

## Method for Macromolecular Colocalization Using Atomic Recombination in Dynamic SIMS

G. Legent,<sup>†,‡</sup> A. Delaune,<sup>†,‡,§</sup> V. Norris,<sup>†,‡,||</sup> A. Delcorte,<sup>‡,⊥</sup> D. Gibouin,<sup>†,‡</sup> F. Lefebvre,<sup>†,‡</sup>  
G. Misevic,<sup>†,‡,||</sup> M. Thellier,<sup>†,‡,||</sup> and C. Ripoll<sup>\*,†,‡,||</sup>

Laboratoire "Assemblages moléculaires: modélisation, et imagerie SIMS" (AMMIS), Faculté des Sciences de l'Université de Rouen, 76821 Mont Saint Aignan Cedex, France, UPRES EA 2123, Faculté de Médecine-Pharmacie de l'Université de Rouen, Boulevard, Gambetta, 76000 Rouen, France, PCPM, Université Catholique de Louvain, 1 Croix du Sud, B1348 Louvain-la-Neuve, Belgium, and "Nanobeams" European Network of Excellence, Epigenomics Programme, Genopole<sup>®</sup>, 91000 Evry, France

Received: October 16, 2007; In Final Form: January 9, 2008

Localizing two or more components of assemblies in biological systems requires both continued development of fluorescence techniques and invention of entirely new techniques. Candidates for the latter include dynamic secondary ion mass spectrometry (D-SIMS). The latest generation of D-SIMS, the Cameca NanoSIMS 50, permits the localization of specific, isotopically labeled molecules and macromolecules in sections of biological material with a resolution in the tens of nanometers and with a sensitivity approaching in principle that of a single protein. Here we use two different systems, crystals of glycine and mixtures of proteins, to show that the formation of recombinant CN secondary ions under Cs bombardment can be exploited to create a new colocalization technique. We show experimentally that the formation of the recombinant  $^{13}\text{C}^{15}\text{N}$  secondary ion between  $^{13}\text{C}$ - and  $^{15}\text{N}$ -labeled macromolecules is indeed an indicator of the distance between the interacting macromolecules and on their shape. We build up a convolution model of the mixing-recombination process in D-SIMS that allows quantitative interpretations of the distance-dependent formation of the recombinant CN. Our results show that macromolecules can be colocalized if they are within 2 nm of one another. We discuss the potential advantages of this new technique for biological applications.

## Introduction

Macromolecular assemblies are central to processes that include metabolism, signaling, and the cell cycle in both prokaryotic and eukaryotic cells.<sup>1–3</sup> One approach to investigate the nature and role of these assemblies is to find out which macromolecules they contain. This often entails recourse to techniques such as electron microscopy using antibodies coupled to gold beads or fluorescence using the green fluorescent protein or its derivatives as tags. In the case of colocalization studies, the limitation on resolution due to the wavelength of light is circumvented by fluorescence resonance energy transfer (FRET), which can determine whether two different species of macromolecules are colocalized to typically within 3–10 nm in an assembly.<sup>4,5</sup>

Secondary ion mass spectrometry (SIMS) involves bombarding the surface of a sample with a primary ion beam and analyzing the secondary ions produced. Coincidence SIMS<sup>6</sup> has been used in studies of inorganic materials to colocalize the atoms giving rise to coincident secondary ions. This technique requires reduction of the intensity of the primary beam to the level where single primary ions can be resolved in time and space; in these conditions, the different secondary ions originating from the same primary ion impact can therefore also be resolved. Coincidence SIMS has the advantage of allowing

colocalization on the 10-nm scale but the disadvantage of being a nonimaging technique.<sup>7,8</sup>

Our ultimate aim is to develop a new method to colocalize two macromolecules using images obtained in dynamic secondary ion mass spectrometry (D-SIMS).<sup>9–12</sup> This technique involves rastering a focused primary ion beam across the sample and analyzing the secondary ions produced by the sputtering; specific images of the distribution of the secondary ions can be obtained for both stable and unstable isotopes<sup>13–15</sup> and hence the distribution of the molecules that contain these isotopes can be determined with a resolution of  $\sim 50$  nm.<sup>16–18</sup> This resolution might seem insufficient for colocalization purposes. However, a process occurs in SIMS that may allow this limitation to be circumvented (and that could become a SIMS equivalent of FRET). This process is that of atomic recombination<sup>9,10,19</sup> which could be exploited to generate a distance-dependent secondary ion. This is because the sputtering of two neighboring molecules, one strongly enriched in  $^{13}\text{C}$  and the other in  $^{15}\text{N}$  (both are rare stable isotopes in biomolecules), produces the recombinant<sup>19,20</sup> secondary ion  $^{13}\text{C}^{15}\text{N}^-$  provided these molecules are close to one another (see Figure 1 for details). This is analogous to the production of a distance-dependent fluorescence in FRET when two labeled molecules with specific fluorophores are located at a distance lower than the Förster distance.<sup>4</sup> The production of a distance-dependent secondary ion should permit the following in a single SIMS imaging experiment: (i) determination of whether two macromolecules are in contact and (ii) localization of their position in a cell with a lateral resolution that depends on the type of D-SIMS instrument used; with a Cameca NanoSIMS 50, this lateral resolution can be a few tens

\* To whom correspondence should be addressed. Tel: +33-235528401. Fax: +33-235522981. E-mail: Camille.Ripoll@univ-rouen.fr.

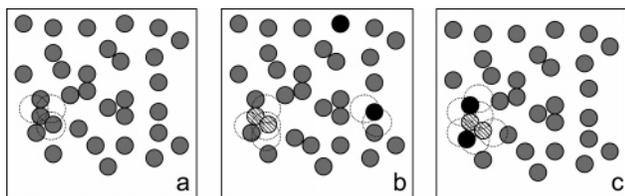
<sup>†</sup> Faculté des Sciences de l'Université de Rouen.

<sup>‡</sup> "Nanobeams" European Network of Excellence.

<sup>§</sup> Faculté de Médecine-Pharmacie de l'Université de Rouen.

<sup>||</sup> Epigenomics Programme, Genopole.

<sup>⊥</sup> Université Catholique de Louvain.



**Figure 1.** Schematic view of the method based on the formation of recombinant secondary ions  $^{13}\text{C}^{15}\text{N}$  to determine the proximity of two macromolecules in a biological sample (resin-embedded cells or tissue). The frames in the figure represent a square area  $50\text{ nm} \times 50\text{ nm}$  (size of a pixel in a NanoSIMS image) on the sample surface. Typically, such an area might contain 30 proteins, represented by filled circles in the figure. The rest of the area is occupied by the embedding resin. Unlabeled proteins (gray) contain solely  $^{12}\text{C}$  and  $^{14}\text{N}$ ,  $^{13}\text{C}$ -labeled proteins (black circles) or  $^{15}\text{N}$ -labeled proteins (cross-hatched circles). The dashed empty circles represent areas from which the recombinant CN ions emerge. In frame a, only unlabeled proteins are present in the area and therefore, under the primary ion bombardment, only the  $^{12}\text{C}^{14}\text{N}$  recombinant ions will be formed. In frame b, the  $^{13}\text{C}$ -labeled and  $^{15}\text{N}$ -labeled proteins are at distances preventing recombination between  $^{13}\text{C}$  and  $^{15}\text{N}$  atoms. In frame c, the labeled proteins are in close contact and therefore  $^{13}\text{C}^{15}\text{N}$  ions will be formed. This is indication that the labeled proteins are in this sample area and in close contact.

of nanometers,<sup>16,17,21</sup> better than that achievable with any fluorescence method.

A full understanding of the recombination process itself would entail characterizing the type of atoms or atomic groups participating in the recombination (e.g.,  $\text{C} + \text{N}^-$  or  $\text{C}^- + \text{N}$ ), their configurations, and several other intertwined parameters. Such characterization, while intrinsically interesting, is difficult. Moreover, it is not essential for an initial evaluation of D-SIMS recombination as the basis for a general method of colocalizing molecules and macromolecules in real biological systems. Rather, evaluation requires the quantitative characterization of the production of distance-dependent secondary ions in D-SIMS.

We chose  $^{13}\text{C}^{15}\text{N}$  as distance-dependent secondary ions, and to generate them, we used two physicochemical model systems, one a mixture of  $^{13}\text{C}$ -labeled,  $^{15}\text{N}$ -labeled, and unlabeled L-glycine and the other a mixture of  $^{13}\text{C}$ -labeled,  $^{15}\text{N}$ -labeled, and unlabeled proteins. We show here that the production of secondary ions by D-SIMS does indeed have promising distance-dependent characteristics.

## Experimental Methods

**Production of Labeled Bacterial Proteins.** Bacteria from *Escherichia coli* strain JM 109 (*endA1*, *recA1*, *gyrA96*, *thi*, *hdsR17* ( $r_k^-$ ,  $m_k^+$ ), *relA1*, *supE44*,  $\Delta$ (*lac-proAB*), [*F'* *traD36*, *proAB*, *lacq19* $\Delta$ M15]; Promega, Charbonnières-les-Bains, France) were grown for 4 days at  $37\text{ }^\circ\text{C}$  with gentle stirring in a liquid M9 medium prepared by adding to 1 L of phosphate-buffered saline (6 g/L  $\text{Na}_2\text{HPO}_4$ , 3 g/L  $\text{KH}_2\text{PO}_4$ , 0.415 g/L NaCl): 1 mL of a 1 M solution of  $\text{MgSO}_4$ , 10 mL of a 0.01 M solution of  $\text{CaCl}_2$ , and 10 mL of a 20% (w/v) glucose solution (these 4 solutions were previously and separately sterilized in an autoclave). The medium is supplemented with 1 g/L  $\text{NH}_4\text{Cl}$  as the sole source of nitrogen. JM 109 is *thi* and requires thiamine; hence, the medium was supplemented with 10 mL/medium L of the vitamin mixture from the MEM vitamin kit (Sigma-Aldrich).

Three isotopic types or classes of proteins were prepared.  $^{15}\text{N}$ -Enriched proteins were obtained from bacteria grown on a medium supplemented with  $^{15}\text{NH}_4\text{Cl}$  1.02 g/L (Isotec/Sigma-Aldrich) as the sole source of nitrogen.  $^{13}\text{C}$ -Enriched proteins were obtained from bacteria grown on U- $^{13}\text{C}$ -glucose (Isotec/Sigma-Aldrich) as the source of carbon. Unlabeled proteins were

in fact  $^{13}\text{C}$ -depleted proteins; they were obtained from bacteria grown on  $^{13}\text{C}$ -depleted glucose (Isotec/Sigma-Aldrich).

Protein extraction was carried out using the protocol as in Roe.<sup>22</sup> Bacteria grown in the appropriate medium were twice pelleted at 5000 rpm and resuspended in phosphate buffer containing glycerol (50 mL of phosphate buffer (0.82 g/mL  $\text{KH}_2\text{PO}_4$ , 21.45 g/mL  $\text{K}_2\text{HPO}_4$ ), 100 mL of glycerol, 17.54 g of NaCl, and water up to 1 L). The pellet was then resuspended in phosphate buffer (1 M, pH 8) containing lysozyme (24  $\mu\text{L}/10\text{ mL}$ ), SDS (sol 1 M, 30  $\mu\text{L}/10\text{ mL}$ ), and antiproteases (24  $\mu\text{L}/10\text{ mL}$ ) but not glycerol. The bacteria suspension was then subjected to a thermal shock by successively cooling the suspension in melting ice (40 min at  $4\text{ }^\circ\text{C}$ ) and placing it in a water bath at  $73\text{ }^\circ\text{C}$  (15 min). This treatment led to bacteria lysis, cell wall digestion, and membrane solubilization. The lysed suspension was then centrifuged at 12 000 rpm for 20 min to pellet cell debris (wall, membranes, DNA). The supernatant was then recovered and filtered using a Centricon (Amicon) with a 50-kDa molecular weight cutoff (MWCO); the filtrate was recovered and filtered using a Centricon 30-kDa MWCO. The retained proteins were then solubilized in ultrapure water (Sigma-Aldrich) to a final concentration of 1.2 mg/mL (determined using a Bradford assay (kit BioRad)). This allowed recovery of the cytosolic soluble proteins with a mass in the range 30–50 kDa.

The isotopic abundances of  $^{13}\text{C}$  or  $^{15}\text{N}$  of the three isotopic types, T, of proteins were then measured (see below) with the Cameca IMS 4F ion analyzer. The isotopic abundances (fractions) are written  $a_{A,T}$ , with  $A = 12, 13, 14, \text{ or } 15$  for  $^{12}\text{C}$ ,  $^{13}\text{C}$ ,  $^{14}\text{N}$ , or  $^{15}\text{N}$  isotopes in  $T = \text{C, N, or } 0$  (zero) for  $^{13}\text{C}$ -,  $^{15}\text{N}$ -enriched proteins or unlabeled proteins (a complete list of symbols is given in Appendix A). The measured values of the abundances were as follows:  $^{13}\text{C}$  abundance in  $^{13}\text{C}$ -enriched proteins,  $a_{13,\text{C}} = 0.911 \pm 0.026$  (mean  $\pm$  standard error);  $^{15}\text{N}$  abundance in  $^{15}\text{N}$ -enriched proteins,  $a_{15,\text{N}} = 0.901 \pm 0.032$ ;  $^{12}\text{C}$  abundance in unlabeled proteins,  $a_{12,0} = 0.997 \pm 0.005$ . Unmeasured isotopic abundances were taken as their natural values, i.e.:  $a_{12,\text{N}} = 0.989$ ;  $a_{14,0} = a_{14,\text{C}} = 0.99634$ .

**Preparation of the Protein Films for SIMS Analysis.** Three stock solutions containing 1.2 mg/mL  $^{15}\text{N}$ -,  $^{13}\text{C}$ -, and unlabeled proteins, respectively, were prepared using ultrapure water purchased from Sigma-Aldrich. Equal volumes of solutions of the labeled proteins were then mixed to prepare a new stock solution: the LP solution. The total concentration of proteins in LP is therefore 1.2 mg/mL (0.6 mg/mL of each type of labeled proteins). The stock solution of unlabeled proteins was named ULP. Then 11 different solutions were prepared by mixing a volume  $v$  (arbitrary units) of LP and  $10-v$  volumes of ULP, with  $v = 0, 1, 2, \dots, 10$ . We defined the concentration factor as  $X = v/10$ ; therefore,  $X = 0$  for ULP and  $X = 1$  for LP and have intermediate values for the mixtures of LP and ULP. To obtain protein films that could be analyzed by SIMS, a small drop ( $\sim 1\text{ }\mu\text{L}$ ) of each solution was deposited on the surface of a mirror-polished stainless steel slab 1 cm in diameter (the slabs were previously cleaned by successive sonication in acetone, ethanol, mQ water (Millipore, Billerica, MA), and absolute ethanol and dried at  $80\text{ }^\circ\text{C}$  in an oven). The slabs were then placed in an oven at  $70\text{ }^\circ\text{C}$  at normal pressure for 1 h. The water was evaporated to give a protein film that remained adherent to the slab surface. The slab was then put in the air lock of the IMS 4F, where it stayed for 12 h at  $20\text{ }^\circ\text{C}$  before being transferred to the sample chamber (this completed the drying and degassing of the sample). Three additional slabs with

drops of pure  $^{15}\text{N}$ - and  $^{13}\text{C}$ -labeled and unlabeled proteins were prepared using the same procedure.

**Preparation of Glycine Samples for SIMS Analysis.**  $^{15}\text{N}$ - and  $^{13}\text{C}$ -labeled L-glycine and unlabeled L-glycine were purchased from Sigma-Aldrich (Milwaukee, WI). Three stock solutions containing 10 mg/mL  $^{15}\text{N}$ -,  $^{13}\text{C}$ -, and unlabeled glycine, respectively, were prepared using ultrapure water purchased from Sigma-Aldrich. Equal volumes of solutions of the labeled glycines were then mixed to prepare a new stock solution: the LG solution. The total concentration of glycine in LG is therefore 10 mg/mL (5 mg/mL of each type of labeled glycine). The stock solution of unlabeled L-glycine was named ULG. Then 11 different solutions were prepared by mixing a volume  $v$  (arbitrary units) of LG and  $10-v$  volume of ULG, with  $v = 0, 1, 2, \dots, 10$ . With the definition of the concentration factor  $X = v/10$ ,  $X = 0$  corresponds to ULG and  $X = 1$  to LG; intermediate values of  $X$  are for the mixtures of LG and ULG. To obtain crystals for analysis by SIMS, a small drop (1  $\mu\text{L}$ ) of each solution was deposited on the surface of a clean silicon wafer (doped with boron for better conductivity) 1 cm in diameter (Siltronix, Archamps, France). The wafers were then placed in an oven at 70  $^{\circ}\text{C}$  at normal pressure during 30 min. The water was evaporated and small glycine crystals, randomly oriented, formed and remained adherent to the silicon surface. The drying was then prolonged for 12 h at 40  $^{\circ}\text{C}$  at reduced pressure (0.1 bar) in the oven. After this treatment, the wafers were coated with  $\sim 40$  nm of gold (Cressington sputter coater, Watford, UK) before SIMS analysis.

Two additional wafers with drops of pure  $^{15}\text{N}$ - and  $^{13}\text{C}$ -labeled L-glycines were prepared using the same procedure. These permitted the SIMS measurement (see next section) of the abundance (isotopic fraction) of  $^{13}\text{C}$  in  $^{13}\text{C}$ -labeled glycine:  $a_{13,\text{C}} = 0.928 \pm 0.021$  (mean  $\pm$  standard error) and of the abundance of  $^{15}\text{N}$  in  $^{15}\text{N}$ -labeled glycine:  $a_{15,\text{N}} = 0.920 \pm 0.015$ . The other isotopic abundances were at their natural values, i.e.:  $a_{12,\text{N}} = a_{12,\text{O}} = 0.989$ ,  $a_{14,\text{C}} = a_{14,\text{O}} = 0.99634$ .

**SIMS Analysis of Glycine and Protein Films.** In principle, SIMS determination of isotopic abundances would take into account different useful yields obtained in different machines for the different isotopes of the same element.<sup>23</sup> In practice, this instrumental mass fractionation has been measured for the CN ion using a Cameca IMS 6F and diamond standards as  $\sim 1\%$ .<sup>24</sup> This fractionation is ignored in this work since we calculate that an instrumental mass fractionation of 5% would alter the recombination diameter (see below) by only 2–3% (not shown), which is negligible at this stage of development of our D-SIMS colocalization method.

SIMS analysis of glycine samples was performed using a NanoSIMS 50 ion analyzer (Cameca). This instrument permits parallel detection and quantification of five chosen secondary ions and is therefore highly accurate in the determination of the isotopic ratios. The NanoSIMS was used in the following conditions: detection of negative secondary ions,  $\text{Cs}^+$  primary ions 16 keV in energy, primary current 1.5 pA, and mass resolving power (MRP) 5500–6000 (the transmission of the NanoSIMS 50 is still maximum at this MRP). Each drop on the silicon disk was analyzed in 10–15 randomly chosen areas 25  $\mu\text{m} \times 25 \mu\text{m}$  in size. To achieve this, the primary beam was slightly defocused (probe diameter estimated at a few  $\mu\text{m}$ ) and rastered over the field divided in  $64 \times 64$  pixels (this allows a large overlap of the probe on neighboring pixels and contributes to averaging the signal). The dwell time was 132  $\mu\text{s}$ /pixel. After the stationary regime was reached (10–20 min with this low primary beam intensity and rather thick gold coverage), 1000

measurements of the intensity ratios  $I(^{13}\text{C}^{14}\text{N})/I(^{12}\text{C}^{14}\text{N})$ ,  $I(^{12}\text{C}^{15}\text{N})/I(^{12}\text{C}^{14}\text{N})$ , and  $I(^{13}\text{C}^{15}\text{N})/I(^{12}\text{C}^{14}\text{N})$  were carried out using the “isotope program” of the NanoSIMS; this took  $\sim 10$  min. The symbols  $I(^{\text{A}}\text{C}^{\text{A}'}\text{N})$ , where  $A = 12$  or  $13$  and  $A' = 14$  or  $15$ , represent the intensity of the current of the secondary  $^{\text{A}}\text{C}^{\text{A}'}\text{N}$  ion. Indeed, even though the measured absolute values of the secondary currents may vary from an area to the other, the ratios of these values vary much less. We therefore calculated the mean value and standard error of the different CN ratios of all the areas analyzed in a given drop.

Two sets of measurements were performed. In set 1, we measured the currents of  $^{12}\text{C}^{14}\text{N}$ ,  $^{12}\text{C}^{15}\text{N}$ , and  $^{13}\text{C}^{15}\text{N}$  secondary ions, and in set 2, we measured those of  $^{12}\text{C}^{14}\text{N}$ ,  $^{13}\text{C}^{14}\text{N}$ , and  $^{13}\text{C}^{15}\text{N}$  (due to the nature of the detection system of the NanoSIMS 50 it is not possible to measure in a single experiment the intensity of the 2 molecular ions  $^{12}\text{C}^{15}\text{N}$  and  $^{13}\text{C}^{14}\text{N}$  along with those of  $^{12}\text{C}^{14}\text{N}$  and  $^{13}\text{C}^{15}\text{N}$ ). Nevertheless, in both sets of experiments, it is possible to calculate the unmeasured ratio from the measured ones. Indeed, the ratio  $[I(^{13}\text{C}^{14}\text{N}) + I(^{13}\text{C}^{15}\text{N})]/[I(^{12}\text{C}^{14}\text{N}) + I(^{12}\text{C}^{15}\text{N})]$  must be equal to the ratio  $\rho$  of the number of  $^{13}\text{C}$  atoms to that of the  $^{12}\text{C}$  atoms in the sample. Therefore, in set 1 we calculated

$$\frac{I(^{13}\text{C}^{14}\text{N})}{I(^{12}\text{C}^{14}\text{N})} = \rho \left( 1 + \frac{I(^{12}\text{C}^{15}\text{N})}{I(^{12}\text{C}^{14}\text{N})} \right) - \frac{I(^{13}\text{C}^{15}\text{N})}{I(^{12}\text{C}^{14}\text{N})}$$

and in set 2

$$\frac{I(^{12}\text{C}^{15}\text{N})}{I(^{12}\text{C}^{14}\text{N})} = \frac{1}{\rho} \left( \frac{I(^{13}\text{C}^{14}\text{N})}{I(^{12}\text{C}^{14}\text{N})} + \frac{I(^{13}\text{C}^{15}\text{N})}{I(^{12}\text{C}^{14}\text{N})} \right) - 1$$

It is easily verified that

$$\rho = \left( \frac{2a_{13,0} + (a_{13,\text{C}} - a_{13,0})X}{2(1 - a_{13,0}) - (a_{13,\text{C}} - a_{13,0})X} \right)$$

where  $X$  is the concentration factor and the  $a_{\text{A},\text{T}}$  are the isotopic abundances defined in the previous section. Numerically,  $\rho = (0.022 + 0.917X)/(1.978 - 0.917X)$ .

Finally, the fraction  $f_{\text{exp}}(^{13}\text{C}^{15}\text{N})$  of  $^{13}\text{C}^{15}\text{N}$  in the mixture of the different molecular CN ions was calculated with:

$$f_{\text{exp}}(^{13}\text{C}^{15}\text{N}) = \frac{\frac{I(^{13}\text{C}^{15}\text{N})}{I(^{12}\text{C}^{14}\text{N})}}{1 + \frac{I(^{13}\text{C}^{14}\text{N})}{I(^{12}\text{C}^{14}\text{N})} + \frac{I(^{12}\text{C}^{15}\text{N})}{I(^{12}\text{C}^{14}\text{N})} + \frac{I(^{13}\text{C}^{15}\text{N})}{I(^{12}\text{C}^{14}\text{N})}}$$

for the different values of the concentration factor  $X$ .

Protein films were analyzed using a Cameca IMS 4F ion analyzer. The setting was as follows: primary ions  $\text{Cs}^+$ , energy 14.5 keV, primary current 300 pA, mass resolving power 4500, analyzed area 100  $\mu\text{m}^2$ , and presputtering: 15 s. With the IMS ion analyzer, the four intensities  $I(^{13}\text{C}^{14}\text{N})$ ,  $I(^{13}\text{C}^{15}\text{N})$ ,  $I(^{12}\text{C}^{14}\text{N})$ , and  $I(^{12}\text{C}^{15}\text{N})$  are acquired sequentially in the same experiment.

## Results

**Two Experimental Model Systems.** The model systems contained two types of  $^{13}\text{C}$ - and  $^{15}\text{N}$ -labeled biological molecules or macromolecules: L-glycine and proteins. The first system consisted of crystals containing various proportions of  $^{13}\text{C}$ -labeled L-glycine,  $^{15}\text{N}$ -labeled L-glycine, and unlabeled L-glycine. The average distances between  $^{15}\text{N}$  atoms and  $^{13}\text{C}$  atoms in the crystal were increased by increasing the proportion of

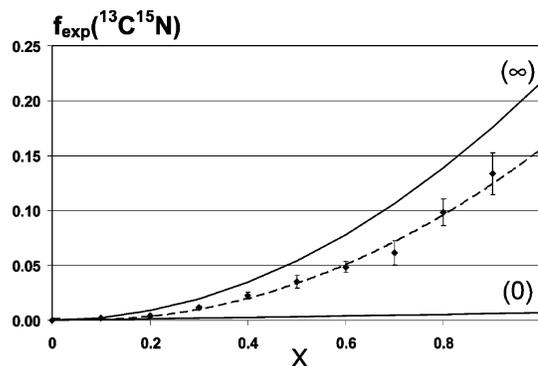
unlabeled glycine. The second model consisted of a thin film of *E. coli* proteins containing different proportions of  $^{13}\text{C}$ - and  $^{15}\text{N}$ -labeled proteins mixed with unlabeled *E. coli* proteins. Two types of SIMS instruments were used in these experiments: a Cameca NanoSIMS 50 (glycine) and a Cameca IMS 4F (proteins).

**Two Limiting Cases of Recombination between C and N Atoms.** In dynamic SIMS, the molecules of the surface are heavily fragmented down to the atomic level. Moreover, the collision cascades of the incoming primary ions in the top few nanometers of the surface of the sample lead to a mixing, i.e., a displacement of the atoms from their original position. On one hand, this creates a sort of “confusion volume”, which ultimately limits the lateral resolution of SIMS images. On the other hand, the production of single atoms (or very small atomic assemblies) is the basis of the recombination process between these atoms when they escape from the surface of the sample. This means that, at the surface under the bombardment, a carbon atom can recombine with an initially distant nitrogen atom. Clearly the probability of recombination between these atoms is a decreasing function of the initial distance between them. The mixing and the resulting surface distribution of the atoms are determined by the very complex process of collision cascades.

Two limiting cases of recombination can be expected. The first case corresponds to recombination occurring between atoms in a surface in which their distribution remains unperturbed despite bombardment by the primary ions. In this case, there is no mixing and the distribution of the atoms (isotopes) at the surface is essentially that existing before the bombardment (ideal static SIMS conditions). Hence, a carbon atom has a high probability of recombination only if this recombination occurs with a nitrogen atom initially (before bombardment) located in its immediate vicinity; i.e., the recombination is essentially between atoms of the same molecule. In this case, it is the unperturbed isotope distribution in the sample that determines the recombination process. We term this first limiting case the “unperturbed recombination” and denote as  $f_0(^{13}\text{C}^{15}\text{N})$  the fraction of  $^{13}\text{C}^{15}\text{N}$  that is produced (see list of symbols in Appendix A). In contrast, the second case corresponds to a complete atomic mixing under the primary beam. This means that any nitrogen atom in the sample has the same probability of recombination with any carbon in the sample whatever the initial (before bombardment) distance between these atoms. We term this second limiting case the “averaged recombination” and denote as  $f_\infty(^{13}\text{C}^{15}\text{N})$  the fraction of  $^{13}\text{C}^{15}\text{N}$  that is produced.

**Recombinant CN Ions in the Case of Glycine Crystals.** The experimental values  $f_{\text{exp}}(^{13}\text{C}^{15}\text{N})$  are given as a function of the concentration factor  $X$  (Figure 2). To show that the measured fraction of  $^{13}\text{C}^{15}\text{N}$  is not merely a function of the proportion of  $^{13}\text{C}$  and  $^{15}\text{N}$  isotopes in the sample but that it depends on the distance between these isotopes, we compared the experimental results to those calculated for the two limiting cases of recombination. Clearly a distance-dependent recombination of  $^{13}\text{C}$  and  $^{15}\text{N}$  will result in  $f_{\text{exp}}(^{13}\text{C}^{15}\text{N})$  values between those of  $f_0(^{13}\text{C}^{15}\text{N})$  and  $f_\infty(^{13}\text{C}^{15}\text{N})$ .

Let us first consider the unperturbed recombination case. Eight isotopic configurations containing different amounts of  $^{12}\text{C}$ ,  $^{13}\text{C}$ ,  $^{14}\text{N}$ , and  $^{15}\text{N}$  are possible for the glycine molecule (the isotopes of H and O are not involved in the formation of the CN recombinant ions). Among these configurations, in the case of local intramolecular recombination, only the molecules with the isotopic configuration  $^{15}\text{NH}_2\text{-}^{12}\text{CH}_2\text{-}^{13}\text{COOH}$ ,  $^{15}\text{NH}_2\text{-}^{13}\text{CH}_2\text{-}^{12}\text{COOH}$ , and  $^{15}\text{NH}_2\text{-}^{13}\text{CH}_2\text{-}^{13}\text{COOH}$  can produce  $^{13}\text{C}^{15}\text{N}$



**Figure 2.** Fraction of  $^{13}\text{C}^{15}\text{N}$  in the mixture of CN recombinant ions in glycine samples plotted as a function of the concentration factor  $X$ . Points: experimental data  $\pm$  standard error; curve (0), unperturbed limiting case  $f_0(^{13}\text{C}^{15}\text{N})$ ; curve ( $\infty$ ), averaged limiting case  $f_\infty(^{13}\text{C}^{15}\text{N})$ . The experimental points are fitted well ( $r^2 = 0.9920$ ) by a quadratic function (dashed line) and are much closer to the quadratic limiting curve  $f_\infty(^{13}\text{C}^{15}\text{N})$  than to the linear curve  $f_0(^{13}\text{C}^{15}\text{N})$ .

recombinant ions. The mean number of molecules with these configurations in one of the three classes of glycine ( $^{15}\text{N}$ -labeled,  $^{13}\text{C}$ -labeled, unlabeled) is proportional to  $a_{15,\text{T}} \cdot a_{12,\text{T}} \cdot a_{13,\text{T}}$ ,  $a_{15,\text{T}} \cdot a_{13,\text{T}} \cdot a_{12,\text{T}}$  and  $a_{15,\text{T}} \cdot a_{13,\text{T}} \cdot a_{13,\text{T}}$ , respectively, where the subscript T, equal N or C or O, is for the three classes of glycine and the  $a_{\text{A},\text{T}}$  are the isotopic abundances of carbon and nitrogen in these three classes (for the notation, see the section on preparation of the crystals of glycines for SIMS analysis in the Experimental Section). The mean number of molecules can be used in this calculation because the SIMS determination of the isotopic abundances is carried out using the simultaneous sputtering of a number of glycine molecules big enough to have an isotopic composition equal to the mean composition. Assume that in glycine molecules, for every set  $N$  of recombinations occurring between a nitrogen atom and a carbon of the molecule,  $pN$  ( $0 < p < 1$ ) occur with the carbon of the carboxyl and  $(1 - p)N$  with the central carbon. Therefore, after sputtering the molecules of configuration  $^{15}\text{NH}_2\text{-}^{12}\text{CH}_2\text{-}^{13}\text{COOH}$ ,  $^{15}\text{NH}_2\text{-}^{12}\text{CH}_2\text{-}^{13}\text{COOH}$ , and  $^{15}\text{NH}_2\text{-}^{13}\text{CH}_2\text{-}^{13}\text{COOH}$ , the fraction of  $^{13}\text{C}^{15}\text{N}$  ions collected is equal to

$$a_{15,\text{T}} a_{12,\text{T}} a_{13,\text{T}} p + a_{15,\text{T}} a_{13,\text{T}} a_{12,\text{T}} (1 - p) + a_{15,\text{T}} a_{13,\text{T}} a_{13,\text{T}} = a_{15,\text{T}} a_{13,\text{T}}$$

Therefore, for a sample with a concentration factor  $X$ , the fraction  $f_0(^{13}\text{C}^{15}\text{N})$  is given by

$$f_0(^{13}\text{C}^{15}\text{N}) = a_{15,\text{N}} \times a_{13,\text{N}} \times \frac{X}{2} + a_{15,\text{C}} \times a_{13,\text{C}} \times \frac{X}{2} + a_{15,0} \times a_{13,0} \times (1 - X)$$

or

$$f_0(^{13}\text{C}^{15}\text{N}) = a_{15,0} \times a_{13,0} + \left[ \frac{(a_{15,\text{N}} \times a_{13,\text{N}} + a_{15,\text{C}} \times a_{13,\text{C}})}{2} - a_{15,0} \times a_{13,0} \right] \times X$$

Hence,  $f_0(^{13}\text{C}^{15}\text{N})$  is a linear function of  $X$ . Numerically, with the experimentally determined values of the C and N isotopic abundances in the labeled glycines

$$f_0(^{13}\text{C}^{15}\text{N}) = 4.026 \times 10^{-5} + 6.718 \times 10^{-3} \times X$$

This function is represented by the straight lower line in Figure 2. Note that in the absence of unlabeled glycine ( $X = 1$ ) the

maximum value of  $f_0(^{13}\text{C}^{15}\text{N})$  is only  $6.758 \times 10^{-3}$  (0.6758%). It would be exactly 0 for a 1:1 mixture of “isotopically pure” labeled glycines (100%  $^{15}\text{N}$ , 100%  $^{12}\text{C}$ ) and (100%  $^{14}\text{N}$ , 100%  $^{13}\text{C}$ ).

Consider now the averaged limiting case. A sample with a concentration factor  $X$  is obtained by mixing  $X$  volumes of the stock solution of labeled glycines, LG, and  $(1 - X)$  volumes of the stock solution of unlabeled glycine, ULG. As the concentration of each of the labeled glycines in the stock solution LG is  $C/2$  and the concentration of the unlabeled glycine in the stock solution ULG is  $C$ , it is easily verified that the mean abundances of  $^{13}\text{C}$  and  $^{15}\text{N}$  in a sample with a concentration factor  $X$  are

$$\overline{a}_{13}(X) = (a_{13,C} + a_{13,N}) \times \frac{X}{2} + a_{13,0} \times (1 - X)$$

and

$$\overline{a}_{15}(X) = (a_{15,C} + a_{15,N}) \times \frac{X}{2} + a_{15,0} \times (1 - X)$$

Therefore, the fraction  $f_{\infty}(^{13}\text{C}^{15}\text{N}) = \overline{a}_{13}(X) \times \overline{a}_{15}(X)$  is

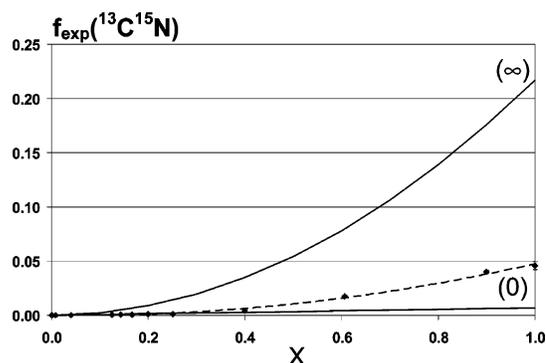
$$f_{\infty}(^{13}\text{C}^{15}\text{N}) = (a_{13,C} + a_{13,N}) \times (a_{15,C} + a_{15,N}) \times \frac{X^2}{4} + (a_{13,0} \times a_{15,0}) \times (1 - X)^2 + \left[ (a_{13,C} + a_{13,N}) \times \frac{a_{15,0}}{2} + (a_{15,C} + a_{15,N}) \times \frac{a_{13,0}}{2} \right] \times X \times (1 - X)$$

a quadratic function of  $X$ . Numerically  $f_{\infty}(^{13}\text{C}^{15}\text{N}) = 0.2101X^2 + 6.718 \times 10^{-3}X + 4.026 \times 10^{-5}$ . In this case, the maximum value of  $f_{\infty}(^{13}\text{C}^{15}\text{N})$  for  $X = 1$  is 0.2168 (21.68%); it would be exactly 25% for a 1–1 mixture of “isotopically pure” labeled glycines (100%  $^{15}\text{N}$ , 100%  $^{12}\text{C}$ ) and (100%  $^{14}\text{N}$ , 100%  $^{13}\text{C}$ ).  $f_{\infty}(^{13}\text{C}^{15}\text{N})$  is plotted as a function of  $X$  in Figure 2.

The experimental values  $f_{\text{exp}}(^{13}\text{C}^{15}\text{N})$  clearly lie between those of  $f_0(^{13}\text{C}^{15}\text{N})$  and  $f_{\infty}(^{13}\text{C}^{15}\text{N})$  (Figure 2). This proves that intermolecular recombination between the  $^{13}\text{C}$  and  $^{15}\text{N}$  isotopes occurred; that is, recombination occurred even when these isotopes were distant from one another (the experimental points would fit the unperturbed recombination curve if only intramolecular recombination had occurred). The intermediate position of the experimental values also shows that the probability of recombination decreased when the distance between the atoms increased (otherwise these values would lie on the averaged recombination curve). Hence, the experimentally determined fraction of  $^{13}\text{C}^{15}\text{N}$  was not merely a function of the proportion of  $^{13}\text{C}$  and  $^{15}\text{N}$  isotopes in the sample but was also a function of the distance between them.

**Recombinant CN Ions in the Case of Protein Films.** To demonstrate that the recombination between  $^{13}\text{C}$  and  $^{15}\text{N}$  in protein films was also distance-dependent, the same approach was adopted as used for the glycine crystals: i.e., the experimental values  $f_{\text{exp}}(^{13}\text{C}^{15}\text{N})$  were compared to the calculated values  $f_0(^{13}\text{C}^{15}\text{N})$  and  $f_{\infty}(^{13}\text{C}^{15}\text{N})$ .

Again, consider first the unperturbed recombination case. In these experiments, the CN secondary ions were collected from an area that contained a large number of proteins, each containing hundreds of nitrogen atoms and thousands of carbon atoms. This means, for example, that in the unperturbed recombination case, a fraction  $a_{13,C} \times a_{15,C}$  of the  $^{13}\text{C}^{15}\text{N}$  ions are collected from a  $^{13}\text{C}$ -labeled protein (a statistical mean value). Therefore, considering the proportion (again a statistical



**Figure 3.** Fraction of  $^{13}\text{C}^{15}\text{N}$  in the mixture of CN recombinant ions in protein samples plotted as a function of the concentration factor  $X$ . Points: experimental data  $\pm$  standard error; curve (0), unperturbed limiting case  $f_0(^{13}\text{C}^{15}\text{N})$ ; curve ( $\infty$ ), averaged limiting case  $f_{\infty}(^{13}\text{C}^{15}\text{N})$ . The experimental points are fitted well ( $r^2 = 0.9958$ ) by a quadratic function (dashed line) and are further from the quadratic limiting curve  $f_{\infty}(^{13}\text{C}^{15}\text{N})$  than in the case of glycine samples (see Figure 2).

mean value) of the three classes of proteins in a film, it is easily verified that

$$f_0(^{13}\text{C}^{15}\text{N}) = a_{15,0} \times a_{13,0} + \left[ \frac{(a_{15,N} \times a_{13,N} + a_{15,C} \times a_{13,C})}{2} - a_{15,0} \times a_{13,0} \right] \times X$$

which is similar to the case of the glycine crystals. Numerically, however,

$$f_0(^{13}\text{C}^{15}\text{N}) = 4.026 \times 10^{-5} + 6.582 \times 10^{-3} \times X$$

As with the glycine crystals, similar relationships for the fractions of  $^{13}\text{C}^{15}\text{N}$  in the averaged recombination case can be derived:

$$f_{\infty}(^{13}\text{C}^{15}\text{N}) = (a_{13,C} + a_{13,N}) \times (a_{15,C} + a_{15,N}) \times \frac{X^2}{4} + (a_{13,0} \times a_{15,0}) \times (1 - X)^2 + \left[ (a_{13,C} + a_{13,N}) \times \frac{a_{15,0}}{2} + (a_{15,C} + a_{15,N}) \times \frac{a_{13,0}}{2} \right] \times X \times (1 - X)$$

Numerically, this equation is  $f_{\infty}(^{13}\text{C}^{15}\text{N}) = 0.2085X^2 + 6.582 \times 10^{-3}X + 4.026 \times 10^{-5}$ .

The experimental values  $f_{\text{exp}}(^{13}\text{C}^{15}\text{N})$  lie between the calculated values  $f_0(^{13}\text{C}^{15}\text{N})$  and  $f_{\infty}(^{13}\text{C}^{15}\text{N})$  as a function of  $X$  (Figure 3). Hence, the fraction of  $^{13}\text{C}^{15}\text{N}$  obtained experimentally was, as with the glycine crystals, not just a function of the proportion of  $^{13}\text{C}$  and  $^{15}\text{N}$  isotopes but also dependent on the distance between these isotopes.

**Phenomenological Model To Quantify the Recombination Process.** Our experimental results show that the recombination of carbon and nitrogen atoms initially contained in two distinct labeled (macro)molecules is a function of the distance between these (macro)molecules. Evaluation of the yield of recombination as a function of the distance between the (macro)molecules requires a quantitative model. Indeed, phenomenological models that permit the experimental determination of operational parameters often prove useful in interpreting SIMS experiments.<sup>25</sup> Quantitative models such as TRIDYN have also been used to simulate processes occurring at surfaces during primary ion bombardment and to compare such data with the experimental results.<sup>26</sup> Here we report related approaches, based on

experiments and simulation, to the process of carbon and nitrogen ion recombination during primary ion bombardment. The results obtained by these approaches are analyzed and compared.

*Modeling Alteration of the Surface under the Primary Beam by Convolution.* Here we develop a phenomenological model in which the local concentration of a given isotope at a point of the sample surface is changed under the bombardment to a 2D Gaussian distribution centered on this point (Gaussian point spread function). It is this new distribution that determines recombination during sputtering. If the initial concentration of a given isotope (mass number  $A$ ) at the surface is modeled by a continuous function,  $C_A(x,y)$ , the new surface concentration obtained after the bombardment,  $C_A^*(x,y)$ , can therefore be modeled by the 2D convolution

$$C_A^*(x,y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} C_A(x',y')G(x-x',y-y') dx' dy'$$

where  $G(x,y)$  is the normalized 2D symmetrical Gauss function defined by

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$

In the following,  $C_A^*(x,y)$ , the beam-altered concentration, is termed the convolved concentration. We assume that  $\sigma$  is identical for all the isotopes that we consider in this work. A justification for this assumption is that the masses of these isotopes are close to one another (range 12–15) and that the energies of the C–C and C–N bonds are similar. Physically, the convolution equation represents the spreading (due to the mixing effect) of atoms initially located at a “point”. The Gauss function shows that 95% of the atoms initially located at a given point on the sample surface are, after mixing, contained in a disk of radius  $2\sigma$ . This allows a quantitative definition of the recombination radius as  $R_c = 2\sigma$  or, indeed, of the recombination diameter as  $\phi_C = 4\sigma$  (other definitions are possible; e.g.,  $R_c = 2.6\sigma$  or  $\phi_C = 5.2\sigma$ , which corresponds to the size of a disk containing 99% of the atoms initially concentrated at its center).  $\sigma$  should be expected to depend on, for example, the chemical composition and crystalline nature of the material.

*Idealized Model of the Initial Distribution of the Molecules at the Sample Surface.* The above 2D convolution allowed a map of the surface concentration  $C_A^*(x,y)$  of each of the carbon and nitrogen isotopes to be calculated from that of the initial concentration  $C_A(x,y)$ . The drying processes used for the sample preparation produced a random distribution and orientation of the different classes of proteins or glycine crystals. The size of the probe used in the SIMS analyses was big enough to allow the simultaneous bombardment of  $10^5$ – $10^7$  (macro)molecules. Under the beam, the three classes of (macro)molecules were therefore distributed in all ways possible. In other words, the measured secondary current was a statistical mean value. Hence, the real, detailed concentration distribution can be replaced in the model by a uniform, statistical distribution of the different types of molecules in the sample.

The shape and the packing of the (macro)molecules in the model was simplified by considering that the sample surface was flat, infinite, and comprised a regular arrangement (like a chessboard) of square molecules of identical size  $\phi$  centered on nodes of coordinates  $(x = i\phi, y = j\phi)$  with  $i, j \in ]-\infty, +\infty[$ . The molecule centered on  $(0,0)$  was termed the central molecule. The size of the proteins was chosen as  $\phi = 4.6$  nm (a typical value for the diameter of a 40-kDs globular protein). The size

of the glycine molecule was chosen as  $\phi = 0.43$  nm. This was because it was mainly the  $\alpha$ -polymorph of glycine that crystallized as monoclinic crystals under our experimental conditions of evaporation of an aqueous solution.<sup>27,28</sup> The crystallographic parameters of the unit cell are as follows:  $a = 0.5105$  nm,  $b = 1.1972$  nm,  $c = 0.5463$  nm, and  $\beta = 111.74^\circ$ .<sup>28,29</sup> The unit cell contains four glycine molecules. From these values, the volume of a glycine molecule was taken as  $1/4$  of the volume of the unit cell, i.e.,  $0.31/4 = 0.0775$  nm<sup>3</sup>, which corresponds to a cube of side  $\phi = 0.43$  nm.

We assumed either that every square area (centered on node  $(i,j)$ ) in the sample surface was actually occupied by a molecule or that some of them were void. We therefore assumed either a tight packing of the molecules with no empty spaces between them or that during the drying some areas remained unfilled with the molecules. In the following,  $\theta$  is the fraction ( $0 \leq \theta \leq 1$ ) of the sample surface occupied by molecules, and again,  $\theta$  is a statistical mean value.

In our model, the concentrations of the carbon,  $C_C$ , and nitrogen,  $C_N$ , elements were assumed to be uniform within a molecule and identical everywhere in the sample. We also assumed that these values remained constant even after the mixing effects due to the bombardment. We thus supposed that the beam produced a pure isotopic mixing at constant elemental density. Initially, before the surface was bombarded, the concentration of a given carbon (mass number  $A = 12$  or  $13$ ) or nitrogen isotope (mass number  $A = 14$  or  $15$ ) in the molecule centered on a node of coordinates  $(x = i\phi, y = j\phi)$  was  $C_{A,ij}$ . This value is constant for  $x \in [(i - 1/2)\phi, (i + 1/2)\phi]$  and  $y \in [(j - 1/2)\phi, (j + 1/2)\phi]$  (i.e., for points on the molecule surface) and depends on the class of the molecule centered on node  $(i,j)$  (N- or C-labeled molecule or unlabeled molecule). A mathematical description of the distribution of molecules on the sample surface is given in Appendix B.

*Calculation of the Convolved Mean Concentration of Isotopes in the Central Molecule.* We show in Appendix C that at any point of coordinates  $(x,y)$  in the sample surface

$$C_A^*(x,y) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \alpha_{ij}(x,y)C_{A,ij}$$

with

$$\alpha_{ij}(x,y) = \frac{1}{4} E_i(x)E_j(y)$$

the function  $E_k(u)$  being defined as

$$E_k(u) = \operatorname{erf}\left(\frac{\sqrt{2}\phi}{2\sigma}\left[\frac{u}{\phi} - k + \frac{1}{2}\right]\right) - \operatorname{erf}\left(\frac{\sqrt{2}\phi}{2\sigma}\left[\frac{u}{\phi} - k - \frac{1}{2}\right]\right)$$

$k$  is a positive, null, or negative integer and

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp(-\xi^2) d\xi$$

is the classical error function. As we assumed a single value for  $\sigma$ , the coefficients  $\alpha_{ij}(x,y)$  are identical for all the isotopes that we consider in this work. Note that by definition,

$$\sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \alpha_{ij}(x,y) = 1$$

and therefore,  $C_A^*(x,y)$ , the convolved local concentration of the isotope of mass number  $A$ , appears as a weighted average

of the concentrations of this isotope in all molecules of the sample surface. Clearly, however, only the molecule that includes the coordinates  $(x, y)$  and the molecules in its vicinity contribute significantly to the local convolved concentration (the  $\alpha_{ij}(x, y)$  coefficients go rapidly to zero as the distance between the molecule located at the distant node  $(i, j)$  and the molecule located at  $x, y$  increases).

Consider now the central molecule (centered on node  $(0, 0)$ ) and a point with coordinates  $(x^0, y^0)$  on the surface of this molecule, i.e.,  $x^0$  and  $y^0 \in [-1/2\phi, 1/2\phi]$ . The convolved concentration for the isotope of mass number  $A$  can be written

$$C_A^*(x^0, y^0) = \alpha_{00}(x^0, y^0) \cdot C_{A,00} + \sum'_{i=-\infty}^{+\infty} \sum'_{j=-\infty}^{+\infty} \alpha_{ij}(x^0, y^0) C_{A,ij}$$

In this equation, the notation  $\sum'$  means that the summation is extended to all  $i$  and  $j$  values except  $i = 0$  and  $j = 0$ . The value  $C_A^*(x^0, y^0)$  depends on the surface distribution of the three classes of molecules surrounding the central molecule. For a given mean isotope concentration  $\overline{C_A}$  in the sample, there are numerous distributions that differ in the precise location of the C-, N-labeled and unlabeled molecules in the sample surface. Consider one of these distributions, labeled  $(k)$ ; the convolved isotope concentration for this particular distribution can be written

$$C_A^*(x^0, y^0)^{(k)} = \alpha_{00}(x^0, y^0) C_{A,00} + \sum'_{i=-\infty}^{+\infty} \sum'_{j=-\infty}^{+\infty} \alpha_{ij}(x^0, y^0) C_{A,ij}^{(k)}$$

Consider now  $N$  of these distributions of the same  $\overline{C_A}$ ; summing all the corresponding  $C_A^*(x^0, y^0)^{(k)}$  gives

$$\sum_{k=1}^N C_A^*(x^0, y^0)^{(k)} = N\alpha_{00}(x^0, y^0) C_{A,00} + \sum_{k=1}^N \sum'_{i=-\infty}^{+\infty} \sum'_{j=-\infty}^{+\infty} \alpha_{ij}(x^0, y^0) C_{A,ij}^{(k)}$$

which after inversion of the order of summation on the right-hand side and division of both sides by  $N$ , becomes

$$\frac{1}{N} \sum_{k=1}^N C_A^*(x^0, y^0)^{(k)} = \alpha_{00}(x^0, y^0) C_{A,00} + \sum'_{i=-\infty}^{+\infty} \sum'_{j=-\infty}^{+\infty} \alpha_{ij}(x^0, y^0) \left[ \frac{1}{N} \sum_{k=1}^N C_{A,ij}^{(k)} \right]$$

For a number of molecules, and therefore for a value of  $N$ , large enough, the left-hand side of this last equation is the mean effective convolved concentration of the isotope  $A$  at the point  $(x^0, y^0)$  in the central molecule (the mean is a  $(k)$ -mean, calculated between the distributions  $(k)$ ). More exactly

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N C_A^*(x^0, y^0)^{(k)} = \overline{C_A^*(x^0, y^0)}$$

and clearly

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N C_{A,ij}^{(k)} = \overline{C_{A,ij}}$$

Therefore, taking into account that  $\sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \alpha_{ij}(x, y) = 1$ ,

$$\overline{C_{A,T}^*(x^0, y^0)} = \alpha_{00}(x^0, y^0) C_{A,T,00} + [1 - \alpha_{00}(x^0, y^0)] \cdot \overline{C_A}$$

We have added the subscript T to signify that any isotopic type T of molecule (T = C, N, O) can be the central molecule. Finally, the convolved isotope concentration in the central molecule appears as a weighted average between the actual isotope concentration in the molecule and the mean isotope concentration in the sample. This conclusion was expected, but the model allows us to estimate a weight that is based on the characteristic mixing process in D-SIMS (note that the coefficient  $\alpha_{00}(x^0, y^0)$  depends on  $\phi/\sigma$ ).

*Calculation of the Mean Fraction  $\langle f(^A C^A N) \rangle$  of a Recombinant Ion in a Sample.* The mean fraction of the recombinant ions  $^A C^A N$  ( $A = 12$  or  $13$  and  $A' = 14$  or  $15$ ) collected at a point  $(x^0, y^0)$  on the surface of a central molecule of class T (with T = C, N for  $^{13}C$ -,  $^{15}N$ -enriched molecules or 0 for unlabeled molecules),  $f_{A,A',T}^*(x^0, y^0)$ , is

$$f_{A,A',T}^*(x^0, y^0) = \frac{\overline{C_{A,T}^*(x^0, y^0)}}{C_C} \times \frac{\overline{C_{A',T}^*(x^0, y^0)}}{C_N}$$

where  $C_C$  and  $C_N$  are the elemental concentrations of carbon and nitrogen, respectively, which are assumed constant (see the subsection on the idealized model of the initial distribution of the molecules at the sample surface). Therefore, with the result of the preceding subsection

$$f_{A,A',T}^*(x^0, y^0) = [\alpha_{00}(x^0, y^0) a_{A,T} + (1 - \alpha_{00}(x^0, y^0)) \overline{a_A}] \times [\alpha_{00}(x^0, y^0) a_{A',T} + (1 - \alpha_{00}(x^0, y^0)) \overline{a_{A'}}]$$

where, with our assumptions, the isotopic fractions are as follows:  $a_{A,T} = C_{A,T}/C_C$ ,  $\overline{a_A} = \overline{C_A}/C_C$  and  $a_{A',T} = \frac{C_{A',T}}{C_N}$ ,  $\overline{a_{A'}} = \frac{\overline{C_{A'}}}{C_N}$ . Note that the coordinate-dependence of  $f_{A,A',T}^*(x^0, y^0)$  is via the coefficient  $\alpha_{00}(x^0, y^0)$  and not via the isotopic fractions which are constant. From the entire surface of the central molecule of class T, we obtain the mean integrated fraction

$\langle f_{A,A',T}^* \rangle$  with

$$\langle f_{A,A',T}^* \rangle = \frac{1}{\phi^2} \int_{-\phi/2}^{\phi/2} \int_{-\phi/2}^{\phi/2} \overline{f_{A,A',T}^*(x^0, y^0)} dx^0 dy^0$$

For the entire sample, the fraction  $\langle f(^A C^A N) \rangle$  is

$$\langle f(^A C^A N) \rangle = \sum_T Q_T \langle f_{A,A',T}^* \rangle$$

the  $Q_T$  being the proportions of molecules of class T in the sample surface (same as in the bulk); i.e.,  $Q_C = Q_N = X/2$  and  $Q_0 = 1 - X$  (homogeneous and uniform surface). With these definitions and some algebra

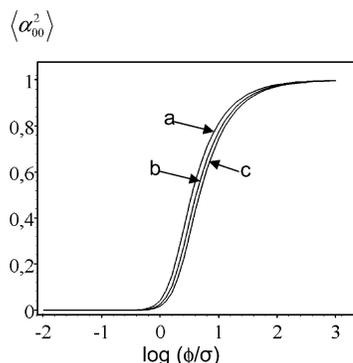
$$\langle f(^A C^A N) \rangle = f_{\infty}(^A C^A N) + \langle \alpha_{00}^2 \rangle [f_0(^A C^A N) - f_{\infty}(^A C^A N)]$$

In this expression, two limiting behaviors discussed above can be recognized:

$$f_{\infty}(^A C^A N) = \overline{a_A a_{A'}}$$

and

$$f_0(^A C^A N) = a_{A,C} a_{A',C} \frac{X}{2} + a_{A,N} a_{A',N} \frac{X}{2} + a_{A,0} a_{A',0} (1 - X)$$



**Figure 4.** Semilogarithmic plot of the mean square  $\alpha$ -coefficient  $\langle \alpha_{00}^2 \rangle$  as a function of the nondimensional variable  $\phi/\sigma$ . The different curves correspond to different values of the packing parameter  $\theta$  (fraction of sample surface occupied by the molecules). Curve a,  $\theta = 1$ ; curve b,  $\theta = 0.8$ ; curve c,  $\theta = 0.6$ .

The mean quadratic  $\alpha$  coefficient being defined by

$$\langle \alpha_{00}^2 \rangle = \frac{1}{\phi^2} \int_{-\phi/2}^{\phi/2} \int_{-\phi/2}^{\phi/2} \alpha_{00}^2(x^0, y^0) dx^0 dy^0$$

(the 2 is not a superscript and means that the square of  $\alpha$  is taken). This coefficient cannot be calculated explicitly from the definition of  $\alpha$  but can be calculated by numerical integration with MAPLE. Figure 4 shows the results of this calculation plotted (in semilogarithmic scale) as a function of the characteristic variable  $\phi/\sigma$  (see the curve for  $\theta = 1$ ; the other curves are discussed below). Two asymptotic behaviors are clearly seen. When  $\phi/\sigma < 1$ , the mean quadratic  $\alpha$  coefficient  $\langle \alpha_{00}^2 \rangle \approx 0$ ; this means that when  $\sigma$  increases ( $f(\text{A}^{\text{C}}\text{A}^{\text{N}}) \rightarrow f_{\infty}(\text{A}^{\text{C}}\text{A}^{\text{N}})$ ), as expected for the averaged behavior. Conversely, when  $\phi/\sigma > 10$ , the mean quadratic  $\alpha$  coefficient  $\langle \alpha_{00}^2 \rangle \approx 1$ ; this means that when  $\sigma$  decreases ( $f(\text{A}^{\text{C}}\text{A}^{\text{N}}) \rightarrow f_0(\text{A}^{\text{C}}\text{A}^{\text{N}})$ ), as expected for the unperturbed behavior. For intermediate values of  $\phi/\sigma$ , the mean fraction  $\langle f(\text{A}^{\text{C}}\text{A}^{\text{N}}) \rangle$  in the sample is simply a weighted average of the two limiting behaviors.

**Simplified Forms of the Model.** In the full expression just derived for  $\langle f(\text{A}^{\text{C}}\text{A}^{\text{N}}) \rangle$ , the quadratic terms in  $X^2$  are numerically dominant. Therefore, neglecting all the linear and constant terms in the expressions of  $f_{\infty}(\text{A}^{\text{C}}\text{A}^{\text{N}})$  and  $f_0(\text{A}^{\text{C}}\text{A}^{\text{N}})$  (and hence putting  $f_0(\text{A}^{\text{C}}\text{A}^{\text{N}})$  to zero as in the case of isotopically pure molecules), the result for  $\langle f(\text{A}^{\text{C}}\text{A}^{\text{N}}) \rangle$  can be simplified to

$$\langle f(\text{A}^{\text{C}}\text{A}^{\text{N}}) \rangle \approx \frac{1}{4} a_{13,\text{C}} a_{15,\text{N}} (1 - \langle \alpha_{00}^2 \rangle) X^2$$

Using a typical test value of  $\langle \alpha_{00}^2 \rangle = 0.5$ , the values of  $\langle f(\text{A}^{\text{C}}\text{A}^{\text{N}}) \rangle$  calculated with the complete and the simplified expressions differ, at least for  $X > 0.3$ , by only a few percent and by less than 2% for  $X \geq 0.7$  (not shown). Our model is not designed to give an accurate value of  $\sigma$  but, rather, to give an estimate of this value sufficient for deciding whether the recombination process in D-SIMS could be exploited to colocalize molecules in biological samples. Therefore, we shall use this simplified expression in the following. A determination to less than 10% of the value of  $\sigma$  is indeed coherent with the above, idealized, structure of the glycine and protein samples.

At this stage, two further modifications of the model may be made. The first modification concerns a loose packing of the molecules (or voids between the small crystals in the glycine samples). Reconsidering the reasoning done in the subsection on calculation of the convolved mean concentration of isotopes in the central molecule yields to the relation

$$\overline{C_{\text{A},\text{T}}^*(x^0, y^0)} = \alpha_{00}(x^0, y^0) C_{\text{A},\text{T},00} + [1 - \alpha_{00}(x^0, y^0)] \overline{C_{\text{A}} \theta}$$

where  $\theta$  is the fraction of surface occupied by the molecules. The introduction of this relationship in the definition of  $f_{\text{A},\text{A}',\text{T}}^*(x^0, y^0)$  gives

$$\overline{f_{\text{A},\text{A}',\text{T}}^*(x^0, y^0)} = \frac{\alpha_{00}(x^0, y^0) C_{\text{A},\text{T}} + [1 - \alpha_{00}(x^0, y^0)] \overline{C_{\text{A}} \theta}}{\alpha_{00}(x^0, y^0) C_{\text{C}} + [1 - \alpha_{00}(x^0, y^0)] C_{\text{C}} \theta} \times \frac{\alpha_{00}(x^0, y^0) C_{\text{A},\text{T}} + [1 - \alpha_{00}(x^0, y^0)] \overline{C_{\text{A}} \theta}}{\alpha_{00}(x^0, y^0) C_{\text{N}} + [1 - \alpha_{00}(x^0, y^0)] C_{\text{N}} \theta}$$

Defining

$$\alpha_{00}'(x^0, y^0) = \frac{\alpha_{00}(x^0, y^0)}{\alpha_{00}(x^0, y^0) + [1 - \alpha_{00}(x^0, y^0)] \theta}$$

and introducing this into  $\overline{f_{\text{A},\text{A}',\text{T}}^*(x^0, y^0)}$  gives

$$\overline{f_{\text{A},\text{A}',\text{T}}^*(x^0, y^0)} = [\alpha_{00}'(x^0, y^0) a_{\text{A},\text{T}} + (1 - \alpha_{00}'(x^0, y^0)) \overline{a_{\text{A}}}] \times [\alpha_{00}'(x^0, y^0) a_{\text{A}',\text{T}} + (1 - \alpha_{00}'(x^0, y^0)) \overline{a_{\text{A}'}}]$$

Note that when  $\theta = 0$ , i.e., when the environment of the central molecule is empty,  $\alpha_{00}'(x^0, y^0) = 1$  and therefore  $\overline{f_{\text{A},\text{A}',\text{T}}^*(x^0, y^0)} = C_{\text{A},\text{T}}/C_{\text{C}} \times C_{\text{A}',\text{T}}/C_{\text{N}}$ , which, as expected in this case, is the unperturbed recombination behavior for the central molecule.

Finally, we obtained for  $\overline{f_{\text{A},\text{A}',\text{T}}^*(x^0, y^0)}$  the same relation as that obtained in the case of tight packing ( $\theta = 1$ ) provided  $\alpha_{00}'(x^0, y^0)$  is substituted for  $\alpha_{00}(x^0, y^0)$ . Therefore, for  $\langle f(\text{A}^{\text{C}}\text{A}^{\text{N}}) \rangle$  the result is

$$\langle f(\text{A}^{\text{C}}\text{A}^{\text{N}}) \rangle \approx \frac{1}{4} a_{13,\text{C}} a_{15,\text{N}} (1 - \langle \alpha_{00}'^2 \rangle) X^2$$

The second modification takes into account the possibility that the recombination mechanism might contribute only partly to the formation of the secondary ions  $\text{A}^{\text{C}}\text{A}^{\text{N}}$ . Let  $\omega$  be the proportion of these secondary ions which are formed by a recombination mechanism ( $0 \leq \omega \leq 1$ ). In the simplified model, a negligible number of  $\text{A}^{\text{C}}\text{A}^{\text{N}}$  ions is locally formed; therefore, the final result for  $\langle f(\text{A}^{\text{C}}\text{A}^{\text{N}}) \rangle$  is simply

$$\langle f(\text{A}^{\text{C}}\text{A}^{\text{N}}) \rangle \approx \frac{\omega}{4} a_{13,\text{C}} a_{15,\text{N}} (1 - \langle \alpha_{00}'^2 \rangle) X^2$$

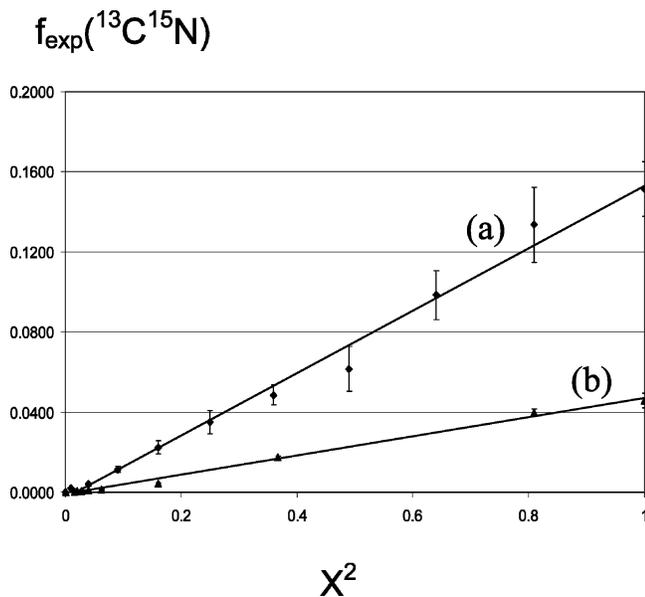
**Determination of a Diameter of Recombination from the Experimental Results.** The final relationship that we derived

$$\langle f(\text{A}^{\text{C}}\text{A}^{\text{N}}) \rangle \approx \frac{\omega}{4} a_{13,\text{C}} a_{15,\text{N}} (1 - \langle \alpha_{00}'^2 \rangle) X^2$$

shows that the plot of the fraction  $f_{\text{exp}}(\text{A}^{\text{C}}\text{A}^{\text{N}})$  as a function of  $X^2$  must be a straight line of slope

$$s = \frac{\omega}{4} a_{13,\text{C}} a_{15,\text{N}} (1 - \langle \alpha_{00}'^2 \rangle)$$

Figure 5 shows that there is an excellent agreement between this prediction of the model and the experimental results for both the glycine and protein samples. In calculating the value of  $\sigma$ ,  $\langle \alpha_{00}'^2 \rangle$  can be estimated from the above slope provided an estimate of  $\omega$  is given and then  $\langle \alpha_{00}'^2 \rangle$  can be obtained provided an estimate of  $\theta$  is given. Hence, the value of  $\sigma$ , can be obtained



**Figure 5.** Plots of the measured fraction of  $^{13}\text{C}^{15}\text{N}$  as a function of  $X^2$ , the square of the concentration factor. Curve a: glycine, slope 0.1516, regression coefficient  $r^2 = 0.9885$ . Curve b: proteins, slope 0.0468, regression coefficient  $r^2 = 0.9938$ .

**TABLE 1: Values of the Diameter of Recombination Determined for Glycine Samples**

$\theta^a$	$\omega^b$	$\phi_c^c = 4\sigma^d$ (nm)	$\phi_c = 5.2\sigma$ (nm)
1	1	0.5 <sub>8</sub>	0.7 <sub>6</sub>
0.8	1	0.6 <sub>6</sub>	0.8 <sub>6</sub>
0.6	1	0.7 <sub>8</sub>	1.0 <sub>1</sub>
1	0.9	0.9 <sub>3</sub>	1.2 <sub>1</sub>
1	0.75	1.3 <sub>8</sub>	1.8 <sub>0</sub>
0.8	0.9	0.8 <sub>0</sub>	1.0 <sub>4</sub>

<sup>a</sup>  $\theta$ , fraction of the sample surface occupied by the molecules. <sup>b</sup>  $\omega$ , proportion of CN secondary ions formed by recombination. <sup>c</sup>  $\phi_c$ , recombination diameter. <sup>d</sup>  $\sigma$ , dispersion in the Gauss function.

**TABLE 2: Values of the Diameter of Recombination Determined for Protein Samples**

$\theta^a$	$\omega^b$	$\phi_c^c = 4\sigma^d$ (nm)	$\phi_c = 5.2\sigma$ (nm)
1	1	1.6 <sub>2</sub>	2.1 <sub>0</sub>
0.8	1	1.8 <sub>7</sub>	2.4 <sub>3</sub>
0.6	1	2.2 <sub>6</sub>	2.9 <sub>3</sub>
1	0.9	2.0 <sub>8</sub>	2.7 <sub>0</sub>
1	0.75	2.5 <sub>3</sub>	3.2 <sub>9</sub>
0.8	0.9	2.0 <sub>8</sub>	2.7 <sub>0</sub>

<sup>a</sup>  $\theta$ , fraction of the sample surface occupied by the molecules. <sup>b</sup>  $\omega$ , proportion of CN secondary ions formed by recombination. <sup>c</sup>  $\phi_c$ , recombination diameter. <sup>d</sup>  $\sigma$ , dispersion in the Gauss function.

from Figure 4 (or more exactly the tabulated values of  $\langle\alpha_{00}^2\rangle$ , not shown). Table 1 for glycine samples and Table 2 for protein samples summarize the results for a range of values of  $\theta$  and  $\omega$ . Note that in these tables the value  $\theta = 0.8$  is very close to  $(1 - \pi/4) = 0.7854$ , which is the value for the packing coefficient for spheres of equal diameter. From Tables 1 and 2, it can be concluded that the recombination diameters are probably in the range 0.6–1.8 nm for the glycine samples and 1.6–3.2 nm for the proteins. In conclusion, final reasonable estimates are  $\phi_c = (1.0 \pm 0.5)$  nm for glycine and  $\phi_c = (2.0 \pm 0.5)$  nm for protein samples.

**Recombinant Yields Give Information about the Surfaces of Macromolecules in Contact.** The recombination process between the atoms of two interacting labeled molecules is clearly

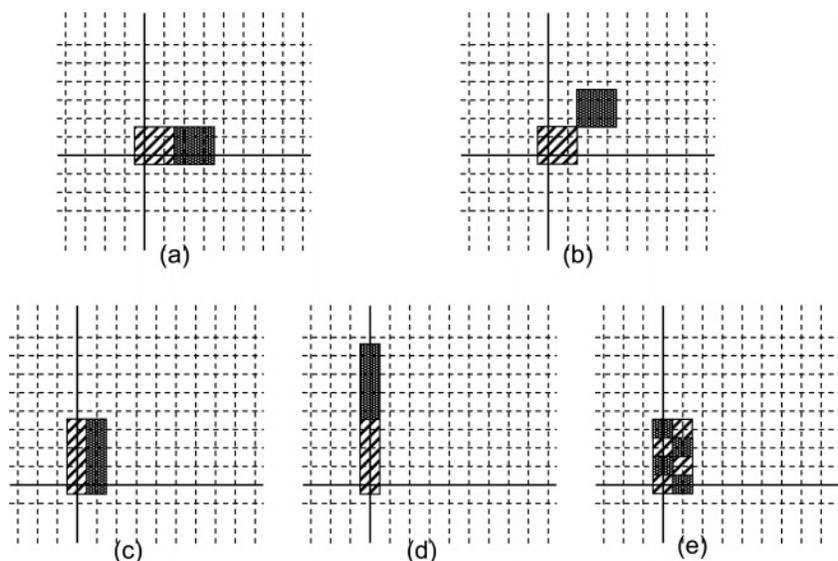
dependent on their shape and on the extent of their contact surface as illustrated in Figure 6 where the “molecules” have the same surface (25 nm<sup>2</sup>) but are of different shapes or have a different contact surface. In the case of recombination between the structures in Figure 6a, the intensity of the secondary  $^{13}\text{C}^{15}\text{N}$  current can be arbitrarily set to 100. In comparison, the current of  $^{13}\text{C}^{15}\text{N}$  is then 6, 213, 44, and 486 in Figure 6b–e, respectively. The simplified helical interaction shown in Figure 6e might correspond to interactions between coiled coil proteins or between two strands of nucleic acids. These values were calculated by convolution (see Appendix D for details) using  $\sigma = 0.5$  nm (recombination diameter 2 nm). In these examples and with the chosen values for  $\phi$  and  $\sigma$ , the intensity of the recombinant ion  $^{13}\text{C}^{15}\text{N}$  is approximately proportional to the contact surface of the two macromolecules. This is an important result since it means that the method might be adapted to give information on the surfaces involved in the interaction between two macromolecules both in vitro and in vivo. (It should be noted that the values in Table 2 were obtained using films composed of proteins that would have had a variety of topologies of contact between them due to stochastic processes during the drying of the samples.)

## Discussion

Both prokaryotic and eukaryotic cells are filled with structures variously termed assemblies, complexes, modules, and hyperstructures. These structures may comprise proteins, nucleic acids, lipids, and inorganic ions and the identification of these constituents is an ongoing challenge. D-SIMS permits a compositional map to be constructed from the secondary ions that result from the fragmentation of the molecules in a sample. Recent improvements in D-SIMS have increased the value of this technique in the localization of nucleic acids, proteins, and inorganic ions within biological systems<sup>16–18</sup> even though SIMS is not, of course, a panacea.

The formation of recombinant molecular ions is a disadvantage for the quantification of elements using static or dynamic SIMS.<sup>10</sup> The complex matrix effects leading to the formation of these ions make elemental quantification only possible after calibration with standards that are believed to produce matrix effects similar to those produced in the samples analyzed.<sup>30</sup> Here, however, we show that the formation of recombinant molecular ions can be a major advantage since they can be used as the basis for a new D-SIMS method for colocalization studies of nitrogen-containing macromolecules (proteins, DNA, RNA, etc.) on the nanometer scale. Other recombinant ions between carbon, oxygen, nitrogen, and hydrogen isotopes might also be used, but this remains to be investigated.

We estimated in this work, a diameter of recombination of C and N atoms in proteins of the order of magnitude of 2 nm; this value is lower than the distances over which the collision cascades occur. Indeed, using the TRIM free software,<sup>31</sup> it can be estimated with polymeric films of composition and density similar to that of the protein films that the lateral straggling distance for projectiles 16 keV in energy is 4 nm and that the depth of penetration can reach 14 nm. It should be noted that these values have been calculated for the static regime and may well be lower in D-SIMS operating in a steady state. Moreover, the lateral straggling distance for organic polymers only decreases by 20% when density is doubled.<sup>32</sup> The collision cascades are at the origin of numerous chemical modifications (recombination, cross-linking, dehydrogenation,<sup>33</sup> and atom displacements (mixing)). It is this modified surface which is then sputtered out. Under the impact of a primary ion, the



**Figure 6.** Examples of idealized C-labeled (100%  $^{13}\text{C}$  and  $^{14}\text{N}$ ) “molecule” (black hatched) and N-labeled (100%  $^{12}\text{C}$  and  $^{15}\text{N}$ ) “molecule” (black dots) in interaction. The molecules are merged in unlabeled molecules (100%  $^{12}\text{C}$  and  $^{14}\text{N}$ ) not represented in the figure. In examples a–e, the molecules have different shapes but identical surfaces; the length of their contact boundary is also different. Cases a and b represent two otherwise identical molecules with a high (a) or low (b) contact surface. Cases c and d: contacts are made either laterally or end-to-end. Case e is a representation of two molecules in helical interaction. The dashed lines of the grid are 2.5 nm apart. The x- and y-axis are represented in continuous line. Each labeled molecule spans 4 squares in the grid; i.e., it has a surface of 25 nm<sup>2</sup>. The intensities of the secondary  $^{13}\text{C}^{15}\text{N}$  currents were calculated by convolution (see Appendix D for details). The intensity for case a is arbitrarily 100, and the other values are 6, 213, 44, and 486 for (b)–(e), respectively. These values were computed using  $\sigma = 0.5$  nm (recombination diameter 2 nm).

recombinant secondary ions, considered in this work, are formed between sputtered atoms, which were located in the top first nanometer beneath this modified surface. This shows that the physical significance of the parameter  $\sigma$  in our model is complex: it integrates the mixing and chemical modification processes and the secondary recombination process. It is why our model is essentially phenomenological. Nevertheless, this model shows that  $\sigma$  is a valid operational parameter for determining whether two molecules or macromolecules are in close contact (rather than expressing  $\sigma$  in terms of specific physicochemical processes).

Intriguingly, the recombination diameter of the protein samples was twice that of the glycine samples. It cannot be totally excluded that this difference is due, at least in part, to the assumptions of the model. Indeed, our model is a continuous model in which the sample surface is described by maps of the local concentrations of the different isotopes. Physically, the local concentration reflects the superficial atom (isotope) density. When the surface of the molecule is very small, as in the case of glycine molecules, the density is the ratio of a small number of atoms (1 for nitrogen and 1 or 2 for carbon) to the surface area of the glycine molecule. Clearly, in this case, the assumption of uniform concentration (before bombardment) at each mathematical point of the exposed molecule surface is questionable. Therefore, the operational value of the size  $\phi$  of the glycine molecule that must be used to calculate  $\sigma$  from the experimentally determined value of  $\phi/\sigma$  may be different from the value we used in this work. This introduces additional uncertainty in the determination of the recombination diameter in the glycine samples. A physical explanation for the difference between the values of the recombination diameters in the glycine and protein samples might be that the packing of the atoms in the crystal is denser than that in the amorphous protein films such that the hydrogen bonds between each glycine molecule and those in its vicinity in the crystal<sup>28</sup> reduce the mixing.

Whatever the difficulties of a precise interpretation of the significance of the parameter  $\sigma$ , the size of the recombination area that we determined is a priori well suited to the determi-

nation of the proximity of two proteins in a cell. Such localization could also be combined with the localization of other isotopically labeled, cellular constituents such as lipids or nucleic acids. The isotopic labeling that the D-SIMS colocalization method needs may be applied in several ways: (1) classical immunocytochemistry might be extended and several, isotopically distinct antibodies to different proteins could be used in the same experiment to show that these proteins are present in the same complex at the same time, (2) isotopically distinct macromolecules might be introduced into permeabilized cells, and (3) synthesis of RNA or protein using drug-resistant polymerases might be combined with addition of isotopically distinct precursors.<sup>34</sup>

Intra-macromolecular conformational changes during protein folding or during interaction of a protein with some structure might also be revealed by D-SIMS colocalization since, if the mean contact surface between interacting macromolecules changes, the yield of recombinant ions also changes. For example, if the mean distance and the surface of contact between two, sufficiently long, amino acid sequences in the protein, one labeled with  $^{13}\text{C}$  and the other with  $^{15}\text{N}$ , change significantly during the folding or the interaction, this should result in an alteration of the proportion of  $^{13}\text{C}^{15}\text{N}$  ions produced (see Figure 6 that suggests that this alteration can be by 2 orders of magnitude just for topological reasons). This possibility is supported by the finding that in polyglycine samples the probability of recombination between C and N atoms is higher when these atoms are adjacent.<sup>20</sup>

The D-SIMS method of colocalization has the following advantages: (i) the possibility of studying colocalization of two labeled molecules in their “native” state, that is, without previous tagging or chemical modification of these molecules, (ii) the alternative or complementary possibility of studying the colocalization of two unlabeled molecules via labeled specific probes directed against these molecules, (iii) a very low limit of detection, (iv) the possibility, using the performances of the NanoSIMS 50 instrument,<sup>16,21</sup> to localize the set of two (or more) interacting macromolecules in the cell with a spatial resolution

of a few tens of nanometers, and (v) the simultaneous detection of assemblies of (macro)molecules labeled by different isotopes provided that several molecular recombinant secondary ions can be used. It may therefore be anticipated<sup>35</sup> that the D-SIMS method will become a valuable tool in colocalization studies in systems derived from both material sciences and biology.

## Conclusions

Localizing two or more components of assemblies in biological systems requires both continued development of fluorescence techniques and invention of entirely new techniques. The latest generation of D-SIMS, the Cameca NanoSIMS 50, permits the localization of specific, isotopically labeled molecules and macromolecules in sections of biological material with a resolution in the tens of nanometers and with a theoretical limit of detection approaching that of a single protein. Here we have shown that the formation of recombinant CN ions under the primary beam allows macromolecules to be colocalized if they are within 2 nm of one another. The advantages of a colocalization method based on D-SIMS would also include (1) the absence of any requirement for addition of tags or for chemical or structural modification of molecules, (2) the possibility of studying the simultaneous colocalization of two or more molecules, and (3) information about the surfaces of the molecules in contact.

## Appendix A

$a_{13,C}; a_{13,N}; a_{13,0}$	fraction of the isotope $^{13}\text{C}$ in $^{13}\text{C}$ -, $^{15}\text{N}$ -labeled and unlabeled molecules, respectively
$a_{15,C}; a_{15,N}; a_{15,0}$	fraction of the isotope $^{15}\text{N}$ in $^{13}\text{C}$ -, $^{15}\text{N}$ -labeled and unlabeled molecules, respectively
$a_{A,T}$	fraction of the isotope of mass $A$ ( $A = 12, 13$ for carbon and $A = 14, 15$ for nitrogen) in a molecule of isotopic type, or class $T$ (with $T = \text{C}, \text{N}$ for $^{13}\text{C}$ -, $^{15}\text{N}$ -enriched molecules or 0 for unlabeled molecules)
$\overline{a_{13}}(X); \overline{a_{15}}(X)$	mean fraction of $^{13}\text{C}$ and $^{15}\text{N}$ in a sample of concentration factor $X$
$C_C; C_N$	constant elemental concentrations of carbon and nitrogen, respectively
$C_A(x,y)$	initial (i.e., before alteration by the primary beam) concentration of the isotope of mass number $A$ at a point with arbitrary coordinates $(x,y)$ on the sample surface
$C_{A,ij}; C_{A,ij}^{(k)}$	constant initial concentration of the isotope of mass number $A$ in the molecule centered on node $(i,j)$ on the sample surface; same, but with the specification of a particular distribution $(k)$ of the different isotopic classes of molecules for a given value of $C_A$ (see below)
$C_A^*(x,y)$	convolved (beam-altered) concentration of the isotope of mass number $A$ at a point with arbitrary coordinates $(x,y)$ on the sample surface
$C_A^*(x^0, y^0)$	convolved concentration of the isotope of mass number $A$ at a point with coordinates $(x^0, y^0)$ on the surface of the central molecule
$C_A^*(x^0, y^0)^{(k)}$	same, but with the specification of a particular spatial distribution $(k)$ of the different isotopic classes of molecules for a given value of $C_A$
$\overline{C_A}$	mean concentration of the isotope of mass number $A$ in a sample
$\overline{C_{A,T}^*(x^0, y^0)}$	$(k)$ -mean convolved (the mean is taken between the distributions $(k)$ ) concentration of the isotope of mass number $A$ at the point $(x^0, y^0)$ in the central molecule

$f_{\text{exp}}(^{13}\text{C}^{15}\text{N})$	experimental value of the fraction of $^{13}\text{C}^{15}\text{N}$ in the mixture of the CN secondary ions
$f_0(^{13}\text{C}^{15}\text{N})$	theoretical value of the fraction of $^{13}\text{C}^{15}\text{N}$ in the mixture of the CN secondary ions in the case of exclusively unperturbed recombination
$f_{\infty}(^{13}\text{C}^{15}\text{N})$	theoretical value of the fraction of $^{13}\text{C}^{15}\text{N}$ in the mixture of the CN secondary ions in the case of averaged recombination
$\overline{f_{A,A',T}^*(x^0, y^0)}$	$(k)$ -mean fraction of the recombinant ions $^A\text{C}^A\text{N}$ ( $A = 12$ or $13$ and $A' = 14$ or $15$ ) collected at a point $(x^0, y^0)$ on the surface of a central molecule of class $T$
$\overline{\langle f_{A,A',T} \rangle}$	mean convolved fraction of $^A\text{C}^A\text{N}$ calculated on the surface of a central molecule of class $T$
$\langle f(^A\text{C}^A\text{N}) \rangle$	mean calculated fraction of $^A\text{C}^A\text{N}$ in a sample; this value is to be compared with $f_{\text{exp}}(^A\text{C}^A\text{N})$
$I(^A\text{C}^A\text{N})$	intensity of the secondary current of the recombinant ion $^A\text{C}^A\text{N}$
$Q_T$	proportion of molecules of class $T$ in a sample
$X$	concentration factor of a mixture obtained after adding $(1 - X)$ volume of a stock solution of unlabeled molecules (concentration $C$ ) to $X$ volume of a stock solution of $^{13}\text{C}$ - and $^{15}\text{N}$ -labeled molecules (both at concentration $C/2$ in the stock solution)
$\alpha_{ij}(x, y)$	weight for the calculation of the convolved concentration
$\alpha_{00}(x^0, y^0)$	weight for the calculation of the convolved concentration in the central molecule
$\alpha_{00}'(x^0, y^0)$	weight for the calculation of the convolved concentration in the central molecule taking into account the packing of the molecules
$\langle \alpha_{00}^2 \rangle$	mean quadratic $\alpha$ coefficient for $\langle f(^A\text{C}^A\text{N}) \rangle$ calculation
$\phi$	size of a molecule
$\phi_C$	recombination diameter
$\theta$	fraction of the sample surface occupied by the molecules
$\rho$	isotopic ratio $^{13}\text{C}/^{12}\text{C}$ in a sample ( $\rho = \overline{a_{13}/a_{12}}$ )
$\sigma$	dispersion in the Gauss function
$\omega$	proportion of CN secondary ions formed by recombination

## Appendix B

**Mathematical Description of the Idealized Initial Distribution of the Molecules at the Sample Surface.** The surface is assumed to be flat and infinite and is a regular array (like a chessboard) of square molecules. The molecules are all of the same size ( $\phi$  is the length of the side of the square) and are centered on points with coordinates  $(i\phi, j\phi)$  with  $i, j \in ]-\infty, +\infty[$ . The molecule centered at the origin ( $i = 0, j = 0$ ) is termed the central molecule. The concentration of a given carbon (mass number  $A = 12$  or  $13$ ) or nitrogen isotope (mass number  $A = 14$  or  $15$ ) in the molecule centered at a point of coordinates  $(i\phi, j\phi)$  is  $C_{A,ij}$ . This value is a constant (it can be zero) for  $x \in [(i - 1/2)\phi, (i + 1/2)\phi]$ , and  $y \in [(j - 1/2)\phi, (j + 1/2)\phi]$  (i.e., for points on the molecule's surface) and zero elsewhere.

Consider the following definitions. The asymmetrical unit step function

$$U_+(\xi) = \begin{cases} 0 & \text{if } \xi \leq 0 \\ 1 & \text{if } \xi > 0 \end{cases}$$

and the unit "crenel" function, in which  $m$  is a positive, null or negative integer

$$K_m(\xi) = U_+ \left( \xi - \left( m - \frac{1}{2} \right) \phi \right) \cdot \left[ 1 - U_+ \left( \xi - \left( m + \frac{1}{2} \right) \phi \right) \right] = \begin{cases} 1 & \text{if } \left( m - \frac{1}{2} \right) \phi < \xi < \left( m + \frac{1}{2} \right) \phi \\ 0 & \text{elsewhere} \end{cases}$$

With these definitions, the concentration of an isotope A in the sample surface at arbitrary coordinates  $(x, y)$  is

$$C_A(x, y) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} C_{A,ij} K_i(x) K_j(y)$$

## Appendix C

**2D Convolution of the Surface Concentration with a Gaussian PSF.** The atomic mixing in our model is described by a 2D convolution of the surface concentration  $C_A(x, y)$  of an isotope of mass number  $A$  with a Gaussian point spread function (PSF)

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$

The new concentration  $C_A^*(x, y)$  is given by

$$C_A^*(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} C_A(x', y') G(x - x', y - y') dx' dy'$$

The symmetrical 2D Gaussian function can be written as the product of two 1D normalized Gaussian functions

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} = \left( \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} \right) \times \left( \frac{1}{\sqrt{2\pi}\sigma} e^{-y^2/2\sigma^2} \right) = g(x)g(y)$$

Introducing in the convolution product the result of Appendix B for  $C_A(x, y)$  gives

$$C_A^*(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left[ \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} C_{A,ij} K_i(x) K_j(y) \right] \times g(x - x') g(y - y') dx' dy'$$

which, taking into account the properties of the functions  $K_i(x)$  and  $K_j(y)$  and inverting the order of integration and summation, becomes

$$C_A^*(x, y) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} C_{A,ij} \left[ \int_{(i-1/2)\phi}^{(i+1/2)\phi} g(x - x') dx' \right] \cdot \left[ \int_{(j-1/2)\phi}^{(j+1/2)\phi} g(y - y') dy' \right]$$

Hence after performing the integrations

$$C_A^*(x, y) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} C_{A,ij} \left[ \frac{1}{2} E_i(x) \right] \left[ \frac{1}{2} E_j(y) \right]$$

The functions  $E_k(u)$  are defined by

$$E_k(u) = \operatorname{erf} \left( \frac{\sqrt{2}\phi}{2\sigma} \left[ \frac{u}{\phi} - k + \frac{1}{2} \right] \right) - \operatorname{erf} \left( \frac{\sqrt{2}\phi}{2\sigma} \left[ \frac{u}{\phi} - k - \frac{1}{2} \right] \right)$$

(with  $k$  integer). In this expression,  $\operatorname{erf}(z)$  is the classical error function

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-\xi^2} d\xi$$

Then with the definition

$$\alpha_{ij}(x, y) = \frac{1}{4} E_i(x) E_j(y)$$

we obtain

$$C_A^*(x, y) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \alpha_{ij}(x, y) C_{A,ij}$$

It is easily verified using MAPLE that

$$\sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \alpha_{ij}(x, y) = 1$$

Physically, this is obvious because in the case of a sample that contains a single class of molecule all concentrations  $C_{A,ij}$  are equal

$$C_{A,ij} = C_A \quad \forall i, j$$

and therefore  $C_A^*(x, y)$  must also be equal to  $C_A$

$$C_A^*(x, y) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \alpha_{ij}(x, y) C_{A,ij} = C_A \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \alpha_{ij}(x, y) = C_A$$

## Appendix D

**Calculation of the  $^{13}\text{C}^{15}\text{N}$  Production from Two Interacting Labeled Molecules of Different Shapes and with Different Contact Surfaces.** In all examples of Figure 6, the  $^{13}\text{C}$  and  $^{15}\text{N}$  atoms are initially (before bombardment) contained in four elementary square areas centered on nodes  $(ij)$  as given below (see Figure 6).

example a:  $^{13}\text{C}$ , 00, 01, 10, 11;  $^{15}\text{N}$ , 20, 21, 30, 31

example b:  $^{13}\text{C}$ , 00, 01, 10, 11;  $^{15}\text{N}$ , 22, 23, 32, 33

example c:  $^{13}\text{C}$ , 00, 01, 02, 03;  $^{15}\text{N}$ , 10, 11, 12, 13

example d:  $^{13}\text{C}$ , 00, 01, 02, 03;  $^{15}\text{N}$ , 04, 05, 06, 07

example e:  $^{13}\text{C}$ , 00, 11, 02, 13;  $^{15}\text{N}$ , 10, 01, 12, 03

In these squares, the initial concentrations of  $^{13}\text{C}$  and  $^{15}\text{N}$  have the constant values  $C_{13}$  and  $C_{15}$ , respectively. Elsewhere these concentrations are null. Therefore, the convolved concentrations of  $^{13}\text{C}$  and  $^{15}\text{N}$ , derived in Appendix C, are simply written as

$$\text{example a: } C_{13}^*(x, y) = (\alpha_{00} + \alpha_{01} + \alpha_{10} + \alpha_{11}) C_{13} \text{ and } C_{15}^*(x, y) = (\alpha_{20} + \alpha_{21} + \alpha_{30} + \alpha_{31}) C_{15}$$

$$\text{example b: } C_{13}^*(x, y) = (\alpha_{00} + \alpha_{01} + \alpha_{10} + \alpha_{11}) C_{13} \text{ and } C_{15}^*(x, y) = (\alpha_{22} + \alpha_{23} + \alpha_{32} + \alpha_{33}) C_{15}$$

example c:  $C_{13}^*(x, y) = (\alpha_{00} + \alpha_{01} + \alpha_{02} + \alpha_{03})C_{13}$  and

$$C_{15}^*(x, y) = (\alpha_{10} + \alpha_{11} + \alpha_{12} + \alpha_{13})C_{15}$$

example d:  $C_{13}^*(x, y) = (\alpha_{00} + \alpha_{01} + \alpha_{02} + \alpha_{03})C_{13}$  and

$$C_{15}^*(x, y) = (\alpha_{04} + \alpha_{05} + \alpha_{06} + \alpha_{07})C_{15}$$

example e:  $C_{13}^*(x, y) = (\alpha_{00} + \alpha_{11} + \alpha_{02} + \alpha_{13})C_{13}$  and

$$C_{15}^*(x, y) = (\alpha_{10} + \alpha_{01} + \alpha_{12} + \alpha_{03})C_{15}$$

Let  $\alpha_{(A)}^{(u)}$  represent the sum of these four  $\alpha$ s for the isotope A (13 or 15) in the example  $u = (a)–(e)$ . At every point on the sample surface, the intensity of the  $^{13}C^{15}N$  secondary ion emission, in each of the examples  $u = (a)–(e)$ , is proportional to  $C_{13}^*(x, y)C_{15}^*(x, y)$ . Therefore, the intensity,  $I^{(u)}(^{13}C^{15}N)$ , for the entire surface is

$$I^{(u)}(^{13}C^{15}N) = K \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} C_{13}^*(x, y)C_{15}^*(x, y) dx dy = KC_{13}C_{15} \cdot \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \alpha_{(13)}^{(u)}(x, y)\alpha_{(15)}^{(u)}(x, y) dx dy$$

where  $K$  is a constant of proportionality. If we choose the example a as a reference and arbitrarily put the corresponding intensity to 100, the value of  $K$  is given by

$$K = \frac{100}{C_{13}C_{15} \cdot \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \alpha_{(13)}^{(a)}(x, y)\alpha_{(15)}^{(a)}(x, y) dx dy}$$

and hence

$$I^{(u)}(^{13}C^{15}N) = \frac{100 \cdot \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \alpha_{(13)}^{(u)}(x, y)\alpha_{(15)}^{(u)}(x, y) dx dy}{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \alpha_{(13)}^{(a)}(x, y)\alpha_{(15)}^{(a)}(x, y) dx dy}$$

With the explicit definitions of the  $\alpha_{ij}$  given in Appendix C, the numerical values of the  $I^{(u)}(^{13}C^{15}N)$  can be calculated with MAPLE. These values are given in the article itself.

**Acknowledgment.** Our SIMS analyzers were purchased with grants from the Région de Haute-Normandie, MIIAT, and Cameca (France). G.L. was supported by a grant from the Région de Haute-Normandie. A.D. was supported by a grant from “Association Vie et Espoir”. This work is a joint activity of the European Network of Excellence “Nanobeams”.

## References and Notes

- (1) Alberts, B. *Cell* **1998**, *92*, 291–294.
- (2) Hartwell, L. H.; Hopfield, J. J.; Leibler, S.; Murray, A. W. *Nature* **1999**, *402* (6761 Suppl), C47–52.
- (3) Norris, V.; den Blaauwen, T.; Doi, R. H.; Harshey, R.; Janniere, L.; Jimenez-Sanchez, A.; Jin, D. J.; Levin, P. A.; Mileykovskaya, E.; Minsky, A.; Misevic, G.; Ripoll, C.; Saier, M.; Skarstad, K.; Thellier, M.

*Annu. Rev. Microbiol.* In press; doi: 10.1146/annu. rev.mi-cro.61.081606.103348.

(4) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*. 2nd ed.; Plenum Publishing Corp.: New York, 1999.

(5) Babu Sekar, R.; Periasamy, A. *J. Cell Biol.* **2003**, *160*, 629–633.

(6) Verkhoturov, S. V.; Rickman, R. D.; Balderas, S.; Schweikert, E. A. *Appl. Surf. Sci.* **2004**, *231–232*, 113–116.

(7) Park, M. A.; Gibson, K. A.; Quinones, K.; Schweikert, E. A. *Science* **1990**, *248*, 988–990.

(8) Diehnelt, C. W.; English, R. D.; Van Stipdonk, M. J.; Schweikert, E. A. *Nucl. Instrum. Methods Phys. Res. B* **2002**, *193*, 883–890.

(9) Castaing, R.; Slodzian, G. *J. Microsc.* **1962**, *1*, 31–38.

(10) Benninghoven, A.; Rudenauer, F. G.; Werner, H. W. *Secondary Ion Mass Spectrometry: Basic Concepts, Instrumental Aspects, Applications and Trends*; John Wiley & Sons: Chichester, U.K., 1987.

(11) Chandra, S.; Smith, D. R.; Morrison, G. H. *Anal. Chem.* **2000**, *72*, 104A–114A.

(12) Thellier, M.; Derue, C.; Tafforeau, M.; Le Sceller, L.; Verdus, M. C.; Massiot, P.; Ripoll, C. *J. Trace Microprobe Tech.* **2001**, *19*, 143–162.

(13) Hindie, E.; Coulomb, B.; Beaupain, R.; Galle, P. *Biol. Cell.* **1992**, *74*, 81–88.

(14) Lhuissier, F.; Lefebvre, F.; Gibouin, D.; Demarty, M.; Thellier, M.; Ripoll, C. *J. Microsc.* **2000**, *198*, 108–115.

(15) Peteranderl, R.; Lechene, C. *J. Am. Soc. Mass Spectrom.* **2004**, *14*, 478–485.

(16) Guerin-Kern, J.-L.; Wu, T.-D.; Quintana, C.; Croisy, A. *Biochim. Biophys. Acta* **2005**, *1724*, 228–238.

(17) Lechene, C.; Hillion, F.; McMahon, G.; Benson, D.; Kleinfeld, A. M.; Kampf, J. P.; Distel, D.; Luyten, Y.; Bonventre, J.; Hentschel, D.; Park, K. M.; Ito, S.; Schwartz, M.; Benichou, G.; Slodzian, G. *J. Biol.* **2006**, *5*, 20.

(18) Kraft, M. L.; Weber, P. K.; Longo, M. L.; Hutcheon, I. D.; Boxer, S. G. *Science* **2006**, *313*, 1948–1951.

(19) Hindie, E.; Blaise, G.; Galle, P. In *Proceedings, 7th International Conference on Secondary Ion Mass Spectrometry, SIMS VII*; “Nanobeams” European Network of Excellence, <sup>5</sup> Epigenomics Programme, Genopole Benninghoven, A., Ed.; John Wiley & Sons: Chichester, U.K., 1990.

(20) McMahon, G.; Saint-Cyr, H. F.; Lechene, C.; Unkefer, C. J. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1181–1187.

(21) Slodzian, G.; Daigne, B.; Girard, F.; Boust, F.; Hillion, F. *Biol. Cell.* **1992**, *74*, 43–50.

(22) Roe, S. *Protein purification techniques*, 2nd ed.; Oxford University Press: New York, 2001.

(23) Lorin, J. C.; Havette, A.; Slodzian, G. In *Proceedings, 3rd International Conference on Secondary Ion Mass Spectrometry, SIMS III*; Benninghoven, A., Giber, J., Laszlo, J., Riedel, M., Werner, H. W., Eds.; Springer-Verlag: Berlin, 1982.

(24) Hauri, E. H.; Wang, J.; Pearson, D. G.; Bulanova, G. P. *Chem. Geol.* **2002**, *185*, 149–163.

(25) Cheng, J.; Wucher, A.; Winograd, N. *J. Phys. Chem. B* **2006**, *110*, 8329–8336.

(26) Philipp, P.; Wirtz, T.; Migeon, H.-N.; Scherrer, H. *Int. J. Mass Spectrom.* **2007**, *261*, 91–99.

(27) Boldyreva, E. V.; Drebuschak, V. A.; Drebuschak, T. N.; Paukov, I. E.; Kovalevskaya, Y. A.; Shutova, E. S. *J. Therm. Anal. Calorim.* **2003**, *73*, 409–418.

(28) Dawson, A.; Allan, D. R.; Belmonte, S. A.; Clark, S. J.; David, W. I. F.; McGregor, P. A.; Parsons, S.; Pulham, C. R.; Sawyer, L. *Cryst. Growth Des.* **2005**, *5*, 1414–1427.

(29) Boldyreva, E. V.; Drebuschak, T. N.; Shutova, E. S. *Z. Kristallogr.* **2003**, *218*, 366–376.

(30) Derue, C.; Gibouin, D.; Lefebvre, F.; Studer, D.; Thellier, M.; Ripoll, C. *Anal. Chem.* **2006**, *78*, 2471–2477.

(31) TRIM software. <http://www.SRIM.org/>.

(32) Biersack, J. In *Beam modification of materials. 2: Ion beam modification of insulators*; Mazzoldi, P., Arnold, G. W., Eds.; Elsevier Science Publishers: Amsterdam, 1987.

(33) Delcorte, A. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3395–3406.

(34) Studier, F. W.; Moffatt, B. A. *J. Mol. Biol.* **1986**, *189*, 113–30.

(35) Legent, G.; Delaune, A.; Tessier, C.; Misevic, G.; Norris, V.; Ripoll, C. *International Meeting SIMS Europe Abstracts*, Münster, Germany, 2006.