

Modeling and Simulation of Aggregation of Membrane Protein LAT with Molecular Variability in the Number of Binding Sites for Cytosolic Grb2-SOS1-Grb2

Ambarish Nag^{1*}, Michael Monine^{2#b}, Alan S. Perelson¹, Byron Goldstein¹

1 Theoretical Biology and Biophysics Group, Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico, United States of America, **2** Theoretical Biology and Biophysics Group, Theoretical Division and Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico, United States of America

Abstract

The linker for activation of T cells (LAT), the linker for activation of B cells (LAB), and the linker for activation of X cells (LAX) form a family of transmembrane adaptor proteins widely expressed in lymphocytes. These scaffolding proteins have multiple binding motifs that, when phosphorylated, bind the SH2 domain of the cytosolic adaptor Grb2. Thus, the valence of LAT, LAB and LAX for Grb2 is variable, depending on the strength of receptor activation that initiates phosphorylation. During signaling, the LAT population will exhibit a time-varying distribution of Grb2 valences from zero to three. In the cytosol, Grb2 forms 1:1 and 2:1 complexes with the guanine nucleotide exchange factor SOS1. The 2:1 complex can bridge two LAT molecules when each Grb2, through their SH2 domains, binds to a phosphorylated site on a separate LAT. In T cells and mast cells, after receptor engagement, receptor phosphorylation is rapidly followed by LAT phosphorylation and aggregation. In mast cells, aggregates containing more than one hundred LAT molecules have been detected. Previously we considered a homogeneous population of trivalent LAT molecules and showed that for a range of Grb2, SOS1 and LAT concentrations, an equilibrium theory for LAT aggregation predicts the formation of a gel-like phase comprising a very large aggregate (superaggregate). We now extend this theory to investigate the effects of a distribution of Grb2 valence in the LAT population on the formation of LAT aggregates and superaggregate and use stochastic simulations to calculate the fraction of the total LAT population in the superaggregate.

Citation: Nag A, Monine M, Perelson AS, Goldstein B (2012) Modeling and Simulation of Aggregation of Membrane Protein LAT with Molecular Variability in the Number of Binding Sites for Cytosolic Grb2-SOS1-Grb2. PLoS ONE 7(3): e28758. doi:10.1371/journal.pone.0028758

Editor: Thomas Höfer, German Cancer Research Center, Germany

Received: July 20, 2011; **Accepted:** November 14, 2011; **Published:** March 1, 2012

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

Funding: This work was supported by the Department of Energy through contract W-7405-ENG-36, and National Institutes of Health grants R37-GM035556 (BG), R37-AI028433 (ASP), R01-RR006555 (ASP) and P01-AI071195 (ASP). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ambarish.nag@nrel.gov

#a Current address: Computational Sciences Center, National Renewable Energy Laboratory, Golden, Colorado, United States of America

#b Current address: Novartis Institutes for Biomedical Research, Inc., Cambridge, Massachusetts, United States of America

Introduction

A ubiquitous method of initiating intracellular signals is through ligand-induced receptor aggregation [1–3]. However the role of aggregation in cell signaling is not restricted to bringing together the cytoplasmic domains of receptors, but occurs in multiple ways among multiple signaling molecules during the course of signaling. Scaffolding and adaptor proteins play important roles in cellular signal transduction [4–7]. The aggregation of scaffolding proteins differs from the aggregation of receptors in a fundamental way. The valence of a scaffolding protein for the species that induces its aggregation depends on its state of phosphorylation. The cytoplasmic domains of scaffolding proteins often have multiple binding motifs that, when phosphorylated, bind one or more types of signaling molecules. The number of phosphorylated binding motifs in a scaffolding protein molecule determines the valency of the molecule for other signaling molecules. Cross-linking of bivalent scaffolding molecules by other bivalent signaling molecules can result in the formation of chain-like aggregates. Cross-linking of multivalent scaffolding molecules by bivalent/multivalent signaling molecules can yield branched aggregates.

In the current article, we focus on the aggregation of the linker for the activation of T cells (LAT), a scaffolding protein that is essential for full mast cell and T cell function [8]. LAT is a membrane-localized adaptor protein which is primarily concentrated in microdomains by palmitoylation [5,9]. It acts as a major signaling hub in the signaling network initiated by the aggregation of FcεRI on mast cells and the aggregation of T cell receptors (TCRs) on T cells. LAT is a single chain protein composed of a transmembrane domain, a two amino acid extracellular domain, and a cytoplasmic domain containing nine tyrosine residues conserved in mice, rats and humans [10,11]. The distal four LAT tyrosines play an essential role for both T cell [12] and mast cell function [13]. These four tyrosines, when phosphorylated, allow the SH2 containing phospholipase C γ (PLC γ) and the adaptor proteins Grb2 and Gads to associate with LAT [12,14–16]. These adaptors [6,17,18] in turn recruit SOS1 and SLP-76 to LAT.

The distal three tyrosines on human LAT, Tyr171, Tyr191 and Tyr226, are located in YXNX sequences that bind the SH2 domain of Grb2 when the tyrosine in the motif is phosphorylated [19]. Hence, the valence of LAT for Grb2 can vary from zero to three depending on the state of LAT phosphorylation and in either

mast cells or T cells that have been activated, there will be heterogeneity in the LAT population with respect to the number of Grb2 binding sites. In mast cells, the distribution of the LAT population over the different Grb2 valence states depends on the concentration of activated Syk molecules [20,21], the kinase responsible for phosphorylating LAT, which in turn depends on the concentration of the extracellular allergen that cross-links IgE-FcεRI complexes. Similarly, In T cells the LAT valence distribution depends on the concentration of activated ZAP-70 molecules [22], which in turn depends on the extent of TCR activation.

The cytosolic adaptor molecule Grb2 consists of one SH2 domain flanked on each side by a Src homology 3 (SH3) domain [23,24]. Both the SH3 domains of a Grb2 molecule can bind to proline-rich regions on a SOS1 molecule to form a 1:1 Grb2-SOS1 complex. Binding of a second Grb2 involving both its SH3 domains to this complex can yield a Grb2-SOS1-Grb2 dimer capable of bridging two LAT molecules [6], each of which has at least one phosphorylated distal tyrosine. Although SOS1 has at least four binding sites for the SH3 domains of Grb2 [25], complexes of SOS1 and Grb2 with more than two Grb2 have not been reported. The interaction of this bivalent ligand (Grb2-SOS1-Grb2) with a trivalent scaffolding protein (fully phosphorylated LAT) offers the possibility of large aggregate formation since cross-linking of a trivalent species by a bivalent ligand can yield highly branched aggregates. Large clusters of LAT have been observed to form in T cells following T cell engagement [6,26] and in mast cells following the aggregation of IgE-FcεRI complexes [27,28]. Experiments have shown that the cytosolic 2:1 Grb2-SOS1 complex acts as a cross-linking agent between LAT molecules, leading to LAT cluster formation in T cells [6,12,29].

Previously, we presented an equilibrium model of LAT oligomerization for homogeneous populations of phosphorylated LAT, being content to compare the oligomerization properties of bi- and trivalent LAT in the presence of Grb2 and SOS1 [30]. The model predicted that over a certain range of concentrations of trivalent LAT, Grb2 and SOS1, a superaggregate of LAT formed in equilibrium with a distribution of smaller LAT aggregates and LAT monomers, i.e., there existed a sol-gel coexistence region [30–32]. We used the model to map the boundaries of the sol-gel region and predict the conditions under which a superaggregate can form, and the fraction of LAT molecules involved in the superaggregate for a given set of concentrations. For concentrations outside this range, a superaggregate (the gel phase) was predicted not to form. As has been shown in general for a bivalent ligand interacting with a bivalent receptor [31], for bivalent LAT there was no concentration range for which a sol-gel coexistence region existed. We now undertake a generalization of our previous equilibrium model to account for the distribution of LAT molecules over different valence states. Our focus is to understand how the presence of mono- and bivalent LAT alters the range of concentrations for which a gel phase can exist. Monovalent LAT will block LAT aggregation while bivalent LAT will prevent branching of LAT aggregates so we expect the presence of bi- and monovalent LAT to significantly reduce the range of concentrations over which a superaggregate can form. The main question that we address in the current article is how sensitive is the superaggregate formation to the presence of lower valence LAT. From the current article, it is evident that we can still expect to observe such aggregate formation for a range of physically relevant parameters. We have also calculated the fraction of the total LAT population in a superaggregate from stochastic simulations of the time evolution of the system. The simulation algorithm is an adaptation of a rule-based kinetic Monte Carlo approach discussed in detail in our previous work [30].

A related but different problem is to obtain an equilibrium model for the aggregation of bivalent receptors on cell surfaces, by a mixture of monovalent, bivalent and trivalent ligand species. Goldstein and Perelson (1984) developed the equilibrium theory for the clustering of bivalent cell surface receptors by trivalent ligands. We extend the theory in Goldstein and Perelson (1984) to include monovalent and bivalent ligand species in the Appendix S3 of this paper.

Methods

Theory

We present an equilibrium theory to calculate the LAT aggregate size distribution when there is differential phosphorylation of the LAT population resulting in heterogeneity in the number of Grb2 binding sites on LAT. We also present agent based simulations to check our equilibrium model [32]. For these simulations we use the rate constants as well as equilibrium constants that characterize the interactions in the model. The parameter values we use in our model are the same as in Nag et al. (2009). How they were obtained are fully discussed in that reference.

Calculation of total partition function for all species containing LAT molecules with variable valency. We first introduce the notation we will use and summarize our previous model. The cytosolic volume and surface area of the cell are represented by V and A respectively. In contrast to our previous paper [30], where the entire LAT population is either bivalent or trivalent, the population of LAT in the current article is distributed between variable valence states for Grb2 - monovalent, bivalent and trivalent. The total cellular concentrations of monovalent, bivalent and trivalent LAT molecules are denoted by M_T , D_T and L_T respectively so that the total numbers of membrane associated trivalent, bivalent and monovalent LAT per cell are AL_T , AD_T , and AM_T respectively. The total concentrations of Grb2 and SOS1 available to interact with LAT and each other are G_T and S_T respectively, so that the total numbers of Grb2 and SOS1 in the cell are VG_T and VS_T respectively. Grb2 contains an SH2 and two SH3 domains. It binds to phosphorylated LAT via its SH2 domain and to SOS1 through its SH3 domains. Grb2 can form both Grb2-SOS1 and Grb2-SOS1-Grb2 complexes [6,33]. The equilibrium concentrations of free monovalent LAT, bivalent LAT, trivalent LAT, Grb2 and SOS1 are M , D , L , G , and S , respectively.

At equilibrium, the formation of the bivalent ligand Grb2-SOS1-Grb2 is characterized by an equilibrium constant K_{GS} and a cooperativity factor σ . The equilibrium constant for the binding of the first Grb2 to a single Grb2 binding site on free SOS1 is K_{GS} . The forward and reverse rate constants for this binding event are k_{+GS} and k_{-GS} , respectively. We take the binding of a second Grb2 molecule to a single Grb2 binding site in a 1:1 Grb2-SOS1 complex to be negatively cooperative with the forward and reverse rate constants for the cross linking event as σk_{+GS} and k_{-GS} . The corresponding equilibrium constant is σK_{GS} . A $\sigma \approx 0.5$ is consistent with the binding studies of Houtman et al. (2006). It follows from the definition of these equilibrium constants that C , the equilibrium concentration of the bivalent ligand Grb2-SOS1-Grb2, is

$$C = \sigma K_{GS}^2 G^2 S. \quad (1)$$

The reactions leading to the formation of this complex are illustrated in Fig. 1a.

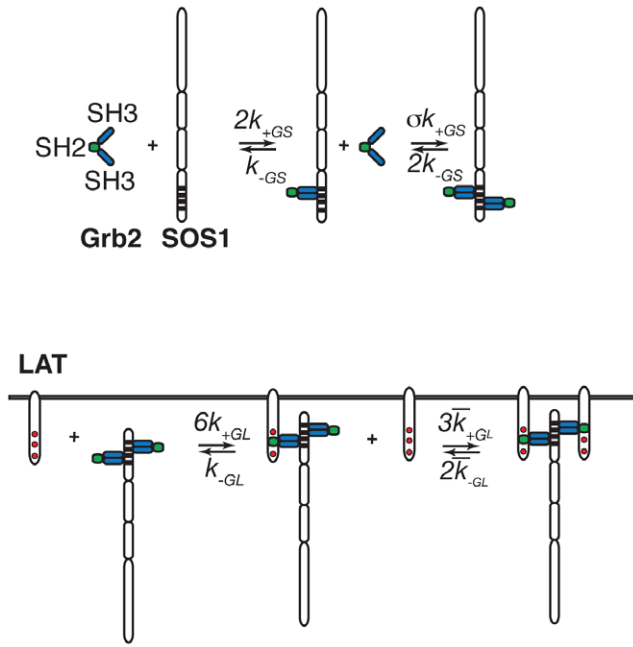


Figure 1. Reactions in the cytosol and just below the plasma membrane. a. The reactions involving the binding SH3 domains of Grb2 to the proline rich domains of SOS1 that result in the formation of a Grb2-SOS1-Grb2 complex. b. and c. Reactions involving the SH2 domain of Grb2 binding to phosphorylated LAT that result in the cross-linking of two trivalent LAT molecules by a Grb2-SOS1-Grb2 complex. doi:10.1371/journal.pone.0028758.g001

For simplicity we treat the three Grb2 binding sites on LAT as if they were identical and noncooperative and take the equilibrium constant for the binding between a LAT phosphotyrosine in a Grb2 binding motif and the SH2 domain of a free Grb2 molecule to be $K_{GL} = (k_{+GL}/k_{-GL})$. We also take K_{GL} to be the same for Grb2, Grb2-SOS1 and Grb2-SOS1-Grb2 binding to a Grb2 binding site on LAT. (In this approximation we are neglecting the observed cooperativity that the binding of SOS1 to Grb2 or SOS1-Grb2 to Grb2 reduces the binding constant for the binding of Grb2 to LAT. For a discussion see Houtman et al. (2007)). We assume that the value of K_{GL} is independent of the valence state of the LAT molecule participating in Grb2 binding, i.e., K_{GL} has the same value irrespective of whether it is a monovalent, bivalent or a trivalent LAT molecule that is binding free Grb2, or Grb2-SOS1, or Grb2-SOS1-Grb2.

As shown in Fig. 1b, the equilibrium cross-linking constant for binding to a phosphotyrosine on a second LAT molecule of the free Grb2 SH2 domain in a Grb2-SOS1-Grb2 species, already bound to a LAT molecule, is taken as $\bar{K}_{GL} = (\bar{k}_{+GL}/\bar{k}_{-GL})$. The value of \bar{K}_{GL} is assumed to be independent of the valency of the two LAT molecules being cross-linked.

Two LAT molecules can be cross-linked by a 2:1 Grb2-SOS1 complex in another way. A 1:1 complex of Grb2 and SOS1 can first bind to a LAT phosphotyrosine and this LAT-bound 1:1 Grb2-SOS1 complex can then form a cross-link with a Grb2 molecule bound to another LAT phosphotyrosine. We denote the forward rate constant for this process by k_{+GS}^{sur} , the reverse rate constant by k_{-GS}^{sur} and the equilibrium constant by $K_{GS}^{sur} = (k_{+GS}^{sur}/k_{-GS}^{sur})$. The principle of detailed balance demands that

$$K_{GS}^{sur} = \sigma \left(\frac{\bar{K}_{GL}}{K_{GL}} \right) K_{GS}. \quad (2)$$

Under the assumptions that $k_{-GS}^{sur} = k_{-GS}$ and $\bar{k}_{-GL} = k_{-GL}$,

$$k_{+GS}^{sur} = \sigma \left(\frac{\bar{k}_{+GL}}{k_{+GL}} \right) k_{+GS}. \quad (3)$$

Since our theory is an equilibrium theory, which predicts the final equilibrium state of the system and not how this state is achieved by the system evolving in time, we do not need to explicitly consider these two different routes of LAT cross-linking in the theoretical derivations that follow. For the purpose of these derivations, we consider only the binding of the free SH2 domain in a LAT-bound Grb2-SOS1-Grb2 to a free phosphotyrosine on a second LAT molecule. On the other hand, for the stochastic simulations discussed later, which follow the evolution of the system in time, we have to explicitly consider the two different routes for LAT cross-linking.

In what follows, we assume that the dimer Grb2-SOS1-Grb2 cannot have its two Grb2 molecules bound simultaneously to the same LAT molecule. More importantly, we exclude all loop formation. If loops can form the number of free sites in an aggregate will be reduced, it will be more difficult to achieve superaggregate formation and the predicted size of the parameter space where a superaggregate can form will be over estimated. Very large aggregates of LAT have been detected in mast cells [27] which suggests that loop formation, if it occurs, is not a dominant effect.

We introduce the following non-dimensional parameters to develop the theory:

$$m = \frac{M}{M_T}, \quad d = \frac{D}{D_T}, \quad l = \frac{L}{L_T}, \quad g = \frac{G}{G_T}, \quad s = \frac{S}{S_T}, \quad (4)$$

$$\alpha = 3\bar{K}_{GL}L_T, \quad (5)$$

$$\epsilon = 2\bar{K}_{GL}D_T, \quad (6)$$

$$\phi = \bar{K}_{GL}M_T, \quad (7)$$

$$c = 2K_{GL}C, \quad (8)$$

$$\chi = K_{GL}G_T, \quad (9)$$

$$\mu = 2K_{GS}S_T, \quad (10)$$

$$\theta = K_{GS}G_T. \quad (11)$$

$$\begin{aligned} \beta &= K_{GL}G + 2K_{GL}K_{GS}GS + 2\sigma K_{GL}K_{GS}^2G^2S \\ &= \chi g + \chi \mu g s + \sigma \chi \mu \theta g^2 s. \end{aligned} \quad (12)$$

$$\Omega = 2\sigma K_{GL} \bar{K}_{GL} K_{GS}^2 G^2 SM = c \bar{K}_{GL} M = \sigma \chi \mu \theta \phi g^2 sm. \quad (13)$$

In our previous paper [30], we derived for a model system comprising only trivalent LAT, Grb2 and SOS1, an expression for the dimensionless partition function q , defined as the sum of the concentrations of all species containing trivalent LAT divided by the total trivalent LAT concentration. This partition function is given by

$$q = w + \frac{1}{\gamma q_0} \left[1 - \frac{1 - (1 - 4\gamma q_0^2)^{\frac{3}{2}}}{6\gamma q_0^2} \right], \quad (14)$$

where w , the concentration of species containing only one trivalent LAT molecule, can be written as $l(1 + \beta)^3$, and q_0 is the linear dimensionless partition function, i.e. it is the sum of concentrations of all possible linear chains of trivalent LAT molecules divided by the total trivalent LAT concentration. We have shown in our previous work [30], that q_0 is given by

$$q_0 = \frac{\frac{3}{2} \delta w^2}{1 - 2\delta w}, \quad (15)$$

where

$$\delta = \frac{\alpha c}{(1 + \beta)^2}, \quad (16)$$

and

$$\gamma = \frac{1}{(3w/2)^2}. \quad (17)$$

In Appendix S1, we show that when the total LAT population is distributed over monovalent, bivalent and trivalent states in contrast to a homogeneous trivalent LAT population in Nag et al. (2009), the dimensionless partition function becomes

$$q = \frac{M_T}{L_T} m(1 + \beta + \Omega/2) + \frac{D_T}{L_T} d(1 + \beta + \Omega)^2 + l(1 + \beta + \Omega)^3 + \frac{1}{\gamma' q_v} \left[1 - \frac{1 - (1 - 4\gamma' q_v^2)^{\frac{3}{2}}}{6\gamma' q_v^2} \right], \quad (18)$$

where

$$\gamma' = \frac{\gamma}{\eta^3}, \quad (19)$$

$$\eta = 1 + \frac{cd}{\alpha l(1 + \beta + \Omega)}, \quad (20)$$

$$q_v = q'_0 \eta^2, \quad (21)$$

$$q'_0 = \frac{\frac{3}{2} \delta' w^2}{1 - 2\delta' w}, \quad (22)$$

and

$$\delta' = \frac{\alpha c}{(1 - cd\epsilon)(1 + \beta + \Omega)^2}. \quad (23)$$

The partition function q is defined as an infinite sum

$$q = \sum_{n=1}^{\infty} w_n, \quad (24)$$

where the surface concentration of species containing n LAT molecules in all possible valence states, normalized by the total concentration of trivalent LAT is denoted as w_n . Expressions for w_n can be obtained by expanding the right hand side of Eq. (18) and summing the terms in that expansion that contain $l^x d^y m^z$ where x, y and z are integers with values greater than or equal to 0, such that $x + y + z = n$. To illustrate, from Eq. (18) we obtained the following expressions for w_n for LAT aggregates up to tetramers, i.e. $n=4$, as:

$$w_1 = l(1 + \beta)^3 + (D_T/L_T)d(1 + \beta)^2 + (M_T/L_T)m(1 + \beta), \quad (25)$$

$$w_2 = (3/2)\alpha c l^2 (1 + \beta)^4 + (3c\epsilon^2 d^2 / 2\alpha)(1 + \beta)^2 + 3cled(1 + \beta)^3 + 3\Omega l(1 + \beta)^2 + (D_T/L_T)2\Omega d(1 + \beta) + (M_T/L_T)(m\Omega/2), \quad (26)$$

$$w_3 = 3\alpha^2 c^2 l^3 (1 + \beta)^5 + (3c^2 \epsilon^3 d^3 / 2\alpha)(1 + \beta)^2 + 6c^2 l\epsilon^2 d^2 (1 + \beta)^3 + (15/2)\alpha c^2 l^2 cd(1 + \beta)^4 + (3c/\alpha)\epsilon^2 d^2 \Omega(1 + \beta) + 6\alpha c l^2 \Omega(1 + \beta)^3 + 9cled\Omega(1 + \beta)^2 + (D_T/L_T)\Omega^2 d + 3\Omega^2 l(1 + \beta), \quad (27)$$

and

$$w_4 = 7\alpha^3 c^3 l^4 (1 + \beta)^6 + (3c^3 \epsilon^4 d^4 / 2\alpha)(1 + \beta)^2 + 10c^3 l\epsilon^3 d^3 (1 + \beta)^3 + 21\alpha^2 c^3 l^3 cd(1 + \beta)^5 + (45/2)(\alpha c^3 l^2 \epsilon^2 d^2 (1 + \beta)^4 + l\Omega^3 + 9\alpha c l^2 \Omega^2 (1 + \beta)^2 + (3c/2\alpha)\epsilon^2 d^2 \Omega^2 + 9cled\Omega^2 (1 + \beta) + 15\alpha c^2 l^2 \Omega(1 + \beta)^4 + (3c^2 \epsilon^3 d^3 \Omega / \alpha)(1 + \beta) + 18c^2 l\epsilon^2 d^2 \Omega(1 + \beta)^2 + 30\alpha c^2 l^2 cd\Omega(1 + \beta)^3). \quad (28)$$

The quantity $W_n = L_T w_n$ is the concentration of all aggregates that contain n LAT molecules, irrespective of their Grb2 valences. This quantity is normalized by the total trivalent LAT concentration to obtain the dimensionless concentration w_n , which is more tractable theoretically. Each separate term in the expression for a given w_n is the normalized concentration for all complexes having the same number of monovalent, bivalent and trivalent LAT. For example, the first term in the expression for w_4 , which is proportional to l^4 , is the normalized concentration of all aggregates of four trivalent LATs. Of the twelve LAT binding sites in these aggregates, six are bound to Grb2-SOS1-Grb2 complexes that cross-linked two LATs while the remaining six are not involved in cross-linking. In general, if there are j binding sites in an aggregate not involved in cross-links, the normalized concentration of the aggregate is proportional to $(1 + \beta)^j$. When $j=0$ all sites on LAT are involved in cross-linking. Two such

aggregates contribute to w_4 , a trivalent LAT bound to three monovalent LAT and an aggregate of two bivalent and two monovalent LAT. Their normalized concentrations are the sixth and eighth terms in the expression for w_4 . Shown in Figure 2 are the aggregates that contribute to the first three terms in w_4 that have only bonds that cross-link two non-monovalent LATs.

Two quantities that will help us characterize the LAT aggregate size distribution are the average number of LATs per aggregate, \bar{s} , and the fraction of the total number of LAT in aggregates, f_{agg} , where

$$\bar{s} = \frac{\sum_{n=2}^{\infty} n w_n}{\sum_{n=2}^{\infty} w_n} = \frac{(L_T + D_T + M_T - w_1 L_T)/L_T}{q - w_1} \quad (29)$$

$$= \frac{1 + D_T/L_T + M_T/L_T - w_1}{q - w_1},$$

and

$$f_{agg} = \frac{L_T + D_T + M_T - w_1 L_T}{L_T + D_T + M_T} \quad (30)$$

$$= (1 + D_T/L_T + M_T/L_T - w_1) \frac{L_T}{L_T + D_T + M_T}.$$

Size distribution of LAT aggregates. The reduced equilibrium concentrations, m , d , l , g and s , and hence the

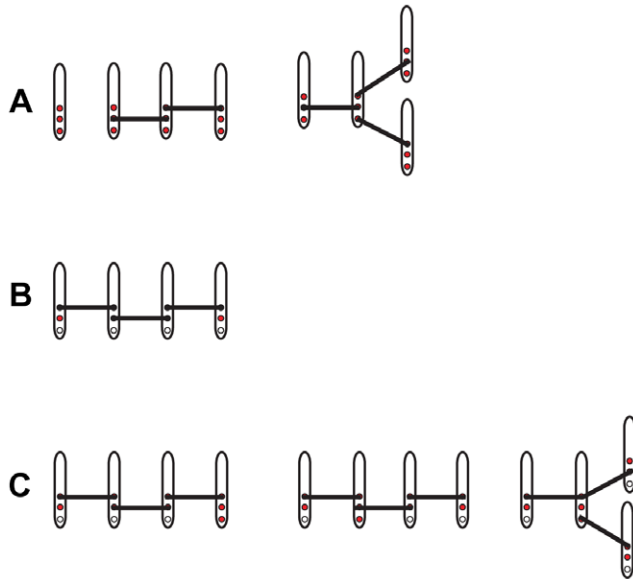


Figure 2. Some of the aggregates that contain four LATs and that only have bonds that participate in cross-linking. The line segment connecting two LAT molecules represents the bivalent ligand Grb2-SOS1-Grb2. **(A)** The total concentration of these two aggregates composed of four trivalent LAT contributes $7x^3c^3l^4$ to the first term of w_4 . The first aggregate in part (A) contributes $6x^3c^3l^4$ and the second aggregate contributes $x^3c^3l^4$ to the first term of w_4 . **(B)** The concentration of this aggregate of four bivalent LAT contributes $3c^3e^4d^4/(2x)$ to the second term of w_4 . **(C)** The total concentration of these three aggregates composed of three bivalent and one trivalent LAT contributes $10c^3l^3d^3$ to the third term of w_4 . The first aggregate in part (C) contributes $3c^3l^3d^3$, the second aggregate contributes $6c^3l^3d^3$ and the third aggregate contributes $c^3l^3d^3$ to the third term of w_4 .

doi:10.1371/journal.pone.0028758.g002

dimensional equilibrium concentrations, can be determined by solving the following conservation equations. These conservation equations are obtained from the partition function, Eq. (18), by noting that the partition function is a sum of normalized concentrations, each term of which is proportional to $m^a d^b l^c g^d s^e$, where the exponents are equal to the number of molecules of the given type in the aggregate, i.e., in an aggregate there would be a monovalent LAT, b bivalent LAT, etc.. For a cell where LAT is distributed on a surface of area A , and the Grb2 and SOS1 that are not associated with LAT are distributed in a volume V , the conservation laws become

$$l \left(\frac{\partial q}{\partial l} \right) = 1, \quad (31)$$

$$\left(\frac{L_T}{D_T} \right) d \left(\frac{\partial q}{\partial d} \right) = 1, \quad (32)$$

$$\left(\frac{L_T}{M_T} \right) m \left(\frac{\partial q}{\partial m} \right) = 1, \quad (33)$$

$$g + \mu g s + \sigma \mu \theta g^2 s + \left(\frac{A L_T}{V G_T} \right) g \left(\frac{\partial q}{\partial g} \right) = 1, \quad (34)$$

and

$$s + 2\theta g s + \sigma \theta^2 g^2 s + \left(\frac{A L_T}{V S_T} \right) s \left(\frac{\partial q}{\partial s} \right) = 1. \quad (35)$$

Once g , s , l , d and m are determined by solving the above conservation equations, one can, for example, evaluate w_n for n between 1 and 4 using Eqs. (25)–(28). The distribution of LAT aggregate sizes can be obtained by plotting $n w_n$ versus n .

Formation of a gel-like phase. Cross-linking of bivalent LAT molecules by a bivalent ligand results in only linear aggregates and cannot yield a gel like state [31]. It has been shown in Nag et al. (2009), that for trivalent LAT, where branched aggregates can form, the possibility arises that a gel-like state, i.e., a superaggregate, can form. In the course of aggregation of trivalent LAT, a point may be reached at which the probability of having an infinite sized aggregate changes from zero to a positive value [31]. This point is called the gel point. We currently investigate the conditions that the gel point should satisfy in the presence of multiple valence states of LAT.

We define M_1 and M_2 as the first and second moments respectively of the LAT aggregate distribution. Then,

$$M_1 = \sum_{n=1}^{\infty} n w_n = l \frac{\partial q}{\partial l} + d \frac{\partial q}{\partial d} + m \frac{\partial q}{\partial m}, \quad (36)$$

and,

$$M_2 = \sum_{n=1}^{\infty} n^2 w_n \quad (37)$$

$$= \left(l \frac{\partial}{\partial l} + d \frac{\partial}{\partial d} + m \frac{\partial}{\partial m} \right) \left(l \frac{\partial q}{\partial l} + d \frac{\partial q}{\partial d} + m \frac{\partial q}{\partial m} \right).$$

The first moment is the fraction of trivalent LAT in finite-sized aggregates. In the absence of infinite aggregates $L_T M_1 = L_T + D_T + M_T$ is the conservation law for LAT. The second moment is the average size of all entities containing trivalent LAT. At the gel point, M_1 equals $(L_T + D_T + M_T)/L_T$ but M_2 diverges indicating the presence of superaggregates. From the validity of the conservation equations Eqs. (31)–(35) and the divergence of M_2 at the gel point, we show in Appendix S2 that at the gel point, the following equation is valid as well.

$$\frac{1 - cd}{2\alpha c l(1 + \beta + \Omega)} = 1 + \left[1 + \frac{cd}{\alpha l(1 + \beta + \Omega)}\right]^{1/2}. \quad (38)$$

For given total numbers of Grb2, SOS1 and monovalent LAT molecules in the cell, we can simultaneously solve Eqs. (31)–(35), (38) to obtain the two phase (sol-gel) region in the parameter space spanned by the number of bivalent LAT molecules and the number of trivalent LAT molecules in the cell.

Alternatively, for given total numbers of Grb2, monovalent and bivalent LAT molecules present in the cell, we can obtain the two phase region in the parameter space spanned by the number of SOS1 molecules and the number of trivalent LAT molecules in the cell.

Stochastic Simulations

We have adapted a rule-based kinetic Monte Carlo (kMC) approach [32,34] to numerically simulate LAT-GRB2-SOS1 interactions. The main advantage provided by the rule-based kMC over standard methods for simulation stochastic dynamics of biochemical systems [35,36] is the avoidance of the combinatorial explosion of the number of possible species and reactions that arises from the interactions of multivalent molecules [37,38]. Particularly, conventional chemical kinetics methods could not be used for dynamical simulation of LAT aggregation in the gel phase regime since the number of equations that would need to be integrated would be effectively infinite. An overview and a detailed description of the method is given in Nag et al. (2009).

The main assumption underlying the kMC model of Nag et al. (2009) is that the binding kinetics at free LAT sites does not depend on the occupancy of other sites on the same LAT molecule. This assumption allows us to condense the reaction network of all possible configurations to a much smaller network of reactants as no local constraints on binding to LAT need to be taken into account. Thus, each binding interaction involving LAT requires sampling only from an array of free LAT sites; complexes of multiple LAT molecules are included automatically. Note that the model of Nag et al. (2009) considers LAT molecules of valence three only. Here, to include variability in the valence of LAT, we modify the data structure of the kMC model of Nag et al. (2009). We create an array of AM_T sites of monovalent LAT molecules, $2AD_T$ sites of bivalent LAT molecules and $3AL_T$ sites of trivalent LAT molecules. By definition, each LAT site is associated with a specific LAT molecule. Therefore, we also create an array of indexed LAT molecules, where molecule index i is used to determine the valence of LAT in accordance with the following rule: if $0 \leq i < AM_T$, the i th molecule is monovalent; if $AM_T \leq i < (AM_T + AD_T)$, the i th molecule is bivalent; if $(AM_T + AD_T) \leq i < (AM_T + AD_T + AL_T)$, the i th molecule is trivalent. Simulation procedure of the modified kMC model follows the steps described in Nag et al. (2009).

We use the simulations to track the aggregate size distribution in time. As discussed earlier in the Theory section, we explicitly consider in the simulations, two different ways by which a cross-

link can form between two LAT molecules, characterized by two different sets of rate constants - i) $\bar{k}_{+GL}, \bar{k}_{-GL}, k_{+GS}, k_{-GS}$ and ii) $k_{+GL}, k_{-GL}, k_{+GS}^{sur}, k_{-GS}^{sur}$. The final output is obtained by averaging over multiple runs. The system size is set to the total number of molecules, and it can also be rescaled by multiplying bimolecular rate constants by a volume factor, F_x , and dividing AL_T, AD_T, AM_T, VG_T and VS_T by F_x .

Results

In this section, we present results obtained from the equilibrium model and stochastic simulations. For convenience in representation, the terms L_T, D_T, M_T, G_T and S_T in the current section and in the related figures are used to refer to the total numbers per cell of trivalent LAT, bivalent LAT, monovalent LAT, Grb2 and SOS1 respectively and not to their concentrations as in the Methods section. The values of kinetic and other model parameters used in the equilibrium model calculations and in the stochastic simulations have been provided in Table 1.

Size distribution of LAT aggregates

Shown in Fig. 3, for two different LAT populations, are the predicted LAT aggregate size distributions from simulation and from the equilibrium theory developed in the previous section. The predictions from simulation and theory for these values of Grb2, SOS1 and LAT are indistinguishable. For both curves the Grb2 and SOS1 populations were set equal to 7.5×10^5 and 3.75×10^5 per cell, respectively, values typical for Jurkat T cells [30]. Unless stated, these are the values we will use in the examples discussed in this section. The solid curve corresponds to a heterogeneous LAT population of 6×10^4 per cell, equally divided among monovalent, bivalent and trivalent LAT. The dotted curve corresponds, to a homogeneous population of 2×10^4 trivalent LAT so that in both cases the concentration of trivalent LAT is the same. From Eqs. (29) and (30) we calculate that for the heterogeneous population of LAT the average LAT aggregate size, $\bar{s} = 3.06$ and the fraction of LAT in aggregates of size five or greater is $w_{\geq 5} = 0.19$. For the homogeneous population of trivalent LAT, $\bar{s} = 3.25$ and $w_{\geq 5} = 0.23$. As we see in this example, the presence of monovalent and bivalent LAT (solid line) modestly increases the amount of LAT not in aggregates ($n = 1$ in Fig. 3) and decreases the amount of LAT in larger aggregates ($n \geq 4$).

Formation of macromolecular aggregates

In this section, we explore the regions of concentration space that allow for the formation of a gel phase where super aggregates can form. In Nag et al. (2009), where we considered only homogeneous LAT populations, we were able to obtain an analytic expression for the fraction f_g of LAT molecules in the gel phase. For a heterogeneous population of LAT the problem is more difficult and we have instead used stochastic simulations to calculate f_g for specific values of the Grb2, SOS1 and LAT concentrations. In the simulations we define the number of LAT molecules in the gel phase as the number of LAT molecules of all three valencies in the largest aggregate averaged over time, once equilibrium has been established. To speed the simulations we reduced the system size, the volume and all the molecular populations, by a factor of 10 ($F_x = 10$). This scaling reduced the size of the superaggregate by a factor of ten compared to the full system, but did not affect the predicted fraction of LAT in the gel phase.

In Fig. 4a we consider three different total concentrations of monovalent LAT with M_T ranging from 2.1×10^4 to 8.4×10^4 molecules/cell. Keeping M_T fixed at each concentration, we

Table 1. Parameters used in the simulations and the equilibrium model calculations.

Parameter	Value(s)
σ	0.5
k_{+GL}	$1.05 \times 10^6 M^{-1} s^{-1}$
k_{-GL}	$0.31 s^{-1}$
$K_{GL} = \frac{k_{+GL}}{k_{-GL}}$	$3.3871 \times 10^6 M^{-1}$
\bar{k}_{+GL}	$5.25 \times 10^{14} mole^{-1} cm^2 s^{-1}$
\bar{k}_{-GL}	$0.31 s^{-1}$
$\bar{K}_{GL} = \frac{\bar{k}_{+GL}}{\bar{k}_{-GL}}$	$1.69355 \times 10^{15} mole^{-1} cm^2$
k_{+GS}	$9.45 \times 10^4 M^{-1} s^{-1}$
k_{-GS}	$0.034412 s^{-1}$
$K_{GS} = \frac{k_{+GS}}{k_{-GS}}$	$2.74614 \times 10^6 M^{-1}$
k_{+GS}^{sur}	$\sigma \times 4.725 \times 10^{13} mole^{-1} cm^2 s^{-1}$
k_{-GS}^{sur}	$0.034412 s^{-1}$
$K_{GS}^{sur} = \frac{k_{+GS}^{sur}}{k_{-GS}^{sur}}$	$\sigma \times 1.37307 \times 10^{15} mole^{-1} cm^2$
V_{Jurkat}	$3.58348 \times 10^{-10} cm^3$
A_{Jurkat}	$8.0 \times 10^{-6} cm^2$

This table furnishes values of parameters used in the simulations and the equilibrium model calculations. σ is a negative cooperativity factor for the binding of a second Grb2 to a Grb2-SOS1 complex (see Eq. (1)) [6]. The equilibrium constants for the binding of Grb2 to one of the three terminal phosphotyrosines on LAT range from $1 - 4 \times 10^6 M^{-1}$ [6,33,45]. In our model the affinities for these three binding sites are identical. We take $K_{GL} = 3.39 \times 10^6 M^{-1}$. The values for k_{+GL} and k_{-GL} are from biacore studies of the binding of the SH2 domain in a Grb2-SOS1 complex to an eleven peptide sequence from the cytoplasmic domain of EGFR that includes the Grb2 binding site pY1068 [48]. The values of these rate constants have no effect on any of the equilibrium results. The value of the equilibrium crosslinking constant \bar{K}_{GL} is estimated from the K_{GL} value as in Nag et al. (2009). The rate constant \bar{k}_{+GL} is calculated using the values of \bar{K}_{GL} and \bar{k}_{-GL} which we assume equals k_{-GL} . The value of k_{+GS} is taken from [49]. The dissociation constants for the binding of the Grb2 SH3 domain to the N-terminal and C-terminal proline-rich regions of SOS1 are 260 nM and 510 nM respectively [6]. In our model we do not distinguish between the two SH3 binding sites on SOS1 and take $K_d = 364.143 nM$, the geometric mean of the two values. The value of k_{-GS} is calculated using the value of k_{+GS} and the geometric mean of the K_d values. The value of k_{+GS}^{sur} is estimated using Eq. (3) and taking $k_{-GS}^{sur} = k_{-GS}$. The diameter of the Jurkat cell, $2a$, has been measured by Rosenbluth et al. [50] to be 11.5 μm with the cytosol taking up about 45% of the total cell volume. $V_{Jurkat} = 0.45 \times 4\pi a^3 / 3$ is the cytosolic volume. The surface area of the Jurkat cell is taken to be approximately twice the area of a sphere of diameter 11.5 μm . In the stochastic simulations, only unimolecular rate constants can be used directly so the solution bimolecular rate constants are scaled by the cytosolic volume V_{Jurkat} and the surface bimolecular rate constants are scaled by A_{Jurkat} . The exact values of the parameters we used are given in the table to allow our results to be reproduced. The accuracy of the parameters that have been determined by experiment is at best two significant figures.
doi:10.1371/journal.pone.0028758.t001

obtain from the equilibrium theory, Eqs. (31)–(35), (38), the boundary of the sol-gel region in the concentration space spanned by the total bivalent and trivalent LAT concentrations, D_T and L_T . As expected, since monovalent LAT acts as a chain terminator, the higher the concentration of monovalent LAT, the smaller the sol-gel region in the D_T - L_T concentration space.

For each D_T value in the two phase region, there exist two critical L_T values which mark the onset of gel formation. Superaggregates cannot form below the lower value L_T^l and above the upper value L_T^u . Below L_T^l , the total concentration of

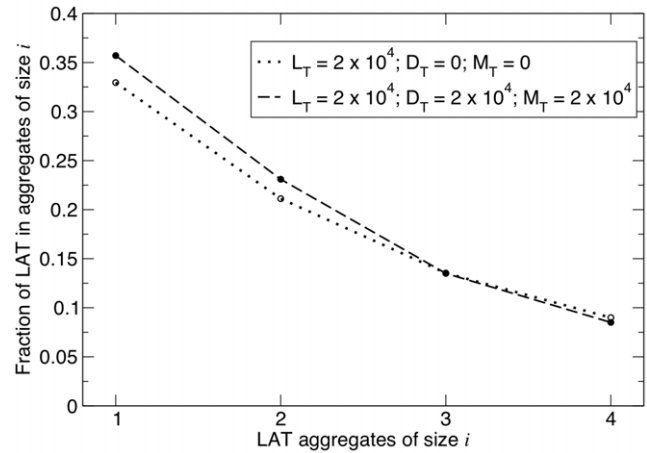


Figure 3. Comparison of size distribution of LAT aggregates for a homogeneous (open circles) and heterogeneous population of LAT (solid circles). For both homogeneous and heterogeneous LAT populations, the number of trivalent LAT is identical and equal to 2×10^4 per cell, but for the heterogeneous population of LAT also present are 2×10^4 monovalent LAT and 2×10^4 bivalent LAT. For both cases the numbers per cell of Grb2 and SOS1 were 7.5×10^5 and 3.75×10^5 respectively. The simulation results were obtained from a single simulation run with $F_x = 1$. The theoretical results were obtained by solving Eqs. (31)–(35) simultaneously for g, s, l, d and m and then calculating w_n from Eqs. (25)–(28). To obtain the fraction of LAT present in an aggregate of a particular size, w_n has to be multiplied by $L_T / (L_T + D_T + M_T)$ since w_n is obtained by dividing W_n by L_T . The open circles connected by dotted lines is for the homogeneous LAT population.
doi:10.1371/journal.pone.0028758.g003

trivalent LAT relative to the total concentrations of Grb2, SOS1, bivalent and monovalent LAT is insufficient to yield a superaggregate. The larger is D_T , the greater is the number of trivalent LAT molecules necessary for the onset of the gel formation, so that L_T^l increases with D_T . Above L_T^u , the number of trivalent LAT molecules is so high that there are not enough cross-linking Grb2-SOS-Grb2 species for superaggregate formation. This shortage is accentuated by the presence of bivalent and monovalent LAT competing with trivalent LAT for the cross-linking species, resulting in L_T^l decreasing with D_T .

In Fig. 4b, for each of the three M_T values in Fig. 4a, we have plotted f_g obtained from simulations as function of L_T . In the simulations, D_T was fixed at 10^5 per cell and G_T and S_T held at their usual values. The vertical lines correspond to the predicted boundaries from the equilibrium theory, i.e., the values of L_T at which f_g goes to zero. The f_g values obtained from the simulations have small but finite non-zero values at the predicted boundaries. This discrepancy arises because the equilibrium theory assumes the size of the system is infinite whereas the simulations are for a finite system. For a finite system there is no true transition from sol to sol plus gel and this is reflected in the way the simulated f_g curves approach the x-axis. Since the simulations were done on a system whose volume was reduced by a factor of ten this discrepancy is accentuated.

In Fig. 4c, we plot the f_g values as a function of D_T , keeping L_T fixed at 1.5×10^5 per cell and G_T , S_T and M_T at the same values as in Fig. 4b. One phase boundary exists for such D_T variation for each fixed M_T value. Again, the horizontal lines in Fig. 4c indicate the predicted boundaries from the equilibrium theory for the three fixed M_T values in Fig. 4a. As in Fig. 4b, the simulated values for f_g are small but not zero at these values, presumably due to finite size effects.

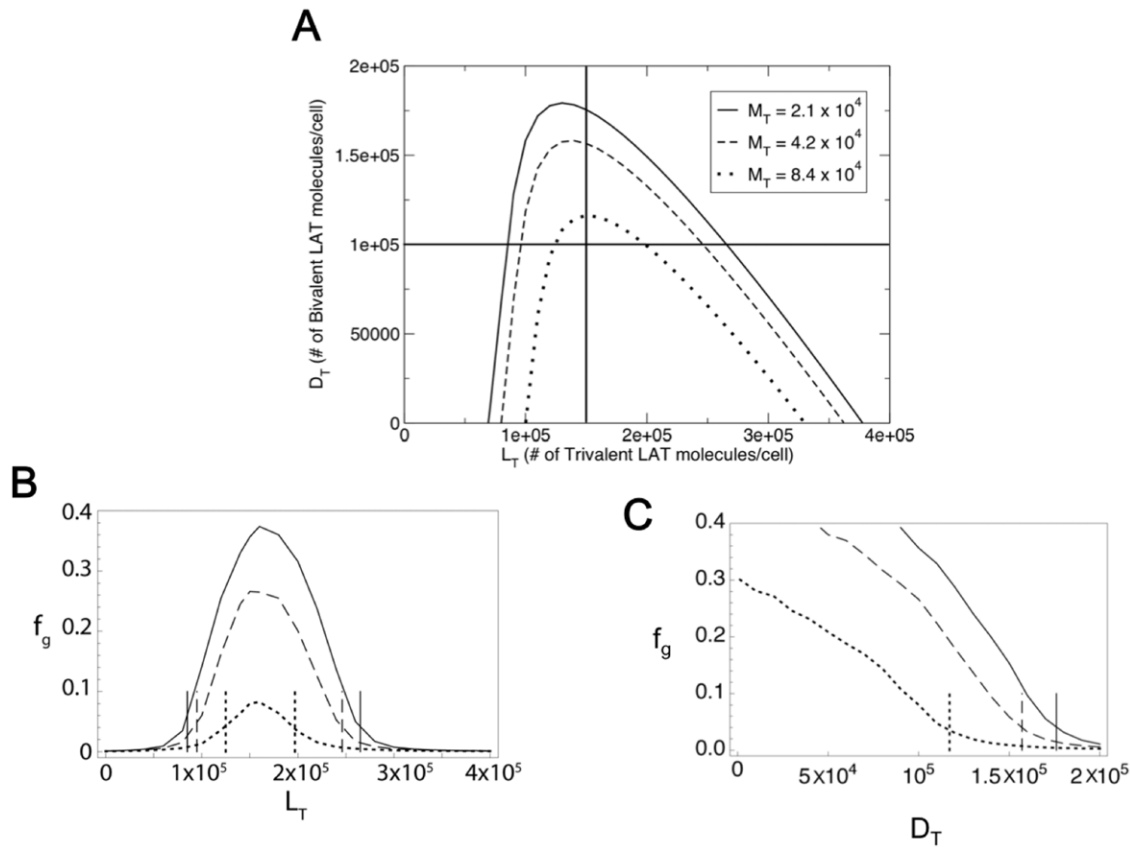


Figure 4. Two phase regions in space spanned by bivalent and trivalent LAT concentrations. (A) Two phase(sol-gel) regions for Grb2 and SOS1 cellular populations of 7.5×10^5 and 3.75×10^5 molecules per cell respectively. The population of monovalent LAT is fixed at three different values, 2.1×10^4 molecules per cell (solid line), 4.2×10^4 molecules per cell (dashed line) and 8.4×10^4 molecules per cell (dotted line). The solid horizontal and vertical lines correspond to the cellular populations of bivalent LAT (D_T) and trivalent LAT (L_T) being fixed at 10^5 and 1.5×10^5 molecules per cell respectively. (B) Plot of f_g , the fraction of the cellular inhomogeneous LAT population present in the gel phase, as a function of the cellular trivalent LAT population L_T , as predicted from stochastic simulations. The cellular bivalent LAT population (D_T) is fixed at 10^5 molecules per cell. In sync with (A), the solid, dashed and dotted lines correspond to fixed cellular monovalent LAT populations (M_T) of 2.1×10^4 , 4.2×10^4 and 8.4×10^4 molecules per cell respectively, and cellular Grb2 and SOS1 populations are fixed at the same values as in (A). The vertical lines correspond to the theoretically predicted boundaries between the one phase (sol) and two-phase (sol-gel) regions obtained by varying L_T along the horizontal line in (a) corresponding to a fixed D_T value of 10^5 molecules per cell. (C) Plot of simulated f_g values as a function of D_T keeping L_T fixed at 1.5×10^5 molecules per cell, and Grb2 and SOS1 fixed at the same values as in (A), for each of the three fixed monovalent LAT populations (M_T) in (A). The vertical lines correspond to the theoretically predicted boundaries between the one phase (sol) and two-phase (sol-gel) regions obtained by varying D_T along the vertical line in (A) corresponding to a fixed L_T value of 1.5×10^5 molecules per cell.
doi:10.1371/journal.pone.0028758.g004

We also examine the sol-gel region in the concentration space spanned by the total SOS1 and trivalent LAT concentrations as we previously did in Nag et al. (2009) for the homogeneous case. In Fig. 5 we fix the total Grb2 and monovalent LAT concentrations at 7.5×10^5 and 4.2×10^4 per cell respectively and consider three cases corresponding to three fixed D_T values in the range 4.0×10^4 – 10^5 per cell. The boundaries of the one phase (sol) and two-phase (sol-gel) regions for these three cases, Fig. 5a, were obtained using the equilibrium theory. As expected since bivalent LAT blocks branching, the higher is the fixed D_T value, the smaller is the two-phase region.

In Fig. 5b, values of f_g obtained from simulations are plotted as a function of L_T , keeping S_T , G_T and M_T fixed at 4×10^5 , 7.5×10^5 and 4.2×10^4 molecules/cell respectively for each of the three fixed D_T values in Fig. 5a. In Fig. 5c, the f_g values are plotted as a function of S_T keeping L_T fixed at 1.5×10^5 and G_T , M_T and D_T at the same values as in Fig. 5b. The horizontal lines indicate the theoretically predicted boundaries between the one phase and two-phase regions for the different D_T values in Fig. 5a.

As in Figs. 4b and 4c, lack of consideration of finite size effects in the theoretical approach results in discrepancies between the theoretically predicted and simulated f_g values at the above-mentioned boundaries.

To get a better feeling for how sensitive LAT superaggregate formation is to the presence of monovalent and bivalent LAT, we consider in Fig. 6, a fixed population of phosphorylated LAT ($M_T + D_T + L_T = 3.45 \times 10^5$ per cell) and show the set of M_T , D_T and L_T values for which a superaggregate can form. The dotted line in Fig. 6 is the boundary between the one-phase (sol) and the two-phase (sol gel coexistence) regions in the concentration space spanned by monovalent and bivalent LAT. The origin ($M_T = 0$, $D_T = 0$, $L_T = 3.45 \times 10^5$ per cell) is in the sol-gel coexistence region. In the absence of monovalent and bivalent LAT, for the values of Grb2 and SOS1 used in Fig. 6, a sol-gel coexistence region is predicted to exist for $5.758 \times 10^4 < L_T < 3.926 \times 10^5$. For a mixture of only monovalent and trivalent LAT (the x-axis), when the monovalent LAT is greater than 54% of the total phosphorylated LAT (point A in Fig. 6), corresponding to

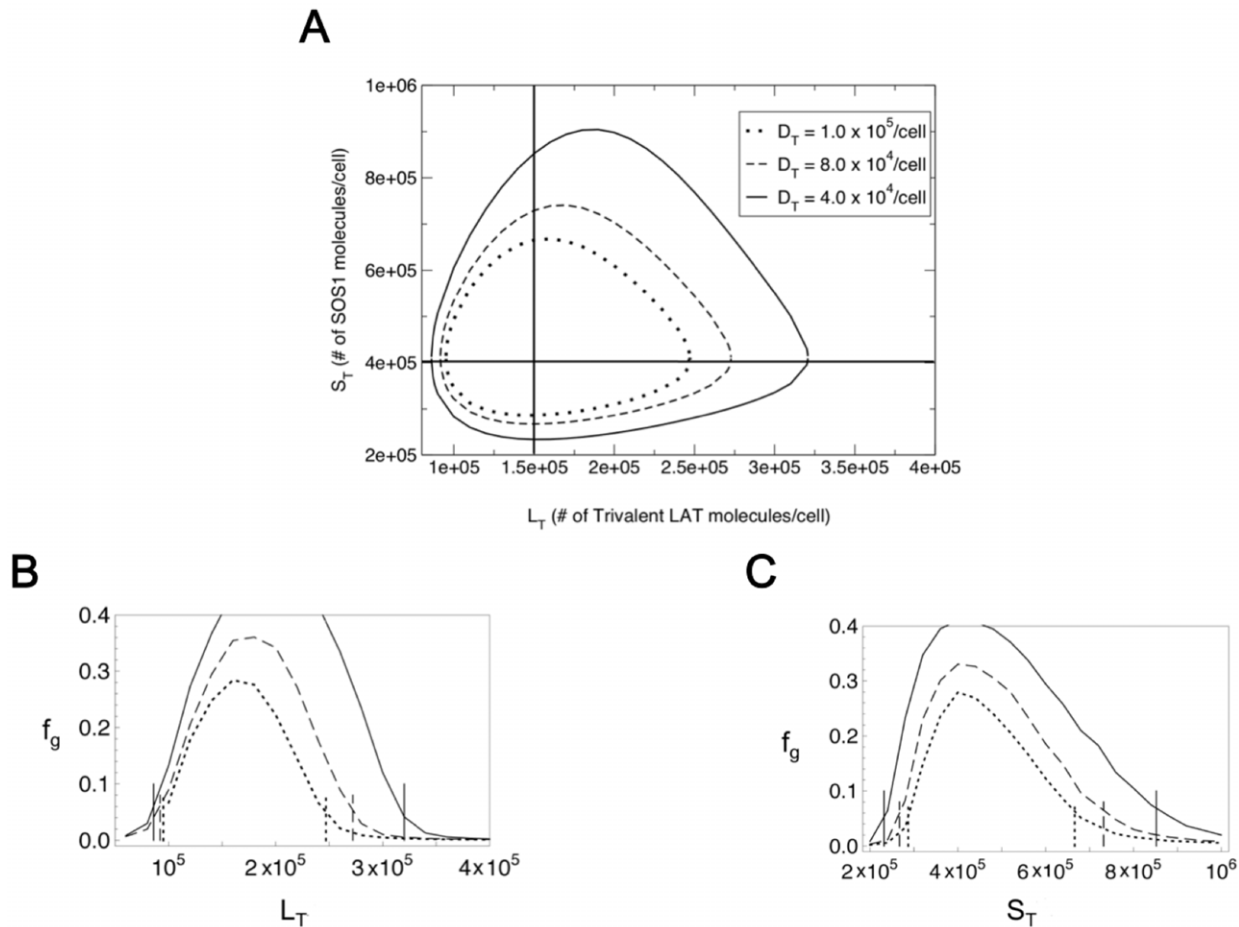


Figure 5. Two phase regions in space spanned by SOS1 and trivalent LAT concentrations. (A) Two phase(sol-gel) regions for Grb2 and monovalent LAT cellular populations of 7.5×10^5 and 4.2×10^4 molecules per cell respectively. The population of bivalent LAT is fixed at three different values, 4×10^4 molecules per cell (solid line), 8×10^4 molecules per cell (dashed line) and 1.0×10^5 molecules per cell (dotted line). The solid horizontal and vertical lines correspond to the cellular populations of SOS1 (S_T) and trivalent LAT (L_T) being fixed at 4×10^5 and 1.5×10^5 molecules per cell respectively. (B) Plot of f_g , the fraction of the inhomogeneous LAT population present in the gel phase, as a function of the trivalent LAT population L_T , as predicted from stochastic simulations. The cellular SOS1 population is fixed at 4×10^5 molecules per cell. As in (A), the solid, dashed and dotted lines correspond to fixed bivalent LAT populations (D_T) of 4×10^4 , 8×10^4 and 1.0×10^5 molecules per cell, and Grb2 and monovalent LAT cellular populations are fixed at the same values as in (A). The vertical lines correspond to the theoretically predicted boundaries between the one phase (sol) and two-phase (sol-gel) regions obtained by varying L_T along the horizontal line in (A) corresponding to a fixed S_T value of 4×10^5 molecules per cell. (C) Plot of simulated f_g values as a function of S_T keeping L_T fixed at 1.5×10^5 molecules per cell, and Grb2 and monovalent LAT fixed at the same values as in (A). The vertical lines correspond to the theoretically predicted boundaries between the one phase (sol) and two-phase (sol-gel) regions obtained by varying S_T along the vertical line in (A) corresponding to a fixed L_T value of 1.5×10^5 molecules per cell. doi:10.1371/journal.pone.0028758.g005

$L_T = 1.59 \times 10^5$ per cell, superaggregate formation is blocked. Similarly, for a mixture of only bivalent and trivalent LAT (the y-axis), superaggregate formation is blocked when the bivalent LAT concentration is greater than 56% of the total LAT (point B in Fig. 6) corresponding to $L_T = 1.52 \times 10^5$ per cell. Monovalent LAT, which terminates aggregate formation, is only slightly more effective in inhibiting superaggregate formation than bivalent LAT, which blocks branching within aggregates, but not the propagation of aggregation. The results shown in Fig. 6 indicate that superaggregate formation can occur in the presence of substantial heterogeneity in the valence of LAT. For this example, superaggregate formation occurs when the sum of the concentrations of monovalent and bivalent LAT is less than 54%.

Discussion

In T cells and mast cells the scaffolding protein LAT links receptors to downstream signaling events [8,29]. LAT has three

terminal tyrosines, each of which when phosphorylated becomes a binding site for the adaptor protein Grb2, which through its SH2 domain, binds to phosphorylated LAT and through its two SH3 domains binds to the proline-rich regions of SOS1 [6]. Two Grb2 molecules can bind to SOS1 to form a Grb2-SOS1-Grb2 dimer, which can bridge two tyrosine phosphorylated LAT molecules. LAT is phosphorylated by Syk family kinases that have been recruited to the cytoplasmic domains of receptors containing phosphorylated immunoreceptor tyrosine-based activation motifs (ITAMs). Thus, extracellular stimuli, by controlling the level of activation of Syk family kinases can modulate the valence of LAT for Grb2 from zero to three. Large aggregates of LAT have been observed in T cells following T cell stimulation with anti-CD3 antibodies [6,26] and in mast cells following the aggregation of IgE-Fc ϵ RI complexes [27,28].

Previously we showed that in the absence of monovalent and bivalent LAT, for a certain range of concentrations of Grb2, SOS1 and trivalent LAT, an equilibrium theory of aggregate formation

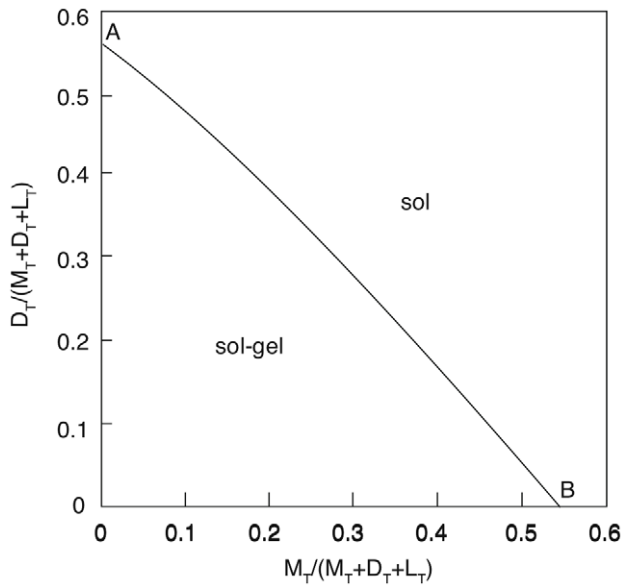


Figure 6. Phase diagram in the concentration space spanned by bivalent and monovalent LAT concentrations. The total LAT (monovalent+bivalent+trivalent) concentration is fixed at 3.45×10^5 molecules/cell and the Grb2 and SOS1 concentrations are 7.5×10^5 molecules/cell and 3.75×10^5 molecules/cell respectively. doi:10.1371/journal.pone.0028758.g006

predicted the Grb2-SOS1-Grb2 mediated formation of a LAT superaggregate [30]. The range of concentrations over which superaggregate formation was predicted to occur corresponded to a sol-gel coexistence region where the sol phase consisted of monomers and small aggregates of LAT, and the gel phase consisted of a single superaggregate of LAT.

In this paper we have presented an equilibrium theory for Grb2-SOS1-Grb2 mediated LAT oligomerization when the LAT population is heterogeneous with respect to its valence for Grb2, extending the equilibrium theory for Grb2-SOS1-Grb2 mediated oligomerization of LAT molecules which are homogeneous with respect to their valence for Grb2 [30]. We have adapted a kinetic Monte Carlo method [32,34] used in our previous studies [30] to simulate the model and assess the importance of finite size effects. These effects were small and the simulations were in good agreement with predictions of the equilibrium theory (see Fig. 3). In the absence of experimentally derived ratios of mono-, bi- and trivalent LAT, we made *ad hoc* but biologically reasonable assumptions about the relative proportions of the different valence states of LAT. We intend to provide an estimate for the relative levels, in stimulated mast cells, of the various LAT valence states in a forthcoming publication (Monine et al. unpublished), using a computational model of FcεRI-mediated early signaling events in mast cells that would include LAT phosphorylation by activated Syk kinase and LAT dephosphorylation by a pool of phosphatases.

The question that most concerned us was, when we include mono- and bivalent LAT in the presence of trivalent LAT, do we still predict there to be a wide range of concentrations for which superaggregate formation can occur? Monovalent LAT terminates the growth step in aggregate formation while bivalent LAT, although it allows for chain elongation, blocks new branch formation. As expected, in all the cases we considered, the presence of monovalent and bivalent LAT reduced the size of the sol-gel coexistence region. In Fig. 4a, for concentrations of Grb2 and SOS1 found in Jurkat T cells, we have mapped the predicted

boundaries of the coexistence region in the space spanned by the bivalent and trivalent LAT concentrations for a set of monovalent LAT concentrations. As can be seen, adding monovalent and bivalent LAT reduces the size of the coexistence region but there is still a substantial range of concentrations where a superaggregate (Fig. 4a) can form that contain significant fractions of the total LAT (Fig. 4b and c). In Fig. 5a we mapped the predicted boundaries of the sol-gel coexistence region in the space of SOS1 and trivalent LAT concentrations, for a set of fixed bivalent LAT concentrations and a single monovalent LAT concentration. Again, although the size of the coexistence region is reduced as the concentration of bivalent LAT is increased, there is still a wide range of concentrations within the boundary of the sol-gel region.

During the initial phase of signaling we expect the distribution of the LAT population over the different valence states to vary with the strength and duration of the external stimulus. In Fig. 6 we consider a fixed concentration of phosphorylated LAT, 3.45×10^5 per cell, and Grb2 and SOS1 concentrations typical of a Jurkat T cell. The value of LAT was chosen so that when all the LAT is trivalent, a substantial fraction of it is predicted to be in a superaggregate. When there is no bivalent LAT present, and we increase monovalent LAT while reducing trivalent LAT, we continue to get gel formation until the concentration of monovalent LAT exceeds 54% of the total. When there is no monovalent LAT present, and we increase bivalent LAT while reducing trivalent LAT, gel formation is completely inhibited when the bivalent LAT exceeds 56% of the total. Surprisingly, monovalent LAT is only slightly more effective than bivalent LAT in reducing the sol-gel coexistence region. In order for superaggregate formation to occur there must be a substantial concentration of trivalent LAT. Although the presence of monovalent and bivalent LAT reduces the sol-gel coexistence region, the set of concentrations of Grb2, SOS1, mono-, bi- and trivalent LAT for which superaggregate formation is predicted remains large and includes concentrations we expect occur in activated Jurkat T cells.

Sol-gel coexistence may have been observed in stimulated RBL-2H3 cells. Using immunogold labeling of mast cell membrane sheets in conjunction with transmission electron microscopy, Wilson et al. (2001a) determined the distribution of LAT clusters formed by cross-linking of FcεRI-bound IgE by an added multivalent antigen (see Fig. 8 in [27]). A minute after the stimulation, a distribution of LAT aggregate sizes was observed with one aggregate much larger than the rest. The outlying aggregate contained over 160 LAT molecules at two minutes after stimulation. This was a lower bound on the number of LAT molecules in the aggregate since the efficiency of labeling was not determined. RBL-2H3 cells have been estimated to contain 1.1×10^6 LAT molecules, although it is not known if all of these are on the plasma membrane. The equilibrium theory predicts that in the sol-gel coexistence region a significant fraction of the total LAT should be in a superaggregate. If the entire cell surface was a single well mixed system, we would therefore expect to see hundreds of thousands rather than hundreds of LAT molecules in a single aggregate. One way for our predictions to be consistent with the observations of Wilson et al. [27] would be if the size of the region where LAT can be treated as well mixed is about a thousand times smaller than the surface area of the cell. The concept of a cell surface being made up of many microdomains, is consistent with single particle tracking experiments on RBL-2H3 cells that show diffusing receptors confined for many seconds in regions with a characteristic length scale of microns [39].

In addition to restrictions imposed by cell surface domains, active mechanisms reduce aggregate size. The aggregation of LAT

is crucial for the propagation of downstream signaling [6], but as signaling propagates, active mechanisms arise to attenuate signaling. On T cells, shortly after receptor phosphorylation and LAT aggregation, LAT is internalized through a mechanism involving the ubiquitin ligase c-Cbl [40]. LAT internalization requires the presence of c-Cbl, although not the ubiquitylation of LAT, while LAT degradation requires LAT ubiquitylation [41]. Like SOS1, c-Cbl can bind to Grb2 SH3 domains through its proline-rich domains and form Grb2-Cbl-Grb2 complexes that can be incorporated into LAT aggregates [6]. If c-Cbl is in low concentration compared to SOS1, c-Cbl will be preferentially incorporated into large aggregates compared to small aggregates. If two LAT molecules in an aggregate bound to a Grb2-Cbl-Grb2 complex can mediate the internalization of an entire aggregate, c-Cbl will efficiently reduce the concentration of surface LAT and limit the growth of large aggregates. Incorporation of LAT internalization reactions into our existing model would serve to lower the size of the superaggregate and might result in better agreement between the sizes of the theoretically predicted and experimentally observed superaggregates.

A major assumption of our current model is that we can ignore binding reactions that lead to the formation of closed structures (loops). Assuming we allow loop formation in our model, there would be a number of conceivable ways to form closed structures. One possible way loops can form would involve binding of the two unbound SH2 domains of the Grb2-SOS1-Grb2 cross-linker to two different phosphotyrosine groups on the same LAT molecule. Such binding events which are feasible for only trivalent LAT and bivalent LAT molecules would convert them to monovalent and zero-valent LAT species respectively. Thus LAT molecules that might participate in aggregate formation by increasing a chain length or by seeding a branch would be converted to species which would either participate in chain termination or not participate in aggregate formation at all. Another way of formation of closed structures is the formation of cross-links between two LAT molecules which are already cross-linked. Formation of such loops would involve the consumption of Grb2-SOS1-Grb2 cross-linkers and unbound LAT phosphotyrosines which could otherwise be instrumental in enhancing the aggregate size. An additional way by which loops can form is the cross-linking of the two ends of a linear or branched chain of LAT molecules which would also result in the non-productive use of cross-linkers and LAT phosphotyrosines so far as increasing the aggregate size is concerned. Hence, it is expected that the addition of loop formation reactions to our existing equilibrium model will limit the size of the superaggregate and reduce the discrepancy between the theoretically predicted and experimentally observed superaggregate sizes. Although we could incorporate certain specific types of loops into the model, experiments to date offer little guidance as to what types of loops are most likely to occur. We expect very large loops to be unlikely because of the multiple ways in which they can open and because the probability of a chain closing to form a loop decreases with increasing chain length. As we have pointed out, we don't expect small loops to have a high probability of forming because if they did, we would not expect to see the observed formation of large aggregates of LAT [27]. This leaves us with a large poorly defined set of possible loop closures to consider. We feel it is premature to try to add loops to the model until there is

data to show that such structures do form during the aggregation of LAT.

The only adaptor molecule in our model that can form a dimer via binding to SOS1 is Grb2. However, at least another adaptor Grap (Grb-2-like accessory protein) is expressed in T cells, which like Grb2 is composed of one central SH2 domain flanked by two SH3 domains [8,42,43]. Grap not only binds the same sites on LAT as GRB2 through its central SH2 domain but also the proline-rich regions of SOS1 through at least one of its two SH3 domains [42,43] and is hence capable of forming a 2:1 complex with SOS1, although no experimental evidence to this effect has been reported. A third member of the Grb2 family of adaptor proteins, Gads, binds phosphorylated LAT tyrosines 171 and 191 which are also binding sites for Grb2 [12,44]. However experimental investigations [6] indicate that Gads does not appear to bind to SOS1 [6,45] and may act as a monovalent inhibitor of gel formation, reducing the range of concentrations over which superaggregate formation can occur [30].

LAT, LAX and LAB form a family of transmembrane adaptor proteins that are widely expressed in lymphocytes [7]. LAB's three distal tyrosines are its main sites of phosphorylation and all are sites for Grb2 binding [46]. LAX also binds Grb2, and five of its ten tyrosines are in Grb2-binding motifs [47]. We have confined our discussion to aggregation of LAT by Grb2-SOS1-Grb2 dimers since LAT is the most studied of these three transmembrane adaptors, but LAB and LAX are possible candidates for superaggregate formation as well. Aggregation is ubiquitous in cell signaling, manifesting itself at many levels. Once signaling is initiated, mechanism that result in the formation of large aggregates compete with opposing mechanisms that limit the growth of aggregates. We expect that models of aggregation, tailored to specific signaling complexes, can contribute to our understanding of the multiple roles of aggregation in regulating cell signaling cascades.

Supporting Information

Appendix S1 Calculation of dimensionless partition function for LAT population with distribution of valence.

(PDF)

Appendix S2 Condition for gel formation for LAT population with mixed valence.

(PDF)

Appendix S3 Calculation of dimensionless partition function for a population of bivalent receptors that can be cross-linked by ligand species of variable valency.

(PDF)

Acknowledgments

We acknowledge Dr. Anurag Sethi for his help in preparing publication-quality figures.

Author Contributions

Conceived and designed the experiments: AN BG. Performed the experiments: AN MM BG. Analyzed the data: AN MM BG. Contributed reagents/materials/analysis tools: AN MM ASP BG. Wrote the paper: AN MM BG.

References

1. Metzger H (1992) Transmembrane signaling: The joy of aggregation. *J Immunol* 149: 1477–1487.
2. Schlessinger J (2000) Cell signaling by receptor tyrosine kinases. *Cell* 103: 211–225.
3. Ozaki K, Leonard WJ (2002) Cytokine and Cytokine Receptor Pleiotropy and Redundancy. *J Biol Chem* 277: 29355–29358.
4. Tkaczyk C, Iwaki S, Metcalfe DD, Gifillan AM (2005) Roles of adaptor molecules in mast cell activation. *Chem Immunol Allergy* 87: 43–58.

5. Rivera J, Gilfillan AM (2006) Molecular regulation of mast cell activation. *J Allergy Clin Immunol* 117: 1214–1225.
6. Houtman JC, Yamaguchi H, Barda-Saad M, Braiman A, Bowden B, et al. (2006) Oligomerization of signaling complexes by the multipoint binding of GRB2 to both LAT and SOS1. *Nat Struct Mol Biol* 13: 798–805.
7. Fuller DM, Zhang W (2009) Regulation of lymphocyte development and activation by the LAT family of adapter proteins. *Immunol Rev* 232: 72–83.
8. Samelson L (2002) Signal transduction mediated by the T cell antigen receptor: the role of adapter proteins. *Annu Rev Immunol* 20: 371–394.
9. Zhang WG, Triple RP, Samelson LE (1998) LAT palmitoylation: Its essential role in membrane microdomain targeting and tyrosine phosphorylation during T cell activation. *Immunity* 187: 239–246.
10. Weber JR, Orstavik S, Torgersen KM, Danbolt NC, Berg SF, et al. (1998) Molecular cloning of the DNA encoding pp36, a tyrosine-phosphorylated adaptor protein selectively expressed by T cells and natural killer cells. *J Exp Med* 187: 1157–1161.
11. Zhang W, Sloan-Lancaster J, Kitchen J, Triple R, Samelson L (1998) LAT the ZAP-70 Tyrosine Kinase Substrate that Links T Cell Receptor to Cellular Activation. *Cell* 92: 83–92.
12. Zhang WG, Triple RP, Zhu M, Liu SK, McGlade CJ, et al. (2000) Association of Grb2, Gads, and phospholipase C-gamma1 with phosphorylated LAT tyrosine residues. Effect of LAT tyrosine mutations on T cell antigen receptor-mediated signaling. *J Biol Chem* 275: 23355–23361.
13. Saitoh S, Arudchandran R, Manetz TS, Zhang W, Sommers CL, et al. (2000) LAT is essential for FcεRI mediated mast cell activation. *Immunity* 12: 525–535.
14. Hartgroves LC, Lin J, Langen H, Zech T, Weiss A, et al. (2003) Synergistic assembly of linker for activation of T cells signaling protein complexes in T cell plasma membrane domains. *J Biol Chem* 278: 20389–20394.
15. Malbec O, Malissen M, Isnardi I, Lesourme R, Mura AM, et al. (2004) Linker for activation of T cells integrates positive and negative signaling in mast cells. *J Immunol* 173: 5086–5094.
16. Braiman A, Barda-Saad M, Sommers CL, Samelson LE (2006) Recruitment and activation of PLC1 in T cells: a new insight into old domains. *EMBO J* 25: 955–962.
17. Bunnell SC, Singer AL, Hong DI, Jacque BH, Jordan MS, et al. (2006) Persistence of cooperatively stabilized signaling clusters drives T-cell activation. *Mol Cell Biol* 26: 7155–7166.
18. Kambayashi T, LaRosa DF, Silverman MA, Koretzky GA (2009) Cooperation of adapter molecules in proximal signaling cascades during allergic inflammation. *Immunol Rev* 232: 99–114.
19. Songyang Y, Shoelson SE, McGlade J, Olivier P, Pawson T, et al. (1994) Specific motifs recognized by the SH2 domains of Csk, 3BP2, fpsf, GRB-2, HCP, Syk, and Vav. *Mol Cell Biol* 14: 2777–2785.
20. Shiue L, Zoller MJ, Brugge JM (1995) Syk is activated by phosphotyrosine-containing peptides representing the tyrosine-based activation motifs of the high affinity receptor for IgE. *J Biol Chem* 270: 10498–10502.
21. Zhang J, Billingsley ML, Kincaid RL, Sirgaganian RP (2000) Phosphorylation of Syk activation loop tyrosines is essential for Syk function: an in vivo study using a specific anti-Syk activation loop phosphotyrosine antibody. *J Biol Chem* 275: 35442–35447.
22. Williams BL, Schreiber KL, Zhang W, Wange RL, Samelson LE, et al. (1998) Genetic evidence for differential coupling of Syk family kinases to the T-cell receptor: reconstitution studies in a ZAP-70-deficient Jurkat T-cell line. *Mol Cell Biol* 18: 1388–1399.
23. Lowenstein EJ, Daly RJ, Batzer AG, Li W, Margolis B, et al. (1992) The SH2 and SH3 domain-containing protein Grb2 links receptor tyrosine kinases to ras signaling. *Cell* 70: 431–442.
24. Yuzawa S, Yokochi M, Hatanaka H, Ogura K, Kataoka M, et al. (1992) Solution structure of Grb2 reveals extensive flexibility necessary for target recognition. *J Mol Biol* 306: 527–537.
25. McDonald CB, Selden KL, Deegan BJ, Farooq A (2009) SH3 domains of Grb2 adaptor bind to pxxp motifs within the Sos1 nucleotide exchange factor in a discriminate manner. *Biochemistry* 48: 4074–4085.
26. Bunnell S, Hong DI, Kardon JR, Yamazaki T, McGlade CJ, et al. (2002) T cell ligation induces the formation of dynamically regulated signaling assemblies. *J Cell Biol* 158: 1263–1275.
27. Wilson BS, Pfeiffer JR, Surviladze Z, Gaudet EA, Oliver JM (2001) High resolution mapping of mast cell membranes reveals primary and secondary domains of FcεRI and LAT. *J Cell Biol* 154: 645–658.
28. Wilson BS, Pfeiffer JR, Oliver JM (2001) FcεRI signaling observed from the inside of the mast cell membrane. *Mol Immunol* 38: 1259–1268.
29. Saitoh S, Odom S, Gomez G, Sommers C, Young HA, et al. (2003) The four distal tyrosines are required for LAT-dependent signaling in FcεRI-mediated mast cell activation. *J Exp Med* 198: 831–843.
30. Nag A, Monine M, Faeder JR, Goldstein B (2009) Aggregation of membrane proteins by cytosolic cross-linkers: Theory and simulation of the LAT-Grb2-SOS1 system. *Biophys J* 96: 2604–2623.
31. Goldstein B, Perelson AS (1984) Equilibrium theory for the clustering of bivalent cell surface receptors by trivalent ligands: Application to histamine release from basophils. *Biophys J* 45: 1109–1123.
32. Yang J, Monine MI, Faeder JR, Hlavacek WS (2008) Kinetic Monte Carlo method for rule-based modeling of biochemical networks. *Phys Rev E* 78: doi: 10.1103/PhysRevE.78.031910.
33. Houtman JC, Brown PH, Bowden B, Yamaguchi H, Appella E, et al. (2007) Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: Application to adaptor protein complexes in cell signaling. *Protein Science* 16: 30–42.
34. Monine MI, Posner RG, Savage PB, Faeder JR, Hlavacek WS (2010) Modeling multivalent ligand-receptor interactions with steric constraints on configurations of cell-surface receptor aggregates. *Biophys J* 98: 48–56.
35. Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* 81: 2340–2361.
36. Cao Y, Li H, Petzold L (2004) Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *J Chem Phys* 121: 4059–4067.
37. Hlavacek W, Faeder J, Blinov M, Perelson A, Goldstein B (2003) The Complexity of Complexes in Signal Transduction. *Biotechnology and Bioengineering* 84: 783–794.
38. Hlavacek WS, Faeder JR, Blinov ML, Posner RG, Hucka M, et al. (2006) Rules for modeling signal-transduction systems. *Sci STKE* 344: re6.
39. Andrews NL, Lidke KA, Pfeiffer JR, Burns AR, Wilson BS, et al. (2008) Actin restricts FcεRI diffusion and facilitates antigen-induced receptor immobilization. *Nature Cell Biol* 10: 774–784.
40. Ballagopalan L, Barr VA, Sommers CL, Barda-Saad M, Goyal A, et al. (2007) c-Cbl-mediated regulation of LAT-nucleated signaling complexes. *Mol Cell Biol* 27: 8622–8636.
41. Balagopalan L, Ashwell BA, Bernot KM, Akpan IO, Quasba N, et al. (2011) Enhanced T-cell signaling in cells bearing linker for activation of T-cell (LAT) molecules resistant to ubiquitylation. *Proc Natl Acad Sci USA* 108: 2885–2890.
42. Feng G, Ouyang Y, Hu D, Shi Z, Gentz R, et al. (1996) Grap is a novel sh3-sh2-sh3 adaptor protein that couples tyrosine kinases to the ras pathway. *J Biol Chem* 271: 12129–12132.
43. Trüb T, Frantz J, Miyazaki M, Band H, Shoelson S (1997) The role of a lymphoid-restricted, grb2-like sh3-sh2-sh3 protein in t cell receptor signaling. *J Biol Chem* 272: 894–902.
44. Cho S, Velikovskiy CA, Swaminathan CP, Houtman JC, Samelson LE, et al. (2004) Structural basis for differential recognition of tyrosine-phosphorylated sites in the linker for activation of t cells (LAT) by the adaptor Gads. *EMBO J* 23: 1441–1451.
45. Houtman JC, Higashimoto Y, Dimasi N, Cho S, Yamaguchi H, et al. (2004) Binding Specificity of Multiprotein Signaling Complexes Is Determined by Both Cooperative Interactions and Affinity Preferences. *Biochemistry* 43: 4170–4178.
46. Koonpaew S, Janssen E, Zhu M, Zhang W (2004) The importance of three membrane-distal tyrosines in the adaptor protein NTAL/LAB. *J Biol Chem* 279: 11229–11235.
47. Zhu MH, Janssen E, Leung K, Zhang W (2002) Molecular cloning of a novel gene encoding a membrane-associated adaptor protein LAX in lymphocyte signaling. *J Biol Chem* 277: 46151–46158.
48. Chook YM, Gish GD, Kay CM, Pai EF, Pawson T (1996) The Grb2-mSos1 Complex Binds Phosphopeptides with Higher Affinity than Grb2. *J Biol Chem* 271: 30472–30478.
49. Sastry L, Lin W, Wong WT, Di Fiore PP, Scoppa CA, et al. (1995) Quantitative analysis of Grb2-Sos1 interaction: the N-terminal SH3 domain of Grb2 mediates affinity. *Oncogene* 11: 1107–1112.
50. Rosenbluth MJ, Lam WA, Fletcher DA (2006) Force microscopy of nonadherent cells: a comparison of leukemia cell deformability. *Biophys J* 90: 2994–3003.