

Computational Design and Experimental Discovery of an Antiestrogenic Peptide Derived from α -Fetoprotein

Karl N. Kirschner,^{*,†} Katrina W. Lexa,[†] Amanda M. Salisburg,[†] Katherine A. Alser,[†] Leroy Joseph,[‡] Thomas T. Andersen,^{*,‡} James A. Bennett,[§] Herbert I. Jacobson,^{||} and George C. Shields^{*,†}

Contribution from the Hamilton College, Department of Chemistry, Center for Molecular Design, 198 College Hill Road, Clinton, New York 13323, Albany Medical College, Center for Cardiovascular Sciences, Albany, New York 12208, Albany Medical College, Center for Immunology and Microbial Disease, Albany, New York 12208, and Albany Medical College, Department of Obstetrics, Gynecology, and Reproductive Science, Albany, New York 12208

Received January 10, 2007; E-mail: gshields@hamilton.edu; kkirschn@hamilton.edu; anderst@mail.amc.edu

Abstract: Breast cancer is the most common cancer among women, and tamoxifen is the preferred drug for estrogen receptor-positive breast cancer treatment. Many of these cancers are intrinsically resistant to tamoxifen or acquire resistance during treatment. Consequently, there is an ongoing need for breast cancer drugs that have different molecular targets. Previous work has shown that 8-mer and cyclic 9-mer peptides inhibit breast cancer in mouse and rat models, interacting with an unsolved receptor, while peptides smaller than eight amino acids did not. We show that the use of replica exchange molecular dynamics predicts the structure and dynamics of active peptides, leading to the discovery of smaller peptides with full biological activity. Simulations identified smaller peptide analogues with the same conserved reverse turn demonstrated in the larger peptides. These analogues were synthesized and shown to inhibit estrogen-dependent cell growth in a mouse uterine growth assay, a test showing reliable correlation with human breast cancer inhibition.

Introduction

Breast cancer is the most common cancer diagnosed in women and is the second leading cause of cancer death among women, closely following lung cancer. In 2006, the American Cancer Society estimated that 212 920 women in the United States will be diagnosed with invasive breast cancer and predicted 40 970 deaths.¹ Tamoxifen is the most widely used drug for the treatment of estrogen receptor-positive breast cancer, acting through competition with estrogen for binding to the estrogen receptor (ER).

However, approximately one-third to one-sixth of ER-positive breast cancers are intrinsically resistant to tamoxifen, and many more acquire resistance to this drug during treatment.² Additionally, tamoxifen stimulates uterine growth, which can lead to uterine cancer in a small percentage of women taking this drug.² Consequently, there is an ongoing need for breast cancer drugs with greater efficacy and fewer side effects.

α -Fetoprotein (AFP) is an embryo specific serum α -globulin glycoprotein that is synthesized primarily by the fetal liver and

circulates through the serum of pregnant women.³ AFP has been reported to bind and transport ligands, including fatty acids, steroids, heavy metal ions, phytoestrogens, drugs, and some toxins.⁴ AFP is a growth-regulating hormone with the capacity to stimulate or inhibit growth.⁵ From fertilization through birth, AFP holds an important physiological role as a developmental promoter for the fetus. More recent reports have shown that AFP has antiestrogenic activity and can inhibit the growth of estrogen dependent cancer.^{2,6–8} These data, combined with epidemiological data showing that elevated levels of AFP are associated with a reduced lifetime risk of breast cancer,⁷ have led to the suggestion that AFP or analogues thereof may be a useful agent for chemoprevention, as well as for the treatment, of breast cancer.^{7,9,10} Festin et al.,¹¹ Eisele et al.,^{12,13} and Mesfin

- (3) Abelev, G. I. *Adv. Cancer Res.* **1971**, *14*, 295–358.
- (4) Mizejewski, G. J. *Exp. Biol. Med.* **2001**, *226*, 377–408.
- (5) Dudich, E.; Semenkova, L.; Gorbatoeva, E.; Dudich, I.; Khromykh, L.; Tatulov, E.; Grechko, G.; Sukhikh, G. *Tumor Biol.* **1998**, *19*, 30–40.
- (6) Jacobson, H. I.; Bennett, J. A.; Mizejewski, G. J. *Cancer Res.* **1990**, *50*, 415–420.
- (7) Bennett, J. A.; Zhu, S. J.; Pagano-Mirarchi, A.; Kellom, T. A.; Jacobson, H. I. *Clin. Cancer Res.* **1998**, *4*, 2877–2884.
- (8) Parikh, R. R.; Gildener-Leapman, N.; Narendran, A.; Lin, H. Y.; Lemanski, N.; Bennett, J. A.; Jacobson, H. I.; Andersen, T. T. *Clin. Cancer Res.* **2005**, *11*, 8512–8520.
- (9) Jacobson, H. I.; Janerich, D. T. *Biological Activities of Alpha-Fetoprotein*; CRC: Boca Raton, FL, 1989.
- (10) Richardson, B. E.; Hulka, B. S.; David-Peck, J. L.; Hughes, C. L.; Van den Berg, B. J.; Christianson, R. E.; Calvin, J. A. *Am. J. Epidemiol.* **1998**, *148*, 719–727.
- (11) Festin, S. M.; Bennett, J. A.; Fletcher, P. W.; Jacobson, H. I.; Shaye, D. D.; Andersen, T. T. *Biochim. Biophys. Acta* **1999**, *1427*, 307–314.

[†] Hamilton College.

[‡] Albany Medical College, Center for Cardiovascular Sciences.

[§] Albany Medical College, Center for Immunology and Microbial Disease.

^{||} Albany Medical College, Department of Obstetrics, Gynecology, and Reproductive Science.

(1) American Cancer Society: Atlanta, 2006; Vol. 2006, American Cancer Society Website, http://www.cancer.org/docroot/MED/content/MED_1_1_Most-Requested_Graphs_and_Figures_2006.asp.

(2) Bennett, J. A.; Mesfin, F. B.; Andersen, T. T.; Gierthy, J. F.; Jacobson, H. I. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2211–2215.

et al.¹⁴ parsed the 70 000 MW AFP into a series of smaller peptides that retained the same antiestrogenic and antibreast cancer activities of the original protein. Their work resulted in the smallest known active analogue, an 8-mer with the sequence EMTVPNPG (AFPep),¹⁴ which consists of amino acids 472–479 from the human AFP sequence. Amino acid substitution studies revealed the importance of specific residues essential for activity.¹⁵ All efforts since then to create an active peptide consisting of fewer than eight residues has resulted in the loss of antiestrogenic activity.^{14,15}

A receptor for AFP on cancer cell membranes has been reported,¹⁶ and binding of AFP (or AFPep) to this receptor signals the cell to inhibit its own growth. The steric/electronic features of the receptor site that permits AFP binding are largely unknown, making rational development of lead compounds difficult. We present here a novel strategy for developing new lead compounds, a strategy that uses molecular dynamics to explore the allowed conformational space of potentially active analogues in solution. Rational drug design using molecular dynamics in this manner involves understanding the conformational space occupied by the active compounds followed by the creation of different compounds that sample the same space. We have used Replica Exchange^{17,18} Molecular Dynamics (REMD) techniques¹⁹ to explore the conformational dynamics of several AFPep analogues. Our REMD results reveal that the peptide's critical region for activity is a four amino acid sequence that adopts a Type I β -turn conformation. We have run REMD simulations on several different four and five amino acid peptides, synthesized those that appeared promising, along with controls, and tested them for activity using an immature mouse uterine growth assay.² Results from the REMD simulations and the experimental activity studies are presented in this paper and show that peptides as short as four amino acid residues retain biological activity.

The REMD technique has been successfully used to obtain energy landscapes for the TRP cage and for the C-terminal β -hairpin of protein G,^{20–22} to find the global minimum for chignolin²³ and to explore unfolding of α -helical peptides as a function of pressure and temperature.²⁴ REMD has also been used to explore intrafacial folding and membrane insertion of designed peptides²⁵ as well as simulations of DNA.^{26,27} Recent

work has focused on use of REMD to determine peptide structure and peptide and protein folding pathways.^{28–85} We

- (12) Eisele, L. E.; Mesfin, F. B.; Bennett, J. A.; Andersen, T. T.; Jacobson, H. I.; Vakharia, D. D.; MacColl, R.; Mizejewski, G. J. *J. Pept. Res.* **2001**, *57*, 539–546.
- (13) Eisele, L. E.; Mesfin, F. B.; Bennett, J. A.; Andersen, T. T.; Jacobson, H. I.; Soldwedel, H.; MacColl, R.; Mizejewski, G. J. *J. Pept. Res.* **2001**, *57*, 29–38.
- (14) Mesfin, F. B.; Bennett, J. A.; Jacobson, H. I.; Zhu, S. J.; Andersen, T. T. *Biochim. Biophys. Acta (Molecular Basis of Disease)* **2000**, *1501*, 33–43.
- (15) DeFreest, L. A.; Mesfin, F. B.; Joseph, L.; McLeod, D. J.; Stallmer, A.; Reddy, S.; Balulad, S. S.; Jacobson, H. I.; Andersen, T. T.; Bennett, J. A. *J. Pept. Res.* **2004**, *63*, 409–419.
- (16) Villacampa, M. J.; Moro, R.; Naval, J.; Failycrepin, C.; Lampreave, F.; Uriel, J. *Biochem. Biophys. Res. Commun.* **1984**, *122*, 1322–1327.
- (17) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141–151.
- (18) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *Biopolymers* **2001**, *60*, 96–123.
- (19) Case, D. A., et al. *AMBER Version 8 Molecular Dynamics Program*; University of California: San Francisco, CA, 2004.
- (20) Zhou, R. H.; Berne, B. J.; Germain, R. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14931–14936.
- (21) Zhou, R. H. *Proteins: Struct., Funct., Genet.* **2003**, *53*, 148–161.
- (22) Zhou, R. H. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13280–13285.
- (23) Seibert, M. M.; Patriksson, A.; Hess, B.; van der Spoel, D. *J. Mol. Biol.* **2005**, *354*, 173–183.
- (24) Paschek, D.; Gnanakaran, S.; Garcia, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6765–6770.
- (25) Im, W.; Brooks, C. L. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6771–6776.
- (26) Murata, K.; Sugita, Y.; Okamoto, Y. *J. Theor. Comput. Chem.* **2005**, *4*, 411–432.
- (27) Murata, K.; Sugita, Y.; Okamoto, Y. *J. Theor. Comput. Chem.* **2005**, *4*, 433–448.
- (28) Nymeyer, H.; Gnanakaran, S.; Garcia, A. E. *Numerical Computer Methods, Pr D*; 2004; Vol. 383, p 119.
- (29) Andreac, M.; Felts, A. K.; Gallicchio, E.; Levy, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6801–6806.
- (30) Bernard, D.; Coop, A.; Mackerell, A. D. *J. Med. Chem.* **2005**, *48*, 7773–7780.
- (31) Malolepsza, E.; Boniecki, M.; Kolinski, A.; Piela, L. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7835–7840.
- (32) Best, R. B.; Clarke, J.; Karplus, M. *J. Mol. Biol.* **2005**, *349*, 185–203.
- (33) Gnanakaran, S.; Garcia, A. E. *Proteins: Struct., Funct., Bioinf.* **2005**, *59*, 773–782.
- (34) Nymeyer, H.; Woolf, T. B.; Garcia, A. E. *Proteins: Struct., Funct., Bioinf.* **2005**, *59*, 783–790.
- (35) Yoshida, K.; Yamaguchi, T.; Okamoto, Y. *Chem. Phys. Lett.* **2005**, *412*, 280–284.
- (36) Baumketner, A.; Shea, J. E. *Biophys. J.* **2005**, *89*, 1493–1503.
- (37) Nguyen, P. H.; Mu, Y. G.; Stock, G. *Proteins: Struct., Funct., Bioinf.* **2005**, *60*, 485–494.
- (38) Snow, C. D.; Sorin, E. J.; Rhee, Y. M.; Pande, V. S. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 43–69.
- (39) Zhang, W.; Wu, C.; Duan, Y. *J. Chem. Phys.* **2005**, 123.
- (40) Ravindranathan, K. P.; Gallicchio, E.; Levy, R. M. *J. Mol. Biol.* **2005**, *353*, 196–210.
- (41) Nishino, M.; Sugita, Y.; Yoda, T.; Okamoto, Y. *FEBS Lett.* **2005**, *579*, 5425–5429.
- (42) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13749–13754.
- (43) Roe, D. R.; Hornak, V.; Simmerling, C. *J. Mol. Biol.* **2005**, *352*, 370–381.
- (44) Nguyen, P. H.; Stock, G.; Mittag, E.; Hu, C. K.; Li, M. S. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 795–808.
- (45) Chen, Z.; Xu, Y. *Proteins: Struct., Funct., Bioinf.* **2006**, *62*, 539–552.
- (46) Wang, Z. X.; Zhang, W.; Wu, C.; Lei, H. X.; Cieplak, P.; Duan, Y. *J. Comput. Chem.* **2006**, *27*, 781–790.
- (47) Chen, J. H.; Im, W. P.; Brooks, C. L. *J. Am. Chem. Soc.* **2006**, *128*, 3728–3736.
- (48) Affentranger, R.; Tavernelli, I.; Di Iorio, E. E. *J. Chem. Theory Comput.* **2006**, *2*, 217–228.
- (49) Okur, A.; Wickstrom, L.; Layten, M.; Geney, R.; Song, K.; Hornak, V.; Simmerling, C. *J. Chem. Theory Comput.* **2006**, *2*, 420–433.
- (50) Baumketner, A.; Bernstein, S. L.; Wyttenbach, T.; Bitan, G.; Teplow, D. B.; Bowers, M. T.; Shea, J. E. *Protein Sci.* **2006**, *15*, 420–428.
- (51) Bunagan, M. R.; Yang, X.; Saven, J. G.; Gai, F. *J. Phys. Chem. B* **2006**, *110*, 3759–3763.
- (52) Jang, S.; Shin, S. *J. Phys. Chem. B* **2006**, *110*, 1955–1958.
- (53) Jang, S.; Kim, E.; Pak, Y. *Proteins: Struct., Funct., Bioinf.* **2006**, *62*, 663–671.
- (54) Zhang, J.; Qin, M.; Wang, W. *Proteins: Struct., Funct., Bioinf.* **2006**, *62*, 672–685.
- (55) Ravindranathan, K. P.; Gallicchio, E.; Friesner, R. A.; McDermott, A. E.; Levy, R. M. *J. Am. Chem. Soc.* **2006**, *128*, 5786–5791.
- (56) Trebst, S.; Troyer, M.; Hansmann, U. H. E. *J. Chem. Phys.* **2006**, 124.
- (57) Larios, E.; Pitera, J. W.; Swope, W. C.; Grubele, M. *Chem. Phys.* **2006**, *323*, 45–53.
- (58) Soto, P.; Baumketner, A.; Shea, J. E. *J. Chem. Phys.* **2006**, 124.
- (59) Yadav, M. K.; Leman, L. J.; Price, D. J.; Brooks, C. L.; Stout, C. D.; Ghadiri, M. R. *Biochemistry* **2006**, *45*, 4463–4473.
- (60) Kmiecik, S.; Kurcinski, M.; Rutkowska, A.; Gront, D.; Kolinski, A. *Acta Biochim. Pol.* **2006**, *53*, 131–143.
- (61) van der Spoel, D.; Seibert, M. M. *Phys. Rev. Lett.* **2006**, 96.
- (62) Baumketner, A.; Bernstein, S. L.; Wyttenbach, T.; Lazo, N. D.; Teplow, D. B.; Bowers, M. T.; Shea, J. E. *Protein Sci.* **2006**, *15*, 1239–1247.
- (63) Naniias, M.; Czaplewski, C.; Scheraga, H. A. *J. Chem. Theory Comput.* **2006**, *2*, 513–528.
- (64) Mu, Y. G.; Nordenskiold, L.; Tam, J. P. *Biophys. J.* **2006**, *90*, 3983–3992.
- (65) Wei, G. H.; Shea, J. E. *Biophys. J.* **2006**, *91*, 1638–1647.
- (66) Wickstrom, L.; Okur, A.; Song, K.; Hornak, V.; Raleigh, D. P.; Simmerling, C. L. *J. Mol. Biol.* **2006**, *360*, 1094–1107.
- (67) Ho, B. K.; Dill, K. A. *PLoS Comp. Biol.* **2006**, *2*, 228–237.
- (68) Li, P. C.; Huang, L.; Makarov, D. E. *J. Phys. Chem. B* **2006**, *110*, 14469–14474.
- (69) Wang, J.; Gu, Y.; Liu, H. Y. *J. Chem. Phys.* **2006**, 125.
- (70) Hughes, R. M.; Waters, M. L. *Curr. Opin. Struct. Biol.* **2006**, *16*, 514–524.
- (71) Baumketner, A.; Shea, J. E. *Theor. Chem. Acc.* **2006**, *116*, 262–273.
- (72) Daura, X. *Theor. Chem. Acc.* **2006**, *116*, 297–306.
- (73) Kokubo, H.; Okamoto, Y. *Mol. Simul.* **2006**, *32*, 791–801.
- (74) Grater, F.; de Groot, B. L.; Jiang, H. L.; Grubmuller, H. *Structure* **2006**, *14*, 1567–1576.
- (75) Li, H. Z.; Li, G. H.; Berg, B. A.; Yang, W. *J. Chem. Phys.* **2006**, 125.
- (76) Baumketner, A.; Shea, J. E. *J. Mol. Biol.* **2006**, *362*, 567–579.
- (77) Liu, P.; Huang, X. H.; Zhou, R. H.; Berne, B. J. *J. Phys. Chem. B* **2006**, *110*, 19018–19022.
- (78) Sugita, Y.; Miyashita, N.; Yoda, T.; Ikeguchi, M.; Toyoshima, C. *Biochemistry* **2006**, *45*, 11752–11761.

demonstrate for the first time that REMD, besides being an excellent technique for probing peptide and protein structure, can be used for a lead compound design.

Methods

We used the AMBER 8 molecular dynamics program package^{19,86} with a new force field that has been created specifically to improve the representation of peptide structure.⁸⁷ Simulations were run in implicit solvent, using the Generalized Born (GB) model⁸⁸ implemented in AMBER⁸⁹ with the default radii under the *multisander* framework.⁸⁹ It has been demonstrated that the use of continuum solvent models in REMD can lead to overstabilization of ion pairs that affect the secondary structure,^{90,91} but since we do not have salt bridges in our peptide structure we avoid this potential problem. Adequate sampling of conformational space was insured through the use of the Replica Exchange algorithm.^{17,18} REMD was developed as a means to overcome local potential energy barriers. Compiled under AMBER 8, one-dimensional REMD explores a generalized canonical ensemble of *N* noninteracting replicas that undergo simulation separately but concomitantly at exponentially related temperatures, with exchanges between replicas occurring at a specified time interval. The consequence of this exchange is that entrapment in local potential energy wells is avoided. Those replicas that do become trapped within a local well at one temperature can escape when transitioned to a higher temperature as part of the exchange process. Thus, accuracy of conformation is maintained through analysis at low temperatures, while simulations at higher temperatures efficiently achieve exploration of the potential energy surface.

The sequences used for the simulations were chosen from the set of previously synthesized active analogues.¹⁵ REMD simulations were run on the cyclic analogues cyclic-[EKTPVNPNG], cyclic-[EKTPVNPNGQ], cyclic-[EMTPVNPNGQ], and the linear analogues EMTPVNPNG and EMTPVNPNG. In addition, REMD simulations were run on the smaller analogues EMTPVNP, MTPVNP, TPVNP, TPVN, and PVNP. All sequences were capped using an acetyl beginning residue and an *N*-methylamine ending residue. Eight different replicas were used, each defined initially with the same input structure. Temperatures were selected to agree with exponential growth such that interchange occurred within the temperature group temp, temp₀ = 265, 304, 350, 402, 462, 531, 610, and 700 K. The same temperatures were used for each peptide simulated and controlled using a weak-coupling algorithm as specified by ntt = 1. The number of exchange attempts between neighboring replicas was set to 1000, and the number of MD steps between exchange attempts was defined as 10 000. Thus, the total length for each simulation with a time step of 0.002 ps was 20 ns.¹⁹ Additional methodological information can be found in a forthcoming paper.⁹²

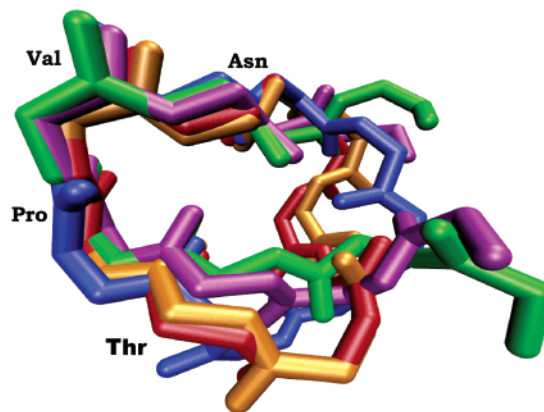


Figure 1. Overlay of the cyclic-[EKTPVNPNG] (red), cyclic-[EKTPVNPNGQ] (blue), cyclic-[EMTPVNPNGQ] (orange), and the EMTPVNPNG (green) and EMTPVNPNG (purple) peptides from REMD simulations. Each structure represents the β -turn motif sampled during the dynamics.

The linear peptide analogues shown in Table 2 were prepared using Fmoc solid-phase synthesis.^{14,15,93} The antiestrogenic activity of each peptide was determined using the immature mouse uterine growth assay as described previously.^{2,7} Intraperitoneal (i.p.) administration of 0.5 μ g of 17 β -estradiol (E₂) to an immature female mouse doubles its uterine weight in 24 h.⁷ Swiss/Webster female mice, weighing 6–8 grams at 13–15 days old, were obtained from Taconic Farms. Mice were grouped so that each group had the same range of body weights. Each group received two sequential i.p. injections spaced 1 h apart. The peptide or a saline solution control was contained in the first injectant, and E₂ or a saline control was contained in the second injectant. 22 h after the second injection, uteri were dissected and weighed immediately. The uterine weights were normalized to mouse body weights (mg of uterine/g of body). Experiments used a minimum of five mice per group, and the mean normalized uterine weight and standard deviations were determined for each group. Percent growth inhibition in a test group was calculated from the normalized uterine weights as given by eq 1.

$$\text{Growth inhibition (\%)} = \frac{(\text{Full E}_2 \text{ stimulation} - \text{E}_2 \text{ stimulation in test group})}{(\text{Full E}_2 \text{ stimulation} - \text{No E}_2 \text{ stimulation})} \times 100\% \quad (1)$$

The significance of differences between groups was evaluated with the nonparametric Wilcoxon ranks sum test (one-sided). Generally, drug-induced growth inhibitions of 20% or greater are significantly different ($p \leq 0.05$) from the group receiving E₂ alone. Each AFPep analogue was evaluated for antiuterotrophic activity in three or more experiments, and the mean growth inhibition \pm the standard error for each analogue is reported in Table 2.

Results

Figure 1 shows the conservation of conformational space of the five active 8-mer and cyclic 9-mer AFP-derived peptides previously shown^{14,15} to have antiestrogenic activity. This figure shows the overlay of the most representative peptide geometries obtained from the conformational family that displays the common reverse turn motif in each REMD simulation. Four amino acids, TPVN, are conserved in the conformational space of the active 8-mers and cyclic 9-mers. The TPVN sequence forms a reverse turn within the longer peptides, a structure that is conserved across all five REMD simulations. In a Type I reverse β -turn, four amino acids form a turn structure, which is

- (79) Kameda, T.; Takada, S. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 17765–17770.
 (80) Lou, H. F.; Cukier, R. I. *J. Phys. Chem. B* **2006**, *110*, 24121–24137.
 (81) Lwin, T. Z.; Luo, R. *Protein Sci.* **2006**, *15*, 2642–2655.
 (82) Periole, X.; Mark, A. E. *J. Chem. Phys.* **2007**, *126*.
 (83) Bu, L. T.; Im, W.; Charles, L. I. *Biophys. J.* **2007**, *92*, 854–863.
 (84) Liwo, A.; Khalili, M.; Czaplewski, C.; Kalinowski, S.; Oldziej, S.; Wachucik, K.; Scheraga, H. A. *J. Phys. Chem. B* **2007**, *111*, 260–285.
 (85) Jang, S.; Kim, E.; Pak, Y. *Proteins: Struct., Funct., Bioinf.* **2007**, *66*, 53–60.
 (86) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668–1688.
 (87) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 712–725.
 (88) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*, 2nd ed.; John Wiley & Sons Ltd: Chichester, U.K., 2004.
 (89) Onufriev, A.; Bashford, D.; Case, D. A. *J. Phys. Chem. B* **2000**, *104*, 3712–3720.
 (90) Zhou, R. H.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12777–12782.
 (91) Geney, R.; Layten, M.; Gomperts, R.; Hornak, V.; Simmerling, C. *J. Chem. Theory Comput.* **2006**, *2*, 115–127.
 (92) Lexa, K. W.; Alser, K. A.; Salisburg, A. M.; Ellens, D. J.; Hernandez, L.; Bono, S. J.; Derby, J. R.; Skiba, J. G.; Feldgus, S.; Kirschner, K. N.; Shields, G. C. *Int. J. Quantum Chem.*, submitted.

- (93) Mesfin, F. B.; Andersen, T. T.; Jacobson, H. I.; Zhu, S.; Bennett, J. A. *J. Pept. Res.* **2001**, *58*, 246–256.

Table 1. Average ϕ/ψ Angle Values for the Second and Third Amino Acids for Five Different Sequences^a

| | | ϕ_2 | ψ_2 | ϕ_3 | ψ_3 |
|-----------------------|-----|----------|----------|----------|----------|
| type I β -turn | | -60 | -30 | -90 | 0 |
| cyclic-[EMTPVNPGQ] | | -68 | -16 | -96 | -9 |
| EMTPVNPG ^b | | -68 | -17 | -94 | -11 |
| TPVNP | | -65 | -19 | -95 | -12 |
| TPVN | | -68 | -19 | -88 | -11 |
| PVNP | 68% | -89 | -8 | -110 | 115 |
| | 27% | -108 | 141 | -104 | 128 |

^a The four residues in boldface define amino acids one through four for each peptide. The ϕ/ψ angle values for the second and third amino acids of a four amino acid sequence are diagnostic for a β -turn structure. ^b The sequence of the AFPep peptide.

defined by the phi (ϕ) and psi (ψ) angles of the second and third amino acids. An ideal Type I β -turn has ϕ/ψ values of -60° and -30° for the second amino acid and ϕ/ψ values of -90° and 0° for the third amino acid in the four amino acid turn sequence. As displayed in Table 1, the cyclic-[EMTPVNPGQ] and EMTPVNPG peptides have average $\phi_{\text{Pro}}/\psi_{\text{Pro}}$ and $\phi_{\text{Val}}/\psi_{\text{Val}}$ values that fall within 13° of a Type I turn. The other three simulations of the 8-mer and 9-mer peptides have similar values (Figure 1).

We ran REMD simulations for the TPVN, TPVNP, and PVNP analogues to explore the conformational space sampled by these four and five amino acid peptides. The TPVN and TPVNP peptides form the same reverse turn seen in the larger, active peptides. As displayed in Table 1, the TPVNP and TPVN analogues have $\phi_{\text{Pro}}/\psi_{\text{Pro}}$ values of approximately -67° and -19° , while the $\phi_{\text{Val}}/\psi_{\text{Val}}$ values are approximately -91° and -12° . In contrast, the average $\phi_{\text{Val}}/\psi_{\text{Val}}$, and $\phi_{\text{Asn}}/\psi_{\text{Asn}}$ values for the two dominant conformations of the PVNP peptide reveal that none of the three structures is a turn (Table 1). The structure that is sampled for 68% of the REMD simulations has $\phi_{\text{Val}}/\psi_{\text{Val}}$ angles of -89° and -8° and $\phi_{\text{Asn}}/\psi_{\text{Asn}}$ angles of -110° and 115° . The structure sampled for 27% of the REMD simulation has $\phi_{\text{Val}}/\psi_{\text{Val}}$ angles of -108° and 141° and $\phi_{\text{Asn}}/\psi_{\text{Asn}}$ angles of -104° and 128° . Consideration of the average asparagine angles alone shows that the PVNP peptide does not form a turn structure.

The representative dynamics of three analogues containing the conserved TPVN sequence can be visualized in Figure 2. The top graphs display the distance between the C_α atoms of the threonine and asparagine residues, serving as a definitive diagnostic for a β -turn. Twenty-five percent of β -turns do not have an intraturn hydrogen bond, so an alternative definition of a β -turn is that the distance between the C_α carbon atoms of amino acid residues one and four in the tetrapeptide sequence is less than 7 \AA .⁹⁴ The TPVN tetrapeptide has a $C_{\alpha(\text{T})} - C_{\alpha(\text{N})}$ distance less than 7 \AA for 64% of the simulation, indicating a β -turn. The TPVNP pentapeptide adopts the turn structure for 74% of the simulation, and the cyclic-[EMTPVNPGQ] peptide is locked into the turn structure for 99% of the simulation. The bottom graphs in Figure 2 show the corresponding three-dimensional plots of ϕ versus ψ . These plots reveal the dynamics of the ϕ/ψ values for proline (red) and valine (green) amino acids throughout the simulation for these three peptides. These plots confirm that these three peptides adopt a β -turn conformation over the course of the simulation.

(94) Chou, K.-C. *Anal. Biochem.* **2000**, *286*, 1–16.

Several 4-mer and 5-mer peptide analogues containing the TPVN sequence, or a similar sequence with hydroxyproline (O) substituted for proline, were synthesized, tested, and compared to the original 8-mer peptides for biological activity. As shown in Table 2, biological activity, as defined by inhibition of estrogen-stimulated growth of an immature mouse uterus, was retained in the TPVN and TOVN 4-mers and even more so in the TPVNP and TOVNO 5-mers.

The OVNO and PVNP analogues, which were thought to represent the pharmacophore region of AFPep,¹⁵ did not show significant activity. Similarly, five amino acid peptides containing the amino or carboxyl end of AFPep did not have biological activity (bottom of Table 2). The 7-mer, EMTPVN, was slightly less active than AFPep and the smaller TPVNP.

The figures reveal why the TPVNP and TPVN analogues are active. All active peptides have a conserved reverse β -turn motif. These β -turns are formed by the TPVN sequence, with a hydrogen bond formed between the carbonyl oxygen of the first amino acid and the amide hydrogen of the fourth amino acid. The conservation of proline in the second position favors the formation of a reverse turn. This proline is conserved in human, gorilla, chimpanzee, horse, rat, and mouse AFP sequences.¹⁵

Discussion

We have used REMD simulations to sample the conformational space of 8-mer and 9-mer AFP-derived peptides that have antiestrogenic and antibreast cancer activity. We discovered that an identical four amino acid sequence had minimal conformational flexibility, suggesting that this region is essential for the biological activities of these peptides. The TPVN, TOVN, TPVNP, and TOVNO sequences were subsequently synthesized and were found to be biologically active. This is a novel finding because Mesfin et al.¹¹ and DeFreest et al.¹⁵ had concluded that the 8-mer, EMTPVNPG, was the minimum number of residues that an AFP-derived peptide can have and still retain significant antiestrogenic activity. This conclusion was based on the findings that the 7-mers, MTPVNPG and EMTPVN, had relatively less biological activity than the 8-mers. These investigators concluded that the 8-mer peptide assumed a pseudo-cyclic conformation, that the middle of the peptide contained its pharmacophore region (the PVNP sequence), and that the exterior residues were essential for holding the peptide in its active conformation. The 7-mers lost activity, which was assumed to result from the loss of hydrogen bonding between the first and eighth amino acid residues, suggesting that full activity must require at least eight amino acid residues. Indeed the 5-mer, EMTPV, was synthesized and found to have no significant activity, which supported their conclusion that the 8-mer sequence was the minimal sequence for biological activity.¹⁴ Furthermore, guided by this hydrogen bond hypothesis, cyclic 9-mers were also synthesized and shown to possess significant biological activity.⁹³ We tested this pseudocyclic conformation hypothesis by examining the percentage of hydrogen bond formation between the first and eighth amino acids for each simulation containing seven or eight amino acid residues. For all 8-mer and 7-mer simulations, hydrogen bond formation between the first and the last amino acids is observed for less than 1% of the simulation if the distance definition for a hydrogen bond is set at 2.3 \AA . This result is contrary to the original hypothesis.

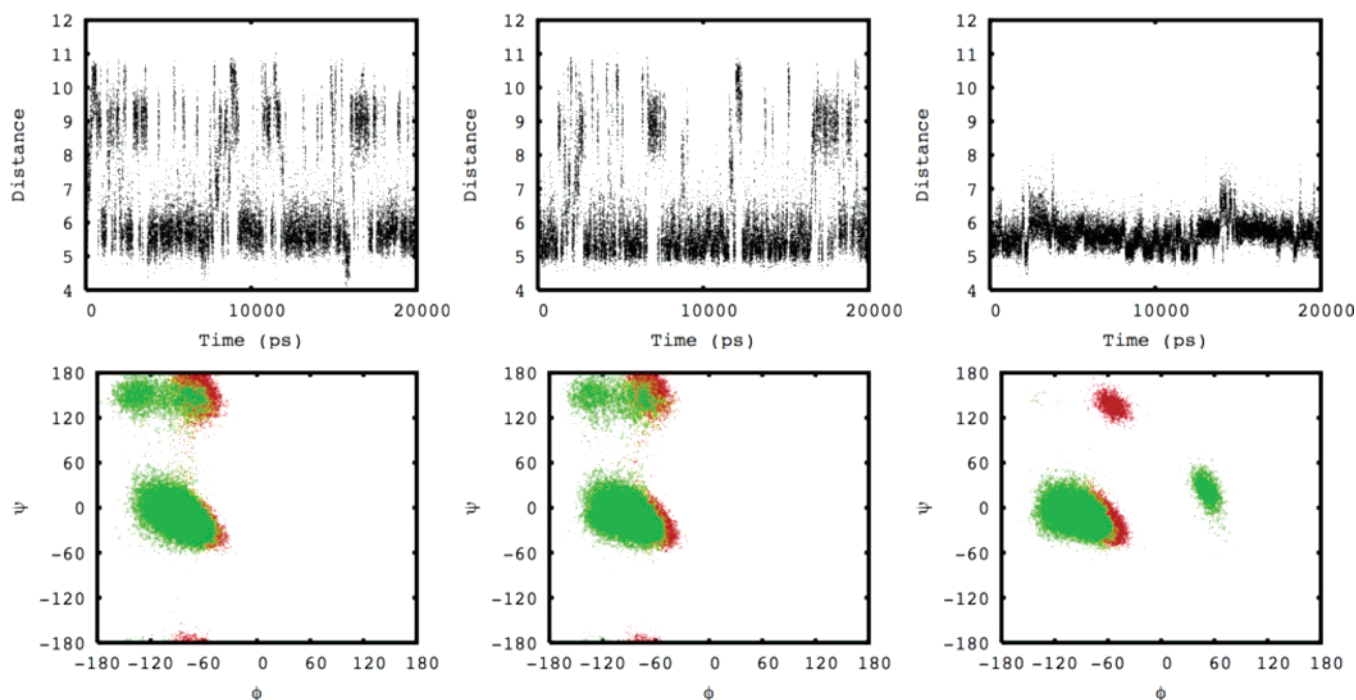


Figure 2. The top graphs depict the distances between the conserved threonine and asparagine C_{α} atoms as a function of simulation time for the TPVN (left), TPVNP (middle), and cyclic-[EMTPVNPQ] (right) peptides. The bottom graphs depict their corresponding ϕ (x -axis) and ψ (y -axis) angles as a function of simulation time for the conserved proline (red) and valine (green) amino acids.

Table 2. Effect of AFP-derived Peptides on Estrogen-stimulated Growth of Immature Mouse Uterus^a

| sequence | % inhibition \pm SE |
|----------|-------------------------|
| EMTPVNPQ | 34 \pm 3 ^b |
| EMTOVNOG | 32 \pm 4 ^b |
| EMTPVNP | 19 \pm 2 |
| TOVNO | 31 \pm 4 ^b |
| TPVNP | 26 \pm 1 ^b |
| TPVN | 24 \pm 5 ^b |
| TOVN | 22 \pm 1 ^b |
| OVNO | 6 \pm 5 |
| PVNP | 5 \pm 3 |
| PGVGQ | 0 |
| EMTPV | 0 |
| EMTOV | 0 |

^a Estradiol and peptide injection into immature mice was carried out as described in Methods. Peptide was injected intraperitoneally at a dose of 1 μ g per mouse. ^b $p < 0.05$ when compared to group stimulated with E_2 alone. Wilcoxon Rank-Sum Test.

The REMD studies reported herein lead to a substantially different conclusion regarding the requirement for peptide activity. These studies indicate that the key region is the TPVN sequence, since this structure retains the turn conformation common in every active peptide; the PVNP peptide does not sample a reverse turn structure (Table 1) and shows insignificant activity (Table 2). The above discovery has profound consequences. First, a 4-mer or a 5-mer peptide that retains biological activity is less expensive to synthesize than an 8-mer or a 9-mer peptide, and these smaller analogues are more druglike. Second, knowing the conformational space occupied by the 4-mer provides insight into the topology of the unknown receptor site for these peptides. Based on this study, the receptor site topology for the AFP analogue peptides is predicted to be a mirror image of a reverse β -turn. We note that the β -turn solution conformation of these peptides may not be retained when bound in the active site. However the high correlation between activity and β -turn conformation coupled with the entropic cost for altering

the conformation upon binding gives us confidence that the receptor topology will accommodate a β -turn conformation. Previous substitutions in the 8-mer sequences reveal that when threonine, leucine, or isoleucine is substituted for valine, biological activity is retained, while substitution of valine with D-valine or alanine results in loss of activity. This implies that the topology of the receptor is stereospecific, and branched amino acids are essential for creation of hydrophobic forces that bind the receptor to the peptide. Finally, it leads to a different conceptual approach for stabilization of these peptides and the development of peptidomimetics. Peptidomimetics can be designed based on the active 4-mer and 5-mer peptides, with REMD used to ensure that the steric and electronic nature of the peptides is retained. Developing a peptidomimetic may not be necessary, for as long as the peptides are bioavailable they have the advantage of a lower probability of side effects compared to peptide mimics. The AFP 8-mer and cyclic 9-mer peptides have already been shown to be nontoxic and bioavailable in mice.² Cancer xenograft assays in mice involved the administration of 8-mer and cyclic 9-mer peptides twice a day for 30 days, during which time tumor growth was significantly inhibited, and there was no change in mouse body weight, cage activity, fur texture, or body temperature. In addition there were no changes in size or appearance of major organs relative to the control group. The uterine growth studies revealed that these peptides, unlike tamoxifen, did not stimulate murine uterine growth; indeed they inhibit the uterine stimulated growth induced by tamoxifen.^{2,8} Thus, the peptides derived from AFP represent a new class of potential breast cancer drugs, which are active through a new, yet to be discovered, receptor.

Because there is excellent correlation between the uterine growth assay and the human breast cancer xenograft assay with regard to AFPep peptide inhibition of estrogen-stimulated growth,^{2,14} the TOVNO, TPVNP, TOVN, and TPVN analogues

that are all active in the uterine growth assay are predicted to inhibit human breast cancer growth. We have begun the xenograft assays on these analogues to evaluate this prediction.

Conclusions

We have applied and demonstrated for the first time that REMD simulations can be used as a novel lead compound design tool. We have shown that REMD predicts a common conformation that is shared between the active linear 8-mer and cyclic 9-mer peptides. The predicted common conformation is a conserved reverse β -turn, and the smaller peptide analogues TOVNO, TPVNP, TOVN, and TPVN also contain the same conserved reverse turn. These analogues are shown to inhibit estrogen-dependent cell growth in a mouse uterine growth assay, through interaction with a yet to be discovered key receptor, and are predicted to inhibit human breast cancer. The 5-mer and 4-mer peptides are new discoveries that may lead to promising new antibreast cancer drugs.

Acknowledgment. We thank Carlos Simmerling and Adrian Roitberg for insightful discussions of the replica exchange methodology and Steve Festin for his contributions. Acknowledgment is made to the donors of the Petroleum Research Fund, administered by the ACS, to the Research Corporation, to the NIH (R01 CA102540 and R15 CA115524), to the New York State Breast Cancer Research and Education fund through Department of Health Contract C017922, to the DOD (BC031067 and W81XWH-05-1-0441), and to Hamilton College for support of this work. This project was supported in part by NSF Grant CHE-0457275 and by NSF Grants CHE-0116435 and CHE-0521063 as part of the MERCURY high-performance computer consortium (<http://mercury.chem.hamilton.edu>). This work was first presented at the 2006 International Symposium on Theory and Computations in Molecular and Materials Sciences, Biology and Pharmacology, on February 26, 2006.

JA070202W