

# Absolute Multiplexed Quantitative Analysis of Protein Expression during Muscle Development Using QconCAT\*

Jenny Rivers‡, Deborah M. Simpson‡, Duncan H. L. Robertson‡, Simon J. Gaskell§, and Robert J. Beynon‡¶

Stable isotope-labeled proteotypic peptides are used as surrogate standards for absolute quantification of proteins in proteomics. However, a stable isotope-labeled peptide has to be synthesized, at relatively high cost, for each protein to be quantified. To multiplex protein quantification, we developed a method in which gene design *de novo* is used to create and express artificial proteins (QconCATs) comprising a concatenation of proteotypic peptides. This permits absolute quantification of multiple proteins in a single experiment. This complete study was constructed to define the nature, sources of error, and statistical behavior of a QconCAT analysis. The QconCAT protein was designed to contain one tryptic peptide from 20 proteins present in the soluble fraction of chicken skeletal muscle. Optimized DNA sequences encoding these peptides were concatenated and inserted into a vector for high level expression in *Escherichia coli*. The protein was expressed in a minimal medium containing amino acids selectively labeled with stable isotopes, creating an equimolar series of uniformly labeled proteotypic peptides. The labeled QconCAT protein, purified by affinity chromatography and quantified, was added to a homogenized muscle preparation in a known amount prior to proteolytic digestion with trypsin. As anticipated, the QconCAT was completely digested at a rate far higher than the analyte proteins, confirming the applicability of such artificial proteins for multiplexed quantification. The nature of the technical variance was assessed and compared with the biological variance in a complete study. Alternative ionization and mass spectrometric approaches were investigated, particularly LC-ESI-TOF MS and MALDI-TOF MS, for analysis of proteins and tryptic peptides. QconCATs offer a new and efficient approach to precise and simultaneous absolute quantification of multiple proteins, subproteomes, or even entire proteomes. *Molecular & Cellular Proteomics* 6:1416–1427, 2007.

From the ‡Proteomics and Functional Genomics Group, Faculty of Veterinary Science, University of Liverpool, Liverpool L69 7ZJ, United Kingdom and §Michael Barber Centre for Mass Spectrometry, School of Chemistry and Manchester Interdisciplinary Biocentre, University of Manchester, Manchester M1 7DN, United Kingdom

Received, December 4, 2006, and in revised form, May 16, 2007

Published, MCP Papers in Press, May 17, 2007, DOI 10.1074/mcp.M600456-MCP200

As the field of proteomics matures as a discipline, there is an increasing realization of the importance of *absolute* as well as *relative* quantification, and considerable effort is being directed toward experimental strategies to achieve this goal. Most commonly, relative protein quantification by mass spectrometry has been based on differential stable isotope labeling implemented by metabolic incorporation (1, 2) or through derivatization strategies such as ICAT (3). The mass-coded abundance tagging method (4) avoids the use of stable isotopes but requires assumptions concerning mass spectrometric response factors. To achieve relative quantification of proteins without isotope labeling or chemical modification steps, quantitative comparisons have been made of equivalent sets of mass spectrometric data by considering peptide detectability in repetitively acquired spectra or by comparing integrated extracted ion chromatograms following liquid chromatography-mass spectrometry analyses (5).

In principle, any of the approaches adopted for relative quantification may also be used for absolute quantification if reference standards are available for all analytes in known amounts. When unknowns and reference standards are co-analyzed, such approaches exploit the well established precept in analytical chemistry of internal standardization in which a known amount of a stable isotope-labeled (or otherwise differentiated) standard is added to the analyte such that the response ratio between analyte and the heavier internal standard can then be used to quantify the unknown. However, for quantification of individual proteins in a proteomics study, the true internal standard would be the corresponding protein expressed in pure and stable isotope-labeled form and quantified. This would be challenging on many fronts, including the expression of a native protein in a heterologous system to effect labeling, purification of the protein, and subsequent mass spectrometric analysis of the complex isotopic profile of the analyte and standard protein. Rather than adopt a protein-based approach, absolute quantification using proteotypic peptides as surrogates for the protein of interest has emerged using stable isotope-labeled peptide internal standards as “signature” or “proteotypic” peptides that are chemically synthesized and incorporate stable isotopes (6–9). Each protein to be quantified requires at least one stable isotope-labeled peptide that must be independently synthesized at relatively

high cost. Moreover each peptide must be separately purified and quantified (10). There is clearly a need for approaches that make this process more streamlined especially if multiple proteins are to be quantified.

We have recently introduced an efficient alternative to the chemical synthesis of multiple stable isotope-labeled peptides (11). In brief, artificial genes are designed *de novo* to direct the synthesis of novel proteins that are assemblies of signature Qpeptides derived from a number of discrete proteins. Usually these Qpeptides are arginine or lysine terminated at the C terminus as they represent and will be internal standards for tryptic peptides derived from digestion of the analyte proteins. Appropriately flanked with added features including an initiator codon, a purification tag, and protective sacrificial regions, the gene is transformed into and expressed in a heterologous system, usually bacterial. The expression strain is grown in a chemically defined medium, uniformly isotopically labeled (for example, using  $^{15}\text{NH}_4\text{Cl}$  as the sole nitrogen source) or containing specific stable isotope-labeled amino acids at a high isotope enrichment such that the artificial protein becomes fully labeled. The artificial protein (termed a “QconCAT” for “quantification concatamers”) is purified by virtue of the affinity tag and quantified using a suitable procedure (12). The QconCAT is mixed with a complex mixture of analyte proteins, and subsequent proteolysis releases both the stable isotope-labeled standard and the cognate peptide from the analyte. The known quantity of standard added can then be used for absolute quantification of the analyte. Because quantification of the QconCAT protein will define in absolute terms the quantity of each of the surrogate peptides, the QconCAT strategy provides an efficient means to multiplex absolute quantification. Tryptic peptides are typically 10–15 amino acids long; thus proteotypic Qpeptides from 50 proteins could be encoded in a protein comprising 500–750 amino acids. The Qpeptides are present, by design, in stoichiometrically known amounts (usually equimolar) so that each analyte peptide (and therefore protein) is simultaneously quantified.

Qpeptides are concatenated in the QconCAT protein out of their normal primary sequence context, and it is formally correct to point out that this different context could influence the quantification data (13). However, this can only occur if either the QconCAT or the analyte proteins are incompletely digested such that the yield of each peptide is incomplete. It is generally accepted that for most general proteases, such as trypsin, the  $K_m$  for proteins and peptides is relatively high, and proteolytic reactions operating at substrate concentrations below this value exhibit pseudo-first order kinetics (14–16). Thus, if the rate of digestion of either the QconCAT or analyte was so low that six or seven reaction half-times could not elapse during the proteolytic reaction, there might be discordance between the yield of the standard and analyte peptide. However, the main determinant of the rate of proteolysis of native proteins is higher order structure, not primary se-

quence context. Tightly folded proteins, particularly those with a high proportion of  $\beta$  sheet, are intrinsically resistant to proteolysis (17, 18). There is no reason, *a priori*, to expect that QconCATs would adopt such tightly folded structures. Indeed their propensity to form insoluble inclusion bodies and their recovery by dissolution in strong chaotropes both mitigate against structural impediments to proteolysis. By contrast, unless care is taken in the prior denaturation of analyte proteins, their higher order structure would almost certainly influence proteolysis and could impact absolute quantification. We stress, however, that the incomplete analyte digestion is as much an issue for quantification using synthetic peptides as those using QconCATs. We address the issue of QconCAT and analyte proteolysis here and show that it is a factor that is readily controlled.

Deployment of a QconCAT experiment has many aspects that must be optimized. We demonstrate the use of a QconCAT for absolute quantification of a group of proteins that demonstrate dramatic changes in expression during development of skeletal muscle in the chicken posthatching. We assessed the scope of the method and the magnitude and sources of variance that the method contains. We confirmed the value of guanidination (19) as a strategy to enhance peptide ion yields in MALDI-TOF MS and showed that effective quantification is attainable and equivalent in both MALDI-TOF and ESI-TOF analyses.

#### EXPERIMENTAL PROCEDURES

**Materials and Reagents**—Trypsin (sequence grade) was obtained from Roche Diagnostics. All other chemicals and solvents (HPLC grade) were purchased from Sigma-Aldrich and VWR International Laboratory Supplies (Leicestershire, UK).

**Proteomics Analysis of Chicken Skeletal Muscle Soluble Fraction**—Chickens (Institut de Sélection Animale (ISA) Brown layer and Ross 308 broiler) were grown to 30 days posthatch, and animals were culled at 1, 3, 5, 10, 20, and 30 days at which time pectoralis muscle was collected (the above procedures were performed at the Roslin Institute, Edinburgh, UK). To isolate the soluble fraction of chicken skeletal muscle, 100 mg of breast tissue was homogenized in 0.9 ml of 20 mM sodium phosphate buffer, pH 7.0, containing protease inhibitors (Complete protease inhibitors, Roche Applied Science). The homogenized sample was centrifuged at  $15,000 \times g$  for 45 min at 4 °C. The supernatant fraction, containing soluble protein, was then removed. This was repeated, homogenizing the insoluble fraction in the same volume of sodium phosphate, and the pooled supernatant fractions were used for all analyses. The total protein concentration of each preparation was measured using a Coomassie Plus Protein Assay (Pierce).

For 1D<sup>1</sup> SDS-PAGE analysis, 10  $\mu\text{g}$  of soluble protein samples (volume, 5–10  $\mu\text{l}$ ) from birds of different strains and ages were each mixed with an equal volume of reducing sample buffer (1 ml of 0.5 M Tris buffer, pH 6.8, 1 ml of glycerol, 0.02 g of SDS, 0.01 g of bromophenol blue, 0.154 g of DTT) and resolved by 12.5% (w/v) SDS-PAGE prior to staining with Coomassie Blue (Bio-Safe, Bio-Rad). Gels were

<sup>1</sup> The abbreviations used are: 1D, one-dimensional; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; AK, adenylate kinase; CV, coefficient of variance.

destained with 10% (v/v) acetic acid, 10% (v/v) methanol.

**Preparation and Purification of QconCAT**—The artificial QconCAT gene (11) was expressed in *Escherichia coli* with a full complement of unlabeled amino acids or in the presence of [ $^{13}\text{C}_6$ ]lysine (100 mg/liter) and [ $^{13}\text{C}_6$ ]arginine (100 mg/liter) as the sole source of these two amino acids. Expression was induced with isopropyl  $\beta$ -D-thiogalactopyranoside, and the cells were harvested by centrifugation at  $1400 \times g$  at 4 °C for 15 min. Inclusion bodies containing QconCAT (as proven by digestion with trypsin and MALDI-TOF MS analysis; data not shown) were recovered by breaking cells using BugBuster Protein Extraction Reagent (Novagen, Nottingham, UK). Inclusion bodies were resuspended in 20 mM phosphate buffer, 6 M guanidinium chloride, 0.5 M NaCl, 20 mM imidazole, pH 7.4. From this solution, [ $^{13}\text{C}_6$ ]lysine/arginine-labeled and unlabeled QconCAT proteins were purified separately by affinity chromatography using a nickel-based resin (HisTrap HP kit, Amersham Biosciences). Following sample loading, HisTrap columns were washed with 20 mM phosphate buffer, pH 7.4, prior to elution of the sample with the same buffer containing a higher concentration of imidazole (20 mM phosphate, 0.5 M NaCl, 500 mM imidazole, 6 M guanidinium chloride, pH 7.4) during which phase fractions (1 ml) were collected. The purified QconCAT was desalted by three rounds of dialysis against 100 volumes of 10 mM ammonium bicarbonate, pH 8.5, for 2 h using fresh buffer each time.

**Proteomics Analysis of QconCAT for Quantification of Chicken Skeletal Muscle Proteins**—The QconCAT protein was diluted to 5  $\mu\text{M}$  in 50 mM ammonium bicarbonate and digested with trypsin (20:1 substrate:protease) at 37 °C for 24 h after which the digest was incubated with additional trypsin (20:1 substrate:protease) to ensure complete digestion. Peptides were analyzed by MALDI-TOF MS (M@LDI, Waters, Manchester, UK). For this, 1  $\mu\text{l}$  of digested material was mixed with an equal volume of  $\alpha$ -cyanohydroxycinnamic acid in 50% (v/v) acetonitrile, 0.1% (v/v) trifluoroacetic acid. This was allowed to dry, and peptides were acquired over the mass range 900–3000  $m/z$ . For each combined spectrum, 20–30 spectra were acquired (laser energy typically 30%) with 10 shots per spectrum and a laser firing rate of 5 Hz. Data were processed using MassLynx software to subtract background noise using polynomial order 10 with 40% of the data points below this polynomial and a tolerance of 0.01. Spectral data were also smoothed by performing two mean smooth operations with a window of three channels.

**Co-digestion of QconCAT and Chicken Skeletal Muscle Soluble Proteins for Quantification**—QconCAT protein was added in a 1:10 (QconCAT:chicken skeletal muscle protein) ratio to chicken skeletal muscle soluble fraction samples taken from both broiler and layer strains at six time points during growth. For each time point, four birds were analyzed. The mixture was diluted 10-fold with 50 mM ammonium bicarbonate, and 10% (v/v) acetonitrile was added prior to addition of trypsin (20:1 substrate:protease). The reaction mixture was incubated at 37 °C for 24 h, after which the digest was incubated with additional trypsin (20:1 substrate:protease) to ensure complete digestion. 1  $\mu\text{l}$  was analyzed by MALDI-TOF MS.

**Monitored Proteolysis of QconCAT and Analyte Proteins**—For QconCAT digestion, 150  $\mu\text{g}$  of protein was digested with trypsin at a ratio of trypsin:protein of 1:20 and 1:100 and stopped at selected time points after addition of enzyme by removing 15  $\mu\text{l}$  (containing 3  $\mu\text{g}$  of protein) and adding to an equal volume of 10% (v/v) formic acid. For analyte protein digestion, 50  $\mu\text{g}$  of protein was digested with trypsin at a ratio of trypsin:protein of 1:20 and stopped at 0 min, 30 min, and 24 h after addition of enzyme by removing 25  $\mu\text{l}$  (containing 6  $\mu\text{g}$  of protein) and adding an equal volume of 10% (v/v) formic acid. The fractions were subsequently stored at  $-20$  °C until the end of the time course. For gel electrophoresis, fractions were dried down in a vacuum centrifuge and reconstituted in 10  $\mu\text{l}$  of reducing sample buffer prior to separation by 12.5% (w/v) 1D SDS-PAGE at 200 V for 45 min.

Analyte proteins were also digested in a solution containing 10% ACN (v/v) and with addition of enzyme following a 1-h incubation of the protein at 60 °C. To quantify proteolysis of analyte proteins, digestion of chicken skeletal muscle soluble proteins in solution with trypsin (as described above) was stopped at various time points during 24-h incubation at 37 °C by removing 20  $\mu\text{l}$  (containing 5  $\mu\text{g}$  of protein) and adding an equal volume of 10% (v/v) formic acid containing 0.5  $\mu\text{g}$  of predigested QconCAT peptides. Each fraction was analyzed by MALDI-TOF MS. This experiment was repeated using protein denatured by incubation at 60 °C for 1 h prior to trypsin addition for comparison.

**Guanidination**—To enhance the signal intensity of lysine-terminated peptides in MALDI-TOF MS, lysine residues were converted to the more basic homoarginine by guanidination (20). This reaction was carried out by drying down the peptide mixture and reconstituting it in 10  $\mu\text{l}$  of 7 M ammonia solution to which was added 5  $\mu\text{l}$  of 0.5 M *O*-methylisourea (in double distilled  $\text{H}_2\text{O}$ ). This was mixed thoroughly and incubated overnight at room temperature prior to drying down and desalting using  $\text{C}_{18}$  ZipTips (Millipore, Watford, UK).

**LC-MS**—Peptide mixtures were analyzed by LC-ESI-Q-TOF MS using an EASY-nLC (Proxeon, Odense, Denmark) nanoflow system coupled to a Q-ToF micro (Waters). Nanoflow HPLC at 200 nl/min was used to resolve peptides (in 0.1% (v/v) formic acid) over a 50-min acetonitrile gradient (0–100%). Peptides were acquired over the mass range 400–2000  $m/z$  with the capillary voltage set at 1900 V, collision energy set at 10 V, and sample cone set at 55 V for the entire 50-min gradient. The same reversed phase separation method was used to collect fractions (200 nl) directly onto a MALDI-TOF target for analysis by LC-MALDI-TOF MS.

**Assessing Analytical and Biological Variance in Quantification**—Ten identical aliquots of a chicken skeletal muscle soluble protein preparation, to each of which was added a known amount of [ $^{13}\text{C}_6$ ]arginine/[ $^{13}\text{C}_6$ ]lysine-labeled QconCAT, were digested in solution with trypsin and analyzed to investigate analytical variance. This was compared with biological variance (four animals at each time point) achieved through quantification by MALDI-TOF MS (with and without guanidination), LC-ESI-Q-TOF MS, and LC-MALDI-TOF MS.

**Comparison of QconCAT Method with Absolute Quantification Using a Stable Isotope-labeled Synthetic Peptide**—Quantification by the QconCAT method was also compared with that achieved using a stable isotope-labeled synthetic peptide to quantify a single analyte protein also represented in the QconCAT. The peptide of sequence LVSWYDNEFGYSNR and mass 1748.77 Da representing the abundant protein GAPDH was synthesized by Sigma-Genosys (Dorset, UK) and was labeled at the arginine residue with both  $^{13}\text{C}_6$  and  $^{15}\text{N}_4$  giving a 10-Da mass offset from the analyte peptide. For quantification, the synthetic peptide was added to broiler chicken skeletal muscle samples corresponding to six time points during growth with four replicate animals at each time point. Quantification data were obtained from analysis by MALDI-TOF MS using the relative intensities of the analyte and standard peaks as with QconCAT analysis.

**Investigation of the Accuracy of Quantification Using QconCAT**—Purified adenylate kinase (AK; Sigma) was added to chicken skeletal muscle soluble fraction from a 30-day broiler. AK was added from 0 to 0.02 nmol resulting in a final protein concentration of 0–300 nmol/g, and the amount of AK in the tissue was quantified by adding 0.015 nmol of QconCAT prior to digestion with trypsin. Proteolysis was allowed to continue for 24 h after which peptides were analyzed by MALDI-TOF MS.

## RESULTS

The QconCAT was designed to include surrogate peptides for 20 chicken skeletal muscle proteins. As chicken skeletal muscle matures posthatch, the protein distribution in the tis-

sue changes dramatically from a large number of proteins that are expressed in similar amounts at hatch to a relatively few high abundance proteins after 30 days of growth (Fig. 1). From previous identification studies (21), the most abundant proteins present in the soluble fraction of chicken skeletal muscle at this stage are predominantly the glycolytic enzymes. Other proteins, notably actin, have disappeared from the soluble

fraction of muscle by 10 days of growth, presumably reflecting repartitioning and assembly of the myofibrillar apparatus. Finally serum proteins are detectable in muscle preparations at hatch but rapidly disappear during development. We ascribe this change to the increased exclusion of interstitial fluid as the muscle develops (22). To measure the absolute concentrations of specific proteins at various time points, we selected a group of 20 to be quantified using a single QconCAT. For each of the proteins, we chose a proteotypic peptide that gave a strong signal in previous MALDI-TOF MS analyses of tryptic digests. The peptides were used to guide construction of the DNA sequence of the QconCAT, which was synthesized, inserted into a pET21a vector, and expressed in *E. coli*. Full details of the design and expression are given elsewhere (12).

For QconCAT expression, a typical bacterial culture of 200 ml was induced at an  $A_{600}$  of 0.6–0.8, which generated 5–10 mg of QconCAT after cell breakage, recovery of inclusion bodies, and affinity chromatography of guanidinium chloride-solubilized protein on 1-ml nickel-nitrilotriacetic acid columns. After induction, the QconCAT protein was visible as a major band in 1D SDS-PAGE of a broken cell preparation (results not shown). After purification, the protein was homogeneous on 1D SDS-PAGE and was used without further purification (results not shown).

QconCAT protein was added in a 1:10 (QconCAT:chicken skeletal muscle protein) ratio to chicken skeletal muscle soluble fraction samples taken from both broiler and layer strains at six time points during growth. For each time point, four birds were analyzed. This ratio was selected pragmatically based on the abundance of the major proteins in chicken skeletal muscle soluble fraction. The influence of dynamic

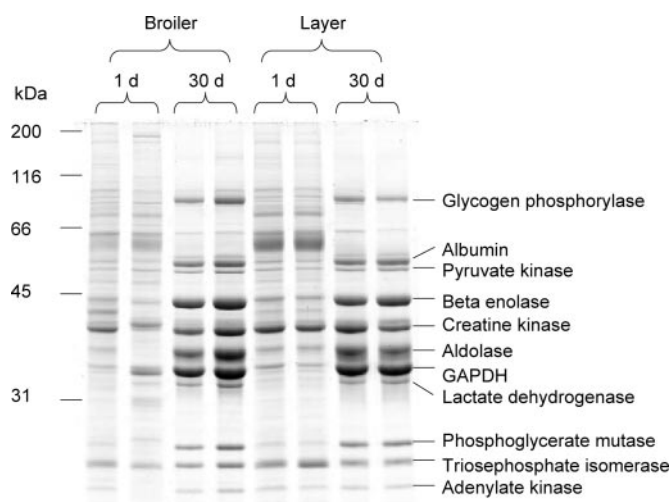


FIG. 1. 1D SDS-PAGE analysis of the soluble fraction of chicken skeletal muscle. Two different birds were compared at 1 and 30 days (*d*) after hatch for each strain. Soluble proteins (10  $\mu$ g; volume, 5–10  $\mu$ l) were mixed in an equal volume with reducing SDS sample buffer, boiled for 5 min, and loaded onto a 12.5% (w/v) large format acrylamide gel prior to staining with Coomassie Blue (Bio-Safe, Bio-Rad). Gels were destained with 10% (v/v) acetic acid, 10% (v/v) methanol. Major proteins were identified by MALDI-TOF peptide mass fingerprinting.

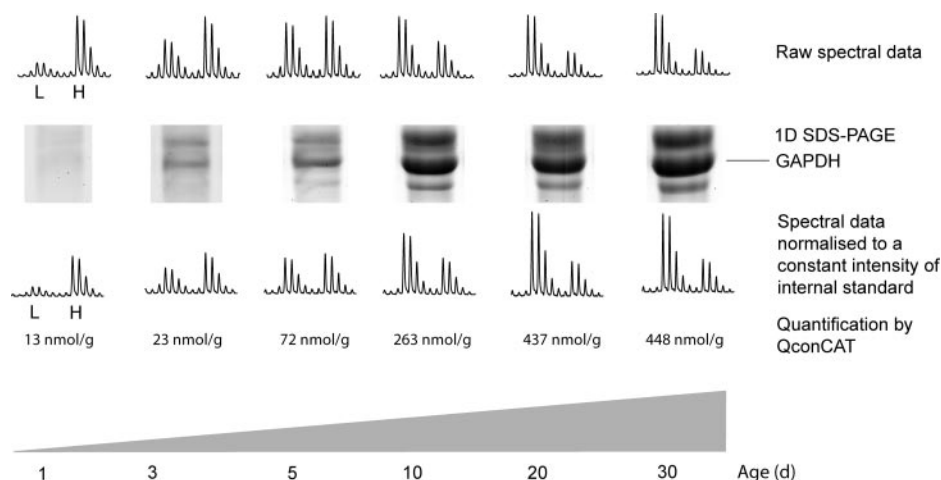
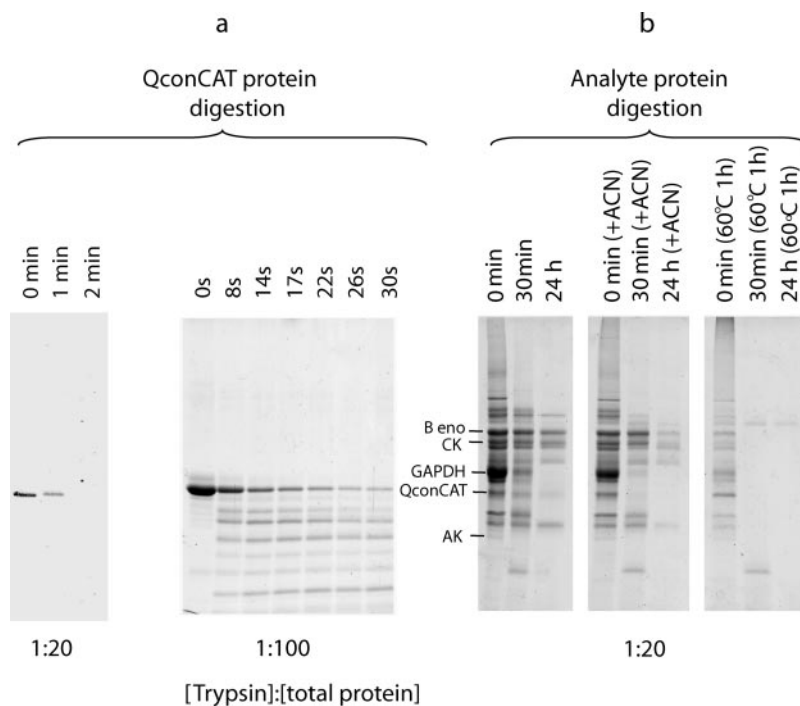


FIG. 2. Quantification of GAPDH expression in chicken skeletal muscle. Soluble muscle proteins were prepared from pectoralis skeletal muscle of birds from 1 to 30 days (*d*) posthatching. Each sample (70  $\mu$ g of protein) was mixed with a constant amount of QconCAT (7  $\mu$ g) and digested to completion with trypsin before analysis by MALDI-TOF mass spectrometry. The change in expression is measured using the relative peak intensity of the analyte and internal standard peptide at each time point. The dramatic increase in protein expression is more apparent when the spectra are normalized to a constant intensity of the internal standard. This change in protein expression is also apparent by 1D SDS-PAGE analysis of chicken skeletal muscle soluble proteins in which a constant 10  $\mu$ g of protein was applied to each lane. The amount of GAPDH at each time point during growth is also expressed as nmol/g of tissue as quantified using QconCAT.

**FIG. 3. Monitored proteolysis of QconCAT and analyte proteins by 1D SDS-PAGE.** QconCAT protein (150  $\mu\text{g}$ ) was digested with trypsin at an enzyme:protein ratio of 1:20 and 1:100. The digestion was stopped at selected time points after addition of enzyme with 10% (v/v) formic acid. Chicken skeletal muscle soluble protein (50  $\mu\text{g}$ ) was digested with trypsin at an enzyme:protein ratio of 1:20 and stopped at 0 min, 30 min, and 24 h after addition of enzyme with 10% (v/v) formic acid. For gel electrophoresis, fractions from QconCAT protein digestion (a) and chicken skeletal muscle soluble protein digestion (b) were dried down in a vacuum centrifuge and reconstituted in 10  $\mu\text{l}$  of reducing sample buffer prior to separation by 12.5% (w/v) 1D SDS-PAGE at 200 V for 45 min. Analyte proteins were also digested in a solution containing 10% ACN (v/v) and with addition of enzyme following a 1-h incubation of the protein at 60 °C. CK, creatine kinase; *B eno*,  $\beta$ -enolase.



range on absolute quantification of proteins in complex biological systems is discussed below. After co-digestion of chicken skeletal muscle soluble fraction and [ $^{13}\text{C}_6$ ]arginine/lysine-labeled QconCAT, MALDI-TOF MS analysis of peptides produced highly complex mass spectra. However, 10 of 20 proteotypic peptides could be identified in the composite spectrum without further sample processing and were therefore used for quantification. For these 10 proteins, for example glyceraldehyde-3-phosphate dehydrogenase (Fig. 2), the change in protein expression can be measured during growth from 1 to 30 days posthatch by converting relative signal intensities of analyte and internal standard peptide ions into absolute quantities of analyte protein expressed as nmol/g of net weight breast muscle tissue.

The QconCAT was completely digested within 2 min such that no intermediate fragments were visible on SDS-PAGE (Fig. 3). When the trypsin was reduced to much lower levels (100:1 substrate:protease) and the digestion reaction was sampled at very short time intervals, there was some evidence for the appearance of partially fragmented intermediates, although MALDI-TOF MS analysis of these bands, once digested with trypsin, demonstrated that each "band" comprised multiple species, consistent with simultaneous tryptic attack on all scissile bonds at very similar rates. MALDI-TOF MS of peptides confirmed rapid digestion with all peptides detected within the first 2 min of digestion (data not shown).

By contrast, if the protein preparation from skeletal muscle was subjected to trypsin digestion at a ratio of 20:1 substrate:protease, many proteins were digested slowly, and even after 24 h, undigested proteins were clearly visible including  $\beta$ -enolase, creatine kinase, and triose-phosphate isomerase. If a

low concentration (10%, v/v) of acetonitrile was included in the digestion reaction, proteolysis was faster. If the protein mixture was denatured by heating to 60 °C for 1 h before digestion, the loss of higher order structure of the substrate proteins meant that the digestion reaction was essentially complete within 30 min.

To demonstrate the importance of complete proteolysis for accurate quantification, we conducted extended digestion reactions with chicken skeletal muscle proteins from 1- and 30-day skeletal muscle. As reported previously and quantified here, these two preparations are dramatically different in the protein expression profiles (Fig. 1), providing different environments for proteolysis. The protein preparations were digested without treatment or after denaturation at 60 °C for 1 h, and the appearance of the analyte peptide used for quantification was determined by the QconCAT methodology; we have previously shown (Fig. 3) that the QconCAT was efficiently and completely digested within 2 min. In all instances, the analyte proteins were digested between 1.3 (AK) and 86 ( $\beta$ -enolase) times faster after denaturation, and in some instances (for example, GAPDH from 1-day muscle) the rate of digestion was very similar (Fig. 4). This is consistent with a model for proteolysis of the native protein in which the initial proteolytic attack exerts a destabilizing effect on the remaining structure such that the rate of proteolysis is increased; the initial proteolysis is effectively rate-limiting. However, in the highly specialized 30-day muscle sample, there was virtually no digestion even after 6 h of proteolysis. Indeed for all proteins studied, the rate of proteolysis of native proteins was diminished in the 30-day muscle sample; we suggest that the acute specialization of this tissue, leading to a

FIG. 4. **Quantification of proteolysis of analyte proteins using QconCAT.**

Chicken skeletal muscle soluble protein (50  $\mu\text{g}$ ) was digested with trypsin at an enzyme:protein ratio of 1:20 and stopped at selected time points with 10% (v/v) formic acid and mixed with 0.5  $\mu\text{g}$  of predigested QconCAT peptides for quantification. Each fraction was analyzed by MALDI-TOF MS. This experiment was repeated using protein denatured by incubating at 60  $^{\circ}\text{C}$  for 1 h prior to trypsin addition for comparison. Data are presented for four individual proteins at both 1 and 30 days (d) after hatch digested over 30 h (for which the first 500 min are shown) with trypsin both with (closed triangles) and without (open triangles) prior denaturation. For each, the rate constant ( $k$ ) for digestion is expressed as  $\text{h}^{-1}$ . CK, creatine kinase; B eno,  $\beta$ -enolase.

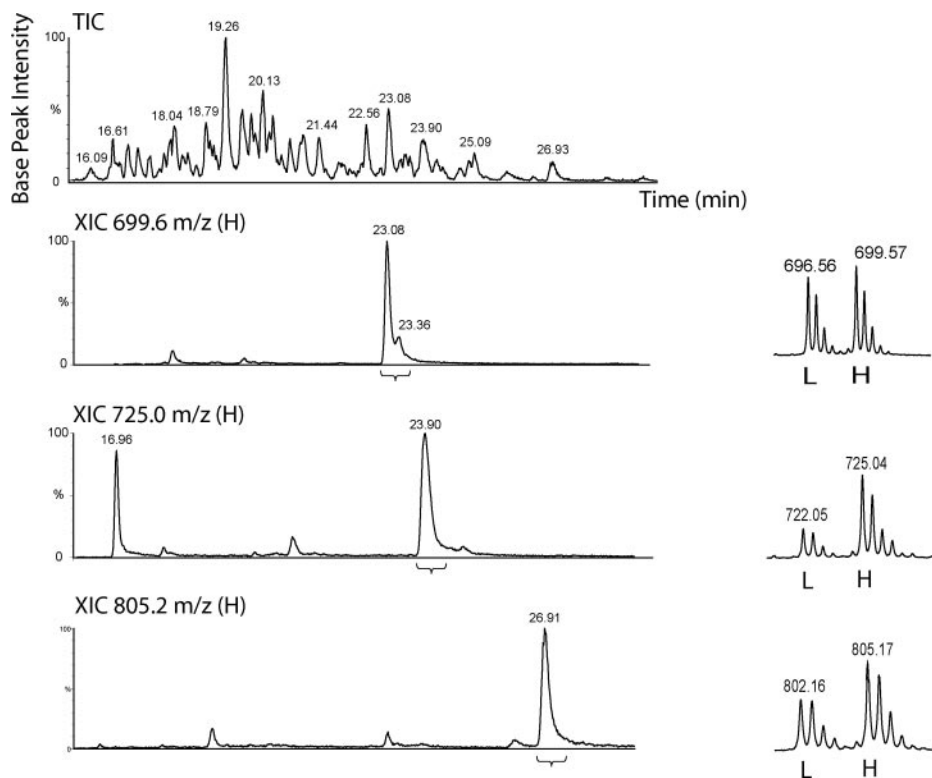
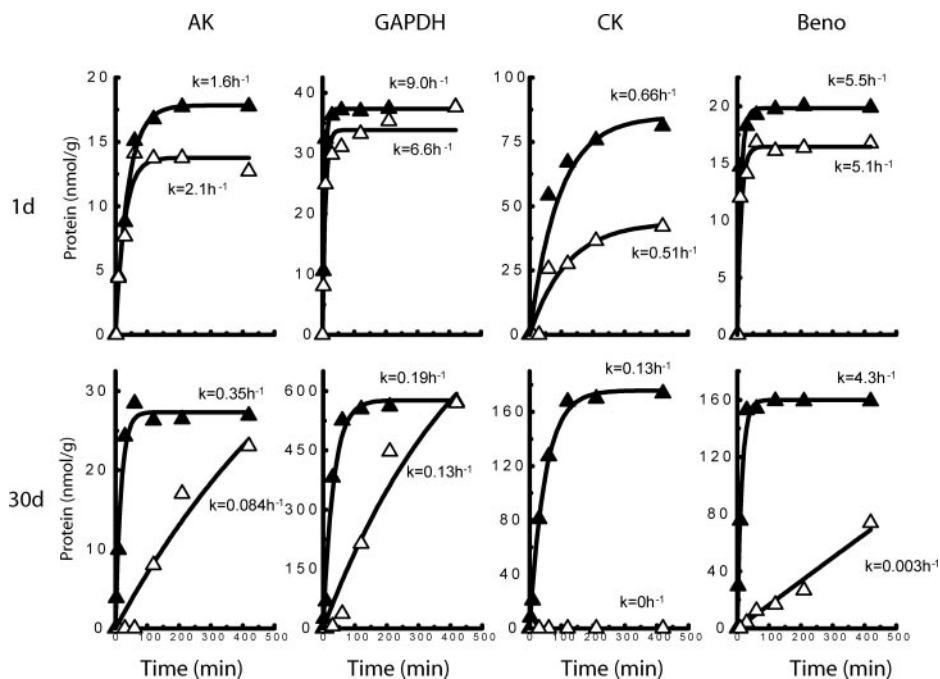


FIG. 5. **Isolation of analyte-standard peptide pairs by reversed phase chromatography.** QconCAT protein (7  $\mu\text{g}$ ) was added to chicken skeletal muscle soluble fraction (70  $\mu\text{g}$  of protein). This mixture was digested with trypsin and analyzed by LC-ESI-Q-TOF mass spectrometry. All peptide pairs for quantification were present as doubly charged ions; there was no evidence of triply charged species. The *upper panel* is the total ion chromatogram (TIC; base peak intensity) of the elution profile from 16 to 29 min. The *lower panels* are the extracted ion chromatograms (XIC) for representative QconCAT peptides of doubly charged ions ( $\beta$ -enolase, 699.6  $m/z$ , eluted at 23.08 min; glycogen phosphorylase, 725.0  $m/z$ , eluted at 23.90 min; and triose-phosphate isomerase, 805.2  $m/z$ , eluted at 25.09 min) with corresponding mass spectra showing analyte and QconCAT peptide ion pairs used for quantification presented as *insets* on the *right*. L, light; H, heavy.

predominance of relatively few proteins, might introduce other factors that impede digestion, such as aggregation into supramolecular assemblies or partial inhibition of the trypsin. In all instances, extended digestion times (greater than 24 h) resulted in complete digestion and the same quantification value irrespective of the initial state of the analyte protein preparation.

Variation in ion signal response is inherent with MALDI-TOF MS analysis (23). In particular, arginine-terminated peptides are known to yield more abundant signals than those terminated with lysine (24). In a complex MALDI-TOF mass spectrum, peptides that are abundant and have a high response factor dominate the spectrum. Theoretically proteolysis of a complex proteome (for illustration, 10,000 proteins) could

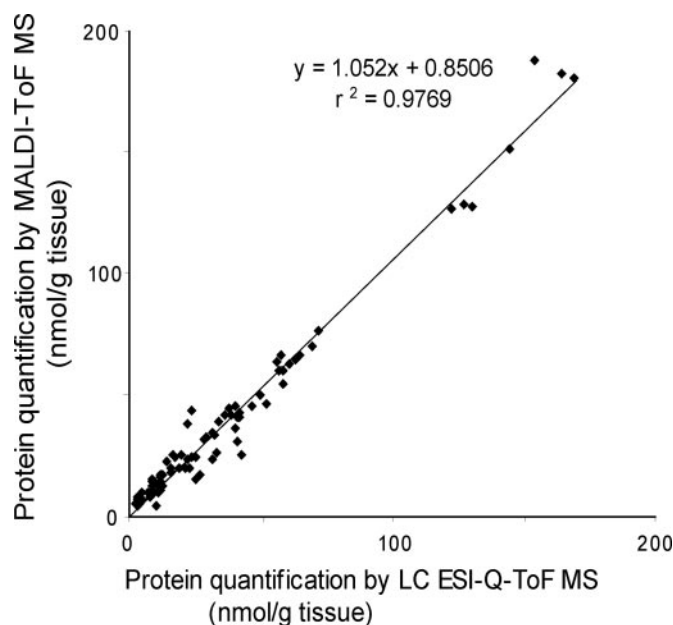


FIG. 6. **Quantification using MALDI-TOF or ESI-MS.** Soluble proteins from chicken skeletal muscle ( $70 \mu\text{g}$ ,  $n = 4$ , covering 1 to 30 days posthatch) were individually mixed with QconCAT protein ( $7 \mu\text{g}$ ) and digested to completion with trypsin. The entire peptide mixture was analyzed by MALDI-TOF MS or by nanoflow reversed phase HPLC prior to ESI-MS, and the absolute tissue content of each of four proteins (triose-phosphate isomerase, glyceraldehyde-3-phosphate dehydrogenase,  $\beta$ -enolase, and  $\alpha$ -actin) was assessed from relative intensities of light (analyte) and heavy (standard) pairs. The absolute amount of each protein was compared using the alternative forms of mass spectrometric analysis.

generate  $10^5$ – $10^6$  peptides (at  $\sim 50$  tryptic peptides per protein), the dynamic range of which will be such that only the most abundant peptides and those that ionize particularly well will be identified. To achieve increased signal intensity from lysine-terminated peptides, guanidination has been used to convert lysine into the more basic homoarginine by reaction with *O*-methylisourea (20). Guanidination of a tryptic digest was effective at increasing the signal intensity of lysine-terminated peptides in the QconCAT and the analyte sample to allow quantification of two more analyte proteins by MALDI-TOF MS. To improve resolution of peptides for quantification, samples were also analyzed by LC-ESI-Q-TOF MS (Fig. 5). The alternative ionization mode coupled with the benefit of separation of peptides by reversed phase chromatography allowed quantification of a further six proteins previously not identified by MALDI-TOF MS and confirmed quantification data for many of those that had previously been analyzed. Extracted ion chromatograms for unlabeled (analyte) and labeled (QconCAT) peptides were used to locate the ions, and the chromatographic boundaries of the coincident pair of peptides were used to delineate the combined mass spectra from which peptides were quantified by mass spectrometric intensities of the doubly charged ions; there was no evidence of multiply charged ions, for example  $[M + 3H]^{3+}$  correspond-

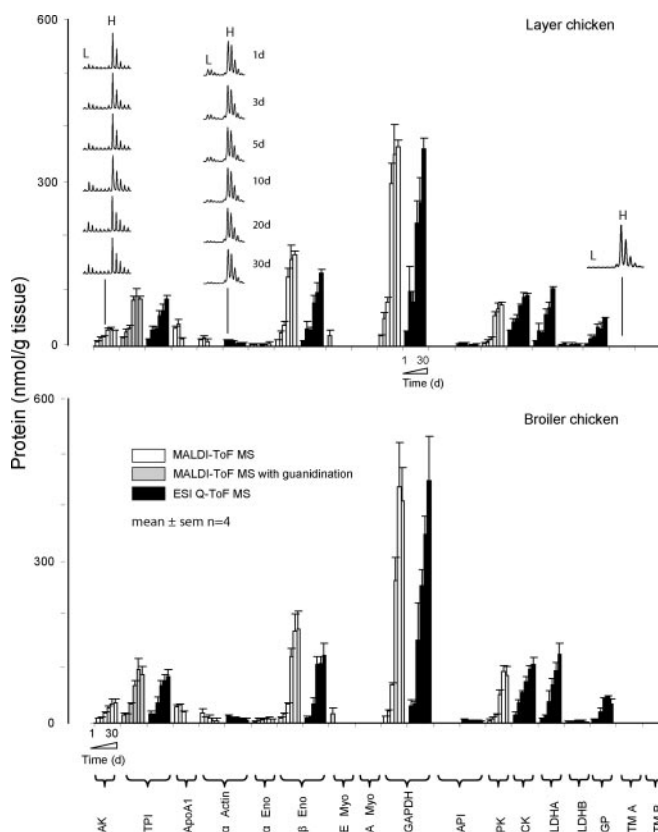


FIG. 7. **Quantification of chicken skeletal muscle protein expression by QconCAT.** Soluble protein derived from broiler and layer chickens ( $70 \mu\text{g}$ ) was mixed with QconCAT protein ( $7 \mu\text{g}$ ) and digested to completion with trypsin. The digests were analyzed by MALDI-TOF MS (with or without guanidination) or LC-ESI-Q-TOF MS. For five proteins (triose-phosphate isomerase,  $\alpha$ -actin,  $\beta$ -enolase, glyceraldehyde-3-phosphate dehydrogenase, and actin polymerization inhibitor) multiple methods were used to quantify a single protein during growth; these data have been plotted separately adjacent to the alternative data set, and these have been grouped below the x axis. Each cluster of data represents six time points during growth (1, 3, 5, 10, 20, and 30 days (d)) for four birds of each strain at each time point. The data are presented as the absolute tissue amount and expressed as mean  $\pm$  S.E. Mass spectra are included for proteins adenylate kinase,  $\alpha$ -actin, and tropomyosin A to highlight the difference in relative signal intensity. For proteins expressed as 0 nmol/g, ions corresponding to analyte peptides were not present in the spectrum (see spectral data for tropomyosin (TM) A). TPI, triose-phosphate isomerase; Eno, enolase; Myo, myosin; API, actin polymerization inhibitor; PK, pyruvate kinase; CK, creatine kinase; LDH, lactate dehydrogenase; GP, glycogen phosphorylase; H, heavy; L, light.

ing to analyte-QconCAT pairs (Fig. 5). Quantification data for four proteins over 30 days of growth obtained by both methods of MALDI-TOF MS and LC-ESI-Q-TOF MS showed excellent agreement such that the correlation coefficient was 0.977 (Fig. 6). All proteins that could be quantified by MALDI-TOF MS (with and without guanidination) and LC-ESI-Q-TOF MS were expressed as nmol/g of pectoralis muscle tissue. The data were obtained during growth from 1 to 30 days post-

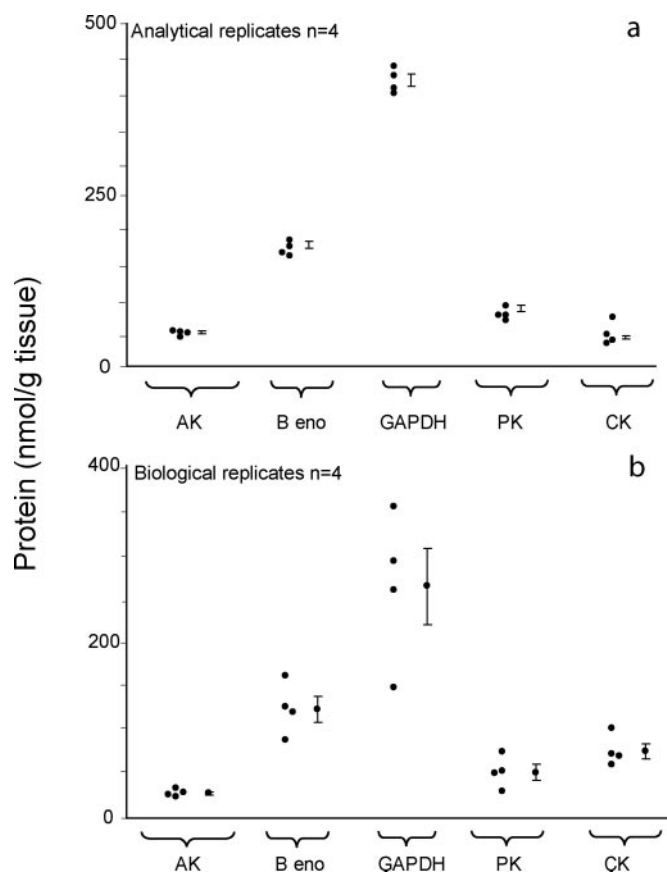


FIG. 8. **Sources of variance in a QconCAT experiment.** Soluble protein from chicken pectoralis muscle (70  $\mu\text{g}$ ) was mixed with QconCAT protein (7  $\mu\text{g}$ ) in four technically replicated experiments and digested to completion with trypsin. For each protein, individual data points are plotted to the left of mean  $\pm$  S.E. for the same bird where  $n = 4$  (a) and for four different birds to demonstrate biological variance (b). *B eno*,  $\beta$ -enolase; *CK*, creatine kinase; *PK*, pyruvate kinase.

hatch for four birds at each time point for chickens of the layer and broiler strains (Fig. 7). Some proteins demonstrated massive pool expansion, whereas others declined to a similar degree, covering a measurable dynamic range across all proteins of 10–550 nmol/g for a single protein (GAPDH) and as low as  $2 \pm 1$  nmol/g ( $\alpha$ -enolase; 1-day broiler). Thus, in a single experiment we were able to assess protein concentrations over a 300-fold range.

To assess variance due to the analytical procedure, four identical protein mixtures (100  $\mu\text{l}$  of chicken skeletal muscle (2.6  $\mu\text{g}/\mu\text{l}$ ) with 9  $\mu\text{l}$  of QconCAT (2.9  $\mu\text{g}/\mu\text{l}$ )) were digested with trypsin, and the surrogate peptides were used to quantify proteins by MALDI-TOF MS. Quantification data were collected and used to assess analytical variance (Fig. 8a). The reproducibility of the method was high, and the variance was similar whether four or 10 replicates were used. In both instances, the analytical variance was significantly lower than that for quantification data measured for four different birds of each strain (Fig. 8b). For example, the analytical variance (CV of 6.0% for  $\beta$ -enolase,  $n = 4$ ) compared favorably to biolog-

ical variance (CV of 24.0% for  $\beta$ -enolase,  $n = 4$ ). Increasing the number of analytical replicates to 10 had very little effect on analytical variance (CV of 6.0% for  $\beta$ -enolase,  $n = 10$ ; data not shown).

For some aspects of quantitative proteomics, MALDI-TOF MS has advantages. Data can be accumulated for a variable number of laser shots, ensuring comparable signal intensities between replicates. Virtually all of the signal resides in the singly charged  $[M + H]^+$  ion, whereas with electrospray ionization, the signal can be distributed over a number of differently charged species. However, for complex analytical mixtures, the complexity of a MALDI-TOF mass spectrum, coupled with a noisy signal base line, can compromise quantification. One approach to simplification of a MALDI-TOF MS analysis relies on prior fractionation of the peptide mixture before deposition of successive fractions on the MALDI target (25). Chicken skeletal muscle with added QconCAT was digested and separated by reversed phase liquid chromatography, and fractions (200 nl) were collected onto a MALDI target at 1-min intervals for analysis by MALDI-TOF MS (Fig. 9). This provided an efficient detection system with peptides fixed in the solid phase for continued interrogation when acquiring data for quantification. LC-MALDI-TOF MS was used for analysis of a single chicken skeletal muscle sample to highlight the potential benefit of this method. This approach allowed quantification of the majority of proteins selected for incorporation into the QconCAT protein and consequently contributed additional information for quantification. Comparing quantification by LC-MALDI-TOF MS with both MALDI-TOF MS and LC-ESI-Q-TOF MS confirmed that all three methods of analysis give consistent and comparable quantification. This quantification can be subtle, for example in monitoring isoform changes from embryonic to adult myosin as well as a change in state from free, soluble protein to that assembled within the myofibrillar apparatus (actin). It is also possible to monitor expression of isoforms of the same enzyme for which Qpeptides differ only in a single amino acid (lactate dehydrogenases A and B).

Although there is nothing formally different between a chemically synthesized peptide and a peptide excised from a QconCAT by proteolysis, we compared the quantification of a single protein (GAPDH, which exhibits a dramatic change in abundance during posthatching development) using the QconCAT-derived peptide and the identical synthetic peptide. The correlation between data obtained using QconCAT and that obtained using the synthetic peptide was high (correlation coefficient, 0.998) (Fig. 10), and quantification data were consistent using either internal standard. A small consistent discrepancy (less than 10%) between the two methods could be attributable to the method of quantification used for the two standards. The discrepancy between the synthetic peptide and the QconCAT was reduced if we used the latter to quantify the former but was still present. We do not have an explanation for this residual discrepancy at present. We are confident, however, that the discrepancy is not attributable to



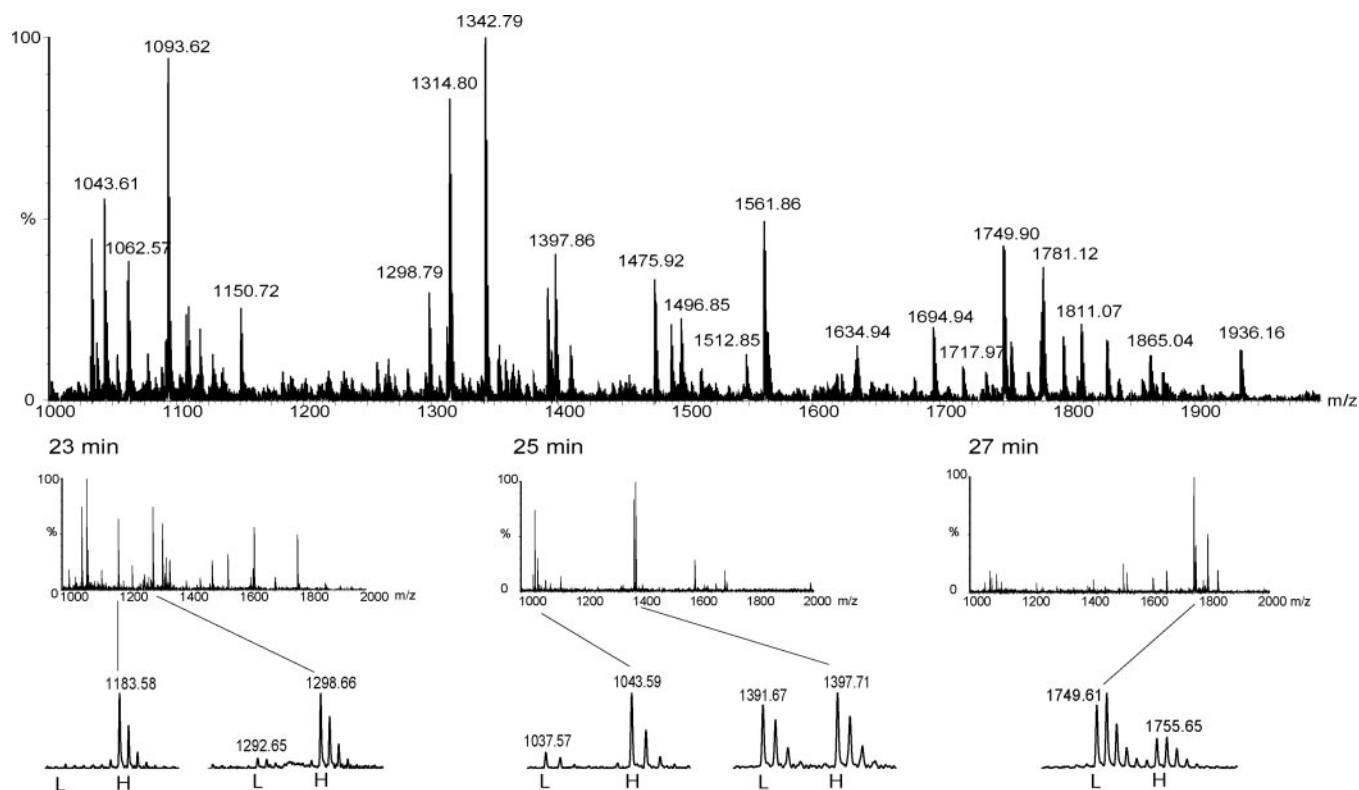


FIG. 9. **Quantification of proteins by LC-MALDI-TOF MS.** QconCAT protein ( $7 \mu\text{g}$ ) was added to a preparation of chicken skeletal muscle soluble fraction ( $70 \mu\text{g}$  of protein) in a ratio of 1:10. This mixture was digested with trypsin and analyzed by LC-MALDI-TOF MS. Peptides were separated over a 50-min reversed phase acetonitrile gradient (0–100%), and fractions of 1 min (200 nl) were collected directly onto a Waters MALDI target. The *upper panel* is the MALDI-TOF mass spectrum of the entire digest; the *lower panels* illustrate three fractions collected from the reversed phase eluate at 23, 25, and 27 min. Representative pairs of analyte-standard peptides are highlighted. *H*, heavy; *L*, light.

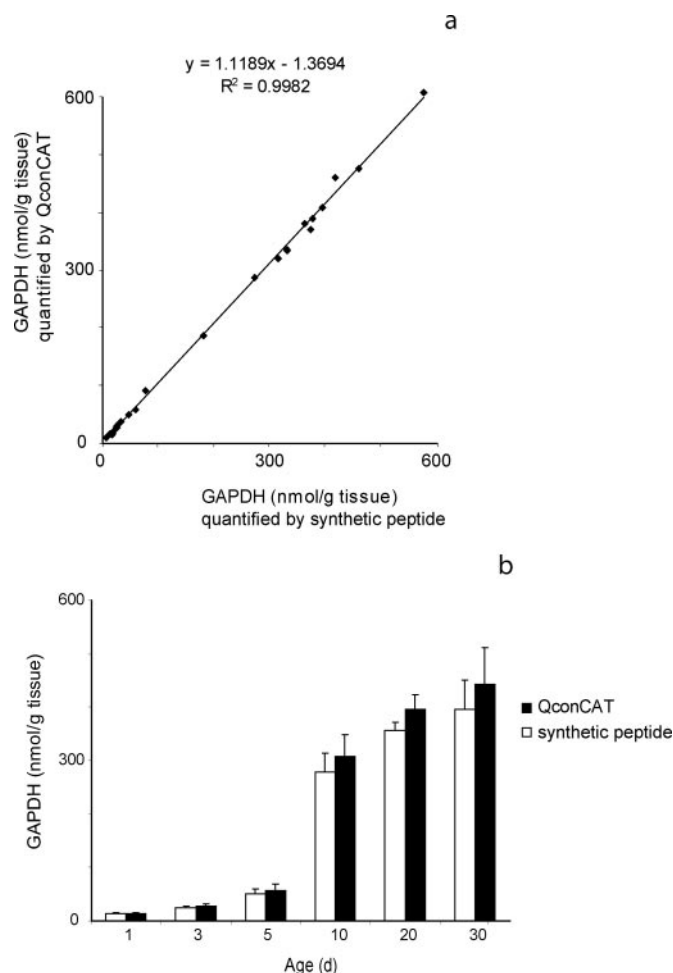
incomplete digestion of the QconCAT (see Figs. 3 and 4). In the case of the QconCAT, we used a protein assay to determine the amount of protein as this was the same method used to quantify total protein in the analyte. For the synthetic peptide, the quantity supplied by the manufacturer is too small for independent quantification, and it was necessary to assume that the quantity in the vial was indeed that specified by the manufacturer. The difference between the two standards was minor compared with the biological variance within the system, would not contribute significant errors, and would be readily controlled by alternative QconCAT quantification strategies (see “Discussion”).

To assess the accuracy of a QconCAT experiment for quantification, we spiked a known amount of AK into chicken skeletal muscle soluble fraction from a 30-day broiler. The amount of AK added was converted into protein concentration as nmol/g tissue and compared with the total concentration of AK in the tissue (nmol/g) as quantified using QconCAT (Fig. 11). As expected, there was a strong correlation ( $R^2 = 0.9992$ ) with a slope of 1, indicating the lack of any systematic quenching effects over an extended dynamic range. Quantification of selected muscle proteins by the QconCAT strategy was also compared with densitometric quantification from 1D SDS-PAGE; the correlation of these methods was poor (data

not shown;  $R^2 = 0.67$ ), although the stain intensity was strongly proportional to the amount of protein loaded on the gel (data not shown;  $R^2 = 0.995$ ). This is most probably due to the different affinity of individual proteins for the stain.

#### DISCUSSION

QconCAT methodology has considerable potential to enhance the scope and scale of quantitative proteomics by multiplexing stable isotope dilution assays using proteotypic peptides as surrogates for the proteins of interest. The specific novelty of the QconCAT approach is derived from the efficient means of simultaneous production of multiple internal standards. Unlike chemical synthesis, biological synthesis *de novo* is not beset by “difficult” peptides (for example those with runs of serine residues or with a large hydrophobic content) that can be problematic to synthesize chemically in high purity. Moreover QconCAT proteins can be labeled using any metabolic precursor from the remarkably inexpensive uniform  $^{15}\text{N}$  labeling using  $^{15}\text{NH}_4\text{Cl}$  as the sole nitrogen source in the medium to specific labeling with  $[^{13}\text{C}_6]\text{Lys}/[^{13}\text{C}_6]\text{Arg}$ , which ensure that, for tryptic proteotypic peptides, each has a constant mass offset of 6 Da. Incorporation of a second labeled amino acid that is variably represented in the QconCAT can also facilitate mass isolation of the standard.



**FIG. 10. Comparison of QconCAT and synthetic peptide for quantification.** For one protein (glyceraldehyde-3-phosphate dehydrogenase), quantification was achieved relative to a QconCAT peptide and the same peptide chemically synthesized. For both methods, the internal standards ( $2 \mu\text{g}$  of QconCAT protein or  $0.05 \mu\text{g}$  of synthetic peptide) were added to  $20 \mu\text{g}$  of chicken skeletal muscle soluble protein prior to digestion with trypsin, and data were acquired using MALDI-TOF MS. The quantification by either method correlated strongly (*upper panel*). The time-dependent developmental expansion of GAPDH (nmol/g of tissue, mean  $\pm$  S.E.,  $n = 4$ ) in broiler was monitored by QconCAT or synthetic peptide (*lower panel*). *d*, days.

The imaginative application of metabolic labeling without the need for resynthesis of the QconCAT gene is an advantage of the approach that has yet to be fully exploited.

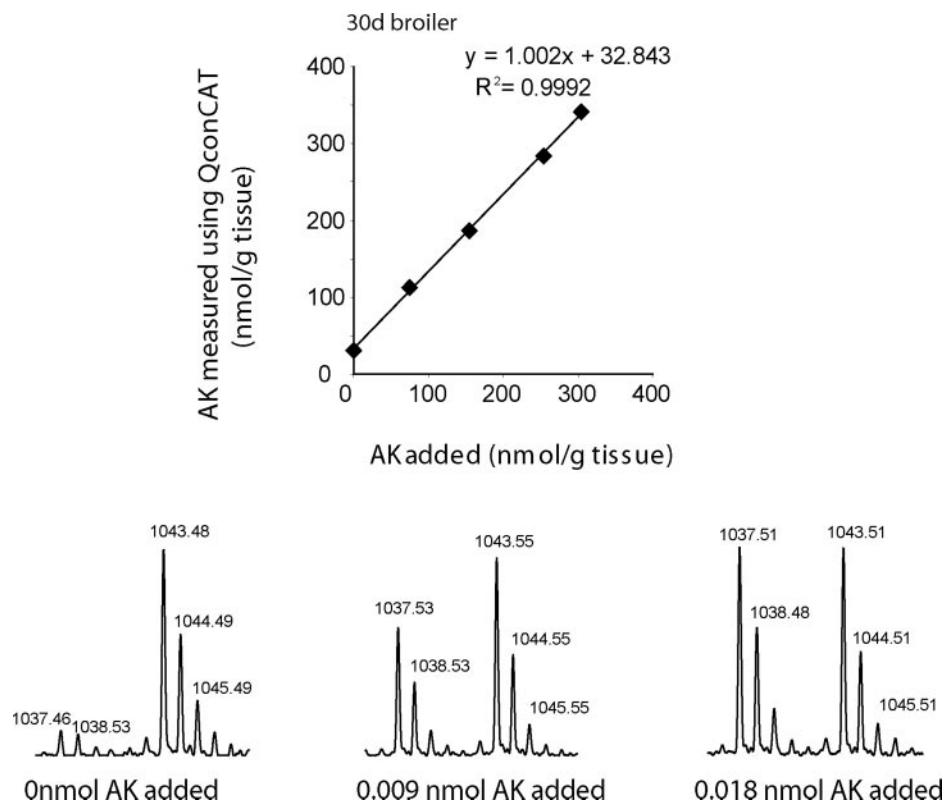
The QconCAT approach is robust to the choice of mass spectrometric method used. Each of the three methods used (MALDI-TOF MS, LC-ESI-Q-TOF MS, and LC-MALDI-TOF MS) allows quantification of individual proteins that are not detected using the alternative techniques, for example adult myosin and pyruvate kinase have only been quantified using MALDI-TOF MS, myosin-binding protein C and phosphoglycerate kinase have only been quantified using LC-MALDI-TOF MS (data not shown), and lactate dehydrogenase A has only been quantified using LC-ESI-Q-TOF MS. The ability to detect

and quantify each peptide incorporated in the original QconCAT protein is very dependent on the analytical context. Although the target ionization method can influence the choice of proteotypic peptides, the opportunity remains to switch to other separation and ionization methods to gain quantification data for large numbers of proteins.

Although a chemically synthesized peptide and a QconCAT peptide are formally equivalent at the analytical stage, we compared the two approaches. Interestingly the two methods gave highly precise estimates of protein levels, but the measured values were different such that there was a consistently lower estimate of protein amount using the synthetic peptide when compared with QconCAT. Quantification of the QconCAT protein was achieved by colorimetric assay using the same method as used for the assessment of total protein concentration in the biological samples. Quantification of the synthetic peptide is based on amino acid analysis conducted by the supplier and was completed separately and prior to analysis with the mixture of analyte proteins. Indeed the quantity of the synthetic peptide supplied (five vials of  $1 \text{ nmol}$ ) was sufficiently low that independent quantification by the end user could be problematical. By contrast, we routinely prepare  $5\text{--}10 \text{ mg}$  (approximately  $250 \text{ nmol}$ ) of the QconCAT used here. The errors introduced by the method of standard quantification are small and, relative to the biological changes we measure here, are not significant. However, future iterations of QconCAT proteins will incorporate a common peptide for internal standard quantification by a synthetic peptide that can be labeled or unlabeled, depending on the labeling status of the QconCAT protein. This peptide, chosen because it ionizes well under MALDI or ESI, could then be used to quantify each QconCAT, normalizing all QconCAT data to a common, absolute standard. This common peptide, which is chemically synthesized, would be required in large amounts, and as such, purification and quantification of this peptide could be conducted to a very high level of confidence. By creating such a “gold standard” for quantification, data from all laboratories using QconCATs (the same or different) could be compared directly.

The application to absolute quantification of multiple proteins within complex biological systems and the adaptable nature of the QconCAT to a variety of analytical systems is clear. For development of strategies for absolute quantification, the QconCAT method provides a reproducible and relatively simple system in which multiple proteins can be quantified using alternative methods of mass spectrometry with chromatographic separation and chemical derivatization. We have made approximate estimates of the costs involved, and to quantify 50 proteins at one Qpeptide per protein, a QconCAT strategy is about 15% of the cost of comparable synthetic peptides and would yield about  $250 \text{ nmol}$  of protein compared with  $5 \text{ nmol}$  of each synthetic peptide. The error in analytical replicates is small, but there can be no “holy grail” target performance of the analytical analyses. First provided that analytical variance can be demonstrated to be substan-

FIG. 11. Investigation of the accuracy of quantification using QconCAT. Purified AK was added to chicken skeletal muscle soluble fraction from a 30-day (*d*) broiler. AK was added from 0 to 0.02 nmol, which resulted in a final protein concentration of 0–300 nmol/g, and the amount of AK in the tissue was quantified by adding 0.015 nmol of QconCAT prior to digestion with trypsin. Proteolysis was allowed to continue for 24 h after which peptides were analyzed by MALDI-TOF MS. The upper panel shows the correlation between AK added and that quantified in the muscle using QconCAT after digestion with trypsin. Spectra showing the change in MALDI-TOF mass spectral signal intensity over the range of protein concentrations used in this experiment are shown beneath.



tially smaller than biological variance (as we have demonstrated here and indeed the normal expectation), it might be argued that there is a much reduced need to perform analytical replicates and that the effort should be directed toward acquisition of greater biological insight by adding new biological replicates.

For identification proteomics, little regard is paid to the completeness of the digestion of the analyte peptide; the goal is to generate sufficient peptides that are readily ionized and/or fragmented for unambiguous identification. Indeed most search engines are tolerant of and include options to match one or more than one “missed cleavage.” However, when the goal shifts to the more demanding task of peptide-based quantification, it is essential that due cognizance is given to the proteolytic reactions that generate the peptides that are to be used for quantification irrespective of the method. The goal has to be complete digestion, and there are several approaches that can be taken to ensure that this has occurred. This should also be checked experimentally. It would also be feasible to embed two Qpeptides for each protein in a single QconCAT or even two different QconCATs to enhance confidence, but we do not subscribe to the view that this is necessary in many instances. Finally our experience with a large number of QconCATs<sup>2</sup> is that they are proteolyzed at rates that are far higher than analyte proteins.

<sup>2</sup> J. Rivers, D. M. Simpson, D. H. L. Robertson, S. J. Gaskell, and R. J. Beynon, unpublished observations.

In all of these QconCAT constructs, we made no attempt to preserve the primary sequence context of the Qpeptides, and it is clear that this is not an important factor in QconCAT design; the selection of suitable proteotypic peptides in the design phase is much more critical. In this regard, the recent work by Aebersold and co-workers (26) points the way toward more effective nomination of Qpeptides for QconCATs.

**Acknowledgments**—We thank the Roslin Institute, Edinburgh, UK for growing birds and collecting muscle samples. We are grateful to Dr. Dominique Rocha, Genus plc, for interest in this work.

\* This work was supported by Biotechnology and Biological Sciences Research Council Grant BB/C007433/1 (to R. J. B. and S. J. G.) and by Genus plc. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed: Proteomics and Functional Genomics Group, Faculty of Veterinary Science, University of Liverpool, Crown St., Liverpool L69 7ZJ, UK. Tel.: 44-151-794-4312; Fax: 44-151-794-4243; E-mail: r.beynon@liv.ac.uk.

#### REFERENCES

- Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **1**, 376–386
- Sechi, S., and Oda, Y. (2003) Quantitative proteomics using mass spectrometry. *Curr. Opin. Chem. Biol.* **7**, 70–77
- Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**, 994–999

4. Cagney, G., and Emili, A. (2002) De novo peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging. *Nat. Biotechnol.* **20**, 163–170
5. Higgs, R. E., Knierman, M. D., Gelfanova, V., Butler, J. P., and Hale, J. E. (2005) Comprehensive label-free method for the relative quantification of proteins from biological samples. *J. Proteome Res.* **4**, 1442–1450
6. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 6940–6945
7. Kuhn, E., Wu, J., Karl, J., Liao, H., Zolg, W., and Guild, B. (2004) Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and <sup>13</sup>C-labeled peptide standards. *Proteomics* **4**, 1175–1186
8. Ishihama, Y., Sato, T., Tabata, T., Miyamoto, N., Sagane, K., Nagasu, T., and Oda, Y. (2005) Quantitative mouse brain proteomics using culture-derived isotope tags as internal standards. *Nat. Biotechnol.* **23**, 617–621
9. Kirkpatrick, D. S., Gerber, S. A., and Gygi, S. P. (2005) The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods (Duluth)* **35**, 265–273
10. Pan, S., Zhang, H., Rush, J., Eng, J., Zhang, N., Patterson, D., Comb, M. J., and Aebersold, R. (2005) High throughput proteome screening for biomarker detection. *Mol. Cell. Proteomics* **4**, 182–190
11. Beynon, R. J., Doherty, M. K., Pratt, J. M., and Gaskell, S. J. (2005) Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. *Nat. Methods* **2**, 587–589
12. Pratt, J. M., Simpson, D. M., Doherty, M. K., J., R., Gaskell, S. J., and Beynon, R. J. (2006) Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nat. Protocols* **1**, 1029–1043
13. Kito, K., Ota, K., Fujita, T., and Ito, T. (2007) A synthetic protein approach toward accurate mass spectrometric quantification of component stoichiometry of multiprotein complexes. *J. Proteome Res.* **6**, 792–800
14. Hubbard, S. J., and Beynon, R. J. (2001) Proteolysis of native proteins as a structural probe, in *Proteolytic Enzymes. A Practical Approach* (Beynon, R. J., and Bond, J. S., eds) pp. 233–264, Oxford University Press, Oxford
15. Hubbard, S. J. (1998) The structural aspects of limited proteolysis of native proteins. *Biochim. Biophys. Acta* **1382**, 191–206
16. Zappacosta, F., Pessi, A., Bianchi, E., Venturini, S., Sollazzo, M., Tramontano, A., Marino, G., and Pucci, P. (1996) Probing the tertiary structure of proteins by limited proteolysis and mass spectrometry: the case of Minibody. *Protein Sci.* **5**, 802–813
17. Hubbard, S. J., Beynon, R. J., and Thornton, J. M. (1998) Assessment of conformational parameters as predictors of limited proteolytic sites in native protein structures. *Protein Eng.* **11**, 349–359
18. Wu, C., Robertson, D. H., Hubbard, S. J., Gaskell, S. J., and Beynon, R. J. (1999) Proteolysis of native proteins. Trapping of a reaction intermediate. *J. Biol. Chem.* **274**, 1108–1115
19. Brancia, F. L., Oliver, S. G., and Gaskell, S. J. (2000) Improved matrix-assisted laser desorption/ionization mass spectrometric analysis of tryptic hydrolysates of proteins following guanidination of lysine-containing peptides. *Rapid Commun. Mass Spectrom.* **14**, 2070–2073
20. Hale, J. E., Butler, J. P., Knierman, M. D., and Becker, G. W. (2000) Increased sensitivity of tryptic peptide detection by MALDI-TOF mass spectrometry is achieved by conversion of lysine to homoarginine. *Anal. Biochem.* **287**, 110–117
21. Doherty, M. K., McLean, L., Hayter, J. R., Pratt, J. M., Robertson, D. H., El-Shafei, A., Gaskell, S. J., and Beynon, R. J. (2004) The proteome of chicken skeletal muscle: changes in soluble protein expression during growth in a layer strain. *Proteomics* **4**, 2082–2093
22. McLean, L., Doherty, M. K., Deeming, D. C., and Beynon, R. J. (2004) A proteome analysis of the subcutaneous gel in avian hatchlings. *Mol. Cell. Proteomics* **3**, 250–256
23. Baumgart, S., Lindner, Y., Kuhne, R., Oberemm, A., Wenschuh, H., and Krause, E. (2004) The contributions of specific amino acid side chains to signal intensities of peptides in matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **18**, 863–868
24. Krause, E., Wenschuh, H., and Jungblut, P. R. (1999) The dominance of arginine-containing peptides in MALDI-derived tryptic mass fingerprints of proteins. *Anal. Chem.* **71**, 4160–4165
25. Mirgorodskaya, E., Braeuer, C., Fucini, P., Lehrach, H., and Gobom, J. (2005) Nanoflow liquid chromatography coupled to matrix-assisted laser desorption/ionization mass spectrometry: sample preparation, data analysis, and application to the analysis of complex peptide mixtures. *Proteomics* **5**, 399–408
26. Mallick, P., Schirle, M., Chen, S. S., Flory, M. R., Lee, H., Martin, D., Ranish, J., Raught, B., Schmitt, R., Werner, T., Kuster, B., and Aebersold, R. (2007) Computational prediction of proteotypic peptides for quantitative proteomics. *Nat. Biotechnol.* **25**, 125–131