JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Molecular Dynamics Simulations Provide Atomistic Insight into Hydrogen Exchange Mass Spectrometry Experiments

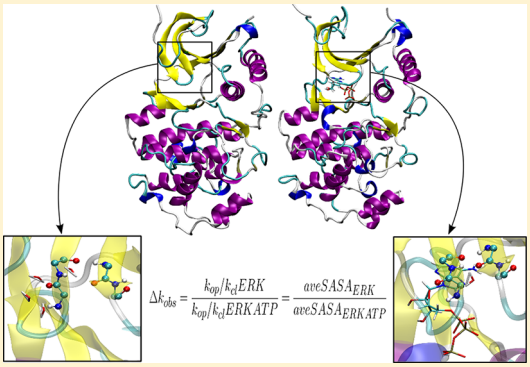
Ariel A. Petruk,[†,§] Lucas A. Defelipe,[†,‡,§] Ramiro G. Rodríguez Limardo,[†,‡] Hernán Bucci,[‡] Marcelo A. Marti,*[,†,‡] and Adrian G. Turjanski*[,†,‡]

[†]INQUIMAE-CONICET, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pabellón 2, Buenos Aires, C1428EHA, Argentina

[‡]Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pabellón 2, Buenos Aires, C1428EHA, Argentina

Supporting Information

ABSTRACT: It is now clear that proteins are flexible entities that in solution switch between conformations to achieve their function. Hydrogen/Deuterium Exchange Mass Spectrometry (HX/MS) is an invaluable tool to understand dynamic changes in proteins modulated by cofactor binding, post-transductional modifications, or protein−protein interactions. ERK2MAPK, a protein involved in highly conserved signal transduction pathways of paramount importance for normal cellular function, has been extensively studied by HX/MS. Experiments of the ERK2MAPK in the inactive and active states (in the presence or absence of bound ATP) have provided valuable information on the plasticity of the MAPK domain. However, interpretation of the HX/MS data is difficult, and changes are mostly explained in relation to available X-ray structures, precluding a complete atomic picture of protein dynamics. In the present work, we have used all atom Molecular Dynamics simulations (MD) to provide a theoretical framework for the interpretation of HX/MS data. Our results show that detailed analysis of protein−solvent interaction along the MD simulations allows (i) prediction of the number of protons exchanged for each peptide in the HX/MS experiments, (ii) rationalization of the experimentally observed changes in exchange rates in different protein conditions at the residue level, and (iii) that at least for ERK2MAPK, most of the functionally observed differences in protein dynamics are related to what can be considered the native state conformational ensemble. In summary, the combination of HX/MS experiments with all atom MD simulations emerges as a powerful approach to study protein native state dynamics with atomic resolution.



INTRODUCTION

Understanding the plasticity of protein's folded state is one of the challenges of modern science. After the great advances of X-ray crystallography, which provided unprecedented information of the so-called structural genomics dogma, "sequence determines structure determines function," it is now clear that proteins are flexible entities that in solution move in a coordinated and specific fashion, switching between conformational states. Moreover, protein conformational dynamics are modulated by cofactor binding, post-translational modifications, and protein−protein interactions that thereby modify and regulate their function. The limiting case of this phenomenon is an allosteric transition as described by the Monod−Wyman−Changeux (MWC) population shift conceptual model.[1] The NMR technique is today's most valuable tool to understand protein dynamics; however its application to proteins of medium to large size is still difficult.[2] As an alternative, Hydrogen/Deuterium Exchange Mass Spectrometry (HX/MS) experiments have emerged as an excellent tool to unravel protein dynamics and conformational changes.[3] However, the obtained information lacks atomistic detail, as usually changes are observed in peptides on the order of 10 amino-acids length. Moreover, changes observed in the HX/MS are usually not directly correlated with protein dynamics but interpreted with respect to available X-ray or NMR structures, thus preventing a complete atomic detail and dynamic understanding of the studied process.

In any given protein, N and O bound (i.e polar) hydrogens are exchanged as protons with the surrounding aqueous solvent spontaneously. In a common HX/MS measurement, the exchange of the backbone amide hydrogens by the heavier deuterium ions is measured, displaying rates ranging from milliseconds to years. The exchange reaction is followed over time as the mass increase of short peptides is determined due to proteolysis by an acidic protease. The final result, therefore, yields for each peptide (i) the total number of exchange hydrogens (#ENH) on the experimental time scale and (ii) three observed exchange rate constants, corresponding to fast, medium, and slow exchange NHs.

Received: June 21, 2012
Published: November 13, 2012

To understand the results of HX/MS experiments, a brief description of the exchange mechanism is needed as shown below in Figure 1.

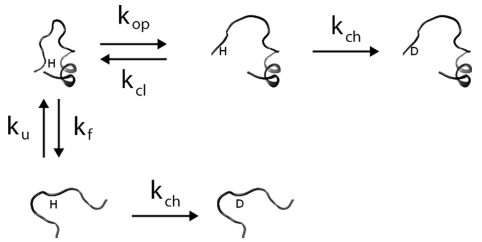


Figure 1. General scheme of protein equilibrium and amide hydrogen exchange. K_u and K_f are the unfolding and folding constants, respectively; the rest of the constants are described in the Introduction.

The observed exchange rate (k_{obs}) for a given backbone amide depends on two factors: (i) the "intrinsic factor," which is directly related to the acidity (i.e., the pka) of the corresponding proton, and (ii) the concentration of OH^- anions in solution (i.e., the environment pH). If the pH is fixed, the observed exchange rate corresponds to the intrinsic rate, which is usually called k_{int}. For folded proteins, it is assumed that totally buried residues do not exchange at all but that, due to protein dynamics, an equilibrium exits for each residue between an open conformation (where the NH exchanges with k_{int}) and a closed conformation where the residue is totally buried and thus unable to exchange. The equilibrium between open and closed conformations is determined by the corresponding rate constants for opening (k_{op}) and closing (k_{cl}). The observed exchange rate (k_{obs}) is thus related to the above-described equilibrium and rate constants as described by eq 1:

$$k_{obs} = \frac{k_{op} * k_{ch}}{k_{cl} + k_{ch} + k_{op}} = \frac{k_{op} * k_{ch}}{k_{cl} + k_{ch}} \quad (1)$$

In this scheme, two limiting cases or regimes appear, termed EX1 and EX2.[3,4] Both of these cases are summarized in Scheme 1.

Scheme 1. Limiting Cases for the Observed Exchange Rate (k_{obs}) Leading to the so Called EX1 and EX2 Regimes[4]

EX1 regimne: $k_{cl} << k_{ch} \Rightarrow k_{obs} = k_{op}$

EX2 regime: $k_{cl} >> k_{ch} \Rightarrow k_{obs} = \frac{k_{op}}{k_{cl}} * k_{ch}$

For native state proteins, evidence based on the pH dependence of k_{ex} confirms the predominance of the EX2 regime. Under these conditions, thus, the observed exchange rate directly measures the equilibrium between the open and closed states of the proteins (see the equation in Scheme 1) and thus provides insight on protein motion. What these conformations actually represent is still a matter of debate, and contributing to its understanding is one of the aims of the present study. In one extreme case, for those hydrogens that are deeply buried from the solvent, i.e., located in the interior of a protein domain, it is clear that open involves complete unfolding, while in the other extreme for completely exposed hydrogens, both open and closed states are meaningless. Between these cases lies the actual measured equilibrium motions, which under conditions where the folded state is mostly populated can be described as native state fluctuations (or breathing motions).[4] In other words, HX/MS experiments under the EX2 regime are able to monitor structural fluctuations from those small enough to involve the breaking of a single hydrogen bond to those large enough to involve complete unfolding. Determining which and to what extent these motions are relevant depends on the specific protein and is still an open issue.

HX/MS experiments are usually performed for the same protein in two different states or conditions (active vs inactive, bound vs unbound), and therefore, the results (i.e., the observed number of exchanges and their three rates) for each peptide are compared between them.[5] In this case, as the intrinsic rate is exactly the same for each residue, independently of the particular condition/state, the observed differences reflect changes in the corresponding open/closed equilibrium, thus reflecting changes in the structure and dynamics of the protein in the different studied conditions/states. The observed differences for each peptide are thus finally rationalized in terms of the available crystallographic structures of the protein. Commonly, the solvent exposure and/or number of hydrogen bonds of each residue or NH moiety are analyzed and compared between each studied state and corresponding structure. However, for those cases where the measured differences in the HX/MS experiments cannot be rationalized or are in apparent contradiction with the available structures, changes in the protein motions (or dynamics) between the states are invoked, although the nature of these changes remains completely unknown. Recently, a method called DXCOREX, based on the COREX algorithm, has been used to predict HX rates.[6] The COREX algorithm can estimate protein flexibility by using a statistical thermodynamic approach that calculates the energetic cost of exposing protein's atoms to solvent.[7] This approach is useful in analyzing protein models but cannot be used to measure the changes due to post transductional modifications, like phosphorylation, or binding of small molecules. Anyway, it is clear that the sole use of static structures to interpret HX/MS data prevents a complete atomic detailed dynamic picture of the obtained differences.

All atom Molecular Dynamics simulations (MD) provide atomistic detailed information of the protein dynamics on the nanosecond to millisecond time scale with state of the art computers.[8] When performed in explicit water, MD simulations have been shown to correctly reproduce protein solvent interactions, including the solvation structure of the different functional groups. Concerning protein motions, during plain MD simulations, as those performed in the present work, these methods have been shown to correctly describe a protein's native state structural fluctuations including small scale side chain rotations or small loop conformational sampling, and also large concerted motion of whole secondary structure domains. MD simulations are able to sample small local unfolding events, as some of those described above related to the mentioned open/closed equilibrium.[9−11] A molecular dynamics approach with a coarse grain Go type potential has been recently used to analyze hydrogen exchange NMR experiments, concluding that protein topology is relevant for exchange kinetics.[12] As already mentioned, the use of explicit solvent methods allows one to

dx.doi.org/10.1021/ct300519v | J. Chem. Theory Comput. 2013, 9, 658−669

directly measure the NH water interactions that determine the H/D exchange rate. Therefore, we expect that MD simulations should be able to predict and allow the rationalization of most of the observed HX/MS rates. It is important to remark that on some occasions large local unfolding events of secondary structure elements are required for protein function, which are difficult to sample using conventional MD simulations. However, by comparing simulations of a given protein in different states (as those measured in the HX/MS experiments), we expect to be able to correctly describe the dynamic changes that produce the observed differences in the exchange rates. In order to test this idea, we have selected the relevant and widely studied Mitogen-Activated Protein Kinase (MAPK) ERK2 protein as an ideal test case.

MAPK cascades are highly conserved signal transduction pathways coupling different extracellular signals to a variety of intracellular responses.[13] The structural characteristics of MAPKs have been studied in great detail with several crystal structures of members of the ERK, p38, and JNK groups solved.[14−19] These studies have provided a molecular understanding of the MAPK domain structure and in particular of the ERK2 protein (Figure S1). ERK2 is composed by two domains, an N-terminal domain formed largely by β-sheets and two helices, αC and αL16, and a C-terminal domain that is mostly helical, with four short β-strands that contain several residues involved in catalysis (Figure S1). A flexible linker between the N- and C-terminal domains allows their reorientation, and therefore it has been suggested that its conformation is important for the enzymatic activity.[20] The catalytic site, where the ATP and the magnesium ions bind, lies between these two domains. The MAPK insert is formed by two short helices located in the C-terminal domain and has been shown to be particularly flexible.[21] As MAPK cascades should transduce signals with high efficiency and fidelity for the cell to achieve normal cellular functions, to understand the molecular basis of these signal transduction cascades is of paramount importance.[22]

Recent studies have shown that dynamic conformational changes occurring on MAPKs upon phosphorylation, binding to ATP or upstream activators, scaffolds, and downstream targets may be important for MAPK activity by regulating the assemblage of the MAPK module.[18−20,23−25] It has been proposed that the binding of MAPKs to the docking motifs of their interacting partners may cause MAPKs to change their conformation and expose their activation loop, thereby becoming more accessible to MAPKKs and phosphatases.[18−20,23−25] Dynamic changes have also been observed in the structure of ERK1/2 upon phosphorylation and ATP binding in sites located far apart from the activation loop.[25] However, most of the results that relate dynamics to function come from X-ray structures that have revealed new conformations but give little information of the actual protein dynamics. On the other hand, important information has been obtained by HX/MS experiments of ERK2 and p38 in different states, active, inactive, with and without ATP, that gives dynamically related valuable information but lacks atomistic detail.[25]

Given the biological relevance of ERK2, its paradigmatic nature as a dynamic and allosterically regulated protein, the available HX/MS data, and our experience in MAPK structural dynamics, we selected this case for the development of a theoretical framework for the structural and dynamic interpretation of HX/MS data using all atom MD simulations. For the sake of this, we performed extended MD simulations of ERK2 in the inactive and active states (with and without ATP) and estimated the number of protons exchanged for each peptide under different conditions and the changes in exchange rates upon activation and ATP binding that are in very good correlation with HX/MS. Detailed intrapeptide analysis revealed that changes in NH solvent interaction inside each peptide are highly residue- and conformational-state-dependent, indicating that the interpretations based on the X-ray structure may be tricky, despite the fact that detailed structural information of the two end states is available. Finally, our data show that at least for ERK2MAPK most of the functionally observed differences in protein dynamics are related to what can be considered the native state conformational ensemble fluctuations. In summary, the combination of HX/MS experiments with all atom MD simulations emerges as a powerful approach to study protein native state dynamics with atomic resolution.

RESULTS

The results are organized as follows: First, we analyze shortly the general characteristics of the MD simulations. Second, we present calculations of several estimates of the number of NH exchanges for each peptide for comparison with the HX/MS experiments. Third, we present a detailed MD HX/MS data combined analysis of the exchange pattern of each peptide. Finally, we consider the changes upon activation and ATP binding in relation with ERK2 dynamics.

General Analysis of the Simulations. Our first analysis concerns the stability of each simulated system. As explained in the Computational Methods section, we performed simulations of ERK2 in four different states corresponding to inactive (ERK), inactive with ATP (ERK-ATP), active-phosphorylated (ERKpp), and active with ATP (ERKpp-ATP), as these correspond to those that have been previously analyzed by HX/MS experiments.[23,25,26] A representation of these four systems is depicted in Figure S1. The Root Mean Square Deviation (RMSD) of all α-carbons using as a reference the initial structure (built as described in the Computational Methods) computed for the last 90 ns is shown in Figure S2. All simulations are relatively stable displaying Cα RMSD values between 0.75 and 2.0 Å for all MDs.

In order to get a first insight into ERK2 dynamics, we computed the average Root Mean Square Fluctuation (RMSF) values for each residue in the four states (shown in Figure S3). The results show that high flexibility is observed in the β1−β2 beta sheets, the activation loop, the MAPKinsert, and the C-terminal loop, in agreement with previous results.[21,25,26] We observe that in the inactive and active states, ATP binding produces a decrease in the mobility of the β1−β2 sheets due to direct protein−ATP interactions. After activation (phosphorylation), both the activation loop and the MAPKinsert regions increase their mobility. In the ERK2 inactive crystal structure, the activation loop is in an extended conformation that folds over the kinase domain, allowing its interaction with the MAPKinsert. Upon activation, the reorientation of the activation loop disrupts the interactions with the MAPK insert, leaving this moiety free to rotate, and thus we observe an increase in its flexibility. Interestingly, both helices that form the MAPKinsert increase their flexibility only in the ERKpp-ATP state (Figure S3), highlighting also an allosteric role for ATP. The increased fluctuations observed in the activation loop upon phosphorylation are mainly localized in the first part of the

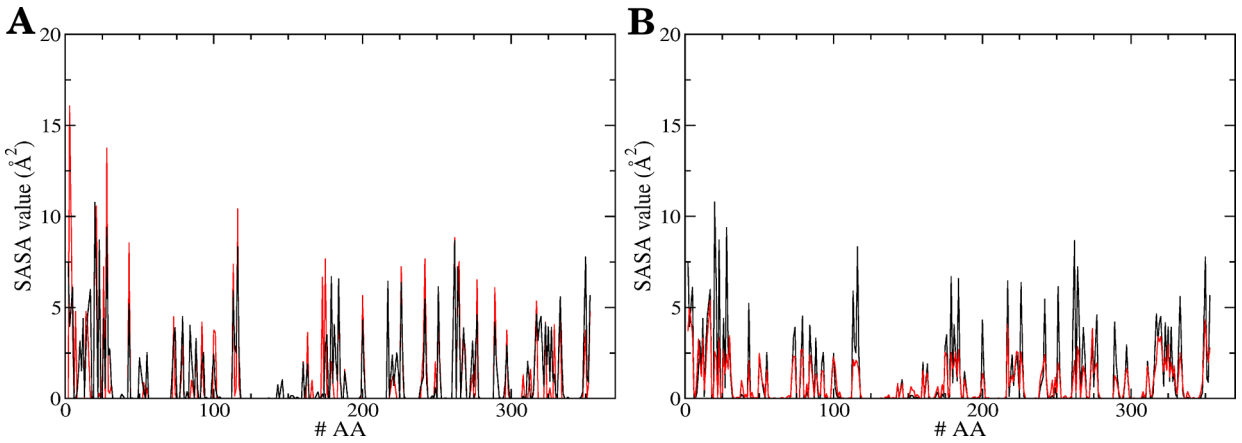


Figure 2. SASA values for each ERK2 residue: (A) aveSASA (black) and xrSASA (red) and (B) aveSASA (black) and stdvSASA (red). We selected as a test case the inactive ERK2 structure without ATP.

loop, before THR183, because in the active state this region is solvent exposed with minimal interactions with the rest of the protein.

The observed global changes in ERK2 dynamics upon activation and ATP binding presented so far are similar to those observed previously by our group for p38γ, suggesting the presence of a common dynamic pattern (and its functionally related changes) of the MAPK fold.[21] We now turn our attention to the analysis of amide hydrogen (NH) solvent interactions in ERK2 and their relation with the HX/MS experimentally derived data, which is the focus for the present work.

Analysis of Solvent Accessible Surface Area. As already mentioned, the initial insight into the interpretation of H/D exchange data is usually obtained from the analysis of NH. Solvent Surface Accessible Area (SASA) was derived from the corresponding crystal structure when available. It should be noted, however, that analyzing one photo of the ensemble that represents the folded protein is unable to account for all protein−solvent interactions, mainly due to the fact that some residues change their solvent exposure in the different conformations of the folded state. In order to analyze the corresponding dynamic effects and compare with HX/MS experiments, we computed, for each residue NH, the SASA corresponding to the initial crystal structure (xrSASA), the average SASA (aveSASA), and the standard deviation (stdvSASA), along each MD simulation. The idea behind the stdvSASA is to estimate the general variation in SASA along the simulation (see Computational Methods).

The results for ERK, depicted in Figure 2, show that although similar, the xrSASA and aveSASA present some differences; the correlation coefficient between both values (as defined in computational methods) is 0.753. We also plotted both values for ERK, ERK-ATP, ERKpp, and ERKpp-ATP systems; the results are shown in Figure S4. Most of the residues have similar aveSASA and xrSASA values; however, key differences are observed in regions that are important for ERK2 function, like the MAPK insert, the DEJL docking site, the activation loop, and the β1− β2 sheets. These results cannot be clearly explained by high order correlated motions, as they are localized only in a couple of residues located in diverse protein regions. The observed differences between the aveSASA when compared to the xrSASA are localized mostly in regions that are flexible and that switch between solvent exposed and buried conformations along the MD. Simulations in explicit solvent are able to capture these subtle but relevant changes, as depicted in Figure S10 for three selected residues. In the figure, we show the case of three residues that belong to the β1−β2 sheet region (see peptide 3 below): residue Y34, aveSASA > xrSASA with low stdvSASA (Figure S10, black solid line); residue G35, aveSASA > xrSASA with a high stdvSASA (Figure S10, red solid line); and residue A33, aveSASA < xrSASA with a high stdvSASA (Figure S10, green solid line). We want to emphasize that despite the fact that these three residues are consecutive in sequence, each of them has a different behavior, ruling out the possibility of explaining the observed differences as a result of the correlated motion of a whole moiety.

Surprisingly, the correlation coefficient between aveSASA and stdvSASA is high (0.878), suggesting that analysis of stdvSASA will lead to similar results as those for aveSASA for this system. On the other hand, the correlation between xrSASA and aveSASA highlights that the simulations are sampling, as expected, fluctuations around the crystal structure but that additional dynamic related information is captured.

In the HX/MS experiments, usually the difference in the exchange rates for the protein in two different states (i.e., inactive vs active, free or with bound substrate) is determined. Therefore, to have a direct comparison with the experimental derived data, we computed the change in aveSASA for each NH between two different states, as shown in Figure S5. The results show, strikingly, that there are pronounced differences in the aveSASA between nearby residues, and that there is no single protein region with at least three consecutive amino acids following the same trend. Despite the fact that we can identify some residues as becoming significantly solvent exposed upon activation, or buried upon ATP binding, a first conclusion is that changes in the structure and dynamics are not due to correlated changes of a whole secondary structure element and therefore require a detailed analysis at the residue level as we have done for each peptide in subsequent sections of the manuscript.

Another common practice in the interpretation of HX/MS experiments is to justify an observed increase in the number of exchanged hydrogens and/or exchange rate, due to an increase in the flexibility of the corresponding protein region, whenever the static picture provided by the X-ray structures shows no difference in the solvent accessible area of the corresponding

dx.doi.org/10.1021/ct300519v | J. Chem. Theory Comput. 2013, 9, 658−669

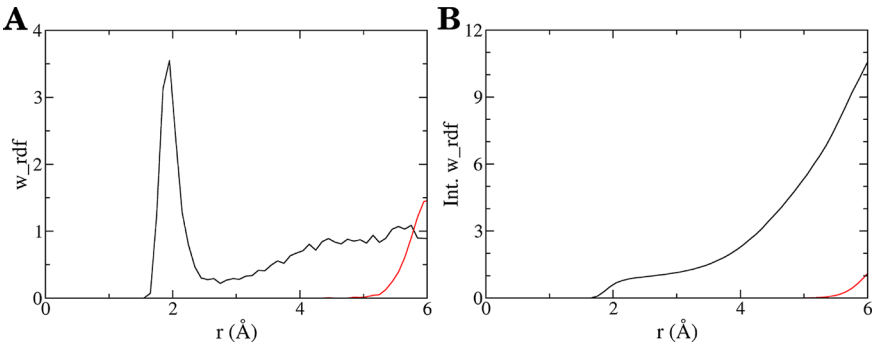


Figure 3. Water−NH interaction. (A) Water radial distribution function (w_rdf), which shows the probability of finding water molecules at a given distance, and (B) the integral of the w_rdf, which shows the average number of water found at a given distance, for the exposed 322 residue (black line) and the buried 154 residue (red line) NH amide of the ERK system.

backbone amides. The increased flexibility would allow the corresponding protein region to sample exposed conformations, not observed in the crystallographic structure, thus justifying the experimental observed trend. Therefore, we sought to analyze how the measured changes in aveSASA correlate with the previously determined RMSF values (which is a common and reliable measure of a given segment flexibility). Interestingly, the calculated cross correlation value between aveSASA and RMSF is only 0.577, showing that higher flexibility is not necessarily related with increased average solvation of the backbone amide. However, as will be shown later, for specific cases, flexibility, as revealed by the MD simulations, can possibly alter the exchange rate, by allowing the sampling of alternative conformations. We now turn our attention to the direct NH water interactions, as possible predictors of the observed rates.

Estimation of the Number of Water Molecules in the First NH Solvation Shell. Exchange of hydrogen atoms depends on their specific interaction with the solvent, i.e., water molecules; we therefore analyze in detail the contact of water molecules with the backbone NH atoms of the whole protein during the MD simulations. Initially, we computed for each NH the corresponding water radial distribution function (w_rdf) averaged over the whole MD simulation as described in Computational Methods. Figure 3 shows two typical examples of the w_rdfs obtained, one corresponding to a highly exposed NH and the other corresponding to a buried residue; both were taken from the ERK system. The figure, as expected, shows that exposed NH has a clear first solvation shell, while for a buried residue the probability of finding a water molecule close to the NH is directly zero.

Using the information obtained from the w_rdfs, we computed the average number of waters in the first solvation shell for all NH, by integrating the corresponding plot up to the first minimum, which is ca. 3 Å. The corresponding results, consisting therefore of the average number of water molecules found in the first solvation shell for each residue NH atom (#WFSS), plotted as a function of residue number for inactive ERK, are presented in Figure 4. For comparison sake, aveSASA values are also shown.

The results from Figure 4 show several interesting features. First, as expected for a secondary amide, the number of waters in the first solvation shell lies between 0 and 2. Again, as was observed for the aveSASA plots, it is evident that in most of the cases significant changes in water exposure occur between nearby residues; a residue may be highly exposed while its neighbor can be almost completely buried. Interestingly,

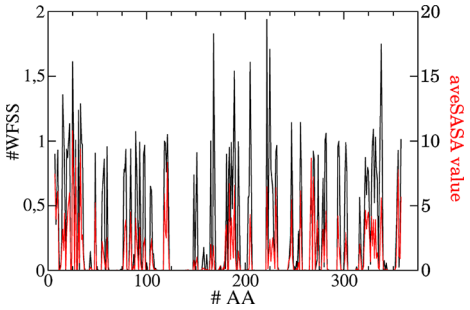


Figure 4. Number of waters in the first NH solvation shell (#WFSS) (black) and the aveSASA value (red) vs residue plot for ERK as obtained from the MD simulations.

aveSASA and #WFSS have an R^2 of 0.95 between them (Figure S6), meaning that both estimates can be used, despite the fact that the #WFSS is a direct estimate of water−protein interactions as compared to the aveSASA. Regardless of their similarity, most of the observed differences correspond to exchanges that are predicted by the aveSASA criteria and are not detected by the explicit solvent interaction analysis. We believe these differences are due to the intrinsic differences in the methodology. For the aveSASA estimation, we use a sphere of radius 1.4 Å in the Connolly algorithm; if we slightly increase the radius to 1.5 Å, we observed that some NH's are not predicted to be exchanged, as seen in the estimation based on the NH−water interaction. However, when looking at the overall results using a 1.5 Å probe, we do not get a better correlation, since some exchanges that were correctly predicted are now lost.

As X-ray structures sometimes have crystallographic waters included in the PDB, we decided to compare the results obtained from the simulation with the crystallographic water placement around the protein structure (Figure S7). The absence of crystallographic waters in the ERK structure (PDB ID: 1ERK) and the absence of ATP bound structures make this calculation impossible for these systems. Nevertheless, we performed this analysis over the ERKpp structure (PDB ID: 2ERK), which has some crystallographic waters. The results show that there is a good correlation with the solvated NH residues in the crystal structure and the dynamics. Most of the residues that have a water molecule in contact with a backbone NH amide on the X-ray structure have a water molecule in the first solvation shell during most of the dynamics. However, since there are only a few water molecules in the X-ray structure, one has to keep in mind that the ERKpp structure

dx.doi.org/10.1021/ct300519v | J. Chem. Theory Comput. 2013, 9, 658−669

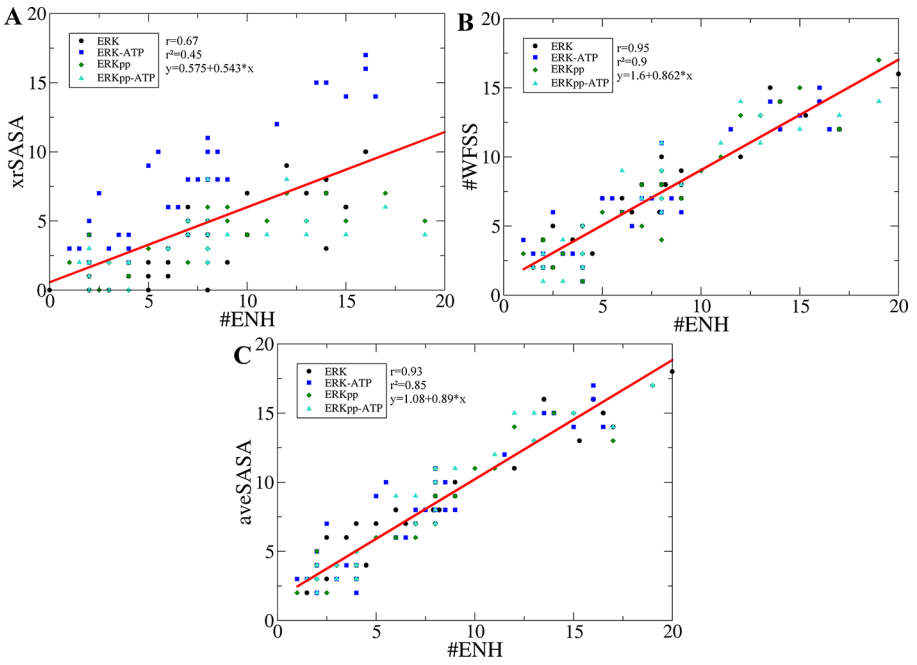


Figure 5. Regression analysis of experimentally determined NH amides that exchanged during the experiment compared with various *in silico* prediction methods. All the experimental data come from Ahn and co-workers.[23,25] (A) Number of amide hydrogens that exchanged per peptide (#ENH) vs the crystal solvent accessible surface area (xrSASA) determined from the corresponding structures, namely, ERK (PDB: 1ERK), ERKpp (PDB: 2ERK), and the model structures of ERK-ATP and ERKpp-ATP. (B) Number of amide hydrogens that exchanged per peptide (#ENH) vs the number of residues with the #WFSS above zero for each peptide (#WFSS) determined by a w_rdf analysis of the last 90 ns MD. (C) Number of amide hydrogens that exchanged per peptide (#ENH) vs the number of residues per peptide that has average SASA (aveSASA) above zero as determined from molecular dynamics simulations. The different colors represent the four systems as shown in the legend, ERK (black), ERK-ATP (blue), ERKpp (green), and ERKpp-ATP (cyan).

does not show crystallographic waters all around the protein, and therefore there is a significant number of residues that interact with solvent molecules in the simulation that do not have a crystallographic water in the X-ray structure.

As for the aveSASA, we also compared the number of water molecules in the first solvation shell of each NH of ERK2 in the different simulated states. The general trend shows that deeply buried regions remain completely buried in most cases but that there are several significant abrupt changes in the solvation of neighbor residues, as already described before for the aveSASA. Therefore, and despite the fact that in some cases a large protein region (i.e more than five aminoacids) becomes more exposed, like residues 185−189 upon ATP binding to inactive ERK2, or less exposed, such as residues 329 to 334 upon activation, in most regions a complex behavior of the solvation pattern is observed.

In summary, our analysis reveals a complex pattern of changes in the solvation structure of the protein amides that precludes a simple interpretation of the observed HX/MS exchanges that usually produce results of whole peptides comprising 5−20 residues. It is clear that the net observed change for a peptide will be an average effect of each individual NH behavior. As we expected, changes in peptide solvation could be explained based on one or two intrapeptide amide changes leaving most of the residues unchanged as has been done previously in several works. On the other hand, we found that, in most peptides, changes are complex involving most of the individual NH groups that on average behave like a change in one or two protons; we expect difficulties in the interpretation of the HX/MS data based solely on the crystallographic structures. We will now analyze the predictive power of both computed values in relation with the data derived from the HX/MS experiments.

Correlation between MD Simulations and Hydrogen Exchange Experiments. To directly relate our data with the HX/MS experiments, we decided to compare the experimentally observed number of exchanged amide hydrogens for each peptide, as determined by Ahn and co-workers,[23,25] with the predicted number based on either the available crystal structures or our simulation derived data. In Figure S8, we show the secondary structure elements together with the sequence and the experimentally analyzed peptides for ERK and ERKpp as performed by Ahn and co-workers. We assumed that a hydrogen would be observed to exchanged in the time course of the experiment if the SASA (determined from the X-ray structure, or averaged over simulation) or the number of waters in the first solvation shell is above zero, meaning that on the nanosecond time scale at least once, the NH group is interacting with a water molecule. The corresponding results are shown in Figure 5.

As shown in Figure 5, we observe a remarkable improvement in the agreement between the experimental observed and predicted values, when we considered the information derived from the MD simulations for all calculated systems. The correlation between both values, measured as the correlation coefficient (R), rises from 0.67 using solely the crystal structures to a striking 0.95 when the simulation data are included. Interestingly, the prediction derived using the data from the first solvation shell performs slightly better. These results not only highlight the potential predictive power of all atom MD simulations for the interpretation of HX/MS data but also show that at least for this protein, the experimental

dx.doi.org/10.1021/ct300519v | J. Chem. Theory Comput. 2013, 9, 658−669

data are capturing mostly the structural and dynamic properties of native state fluctuations that can be sampled reasonably well using all atom MD simulations on the nanosecond time scale and that no major unfolding events are required to understand the observed trend. The exchange described in per residue level detail can be seen in Figure S9. Residues that are predicted to exchange are shown in red, and the ones that are buried are shown in blue. As expected, residues belonging to the hydrophobic core show no NH exchange, whereas more exposed residues do.

Given the observed improvement in the general agreement between the experimentally observed and predicted number of exchanged hydrogens, we will now discuss the general changes in ERK2 dynamics upon activation and ATP binding and will then deepen the analysis and interpretation of the changes in the HX/MS pattern for each peptide upon ERK2 ATP binding and/or activation (phosphorylation).

General Changes in ERK2 upon Activation and ATP Binding. Detailed analysis of the results shows that there are significant changes in the mobility of the protein upon activation and ATP binding. Changes upon activation, that are observed both in the ATP bound and unbound states, correspond mainly to the MAPKinsert moiety and the activation loop. The flexibility of the MAPKinsert increases upon activation, probably due to the disruption of the interactions with the activation loop (Figure 6). Interestingly,

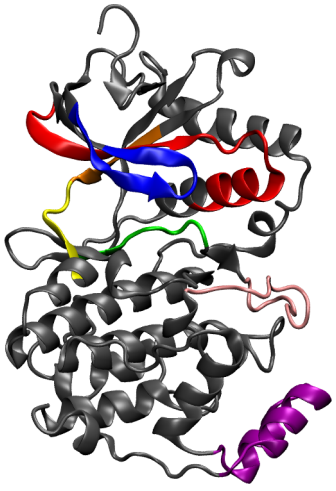


Figure 6. Important regions as determined from HX-MS data. ERK peptides 3 (blue), 6 (red), 12 (orange), 13 (yellow), and 18/19 (green) and Activation loop (pink) and MAPKinsert (purple) domains. We show the ERK system structure.

increased flexibility is observed in the first region of the activation loop, residues 175 to 180, when Thr183 and Tyr185 residues are phosphorylated. The ERK2 activation loop is longer than other MAPKs like p38s, and only the last part of the loop, residues 181 to 186, has strong interactions with rest of the protein, leaving the first six residues facing the solvent and quite flexible. Residues 181 to 186 are rigid in both states. Despite the fact that in the inactive state the loop is extended, both Phe181 and Leu182 are positioned inside a hydrophobic groove formed by the MAPKinsert and helixG, and they are quite rigid. In the active state, phosphorylation of the two conserved residues leads to strong interactions with nearby lysine and arginine residues, preventing the movement of both Phe181 and Leu182. General changes upon ATP binding are observed in the $\beta1-\beta2$ beta sheets, αD$-\alpha$E loop (docking site), and the second helix of the MAPKinsert. The decreased flexibility observed when ATP is bound in the region belonging to the $\beta1-\beta2$ sheets can be explained based on the direct interactions with the cofactor. Changes in flexibility in the docking region and in the MAPKinsert are not due, as expected, to ATP direct interactions but to contacts of the adenine moiety of ATP with the hinge region and the activation loop that connects to both regions.

Detailed Analysis Changes in the HX/MS Pattern upon ATP Binding to ERK2. In the N-terminal domain of ERK2, there are two beta sheets, $\beta1$ and $\beta2$ (Figure 6), separated by a short glycine rich loop. This key region, which is included in peptide 3 (res. 27−38, SYIGEGAYGMVC), is involved in the interaction with ATP as shown in other protein kinases[27] and experiences significant changes in the H/D experiments upon both ATP binding and activation. Changes upon activation are regulated by the allosteric conformational rearrangements, but changes dependent on ATP binding could be also explained by direct interactions with the cofactor.

Upon ATP binding to inactive ERK2,[23,25] for peptide 3 a one residue decrease in the number of exchanged NH is observed in the HX/MS experiments. As already explained, although particular observed rates of HX cannot be predicted, changes in the exchange rate (Δk_{obs}) should reflect the changes in the NH solvation pattern as determined from the aveSASA computed from the MD data. To estimate the corresponding change, we thus computed the ratio of the aveSASA between the two states, which in this case corresponds to ERK2 in the free and ATP bound states. This relation is shown in eq 2.

$$\Delta k_{obs} = \frac{\frac{k_{op}}{k_{cl}}ERK}{\frac{k_{op}}{k_{cl}}ERKATP} = \frac{aveSASA_{ERK}}{aveSASA_{ERKATP}} \quad (2)$$

To visualize the change in the solvation pattern of peptide 3, Figure 7 shows the logarithm of the ratio of aveSASA between the free and ATP bound states. Thus, a positive value represents a decrease in the aveSASA upon ATP binding, which should reflect a decrease in the observed exchange rates, whereas a negative value represents the opposite.

In this case, residues 35, 36, and 37 hide completely from the solvent upon ATP binding, which can easily be predicted by inspection of the protein structure. However, residues 30 and 32 become significantly exposed only in the ATP bound state. In Figure 7B,C, we show that upon ATP binding the NH of Gly30 loses its interaction with Val37 carbonyl, which modifies its position due to ATP binding, and therefore Gly30 NH interacts with water. A simple rule of thumb suggests that when a looking at the whole peptide, a significant decrease in the exchange rate of one residue is expected, consistent with the observed data.

While peptide 3 showed the above-described changes, peptide 4 (res. 39−44 SAYNDL) did not. In agreement with these results, we do not see any changes in the solvation pattern for the corresponding residues.

Peptide 6 (res. 45−69, NKVRVAIKKISPFEHQTYCQRTLRE) showed a decrease in the exchange rates in the nucleotide-bound form of ERK and ERKpp. The MD shows that in ERK, Ser55 NH decreases both the number of waters and aveSASA upon ATP binding, but both values do not show any difference after ATP incorporation to ERKpp. Ser55 performs hydrogen bonding with the carboxyl backbone of

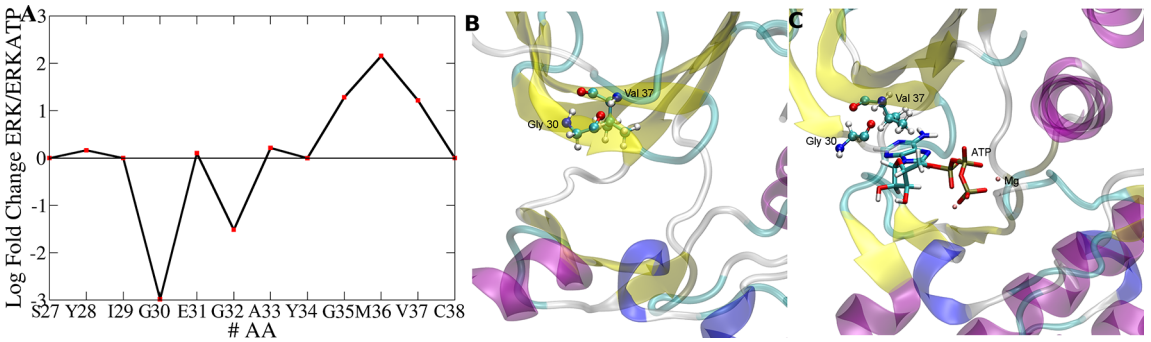


Figure 7. (A) Differential solvation of ERK peptide 3 (residues 27 to 38) in the absence or presence of bound ATP. The results are shown as the log of fold change between ERK/ERKATP aveSASA. (B and C) Structures of ERK and ERK-ATP respectively showing Gly 30 and Val 37 differential positioning.

Tyr34, helping the rotation of the N-terminal β-sheets. This is an interesting result because it highlights that the $\beta1-\beta2$ region rotates not in the active state but upon ATP binding to the inactive state. Thus, combination of the H/D experiments and the simulations suggests that ATP stabilizes the N-terminal domain in a conformation suitable for catalysis.

Peptides 12 and 13 are also important in MAPK flexibility because they include the hinge region that regulates the orientation of the N-terminal domain with respect to the C-terminal domain. We have previously shown that ATP binding to p38γ stabilizes the active conformation in part due to direct interactions with the hinge region.[21] The backbone carbonyl oxygen atom of Asp104 forms a hydrogen bond with the N6 of the adenine ring (Figure S11). The Met106 NH backbone also forms a hydrogen bond with the N6 of the adenine ring, and the acidic group of Asp109 forms hydrogen bonds with the OH2 of the ribose ring (Figure S11). Similar interactions are observed both in the inactive and active states and are maintained along the simulation, indicating that this region has similar interactions with the cofactor. The HX/MS data for this region show that in peptide 12 (res. 101−105, IVQDL) there is a decrease in the exchange rates and a reduction in the number of exchanges upon ATP binding. The MD results show that, due to ATP incorporation into ERK, the Leu105 amide group decreases both the number of water molecules and the aveSASA. Despite the fact that the Leu105 backbone is facing the solvent in the ATP bound state, its NH group interacts with the Asp104 acidic group and therefore significantly reduces its direct interaction with the solvent (Figure S11, panels A and B). This change can be ascribed to a change in conformation in the hinge region induced by the direct interaction of the Asp104 carboxyl moiety with ATP and to the mentioned rotation of the N-terminal domain.

Peptide 13 (res. 106−110, METDL) showed decreased hydrogen exchange rates upon nucleotide binding for both ERK and ERKpp. MD shows that after ATP incorporation to ERKpp, Met106 and the Glu107 NH groups decrease both the number of water molecules and aveSASA, in agreement with the observed results (Figure S11, panels C and D). For inactive ERK, similar but smaller changes are also observed. Interestingly, in the unbound inactive state, we identified two conformations for the side chain of Met106, one facing the solvent and leaving the NH group free to interact with water molecules and a second one where the side chain of Met106 occupies the ATP pocket, impeding the water molecules to interact with its backbone. The interconversion between these two conformations is slow as compared to the simulation time scale, and probably longer computational simulations are necessary to accurately measure each conformation population. In any case, we found that direct interaction with water is only possible in the ATP unbound state, explaining the difference observed in the HX/MS experiments.

No difference in the exchange pattern is observed in the HX/MS experiments for peptide 12/ME (res. 101−107 IVQDLME). In agreement, we do not observed any difference in the solvent interaction of residues 108−110 (TDL of peptide 13) when we compare the four states. However, residues Met106 and Glu107 are the ones responsible for the changes observed in peptide 13, and they seem to compensate for the changes observed for peptide 12, highlighting once again the complexity of analyzing the data at the peptide level.

Peptide 18/19 (res. 162−168, KICDFGL) showed a decreased hydrogen exchange in the ERKpp case upon ATP binding, which could be modeled by a 34-fold decrease in the rate of an intermediate exchanging amide upon nucleotide binding.[23,25] Slight changes were observed for ERK. The MD simulations depict a complex picture, after ATP incorporation to ERK Cys164 increases both the number of water and SASA values, but there is a 3-fold decrease in Gly167. On the other hand, in ERKpp, the Ile163 NH group increase its interaction with the solvent upon ATP binding, and Gly167 decreases. This complex behavior makes simple interpretations of HX/MS difficult, especially in highly flexible regions.

Detailed Analysis in the HX/MS Pattern due to ERK2 Phosphorylation Mediated Activation. In the HX/MS experiments, peptide 3 (res. 27−38, SYIGEGAYGMVC) showed a conversion of a single amide with a slow exchange rate into a fast exchange upon activation. As the analysis of the crystal structures does not allow an explanation based on conformational changes, an increase in mobility in this region when the protein gets phosphorylated was proposed. The RMSF results for this region show a significant reduction in mobility in residues Gly32, Ala33, and Tyr34 upon activation. However, the aveSASA indicates an increase upon activation in Glu31 and a less significant decrease in Tyr28. Interestingly, when we observe the SASA along the dynamics for Gly32, only one exposed conformation appears in the active state, but two conformations, one with low SASA and one with higher SASA, appears in the inactive state. The number of water molecules shows no significant difference for this residue, indicating that exchange is possible, but oscillations in the conformation of the amide could account for the observed differences. One difference that we observe is the loss of a water molecule

dx.doi.org/10.1021/ct300519v | J. Chem. Theory Comput. 2013, 9, 658−669

interaction with Gly35 upon activation, something that is not seen in the experiments.

Peptide 12 (res. 101−105, IVQDL) showed a decrease in the rate of hydrogen exchange upon activation consistent with the conversion of one intermediate amide to a slow one, a result that cannot be explained by a conformational change as judged from the X-ray structures. In the simulations, we observe that both Asp104 and Leu105 reduce their interactions with the solvent as judged by a reduction in the number of water molecules observed in the first solvation shell and in the aveSASA. Interestingly, the total number of hydrogen atoms exchanged should not change, as observed in the experiments, because in both states there is a conformation where exchange is possible. However, dynamic changes upon activation lead to a change in population between the open and closed states, favoring the latter in the active case. If we traduce the change in populations to ΔG, we obtain a stabilization of the closed state by 0.5−1.0 kcal/mol, similar to the experimental observed values.

Peptide 13 (res. 106−110, METDL), in contrast to peptide 12, showed an increase of the hydrogen exchange rate upon activation. In agreement, the molecular dynamics simulations show an increase in the number of waters and in the aveSASA of residues Met106 and Glu107 and to a lesser extent Thr108 (Figure S11, panel E and F). Again, as in peptide 12, we identified two conformations, and activation produces a shift in the populations. I,n this case an open conformation is favored, consistent with the observed increase in exchange.

The loop that is included in peptides 12/13 is important for the orientation between the C-terminal and N-terminal domains. Analyzing its flexibility is important for understanding protein function. The more flexible region is in the middle of the loop, residues 104 to 107. However, we observe a net change in flexibility in Asp104, adopting a rigid structure upon activation, which could be important for ATP binding, as it will reduce the entropy associated with ligand binding. On the other hand, Glu107 increases its flexibility upon activation. What we observed is an increased exchange between the two different conformations that are involved in the regulation of the domain−domain orientation upon activation.

In summary, both the predicted number of exchanged hydrogens for each peptide and the changes in the solvation pattern of individual NH atoms analyzed for particular peptides show good correlation with the HX/MS experiments. We have been able to reproduce the values for the number of exchanged hydrogens in each peptide and the changes in the number and rates observed for ERK2 in the different studied states, respectively. Therefore, the MD results allow an atomic detailed structural and dynamic interpretation of the experimentally observed trends, thus providing further insight into the structural and dynamic properties of the studied protein and its associated functional changes.

DISCUSSION

All Atom MD As Powerful Tool to Interpret and Analyze HX/MS Data. As described in the Introduction, the first result of common HX/MS experiments consists of (i) the total number of exchange hydrogens for each peptide on the experimental time scale and (ii) the three (slow, medium, and fast) observed exchange rate constants. Since, usually HX/MS experiments are performed for the same protein under two different conditions under the EX2 regime, the observed differences should reflect differences in the corresponding open−closed dynamic equilibrium, between the studied states. This equilibrium involves changes that range between a local unfolding of specific secondary structure elements to structural fluctuations of the native state structure. What is actually measured in the HX/MS experiments is not always clear, and MD can yield valuable and detailed information provided a good correlation with the measurements is obtained. In the present work, we have shown that all atom explicit water MD simulations are able to reasonably predict the number of exchanged hydrogens for each peptide, based on the whole dynamic averaged Solvent Accessible Surface Area (aveSASA) and/or the number of waters in the first solvation shell (#WFSS) of each NH amide. This result strongly suggests that most of the measured exchanges are due to completely exposed NH atoms or require only native state fluctuations that involve breaking of a small number of hydrogen bonds. Therefore, local unfolding events of secondary structure elements seem not to play a major role in relation to the common HX/MS experiments, at least for ERK2.

The comparison of the aveSASA and #WFSS taken from the MD for the same protein in different states also allows rationalization of the experimentally observed changes in the number of exchanged hydrogens and their rates. Moreover, the changes can be traced down to both, changes in the equilibrium structure (or the native state ensemble), as can be determined from the comparison of available X-ray structures for some cases, as well as those changes due to changes in the protein motions that cannot be revealed by comparison of the equilibrium structure. These changes usually correspond to either a significant increase/decrease in the region flexibility or the presence of an alternative less populated conformation (that could be assigned as the open state) whose probability equilibrium is shifted in the different states.

Finally, our results show that for many peptides, the predicted exchange pattern is very complex, with neighboring residues having significant differences in their solvent interactions, and thus the rather simple interpretation related to the change in the number of exchanged hydrogens can be incomplete or even misleading.

What Does the Combination of HX/MS and MD Results Reveal Concerning ERK2 and MAPK Functional Dynamics? Even though MAPKs structure has been extensively studied, we still do not have a clear picture of how their function is regulated at the molecular level.[14,28,29] Moreover, several studies pointed out that protein dynamics and allosterism play an important role in the regulation of the MAPK function.[19,21,23] Most of this knowledge is derived from X-ray crystallographic structures of different MAPKs in different states.[17,19,28−30] HX/MS experiments have been performed to study several MAPKs and their associated process of activation, substrate binding, and interaction with upstream/downstream regulators,[20,23−25] and the observed changes in the exchange pattern were interpreted in terms of the available structures. Interestingly, for several regions the results could not be rationalized or were in clear contradiction with the observations derived from the structures, and were therefore interpreted as the result of significant dynamic changes along the studied process. However, no detail concerning the nature of the proposed dynamic role could be presented. In the present work, we have performed a new analysis/interpretation of the HX/MS data, in light of the results obtained from all atom MD simulations of ERK2 in the different states.

Our results also allow the interpretation of the HX/MS results that invoked changes in protein dynamics, mainly, the reduction in exchange for the F and G helices (also observed for p38α) and the increase in the MAPK insert exchange.[20,23−25] Our data show that phosphorylation significantly reduces F, G, and H helices' mobility and that significant changes are observed in the MAPK insert dynamics, in agreement with the HX/MS data.

General Conclusions of the Simulations. By performing all atom MD simulations of ERK2 in different states, we have shown that the whole dynamic aveSASA and #WFSS of each NH amide are both good predictors of the number of protons exchanged for each protein region (peptide) as determined in HX/MS experiments. We also show that the aveSASA and #W allow a rationalization of the observed exchange rates and the experimentally observed changes in the rates associated with different protein conditions. Finally, we propose that at least for ERK2MAPK, most of the functionally observed differences in protein dynamics are related to what can be considered the native state conformational ensemble. In summary, the combination of HX/MS experiments with all atom MD simulations emerges as a powerful approach to study protein native state dynamics with atomic resolution.

COMPUTATIONAL METHODS

Setup of the System. Initial coordinates for ERK2 in the different studied states were retrieved from the protein Data Bank. The corresponding PDB codes are for inactive unphosphorilated ERK (PDB ID 1ERK) and for active phosphorylated ERK (PDB ID 2ERK). The ATP bound structures were built from inactive and active ERK structures, adding the ATP *in silico*. To build the ATP bound structures, we used two sources of information. First, there is one unphosphorylated ERK2 structure with bound ATP (PDB ID 1GOL). The structure, however, corresponds to a single point mutant K52R and hosts only one Mg^{2+} inside the ATP bound active site. Thus, although it correctly "docks" the ATP inside the ERK2 ATP binding groove, the ATP phosphates are not properly oriented in an active structure. The best active-like structure of a MAPK (i.e., with the two required Mg^{2+} ions and ATP bound) is that of activated p38γ (PDB ID 1CM8). Thus, we used both these structures to build the wild-type two Mg^{2+} ATP bound ERK2 structure both in the inactive and active conformation. It is important to remark that we have previously conducted a QM/MM study[31] where we have clearly shown that we were able to model the ATP bound structure and obtained results related to the reaction mechanism, in clear agreement with experiments. The high structural similarity, in particular, of the active site, and phosphorylation mechanisms of human MAPKs, plus the fact that the built structure is stable during long molecular dynamics simulations, supports the quality and reliability of the ATP bound structures (Figure S2).

For all structures, hydrogens were added with the tleap module of the AMBER 9 package of programs.[9] Standard protonation states were assigned to tritable residues (D and E are negatively charged; K and R positively charged) Histidine protonation was assigned favoring formation of hydrogen bonds in the crystal structure. Each protein was then immersed in a truncated octahedral box of TIP3P waters. The system consisted of the protein plus ca. 10.000 water molecules, and two MG ions and ATP for ERKpp-ATP and ERK-ATP.

MD Simulations Parameters. Each system was first optimized using a conjugate gradient algorithm for 2000 steps. This optimization was followed by 100-ps-long constant volume MD simulation, where the temperature of the system was slowly raised to 300 K. The heating was followed by a 100-ps-long constant temperature and constant pressure MD simulation to equilibrate the system's density. During these processes, the protein CA atoms were restrained by a 10 and 1 kcal/mol harmonic potential for each process, respectively. Pressure and temperature were kept constant with the Berendsen barostat and thermostat, respectively.[32] All simulations were performed with periodic boundary conditions; the SHAKE algorithm was used to retain the hydrogens at equilibrium bond lengths, allowing the use of a 2 fs time step. The water model corresponds to TIP3P waters. All protein residue parameters correspond to the Amber 99SB force field.[33] ATP parameters were taken from Meagher et al.[34] All parameters are the same as those used in our previous study of MAPK p38γ.[21] For each system, 100-ns-long constant temperature and pressure MD simulations were performed without restraints.

w-rdf and Number of Solvating Water Calculations. To analyze the amide backbone (NH) solvent interactions, we computed the water radial probability distribution function (w_rdf) for each NH atom over the whole simulation. For the sake of this, the number of water molecules on all production snapshots at a distance "r" from the NH was counted and normalized (i) by the number of analyzed snapshots and (ii) with respect to the bulk water density (1 g/cm^3). Thus, the w_rdf units are cm^3/g. The NH−water distance was defined as the distance between the amide hydrogen and each water oxygen atom. As expected, all w_rdf's for relatively exposed hydrogens showed the same peak and first minima position corresponding to the optimal hydrogen bond distance. Integrating the w_rdf up to the first minimum allows for computing the average number of water molecules in the first solvation shell for each amide hydrogen during the simulation.

Root Mean Square Fluctuation (RMSF) Calculations. The RMSF is a measure of the average deviation between the position of particle i along the simulation and the average position of particle i. All α carbons of the protein were aligned, and the RMSF was measured for each residue using the following equation:

$$\mathrm{RMSF} = \sqrt{\frac{1}{T}\sum_{t_j=1}^{T}(x_i(t_j) - x_i)^2}$$

where T is the simulation time over which one wants to compute the deviation, in our case the last 90 ns of the MD simulations, and x_i is the average position of the amino acid i.

Solvent Accessible Surface Area (SASA) Calculations. We computed the solvent accessible surface area by using the corresponding plug-in based on the Connolly surface algorithm as implemented in the VMD program.[35] The radius of the prove used was 1.4 Å. Three different SASA values were measured: (i) the initial crystal structure (xrSASA), (ii) the average SASA (aveSASA) of the MD, where we have computed the SASA of each snapshot of the simulation, taken every 10 ps, for the last 90 ns of the MD and computed the average value, and (iii) the standard deviation of the aveSASA (stdvSASA). Since proteins are flexible entities, each residue can be buried or exposed most of the time; i.e., a constant value of the SASA is obtained along the MD simulation, and therefore the residue will have a low stdvSASA. On the other hand, the residue can

oscillate between conformations that may have different SASAs; i.e., a variable SASA value is obtained along the MD simulation, and therefore the residue will have a high stdvSASA. In this sense, aveSASA along the MD gives information about the average behavior of the protein in solution, and stdvSASA tells us if this behavior is due to one conformation or multiple conformations that have significantly different SASA values. As shown by our results, ave and stdvSASA are not necessarily correlated. In Figure S10, we show three paradigmatic cases: (i) residue Y34 where aveSASA > xrSASA with low stdvSASA (Figure S10, black solid line), (ii) residue G35 where aveSASA > xrSASA with a high stdvSASA (Figure S10, red solid line), and (iii) residue A33 where aveSASA < xrSASA with a high stdvSASA (Figure S10, green solid line).

Number of Amide Hydrogens That Exchanged Per Peptide (#ENH). We estimated the #ENH per peptide by using two outcomes from the simulations, the aveSASA and #WFSS per residue. We decided that a residue will be observed to exchange if, during 90 ns of simulation time, at least in one of the analyzed snapshots (we took a snapshoot every 10 ps), we observe that the NH interacts with a water molecule as previously defined. Once we decide a residue is expected to exchange, we sum over all residues of the corresponding peptide and calculate the #ENH.

Correlation Coefficient. The correlation coefficient between any two obtained parameters vs residue plot was computed according to the following equation:

$$\sigma(A, B) = \frac{\sum_i^N A_i \times B_i}{\sqrt{\sum (A_i)^2 \times \sum (B_i)^2}}$$

Figure Rendering. All figures were produced using a built-in pymol[36] ray tracing module or using a vmd tachyon ray tracing module[35] and compiled using GIMP.

■ ASSOCIATED CONTENT

Supporting Information

Tables S1−S3, hydrogen exchange data as predicted by various methods. Figure S1, structure of MAPK ERK2. Figure S2, RMSD for the last 90 ns of MD simulations. Figure S3, RMSF of the last 90 ns of MD simulations. Figure S4, aveSASA and xrSASA of the four states of ERK plotted over the structure. Figure S5, differences in aveSASA between the four states. Figure S6, regression analysis between aveSASA and #WFSS. Figure S7, number of waters calculated from MD and from X-ray structure. Figure S8, sequence and peptide diagram of ERK and ERKpp. Figure S9, predicted exchange computed by aveSASA plotted over the structure. Figure S10, SASA over time of the four ERK states. Figure S11, changes of peptides 12 and 13 shown with atomistic detail. This information is available free of charge via the Internet at http://pubs.acs.org/

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: marcelo@qi.fcen.uba.ar, adrian@qi.fcen.uba.ar.

Author Contributions

§Both authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the anonymous reviewers for comments that have significantly improved this manuscript. A.A.P. is a CONICET postdoctoral fellow, L.A.D. and R.G.R.L. are CONICET doctoral fellows. A.G.T. and M.A.M. are members of the CONICET. This work was supported by PRH PICT No. 2009-00121 awarded to A.G.T., PICT-No. 2010-2805, and Universidad de Buenos Aires CyT No. 20020110100061 awarded to A.G.T. and M.A.M. Computer power was generously provided by CeCAR at FCEN-UBA and SimOne at INSIBIO-CONICET.

■ REFERENCES

(1) Monod, J.; Wyman, J.; Changeux, J. P. On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.* **1965**, *12*, 88−118.

(2) Grzesiek, S.; Sass, H. J. From biomolecular structure to functional understanding: new NMR developments narrow the gap. *Curr. Opin. Struct. Biol.* **2009**, *19*, 585−95.

(3) Konermann, L.; Pan, J.; Liu, Y. H. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem. Soc. Rev.* **2011**, *40*, 1224−34.

(4) Hoofnagle, A. N.; Resing, K. A.; Ahn, N. G. Protein analysis by hydrogen exchange mass spectrometry. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 1−25.

(5) Chalmers, M. J.; Busby, S. A.; Pascal, B. D.; West, G. M.; Griffin, P. R. Differential hydrogen/deuterium exchange mass spectrometry analysis of protein-ligand interactions. *Expert Rev. Proteomics* **2011**, *8*, 43−59.

(6) Liu, T.; Pantazatos, D.; Li, S.; Hamuro, Y.; Hilser, V. J.; Woods, V. L., Jr. Quantitative assessment of protein structural models by comparison of H/D exchange MS data with exchange behavior accurately predicted by DXCOREX. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 43−56.

(7) Hilser, V. J.; Garcia-Moreno, E. B.; Oas, T. G.; Kapp, G.; Whitten, S. T. A statistical thermodynamic model of the protein ensemble. *Chem. Rev.* **2006**, *106*, 1545−58.

(8) Klepeis, J. L.; Lindorff-Larsen, K.; Dror, R. O.; Shaw, D. E. Long-timescale molecular dynamics simulations of protein structure and function. *Curr. Opin. Struct. Biol.* **2009**, *19*, 120−7.

(9) Case, D. A.; Cheatham Iii, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668−1688.

(10) Brooks, B. R.; Brooks, C. L., 3rd; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545−614.

(11) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701−18.

(12) Craig, P. O.; Latzer, J.; Weinkam, P.; Hoffman, R. M.; Ferreiro, D. U.; Komives, E. A.; Wolynes, P. G. Prediction of native-state hydrogen exchange from perfectly funneled energy landscapes. *J. Am. Chem. Soc.* **2011**, *133*, 17463−72.

(13) Turjanski, A. G.; Vaque, J. P.; Gutkind, J. S. MAP kinases and the control of nuclear events. *Oncogene* **2007**, *26*, 3240−53.

(14) Bellon, S.; Fitzgibbon, M. J.; Fox, T.; Hsiao, H. M.; Wilson, K. P. The structure of phosphorylated p38gamma is monomeric and reveals a conserved activation-loop conformation. *Structure* **1999**, *7*, 1057−1065.

(15) Goldsmith, E. J.; Cobb, M. H.; Chang, C. I. Structure of MAPKs. *Methods Mol. Biol. (N. Y., NY, U. S.)* **2004**, *250*, 127−144.

(16) White, A.; Pargellis, C. A.; Studts, J. M.; Werneburg, B. G.; Farmer, B. T., 2nd. Molecular basis of MAPK-activated protein kinase 2:p38 assembly. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 6353−8.
(17) Liu, S.; Sun, J. P.; Zhou, B.; Zhang, Z. Y. Structural basis of docking interactions between ERK2 and MAP kinase phosphatase 3. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 5326−5331.
(18) Shaw, D.; Wang, S. M.; Villasenor, A. G.; Tsing, S.; Walter, D.; Browner, M. F.; Barnett, J.; Kuglstatter, A. The crystal structure of JNK2 reveals conformational flexibility in the MAP kinase insert and indicates its involvement in the regulation of catalytic activity. *J. Mol. Biol.* **2008**, *383*, 885−93.
(19) Zhou, T.; Sun, L.; Humphreys, J.; Goldsmith, E. J. Docking interactions induce exposure of activation loop in the MAP kinase ERK2. *Structure* **2006**, *14*, 1011−1019.
(20) Hoofnagle, A. N.; Stoner, J. W.; Lee, T.; Eaton, S. S.; Ahn, N. G. Phosphorylation-dependent changes in structure and dynamics in ERK2 detected by SDSL and EPR. *Biophys. J.* **2004**, *86*, 395−403.
(21) Rodriguez Limardo, R. G.; Ferreiro, D. N.; Roitberg, A. E.; Marti, M. A.; Turjanski, A. G. p38gamma activation triggers dynamical changes in allosteric docking sites. *Biochemistry* **2011**, *50*, 1384−95.
(22) Tanoue, T.; Nishida, E. Molecular recognitions in the MAP kinase cascades. *Cell Signal* **2003**, *15*, 455−62.
(23) Hoofnagle, A. N.; Resing, K. A.; Goldsmith, E. J.; Ahn, N. G. Changes in protein conformational mobility upon activation of extracellular regulated protein kinase-2 as detected by hydrogen exchange. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 956−961.
(24) Lee, T.; Hoofnagle, A. N.; Kabuyama, Y.; Stroud, J.; Min, X.; Goldsmith, E. J.; Chen, L.; Resing, K. A.; Ahn, N. G. Docking motif interactions in MAP kinases revealed by hydrogen exchange mass spectrometry. *Mol. Cell* **2004**, *14*, 43−55.
(25) Lee, T.; Hoofnagle, A. N.; Resing, K. A.; Ahn, N. G. Hydrogen exchange solvent protection by an ATP analogue reveals conformational changes in ERK2 upon activation. *J. Mol. Biol.* **2005**, *353*, 600−612.
(26) Emrick, M. A.; Lee, T.; Starkey, P. J.; Mumby, M. C.; Resing, K. A.; Ahn, N. G. The gatekeeper residue controls autoactivation of ERK2 via a pathway of intramolecular connectivity. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 18101−6.
(27) Kornev, A. P.; Taylor, S. S. Defining the conserved internal architecture of a protein kinase. *Biochim. Biophys. Acta* **2010**, *1804*, 440−4.
(28) Zhang, F.; Strand, A.; Robbins, D.; Cobb, M. H.; Goldsmith, E. J. Atomic structure of the MAP kinase ERK2 at 2.3 A resolution. *Nature* **1994**, *367*, 704−711.
(29) Canagarajah, B. J.; Khokhlatchev, A.; Cobb, M. H.; Goldsmith, E. J. Activation mechanism of the MAP kinase ERK2 by dual phosphorylation. *Cell* **1997**, *90*, 859−869.
(30) Robinson, M. J.; Harkins, P. C.; Zhang, J.; Baer, R.; Haycock, J. W.; Cobb, M. H.; Goldsmith, E. J. Mutation of position 52 in ERK2 creates a nonproductive binding mode for adenosine 5′-triphosphate. *Biochemistry* **1996**, *35*, 5641−5646.
(31) Turjanski, A. G.; Hummer, G.; Gutkind, J. S. How mitogen-activated protein kinases recognize and phosphorylate their targets: A QM/MM study. *J. Am. Chem. Soc.* **2009**, *131*, 6141−8.
(32) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular-Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(33) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712−725.
(34) Meagher, K. L.; Redman, L. T.; Carlson, H. A. Development of polyphosphate parameters for use with the AMBER force field. *J. Comput. Chem.* **2003**, *24*, 1016−1025.
(35) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(36) *The PyMOL Molecular Graphics System*, version 1.3r1; Schrodinger, LLC: Cambridge, MA, 2010.