ELSEVIER

Il Farmaco 54 (1999) 202–212

IL FARMACO

DNA–psoralens molecular recognition using molecular dynamics

Raffaella Boggia [a,*], Marta Fanciullo [b], Livia Finzi [b], Ottaviano Incani [b], Luisa Mosti [c]

[a] *Dipartimento di Chimica e Tecnologie Farmaceutiche ed Alimentari, Università di Genova, Via Brigata Salerno (ponte), I-16147 Genoa, Italy*

[b] *Tecnofarmaci S.C.p.A., Laboratorio di Chimica Computazionale, Via del Mare 87, I-00040 Pomezia, Rome, Italy*

[c] *Dipartimento di Scienze Farmaceutiche, Università di Genova, Viale Benedetto XV 3, I-16132 Genoa, Italy*

Received 1 December 1998; accepted 10 February 1999

Abstract

Furocoumarins are an important class of compounds for photochemotherapy used in the treatment of numerous diseases characterised by hyperproliferative conditions. Their photosensitising activity has been related to the ability to form covalent linkage with the pyrimidine bases of DNA upon UV-A irradiation. Using the published experimental data of the 3D-structure of the furan-side monoadduct between the 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen and the DNA oligonucleotide d(GCG-TACGC)$_2$ as starting point, a computational procedure to study the dark intercalation complexes by molecular dynamics was built and then preliminary validated on another psoralen. This docking procedure provides a tool to better understand the mechanism of the DNA-intercalation of these linear furocoumarins and it can be useful in the design of new photochemotherapeutic agents with an improved therapeutic profile.

Keywords: Clustering; DNA; Docking; Molecular dynamics; Psoralens

1. Introduction

Psoralens are naturally occurring or synthetic compounds with strong photobiological and phototherapeutic activities. In the living system their principal target is DNA, to which they are able to bind covalently upon UV-A irradiation. Due to their DNA-photobinding properties, psoralen derivatives have been used in the treatment of numerous diseases characterised by hyperproliferative conditions or by lack of skin pigmentation [1] and, more recently, in extracorporeal form of photochemotherapy (photophoresis) for cutaneous T-cell lymphoma [2]. Furthermore, psoralens are now under investigation for both antibacterial and antiviral activities [3]. However, together with highly therapeutic efficacy, some side effects, such as skin phototoxicity, mutagenesis and risk of skin cancer, have been observed [4].

It has also been shown that the action of the linear furocoumarins is due to the fact that: after having intercalated into nucleic acids and consequently formed a molecular complex with the macromolecules, they undergo cycloadditions with adjacent pyrimidine bases when activated by UV-A light. Thymidine is the primary nucleoside in DNA at which photoreaction occurs. In this photoreaction, the pyrimidine bases always react with their 5,6 double bond; while two reactive sites are present in the psoralens, i.e. the 4′,5′ (furan) and the 3,4 (pyrone) double bonds. Therefore, different types of cycloadducts can be formed: mono- (furan-side or pyrone-side) and diadducts (cross-link) [5].

In particular, if the geometrical configuration is favourable, the psoralens linked by furan-side can absorb a second photon and can form an inter-strand cross-link by reacting with a pyrimidine belonging to the opposite strand of the DNA. On the contrary, pyrone-side monoadducts do not absorb light at wavelengths above 320 nm and cannot be driven to inter-strand cross-link with long-wavelength UV light [6,7].

* Corresponding author. Tel.: +39-010-353 2636; fax: +39-010-353 2684.

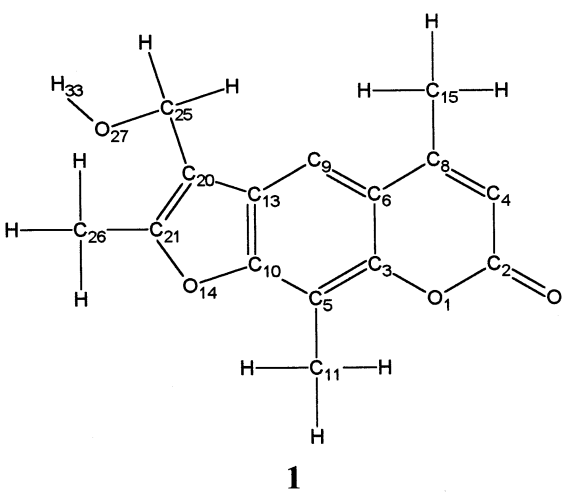
E-mail address: raffa@anchem.unige.it (R. Boggia)

0014-827X/99/$ - see front matter
PII: S0014-827X(99)00012-9

Recently the structures of the furan-side monoadduct (MAf) between the 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen (HMT) (**1**) and the DNA oligonucleotide d(GCGTACGC)$_2$ and the HMT inter-strand cross-link diadduct of the same oligonucleotide have been determined by ^{1}H NMR spectroscopy and their co-ordinates have been deposited in the Brookhaven Protein Data Bank (PDB codes: 203D and 204D) [7].

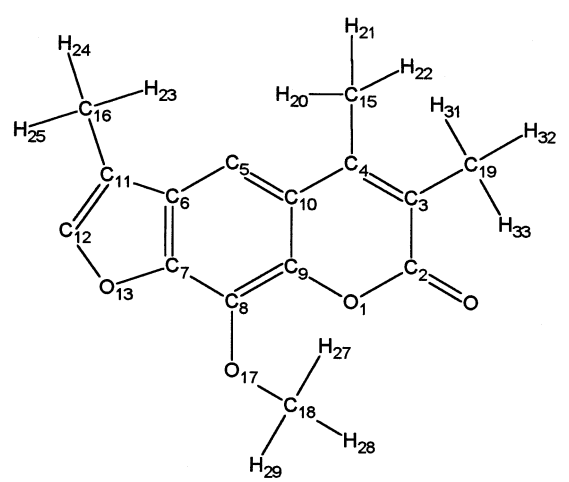
HMT is a bifunctional synthetic psoralen derived from the natural product 4,5′,8-trimethylpsoralen (TMP) and characterised by a greatly enhanced water solubility with respect to TMP.

Since the geometrical arrangements allowed during the non-covalent binding (in the dark) strongly influence the subsequent photoreactions, we have built a computational procedure to study the intercalation complexes between the d(GCGTACGC)$_2$ oligonucleotide and furocoumarins. The results obtained have been compared to experimental data, keeping in mind that they can be a preliminary important approach to the study of the UVA-light activated complexes.



1

Fig. 1. Chemical structure of the bifunctional 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen (HMT) (**1**). The atomic numbering is the convention used.



2

Fig. 2. Chemical structure of the monofunctional 3,4,4′-trimethyl-8-methoxyfurocoumarin (**2**). The atomic numbering is the convention used.

Table 1
Atom names and partial atomic charges of 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen (**1**)

Atom [a]	Charge
O1	−0.3906
C2	0.9153
C3	0.2547
C4	−0.6254
C5	−0.0827
C6	−0.1885
O7	−0.5543
C8	0.3502
C9	−0.2204
C10	0.1164
C11	−0.0588
H12	0.2061
C13	0.0047
O14	−0.1692
C15	−0.2993
H16	0.1971
H17	0.0487
H18	0.0352
H19	0.0514
C20	−0.3040
C21	0.2407
H22	0.1127
H23	0.0817
H24	0.0863
C25	0.4759
C26	−0.3071
O27	−0.6067
H28	−0.0603
H29	−0.0048
H30	0.1256
H31	0.1046
H32	0.0960
H33	0.3688

[a] Atom numbering conventions in Fig. 1.

In order to test our computational procedure another linear furocoumarin has been taken into account: the 3,4,4′-trimethyl-8-methoxypsoralen (**2**). Molecule **2** has recently been reported in literature [8] as a monofunctional compound able to form only furan-side monoadduct. For this compound neither X-ray nor NMR data concerning the intercalation complexes are available, however, our computational model is in agreement with the biological data.

The theoretical procedure proposed in this paper is able to describe the different behaviour in the dark of the bifunctional agent **1** and of the monofunctional **2**, so it could be applicable to other analogues, under the hypothesis that they recognise the DNA in a similar way.

Therefore, the present approach can represent the first step of a more general project that involves the use of several theoretical methods as a tool to assist the design of new potential photochemotherapeutic agents.

2. Experimental

2.1. Model building of the ligands

The furocoumarins **1** and **2** were designed using the interactive model building of the package MacroModel version 5.5 [9], their chemical structures are reported in Figs. 1 and 2, respectively.

A complete geometry optimisation of the two molecules was obtained using the MM2* force-field implemented in MacroModel and the TNCG (Truncated Newton Conjugate Gradient) optimisation method until the energy gradient was < 0.05 kcal/mol per Å.

Table 2
Atom names and partial atomic charges of 3,4,4′-trimethyl-8-methoxyfurocoumarin (**2**)

Atom [a]	Charge
O1	−0.2841
C2	0.7459
C3	−0.2634
C4	0.1502
C5	−0.2847
C6	0.0370
C7	0.0250
C8	0.2820
C9	0.0304
C10	−0.0940
C11	0.0013
C12	−0.1202
O13	−0.1472
O14	−0.5179
C15	−0.1108
C16	−0.1511
O17	−0.3152
H18	0.2233
H19	0.1933
H20	0.0747
H21	0.0354
H22	0.0354
H23	0.0644
H24	0.0549
H25	0.0549
C26	0.0814
H27	0.0530
H28	0.0423
H29	0.0423
C30	−0.1085
H31	0.0470
H32	0.0470
H33	0.0756

[a] Atom numbering conventions in Fig. 2.

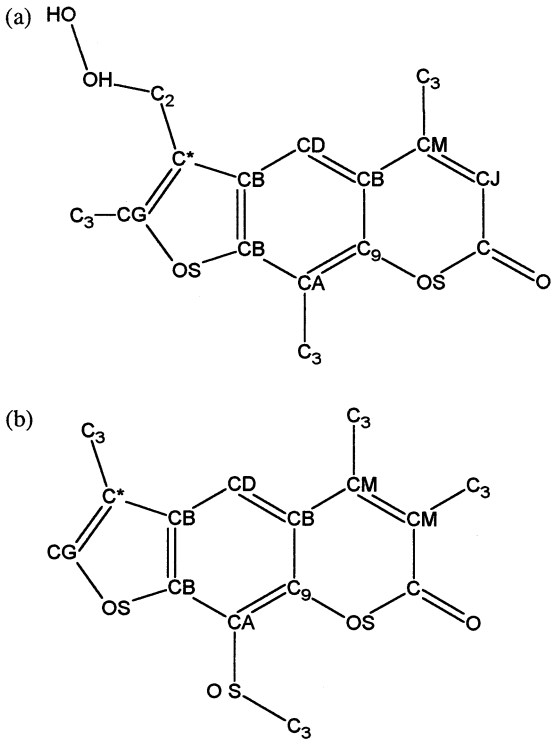


Fig. 3. (a) 4′-(Hydroxymethyl)-4,5′,8-trimethylpsoralen (**1**)—AMBER topology. (b) 3,4,4′-Trimethyl-8-methoxyfurocoumarin (**2**)—AMBER topology.

Afterwards they were studied by ab initio techniques using the GAUSSIAN-94 [10] package (run on an Alpha AXP-3000/500 cluster, while all the other calculations were performed on a SGI-Indigo2 workstation) both to fully optimise their conformation and to derive the partial atomic charges. The geometries were preliminary determined at the SCF level by analytical gradient techniques, using a STO-3G and a 3-21G basis set, and then optimised using a 6-31G* basis set.

The electrostatic potential derived charges were calculated on the fully optimised structures using the 6-31G* basis set and the Chelp schemes and they are reported in Tables 1 and 2 (for **1** and **2**, respectively). The geometrical data and the charges have been used to parameterise the psoralen moiety inside the AMBER 3.0A package [11].

It was also necessary to define a new atom type inside the AMBER force-field which we refer to as *C*9 (sp^2 aromatic carbon at the junction between six- and six-membered rings) the same as to define the aromatic junction carbon in between five- and six-membered rings as AMBER sp^2.

Fig. 3 and Table 3, respectively, report the AMBER atom type used and the additional bonds, bond angles and torsion parameters for psoralens.

In order to verify the reliability and the usefulness of the additional parameters, we performed an energy minimisation of the two ligands and we compared the

results to the ab initio geometries. The obtained RMS for molecules **1** and **2** were 0.3 and 0.2 Å, respectively, being even lower when the hydroxymethyl group of **1** and the methoxy moiety of **2** were not included in the comparison (these substituents are free to rotate without affecting sensibly the molecular energy).

Besides, we submitted the two ligands to a molecular dynamics simulation of 300 ps at 300 K (after heating the system from 0 to 300 K) in order to verify the behaviour of the modified force-field at the conditions subsequently used in the docking phase.

2.2. Model building of a canonical octanucleotide alone

The duplex d(5′-GCGTACGC-3′) oligonucleotide simulation involves in vacuo energy minimisation and molecular dynamics (MD) calculations using the AMBER 3.0A package [11].

In order to assess the reliability of our simulation we have undertaken MD runs of 1.2 ns on the canonical octamer d(5′-GCGTACGC-3′)$_2$ starting from a canonical B-DNA structure [12] created using the NUCGEN module of AMBER (see Fig. 4(a)).

Since our main goal is the docking rather than the dynamics of DNA alone, it seems too computationally-heavy to us to perform a nanosecond MD simulation on duplex d(5′-GCGTACGC-3′) with full representation of DNA charges, solvent and long range electrostatics as recently described by Cheathan and Kollman [13]. Thus some approximations were necessarily applied. At first, phosphate charges were reduced so as to have a slightly negative charge on each nucleotide (-0.32 e) and a distance-dependent dielectric was chosen. This was done according to what is reported in literature [14] to simulate the effect of solvent and shielding, since the calculations were performed in vacuo. However, since under these conditions the helix was not retained, we decided to follow the application of Srinivasan at al. [15] and Singh et al. [16].

We configure the system as a fully anionic oligonucleotide with monovalent cations placed initially at positions bifurcating the O–P–O angle at 6.0 Å from the phosphorus atom. The cations were assigned a charge of $+1$ electron, a Lennard–Jones radius parameter of 5 Å with a well depth of 0.12 kcal/mol, and a mass of 131 amu, which is a representation of hexahydrates Na^+. The effect of hydration was included implicitly with a distance-dependent dielectric model ($\varepsilon = R$).

The starting structure was energy minimised using steepest descent followed by conjugate gradient method until the energy gradient was < 0.1 kcal/mol.

A restraint was applied to counter-ions with a force constant of 20 kcal/mol per Å; different from what has been reported in literature [15], the restraints applied to counter-ions were necessary to preserve the helix geometry.

Table 3
Additional bond, bond angle, torsion and improper torsion for psoralens introduced in AMBER force-field [11]

Bonds [a]	K_r	r_{eq}	General torsion [b]	$Vn/2$	γ	n
C*–C3	317.0	1.510	X–CA–CB–X [c]	16.3	180.0	2
CA–OS	450.0	1.364	X–CB–CB–X	16.3	180.0	2
C9–CA	469.0	1.404	X–CB–CD–X	16.3	180.0	2
C9–CB	520.0	1.370	X–OS–C9–X	16.3	180.0	2
C9–OS	428.0	1.360	X–C9–CA–X	16.3	180.0	2
Angles [d]	K_θ	θ_{eq}	Specific torsion [b]	$Vn/2$	γ	n
CB–CA–OS	70.0	120.0	CM–C–OS–C9	5.4	0.0	2
CB–CM–CM	85.0	120.0	CM–CM–CB–C9	5.4	0.0	2
CM–C–OS	70.0	117.2	CB–CA–OS–C3	1.8	180.0	2
C3–C*–CB	70.0	120.0	CD–CB–CM–CM	5.3	180.0	2
C3–C*–CG	70.0	120.0	C9–CA–OS–C3	1.8	0.0	2
C3–CM–CM	85.0	120.0				
CA–OS–C3	65.0	113.5				
OS–CA–C9	70.0	120.0				
			Improper torsion [b]	$Vn/2$	γ	n
			CB–C9–CA–OS	14.0	180.0	2
			CB–CG–C*–C3	14.0	180.0	2
			C–CM–CM–C3	14.0	180.0	2
			C3–CM–CB–CM	14.0	180.0	2
			O–C–CM–OS	14.0	180.0	2

[a] K_r (kcal/mol Å^2), stretching force constant; r_{eq} (Å), equilibrium bond length.
[b] $Vn/2$ (kcal/mol), torsional force constant; γ (°), phase; n, periodicity.
[c] X means any atom.
[d] K_θ (kcal/mol rad^2), bending force constant; θ_{eq} (°), equilibrium bond angle.

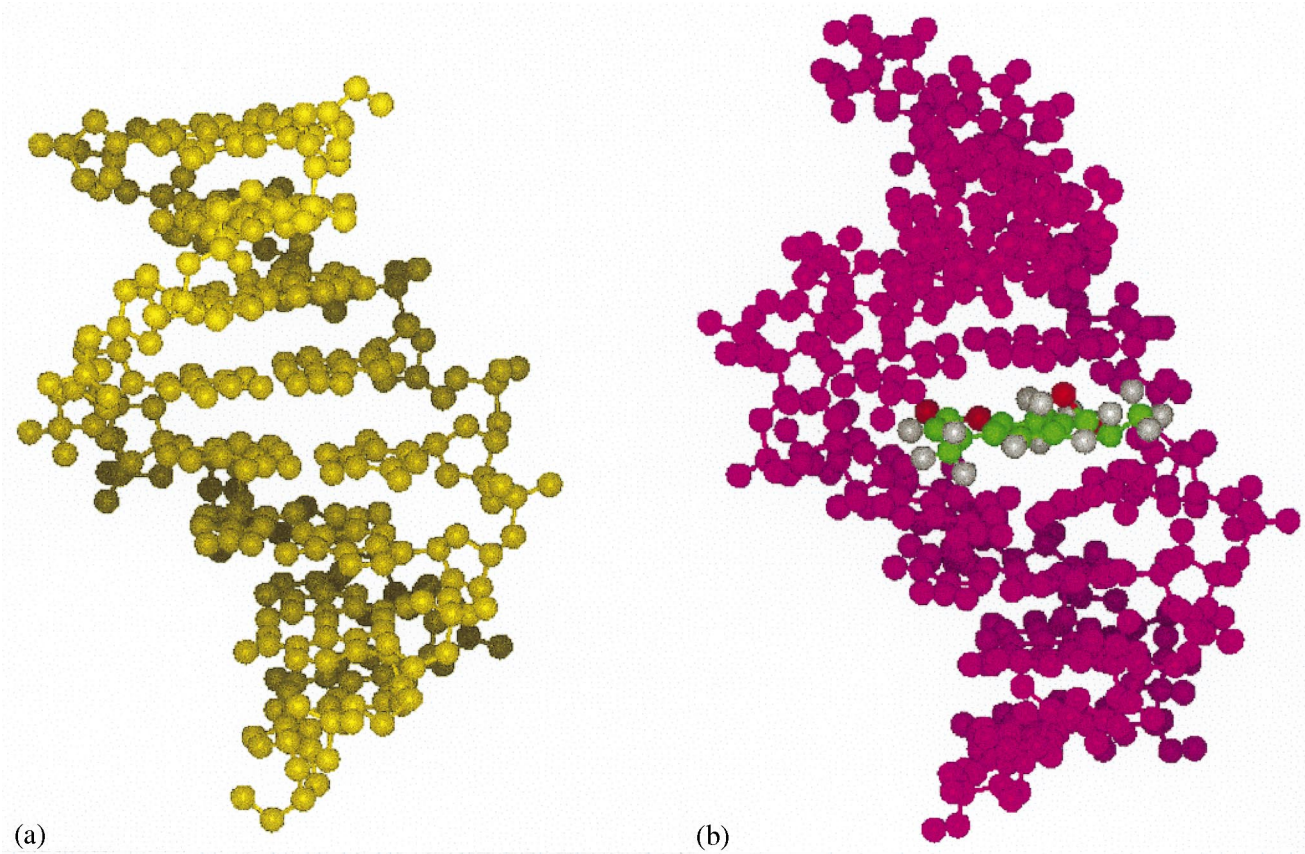


Fig. 4. (a) B_I-DNA-like d(GCGTACGC)$_2$ oligonucleotide (yellow). (b) MAf structure: oligonucletide (pink)-psoralen (coloured by atom type) experimental complex.

The minimised structure was then heated from 0 to 300 K in steps of 20 K for 15 ps with the MD velocities being reassigned after each step from a Maxwell–Boltzmann distribution. A 1.0 fs integration step was chosen. The MD simulation was continued for 1.2 ns at 300 K during which 400 structures have been sampled and energy minimised.

The results indicated this dynamical model to be a quite stable double helix which lies at 2.9 Å RMS deviation from the canonical B-form and this RMS value is reduced to 1.7 Å when the first and last base pairs of the oligonucleotide were excluded from the comparison.

It is important to underline that our final target is to apply this theoretical model to the distorted octamer derived from the ^{1}H NMR data, which consist of 40 structures. These 40 structures are based on 328 interproton distance restraints from NOE measurements, 18 hydrogen bond distance restraints, and 50 restraints to enforce chirality in the deoxyribose residues. The first 20 structures were generated from B-form DNA starting co-ordinates, and the second 20 structures were generated from A-form DNA starting co-ordinates [7].

The RMS deviation of the co-ordinates for the 40 structures for the MAf is 1.12 Å and this value represents for us a reference parameter for all theoretical calculations.

The analysis of the typical helix parameters was made using the CURVES software package [17].

The results are reported graphically in Fig. 5, showing the helical axis system, the base positions (represented by rectangles) and backbone ribbons for a sequence of structures sampled at intervals of 150 ps during the dynamical trajectory (the numerical output of the program is available on request). It is clear that our theoretical model maintains an overall B-form structure over the course of the MD trajectory, although some deviations from the canonical form are evident.

2.3. Model building of HMT-d(GCGTACGC)$_2$ non-covalent complex

One of the MAf B-form (see Fig. 4(b)) has been used as starting point after the deletion of the ligand co-ordinates. Thus the DNA oligonucleotide d(GCGTACGC)$_2$ has been configured as a fully anionic oligonucleotide with 14 hexahydrates Na^+, following the procedure described above. The energy minimisation of the octamer allowed a geometrical rearrangement of the

thymine involved in the covalent adduct, giving the plausible non-covalent oligonucleotide structure as final result.

In order to verify the capability of our procedure to dock a linear psoralen into double stranded nucleic acids, we performed some calculations using different inter-molecular (psoralen–DNA) orientations (configurations) as starting points.

The HMT molecule has been graphically incorporated between the 4th and 5th base pairs in three different ways (see Fig. 6) in which HMT is:

A. located and oriented in quite a good way to promote the full intercalation inside the DNA (yellow);
B. oriented as in A, but the psoralen residue is located 5 Å out of the major groove (green);
C. located as in A, but oriented in a perpendicular way with respect to A (pink).

Energy minimisation cycles were performed. At first an harmonic restraint (20 kcal/mol per Å) has been applied to the psoralen moiety, to eliminate bad contacts due to the manual graphical phase; then the system has been fully relaxed (until the energy gradient was < 0.1 kcal/mol.) following the same procedure described above. Three MD simulations each of 300 ps at 300 K were performed during which 99 structures have been sampled and energy minimised.

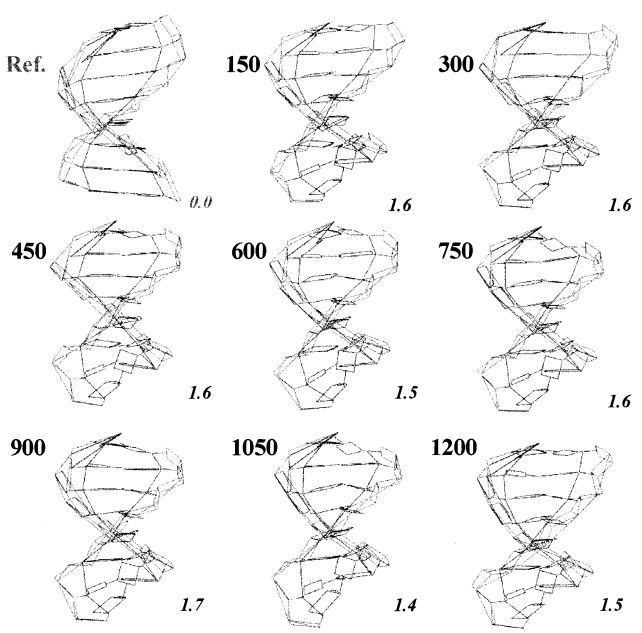


Fig. 5. Graphic output of CURVES showing the helical axis system for a sequence of snapshots of the d(GCGTACGC)$_2$ oligonucleotide during 1200 ps of molecular dynamic simulation. The first structure named as Ref. (reference structure) is the canonical B_I-DNA form used as starting configuration. The time at which the snapshot was taken is shown in the upper left-hand corner of each panel. The calculated RMS deviation of the structure from the Ref. is given in the lower right-hand corner of each panel (excluding the first and the last base pairs from the comparison).

In each simulation the ligand was allocated into the helix, independently of the starting orientation. So we decided to use the less favourite configuration **C** to continue the MD simulation for another 300 ps sampling and minimising another 99 structures, for a total of 600 ps and 198 structures.

2.4. Model building of 3,4,4′-trimethyl-8-methoxypsoralen-d(GCGTACGC)$_2$ non-covalent complex

In order to test our computational procedure, another linear furocoumarin has been taken into account: the 3,4,4′-trimethyl-8-methoxypsoralen (**2**). It is very similar to **1**, but it is a monofunctional agent. Due to the structure similarity to **1** we assumed that the DNA undergoes comparable conformational and dynamic changes, so that the use of the same computational procedure is justified.

Molecule **2** has been graphically incorporated between the 4th and 5th base pairs of the octamer similarly to the HMT configuration **C** and all the steps of the computational procedure already described were repeated.

A MD simulation of 600 ps at 300 K was performed during which 198 structures have been sampled and minimised.

3. Results and discussion

MD simulations were analysed using the clustering procedure, as implemented in an in-house software (FRAMES [18]) described below. It is important to underline that the FRAMES algorithm does not work according to any classification–clustering technique commonly used in chemometrics. In the current implementation, the clustering is not hierarchical.

To monitor the geometry of the reversible psoralen–DNA complexes, obtained by MD, the parameters (distances and angles) shown in Fig. 7 were measured on each sampled structure to assess the furan-side and/or the pyrone-side intercalation-adduct. The MD configurations were a priori classified as follows [6,8]:

- CROSS—favourable geometrical arrangements for inter-strand cross-links adducts (all the monitored distances are in the range 3.2–4.0 Å and all the angles in the range 60–120°);
- FUR—favourable geometrical arrangements for furan-side monoadducts (the distances dF1, dF2 are in the range 3.2–4.0 Å, while the distances dP1, dP2 more than 4.0 Å and the angles aF1, aF2 in the range 60–120°);
- PYR—favourable geometrical arrangements for pyrone-side monoadducts (the distances dP1, dP2 are in the range 3.2–4.0 Å, while the distances dF1, dF2 are more than 4.0 Å and the angles aP1, aP2 in the range 60–120°).

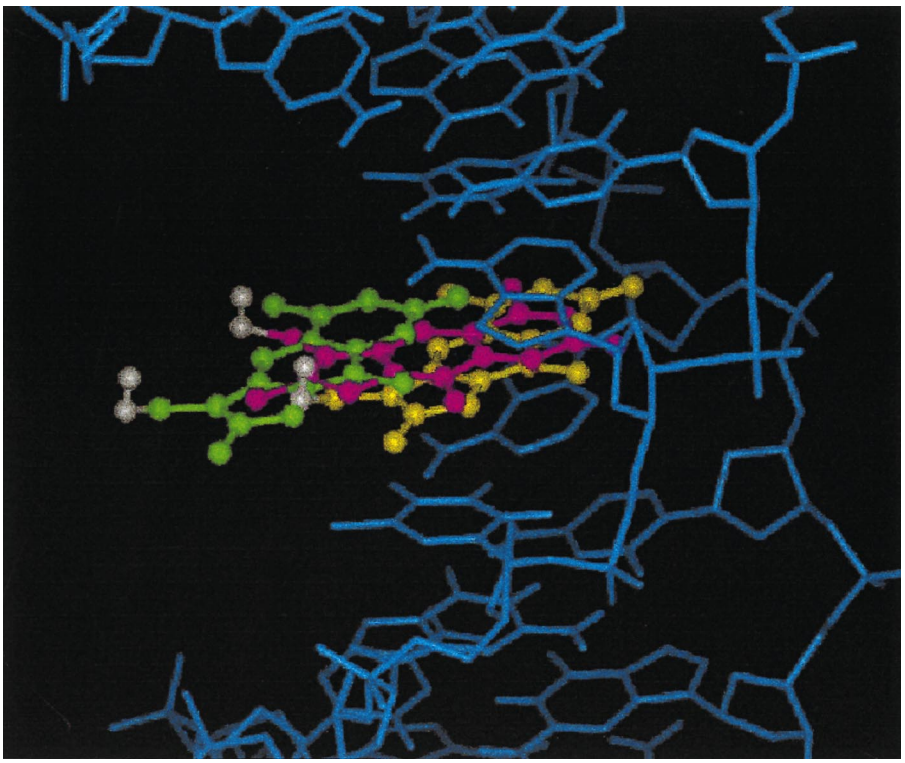


Fig. 6. The HMT moiety has been graphically incorporated between the 4th and the 5th base pairs in three different ways: (A) in which HMT is located and oriented in a quite good way to promote the full intercalation inside the DNA (yellow); (B) in which HMT is oriented as in A, but the psoralen residue is located 5 Å out of the major groove (green); (C) in which HMT is located as in A, but oriented in a perpendicular way with respect to A (pink).

The configurations were assigned to the appropriate class only if all the cited geometrical criteria were satisfied; on the contrary the configurations were unassigned, and named OTHER. The results of this classification are reported at the beginning of Tables 4 and 5.

The clustering works according to a geometrical criterion (in this case the distances and angles already defined) coupled with energetical considerations. In this work the non-bonded inter-molecular energy between the oligonucleotide and the psoralen (as computed by the Anal module of AMBER) was considered:

$$E_{\text{interaction}} = E_{\text{complex}} - E_{\text{DNA}} - E_{\text{psoralen}} \quad (1)$$

Three different types of configurations are defined as follows:

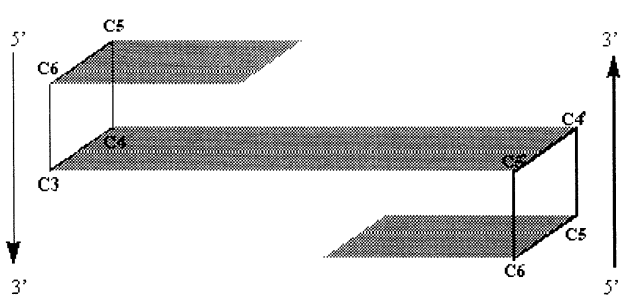
– the i-th configuration is defined as 'representative' if the following criterion is satisfied:

$$E_i < E_j, \ \forall \text{object } j \ (j \neq i) \text{ with } \mathrm{d}x_{ij} < \mathrm{d}x^\circ \quad (2)$$

where $\mathrm{d}x_{ij}$ is the difference in the geometrical parameters x between the i-th and j-th objects and $\mathrm{d}x^\circ$ is the corresponding user-defined threshold, whereas E_i and E_j are the corresponding energies. In the above expression, the geometrical criteria must be satisfied for each geometrical parameter x (in our case x stands for any distance and angle reported in Fig. 7);

– the j-th configuration is defined 'collapsed' if at least one representative configuration i $(i \neq j)$ exists such that:

$$E_j > E_i \ \ \& \ \ \mathrm{d}x_{ij} < \mathrm{d}x^\circ \quad (3)$$



Distances, d's in Å, angles, a's in degrees.

Furan-side	**dF1: C4'-C5**
	dF2: C5'-C6
	aF1: C5-C4'-C5'
	aF2: C5'-C6-C5
Pyrone-side	**dP1: C3-C6**
	dP2: C4-C5
	aP1: C3-C4-C5
	aP2: C3-C6-C5

Fig. 7. Definition of geometrical parameters used for the intercalation-adducts classification (see text). The large slab represents the psoralen while the smaller ones are for thymine bases.

Table 4
DNA-molecule **1** docking analysis by FRAMES [18]

A priori class analysis

Class name	No. of matching configurations
FUR	166
PYR	11
CROSS	11
OTHER	10

Clustering report

Configuration type	No.	%
Representative	12	6.06
Collapsed	147	74.24
Uncollapsed	39	19.70
Total	198	100.00

Report of representative configurations [a]

Entry	*N*	Class	*E*	*d*F1	*d*F2	*d*P1	*d*P2	*a*F1	*a*F2	*a*P1	*a*P2
1	22	FUR	0.000	3.6	3.4	4.2	4.1	76.5	85.1	89.6	88.2
2	15	CROSS	1.840	3.8	3.7	3.8	3.8	92.0	96.3	119.1	117.2
3	14	FUR	2.530	4.0	3.9	4.2	4.1	95.9	98.0	128.0	116.3
4	9	FUR	5.180	4.0	3.8	4.2	4.3	88.8	97.1	100.5	104.2
5	37	FUR	5.970	3.7	3.5	4.3	4.8	70.1	78.3	57.5	79.0
6	21	PYR	6.260	4.4	4.2	3.9	3.8	87.4	98.3	109.6	103.4
7	11	FUR	6.260	3.8	3.6	4.5	4.6	81.0	88.5	82.5	85.8
8	17	PYR	6.660	4.6	4.4	3.9	3.9	99.6	106.8	119.5	118.1
9	32	FUR	7.160	3.8	3.7	4.5	5.1	68.0	70.1	57.9	80.6
10	38	FUR	7.600	3.9	3.6	4.1	4.4	67.7	79.5	62.5	73.4
11	48	FUR	7.620	3.7	3.5	4.1	4.4	73.2	80.9	62.0	76.4
12	78	OTHER	9.330	4.0	3.7	5.0	5.7	59.2	72.2	47.6	78.1

[a] *N*, number of the sampled configuration (MD run); *E*, inter-molecular energy (kcal/mol); *d*, distances (Å); *a*, angles (°) (see Fig. 7).

– configurations defined as 'uncollapsed' are those satisfying neither Eq. (2) nor Eq. (3).

This classification-clustering technique is based on a logical criterion and it uses the energy as a driving force to agglomerate objects (configurations) into clusters. It allows us to evidence the highest significative configurations avoiding the agglomerative union of objects (configurations) located outside the user-defined geometrical criteria.

The configurations are joined in a cluster after the detection of one representative. Fig. 8 illustrates this agglomeration procedure in which a hypothetical molecular energy surface is reported versus a generic geometrical parameter (x).

The **R** labels indicate the representative configurations, which do not have any other configurations at lower energy close to them according to the defined geometrical criteria. By **C** are named the collapsed ones, which are close (from a geometrical point of view) to a representative configuration at lower energy. Finally, **U** indicates the uncollapsed, which are close to a collapsed configuration, but do not satisfy the criteria to be assigned to representative configurations.

Following the above rules, the intercalation adducts are assigned to a cluster in such a way that each cluster contains one representative and the fall-down collapsed configurations.

For what concerns molecule **1,** we can observe (see Table 4) that all the possible a priori classes of reversible intercalation-adducts were sampled (Figs. 7 and 9). The minimum energy (inter-molecular energy) configuration is a FUR, while a CROSS is 1.8 kcal/mol above. A favourable geometry for pyrone-side monoadducts (PYR) has been observed but at higher energy (6.3 kcal/mol above).

The obtained results are in agreement with the experimental data:

- The FUR retrieval as global minimum orientation is justified by the use of the furan-side monoadduct structure as starting geometry.
- Nevertheless the bifunctionality of HMT is confirmed by both CROSS and PYR reversible configurations retrieval.

Results regarding the docking between DNA and molecule **2** are in agreement with the biological data. In fact, as one can see in Table 5, only FUR geometry

Table 5
DNA-molecule **2** docking analysis by FRAMES [18]

A priori class analysis

Class name	No. of matching configurations
FUR	52
PYR	0
CROSS	0
OTHER	146

Clustering report

Configuration type	No.	%
Representative	13	6.57
Collapsed	170	85.86
Uncollapsed	15	7.58
Total	198	100.00

Report of representative configurations [a]

Entry	*N*	Class	*E*	*d*F1	*d*F2	*d*P1	*d*P2	*a*F1	*a*F2	*a*P1	*a*P2
1	13	FUR	0.000	3.5	4.0	6.1	5.5	83.3	60.8	125.2	94.4
2	134	FUR	0.110	3.6	3.9	6.7	6.0	77.2	65.0	131.2	95.6
3	58	OTHER	0.140	3.5	4.3	6.7	6.0	88.5	59.2	123.3	93.1
4	67	FUR	0.340	3.7	4.0	7.0	6.2	78.7	64.3	133.9	93.8
5	59	FUR	0.720	3.5	4.0	6.4	5.8	84.9	61.7	127.4	95.6
6	49	OTHER	0.920	4.0	4.2	7.0	6.4	84.2	76.5	130.0	101.7
7	18	FUR	1.830	3.8	3.7	5.8	5.0	67.3	68.2	132.9	89.6
8	35	OTHER	2.200	3.5	4.3	7.1	6.4	90.6	55.5	126.7	95.1
9	30	OTHER	2.760	3.6	4.8	6.6	6.0	100.5	52.6	123.7	93.2
10	29	OTHER	3.420	4.0	4.1	6.8	6.0	74.8	73.8	138.1	94.6
11	31	FUR	4.770	3.8	3.9	6.1	5.3	70.1	69.3	127.3	89.9
12	125	OTHER	5.380	3.4	4.1	7.1	6.7	85.5	55.6	107.8	89.5
13	147	OTHER	7.430	3.4	4.1	5.3	4.6	81.2	55.8	119.2	86.4

[a] *N*, number of the sampled configuration (MD run); *E*, inter-molecular energy (kcal/mol); *d*, distances (Å); *a*, angles (°) (see Fig. 7).

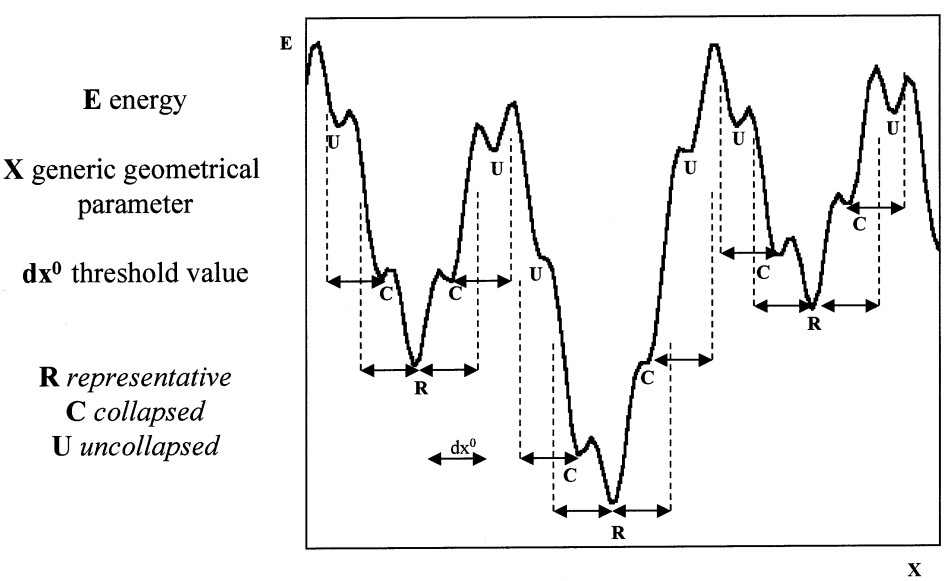


Fig. 8. FRAMES agglomeration procedure. A hypothetical molecular energy surface is reported vs. a generic geometrical parameter (*x*).

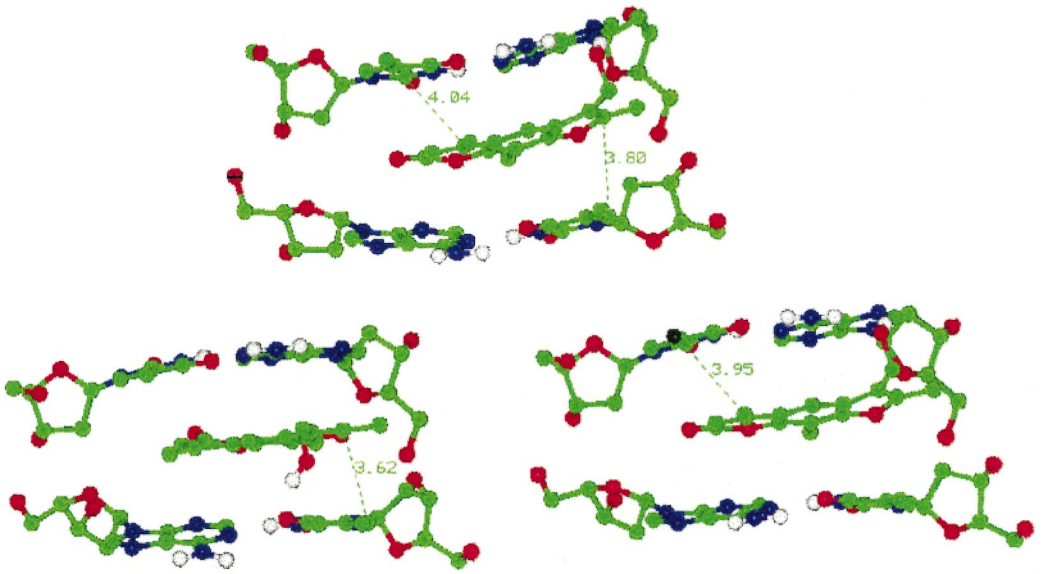


Fig. 9. The three intercalation-adducts: (top) CROSS geometry favourable for inter-strand cross-links; (bottom) FUR geometry favourable for furan-side monoadducts (right) and PYR geometry favourable for pyran-side monoadducts (left).

has been sampled during the MD simulation while CROSS and PYR ones were never retrieved. This fact seems to confirm the hypothesis made by Gia et al. [8] on the basis of the photobiological data according to which the introduction of a methyl group in position 3 on the pyrone ring gives rise to monofunctional agents prevalently able to photoreact on the furan-side.

4. Conclusion

Since the first step of the dark DNA-psoralens intercalation seems to determine the subsequent photoreactions, our methodology represents a computational tool to partially understand the mechanism of action of these photochemotherapeutic agents.

In fact, even if there is the possibility for molecules, which exhibit unfavourable geometry during the dark intercalation, to undergo the cycloaddition to the DNA during the light activation, favourable orientations determine a much greater probability of photoreaction.

We can conclude that our theoretical procedure is able to describe the different dark behaviour of the bifunctional agent **1** and of the monofunctional **2** and it could not only be used for other analogue compounds, but also represent a preliminary approach towards the photoactivated DNA–furocoumarins molecular recognition.

Acknowledgements

The Italian funding agencies of MURST and CNR are gratefully acknowledged for financial support. We wish to thank Dr Richard Lavery (Laboratoire de Biochimie Théorique, CNRS, Paris) for having allowed us to use the package CURVES 5.1.

References

[1] J.A. Parrish, R.S. Stern, M.A. Pathak, T.B. Fitzpatrick, Photochemotherapy of skin diseases, in: J.D. Regan, J.A. Parrish (Eds.), The Science of Photomedicine, Plenum, New York, 1982, pp. 595–624.
[2] F.P. Gasparro, Extracorporeal Photochemotherapy: Clinical Aspects and the Molecular Basis for Efficacy, R.G. Landes, Austin, TX, 1994.
[3] R.P. Goodrich, N.R. Yerram, B.H. Tay-Goodrich, P. Forster, M.S. Platz, C. Kasturi, S.C. Park, J.N. Aebischer, S. Rai, L. Kulaga, Selective inactivation of viruses in the presence of human platelets: UV sensitisation with psoralen derivatives, Proc. Natl. Acad. Sci. USA 91 (1994) 5552–5556.
[4] R.S. Stern, R. Lange, Non-melanoma skin cancer occurring in patients treated with PUVA five to ten years after first treatment, J. Invest. Dermatol. 91 (1988) 120–124.
[5] J.W. Tessman, S.T. Isaacs, 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 24 (1985) 1669–1676.
[6] J. Demaret, J. Ballini, P. Vigny, Molecular mechanics and dynamics study of DNA–furocoumarins complexes: effect of methylation of the angular derivatives on the intercalation geometry, J. Comput. Aided Mol. Des. 7 (1993) 683–698.
[7] P.H. Spielmann, T.J. Dwyer, J.E. Hearst, D.E. Wemmer, Solution structures of psoralen monoadducted and cross-linked DNA oligomers by NMR spectroscopy and restrained molecular dynamics, Biochemistry 34 (1995) 12937–12953.
[8] O. Gia, A. Anselmo, A. Pozzan, C. Antonello, S. Marciani Magno, E. Uriarte, Some new methyl-8-methoxypsoralens: synthesis, photobinding to DNA, photobiological properties and molecular modelling, Farmaco 52 (1997) 389–397.

[9] MacroModel, Version 5.5, Columbia University, New York, 1996.
[10] M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M.W. Gill, B.G. Johnson, M.A. Robb, J.R. Cheeseman, T.A. Keith, G.A. Petersson, J.A. Montgomery, K. Raghavachari, M.A. Al-laham, V.G. Zakrzewski, J.V. Ortiz, J.B. Foresman, J. Cioslowski, B.B. Stefanov, A. Nanayakkara, M. Challacombe, C.Y. Peng, P.Y. Ayala, W. Chen, M.W. Wong, J.L. Andres, E.S. Replogle, R. Gomperts, R.L. Martin, D.J. Fox, J.S. Binkley, D.J. Defrees, J. Baker, J.P. Stewart, M. Head-Gordon, C. Gonzalez, J.A. Pople, GAUSSIAN-94 (Revision A.1), Gaussian Inc., Pittsburgh, PA, 1995.
[11] S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona, S. Profeta Jr., P. Weiner, AMBER, Version 3.0, J. Am. Chem. Soc. 106 (1984) 765–784.
[12] S. Arnott, P.J. Campbell-Smith, R. Chandrasekaran, Atomic coordinates and molecular conformations for DNA–DNA, RNA–RNA and DNA–RNA helices, in: G.D. Fasman (Ed.), CRC Handbook of Biochemistry and Molecular Biology, vol. II, CRC Press, Cleveland, OH, 1975, pp. 411–423.
[13] T.E. Cheatham III, P.A. Kollman, Observation of the A-DNA to B-DNA transition during unrestrained molecular dynamics in aqueous solution, J. Mol. Biol. 259 (1996) 434–444.
[14] B. Tidor, K.K. Irikura, B.R. Brooks, M. Karplus, Dynamics of DNA oligomers, J. Biom. Struct. Dyn. 1 (1983) 231–252.
[15] J. Srinivasan, J.M. Withka, D.L. Beveridge, Molecular dynamics of an in vacuo model of duplex d(CGCGAATTCGCG) in the B-form based on the AMBER 3.0 force field. 58 (1990) 533–547.
[16] U.C. Singh, S.J. Weiner, P. Kollman, Molecular dynamics simulations of d(CGCGA) × d(TCGCG) with and without 'hydrated' counterions, Proc. Natl. Acad. Sci. USA 82 (1985) 755–759.
[17] R. Lavery, H. Sklenar, CURVES, Version 5.1, Laboratoire de Biochimie Théorique, CNRS, Paris, 1996.
[18] O. Incani, FRAMES: Conformational Analysis Tools, unpublished material.