Do Collective Atomic Fluctuations Account for Cooperative Effects?
Molecular Dynamics Studies of the U1A−RNA Complex

Bethany L. Kormos, Anne M. Baranger,* and David L. Beveridge*

Chemistry Department and Molecular Biophysics Program, Wesleyan University, Middletown, Connecticut 06459

Received January 26, 2006; E-mail: dbeveridge@wesleyan.edu

The protein U1A is a component of the U1 snRNP, a subunit of the spliceosome, which splices most eukaryotic pre-mRNA.[1,2] U1A binds to RNA using an RNA recognition motif (RRM), which is one of the most common RNA binding domains.[3] Extensive structural determinations[4−7] and biophysical experiments[8−15,17−23] have been used to investigate RNA binding by the N-terminal RRM of U1A, which has become a paradigm for recognition by single RRMs. In this communication, we describe an analysis of correlated inter-residue fluctuations within the complex formed between U1A and stem loop 2 of U1 snRNA based on molecular dynamics (MD) simulations. The results support the hypothesis that collective atomic fluctuations are indicative of cooperative effects and thermodynamic coupling.

A series of recent studies[8−15] have provided biophysical evidence for three cooperative networks of amino acid residues in U1A that play a role in binding and specificity. Showalter and Hall reported a study of cooperativity in the U1A−RNA complex, including an analysis of couplings based on reorientational eigenmode dynamics.[23] Moreover, Crowder et al.[16] have reported statistical covariance analysis[24] of 330 aligned RRM sequences (including U1A) that revealed an extensive network of positional pairs, some not directly in contact with bound RNA, that may influence RNA binding affinity and specificity. These studies indicate that RNA recognition not only is a result of direct interactions but also involves long-range cooperative effects not readily identified by a visual inspection of the crystal structure. We have previously reported both experimental and molecular dynamics (MD) simulations to probe structural adaptations and the energetic origins of affinity and specificity in the U1A−stem loop 2 RNA complex.[11−13,21,22] We report herein theoretical calculations aimed at accounting for, understanding, and predicting cooperative effects from MD simulations.

A MD trajectory (10 ns) has been performed on the U1A−RNA complex,[4] including explicit consideration of water and counterions using the AMBER suite of programs[25] and the parm96 force field.[26] The calculation of cross-correlations between atomic fluctuations from MD simulations is based on methods described previously[27,28] and was used in this laboratory in early theoretical studies of HIV-1 protease.[28−30] The elements of the covariance matrix of atomic fluctuations (C_{ij}, defined in Supporting Information) are presented on a two-dimensional grid.[28,31] For ease of interpretation, the results are reduced to display correlations between amino acid or nucleotide subunits by averaging over backbone and all component atoms, respectively.

The covariance matrix for fluctuations as calculated from MD for the U1A−RNA complex is shown in Figure 1. The dispersion of the data is a reflection of the superposition of results from each of the individual thermally accessible structures that comprise the equilibrium Boltzmann distribution. A considerable amount of collective fluctuations are revealed by the map and correspond to both through-space as well as contact interactions.

The regions of the grid which correspond to amino acid residues and nucleotides that have been experimentally implicated in
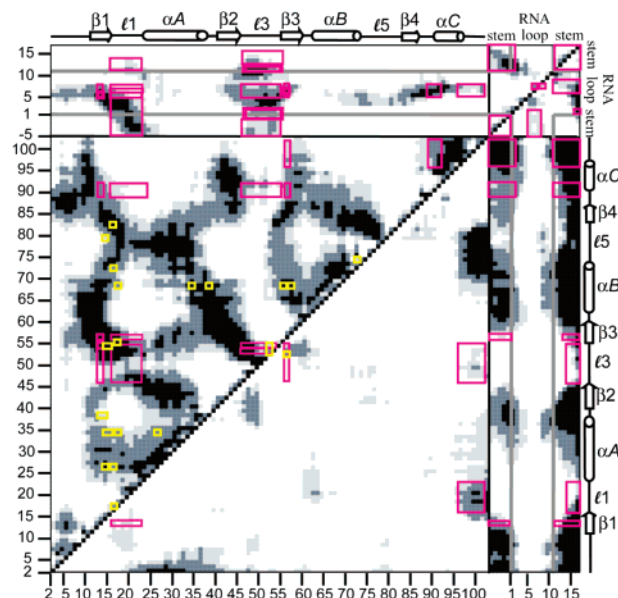

Figure 1. MD calculated collective atomic fluctuations for the U1A−RNA complex reduced to a residue by residue basis. Magnitudes of calculated correlations are indicated via a gray scale with black, slate gray and light gray regions corresponding to strong ($C_{ij} = \pm 0.75-1.00$), moderate ($C_{ij} = \pm 0.50-0.75$), and weak ($C_{ij} = \pm 0.25-0.50$) cross-correlations, respectively. Upper triangle: positive correlations. Lower triangle: negative correlations. Magenta boxes indicate regions of the grid corresponding to residues and nucleotides implicated by biophysical studies as being involved in cooperative networks.[8−15] Yellow boxes indicate positions of residues revealed to exhibit high covariance from the Crowder et al.[16] study of 330 RRMs.

cooperative networks of interactions are indicated in magenta-colored rectangular boxes. Considering both contact and through-space effects, the residues involved in strong collective atomic fluctuations correspond well to those observed to be a part of individual cooperative networks.[8−15,23] Beyond this, the MD results show cross-correlations predicting that the three observed networks are all components of one larger hyper-network. This reflects the overall dynamical integrity of the RRM and identifies correlated residues that have not yet been investigated by mutational studies.

The areas in Figure 1 corresponding to little or no collective fluctuations, indicated by white space on the map, are important to further clarification of the cooperative networks. For example, loop 3 has been shown to be involved in U1A−RNA cooperative networks; however, not all of the loop 3 residues reveal correlated fluctuations with all of the other residues. The Lys50−Gln54 residues of loop 3 are strongly correlated to Tyr13, while the Ser46−Leu49 and Met51 residues of loop 3 are strongly correlated with Gly53 and Gln54. These results and those in ref 23 provide an encouraging indication of the potential for using information on collective atomic fluctuations to map out and also predict cooperativity in protein−nucleic acid systems. An enlarged figure
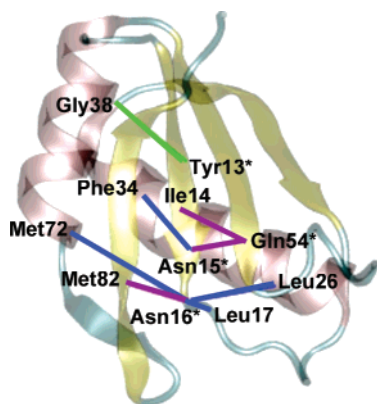
Figure 2. Cooperative interactions related to the U1A−RNA binding interface predicted from both MD calculated fluctuations and statistical covariance analysis for U1A. Residues directly in contact with RNA are indicated by an asterisk ∗. The magnitude of fluctuational covariance between residues corresponds to the line coloring: purple, $C_{ij} = \pm0.75-1.00$; blue, $C_{ij} = \pm0.50-0.75$; green, $C_{ij} = \pm0.25-0.50$.

and detailed account of the correlations are provided as Supporting Information.

To provide validation of the link between cooperative effects and cross-correlated positional fluctuations in a completely independent manner, we have compared our results to those of Crowder et al.,[16] who performed a statistical covariance analysis of 330 proteins containing RRMs and identified a network of covariant amino acid residues present within those RRMs. The Crowder et al. results revealed a number of positions of high covariance. We have indicated these positions for U1A in Figure 1 using yellow boxes for a comparison of the Crowder et al. results with those obtained in this study. There is a remarkable correspondence, with 96% of the calculated fluctuations in agreement with instances of high positional covariance. We find that 76% of the U1A residues that were identified in the Crowder et al. statistical covariance analysis exhibit fluctuational cross-correlations stronger than 0.50 in the MD results.

A preliminary consideration of the predictive aspect of this analysis is shown in Figure 2, based on a 3D structural representation of U1A with the extent of cross-correlated fluctuations between covarying residues indicated. This type of presentation has been used in the analysis of allosteric effects in proteins by Young et al.[32] Fluctuations of residues that covary with those directly in contact with RNA (indicated by an asterisk) are shown in Figure 2. While some of these have been examined experimentally, others such as Ile14 and Met82 have not yet been subjected to mutational analysis and are thus candidates for future experimental work. Note also that some but not all collective fluctuations indicated in Figure 1 appear as statistical covariance in the 330 RRMs, so that the predictive power of MD fluctuation analysis and the relationship between results specific to U1A and generic to RRMs remains to be fully clarified.

In summary, for the U1A−RNA complex, we have found that the results on collective atomic fluctuations from MD simulations agree well with experimentally observed cooperativities in U1A and the results of a statistical covariance analysis of 330 RRMs. The combination of MD and corresponding atomic fluctuation analysis is thus indicated to be a new and unique method for understanding and predicting cooperativity at the molecular level in protein−RNA binding. Implicit in our results is a link between MD simulation and structural bioinformatics. We have observed only a correspondence between collective atomic fluctuations and cooperativity, and a causal connection remains to be established.

However, the idea that thermodynamic coupling is linked to collective motions and cooperativity is quite plausible. This approach can be adapted to the study of cooperativity in protein−DNA, protein−protein, or any system on which MD can be successfully performed.

Supporting Information Available: Simulation details, cross-correlation calculations, MD calculated fluctuational covariance map with cooperative networks clearly labeled, and complete ref 25. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Green, M. R. *Annu. Rev. Cell Biol.* **1991**, *7*, 559−599.
(2) Stark, H.; Dube, P.; Lührmann, R.; Kastner, B. *Nature* **2001**, *409*, 539−542.
(3) Maris, C.; Dominguez, C.; Allain, F. H.-T. *FEBS Lett.* **2005**, *272*, 2118−2131.
(4) Oubridge, C.; Ito, N.; Evans, P. R.; Teo, C. H.; Nagai, K. *Nature* **1994**, *372*, 432−438.
(5) Allain, F. H.; Howe, P. W.; Neuhaus, D.; Varani, G. *EMBO J.* **1997**, *16*, 5764−5772.
(6) Mittermaier, A.; Varani, L.; Muhandiram, D. R.; Kay, L. E.; Varani, G. *J. Mol. Biol.* **1999**, *294*, 967−979.
(7) Varani, L.; Gunderson, S. I.; Mattaj, I. W.; Kay, L. E.; Neuhaus, D.; Varani, G. *Nat. Struct. Biol.* **2000**, *7*, 329−335.
(8) Kranz, J. K.; Hall, K. B. *J. Mol. Biol.* **1998**, *275*, 465−481.
(9) Kranz, J. K.; Hall, K. B. *J. Mol. Biol.* **1999**, *285*, 215−231.
(10) Showalter, S. A.; Hall, K. B. *J. Mol. Biol.* **2002**, *322*, 533−542.
(11) Nolan, S. J.; Shiels, J. C.; Tuite, J. B.; Cecere, K. L.; Baranger, A. M. *J. Am. Chem. Soc.* **1999**, *121*, 8951−8952.
(12) Shiels, J. C.; Tuite, J. B.; Nolan, S. J.; Baranger, A. M. *Nucleic Acids Res.* **2002**, *30*, 550−558.
(13) Tuite, J. B.; Shiels, J. C.; Baranger, A. M. *Nucleic Acids Res.* **2002**, *30*, 5269−5275.
(14) Katsamba, P. S.; Bayramyan, M.; Haworth, I. S.; Myszka, D. G.; Laird-Offringa, I. A. *J. Biol. Chem.* **2002**, *36*, 33267−33274.
(15) Law, M. J.; Chambers, E. J.; Katsamba, P. S.; Haworth, I. S.; Laird-Offringa, I. A. *Nucleic Acids Res.* **2005**, *33*, 2917−2928.
(16) Crowder, S.; Holton, J.; Alber, T. *J. Mol. Biol.* **2001**, *310*, 793−800.
(17) Scherly, D.; Kambach, C.; Boelens, W.; van Venrooij, W. J.; Mattaj, I. W. *J. Mol. Biol.* **1991**, *219*, 577−584.
(18) Jessen, T. H.; Oubridge, C.; Teo, C. H.; Pritchard, C.; Nagai, K. *EMBO J.* **1991**, *10*, 3447−3456.
(19) Hall, K. B.; Stump, W. T. *Nucleic Acids Res.* **1992**, *20*, 4283−4290.
(20) Rimmele, M. E.; Belasco, J. G. *RNA* **1998**, *4*, 1386−1396.
(21) Blakaj, D. M.; McConnell, K. J.; Beveridge, D. L.; Baranger, A. M. *J. Am. Chem. Soc.* **2001**, *123*, 2548−2551.
(22) Pitici, F.; Beveridge, D. L.; Baranger, A. M. *Biopolymers* **2002**, *65*, 424−435.
(23) Showalter, S. A.; Hall, K. B. *Biophys. J.* **2005**, *89*, 2046−2058.
(24) Lockless, S. W.; Ranganathan, R. *Science* **1999**, *286*, 295−299.
(25) Case, D. A.; et al. *AMBER 8*; University of California: San Francisco, 2004.
(26) Kollman, P. A.; Dixon, R.; Cornell, W.; Fox, T.; Chipot, C.; Pohorille, A. In *Computer Simulation of Biomolecular Systems*; Wilkinson, A., Weiner, P., van Gunsteren, W. F., Eds.; Kluwer: London, 1997; Vol. 3, pp 83−96.
(27) Ichiye, T.; Karplus, M. *Proteins* **1987**, *2*, 236−259.
(28) Swaminathan, S.; Harte, W. E.; Beveridge, D. L. *J. Am. Chem. Soc.* **1991**, *113*, 2717−2721.
(29) Harte, W. E., Jr.; Swaminathan, S.; Mansuri, M. M.; Martin, J. C.; Rosenberg, I. E.; Beveridge, D. L. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8864−8868.
(30) Harte, W. E., Jr.; Swaminathan, S.; Beveridge, D. L. *Proteins: Struct., Funct., Genet.* **1992**, *13*, 175−194.
(31) Ichiye, T.; Karplus, M. *Proteins: Struct., Funct., Genet.* **1991**, *11*, 205−217.
(32) Young, M. A.; Gonfloni, S.; Superti-Furga, G.; Roux, B.; Kuriyan, J. *Cell* **2001**, *105*, 115−126.

JA0606071