highlighting the optimal use of finite membrane surface area and protein crowding; **prediction 4,** *in silico* P/O ratios, plotted as a function of growth rate, will have an inflection point that corresponds with the experimental onset of acetate overflow; and **prediction 5,** the most efficient biomass yield on glucose will correspond with the inflection point of P/O number vs. growth rate, providing a basis for defining an optimal glucose phenotype. These evaluations of *sMSA* theory used the same metabolic model with the same enzyme parameters as applied for MG1655. However, strain-specific SA:V ratios (Fig. 2B) and consensus fluxes were applied (Fig. 6A).

Experimental and theoretical *E. coli* NCM3722 data were assembled to develop consensus fluxes (6, 40, 54). The majority of the NCM3722 data derived from a recombinant strain with an engineered glucose uptake system (6). Briefly, the strain had the native PtsG glucose transport component deleted, and in place, a genetic construct with a titratable PtsG protein was used to control growth rate. The recombinant NCM3722 strain was grown in batch, and the glucose transporter was saturated (enzyme K_M ~ 5μm, initial medium glucose concentration S ~ 11 mM) (Fig. 6B). The recombinant NCM3722 phenotype was applied to the *sMSAc-FBA* simulations using a PtsG enzyme saturation factor Φ_{PtsG} = 1 (Fig. 6B). This was in contrast with data for MG1655 which was collected from wild type cultures with growth rates controlled by substrate concentration modulated by chemostat dilution rates (6, 35, 40, 45, 46). MG1655 simulations accounted for changing PtsG enzyme saturation as a function of growth rate (Fig. 6B).

Testing prediction 1. NCM3722 data demonstrated trends analogous to MG1655, including the existence of a Pareto tradeoff between biomass synthesis and the capacity to generate maintenance energy. At a protein crowding of f_{MSA} = 0.07, the predicted μ_{max} of 0.95 h^{-1} correlated closely with experimental μ_{max} of 0.97 ± 0.06 h^{-1} (red shaded area, Fig. 6C, supplementary data S6). Both strains had similar predicted protein crowding at their respective

μ_{max} (f_{MSA} = 0.07) despite a 40% difference in maximum growth rate and substantial differences in SA:V ratios as a function of growth rate (Fig. 2B). Interestingly, strains MG1655 and NCM3722 are predicted to have a similar potential for q_{ATP} at their respective μ_{max} (37 ± 5 and 37 ± 4 mmol ATP (g cdw h)$^{-1}$) despite the faster growth rate of NCM3722 (Fig. 2A).

Testing prediction 2. Altering the NCM3722 SA:V ratio while holding f_{MSA} = 0.07 constant caused the predicted μ_{max} to deviate from the experimental value. The experimental SA:V was varied from +25% to -25% (Fig. 2B), and the μ vs. q_{ATP} analysis was repeated (Fig. 6D, supplementary data S6). Additionally, the experimental MG1655 SA:V ratio was applied to the NCM3722 simulations (Fig 6D). The predicted μ_{max} changed in a predictable manner; increasing SA:V predicted faster growth before the Pareto tradeoff occurred. Applying an MG1655 SA:V ratio to the NCM3722 fluxes predicted a μ_{max} of 0.75 h^{-1}, a substantial deviation from the experimental μ_{max} of 0.97 h^{-1}.

Testing prediction 3. *sMSAc-FBA* predictions of NCM3722 P/O numbers were quantitatively consistent with experimental values when experimental NCM3722 SA:V values were applied (Fig. 6E, supplementary data S6). Analogous to the MG1655 analysis, the predictions suggest the submaximal P/O numbers are actually optimal uses of the constrained *sMSA*. Predicted P/O numbers varied substantially when the SA:V was increased or decreased (Fig. 6E). Applying the experimental MG1655 SA:V ratios with experimental NCM3722 fluxes did not accurately predict strain phenotype (Fig. 6E, supplementary data S6). The P/O number increased with decreasing growth rate, again suggesting *sMSA* is a considerable constraint only at high growth rates.

Testing prediction 4. The inflection point in the NCM3722 P/O number vs. μ relationship predicted the onset of overflow metabolism. The analysis predicted acetate overflow at μ = 0.68 h^{-1}, which correlated well with the experimental onset at 0.75 ±0.05 h^{-1}. The P/O number vs. μ trends for seven different SA:V trends were considered (Fig. 6E, supplementary data S6). The trends demonstrate inflection points that were identified by taking the second derivative of the empirical polynomial fits. Altering the SA:V altered the predicted onset in a predictable manner. A summary of the predicted maximum growth rates and acetate overflow metabolism is presented in Fig. 6F.

Testing prediction 5. Experimental NCM3722 biomass yields on glucose were plotted along with the P/O numbers as a function of growth rate (Fig. 6E, supplemental data S1, S6). Analogous to MG1655 data, the maximal biomass yield corresponded closely with the inflection point for the P/O vs. μ relationship (0.66 vs. 0.68 h^{-1}, respectively). Finally, it should be noted that the *sMSAc-FBA* predictions were remarkably accurate despite the majority of the NCM3722 data being measured using a recombinant strain with an altered membrane proteome.

DISCUSSION

The rules of cell design remain elusive despite decades of research. For example, the design rules that link cell geometry and membrane protein capacity with central metabolism have not been defined previously on a quantitative basis for different strains, growing at different rates, with different SA:V ratios. Bacteria display exquisite control over cellular geometry and membrane properties which are proposed here to be biologically significant as these properties are major mediators of possible phenotypes (1, 3, 18). *E. coli* cell length and diameter increase with growth rate, reducing the SA:V ratio (Fig. 2B, 2C) (2, 12). However, as shown here, the reduction in SA:V is offset by an increase in the capacity of the membrane to host proteins (Fig 3A, supplementary data S3). This property is hypothesized to be critical to competitive phenotypes; if membrane protein content were not essential to fitness, this capacity would likely remain independent of, or decrease with increasing growth rate (18). This membrane property was integrated into the new, molecular level theory which combined metabolism, cell geometry, and membrane protein crowding. *sMSA* theory was used to predict maximum growth rates, acetate overflow metabolism, electron transport chain efficiency, and maintenance energy fluxes (6, 24, 40). The predictions were compared to experimental phenomics data, membrane proteomics data, and MG1655 strains with titrated SA:V ratios (Fig. 5 and 6). In all cases, the theory was remarkably accurate without accounting for either cytosolic volume or cytosolic macromolecular crowding (6, 12, 24).

Much effort has documented the importance of volume-associated proteome investment and molecular crowding on phenotype (6, 12, 24). Volume-based processes are certainly central to phenotype (5, 6, 9), but we hypothesize that both surface area- and volume-associated processes are critical and concurrently influence phenotype (1). Surface area requirements for

membrane enzymes do not correlate well with volume proxies like enzyme mass (Fig. 1C, supplementary data S1); membrane surface area requirements are instead proposed to be distinct biophysical properties which impose unique constraints (24, 40). Tight regulation of cellular geometry and therefore SA:V ratio would arguably be futile if surface area-associated substrate transport and volume-associated protein synthesis were not also tightly regulated. If a single geometric aspect were limiting, evolution could select cells with changes in either cell length or cell diameter to overcome the surface area or volume limitation, respectively. A potential strategy to overcome the limits of geometry is to evolve membranes with higher protein hosting capacities. Mitochondria are eukaryotic organelles specialized in ATP generation (12). While the geometries of mitochondria are similar to bacteria, their membranes can host proteins at areal densities approximately two-fold higher than an *E. coli* membrane (24 ± 5.6 % vs. 40-50%) (4, 12, 55) (reviewed in supplementary data S4) (Fig. 7A). Mitochondria membranes have likely evolved the capacity to host high protein fractions while maintaining membrane integrity because they persist in the relatively safe environment of the buffered and relatively safe eukaryotic cytosol. The role of cell geometries in different bacterial morphologies, *e.g.* bacilli vs. cocci, or the membrane protein capacities of Gram-negative vs. Gram-positive membrane architectures are open questions that can be addressed using *sMSAc-FBA* (12).

sMSAc-FBA predictions assumed an idealized scenario including 1) all enzymes (except PtsG in MG1655 simulations) were saturated for substrate including ETC components and O_2 cytochromes, 2) cells produced optimal levels of each enzyme with no overproduction, 3) all protein complexes were perfectly assembled with exact subunit ratios, 4) all enzymes were flawlessly embedded in the membrane, 5) no allosteric regulation of enzyme activity occurred based on membrane curvature (3, 56), and 6) the enzyme k_{cat} numbers increased two-fold for every 10°C increase in temperature (Q_{10} = 2) (48). The curated proteomics data provided a basis for estimating deviations from this idealized scenario. Using a protein offset (Δ, Box 1) of ~3% quantitatively aligns the *sMSAc-FBA* simulations with the experimental proteomics data without altering any presented interpretations or conclusions. The presented study implicitly included a perturbation analysis of Δ. For example, perturbing Δ for batch growth can be analyzed by examining results for f_{MSA} values larger or smaller than 0.07 (Fig. 5C, 6B).

Computational biology often assumes cells will adopt phenotypes that tend toward a maximum growth rate. This computational growth rate is, in actuality, a relative maximum growth rate and not an absolute maximum growth rate. Fig. 7B plots the experimental MG1655 batch growth rate relative to the possible growth rates based on the presented *sMSA* theory. The experimental data maps to the maximization of biomass yield for a fixed maintenance energy demand not the maximization of growth rate. If the cell reduces maintenance energy demand, it could theoretically increase its growth rate while maintaining the same yield. For example, reducing the q_{ATP} requirement by 8-10 mmol ATP (g cdw h)$^{-1}$ could theoretically increase growth rate by 0.1 h^{-1}. In fact, adaptive laboratory evolution can be used to create *E. coli* strains that grow faster than wild type cells; additionally, some gene deletions can increase growth rate (57, 58). Minimizing maintenance energy demand through reductions in futile cycling may contribute to the phenotypes. These strains increase growth rates, but likely at the expense of robustness. The presented work provides a quantitative framework to assess the changes.

Enzyme k_{cat} numbers vary substantially between published studies (supplementary data S1) (59). The presented study performed a literature survey and applied the same median values to simulations for MG1655 and NCM3722 (supplementary data S4). The robustness of the presented results was analyzed by perturbing the MRE_i parameter which accounted for enzyme properties including k_{cat}, degree of saturation, and surface area requirements (Box 1). The MRE_i parameter was randomly and independently perturbed within ±95% of the base value. Experimental fluxes were also perturbed randomly and independently within a standard deviation of the measured flux. Each scenario was analyzed with 100 simulations (Fig. 7C) (supplementary data S8). Capacity for q_{ATP} was robust to perturbations. Calculated q_{ATP} varied ≤ 10% for random perturbations of all MRE_i parameters up to ±95% of the base value (Fig. 7C) (supplementary data S8). *sMSA* theory is robust to the presented assumptions, likely based on the compensating nature of the redundant and parallel structures in metabolic networks (60). Given the robustness of the results, it is probable the parameters can be used in other animalcules with reasonable accuracy.

Additional robustness analyses examined perturbing one MRE_i parameter at a time to assess the effect on q_{ATP}. These analyses considered five membrane-associated enzymes which accounted for >80% of the predicted surface area during batch growth (ATP synthase, glucose transporter

(PtsG), ammonium transporter (AmtB), NADH dehydrogenase (Nuo), and cytochrome oxidase (Cyo); supplementary data S8). Each enzyme MRE_i was perturbed ±35% to increase or decrease the *sMSA* requirements, respectively. ATP synthase accounted for 45% of the surface area utilized by central metabolism enzymes and was most sensitive to MRE_i perturbations. Increasing the ATP synthase MRE_i parameter 35% while holding all other MRE_i parameters constant decreased the potential for q_{ATP} fluxes by ~26%, whereas decreasing the MRE_i parameter 35% increased the potential for q_{ATP} fluxes by ~8%. Perturbing the MRE_i parameter for the other four enzymes individually ±35% did not change the capacity for q_{ATP} fluxes more than ±5% (supplementary data S8).

Biological optimality principles are of keen interest as they illuminate the (pseudo) rules of life and are readily implemented using systems biology methods. The presented study demonstrates that the commonly observed submaximal P/O numbers are likely not suboptimal strategies. Instead, they represent an optimal tradeoff between surface area requirements for substrate transporters and ETC components. Other commonly observed submaximal, yet likely optimized, phenotypes include overflow metabolisms occurring at both high (6, 61) and low growth rates (45), use of the Entner-Doudoroff (ED) glycolysis rather than Embden-Meyer-Parnas (EMP) glycolysis (60), and use of a partial citric acid cycle under nutrient limitation (62). Numerous *in silico* studies have replicated the overflow and ED vs. EMP glycolysis phenotypes using a variety of criteria (6, 14-16, 40, 60, 63). It is probable that many of these criteria are concurrently relevant to cell geometry (24, 40). All the criteria explicitly or implicitly define tradeoffs between resource sufficiency and resource scarcity and can be viewed through a lens of cellular economics. The relative value of a limiting resource is higher than for resources present in excess. Therefore, fitness is increased when use of the limiting resource is optimized (37, 60). The presented *sMSA* theorical framework quantifies a metabolic resource defined by cell geometry and membrane protein crowding and elucidates cell design rules which govern selection of competitive phenotypes.

MATERIALS AND METHODS

Enzyme parameters

Computational parameters including enzyme k_{cat} (*i.e.* turnover) numbers, enzyme surface area, cellular water content, and cell density for *E. coli* were estimated using literature reviews. Parameter values and references are provided in supplementary data S4. k_{cat} numbers were temperature corrected to 37°C using a Q_{10} number of 2 (48). Analyses assumed all membrane-associated enzymes were saturated ($\varepsilon_i = 1$) during batch growth, as typical concentrations of medium components during batch growth are ~2 orders of magnitude greater than the average K_M values of *E. coli* enzymes, ~0.1 mM. During simulation of MG1655 chemostat growth, the saturation parameter (ε_i) for glucose transport (PtsG) was adjusted based on dilution rate.

Metabolic model

The *E. coli* metabolic model was based on published models (37, 38). Every model reaction was balanced for atoms and electrons. The biomass reaction was constructed using theory developed by Neidhardt *et al.,* as described previously (37, 64). The biomass macromolecular composition on a dry mass basis was 78% protein, 10% RNA, 6% DNA, and 6% lipid and polysaccharide (64), the elemental composition was $C_1H_{1.96}O_{0.52}N_{0.28}P_{0.03}$, and the degree of reduction was 4.24 oxidizable electrons per carbon mole of biomass (NH_3 basis). The biomass ATP demand for monomer synthesis and macromolecule polymerization was 36.9 mol ATP per kg dry biomass (37). The biomass reaction did not account for any maintenance energy requirements as those fluxes were calculated separately. The maximum potential for maintenance energy fluxes (q_{ATP}) was calculated from experimental data by constraining exchange fluxes to consensus data and maximizing flux through the ATP hydrolysis reaction (reaction identifier: *ATPm*). Thus, the calculated maintenance energy fluxes account for both growth- and nongrowth-associated ATP (GAM, NGAM). The same *E. coli* model was used for both *E. coli* strain MG1655 and NCM3722 simulations while applying strain-appropriate SA:V ratios and consensus data (supplementary data S1, S2, S4, S5). Unless otherwise noted, enzyme saturation (ε_i), enzyme over capacity factor (ω_i), and allosteric regulation based on membrane properties (σ_i) were set to 1.

The metabolic model is provided in supplementary data S5. All flux balance analysis simulations used COBRA Toolbox (39) with Gurobi as the solver. FBA is a linear algebra based method of interrogating stoichiometric metabolic models (65). By assuming a (quasi) steady state, FBA

identifies a network flux distribution that optimizes an objective function while applying user defined constraints such as bounds on the flux permitted through particular reactions (65). The FBA linear optimization problem is often formalized as:

$$\text{maximize } Z = c^T v$$

$$\text{subject to } Sv = 0 \; and \; lb_i \leq v_i \leq ub_i$$

where *S* is the stoichiometric matrix and defines the stoichiometry of all metabolites in all reactions, *v* is the vector of reaction fluxes, and *c* is the vector specifying the objective function *Z* such as the flux toward maintenance energy. Fluxes through each reaction v_i were constrained by assigned lower (lb_i) and upper flux (ub_i) bounds (65).

A summary of the code used in the current study can be found in supplementary data S9. Additional *sMSAc-FBA* analyses and code can be found at (66).

Data Availability: All MATLAB code, models, and supplementary data can be found at github.com/rosspcarlson/carlson2025_sMSAcFBA or DOI: 10.5281/zenodo.16876885 .

Author Contributions: R.P.C. conceived project, performed theory synthesis, developed code, analyzed data, wrote document, T.G., M.G.B., W.R.H., and R.M. analyzed data, developed interpretation, read and edited document, A.E.B. conceived project, developed code, analyzed data, read and edited document.

Competing Interest Statement: The authors have no competing interests.

ACKNOWLEDGEMENTS

R.P.C. would like to acknowledge Kristopher Hunt, Thomas Cronimus, Terrance Hwa, Adrienne Arnold, and Campbell Putnam for insightful discussions and editing suggestions.

R.P.C. was supported in part by NIH award U01EB019416 and AFOSR award FA9550-23-1-0589. A.E.B. was supported in part by USDA National Institute of Food and Agriculture grant no. 2021-67013-34537. T.G. was supported in part by NSF award DMS 1361240

Figure 1. *E. coli* membrane and enzyme properties. A) The inner membrane enables critical cellular functions including substrate transport and oxidative phosphorylation. Membrane surface area is finite. **B)** The physical dimensions of membrane-associated enzymes influence both two- and three-dimensional molecular crowding. ATP synthase image modified from Wiki Commons. **C)** Membrane surface area requirements and enzyme volume for 22 central metabolism enzymes do not correlate well. Data available in supplementary data S1 and S4.

Figure 2. *E. coli* K-12 strains MG1655 and NCM3722 have distinct phenotypes and geometries. A) Comparison of strains MG1655 and NCM3722 based on maximum specific growth rate on glucose minimal medium, growth rate at which acetate overflow initiates on glucose minimal medium, and biomass yield on glucose (g biomass (g glucose)$^{-1}$). **B)** Experimental surface area to volume (SA:V) ratios for *E. coli* K-12 strains MG1655 and NCM3722 as a function of specific growth rate. *E. coli* geometry data from Si *et al.* 2017 (2). **C)** Graphical representation of differences in cell geometry (cell diameter and length) as a function of *E. coli* strain (blue = MG1655, red = NCM3722) and specific growth rate.

Figure 3. Dynamic membrane proteome of *E. coli* MG1655 as a function of growth rate and environment. A) The capacity of the *E. coli* MG1655 inner membrane to host proteins, expressed here as nm^2 occupied by central metabolism enzymes per volume (μm^{-3}). Data from *E. coli* cultures grown on glucose salts medium. Proteomics data from (33-36). Enzyme properties and calculations can be found in supplementary data S1, S3, and S4. Analysis was limited to cultures grown on glucose minimal data. Calculations were not possible for complex medium because the consumption rates of the numerous substrates were unknown. **B)** Glucose transporter (PtsG) copy number per cell volume as a function of growth rate and growth medium. Gray data points are from glucose negative medium. Proteomics data from (33-36). Supplementary data S3. **C)** Alternative, high affinity sugar transporter copy number per cell volume as a function of growth rate during chemostat growth on glucose containing medium. The specific substrate for each transporter was not present in the medium. Proteomics data from (33-36). Supplementary data S3.

Figure 4. Predictive growth rate – biomass yield phenotype space for *E. coli* strain MG1655 constrained by membrane surface area and membrane protein crowding. **A)** Maintenance

energy flux (mmol ATP (g cdw h)$^{-1}$) shown by the color gradient (protein crowding, f_{MSA} = 0.07). Black contour lines are maintenance energy isoclines. Maintenance energy flux accounts for ATP in excess of ATP used for macromolecule polymerization. **B)** Color gradient represents the P/O number which quantifies the efficiency of oxidative phosphorylation. **C)** Color gradient quantifies the areal density of glucose transporter PtsG (complexes μm^{-2}). **D)** Color gradient quantifies areal density of the electron transport chain component Nuo (complexes μm^{-2}). **E)** Color gradient quantifies areal density of the electron transport chain component Ndh II (complexes μm^{-2}). **F)** Color gradient quantifies overflow metabolism $q_{acetate}$ (mmol acetate (g cdw h)$^{-1}$).

Figure 5. *E. coli* MG1655 phenotypic properties constrained by membrane surface area and protein crowding. A) Experimental data from *E. coli* K-12 strains MG1655 used to build consensus fluxes as a function of growth rate. Experimental data references can be found in supplementary material S1. **B)** Pseudocode of flux balance analysis algorithm used to predict phenotypes constrained by *sMSA* theory and experimental consensus fluxes. **C)** An example of the nexus of experimental consensus fluxes and *sMSA* theory defines three regions based on the slope between the capacity to support biomass fluxes and maintenance energy fluxes. The negative slope defines a Pareto tradeoff. **D)** Maintenance energy potential of the *E. coli* MG1655 consensus fluxes as a function of membrane protein crowding quantified by f_{MSA}. A series of simulations were performed enforcing different membrane protein crowding (f_{MSA} = 0.02-0.09, or 2-9% of occupied membrane area). Error bars represent standard deviation of 100 simulations; see Figure 5B and main text. A Pareto tradeoff between growth rate and maintenance energy flux occurs at the experimental μ_{max} for a f_{MSA} of 0.07. Shaded area quantifies experimental maximum growth rate and standard deviation. **E)** Predicted *E. coli* MG1655 P/O numbers (solid blue points) plotted with experimentally determined P/O numbers (striped points) as a function of growth rate. Predicted P/O numbers have an inflection point denoting a change in scarcity of the membrane surface area. Overflow metabolism occurs at severe surface area scarcity, right of inflection point. The P/O number inflection corresponds with higher experimental biomass yields on glucose. Biomass yield maximum highlighted with (*). Error bars represent standard deviation of 100 simulations using strain-specific consensus fluxes. Experimental proteomics data from (33-36). **F)** Experimental testing of *sMSA* theory using

an MG1655 strain engineered for titratable MreB expression to control cell SA:V ratios. Increasing strain SA:V ratio resulted in faster growth rates on glucose medium supplemented with amino acids (gray data). The total volume (purple data) of the MG1655 strain expressing titratable levels of MreB decreased with growth rate. Data from (2).

Figure 6. Testing *sMSA* predictions using data from *E. coli* NCM3722. A) Experimental data from *E. coli* K-12 NCM3722 used to build consensus fluxes as a function of growth rate. Experimental data references can be found in supplementary material S1. **B**) Relationship between glucose transporter (PtsG) saturation (dotted lines) and as a function of growth rate for wild type MG1655 and recombinant NCM3722. Steady state glucose concentration vs. growth rate for wild type MG1655 and representative batch growth glucose concentration for recombinant NCM3722 cultures. NCM3722 cultures started at approximately 11 mM (off the chart). **C)** Maintenance energy potential of the *E. coli* NCM3722 consensus fluxes as a function of membrane protein crowding quantified by f_{MSA}. A series of simulations were performed enforcing different membrane protein crowding (f_{MSA} = 0.02-0.08, or 2-8% of occupied membrane area). Error bars represent standard deviation of 100 simulations; see Figure 5B and main text. A Pareto tradeoff between growth rate and maintenance energy flux occurs at the experimental μ_{max} for a f_{MSA} of 0.07. Shaded area quantifies experimental maximum growth rate and standard deviation. **D)** Effect of applied SA:V ratios on the relationship between biomass synthesis and maintenance energy potential. Simulations fixed f_{MSA} at 0.07 and varied the SA:V from the experimentally determined relationship (±25% of the experimental values and the MG1655 SA:V relationship). Changing the applied SA:V changed the predicted μ_{max} . Data points represent the mean of 100 simulations. **E)** Predicted P/O numbers as a function of varying SA:V rations and growth rate plotted with experimentally determined P/O number (supplementary data S7). Predicted P/O numbers have an inflection point denoting the change in scarcity of the membrane surface area. Overflow metabolism occurs at severe surface area scarcity, right of inflection point. The P/O number inflection corresponds with higher experimental biomass yields on glucose. Biomass yield maximum highlighted with (*). Experimental proteomics data from (46). **F)** Summary of the predicted μ_{max} and overflow initiation point for *E. coli* NCM3722 when the SA:V ratios were varied. Values determined from Figure 6D and 6E respectively.

Figure 7. Related sMSA theory topics and implications. A) Quantitative representation of 24% and 50% of the membrane surface area being occupied by proteins. Areas are representative of bacteria *E. coli* and eukaryotic mitochondria inner membranes. **B)** Illustration of MG1655 plotted in a biomass synthesis rate, biomass yield on glucose, and maintenance energy generation space. Experimental data suggests *E. coli* maximizes biomass yield for a specific maintenance energy requirement. Reported μ_{max} is a relative value as decreasing the requirement for maintenance energy could increase the biomass synthesis rate (growth rate). **C)** Perturbation analysis of enzyme parameters and their effect on the potential for maintenance energy (q_{ATP}) production. Experimental fluxes from batch growth of *E. coli* MG1655 on glucose salts media with f_{MSA} = 0.07. Enzyme parameters were independently perturbed within ±95% (abscissa) while experimental fluxes were simultaneously and independently perturbed within ± one standard deviation. Each point is the mean of 100 simulations, and error bars represent one standard deviation.

Summary of Supplementary Materials

Files available at: github.com/rosspcarlson/carlson2025_sMSAcFBA

Supplementary Data S1: Literature survey of extracellular flux data for *E. coli* K-12 strains MG1655 and NCM3722 including consensus data sets. Comparison of enzyme volume and membrane surface area requirements for 22 central metabolism enzymes.

Supplementary Data S2: Literature values for cell geometry as a function of growth rate for *E. coli* K-12 strains MG1655 and NCM3722.

Supplementary Data S3: *E. coli* K-12 membrane proteomics data from three independent studies collated by Belliveau et al. 2021.

Supplementary Data S4: Literature review of enzyme and biophysical parameters used in study.

Supplementary Data S5: *In silico*, stoichiometric model and associated data files.

Supplementary Data S6: Summary of simulation output for *E. coli* K-12 strains MG1655 and NCM3722.

Supplementary Data S7: Summary of experimental and in silico P/O number calculations.

Supplementary Data S8: Analysis of maintenance energy fluxes as a function of growth rate, membrane surface area, protein hosting capacity, and P/O number. Perturbation analysis of enzyme *MRE* parameters.

Supplementary Data S9: Summary of MATLAB code used for each figure in study.

References

1. Harris LK & Theriot JA (2018) Surface Area to Volume Ratio: A Natural Variable for Bacterial Morphogenesis. *Trends Microbiol* 26(10):815-832.
2. Si F*, et al.* (2017) Invariance of Initiation Mass and Predictability of Cell Size in Escherichia coli. *Curr Biol* 27(9):1278-1287.
3. Shi H, Bratton BP, Gitai Z, & Huang KC (2018) How to Build a Bacterial Cell: MreB as the Foreman of E. coli Construction. *Cell* 172(6):1294-1305.
4. Lowe M, Kalacheva M, Boersma AJ, & Kedrov A (2020) The more the merrier: effects of macromolecular crowding on the structure and dynamics of biological membranes. *Febs J* 287(23):5039-5067.
5. Beg QK*, et al.* (2007) Intracellular crowding defines the mode and sequence of substrate uptake by Escherichia coli and constrains its metabolic activity. *Proceedings of the National Academy of Sciences of the United States of America* 104(31):12663-12668.
6. Basan M*, et al.* (2015) Overflow metabolism in Escherichia coli results from efficient proteome allocation. *Nature* 528(7580):99-104.
7. Vazquez A & Oltvai ZN (2016) Macromolecular crowding explains overflow metabolism in cells. *Sci Rep* 6:31007.
8. van den Berg J, Boersma AJ, & Poolman B (2017) Microorganisms maintain crowding homeostasis. *Nature reviews. Microbiology* 15(5):309-318.
9. Ellis RJ (2001) Macromolecular crowding: obvious but underappreciated. *Trends Biochem Sci* 26(10):597-604.
10. O'Brien EJ, Lerman JA, Chang RL, Hyduke DR, & Palsson BO (2013) Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Molecular systems biology* 9:693.
11. Carlson RP (2007) Metabolic systems cost-benefit analysis for interpreting network structure and regulation. *Bioinformatics* 23(10):1258-1264.
12. Lynch M & Marinov GK (2017) Membranes, energetics, and evolution across the prokaryote-eukaryote divide. *Elife* 6.
13. Phillips R & Milo R (2009) A feeling for the numbers in biology. *Proceedings of the National Academy of Sciences of the United States of America* 106(51):21465-21471.
14. Bruggeman FJ, Planque R, Molenaar D, & Teusink B (2020) Searching for principles of microbial physiology. *FEMS Microbiol Rev* 44(6):821-844.
15. Szenk M, Dill KA, & de Graff AMR (2017) Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the Membrane Real Estate Hypothesis. *Cell Syst* 5(2):95-104.
16. Zhuang K, Vemuri GN, & Mahadevan R (2011) Economics of membrane occupancy and respiro-fermentation. *Molecular systems biology* 7:500.
17. Papanastasiou M*, et al.* (2013) The Escherichia coli peripheral inner membrane proteome. *Mol Cell Proteomics* 12(3):599-610.
18. Guigas G & Weiss M (2016) Effects of protein crowding on membrane systems. *Biochim Biophys Acta* 1858(10):2441-2450.
19. Hellingwerf K*, et al.* (1982) Energetics of microbial growth: an analysis of the relationship between growth and its mechanistic basis by mosaic non-equilibrium thermodynamics. *FEMS Microbiol. Lett* 15(1):7-17.
20. Pirt SJ (1965) The maintenance energy of bacteria in growing cultures. *Proc R Soc Lond B Biol Sci* 163(991):224-231.

21. van Bodegom P (2007) Microbial maintenance: a critical review on its quantification. *Microbial ecology* 53(4):513-523.
22. Carlson R & Srienc F (2004) Fundamental Escherichia coli biochemical pathways for biomass and energy production: creation of overall flux states. *Biotechnol Bioeng* 86(2):149-162.
23. Hoehler TM & Jorgensen BB (2013) Microbial life under extreme energy limitation. *Nature reviews. Microbiology* 11(2):83-94.
24. Goelzer A & Fromion V (2011) Bacterial growth rate reflects a bottleneck in resource allocation. *Biochim Biophys Acta* 1810(10):978-988.
25. Sueki A, Stein F, Savitski MM, Selkrig J, & Typas A (2020) Systematic Localization of Escherichia coli Membrane Proteins. *mSystems* 5(2).
26. Budin I*, et al.* (2018) Viscous control of cellular respiration by membrane lipid composition. *Science* 362(6419):1186-1189.
27. Sun J, Rutherford ST, Silhavy TJ, & Huang KC (2022) Physical properties of the bacterial outer membrane. *Nature reviews. Microbiology* 20(4):236-248.
28. Dupuy AD & Engelman DM (2008) Protein area occupancy at the center of the red blood cell membrane. *Proceedings of the National Academy of Sciences of the United States of America* 105(8):2848-2852.
29. Lyons E, Freeling M, Kustu S, & Inwood W (2011) Using genomic sequencing for classical genetics in E. coli K12. *Plos One* 6(2):e16717.
30. Volkmer B & Heinemann M (2011) Condition-dependent cell volume and concentration of Escherichia coli to facilitate data conversion for systems biology modeling. *Plos One* 6(7):e23126.
31. Oldewurtel ER, Kitahara Y, & van Teeffelen S (2021) Robust surface-to-mass coupling and turgor-dependent cell width determine bacterial dry-mass density. *Proceedings of the National Academy of Sciences of the United States of America* 118(32).
32. Lendenmann U & Egli T (1998) Kinetic models for the growth ofEscherichia coli with mixtures of sugars under carbon-limited conditions. *Biotechnol Bioeng* 59(1):99-107.
33. Belliveau NM*, et al.* (2021) Fundamental limits on the rate of bacterial growth and their influence on proteomic composition. *Cell Syst* 12(9):924-944 e922.
34. Schmidt A*, et al.* (2016) The quantitative and condition-dependent Escherichia coli proteome. *Nat Biotechnol* 34(1):104-110.
35. Valgepea K, Adamberg K, Seiman A, & Vilu R (2013) Escherichia coli achieves faster growth by increasing catalytic and translation rates of proteins. *Mol Biosyst* 9(9):2344-2358.
36. Peebo K*, et al.* (2015) Proteome reallocation in Escherichia coli with increasing specific growth rate. *Mol Biosyst* 11(4):1184-1193.
37. Carlson RP (2009) Decomposition of complex microbial behaviors into resource-based stress responses. *Bioinformatics* 25(1):90-97.
38. Schuetz R, Zamboni N, Zampieri M, Heinemann M, & Sauer U (2012) Multidimensional optimality of microbial metabolism. *Science (New York, N.Y.)* 336(6081):601-604.
39. Heirendt L*, et al.* (2019) Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat Protoc* 14(3):639-702.
40. Mori M, Hwa T, Martin OC, De Martino A, & Marinari E (2016) Constrained Allocation Flux Balance Analysis. *PLoS computational biology* 12(6):e1004913.

41. Baldazzi V, Ropers D, Gouze JL, Gedeon T, & de Jong H (2023) Resource allocation accounts for the large variability of rate-yield phenotypes across bacterial strains. *Elife* 12.
42. Grauel A*, et al.* (2021) Structure of Escherichia coli cytochrome bd-II type oxidase with bound aurachin D. *Nat Commun* 12(1):6498.
43. Borisov VB*, et al.* (2011) Aerobic respiratory chain of Escherichia coli is not allowed to work in fully uncoupled mode. *Proceedings of the National Academy of Sciences of the United States of America* 108(42):17320-17324.
44. Anraku Y (1988) Bacterial electron transport chains. *Annual review of biochemistry* 57(1):101-132.
45. Folsom JP & Carlson RP (2015) Physiological, biomass elemental composition and proteomic analyses of Escherichia coli ammonium-limited chemostat growth, and comparison with iron- and glucose-limited chemostat growth. *Microbiology* 161(8):1659-1670.
46. Mori M*, et al.* (2021) From coarse to fine: the absolute Escherichia coli proteome under diverse growth conditions. *Molecular systems biology* 17(5):e9536.
47. Beck A, Hunt K, Bernstein HC, & Carlson R (2016) Interpreting and designing microbial communities for bioprocess applications, from components to interactions to emergent properties. *Biotechnology for Biofuel Production and Optimization*, eds Eckert CA & Trinh CT (Elsevier, Amsterdam), pp 407-432.
48. Elias M, Wieczorek G, Rosenne S, & Tawfik DS (2014) The universality of enzymatic rate-temperature dependency. *Trends Biochem Sci* 39(1):1-7.
49. Taymaz-Nikerel H, Borujeni AE, Verheijen PJ, Heijnen JJ, & van Gulik WM (2010) Genome-derived minimal metabolic models for Escherichia coli MG1655 with estimated in vivo respiratory ATP stoichiometry. *Biotechnol Bioeng* 107(2):369-381.
50. Westerhoff HV, Hellingwerf KJ, & Van Dam K (1983) Thermodynamic efficiency of microbial growth is low but optimal for maximal growth rate. *Proceedings of the National Academy of Sciences of the United States of America* 80(1):305-309.
51. Werner S, Diekert G, & Schuster S (2010) Revisiting the thermodynamic theory of optimal ATP stoichiometries by analysis of various ATP-producing metabolic pathways. *J Mol Evol* 71(5-6):346-355.
52. Ojkic N, Serbanescu D, & Banerjee S (2019) Surface-to-volume scaling and aspect ratio preservation in rod-shaped bacteria. *Elife* 8.
53. Young KD (2006) The selective value of bacterial shape. *Microbiology and molecular biology reviews : MMBR* 70(3):660-703.
54. Mori M, Cheng C, Taylor BR, Okano H, & Hwa T (2023) Functional decomposition of metabolism allows a system-level quantification of fluxes and protein allocation towards specific metabolic functions. *Nature Communications* 14(1):4161.
55. Schlame M (2021) Protein crowding in the inner mitochondrial membrane. *Biochim Biophys Acta Bioenerg* 1862(1):148305.
56. Andersen OS & Koeppe RE, 2nd (2007) Bilayer thickness and membrane protein function: an energetic perspective. *Annu Rev Biophys Biomol Struct* 36:107-130.
57. LaCroix RA*, et al.* (2015) Use of adaptive laboratory evolution to discover key mutations enabling rapid growth of Escherichia coli K-12 MG1655 on glucose minimal medium. *Appl Environ Microbiol* 81(1):17-30.

58. Wang G, *et al.* (2023) Recent progress in adaptive laboratory evolution of industrial microorganisms. *J Ind Microbiol Biotechnol* 50(1).
59. Chen Y & Nielsen J (2021) In vitro turnover numbers do not reflect in vivo activities of yeast enzymes. *Proceedings of the National Academy of Sciences of the United States of America* 118(32).
60. Carlson RP & Taffs RL (2010) Molecular-level tradeoffs and metabolic adaptation to simultaneous stressors. *Curr Opin Biotech* 21(5):670-676.
61. El-Mansi M (2023) Control of central metabolism's architecture in Escherichia coli: An overview. *Microbiol Res* 266:127224.
62. Fischer E & Sauer U (2003) A novel metabolic cycle catalyzes glucose oxidation and anaplerosis in hungry *Escherichia coli*. *The Journal of Biological Chemistry* 278:46446-46451.
63. Carlson RP, *et al.* (2018) Competitive resource allocation to metabolic pathways contributes to overflow metabolisms and emergent properties in cross-feeding microbial consortia. *Biochemical Society Transactions* 46(2):269-284.
64. Carlson R & Srienc F (2004) Fundamental Escherichia coli biochemical pathways for biomass and energy production: identification of reactions. *Biotechnol Bioeng* 85(1):1-19.
65. Orth JD, Thiele I, & Palsson BØ (2010) What is flux balance analysis? *Nat Biotechnol* 28(3):245-248.
66. Carlson RP, *et al.* (2024) Cell Geometry and Membrane Protein Crowding Constrain Growth Rate, Overflow Metabolism, Respiration, and Maintenance Energy. *bioRxiv*.

Box and Figures

Box 1: The membrane surface area required for a phenotype can be tabulated by considering enzyme fluxes and enzyme properties. Literature reviews of enzyme parameters used in this study can be found in supplementary data S4. The surface area required for enzyme *i*, per g cdw, is calculated as:

$$SA_{Ei} = \left[\frac{v_i \gamma_i}{\tau_i \varepsilon_i}\right]\left[\frac{\omega_i}{\sigma_i}\right]\left[\frac{mol}{10^3\ mmol}\right]\left[N_{Avo}\right]\left[\frac{h}{3600\ s}\right] \qquad \textit{eqn. 1}$$

SA_{Ei} = surface area occupied by enzyme *i*, units: nm^2 (g cdw)$^{-1}$
v_i = flux of reaction *i*, units: mmol (g cdw h)$^{-1}$
γ_i = membrane area occupied by single enzyme *i*, units: nm^2 (enzyme)$^{-1}$
τ_i = enzyme *i* turnover number (*a.k.a.* k_{cat}), units: molecules (s*enzyme)$^{-1}$, where s is seconds
ε_i = enzyme *i* active site saturation (S (S +K_M)$^{-1}$, where S is substrate concentration), unitless (0 - 1)
σ_i = membrane allosteric regulation , unitless (0 - 1)
ω_i = enzyme overcapacity factor , unitless (≥1)
N_{Avo} = Avogadro's number, units: molecules mole^{-1}

Membrane protein crowding is quantified as the total membrane surface area used by central metabolism enzymes. It is determined by summing the contributions from all active enzymes and applying biophysical parameters (reviewed in supplementary data S4). *Eqn. 2* quantifies membrane protein crowding as the fraction of membrane surface area (f_{MSA}) occupied by active central metabolism enzymes. Conversions include, from left to right, water content of live cells, density of live cells, growth rate dependent SA:V ratio of cells (c_1*μ + c_2), and unit conversions. δ_i accounts for nonidealities in surface area usage including misfolded proteins or misassembled protein complexes:

$$f_{MSA} = \sum_i \left(SA_{Ei}\left[\frac{g\ cdw}{3.85\ g\ wet\ cell}\right]\left[\frac{1.1\ g\ wet\ cell}{cm^3\ wet\ cell}\right]\left[\frac{\mu m^3\ wet\ cell}{(c_1 * \mu + c_2)\ \mu m^2}\right]\left[\frac{cm}{10^4\ \mu m}\right]\left[\frac{cm^2}{10^{14}\ nm^2}\right] + \delta_i\right) \qquad \textit{eqn. 2}$$

Eqn. 2 can be arranged to define the surface area resource MRE_i (membrane real estate for enzyme *i*) which is incorporated into the presented metabolic model. θ is a constant comprised of biophysical parameters and conversion factors (note: c_1 has units of h, c_2 is unitless):

$$f_{MSA} - \sum_i \left(\left[\frac{v_i \gamma_i}{\tau_i \varepsilon_i}\right]\left[\frac{\omega_i}{\sigma_i}\right]\left[\frac{\theta}{c_1 * \mu + c_2}\right] + \delta_i\right) - \sum_i \left(v_i MRE_i\left[\frac{\theta}{c_1 * \mu + c_2}\right] + \delta_i\right) \qquad \textit{eqn. 3}$$

with θ = 0.048 (molecules*h*g cdw)(mmol*s*nm^2)$^{-1}$.

Summing δ_i over all the enzymes quantifies the f_{MSA} nonideality offset Δ.

$$(f_{MSA} - \Delta)\left[\frac{c_1 * \mu + c_2}{\theta}\right] - \sum_i v_i MRE_i \qquad \textit{eqn. 4}$$

Which can be arranged as:

$$(f_{MSA} - \Delta)\left[\frac{c_2}{\theta}\right] = \sum_i v_i MRE_i - (f_{MSA} - \Delta)\left[\frac{c_1 * \mu}{\theta}\right] \qquad \textit{eqn. 5}$$

Collecting constants results in the linear constraint accounting for finite surface area:

$$sMSA' - \sum_i v_i MRE_i' - \beta' * \mu \qquad \textit{eqn. 6}$$

Converting the units of eqn. 6 for interpretability, multiplying by: 1 molecule (s*mmol)$^{-1}$:

$$sMSA = \sum_i v_i MRE_i - \beta * \mu \qquad \textit{eqn. 7}$$

sMSA = specific membrane surface area available for enzymes, nm^2 (g cdw)$^{-1}$

MRE_i = membrane area used by active enzyme *i*, s*nm^2 molecules^{-1}

β = growth rate dependent change in surface area due to change in SA:V, h*nm^2 (g cdw)$^{-1}$

The constraint was applied by including an appropriate MRE_i cost to membrane-associated reactions and incorporating the β parameter into the biomass synthesis reaction. Note, the SA:V slope, c_1, is negative making the contribution of $\beta * \mu$ positive in *eqn. 6*. See supplementary data S4 and S5 for parameter values and literature reviews.

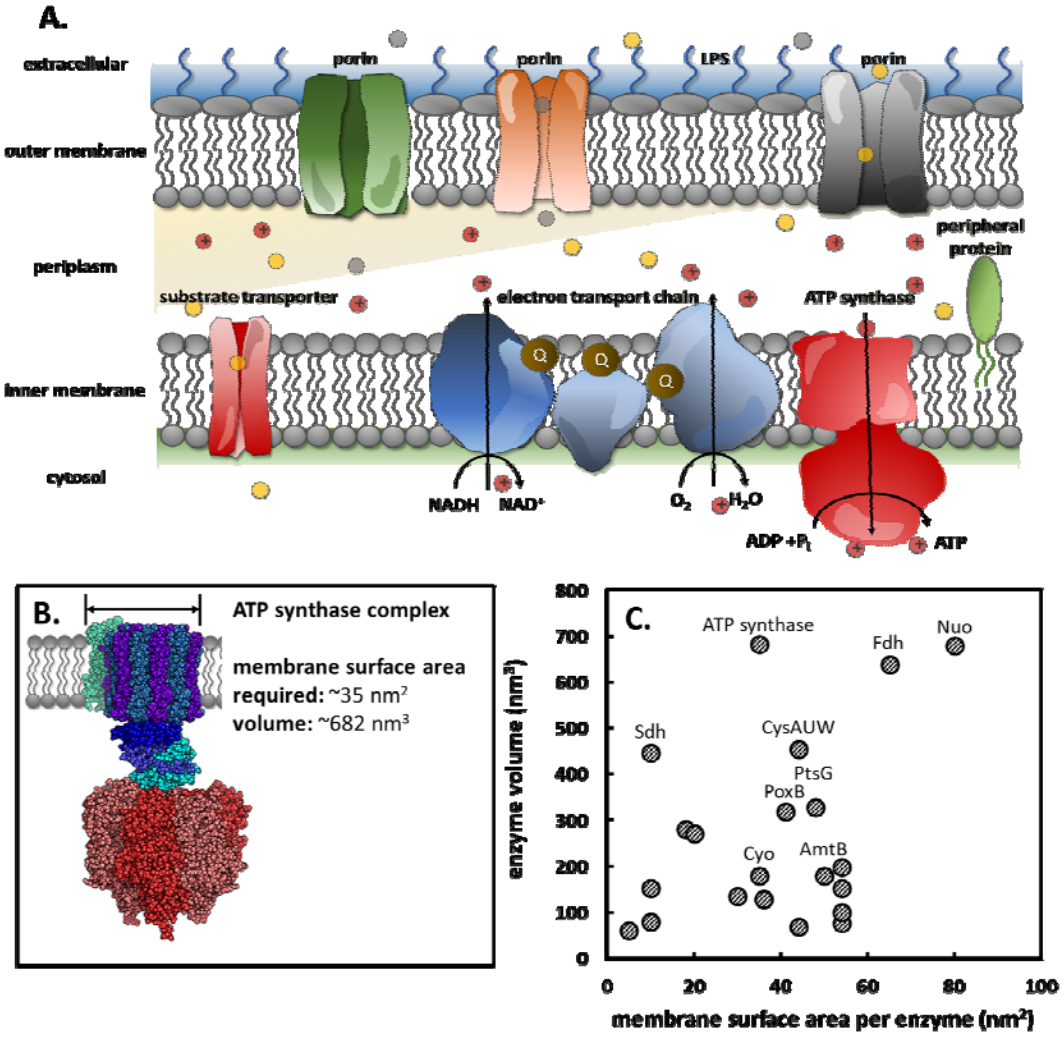

A.
extracellular
porin
porin
LPS
porin
outer membrane
periplasm
peripheral protein
substrate transporter
electron transport chain
ATP synthase
Q
Q
Q
inner membrane
cytosol
NADH
NAD+
O2
H2O
ADP +Pi
ATP
B.
ATP synthase complex
membrane surface area required: ~35 nm2
volume: ~682 nm3
C.
ATP synthase
Fdh
Nuo
Sdh
CysAUW
PtsG
PoxB
Cyo
AmtB
enzyme volume (nm3)
membrane surface area per enzyme (nm2)
0
100
200
300
400
500
600
700
800
0
20
40
60
80
100

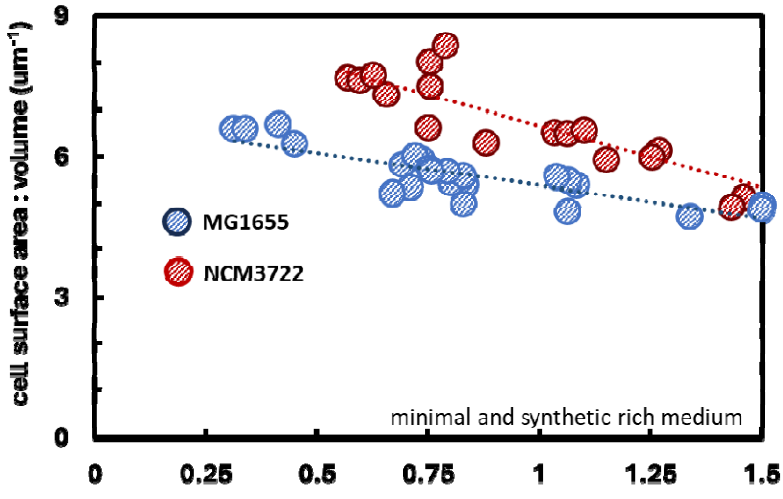
Figure 1

A.

strain	μ_{max} (h^{-1})	overflow (h^{-1})	yield ($g\ g^{-1}$)
MG1655	0.69± 0.02	≥0.4 ± 0.1	0.4 ± 0.02
NCM3722	0.97± 0.06	≥0.75 ± 0.05	0.45 ± 0.04

Experimental data references can be found in supplementary material S1.

B.



C.

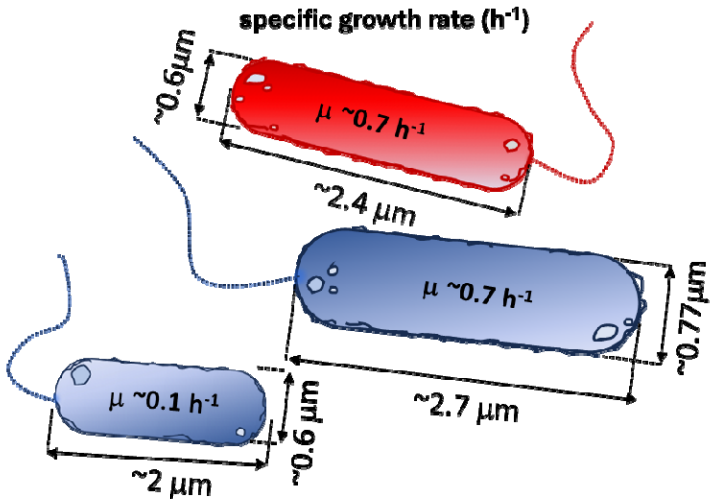


Figure 2

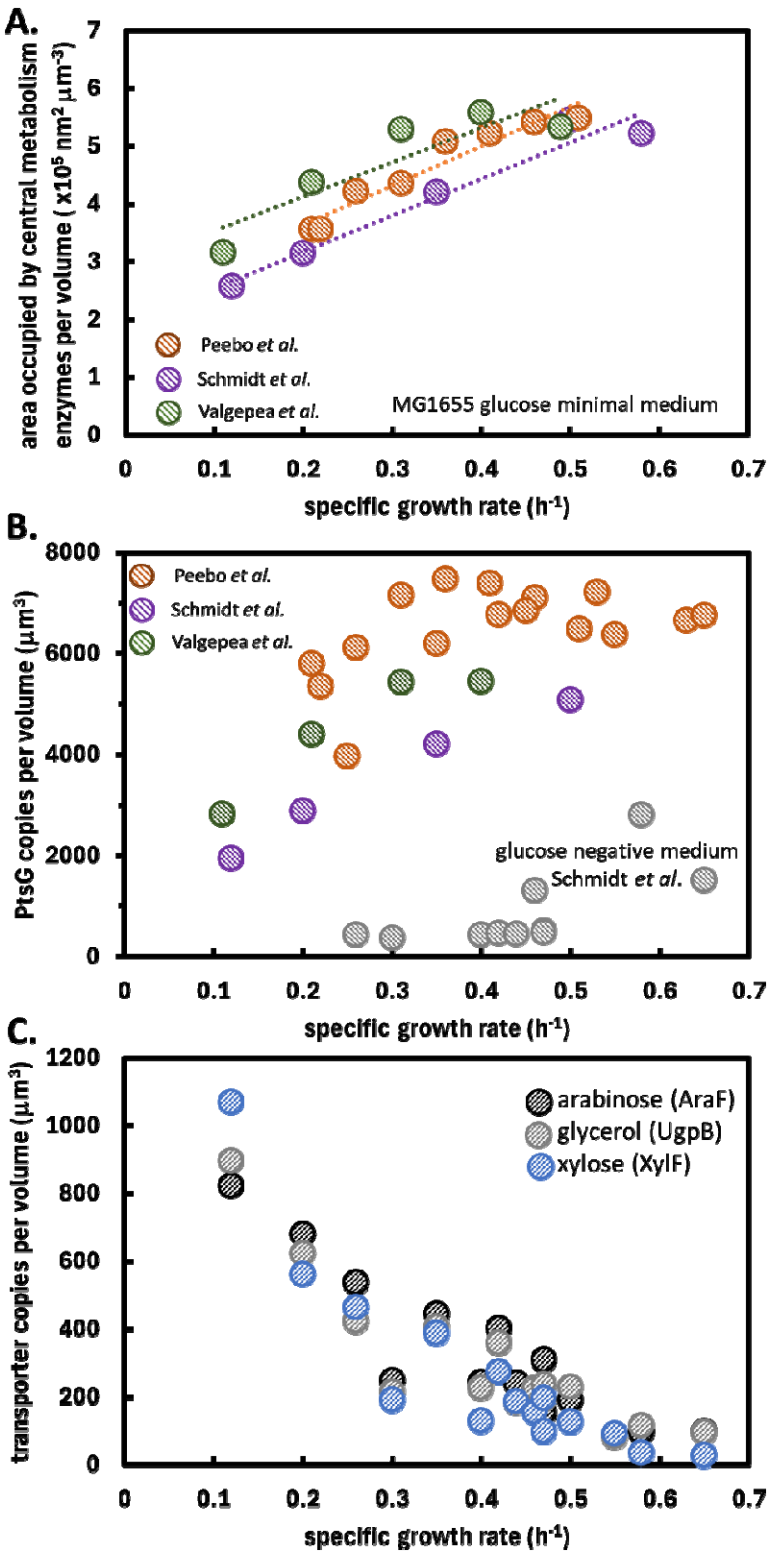

A.
area occupied by central metabolism enzymes per volume (x10^5 nm^2 µm^-3)
Peebo et al.
Schmidt et al.
Valgepea et al.
MG1655 glucose minimal medium
specific growth rate (h^-1)
B.
PtsG copies per volume (µm^3)
Peebo et al.
Schmidt et al.
Valgepea et al.
glucose negative medium
Schmidt et al.
specific growth rate (h^-1)
C.
transporter copies per volume (µm^3)
arabinose (AraF)
glycerol (UgpB)
xylose (XylF)
specific growth rate (h^-1)

Figure 3

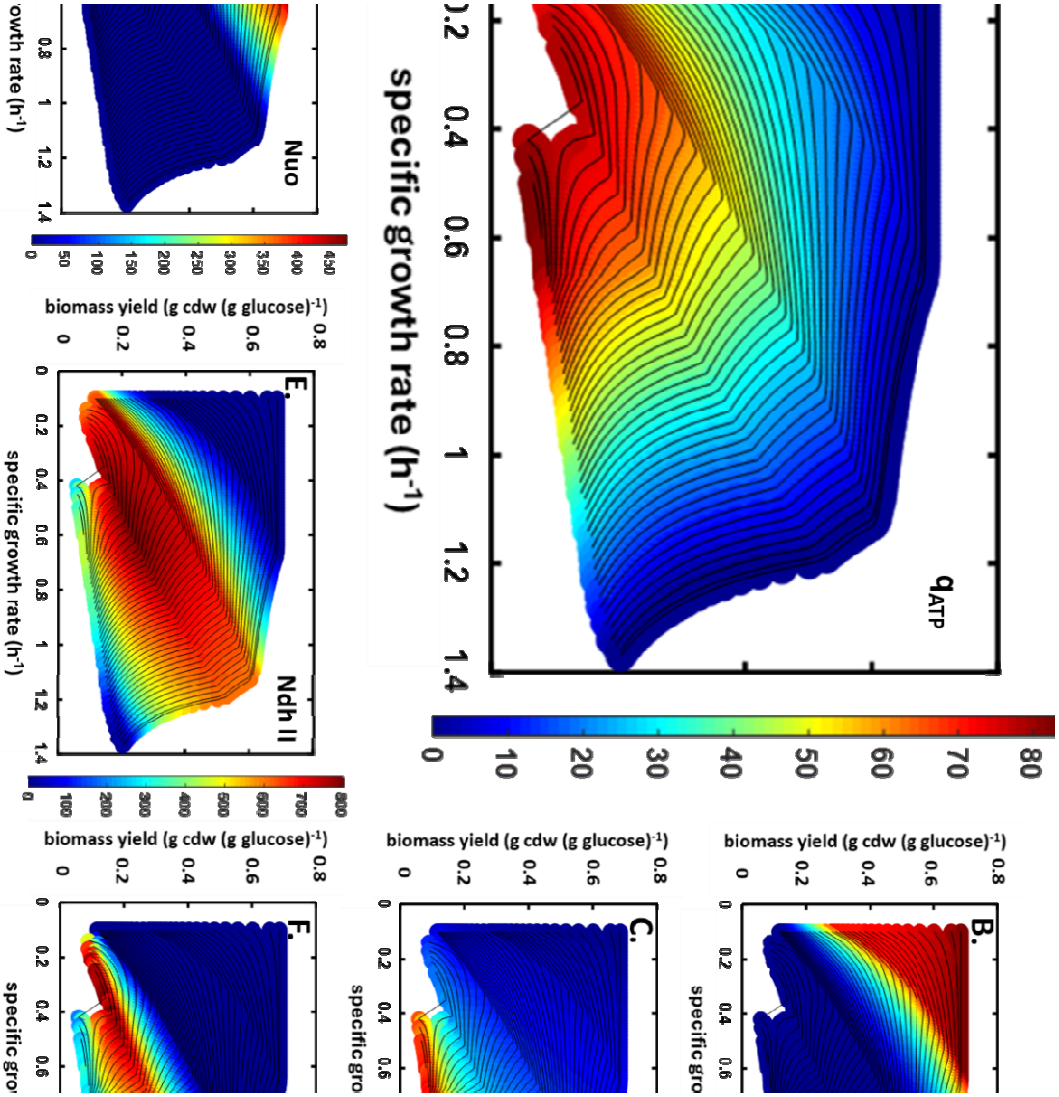

qATP
specific growth rate (h-1)
B.
C.
E.
F.
Nuo
Ndh II
biomass yield (g cdw (g glucose)-1)
specific growth rate (h-1)

Figure 4

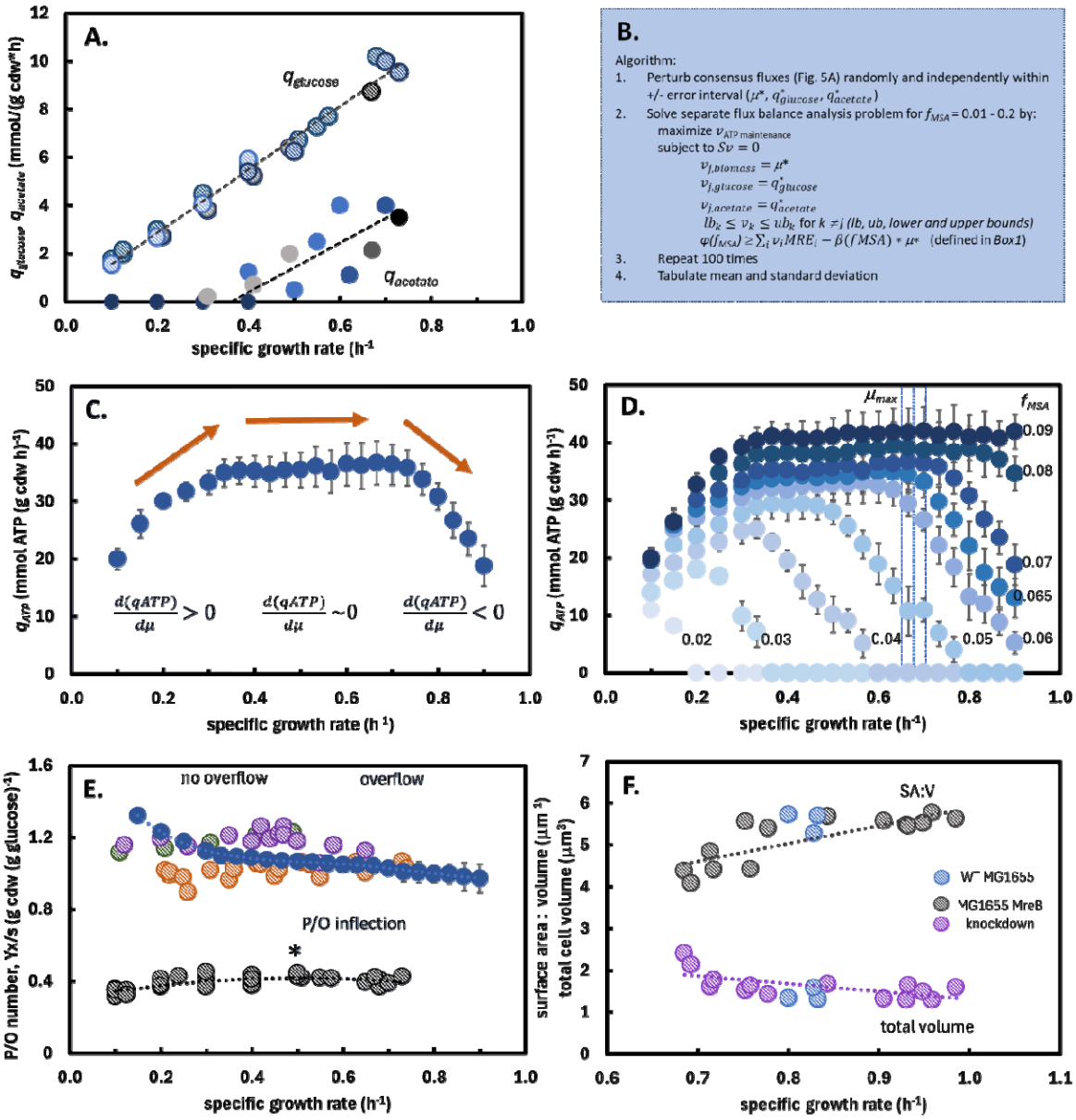
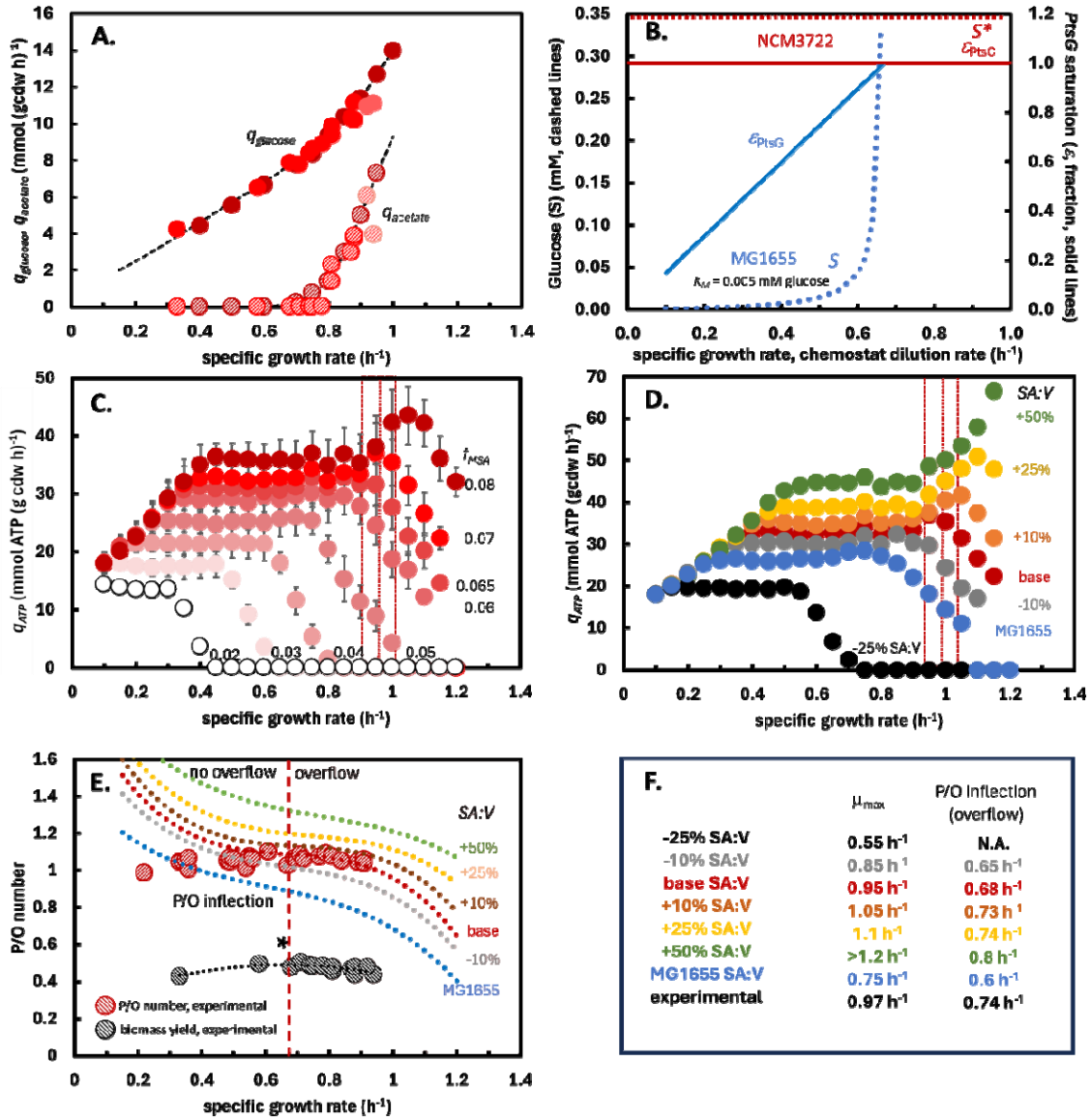

A.
$q_{glucose}$
$q_{acetate}$
$q_{glucose}$, $q_{acetate}$ (mmol/(g cdw*h))
specific growth rate (h-1)
B.
Algorithm:
1. Perturb consensus fluxes (Fig. 5A) randomly and independently within +/- error interval (μ^*, $q^*_{glucose}$, $q^*_{acetate}$)
2. Solve separate flux balance analysis problem for f_{MSA} = 0.01 - 0.2 by:
maximize $v_{ATP\ maintenance}$
subject to $Sv = 0$
$v_{J,biomass} = \mu^*$
$v_{J,glucose} = q^*_{glucose}$
$v_{J,acetate} = q^*_{acetate}$
$lb_k \leq v_k \leq ub_k$ for $k \neq j$ (lb, ub, lower and upper bounds)
$\varphi(J_{MSA}) = \sum_i v_i MRE_i - \beta(f_{MSA}) * \mu^*$ (defined in Box1)
3. Repeat 100 times
4. Tabulate mean and standard deviation
C.
q_{ATP} (mmol ATP (g cdw h)-1)
$\frac{d(qATP)}{d\mu} > 0$
$\frac{d(qATP)}{d\mu} \sim 0$
$\frac{d(qATP)}{d\mu} < 0$
specific growth rate (h-1)
D.
μ_{max}
f_{MSA}
0.09
0.08
0.07
0.065
0.06
0.05
0.04
0.03
0.02
q_{ATP} (mmol ATP (g cdw h)-1)
specific growth rate (h-1)
E.
no overflow
overflow
P/O inflection
*
P/O number, Yx/s (g cdw (g glucose)-1)
specific growth rate (h-1)
F.
SA:V
W- MG1655
MG1655 MreB knockdown
total volume
surface area : volume (µm-1)
total cell volume (µm3)
specific growth rate (h-1)

Figure 5



F.

	μ_{max}	P/O inflection (overflow)
-25% SA:V	0.55 h^{-1}	N.A.
-10% SA:V	0.85 h^{-1}	0.65 h^{-1}
base SA:V	0.95 h^{-1}	0.68 h^{-1}
+10% SA:V	1.05 h^{-1}	0.73 h^{-1}
+25% SA:V	1.1 h^{-1}	0.74 h^{-1}
+50% SA:V	>1.2 h^{-1}	0.8 h^{-1}
MG1655 SA:V	0.75 h^{-1}	0.6 h^{-1}
experimental	0.97 h^{-1}	0.74 h^{-1}

Figure 6

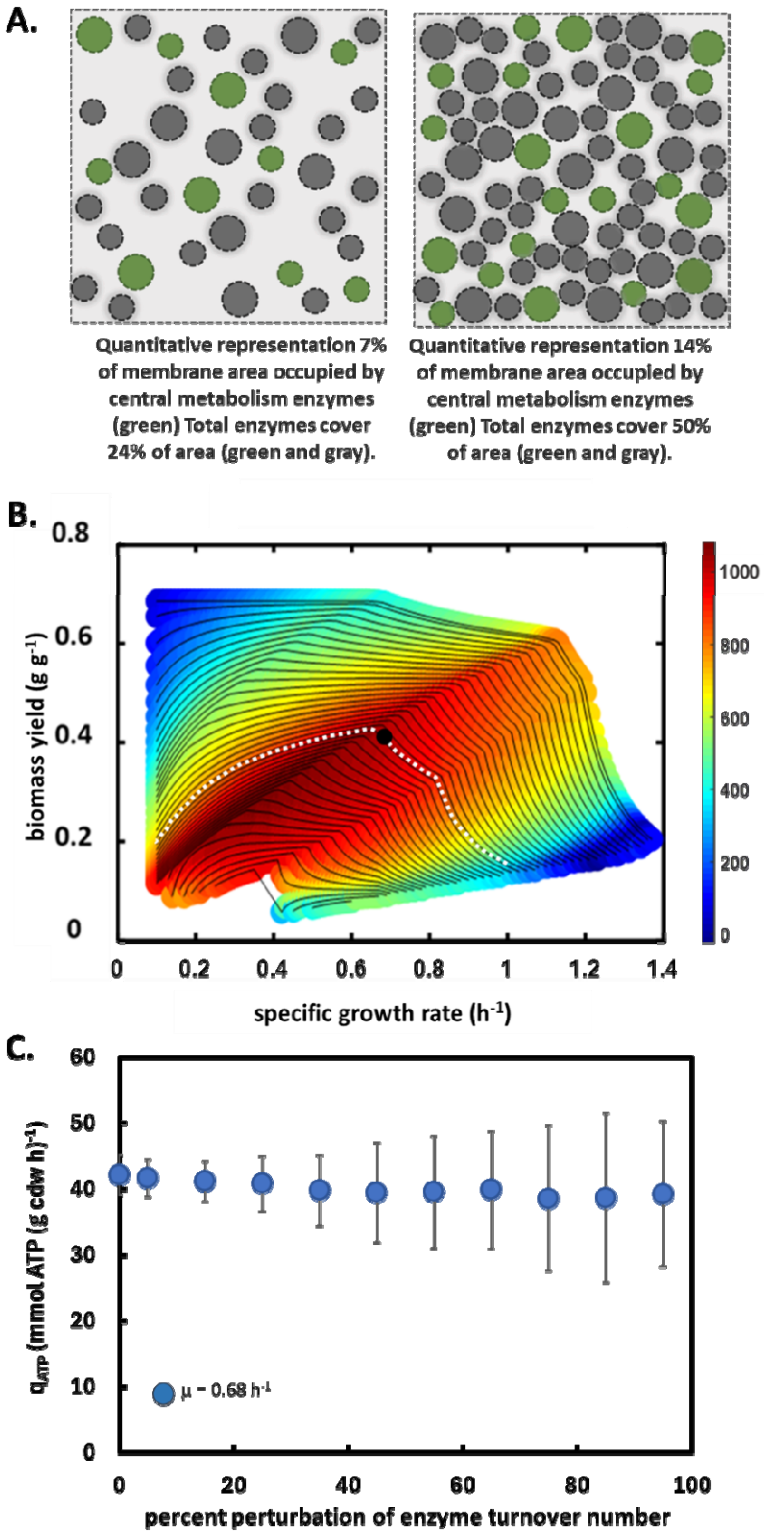

A.
Quantitative representation 7% of membrane area occupied by central metabolism enzymes (green) Total enzymes cover 24% of area (green and gray).
Quantitative representation 14% of membrane area occupied by central metabolism enzymes (green) Total enzymes cover 50% of area (green and gray).
B.
biomass yield (g g-1)
specific growth rate (h-1)
C.
qATP (mmol ATP (g cdw h)-1)
μ = 0.68 h-1
percent perturbation of enzyme turnover number

Figure 7