Supporting Information

SYNTHESIS AND THEORETICAL STUDIES OF THE CONFORMATIONAL BEHAVIOUR OF N-VINYLCAPROLACTAM/N-VINYLIMIDAZOLE COPOLYMERS IN SELECTIVE SOLVENT

A. I. Barabanova[1], A. V. Vorozheykina[1], M. K. Glagolev[1*], P. V. Komarov[1,2*], A. R. Khokhlov[1,3]

[1]Institute of Organoelement Compounds RAS, Vavilova st., 28, 119991, Moscow, Russia
[2]General Physics Department, Tver State University, Sadovy st., 35, 170002, Tver, Russia
[3]Faculty of Physics, M. V. Lomonosov Moscow State University, Leninskie Gory, 119991, Moscow, Russia
** E-mail: mikhail.glagolev@gmail.com (M.K.G.); pv_komarov@mail.ru (P.V.K.)*

Here we present auxiliary information and a set of additional data not included in the main article.

S1. The composition and molecular-mass characteristics of VCL-VI copolymers

The composition of the copolymers was estimated by ^{1}H NMR spectroscopy. The ^{1}H NMR spectra were recorded using a Bruker Avance 600 spectrometer at 600.22 MHz (1H) with D_2O as solvent. The accuracy of the chemical shift determination was at least 0.001 ppm.

^{1}H NMR spectra of copolymers VCL-VI with various compositions are shown in Figures S1-S6. The singlet (1) at 4.46 - 3.97 ppm and the multiplet (2, 5 and 6) at 2,00 -1,00 ppm characterise one proton of the -CH group of the VCL units of the copolymer backbone, four protons of the $-CH_2$ groups of the backbone of the VCL and VI units and six protons of the $-CH_2$ groups in the caprolactam ring, respectively (Figure S1). The signal (3) at 2.48 - 1.92 ppm ($-COCH_2-$) and the signals (4 and 7) between 3.95 - 3.36 and 3.35 - 2.55 ppm ($-NCH_2$, $-NCCH_2CH_2CH_2$) were attributed to the caprolactam ring and the -CH group of the VI units in the copolymer backbone. The multiplet (8, 9 and 10) at 7.75 - 6.46 ppm due to the protons of the imidazole ring (Figure S1).

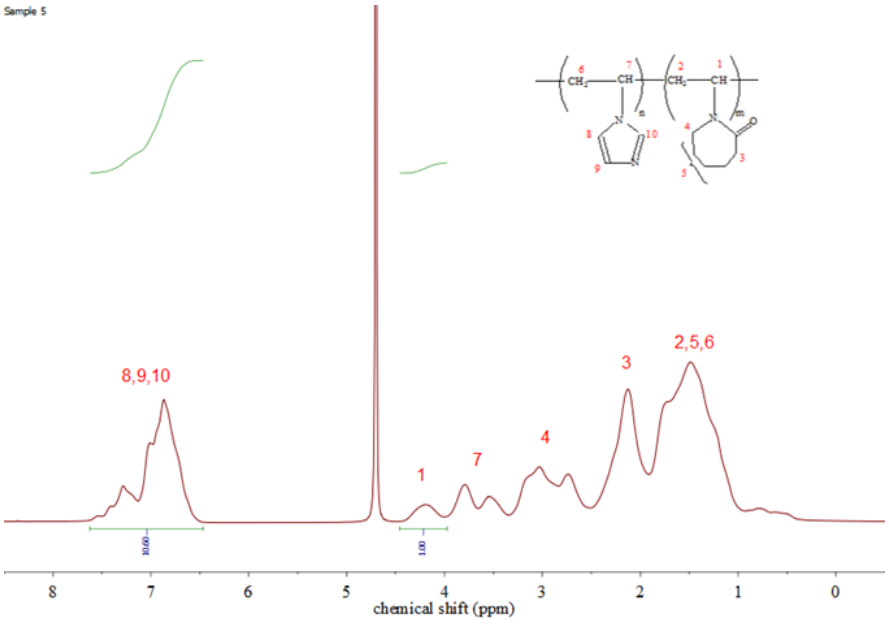

Figure S1. ^{1}H NMR-spectrum of sample 5. D_2O.

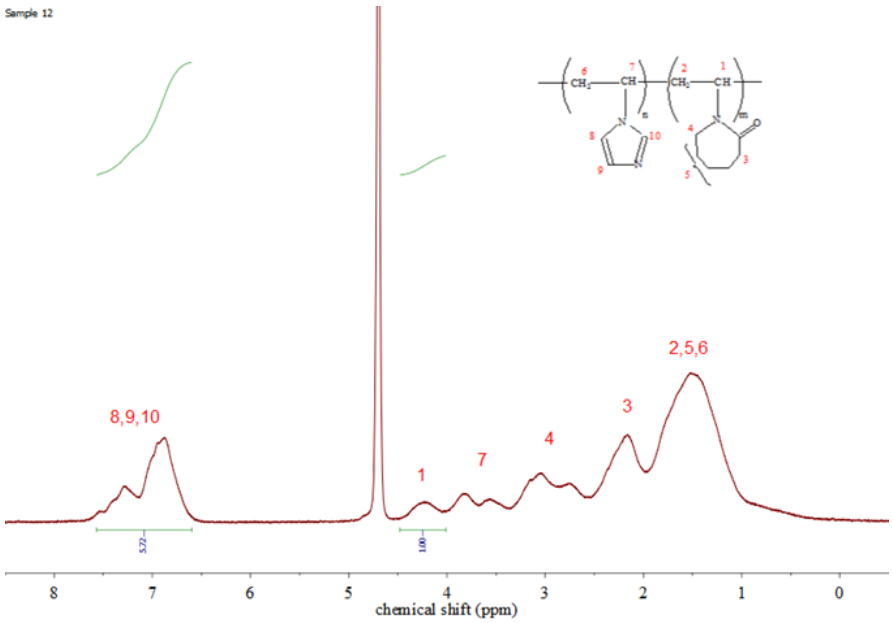

Figure S2. ^{1}H NMR-spectrum of sample 12. D_2O.

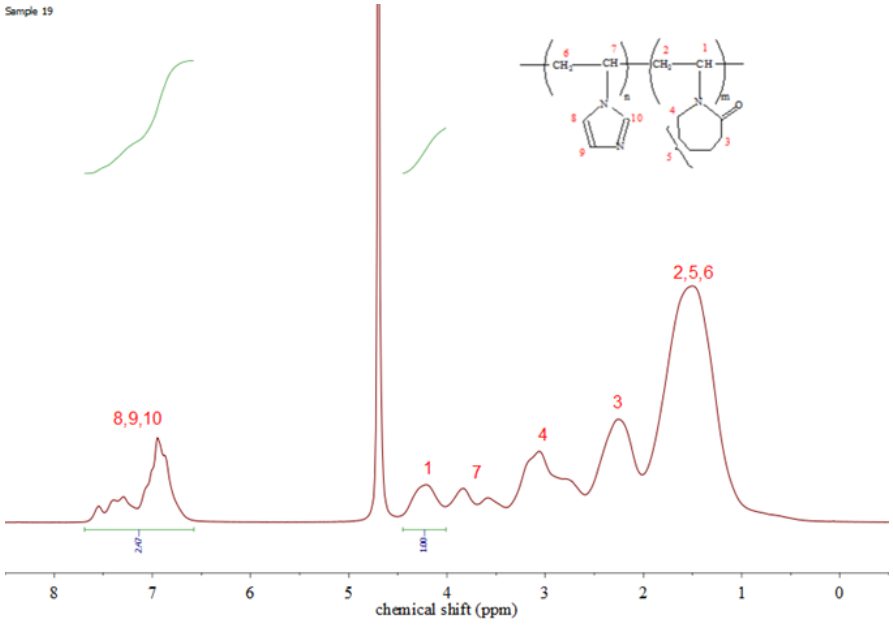

Figure S3. ^{1}H NMR-spectrum of sample 19. D_2O.

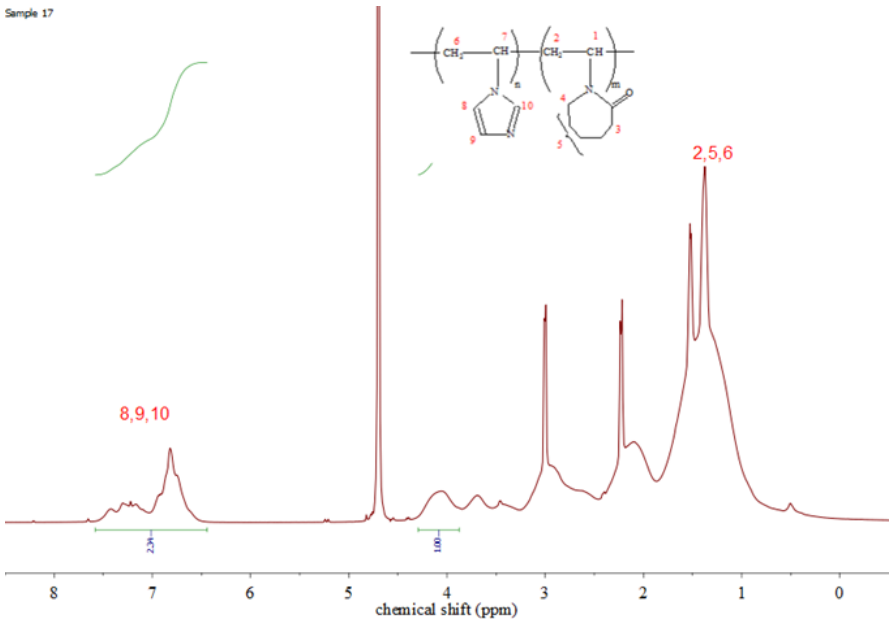

Figure S4. ^{1}H NMR-spectrum of sample 17. D_2O.

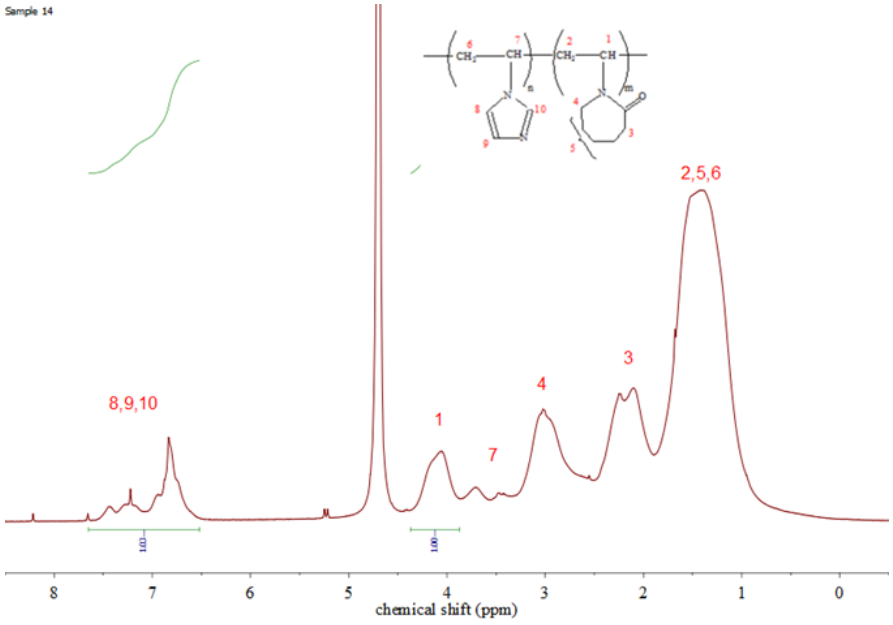

Figure S5. ^{1}H NMR-spectrum of sample 14. D_2O.

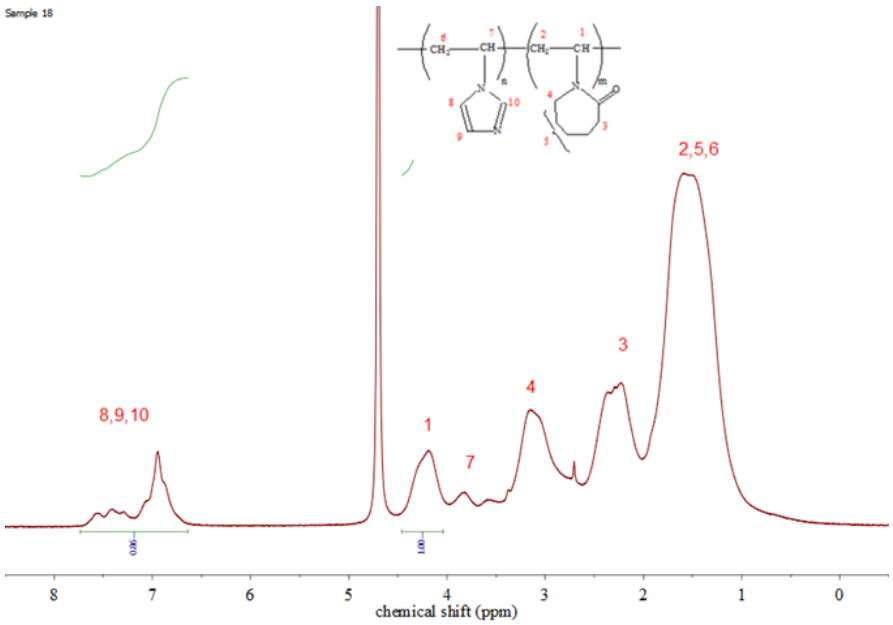

Figure S6. ^{1}H NMR-spectrum of sample 18. D_2O.

The copolymer compositions (mol%) were determined from the integral intensity of the signals of three protons of the imidazole ring in the VI units and one proton of the -CH group in where H_{VI} and H_{VCL} are the integral intensities of the signals of three protons of the imidazole group in the VI units and one -CH- proton in the VCL units. the VCL units:

$$[VI] = \frac{H_{VI}}{3H_{VCL} + H_{VI}} \times 100 \quad , \quad [VCL] = \frac{H_{VCL}}{\frac{H_{VI}}{3} + H_{VCL}} \times 100$$

The average molecular masses and molecular mass distribution of the copolymers were determined by GPC with polystyrene standards using an Agilent 1200 chromatograph equipped with a refractometric detector and a PLmixC column. A 0.03 M solution of LiCl in N-methylpyrrolidone was used (flow rate 0.5 mL/min, temperature 50°C).

The copolymerisation conditions and GPC data are summarised in Table 1.

Table 1. Composition and molecular-mass characteristics of the copolymers.

Sample	Comonomers feed composition, mol. %		q, %	Copolymer composition, mol %*		GPC	
	VCL	VI		VCL	VI	$M_n \times 10^{-3}$, g/mol	Đ
1	55	45	3	20.8	79.2	107	1.63
2			2.7	21.4	78.6	111	1.59
3			4.9	22.4	77.6	128	1.76

4			1.1	23.2	76.8	92	1.59
5			8.5	22.0	78.0	141	1.75
6			16	21.5	78.5	-	-
7			26	25	75	115	3.17
8	70	30	2.4	41	59	125	2.10
9			1.8	39	61	74	2.20
10			16.9	39.4	60.6	214	2.35
11			42.7	41	59	252	3.35
12			1.9	35	65	124	2.02
13	85	15	8.7	57.3	42.7	153	2.83
14			52	74.8	25.2	370	3.53
15			1.4	58	42	195	2.34
16			4.3	59.9	40.1	184	2.00
17			22.9	56	44	300	2.34
18			90.1	78	22	310	3.38
19			6.9	55	45	217	2.12

S2. Computer modelling methods used

We use established methods of computer simulation that are most appropriate for the purposes of our study. The kinetic Monte Carlo method was chosen to simulate the in silico copolymerisation of N-vinylcaprolactam and N-vinylimidazole units. Its theoretical basis is well described in the literature [S1–S3]. For the design of our simulation program, we used a Monte Carlo approach from Ref. [S4]. Then, to predict the behaviour of the virtually synthesised poly-VCL-b-VCL-VI copolymer chains, a widely used coarse-grained approach based on Langevin dynamics was used [S5–S7]. This method is a mesoscale modelling technique that allows one to quickly predict the conformational behaviour of macromolecules based on the fundamental ideas of their structure and intermolecular interactions. We have already used this method to study several molecular systems [S8,S9].

S2.1. In silico synthesis protocol using kinetic Monte-Carlo

The algorithm of the simulated copolymerization was designed using the standard kinetic Monte-Carlo approach. It was implemented using Python3 programming language. The flowchart of the algorithm is shown in Figure S7. The Mersenne Twister [S10] algorithm from the Python3 built-in random module was used for generating all of the random numbers used.

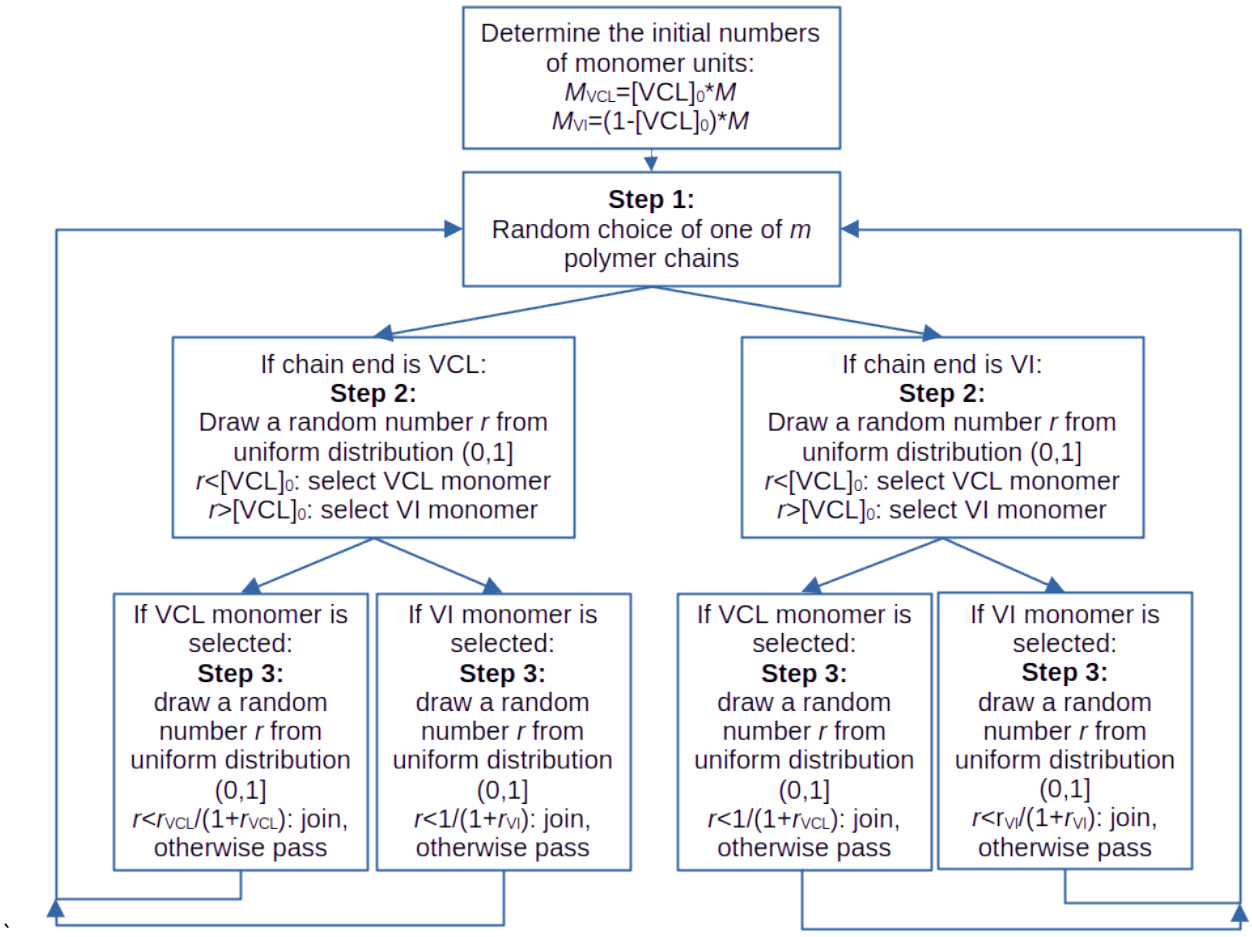

Figure S7. Flowchart of the kinetic Monte-Carlo model of virtual synthesis of VCL-VI copolymer under the assumption of constant local concentrations of VCL and VI units in the vicinity of the reaction centres.

The input parameters of the synthesis simulation included the total number M of monomer units in the virtual reactor, the number m of chains that were grown in parallel, the molar fractions of VCL and VI monomer units, $[VCL]_0$ and $1-[VCL]_0$, respectively, and the copolymerization constants r_{VCL} and r_{VI}.

Before the start of the chain growth cycle, the total numbers of the VCL and VI monomer units in the implied reactor were calculated as $M\times[VCL]_0$ and $M\times(1-[VCL]_0)$. The chains had their initial monomer units distributed proportionally to the fractions of the units in the simulated reaction mixture: $m\times[VCL]_0$ chains had a VCL monomer as the first unit, while $m\times(1-[VCL]_0)$ chains had a VI monomer as the first unit.

After the initialization, the chain growth loop was run until either all of the monomer units in the reactor were exhausted or 5,000,000 iterations were reached. At the first step of each iteration, one of the m growing chains was randomly selected with equal probability. At the second step, the type of the monomer unit trying to join the chain was selected by generating a random number in the range [0,1). If the random number was less than $[VCL]_0$, the VCL monomer was selected; otherwise, the VI monomer was selected. This corresponded to the condition of constant local concentrations of the comonomers in the reaction zone. Depending on the type of monomer unit at the end of the growing chain and the type of monomer fetched from the implied reactor, the joining of the comonomer unit to the selected chain took place with the following probabilities: VCL to VCL: $r_{VCL}/(1+r_{VCL})$; VI to VCL: $1/(1+r_{VCL})$; VCL to VI: $1/(1+r_{VI})$; VI to VI: $r_{VI}(1+r_{VI})$. The experimentally determined values of r_{VCL}=0.19 and r_{VI}=5.21 were used in the simulation. If the joining took place, the joined comonomer was removed from the implied reactor. When one of the monomers in the implied reactor had been fully consumed, only the joining of the remaining monomer was simulated. When all comonomers have been consumed or 5,000,000 Monte-Carlo steps have been reached, the reaction was terminated.

S2.1. Simulation protocol using Langevin dynamics

The first five monomer unit sequences generated by the MC model for each monomer feed composition were used to generate initial chain conformations by a self-avoiding random walk using the create_configuration.py tool from the MOUSE2 package [S11].

The behaviour of the chains in the selective solvent was studied using the simulation protocol established in our earlier studies of solvophobic [S8] and amphiphilic copolymers [S9]. Each of the generated chain conformations was placed in a cubic simulation cell with an edge length of a = 1000 (infinitely dilute solution) to study their evolution in the implicit solvent. All simulations were performed using Langevin dynamics within the LAMMPS simulation package [S12] according to the simulation protocol described in [S9].

The excluded volume interactions were represented by the repulsive part of the Lennard-Jones potential:

$$U_{ev}=4\varepsilon((\sigma/r)^{12}-(\sigma/r)^{6})+1\ ,\ r<r_c\ , \qquad (S1)$$

here the cut-off radius $r_c = 2^{1/6}$.

The hydrophobic interactions between the VCL monomer units, induced by the implicit solvent, were represented by the Yukawa-type potential:

$$U_{solv}=A(\exp(-kr_{ij})/r_{ij}-\exp(-kr_{solv})/r_{solv}),\ r<r_{solv}\ , \qquad (S2)$$

where the cut-off radius r_{solv}=4, the parameter k = 0.23 and the factor A was 0 during initial relaxations for 40,000,000 time steps and then decreased by ΔA = -0.01 every 2,000,000 time steps until the value of A = -1 was reached. The temperature of the system T = 1 was maintained by the Langevin thermostat.

S3. Properties of obtained nanostructures

To analyse the globular configurations formed in the solvent, which is poor for the VCL units, we calculated the radius of gyration and average block lengths for the VCL and VI monomer units.

1. The radius of gyration is calculated using the following equation:

$$R_g = (1/N \cdot \sum(r_i - r_{c.m.})^2)^{1/2}, \quad \text{(S3)}$$

where r_i is the position of the i-th bead and $r_{c.m.}$ is the position of the centre of mass. The radii of gyration were also calculated taking into account only VCL or VI monomer units:

$$R_{g,VCL} = (1/N \cdot \sum(r_{i,VCL} - r_{c.m.,VCL})^2)^{1/2}, \quad \text{(S4)}$$

$$R_{g,VI} = (1/N \cdot \sum(r_{i,VI} - r_{c.m.,VI})^2)^{1/2} \quad \text{(S5)}$$

2. Average block lengths for VCL and VI monomer units were calculated excluding the VCL tail of the copolymer chain:

$$<b_{VCL}> = (\sum b_{VCL})/n_{blocks,VCL}, \quad \text{(S6)}$$

$$<b_{VI}> = (\sum b_{VI})/n_{blocks,VI}, \quad \text{(S7)}$$

where b_{VCL} and b_{VI} are the lengths of the continuous VCL and VI blocks in the part of the polymer chain excluding the long VCL block formed after consumption of the VI monomers from the monomer feed, and $n_{blocks,VCL}$ and $n_{blocks,VI}$ are the respective numbers of these blocks.

The numerical simulation data presented in the manuscript were averaged over all five statistically independent samples.

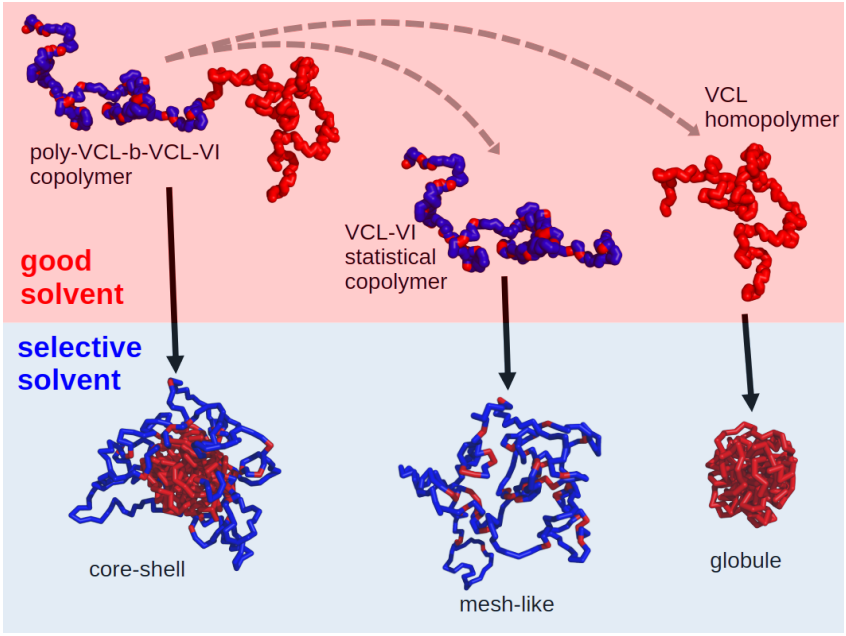

Figure S8. Conformations of the VCL-VI copolymer and its constituent parts in good (top row) and poor (bottom row) solvents. The whole copolymer synthesised according to the procedure is shown on the left, the copolymer block with the VCL tail removed is shown in the centre, and the isolated VCL tail is shown on the right.

References

S1. J. J. González and P. C. Hemmer, *The Journal of Chemical Physics*, 1977, **67**, 2496–2508.

S2. K. Platkowski and K.-H. Reichert, *Polymer*, 1999, **40**, 1057–1066.

S3. R. Szymanski and S. Sosnowski, *Macro Theory & Simulations*, 2018, **27**, 1800015.

S4. A. A. Gavrilov and A. V. Chertovich, *Macromolecules*, 2017, **50**, 4677–4685.

S5. P. G. Khalatur, in *Polymer Science: A Comprehensive Reference*, Elsevier, 2012, pp. 417–460.

S6. E. Paquet and H. L. Viktor, *BioMed Research International*, 2015, **2015**, 1–18.

S7. A. Liwo, C. Czaplewski, A. K. Sieradzan, A. G. Lipska, S. A. Samsonov and R. K. Murarka, *Biomolecules*, 2021, **11**, 1347.

S8. M. K. Glagolev and V. V. Vasilevskaya, *Polym. Sci. Ser. C*, 2018, **60**, 39–47.

S9. A. A. Abramova, M. K. Glagolev and V. V. Vasilevskaya, *Polymer*, 2022, **253**, 124974.

S10. M. Matsumoto and T. Nishimura, *ACM Trans. Model. Comput. Simul.*, 1998, **8**, 3–30.

S11. M. K. Glagolev, A. A. Glagoleva, V. V. Vasilevskaya, *Journal of Supercomputing Frontiers and Innovations*, 2023, **10**, 73-87

S12. A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. In 'T Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott and S. J. Plimpton, *Computer Physics Communications*, 2022, **271**, 108171.