Goal:

Claims:

1. mRNA synthesis rates is dependent on media.
   1. Support: When mRNA synthesis per genome equivalent of DNA was compared in six cultures (3 minimal and 3 supplemented with all 20 AAs), there were 2 distinct mRNA synthesis rates – one for the minimal media cultures and 1 for the supplemented cultures.
2. For cultures grown in presence of all AAs, independent of growth rate, the rate of ribosome function is near maximal
3. There are about 9 tRNA molecules per ribosome. During slow cell growth in nutritionally poor media, # of tRNA per ribosome is somewhat higher (12 @ µ=0.4 to a constant plateau of ~7 at growth rates faster than µ=1). [**Dong]**
   1. Higher tRNA ratio during slow growth is because 1) instability of newly made rRNA at slow growth rates
      1. During growth in minimal media @ µ=0.67, accumulation of tRNA relative to rRNA was 10-15% higher than during twofold faster growth in glucose minimal media; so possible that 10-15% correction to tRNA:ribosome ratio needed for lower growth rates near 0.6 (i.e., 10-15% more tRNA per ribosome than reported by Bremer?).
   2. Increase in ratio of rRNA to tRNA genes at fast growth rate due to increased chromosome branching associated w/ shorter intevrals btwn chromosome replication. An analysis of location of seven rrn transcription units and 81 different tRNA genes on E. Coli indicate that average rrn unit is closer to the origin of replication, and so dosage of rrn genes increases slightly faster than tRNA genes w/ increased chromosome branching (btwn µ=0.6 and 3.0, rRNA/tRNA is expected to increase by ~15%).
4. At least some of the tRNAs reading the most abundant codons is coordinate with the synthesis of ribosomes, whereas at least some of the tRNAs reading minor codons are more highly expressed at slow growth rates where they would be required for synthesis of proteins w/ higher frequency of minor codons.
   1. The abundance of two leucine and 3 methionine isoacceptor tRNAs relative to # of ribosomes as a function of growth remained relatively constant btwn µ=0.5 to 2.1. In contrast, Leu2,Leu4, and Leu5 (responsible for minor codon reading that are infrequent in highly expressed proteins) tRNAs to ribosomes decrease 5-8 fold as growth rate increased from µ=0.5 to 2.1
5. Since a substantial fraction of cell volume is taken up by nucleoid, the actual space for biochemical reactions within cell is much smaller than the cell volume. Large cenzyems complexes like RNAP and ribosomes might not move freely between DNA strands within the nucleoid.
6. For DNA, ribosomes, and proteins, the rate of synthesis during periods of balanced growth are essentially equal to rates of accumulation since turnover is negligible. For total RNA, however, instantaneous synthesis rate is substantially higher than accumulation rate bcz of instability of mRNA and spacer sequences in primary rRNA and tRNA transcripts.
7. Strength of rRNA and tRNA promoters is lower than strength of many mRNA promoters. Since mRNA promoters become saturated, the stronger mRNA promoters have higher activity than stable RNA promoters during growth in poor media, but @ higher conc. of free RNAP during growth in rich media, stable RNA promoters have higher activity despite lower strength.

**Polypeptide-Chain-Elongation Rate in Escherichia coli B/r as a function of growth rate**

**Goal:**

**Claim**:

1. The peptide-chain-elongation rate is 17 aa/s and independent of the bacterial growth rate in fast-growing bacteria (µ=1.3 and 2 doublings/hr), and significantly decreased to 12aa/s in slowly growing µ=0.67 doublings/hr)

**Evidence:**

**Important assumptions/limitations:**

1. This study tells us about the average translation rate of mRNA encoding proteins of specific sizes. It’s possible that certain proteins are translated much faster/slower than what was found here.
2. Elongation rate estimates are probably accurate to within 15% (due to possible error with MW estimates for each band).
3. The 3s sampling time means that small polypeptides couldn’t be measured since break-time would be missed (<~12k daltons).
4. Less and less nascent protein is made w/ increasing MW, so less distinct breakpoint (so only polypeptides <42kD used)

**Methods:**

E. coli B/r grown at 37 C in supplemented medium C w/ E460 monitored.

A. 0.2% succinate @ pH 6.8 supplemented led to µ=0.67

B. 0.2% glucose supplemented led to µ=1.3

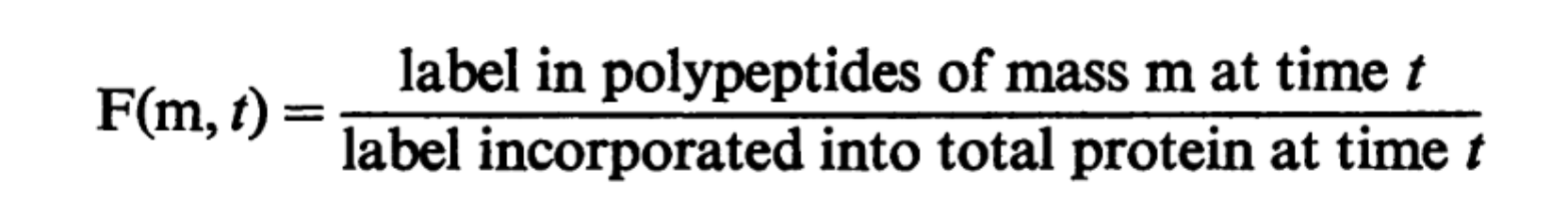
C. 0.2% glucose + 19 aas (not leucine) at initial concentration 20µg/mL led to µ=2.0

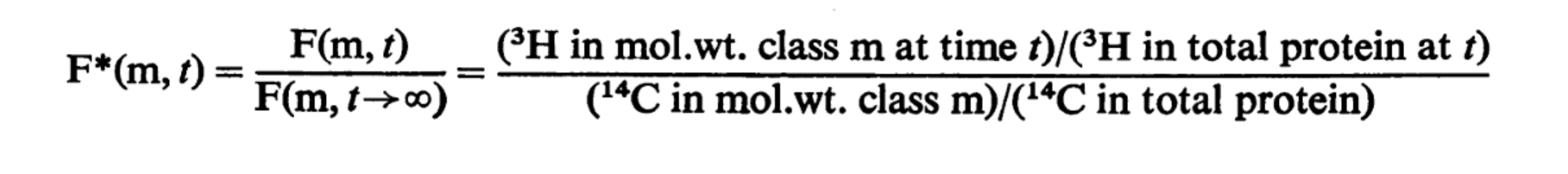
1. At E460=0.2, 15mL of exponential phase culture was labeled with 100µL of Carbon-14 leucine (final leucine conc = 5µM).
2. At E460=0.6, the 15mL culture was placed in auto-sampling device @ 37C, and H-3 leucine was rapidly mixed into the culture to a final conc. of 5µM. Samples (~0.5mL) were taken automatically at 3s intervals for 1 min into test tubes in a CO2/methanol bath so that samples were frozen completely within 1s 🡪 samples stored at -70C
3. Each sample was then thawed at 0C, pelleted, washed, and lysed, and then applied to a polyacrylamide gel, which was electrophoresed for 3-4hr at 25mA. Slabs were fixed and then stained with Coomassie Brilliant Blue & proteins w/ known molecular weight were co-electrophoresed, and then the gel was dried under heat and vacuum.
4. A band was cut in stained border (because 14C is localized within the stained band, but H-3 which was added only at the end is distributed both in the stained bands and in intervening unstained areas, where there are only nascent polypeptides), and then burnt in an oxidizer. 3-H20 was trapped and the radioactive condensates were dissolved. Recoveries of both 3H and 14C were always above 90%.
5. For total incorporation assays, small amount of the lysate (from step 3) of each sample was precipitated, collected on nitrocellulose filter, and dried 🡪 oxidization same as step 4 (just no band needed, since this is overall incorporation)

Calculations:

1. For each given MW, F should initially increase linearly until a time tm equal to the average synthesis time of a polypeptide of MW m (since initially, 0 completed polypeptides had radioactivity, and as proteins are completed –each ribosome having different extents of incomplete protein -- radioactivity goes up linearly until every ribosome has completed a full protein w/ radioactivity incorporated). At this point, the overall ratio of radioactive incorporated proteins to non-radioactive incorporated proteins (F\*) will slowly increase and approach 1.

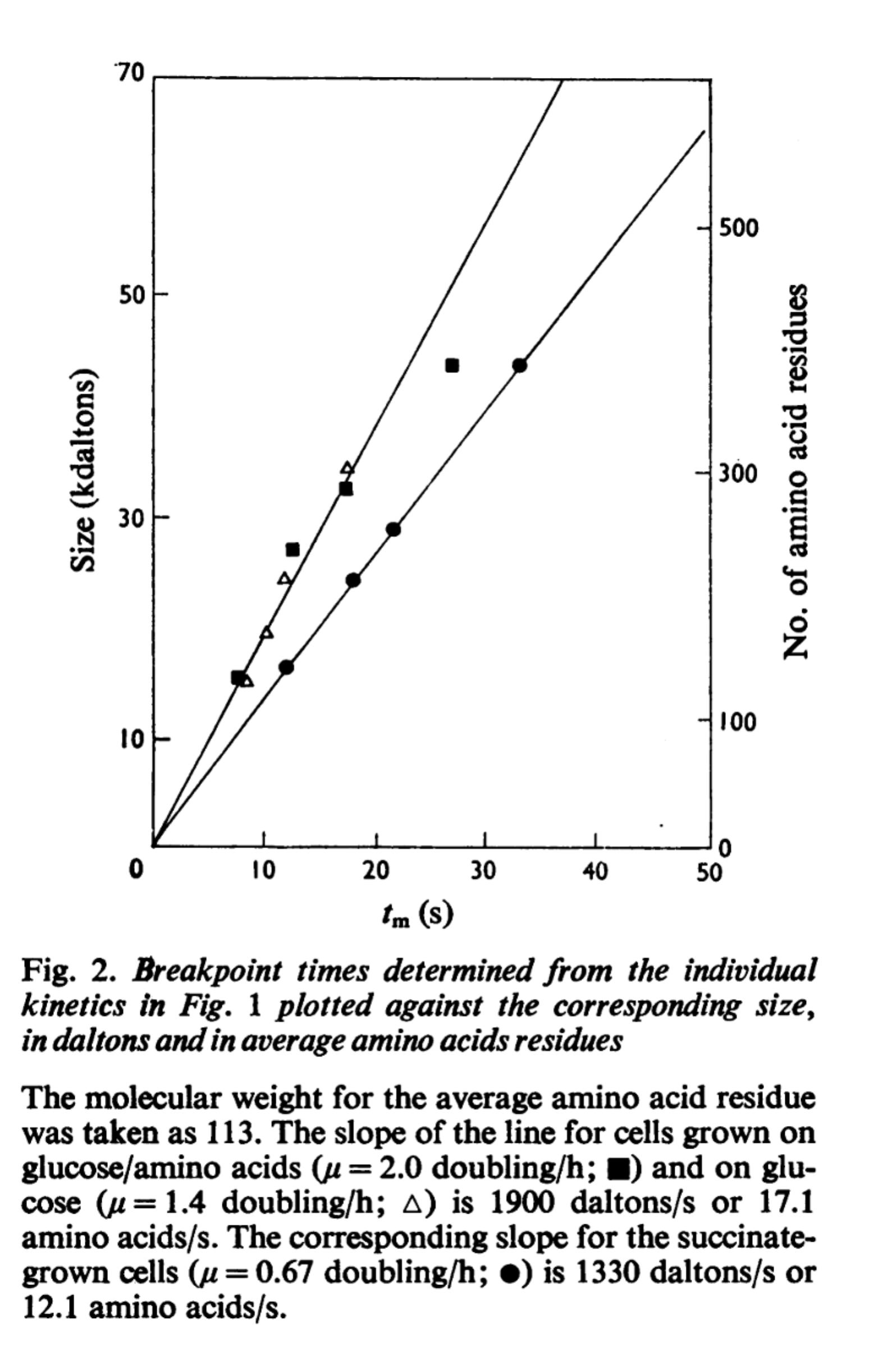
The denominator is a normalization since 3-H c.p.m will be different from 14C c.p.m. By normalization, F will approach 1 over time. 3H in total protein will keep going up nonlinearly since it’s approaching the relative amount of 14C total in protein. Thus, for a sufficiently large MW protein, 3H cpm in the MW class will grow more slowly than 3H c.p.m in total protein.





1. A best-fit line was qualitatively drawn for each F\* of a given MW with the condition that the theoretical increase in F\* is 0.5 between zero time and tm. Then, breakpoint times were plotted for different sizes (for some reason they flipped x and y independent dependent axes). mass of protein/tm gives the average elongation rate assuming all ribosomes are active.
2. Dennis & Bremer 1974 calculated 13.5,13.5, and 9.5 aa/s per 70S ribosome for µ=2.1, 1.4, and 0.67 respectively. Thus, given the elongation rates calculated in this paper, the authors hypothesize that ~80% of ribosomes appear to be active at all growth rates. **Shouldn’t they be dividing cp by B?** I think what the 1974 paper is calculating is aa/s per all 70S ribosomes (inactive and active). So the current paper aa/s can be reached by assuming that only 80% are active.

**Important figure:**



Scheme 1e only parabolic after tm because of normalization; else should be a flat line I believe for F(t).

1. Relevance to project

Include a short description of why this paper matters (in the context of the project) & what your takeaways are.

1. Claims

List the major claims of the paper

1. Support of claims

List the evidence that the authors provide for each of their claims.

1. Similarities & Applications to our system [Not needed?]
2. Parameters

List parameters that the authors vary

1. Methods

Thoroughly enumerate the methodology used in the paper. Details of a particular experimental or computational technique can be included in a footnote or in the glossary if an extended description is warranted.

1. Analysis

Provide a thorough description (including any relevant equations) of how the authors analyzed their results

1. Assumptions & Shortcomings

Include assumptions in paper and any shortcomings that could be improved upon.

1. Important figures & tables

List up to a few core figures & tables

1. Key references

Include any references that play a critical role in the paper’s arguments

Aponte-Rivera & Zia 2018

1. Relevance to project

This paper

1. Claims

List the major claims of the paper

1. Support of claims

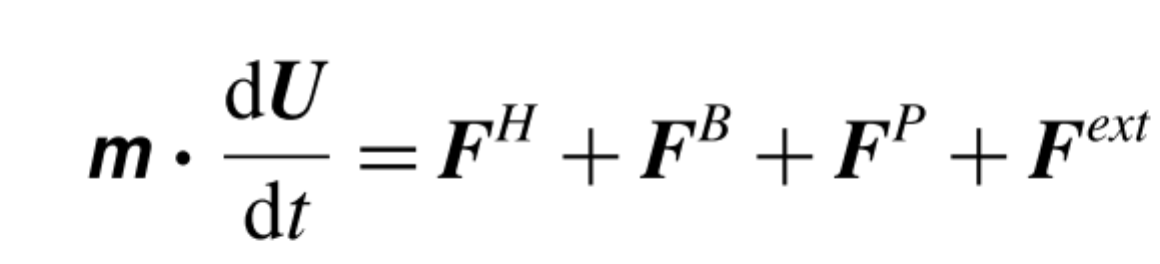
List the evidence that the authors provide for each of their claims.

1. Similarities & Applications to our system [Not needed?]
2. Parameters

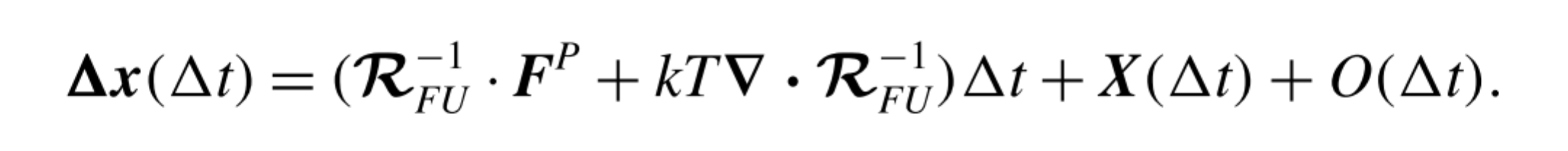
List parameters that the authors vary

1. Methods

N-body Langevin equation:



Integrating this equation over a time long compared w/ particle inertia relaxation time (overdamped limit) but short compared w/ the time for structural configurations to take place gives:



Short-time self-diffusivity was determined by connection to hydrodynamic mobility via the fluctuation-dissipation theorem:



Mufaa refers to self-mobility of particle. The mobility is normalized by the mobility of an isolated particle 1/6piettaa, and D0 is normalized by stokes-einstein coefficient, D0 = kT/(6piettaa).

[The idea here is that at short-time, the motion of a sphere is not influenced by any other sphere – this is an approximation, since the continuum approximation says that hydrodynamic interactions are instantaneously transmitted to all N particles].

Since ensemble-averaged mobility tensor (and thus short time diffusivity) is shown in this paper to be position-dependent and anisoropic, they needed a way to measure this. In an unbound suspension, self-diffusion is independent of structure and position, but any confinement leads to spatial variations in structure and to changes in the length scale over which ordered structure appears, which in turn influence particle motion and diffusion

To measure short-time self-diffusivity, particles are assigned stochastic forces with zero mean and identity covariance. Particle velocities are then calculated for a given configuration. The full mobility tensor for one realization is then given by the outer product of the forces and velocities, and averaging over many configurations then yields the short-time self-diffusivity. To capture the anisotropy of the mobility tensor, the outer product is projected onto an orthogonal basis corresponding to diffusion along and perpendicular to the cavity radius.



[Without this projection onto an the orthogonal perpendicular and along cavity basis, I would expect the mobility tensor to average out to be isotropic…? Meaning, without projection, anisotropy wouldn’t be captured?].

Position dependence of short-time self-diffusivity is captured using (100) radial bins as annular shells each separated by a delta y from center of the cavity; statistics of particles are averaged within a bin 🡪 get a radially dependent diffusion tensor.

1. Analysis

Radial distribution function g(r,ϕ) was used (In dilute limit in an unbound suspension, g(r,ϕ≪1)→1 for all r. At finite concentration, as r→infinity, g(r,ϕ) → 1. Recall that rdf already includes a normalization by the expected rdf for an ideal gas (homogenous, isotropic system…such as in dilute limit).

1. Assumptions & Shortcomings
2. Why did they choose λc = 0.05 to be the smallest particle size?
3. Important figures & tables

List up to a few core figures & tables

1. Key references

Include any references that play a critical role in the paper’s arguments

Ando & Skolnick, 2010 – Crowding and hydrodynamic interactions likely dominate in vivo macromolecular motion

1. Relevance to project
2. Claims
3. Support of claims
4. Similarities & Applications to our system [Not needed?]
5. Parameters
6. Methods
7. Analysis
8. Assumptions & Shortcomings
9. Important figures & tables
10. Key references

Diffusion, Crowding & Protein Stability in a Dynamic Molecular Model of the Bacterial Cytoplasm, McGuffee & Elcock, 2010

1. Relevance to project

This paper builds a single-molecule resolution model of the cytoplasm that captures the 50 most abundant macromolecules in E. Coli (accounting for ~85% of the cytoplasm’s characterized protein content by weight). They present many good methods for characterizing & modeling molecule motion in cytoplasm, including characterizing charge + both translational and rotational diffusivity. They attempt to recreate the long-time diffusivity of GFP via their model, and are able to do so using a fitted charge parameter.

It’s a fantastic paper showing the pipeline of parameter setup, modeling, and analysis that we’ll want to go through for our modeling of translation.

1. Claims

Claim 1.

Claim 2.

1. Support of claims
2. Similarities & Applications to our system [Not needed?]
3. Parameters
4. Methods
5. Analysis
6. Assumptions & Shortcomings

* The authors miss that the experimentally measured translation length scales for GFP are on the order of ~1µm & timescale of 10s of milliseconds; thus the “long-time diffusivity” determined by the author using a ~80nm sided voxel after 15µs should not necessarily fit with the experimentally measured diffusivity (unless 15µs and 0.1 of a cell length is sufficient to reach the average behavior shown experimentally at lower resolution). At what timescales does caging happen?

1. Important figures & tables
2. Key references

Schmidt 2016 – The quantitative and condition-dependent Escherichia coli proteome

1. Relevance to project

This paper measures the abundances (and concentrations) of half of the E. Coli BW25113 proteome in 22 different defined experimental conditions (media). Temperature defined, and tested 2 different conditions for 2 additional strains of E. coli (MG1655 and NCM3722). using state-of-the-art quantitative mass spectrometry. Thus, the paper provides high quality protein abundances for our translation model.

1. Claims
2. Determined protein abundance levels for ~55% of E. coli predicted open reading frames (>2300 proteins) and >95% of proteome mass including quantitative information on membrane and ribosomal proteins under 22 different conditions.
3. Data set is an unbiased representation of E. coli proteome that includes very hydrophobic proteins – with highly reproducible and accurate protein concentrations determined for 22 growth conditions w/ strain BW25113.
4. Also, two additional E. coli strains (MG1655 and NCM3722) were tested. Highly similar protein levels were found in these 2 strains and BW25113 (Supp fig 10 and sup table 10)
5. Identified 11 different (3 novel) types of post-translation modifications, including 318 novel ones, predominantly N-acetylations and methylations, which had not been previously reported in E. Coli
6. Uncover growth rate-dependent proteome rearrangements. Linear increase in % of total mass w/ increasing growth rate for translation molecules (Fig 2e).
7. During stationary phase, there is more bulk of protein in the periplasm/membrane relative to cytoplasm (Fig 4b). More generally, generally, protein mass fraction of cytosolic proteins significantly increases with growth rate while periplasmic proteins decrease (even considering geometric alterations from increased cell volumes w/ faster growth rate – sup fig 13).
8. Support of claims
9. Similarities & Applications to our system [Not needed?]
10. Parameters

Varied media (22 different conditions – see table S23) and strain (2 strains for 2 different conditions).

1. Methods
2. Aliquots of all samples taken from the different conditions were subjected to shotgun liquid chromatography –MS analysis to determine their condition dependent intensities by label-free quantification. Samples for proteome analyses were taken from cells that were grown until they reached ten divisions in exponential state grown at 37C. For MS analysis, DOC buffer (sodium deoxycholate buffer) was chosen to be used as lysis buffer to lyse cells (for dataset 2), 120-min LC gradient optimal to find unique peptides, and fractionation of Off-gel electrophoresis peptides into 4 pools the best compromise between comprehensiveness and analytical efforts (more fractions = more runtime but also more comprehensive). [Fig S1-S3]
3. Two LC-MS datasets were generated with different sets of methods, and then pooled (Dataset 1 was obtained first; then dataset 2 was obtained with different methods learning from dataset 1). Dataset 1 = CID MS/MS acquisition mode (refers to a specific peptide fragmentation mode), OGE-fractionation, 99 LC-MS runs, 1 biological replicate; Dataset 2 = HCD MS/MS acquisition mode, No OGE-fractionation, 72 LC-MS runs, 3 biological replicates. One major tradeoff between the two sets is between fractionation and replicates (see Supp Fig 3). In the second dataset, the lysis buffer was changed from urea to DOC to better analyze membrane proteins, and this led to 3x more quantified membrane proteins. **To provide statistically controlled quantitative values, condition and growth rate-dependent quantitative results are solely based on data set 2 (that had triplicates)**
4. Absolute protein abundances were estimated with iBAQ, The **iBAQ** (Intensity based absolute quantification) corresponds to the sum of all the peptides intensities divided by the number of observable peptides of a protein.
5. More accurately quantified 41 proteins subset to establish a calibration for MS intensities determined for all identified proteins. Quantified via stable isotope dilution (SID) and selected reaction monitoring (SRM) LC-MS/MS analysis. MS/MS refers to Tandem mass spectrometry, which involves multiple steps of mass spectrometry selection, with some form of fragmentation (such as collision-induced) occurring in between the stages to detect molecules and then their fragments. Note, SID requires doping of the target molecule (a heavy isotope version), so the proteins had to have been purified (thus this sort of quantification is not doable for all proteins).
6. Using cell numbers determined from flow cytometry analyses and condition-dependent cell volumes, accurate protein abundances per cell and per cell volume were calculated.

Conditions tested: See table S23.

1. Analysis
2. Assumptions & Shortcomings
3. Biggest assumptions: a. They assume cellular protein concentration is constant across the different growth conditions. They do this to control for variations in protein extraction efficiency—they used the total protein mass per cell (summed masses of all quantified proteins) determined in triplicate for glucose experiment by LC-MS approach, and by assuming volumetric protein concentration is condition independent, adjusted the total protein mass per cell for each condition according to precisely measured cellular volumes.
   1. NOTE: In their supplementary note 3, they note that in their more recent work (Radzikoski), when they used super resolution microscopy to measure volumes of E. coli strain K12 BW25113**, they found that their PLoS ONE study (Volkmer and Heinemann) overestimates cell volumes by ~1.5-2x for the few conditions tested.** If this holds for different cell types as well, this represents a systematic error that overestimates the protein counts per cell (since the authors claimed concentration was same, they increased protein counts for a given cell to fit that concentration; but if the volume of cell was smaller than they thought, they put too high of a count). We can adjust for this error for the volumes that were more closely measured in Radzikoski. If we assume all cell volumes were overestimated about the same, a correction factor of need not be used to get correct abundances! Since glucose grown E. coli volume used was 3.2 fL (Volkmer), while the higher resolution imaging showed 2.15 fL (Radzikoski), the correct concentration of glucose media E. coli should be 280fg/cell (see note 3) divided by 2.15 fL = 280/2.15 = 130 fg/fL as opposed to the underestimated concentration using Volkmer volume. If we want to get correct concentrations for different conditions, we need to assume that the Volkmer paper was off by ~a factor of 2 overestimate of cell volume and adjust accordingly 🡪 paper now claiming that all cells stay at ~130fg/fL protein concentration regardless of media in exponential growth.
   2. **Assumption that cellular protein concentration is constant across diff growth conditions in order to use the glucose condition measured protein mass per cell** (which they found consistent – see note 3) for other conditions for normalization. They cite Growth rate-dependent global effects on gene expression in bacteria (Hwa 2009). Figure 2 in that paper shows that protein/cell remains fairly constant across growth rate, and that mRNA/cell increases across growth rate. HOWEVER, their graph still shows a slight increase in protein/cell with increased growth rate (~30% maybe). This means conditions with growth rate faster than glucose have their abundances UNDERESTIMATED, and growth rates slower than glucose have their abundances OVERESTIMATED – within a ~30% upper-bound incorrect estimation if Hwa 2009 is to be regarded as correct. If glucose media abundances are used, the abundances are more trustworthy since they repeatedly got the same mass of protein with this condition and thus used it as the normalization media. NOTE – Dennis and Bremer 1974 (Macromolecular composition during steady-state growth of E. coli B/r ) in Figure 1c show that, above ~1.5 doublings/hr, protein/DNA ratio is constant. From ~0.6 doublings/hr to ~1.4 doublings/hr, increase from ~2.5 to 4\*10^8 aa/genome i.e. protein/DNA ratio. BUT, this is per genome not directly clear how that translates to per cell. Dennis & Bremmer later correct themselves as well in Dennis & Bremmer 2008 – they don’t calculate protein concentration, but the calculate protein/cell and protein/origin. Protein/origin is probably more accurate correlation to protein concentration – and they findBut, given that glucose is <1.4 doublings/hr, this assumption of constant concentration of protein to jump from glucose to others is questionable.
4. Their leave one out validation (Supplementary Figure 5) – I assume they are leaving out one of the triplicates -- showed that estimated abundances w/ iBAQ were on average across proteins could be off for a given condition by ~1.5x (1.38), and possibly higher (large range). The direct correlation of the 41 protein copies/cell that were accurately measured by SID to estimations via iBAQ showed pretty good correlation however; R^2 = 0.84; median fold error = 1.38. Additionally, the three biological replicates correlated very well in estimated absolute protein abundances (R^2=0.98, tested for glucose media and chemostat conditions – Supp fig 7).
5. Steady state is achieved after ten divisions in exponential phase, and thus the abundances found represent steady state abundances of E. coli in steady state (not including stationary phase conditions).
6. Important figures & tables

Fig 2e. Fraction of total protein mass from translation molecules as a function of growth rate (lambda/hr) – shows linear increase.

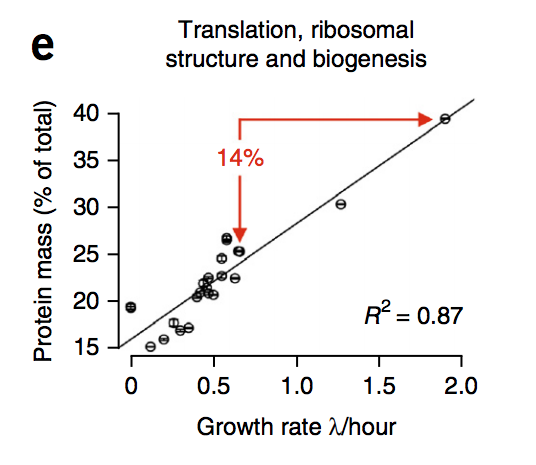


Fig 4b. Volume fraction of different cell compartments at the 3-d stationary phase and LB growth conditions based on protein mass fraction assuming constant volumetric protein concentration across conditions.

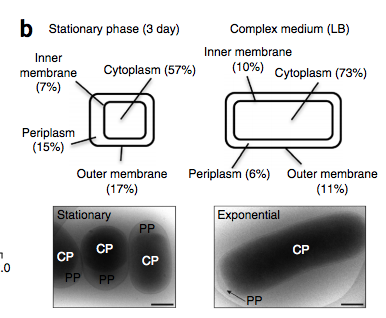
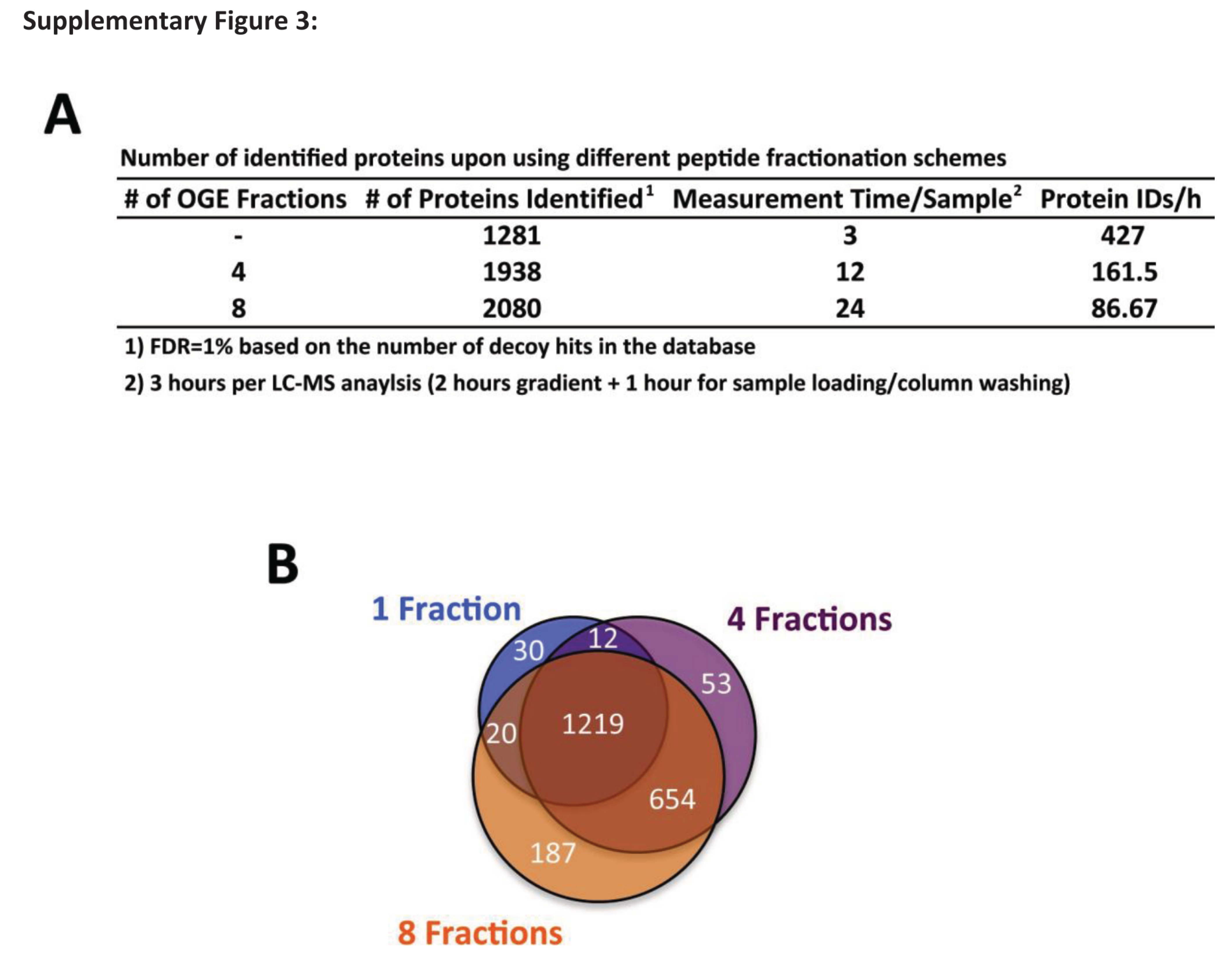


Fig S3. Effect of fractionation on number of unique proteins discovered. Evaluating more fractions (each fraction represents proteins at a different pI) separately leads to the discovery of more unique proteins, presumably because greater resolution within a set of proteins at a certain pI (need to confirm this last statement). Of course, evaluating 8 fractions separately takes (presumably) 8x more runs than just a whole pool with no fractionation, so there’s a tradeoff.



Supplementary Figure 5: The correlation of the actual cellular abundances of 41 selected proteins (in copies/cell, see Supplementary Tables 1-3 for details) determined by selected reaction monitoring and stable isotope dilution (SRM/SID)3,8 and the intensity-based absolute quantification (iBAQ) values2,5 determined by label-free quantification (both in logarithmic scale) from dataset 2 (A). (C) = Fold errors determined by leave one out cross validation (LOOCV) as box plots for all growth conditions included in data set 2. The black bar indicates the median fold error. (D). **My own note: Measurements seem to be qualitatively better for higher protein abundances (i.e., >2k protein copies/cell)**

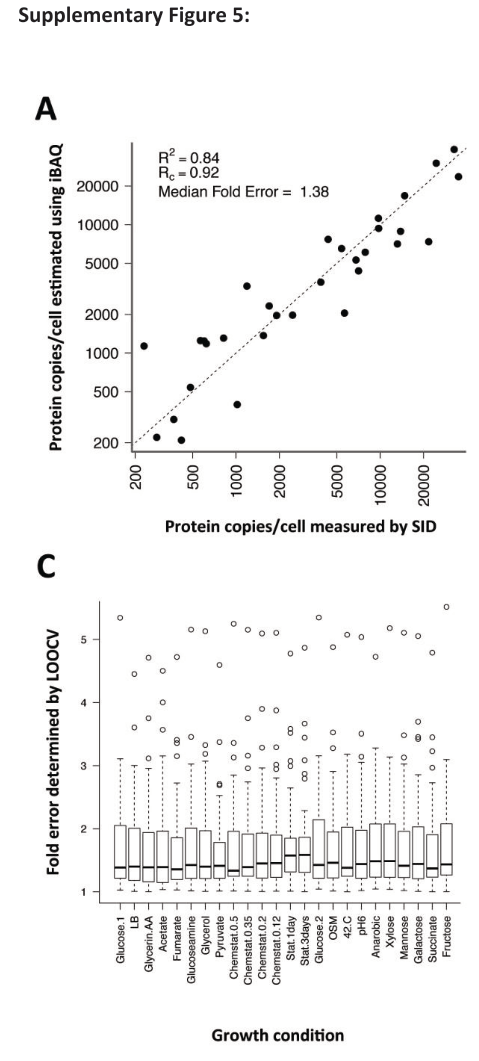
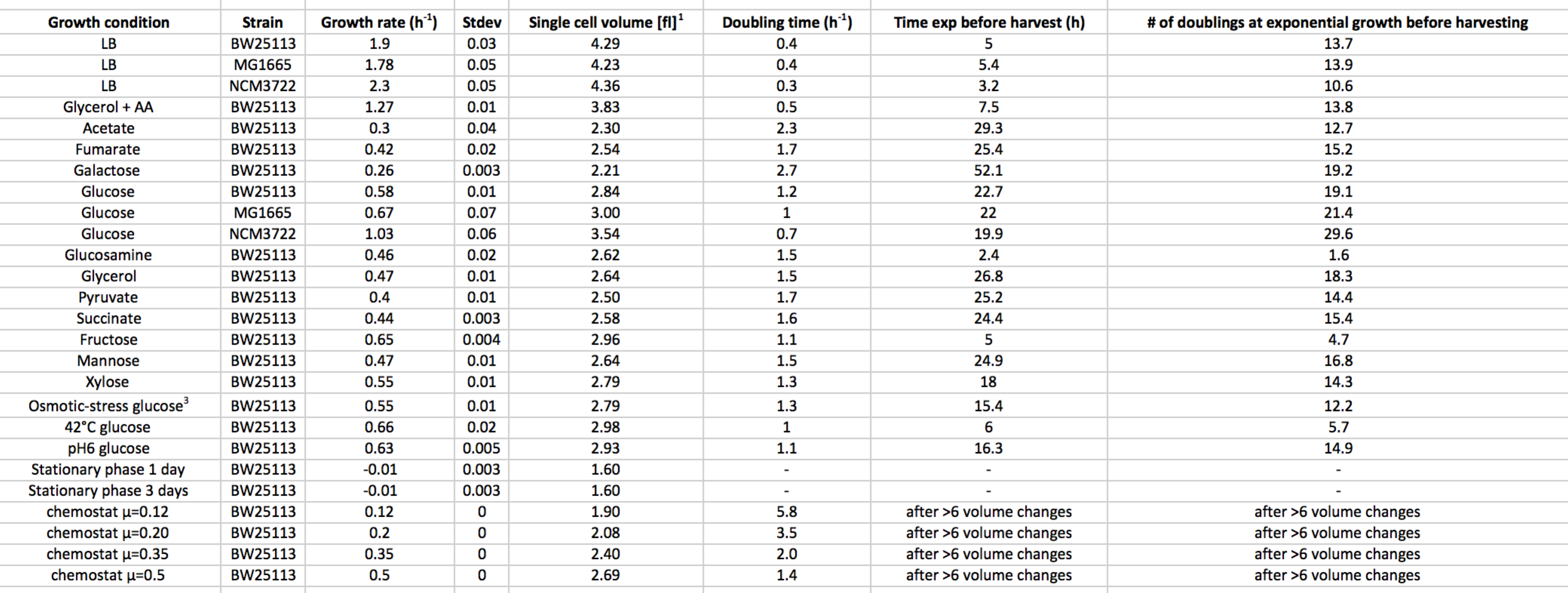


Table S23. The different conditions, and their associated growth rates, that mass spec was performed at. Check actual table for more information such as what OD and # of doublings before harvest. I believe doubling time should be h not h^-1, and that this is a typo.



1. Key references

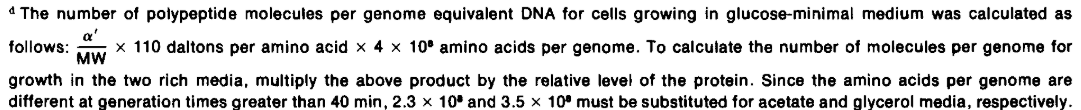
Pedersen,…, Neidhardt 1978 – Patterns of Protein Synthesis in E. coli: a Catalog of the Amount of 140 Individual Proteins at Different Growth Rates

1. Relevance to project

This paper measures the abundance of translation proteins (amongst others) in E. coli B/r in 5 different media during steady state (exponential growth) via 2-D gel electrophoresis & radiolabeling.

1. Claims
2. Support of claims
3. Similarities & Applications to our system [Not needed?]
4. Parameters
5. Methods
6. The absolute amount of each of 140 proteins was first determined in one medium (glucose+minimal media), followed by determination of the relative amount of protein at each different growth rate.
7. First, to avoid error caused by a different aa composition of different proteins, they measured absolute proteins w/ uniform 14C-glucose labeling in minimal media (presumably so 14C will be used in central metabolism to make all amino acids with 14C labeling). Protein spots were run on a 2D gel and then protein spots were cut out, and their radioactivity was measured and normalized to radioactivity in EF-G. EF-G was normalized with presumably because the authors had in a previous paper found that recovery of EF-G is ~0.94. Thus, they assumed that each protein has similar recovery to that of EF-G. They assumed that the fraction of total protein which is ribosomal proteins is 0.137 (which was measured same strain in glucose media by Dennis & Bremer) and that there is 1 EF-G molecule per ribosome in order to then set the weight fraction alpha of EF-G (which was used as the normalization for every other protein; unclear how they got # of ribosomes from 0.137 weight fraction). See assumptions section of this review for analysis of these assumptions
8. Measurement at different growth rates was done via relative measurement against the glucose media measurement. Two methods were done to cross-check.
   1. In the first method, 14C-leucine labeled reference culture was mixed with 3H-leucine other media cultures and then the 3H/14C ratio was measured. Presumably, for glucose media labeled reference into glucose media, the 3H/14C ratio should be the same for all proteins = total protein ratio, and this is what they found within “experimental error” for majority of proteins.
   2. In the second method reference glucose medium culture was labled with 3H-leucine + 3H-isoleucine, and then mixed with cells growth at various growth rates and labeled with 35-SO4. The 3H/35S ratio was then determined for total protein and each protein spot to get relative differences in abundance (relative protein spots can be normalized based on overall difference between total proteins in different media)
   3. The purity of protein spots was tested by measuring ratio of (met+cys)/(leu+ile) for each growth condition by labeling cells in separate cultures of the same medium with 35-SO4 and 3-H-leucine+3-H-isoleucine and then mixing the cells. This amino acid ratio is characteristic for each protein, and so if a spot is composed of not just the one protein, we expect that the ratio will vary with growth rate = bad. They found the ratios stayed mostly the same.
   4. They found that relative amount of protein measured by the two labeling methods were reproducible to within ~10%.
   5. Number of polypeptide molecules per genome equivalent DNA for cells in glucose-minimal media was calculated as: alpha/MW \* 110daltons per amino acid \* 4\*10^8 amino acids/genome, where alpha = weight fraction of total protein (relative to all protein) in glucose medium. Note: Since aa per genome are different at generation times greater than 40 minutes, 2.3\*10^8 and 3.5\*10^8 were substituted for acetate and glycerol media (Amino acids/genome @ diff growth rates from Dennis & Bremer 1974).
9. While most of the proteins were unlabeled that they found, A few specific polypeptides were located by their migration with marker polypeptides (pure protein or part of a purified protein complex). For example, S1, S2, L7, L12 ribosomal proteins; EF-Tu, EF-Ts, EF-G, and ten aminoacyl-tRNA synthetases
10. Analysis

* The total weight fraction of proteins found was ~0.53 in acetate to 0.46 in rich media.



alpha prime calculated using assumption of one EF-G molecule present per ribosome (Gordon 1970); that the fraction of total protein which is ribosomal proteins is 0.137 [determined in this E. coli strain in glucose medium by Dennis & Bremer 1974] and that recovery of each protein is similar to that of EF-G. They use EF-G in assumption since they determined EF-G recovery previously to be 0.94 (Neidhardt et al., 1977). Thus, from their gel, they can get the mass of EF-G; divide it by 0.94 to get the expected actual mass of EF-G, and then divide by the molar mass of EF-G to get the number of EF-G. THEN, using the 1 EF-G:1 ribosome assumption, we know that there’s an equivalent number of ribosomes. THEN, since we know how many amino acids are in the ribosomal proteins per ribosome (7336; Wittmann HG. 1982 cited by Dennis and Bremmer 2008), we can calculate the total expected mass of ribosomal proteins. With the assumption that ribosomal proteins make up 0.137 fraction of total protein mass, we now have an absolute mass that’s equivalent to 0.137 fraction of total protein mass, and we have an alpha (weight fraction) for EF-G relative to entire protein mass: (0.137\*(MW of EF-G/MW of ribosomal proteins) connected to the recovered gel mass of EF-G. Thus, all other gel mass proteins can be converted to predicted actual MW in cell (at least a minimum, in case they are actually recovered at less rate than EF-G), and then can calculate the # of molecules for each molecule in the gel.

1. Assumptions & Shortcomings

The fraction of total protein represented by each protein was calculated using the assumption that one EF-G molecule is present per ribosome & that the fraction of total protein which is ribosomal proteins is 0.137 and that the recovery of each protein is similar to that of EF-G

* 1. **One EF-G molecule present per ribosome (Gordon 1970).** I think this is the weakest assumption that is made. **Based on Schmidt 2016, we expect ~15k ribosomes but ~40k EF-G for Glucose media and strain BW25113.** If we use this assumption to say that there’s ~2.5 EF-G:1 ribosome, then predicted EF-G should be 2.5x higher. This means that alpha would be 2.5x higher for EF-G (and every other molecule)? But this would go over 1 which is a problem. Also, Dennis and Bremmer 2008 think EF-G:Ribosome ratio is 0.8 referencing this Neidhardt paper…so their table must be off too.
  2. Recovery of EF-G was previously determined to be about 0.94 (Neidhardt 1977). Recovery of other proteins might not be as high as that for EF-G, so calculated numbers should be treated as a minimum. They suspect that recovery is still satisfactory in general since each protein (Except ribosomal protein) appears as resolved spots on gel that are consistent between gels.
  3. Ribosomal proteins make up 0.137 of protein mass for the given strain in glucose medium (Dennis and Brenner 1974 is reference).
  4. Number of amino acids per genome is 4\*10^8 amino acids per genome for generation times faster than 40 min; and 2.3\*10^8 for acetate media and 3.5\*10^8 for glycerol media i.e. slower (Amino acids/genome @ diff growth rates from Dennis & Bremer 1974). BUT, in a newer paper by Dennis & Bremer with improved techniques, they found that aa/genome = 4.8\*10^8 for t\_d=100; 6.0\*10^8 for t\_d=60min; 6.6\*10^8 for t\_d=40 min….and around that bust still for faster (Dennis and Bremmer 2008 referencing Dennis & Bremmer 1998). **This would mean that all number of polypeptides were underestimated by ~6\*10^8/4\*10^8 = 1.5x. Or maybe not. Need to read the 1998 paper to double check.**

1. Important figures & tables
2. Key references