Rapid Identification of Probiotic Lactobacillus Biosurfactant Proteins by ProteinChip Tandem Mass Spectrometry Tryptic Peptide Sequencing

Gregor Reid,[1,2,3] Bing Siang Gan,[1,2,4,5,6] Yi-Min She,[7] Werner Ens,[7] Scot Weinberger,[8] and Jeffrey C. Howard[1,2,3,6,9]*

Lawson Health Research Institute[1] and Hand and Upper Limb Centre, Cell and Molecular Biology Laboratory,[6] St. Joseph's Health Care, and Departments of Surgery,[2] Microbiology and Immunology,[3] Biochemistry,[9] Pharmacology and Toxicology,[4] and Medical Biophysics,[5] University of Western Ontario, London, Ontario, and Department of Physics and Astronomy, University of Manitoba, Winnipeg, Manitoba,[7] Canada, and Ciphergen Biosystems, Inc., Fremont, California[8]

A novel ProteinChip-interfaced tandem mass spectrometer was employed to identify collagen binding proteins from biosurfactant produced by *Lactobacillus fermentum* RC-14. On-chip tryptic digestion of the captured collagen binding proteins resulted in rapid sequence identification of five novel tryptic peptide sequences via collision-induced dissociation tandem mass spectrometry.

Advances within the field of mass spectrometry (MS) have greatly accelerated proteomic research. Improvements in the ability of MS to rapidly identify and characterize proteins from complex biological materials have made it one of the most powerful tools for studying the function of genes (3, 5). One recent and exciting advance in MS instrumentation is the development of a tandem MS (MS/MS) system with a matrix-assisted laser desorption ionization source (2) that is capable of surface-enhanced laser desorption ionization (SELDI) ProteinChip analysis (4). Recently, the SELDI ProteinChip time-of-flight MS technique (SELDI-TOF-MS) was used to analyze the collagen binding (Cnb) protein content of biosurfactants (BSF) produced by probiotic strains of lactobacilli (1). As an extension to this methodology, we now report the use of a SELDI ProteinChip-interfaced tandem quadrupole TOF instrument (MALDI QSTAR, MDS Sciex, Concord, Ontario, Canada), termed ProteinChip LDI-Qq-TOF-MS (4), to perform MS/MS tryptic peptide sequencing of *Lactobacillus fermentum* RC-14 BSF Cnb proteins. The BSF produced by probiotic strains *Lactobacillus rhamnosus* GR-1 and *L. fermentum* RC-14 have been shown to inhibit the adhesion of several bacterial pathogens in vitro (6). However, only the RC-14-produced BSF has been shown to inhibit *Staphylococcus aureus* adhesion to surgical implants in vivo (7). In light of these findings, we have attempted in this study to identify potential RC-14 BSF factors that may be responsible for the observed antistaphylococcal activity.

As shown in Fig. 1, the ProteinChip LDI-Qq-TOF-MS machine consists of three quadrupole chambers and a TOF analyzer. Charged ions produced by the SELDI process at the ProteinChip interface enter the initial quadrupole (q0), where they are cooled by collisions with inert gas and ultimately guided into the quadrupole mass filter compartment (Q1).

When the device is operated in single-MS mode, quadrupoles Q1 and q2 are run in radio frequency-only mode to allow direct transmission of formed ions into the TOF analyzer. For MS/MS analysis, individual precursor species (present mass-to-charge [m/z] limit, ~3,000 Da) are selected by quadrupole mass filter Q1 and are transmitted into quadrupole assembly q2, which in this case functions as a collision cell. Transmitted ions collide with inert gas to undergo collision-induced dissociation (CID), which produces three primary characteristic fragment series: b ions, for which the charge is retained on the amino terminus; y ions, for which the charge is retained on the carboxy terminus; and internal product ions. CID fragments of the precursor ion are then analyzed by TOF-MS, which records the m/z distribution of the resulting mixture of product ions. Computer-assisted analysis of the resulting spectra helps to identify the peptide by using the obtained fragmentation spectrum to mine protein, cDNA, or expressed sequence tag databases. In the event of artifactual or nascent modifications of amino acids, these results may prove to be unfruitful. In such instances, the fragmentation spectrum is scrutinized for de novo sequence information among manifesting ion series by using either manual or computer-assisted algorithms.

To unambiguously identify the Cnb proteins by on-chip MS/MS tryptic peptide sequencing, we first captured the proteins by using preactivated surface arrays (PS1; Ciphergen Biosystems, Fremont, Calif.) containing either immobilized collagen type I (Cn-I) or Cn-III. Identical preparation and analysis procedures were performed for both types of collagen cross-linked PS1 arrays. Briefly, Cn-I or -III (Sigma, St. Louis, Mo.) was covalently cross-linked using the iminodiacetate chemistry of the PS1 ProteinChip by incubating each spot on the array with 3 μl of an aqueous solution containing 1 μg of collagen for 1 h at room temperature (RT) in a humidified chamber (HC). Any residual non-cross-linked areas on the arrays were inactivated by incubating each of the spots with 3 μl of an aqueous solution containing 1 M Tris (pH 7.5) for 30 min (RT, HC). After the arrays were washed in phosphate-buffered saline and high-pressure liquid chromatography-

* Corresponding author. Mailing address: Lawson Health Research Institute, St. Joseph's Health Care, Room G518, 268 Grosvenor St., London, Ontario N6A 4V2, Canada. Phone: (519) 646-6000, ext. 64167. Fax: (519) 646-6049. E-mail: jhoward@lri.sjhc.london.on.ca.
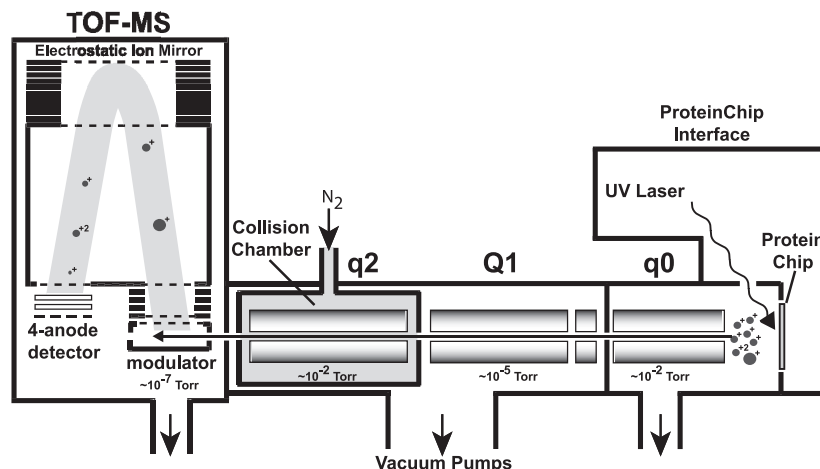
FIG. 1. ProteinChip LDI-Qq-TOF-MS system. ProteinChip SELDI is supported by a UV laser source. Ions created in q0 are injected into the quadrupole mass filter (Q1), where precursor ions are selected for subsequent CID fragmentation within the gas collision chamber (q2). Parent and product ions are then injected orthogonally into a modulator region of the TOF-MS compartment, where they are accelerated into an ion mirror (reflectron) and detected via a four-anode detector.
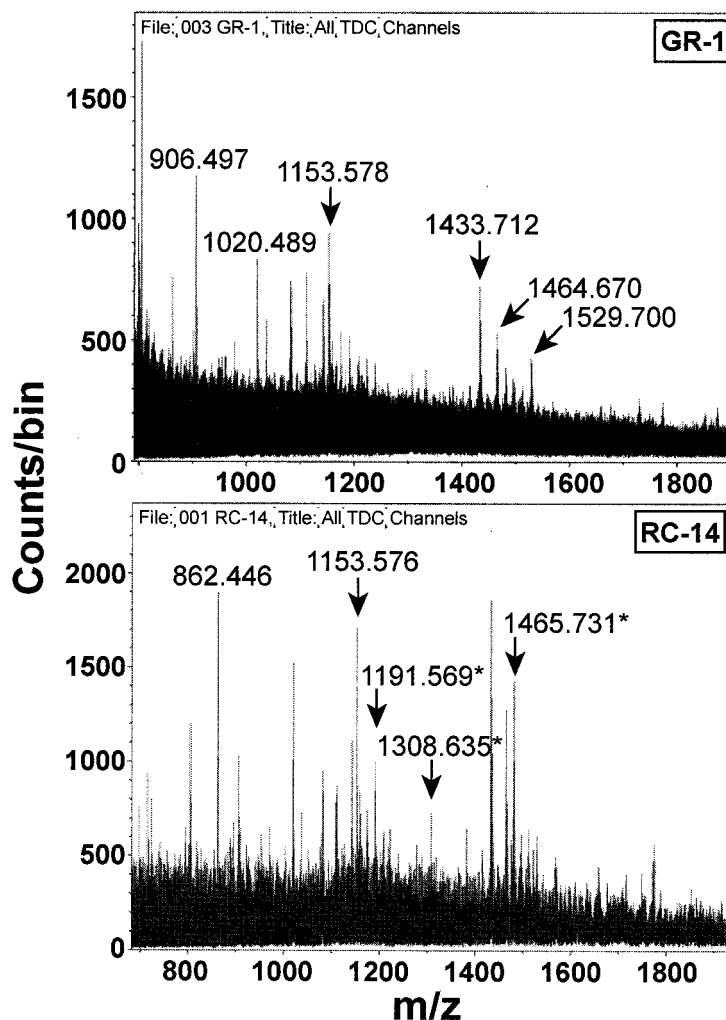

FIG. 2. Mass analysis of trypsin-digested Cnb proteins from *Lactobacillus* GR-1 (upper panel) and RC-14 BSF (lower panel). BSF Cnb proteins isolated from *L. fermentum* RC-14 and *L. rhamnosus* GR-1, using Cn-I cross-linked PS1 arrays, were on-chip digested with sequencing-grade modified trypsin. The resulting tryptic fragments were analyzed by the ProteinChip LDI-Qq-TOF instrument in single-MS mode to identify peptides that are unique to RC-14. These precursor fragments (*) were then chosen for CID-MS/MS analysis.
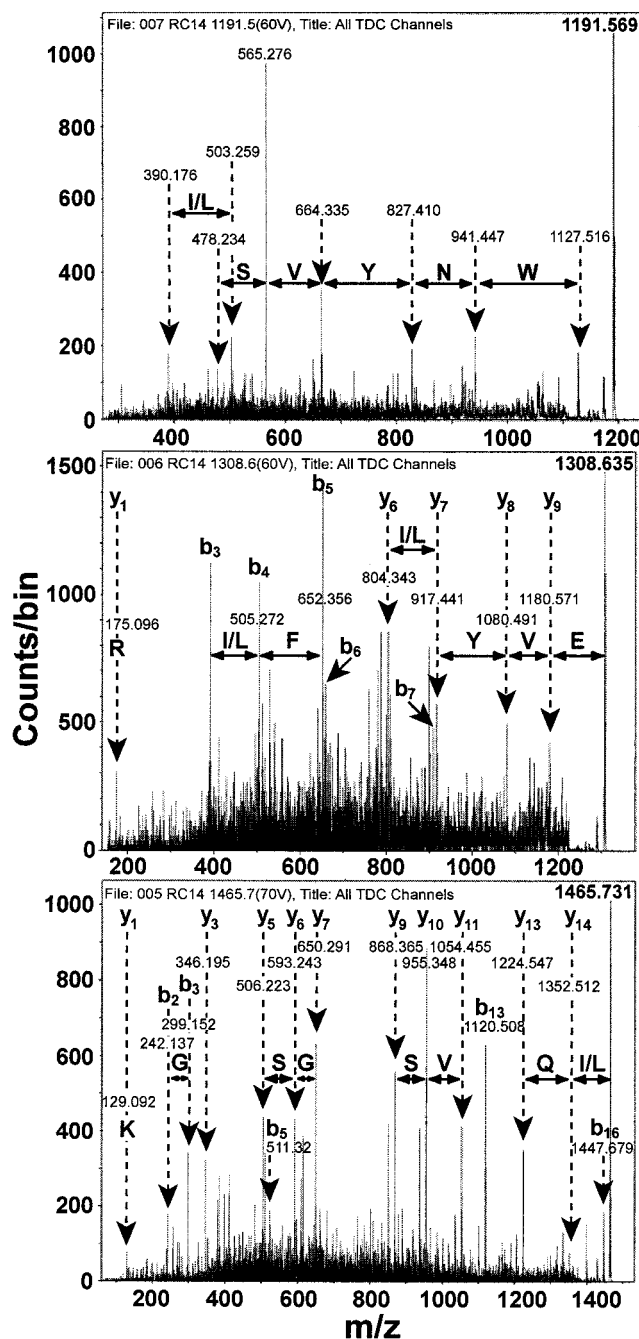
FIG. 3. CID-MS/MS analysis of tryptic fragments of Cnb proteins from *L. fermentum* RC-14 BSF. Tryptic peptides at *m/z* 1,191.569, 1,308.635, and 1,465.731 Da were subjected to CID fragmentation. The resulting MS/MS spectra were analyzed with the aid of the Predict Fragment routine from the SCIEX program BioMultiView in order to assign the b and y ion peaks. The amino acid sequence of each fragmented tryptic peptide is determined by calculating the differences in the masses of consecutive b or y ions (e.g., $y_{14} - y_{13}$).

grade H_2O (Sigma), each spot was incubated with 3 μl of a phosphate-buffered saline solution containing 1 μg of BSF (*L. fermentum* RC-14 or *L. rhamnosus* GR-1) (6) for 1 h (RT, HC). After the arrays were briefly washed with high-pressure liquid chromatography-grade H_2O, the captured Cnb proteins were subjected to on-chip digestion with 25 ng of sequencing-grade modified trypsin (Promega, Madison, Wis.) for 2 h (RT, HC). Limited trypsin digestion of the Cnb proteins was suffi-

cient for protein identification via MS/MS sequencing. After digestion, the arrays were air dried, spotted with 1 μl of a saturated solution of matrix (either α-cyano-4-hydroxycin-namic acid or sinapinic acid [obtained from Ciphergen Biosystems] prepared in an aqueous solution containing 50% aceto-nitrile and 0.5% trifluoroacetic acid), and allowed to air dry.

Initially, tryptic peptide maps of the Cnb proteins were generated with the ProteinChip LDI-Qq-TOF-MS in single-MS

TABLE 1. ProteinChip LDI-Qq-TOF-MS tryptic peptide sequence analysis

PS1 ProteinChip	Peptide mass (Da)	Amino acid sequence[a]	Database homology
Cn-I	1,191.569	LQGLVS[FA]GSCTCAQK	Yo1026cp (*S. cervevisiae*), probable membrane protein ($E = 31^b$)
	1,308.635	EVYLFQE[QP]R	NSH[c]
	1,465.731	WNYVS[d]	NSH
Cn-III	1,263.696	RAFALLSNWPVK	Cyclic β-1,2-glucan synthase (*Mesorhizobium loti*) ($E = 25$)
	1,329.685	GLQGGDL[SH][SH]R	NSH

[a] Underlining indicates, leucine-isoleucine uncertainty (identical mass), double underlining indicates potential GG = N (identical mass), and brackets indicate tentative order of sequence.
[b] Blastp search E value.
[c] NSH, no significant homologies.
[d] W oxidation prevented further sequence analysis.

mode. This allowed identification of unique tryptic peptides for subsequent CID-MS/MS sequencing. As shown in Fig. 2, using the Cn-I PS1 array, three tryptic fragments were chosen for CID-MS/MS sequencing. MS/MS spectra of the product ions for each of the three tryptic fragments (m/z = 1,191.569, 1,308.635, and 1,465.731 Da) were analyzed to assign the b and y ions with the aid of the Predict Fragment routine from the SCIEX program BioMultiView (Fig. 3).

In total, five tryptic peptide fragments analyzed by CID-MS/MS generated amino acid sequence information that was used to search available protein databases (Table 1). For three of the sequenced peptides, no significant homologies were found. Homology matches found for the remaining two sequences were from distinct bacterial or nonbacterial species with predicted Blastp E values of ≥ 25.

In conclusion, the ProteinChip MS/MS tryptic peptide sequencing technique is a very powerful method of rapidly identifying proteins. In particular, the use of biologically modified probe surfaces (i.e., PS1 or PS2 arrays) in combination with LDI-Qq-TOF-MS will permit rapid analysis and identification of physiologically important biomolecules (e.g., proteins, DNA, and RNA, etc.) that bind to proteins. The finding of five novel peptides from *L. fermentum* RC-14 may prove to be significant with respect to its probiotic activity.

REFERENCES

1. **Howard, J. C., C. Heinemann, B. J. Thatcher, B. Martin, B. S. Gan, and G. Reid.** 2000. Identification of collagen-binding proteins in *Lactobacillus* spp. with surface-enhanced laser desorption ionization–time-of-flight ProteinChip technology. Appl. Environ. Microbiol. **66**:4396–4400.
2. **Loboda, A. V., A. N. Krutchinsky, M. Bromirski, W. Ens, and K. G. Standing.** 2000. A tandem quadrupole/time-of-flight mass spectrometer with a matrix-assisted laser desorption/ionization source: design and performance. Rapid Commun. Mass. Spectrom. **14**:1047–1057.
3. **Mann, M., R. C. Hendrickson, and A. Pandey.** 2001. Analysis of proteins and proteomes by mass spectrometry. Annu. Rev. Biochem. **70**:437–473.
4. **Merchant, M., and S. R. Weinberger.** 2000. Recent advancements in surface-enhanced laser desorption/ionization-time of flight-mass spectrometry. Electrophoresis **21**:1164–1177.
5. **Pandey, A., and M. Mann.** 2000. Proteomics to study genes and genomes. Nature **405**:837–846.
6. **Reid, G., C. Heinemann, M. Velraeds, H. C. van der Mei, and H. J. Busscher.** 1999. Biosurfactants produced by *Lactobacillus*. Methods Enzymol. **310**:426–433.
7. **Reid, G., J. Howard, and B. S. Gan.** 2001. Can bacterial interference prevent infection? Trends Microbiol. **9**:424–428.