Structure Determination and Dynamics of Peptides Overlapping the Catalytic Hairpin of the Ras-Specific GEF Cdc25^{Mm}†

Roberto Consonni,*,‡ Ivana Arosio,‡ Teresa Recca,‡ Renato Longhi,§ Giorgio Colombo,§ and Marco Vanoni||

Istituto per lo Studio delle Macromolecole, Lab NMR, CNR, v. Ampere 56, I-20131 Milan, Italy, Istituto di Chimica del Riconoscimento Molecolare, CNR, via M. Bianco 9, 20131 Milan, Italy, and Dipartimento di Biotecnologie e Bioscienze, Università Milano-Bicocca, Piazza della Scienza 2, I-20126, Milan, Italy

Received March 12, 2003; Revised Manuscript Received July 7, 2003

ABSTRACT: Ras proteins are small G proteins playing a major role in eukaryotic signal transduction. Guanine nucleotide exchange factors (GEF) stimulate GDP/GTP exchange, resulting in the formation of the active Ras-GTP complex. In mammalian cells, two major Ras-specific GEF exist: Sos-like and Cdc25-like. To date, structural data are available only for Cdc25^{Mm}. We designed and synthesized Cdc25^{Mm}-derived peptides spanning residues corresponding to the hSos1 HI helical hairpin that has been implicated in the GEF catalytic mechanism. NMR experiments on a chemically synthesized Cdc25$^{Mm}{}_{1178-1222}$ peptide proved that helix I readily reaches a conformation very similar to the corresponding helix in hSos1, while residues corresponding to helix H in hSos1 show higher conformational flexibility. Molecular dynamics studies with the appropriate solvent model showed that different conformational spaces are available for the peptide. Since helix H is making several contacts with Ras and a Cdc25$^{Mm}{}_{1178-1222}$ peptide is able to bind nucleotide-free Ras in a BIAcore assay, the peptide must be able to obtain the proper Ras-interacting conformation, at least transiently. These results indicate that rational design and improvement of the Ras-interacting peptides should take into account conformational and flexibility features to obtain molecules with the appropriate biochemical properties.

Ras proteins (*1*) are intracellular switches whose activation state (i.e., their binding to GDP and GTP, off and on state, respectively) couples extracellular stimuli to cellular response machinery (reviewed in refs *2*−*5*). Ras proteins are endowed with intrinsic GTPase and guanine nucleotide exchange activities. These basal low activities are stimulated by GTPase activating proteins (GAP)[1] and guanine nucleotide exchange factors (GEF), respectively. Altering this fine balance by the deregulation of either GAP or GEF activity may result in the hypo- or hyperactivation of downstream pathways, so that, for instance, overexpression of a GEF or inactivation of a GAP may result in severe phenotypic consequences (reviewed in ref *6*).

In mammalian cells, two major Ras-specific GEF classes have been identified: Sos-like, such as hSos1, and CDC25-like, such as Cdc25^{Mm} (also called Ras-Grf1) or its rat and human homologue. The two classes differ in their structural organization and tissue distribution and are involved in different signal transduction pathways (reviewed in ref *7*). The crystal structure (*8*) of human Ras complexed with the Ras guanine-nucleotide-exchange-factor region of a human Sos-like GEF (hSos1-p21^{Ras} complex, PDB entry 1BKD) has recently been solved and has shown that Sos interacts with Ras through the so-called catalytic domain, consisting of a series of helical hairpins packed against each other, formed by helices αA−αK. It is currently accepted that the insertion of the helical hairpin HI of hSos1 into the so-called switch 1 of Ras induces a dramatic conformational change of this region, resulting in opening the Ras nucleotide binding site and inducing conformational changes in the switch 2 region. The hairpin HI of hSos1 (Figure 1) seems thus to play a major role in the GEF catalytic mechanism, although other regions may also be important for GEF function, as suggested by structural and mutational data (*8*, *9*). More recently, the possibility that a second GTP-bound Ras molecule may act as an allosteric effector to regulate Sos activity has been reported (*10*).

CD and NMR can give a reasonably accurate picture of the conformational and dynamical behavior of small peptides in solution, providing information on their preferential secondary structure characteristics and their relative stability within a range of related sequences. CD, in particular, and NMR can provide information in terms of average properties of the systems under study. The atomic details determining conformational transitions leading to the instability of a particular secondary structure can be obtained via the application of molecular dynamics (MD) techniques.

MD simulations proved to be able to characterize both the range of alternative states that can be sampled by a

† This work was supported by grants from the Italian National Research Council (Target project on Biotechnology, to R.C. and M.V.) and the Scientific Foundation "Antonio De Marco".

* Corresponding author. E-mail: roberto.consonni@ismac.cnr.it.
‡ Istituto per lo Studio delle Macromolecole.
§ Istituto di Chimica del Riconoscimento Molecolare.
|| Università Milano-Bicocca.

[1] Abbreviations: p21^{ras}, protein encoded by the human *ras (proto)-oncogene*; GEF, guanine nucleotide exchange factor; GAP, GTPase activating protein; GST, glutathione *S*-transferase; CD, circular dichroism; GE-NOESY, gradient enhanced nuclear Overhauser spectrosopy; TOCSY, total correlation spectroscopy; CVFF, consistent valence force field.

10.1021/bi0344026 CCC: $25.00
Published on Web 09/30/2003

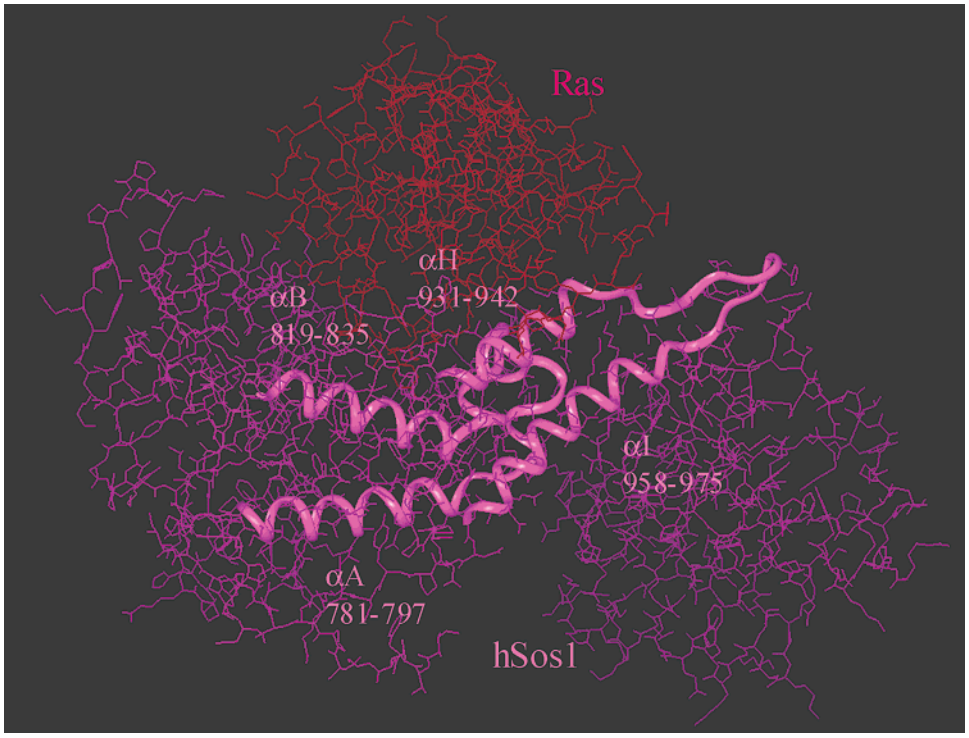

FIGURE 1: Crystal structure (PDB entry 1BKD) of the hSos1-p21^{Ras} complex. HI and AB hairpins are highlighted with hSos1 numbering.

peptide under specific conditions and the dynamics of the processes of conformational variation by sampling relevant (low energy) regions of conformational space. In cases that are characterized by a limited number of low energy conformations in rapid equilibrium (on the MD time scale), MD simulations have been demonstrated to provide highly detailed information on the nature of alternative states in solution and the dynamics of spontaneous folding/unfolding (*11*−*13*). In particular, Daura et al. have been able to demonstrate the reversible folding of a short helix forming a β-heptapeptide in methanol from an arbitrary starting structure under a variety of conditions within 50 ns (*14*−*16*). These studies clearly showed how the study of small peptides could help better understand the nature of the unfolded state and the mechanism of protein folding. MD simulations with the appropriate solvent model for TFE were also used to rationalize the effect of cosolvents on the formation and stabilization of secondary structure motifs in designed peptides (*17*). A variety of other computational studies on helix-forming peptides have also appeared, and factors affecting helix formation and stability, at least in general terms, are well-characterized (*18*−*20*).

The catalytic domains of hSos1 and Cdc25^{Mm} display extensive primary structure homology, and most likely the general fold is conserved. Thus, in the absence of the complete structure of the Cdc25^{Mm} protein, the structural elucidation of the GEF regions involved in Ras recognition and interaction becomes challenging and very important. We designed and synthesized wild type and mutant peptides overlapping the hypothetical HI hairpin sequence of Cdc25^{Mm} (one of the major regions of interaction with Ras, according to the structure of the p21^{Ras}/hSos1 complex), to obtain structural information by means of spectroscopic methods and computational approaches. The possible interference of the N-terminus positive charge in the noncapped peptides has been taken into account by also analyzing the N,C-capped peptides (N$^{\alpha}$-acetyl, C-carboxyamidated peptides), and no significant differences between the two peptide types have been detected. NMR experiments on a chemically synthesized Cdc25$^{Mm}{}_{1178-1222}$ peptide proved that helix I readily reaches a conformation very similar to the corresponding helix in hSos1, while residues corresponding to helix H in hSos1 showed higher conformational flexibility. Molecular dynamics studies with the all-atom TFE solvent model showed that different conformational spaces are available for the peptide. Since helix H made several contacts with Ras, and a Cdc25$^{Mm}{}_{1178-1222}$ peptide is able to bind nucleotide-free Ras in a BIAcore assay, the peptide must be able to obtain the proper Ras-interacting conformation, at least transiently. These results indicated that the rational design and improvement of Ras-interacting peptides should take into account conformational and flexibility features to obtain molecules with the appropriate biochemical properties.

MATERIALS AND METHODS

Design and Synthesis of Peptides. In Figure 1, the AB and HI helical hairpins of hSos1 interacting with the Ras molecule are highlighted. Peptides overlapping the hairpin HI of the hSos1 sequence were designed. A 38mer peptide, in which the N-terminal helical part was truncated, (Cdc25$^{Mm}{}_{1185-1222}$) and three 45mer peptides (Cdc25$^{Mm}{}_{1178-1222}$, Cdc25$^{Mm}{}_{1178-1222}{}^{T1184A}$, and Cdc25$^{Mm}{}_{1178-1222}{}^{T1184E}$), were prepared by solid-phase synthesis, to perform spectroscopic studies using CD and NMR techniques. The 45mer mutant peptides were synthesized because it has been reported that the T1184E mutation turns the CDC25^{Mm} molecule into a dominant negative protein that can trap the Ras molecule in the nucleotide-free state, thus effectively down-regulating Ras both in vivo and in vitro (*21*). A T935A mutation in the corresponding position of hSos1, on the contrary, did not appear to affect hSos1 binding to Ras or the GEF activity of the mutant (*9*).

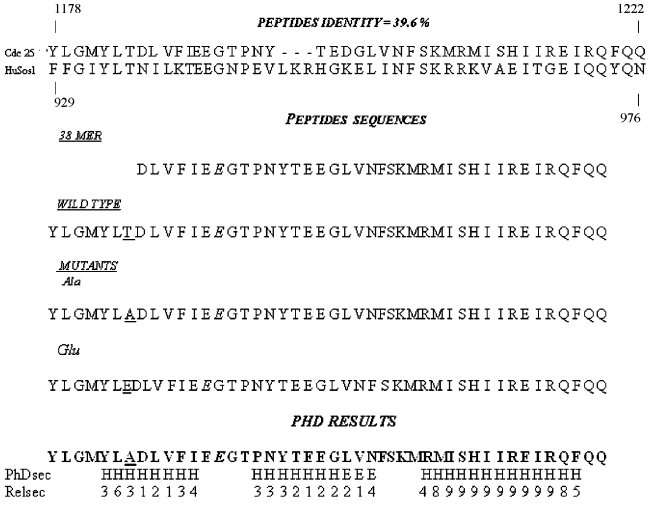

FIGURE 2: Sequence alignments for hSos1 and CDC25^{Mm} proteins according to the LALIGN program (ref *24*), amino acid sequences of the synthetic peptides, and secondary structure prediction according to the PHD prediction algorithm (ref *25*).

Sequence alignment performed at an ISREC server (www.ch.embnet.org) gave 39.6% identity (*22*) between the hSos1 and the Cdc25^{Mm} HI hairpins, strongly suggesting that this region should share the same overall fold. The sequence of the Cdc25^{Mm}-derived 45- and 38mer peptides are represented in Figure 2, together with the sequence alignment of helices HI of hSos1 and CDC25^{Mm}, spanning residues 1178−1222 (CDC25^{Mm} numbering, 929−976 hSos1 numbering), and finally, the secondary structure prediction result (PHD prediction algorithm at EMBL) (*23*) is represented in the same figure.

Residues 1180−1191 overlap helix H, and residues 1208−1220 overlap helix I, the latter being only partially present in the 38mer. It is interesting to note that the reliability of the predicted helices for the CDC25^{Mm} sequence is very high only for helix I, while the opposite holds for helix H. Poor reliability of helical propensity was also found for hSos1 helix H (data not shown), although the helix is present in the hSos1/Ras complex.

The three Cdc25$^{Mm}{}_{1178-1222}$ 45mer peptides (wild type and T1184E and T1184A mutants) and the truncated 38mer Cdc25$^{Mm}{}_{1185-1222}$ peptide were synthesized by solid-phase methods by using an automated Applied Biosystem Model 433A peptide synthesizer by using N^{α}-[9-fluorenil (methoxycarbonyl)]-(Fmoc) protected amino acids and *N*-[(dimethylamino)-1*H*-1,2,3-triazole-[4,5-*b*]pyridine-1-ylmethylene]-*N*-methylmethanaminium hexafluorophosphate *N*-oxide (HATU, PE Biosystems, Inc., Warrington, UK). The peptide chain was assembled stepwise on 2-chlorotritil resin. Because of the known problem of the possible intermolecular cyclization involving residues Asp-Gly and/or Asp-Ser, Glu conservatively substituted the Asp residue proceeding Gly. Cleavage from the resin was performed with a mixture of K including 2% triisopropylsilane (*24*). The peptides were purified to near homogeneity (purity > 96%) using semipreparative reversed-phase high-performance liquid chromatography (RP-HPLC) on a Vydac C4 column (19 × 250 mm) using an acetonitrile/water gradient containing 0.1% trifluoroacetic acid, and the peptides were recovered in 30−40% yield. The mass of the peptides was measured by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry on a Voyager-RP Biospectrometry Workstation (PerSpective Biosystem, Inc.). The calculated and experimental molecular weights (MH^+) were coincident.

BIAcore Analysis. The wild type fusion protein between glutathione *S*-transferase (GST) and the catalytic domain of CDC25^{Mm} (CDC25$^{Mm}{}_{976-1262}$) was purified by glutathione-sepharose chromatography (Pharmacia) as previously described (*25*, *26*). Recombinant N-terminal His-tagged p21^{Ras} protein (GDP-bound) was purified from an *Escherichia coli* strain harboring a pQE(Qiagen)-derived plasmid according to the manufacturer's suggestions and directly charged to an activated CMS sensor chip research grade as described by the producer (BIAcore AB). Peptides or the control CDC25$^{Mm}{}_{976-1262}$ catalytic domain (as a GST fusion protein) were diluted immediately before use in buffer C (10 mM Na HEPES, pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 0.005% Tween 20). The peptide buffer always contained a final concentration of 0.3% TFE. Hereafter, samples were injected, and their binding to Ras was measured as a change in the resonance signal in time. Subsequently, the dissociation of the complex was studied by passing buffer C over the lane. This way, the binding and dissociation kinetics of the peptide to and from Ras was measured. At the end of the measurement, the bound analyte was dissociated by passing 5 mM NaOH (5 μL, 10 μL/min) followed by 5 mM HCl (5 μL, 10 μL/min) and buffer C (10 μL/min) until a steady baseline was obtained. The association (k_{on}) and dissociation (k_{off}) rate constants were calculated using the BIAlogue 3.1 software (BIAcore BA).

Circular Dichroism Analysis. CD spectra were collected using a Jasco J600 spectropolarimeter. Spectra were run at 22 °C at polypeptide concentration of 0.1 and 2 mg/mL with a cuvette path length of 0.1 and 0.005 cm, respectively. All spectra were baseline corrected by subtracting the buffer spectra. The molar ellipticity was calculated according to

$$\theta_\lambda = \frac{[\theta]_\lambda 100}{[(N-l)lc]}$$

where $[\theta]_\lambda$ is the measured ellipticity (in degrees) at a wavelength λ, N is the number of residues, l is the cuvette path length (in cm), and c is the protein concentration (in mol/L). The recorded spectra showed no dependence from peptide concentrations.

NMR Structure Determination. NMR samples were prepared dissolving 9 mg of peptide in pure TFE at pH 3.5. ^{1}H multidimensional NMR spectra were recorded with a Varian 600 and Bruker DMX 500 spectrometer equipped with a z-gradient coil with a proton frequency of 599.90 and 500.13 MHz, respectively. Bidimensional spectra (DQF-COSY, GE-TOCSY, and GE-NOESY) were recorded with standard pulse sequences, in the phase sensitive mode, with quadrature detection in both dimensions by using a time proportional phase method and pulse field gradients for solvent suppression (*27*, *28*). All spectra were processed on Octane (Silicon Graphics) by using Felix_ND and Discover modules from the Molecular Simulation software. Spin system identification and sequential resonance assignments of ^{1}H resonances were carried out in homonuclear 2-D spectra using standard methodologies (*29*).

The volumes of all assigned NOE cross-peaks from the NOESY spectra were measured using the FELIX integration

routine. NOE intensities were converted to distances as described previously (*30*). The β-geminal proton resonances of His1212 were used to calibrate the conversion of volumes into distances. Distance constraints were divided into three classes (strong, medium, and weak) and converted into the corresponding upper-limit distance constraints of 2.7, 3.5, and 5 Å, respectively. Pseudoatom distance corrections were performed for nonstereospecifically assigned methylene protons, aromatic ring protons, and methyl groups (*29*).

3-D structures were generated using a simulated annealing method implemented within the InsightII program. The standard Nilges et al. (*31*) protocol was followed but with some modifications. The CVFF force field and extended peptide conformation as a starting structure were used. The maximum force of the distance constraint potential was set to 15 kcal mol^{-1} Å^{-2}, while the upper and lower bound force constant was set to 1 kcal mol^{-1} Å^{-2}. Simulations were carried out in five consecutive stages: (i) 100 steps of Powell energy minimization to remove bad nonbonded contacts; (ii) 15 ps of dynamics at 1000 K with normal van der Waals radii and the repulsive force constant fixed at 0.001 kcal mol^{-1} Å^{-4}; (iii) 10 ps of dynamics at 1000 K with the repulsive force constant incremented at 0.01 kcal mol^{-1} Å^{-4}; (iv) cooling to 300 K during 13 ps (13 steps of 1 ps with 8.3 K cooling/step) with the repulsive force constant of 0.25 kcal mol^{-1} Å^{-4} and van der Waals radii scaled by 0.85; and (v) energy minimization with 1200 steps of the steepest descent algorithm followed by 1200 steps of the conjugate gradient.

A family of structures was calculated after the assignment was completed, and some structures were selected on the basis of either large structural deviations from the mean coordinates or the number of violations and excluded from further refinement. The last simulation contained 50 structures out of which low energy structures have been selected: by averaging superimposed coordinates of these selected structures, the final RMSD was calculated.

Molecular Dynamics Simulations in Explicit TFE. Three different simulations of the best NMR-characterized Cdc25$^{Mm}{}_{1178-1222}{}^{T1184A}$ (45mer) and Cdc25$^{Mm}{}_{1185-1222}$ (38mer) peptides were performed in the presence of explicit organic solvent molecules by using a modified force field consisting of new parameters. This approach gave the possibility of reproducing the thermodynamic properties of TFE (*32*). Two MD simulations of Cdc25$^{Mm}{}_{1178-1222}{}^{T1184A}$, labeled 45A and 45B, used two different starting structures, while the simulation of Cdc25$^{Mm}{}_{1185-1222}$ started from an all helical conformation.

The starting structures for simulations 45A and 38A in explicit TFE were the averaged NMR structures obtained after the NMR refinement, while 45B was started from the structure obtained after 2 ns of simulation 45A, imposing a new set of velocities obtained from a Maxwellian distribution at the desired initial temperature. The peptides were protonated to give a zwitterionic form (with a charged N- and C-terminal) in line with the experimental conditions at which the peptide was studied, and the total charge on the peptides was +3 with no counterions added. The peptides were solvated with TFE in a periodic truncated octahedron large enough to contain the peptide and 0.8 nm of solvent on all sides. All solvent molecules within 0.15 nm of any peptide atom were removed. The resulting system was composed of 496 peptide atoms and 2918 TFE molecules in the case of the 45mer, while it was composed of 422 peptide atoms and 1745 TFE molecules in the case of the 38mer. The system was subsequently energy minimized with a steepest descent method for 100 steps. To compare the dynamical behavior of the peptides, three different simulations of 30 ns each were performed. In all simulations, the temperature was maintained close to the intended values by weak coupling to an external temperature bath (*33*) with a coupling constant of 0.1 ps. The peptide and the rest of the system were coupled separately to the temperature bath. The GROMOS96 force field (*34*, *35*), augmented with the new TFE parameters by Fioroni et al. (*32*), was used. The LINCS algorithm (*36*) was used to constrain all bond lengths. To extend the time range that could be simulated, dummy atoms were used to model the polar hydrogen atoms, permitting a time step of 4 fs.

A twin-range cutoff was used for the calculation of the nonbonded interactions. The short-range cutoff radius was set to 0.8 nm and the long-range cutoff radius to 1.4 nm for both Columbic and Lennard−Jones interactions. The cutoff values are the same as those used for the GROMOS96 force field parametrization (*34*). Interactions within the short-range cutoff were updated every time step, whereas interactions within the long-range cutoff were updated every five time steps together with the pairlist. All atoms were given an initial velocity obtained from a Maxwellian distribution at the desired initial temperature. The density of the system was adjusted, performing the first equilibration runs at NPT (constant number of particles, pressure, and temperature) conditions by weak coupling to a bath of constant pressure $P_0 = 1$ bar and coupling time $\tau_P = 0.5$ ps (*33*). All the simulations were equilibrated by 50 ps of MD runs with position restraints on the peptide to allow relaxation of the solvent molecules. Other 50 ps runs followed these first equilibration runs without position restraints on the peptide. The production runs using NVT (constant number of particles, volume, and temperature) conditions, after equilibration, were 30 ns long. All the MD runs and the analysis of the trajectories were performed using the GROMACS software package.

RESULTS AND DISCUSSION

Chemically Synthesized HI Hairpin Binds to Ras. Plasmon surface resonance was used to ascertain if the HI hairpin was able to bind to Ras. We therefore coupled Ras directly to the sensor chip, after which the nucleotide was released from Ras in the presence of a high excess of EDTA. The binding and release of the HI hairpin and GST-CDC25$^{Mm}{}_{976-1262}$ fusion protein to and from this chip was followed. Since it is more soluble and less prone to aggregation problems that might hamper quantitative determinations, the peptide Cdc25$^{Mm}{}_{1178-1222}{}^{T1184A}$ was used in quantitative BIAcore binding experiments. The calculated association rate constant k_{on} for the HI hairpin and the GST-CDC25$^{Mm}{}_{976-1262}$ fusion protein are reported in Table 1. Kinetic and thermodynamic constants for the GST-CDC25$^{Mm}{}_{976-1262}$ fusion protein agree with those previously reported for the unfused CDC25$^{Mm}{}_{976-1262}$ protein using a different strategy of binding to the chip (*37*, *38*). The binding of the HI hairpin to the Ras protein indicates that the peptide (or at least a fraction of it) is able to attain a conformation compatible with the interaction with Ras. Notably, while the

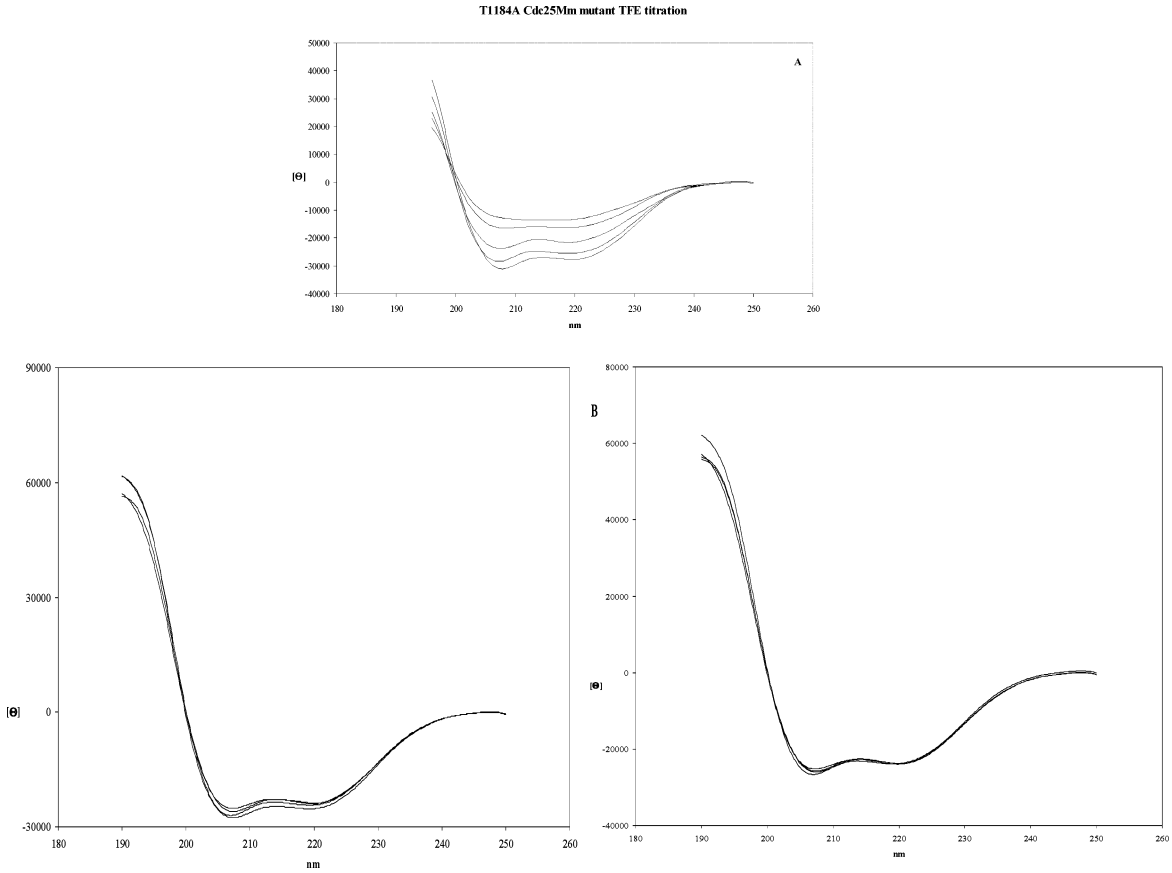

FIGURE 3: (A) Cdc25$^{Mm}_{1178-1222}{}^{T1184A}$ mutant at different TFE concentrations: from the bottom (deeper minima at 208 and 222 nm) to the top curve, the TFE concentration is 98, 50, 20, 10, and 2%. The maximum values around 196 nm, from the top to the bottom, correspond to 98, 50, 20, 10, and 2% TFE concentration. All the other peptides showed the same profiles and have been omitted for clarity. (B) CD spectra of Cdc25$^{Mm}_{1178-1222}$ wt, 38mer, Cdc25$^{Mm}_{1178-1222}{}^{T1184A}$, and Cdc25$^{Mm}_{1178-1222}{}^{T1184E}$ mutants at 0.1 mg/mL 50% TFE concentration and 2 mg/mL 50% TFE concentration (left and right, respectively).

Table 1: Equilibrium and Rate Constants Derived from BIAcore Experiments for GST-CDC25$^{Mm}_{976-1262}$ and the HI Catalytic Hairpin

sample	k_{on} (10^4 M^{-1} s^{-1})	k_{off} (10^{-5} sec^{-1})	K_D (nM)
Cdc25$^{Mm}_{1178-1222}{}^{T1184A}$	4.9 ± 0.3	200 ± 20	41
GST-CDC25$^{Mm}_{976-1262}$	4.0 ± 3.0	4.7 ± 0.3	1.2

association constants of the Cdc25$^{Mm}_{1178-1222}{}^{T1184A}$ peptide and of the intact catalytic domain agree within experimental error, the peptide dissociated from Ras with much faster kinetics (a near 40-fold difference). This result is not unexpected because only a small fraction of the Ras/Gef contacts involves the HI hairpin and is more fully discussed in the Conclusion. Results with capped peptides gave results superimposable to those obtained with uncapped peptides.

Structural Analysis of the GEF-Derived Peptides. The structural properties of the HI hairpin were then investigated by structural and computational methods. First, the synthetic peptides were investigated by CD at pH 3.5 using different solvent conditions to define the better peptide conformational features. The Cdc25$^{Mm}_{1178-1222}$ wild type and mutant peptides and the 38mer Cdc25$^{Mm}_{1185-1222}$ peptide were investigated upon the addition of TFE in the range of 2−98%. Progressively deeper minima at 208 and 222 nm were observed by increasing the TFE concentration (Figure 3A).

The corresponding CD profiles were consistent with a high helical content, only in a mixture of TFE/water over 50% TFE. The peptides remained poorly structured and probably aggregated (due to their poor solubility in water) at a low TFE concentration, and moreover, the four peptides analyzed showed a very similar behavior in the TFE titration. Finally, the elipticity of all peptides in 50% TFE resulted independently from the peptide concentration, as shown in Figure 3B.

Conventional bidimensional NMR spectra have been recorded for the spin system and sequential assignments for all peptides. Analysis of the GE-NOESY NMR spectra of the Cdc25$^{Mm}_{1185-1222}$ 38mer peptide and the Cdc25$^{Mm}_{1178-1222}{}^{T1184A}$ mutant in pure TFE revealed the typical medium- and short-range proton contacts for a helix only for residues 1208−1219 of the peptide (corresponding to helix I in the intact GEF catalytic domain, residues 958−975, hSos1 numbering). In particular, the NOE contacts obtained for the Cdc25$^{Mm}_{1178-1222}{}^{T1184A}$ mutant were used for restrained molecular dynamic simulations (RMD), which gave a conformer family consisting of 39 structures with a backbone RMSD of 1.21 ± 0.28 Å for helix I (Figure 4A).

Even in these conditions, residues 1180−1191 of the Cdc25$^{Mm}_{1178-1222}{}^{T1184A}$ peptide, corresponding to helix H in the intact GEF catalytic domain (residues 931−942, hSos1 numbering), remained mostly disordered with a low helical tendency, confirmed by a high backbone RMSD alignment (1.92 ± 0.28 Å) obtained for these residues spanning from residue 1183 up to residue 1190.

The poor chemical shift dispersion present in the spectra of the peptides gave many fully overlapped cross-peaks, and

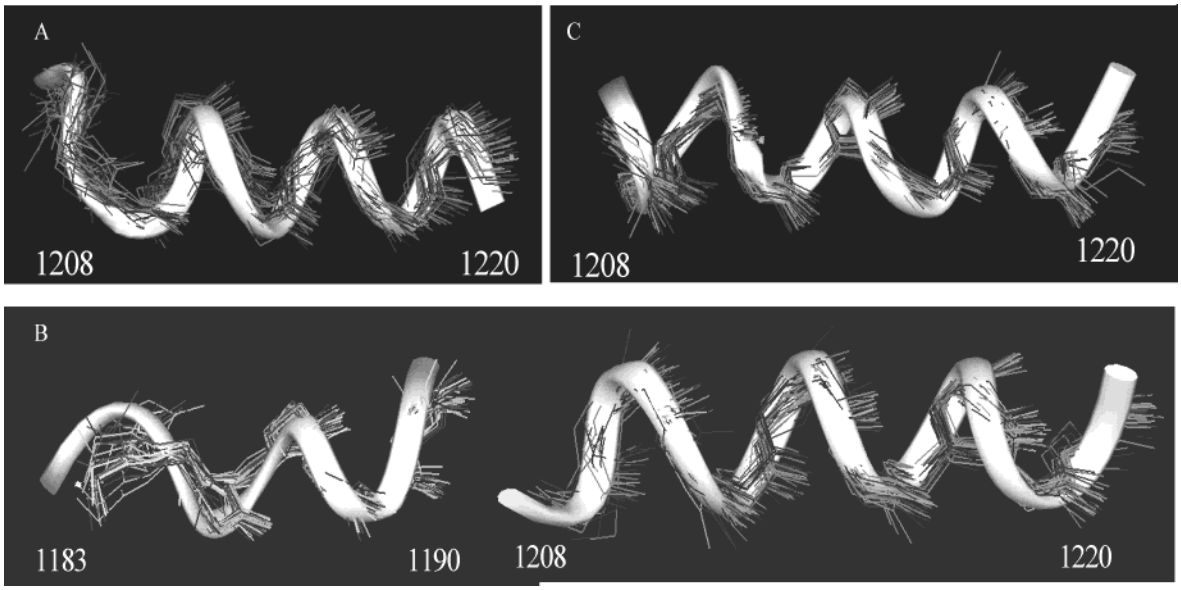

FIGURE 4: (A) Ribbon representation of the 39 overlapped conformers for Cdc25$^{Mm}_{1178-1222}{}^{T1184A}$, obtained with the use of a nonambiguous NOE dataset. Only helical region I (residues 1208−1220) is represented. (B) Ribbon representation of the 35 overlapped conformers for Cdc25$^{Mm}_{1178-1222}{}^{T1184A}$, obtained with the use of an ambiguous NOE dataset. Helical regions I (residues 1208−1220) and H (1183−1190) are represented. (C) Ribbon representation of the 42 overlapped conformers for the 38mer Cdc25$^{Mm}_{1185-1222}$ peptide. Residues 1208−1220 are represented.

Table 2: Backbone RMSD Values for Residues 1208−1220 and NOE Distance Restraints Used[a]

peptide	restraints	sequential	medium	intraresidue	long	RMSD$_{bb}$ H[b]	RMSD$_{bb}$ I[b]
38mer	517	152	120	243	2	0.63 ± 0.13	1.81 ± 0.71
T1184A	271	87	44	140	0	1.21 ± 0.28	1.92 ± 0.65
T1184A′	380	118	101	161	0	0.78 ± 0.16	0.71 ± 0.19

[a] Mutant T1184A′ has been determined by using an ambiguous NOE dataset. [b] RMSDbb indicates deviations for the backbone atoms.

this lead to an ambiguity assignment problem. We decided to use these ambiguous NOE cross-peaks in RMD simulations, to obtain more defined structures: in the case of the Cdc25$^{Mm}_{1178-1222}{}^{T1184A}$ mutant NOESY spectra, RMD yielded a family of 35 best structures displaying helices H and I with a backbone RMSD of 0.71 ± 0.19 and 0.78 ± 0.16 Å, respectively, as shown in Figure 4B.

The Cdc25$^{Mm}_{1185-1222}$ 38mer peptide resulted as the most soluble one and did not present ambiguity in the assignment. The RMD gave a family of 42 conformers with a backbone RMSD of 0.63 ± 0.16 Å as represented in Figure 4C. Also, in this shorter peptide residues 1208−1220 constitute the only helix present. The RMSD comparison between these different conformers' family is summarized in Table 2.

The Cdc25$^{Mm}_{1178-1222}$ and Cdc25$^{Mm}_{1178-1222}{}^{T1184E}$ peptides displayed very low solubility even in pure organic solvent, thus leading to very poor NMR spectra and preventing a good quality structure determination by NMR. The confirmation of the aggregation present in our solution was confirmed by applying the Esposito and Pastore method (*39*), where the ratio between the isolated cross-peak and the corresponding diagonal peak in the NOESY spectrum evaluates the correlation time for the solved molecule. In our peptides, and in particular for Cdc25$^{Mm}_{1178-1222}{}^{T1184A}$, we obtained 6.5 ns, a typical value for a much bigger protein, thus indicating the presence in solution of at least a trimeric peptide aggregate. The conformers' family obtained for these last peptides was very randomized and not well-confined with a small backbone RMSD value but still displayed a high helical tendency only for the putative helix I (data not shown).

MD Analysis of the GEF-Derived Peptides in the Presence of Explicit Organic Solvent. In general, the structural information obtained for these peptides was puzzling since the contacts between the HI hairpin and the Ras also involved helix H. It is worth noting that the PHD algorithm predicts a very poor helical propensity for helix H, which is nevertheless present in the solved Ras/Sos complex (Figure 1).

To reproduce and follow the peptide behavior, MD simulations of the best NMR-characterized Cdc25$^{Mm}_{1178-1222}{}^{T1184A}$ peptide have been performed in the presence of explicit organic solvent molecules by using a modified force field (*32*). The MD simulations of Cdc25$^{Mm}_{1178-1222}{}^{T1184A}$ used two different starting structures. For simulation 45A, the starting structure was that obtained by imposing ambiguous NMR restraints from the NOE data set, where residues 1183−1190 and 1208−1220 were in a helical conformation. For simulation 45B, the starting structure was the one obtained after 2 ns of simulation 45A; a new set of velocities was imposed (see Materials and Methods) to speed up the exploration of the large conformational space available for the peptide.

In analogous conditions, MD simulations proved to be a suitable tool to understand the atomic determinants of peptide folding and stabilization via solvent effects (*17*, *40*). The analysis of the secondary structure content assigned using the DSSP algorithm (*41*) as a function of time for simulation 45A and 45B is reported in Figure 5. The results for simulation 45A (Figure 5, top) clearly showed that the

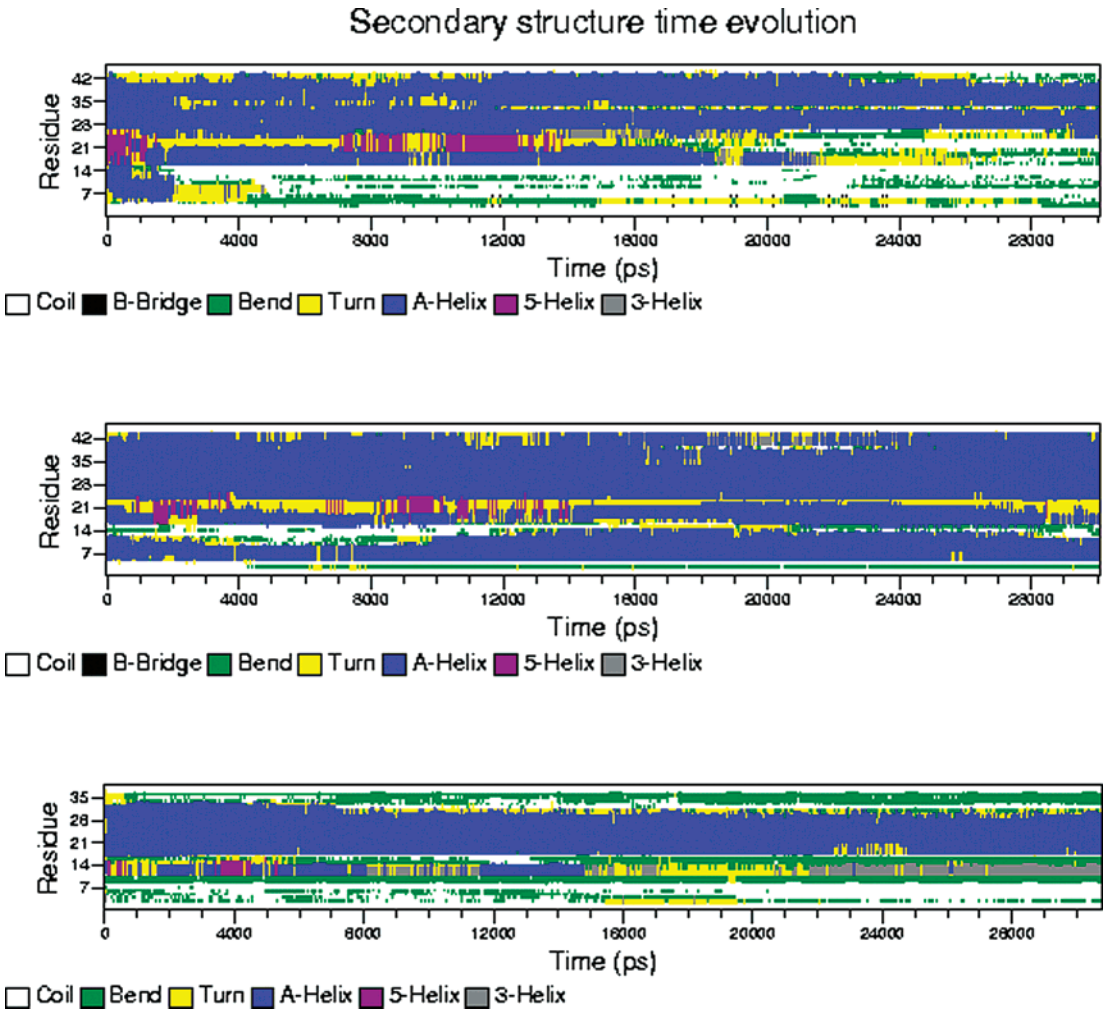

FIGURE 5: Top: T1184A mutant secondary structure time evolution for simulation 45A. Middle: T1184A mutant secondary structure time evolution for simulation 45B. Bottom: 38mer peptide secondary structure time evolution.

helical conformation was maintained over residues Tyr1178 to Pro1194 only for a very limited amount of time, around 2 ns; after this time, a more disordered conformation was assumed. On the other hand, simulation 45B, starting with both different initial conditions and velocity sets, showed that the global helical fold can be maintained for all of the simulation time, without any fraying of the N-terminal part of the H helix (Figure 5, middle). This phenomenon was most likely due to the intrinsic high flexibility of the peptide, which allowed the population of different conformational families.

The results of the MD simulations, where two basically helical starting structures for the T1184A mutant have been used, showed that the N-terminal part of the peptide is extremely flexible and samples a range of conformations in which the helix H is present only for a limited amount of time (first 2 ns) in the case of simulation 45A.

To clarify the conformational behavior in the force field simulation, a clustering analysis was run on the trajectories obtained from both simulations, according to a procedure already followed by our group (*40*). The number of different structure clusters was much higher in the case of simulation 45A (119 vs 60 clusters), indicating that in this case the peptide was sampling a wider range of conformations: this higher conformational flexibility is most likely due to the partial unfolding of the N-terminal region, which, after losing the constraints imposed by the helical geometry, is able to sample a range of disordered conformations. Figure 6 shows a set of representative conformations, extracted from the trajectories, as compared to the central structure of the most populated clusters for both cases. It is evident that the different starting conditions of the two simulations determined the sampling of different regions of conformational space, which were both accessible to the peptide $Cdc25^{Mm}{}_{1178-1222}{}^{T1184A}$ in TFE at 300 K.

MD calculations on the $Cdc25^{Mm}{}_{1185-1222}$ 38mer peptide showed and confirmed a trend that was similar to the one observed previously. Simulation 38, starting from the helical structure obtained imposing ambiguous NOE restraints, showed that the helical content of the N-terminal part was not stable in the simulation conditions, as expected. The representative structures obtained from the clustering algorithm applied to this simulation and representative structures from the simulation are reported in Figures 5 and 6, bottom. The combination of the MD data showed that the helical conformation in the H portion of the isolated HI hairpin can actually be present, but its interconversion to an extended structure could be too fast to be measured on the NMR time scale.

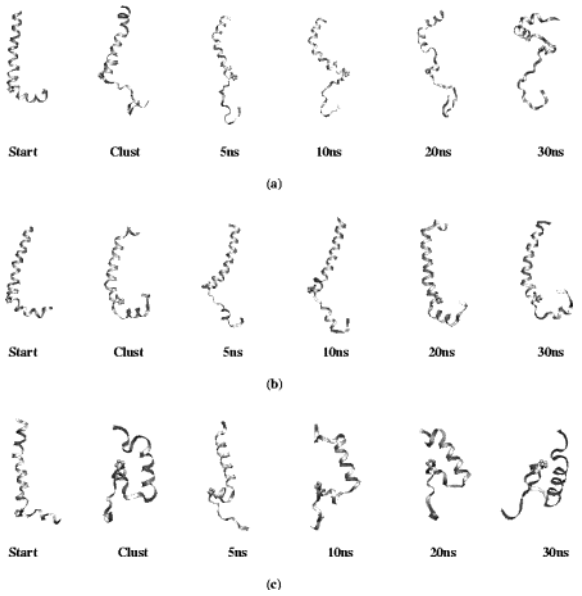

FIGURE 6: Pictorial representation of the 3-D structures obtained for the peptides in the three simulations. From top to bottom: simulation 45A, 45B, and38. Start labels the NMR defined starting structure, Clust the representative structure of the most populated cluster of conformers. The other conformations represent the structural evolution at different time ranges.

CONCLUSIONS

GEFs are the most common link between extracellular stimuli and Ras protein activation. The HI helical hairpin has been implied as the major structural element responsible for Ras activation by the GEF (*8*). We have used a combination of computational, spectroscopic, and biophysical tools to investigate the properties of synthetic peptides derived from the GEF Cdc25^{Mm}, corresponding to the HI hairpin of hSos1. The NMR experiments proved that while helix I readily forms under our experimental conditions, a much higher degree of flexibility is observed in the N-terminal region (i.e., residues 1180−1191, corresponding to helix H in hSos1). The PHD algorithm predicts a low propensity for helix H of both Cdc25^{Mm} and hSos1, despite the fact that helix H makes several contacts with Ras in the published hSos1/Ras complex. In the region corresponding to helix H, in the Cdc25^{Mm}-derived peptide, a lack of typical helix NOE contacts is observed. MD simulations have been run in the presence of an explicit solvent to elucidate the dynamical behavior of the peptides, and the stability of the Ras-interacting helix H, in an environment reproducing NMR experimental conditions. Consistent with NMR data, analysis of the MD trajectories of the analyzed peptides showed that they consist of a stable, although flexible, helical part involving residues 1202−1219 and of a less structured tail spanning the N-terminal part of the peptide. The simulations were still too short with respect to the NMR time scale and to provide access to the spontaneous conformational interconversion between the structures sampled in 45A and those sampled in 45B. However, our simulations showed that the use of a different set of starting conditions can actually give access to different regions of conformational space. Importantly, the partial unfolding starting at the N-terminal domain of Cdc25$^{Mm}{}_{1178-1222}{}^{T1184A}$ confirmed the low confidence value for the predicted helical propensity of the N-terminal region of the peptide and also the uncertainty associated with the ambiguous NOE's restraints used to impose the helical conformation in this part of the molecule. Furthermore, the MD simulations showed a conformational transition for the helix H to an extended structure, which could be considered the starting point for the peptide aggregation, a distinct feature of all the peptides especially at high concentration. The transition to the extended or β-sheetlike structures has been considered by several researchers to be the starting point for peptide aggregation and the fibril formation, a factor in a wide variety of pathological disorders, due to the insolubility of the resulting superstructure (*42*, *43*).

BIAcore data indicate that the HI peptide is able to bind with significant affinity to nucleotide-free Ras. The kinetic association constant is not dissimilar from that observed for the catalytic domain in control experiments, indicating that the HI peptide, despite the high flexibility of helix H, binds Ras almost as fast as the whole catalytic domain. Since the HI hairpin makes contact with Ras almost exclusively through helix H, this finding suggests that residues 1180−1191 may be highly flexible not only in the isolated peptide but, at least to a certain extent, in the complete catalytic domain as well. Such flexibility may be instrumental in allowing the interaction of the GEF molecule with nucleotide-bound Ras. In fact, a GEF would not be able to bind nucleotide-bound Ras, maintaining the structure observed in the nucleotide-free Ras/GEF complex, as can be seen by substituting either GDP-Ras or GTP-Ras for nucleotide-free Ras in the hSos1/nucleotide-free Ras complex (L. De Gioia, personal communication). Once bound to Ras, however, the HI hairpin dissociates much more readily than the complete catalytic domain as indicated by the large difference in k_{off}. NMR and MD results suggest that the H helix does not completely stabilize upon binding, unlike the situation seen in the crystal structure of the complex between the Ras and the Sos catalytic domain. Thus, the difference in k_{off} between the isolated hairpin and the catalytic domain is likely due to both intramolecular stabilization of the H helix in the complete catalytic domain and to the increased interaction area (i.e., the total number of intermolecular bonds) in the catalytic domain as compared to the isolated HI hairpin.

Our studies showed that in designing peptides, the conformational and flexibility features should be taken into account to obtain sequences with the appropriate characteristics for binding. Moreover, the combination of NMR and MD studies with the appropriate solvent model can give important insights into the conformational preferences of the different molecule parts, a very important point for the future rational design and improvement of related peptides. Non-nucleotide inhibitors of the Ras nucleotide exchange process have been already identified (*44*, *45*), and recently, a complex inhibitor/Ras protein has been characterized by NMR spectroscopy (*46*). The definition of structural and functional epitopes (*47*) at the Ras/GEF interface may ultimately lead to the downsizing (*48*) of Ras/GEF interfering peptides until they can be used as lead compounds in the development of drugs able to dock at the Ras/GEF interface, possibly down-regulating Ras-dependent signal transduction.

ACKNOWLEDGMENT

The authors wish to thank Dr. Sonia Fantinato for the kind gift of purified Ras and Cdc25^{Mm}/GST proteins, Elena

Accardo for excellent technical assistance with BIAcore assays, L. De Gioia for sharing unpublished results, and Dr. Marco Tatò for the use of the 600 MHz spectrometer.

REFERENCES

1. Lowy, D. R., and Willumsen, B. M. (1993) *Annu. Rev. Biochem. 62*, 851−891.
2. Campbell, S. L., Khosravi-Far, R., Rossman, K. L., Clark, G. J., and Der, C. J. (1988) *Oncogene 17*, 1395−1413.
3. Bos, J. L. (1989) *Cancer Res. 49*, 4682−4689.
4. Boettner, B., and Van Aelst, L. (2002) *Genes Dev. 16*, 2033−2038.
5. Downward, J. (1998) *Curr. Opin. Genet. Dev. 8*, 49−54.
6. Downward, J. (2003) *Nat. Rev. 3*, 11−22.
7. Quilliam, L. A., Rebhun, J. F., and Castro, A. F. (2002) *Prog. Nucleic Acids Mol. Biol. 71*, 391−444.
8. Boriack-Sjodin, P. A., Margarit, M. S., Bar-Sagi, D., and Kuriyan, J. (1998) *Nature 394*, 337−343.
9. Hall, B. E., Yang, S. S., Boriack-Sjodin, P. A., Kuriyan, J., and Bar-Sagi, D. (2001) *J. Biol. Chem. 276*, 2762−2767.
10. Margarit, S. M., Sondermann, H., Hall, B. E., Nagar, B., Hoelz, A., Pirruccello, M., Bar-Sagi, D., and Kuriyan, J. (2003) *Cell* 112, 685−695.
11. Shakhnovich, E. I. (1997) *Curr. Opin. Struct. Biol. 7*, 29−40.
12. Lazaridis, T., and Karplus, M. (1997) *Science 278*, 1928−1933.
13. Duan, Y., and Kollman, P. A. (1998) *Science 282*, 740−749.
14. Daura, X., Jaun, B., Seebach, D., van Gunsteren, W. F., and Mark, A. E. (1998) *J. Mol. Biol. 280*, 925−932.
15. Daura, X., Gademann, K., Jaun, B., Seebach, D., van Gunsteren, W. F., and Mark, A. E. (1999) *Angew. Chem. Intl. Ed. 38*, 236−240.
16. Daura, X., van Gunsteren, W. F., and Mark, A. E. (1999) *Proteins 34*, 269−280.
17. Roccatano, D., Colombo, G., Fioroni, M., and Mark, A. E. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*(19), 12179−12184.
18. Lacroix, E., Viguera, A., and Serrano, L. (1998) *J. Mol. Biol. 284*, 173−191.
19. Munoz, V., Cronet, P., Lopez-Hernandez, E., and Serrano, L. (1996) *Fold. Des. 1*, 167−178.
20. Zhou, Y., and Karplus, M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 14429−14432.
21. Vanoni, M., Bertini, R., Sacco, E., Fonatanella, L., Rieppi, M., Colombo, S., Martegani, E., Carrera, V., Moroni, A., Bizzarri, G., Sabbatici, V., Catozzo, M., Colagrande, A., and Alberghina, L. (1999) *J. Biol. Chem. 274*, 36656−36662.
22. Huang, X., and Miller, W. (1991) *Adv. Appl. Math. 12*, 373−381.
23. Rost, B., and Sander, C. (1993) *J. Mol. Biol. 232*, 584−599.
24. King, D. S., Fields, C. G., and Fields, G. B. (1990*) Int. J. Pept. Protein Res. 42*, 255−266.
25. Martegani, E. Vanoni, M., Zippel, R., Cocceti, P., Brambilla, R., Ferrari, C., Sturani, E., and Alberghina, L. (1992) *EMBO J. 11*, 2151−2157.
26. Jacquet, E., Vanoni, M., Ferrari, C., Alberghina, L., Martegani, E., and Parmeggiani, A. (1992) *J. Biol. Chem. 267*, 24181−24183.
27. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−666.
28. Sklenar, V., Piotto, M., Lappik, R., and Saudek, V. (1993) *J. Magn. Reson., Ser. A 102*, 241−245
29. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
30. Baleja, J. D., Marmorstein, R., Harrison, S. C., and Wagner, G. (1992) *Nature 356*, 450−453.
31. Nilges, M., Gronenborg, A. M., Brünger, A. T., and Clore, G. M. (1988) *Protein Eng. 2*, 27−38.
32. Fioroni, M., Burger, K., Mark, A., and Roccatano, B. (2000) *J. Phys. Chem. B 104*, 12347−12354.
33. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., Di Nola, A., and Haak, J. R. (1984) *J. Chem. Phys. 81*, 3684−3690.
34. Van Gunsteren, W. F., Daura, X., and Mark, A. E. (1998) *Encycl. Comput. Chem. 2*, 1211−1216.
35. Van Gunsteren, W. F., Billeter, S. R., Eising, A. A., Hunenberger, P. H., Kruger, P., Mark, A. E., Scott, W. R. P., and Tironi, I. G. (1996) *Biomolecular simulation: The GROMOS96 manual and user guide*, vdf Hochschulverlag, ETH Zurich, Switzerland.
36. Hess, B., Bekker, H., Fraaije, J., and Berendsen, H. J. C. (1997) *J. Comput. Chem. 18*, 1463−1472.
37. Lenzen, C., Cool, R. H., Prinz, H., Kuhlmann, J., and Wittinghofer, A. (1998) *Biochemistry 37*, 7420−7430.
38. Carrera, V., Moroni, A., Martegani, E., Volponi, C., Cool, R. H., Alberghina, L., and Vanoni, M. (1998) *FEBS Lett. 440*, 291−296.
39. Esposito, G., and Pastore, A. (1988) *J. Magn. Reson. 76*, 331−336.
40. Colombo, G., Roccatano, D., and Mark, A. E. (2002) *Proteins: Struct., Funct., Genet. 46*, 380−392.
41. Kabsch, W., and Sander, C. (1983) *Biopolymers 22*, 2576−2637.
42. Kelly, J. (1997) *Structure 5*, 595−600.
43. Booth, D. (1997) *Nature 385*, 787−793.
44. Taveras, A. et al. (1997) *Bioorg. Med. Chem. 5*, 125−133.
45. Ganguly, A. K. et al. (1997) *Bioorg. Med. Chem. 5*, 817−820.
46. Ganguly, A. K. et al. (1998) *Biochemistry 37*, 15631−15637.
47. Wells, J. A. (1995) *Biotechnology 13*, 647−651.
48. Cunningham, B. C., and Wells, J. A. (1997) *Curr. Opin. Struct. Biol. 7*, 457−462.

BI0344026