Research Letter

Theoretical Study of Sequence Selectivity and Preferred Binding Mode of Psoralen with DNA

Patricia Saenz-Méndez,^{1,2} Rita C. Guedes,³ Daniel J. V. A. dos Santos,³ and Leif A. Eriksson¹

¹ Department of Natural Sciences and Örebro Life Science Center, Örebro University, 701 82 Örebro, Sweden

² Faculty of Chemistry, University of Uruguay (UdelaR), Montevideo 11800, Uruguay

³ Faculty of Pharmacy, University of Lisbon, Prof. Gama Pinto Avenue, 1649-019 Lisbon, Portugal

Correspondence should be addressed to Leif A. Eriksson, leif.eriksson@nat.oru.se

Received 11 July 2007; Accepted 10 October 2007

Recommended by L. Serrano-Andrés

Psoralen interaction with two models of DNA was investigated using molecular mechanics and molecular dynamics methods. Calculated energies of minor groove binding and intercalation were compared in order to define a preferred binding mode for the ligand. We found that both binding modes are possible, explaining the low efficiency for monoadduct formation from intercalated ligands. A comparison between the interaction energy for intercalation between different base pairs suggests that the observed sequence selectivity is due to favorable intercalation in 5'-TpA in (AT)_n sequences.

Copyright © 2007 Patricia Saenz-Méndez et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

The study of the interaction of low molecular weight agents with DNA is essential for a deeper understanding of the biochemistry of the cell and also in the rational design of compounds with desired pharmacological properties. Among the vast amount of molecules that are known or proposed to interact with DNA we have in the current study focused our attention to furocoumarins.

Furocoumarins are tricyclic heterocycles found in a large number of plants [1]. Psoralens (Figure 1) are the best known of these chemical agents and have been employed in the treatment of several skin diseases such as psoriasis, vitiligo, and mycosis fungoides, as well for disorders related to the immune system such as cutaneous T-cell lymphoma [2]. Even though extracts of *Ammi majus* (Bishop's weed) were used by the Egyptians (1550 BC) in the treatment of vitiligo, the first concise report of the combined use of psoralen derivatives and UV light dates back to 1974 [3], in which Parrish et al. coined the term *photochemotherapy* to describe the interaction of UVA light and drugs.

The photochemical reactions of psoralens with DNA have been described in detail [1, 4, 5], and is proposed to proceed via a stepwise mechanism. First, the molecule inter-

calates between the base pairs in double stranded DNA interacting with the π -stack of the nucleobases. After irradiation with UVA light, the photoexcited furocoumarins react with DNA and form covalent adducts through a [2 + 2] pericyclic reaction. The photoaddition takes place mostly between the psoralen 4', 5' double bond (furan ring) and the C5, C6 double bond of a pyrimidine (usually thymine). The furan-side adduct can absorb an additional UVA photon, forming an interstrand cross-linked derivative through photoaddition at the 3, 4 double bond in the pyrone ring [6–8].

Furocoumarin photoaddition takes place predominantly with thymine. The sequence specificity has been investigated using DNA sequencing methodology. The results revealed that thymines in a GC environment have weak reactivity, while adjacent thymines are better targets. In addition, there is a strong preference for 5'-TpA sites compared with 5'-ApT, and (AT)_n sequences are the most reactive towards photoaddition [9, 10].

There is however no information regarding the factors controlling the sequence selectivity. One possibility is that the preference is due to favorable intercalation into this site. Another explanation is that the molecules are docked between all different base pairs, but photoaddition occurs preferentially in a special environment. A third possibility could be

that psoralens prefer both intercalation and photoreaction with a 5'-TpA site in an AT environment.

In this context, computational techniques are an exceptional tool to explore these different factors. Molecular docking is a computational method for predicting ligand-receptor binding when both the binding site and binding mode are unknown [11]. Even though DNA has an exceptionally well-characterized structure, there is a very small number of DNA-ligand docking studies [12–16], and to the best of our knowledge there is only scarce information on docking involving this kind of ligand systems [17].

The aim of this study is to obtain complexes of DNA-models and psoralen in which the ligand in principle could act either as intercalator or minor groove binder. We used two different DNA fragments, d(CCTTGCTACCTT)₂ and d(TATATATATATA)₂, as targets for intercalation and minor groove complexes, in order to evaluate the preferential interaction between different base pairs and explore the proposed mechanism. Obtaining detailed information on the possible interaction sites will assist in the development of new and more efficiently intercalating photoactive compounds.

2. METHODS

Duplex B-DNA was constructed with the Molecular Operating Environment (MOE) software [18] and was minimized to 0.00001 kcal mol⁻¹ Å⁻¹ using the CHARMM27 force field [19]. After minimization, the model was subjected to molecular dynamics (MD) simulations [20] employing an NVT ensemble [21] and the CHARMM27 potential. After the MD simulation, a final energy minimization was performed. The resulting DNA-model was used in the docking studies.

Because of the ring-fused planar structure of furocoumarin, no ligand conformational search is needed. A single energy minimization calculation was performed until the root-mean-square (RMS) gradient was 0.00001 kcal mol⁻¹ Å⁻¹ using the same force field.

Manual docking of psoralen to the DNA was performed, in which different orientations for the ligand were considered. The DNA model with the ligand was energy minimized, subjected to MD simulation using the same settings as above except for a longer (250 picoseconds) heating interval, and finally energy minimized, using in all the stages the CHARMM27 force field. All trajectories were rapidly equilibrated during the first 200 picoseconds of the production runs. Hence, a total 1 nanosecond production simulation was considered sufficient to allow for structural relaxation. The potential energy of the final complex was calculated (E_{complex}). The ligand was then moved away from the DNA stack, the noninteracting system energy minimized, and the potential energy of the separated moieties calculated ($E_{\text{DNA}} + E_{\text{ligand}}$). Subtracting this from E_{complex} gives the interaction energy (IE) reported herein.

Throughout the calculations, the surrounding was modeled through the distance-dependent dielectric model of bulk water [18, 22].

3. RESULTS AND DISCUSSION

3.1. Psoralen binding modes

El-Gogary and El-Gendy [17] employed spectrophotometric DNA titration to calculate the intercalation affinity. The spectrophotometric data for 8-methoxypsoralen bound to calf-thymus DNA showed that even though the molecule could intercalate, a large fraction (approximately 30%) was also bound to the surface of DNA. No theoretical studies regarding different binding modes of psoralens are available. To this end, the ability of psoralen to intercalate between base pairs in d(CCTTGCTACCTT)₂ and d(TATATATATATA)₂ sequences, as well to bind to the minor groove, was therefore examined in the current study. The results are shown in Tables 1 and 2.

Psoralen does not present different geometric conformations due to its rigid nature. However, the molecule is asymmetric and may therefore interact with different energy in different poses. Four different orientations for docking into the DNA stack were considered in all systems, numbered as indicated in Figure 1.

First, we examine the ability to intercalate between different base pairs, GC, CT, and TA following the procedure described in Section 2. Each pose was inserted between the considered base pair. In some cases the ligand moved away from the intercalation site to interact with the minor groove, the major groove, or interacting only with one strand of DNA ("single-strand insertion," Table 1). In addition, we calculated the interaction energy when psoralen was placed initially in the minor groove. In this case, all poses remained bound to this region after the MD simulation (Table 2).

In all cases, the interaction between the ligand and the macromolecule lowers the total energy considerably, leading to stable complexes. Our results show that intercalation and minor groove binding are competing processes, in full agreement with the experimental information. When psoralen interacted with the d(CCTTGCTACCTT)₂ fragment, the most favorable docking energies for intercalation were obtained for pose 2 intercalated between CT (−36.8 kcal mol⁻¹) and pose 3 between TA (−36.7 kcal mol⁻¹). Minor groove binding into that DNA model was found to be slightly more favorable for poses 1 and 3, yielding interaction energies of −39.1 and −38.5 kcal mol⁻¹, respectively. When considering the poly-TA fragment, intercalation of pose 2 in the 5'-TpA site led to the most stable complex with an interaction energy of −43.7 kcal mol⁻¹. Minor groove binding energies were a few kcal mol⁻¹ less favorable compared with the nonpoly-TA fragment (best docking energy −34.4 kcal mol⁻¹ for pose 1). Those results show that even though minor groove binding is possible in both DNA strands, it is comparatively less favored in poly-TA fragments.

The current docking studies also explain other experimental facts. Tessman et al. [6] studied the photochemical reaction of psoralen with calf-thymus DNA. They observed a relatively low efficiency for the conversion of noncovalent complexes to monoadducts (approximately 11%). They suggested that the probability of having the appropriate excited state with the correct geometry alignment for photoaddition

TABLE 1: Calculated docking energies $IE = E_{\text{complex}} - (E_{\text{DNA}} + E_{\text{ligand}})$ (kcal mol⁻¹) obtained for different orientations of psoralen interacting with d(CCTTGCTACCTT)₂ and d(TATATATATATA)₂ fragments.

d(CCTTGCTACCTT) ₂ fragment			
Orientation	Docking mode	Docking site	IE
1	Single strand insertion	GC	-38.0
2	Major groove binding	GC	-32.4
3	Single strand insertion	GC	-51.1
4	Major groove binding	GC	-35.5
1	Intercalation	CT	-32.3
2	Intercalation	CT	-36.8
3	Intercalation	CT	-35.5
4	Intercalation	CT	-36.5
1	Intercalation	TA	-36.5
2	Intercalation	TA	-32.7
3	Intercalation	TA	-36.7
4	Intercalation	TA	-30.2
d(TATATATATATA) ₂ fragment			
Orientation	Docking mode	Docking site	IE
1	Intercalation	5'-TpA	-24.1
2	Intercalation	5'-TpA	-43.7
3	Intercalation	5'-TpA	-18.2
4	Intercalation	5'-TpA	-21.4
1	Major groove binding	5'-ApT	-28.7
2	Major groove binding	5'-ApT	-29.5
3	Major groove binding	5'-ApT	-26.4
4	Single strand insertion	5'-ApT	-33.7

TABLE 2: Calculated docking energies $IE = E_{\text{complex}} - (E_{\text{DNA}} + E_{\text{ligand}})$ (kcal mol⁻¹) obtained for different orientations of psoralen for minor groove binding in d(CCTTGCTACCTT)₂ and d(TATATATATATA)₂ fragments.

d(CCTTGCTACCTT) ₂ fragment	
Orientation	IE
1	-39.1
2	-37.7
3	-38.5
4	-35.5
d(TATATATATATA) ₂ fragment	
Orientation	IE
1	-34.4
2	-31.7
3	-32.7
4	-32.7

is quite low, which could be due to the fact that intercalated psoralen is in dynamic equilibrium with nonbound ligand outside the helix. From our results, we instead propose an alternative explanation, namely, that the low efficiency is a consequence of the two competing binding modes, that is, intercalation versus minor groove binding.

FIGURE 1: Molecular structure of psoralen and the four different docking poses investigated for intercalation.

3.2. Psoralen sequence selectivity

The use of two different sequences of bases (d(CCTTGCTACCTT)₂ and d(TATATATATATA)₂) in our docking studies enables us to evaluate the sequence selectivity. The two strands in principle have the same potential to react with psoralen and form photoadducts. However, the experimental studies reveal a different behavior between the two [9, 10]. Psoralen preferentially photoreact in (AT)_n sequences and especially in 5'-TpA sites. However, it is difficult from those results to conclude whether the preference is a consequence of preferential intercalation, photoreaction, or both.

Thus, we decided to study the ability of psoralen to intercalate between different base pairs in two DNA models (Table 1). In all cases the same four different orientations were considered (Figure 1). Each pose was inserted between the base pair, minimized and subjected to MD simulation. In some cases, the ligand did not remain intercalated, but moved to either the minor or major groove, or became inserted between two bases interacting with one strand of DNA only.

When the ligand interacted with the d(CCTTGCTACCTT)₂ fragment, all three environments (GC, CT, TA) behaved differently. We were unable to get psoralen intercalated between GC, while intercalation between CT turned out to be almost as favorable as between TA. When the ligand is inserted (GC base pair) it interacts with only one strand of DNA, acting as hydrogen bond acceptor via its carbonyl oxygen atom with the cytosine amino group (Figure 2). Such interaction stabilizes the insertion of the ligand to the GC tract. Figure 2 also shows the van der Waals contact surface for pose 3 inserted between GC.

The results of the docking studies with the poly-TA model were completely different. Psoralen intercalates to give stable complexes in 5'-TpA sites that furthermore yields the most favorable interaction energy (-43.7 kcal mol⁻¹). However, when considering 5'-ApT sites, the ligand became bound to the minor or major groove or single-strand inserted, instead of remaining intercalated after MD simulation.

The most stable complex has the ligand perfectly intercalated between T and A, interacting only with those base pairs. During the MD simulation the psoralen was however rotated

FIGURE 2: (a) Space-filling model of psoralen interacting with DNA bases for orientation 3 inserted between GC. (b) van der Waals contact surface between ligand and DNA.

FIGURE 3: (a) Space-filling model of psoralen interacting with DNA bases for orientation 2 intercalated in the 5'-TpA site. (b) van der Waals contact surface between ligand and DNA.

90 degrees from the initial pose to the final one, as is shown in Figure 3. In addition, the van der Waals surface is closed around the ligand (Figure 3).

The topology of the two sites (5'-TpA and 5'-ApT) is very different in our model. The base stacking is not the same in the two sequences; in the former the base pairs are per-

fectly parallel while in the latter they are not. These results are in agreement with experimental information. McClellan et al. [23] have found a nonstandard B-helix for alternating (AT)_n sequences, as a consequence of poor stacking between the base pairs. The same authors mentioned that (AT)_n sequences are sites of great instability, present in promoters where DNA replication is initiated, a process in which unstacking is required. The unstacking and instability compared with nonpoly-TA sequences could lead to a better accessibility of the C5–C6 double bond of thymine due to easier intercalation. Our results suggest that intercalation is indeed favored in (AT)_n sequences and especially in 5'-TpA sites.

4. CONCLUSIONS

A computational docking study for the photochemical agent psoralen was performed for the first time, employing two different DNA models (d(CCTTGCTACCTT)₂, and d(TATATATATATA)₂). We have analyzed different binding modes and suggest that the low efficiency observed for the conversion of the noncovalent complexes to monoadducts are a consequence of two competing binding modes, intercalation and minor groove binding.

It is well known that strong sequence selectivity for furocoumarin photoaddition exists. From our results we can conclude that the preference is due to favorable intercalation in 5'-TpA sites in AT environments (poly-TA DNA). Albeit this will control the distribution in different environments, it does however not exclude that photoaddition may take place preferentially in those sites also. We are currently combining additional docking studies with quantum chemical calculations to take these factors into account.

ACKNOWLEDGMENTS

The Swedish Science Research Council (VR) and Sparbanksstiftelsen Nya are gratefully acknowledged for financial support.

REFERENCES

- [1] T. F. Anderson and J. J. Voorhees, "Psoralen photochemotherapy of cutaneous disorders," *Annual Review of Pharmacology and Toxicology*, vol. 20, pp. 235–257, 1980.
- [2] P. S. Song and K. J. Tapley, "Photochemistry and photobiology of psoralens," *Photochemistry and Photobiology*, vol. 29, no. 6, pp. 1177–1197, 1979.
- [3] J. A. Parrish, T. B. Fitzpatrick, L. Tanenbaum, and M. A. Pathak, "Photochemotherapy of psoriasis with oral methoxsalen and longwave ultraviolet light," *The New England Journal of Medicine*, vol. 291, no. 23, pp. 1207–1211, 1974.
- [4] G. D. Cimino, H. B. Gamper, S. T. Isaacs, and J. E. Hearst, "Psoralens as photoactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry," *Annual Review of Biochemistry*, vol. 54, pp. 1151–1193, 1985.
- [5] N. Kitamura, S. Kohtani, and R. Nakagaki, "Molecular aspects of furocoumarin reactions: photophysics, photochemistry, photobiology, and structural analysis," *Journal of Photochemistry and Photobiology C: Photochemistry Reviews*, vol. 6, no. 2-3, pp. 168–185, 2005.

- [6] J. W. Tessman, S. T. Isaacs, and J. E. Hearst, "Photochemistry of the furan-side 8-methoxypsoralen-thymidine monoadduct inside the DNA helix. Conversion to diadduct and to pyrone-side monoadduct," *Biochemistry*, vol. 24, no. 7, pp. 1669–1676, 1985.
- [7] M. A. Pathak and T. B. Fitzpatrick, "The evolution of photochemotherapy with psoralens and UVA (PUVA): 2000 BC to 1992 AD," *Journal of Photochemistry and Photobiology B: Biology*, vol. 14, no. 1-2, pp. 3–22, 1992.
- [8] I. M. Schmitt, S. Chimenti, and F. P. Gasparro, "Psoralen-protein photochemistry—a forgotten field," *Journal of Photochemistry and Photobiology B: Biology*, vol. 27, no. 2, pp. 101–107, 1995.
- [9] E. Sage and E. Moustacchi, "Sequence context effects on 8-methoxypsoralen photobinding to defined DNA fragments," *Biochemistry*, vol. 26, no. 12, pp. 3307–3314, 1987.
- [10] V. Boyer, E. Moustacchi, and E. Sage, "Sequence specificity in photoreaction of various psoralen derivatives with DNA: role in biological activity," *Biochemistry*, vol. 27, no. 8, pp. 3011–3018, 1988.
- [11] I. Halperin, B. Ma, H. Wolfson, and R. Nussinov, "Principles of docking: an overview of search algorithms and a guide to scoring functions," *Proteins*, vol. 47, no. 4, pp. 409–443, 2002.
- [12] B. G. Feuerstein, N. Pattabiraman, and L. J. Marton, "Spermine-DNA interactions: a theoretical study," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 83, no. 16, pp. 5948–5952, 1986.
- [13] K. Gao, B. Tao Fan, N. El Fassi, et al., "Comparative study of activities between verbascoside and rutin by docking method," *QSAR & Combinatorial Science*, vol. 22, no. 1, pp. 18–28, 2003.
- [14] T. Tuttle, E. Kraka, and D. Cremer, "Docking, triggering, and biological activity of dynemicin A in DNA: a computational study," *Journal of the American Chemical Society*, vol. 127, no. 26, pp. 9469–9484, 2005.
- [15] K. V. Miroshnychenko and A. V. Shestopalova, "Flexible docking of DNA fragments and actinocin derivatives," *Molecular Simulation*, vol. 31, no. 8, pp. 567–574, 2005.
- [16] R. Rohs, I. Bloch, H. Sklenar, and Z. Shakked, "Molecular flexibility in *ab initio* drug docking to DNA: binding-site and binding-mode transitions in all-atom Monte Carlo simulations," *Nucleic Acids Research*, vol. 33, no. 22, pp. 7048–7057, 2005.
- [17] T. M. El-Gogary and E. M. El-Gendy, "Noncovalent attachment of psoralen derivatives with DNA: Hartree-Fock and density functional studies on the probes," *Spectrochimica Acta, Part A*, vol. 59, no. 11, pp. 2635–2644, 2003.
- [18] "Molecular Operating Environment (MOE)," 2005.06; Chemical Computing Group Inc.; Montreal, Canada, 2005.
- [19] A. D. Mackerell Jr., M. Feig, and C. L. Brooks III, "Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulation," *Journal of Computational Chemistry*, vol. 25, no. 11, pp. 1400–1415, 2004.
- [20] MD conditions: duration: 1.0 ns, temperature: 300 K, heat: 100 ps, initial temperature: 150 K, time step: 0.002 ps, sample period: 0.5 ps, temperature response: 0.1 ps, pressure response: 0.1 ps.
- [21] S. D. Bond, B. J. Leimkuhler, and B. B. Laird, "The Nosé-Poincaré method for constant temperature molecular dynamics," *Journal of Computational Physics*, vol. 151, no. 1, pp. 114–134, 1999.
- [22] J. Guenot and P. A. Kollman, "Molecular dynamics studies of a DNA-binding protein—2: an evaluation of implicit and explicit solvent models for the molecular dynamics simulation of the Escherichia coli trp repressor," *Protein Science*, vol. 1, no. 9, pp. 1185–1205, 1992.
- [23] J. A. McClellan, E. Palecek, and D. M. Lilley, "(A-T)_n tracts embedded in random sequence DNA—formation of a structure which is chemically reactive and torsionally deformable," *Nucleic Acids Research*, vol. 14, no. 23, pp. 9291–9309, 1986.