Applying Dominant Balances to a Model of Nuclear Import

Student Author: Benjamin Tseng
Primary Investigator: Michael Brenner
Graduate Student Advisor: Natalie Arkus

Jan 8, 2007

Abstract

Mathematical models of complex biological systems are oftentimes exceedingly complex despite relatively simple system dynamics. This report investigates one complex model of protein nuclear import, a process which has been experimentally shown to depend on the large concentration gradient of Ran-GTP between the nucleus and cytoplasm. It shows that the method of dominant balances can be used to effectively study the system's behavior, including determining a simple, yet accurate relationship which yields the numerical value of the Ran-GTP gradient.

1 Introduction

1.1 Dominant Balances

Recent advances in computation coupled with the increased availability of biochemical and genetic information have given mathematical models an increasingly important role in biological studies. The exceedingly complex nature of many of these systems, however, has prevented their use in developing deep biological insights. Even "simple" biological models consisting solely of mass conservation and ordinary differential equations have proven to be remarkably difficult to study due to lack of knowledge of all relevant biological parameters, the absence of explicit analytical solutions, and the ability of systems with as few as three dependent variables to yield vastly complex behavior.

Several strategies have been developed to address these issues. Metabolic reaction networks with limited parameter information have been amenable to study by semi-qualitative methodologies which employ stoichiometric relationships in those networks to make predictions about the behavior of biological systems [4, 11]. Statistical methods to create an "ensemble model" of all models capable of explaining experimental data have also been able to explore the full range of behavior of a system [1]. These methodologies, however, all fall prey to the fundamental limitation that increased simplicity and clarity comes at the loss of quantitative accuracy.

This work employs the technique of dominant balances as a means to derive simple, but accurate insights of complex biological models. The technique is a very simple one. In any multi-component system, one does not expect that all components are equally important in determining the final quantitative value. Dominant balances simply refers to simplifying a system down to only the most important components. Take the following example:

$$1 = \frac{8000 + 1900 + 100}{10100 - 100} \tag{1}$$

$$\approx \frac{8000 + 1900}{10100} \approx 0.98 \tag{2}$$

While equation (1) uses 5 terms, (2) achieves fundamentally the same relationship with 2 fewer terms (although the accuracy of such approximations must always be validated before proceeding), thus achieving greater simplicity (especially if the 2 terms involved multiple variables and/or were nonlinear) without sacrificing much quantitative accuracy (2%).

Additionally, dominant balances can also be applied to reduce the number of differential equations in a system. If there exists a separation of timescales (certain components change much slower) or a separation of quantity-scales (certain quantities are much smaller), then one can approximate those differential equations as being 0 (or, more formally, making a zeroth order approximation via perturbation analysis). This enables one to convert ordinary differential equations into solvable algebraic relationships.

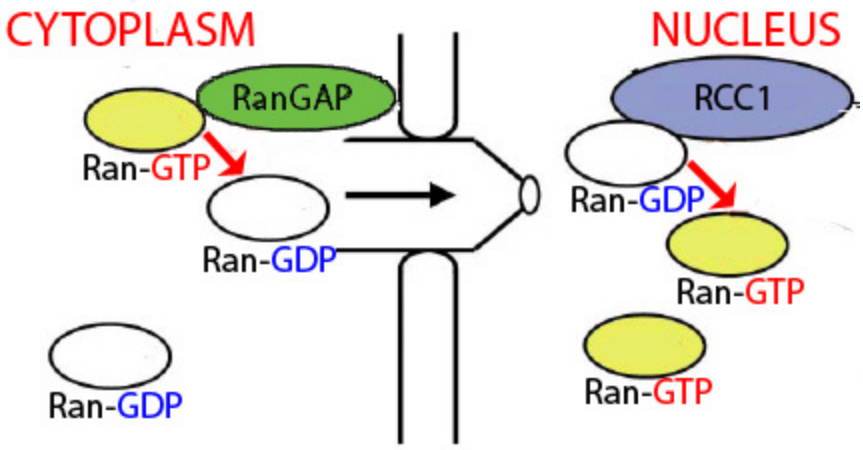


Figure 1: Establishment of the Ran-GTP Gradient

1.2 Nuclear Import and the Ran-GTP Gradient

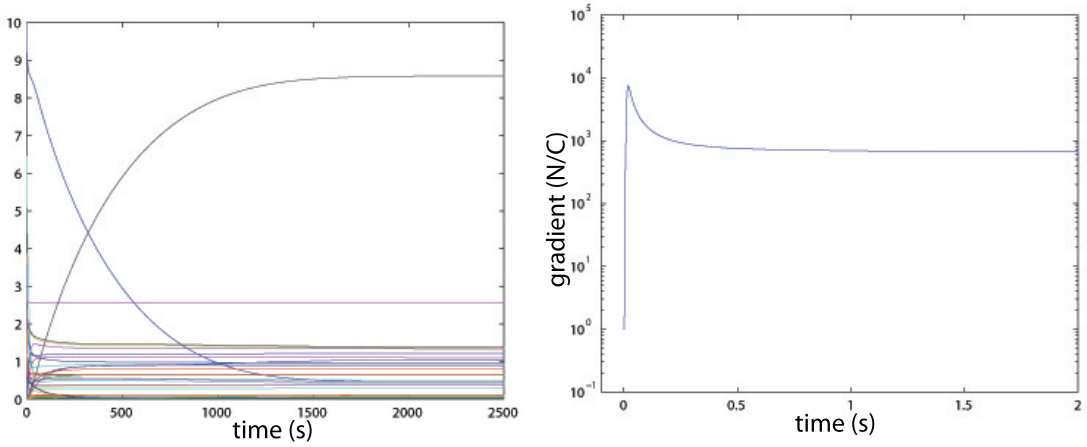
This report will investigate the use of dominant balances on a model of nuclear import published by Gregory Riddick and Ian G. Macara [9]. Protein cargo is targeted for nuclear import by the presence of a Nuclear Localization Sequence. This leads to subsequent binding by a karyopherin protein such as Importin α-β or transportin [5]. This karyopherin-cargo complex is then believed to associate with the proteins of the nuclear pore complex to facilitate diffusion across the pore and into the nucleus [3]. Subsequently, the N-terminal Ran binding domain on the karyopherin protein binds to the GTPase Ran which is typically bound to GTP in the nucleus. Association with Ran-GTP dissociates the cargo-karyopherin complex. The new Ran-GTP-karyopherin complex then diffuses to the cytoplasm whereby Ran-GTPase activity hydrolyzes the GTP, producing Ran-GDP, and a liberated karyopherin protein. Ran-GDP is then brought back into the nucleus by the protein NTF2.

This process is greatly dependent on the concentration gradient of free Ran-GTP in the cell, whereby almost all Ran-GTP is located within the nucleus and all Ran-GDP is in the cytoplasm [8] (Figure 1). This concentration gradient is believed to be maintained by two factors. The first is that RanGAP and RanBP1, two factors needed for Ran-GTP hydrolysis, are found primarily in the cytoplasm. The second is that RCC1, a factor which, in a four-step process, causes Ran to exchange GDP for GTP, resides mainly in the nucleus [6].

The work done here employs dominant balances to study which parameters in the model set the Ran-GTP gradient.

2 Model

The Macara model is a system of 61 coupled reactions determining the concentrations of 57 proteins/complexes. It assumes two separate but well-mixed compartments: the nucleus and the cytoplasm. Of the reactions, 4 are Michaelis-Menten enzyme



(a) Macara Model (b) Ran-GTP Gradient (N/C)

Figure 2: Solution to the Macara Model

kinetics reactions involving RanGAP and 20 model transport across the nuclear pore complex as approximated by simple diffusion using assumed or experimentally deduced permeabilities.

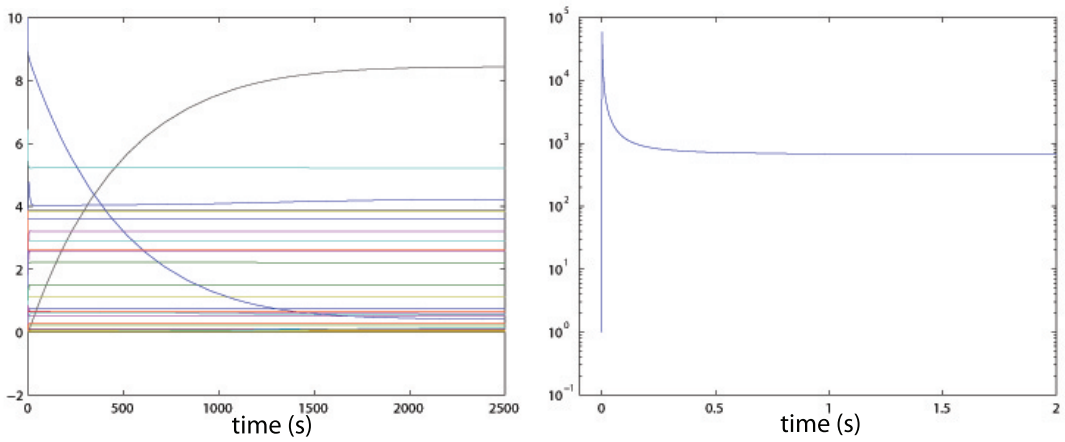
The model employs Klebe et al's four-step mechanism to describe RCC1's nucleotide exchange activity [6]. It assumes that the cargo is mostly bound to Importin α-β, a binding reaction which is modeled as a two-step process. Importin α export in this model is mediated by the karyopherin CAS [7]. The model also assumes the existence of separate karyopherins and models them as generic karyopherin entities which undergo similar reactions as their nongeneric counterparts.

The model was originally scripted for the biochemical simulation package Jarnac and was converted into a form usable by MATLAB. The system of 57 ordinary differential equations were solved by MATLAB's built-in *ode15s* solver for stiff differential equations. The model was run for a period of 2500 seconds and all available data was collected. All modifications to the model were also solved using *ode15s* for the same period of time. A graph of the behaviors of each of the variables and the Ran-GTP gradient, or the ratio between nuclear Ran-GTP to cytoplasmic Ran-GTP, is shown in Figure 2.

3 Results

3.1 Initial Attempts at Reduction

Figure 2(a) shows that the Macara model, despite the abundance of variables and reaction fluxes to consider, exhibits fairly simple behavior. There appears to be two key components moving at a timescale and concentration-scale significantly larger than the others. These two components correspond to the cytoplasmic and the nuclear cargo. Figure 2(b) also shows that the Ran-GTP gradient is set very quickly (to a final value of approximately 645) despite the initial gradient values being vastly



(a) Macara Model (b) Ran-GTP Gradient (N/C)

Figure 3: Reduced Macara Model

different (the vertical axis in the graph is on a log scale).

These two observed patterns are remarkably robust. Eliminating individual differential equations (by approximating them as 0) or the effect of individual reaction fluxes (by setting them to 0), except for in a few instances, do not significantly impact the gradient or the general pattern of cargo transport. It was even possible to approximate more than 30 of the differential equations as algebraic equations and still retain proper cargo transport and Ran-GTP gradient value (at approximately 650) (Figure 3).

The gradient value itself appears to be particularly robust. Even when cargo transport is severely impacted, the final gradient level still rarely shows any significant alteration. Table 1 lists the modifications to the system that resulted in notable changes in the gradient value, suggesting that equation 47 (concentration of RanBP1), reaction flux 9 (diffusion of Ran-GTP across the nuclear membrane), and reaction fluxes 41 and 61 (GTPase activity on cytosolic Ran-GTP) play a major role in determining the Ran-GTP gradient. The table also shows modifications which do not appear to significantly impact the gradient value but that dramatically affect the nuclear and cytoplasmic Ran-GTP concentrations (and in the case of flux 25 even making them negative!), suggesting that while the compartment concentrations of Ran-GTP can be altered, the ratio between the two is a conserved quantity controlled by some aspect of the system which is not easy to modify.

3.2 Gradient Determination

Of the three implicated reaction fluxes, two appear to be significantly more important as setting them to 0 induces a greater change in the gradient value. These two fluxes are both components of the ordinary differential equation which controls the

modification	gradient (N/C)	Nuc Ran-GTP	Cyt Ran-GTP
Original	645.4926	0.9628	0.0015
eqn 32 = 0	646.8339	1.43E-09	2.21E-12
eqn 47 = 0	486.0025	7.31E-04	1.50E-06
flux 9 = 0	4.09E+07	1.0137	2.48E-08
flux 24 = 0	646.6869	1.60E-16	2.47E-19
flux 25 = 0	646.6844	-9.64E-17	-1.49E-19
flux 26 = 0	646.6831	3.21E-16	4.96E-19
flux 41 = 0	463.6063	0.1008	2.17E-04
flux 61 = 0	29.2978	0.958	0.0327

Table 1: Approximations Leading to Gradient Value Changes

concentration of cytoplasmic Ran-GTP:

$$\begin{aligned}\frac{d[\text{RanGTP}]_c}{dt} &= (0.03 \cdot [\text{RanGTP}]_n - 0.03 \cdot [\text{RanGTP}]_c) \\ &- (0.096 \cdot [\text{RanGTP}]_c \cdot [\text{Imp } \beta]_c - 4.8 \times 10^{-6} \cdot [\text{RanGTP-Imp } \beta]_c) \\ &- (10^{-7} \cdot [\text{Cas}]_c \cdot [\text{RanGTP}]_c \cdot [\text{Imp } \alpha]_c - 10^{-10} \cdot [\text{RanGTP-Cas-Imp } \alpha]) \\ &- \frac{20.1 \cdot [\text{RanGAP}] \cdot [\text{RanGTP}]_c}{0.7 + [\text{RanGTP}]_c} \\ &- (0.3 \cdot [\text{RanGTP}]_c \cdot [\text{RanBP1}]_c - 4 \times 10^{-4} \cdot [\text{RanGTP-RanBP1}]_c) \quad (3)\end{aligned}$$

Equation (3) is not amenable to a simple analysis. However, if one plugs in the final concentrations of each of those components, one obtains:

$$\begin{aligned}\frac{d[\text{RanGTP}]_c}{dt} &= (0.02888 - 4.4747 \times 10^{-5}) \\ &- (0.0001147 - 4.4192 \times 10^{-7}) \\ &- (3.9599 \times 10^{-12} - 1.0525 \times 10^{-12}) \\ &- 0.02757 \\ &- (0.001152 - 2.9669 \times 10^{-8}) \quad (4)\end{aligned}$$

Noting that as the system approaches equilibrium the rate of change of cytoplasmic Ran-GTP approaches 0 and that two terms appear to dominate equation (4) one obtains the dominant balance:

$$0 \approx 0.02888 - 0.02757 \quad (5)$$

Equation (5) presents all the necessary information captured by the relationship in (4) and thus the dominant balance at near-equilibrium includes two terms:

$$0 \approx 0.03 \cdot [\text{RanGTP}]_n - \frac{20.1 \cdot [\text{RanGAP}] \cdot [\text{RanGTP}]_c}{0.7 + [\text{RanGTP}]_c} \quad (6)$$

Separating terms and noting that cytoplasmic Ran-GTP concentrations are very small throughout the course of the model one obtains an approximation for nuclear

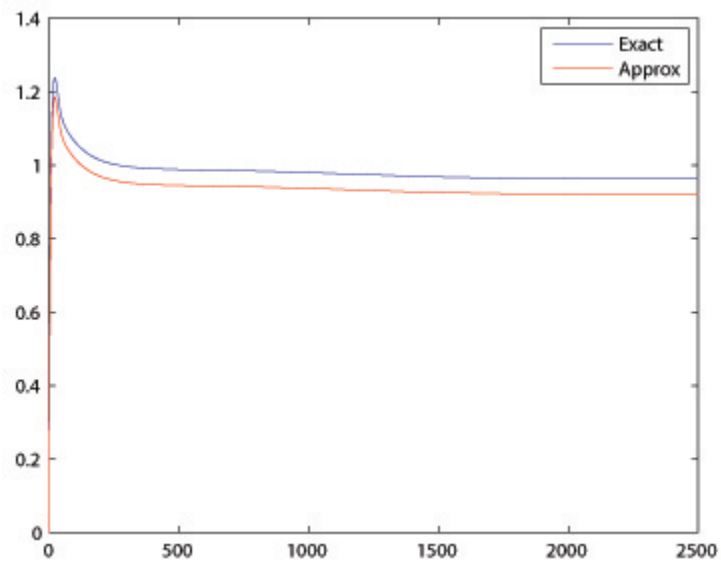


Figure 4: Graphic Validation of Nuclear Ran-GTP Dominant Balance

Ran-GTP levels:

$$[\text{RanGTP}]_n \approx \frac{20.1 \cdot [\text{RanGAP}] \cdot [\text{RanGTP}]_c}{0.03 \cdot 0.7} \tag{7}$$

Equation (7) is shown in Figure 4 to not only have approximately the same value as the exact solution but to follow the same pattern throughout the course of the model, suggesting that this dominant balance is appropriate for the entire timecourse of the model. The validity of this approximation thus also allows one to derive an elementary formula for the Ran-GTP gradient (note: the concentration RanGAP is an exogenously specified, unvarying parameter):

$$\frac{[\text{RanGTP}]_n}{[\text{RanGTP}]_c} \approx \frac{20.1 \cdot [\text{RanGAP}]}{0.03 \cdot 0.7} = \frac{k_{cat} \cdot [\text{RanGAP}]}{P(\text{RanGTP}) \cdot K_m(\text{no RanBP1})} = 617.5115 \tag{8}$$

As the model gives a gradient value of approximately 645, thus the approximation gives an error of less than 5%!

3.3 Ran-GTP Determination

The next most logical step was to apply dominant balances towards finding an explicit approximation for the concentration of nuclear Ran-GTP, and thus, using the explicit form for the gradient derived in Equation (8), determine an explicit form for the concentration of cytoplasmic Ran-GTP. Table 2 shows the size of all 24 terms in the ordinary differential equation controlling nuclear Ran-GTP concentration at near-equilibrium.

	term 1	term 2
flux 9	0.028884	4.47473E-05
flux 26	35.77643	35.58247182
flux 31	0.0081284	4.1166E-13
flux 32	0.0131608	0.005031922
flux 34	1.02E-08	4.39074E-11
flux 37	2.793E-12	1.36145E-22
flux 39	3.375E-18	2.82963E-20
flux 48	0.0008985	4.78256E-12
flux 51	2.366E-12	2.51591E-20
flux 54	0.0596482	5.30924E-06
flux 55	0.0023119	4.42148E-14
flux 56	0.0860152	6.43236E-06

Table 2: Size of Terms in Equation Controlling Nuclear Ran-GTP

As before, there are two terms which appear to dominate the equation.

$$\begin{aligned} 0 &\approx 55 \cdot [\text{RCC1-RanGTP}]_n - 100 \cdot [\text{RCC1}] \cdot [\text{RanGTP}]_n \qquad (9) \\ &= 35.7764 - 35.5825 \\ \text{RanGTP}_n &\approx \frac{55 \cdot [\text{RCC1-RanGTP}]}{100 \cdot [\text{RCC1}]} \qquad (10) \end{aligned}$$

The objection may be raised that the numerical difference in Equation (9) is approximately 0.2 which is significantly different enough from 0 to warrant including additional terms in the dominant balance. While this is a valid concern, one need only consider that both numerator and denominator in Equation (10) are sufficiently larger than all the other terms in Table 2 that inclusion of additional terms will not greatly affect the approximation. Furthermore, Figure 5 shows that the approximation is very good, also suggesting that the dominant balance made was an appropriate one.

The equation for nuclear Ran-GTP involves two variables. Unfortunately, unlike in the case of RanGAP, neither are unvarying parameters. However, RCC1 and the RCC1-RanGTP complex are two of the four species of RCC1 complexes. Because the total amount of RCC1 must remain constant, mass conservation allows one to develop a relationship between RCC1 and RCC1-RanGTP. Investigating the mass conservation of RCC1 species and applying dominant balances:

$$\begin{aligned} \frac{d\sum \text{RCC1}_x}{dt} &= 0 \\ \sum \text{RCC1}_x &= [\text{RCC1}] + [\text{RCC1-RanGTP}] + [\text{RCC1-RanGDP}] + [\text{RCC1-Ran}] \\ 1.1111 &= 0.3696 + 0.6505 + 0.0445 + 0.0465 \\ &\approx 0.3696 + 0.6505 \\ \sum \text{RCC1}_x &\approx [\text{RCC1}] + [\text{RCC1-RanGTP}] \qquad (11) \end{aligned}$$

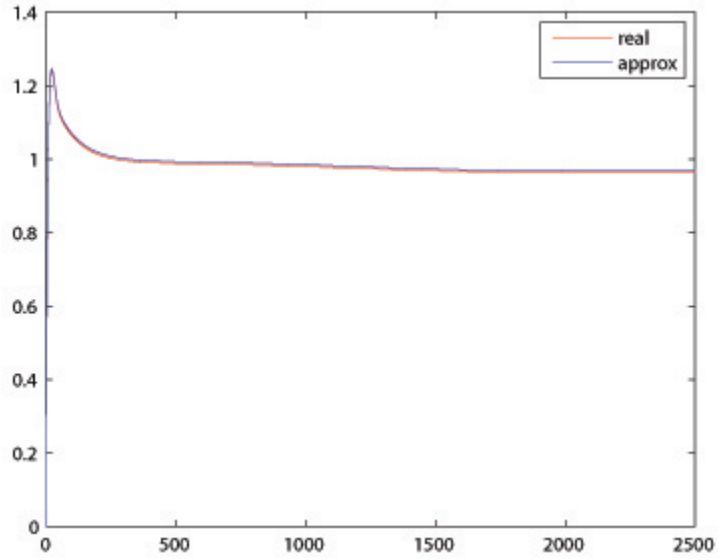


Figure 5: Graphic Validation of Nuclear Ran-GTP Dominant Balance

Applying equation (11) to (10), one thus obtains the relationship:

$$[\text{RanGTP}]_n \approx \frac{0.55 \cdot [\text{RCC1-RanGTP}]}{1.1111 - [\text{RCC1-RanGTP}]} = \frac{0.55 \cdot (1.1111 - [\text{RCC1}])}{[\text{RCC1}]} \tag{12}$$

Obtaining an explicit solution via dominant balances for RCC1-RanGTP or for RCC1, however, has proven to be elusive, as the most appropriate dominant balance for RCC1 has simply been a reformulation of equation (10). The most appropriate dominant balance for RCC1-RanGTP, however, can be expressed in terms of a series of parameters and RCC1-Ran:

$$[\text{RCC1-RanGTP}] \approx \frac{0.6 \cdot [\text{RCC1-Ran}] \cdot [\text{GTP}]_n}{19} \tag{13}$$

The most appropriate dominant balance for RCC1-Ran, on the other hand, is a reformulation of (13), rendering the problem thus far intractable.

4 Discussion

The application of dominant balances was successful in achieving a simple and accurate representation of the Ran-GTP gradient. Although we have been thus far unable to achieve an explicit formulation for the concentration of nuclear Ran-GTP in terms of only parameters, the application of dominant balances reveals very simple intuitions to explain both the gradient value and the equilibrium concentrations of nuclear and cytoplasmic Ran-GTP.

Equation (6) shows that the critical dominant balance for cytoplasmic Ran-GTP concentration is a balance between diffusion of Ran-GTP across the nuclear membrane and its conversion to Ran-GDP by RanGAP. This is biologically startling for

Reaction	Parameters	Reaction Rate
Hydrolysis of RanGTP-RanBP1	$k_{cat} = 20.1$, $K_M = 0.1$	0.009611
Hydrolysis of RanGTP-Importin β-RanBP1	$k_{cat} = 2$, $K_M = 0.1$	0.070764
Hydrolysis of RanGTP-Karyopherin-RanBP1	$k_{cat} = 21.2$, $K_M = 0.1$	0.086009
Hydrolysis of RanGTP	$k_{cat} = 20.1$, $K_M = 0.7$	0.027573

Table 3: RanGAP Enzymatic Rates at Equilibrium

three reasons. The first is that the conventional wisdom of Ran-GTP movement from the nucleus to the cytoplasm is much more heavily dependent on association of Ran-GTP with karyopherin proteins [5, 8, 9]. This dominant balance stipulates that the dominant "source" of Ran-GTP in the cytoplasm is not from dissociation of the various karyopherin-Ran-GTP complexes, but from what the model refers to as "RanGTP Leakage into the Cytoplasm" (see supplemental information for [9]). This appears to be a consequence of the very small dissociation constants for the various complexes (see equation 3).

The second startling factor is that the critical source for the "loss" of cytoplasmic Ran-GTP lies in the GTPase activity stimulated by RanGAP not in the presence of RanBP1. This is surprising because RanBP1 has been established to be a critical cofactor for the GTPase stimulating activity of RanGAP by stabilizing the Ran-GTP-RanGAP interaction (lowering the effective Michaelis-Menten constant) [2]. Table 3 supports this as it shows that the most dominant GTP hydrolysis reactions occur in conjunction with RanBP1. However, it also reveals that the hydrolysis of Ran-GTP in the presence of RanBP1 is significantly smaller than the hydrolysis of Ran-GTP, a consequence of the very low amount of RanGTP-RanBP1 complex in the cytoplasm.

The third interesting observation is that while it is widely believed that RCC1 plays a major role in establishing the value of the gradient, equation (8) reveals little dependence of the gradient value on any parameter involved with RCC1. Instead, the equation reveals that the only parameters that matter are the enzymatic parameters of the RanGAP-assisted GTPase reaction independent of RanBP1 and the ability of Ran-GTP to diffuse into the cytoplasm from the nucleus. The critical determinant of the gradient, then, according to this dominant balance, lies completely in the cytoplasm between Ran-GTP diffusion and hydrolysis.

An astute observer should question this, as it is clear that RCC1 ought to play a major role in the system. This is clear from various studies, computational and experimental [9, 10], as import levels and rates have been shown to be very sensitive to RCC1 concentrations. The insight then, is from the as yet unsolved equation (12) which shows that the nuclear Ran-GTP concentration is dependent on its association/dissociation with RCC1. Thus, what this system of dominant balances has shown (and verified in Figure 5) that RCC1 is that the critical determinant of the nuclear Ran-GTP concentration and the relationship between the ability of Ran-GTP to diffuse across the nuclear membrane and RanGAP's hydrolysis of cytoplasmic Ran-GTP is the critical determinant of the Ran-GTP gradient value.

It should be noted that these observations are not necessarily valid biological insights as they are contingent on the validity of the model in question. However, these dominant balances provide testable hypotheses pertaining to the system. It should be relatively simple to modulate specific properties such as the amount of RanGAP via introduction of destabilizing protein tags or modification of transcriptional control sequences or the enzymatic properties of RanGAP via amino acid substitutions in RanGAP or in Ran. One would expect such changes to impact the Ran-GTP gradient if this model were correct. On the other hand, similar modifications to RCC1 would be expected to have a minimal effect on the gradient value, but would be expected to alter the nuclear concentration of Ran-GTP. Negative hypotheses can also be easily formulated and tested. While one would expect the nuclear pore permeability of Ran-GTP to impact the gradient, one would not expect, within a reasonable order of magnitude, that changes to CAS or NTF2 would have a significant impact on the gradient value.

Of particular mathematical note is the difficulty of finding an explicit solution for the concentration of nuclear Ran-GTP. At first glance, this problem appears to be due to the use of Klebe et al's four-step model for the mechanism of RCC1's mediation of Ran nucleotide exchange [6]. The four-step process involves RCC1 cycling through being complexed with Ran-GTP, Ran-GDP, and Ran unbound to any nucleotide. Riddick and Macara in formulating this model address experimental evidence for this as shown by the sensitivity of import rates to RCC1 levels whereby increases and decreases in RCC1 affect the balance of the four RCC1-Ran complexes [9]. This cyclical nature of the mechanism has made it somewhat difficult to apply dominant balance to obtain an explicit form for any of the four complexes as a result of the dominant balance for each complex being dependent on the concentrations of the other complexes. More analysis will have to be done in this arena to obtain an explicit expression for nuclear Ran-GTP dependent only on parameters given by the system.

Dominant balances is a simple means to obtain an intuitive understanding of a mathematical system without sacrificing a great deal of accuracy. In this report, we show that dominant balances can be applied towards a complex system of equations to derive an intuitive understanding of how two particular numbers, the Ran-GTP nuclear to cytoplasmic ratio and the nuclear Ran-GTP concentration, are derived. While an explicit determination of the latter quantity only in terms of unvarying parameters was not produced, the determination of the former showed that the simple understanding of the system resulted in a negligible error. In both cases, insights into the system were arrived upon which were not possible in the original formulation, and these insights can be easily tested experimentally to validate or invalidate the model and to derive a new understanding of nuclear import. This, and other work done in our group, suggests that dominant balances may be a means to understand a wide range of complex biological systems.

References

[1] D. Battogtokh, D. K. Asch, M. E. Case, J. Arnold, and H-B. Schuttler. An ensemble method for identifying regulatory circuits with special reference to the qa gene cluster of Neurospora crassa. *Proc Natl Acad Sci U S A*, 99(26):16904–16909, Dec 2002.

[2] F. R. Bischoff, H. Krebber, E. Smirnova, W. Dong, and H. Ponstingl. Co-activation of RanGTPase and inhibition of GTP dissociation by Ran-GTP binding protein RanBP1. *EMBO J*, 14(4):705–715, Feb 1995.

[3] Y. M. Chook and G. Blobel. Karyopherins and nuclear import. *Curr Opin Struct Biol*, 11(6):703–715, Dec 2001.

[4] J. S. Edwards, R. U. Ibarra, and B. O. Palsson. In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. *Nat Biotechnol*, 19(2):125–130, Feb 2001.

[5] Amnon Harel and Douglass J Forbes. Importin β: conducting a much larger cellular symphony. *Mol Cell*, 16(3):319–330, Nov 2004.

[6] C. Klebe, H. Prinz, A. Wittinghofer, and R. S. Goody. The kinetic mechanism of Ran–nucleotide exchange catalyzed by RCC1. *Biochemistry*, 34(39):12543–12552, Oct 1995.

[7] U. Kutay, F. R. Bischoff, S. Kostka, R. Kraft, and D. Grlich. Export of importin α from the nucleus is mediated by a specific nuclear transport factor. *Cell*, 90(6):1061–1071, Sep 1997.

[8] B. Booth Quimby and Mary Dasso. The small GTPase Ran: interpreting the signs. *Curr Opin Cell Biol*, 15(3):338–344, Jun 2003.

[9] Gregory Riddick and Ian G Macara. A systems analysis of importin-α-β mediated nuclear protein import. *J Cell Biol*, 168(7):1027–1038, Mar 2005.

[10] Alicia E Smith, Boris M Slepchenko, James C Schaff, Leslie M Loew, and Ian G Macara. Systems analysis of Ran transport. *Science*, 295(5554):488–491, Jan 2002.

[11] Ralf Steuer, Thilo Gross, Joachim Selbig, and Bernd Blasius. Structural kinetic modeling of metabolic networks. *Proc Natl Acad Sci U S A*, 103(32):11868–11873, Aug 2006.