IL NUOVO CIMENTO Vol. 35 C, N. 5 Suppl. 1 Settembre-Ottobre 2012
DOI 10.1393/ncc/i2012-11331-5

Colloquia: PR PS BB 2011

Symmetries, functional theory and simulation models for macromolecules and polymers

G. Maino(*)

ENEA - via Martiri di Montesole 4, 40128 Bologna, Italy
University of Bologna - Ravenna Campus, Italy

ricevuto il 30 Settembre 2011; approvato l' 1 Dicembre 2011

Summary. — Functional techniques have proven to be very powerful tools in the treatment of integro-differential equations which rule the dynamics of crystallographic structures or of the Free Electron Laser (FEL), allowing for a derivation of the relevant solutions in closed form by means of Bessel-Clifford functions of half-integer order which are easily computable. Therefore, this approach coupled to algebraic methods can be extended to other physical and biochemical fields, where the dynamics of the condensed-matter system is determined by integro-differential equations of Volterra type, satisfying very general conditions, thus allowing calculations that are otherwise both complex and cumbersome. This paper shows that it is possible to reduce very complex problems such as those of protein folding or polymer/dendrimer dynamics to a series of manageable steps, starting from the description of fundamental structural patterns to their mutual interaction and the computation of the dynamics up to the formation of the final equilibrated 3D structure.

PACS 64.70.km – Polymers.
PACS 36.20.Kd – Electronic structure and spectra.
PACS 82.35.Lr – Physical properties of polymers.
PACS 87.14.E- – Proteins.

1. – Introduction

Lie groups and relevant algebraic methods have been extensively investigated and applied to the spectroscopy of atoms and molecules [1] with significant results in the understanding of roto-vibrational spectra of polyatomic molecules by taking into account in a very simple yet realistic way anharmonicities in both energy spectra and interactions

(*) E-mail: giuseppe.maino@unibo.it

among different modes. The algebraic model allows to calculate not only energy spectra but also transition intensities, namely infrared, Raman and Franck-Condon ones [2-4]. The same approach can be used to describe more complex molecules consisting of small subunits exhibiting some kind of symmetry.

Symmetry plays a major role in physical chemistry for describing the static equilibrium structure of atoms and molecules and many properties as well as transition probabilities of their excited electronic, vibrational and rotational states [5]. As an example, protein structures can be approximately classified using the crystallographic point groups such as cyclic groups, dihedral groups and cubic groups, translational symmetry and the functions related to a given symmetry. Finally, helical symmetry is a crucial ingredient of secondary protein structure, not only of DNA and RNA shapes [6-9].

While the theoretical methods of group theory are suitable for the description of small and medium-size molecules [1-5], complex symmetries in proteins have been investigated by complementary approaches such as the mathematical theory of knots [10], through a contact map which represents the matrix form of a polymer graph of a particular chain conformation.

Approximate shapes of secondary structures of proteins, namely alpha-helices and beta-sheets or beta-barrels, have been recently explained, for instance, as the best families of simple approximating sets [11], by using the symmetry methods firstly introduced by mathematician Gromov [12].

On our side, some years ago we introduced [13-15] functional techniques for the treatment of integro-differential equations which rule the dynamics of the Free Electron Laser (FEL). These techniques allow a derivation of the relevant solutions in closed form by means of analytic expressions depending on Bessel-Clifford functions of half-integer order which are easily computable [16]. Therefore, our mathematical method provides a simple yet precise framework for the description of the complex dynamics of a FEL in the unsatured regime, a problem that generally requires the numerical solution of a coupled set of Maxwell-Lorentz equations governing the interaction between the electron beam and the electromagnetic field and, consequently, a large computational effort.

The same approach can be applied to other physical chemical or biological complex systems, whose dynamics is ruled by integro-differential equations of Volterra type, satisfying very general conditions, thus allowing calculations that are otherwise both difficult and cumbersome. In particular, integro-differential equations of Volterra type whose integral kernel can be developed in Fourier series, as only assumption, admit simple analytic solutions which are easily computable even for the three-dimensional case. By this way it is possible to reduce very complex problems such as those of protein folding or polymer dynamics to a series of manageable steps, starting from the description of fundamental structural patterns—thanks to exploitation of group theory techniques provided the underlying symmetries—to their mutual interaction and the computation of the relevant dynamics up to the formation of the final equilibrated 3D structure, as shown in sect. **2**.

Integro-differential equations of this kind describe, for instance, the dynamics of chemical reactions in condensed phase extending kinetics theory to a high level of structural complexity when compared with reactions in the gas phase [17]. The theory of condensed phase processes meets the additional difficulty of having to deal with solvent interactions, but the detailed understanding of their effects is crucial in many areas of chemistry, biology and related sciences.

Molecular dynamics of proteins is one of the most important fields of application which this formalism can deal with. For the interpretation of the structural and functional

properties of proteins, it is essential to have a knowledge of the average positions of the atoms, and the magnitudes and time scales of the fluctuations about the average positions in a large number of conformational substates continuously converting each other. The characteristic times of the considered processes are not very different and non-Markovian effects must be properly included. Then, the integral kernel propagates the conseguences of dynamics began in the past to the actual time and accounts for the "memory" of the system as far as its history is concerned. A preliminary application of this approach to the structure and dynamics of a particular class of macromolecules (dendrimers) is described in sect. **3**.

2. – Analytical solutions of integro-differential equations

Kramers' model represents a classical framework for dealing with thermally activated transitions in bistable potentials [18, 19]. The chemical reaction is then modelled by a transition of a reaction coordinate, x, in a double-well potential, $V(x)$. A Langevin-type equation rules the Brownian motion of the reaction coordinate, including a stochastic force; this equation is then replaced by an integro-differential equation, as shown in ref. [17].

Suitable kernels, including memory effects, can model the solvent dynamics or intramolecular interactions in complex structures which affect the primary reactive process. As an example, it is worth mentioning the protein dynamics where the effective friction acting on the ligands in the interior of the protein itself depends on both the intrinsic properties of the considered macromolecule and environment (solvent).

Formally, the integro-differential equation of ref. [17] can be replaced by a corresponding differential equation of third order with suitable initial conditions, and then solved with classical numerical tools. However, since it corresponds to an ill-posed problem according to Hadamard's well-known definition, this is just a formal result because of the intrinsic, unpredictable large numerical instabilities associated with the "classical" solution, the degree of ill-posedness depending on the decay rate of the kernel transform, the time derivative of the stochastic force which in general cases implies strong oscillations, and the smoothness of the searched solution. The functional technique here proposed avoids these difficulties and allows us to develop a simple, intrinsically stable, algorithm which gives the exact solution of the integro-differential equation in any possible case of physical interest.

From a mathematical point of view, according to the one-dimensional model previously described (the extension to the 3D case is straightforward), the starting point is a linear integro-differential equation, whose kernel fulfills some very general requirements, namely to be a continuous and piecewise smooth function. The piecewise smooth conditions are not severe requirements from a physical point of view; kernel functions relevant to chemical reactions must have continuous behaviour allowing for a finite number of discontinuities.

The resulting equation has the same form than the single-mode FEL integro-differential equation, for instance, ruling its dynamics in the high-gain regime where transverse and longitudinal effects are negligible [16]. The only difference lies in the presence of a summation because of the Fourier expansion of the relevant kernel. One can then apply the same functional analysis technique developed in refs. [14-16].

Since our resulting equation is a Volterra integral equation of the second kind with continuous kernel of the evolution type, it admits one and only one continuous solution whatever the values of the expansion parameters. The solution is then recovered as

the limit of a Neumann series expansion [20-22]. Finally, due to the particular form of the integral operator, the relevant solution is given in closed analytical form, thus dropping the numerical evaluation of the multiple integrals. In fact, the analytical closed-form solution can be expressed in terms of modified incomplete gamma functions of purely imaginary argument [23], which can be also easily evaluated by means of efficient algorithms.

Functional analysis techniques in combination with Lie algebraic methods, to fully exploit the symmetries of the considered macromolecules and reduce the number of degrees of freedom in realistic calculations, may be thus introduced as useful complement to molecular-dynamics methodologies for proteins in order to investigate local and collective motions of macromolecules.

The identification of the relationship between gene and desease is a major achievement of modern molecular biology and remarkable advances have been obtained in the last decades, up to the start of the Human Genome project. However, from a practical—clinical and pharmacological—point of view, difficulties in applications arise from the complexity of reconstructing the three-dimensional (3D) spatial structure of the protein from the sequential data representing the constituent aminoacids [24].

Computational tools can bridge the gap between sequence and protein 3D-structure based on the notion that information is to be retrieved from the data bases and that knowledge-based methods can help in approaching a solution of the protein folding problem. To this aim, in the past it has implemented neural network-based predictors to derive, for instance, the secondary structure of globular and membrane proteins, the topology of membrane proteins and porins and stable alpha helical segments suited for protein design. This neural system is able to predict the secondary structure of any protein starting from its amino acid contents with an average accuracy, $Q_3 = 74.1\%$ [25]. Moreover, it has proved that larger sets of proteins given as input to the network system improve the accuracy of the prediction and verified that the evolutionary information and the filtering through the second network increase the performance of the predictor.

Starting from these theoretical data about secondary structure, it is possible to gain a useful insight into the symmetry patterns of otherwise unknown protein structures. An algebraic description of small molecular units can be then used to obtain suitable models of protein pieces that interact through the dynamical processes modelled by the previous functional-theory approach. As a whole, a dramatic reduction in the numerical complexity of the protein folding problem can be attained and in some cases simple solutions can be carried out that provide a convenient physical framework for the understanding of mechanisms responsible for the final 3D structure and the relevant functional properties. Similar remarks hold for the formation of polymer structures; preliminary results have been obtained for the generations of dendrimers where a phase transition from normal to fractal dimension has been found [26, 27].

3. – Molecular-dynamics analysis of the structure of dendrimers

In last years, many progresses have been made in the analysis of structural properties of a special class of branched polymers, named dendrimers. Owing to their extremely particular structure and physical behaviour, these macromolecules have important applications in many fields such as material science, pharmacology, medicine (for instance, improvement of cancer therapy), nanochemistry, viscosity modification and biochemistry (for example, micelle, liposome and biomolecule mimics) [28-33].

TABLE I. – *Number of terminal amino groups, NZ, number of monomer repeat units, NRU, and number of branch cells, NBC, for starburst PAMAM dendrimers, as a function of the generation number, g. The generation $g = 0$ corresponds to the initial core.*

g	NZ	NRU	NBC
0	3	3	0
1	6	9	3
2	12	21	9
3	24	45	21
4	48	93	45
5	96	189	93
6	192	381	189
7	384	765	381
8	768	1533	765
9	1536	3069	1533
10	3072	6141	3069

The internal structure of dendrimers has been investigated from both a theoretical and experimental point of view. Experimental studies of structural properties of dendrimers have been carried out by means of different techniques such as size exclusion chromatography (SEC) [34], nuclear magnetic resonance spectroscopy (NMR) [35, 36], small-angle neutron scattering (SANS) [37], small-angle X-ray scattering (SAXS) [38, 39] and rheology [40].

Experimental data support the existence of a structural change from an open shape to a spherical one as dendrimer growths [41]. This transition in symmetry shape is also indicated by molecular dynamics (MD) simulations [42, 43]. However, analytical calculations, performed within the framework of the self-consistent mean field (SCMF) model [44], have provided contradictory results in comparison to MD simulations [42], the former predicting the existence of cavities and voids near to the centre of dendrimer (hollow inner core) with the end groups of the dendritic polymer confined to its surface and the latter indicating a dense core shell with a distribution of the end groups throughout the molecule.

In this work, as an example of present possibilities of these model simulations, some results of MD simulations are presented, carried out for polyamidoamine (PAMAM) dendrimers. We have focused our attention on the study of structural (radius of gyration, structure function, and so on) and thermodynamics (for instance, energies and temperature) properties of these non-ionic dendrimers.

By comparing the data obtained from our MD simulations with the literature ones, we found a structural transition from an internal self-similar structure to a spherical one. These results confirm the PAMAM dendrimer structure proposed by experiments performed by means of SAXS and quasi elastic light scattering (QELS) techniques.

The first ten generations ($g = 1$–10) of non-ionic poly(amidoamine) (PAMAM) starburst dendrimers with an ammonia-central core are shown in figs. 1-4. Table I shows the main characteristics of each theoretical configuration.

Dendrimers are three-dimensional, monodisperse, highly ordered and covalent macromolecules. They differ from hyperbranched polymers by their extremely regular archi-

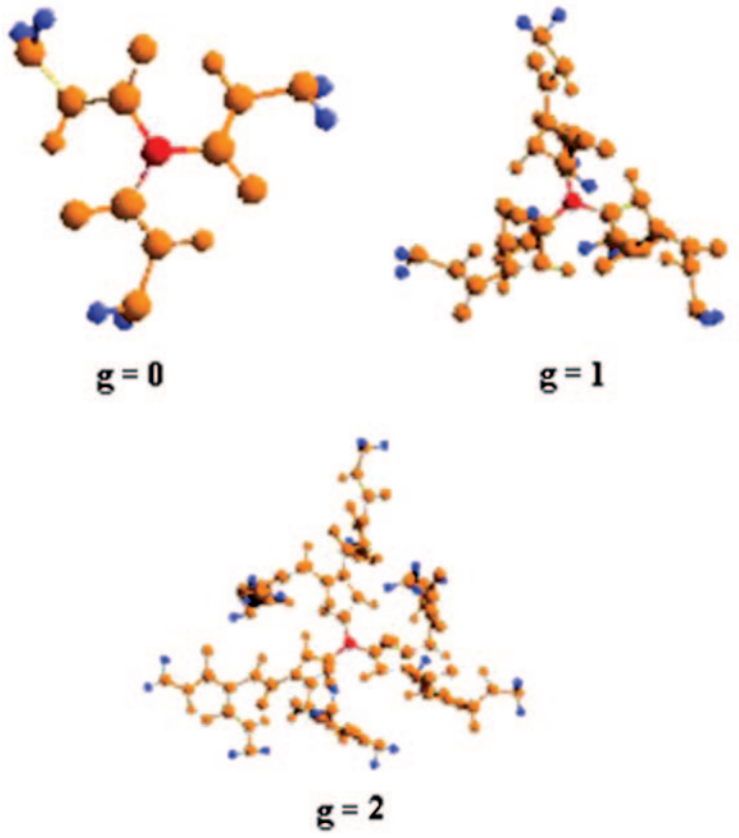

Fig. 1. – Computer-simulated molecular graphics of the zero, first and second generations ($g = 0, 1$ and 2) of starburst PAMAM dendrimer with an ammonia core ($N_c = 3$).

tecture. Indeed, hyperbranched polymers are produced in a non-iterative polymerization procedure, while the synthetic building of dendrimers proceeds by iterative reaction sequences beginning from a central multifunctional initiator core, with multiplicity (number of reactive sites) $N_c > 1$, by means of systematic introduction of monomer repeat units (branch cells) with branch-juncture multiplicity $N_b > 1$.

The final structure is a concentric assembly of branch cells which fills the space following a geometric growth (generation). However, dendrimers cannot grow indefinitely because the number of monomer units increases exponentially with generation, while the volume that the dendritic molecule can fill only grows as g^3. As a consequences, there exists a dendrimer limited generation beyond which the structure of the molecule becomes irregular. This critical generation number is referred to as dendrimer dense packing [44].

As for symmetry considerations, also the physical properties of dendrimers (radius of gyration, radial density profile, structure function, and so on) can be affected by the symmetry of the interior branch cells. For instance, when the dendrimer shows a symmetrical internal structure, the connectivity paths from the interior core to the end groups are identical. As a result, the terminal groups of the dendritic molecule lie on a sphere (starburst topology).

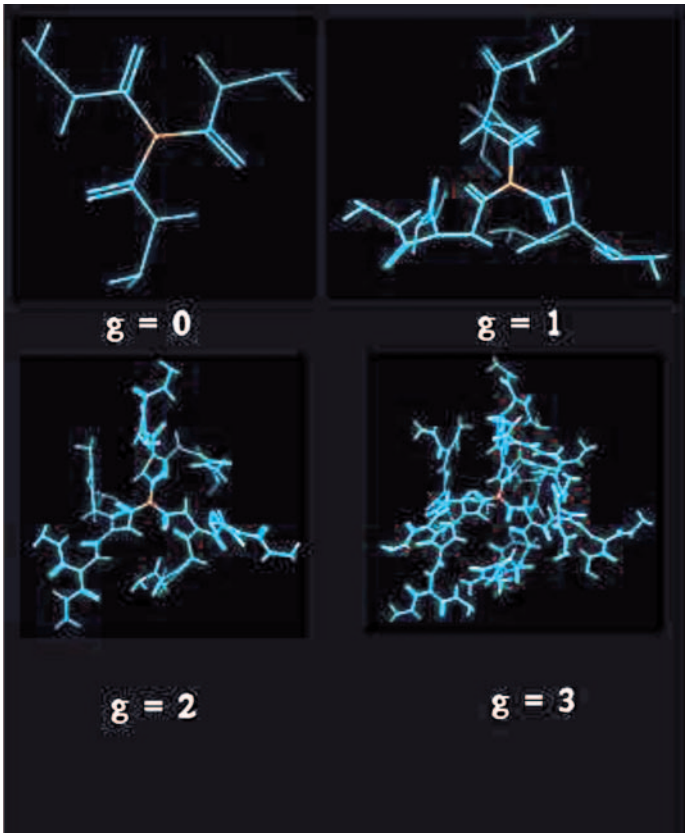

Fig. 2. – Simulations of the first three generations ($g = 0$–3) of PAMAM starburst dendrimers. The generation $g = 0$ is the ammonia core.

On the contrary, when the dendrimer branch cells are unsymmetrical, two identical connectivity paths between the interior core and the final groups (Denkewalter topology) do not exist [45]. Unlike starburst dendrimers, the surfaces of Denkewalter structures have higher fractal dimension and interior branch cells less ordered.

4. – Concluding remarks

Prediction of ternary and quaternary protein structure is a major challenge and lies besides the computational capability of the present parallel computers. Even if in the future great improvements will be possible for computing power, numerical calculations only do not provide a complete and clear understanding of macromolecular structures. As in the case of atomic nuclei, (dynamical) symmetries and group-theoretical approaches—like the interacting boson-fermion models and the vibron model [1]—allow a better understanding of mechanisms underlying the nuclear or molecular dynamics and the formation of particular structures and shapes, that are not evident from direct calculations.

A suitable combination of algebraic, functional analytical and simulation (or numerical) methods possibly represents a key to deal with the prediction of macromolecule and protein structures and their interaction properties with the surroundings. In the previous

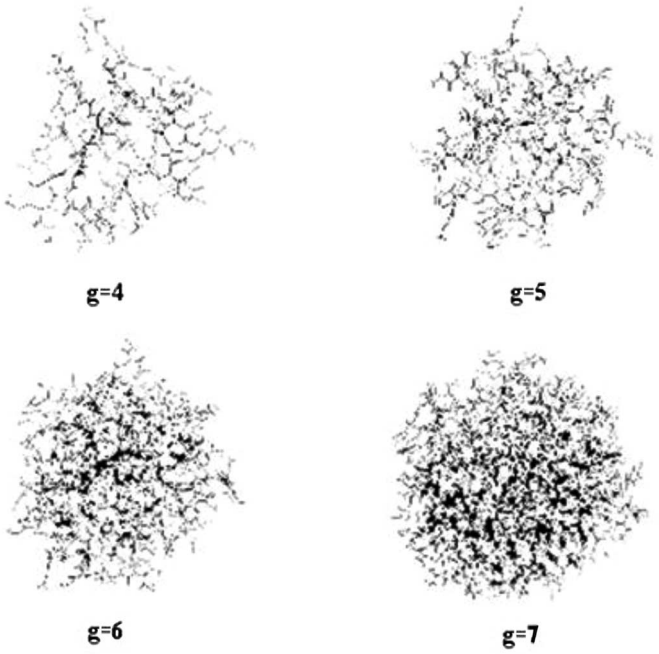

Fig. 3. – As fig. 2, but for generations $g = 4$–7.

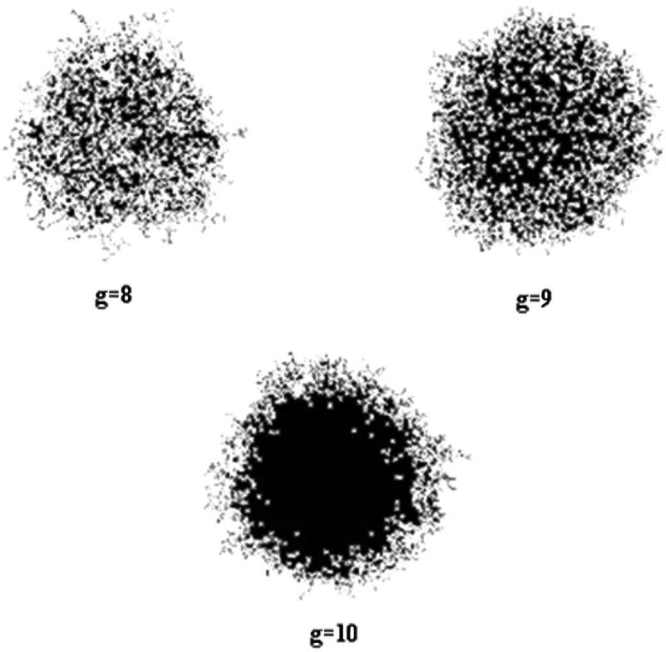

Fig. 4. – As fig. 2, but for generations $g = 8$–10.

sections, a few examples have shown the richness and potential of these techniques, once applied to the study of large complex systems of interacting particles or substructures. The present line of investigation concerns the way in which the analytical descriptions of the simple parts of the whole molecular structure can be inserted in molecular-dynamics calculations or other simulation models in order to make the problem of the 3D structure prediction affordable within the capabilities of the actual computers.

REFERENCES

[1] Iachello F. and Levine R. D., *Algebraic Theory of Molecules* (Oxford University Press, Oxford) 1995.
[2] Oss S., *Adv. Chem. Phys.*, **93** (1996) 455.
[3] Frank A. and van Isacker P., *Algebraic Methods in Molecular and Nuclear Structure* (John Wiley, New York) 1994.
[4] Iachello F. and Oss S., *Eur. Phys. J. D*, **19** (2002) 307.
[5] Kellman M. E., *Proc. Natl. Acad. Sci. USA*, **93** (1996) 14287.
[6] Goodsell D. S. and Olson A. J., *Annu. Rev. Biophys. Biomol. Struct.*, **29** (2000) 105.
[7] Levitt M. and Chothia C., *Nature*, **261** (1976) 552.
[8] Richardson J. S., *Adv. Protein Chem.*, **34** (1981) 167.
[9] Thornton J. M., Jones D. T., Macarthur M. W., Orengo C. M. and Swindells M. B., *Philos. Trans. R. Soc. London Ser. B*, **348** (1995) 71.
[10] Chen S. and Dill K. A., *J. Chem. Phys.*, **104** (1996) 5964.
[11] Nava J. and Kreinovich V., *Symmetry*, **4** (2012) 15.
[12] Gromov M., *Bull. Am. Math. Soc.*, **48** (2011) 229.
[13] Dattoli G., Lorenzutta S., Maino G. and Torre A., *Radiat. Phys. Chem.*, **47** (1996) 183.
[14] Maino G., Lorenzutta S., Dattoli G., Torre A. and Tocci D., *Nuovo Cimento D*, **16** (1994) 1325.
[15] Dattoli G., Lorenzutta S., Maino G., Tocci D. and Torre A., in *Dynamics of Transport in Plasmas and Charged Beams*, edited by Maino G. and Ottaviani M. (World Scientific Publ. Co., Singapore) 1996, pp. 202.
[16] Dattoli G., Lorenzutta S., Maino G. and Torre A., *Ann. Num. Math.*, **2** (1995) 211.
[17] Bonifazzi C., Maino G. and Paganetto G., *Nuovo Cimento D*, **20** (1999) 2041.
[18] Kramers H. A., *Physica*, **VII** (1940) 284.
[19] Compiani M., *J. Chem. Phys.*, **98** (1993) 602.
[20] Deufelhard P. and Hairer F. (Editors), *Numerical Treatment of Inverse Problems in Differential and Integral Equations* (Birkhafiser, Boston) 1983.
[21] Budak B. M. and Fomin S. V., *Multiple Integrals, Field Theory and Series* (MIR Publ., Moscow) 1973.
[22] Brunner H. and van der Houwen P. J., *The Numerical Solution of Volterra Equations* (North Holland, Amsterdam) 1986.
[23] Slater L. J., *Confluent Hypergeometric Functions* (Cambridge University Press, Cambridge) 1960.
[24] Anfinsen C. B., *Science*, **181** (1973) 223.
[25] Fini F., Maino G., Casadio R. and Fariselli P., L'informazione evolutiva nella predizione della struttura secondaria delle proteine, rapporto ENEA RT/2006/47/FPN, 2006, 92 pp.
[26] Mallamace F., Canetta E., Lombardo D., Mazzaglia A., Romeo A., Monsu Scolaro L. and Maino G., *Physica A*, **304** (2002) 235.
[27] Canetta E. and Maino G., *Nucl. Instrum. Methods Phys. Res. B*, **213** (2004) 71.
[28] Newkome G. R., Moorefield C. N. and Voegtle F., *Dendritic Molecules. Concepts-Synthesis-Perspectives* (VCH Publishers, New York) 1995.

[29] Bosman A. W., Jansenn H. M. and Meijer E. W., *Chem. Rev.*, **99** (1999) 1665.
[30] Grayson S. M. and Fréchet M. J., *Chem. Rev.*, **101** (2001) 3819.
[31] Berresheim A. J., Müller M. and Müller K., *Chem. Rev.*, **99** (1999) 1747.
[32] Newkome G. R., He E. and Moorefield C. N., *Chem. Rev.*, **99** (1999) 1689.
[33] Tomalia D. A., Naylor A. M. and Goddard W. A. III, *Angew. Chem. Int. Ed. Engl.*, **29** (1990) 138.
[34] Bu L., Nonidez W. K., Mays J. W. and Tan N. B., *Macromolecules*, **33** (2000) 4445.
[35] Malyarenko D. I., Vold R. L. and Hoatson G. L., *Macromolecules*, **33** (2000) 1268.
[36] Choi M., Niu Y., Youngs W. J. and Rinaldi P. L., *J. Am. Chem. Soc.*, **123** (2001) 4670.
[37] Ramzi A., Scherrenberg R., Brackman J., Joosten J. and Mortensen K., *Macromolecules*, **31** (1998) 1621.
[38] Mallamace F., Canetta E., Lombardo D., Mazzaglia A., Romeo A., Monsù Scolaro L. and Maino G., *Physica A*, **304** (2002) 235.
[39] Prosa T. J., Bauer B. J. and Amis E. J., *Macromolecules*, **34** (2001) 4897.
[40] Uppuluri S., Morrison F. A. and Dvornic P. R., *Macromolecules*, **33** (2000) 2551.
[41] Prosa T. J., Bauer B. J., Tomalia D. A. and Scherrenberg R., *J. Polym. Sci., Part B: Polym. Phys.*, **35** (1997) 2913.
[42] Lescanec R. L. and Muthukumar M., *Macromolecules*, **23** (1990) 2280.
[43] Murat M. and Grest G. S., *Macromolecules*, **29** (1996) 1278.
[44] de Gennes P. G. and Hervet H., *J. Phys. (Paris) Lett.*, **44** (1983) L351.
[45] Denkewalter R. G., Kolc J. F. and Lukasavage W. J., *Chem. Abstr.*, **100** (1984) 103.