

Association Rate Constants of Ras-Effector Interactions Are Evolutionarily Conserved

Christina Kiel^{1*}, Dorothee Aydin², Luis Serrano³

1 EMBL-CRG Systems Biology Unit, Centre de Regulacio Genomica, Barcelona, Spain, **2** EMBL-European Molecular Biology Laboratory, Heidelberg, Germany, **3** ICREA Professor, EMBL-CRG Systems Biology Unit, Centre de Regulacio Genomica, Barcelona, Spain

Abstract

Evolutionary conservation of protein interaction properties has been shown to be a valuable indication for functional importance. Here we use homology interface modeling of 10 Ras-effector complexes by selecting ortholog proteins from 12 organisms representing the major eukaryotic branches, except plants. We find that with increasing divergence time the sequence similarity decreases with respect to the human protein, but the affinities and association rate constants are conserved as predicted by the protein design algorithm, FoldX. In parallel we have done computer simulations on a minimal network based on Ras-effector interactions, and our results indicate that in the absence of negative feedback, changes in kinetics that result in similar binding constants have strong consequences on network behavior. This, together with the previous results, suggests an important biological role, not only for equilibrium binding constants but also for kinetics in signaling processes involving Ras-effector interactions. Our findings are important to take into consideration in system biology approaches and simulations of biological networks.

Citation: Kiel C, Aydin D, Serrano L (2008) Association Rate Constants of Ras-Effector Interactions Are Evolutionarily Conserved. *PLoS Comput Biol* 4(12): e1000245. doi:10.1371/journal.pcbi.1000245

Editor: Christopher Rao, University of Illinois at Urbana-Champaign, United States of America

Received: July 7, 2008; **Accepted:** November 6, 2008; **Published:** December 19, 2008

Copyright: © 2008 Kiel et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by EU INTERACTION PROTEOME (Grant No. LSHG-CT-2003-505520).

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

* E-mail: christina.kiel@crge.es

Introduction

Protein-protein interactions are the central elements in all signal transduction processes. The life times of protein complexes as well as regulatory processes need to be tightly controlled for proper systems functioning. Affinities are used to characterize the strength of protein interactions and the affinities between proteins involved in signaling processes have been shown to correlate with the activities (output/response) in such signal transduction processes [1,2]. In the majority of the cases, affinities between proteins and protein-ligands are determined using equilibrium binding methods, like isothermal titration calorimetry and fluorescence based methods, while rate constants of association and dissociation are only rarely determined. However, correlations of either association or dissociation rate constants with *in vivo* activity suggest that kinetic properties play a role in the cellular context [3–7]. As the affinity (K_d) can be described as the ratio between the dissociation (k_{off}) and association (k_{on}) rate constants, different ratios of k_{on} and k_{off} values can give rise to similar affinities. Kinetic rate constants have been shown to be important for signal transduction, however to which extent kinetics influence signaling might depend on the actual network and network topology. We could speculate that fast k_{on} and k_{off} values could result in rapid activation and deactivation upon short pulses of a stimulus, while slow ones could filter noise and result in prolonged signaling. If this is true it might open new aspects of cellular signal transduction regulation and could probably lead to conceptually new strategies in drug design. It is likely that the answer will depend on the network topology: rate constants might be important in some signaling branches, in others not.

Evolutionary conservation of protein composition and biochemical properties is usually a valuable indication for the cellular importance of a specific protein complex. In this study we have selected the Ras-effector complex formation, in order to analyze whether kinetic rate constants are evolutionary conserved. Ras proteins belong to the Ras superfamily of small GTPases and they have key roles in various signal transduction pathways, like proliferation and differentiation [8]. They act as molecular switches by cycling between an active GTP-bound and an inactive GDP-bound state [9,10]. Active Ras (Ras·GTP) can interact with effector molecules such as the Ser/Thr kinase Raf. The resulting Raf activation triggers the MAP kinase pathway, which leads to the transcription of target genes in the nucleus [11,12]. Other Ras·GTP binding effector proteins that have been identified are the PI3-kinase, members of the RalGDS family, and AF6 [13–16]. Effector proteins bind to Ras·GTP via a common domain with a ubiquitin-like topology [17–22], and various structures of effector domains in complex with Ras proteins have revealed a similar binding mode that involves mainly two antiparallel β -sheets of the RBD and Ras, respectively [23–29].

As Ras-effector protein interactions play a key role in cells, pathways involving Ras-effector interactions can be assumed to be at least partially conserved during evolution. In this study we analyzed whether the affinities and the association rate constants are conserved for 10 Ras-effector complexes in 12 different species, including worms, flies, fishes, and mammalian organisms. We used homology interface modeling and energy calculations, using FoldX 2.8 (<http://foldx.crg.es/>) [30,31] in order to model Ras-effector interactions of proteins from different organisms. FoldX uses an algorithm based on the original work of Schreiber

Author Summary

Cellular signal transduction processes are based on protein interactions. Proteins can either associate transiently with each other or form stable complexes, and the strength of the interaction is described by the affinity (the affinity is the ratio between the rate of dissociation and association). Protein complexes with similar affinities can bind and dissociate with different rates, and these rates describe the kinetic properties of protein binding. These kinetic rates are important for signaling; however, to what extent individual changes in such rate constants are biologically important or whether the affinity is more crucial might be different in different signaling processes. In this study we analyze whether association rates are conserved during evolution, because evolutionary conservation of protein biochemical properties is usually a valuable indication of its importance. We analyzed the binding of Ras proteins to effector domains, which are central proteins in many signal transduction pathways, in different organisms. On the basis of homology modeling and energy calculations we find that association rates are conserved, although the sequence similarity decreases compared to the human protein. Our finding should encourage further analysis of the importance of kinetics for cellular signal transduction.

and co-workers to calculate relative changes in k_{on} which has been validated experimentally numerous times [32]. Homology modeling was performed in a similar way as done in a previous study, on a genome-wide level for all human Ras-effector complexes [33,34].

Results

Importance of Electrostatic Charge Complementarity for Ras-Effector Association Kinetics

Binding of effector proteins to Ras proteins is mediated via a domain with an ubiquitin-like topology [35]. Members of the ubiquitin domain superfamily are the RA, the RBD, the PI3Krbd, the UBQ and the B41/ERM domain families [36]. However, the binding of Ras to effector domains does not depend on the fold itself, but rather on certain amino acid residues on the surface that are crucial for binding.

An important observation found in Ras-effector complex structures is the high charge complementarity between the proteins of the complex, where Ras is mainly negatively charged and the effector RBDs are mainly positively charged [23–29]. Various studies have shown that a strong electrostatic surface complementarity in a protein complex enhances the association rate constant by forming of a low affinity encounter complex before the final high affinity complex is formed [37–42]. The complex formation itself is promoted by electrostatic steering which stabilizes the transition state by decreasing the energy barrier for association [42,43]. In agreement with this concept of electrostatic steering and encounter complex formation, the association rate constants between Ras and effector domains were found to be fast (reviewed in [44]). Interestingly, the variance in binding energies when comparing different Ras-effector complexes is mainly the consequence of different association rate constants, while the dissociation rate constants are in a similar range [45–47]. For example, RafRBD is highly positively charged in its Ras binding region, and here the association rate constant was found to be very high in complex with the mainly negatively charged Ras proteins. In contrast, RalGDS

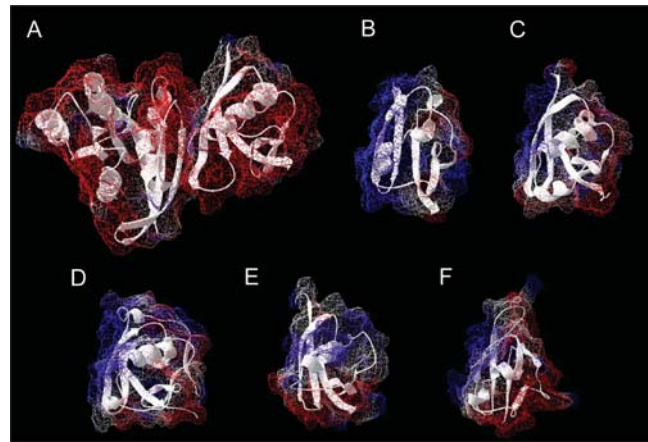


Figure 1. Electrostatic surface representation of Ras effector complex interfaces. (A) The Ras-RalGDS complex (pdb entry: 1lfd), and they single domains of (B) Rgl1 (pdb-entry: 1ef5), (C) Rgl2 (pdb-entry: 1rlf), (D) Grb7 (pdb-entry: 1wgr), (E) AF6-RA1 (pdb-entry: 1wxa), and (F) PLCeRA2 (pdb-entry: 2c5l) are shown. doi:10.1371/journal.pcbi.1000245.g001

has a mixed charged distribution (Figure 1A), and the k_{on} in complex with Ras is much lower. Interestingly, introducing positively charged residues at the edge of the interface of RalGDS can change binding kinetics and these RalGDS mutants were shown to bind “Raf-like” to Ras [43]. In Figure 1B–F we show the electrostatic surface potentials of several other RA/RB domains, which can bind to Ras (Rgl1, Rgl2, Grb7, AF6_RA1, PLCe_RA2) and for which structures have been solved, either by NMR or X-Ray, and we orient them similar as the RA domain of RalGDS in complex with Ras. In all cases the interface surface areas have a strong positively charged electrostatic potential, which suggests that association kinetics are important for these RA/RBD domains as well.

FoldX Electrostatic Interactions and Association Rate Constants Correlate with Experimental Association Rate Constants in Ras/Raf and Ras/RalGDS Complexes

Although the algorithm developed by Schreiber and co-workers implemented in FoldX (<http://foldx.crg.es/>) [30,31], has been validated experimentally on many different proteins, still it is a prediction method and as such needs some validation on the particular system under study. For this, we have selected the Ras-Raf complex and calculated k_{on} values (ΔG_{kon}) at different salt concentrations, ranging from 0 to 800 mM NaCl (corresponding to an ionic strength of ~50 to 850 mM in 50 mM Tris-buffer), and compared these results with experimental k_{on} values measured at different ionic strength using stopped-flow (Table S1; [48]). The experimental k_{on} values range from 7.4 to 60 $\mu\text{M}^{-1} \text{s}^{-1}$ and an excellent correlation with calculated association rate constants was observed ($R = 0.99$) (Figure 2A). Further, we used FoldX in order to generate *in silico* a series of mutations of charged residue in RalGDS, located either in the binding site, or at the edge of the binding site, and we calculated binding energies as well as association rate constants using the Ras-Ral complex. When comparing these results with experiments [43], we find again a very good correlation between experimental and calculated k_{on} values (Figure 2B) ($R = 0.89$), with the slopes of the two correlations (ionic strength and mutants) been similar. This indicates that absolute values of association rate constants can be reliably calculated over a wide range for different ionic strengths and mutations of Ras-effector complexes.

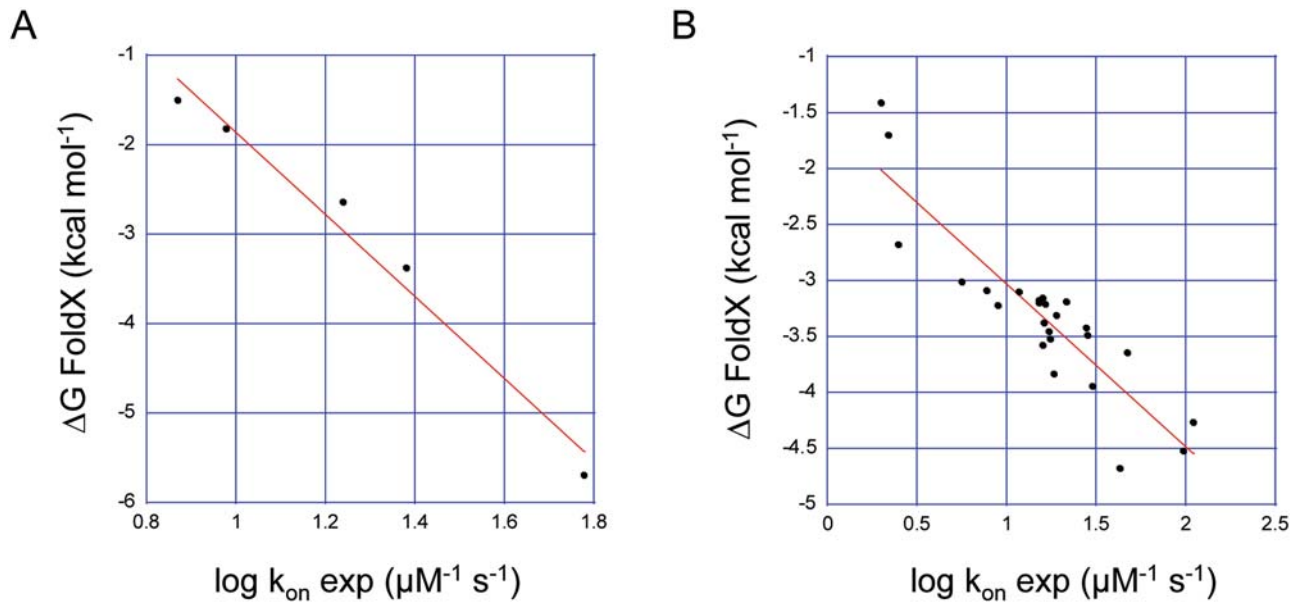


Figure 2. Calculation of association rate constants using FoldX and comparison with experimental association rate constants. (A) Ras-Raf complex at different ionic strength (salt concentrations from 0 to 800 mM NaCl) (Table S1, [48]). The correlation coefficient is 0.987. (B) The Ras-RalGDS complex. Mutations within the RalGDS domain were introduced, and compared with experimental data (Table S1, [43]). The correlation coefficient is 0.892.

doi:10.1371/journal.pcbi.1000245.g002

Ortholog Prediction of Ras and Selected Effector Proteins

We have selected proteins containing RA, RBD, PI3Krbd, and B41 domains, similar as in our previous genome-wide Ras-effector homology interface modeling study [34] (Figure 3A), for which binding to Ras has been shown experimentally (Table 1). These include the different isoforms of the Raf kinases, RalGDS, and the related proteins, Rgl1, and Rgl2. Other Ras binding domains are the PI3K-p110 gamma subunit, and Krit. In the following we will often refer to members of the ubiquitin superfamily as UBDs, without differentiating between RA, RBD, PI3Krbd, or B41.

In order to derive Ras and effector protein “interactions” from organisms representing the major eukaryotic branches, we have selected the following species (Figure 3B): *Homo sapiens* (*hs*) and *Mus musculus* (*mm*) were chosen for mammals, *Gallus gallus* (*gg*) for birds, *Xenopus tropicalis* (*xt*) for amphibians, *Fugu rubripes* (*fr*), and *Dario rerio* (*dr*) for fishes, *Drosophila melanogaster* (*dm*), *Drosophila pseudoobscura* (*dp*), *Anopheles gambiae* (*ag*) and *Apis mellifera* (*am*) for arthropods, and *Caenorhabditis elegans* (*ce*) and *Caenorhabditis briggsae* (*cb*) for nematodes. The orthologs were predicted by using the ENSEMBL (<http://www.ensembl.org>) [49] and the IMPARANOID databases (<http://inparanoid.cgb.ki.se/>) [50,51]. Domains were predicted using SMART [52,53] and the sequences were aligned automatically and by manual curation taking structural information into account [35] (Figure S1). Depending on the organism, between 22% and 78% of all human proteins orthologs were identified. When taking into account that certain proteins in lower organisms are orthologs of more than one human protein, e.g., RalGDS of *C. elegans* is also an ortholog of Rgl1 and Rgl2, the number of orthologs in different organisms ranges from 33% to 100%. The alignments of the UBDs of Ras effector proteins show a high similarity within orthologs and often also between different proteins of the same domain family. Furthermore, the similarity within the secondary structures of the RBD is higher than within the loops, indicating a conservation of the binding mode. The sequence identity of ortholog proteins (for detailed description see method) ranges between 100 and ~20% (Table S2). However, in

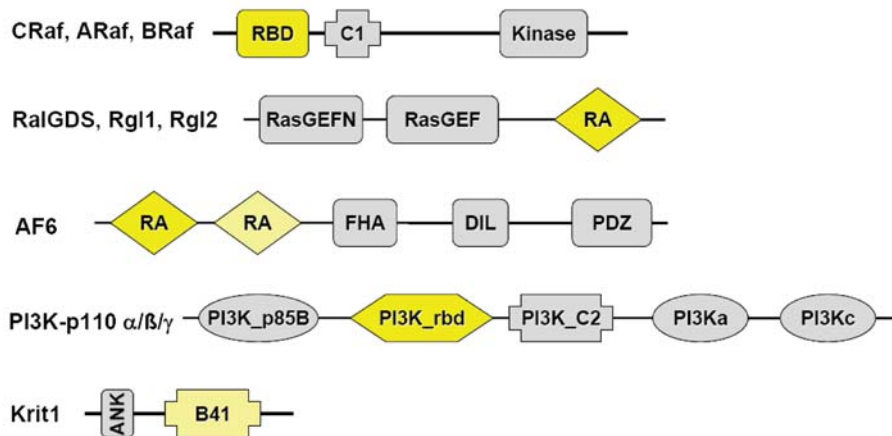
the majority of the cases the sequence identity decreases to ~30/40%. The only exceptions are the different PI3kinase p110 isoforms, where a drastic drop in sequence identity is observed for the corresponding orthologs/isoforms in *C. elegans*/*C. briggsae*.

The sequences of proteins that have a key role in cells are usually highly conserved among all organisms. In accordance with this, the sequences of Ras proteins were found to be nearly identical, especially in the effector binding region (Figure S2). The three Ras proteins, H-Ras, N-Ras and K-Ras could only be found in vertebrates, for arthropods and nematodes there is only one Ras protein which is most likely to be an ortholog of H-Ras. Due to the similarity in the effector binding region, only HRas was modeled (here termed as Ras).

Homology Interface Modeling of Ortholog Ras-Effector Complexes

The first three secondary structure elements ($\beta 1$, $\beta 2$, and $\alpha 1$) of the ubiquitin-like domain determine the interaction surface towards Ras and they have the largest impact on binding energy of the complex [33]. In those cases in which a crystal structure of Ras in complex with a RBD domain was available we use the structure to model the ortholog sequences (Ras-Raf, Ras-Ral, Ras-PI3 Kinase, Ras-Byr). For the rest we used the templates modeled in our previous study [34] (Table S3) that were validated experimentally by pull-down experiments (for details see methods). Only those UB domains that could be reliably modeled were selected (e.g., no van der waals' clashes above a fixed threshold of 2 kcal/mol). The species were then grouped into human (*hs*), mouse (*mm*), birds (*gg*), amphibians (*xt*), insects (*dm*, *dp*, *ag*, *am*) and nematodes (*ce*, *cb*). The mean of ΔG and ΔG_{kon} within each group was calculated and taken as value for the complete group. By grouping the different organisms, the problem of missing sequences can be solved for many proteins and mean values as well as standard deviation of ΔG and ΔG_{kon} can be calculated (the results do not change if we consider individual organisms, data not shown).

A



B

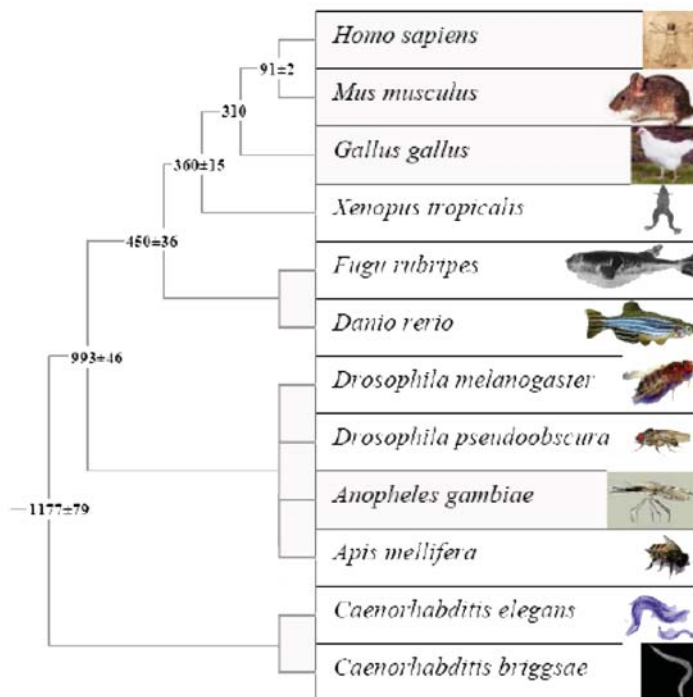


Figure 3. Selected effector domains and the phylogeny of used model organisms. (A) Domain architecture of selected effector proteins. The ubiquitin-like domains mediating binding to Ras (RA, RBD, PI3K_rbd, and B41) are shown in yellow, other domains in grey. The domain prediction was done using SMART [52,53] using the respective sequence from *Homo sapiens*. (B) The relationship and divergence time (million of years ago (Mya) \pm standard error) of used model organisms are shown. The divergence times for the chordate – arthropod divergence and the one for divergence of nematodes from the lineage leading to chordates and arthropods were derived from Wang et al 1999 [61]. The others were taken from Hedges 2002 [62]. doi:10.1371/journal.pcbi.1000245.g003

The results for all interaction energies (ΔG_{int} FoldX) and contribution of association rate constants (ΔG_{kon} FoldX) were plotted against the divergence time (Table S4 and Figure S3). While the sequence similarity decreases with increasing divergence time, the interaction energies as well as the association rate constants are conserved. A selection of representative results is

shown in Figure 4. A comparison of the mean values for all interaction energies and k_{on} values calculated for a particular Ras-effector complex in different organisms shows that the standard deviations are in the majority of the complexes small (Table S5 and Figure 5A). Interestingly, the interaction energies correlate with the association rate constant contribution (Figure 5B)

Table 1. Selected Ras binding domains, structural information, and template structures used for modeling.

Effector RBD	PDB Single Domain	PDB Complex with Ras Protein	Template Structure Used for Modeling
AF6 RA1	1wxa		1lfd
AF6 RA2			1lfd
Araf	1wxm		1gua
Craf	1rfa, 1rrb	1gua	1gua
Braf			1gua
Krit1			1lfd
PI3K p110 gamma		1he8	1he8
RalGDS	1lxd, 2rgf	1lfd	1lfd
Rgl1	1ef5		1lfd
Rgl2	1rlf		1lfd

doi:10.1371/journal.pcbi.1000245.t001

($R = 0.71$). This indicates that also for the so far kinetically uncharacterized UBD domains, the changes in ΔG are mainly a consequence of changing k_{on} . Thus, this could be underlying binding principle for the complete Ras-effector family.

In order to demonstrate that large changes in association rate constants would have been possible theoretically, we have selected the human Ras-RalGDS complex as an example for an *in silico* mutagenesis using FoldX. By either introducing positively or negatively charged residues at all positions at the surface of RalGDS, the FoldX-kon contribution could be increased from -3.65 kcal/mol to -7.6 kcal/mol or decreased to -0.47 kcal/mol, respectively (data not shown).

Simulation of a Minimal Ras-Effector Network

In order to analyze whether compensating changes in k_{on} and k_{off} can influence signal transduction, we used *in silico* simulations of a sub-network within the EGF signal transduction pathway. Activation of proteins following EGF stimulation is one of the most studied signaling systems, which involves the Ras-CRaf interaction as central elements, and numerous simulation models exist, which are able to correctly predict different aspects of EFG signaling found experi-

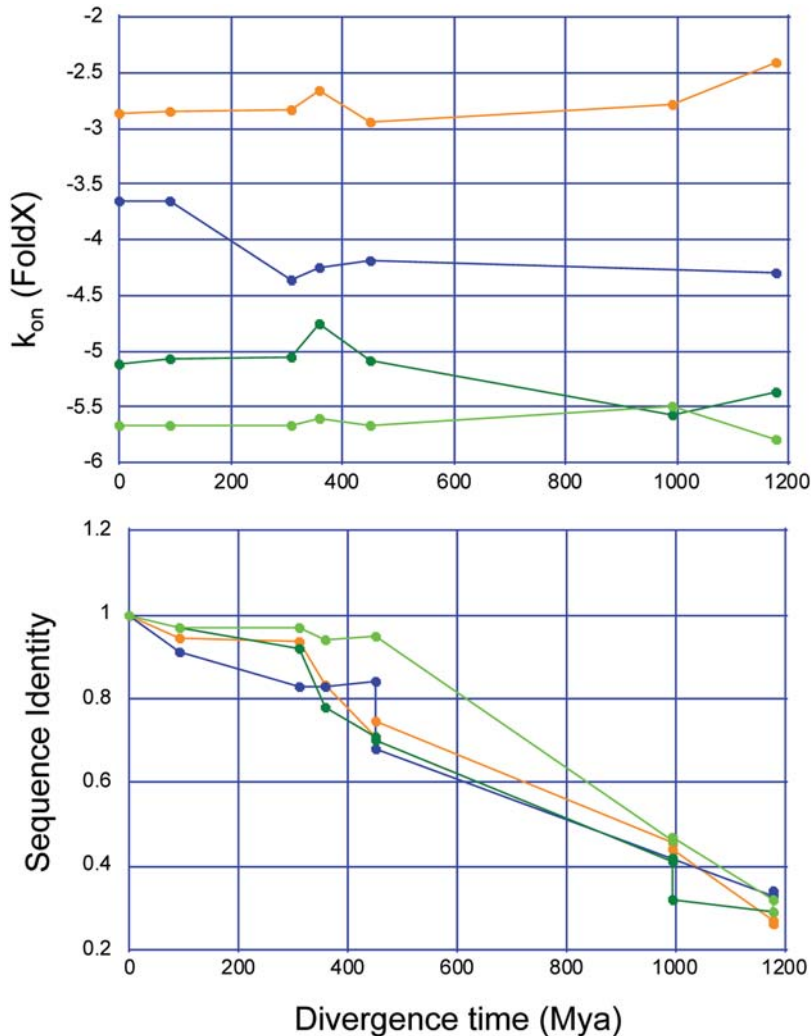


Figure 4. Homology interface modeling of ortholog Ras-effector complexes. Representative Results for predicted association rate constants. Association rate constants are conserved for Ras-effector complexes in different organisms with increasing divergence time, while the sequence identity is decreasing. Color code: AF6_RA2 (orange), RalGDS (dark blue), Braf (light blue), CRaf (dark green), and AF6_RA1 (light green). doi:10.1371/journal.pcbi.1000245.g004

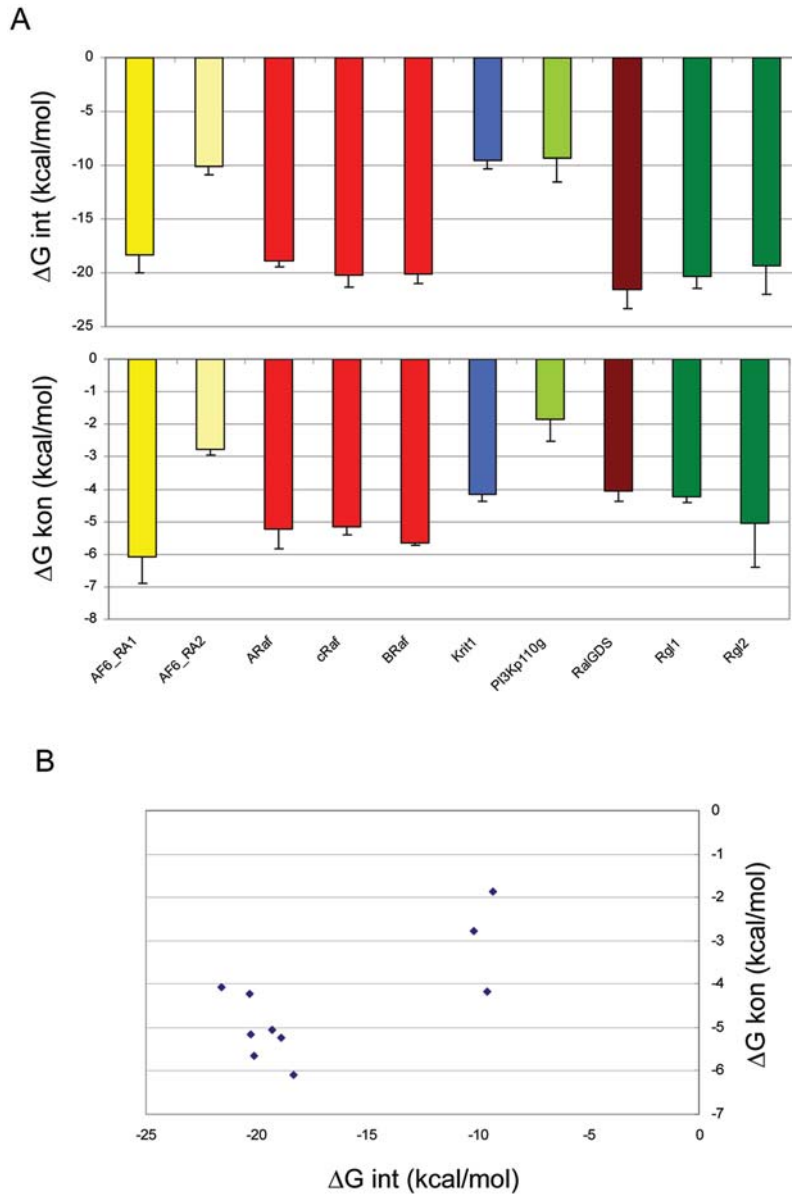


Figure 5. Mean values and standard deviation for calculated interaction energies (ΔG_{int}) and association rate constants (ΔG_{kon}). (A) Results from UBDs from the same protein or from the same protein family were shown in similar color. The results were taken from Table S5. (B) Correlation of calculated interaction energies with calculated association rate constants. The results were taken from Table S5, and the correlation coefficient is 0.71. doi:10.1371/journal.pcbi.1000245.g005

mentally [54–59]. Based on these earlier models we have constructed a minimal network involving Ras and Raf kinase (Figure 6A and Table S6). This minimal model involves activation of GEF upon stimulation (A), which results in activation of Ras (RasT = RasGTP). Subsequent binding of Raf to RasT activates Raf (Raf_{act}), which in turn leads to activation of a downstream target (X). Negative regulation was introduced by the GAP catalyzed hydrolysis reaction of RasGTP to RasGDP (RasD). We simulated this network by first applying a constant stimulus of “A” for 500 seconds using the wild type k_{on} and k_{off} values for the Ras-Raf interaction. Then we simulated the network with either 10-fold higher k_{on} and k_{off} rate constants, or 10 fold lower k_{on} and k_{off} (Figure 6B). Minor changes are observed when following X over time (activation peak) for the simulation with 10-fold higher k_{on} and k_{off} compared to the wild type situation. Only the simulation of 10-fold lower k_{on} and k_{off} resulted in a slightly smaller activation peak. However, when simulating the

network by applying a pulse of stimulation, of 10 s of “A” (and then removing the stimulus), large changes in the activation peak are observed, with a higher maximum for the simulation of 10-fold higher k_{on} and k_{off} values for the Ras-Raf interaction (Figure 6C). This shows, that under certain cellular conditions, like short pulse of activation, large changes in activation are expected for mutants with similar affinity, but changed and compensating effects on k_{on} and k_{off} . Thus, kinetic properties can be crucial, and in the case of Ras-effector interactions, association kinetics will be important to result in sufficient activation, when the system is activated by applying a pulse.

Discussion

The complex formation of Ras and effector proteins is driven by high association rate constants and only moderate dissociation rate constants [45–47]. Further, changes in affinity are mainly the consequence of changed association rate constants. Association

rate constants can be influenced by mutating charged residues at the edge of the interface [32,43]. If electrostatic interactions and association rate constants are important for the biological function of the cell, they should be conserved during the course of evolution. Using homology modeling and energy calculation covering a wide-range of sequences, and relating the output to the sequence conservation, we found that interaction energies as well as the electrostatic contributions and the association rate constants are conserved as well. While the sequence identity decreases with divergence time between the selected organisms, no trend could be found for the interaction energy and energies related to the electrostatics and k_{on} , although theoretically it should be possible, when sampling the possible contributions of k_{on} at different amino acid positions (Figure S4).

Biologically, electrostatic interactions within Ras-effector complex interfaces could be functionally important, because they are the basis for the observed dynamic behavior, as observed in the case of Ras binding to the Raf kinase effector protein: The Ras-RafRBD complex formation is characterized by both high association and dissociation rate constants (k_{on} and k_{off}), leading to affinities ($K_d = k_{off}/k_{on}$) in the range of 1 to 0.05 μM , under physiological conditions (this relatively low affinity seems to be functionally sufficient, since Ras is attached to the membrane via a lipid modification). The high k_{on} values provide the possibility to have a fast dissociation of the complex, while still having a reasonable tight binding complex (the lifetime of the complex between Ras and RafGDS, for example, is 0.1 s-1; see reference [47]). As Ras signaling depends very crucially on a strict control through regulating proteins like GAPs (GTPase activating

proteins) and GEFs (guanine nucleotide exchange factors), this fast dissociation allows regulatory proteins to access and act.

We assume that electrostatics contributions and binding kinetics could be important in other Ras signaling pathways, since association rate constants were found to be conserved during evolution, as demonstrated in this study for 10 effector domains. Further *in vivo* analysis will be needed to prove this hypothesis. These experiments could be performed by designing mutant variants, which are expected to have similar affinities, but changed association and dissociation rate constants. These protein variants could be expressed in cells and the effect on signal transduction monitored, e.g., after different pulses of stimulation.

It is expected that the effect of changing rate constants depends also on the network topology (negative feedback, feed forward inhibition, etc). This knowledge will be important for systems biology and simulation approaches, in order to know, at which positions in the network affinities will be sufficient, while for other accurate rate constants will be crucial for correct prediction. Further, it could open conceptually new aspects in drug design.

Methods

Selected Species, Orthologs, and Domain Prediction

Proteins from the following species were used in order to get a good representation of all branches: *Homo sapiens* and *Mus musculus* (mammals), *Gallus gallus* (birds), *Xenopus tropicalis* (amphibians), *Fugu rubripes* and *Danio rerio* (fishes), *Drosophila melanogaster*, *Drosophila pseudoobscura*, *Anopheles gambiae* and *Apis mellifera* (arthropods), and *Caenorhabditis elegans* and *Caenorhabditis briggsae* (nematodes). Only

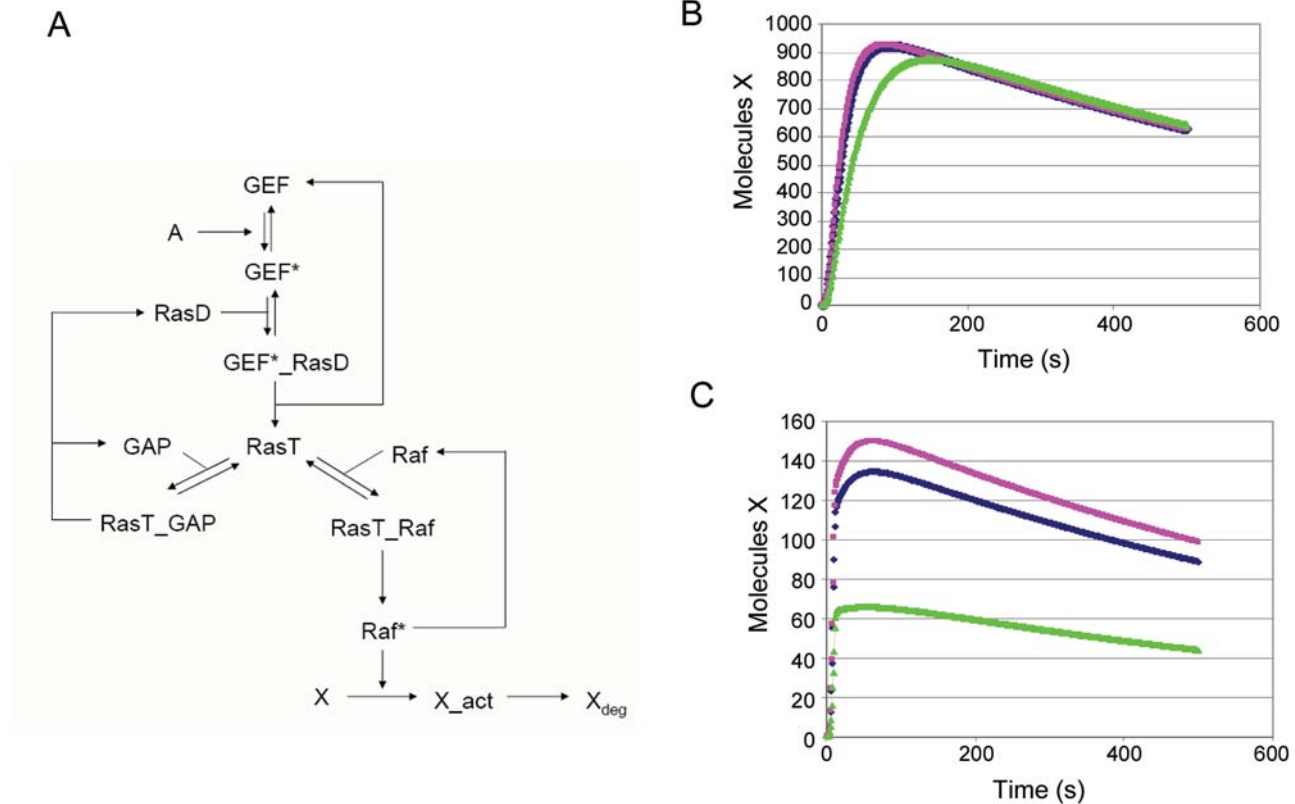


Figure 6. SmartCell Simulations of a minimal network within EGF signal transduction. (A) Schematic diagram of the reactions involved. For details see text and Table S5. Simulation of the network using the WT affinity of the Ras-Raf complex (blue line), 10 fold higher k_{on} and k_{off} rates (pink line), or 10 fold lower k_{on} and k_{off} values (green line). Either a constant stimulus was applied for 500 seconds (B) or the system was activated for 10 seconds and then the stimulus was removed (C).

doi:10.1371/journal.pcbi.1000245.g006

two RA domain containing were retrieved from *Saccharomyces cerevisiae* (sc), because these proteins are involved in a different pathway. For each protein, the human ENSEMBL protein ID was retrieved from ENSEMBL (<http://www.ensembl.org>) [49]. The orthologs were predicted by using the ENSEMBL database for *Xenopus tropicalis*, or the INPARANOID database (<http://www.inparanoid.cgb.ki.se>) [50,51]. ENSEMBL [49] classifies the prediction based on the BLAST results. Only those orthologs were chosen that were a unique best reciprocal hit in both directions. As the INPARANOID database [50,51] provides more information about orthologs of members of protein families, e.g., PI3K p110, the prediction was preferentially used. The sequences retrieved from ENSEMBL or INPARANOID were analyzed using SMART (<http://smart.embl-heidelberg.de>) [52,53], in order to determine the domain architecture of the protein and the domain sequences.

Homology Modelling and Energy Calculations Using FoldX

For modeling of Ras-binding domains in complex with Ras proteins, we have taken the pdb-files of the following Ras effector complexes: Ras-RalGDS (pdb-entry: 1LFD), Ras-PI3Kinase (pdb-entry: 1HE8), and Raps-Raf (pdb-entry: 1GUA). Different template structures have been generated by deleting certain parts in the complex; the decision was mainly based on the alignment used to model the different binding domains. The ortholog sequences for one protein were aligned using standard automatic alignment tools, since sequence homology is high. However, the alignment of different effector domains from different families (RA, RBD, PI3K_{rbd}, B41), was done based on manual curated structural-based sequence alignments as discussed in detail in a previous publication [35]. Basically two kinds of template structures have been generated (Table S3): a short version, where all secondary structure elements and loops (apart from β 1, β 1, β 2, α 1) were deleted, as this is the part mainly contributing to the binding energy (similar as done in our previous study [33,34]). In addition 'long template' structures have been generated. We could not model loop regions in those cases where the loops were not of the same length. For having a proline at the beginning of β -strand 1 (position 26 in RalGDS, position 229 in PI3K, position 66 in Raf and position 81 in spByr2), we prepared special template structures by moving the backbone slightly, after introducing the proline at these positions (we checked that the proline was in acceptable dihedral angles and that the main chain CO group was still H-bonded to Ras). These template structures were then used to model the complex structures for AF6_RA2. The homology modelling was done as described before [33,34]. The homology modeling was done based on the sequence alignment (Figure S1 and Figure S2), using different template structures using the design option in a new version of FoldX 2.8 [30,31]. During this design procedure, FoldX is testing different rotamers and allows neighbor side chains to move. After reconstruction, all models have been

passed through an additional optimization step by using the repair function of FoldX (detailed description in [33,34]). Energy calculations of Ras-effector complexes have been done using FoldX as described before (<http://fold-x.crg.es>) [30,31].

Simulations of a Minimal Ras-Effector Network

A model was generated based on previous models of EGF signal transduction (see Table S6). Simulations were performed using the SmartCell software (<http://www.smartcell-crg.es>) [60] using ordinary differential equations.

Supporting Information

Figure S1 Alignment of UB domains

Found at: doi:10.1371/journal.pcbi.1000245.s001 (0.05 MB PDF)

Figure S2 Alignment of Ras proteins

Found at: doi:10.1371/journal.pcbi.1000245.s002 (0.03 MB PDF)

Figure S3 Diagrams of DGint and DGkon FoldX values plotted against the divergence time

Found at: doi:10.1371/journal.pcbi.1000245.s003 (0.07 MB PDF)

Figure S4 FoldX mutational scanning of RafRBD and RalGDS-RA. (A) Ras-RafRBD (pdb-entry 1GUA). (B) Ras-RalGDS (pdb entry 1LFD). Effect of all residues in Raf-RBD or RalGDS-RA on the contribution of DG kon as calculated by FoldX. Either positively charged residues were mutated to alanine (red) or negatively charged or neutral residues were mutated to lysine using FoldX and the difference compared to the WT DGkon was calculated and plotted for every amino acid position.

Found at: doi:10.1371/journal.pcbi.1000245.s004 (0.03 MB PDF)

Table S1 Experimental and calculated association rate constants

Found at: doi:10.1371/journal.pcbi.1000245.s005 (0.03 MB PDF)

Table S2 Divergence times and sequence identities for Ub domains

Found at: doi:10.1371/journal.pcbi.1000245.s006 (0.02 MB PDF)

Table S3 Template structures used for homology modelling

Found at: doi:10.1371/journal.pcbi.1000245.s007 (0.01 MB PDF)

Table S4 FoldX results for all homology models

Found at: doi:10.1371/journal.pcbi.1000245.s008 (0.01 MB PDF)

Table S5 Mean values and STDEV for homology models for ortholog complexes

Found at: doi:10.1371/journal.pcbi.1000245.s009 (0.01 MB PDF)

Table S6 Modelling parameters

Found at: doi:10.1371/journal.pcbi.1000245.s010 (0.05 MB PDF)

Author Contributions

Conceived and designed the experiments: LS. Performed the experiments: CK DA. Analyzed the data: CK DA LS. Wrote the paper: CK LS.

References

- Block C, Janknecht R, Herrmann C, Nassar N, Wittinghofer A (1996) Quantitative structure-activity analysis correlating Ras/Raf interaction in vitro to Raf activation in vivo. *Nat Struct Biol* 3: 255–251.
- Pearce KHJ, Cunningham BC, Fuh G, Teeri T, Wells JA (1999) Growth hormone binding affinity for its receptor surpasses the requirements for cellular activity. *Biochemistry* 38: 81–89.
- Piehler J, Roisman LC, Schreiber G (2000) New structural and functional aspects of the IFN-receptor interaction revealed by comprehensive mutational analysis of the binding interface. *J Biol Chem* 275: 40425–40433.
- Wang Y, Shen BJ, Sebald W (1997) A mixed-charge pair in human interleukin 4 dominates high-affinity interaction with the receptor α chain. *Proc Natl Acad Sci U S A* 94: 1657–1662.
- Batista FD, Neuberger MS (1998) Affinity dependence of the B cell response to antigen: a threshold, a ceiling, and the importance of off-rate. *Immunity* 8: 751–759.
- Foote J, Eisen HN (1995) Kinetic and affinity limits on antibodies produced during immune response. *Proc Natl Acad Sci U S A* 92: 1254–1256.
- Botti SA, Felder CE, Sussmann JL, Silman I (1998) Electrotactins: a class of adhesion proteins with conserved electrostatic and structural motifs. *Protein Eng* 1: 415–420.
- Vetter IR, Wittinghofer A (2001) The guanine nucleotide-binding switch in three dimensions. *Science* 294: 1299–1304.
- Bourne HR, Sanders DA, McCormick F (1990) The GTPase superfamily: a conserved switch for diverse cell functions. *Nature* 348: 125–132.

10. Bourne HR, Sanders DA, McCormick F (1991) The GTPase superfamily: conserved structure and molecular mechanism. *Nature* 349: 117–127.
11. Rapp UR, Goldsborough MD, Mark GE, Bonner TI, Groffen J, et al. (1983) Structure and biological activity of v-raf, a unique oncogene transduced by a retrovirus. *Proc Natl Acad Sci U S A* 80: 4218–4222.
12. Vojtek AB, Hollenberg SM, Cooper JA (1993) Mammalian Ras interacts directly with the serine/threonine kinase Raf. *Cell* 74: 205–214.
13. Rodriguez-Viciana P, Warne PH, Dhand R, Van Haesebroeck B, Gout I, et al. (1994) Phosphatidylinositol-3-OH kinase as a direct target of Ras. *Nature* 370: 527–532.
14. Hofer F, Fields S, Schneider C, Martin GS (1994) Activated Ras interacts with the Ral guanine nucleotide dissociation stimulator. *Proc Natl Acad Sci U S A* 91: 11089–11093.
15. Kikuchi A, Demo SD, Ye Z-H, Chen Y-W, Williams LT (1994) RalGDS family members interact with the effector loop of ras p21. *Mol Cell Biol* 14: 7483–7491.
16. Kuriyama M, Harada N, Kuroda S, Yamamoto T, Nakafuku M, et al. (1996) Identification of AF-6 and Canoe as putative targets for Ras. *J Biol Chem* 271: 607–610.
17. Emerson SD, Madison VS, Palermo RE, Waugh DS, Scheffler JE, et al. (1995) Solution structure of the Ras-binding domain of c-Raf-1 and identification of its Ras interaction surface. *Biochemistry* 34: 6911–6918.
18. Huang L, Weng X, Hofer F, Martin GS, Kim SH (1997) Three-dimensional structure of the Ras-interacting domain of RalGDS. *Nat Struct Biol* 4: 609–615.
19. Geyer M, Herrmann C, Wohlgemuth S, Wittinghofer A, Kalbitzer HR (1997) Structure of the Ras-binding domain of RalGEF and implications for Ras binding and signaling. *Nat Struct Biol* 4: 694–699.
20. Walker EH, Perisic O, Ried C, Stephens L, Williams RL (1999) Structural insights into phosphoinositide 3-kinase catalysis and signaling. *Nature* 402: 313–320.
21. Steiner G, Kremer W, Linnemann T, Herrmann C, Geyer M, et al. (2000) Sequence-specific resonance assignment of the Ras-binding domain of AF6. *J Biomol NMR* 18: 73–74.
22. Gronwald W, Huber F, Grünewald P, Spörner M, Wohlgemuth S, et al. (2001) Solution Structure of the Ras-binding domain of the protein kinase Byr2 from *Schizosaccharomyces pombe*. *Structure* 9: 1029–1041.
23. Nassar N, Horn G, Herrmann C, Scherer A, McCormick F, et al. (1995) The 2.2 Å crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with Rap1A and a GTP analogue. *Nature* 375: 554–560.
24. Nassar N, Horn G, Herrmann C, Block C, Janknecht R, et al. (1996) Ras/Rap effector specificity determined by charge reversal. *Nat Struct Biol* 3: 723–729.
25. Huang L, Hofer F, Martin GS, Kim SH (1998) Structural basis for the interaction of Ras with RalGDS. *Nat Struct Biol* 5: 422–426.
26. Vetter IR, Linnemann T, Wohlgemuth S, Geyer M, Kalbitzer HR, et al. (1999) Structural and biochemical analysis of Ras-effector signaling via RalGDS. *FEBS Lett* 451: 175–180.
27. Pacold ME, Suire S, Perisic O, Lara-Gonzalez S, Davis CT, et al. (2000) Crystal structure and functional analysis of Ras binding to its effector phosphoinositide 3-kinase gamma. *Cell* 103: 931–943.
28. Scheffzek K, Grünewald P, Wohlgemuth S, Kabsch W, Tu H, et al. (2001) The Ras-Byr2RBD complex: structural basis for Ras effector recognition in yeast. *Structure* 9: 1043–1050.
29. Bunney TD, Harris R, Gandarillas NL, Josephs MB, Roe SM, et al. (2006) Structural and mechanistic insights into ras association domains of phospholipase C epsilon. *Mol Cell* 21: 495–507.
30. Guerois R, Nielsen JE, Serrano L (2002) Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J Mol Biol* 320: 369–387.
31. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, et al. (2005) The FoldX web server: an online force field. *Nucleic Acids Res* 33: W382–W388.
32. Selzer T, Albeck S, Schreiber G (2000) Rational design of faster associating and tighter binding protein complexes. *Nat Struct Biol* 7: 537–541.
33. Kiel C, Wohlgemuth S, Rousseau F, Schymkowitz J, Ferkinghoff-Borg J, et al. (2005) Recognizing and defining true Ras binding domains II: in silico prediction based on homology modelling and energy calculations. *J Mol Biol* 348: 759–775.
34. Kiel C, Foglierini M, Kuemmerer N, Beltrao N, Serrano L (2007) A genome-wide Ras-effector interaction network. *J Mol Biol* 370: 1020–1032.
35. Kiel C, Serrano L (2006) The ubiquitin domain superfold: structure-based sequence alignments and characterization of binding epitopes. *J Mol Biol* 355: 821–844.
36. Orengo CA, Jones DT, Thornton JM (1994) Protein superfamilies and domain superfolds. *Nature* 372: 631–634.
37. Berg OG, von Hippel PH (1985) Diffusion-controlled macromolecular interactions. *Annu Rev Biophys Chem* 14: 131–160.
38. Northrup SH, Erickson HP (1992) Kinetics of protein-protein association explained by Brownian dynamic computer simulation. *Proc Natl Acad Sci U S A* 89: 3338–3342.
39. Vijayakumar M, Wong K-Y, Schreiber G, Fersht AR, Szabo A, et al. (1998) Electrostatic enhancement of diffusion-controlled protein-protein association: comparison of theory and experiment on barnase and barstar. *J Mol Biol* 278: 1015–1024.
40. Camacho CJ, Weng Z, Vajda S, DeLisi C (1999) Free energy landscapes of encounter complexes in protein-protein association. *Biophys J* 76: 1166–1178.
41. Camacho CJ, Kimura SR, DeLisi C, Vajda S (2000) Kinetics of desolvation-mediated protein-protein binding. *Biophys J* 78: 1094–1105.
42. Selzer T, Schreiber G (2001) New insights into the mechanism of protein-protein association. *Proteins* 45: 190–198.
43. Kiel C, Selzer T, Shaul Y, Schreiber G, Herrmann C (2004) Electrostatically optimized Ras-binding Ral guanine dissociation stimulator mutants increase the rate of association by stabilizing the encounter complex. *Proc Natl Acad Sci U S A* 101: 9223–9228.
44. Kiel C, Serrano L (2007) Affinity can have many faces: Thermodynamic and kinetic properties of Ras-Effector Complex formation. *Curr Chem Biol* 1: 215–225.
45. Sydor JR, Engelhard M, Wittinghofer A, Goody R, Herrmann C (1998) Transient kinetic studies on the interaction of Ras and the Ras-binding domain of c-Raf-1 reveal rapid equilibrium of the complex. *Biochemistry* 37: 14292–14299.
46. Linnemann T, Geyer M, Jaitner BK, Block C, Kalbitzer HR, et al. (1999) Thermodynamic and kinetic characterization of the interaction between the Ras binding domain of AF6 and members of the Ras subfamily. *J Biol Chem* 274: 13556–13562.
47. Linnemann T, Kiel C, Herter P, Herrmann C (2002) The activation of RalGDS can be achieved independently of its Ras binding domain: implications for an activation mechanism in Ras effector specificity and signal distribution. *J Biol Chem* 277: 7831–7837.
48. Kiel C (2003) Investigation of Ras-effector complexes with changed electrostatic properties. Ph.D. thesis, Ruhr-Universität Bochum, Germany.
49. Hubbard TJ, Aken BL, Beal K, Ballester B, Caccamo M, et al. (2007) Ensembl 2007. *Nucleic Acids Res* 35: D610–D617.
50. Remm M, Storm CEV, Sonnhammer ELL (2001) Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J Mol Biol* 314: 1041–1052.
51. O'Brien KP, Remm M, Sonnhammer ELL (2005) Inparanoid: a comprehensive database of eukaryotic orthologs. *Nucleic Acids Res* 33: D476–D480.
52. Schultz J, Milpetz F, Bork P, Ponting CP (1998) Smart, a simple modular architecture research tool – identification of signalling domains. *Proc Natl Acad Sci U S A* 95: 5857–5864.
53. Letunic I, Goodstadt L, Dickens NJ, Doerks T, Schultz J, et al. (2002) Recent improvements to the SMART domain-based sequence annotation resource. *Nucleic Acid Res* 30: 242–244.
54. Kholodenko BN, Demin OV, Moehren G, Hoek JB (1999) Quantification of short term signaling by the epidermal growth factor receptor. *J Biol Chem* 274: 30169–30181.
55. Schoeberl B, Eichler-Jonsson C, Gilles ED, Mueller G (2002) Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat Biotechnol* 20: 370–375.
56. Yamada S, Taketomi T, Yoshimura A (2004) Model analysis of difference between EGF pathway and FGF pathway. *Biochem Biophys Res Commun* 314: 1113–1120.
57. Sasagawa S, Ozaki Y, Fujita K, Kuroda S (2005) Prediction and validation of the distinct dynamics of transient and sustained ERK activation. *Nat Cell Biol* 7: 365–373.
58. Kiyatkin A, Aksamitiene E, Markevich NI, Borisov N, Hoek JB, et al. (2006) Scaffolding protein Grb2-associated binder 1 sustains epidermal growth factor-induced mitogenic and survival signaling by multiple positive feedback loops. *J Biol Chem* 281: 19925–19938.
59. Birtwistle MR, Hatakeyama M, Yumoto N, Ogunnaike BA, Hoek JB, et al. (2007) Ligand-dependent responses of the ErbB signaling network: experimental and modeling analyses. *Mol Syst Biol* 3: 144.
60. Ander M, Beltrao N, Di Ventura B, Ferkinghoff-Borg J, Foglierini M, et al. (2004) SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: analysis of simple networks. *Syst Biol* 1: 129–138.
61. Wang DY, Kumar S, Hedges SB (1999) Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. *Proc Biol Soc* 266: 163–171.
62. Hedges SB (2002) The origin and evolution of model organisms. *Nat Rev Genet* 3: 838–849.