Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes

Julie M Pratt[1], Deborah M Simpson[1], Mary K Doherty[1], Jenny Rivers[1], Simon J Gaskell[2] & Robert J Beynon[1]

[1]Department of Veterinary Preclinical Sciences, University of Liverpool, Crown Street, Liverpool, L69 7ZJ, UK. [2]Michael Barber Centre for Mass Spectrometry, School of Chemistry, University of Manchester, Manchester M13 9PL, UK. Correspondence should be addressed to R.B. (r.beynon@liv.ac.uk), website URL http://www.liv.ac.uk/pfg.

Published online XX XX 2006; doi:10.1038/nprot.2006.129

An important area of proteomics involves the need for quantification, whether relative or absolute. Many methods now exist for relative quantification, but to support biomarker proteomics and systems biology, absolute quantification rather than relative quantification is required. Absolute quantification usually involves the concomitant mass spectrometric determination of signature proteolytic peptides and stable isotope-labeled analogs. However, the availability of standard labeled signature peptides in accurately known amounts is a limitation to the widespread adoption of this approach. We describe the design and synthesis of artificial QconCAT proteins that are concatamers of tryptic peptides for several proteins. This protocol details the methods for the design, expression, labeling, purification, characterization and use of the QconCATs in the absolute quantification of complex protein mixtures. The total time required to complete this protocol (from the receipt of the QconCAT expression plasmid to the absolute quantification of the set of proteins encoded by the QconCAT protein in an analyte sample) is ~29 d.

INTRODUCTION

Most proteomics studies to date have delivered relative quantification, expressing the changes in the amount of proteins in the context of a second cellular state or control sample[1–3]. Such studies do not facilitate the generation of large databases of results, with data not being transferable between different laboratories.

If proteomics is to support the emergent fields of protein biomarker discovery (whether in medical diagnostics or toxicology for drug discovery) or to provide the rigorous data needed for systems biology, absolute quantification is needed. Absolute quantification relies on the well-established precepts of stable isotope dilution, specifically the use of labeled peptide internal standards that are characterized and quantified by mass spectrometry (MS)[4,5]. These internal standards are currently synthesized *de novo* by chemical methods. This requires the individual synthesis in labeled form, purification and quantification, of each peptide for use as an internal standard. Complex studies would require the synthesis of large numbers of peptides at significant cost, and each would have to be quantified individually.

We describe here a method for the design, expression and use of artificial proteins (QconCATs) that are concatamers of Q peptides, generated by chemical or endoproteolytic cleavage, for a group of proteins under study[6]. The QconCAT proteins are expressed in *Escherichia coli* and are readily labeled with stable isotopes by growth in the presence of stable isotope–labeled precursors. The labeled QconCAT proteins are then purified, quantified and added to complex protein mixtures in known amounts. Endoproteolytic (and/or chemical) fragmentation of the QconCAT–analyte mix releases each of the QconCAT peptides in a strict stoichiometry of 1:1, and MS analysis allows the quantification of each represented peptide of the analyte (see **Fig. 1** for a schematic of the overall process). In the specific protocol described here we focus on the use of trypsin as the endoproteinase, but other endoproteinases and/or chemical cleavage could be used as an alternative method of generating suitably sized peptides from the analyte and identical peptides, to act as internal standards, from the labeled QconCAT protein.

Unlabeled QconCAT proteins could also provide the basis for absolute quantification if differential isotope labeling via derivatization of proteins or peptides is incorporated in the analytical procedure. Use of the same QconCAT protein for each quantification experiment allows direct comparison of results between each experiment and between laboratories, and will facilitate the construction of the large data sets needed for toxicological and diagnostic studies in which even control samples vary considerably. Later, we will describe a refinement to the technique that even allows the quantification of the QconCAT standard between laboratories. Because the quantification data are in absolute terms—expressed, for example, as picomoles of protein per gram of tissue or per number of cells—knowledge of the number of molecules of a protein present per cell in a given state will underpin the generation of testable mathematical models in systems biology.

APPLICATIONS

QconCATs can be applied to the absolute quantification of any peptide that can be reproducibly generated by endoproteolytic or chemical cleavage from any sample source. Here we discuss only a sampling of the potential applications.

General

- In comparative proteomics the shift from relative to absolute quantification permits comparison of results, not only between different cellular states within an experiment, but also between different experiments and different laboratories.
- An unlabeled QconCAT could be subjected to labeling *in vitro* using one of the many reagents that have been advocated for comparative proteomics, enhancing all of these technologies to absolute quantification.

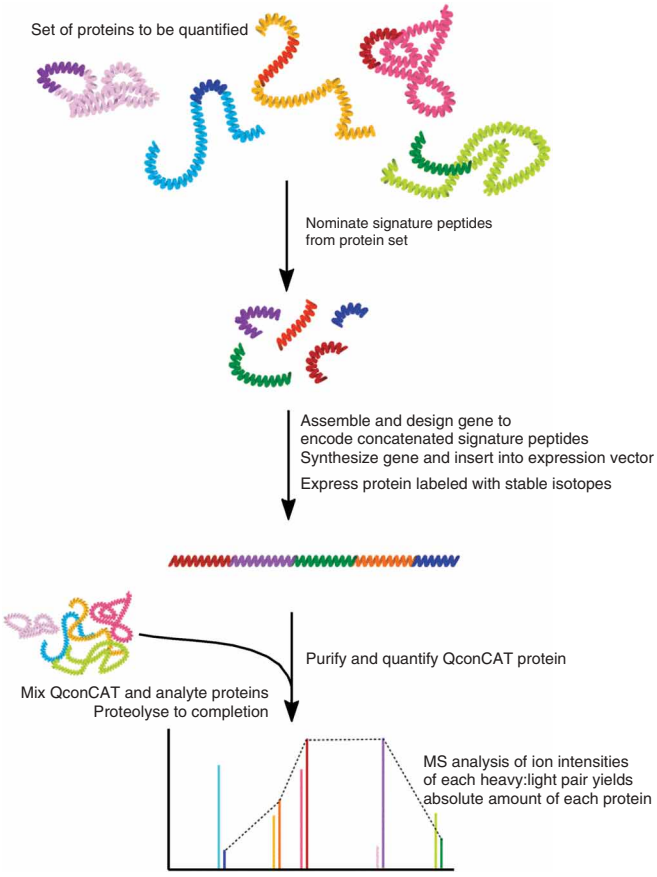

Figure 1 | General principle of a QconCAT quantification experiment.

- Systems biology requires absolute quantification, including the definition of processes in molecular terms and the generation of data to allow mathematical modeling.
- Determination of the absolute stoichiometry of components in subcellular particles requires absolute quantification.
- Analysis of isoform expression patterns (e.g., cytochrome P450s) is possible using this process. Each QconCAT peptide would report on a peptide that contains the variant amino acid(s), permitting one to distinguish among several variants.
- QconCATs could be used to quantify many or all proteins expressed in a control strain or cell culture grown under standardized conditions. Subsequently, the same strain or cell culture would be labeled by growth in the presence of stable isotope precursors under the same conditions. These fully labeled cells would then be used as quantification standards in comparative studies. The indirect quantification approach has the advantage that, for the second quantification phase, any labeled peptide (not only those embedded in QconCATs) could be used, providing the opportunity for alternative enrichment and quantification approaches.

Medical

- Diagnostic/protein biomarker validation and analysis could be facilitated by QconCATs for accurately measuring changes in the protein components of body fluids and thereby adding another dimension to computer training sets for identifying patterns indicative of disease.
- For routine diagnostic tests, QconCATs could allow determination of levels of key biomarkers in clinical samples including blood, urine, cerebrospinal fluid (CSF), synovial fluid and bronchoalveolar lavage.
- QconCATs could allow the monitoring of changes in protein levels in response to exposure to drugs, as well as identifying changes relating to toxicology and routine monitoring of protein biomarkers to accelerate the process of drug discovery.

Agricultural

This technology could be used in routine determination of levels of proteinaceous contamination from other sources in foodstuffs, for example, signatures from genetically modified sources. It would also facilitate monitoring of the effect of pesticides and herbicides.

Protein chemistry

QconCATs introduce the concept of 'designer protein' to proteomics. Specific QconCAT proteins could therefore be used to assess the properties of defined peptides in MS and MS/MS analysis.

Post-translational modifications

The use of QconCATs to study some post-translational modifications is limited, because only the unmodified analyte peptide can be quantified using a Q-peptide. Nevertheless, we can envisage two methods of using QconCATs to provide a preliminary study of post-translational modifications. One approach involves selecting a minimum of two peptides for each protein of interest; one peptide is not modified and the other is modifiable. Comparison of the quantification of the two peptides should allow the determination of the amount of modified peptide present. The second approach is to select the modified peptide as the Q-peptide to report on the protein and carry out the quantification before and after treatment of the sample to remove post-translational modifications—for example, alkaline phosphatase to hydrolyze phosphopeptides. Quantification before and after treatment will reveal the quantity of protein in the modified state. These methods are not, however, immediately applicable to the unraveling of more complex patterns of post-translational modification, in which several sites of modification are possible on the same peptide.

For some modifications, such as the primary sequence change elicited by limited proteolysis (such as apoptosis), it might be argued that QconCAT-type approaches are the only ways to accurately quantify the proteolysed and unproteolysed variants of the target analyte.

EXPERIMENTAL DESIGN

There are two stages in a QconCAT experiment: the design phase and the implementation phase leading to the construction of a QconCAT (**Fig. 2**). Once the design exists as a DNA sequence it can then be inserted into a suitable vector, transferred to a suitable expression system, expressed, labeled and used in quantification experiments. The first stage reflects the analytical and decision-making processes that encompass the nomination of candidate proteins and peptides, followed by their realization as a codon-optimized sequence in which opportunities for stable RNA secondary structure formation are minimized. A Gantt chart outlining the timeline for the procedure is given in **Figure 3**.

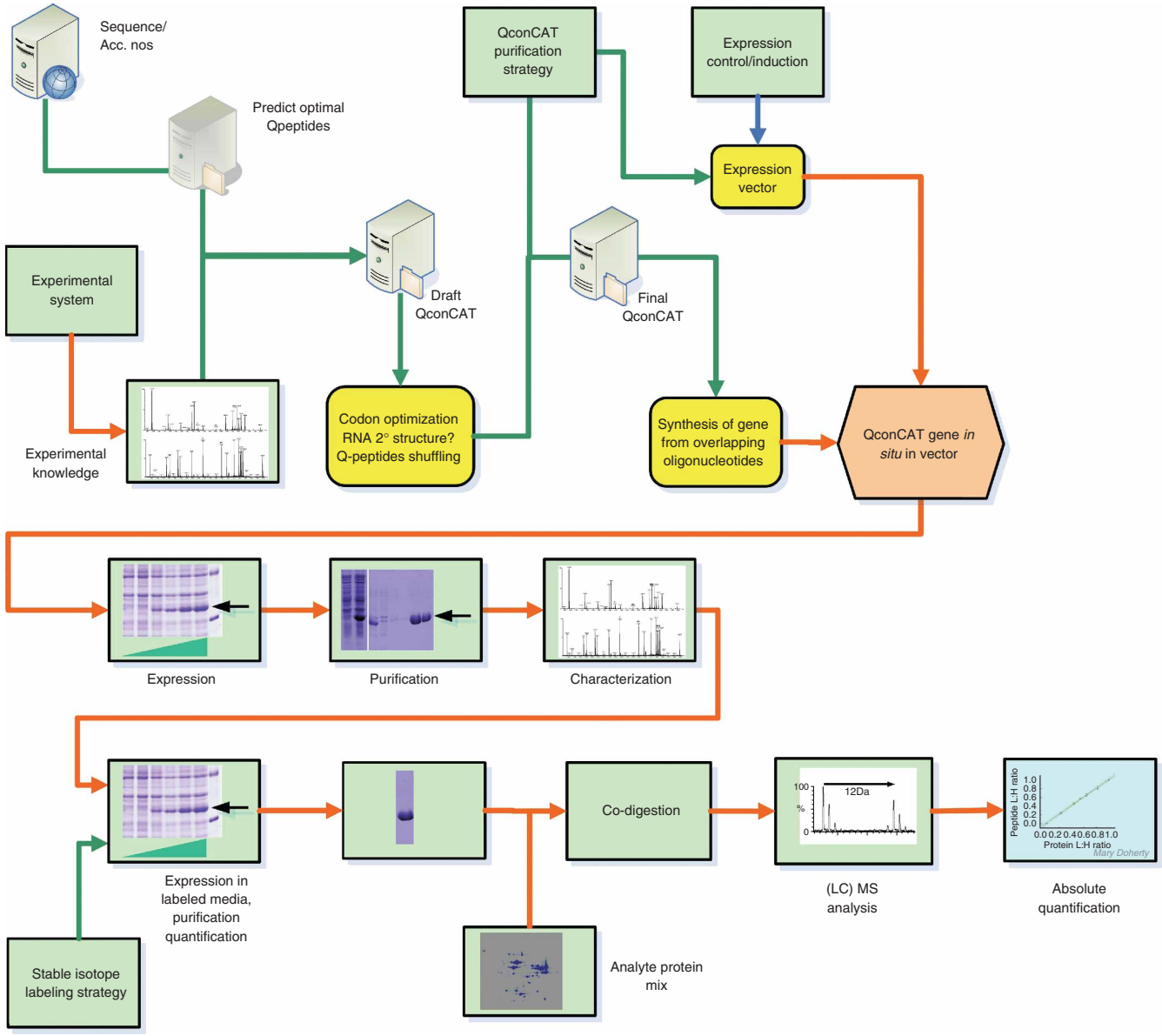

Figure 2 | Overall workflow for QconCAT quantification. A QconCAT experiment consists of the design phase and implementation phase.

Design of QconCAT genes and construction of plasmids encoding QconCAT proteins

To quantify the group of proteins, a unique peptide (a Q-peptide) must be selected as a surrogate 'signature' peptide for each protein to be quantified. This selection depends on the inherent properties of the peptide and the cleavage method chosen for the analysis of the analyte. There are no restrictions on the protein-containing samples that can be quantified using QconCAT technology. However, in the absence of complete genome information, full sequencing of the signature peptide (probably by tandem MS) would be required before incorporation of this peptide into a designed QconCAT.

Selection of cleavage method

Most proteomics studies have used trypsin as the method of cleavage of proteins into peptides suitable for MS analysis. Trypsin has the advantage of a rigorously expressed sequence specificity (cleaves C-terminal to arginyl and lysyl residues, except arginyl-prolyl and lysyl-prolyl sequences, which are not cleaved) and minimal autoproteolysis (recombinant trypsin). For many proteins cleavage with trypsin generates a good distribution of peptides within the range 600–4,000 Da, ideal for analysis by MS. Because only one peptide is required in a QconCAT protein to represent each protein under study, if trypsin does generate a suitably sized peptide, with the additional properties described in the next section, trypsin will be the endoproteinase of choice. Moreover, tryptic peptides readily generate doubly charged ions ($[M+2H]^+$) that extend the usefulness from MS to MS/MS analysis and increase the scope of quantification to include monitoring of single or multiple reactions. However, for integral membrane proteins, with several membrane-spanning regions, there is an under-representation of lysyl and arginyl residues. In addition, cleavage sites may be

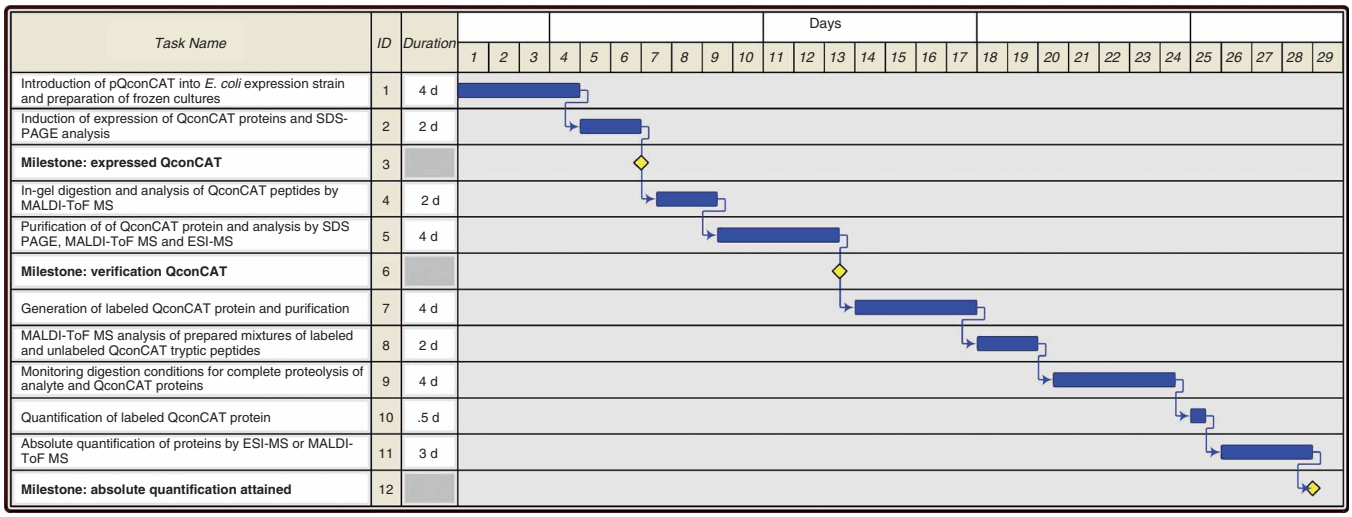

Figure 3 | Gantt chart for a typical QconCAT experiment.

inaccessible to endoproteinases because of steric hindrance from the lipid bilayer or localization of cleavable sequences within the lumen of membrane vesicles. More efficient study of these proteins requires alternative solubilization[7,8] and cleavage strategies[9]. Methods that have been developed include (i) chemical cleavage at methionine residues using cyanogen bromide[7], in combination with trypsin[9]-chymotrypsin (which cleaves C-terminally at phenylalanyl, tyrosyl and tryptophanyl residues), and (ii) staphylococcal peptidase I (which cleaves at glutamyl residues), either in combination or singly[9]. However, for some of these peptidases, the specificity is sufficiently relaxed that preliminary experiments will be needed to ensure exclusive release of a suitable peptide (with no overlapping or alternative cleavages). The cleavage strategy should therefore be decided and tested for the group of proteins to be quantified before QconCATs are designed, so that the appropriate peptides can be selected for inclusion to match the planned analyte protein fragmentation strategy. A relaxed specificity can be tolerated provided that the same peptide is generated quantitatively from both QconCAT and analyte; in practice this might be difficult to achieve.

Selection of Q-peptides

Q-peptides are selected for uniqueness of mass, propensity to ionize and detectability in MS, the presence of specific amino acid residues required for labeling and the absence of amino acid residues that can cause problems (cysteine, methionine) (see http://www.qconcat.com for additional guidelines). The propensity of a selected peptide to ionize is often known, because the sample has been analyzed by MS previously. It is not yet possible to predict ion intensities using knowledge-based approaches, but this may make peptide selection easier in the future. At the moment we favor a pragmatic solution. The inclusion of two Q-peptides per protein might reduce the risk associated with poorly ionizing or otherwise difficult peptides. In addition, an improvement in ion intensity of peptides lacking arginine, but carrying lysine, can be achieved by guanidination, and this can be included as a relatively simple step before analysis by MS[10–12].

In some instances it will be impossible to avoid inclusion of peptides containing cysteinyl and methionyl residues, for example. For these QconCATs it is essential to be aware of the need to maintain conditions that ensure that the oxidation state of the Q-peptide is the same as the analyte, because any discrepancy between the two would compromise the quantification. Finally, for those groups that have been using individual chemically synthesized peptides, for comparison of quantifications using QconCATs with chemically synthesized stable isotope-labeled peptides, a peptide available in purified form could be included in the QconCAT sequence and its behavior as part of a QconCAT compared to that when present in an analyte mixture, as a purified, quantified peptide.

Construction of QconCAT genes

The Q-peptide sequences once selected are randomly concatenated *in silico* and used to direct the design of a gene, codon-optimized for expression in *E. coli*. The predicted transcript is analyzed for RNA secondary structure that might diminish expression, and if undesirable secondary structure features are present, the order of the peptides is altered. N-and C- terminal sequences are added as sacrificial structures, protecting the assembly of true Q-peptides from exoproteolytic attack during expression. Additional peptide sequences are added to provide an initiator methionine and, if appropriate, a single cysteinyl residue as an alternative means of quantification. The QconCAT genes are synthesized by companies (e.g., Entelechon GmbH) who specialize in gene synthesis for the generation of recombinant proteins. The construction and synthesis of the gene to encode a QconCAT is not described here (see **Supplementary Methods**[6]), and the protocol assumes that the experimenter is starting with a QconCAT plasmid.

Expression and purification of QconCAT proteins

Once an expression plasmid encoding the QconCAT protein has been generated (pQconCAT), the plasmid is introduced into an appropriate *E. coli* expression strain. A single transformant is then grown in rich medium and the expression of the QconCAT protein is induced and analyzed by SDS-PAGE; finally, the QconCAT band is subjected to in gel proteolytic digestion and analysis by MS. This simple characterization allows rapid confirmation of adequate QconCAT protein expression and also allows an early evaluation of the signal intensities of the chosen Q-peptides. These preliminary characterizations of the QconCATs are advisable, before

proceeding to labeling with relatively expensive heavy isotope–labeled precursors.

Because QconCAT proteins are concatenations of short peptides (typically 10–20 amino acids) from different proteins, it is difficult to predict their properties. However, it might be anticipated that expression and purification of QconCAT proteins should be straightforward, because these proteins do not have to be functionally folded. Furthermore, they can be purified and used in denatured form, because the only requirement is that they be susceptible to digestion by endoproteolytic enzymes (trypsin, endopeptidase ArgC, endopeptidase LysC, chymotrypsin, endopeptidase Asp-N, staphylococcal peptidase I) or by chemical cleavage. Therefore, expression of a QconCAT does not bring the complexities that are sometimes encountered when a protein must be heterologously expressed and folded in the correct form.

Stable isotope–labeling strategy

Labeled QconCAT proteins are generated by heterologous expression, most simply by growth of the *E. coli* host carrying the pQconCAT in defined medium supplemented with an appropriate stable isotope–labeled precursor. An alternative method would be to use QconCAT expression plasmid DNA to prime a cell-free coupled transcription/translation system, supplemented with the chosen heavy isotope precursor(s). Generation of QconCAT proteins in the latter system may be more compromised in terms of sequence fidelity and also gives lower yields. However, constructs that are toxic to *E. coli* or are extremely unstable (as a result of proteolysis) may be expressed *in vitro* when they cannot be expressed in *E. coli*.

Uniform labeling with ^{13}C or ^{15}N would ensure that every QconCAT peptide is comprehensively labeled; however, each labeled (heavy) peptide will be separated in mass by differing amounts from its unlabeled (light) counterpart, complicating the mass spectroscopic analysis. If Q-peptides are selected that contain at least one instance of a specific amino acid, that amino acid can be incorporated in stable isotope–labeled form. Perhaps the preferred strategy reflects that because most QconCAT proteins are assemblies of tryptic peptides, incorporation of [$^{13}C_6$]lysine and [$^{13}C_6$]arginine would ensure that most Q-peptides would be singly labeled and the mass offset between heavy and light peptides would be a constant 6 Da. Q-peptides that contain internal Arg-Pro or Lys-Pro sequences will of course contain two instances of the labeled amino acid, yielding a mass offset of 12 Da. We advocate the use of ^{13}C-labeled amino acids instead of ^{2}H-labeled amino acids, because of the difference in elution time of peptides containing ^{2}H-labeled amino acids in liquid chromatography and the probable decrease in quantification accuracy caused[13]. Alternatively, unlabeled QconCATs could be labeled *in vitro* using one of the many reagents that have been advocated for comparative proteomics.

Extent of proteolysis

The next stage is to determine the conditions for complete digestion of both the analyte and QconCAT proteins. First, a true internal standard will, apart from the stable isotope labeling, be otherwise exactly the same as the analyte. In the QconCATs, this is clearly not the case, because each surrogate peptide in the concatamer is in a sequence context that is different from the same peptide in the analyte protein. This means that the sequence contexts of the scissile bonds differ, and rates of hydrolysis could therefore be different for the QconCAT and the analyte proteins. Our experience with QconCATs thus far is that they lack higher order structure and are readily and rapidly digested by endopeptidases such as trypsin or endopeptidase LysC. Proteolytic digestion of the analyte proteins, on the other hand, is mostly impeded by the presence of higher order structure, which restricts access to proteases[14–18]. However, provided that the analyte proteins are extensively denatured and that the reaction is allowed sufficient time to proceed to completion, this need not be a serious issue. We have used the following methods to denature protein mixtures in anticipation of proteolysis for quantification.

- Precipitation with an acid such as 10% (wt/vol) trichloroacetic acid. The ether-washed pellet (analyte plus QconCAT) can then be proteolysed back into solution.
- Heat treatment (95 °C for 10 min). Again, precipitated material can be resolubilized by proteolysis.
- Treatment with a chaotrope such as urea or guanidinium hydrochloride, which would then have to be removed before proteolysis.

Acetone precipitation may not be a sufficiently potent denaturation step to enhance proteolysis of many proteins—indeed, it has been used as a precipitation step in protein purification protocols where activity is preserved.

If the analyte proteins contain disulfide bonds, it is also necessary to reduce and alkylate cysteinyl residues. If the QconCAT must contain cysteinyl residues as well, then the analyte-QconCAT should be mixed before reduction/alkylation.

The rapid digestion (usually within 1 min at typical substrate-to-protease ratios of 50:1 for trypsin) of the QconCAT means that it can also be used as a reference in pilot studies designed to assess the degree of proteolysis of the analyte mixture. We would stress that the requirement for complete digestion has rarely been specifically addressed in any quantification studies, and that this step is critical to any such experiment, whether relative or absolute. Even single-protein quantification by chemical synthesis requires complete hydrolysis of the analyte. For identification proteomics, complete digestion may not be required, and the extent of digestion may not therefore be routinely assessed, but we must emphasize that incomplete digestion means incorrect quantification.

Quantification of the QconCAT

Because the absolute quantification strategy involves relative quantification against an accurately determined internal standard, it is worth emphasizing the principles that are involved.

Absolute quantification with a QconCAT is only as good as the quantification of the QconCAT itself. The first step of any QconCAT experiment is thus the determination of the concentration of His-tag purified QconCAT. Several approaches are possible.

- Determination of a QconCAT by protein assay. It is possible to determine the QconCAT concentration by protein assay, and this will provide an acceptable degree of accuracy for many purposes. However, different methods of protein assay yield different results. Furthermore, it is not possible to use the exact protein as a reference, and often a protein such as BSA must be used. Of the different protein assays available, the biuret method, which is insensitive to the nature of the protein being assayed, is preferred. This method can consume a lot of protein (1 mg for a 1-ml assay), but microbiuret methods are available. Moreover, new spectrophotometric

instruments are capable of operating on vanishingly small volumes, and thus sample consumption might be less of a problem.

- Determination of QconCAT by densitometry. It is possible to run the QconCAT on a one-dimensional SDS-PAGE gel and, in parallel lanes, to include known amounts of a protein standard. After staining, densitometry of the lanes allows interpolation of the amount of the QconCAT relative to the standard; thus, its concentration can be assessed. However, this method also suffers from different response factors to staining protocols.
- Determination of QconCAT by Kjeldahl assay. This approach measures total nitrogen by a complex and fairly hazardous process, and although it probably delivers excellent results, its use is not warranted.
- Determination of QconCAT by amino acid analysis. Complete acid hydrolysis (6 M HCl, 110 °C, oxygen-free environment) of a QconCAT would release free amino acids, which can then be assessed by any one of the many methods for amino acid analysis. In nearly every method the amino acid content of the QconCAT will itself be assessed relative to a set of standards, which must obviously be carefully prepared. However, acid hydrolysis affects many amino acids: valine and isoleucine bonds are less easily hydrolyzed, threonine and serine are partially destroyed, methionine can be oxidized during acid hydrolysis, asparagine and glutamine are both converted to the respective acidic residue, and tryptophan is completely destroyed.
- Assay based on specific amino acids in the QconCAT. There is a series of assays that could be used to determine the concentration of specific amino acids in the intact QconCAT. For example, a reagent reactive to primary amino groups (lysine and the α-amino group of the N terminus) such as fluorescamine or *o*-phthaldialdehyde-*N*-acetylcysteine could be used. The advantage of this approach is that fluorescence-based assays are sensitive, there will usually be several reaction sites in a QconCAT and the fluorescence yield of all primary amines is very similar—meaning that a simple amine reagent can be used as a standard.

Colorimetric assays for specific amino acids in intact proteins are not common. We have previously included a single cysteine residue in a QconCAT to facilitate determination by Ellman's reagent (DTNB, dithio *bis*-2-nitrobenzoic acid). However, the extinction coefficient of DTNB is 13,600 M^{-1} cm^{-1} at 412 nm, thus, it cannot be used convincingly with a QconCAT concentration of <0.01 mM ($A_{412} = 0.136$). A typical QconCAT (molecular weight ~40 kDa) would, at 0.01 mM, be 0.4 mg ml^{-1}, and this method therefore consumes a lot of QconCAT.

- Assay based on a common synthetic peptide. An alternative strategy for quantification of QconCATs would be the purchase of an accurately quantified synthetic peptide (in a less expensive unlabeled form, because we are interested in quantifying the labeled QconCAT). Again, we stress that this quantification is only as good as the quantification of the synthetic peptide. However, it would be possible to design a common peptide into every QconCAT and to quantify every protein relative to the same synthetic peptide using MS. For example, the His-tag peptide in the construct described here gives a good mass-spectrometric signal and could be made common to every artificial protein. A common quantification protocol could then be widely disseminated and diminish problems of different peptide or protein quantification.

The quantification of any protein is not without problems, and QconCAT technology brings the same challenges. However, it is essential to recognize the importance of quantification of the QconCAT in the correct analytical context. The QconCAT method has, for any one peptide, an experimental coefficient of variance of ~2%, and this therefore sets the limit on the precision that is available for quantification. However, in many experiments the biological variance will be substantially greater than any analytical step, and again, this defines the quality of the protein quantification that is required. It might be argued that the goal of any analytical method is to make the analytical variance sufficiently small relative to biological variance that the emphasis should always be on biological variation, and that biological replicates are chosen over technical replicates.

We would venture to suggest that for many experiments a simple colorimetric protein assay is acceptable (either dye binding or biuret), but that for extremely high-precision work, determination of several amino acids or total nitrogen might be best. For many experiments such stringent quantification methods might be considered to be unnecessary. Finally, once a QconCAT has been quantified, there is merit in then determining the A_{280} (1%) of the protein, so that future quantification can be based on simple, nondestructive spectrophotometric measurement.

Comparison between absolute quantification by QconCATs or synthetic peptides

So far a direct comparison between peptides prepared by chemical synthesis and by QconCAT has not been undertaken, so we can only discuss the probable pros and cons of the two methods. First, if only a small number of proteins (e.g., ≤10 proteins) is to be quantified, the synthetic peptide approach is probably more appropriate and economical. For the absolute quantification of larger numbers of proteins, because each constituent peptide of a QconCAT is assayed simultaneously, the method is far superior to the quantification of several synthetic peptides in independent experiments, although this feature increases the demand for good QconCAT protein quantification. If the proteins in the group under study differ in concentration by orders of magnitude, it is perhaps easier to add varying amounts of internal standard if they are available as synthetic peptides. However, when designing several QconCATs to cover the range of proteins of interest, it is sensible to group together those proteins of similar abundance within the same QconCAT and adjust the amount of each QconCAT added accordingly. One advantage of a QconCAT is that once the gene has been constructed, a range of labeling methods is available (e.g., Arg, Lys labeling, or complete labeling with ^{15}N), whereas with synthetic peptides the choice of label must be made in advance of the one-shot synthesis.

Laboratories specializing in proteomics and MS alone

The protocol described below is a step-by-step walk through the whole process from the design of the QconCAT to its use in the absolute quantification of complex protein mixtures. This necessarily covers a range of different techniques, which may not be within the capabilities of a typical proteomics laboratory. For

those laboratories at which work with genetically manipulated organisms is not permitted or the required expertise is not present, purified, labeled QconCATs can be obtained commercially (http://www.qconcat.com). Once the group of proteins that you wish to quantify has been decided and the set of surrogate peptides elected, a QconCAT gene can be designed, synthesized, expressed in labeled form, characterized and delivered to the client for immediate use in absolute quantification experiments (http://www.qconcat.com).

Sources of proteinaceous samples

Potential analyte sources include human, animal, plant (grasses, shrubs, trees, algae, food), microorganisms, body fluids (blood, serum, CSF, bronchial lavage, semen, vaginal secretion, tears, saliva, sweat, sputum and urine), as well as feces, tissues, cells, cell lines, hair, food, soil, water (as from rivers or the ocean) and sewage. In this exemplar protocol we have included the quantification of 20 proteins in the soluble fraction of chicken skeletal muscle during development immediately after hatching.

MATERIALS

REAGENTS

- $^{15}NH_4Cl$ (99% atom percent excess) (CIL Inc.)
- L-arginine hydrochloride (U-$^{13}C_6$, 98%; CIL Inc.)
- L-lysine hydrochloride (U-$^{13}C_6$, 98%; CIL Inc.)
- 20 mM phosphate
- 20 mM, 500 mM imidazole
- 6 M guanidinium chloride
- Protease inhibitors (Complete Protease Inhibitors; Roche)
- Coomassie Plus Protein Assay (Pierce)
- *E. coli* BL21(λDE3) (Stratagene; Promega; Genlanatis)
- Plasmid vector pET21a (Novagen, Merck)
- Luria broth (LB; Merck)
- Luria agar (LA)
- HCl
- NaOH
- Ampicillin sodium salt (Sigma cat. no. A9518-5G)
- Isopropyl β-D-thiogalactopyranoside (IPTG) (Sigma cat. no. 16758-1G)
- Glycerol
- Bromophenol blue
- Ammonium bicarbonate solution
- Trypsin, sequencing grade (Roche)
- Formic acid
- Acetonitrile (ACN)
- Iodoacetamide
- Trifluoroacetic acid (TFA)
- α-cyano- 4-hydroxycinnamic acid
- Bugbuster Protein Extraction Reagent (Novagen; EMD Biosciences)
- Lysozyme (from chicken egg white; Sigma cat. no. L-7651)
- HisTrap HP (GE Healthcare UK, Ltd.)
- StrataClean Resin (Stratagene)
- 7 M ammonium hydroxide
- Sodium phosphate buffer, pH 7.0
- $Na_2HPO_4 \cdot 7H_2O$
- KH_2PO_4
- NH_4Cl
- $MgSO_4$
- Glucose
- Thiamine
- $CaCl_2$
- Amino acids

REAGENT SETUP

Bacterial strains *E. coli* BL21(λDE3). BL21 has the following genotype: F^--, *omp*T, *hsd*S_B ($r_B^- m_B^-$) *gal*, *dcm*. Frozen competent cells of BL21(λDE3) can be obtained from many suppliers, including Stratagene, Promega and Genlanatis. BL21(λDE3) has a recombinant λ phage carrying the T7 RNA polymerase, stably integrated into the chromosome. ! **CAUTION** Use good microbiological practice in manipulating and disposing of this *E. coli* laboratory strain. BL21(λDE3) does not require additional amino acids to grow in minimal medium (MM). If you are using a different host strain, check the genotype and add amino acids as appropriate for growth. If you are unsure about the requirements, the supplier of the strain (or transformation-competent cells) should be able to provide you with a recipe for defined growth medium for a particular strain.

Plasmids The recombinant plasmids encoding the QconCAT proteins are called pQconCATs, and these carry QconCAT genes cloned into the *Nde*1-*Hin*dIII restriction sites of pET21a, although other restriction sites are available.

Tris-EDTA buffer Prepare 100 ml 10 mM Tris HCl, 1 mM EDTA, pH 8.0 (TE). Sterilize by autoclaving for 15 min at 121 °C, and store at room temperature (RT, 20–25 °C).

LB Dissolve 25 g of LB powder in 1 liter of distilled water. The pH should be 7.0 ± 0.2 at 25 °C; if it is not, adjust with HCl or NaOH as appropriate. Sterilize by autoclaving for 15 min at 121 °C.

Ampicillin sodium salt Prepare a 20 mg ml^{-1} solution in sterile distilled water, freeze in 1-ml aliquots and store at –20 °C for several months, or store at 4 °C for no longer than 1 week. For the growth of BL21(λDE3)-pQconCAT strains use a final concentration of 50 μg ml^{-1} to maintain selection for the recombinant plasmid.

IPTG Prepare a small volume of a 1 M solution in sterile distilled water.

2× SDS-PAGE sample buffer 0.5M Tris HCl, pH 6.8 (2.5 ml, for a final concentration of 0.125 M); 10% (wt/vol) SDS (4.0 ml, for a final concentration of 4% (wt/vol)); glycerol (2.0 ml (density 1.26 g ml^{-1}), for a final concentration of 20% (vol/vol)); DTT (0.31 g, for a final concentration of 0.2 M); bromophenol blue 0.5% (wt/vol), (40 μl). Add double-distilled water to the 2× SDS-PAGE sample buffer to a final volume of 10 ml, divide into 1-ml aliquots and store at –20 °C. ! **CAUTION** DTT solid is harmful by inhalation, in contact with skin and if swallowed. It is irritating to eyes, skin and the respiratory system.

Tryptic digestion solution 25 mM and 50 mM ammonium bicarbonate solution (no pH adjustment needed); Trypsin, sequencing grade. Prepare 10 ml of 10% (vol/vol) formic acid in distilled water for use in time course experiments. ! **CAUTION** Formic acid is harmful by inhalation; it causes severe burns.

Solution for in-gel digestion and reduction and alkylation of cysteine-containing proteins Prepare 20 ml of a 2:1 mixture of 25 mM ammonium bicarbonate–ACN.

DTT Prepare 10 ml of 10 mM DTT in 25 mM ammonium bicarbonate for in-gel digestion. Prepare 10 ml of 100 mM DTT stock solution in 25 mM ammonium bicarbonate for addition to analyte-QconCAT mixtures to a final concentration of 10 mM.

Iodoacetamide Prepare 10 ml of 55 mM iodoacetamide in 25 mM ammonium bicarbonate (cover the bottle with foil to exclude light) for in-gel digestion. ! **CAUTION** Iodoacetamide is toxic if swallowed. It may cause sensitivity by inhalation and skin contact. Do not breathe the dust. Wear suitable protective clothing and gloves.

Matrix for matrix-assisted laser desorption-ionization–time of flight (MALDI-ToF) MS Prepare 50 ml of 50% (vol/vol) acetonitrile–0.1% (vol/vol) TFA; store at RT. Prepare a fresh saturated solution of ~10 mg of α-cyano-4-hydroxycinnamic acid in 1 ml 50% (vol/vol) ACN–0.1% (vol/vol) TFA. ! **CAUTION** ACN is highly flammable and is harmful by inhalation, contact with skin and if swallowed. It is also irritating to the eyes. TFA is harmful by inhalation and causes severe burns. α-cyano-4-hydroxycinnamic acid is irritating to the eyes, respiratory system and skin.

Solution for electrospray ionization (ESI) MS Prepare 500 ml 50% (vol/vol) ACN–1% (vol/vol) formic acid. ! **CAUTION** Formic acid is harmful by inhalation; it causes severe burns.

Guanidinium chloride binding buffer 20 mM phosphate, pH 7.4; 0.5 M NaCl; 20 mM imidazole; 6 M guanidinium chloride.

Elution buffer 20 mM phosphate, pH 7.4; 500 mM imidazole; 6 M guanidinium chloride.

Dialysis buffer 10mM ammonium bicarbonate, pH 8.5–1 mM DTT; the pH is adjusted with a solution of 7 M ammonium hydroxide.

Protein assay Coomassie Plus Protein Assay (Pierce); BSA for the protein calibration curve.

Preparation of chicken skeletal muscle soluble fraction Prepare 500 ml 20 mM sodium phosphate buffer, pH 7.0. Add one tablet Complete Protease Inhibitors to 20 ml of 20 mM sodium phosphate buffer, pH 7.0.

Bacterial growth media The minimal medium (MM) is a chemically defined mixture into which the stable isotope–labeled precursor to be used as a label can be added. For labeling with stable isotope–labeled amino acids, the following sterile stock solutions are made.

5× M9 salts To 90 ml of distilled water add the following, sequentially, allowing each salt to dissolve before adding the next: 6.4 g $Na_2HPO_4 \cdot 7H_2O$; 1.5 g KH_2PO_4; 0.25 g NaCl; 0.5 g NH_4Cl. Make up to 100 ml with distilled water, and autoclave at 121 °C for 15 min. Also prepare: 1 M $MgSO_4$ (autoclave); 20% (wt/vol) glucose (filter sterilize); 0.5% (wt/vol) thiamine (filter sterilize); 0.1 M $CaCl_2$ (autoclave).

Amino acids Prepare a 10 mg ml^{-1} mixture including all the amino acids except those to be used for labeling. Vortex-mix the suspension vigorously (not all amino acids will dissolve fully), and immediately dispense in 2-ml aliquots and store at –20 °C (stable for several months). If you intend to use a variety of labeled amino acids in future experiments, omit all of these from this mixture so that you only have to make it once.

MM with amino acids Prepare MM by mixing the following volumes of the above stock solutions: 5× M9 salts (40 ml); 1 M $MgSO_4$ (0.2 ml); 20% (wt/vol) glucose (2 ml); 0.1 M $CaCl_2$ (0.2 ml); 0.5% (wt/vol) thiamine (20 µl); amino acid mix lacking chosen labeled amino acid(s) (2.0 ml); sterile water to make 200 ml.

Labeled amino acid–supplemented MM The medium used for growth should contain all 20 amino acids. The amino acid(s) chosen for labeling is weighed and added as a solid to the above MM plus amino acids. If you have left out other amino acids you should add these individually in unlabeled form.

Therefore, add 10 mg of either labeled or unlabeled amino acid per 100 ml of the MM with amino acids, filter sterilize and store at 4 °C for as long as 48 h.

MM for labeling with $^{15}NH_4Cl$ For $^{15}NH_4Cl$ as the labeled precursor, it is essential that you choose an *E. coli* strain (like BL21(λDE3)) that does not require amino acid supplements to grow. Check the genotype of your host strain, and select a different host strain if appropriate. *E. coli* will grow more slowly in the absence of amino acids, so times required for growth and induction will be extended ~50%.

^{15}N-labeled 5× M9 salts Na_2HPO_4 (3.39 g; anhydrous M_r 141.96); KH_2PO_4 (1.5 g); NaCl (0.25 g); $^{15}NH_4Cl$ (0.5 g); distilled water to make 100 ml. Autoclave in 20-ml aliquots.

MM containing ^{15}N Mix the solutions as follows and store the medium at 4 °C for as long as 48 h: ^{15}N-labeled 5× M9 salts (20 ml); 1 M $MgSO_4$ (0.1 ml); 20% (wt/vol) glucose (1 ml); 0.1 M $CaCl_2$ (0.1 ml). Add sterile distilled water to make 100 ml. Note: thiamine is not added, because this is a possible source of unlabeled nitrogen (^{14}N).

EQUIPMENT

- Standard spectrophotometer (for absorbance readings in the visible range, including 600 nm)
- Peristaltic pump
- MALDI-ToF MS instrumentation (e.g., Waters-Micromass Q-ToF micro mass spectrometer)
- ESI MS instrumentation
- HisTrap column (GE Healthcare)
- Multiskan plate reader (Thermo Electron)
- Ascent software
- Syringe pump (Harvard)
- MaxEnt 1 module of the MassLynx software
- Jouan centrifugal evaporator (Thermo Electron)

PROCEDURE

Transformation of expression host *E. coli* BL21(λDE3) with a QconCAT plasmid

1| Dissolve the QconCAT plasmid (pQconCAT), which is typically delivered as a lyophilized powder, in ~100 µl Tris EDTA buffer and store it at –20 °C.

2| Prepare a 1 ng/µl solution of the pQconCAT plasmid in Tris EDTA buffer.

3| Prepare BL21(λDE3) cells competent for transformation (using basic transformation methods; see ref. 19), or purchase frozen competent cells and introduce the pQconCAT plasmid by transformation using standard procedures (as described in the notes that accompany frozen competent cells).

! CAUTION Use good microbiological practice.

4| Select colonies on LA plates with 50 µg ml^{-1} ampicillin by growing overnight at 37 °C.

5| Take a single colony and streak onto a fresh LA plate with 50 µg ml^{-1} ampicillin. Store the freshly streaked plate at 4 °C for as long as a month.

Preparation of frozen cultures of BL21(λDE3)-pQconCAT

6| Inoculate 10 ml of LB–50 µg ml^{-1} ampicillin, with a single colony of BL21(λDE3)-pQconCAT, and grow overnight at 37 °C with shaking.

7| Add sterile glycerol to 30%, mix well, aliquot 1 ml into sterile microcentrifuge tubes and store at –70 °C.

■ **PAUSE POINT** Frozen cells may be kept in this way for several years.

8| To prepare a fresh plate, place an aliquot of frozen culture on ice and thaw a small amount of the top; use 20 µl to streak a fresh LA plate containing ampicillin, return the frozen culture to the freezer and incubate the plate overnight at 37 °C.

■ **PAUSE POINT** Plates can be kept for as long as a month under refrigeration, but it is best to inoculate cultures using a colony from a fresh plate.

Induction of expression of QconCAT proteins in BL21DE3-pQconCAT and confirmation of concatenated tryptic peptides by SDS-PAGE and MALDI-ToF MS

9| Using a single colony of BL21(λDE3)-pQconCAT, inoculate 10 ml LB containing ampicillin (50 µg ml^{-1}) and incubate overnight at 37 °C with shaking.

10| Transfer 500 µl of the overnight culture to 50 ml of prewarmed (to 37 °C) fresh LB medium (a 1:100 dilution) containing ampicillin (50 µg ml^{-1}), and incubate the culture with shaking. Remove 1-ml samples at hourly intervals, and determine the absorbance at 600 nm using a spectrophotometer.

11| When an A_{600} of 0.6–0.8 is reached (usually ~2.5 h) add 50 µl of 1 M IPTG (final concentration of ~1 mM) to induce expression of the QconCAT protein.

12| Remove 1-ml samples at time 0 and then every 1–2 h (up to 6 h), measure the A_{600} immediately, transfer the equivalent of 0.6–0.8 A_{600} of cells to labeled microcentrifuge tubes and hold on ice.

13| Centrifuge the samples at 10,000 r.p.m. in a microfuge at 4 °C for 10 min. Remove and discard all the supernatant using a micropipette, and suspend the pellet in 100 µl of distilled H_2O by vigorous vortexing.
■ **PAUSE POINT** If the intention is to perform the SDS-PAGE analysis at a future date, place at –20 °C at this stage.

14| Transfer the remainder of the culture to a preweighed 50-ml centrifuge tube, and centrifuge at 1,450 *g* for 10 min at 4 °C.

15| Decant the supernatant, and weigh the tube again to determine the wet weight of the cell pellet. Freeze at –20 °C until required for purification of QconCAT after confirmation by SDS-PAGE of sufficient levels of induction.

16| Add 100 µl double-strength SDS sample buffer to samples from Step 13, mix well and place in a boiling water bath for 4 min. Analyze 20 µl of each sample by SDS-PAGE mini-gel, carry out electrophoresis, and stain and destain the gel (e.g., with Coomassie blue).

17| The QconCAT protein should appear as a clearly visible band that is either not present or much fainter in the time 0 sample. Check that the molecular weight is close to that calculated for the QconCAT, and if all is as expected proceed to the next stage.

Analysis of QconCAT by MALDI-ToF MS after in-gel digestion of SDS-PAGE gel band with trypsin

18| Using a Pasteur pipette, cut a plug of gel from the band corresponding to the QconCAT protein, transfer the plug to a 1.5-ml microcentrifuge tube, add 25 µl of 25 mM ammonium bicarbonate and incubate at 37 °C for 15 min. Discard any liquid.

19| Add 25 µl of 25 mM ammonium bicarbonate–ACN (2:1), and incubate at 37 °C for 15 min. Discard any liquid.
! **CAUTION** ACN is flammable and toxic.

20| Add 25 µl of 25 mM ammonium bicarbonate, and incubate at 37 °C for 15 min. Discard any liquid. Repeat Steps 19 and 20 using alternate washes until the plug is fully destained. Note: If cysteine residues are present, alkylate the QconCAT protein by performing these three steps: (i) Add 25 µl of 10 mM DTT in 25 mM ammonium bicarbonate, and incubate at 56 °C for 60 min. Discard any liquid. (ii) Add 25 µl 55 mM iodoacetamide in 25 mM ammonium bicarbonate, and incubate in the dark at 37 °C for 45 min. Discard any liquid. (iii) Repeat alternate washes (Steps 19 and 20), finishing with a wash in 25 mM ammonium bicarbonate–can (2:1). These three steps can be omitted if cysteine residues are not present in the QconCAT protein.

21| Add 10 µl of 25 mM ammonium bicarbonate containing 0.1 mg ml^{-1} trypsin giving a final concentration of 12.5 ng $µl^{-1}$.

22| Hold samples on ice for 30 min before incubation at 37 °C overnight. Ensure that the gel plug is covered with liquid; if it is not, add 25 mM ammonium bicarbonate until covered.

23| Spot 1 µl of digest on a MALDI target, and overlay with 1 µl of matrix.

24| Acquire spectra using a MALDI-ToF mass spectrometer. Acquire mass spectra over the range 800—3,500 *m/z*.

Evaluation of MALDI ToF results

25| Analyze spectra. Peaks due to each peptide in the QconCAT with masses larger than 800 *m/z* should be clearly visible. Confirm that all the expected peaks are present, and confirm that digestion has gone to completion—meaning there are no peaks due to missed cleavages.

Purification and analysis of QconCAT proteins

26| Thaw the induced cells from Step 15. The QconCAT is generally present in inclusion bodies, but it is important to confirm this. Inclusion bodies are first recovered by breaking cells using BugBuster Protein Extraction Reagent (Novagen). The following method is adapted from the Novagen protocol.

27| For ≤1 g of wet cell pellet add 2.5 ml of BugBuster (BB) and, to ensure good resuspension, place cells on a rocker platform at RT for 15 min. Remove 20 µl for analysis by SDS-PAGE (total fraction).

28| Centrifuge the cells at 16,000*g* for 20 min at (4 °C). Set the braking speed low to give a gentle rotor deceleration.

29| Collect the supernatant by pipetting carefully to a fresh tube, transfer 20 µl to a labeled microcentrifuge tube for analysis (soluble fraction) and store the remainder at –20 °C. Resuspend the pellet in 2.5 ml of BB by drawing the cells up into a Pasteur pipette and by gentle vortexing.
▲ **CRITICAL STEP** Thorough resuspension is critical to obtaining a preparation of high purity.

30| Add 50 µl of lysozyme (10 mg ml^{-1} in BB), mix gently by vortexing and incubate at RT for 5 min.

31| Add 15 ml of 1:10 dilution of BB in distilled water, and mix by vortexing for 1 min.

32| Centrifuge at 15,000*g* for 15 min at 4 °C.
▲ **CRITICAL STEP** This is a much higher speed than recommended in the Novagen protocol, but it generates a firmer pellet, which facilitates handling. Carefully decant the supernatant and discard.

33| Resuspend the pellet of inclusion bodies in 20 ml of 1:10 BB, and mix by vortexing (low speed), then centrifuge at 15,000*g* for 15 min at 4 °C. Discard the supernatant.

34| Repeat Step 33 twice more, finishing by centrifuging at 16,000*g* for 15 min. Store the pellet of inclusion bodies at –20 °C in the 50-ml centrifuge tubes.
■ **PAUSE POINT** Inclusion body pellets can be stored at –20 °C indefinitely.

35| Analyze the total fraction (Step 27) and soluble fraction (Step 29) by SDS-PAGE.

36| Add 20 µl of 2× SDS-PAGE sample buffer to each sample, heat in a boiling-water bath for 4 min and load 20 µl onto a gel.

37| From the Coomassie blue–stained gel determine the location of the QconCAT and, if largely absent from the soluble fraction (compared with the total fraction), proceed with the purification of the QconCAT protein from the inclusion bodies.

38| Resuspend the inclusion bodies in 20 ml of binding buffer at room temperature. (If the QconCAT is in the soluble fraction, add this fraction to 20 ml of binding buffer.)

39| Centrifuge at 8,000 r.p.m. (~5,000*g*) for 5 min at room temperature, and collect the supernatant in a fresh tube. Place 20 µl of the supernatant into a microcentrifuge tube; this sample is the starting material (SM) for SDS-PAGE analysis.

40| Apply the remainder of the supernatant, with the aid of a peristaltic pump set at a flow rate of 0.25 ml min^{-1}, to a 1-ml HisTrap HP column equilibrated in the same buffer.

41| Collect the unbound sample in a 25-ml Universal tube. Place 200 µl of the eluate into a microcentrifuge tube and hold at room temperature; this sample is the unbound material (UM).

42| Wash the column with 10 ml of 20 mM phosphate, pH 7.4, 20 mM imidazole, 0.5 M NaCl, 6 M guanidinium chloride using an increased flow rate of 0.5 ml min^{-1}, and collect the wash. Transfer 200 µl of the wash into a microcentrifuge tube, and hold at room temperature; this sample is the wash (W).

43| Elute the bound protein with 5 ml 20 mM phosphate, pH 7.4, 500 mM imidazole, 0.5 M NaCl, 6 M guanidinium chloride at a flow rate of 0.25 ml min^{-1}, and collect five 1-ml fractions. Place 20 µl of each fraction into a microcentrifuge tube, and hold at room temperature; this sample is the eluted bound material (EBM). Hold the eluted fractions on ice.
■ **PAUSE POINT** Samples can be stored at –20 °C.

44| To identify the fractions containing the QconCAT, prepare each of the 20-µl fractions for SDS-PAGE analysis. Collect the samples SM, UM, W and EBM (Steps 39–43), and add 10 µl of a StrataClean Resin (Stratagene) bead suspension to each tube (use the same volume of beads for the 200-µl W and UM samples).
▲ **CRITICAL STEP** The guanidinium chloride interferes with the SDS-PAGE and so must be removed.

45| Vortex each sample for 1 min, centrifuge for 2 min at 230 *g* at RT, and remove and discard the supernatant.

46| Resuspend pellets in 1 ml of distilled H_2O, vortex briefly to mix and centrifuge at 2,000 r.p.m. for 2 min. Discard the supernatant.

47| Add 10 µl of 2× SDS sample buffer to the beads, boil for 4 min and then load both sample and beads onto the gel.

48| Stain (with Coomassie blue) and destain the gel.

49| Pool the eluant fractions containing pure QconCAT protein, and remove denaturant by dialyzing against 100 volumes of 10 mM ammonium bicarbonate, pH 8.5, 1 mM DTT, at 4 °C, for 2 h. Repeat twice more with fresh buffer.

50| Determine the protein concentration. For this one can use Coomassie Plus Protein Assay (Pierce) and BSA as the standard, and a LabSystems Multiskan plate reader using Ascent software.

51| Prepare 1:50 and 1:100 dilutions of the purified QconCAT protein with a protein assay ranging from 0 to 50 µg. The concentration of the QconCAT protein is likely to be in the range of 0.2–4 mg ml^{-1}.

Analysis of QconCAT proteins using ESI MS (Q-ToF)

52| Dilute purified and dialyzed QconCAT protein directly into 50% (vol/vol) ACN–1% (vol/vol) formic acid to a final concentration of 60–100 fmol μl^{-1}.

53| Directly infuse the sample into the source at a flow rate of 0.5 µl min^{-1} using a syringe pump (Harvard). The capillary voltage may be set between 1,600 and 2,100 V and data acquired over a mass range of 400–1,500 *m/z* with a scan/interscan speed of 2.4/0.1 s.

54| Combine the scans and subtract the spectra before transformation using the MaxEnt 1 module of the MassLynx software. This permits deconvolution of the spectrum into an intact mass measurement within 2 Da of the predicted mass.

Growth of BL21(λDE3)-pQconCAT in minimal medium and labeling with stable isotopes

55| Using sterile technique, pipette 10 ml of MM without amino acids into a sterile 50-ml conical flask, and add 25 µl of 20 mg ml^{-1} ampicillin in sterile distilled water.

56| Inoculate with a single colony of BL21(λDE3)-pQconCAT, and incubate overnight at 37 °C with shaking.

57| Using sterile technique, transfer 50 ml MM plus amino acids and 50 ml MM plus labeled amino acids to two sterile 250-ml conical flasks. Add 125 µl of 20 mg ml^{-1} ampicillin in sterile distilled water to each, and warm to 37 °C.
▲ **CRITICAL STEP** Amino acids are included in this step to increase the growth rate in MM. This shortens the time before induction can commence and prevents the recycling of labeled amino acids.

58| Determine the A_{600} of the overnight culture and, using sterile technique, transfer enough culture to each flask to give a starting A_{600} of 0.06–0.1.

59| Monitor the A_{600} of the cultures at hourly intervals until an A_{600} of 0.6–0.8 is reached (this should take 3.5–4.5 h depending on the starting A_{600}, the *E. coli* strain and the growth medium used).

60| Add 50 µL 1 M IPTG to each flask to induce protein expression. Remove a 1-ml sample at 0 and 5 h after induction, measure the A_{600} and prepare for gel analysis exactly as described in Steps 13, 16 and 17.

61| Centrifuge the remainder of the 5-h induced cells in preweighed centrifuge tubes, at 1,450 *g* for 15 min at 4 °C, decant and discard the supernatant, record the weight of the cell pellets and store the cell pellets in the centrifuge tubes at –20 °C.

62| Process exactly as described in Steps 26–53, for purification and analysis.

Comparison of labeled and unlabeled QconCAT tryptic peptide profiles

63| Mix unlabeled (L, light) and labeled (H, heavy) QconCAT protein to give a total of 10 µg (e.g., L/H ratios of 10:0, 9:1, 7:3, 5:5, 3:7, 1:9 and 0:10), and add 50 mM ammonium bicarbonate to a volume of 20 µl.

64| Add trypsin in a ratio of trypsin to QconCAT of 1:50, and incubate at 37 °C overnight.

65| Analyze four 1-µl fractions by MALDI-ToF MS as described in Steps 23 and 24.

66| Measure the peak intensities, and confirm that the quantification is as expected from the prepared ratios.

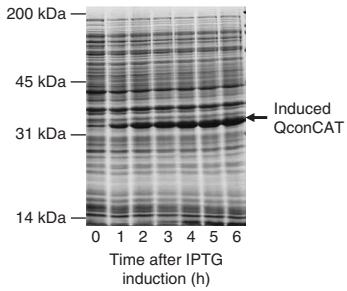

Figure 4 | Induction of QconCAT protein in BL21(λDE3) cells grown in Luria broth. BL21(λDE3) cells were grown to an A_{600} of 0.8, and IPTG added to a final concentration of 1 mM. Cells equivalent to 0.08 A_{600} units were analyzed by SDS-PAGE and Coomassie blue staining.

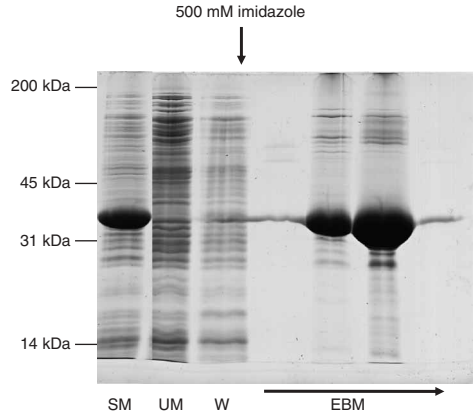

Figure 5 | HisTrap purification of QconCAT. SM, starting material (solubilized inclusion bodies), UB, unbound material (flowthrough); W, wash fraction; EBM, eluted bound material (sequential 1-ml fractions). The middle two fractions containing the eluted QconCAT protein are pooled for use.

Determination of conditions required for complete digestion of the intended analyte and confirming the digestion of analyte and QconCAT protein mixtures

67| Label 12 0.5-ml microcentrifuge tubes, pipette 10 µl of 10% (vol/vol) formic acid into each tube and hold on ice.

68| Dilute ~20 µg of analyte protein 1:10 with 50 mM ammonium bicarbonate solution. This typically gives a volume of 200–300 µl.

69| Add trypsin in a ratio of trypsin to analyte of 1:50, mix and immediately transfer 12 µl into the first tube previously prepared in Step 67.

70| Incubate the analyte-trypsin mixture at 37 °C, and take 12-µl samples over a 24-h period. Suitable time intervals are 0, 1, 2, 5, 10, 30, 60, 90, 120, 240, 480 and 1,440 min. Store samples at –20 °C.

71| Thaw the samples, and spot 1 µl of each sample onto two positions on a MALDI target; that is, perform in duplicate. Analyze by MALDI-ToF MS exactly as described in Steps 23 and 24. If digestion is complete, the peptides of the proteins under study (i.e., chosen for inclusion in the QconCAT) should be readily seen. At early time points partial digestion products will be seen for some proteins; this analysis allows the monitoring of these more slowly digested portions of the protein, which should be fully digested within 24 h if conditions are appropriately optimized.

72| Dry down the remaining sample (10 µl) for 2 h by heating under vacuum (Jouan centrifugal evaporator (Thermo Electron)), and resuspend the residue directly in 10 µl 1× SDS sample buffer.

73| Heat samples in a boiling-water bath for 4 min, and load the entire sample onto an 12.5 % SDS-PAGE gel. After electrophoresis stain the gel with Coomassie blue; no protein should be visible on the gel after 24 h of digestion.

74| If digestion of the analyte is incomplete, increase the protein-to-trypsin ratio from 50:1 to 20:1; also, an organic solvent such as ACN can be added (e.g., to 10%). Complete digestion of the analyte proteins is measured by the absence of all bands in an SDS-PAGE gel.

75| Once you have optimized digestion conditions for the analyte, repeat the digestion but add the labeled QconCAT protein in a 1:10 ratio with the analyte.

76| Analyze 1 µl by MALDI-ToF MS, and confirm the complete digestion of the QconCAT when present in a mixture with the analyte.

Quantification of QconCAT proteins for use in absolute quantification

77| Quantify the labeled QconCAT by a method of suitable accuracy for the intended experiment (see section on "Quantification of the QconCAT" in the EXPERIMENTAL DESIGN section). So far we have used a dye-binding assay to determine the protein concentration of both the QconCAT and the analyte proteins. This method gives acceptable accuracy and does not require the use of significant quantities of the

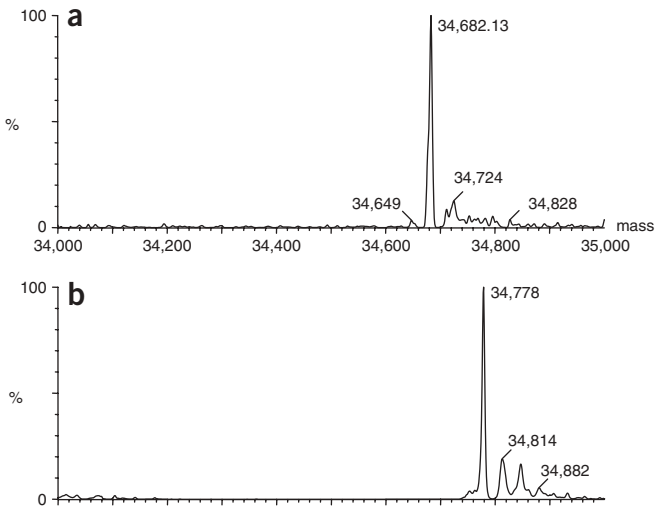

Figure 6 | ESI-MS to measure the intact mass of QconCATs. Purified and dialyzed, labeled and unlabeled, QconCAT protein was diluted into 50% (vol/vol) ACN–1% (vol/vol) formic acid to a final concentration of 60–100 fmol μl^{-1} and infused directly into the source of a Q-ToF MS at a flow rate of 0.5 µl min^{-1} using a syringe pump (Harvard). The capillary voltage was set between 1,600 and 2,100 V and data acquired over a mass range of 400–1,500 *m/z* with a scan/interscan speed of 2.4/0.1 s. Scans were combined and the spectra subtracted before transformation using the MaxEnt 1 module of the MassLynx software, which allows deconvolution of the spectrum into an intact mass measurement within 2 Da of the predicted mass. (**a**) Unlabeled QconCAT (predicted mass: 34,684); (**b**) QconCAT labeled with [$^{13}C_6$]lysine (predicted mass 34,780).

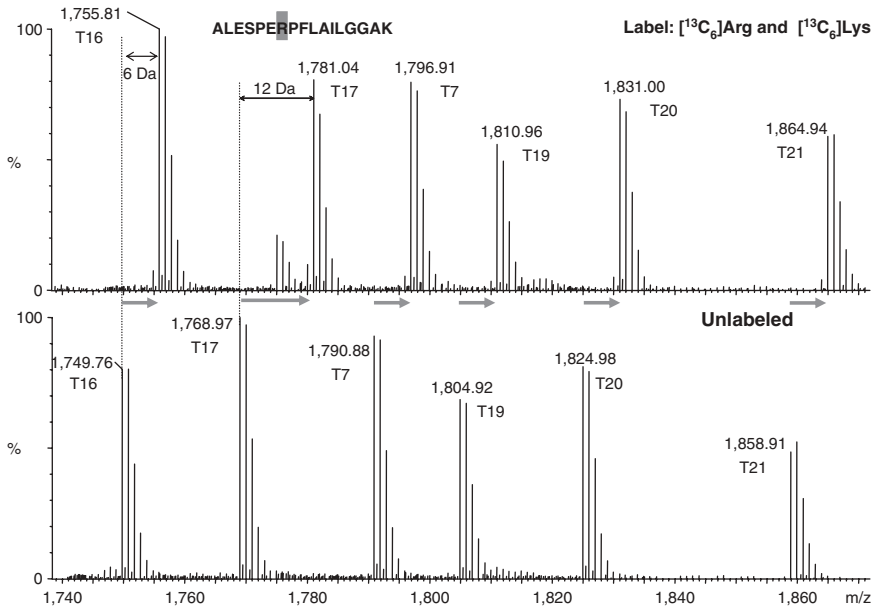

Figure 7 | Incorporation of label into peptides derived from QconCATS. QconCAT proteins labeled with [$^{13}C_6$]lysine and [$^{13}C_6$]arginine (upper spectrum) and unlabeled (lower spectrum) were digested with trypsin and each digestion analyzed directly by MALDI-ToF MS. Only a portion of the spectrum is shown so that the mass shift of the labeled peptides can be seen. Peptide T17 differs from its unlabeled counterpart by 12 Da as a result of the uncleaved arginine-proline sequence in the peptide; all the other peptides show a shift to 6 Da heavier in the labeled peptide.

purified QconCAT. We are currently exploring the use of other methods that are more accurate but utilize small amounts of the purified QconCAT.

Use of a QconCAT protein in absolute quantification: quantification of proteins in chicken skeletal muscle

78| Take 100 mg of chicken breast tissue from chickens of different ages (for example), and homogenize in 1 ml 20 mM sodium phosphate buffer, pH 7.0, containing protease inhibitors.

79| Centrifuge at 15,000*g* for 45 min at 4 °C. Collect the soluble fraction, and store at –20 °C in 100-µl aliquots.

80| Determine the total protein concentration in each sample of chicken skeletal muscle soluble fraction (CSM) and the concentration of the labeled purified QconCAT protein using a Coomassie Plus Protein Assay.

81| Dilute samples 1:10 with 50 mM ammonium bicarbonate, and mix the CSM and labeled QconCAT in known concentrations. For initial pilot studies a ratio of 1:10 (QconCAT/CSM) is suitable, and total protein in the range of 20 µg.

82| Add ACN to 10% and trypsin in a 20:1 (protein to enzyme) ratio, and incubate at 37 °C for 24 h.

83| Spot four 1-µl (four replicates) of each digest onto a MALDI target, and overlay with 1 µl of matrix.

84| Acquire spectra, for example using a MALDI-ToF mass spectrometer, over the range 800—3,500 *m/z*, or use a lower *m/z* maximum if peptides under study fall well below this maximum mass.

85| Record the intensity of the monoisotopic peak for each analyte peptide and the corresponding QconCAT peptide, and convert these values into nanomoles of protein per gram of tissue using the known concentration of the QconCAT protein and CSM. Determine the standard error for each ratio.

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 1**.

TABLE 1 | Troubleshooting the QconCAT design, expression and labeling steps

PROBLEM	POSSIBLE REASON	SOLUTION
Expression of the QconCAT is very poor	Multiple potential reasons see below:	If expression is low but just detectable, sufficient QconCAT protein can be obtained by processing larger volumes of induced cultures.
	Gene sequence incorrect	Check the gene sequence. Look at transcript and predicted translation product. If incorrect, the gene must be reconstructed.
	Protein instability	Vary the induction conditions (times, temperature) to try to improve expression. Try a different host *E. coli* strain.
	Problems with mRNA secondary structure occluding expression	Monitor RNA production upon induction to confirm that transcription is proceeding well. If this is not the problem, alter the order of the peptides near the N terminus of the QconCAT to generate a new gene construct.

TABLE 1 | Troubleshooting table (continued).

	Protein toxicity	If the expression problem is due to toxicity the cells are likely to lyse upon induction; monitor the A_{600} of the culture during the induction phase. Alter the order of the peptides near the N terminus of the QconCAT to generate a new gene construct.
Poor ion intensities of selected Q-peptides	Many of the peptides terminate with lysine	Specific to MALDI-ToF MS. Investigate whether guanidination of the QconCAT peptides overcomes the problem[10–12].
	Peptide ionizes poorly by chosen MS ionization method	Try alternative ionization methods. Peptides that are known to yield poor signals in MALDI-ToF may perform well in ESI MS.
	Peptide chosen ionizes very weakly using any MS approach	Select another peptide to represent the protein in question. This requires generating a new gene. Fortunately this only involves changing a few of the oligonucleotides used for gene synthesis, and a new version can be rapidly and economically generated.
Incomplete labeling of QconCAT protein with stable isotopes	Precursor pool is incompletely labeled. Either the isotope has a relative isotope abundance less than 0.99 or endogenous, unlabeled amino acids are diluting the labeled precursor pool.	
	Minimal medium is contaminated with unlabeled versions of heavy-isotope precursors.	Re-prepare minimal medium from pure reagents.
	Too large an inoculum of the overnight culture grown in MM plus all amino acids was used to start the culture for the labeling step.	Decrease the size of the inoculum.
	Insufficient heavy isotope–labeled precursor is added, forcing *E. coli* to synthesize amino acids *de novo*, which are not labeled.	Add more heavy-isotope precursor for growth and induction, or use an auxotroph for the amino acid.
Incomplete digestion of analyte proteins and QconCAT	Insufficient trypsin	Increase ratio of trypsin to analyte proteins. Alter digestion conditions to enhance proteolysis. Pretreat analyte proteins to increase susceptibility to digestion.
	Conditions suboptimal	Add organic solvent (e.g., CAN) to 10%.
	Proteins are tightly folded	Include a denaturant (e.g., urea 2–4 M), and increase the trypsin concentration because the denaturant will diminish trypsin activity.
	Proteins are very hydrophobic membrane proteins, with cleavage sites unavailable due to steric hindrance or sequestration within membrane vesicles	Solubilize membrane preparations using, for example, organic solvents, organic acids or detergents[7–9] (e.g., nonionic detergent such as Rapigest; Waters (http://www.waters.com)).

ANTICIPATED RESULTS

During the course of the induction of expression of QconCAT proteins, a band corresponding to the expected molecular weight of the QconCAT protein should be readily seen to increase in intensity, confirming good levels of expression (**Fig. 4**).

Purification and SDS-PAGE analysis confirms that the QconCAT was readily solubilized from the inclusion bodies, bound efficiently to the HisTrap column (i.e., was not present in the unbound fraction) and was the major band in the bound eluted fractions (**Fig. 5**).

ESI MS analysis of the intact QconCAT protein allows comparison of the experimentally derived molecular weight and the calculated molecular weight (**Fig. 6a**). Loss of the N-terminal methionine can be predicted, and the degree to which other potential modifications to the primary structure have occurred can be evaluated (e.g., oxidation of methionine). Analysis of the stable isotope–labeled QconCAT protein allows evaluation of the labelling step (**Fig. 6b**). Protein peaks that are present between the unlabeled and labeled protein predicted masses could indicate incomplete labeling with heavy isotope precursors.

MALDI-ToF analysis of labeled and unlabeled QconCATs (in this case labeled with [$^{13}C_6$]lysine and [$^{13}C_6$]arginine) allows confirmation of the consistent 6-Da difference in mass between labeled and unlabeled forms of each peptide, unless an uncleavable sequence (e.g., RP) is present, as is the case with peptide T17, which differs from its unlabeled counterpart by 12 Da (**Fig. 7**).

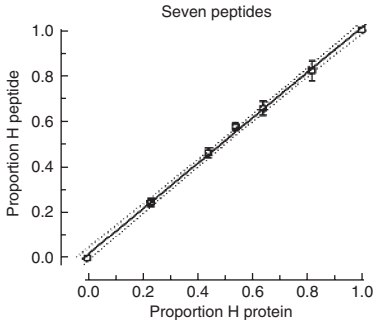

Figure 8 | Quantification by QconCATs. Unlabeled (L, light) and uniformly labeled with ^{15}N (H, heavy) QconCAT proteins were separately purified, quantified and mixed in different ratios, before tryptic digestion and measurement of peptide intensities by MALDI-ToF MS. The measured proportion of H peptide is plotted relative to the proportion of H protein in the mixture for three replicates of each of seven peptides (right); error bars ± s.e.m., $n = 7$. The dotted lines define the 95% confidence limits of the fitted line. (Adapted from reference 6.)

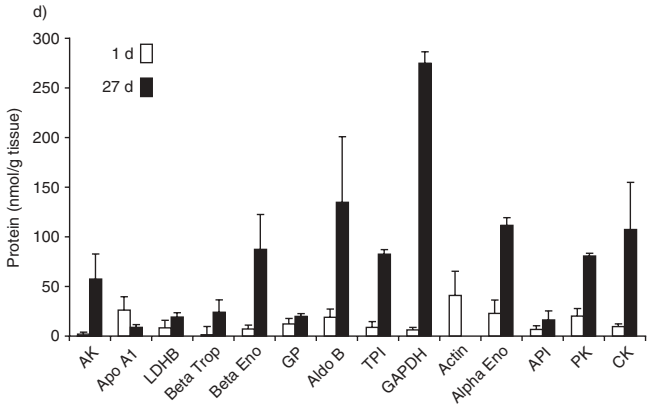

Figure 9 | Quantification of biological samples by QconCAT. A preparation of soluble proteins from 100 mg of chicken skeletal muscle at 1 d and 27 d was mixed with 290 μg ^{15}N-labeled QconCAT, digested with trypsin overnight and analyzed by MALDI-ToF. The intensity of the monoisotopic peak for the analyte peptide and the corresponding QconCAT peptide were recorded and the data converted into nanomoles of protein per gram of tissue, to give the absolute amount of each protein. Error bars, ± s.e.m.; $n = 3$. For details of the individual proteins, see reference 6. (Adapted from reference 6.)

Statistical analysis of labeled and unlabeled QconCAT proteins, mixed in known proportions, digested and analyzed by MALDI-ToF MS (**Fig. 8**) allows evaluation of the reproducibility of the MALDI-ToF analysis. In addition, the behavior of each peptide within the QconCAT can be assessed and the 95% confidence limits, of the line fitted to the data and visualized.

The data shown in **Figure 9** are adapted from ref. 6 and show the absolute quantification values, as well as standard errors, that can be obtained using QconCAT proteins as internal standards. Results are presented as nanomoles of each protein per gram of tissue.

Note: Supplementary information is available via the HTML version of this article.

ACKNOWLEDGMENTS This work was supported by the Biotechnology and Biological Sciences Research Council, Swindon, UK.

COMPETING INTERESTS STATEMENT The authors declare competing financial interests (see the HTML version of this article for details).

Published online at http://www.natureprotocols.com/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Julka, S. & Regnier, F.E. Recent advancements in differential proteomics based on stable isotope coding. *Brief. Funct. Genomic. Proteomic.* **4**, 158–177 (2005).
2. Ong, S.E. & Mann, M. Mass spectrometry–based proteomics turns quantitative. *Nat. Chem. Biol.* **1**, 252–262 (2005).
3. Yan, W. & Chen, S.S. Mass spectrometry–based quantitative proteomic profiling. *Brief. Funct. Genomic. Proteomic.* **4**, 27–38 (2005).
4. Gerber, S.A., Rush, J., Stemman, O., Kirschner, M.W. & Gygi, S.P. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. USA* **100**, 6940–6945 (2003).
5. Kirkpatrick, D.S., Gerber, S.A. & Gygi, S.P. The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods* **35**, 265–273 (2005).
6. Beynon, R.J., Doherty, M.K., Pratt, J.M. & Gaskell, S.J. Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. *Nat. Methods* **2**, 587–589 (2005).
7. Wu, C.C. & Yates, J.R. The application of mass spectrometry to membrane proteomics. *Nat. Biotechnol.* **21**, 262–267 (2003).
8. Lu, X. & Zhu, H. A novel proteomic approach for high throughput analysis of membrane proteins. *Mol. Cell Proteomics* **4**, 1948–1958 (2005).
9. Fischer, F. & Poetsch, A. Protein cleavage strategies for an improved analysis of the membrane proteome. *Proteome Sci.* **4**, 1–12 (2006).
10. Brancia, F.L. *et al.* A combination of chemical derivatisation and improved bioinformatic tools optimises protein identification for proteomics. *Electrophoresis* **22**, 552–559 (2001).
11. Beardsley, R.L. & Reilly, J.P. Optimization of guanidination procedures for MALDI mass mapping. *Anal. Chem.* **74**, 1884–1890 (2002).
12. Thevis, M., Ogorzalek Loo, R.R. & Loo, J.A. In-gel derivatization of proteins for cysteine-specific cleavages and their analysis by mass spectrometry. *J. Proteome Res.* **2**, 163–172 (2003).
13. Beynon, R.J. & Pratt, J.M. Metabolic labelling of proteins for proteomics. *Mol. Cell Proteomics* **4**, 857–8872 (2005).
14. Ellison, D., Hinton, J., Hubbard, S.J. & Beynon, R.J. Limited proteolysis of native proteins: the interaction between avidin and proteinase K. *Protein Sci.* **4**, 1337–1345 (1995).
15. Hubbard, S.J. The structural aspects of limited proteolysis of native proteins. *Biochim. Biophys. Acta* **1382**, 191–206 (1998).
16. Hubbard, S.J., Beynon, R.J. & Thornton, J.M. Assessment of conformational parameters as predictors of limited proteolytic sites in native protein structures. *Protein Eng.* **11**, 349–359 (1998).
17. Wu, C., Robertson, D.H., Hubbard, S.J., Gaskell, S.J. & Beynon, R.J. Proteolysis of native proteins. Trapping of a reaction intermediate. *J. Biol. Chem.* **274**, 1108–1115 (1999).
18. Hubbard, S.J. & Beynon, R.J. Proteolysis of native proteins as a structural probe. In *Proteolytic Enzymes. A Practical Approach* (eds. Beynon, R.J. & Bond, J.S.) 233–264 (Oxford University Press, Oxford, 2001).
19. Sambrook, J. & Russell, D.W. (eds.) Preparation and transformation of competent *E. coli* using calcium chloride (Protocol 25). In *Molecular Cloning: A Laboratory Manual* 3rd Edn. Vol 1. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001).

QUERY FORM

NPROT	
Manuscript ID	[Art. Id: 129]
Author	
Editor	
Publisher	

AUTHOR:

The following queries have arisen during the editing of your manuscript. Please answer queries by making the requisite corrections directly on the galley proof. It is also imperative that you include a typewritten list of all corrections and comments, as handwritten corrections sometimes cannot be read or are easily missed. Please verify receipt of proofs via e-mail

Query No.	*Nature of Query*
.	No Queries