PROTEINS: Structure, Function, and Bioinformatics 56:322–337 (2004)

Change in Protein Flexibility Upon Complex Formation: Analysis of Ras-Raf Using Molecular Dynamics and a Molecular Framework Approach

Holger Gohlke,[1] Leslie A. Kuhn,[2] and David A. Case[1]*

[1]*Department of Molecular Biology, The Scripps Research Institute, La Jolla, California*

[2]*Department of Biochemistry and Molecular Biology and Center for Biological Modeling, Michigan State University, East Lansing, Michigan*

ABSTRACT **Changes in flexibility upon protein–protein complex formation of H-Ras and the Ras-binding domain of C-Raf1 have been investigated using the molecular framework approach FIRST (Floppy Inclusion and Rigid Substructure Topology) and molecular dynamics simulations (MD) of in total ~35 ns length. In a computational time of about one second, FIRST identifies flexible and rigid regions in a single, static three-dimensional molecular framework, whose vertices represent protein atoms and whose edges represent covalent and non-covalent (hydrogen bond and hydrophobic) constraints and fixed bond angles within the protein. The two methods show a very good agreement with respect to the identification of changes in flexibility in both binding partners on a local scale. This implies that flexibility can be successfully predicted by identifying which bonds limit motion within a molecule and how they are coupled. In particular, as identified by MD, the β-sheet in Raf shows considerably more pronounced orientational correlations in the bound state compared to the unbound state. Similarly, FIRST assigns the β-sheet to the largest rigid cluster of the complex. Interestingly, FIRST allows us to identify that interactions across the interface (but not conformational changes upon complex formation) result in the observed rigidification. Since regions of the β-sheet of Raf that do not interact directly with Ras become rigidified, this also demonstrates the long-range aspect to rigidity percolation. Possible implications of the change of flexibility of the Ras-binding domain of Raf on the activation of Raf upon complex formation are discussed. Finally, the sensitivity of FIRST results with respect to the representation of non-covalent interactions used as constraints is probed. Proteins 2004; 56:322–337.** © 2004 Wiley-Liss, Inc.

Key words: protein–protein interactions; flexibility/rigidity; mobility/dynamics; correlated motions; FIRST; induced fit; conformational change; allostery

INTRODUCTION

It has long been recognized that protein flexibility is important for a wide range of biological phenomena, such as enzymatic reaction and control.[1] In fact, some studies have found a correlation between the onset of protein activity and a changeover from harmonic to anharmonic protein dynamics,[2,3] although the exact temperature range of the dynamical transition is still being discussed.[4] Protein flexibility also influences stability[5,6] and folding.[7–9]

In the case of protein–protein binding, flexibility of the binding partners provides the origin for their plasticity, enabling them to conformationally adapt to each other. As such, examples of considerable "induced fit" have been shown to play a dominant role in protein–protein binding interfaces.[10–12] Perhaps even more important than flexibility per se are *changes* in the flexibility upon complex formation. Although usually recognition sites become less flexible in contact with the binding partner, influences on the flexibility of residues distant from the epitope are also seen. This can be explained in that perturbations at the binding site can be propagated to remote locations by altering the dynamic network of interactions in proteins.[13–15] Finally, it is increasingly recognized that complex formation between proteins and other proteins, DNA, or small molecules may lead to an increase in configurational entropy (which is related to an increase in flexibility of the system), compensating for the loss of translational and rotational entropy upon association,[16–23] as proposed by Steinberg and Scheraga[24] almost 40 years ago. For some cases, transfer of flexibility to other protein parts (within one binding partner) has been described, leading to a redistribution of protein configurational entropy, also potentially reducing the total entropy loss.[25–30]

Two classes of computational approaches are available to understand flexibility (or its opposite, rigidity) in proteins without doing expensive molecular dynamics (MD) simulations. First, *different* conformational states (for example different conformations determined by crystallography or NMR spectroscopy) can be compared to locate

Grant sponsor: the National Institutes of Health; Grant number: RR12255; Grant sponsor: the National Institutes of Health, Mathematics in Biology; Grant number: GM67249

*Correspondence to: David A. Case, Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037. E-mail: case@scripps.edu.

Received 9 September 2003; Accepted 31 December 2003

Published online 14 May 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20116

rigid and flexible regions (such as domains and hinges) within the molecule.[31–35] Second, flexible and rigid protein parts can also be predicted using a *single* protein conformation by either structurally identifying domain regions and hinge joints,[6,36–39] by methods related to normal mode analysis,[28,40–44] or by determining spatial variations in local packing density, which have been shown to be quantitatively related to root-mean-square fluctuations of proteins.[45]

Here, we use flexibility concepts that are well grounded in mathematics,[46] engineering,[47] and solid state physics.[48] Laman's Theorem[49] (which precisely defines the rigid regions and degrees of freedom for bonds in a 2D framework) and the absence of any counterexamples in twenty years of applications of the resulting Molecular Framework Conjecture[50] (which extends Laman's work to the subset of all 3D networks with molecule-like properties, such as fixed bond coordination angles) allow us to apply ideas from structural engineering that relate the stability of a network of joints (e.g., atoms) connected by struts (e.g., bonds) to the average number of struts at the joints (e.g., the mean coordination, or number of bonds, for atoms in the network).[47] Proteins can be considered as molecular frameworks[50] when the covalent bonds and strong hydrogen bonds and hydrophobic interactions are modeled appropriately as distance constraints between atoms.[9,51] In the past few years, new computational algorithms have been developed[52,53] that also allow a detailed *local* determination of the rigidity of networks to be made.[48] Thus, the number and spatial distribution of bond-rotational degrees of freedom can be related to regions of rigidity and flexibility. A fast combinatorial algorithm, the pebble game, has been developed for counting these degrees of freedom in 2D and 3D bond networks and has been applied to very large bond networks in amorphous materials.[48,52,54]

The FIRST (Floppy Inclusion and Rigid Substructure Topology) software is an implementation of the pebble game along with code that deduces and represents the protein covalent and non-covalent bond network, including angular constraints, given a Protein Data Bank[55] file as input. Previous work on a series of proteins characterized as flexible by crystallography or NMR has shown that FIRST accurately identifies flexible regions, from a single, static structure of the protein.[51,56] Rigid regions as well as collectively and independently moving regions are also identified. FIRST analyses have also been used successfully to relate the unfolding of a protein to its loss of structural stability,[9] to identify protein folding cores and pathways,[57] to identify hinge and loop motions that are essential for biological function,[56] and to provide the starting point for simulating the motions of flexible protein regions[58] (M.I. Zavodsky, M. Lei, M.F. Thorpe, L.A. Kuhn, Proteins, in press).

In this study, we apply FIRST analyses to the complex of H-Ras (166 residues) and the Ras-binding domain of C-Raf1 (76 residues) (Fig. 1) to investigate the influence of protein–protein complex formation on the intrinsic flexibility of both binding partners. We compare the results with those obtained from molecular dynamics simulations.

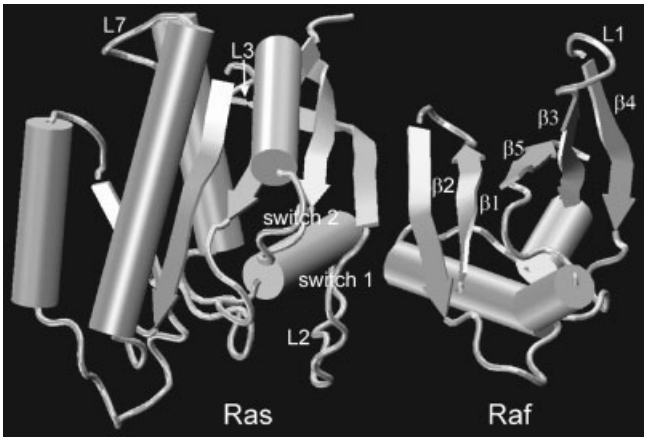


Fig. 1. Modeled Ras–Raf complex structure (see Gohlke et al.[75] for further details). Secondary structure elements discussed in the text are labeled. All molecular graphics figures in this study were prepared with the program VMD.[106]

Ras is an essential component of signal transduction pathways.[59] On going from an inactive (GDP-bound) state to an active (GTP-bound) state, conformational changes mainly occur in two regions of Ras, designated "switch 1" (corresponding to residues 30–37 of loop L2) and "switch 2" (corresponding to residues 60–76 of loop L4 and helix α2).[60,61] Being activated by "upstream" signals, the protein interacts with "downstream" effectors such as Raf mostly via the "effector region" (residues 32–40).[62] Molecular structures for the unbound proteins[61,63–65] or complexes closely related to Ras-Raf[66,67] have been determined either by X-ray crystallography or by NMR spectroscopy. They show that the Ras-effector interactions mainly occur due to a formation of an inter-protein β-sheet, resulting in an interface size of approximately 1200 $Å^2$ [59] and only small to moderate changes of the protein structure upon binding[59] [e.g., 1.7 Å C_α root-mean-square deviation between Raf in the unbound state (PDB code: 1RRB) and Raf in the complex Rap1A-Raf (PDB code: 1GUA)]. Ras-effector interactions have been repeatedly investigated by means of theoretical methods, in particular molecular dynamics simulations.[68–74] Yet, until now, the influence of complex formation on the flexibility of both proteins has not been investigated. Recently, we studied structural determinants of the binding free energy of Ras-Raf by means of absolute binding free energy calculations and free energy decomposition.[75] For Raf, we identified pathways of interacting residues that originate in the binding epitope and protrude through the protein structure. From this, an influence on the protein flexibility in particular of Raf due to complex formation may be anticipated, which may also provide an explanation for the activation of Raf upon Ras binding.

METHODS

Molecular Dynamics Simulations

The simulations and the structural properties of Raf, Ras, and Ras-Raf have been described in detail else-

where.[75] Briefly, the AMBER 7 suite of programs[76] together with the Cornell et al. force field[77] and the TIP3P model for water[78] were used for all three simulations. Starting structures for the human proteins were taken from the Protein Data Bank (PDB codes: 1rrb, 121p, 1gua) and modified to achieve consistency with respect to the biological source and the number of amino acids (see Gohlke et al.[75] for details). In particular, since to-date no experimentally determined structure for the Ras-Raf complex has been reported, the Ras-Raf complex has been modeled by superimposing the Ras structure from 121p onto the Raps structure from 1gua, as described in Terada et al.[79] Bonded parameters for the triphosphate moiety of GTP were taken from Leach and Klein[80] and atomic partial charges for GTP^{4-} were derived using the RESP procedure.[81] Non-bonded parameters for Mg^{2+} were taken from Aqvist.[82] The particle mesh Ewald (PME) method was used with a direct-space non-bonded cutoff of 9 Å. Bond lengths involving bonds to hydrogens were constrained with SHAKE and the time step for all MD simulations was 2 fs. After equilibration, 2 ns of unconstrained MD in the canonical ensemble (NVT) at 300 K with a time constant of 2.0 ps for heat bath coupling were performed before 500 snapshots for FIRST analyses were extracted at time intervals of 20 ps from production runs of 10 ns length each. All counterions and water molecules except the two closest to the Mg^{2+} ion in Ras and Ras-Raf were stripped from the structures.

To analyze dynamic properties of the molecules, snapshots saved at 100 fs intervals along the 10 ns production runs were used. The root-mean-square amplitude of motion about the mean position of atom i is determined from the MD trajectory by

$$\langle u_i^2 \rangle^{1/2} = \left[\frac{1}{N} \sum_{j=1}^{N} \Delta x_i(t_j)^2 \right]^{1/2} \quad (1)$$

where $\Delta x_i(t_j) = x_i(t_j) - \bar{x}_i$ with $x_i(t_j)$ is the coordinate vector of atom i at time step t_j and $\bar{x}_i$ is the mean position of atom i during the sample period. N is the number of samples.

Experimental root-mean-square fluctuations about the mean position of atom i can be estimated from the crystallographic temperature factor B_i:[83]

$$\langle u_i^2 \rangle^{1/2} = (3B_i/8\pi^2)^{1/2} \quad (2)$$

Equal-time cross-correlations of the atomic fluctuations are obtained by

$$C(i,j) = \langle \Delta x_i(t_k) \cdot \Delta x_j(t_k) \rangle / (\langle \Delta x_i(t_k)^2 \rangle^{1/2} \langle \Delta x_j(t_k)^2 \rangle^{1/2}), \quad (3)$$

where $\langle \cdots \rangle$ indicates an average over the sample period.[84] For completely correlated motions, $C(i,j) = 1$, and for completely anticorrelated motions, $C(i,j) = -1$.

For the calculation of $\Delta x_i(t_j)$, global translational and rotational differences between the structures along the trajectory need to be removed by least-squares fitting. It has been noted that the choice of the atom set used for fitting can have an influence on the picture of internal motions.[85–87] Hence, we tested this influence by comparing results obtained using either all C_α atoms or only those in secondary structure elements for the fitting. Since no qualitatively different answers were found for both cases (data not shown), we report results here obtained by including all C_α atoms into the least-squares fitting.

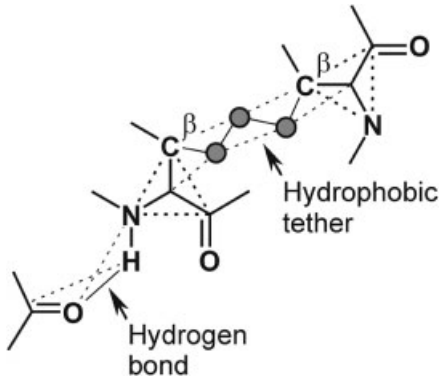


Fig. 2. Bond-bending network where vertices represent protein atoms and edges represent covalent (thick lines) and non-covalent (thin lines) constraints within the molecule. Constraints between next-nearest neighbors (broken lines) define coordination angles between bonded atoms. Hydrogen bonds are modeled by a bond between the hydrogen and the acceptor atom and two additional angular constraints associated with these atoms. Hydrophobic interactions are modeled by a flexible linkage ("hydrophobic tether") consisting of three additional vertices (gray spheres). All single bonds may undergo dihedral rotation unless they are locked by the network of bonds in which they participate.

Finally, residues are considered to be in the binding interface if at least 10% of their solvent-accessible surface in the unbound state is buried upon complex formation. Assignments of secondary structure elements refer to the unbound proteins and were taken from the PDBsum database.[88]

FIRST Analysis

FIRST identifies flexible and rigid regions in three-dimensional bond molecular frameworks. The algorithm and the underlying mathematical rigidity theory have been detailed elsewhere.[50–52,56] Here, we describe the approach briefly to introduce the terminology used below. FIRST applies the pebble game algorithm[52] to identify and count the bond-rotational degrees of freedom in a directed graph, whose vertices represent protein atoms and whose edges represent covalent and non-covalent (hydrogen-bond and hydrophobic) constraints within the protein[9,51,56] (Fig. 2). Flexibility in this network results from dihedral rotations of bonds that are not locked in by other bonds ("hinge joints"). Each bond is assigned by FIRST to be part of either a rigid cluster or a flexible (under-constrained) region. A rigid cluster forms a collection of interlocked bonds in which no relative motion can be achieved without a cost in energy. If a rigid cluster does not contain redundant bond constraints, it is minimally rigid or "isostatic." Conversely, if a rigid cluster contains redundant bond constraints, stress is introduced within this region. Such an over-constrained region is more stable than an isostatic (just rigid) region, in that it remains rigid even if

one of the bonds is broken. Under-constrained regions typically are flexible links between rigid clusters. A number of degrees of bond-rotational freedom (so called "floppy modes") are associated with each under-constrained region. Note that this number of internal degrees of freedom within that region is usually much smaller than the actual number of rotatable bonds, because not all the rotatable dihedral angles associated with hinge joints are independent (as they are usually part of a ring of constraints formed by covalent and non-covalent interactions). Finally, FIRST identifies distinct collective motions, each of which consists of coupled rotatable bonds. These motions occur within a particular under-constrained region without affecting internal coordinates outside of this region.

A continuous flexibility index f_i has been defined[56] that characterizes the degree of flexibility of the ith bond in the network [Eq. (4)]. Going beyond the qualitative distinction of regions as over-constrained (stressed), isostatic (just rigid), or under-constrained (flexible), this index allows quantification of *how much more* flexible an under-constrained region is compared to an isostatically rigid region or *how much more* stable an over-constrained region is.

$$f_i = \begin{cases} \dfrac{F_j}{H_j} & \text{in an underconstrained region} \\ 0 & \text{in an isostatically rigid region} \\ -\dfrac{R_k}{C_k} & \text{in an overconstrained region} \end{cases} \quad (4)$$

Here, F_j and H_j are the number of independent degrees of bond-rotational freedom and the number of potentially rotatable bonds (independent or not) within the jth under-constrained region (i.e., the region containing bond i). R_k and C_k are the number of redundant bonds and the total number of bonds, respectively, in the kth overconstrained region (again, the region containing bond i). Since $F_j = H_j$, it follows that $0 < f_i = 1$ for underconstrained regions. Similarly, for overconstrained regions, $R_k = C_k$, and f_i is bounded by 0 and -1.

In previous studies with FIRST ,[9,51,56,57] protein structures determined by X-ray crystallography have been used. Here, we perform FIRST analyses for a series of snapshots extracted from MD trajectories. Instead of providing a unique assignment for each atom as part of a rigid or flexible region based on a single input structure, we can now define the probability $P_j(i)$ that atom i belongs to the jth largest cluster of mutually rigid atoms.

$$P_j(i) = \frac{n_j(i)}{N} \quad (5)$$

$n_j(i)$ is the number of occurrences of atom i as part of the jth rigid cluster, determined over all N snapshots. $P_j(i)$ is expected to provide a more accurate picture of local rigidity compared to the "all-or-nothing" answer given by only one input structure. This is expected to be particularly valuable for substructures that are close to isostatic and, thus, may change from flexible to rigid or vice versa upon formation or breaking of one or a few non-covalent bonds as a result of the inherent mobility of proteins.

Along these lines, a flexibility index $\Phi(i)$ for C_α atom i is calculated by averaging over all flexibility indices $f_{i,j}(k)$ [Eq.(4)] of the two backbone bonds j originating from this atom as well as averaging over all snapshots k.

$$\Phi(i) = \frac{1}{2N}\sum_{j=1}^{2}\sum_{k=1}^{N} f_{i,j}(k) \quad (6)$$

The sum over j is necessary because the flexibility index is a property of bonds, not of atoms. In most cases, the N-C_α and C_α-C′ bonds will belong to the same under- or over-constrained region. If this is not the case, the flexibility index $\Phi(i)$ assigned to the ith C_α is the average of the two backbone bonds that connect to it.

Applying FIRST to Proteins

The bond network used as input to FIRST is completely defined by bond constraints between atoms as well as next-nearest neighbor constraints that define the coordination angles between bonded atoms (Fig. 2).[51] The goal is to identify macroscopically significant flexibility, which is in general associated with low-frequency structural fluctuations, rather than the small-amplitude, high-frequency motions that arise from bond stretching and bending. By including constraints into the network that represent strong forces between atoms, the high-frequency motions can be effectively quenched. Here, covalent and hydrogen bonds, salt bridges, and hydrophobic interactions are considered to be strong forces.

Bond lengths, represented as distance constraints between the bonded atoms, and bond coordination angles, represented as distance constraints between next-nearest neighbors of the central atom, are set to their values observed in the input protein structure. Additional constraints are set to restrict the configuration of double or partial double bonds, such as peptide bonds (see Rader et al.[9] for details). Hydrogen bonds have a prominent role for secondary structure formation and stabilization of tertiary structure due to interactions betweens parts of the protein that are distant in sequence. Thus, they structurally stabilize, or rigidify, large parts of the protein. Hydrogen bonds are modeled by a bond between the hydrogen and the acceptor atom and two additional angular constraints associated with these atoms, removing three degrees of bond-rotational freedom from the system.[56] Here, we use donor atom–hydrogen-acceptor atom geometries as given in snapshots extracted from molecular dynamics trajectories (see below) to define hydrogen bond strengths. The latter are obtained by a modification of the hydrogen-bond energy function of Mayo and coworkers,[89] which has been used in previous FIRST studies,[9,57] adjusting the angular dependence to favor more linear hydrogen bonds and increasing the well depth and distance-dependence as well as removing the angular term for salt bridges, due to their substantial Coulombic component.

$$E_{HB} = V_0\left\{5\left(\frac{d_0}{d}\right)^{12} - 6\left(\frac{d_0}{d}\right)^{10}\right\}A(\theta,\phi,\varphi) \quad (7)$$

where

$V_0 = 8$ kcal mol^{-1}; $d_0 = 2.8$Å

sp^3 donor – sp^3 acceptor	$A = \cos^2\theta\exp(-(\pi - \theta)^6)\cos^2(\phi - 109.5)$
sp^3 donor – sp^2 acceptor	$A = \cos^2\theta\exp(-(\pi - \theta)^6)\cos^2\phi$
sp^2 donor – sp^3 acceptor	$A = \cos^4\theta\exp(-2(\pi - \theta)^6)$
sp^2 donor – sp^2 acceptor	$A = \cos^2\theta\exp(-(\pi - \theta)^6)\cos^2(\max[\phi,\varphi])$

Thus, the quality of a hydrogen bond depends on the deviation of its donor-acceptor distance d from an optimal value d_0 as well as the angular function A, which is dependent on the hybridization of donor and acceptor and the angles θ (angle between donor atom— hydrogen-acceptor atom), ϕ (angle between hydrogen atom-acceptor—base atom bonded to acceptor), and φ (angle between normals of two planes defined by the sp^2 centers). If ϕ is less than 90°, the supplement of the angle is used. V_0 is the well depth of the interaction.

Salt bridges are defined to occur between the negatively charged groups of glutamate, aspartate, or the C-terminal carboxylate and the positively charged groups of histidine, lysine, arginine, or the N-terminal amino group.

$$E_{SB} = V_s\left\{5\left(\frac{d_s}{d+a}\right)^{12} - 6\left(\frac{d_s}{d+a}\right)^{10}\right\} \qquad (8)$$

where

$$V_S = 10 \text{ kcal mol}^{-1};\ d_S = 3.2\text{Å};\ a = 0.375\text{Å}$$

The values for V_S (well-depth of the salt bridge), d_S (optimal distance between donor and acceptor atoms), and a were selected such that the computed energies matched experimental ones.[90] The distance between donor and acceptor atoms is denoted by d.

Knowing the energies of hydrogen bonds and salt bridges (from now on, together referred to as hydrogen bonds), these values can be used to select hydrogen bonds for inclusion as constraints into the network. On going from a high hydrogen-bond energy threshold value (thus including many hydrogen bonds, resulting in few but large rigid clusters) to a low energy threshold (thus including only a few hydrogen bonds, resulting in a floppy protein with many small rigid clusters), the flexibility characteristics of the protein will change. Using $E_{cut} = -kT = -0.6$ kcal mol^{-1} at $T = 298$ K appears a reasonable choice. However, because of the approximate energy function, which does not consider the environmental dependence of the hydrogen bond strength, we repeated the FIRST analyses, considering hydrogen bonds with energy values below -0.1 kcal mol^{-1} (which had been used in earlier studies[56]) or -1.0 kcal mol^{-1}. A comparison of the results indicates that the outcome of FIRST is robust with respect to the choice of E_{cut} (see Appendix).

Finally, hydrophobic interactions between carbon and/or sulfur atoms are taken into account, if the distance between these atoms is less than the sum of their van der Waals radii (C: 1.7 Å, S 1.8 Å) plus 0.25 Å. To model the fact that hydrophobic interactions generally allow the atoms to move with respect to each other somewhat, yet not pull apart, hydrophobic tethers are included in the network. Each tether is a flexible linkage that removes two degrees of bond-rotational freedom from the system.[9] In addition, the influence of omitting all hydrophobic interactions in FIRST, while maintaining hydrogen bonds of a strength of at least $E_{cut} = -0.1$ kcal mol^{-1}, was tested (see Appendix).

Non-bonded interactions between the GTP ligand of Ras and the surrounding protein were treated analogously to interactions between protein atoms. Interactions between the Mg^{2+} ion and its six nearest neighbors were modeled as covalent bonds, which is in agreement with the observation that these interactions persist throughout the simulation time. To test the influence of a structural (persistently bound) water identified by MD simulation in the interface of Ras-Raf, FIRST calculations for the complex were performed either including or excluding the additional hydrogen bond constraints between the water and the proteins. Negligible differences in the flexibility characteristics of Ras-Raf were found in these cases (see Appendix).

In this study, snapshots extracted from molecular dynamics trajectories are used as input structures for FIRST analyses. Using multiple conformations and hydrogen-bond configurations for each protein allows one to assess how strongly FIRST results depend on the input data.

RESULTS AND DISCUSSION

Root-Mean-Square Atomic Fluctuations From MD for Unbound Ras and Raf and Comparison to Experiment

Root-mean-square atomic fluctuations of C_α atoms of unbound Ras and Raf as obtained from molecular dynamics simulations [Eq. (1)] are given in Figure 3(a) and Figure 4(a), respectively. The MD results are compared to characteristics of atomic motions as obtained from experiment. Figure 3(b) depicts root-mean-square atomic fluctuations calculated from crystallographic temperature factors [Eq. (2)] of four unbound Ras structures (PDB codes: 121p, 1ctq, 1qra, 5p21). Figure 4(b) shows averaged root-mean-square displacements of C_α atoms with respect to the average structure obtained for an ensemble of 30 NMR structures of unbound Raf (PDB code: 1rfa).

The average atomic fluctuations of unbound Ras as obtained by MD simulations (0.71 Å) are in good agreement with those found from the crystallographic tempera-

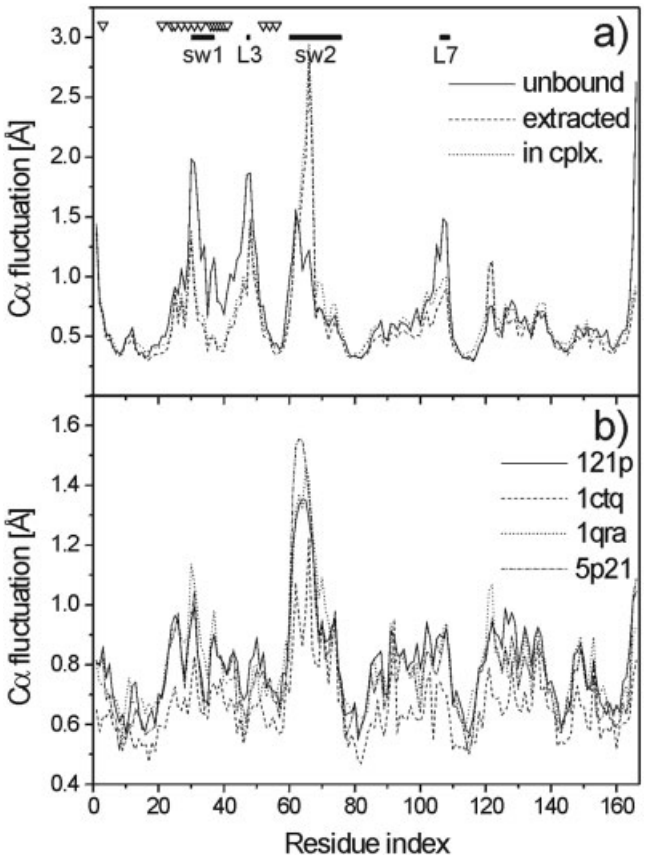


Fig. 3. **a**: Root-mean-square atomic fluctuations of C_α atoms of Ras in unbound (solid line) and bound (using least-squares fitting of bound, yet extracted, Ras conformations only: dashed line; using least-squares fitting of the whole complex Ras–Raf: dotted line) state. Residues in the interface are marked with ▽. The switch 1 and switch 2 regions and secondary structure elements discussed in the text are indicated by bold lines. **b**: C_α atomic fluctuations as calculated by [Eq. (2)] from temperature factors of four different X-ray crystal structures of Ras bound to GTP analogs (PDB code: 121p, 1ctq, 1qra, 5p21).

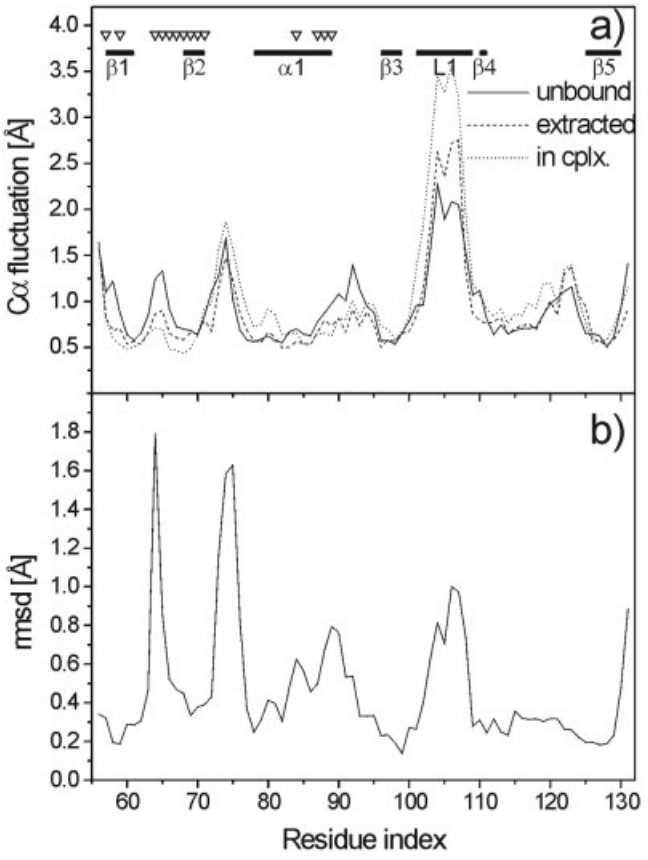


Fig. 4. **a**: Root-mean-square atomic fluctuations of C_α atoms of Raf in unbound (solid line) and bound (using least-squares fitting of bound, yet extracted, Raf conformations only: dashed line; using least-squares fitting of the whole complex Ras-Raf: dotted line) state. Residues in the interface are marked with ▽. Secondary structure elements discussed in the text are indicated by bold lines. **b**: Root-mean-square displacements with respect to the average structure obtained for a set of 30 NMR structures of the PDB entry 1rfa.

ture factors (0.66–0.80 Å). In addition, the overall trends between computed and experimental fluctuations are similar, although the fine details do not compare perfectly. Thus, the switch 1 region of unbound Ras, residues 45–50 (loop L3), and residues 105–108 (loop L7) are predicted to be more mobile compared to experiment. This results in correlation coefficients of the calculated fluctuations and the experimental ones of only 0.50–0.55, whereas the correlation coefficients of fluctuations obtained from the four different crystal structures are > 0.80. In the case of unbound Raf, the location of maxima of computed fluctuations again compare favorably with the ones of root-mean-square displacements determined from the ensemble of NMR structures. However, calculated fluctuations for regions around residues 65 and 75, respectively, are predicted to be smaller than the ones of loop L1, whereas an opposite trend is observed in the experimental data.[63]

Hence, in both Ras and Raf, overall trends between computed and experimentally determined fluctuations are in fair agreement, taking into account that the time-scales of averaging over structural and dynamical properties differ by orders of magnitude between crystallographic experiments and MD simulations and that calculated fluctuations and the ones obtained from crystallographic temperature factors include different contributions to the atomic motions.[28,91–93]

Influence of Complex Formation on Atomic Fluctuations of Ras and Raf

In Figure 3(a) and Figure 4(a), atomic fluctuations of C_α atoms are also depicted for the Ras-Raf complex ("in cplx."). In the latter case, the fluctuations describe not only internal motions of the binding partners, but also include information about (rigid-body) motions of the two proteins with respect to each other. To obtain information solely about internal motions of the proteins in the bound state, we also calculated atomic fluctuations after least-squares fitting conformations of bound Ras or Raf only, extracted from the complex structure ("extracted"). It should be noted that to date, no experimentally determined structure has been reported for the Ras-Raf complex.

Restrictions of C_α atomic fluctuations upon complex formation are observed for residues in the interface of Ras

and Raf (marked by ▽ in Figure 3(a) and Figure 4(a)). This is particularly obvious for residues 32–40 of the "effector region" of Ras and the strand β1 and the C-terminal end of helix α1 in Raf. Loop L3, which is adjacent to the interface, also shows a decrease in the fluctuations. Surprisingly, the loop L7 at the C-terminal end of helix α3 in Ras (residues 105–108) also shows reduced mobility, although it is more than 20 Å apart from the binding epitope.

In contrast, residues 64–67 of Ras (within the switch 2 region) as well as loop L1 of Raf show considerably increased fluctuations upon complexation. For Ras, this finding is nearly independent of whether rigid-body motions between the two proteins are included ("in cplx.") or not ("extracted"). In the case of Raf, however, there is a contribution due to rigid-body motions to the fluctuations of L1 in the complex ("in cplx."), as indicated when comparing these fluctuation values to those where only internal motions of Raf in the bound state are taken into account ("extracted"). Such an overall motion of the effector protein with respect to Ras has also been described by Zeng et al.[69] In both cases, complex formation is accompanied with changes in the flexibility at sites that do not interact directly with the respective binding partner.

The increased movements in the first part of switch 2 of bound Ras can be largely attributed to the occurrence of two conformations, whereby the conformational modification occurs after approximately 5.6 ns of simulation time. This transition is manifested in the change of ψ angles of residues M67 (−15° → −40°) and R68 (−25° → −50°), whereas all other backbone dihedrals in the switch 2 region remain essentially unchanged over the course of the simulation (data not shown). In fact, considerable variations in the conformation of switch 2 have also been found for GTP-analog bound Ras in structures determined from different crystal forms as well as for crystallographically independent Ras molecules of one X-ray structure,[60,94] which points to the inherent flexibility of this region. Thus, it is not clear whether complex formation of Ras with Raf leads to the increased mobility of the L4 part of switch 2 compared to the case of unbound Ras or whether a conformational transition of this region has not yet been observed for unbound Ras due to an insufficient simulation time.

In the case of increased fluctuations of loop L1, we have shown recently by free energy decomposition for the Ras-Raf system that the binding event of the two proteins leads to changes in "effective energies" (gas-phase energies plus solvation free energies) for residues apart from the binding interface and that these changes are mostly regional, not global.[75] In particular, we have identified pathways of energetic coupling that originate in the epitope and percolate through the Raf structure. The longest pathway through Raf is found in the case of unfavorable interactions between two residues; it extends from α1 over β2, β1, and β5 into the region of β3 and β4 adjacent to the loop L1. Interestingly, this β-sheet region does not only show strongly correlated motions after complex formation, but it is also identified by FIRST to become stressed (over-constrained) upon binding (see below). Furthermore, since this pathway extends from the binding interface into the region between L101 and K106 and shows significant interactions between the residues, observed differences in the flexibility of L1 in bound Raf compared to unbound Raf may be largely influenced by it.

Changes in Correlated Motions Upon Complex Formation

Orientational correlations between motions of C_α atoms have been calculated from Eq. (3) and are displayed color-coded in Figure 5 and Figure 6(a). The correlation values vary in the range between −1 (dark blue) and 1 (dark red), the lower and upper bound for fully anti-correlated and correlated motions, respectively. A value of zero (green) indicates uncorrelated motions. The two axes of the maps in Figure 5 and Figure 6(a) represent residue indices for Ras and Raf, respectively. The upper left triangles correspond to motions of the unbound proteins, while the lower right triangles show correlations for bound Ras and Raf. In the latter cases, the proteins have only been superimposed on the Ras or Raf part of Ras-Raf, respectively; thus, rigid-body motions of both binding partners with respect to each other are not included in the analyses. Hence, comparing upper and lower triangles reveals changes due to complex formation in internal motions of the proteins only.

The map for unbound Ras has little structure: correlated motions are solely revealed between the C-terminal end of β2, the loop region L3, and the N-terminal end of β3 (residues 38 to 54), whereas anti-correlated motions occur between helix α1 and loop L3 as well as between loop L3 and the N-terminal region of helix α5. In the bound state, these (anti-)correlations become less pronounced or even vanish. In contrast, weak anti-correlated motions can now be observed between residues in the effector region and those of the loop L4 part of switch 2. Overall, complex formation does not seem to have much influence on the orientational correlation between motions of Ras.

A different picture emerges from the correlation map of Raf. In the unbound state, correlated motions can be observed here in particular between the C-terminal end of β1 and β5 as well as between β3 and β5. Weakly anti-correlated motions emerge between the turn between β1 and β2 and the loop L1 region. In the bound state, these orientational correlations become considerably more pronounced. Now, the β-sheet shows a strongly correlated movement, with particularly intense couplings between β1, β2, and β5 as well as between β3, β4, and β5. In turn, loop L1 moves in an anti-correlated fashion with respect to the β-sheet [Fig. 6(b)]. In the case of Raf, complex formation, thus, does not only affect the magnitude of fluctuations of individual residues as described in the previous section but also the directionalities of motions of subsets of residues with respect to each other.

Global Changes in Flexibility Identified by FIRST Analyses

While considerable computational resources are needed to identify flexible or rigid regions of a protein by means of

TABLE I. Statistics on FIRST Calculations for Raf, Ras, and Ras-Raf

	Raf		Ras		Ras-Raf
Flexibility measures	Unbound	Bound[a]	Unbound	Bound	
Number of atoms in the largest rigid cluster	318.1 (11.8)	319.2 (11.6)	1889.1 (8.1)	2011.9 (8.0)	2717.3 (14.1)
Number of flexible regions exhibiting collective motions	3.3 (0.1)	3.9 (0.1)	15.1 (0.1)	14.0 (0.1)	19.7 (0.2)
Number of atoms in the largest region exhibiting collective motions	282.1 (5.3)	271.3 (5.7)	86.8 (3.0)	64.8 (2.2)	143.7 (5.7)
Number of bond-rotational degrees of freedom	138.9 (0.6)	143.7 (0.6)	226.5 (0.7)	214.0 (0.7)	327.3 (0.8)

†The values were averaged over 500 snapshots extracted from MD simulations. The hydrogen bond energy cutoff was set to $E_{cut} = -0.6$ kcal mol^{-1}. Values in parentheses are the standard error in the mean, obtained by dividing the standard deviation by the square root of the number of snapshots.

[a]Raf or Ras conformations extracted from the complex, considered without their binding partners.

MD simulations, FIRST can predict the intrinsic flexibility of a macromolecule from a single 3D structure in about one second. Several applications of this approach using experimentally determined protein structures as input have been presented.[9,56,57] Here, we performed FIRST analyses on ensembles of 500 snapshots of unbound and bound Ras and Raf, extracted from MD trajectories, and compared the results to the dynamical information obtained directly from the simulation.

Statistics on the number and size of rigid clusters, collective motions, and floppy modes as determined by FIRST calculations for Ras, Raf, and Ras-Raf are given in Table I. Also considering the FIRST flexibility predictions for Ras or Raf from the Ras-Raf complex while removing the partner (Raf or Ras) allows us to assess to what extent conformational changes upon complexation contribute to changes in the flexibility, compared to the contribution of additional interactions across the binding interface of the complex. For each flexibility measure in Table I, the values in brackets denote the standard error in the mean. These values have been obtained by dividing the standard deviation of the distributions by the square-root of the number of snapshots, because the snapshots are statistically independent (as determined from correlation functions of the time series of data points; data not shown).[95] In all cases, the errors in the mean are at most 4%. Hence, with respect to global changes in flexibility, FIRST results are rather insensitive to which MD snapshot is analyzed, as has also been found when different experimental 3D structures have been used.[56]

The mean values of the size of the largest rigid cluster and the number of independent hinge joints reveal *global* changes in the flexibility of the binding partners due to protein–protein association. The size of the largest rigid cluster in the case of Ras-Raf is 2717 atoms, whereas the sum of the sizes of the largest rigid clusters for unbound Ras and Raf amounts to only 2207 atoms. The latter number would be expected if complex formation did not affect the conformation or flexibility of the two proteins. Since covalent bonding is identical for both the unbound proteins and the complex, non-covalent interactions across the interface plus conformational changes between unbound and bound Ras and Raf must lead to an extension of rigid regions in the proteins upon Ras–Raf association. Along these lines, the number of independent bond-rotational degrees of freedom in Ras–Raf (327) is smaller by 38 than the sum of these values for the unbound Ras and Raf (365), again revealing that complex formation overall decreases the internal degrees of freedom of the binding partners. This picture is consistent with a decrease in vibrational entropy upon formation of the complex, as determined by normal-mode calculations.[75,96]

It is important to note that differences arising from conformational changes of the proteins upon complex formation are overlaid on those due to direct interactions across the interface. As such, the unbound to bound transition in the case of Ras already accounts for an increase in the size of the largest rigid cluster of 123 atoms, whereas the bound Raf has more independent bond-rotational degrees of freedom (143) than unbound Raf (139), i.e., Raf in the bound conformation is slightly more flexible than in the unbound state.

Local Flexibility Characteristics as Determined by FIRST and Comparison to MD Results

By means of rigid cluster decomposition, FIRST identifies structurally rigid regions in a protein as well as flexible bonds between them. FIRST thus provides information about *local* flexibility characteristics, which are dependent on the density and placement of (covalent and noncovalent) interactions between atoms in the network. Figures 7–9 depict results of the rigid cluster decomposition for Ras, Raf, and Ras–Raf, respectively. In part (a) of the figures, the probability of being part of the largest rigid cluster for C_α atoms of the proteins is color-coded, ranging from dark blue, representing a probability of one, to dark red, representing a probability of zero. While these probabilities were obtained by FIRST analyses over 500 snapshots for each protein, the (b) panels depict rigid clusters identified for a representative snapshot from these ensembles. In the latter case, dark blue represents the largest cluster found for this snapshot.

A visual inspection of the rigid cluster decomposition obtained for several snapshots of the proteins revealed for Ras and Ras–Raf that the results do not change significantly along the trajectory, which is in agreement with the finding that large parts of these proteins are overconstrained (i.e., some of the constraints in the framework are redundant; see below). Hence, fluctuations in the constraints do not have a noticeable effect on the outcome of the FIRST analysis. For Raf, however, the rigid cluster

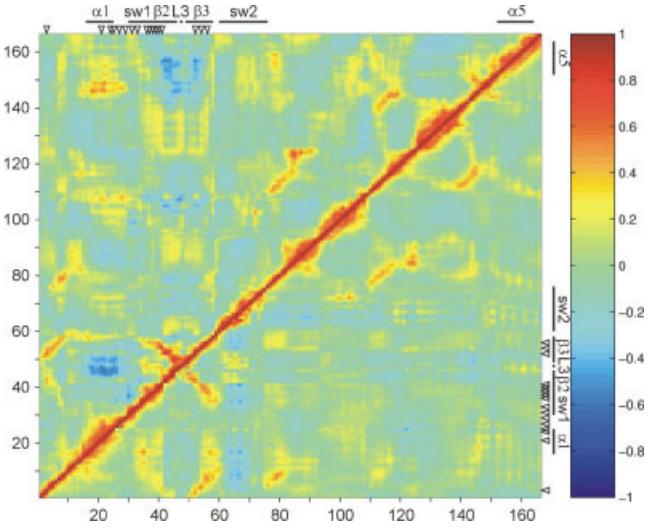
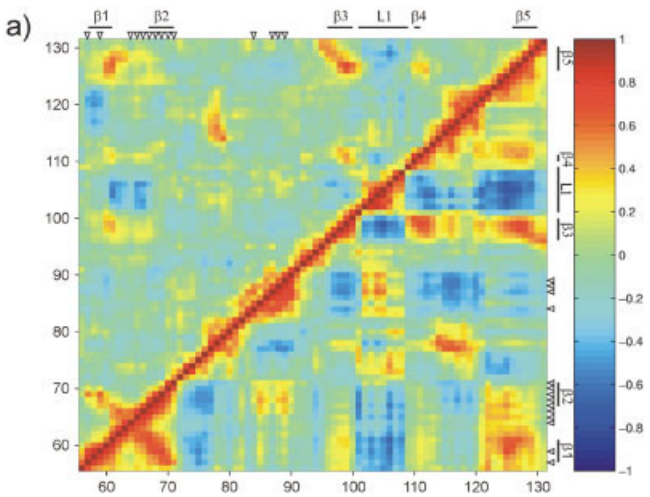


Fig. 5.



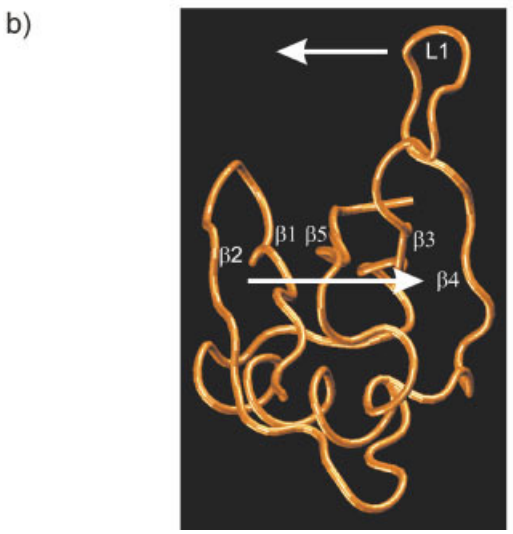
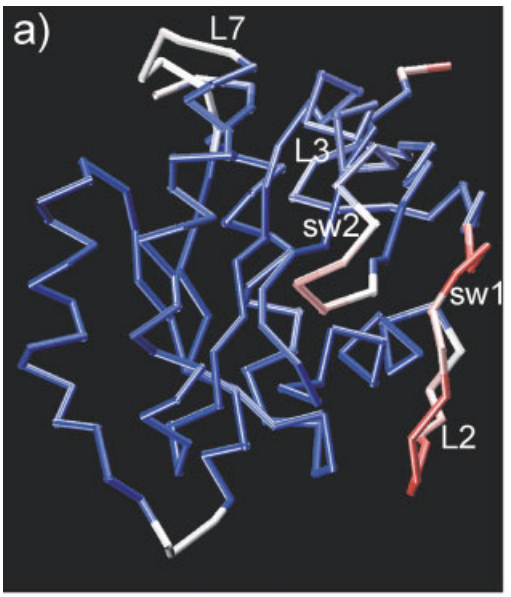


Fig. 6.



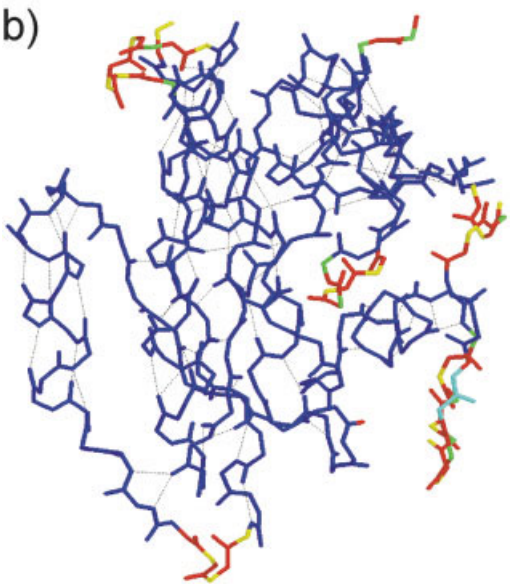


Fig. 5. Cross-correlation map of C_α atomic fluctuations of Ras. The two axes refer to residue indices. Positive correlations are indicated in red, negative correlations in blue. See also the color scale to the right. The correlations for the unbound protein are shown in the upper left triangle. The correlations for the bound Ras molecule are given in the lower right triangle. In the latter case, the molecule was least-squares fitted only onto itself, i.e., relative motions between Ras and Raf in the Ras–Raf complex are not contained in the map. Residues in the interface are marked with ▽. The switch 1 and switch 2 regions and secondary structure elements discussed in the text are indicated by bold lines.

Fig. 6. Cross-correlation map of C_α atomic fluctuations of Raf (**a**). Secondary structure elements discussed in the text are indicated by bold lines. For further details see Figure 5. In part (**b**), a schematic representation of the anti-correlated motion between the β-sheet consisting of β1 to β5 and the loop L1 region that occurs in the bound state of Raf is given.

Fig. 7. **a**: Color-coded projection of the probability of being part of the largest rigid cluster for C_α atoms of unbound Ras. Probability values were determined from 500 snapshots extracted from a MD trajectory using Eq. (5). Dark red color codes for zero probability, white color for a probability of 0.5, and dark blue color for a probability of 1.0. **b**: Rigid cluster decomposition of unbound Ras. Rigid clusters are colored according to their size in blue, cyan, black, yellow, red, and green. Hydrogen bonds between backbone atoms are depicted as black dotted lines.

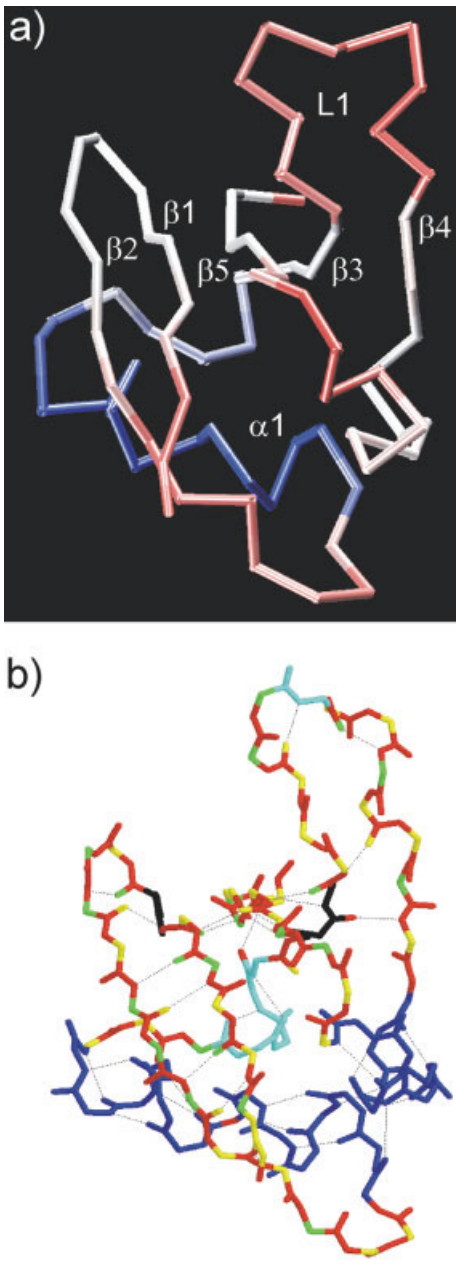


Fig. 8. **a**: Color-coded projection of the probability of being part of the largest rigid cluster for C_α atoms of unbound Raf. **b**: Rigid cluster decomposition of unbound Raf. For further information see Figure 7.

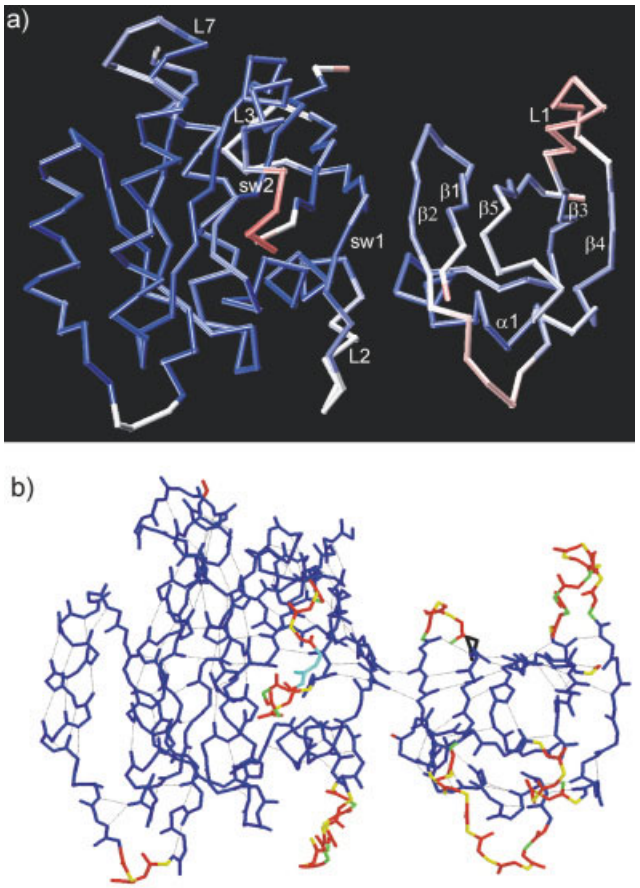


Fig. 9. **a**: Color-coded projection of the probability of being part of the first rigid cluster for C_α atoms of the complex Ras–Raf. **b**: Rigid cluster decomposition of Ras–Raf. For further information see Figure 7.

decomposition of single snapshots may lead to different conclusions in some cases compared to those obtained by averaging over all snapshots [Fig. 8(a)]. This corresponds to the notion that Raf is metastable with respect to its flexibility characteristics (see below), and, thus, addition or removal of a constraint may influence the rigidity analysis. Consequently, while averaging over 500 snapshots mostly gives a more realistic delineation of where rigid regions end and flexible regions start in the case of Ras and Ras–Raf, for metastable systems such as Raf, considering only a single input structure may be misleading.

In addition, the averaged flexibility indices Φ [Eq. (6)] for C_α atoms of Ras and Raf in unbound state or as part of the Ras–Raf complex are given in Figure 10. Values of Φ as obtained for Ras and Raf in bound conformation, yet extracted from the complex, are shown, too. The Φ index allows us to quantify the flexibility or rigidity of each residue across the MD trajectory. Finally, Table II summarizes changes in local flexibility characteristics of the proteins as determined by MD or FIRST, indicating that the results of both methods agree in 5 out of 6 regions.

Ras and Raf in Unbound State

Figure 7 indicates that most of the Ras protein is part of the largest rigid cluster, exceptions being loop L2 and the switch 1 region, switch 2, and (less pronounced) loop L7. In contrast, for Raf, FIRST analysis reveals that only helix α1 (Fig. 8) shows a high probability of being part of the largest rigid cluster. Probabilities of approximately 0.5 are found for the β1–bend–β2 region, the C-terminal part of β5 and the β3 and β4 part, whereas the region between β2 and α1 and the loop L1 hardly ever participate in the largest rigid cluster. In that respect, the Ras-binding domain of Raf investigated here appears to be very flexible. This result does not change if the probability plots for the second

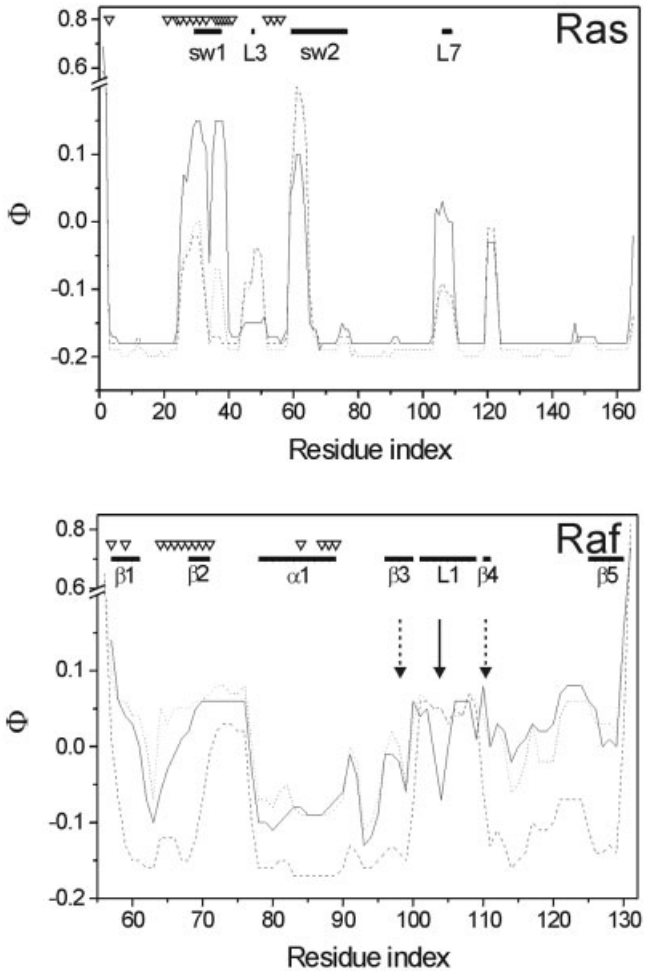


Fig. 10. Average flexibility index of Ras (top) and Raf (bottom) C_α atoms as determined by Eq. (6) (straight line: unbound Ras or Raf; dashed line: bound Ras or Raf; dotted line: Ras or Raf in bound conformation, extracted from the complex). Residues in the interface are marked with ▽. The switch 1 and switch 2 regions (Ras) and secondary structure elements discussed in the text are indicated by bold lines. In the case of Raf, the straight line arrow indicates part of the loop L1 region where conformational changes due to complex formation, but not additional interactions across the interface, lead to the changes in flexibility. In contrast, broken line arrows show that interactions across the interface result in the observed rigidification of the β-sheet, even for β3 and β4 that are most distant to the interface.

largest or smaller rigid clusters (data not shown) are considered. In all cases, the β-sheet β1–β5 never forms a rigid cluster by itself.

Bound Ras

Complex formation leads to small alterations in the flexibility of Ras (Fig. 9), as identified by FIRST. As one might have expected due to its proximity to the binding interface, the switch 1 region and loop L2 become more rigid, but so does loop L7 at the C-terminal end of helix α3, which is distant from the epitope. In contrast, the switch 2 region now shows a smaller probability of being part of the largest rigid cluster compared to the unbound Ras, i.e., FIRST identifies this region to be more flexible. A comparison of these results and the corresponding averaged flexibility indices of Ras C_α atoms (Fig. 10) to changes in the MD-determined atomic fluctuations of the protein on going from the unbound to the bound state (Figs. 3, 4) reveals a convincing agreement. Not only is the restriction of flexibility in the case of residues of and adjacent to the switch 1 region identified by both methods, but also the decrease of atomic fluctuations for loop L7 coincides with the transition of this region from under-constrained to over-constrained as determined by FIRST (see Fig. 10). Along these lines, FIRST and MD agree in finding increased flexibility for the switch 2 region. The almost identical Φ values found for this region in Ras–Raf and for Ras extracted from the complex also reveal that the change in flexibility of switch 2 results from conformational changes upon binding, whereas interactions across the interface do not have a direct influence.

TABLE II. Local Flexibility Changes Upon Formation of the Ras-Raf Complex

Region	Method[a]		Agreement[b]
	MD	FIRST	
Ras			
Switch1	Decreased	Decreased	+
L3	Decreased	Increased	−
Switch2	Increased	Increased	+
L7	Decreased	Decreased	+
Raf			
β-sheet	Decreased[c]	Decreased	+
L1	Increased	Increased	+

[a]The change in flexibility is reported with respect to the unbound proteins. Unless otherwise stated, in the case of MD, changes in flexibility are determined from changes in the root-mean-square atomic fluctuations.
[b]A "+" indicates that MD and FIRST agree with respect to the change in flexibility, a "−" indicates disagreement.
[c]Determined from changes in correlated motions. For further details see text.

FIRST results, however, deviate from MD with respect to loop L3. Here, the simulation shows a slight decrease in the fluctuations, whereas FIRST identifies this region to be more flexible in the bound state. At first sight, one may anticipate that this discrepancy arises from long-range (electrostatic) forces, which are taken into account in the MD simulation, but are not considered by FIRST. In the latter case, only short-range covalent and non-covalent constraints are considered, and any long-range effect is due to the percolation of flexibility or rigidity in the structural network. However, a pair-wise decomposition of "effective interaction energies" (i.e., differences in gas-phase plus solvation free energies occurring upon complex formation) for the Ras–Raf system has revealed that the resulting interaction matrix is rather "sparse" and that sizeable interaction energies occur only between spatially adjacent residues.[75] This has even been found for interactions between charged residues of the interface region, which interact most with their immediate counterparts of the binding partner. Thus, although it cannot be ruled out in general that not considering long-range electrostatic interactions in FIRST may affect the outcome of the

flexibility analysis, this effect is likely not the reason for the observed difference between MD and FIRST results for loop L3.

Bound Raf

Compared to Ras, protein–protein association significantly changes the flexibility properties of Raf in that now the β-sheet of the effector protein becomes part of the largest rigid cluster of the complex (Fig. 9). Conversely, the flexibility of the loop L1 is influenced in that the region now shows a lower probability of being part of the largest rigid cluster than in the unbound state. In particular, the very similar Φ values of this region found for Ras–Raf and for Raf extracted from the complex (see straight line arrow in Figure 10) demonstrate that conformational changes due to complex formation, but not additional interactions across the interface, lead to the changes in flexibility. In contrast, in the case of the β-sheet, interactions across the interface result in the observed rigidification, as is revealed by the lower Φ values found in Ras–Raf compared to those computed for the extracted Raf molecule (see e.g., broken line arrows in Figure 10). This clearly demonstrates the long-range aspect to rigidity percolation,[52,97,98] because changes of the interaction network at one side of the protein (here: the interface region) affect rigidity throughout the molecule (here: β3, β4, and β5).

Furthermore, it makes sense that changes in the flexibility upon association occur predominantly on the Raf side, because in the unbound state, large parts of the molecule are close to isostatically rigid. This means that small changes in the number and distribution of constraints can have a large effect on the stability of these regions, compared to highly over-constrained regions, where the breaking or forming of a few bonds does not grossly change the overall interaction network. This is in agreement with a recent study proposing that for the propagation of binding effects to distal regions of a protein, a significant fraction of residues with low structural stability in the uncomplexed binding site is necessary.[99] Finally, it is interesting to note that for Raf it has been consistently proposed that a conformational change between regulatory and kinase domains is a requirement for activation.[100,101] It is hence possible that the metastability in terms of flexibility/rigidity of the Ras-binding domain of Raf triggers such a change upon association with Ras.

Again, in comparing FIRST results with changes in the fluctuations as determined by MD, very good agreement is found for the decrease in flexibility especially for β2 as well as the increase in flexibility for the loop L1 region of Raf. However, the reduction in flexibility revealed for β3, β4, and β5 (Fig. 10) by FIRST is not reflected in the changes of atomic fluctuations for these residues. Yet, when comparing results obtained by MD and FIRST with respect to the flexibility of macromolecules, one needs to remember that FIRST analyzes a static bond network, involving only virtual atomic displacements. As such, a rigid cluster identified by FIRST is a collection of atoms in which no *relative* atomic motions occur. Conversely, flexible regions are sets of atoms that may move relative to one another. Flexibility information from MD simulations, however, derives from the evolution of the system with time, and a rigid cluster may well move in total, leading to the notion that its atoms are mobile if atomic displacements are used as metric. Hence, "flexibility" as determined by FIRST needs to be distinguished from "mobility" as determined by MD. In that respect, it seems more appropriate to compare FIRST results to analyses of correlated motions from MD simulations, in that regions whose motions are strongly correlated will correspond to rigid clusters as identified by FIRST analyses. Along these lines, the above mentioned strongly correlated motions found for the β-sheet in bound Raf, which appear less distinct in the unbound effector, very well reflect the change in rigidity identified by FIRST for these secondary structure elements.

CONCLUSION

We have investigated binding-induced changes in flexibility upon protein–protein complex formation of Ras-Raf, using a distance constraint approach, FIRST, and MD simulations. From a methodological point of view, it is encouraging that a comparison of the results demonstrates good agreement with respect to the identification of changes in flexibility in both binding partners on a local scale, because the computational time requirement for FIRST is several orders of magnitude smaller than for the MD simulation. This is both true for the unbound proteins and the Ras–Raf complex, such that *changes* in the flexibility upon protein–protein complex formation identified by both methods also correspond in 5 out of 6 regions (Table II). This agreement is by no means self-evident, although snapshots extracted from the MD trajectories are used as structural basis for the FIRST calculations. Namely, while atomic motions along a trajectory are governed by the continuous spectrum of forces exerted by surrounding atoms, in FIRST the constraints in the bond network are "all-or-nothing"—a bond is either present or absent. Especially in the case of non-covalent interactions, this leads to the need to distinguish forces sufficiently strong from weaker ones. The present results indicate that the simple model in FIRST provides essentially the same picture of dynamics over a moderate timescale as the much more detailed representation of covalent and non-covalent interactions in a widely used molecular mechanics force field. Phrased differently, the agreement indicates that dynamics can be predicted by the coupling between bonds that limits motion within a molecule.

With respect to the robustness of FIRST, we have shown that the analysis results do not strongly depend on the choice of the energy criterion for whether a hydrogen bond is included or not. However, hydrophobic interactions may be critical for a complete description of the flexibility characteristics of the proteins investigated. Finally, it is encouraging that FIRST results are remarkably consistent with respect to different input structures (at least if the investigated system is not metastable). This finding suggests that fewer than the 500 snapshots used in this study are sufficient for FIRST to give reasonable results on predicting flexibility, which should still correlate well with

the results of a long molecular dynamics trajectory. As an extreme example, FIRST results obtained from a single structure as input agreed well with experimentally determined mobility (as measured by crystallographic temperature factors)[56] and with protein folding cores identified by hydrogen/deuterium exchange NMR,[57] in which the experimental measures are obtained by averaging over ensemble and time scales that are out of reach for current MD simulations. Along these lines, FIRST results obtained from single structures were also remarkably consistent between different conformations.[56] This is related to the fact that the results of flexibility analysis by FIRST are remarkably independent of the conformation(s) used,[56] as long as the non-covalent bond network remains quite similar. In contrast, flexibility as determined by MD is influenced by which part of the configurational space of the molecule is accessible within the simulation time (which may have led to the underestimation of flexibility of the switch 2 region of unbound Ras). One should note, however, that for both, MD and FIRST, good experimental (starting) structures of unbound and bound molecules are a prerequisite to reach definite conclusions regarding biological significance of flexibility changes, because otherwise it is difficult to ensure that equilibrium conformations are being analyzed.

From a structural point of view, an interesting feature of FIRST is that it allows us to determine whether changes in the flexibility of the binding partners occur because of conformational transitions within the individual molecules upon complex formation and/or whether interactions across the interface exert a direct influence. The latter was found to affect the rigidity of the β-sheet of Raf, which also demonstrated the long-range effect of rigidity/flexibility percolation in bond networks. Such an effect could not be detected by methods that determine atomic fluctuations in proteins solely from spatial variations in *local* packing density.[45]

In terms of biology, perhaps the most important finding is the considerable decrease in flexibility of Raf upon binding. This change occurs predominantly in a region of the β-sheet where pathways of interacting residues originating in the binding epitope and percolating through the structure have been identified before by free energy decomposition.[75] It has been repeatedly observed that regions involved in molecular recognition exhibit an increased flexibility; prominent examples are the antigen binding loops of immunoglobulins[102,103] and the recognition loops of the HIV-1 protease.[104] Changes in protein dynamics due to a propagation of the binding effect through the structures may then lead to allosteric effects.[99,105] It is hence possible that the metastability in terms of flexibility/rigidity of the Ras-binding domain of Raf is a prerequisite for the activation of this protein by binding to Ras.

Considering the success of FIRST in describing the influence of binding on the flexibility of the binding partners and its computational efficiency, this method shall provide a valuable tool for gaining an understanding of the interplay between binding, flexibility, and function for even more complex macromolecular assemblies.

ACKNOWLEDGEMENTS

The present study was supported by NIH Grant RR12255 to D.C. and by NIH Mathematics in Biology grant GM67249 to L.K. H.G. gratefully acknowledges a Feodor-Lynen fellowship awarded by the Alexander-von-Humboldt foundation, Germany. Those interested in using FIRST can access the webserver version at http://firstweb.asu.edu or contact L.K. (KuhnL@msu.edu) regarding source-code distribution. Those interested in using the AMBER7 suite of programs will find ordering information at http://amber.scripps.edu.

REFERENCES

1. Koshland DE. Protein flexibility in enzyme action and enzyme control. Science 1967;156:540.
2. Rasmussen BF, Stock AM, Ringe D, Petsko GA. Crystalline ribonuclease A loses function below the dynamical transition at 220 K. Nature 1992;357:423–424.
3. Ferrand M, Dianoux AJ, Petry W, Zaccai G. Thermal motions and function of bacteriorhodopsin in purple membranes: Effects of temperature and hydration studied by neutron scattering. Proc Natl Acad Sci USA 1993;90:9668–9672.
4. Hayward JA, Smith JC. Temperature dependence of protein dynamics: computer simulation analysis of neutron scattering properties. Biophys J 2002;82:1216–1225.
5. Tang KE, Dill KA. Native protein fluctuations: the conformational-motion temperature and the inverse correlation of protein flexibility with protein stability. J Biomol Struct Dyn 1998;16:397–411.
6. Vihinen M, Torkkila E, Riikonen P. Accuracy of protein flexibility predictions. Proteins 1994;19:141–149.
7. Micheletti C, Lattanzi G, Maritan A. Elastic properties of proteins: insight on the folding process and evolutionary selection of native structures. J Mol Biol 2002;321:909–921.
8. Ferreira ST, De Felice FG. Protein dynamics, folding, and misfolding: from basic physical chemistry to human conformational diseases. FEBS Lett 2001;498:129–134.
9. Rader AJ, Hespenheide BM, Kuhn LA, Thorpe MF. Protein unfolding: rigidity lost. Proc Natl Acad Sci USA 2002;99:3540–3545.
10. Atwell S, Ultsch M, De Vos AM, Wells JA. Structural plasticity in a remodeled protein–protein interface. Science 1997;278:1125–1128.
11. Arkin MR, Randal M, DeLano WL, Hyde J, Luong TN, Oslob JD, Raphael DR, Taylor L, Wang J, McDowell RS, Wells JA, Braisted AC. Binding of small molecules to an adaptive protein–protein interface. Proc Natl Acad Sci USA 2003;100:1603–1608.
12. DeLano WL, Ultsch MH, de Vos AM, Wells JA. Convergent solutions to binding at a protein–protein interface. Science 2000;287:1279–1283.
13. Schymkowitz JW, Rousseau F, Wilkinson HR, Friedler A, Itzhaki LS. Observation of signal transduction in three-dimensional domain swapping. Nat Struct Biol 2001;8:888–892.
14. Freire E. Can allosteric regulation be predicted from structure? Proc Natl Acad Sci USA 2000;97:11680–11682.
15. Pan H, Lee JC, Hilser VJ. Binding sites in Escherichia coli dihydrofolate reductase communicate by modulating the conformational ensemble. Proc Natl Acad Sci USA 2000;97:12020–12025.
16. Yu L, Zhu CX, Tse-Dinh YC, Fesik SW. Backbone dynamics of the C-terminal domain of Escherichia coli topoisomerase I in the absence and presence of single-stranded DNA. Biochemistry 1996;35:9661–9666.
17. Olejniczak ET, Zhou MM, Fesik SW. Changes in the NMR-derived motional parameters of the insulin receptor substrate 1 phosphotyrosine binding domain upon binding to an interleukin 4 receptor phosphopeptide. Biochemistry 1997;36:4118–4124.
18. Zidek L, Novotny MV, Stone MJ. Increased protein backbone conformational entropy upon hydrophobic ligand binding. Nat Struct Biol 1999;6:1118–1121.
19. Loh AP, Pawley N, Nicholson LK, Oswald RE. An increase in side chain entropy facilitates effector binding: NMR characterization of the side chain methyl group dynamics in Cdc42Hs. Biochemistry 2001;40:4590–4600.

20. Arumugam S, Gao G, Patton BL, Semenchenko V, Brew K, Van Doren SR. Increased backbone mobility in β-barrel enhances entropy gain driving binding of N-TIMP-1 to MMP-3. J Mol Biol 2003;327:719–734.
21. Forman-Kay JD. The "dynamics" in the thermodynamics of binding. Nat Struct Biol 1999;6:1086–1087.
22. Kay LE, Muhandiram DR, Wolf G, Shoelson SE, Forman-Kay JD. Correlation between binding and dynamics at SH2 domain interfaces. Nat Struct Biol 1998;5:156.
23. Kay LE, Muhandiram DR, Farrow NA, Aubin Y, Forman-Kay JD. Correlation between dynamics and high affinity binding in an SH2 domain interaction. Biochemistry 1996;35:361–368.
24. Steinberg IZ, Scheraga HA. Entropy changes accompanying association reactions of proteins. J Biol Chem 1963;238:172–181.
25. Stivers JT, Abeygunawardana C, Mildvan AS, Whitman CP. ^{15}N NMR relaxation studies of free and inhibitor-bound 4-oxalocrotonate tautomerase: Backbone dynamics and entropy changes of an enzyme upon inhibitor binding. Biochemistry 1996;35:16036–16047.
26. Hodsdon ME, Cistola DP. Ligand binding alters the backbone mobility of intestinal fatty acid-binding protein as monitored by 15N NMR relaxation and 1H exchange. Biochemistry 1997;36:2278–2290.
27. Lee AL, Kinnear SA, Wand AJ. Redistribution and loss of side chain entropy upon formation of a calmodulin-peptide complex. Nat Struct Biol 2000;7:72–77.
28. Canino LS, Shen T, McCammon JA. Changes in flexibility upon binding: application of the self-consistent pair contact probability method to protein–protein interactions. J Chem Phys 2002;117:9927–9933.
29. Tsai C-J, Kumar S, Ma B, Nussinov R. Folding funnels, binding funnels, and protein function. Protein Sci 1999;8:1181–1190.
30. Tsai C-J, Ma B, Nussinov R. Folding and binding cascades: Shifts in energy landscapes. Proc Natl Acad Sci USA 1999;96:9970–9972.
31. Nichols WL, Rose GD, Ten Eyck LF, Zimm BH. Rigid domains in proteins: An algorithmic approach to their identification. Proteins 1995;23:38–48.
32. Siddiqui AS, Barton GJ. Continuous and discontinuous domains: an algorithm for the automatic generation of reliable protein domain definitions. Protein Sci 1995;4:872–884.
33. Boutonnet NS, Rooman MJ, Wodak SJ. Automatic analysis of protein conformational changes by multiple linkage clustering. J. Mol. Biol. 1995;253:633–647.
34. Krebs WG, Gerstein M. The morph server: a standardized system for analyzing and visualizing macromolecular motions in a database framework. Nucleic Acids Research 2000;28:1665–1675.
35. Shatsky M, Nussinov R, Wolfson HJ. Flexible protein alignment and hinge detection. Proteins 2002;48:242–256.
36. Holm L, Sander C. Parser for protein folding units. Proteins 1994;19:256–268.
37. Zehfus MH, Rose GD. Compact units in proteins. Biochemistry 1986;25:5759–5765.
38. Karplus PA, Schulz GE. Prediction of chain flexibility in proteins. Naturwissenschaften 1985;72:212–213.
39. Maiorov V, Abagyan R. A new method for modeling large-scale rearrangements of protein domains. Proteins 1997;27:410–424.
40. Case DA. Molecular dynamics and normal mode analysis of biomolecular rigidity. In: Thorpe MF, Duxbury PM, eds. Rigidity Theory and Applications. New York: Kluwer Academic/Plenum Publishers; 1999. pp 329–344.
41. Bahar I, Atilgan AR, Erman B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. Fold Des 1997;2:173–181.
42. Ming D, Kong Y, Wu Y, Ma J. Substructure synthesis method for simulating large molecular complexes. Proceedings of the National Academy of Sciences of the United States of America 2003;100:104–109.
43. Ming D, Kong Y, Wakil SJ, Brink J, Ma J. Domain movements in human fatty acid synthase by quantized elastic deformational model. Proc Nat Acad Sci USA 2002;99:7895–7899.
44. Tama F, Wriggers W, Brooks CL, 3rd. Exploring global distortions of biological macromolecules and assemblies from low-resolution structural information and elastic network theory. J Mol Biol 2002;321:297–305.
45. Halle B. Flexibility and packing in proteins. Proc Natl Acad Sci USA 2002;99:1274–1279.
46. Graver J, Servatius B, Servatius H. Combinatorial rigidity (Graduate studies in mathematics). Providence, RI: American Mathematical Society; 1993.
47. Maxwell JC. On the calculation of the equilibrium and stiffness of frames. Philos Mag 1864;27:294–299.
48. Thorpe MF, Jacobs DJ, Chubynsky NV, Rader AJ. Generic rigidity of network glasses. In: Thorpe MF, Duxbury PM, eds. Rigidity theory and applications. New York: Kluwer Academic/Plenum Publishers; 1999. pp 239–277.
49. Laman G. On graphs and rigidity of plane skeletal structures. J Eng Math 1970;4:331–340.
50. Tay T-S, Whiteley W. Recent advances in generic rigidity of structures. Struct Topol 1985;9:31–38.
51. Jacobs DJ, Kuhn LA, Thorpe MF. Flexible and rigid regions in proteins. In: Thorpe MF, Duxbury PM, eds. Rigidity theory and applications. New York: Kluwer Academic/Plenum Publishers; 1999.
52. Jacobs DJ, Thorpe MF. Generic rigidity percolation: The pebble game. Phys Rev Lett 1995;75:4051–4054.
53. Jacobs DJ, Hendrickson B. An algorithm for two dimensional rigidity percolation: The pebble game. J Comput Phys 1997;137:346–365.
54. Jacobs DJ, Thorpe MF. Generic rigidity percolation in two dimensions. Phys Rev E 1996;53:3683–3693.
55. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The protein data bank. Nucl Acids Res 2000;28:235–242.
56. Jacobs DJ, Rader AJ, Kuhn LA, Thorpe MF. Protein flexibility predictions using graph theory. Proteins 2001;44:150–165.
57. Hespenheide BM, Rader AJ, Thorpe MF, Kuhn LA. Identifying protein folding cores from the evolution of flexible regions during unfolding. J Mol Graph Model 2002;21:195–207.
58. Thorpe MF, Lei M, Rader AJ, Jacobs DJ, Kuhn LA. Protein flexibility and dynamics using constraint theory. J Mol Graph Model 2001;19:60–69.
59. Wittinghofer A, Waldmann H. Ras—a molecular switch involved in tumor formation. Angew Chem Int Ed Engl 2000;39:4193–4214.
60. Sprang SR. G protein mechanisms: insights from structural analysis. Annu Rev Biochem 1997;66:639–678.
61. Milburn MV, Tong L, deVos AM, Brunger A, Yamaizumi Z, Nishimura S, Kim SH. Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic Ras proteins. Science 1990;247:939–945.
62. Wittinghofer A, Nassar N. How Ras-related proteins talk to their effectors. Trends Biochem Sci 1996;21:488–491.
63. Emerson SD, Madison VS, Palermo RE, Waugh DS, Scheffler JE, Tsao KL, Kiefer SE, Liu SP, Fry DC. Solution structure of the Ras-binding domain of c-Raf-1 and identification of its Ras interaction surface. Biochemistry 1995;34:6911–6918.
64. Pai EF, Kabsch W, Krengel U, Holmes KC, John J, Wittinghofer A. Structure of the guanine-nucleotide-binding domain of the Ha-Ras oncogene product p21 in the triphosphate conformation. Nature 1989;341:209–214.
65. Pai EF, Krengel U, Petsko GA, Goody RS, Kabsch W, Wittinghofer A. Refined crystal structure of the triphosphate conformation of H-Ras p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. EMBO J 1990;9:2351–2359.
66. Nassar N, Horn G, Herrmann C, Scherer A, McCormick F, Wittinghofer A. 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 1995;375:554–560.
67. Nassar N, Horn G, Herrmann C, Block C, Janknecht R, Wittinghofer A. Ras/Rap effector specificity determined by charge reversal. Nat Struct Biol 1996;3:723–729.
68. Chen JM, Monaco R, Manolatos S, Brandt-Rauf PW, Friedman FK, Pincus MR. Molecular dynamics on complexes of Ras-p21 and its inhibitor protein, Rap-1A, bound to the Ras-binding domain of the Raf-p74 protein: Identification of effector domains in the Raf protein. J Protein Chem 1997;16:619–629.
69. Zeng J, Treutlein HR, Simonson T. Molecular dynamics simulations of the Ras:Raf and Rap:Raf complexes. Proteins 1999;35:89–100.
70. Zeng J, Treutlein HR, Simonson T. Conformation of the Ras-

binding domain of Raf studied by molecular dynamics and free energy simulations. Proteins 1998;31:186–200.
71. Zeng J, Fridman M, Maruta H, Treutlein HR, Simonson T. Protein–protein recognition: an experimental and computational study of the R89K mutation in Raf and its effect on Ras binding. Protein Sci 1999;8:50–64.
72. Muegge I, Schweins T, Warshel A. Electrostatic contributions to protein–protein binding affinities: application to Rap/Raf interaction. Proteins 1998;30:407–423.
73. Chen JM, Monaco R, Manolatos S, Brandt-Rauf PW, Friedman FK, Pincus MR. Molecular dynamics on complexes of Ras-p21 and its inhibitor protein, Rap-1A, bound to the Ras-binding domain of the Raf-p74 protein: identification of effector domains in the Raf protein. J Protein Chem 1997;16:619–629.
74. Chen JM, Rijhwani K, Friedman FK, Hyde MJ, Pincus MR. Identification, using molecular dynamics, of an effector domain of the Ras-binding domain of the Raf-p74 protein that is uniquely involved in oncogenic Ras-p21 signaling. J Protein Chem 2000;19: 545–551.
75. Gohlke H, Kiel C, Case DA. Insights into protein–protein binding by free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. J Mol Biol 2003;330:891–913.
76. Case DA, Pearlman DA, Caldwell JW, Cheatham III TE, Wang J, Ross WS, Simmerling CL, Darden TA, Merz KM, Stanton RV, Cheng AL, Vincent JJ, Crowley M, Tsui V, Gohlke H, Radmer RJ, Duan Y, Pitera J, Massova I, Seibel GL, Singh UC, Weiner PK, Kollman PA, Amber F. University of California: San Francisco, 2002.
77. Cornell WD, Cieplak CI, Bayly IR, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 1995;117:5179–5197.
78. Jorgensen WL, Chandrasekhar J, Madura J, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.
79. Terada T, Ito Y, Shirouzu M, Tateno M, Hashimoto K, Kigawa T, Ebisuzaki T, Takio K, Shibata T, Yokoyama S, Smith BO, Laue ED, Cooper JA. Nuclear magnetic resonance and molecular dynamics studies on the interactions of the Ras-binding domain of Raf-1 with wild-type and mutant Ras proteins. J Mol Biol 1999;286:219–232.
80. Leach AR, Klein TE. A molecular dynamics study of the inhibition of chicken dihydrofolate reductase by a phenyl triazine. J Comput Chem 1995;16:1378–1393.
81. Bayly CI, Cieplak P, Cornell WD, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for determining atom-centered charges: The RESP model. J Phys Chem 1993;97:10269–10280.
82. Aqvist J. Modeling ion ligand interactions in solutions and biomolecules. Theochem J Mol Struct 1992;88:135–152.
83. McCammon JA, Harvey SC. Dynamics of proteins and nucleic acids. Cambridge: Cambridge University Press; 1987.
84. Ichiye T, Karplus M. Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. Proteins 1991;11:205–217.
85. Abseher R, Nilges M. Are there non-trivial dynamic cross-correlations in proteins? J Mol Biol 1998;279:911–920.
86. Hunenberger PH, Mark AE, van Gunsteren WF. Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations. J Mol Biol 1995;252:492–503.
87. Karplus M, Ichiye T. Comment on a "fluctuation and cross correlation analysis of protein motions observed in nanosecond molecular dynamics simulations" J Mol Biol 1996;263:120–122.
88. Laskowski RA. PDBsum: summaries and analyses of PDB structures. Nucl Acid Res 2001;29:221–222.
89. Dahiyat BI, Gordon DB, Mayo SL. Automated design of the surface positions of protein helices. Protein Sci 1997;6:1333–1337.
90. Dong X, Tsai C-J, Nussinov R. Hydrogen bonds and salt bridges across protein–protein interfaces. Protein Eng 1997;10:999–1012.
91. Sternberg MJ, Grace DE, Phillips DC. Dynamic information from protein crystallography. An analysis of temperature factors from refinement of the hen egg-white lysozyme structure. J Mol Biol 1979;130:231–252.
92. Simonson T, Brunger AT. Thermodynamics of protein–peptide interactions in the ribonuclease-S system studied by molecular dynamics and free energy calculations. Biochemistry 1992;31: 8661–8674.
93. Garcia AE, Krumhansl JA, Frauenfelder H. Variations on a theme by Debye and Waller: from simple crystals to proteins. Proteins 1997;29:153–160.
94. Milburn MV, Tong L, deVos AM, Brunger A, Yamaizumi Z, Nishimura S, Kim SH. Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic Ras proteins. Science 1990;247:939–945.
95. Bevington PR. Data reduction and error analysis for the physical sciences. New York: McGraw-Hill; 1969.
96. Gohlke H, Case DA. Converging free energy estimates: MM-PB(GB)SA studies on the protein–protein complex Ras-Raf. J Comput Chem 2004;25:238–250.
97. Yilmaz LS, Atilgan AR. Identifying the adaptive mechanism in globular proteins: fluctuations in densely packed regions manipulate flexible parts. J Chem Phys 2000;113:4454–4464.
98. Baysal C, Atilgan AR. Elucidating the structural mechanisms for biological activity of the chemokyne family. Proteins 2001;43:150–160.
99. Freire E. The propagation of binding interactions to remote sites in proteins: Analysis of the binding of the monoclonal antibody D1.3 to lysozyme. Proc Natl Acad Sci USA 1999;96:10118–10122.
100. McCormick F, Wittinghofer A. Interactions between Ras proteins and their effectors. Curr Opin Biotech 1996;7:449–456.
101. Zhang B-H, Guan K-L. Regulation of the Raf kinase by phosphorylation. Exp Lung Res 2001;27:269–295.
102. Heymann B, Grubmuller H. Molecular dynamics force probe simulations of antibody/antigen unbinding: entropic control and nonadditivity of unbinding forces. Biophys J 2001;81:1295–1313.
103. Williams DC, Jr., Benjamin DC, Poljak RJ, Rule GS. Global changes in amide hydrogen exchange rates for a protein antigen in complex with three different antibodies. J Mol Biol 1996;257: 866–876.
104. Bahar I, Atilgan AR, Demirel MC, Erman B. Vibrational dynamics of folded proteins: Significance of slow and fast modes in relation to function and stability. Phys Rev Lett 1998;80:2733–2736.
105. Cooper A, Dryden DT. Allostery without conformational change. A plausible model. Eur Biophys J 1984;11:103–109.
106. Humphrey W, Dalke A, Schulten K. VMD—visual molecular dynamics. J Mol Graph 1996;14:33–38.

APPENDIX

Sensitivity of FIRST Results to the Representation of Non-Covalent Interactions

The inclusion of hydrogen bonds as constraints into the network depends on their strength with respect to a chosen threshold value, E_{cut}. The hydrogen bond strength is thereby evaluated using an approximate energy function [Eqs. (7, 8)]. To test the sensitivity of the outcome of FIRST calculations on the choice of E_{cut}, FIRST analyses were repeated for Ras, Raf, and Ras–Raf with $E_{cut} = -0.1$ and -1.0 kcal mol^{-1}, respectively. Correlation coefficients of the thus computed flexibility indices with the ones obtained with $E_{cut} = -0.6$ kcal mol^{-1} are given in Table AI, together with the average differences in the number of included hydrogen bonds with respect to the networks for $E_{cut} = -0.6$ kcal mol^{-1}. Although these differences amount to approximately 5–10% of the total number of hydrogen bonds in all cases, all correlation coefficients are larger than 0.970 (with the exception of Ras–Raf in the case of $E_{cut} = -1.0$ kcal mol^{-1}), indicating that the outcome of FIRST is robust with respect to the choice of E_{cut}. Along these lines, inclusion of one water molecule in the binding interface of Ras–Raf, which has been identified as "structural water" by MD simulation,[75] yields negligible changes in the flexibility index for Ras–Raf C_α atoms, as indicated by a correlation coefficient of 0.999.

TABLE AI. Correlation of Flexibility Indices†

FIRST analyses	Raf[a]	Ras[a]	Ras-Raf
$E_{cut} = -0.1$ kcal mol^{-1}, HP included[b]	0.979 (9.3)	0.977 (29.6)	0.975 (40.0)
$E_{cut} = -1.0$ kcal mol^{-1}, HP included[b]	0.982 (−6.5)	0.985 (−17.6)	0.955 (−26.4)
$E_{cut} = -0.6$ kcal mol^{-1}, HP included, structural water included[b]	—	—	0.999 (1.3)
$E_{cut} = -0.1$ kcal mol^{-1}, no HP[b]	0.888 (−31.1)	0.916 (−69.7)	0.912 (−95.5)

†The flexibility indices were determined for 500 snapshots extracted from MD using Eq. (6). Pearson's correlation coefficients are given for comparison of flexibility indices calculated for modified bond networks to the one obtained with standard parameters ($E_{cut} = -0.6$ kcal mol^{-1}; including hydrophobic interactions; no water in the binding interface). The values in brackets are the average differences in the number of non-covalent interactions in the molecules with respect to the standard calculation.
[a]Unbound molecules.
[b]HP: hydrophobic interactions.

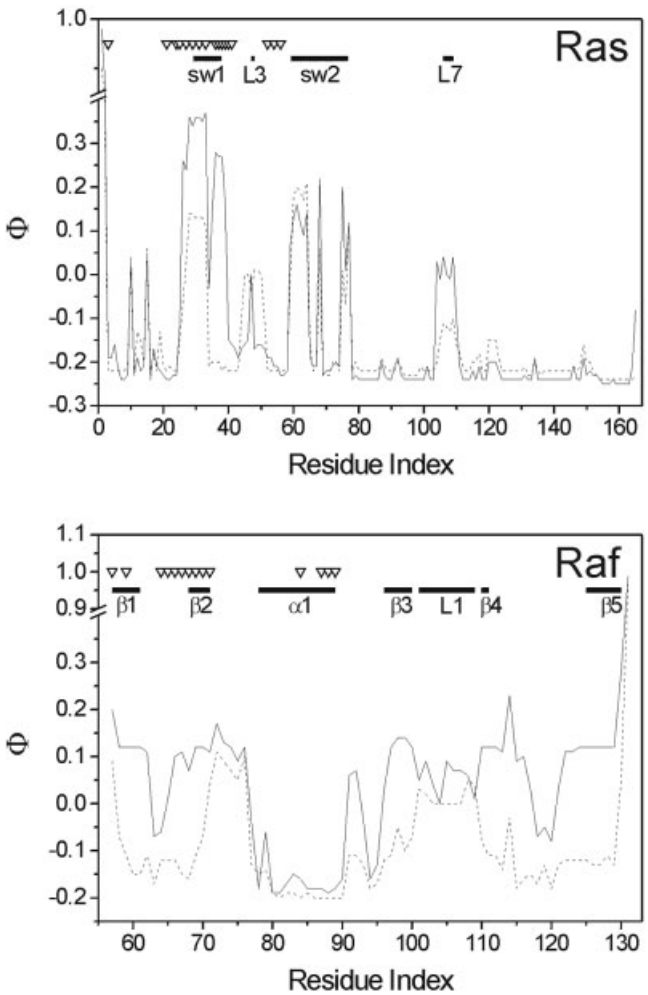


Fig. A1. Average flexibility index of Ras (top) and Raf (bottom) C_α atoms as determined by Eq. (6) (straight line: unbound Ras or Raf; dashed line: bound Ras or Raf) omitting hydrophobic interactions and using a threshold value $E_{cut} = -0.1$ kcal mol^{-1}. For further information see Figure 10.

In early studies with FIRST,[51,56,57] hydrophobic interactions were not included as constraints into the network. Since the number of hydrophobic interactions generally amounts to about half of the number of hydrogen bonds, they may influence the stability of the binding partners considerably. To test the effect of omitting these interactions from the analyses of flexibility in our case, FIRST calculations were repeated for the unbound proteins and the Ras–Raf complex only using hydrogen bonds as non-covalent interactions. Since in this case a threshold value of $E_{cut} = -0.1$ kcal mol^{-1} has been found to provide maximum commonality in the hydrogen bonds for multiple, independently determined structures,[56] this value was also applied here. The correlation coefficients for the comparison of the obtained Φ values to the ones computed with inclusion of hydrophobic interactions (and using $E_{cut} = -0.6$ kcal mol^{-1}) indicate differences in particular in the case of Raf (Table AI). Some characteristic changes in the flexibility indices upon complex formation as found by FIRST with inclusion of hydrophobic interactions (Fig. 10), which agree with results from the MD simulations, are now no longer observed (Fig. A1). As such, the increase in the flexibility of the loop L1 region of Raf is not revealed anymore, and the increase in the flexibility of part of the switch 2 region of Ras becomes less pronounced. The differences thus indicate that hydrophobic interactions (although modeled simply) may be critical for a complete description of the flexibility characteristics of the proteins investigated, and they are known to be critical for the prediction of folding determinants.[57] The fact that the increase in the rigidity of the β-sheet of Raf (as manifested by the more strongly correlated motions in the Ras-Raf complex; see above) is equally well represented with and without hydrophobic interactions (yet, in the latter case, substituted by an increased number of hydrogen bonds due to $E_{cut} = -0.1$ kcal mol^{-1}) indicates the interplay between modeling hydrogen bonds and hydrophobic interactions. Finding the appropriate balance between these interactions is thus crucial for an accurate representation of the flexibility characteristics of proteins by FIRST. This balance is apparently achieved by including all hydrophobic tethers and all hydrogen bonds with a strength of at least -0.6 kcal mol^{-1}, as used for the results described in the text.