

Sequence Constraints in a Glycine–Lactic Acid Dopolymer Determined by Matrix-assisted Laser Desorption/Ionization Mass Spectrometry

Maurizio S. Montaudo*

Istituto per la Chimica e la Tecnologia dei Materiali Polimerici, Consiglio Nazionale delle Ricerche. Viale A. Doria, 6 - 95125 Catania, Italy

Received 4 January 1999; Revised 18 January 1999; Accepted 10 January 1999

The matrix-assisted laser desorption/ionization (MALDI) time-of-flight (TOF) mass spectrum was recorded for a copolymer sample obtained by ring-opening copolymerization of dilactide and 6-methyl-2,5-morpholinedione (the cyclic glycine-lactic dimer). MALDI-TOF mass spectrometry (MS) has been used for the analysis of an unfractionated copolymer sample whereas previous studies by MALDI-TOF MS have solely concerned narrow fractions of copolymer. The sequence distribution which results from the copolymerization of the glycine–lactic copolymer is such that a sequence with two consecutive glycine units cannot be formed. The experimental mass spectrometric intensities were compared with the predicted intensities using a model for the sequence distribution (Markoff model) which has no constraints, it allows for sequences with two consecutive glycine units. The copolymer sequence is determined by least-squares best-fit minimization. The composition determined by MALDI-TOF MS is in good agreement with NMR results. Furthermore, the resulting distribution is Markoffian and the probability matrix elements associated with it come quite close to those predicted. In fact, the synthetic route imposes a sequence constraint on the copolymer chain so that two consecutive glycine units are forbidden.

The determination of copolymer sequences can be achieved by mass spectrometry (MS) and nuclear magnetic resonance (NMR) since the intensity of the peaks which appear in each reflect the sequence of the copolymer.^{1–18} The use of NMR for sequence determination is quite common.¹ However, in order to use NMR the peaks which appear at different chemical shifts must be reliably assigned and this represents an error-prone task which may produce an incorrect sequence determination.

In order to avoid such potential limitations, MS can be used. The problem of decoding the intensity of peaks appearing in the mass spectra of copolymers and of relating them to the comonomers sequence has already been studied, providing a method for deducing the sequence distributions and composition of comonomers in copolymers.^{2–18}

Chain statistics (Bernoullian, first order Markoffian)

allow any arrangement of comonomer units along the chain to be generated. Starting from any sequence, a theoretical mass spectrum can be obtained, based on the assignment of each spectroscopic peak to a set of sequential arrangements of monomers.

Looking at copolymers by MS, the relative abundance of all the co-oligomers of a defined chain length reflects the composition and sequence present in the copolymer. This means that the possibility exists to build a theoretical mass spectrum for any given copolymer sequence, which can be compared with the experimental mass spectrum corresponding to the copolymer sample under investigation (MACO4).²³

Since each series of oligomers (dimers, trimers, etc.) allows the calculation of the copolymer composition and sequence distribution, and since the mass spectra often provide information up to decamers and beyond, MS provides an excellent method to evaluate these two quantities with good precision.

Other quantities that can be determined from the mass spectral intensities in copolymers are: average length of long blocks of like monomers, degree of randomness, reactivity ratios and the compositional heterogeneity index, which indicates if the composition changes as the length of the macromolecular chain grows.^{2–18}

Recently, several homo- and co-polydepsipeptides have been synthesized and these materials are currently under consideration as biodegradable materials.^{20–25} Ring-opening copolymerization of dilactide and 6-methyl-2,5-morpholinedione (the cyclic glycinelactic acid dimer) affords^{23–25} a series of co-polydepsipeptides corresponding to Scheme 1.

The synthetic scheme (Scheme 1) yields a quasi-random copolymer with a peculiar sequence distribution, since it deviates from the standard form in which the two repeat units (glycine and lactic acid residues) are found at random along the copolymer chain. The deviation is due to chains having sequences containing two consecutive glycine units, which are not produced. This deviation is small and is therefore difficult to detect.

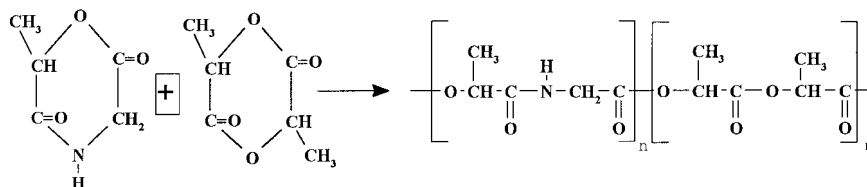
In this paper, the MALDI-TOF mass spectrum of a sample of glycine(24)–lactic acid(76) copolymer is reported which was obtained using the reaction depicted in Scheme 1. Sequence information was extracted from mass spectral peak intensities.²³ It is demonstrated that the MACO4 method²³ is able to detect subtle deviations within a sequence in which lactic and glycine residues alternate at random along the chain.

*Correspondence to: M. S. Montaudo, Istituto per la Chimica e la Tecnologica dei Materiali Polimerici, Consiglio Nazionale delle Ricerche, Viale A. Doria, 6-95125 Catania, Italy.

Contract/grant sponsor: Italian Ministry for University and Scientific and Technological Research (MURST).

Contract/grant sponsor: Progetto Finalizzato Materiali Speciali per Tecnologie Avanzate II (CNR, Rome).

Contract/grant sponsor: National Council of Research (CNR, Rome).



COPOLYMER SEQUENCE

According to chain statistics, the sequence distribution of a two-component copolymer is completely defined when the four elements P_{GG} , P_{GL} , P_{LG} , P_{LL} of the probability matrix are determined. Matrix elements are related by the following conditions:

$$P_{GG} + P_{GL} = 1 \quad P_{LG} + P_{LL} = 1 \quad (1)$$

The mass spectral intensity, $I(L_m G_n)$, of the oligomer $L_m G_n$ is given by:²³

$$I(L_m G_n) = \sum \phi (P_{GG})^{n1} (P_{GL})^{n2} (P_{LG})^{n3} (P_{LL})^{n4} / (P_{GL} + P_{LG}) \quad (2)$$

where ϕ is an integer number and Σ indicates summation over $n1$, $n2$, $n3$ and $n4$. The above equation can be used to compute the molar fraction, $I(L)$ and $I(G)$, of the two monomers in the copolymer:

$$I(L) = \frac{P_{GL}}{P_{GL} + P_{LG}} \quad I(G) = \frac{P_{LG}}{P_{GL} + P_{LG}} \quad (3)$$

The number average length of long LLLLLL blocks, n_L , and

the number average length of long GGGGG blocks, n_G , can be computed using the formula:

$$n_L = 1/P_{LG} \quad n_G = 1/P_{GL} \quad (4)$$

The synthetic route (Scheme 1) adopted to produce the copolymer is such that oligomers having two consecutive glycine units are not present. Thus

$$P_{GG} = 0 \quad P_{GL} = 1 \quad (5)$$

This implies that oligomers $L_m G_n$ with $m > n$ cannot be formed.

EXPERIMENTAL

Materials

The copolymer sample of 6-methyl-2,5-morpholinedione and lactide analyzed here (copolymer **1**) was provided by Dr. P. J. Dijkstra (Twente University of Technology, The Netherlands). The molar fraction of lactide in the feed was 0.50. ¹H-NMR data indicate that the molar fraction of glycine units in the copolymer is 0.24.

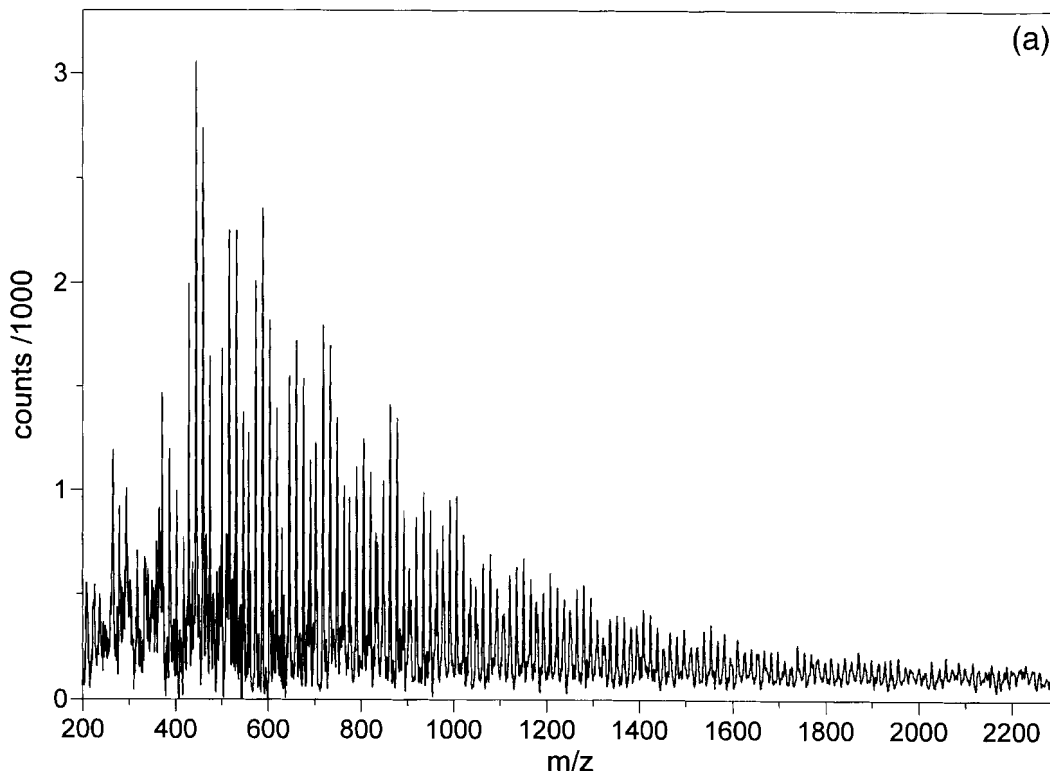


Figure 1a. MALDI-TOF mass spectrum of the copolymer sample.

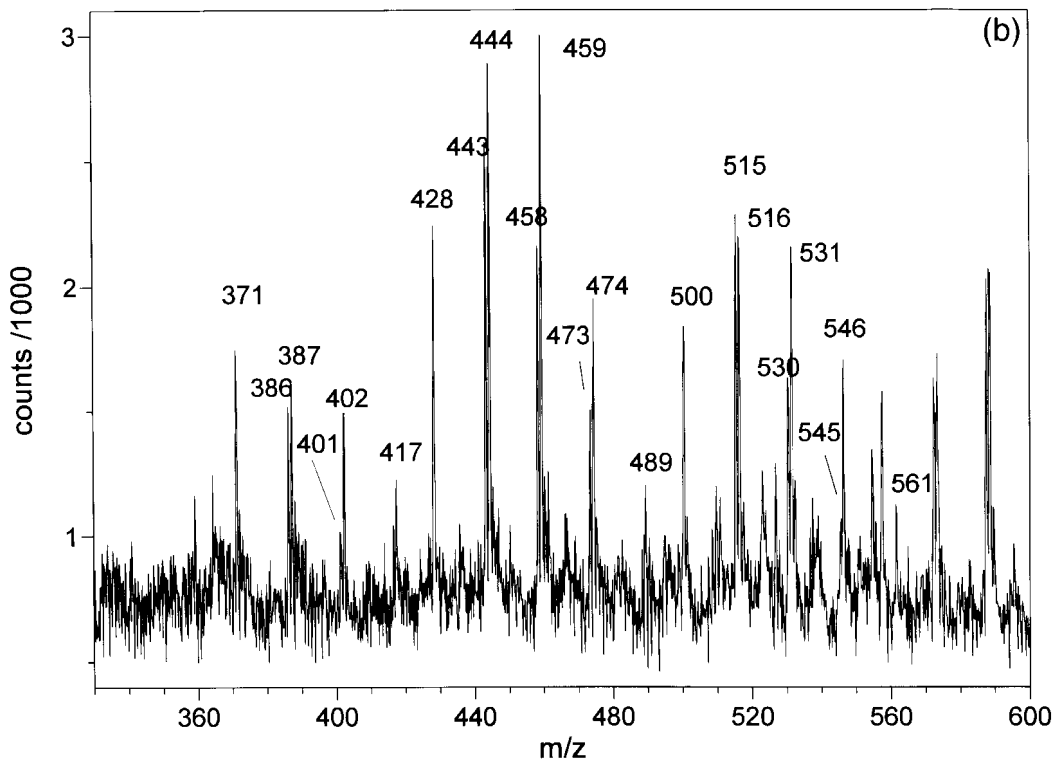


Figure 1b. Expansion of the MALDI-TOF mass spectrum of the copolymer sample in the region of pentamers, hexamers and heptamers.

MALDI-TOF-MS

A Bruker Reflex mass spectrometer (Billerica, MA, USA) was used to obtain the MALDI-TOF mass spectra. The

spectrometer was equipped with a nitrogen laser (337 nm, 5 ns) (LSI, Newton, MA, USA), a transient recorder (time base 4 ns) and a dual-plate microchannel detector. The accelerating and the reflectron voltages were 15000 and

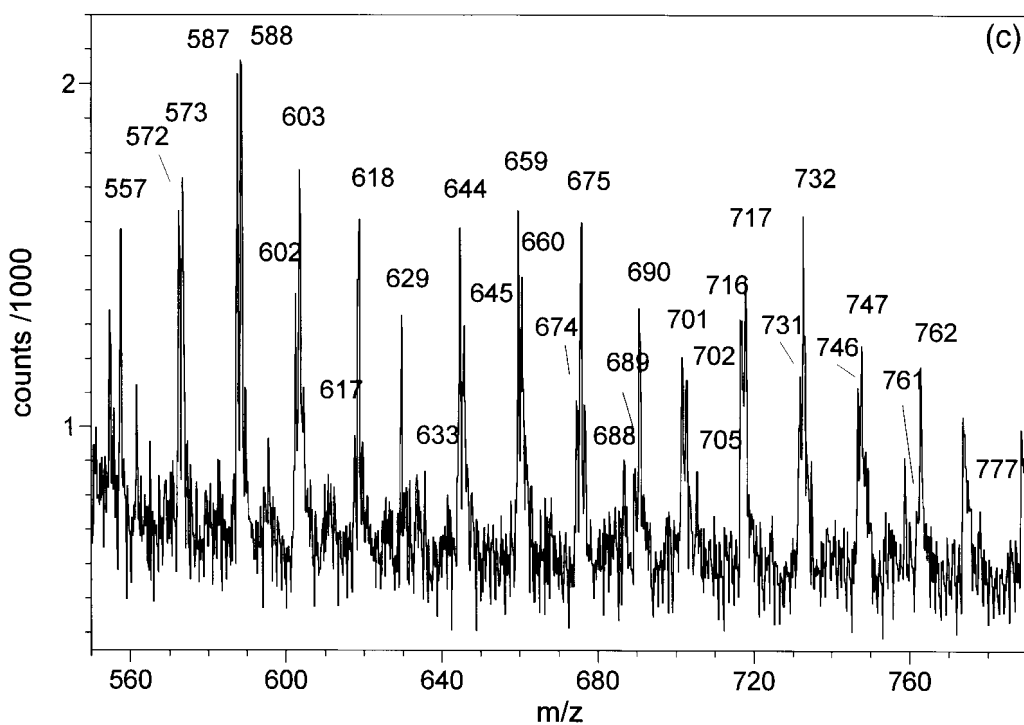


Figure 1c. Expansion of the MALDI-TOF mass spectrum of the copolymer sample in the region of octamers, nonamers and decamers.

Table 1. Assignments of mass spectral peaks appearing in the MALDI-TOF spectrum shown in Fig. 1(a)

Assign. ^a	Mass ^b	I _{exp.} ^c	I _{calc.} ^d								
L3G2Na+	371	410	241	L6G5K+	774	140	100	L10G5Na+	1046	52	43
L4G1Na+	386	359	404	L10K+	777	66	102	L9G6K+	1047	34	41
L3G2K+	387	461	303	L8G3Na+	788	216	228	L14Na+	1049	9	10
L5Na+	401	243	270	L7G4K+	789	186	211	L13G1K+	1050	22	49
L4G1K+	402	372	448	L9G2Na+	803	202	254	L11G4Na+	1061	60	66
L5K+	417	154	270	L8G3K+	804	190	315	L10G5K+	1062	45	73
L3G3Na+	428	950	291	L6G6Na+	815	138	27	L14K+	1065	9	10
L4G2Na+	443	999	732	L10G1Na+	818	101	170	L8G8Na+	1073	45	5
L3G3K+	444	680	420	L9G2K+	819	191	314	L12G3Na+	1076	60	74
L5G1Na+	458	551	980	L7G5Na+	830	172	79	L11G4K+	1077	52	101
L4G2K+	459	820	915	L6G6K+	831	78	54	L9G7Na+	1088	60	16
L6Na+	473	238	547	L11Na+	833	108	51	L8G8K+	1089	22	12
L5G1K+	474	648	1089	L10G1K+	834	117	189	L13G2Na+	1091	48	57
L6K+	489	366	547	L8G4Na+	845	300	165	L12G3K+	1092	60	101
L4G3Na+	500	666	279	L7G5K+	846	129	137	L10G6Na+	1103	98	38
L5G2Na+	515	641	561	L11K+	849	126	51	L9G7K+	1104	30	33
L4G3K+	516	512	396	L9G3Na+	860	248	246	L14G1Na+	1106	22	27
L6G1Na+	530	450	626	L8G4K+	861	177	255	L13G2K+	1107	40	70
L5G2K+	531	699	699	L10G2Na+	875	175	248	L11G5Na+	1118	102	69
L7Na+	545	170	299	L9G3K+	876	143	339	L10G6K+	1119	49	70
L6G1K+	546	417	695	L7G6Na+	887	116	32	L15Na+	1121	10	6
L4G4Na+	557	486	137	L11G1Na+	890	83	150	L14G1K+	1122	28	30
L7K+	561	180	299	L10G2K+	891	103	306	L12G4Na+	1133	110	96
L5G3Na+	572	600	368	L8G5Na+	902	115	81	L11G5K+	1134	51	116
L4G4K+	573	403	222	L7G6K+	903	78	63	L15K+	1137	22	6
L6G2Na+	587	473	616	L12Na+	905	55	42	L13G3Na+	1148	81	99
L5G3K+	588	424	517	L11G1K+	906	83	167	L12G4K+	1149	55	146
L7G1Na+	602	420	589	L9G4Na+	917	200	152	L9G8K+	1161	1	73
L6G2K+	603	582	766	L8G5K+	918	91	140	L14G2Na+	1163	51	71
L8Na+	617	118	246	L12K+	921	55	42	L13G3K+	1164	41	136
L7G1K+	618	328	655	L10G3Na+	932	174	203	L10G7K+	1176	10	74
L5G4Na+	629	293	141	L9G4K+	933	97	232	L15G1Na+	1178	30	31
L8K+	633	162	246	L7G7Na+	944	40	7	L14G2K+	1179	28	88
L6G3Na+	644	350	314	L11G2Na+	947	116	186	L11G6K+	1191	33	59
L5G4K+	645	306	224	L10G3K+	948	116	279	L16Na+	1193	1	6
L7G2Na+	659	416	451	L8G6Na+	959	78	22	L15G1K+	1194	17	35
L6G3K+	660	407	439	L7G7K+	960	30	16	L12G5K+	1206	64	43
L8G1Na+	674	315	378	L12G1Na+	962	92	103	L16K+	1209	5	6
L7G2K+	675	425	559	L11G2K+	963	156	229	L9G9K+	1218	1	4
L5G5Na+	686	228	61	L9G5Na+	974	84	50	L13G4K+	1221	80	35
L9Na+	689	118	140	L8G6K+	975	68	43	L10G8K+	1233	6	21
L8G1K+	690	206	419	L13Na+	977	45	26	L14G3K+	1236	64	42
L6G4Na+	701	350	171	L12G1K+	978	116	115	L11G7K+	1248	22	31
L5G5K+	702	131	110	L10G4Na+	989	80	84	L15G2K+	1251	33	37
L9K+	705	120	140	L9G5K+	990	74	85	L12G6K+	1263	48	37
L7G3Na+	716	420	327	L13K+	993	87	26	L16G1K+	1266	10	28
L6G4K+	717	258	268	L11G3Na+	1004	68	102	L13G5K+	1278	71	43
L8G2Na+	731	267	411	L10G4K+	1005	80	128	L17K+	1281	1	1
L7G3K+	732	302	454	L8G7Na+	1016	33	8	L14G4K+	1293	71	33
L9G1Na+	746	245	306	L12G2Na+	1019	46	86	L15G3K+	1308	48	47
L8G2K+	747	232	509	L11G3K+	1020	68	140	L16G2K+	1323	22	45
L6G5Na+	758	92	57	L9G6Na+	1031	45	21	L17G1K+	1338	6	32
L10Na+	761	79	102	L8G7K+	1032	8	17	L18K+	1353	1	1
L9G1K+	762	219	339	L13G1Na+	1034	22	44				
L7G4Na+	773	172	136	L12G2K+	1035	46	106				

^a Lactic and glycine in the oligomer backbone L = O-CH(CH₃)-CO and G = NH-CH₂-CO.

^b Observed mass of the ion. Ions are due to oligomers of the type H-(L)_m-(G)_n-OH.

^c Observed intensity.

^d Calculated intensity.

15700 V, respectively. The laser irradiance was slightly above threshold (ca. 10⁶ W/cm²). Ions below *m/z* 250 were removed by pulsed deflection and 80 transient were summed. Sample preparation was as follows: 0.1 mL of a

5 g/L solution of the copolymer sample in tetrahydrofuran (THF) was added to 0.1 mL of a 500 mM solution of 2,5-dihydroxy benzoic acid (DHB) in THF. 1 μL was then placed on the probe tip.

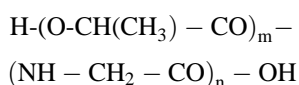
MACO4 calculations

The intensities of mass spectra of copolyesters were used to determine the sequence distribution of the copolymer samples by using the MACO4 computer program.^{2,3} Theoretical mass spectra intensities were computed assuming a first-order Markoff distribution. The theoretical intensities were then compared with the experimental ones. In order to perform the minimization, the program uses MINPACK which belongs to the ARGONNE library (see www.anl.gov). It computes the euclidean norm of the residuals vector. In addition it gives as an output the compositional values for which a minimum occurs, the norm and the covariance matrix. MACO4 uses the norm to compute the agreement factor³ and the elements, on the principal diagonal of the covariance matrix, to compute the error. The overall theoretical intensity of a peak is then spread to all the isotopes which are actually seen in the spectrum. Finite resolution is taken into account to a standard procedure²⁷.

RESULTS AND DISCUSSION

Spectral features

Figures 1(a)–1(c) report the MALDI-TOF mass spectrum of copolymer **1**. The mass spectrum displays about 163 well-defined mass spectral peaks between 350 and 1400 Da. Weak peaks are also present between 1400 and 2000 Da. The 163 peaks in the mass spectrum of Fig. 1(a) were assigned to oligomers bearing the same end groups, namely



II

and Table 1 reports the detailed peak assignment. (Mass spectral) peaks given in Table 1 range from pentamers to 18-mers. Each single oligomer produces two mass spectral peaks, namely a sodium adduct and a potassium adduct. The highest masses seen are the lactic acid 18-mer at 1337 Da and a peak at 1322 Da which corresponds to an oligomer with 17 lactic residues and one glycine residue.

Pentamers, heptamers and other odd-membered oligomers are present in the sample. The polymerization reaction (Scheme 1) should produce exclusively even-membered oligomers. Possible explanations for the presence of odd oligomers may be the hydrolysis of the polymer chain or an ester–ester exchange reaction occurring during polymerization. This phenomenon has already been observed in poly(lactic) produced by ring opening of the dilactide.²⁶

Figure 1(b) reports an enlarged portion of the spectrum, namely the spectral region with pentamers, hexamers and heptamers. The peak at 401 is the lactic pentamer. The lactic hexamer falls at 473 Da and is more than twice the intensity of the pentamer. The most intense peak in the mass spectrum is at 443 Da. It corresponds to an oligomer with 5 lactic residues and 2 glycine residues.

Figure 1(c) reports octamers, nonamers and decamers. It can be seen that the mass range covered by octamers does not superpose that of the mass range covered by nonamers. The same is not true for decamers (refer to Fig. 1(c) and Table 1).

In the 200 MHz ¹H-NMR spectrum of copolymer **1**, the methyl peak appears split into a multiplet, reflecting triads,

pentads, etc. In principle, the methyl absorptions in ¹H-NMR should allow the determination of copolymer sequence. However, in the absence of reliable peak assignment, the NMR yielded only the copolymer composition. (molar ratio glycine/lactic acid = 24:76).

MACO4 best-fit

Mass spectral peak intensities in Table 1 were given as input to the MACO4 program^{2,3} and a least-squares best-fit minimization was launched in order to find the four elements P_{GG}, P_{GL}, P_{LG}, P_{LL} which simulate the input data.

The minimization process converged towards the values P_{GG} = 0.04, P_{GL} = 0.96, P_{LG} = 0.29 and P_{LL} = 0.71. The composition associated with these four values (Eqn 3) is I(L) = 0.77, which compares well with the composition determined by NMR, namely I(L) = 0.76. Furthermore, the values of P_{GG} and P_{GL} resulting from the minimization (P_{GG} = 0.04, P_{GL} = 0.96) come extremely close to the expected values, as displayed in Eqn 5 (P_{GG} = 0, P_{GL} = 1). Similar calculations have been performed before^{17,18} on narrow fractions of different copolyesters (obtained by size exclusion chromatography (SEC)).

In this experiment the analysis was applied directly to the unfractionated sample, despite the fact that it is well known that MALDI-TOF underestimates the molar masses of polydisperse polymers.¹⁹ In fact, the mass spectrometric method used to calculate the composition and sequence of copolymers performs the best match of the theoretical and experimental mass peak intensities within the same oligomer series (dimers, trimers, ... decamers).^{2–18}

Therefore, the method does not suffer for the non-linear response of the MALDI detector, which is negligible for small mass ranges and becomes important when examining the whole mass range in polymer samples, since it is responsible for the well-known underestimation of the molar masses of polydisperse polymers.¹⁹

Acknowledgements

Partial financial support from the Italian Ministry for University and for Scientific and Technological Research (MURST), from progetto finalizzato materiali speciali per tecnologie avanzate II (CNR, Rome), and from the National Council of Research (CNR, Rome) is gratefully acknowledged.

REFERENCES

1. J. C. Randall, *Polymer Sequences Determination*, Academic Press, New York, (1977).
2. M. S. Montaudo, G. Montaudo and A. Ballistreri, *Macromolecules* **24**, 5051 (1991).
3. M. S. Montaudo and G. Montaudo, *Macromolecules* **25**, 4264 (1992).
4. G. Montaudo, C. Puglisi and F. Samperi, *Polym. Bull.* **21**, 483 (1989).
5. G. Montaudo, *Rapid Comm. Mass. Spectrom.* **5**, 95 (1991).
6. G. Montaudo, E. Scamporrino and D. Vitalini, *Macromolecules*, **22**, 623 (1989).
7. G. Montaudo, M. S. Montaudo, E. Scamporrino and D. Vitalini, *Macromolecules* **25**, 5099 (1992).
8. A. Ballistreri, D. Garozzo, M. Giuffrida, G. Impallomeni and G. Montaudo, *Macromolecules*, **22**, 2107 (1989).
9. G. Montaudo, *Trends Polym. Sci.* **4**, 81 (1996).
10. A. Ballistreri, G. Impallomeni, G. Montaudo, R. W. Lenz, B. Kim and R. C. Fuller, *Macromolecules* **23**, 5059 (1990).
11. G. Montaudo, M. S. Montaudo, E. Scamporrino and D. Vitalini, *Makromol. Chemie* **194**, 993 (1993).

12. D. Garozzo, M. Giuffrida, G. Montaudo and R. W. Lenz, *J. Polym. Sci., Polym. Chem. Ed* **25**, 271 (1987).
13. A. Ballistreri, D. Garozzo, M. Giuffrida, G. Montaudo and M. S. Montaudo, *Macromolecules* **24**, 1231 (1991).
14. G. Montaudo, E. Scamporrino and D. Vitalini, *Macromolecules* **24**, 376 (1991).
15. G. Montaudo, M. S. Montaudo, C. Puglisi and F. Samperi, *Rapid Commun. Mass Spectrom.* **12**, 519 (1998).
16. R. Abate, A. Ballistreri, D. Garozzo, G. Impallomeni, G. Montaudo, G. Critchley and K. Tanaka, *Rapid Commun. Mass Spectrom.* **7**, 1033 (1993).
17. G. Montaudo, D. Garozzo, M. S. Montaudo, C. Puglisi and F. Samperi, *Macromolecules* **28**, 7983 (1995).
18. G. Montaudo, M. S. Montaudo, C. Puglisi and F. Samperi, *Rapid Commun. Mass Spectrom.* **19**, 445 (1998).
19. G. Montaudo, M. S. Montaudo, C. Puglisi and F. Samperi, *Rapid Commun. Mass Spectrom.* **9**, 453 (1995).
20. D. Wang and X. D. Feng, *Macromolecules* **31**, 3824 (1998).
21. T. Ouchi, H. Seike, T. Nozaki and Y. Ohya, *J. Polym. Sci. Part A Polym. Chem.* **36**, 1283, (1998).
22. H. Höcker, *Makromol. Symp.* **130**, 161 (1998).
23. G. John, S. Tsuda and M. Morita, *J. Polym. Sci. Part A Polym. Chem.* **35**, 1901 (1997).
24. J. Helder, J. Feijen, S. J. Lee and S. W. Kim, *Macromol. Chem. Rapid Commun.* **7**, 193 (1986).
25. J. Helder, F. E. Kohn, S. Sato, J. W. van der Berg and J. Feijen, *Makromol. Chem. Rapid Comm.* **6**, 9 (1985); J. Helder, *PhD Thesis* Twente University of Technology, The Netherlands (1988).
26. G. Montaudo, M. S. Montaudo, C. Puglisi, F. Samperi, N. Spassky, M. Wisniewski and A. LeBorgue, *Macromolecules* **29**, 6461 (1996).
27. A. Yergey, *Int. J. Mass Spectrom. Ion Phys.* **52**, 337 (1983); B. Kubinyi, *Anal. Chim. Acta* **107**, 247 (1991); Senko: ISOPRO computer program version 2.0 Cornell University; D. Rockwood, *Rapid Commun. Mass Spectrom.* **9**, 103 (1995).