Molecular Crystals and Liquid Crystals

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gmcl20

Application of Genetic Algorithms to Control Synthesis of Block Copolymers Containing PDMS - CO - PMAA Sequences

Teodora Rusu* [a] & Oana Gogan [b]

[a] "P. Poni" Institute of Macromolecular Chemistry, Iasi

[b] ARTINFO SRL, Roumanie branch of ARTI INFOMATICHE, Iasi, Italy

Version of record first published: 18 Oct 2010

To cite this article: Teodora Rusu* & Oana Gogan (2004): Application of Genetic Algorithms to Control Synthesis of Block Copolymers Containing PDMS - CO - PMAA Sequences, Molecular Crystals and Liquid Crystals, 416:1, 155-164

To link to this article: http://dx.doi.org/10.1080/15421400490478902

Mol. Cryst. Liq. Cryst., Vol. 416, pp. 155/[411]–164/[420], 2004

ISSN: 1542-1406 print/1563-5287 online
DOI: 10.1080/15421400490478902

Taylor & Francis
Taylor & Francis Group

APPLICATION OF *GENETIC ALGORITHMS* TO CONTROL SYNTHESIS OF BLOCK COPOLYMERS CONTAINING PDMS – CO - PMAA SEQUENCES

*Teodora Rusu**
"P. Poni" Institute of Macromolecular Chemistry, Iasi

Oana Marlena Gogan
ARTINFO SRL, Iasi, Roumanie branch of ARTI INFOMATICHE, Italy

In the last years a lot of scientist have reached to the conclusion that the Artificial Intelligence Methods can improve/facilitate the design of new macromolecules with desired/imposed properties.

This paper deals with the use of a method that combine a Genetic Algorithms (GA) based strategy with a Tabu Search Method (TS) in order to identify and optimize the molecular ratio between polydimethylsiolxane/poly(methacrylic acid) (PDMS/PMAA) sequences related to the water delivery properties of the final network.

Keywords: *Genetic Algorithms*; optimization; PDMS/PMAA; *Tabu Search Method*; water delivery proprieties

INTRODUCTION

One of the challenges in modern chemistry is the problem of designing new molecules with desired properties. The traditional approaches to this problem are usually expensive and time-consuming iterative process with the scientist or engineer hypothesizing a compound, synthesizing the material, testing for desired properties, and redesigning the candidate if the desired properties are not met. The use of *Artificial Intelligence Methods* (*Genetic Algorithms, Neuronal Network, Monte Carlo Methods, etc*.) has been tested in the last time and they have become a good opportunity to simplify the designee process of new macromolecular compounds with imposed properties [1–3].

*Corresponding author. Tel.: 40-232-217454, Fax: 40-232-211299, E-mail: teia@tuiasi.ro

This paper describes early results using genetic software techniques to automatically design macromolecules with imposed properties under the control of the fitness function. The fitness function must be capable of determining which of two arbitrary molecules is better for a specific task (water delivery test). For a experimental set of data, obtained form a group of copolymers containing polydimethylsiloxane (PDMS) and poly (methacrylic acid) (PMAA) sequences, the software begins by generating a new population of random molecules. The individual molecules in this population are then evolved the fitness function by randomly combining parts of the better existing molecules to create new macromolecules with the molecular ratio between the two sequences adapted to the best water delivery properties of the resulting network. These new molecules then replace some of the less fit molecules in the population. We apply a genetic crossover operator to molecules represented selected by a *Tabu Search Method*. The crossover alone, operating according to the *Tabu Search Method* restrictions, evolve any possible molecule given to an appropriate fitness function.

METHODS

As show by Venkat Venkatasubramanian [4] the designee of new macromolecular compounds deals with an evaluation schema as shown in Figure 1.

After a number of iterations the population consists of individuals that are well adapted in terms of the fitness function. The method used in this paper deals with a classical GA method to which the genetic crossover operator and the evaluation of the fitness function is made by an adapted *Tabu Search Method* (*TS*).

The *Tabu Search Method* is based on principles of intelligent search and it is an optimization method that uses a form of short-term memory to keep a search from becoming trapped in a local minima. A tabu list is formed that keeps track of recent solutions. At each iteration in the optimization process, solutions are checked against the tabu list. A solution that is on the list will not be chosen for the next iteration (unless it overrules its tabu condition by what is called an aspiration condition.) The tabu list forms the core of TS and keeps the process from cycling in one neighborhood of the solution space.

At each iteration, a steepest-descent solution that does not violate the tabu condition is selected. If no non-tabu improving solution exists, the best non-improving solution is taken. The combination of memory and gradient descent allows for diversification and intensification of the search. Local minima in the search space are avoided while good areas are well explored.

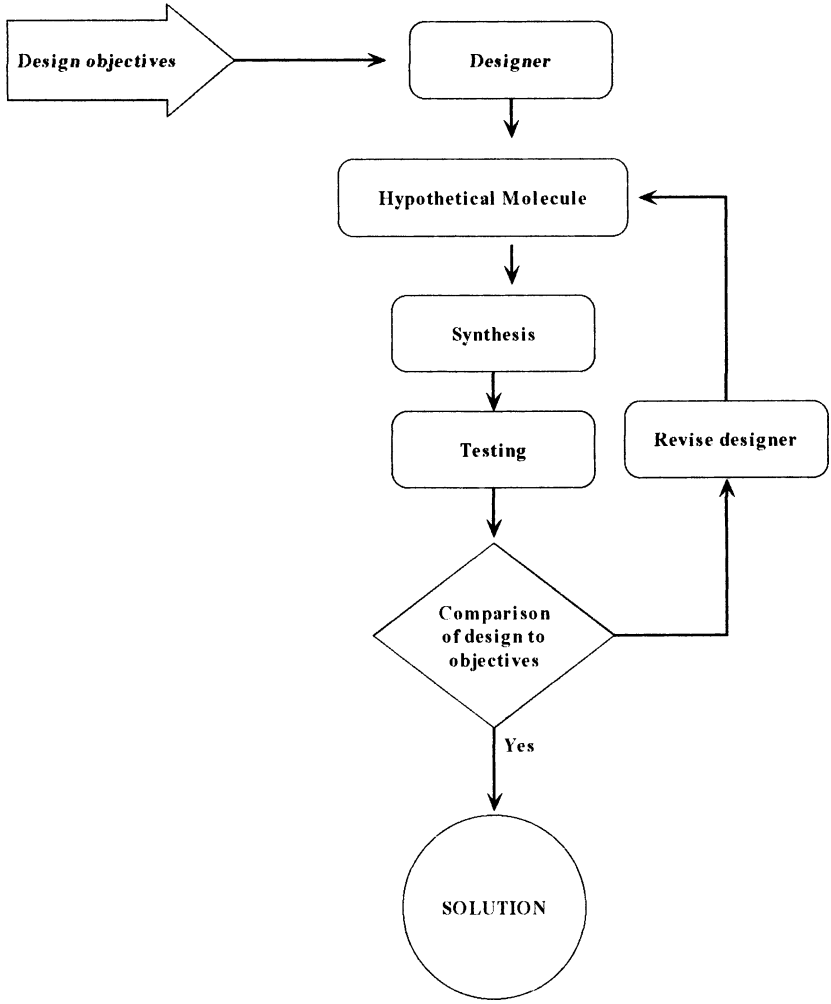


FIGURE 1 Designee of new macromolecular compounds using the *GA method*.

Two bits of pseudo-code presented in Schemes 1 and 2 show the basic of the TS method used in this paper. The first, in Scheme 1, is the overhead procedure. It controls the number of iterations, updating of the best solution so far, and controls the tabu list. The second, in Scheme 2, finds the next move in the search by a swapping procedure. All the possible swaps in the sequence are tried, and the best non- tabu swap is chosen. The routine shown finds the best move from the current location in the search space to a neighboring position. An alternative is to find the first

```
tabu_search ()
{
    for (i = 1; i <= # iterations; i++)
    {
        value = best_move ();
        make best_move;
        make best_move tabu ;
        if (value < global_best)
        {
            global_best = value
        }
    }
}
```

SCHEME 1 Basic code for the TS method.

location in the neighborhood that is an improvement over the current one. These two possibilities are termed *best improving* and *first improving* respectively.

```
best_move ()
{
    for (i = 1; i < n; i++)
    {
        for (j = i + 1; j <= n; j++)
        {
            swap (sequence[i,j]);
            value = evaluate (sequence);
            if (tabu [i][j] &&
                value > global_best)
            {
                continue;
            }
            if (value < best_so_far)
            {
                best_so_far = value;
                best_move = [i,j];
            }
        }
    }
    return best_so_far;
}
```

SCHEME 2 Basic code for the TS method.

The same evaluation function used with GAs, was used for TS. The first improving move scheme was employed.

RESULTS AND DISCUSSIONS

The initial experimental data were based on six selected PDMS – co – PMAA copolymers to whom the water delivery capacity was tested in relation with the molecular ratio between the hydrophilic/hydrophobic sequences.

Synthesis of Copolymers

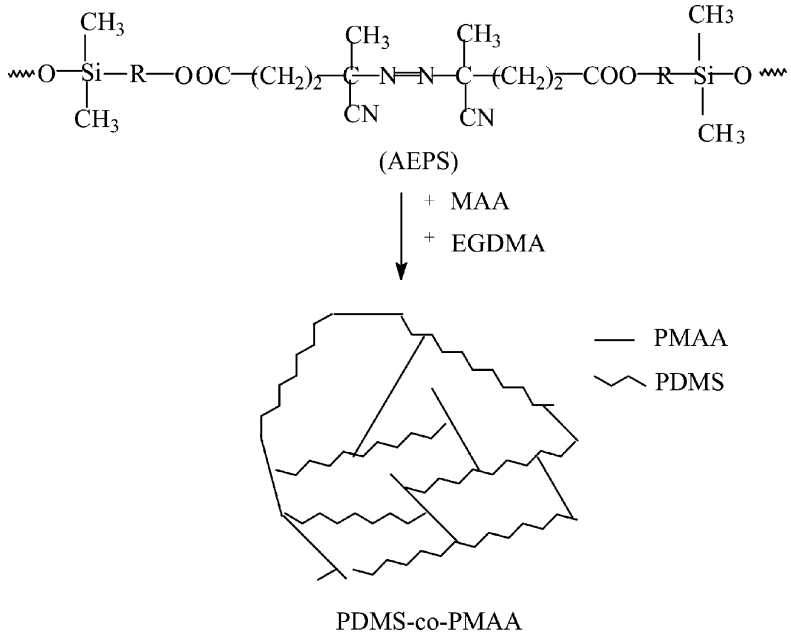
The synthesis of azoester containing polydimethylsiloxanes (AEPS) with different molecular weights of the siloxane sequences and different azo groups contents was realized according to a procedure previously described [5].

The synthesis of copolymers was achieved by the radical polymerization of methacrylic acid (MAA) in the presence of azoesther polydimethylsiloxane macroinitiators (AEPS) and of ethyleneglicole dimethacrylate (EGDMA) (1% mol. vs. MAA) as crosslinking agent. The reaction was carried out in toluene (total concentration 25%, by weight) under N_2, by maintaining the reaction mixture for 20 hours at 80°C in sealed ampoules. Control experiment of heating of a 25% MMA solution in toluene at 80°C for 20 hours showed that no pure thermal initiated polymerization of this monomer takes place in these conditions. The reaction scheme is presented in Scheme 3 and the synthesis data are summarized in Table 1.

To perform the water delivery test we have selected six copolymers samples as show in Table 2.

The water delivery tests were performed as follows: the same amount of samples from each of the six copolymers have been let to achieved the swelling equilibrium by keeping them in an excess of water for three days. After that the excess of water was removed by filtration, the swollen sample was submitted to control dryness by thermosetting the samples at 100°C and weighing them from two to two hours. Figure 2 presents the dependence of the percent of residual water from the samples as a function of the time and of the molar ratio between the two polymeric blocks.

The experimental data for the selected samples were used as parents in the GA method. Genetic algorithms allow computers to generate new, random data. When combined with a fitness function, only the best data are randomly exchanged and mutated until a new generation of data is produced. The first step, however, is to convert the genetic representation of the molecules into the molecular structure, similar to the way proteins



SCHEME 3 Synthesis scheme of PDMS – co - PMAA networks.

TABLE 1 Synthesis Data for PDMS – co - PMAA Obtained by the Radical Polymerization of MMA in the Presence of AEPS

	Initial mixture			Final mixture			
Mn_{-PDMS-}	SiO	AMA	SiO/AMA	SiO unreacted	MAA unreacted	η(%)	SiO/AMA final
5 040	0.897	1.794	0.50	0.888	1.794	85.9	0.495
5 040	0.928	4.641	0.20	0.816	4.641	91.5	0.183
5 040	1.213	1.214	1.00	0.951	1.214	78.3	0.783
5 040	0.959	0.377	2.00	0.674	0.377	89.4	1.788
6 620	0.928	1.857	0.50	0.813	1.857	87.6	0.438
6 620	1.288	6.442	0.20	1.005	6.442	77.9	0.156
6 620	1.161	1.161	1.00	1.098	1.161	94.6	0.946
6 620	1.031	0.515	2.00	0.897	0.515	87.1	1.742
14 060	1.029	2.058	0.50	0.895	2.058	86.9	0.435
14 060	1.387	6.934	0.20	1.109	6.934	80.2	0.160
14 060	1.203	1.203	1.00	0.908	1.203	75.5	0.755
14 060	0.849	0.425	2.00	0.768	0.425	90.4	1.808

TABLE 2 Selected Copolymers for Water Delivery Tests

Initial mixture			Network*		
Mn_{PDMS} (from AEPS)	SiO/MAA Molar ratio	Sample	Yield, %	SiO/MAA** Molar ratio	Aspect
5040	0.5	I.1	85.9	0.42	Gel
6620	0.5	I.2	87.6	0.44	Gel
14060	0.5	I.3	86.9	0.43	Gel
5040	2.0	II.1	89.4	1.63	Gel
6620	2.0	II.2	87.1	1.74	Dense gel
14060	2.0	II.3	90.4	1.81	Dense gel

*Polymerization in sealed ampoules; 80°C; 20 hours; solvent, toluene (total concentration 25%); EGDMA, 1% molar against MAA.

**Determined from elemental analysis (Si content).

are made based on DNA structure [6]. Once the molecules are built, the fitness function is applied. The fitness function must be very robust since the randomly generated initial molecules rarely make chemical sense.

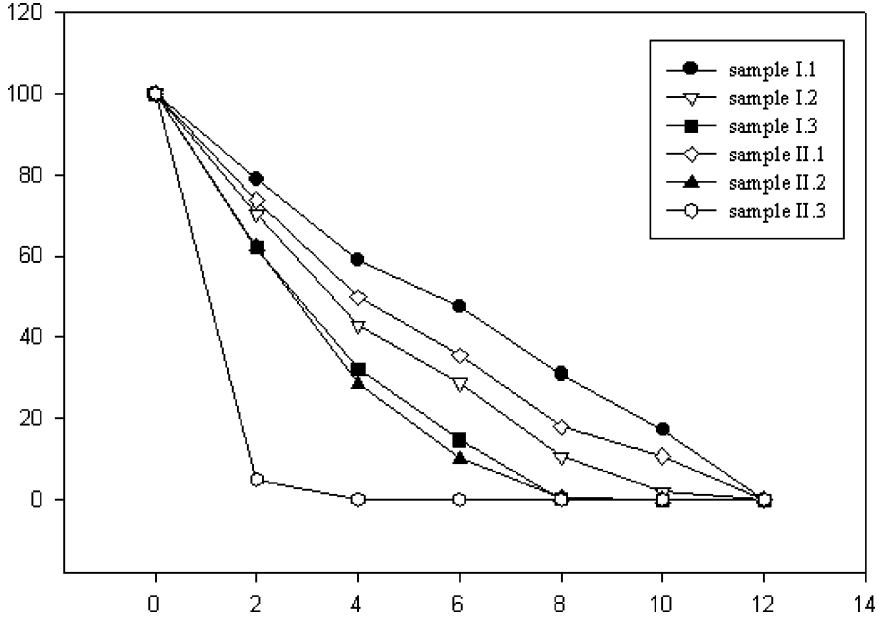


FIGURE 2 Water delivery tests for the selected PDMS – co – PMAA samples.

Table 3 summarize the schemata dates used in the GA for three parents selected as possessing best water delivery capacity from the tested copolymers. Four new initial individuals were randomly generated. The bit strings of the individuals are decoded to unsigned integer values. The fitness function f $(i) = i^2$ is used to assign a fitness value to each individual in Table 3. Depending on their relative fitness values the reproduction probability (between 0% and 100%) for each individual is calculated and converted into the number of expected successors. Then the *Tabu Search algorithm* is used to perform the selection based on the reproduction probability. Table 4 presents the reproduction performance for the given system.

The individuals selected for reproduction have been replicated according to their relative fitness function. Crossover sites and mating partners have been assigned randomly. After performing crossover the new fitness values of the individuals in the new population was calculated.

Both *GA*s and *TS* worked in order to find the best fit solutions to the design problem. Genetic operators emphasizing order converged faster than those emphasizing adjacency. This evidence does however shed additional light on the nature of the instruction scheduling process by providing more controlled, empirical evidence. The ordering of the placement of nodes by the genetic algorithm mirrors the approach used by human coders. The nodes with the greatest impact on final schedule length are placed first, with those having lesser impact placed later. The order of placement that ensures validity is also reflected. Figure 3 presents the performance of the proposed method in relation to the fitness function and the local twist.

TABLE 3 Genetic Algorithm Data

Mn_{PDMS}	Schemata code	Pattern	Individual	Average schemata fitness
5040	I.1	00###	3	4
6620	I.2	1####	2.4	558.5
14060	I.3	#1#1#	1	100

Individual	Bit string	Integer value	Fitness	Reproduction probability $\frac{f(i)}{\sum f}$	Expected count $\frac{f(i)}{\bar{f}}$	Actual count (TS)
1	01010	10	100	8.2%	0.33	1
2	10101	21	441	36.1%	1.45	1
3	00010	2	4	0.3%	0.01	0
4	11010	26	676	55.4%	2.22	2
Sum			1221	100.0%	4	4
Average			305.25			
Max			676			

TABLE 4 Reproduction Parameters

Mating pool after reproduction with crossover site	Mating parameter	Crossover site	New population	Integer value	Fitness value $f(i) = i^2$
010|10	3	3	01010	10	100
10|101	4	2	10010	18	324
110|10	1	3	11010	26	676
11|010	2	2	11101	29	841
Sum					1941
Average					485.25
Max					841

Schemata	After reproduction		After crossover		
Expected count $\frac{\sum f(H)}{f}$	Actual count	Individual	New count expected	Actual count	Individual
0.01	0	–	0.00	0	–
1.83	3	2,3,4	2.54	3	2,3,4
0.33	3	1,3,4	1.60	2	1,3

In this study, both the swap and insert operators demonstrated very similar behavior, pointing to the fact that absolute order is not of ultimate importance for this particular problem. The proposed method was able to avoid local minima and find competitive solutions.

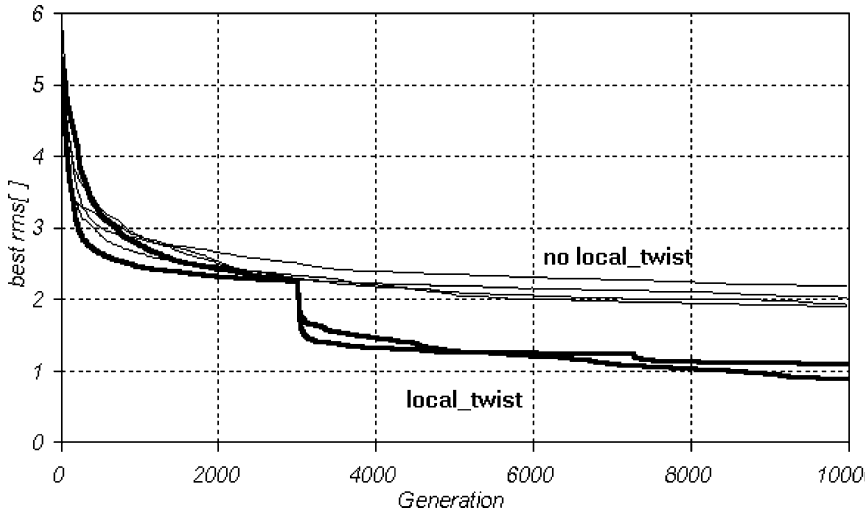


FIGURE 3 Performance for the GA with TS operator method.

CONCLUSION

In this paper we have presented a hybrid method to design macromolecular compounds with imposed properties. The method described here is composed by a classical *Genetic Algorithm*, connected with a *Tabu Search Algorithm*, used to define the crossover operator and estimate the fitness function.

Tabu search operator leads to a better solutions than the best genetic operators. The key advantage of genetic algorithms is that they not only use the computer as a tool for modeling and understanding properties, but also enable the computer to develop new structures and determine if those structures serve a specific purpose. This allows discoveries of a random, almost accidental nature. Genetic algorithms show the potential for the discovery of new structures able to perform imposed tasks. This advantage may increase the development of chemistry in the coming decades.

REFERENCES

[1] Holland, J. H. (1992). *Adaptation in Natural and Artificial Systems*, 2nd Ed., MIT Press.
[2] Davis, L. (Ed.), (1991). *Handbook of Genetic Algorithms*, Van Nostrand Reinhold: New York.
[3] Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley.
[4] Venkatsubramanian, V. *et al*. (1996). Computer-Aided Molecular Design Using Neural Networks and Genetic Algorithms. In: *Genetic Algorithms in Molecular Modeling*, Devillers, J. (Ed.), Academic Press: New York.
[5] Rusu, T., Ioan, S., & Buraga, S. C. (2001). Amphyphylic copolymers as water delivery systems. *Eur. Polym. J.*, *37*, 2005.
[6] Rusu, T. & Gogan, O. M. (2002). Méthodes d'utilisation des Algorithmes Génétiques dans la génération assistée sur ordinateur des composes macromoléculaires. *Proceedings of CoFrRoCA*, 129–132.