Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 10695–10698, September 1999
Biophysics

Toward the bilayer proteome, electrospray ionization-mass spectrometry of large, intact transmembrane proteins

J. P. WHITELEGGE*†‡§, J. LE COUTRE‡¶, J. C. LEE¶, C. K. ENGEL‖, G. G. PRIVÉ‖, K. F. FAULL†**, AND H. R. KABACK¶

*Department of Chemistry and Biochemistry, ¶Howard Hughes Medical Institute, Departments of Physiology, and Microbiology and Molecular Genetics, and the Molecular Biology Institute, **Department of Psychiatry and Biobehavioral Sciences, and †The Pasarow Mass Spectrometry Laboratory, University of California, Los Angeles, CA 90095; and ‖Ontario Cancer Institute, and Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada M2G 2M9

Contributed by H. R. Kaback, July 20, 1999

ABSTRACT **Genes encoding membrane proteins comprise a substantial proportion of genomes sequenced to date, but ability to perform structural studies on this portion of the proteome is limited. Electrospray ionization-MS (ESI-MS) of an intact protein generates a profile defining the native covalent state of the gene product and its heterogeneity. Here we apply ESI-MS technology with accuracy exceeding 0.01% to a hydrophobic membrane protein with 12-transmembrane α-helices, the full-length lactose permease from *Escherichia coli*. Furthermore, ESI-MS is used to titrate reactive thiols with *N*-ethylmaleimide. Treatment of the native protein solubilized in detergent micelles reveals only two reactive thiols, and both are protected by a substrate analog.**

Of the genomes sequenced thus far, 35% of the gene products are known or predicted to be polytopic transmembrane proteins. These proteins catalyze a multitude of essential functions, from oxidative phosphorylation and electron transfer to transport of molecules in and out of cells to signal transduction across membranes to bacterial motility. Many are important with regard to human disease (e.g., cystic fibrosis), and at least two of the most widely prescribed drugs in the U.S., Prozac and Prilosec, are targeted to membrane transport proteins. Although advances in molecular biology and biochemistry over the last 15 years have led to the characterization, purification, and modification of this class of proteins, only a handful have been crystallized in a form allowing for structure determination at atomic resolution that is essential for understanding mechanism of action. Furthermore, many membrane proteins require conformational flexibility to function, making it imperative to obtain dynamic information to fully understand function.

Unfortunately, protein primary structure predicted from gene sequencing is biased because of sequencing errors, RNA editing, and alternative splicing, as well as a range of post-translational modifications that often play essential roles in function. One way to monitor the proteome is through use of MS (1). The electrospray ionization-MS (ESI-MS) of an intact protein provides a unique image of native covalent state and heterogeneity, because of the accuracy and resolution of the technique (2–5). The measured mass sometimes agrees with the theoretical mass based on the gene sequence, but most often demands modification of the predicted structure until coincidence is achieved. Thus, in a first step to obtain proteomic information, the native mass of a protein is a vital statistic, but is found only in a few database entries. In this study, size-exclusion HPLC in acidic aqueous chloroform/methanol is coupled with online ESI-MS (size exclusion chromatography-MS) to obtain the mass spectrum of a 12-transmembrane helix transport protein, the lactose permease of *Escherichia coli*.

Lactose permease is a paradigm for secondary active transport proteins that convert the free energy stored in an electrochemical proton gradient into work in the form of a solute concentration gradient (6). The *lacY* gene has been cloned and sequenced, and each residue has been mutagenized. The permease has been overexpressed, purified, reconstituted into proteoliposomes, and shown to be active as a monomer. A battery of site-directed techniques has been developed to determine helix packing and the relationship between the six irreplacable amino acid side chains in the molecule (7). In addition, two-dimensional crystallography, attenuated total reflection-Fourier transform IR spectroscopy, and other biophysical studies indicate that the protein is loosely arranged to form a central aqueous pore that presumably represents the translocation pathway (8–10). However, the recombinant protein used for structural analyses has not been studied with respect to covalent structure or the occurrence of noncovalent, tightly bound lipids.

MATERIALS AND METHODS

Cell Growth, Membrane Preparation, and Purification of Native Permease. *E. coli* XL 1 blue was grown in fermentors by using enriched LB medium (90 g tryptone, 90 g NaCl, 45 g of yeast extract, 35 g KH_2PO_4, 50 g K_2HPO_4, trace metals in 10 liters of water). The fermentors were inoculated with an overnight culture of freshly transformed cells bearing the PTC189 + $6His_{417}$ expression vector encoding wild-type lacY with a C-terminal 6-His affinity tag. Cells were induced with 1 mM isopropyl β-D-thiogalactoside at OD_{600} 1.0 and harvested at an OD of approximately 2.0. Concentrated cells were washed once in cold 50 mM KPi (pH 7.5) and a second time in the same buffer plus 10 mM $MgSO_4$. To prepare membranes, cells were resuspended in cold, 50 mM KPi (pH 7.5), 10 mM $MgSO_4$, 0.5 mM Pefabloc, 10 mM β-mercaptoethanol (BME), and 30 μg/ml DNase. Cells were broken by using two consecutive passes through an Emulsiflex (Avestin, Ottawa, Canada) homogenizer (15,000–20,000 psi) and unbroken cells were removed by centrifugation (15 min at 10,000 $\times$ *g*). A membrane fraction was obtained from the supernatant by centrifugation (2 h at 150,000 $\times$ *g*). To remove peripheral membrane proteins, the membrane fraction was washed with urea (5 M urea, 50 mM KPi, pH 7.5, 0.5 mM Pefabloc, 10 mM BME), stirred for 30 min on ice,

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

Abbreviations: ESI-MS, electrospray ionization-MS; NEM, *N*-ethylmaleimide.

‡J.P.W. and J.l.C. contributed equally to this work.
§To whom reprint requests should be addressed at: Department of Chemistry, University of California, 405 Hilgard Avenue, Los Angeles, CA 90095. E-mail: jpw@chem.ucla.edu.

and centrifuged for 2 h at 170,000 × *g*. The supernatant was discarded, and the membranes were resuspended in cold 50 mM KPi (pH 7.2), 0.2 M NaCl, 10 mM BME, 0.25 mg/ml phospholipid (phosphatidylethanolamine/phosphatidylglycerol; 3:1, wt/wt) (1.0 ml buffer/g cells), and *n*-dodecyl-β-D-maltoside (DDM) was added to a final concentration of 2% while stirring on ice for 45 min. After removing insoluble material by centrifugation at 150,000 × *g* for 30 min, the supernatant was combined with washed Ni/nitrilotriacetic acid resin (Qiagen, Chatsworth, CA) (0.35 ml of resin/1.0 g of cells) and rotated for 2 h at 4°C. Using a glass column, the loaded Ni resin was washed thoroughly with 50 mM KPi (pH 7.2), 10% glycerol, 0.16% DDM, and subsequently with up to 50 mM imidazole in the same buffer. The protein was eluted upon increasing the concentration of imidazole to 200 mM and dialyzed overnight against 50 mM KPi (pH 7.5), 0.016% DDM, 1 mM EDTA. A recombinant form of the permease containing cytochrome b_{562} fused into the middle cytoplasmic loop and named "red" permease was purified with a similar protocol.

Sample Preparation for MS. All permease samples used for MS experiments displayed substrate binding activity by flow dialysis (28). Typically, aliquots of permease in detergent suspension were precipitated with $CHCl_3$/MeOH, essentially as described by Wessel and Flügge (29). An aliquot (100 μl) of the aqueous protein solution at 1–2 mg/ml protein was diluted 1:3 (vol/vol) with MeOH and mixed briefly. $CHCl_3$ (100 μl) was added, and mixing yielded a single phase. Phase separation was initiated by addition of water (200 μl) and vigorous mixing. The phases were separated by centrifugation (10,000 × *g*, for 2 min), yielding precipitated protein at the interface. The bulk of the upper aqueous methanol phase then was removed, and methanol (300 μl) was added. After mixing, the pellet remained insoluble in a single-phase mixture before it was recovered by centrifugation (10,000 × *g*, 1 min). The pellet was dried for 1–2 min with the tube inverted before dissolving the sample in 50 μl of 90% formic acid. Samples were subjected immediately to HPLC. For labeling with *N*-ethylmaleimide (NEM) the native, solubilized protein was incubated at a final concentration of 1 mM with NEM, and the reaction was terminated by addition of DTT to a final concentration of 10 mM.

HPLC Chromatography. For final purification, HPLC was used before ESI-MS in an in-line setup. The "white" permease was purified by using isocratic elution (0.5 ml/min), with a size exclusion column (G2000SW, TosoHaas, Montgomeryville, PA, 10 μM, 300 × 7.8 mm) in an aqueous organic solvent mixture (30). A mixture of degassed $CHCl_3$/MeOH/1% aqueous formic acid (4:4:1; vol/vol) was used as solvent. The column was equilibrated extensively and tested with blank injections of 90% formic acid before experiments. For "red" permease, a polystyrene/divinylbenzene column (PLRP/S, Polymer Labs, Amherst, MA) was substituted. Separations were performed at 40°C by using a modified ABI120A dual syringe pump machine equipped with a postdetector splitter for back-pressure regulation and fraction collection.

ESI-MS. ESI-MS analysis was performed by using a Perkin–Elmer Sciex API III+ triple quadrupole instrument operating in positive ion mode as described (2). An orifice potential of 75 V was found optimal for *E. coli* lactose permease.

RESULTS AND DISCUSSION

To achieve mass spectrometric analysis of a membrane protein, it is necessary to remove detergents and lipids used for solubilization (11–17). Chloroform/methanol/water phase separation was used to precipitate the permease, thereby removing detergent and residual phospholipid. The aggregated pellet can be dissolved in 90% formic acid, but is still accompanied by tightly bound lipids and residual detergent requiring further purification of the polypeptide before ionization. Because a single-phase mixture of chloroform/methanol/1% aqueous formic acid (4:4:1, vol/vol) effectively solubilizes certain intrinsic membrane proteins (18), the permease was transferred to this solvent mixture by using size-exclusion chromatography (Fig. 1). Both the UV and the total ion chromatogram show separation of material, which is not visualized by SDS/PAGE. On-line ESI-MS of the permease peak reveals a clean spectrum (Fig. 2) with an excellent signal-to-noise ratio.

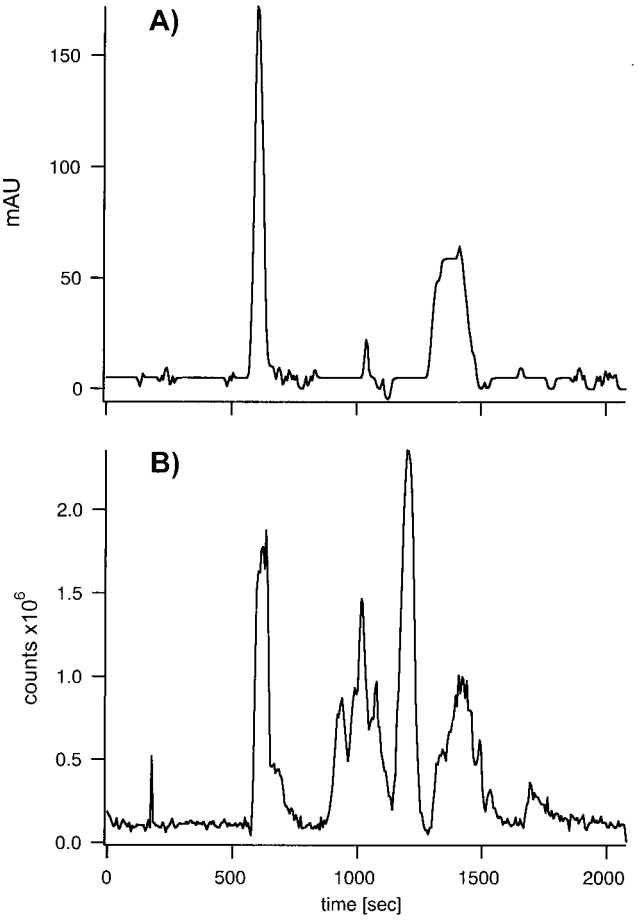

FIG. 1. Chromatographic purification of lactose permease by size-exclusion chromatography in aqueous organic solvent. Purified lactose permease (200 μg protein) was separated from nearly all noncovalently associated lipid and detergent to perform effective ESI-MS. The permease/detergent micelle was disrupted by chloroform/methanol phase separation with the protein precipitating at the interface. Aggregated protein was recovered by centrifugation and dissolved in 90% aqueous formic acid (50 μl), but was still accompanied by residual detergent and lipid. Size-exclusion chromatography in chloroform/methanol/1% aqueous formic acid (4:4:1, vol/vol) was used to separate the permease from these contaminants by using a silica column (G2000SW, 7.8 × 300 mm, TosoHaas) at 40°C and 0.5 ml/min. (*A*) Elution profile measured at 280 nm (mAU). Permease eluted at 600 sec as a single peak. The eluent is directed to the mass spectrometer, scanning the *m/z* range 600 to 2,400 every 6 sec. (*B*) Total ion chromatogram showing the signal detected for each scan versus time. Both chromatograms show separation from residual UV absorbing and ion-generating material.

The acidified aqueous organic solvent mixture is highly compatible with electrospray ionization, and permease ions carrying from 20 to 50 positive charges (H^+) are detected (Fig. 2). The ions fall into three distinct classes, which may indicate

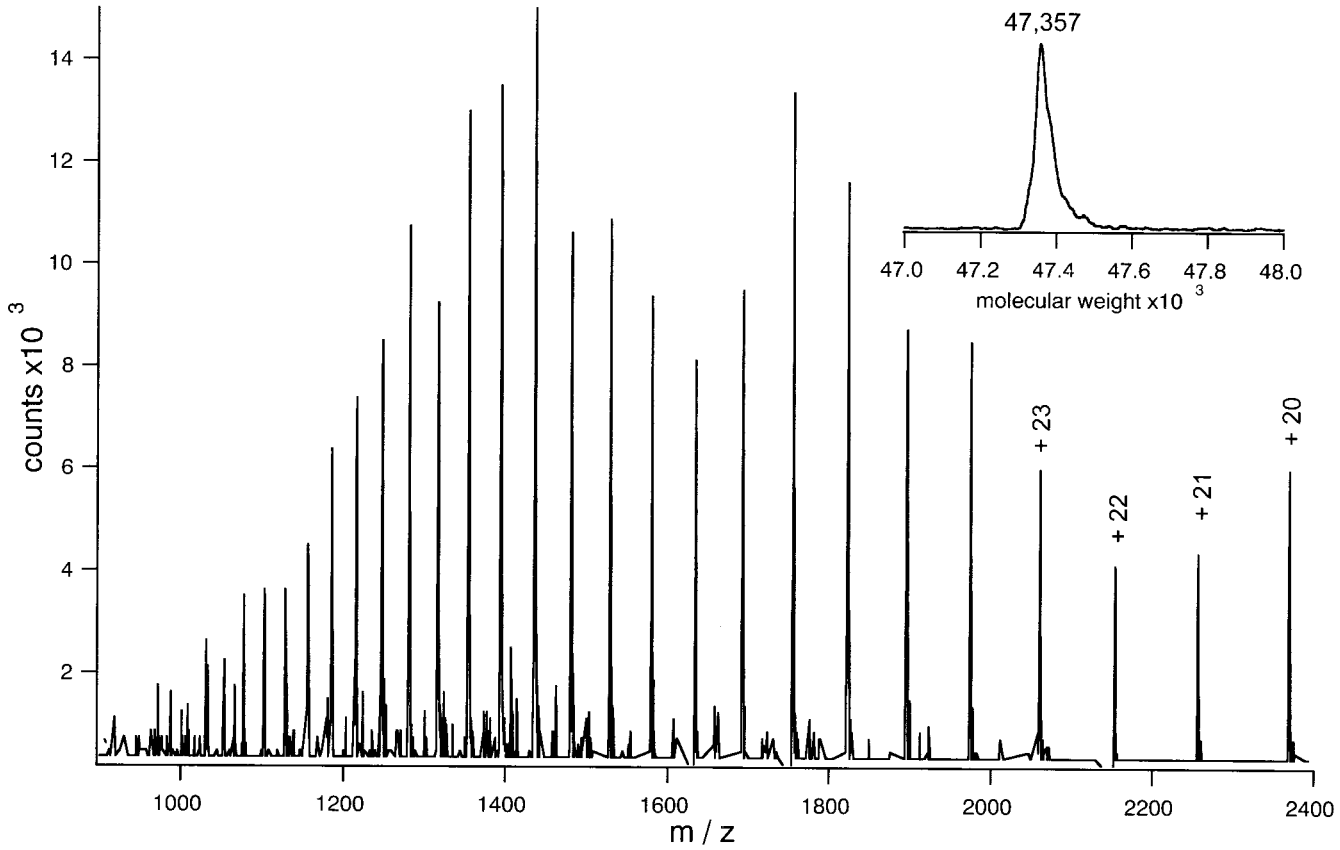

FIG. 2. ESI mass spectrum of lac permease collected upon elution of the 600-sec UV peak. ESI generates multiply charged ions, each with different mass/charge (m/z) ratios through the addition of variable numbers of protons. In the case of lactose permease, it was necessary to generate ions with at least 20 protons such that their m/z was in the scanning range of the mass spectrometer. Spectra from several scans across the peak were added to generate the spectrum shown. Typically ions carrying from 20 to 50 positive charges were detected, appearing to fall into at least three distributions, suggestive of different tertiary structures in the gas phase. The measured mass is calculated from the m/z of these ions after assigning the charge states. A computer-generated reconstruction of the zero-charged protein (Hypermass, PE Sciex, Toronto) is shown (*Inset*).

different conformational states in the gas phase (19). Stable elements of secondary and tertiary structure have been observed for some membrane proteins in aqueous organic solvent mixtures (20). The molecular weight reconstruction of the uncharged permease with a 6-His tag (Fig. 2, *Inset*) reveals that the *lacY* gene product is homogeneous and that the computed mass (47,357) is within 0.01% of that calculated from the DNA sequence with a formyl group on the initiating methionine (Table 1). Previous studies (21) show that the formyl group is removed from native permease but overexpression may overwhelm deformylase activity. Importantly, the MS spectra indicate formylation when the protein is solubilized in trifluororacetic acid (not shown), and a recombinant form of the permease containing cytochrome b_{562} fused into the middle cytoplasmic loop (22, 23) also is formylated. Finally the protein is amenable to Edman degradation only after N-terminal deblocking.

Table 1. Calculated and measured masses for lactose permease

E. coli lactose permease	Calculated mass,‡ Da	Measured mass, Da	Difference (%)
Native (+6 His tag)*	47,354	47,357	(<0.01)
"Red" (+10 His tag)†	60,458	60,457	(<0.01)

*Native lactose permease sequence with a six-His tag and *N*-formylation of Met1.
†Permease construct with cytochrome b_{562} in the middle cytoplasmic loop, a 10-His tag at the C terminus and N-formylation of Met1.
‡Average masses are used based on natural isotopic abundance.

MS has the power to monitor subtle covalent modifications that often modulate regulatory and functional plasticity (2, 24–26), and it is possible to observe chemical modifications with unprecedented detail. To probe the eight native Cys residues in lactose permease, the protein was alkylated with *N*-ethylmaleimide in the presence or absence of the substrate analog D-galactopyranosyl β-D-thio-galactopyranoside (TDG) (Fig. 3). Although none of the thiols are involved in disulfide bonds, ESI-MS demonstrates that only two are alkylated (each NEM addition increases the mass by 125 Da), and slowing of the reaction indicates different reaction rates for each thiol (Fig. 3*B*). Substrate-protected NEM labeling of only two native Cys residues, Cys-148 and Cys-355, has been described (7). ESI-MS detects a dominant amount of material that remains unmodified or labeled with only one NEM when the reaction is performed in the presence of TDG (Fig. 3*D*), demonstrating that one of the Cys residues is only partially protected by substrate. Labeling of Cys-148, a component of the binding site that interacts hydrophobically with the galactosyl moiety of the substrate, is inhibited completely by substrate. On the other hand, substrate protection of Cys-355 against alkylation reflects a long-range conformational change caused by binding of substrate (7).

It is demonstrated here that large hydrophobic intrinsic membrane proteins are completely amenable to mass spectrometric analysis. MS will play a vital role in proteomics with ionization of intact proteins increasingly becoming an

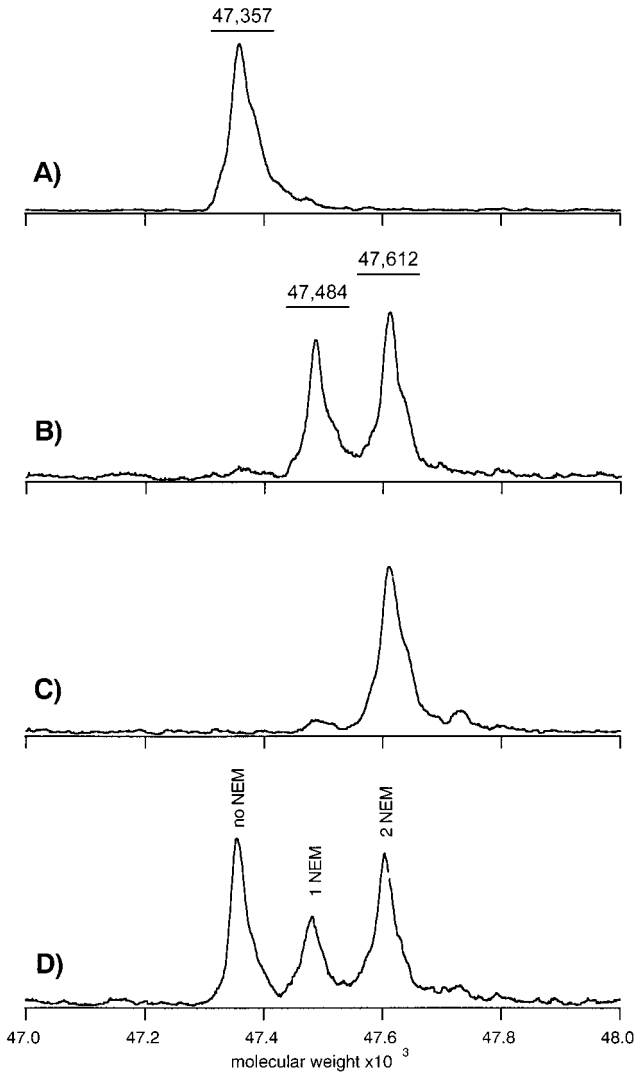

FIG. 3. NEM modification and substrate protection by D-galactopyranosyl β-D-thio-galactopyranoside (TDG). ESI-MS is used to quantitate the addition of NEM groups to Cys residues of the solubilized protein. Each NEM addition increases the mass of the protein by 125 Da. (*A*) Unmodified permease. (*B*) Ten-minute labeling with NEM on ice. (*C*) Ten-minute labeling with NEM at room temperature. (*D*) same as *C* after preincubation with a saturating concentration (50 mM) of the substrate analog TDG. Of eight thiol groups in lactose permease only two are modified under the conditions used. Preincubation with substrate reveals differential protection of these Cys residues against alkylation with NEM (see text).

alternative to working with peptide fragments (27). Combination of ESI-MS with suitable chromatography should allow rapid analysis of complex protein mixtures such as entire membrane fractions with high accuracy within less than 1 hr.

The W.M. Keck Foundation provided support toward instrument purchase. The work was also supported in part by National Institutes of Health Grant R01 DK51131-01 (to H.R.K.). G.G.P. is supported by a Grant from the Medical Research Council of Canada. C.K.E. is a postdoctoral fellow of the Deutsche Forschungsgemeinschaft.

1. Lander, E. S. (1996) *Science* **274,** 536–539.
2. Whitelegge, J. P., Gundersen, C. B. & Faull, K. F. (1998) *Protein Sci.* **7,** 1423–1430.
3. Ashton, D. S., Beddell, C. R., Green, B. N. & Oliver, R. W. (1994) *FEBS Lett.* **342,** 1–6.
4. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. (1989) *Science* **246,** 64–71.
5. Covey, T. R., Bonner, R. F., Shushan, B. I. & Henion, J. (1988) *Rapid Commun. Mass Spectrom.* **2,** 249–256.
6. Kaback, H. R. (1976) *J. Cell Physiol.* **89,** 575–593.
7. Frillingos, S., Sahin-Tóth, M., Wu, J. & Kaback, H. R. (1998) *FASEB J.* **12,** 1281–1299.
8. Zhuang, J., Privé, G. G., Verner, G. E., Ringler, P., Kaback, H. R. & Engel, A. (1999) *J. Struct. Biol.* **125,** 63–75.
9. le Coutre, J., Narasimhan, L. R., Patel, C. K. & Kaback, H. R. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 10167–10171.
10. le Coutre, J., Kaback, H. R., Patel, C. K. N., Heginbotham, L. & Miller, C. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 6114–6117.
11. Bouchon, B., Van Dorsselaer, A. & Roitsch, C. (1993) *Biol. Mass Spectrom.* **22,** 358–360.
12. Hufnagel, P., Schweiger, U., Eckerskorn, C. & Oesterhelt, D. (1996) *Anal. Biochem.* **243,** 46–54.
13. le Maire, M., Deschamps, S., Møller, J. V., Le Caer, J. P. & Rossier, J. (1993) *Anal. Biochem.* **214,** 50–57.
14. Rosinke, B., Strupat, K., Hillenkamp, F., Rosenbusch, J., Dencher, N., Kruger, U. & Galla, H. J. (1995) *J. Mass Spectrom.* **30,** 1462–1468.
15. Schaller, J., Pellascio, B. C. & Schlunegger, U. P. (1997) *Rapid Commun. Mass Spectrom.* **11,** 418–426.
16. Schindler, P. A., Van Dorsselaer, A. & Falick, A. M. (1993) *Anal. Biochem.* **213,** 256–263.
17. Barnidge, D. R., Dratz, E. A., Jesaitis, A. J. & Sunner, J. (1999) *Anal. Biochem.* **269,** 1–9.
18. Fearnley, I. M. & Walker, J. E. (1996) *Biochem. Soc. Trans.* **24,** 912–917.
19. Konermann, L. & Douglas, D. J. (1997) *Biochemistry* **36,** 12296–12302.
20. Arkin, I. T., Russ, W. P., Lebendiker, M. & Schuldiner, S. (1996) *Biochemistry* **35,** 7233–7238.
21. Ehring, R., Beyreuther, K., Wright, J. K. & Overath, P. (1980) *Nature (London)* **283,** 537–540.
22. Privé, G. G., Verner, G. E., Weitzman, C., Zen, K. H., Eisenberg, D. & Kaback, H. R. (1994) *Acta Crystrallogr.* **50,** 375–379.
23. Privé, G. G. & Kaback, H. R. (1996) *J. Bioenerg. Biomembr.* **28,** 29–34.
24. Burlingame, A. L., Boyd, R. K. & Gaskell, S. J. (1996) *Anal. Chem.* **68,** 599R–651R.
25. Chait, B. T. & Kent, S. B. (1992) *Science* **257,** 1885–1894.
26. Roepstorff, P. (1997) *Curr. Opin. Biotechnol.* **8,** 6–13.
27. McLafferty, F. W., Kelleher, N. L., Begley, T. P., Fridriksson, E. K., Zubarev, R. A. & Horn, D. M. (1998) *Curr. Opin. Chem. Biol.* **2,** 571–578.
28. Sahin-Tóth, M., le Coutre, J., Kharabi, D., le Maire, G., Lee, J. C. & Kaback, H. R. (1999) *Biochemistry* **38,** 813–819.
29. Wessel, D. & Flügge, U. I. (1984) *Anal. Biochem.* **138,** 141–143.
30. Whitelegge, J. P., Jewess, P., Pickering, M. G., Gerrish, C., Camilleri, P. & Bowyer, J. R. (1992) *Eur. J. Biochem.* **207,** 1077–1084.