UNIVERSITAT POLITÈCNICA DE CATALUNYA
DEPARTAMENT D'ENGINYERIA QUÍMICA

APPLICATION OF MOLECULAR SIMULATION TECHNIQUES TO THE DESIGN OF NANOSYSTEMS

Francisco Rodríguez Ropero

Advisors: Dr.Carlos Alemán Llansó y Dr. David Zanuy Gómara

Barcelona, September 2009

OBJECTIVES

(1) Improve the stability of tubular nanodevices based on highly regular β-helical motifs excised from naturally occurring proteins by replacing targeted natural residues by conformationally constrained synthetic amino acids.

(2) Model and characterize a dendronized polymethacrylate carrying second generation chiral 4-aminoproline-based dendrons and a thiophene based dendronized polymer in order to explain, in basis of their structure, the extraordinary stiffness of the first and the great conductivity exhibited by the latter in previous experimental studies.

(3) Study of the interactions of thiophene based conducting polymers. The interaction of these polymers with the solvent medium as well as their intermolecular interactions such as π-stacking interactions play a key role when designing new devices based on conducting polymers.

(4) Exploit the intrinsic tubular geometry of natural motifs extracted from certain proteins together with the ability of π-conjugated systems to promote charge transfer to model a tubular electronic conductor based on natural peptide sequences by replacing natural amino acids by synthetic amino acids owning an aromatic side chain.

(5) Study of the role played by the solvent molecules in the actuation mechanism in the molecular actuator poly(calix[4]arene bis-bithiophene) by the examination of the structure and dynamics of the dichloromethane solvent around the calix[4]arene units.

(6) Investigation of the selectivity of a polythiophene functionalized with calix[4]arene nanosensor at the microscopic level which experimentally presents a great selectivity towards Na^+ with respect to K^+ and Li^+ alkaline cations.

ABBREVIATIONS

1hv9	residues 296-329 of N-acetylglucosamine 1-phosphate uridyltransferase GlmU, C-terminal domain from *Eschericia coli*
1krr	Residues 131-165 of the chain A from the *Escherichia coli* galactoside acetyltransferase
Ac_3c	1-aminocyclopropane-1-carboxylic acid
Ac_4c	1-aminocyclobutane-1-carboxylic acid
Ac_5c	1-aminocyclopentane-1-carboxylic acid
Ac_6c	1-aminocyclohexane-1-carboxylic acid
Ac_nc	1-aminocycloalkane-1-carboxylic acids
Aib	α-aminoisobutyric acid
AIBN	2,2'-azobisisobutyronitrile
AIMI	Aromatic Intermolecular Interaction Model
AM1	Austin Model number 1
B3LYP	Becke's three parameter hybrid functional with the Lee, Yang and Parr expression.
Boc	*tert*-butoxycarbonyl
BSSE	Basis Set Superposition Error
c_3Dip	1-amino-2,2-diphenyl-cyclopropanecarboxylic acid
CC	Coupled Cluster
CCSD(T)	Coupled cluster singles, doubles and triples method being the contribution of triple excitations represented in an approximate way
CD	Circular Dichroism
CP	Counterpoise
DFT	Density Functional Theory
DSC	Differential Scanning Calorimetry
Dϕg	Diphenylglycine
ESP	Electrostatic potential
FF	Force Field
GGA	Generalised Gradient Approximation
$\hat{H}$	Hamiltonian operator
HF	Hartree Fock
HOMO	Highest Occupied Molecular Orbital
IP	Ionization Potential

KT	Koopman's Theorem
LCAO	Linear Combination of Atomic Orbitals
LDA	Local Density Approximation
LSCF	Local Self-Consistent Field
LSDA	Local Spin Density Approximation
LUMO	Lowest Unoccupied Molecular Orbital
MD	Molecular Dynamics
MEP	Molecular Electrostatic Potential
MGn	n^{th} Generation monomer (dendron)
MM	Molecular Mechanics
MP	Møller Plesset
NA9PH	9-(2,3-dihydro-2-phenyl-1H-benz[e]inden-1-ylidene)-9H-fluorene
NA9PH(S)	NA9PH Singlet state
NA9PH(T)	NA9PH Triplet state
NPT	Isobaric-isothermal ensemble
NVT	Canonical ensemble
OR	Optical Rotation
PCM	Polarizable Continuum Model
PDB	Protein Data Bank
PG_n	n^{th} Generation polymer (dendronized polymer)
PM3	Parameterized Model number 3
PME	Particle Mesh Ewald
PMF	Potential Mean Force
PMMA	Poly(methyl methacrylate)
QM	Quantum Mechanics
QM/MM	Quantum mechanics/Molecular mechanics
RESP	Restrained Electrostatic Potential
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
r.u.	Repeat Unit
SCF	Self Consistent Field
SCRF	Self Consistent Reaction Field
SDTQ	Single, Double, Triple and Quadruple Excitations

SE	Semiempirical
SLBO	Strictly Localized Bond Orbitals
STO	Slater Type Orbital
TD	Time-dependent
Tie	β-3-thienylalanine
UB3LYP	Unrestricted B3LYP
UHF	Unrestricted Hartree Fock
UMP2	Unrestricted MP2
UPM3	Unrestricted PM3
wt	Wild type
ZPVE	Zero-Point Vibrational Energies
βAla	β-Alanine
ΔAla	Dehydroalanine

TABLE OF CONTENTS

1. INTRODUCTION

1.1. A brief description of nanotechnology: history and current state

Nanotechnology is a multidisciplinary branch of science and technology that involves a wide range of different fields such as chemistry, materials science, physics or chemical engineering whose goal is the production of new functional materials and devices through the control of their organization at the atomic and molecular scale.

What differences nanotechnology from traditional chemistry and physics can be summarized in three points:[1]

- Analysis and control of the matter at the atomic and molecular level focusing in individual atoms.
- The appearance of novel physical properties because of the nanoscopic dimensions.
- The possibility of generating new complex functional systems with novel properties.

Professor Norio Taniguchi defined for the first time the term "Nanotechnology" in 1974 as the production technology to get the extra high accuracy and ultra fine dimensions, i.e. the preciseness and fineness on the order of 1 nanometer.[2] But the concept of nanotechnology is much older. Specifically, it dates back 50 years ago when the Nobel laureate Richard Feynman gave his famous lecture entitled "There's Plenty of Room at the Bottom" at the annual meeting of the West Coast Section of the American Physical Society held at the California Institute of Technology. In such lecture Richard Feynman envisioned the possibility of building extremely small robots, information storage devices or motors by maneuvering individual atoms.[3] In spite of the revolutionary ideas exposed by Richard Feynman it was not until late 1970s and early 1980s when nanotechnology gained a certain attention in the scientific community.[4]

Dr. K. Eric Drexler is considered the pioneer of modern molecular nanotechnology. He borrowed the word from Prof. Taniguchi and the primordial ideas of Prof. Feynman to explain the possibility of building machines on the scale of molecules, nanomotors, robot arms or nanocomputers.[5] Drexler advanced the proposal that the molecular machinery found in living systems demonstrates the feasibility of doing advanced molecular engineering to produce complex, artificial molecular machines.[6]

A weighty milestone for the development of nanotechnology was the invention of the Scanning Tunneling Microscope (STM) in 1981 by Nobel laureates Gerd Binning and Heinrinch Rohrer. STM afforded the possibility of viewing individual atoms for the

very first time.[7] Many progresses were done in the following years such as the discovery of the buckminsterfullerene molecule,[8] the invention of the Atomic Force Microscope (AFM),[9] the first engineered protein[10] or the finding of the carbon nanotubes.[11]

All these advances in molecular engineering encouraged scientists to set up formal organizations, being the first one the Foresight Institute founded by K. Eric Drexler and Christine Peterson in 1986 in Palo Alto, California. Foresight Institute together with the Global Business Network sponsored the first comprehensive conference on nanotechnology held in Palo Alto in October 1989 where a broad range of technologies that contribute to nanotechnology as well as several perspectives for the future were presented.[12]

In early 90s it was clear that nanotechnology would play a key role in the development of the future world and first national funding programs in nanotechnology arose in Japan[13] and in the USA,[14] which summed up $185 million into a project over 10 years and $40 million respectively. Investment in nanotechnology has increased exponentially in the last years. According to the "The Global Market for Nanotechnology" edited in November 2008 by Nanoposts (University of Stirling, Innovation Park) worldwide investment in nanotechnology research in 2008 was approximately $12.3 billion, $6.7 billion coming from public funding programs, being the U.S.A., the E.U. and Japan on the top of the list. In this sense the 2009 North American Budget provides $1.5 billion for the National Nanotechnology Initiative (NNI).[15] The Seventh Research Framework Program (FP7) of the E.U. includes specific areas devoted to nanosciences and nanotechnologies with a €3.5 billion portion of the budget to be invested from 2007 to 2013.[16] Public investment in Japan rose up to $1.06 billion in 2007.[17]

The enormous growth achieved by nanotechnology is clearly reflected in the amount of scientific papers published in the last years. A broad bibliographical research in ISI Web of Science of papers dealing with the topic "nanotechnology" reflects the following convincing results: less than 700 publications up to year 2000 and around 10000 publications in the present decade. All the progresses made in nanotechnology research have been translated into direct applications in different fields being electronics and computer science[18] and biomedicine,[19] where the most striking advances have been done.

The current trend is the mixture of nanotechnology and supramolecular chemistry, i.e. using the preciseness of the first with concepts such as hydrogen bonding, van der Waals interactions or π-π interactions –leading to molecular self-assembly– of the latter to build more complex structures.[20]

1.2. Presence of nanotechnology: applications

As previously stated nanotechnology has jumped from research laboratories to our daily life and today, according to the Project on Emerging Nanotechnologies organized by the Pew Charitable Trust and the Woodrow Wilson International Center for Scholars, the number of products on the market that incorporate nanotechnology is steadily increasing. At present more than 800 manufacturer-identified nanotechnology-based consumer products are available on the market. They include numerous cosmetics, electronic devices, hardware utilities, silver based anti-microbial products, coatings or food additives just to mention few examples.[21] Let us briefly summarize some actual trends and uses of nanotechnology in some fields.

Electronics and computer science

The development of nanoscale molecular-electronics together with new nanolithography techniques has contributed enormously to construct nanoscale circuits. In this respect research in crossbar switch based electronics mostly made from carbon nanotubes and nanowires building blocks has enabled the appearance of ultra high density data storage devices.[18,22,23] Likewise, the advances in semiconductor nanowires have helped the improvement of complementary metal oxide semiconductors (CMOS), which are used in microprocessors, microcontrollers, static RAM and various digital logic circuits.[23] In this sense hybrid crossbar/CMOS circuits have been vastly developed due to the ultrahigh device density of crossbar structure and the flexibility offered by CMOS circuitry.[18] The most spread hybrid/CMOS circuit is the CMOS/nanowire/molecular hybrid (CMOL).[24] CMOL memory with 32 nm CMOS technology can store 1 terabit of data in a 2cm x 2cm device.[25] Recently high-density crossbar arrays based on Si memristive system have been developed, which is useful for ultrahigh performance memory and logic applications.[26]

Biomedicine

The potential applications of nanotechnology for diagnosis, prevention, and treatment of diseases are currently very broad. Nanotechnology applied to medical diagnosis has a potentially higher sensitivity and selectivity than classical methods.[19] Regarding diagnosis several improvements have been done up to date like the development of antibody-based microfluidics system to detect different cancers,[27] microfluidic chip cells for the detection of allergic responses,[28] nanosensors,[29] quantum dots for investigating cells or low abundance antigens[30] or metallic and semiconductor nanoclusters and nanopowders as markers in nuclear imaging techniques.[31] Nanotechnology has greatly aided the development of new drug delivery systems such as nanovectors,[32] liposomal drug carriers,[33] dendronized polymers[34] or other polymer based delivery systems.[35] Nanotechnology has also enhanced the scope of regenerative medicine. Some examples of the utilization of nanotechnology in the manipulation of stem cells,[36] in tissue engineering[37] or regeneration within the central nervous system[38] are reported.

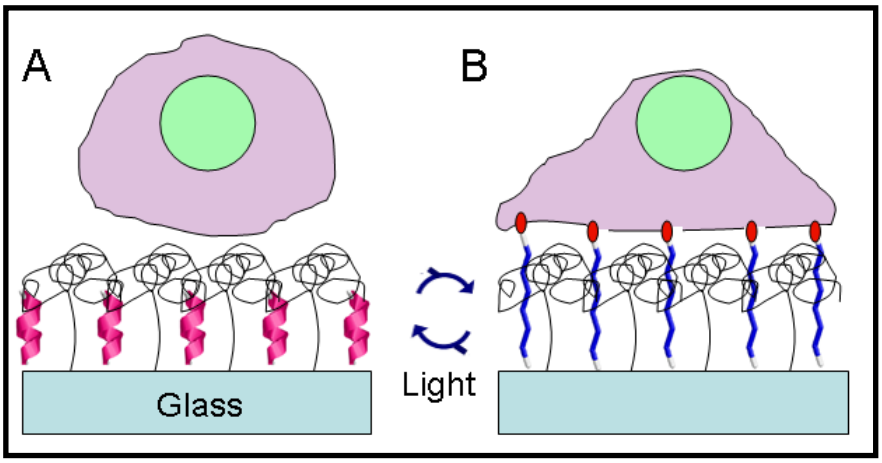


Figure 1. Control of cell adhesion through changes in the conformation of peptide foldamers.

Other applications

The bearing of nanotechnology applications is extremely wide and ranges – besides electronics and biomedicine – from catalysis[39] to cosmetics.[40] Under a catalyst point of view the two major improvements stemmed from nanotechnology are the use of supported metal nanoparticles and the tuning of the catalyst surface at the atomic level.

These solutions have helped the development of high activity and selectivity catalysts.[39] Nanotechnology also offers new horizons and possibilities in water treatment. Development of nanostructured catalytic membranes, nanosorbents, nanocatalysts and bioactive nanoparticles for treatment, purification and disinfection of surface water, groundwater and water contaminated by toxic metal ions is widely reported in the literature.[41] Beyond all these previously stated applications nanotechnology has an incipient presence in the development of machine tools, insulators, motor vehicles and aircrafts, batteries, high resolution screens, light-emitting diodes or molecular actuators, just to mention some examples.[42]

1.3. Role of Computational Chemistry in the development of nanodevices

Modeling and theory are becoming vital to designing and improving nanodevices. The intrinsic nature of nano and supramolecular scale that involves tens, hundreds and thousands of atoms makes computational chemistry the perfect ally to design new devices and predict their properties. Computational chemistry provides the perfect tools to describe the electronic structure and the dynamic behavior as well as the properties derived from them through quantum mechanics and classical mechanics formalisms. Quantum Mechanics (QM) offer the most accurate description of the matter as it describes its fundamental behavior at the atomic and molecular level. The use of QM requires solving Schrödinger equation using some approximations, for instance, *ab initio* or Density Functional Theory (DFT) methods. Because of the computational expensiveness of QM methods, QM is in practice only applicable to small systems, i.e. few dozens of atoms.

Classical mechanics or Molecular Mechanics (MM) concerns the use of simple potential-energy functions such as Coulombic potentials or harmonic oscillators to model molecular systems. In spite of not being as rigorous as quantum mechanics, MM approaches allow the inclusion of a bigger quantity of atoms describing the whole system in a reliable manner. MM approaches are widely applied in molecular structure refinement, Molecular Dynamics (MD) simulations, Monte Carlo simulations and ligand-docking simulations. A more detailed description of QM and MM based methods is included in the chapter devoted to Methodology.

When the number of atoms to be treated in a system becomes too large coarse graining approximations are very useful. These approximations consist in the use of pseudo-

atoms to represent groups of atoms. Pseudo-atoms must be parameterized so that they can reproduce as much as possible the experimental data available or the results derived from all atom classical simulations.

Since the first simulation of a protein in 1976[43] computational chemistry has become an essential tool in protein design,[44] polymer science,[45] nano-science,[46] drug design[47] or drug delivery systems design.[48]

QM calculations are useful to determine properties derived from the electronic structure such as charge density, spin density, electronic affinity, ionization potential or band-gap as well as an insight of supramolecular interactions.[49] QM calculations can also provide the necessary data to set a force field for MM calculations, i.e. the functional form and parameter sets used to define the potential energy of a classical system of particles.[50]

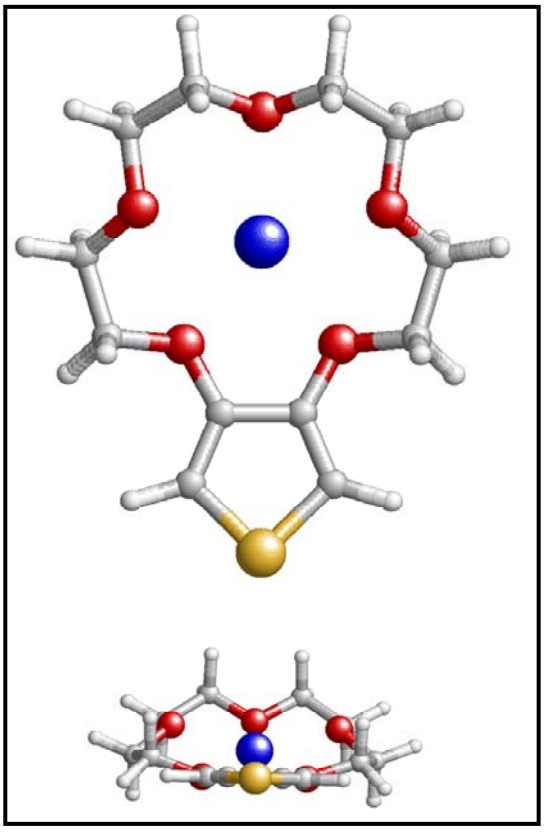


Figure 2. Axial and equatorial projections of a model molecule for a 5-ether functionalized polythiophene based Na^+ nanosensor

The amount of examples available in the bibliography of the utilization of computational chemistry to study nanodevices is really vast so only a brief survey is given, emphasizing in systems that are similar to those presented in this thesis. Several

papers where MD simulations have been employed to model and characterize different kinds of nanotubes can be found in the literature, being carbon nanotubes and their derivatives the most studied ones.[46] Thus, there are studies ranging from the determination of their structure and mechanical properties[51] to the diffusion of different molecules through them or derivatives of them.[46,52] Peptide based nanotubes are gaining a certain attention due to their applicability in biomedicine as drug carriers and nanowires.[53,54,55] *In silico* methods have helped the design of new and more efficient drug and gene delivery systems predicting their required composition, geometry, dimensions and correct dosage.[56] One of the most promising families of drug delivery systems is that concerning dendronized polymers.[57] Both experimental and computational research in nanoactuators and nanosensors are widely reported in the literature. Among them, those based in conducting polymers are flourishing because of their great transport properties, electrical conductivity or rate of energy migration which provide amplified sensitivity in nanosensors and a rapid response in nanoactuators.[58,59] Combining QM and MM based methods provides a good outlook of the mechanisms that rule the sensory and actuation behavior of these devices.[60,61]

All in all, the predictive essence of computational chemistry makes it exceptionally helpful to rapidly make decisions, prototype, improve, innovate and learn from failure, which means saving of time and money in experimental work.

1.4. References

[1] Theis T., Parr D., Binks, P., Ying J., Drexler K.E., Schepers E., Mullis K., Bai C., Boland J.J., Langer R., Dobson, P., Rao C.N.R., Ferrari, M. *Nat. Nanotechnol.*, **2006**, *1*, 8.

[2] Taniguchi N. "On the Basic Concept of Nano-Technology", **1974**, *Proc. Intl. Conf. Prod.* Part II, Japan Society of Precision Engineering, Tokyo.

[3] Feynman R. P. *Engineering and Science*, **1960**, *23*, 22.

[4] Drexler K.E. "Nanosystems: Molecular Machinery, Manufacturing, and Computation", **1992**, John Wiley & Sons Inc., New York.

[5] Drexler K.E. "Engines of Creation: The Coming Era of Nanotechnology", **1986**, Anchor Books, New York.

[6] Drexler K.E. *Proc. Natl. Acad. Sci.*, **1981**, *78*, 5275.

[7] Binning G., Rohrer H., Gerber C., Weibel, E. *Phys. Rev. Lett.*, **1982**, *49*, 57.

[8] Kroto H.W., Heath J.R., O'Brien S.C., Curl R.F., Smalley R.E. *Nature*, **1985**, *318*, 162.

[9] Binning G., Quate C.F., Gerber C. *Phys. Rev. Lett.*, **1987**, *56*, 930.

[10] Ho S.P., DeGrado W.F. *J. Am. Chem. Soc.*, **1987**, *109*, 6751.

[11] Ijima S., Ichihashi T. *Nature*, **1993**, *363*, 603.

[12] Crandall B.C., ed. "Nanotechnology: molecular speculations on global abundance", **1996**, The MIT Press, Cambridge, Massachusets.

[13] Swinbanks D. *Nature*, **1991**, *352*, 650.

[14] Amato I. *Science*, **1991**, *254*, 1302.

[15] "National Nanotechnology Initiative FY2009 Budget and Highlights": http://www.nano.gov/NNI_FY09_budget_summary.pdf

[16] CORDIS: Nanotechnology: EU Funding: http://cordis.europa.eu/nanotechnology/src/eu_funding.htm

[17] Nordan M.M. "Change Required for the National Nanotechnology Initiative as Commercialization Eclipses Discovery", **2008**, Lux Research Inc.

[18] Lu W., Lieber C.M. *Nat. Nanotechnol.*, **2007**, *6*, 841.

[19] Riehemann K., Schneider S.W., Luger T.A., Godin B., Ferrari M., Fuchs H. *Angew. Chem. Int. Ed.*, **2009**, *48*, 872.

[20] Kay E.R., Leigh D. A., Zerbetto F. *Angew. Chem. Int. Ed.*, **2007**, *46*, 72.

[21] Project on Emerging Nanotechnologies, Woodrow Wilson International Center for Scholars. **2009**. A Nanotechnology Consumer Products Inventory: http://www.nanotechproject.org/inventories/consumer

[22] Chen Y., Jung G-Y., Ohlberg D.A.A., Li X., Stewart D.R., Jeppesen J.O., Nielsen K.A., Stoddart J.F., Williams R.S. *Nanotechnology*, **2003**, *14*, 462.

[23] Szaciłowski K. *Chem. Rev.*, **2008**, *108*, 3481.

[24] Strukov D.B., Likharev K.K. *Nanotechnology*, **2005**, *16*, 888.

[25] Strukov D.B., Likharev K.K. *Nanotechnology*, **2005**, *16*, 137.

[26] Jo S.H., Kim K-H., Lu W. *Nano. Lett.*, **2009**, *9*, 870.

[27] Du Z., Colls N., Cheng K.H., Vaughn M.W., Gollahon L. *Biosens. Bioelectron.*, **2006**, *21*, 1991.

[28] Matsubara Y., Murakami Y., Kobayashi M., Morita Y., Tamiya E. *Biosens. Bioelectron.*, **2004**, *19*, 741.

[29] Fritz J., Baller M. K., Lang H. P., Rothuizen H., Vettiger P., Meyer E., Guntherodt H., Gerber C., Gimzewski J. K. *Science*, **2000**, *288*, 316.

[30] Alivisatos A. P., Gu W., Larabell C. *Annu. Rev. Biomed. Eng.*, **2005**, *7*, 55.

[31] Kinge S., Crego-Calama M., Reinhoudt D. N. *ChemPhysChem*, **2008**, *9*, 20.

[32] Sakamoto J., Annapragada A., Decuzzi P., Ferrari M. *Expert Opin. Drug Delivery*, **2007**, *4*, 359.

[33] Torchilin V. P. *Nat. Rev. Drug Discovery*, **2005**, *4*, 145.

[34] Costantino L., Gandolfi F., Bossy-Nobs L., Tosi G., Gurny R., Rivasi F., Vandelli M.A., Forni F. *Biomaterials*, **2006**, *27*, 4635.

[35] Desai N., Trieu V., Yao Z., Louie L., Ci S., Yang A., Tao C., De T., Beals B., Dykes D., Noker P., Yao R., Labao E., Hawkins M., Soon-Shiong P. *Clin. Cancer Res.*, **2006**, *12*, 1317.

[36] Solanki A., Kim J.D., Lee K.B. *Nanomedicine*, **2008**, *3*, 567.

[37] Zhang L.J., Webster T.J. *Nano Today*, **2009**, *4*, 66.

[38] Tysseling-Mattiace V. M., Sahni V., Niece K. L., Birch D., Czeisler C., Fehlings M. G., Stupp S. I., Kessler J. A. *J. Neurosci.*, **2008**, *28*, 3814.

[39] Zecchina A., Groppo E., Bordiga S. *Chem. Eur. J.*, **2007**, *13*, 2440.

[40] Nohynek G.J., Lademann J., Ribaud C., Roberts M.S. *Crit. Rev. Toxicol.*, **2007**, *37*, 251.

[41] Theron J., Walker J.A., Cloete T.E. *Crit. Rev. Microbiol.*, **2008**, *34*, 43.

[42] Wilson M., Kannangara K., Smith G., Simmons M., Raguse B. "Nanotechnology: Basic Science and Emerging Technologies", **2002**, Chapman & Hall/CRC, Boca Raton.

[43] McCammon J.A., Gelin B.R., Karplus M. *Nature*, **1977**, *267*, 585.

[44] Adcock S.A., McCammon J.A. *Chem. Rev.*, **2006**, *106*, 1589.

[45] Kremer K. *Macromol. Chem. Phys.*, **2003**, *204*, 257.

[46] Rafii-Tabar H. *Phys. Rep.-Rev. Sec. Phys. Lett.*, **2004**, *390*, 235.

[47] Jenwitheesuk E., Horst J.A., Rivas K.L., Van Voorhis W.C., Wesley C., Samudrala R. *Trends Pharmacol. Sci.*, **2008**, *29*, 62.

[48] Arifin D.Y., Lee L.Y., Wang C.H. *Adv. Drug Deliv. Rev.*, **2006**, *58*, 1274.

[49] Min S.K., Lee E.C., Lee H.M., Kim D.Y., Kim D., Kim K.S. *J. Comput. Chem.*, **2007**, *29*, 1208.

[50] Cornell W.D., Cieplak P., Bayly C.I., Gould I.R., Merz K.M. Jr., Ferguson D.M., Spellmeyer D.C., Fox T., Caldwell J.W., Kollman P.A. *J. Am. Chem. Soc.*, **1995**, *117*, 5179.

[51] Alexiadis A., Kassinos S. *Chem. Rev.*, **2008**, *108*, 5014.

[52] Liu Y.C., Wang Q. *Phys. Rev. B*, **2005**, *72*, 085420.

[53] Teranishi M., Okamoto H., Takeda K., Nomura K., Nakano A., Kalia R.K., Vashishta P., Shimojo F. *J. Phys. Chem. B*, **2009**, *113*, 1473.

[54] Haspel N., Zanuy D., Alemán C., Wolfson H., Nussinov R. *Structure*, **2006**, *14*, 1137.

[55] Haspel N., Zanuy D., Zheng J., Alemán C., Wolfson H., Nussinov R. *Biophys. J.*, **2007**, *93*, 245.

[56] Haworth I.S. *Adv. Drug Deliv. Rev.*, **2006**, *58*, 1271.

[57] Gibson S.E., Rendell J.T. *Chem. Commun.*, **2008**, 922.

[58] Swager, T.M. *Acc. Chem. Res.*, **1998**, *31*, 201.

[59] Otero T.F., Cortés M. *Adv. Mater.*, **2003**, *15*, 279.

[60] Casanovas J., Zanuy D., Alemán, C. *Angew. Chem. Int. Ed.*, **2006**, *45*, 1103.

[61] Casanovas J., Preat J., Zanuy D., Alemán C. *Chemistry Eur. J.*, **2009**, *15*, 4676.

2. METHODOLOGY

2.1. Computational chemistry

All the work included in the present thesis has been performed using computational methods. Computational chemistry can be roughly defined as a set of theoretical methodologies for investigating chemical problems on a computer. Thus progresses in computational chemistry have befallen closely with the headways in computer science. As previously stated computational chemistry faces chemical problems according to two formalisms: QM and MM. QM calculations require a big amount of CPU time and, since it increases exponentially with the size of the system, only small systems are affordable by using QM formalism. Bigger systems are studied using MM methods.

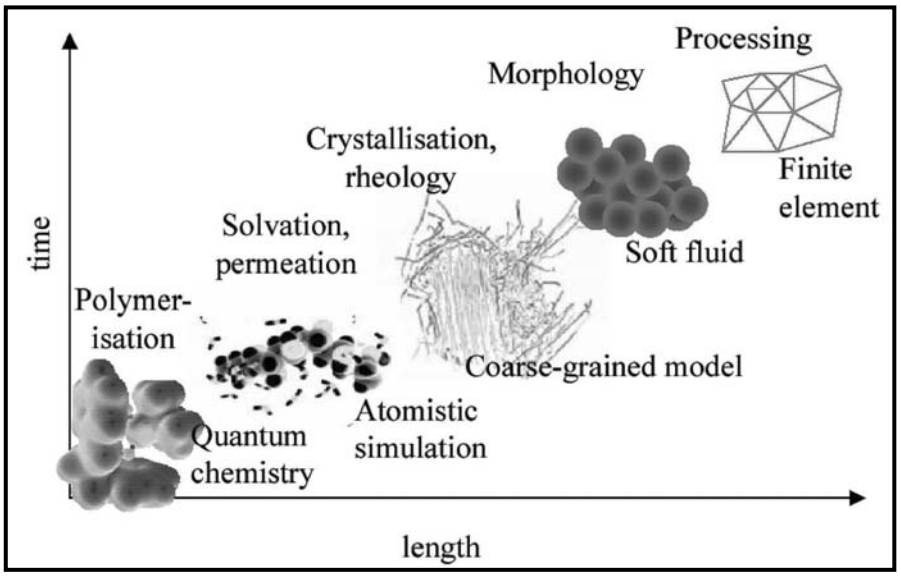


Figure 1. Qualitative CPU time needed to study *in silico* different systems with respect to their size

Since each work presented in this thesis contains a specific section devoted to the methods, only a general overview of QM based methods and MM is presented in this chapter. A deeper description of these methods can be found in many text books, for instance "Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics" by Errol Lewars.[1]

2.2. Quantum Mechanical Methods

Quantum mechanics deals with the motion of electrons under the influence of the electromagnetic force exerted by nuclear charges. Most QM methods assume that nuclei and electron are distinguished from each other and both electron-electron and electron-nucleus interactions are explicitly described. These interactions, which are ruled by nuclear and electron charges and electron motions, determine the spatial distribution of nuclei and electrons and their energies. Electrons are considered as wave-like particles, which are represented by a set of wave functions obtained by solving the Schrödinger equation (equation 2.1.):

$$\hat{H}\,\psi = E\,\psi \qquad (2.1.)$$

where $\hat{H}$ is the Hamiltonian operator containing the kinetic and potential energy of the nuclei and electrons, E is the energy of the system and ψ is the wave function that describes the molecular orbitals.

Solving a QM problem implies solving the Schrödinger equation, which is only affordable by using some approximations. In this sense *Ab initio* calculations, semiempirical calculations and calculations based in the Density Functional Theory (DFT) are extensively widespread to solve QM problems.

2.2.1. *Hamiltonian operator*

Hamiltonian ($\hat{H}$) is defined as the operator that contains all the operations associated with the kinetic and potential energies. The general expression of the Hamiltonian for a system with M nuclei and N electrons described respectively by position vectors R_A and r_i is given in equation 2.2.

$$\hat{H} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>1}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \qquad (2.2.)$$

where M_A is the relation of the mass of the nucleus A with respect to the mass of the electron, Z_A is the atomic number of the nucleus A and ∇_i^2 and ∇_A^2 are the Laplace operators referred to the differentiation between the coordinates of the electron i and

nucleus *A* respectively. In equation 2.2. the first two terms are related to the kinetic energy of the electrons and nuclei respectively, the third term defines the electrostatic attraction between nuclei and electrons and the last two terms are, respectively, the electrostatic repulsion between electrons and nuclei. Taking into account the Born-Oppenheimer approximation, which states that nuclei in a molecule are stationary with respect to the electrons because of their higher mass, the second and fifth terms of equation 2.2. can be neglected. These simplifications lead to the so called electronic Hamiltonian (equation 2.3.).

$$\hat{H}_{el} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>1}^{N} \frac{1}{r_{ij}} \qquad (2.3.)$$

2.2.2. *Basis sets*

Molecular orbitals can be represented as linear combinations of m basis functions (equation 2.4.).[2]

$$\psi_i = \sum_{s=1}^{m} c_{si}\varphi_s \qquad (2.4.)$$

where ψ_i is the ith molecular orbital, φ_s is the sth basis function and c_{si} are weighting coefficients that must be adjusted to get the best molecular orbital.

This approach to generate molecular orbitals is called Linear Combination of Atomic Orbitals (LCAO) in spite that these basis functions are not necessarily atomic orbitals but any set of mathematical functions whose linear combination yields useful representations of molecular orbitals. The linear combination of basis functions, which generate molecular orbitals, is called basis set.

Slater functions (equation 2.5.) are good approximations to atomic wave functions, but large computer time is required to evaluate them. Using Gaussian functions (equation 2.6.) reduce considerably the CPU time.

$$\varphi = a \cdot \exp(-br) \qquad (2.5.)$$

$$\varphi = a \cdot \exp(-br^2) \qquad (2.6.)$$

The total wave function ψ is expressed as a Slater determinant of spin orbitals α and β which for a system containing 2n electrons is:

$$\Psi = \frac{1}{\sqrt{(2n)!}} \begin{vmatrix} \psi_1(1)\alpha(1) & \psi_1(1)\beta(1) & \psi_2(1)\alpha(1) & \psi_2(1)\beta(1) & \cdots & \psi_n(1)\beta(1) \\ \psi_1(2)\alpha(2) & \psi_1(2)\beta(2) & \psi_2(2)\alpha(2) & \psi_2(2)\beta(2) & \cdots & \psi_n(2)\beta(2) \\ \vdots & \vdots & \vdots & \vdots & & \\ \psi_1(2n)\alpha(2n) & \psi_1(2n)\beta(2n) & \psi_2(2n)\alpha(2n) & \psi_2(2n)\beta(2n) & & \psi_n(2n)\beta(2n) \end{vmatrix} \quad (2.7.)$$

Next a brief survey of different kinds of basis sets is given.

STO-nG basis sets

This kind of basis sets are characterized by the use of *n* Gaussian primitive functions for each basis functions in which coefficients *a* and *b* (eq. 2.6.) are fitted so that they approximate the corresponding Slater function.[3] In this sense one can distinguish between STO-3G, STO-4G or STO-6G basis sets in which three, four and six Gaussian functions have been used, respectively.

Split-valence basis sets

In split-valence basis sets while core electrons are described using one single basis function, valence electrons are described by two or more basis functions. In the present thesis two kinds of split-valence basis sets have been used: Pople basis sets and Dunning basis sets. Split-valence double-zeta Pople basis sets nomenclature follows the *K-MNG* scheme. *K* is the number of Gaussian functions used to represent the basis functions of the core electrons. *M* and *N* indicate that each valence orbital is split into two parts being each part represented by M and N Gaussians respectively. Examples of these basis sets are 3-21G[4] or 6-31G.[5] Split-valence triple-zeta Pople basis set such as 6-311G, in which each valence orbital is split into three parts, have been used in the present thesis. The addition of polarization and/or diffuse functions to the basis set leads to a more realistic electron distribution. Polarization functions are added by supplementing basis functions with d, p or f orbitals so that electron distribution is polarized, *i.e.* displaced along a particular direction. Polarization functions in Pople basis sets are denoted in parenthesis such as, for instance, 6-31G(d,p) basis set. Diffuse functions are added to reproduce in a better way the behavior of electrons that are at a larger distance from the nuclei. Thus, diffuse functions, which are Gaussian functions

with small values of coefficient *b* of equation 2.6., are recommended to be used to study molecules with heteroatoms, anions and electronically excited molecules. Diffuse functions are denoted with a plus sign (+) when they added only to heavy atoms and with two plus signs (++) when they are also added to helium and hydrogen atoms.

Dunning basis sets, also called correlation consistent basis sets, are constituted by adding shells of functions to a core set of atomic Hartree-Fock functions. Each function in a shell contributes approximately the same amount of correlation energy. Dunning basis sets follow the nomenclature *cc-pVXZ*, being X the degree of polarization. The addition of diffuse functions to Dunning basis sets is denoted with the prefix "aug".[6]

2.2.3. *Closed-shell and open-shell systems*

Closed-shell systems are those where all orbitals are doubly occupied. In other words, all electrons are split in pairs and, in each pair, electrons have opposite spins and occupy the same spatial orbitals. These systems can be handled by using the so called restricted methods, which use a single molecular orbital twice (one multiplied by the α spin function and the other multiplied by the β spin function) in the Slater determinant (eq. 2.7.).

Conversely open-shell systems are those where some of the electrons are unpaired, leading to singly occupied molecular orbitals. These systems can be treated by using unrestricted methods, which use different orbitals for different spins.

2.2.4. *Ab initio calculations*

Ab initio calculations are grounded on solving the Schrödinger equation (eq. 2.1.) without using any empirical adjustment, *i.e.* first principles. Thus *ab initio* calculations are based only on basic physical theory. The simplest *ab initio* method is the Hartree-Fock (HF) method or Self Consistent Field (SCF) procedure.[7] The HF wave function is approximated to a single Slater determinant of spin orbitals (eq. 2.7.). Schrödinger equation is solved through an iterative process applying the variational principle, *i.e.* the energy of any approximate wave function is always larger than the exact energy. Since one electron wave functions are used, electron-electron interactions are neglected and each electron interacts only with the mean field arising from all the remaining electrons, which is described by the Fock operator. The best approximate wave function is

obtained by varying iteratively the weighting coefficients of the atomic orbitals (c_{si} in eq. 2.4.) until the energy expectation value, $E(\psi)$, of the approximate wave function is minimized.

$$E(\psi) = \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} \tag{2.8.}$$

The major weakness of HF method is the fact that it does not describes electron correlation properly since each electron is considered to move in an electrostatic field whereas in reality electrons repeal each other. This deficiency makes HF method not reliable to get accurate estimates of systems where electron correlation plays an important role such as π-stacking interactions or chemical reactions. Post-HF methods correct the aforementioned deficiency. In this thesis two post-HF methods have been used: Møller Plesset (MP) procedure[8] and Coupled Cluster (CC) method.[9] Møller Plesset is based on perturbation theory. The philosophy behind MP methods is that a correction term handles electron correlation by promoting electrons from occupied to virtual molecular orbitals giving electrons more room to move and thus making it easier for them to avoid one another. The Hamiltonian operator is defined as the addition of a perturbation operator ($\hat{V}$) to the unperturbed HF Hamiltonian ($\hat{H}^0$):

$$\hat{H} = \hat{H}^0 + \lambda \cdot \hat{V} \tag{2.9.}$$

where λ is a dimensionless parameter taking values compressed between 0 and 1. Thus both wavefunction and energy are described by this perturbed Hamiltonian operator.

$$\psi = \psi^{(0)} + \lambda \cdot \psi^{(1)} + \lambda^2 \cdot \psi^{(2)} + \ldots + \lambda^n \cdot \psi^{(n)} \tag{2.10.}$$

$$E = E^{(0)} + \lambda \cdot E^{(1)} + \lambda^2 \cdot E^{(2)} + \ldots + \lambda^n \cdot E^{(n)} \tag{2.11.}$$

where $\psi^{(n)}$ is the n^{th} correction of the wave function related to the number of virtual spin orbitals and $E^{(n)}$ is the n^{th} correction energy term. Since HF energy is the sum of terms $E^{(0)}$ and $E^{(1)}$ electronic correlation corrections are taken into account from the second term (MP2). In the present thesis MPn (with n=2, 3 and 4) have been employed.

The Coupled Cluster method expresses the correlated wave function as a sum of the HF ground state determinant plus determinants representing the promotion of electrons to virtual molecular orbitals:

$$\psi = \left(1 + \hat{T} + \frac{\hat{T}^2}{2!} + \frac{\hat{T}^3}{3!} + \cdots + \frac{\hat{T}^n}{n!}\right) \cdot \psi_{HF} = e^{\hat{T}} \cdot \psi_{HF} \qquad (2.12.)$$

where $\hat{T} = \hat{T}_1 + \hat{T}_2 + \ldots + \hat{T}_n$ and the operators $\hat{T}_n$ are excitation operators and have the effect of promoting n electrons into virtual spin orbitals. In the present thesis CCSD(T) has been employed, that means coupled cluster singles, doubles and triples method being the contribution of triple excitations represented in an approximate way.

2.2.5. *Semiempirical calculations*

Semiempirical (SE) methods are based on the Schrödinger equation (eq. 2.1.) and experimental data. In spite of being computationally cheaper than *ab initio* methods a very accurate parameterization is needed to get reliable results. The main differences between SE and *ab initio* methods lie basically in four points:

- SE methods treat only valence or π electrons
- The basis functions correspond to atomic orbitals while in *ab initio* calculations this is only true when minimal basis set are used, *i.e.* the smallest number of functions that can accommodate all the electrons.
- Both core and the two electron repulsion integrals are not calculated from first principle but using empirical data.
- Simplifications in the mathematical model by assuming some matrices as unit matrices.

In the present thesis PM3 (Parameterized Model number 3) has been used.[10] PM3 method follows the same formalism as the previous SE method AM1[11] and differences between them arise in the parameterization employed as well as in the way to describe functions for the core repulsion.

2.2.6. *Density Functional Theory calculations*

The Density Functional Theory (DFT) is based on the theorems of Hohenberg-Kohn[12] which state that the ground state energy of a non degenerate electronic system as well as its electronic properties are solely defined by its electron density. In consequence DFT does not use the wave function but an electron probability density function *ρ(x,y,z)*, which refers to the probability of finding an electron in a volume element *dxdydz* centered on a point with coordinates (x, y, z). Assuming *r* as the position vector of the point with coordinates (x, y, z) and taking into account Born and Pauli interpretation, which states that the square of a one electron wave function (Kohn-Sham orbitals, ψ_i in eq. 2.13.) at any point is the probability density at that point, electron density can be written as follows:

$$\rho(r) = \sum_{i=1}^{n} \left|\psi_i(r)\right|^2 \qquad (2.13.)$$

where *n* is the number of occupied molecular orbitals. Electronic energy *E(ρ)* is calculated as a simple sum of different contributions that depend on the electronic density:

$$E(\rho) = E^T(\rho) + E^V(\rho) + E^J(\rho) + E^{XC}(\rho) \qquad (2.14.)$$

being $E^T(\rho)$ the kinetic energy, $E^V(\rho)$ the term containing the potential attractive energy electron-nucleus and the repulsive term between nuclei, $E^J(\rho)$ the Coulombic repulsion between electrons and $E^{XC}(\rho)$ the interchange-correlation energy. The latest term is the only one that is not determined directly because of its unknown mathematical formulation. Usually $E^{XC}(\rho)$ is described as a sum of an exchange term $E^X(\rho)$ and another of electronic correlation $E^C(\rho)$.

$$E^{XC}(\rho) = E^X(\rho) + E^C(\rho) \qquad (2.15.)$$

The exchange term can be calculated by using approximations that applies a homogeneous electron density, such as the Local Density Approximation (LDA)[13] and the Local Spin Density Approximation (LSDA),[14] or by using gradient corrected functionals such as the so called Generalised Gradient Approximation (GGA) methods. Examples of functionals using a homogeneous electron density are the VWN[15] or the local correlation functional of Perdew (PL).[16] Several GGA based functionals are

reported such as Becke95 (B95),[17] Perdew 86 (P),[18] Perdew-Burke-Ernzerhof (PBE),[19] Perdew-Wang 91 (PW91)[20] and the extremely popular Lee-Yang-Parr (LYP).[21] The exchange-correlation term can also be calculated using the so called hybrid density functionals, which combine a conventional GGA method with a percentage of Hartree-Fock exchange. Examples of hybrid density functionals include B3LYP,[21,22] B3PW91,[20,22] MPW1K,[23] O3LYP[21,24] and X3LYP.[20,21,25]

2.2.7. *Solvent effects*

Solvent effects in QM calculations have been evaluated by means of continuum models, *i.e.* while the solvent is described as an infinite dielectric medium the solute is treated at the QM level. In the present thesis the so called Polarizable Continuum Model (PCM) has been employed.[26] PCM is a Self Consistent Reaction Field (SCRF) method, *i.e.* implies the use of a reaction potential to be self-consistently solved together with the solute charge.

PCM method involves the generation of a solvent cavity from spheres centered at each atom of the molecule and the calculation of virtual point-charges on the cavity surface representing the polarization of the solvent. The magnitude of these charges is proportional to the derivative of the solute electrostatic potential at each point calculated from the molecular wave function.

Within PCM method the transfer of a given solute from the gas phase into solution can be portioned into three steps:

- Creation of the solute cavity inside bulk solvent.
- Generation of the van der Waals particle inside the cavity.
- Generation of the solute charge distribution in solution.

Thus the net molecular free energy of a system in solution (ΔG_{solv}) is the sum of these three contributions:

$$\Delta G_{solv} = \Delta G_{cav} + \Delta G_{vdW} + \Delta G_{elec} \qquad (2.16.)$$

where ΔG_{cav} is the work involved to increase the cavity, ΔG_{vdW} is the contribution due to the van der Waals interactions and ΔG_{elec} is the electrostatic component of the ΔG_{solv} and it corresponds to the work required in the polarization process. Interaction potential between solvent and solute is introduced in the solute electronic Hamiltonian as a

perturbation operator ($\hat{V}_R$). Thus the Schrödinger equation, being $\hat{H}^0$ the solute's Hamiltonian operator, is expressed as follows:

$$\left(\hat{H}^0 + \hat{V}_R\right)\psi = E\,\psi \qquad (2.17.)$$

.

2.3. Molecular dynamics simulations

Molecular dynamics (MD) simulations reproduce the dynamical evolution of a system at a molecular level by means of MM. MM methods are based on mathematical models that assume a molecule as a collection of balls and springs representing respectively atoms and bonds. Each particle is described by a radius, hardness and net charge.

2.3.1. *Force-Field*

Force-field (FF) is the form of the mathematical expression and the parameters needed to calculate the potential energy of a system under the MM formalism. FF includes terms referring to both bonding and non-bonding interactions. In the present thesis AMBER FF has been employed.[27]

$$E(r) = \sum_{bonds} k_{stretcht}(r - r_{eq})^2 + \sum_{angles} k_{bend}(\theta - \theta_{eq})^2 + \sum_{dihedrals} k_{torsion}[1 + \cos(n \cdot \phi - \gamma)] + \\ + \sum_{i<j}\left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right] + \sum_{i<j}\frac{q_i q_j}{4\pi\varepsilon_o r_{ij}} \qquad (2.18.)$$

As can be seen in equation 2.18. total energy has five terms which are, respectively:

- The bond stretching term in which the stretching force constant ($k_{stretch}$) and the distance between two bonded atoms and its equilibrium bond distance (r and r_{eq} respectively) are included.
- The angle bending term including the bending force constant (k_{bend}) and the angle between three consecutive atoms and its equilibrium value (θ and θ_{eq}).
- The torsional term contains a dihedral constant ($k_{torsion}$) setting the energy barrier for the rotation profile, the actual dihedral angle (ϕ) and the equilibrium one (γ).
- The van der Waals non bonded term, a mathematical model following the Lenard-Jones potential that takes into account the attractive and repulsive forces between two atoms being the r_{ij} the distance between them. A_{ij} and B_{ij} are parameters that depend of the pair of atoms referring to hard core repulsion and dispersive attraction respectively.
- The electrostatic interactions are taken into account in the last term, where q_i and q_j are the point charges of atoms i and j respectively, r_{ij} is the distance between them and ε_0 is vacuum permittivity.

All the constants in 2.18. as well as all the equilibrium values are taken from AMBER libraries. All the terms not found in such libraries must be parameterized either using experimental data or by means of *ab initio* or DFT calculations. In the present thesis charges have been obtained from QM *ab initio* calculations according to either the so-called Restrained Electrostatic Potential (RESP)[28] or by fitting the quantum mechanical and the Coulombic Molecular Electrostatic Potential (MEP).[29]

2.3.2. *Mathematical model of molecular dynamics*

Given an atom i with mass m_i and considering that its position is described by a three dimensional vector r_i, its motion is ruled by Newton's law

$$\frac{dv_i(t)}{dt} = \frac{F_i}{m_i} \tag{2.19.}$$

$$v_i(t) = \frac{dr_i(t)}{dt} \tag{2.20.}$$

where v_i and F_i are respectively the velocity and the force acting on the atom i in a given moment. F_i can be obtained through equation 2.21.

$$F_i = -\frac{\partial E(r_N)}{\partial r_i} \tag{2.21.}$$

where $E(r_N)$ is given by eq. 2.18. and N is the number of particles in the system.

Equations 2.19. and 2.20. form a system of N coupled differential equations that might be solved numerically. This system can be integrated step by step by the so called leap-frog Verlet algorithm,[30] which is founded on a Taylor expansion of the position r_i at time $t_n=t_0+n\Delta t$, leading to equations 2.22. and 2.23.

$$v_i\left(t_n + \frac{\Delta t}{2}\right) = v_i\left(t_n - \frac{\Delta t}{2}\right) - \frac{\Delta t}{m_i}\frac{\partial E(r_N)}{\partial r_i} + o\left(\Delta t^3\right) \tag{2.22.}$$

$$r_i(t_n + \Delta t) = r_i(t_n) + v_i\left(t_n + \frac{\Delta t}{2}\right)\Delta t + o\left(\Delta t^4\right) \tag{2.23.}$$

Time step Δt might be small enough to simulate those movements with the highest frequencies, which normally are the bond vibrations. Consequently time step takes

values of the order of the femtosecond. Small time steps lead to more expensive simulations.

The CPU time needed for a MD simulation depends on several factors as the number of the explicit particles in the system, the time step or the cut-off, *i.e.* the maximum distance in which non bonded interactions are evaluated. In order to speed up the MD simulation some simplifications might be done such as freezing the fastest modes of vibration by constraining the bonds to hydrogen atoms to fixed lengths. In this sense, SHAKE or constrained Verlet method[31] has been used in this thesis.

2.3.3. *Thermodynamic ensembles*

An ensemble is a collection of all possible systems that have differing microscopic states but belong to a single macroscopic or thermodynamic state. There are four ensembles that are used in MD simulations.[32]

- The canonical or NVT ensemble, whose thermodynamic state is characterized by a fixed number of atoms *N*, fixed volume *V* and fixed temperature *T*.
- The isobaric-isoenthalpic or NPH ensemble, where the number of atoms *N*, the pressure *P* and enthalpy *H* are fixed.
- The isobaric-isothermal or NPT ensemble, with fixed values of number of atoms *N*, pressure *P* and temperature *T*.
- The microcanonical or NVE ensemble, which corresponds to a closed or isolated system since energy *E*, besides the number of atoms *N* and volume *V*, is fixed.

In the present thesis NVT and NPT ensembles have been employed. Temperature is mainly set by the average kinetic energy of the molecules of the system. Consequently temperature can be controlled by adjusting the velocities of each atom. Berendsen thermostat rescales velocities of atoms assuming a scale factor λ defined as:

$$\lambda = \left[1 + \frac{\Delta t}{\tau_T} \left(\frac{T}{T_0} - 1 \right) \right]^{\frac{1}{2}} \qquad (2.24.)$$

where τ_T is the heat bath coupling time and T_0 is the set point temperature. Concerning the pressure, a fixed value of it implies that volume must be able to fluctuate by dynamically adjusting the size of the unit cell and rescaling all atomic coordinates. In the case of Berendsend barostat, a scale factor μ is used to scale lengths in the system.

$$\mu = \left[1 - \frac{\Delta t}{\tau_P} \left(P - P_0 \right) \right]^{\frac{1}{3}} \qquad (2.25.)$$

where τ_P is the pressure relaxation time of the barostat and P_0 is the set point pressure. Equations 2.24. and 2.25. define the so called Berendsen thermobarostat.[33]

2.3.4. *Periodic boundary conditions and Particle Mesh-Ewald Summation*

MD simulations are usually performed in simulation boxes containing the solute and solvent atoms. In order to represent an infinite system the whole system should be replicated periodically in all directions so atoms outside the simulation box are simply images of the atoms simulated in that box. Periodic boundary conditions ensure that all simulated atoms are surrounded by neighboring atoms, either images or not. This condition guarantees that atoms moving out of the box at one side are able to enter again inside the box at the opposite side because replicas of each particle in all duplicated boxes move exactly the same way. Minimum image convention avoid duplicate interactions between atoms *i* and *j* by taking into account only the interaction of atom *i* with the closest *j* atom, either original or copy. Thus, periodic boundary conditions are useful to correct potential errors in the van der Waals non bonded term. Concerning the Coulombic non bonded term, the main problem to evaluate it correctly is that a sudden cut-off lead to large errors. This entanglement can be solved with the use of the so called Particle Mesh-Ewald summation (PME),[34] which calculate the infinite electrostatic interactions by splitting the summation into short and long range parts. For PME, the cut-off only determines the balance between the two parts, and the long-range part is treated by assigning charges to a grid that is solved in reciprocal space through Fourier transforms.

2.4. QM/MM calculations

QM/MM methods are based in splitting the system in two parts, a small one of special interest described at the QM level and the rest of the system described at the MM level. QM/MM philosophy was firstly introduced by Warshel and Levitt[35] and developed by Field, Bash and Karplus.[36] QM/MM methods differ, besides the classical force field and the QM level of theory employed, in how the two parts are connected. There are two way to connect QM and MM parts:

- The so called link atom scheme, which uses a monovalent atom to cap the unsaturated QM atoms. Normally hydrogen atom is used as link atom.[37]
- Methods based in the use of localized orbitals in the boundaries separating the QM and MM parts. One example is the Local Self-Consistent Field (LSCF) method[38] in which the bonds connecting both parts are represented by a set of strictly localized bond orbitals (SLBOs) that are determined by calculations on small model compounds and assumed to be transferable.

Within the framework of QM/MM calculations the total Hamiltonian operator is defined as the sum of the QM, QM/MM and MM Hamiltonians:

$$\hat{H} = \hat{H}^{QM} + \hat{H}^{QM/MM} + \hat{H}^{MM} \tag{2.26.}$$

where $\hat{H}^{QM/MM}$ is expressed as the sum of electron-charge, nuclei-charge and van der Waals interaction potentials between QM and MM atoms:

$$\hat{H}^{QM/MM} = V_{elec}^{QM/MM} + V_{nucl}^{QM/MM} + V_{vdW}^{QM/MM} \tag{2.27.}$$

Equation 2.27. can be expanded for a system containing *N* QM atoms and *M* MM atoms as follows:

$$\hat{H}^{QM/MM} = -\sum_{i=1}^{N}\sum_{j=1}^{M}\frac{q_j}{r_{ij}} + \sum_{i=1}^{N}\sum_{j=1}^{M}\frac{Z_i q_j}{r_{ij}} + \sum_{i=1}^{N}\sum_{j=1}^{M}\left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right] \tag{2.28.}$$

Thus energy in a QM/MM system is defined as shown in Equation 2.29.

$$E = E^{QM} + E^{MM} + E^{QM/MM} \tag{2.29.}$$

where $E^{QM/MM}$ is defined as:

$$E^{QM/MM} = \langle \psi | - \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{q_j}{r_{ij}} | \psi \rangle + V_{nucl}^{QM/MM} + V_{vdW}^{QM/MM} \quad (2.30.)$$

where the first term is the electronic interaction within the Electronic Emboding scheme.

.

2.5. References

[1] Lewars E. “Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics”, **2004**, Kluwer Academic Publishers, Dordrecht.

[2] a) Roothaan C.C.J. *Rev. Mod. Phys.* **1951**, *23*, 69. b) Hall G.G., *Proc. Roy. Soc.* **1951**, *A205*, 541.

[3] Hehre W.J., Stewart R.F., Pople J.A. *J. Chem. Phys.* **1969**, *51*, 2657.

[4] Binkley J.S., Pople J.A., Hehre W.J. *J. Am. Chem. Soc.* **1980**, *102*, 939.

[5] Ditchfield R., Hehre W.J., Pople J.A. *J. Chem. Phys.* **1971**, *54*, 724.

[6] a) Dunning T.H. *J. Chem. Phys.* **1989**, *90*, 1007 b) Kendall R.A., Dunning T.H., Harrison R.J. *J. Chem. Phys.* **1992**, *96*, 6796 c) Woon D.E., Dunning T.H. *J. Chem. Phys.* **1993**, *98*, 1358 d) Peterson K.A., Woon D.E., Dunning T.H. *J. Chem. Phys.* **1994**, *100*, 7410 e) Wilson A., van Mourik T., Dunning T.H. *J. Mol. Struct. (Theochem)* **1997**, *388*, 339.

[7] a) Hartree D.R. *Proc. Cambridge Phil. Soc.* **1928**, *24*, 89 b) Slater J.C. *Phys. Rev.* **1930**, *35*, 210 c) Fock V. *Z. Physik* **1930**, *61*, 126.

[8] a) Møller C., Plesset M.S. *Phys. Rev.* **1934**, *46*, 618 b) Binkley J.S., Pople J.A. *Int. J. Quantum Chem.*, **1975**, *9*, 229.

[9] a) Coester F., Kümmel H. *Nucl. Phys.* **1960**, *17*, 477 b) Paldus J., Čížek J. Shavitt I. *Phys. Rev. A* **1972**, *5*, 50.

[10] a) Stewart J.J.P. *J. Comp. Chem.* **1989**, *10*, 209 b) Stewart J.J.P. *J. Comp. Chem.* **1989**, *10*, 221 c) Stewart J.J.P. *J. Comp. Chem.* **1991**, *12*, 320.

[11]Dewar M.J.S., Zoebisch E.G., Healy E.F., Stewart J.J.P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

[12] Hohenberg P., Kohn W. *Phys. Rev. B* **1964**, *136*, B864.

[13] Slater J.C. *Int. J. Quantum Chem. Symp.* **1975**, *9*, 7.

[14] Painter G.S. *Phys. Rev. B* **1981**, *24*, 4264.

[15] Vosko S.H., Wilk L., Nusair M. *Can. J. Phys.* **1980**, *58*, 1200.

[16] Perdew J.P., Zunger A. *Phys. Rev. B* **1981**, *23*, 5048.

[17] Becke A.D. *J. Chem. Phys.* **1996**, *104*, 1040.

[18] Perdew J.P. *Phys. Rev. B* **1986**, *33*, 8822.

[19] Perdew, J.P., Burke, K., Ernzerhof M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

[20] Perdew J. P., Wang Y. *Phys. Rev. B* **1992**, *45*, 13244.

[21] Lee C., Yang W., Parr R.G. *Phys. Rev. B* **1988**, *37*, 785.

[22] Becke A.D. *J. Chem. Phys.* **1993**, *98*, 1372.

[23] a) Adamo C., Barone V. *J. Chem. Phys.* **1998**, *108*, 664 b) Lynch B.J., Fast P.L., Harris M., Truhlar D.G. *J. Phys. Chem. A* **2000**, *104*, 4811 c) Lynch B.J., Zhao Y., Truhlar D.G. *J. Phys. Chem. A* **2003**, *107*, 1384.

[24] Schultz N.E., Zhao Y., Truhlar D.G. *J. Phys. Chem. A* **2005**, *109*, 11127.

[25] a) Becke A.D. *Phys. Rev. A* **1988**, *38*, 3098 b) Xu X., Goddard W.A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 2673.

[26] a) Miertuš S., Scrocco E., Tomasi J. *Chem. Phys.* **1981**, *55*, 117 b) Miertuš S., Scrocco E., Tomasi J. *Chem. Phys.* **1982**, *65*, 239.

[27] a) Cornell W.D., Cieplak P., Bayly C.I., Gould I.R., Merz K.M., Ferguson D.M., Spellmeyer D.C., Fox T., Caldwell J.W., Kollman P.A. *J. Am. Chem. Soc.* **1995**, *117*, 5179 b) Wang J., Cieplak P., Kollman P.A. *J. Comput. Chem.* **2000**, *21*, 1049.

[28] a) Bayly C.I., Cieplak P., Cornell W., Kollman P.A. *J. Phys. Chem.***1993**, *97*, 10269 b) Cornell W.D., Cieplak P., Bayly C.I., Kollmann P.A. *J. Am. Chem. Soc.* **1993**, *115*, 9620.

[29] Alemán C., Luque F.J., Orozco M. *J. Comput. Aided Mol. Design* **1993**, *7*, 721.

[30] Verlet J. *Phys. Rev.* **1967**, *159*, 98.

[31] a) van Gunsteren W.F., Berendsen H.J.C. *Mol. Phys.* **1977**, *34*, 1311 b) Ryckaert J. P., Ciccotti G., Berendsen H.J.C. *J. Comput. Phys.***1977**, *23*, 327.

[32] Adcock S.A., McCammon J.A. *Chem. Rev.* **2006**, *106*, 1589.

[33] Berendsen H.J.C., Postma J.P.M., van Gunsteren W.F., DiNola A., Haak J.R. *J. Chem. Phys.* **1984**, *81*, 3684.

[34] Darden T., York D., Pedersen L. *J. Chem. Phys.* **1993**, *98*, 10089.

[35] Warshel A., Levitt M. *J. Mol. Biol.* **1976**, *103*, 227.

[36] Field M., Bash P., Karplus M. *J. Comp. Chem.* **1990**, *11*, 700.

[37] Maseras F., Morokuma K. *J. Comput. Chem.* **1995**, *16*, 1170.

[38] Ferré N., Assfeld X., Rivail J.L. *J. Comput. Chem.* **2002**, *23*, 610.

3. NANOTUBES BASED ON NATURAL PEPTIDE SEQUENCES

3.1. Testing β–helix terminal coils stability by targeted substitutions with non-proteogenic amino acids: A Molecular Dynamics study*

3.1.1. *Introduction*

An important aim of nanotechnology is to design new microscopic devices at the molecular level, necessitating the tuning of the structural and physicochemical features of its molecular parts. An increasingly common strategy is to combine polymer science with advanced organic synthesis procedures.[1] Nanoconstructs are usually built by assembling of macromolecular components, modules or building blocks, either via covalent linkages or self assembly processes.[2] However, detailed knowledge of the principles that govern the organization of the nanostructures is still lacking: currently, successful strategies are largely based on the assumption that isolated modules inherently present an acceptable degree of ordered organization, sufficient to guarantee and promote the correct formation of the nanoconstruct.[3] This restricts the type of building blocks that can be selected for nanodesign, since the inherent capacity for self-organization needs to be a preferred feature. Building blocks of natural origin appear as a logical choice in this respect, since subcellular biological structures exist and operate at the nano scale: the molecular machinery that makes life possible is inherently based on self assembled structures that perform specific and extraordinarily complex functions.[4] Creation of nano-sized structures follows the same principles that control the folding and the stability of interacting biomolecules.[5] Whether in natural structures or in nanoconstructs, the final product is the last stage in a hierarchical chain of chemical events that initiate with the selection of those building blocks with molecular conformations that favor a particular assembly.[6] In such a scheme, the stabilization of conformational motifs of natural building blocks extracted from their natural environment, is an important task.

In recent years the use of segments extracted from proteins as molecular scaffolds to build tubular nanoconstructs has been explored.[7] For this purpose, several efforts have been centered on the β–helix fold: a conformational arrangement that contains a repetitive helical strand-loop motif, where each repeat contributes a strand segment to one or more parallel β sheet(s). Systems containing left-handed β-helical building blocks are of special interest since this arrangement is highly repetitive and

* Results presented in this chapter are published in *J. Struct. Biol.* **2007**, *160*, 177

symmetrical, which should allow the formation of nanofibers without performing many structural manipulations.[8] As compared with right-handed β-helices, left-handed β-helices exhibit limited variability of shape, size and sequence,[9] displaying the characteristic equilateral triangular shape. It was demonstrated that the nanofibril constructed by four stacked replicas of a two-turn repeat building block (i.e. a β-helix terminal coil) from the *Escherichia coli* galactoside acetyltransferase (PDB code 1krr, chain A) exhibits remarkable stability under different simulation conditions, including temperature stress and various ionic strengths.[7] Results indicated that the formation of a nanofibril by self-assembling is granted if the native conformation of the building block is preserved once it is excised from the whole protein. Yet, this necessary condition is difficult to accomplish, since even in the best scenario, the desired conformation would only be part of a set of conformers that define the structure of the building block in solution.

Previous contributions toward the engineering of nanotubular architectures was to shift the population balance toward the desired conformation by selective stabilization of key positions along the peptide chain.[7,10] The conformational preferences of the building blocks, extracted from whole β-helix proteins were first probed *in silico*; next, positions observed to be most sensitive to conformational disorder were stabilized. This was accomplished by replacing the chosen positions by synthetic amino acids with restricted conformational space.[10] To date, the most flexible regions of the β-helix capping coil, the turns, have been the replacement targets since such substitutions would contribute substantially to destabilize the whole conformation. Overall, the results of these investigations have confirmed this assumption.[7,10] However, in some cases the connection between the β-sheet segments becomes too rigid if the flexibility of the loop is reduced.[10b] An alternative approach would be to fix the β-sheet segments by targeted replacements; however, without altering the flexibility of the loop.

In order to choose the proper target sites for the stabilization of the β-sheet segments in the wild type sequence, some previous considerations were taken into account:

(i) The residue/s to be replaced must not hold any essential role in the overall motif organization. Specifically, in the case of the β-helix conformation, the side chain should not be involved in the hydrophobic core of the helix center as frequently happens with the residues located in β-sheets;

(ii) The targeted position should preferably be in the center of the β-sheet to avoid undesirable conformational influence in the adjacent loops;

(iii) The residue type should preferably be hydrophobic, since polar groups may induce competitive interactions with other side chains destabilizing the whole motif more easily than hydrophobic side chains, due to the lager range of action of polar interactions;

(iv) The size of the targeted side chain should be voluminous enough to allow the replacement by a conformationally restrained amino acid, which is quite large. Thus, conformational restrictions are usually introduced by modifying the main chain or the side chain of conventional α-amino acids.

In this chapter the investigated system is the aforementioned protein segment excised from the chain A of *E. coli* galactoside acetyltransferase, which comprises residues 131-165, hereafter referred to solely as 1krr. If the considerations exposed above are applied to 1krr, the most suitable positions to be target for replacement are the valines in positions 145 and 157 (Figure 1). These residues, located in the middle of their respective β-sheets, do not project inward, toward the helix inner cavity; hence, reducing the hydrophobicity of the site may be expected to increase the sheet solubility and, therefore, the conformational stability.

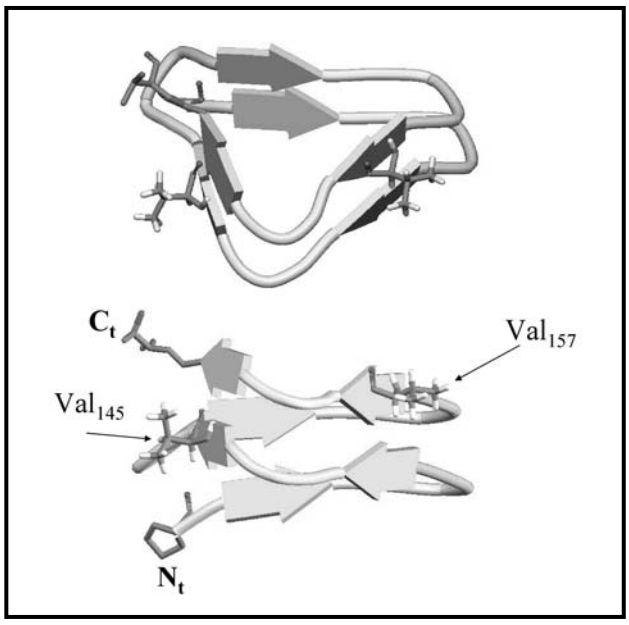
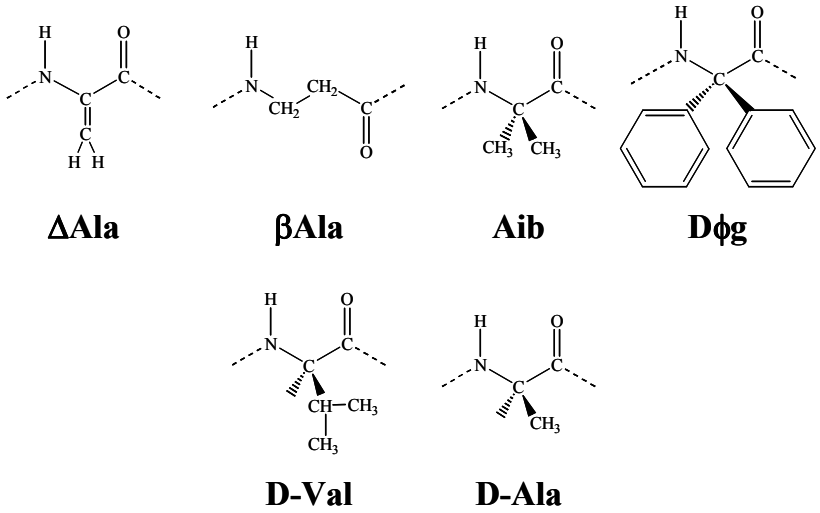


Figure 1. Equatorial (top) and axial (bottom) partial projections of the β-helical building block extracted for the crystal structure of $1krr_{131-165}$ segment. The motif is represented by arrows with the exception of the sites targeted for residue replacement, depicted by all atoms of both Val residue using a stick representation. For clarity, N_t and C_t termini are also marked.

Six different synthetic amino acids have been tested as possible stabilizing agents (Scheme 1). These series were chosen because they are relatively easy to synthesize, their conformational properties have already been characterized (in most cases both theoretically and experimentally) and in most cases have already been used in conformational studies of polypeptides (see below for details).



Scheme 1

On the other hand, although the conformational properties of three of them are more restricted than those of proteogenic residues, the characteristic shared by all the amino acids probed is the ability to favor extended or pseudoextended conformations under proper conditions. At the same time, it was attempted to cover a broad range of chemical modifications that can be introduced into proteonogenic amino acids, which justifies the variability of the tested amino acids.

The results presented in this work clearly demonstrate that the stability of β-helix terminal coils is extremely site sensitive, which implies that the conformational properties of the amino acids are not the sole determinants in enhancing the stability of peptide segments folded in β-sheets. This is illustrated by the fact that only when the conformational alteration induced by the synthetic amino acids is combined with the formation of new mid- and long-range interactions, the overall stability of the β-helix capping coil can be enhanced. In spite of the steric hindrance imposed by the voluminous phenyl side chains, diphenylglycine (Dϕg) has been found to be the best option for the enhancement of the building block stability. This is because the side

chains form new non-covalent interactions with neighboring residues, which provides rigidity to the β–sheet segments in the fold.

3.1.2. *Survey of Synthetic Amino Acids*

Four synthetic amino acid types, which differ significantly in their chemical constitution, have been considered for targeted replacements. This section describes the more important characteristics of each amino acid type.

Dehydroalanine (ΔAla)

This is an α,β-unsaturated amino acid commonly found in a number of naturally occurring peptide antibiotics of bacterial origin.[11] The conversion of the tetrahedral saturated to a trigonal dehydro residue with a shortening of the C^{α}-C^{β} distance introduces strong and specific steric effects, which influence its conformational behavior. Thus, the double bond of the side chain enhances the stability of the extended and (pseudo)extended conformations in both single residues.[12]

α,α-Dialkylated amino acids: α-aminoisobutyric acid (Aib) and diphenylglycine (Dϕg)

Substitution of the α-carbon hydrogen by an alkyl group induces drastic changes in the conformational properties of amino acids, which are frequently associated with the incorporation of strong conformational constraints. The simplest α,α-amino acid is Aib, in which a methyl group replaces the α-carbon hydrogen of Ala, while Dϕg should be considered as a Gly derivative with two bulky phenyl groups replacing the corresponding α-carbon hydrogen atoms. It was recently found that Dϕg prefers an extended conformation,[13] although depending on the peptide sequence, it can also be part of helix arrangements, such as 3_{10} helix organization.[14] In contrast, Aib promotes folded conformations,[15] hence this residue is used in this work for comparison with Dϕg.

β-Alanine (βAla)

In general, β-amino acids, which present two carbon atoms between the amino and carbonyl moieties, have been shown to stabilize β-sheet secondary structures in short peptides.[16] Moreover, βAla, which is the simplest β-amino acid, forms poly(β-alanine)

(also denoted nylon 3) upon polycondensation that exists in two crystal forms, the molecular chains adopting a fully extended conformation in both structures.[17] A study of the intrinsic conformational properties of βAla revealed that at least two of the three flexible angles of this residue tend to adopt an extended or semi-extended conformation.[18]

D-Amino acids

Inversion of the natural chirality of the α-carbon atom position was shown to be an excellent strategy to stabilize β-sheet motif at the edges of β-turns.[19] Here, the consequences of breaking the inner symmetry of the peptide backbone have been evaluated by introducing D-Val or D-Ala instead of the targeted L–Val sites.

3.1.3. *Methods*

19 Independent simulations were performed corresponding to all the studied targeted substitutions plus the wild type sequence (*wt*). For each family of non proteogenic amino acids, two different targeted replacements were considered (positions 145 and 157). A summary of all the studied cases is given in Table 1. The simulated building block corresponds to the segment 131-165 of 1krr, with the simulation initiating from the main chain conformation of the crystal structure coordinates for each studied case. Each simulated peptide was conveniently blocked at the N_t and C_t termini by acetyl and N-methylamide groups, respectively. All simulated systems were placed in the center of a cubic simulation box (56×56×56 $Å^3$) filled with explicit water molecules, which were represented using the TIP3 model.[20] A positively charged sodium atom was added to the simulation box to reach electric neutrality (all considered systems had a negative net charge at neutral pH). Overlapping water molecules were removed and the systems subsequently used the lowest number of water molecules among all built systems (5656 water molecules).

Force-field specifics.

All the atoms of the system were explicitly considered and the energy was computed using the AMBER force-field.[21] The parameters for all the residues with exception of those that are non-proteogenic were taken from the AMBER libraries. Among the latter

residue types, only βAla atomic partial charges were parameterized *de novo*. For all the other studied synthetic residues, the necessary parameters were extracted from the literature. Recent studies indicated that the backbone conformation of ΔAla alters the hybridization of the α-carbon of this residue, which can transform from sp^2 to sp^3.[12a]

Table 1. Summary of all the simulated segments. The wild type sequence and the mutation sites are specified. Notice that the 3 letters code has been used for the non-proteogenic residues since they have never been named with a single letter code.

Series	Simulation	Sequence
Wild type	wt	PITIGNNVWIGSHVVINPGVTIGDNSVIGAGSIVT
Dehydroalanine	Val145ΔAla$_{sp2}$	PITIGNNVWIGSHVΔAlaINPGVTIGDNSVIGAGSIVT
	Val145ΔAla$_{sp3}$	PITIGNNVWIGSHVΔAlaINPGVTIGDNSVIGAGSIVT
	Val157ΔAla$_{sp2}$	PITIGNNVWIGSHVVINPGVTIGDNSΔAlaIGAGSIVT
	Val157ΔAla$_{sp3}$	PITIGNNVWIGSHVVINPGVTIGDNSΔAlaIGAGSIVT
D-Amino acids	Val145D-Val	PITIGNNVWIGSHVD-ValINPGVTIGDNSVIGAGSIVT
	Val145D-Ala	PITIGNNVWIGSHVD-AlaINPGVTIGDNSVIGAGSIVT
	Val145Ala[a]	PITIGNNVWIGSHVAINPGVTIGDNSVIGAGSIVT
	Val157D-Val	PITIGNNVWIGSHVVINPGVTIGDNSD-ValIGAGSIVT
	Val157D-Ala	PITIGNNVWIGSHVVINPGVTIGDNSD-AlaIGAGSIVT
	Val157Ala[a]	PITIGNNVWIGSHVVINPGVTIGDNSAIGAGSIVT
β-Alanine	Val145βAla	PITIGNNVWIGSHVβAlaINPGVTIGDNSVIGAGSIVT
	Val145Gly[b]	PITIGNNVWIGSHVGINPGVTIGDNSVIGAGSIVT
	Val157βAla	PITIGNNVWIGSHVVINPGVTIGDNSβAlaIGAGSIVT
	Val157Gly [b]	PITIGNNVWIGSHVVINPGVTIGDNSGIGAGSIVT
α,α-Dialkyl amino acids	Val145Aib	PITIGNNVWIGSHVAibINPGVTIGDNSVIGAGSIVT
	Val145Dϕg	PITIGNNVWIGSHVDϕgINPGVTIGDNSVIGAGSIVT
	Val157Aib	PITIGNNVWIGSHVVINPGVTIGDNSAibIGAGSIVT
	Val157Dϕg	PITIGNNVWIGSHVVINPGVTIGDNSDϕgIGAGSIVT

[a] Mutation designed to assess the real effect of the substitution of Val by D-Ala. If both D-Ala and Ala favor the sheet stabilization, the chirality inversion is unlikely to be the cause of stabilization. [b] Mutation designed to asses the effect of the substitution of Val by βAla. If βAla favors the sheet stabilization, the effect may be due to either the loss of side chain or the increase in the main chain length, being the former possibility if Gly also stabilizes the motif.

Accordingly, two sets of recently developed force-field parameters were used to describe the ΔAla residue (ΔAla_{sp2} and ΔAla_{sp3} in Table 1,[12c]). Aib bonding and non-bonding parameters were obtained from Preto and coworkers[22] with exception of the atomic partial charges, which were taken from the work of Improta et al.[23] Regarding the Dϕg residue, it has been recently reported a set of parameters compatible to the AMBER force-field.[13]

In order to properly compute the electrostatic energy associated with the interactions produced by βAla, a new set of atomic partial charges was parameterized. For this purpose, geometry optimizations at the *ab initio* Hartree-Fock quantum mechanical level combined with the 6-31G(d) basis set,[24] *i.e.* HF/6-31G(d) level, were performed on the most representative conformations of the βAla dipeptide.[18] For each conformation, the molecular electrostatic potential (MEP) was computed from the wave function of the optimized geometry. Finally, atomic electrostatic parameters were derived for each conformation by fitting the quantum mechanical MEP values into partial charges centered in the nuclei through a Levenberg-Marquardt nonlinear optimization procedure. The final set of charges was derived from the Boltzmann distribution provided by the energies of all of the calculated conformations. The parameters provided at the HF/6-31G (d) level of theory are fully consistent with those included in the AMBER libraries. All quantum mechanical calculations were performed using the Gaussian 03 computer program.[25]

Simulation details

All MD simulations were performed using NAMD program.[26] Atom pair distance cutoffs were applied at 14.0 Å to compute van der Waals interactions. The electrostatic interactions were evaluated using the non-truncated electrostatic potential by means of Ewald Summations. The real space term was determined by the van der Waals cut off (14.0 Å), while the reciprocal term was estimated by interpolation of the effective charge into a charges mesh, with maximum space between grid points 1 Å, i.e. the particle-mesh Ewald (PME) method.[27] Bond lengths were constrained using the SHAKE algorithm,[28] the numerical integration step being 2 fs.

Before the production series, the thermodynamic variables of the system were equilibrated. The energy of each system was initially minimized to relax conformational and structural tensions using the conjugate gradient method for $5 \cdot 10^3$ steps. Next,

different consecutive rounds of short MD runs were performed in order to equilibrate the density, temperature and pressure: first, the solvent and the counterion were thermally relaxed by three consecutives runs, whereas protein parts were kept frozen. Thus, 0.5 ns of NVT-MD at 500 K were used to homogeneously distribute the solvent components in the box. Next, 0.5 ns of isothermal and 0.5 ns isobaric relaxation were run. Finally, all the atoms of the system were submitted to 0.15 ns of steady heating until the target temperature was reached (298 K), 0.25 ns of NVT-MD at 298 K (thermal equilibration) followed by 0.5 ns of density relaxation (NPT-MD). Both temperature and pressure were controlled by the weak coupling method, the Berendsen thermo-barostat[29] using a time constant for heat bath coupling and a pressure relaxation time of 1 ps. The coordinates of the NPT-MD production runs, which were 10 ns long, were saved every 1000 steps (2 ps intervals) for subsequent analysis.

3.1.4. *Results and discussion*

Dehydroalanine (ΔAla): Variation in the α-carbon hybridization

A remarkable feature was observed in all cases: rigidifying the backbone by insertion of the double bond $C^{\alpha}=C^{\beta}$ induces an important destabilization in the overall three-dimensional organization. Nonetheless, the specific distortions depend on the geometric restrictions imposed by the equilibrium geometric parameters, being different for simulations based on ΔAla_{sp2} and ΔAla_{sp3} force-fields (see Methods). In the ΔAla_{sp3} force-field the equilibrium bond angle $\angle N\text{-}C^{\alpha}\text{-}C$ corresponds to a tetrahedral value (109.5º,[12a]). It was observed that, when the α-carbon atom adopts a sp^3 hybridization, the conformational preferences of ΔAla impose structural restrictions that rapidly induce a global loss of coherent fold, as was also detected in ΔAla-homopeptides.[12a] The aforementioned bond angle strongly favors the formation of extended conformations with φ,Ψ≈ 180º,180° (Figure 2a), which do not allow the proper organization of the β-sheet secondary structure due to topological constrains. Thus, dihedral angles associated with the extended conformation impose a trans arrangement between the hydrogen atom of the amide moiety and the oxygen atom of the next peptide bond. This topology is incompatible with the formation of parallel β-sheets, which require a cis disposition between these two groups. Detailed inspection of the time dependent structural evolution reveals that the ΔAla residue reaches the extended conformation after 2 and 3

ns for Val145ΔAla$_{sp3}$ and Val157ΔAla$_{sp3}$ simulations, respectively, inducing a very rapid lose of global fold coherency.

In the ΔAla$_{sp2}$ force-field the equilibrium bond angle ∠N-C$^{\alpha}$-C corresponds to a trigonal configuration (119.4º,[12a]). The influence of the ΔAla$_{sp2}$ force-field on the tested folding motif differs from that obtained with the ΔAla$_{sp3}$ force-field. Independently of the targeted position, the local organization of the β-sheet that hosts the synthetic residue is rigidified for few ns suggesting that the β-helix based building block retains the stability. However, for the Val157ΔAla$_{sp2}$ simulation the local reduction of the sheet flexibility induces adverse long range effects on the global organization, leading to collapse (Figure 2b). The loss of the global fold is the consequence of increasing the conformational flexibility of those sheet segments facing the fragment with the substituted residue (assuming that the terminal coil is an equilateral triangle, the perturbed sides would be those joined by the vertex that faces the triangle side that holds the ΔAla). At the same time, the preservation of hydrogen bonds is compromised by increasing the rigidity of the peptide backbone. Thus, its lower conformational flexibility compromises the sheet capacity to compensate for the structural fluctuations that the whole peptide undergoes under thermal agitation. All these factors were minimized for Val145ΔAla$_{sp2}$, where the temporal evolution of the mutated building block is similar to that of the wild type (Figure 2b). Thus, although the β-sheet was also rigidified by ΔAla, the conformational flexibility of the remaining secondary structures was less affected by the targeted substitution than in Val157ΔAla$_{sp2}$. In spite of this, the poor stability of the wild type building block was not enhanced by the selective incorporation of ΔAla residues.

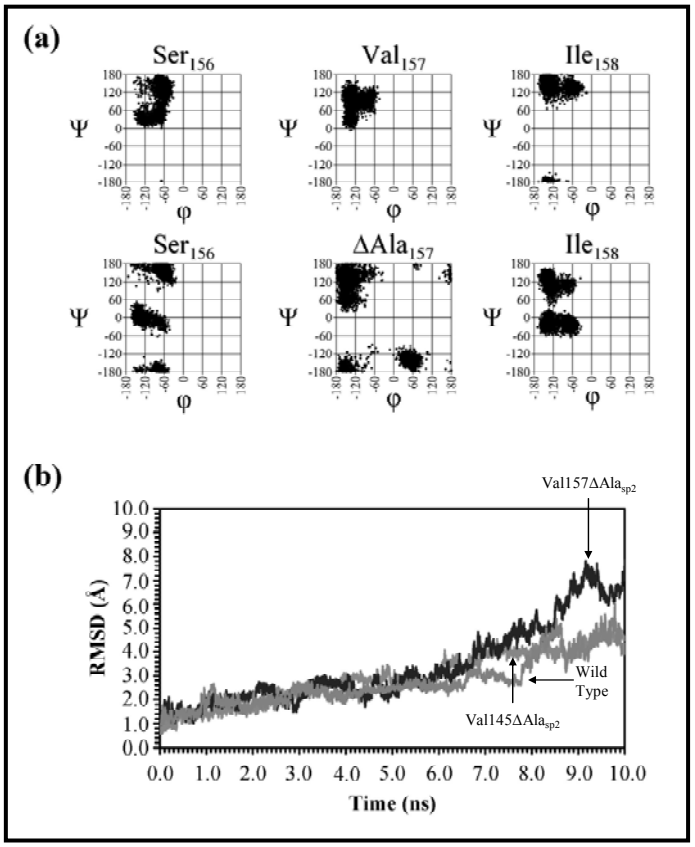


Figure 2. (a) Ramachandram plot over 10 ns of critical residues for $1krr_{131\text{-}165}$ wild type (top) and $Val157\Delta Ala_{sp3}$ (bottom) systems. Notice that the plot for $Val145\Delta Ala_{sp3}$ is analogous to that of $Val157\Delta Ala_{sp3}$. (b) Time-dependent evolution of the main chain root mean squared deviation (RMSD) for $Val145\Delta Ala_{sp2}$, $Val157\Delta Ala_{sp2}$ and wild type systems. Each magnitude was computed by comparing with the last snapshot of the equilibration.

D-Amino acids: Change in the α-carbon chirality

The effect of inverting the chirality of the α-carbon atom seems to promote local stabilization in the sheet organization, although it is not possible to obtain a general trend. Detailed analyses revealed that the residue replacements are case sensitive: depending on the position, the building block stability is improved or drastically worsened. In some cases the inversion of the chirality leads toward moderate stabilization, while in others the local topology of the motif does not allow the exchange between the α-carbon hydrogen and the side chain. This apparent site dependence is probably the consequence of the directionality in the β-helix organization, which locates each residue involved in a β-sheet with both N-H and C=O groups pointing alternatively either toward the helical growth direction or toward the opposite direction. This feature will determine the final influence that the D-amino acid holds on the sheet organization.

The first described effect, the apparent stabilization of the building block, is observed upon substitutions of Val145. In the wild type sequence, this residue has polar N-H and C=O groups pointing toward the helix direction. This topological arrangement apparently facilitates the insertion of any residue with inverted chirality at the α-carbon atom. Even though the reduction of the side chain volume decreases the local stability of the sheet, at the end of the simulations the sheet shows site acceptable adaptation to both Val145D-Ala and Val145D-Val replacements (Figure 3a). Val145D-Val demonstrates a significantly better stability than the wild type sequence under the simulated conditions, assisting in maintaining the β-helix capping coil by rigidifying the targeted position as intended. The specific role of the chirality inversion is confirmed by comparing the Val145D-Ala with Val145Ala replacement, where the latter induces a global destabilization after less than 4 ns.

On the other hand, D-replacements of Val157 always increased the local steric hindrance, which leads to a partial or complete loss of the folded building block. This evolution is a consequence of the previously mentioned β-sheet topological restraint: at this site, the α-carbon hydrogen is set in axial orientation with respect to the helical axis. The insertion of the D-amino acid clearly promotes a local disruption in the sheet conformation due to steric contacts in that region of the space. The final result was similar for all the studied cases even though the major unfolding was detected in Val157D-Ala (Figure 3b). In the other replacement, Val157D-Val, the most remarkable distortions appear in the segment that originally formed the β-sheet in which residue 157 is involved. If compared with Val157D-Ala, the lower RMSD difference with an ideal β-helix motif is a consequence of maintaining the apparent shape of the motif. Thus, the smaller conformational fluctuations are not a significant improvement: the random disruption of the native organization led to the spontaneous formation of an α-helix at the segment that held the lower strand of the lost sheet, giving an apparent image of conformational stability (Figure 4c).

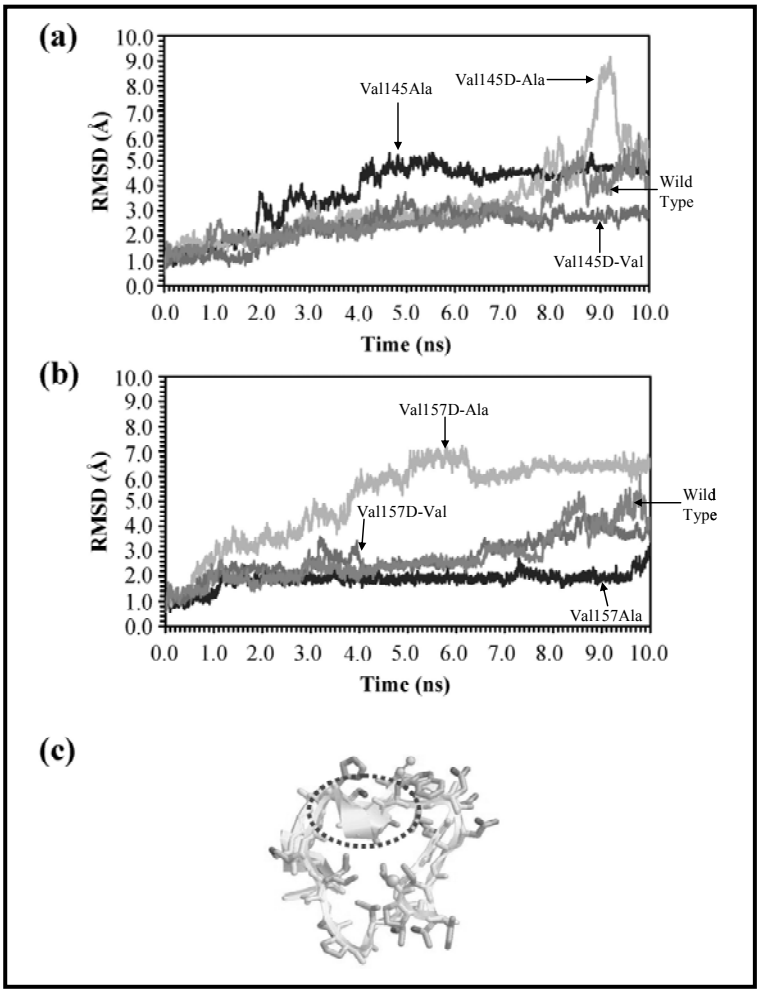


Figure 3. (a) Time-dependent evolution of the backbone root mean squared deviation (RMSD) for Val145Ala, Val145D-Ala, Val145D-Val and wild type systems. (b) The time dependence of the evolution of the backbone root mean squared deviation for Val157Ala, Val157D-Ala, Val157D-Val and wild type systems. (c) Schematic representation of the last snapshot of the Val157D-Ala simulation. The spontaneous formation of a small α-helix at the C_t is marked (see text).

β-Alanine (βAla): Incorporation of an additional carbon atom to the backbone. Insertion of omega amino acids in the β-sheet apparently produces a similar effect as that observed when D amino acids were inserted. Depending on the residue position the substitution of Val by βAla either stabilizes the motif (Val157βAla) or induces a global collapse (Val145βAla). This behavior is consistent with the overall conformational properties of the βAla residue. In Val145βAla system the total loss of the β-helix building block is preceded by a local destabilization of the sheet where the βAla residue is located: the increased residue length precludes the formation of one of the two hydrogen bonds of each residue in a parallel β-sheet arrangement. This event allows the

β-amino acid to freely explore other conformational arrangements until it reaches a stationary conformation. Thus, when the organization of the β-strand remains close to the original conformation, the βAla conformation stays close to its theoretical absolute conformational minimum.[18,30] When the sheet falls apart, this residue explores high energy conformations (only favored in monosubstituted β-amino acids[30]) and the overall organization of the polypeptide is disrupted. However, the β-Ala is free to recover its most stable conformation[18,30] when the system reaches a stable (but not regular) organization. The confirmation that this substitution is responsible for the destabilization of the β-helix motif comes from comparing with Val145Gly, where only removing the side chain does not dramatically affect the organization of both the sheet and the overall motif.

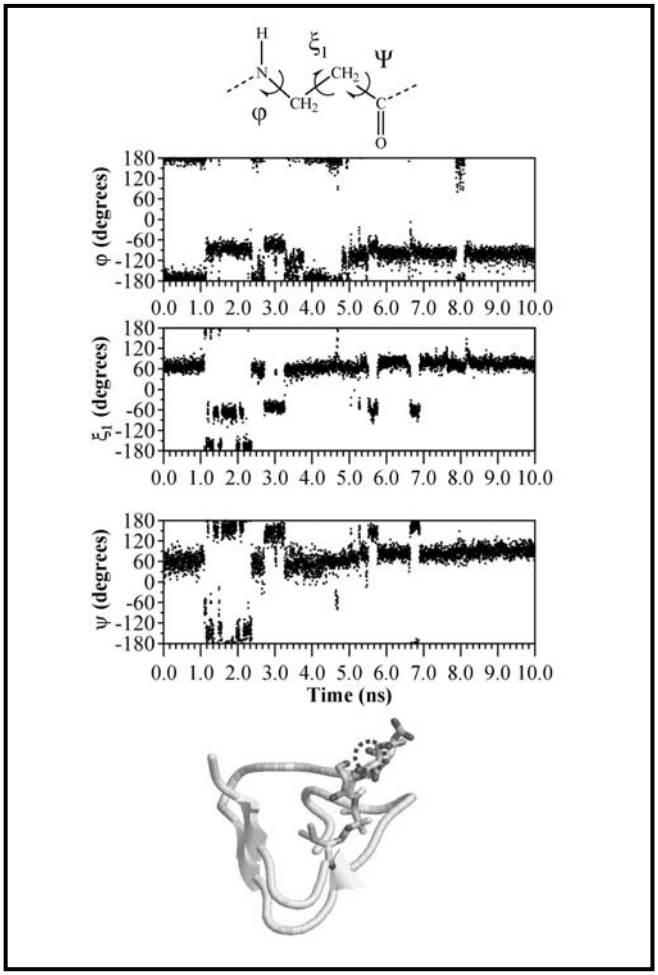


Figure 4. Top: Schematic representation of the βAla residue, illustrating the main chain dihedral angles. Middle plots: temporal evolution of the three flexible backbone dihedral angles of the βAla residue (φ,ξ_1,Ψ) along the Val157βAla simulation. Bottom: schematic representation of the final organization, where a new turn is formed. Only the residues involved in the arrangement are explicitly depicted; the hydrogen bond formed by the βAla residue is marked.

On the other hand, the noticeable reduction of the conformational flexibility observed in the Val157βAla case, can only be directly attributed to the conformational properties of βAla: almost 80% of the total simulated time, the βAla residue maintains the conformation of lowest energy (Figure 4a). This conformation sufficiently suits the topology of a β-strand consisting of alpha amino acids. Moreover, each time the residue tends to explore the same destabilizing conformation observed in the previous case, the residue rapidly reverses its conformational state back to the most stable conformation. The additional flexibility introduced by the extra methylene group aids the structural reorganization. Thus, the stability of the adjacent turn region is enhanced by the formation of a new hydrogen bond between the βAla and a neighbor residue (Figure 4b).

To conclude, the overall results point to a clear position dependence: when the β-amino acid is placed in the middle of the sheet it does not stabilize the β-helix. However, when it is set at the edge of the strand, it promotes the stabilization of the adjacent turn without any penalty on the sheet global coherency.

α,α-Dialkylated amino acids: substitution of the α-carbon hydrogen by an alkyl group. The behavior detected in the last series of simulations reflected the general trends already seen in the previous cases. The response to the substitutions depends on the position, and the aromatic side chains have a positive influence as motif stabilizer when introduced in the position 145. However, when two methyl groups are attached to the α-carbon in position 145 there is a quick conformational change that propagates through the entire strand. This behavior is mainly due to the conformational preferences of Aib that has a strong tendency to evolve toward folded conformations.[13] This local distortion induces the rupture of the β–sheet in which residue 145 was located. Although the final organization is able to maintain certain resemblance to the native state, Aib substitution clearly reduces the strand stability. On the other hand, if the side chains of the quaternary atom in position 145 correspond to two phenyl groups, the effect is significantly different. Two different factors make this substitution enhance both the local stability and the global coherency of the β-helical motif. First, the inherent conformational preferences of Dϕg residue induce a fast reorganization of the main chain conformation for the segment 144-146 (Figure 5a). This change reflects the preferred conformation for this residue type that tends to adopt extended conformations

over folded ones.[13] The abrupt readjustment introduces a distortion in the β-strand that, in this particular case, is compensated by the formation of strong side chain-side chain interactions between Ile_{163} and the phenyl lateral groups (in contrast to what occurred when ΔAla induced the formation of the same main chain conformation). The strength of these interactions is enhanced by the partial exposure to the solvent of both Ile_{163} and Dϕg side chains (Figure 6b). Once the stationary point is reached, the organization remains stable to the same extent as the wild type sequence under the same conditions.

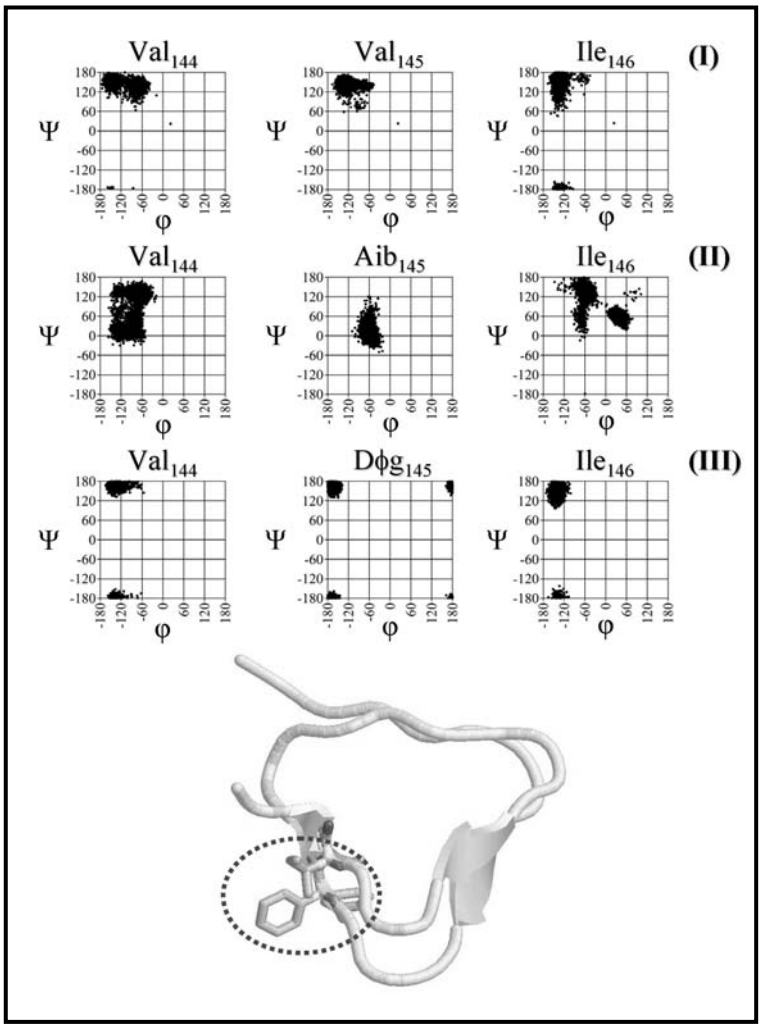


Figure 5. Top: Ramachandram plot over 10 ns of critical residues for wild type (I), Val145Aib (II) and Val145Dϕg (III) systems. Bottom: detailed schematic representation of a representative organization stabilized by the interaction between the Dϕg and Ile_{163} side chains.

Second, the substitution of the Val_{157} by either Aib or Dϕg only induces the loss of a regular organization, as shown in Figure 6. This unfavorable influence is especially remarkable in the Aib replacement. In both cases, the non-proteogenic residue imposes the formation of a pseudo-turn in the former β-sheet. This drastic structural change cannot be compensated and makes the β-helix motif unstable. In the case of Dϕg, the

motif disruption takes place over a longer period of time due to the formation of stationary local conformations that allow the phenyl side chains to temporary interact with neighboring groups. However, those organizations are not stable enough because of the lack of short-range interactions between voluminous hydrophic side chains. Therefore, the outcome of quaternary substitutions close to C_t always involves the loss of β-sheet conformation.

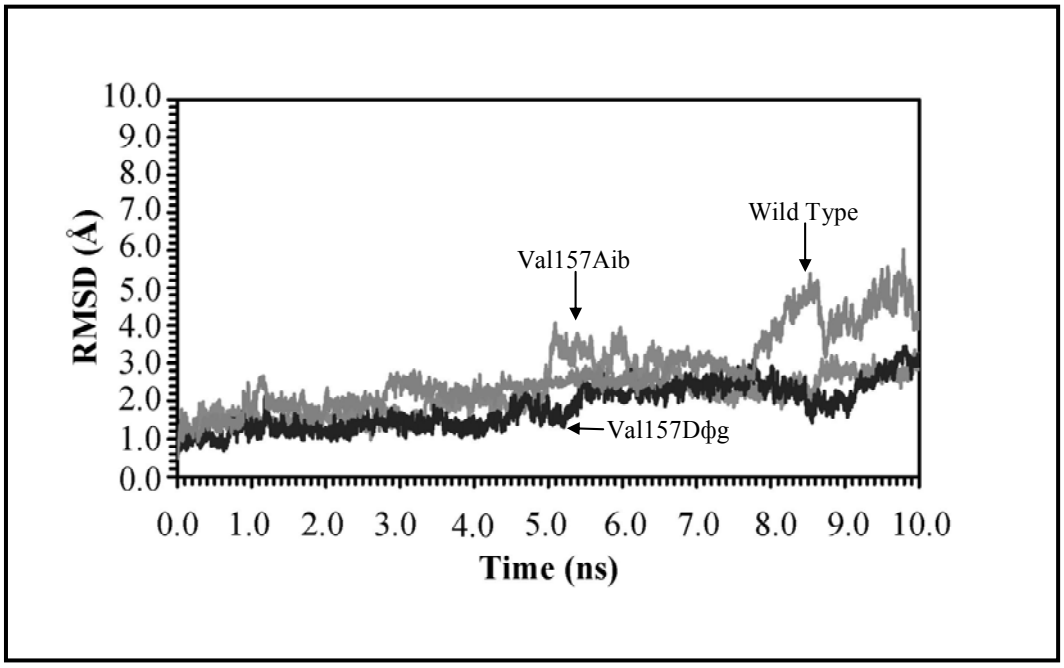


Figure 6. Time-dependent evolution of the backbone root mean squared deviation (RMSD) for Val157Dϕg, Val157Aib and wild type systems.

3.1.5. *Conclusions*

The perspective gained in this work, combined with data from previous investigations[6,7,10] indicates that the stabilization of the β-sheets in β-helical building blocks can not be understood only as the outcome of a decrease in the flexibility of the target sites. This conclusion points to a clear difference with respect to stabilizing effects of the segments that form turns in these motifs.[7,10] In the latter cases, most of the substitutions were successful; with the steric hindrance introduced in the local environment that hosted the replacement being the only limiting factor. In contrast, in β-sheets, the complexity of the problem increases. First, a reduction of the local flexibility in the targeted site might be neither sufficient nor efficient. For instance, ΔAla proved to restrict the mobility of the target β-sheets, yet the overall stability of the β-helix terminal coil was drastically reduced. Therefore, the conformational properties of the

amino acids are not the sole determinants of the stability in segments folded in β-sheets. The reduction of disorder is solely achieved when the replacement does not alter the energy balance between attractive interactions and steric hindrance. However, reaching that balance is not a sufficient condition to make synthetic residues suitable candidates to nucleate the formation of β-helix conformational motif is such segments. Only, those cases in which the substitutions promote the formation of new local conformational arrangements, the stabilization of a β-helix conformation in the capping coil segment is accomplished. For instance, Val157βAla induces a local reorganization of the strand that leads to the formation of a longer turn. This new organization reduces the conformational fluctuations of the β-helix organization.

Second, results show clear site dependence. The selective replacement by similar or even identical non-proteogenic amino acid introduces opposite effects on the local and on the overall stability of the building block with a β-helix arrangement. Substitutions of Val_{145} enhance the motif stability when the newly inserted residue does not increase the local steric hindrance. If combined with the formation of new side chain interactions, the goal is achieved. This was the case of Val145D-Val and Val145Dϕg mutations. However, the lack of new interactions invariably leads to a loss of the β-helix motif, even in those cases that showed acceptable stabilization with residues of the same family as Val145D-Ala and Val145Aib.

Among the surveyed amino acids, Dϕg has been revealed as the most suitable option to stabilize the β-sheet segments in a β-helical building block. In spite of the steric hindrance imposed by the voluminous phenyl side chains, they generally are able to form new non-covalent interactions with neighboring residues, which helps fixing the desired fold. Nonetheless, that residue does not totally fulfill the requirements since the stabilization is always preceded by structural rearrangement of the initial motif. Therefore, target replacement in β-sheet segments requires specific efforts to design new synthetic amino acids that enhance the stability of this conformation. In summary, results show that reducing the mobility of a β-sheet does not guarantee to maintain the shape of the selected building block. Nonetheless, this simple strategy is still valid and efficient when turn sites are targeted.

3.1.6. *References*

[1] Hawker C.J., Wooley K.L. *Science* **2005**, *309*, 1200.

[2] a) Percec V., Dulcey A.E., Balagurusamy V.S., Miura Y., Smidrkal J., Peterca M., Nummelin S., Edlund U., Hudson S.D., Heiney P.A., Duan H., Magonov S.N., Vinogradov A.S. *Nature* **2004**, *430*, 764. b) Reches M., Gazit E. *Science* **2003**, *300*, 625. c) Matsuura K., Murasato K., Kimizuka N. *J. Am. Chem. Soc.* **2005**, *127*, 10148.

[3] Padilla J.E., Colovos C., Yeates T.O. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 2217.

[4] Alemán C., Zanuy D., Jiménez A.I., Cativiela C., Haspel N., Zheng J., Casanovas J., Wolfson H., Nussinov R. *Phys. Biol.* **2006**, *3*, S54.

[5] a) Seeman N.C.Q. *Rev. Biophys.* **2006**, *6*, 1. b) Feldkamp U., Niemeyer C.M. *Angew. Chem. Int. Ed.* **2006**, *45*, 1856. c) Zhao X.J., Zhang S.G. *Chem. Soc. Rev.* **2006**, *35*, 1105.

[6] Tsai C.J., Zheng J., Zanuy D., Haspel N., Wolfson H., Alemán C., Nussinov R. *Proteins* **2007**, *68*, 1.

[7] Haspel N., Zanuy D., Alemán C., Wolfson H., Nussinov R. *Structure* **2006**, *14*, 1137.

[8] Main E.R.G., Lowe A.R., Mochrie S.G.J., Jackson S.E., Regan L. *Curr. Opin. Struct. Biol.* **2005**, *15*, 464.

[9] Jenkins J., Mayans O., Pickersgill R. *J. Struct. Biol.* **1998**, *122*, 236.

[10] a) Zanuy D., Jiménez A.I., Cativiela C., Nussinov R., Alemán C. *J. Phys. Chem. B*. **2007**, *111*, 3236. b) Zheng J., Zanuy D., Haspel N., Tsai C.J., Alemán C., Nussinov R. *Biochemistry* **2007**, *46*, 1205.

[11] Singh T.P., Kaur P. *Prog. Biophys. Molec. Biol.* **1996**, *66*, 141.

[12] a) Alemán C., Casanovas, J. *Biopolymers* **1995**, *36*, 71. b) Crisma M., Formaggio F., Toniolo C., Yoshikawa T., Wakamiya, T. *J. Am. Chem. Soc.* **1999**, *121*, 3272. c) Zanuy D., Casanovas J., Alemán C. *Biophys. Chem.* **2002**, *98*, 301.

[13] Casanovas J., Zanuy D., Nussinov R., Alemán C. *J. Org. Chem.* **2007**, *72*, 2174.

[14] a) Pavone V., Lombardi A., Saviano M., Di Blasio B., Nastri F., Fattorusso R., Zaccaron L., Maglio O., Yamada T., Omote Y., Kuwatam S. *Biopolymers* **1994**, *34*, 1595. b) Toniolo C., Crisma M., Fabiano N., Melchiorri P., Negri L., Krause J. A., Eggleston D.S., *Int. J. Pept. Prot. Res.* **1994**, *44*, 85.

[15] Alemán C. *J. Phys. Chem. B* **1997**, *101*, 5046.

[16] a) Karle I.L., Gopi H.N., Balaram P. *Prod. Nat. Acad. Sci. U.S.A.* **2001**, *98*, 3716. b) Roy R.S., Gopi H.N., Raghothama S., Karle I.L., Balaram P. *Chem. Eur. J.* **2006**, *12*, 3295.

[17] Muñoz-Guerra S., Fernández-Santín J.M., Rodríguez-Galán A., Subirana J.A. *Polym. Sci, Polym. Phys. Ed.* **1985**, *23*, 733.

[18] Alemán C., León S. *J. Mol. Struct.* **2000**, *505*, 211.

[19] Rai R., Raghothama,S., Balaram P. *J. Am. Chem. Soc.* **2006**, *128*, 2675.

[20] Jorgensen W.L., Chandrasekhar J., Madura J.D., Impey R.W., Klein M.L. *J. Chem. Phys***. 1983**, *79*, 926.

[21] Wang J., Cieplak P., Kollman P.A. *J. Comput. Chem.* **2000**, *21*, 1049.

[22] Preto M.A.C., Melo A., Costa S.P.G., Maia H.L.S., Ramos M.J. *J. Phys. Chem. B* **2003**, *107*, 14556.

[23] Improta R., Rega N., Alemán C., Barone V. *Macromolecules* **2001**, *34*, 7550.

[24] Harihara P.C., Pople J.A. *Chem. Phys. Lett.* **1972**, *16*, 217.

[25] Frisch M.J., Trucks G.W., Schlegel H.B., Scuseria G.E., Robb M.A., Cheeseman J.R., Montgomery J.A., Jr., Vreven T., Kudin K.N., Burant J.C., Millam J.M., Iyengar S.S., Tomasi J., Barone V., Mennucci B., Cossi M., Scalmani G., Rega N., Petersson G.A., Nakatsuji H., Hada M., Ehara M., Toyota K., Fukuda R., Hasegawa J., Ishida M., Nakajima T., Honda Y., Kitao O., Nakai H., Klene M., Li X., Knox J.E., Hratchian H.P., Cross J.B., Bakken V., Adamo C., Jaramillo J., Gomperts R., Stratmann R.E., Yazyev O., Austin A.J., Cammi R., Pomelli C., Ochterski J.W., Ayala P.Y., Morokuma K., Voth G.A., Salvador P., Dannenberg J.J., Zakrzewski V.G., Dapprich S., Daniels A.D., Strain M.C., Farkas O., Malick D.K., Rabuck A.D., Raghavachari K., Foresman J.B., Ortiz J.V., Cui Q., Baboul A.G., Clifford S., Cioslowski J., Stefanov B.B., Liu G., Liashenko A., Piskorz P., Komaromi I., Martin R.L., Fox D.J., Keith T., Al-Laham M.A., Peng C.Y., Nanayakkara A., Challacombe M., Gill P.M.W., Johnson B., Chen W., Wong M.W., Gonzalez C., Pople J.A. Gaussian 03, Revision C.02, Gaussian, Inc., Wallingford CT, **2004**.

[26] Kale L., Skeel R., Bhandarkar M., Brunner R., Gursoy A., Krawetz N., Phillips J., Shinozaki A., Varadarajan K., Schulten K. *J. Comput. Phys.* **1999**, *151*, 283.

[27] Darden T., York D., Pedersen L. *J. Chem. Phys.* **1993**, *98*, 10089.

[28] Ryckaert J.P., Ciccotti G., Berendsen H.J.C. *J. Comput. Phys.* **1977**, *23*, 327.

[29] Berendsen H.J.C., Postma J.P.M., van Gunsteren W.F., DiNola A., Haak J.R. *Chem. Phys.* **1984**, *81*, 3684.

[30] Wu Y.D., Wang D.P. *J. Am. Chem. Soc.* **1998**, *120*, 13485.

3.2. Stability of tubular structures based on β-helical proteins: self-assembled versus polymerized nanoconstructs and wild-type versus mutated sequences*

3.2.1. *Introduction*

Complex multi-molecular structures can be usefully designed from biological macromolecules[1,2] via non-covalently self-assembled 'foldamer' building blocks.[3] Self-assembly of peptides and proteins,[4-12] DNA[13,14] and RNA segments[15] into novel nanostructures have been reported. In these nanostructures, specially in those derived from proteins, it was proposed that suitable building blocks are composed by fragments able to retain in solution a conformation similar to that observed in the native protein structure.[11,12] Thus, manipulation of such pre-existing protein foldamers may provide stable nanostructures through a favorable association process.[16-18]

Previously, computer simulation methods were used to design tubular nanostructures based on β-helical proteins.[19] 17 building blocks were selected from native left-handed β-helical proteins by slicing β-helices into two-turn repeat units. Four copies of each structural unit were stacked one atop the other, with no covalent linkage between them. The stability of these self-assembled tube organizations was investigated through relatively short (20-40 ns) Molecular Dynamics (MD) simulations. Constructs which preserve their organization in the simulation are candidates for experiments. It was observed that a structural model based on the self-assembly of a two-turn repeat motif from *E. coli* galactoside acetyltransferase (PDB code 1krr, chain A) produced a very stable tube. Mutational studies revealed that the stability of the self assembly is enhanced when the mobility in the loop regions is reduced or unfavorable electrostatic interactions are eliminated.[19-21] In addition, technological applications as the formation of nanofibers to transfer charge through p-electron stacking or through H^+ transfer have been proposed.[22]

Here, the goal is to design through atomistic computer simulations a stable tubular structure using naturally occurring repeat building blocks. Current techniques do not allow simulations of the self-assembly processes. However, the systematic use of MD to screen potential assembly arrangements has proven to be an efficient and reliable tool to predict the ultrastructure of amyloid fibrils by discarding unfavorable association modes in small model systems.[23-26] In the present work the stability of self-assembled versus

* Results presented in this chapter are published in *Biomacromolecules* **2007**, *8*, 3135

covalently-linked repeats is studied. Regan and coworkers have recently shown that copies of structural units covalently linked by spacers lead to stable supramolecular structures.[16,18] Accordingly, it is investigated if this strategy can be extended to tubular nanostructures based on polymerized β-helical repeats using MD simulations. In addition, it is examined the ability of a given conformationally restricted synthetic residue to enhance the stability of the nanotubular construct. Specifically, 1-aminocyclopentane-1-carboxylic acid (Ac_5c) is the selected amino acid, a simple cyclic α,α-dialkylated amino acid with strong stereochemical constraints induced by the cyclopentane ring (Figure 1a). A distinct structural feature of Ac_5c is that the backbone dihedral angles φ,ψ are not as geometrically restricted by the five-membered cycle as in Proline, even though the steric interactions produced by this bulky side group constrain its conformational flexibility.[27] Taking into account that the stability of the nanotubular construct is enhanced by the incorporation of proline in the loop region,[19] the present study is devoted to analyze if Ac_5c is a good candidate to increase the stability of both the repeats and the tubes formed by self-assembling and by polymerization.

On the other hand, incorporation of synthetic residues into naturally occurring building blocks may also increase resistance towards in vivo proteolysis, due to the low selectivity of proteolytic enzymes with respect to the newly incorporated amino acids; while at the same time, the presence of new functional groups may also redesign their biological function.[28,29]

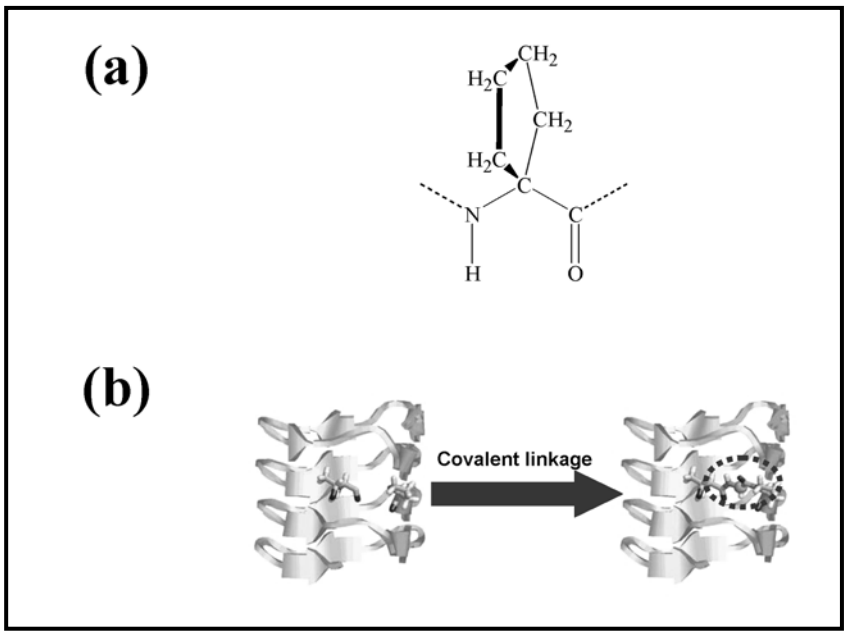


Figure 1. (a) Chemical structure of the Ac5c residue. (b) Formation of a covalent link between two building blocks of 1krr. The Gly residue used as linker between the two building blocks is specifically displayed.

3.2.2. Description of the β-helical building blocks and the Ac_5c synthetic residue

Previously[19] left-handed β-helical building blocks were selected for nanotubular structural design according to the following criteria: (i) they contain highly repetitive, symmetrical building blocks allowing formation of nanofibers without performing many structural manipulations;[16] (ii) they mostly occur in or near active or binding sites, thus likely to be of functional importance;[30-36] and (iii) as compared with right-handed β-helices, left-handed β-helices exhibit small variability of shape, size and sequence.[37] Left-handed β-helices display an equilateral triangular shape and highly repetitive sequence while right-handed β-helices are less regular. It was found that a nanotube constructed of four stacked replicas of the left-handed β-helix formed by residues 131-165 of 1krr exhibited remarkable stability under different simulated conditions, including temperature increase and addition of ions.[19] On the other hand, the less stable model was obtained from four self-assembled copies of the β-helix formed by residues 296-329 of N-acetylglucosamine 1-phosphate uridyltransferase GlmU, C-terminal domain from *E. coli* (PDB code 1hv9). Thus, nanoconstructs formed by this repeat are good systems to test stabilization strategies, and models constructed from both 1krr and 1hv9 have been used to compare the stability of the nanotubes produced by self-assembly versus covalent polymerization. A description of the sequences used to create the models based on 1krr and 1hv9 repeats is provided in Table 1. Each sequence contains a repetitive helical strand-loop motif, where the peptide backbones alternate between β-strands and loops. X-ray crystal structures of both 1krr and 1hv9 show that these two protein fragments have an almost perfect equilateral triangular shape, with each side being ~18 Å.

Table 1. Sequences of the wild type 1krr and 1hv9 building blocks used to construct the nanotubes. For each fragment the mutated residues are highlighted with both bolding and italics.

PDB	*Protein name*	*Residues*	*Sequence*
1krr	*Galactoside acetyltransferase from e. coli*	*131-165*	*PITIGNNVWIGSHVVIN***P** **G***VTIGDNSVIG***A***GSIVT*
1hv9	N-acetylglucosamine 1-phosphate uridyltransferase GlmU, C-terminal domain from e. coli.	296-329	*CVIKNSVIG***DD***CEISPY TVVEDANLAA***A***CTIGPF*

Studies of peptides containing the Ac_5c residue have indicated that this synthetic amino acid presents a restricted conformational space with a high propensity to adopt folded conformations.[27,38,39] This conformational characteristic makes Ac_5c a potential candidate to reduce the conformational freedom of the β-helix building blocks if it is introduced in the most mobile regions, *i.e.* the folded loops, replacing natural amino acids. Furthermore, the strain energy associated to the cyclopentane ring in Ac_5c is significantly lower than that of other constrained amino acids with cyclic side chains recently investigated.[20,21] Accordingly, Ac_5c is expected to adapt its folded conformation within the targeted position more easily than the corresponding analogue with a cyclopropane ring. Three Ac_5c-single mutations were considered in this work for each β-helix building block, which were selected based on a previous study (see below).[19] The mutated positions for the 1krr and 1hv9 fragments are displayed in Table 1.

The results presented in this chapter demonstrate the promising utility of the conformationally restricted amino acids as key pieces to bias the conformational preferences of naturally occurring building blocks. In all cases, favorable formation of coherently organized nanostructures was observed when Ac_5c is introduced in selected positions of the β-helical building block, independently of the conditions studied and the protein segment chosen for the experiments. Thus, nanotubes built by the assembly of the 1krr motifs get extra stabilization when Gly 149 is substituted by Ac_5c. Analogous results are obtained when Ac_5c replaces Asp at the position 306 of 1hv9. On the other hand, the promotion of polymerization as compared to self-assembly always increases the thermal stability of the constructs, even when the assembly is not favored as in 1hv9.

3.2.3. *Methods*

Computational details

MD simulations were performed using the NAMD program.[40] Each simulated system was placed in the center of an orthorhombic simulation box filled with explicit water molecules, which were represented using the TIP3 model.[41] Positively charged sodium atoms were added to the simulation box in the required amount to reach electric neutrality (all considered building blocks had negative net charge at neutral pH). All

atoms of both building blocks and the nanotubes were considered explicitly. The details for each simulated system are provided in Table 2.

Table 2. Details of the simulated systems[a] studied in this work: total number of particles (N_{atoms}), number of solvent molecules (N_{Water}), number of sodium atoms (N_{Na+}) and size of the simulation box at the beginning of the simulation.

System	N_{atoms}	N_{Water}	N_{Na+}	Initial box size(Å)	Simulation Conditions
Nanotubes					
Self-assembed 1krr	56238	18040	4	80×80×110	NPT 298K NVT 350K
Self-assembled 1hv9	56084	18032	16	80×80×110	NPT 298K NVT 350K
Polymerized 1krr	56226	18041	4	80×80×110	NPT 298K NVT 350K
Polymerized 1hv9	56129	18059	16	80×80×110	NPT 298K NVT 350K
Building blocks					
1krr	17474	5656	1	56×56×56	NPT 298K
1hv9	17465	5656	4	56×56×56	NPT 298K

[a] Mutated systems in which proteogenic residues have been replaced by Ac_5c have not been explicitly specified since the variation in number of particles with respect to the wild-type system is very small.

The energy was calculated using the AMBER force-field,[42,43] with the required parameters taken from the AMBER libraries for all the residues with the exception of Ac_5c. A recently developed parametrization[27] that is fully consistent with AMBER parameters was used for this conformationally constrained residue. Atom pair distance cutoffs were applied at 14.0 Å to compute the van der Waals interactions. The electrostatic interactions were computed using the non-truncated electrostatic potential with Ewald Summations.[44] The real space term was determined by the van der Waals cut off (14.0 Å), while the reciprocal term was estimated by interpolation of the effective charge into a charges mesh with a grid thickness 5 points per volume unit, *i.e.*

particle-mesh Ewald (PME) method.[44] Bond lengths were constrained using the *SHAKE* algorithm,[45] with a numerical integration step of 2 fs.

Before the production series, the thermodynamic variables of the system were equilibrated. The energy of each system was initially minimized to relax conformational and structural tensions using the conjugate gradient method for at least $5 \cdot 10^3$ steps, *i.e.* the system was subjected to minimization until the difference in energy was less than a threshold of 10^{-6} kcal/mol or the norm of the gradient for two successive steps in the minimization was less than 0.1 kcal/mol·Å. Next, different consecutive rounds of short MD runs were performed in order to equilibrate the density, temperature and pressure. First, solvent and charged sodium atoms were thermally relaxed by three consecutives runs, while the protein parts were kept frozen. 0.5 ns of NVT-MD at 500 K were used to homogeneously distribute the solvent and ions in the box. Second, 0.5 ns of isothermal and 0.5 ns isobaric relaxation were run. Finally, all the atoms of the system were submitted to 0.15 ns of steady heating until the target temperature was reached (298 K), 0.25 ns of *NVT*-MD at 298 K (thermal equilibration) followed by 0.5 ns of density relaxation (*NPT*-MD). Both temperature and pressure were controlled by the weak coupling method, the Berendsen thermo-barostat,[46] using a time constant for heat bath coupling and a pressure relaxation time of 1 ps. The end of the density relaxation simulation was the starting point of the molecular simulations presented in this work. All the building block systems were simulated at 298 K and constant pressure of 1 atm. On the other hand, all the nanotube systems, both self assembled and covalently linked models, were simulated using the aforementioned conditions to assess their stability at physiological conditions and at higher temperature to partially study their kinetic stability. Thus, in those simulations the end of NPT relaxation was used as starting point for an NVT simulation at 350 K (more than 50 degrees over standard conditions). The coordinates of all the production runs, which were 10 ns long, were saved every 500 steps (1 ps intervals) for subsequent analysis.

Structural analyses

The conformational stability and conservation of the nanostructures was measured by calculating: (i) the evolution of the backbone root mean square deviation (RMSD) through the simulation relative to the initial structure; and (ii) the root mean square fluctuation (RMSF) of individual residues averaged over the whole simulation. Both

RMSD and RMSF were computed with respect to the backbone atoms (-N-C^{α}-C-). The distribution of the backbone dihedral angles φ,ψ was used to show the mobility of the loop regions. The conformation of a given residue is more restrained when the φ,ψ distribution is narrower.

3.2.4. *Results and discussion*

Self-assembled versus polymerized nanotubes derived from wild type 1krr and 1hv9

The overall stability of the self-assembled and polymerized nanotubes of 1krr and 1hv9 was measured by computing the backbone RMSD with respect to the initial structure. Analysis of the RMSDs for simulations performed at 298 K, which are displayed in Figure 2a, indicates that the self-assembled structure of 1krr is very stable (average RMSD= 1.43±0.12 Å), while the self-assembled construct of 1hv9 exhibits significant conformational distortions that lead to an increase of the RMSD beyond 2 Å after 1 ns of simulation (average RMSD= 2.18±0.21 Å). Figure 3a compares the two self-assembled structures following 10 ns of simulation, while Figure 2c plots the RMSF of these two systems. The low stability of the 1hv9 structure is mainly due to the distortions of the loop regions, especially those formed by segments 304-306 and 321-323. In contrast, no significant fluctuation was found for the self-assembled structure of 1krr after 10 ns. Although no qualitative change is detected when these results are compared with those derived from simulations at 350K (Figure 2b), self-assembled structures show significant sensitivity towards thermal strain. Thus, the self-assembled structure of 1krr exhibits much higher structural fluctuations at 350 K than at 298 K, even though the core of the assembly remains stable in both cases. Inspection to the RMSF clearly shows that the main distortions are located at the building blocks of the assembly edges, suggesting that this greater flexibility might mostly be a reflection of the finite size of the system. On the other hand, the self-assembled construct based on the 1hv92 building block exhibits the same behavior as previously seen at 298K. However, at 350 K the destabilization occurs in a much shorter time lapse (the RMSD gets to 2.5 Å even before reaching 1 ns).

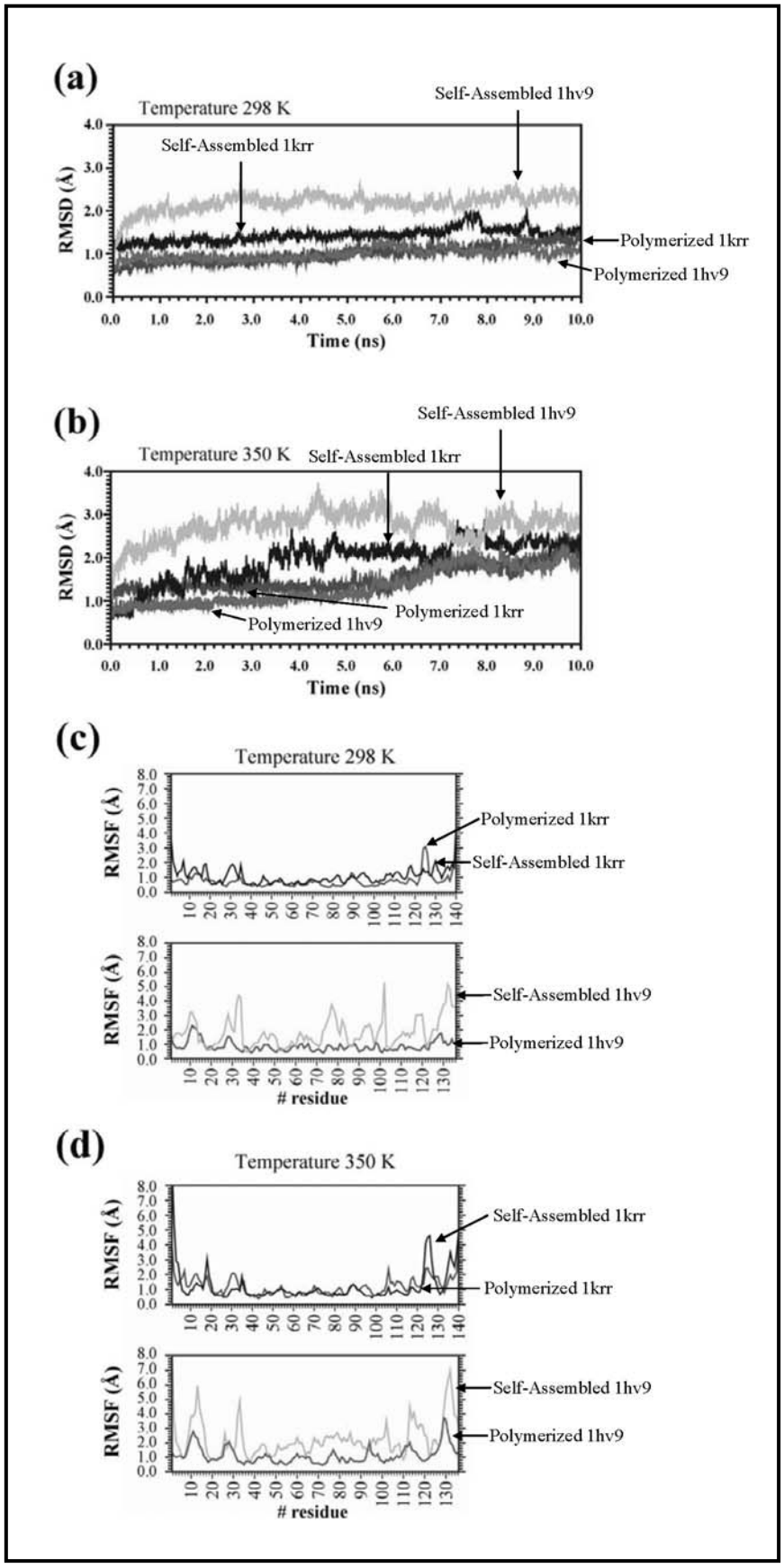


Figure 2. Evolution of the backbone RMSD of the simulated wild type nanotube models with respect to the initial structure at 298 (a) and 350 K (b): self-assembled 1krr, self-assembled 1hv9, polymerized 1krr and polymerized 1hv9. (c) RMSF of the 1krr (top) and 1hv9 (bottom) polymerized nanotubes in comparison with their corresponding self-assembled ones at 298 (c) and 350 K (d)

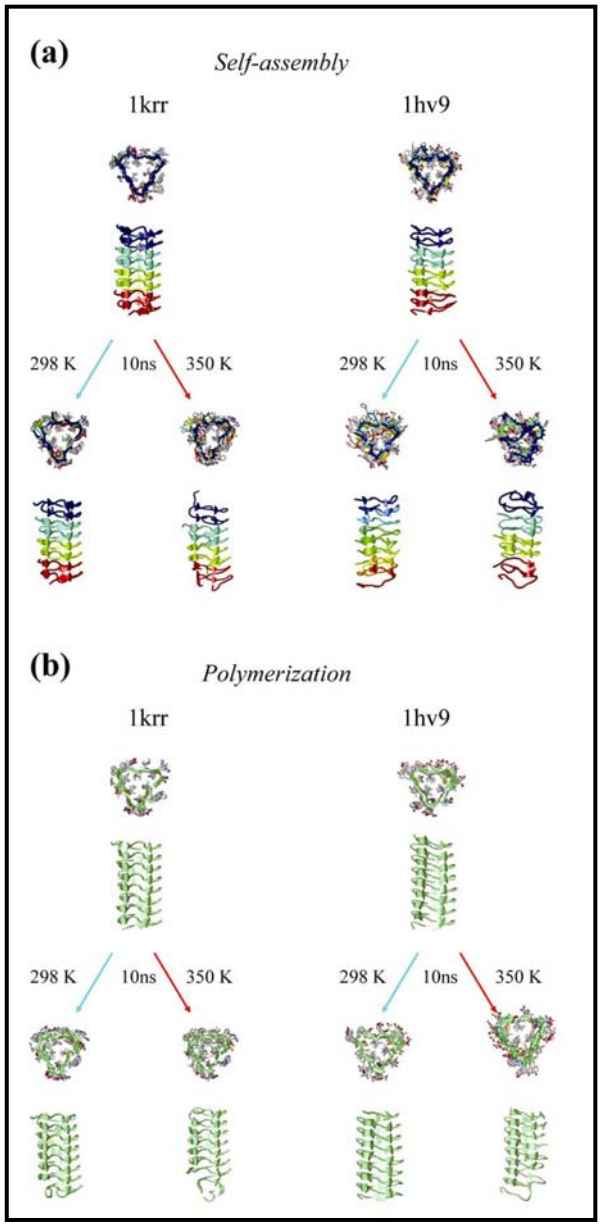


Figure 3. Structure of self-assembled (a) and polymerized (b) nantubes obtained from both 1krr and 1hv9 at the beginning and after 10 ns of MD simulation. In all cases the arrows pointing to the left and right of the starting structures precede the organization sustained at 298 and 350 K, respectively. Note that for the assembled structures each building block is highlighted with different color.

Moreover, these results are in excellent agreement with reference 19, even though different MD conditions were used. Previous simulations were performed with the CHARMM22 force-field using a simulation box of dimensions 50×50×70 Å with *NVT* conditions at higher temperature (330 K), while the present work was carried out with AMBER force-field[42,43] using an initial box size of 80×80×110 Å and the density was previously relaxed at 298 K and 1 atm (*NPT* conditions), while a 350 K simulations

were run in parallel. Hence, it can be concluded that the predictions do not depend on the force-field or on the simulation conditions.

The polymerized nanotubes of 1krr and 1hv9 were constructed by linking covalently the corresponding building blocks. To successfully link these polymeric nanostructures, the covalent linkage of several identical building blocks should guarantee the formation of the desired fold. To provide a homogeneous and regular β-helix fold, a Gly residue which is not expected to alter the chemical nature of the nanoconstruct due to the lack of side chain, was used. Figure 1b shows the position of the Gly linker between two consecutive 1krr building blocks. In contrast, no linker was needed for the polymerized nanotube of 1hv9, as a regular β-helix arrangement was obtained when the building blocks are linked directly. This structural difference between the nanoconstructs formed by 1krr and 1hv9 was found to be crucial for understanding the stabilities of both self-assembled and polymerized tubes (see below).

The evolution of the RMSDs along the simulations for the polymerized nanotubes is included in Figures 2a and 2b (simulations at 298 and 350 K). The structures obtained at the end of each trajectory are displayed in Figure 3b. Finally, Figures 2c and Figures 2d compare the residue-based fluctuation of the polymerized structures with their corresponding self-assembled ones, at 298K and 350K, respectively. For both 1krr and 1hv9 the polymerized structures are considerably more stable than the self-assembled ones, even though the most significant improvement was obtained for the construct derived from the 1hv9 building block. The average RMSDs are 0.98±0.20 Å and 0.98 ± 0.13 Å for the tubes formed by polymerizing the 1krr and 1hv9, respectively, showing a reduction of about and 1.2 Å with respect to the self-assembled analogues. The behavior exhibited by both the polymerized and the self-assembled tubes at 350K is qualitatively analogous to that previously described at 298K. However, in all cases the structural fluctuations increase with the temperature. Figure 2d illustrates such behavior, as can be observed by the homogeneous increment of the structural fluctuations along all the studied sequences when temperature is raised.

On the other hand, a detailed comparison of Figures 2a and 2b illustrates how covalent linkages between building blocks reduce the unfolding velocity, *i.e.* polymerization provides not only thermal stabilization but also enhances the kinetic stability of the nanoconstructs. While self-assembled models rapidly lose global organization, those

that were covalently linked show similar behavior to that exhibited at 298 K (Figure 3b).

A detailed analysis of the snapshots recorded during the simulation allows explaining the differences found between the nanostructures formed by these two β-helical motifs. As mentioned above, no linker was required to join two consecutive building blocks of 1hv9 indicating that the distance between the end of a repetitive unit and the beginning of the next one is close to that typically found for a chemical bond. This distance, which is excessively short for self-assembling, together with the low inherent conformational stability of the 1hv9 building block, precludes the stabilization of the self-assembled nanoconstruct and induces the unfolding of the loop regions. On the other hand, a larger distance between consecutive building blocks was obtained for 1krr. This distance was suitable for the stabilization of the tube formed by the self-assembled building blocks since optimum nonbonding interactions were obtained between successive repeating units. Finally, the behavior of polymerized nanoconstructs can be straightforwardly explained: the stability of regular and homogeneous conformations increases with the size of the chain.

Overall these results allow us to conclude that the stability of the nanotubes depends on how efficient is the match between consecutive building blocks. Homogeneous β-helix folds can be obtained by forming stable assemblies of building blocks aligned in a complementary (self-assembled) manner or by introducing suitable linkers to retain the backbone hydrogen bonds and the electrostatic interaction in the β-sheet and loop regions.

Stability of building block mutants of 1krr and 1hv9

Three positions of 1krr and 1hv9 were selected as suitable candidates for substitution by Ac_5c according to the following criteria: i) the residues are located in the loop regions, which display higher mobility than the β sheets and present a turn-conformation similar to that preferred by Ac_5c; ii) the side chain of the residues is outward-pointing avoiding unfavorable steric interactions between the cyclopentane group of the Ac_5c substitution and the side chains of the inward-pointing residues; and iii) the charged side chain of the Asp-305 and Asp-306 residues produces electrostatic repulsions in the loop regions of 1hv9. Thus, targeting flexible and electrostatically disfavored loop regions should have direct impact on the structural stability of the nanoconstructs.

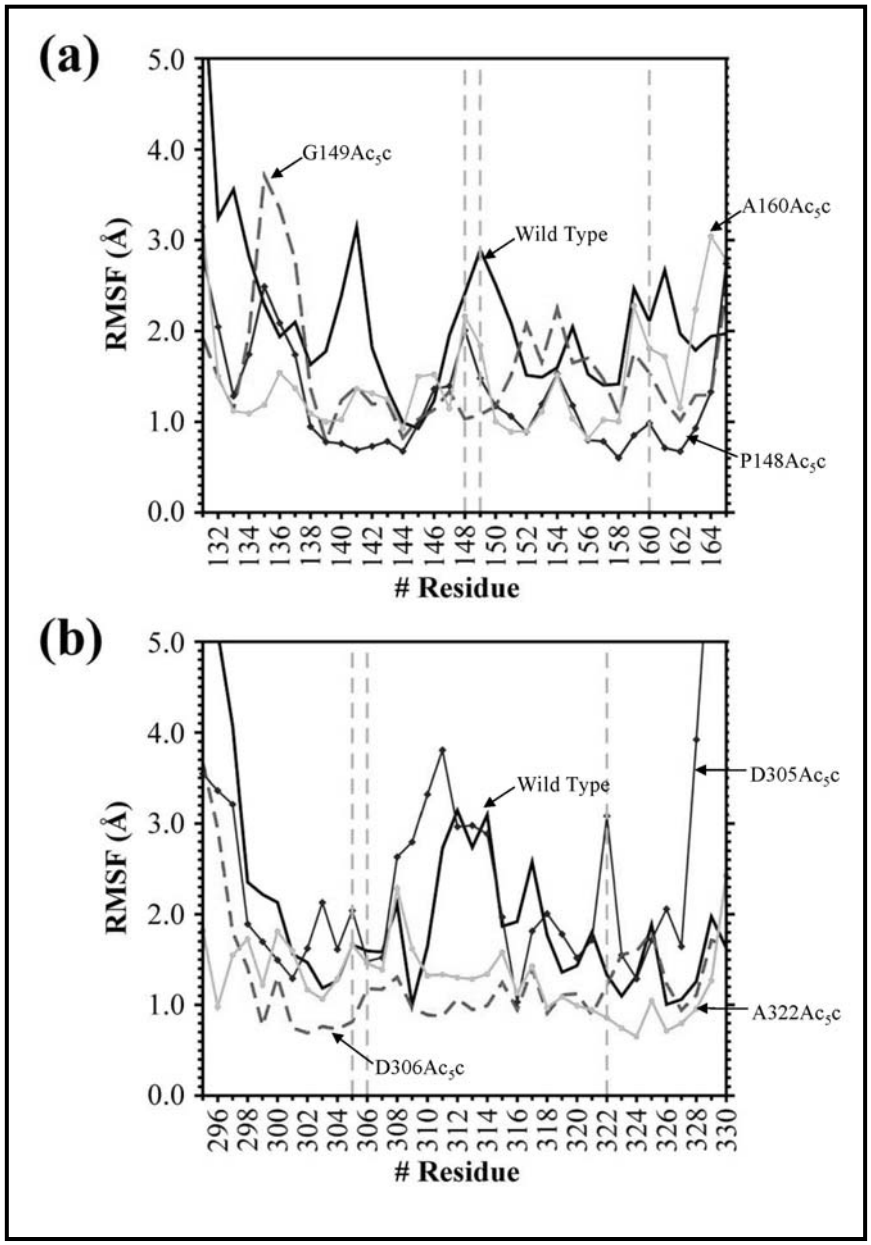


Figure 4. Comparison of the RMSF for the wild type and mutated building blocks. (a) Building blocks based on 1krr: wild type, P148Ac_5c, G149Ac_5c and A160Ac_5c. (b) Building blocks based on 1hv9: wild type, D305Ac_5c, D306Ac_5c and A322Ac_5c. Gray dashed lines indicate the position of the substitutions.

For the 1krr building block the Pro, Gly and Ala residues at positions 148, 149 and 160 are substituted by Ac_5c one at a time (Table 1), with the three corresponding mutants denoted P148Ac_5c, G149Ac_5c and A160Ac_5c, respectively. The single mutations performed on the 1hv9 building block at Asp-305, Asp-306 and Ala-322 were denoted D305Ac_5c, D306Ac_5c and A322Ac_5c, respectively. Figure 4 compares the RMSF of the wild type building blocks with the corresponding mutants, while Figure 5 displays for each case the structure obtained after 10 ns of MD simulation. As can be seen, for the

1krr mutants the highest local improvement was obtained for G149, the averaged RMS of residue 149 decreasing by about 2 Å. Thus, the substitution of a flexible Gly residue by an Ac_5c, which is constrained to adopt turn conformations, stabilizes the bend architecture of the loop by reducing its mobility. This is reflected in Figure 6a, which shows the Ramachandran plots of the substituted positions and the adjacent residues. As can be seen, the conformations of the Pro-148 and Val-150 residues are significantly more restrained in the mutant than in the wild type.

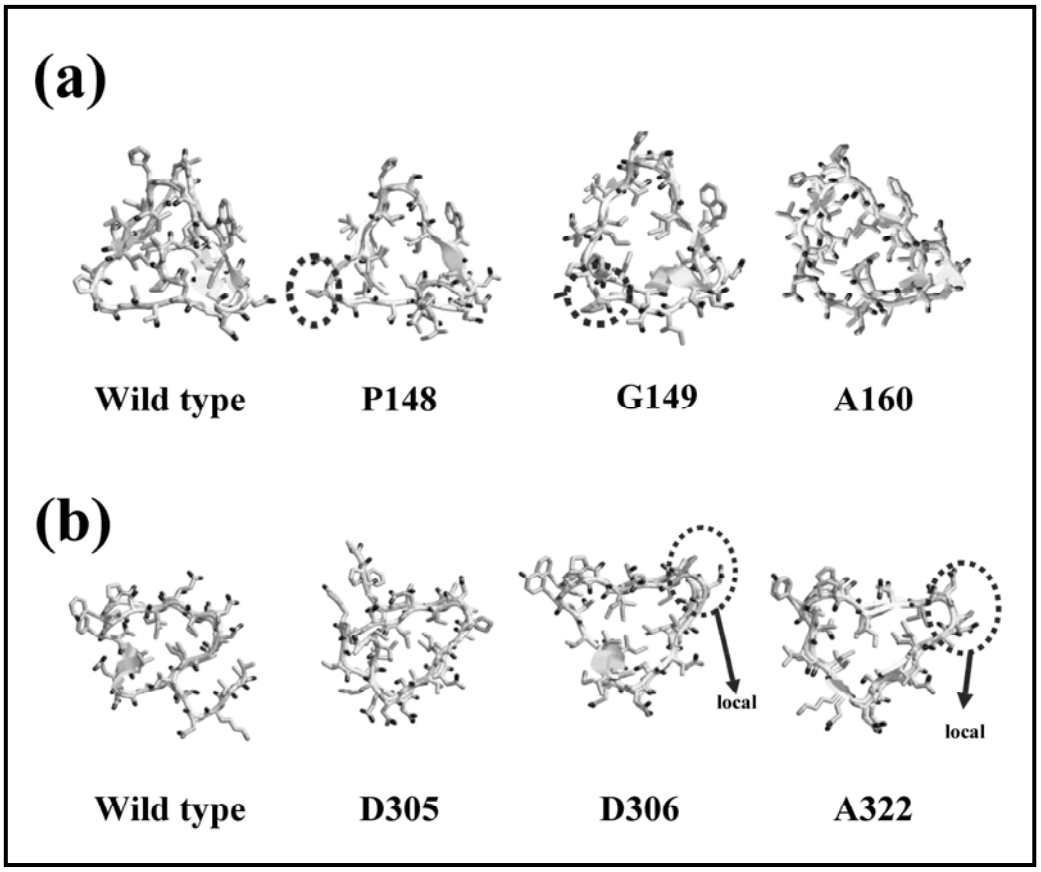


Figure 5. Structure of the wild type and mutated building blocks of 1krr (a) and 1hv9 (b) after 10 ns of MD simulation. The hydrogen atoms have been omitted for clarity and the backbone has been represented by solid shapes (arrows indicate sheet conformation). Circled zones correspond to explicit structural effects that are highlighted in the text.

The P148Ac_5c and A160Ac_5c mutants also display a stabilization of the substituted position, although this local improvement was smaller than 0.5 Å (Figure 4a). In the case of P148Ac_5c this is an expected result. Since Pro also prefers turn conformations,[47] this substitution is less important than that of G149Ac_5c. Regarding A160Ac_5c, the snapshot recorded at the end of the simulations (Figure 5a) reveals a significant distortion in the turn at positions 148-150. This segment is located between both N- and C -termini edges, implying that it is susceptible to undergo conformational changes

when those edges start fraying. Figure 5a also evidences the reorganization of the turn between residues 159-161, which involves the substituted position. Thus, in A160Ac$_5$c the conformational rigidity of this loop is enhanced inducing a global chain effect that causes the edges to collapse. In the wild type case, the loop spontaneously evolved towards an alternative organization that favors the preservation of the β–helix motif (Figure 5a). In summary, the flexibility of the loop containing Ala-160 is crucial for the global organization of the building block.

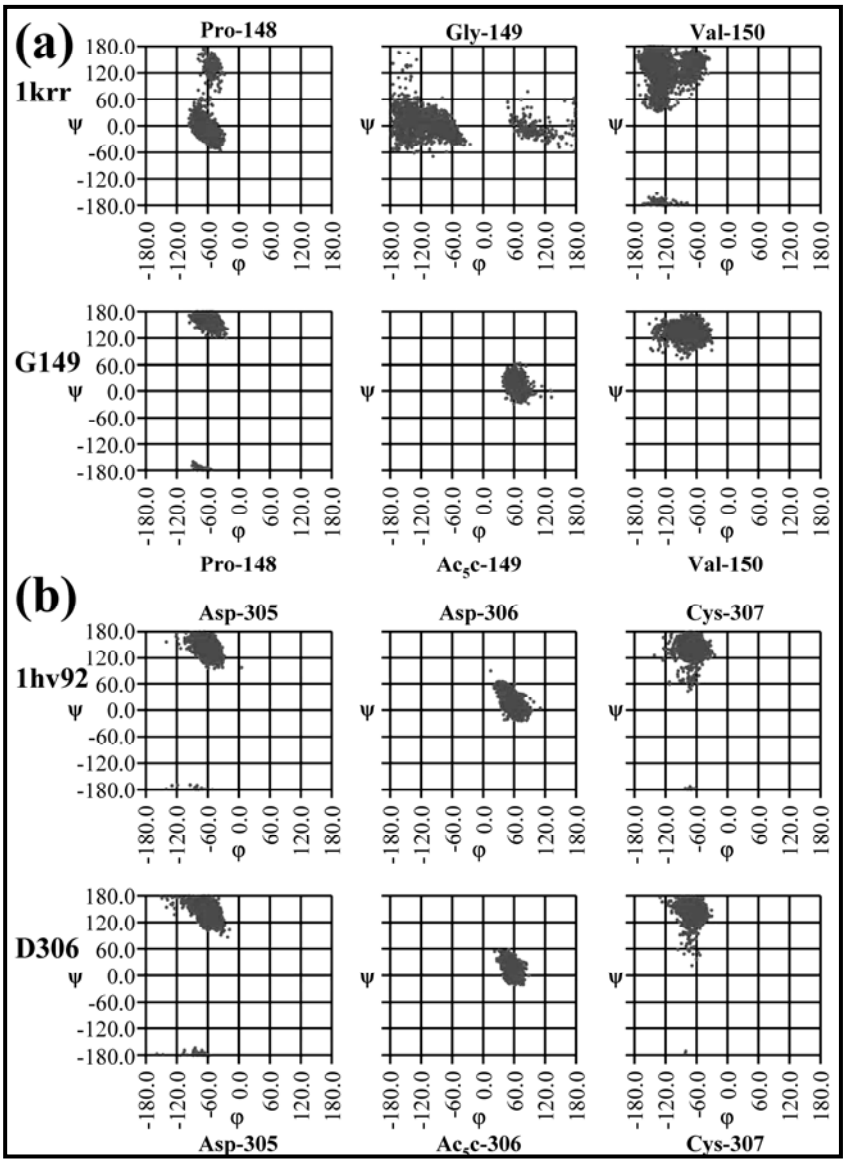


Figure 6. Comparison of the backbone torsion angle (φ, ψ) distributions for the more significant building blocks. The Ramachandran plots correspond to the targeted position and its adjacent residues in (a) 1krr wild type and G149Ac$_5$c mutant and (b) 1hv9 wild type and D306Ac$_5$c mutant.

Results for the wild type building block of 1hv9 and its three mutants are displayed in Figure 4b and 5b. As can be seen, substitution at Asp-305 does not provide neither local nor global improvement. The structural distortion is significant and similar for both the wild type and the D305Ac_5c mutant, with the β-helix disrupted in both cases. In contrast, a notable structural stability is displayed by the D306Ac_5c mutant. In this case, the RMSF is considerably smaller not only at the substitution site but also at all other positions. Thus, the initial β-helix conformation is retained without apparent distortions after 10 ns of MD simulation. Inspection of the Ramachandran plots of the substituted position and the adjacent residues, displayed in Figure 6b, reveals a significant resemblance between the wild type and the D306Ac_5c mutant. Even though the substitution at position Asp-306 by Ac_5c eliminates the electrostatic repulsion with Asp-305, the backbone constraints associated with the Ac_5c residue have a significantly smaller effect in the fluctuations at the loop than at the other positions. As a consequence, a substantial global stabilization of the building block is obtained. Finally, the A322Ac_5c mutant shows acceptable local and global stabilization, which is a consequence of the restrictions imposed by the Ac_5c on the backbone conformation.

Stability of the self-assembled and polymerized nanotubular mutants of 1krr and 1hv9

G149Ac_5c and D306Ac_5c were selected as building blocks for self-assembled and polymerized nanotubes. These two mutants provided the largest local and global stabilization for the β-helix motifs derived from 1krr and 1hv9, respectively. The strategy used to build the structures of both self-assembled and polymerized nanotubular mutants was identical to that described above for the wild-type nanoconstructs. The polymerized nanotube of G149Ac_5c only formed a homogeneous β-helix when a Gly was introduced as linker residue. Figure 7 compares the evolution of the RMSD calculated for nanotubes of the mutated building blocks with those obtained using wild-type repeats, while Figure 8 compares the RMSF. In Figure 7, the simulations performed at 350 K are also included, while Figure 8 only depicts the RMSF at 298 K (see text below).

As can be seen, the RMSD of the G149Ac_5c self-assembled system is considerably smaller than its corresponding wild type system. At 298 K, the backbone RMSD of the self-assembled G149Ac_5c remains close to 1 Å for the whole simulation indicating the stability of this self-assembled organization is remarkably high. No significant

difference was found in the RMSF and Ramachandram plots of all the 1krr based tubes when temperature is raised to 350 K. Thus, further results related to the tubes' ultra-structure will only be described for the simulations performed at 298 K.

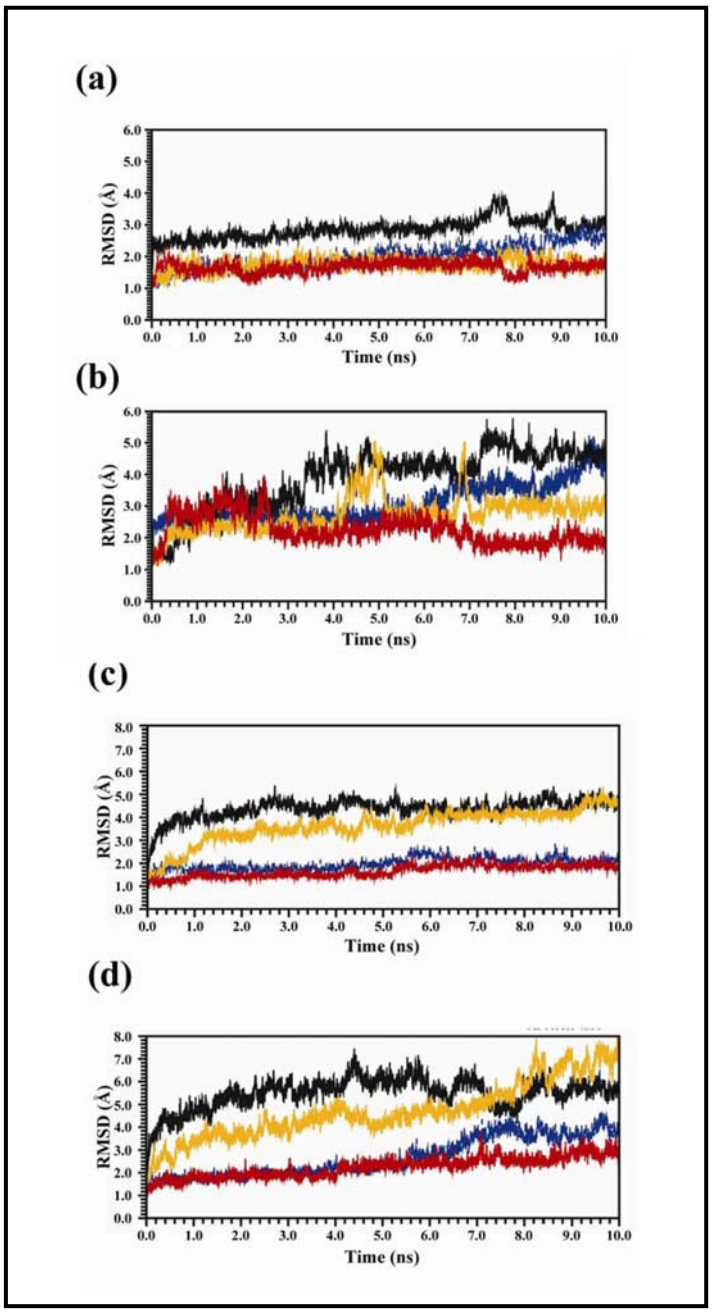


Figure 7. Evolution of the backbone RMSD of the simulated nanotube models based upon both wild type sequences and mutants with respect to the initial structures. In the plot are compared assemblies and polymerized structures for each case. For 1krr based nanotubes: wild type self-assembled nanotube (black line), wild type polymerized nanotube (blue line), G149Ac$_5$c self-assembled nanotube (orange line) and G149Ac$_5$c polymerized nanotube (red line) simulated at (a) 298 K and (b) 350 K. For 1hv9 based nanotubes: wild type self-assembled nanotube (black line), wild type polymerized nanotube (blue line), D306Ac$_5$c self-assembled nanotube (orange line) and D306Ac$_5$c polymerized nanotube (red line) simulated at (c) 298 K and (d) 350 K.

The aforementioned G149Ac$_5$c behavior clearly demonstrates that the substitution of Gly-149 by Ac$_5$c not only reduces the flexibility of the building block but also enhances its ability to retain the assembled structure constructed using the mutated subunits. The latter feature is clearly evidenced by the Ramachandran plots displayed in Figure 9a, which reflect not only the low mobility of the whole mutated loops but also the remarkable conformational similarity among the four Ac$_5$c residues contained in the different subunits of the self-assembled G149Ac$_5$c. This result is particularly important since although self-assembled nanotubes using 1krr repeats present an intrinsic stability, it can be significantly increased by restricting the conformational freedom at a specific position within the most mobile loop.

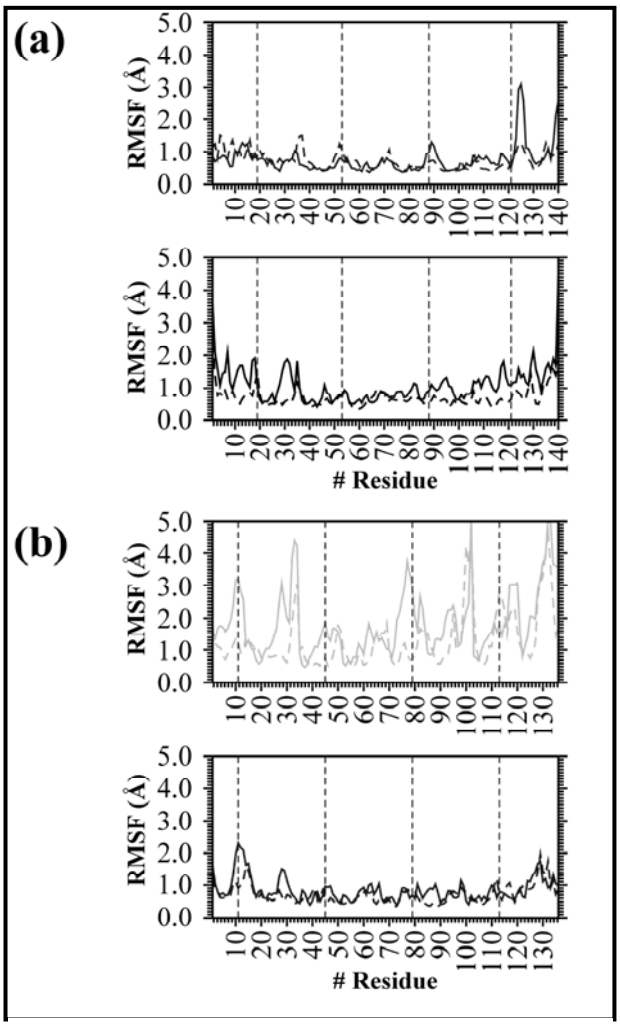


Figure 8. Comparison of RMSF values for the wild type and mutated sequences in both self-assembled (top plots) and polymerized (bottom plots) nanotubes, simulated at 298K. (a) Nanotubes based on 1krr: On the top, the self-assembled wild type (black line) and the G149Ac$_5$c (black-dashed line) building blocks; *on the bottom*, the polymerized wild type (black line) and G149Ac$_5$c (black-dashed line) building blocks. (b) Nanotubes based on 1hv9: On the top, the self-assembled wild type (grey line) and D306Ac$_5$c (grey-dashed line) building blocks; *on the bottom* the polymerized wild type (black line) and D306Ac$_5$c (black-dashed line) building blocks.

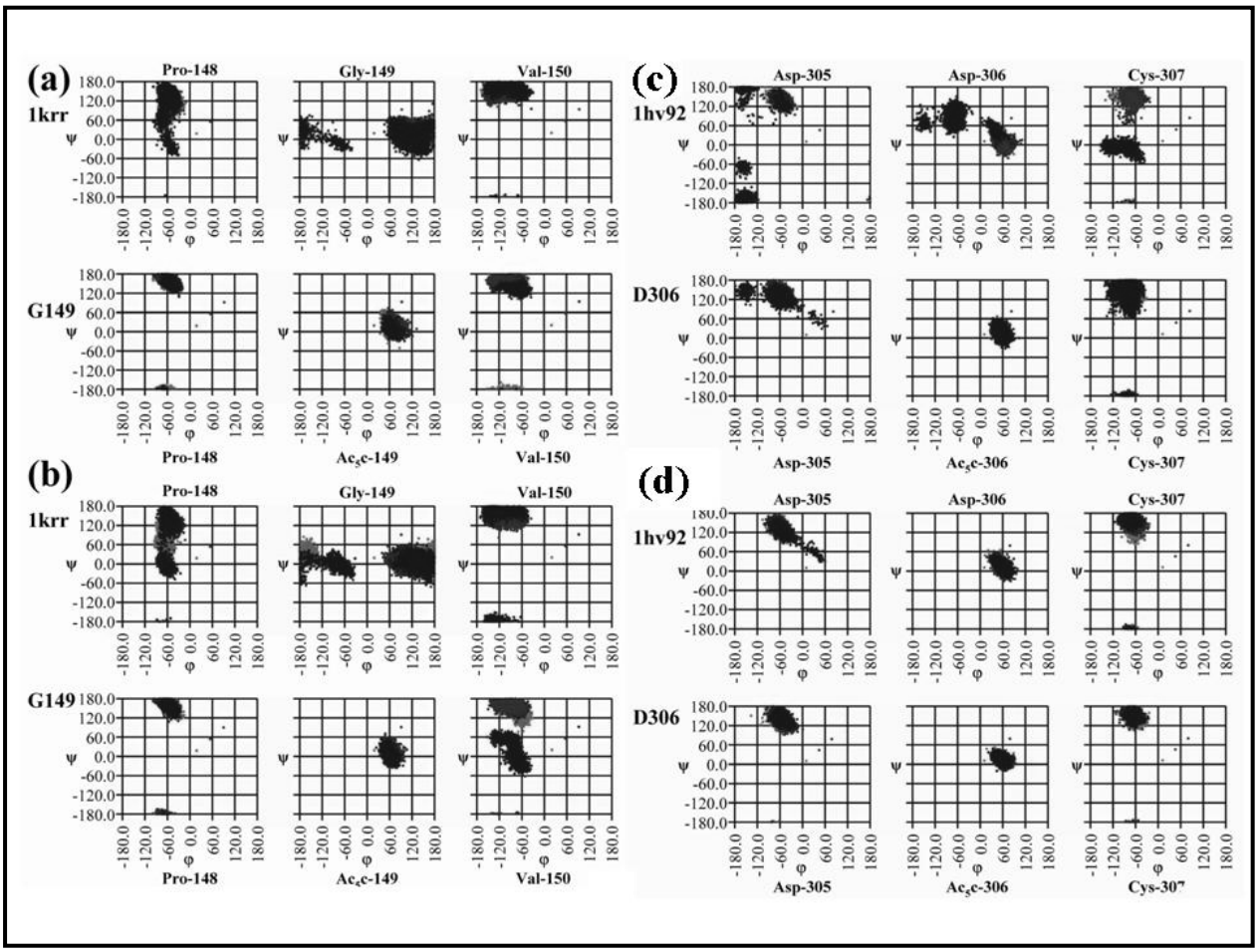


Figure 9. Comparison of the backbone torsion angle (φ, ψ) distributions of the targeted segments for the nanotubes constructed with the 1krr and 1hv9 building blocks simulated at 298 K. The Ramachandran plots correspond to the mutated position and its adjacent residues in the self-assembled (a) and polymerized (b) nanotubes derived from the 1krr wild type and G149Ac$_5$c building blocks and to the mutated position and its adjacent residues in the self-assembled (c) and polymerized (d) nanotubes derived from the 1hv9 wild type and D306Ac$_5$c building blocks. Different colors have been used for each building block.

On the other hand, Figure 7a suggests a small stabilization when this mutation is introduced in the polymerized tubes, even though the RMSFs are very low and similar to the G149Ac$_5$c and 1krr polymerized systems. Indeed, the Ramachandran plots displayed in Figure 9b indicates that the conformation is more regular for the mutated tube than for the wild type one. As mentioned above, introduction of Ac5c in position 149 reduces the conformational freedom of the segment enhancing the stability of the conformational motif. Hence, introduction of conformationally restrained residues at suitable positions appears a good strategy to stabilize not only the self-assembed nanoconstructs but also polymerized ones. Comparison of the self-assembled and polymerized nanotubes of G149Ac$_5$c reveals that both the RMSD and RMSF are very similar for the two systems. Furthermore, the conformations of the Ac$_5$c and its

neighboring residues do not depend on the self-assembled or polymerized nature of the nanotube, being almost identical in both cases. Thus overall these results allow us to conclude that G149Ac$_5$c is a potentially good mutant of 1krr.

Comparison of the self-assembled and polymerized nanostructures derived from D306Ac$_5$c with those constructed using the 1hv9 wild type building block is provided in Figures 7c and 7d, at two different temperatures. Once again, the increment of temperature does not lead to meaningful differences with respect to the simulations performed at 298 K. In order to reduce the background noise introduced by increasing thermal vibrations, ultra-structural details will be provided for the 298 K simulations.

Figure 8b shows the RMSF for all the 1hv9 based nanotubes studied at 298 K. As can be seen, substitution in the self-assembled tube does not produce any significant improvement. Thus, the RMSD of the D306Ac$_5$c and 1hv9 self-assembled tubes are very similar after 10 ns of MD simulation. Indeed, inspection to the RMSF values reveals that the mutation at the Asp-306 produces a significant fraying at the C-terminal region of all the interacting subunits, even though an improvement is clearly observed not only at the substituted position but also at the residues located at the N-terminal and central regions. The reduced conformational flexibility of the mutated loop is illustrated in Figure 9c, which compares the Ramachandran plots of residues 305, 306 and 307 contained in the four self-assembled subunits. These results indicate that although electrostatic repulsions were removed in the mutated building block, they appear again when the Asp-305 residues of the stacked D306Ac$_5$c units interact.

Finally, inspection of the RMSD (Figure 7c), RSMF (Figure 8b) and Ramachandran plots of the residues contained in the mutated loop (Figure 9d) of the D306Ac$_5$c and 1hv9 polymerized tubes shows that mutation by Ac$_5$c does not produce any relevant variation in the stability of the nanoconstruct. This is because the stability of the polymerized nanostructure constructed using the 1hv9 wild type sequence is very significant. Thus, the unfavorable effects associated with both the conformational flexibility of the loops and the electrostatic repulsions between neighboring Asp are minimized when the building blocks are covalently linked, the influence of the mutation by itself being almost negligible.

3.2.5. *Conclusions*

The present simulations confirm the intrinsic stability of self-assembled 1krr, which should be attributed to the stabilizing nonbonded interactions of the complementary arrangement of the building blocks. On the other hand, the high stability of the polymerized nanotubes constructed by linking the building blocks through covalent bonds indicates that the main structural principles of polymer chemistry and physics are valid for repeating units made of relatively large fragments of proteins. Thus, polymerization should be considered as a good alternative to self-assembling for the construction of stable nanotubes, especially for 1hv9 building block.

Analysis of the mutated building blocks provides some conclusions with respect to the substitution by the synthetic residue. For the 1krr the substitution of Gly-149 by Ac_5c has reduced significantly the conformational mobility not only at the mutated position but also of the adjacent positions. On the other hand, for 1hv9 the substitution of Asp-306 by Ac_5c induced a remarkable structural stability, which has been attributed to the elimination of unfavorable electrostatic interactions.

The simulation of the self-assembled and polymerized nanotubes constructed using G149Ac_5c and D306Ac_5c building blocks allow us to derive the following conclusions. First, the G149Ac_5c is a promising mutant of 1krr as revealed by both the significant stabilization found for the self-assembled system and the results obtained for the polymerized tube. Second, the introduction of residues with restrained conformations at strategic positions within the more flexible regions of the β-helix is a good strategy to improve the stability of the nanotubes. Finally, the stabilization induced in a building block by the substitution of a given residue does not guarantee that the same effect will occur in the corresponding nanostructure.

3.2.6. *References*

[1] Ferrari M. *Nat. Rev. Cancer* **2005**, *5*, 161.

[2] Ferrari M. *Curr. Opin. Chem. Biol.* **2005**, *9*, 343.

[3] Alemán C., Zanuy D., Jiménez A.I., Cativiela C., Haspel N., Zheng J., Wolfson H, Nussinov R. *Phys. Biol.* **2006**, *3*, S54.

[4] Claussen R.C., Rabatic B.M., Stupp S.I. *J. Am. Chem. Soc.* **2003**, *125*, 12680.

[5] Percec V, Dulcey A.E., Balagurusamy V.S., Miura Y., Smidrkal J., Peterca M., Nummelin S., Edlund U., Hudson S.D., Heiney P.A., Hu D.A., Magonov S.N., Vinogradov S.A. *Nature* **2004**, *430*, 764.

[6] Rajagopal K., Schneider J.P. *Curr. Opin. Struct. Biol.* **2004**, *14*, 480.

[7] Valery C., Paternostre M., Robert B., Gulik-Krzywicki T., Narayanan T., Dedieu J.C., Keller G., Torres M.L., Cherif-Cheikh R., Calvo P., Artzner F. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10258.

[8] Rathore O., Sogah D.Y. *J. Am. Chem. Soc.* **2001**, *123*, 5231.

[9] Vauthey S., Santoso S., Gong H., Watson N., Zhang S. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5355.

[10] Yokoi H., Kinoshita T., Zhang S. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 8414.

[11] Tsai C.-J., Maizel J.V.Jr., Nussinov R. *Proc. Natl. Acad. Sc*i. *USA* **2000,** *97*, 12038.

[12] Tsai H.-H., Tsai C.-J., Ma B., Nussinov R. *Protein Sci***. 2004**, *13*, 2753.

[13] Liao S., Seeman N.C. *Science* 2004, *306*, 2072.

[14] Yan H., Park S.H., Finkelstein G., Reif J.H., LaBean T.H. *Science* **2003**, *301*, 1882.

[15] Chworos A., Severcan I., Koyfman A.Y., Weinkam P., Oroudjev E., Hansma H.G., Jaeger L. *Science* **2004**, *306*, 2068.

[16] Main E.R.G., Lowe A.R., Mochrie S.G.J., Jackson S.E., Regan L. *Curr. Opin. Struct. Biol.* **2005**, *15*, 464.

[17] Kajander T., Cortajarena A.L., Main E.R.G., Mochrie S.G.J., Regan L. *J. Am. Chem. Soc.* **2005**, *127*, 10188.

[18] Main E.R.G., Stott K., Jackson S.E., Regan L. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5721.

[19] Haspel N., Zanuy D., Alemán C., Wolfson H., Nussinov R. *Structure* **2006**, *14*, 1137.

[20] Zheng J., Zanuy D., Haspel N., Tsai C.-J., Alemán C., Nussinov R. *Biochemistry* **2007**, *46*, 1205.

[21] Zanuy D., Jiménez A.I., Cativiela C., Nussinov R., Alemán C. *J. Phys. Chem. B* **2007**, *111*, 3236.

[22] Haspel N., Zanuy D., Zheng J., Alemán C., Wolfson H., Nussinov R. *Biophys. J.* **2007**, *93*, 245.

[23] Zanuy D., Ma B., Nussinov R. *Biophys. J.* **2003**, *84*, 1884.

[24] Haspel N., Zanuy D., Ma B., Wolfson H., Nussinov R. *J. Mol. Biol.* **2005**, *345*, 1213.

[25] Zanuy D., Nussinov R. *J. Mol. Biol.* **2003**, *329*, 565.

[26] Zanuy D., Gunasekaran K., Lesk A.M., Nussinov R. *J. Mol. Biol.* **2006**, *358*, 330.

[27] Alemán C., Zanuy D., Casanovas J., Cativiela C., Nussinov R. *J. Phys. Chem. B*, **2006**, *110*, 21264.

[28] Cheng R.P., Gellman S.H., DeGrado W.F. *Chem. Rev.* **2001**, *101*, 3219.

[29] Xie J.M., Schultz P.G. *Curr. Opin. Chem. Biol.* **2005**, *9*, 548.

[30] Pye V.E., Tingey A.P., Robson R.L., Moody P.C.E. *J. Biol. Chem.* **2004**, *279*, 40729.

[31] Govaerts C., Wille H., Prusiner S.B., Cohen F.E. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 8342.

[32] Kreisberg J.F., Betts S.D., King J. *Protein Sci.* **2000**, *9*, 2338.

[33] Adzhubei A.A., Sternberg M.J.E. *J. Mol. Biol.* **1993**, *229*, 472.

[34] Kostrewa D., D'Arcy A., Takacs B., Kamber M. *J. Mol. Biol.* **2001**, *305*, 279.

[35] Sugantino M., Roderick S.L. *Biochemistry* **2002**, *41*, 2209.

[36] Iverson T.M., Alber B.E., Kisker C., Ferry J.G., Rees D.C. *Biochemistry* **2000**, *39*, 9222.

[37] Jenkins J., Mayans O., Pickersgill R. *J. Struct. Biol.* **1998**, *122*, 236.

[38] Toniolo C., Crisma M., Formaggio F., Peggion C. *Biopolymers (Pept. Sci.)* **2001**, *60*, 396.

[39] Benedetti E. *Biopolymers (Pept. Sci.)* **1996**, *40*, 3.

[40] Phillips J.C., Braun R., Wang W., Gumbart J., Tajkhorshid E., Villa E., Chipot C., Skeel R.D., Kale L., Schulten K. *J. Comput. Chem.* **2005**, 26, 1781.

[41] Jorgensen W.L., Chandrasekhar J., Madura J.D., Impey R.W., Klein M.L. *J. Chem. Phys.* **1983**, *79*, 926.

[42] Wang J., Cieplak P., Kollman,P.A. *J. Comput. Chem.* **2000**, 21, 1049.

[43] Cornell W.D., Cieplak P., Bayly C.I., Gould I.R., Merz K.M., Ferguson D.M., Spellmeyer D.C., Fox T., Caldwell J.W., Kollman P.A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

[44] Darden T., York D., Pedersen L. *J. Chem. Phys.* **1993**, *98*, 10089.

[45] Ryckaert J.P., Ciccotti G., Berendsen H.J.C. *J. Comput. Phys.* **1977**, *23*, 327.

[46] Berendsen H.J.C., Postma J.P.M., van Gunsteren W.F., DiNola A., Haak J.R. *J. Chem. Phys.* **1984**, *81*, 3684.

[47] a) Mac Arthur M.W., Thornton J. M. *J. Mol. Biol.* **1991**, *218*, 397. b) Gibbs A. C., Bjorndahl T. C., Hodges R. S., Wishart D. S. *J. Am. Chem. Soc.* **2002**, *124*, 1203. c) Richardson J. S., Richardson D. C. "Principles and patterns of protein conformation", in: Prediction of Protein Structure and the Principles of Protein Conformation, Fasman, G.D. (Ed.), Plenum, New York, 1989, pp. 1-98.

3.3. Application of 1-aminocyclohexane carboxylic acid to protein nanostructure computer design*

3.3.1. *Introduction*

Synthetic α-amino acids 1-aminocycloalkane-1-carboxylic acids (Ac_nc, where *n* indicates the size of the ring) are the result of $C^{\alpha}\leftrightarrow C^{\alpha}$ cyclization whereby dialkylated glycine residues with cyclic side chains are formed. The conformational properties and electronic characteristics of *N*-acetyl-*N'*-methylamide derivatives of the cyclopropane[1] (Ac_3c), cyclobutane[2] (Ac_4c) and cyclopentane[3] (Ac_5c) analogues were fully characterized using ab initio and Density Functional Theory (DFT) quantum chemical calculations. The conformational flexibility of the backbone increases with the size of the ring. In the Ac_5c dipeptide each accessible backbone conformation is compatible with different arrangements of the cyclopentane ring.[3] The pseudorotational cyclopentane puckering strongly depends on the peptide backbone structure, that is, the arrangement of the cycle is determined by interactions between the side chain and the backbone.

When re-engineering protein modules via targeted replacements with synthetic amino acids,[4] the stability of self-assembled β-helical-based constructs, which should be considered as nanotubes or nanofibers, increases when the mobility of the loop regions is reduced by incorporating conformationally restricted amino acids.[5,6] Previously it was examined the stability of β-helical nanoconstructs following introduction of Ac_3c and 1-amino-2,2-diphenyl-cyclopropanecarboxylic acid (c_3Dip), an Ac_3c analogue bearing two phenyl rings at geminal positions, into flexible regions of the β-helical building blocks. Position-specific mutations indicated that when Ac_3c is introduced in loop regions it is able to enhance the stability of the nanoconstucts due to its strong tendency to adopt a turn structure, while c_3Dip is unsuccessful due to the steric effects induced by the bulky side chain.[5,6]

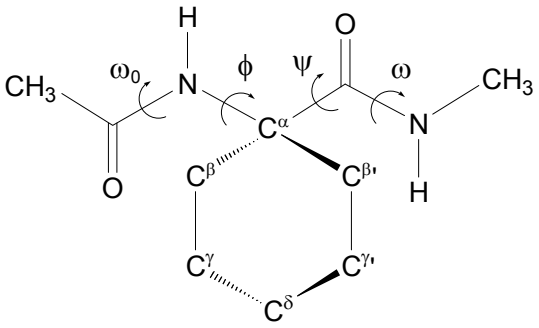
This chapter is devoted to a complete study about the stability of β-helical-based nanostructures when 1-aminocyclohexane-1-carboxylic acid (Ac_6c) is introduced in the loop region proteins. Early studies on small linear Ac_6c-containing peptides indicated that this amino acid tends to be involved in β-turns suggesting that this synthetic amino

* Results presented in this chapter are published in *J. Chem. Inf. Model.* **2008**, *48*, 333

acid is a potential candidate to increase the stability of nanoconstructs. Crystal state structural analyses indicated that this residue can occupy either the corner position (*i+1*, *i+2*) of type I β-bend or the *i+2* position of type II β-bends,[7-10] that is, the backbone conformation adopted by the Ac_6c is located in the helical region (φ,ψ≈ ±55º,±30º). Similarly, ^{1}H-NMR and FTIR studies in solution indicated that small oligomers of Ac_6c tend to form (incipient) helical secondary structures.[11] In cyclic peptides the Ac_6c residue adopts conformations in the helical region leading to β-bends.[12] On the other hand, Ac_6c was recently introduced in position 8 of bradykinin and positions 6, 7 and 8 of its B_2 receptor antagonist to reduce the flexibility of the peptides, thus forcing the peptide backbone and side chains to adopt specific orientations.[13] Interestingly, the Ac_6c substitution resulted in an increase in B_2 antagonistic activity, which offered new possibilities for designing new potent and selective B_2 blockers. Biological and pharmaceutical properties of several Ac_6c-containing peptides have been evaluated, some showing highly potent antidiuretic activity.[14] However, in spite of all these studies, no accurate description about the intrinsic conformational preferences of Ac_6c has been previously reported.

The aim of this work is to report the overall of the results of a four-steps project that involve: (i) the characterization of the conformational properties of Ac_6c; (ii) the influence of the environment in such properties; (iii) the development of force-field parameters for this synthetic amino acid; and finally (iv) a complete study about how the incorporation of Ac_6c within the more flexible regions of the β-helical building blocks affects to the stability of self-assembled nanoconstructs. The chapter has been organized as follows. First, the computational procedures used to examine the intrinsic conformational properties of Ac_6c and the stability of protein nanostructures that incorporate this constrained amino acid at the more flexible regions of the protein building blocks are presented. The minimum energy conformations and the Ramachandran maps of the *N*-acetyl-*N'*-methylamide derivative of Ac_6c (Scheme 1), hereafter denoted Ac-Ac_6c-NHMe, calculated in the gas-phase and solution (chloroform, methanol and water solvents) are compared and discussed. After this, a set of force-field parameters has been developed for Ac_6c to allow molecular dynamics simulations (MD). The suitability of such parameters has been checked by comparing the conformational preferences predicted for Ac-Ac_6c-NHMe by DFT calculations and MD simulations. Next, it is examined if Ac_6c is a good candidate to increase the

stability of both the β-helical protein building blocks and the nanoconstructs formed by self-assembling of such repeats. For this purpose the considered building blocks are the left-handed β-helices formed by residues 131-165 of *E. coli* galactoside acetyltransferase (PDB code 1krr, chain A) and by residues 296-329 of N-acetylglucosamine 1-phosphate uridyltransferase GlmU, C-terminal domain from *E. coli* (PDB code 1hv9) that, previously, were found to form the most and least stable protein nanotubes, respectively.[15] Finally, to further address the potential use of conformationally restricted amino acids to stabilize natural building block foldamers, the stabilizing effect produced by Ac_3c and Ac_6c substitutions are compared.

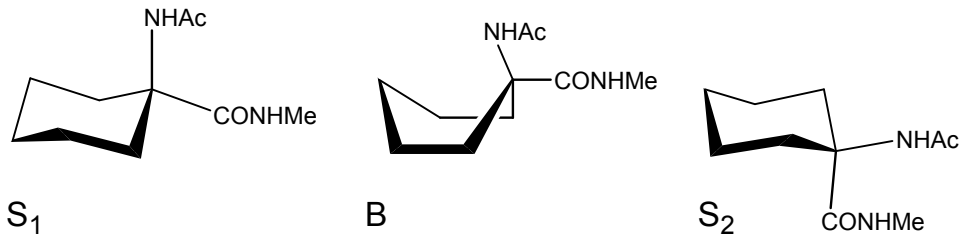


Scheme 1

3.3.2. *Methods*

Quantum chemical calculations were carried out using the Gaussian 03 program.[16] All the energy minima of Ac-Ac_6c-NHMe were localized using a systematic conformational analysis strategy. Specifically, because each flexible backbone dihedral angle is expected to have three minima, *i.e. gauche*$^+$ (60º), *trans* (180º) and *gauche*$^-$ (–60º), the number of minima that may be anticipated for the potential energy hypersurface E= E(φ,ψ) of Ac-Ac_6c-NHMe is 3^2= 9. Furthermore, due to the cyclic nature of the side chain, two chair and one boat (labeled as S_1, S_2 and B, respectively, in Scheme 2) arrangements of the cyclohexane side chain have been considered for each backbone minimum energy conformation. Another boat conformation can be obtained by exchanging the positions of the NHAc and CONHMe moieties. However, this alternative boat arrangement is expected to be of the highest energy and, therefore, it has not been included in the conformational analysis. Accordingly, such 9×3= 27 structures were considered as starting points for complete geometry optimization.

This systematic strategy permitted satisfactory exploration of the potential energy hypersurfaces not only of small dipeptides[1-3] but also of flexible organic molecules.[17] Frequency analyses were carried out to verify the nature of the minimum state of all the stationary points obtained and to calculate the zero-point vibrational energies (ZPVE) and both thermal and entropic corrections. These statistical terms were used to compute the conformational Gibbs free energies in the gas-phase at 298 K (ΔG^{gp}). All the calculations were carried out using the B3LYP[18,19] functional and the 6-31+G(d,p) basis set.[20]



Scheme 2

A complete exploration of the potential energy surface E(φ,ψ) was performed by mapping the Ramachandran plot of Ac-Ac$_6$c-NHMe. Calculations were performed on a grid of points on the (φ,ψ) space at 30º intervals, ω_1 and ω_2 being initially positioned at 180º in all cases. Due to the achiral nature of the molecule, only half of the map was computed since E(φ,ψ)=E(−φ,−ψ). At each point of the grid the geometry was optimized in the gas-phase by keeping the dihedral angles φ and ψ constrained during the minimization process.

To calculate the Ramachandran map of Ac-Ac$_6$c-NHMe in solution and to obtain estimation of the solvation effects on the relative stability of the different minima, single point calculations were also conducted on the B3LYP/6-31+G(d,p) optimized structures using a Self-Consistent Reaction Field (SCRF) model. SCRF methods treat the solute at the quantum mechanical level, while the solvent is represented as a dielectric continuum. Specifically, the Polarizable Continuum Model (PCM) developed by Tomasi and co-workers was chosen to describe the bulk solvent.[21,22] The PCM method involves the generation of a solvent cavity from spheres centered at each atom in the molecule and the calculation of virtual point charges on the cavity surface representing the polarization of the solvent. The magnitude of these charges is

proportional to the derivative of the solute electrostatic potential at each point calculated from the molecular wavefunction. The point charges may, then, be included in the one-electron Hamiltonian, thus inducing polarization of the solute. An iterative calculation is carried out until the wavefunction and the surface charges are self-consistent. PCM calculations were also performed in the framework of the DFT B3LYP/6-31+G(d,p) level using the standard protocol and considering the dielectric constants of chloroform (ε= 4.9), methanol (ε= 32.63) and water (ε= 78.4). The conformational free energies in solutions ($\Delta G^{\#sol\#}$, where #sol# refer to the solvent) were computed using the classical thermodynamics scheme: the free energies of solvation provided by the PCM model were added to the ΔG^{gp}.

Molecular Dynamics (MD) simulations were performed using the NAMD program.[23] Each simulated system was placed in the center of a simulation box filled with explicit water molecules, which were represented using the TIP3P model.[24] Positively charged sodium atoms were added to the simulation box in the required amount to reach electric neutrality (all considered building blocks had negative net charge at neutral pH). All atoms of both building blocks and nanotubes were considered explicitly.

The energy was calculated using the AMBER force-field,[25,26] with the required parameters taken from the AMBER libraries for all the residues with the exception of Ac_6c. Force-field parameters for Ac_6c were developed using the same strategy previously employed for Ac_5c.[3] As was be demonstrated, the parameters provided by such strategy are fully consistent with AMBER ones. Atom pair distance cutoffs were applied at 14.0 Å to compute the van der Waals interactions. The electrostatic interactions were computed using the non-truncated electrostatic potential with Ewald Summations.[27] The real space term was determined by the van der Waals cut off (14.0 Å), while the reciprocal term was estimated by interpolation of the effective charge into a charges mesh with a grid thickness 5 points per volume unit, *i.e.* particle-mesh Ewald (PME) method.[27] Bond lengths involving hydrogen atoms were constrained using the *SHAKE* algorithm,[28] with a numerical integration step of 2 fs. The initial edge of the cubic box for simulations of wild type and mutated building blocks was 56.6 Å, the total number of particles considered explicitly ranging from 17465 to 17487. For simulations of the self-assembled nanotube, the initial dimensions of the tetragonal box were (72.0 x 72.0 x 112.0) $Å^3$ and the number of explicit particles ranged from 56084 to 56196.

Before the production series, the thermodynamic variables of the system were equilibrated. The energy of each system was initially minimized to relax conformational and structural tensions using the conjugate gradient method for $5 \cdot 10^3$ steps. Next, different consecutive rounds of short MD runs were performed in order to equilibrate the density, temperature and pressure. First, solvent and charged sodium atoms were thermally relaxed by three consecutives runs, while the protein parts were kept frozen: 0.5 ns of NVT-MD at 500 K were used to homogeneously distribute the solvent and ions in the box. Second, 0.5 ns of isothermal and 0.5 ns isobaric relaxation were run. Finally, all the atoms of the system were submitted to 0.15 ns of steady heating until the target temperature was reached (298 K), 0.25 ns of *NVT*-MD at 298 K (thermal equilibration) followed by 0.5 ns of density relaxation (*NPT*-MD). Both temperature and pressure were controlled by the weak coupling method, the Berendsen thermo-barostat,[29] using a time constant for heat bath coupling and a pressure relaxation time of 1 ps. The end of the density relaxation simulation was the starting point of the molecular simulations presented in this work. All the simulations were performed at 298 K and constant pressure of 1 atm. The coordinates of all the production runs, which were 10 ns long, were saved every 500 steps (1 ps intervals) for subsequent analysis.

3.3.3. *Results and discussion*

Intrinsic Conformational Properties of Ac_6c

Geometry optimizations in the gas-phase at the B3LYP/6-31+G(d,p) level of the 27 structures considered as starting geometries (see Methods), led to 11 different minimum energy structures for Ac-Ac_6c-NHMe. Table 1 summarizes the more relevant structural data of such minima, which are depicted in Figure 1. Within this context, it should be mentioned that the criterion used to accept the formation of an intramolecular hydrogen bond was: d(H···O) < 2.5 Å.

The two minima more stable in the gas-phase, I and II, correspond to a C_7 (γ-turn), in which the backbone dihedral angles φ,ψ define a seven-membered intramolecular hydrogen bonded ring with parameters d(H···O)= 1.907 (I) / 1.938 (II) Å and ∠N–H···O= 151.7º (I) / 147.9º (II). In these minima, which differ in 0.2 kcal/mol, the cyclohexane ring adopts a S_1 and S_2 conformation, respectively. Thus, in this case the two chair arrangements show similar stabilities. However, minimum VIII, in which the

C_7 backbone conformation and the B arrangement for the cyclohexane ring are combined, is disfavored by 6.0 kcal/mol with respect to I.

Table 1. Backbone Torsion Angles,[a] Backbone Conformation,[b] Side Chain Conformation[c] and Relative Energy (ΔE)[d] for the Conformational Energy Minima of Ac-Ac$_6$c-NHMe at the B3LYP/6-31+G(d,p) Level of Theory.

#	ω_0	φ	ψ	ω	Backbone	Side Chain	ΔE
I	-175.4	-73.5	55.4	179.2	C_7	S_1	0.0[e]
II	-177.3	-73.2	69.1	-175.5	C_7	S_2	0.2
III	169.8	70.4	19.5	-177.7	α	S_1	2.5
IV	-170.2	59.6	-122.2	174.6	P_{II}	S_1	2.6
V	-178.4	178.8	174.1	179.5	C_5	S_2	3.4
VI	168.1	63.4	35.1	-176.8	α	S_2	3.8
VII	-171.1	-177.8	50.2	175.3	P_{II}	S_2	5.6
VIII	174.5	73.7	-51.7	-177.8	C_7	B	6.0
IX	179.6	-179.8	-179.6	180.0	C_5	S_1	6.2
X	177.9	178.1	-169.5	-177.7	C_5	B	7.5
XI	-167.7	59.6	-134.6	177.2	P_{II}	B	9.0

[a] In degrees. [b] The backbone conformation is defined by the dihedral angles φ and ψ (see text). [c] S_1, S_2 and B refer to chair-1, chair-2 and boat conformations, respectively, of the cyclohexane ring (see Scheme 2). [d] In kcal/mol. [e] E= -651.954877 a.u.

Minimum III corresponds to an α-helical conformation with the cyclohexane arranged in S_1. No intramolecular hydrogen bonding interaction is present in this structure, which is 2.5 kcal/mol less stable than the global minimum I. The analogous α-helical minimum but with the cyclohexane arranged in S_2, VI, is destabilized by 1.3 kcal/mol with respect to III. In this case, no B arrangement was compatible with the α-helical backbone conformation. Comparison of these results with those obtained for I and II indicates that the $S_1 \leftrightarrow S_2$ equilibrium is significantly influenced by the backbone

conformation. This influence becomes more apparent in minima IV, VII and XI, in which the P_{II} backbone arrangement is compatible with the S_1, S_2 and B conformations, respectively. The chair conformation with the CONHMe moiety in an equatorial position (S_1) is more stable than the structure that adopts a S_2 conformation by 3.0 kcal/mol, the B arrangement being destabilized by 6.4 kcal/mol.

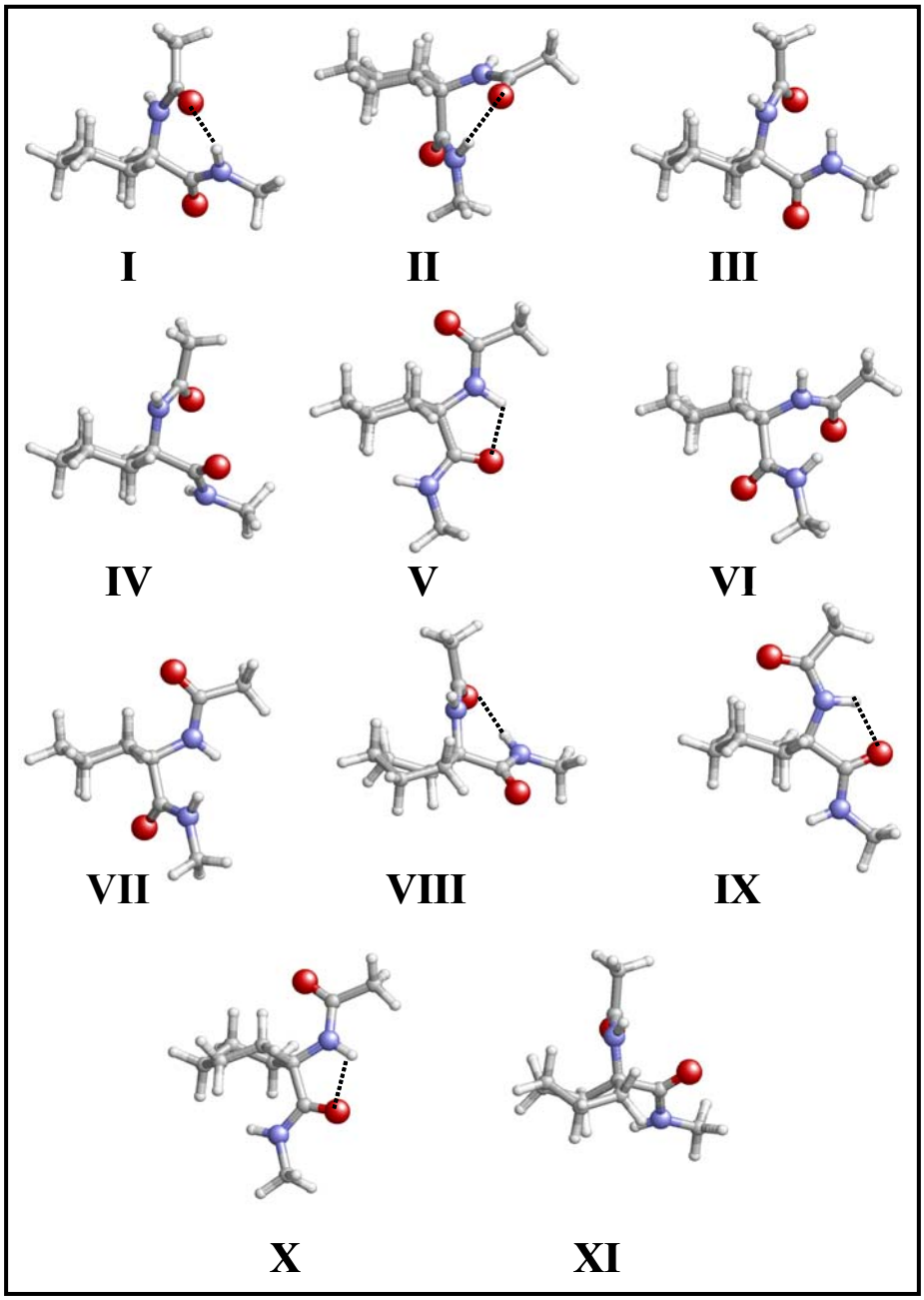


Figure 1. Molecular structure of the minimum energy conformations characterized for the Ac-Ac_6c-NHMe dipeptide at the B3LYP/6-31+G(d,p) level. Intramolecular hydrogen bonds are indicated by dashed lines. The structural characteristics of each minimum are described in Table 1.

Finally, the least favored backbone conformation corresponds to the C_5, which is displayed by minima V, IX and X. In these structures the backbone dihedral angles φ,ψ are arranged in trans, defining a five-membered intramolecular hydrogen bonded ring with parameters *d*(H···O)= 1.974 (V) / 1.922 (IX) / 2.021 (X) Å and ∠N–H···O= 114.2º (V) / 117.9º (IX) / 112.5º (X). The V, IX and X minima, which show steric repulsions between the amide groups and the β-methylene hydrogen atoms, are destabilized by 3.4, 6.2 and 7.5 kcal/mol, respectively, with respect to the global minimum I. For the C_5 the S_2 chair conformation is the most stable arrangement of the cyclohexane ring. Table 2 lists the values of ΔG^{gp} for the 11 minima. As can be seen, the influence of the ZPVE, thermal and entropic corrections do not introduce significant changes in the stability of these structures, the maximum difference between ΔE and ΔG^{gp} being 1.3 kcal/mol (minimum V).

Table 2. Conformational Free Energies in the Gas-phase, Chloroform, Methanol and Aqueous Solutions for the Energy Minima of Ac-Ac$_6$c-NHMe determined at the B3LYP/6-31+G(d,p) Level of Theory.

#	ΔG^{gp}	ΔG^{CHCl3}	ΔG^{MeOH}	ΔG^{H2O}
I	0.0[a]	0.0	1.5	1.8
II	0.6	1.1	2.4	3.1
III	1.6	1.5	0.3	0.7
IV	1.8	1.5	1.7	2.1
V	2.1	1.8	5.2	3.4
VI	3.5	1.8	0.0	0.0
VII	4.5	3.3	2.4	0.0
VIII	5.9	4.2	3.2	3.8
IX	5.7	4.5	3.3	3.9
X	7.4	5.3	4.3	4.1
XI	8.0	5.9	4.4	4.4

[a] G= -651.717012 a.u.

Figure 2 represents the Ramachandran map calculated for Ac-Ac$_6$c-NHMe in the gas-phase. The map was obtained by considering the S_1 as the starting conformation for the cyclohexane ring of all the calculated structures on the (φ,ψ) space. Although this chair

conformation was frequently retained after geometry optimization, in some cases the S_1 transform into the S_2 or, even, the boat arrangements. As expected, the four backbone conformations listed in Table 1 correspond to low energy regions in the map. Comparison of the map calculated for Ac-Ac_6c-NHMe with those previously reported for Ac-Ac_4c-NHMe[2] and Ac-Ac_3c-NHMe[1] evidence that the conformational flexibility of Ac_nc-containing dipeptides increases with the size of the ring. Thus, although Ac_6c is a constrained amino acid, the number of low energy regions is significantly smaller for Ac_4c and Ac_3c.

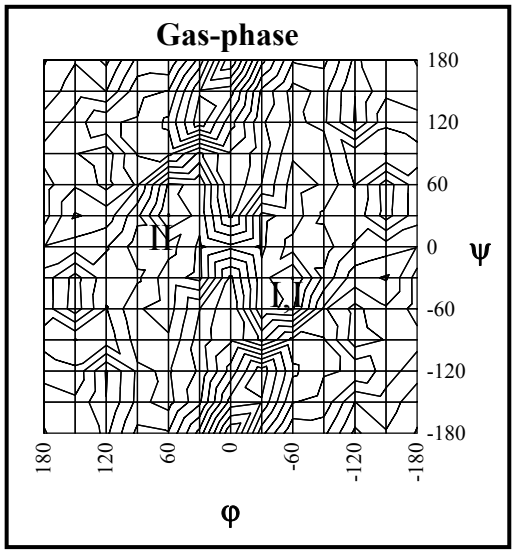


Figure 2. Ramachandran map of the Ac-Ac_6c-NHMe dipeptide in gas-phase computed at the B3LYP/6-31+G(d,p) level of theory. Energies (in kcal/mol) are relative to the lowest energy minimum. Contours are drawn every two kilocalories per mole. The position of the minima with lower free energies (see Table 2) is indicated.

The Ramachandran maps calculated in chloroform, methanol and water solutions are displayed in Figure 3. The topology of these maps is different from that obtained in the gas-phase indicating that the solvent plays a crucial role in the conformational preferences of Ac_6c. Specifically, Figure 3 reveals that the relative energy range that separates the most and least favored conformations is drastically reduced by the solvent. This is also indicated in Table 2, which compares the conformational free energies in solutions for the 11 minimum energy structures for Ac-Ac_6c-NHMe. Interestingly, although the relative free energy order of these structures is very similar in the gas-phase and in a chloroform solution, the free energy range for the first 6 minima is significantly narrower in the latter organic solvent, *i.e.* they are separated by 3.5 and 1.8

kcal/mol in the gas-phase and chloroform solution, respectively. The effect of the solvent is even more pronounced in methanol and water, in which a substantial change in the relative conformational free energy order is detected. Thus, minimum VI is the most favored in these polar environments. Overall, these results indicate that the conformational restrictions imposed by the cyclohexane ring are less severe in solution than in the gas-phase, the role of the solute-solvent interactions being crucial for the conformational preferences of Ac_6c in the former environment.

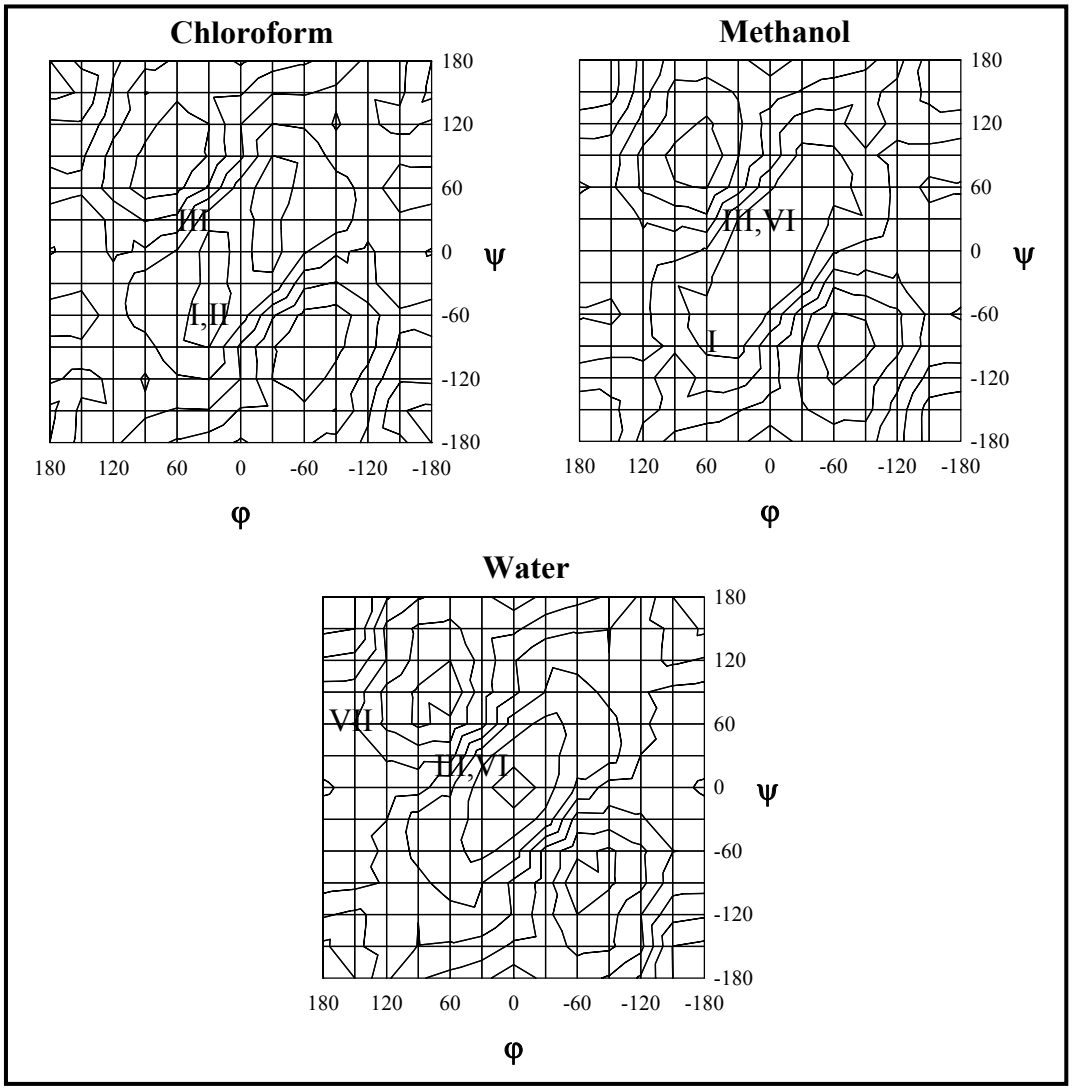


Figure 3. Ramachandran maps of the Ac-Ac_6c-NHMe dipeptide in chloroform, methanol and water solutions computed at the B3LYP/6-31+G(d,p) level of theory. Energies (in kcal/mol) are relative to the lowest energy minimum. Contours are drawn every two kilocalories per mole. The position of the minima with lower free energies (see Table 2) is indicated in each case.

Force-Field Parameterization

Electrostatic charges have been the only force-field parameters specifically developed for Ac_6c, the stretching, bending, torsional and van der Waals parameters being directly transferred from the Amber force-field. [25,26]

Atomic charges for all eleven minimum energy conformations listed in Table 1 were calculated by fitting the HF/6-31G(d) quantum mechanical and the Coulombic molecular electrostatic potentials (MEPs) to a large set of points placed outside the nuclear region. It should be noted that the electrostatic parameters derived at this level of theory are fully compatible with the current Amber force-field.[25,26] Electrostatic potential (ESP) fitting atomic centered charges for the Ac_6c residue were derived by weighting the charges calculated for the eleven minimum energy conformations according to the Boltzmann populations. The weights were given by the standard Boltzmann formula using the ΔG^{gp} values listed in Table 2, *i.e.* in practice only minima I and II contributed significantly. The resulting electrostatic parameters are displayed in Figure 4.

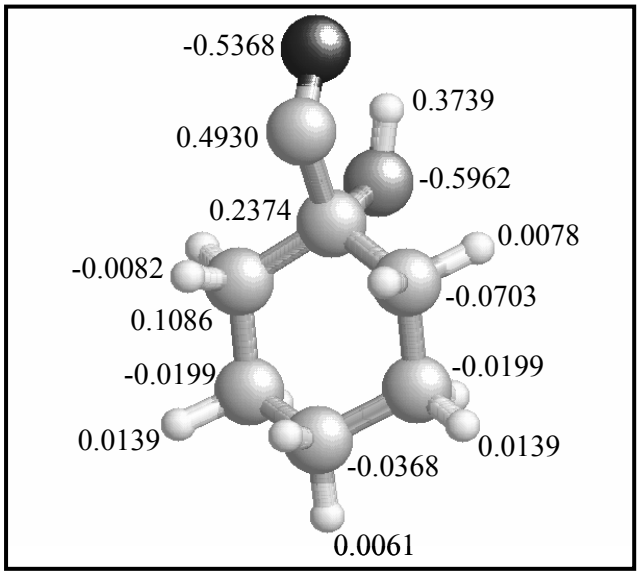


Figure 4. Electrostatic parameters determined for the Ac_6c residue.

An important test to check the reliability of force-field parameters developed for amino acids is the attempt to reproduce both the energetically accessible conformations of dipeptides and their flexibility.[3,25,26] In order to ascertain how the force-field parameters describe these properties Ac-Ac_6c-NHMe, MD simulations were performed at 298 K in the gas-phase. The three conformations of lower energy were used as starting point,

each trajectory being 8 ns long. Figure 5 represents the accumulated Ramachandran plot of the three trajectories for the Ac_6c dipeptide. It is worth noting the C_7 and α are the populated regions, in excellent agreement with the ΔG^{gp} values derived from DFT calculations (Table 2). On the other hand, as expected a chair conformation was systematically detected for the cyclohexane ring.

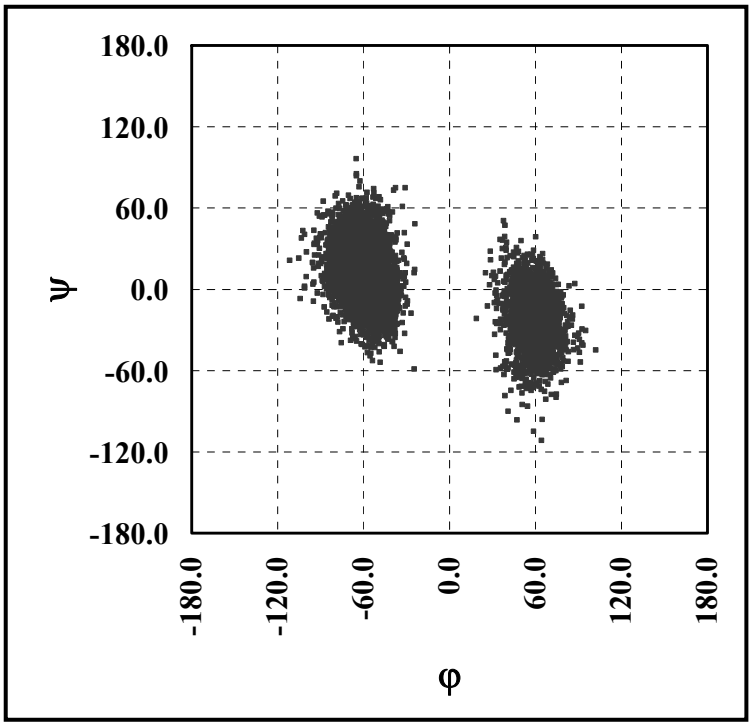


Figure 5. Accumulated Ramachandran plot for Ac-Ac6c-NHMe derived from MD trajectories in the gas-phase.

Building Block Mutants of 1krr and 1hv9

DFT calculations on Ac-Ac_6c-NHMe indicated that Ac_6c presents a high propensity to adopt folded conformations. This conformational characteristic makes this synthetic amino acid a potential candidate to reduce the conformational freedom of the β-helix building blocks if it is introduced in the most mobile regions, *i.e.* the folded loops, replacing natural amino acids. The strain energy associated with the cyclohexane ring in Ac_6c is significantly lower than that of other recently investigated constrained amino acids with cyclic side chains, Ac_3c and c_3Dip.[5,6] Accordingly, Ac_6c is expected to adapt its folded conformation within the targeted position more easily than the corresponding analogue with a cyclopropane ring. Two Ac_6c-single mutations were considered for the 1krr and 1hv9 β-helix building blocks.[15] It was found that a nanostructure constructed by four stacked replicas of the left-handed β-helix formed by residues 131-165 of 1krr

exhibited remarkable stability under different simulated conditions, including temperature increase and addition of ions.[15] On the other hand, the less stable model was obtained from four self-assembled copies of the β-helix formed by residues 296-329 of 1hv9. Thus, nanoconstructs formed by the latter repeat are good systems to test stabilization strategies. A description of the sequences of 1krr and 1hv9 building blocks is provided in Table 3.

Table 3. Sequences of the wild type 1krr and 1hv9 building blocks used to construct the nanotubes. For each fragment the mutated residues are highlighted with both bolding and italics.

PDB	*Protein name*	*Residues*	*Sequence*
1krr	Galactoside acetyltransferase from e. coli	131-165	PITIGNNVWIGSHVVINP ***G***VTIGDNSVIG***A***GSIVT
1hv9	N-acetylglucosamine 1-phosphate uridyltransferase GlmU, C-terminal domain from e. coli.	296-329	CVIKNSVIGD***D***CEISPY TVVEDANLA***A***ACTIGPF

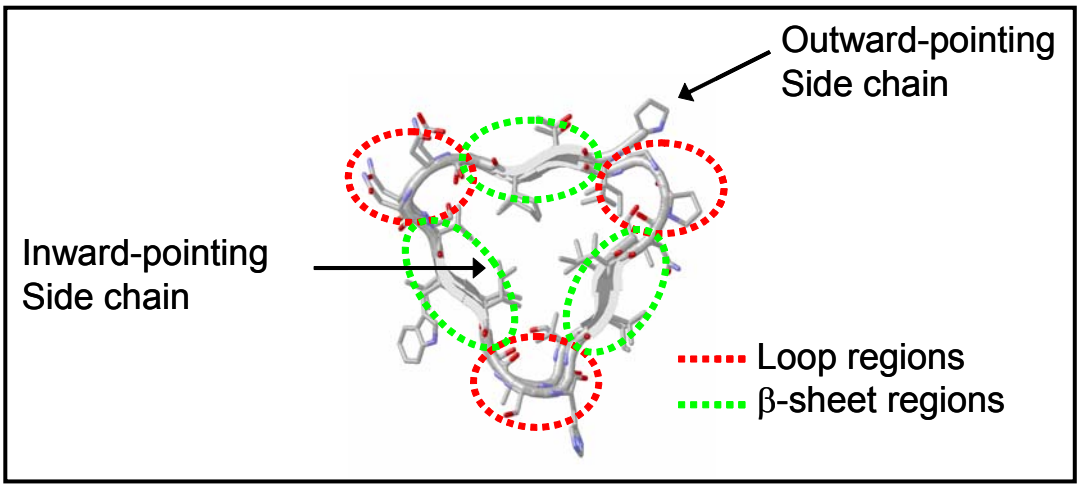


Figure 6. Schematic picture showing the loop and β-sheet regions of the β-helix as well as residues with outward- and inward-pointing side chains.

Two positions of 1krr and 1hv9 building blocks were selected as suitable candidates for substitution by Ac_6c according to the following criteria (Figure 6): i) the residues are

located in the loop regions of the β-helix, which display higher mobility than the β sheets and present a folded conformation similar to that preferred by Ac_6c; ii) the side chain of the residues is outward-pointing avoiding unfavorable steric interactions between the cyclohexane group of the Ac_6c substitution and the side chains of the inward-pointing residues. Targeting flexible disfavored loop regions should have direct impact on the structural stability of the nanotubes. For the 1krr building block the Gly-149 and Ala-160 residues were substituted by the Ac_6c one at a time (Table 3), the two corresponding mutants being denoted G149Ac_6c and A160Ac_6c, respectively. The single mutations performed on the 1hv9 building block at Asp-306 and Ala-322 were denoted D306Ac_6c and A322Ac_6c, respectively. The charged side chain of Asp-306 produces electrostatic repulsions with Asp-305 in the loop region of 1hv9.[15]

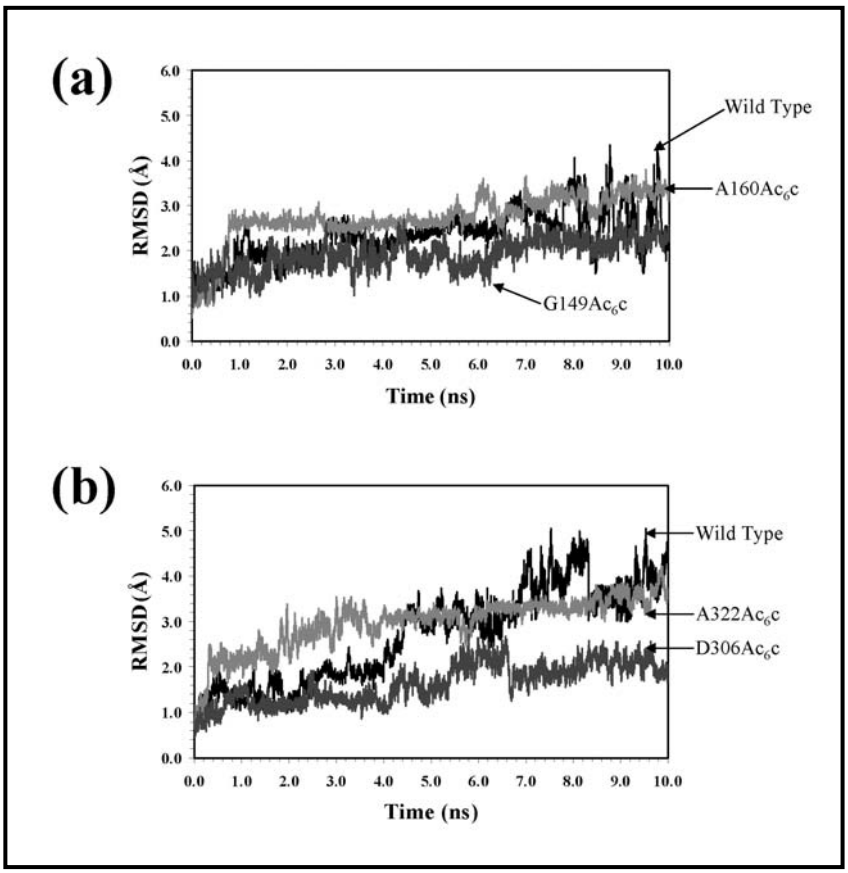


Figure 7. Comparison of the RMSD (in Å) for the wild type and mutated building blocks. (a) Building blocks based on 1krr: wild type, G149Ac_6c and A160Ac_6c. (b) Building blocks based on 1hv9: wild type, D306Ac_6c and A322Ac_6c.

Figure 7 compares the root mean square deviation (RMSD) of the wild type with the corresponding mutants, while Figure 8 displays for each case the initial structure with

that obtained after 10 ns of MD simulation. As can be seen, replacement of Gly-149 by Ac_6c produced an improvement in the RMSD of the 1krr building block. Such improvement is clearly observed in Figure 9, which shows the root mean square fluctuation (RMSF) of individual residues averaged over the whole simulation. Analysis of the RMSF reveals a considerable local improvement, the averaged RMSD of residue 149 decreasing by about 1.5 Å. Accordingly, substitution of a flexible Gly residue by an Ac_6c, which is constrained to adopt turn conformations, stabilizes the bend architecture of the loop by reducing its mobility. This feature is also evidenced in the Ramachandran plots (Figure 10). In contrast, neither local nor global improvement was observed for the A160Ac_6c mutant. The RMSD of residue 160 increased by 1.17 Å after mutation (Figure 9). A detailed examination of the snapshot recorded for this mutant at the end of the simulation (Figure 8) reveals a significant distortion in the turn at positions 159-161, indicating that the rigidity introduced by the synthetic residue in the loop causes the disruption of the β–helix motif. This feature suggests that in order to retain the global organization of the building block it is important to preserve some flexibility in this loop.

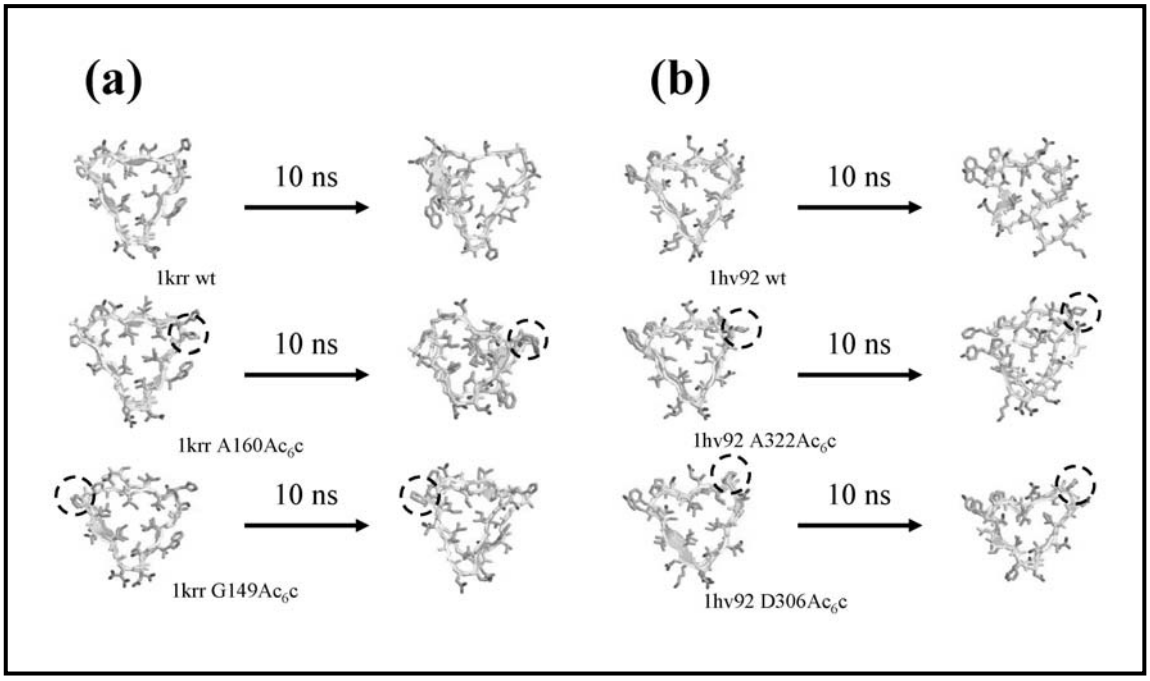


Figure 8. Structure of the wild type and mutated building blocks of 1krr (a) and 1hv9 (b) after 10 ns of MD simulation. Circles indicate the position of Ac_6c in each mutated building block. The hydrogen atoms have been omitted for clarity and the backbone has been represented by solid shapes (arrows indicate sheet conformation).

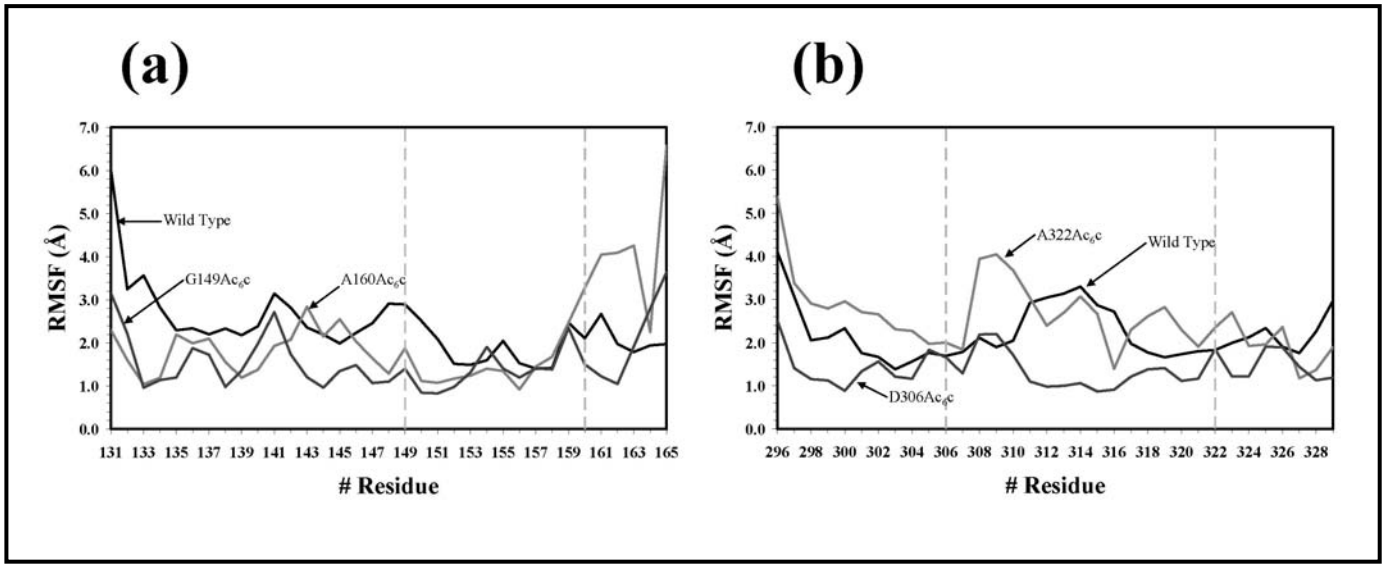


Figure 9. Comparison of the RMSF for the wild type and mutated building blocks. (a) Building blocks based on 1krr: wild type, G149Ac$_6$c and A160Ac$_6$c. (b) Building blocks based on 1hv9: wild type, D306Ac$_6$c and A322Ac$_6$c. Dashed lines indicate the position of the substitutions.

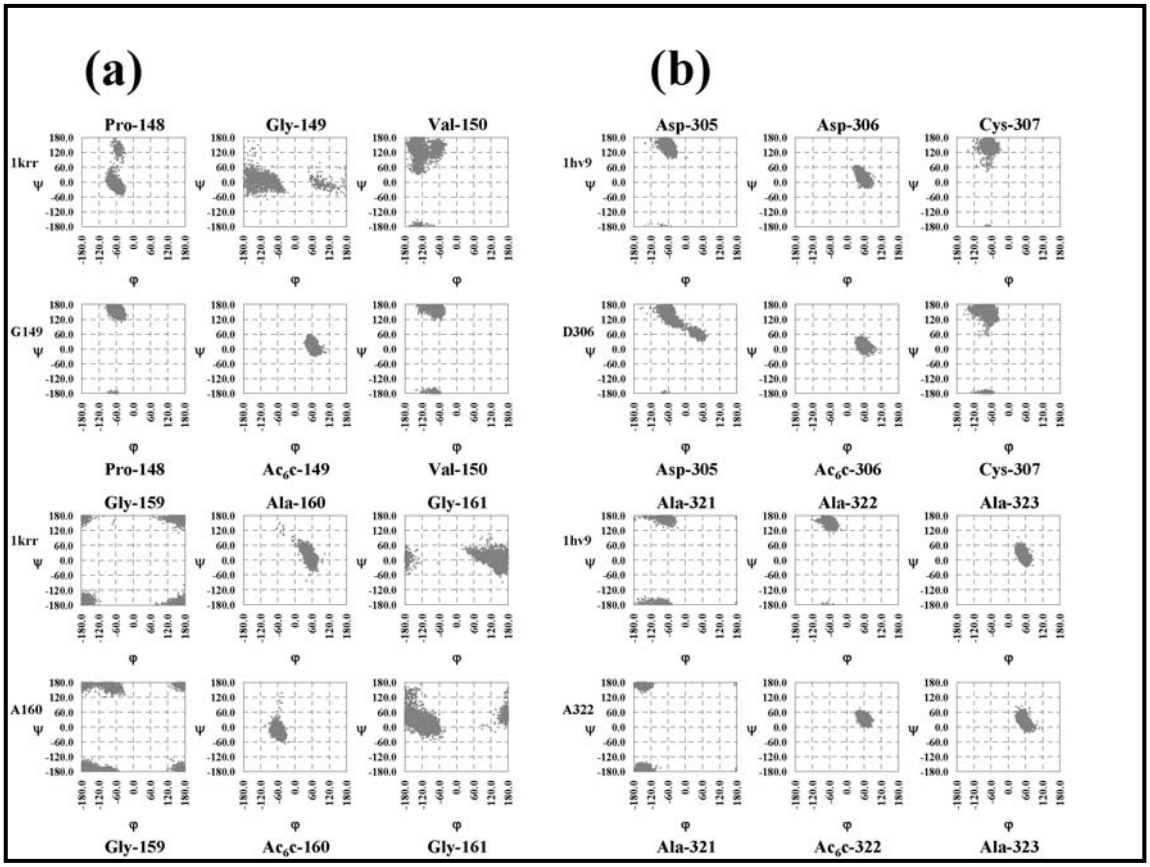


Figure 10. Comparison of the backbone torsion angle (φ,ψ) distributions for the building blocks. The Ramachandran plots correspond to the targeted position and its adjacent residues in (a) 1krr wild type, G149Ac6c and A160Ac6c mutants and (b) 1hv9 wild type, D306Ac6c and A322Ac6c mutants.

For the wild type building block of 1hv9 and its mutants, results indicate that the D306Ac$_6$c mutant shows a notable structural stability. In this case, the RMSF is considerably smaller at many positions along the chain. The initial β-helix conformation is retained without apparent distortions after 10 ns of MD simulation. Inspection of the Ramachandran plots of the substituted position and the adjacent residues (Figure 10) reveals a significant resemblance between the wild type and the D306Ac$_6$c mutant. Analysis of the RMSF indicates that even though the substitution at position Asp-306 by Ac$_6$c eliminates the electrostatic repulsion with Asp-305, the backbone constraints associated with the Ac$_6$c residue have a significantly smaller effect in the fluctuations at the loop than at the other positions along the chain. As a consequence, a substantial global stabilization of the building block is obtained. By contrast, substitution at Ala-322 does not provide neither local nor global improvement. The structural distortion is significant for both the wild type and the A322Ac$_6$c mutant, with the β-helix disrupted in both cases.

Analysis of the conformation adopted by the cyclic side chain of Ac$_6$c through the MD simulations of 1krr and 1hv9 mutants reveals a certain degree of mobility. Although the S_1 is clearly the predominant conformation in all cases (population of 67%), relatively high populations of both S_2 and B arrangements are detected (21% and 12%, respectively). However, the apparition of the latter arrangements should be attributed to relaxation effects associated to the dynamics of the building block, rather than to a thermodynamic equilibrium. This feature indicates that the flexibility of the cyclohexane ring plays a crucial role in the adaptation of the Ac$_6$c residue to the bend organization of the loop. Thus, although the intrinsic stability of wild type building blocks increases by restricting the conformational freedom at a specific position within the most mobile loop, the synthetic residue introduced for this purpose should retain some flexibility (for example in the side chain, as Ac$_6$c) to avoid unfavorable strain effects that disturb this architecture.

Assembled Mutants of 1krr and 1hv9

G149Ac$_6$c, A160Ac$_6$c, D306Ac$_6$c and A322Ac$_6$c were used as building blocks for nanoassemblies, which were constructed by stacking four copies of each repeat one atop the other, with no covalent linkage between them. Figure 11 compares the evolution of the RMSD calculated for nanotubes of the mutated building blocks with those obtained

using wild-type repeats, while Figure 12 depicts the RMSF of all the investigated systems. Figure 13 shows structure of the nanoconstructs after 10 ns of MD.

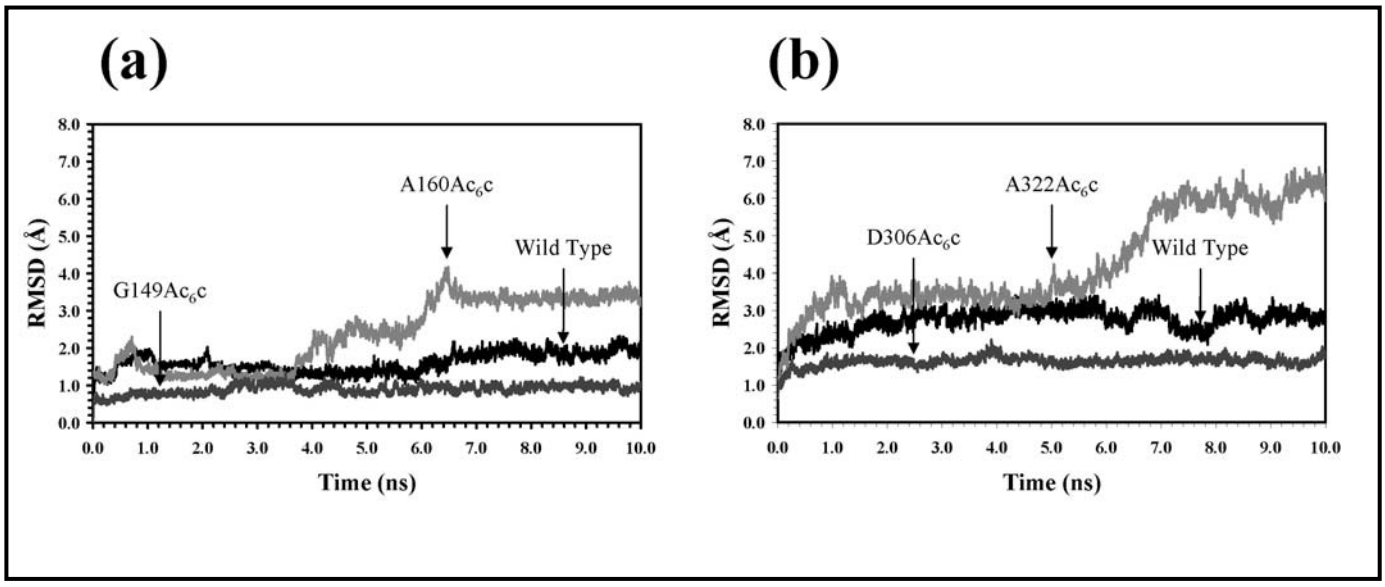


Figure 11. Evolution of the backbone RMSD (in Å) of the simulated nanotube models based on the self-assembly of both wild type and mutated building blocks. (a) Building blocks based on 1krr: wild type, G149Ac$_6$c and A160Ac$_6$c. (b) Building blocks based on 1hv9: wild type, D306Ac$_6$c and A322Ac$_6$c.

As expected from the results of the single building block, the RMSD of the G149Ac$_6$c self-assembled system is significantly smaller than that derived from 1krr wild type repeat. At 298 K, the backbone RMSD of the self-assembled G149Ac$_6$c remains below 1 Å for the entire simulation indicating the stability of this organization is remarkably high. This behavior demonstrates that the substitution of Gly-149 by Ac$_6$c not only reduces the flexibility of the building block but also enhances its ability to retain the assembled nanostructure constructed using the mutated subunits. This feature is clearly indicated by the Ramachandran plots (Figure 14), which reflect not only the low mobility of the whole mutated loops but also the remarkable conformational similarity among the four Ac$_6$c residues contained in the different subunits of the self-assembled G149Ac$_6$c. On the other hand, inspection to the RMSD and RMSF of the A160Ac$_6$c mutant indicates that self-assembling does not provide any improvement of the local organization with respect to that observed for the building block. Indeed, Figure 13 shows that the fourth building block is completely unfolded after 10 ns of MD simulation. To conclude, restricting the conformational freedom at a specific position within the most mobile loop significantly enhances the intrinsic stability of 1krr self-assembled nanotubes.

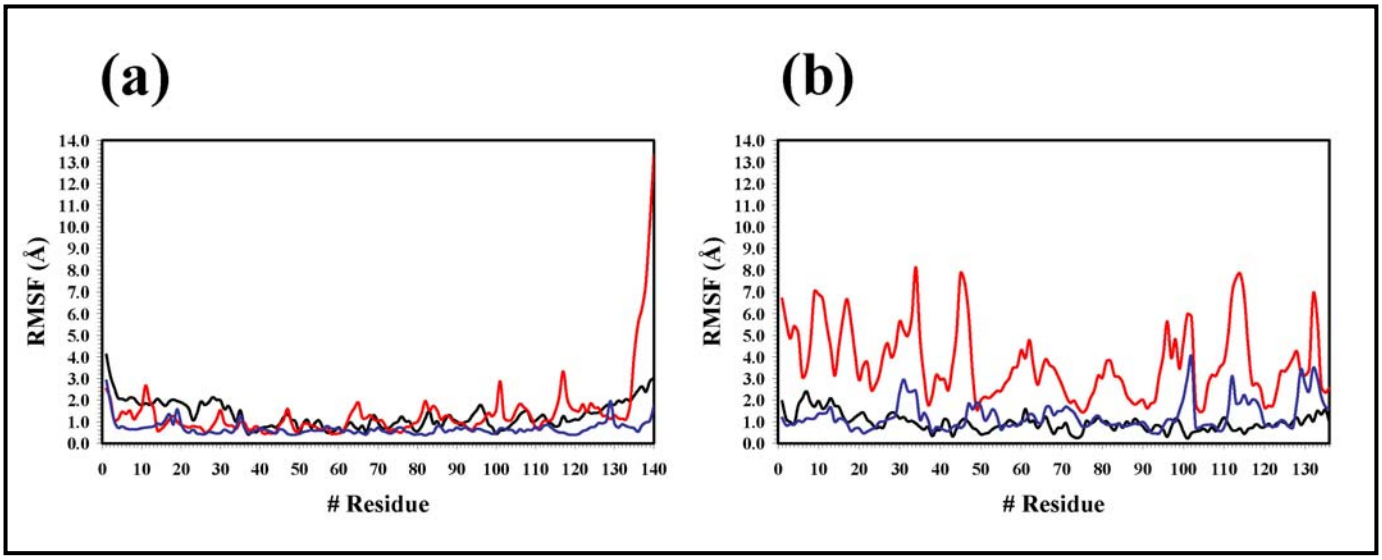


Figure 12. Comparison of the RMSF for the wild type and mutated sequences in self-assembled nanotubes. (a) Nanotubes based on 1krr: wild type (black), G149Ac$_6$c (blue) and A160Ac$_6$c (red). (b) Nanotubes based on 1hv9: wild type (black), D306Ac$_6$c (blue) and A322Ac$_6$c (red).

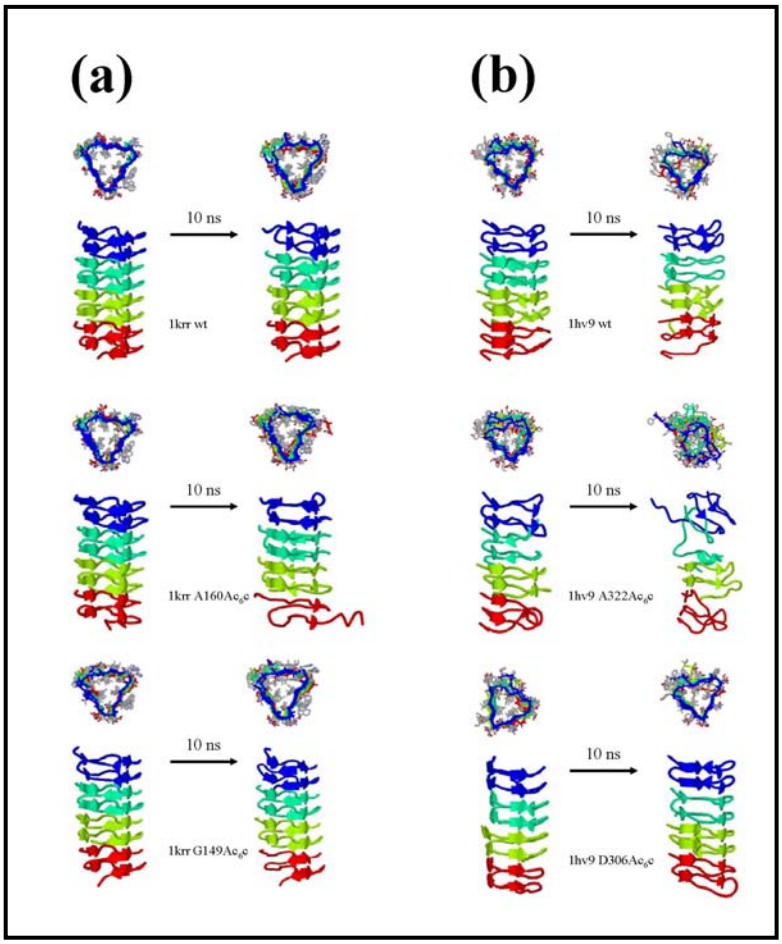


Figure 13. Structure of the nanotube models based on the self-assembly of wild type and mutated building blocks of 1krr (a) and 1hv9 (b) after 10 ns of MD simulation. The hydrogen atoms have been omitted for clarity and the backbone has been represented by solid shapes (arrows indicate sheet conformation). Stacked β-helix building blocks have been represented by different colors.

Comparison of the self-assembled nanostructure derived from D306Ac$_6$c with that constructed using the 1hv9 wild type building block is provided in Figure 10b. As can be seen, substitution in the self-assembled tube produces a slight improvement. After 10 ns of MD simulation the RMSD of the D306Ac$_6$c self-assembled tubes is about 1 Å smaller than that of the 1hv9 nanoconstruct. However, inspection of the RMSF values reveals that the mutation at the Asp-306 produces a partial fraying at the C-terminal region of all the interacting subunits, even though a slight improvement is observed at some residues located at the N-terminal and central regions. These results indicate that although electrostatic repulsions were removed in the mutated building block, they partly reappear when the Asp-305 residues of the stacked D306Ac$_6$c units interact. Finally, the results obtained for the A322Ac$_6$c self-assembled nanotube clearly indicate that mutation at the Ala-322 position destabilize not only the building block but also the nanoconstruct. This is probably consequence of the steric interactions produced by the cyclohexane side chain. Figure 13 reveals that the self-assembled organization is completely lost after 10 ns of MD simulation.

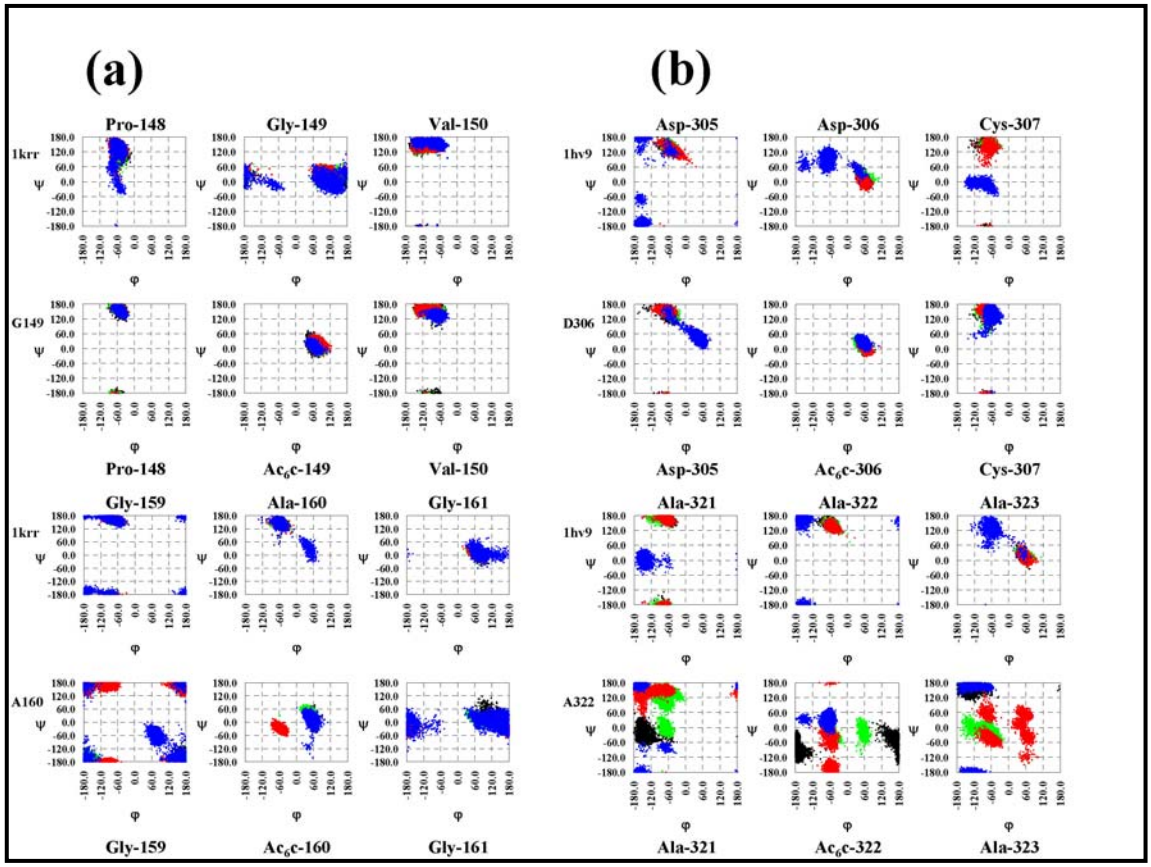


Figure 14. Comparison of the backbone torsion angle (φ,ψ) distributions of the targeted segments for the nanotubes constructed with the 1krr and 1hv9 building block simulated at 298 K. The Ramachandran plots correspond to the targeted position and its adjacent residues in (a) 1krr wild type, G149Ac6c and A160Ac6c mutants and (b) 1hv9 wild type, D306Ac6c and A322Ac6c mutants. Different colors have been used for each building block.

Comparison of the Stabilizing Effect Produced by Ac_3c and Ac_6c Substitutions

In a recent study,[5] MD simulations showed that when Ac_3c is introduced in the loop regions of 1krr and 1hv9, it is able to enhance the stability of the nanostructures due to its highly strained backbone and strong tendency to adopt turn conformations. It was found that substitutions of the middle Ala by Ac_3c in different Gly-Ala-Gly and Ala-Ala-Ala motifs lead to remarkable stability, which implies that these motifs are potential targets for Ac_3c-mutations. Although the conformational preferences of Ac_3c are severely restrained by the strain of the cyclopropane ring,[1] the whole mutated motif self-adapts to the regular β-helical organization by altering the backbone dihedral angles of the neighboring Gly and Ala residues. The latter residues act as a flexible hinge to absorb the tension induced by the cyclopropane ring in the loop, *i.e.* their flexible backbone allows a conformational adjustment.

In this work better results were obtained for the G149Ac_6c and D306Ac_6c mutants, which form self-assembled nanotubes with higher stability than those derived from 1krr and 1hv9 wild type, respectively. In these mutants the middle position of Pro-Gly-Val and Asp-Asp-Cys motifs is substituted by Ac_6c. The introduction of Ac_6c overcomes the limitations detected in Ac_3c-mutants. This should be attributed to the enhancement of conformational freedom provided by the cyclohexane ring with respect to the cyclopropane. The backbone of Ac_6c is more flexible than that of Ac_3c and different arrangements of the cyclohexane ring are compatible with each backbone conformation.

3.3.4. *Conclusions*

The intrinsic conformational preferences of Ac_6c have been examined using DFT calculations in the gas-phase at the B3LYP/6-31+G(d,p) level. Results indicated that Ac-Ac_6c-NHMe tends to adopt folded conformations, as is usual in Ac_nc dipeptides, even though its conformational freedom is higher than that of the corresponding Ac_3c, Ac_4c and Ac_5c analogues. This partial flexibility is mainly the consequence of the lack of strain in the cyclohexane ring. The influence of the environment on these conformational preferences has been examined using the PCM model to represent chloroform, methanol and aqueous solutions. Results indicated that the conformational freedom increases with the polarity of the solvent. Calculations lead to conclude that even though Ac_6c is a conformationally constrained amino acid its rigidity is

significantly lower than that of other related compounds, *i.e.* Ac_nc with $n < 6$, making it a potential candidate to stabilize nanoconstructs.

Here, Ac_6c was introduced in loop regions of 1krr and 1hv9 building blocks to examine the effect of a conformational confinement on the stability of protein nanostructures. It was observed that (i) for the more stable 1krr the substitution of Gly-149 by Ac_6c has further reduced significantly the conformational mobility not only at the mutated position but also of adjacent positions significantly stabilizing the self-assembled tube. Thus, $G149Ac_6c$ is a promising mutant of 1krr; (ii) substitution of Asp-306 by Ac_6c in the wild type unstable 1hv9 induces a remarkable structural stability especially in the building block, which are attributed to the partial elimination of unfavorable electrostatic interactions. Finally, (iii) comparison of these results with those reported previously[5] indicates that the ability of Ac_nc constrained amino acids to adapt their folded conformations within the targeted position improves with the size of the ring. The rigidity introduced by the strain energy of the cyclopropane ring makes this adaptation process more difficult for Ac_3c-mutants than for Ac_6c-mutants.

3.3.5. *References*

[1] Alemán C., Jiménez A. I., Cativiela C., Pérez, J. J., Casanovas J. *J. Phys. Chem. B* **2002**, *106*, 11849.

[2] Casanovas J., Zanuy D., Nussinov R., Alemán C. *Chem. Phys. Lett.* **2006**, *429*, 558.

[3] Alemán C., Zanuy D., Casanovas J., Cativiela C., Nussinov R. *J. Phys. Chem. B* **2006**, *110*, 21264.

[4] Alemán C., Zanuy D., Jiménez A. I., Cativiela C., Haspel N., Zheng J., Casanovas J., Wolfson H., Nussinov R. *Phys. Biol.* **2006**, *3*, S54.

[5] Zheng J., Zanuy D, Haspel N., Tsai C.-J., Alemán C., Nussinov R. *Biochemistry* **2007**, *46*, 1205.

[6] Zanuy D., Jiménez A. I., Cativiela C., Nussinov R., Alemán C. *J. Phys. Chem. B* **2007**, *111*, 3236.

[7] Bardi R., Piazzesi A. M., Toniolo C., Sukumar M., Raj P., Balaram P. *Int. J. Pepti. Protein Res.* **1985**, *25*, 628.

[8] Paul P. K. C., Sukumar M., Bardi R., Piazzesi A. M., Valle G., Toniolo C., Balaram P. *J. Am. Chem. Soc.* **1986**, *108*, 6363.

[9] Pavone V., Benedetti E., Barone V., Di Blasio B., Lelj F., Pedone C., Santini C., Crisma M., Bonora G. M., Toniolo C. *Macromolecules* **1988**, *21*, 2064.

[10] Fabiano N, Valle G., Crisma M., Toniolo C., Saviano M., Lombardi A., Isernia C., Pavone V., DiBlasio B., Pedone C., Benedetti E. *Int. J. Pept. Protein Res.* **1993**, *42*, 459.

[11] Crisma M., Bonora G. M., Toniolo C., Bavoso A., Benedetti E., Di Blasio V., Pavone C., Pedone C. *Macromolecules* **1988**, *21*, 2071.

[12] Saviano M., Isernia C., Rossi F., Di Blasio B., Iacovino R., Mazzeo M., Pedone C., Bendetti E. *Biopolymers* **2000**, *53*, 189.

[13] Dawidowska O., Wierzba T. H., Prahl A., Kowalczyk W., Gawinski L., Plackova M., Slaninova J., Lammerk B. *J. Med. Chem.* **2005**, *48*, 8055.

[14] Kowalczyk W., Prahl A., Derdowska I., Dawidowska O., Slaninova J., Lammek B. *J. Med. Chem.* **2004**, *47*, 6020.

[15] Haspel N., Zanuy D., Alemán C., Wolfson H., Nussinov R. *Structure* **2006**, *14*, 1137.

[16] Frisch M.J. et al. *Gaussian 03,* Revision B.02, Gaussian, Inc.: Pittsburgh PA, **2003**.

[17] Alemán, C., Casanovas, J., Hall, Jr., H. K. *J. Org. Chem.* **2005**, *70*, 7731.

[18] Becke A. D. *J. Chem. Phys.* **1993**, *98*, 1372.

[19] Lee C., Yang W., Parr R. G. *Phys. Rev. B* **1993**, *37*, 785.

[20] McLean A. D., Chandler G. S. *J. Chem. Phys.* **1980**, *72*, 5639.

[21] Miertus M., Scrocco E., Tomasi J. *Chem. Phys.* **1981**, *55*, 117.

[22] a) Tomasi J., Persico M. *Chem. Rev.* **1994**, *94*, 2027. b) Tomasi J., Mennucci B., Cammi R. *Chem. Rev.* **2005**, *105*, 2999.

[23] Phillips J.C., Braun R., Wang W., Gumbart J., Tajkhorshid E., Villa E., Chipot C., Skeel R.D., Kale L., Schulten K. *J. Comput. Chem.* **2005**, *26*, 1781.

[24] Jorgensen W.L., Chandrasekhar J., Madura J.D., Impey R.W., Klein M.L. *J. Chem. Phys.* **1983**, *79*, 926.

[25] Wang J., Cieplak P., Kollman P.A. *J. Comput. Chem.* **2000**, *21*, 1049.

[26] Cornell W.D., Cieplak P., Bayly C.I., Gould I.R., Merz K.M., Ferguson D.M., Spellmeyer D.C., Fox T., Caldwell J.W., Kollman P.A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

[27] Darden T., York D., Pedersen L. *J. Chem. Phys.* **1993**, *98*, 10089.

[28] Ryckaert J.P., Ciccotti G., Berendsen H.J.C. *J. Comput. Phys.* **1977**, *23*, 327.

[29] Berendsen H.J.C., Postma J.P.M., van Gunsteren W.F., DiNola A., Haak J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

4. DENDRONIZED POLYMERS

4.1. A rigid, chiral, dendronized polymer with a thermally stable, right-handed helical conformation[†,‡]

4.1.1. *Introduction*

Chiral helical polymers have received considerable interest recently for mimicking biological helices and developing novel chiral materials.[1] They have been prepared either by achiral polymerization of optically active monomers such as acetylenes,[2] isocyanates,[3] and silanes,[4] or by helix-senseselective polymerization of achiral monomers (carbodiimides[5] and methacrylates[6]) and chiral monomers (isocyanides[7] and methacrylamides[8]). The former helical polymers are dynamic in solution, and right-handed helices can change into left-handed ones and vice versa under certain conditions, while the helical conformations of the latter are more persistent. Recently, dendronized polymers[9] with polyacetylene backbones were reported by Masuda *et al.*[10] and Percec *et al.*[11] to adopt helical conformations with limited thermal stability. In a related recent publication[12] a helical (nondendronized) polyacetylene was reported for which the thermal stability of its helix could be increased significantly by increasing the size of the substituent that followed the chiral center near the backbone.[13]

The motivation of the present research was to prove if it would be possible to use the "thickening" of polymer chains that occurs in dendronized polymers to render them rigid and, at the same time, chiral with an unprecedented thermally stable helical conformation. Prof. Dr. A. Zhang and Prof. Dr. A.D. Schlüter attempted this by polymerizing chiral 4-aminoproline-based second generation macromonomer **MG2** (Figure 1), with its rather compact and chiral second-generation dendron, and it was of major interest whether such a monomer would be polymerizable. If so, it was expected that the dendrons would be forced into a tight packing around the backbone which should lead to rigidification and helix-sense creation. This chapter reports the results of the spontaneous polymerization of **MG2** leading to helical dendronized polymer **PG2** together with a spectroscopic and molecular dynamics (MD) simulation analysis on the characteristics of this helix, including stability aspects. A detailed structural model of the helix is also provided.

[†] Results presented in this chapter are published in *Chem. Eur. J.* **2008**, *14*, 6924.

[‡] Experimental work has been performed by Prof. Dr. A. Zhang and Prof. Dr. A. D. Schlüter (Institute of Polymers, Department of Materials ETH Zürich). UV spectra performed by Prof. Dr. E. W. Meijer (Laboratory of Macromolecular and Organic Chemistry Eindhoven University of Technology).

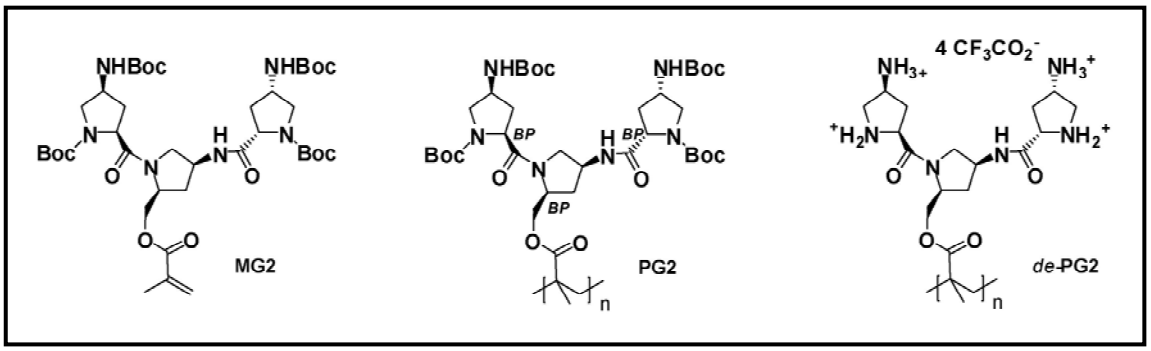


Figure 1. Chemical structures of second-generation, 4-aminoproline-based, Boc-protected dendronized methacrylate monomer **MG2**, the corresponding polymer **PG2**, and its deprotected derivative *de*-**PG2**. BP stands for branching points, and the number of atoms between them (here 2 and 4) is used in the text as a rough measure of dendron compactness.

4.1.2. *Computational methods*[§]

Quantum chemical calculations on **2c** (see Figure 2) were performed at the HF/3-21G,[23] HF/6–31G(d),[24] and B3LYP/6–31G(d)[25] levels with the Gaussian 03 program suite.[29] Atomistic MD simulations of isotactic, syndiotactic, and heterotactic models of **PG2** were performed with the NAMD program.[30] For NVT-MD in chloroform and methanol solutions, the simulated system was placed in the center of an orthorhombic simulation box filled with explicit solvent molecules. Chloroform was represented by the four-particle OPLS model,[31] while the all-atom model reported by Caldwell and Kollman was used for methanol.[32] In all cases **PG2** was simulated by considering explicitly all atoms of a 20 r.u. model molecule, which was blocked at the ends with methyl groups. The dimensions of the orthorhombic box for simulations in chloroform and methanol solutions were (99.31 x 99.31 x 115.00) and (105.65 x 105.65 x 115.25) $Å^3$, respectively, while the total number of particles considered explicitly was 35496 (8267 chloroform molecules) and 115888 (18910 methanol molecules). The energy was calculated with the AMBER force field,[26] and all parameters required for **PG2** were taken from the AMBER libraries with the exception of the electrostatic ones. Atomic charges were explicitly developed by using a previously developed procedure.[33] For this purpose, the molecular electrostatic potential was calculated at the HF/6-31G(d) level for a model system constituted by 3 r.u.s. Atom-pair distance cutoffs were applied at 12.0 Å to compute the van der Waals interactions. The electrostatic interactions were computed by using the nontruncated electrostatic potential with Ewald summations.[34] The real-space term was determined by the van der Waals cut off (12.0 Å).

[§] Syntheses and characterization are available in *Chem. Eur. J.* **2008**, *14*, 6924.

Bond lengths involving hydrogen atoms were constrained by using the SHAKE algorithm[35] with a numerical integration step of 2 fs. Before the MD production series, the thermodynamic variables of the system were equilibrated. The energy of each system was initially minimized to relax conformational and structural tensions by using the conjugate gradient method for 57103 steps. Next, the dendronized polymer was placed in a previously equilibrated solvent box. Different consecutive rounds of short MD runs were performed to equilibrate the variable thermodynamic magnitudes. First, solvent was thermally relaxed by three consecutives runs, while the dendronized polymer was kept frozen: 0.2 ns of NVT-MD at 500 K was used to homogeneously distribute the solvent in the box. Second, 0.15 ns of isothermal relaxation at 298 K was run. Finally, all atoms of the system were submitted to 0.25 ns of NVT-MD at 298 K (thermal equilibration). Temperature was controlled by the weak coupling method, the Berendsen thermostat,[36] with a time constant for heat-bath coupling of 1 ps. The end of the thermal relaxation simulation was the starting point of the molecular simulations presented in this work. All simulations were performed at 298 K. The coordinates of all the production runs, which were between 6 and 50 ns long, were saved every 500 steps (1 ps intervals) for subsequent analysis.

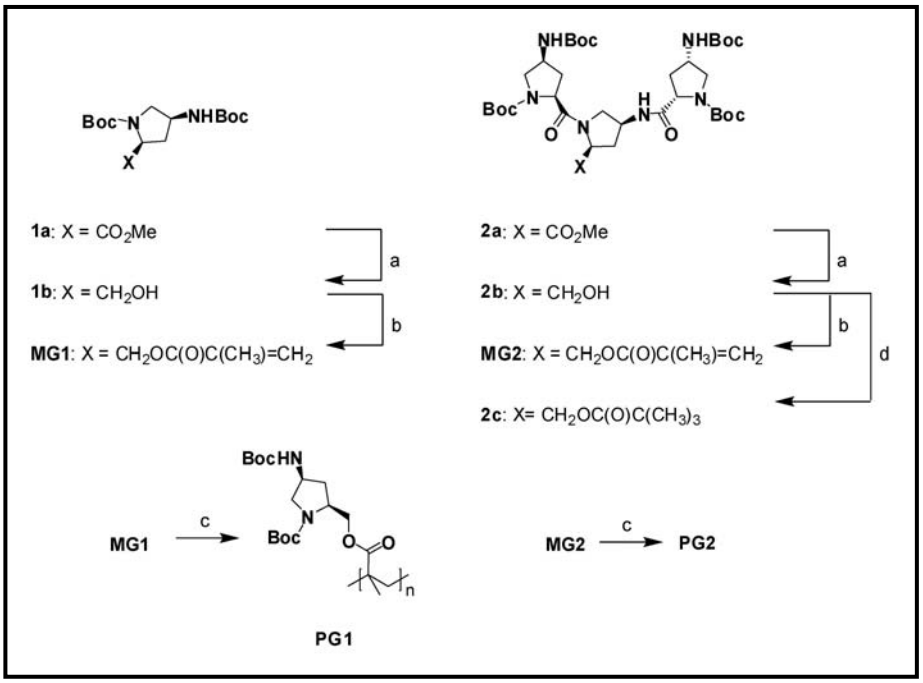


Figure 2. Synthesis of macromonomers **MG1** and **MG2**, as well as **PG1** and **PG2**. a) $LiBH_4$, THF, -10ºC, 12 h (**1b**: 92%, 2b: 91%); b) MAC, DMAP, NEt_3, THF, 0ºC, 12 h (95% and 92% for **MG1** and **MG2**, respectively); c) DMF, 60ºC, 12 h; d) trimethylacetyl chloride, NEt3, DMAP, dichloromethane, 0ºC, 6 h (92%). MAC=methacryloyl chloride, DMAP=4-dimethylaminopyridine.

The conformational stability and conservation of the helix when the **PG2** model changes from isotactic to syndiotactic and heterotactic was measured by calculating the evolution of the backbone root mean square deviation (RMSD) through the simulation relative to the idealized helix proposed for the heterotactic model and the root mean square fluctuation (RMSF) of individual repeat units averaged over the whole simulation. To obtain the helix parameters for idealized models, the backbone dihedral angles were averaged with consideration of the 18 central repeat units, that is, the dihedral angles of the repeat units located at the end were omitted because of the frying. The end-to-end distance was obtained by measuring the distance between the carbon atoms of the blocking methyl groups of each recorded microstructure.

4.1.3. *Results and discussion*

Syntheses

To achieve the goals mentioned above, an ideal dendron should be chiral and compact. Dendrons constructed from 4-aminoproline branching units therefore accomplish these requirements. Taking the atoms between the branching points as a simple measure for compactness,[14] **MG2** should result in a polymer with a very compact dendritic layer that forces consecutive dendrons into register. Because of the chiral nature of the monomer, this could give rise to a biased helix sense of **PG2**. Recently, various such second-generation, diastereomeric dendrons were made available on the gram scale through fully optimized procedures.[15] This includes the dendron contained in **MG2** with its three all-(2S,4S)-configured branching units (**2a** in Figure 2), and that was used for the present study. Figure 2 shows the synthetic sequence to **MG2** and, for comparison purposes, also that to **MG1**. In the first step esters **1a** and **2a** were reduced with $LiBH_4$ to alcohols **1b** and **2b**, which were then esterified with methacryloyl chloride to give the corresponding monomers (**MG1**: 87%; **MG2**: 84%). Both monomers were obtained on a multigram scale as analytically pure materials and characterized by ^{1}H and ^{13}C NMR spectroscopy, as well as high resolution mass spectrometry. After careful purification by column chromatography, they were left to stand as highly concentrated DMF solutions at 60ºC without addition of any radical initiator, under which conditions even **MG2** started to polymerize spontaneously.[16] After a few hours the solutions had turned solid, and the polymers **PG1** and **PG2** were isolated by column chromatography on silica gel. For comparison, the polymerization of **MG2** was also conducted in the presence of

radical initiator 2,2'-azobisisobutyronitrile (AIBN). All polymers were obtained as colorless solids. Their molar masses were determined by gel permeation chromatography (DMF with 1% LiBr as eluent, two-angle light-scattering detection, 45ºC), and the results are compiled in Table 1. Polymers **PG1** and **PG2** were obtained with the highest molar masses when polymerization was carried out in concentrated solutions spontaneously (Table 1, entries
1–4). In the presence of 2,2'-azobisisobutyronitrile (AIBN) the molar mass of **PG2** was lower by roughly a factor of two. The molar masses are by far the highest ever obtained for peptidic dendronized polymers even with the less compact amino acids lysine and glutamic acid as repeat units (r.u.s).[17] As is typically observed for dendronized polymers prepared by the macromonomer approach, the average chain lengths of **PG2** were shorter than those of **PG1**.

Table 1. Conditions for and results of the polymerizations of **MG1** and **MG2**.

Entry	Polymerization conditions[a]			Yield (%)	GPC[b]			
	Monomer	Concetration [mol·L^{-1}]	Time [h]		$M_n \cdot 10^{-6}$	DP·10^{-3}	PDI	T_g [ºC]
1	**MG1**	1.59	5	80	0.76	2.04	2.80	125
2[c]	**MG1**	2.67	6	52	5.00	13.4	3.66	129
3	**MG2**	0.64	48	50	0.77	0.95	1.94	>200
4	**MG2**	0.79	25	60	1.11	1.37	2.24	>200
5[d]	**MG2**	0.67	24	52	0.38	0.47	2.23	>200

[a] Polymerization carried out at 60º in DMF. [b] All GPC measurements were done in DMF at 80ºC. [c] Polymerization in bulk. [d] With 13.8m AIBN as initiator.

Polymer rigidity

A first indication of unusual stiffness of PG2 came from NMR spectroscopy. Although **PG1** and **PG2** with *tert*-butoxycarbonyl (Boc) protecting groups have good solubility in $CDCl_3$ and CD_3OD, only the former polymer gave a reasonable ^{1}H NMR spectrum at room temperature in which all expected signals were visible (Figure 3 a, top). For the latter, only a broad signal for the peripheral Boc groups was recorded and most of the other signals were absent (Figure 3 a, bottom).[18] The same phenomenon was observed for the ^{13}C NMR spectrum of **PG2**, in which again practically all signals were absent, except for those of the Boc groups (Figure 3b, top). The solid-state CPMAS ^{13}C NMR

spectrum of **PG2** showed all expected signals (Figure 3b, bottom). It seems that the reduced mobility of the magnetic nuclei due to the high rigidity of **PG2** is responsible for these effects, which were not encountered to this extent with other dendronized polymers. Differential scanning calorimetry (DSC) measurements gave T_g values for **PG1** and **PG2** of 129 and above 200ºC (beyond decomposition), respectively, which are the highest so far observed for (less compact) dendronized polymers. Typically, T_g values are in the range of 55-80ºC.[14c] Compared to parent poly(methyl methacrylate), PMMA (T_g=105 ºC), the T_g of PG2 is at least 95ºC higher, which indicates an increase in rigidity, though the glass transition is not only determined by chain stiffness. Due to the low resolution of the proton spectrum of **PG2**, the backbone tacticity was estimated by using its analogue *de*-**PG2**. In contrast to PMMAs from conventional radical polymerization in DMF, the backbones of which are dominated by syndiotactic units (ca. 62%),[19] *de*-**PG2** contains mostly heterotactic (ca. 63%) and a smaller amount of syndiotactic units (ca. 32%).

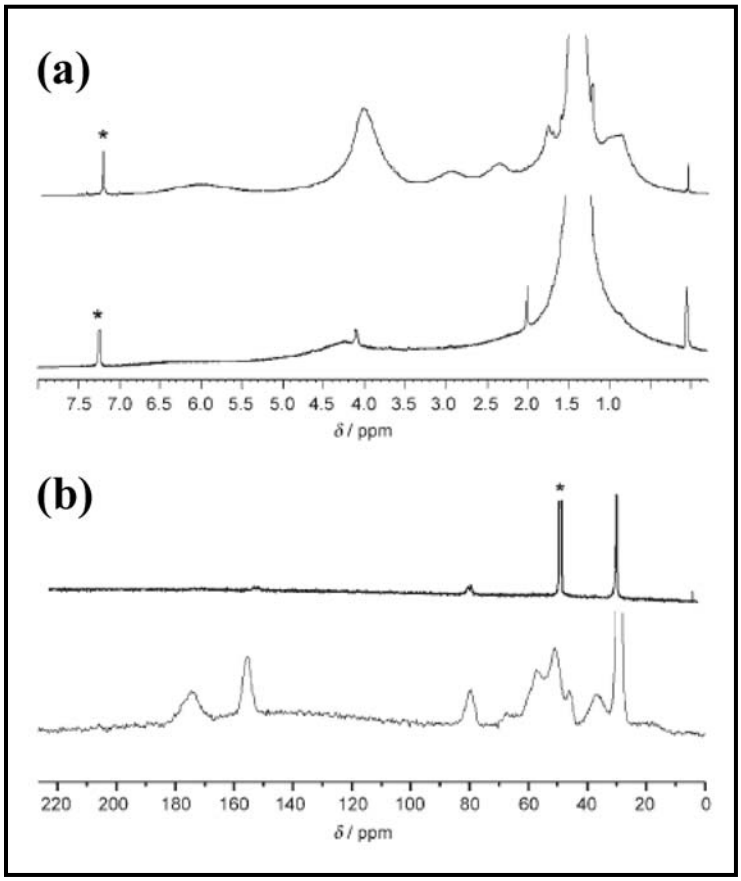


Figure 3. (a) ^{1}H NMR spectra of **PG1** (top) and PG2 (bottom) in $CDCl_3$ and (b) ^{13}C NMR spectra of PG2 in CD_3OD (top) and in the solid state (bottom) at room temperature. Solvent peaks are marked with asterisks.

It is noteworthy that the ester bonds of **PG2** through which the dendrons are attached to the backbone cannot be easily saponified. Treatment with lithium or sodium hydroxide in refluxing MeOH for 24h left this polymer completely unchanged, whereas other less compact dendronized polymers (of even higher generation) can be efficiently dedendronized under these conditions.[20] This is an indication of the compactness of the dendritic layer of **PG2**.

Finally, MD simulations in both chloroform and methanol solutions of a model polymer chain consisting of 20 repeat units were performed to investigate the rigidity of **PG2** at the atomistic level. Simulations were performed with explicit solvent molecules and a helical conformation compatible with available experimental data, which was derived from a systematic conformational search (see below). The end-to-end distance remained unchanged at (46.83±0.26) and (46.32±0.45) Å in chloroform and methanol, respectively, during the whole simulation time range of 50 ns (Figure 4). The high stiffness of PG2 was also supported by the constant cross-sectional diameter (average values: 29.16±0.30 and 29.10±0.44 Å in chloroform and methanol, respectively). Consistently, the temporal evolution of this property shows that no significant fluctuation occurred along the trajectories (Figure 4).

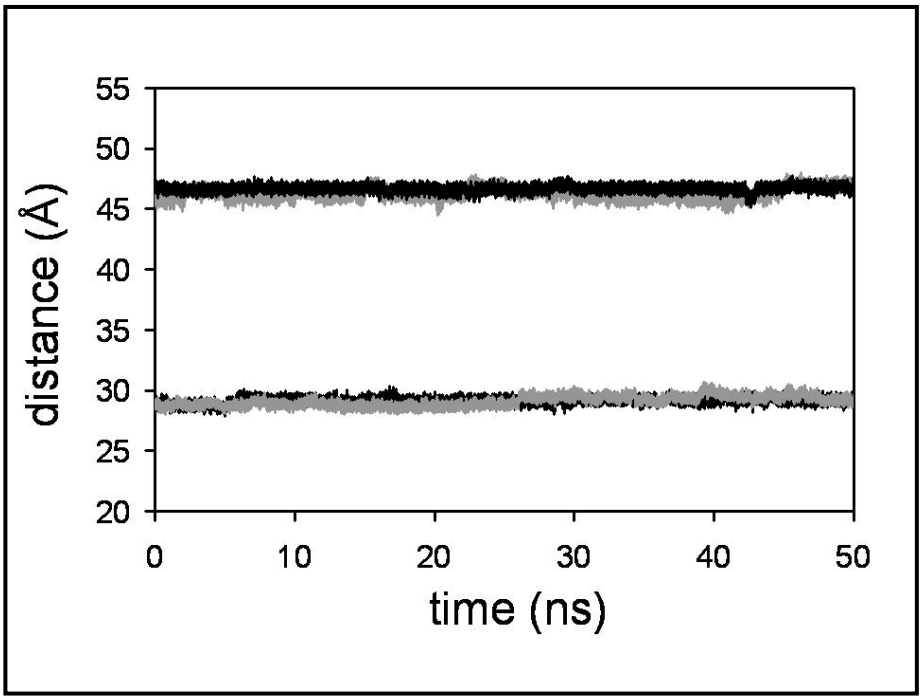


Figure 4. Time-dependent evolution of the end-to-end distance (upper data set) and the cross-sectional diameter (lower data set) in chloroform (black lines) and methanol (gray lines) solution. The lines representing the two environments overlap.

Secondary structure and helix conformation

The secondary structures of the polymers were investigated by optical rotation (OR) and circular dichroism (CD) spectroscopic measurements (Table 2, Figure 5). The OR measurements were performed with protected polymers **PG1** and **PG2**, whereas the CD spectra were additionally recorded also for both the deprotected and thus positively charged *de*-**PG1** and *de*-**PG2**. The OR of **PG1** was practically identical to that of **MG1**, whereas for **PG2** it was slightly less intense compared to that of **MG2**, irrespective of whether CH_3OH or $CHCl_3$ was used as solvent. This decrease suggests a different secondary structure for the polymer compared to the monomer and is most likely due to restricted conformational freedom in **PG2**.[21] In other words, **PG1** does not adopt an ordered secondary structure, while **PG2** may do so, though the difference in OR between **MG2** and **PG2** is rather small. Such small differences were interpreted by Masuda *et al.*[12] in terms of high rigidity of the polymer backbone.

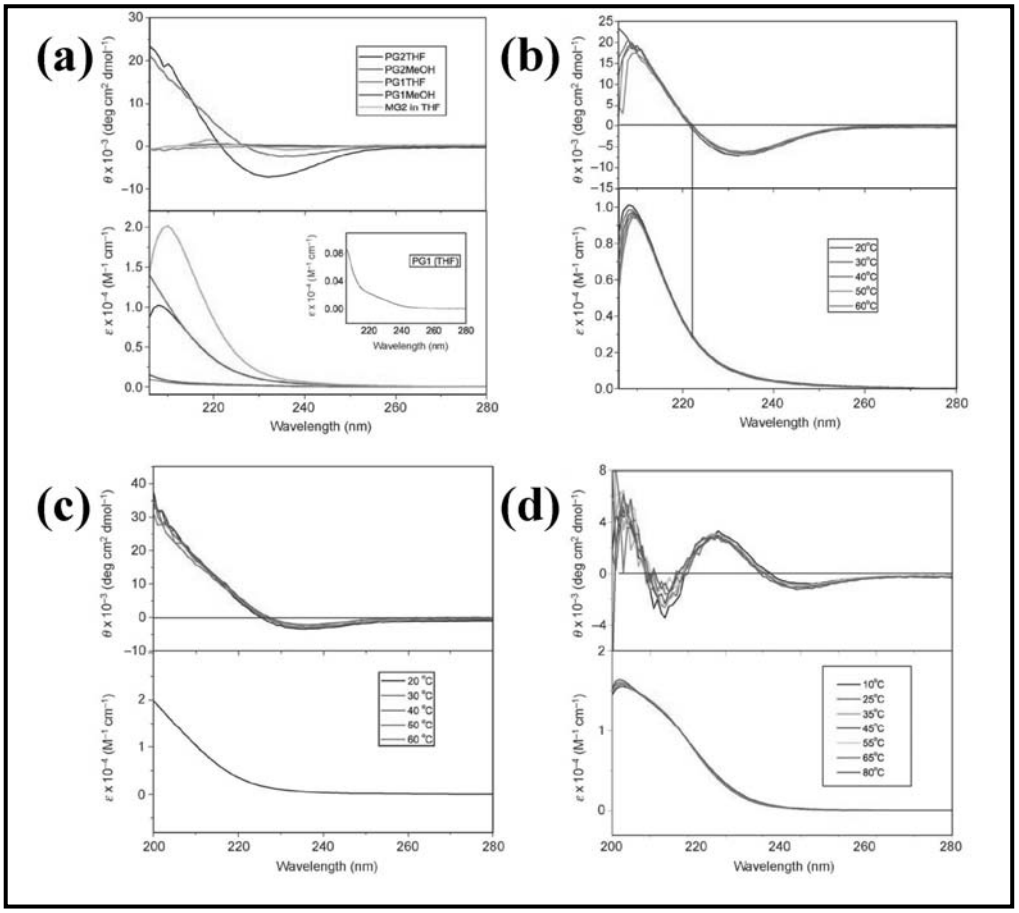


Figure 5. CD (top) and UV spectra (bottom) of (a) **MG2** in THF and polymers **PG1** and **PG2** in MeOH and THF at room temperature (inset: UV spectrum of **PG1** in THF), (b) **PG2** in THF, (c) **PG2** in MeOH, and (d) *de*-**PG2** in water at different temperatures.

Table 1. Specific rotations of **MG1**, **MG2**, **PG1** and **PG2**.[a]

Entry	Concentration [g·dL^{-1}]	$[\alpha]_D^{25}$	
		MeOH	$CHCl_3$
MG1	1.48	-30.2831	-
PG1	1.25	-31.6560	-
MG2	1.13 (1.02 [b])	-37.0973	-52.5961
PG2	0.95 (1.14 [b])	-22.3116	-38.9009

[a] All measurements were done at 25 °C with a 1 dm cuvette in MeOH.
[b] For chloroform.

Figure 5 summarizes the results of all CD and UV spectroscopic measurements performed with **PG1**, **PG2**, and *de*-**PG2** in different solvents and at different temperatures. Figure 5a shows the normalized CD spectra of **PG1** and **PG2** in MeOH and THF. The former did not exhibit a significant absorption in the range of λ= 200-260 nm, whereas the latter has a negative Cotton effect at λ = 232 nm. This is a strong indication that **PG2**, but not **PG1**, adopts an ordered secondary structure. The CD spectrum of **MG2** in THF was also recorded for comparison and showed only a weak negative signal at 237 nm (Figure 5a), which again suggests that **PG2** adopts an ordered structure in solution. The Cotton effect of **PG2** was stronger in THF than in MeOH. Figure 5b and c show the temperature-dependent CD and UV spectra of **PG2** in THF and MeOH, respectively. In the temperature range of 20-60ºC, they turned out to be practically superimposable, indicative of a fully retained secondary structure under these conditions. Figure 5d shows the behavior of *de*-**PG2** in water at pH 4 in the temperature range of 10-80ºC. The appearance of its CD and UV spectra differs from the above because of the different chromophore. Despite the considerable loss of mass associated with deprotection, the secondary structure is basically retained, though a somewhat more pronounced temperature dependence is observed as compared with protected congener **PG2**.

These results point towards a stable and chiral secondary structure of **PG2**, which even seems to be largely maintained in the deprotected form. The nature of the secondary structure was addressed by computations combining quantum chemical methods and MD simulations based on isotactic and heterotactic model polymers with 20 repeat units.[22] Initially, the more favorable arrangements of the model compound **2c** (Figure 2), resembling a large part of the polymer repeat unit, were examined by exploring its potential-energy hypersurface. More specifically, a rigorous multidimensional conformational analysis was performed by considering that three minima are expected

for each flexible dihedral angle involved in the fragment of **2c** containing four aminoproline residues. Twenty-seven possible structures were considered as starting points for complete geometry optimization at the HF/3-21G level,[23] and the 18 minima of lower energy were reoptimized at the HF/6-31G(d) level;[24] the remaining HF/3-21G minima were separated from the lowest-energy minimum by more than 10 kcal·mol^{-1}. Finally, single-point calculations at the B3LYP/6-31G(d) level[25] were carried out on the resulting HF/6-31G(d) geometries to obtain reliable energy estimates.

The two most stable conformations of **2c** were considered in a systematic study devoted to building sterically accessible helical models. For this purpose, the backbone dihedral angles of the dendronized methacrylate repeat unit, hereafter denoted χ_1 and χ_2, were varied systematically in steps of 30º within an isotactic model system of **PG2** consisting of 20 repeat units, which were kept with identical conformation. The structures without apparent steric contacts were subjected to energy minimization with the Amber force field.[26] Surprisingly, 80 helical conformations were able to retain their regularity and homogeneity, that is, similar (±20º) values of χ_1 and χ_2 for the 20 repeat units. Next, the 80 structures were submitted to 0.5 ns of MD at 298 K in the absence of environment. Only eight models maintained the secondary structure, and the helical conformation of the remaining 72 models was rapidly lost owing to thermal vibrations. After this, a NVT-MD in chloroform solution was run for 6 ns for each of the eight structures. These simulations provided two models with regular and homogeneous secondary structure and, in addition, similar energies. Finally, 6 ns NVT-MD simulations in methanol solution were performed for these two structures. Although the regularity was maintained again and their internal energies were very similar in the two cases, the solute–solvent interactions were significantly more favorable for one of the two helical models.

It consists of a right-handed helix with average backbone dihedral angles of χ_1= (-159.5±2.8)º and χ_2= (-52.2±2.1)º (Figure 6a). The characteristic helical parameters derived for this structure considering an idealized system formed by identical repeat units (Figure 6a) are 2.65 repeat units per turn, 0.71 Å internal radius, and 2.12 Å rise per repeat unit. This structure not only remains stable but also is very rigid in solution in both chloroform and methanol (Figure 4). These unique properties should be attributed not only to the size of the dendron but also to the hydrogen-bonding network detected in the secondary structure, which extends along the whole polymer chain. Thus, the NH

groups of the 4-aminoproline residues form this kind of interaction with the amide oxygen atoms of either the neighboring repeat unit (Figure 6b, right) or the same repeat unit (Figure 6b, left), which enhances the stability and stiffness of the helical conformation.

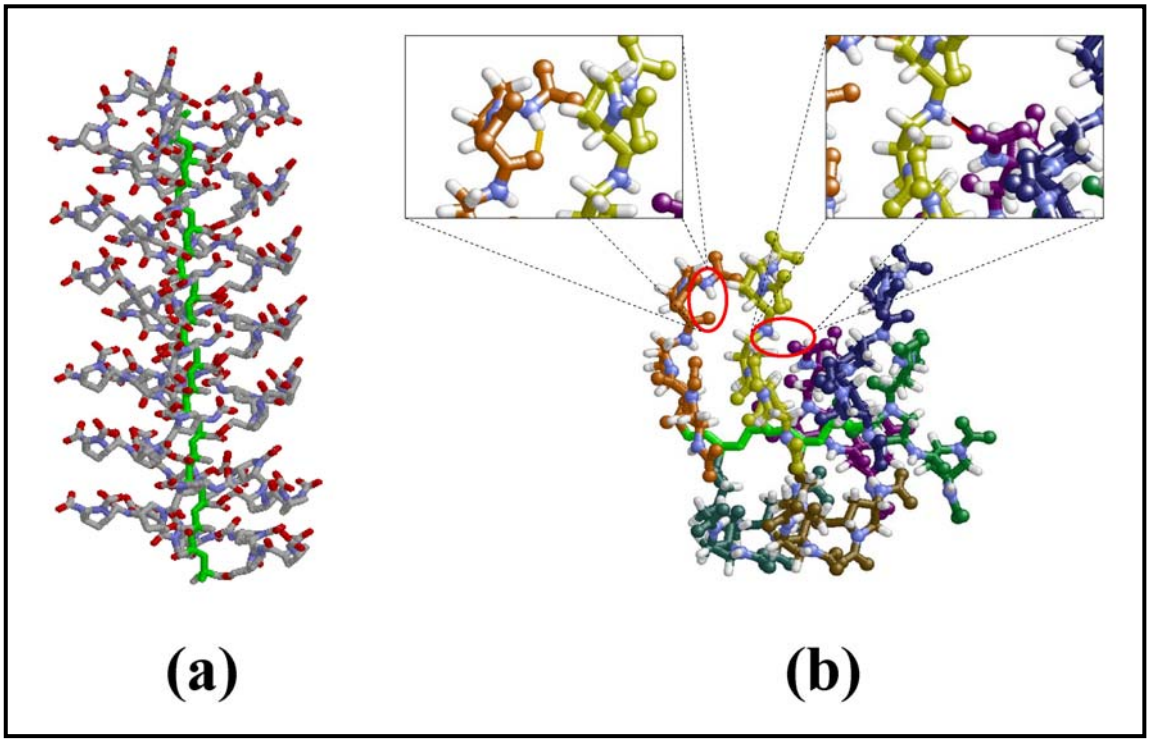


Figure 6. (a) Axial projection of the right-handed helical model of **PG2**. *tert*-Butyl groups and hydrogen atoms have been removed for clarity. The backbone is represented by a solid green line. (b) Details of the N-H···O=C hydrogen-bonding network detected in the helical secondary structure. The magnified view on the right shows the hydrogen bond (red line) between different r.u.s, and that on the left the hydrogen bond (yellow line) between groups of the same r.u.

To ascertain whether this helical model is only valid for isotactic **PG2** or does not depend on the backbone tacticity, a complementary study was performed on syndiotactic and heterotactic models for a polymer chain of 20 repeat units. Table 3 lists the 16 sequences used for such study, where A and B refer to repeat units that differ in the configuration in the chiral backbone atom. Isotactic, syndiotactic, and heterotactic models were analyzed. Initially, the helix found for the isotactic A20 model was imposed in the other 15 sequences. After energy minimization, a regular and homogeneous helical conformation was retained in all cases. Next, each minimized structure was subjected to 10 ns of NVT-MD simulation in chloroform solution, equilibration, and thermalization of the simulation box.

Table 3. Average values of the backbone root-mean-square deviation[a] ($RMSD_{Av}$) calculated with respect to the idealized helical model proposed after a detailed conformational search on an isotactic model, backbone dihedral angles[a] (χ_1 , χ_2), cross sectional diameter[a] (D), end-to-end distance[a] ($d_{e\text{-}e}$), internal energy of the helix[b] relative to the most stable model (ΔE_{den}) and potential energy of the whole simulated system[b] (dendronized polymer and explicit solvent molecules) relative to the most stable system (ΔE_{den+s}).

# Model[c]	$RMSD_{Av}$ [Å]	χ_1 , χ_2 [º]	D [Å]	$d_{e\text{-}e}$ [Å]	ΔE_{den} [kcal·mol^{-1}·(r.u.)$^{-1}$]	ΔE_{den+s} [kcal·mol^{-1}·(r.u.)$^{-1}$]
A_8-B_4-A_8	1.875±0.288	-169.0±1.9,-110.8±8.1	30.11±0.23	37.88±0.80	2.2	-0.6
A_7-B-A_4-(B-A)$_2$-A_2-B-A	1.368±0.256	-165.2±2.0, -67.9±2.2	30.33±0.20	45.12±0.51	0.3	-2.5
A_2-B-A_6-B-A-B-A_3-B-A_4	1.676±0.512	-166.6±2.8, -90.5±7.3	30.05±0.29	40.16±1.45	0.0	0.0
A_2-B-A_6-B_3-A_3-B-A_4	1.154±0.188	-173.6±3.3, -94.0±1.7	30.42±0.34	45.93±0.50	2.8	-0.2
A-(A-B)$_2$-A_4-B_3-A_3-B-A_4	1.734±0.215	-168.1±1.9, -67.8±1.5	30.43±0.26	44.51±0.32	2.4	-1.9
A-(A-B)$_2$-A_3-B_4-A_3-B-A_4	1.049±0.343	-163.3±2.3, -70.6±1.4	29.84±0.27	45.74±0.50	0.8	-1.9
A_2-B_3-A_3-B_4-A_3-B-A_4	1.457±0.232	-169.7±2.8, -67.5±1.4	29.18±0.31	41.23±0.69	3.1	-2.1
A_2-B_3-A_2-B_5-A_3-B-A_4	1.037±0.142	-167.5±2.5, -67.8±1.5	29.44±0.20	44.67±0.41	2.6	-2.5
A_2-B_3-A_2-B_6-A_2-B_2-A_3	1.009±0.186	-161.1±2.5, -70.2±1.6	29.56±0.21	42.85±0.89	3.2	2.8
A_2-B_3-A_2-B_6-A_2-B_2-A_2-B	0.888±0.123	-163.5±2.1, -68.1±1.4	29.67±0.48	44.14±0.45	1.3	-0.2
B-A-B_3-A_2-B_6-A_2-B_2-A_2-B	1.131±0.142	-164.8±2.4, -68.7±1.5	30.53±0.36	44.08±0.76	3.7	-1.9

[a] Averaged over the whole simulation. [b] Averaged over the last 50 recorded snapshots. [c] A and B refer to repeat units that differ in the configuration in the chiral backbone atom. [d] Used to derive the idealized helical model of A_{20}, which was used as reference.

The helical structure was stable for all the investigated sequences, as evidenced by the small values of the averaged (over the whole simulation) root-mean-square deviations ($RMSD_{Av}$) of the backbone, which were calculated relative to the right-handed helix proposed for model A_{20}. Thus, the $RMSD_{Av}$ values (Table 3) were smaller than 2 Å in all cases, independent of backbone tacticity. Furthermore, the $RMSD_{Av}$ value obtained for the syndiotactic sequence $(A\text{-}B)_{10}$ is extremely low ((0.744±0.148) Å), whereas the values for the sequences formed by short blocks of A and B repeat units are also notably low (~1 Å).

Figure 7a compares the temporal evolution of the backbone RMSD for selected models: syndiotactic $(A\text{-}B)_{10}$, isotatic B_{20} and, heterotactic $A\text{-}(A\text{-}B)_2\text{-}A_3\text{-}B_4\text{-}A_3\text{-}B\text{-}A_4$ and $A_8\text{-}B_4\text{-}A_8$. Although in some cases, for example, $A\text{-}(A\text{-}B)_2\text{-}A_3\text{-}B_4\text{-}A_3\text{-}B\text{-}A_4$, minor conformational rearrangement occurs after 4-5 ns, in general the RMSD converges after a few hundred picoseconds. Practically identical behavior was found for the remaining 11 models. Thus, for each model the RMSD remains close to the average value (Table 3) for the whole simulation, that is, the stability of the proposed helical model is remarkably high and does not depend on backbone tacticity. Figure 7b displays the backbone root-mean-square fluctuation (RMSF) of the individual repeat unit averaged over the whole simulation for the four models selected above. Although, in general, no significant fluctuation was found for the 20 repeat units, the distortions were larger for the model $A_8\text{-}B_4\text{-}A_8$. Inspection of the RMSF obtained for this model clearly indicates that the main distortions are located in the short B block and the extremes of the molecule. In contrast, the low and uniform RMSF values for all 20 repeat units of the $A\text{-}(A\text{-}B)_2\text{-}A_3\text{-}B_4\text{-}A_3\text{-}B\text{-}A_4$, B_{20}, and $(A\text{-}B)_{10}$ models, especially for the last named, evidences a remarkable conservation of the right-handed helix. These results clearly demonstrate that the helical structure proposed for the isotactic polymer does not depend on backbone tacticity; this structural motif is also compatible with syndiotactic and heterotactic systems. However, both RMSD and RMSF profiles suggest that such a helical conformation is particularly stable when the A and B repeat units alternate frequently, that is, ideally in a syndiotactic arrangement.

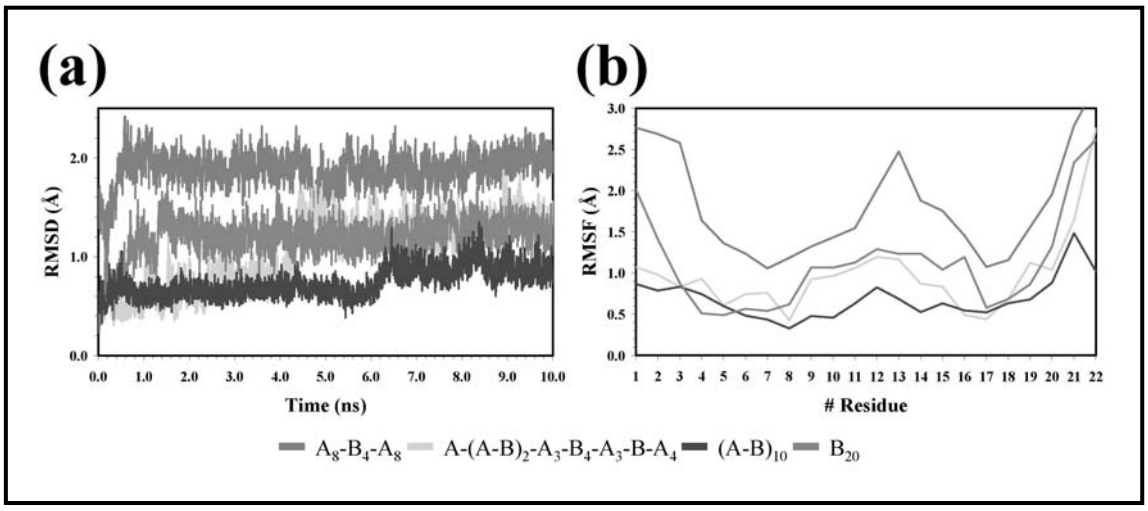


Figure 7. (a) Evolution of the backbone RMSD for four selected models of dendronized polymer **PG2** with respect to the right-handed helix proposed for the isotactic model A_{20}. (b) RMSF calculated for the 20 r.u.s considered in the four selected models.

Table 3 includes the average values of the structural parameters determined for the 16 sequences examined: diameter of the helix *D* and end-to-end distance $d_{e\text{-}e}$. These parameters are in agreement with those obtained for the isotactic model used in the conformational search strategy. Furthermore, the very small standard deviations of the resulting values in all cases indicate that the reduced mobility and high rigidity found for **PG2** is also independent of backbone tacticity. On the other hand, the decrease in the χ_1 and χ_2 values produced by incorporation of B repeat units into the molecular chain should be attributed to packing of the dendrons, which is less compact in syndiotactic and heterotactic systems than in isotactic ones. However, the hydrogen-bonding network (Figure 6b) is preserved in all cases. Consideration of all the data recorded from the simulations of the 13 B containing heterotactic models allows one to propose a right-handed helix with averaged backbone dihedral angles of χ_1=(-171.3±2.4) Å and χ_2=(-64.2±1.6) Å. The characteristic helical parameters of an idealized system that adopts such a conformation are 3.03 repeat units per turn, 0.79 Å internal radius, and 2.09 Å rise per repeat unit. These results clearly reflect that the helical conformation predicted for isotactic PG2 is close to that obtained for heterotactic models. The most notable difference between the conformations obtained for isotactic and heterotactic PG2 is the number of residues per turn, which is higher for the latter. The larger number of residues per turn in the heterotactic models results in enlargement of the helix pitch and therefore a decrease in the compactness of the dendritic side groups with respect to the isotactic model. Differences in the packing of the dendrons

are clearly evidenced in Figure 8a, which compares the equatorial projections of the two idealized helical conformations. Moreover, the helical conformation proposed for the heterotactic model is fully compatible with the growth of the polymer chain during the polymerization process. Thus, the low steric hindrance between the incoming dendron and the helix (Figure 8b) supports helix formation during polymerization.

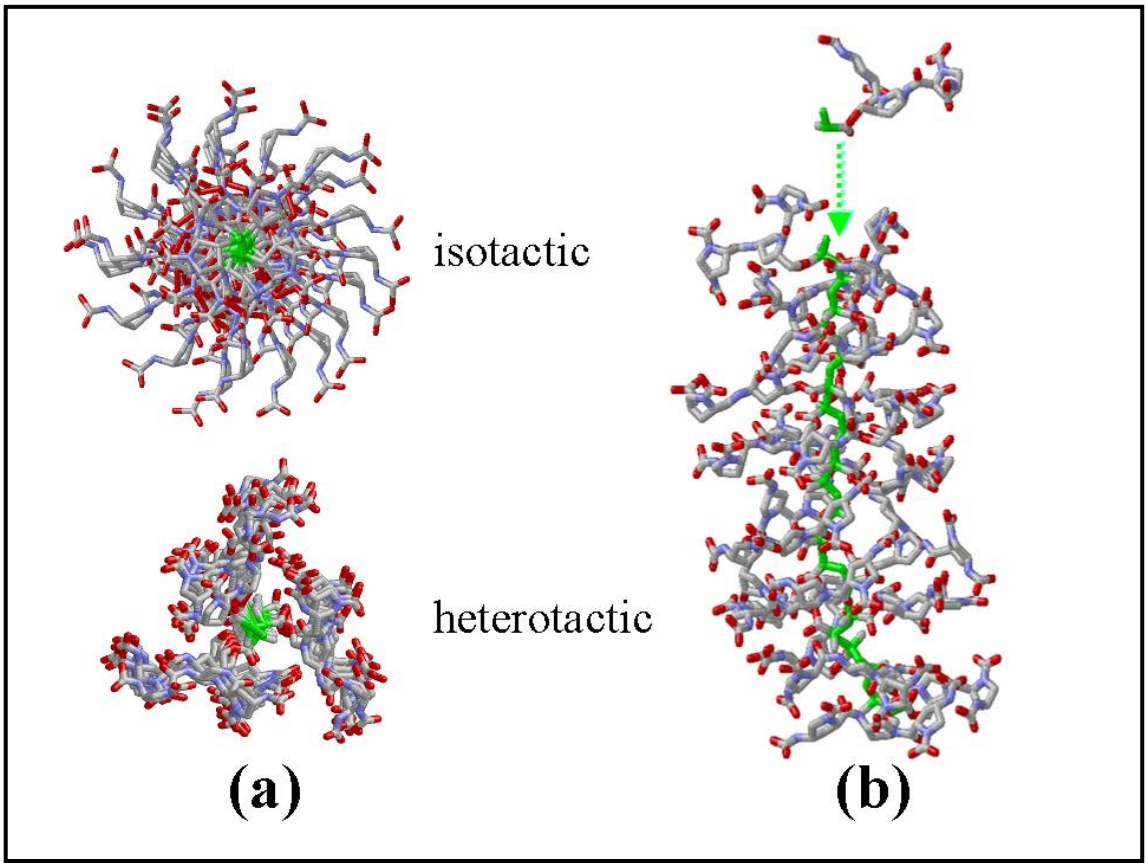


Figure 8. a) Equatorial projection of the right-handed helices proposed for isotactic and heterotactic **PG2**. b) Schematic showing helix formation of growing heterotactic **PG2** chain with incoming dendron. *tert*-Butyl groups and hydrogen atoms have been removed for clarity. The backbone is represented by a solid green line in all cases.

To gain insight into the energetic stability of the different sequences, a simple energy analysis was performed on the last 50 snapshots recorded during the simulations. Specifically, the averaged internal energy of the **PG2** helix (E_{den}) and the potential energy of the whole simulated system (E_{den+s}) were calculated, whereby the latter includes both the dendronized model molecule and the explicit solvent molecules. Relative values of E_{den} and E_{den+s} are included in Table 3. Although solute···solvent and solvent···solvent interactions alter the relative energy order derived from the analysis of the internal energy of the helix, the values of ΔE_{den} and ΔE_{den+s} lie within intervals of

only 3.7 and 5.3 kcal·mol^{-1} per repeat unit, respectively. Although such narrow intervals do not allow discrimination between the different models, they provide more evidence of both the significant stability of the helical conformation and its independence with respect to backbone tacticity.

4.1.4. *Conclusions*

Attachment of compact, chiral second-generation dendrons with a low conformational flexibility to polymethacrylates forces the dendrons into tight register and thus not only makes the polymer highly rigid but also induces a conformation with a right-handed helical sense that is stable over a wide temperature range in both MeOH (ε=33, protic) and THF (ε =6, aprotic). Molecular dynamics simulations provided a right-handed helix model that is stable for isotactic, syndiotactic, and heterotactic **PG2**. Thus, backbone tacticity mainly affects the packing of the dendrons, which is less compact for syndiotactic and heterotactic **PG2** than for the isotactic system. Although the main trends of the secondary structure are not altered, the backbone tacticity modulates the characteristic helical parameters: the helix pitch and the number of r.u.s per turn of 6.33 Å and 3.03, respectively, for the heterotactic polymer decrease to 5.62 Å and 2.65 for the isotactic polymer. This allows the conclusion that helix formation during polymerization is mainly driven by dendron / dendron interaction and not by backbone tacticity. The backbone seems to be "decoupled" from the dendritic layer. The interdendron interaction, which is supported by interdendron hydrogen bonds, is strong enough to cope with any eventual structural constraints originating from backbone stereochemistry. This result differs from the findings of Okamoto and Nakano on trityl polymethacrylates, in which backbone tacticity is essential for the polymer to attain a helical conformation.[22] In their systems helical organization is believed to lead to an extended trityl-based chromophore, whereas in the present case the chromophores appear to be more localized.

Because of the good availability of the monomer, **PG2** could, in principle, be easily prepared on a multigram scale and thus be used for further investigations such as formation of self-assembled hybrid structures with biological helices, as well as covalent decoration on the basis of selective addressability[27] of its peripheral protected primary and secondary amino groups.[28] Finally, with other diastereomers of **2a** at hand,

an obvious next task to see whether not only right-handed but also left-handed helices can be obtained with the present concept.

4.1.5. *References*

[1] a) Okamoto T., Nakano T. *Chem. Rev.* **1994**, *94*, 349. b) Nakano T., Okamoto Y. *Chem. Rev.* **2001**, *101*, 4013. c) Cornelissen J.J.L.M., Rowan A.E., Nolte R.J.M., Sommerdijk N.A.J.M. *Chem. Rev.* **2001**, *101*, 4039. d) Green M.M., Cheon K.-S., Yang S.-Y., Park J.-W., Swansburg S., Liu W. *Acc. Chem. Res.* **2001**, *34*, 672. e) Yashima E., Maeda K., Nishimura T. *Chem. Eur. J.* **2004**, *10*, 42. Rudick J.G., Percec V. *New J. Chem.* **2007**, *31*, 1083. f) Palmans A.R.A., Meijer E.W. *Angew. Chem.* **2007**, *119*, 9106. *Angew. Chem. Int. Ed.* **2007**, *46*, 8948. g) Yashima E., Maeda K. *Macromolecules* **2008**, *41*, 3.

[2] a) Sakurai S.-I., Okoshi K., Kumaki J., Yashima E. *J. Am. Chem. Soc.* **2006**, *128*, 5650. b) Zhou J.-L., Chen X.-F., Fan X.-H., Lu C.-X., Zhou Q.-F. *J. Polym. Sci., Part A Polym. Chem.* **2006**, *44*, 6047. c) Cheuk K.K.L., Lam J.W.Y., Li B.-S., Xie Y., Tang B.-Z. *Macromolecules* **2007**, *40*, 2633. d) Zhao H., Sanda F., Masuda T. *J. Polym. Sci. Part A Polym. Chem.* **2007**, *45*, 1691. e) Fujii T., Shiotsuki M., Inai Y., Sanda F., Masuda T., *Macromolecules* **2007**, *40*, 7079. f) Goh M., Kyotani M., Akagi K. *J. Am. Chem. Soc.* **2007**, *129*, 8519.

[3] a) Mayer S., Zentel R. *Prog. Polym. Sci.* **2001**, *26*, 1973. b) Green M.M., Peterson N.C., Sato T., Teramoto A., Cook R., Lifson S. *Science* **1995**, *268*, 1860. c) Green M.M., Zanella S., Gu H., Sato T., Gottarelli G., Jha S.K., Spada G.P., Schoevaars A.M., Feringa B.L., Teramoto A. *J. Am. Chem. Soc.* **1998**, *120*, 9810. d) Green M.M., Park J.W., Sato T., Teramoto A., Lifson S., Selinger R.L.B., Selinger J.V. *Angew. Chem.* **1999**, *111*, 3328. *Angew. Chem. Int. Ed.* **1999**, *38*, 3138. e) Pijper D., Feringa B.L., *Angew. Chem.* **2007**, *119*, 3767. *Angew. Chem. Int. Ed.* **2007**, *46*, 3693.

[4] a) Koe J.R., Fujiki M., Nakashima H., Motonaga M. *Chem. Commun.* **2000**, 389. b) Sanji T., Takase K., Sakurai H. *J. Am. Chem. Soc.* **2001**, *123*, 12690. c) Fujiki M. *Macromol. Rapid Commun.* **2001**, *22*, 539. d) Ohira A., Kim S.-Y., Fujiki M., Kawakami Y., Naito M., Kwak G., Saxena A. *Chem. Commun.* **2006**, 2705.

[5] a) Goodwin A., Novak B.M. *Macromolecules* **1994**, *27*, 5520. b) Tian G., Lu Y., Novak B.M. *J. Am. Chem. Soc.* **2004**, *126*, 4082. c) Tang H.-Z., Novak B.M., He J., Polavarapu P.L. *Angew. Chem.* **2005**, *117*, 7464. *Angew. Chem. Int. Ed.* **2005**, *44*, 7298.

[6] a) Nakano T., Shikisai Y., Okamoto Y. *Polym. J.* **1996**, *28*, 51. b) Nakano T., Okamoto Y. *Macromolecules* **1999**, *32*, 2391. c) C.A. Müller, Hoffart T., Holbach M., Reggelin M. *Macromolecules* **2005**, *38*, 5375.

[7] a) Cornelissen J.J.L.M., Fischer M., Sommerdijk N.A.J.M., Nolte R.J.M. *Science* **1998**, *280*, 1427. b) Cornelissen J.J.L.M., Donners J.J.J.M., de Gelder R., Graswinckel W.S., Metselaar G.A., Rowan A.E., Sommerdijk N.A.J.M., Nolte R.J.M. *Science* **2001**, *293*, 676. c) Takei F., Hayashi H., Onitsuka K., Kobayashi N., Takahashi S. *Angew. Chem.* **2001**, *113*, 4216. *Angew. Chem. Int. Ed.* **2001**, *40*, 4092. d) Cornelissen J.J.L.M., Sommerdijk N.A.J.M., Nolte R.J.M. *Macromol. Chem. Phys.* **2002**, *203*, 1625. e) Wezenberg S.J., Metselaar G.A., Rowan A.E., Cornelissen J.J.L.M., Seebach D., Nolte R.J.M. *Chem. Eur. J.* **2006**, *12*, 2778. f) Schwartz E., Kitto H.J., de Gelder R., Nolte R.J.M., Rowan A.E., Cornelissen J.J.L.M. *J. Mater. Chem.* **2007**, *17*, 1876. g) Hase Y., Mitsutsuji Y., Ishikawa M., Maeda K., Okoshi K., Yashima E. *Chem. Asian J.* **2007**, *2*, 755. h) Onouchi H., Okoshi K., Kajitani T., Sakurai S.-I., Nagai K., Kumaki J., Onitsuka K., Yashima E. *J. Am. Chem. Soc.* **2008**, *130*, 229.

[8] a) Hoshikawa N., Hotta Y., Okamoto Y. *J. Am. Chem. Soc.* **2003**, *125*, 12380. b) Azam A.K.M.F., Kamigaito M., Okamoto Y. *J. Polym. Sci. Part A, Polym. Chem.* **2007**, *45*, 1304. c) Miyake G.M., Mariott W.R., Chen E.Y.-X. *J. Am. Chem. Soc.* **2007**, *129*, 6724.

[9] a) Schlüter A.D., Rabe J.P. *Angew. Chem.* **2000**, *112*, 860. *Angew. Chem. Int. Ed.* **2000**, *39*, 864. b) Zhang A., Shu L., Bo Z., Schlüter A.D. *Macromol. Chem. Phys.* **2003**, *204*, 328. c) Zhang A. *Prog. Chem.* **2005**, *17*, 157. d) Frauenrath H. *Prog. Polym. Sci.* **2005**, *30*, 325. e) Schlüter A.D. *Top. Curr. Chem.* **2005**, *245*, 151.

[10] Zhao H., Sanda F., Masuda T. *Macromol. Chem. Phys.* **2006**, *207*, 1921.

[11] a) Percec V., Rudick J.G., Peterca M., Wagner M., Obata M., Mitchell C.M., Cho W.-D., Balagurusamy V.S.K., Heiney P.A. *J. Am. Chem. Soc.* **2005**, *127*, 15257. b) Percec V., Rudick J.G., Peterca M., Staley S.R., Wagner M., Obata M., Mitchell C.M., Cho W.-D., Balagurusamy V.S.K., Lowe J.N., Glodde M., Weichold O., Chung K.J., Ghionni N., Magonov S.N., Heiney P.A. *Chem. Eur. J.* **2006**, *12*, 5731. c) Percec V., Aqad E., Peterca M., Rudick J.G., Lemon L., Ronda J.C., De B.B., Heiney P.A., Meijer E.W. *J. Am. Chem. Soc.* **2006**, *128*, 16365.

[12] Zhang W., Shiotsuki M., Masuda T. *Macromol. Rapid Commun.* **2007**, *28*, 1115.

[13] Suzuki Y., Shiotsuki M., Sanda F., Masuda T. *Macromolecules* **2007**, *40*, 1864.

[14] A simple measure of the compactness of dendrons is the number of atoms between the branching points, which for PG2 are denoted as BP in Figuree 1. Note that for nonsymmetric dendrons the branches can consist of many different atoms. For the

dendronized polymer PG2 the numbers of atoms between BPs are 2,4 (see Figure 1). 3,3: a) Grayson S.M., Fréchet J.M.J. *Macromolecules* **2001**, *34*, 6542. 3,3: b) Malkoch M., Carlmark A., Woldegiorgis A., Hult A., Malmström E.E. *Macromolecules* **2004**, *37*, 322. 7,7: c) Zhang A., Okrasa L., Pakula T., Schlüter A.D. *J. Am. Chem. Soc.* **2004**, *126*, 6658. 8,8: d) Kasëmi E., Zhuang W., Rabe J.P., Fischer K., Schmidt M., Colussi M., Keul H., Yi D., Cölfen H., Schlüter A.D. *J. Am. Chem. Soc.* **2006**, *128*, 5091.

[15] Zhang A., Schlüter A.D. *Chem. Asian J.* **2007**, *2*, 1540.

[16] High purity is a prerequisite for spontaneous polymerization. It is reasonable to assume that the polymerization is initiated by some inadvertently formed radical: Zhang A., Zhang B., Wächtersbach E., Schmidt M., Schlüter A.D. *Chem. Eur. J.* **2003**, *9*, 6083.

[17] For peptidic dendronized polymers, see: a) Draheim G., Ritter H. *Macromol. Chem. Phys.* **1995**, *196*, 2211. b) Niggemann M., Ritter H. *Acta Polym.* **1996**, *47*, 351. c) Niggemann M., Ritter H. *J. Macromol. Sci. Pure Appl. Chem.* **1997**, *A34*, 1325. d) Bilibin A., Zorin I., Saratovsky S., Moukhina I., Egorova G., Girbasova N. *Macromol. Symp.* **2003**, *199*, 197. e) Girbasova N., Aseyev V., Saratovsky S., Moukhina I., Tenhu H., Bilibin A. *Macromol. Chem. Phys.* **2003**, *204,* 2258. f) Zhuravel M.A., Davis N.E., Nguyen S.T., Koltover I. *J. Am. Chem. Soc.* **2004**, *126*, 9882. g) Luebbert A., Nguyen T.Q., Sun F., Sheiko S.S., Klok H.-A. *Macromolecules* **2005**, *38*, 2064. h) Lee C., J. Frechet J.M. *Macromolecules* **2006**, *39*, 476.

[18] The NMR spectrum of PG2 in [D_6]DMF at 25ºC was very similar to that in $CDCl_3$, while a rather well-resolved spectrum was obtained in [D_6]DMF at 80ºC.

[19] Hatada K., Kitayama T., Ute K. *Prog. Polym. Sci.* **1988**, *13*, 189.

[20] Al-Hellani R., Schlüter A.D. *Helv. Chim. Acta* **2006**, *89*, 2745.

[21] Jansen J.F.G.A., Peerlings H.W.I., de Brabander-van den Berg E.M.M., Meijer E.W. *Angew. Chem.* **1995**, *107*, 1321. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1206.

[22] From the work of Okamoto and Nakano on trityl polymethacrylates with extended chromophores associated with the trityl groups, it is known that polymethacrylates must have high backbone tacticity in order to adopt a helical conformation: a) Nakano T., Mori M., Okamoto Y. *Macromolecules* **1993**, *26*, 867. b) Nakano T., Okamoto Y. *Macromol. Rapid Commun.* **2000**, *21*, 603.

[23] Binkley J.S., Pople J.A., Hehre W.J. *J. Am. Chem. Soc.* **1980**, *102*, 939.

[24] Hariharan P.C., Pople J.A. *Theor. Chim. Acta* **1973**, *28*, 213.

[25] a) Becke A.D. *J. Chem. Phys.* **1993**, *98*, 1372. b) Lee C.T., Yang W., Parr R.G. *Phys. Rev. B* **1988**, *37*, 785.

[26] Cornell W.D., Cieplak P., Bayly C.I., Gould I.R., Merz K.M., Ferguson D.M., Spellmeyer D.C., Fox T., Caldwell J.W., Kollman P.A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

[27] Al-Hellani R., Schlüter A.D. *Macromolecules* **2006**, *39*, 8943.

[28] a) Nigam S.C., Mann A., Taddei M., Wermuth C.-G. *Synth. Commun.* **1989**, *19*, 3139. b) Ravinder K., Reddy A.V., Mahesh K.C., Narasimhulu M., Venkateswarlu Y. *Synth. Commun.* **2007**, *37*, 281.

[29] Frisch M.J. et al. *Gaussian 03,* Revision B.02, Gaussian, Inc.: Pittsburgh PA, **2003**.

[30] Phillips J.C., Braun R., Wang W., Gumbart J., Tajkhorshid E., Villa E., Chipot C., Skeel R.D., Kale L., Schulten K. *J. Comput. Chem.* **2005**, *26*, 1781.

[31] Jorgensen W.L., Chandrasekhar J., Madura J.D., Impey R.W., Klein M.L. *J. Chem. Phys.* **1983**, *79*, 926.

[32] Caldwell J.W., Kollman P.A. *J. Phys. Chem.* **1995**, *99*, 6208.

[33] Alemán C., Luque F.J., Orozco M. *J. Comput. Aided Mol. Design* **1993**, *7*, 721.

[34] Darden T., York D., Pedersen L. *J. Chem. Phys.* **1993**, *98*, 10089.

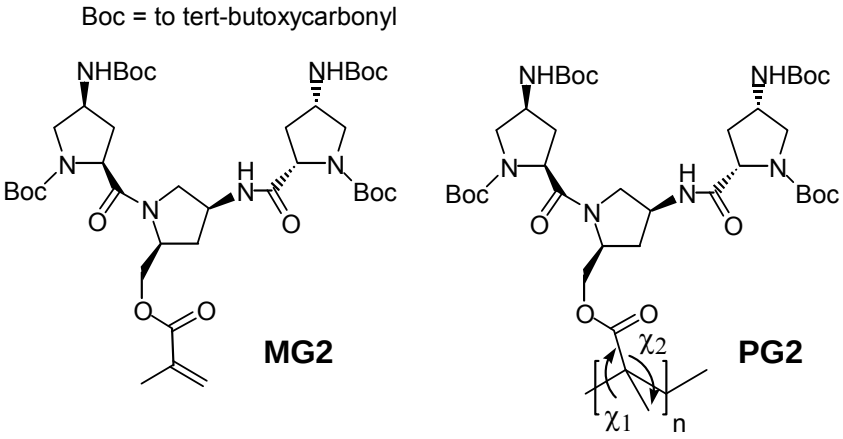
[35] Ryckaert J.P., Ciccotti G., Berendsen H.J.C. *J. Comput. Phys.* **1977**, *23*, 327.

[36] Berendsen H.J.C., Postma J.P.M., van Gunsteren W.F., Dinola A., Haak J.R. *J. Chem. Phys.* **1984**, *81*, 3684.

4.2. Helical dendronized polymers carrying chiral 4-aminoproline second-generation dendrons: Atomistic view and driving forces*

4.2.1. *Introduction*

Among novel chiral materials, helical polymers deserve special attention because of their ability to adopt secondary structures that approach those typically found in biomacromolecules.[1] Dendronized polymers, a subclass of comb-like polymers that consist of a polymeric backbone with dendrons attached to all the repeat units,[2] are particularly interesting because the size of the side chains can impart a significant degree of rigidity to the secondary structure. Helical dendronized polymers with polyacetylene backbones were recently reported by Masuda[3] and Percec;[4] their molecular conformations were found to have limited thermal stability. A helical dendronized polymer that was obtained from the spontaneous polymerization of a chiral 4-aminoproline-based second generation macromonomer, **MG2** (Scheme 1) is reported in the previous chapter.[5] It is worth noting that the compact nature of the proline-like dendrons makes the resulting dendronized polymer PG2 especially rigid. Such polymers are of interest for applications including organo catalysis, supramolecular constructions and chiral separations.



Scheme 1

Computer atomistic simulations based on Molecular Dynamics (MD) allowed propose a detailed right-handed helical model for **PG2** (Figure 1).[5] This secondary structure, which is stabilized by interdendron and intradendron hydrogen bonding networks, is stable irrespective of whether PG2's backbone is isotactic, syndiotactic or heterotactic.

* Results presented in this chapter have been accepted for publication in *J. Phys. Chem. B* **2009**.

The backbone tacticity only affects the packing of the dendrons, thus, modulating the structural parameters of the helix, *i.e.* the helix pitch and the number of repeat units per turn are 6.33 Å and 3.03, respectively, for the heterotactic polymer, and 5.62 Å and 2.65 for the isotactic one. Furthermore, spectroscopic studies evidenced that the helix is stable in a wide range of temperature and in solvents with very different polarities,[5] *e.g.* methanol (ε= 33, protic) and tetrahydrofuran (ε= 6, aprotic).

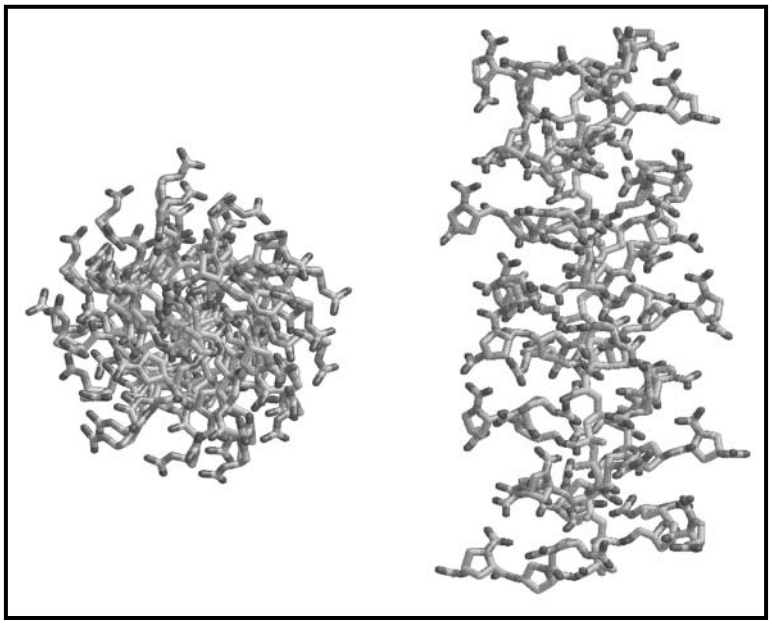
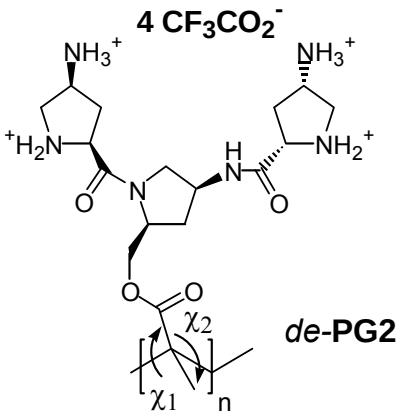


Figure 1. Equatorial (left) and axial (right) projections of the right-handed helix proposed for **PG2**. *tert*-Butyl groups and hydrogen atoms have been removed for clarity.

On the other hand, the deprotected polymer (*de*-**PG2**), which was obtained by adding trifluoroacetic acid to a solution of **PG2,** was observed to present also a helical conformation in polar solvents like methanol and water.[5] This was a striking feature because this polymer presents a high concentration of positive charges (4 per repeat unit) located in the dendritic side groups (Scheme 2). In spite of *de*-**PG2** being able to retain an ordered secondary structure, circular dichroism (CD) and UV spectroscopic measurements evidenced a more pronounced temperature dependence as compared with the protected **PG2**. Preliminary MD simulations on an isotactic model indicated that *de*-**PG2** is able to retain the right-handed helical structure proposed for **PG2**,[5] although no systematic study about its preferences has been presented yet.



Scheme 2

In this work atomistic MD simulations are used to examine the response of the right-handed helical conformation proposed for **PG2** against temperature and size of the molecular model (number of repeat units), two essential factors that were not considered before. It is worth noting that optical rotation and CD spectroscopic measurements evidenced that **PG2** exhibits an ordered secondary structure in the temperature range 293-333 K.[5] Specifically, the thermal stability of the proposed helix has been microscopically examined considering simulations at 9 different temperatures, while the effects of the chain length on the structural parameters of this conformation have been analyzed at room temperature using models constituted by 15, 20 and 60 repeat units. After this, the development of a detailed atomistic model for the helix of *de*-**PG2** is presented. This has been achieved by applying a systematic conformational search that combines quantum mechanical calculations, force-field energy minimizations and MD trajectories on an isotactic polymer chain. Next, the influence of the solvent, temperature and backbone tacticity on the stability of the helix proposed for the deprotected polymer has been studied. Finally, a comparison between the helical models proposed for **PG2** and *de*-**PG2** is presented.

4.2.2. *Methods*

All atomistic MD simulations were performed using the NAMD program.[6] Simulations of **PG2** were performed in chloroform solution considering isotactic models formed by 15, 20 and 60 repeat units that were blocked at the ends with methyl groups, whereas simulations of *de*-**PG2** were carried out in both methanol and aqueous solutions using a molecular model with 20 repeat units. For *NVT*-MD, the simulated systems were placed

in the center of an orthorhombic simulation box filled with explicit solvent molecules. Chloroform, methanol and water were described using the four-particles OPLS,[7] the Caldwell and Kollman[8] and TIP3P[9] models, respectively. The dimensions of the orthorhombic box for simulations of **PG2** with models formed by 15, 20 and 60 repeat units were (71.34 × 71.34 × 83.64) $Å^3$, (99.31 × 99.31 × 115.00) $Å^3$ and (99.31 × 99.31 × 251.32) $Å^3$, respectively, while the total number of particles considered explicitly was 14391 (3142 chloroform molecules), 35496 (8267 chloroform molecules) and 80784 (18379 chloroform molecules). For the **PG2** model with 20 repeat units, the influence of the temperature on the stability of the right-handed helical conformation was studied performing simulations at T= 263, 273, 283, 293, 313, 333, 373, 473 and 773 K. On the other hand, all the simulations on *de*-**PG2** were performed using 80 CF_3COO^- explicit molecules to neutralize the positive charges introduced by the 20 repeat units. The number of explicit particles in methanol and aqueous solution was 82214 (13391 methanol molecules) and 101909 (33347 water molecules), respectively, while the dimensions of the simulation box were (93.14 × 93.14 × 107.84) $Å^3$ and (95.67 × 95.67 × 109.31) $Å^3$.

Energy was calculated using the AMBER force-field potential,[10] with all the required parameters for **PG2** and *de*-**PG2** transferred from the AMBER libraries with exception of the electrostatic ones. Electrostatic charges centered on the atoms were explicitly developed by fitting the quantum mechanics molecular electrostatic potential, which was calculated at the HF/6-31G(d) level for a model system constituted by 3 chemical repeat units, to the classical one.[11] Atom pair distance cutoffs were applied at 16.0 Å to compute the van der Waals interactions. However, to avoid discontinuities in the potential energy function, Van der Waals energy function was forced to slowly converge to 0, applying a smoothing factor from a distance of 12.0 Å. The electrostatic interactions were evaluated by computing Ewald Summations using the Particle Mesh Ewald method (PME).[12] The real space term was determined by the van der Waals cut off (16.0 Å). Bond lengths involving hydrogen atoms were constrained using the *SHAKE* algorithm[13] with a numerical integration step of 2 fs.

Before the MD production series, the thermodynamic variables of the system were equilibrated. The energy of each system was initially minimized to relax conformational and structural tensions using the conjugate gradient method for $5 \cdot 10^3$ steps. Next, the dendronized polymer was placed in a previously equilibrated solvent box. Different

consecutive rounds of short MD runs were performed in order to equilibrate the variable thermodynamic magnitudes. First, solvent was thermally relaxed, while the dendronized polymer was kept frozen: 0.2 ns of *NVT*-MD at 500 K were used to homogeneously distribute the solvent in the box. Second, 0.15 ns of isothermal relaxation at 298 K were run. Finally, all the atoms of the system were submitted to 0.25 ns of *NVT*-MD at the working temperature (thermal equilibration). Temperature was controlled by the weak coupling method, the Berendsen thermostat,[14] using a time constant for heat bath coupling of 1 ps. The end of the thermal relaxation simulation was the starting point of the molecular simulations presented in this work. The coordinates of all the production runs, which were from 6 and 50 ns long, were saved every 500 steps (1 ps intervals) for subsequent analysis.

Quantum mechanical calculations were carried out using the unrestricted formalism of both the Hartree-Fock method (UHF) and Density Functional Theory method based on the Becke's three-parameter hybrid functional[15] with the Lee, Yang and Parr[16] expression for the non-local correlation (UB3LYP), the 6-31G(d) basis set being used in all cases.[17] Specifically, geometry optimizations were performed at the UHF/6-31G(d) level, while single point energy calculations at the UB3LYP/6-31G(d) on such optimized geometries were carried to re-evaluate the energy. The Gaussian 03 program suite[18] was used for all the quantum mechanical calculations.

4.2.3. *Results and discussion*

Thermal stability of the ***PG2*** *helix.*

Figures 2a and 2b show the time-dependent evolution of the cross-sectional diameter (D_{cs}) and the end-to-end distance ($R_{e\text{-}e}$), respectively, obtained in chloroform solution when the temperature ranges from 263 to 773 K. As can be seen, D_{cs} values remain around 29-30 Å in the whole range of temperatures, the variation associated of the standard deviations obtained for $\overline{D}_{cs}$ being very small, *i.e.* from 0.18 (T= 263 K) and 0.44 Å (T= 773 K). The $\overline{R}_{e-e}$ ranges from 45.7 (T 773 K) to 48.59 Å (T= 313 K), although significant fluctuations were obtained in the trajectory obtained at the highest temperature (Figure 2b).

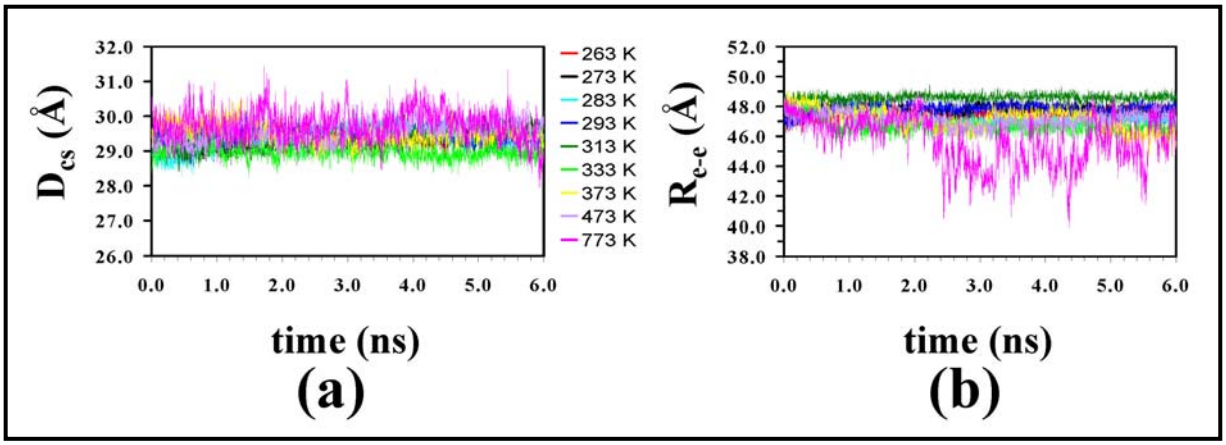


Figure 2. Time-dependent evolution of (a) the cross sectional diameter (D_{cs}) and (b) the end-to-end distance (R_{e-e}) for the right-handed helix of **PG2** at nine simulated temperatures.

Inspection of the temporal evolution of the backbone root mean square deviation (RMSD) for the nine simulated temperatures revealed that the conformational distortions underwent by the helix are notably low in all cases. Thus, the averaged RMSD was lower than 1.4 Å for temperatures ranging from 263 K to 473 K, while it was 1.980±0.445 Å at 773 K. On the other hand, the backbone root-mean square fluctuation (RMSF) of the individual repeat units averaged over the whole trajectory for the nine simulated temperatures is displayed in Figure 3. No significant fluctuation was found for the 15 central repeat units, the main distortions being located at the two extremes of the molecule, *i.e.* effects due to the limited size of the model. Results displayed in Figures 2 and 3 clearly support the remarkable thermal stability of the secondary structure proposed for **PG2**. This behavior is fully consistent with the inter-dendron hydrogen bonding network found in this helical conformation, which was also preserved in syndiotactic and heterotactic systems.[5]

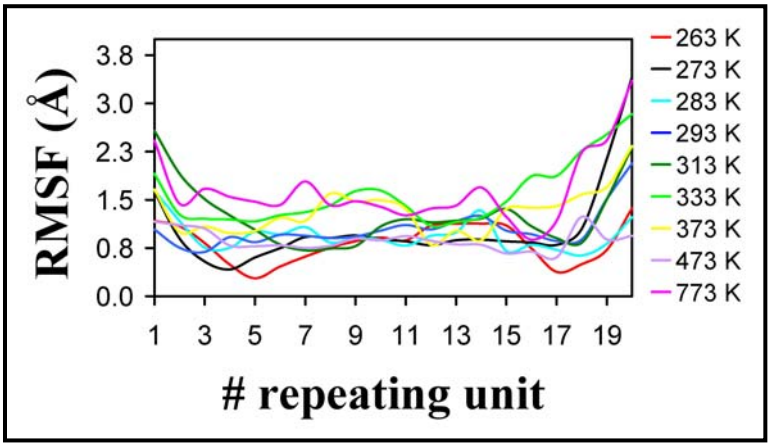


Figure 3. Root mean square fluctuation (RMSF) of the individual 20 repeat units considered in the modeling of the right-handed helix of PG2 at the nine simulated temperatures. Results have been derived from 6 ns MD trajectories.

It is well known that in solution helical structures adopted by conventional polypeptides[19] and synthetic polymers[20] present small elongation and compression movements along the helix axis, which are caused by the breathing of the weak non-bonded intramolecular interactions and usually involve the interface of rigidly bound subgroups, *i.e.* the turns that form this secondary structure. Thus, the helix behaves as an ideal spring, the degree of flexibility associated to such conformational movement being defined by the spring constant k_h. In order to examine if the helix adopted by **PG2** behaves like a secondary structure of conventional polypeptides or is rather to be considered shape-persistent without any flexibility along the helical axis, the variation of k_h with the temperature has been estimated considering that it is inversely proportional to the second moment of the $R_{e\text{-}e}$ distribution:[21]

$$k_h \propto \frac{1}{\overline{R_{e-e}^2} - \overline{R}_{e-e}^2} \qquad (5.1.)$$

where $\overline{R_{e-e}^2}$ refers to the average of R_{e-e}^2. Figure 4a shows the variation of $\left(\overline{R_{e-e}^2} - \overline{R}_{e-e}^2\right)^{-1}$ as a function of the temperature. It is worth noting that in all cases the inverse second moment is small, decreasing to almost negligible values for T ≥ 333 K. The small variation of the effective stiffness (shape-persistence) against the temperature clearly indicates that the behavior of **PG2** is close to that of a helical rigid rod.

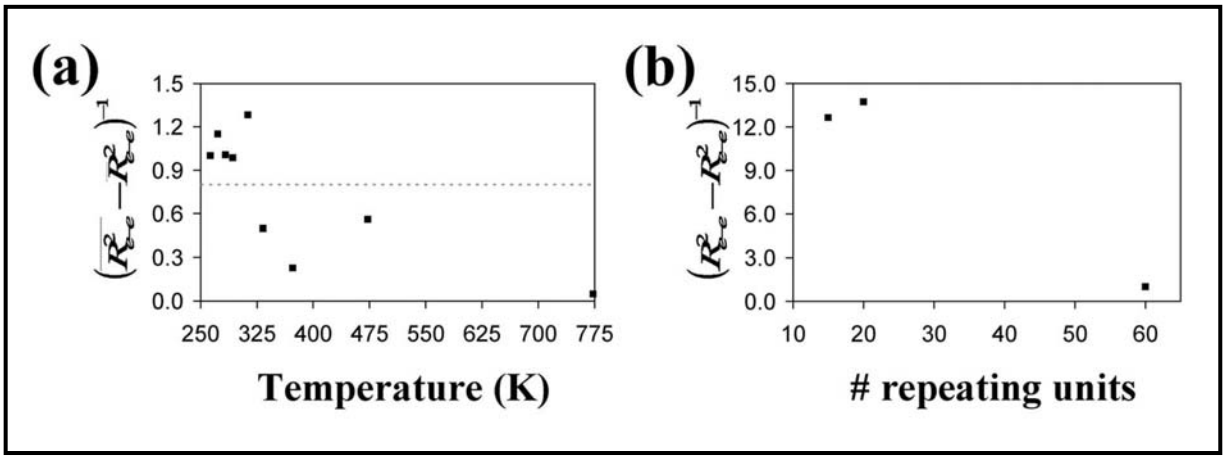


Figure 4. Temperature (a) and size (b) dependence of the inverse second moments of the $R_{e\text{-}e}$ distribution for PG2, scaled by the inverse second moment obtained at T= 263 K and 60 repeat units, respectively. The gray line in (a) reflects the different behavior displayed at T < 333 K and T ≥ 333 K. Results presented in (a) and (b) have been derived from 6 ns and 50 ns trajectories, respectively.

*Structural parameters of the **PG2** helix.*

Figure 5 shows the temporal evolution of D_{cs} and $R_{e\text{-}e}$ obtained in chloroform solution at 298 K for chains of **PG2** involving different numbers of repeat units. The values of $\overline{D}_{cs}$ are 29.76±0.15, 29.16±0.30 and 29.16±0.43 Å for the systems containing 60, 20 and 15 repeat units, respectively. It is worth noting that, even though the D_{cs} is practically independent of the molecular length, the packing of the dendrons is slightly denser for the chains formed by 20 and 60 repeat units than for those involving only 15 repeat units. Thus, the reduction detected in both the amplitude of the fluctuations during the simulation time range of 50 ns and the standard deviation of the averages is consistent with an enhancement of the rigidity in the direction perpendicular to the axis of the helix. In contrast, the largest fluctuations of $R_{e\text{-}e}$ were obtained for the system containing 60 repeat units. This is because the impact of the conformational defects detected at specific dendron units, whose frequency increases with the molecular size, is larger in the structural parameters that involve the whole axis of the helix than in those related with the projection of the cross-section. This is also reflected in Figure 4b, which shows the variation of $\left(\overline{R_{e-e}^2} - \overline{R}_{e-e}^2\right)^{-1}$ against the number of repeat units. Thus, the lowest value of the inverse second moment was obtained for the largest chain. Both the temporal evolution of the RMSD and the RMSF obtained for the systems described using 60, 20 and 15 repeat units confirm this evidence. On the other hand, the consistency of the $\overline{R}_{e-e}$ values (138.23±1.03, 46.83±0.26 and 33.75±0.35 Å for systems containing 60, 20 and 15 repeat units, respectively) also point towards a significant

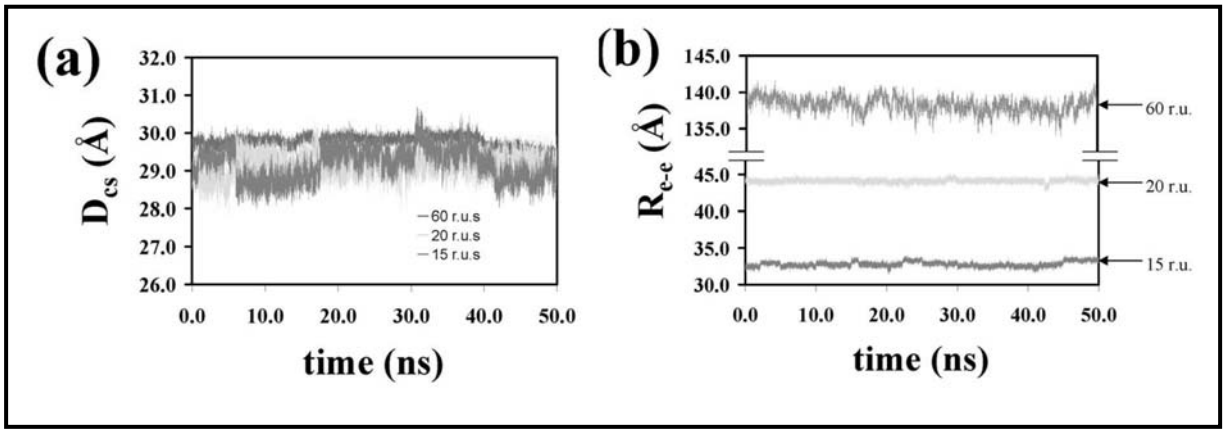


Figure 5. Time-dependent evolution of (a) the cross sectional diameter (D_{cs}) and (b) the end-to-end distance ($R_{e\text{-}e}$) for the right-handed helix of PG2 when simulations involve molecular chains formed by 60, 20 and 15 repeat units.

The rigidity of **PG2** has been also examined by calculating the persistence length (L_p), a very useful microscopic parameter that indicates the distance over which the direction of the chain persists characterizing polymer chain flexibility/stiffness. This parameter is defined as:[22]

$$L_p = \frac{R_{e-e}^2}{2L_c} + \frac{b}{2} \tag{5.2.}$$

where L_c is the contour length and b is the average bond length, *i.e.* $b = L_c/N_b$ where N_b is the number of bonds. The latter expression predicts that for an idealized extended conformation: $R_{e\text{-}e}=L_c$ and $L_p \approx L_c/2$. In this work L_p has been calculated applying the expression developed by Hsiao[23] to the coordinates of the snapshots recorded from the MD trajectories:

$$L_p = \frac{1}{2b} \sum_{i=0}^{(N/2)-1} \left\langle \mathbf{b}_{N/2} \cdot \mathbf{b}_{(N/2)-i} + \mathbf{b}_{N/2} \cdot \mathbf{b}_{(N/2)+i} \right\rangle \tag{5.3.}$$

where N is the number of bonds and $\mathbf{b}_{N/2}$ is the length of the central bond. The values of L_p determined using the MD simulations of the system formed by 20 and 60 repeating units are 14.39 and 42.25 Å, which are about 20% smaller than those predicted by Eqn (1), *i.e.* 18.98 and 52.11 Å, respectively. On the other hand, it is well known that the ratio $\left(\frac{R_{e-e}}{R_g}\right)^2$, where R_g is the radius of gyration, is 12 for rigid macromolecules.[24] The values derived for R_g and $\left(\frac{R_{e-e}}{R_g}\right)^2$ using MD coordinates are 11.37 Å (R_g= 14.35 Å) and 11.44 (R_g= 40.86 Å) for the models containing 60 and 20 repeating units, respectively. These results, which also reflect the rigid-rod like character of **PG2**, are fully consistent with those reported in the previous section. Specifically, evaluation of $\left(\frac{R_{e-e}}{R_g}\right)^2$ using the MD trajectories performed at different temperatures indicates a very narrow interval of variation, *i.e. from* 11.10 (263 K) to 10.64 (773 K).

*Atomistic model for the helical conformation of de-**PG2**.*

The procedure used to model the helix of *de*-**PG2**, which is similar to that previously used for **PG2**, is schematically summarized in Figure 6. Initially, the most stable arrangements of the dendritic side group were explored by considering that each of the flexible dihedral angles $\{\psi_0, \varphi_1, \psi_1\}$ of a model molecule, denoted *de*-**m2** in Figure 6, is expected to have three minima. Accordingly, the 27 minima that may be anticipated for the potential energy hypersurface $E=E(\{\psi_0, \varphi_1, \psi_1)$ were employed as starting points for complete geometry optimizations at the UHF/6-31G(d) level. After this, the energy of all the resulting structures was re-evaluated by carrying out single point calculations at the UB3LYP/6-31G(d) level.

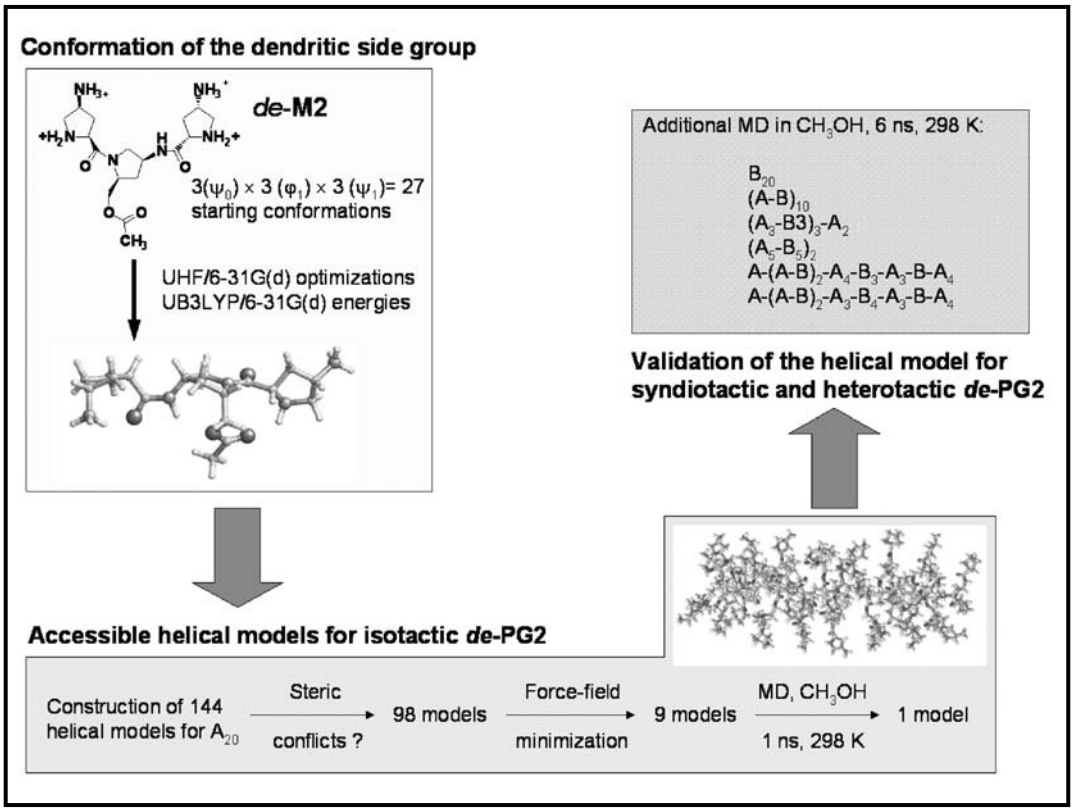


Figure 6. Procedure used to model the helix of de-PG2. This involves three different and well-defined stages: (i) quantum mechanical calculations on *de*-**M2** to investigation the conformation of the dendritic side group; (ii) systematic conformational search using molecular mechanics and molecular dynamics force-field simulation to ascertain the most stable helix for isotactic *de*-**PG2** using a model with 20 repeat unit; and (iii) molecular dynamics simulations to check that the stability of the helix found in the previous stage is maintained in syndiotactic and heterotactic *de*-**PG2.**

The optimized dihedral angles of the most stable conformation, which is displayed in Figure 6, are ψ_0= -110.8º, φ_1= 123.7º and ψ_1= 139.9º. As can be seen, it corresponds to a pseudo-extended arrangement with two of the three 4-aminoproline groups located in

a same plane, *i.e.* those connected through the carbonyl group. This conformation allows minimize the repulsive electrostatic interactions between the positive charges. The other 10 optimized structures, *i.e.* many of the starting points converged to the same conformation after geometry optimization, resulted disfavored by more than 2.7 kcal·mol^{-1} with respect to the lowest energy one. This represents a significant change with respect to the conformation found in the previous chapter for the corresponding protected model molecule,[5] which showed ψ_0= 163.1º, φ_1= -62.4º and ψ_1= 137.2º. According to these results, the packing of the dendrons is expected to be different in the helical conformations of *de*-**PG2** and **PG2**. Figure 7, which displays the superposition of the model molecules used to mimic the side group of the two systems, evidences these differences.

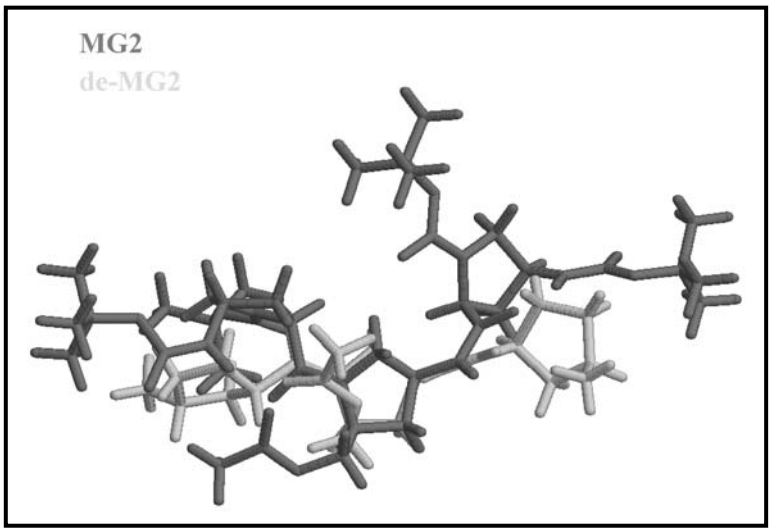


Figure 7. Superposition of the lowest energy minimum found by quantum mechanical calculations for the model molecules used to describe the side group of *de*-**PG2** (green) and **PG2** (red).

The lowest energy conformation of the dendritic side chain was used in a systematic search of the sterically and electrostatically accessible helical structures of *de*-**PG2** (Figure 6). For this purpose, an isotactic molecular model constituted by 20 dendronized methacrylate repeat units, hereafter denoted A_{20}, was built, the backbone dihedral angles of all units (χ_1 and χ_2 in Scheme 2) being varied simultaneously in steps of 30º, *i.e.* the conformation was kept identical in all the repeat units. Structures without apparent steric conflicts were subject to energy minimization with the AMBER force-field.[10] Interestingly, many of these initial helical models were not able to retain a regular and homogeneous conformation due to the strong electrostatic repulsions among the positive charges of neighboring dendritic side groups. Indeed, only 9 helical

conformations showed repeat units with similar values for the dihedral angles χ_1 and χ_2, *i.e.* the maximum allowed differences were ±30º. This represents a drastic reduction with respect to the 80 helical conformations retained for **PG2** after energy minimization evidencing that, as expected, the effects of the electrostatic repulsions (long-range interactions) are more pronounced than those of the steric ones (short-range interactions). Next, the 9 helical structures were submitted to 1 ns of MD at 298 K in methanol solution. Only one model maintained the secondary structure, the helical conformation of the remaining 8 models being rapidly lost. In these models, electrostatic repulsions appeared when atoms were displaced from their minimum energy positions by thermal vibrations at room temperature.

In order to characterize the structural parameters of the regular and homogeneous helical conformation found for isotactic *de*-**PG2**, 10 ns *NVT* MD simulations were conducted in methanol and aqueous solutions at T= 298 K. Results were similar in both environments indicating that the helix is stable in these two polar solvents, which is in good agreement with experimental data.[5] The secondary structure consists of a right-handed helix with average backbone dihedral angles {χ_1= -166.9±1.9º, χ_2= -49.3±1.8º} and {χ_1= -164.8±3.2º, χ_2= -51.4±4.1º} in methanol and aqueous solutions, respectively. The characteristic helical parameters derived for this structure, which have been determined considering idealized structures formed by identical repeat units, are 2.66/2.68 repeat units per turn, 0.77/0.73 Å internal radius and 2.09/2.10 Å rise per repeat unit in methanol/aqueous solution. A comparison with the structural parameters of the helix found for **PG2** in chloroform and methanol solutions is provided in Table 1, revealing that the arrangement of the backbone is very similar in the two polymers. On the other hand, the macromolecular parameters calculated for *de*-**PG2** (L_p= 13.76 Å, R_g= 13.72 Å and $\left(\frac{R_{e-e}}{R_g}\right)^2$ = 11.37) were very similar to those obtained for a chain of **PG2** containing the same number of repeating units evidencing similar stiffness at room temperature.

Table 1. Structural parameters of the helices found for isotactic models of *de*-**PG2** and **PG2**.

	de-**PG2**		**PG2**	
	Water	Methanol	Methanol	Chloroform
χ_1 (°)	-166.9±1.9	-164.8±3.2	-162.1±1.7	-159.5±2.8
χ_2 (°)	-49.3±1.8	-51.4±4.1	-51.6±3.4	-52.2±2.1
Rise per repeat unit (Å)	2.10	2.09	2.11	2.12
Internal diameter (Å)	1.46	1.54	1.45	1.42
Repeat units per turn	2.68	2.66	2.66	2.65

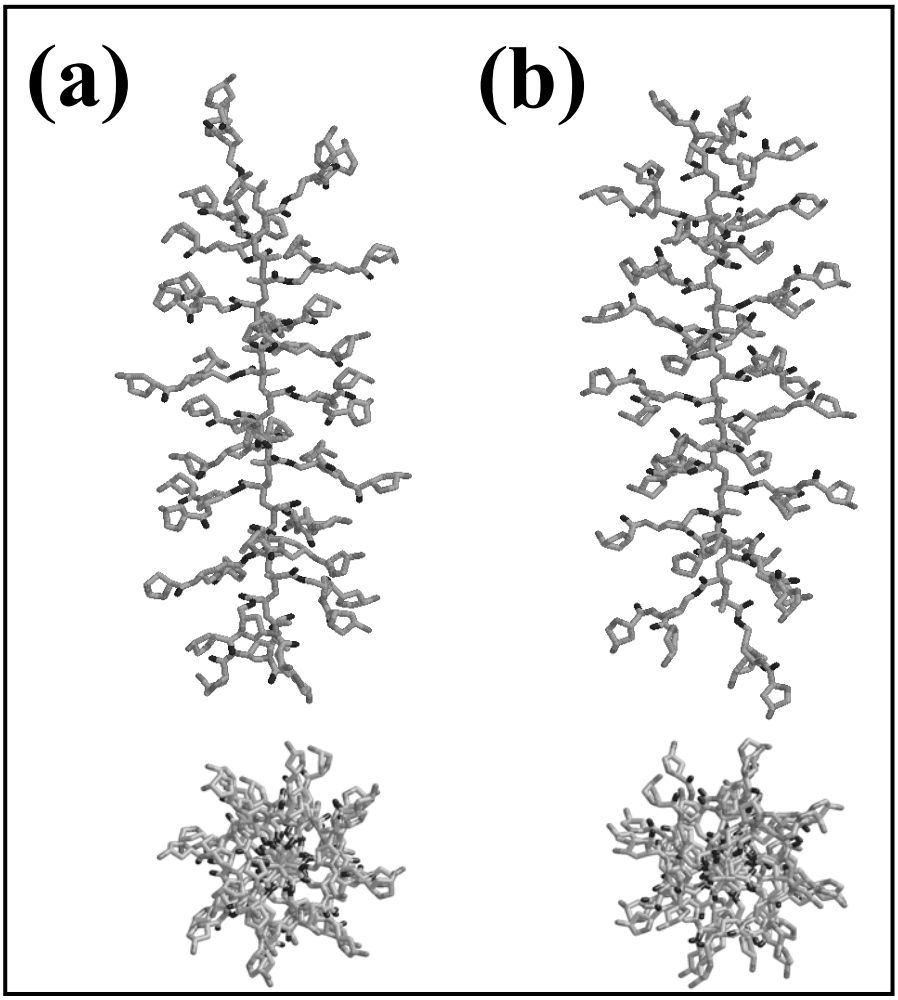


Figure 8. Axial (top) and equatorial (bottom) projections of the right-handed helix proposed for *de*-**PG2** in methanol (a) and aqueous (b) solutions. Hydrogen atoms have been removed for clarity.

Figure 8 compares the helical models obtained for *de*-**PG2** in methanol and aqueous solution. As can be been, these two structures are similar even in the spatial orientation of the dendronized side groups, although this is considerably different from that found for the protected polymer. A detailed analysis of the structures displayed in Figure 8

reveals the lack of interdendron hydrogen bonds, which represents a significant difference with respect to the conformation found for **PG2**. In *de*-**PG2** dendrons are forced in a tight register within the helix imparting rigidity to the backbone, even though the lack of attractive interdendron interactions does not preclude the thermal stability of the secondary structure. Although *NVT* MD simulations on *de*-**PG2** at high temperatures developed using a heating rate of 4 K/ps led the helix to unfold within a few ns, simulations conducted using a less aggressive heating rate (0.25 K/ps) evidenced that the helix is retained at high temperatures when the kinetic energy is allowed to be distributed without perturb the delicate electrostatic equilibrium. The temperature dependence of the helix proposed for *de*-**PG2** is displayed in Figure 9a, which compares the RMSD found at 298, 498 and 698 K. The interdendron repulsive electrostatic interactions, which are minimized at room temperature, remain very small at 498 K, the averaged RMSD being 1.647±0.324 and 1.685±0.391 Å, respectively. This structural parameter increases to 3.169±1.078 Å at 698 K evidencing that, although the thermal fluctuations are significantly enhanced, they are not strong enough to disrupt the regular secondary structure.

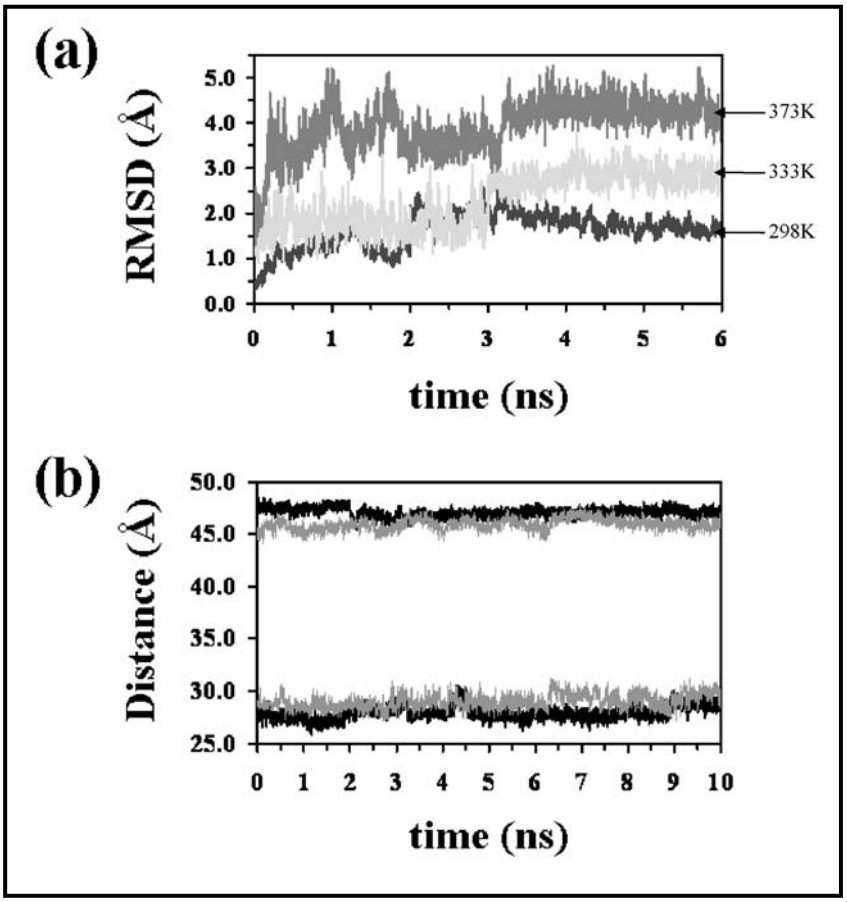


Figure 9. For the helix conformation of isotactic *de*-**PG2**: time dependent evolution of (a) the backbone RMSD in methanol solution at 298, 333 and 373 K, and (b) the end-to-end distance (upper data set) and the cross sectional diameter (lower data set) in methanol (black lines) and aqueous (gray lines).

The rigidity of the helical model proposed for *de*-**PG2** is reflected in Figure 9b, which shows the temporal evolution of D_{cs} and $R_{e\text{-}e}$ along 10 ns MD simulations in explicit methanol and aqueous solutions. As can be seen, these two structural parameters remained practically unchanged during the whole simulation period, the averaged values being $\overline{D}_{cs}$ = 27.93±0.68/29.02±0.69 Å and $\overline{R}_{e-e}$ = 47.05±0.44/45.91±0.54 Å in methanol/aqueous solution. The lack of significant fluctuations in these structural parameters, which are similar to those reported for **PG2** in chloroform ($\overline{D}_{cs}$ = 27.93±0.68 Å and $\overline{R}_{e-e}$ = 46.83±0.26 Å) and methanol ($\overline{D}_{cs}$ =28.78±0.55 Å and $\overline{R}_{e-e}$ = 46.32±0.45 Å),[5] indicate that the secondary structure of *de*-**PG2** also presents a remarkable stiffness, independently of the solvent.

*Influence of the backbone tacticity on the helix model of de-**PG2***.

In order to ascertain the effect of the backbone stereochemistry on the stability of the proposed helix, additional *NVT* MD simulations in methanol solution were performed at T= 298 K considering isotactic, syndiotactic and heterotactic model systems formed by 20 repeat units (Figure 6). Specifically, the sequences of the seven systems investigated are: A_{20}, B_{20}, $(A\text{-}B)_{10}$, $(A_3\text{-}B_3)_3\text{-}A_2$, $(A_5\text{-}B_5)_2$, $A\text{-}(A\text{-}B)_2\text{-}A_4\text{-}B_3\text{-}A_3\text{-}B\text{-}A_4$ and $A\text{-}(A\text{-}B)_2\text{-}A_3\text{-}B_4\text{-}A_3\text{-}B\text{-}A_4$, where A and B refer to repeat units that differ in the configuration in the chiral backbone atom.

Table 2. Average values[a] for the backbone root mean-square deviation ($RMSD_{Av}$; in Å), cross sectional diameter ($\overline{D}_{cs}$; in Å) and end-to-end distance ($\overline{R}_{e-e}$; in Å) for the isotactic, syndiotactic and heterotactic models of *de*-**PG2**.

Sequence[b]	**$RMSD_{Av}$**[c]	$\overline{D}_{cs}$	$\overline{R}_{e-e}$
A_{20}	-	27.93±0.69	47.05±0.44
B_{20}	1.286±0.272	29.16±0.53	44.47±0.67
$(A\text{-}B)_{10}$	1.408±0.286	28.80±0.46	39.61±1.29
$(A_3\text{-}B_3)_3\text{-}A_2$	1.104±0.314	29.38±0.77	45.97±0.57
$(A_5\text{-}B_5)_2$	1.493±0.303	29.05±0.48	44.69±0.68
$A\text{-}(A\text{-}B)_2\text{-}A_4\text{-}B_3\text{-}A_3\text{-}B\text{-}A_4$	0.813±0.207	28.81±0.60	45.78±0.51
$A\text{-}(A\text{-}B)_2\text{-}A_3\text{-}B_4\text{-}A_3\text{-}B\text{-}A_4$	1.339±0.305	28.96±0.46	46.09±0.61

[a] Averages were obtained over the whole simulation. [b] A and B refers to repeat units that differ in configuration of the chiral backbone atom. [c] Calculated with respect to the model proposed for A_{20}.

Table 2 lists the average of the backbone RMSD, $\overline{D}_{cs}$ and $\overline{R}_{e-e}$ values found for the seven investigated sequences, while Figure 10 represents the temporal evolution of D_{cs} and $R_{e\text{-}e}$. As can be seen, in all cases the RMSDs were lower than 1.7 Å evidencing that the proposed helix is stable not only for the isotactic models but also for syndiotactic and heterotactic ones. However, a detailed inspection of the resulting secondary structures reveals that the values of backbone dihedral angles χ_1 and χ_2 are exchanged for repeat units A and B. Thus, the average values obtained for B_{20} are {χ_1= -48.1±1.4º, χ_2= -166.9±1.9º}, which are perfectly exchanged with respect to those found for A_{20} (see previous section). Furthermore, this particular feature is systematically observed in the B repeat units of the syndiotactic and heterotactic sequences. As can be seen in Table 2 and Figure 10a, the influence of this exchange in the D_{cs} of helix is very small, the range of variation of the $\overline{D}_{cs}$ values being of only 1.45 Å (4.9%). For the $R_{e\text{-}e}$, the impact of the dependence of χ_1 and χ_2 on the chirality of the repeat unit is more significant (Figure 10b). Thus, $\overline{R}_{e-e}$ ranges from 39.61 Å (syndiotactic) to 47.05 Å (isotactic), which corresponds to a maximum variation of 16%. Accordingly, the re-adaptation in the packing of the dendritic side chains to miniminize the repulsive interactions has a larger effect on the structural parameters defined along the helical axis.

The overall of the results presented in this section indicates that the helix model found for the isotactic polymer is also stable for syndiotactic and heterotactic systems, even though the microscopic description of the backbone dihedral angles depends on the chirality of the repeat unit. Obviously, the changes produced by the backbone tacticity in the dihedral angles χ_1 and χ_2 influences the packing of the dendritic side groups. This is reflected by the projections of the different models after MD relaxation, which show that dendron···dendron interactions are different for isotactic, syndiotactic and heterotactic systems. In spite of this, in all cases the driving force that determines the packing of the dendrons is the electrostatic repulsion between the positive charges, which tends to be minimized in the helix. On the other hand, analysis of the relative potential energies indicated that the isotactic B_{20} and some of the heterotactic examined sequences, *e.g.* $(A_3\text{-}B_3)_3\text{-}A_2$ are almost isoenergetic.

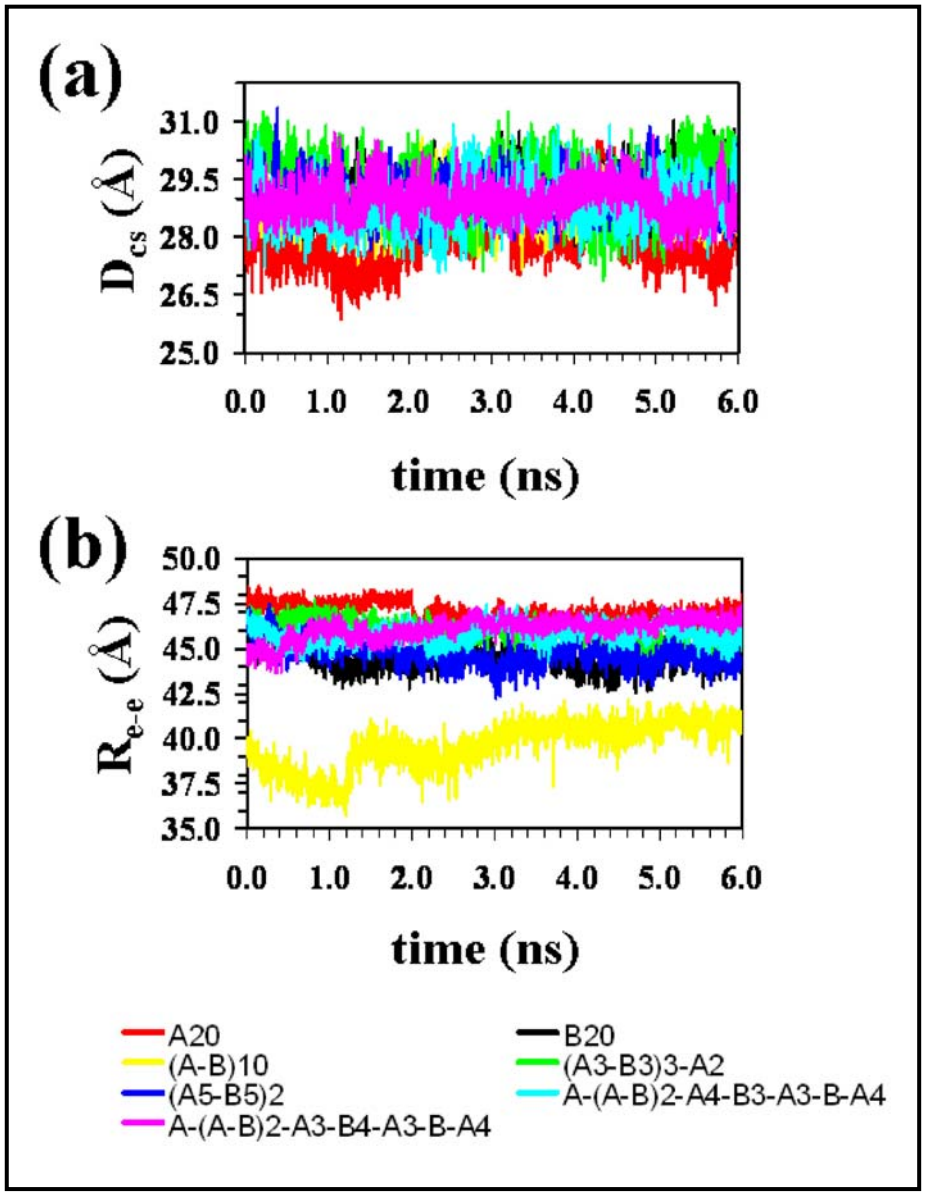


Figure 10. Time-dependent evolution of (a) the cross sectional diameter (D_{cs}) and (b) the end-to-end distance (R_{e-e}) for the right-handed helix of isotactic, syndiotactic and heterotactic **PG2**.

de-**PG2** is prepared by deprotecting **PG2** through the addition of trifluoroacetic acid.[5] CD-spectroscopic information indicated that the secondary structure of **PG2** is basically retained during this process in spite of the charged nature of *de*-**PG2**. This is fully consistent with the right-handed helices provided by atomistic simulations for the two systems. Thus, the conformation of the backbone is similar for **PG2** and *de*-**PG2**, the main difference between the two secondary structures being the packing of the dendritic side chains. In the protected polymer the helix, which is induced by the tight register between consecutive dendrons, is stabilized by a network of hydrogen bonds that involve the NH and CO groups of the neighboring repeat units. In contrast, the

secondary structure of the deprotected system is suitable to minimize the electrostatic repulsions, even though it is also induced by the size of the dendrons that enforce a very compact dendritic layer. It is worth noting that the conformational transformation between the two helices can be easily achieved by re-arranging the dendritic side chains. This is schematically illustrated in Figure 11, which shows the superposition of backbone and one dendritic side chain for the helices of **PG2** and *de*-**PG2**.

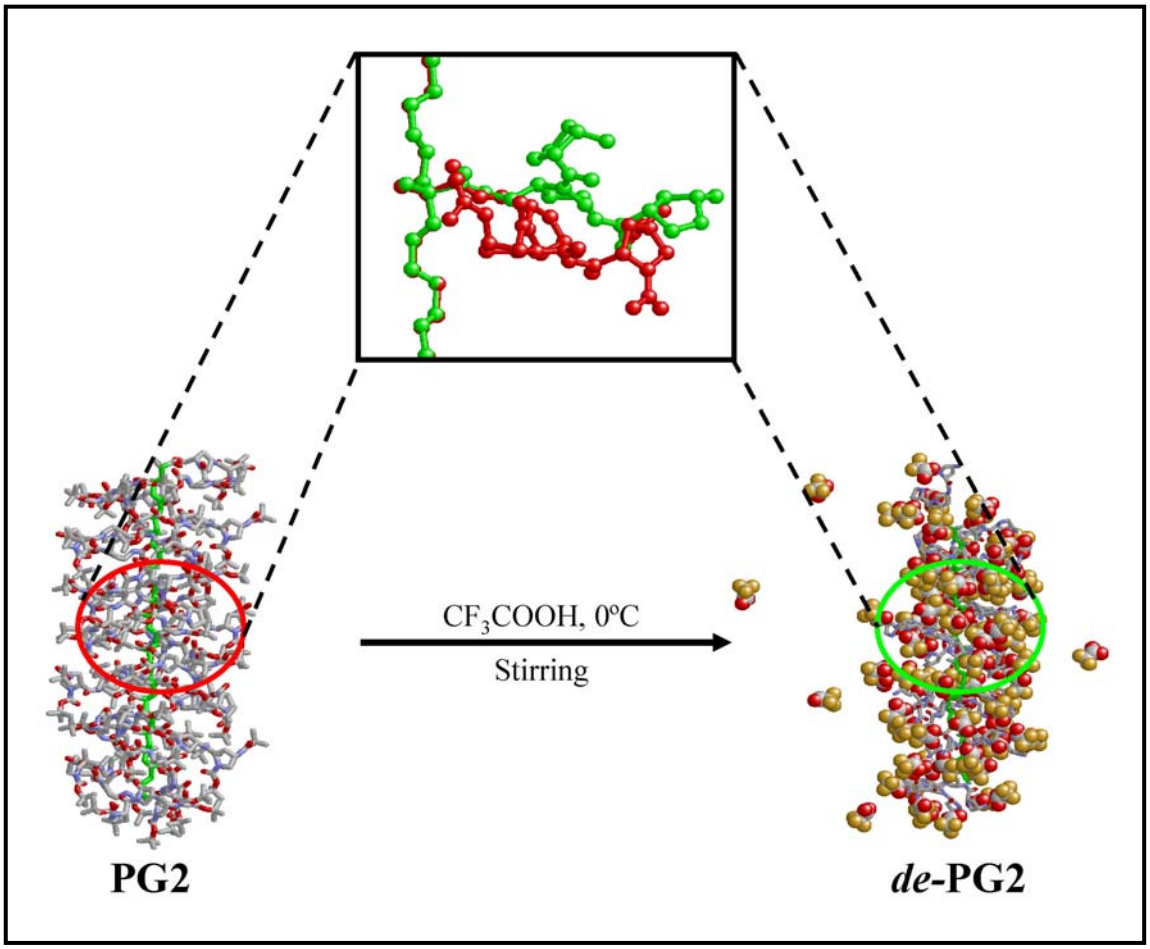


Figure 11. Axial projection of the right handed helical models of **PG2** and *de*-**PG2**. *tert*-Butyl groups and hydrogen atoms have been removed for clarity. The backbone is represented by a solid line and the CF_3COO^- counterions around the *de*-**PG2** helix are described using CPK filled balls. The inset shows the superposition of the helical backbone and one dendritic side chain (**PG2** and *de*-**PG2** in red and green, respectively).

4.2.4. *Conclusions*

MD simulations have been performed to determine and characterize the regular secondary structures of chiral second-generation dendronized polymethacrylates. Specifically, simulations have been used to explain the thermal stability of the helix found for **PG2**, which behaves as a helical rigid rod.

On the other hand, a systematic conformational search procedure, which is based on the combination of (i) quantum mechanical calculations on the dendritic side chain, (ii) force-field energy minimizations of a polymer chain using all the possible regular conformations as starting points, and (iii) classical MD simulations of the minimized models that retain the initial regularity, has been applied to determine the helix of *de*-**PG2**. In the secondary structure of this positively charged system, the backbone adopts an arrangement similar to that found for **PG2**, the main difference between the protected and the deprotected polymers referring to the dendritic side groups. Both the elongated conformation and the packing of the dendrons in the helix of *de*-**PG2** are defined by the repulsive interactions between the positive charges. However, such interactions, which are miniminized in the helix at room temperature, do not produce an appreciable reduction in the thermal stability of the deprotected polymer with respect to the protected one. Thus, the tight register allows retain the secondary structure if the kinetic energy is distributed without alter the electrostatic equilibrium associated to the positive charges.

MD simulations have proved that the helix of *de*-**PG2** is compatible with isotactic, syndiotactic and heterotactic sequences. Repeat units with different configuration in the chiral backbone atom show exchanged values in the dihedral angles χ_1 and χ_2, even though this variation does not alter the general aspect and main structural characteristics of the helix. It is worth noting that models proposed for **PG2** and *de*-**PG2** are fully compatible with reported experimental finding, the most significant being the transformation of the former into the latter upon addition of acid without disrupting the regular secondary structure.

4.2.5. *References*

[1] a) Nakano T., Okamoto Y. *Chem. Rev.* **2001,** *101,* 4013. b) Cornelissen J.J.L.M., Rowan A.E., Nolte R.J.M., Sommerdijk N.A.J.M. *Chem. Rev.* **2001,** *101,* 4039. c) Green M.M., Cheon K.-S., Yang S.-Y., Park J.-W., Swansburg S., Liu W. *Acc. Chem. Res.* **2001**, *34*, 672. d) Palmans A.R.A., Meijer E.W. *Angew. Chem., Int. Ed.* **2007**, *46*, 8948.

[2] a) Zhang A. *Prog. Chem.* **2005**, *17*, 157. b) Frauenrath H. *Prog. Polym. Sci*. **2005**, *30*, 325, c) Schlüter A.D. *Top. Curr. Chem*. **2005**, *245*, 151. d) Schlüter A.D., Raben J.P. *Angew. Chem. Int. Ed.* **2000**, *39*, 864. e) Gibson S.E., Rendell J.T. *Chem. Commun.* **2008**, 922.

[3] Zhao H., Sanda F., Masuda T. *Macromol. Chem. Phys.* **2006**, *207*, 1921.

[4] a) Percec V., Rudick J.G., Peterca M., Wagner M., Obata M., Mitchell C.M., Cho W.-D., Balagurusamy V.S.K., Heiney P.A. *J. Am. Chem. Soc.* **2005**, *127*, 15257. b) Percec V., Rudick J.G., Peterca M., Staley S.R., Wagner M., Obata M., Mitchell C.M., Cho W.-D., Balagurusamy V.S.K., Lowe J.N., Glodde M., Weichold O., Chung K.J., Ghionni N., Magonov S.N., Heiney P.A. *Chem. Eur. J.* **2006**, *12*, 5731. c) Percec V., Aqad E., Peterca M., Rudick J.G., Lemon L., Ronda J.C., De B.B., Heiney P.A., Meijer E.W. *J. Am. Chem. Soc.* **2006**, *128*, 16365. d) Percec V., Rudick J.G., Peterca M., Heiney P.A. *J. Am. Chem. Soc.* **2008**, *130*, 7503.

[5] Zhang A., Rodríguez-Ropero F., Zanuy D., Alemán C., Meijer E.W., Schlüter A.D. *Chem. Eur. J.* **2008**, *14*, 6924.

[6] Phillips J.C., Braun R., Wang W., Gumbart J., Tajkhorshid E., Villa E., Chipot C., Skeel R.D., Kale L., Schulten K. *J. Comput. Chem.* **2005**, *26*, 1781.

[7] Jorgensen W.L., Chandrasekhar J., Madura J.D., Impey R.W., Klein M.L. *J. Chem. Phys.* **1983**, *79*, 926.

[8] Caldwell J.W., Kollman P.A. *J. Phys. Chem.* **1995**, *99*, 6208.

[9] Jorgensen W.L., Chandrasekhar J., Madura J.D., Impey R.W., Klein M.L. *J. Chem. Phys.* **1983**, *79*, 926.

[10] Cornell W.D., Cieplak P., Bayly C.I., Gould I.R., Merz K.M., Ferguson D.M., Spellmeyer D.C., Fox T., Caldwell J.W., Kollman P.A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

[11] Alemán C., Luque F.J., Orozco M. *J. Comput. Aided Mol. Design* **1993**, *7*,721.

[12] Darden T., York D., Pedersen L. *J. Chem. Phys.* **1993**, *98*, 10089.

[13] Ryckaert J.P., Ciccotti G., Berendsen H.J.C. *J. Comput. Phys.* **1977**, *23*, 327.

[14] Berendsen H.J.C., Postma J.P.M., van Gunsteren W.F., Dinola A., Haak J.R. *J. Chem. Phys.* **1984**, *81*, 3684.

[15] Becke A.D. *J. Chem. Phys.* **1993**, *98*, 1372.

[16] Lee C., Yang W., Parr R.G. *Phys. Rev. B* **1988**, *37*, 785.

[17] Hariharan P.C., Pople J.A. *Theo. Chim. Acta* **1973**, *28*, 213.

[18] Frisch M.J. et al. *Gaussian 03,* Revision B.02, Gaussian, Inc.: Pittsburgh PA, **2003**.

[19] Brown K.G., Erfurth S.C., Small E.W., Peticolas W.L. *Proc. Natl. Acad. Sci.* **1972**, *69*, 1467.

[20] Zhang J., Beshra A., Domb A.J., Ozaki Y. *Eur. Polymer J.* **2007**, *43*, 3016.

[21] Rubinstein M., Colby R.H. *Polymer Physics*, Oxford University Press, Oxford, **2003**.

[22] a) Flory P.J. *Statistical Mechanics of Chain Molecules,* Wiley, New Cork, **1969**. b) Grosberg A.Y., Khokhlov A.R. *Statistical Physics of Macromolecules*, American Institute of Physics, New York, **1994**.

[23] Hsiao P.-Y. *Macromolecules* **2006**, *39*, 7125.

[24] Cifra P. *Polymer* **2004**, *45*, 5995.

4.3. Electronic characterization of all-thiophene conducting dendrimers: Molecules and assemblies*

4.3.1. *Introduction*

Over the past decades the interest towards conducting polymers has attracted the attention of many researchers because of their wide range of technological applications in fields such as electronics, biomedical engineering or optics.[1] Among conducting polymers, materials based on thiophene have settled as a prominent family because of the high stability of their doped and undoped states, the easiness to tune their structure and their controllable electrochemical behavior, giving rise to diverse applications in the areas of microelectronics, electrode materials, optoelectronics and sensors.[2] Electronic conduction in polythiophene derivatives is founded on their characteristic polyconjugated nature, which promotes both the intramolecular and intermolecular delocalization of π-electrons.[3] Although the synthesis and investigation of a number of conventional thiophene derivatives have been vastly exploited in the last decade,[4] investigation on all-thiophene based dendrimers is currently one of the most challenging achievements.[5-9]

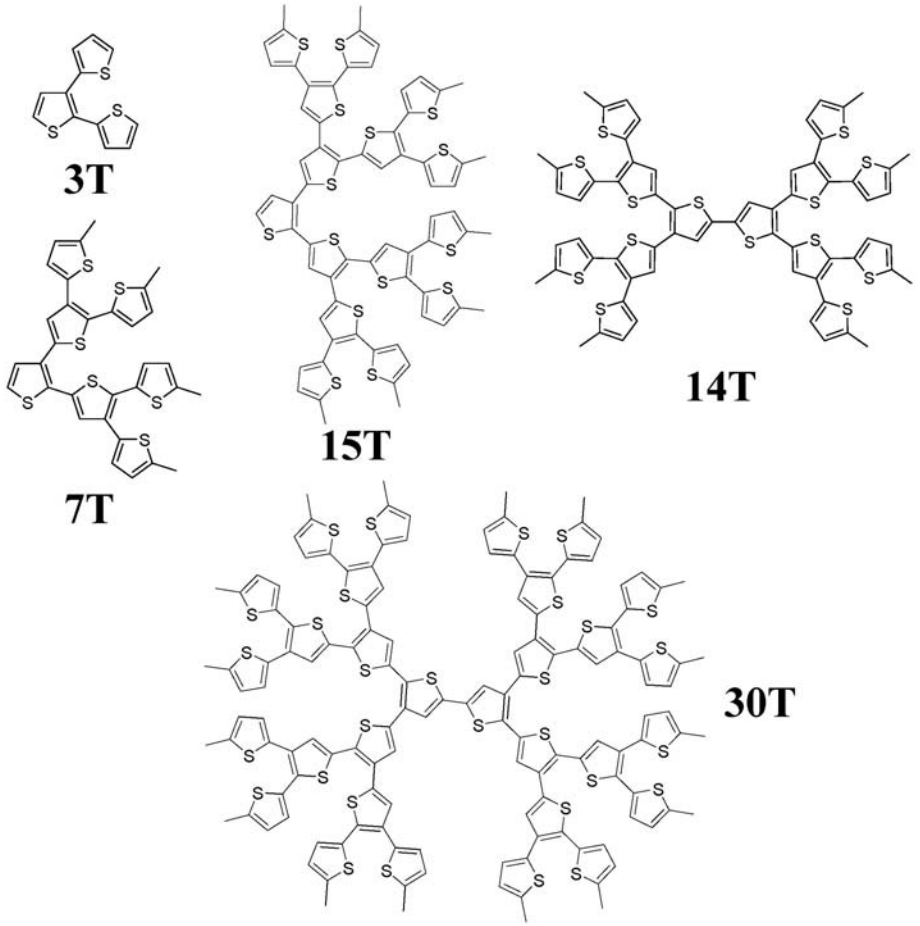
Dendrimers are ideally perfect monodisperse macromolecules composed by a multifunctional core unit where a defined number of units called *dendrons* are attached leading to a regular and highly branched three-dimensional architecture. Three topological regions coexist in dendrimers, *i.e.* the inner dense core, the dendritic region around the core and the external surface, being the number of repeating units between the inner core and a terminal unit defined as *dendron generation*.[10] The peculiar architecture of dendrimers and the possibility of easily tuning their surface by adding different functional groups make them attractive materials to be used in a great variety of fields such as in biomedical applications,[11] catalysis,[12] sensing,[13] or optoelectronics.[14] In the latter field the electronic properties of conducting polymers combined with both the great shape stability and the ability to self-assemble of dendrimers,[15] offer new possibilities for designing new optoelectronic devices.[16]

Given the aforementioned assumptions, all-thiophene based dendrimers are considered as a very promising kind of conducting material. They were synthesized and characterized for the first time by Advincula and co-workers.[5] Specifically, these

* Submitted for publication

researchers prepared a series of *n*T dendrimers, where *n* refers to the number of thiophene rings (T), with *n*= 3, 6, 7, 14, 15 and 30 using metal-mediated coupling reactions through a convergent approach. The potential use of such dendrimers in optoelectronics and molecular electronics was evidenced by their very broad absorption spectra and the ability of the bigger dendrimers to self-assemble on different solid substrates forming nanowires and 2-D crystalline structures. On the other hand, Bäuerle and co-workers[6] synthesized different all-thiophene dendrimers containing up to 90 thiophene rings (90T) with a divergent/convergent approach to facilitate the inclusion of functionalities in the external surface of the conducting dendrimer. More recently, femtosecond transient absorption measurements showed the suitability of all-thiophene dendrimers as light harvesting antennae for solar cells devices.[7] Concerning their optical properties, non-linear optical and excited state dynamical investigations evidenced a delocalization of the excitation throughout all the thiophene units and ultrafast energy transfer to the longest branch of the dendrimer.[7] Moreover, all-thiophene dendrimers have been demonstrated to be good sensors for entangled photons at very low flux making them suitable materials for imaging of biological systems, quantum optical applications and remote sensing.[9]

Given the importance of the electronic properties in the applications of all-thiophene dendrimers, theoretical studies based on quantum mechanics (QM) are necessary not only to provide comprehensive understanding at the microscopic level but also to gain deeper insight on the intrinsic characteristics of these materials. In this chapter QM methods are used to study the structural and electronic properties of *n*T dendrimers with *n*=3, 7, 14, 15 and 30 (Scheme 1). Since the electrical properties of poly(thiophene) derivatives are promoted by p-doping, *i.e.* an oxidation process,[17] QM calculations have been performed considering both the neutral and the oxidized states of the dendrimers under study. Special attention has been paid to the evaluation of the lowest π-π^* transition energy, which has been determined using not only a conventional QM method but also a sophisticated time-dependent approach. Furthermore, calculations based on a theoretical procedure that combines quantum mechanics and molecular mechanics (QM/MM calculations) have been used to examine the supramolecular assembly of the 30T dendrimer.

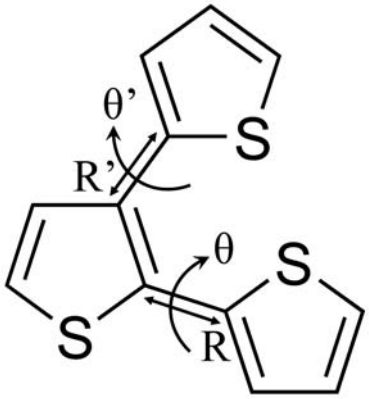


Scheme 1

4.3.2. *Methods*

Quantum Mechanics

QM calculations were carried out using the Gaussian 03 computer program.[18] Complete geometry optimizations were performed at the B3LYP[19,20] level combined with the 6-31G(d) basis set,[21] *i.e.* B3LYP/6-31G(d). The minimum energy arrangements of 3T in both neutral and oxidized states were determined using a systematic conformational search strategy. For this purpose, the potential energy hypersurface of this compound was defined as E= E(θ,θ'), where θ and θ' refer to the flexible inter-ring dihedral angles (Scheme 2). The dihedral angles θ and θ' were scanned in steps of 30º between 0º and 360º, which led to build 12×12= 144 structures. All these structures were calculated using a flexible rotor approximation. Thus, each structure was submitted to a constrained geometry optimization in which the inter-ring dihedral angles θ and θ' were kept fixed at the initial values.



Scheme 2

The most stable arrangement of 3T was used to build up the 7T, which after complete geometry optimization was used to construct the 15T. The initial structures of 14T and 30T were obtained by dimerizing the optimized structures of 7T and 15T, respectively, the dihedral angle associated to such dimerization ($\theta_{D\text{-}D}$) being initially arranged at 180º. It is worth noting that the lowest energy minimum of α,α'-bithiophene corresponds to the *anti-gauche* ($\theta \approx$ 150º) and anti ($\theta =$ 180º) conformations in the neutral[22] and oxidized[23] state, respectively. As was performed for 3T and 7T, the molecular geometries of 14T, 15T and 30T were optimized without any constraint. The structures of the dendrimers in the oxidized state, hereafter identified as $3T^{\cdot+}$, $7T^{\cdot+}$, $14T^{\cdot+}$, $15T^{\cdot+}$ and $30T^{\cdot+}$, was identical to that described for the systems in the neutral state. The restricted formalism was considered for calculations on neutral dendrimers (closed-shell systems), while for oxidized $n\mathrm{T}^{\cdot+}$ systems the unrestricted formalism (UB3LYP) was used. For all the calculated dendrimers *n*T and $n\mathrm{T}^{\cdot+}$, with exception of those with $n =$ 30, frequency analysis were carried out to verify the nature of the minimum state of all the stationary points obtained by geometry optimization.

The lowest π-π^* transition energy (ε_g) in the neutral and oxidized states was obtained using two different strategies. In the first one, ε_g was estimated using the

Koopman's theorem (KT) as the difference between the energies of the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO),[24] *i.e.* $\varepsilon_g = \varepsilon_{LUMO} - \varepsilon_{HOMO}$. The second strategy consists on the evaluation of ε_g using time-dependent density functional theory (TD-DFT).[25] More specifically, electronic excitations were evaluated at the (U)B3LYP/6-31G(d) level using the previously optimized geometries. On the other hand, the ionization potential of each dendrimer (IP) was estimated using the KT, that is, relating the IP to the energy of the HOMO.[24] However, more accurate IP values were obtained using the energies of the fully relaxed neutral and oxidized species. This approach, usually denoted ΔSCF, takes into account the relaxation energy of the ionized state, which can be calculated as the difference between KT and ΔSCF IPs.[26,27]

Quantum Mechanics / Molecular Mechanics

All QM/MM calculations were performed on $30T^{\cdot+}$ and 30T dendrimers assuming two ideal situations: (i) a dendrimer surrounded by four replicas placed in parallel stacking, *i.e.* sandwich arrangement; and (ii) a dendrimer surrounded by four replicas in T-shaped arrangement. Initially, these arrangements were built using the geometry of $30T^{\cdot+}$ optimized at the UB3LYP/6-31G(d) level. Accordingly, five dendrimers were involved in each supramolecular arrangement: $(5{\cdot}30T)^{\cdot+}$. The sandwich and T-shaped arrangements of $(5{\cdot}30T)^{\cdot+}$ were submitted to constrained geometry optimizations using MM, in which the intermolecular parameters were the only allowed to vary, *i.e.* the molecular geometry of $30T^{\cdot+}$ was kept fixed in all cases. MM minimizations were performed using the NAMD computer package.[28] The optimized geometries were used for single point QM/MM calculations of both the oxidized $(5{\cdot}30T)^{\cdot+}$ and the corresponding neutral system, *i.e.* 5·30T, which allowed to describe a complete redox

process. In these calculations, which were carried out using the Gaussian 03 computer program,[18] the central molecule was described at the (U)B3LYP/6-31G(d) QM level while the four surrounding replicas were represented as simple point charges. The molecular electrostatic potential (MEP) was computed at the (U)HF/6-31G(d) level for $30T^{\cdot+}$ and 30T. These calculations, which were performed on a large set of points located outside the nuclear region, were used to obtain the electrostatic charges applying an early developed procedure.[29]

4.3.3. *Results and discussion*

3T has been considered as the building block of the *n*T with $n > 3$ dendrimers studied in the present work. Accordingly, a good description of this small compound becomes essential to build the rest of the dendrimers applying a bottom-up approach. Figure 1 displays the potential energy hypersurfaces E= E(θ,θ') calculated for 3T and $3T^{\cdot+}$ at the (U)B3LYP/6-31G(d). The global minimum of the neutral system appears at θ,θ'= 125.2º, -37.9º (Figure 2a), which upon oxidation moves towards θ,θ'= 26.6º, 28.1º (Figure 2b). Comparison of the two hypersurfaces reveals that the rotational freedom is significantly higher for 3T than for $3T^{\cdot+}$. It is worth noting that planar conformations of the neutral building block, *i.e.* those with θ and θ' taking values of 0º or 180º, are severely destabilized (by more than 10 kcal/mol), which must be attributed to steric hindrance between the first and third thiophene rings. For $3T^{\cdot+}$ rotations around the α-α' and α'-β'' linkages are clearly disfavored, which leads to a very rigid structure. This behavior is due to the formation of a quinoid structure in the oxidized specie that is clearly evidenced by the shortening of the inter-ring lengths, *e.g.* *R*= 1.457 / 1.422 Å and *R'*= 1.466 / 1.432 Å for the lowest energy minimum of 3T / $3T^{\cdot+}$. Thus, rotation around a double bond involves a high energy penalty.

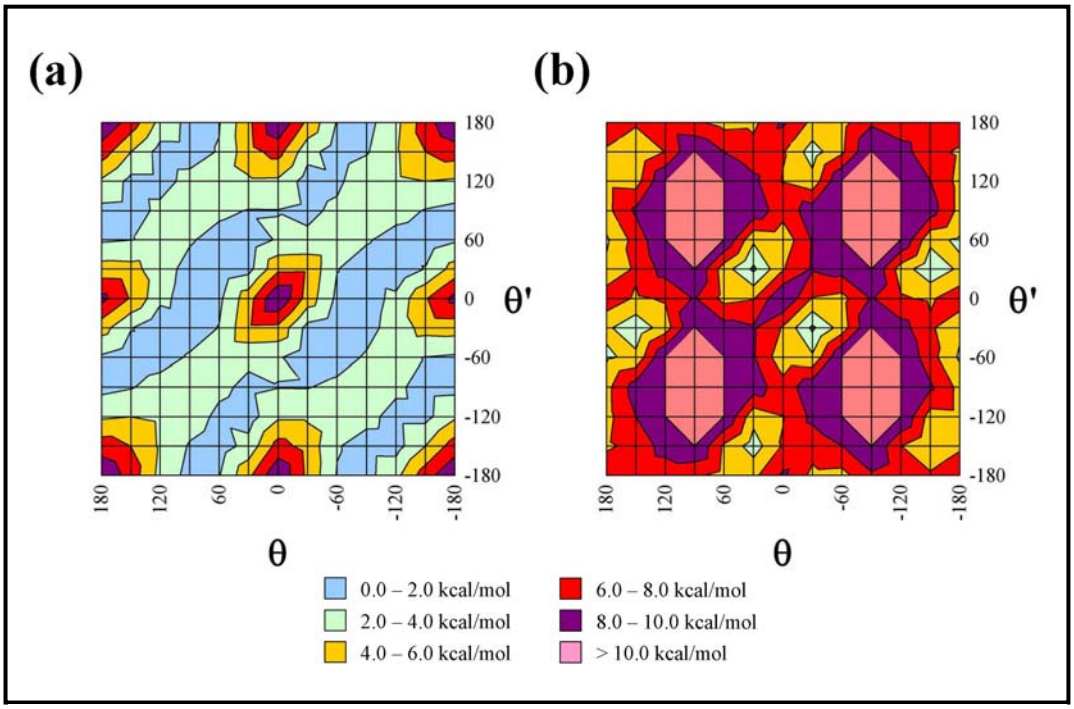


Figure 1. Potential energy hypersurface E=E(θ,θ') of (a) 3T and (b) $3T^{\bullet+}$ computed at the (U)B3LYP/6-31G(d) level.

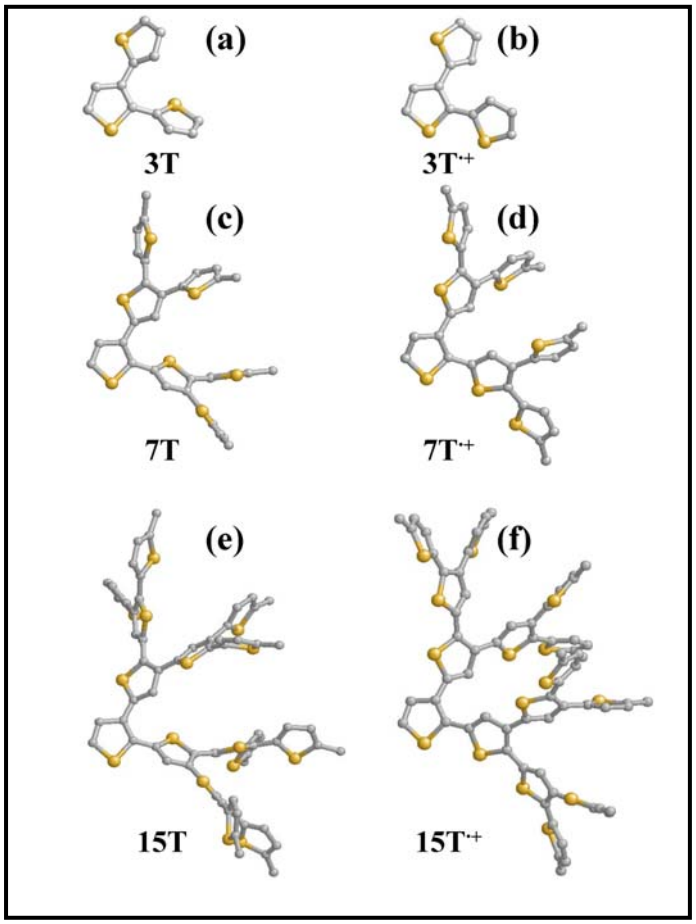


Figure 2. Optimized structures of (a) 3T, (b) $3T^{\bullet+}$, (c) 7T, (d) $7T^{\bullet+}$, (e) 15T and (f) $15T^{\bullet+}$ calculated at the (U)B3LYP/6-31G(d) level. Hydrogen atoms have been omitted for clarity.

The optimized structures of 3T and $3T^{\bullet+}$ were used as a building blocks to construct the third generation dendrimers, 7T and $7T^{\bullet+}$, respectively. Similarly, the optimized geometries of 7T and $7T^{\bullet+}$, which are displayed in Figures 2c and 2d, were used to build 15T and $15T^{\bullet+}$, respectively. The structures obtained for these fourth generation dendrimers after complete geometry optimization are displayed in Figures 2e and 2f, respectively. Average values for the dihedral angles θ and θ' as well as for the inter-ring bond lengths R and R' (see Scheme 2) for the neutral and oxidized dendrimers are reported in Table 1. As can be seen, the geometrical parameters predicted for 3T, 7T and 15T are very similar suggesting a regular growing pattern. Thus these three dendrimers are fully superimposable, as shown in Figure 3a. On the other hand, oxidized dendrimers undergo substantial deviation from ideality, which increase with the molecular size. In spite of this behavior, that is mainly due to steric hindrance between neighboring sulfur atoms, Figure 3b shows that the structures of the three oxidized dendrimers can be also superposed. It is also noticeable that the average inter-ring length R is slightly larger than that found in 2,2',5',2''-terthiophene, in which the two inter-ring linkages are α-α', at the same level of theory (1.448 Å). This feature is consequence of both the steric hindrance between neighboring thiophene rings and the different distribution of charges. Comparison between the inter-ring bond lengths predicted for nT and $n\mathrm{T}^{\bullet+}$ reveals that, as expected, the differences (ΔR and $\Delta R'$) decreases with n, *e.g.* ΔR / $\Delta R'$= 0.035 / 0.034 Å for n= 3 and ΔR / $\Delta R'$= 0.012 / 0.003 Å for n= 15.

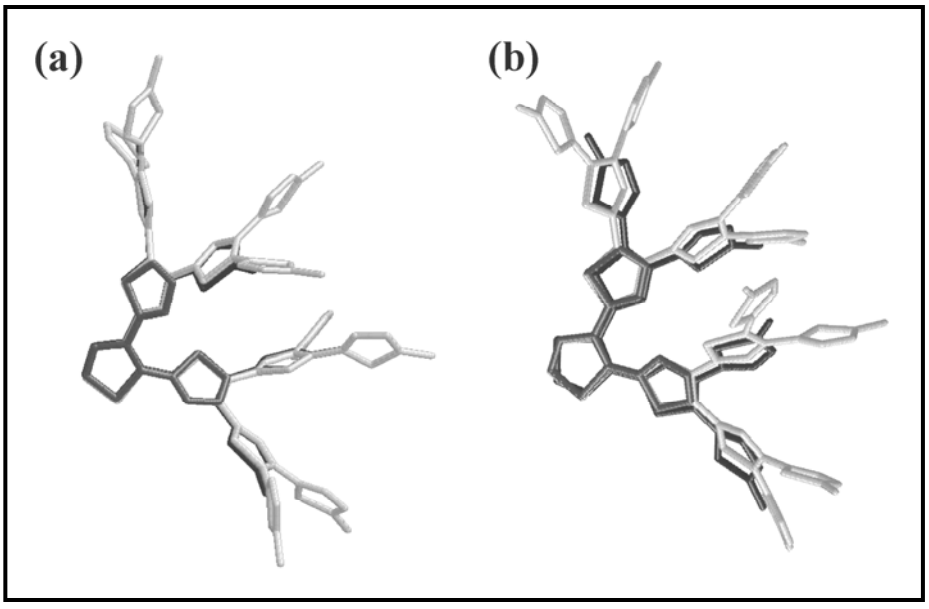


Figure 3. Superposition of the optimized structures of: (a) 3T, 7T and 15T; and (b) $3T^{\bullet+}$, $7T^{\bullet+}$ and $15T^{\bullet+}$. Hydrogen atoms have been omitted for clarity. Each dendrimer is represented using different colors.

Table 1. Geometric parametersa for nT (with n= 3, 7 and 15) dendrimers in the neutral (plain text) and oxidized (italic text) states obtained from complete geometry optimizations at the (U)B3LYP/6-31G(d) level.

	θ (º)	θ' (º)	R (Å)	R'(Å)
3T	125.2	-37.9	1.457	1.466
	26.6	*28.1*	*1.422*	*1.432*
7T[b]	128.6 ± 1.2	-38.5 ± 0.7	$1.458 \pm 1.5 \cdot 10^{-3}$	$1.466 \pm 1.5 \cdot 10^{-3}$
	25.4 ± 3.0	*35.0 ± 5.6*	*$1.435 \pm 5.8 \cdot 10^{-4}$*	*$1.455 \pm 8.7 \cdot 10^{-4}$*
15T[c]	129.6 ± 1.4	-40.1 ± 0.6	$1.457 \pm 2.1 \cdot 10^{-3}$	$1.465 \pm 1.3 \cdot 10^{-3}$
	29.9 ± 4.0	*38.4 ± 2.6*	*$1.445 \pm 4.5 \cdot 10^{-3}$*	*$1.462 \pm 3.8 \cdot 10^{-3}$*

[a] See Scheme 2. [b] Averages over the three values measured for each geometric parameter. [c] Averages over the seven values measured for each geometric parameter.

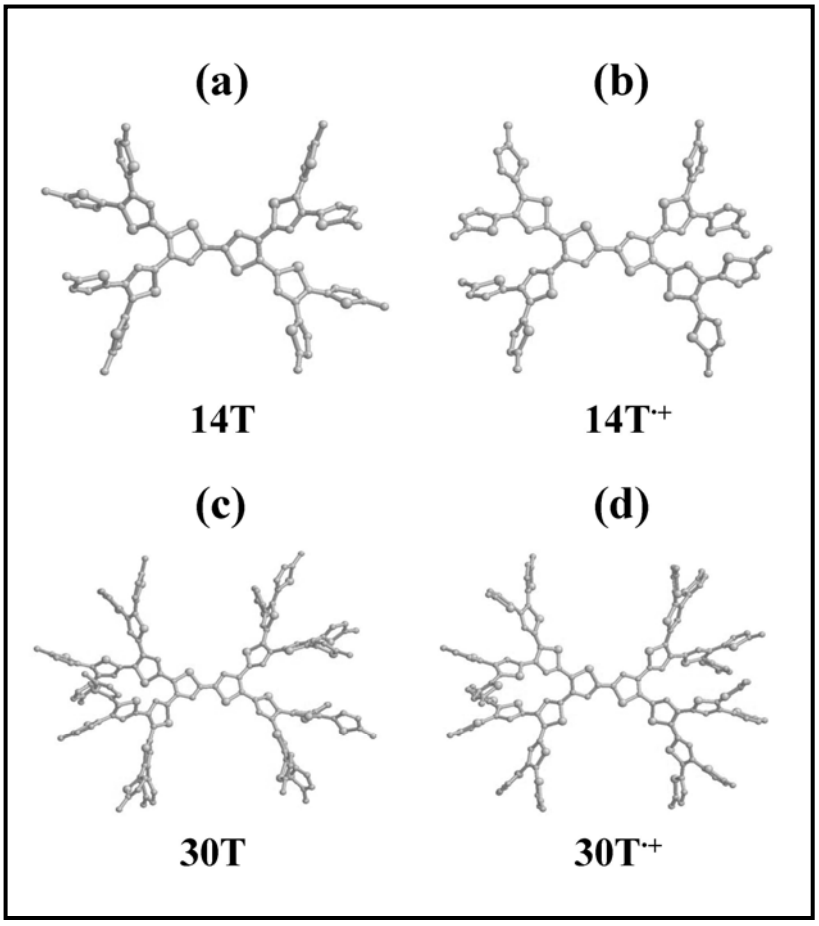


Figure 4. Optimized structures of (a) 14T, (b) $14T^{\bullet+}$, (c) 30T and (d) $30T^{\bullet+}$ calculated at the (U)B3LYP/6-31G(d) level. Hydrogen atoms have been omitted for clarity.

Figure 4 shows the optimized structures of dendrimers 14T and 30T in both neutral and oxidized states, which were built upon *trans* α-α' dimerization of dendrimers 7T and 15T, respectively. Inspection of the most relevant averaged geometric parameters for the neutral dendrimers (Table 2) shows a complete agreement with those of the parent structures 7T and 15T. However, steric hindrance caused again structural deviations from ideality in $14T^{\cdot+}$ and $30T^{\cdot+}$ leading to an increase up to 2.1 % in the interthiophene bond *R'*. The inter-dendron dihedral angle, $\theta_{D\text{-}D}$ in Table 2, in the neutral dendrimers deviates slightly from 158.3º (8.9º and 4.0º for 14T and 30T, respectively), which is the inter-ring dihedral angle associated to the rotation of the α-α' linkage in 2,2'-bithiophene predicted at the same level of theory. On the contrary, in $14T^{\cdot+}$ and $30T^{\cdot+}$ both steric hindrance and charge distribution lead to loss the planar structure, which is the most stable arrangement in oxidized 2,2'-bithiophene.

Table 2. Geometric parameters[a,b] for *n*T (with *n*= 14 and 30) dendrimers in the neutral (plain text) and oxidized (*italic text*) states obtained from complete geometry optimizations at the (U)B3LYP/6-31G(d) level.

	θ (º)	θ' (º)	R (Å)	R'(Å)	$\theta_{D\text{-}D}$ (º)
14T[c]	131.5 ± 2.8	-40.6 ± 2.2	$1.455 \pm 1.4 \cdot 10^{-3}$	$1.466 \pm 1.9 \cdot 10^{-3}$	167.2
	28.2 ± 4.8	*39.8 ± 2.4*	*$1.443 \pm 7.5 \cdot 10^{-3}$*	*$1.464 \pm 3.9 \cdot 10^{-3}$*	*167.0*
30T[d]	131.5 ± 2.7	-40.8 ± 1.9	$1.456 \pm 1.9 \cdot 10^{-3}$	$1.466 \pm 1.5 \cdot 10^{-3}$	162.3
	33.6 ± 5.2	*40.1 ± 1.7*	*$1.449 \pm 5.4 \cdot 10^{-3}$*	*$1.466 \pm 3.0 \cdot 10^{-3}$*	*168.4*

[a] See Scheme 2. [b] The inter-dendron dihedral angle has been defined by the S-C-C-S sequence. [c] Averages over the six values measured for each geometric parameter. [c] Averages over the fourteen values measured for each geometric parameter.

The IPs predicted by the KT and ΔSCF approaches for all the calculated dendrimers, which have been normalized considering an average per thiophene ring, are reported in Table 3. IPs predicted by the KT are slightly higher than those predicted by the ΔSCF, being this difference mainly attributed to the fact that the latter approach takes into account the molecular relaxation of the ionized state. Both KT and ΔSCF IPs indicate that the energy needed to oxidize *n*T dendrimers decrease when *n* grows, which must be attributed to the increase of the electron delocalization with the molecular size. The same behavior was found in linear α-oligothiophenes.[30] Similarly, the IPs calculated for 14T and 30T are lower than those of their parent monomers, 7T and 15T, respectively.

Table 3. IPs calculated using the KT and ΔSCF approaches for dendrimers 3T, 7T, 15T, 14T and 30T. In order to facilitate the comparison, values have been normalized by considering the average per thiophene ring.

KT (eV·ring^{-1})		ΔSCF (eV·ring^{-1})	
3T	3.21	**3T**	2.26
7T	1.08	**7T**	0.84
14T	0.48	**14T**	0.39
15T	0.44	**15T**	0.37
30T	0.20	**30T**	0.18

The ε_g values estimated for *n*T and *n*T$^{\cdot+}$ dendrimers using both the KT and TD-DFT calculations, in which HOMO–LUMO excitations are considered, are listed in Table 4, available experimental data[5b] being also included for comparison. In all cases the values estimated using the KT are overestimated with respect to the TD-DFT values. Comparison between the TD-DFT ε_g values and the experimental ones, which are available for the neutral dendrimers, reveals an excellent agreement. Furthermore, this notable concordance, that is especially remarkable for 14T and 30T dimers, increases with the molecular size. This indicates that the electronic structure of the neutral dendrimers is satisfactorily described by the TD-B3LYP/6-31G(d) methodology.

Table 4. Lowest π-π^* transition energy (ε_g; in eV) of *n*T and *n*T$^{\cdot+}$ dendrimers obtained using the KT and TD-DFT methodologies at the (U)B3LYP/6-31G(d) level. Experimental values for the neutral systems, which were obtained using spectroscopy and oxidation-reduction potentials (in parenthesis for 14T and 30T), have been included for comparison. n_α refers to the longest α-conjugated chain of a given dendrimer.

		Neutral			Oxidized	
	n_α	KT	TD-DFT	Exp.[a]	KT	TD-DFT
3T	2	4.13	3.66	3.33	1.55	1.25
7T	3	3.51	3.05	2.82	0.84	0.80
15T	4	3.24	2.82	2.62	0.71	0.62
14T	6	2.95	2.60	2.48 (2.46)	0.54	0.43
30T	8	2.86	2.51	2.50 (2.51)	0.34	0.34

[a] Data from reference 5b. Methyl groups have been used as terminators in this work, while hexyl groups were used in reference 5b.

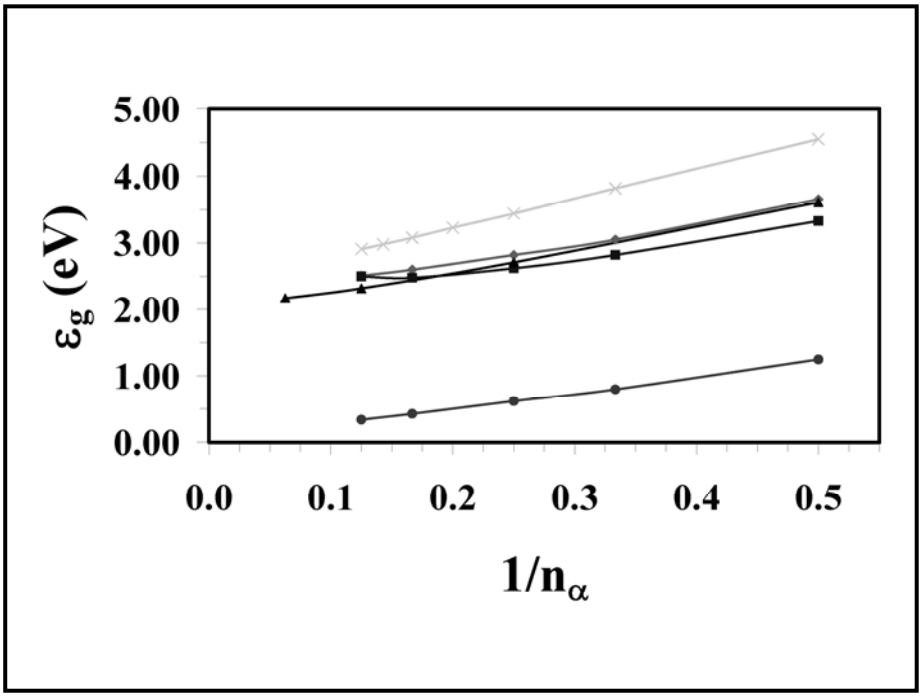


Figure 5. Evolution of the π- π * lowest transition energy (ε_g; in eV) plotted against the inverse number of the longest α-conjugated chain present in each dendrimer (1/nα). The symbol code corresponds to: (i) ε_g calculated using the TD-B3LYP/6-31G(d) method for nT (diamonds); (ii) ε_g calculated using the TD-B3LYP/6-31G(d) method for $nT^{\bullet+}$ (circles); (iii) experimental ε_g values (taken from reference 5b) for $nT^{\bullet+}$ (squares); (iv) ε_g calculated using the KT at the B3LYP/6-31G(d) level (taken from reference 27) for linear α-oligothiophenes in the neutral state (crosses); and (v) experimental ε_g values (taken from reference 6) for linear α-oligothiophenes in the neutral state (triangles).

Figure 5 represents the variation of the ε_g calculated at the TD-B3LYP/6-31G(d) for both *n*T and $n\mathrm{T}^{\bullet+}$ against $1/n_\alpha$, where n_α corresponds to the longest α-conjugated chain present of a given dendrimer. Additionally, available experimental data for *n*T,[5b] as well as the ε_g values for linear α-oligothiophenes determined theoretically applying the KT[31] to the B3LYP/6-31G(d) energies and experimentally from the onset of the longest wavelength absorption band,[6] are also plotted in Figure 5 for comparison. In all cases, the ε_g decreases when the size of the α-conjugated chain increases. It should be noted that the overestimation of the ε_g values estimated for α-oligothiophenes using the KT with respect to the experimental values decreases with the molecular size, this behavior being also detected in *n*T dendrimers. In the latter systems, this should be attributed to the loss of planarity induced by the high branching. Comparison of the two experimental sets confirms this assumption for the two dimers, 14T and 30T. A linear regression analysis ($\varepsilon_g = a \cdot 1/n_\alpha + b$) using the ε_g predicted by TD-DFT calculations for

neutral dendrimers was performed to extrapolate the value of an idealized dendrimer containing infinite thiophene units. The result, ε_g = 2.08 eV, indicates a reduction of 1.58 eV with respect to the value obtained for 3T. Regarding $n\mathrm{T}^{\cdot+}$, ε_g also shows a clear linear dependence with $1/n_\alpha$ (Figure 5). As it can be seen, the ε_g predicted for oxidized dendrimers is significantly smaller than that of the neutral ones indicating the good performance of these systems as electronic conductors upon oxidation. This feature is especially remarkable for $30\mathrm{T}^{\cdot+}$, which shows a ε_g value that is 7.4 times smaller than that of 30T.

Given the aforementioned promising electronic properties of $n\mathrm{T}^{\cdot+}$ dendrimers, the modeling of their supramolecular assemblies is very desirable. For this purpose, calculations were performed on the largest oxidized dendrimer, *i.e.* $30\mathrm{T}^{\cdot+}$, which is the paradigmatic case. Specifically, the stacked was considered, also denoted sandwich, and T-shaped ideal arrangements of five molecules, $(5{\cdot}30\mathrm{T})^{\cdot+}$, where a central dendrimer was surrounded by four replicas placed with coplanar and perpendicular orientations, respectively. The optimized geometry of the individual dendrimers was used to construct the assemblies. Figures 6a and 7a depict the sandwich and T-shaped assemblies of $(5{\cdot}30\mathrm{T})^{\cdot+}$, respectively, after energy minimization using MM. Due to both the lack of planarity and the large dimensions of $30\mathrm{T}^{\cdot+}$, the relative orientation between pairs of thiophene rings located at neighboring dendrimers is mainly perpendicular. This detail is illustrated in Figures 6b and 7b for the sandwich and T-shaped supramolecular structures, respectively.

Table 5. Ionization potential (IP) and lowest π-π^* transition energy ($\varepsilon_{g,neutral}$) of 5·(30T) and lowest π-π^* transition energy ($\varepsilon_{g,oxidized}$) of $(5{\cdot}30\mathrm{T})^{\cdot+}$ derived from QM/MM calculations considering the sandwich and T-shaped arrangements. In order to facilitate the comparison, the IP values have been normalized by considering the average per thiophene ring.

	IP ($\mathrm{eV{\cdot}ring^{-1}}$)	$\varepsilon_{g,neutral}$ (eV)	$\varepsilon_{g,oxidized}$ (eV)
Sandwich	0.17	2.32	0.32
T-Shaped	0.16	2.33	0.32

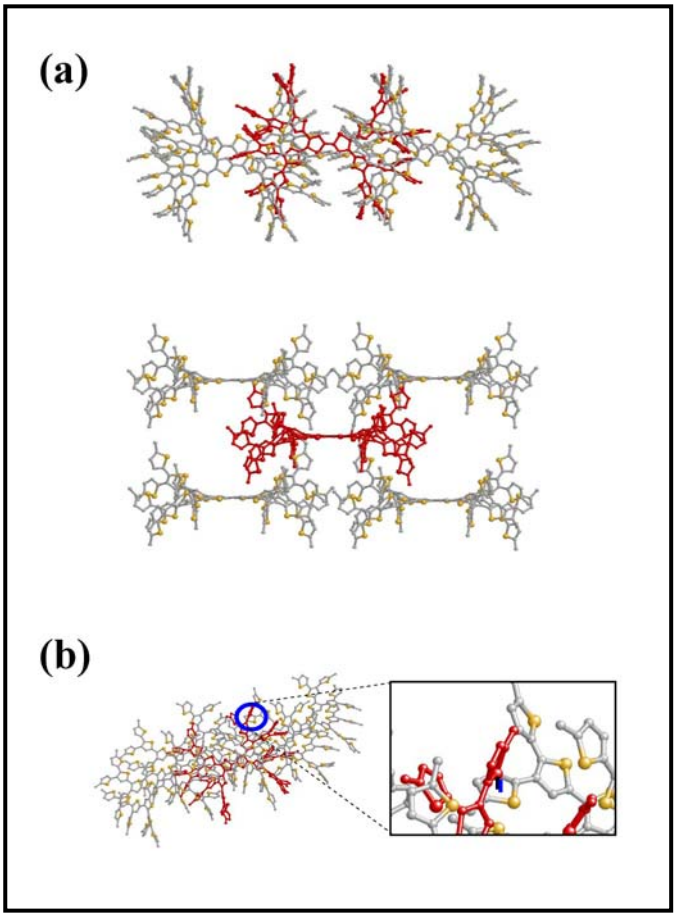


Figure 6. Sandwich structure of $(5{\cdot}30T)^{\cdot+}$: (a) equatorial (top) and axial (bottom) projections after energy minimization using MM; and (b) view of the perpendicular stacking (indicated by a blue line in the magnified picture) of the thiophene rings located at neighboring dendrimers. The dendrimer placed in the center is depicted in red. Hydrogen atoms have been removed for clarity.

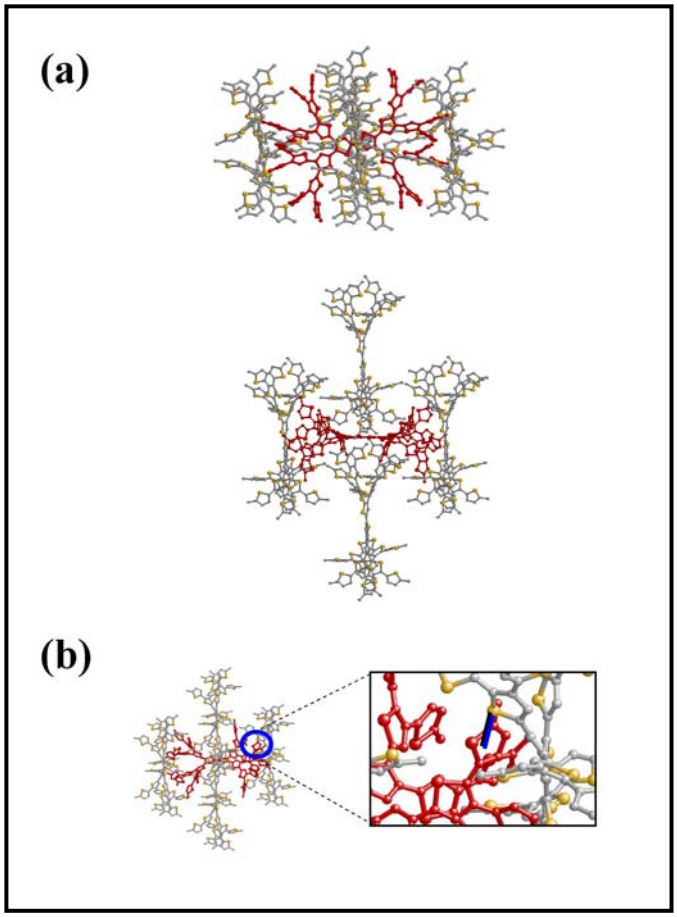


Figure 7. T-shaped structure of $(5{\bullet}30T)^{\bullet+}$: (a) equatorial (top) and axial (bottom) projections after energy minimization using MM; and (b) view of the perpendicular stacking (indicated by a blue line in the magnified picture) of the thiophene rings located at neighboring dendrimers. The dendrimer placed in the center is depicted in red. Hydrogen atoms have been removed for clarity.

The geometries displayed in Figures 6 and 7 were used to carry out single point QM/MM calculations not only of the oxidized system $(5{\cdot}30T)^{\cdot+}$ but also of the neutral one, 5·(30T), as was described in Methods. It is worth noting that, as the energy difference between the neutral and oxidized states of the part described using QM was calculated, both the MM and the QM/MM van der Waals contributions vanish because the geometries used for the two electronic states were identical. Thus, the only term that was not canceled in such energy difference, apart from the QM energy, corresponds to the QM/MM electrostatic interaction since the polarization of the wave-function is a non-linear process.[32] Results obtained from these calculations are summarized in Table 5. As can be seen, the two arrangements show similar ε_g and IP values, the most noticeable result being the reduction of ε_g for the neutral system with respect to the

isolated dendrimer. Thus, the ε_g predicted for 30T using the KT was 2.86 eV (Table 4), whereas the value of this property decreases to 2.32-2.33 eV, depending on the supramolecular arrangement, when the intermolecular electronic delocalization is allowed. Thus, thiophene-thiophene stacking interactions promote this intermolecular phenomenon, which is added to the intrinsic intramolecular electronic delocalization, narrowing the energy difference between HOMO and LUMO, *i.e.* reducing the ε_g.[33] Accordingly, QM/MM calculations confirm the ability of the investigated dendrimers to form aggregates and their good performance when assembled in supramolecular arrangements. Thus, they should be considered as novel conducting polymers with potential electrooptical applications. On the other hand, it should be noted that, if longer alkyl chains are used as terminal groups (in the present chapter methyl as terminal groups have been used), only the T-shaped arrangement is no possible due to the steric hindrances. Moreover, longer alkyl groups would stabilize the overall sandwich arrangement because of the favorable van der Waals interactions between neighboring molecules.

4.3.4. *Conclusions*

QM calculations have been used to get an accurate description of the molecular and electronic structure of a series of all thiophene dendrimers in both the neutral and oxidized states. Independently of the electronic state, minimum energy structures were found to present repulsive steric interactions between the neighbouring thiophene rings, which induce significant distortions from the co-planarity of thiophene rings. Within this context, the regular growing pattern of neutral dendrimers is especially remarkable since it favours the formation of highly regular architectures. On the other hand, the highest generation dendrimers exhibited the lowest IP per ring, which is consequence of the largest electron delocalization. Moreover, ε_g values derived from TD-DFT calculations, which were in excellent agreement with available experimental data for very similar all thiophene dendrimers, presented a reciprocal interdependence with the inverse of the longest α-conjugated chain of the dendrimer, *i.e.* ε_g decreases when the generation increases. Extrapolation of the results derived from TD-DFT calculations in the neutral dendrimers allowed to predict a ε_g value of 2.08 eV for an idealized infinite

generation dendrimer. Oxidized dendrimers presented very small ε_g values suggesting that they behave as excellent electronic conductors upon p-doping.

QM/MM calculations on systems formed by five dendrimers reflected the ability of these materials to form supramolecular aggregates in both sandwich and T-shaped arrangements, the latter being only feasible when short chains are used as terminators. Furthermore, these calculations have evidenced the influence of the intermolecular electronic delocalization, which reduces considerably the ε_g. The overall of these results offers a deeper insight on the electronic structure that controls the electrooptical properties characteristic of all thiophene dendrimers and indicates that they are a very promising family of conducting materials.

4.3.5. *References*

[1] a) Otero T. F., Sariñena J. M. *Adv. Mater.* **1998**, *10*, 491. b) Dai L., Winkler B., Dong L., Tong L., Mau A. W. H. *Adv. Mater.* **2001**, *13*, 915. c) Tour J. M. *Acc. Chem. Res.* **2000**, *33*, 791.

[2] Roncali J. *Chem. Rev.* **1992**, *92*, 711.

[3] Chan H. S. O., Ng S., C. *Prog. Polym. Sci.* **1998**, *23*, 1167.

[4] Mishra A., Ma C.-Q., Bäuerle P. *Chem. Rev.* **2009**, *109*, 1141.

[5] a) Xia C., Fan X., Locklin J., Advincula R. C. *Org. Lett.* **2002**, *4*, 2067. b) Xia C., Fan X., Locklin J., Advincula R. C., Gies A., Nonidez W. *J. Am. Chem. Soc.* **2004**, *126*, 8735.

[6] Ma C.-Q., Mena-Osteritz E., Debaerdemaeker T., Wienk M. M., Janssen R. A. J., Bäuerle P. *Angew. Chem. Int. Ed.* **2007**, *46*, 1679

[7] Ramakrishna G., Bhaskar A., Bäuerle P., Goodson III T. *J. Phys. Chem. A* **2008**, *112*, 2018.

[8] Benincori T., Capaccio M., De Angelis F., Falciola L., Muccini M., Mussini P., Ponti A., Toffanin S., Traldi P., Sannicol F. *Chem. Eur. J.* **2008**, *14*, 459.

[9] Harpham M. R., Süzer Ö., Ma C.-Q., Bäuerle P., Goodson III T. *J. Am. Chem. Soc.* **2009**, *131*, 973.

[10] a) Bosman A. W., Janssen H. M., Meijer E. W. *Chem. Rev.* **1999**, *99*, 1665. b) Zhang A. *Prog. Chem.* **2005**, *17*, 157. c) Frauenrath H. *Prog. Polym. Sci*. **2005**, *30*, 325, d) Schlüter A. D. *Top. Curr. Chem*. **2005**, *245*, 151. e) Schlüter A. D., Rabe J. P. *Angew. Chem. Int. Ed.* **2000**, *39*, 864. f) Ballauff M., Likos C. N. *Angew. Chem. Int. Ed.* **2004**, *43*, 2998.

[11] Jang W.-D., Selim K. M. K., Lee C.-H., Kang I.-K. *Prog. Pol. Sci.* **2009**, *34*, 1.

[12] Caminade A.-M., Servin P., Laurent R., Majoral J.-P. *Chem. Soc. Rev.* **2008**, *37*, 56.

[13] Astruc D., Ornelas C., Ruiz J. *Acc. Chem. Res.* **2008**, *41*, 841.

[14] Percec V., Glodde M., Bera T.K., Miura Y., Shiyanovskaya I., Singer K. D., Balagurusamy V. S. K., Heiney P. A., Schnell I., Rapp A., Spiess H. W., Hudson S. D., Duan H. *Nature* **2002**, *419*, 384.

[15] a) Percec V., Ahn C. H., Ungar G., Yeardley D. J. P., Möller M., Sheiko S. S. *Nature* **1998**, *391*, 161. b) Percec V., Imam M. R., Peterca M., Wilson D. A., Heiney P. A. *J. Am. Chem. Soc.* **2009**, *131*, 1294.

[16] Sonar P., Benmansour H., Geiger T., Schlüter A .D. *Polymer* **2007**, *48*, 4996.

[17] Graf D. D., Campbell J. P., Miller L. L., Mann K. R. *J. Am. Chem. Soc.* **1996**, *118*, 5480.

[18] Frisch M.J. et al. *Gaussian 03,* Revision B.02, Gaussian, Inc.: Pittsburgh PA, **2003**.

[19] Becke A. D. *J. Chem. Phys.* **1993**, *98*, 1372.

[20] Lee C., Yang W., Parr R. G. *Phys. Rev. B* **1993**, *37*, 785.

[21] Hariharan P. C., Pople J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

[22] Rodríguez-Ropero F., Casanovas J., Alemán C. *Chem. Phys. Lett.* **2005**, *416*, 331.

[23] Casanovas J., Alemán C. *J. Phys. Chem. C* **2007**, *111*, 4823.

[24] Koopmans T. *Physica* **1934**, *1*, 104.

[25] Runge E., Gross E.K.U. *Phys. Rev. Lett.* **1984**, *52*, 997.

[26] Casanovas J., Ricart J. M., Rubio J., Illas F., Jiménez-Mateos J. M. *J. Am. Chem. Soc.* **1996**, *118*, 8071.

[27] Casanovas J., Alemán C. *J. Phys. Chem. A* **2004**, *108*, 1440.

[28] Phillips J.C., Braun R., Wang W., Gumbart J., Tajkhorshid E., Villa E., Chipot C., Skeel R.D., Kale L., Schulten K. *J. Comput. Chem.* **2005**, *26*, 1781.

[29] Alemán C., Casanovas J. *J. Chem. Soc. Perkin Trans. 2* **1994**, 563.

[30] Johansson E., Larsson S. *Synthetic Met.* **2004**, *144*, 183.

[31] Alemán C., Armelin E., Iribarren J.I., Liesa F., Laso M., Casanovas J. *Synthetic Met.* **2005**, *149*, 151.

[32] Jacquemin D., Perpète E.A., Laurent A.D., Assfeld X., Adamo C. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1258.

[33] Rodríguez-Ropero F., Casanovas J., Alemán C. *J. Comput. Chem.* **2007**, *29*, 69.

5. THEORETICAL STUDY OF π-CONJUGATED SYSTEMS

5.1. Solvation of 2,2'-bithiophene: influence of the first solvation shell in the properties of π-conjugated systems*

5.1.1. *Introduction*

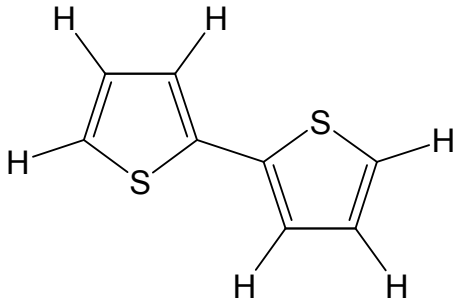
Conducting polymers are among the more studied novel families of polymers because of their wide range of potential applications in different fields. Some relevant applications are their use in electrochromic devices, light-emiting diodes (LEDs), artificial muscles, thin-film transistors and electrolytic capacitors.[1-6] Conducting polymers present a characteristic polyconjugated nature, which allows the intramolecular delocalization of π-electrons.[4] In particular, polythiophene and some of its derivatives have been subject of a major research activity due to their electric and nonlinear optical properties, excellent environmental stability and relatively low cost of production.[4-6]

Quantum mechanical calculations have been frequently used to investigate the molecular and electronic structure in the gas phase of conducting polymers through properly chosen model systems.[7-13] In contrast, the properties and behavior of these compounds in solution remain almost unexplored.[14,15] There are three basic strategies for modeling molecular interactions in solution using quantum mechanics.[16,17] The first one involves the discrete models, in which both solvent and solute molecules are explicitly treated allowing specific solute-solvent interactions. The second approach is based on continuum models. In these strategies the solvent is represented as a continuum dielectric medium that interacts with the solute. The most popular are the Self-Consistent Reaction-Field (SCRF) methods. The third strategy, which brings the advantages of the two aforementioned methods together, is called combined discrete/continuum model. Thus, this model considers both the electronic effects associated to specific solute-solvent interactions and the electric polarization of the bulk solvent.[18,19]

This chapter presents a systematic study about the influence of the solvation in the conformational preferences of 2,2'-bithiophene, abbreviated 2-thp (Scheme 1). For this purpose calculations were performed in water and acetonitrile using the three stated above methods. It should be noted that 2-thp is the simplest model molecule of polythiophene, while acetonitrile is the organic solvent usually employed in the

* Results presented in this chapter are published in ***Chem. Phys. Lett.*** **2005,** ***416***, 331

electrogeneration of polythiophenes.[20,21] On the other hand, results in aqueous solution will be especially relevant for future studies on model compounds of recently developed water-soluble polythiophenes.[22,23]



Scheme 1

5.1.2. *Methods*

Gas phase calculations

Ab initio calculations were performed at the HF/6-311+G(d,p) level. To study the internal rotation of 2-thp, the inter-ring dihedral angle (θ) defined in Scheme 1, was scanned in steps of 30º between θ = 180º (anti conformer) and θ = 0º (syn conformer).

Solvation models

The three solvation models used in this work were applied to both water and acetonitrile solvents. For the discrete model, complexes were generated by placing two explicit solvent molecules at the more hydrophilic regions of the solute gas-phase optimized structure. The generated complexes were submitted at complete geometry optimization at the HF/6-311+G(d,p) level.

The Polarizable Continuum Model (PCM) developed by Tomasi and coworkers[17] was chosen for SCRF calculations. This method involves the generation of a solvent cavity from spheres centered at each atom of the molecule and the calculation of virtual point charges on the cavity surface representing the polarization of the solvent. The magnitude of these charges is proportional to the derivative of the solute electrostatic potential at each point calculated from the molecular wave function. The PCM was applied within the ab initio HF/6-311+G(d,p) framework considering two different approximations. First, molecular geometries optimized in the gas-phase were used for

SCRF calculations. Second, the influence of the solvent-induced geometrical relaxation was analyzed by re-optimizing the geometries in solution.

The combined discrete/SCRF model was applied by performing PCM calculations within the HF/6-311+G(d,p) framework on the complexes optimized previously for the discrete solvation.

To study the internal rotation of 2-thp in solution, the inter-ring dihedral angle θ was scanned as described previously for the gas-phase.

All the calculations were performed using the Gaussian 98 and Gaussian 03 computer programs.

5.1.3. *Results and discussion*

Gas phase calculations

Figure 1 shows the rotational profile calculated for 2-thp in the gas-phase, while the energies and dihedral angles of the minima and saddle points obtained for this compound are compiled in Table 1. The global minimum appears at θ= 147.2º (anti-gauche conformation), which is in excellent agreement with results previously reported in the literature.[7] Similarly, the stability and position of the local minimum, which corresponds to the syn-gauche conformation (θ= 43.8º), and the barriers associated to the anti (θ= 180º), gauche-gauche (θ= 90º) and syn (θ= 0º) conformations are fully consistent with results previously obtained at the HF level.[7,11] The application of more sophisticated methods provides more precise estimations of the 2-thp conformational properties but it should be noted that the aim of this chapter is to analyze the ability of different methods to describe solvent-induced effects, which are usually omitted, in the π-conjugated systems used to model conducting polymers rather than to re-investigate the properties of these molecules.

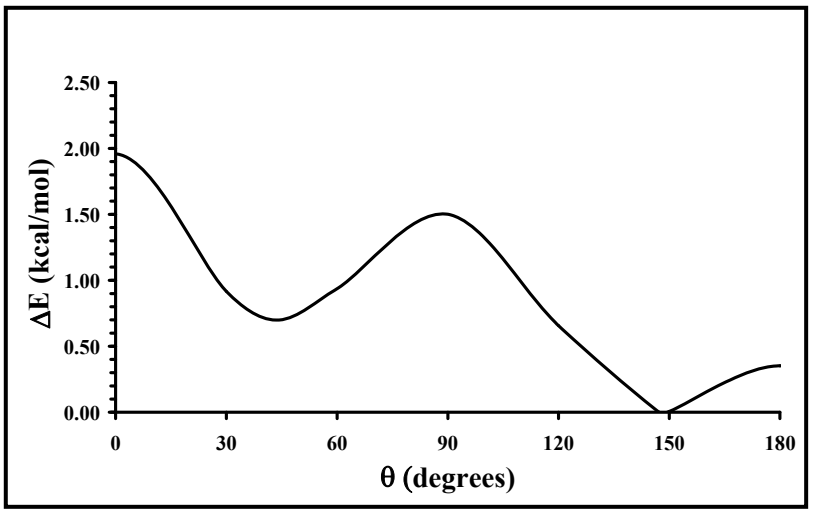


Figure 1. Evolution of the gas-phase relative energy as a function of the inter-ring dihedral angle, θ, for 2-thp

Table 1. Energies (in kcal/mol) and torsional angles (θ, in degrees) for the minima and saddle points predicted in different environments for 2-thp. Calculations in solution were performed using the discrete, continuum SCRF and combined discrete/SCRF models.

Environment	Model	syn (θ=0º)	syn-gauche	gauche-gauche (θ=90º)	anti-gauche	anti (θ=180º)
Gas-phase		2.0	0.7 (θ=43.8º)	1.5	0.0 (θ=147.2º)	0.3
Aqueous solution	Discrete	1.8	0.0 (θ=40.8º)	4.7	0.6 (θ=150.4º)	4.3
Aqueous solution	SCRF (PCM)	1.1	0.6 (θ=37.6º)	2.2	0.0 (θ=152.4º)	0.2
Aqueous solution	Discrete/SCRF	1.5	0.8 (θ=40.8º)	6.0	0.0 (θ=150.4º)	0.3
Acetonitrile solution	Discrete	3.4	0.0 (θ=43.4º)	4.1	2.7 (θ=150.0º)	2.6
Acetonitrile solution	SCRF (PCM)	1.2	0.6 (θ=38.6º)	2.1	0.0 (θ=150.4º)	0.2
Acetonitrile solution	Discrete/SCRF	1.9	0.8 (θ=43.4º)	1.8	0.0 (θ=150.0º)	0.5

Calculations with explicit solvent molecules

Figure 2 shows the rotational profiles calculated for 2-thp·2H_2O and 2-thp·2CH_3CN complexes, the more important information about the minima and saddle points being included in Table 1. The global minimum predicted in both cases corresponds to the syn-gauche conformation revealing that solute-solvent specific interactions induce a drastic change in the relative stability of the two minima. Furthermore, the interaction between 2-thp and the two water molecules retains the degree of folding detected in the gas-phase for the anti-gauche minimum, i.e. θ changes from 147.2º (gas-phase) to 150.4º. Conversely, the interaction of 2-thp with two CH_3CN molecules enhances the

planarity of the anti-gauche minimum, the value of θ increasing by more than 30º with respect to the gas-phase. Regarding to the barriers, the incorporation of two explicit solvent molecules produces a destabilization of the gauche-gauche and anti conformations, especially in aqueous solution, while the influence of the solvent molecule in the syn conformation is less pronounced.

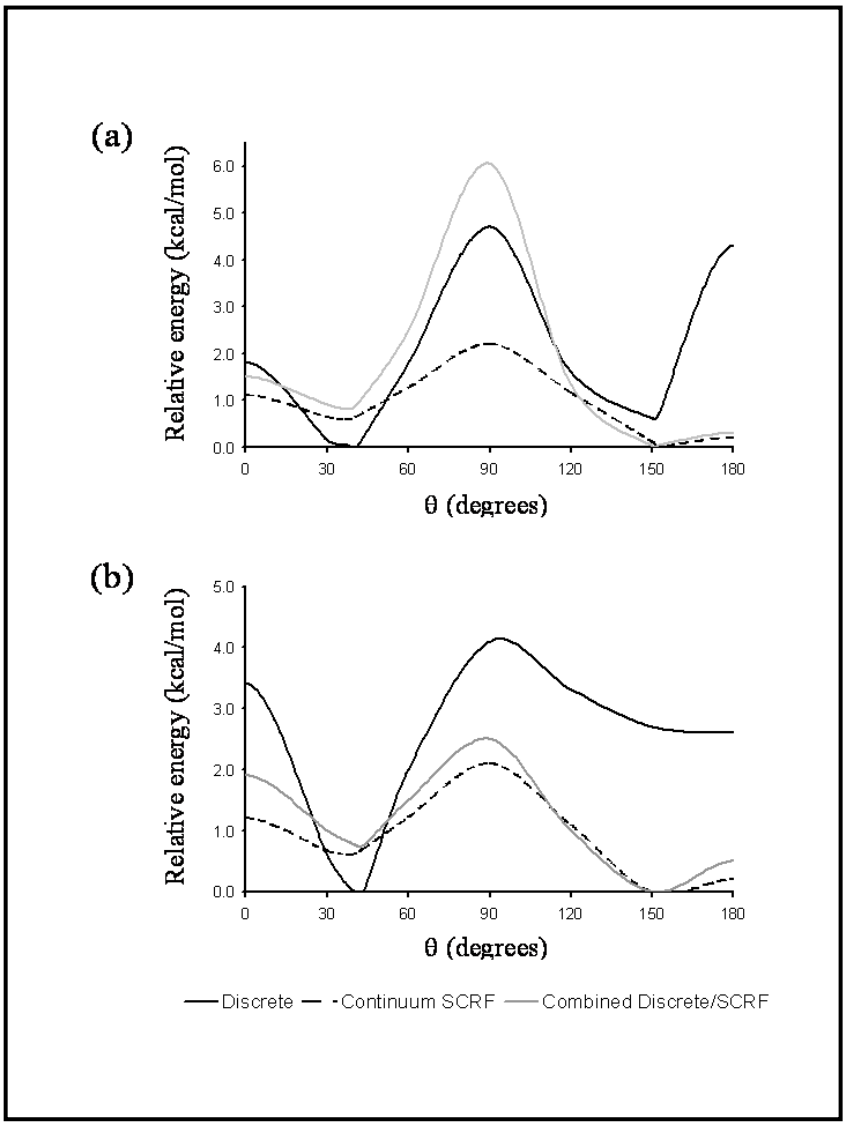


Figure 2. Evolution of the relative energy in (a) aqueous and (b) acetonitrile solutions as a function of the inter-ring dihedral angle, θ, for 2-thp. Results obtained using discrete, continuum SCRF and combined discrete/SCRF solvation models are displayed.

The binding energy, ΔE_b, of each calculated complex was computed as the difference between the energy of the complex and the sum of the energies calculated for 2-thp and the two solvent molecules (Eq. (8.1)), the basis set superposition error being corrected by applying the counterpoise method.[24]

$$\Delta E_b = E(\text{complex}) - E(\text{explicit solvent molecules}) \quad (8.1.)$$

Figure 3 shows the variation of ΔE_b against the inter-ring dihedral angle θ for the optimized geometries of 2-thp·$2H_2O$ and 2-thp·$2CH_3CN$ complexes. Interestingly, the most favorable ΔE_b value corresponds to the most stable conformation, *i.e.* the syn-gauche minimum, in both cases. Indeed, the results displayed in Figure 2 and 3 suggest that the energies of the rotational profile predicted by the discrete model depends on ΔE_b. These results clearly indicate that the specific interactions between the solute and the solvent molecules localized at the first solvation shell are fundamental to describe the conformational properties of π-conjugated molecules like that studied in this work.

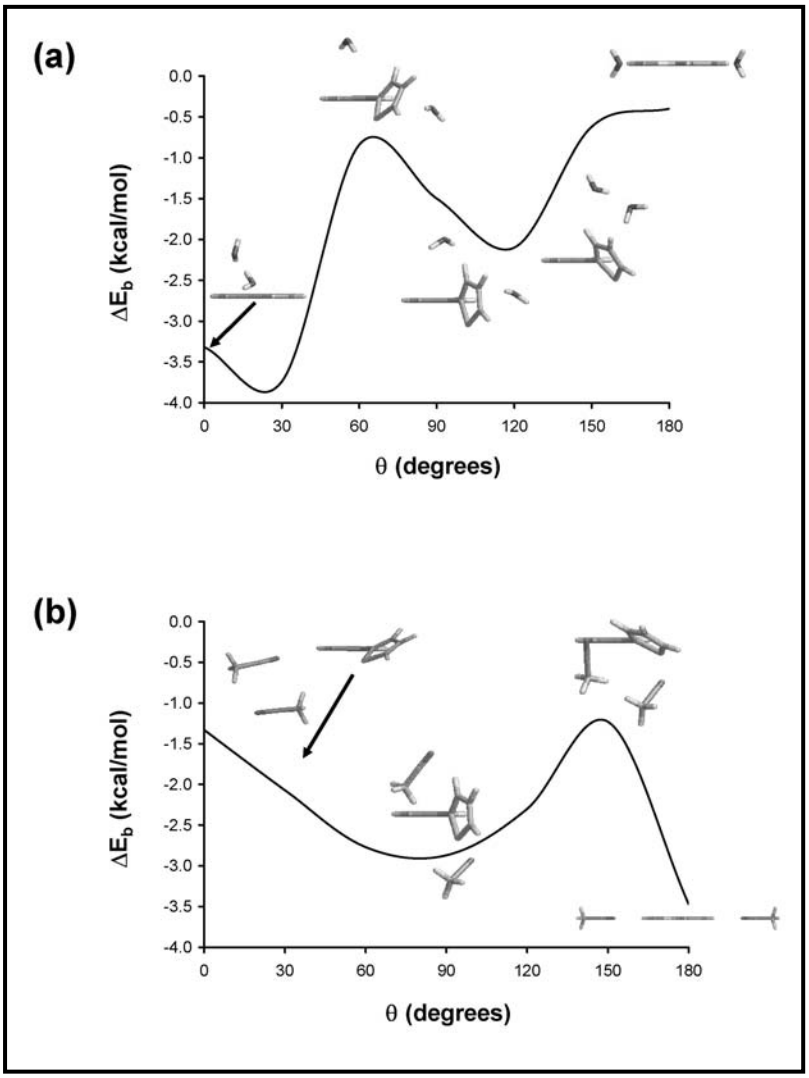


Figure 3. Evolution of the binding energy, ΔE_b, as a function of the inter-ring dihedral angle, θ, for 2-thp·$2H_2O$ (a) and 2-thp·$2CH_3CN$ (b) complexes. The more relevant complexes are displayed.

SCRF calculations

The internal rotation of 2-thp was computed in both aqueous and acetonitrile solutions (ε= 78.39 and 36.64, respectively) using the PCM method. Initially, the free energies of solvation (ΔG_{sol}) were calculated using the gas-phase geometries. After this, the ΔG_{sol} values were recomputed but optimizing all the molecular geometric parameters (bond lengths and bond angles) in the solvent-phase, while the dihedral angles were kept fixed at the gas-phase values. The resulting values were almost identical to those obtained using the gas-phase geometries indicating that the change of molecular geometry upon solvation has a negligible effect on the thermodynamic parameters. Thus, considering the expensiveness of the solvent phase geometry optimizations the use of gas-phase optimized geometries seems fully justified in future studies.

Figure 2 includes the energy profiles for the rotation of the inter-ring dihedral angle in water and acetonitrile provided by the PCM method. The energies in solution were obtained by the classical thermodynamical scheme, *i.e.* adding the ΔG_{sol} energies to the gas-phase energies. As can be seen, the rotational profiles were similar for the two solvents from both qualitative and quantitative points of views. This is a consistent result since the variation of the 2-thp dipole with θ is relatively small, *i.e.* the dipole moment varies from 0.01 to 2.40 Debyes when θ increases from 0º to 180º, and therefore, the variation of ΔG_{sol} with the conformation is expected to be also small. This is because for the chosen solvents and within this SCRF method, solute-solvent interactions are dominated by the electrostatic contribution.

Interestingly, the anti-gauche was predicted as the most stable conformation in both solvents (θ= 150.4º and θ= 152.4º for acetonitrile and water respectively), the syn-gauche being a local minimum that is destabilized by about 0.6 kcal/mol with respect to the global minimum. These results are fully consistent with those previously reported by Hernández and Navarrete.[14] These authors used the Onsager SCRF approximation[26] to predict that, in acetonitrile the anti-gauche minimum (θ= 147.6º) is about 0.6 kcal/mol stable than the syn-gauche one. These results are in contradiction with those obtained using the discrete solvation model with explicit solvent molecules. Thus, electric polarization of the bulk solvent and solute-solvent specific interactions produce different effects on the conformational preferences of 2-thp.

Inspection to the energy barriers listed in Table 1 shows that the PCM predicts that the gauche-gauche is about 0.6 kcal/mol less stable in solution than in the gas-phase, the

stability of the anti is similar in the two environments, and the syn is about 0.8 more favored in solution than in the gas-phase. The barriers in acetonitrile solution provided by the Onsager model[14] were, in general, similar to those calculated in this work with the PCM method, the largest difference being 0.7 kcal/mol.

Discrete/SCRF calculations

The rotational profiles derived from combined discrete/SCRF calculations are included in Figure 2, while the more important information about the minima and saddle points is included in Table 1. Calculations were performed applying the PCM method to the optimized geometries of 2-thp·$2H_2O$ and 2-thp·$2CH_3CN$ complexes. The energy associated to each conformation in solution was estimated as the sum of the electronic energy calculated for complex in gas-phase and the ΔG_{sol} provided by the PCM method. The profile computed in aqueous solution indicates the anti-gauche is the global, the syn-gauche being 0.8 kcal/mol less stable. This is in excellent agreement with the results provided by SCRF calculations, while the discrete model predicted the opposite order of stability. The highest energy barrier predicted by the combined discrete/SCRF model corresponds to the gauche-gauche, which was unfavored by 6.0 kcal/mol with respect to the global minimum. This value differs by 1.3 and 4.8 kcal/mol with respect to those calculated using the discrete and PCM models, respectively. Accordingly, the first solvation shell effected provided by the discrete model are crucial to describe the gauche-gauche barrier, while the anti and syn barriers are reasonably reproduced by the continuum model.

The results obtained in aqueous solution show that the description of the bulk solvent through a continuum dielectrics leads to different conformational preferences compared with those obtained using explicit solvent molecules. However, it should be noticed that the combined discrete/SCRF model takes into account both solute-water specific interactions and the solute polarization induced by the bulk solvent.

Regarding to the rotational profile in acetonitrile solution, the combined discrete/SCRF model predicts that the anti-gauche is the lowest energy minimum, the syn-gauche being destabilized by 0.8 kcal/mol. The gauche-gauche results 1.8 kcal/mol unfavored with respect to the global minimum, being again the highest energy barrier. As can be seen, the solvent effects predicted by the discrete/SCRF model are very similar to those

calculated using PCM continuum method suggesting that speficic solute-solvent interactions are not crucial in this case.

5.1.4. *Conclusions*

The results presented in this chapter suggest that the choice of the solvation model for studying the conformational preferences of polythiophenes in solution mainly depends on the solvent. For typical organic solvents, continuum SCRF calculations are able to provide satisfactory results, as have been demonstrated in the present study for acetonitrile. Thus, no specific interaction between the solute and the solvent molecules located at the first solvation shell is formed in these cases. On the contrary, more sophisticated combined discrete/SCRF models are required to provide good results in aqueous solution. This method, which requires huge amounts of computer resources, considers both the effects produced by the specific interactions associated to the first solvation shell and the solute polarization induced by the bulk solvent.

5.1.5. *References*

[1] Toribio T.F., Sansiñena J.M. *Adv. Mater.* **1998**, *10*, 491.

[2] Saraç A.S., Sönmez G., Cebeci F.Ç. *J. Appl. Electrochem.* **2003**, *33*, 295.

[3] Granström M., Inganäs O. *Adv. Mater.* **1995**, *7*, 1012.

[4] Roncali J. *Chem. Rev.* **1992**, *92*, 711.

[5] Heeger A.J. *Synth. Met.* **1993**, *57*, 3471.

[6] Goenendaal L.B., Jonas F., Freitag D., Pielartzik H., Reynolds J.R. *Adv. Mater.* **2000**, *12*, 481.

[7] Hernández V., López-Navarrete J.T. *J. Chem. Phys.* **1994**, *101*, 1369.

[8] Hutchison G.R., Zhao Y-J., Delley B., Freeman A.J., Ratner M.A., Marks T.J. *Phys. Rev. B* **2003**, *68*, 035204.

[9] Cornil J., Gueli I., Dkhissi A., Sancho-García J.C., Hennebicq E., Calbert J.P., Lemaur V., Beljonne D., Brédas J.L. *J. Chem. Phys.* **2003**, *118*, 6615.

[10] Tamm T., Tamm J., Karelson M. *Synth. Met.* **2005**, *149*, 47.

[11] Alemán C., Casanovas J. *J. Phys. Chem. A* **2004**, *108*, 1440.

[12] Brédas J.L., Beljonne D., Cornil J., Calbert J.Ph., Shuai Z., Silbey R. *Synth. Met.* **2002**, *125*, 107.

[13] Yang S., Olishevski P., Kertesz M. *Synth. Met.* **2004**, *141*, 171.

[14] Hernández V., López-Navarrete J.T. *Synth. Met.* **1996**, *76*, 221.

[15] Scherlis D.A., Marzari N. *J. Phys. Chem. B* **2004**, *108*, 17791.

[16] Alemán C. *Chem. Phys.* **1999**, *244*, 151.

[17] Tomasi J., Persico M., *Chem. Rev.* **1994**, *94*, 2027.

[18] Alemán C. *Chem. Phys. Lett.* **1999**, *302*, 461.

[19] Contador J.C., Aguilar M. A., Olivares del Valle F. J. *Chem. Phys.* **1998**, *232*, 151.

[20] Alemán C., Oliver R., Brillas E., Casanovas J., Estrany F. *Chem. Phys.* **2005**, *314*, 1.

[21] Brillas E., Oliver R., Estrany F., Rodríguez E., Tejero S. *Electrochim. Acta* **2002**, *47*, 1623.

[22] Kim B.S., Fukuola H., Gong J. P., Osada Y. *Eur. Polym. J.* **2001**, *37*, 2499.

[23] Rasmussen S.C., Pickens J.C., Hutchinson J. E. *Macromolecules* **1998**, *31*, 933.

[24] Boys S.F., Bernardi F. *Mol. Phys.* **1970**, *19*, 533.

[25] Wong M.M., Wiberg K.B., Fisch M.J. *J. Am. Chem. Soc.* **1992**, *114*, 1645.

5.2. *Ab initio* calculations on π-stacked thiophene dimer, trimer and tetramer: structure, interaction energy, cooperative effects and intermolecular electronic parameters*

5.2.1. *Introduction*

Conducting polymers are among the more studied novel families of materials because of their wide range of potential technological applications, which involve very different fields such as electronics, biomedical engineering or optics.[1] In particular thiophene-based conducting polymers have been widely studied because of their electric and nonlinear optical properties, excellent environmental stability and relatively low cost of production.[2] Polythiophene derivatives present a characteristic polyconjugated nature, which allows the intramolecular delocalization of π-electrons and, consequently, the reduction of the band-gap.[3] The electronic delocalization along the polymer chain has been widely investigated. Theoretical studies based on the application of ab initio quantum mechanical methods to model oligomers typically containing from 2 to 8 thiophene-rings evidenced that electronic delocalization and electrical conductivity increase with the size of the system.[4]

In addition to this intramolecular effect, polythiophene derivatives present electronic intermolecular delocalization, which contributes to electron transport from chain to chain in the conduction process.[5,6] Nevertheless, the number of studies devoted to investigate and understand the electronic interaction between thiophene rings belonging to different polymers chains, or equivalently thiophene rings belonging to the same chain but very far in the space, is very scarce.[5,6] Furthermore, it should be noted that very sophisticated ab initio calculations, which need huge amount of computational resources, are required to provide an accurate theoretical description of the interactions that take place between conjugated π-systems placed either in parallel (cofacial), tilted or herringbone (slipped) arrangement.[5,7]

π-Stacking interactions, *i.e.* attractive noncovalent interactions between π systems, play an important role in the stabilization of many different assemblies that contain aromatic moieties.[5,8-11] *Ab initio* calculations including electron correlation corrections have been widely used by theoretical researchers to study the π-stacking interactions.[5,6e,7,8c,12,13] Coupled cluster calculations with single and double substitutions with noniterative triple

* Results presented in this chapter are published in *J. Comput. Chem.* **2007, *29,*** 69

excitations,[14] or CCSD(T), have been reported as the best method to describe the intermolecular interactions in systems involving aromatic π-stacking.[7a,d,12] However, CCSD(T) calculations require huge amount of computational resources, which restricts the applicability of this method to small aromatic systems.[5,7a,d,12] Due to this difficulty, in practice the π-stacking of aromatic molecules is studied using the second order Møller-Plesset perturbation (MP2) calculations,[15] even although it is well known that this method tends to overestimate the strength of such interaction.[6e,8c,7d-f,12b-d,13]

In an early study Tsuzuki *et al.*[5] studied the origin of the π-stacking interaction of the thiophene dimer using very large basis sets near saturation and estimating the MP2 and CCSD(T) level interaction energies at the basis set limit. They concluded that the interaction energy of the perpendicular dimer is favored with respect to that of the parallel dimer by 1.4 kcal/mol. The same authors also studied the origin of attraction and directionality of the π-stacking in the benzene dimer.[7b] The interaction energy of the parallel dimer was predicted to be lower than that of the perpendicular dimer by about 1.0 kcal/mol, both the dispersion and electrostatic interactions being important for the directionality of the benzene dimer. In opposition, a recent study revealed that the stacked toluene dimers are substantially more stable than the perpendicular toluene dimer.[12e] On the other hand, MP2 and CCSD(T) calculations have been used to examine the π-stacking of nucleic acid bases, which are essential in DNA and RNA architectures.[16]

In this chapter, investigation of the π-stacking interaction in thiophene dimer, trimer and tetramer is investigated. For this purpose, the potential energy surfaces of systems containing two, three and four thiophene rings were explored at the MP2, MP3 and MP4 levels. The 6-31+G(d,p) basis set [17] has been used for MP2 and MP3 calculations, while the small 6-31G basis set has been used to evaluate the fourth order contribution. Calculations on the dimer were also performed at the CCSD(T) level combined with the 6-31+G(d,p) basis set. In all cases, calculations have been restricted to the planar sandwich (π-stacked) arrangement of rings, which is the most relevant for conducting polymers as polythiophene.[6a,b,c,18,19] Thus, in a very recent study it was found that amorphous polythiophene can be described as a packing of elongated molecular chains more or less aligned in the same direction, in which the thiophene rings close in the space but belonging to different chains tend to adopt π-stacked arrangements.[19] Details about the influence of the number of thiophene rings on the structure, interaction

energies, nonadditive effects in the interaction energies and the lowest transition energy are given and discussed.

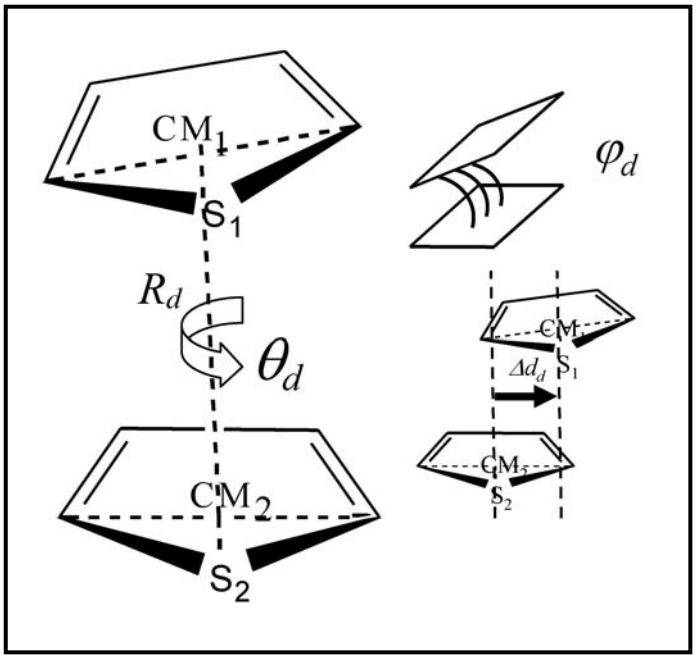


Figure 1. Structure of the π-stacked thiophene dimer. Relevant intermolecular parameters are indicated: R_d corresponds to the distance between the centers of masses of the two thiophene monomers, θ_d is the torsion angle defined by the sulfur atoms and the center of mass of the two monomers, i.e. S···CM_1···CM_2···S, and φ_d is the angle formed by the two planar rings.

5.2.2. *Computational methods*

All calculations were performed using the Gaussian 03 computer program.[20] Electron correlation was accounted for at the MP2, MP3, MP4(SDTQ)[15] (hereafter abbreviated MP4) and CCSD(T)[14] levels. All geometry optimizations were performed at the MP2 level combined with the 6-31+G(d,p) basis set.[17] The geometry of an isolated thiophene molecule was fully optimized and used for the subsequent calculations on π-stacked dimers, trimers and tetramers. The planar sandwich arrangement considered for calculations on thiophene dimer is schematized in Figure 1, where the distance R_d refers to the center of mass of the two thiophene monomers, θ_d is the virtual dihedral angle defined by the sulfur atoms and the center of mass of the two monomers, *i.e.* S···CM_1···CM_2···S, and φ_d corresponds to the angle formed by the two thiophene rings. Accordingly, R_d, θ_d and φ_d indicates the separation between the two thiophene rings, their relative orientation (the rings are parallel and antiparallel when θ_d= 0º and 180º,

respectively) and the degree of tilting (the two rings are coplanar when φ_d= 0º, while φ_d=180º corresponds to a perpendicular T-shaped configuration), respectively. Constrained geometry optimizations on the thiophene dimer were performed by spanning the dihedral θ_d from 0º to 180º in steps of 30º, the intermolecular parameters R_d and φ_d being optimized for each value of θ_d while the monomer geometries were kept fixed. Furthermore, an additional geometry optimization was carried out to relax all the intermolecular geometric parameters using as starting point the lowest energy θ_d value. Single point calculations at the MP3/6-31+G(d,p), MP3/6-31G, MP4/6-31G and CCSD(T)/6-31+G(d,p) levels were performed on all the thiophene dimer arrangements using the geometries optimized at the ab initio MP2/6-31+G(d,p) level. The MP4 energy contribution estimated with the 6-31G basis set, which was calculated as [E(MP4/6-31G) – E(MP3/6-31G)], was added to the MP3/6-31+G(d,p) energy, the label used to identify this level of theory being MP4#/6-31+G(d,p).

The lowest energy structure of the π-stacked thiophene dimer was used as starting point for calculations on thiophene trimer. Thus, a third thiophene ring was placed over the relaxed thiophene dimer considering a planar sandwich arrangement with respect to the first ring. The structure of the initial thiophene dimer was kept fixed during all the calculations on the trimer. The position of the third thiophene with respect to the first ring of the dimer was defined by R_t, θ_t and φ_t, whose meaning is identical to that described above for the dimer (Figure 1). Constrained geometry optimizations on the thiophene trimer were performed by spanning the dihedral θ_t from 0º to 360º in steps of 30º, the intermolecular parameters R_t and φ_t being optimized for each value of θ_t while the internal geometry of the thiophene ring was kept also fixed. As occurs for the dimer, the lowest energy thiophene trimer was obtained by optimizing the three intermolecular parameters R_t, θ_t and φ_t. The resulting structure was used to generate the thiophene tetramer by adding a new thiophene ring in a π-stacked arrangement with respect to the first thiophene. The position of the fourth thiophene was defined with respect to the first ring of the trimer by using the intermolecular parameters R_q, θ_q and φ_q, whose meaning is analogous to those described above for the dimer and the trimer. The calculations performed on the thiophene tetramer were similar to those described above for the thiophene trimer.

Single point calculations at the MP3/6-31+G(d,p) and MP4#/6-31+G(d,p) levels were performed on all the thiophene trimer arrangements using the geometries optimized at the ab initio MP2/6-31+G(d,p) level. Furthermore, a single point energy calculation at the MP3/6-31+G(d,p) level was performed on the minimum energy arrangement of the thiophene tetramer obtained at the MP2/6-31+G(d,p) level. It should be noted that this calculation, which required a huge amount of computer resources, deals with the three-body part of the dispersion nonadditivity in the tetramer.

The interaction energies (ΔE_{Int}) of the thiophene dimer, trimer and tetramer were calculated by correcting the basis set superposition error (BSSE), which comes from the adoption of a limited basis set, with the counterpoise method (CP).[21] In the latter procedure the total energy of each thiophene monomer is calculated including the basis set of the other monomers in the complex. The cooperative energy (ΔE_{coop}) for the trimer and the tetramer was estimated as the difference between the quantum mechanical and the expected interaction energies. The expected interaction energy (ΔE^{e}_{Int}) was supplied by the sum of the counterpoise-corrected ΔE_{Int} of all the dimers contained in the complex. Accordingly, ΔE_{coop} provides an evaluation of the many-body (non-aditive) effects.

$$\Delta E_{coop} = \Delta E_{Int} - \Delta E^{e}_{Int} \tag{8.2.}$$

$$\Delta E_{Int} = E(complex) - \sum E(monomers) \tag{8.3.}$$

$$\Delta E^{e}_{Int} = \sum_{dimers} \Delta E_{Int} \tag{8.4.}$$

5.2.3. *Results and discussion*

Figure 2a depicts the MP2/6-31+G(d,p) relative energy of the thiophene dimer as a function of the dihedral angle θ_d. Surprisingly, the dependence of the energy of the dimer on the relative orientation of the two rings is very small. Thus, the parallel (θ_d= 0º) and antiparallel (θ_d= 180º) arrangements differ by only 0.2 kcal/mol. The minimum energy arrangement, which was derived from optimization of all the intermolecular parameters, corresponds to θ_d= 57.2º. However, this configuration is favored by only 0.4 kcal/mol with respect to the least stable (θ_d= 150º). Furthermore, the ideal sandwich

arrangement used as starting geometry changes in all cases to a displaced π-stacked configuration, the displacement of the center of mass of the second thiophene with respect to that of the first one (Δd_d) ranging from 1.132 (θ= 120º) to 1.845 Å (θ= 30º). The displacement of the rings is clearly illustrated for the minimum energy structure in Figure 3a, where Δd_d= 1.387 Å. The minimum energy structure found for the thiophene dimer is similar to those found for other complexes formed by two polar aromatic molecules.[22]

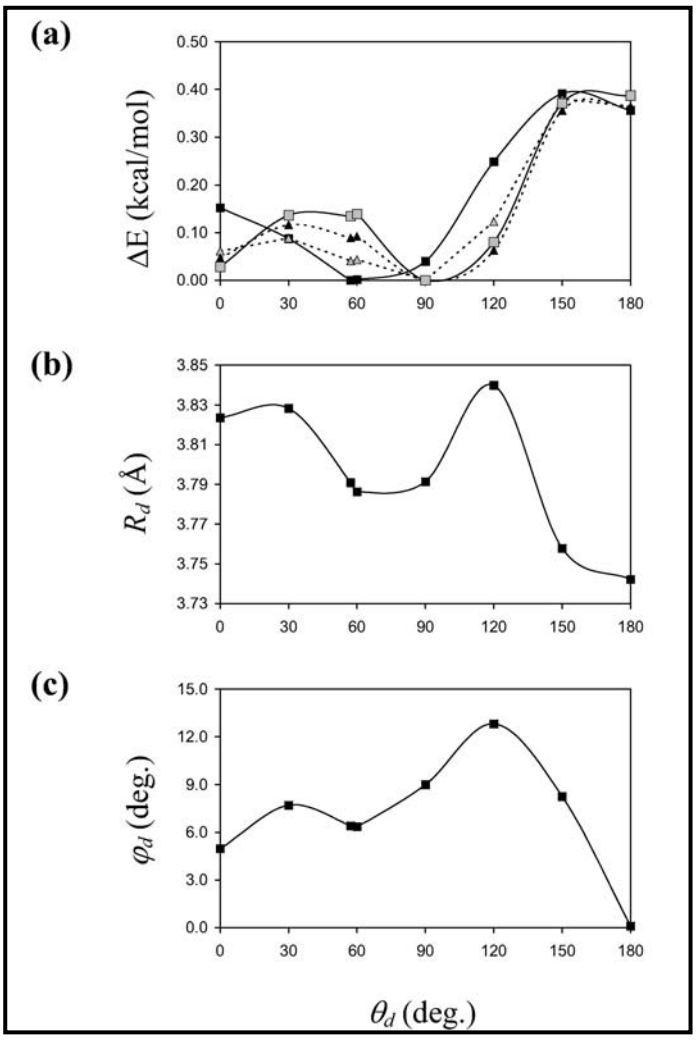


Figure 2. MP2/6-31+G(d,p) (black squares-solid line), MP3/6-31+G(d,p) (black triangles-dashed line), MP4#/6-31+G(d,p) (gray squares-solid line) and CCSD(T)/6-31+G(d,p) (gray triangles-dashed line) relative energies (a), R_d (b) and φ_d (c) of the thiophene dimer as a function of the dihedral angle θ_d.

Figures 2b and 2c represent the values of R_d and φ_d for the optimized dimers, respectively, against the dihedral angle θ_d. As can be seen, R_d ranges from 3.742 Å (θ_d= 180º) to 3.840 Å (θ_d= 120º), the value of the minimum energy structure being 3.791 Å.

The latter value is similar to that predicted for the antiparallel thiopene dimer[5] and the parallel benzene dimer[7b] using the CCSD(T) method combined with the aug(d)-6-311G(d) basis set. On the other hand, it is worth noting that the degree of tilting is relatively small independently of θ_d, *i.e.* φ_d ranges from 0.1º (θ_d= 180º) to 12.8º (θ_d= 120º). However, it should be noted that the potential energy surface is very shallow and, therefore, no quantitative discussion about the degree of tilting is possible. The parameters obtained for the lowest energy minimum of the thiophene dimer are compiled in Table 1.

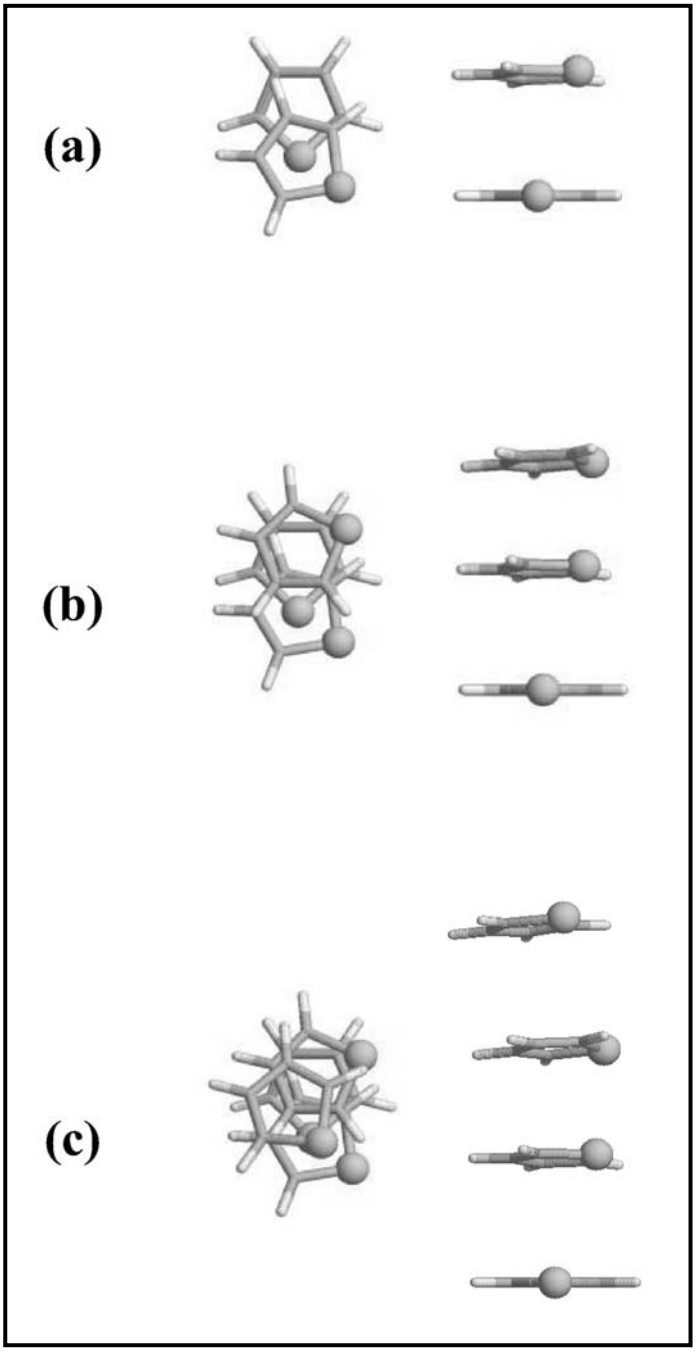


Figure 3. Lowest energy arrangement of the π-stacked thiophene dimer (a), trimer (b) and tetramer (c): equatorial view (left) and axial view (right).

Figure 4 represents the counterpoise-corrected ΔE_{Int} as a function of the dihedral angle θ_d. In this case there is a quantitative discrepancy between the results provided by the four theoretical levels. Thus, ΔE_{Int} ranges from -1.55 to -1.76 kcal/mol at the MP2 level, from +0.43 to +0.71 kcal/mol at the MP3 level, from -0.16 to +0.06 kcal/mol at the MP4# level, and from -0.12 to -0.34 kcal/mol at the CCSD(T) level. These results are consistent with observations previously reported: MP2 significantly overestimates the ΔE_{Int} compared to CCSD(T).[5,7b,e,12a,e] On the other hand, MP3 tends to underestimate the strength of the interaction with respect to MP4# and CCSD(T), the results provided by the later two levels being fully consistent. In spite of this discrepancy, it should be noted that the four methods predict that the influence of the relative orientation of thiophene rings is very small, showing a qualitative agreement.

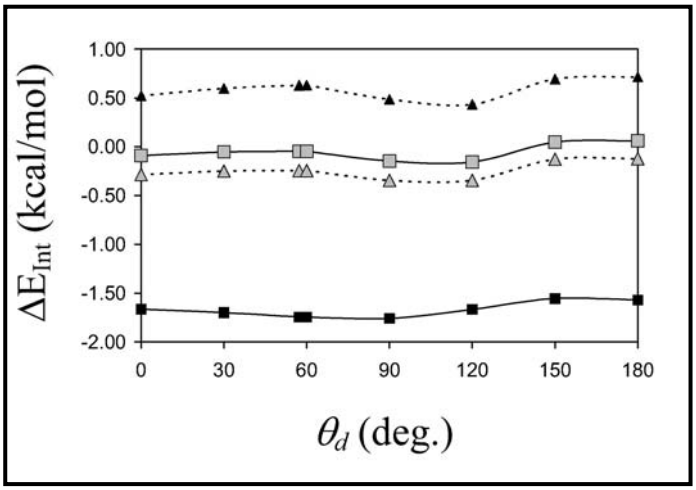


Figure 4. MP2/6-31+G(d,p) (black squares-solid line), MP3/6-31+G(d,p) (black triangles-dashed line), MP4#/6-31+G(d,p) (gray squares-solid line) and CCSD(T)/6-31+G(d,p) (gray triangles-dashed line) interaction energies of the thiophene dimer as a function of the dihedral angle θ_d.

The complete basis set MP2 value of ΔE_{Int} has been obtained for the lowest energy minimum of the thiophene dimer by a simple extrapolation of the interaction energies calculated with the Aug-cc-pVXZ (X= D, T and Q) basis sets versus 1/*N*, where *N* is the number of contracted atomic orbitals. It was found that the complete basis set MP2 value is -5.09 kcal/mol, which indicates that the ΔE_{Int} is significantly underestimated by the 6-31+G(d,p) basis set. On the other hand, inspection to the literature reveals that the ΔE_{Int} provided by Tsuzuki *et al.*[5] for the π-stacked thiophene dimer is -1.71 kcal/mol, this value being estimated by using an aromatic intermolecular interaction model (AIMI) based on CCSD(T) interaction energies at the basis set limit. It should be noted

that this value is in excellent agreement with the MP2/6-31+G(d,p) estimation. This is due to a cancellation of errors: the enlargement of the stabilization MP2 energy when the basis set evolves from 6-31+G(d,p) to the CBS is compensated by the repulsive correction term provided by the CCSD(T)/6-31+G(d,p) method. Moreover, the ΔE_{Int} values estimated for the π-stacked benzene[7b] and toluene[12e] dimers using the same CCSD(T)/basis set limit-based AIMI model were -1.48 and -2.18 kcal/mol, respectively. Unfortunately, the application of the computational demanding CCSD(T) procedures with a basis set larger than the 6-31+G(d,p) is a major obstacle to study complexes formed by two or more aromatic molecules.

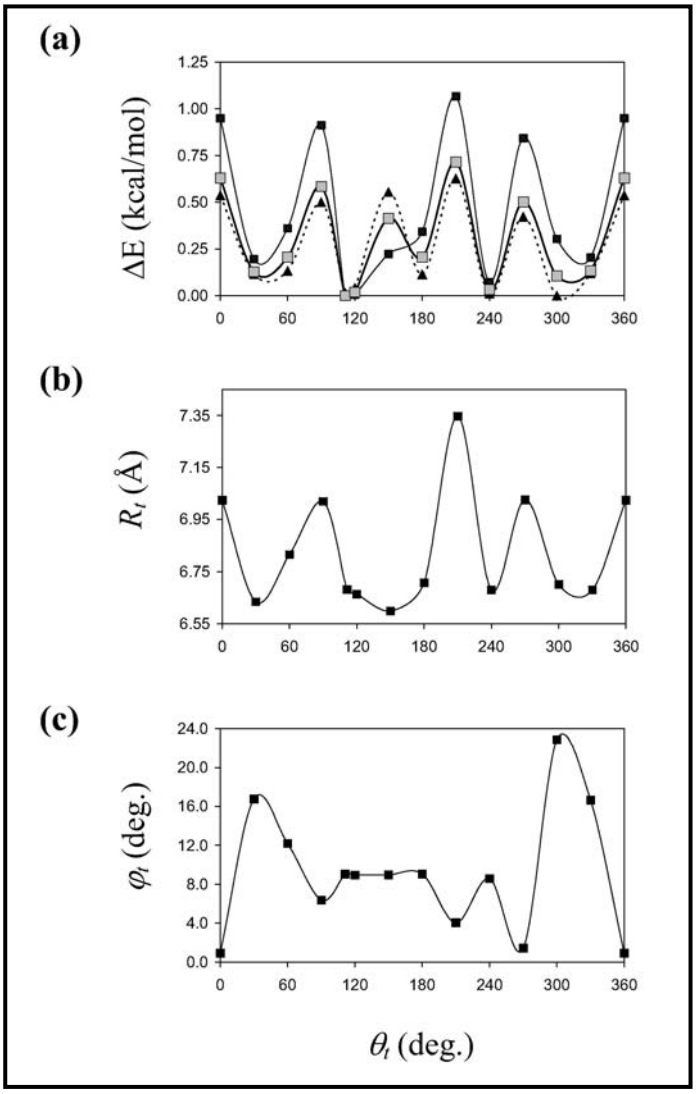


Figure 5. MP2/6-31+G(d,p) (black squares-solid line), MP3/6-31+G(d,p) (black triangles-dashed line) and MP4#/6-31+G(d,p) (gray squares-solid line) relative energies (a), R_t (b) and φ_t (c) of the thiophene trimer as a function of the dihedral angle θ_t.

The lowest energy structure of the dimer was used to build the trimer by adding a new thiophene above the second and aligning it with respect to the first ring, *i.e.* the initial displacement of the center of mass of the third thiophene with respect to that of the first one (Δd_t) was zero. Figure 5a represents the relative energies of the thiophene trimer calculated at the MP2/6-31+G(d,p), MP3/6-31+G(d,p) and MP4#/6-31+G(d,p) levels as a function of the dihedral angle θ_t. It should be noted that the dependence on the relative orientation of the third thiophene ring with respect to the first one is higher than that predicted for the dimer (Figure 2a), especially at the MP2 level. Thus, although low-energy orientations are clearly identified at around $\theta_t \approx 30º$, 120º, 180º, 240 and 330º, the trimers with $\theta_t \approx$ 0º, 90º, 150º, 210º and 270º are unfavored by about 0.5-1.0 kcal/mol depending on the computational method. Thus, the relative energies calculated at the MP2 level are considerably overestimated with respect those predicted using the MP3 and MP4# methods at these high-energy θ_t values. This dependence is also evidenced by the variation of R_t against θ_t, which is showed in Figure 5b. As can be seen, the lower R_t values (~ 6.65 Å) correspond to the low-energy θ_t values, while R_t is higher than 7 Å for the less favored orientations. Indeed, the largest R_t value, 7.35 Å, corresponds to the least favored orientation with $\theta_t = 210º$. Regarding to the tilting of the third ring, the behavior is more complex (Figure 5c), even although no quantitative discussion is possible at this level of theory. On the other hand, as previously occurs for the dimer, the ideal π-stacked configuration of the third thiophene with respect to the first changed to a displaced π-stacked arrangement, the maximum value of Δd_t (2.279 Å) being detected for $\theta_t = 210º$.

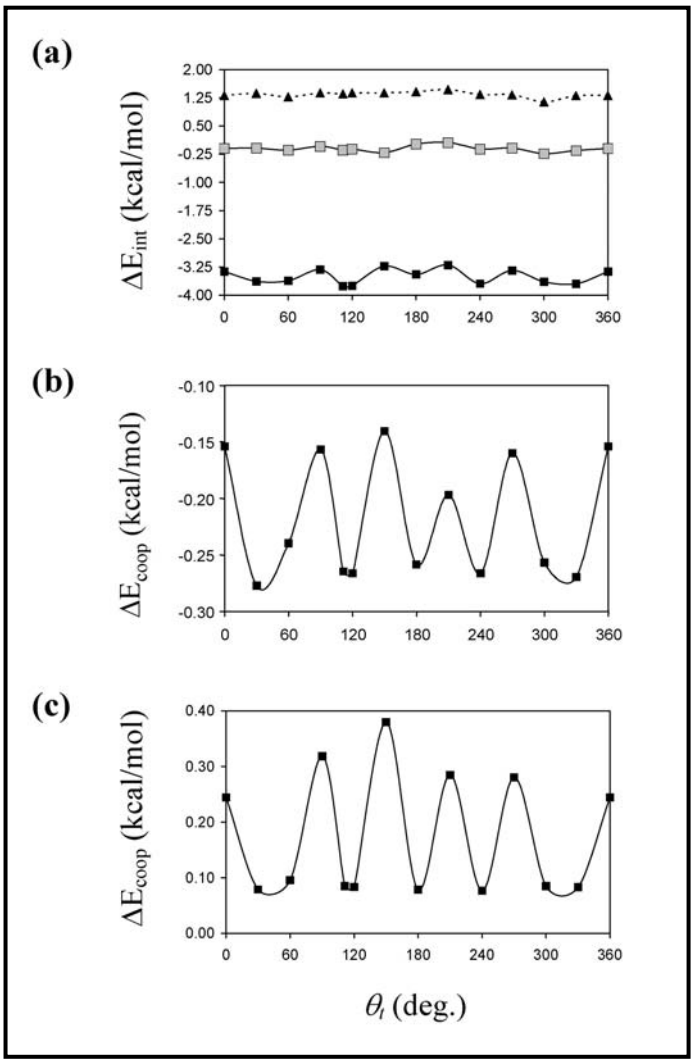


Figure 6. MP2/6-31+G(d,p) (black squares-solid line), MP3/6-31+G(d,p) (black triangles-dashed line) and MP4#/6-31+G(d,p) (gray squares-solid line) interaction energies (a) and cooperative energy at the MP2/6-31+G(d,p) (b) and HF/6-31+G(d,p) (c) levels for the thiophene trimer as a function of the dihedral angle θ_t.

The lowest energy arrangement was found at θ_t= 111.3º (Table 1), which represents a rotation of about 54º with respect to the second thiophene ring. The latter value is close to the θ_d predicted for the lowest energy structure of the dimer. Accordingly, a progressive rotation of about 55-60º is predicted for each thiophene ring that incorporates to the complex. This is reflected in Figure 3b, which shows the relative orientation of the thiophene rings in the lowest energy structure of the trimer.

The counterpoise-corrected ΔE_{Int} values, which are displayed in Figure 6a, evidence a discrepancy between the three computational methods. The MP2 values range from -3.20 to -3.76 kcal/mol depending on the value of θ_t, the latter value being that obtained for the lowest energy trimer (Table 1). This range of variation is about two times larger

than that previously observed at the same level of theory for the thiophene dimer (Figure 4), suggesting that the relative orientation of the third ring affects not only to the relative energy of the trimer but also to the strength of the binding. However, MP3 calculations, which are able to describe the non-additivity dispersion in trimers, predict a repulsive interaction, ΔE_{Int} values ranging from +1.14 to +1.47 kcal/mol. This behavior is fully consistent with that displayed in Figure 4. Finally, the values of ΔE_{Int} predicted at the MP4# level reflect a very weak attractive interaction for all the orientations, *i.e.* $\Delta E_{Int} \approx$ -0.15 kcal/mol, with exception of θ_t= 180º and 210º for which no-interaction was detected.

On the other hand, it should be noted that the ΔE_{Int} values calculated for the trimer at the MP2/6-31+G(d,p) level (Figure 6a) are consistent with attractive many body effects, suggesting that intermolecular π-stacking interactions are cooperative. This is reflected in Figure 6b, which represents the ΔE_{coop} predicted at this level of theory as a function of θ_t. As can be seen, the interaction provided by MP2 calculations is stronger than that derived from the addition of dimer interactions suggesting that the strength of the intermolecular π-stacking interaction is enhanced upon increasing the number of thiophene rings in the complex. However, it should be emphasized that the MP2 level of theory only includes non-additive electrostatic (induction) contributions, the n-body effects predicted by this method being only reliable for systems with insignificant dispersion contribution. Obviously, this is not the case of stacked thiophene arrangements as those studied in this work. Accordingly, MP3 and MP4# results (Figure 6a) evidence that the n-body effects become positive or almost completely cancelled by adding of the non-additive dispersion component. Furthermore, it was also found that the relative energies and the counterpoise-corrected ΔE_{Int} calculated at the MP4(SDQ)/6-31+G(d,p) level (data not shown) for both the dimer and the trimer are in excellent (quantitative) agreement with the MP3 results displayed in Figures 2a, 4a, 5a and 6a, this feature being fully consistent with the previous discussion. On the other hand, Figure 6c shows the ΔE_{coop} calculated at the HF/6-31+G(d,p) level as a function of θ_t. As it was expected, the increase of the number of thiophene rings in the complex has an unfavorable effect since induction is the main contribution at the HF level.

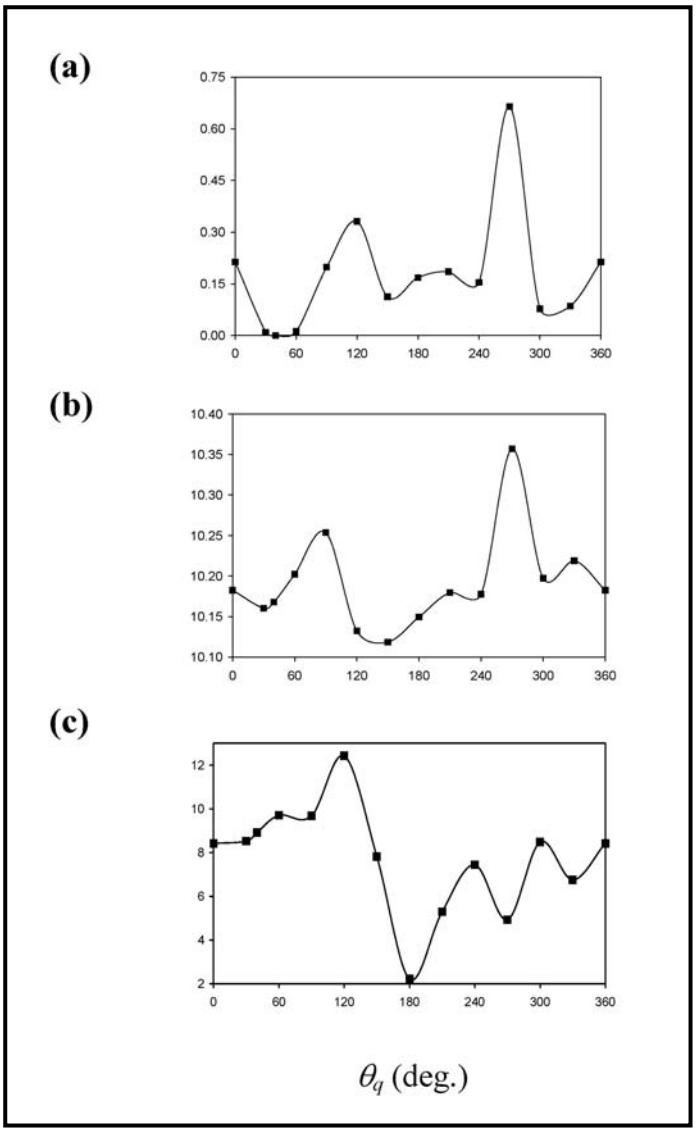


Figure 7. MP2/6-31+G(d,p) relative energy (a), R_q (b) and φ_q (c) of the thiophene tetramer as a function of the dihedral angle θ_q.

The minimum energy structure of the thiophene trimer was used to generate a tetramer by placing a thiophene above the third ring and considering an ideal sandwich arrangement with respect to the first ring, *i.e.* the initial displacement of the center of mass of the fourth thiophene with respect to that of the first one (Δd_q) was zero. Figure 7 represents the MP2/6-31+G(d,p) relative energy of the thiophene tetramer, R_q and φ_q as a function of the dihedral angle θ_q. As can be seen, the behavior of the tetramer is similar, from a qualitative point of view, to that predicted for the dimer at the same theoretical level. Thus, the relative energies are smaller than 0.33 kcal/mol with exception of θ_q= 270º (0.66 kcal/mol) indicating that the orientation of the fourth ring does not play a crucial role on the stability of the tetramer. Unfortunately, single point

calculations at the MP3/6-31+G(d,p) and MP4#/6-31+G(d,p) levels on the optimized geometries were not possible in this case due to the large size of this π-stacked complex. On the other hand, the range of variation of both R_q and φ_d are relatively small (from 10.160 to 10.357 Å and from 2.2º to 12.4º, respectively), similar trends being also detected for R_d and φ_d in the dimer (Figures 2b and 2c, respectively).

Figure 3c shows the lowest energy arrangement of the thiophene tetramer, which was reached at θ_q= 39.9º (Table 1). As can be seen, the relative orientation of the fourth ring in the tetramer is similar to that achieved for the second ring in the dimer suggesting the existence of some degree of periodicity. On the other hand, a parallel displacement of the center of mass of the fourth ring with respect to that of the first one (Δd_q) was detected in all the optimized structures, the parameter ranging from 0.460 Å (θ_q= 60º) to 1.867 Å (θ_q= 270º). The value of Δd_q for the minimum energy structure was 0.656 Å.

The counterpoise-corrected ΔE_{Int} values calculated at the MP2/6-31+G(d,p) level as a function of the dihedral angle θ_q are represented in Figure 8a. As can be seen, the influence of the relative orientation of the fourth thiopene ring is not significant, ΔE_{Int} ranging from –5.51 (θ_q= 270º) to –5.79 (θ_q= 60º). This range of variation (0.28 kcal/mol) is very similar to that obtained for the dimer at the same theoretical level (0.21 kcal/mol) and considerably lower than that predicted for the trimer (0.56 kcal/mol). Furthermore, it should be noted that the addition of the fourth thiophene ring to the trimer produces an increase of more than 2 kcal/mol in the MP2 estimation of ΔE_{Int}, this value being larger than the ΔE_{Int} obtained for the dimer. This feature indicates that induction interactions, which are rightly reproduced at this level of theory, contribute to the stability of the system. However, as previously occurred for the trimer, the role played by the dispersion contribution is essential also for the tetramer. The value of ΔE_{Int} estimated at the MP3/6-31+G(d,p) level for the minimum energy arrangement of the tetramer, *i.e.* that with θ_q= 39.9º, is +2.10 kcal/mol (Table 1) evidencing that the three-body part of non-additive dispersion interaction is positive.

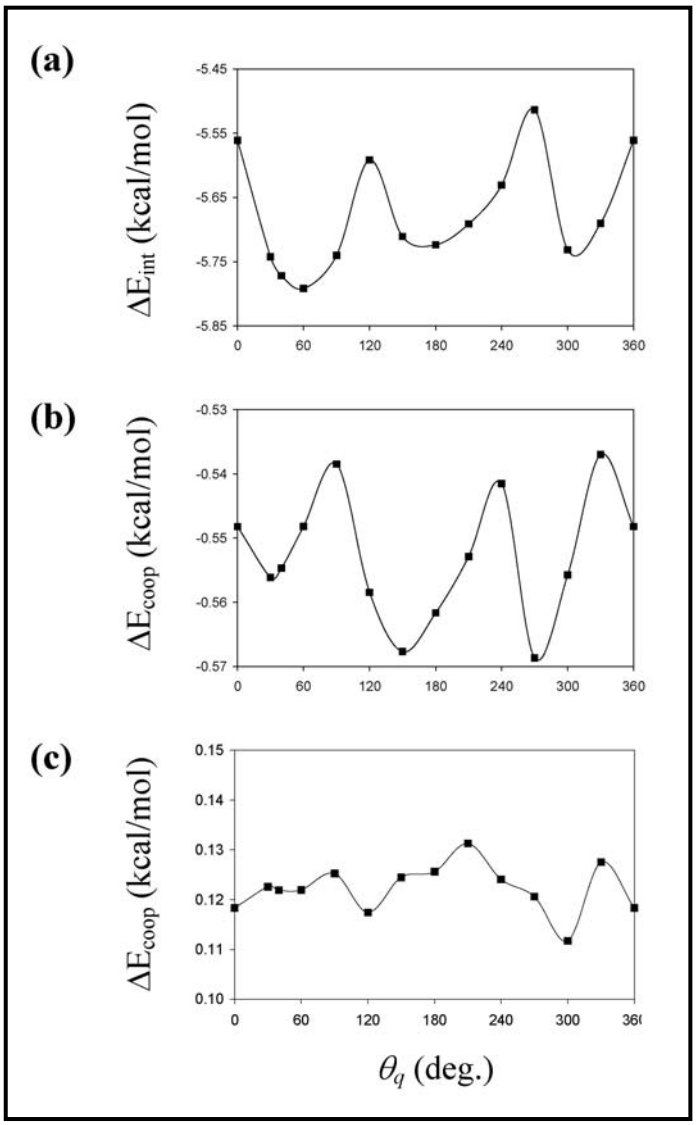


Figure 8. Interaction energy at the MP2/6-31+G(d,p) level (a) and cooperative energy at the MP2/6-31+G(d,p) (b) and HF/6-31+G(d,p) (c) levels for the thiophene tretramer as a function of the dihedral angle θ_q.

Figure 8b represents the variation of the MP2 estimation of ΔE_{coop}, which ranges from –0.54 to –0.57 kcal/mol, against θ_q. As expected, the cooperative effects predicted for the tetramer at this level of theory are higher than those predicted for the trimer (Figure 6b) by more than two times indicating that, although the three-body electrostatic interactions between the 1^{st}-3^{rd} and 2^{nd}-4^{th} rings provide the most significant contribution to ΔE_{coop}, there is also a very small electrostatic contribution of the four-body interaction between the 1^{st}-4^{th} rings. However, as occurred above for the trimer, n-body non-additive dispersion contributions are expected to compensate these favorable

electrostatic interactions. The repulsive cooperative effects predicted by HF/6-31+G(d,p) are displayed in Figure 8c.

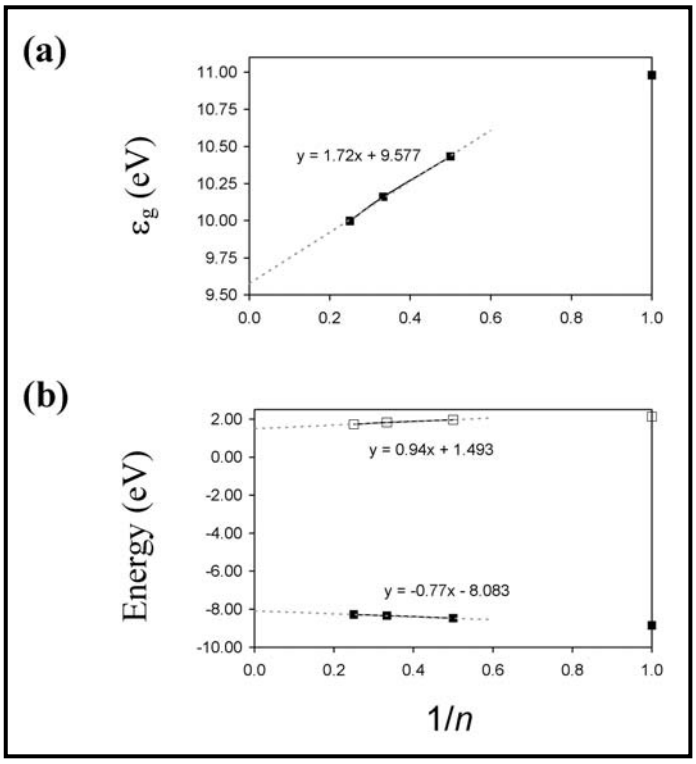


Figure 9. Evolution of the lowest transition energy (a) and HOMO and LUMO energies (b) with inverse number of thiophene rings (1/n) using the lowest energy structure of each complex.

The formation of discrete π-stacked thiophene aggregates is expected to facilitate the intermolecular electron transfer.[6] Thus, although higher conductivities are found for polythiophene derivatives in the oxidized state, the presence of intermolecular π-stacked structures in these neutral conjugated chain polymers is expected to affect the electronic parameters related with the electron transport.[6b,c] Figure 9a represents the variation of the lowest transition energy (ε_g), which was approximated as the difference between the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) energies, with the inverse of the number of thiophene rings in the lowest energy π-stacked complexes (1/*n*). It is worth nothing that ε_g decreases with *n*, a linear behaviour being followed for $n \geq 2$. A linear regression analysis, excluding the value for *n*= 1, was performed to extrapolate the ε_g value for an infinite π-stacked aggregate. The result, ε_g= 9.58 eV (Figure 9a), indicates a reduction of 1.40 eV with respect to the value calculated for a single thiophene ring, *i.e.* ε_g= 10.98 eV for *n*=1. On the other

hand, within a given complex the range of variation of ε_g with the relative orientation of the rings was small for the thiophene dimer (from 10.33 to 10.60 eV), trimer (from 10.03 to 10.19 eV) and tetramer (from 9.96 to 10.03 eV).

Figure 9b shows the variation of the HOMO and LUMO energies with 1/*n*. As can be seen, the stability of the HOMO decreases with the *n*, while that of the LUMO increases with the size of the aggregate. Thus, the reduction of ε_g is consequence of the dependence of both the HOMO and LUMO energies on the π-stacking interactions. The values of the orbitalic energies extrapolated for an infinite π-stacked complex, which were obtained through a linear regression of the data predicted for $n \geq 2$, are -8.08 and 1.49 eV for the HOMO and LUMO, respectively. Thus, the latter becomes 0.64 eV more stable than in a single thiophene ring, while the latter destabilizes 0.76 eV.

5.2.4. *Conclusions*

Ab initio quantum chemical methods have been used to investigate both the structure and the energetics of the thiophene π-stacked dimer, trimer and tetramer. The dimer exhibits a slightly displaced π-stacked geometry with a very small tilting, the interaction energy having a small dependence on the relative orientation of the two rings. The importance of considering more than two thiophene rings in this kind of studies has been already revealed in the trimer analysis. Thus, geometry optimizations on the trimer showed a higher dependence on the relative orientation of the third thiophene ring with respect to the first one than that predicted for the dimer. Moreover, the influence of many-body electrostatic and dispersion contributions in the strength of the π-stacking interaction is essential. Furthermore, calculations on the tetramer suggest that, although the relative orientation of the rings does not determine drastically the stability of the system, the orientation presents a preferred degree of periodicity. The overall of the results also reveals that the formation of π-stacked thiophene aggregates may contribute to improve the properties related with electron transport by reducing the lowest transition energy.

5.2.5. *References*

[1] a) Toribio T. F., Sansiñena J.M. *Adv. Mater.* **1998**, *10*, 491. b) Saraç A.S., Sönmez G., Cebeci F.Ç. *J. Appl. Electrochem.* **2003**, *33*, 295. c) Granström M., Inganäs O. *Adv. Mater.* **1995**, *7*, 1012. d) Roncali J. *Chem. Rev.* **1992**, *92*, 711. e) Casanovas J., Zanuy D., Alemán C. *Angew. Chem.* **2006**, *7*, 1103. f) Goenendaal L.B., Jonas F., Freitag D., Pielartzik H., Reynolds J.R. *Adv. Mater.* **2000**, *12*, 481. g) Ocampo C., Armelin E., Liesa F., Alemán C., Ramis X., Iribarren J.I. *Prog. Org. Coat.* **2005**, *53*, 217.

[2] a) Heeger A.J. *Synth. Met.* **1993**, *57*, 3471. b) Casado J., Ortiz R.P., Delgado M.C.R., Azumi R., Oakley R.T., Hernández V., Navarrete J.T.L. *J. Phys. Chem. B* **2005**, *109*, 10115. c) Lukeš V., Breza M., Végh D., Hrdlovič P., Krajčovič J., Laurinc V. *Synth. Met.* **2005**, *129*, 85. d) Alvarado Y.J., Cubillan N., Labarca P.H., Karma A., Arrieta F., Castellano O., Soscún H. *J. Phys. Org. Chem.* **2002**, *15*, 154. e) Majumdar H.S., Botta C., Bolognesi A., Pal A. *J. Synth. Met.* **2005**, *148*, 175. f) Iosip M.D., Destri S., Pasini M., Porzio W., Pernstich K.P., Batlogg B. *Synth. Met.* **2005**, *146*, 251. g) Chan H.S.O., Ng S.C. *Prog. Polym. Sci.* **1998**, *23*, 1167. g) Jen A.K.Y., Cai Y.M., Bedworth P.V., Marder S.R. *Adv. Mater.* **1997**, *9*, 132. h) Barbarella G., Melucci M., Sotgiu G. *Adv. Mater.* **2005**, *17*, 1581.

[3] a) Oyaizu K., Iwasaki T., Tsukahara Y., Tsuchida E. *Macromolecules* **2004**, *37*, 1257. b) Kobayashi M., Chen J., Chung T.C., Moraes F., Heeger A. J., Wudl F. *Synth. Met.* **1984**, *9*, 77. c) Chung T.C., Kaufman J.H., Heeger A. J., Wudl F. *Phys. Rev. B* **1984**, *30*, 702. d) Tachibana M., Tanaka S., Yamashita Y., Yoshizawa K. *J. Phys. Chem. B* **2002**, *106*, 3549. e) Lee B., Seshadri V., Sotzing G.A. *Langmuir* **2005**, *21*, 10797.

[4] a) Chakraborty D., Lagowski J.B. *J. Chem. Phys.* **2001**, *115*, 184. b) Rubio M., Merchán M., Ortí E. *Chem. Phys. Chem.* **2005**, *6*, 1357. c) Puschnig P., Ambrosch-Draxl C. *Synth. Met.* **2001**, *119*, 245. d) van der Horst J.W., Bobbert P. A., de Jong P. H.L., Michels M.A.J., Brocks G., Kelly P.J. *Phys. Rev. B* **2000**, *61*, 15817. e) van der Horst J.W., Bobbert P.A., Michels M.A.J., Brocks G., Kelly P.J. *Phys. Rev. Lett.* **1999**, *83*, 4413.

[5] Tsuzuki S., Honda K., Azumi R. *J. Am. Chem. Soc.* **2002**, *124*, 12200.

[6] a) Miller L., Mann K.R. *Acc. Chem. Res.* **1996**, *29*, 417. b) Skotheim T.A. in *Handbook of Conducting Polymers*, M. Deker, New York, **1986**. c) Brédas J.L., Silbey R. in *Conjugated Polymers*, Kluwer, Dordrecht, The Netherlands, **1991**. d) Heeger A.J.,

Kilvelson S., Schrieffer J.R., Su W.P. *Rev. Mod. Phys.* **1988**, *60*, 781. e) Pickholz M., dos Santos M.C. *J. Mol. Struct. (Theochem)* **2005**, *717*, 99. f) Geskin V.M., Brédas J.L. *Chem. Phys. Chem.* **2003**, *4*, 498.

[7] a) Willnauer C., Birkenheuer U. *J. Chem. Phys.* **2004**, *120*, 11910. b) Tsuzuki S., Honda K., Uchimaru T., Mikami M., Tanabe K. *J. Am. Chem. Soc.* **2002**, *124*, 104. c) Hobza P., Šponer J. *J. Am. Chem. Soc.* 2002, 124, 11802. d) Tauer T.P., Sherrill C.D. *J. Phys. Chem. A* **2005**, *109*, 10475. e) Sinnokrot M.O., Valeev E.F., Sherrill C.D. *J. Am. Chem. Soc.* **2002**, *124*, 10887. f) Hutchison G.R., Ratner M.A., Marks T.J. *J. Am. Chem. Soc.* **2005**, *127*, 16866.

[8] a) Sirringhaus H., Brown P.J., Friend R.H., Nielsen M.M., Bechgaard K., Langeveld-Voss B.M.W., Spiering A.J.H., Janssen R.A.J., Meijer E.W., Herwig P., de Leeuw D.M. *Nature* **1999**, *401*, 685. b) Bjørnholm T., Greve D.R., Reitzel N., Hassenkam T., Kjaer K., Howes P.B., Larsen N.B., Bøgelund J., Jayaraman M., Ewbank P.C., McCullough R.D. *J. Am. Chem. Soc.* **1998**, *120*, 7643. c) Scherlis D.A., Marzari N. *J. Am. Chem. Soc.* **2005**, *127*, 3207. d) Yamamoto T., Kokubo H., Morikita T. *J. Polym. Sci. Pol. Phys.* **2001**, *39*, 1713.

[9] a) Marco E., Negri A., Luque F.J., Gago F. *Nucleic Acids Res*. **2005**, *33*, 6214. b) Isaksson J., Acharya S., Barman J., Cheruku P., Chattopadhyaya J. *Biochemistry* **2004**, *43*, 15996. c) Ieong S., Kao M.Y., Lam T.W., Sung W.K., Yiu S.M. *J. Comput. Biol.* **2003**, *10*, 981. d) Kim T.W., Kool E.T. *J. Org. Chem.* **2005**, *70*, 2048. e) Mignon P., Loverix S., Steyaert J., Geerlings P. *Nucleic Acids Res.* **2005**, *33*, 1779.

[10] Müller-Dethlefs K., Hobza P. *Chem. Rev.* **2000**, *100*, 143.

[11] a) Lane J.D., Lucocq J., Pryde J., Barr F., Woodman P.G., Allan V.J., Lowe M. *J. Cell. Biol.* **2002**, *156*, 495. b) Shiels J.C., Tuite J.B., Nolan S.J., Baranger A.M. *Nucleic Acids Res*. **2002**, *30*, 550. c) Blakaj D.M., McConnell K.J., Beveridge D.L., Baranger A. M. *J. Am. Chem. Soc.* **2001**, *123*, 2548.

[12] a) Hobza P., Selzle H.L., Schlag E.W. *J. Phys. Chem.* **1996**, *100*, 18790. b) Tsuzuki S., Uchimaru T., Matsumura K., Mikami M., Tanabe K. *Chem. Phys. Lett.* **2000**, *319*, 547. c) Jurecka P., Hobza P. *J. Am. Chem. Soc.* **2003**, *125*, 15608. d) Sponer J., Hobza P. *Chem. Phys. Lett.* **1997**, *267*, 263. e) Tsuzuki S., Honda K., Uchimaru T., Mikami M. *J. Chem. Phys.* **2005**, *122*, 144323.

[13] a) Jaffe R.L., Smith G.D. *J. Chem. Phys.* **1996**, *105*, 2780. b) Ye X.Y., Li Z.H., Wang W.N., Fan K.N., Xu W., Hua Z.Y. *Chem. Phys. Lett.* **2004**, *397*, 56.

[14] Pople J.A., Head-Gordon M., Raghavachari K. *J. Chem. Phys.* **1987**, *87*, 5968.

[15] Møller C., Plesset M.S. *Phys. Rev.* **1934**, *46*, 618.

[16] a) Mao L., Wang Y., Liu Y., Hu X. *J. Mol. Biol.* **2004**, *336*, 787. b) Jurečka P., Hobza P. *J. Am. Chem. Soc.* **2003**, *125*, 15608. c) Šponer J., Jurečka P., Marchan I., Luque F.J., Orozco M., Hobza P. *Chem. Eur. J.* **2006**, *12*, 2854.

[17] Frich M.J., Pople J.A., Binkley J.S. *J. Chem. Phys.* **1984**, *80*, 3265.

[18] Zhang G., Pei Y., Ma J., Yin K., Chen C.L. *J. Phys. Chem. B* **2004**, *108*, 6988.

[19] Curcó D., Alemán C. *J. Comput. Chem.* **2007**, *28*, 1743

[20] Frisch M.J. *et al*. *Gaussian 03,* Revision B.02, Gaussian, Inc.: Pittsburgh PA, **2003**.

[21] Boys S.F., Bernardi F. *Mol. Phys.* **1970**, *19*, 553.

[22] a) Piacenza M., Grimme S. *J. Am. Chem. Soc.* **2005**, *127*, 14841. b) Piacenza M., Grimme S. *Chem. Phys. Chem.* **2005**, *6*, 1554.

5.3. On the π-dimerization of oxidized thiophene oligomers: a theoretical investigation*

5.3.1. *Introduction*

Attractive nonconvalent interactions that take place between conjugated π-systems, *i.e.* π-stacking interactions, play a central role in the stabilization of many different assemblies that contain aromatic moieties, *e.g.* conjugated polymers based systems, consecutive base pairs in DNA/RNA, host-guest complexes, etc.[1,2] Within this context, the π-stacking interactions between thiophene rings belonging to different molecules, or equivalently, to the same polymer molecule but very far in the sequence are particularly important for the many of the relevant technological properties of oligothiophene, polythiophene and their derivatives.[2] For these reasons, a number of quantum mechanical approaches have been applied to model and understand the phenomena associated to the π-stacking of neutral aromatic molecules, including thiophene-derivatives.[3,4] Coupled cluster with single and double substitutions with noniterative triple excitations (CCSD(T)) has been reported as the best method to describe the intermolecular interactions in systems involving aromatic π-stacking. However, the applicability of this method is restricted to very small π-conjugated systems and to simple energy evaluations since it requires huge amount of computational resources. In practice the π-stacking of aromatic molecules is studied using second order Møller-Plesset perturbation (MP2) calculations, even although it is well known that this method tends to overestimate the strength of such interaction.

In an early study devoted to examine the origin of the π-stacking interaction of the neutral thiophene dimer, Tsuzuki *et al.*[3a] concluded that the interaction energy of the perpendicular dimer is favored with respect to that of the parallel dimer by 1.4 kcal/mol. More recently, π-stacked complexes formed by two, three and four neutral thiophene rings were investigated.[3b] Analysis of the optimized geometries obtained for the trimer and tetramer revealed that the orientation of the rings presents a preferred degree of periodicity. Furthermore, negligible or even slightly positive n-body effects were predicted for the π-stacked thiophene arrangements studied in such work. On the other hand, in the last decade experimental evidences have demonstrated that thiophene

* Submitted for publication

derivatives in the radical cation state form π-dimers.[2c,5] These π-stacked aggregates of radical cations, which are very stable charge carrying species, have important implications for the mechanism of charge transport in oxidized conducting polymers. Recently, the π-stacking of positively charged thiophene-containing molecules was modeled using MP2 and Density Functional Theory (DFT) quantum mechanical calculations.[6] However, in spite of the significant chemical and technological interest of the aggregates formed by π-stacked oxidized thiophene derivatives, full geometry optimizations were not considered in such studies, *i.e.* formation of aggregates was imposed by fixing the intermolecular distance.

In this chapter a complete and systematic theoretical study about the ability of different quantum mechanical methods to describe the existence of π-stacked aggregates involving thiophene rings in the radical cation state is presented. Specifically, the formation of complexes involving thiophene and 2,2'-bithiophene radical cations have been investigated using UHF, DFT and UMP2 calculations. The effects of complexation in the molecular geometry, atomic charge distribution, spin densities and the lowest $\pi-\pi^*$ transition energy of the charged species are discussed.

5.3.2. *Computational methods*

All calculations were performed using the Gaussian 03 computer program.[7] Geometry optimizations were performed using UHF, DFT and UMP2[8] methods combined with the 6-31G(d)[9] and 6-311+G(d,p)[10] basis sets. Specifically, DFT calculations were carried out using the following combinations: the Becke's three-parameter hybrid functional (B3)[11] with the Lee, Yang and Parr (LYP) expression for the nonlocal correlation[12] (UB3LYP); the modified Perdew-Wang exchange[13] (PW91) with the PW91 gradient-corrected correlation[14] (UMPW1PW91); the PW91 exchange functional[14] with the LYP correlation functional[12] (UPW91LYP); and the Becke's "half-and-half" functional (UBHandH), which is an *ad hoc* mixture of exact (UHF) and local density approximation exchange, coupled with the LYP expression for the correlation energy.[12] Complete geometry optimizations were performed in all cases, *i.e.* isolated molecules and both π-stacked dimers and trimers.

Atomic charges and spin densities were determined using the Mulliken population analysis. It should be taken into account that the values provided by this procedure

cannot be interpreted from a quantitative point of view, even although they are useful to compare the electronic structure of the investigated molecules when they are isolated and involved in π-stacked complexes. The π-π^* lowest transition energy (ε_g) was approximated as the difference between the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) energies, *i.e.* $\varepsilon_g = \varepsilon_{LUMO} - \varepsilon_{HOMO}$.

5.3.3. *Results and discussion*

The geometric parameters used to describe the different arrangements considered for calculations on complexes formed by two or more radical cations are schematized in Figure 1. The distance R_d indicates the separation between the two molecules, being defined with respect to the center of mass of each molecule. The angle θ_d indicates the relative orientation of the two molecules (the molecules are parallel and antiparallel when θ_d= 0º and 180º, respectively). This angle was defined considering the center of masses of each ring and the center of the inter-ring bonds for complexes formed by thiophene and 2,2'-bithiophene, respectively. Finally, φ_d refers to the angle formed by the planes defined by two molecules, which evidences the degree of tilting (the two rings are coplanar when φ_d= 0º, while φ_d=90º corresponds to a perpendicular T-shaped configuration).

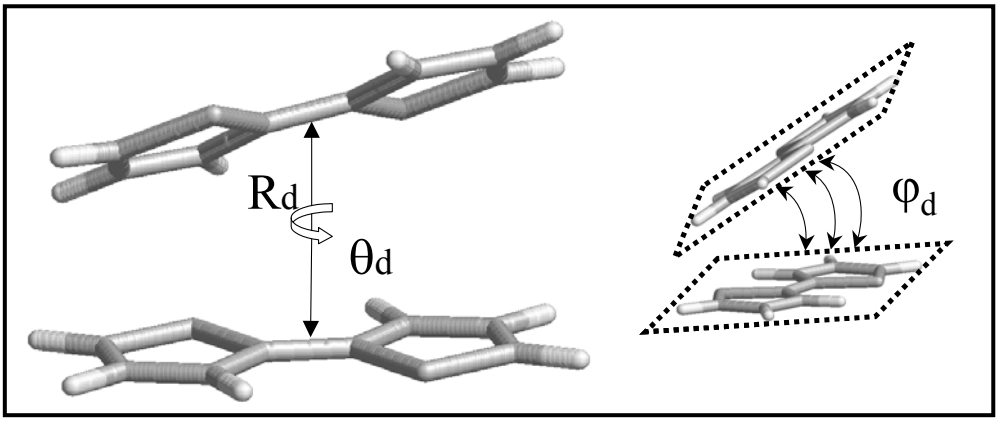


Figure 1. Definition of the intermolecular parameters used in this work (see text): Rd corresponds to the distance between the centers of masses of the two molecules, □d is the angle that defines the relative orientation of the two molecules and □d indicates the degree of tilting between two planar molecules.

Structures of a complex formed by two thiophene rings in the radical cation state (2-$T^{\bullet+}$) were constructed considering the two rings arranged in a π-stacked configuration (φ_d= 0º) with θ_d= 0º, 90º and 180º, the initial value for the inter-ring distance R_d being 3.7 Å. Such three structures were used as starting points for complete geometry optimizations using the UHF, UMP2, UB3LYP, UMPW1PW91, UPW91LYP and UBHandH methods combined with both the 6-31G(d) and 6-311+G(d,p) basis sets. The four DFT methods led to a separation of the fragments in all cases, even although some of them were successfully used in previous works to study π-stacking interactions. Specifically, the BHandH hybrid functional was found to describe correctly π-stacked aromatic complexes formed by benzenes, pyridines and DNA bases,[15] while π-stacking interactions between cytosine and single-walled carbon nanotubes were successfully evaluated in a recent work using the PW91LYP functional.[16] On the other hand, charge transfer complexes formed by thiophene and metallic atoms were reasonably described using the UMPW1PW91 method.[17] On the other hand, UHF and UMP2 optimizations led to a stable π-stacked dimer when the starting point was that with θ_d= 180º. The minimum nature of the π-stacked dimers provided by UHF (both basis sets) and UMP2 (6-31G(d) basis set only) calculations was confirmed by frequency calculations. The spin contamination was very low in all cases, that is, the highest overestimation was 2%.

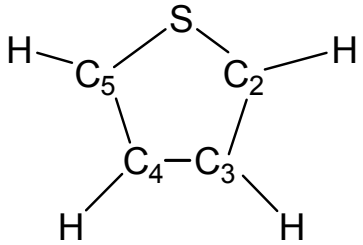
Table 1. Intermolecular geometric parameters and energy for the π-stacked dimer of 2-$T^{\bullet+}$ obtained from full geometry optimizations.

Level	R_d (Å)	θ_d (º)	φ_d (º)	E (a.u.)
UHF/6-31G(d)	3.042	150.6	6.7	-1101.8500459
UHF/6-311+G(d,p)	3.064	152.1	9.3	-1101.9631447
UMP2/6-31G(d)	2.921	180.0	0.0	-1103.1178556
UMP2/6-311+G(d,p)	2.909	180.0	0.0	-1103.3377307

The most relevant structural parameters of the π-stacked dimers converged at the UHF and UMP2 levels are listed in Table 1. As can be seen, the geometries produced by such two theoretical methods present some significant differences in the intermolecular parameters, which are independent of the size of the basis set. First, the inter-ring distance R_d predicted at the UHF level is larger (~ 0.15 Å) than that calculated using the

UMP2 method. Second, in the UHF minima the relative orientation of the two thiophene rings deviates by ~30º from the ideal antiparallel arrangement, the latter being identified in the minima obtained at the UMP2 level. Finally, the UHF π-stacked dimers show a tilt angle of ~7-9º, while the UMP2 method provides ideally coplanar geometries.

Table 2. Bond lengths and bond angles for the thiophene ring in the radical cation state calculated for 1-$T^{\bullet+}$ and 2-$T^{\bullet+}$ π-stacked complexes at different levels of theory.



	UHF/6-31G(d)[a]		UMP2/6-31G(d)		UHF/6-311+G(d,p)[a]		UMP2/6-311+G(d,p)	
#	1-$T^{\bullet+}$	2-$T^{\bullet+}$	1-$T^{\bullet+}$	2-$T^{\bullet+}$	1-$T^{\bullet+}$	2-$T^{\bullet+}$	1-$T^{\bullet+}$	2-$T^{\bullet+}$
S-C_2	1.714	1.720	1.709	1.718	1.713	1.724	1.705	1.712
C_2-C_3	1.357	1.374	1.429	1.402	1.355	1.373	1.432	1.404
C_3-C_4	1.497	1.440	1.374	1.425	1.496	1.439	1.375	1.425
C_2-H	1.071	1.073	1.086	1.088	1.071	1.074	1.086	1.088
C_3-H	1.072	1.074	1.085	1.089	1.073	1.075	1.084	1.088
$\angle SC_2C_3$	110.6	111.8	113.4	111.8	110.5	111.9	113.5	111.9
$\angle C_2C_3C_4$	112.1	111.9	111.5	112.0	112.1	112.1	111.3	111.9
$\angle HC_2S$	120.5	120.6	120.2	120.2	120.4	120.5	120.0	120.1
$\angle HC_3C_2$	124.4	123.7	122.5	123.2	124.4	123.6	122.6	123.1

[a] As in the UHF minima of 2-$T^{\bullet+}$ the relative orientation of the two thiophene rings deviates deviate 30º from the ideal antiparallel arrangement (see Table 1), the molecular geometry was not symmetric. Accordingly, listed bond lengths and bond angles have been averaged for 2-$T^{\bullet+}$.

Table 2 compares the geometric parameters, *i.e.* bond lengths and bond angles, calculated for an isolated thiophene ring in the radical cation state (1-$T^{\bullet+}$) with those obtained for the monomers in the 2-$T^{\bullet+}$ π-stacked complex. It should be noted that the geometric parameters predicted by the UHF method for the 2-$T^{\bullet+}$ complex were not symmetric due to distortion with respect to the ideal coplanar arrangement (Table 1), the

values listed in Table 2 being an average. Independently of the basis set, UHF calculations predict that the structure of 1-$T^{\bullet+}$ is similar to that typically found for neutral thiophene, *i.e.* a benzenoid-like structure with C_2-C_3 and C_4-C_5 bond lengths shorter than the C_3-C_4 one, while the opposite behavior is provided by UMP2 calculations, *i.e.* the largest bond length corresponds to C_3-C_4. The UMP2 geometries are in good agreement with previous calculations on charged thiophene-containing oligomers,[18] while the UHF method seems to overestimate the benzenoid-like structure of thiophene even in the radical cation state. Examination to the geometric parameters of 2-$T^{\bullet+}$ indicates that the electronic delocalization through the rings is enhanced by both UHF and UMP2 methods. Thus, the difference between the bond lengths of C_2-C_3 and C_3-C_4 is significantly larger for 1-$T^{\bullet+}$ (0.14 and 0.05 Å at the UHF and UMP2 level, respectively) than for 2-$T^{\bullet+}$ (0.06 and 0.02 Å at the UHF and UMP2 level, respectively).

Table 3. Mulliken charge[a] and spin density[a] obtained from UMP2/6-31G(d) and UMP2/6-311+G(d,p) calculations on 1-$T^{\bullet+}$ and the π-stacked complex of 2-$T^{\bullet+}$.

	UMP2/6-31G(d)		UMP2/6-311+G(d,p)	
#	1-$T^{\bullet+}$	2-$T^{\bullet+}$	1-$T^{\bullet+}$	2-$T^{\bullet+}$
	Charge			
S	0.546	0.630	0.188	0.532
C_2/C_5	0.044	-0.025	0.334	0.296
C_3/C_4	0.183	0.210	0.072	-0.062
	Spin Density			
S	-0.230	0.328	-0.286	0.414
C_2/C_5	0.662	0.151	0.691	0.162
C_3/C_4	-0.047	0.185	-0.048	0.131

[a] Atomic charges and spin densities with hydrogen atoms summed into heavy atoms.

Table 3 lists the distributions of Mulliken charges and spin densities predicted for both the 1-$T^{\bullet+}$ and the 2-$T^{\bullet+}$ π-stacked complex at the UMP2/6-31G(d) and UMP2/6-311+G(d,p) levels. As can be seen, calculations with the two basis set indicate that the spin density is moved from the C_2 and C_5 to the sulfur upon formation of the π-stacked complex. In contrast, the distribution of charges is significantly influenced by the basis

set. Thus, UMP2/6-31G(d) calculations on both 1-$T^{\bullet+}$ and 2-$T^{\bullet+}$ predict that the charge is mainly localized in the sulfur atom, *i.e.* the charge distributions of the complex and the isolated monomer are very similar. However, the charge of 1-$T^{\bullet+}$ is fundamentally concentrated in C_2 and C_5 at the UMP2/6-311+G(d,p), being moved to the sulfur atom in the π-stacked complex.

On the other hand, a number of structures of 2-$T^{\bullet+}$ with the two rings arranged in a T-shaped configuration (φ_d= 90º) were constructed considering not only different values of θ_d (0º, 90º, 180º and 270º) but also eight different orientations of the ring that is bisecting the π-cloud. Geometry optimizations were performed with the same computational levels used above for the π-stacked complexes. Independently of both the theoretical method and the basis set, a separation of the two thiophene rings was obtained for all the starting arrangements. Furthermore, π-stacked structures of a complex formed by three thiophene rings (3-$T^{\bullet+}$) were constructed using the geometry of the π-stacked dimer of 2-$T^{\bullet+}$ obtained at the UMP2/6-311+G(d,p) level and arranging the third ring in both parallel and antiparalel with respect to the second one. Geometry optimization of 3-$T^{\bullet+}$ structures at the UHF and UMP2 levels combined with the 6-31G(d) and 6-311+G(d,p) basis sets led to a separation of the three thiophene rings in all cases. Accordingly, π-stacked complexes involving thiophene rings in the radical cation were found to be stable at the dimer level only.

π-Stacked structures of a complex formed by two molecules of 2,2'-bithiophene in the radical cation state (2-2$T^{\bullet+}$) were constructed considering the parallel and antiparallalel relative arrangements, the initial inter-ring distance calculated with respect to the center of masses of each molecule being 3.0 Å. Geometry optimizations of these structures using the UHF, UB3LYP, UMPW1PW91, PW91LYP and BHandH methods combined with the 6-31G(d) and 6-311+G(d,p) basis sets led to the separation of the two molecules in all cases. In opposition, full geometry optimizations at the UMP2/6-31G(d) and UMP2/6-311+G(d,p) levels converged into two stable complexes: the parallel and the antiparallel π-stacked dimers. It is worth noting that this in an important difference with respect to 2-$T^{\bullet+}$, for which the antiparallel π-stacked dimer was the only minimum characterized. On the other hand, spin contamination for 2-2$T^{\bullet+}$ complexes was very low, *i.e.* lower than 2%.

Table 4 lists the energies and inter-molecular parameters of the minima characterized for 2-2$T^{\bullet+}$, while Figure 2 displays the two complexes optimized at the UMP2/6-

311+G(d,p) level. As can be seen, the intermolecular parameters of the parallel π-stacked dimer (Figure 2a) remains relatively close to the ideal values, especially when the largest basis set is used, *i.e.* the most significant distortion corresponds to θ_d that deviates ~16º. In contrast, for the antiparallel π-stacked dimer (Figure 2b) the distortion of θ_d increases to ~38º. On the other hand, inspection to the R_d values reveals two significant features: (i) the intermolecular distance is noticeably larger for 2-2T$^{\bullet+}$ than for 2-T$^{\bullet+}$; and (ii) the intermolecular distance is shorter for the antiparallel dimer than for the parallel one. These results indicate that the distance between the two molecules increases with the delocalisation of the positive charge within each molecule. Furthermore, the stability of the π-stacked dimer increases when R_d decreases. Thus, the antiparallel arrangement is more stable than the parallel one by 1.8 kcal/mol. Finally, it should be noted that the displacement of the center of mass of the second molecule with respect to that of the first one is negligible for both the parallel and antiparallel π-stacked dimers.

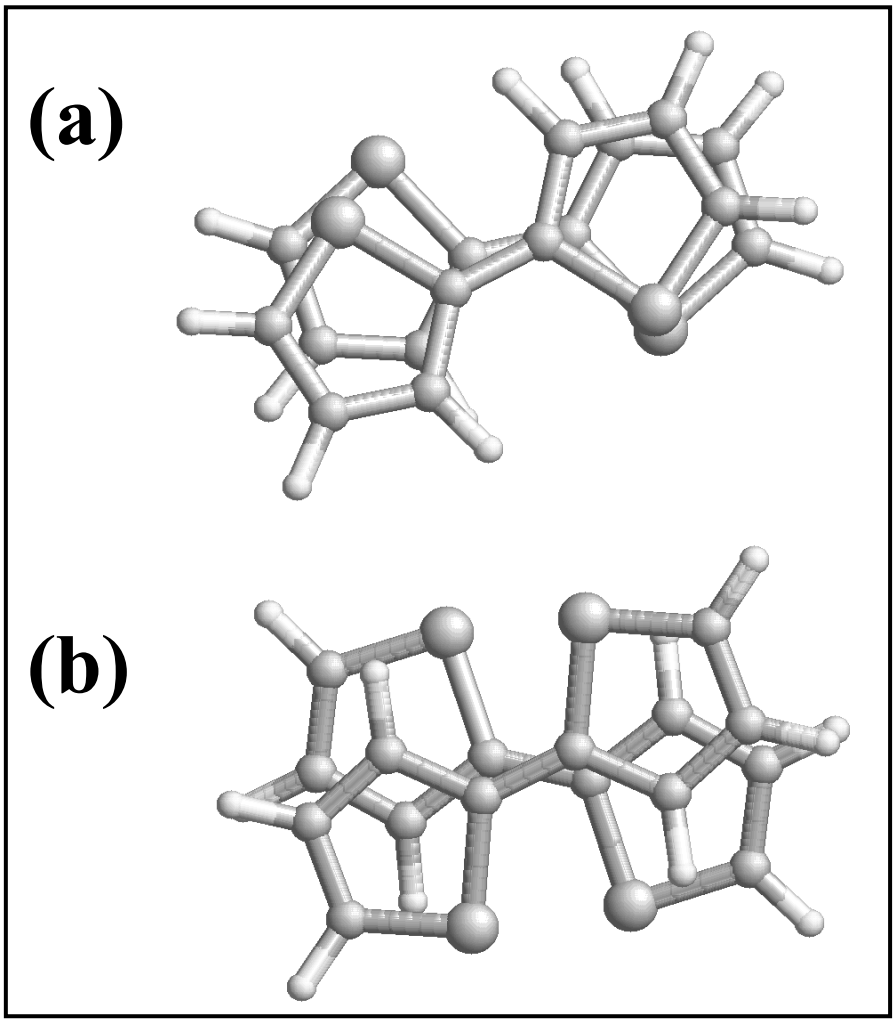


Figure 2. (a) Parallel and (b) antiparallel π-stacked dimers obtained for 2-2T$^{\bullet+}$.

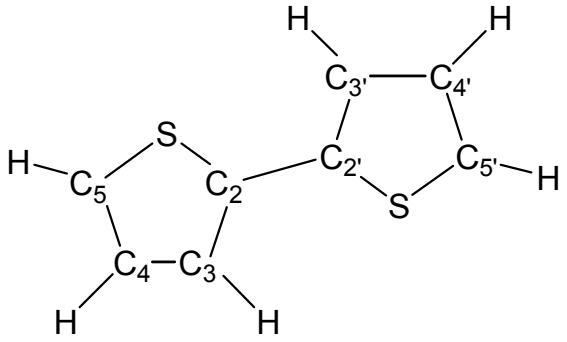
Table 4. Intermolecular geometric parameters and relative energy (ΔE) for the π-stacked dimers of 2-2T$^{\bullet+}$ obtained from full geometry optimizations.

Level	Starting arrangement	R_d (Å)	θ_d (º)	φ_d (º)	ΔE (kcal/mol)
UMP2/6-31G(d)	Parallel	3.708	35.6	2.5	0.5
	Antiparallel	3.642	140.6	1.7	0.0[a]
UMP2/6-311+G(d,p)	Parallel	4.098	15.7	3.6	1.8
	Antiparallel	3.416	141.8	2.2	0.0[b]

[a] E= -2204.7927028 a.u. [b] E= -2205.217389 a.u.

In Table 5 the bond lengths and bond angles obtained for the π-stacked dimers of 2-2T$^{\bullet+}$ are compared with those calculated for an isolated molecule of 2,2'-bithiophene in the radical cation state (1-2T$^{\bullet+}$). Interestingly, in this case the geometric parameters predicted for 1-2T$^{\bullet+}$ and 2-2T$^{\bullet+}$ are very similar independently of both the parallel and antiparallel arrangement of the π-stacked complexes and the basis set used for the calculations. Moreover, comparison of the inter-ring bond length, which is the most relevant geometric parameter, shows that the largest difference is 0.007 and 0.005 Å at the UMP2/6-31G(d) and UMP2/6-311+G(d,p) levels, respectively. The Mulliken spin densities and atomic charges calculated for 1-2T$^{\bullet+}$ and 2-2T$^{\bullet+}$ are compared in Table 6. As can be seen, the spin densities of both the parallel and antiparallel π-stacked complexes are very similar to that obtained for the isolated 1-2T$^{\bullet+}$, which is fully consistent with the small changes detected in the geometric parameters upon complexation. Inspection of the Mulliken atomic charges computed at the UMP2/6-31G(d) shows the same feature, even although the charges obtained with the 6-311+G(d,p) does not maintain this consistency. Thus, although UMP2/6-311+G(d,p) charges were similar for the parallel and antiparallel π-stacked complexes, they differ from those obtained for the isolated 1-2T$^{\bullet+}$. In the latter molecule, the charge is mainly concentrated on the C_3 and $C_{3'}$ atoms, which suggest that the differences between 1-2T$^{\bullet+}$ and 2-2T$^{\bullet+}$ are mainly due to the limitations of the Mulliken analysis.

Table 5. Bond lengths and bond angles for the thiophene ring in the radical cation state calculated for 1-2T$^{\bullet+}$ and 2-2T$^{\bullet+}$ π-stacked complexes at different levels of theory.[a]



	UMP2/6-31G(d)			UMP2/6-311+G(d,p)		
#	1-2T$^{\bullet+}$	2-2T$^{\bullet+}$ (parallel)	2-2T$^{\bullet+}$ (antiparallel)	1-2T$^{\bullet+}$	2-2T$^{\bullet+}$ (parallel)	2-2T$^{\bullet+}$ (antiparallel)
S-C_2	1.760	1.758	1.761	1.757	1.760	1.760
C_2-C_3	1.406	1.409	1.399	1.408	1.410	1.408
C_3-C_4	1.375	1.379	1.376	1.378	1.378	1.377
C_4-C_5	1.390	1.396	1.392	1.395	1.396	1.397
C_5-S	1.673	1.669	1.673	1.666	1.667	1.668
C_3-H	1.087	1.087	1.087	1.086	1.086	1.087
C_4-H	1.084	1.084	1.085	1.083	1.084	1.084
C_5-H	1.086	1.086	1.087	1.085	1.086	1.087
C_2-$C_{2'}$	1.405	1.412	1.406	1.410	1.415	1.410
$\angle SC_2C_3$	109.7	109.7	109.7	109.8	109.6	109.8
$\angle C_2C_3C_4$	113.4	113.3	113.4	113.3	113.4	113.2
$\angle C_3C_4C_5$	111.7	111.5	111.7	111.5	111.4	111.6
$\angle C_4C_5S$	114.2	114.3	114.1	111.4	114.4	114.2
$\angle SC_2C_{2'}$	121.4	121.1	121.2	121.1	120.5	120.9
$\angle HC_3C_2$	122.9	123.0	122.9	122.9	122.8	125.6
$\angle HC_4C_3$	125.3	125.5	125.4	125.5	125.4	126.0
$\angle HC_5C_4$	126.0	125.9	126.1	125.9	125.8	120.9

[a] Bond lengths and bond angles have been averaged for 2-2T$^{\cdot+}$.

Table 6. Mulliken charge and spin density obtained from UMP2/6-31G(d) and UMP2/6-311+G(d,p) calculations on 1-2T$^{\bullet+}$ and the π-stacked complex of 2-2T$^{\bullet+}$.[a]

	UMP2/6-31G(d)			UMP2/6-311+G(d,p)		
	1-2T$^{\bullet+}$	2-2T$^{\bullet+}$ (parallel)	2-2T$^{\bullet+}$ (antiparallel)[b]	1-2T$^{\bullet+}$	2-2T$^{\bullet+}$ (parallel)	2-2T$^{\bullet+}$ (antiparallel)[b]
	Charge					
S/S'	0.462	0.451/0.460	0.458	-0.337	0.181/0.396	0.320
$C_2/C_{2'}$	-0.197	-0.210/-0.195	-0.204	-0.171	0.847/0.968	0.744
$C_3/C_{3'}$	0.166	0.170/0.159	0.161	1.245	-0.235/-0.362	-0.137
$C_4/C_{4'}$	0.112	0.111/0.127	0.116	-0.594	-0.455/-0.303	-0.263
$C_5/C_{5'}$	-0.043	-0.048/-0.025	-0.031	0.357	0.038/-0.075	-0.164
	Spin Density					
S/S'	-0.116	-0.113/-0.123	-0.116	-0.161	-0.203/-0.212	-0.156
$C_2/C_{2'}$	-0.015	-0.047/0.038	-0.009	0.015	0.031/0.048	0.058
$C_3/C_{3'}$	0.411	0.416/0.376	0.398	0.392	0.419/0.396	0.328
$C_4/C_{4'}$	-0.392	-0.399/-0.368	-0.380	-0.357	-0.385/-0.381	-0.324
$C_5/C_{5'}$	0.612	0.612/0.608	0.607	0.611	0.637/0.650	0.594

[a] Atomic charges and spin densities with hydrogen atoms summed into heavy atoms. The distribution of atomic charges and spin density was symmetric for the antiparallel π-stacked complex.

The formation of π-stacked dimers involving radical cations is expected to facilitate the intermolecular electron transfer.[2,5] Thus, the presence of intermolecular π-stacked dimers containing charged conjugated fragments of polymers is expected to affect the electronic parameters related with the high conductivities of polythiophene derivatives in the oxidized state. Table 7 compares the lowest $\pi-\pi^*$ electron transition energy (ε_g), of all the calculated dimers with that of the corresponding monomers. As can be seen, ε_g is smaller for the π-stacked dimer than for the monomer in all cases. However, such reduction is much more important when the positive charge of each monomer is located in a single thiophene ring than when it is distributed between two thiophenes, as in 2-2T$^{\bullet+}$. These results suggest that in thiophene-containing charged systems, the

intermolecular delocalization of π-electrons is significantly smaller than the corresponding intramolecular phenomenon.

Table 7. Lowest transition energy (ε_g; in eV) for the π-stacked dimers of 2-$T^{\bullet+}$ and 2-2$T^{\bullet+}$ as well as their corresponding monomers.

System	Description / Geometry	6-31G(d)	6-311+G(d,p)
1-$T^{\bullet+}$	Isolated thiophene in the radical cation state	9.53	9.40
2-$T^{\bullet+}$	π-stacked dimer formed by two charged thiophene rings	6.37	6.38
1-2$T^{\bullet+}$	Isolated 2,2'-bithiophene in the radical cation state	8.14	8.01
2-2$T^{\bullet+}$/p	π-stacked dimer formed by two charged 2,2'-bithiophene molecules arranged in parallel	7.98	7.64
2-2$T^{\bullet+}$/a	π-stacked dimer formed by two charged 2,2'-bithiophene molecules arranged in antiparallel	7.91	7.72

5.3.4. *Conclusions*

In summary, this chapter reports a systematic study based on full geometry optimizations about the formation of π-stacked complexes by thiophene-containing molecules in the radical cation state. Results allow obtain the following conclusions: (i) UHF and DFT calculations are not suitable for this kind of systems; (ii) π-stacked complexes are formed at the dimer level, aggregation of three molecules being not found for 3-$T^{\bullet+}$; (iii) dimers arranged in T-shaped are not stable; (iv) the variation of the intermolecular distances indicates that the stability of the dimers increases with the localization of the positive charge within each molecule; and (v) within the context of charge transport in oxidized conducting polymers, the intermolecular phenomena are small with respect to the intramolecular ones.

5.3.5. *References*

[1] a) Müller-Dethlefs K., Hobza P. *Chem. Rev.* **2000**, *100*, 143.

[2] a) Handbook of Conducting Polymers, 2nd ed, Skotheim T.A., Reynolds J.R., Elsenbaumer R.L. Eds., Marcel Dekker: New York, **1997**. b) Handbook of Organic Conductive Molecules and Polymers, Nalwa H.S. Ed., Wiley: New York, **1997**, Vols. 1-4. c) Miller L., Mann K.R. *Acc. Chem. Res.* **1996**, *29*, 417.

[3] a) Tsuzuki S., Kazumasa H., Azumi R. *J. Am. Chem. Soc.* **2002**, *124*, 12200. b) Rodríguez-Ropero F., Casanovas J., Alemán C. *J. Comput. Chem.* **2008**, *29*, 69.

[4] a) Willnauer C., Birkenheuer U. *J. Chem. Phys.* **2004**, *120*, 11910. b) Tsuzuki S., Honda K.R., Uchimaru T., Mikami M., Tanabe K. *J. Am. Chem. Soc.* **2002**, *124*, 104. c) Hobza P., Šponer J. *J. Am. Chem. Soc.* **2002**, *124*, 11802. d) Tauer T.P., Sherrill C.D. *J. Phys. Chem. A* **2005**, *109*, 10475. e) Hobza P., Selzle H.L., Schlag E.W. *J. Phys. Chem.* **1996**, *100*, 18790. f) Tsuzuki S., Uchimaru T., Matsumura K., Mikami M., Tanabe K. *Chem. Phys. Lett.* **2000**, *319*, 547. g) Jurecka P., Hobza P. *J. Am. Chem. Soc.* **2003**, *125*, 15608. i) Sponer J., Hobza P. *Chem. Phys. Lett.* **1997**, *267*, 263. j) Tsuzuki S., Honda K., Uchimaru T., Mikami M. *J. Chem. Phys.* **2005**, *122*, 144323.

[5] a) Hill M.G., Mann K.R., Miller L.L., Penneau J.F., Zinger B. *Chem. Mater.* **1992**, *4*, 1106. b) Zinger B., Mann K.R., Hill M.G., Miller L.L. *Chem. Mater.* **1992**, *4*, 1113. c) Yu Y., Gunic E., Zinger B., Miller L.L. *J. Am. Chem. Soc.* **1996**, *118*, 1013. d) Graf D.D., Campbell J.P., Miller L.L., Mann K.R. *J. Am .Chem. Soc.* **1996**, *118*, 5480. e) Yamazaki D., Nishinaga T., Tanino N., Komatsu K. *J. Am. Chem. Soc.* **2006**, *128*, 14470.

[6] a) Scherlis D., Fattebert J.-L., Marzari N. *J. Chem. Phys.* **2006**, *124*, 194902. b) Pickholz M., Dos Santos M.C. *J. Mol. Struct. (Theochem)* **2005**, *717*, 99. c) Scherlis D., Marzari N. *J. Phys. Chem. B* **2004**, *108*, 17791.

[7] Frisch M.J., *et al.* Gaussian 03, Revision B.02, Gaussian, Inc.: Pittsburgh PA, **2003**.

[8] Møller C., Plesset M.S. *Phys Rev* **1934**, *46*, 618.

[9] Hariharan P.C., Pople J.A. *Chem. Phys. Lett.* **1972**, *16*, 217.

[10] Frich M.J., Pople J.A., Binkley J.S. *J. Chem. Phys.* **1984**, *80*, 3265.

[11] Becke A.B. *J. Chem. Phys.* **1993**, *98*, 1372.

[12] Lee C., Yang W., Parr R.G. *Phys. Rev. B* **1988**, *37*, 785.

[13] Adamo C., Barone V. *J. Chem. Phys.* **1998**, *108*, 664.

[14] Perdew J.P., Wang Y. *Phys. Rev.* **1992**, *45*, 13244.

[15] Waller M.P., Robertazzi A., Platts J.A., Hibbs D.E., Williams P.A. *J. Comput. Chem.* **2006**, *27*, 491.

[16] Wang Y., Bu Y. *J. Phys. Chem. B* **2007**, *111*, 6520.

[17] Alemán C., Curcó D., Casanovas J. *Phys Rev E* **2005**, *72*, 026704.

[18] a) Alemán C., Julià L. *J. Phys. Chem.* **1996**, *100*, 14661. b) Casanovas J., Alemán C. *J Phys Chem C* **2007**, *111*, 4823.

5.4. Modeling an electronic conductor based on natural peptide sequences*

5.4.1. *Introduction*

One of the most important achievements of nanotechnology is the design of new materials and devices at the molecular level, which have helped enormously the development of areas such as biomedicine or electronics. Within this context, nanotubes have been settled as an important family of nanoconstructs because of their applicability in different fields ranging from medicine, *e.g.* drug and gene delivery systems,[1] to electronics, *e.g.* nanowires,[2]. Among them, peptide-based nanotubes are gaining broad attention within the scientific community because of the potentiality of peptide segments to self-assemble, the possibility of adding functionalities in specific positions of the nanotube, *i.e.* by tuning certain amino acids, and their high biocompatibility.[3]

The building of nanotubes based on natural left-handed β-helical proteins motifs has been surveyed.[4] The tubular nature of left-handed β-helical proteins makes them excellent candidates to be used as building blocks for the construction of nanotubes by self-assembling directly these repeats without the need to introduce major structural changes. Haspel and *et al.*[4] selected up to 17 basic two-turn repetitive motifs as starting points to construct 17 nanotubes by stacking of replicas of these motifs, the stability of which was studied *in silico*. Results led to conclude that nanotubes constructed from residues 131 to 165 from the *Escherichia coli galactoside acetyltransferase* (PDB code 1krr, chain A) display remarkable stability under different simulation conditions, including temperature stress and various ionic strengths.[4]

Several studies centering their attention in 1krr based nanotubes have been performed, mainly focused on the enhancement of their structural stability. Thus, loop regions present a great mobility that can disrupt the tubular structure. This inconvenience was overcome replacing targeted residues located in the loop regions by synthetic amino acids with low conformational mobility such as 1-aminocyclopropanecarboxylic acid[5], 1-aminocyclopentane-1-caboxylic acid[6], 1-aminocyclohexanecarboxylic[7] acid and other non-proteinogenic amino acids.[8] On the other hand, 1krr nanotubes are characterized by an inner hydrophobic core containing valine and isoleucine, making them inappropriate to be used as nanowires or molecule carriers. Seeing that the narrowness of their hollow inner space makes them unsuitable to transfer large molecules, their use as nanowires

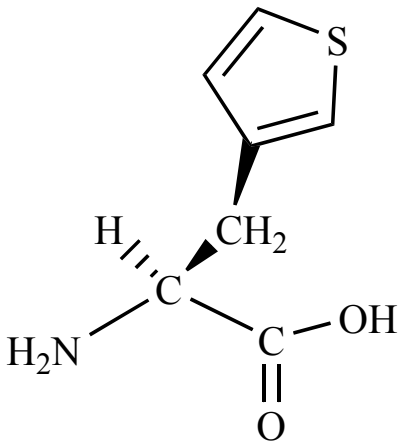
* Results presented in this chapter are published in *Biomacromolecules*. **2009***, 10*, 2338

seems to be their natural application. Thus the only limitation that should be overcome to make 1krr based nanotubes useful for the transfer of charge is the modification of their internal core enabling proper conditions for charge transfer. In a previous work it was shown by means of molecular dynamics (MD) simulations the feasibility of modifying the residue composition and the hydrophobicity of the interior core of a 1krr based nanotube by replacing specific residues in the inner core by histidine (His) residues. This creates a ladder of π-stacking residues while maintaining the original tubular structure throughout the simulation time.[9]

In this chapter MD and quantum mechanics/molecular mechanics (QM/MM) simulations are used to study the suitability of a nanotube built from modified 1krr motifs covalently linked to work as a charge transfer nanodevice. For this purpose, a ladder of π-electron-rich functional groups has been created through targeted substitutions of natural amino acids by synthetic ones. In order to fulfill this goal two necessary conditions are required:

(i) The mutated nanotubes must retain their tubular structure under all experimental conditions.

(ii) The synthetic amino acids as well as the positions they occupy in the nanotube must create a ladder of π-stacking.



Scheme 1

To satisfy such requirements, β-3-thienylalanine, hereafter abbreviated Tie (see Scheme 1), was selected as non-proteinogenic amino acid to construct the mutated nanotubes. This choice is based on the following considerations:

(i) Both the chemistry of Tie, as well as the accessibility of its synthesis, were reported a long time ago.[10]

(ii) The side chains of Tie and His show similar size suggesting that no drastic steric hindrance will occur if the former residue is introduced in the positions used by Haspel *et al.*[4] for the latter one.

(iii) Thienyl groups form a ladder of π-stacking when placed close to one another.[11]

5.4.2. *Methods*

Quantum Mechanics Calculations of N-Acetyl-N'-methylamide Derivative of β-3-Thienylalanine.

The *N*-acetyl-*N'*-methylamide derivative of β-3-thienylalanine, hereafter denoted as Ac-Tie-NHMe (Figure 1), was used to study the conformational flexibility of the selected synthetic amino acid. All quantum mechanics calculations, which were carried out using the Gaussian 03 program,[12] were performed using the B3LYP[13,14] functional combined with the 6-31+G(d,p) basis set,[15] *i.e.* B3LYP/6-31+G(d,p). As the main purpose of these calculations was to ascertain the intrinsic conformational properties of Tie, geometry optimizations were carried out in the gas-phase. On the other hand, the molecular electrostatic potential (MEP) was computed at the HF/6-31G(d)[16] level for selected minimum energy conformations of Ac-Tie-NHMe. These calculations, which were used to develop classical electrostatic parameters for the Tie residue (see below), were performed on a large set of points located outside the nuclear region.

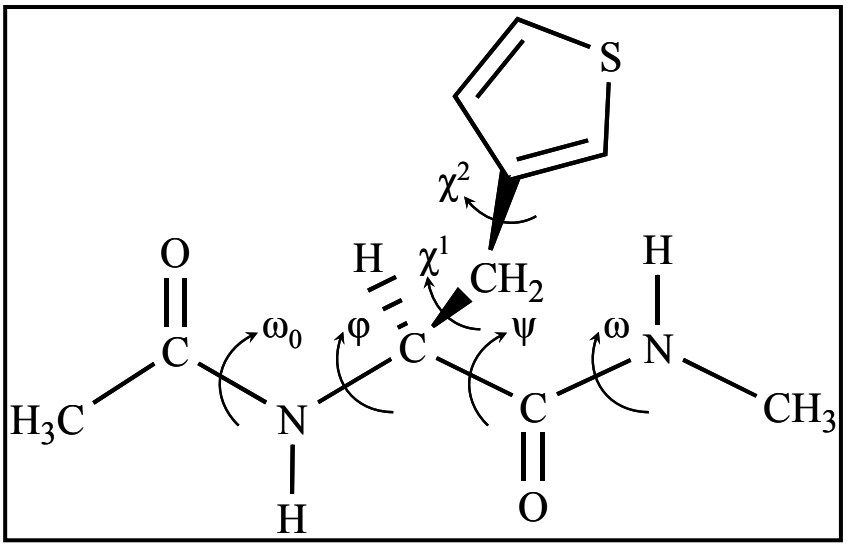


Figure 1. Structure of Ac-Tie-NHMe indicating the backbone and side chain dihedral angles.

Molecular Dynamics Simulations of the Nanotubes

Molecular dynamics (MD) simulations were performed using the NAMD computer package.[17] The energy was calculated with the AMBER force field[18,19] and all parameters required for the mutated 1krr based nanotubes were taken from AMBER libraries with the exception of the electrostatic ones for the Tie residue. Atomic centered charges were calculated by fitting the quantum mechanical and the Coulombic MEPs using a previously developed procedure.[20]

Simulations were performed using the NPT ensemble in an orthorhombic periodic box, periodic boundary conditions being applied through the nearest image convention. All the atoms were considered explicitly in each simulation. Water molecules were represented using the TIP3 model.[21] Chloride and sodium ions were added (28 ions) to maintain a physiological ionic strength of approximately 0.8% w/w and to ensure the neutrality of the whole system. In all cases the initial dimensions of the simulation box was (55.0 × 55.0 × 75.0) $Å^3$, the number of explicit particles ranging from 17221 to 17233. Atom-pair distance cutoffs were applied at 14.0 Å to compute the van der Waals interactions. The electrostatic interactions were computed by using the non-truncated electrostatic potential with Ewald summations.[22] Bond lengths involving hydrogen atoms were constrained by using the SHAKE algorithm[23] with a numerical integration step of 2 fs.

Before the production series, the thermodynamic variables of each system were equilibrated. The energy was initially minimized to relax conformational and structural tensions using the conjugate gradient method for $5 \cdot 10^3$ steps. Next, different consecutive rounds of short MD runs were performed in order to equilibrate the density, temperature, and pressure. First, solvent and ions were thermally relaxed, while the nanoconstructs were kept frozen: 0.5 ns of NVT-MD at 500 K were used to homogeneously distribute the solvent and ions in the box. Second, 0.5 ns of isothermal and 0.5 ns isobaric relaxation were run. Finally, all the atoms of the system were submitted to 0.15 ns of steady heating until the target temperature was reached (298 K), 0.25 ns of NVT-MD at 298 K (thermal equilibration) followed by 0.5 ns of density relaxation (NPT-MD). Both temperature and pressure were controlled by the weak coupling method, the Berendsen thermobarostat,[24] using a time constant for heat bath coupling and a pressure relaxation time of 1 ps. The end of the density relaxation simulation was the starting point of the molecular simulations presented in this work.

All the simulations were performed at 298 K and constant pressure of 1 atm. The coordinates of all the production runs, which were 10 ns long, were saved every 500 steps (1 ps intervals) for subsequent analysis.

Quantum Mechanics / Molecular Mechanics Calculations

QM/MM calculations were performed in the neutral and oxidized state in selected snapshots of the MD trajectory. While the side chain of Tie was described at the quantum mechanical level, the rest of the nanowire as well as the first solvation core were represented as point charges. QM region was connected to the MM part by the so-called "link atom" scheme.[25] The B3LYP[13,14] functional combined with the 6-311+G(d) basis set was used in the QM part. All QM/MM calculations were performed using a locally modified version of Gaussian 03 computer program.[12]

5.4.3. *Results and discussion*

Conformational properties and force-field parameterization of β-3-Thienylalanine

The similarity of β-3-Thienylalanine with phenylalanine (Phe) suggests a similar conformational behavior. Thus, the initial Ac-Tie-NHMe structures were built using the five minimum energy conformations reported by Chas *et al.*[26] for Ac-Phe-NH_2. Full geometry optimizations at the B3LYP/6-31+G(d,p) level led to the results displayed in Table 1 and Figure 2.

Table 1. Backbone and Side Chain Torsion Angles (see Figure 1),[a] Backbone Conformation[b] and Relative Energy (ΔE)[c] for the Minimum Energy Conformations of Ac-Tie-NHMe at the B3LYP/6-31+G(d,p) Level.

#	ω_0	φ	ψ	ω	χ^1	χ^2	Backbone	ΔE
I	-174.7	-82.9	59.8	179.0	44.8	73.8	$C_{7,eq}$	0.0[d]
II	-172.8	-84.4	74.4	-175.9	-55.2	118.0	$C_{7,eq}$	0.2
III	177.81	-156.6	157.2	174.4	-170.2	70.2	C_5	0.4
IV	-179.2	-82.9	80.9	-174.7	-165.4	100.0	$C_{7,eq}$	0.4
V	172.9	74.3	-53.5	-178.4	-59.1	109.2	$C_{7,ax}$	1.7

[a] In degrees. [b] The backbone conformation is defined by the dihedral angles φ and ψ. [c] In kcal/mol. [d] E= -1047.713098 a.u.

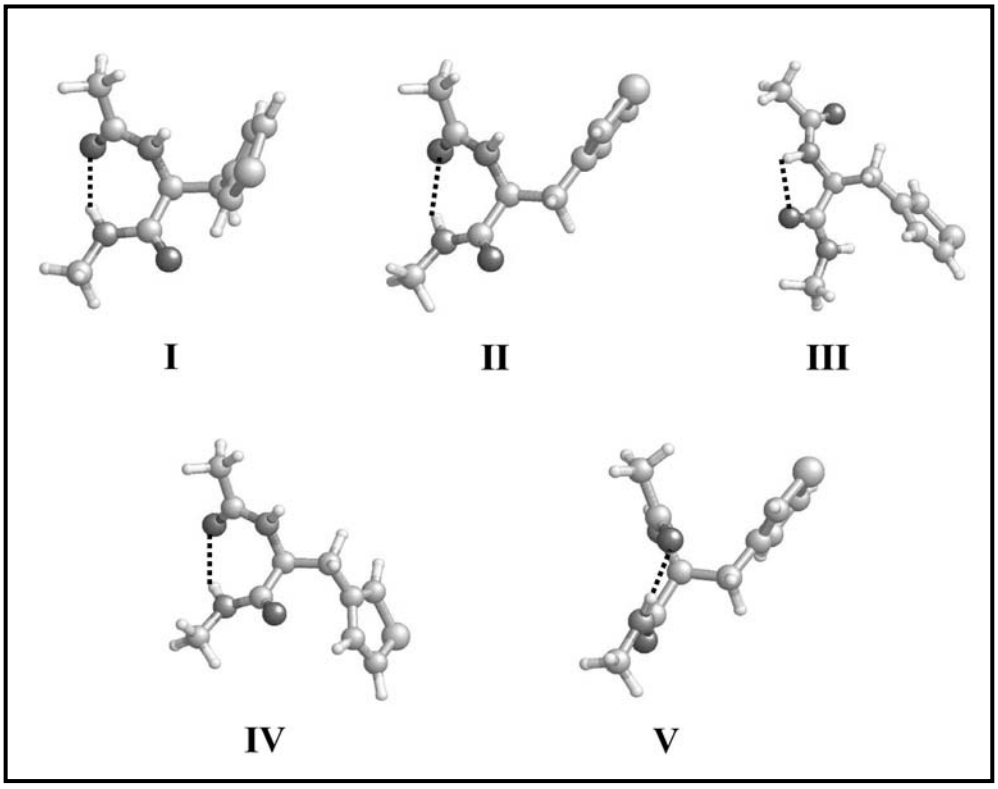


Figure 2. Minimum energy conformations calculated for Ac-Tie-NHMe at the B3LYP/6-31+G(d,p) level. Intramolecular hydrogen bonds N-H•••O=C (dashed lines) are indicated. The geometric parameters associated with the intramolecular interactions are given in the text. The labels used to describe the different minima correspond to those indicated in Table 1.

The most stable minimum in the gas phase (**I**) shows a $C_{7,eq}$ backbone conformation (Figure 2a), *i.e.* a γ-turn in which the backbone forms a seven-membered intramolecular hydrogen bonded ring. It is worth noting that the four minima of lower energy are separated by less than 0.5 kcal/mol. Among the three minima with a $C_{7,eq}$ conformation, **I** exhibits the hydrogen bond with the best geometric parameters: d(H···O) = 2.003 Å and ∠N–H···O = 148.2º. The hydrogen bonding parameters of the other two $C_{7,eq}$ backbone arrangements are: d(H···O) = 2.046 Å and ∠N–H···O = 144.4º for **II** (Figure 2b)**,** and d(H···O) = 2.152 Å and ∠N–H···O = 138.7º for **IV** (Figure 2d). Minimum **III** (Figure 2c), which is stabilized by a five-membered hydrogen bonded ring (C_5), shows the weakest intramolecular interaction as reveals the poor geometric parameters associated to the intramolecular hydrogen bond: d(H···O) = 2.186 Å and ∠N–H···O = 106.9º. Finally, minimum **V** (Figure 2e) adopts a $C_{7,ax}$ backbone conformation with hydrogen bonding parameters d(H···O) = 1.931 Å and ∠N–H···O = 150.2º. This structure is destabilized because of the steric strain associated with axial position of the side group in the seven membered hydrogen bonded ring characteristic of this

conformation. The N-H group of the Tie residue interacts with the π-electron cloud of the thienyl side chain in both **I** and **II**, the distance between the hydrogen atom of the NH moiety and the center of masses of the aromatic ring (*D*) being 3.150 and 3.648 Å, respectively. On the other hand, minimum **III** also shows N-H···π interaction with *D*= 3.097 Å, even although in this case it involves the N-methylamine blocking group rather than the NH of the Tie residue, *i.e.* the latter participates in the formation of the five-membered hydrogen bonded ring. All these results are fully consistent with those obtained for Ac-Phe-NH_2 by Chass *et al.*[26]

Electrostatic charges have been the only force-field parameters specifically developed for Tie residue, the stretching, bending, torsional, and van der Waals parameters being directly transferred from the Amber force-field.[18,19] Atomic centered charges of the five minima were derived from the quantum mechanical and the Coulombic MEPs.[20] Electrostatic charges, which are displayed in Figure 3, were obtained by weighting these charges according to the Boltzmann population of each minima.[27,28] It should be noted that electrostatic force-field parametrization using a strategy based on weighted multiple conformations through a Boltzmann distribution in the gas-phase has been demonstrated to be specially successful for non-proteogenic residues.[28-30]

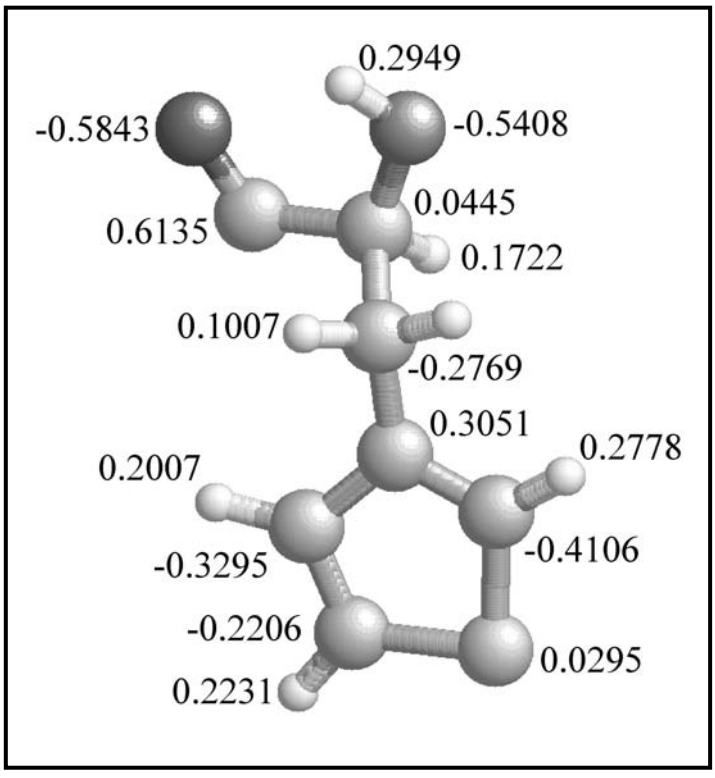


Figure 3. Atom centered electrostatic charges computed for the β-3-thienylalanine residue.

Stability of the mutated 1krr nanotubes

As previously stated, the non-proteinogenic amino acid Tie was introduced in the positions that ensure the formation of a ladder of π-stacking preserving the tubular structure of the system. Due to the similarity of Tie and His side chain, the three more stable His-mutated nanotubes reported by Haspel et al.[9] were considered to build the 1krr based nanotubes of the present study. These mutants, which are briefly described in Table 2, are the following: (a) I134TIE-I152TIE, in which the Ile134 and Ile152 residues of each 1krr building block are substituted by Tie; (b) I140TIE-I158TIE, in which the Ile140 and Ile158 residues are substituted by Tie; and (c) I146TIE-V164TIE, in which the Ile146 and Val164 residues are substituted by Tie.

Table 2. Name and sequence of all the studied systems.

Name	Sequence	Replaced Residues
1krr wild type	Ac-(PITIGNNVWIGSHVVINP GVTIGDNSVIGAGSIVTG PITIGNNVWIGSHVVINP GVTIGDNSVIGAGSIVTG PITIGNNVWIGSHVVINP GVTIGDNSVIGAGSIVTG PITIGNNVWIGSHVVINP GVTIGDNSVIGAGSIVT)-NHMe	-
I134TIE I152TIE	Ac-(PIT***Tie***GNNVWIGSHVVINP GVT***Tie***GDNSVIGAGSIVTG PIT***Tie***GNNVWIGSHVVINP GVT***Tie***GDNSVIGAGSIVTG PIT***Tie***GNNVWIGSHVVINP GVT***Tie***GDNSVIGAGSIVTG PIT***Tie***GNNVWIGSHVVINP GVT***Tie***GDNSVIGAGSIVT)-NHMe	Ile134 → Tie Ile152 → Tie
I140TIE I158TIE	Ac-(PITIGNNVW***Tie***GSHVVINP GVTIGDNSV***Tie***GAGSIVTG PITIGNNVW***Tie***GSHVVINP GVTIGDNSV***Tie***GAGSIVTG PITIGNNVW***Tie***GSHVVINP GVTIGDNSV***Tie***GAGSIVTG PITIGNNVW***Tie***GSHVVINP GVTIGDNSV***Tie***GAGSIVT)-NHMe	Ile140 → Tie Ile158 → Tie
I146TIE V164TIE	Ac-(PITIGNNVWIGSHVV***Tie***NP GVTIGDNSVIGAGSI***Tie***TG PITIGNNVWIGSHVV***Tie***NP GVTIGDNSVIGAGSI***Tie***TG PITIGNNVWIGSHVV***Tie***NP GVTIGDNSVIGAGSI***Tie***TG PITIGNNVWIGSHVV***Tie***NP GVTIGDNSVIGAGSI***Tie***T)-NHMe	Ile146 → Tie Val164 → Tie

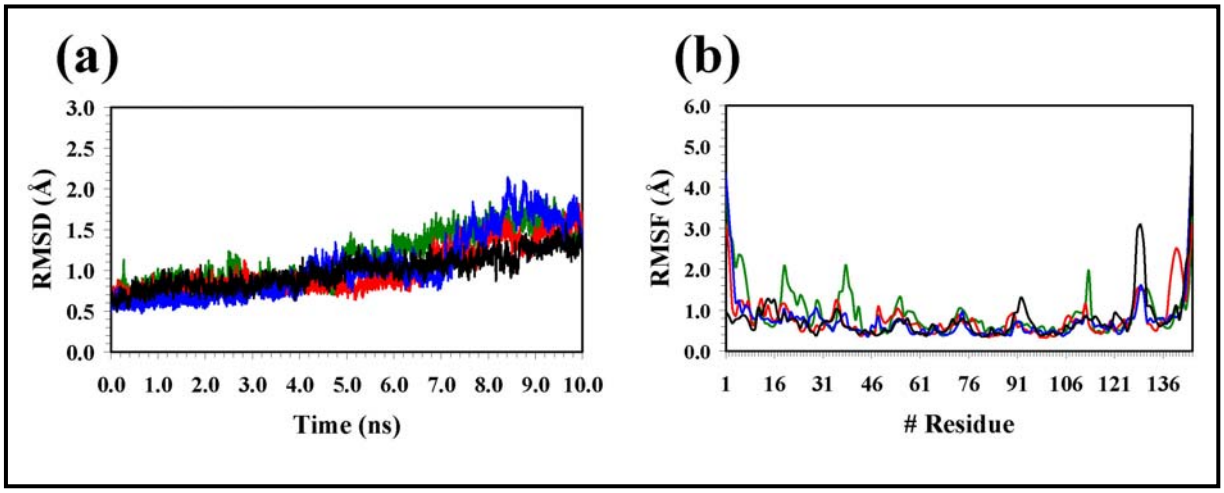


Figure 4. Comparison of the RMSD (a) and RMSF (b) for the wild type 1krr nanotube and the three mutants studied in this work (see Table 2). Both RMSD and RMSF are in Å. The color code is the following: wild type (black), I134TIE-I152TIE mutant (green), I140TIE-I158TIE mutant (red) and I146TIE-V164TIE mutant (blue).

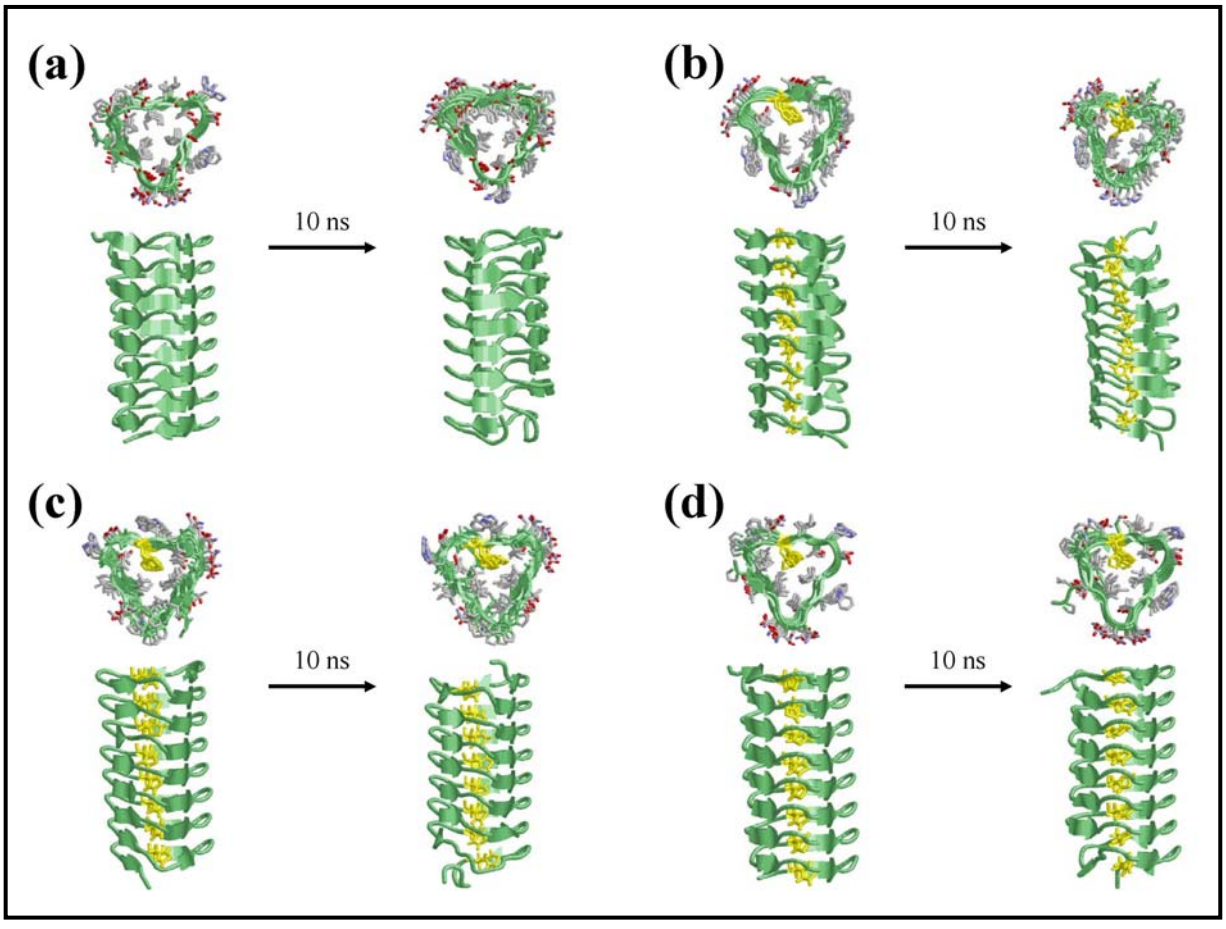


Figure 5. Structure of the nanotube for wild type 1krr nanotube (a), I134TIE-I152TIE (b), I140TIE-I158TIE and (d) I146TIE-V164TIE (d) at the beginning and at the end of the MD simulation. Hydrogen atoms have been omitted for clarity, and the backbone has been represented by solid shapes (arrows indicate sheet conformation). The side chain of the Tie residues has been represented in yellow color.

The conformational stability of the mutated nanotubes was measured by calculating both the evolution of the backbone root-mean-square deviation (RMSD) throughout the simulation time and the root-mean-square fluctuation (RMSF) of individual residues averaged over the whole simulation. RMSD and RMSF were calculated with respect to the N, C^{α} and C(=O) backbone atoms of each residue. In order to check whether the positions chosen for the mutations are good to keep the overall stability of the nanotube, RMSD and RMSF have been compared with the results obtained for the wild type 1krr-nanotube (Figure 4). A detailed inspection of the RMSD and RMSF plots for the three mutated systems compared with the wild type shows that I140TIE-I158TIE and I146TIE-V164TIE mutants enhance the global stability of the nanotube. These later two mutants present a high stabilization in the amino acids located in the center of the nanotubes. This feature is also reflected in Figure 5, which depicts the structure of the nanotubes after 10 ns of MD simulation. The distribution of the φ,ψ dihedral angles evidence that no major perturbation is introduced neither in the mutated positions nor in their adjacent residues for both I140TIE-I158TIE and I146TIE-V164TIE mutants mostly due to the fact that replaced amino acids in the wild type adopt the same $C_{7,eq}$ backbone conformation as Ac-Tie-NHMe. For comparative reasons with Haspel et al.[9] the I146TIE-V164TIE mutated nanotube was selected as a reference structure for further simulations.

Interactions analysis

The mutated 1krr nanotubes present π-stacking interactions involving the aromatic thienyl side chains of the Tie residues located in equivalent positions, i.e. consecutive strands in the inner core of the nanoconstruct. This feature leads to the formation of a ladder of π-stacking, which can promote charge transfer. In the mutants up to seven possible π-stacking interactions can occur, the strength of such interactions depending on the distance between consecutive thienyl rings and their relative orientation (R_d and φ_d, respectively, in Figure 6). Two ideal arrangements can be distinguished for consecutive aromatic thienyl side groups: planar sandwich and T-shaped arrangements, in which the rings are placed coplanar and perpendicular respectively. While Figure 7 depicts the evolution of the distance between the centers of masses of the thienyl side groups of two consecutive Tie residues (R_d), Figure 8 represents the percentage of tilting between these aromatic rings (φ_d). The average R_d value resulting from the MD

trajectory is 4.91 Å ± 0.44 Å, this value being very similar to those obtained when using His instead of Tie.[9] This agreement suggests a similar performance in terms of stability due to the resemblance of the size of both side chains.

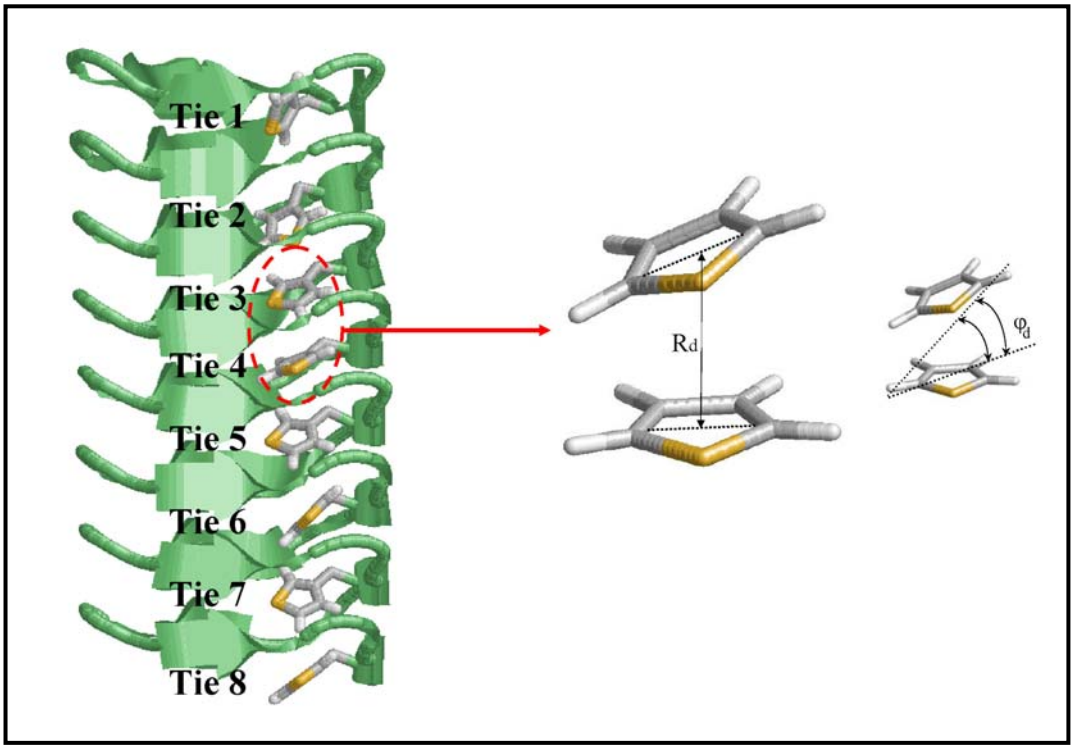


Figure 6. Nanotubular structure of I146TIE-V164TIE mutant, in which the side chains of Tie residues have been numbered and depicted in sticks (left side). Relevant inter-residue parameters are indicated (right side): distance between the centers of masses of two consecutive thienyl rings (Rd) and the angle formed by the planes containing such two rings (φd).

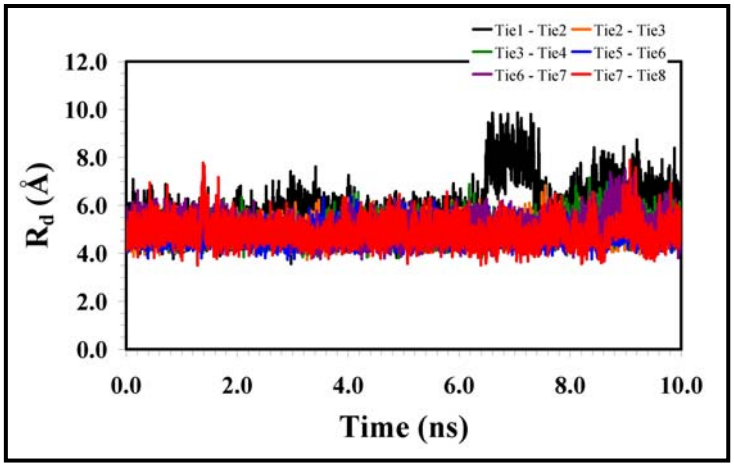


Figure 7. Evolution of the distance between the centers of masses of the side chains of two consecutive Tie residues (Rd) throughout the 10 ns of MD simulation for the I146TIE-V164TIE mutant.

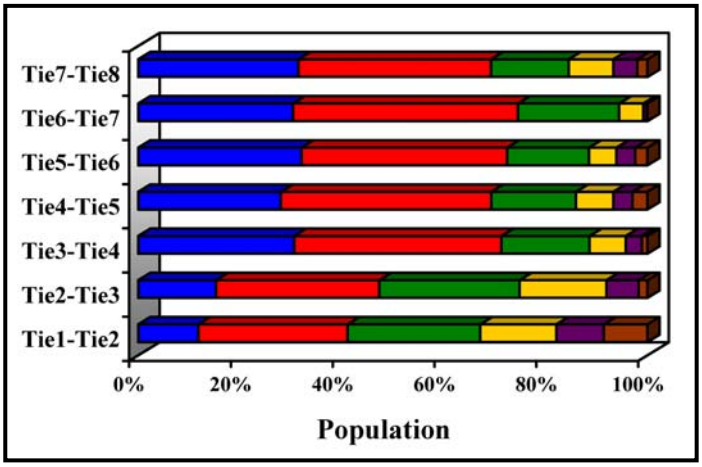


Figure 8. Percentage of tilting angle between the thienyl side groups of two consecutive Tie residues (φd) for the whole 10 ns of MD trajectory: 0º – 15º (blue), 15º – 30º (red), 30º – 45º (green), 45º – 60º (yellow), 60º – 75º (purple) and 75º – 90º (brown).

On the other hand, the average φ_d stemmed from the analysis of the arrangements recorded along whole trajectory is 26.41º ± 16.98º. This average value combined with the results displayed in Figure 8 reveal that thienyl side chains prefer a parallel or slightly tilted arrangement instead of a perpendicular arrangement. This prediction is fully consistent with previous reports,[31] which showed the preference of β-substituted thiophenes to adopt slipped-parallel orientations, *i.e.* π-stacked structures, due to the steric repulsion of the β-substituents. Parallel stacking is stronger than T-shaped stacking because it favors a better overlapping of the π-electrons of the rings. This assertion makes the parallel π-stacking the most relevant for polythiophene based conducting systems.[32] Indeed, a previous study in amorphous polythiophene, *i.e.* dense packing of polymer chains in the solid state, revealed that rings belonging to different chains tends to be separated by less than 6.0 Å suggesting some degree of intermolecular order at distances under this threshold.[33] In the case of the I146TIE-V164TIE mutated nanotube two consecutive thienyl side chains are separated by less than 6.0 Å in 94.26% of the whole MD trajectory. Furthermore, this separation distance co-exists with tilting angles below 20º in 39.17% of the whole trajectory. Accordingly, these results clearly evidence the formation of a ladder of π-stacking in the inner core of the nanotube.

Charge transfer in I146TIE-V164TIE mutated 1krr nanotubes: their use as nanowires

As mentioned in the Introduction section, the stability of the tubular structure together with the formation of a ladder of π-stacking in the inner core suggests the possible existence of charge transfer processes in the I146TIE-V164TIE mutant, which is a potential candidate to be used as a nanowire. In order to examine this hypothesis, QM/MM calculations were performed for both the neutral and the oxidized states. Thus, it is well known that the electrical properties of thiophene based polymers are promoted by p-doping, which allows inter-chain charge transfer.[34]

QM/MM calculations were performed over 51 snapshots that were regularly extracted from the MD trajectory, i.e. the snapshot used as starting point in the simulation plus 50 snapshots at intervals of 0.2 ns. Eight single point QM/MM calculations were performed for each selected snapshot. The system was described as follows in each of such calculations: the QM part consisted on a single thienyl side chain, while the rest of the nanotube and the first solvation shell were represented by point charges that polarize

the QM wave function. For each snapshot the π-electron deficiency was assigned to the thienyl group that presents the lowest energy difference between the neutral and the cationic state among the eight calculations. It is worth noting that as the energy difference between the neutral and cationic states of the QM part are computed, both the MM and the QM/MM van der Waals contributions vanish. The only term that was not canceled in such energy difference, apart from the QM energy, correspond to the QM/MM electrostatic interaction since the polarization of the wave-function is a non-linear process.[35]

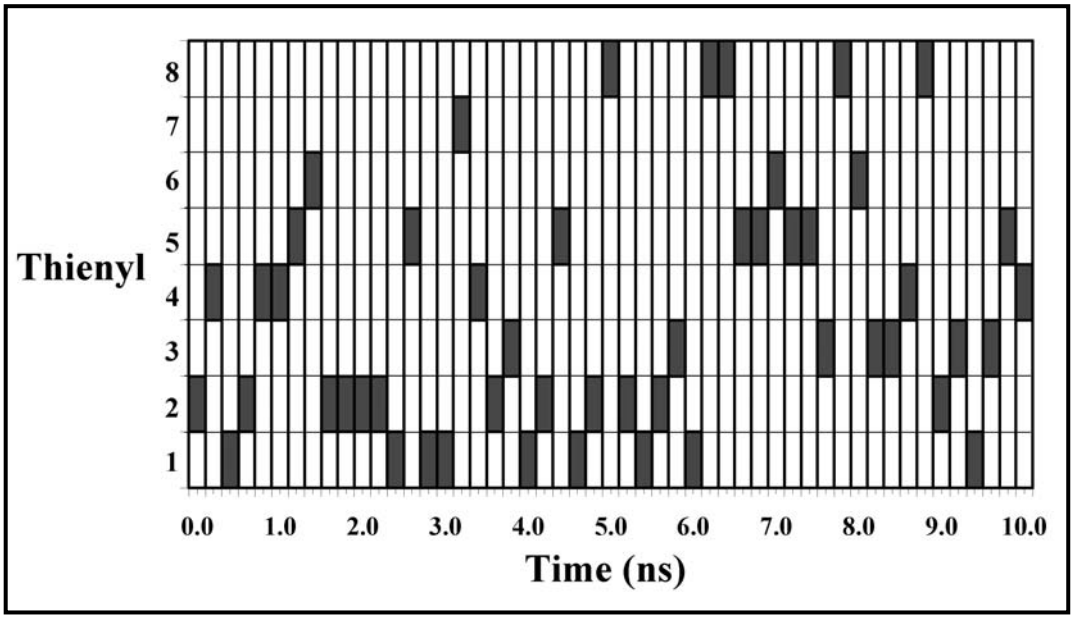


Figure 9. Description of thienyl group hosting the electronic defect along the 10 ns of MD trajectory. In the vertical axis each box corresponds to one of the eight thienyl groups. Gray boxes indicate the thienyl ring hosting the charge defect in a given snapshot. Numeration of thienyl rings corresponds to that showed in Figure 6.

As can be seen in Figure 9, the thienyl group that hosts the π-electron deficiency varies throughout the whole trajectory. This fact evidences the charge transfer between the side chains of the different Tie residues. The averaged stacking parameters of the thienyl groups when hosting the electronic defect are R_d= 4.84 Å ± 0.41 Å and φ_d= 29.27º ± 18.81º. Comparison of these values with those described in the previous section indicates that the conditions for charge transfer are optimum through the whole MD trajectory. Accordingly, QM/MM calculations confirm the good performance of the mutated nanotube to transfer charge and to be used as a nanowire.

5.4.4. *Conclusions*

This work shows that the substitution of natural amino acids by Tie at specific positions of the 1krr based nanotube enables the formation of a π-stacking ladder without disrupting the regular structure, evidencing that this system is a potential candidate for applications related with charge transfer processes, *e.g.* nanowires.

Initially, three mutants of the 1krr motif were selected to construct three nanotubes by joining them through covalent linkages. After 10 ns of NPT-MD in water, it was found that mutant I134TIE-I152TIE was not able to retain its initial tubular structure while I140TIE-I158TIE and I146TIE-V164Tie not only achieved this requirement but improved it compared with the system with the wild-type sequence. However, for comparative reasons further analyses were performed only with the I146TIE-V164TIE mutant. Analysis of the π-stacking interactions between Tie side chains in the inner core of the I146TIE-V164TIE mutant indicated that the average geometric parameters associated to the π-stacking ladder are: R_d= 4.91 Å ± 0.44 Å and φ_d=26.41º ± 16.98º. On the other hand, QM/MM calculations in selected snapshots, which were considered in an oxidized state, evidenced a complete delocalization of the thienyl group hosting the π-electronic defect throughout the whole trajectory, confirming the ability of this nanoconstruct to work as a nanowire.

5.4.5. *References*

[1] Riehemann K.; Schneider S.W.; Luger T.A.; Godin B.; Ferrari M.; Fuchs H. *Angew. Chem. Int. Ed.* **2009**, *48*, 872.

[2] Haydena O.; Agarwalb R.; Lu W. *Nano Today* **2008**, *3*, 12.

[3] Gao X.; Matsui H. *Adv. Mater.***2005**, *17*, 2037.

[4] Haspel N.; Zanuy D.; Alemán C.; Wolfson H.; Nussinov R. *Structure* **2006**, *14*, 1137.

[5] Zheng J.; Zanuy D.; Haspel N.; Tsai C-J.; Alemán C.; Nussinov R. *Biochemistry* **2007**, *46*, 1205.

[6] Zanuy D.; Rodríguez-Ropero F.; Haspel N.; Zheng J; Nussinov R.; Alemán C. *Biomacromolecules* **2007**, *8*, 3135.

[7] Rodríguez-Ropero F.; Zanuy D.; Casanovas J.; Nussinov R.; Alemán C. *J. Chem. Inf. Model.* **2008**, *48*, 333.

[8] Zanuy D.; Rodríguez-Ropero F.; Nussinov R.; Alemán C. *J. Struct. Biol.* **2007**, *160*, 177.

[9] Haspel N.; Zanuy D.; Zheng J; Alemán C.; Wolfson H.; Nussinov R. *Biophys. J.* **2007**, *93*, 245.

[10] Campaigne E.; Bourgeois R.C.; Garst R.; McCarthy W.C., Patrick R.L.; Day H.G. *J. Am. Chem. Soc.***1948**, *70*, 2611.

[11] Rodríguez-Ropero F.; Casanovas J.; Alemán C. *J. Comput. Chem.* **2007**, *29*, 69.

[12] Gaussian 03, Revision B.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, Jr., T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; C. Strain, M.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.;

Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Pittsburgh PA, **2003**.

[13] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.

[14] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1993**, *37*, 785.

[15] McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639.

[16] Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

[17] Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781.

[18] Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

[19] Wang J.; Cieplak P.; Kollman P.A. *J. Comput. Chem.* **2000**, *21*, 1049.

[20] Alemán, C.; Luque, F. J.; Orozco, M. *J. Comput. Aided Mol. Design* **1993**, *7*, 721.

[21] Jorgensen, W. L.; Chandrasekhar J.; Madura J. D.; Impey R.W.; Klein M. L.. *J. Chem. Phys.* **1982**, *79*, 926.

[22] Darden, T.; York D.; Pedersen L. *J. Chem. Phys.***1993**, *98*, 10089.

[23] Ryckaert, J.P.; Ciccotti, G.; Berendsen, H.J.C. *J. Comput. Phys.* **1977**, *23*, 327.

[24] Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

[25] Maseras, F.; Morokuma K. *J. Comput. Chem.***1995**, *16*, 1170.

[26] Chass G.A.; Mirasol R.S.; Setiadi D.H.; Tang T-H.; Chin W.; Mons M.; Dimicoli I.; Dognon J-P.; Viskolcz B.; Lovas S.; Penke B.; Csizmadia I.G. *J. Phys. Chem. A* **2005**, *109*, 5289.

[27] Reynolds, C. A.; Essex, J. W.; Richards, W. G. *J. Am. Chem. Soc.* **1992**, *114*, 9075.

[28] Alemán, C.; Casanovas, J. *J. Chem. Soc., Perkin Trans 2* **1994**, 563.

[29] Casanovas, J.; Zanuy, D.; Nussinov, R.; Alemán, C. *J. Org. Chem.* **2007**, *72*, 2174.

[30] Casanovas, J.; Zanuy, D.; Nussinov, R.; Alemán, C. *Chem. Phys. Lett.* **2006**, *429*, 558.

[31] Tsuzuki, S.; Honda, K.; Azumi, R. *J. Am. Chem. Soc.* **2002**, *124*, 12200.

[32] Zhang, G.; Pei, Y.; Ma, J.; Yin, K.; Chen, C. L. *J. Phys. Chem. B* **2004**, *108*, 6988.

[33] Curcó, D.; Alemán, C. *J. Comput. Chem.* **2007**, *28*, 1743.

[34] Graf D.D.; Campbell J.P.; Miller L.L.; Mann K.R. *J. Am. Chem. Soc.* **1996**, *118*, 5480.

[35] Jacquemin D.; Perpète E.A.; Laurent A.D.; Assfeld X.; Adamo C. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1258.

6. MOLECULAR ACTUATORS AND SENSORS BASED ON CONDUCTING POLYMERS

6.1. Molecular Dynamics of a Calix[4]arene-containing polymer in dichloromethane solution: Ability of the solvent molecules to fill the cavity of the macrocycle*

6.1.1. *Introduction*

Calix[4]arenes are synthetic macrocyclic molecules consisting of four phenol or anisole rings connected via methylene bridges. These systems present a significant conformational flexibility, which arises from an oxygen-through-the-annulus rotation mechanism, *i.e.* the methylene groups act as hinges. The most relevant conformations are the *cone* with all the rings oriented in the same direction, and the partial-*cone* (*paco*), in which one ring is rotated with respect to the other three. Thus, all the other possible states of calix[4]arenes are combinations of such two conformations, *i.e.* alternated conformations with two rings rotated with respect to the other two. The cone-to-paco rotational isomerism has been investigated using not only NMR spectroscopy[1-3] but also theoretical methods based on both quantum mechanics[4-6] and molecular mechanics.[7,8]

The cavity of the cone conformation of calix[4]arenes posses a π-rich character defined by the four aromatic rings, which is suitable for the inclusion of neutral guests of complementary size. Indeed, inclusion of organic molecules, such as chloroform and nitrobenzene, in one of the most studied families of calix[4]arenes, *p-tert*-butylcalix[4]arene derivatives, *i.e*. a calix with *tert*-butyl groups at the upper rim, have been observed by X-ray structure determination[9-11] and NMR.[10,12,13] Furthermore, a recent Molecular Dynamics (MD) study of *p-tert*-butylcalix[4]arene in chloroform solution evidenced how one solvent molecule is firmly captured in the cavity of the *cone* conformation.[8] The immobilization of such chloroform inside the cavity was proved to stabilize the cone conformation influencing the conformational inversion rate of the cone-to-paco isomerization process. The overall of these observations suggests that the solvent molecules are expected to have a pronounced effect on the dynamics of calix[4]arenes, even when their cavity is smaller than that of *p-tert*-butylcalix[4]arene.

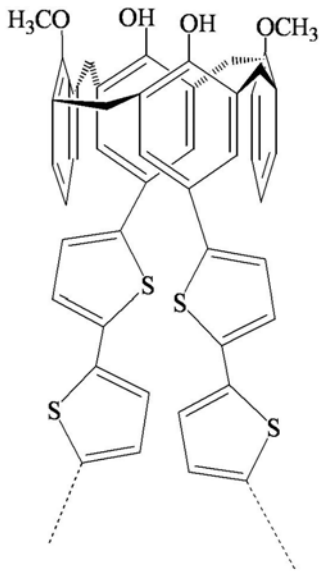
The distribution of the solvent molecules around the *p-tert*-butylcalix[4]arene was recently reported by Briels and co-workers.[8] These authors used 1 ns MD trajectories of a system constituted by this macrocycle and 386 explicit chloroform molecules to

* Results presented in this chapter are published in ***J. Comput. Chem.*** **2008,** ***29,*** 1233

describe the structure of the first solvation shell. However, these trajectories were too short to provide dynamical information, the analysis being restricted to provide a static description of the density distribution of the solvent molecules.

In recent studies it has been examined the actuation mechanism of poly(calix[4]arene bis-bithiophene), an artificial muscle constituted by small thiophene oligomers (quarterthiophene) separated by calix[4]arene scaffolds (25,27-dihydroxy-26,28-dimethoxy calix[4]arene).[14,15] In particular, the contraction process of a model molecule formed by 6 calix[4]arene units and 5 quarterthiophene units was investigated at the miscroscopic level using MD simulations in dichloromethane solution.[15] Results derived from a 1-μs trajectory revealed not only the non-concerted nature of the actuation mechanism but also the large influence of the solvent on the contraction, which reduce the rate of the process. However, no analysis of this large MD trajectory was performed to understand the dynamics of the interaction between the calix[4]arene scaffolds and the solvent molecules contained in their first solvation shell.

In this chapter a detailed structural and dynamical analyses of the solvent in the first solvation shell around the calix[4]arene scaffolds contained in poly(calix[4]arene bis-bithiophene) is presented. For this purpose, the coordinates of the $5 \cdot 10^5$ snapshots recorded from the 1-μs trajectory developed were employed to examine the actuation mechanism of the latter system.[15] Particular attention has been paid to the structure and dynamics of the captured dichloromethane molecules.



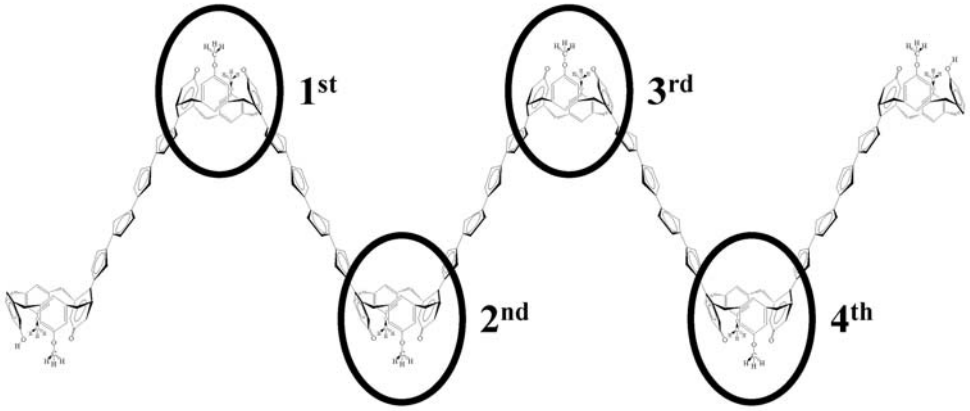
Scheme 1

6.1.2. *Methods*

The molecular system

The chemical structure of the repeating unit of the poly(calix[4]arene bis-bithiophene) is represented in Scheme 1. The model system used to study this molecular actuator was constituted by 6 calix[4]arene units and 5 quarterthiophene fragments, which was placed inside a previously equilibrated box with 10842 molecules of dichloromethane. After removing the solvent molecules overlapping any atom of the polymer, the final system contained 32526 explicit atoms. Figure 1, which shows schematic representations of the simulation box, provides information about the dimensions and conformation of the model molecule at the beginning and end of the simulation.

Analyses have been performed considering the four central calix[4]arene scaffolds of the model molecule (Scheme 2). These calix[4]arene units present identical composition: two phenol rings with quartherthiophene substituents at the *para* position with respect to the hydroxyl group and two unsubstituted anisole rings. Thus, the calix[4]arene units located at the end of the model molecule have not been considered in the analyses because the quartherthiophene substituent is incorporated to only one anisole ring.



Scheme 2

The results provided in this chapter should be considered equivalent to those derived from four independent 1-μs trajectories on a single calix[4]arene scaffold since, as was mentioned above, it was found that all the actuating units contained in the model molecule act as independent entities.[15]

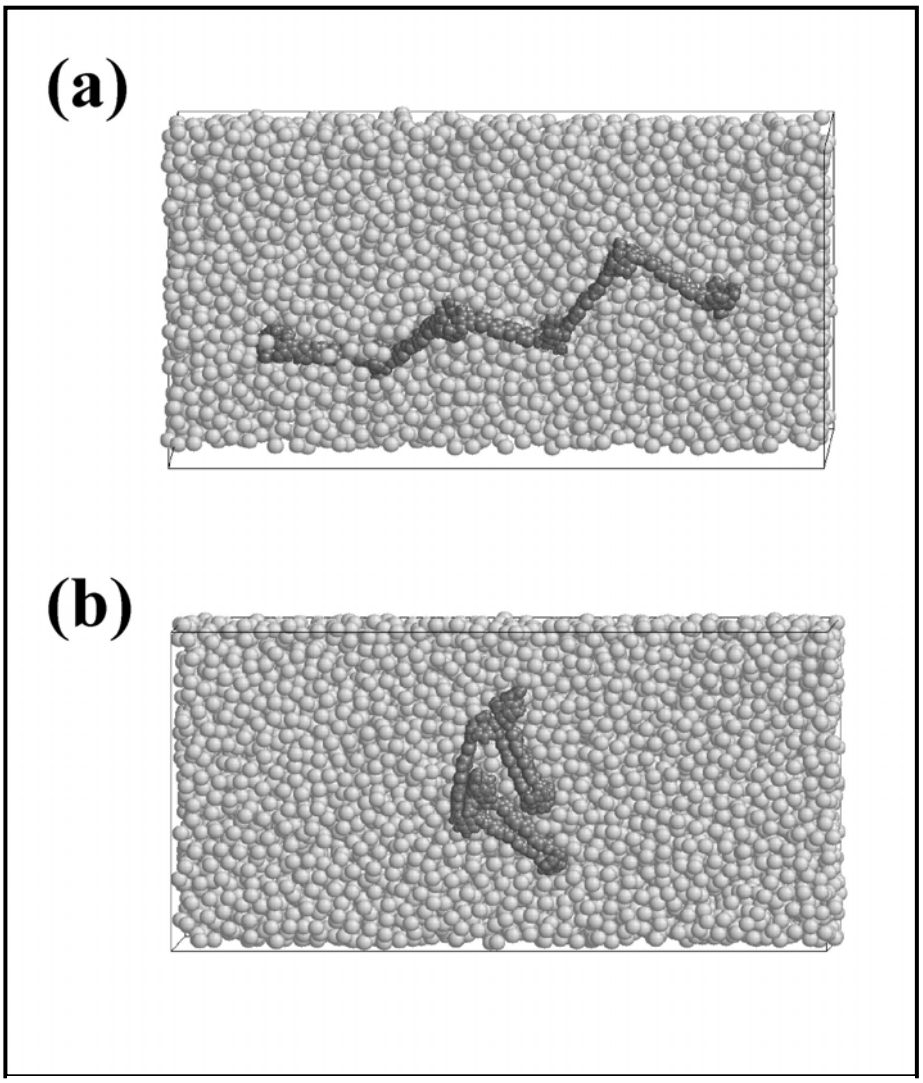


Figure 1. Schematic representation of the simulation box at the beginning (a) and end (b) of the MD simulation. The model molecule is represented in dark gray, while solvent molecules are indicated by light gray solid balls.

Computational details

The 1-μs MD was performed using the high performance parallel work program NAMD 2.5,[16] which was run on 512 CPUs at the MareNostrum supercomputer of the Barcelona Supercomputing Center (BSC). The bonding and van der Waals force-field parameters for the calix[4]arene units and thiophene rings were taken from literature,[17-19] while the electrostatic parameters were specifically developed using quantum mechanical calculations.[14] Finally, the OPLS3 model was used to represent dichloromethane.[20]

Van der Waals interactions were evaluated by applying a cutoff distance at 14 Å. while electrostatic interactions were computed using the Particle-Mesh Ewald (PME) method.[21] The real space term was determined by the van der Waals cut off (14 Å),

while the reciprocal term was estimated by interpolation of the effective charge into a charges mesh with a grid thickness 5 points per volume unit. Bond lengths of the calix[4]arene units were constrained using the *SHAKE* algorithm.[22]

Before starting the production run series, the molecular system was allowed to spontaneously reach a steady state by two consecutive *NVT*-MD (0.5 ns) and *NPT*-MD (1.0 ns) runs, starting from a previously relaxed ensemble (simple energy minimization). The temperature and simulation box size were controlled by the weak coupling method, the Berendsen thermo-barostat,[23] in order to maintain the experimental conditions. The *NPT*-MD production run lasted up to 1-μs long. The numerical integration step was 2 fs and the coordinates were saved every 2 ps intervals for subsequent analysis. More technical details about the simulations presented in this work are reported in reference 15.

6.1.3. *Results and discussion*

Structure of the solvent

Figure 2 shows the radial distribution function of solvent molecules around the four calix[4]arene units, which were calculated considering the carbon of the dichloromethane as reference. The function was calculated relative to the center of mass of each macrocycle, *i.e.* without considering the quarterthiophene substituents, using the snapshots recorded through the 1-μs MD. It is worth noting that in all cases a peak followed by a region with smaller probability is located at 3.5-4.0 Å. This suggests that one dichloromethane molecule is located inside the cavity of each calix[4]arene unit during a significant period of time. Furthermore, the profiles displayed in Figure 2 reveal that the behavior of the four calix[4]arene units differs significantly. Thus, the first of the 1st unit is high and sharp, while the same peak is smaller for the 2nd unit. Finally, the first peak of the 3rd and 4th units is broad and small, respectively. As will be showed below, these differences are related with the individual dynamics of the each calix[4]arene unit but also with the whole dynamics of complete model used to represent the poly(calix[4]arene bis-bithiophene) chain.

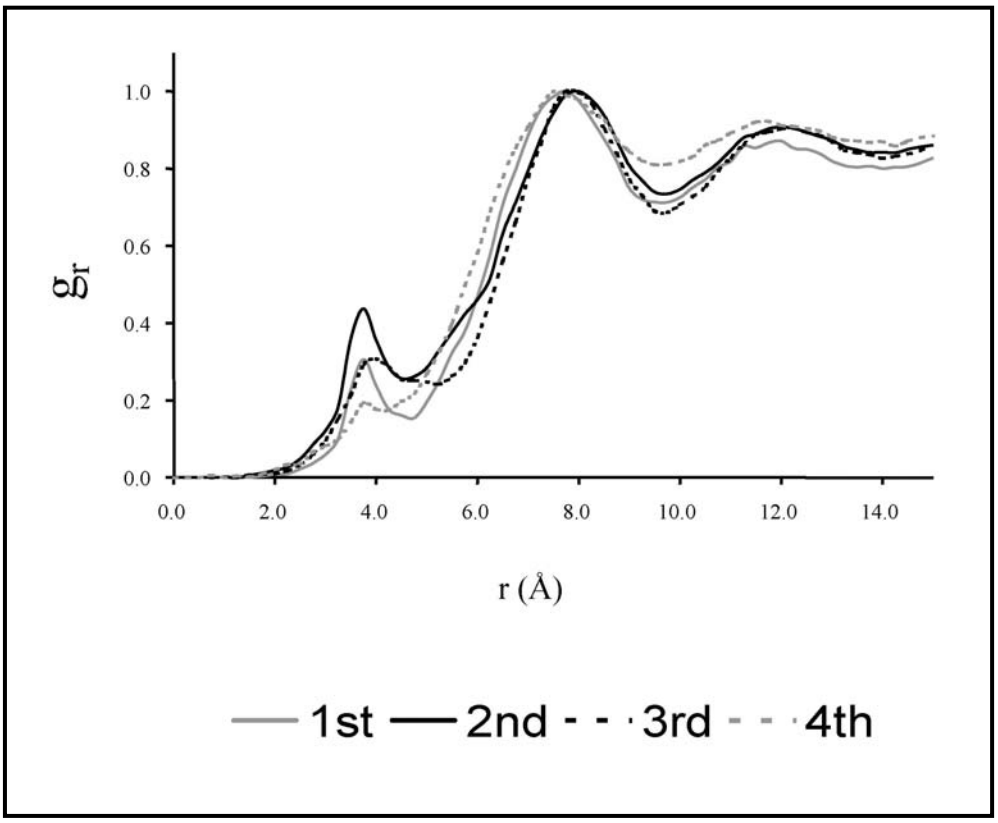


Figure 2. Radial distribution function of the carbon atom of dichloromethane relative to the center of mass of the 1st, 2nd, 3rd and 4th calix[4]arene unit of the model molecule studied in this work.

Figure 3 compares for each calix[4]arene unit the radial distribution functions of dichloromethane solvent molecules along five consecutive 200 ns periods, in which the 1-μs MD simulation has been divided. It is worth noting that the shape of the function at distances smaller than ~5 Å changes drastically along the simulation suggesting that the structure of the solvent around the calix[4]arene units, particularly inside the cavity, varies with the structure of the poly(calix[4]arene bis-bithiophene). On the other hand, analyses of Figures 1 and 2 reveals that the first peak of the distribution function is in average centered at r= 3.75 Å. This distance, which is defined by the center of mass of the macrocycle and the carbon atom of the dichloromethane, has been used in next sections to identify the solvent molecules that are inside the cavity of each calix[4]arene unit within a given snapshot.

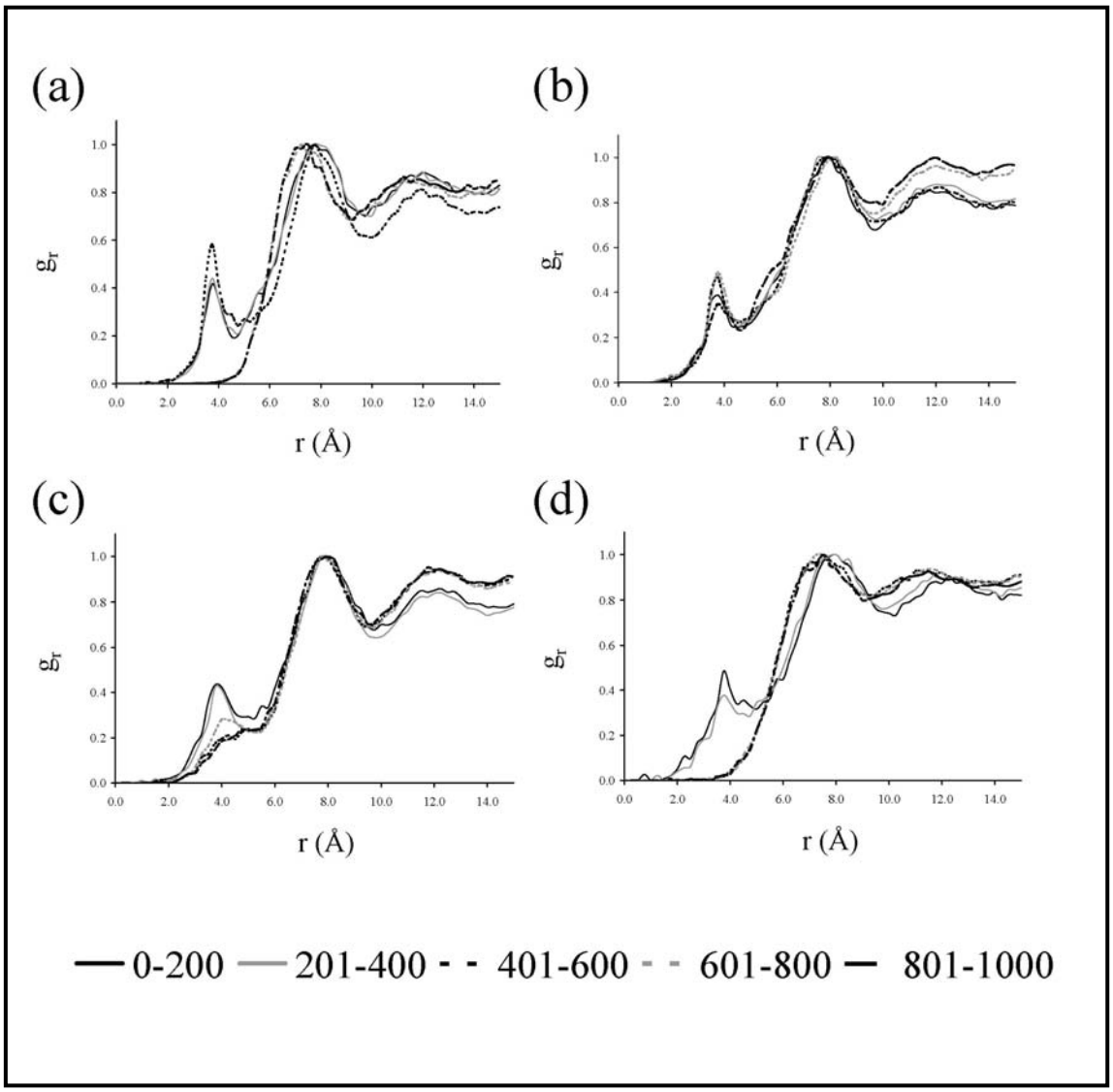


Figure 3. Radial distribution functions of the carbon atom of dichloromethane relative to the center of mass of the (a) 1st, (b) 2nd, (c) 3rd and (d) 4th calix[4]arene unit along each of the five 200 ns periods considered in the 1-μs MD simulation.

Molecules filling the cavity

Analysis of the coordinates stored every 2 ps confirmed that the cavity defined by each calix[4]arene unit in cone conformation is filled by a dichloromethane molecule with significant frequency. Thus, the sum of times associated to the stored configurations indicate that solvent molecules were within the cavity of the 1st, 2nd, 3rd and 4th calix[4]arene unit during 124.3, 235.9, 38.8 and 47.2 ns, respectively, of the whole simulation. These times, hereafter denoted filling period (t_{fill}), suggest that the behavior of the four units is not concerted. Indeed, the t_{fill} values are intimately correlated with

the dynamics of the poly(calix[4]arene bis-bithiophene) actuator, which was described in detail previously.[15]

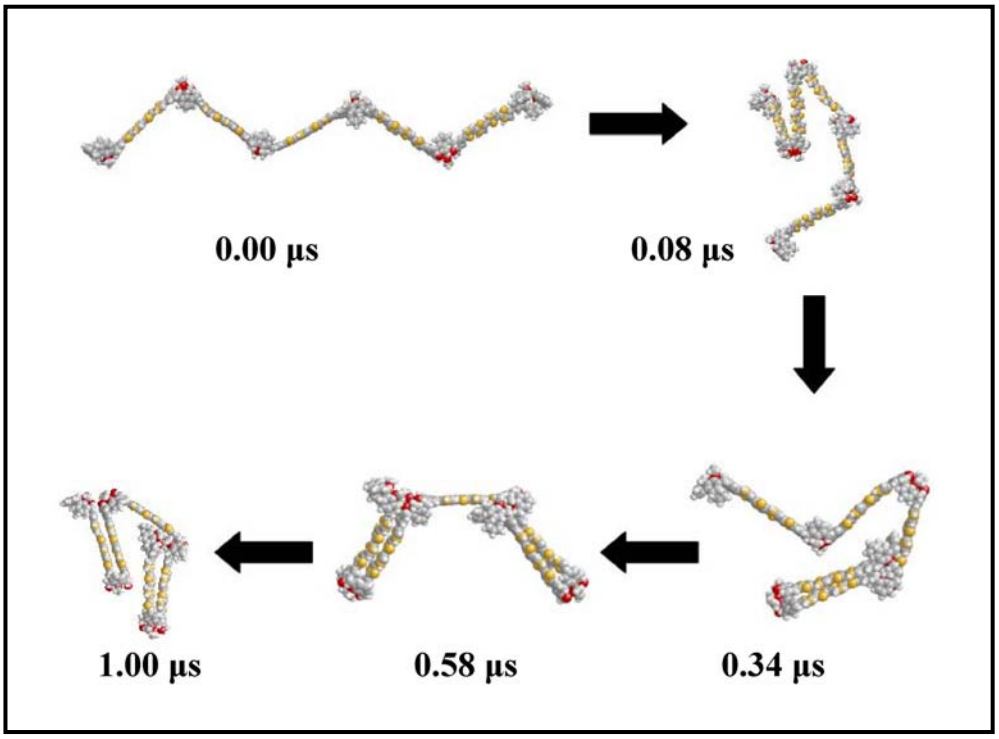


Figure 4. Atomistic representation of selected snapshots to illustrate the contraction of the actuator through the 1µs MD simulation.

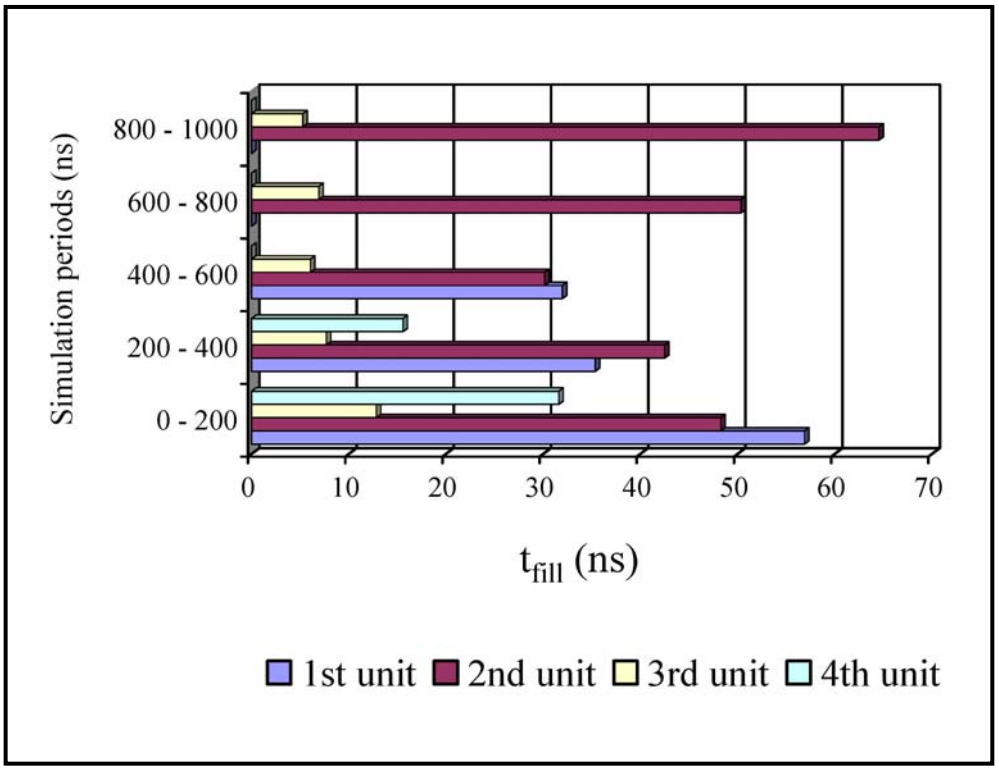


Figure 5. Time with solvent molecules filling the cavity of the four calix[4]arene units (t_{fill}) along each of the five 200 ns periods considered in the 1-µs MD simulation.

Figure 4 shows selected snapshots of the simulated poly(calix[4]arene bis-bithiophene) actuator, while Figure 5 depicts the value of t_{fill} in the five 200 ns periods used to divide the whole simulation. The 3rd calix[4]arene unit, which shows the lowest t_{fill}, underwent a spontaneous and irreversible, at least in the 1 μs range, cone-to-paco rotational isomerism. After this conformational transition, the contribution to t_{fill} was very small indicating that the filling of the cavity is easier for the cone than for the paco conformation. The 4th calix[4]arene scaffold is contained in an actuating unit that was fully contracted after about 300 ns. Thus, after such process the two quarterthiophene fragments attached to such calix[4]arene scaffold were so close that no solvent molecule was allowed inside the cavity. On the other hand, the actuating unit associated to 1st calix[4]arene was fully contracted after about 550 ns, no other contribution to t_{fill} being detected after this process. Finally, the 2nd calix[4]arene was always accessible to the solvent, the contribution to t_{fill} being approximately constant during the whole simulation. This is because one of the quarterthiophene substituents attached to it was also shared by the 3rd calix[4]arene unit, the latter being unable to contract due to the conformational transition mentioned above.

The interplay between the cavity of each calix[4]arene unit and the solvent molecules is understood by examining the distribution of the solvent molecules that are inside the cavity. In Figure 6 scatter plots are given of the positions of the carbon atoms of the dichloromethane molecules that are separated by less than 3.75 Å from the center of mass of macrocycle. Note that these plots give information about the preferred positions of the solvent molecule. As can be seen, a four-fold symmetry related with the symmetry of the macrocycle is visible for the positions of the solvent in the cavities of the 1st, 2nd and 4th calix[4]arene units, with solvent molecules preferably oriented towards the center of each phenol/anisole ring. Obviously, this situation is less clear for the 4th calix[4]arene unit than for the other two since, as was mentioned above, t_{fill} was very short in this case. A completely different situation is showed by the 3rd calix[4]arene unit, in which the only region of high solvent density lies facing one of the anisole rings, *i.e.* the four-fold symmetry detected for the 1st, 2nd and 4th units transforms into a two-fold symmetry. This must be attributed to the distortion of the macrocycle produced during the first steps of the cone-to-paco rotational isomerism, which starts after a few ns of MD simulation. Indeed, inspection to the density distribution for the

paco arrangement, which was preserved during the last ~920 ns of simulation, reveals that the solvent is located at the center of the cavity.

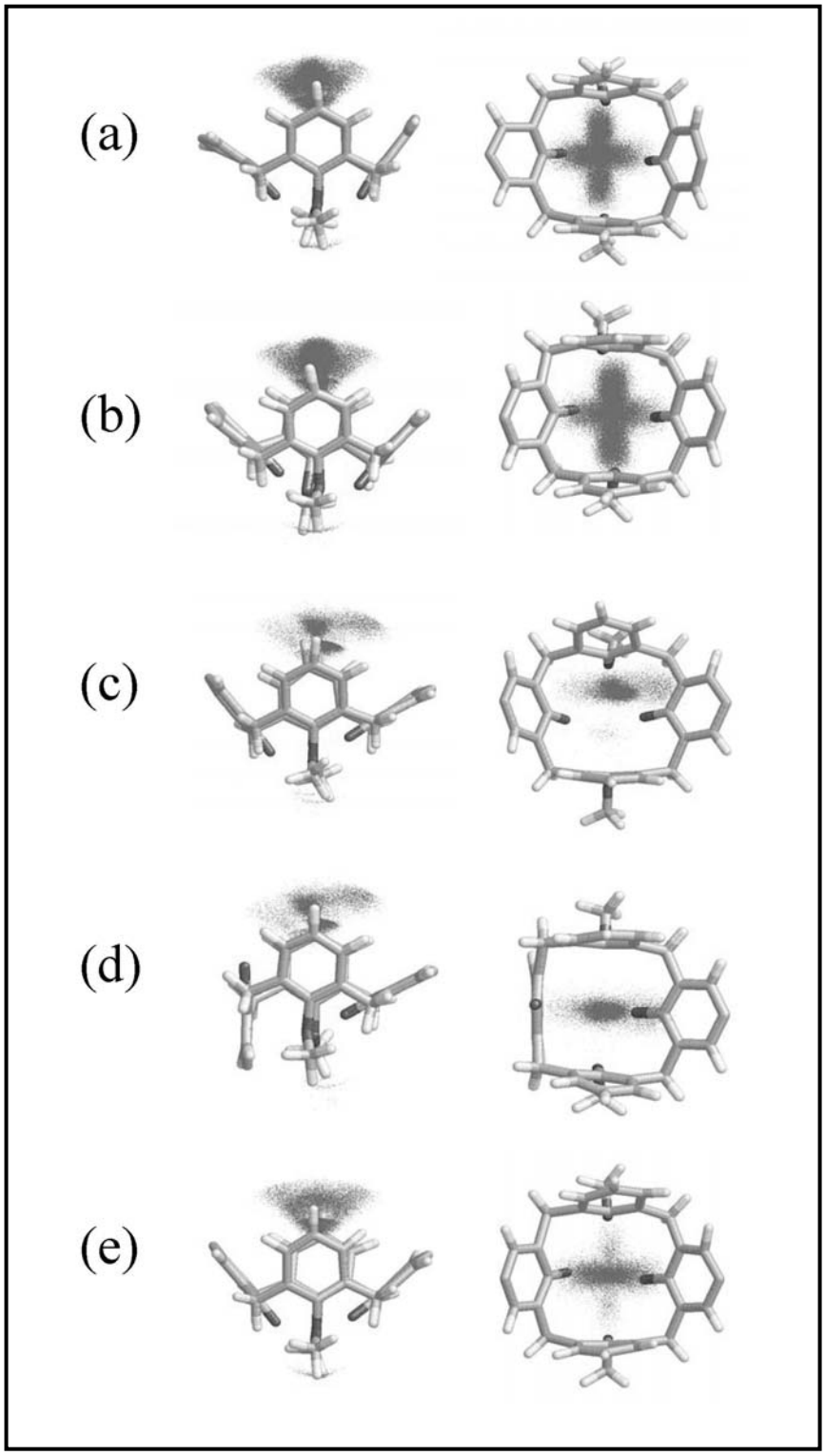


Figure 6. Scatter plots of the positions of the carbon atoms of the dichloromethane molecules that are separated by less than 3.75 Å from the center of mass of each calix[4]arene unit. The plots show to the distribution of the solvent molecules filling the cavity of the: a) cone conformation of the 1st unit; (b) cone conformation of the 2nd unit; (c) cone conformation of the 3rd unit; (d) paco conformation of the 3rd unit; and (e) cone conformation of the 4th unit. The left column gives a side view, and the right column a top view. The configurations shown for the 1st, 2nd and 4th macrocyles are averaged over the whole simulation, while that shows for the 3rd unit corresponds to the average before and after the conformational transition, respectively.

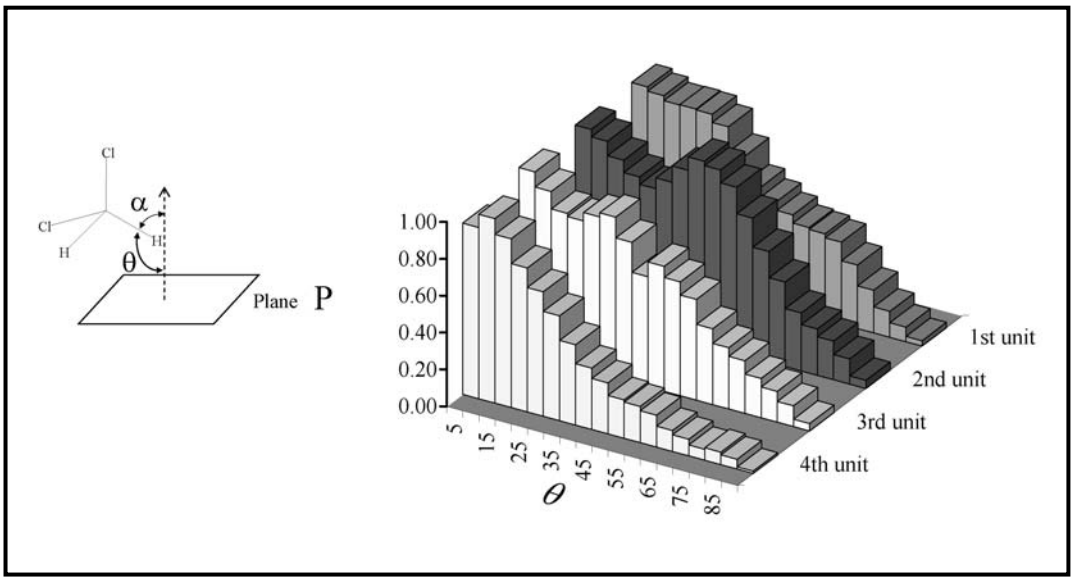


Figure 7. Probability distribution of the angle θ (in degrees) describing the orientation of the dichloromethane molecule located inside the cavity of the four calix[4]arene units (see text). The distribution has been normalized with respect to the solvent molecules for the most populated θ value.

The orientation of the dichloromethane molecule inside the cavity has been described by the angle θ, which defines the incidence angle on the plane formed by the center of mass of the phenol and anisole rings of the bond formed by the chloride atom closest to the center of mass of calix[4]arene and the carbon atom of dichloromethane. More specifically, θ was calculated as 90º-α, where α is the angle by the vector normal to the plane and the C-Cl bond. Note that the C-Cl bond is perpendicular to the plane defined by the four rings of the macrocycle when θ= 90º. Figure 7 shows the probability distribution of θ for the four calix[4]arene units. It is worth noting that the perpendicular orientation was seldom found, its proportion being almost negligible. Indeed, the θ values preferred by the 1st and 4th calix[4]arene suggests that one of the hydrogens of the dichloromethane is pointing towards the cavity. Unfortunately, the hydrogen atoms are not explicitly represented in the OPLS model of dichloromethane and, therefore, this fact has been inferred. The preferred value of the 2nd calix[4]arene ring is $\theta \approx$ 45º suggesting that neither the C-Cl nor the C-H bond points towards the cavity, *i.e.* a tilted orientation of the dichloromethane is favored. Finally, the θ values found for the 3rd calix[4]arene spreads out from 0º to 65º. Thus, although parallel orientation is still slightly preferred, the orientation is also affected by the particular conformation behavior of the 3rd ring.

Dynamics of the Solvent Molecules

In order to ascertain if the solvent molecule that fills the cavity of the macrocycle resides during a long period or, in opposition, is easily substituted by another molecule of the bulk, the 1-μs trajectory was analyzed in detail. Interestingly, it was found that the number of different solvent molecules that were located inside the cavity (N) of the 1st, 2nd, 3rd and 4th calix[4]arene units is 1040, 1640, 1682 and 654, respectively, the average residence times being 0.12, 0.14, 0.02 and 0.07 ns. The residence time was defined as the time that a given molecule spends continuously within the cavity. However, these average values lack meaning since the residence time (t_{res}) range from very low values (typically $2 \cdot 10^{-3}$ ns) to relatively high values, the highest t_{res} being 1.73, 1.55, 0.23 and 0.72 ns for the 1st, 2nd, 3rd and 4th calix[4]arene unit, respectively. Figure 8 represents the t_{res} for the N molecules of each calix[4]arene unit.

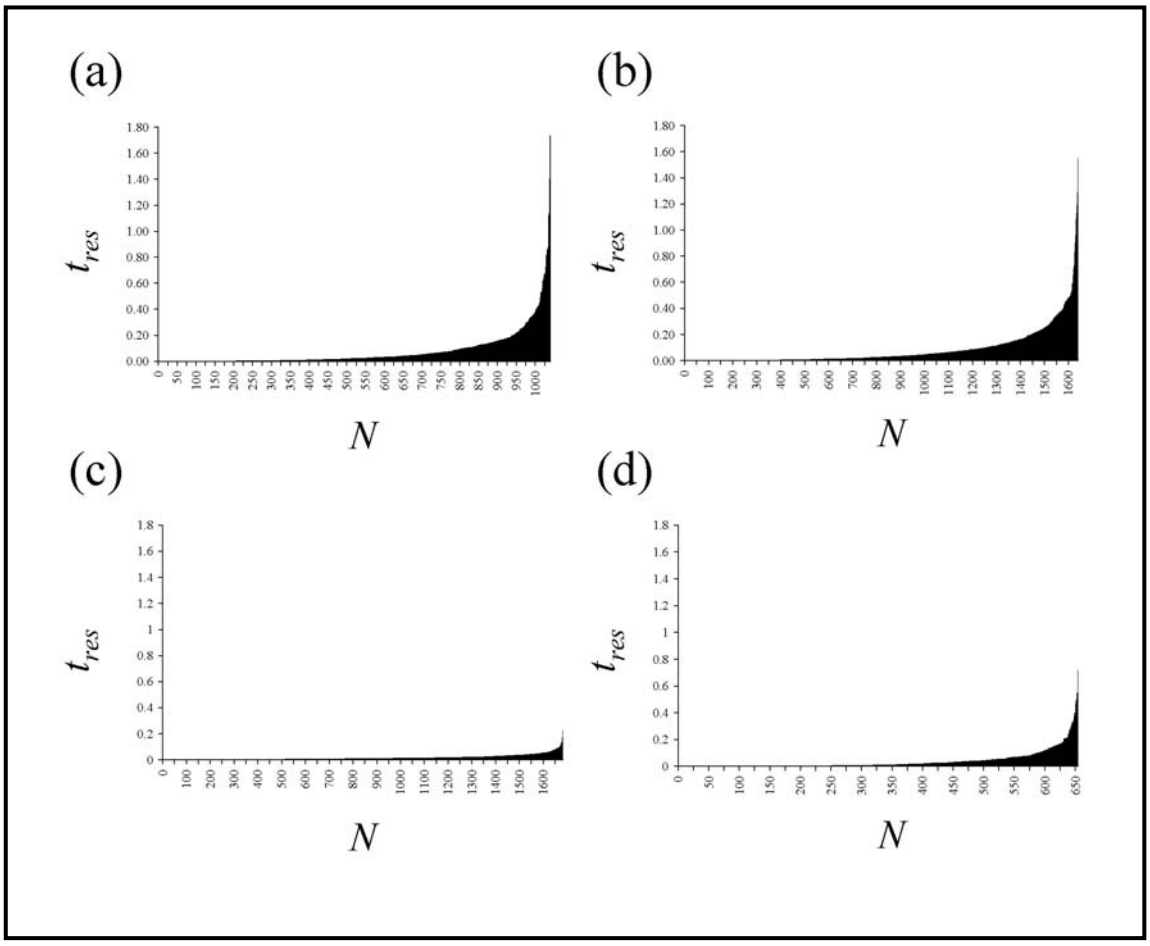


Figure 8. Representation of the residence time (t_{res}; in ns) for the N dichloromethane molecules contained inside the cavity of the (a) 1st, (b) 2nd, (c) 3rd and (d) 4th calix[4]arene unit. Molecules have been ordered in increasing order.

As can be seen, the N with t_{res} > 0.1 ns is 231, 387, 14 and 67 for the 1st, 2nd, 3rd and 4th calix[4]arene unit, respectively, while the value of N with t_{res} > 0.5 ns is 27, 32, 0 and 3. These results indicate that, although the N value is very high and the cavity of the macrocycle is frequently filled with a dichloromethane molecule, as showed above, only a few solvent molecules are actually captured in the cavity, *i.e.* only a few molecules show residence times greater than 0.5. Thus, the most frequent situation is that the dichloromethane inside the cavity is rapidly substituted by another molecule of the bulk. The latter feature is specially apparent for the 3rd calix[4]arene unit, in which the cavity of the paco conformation is too small to capture a solvent molecule.

In a very recent study based on relatively short MD simulations, *i.e.* about 1 ns each, Tolpekina *et al.*[8] revealed that the *p-tert*-butylcalix[4]arene in the cone arrangement captures a chloroform molecule inside its cavity that hardly ever escapes. Thus, the same solvent molecule was observed to reside inside the cavity during the whole simulation. This was in agreement with X-ray diffraction experiments on the solid state, which evidenced the inclusion of a chloroform molecule in the cavity of *p-tert*-butylcalix[4]arene.[12] It should be noted that X-ray crystallography is only able to detect solvent molecules with very high residence times. In contrast, the paco conformation was not capable of immobilizing a solvent molecule. On the other hand, simulations on the most simple calix[4]arene, *i.e.* the unsubstituted compound, showed that the chloroform molecules hover just above the macrocycle, but this molecule was exchanged regularly and rapidly for another chloroform.[8] It is worth noting that the results described in this work for the calix[4]arene units of poly(calix[4]arene bis-bithiophene) are fully consistent with those reported by Tolpekina *et al.*[8] Thus, in general the solvent molecules resides less time in the cavity defined by such units than in that of *p-tert*-butylcalix[4]arene. This is because the encapsulation of the solvent molecule is easier in the cavity of the latter compound, which is formed by the four bulky *tert*-butyl substituents. On the other hand, as expected from the results of Tolpekina *et al.*[8] the cavity of the calix[4]arene unit in paco arrangement is too small for an inclusion.

Table 1. Distribution of the solvent molecules that form inclusion complexes with two, three and four calix[4]arene units during the 1 μs MD simulation: Ns indicates the number of solvent molecules involved in each case, while 1st, 2nd, 3rd and 4th correspond to the label of the calix[4]arene units (see Scheme 2).

2 calix[4]arene units		3 calix[4]arene units		4 calix[4]arene unit	
Units	N_s	Units	N_s	Units	N_s
$1^{st} - 2^{nd}$	159	$1^{st} - 2^{nd} - 3^{rd}$	24	$1^{st} - 2^{nd} - 3^{rd} - 4^{th}$	3
$1^{st} - 3^{rd}$	178	$1^{st} - 3^{rd} - 4^{th}$	20		
$1^{st} - 4^{th}$	82	$2^{nd} - 3^{rd} - 4^{th}$	25		
$2^{nd} - 3^{rd}$	261				
$2^{nd} - 4^{th}$	105				
$3^{rd} - 4^{th}$	126				

In order to ascertain if the dichloromethane molecules that fill the cavities are able to migrate from one calix[4]arene unit to another, the trajectories of the solvent molecules with $t_{res} \geq 2 \cdot 10^{-3}$ ns were examined in detail. Surprisingly, it was found that 983 solvent molecules interact with two or more calix[4]arene units along the 1-μs MD simulation. More specifically, 911, 69 and 3 dichloromethane molecules were detected in inclusion complexes formed by 2, 3 and 4 different calix[4]arene units, respectively. The distribution of these solvent molecules is provided in Table 1. The migration of a dichloromethane from one calix[4]arene unit to the others is illustrated in Figure 9, which shows the position of a given solvent molecule at different times through the simulation. As can be seen, in less than 0.20 ns the solvent molecule is able to be inside the cavities of the four macrocycles. However, it should be noted that this is physically unfeasible without the conformational contraction undergone by the polymer. Thus, assuming that the coefficient of diffusion for dichloromethane is within the range typically observed for organic solvents ($D \approx 10^{-5}$ $cm^2 \cdot s^{-1}$), Einstein equation ($D = \langle x^2 \rangle / 6t$) predicts that the maximum displacement for a solvent molecules in 0.2 ns is about 10 Å. This distance is significantly lower than that between the cativities of the two closest calix[4]arene units if the actuator is not contracted. Thus, the same solvent molecule cannot visit the cavities of the four macrocycles in such short period of time unless the contraction of the actuator reduces the distance to run over, *i.e.* the cooperation between the solvent-polymer systems is required.

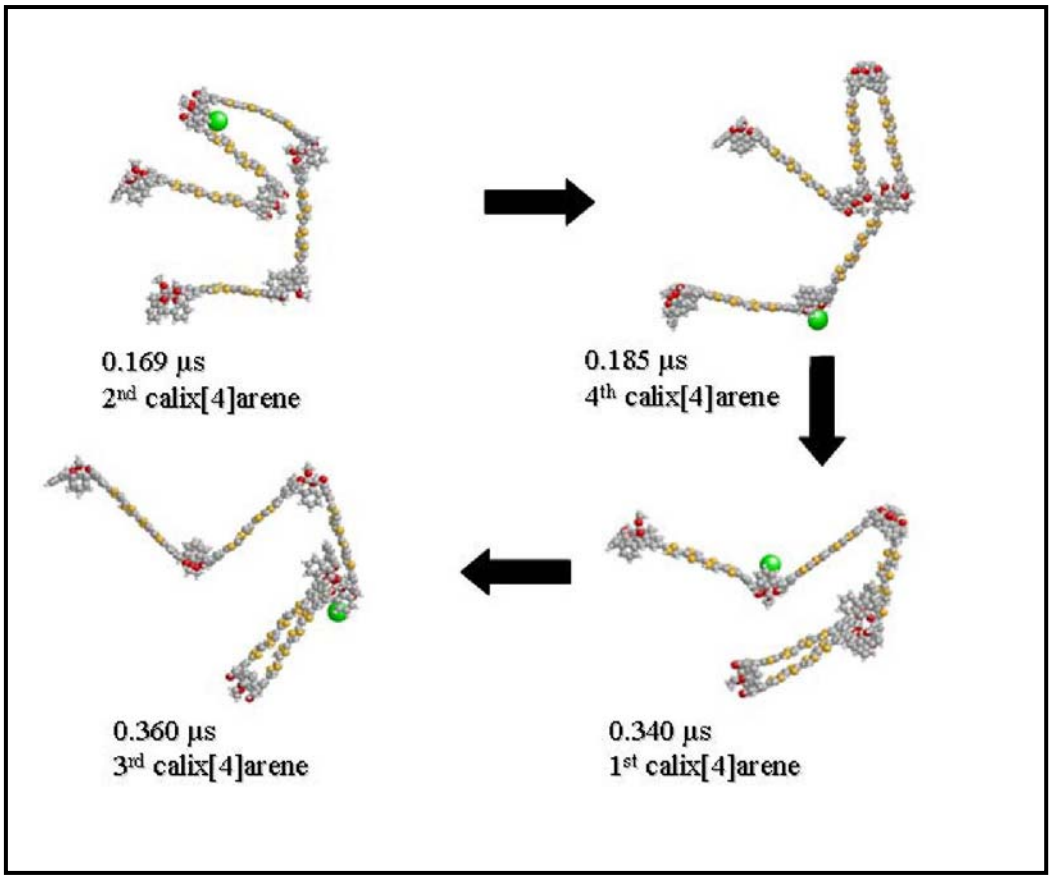


Figure 9. Dynamics of a given dichloromethane molecule that is able to be inside the cavity of the four calix[4]arene units along the MD simulation. The picture shows the temporal evolution of such solvent molecule in a relatively short period of time. Both the time of the recorded snapshot and the calix[4]arene unit that contains the solvent molecule inside its cavity is indicated in each case.

6.1.4. *Conclusions*

Analysis of the snapshots recorded during the 1-μs MD simulation of poly(calix[4]arene bis-bithiophene) in dichloromethane solution allows us to draw the following conclusions about the solvation of calix[4]arene units:

- Solvent molecules have been frequently found in the cavity of the macrocycles indicating the formation of inclusion complexes. The distance between the center of mass of the calix[4]arene and carbon atom of the dichloromethane is in average 3.75 Å.
- The position of the solvent molecules located inside the cavity of the cone conformation follows a four-fold symmetry with a maximum oriented towards the center of the six-membered rings contained in the macrocycle.

- In general the dichloromethane contained in the inclusion complex is arranged with one of the C-H bonds pointing towards the cavity of the macrocycle, even although a tilted arrangement is also possible.
- The solvent molecules located inside the cavity are, in general, rapidly substituted by other molecules of the bulk. However, in selected cases the residence time is higher than 0.5 ns indicating that the dichloromethane has been captured. Furthermore, a significant number of dichloromethane molecules migrate from the cavity of one macrocycle to the cavity of another.

6.1.5. *References*

[1] Ikeda, A., Shinkai, S. *Chem. Rev.* **1997**, *97*, 1713.

[2] Gutsche, C.D., Bauer, L.J. *J. Am. Chem. Soc.* **1985**, *107*, 6052.

[3] Danila, C., Bohmer, V., Bolte, M. *Org. Biomol. Chem.* **2005**, *3*, 3508.

[4] Alemán, C., den Otter, W.K., Tolpekina, T.V., Briels, W.J. *J. Org. Chem.* **2004**, *69*, 951.

[5] Alemán, C., Casanovas, J. *J. Phys. Chem. A* **2005**, *109*, 8049.

[6] Alemán, C., Zanuy, D., Casanovas, J. *J. Org. Chem.* **2006**, *71*, 6952.

[7] den Otter, W.K., Briels, W.J. *J. Am. Chem. Soc.* **1998**, *120*, 13167.

[8] Tolpekina, T.V., den Otter, W.K., Briels, W.J. *J. Phys. Chem. B* **2003**, *107*, 14476.

[9] Ungaro, R., Pochini, A., Andreetti, G.D., Sangermano, V. *J. Chem. Soc., Perkin Trans 2* **1984**, 1979.

[10] Brouwer, E.B., Enright, G.D., Ripmeester, J.A. *Supramol. Chem.* **1996**, *7*, 7.

[11] León, S., Leigh, D.A., Zerbetto, F. *Chem. Eur. J.* **2002**, *8*, 4854, and references therein.

[12] Benevelli, F., Bond, A., Duer, M., Klinowski, J. *Phys. Chem. Chem. Phys.* **2000**, *2*, 3977.

[13] Arduini, A., Giorgi, G., Pochini, A., Secchi, A., Ugozzoli, F. *Tetrahedron* **2001**, *57*, 2411.

[14] Casanovas, J., Zanuy, D., Alemán, C. *Angew. Chemie Int. Ed.* **2006**, *45*, 1103.

[15] Zanuy, D., Alemán, C. *Chem. Eur. J.* **2007**, *13*, 2695.

[16] Kale, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K., Schulten K. *J. Comp. Phys.*, **1999**, *151*, 283.

[17] Baaden, M., Wipff, G., Yaftian, M.R., Burgard, M., Matt D. *J. Chem. Soc., Perkin Trans 2* **2000**, 1315.

[18] Blas, J.R., Márquez, M., Sessler, J.L., Luque, F.J., Orozco M. *J. Am. Chem. Soc.* **2002**, *124*, 12796.

[19] Wang, J., Cieplak, P., Kollman, P.A. *J. Comput. Chem.* **2000**, *21*, 1049.

[20] Jorgensen, W.L., Briggs, J.M., Contreras, M.L. *J. Phys. Chem.* **1990**, *94*, 1683.

[21] Darden, T., York, D., Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.

[22] Ryckaert, J.P., Ciccotti, G., Berendsen, H.J.C. *J. Comput. Phys.* **1977**, *23*, 327.

[23] Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., Haak, J.R. *J. Chem. Phys.* **1984**, *81*, 3684.

6.2. Sensing mechanism of Calix[4]arene-substituted polythiophene ion receptor: Effects of the selectivity on the molecular rigidity*

6.2.1. *Introduction*

Electron-rich thiophene-based oligomers and polymers combined with versatile calix[4]arenes provide novel interesting materials with important technological properties. For example, poly(calix[4]arene quaterthiophene) was found to behave as a molecular actuator,[1,2] the electronic and dynamic aspects associated to its actuation mechanism being investigated both theoretically[3-5] and experimentally.[6] On the other hand, poly(thiophene)s functionalized with calix[4]arenes (Figure 1) have been found to be highly sensitive ion receptors.[7,8] These materials, which involve four thiophene rings per repeat unit, display exceptional selectivity for the Na^+ ion, while the chemosensory response is minimum on exposure to Li^+ and K^+. The changes induced by the former ion in the properties of the material were essentially attributed to electrostatic factors, the role played by conformational perturbations being assumed to be much lesser.

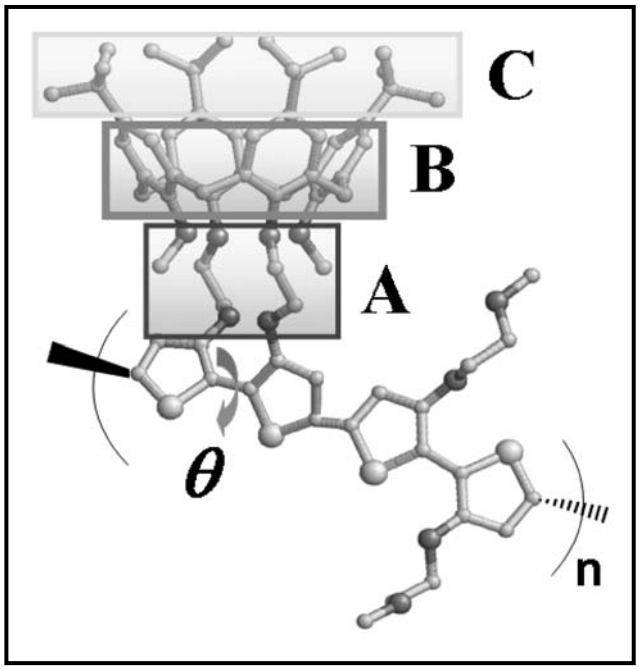


Figure 1. Chemical structural unit of the poly(thiophene) functionalized with calix[4]arene studied in this work. The three regions of the receptor that may be involved in the interaction of this system with alkali ions have been labeled as A, B and C. The dihedral angle θ refers to the relative arrangement of two thiophene rings connected not only by the conventional α,α-linkage but also by the calix[4]arene unit.

* Results presented in this chapter are published in *J. Phys. Chem. B* **2009**, *113*, 8284

Inspection to the chemical unit responsible of the sensory response of poly(thiophene)s functionalized with calix[4]arene reveals that three regions (A, B and C in Figure 1) within the cavity of the receptor may be involved in ionoresistive behavior of this system. The first is formed by the six oxygen atoms while the second and third, which are large enough to capture solvent molecules,[9-12] are defined by the four phenolic units and the *tert*-butyl groups, respectively. Although one may expect that the preferences of metallic ions by these receptor sites are directly related with the specificity of this ionophoric conjugate polymer, no microscopic information about these mechanistic aspects has been reported. This chapter provides a comprehensive computational study about the sensing mechanism of this conducting polymer, which is based on the application of different simulation techniques to obtain accurate and detailed microscopic information about its response to Na^+, Li^+ and K^+.

6.2.2. *Methods*

Quantum Mechanical Calculations

All quantum mechanical calculations were performed using the Gaussian 03 computer program.[13] The internal rotation of the model compound formed by 2,2'-bithiophene functionalized with a calix[4]arene macrocycle that connects the two thiophene rings was studied by scanning the inter-ring dihedral angle S-C-C-S (θ; see Figure 1) in steps of 30º between θ= 0º and 360º. A flexible rotor approximation was used, each point of the path being obtained from a geometry optimization of the molecule at a fixed value of θ. Furthermore, the minimum energy conformation was determined from full optimization using a gradient method. The rotational profiles and the geometry of the minimum energy conformation were calculated at both the HF/6-31G(d)[14] and the B3LYP/6-31+G(d,p)[15,17] levels.

Molecular Dynamics Simulations

Classical atomistic MD simulations were performed using the NAMD 2.6 program,[18] an integration step of 2 fs being selected. The bond lengths involving hydrogen atoms were kept at their equilibrium distances using the SHAKE algorithm.[19] Atom pair distance cut-offs were applied at 14.0 Å to compute the van der Waals interactions. Electrostatic interactions were extensively computed by means of Ewald summations. The real space

term was defined by the van der Waals cut-off, while the reciprocal space was computed by interpolation into an infinite grid of points (Particle Mesh Ewald) with maximum space grid points being 1.0 Å.[20]

A set of partial charged was derived explicitly for the investigated poly(thiophene) functionalized with calix[4]arene. For this purpose, the molecular electrostatic potential (MEP) was computed from the HF/6-31G(d) wave function of the optimized geometries. Atomic electrostatic parameters were derived by fitting the quantum mechanical MEP values into partial charges centered in the nuclei through a Levenberg-Marquardt non-linear optimization procedure. Equilibrium bond lengths and angles for MD simulations were derived from the minimum energy conformation optimized at the B3LYP/6-31+G(d,p) level. The torsional potential associated to the inter-ring dihedral angle θ was derived from the energy profiles calculated using quantum mechanical methods by applying a methodology early reported.[21,22] The rest of the force-field parameters for the polymer and the ions were extrapolated from the Amber libraries[23]. Tetrahydrofurane (ρ= 0.8892 g/cm^3) was described using a united atom model explicitly developed for this work. In this model, which is able to reproduce very satisfactorily the main experimental properties of the solvent, atomic parcial charges were parametrized *de novo* as aforementioned for the poly(thiophene) functionalized with calix[4]arene, equilibrium bond lengths and angles were derived from geometry optimizations at the HF/6-31+G(d) level, and the rest of force-field parameters extrapolated from Amber libraries.[23]

The constructed polymer chain, which was built considering four repeat units, was placed in the center of a cubic simulation box (*a*= 100.0 Å) filled with solvent molecules, which was previously equilibrated in *NPT* conditions (temperature 298 K and 1 atm of pressure). Once inside the box, those solvent molecules that overlapped any atom of the polymer were removed. This system was used for md0, while for md1-md5 an alkali ion was introduced in the cavity of two of the receptors (see text to identify the receptors as well as the chemical nature of alkali ion for each md#). The electric neutrality was reached by adding two Cl^- counterions to the simulation box.

Before any MD trajectory was run, $5 \cdot 10^3$ steps of energy minimization were performed in order to relax conformational and structural tensions. Next, different consecutive rounds of short MD runs were performed in order to equilibrate the density, temperature, and pressure. First, solvent and anions were thermally relaxed by three

consecutives runs, while the polymer and alkali cations were kept frozen: 0.5 ns of *NVT*-MD at 500 K were used to homogeneously distribute the solvent and anions in the box. Second, 0.5 ns of isothermal and 0.5 ns isobaric relaxation were run. Finally, all the atoms of the system were submitted to 0.15 ns of steady heating until the target temperature was reached (298 K), 0.25 ns of *NVT*-MD at 298 K (thermal equilibration) followed by 0.5 ns of density relaxation (*NPT*-MD). Both temperature and pressure were controlled by the weak coupling method, the Berendsen thermobarostat,[24] using a time constant for heat bath coupling and a pressure relaxation time of 1 ps. The end of the density relaxation simulation was the starting point of the molecular simulations presented in this work. All the simulations were performed at 298 K and constant pressure of 1 atm. The coordinates of all the production runs were saved every 500 steps (1 ps intervals). The *NPT*-MD production runs were 30 ns long for systems containing Na^+ and withouth ions and 10 ns long for systems with Li^+, K^+.

6.2.3. *Results and discussion*

Figure 2 shows the potential energy curve E=E(θ), where θ is the inter-ring dihedral angle that defines the relative arrangement of two thiophene rings connected not only by the conventional α,α-linkage but also by the calix[4]arene unit (Figure 1). The energy profiles, which were calculated using different quantum mechanical calculations, show that the accessible conformations are restricted to the *syn-gauche*$^+$ and *gauche-gauche*$^+$, *i.e.* the energy relative to the lowest energy minimum (θ= 53º) is lower than 3.5 kcal/mol for θ values ranging between 30º and 120º. The energy penalty associated to conformations that involve θ values out of this region is very high because the functionalization affects two consecutive thiophene rings. These results are in very good agreement with the structures experimentally determined for this specific model compound using both X-ray crystallography (θ= 68º) and the NMR spectroscopy,[7] the latter suggesting the existence of several conformers separated by small energy differences. On the other hand, the rotation of θ provokes geometric strain that is alleviated by deforming not only bond angles but also some bond lengths like C^{α}-$C^{\alpha'}$ inter-thiophene, which is essential for the π-conjugated system.

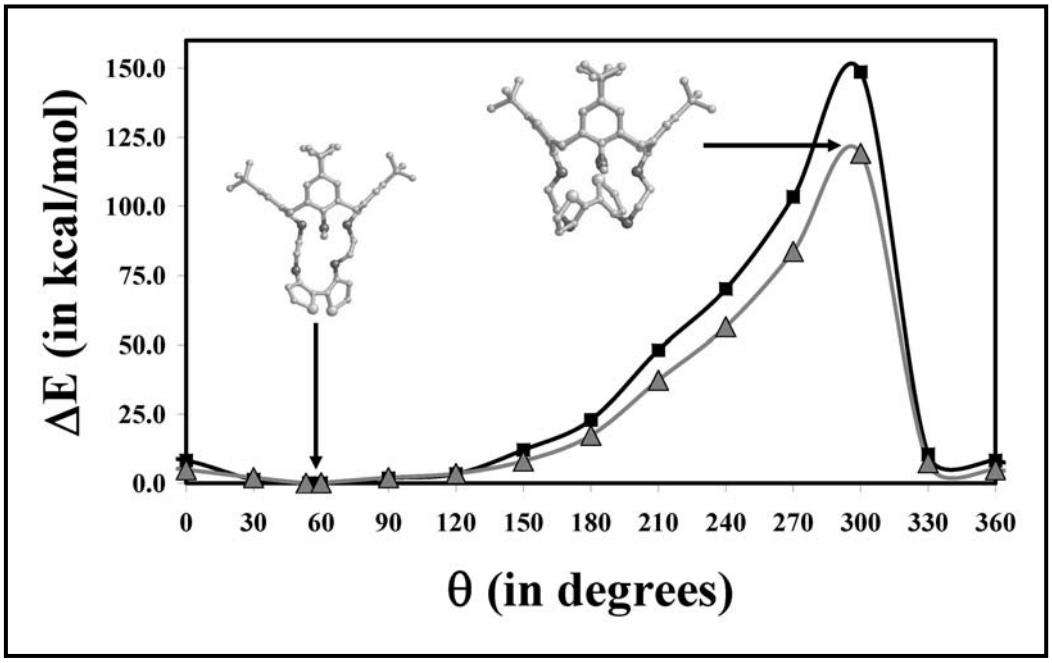


Figure 2. Potential energy curve for the internal rotation of the model molecule of poly(thiophene) functionalized calix[4]arene (n=1 in Figure 1) in the neutral state as a function of the inter-ring dihedral angle. Profiles were derived from HF/6-31G(d) (black squares) and B3LYP/6-31+G(d,p) (grey triangles) geometry optimizations. In each case energies are relative to the global minimum.

Quantum mechanical results show that the conformational flexibility associated to the interval of low-energy θ values is still high enough to play a crucial role in the response of this ion-sensitive receptor. In order to ascertain this as well as the essential features of the own sensing mechanism, the geometries and energies calculated for the profiles displayed in Figure 2 were used to derive torsional (Table 1) and electrostatic (Figure 3) force-field parameters for classical Molecular Dynamics (MD) simulations. These parameters, combined with others directly transferred from the Amber libraries[23] were used to conduct atomistic MD simulations of system constituted by a molecular chain containing four ionophores (n= 4 in Figure 1), two NaCl molecules and 6827 explicit tetrahydrofurane (THF) molecules, i.e. 89437 explicit particles.

Table 1. Torsional force-field potential to describe the rotation between two consecutive thiophene rings in the poly(thiophene) functionalized with calix[4]arene. The potential was derived using the rotational profile calculated at the B3LYP/6-31+G(d,p) level.

Torsion	Periodicity (n)	Number of paths	$V_n/2$ (kcal/mol)	γ
-C-C-	1	4	1.02	180.0
	2	4	1.08	0.0
	3	4	0.31	180.0
	6	4	0.13	0.0

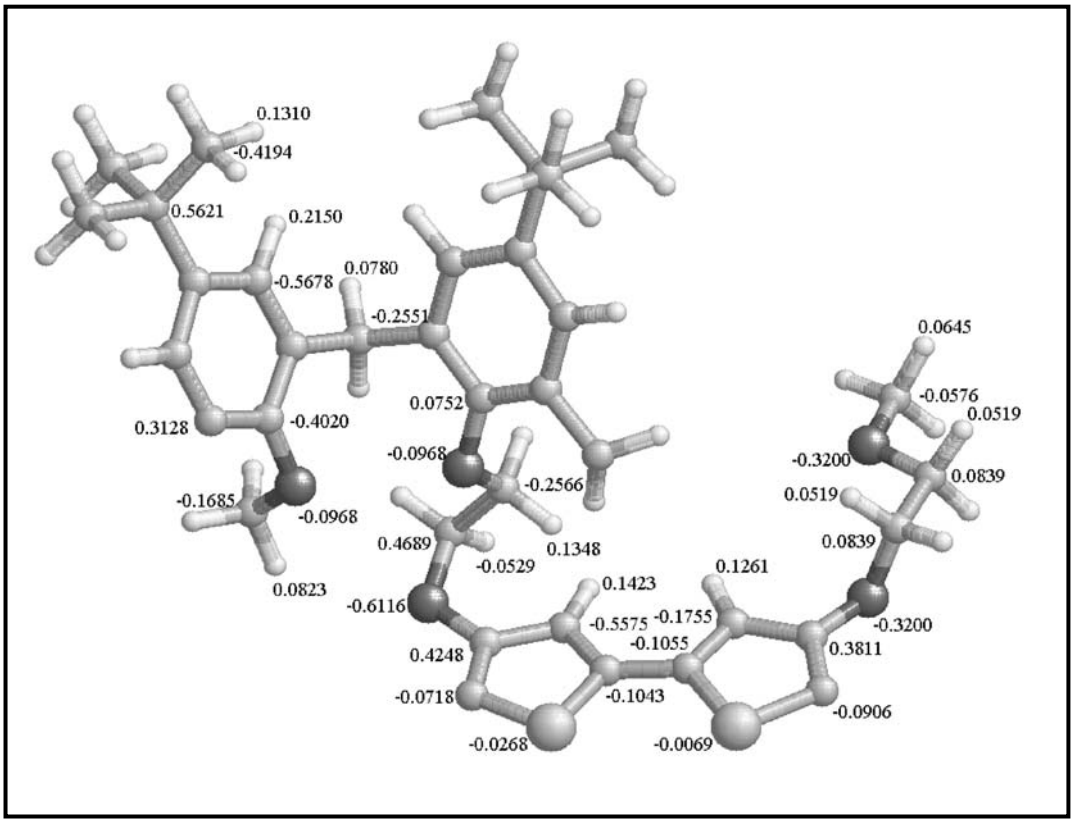


Figure 3. Electrostatic parameters used in the Molecular Dynamics simulations of the polythiophenes functionalized with calix[4]arenes. Charges were derived by fitting the quantum mechanical molecular electrostatic potential calculated at the HF/6-31G(d) level to the Coulombic molecular electrostatic potential.

Three independent *NPT* MD simulations were performed, which differ in the starting positions for the Na^+ ions: (i) the cavities of the 1st and 4th calix[4]arene units (md1); (ii) the cavities of the 1st and 3rd units (md2); and (iii) those of the 1st and 2nd units (md3). In the two latter the ions were retained within the cavities around their equilibrium positions for almost the whole 30 ns trajectory, *i.e.* the largest period with a Na^+ out of the cavity was ~0.9 ns (md2). In contrast, in md1 the Na^+ located in the 1st cavity escaped periodically for intervals that range from 2.0 to 6.7 ns. Interestingly, in all cases the ions were located in the A region of the cavity labeled, being directly coordinated by the six ether oxygens of the macrocycle. These features are reflected in Figure 4, which shows the temporal evolution of distances between each Na^+ and the center of masses of the regions A, B and C (d_A, d_B and d_C, respectively) for md1. The same representations are provided for md2 and md3 in Figures 5 and 6 respectively. As can be seen, in all cases $d_A < d_B < d_C$, which evidence the electrostatic nature of the interaction between the receptor and the ion. This is fully consistent with the voltammetric and spectroscopic responses of the Na^+ complexed material, which were explained assuming that the oxygen-rich binding site interacts electrostatically with the ion.

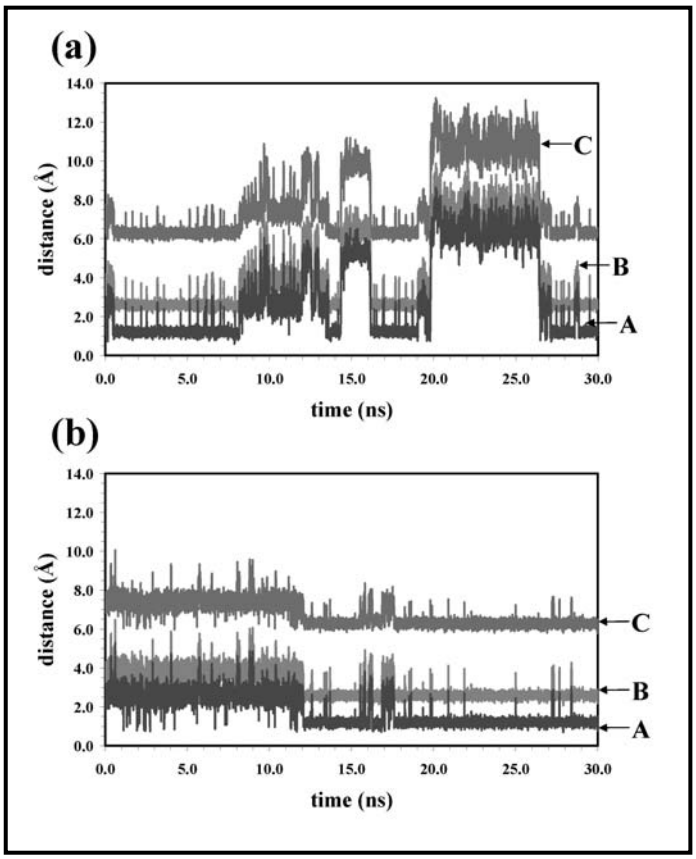


Figure 4. Temporal evolution in md1 of the distance between the Na^+ and the center of masses of regions A, B and C of the receptor (see Figure 1). In md1 ions were located filling the cavities of the first (a) and fourth (b) calix[4]arene units.

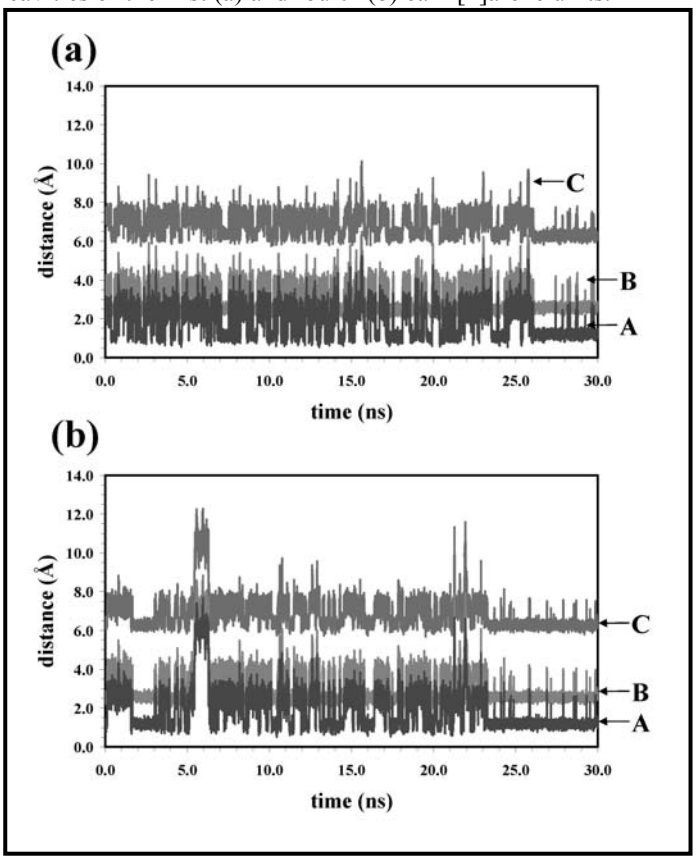


Figure 5. Temporal evolution in md2 of the distance between the Na^+ and the center of masses of regions A, B and C of the receptor (see Figure 1). In md2 ions were located filling the cavities of the first (a) and third (b) calix[4]arene units.

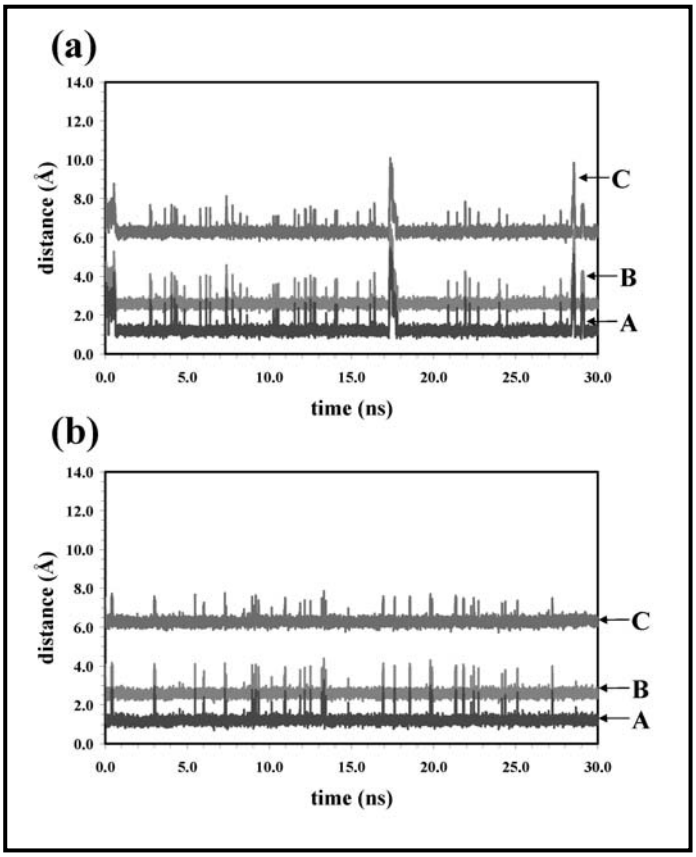


Figure 6. Temporal evolution in md3 of the distance between the Na^+ and the center of masses of regions A, B and C of the receptor (see Figure 1). In md3 ions were located filling the cavities of the first (a) and second (b) calix[4]arene units.

Interestingly, the inter-ring dihedral angles θ underwent significant fluctuations during the dynamics, even though the *syn-gauche* arrangement was retained in all cases. Thus, the average value derived from md1 is $\overline{\theta} = 55.2º \pm 6.7º$, the values ranging from 19.1º to 81.1º. Similar intervals of variation and average values were obtained for md2 and md3, *i.e.* $\overline{\theta} = 54.5º \pm 6.7º$ and $56.1º \pm 6.9º$, respectively. However, the most remarkable feature is that the intrinsically restricted conformational flexibility of the polymer backbone, which is represented in Figure 8a for md1, is reduced even more by the accommodation of Na^+ in the region A of the ionophore. Thus, the standard deviations of $\overline{\theta}$ calculated for each of the four repeat units of md1 ($54.5º \pm 6.8º$, $55.3º \pm 6.7º$, $54.8º \pm 6.7º$ and $56.3º \pm 6.6º$ for the 1st, 2nd, 3rd and 4th repeat unit, respectively) are slightly lower that those derived from a new simulation (md0) without alkali ions ($54.5º \pm 7.0º$, $54.3º \pm 7.2º$, $55.5º \pm 7.4º$ and $54.8º \pm 7.5º$ for the 1st, 2nd, 3rd and 4th repeat unit). These results provide a microscopic view of the experimental observations.[7,8] Accordingly, the enhancement of

the effective conjugation length observed on exposure to Na^+ should be attributed to an increase of the molecular rigidity.

Additional *NPT* MD simulations were conducted on systems formed by four repeat units of the calix[4]arene ion receptor (*n*= 4 in Figure 1) and two molecules of LiCl (md4) or KCl (md5), the cations being initially accommodated at the 1st and 4th ionophores in both cases. Inspection to the temporal evolution of the distances between alkali ions and the center of masses of the regions A, B and C (Figure 7) reveals significant differences with respect to the results obtained for the system with Na^+. Specifically, Li^+ and K^+ are trapped in the region of the ionophore defined by the *tert*-butyl groups rather that on the cavity defined by the ether oxygen atoms, *i.e.* $d_C < d_B < d_A$. This result is fully consistent with recent observations about the interaction of alkali ions with crown-ether macrocycles of similar size,[25] which indicated that the coordination with the oxygen atoms is preferred by Na^+. The preference of Li^+ and K^+ by region C indicates that the interaction with the receptor is not only weaker but also much less specific than that predicted for Na^+. These features are in good agreement with experimental observations, the binding affinity measured for Na^+ being approximately 40 times greater than for Li^+ and 100 times stronger than for K^+.[7]

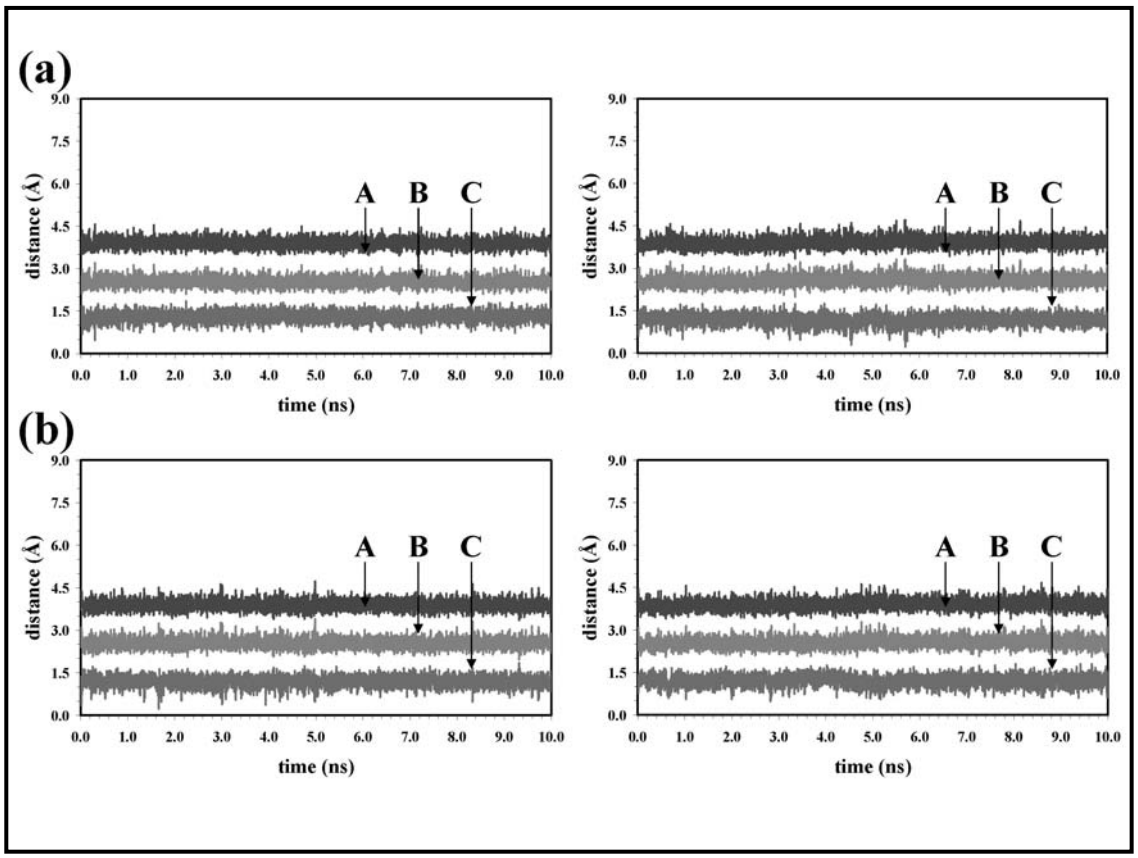


Figure 7. Temporal evolution in simulation md4 (a) and md5 (b) of the distance between the ion (Li^+ and K^+, respectively) and the center of masses of the atoms involved in the regions A, B and C of the receptor (see Scheme 1). In both md4 and md5 Li^+ and K^+ ions were located filling the cavities of the first (left) and fourth (right) calix[4]arene receptors.

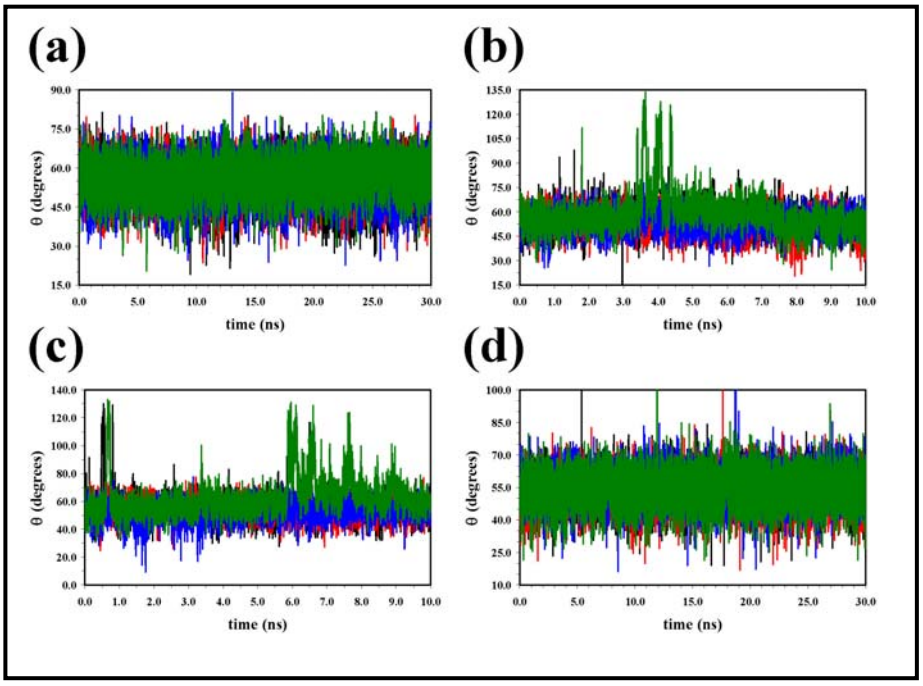


Figure 8. Temporal evolution in simulations with (a) Na^+ (md1), (b) Li^+ (md4) and (c) K^+ (md5), as well as (d) without alkali ions (md0), of the inter-ring dihedral angle θ for the four ionophores (1st black, 2nd red, 3rd blue and 4th green) contained in the polymer chain.

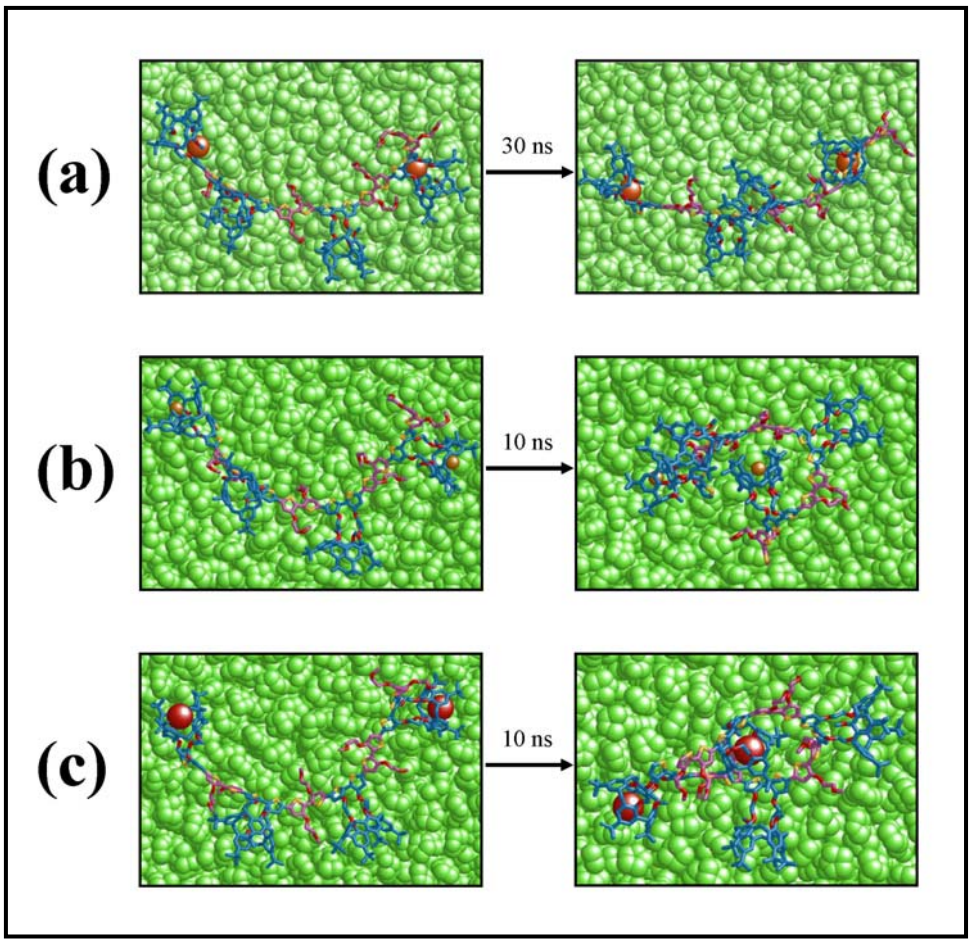


Figure 9. First and last snapshot recorded for the MD simulations with: (a) Na^+ (md1); (b) Li^+ (md4); and K^+ (md5). The shortening of the molecular chain in the systems containing Li^+ and K^+ ions (region C of the cavities in the 1st and 4th units) is clearly evidenced, while the polymer chain remains extended when the Na^+ ion are trapped in the region A of the receptors (1st and 4th units). In all cases the solvent molecules are represented by green balls.

On the other hand, inspection to the temporal evolution of the inter-ring dihedral angles θ indicates that the Li^+ ions of md4 enhance the conformational flexibility of the repeat units that accommodate the alkali ions. This is clearly evidenced not only by the standard deviation of the averages determined for the different repeat units, *i.e.* 55.1º±9.9º, 53.4º±6.2º, 54.0º±6.7º and 58.9º±12.1º for the 1st, 2nd, 3rd and 4th repeat unit, respectively, but also by the wide range of variation detected for the 1st and 4th repeat units (Figure 8b). This effect is more pronounced in md5, in which the K^+ ions produce a notable dispersion in the θ values for the 1st and 4th repeat units (Figure 8c): the average value is 55.3º±10.1º, 54.5º±6.6º, 52.5º±6.3º and 62.6º±13.5º for the 1st, 2nd, 3rd and 4th repeat unit. The conformational flexibility detected in presence of Li^+ and K^+ is also manifested by the average of the end-to-end distance, which is 31.53±9.28 (md4) and 23.46±7.94 Å (md5), respectively. These values are significantly lower than those derived from simulations with Na^+, independently of the receptor that accommodate the ion (52.92±1.85, 51.54±2.94 and 50.65±5.85 Å for md1, md2 and md3, respectively) and from the simulation performed in absence of alkali ions (53.35±6.75 Å). The remarkable influence of the Li^+ and K^+ ions on the conformation of the polymer is clearly illustrated in Figure 9, which compares the molecular arrangements recorded in the initial and final snapshots of md1, md4 and md5.

6.2.4. *Conclusions*

The overall of the calculations presented in this chapter evidences that the binding of Na^+ to the receptor of poly(thiophene) functionalized with calix[4]arene reduces the conformational flexibility of the backbone, which should be attributed to the electrostatic interactions between the ion and the oxygen atoms of the bottom part of the cavity. In contrast, Li^+ and K^+, which interact with the *tert*-butyl groups of the receptor, slightly enhance such conformational flexibility. These conformational changes are in agreement with experimental data[7,8] explaining the selectivity of this material for Na^+ over Li^+ and K^+.

6.2.5. *References*

[1] Vigalok A., Swager T. M. *Adv. Mater.* **2002**, *14,* 368.

[2] Yu H.-h., Xu B., Swager T.M. *J. Am. Chem. Soc.* **2003**, *125*, 1142.

[3] Scherlis D.A., Marzari N. *J. Phys. Chem. B* **2004**, *108*, 17791.

[4] Casanovas J., Zanuy D., Alemán C. *Angew. Chem. Int. Ed.* **2006**, *45*, 1103.

[5] Zanuy D., Alemán C. *Chem. Eur. J.* **2007**, *13*, 2695.

[6] Song C., Swager T.M. *Org. Lett.* **2008**, *10*, 3575.

[7] Marsella M.J., Newland R.J., Carrol P.J., Swager T.M. *J. Am. Chem. Soc.* **1995**, *117*, 9842.

[8] Crawford K.B., Goldfinger M.B., Swager T.M. *J. Am. Chem. Soc.* **1998**, *120*, 5187.

[9] Benevelli F., Bond A., Duer M., Klinowski J. *Phys. Chem. Chem. Phys.* **2000**, *2*, 3977.

[10] Tolpekina T.V., den Otter W.K., Briels W.J. *J. Phys. Chem. B* **2003**, *107*, 14476.

[11] Alemán C., Tolpekina T.W., den Otter W.K., Briels W.J. *J. Org. Chem.* **2004**, *69*, 951.

[12] Rodríguez-Ropero F., Zanuy D., Alemán C. *J. Comput. Chem.* **2008**, *29*, 1233.

[13] Frisch M.J., Trucks G.W., Schlegel H.B., Scuseria G.E., Robb M.A., Cheeseman J.R., Montgomery J.A., Jr., Vreven T., Kudin K.N., Burant J.C., Millam J.M., Iyengar S.S., Tomasi J., Barone V., Mennucci B., Cossi M., Scalmani G., Rega N., Petersson G.A., Nakatsuji H., Hada M., Ehara M., Toyota K., Fukuda R., Hasegawa J., Ishida M., Nakajima T., Honda Y., Kitao O., Nakai H., Klene M., Li X., Knox J.E., Hratchian H.P., Cross J.B., Bakken V., Adamo C., Jaramillo J., Gomperts R., Stratmann R.E., Yazyev O., Austin A.J., Cammi R., Pomelli C., Ochterski J.W., Ayala P.Y., Morokuma K., Voth G.A., Salvador P., Dannenberg J.J., Zakrzewski V.G., Dapprich S., Daniels A.D., Strain M.C., Farkas O., Malick D.K., Rabuck A.D., Raghavachari K., Foresman J.B., Ortiz J.V., Cui Q., Baboul A.G., Clifford S., Cioslowski J., Stefanov B.B., Liu G., Liashenko A., Piskorz P., Komaromi I., Martin R.L., Fox D.J., Keith T., Al-Laham M.A., Peng C.Y., Nanayakkara A., Challacombe M., Gill P.M.W., Johnson B., Chen W., Wong M.W., Gonzalez C., Pople J.A. *Gaussian 03*, revision C.02, Gaussian, Inc.: Wallingford, CT, **2004**.

[14] Hariharan P.C., Pople J.A. *Theor. Chim. Acta* **1973**, *23*, 213.

[15] Becke A.D. *J. Chem. Phys.* **1993**, *98*, 1372.

[16] Lee C., Yang W., Parr R.G. *Phys. Rev. B* **1993**, *37*, 785.

[17] Frich M.J., Pople J.A., Krishnam R., Binkley J.S. *J. Chem. Phys.* **1984**, *80*, 3264.

[18] Phillips J.C., Braun R., Wang W., Gumbart J., Tajkhorshid E., Villa E., Chipot C., Skeel R.D., Kale L., Schulten K. *J.Comput. Chem.* **2005**, *26*, 1781.

[19] Ryckaert J.P., Ciccotti G., Berendsen H.J.C. *J. Comput. Phys.* **1977**, *23*, 327.

[20] Darden T., York D., Pedersen L. *J. Chem. Phys.* **1993**, *98*, 10089.

[21] Alemán C., Canela E.I., Franco R., Orozco M. *J. Comput. Chem.* **1991**, *12*, 664.

[22] Alemán C., Puiggalí J. *J. Org. Chem.* **1995**, *60*, 910.

[23] Cornell W.D., Cieplak P., Bayly C.I., Gould I.R., Merz Jr. K.M., Ferguson D.M., Spellmeyer D.C., Fox T., Caldwell J.W., Kollman P.A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

[24] Berendsen H.J.C., Postma J.P.M., van Gunsteren W.F., DiNola A., Haak J.R. *J. Chem. Phys.* **1984**, *81*, 3684.

[25] Casanovas, J., Preat, J., Zanuy, D., Alemán, C. *Chem. Eur. J.* **2009**, *15*, 4676.

6.3. Controlled isomerization of a light-driven molecular motor: A theoretical study*

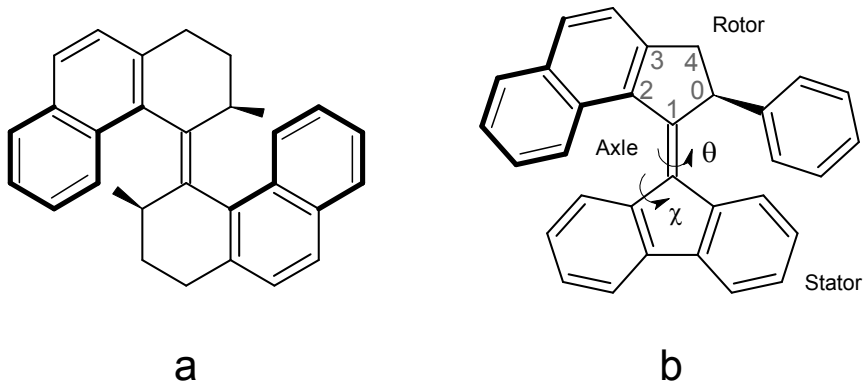
6.3.1. *Introduction*

Molecular machines are widespread in biological systems where a large collection of dedicated protein machines are performing complicated tasks such as cell division or intracellular transport. This kind of machine has inspired both (i) the use of these biological machines in artificial environments outside the cell; and (ii) the design of new synthetic motors at the same scale.[1] The building of artificial molecular motors and machines is essential for the development of bottom-up strategies aimed to construct nanoscale devices, a big challenge at the intersection of chemistry, physics, and biology knowledge. Over the past decade, the construction of new artificial molecular motors as well as the demonstration that energy consumption can produce mechanical work by inducing controlled and unidirectional motions that allow objects to be moved at large scale has been reported.[2,3]

One particular class of artificial molecular motors involves those that are light-driven. There are several advantages as compared to chemical or electrochemical means of stimulation: easy control of energy supply, selectivity, and precise energy application at nanometer resolution.[4] The first reported lightdriven molecular motor was based on a symmetric biphenanthrylidene (Scheme 1a).[5] In this motor, the unidirectional rotary motion around the central double bond, which separates the lower and upper halves of the molecule (hereafter denoted the stator and the rotor, respectively), is reached by two olefin photochemical isomerizations, each one followed by a thermal conversion. A dramatic increase of the speed of rotation was achieved when the six-membered ring of the rotor is substituted by a five-membered ring. This structural modification results in a reduction of the free energy of activation for the two irreversible thermal steps in the rotary process, which is directly correlated with the speed of rotation.6 To accelerate further the rotational speed, a second generation of light-driven molecular motors was recently introduced by Feringa and co-workers.[6-8] In these new motors, the stator was substituted by a fluorine moiety.[6-8] Interestingly, one of the molecular motors belonging to the second generation (Scheme 1b) was able to move large objects, such as a micrometer-sized glass rod, when it was used dispersed in a liquid crystalline film.[9,10]

* Results presented in this chapter are published in *J. Phys. Chem. C* **2009**, *113*, 3574

Specifically, it was found that ultraviolet irradiation at 365 nm induces a photochemical isomerization around the central double bond (motor axle) with inversion of molecular helicity (PfM). Subsequently, a thermal step with another helix-inversion (MfP) at 20 °C completes a 180° rotation. Accordingly, successive steps of irradiation and thermal relaxation induce the movement of the rotor with respect to the stator. Furthermore, in a very recent study, another fluorenyl motor was used to control the preferred helical twist sense of a poly(n-hexyl isocyanate), which was covalently attached to the polymer's terminus.[11]



Scheme 1

The present chapter, enclosed in the frame of a project devoted to control the conformational transitions between the folded states of individual peptide molecules by coupling a light-driven molecular motor to the main chain and/or side chain of selected residues, aims the characterization of the rotational profile of 9-(2,3-dihydro-2-phenyl-1H-benz[e]inden-1-ylidene)-9H-fluorene (Scheme 1b), abbreviated NA9PH, using theoretical methods. Thus, both the demonstrated ability to move big objects and the simplicity to induce controlled rotation under ultraviolate irradiation reinforce this molecule as a potential inductor of such conformational transitions. Within this context, Assfeld and co-workers[42] used the QM/MM-SCRF scheme to investigate the conformational changes in a peptide induced by a dye attached as a side chain. The rotational profile of NA9PH has been investigated in detail using two different but complementary approaches. First, quantum mechanical calculations in the gas phase and implicit chloroform solvent have been performed to determine the free energy profiles for the complete rotation cycle. After this, semiempirical QM/MM molecular dynamics

simulations[12] have been carried out using the umbrella sampling methodology[13] and considering explicit chloroform molecules to describe the solvent. Results have been compared to available experimental data reported in previous investigations.[7]

6.3.2. *Methods*

Gas-phase calculations

Starting molecular geometries were obtained from experimental crystallographic data[7] and full geometry optimizations in the gas phase, which were performed at different quantum mechanical levels, for the singlet and triplet states, hereafter denoted NA9PH(S) and NA9PH(T), respectively. Specifically, geometries were initially optimized using the PM3[14,15] semiempirical method and subsequently reoptimized using the density functional theory (DFT) method. All DFT calculations were carried out using the Becke's three parameter hybrid functional (B3)[16] with the Lee, Yang, and Parr (LYP)[17] expression for the nonlocal correlation, combined with the 6-31G(d) and 6-31+G(d,p) basis sets.[18] The restricted and unrestricted formalism of the semiempirical (PM3 and UPM3) and DFT methods (B3LYP and UB3LYP) was considered for the singlet and triple states, respectively. Frequency calculations of all of the stationary points were performed to calculate the zero-point vibrational energies (ZPVE) and both the thermal and the entropic corrections. All computed free energies were obtained using the unscaled harmonic vibrational frequencies.

After geometry optimization, the internal rotation of NA9PH(S) and NA9PH(T) was investigated by scanning the rotational dihedral angle θ (Scheme 1b) from 0° to 360° in steps of 15° using the flexible rotor approximation. Thus, each point of the path was derived from a constrained geometry optimization, in which the rotational dihedral angle θ was fixed. A better description of the rotational energy profiles was achieved by performing single point energy calculations at the second-order Møller-Plesset perturbation level[19] with the 6-31G(d) basis set, that is, MP2/6-31G(d) and UMP2/6-31G(d) levels for single and triplet states, respectively. For the stationary points, free energies were determined by combining the electronic energies computed at the UMP2/6-31G(d) level with the thermodynamic corrections derived from the UB3LYP/6-31G(d) frequencies.

All quantum mechanical calculations presented in this chapter were performed using the Gaussian 03 computer program.[20]

Solvation models

The influence of the chloroform solvent on the energy profiles associated with the rotation of the dihedral angle θ has been investigated using two different models. The first one is a self-consistent reaction-field (SCRF) model, which describes the solute at the quantum mechanical level, while the solvent is represented as a dielectric continuum. Specifically, the polarizable continuum model (PCM) developed by Tomasi and co-workers was chosen to describe the solvent.[21,22] The PCM represents the polarization of the liquid by a charge density appearing on the surface of the cavity created in the solvent, that is, the solute/solvent interface. PCM calculations were performed in the framework of the UB3LYP/6-31G(d) level using the standard protocol, considering the dielectric constant of chloroform (ε = 4.9) and the geometries optimized in the gas phase. Thus, previous studies indicated that solute geometry relaxations in solution and single point calculations on the optimized geometries in the gas phase give very similar free energies of solvation.[23-25] In the second model, which was applied in mixed semiempirical quantum mechanics/molecular mechanics (QM/MM) calculations, the solvent was described with explicit chloroform molecules using the model reported by Kollman and co-workers.[2]

QM/MM calculations

QM/MM calculations were performed using the umbrella sampling[27] procedure. For this purpose, a starting configuration of the system under study was prepared for each value of the dihedral angle θ scanned in the gas phase. Specifically, the NA9PH was constructed using the LEaP program from the Amber 9 suite,[28] and solvated in a box with enough chloroform molecules to allow ~30 Å of buffer region around the solute. The number of chloroform solvent molecules contained in the simulation box ranged between 2530 and 2660 depending on the initial NA9PH geometry. The general Amber force field parameters[29] were used for all classical calculations, the RESP charges (restrained electrostatic potential)[30,31] being explicitly derived for value of the dihedral angle θ. The electrostatic potential obtained at the Hartree-Fock level of theory with the 6-31G(d) basis set[18] was used to fit the RESP atomic partial charges.

Initially, each starting configuration was relaxed for 100 ps (2 fs time step). After this, the system was heated from 0 to 300 K during a 200 ps of simulation using the Langevin thermostat with a collision frequency of 1 ps-1, and then brought to 1 atm

pressure for another 500 ps using a pressure relaxation time of 1 ps. Next, the NA9PH was transformed into a quantum molecule (PDDG/PM3[32] and DFTB/SCC-DFTB[33,34] semiempirical quantum mechanical description) using the QM/MM module available in Amber 9. The system was then equilibrated by running another 50 ps with a NPT ensemble simulation. In all of these simulations, the geometry of the NA9PH molecule was kept restrained using a force constant of 100 kcal/mol. The SHAKE algorithm[35] was used to constrain the covalent bonds involving hydrogen atoms in all classical and QM/MM simulations. The equilibrated systems were used as starting points for the determination of the potential of mean force (PMF).

The PMF provided by the umbrella sampling[27] procedure was calculated using a total of 26 windows. A harmonic biasing potential of 30 kcal/mol was applied as umbrella potential with a minimum located at each scanned value of the dihedral angle θ. QM/MM production runs of 0.5 ns using both the PDDG-PM3[32] and the SCC-DFTB[33,34] Hamiltonians for the solute were performed to extract 25 000 bins and used to acquire statistics at each window. The total QM/MM simulation time used to determine the whole PMF was 13 ns. The bins obtained from the 26 windows were unbiased and combined using the weighted histogram analysis method (WHAM)[36-38] to generate the PMF.

6.3.3. *Results and discussion*

Equilibrium geometries in the gas phase

Complete geometry optimizations of NA9PH(S) and NA9PH(T) were carried out in the gas phase, with the results summarized in Table 1.

Table 1. Relevant Geometrical Parameters (See Text) for the Most Stable Structures of NA9PH(S) and NA9PH(T) Obtained in the Gas Phase Using Different Theoretical Procedures

	NA9PH(S)				NA9PH(T)		
	PM3	**B3LYP 6-31/G(d)**	**B3LYP 6-31+G(d)**	**exp.[a]**	**PM3**	**UB3LYP 6-31/G(d)**	**UB3LYP 6-31+G(d)**
θ [b]	7.9	16.1	16.3	13.4	90.0	92.5	91.5
d [c]	1.348	1.369	1.370	1.357	1.444	1.466	1.466
χ [b]	10.2	3.5	3.6	-1.3	0.4	-3.7	-3.6
P [b,d]	105.4(T_1^0)	94.0(E^0)	94.3(E^0)	97.8(E^0)	144.0(T_1^2)	73.5(T_4^0)	74.2(T_4^0)
τ_m [b]	28.5	33.3	33.2	31.7	0.2	7.5	7.5
ΔG [e]	-	0.0	0.0	-	-[f]	23.5	22.8

[a] From reference 7. [b] In degrees. [c] Bond length (in Å) associated to the molecular axle. [d] The conformation of the five-membered ring belonging to the rotor (Scheme 1b) is indicated in parenthesis. Both the phase angle (***P***) and the phase angle (τ_m) have been determined using the model reported in reference 39. [e] In kcal·mol^{-1}. [f] The PM3 semiempirical method describes poorly the relative stability between the surfaces calculated for the two electronic states (see text)

Geometry optimization of NA9PH(S) at the B3LYP/6-31G(d) and B3LYP/6-31+G(d) levels yields results close to those determined by X-ray crystallography,[7] the latter being included in Table 1. The optimized geometry values do not show a significant difference when a diffuse function is included in the basis set, especially on the scanned dihedral angle. Thus, the values found for the dihedral angle θ (Scheme 1b) differ by only 2.7°, and the phenyl substituent at the stereogenic center was rightly predicted to prefer a pseudoaxial arrangement. Specifically, the isomer with the phenyl substituent and the naphthalene moiety located on the same side with respect to the plane defined by the fluorene ring is the most stable. Hereafter, this isomer will be denoted "Ax". Furthermore, other important geometrical parameters such as the bond distance associated with the axle that separates the rotor and the stator (see Scheme 1b), the amplitude (τ_m) and phase angle (P) that define the pseudorotational preferences of the five-membered ring in the rotor,[39] and the dihedral angle defined by the molecular axle and the fluorene (χ, in Scheme 1b) also show good agreement with the values determined experimentally.[7]

On the other hand, the dihedral angles θ obtained for NA9PH(T) at the UB3LYP/6-31G(d) and B3LYP/6-31+G(d) levels are 92.5° and 91.5°, respectively. Not much difference is observed with the inclusion of a diffuse function. In this structure, the phenyl substituent at the stereogenic center prefers pseudoequatorial arrangement, with the phenyl and naphthalene moieties lying on opposite sides of the planar fluorine. The bond associated with the molecular axle of this configuration, hereafter identified as "Eq", is 0.097 Å larger than that of the Ax in NA9PH(S). This increment of the bond length will allow the upper naphthalene moiety of NA9PH(T) to slip over the fluorine more easily during the NA9PH half-rotation. It should be taken into account that the bond length reported in Table 1 is larger than that previously calculated by Vicario et al.[7] for a less stable NA9PH(T) structure ($d_{C\text{-}C}$ = 1.375 Å). On the other hand, the pseudorotational parameters reported in Table 1 for NA9PH(T) indicate that the five-membered ring of the rotor adopts a twisted conformation (T_4^0) with a very low amplitude, τ_m = 7.5°. Finally, it is worth noting that DFT calculations predict that the free energy difference between the singlet and triplet states is 23.5 kcal mol^{-1}, this difference being reduced to 19.9 kcal mol^{-1} by applying single point calculations at the UMP2/6-31G(d) level.

Structural parameters obtained using the PM3 semiempirical method are, in general, in good agreement with those provided using the B3LYP functional (Table 1). Thus, the only remarkable discrepancy appeared on the conformation of the five membered ring contained in the rotor. For the NA9PH(S), this ring was predicted to adopt the twisted T_1^0 and envelope E^0 conformations (see Scheme 1b for atom numbering) at the PM3 and B3LYP/6-31G(d) levels, respectively, the latter puckering being that observed experimentally.[7] Similarly, the two methods differ in the twisted conformation predicted for the NA9PH(T), that is, T_1^2 and T_4^0 at the UPM3 and UB3LYP/6-31G(d) levels, respectively. Furthermore, this semiempirical procedure fails in the relative stability between the triplet and singlet states. Thus, UPM3 calculations predict similar energies for the surfaces associated with these two electronic states, while the *ab initio* and DFT methods used in this work indicate that the potential energy surface of the triplet state is significantly destabilized with respect to that of the singlet state. Despite these limitations, the overall results allow us to conclude that the PM3 Hamiltonian provides a satisfactory qualitative description of the structural properties of NA9PH using very reduced computational resources. This feature is essential for QM/MM calculations, which are very demanding from a computational point of view.

Rotational profiles in the gas phase

The rotational profiles obtained in the gas-phase by scanning the rotational dihedral angle θ at UB3LYP/6-31G(d) level are displayed in Figure 1, while the relative energies and free energies of the minima (ΔE and ΔG, respectively) and the barriers ($\Delta E^{\ddagger}$ and $\Delta G^{\ddagger}$, respectively) are listed in Table 2. The profiles calculated for both NA9PH(S) and NA9PH(T) present two equivalent minima that follow a periodicity of 180 degrees. In the profile calculated for NA9PH(S), the barrier occurs at θ= 100.3º with a relative free energy of $\Delta G^{\ddagger}_S$= 31.4 kcal•mol-1. On the other hand, UB3LYP/6-31G(d) calculations on NA9PH(T) provides the barrier at θ= 182.1º with $\Delta G^{\ddagger}_T$= 17.5 kcal mol^{-1}. The latter value shows a difference of ~5 kcal mol^{-1} with respect to estimations previously reported in the literature,[7] which were obtained using a similar theoretical methods.

Table 2. Energy and free energy barrier ($\Delta E^{\ddagger}$ and $\Delta G^{\ddagger}$, respectively; in kcal·mol^{-1}) for the rotation of the dihedral angle θ[a]

	$\Delta E^{\ddagger}_S$	$\Delta E^{\ddagger}_T$	$\Delta G^{\ddagger}_S$	$\Delta G^{\ddagger}_T$	ΔG_{ST}
Gas-phase					
(U)PM3	28.5	13.0	25.7	15.9	-[b]
(U)B3LYP/6-31G(d)	32.6	14.8	31.4	17.0	23.5
(U)MP2/6-31G(d)	31.4	10.6	33.0[b]	13.6[b]	19.9[c]
Chloroform solution					
PCM-UB3LYP/6-31G(d)	31.2	14.2	29.5	15.9	23.4
QM/MM PDDG-PM3	-	-	22.0	-	-
QM/MM SCC-DFTB	-	-	21.3	-	-
$\Delta G^{\ddagger}$ exp. [d]				21.0	

[a] Calculations were performed for the singlet and triplet electronic states of NA9PH. The free energy difference between the minimum energy conformations of the two states (ΔG_{ST}; in kcal mol-1) is also displayed. Both gas phase and chloroform solution environments are considered. [b] The PM3 semiempirical method describes poorly the relative stability between the surfaces calculated for the two electronic states (see text). [c] Obtained by adding the (U)B3LYP/6-31G(d) thermodynamic corrections to the electronic energies calculated at the (U)MP2/6-31G(d) level. [d] From ref 7.

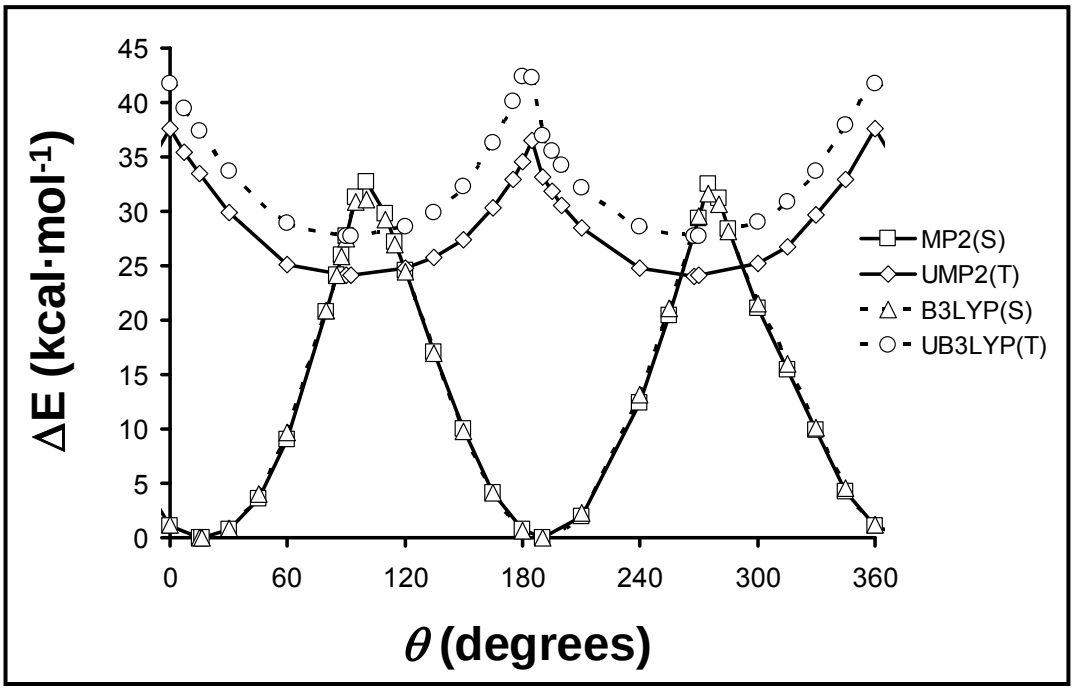


Figure 1. Rotational profiles in the gas-phase calculated for the singlet (S) and triplet (T) states of NA9PH. Calculations were performed at the (U)B3LYP/6-31G(d) and (U)MP2/6-31G(d) levels (see text).

To provide a more accurate estimation of the rotational barrier, single point calculations were developed at the UMP2/6-31G(d) level using the UB3LYP/6-31G(d) geometries, with the relative free energies between stationary points being obtained by adding to the electronic energy the thermodynamic corrections derived at the latter level of theory. The resulting profiles, which are included in Figure 1, are in good agreement with the DFT ones, especially those obtained for the single state, that is, $\Delta G^{\ddagger}_{S}$= 33.0 and 31.4 kcal mol^{-1} at the UMP2/6-31G(d) and UB3LYP/6-31G(d) levels (Table 2), respectively. For the triplet state, UMP2/6-31G(d) calculations provide a stabilization of ~3.5 kcal mol^{-1}. This is reflected not only by the barrier $\Delta G^{\ddagger}_{T}$, which is 3.4 kcal mol^{-1} smaller than that derived from the UB3LYP/6-31G(d) calculations (Table 2), but also by the difference between the energies of the minima obtained for the singlet and triplet states (ΔG_{ST} in Table 2).

On the other hand, the rotational profiles derived from UPM3 semiempirical calculations were very similar to those displayed in Figure 1, the free energy of the barrier being 25.7 and 15.9 kcal mol^{-1} for NA9PH(S) and NA9PH(T), respectively. However, this semiempirical procedure fails in the relative stability between the triplet and singlet states. Despite the limitation detected for this semiempirical method, the different theoretical estimations reported in Table 2 are in good agreement with the experimental value of the free energy barrier of activation for the irreversible thermal helix inversion, at the second step of NA9PH half-rotation movement ($\Delta^{\ddagger}G^{o}$ = 21.0 kcal mol^{-1}).[7]

Figure 2 shows the evolution of the dihedral angle χ against the scanned dihedral angle θ (Scheme 1b) for both NA9PH(S) and NA9PH(T). These profiles provide an approximated template of the flexibility associated with the coupling between the double bond that connects the rotor and the stator and the plane defined by fluorene moiety. Interestingly, there is a bending deformation in the fluorene as well as a deviation from the planarity of the two carbons involved in the double bond with respect to the fluorene plane, these distortions being larger when the rotated dihedral angle is closer to the barriers, that is, the sign of χ undergoes a change during the semirotation. Moreover, the two electronic states show a similar behavior, and the largest values reached by χ are of ~45° and ~30° for NA9PH(S) and NA9PH(T), respectively. The distortion is smaller for NA9PH(T) than for NA9PH(S) because, as

was discussed in the previous section, the bond length that connects the rotor and the stator is larger for the former than for the latter.

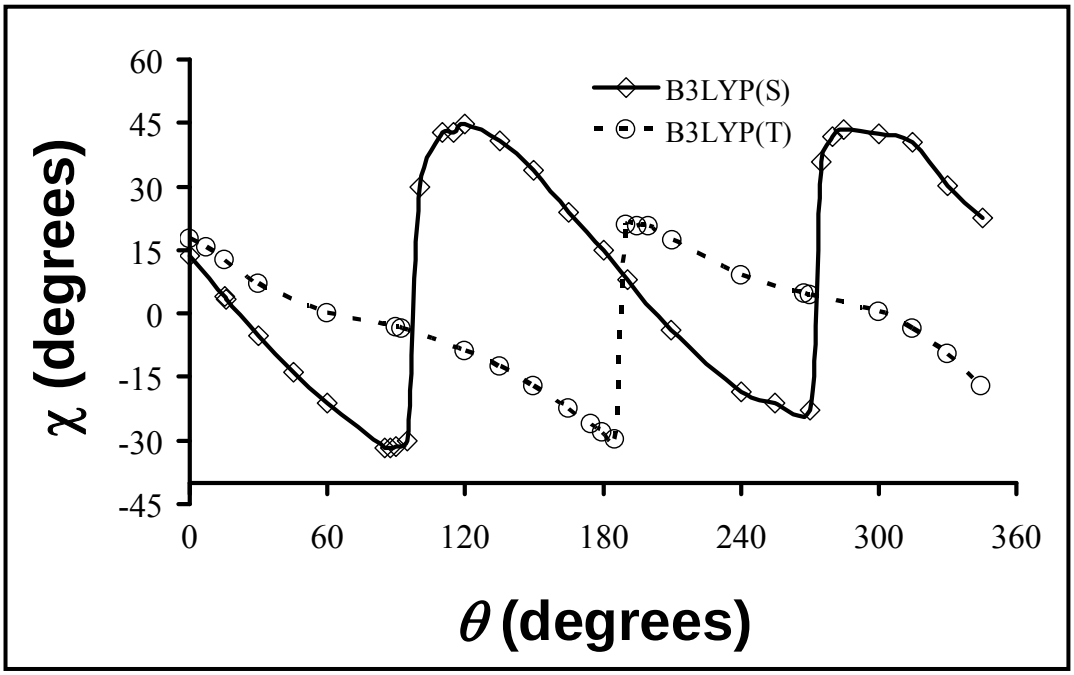


Figure 2. Evolution of the dihedral angle χ as a function of the rotational dihedral angle θ, in the gas phase and using the (U)B3LYP/6-31G(d) level of calculation.

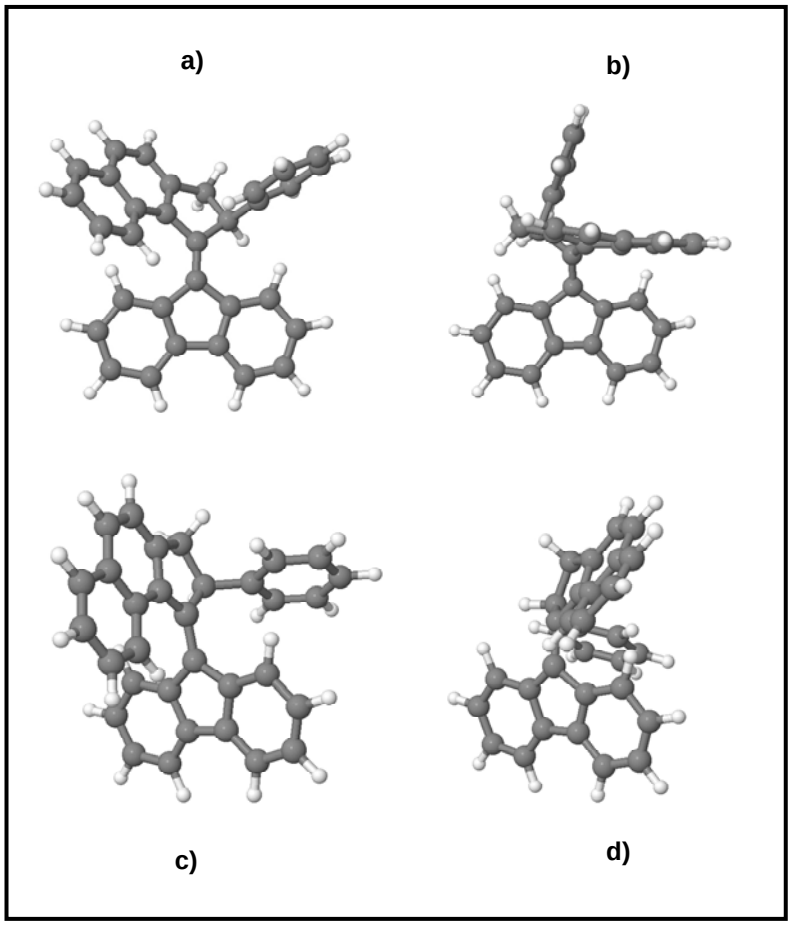


Figure 3. Molecular structures of the minima and barriers of the rotational profiles calculated in the gas phase at the (U)B3LYP/6-31G(d) level for NA9PH: (a) minimum of NA9PH(S); (b) barrier of NA9PH(S); (c) minimum of NA9PH (T); and (d) barrier of NA9PH(T).

Figure 3a and b displays the molecular geometries of the minimum and the maximum found in the rotational profile calculated for NA9PH(S) rotational profile. A close inspection of these geometries suggests a $sp^2 \rightarrow sp^3$ hybridization transition at the atoms that connect the rotor and the stator, which allows one to minimize the steric repulsion associated with the evolution from the minimum to the maximum. The deviation from planarity calculated for the C^1 atom of the rotor (Scheme 1b for numbering) in the lowest energy structure is 7.3°, which is in excellent agreement with the experimental value (5.25°).[7] The biggest deviation, 24°, is located close to the maximum of the gas-phase energy profile, where the sign of the dihedral angle shows that the planarity is inverted, indicating an intermediate conformation. The flexibility in the five-membered ring of the rotor is also detected through the puckering variables used to define the conformation of the cycle. Specifically, a conformational transition $E^0 \rightarrow T_1^0 \rightarrow E^0$ was predicted for this ring. On the other hand, the structures associated with the minimum and maximum of the rotational profile calculated for NA9PH(T) are depicted in Figure 3c and d, respectively. The dihedral angle χ obtained for the triplet state should be mainly attributed to the distortion of the bond that connects the rotor and the stator, the maximum deviation from planarity determined for the C^1 atom being 23°.

Equilibrium geometries in chloroform solution

The influence of the solvent on the geometry of the minimum and the barrier was considered by running 2.5 ns of NPT MD trajectories in explicit chloroform solution under the QM/MM scheme, where NA9PH and chloroform molecules were quantum and classical, respectively. The starting geometries for NA9PH were those derived for the minimum and the barrier from the corresponding rotational profiles (see next section). Simulations of the barrier and the minimum were developed with and without restraints at the dihedral angle θ. As was stated in the Computational Methods, QM/MM simulations were run with AMBER 9.0. Unfortunately, this package does not handle unpaired electrons correctly because it does not include the code for UHF calculations; that is, it calculates the energy of the individual electrons but not the difference between having two unpaired versus one paired and an empty orbital.[40] Accordingly, all QM/MM simulations were run on NA9PH(S), the influence of the explicit solvent molecules on the triplet state not being considered. Time-averages of the relevant

geometrical parameters obtained for the more representative structures of NA9PH(S) using both PDDG-PM3 and SCC-DFTB Hamiltonians are reported in Table 3.

Table 3. Time-averaged geometries derived from QM/MM calculations in chloroform solution for the minimum and the barrier of NA9PH(S)[a]

	Minimum		Barrier	
	PDDG-PM3	**SCC-DFTB**	**PDDG-PM3**	**SCC-DFTB**
θ [b]	13.3 ± 0.1	14.3 ± 0.1	98.4 ± 0.0	100.1 ± 0.1
d [c]	1.351 ± 0.000	1.383 ± 0.000	1.371 ± 0.000	1.415 ± 0.000
χ [b]	5.0 ± 0.2	5.2 ± 0.2	17.8 ± 0.1	16.2 ± 0.1
P [b,d]	104.3 ± 0.2 (T_1^0)	97.2 ± 0.2 (E^0)	-54.0 ± 1.1(E^1)	-75.1 ± 0.8 (T_1^0)
τ_m [b]	29.3 ± 0.1	26.2 ± 0.1	15.0 ± 0.1	18.1 ± 0.1

[a] The standard error of the mean is also displayed. [b] In degrees. [c] Bond length (in Å) associated with the molecular axle. [d] The conformation of the five-membered ring belonging to the rotor (Scheme 1b) is indicated in parentheses. Both the phase angle (P) and the amplitude (τ_m) have been determined using the model reported in ref 39.

As can be seen, the geometry of the minimum in solution is very similar to that obtained in the gas phase (Table 1). Thus, the dihedral angles θ and χ predicted by the PDDG-PM3 method differ by only 5.4° and 5.2°, respectively, with respect to those obtained using the PM3 in the gas phase, while the differences between SCC-DFTB and B3LYP methods are 1.8° and 1.7°. On the other hand, the puckering predicted for the minimum in the gas phase was essentially identical to that obtained in chloroform solution; that is, PDDG-PM3 and PM3 predict the T_1^0 conformation, whereas SCC-DFTB and B3LYP prefer the E^0 one.

Table 3 also shows the more relevant structural parameters obtained for the barrier in chloroform solution. In the gas phase, this structure was predicted to be at θ = 95° by both the B3LYP and the PM3 methods, in solution the deviation with respect to this value ranging between 3° and 5°. Moreover, comparison between the geometries obtained for the minimum and the barrier in chloroform solution reflects distortions similar to those discussed above for the gas phase. Specifically, the length of the bond that connects the rotor and the stator and the dihedral angle χ increase, while the

variation of the pseudorotational angle *P* evidences a conformational change of the five membered ring located at the rotor.

Rotational profiles in chloroform solution

The effects of the solvent on the rotational profile have been taken into account using implicit (singlet and triplet states) and explicit (singlet state) solvation models. The PMFs calculated using the PCM solvation model within the UB3LYP/6-31G(d) framework are displayed in Figure 4, which includes the profiles obtained in the gas phase at the same computational level for comparison, while Figure 5 details the profiles derived from QM/MM trajectories. The free energy barriers in solution obtained using PCM and QM/MM calculations are listed in Table 2.

As can be seen in Figure 4, implicit solvation effects do not introduce significant changes in the rotational profile of NA9PH(S), the maxima and minima predicted in chloroform solution being very close to those in the gas phase. However, a stabilization of 1.9 and 1.1 kcal mol-1 in the rotation free energy barrier of NA9PH(S) and NA9PH(T), respectively, was detected. Thus, the barriers predicted by the implicit salvation model are $\Delta^{\ddagger}G_S$ = 29.5 kcal mol^{-1} and $\Delta^{\ddagger}G_T$ = 15.9 kcal mol^{-1}, respectively. On the other hand, inspection of the QM/MM results obtained using the PDDG-PM3 and SCC-DFTB Hamiltonians reveals a free energy barrier of 22.0 and 21.3 kcal mol^{-1}, respectively, for NA9PH(S).

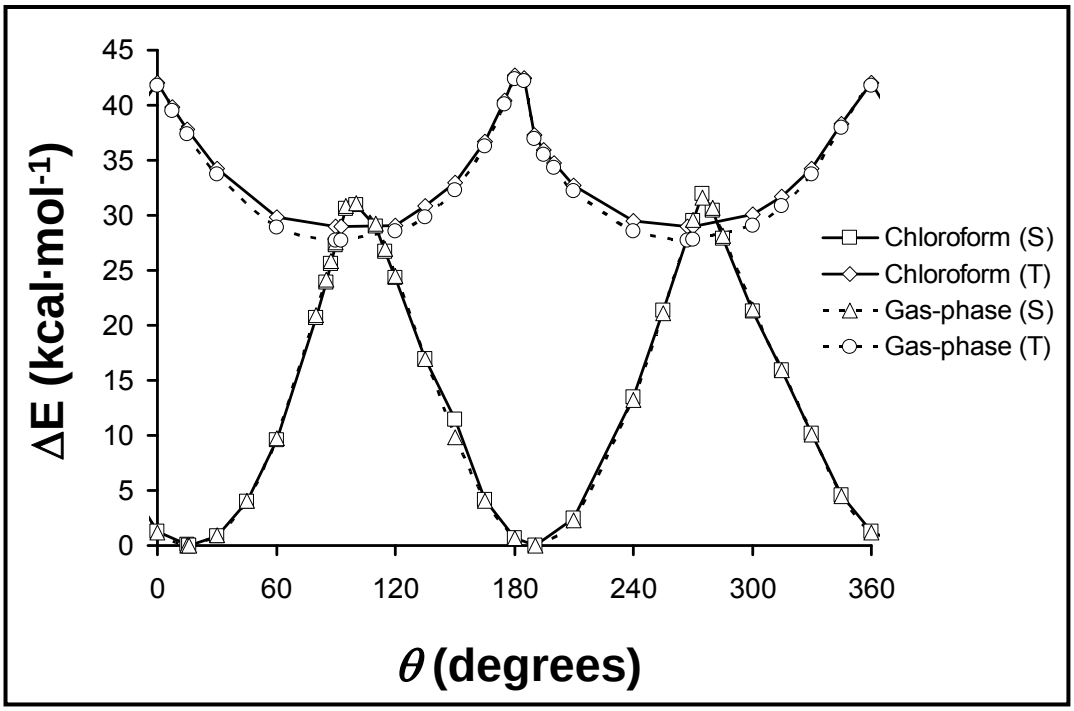


Figure 4. Rotational profiles in chloroform solution calculated for the singlet (S) and triplet (T) states of NA9PH using the PCM method in the framework of the (U)B3LYP/6-31G(d) level. Rotational profiles obtained in the gas phase using the same theoretical level have been included for comparison.

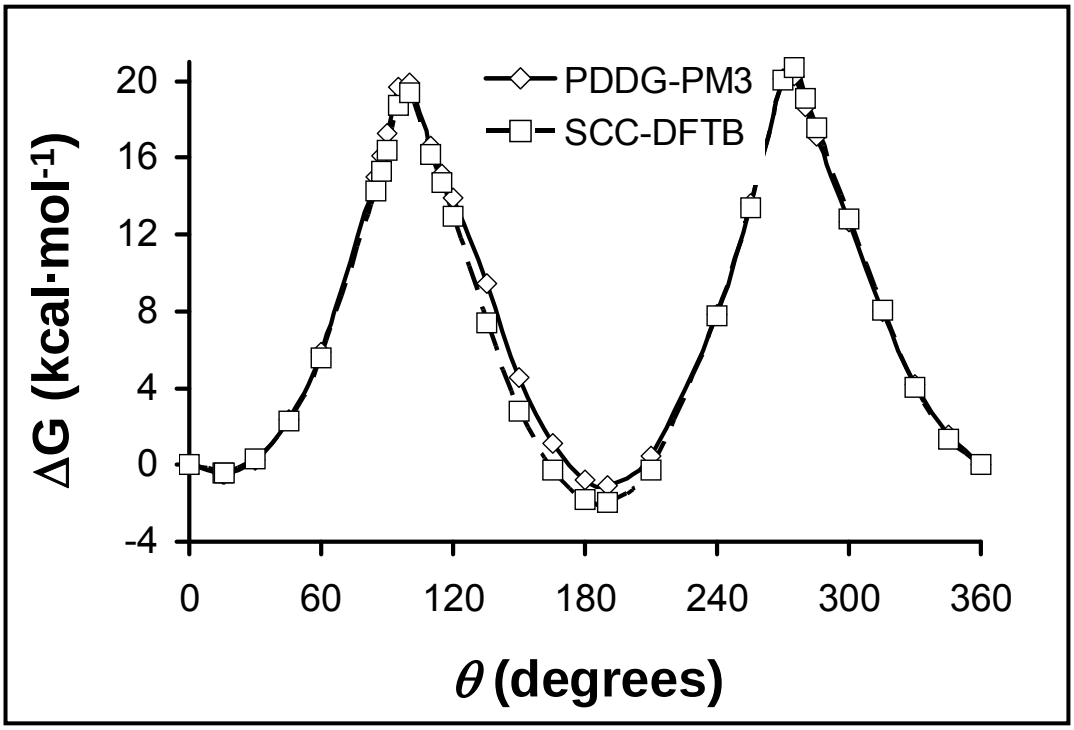
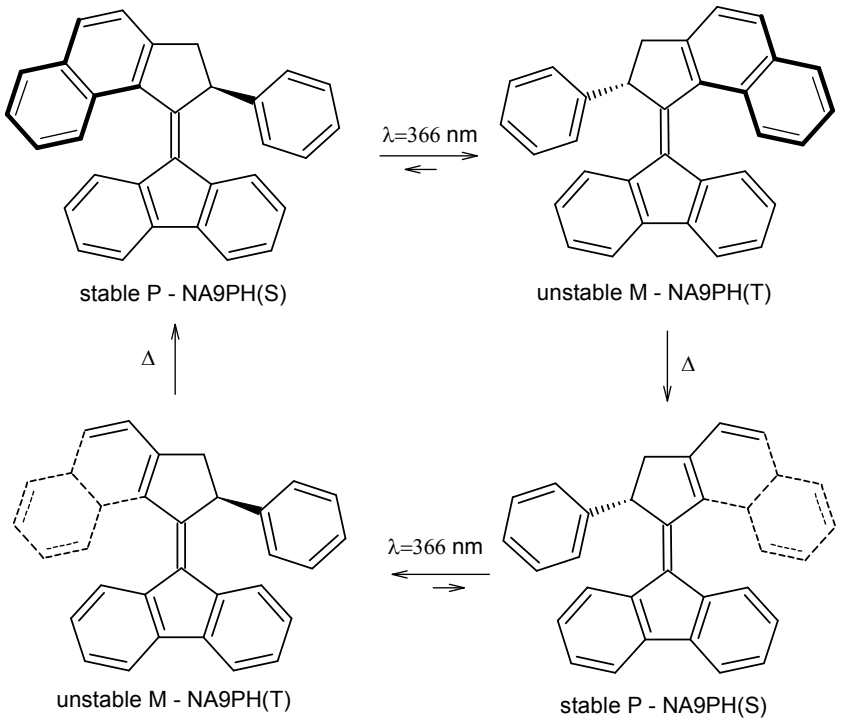


Figure 5. Rotational profiles in chloroform solution calculated for the singlet (S) state of NA9PH using QM/MM with umbrella sampling at PDDG-PM3 and SCC-DFTB levels.

A detailed analysis of the geometries recorded from QM/MM trajectories shows structural features similar to those found in the gas phase. Specifically, in the region close to the barrier, the carbon atoms involved in the bond that connects the rotor and the stator transform their hybridization from sp^2 to sp^3, facilitating the rotation of naphthalene over the fluorene. This is evidenced by the increase of the average values of the bond length that involve such carbon atoms and the dihedral angle χ. On the other hand, the average of phase angle P is completely unmeaning, even though the standard error of the mean shown in Table 3 is low, because the standard deviation obtained from statistical treatment shows a very large value (75° and 56° with the PDDG-PM3 and SCC-DFTB Hamiltonians, respectively). This indicates that in the region of the barrier, a wide set of conformations, very close in energy, are visited by the five membered ring of the rotor during the restrained QM/MM MD trajectories. It should be noted that, typically, complex systems are dense in saddle points, and, therefore, transition states are not necessarily identified with such saddle points but ensembles of transition states.[41]

Switching mechanism

In a previous work of Vicario et al.,[7] the rotation of the double bond that connects the rotor and the stator in NA9PH was described through a two-step mechanism (see Scheme 2). The first step consists of a photochemical isomerization with inversion of molecular helicity (P → M). In this unstable NA9PH structure, which involves a conformational change of the five-membered ring located at the rotor, the exocyclic phenyl substituent is forced to adopt a strained pseudoequatorial arrangement. In the second step, the strain of the previous unstable structure is released through an irreversible thermal helix-inversion (M → P), which produces a stable NA9PH structure. It is worth noting that there is a clear correspondence between this mechanism and the minima obtained in the rotational profiles calculated in this work. Specifically, the minimum of NA9PH(S) corresponds to the stable P structure, while the minimum of NA9PH(T) refers to the unstable M structure. The optimized structures reported in this work evidence that the rotation is facilitated by the elongation of the bond between the rotor and stator, that is, reduction of the double bond character. The barrier determined experimentally, 21 kcal mol^{-1},[7] is in good agreement with the theoretical estimations reported in this work, which unfortunately are limited by the relatively low accuracy of the computational levels used to describe this large molecule.



Scheme 2

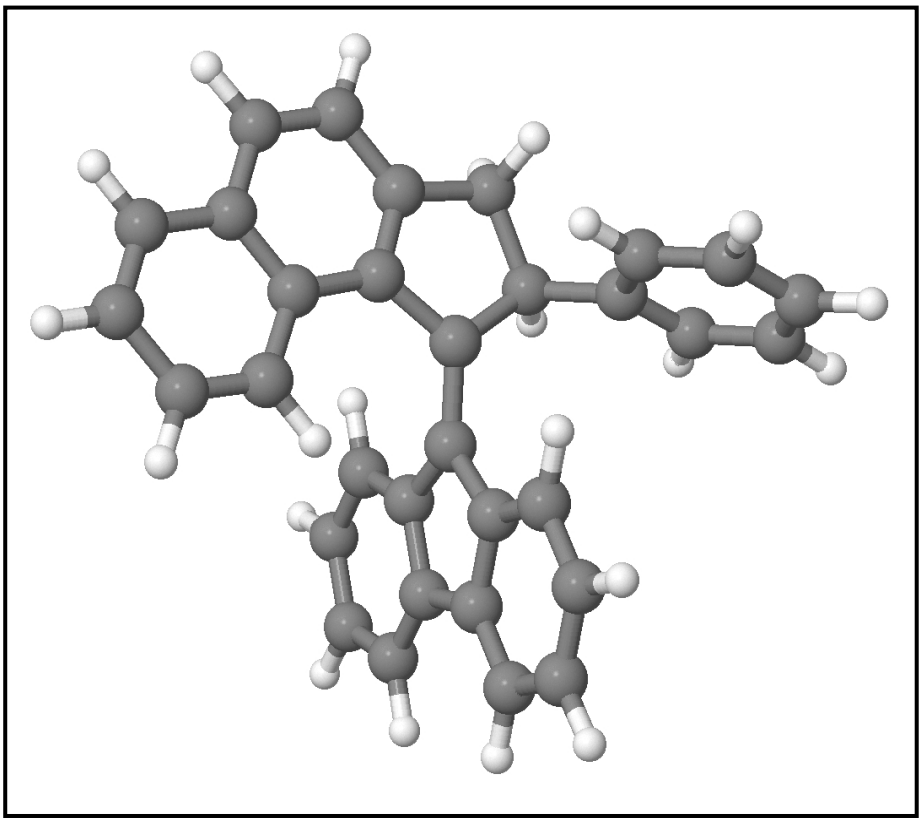


Figure 6. Structure obtained for the conical intersection between the two electronic states of NA9PH.

On the other hand, the switch from NA9PH(T) to NA9PH(S) is expected to occur after crossing the rotational barrier of the former through a conical intersection between the surfaces of the two states. To identify qualitatively the existence of this conical intersection between the ground and excited states, a geometry optimization of NA9PH in the gas phase using CASSCF(2,4)/STO-3G and CASSCF(4,6)/STO-3G has been carried out, with a total of 16 and 225 configurations, respectively. An intermediate structure, which is displayed in Figure 6, was located at about $\theta \approx 73°$ in both levels of calculation used (see Table 4). In this structure, the naphthalene moiety and the 5-membered ring of the rotor (τ_m = 1.6° and 1.6°, respectively) are planar, while the carbon atom of the fluorine involved in the bond connecting the rotor and the stator shows a small but non-negligible degree of bending; that is, it deviates 15.7° and 15.9° from planarity, respectively.

Table 4. Relevant geometrical parameters (see text) for the most stable structures of NA9PH Singlet-Triplet conical intersection optimized in the gas phase using CASSCF(2,4)/STO-3G and CASSCF(4,6)/STO-3G theoretical procedures

	CASSCF/STO-3G	
	(2,4)	**(4,6)**
θ [a]	73.0	72.6
d [b]	1.490	1.491
χ [a]	-4.35	-3.87
P [a,c]	-81.7	-80.3
τ_m [a]	1.56	1.68

[a] In degrees. [b] Bond length (in Å) associated with the molecular axle. [c] The conformation of the five-membered ring belonging to the rotor (Scheme 1b) is indicated in parentheses. Both the phase angle (P) and the amplitude (τ_m) have been determined using the model reported in ref 39.

6.3.4. *Conclusions*

In this chapter, several different theoretical methods have been applied to study the rotational profiles of the NA9PH light-driven nanomachine. The singlet and triplet states of NA9PH have been considered at the quantum mechanical level, while the solvent effect was described using two different procedures: (i) implicitly using the SCRF approach; and (ii) explicitly under the QM/MM scheme. UB3LYP/6-31G(d) estimate of the free energy barrier in the gas phase for the singlet and triple states was 31.4 and 17.0 kcal mol^{-1} (33.0 and 13.6 kcal mol^{-1} at the UMP2/6-31G(d) level), respectively. These values are in good agreement with the energy barrier experimentally determined for the thermal helix inversion (excited state) of NA9PH, that is, 21.0 kcal mol^{-1}. The incorporation of solvent effects through the PCM method produces a reduction in the free energy barriers of few kilocalories, while the umbrella sampling study under the QM/MM SCC-DFTB scheme yields a free energy barrier of 21.3 kcal mol^{-1} for the ground state. Analysis of barrier shows that the naphthalene and the five-membered ring of the rotor are planar, even though the bond that connects the rotor and the stator presents a significant deviation from the planarity in the atom belonging to the fluorene. This distortion allows an easy pass of the rotor over the stator.

6.3.5. *References*

[1] van den Heuvel M.G.L., Dekker C. *Science* **2007**, *317*, 333.

[2] Browne W.R., Feringa B.L. *Nat. Nanotechnol.* **2006**, *1*, 25.

[3] Kay E.R., Leigh D.A., Zerbetto F. *Angew. Chem., Int. Ed.* **2007**, *46*, 72.

[4] Credi A. *Aust. J. Chem.* **2006**, *59*, 157.

[5] Koumura N., Zijlstra R.W.J., van Delden R.A., Harada N., Feringa B.L. *Nature* **1999**, *401*, 152.

[6] Vicario J., Meetsma A., Feringa B.L. *Chem. Commun.* **2005**, 5910.

[7] Vicario J., Walko M., Meetsma A., Feringa B.L. *J. Am. Chem. Soc.* **2006**, *128*, 5127.

[8] Pollard M.M., Klok M., Pijper D., Feringa B.L. *Adv. Funct. Mater.* **2007**, *17*, 718.

[9] Eelkema R., Pollard M.M., Vicario J., Katsonis N., Ramon B.S., Bastiaansen C.W.M., Broer D.J., Feringa B.L. *Nature* **2006**, *440*, 163.

[10] Eelkema R., Pollard M.M., Katsonis N., Vicario J., Broer D.J., Feringa B.L. *J. Am. Chem. Soc.* **2006**, *128*, 14397.

[11] Pijper D., Jongejan M.G.M., Meetsma A., Feringa B.L. *J. Am. Chem. Soc.* **2008**, *130*, 4541.

[12] Senn H.M., Thiel W. QM/MM methods for biological systems. In *Atomistic Approaches in Modern Biology*, M., R., Ed., Springer: Berlin, **2007**, Vol. 268, p 173.

[13] Torrie G.M., Valleau J.P. *Chem. Phys. Lett.* **1974**, *28*, 578.

[14] Stewart J.J.P. *J. Comput. Chem.* **1989**, *10*, 209.

[15] Stewart J.J.P. *J. Comput. Chem.* **1989**, *10*, 221.

[16] Becke A.D. *J. Chem. Phys.* **1993**, *98*, 1372.

[17] Lee C., Yang W., Parr R.G. *Phys. Rev. B* **1988**, *37*, 785.

[18] a) Hariharan P.C., Pople J.A. *Theor. Chim. Acta* **1973**, *28*, 213. b) McLean A.D., Chandler G.S. *J. Chem. Phys.* **1980**, *72*, 5639.

[19] Møller C., Plesset M.S. *Phys. Rev.* **1934**, *46*, 618.

[20] Frisch M.J., Trucks G.W., Schlegel H.B., Scuseria G.E., Robb M.A., Cheeseman J.R., Montgomery J.A., Jr., Vreven T., Kudin K.N., Burant J.C., Millam J.M., Iyengar S.S., Tomasi J., Barone V., Mennucci B., Cossi M., Scalmani G., Rega N., Petersson G.A., Nakatsuji H., Hada M., Ehara M., Toyota K., Fukuda R., Hasegawa J., Ishida M., Nakajima T., Honda Y., Kitao O., Nakai H., Klene M., Li X., Knox J.E., Hratchian H.P., Cross J.B., Bakken V., Adamo C., Jaramillo J., Gomperts R., Stratmann R.E., Yazyev O., Austin A.J., Cammi R., Pomelli C., Ochterski J.W., Ayala P.Y., Morokuma

K., Voth G.A., Salvador P., Dannenberg J.J., Zakrzewski V.G., Dapprich S., Daniels A.D., Strain M.C., Farkas O., Malick D.K., Rabuck A.D., Raghavachari K., Foresman J.B., Ortiz J.V., Cui Q., Baboul A.G., Clifford S., Cioslowski J., Stefanov B.B., Liu G., Liashenko A., Piskorz P., Komaromi I., Martin R.L., Fox D.J., Keith T., Al-Laham M.A., Peng C.Y., Nanayakkara A., Challacombe M., Gill P.M.W., Johnson B., Chen W., Wong M.W., Gonzalez C., Pople J.A. *Gaussian 03*, revision C.02, Gaussian, Inc.: Wallingford, CT, **2004**.

[21] Miertus M., Scrocco E., Tomasi J. *Chem. Phys.* **1981**, *55*, 117.

[22] Tomasi J., Persico M. *Chem. Rev.* **1994**, *94*, 2027.

[23] Iribarren J.I., Casanovas J., Zanuy D., Alemán C. *Chem. Phys.* **2004**, *302*, 77.

[24] Jang Y.H., Goddard W.A., III, Noyes K.T., Sowers L.C., Hwang S., Chung D.S. *J. Phys. Chem. B* **2003**, *107*, 344.

[25] Hawkins G.D., Cramer C.J., Truhlar D.G. *J. Phys. Chem. B* **1998**, *102*, 3257.

[26] Cieplak P., Caldwell J., Kollman P. *J. Comput. Chem.* **2001**, *22*, 1048.

[27] Torrie G.M., Valleau J.P. *J. Comput. Chem.* **1977**, *23*, 187.

[28] Case D.A., Darden T.A., Cheatham T.E.I., Simmerling C.L., Wang J., Duke R.E., Luo R., Merz K.M., Pearlman D.A., Crowley M., Walker R.C., Zhang W., Wang B., Hayik S., Roitberg A.E., Seabra G., Wong K.F., Paesani F., Wu X., Brozell S., Tsui V., Gohlke H., Yang L., Tan C., Mongan J., Hornak V., Cui G., Beroza P., Mathews D.H., Schafmeister C., Ross W.S., Kollman P.A. *AMBER 9*, University of California: San Francisco, CA, **2006**.

[29] Wang J., Wolf R.M., Caldwell J.W., Kollman P.A., Case D.A. *J. Comput. Chem.* **2004**, *25*, 1157.

[30] Bayly C.I., Cieplak P., Cornell W., Kollman P.A. *J. Phys. Chem.* **1993**, *97*, 10269.

[31] Cornell W.D., Cieplak P., Bayly C.I., Kollmann P.A. *J. Am. Chem. Soc.* **1993**, *115*, 9620.

[32] Repasky M.P., Chandrasekhar J., Jorgensen W.L. *J. Comput. Chem.* **2002**, *23*, 1601.

[33] Elstner M., Porezag D., Jungnickel G., Elsner J., Haugk M., Frauenheim T., Suhai S., Seifert G. *Phys. Rev. B* **1998**, *58*, 7260.

[34] Elstner M., Hobza P., Frauenheim T., Suhai S., Kaxiras E. *J. Chem. Phys.* **2001**, *114*, 5149.

[35] Ryckaert J.P., Ciccotti G., Berendsen H.J.C. *J. Comput. Phys.* **1977**, *23*, 327.

[36] Kumar S., Rosenberg J.M., Bouzida D., Swendsen R.H., Kollman P.A. *J. Comput. Chem.* **1992**, *13*, 1011.

[37] Kumar S., Rosenberg J.M., Bouzida D., Swendsen R.H., Kollman P.A. *J. Comput. Chem.* **1995**, *16*, 1339.

[38] Roux B. *Comput. Phys. Commun.* **1995**, *91*, 275.

[39] Altona C., Sundaralingam M. *J. Am. Chem. Soc.* **1972**, *94*, 8205.

[40] Walker R.C. Personal communication, 2008.

[41] Bolhuis P.G., Dellago C., Chandler D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5877.

[42] Loos P.F., Preat J., Laurent A.D., Michaux C., Jacquemin D., Perpète E.A., Assfeld X. *J. Chem. Theory Comput.* **2008**, *4*, 637.

7. SUMMARY AND DISCUSSION OF THE RESULTS

The leitmotif of the present thesis is clearly summarized in its title: *Application of molecular simulation techniques to the design of nanosystems*. Throughout the previous chapters it has been clearly proven the suitability of such techniques in the design and improvement of nanodevices, as well as the prediction of their properties. As stated in the Introduction the inherent nature of nano and supramolecular scale, involving the operation of individual atoms, makes molecular simulation techniques the perfect tool to achieve this goal. All the results presented in this thesis have been grouped in four general blocks:

- Nanotubes based on natural peptide sequences
- Dendronized polymers
- Theoretical study of π-conjugated systems
- Molecular actuators and sensors based on conducting polymers

When designing a new nanodevice it is very common to survey naturally occurring structures with a natural conformational behavior favoring the needed structural shape and use them as starting point to build the new nanoconstruct. These naturally occurring structures can be modified in order to improve their stability and get the desirable properties. It is worth noting that the use of naturally occurring structures, besides favoring a specific shape, increases the chances of biocompatibility, which is a decisive factor when designing nanoconstructs for biotechnological purposes.

This philosophy has been clearly reported in Chapter 3, in which motifs taken from left-handed β-helical proteins have been used to design tubular nanoconstructs. Such motifs possess a natural tubular shape that makes them useful to build up nanofibers. The key point of these nanoconstructs to be used either as matter or charge carriers is the ability to retain their tubular geometry, which can be enhanced by reducing the mobility of the more flexible parts, *i.e.* the loop regions. In basis of the results obtained over 1krr and 1hv9 based tubular nanoconstructs, it is possible to increase the rigidity of such nanoconstructs by replacing targeted amino acids located in the more flexible parts of these motifs by synthetic amino acids with low conformational flexibility. This replacement, besides reducing the conformational mobility of the targeted and neighboring positions, must avoid the destabilization of the overall structure due to either steric hindrance or new adverse side chain interactions. Diphenylglicine and synthetic α-amino acids 1-aminocycloalkane-1-carboxylic acids (Ac_nc, where n indicates the size of the ring) have been proven as good candidates to stabilize tubular

nanoconstructs. In the latter kind of synthetic amino acids results led to conclude that the ability to adapt their folded conformations and fit better in the mutated structure improves with the size of the ring.

The same idea of using naturally occurring structures in terms of biocompatibility lies behind the design of nanodevices based on natural amino acids. One illustrative example of it is the design of dendronized polymethacrylates carrying second generation chiral 4-aminoproline-based dendrons. These dendronized polymers are characterized by a helical secondary structure which can be either left or right handed. The helix reported in the present thesis is right handed but its sense might vary depending on the chirality of the 4-aminoproline-based dendritic units. The great rigidity displayed by this polymer is explained in basis of the hydrogen-bonding network which plays the main role in retaining its secondary structure even at high temperatures and independently of the backbone tacticity. In the case of deprotected dendronized polymers, where each dendritic unit contains 4 positive charges, dendritic side chains are arranged so that the repulsive interactions between the positive charges of neighboring repeat units is minimized leading to the formation of a helical structure with the same sense as the neutral one. Helical structure of deprotected polymers is more temperature dependent because the atomic vibrations produced upon increasing the temperature enhance the electrostatic repulsions inducing the disruption of the secondary structure.

The second great family of systems presented in this thesis concerns conducting polymers and their interactions. The effect of solvent upon the conformational preferences of conducting polymers is very important due to the strong geometry dependence in their electronic properties. In this sense it has been demonstrated that the choice of the solvation method to study such effects depends on the nature of the solvent. While the PCM method provides accurate enough results for typical organic solvents, combined discrete/continuum methods are required when the solute is in aqueous medium. Both intermolecular and intramolecular electronic delocalization play a crucial role in the transfer of charge. In this sense study of π-stacking interactions in thiophene containing systems evidences that the formation of π-stacked aggregates promotes electron transport between different rings. Combination of this feature together with the stated above conditions to enhance the rigidity of tubular nanoconstructs were employed to design a nanowire based on 1krr motifs. Such

electronic conductor has been designed by replacing targeted amino acids by β-3-thienylalanine, a synthetic amino acid with a thienyl side chain able to create a ladder of π-stacking.

Intramolecular charge transport in polythiophene chains lies behind the sensory and actuation mechanisms of the, respectively, sensor and actuator presented in this thesis. Both sensor and actuator are thiophene-based polymers combined with versatile calix[4]arenes units. Calix[4]arenes are synthetic macrocycles consisting of four phenol or anisole rings connected via methylene bridges that can host different guest molecules leading to conformational rearrangement of the whole device making it useful to be employed as a sensor or actuator. Sensory mechanism of the nanosensor presented in this thesis, which evidences a great selectivity towards sodium cation with respect to the rest of alkaline cations, is based on the reduction of the conformational flexibility of the polythiophene backbone by reason of the binding of the cation to the sensor unit. As regards the nanoactuator, interactions between solvent molecules and the calix[4]arenes units play a fundamental role in the actuation mechanism. Thus combination of the great transport properties, electrical conductivity and rate of energy migration of conducting polymers provide amplified sensitivity in nanosensors and a rapid response in nanoactuators. Similarly *in silico* studies have clarified the mechanism of a light driven nanomachine based on aromatic molecules evidencing that the distortion of the bond that connects the rotor and the stator allows the pass of the rotor over the stator. Putting together intramolecular and intermolecular electronic delocalization in conducting polymers has helped to explain the charge transfer mechanism in thiophene based dendrons leading to conclude that it takes place thanks to both the intramolecular π conjugated system and the π and T stacking of the different thiophene units.

A general overview of all the research work displayed throughout all the previous chapters corroborates the increasing weight of nanotechnology in the present and future development of science and technology because of its promising applications as well as the strength of molecular simulation techniques to speed up all the progresses in this field.

8. CONCLUSIONS

(1) The ability of non-proteogenic amino acids to stabilize small building block motifs is site-dependent. As far as the replacement does not alter the energy balance between attractive non-covalent interactions and steric hindrance, synthetic residues are suitable candidates to nucleate β-helix formation.

(2) Incorporation of a highly constrained synthetic α-amino acid such as Ac_5c or Ac_6c within the more flexible regions of the β-helical building blocks is an excellent strategy to enhance the stability of the resulting nanotubes. Thus, when a single mutation is performed in the loop region of the β-helix, the bend architecture of the whole loop is stabilized since the conformational mobility is reduced not only at the mutated position but also at the adjacent positions. Comparison of all the results obtained with other Ac_nc mutated nanotubes indicates that the ability to adapt to the targeted position improves considerably with the ring size, *i.e.*, when the rigidity introduced by the strain of the ring decreases.

(3) Formation of an interdendron hydrogen bonding network in dendronized polymethacrylates carrying second generation chiral 4-aminoproline-based dendrons favors the formation of a very rigid structure with a high stability in a wide range of temperatures and different media. This fact leads to a helical conformation, which is independent of the backbone tacticity, whose sense might be varied by changing the chirality of the dendritic units.

(4) Dendritic units in deprotected derivatives of dendronized polymethacrylates carrying second generation chiral 4-aminoproline-based dendrons are arranged so that the repulsive interactions between the positive charges of neighboring repeat units are minimized, which leads to the formation of a helical structure whose characteristic helical parameters are very similar to those found in the protected dendronized polymer. However, the atomic vibrations produced upon increase the temperature enhance the electrostatic repulsions inducing the disruption of the secondary structure confirming its lowest stability compared with the protected analogue.

(5) Exceptional electronic conductivity in thiophene based dendrons is explained on the basis of the contribution of both the intramolecular π conjugated system and the intermolecular π and T stacking between different thiophene units. The energy difference between HOMO and LUMO, *i.e.* the band-gap, is inversely

proportional to the dendritic generation degree of the polymer enhancing the intramolecular delocalization of π-electrons as the dendron grows. Additionaly, highest generation dendrimers exhibit a very low ionization potential per ring due to the large electron delocalization.

(6) Study of the rotation profile of internal rotation of 2,2'-bithiophene in different solvents points that the polarizable continuum model provides a satisfactory description of the solvent effects in acetonitrile solution. However, combined discrete/self-consistent reaction-field calculations are more appropriated in aqueous solution, where both the first-solvation shell effects and the polarization induced by the bulk solvent affect the rotational profile of 2,2'-bithiophene.

(7) Cancellation of favorable induction contribution by the destabilizing dispersion component causes negligible or even slightly positive n-body effects for stacked thiophene arrangements in the neutral state. Simultaneous destabilization of the HOMO and stabilization of the LUMO makes the lowest transition energy vary inversely with the size of the complex.

(8) π-Stacked complexes of thiophene-containing molecules in the radical cation state are formed at the dimer level, being the T-shaped arrangement unstable. The variation of the intermolecular distances in such complexes indicates that the stability of the dimers increases with the localization of the positive charge within each molecule. Within the context of charge transport in oxidized conducting polymers, the intermolecular phenomena are small with respect to the intramolecular ones.

(9) MD simulations showed that a rational design of the targeted replacements of natural amino acids at selected positions in a nanotube based on natural peptide sequences by β-3-thienylalanine enhances the stability of a tubular nanostructure and forms a π-stacking ladder in the inner core. QM/MM calculations on the oxidized nanotube evidenced the delocalization of the π-electron deficiency between the thienyl side groups of the different synthetic residues corroborating its potential use as a nanowire.

(10) Results of the dynamics of the dichloromethane solvent around the calix[4]arene units contained in the molecular actuator poly(calix[4]arene bis-bithiophene) over 1μs MD simulation indicate that a solvent molecule fills the cavity associated to the cone conformation of the macrocycle during a

significant period of time, especially when the actuator is not contracted. The position of such solvent molecule presents a fourfold symmetry with a maximum orientation toward the center of each ring contained in the calix[4]arene.

(11) Frequently, the solvent molecule located inside the cavity is rapidly exchanged for another molecule of the bulk. Thus, the number of dichloromethane molecules that reside more than 0.5 ns inside the cavity is relatively small. Finally, it has been detected that a significant number of solvent molecules are able to migrate from the cavity of one calix[4]arene to the cavity/ies of other/s, suggesting that the dynamics of the bulk solvent is important for the formation of these inclusion complexes.

(12) Evaluation of the sensing response of polythiophene functionalized nanosensor with calix[4]arene to Li^+, Na^+, and K^+ indicates that Na^+ interacts electrostatically with the oxygen-rich site of the receptor reducing the conformational flexibility of the polymer backbone. In contrast, Li^+ and K^+ interact with the tert-butyl groups of the macrocycle enhancing such conformational flexibility. These features are consistent with available experimental data explaining the selectivity of this material for Na^+ over Li^+ and K^+.