

DGMS40 2007 - Workshop Quantitative Proteomics 11. März, Bremen



**QconCAT – eine neue Methode zur parallelen
absoluten Quantifizierung von Proteinen
in der Massenspektrometrie**

Proteomics 2007 (1)

Western blot

ELISA (Enzyme Linked Immunosorbent Assay)

RIA (Radio Immunosorbant Assay)

FACS analysis (cell based)

Luminex (bead based)

DIGE (2 Dimensional Gel Electrophoresis)

MALDI-TOF identification of proteins

Proteomics 2007 (2)

Multidimensional Liquid Chromatography

ICAT - derivatisation of cysteine residues

**ITRAQ- Labeling of N-termini, multiplexing
limited to six samples**

SILAC

ICPL

**Isotope labeled quantitation peptides –
AQUA-peptides**

Proteomics 2017 (1)

Multiplex parallel quantitation of proteins by MS

- Used for diagnosis, prognosis and individual therapy control
- Wide applications in controlling processes in biotechnological production
- Important tool for systems biology (whole proteomes available)
- In scientific groups used to analyse changes in protein complexes, cell organelles and membrane molecules (e. g. lipid rafts)

Proteomics 2017 (2)

Analyses will be possible using

- **Unmodified proteolytic peptides**
- **Natural peptides or protein fragments**
- **Phosphorylated peptides**
- **N- and O- glycosylated peptides**
- **Other modifications (ubiquitination, suomylation, acetylation etc.)**

QconCAT used as routine method to verify effects of RNAi therapy

AQUA – chemical synthesis (SIGMA-ALDRICH)

- enables **absolute protein quantitation** using **stable isotope labeled peptides** as an internal standard and HPLC-MS
- **Step 1: AQUA Peptide Selection**
Select an optimal tryptic peptide and stable isotope amino acid from the sequence of your protein of interest , Order synthetic AQUA Peptide, Optimize LC-MS/MS
- **Step 2: Implementation**
Extract protein from biological samples and add known quantity of AQUA Peptide , Digest , Analyze by LC-MS/MS to quantitate protein of interest

Gerber, Scott A, et al. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. PNAS. June 10, 2003. Vol 100. No 12. p 6940-6945

Problems

- **Absolute quantification expensive**
- **Chemical synthesis of reference peptides not always feasible**
- **Sources of errors due to different treatment of reference peptide and target proteins**

QconCAT – biosynthesis of multiple peptides

- Absolute **Q**uantification
- Based on mass tagged internal standards (Q-peptides)
- One Q-peptide per analyte protein
- Multiple **Q**-peptides **conCAT**enated into a synthetic gene, expressed as a heterologous QconCAT protein
- Easy construction of large sets of mass tag peptides
- Tryptic digest of reference and target peptides in the *same* tube
- Separation of reference and target peptides in the same solution

The method 1 - selection of protein targets

MS-derived peptides

DIGE-derived proteins

Western blot

ELISA

DNA-Array experiments

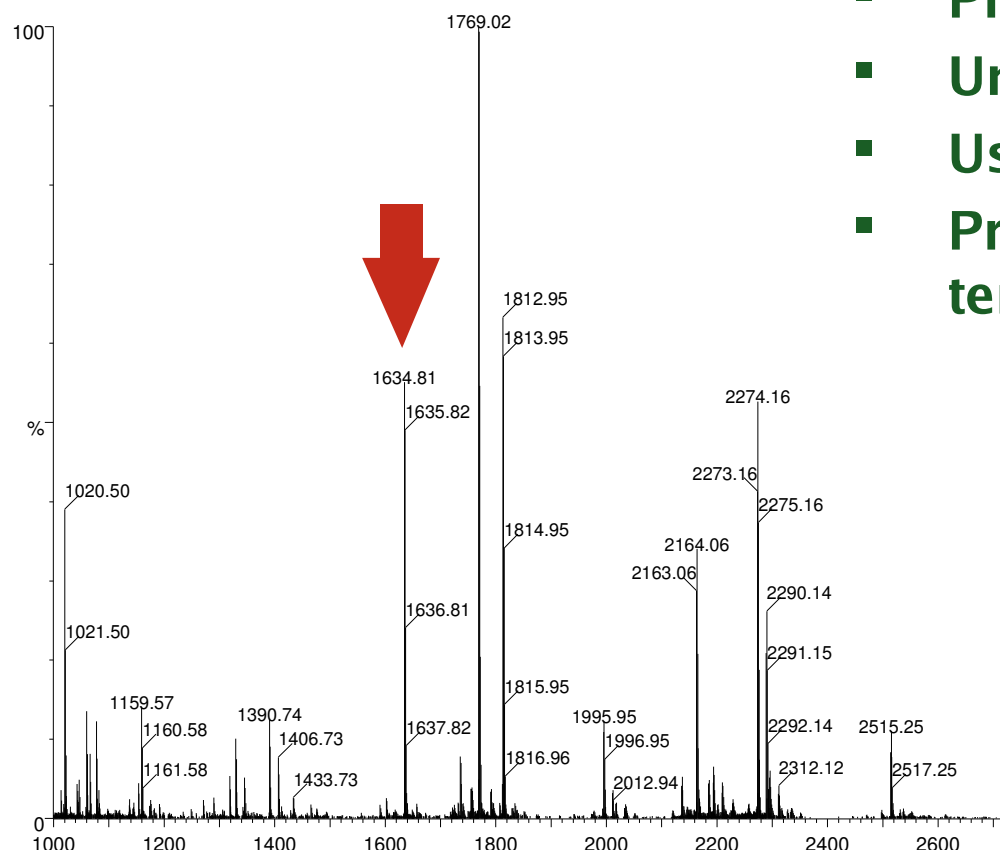
Genetic linkage analysis

SNP analysis (Small nucleic acid polymorphisms)

etc.

The method 2 - selection of Q-peptides (1)

- One per analyte protein
- Propensity to ionise
- Uniqueness in set
- Useable mass
- Preferably arg-terminated



The method 2 - selection of Q-peptides (2)

- **Filter the list of peptides from the discovery data**

- Acceptable Peptides:

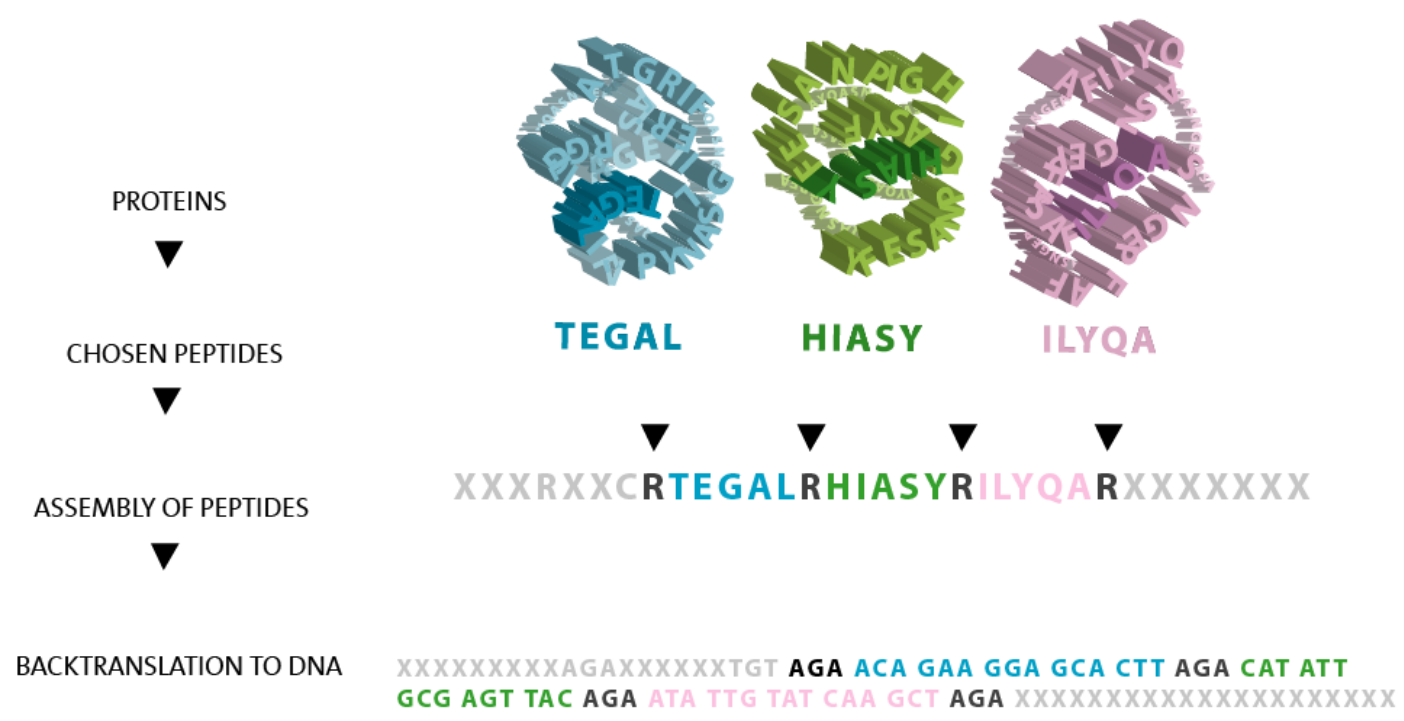
- Peptides should be 6-20 amino acids (preferably 8-15)
- Eliminate peptides containing chemically reactive residues (C, M, W)
- Eliminate chemically unstable residues (NG, DG, QG, N-term N, N-term Q)
- Eliminate peptides next to “ragged ends” (sequential predicted trypsin cleavage site, i.e. AGH**DR**KGLILFSDR)
- Eliminate “LC-incompatible” peptides (high or low predicted retention times- i.e., 8-30 minutes, preferably 10-20 minutes)
- Eliminate peptides containing an arginine adjacent to proline (potential missed cleavage)
- Eliminate known sites of post-translational modifications (Human Protein Reference Database, www.hrpdb.org)
- The group of peptides that meet these criteria should be BLASTed to ensure that they are unique to the target protein

Introduced Mass Shifts By Heavy Amino Acid(s)

Stable Isotope Labeled Amino Acid Options

Stable Isotope Labeled Amino Acid (Fully Labeled, ^{13}C, ^{15}N > 98 atom%)	Mass Difference between Native Peptide and AQUA Peptide
L-Arginine- $^{13}\text{C}_6, ^{15}\text{N}_4$	10 Daltons
L-Isoleucine- $^{13}\text{C}_6, ^{15}\text{N}$	7 Daltons
L-Leucine- $^{13}\text{C}_6, ^{15}\text{N}$	7 Daltons
L-Lysine- $^{13}\text{C}_6, ^{15}\text{N}_2$	8 Daltons
L-Phenylalanine- $^{13}\text{C}_9, ^{15}\text{N}$	10 Daltons
L-Proline- $^{13}\text{C}_5, ^{15}\text{N}$	6 Daltons
L-Valine- $^{13}\text{C}_5, ^{15}\text{N}$	6 Daltons

The method 3 - creation of the concatamer



The method 4 - typical QconCAT gene

CATATCGCGG GTAAA GTTAT CCGTG GCTTC CTGAT CGACG GTTAC CCGCGT GTAGT TCTGG CITAT GAACC GGTGT GGGCC ATCGG CACCG GTAAA AAC CTG GCGCC ATACT CCGAT GAA
 M A G K V I R G F L I D G Y P R V V L A Y E P V W A I G T G K N L A P Y S D E

M

Adenyla te kinase

Triosephosphate isomerase

Apolipoprotein AI

CTGCGCGGC GATCA GCTGT TTACCG CGACT GAAGGT CGTAG CTATG AACTGC CGGAT GGTTCAGGTGATTACTAT TGGTA ACGAA CGTCAG GTTGT GGAAA GC C C T ACGAA GTTATC CGT
 L R G D Q L T T A T E G R S Y E L P D G Q V I T I G N E R Q V V E S A Y E V I R

Myosin binding C

Alpha actin

Lactate dehydrogenase B

CTGAT TACCG GCGAA CAGCT GGTGT AAATT TATCG TGCTA CCGAC GCAGAA TCTGA AGTTG CCAGC CTGAA CCGTT CCCTG GAAGA TCAGC TGTCC GAAATC AAGT GCTGT ATCCA AAC
 L I I G E Q L G E I Y R A I D A E S E V A S L N R S L E D Q L S E I K V L Y P N

Beta enolase

Tropomyosin A

Embryonic myosin

GATAACTTCTTTGAAGGCAAAGGCATCCTGGCAGCGGACGAATCCGTTGGCACCATGGGTAACCGCGCCACCGACGCTGAAGCTGAAGTTGCGTCCCTGAACCGTCTGCAGAACGAAGTT
 L N F F E G K G I L A A D E S V G I M G N R A I D A E A E V A S L N R L Q N E V

Glycoge n phosphorylase

Aldolase B

Tropomyosin B

GAAGA CCTGA TGGTT GATGT TGAAC GCCTG GTTTC TTGGT ACGAC AACGAG TTCGG TTA CT CCAAC CGTGC TCTGG AATCT CCGGA ACGCC CGTTC CTG GCG ATCCT GG GTG CCGCT AAA
 E D L M V D V E R L V S W Y D N E F G Y S N R A L E S P E R P F L A I L G G A K

Adult myosin HC

GAPDH

Phosphoglycerate kinase

CAGGT AGTGG ACTCT GCCTA CGAAG TTATC AAAGC AGCTG TACCG AGCGGC CCGTC TACCG GTATT TACGA AGCAC TGGAA CTGCG TCTGC TGCCG TCTGAA TCTGC CTTGC TGCCG GCA
 C V V D S A Y E V I K A A V P S G A S T G I Y E A L E L R L L P S E S A L L P A

La ctate dehydrogena se A

Alpha enolase

Actin polymerisation inhibitor

CCGGG TTCTC CGTAC GGCCG CTTTG GTGTG GAACA GAACG TTGAT ATGGTG TTCGC GTCTT TCATC CGTGG TACTG GTGGC GTTGA TACCG CTGCA GTAAGT GCGGT ATTTCG ACATC AGC
 F G S P Y G R F G V E Q N V D M V F A S F I R G T G G V D T A A V G A V F D I S

Pyruvate kinase

Creatinine kinase

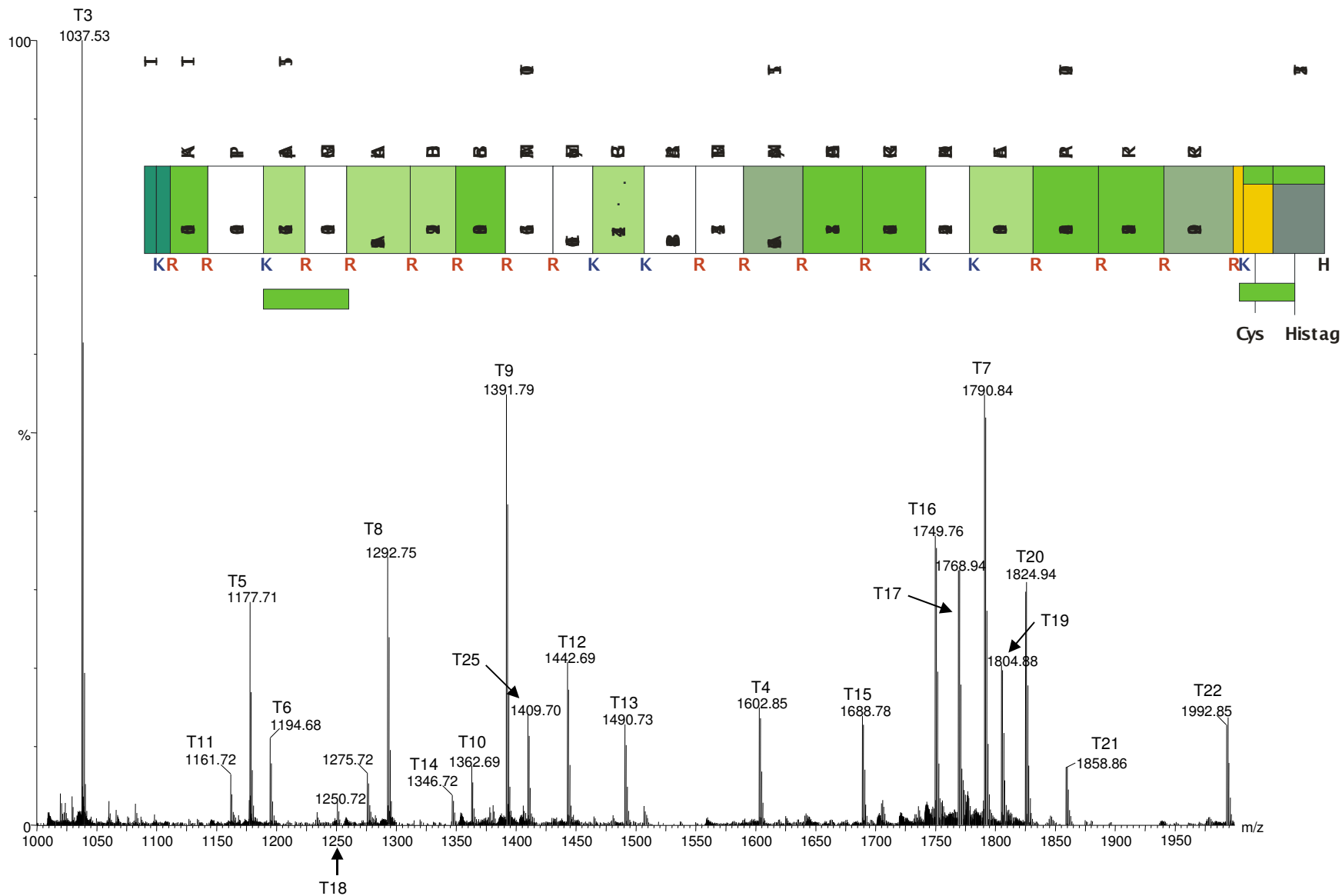
AACGC AGACC GTGCG GGTAA AGTTA TCTGC TCTGC GGAAG GATCC AAGCTT GCGGC CGCAC TCGAG CACCA CCACC ACCAC CACTG A
 N A D R A G K V I C S A E G S K L A A A L E H H H H H H *

CACCA CCACC ACCAC CACTG A
 H H H H H H *

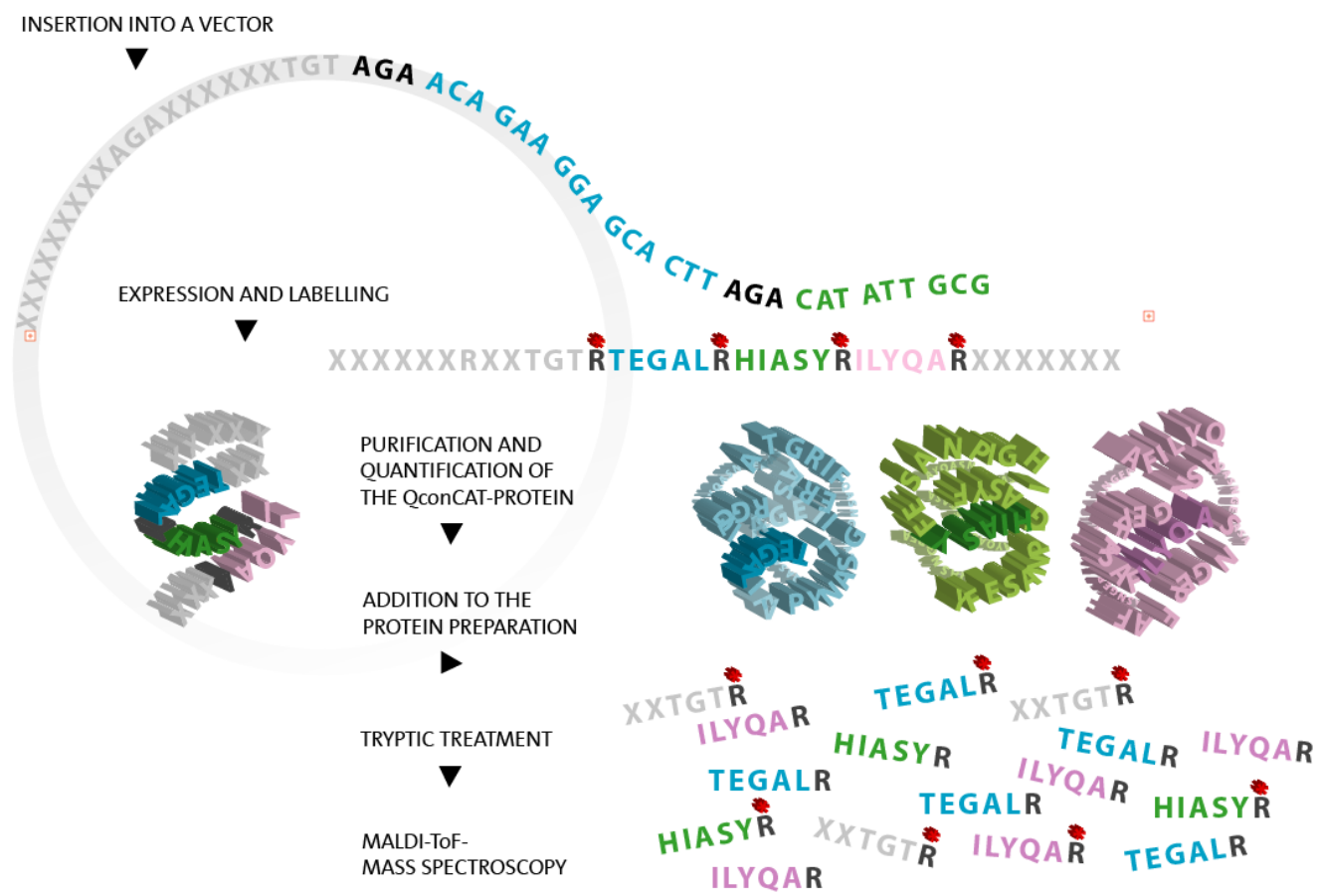
Quantification

Purification

MALDI-ToF of unlabelled QconCAT

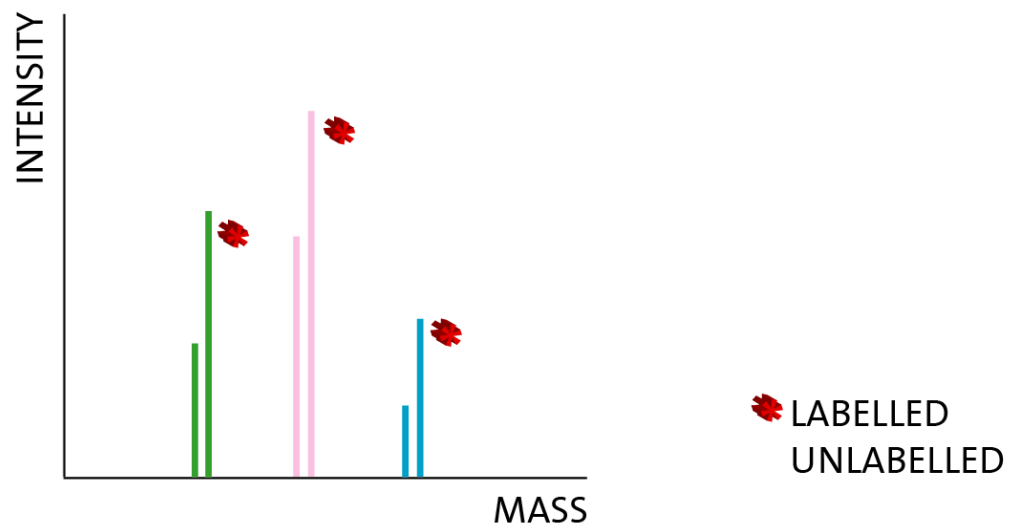


The method 5 - expression, labelling, digest

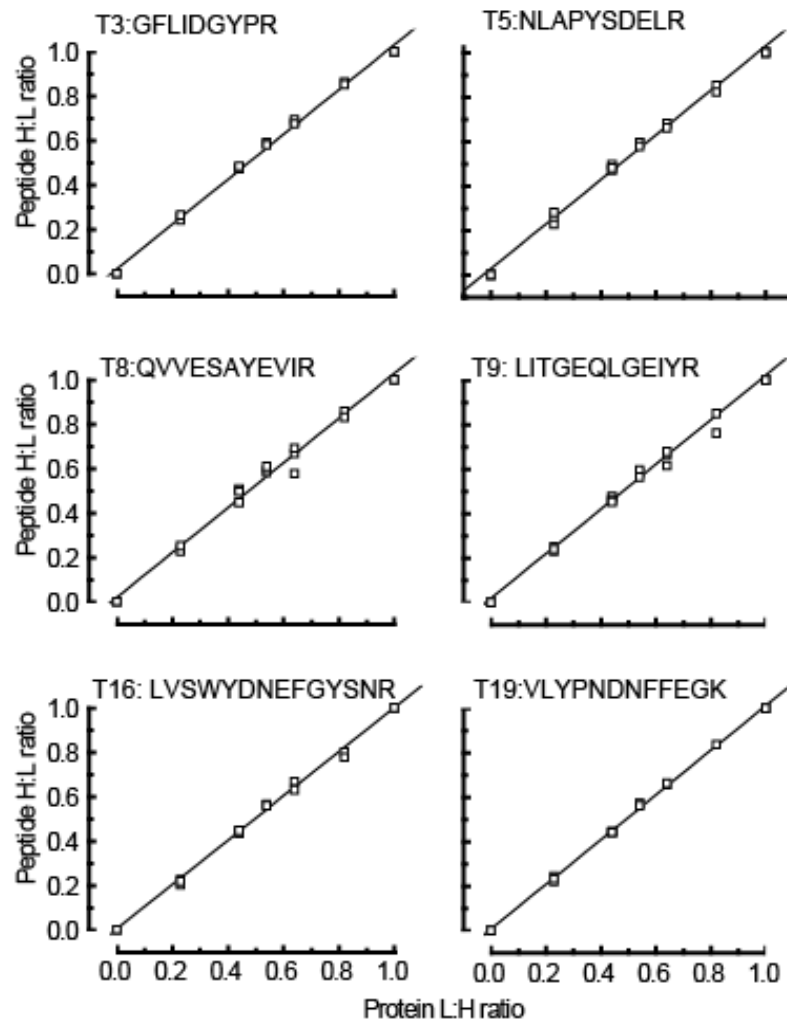


The method 6 - analysis

Mass spectrometry



The method 7 – proof of principle



- 15N-labelled and unlabelled QconCAT proteins
- Mixed in different ratios
- Proteolysed before MALDI-ToF

Advantages

- Chemically inaccessible peptides can be produced biosynthetically
- Absolute quantification at reasonable costs
- High fault tolerance, e.g. against preparation artifacts
- Low costs for isotope labelling
- Re-usability of the concatamer gene
- 'Themed' peptide sets can be provided off the shelf as quantification kits
- Realistic approach to quantify whole proteomes

Applications

- **Highly accurate quantification of large sets of analyte proteins**
- **Quantitative analysis of proteomes**
- **Comparison of results between labs**
- **High-precision modelling in systems biology**
- **Quantification of phosphorylation states and other covalent modifications (in future)**
- **Diagnosis**
- **Food quality control, environmental research**

Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides

Robert J Beynon¹, Mary K Doherty¹, Julie M Pratt¹ & Simon J Gaskell²

Absolute quantification in proteomics usually involves simultaneous determination of representative proteolytic peptides and stable isotope-labeled analogs. The principal

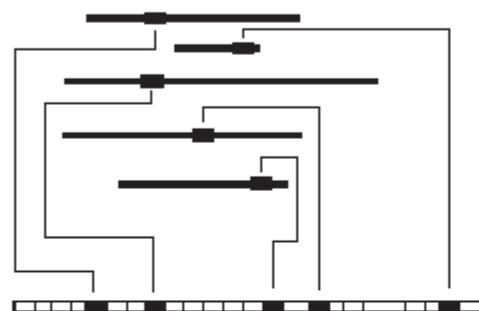
Inventors

- Prof. Rob Beynon, University of Liverpool
- Prof. Simon Gaskell, University of Manchester
- Dr. Julie Pratt, independent consultant

Patent application

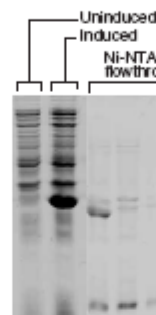
- Entelechon GmbH

a



Assembly of signature Q peptides into QCAT protein

b



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

Julie M Pratt¹, Deborah M Simpson¹, Mary K Doherty¹, Jenny Rivers¹, Simon J Gaskell² & Robert J Beynon¹

¹Department of Veterinary Preclinical Sciences, University of Liverpool, Crown Street, Liverpool, L69 7ZJ, UK. ²Michael 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>.

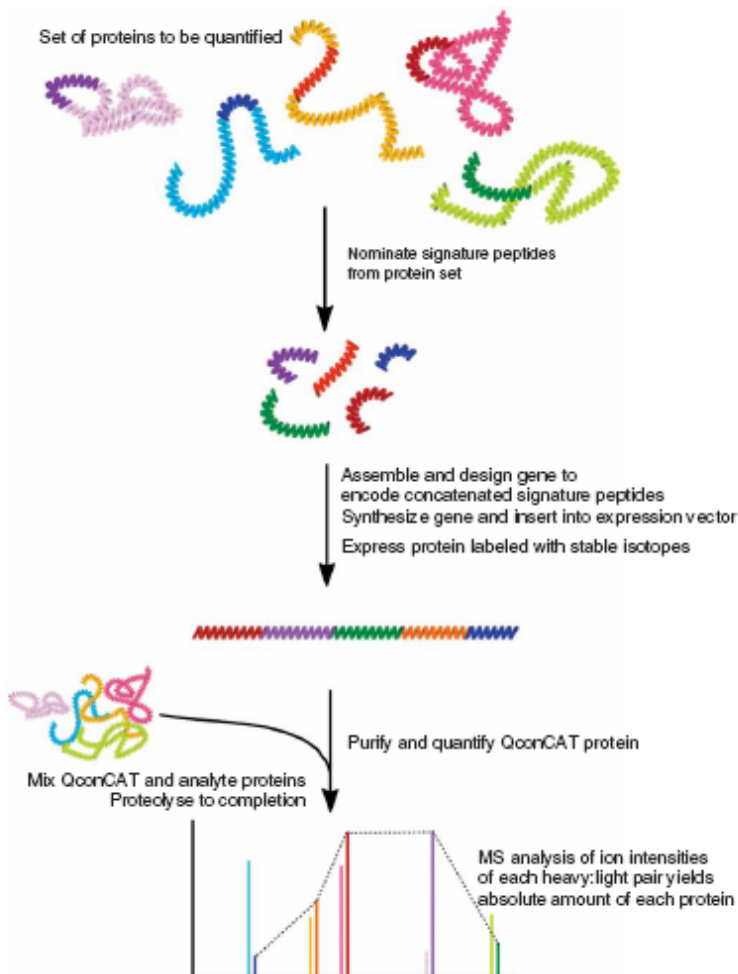


Figure 1 | General principle of a QconCAT quantification experiment.

More information

www.qconcat.com

Entelechon GmbH
www.entelechon.com

giegerich@entelechon.com

+49 (0)941 / 69 818 17

Cooperations 2007

Rob Beynon, Julie Pratt, Liverpool

Simon Gaskell, Glasgow

Ruedi Aebersholt, Bernd Wollscheid,

Vinzenz Lange, Zürich

**Saleh Ibrahim, Martin Eggert, Gerhard
Neeck, Rostock**

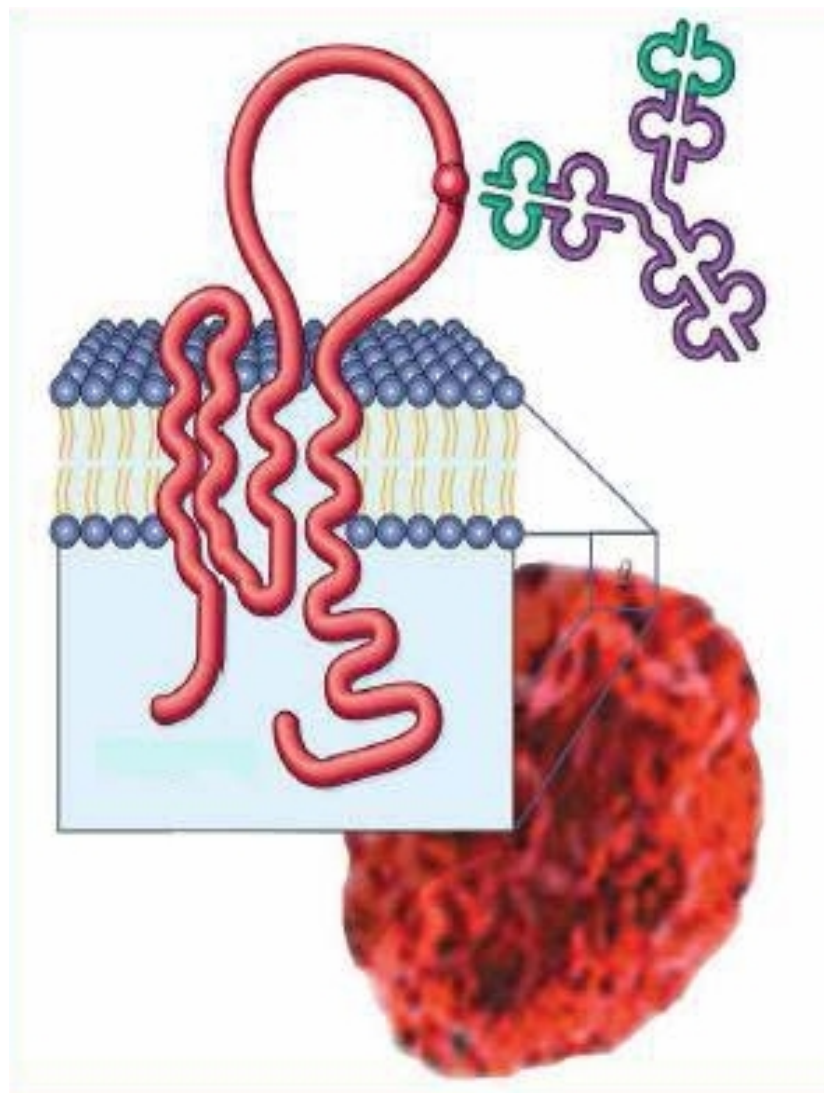
Peter Öfner, Regensburg

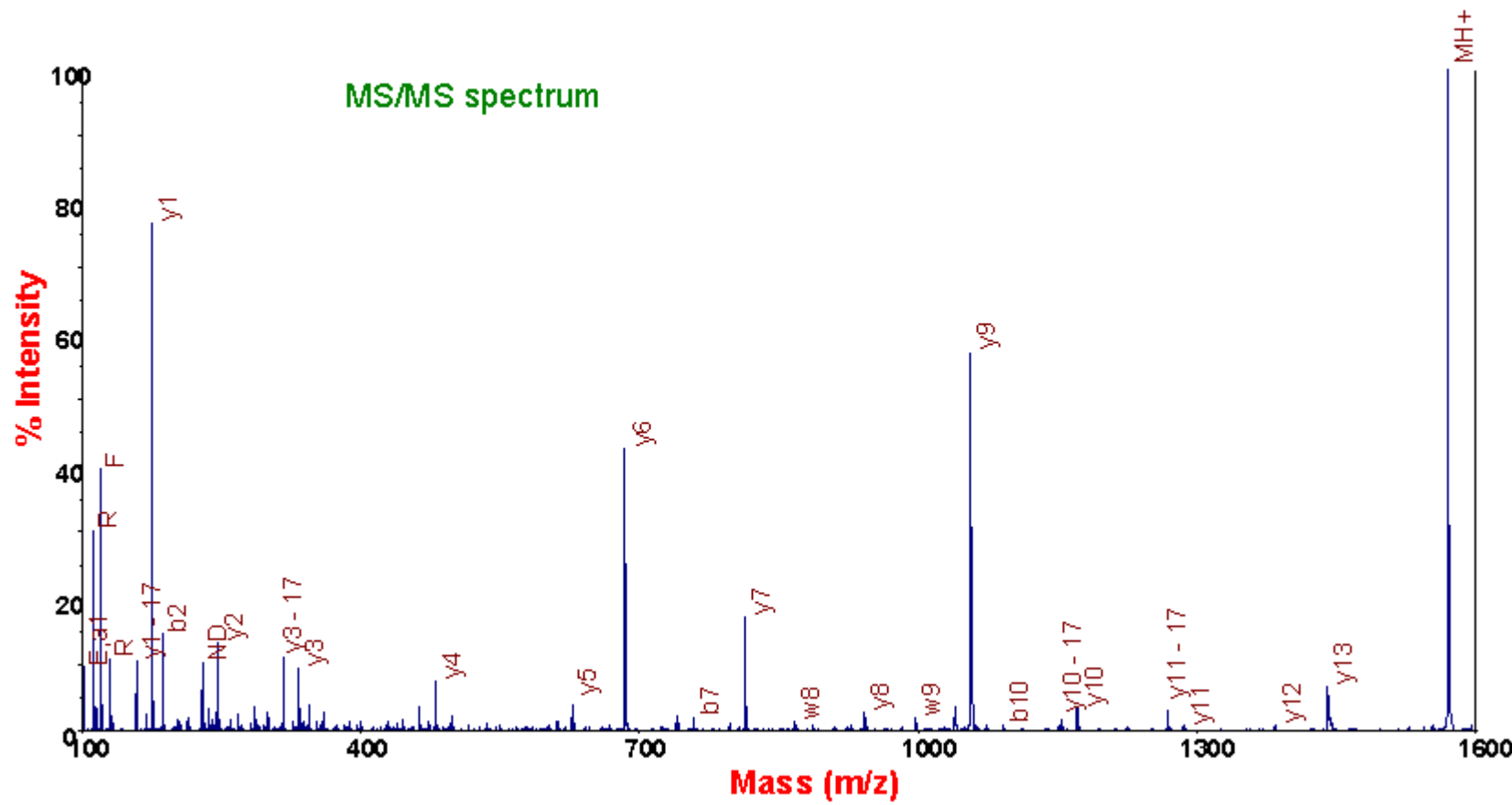
Panatecs, Tübingen

OMX, München

Proteome Factory, Berlin









**A new Method for the multiplexed
absolute Quantification of Proteins**

Protein quantification - existing approaches

- Switch from gel-based methods to isotope-labeled derivatisation methods
- DIGE (2D differential gel electrophoresis)
- ICAT - derivatisation of cysteine residues
- iTRAQ - derivatisation of N-terminus
- Comparable accuracy
- Specific drawbacks:
 - DIGE: Complex procedure, different migration rates for isoforms
 - ICAT: Requirement for cys residues
 - iTRAQ: Multiplexing limited to four independent samples

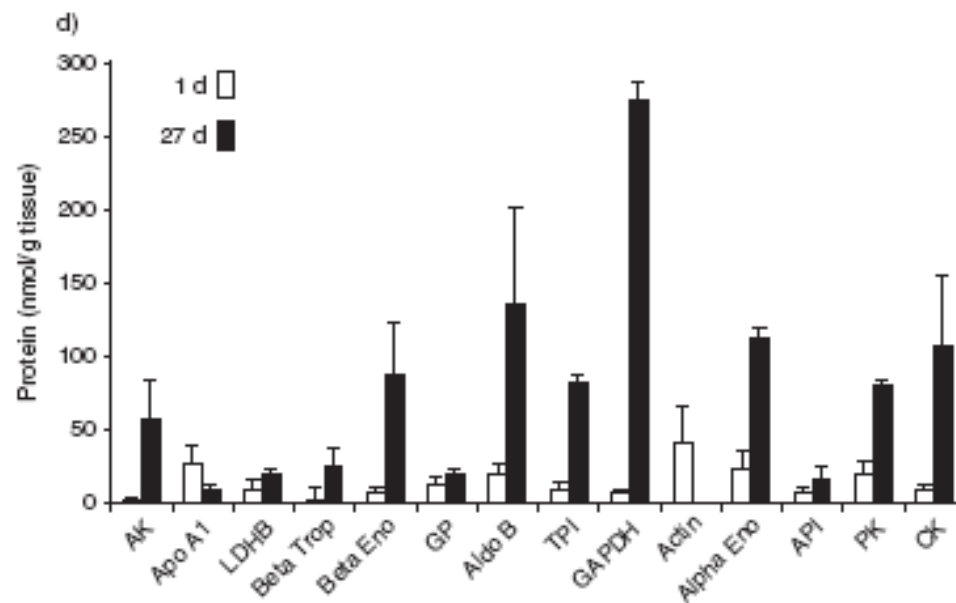


Figure 9 | Quantification of biological samples by QconCAT. A preparation of soluble proteins from 100 mg of chicken skeletal muscle at 1 d and 27 d was mixed with 290 μg ^{15}N -labeled QconCAT, digested with trypsin overnight and analyzed by MALDI-ToF. The intensity of the monoisotopic peak for the 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, \pm s.e.m.; $n = 3$. For details of the individual proteins, see reference 6. (Adapted from reference 6.)

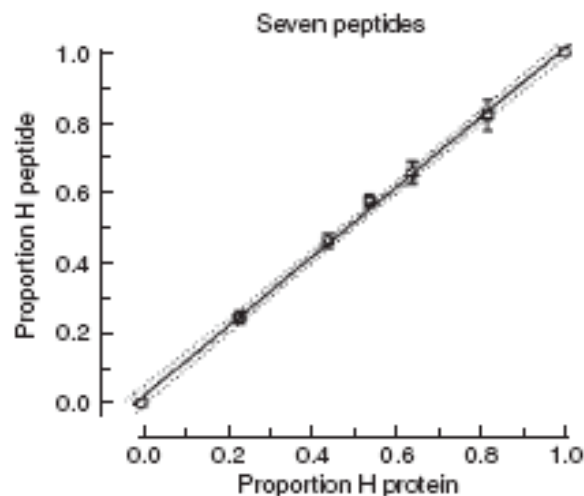
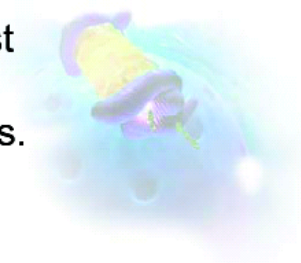


Figure 8 | Quantification by QconCATs. Unlabeled (L, light) and uniformly labeled with ^{15}N (H, heavy) QconCAT proteins were separately purified, quantified and mixed in different ratios, before tryptic digestion and measurement of peptide intensities by MALDI-ToF MS. The measured proportion of H peptide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

Quantitative Mass Spectrometry

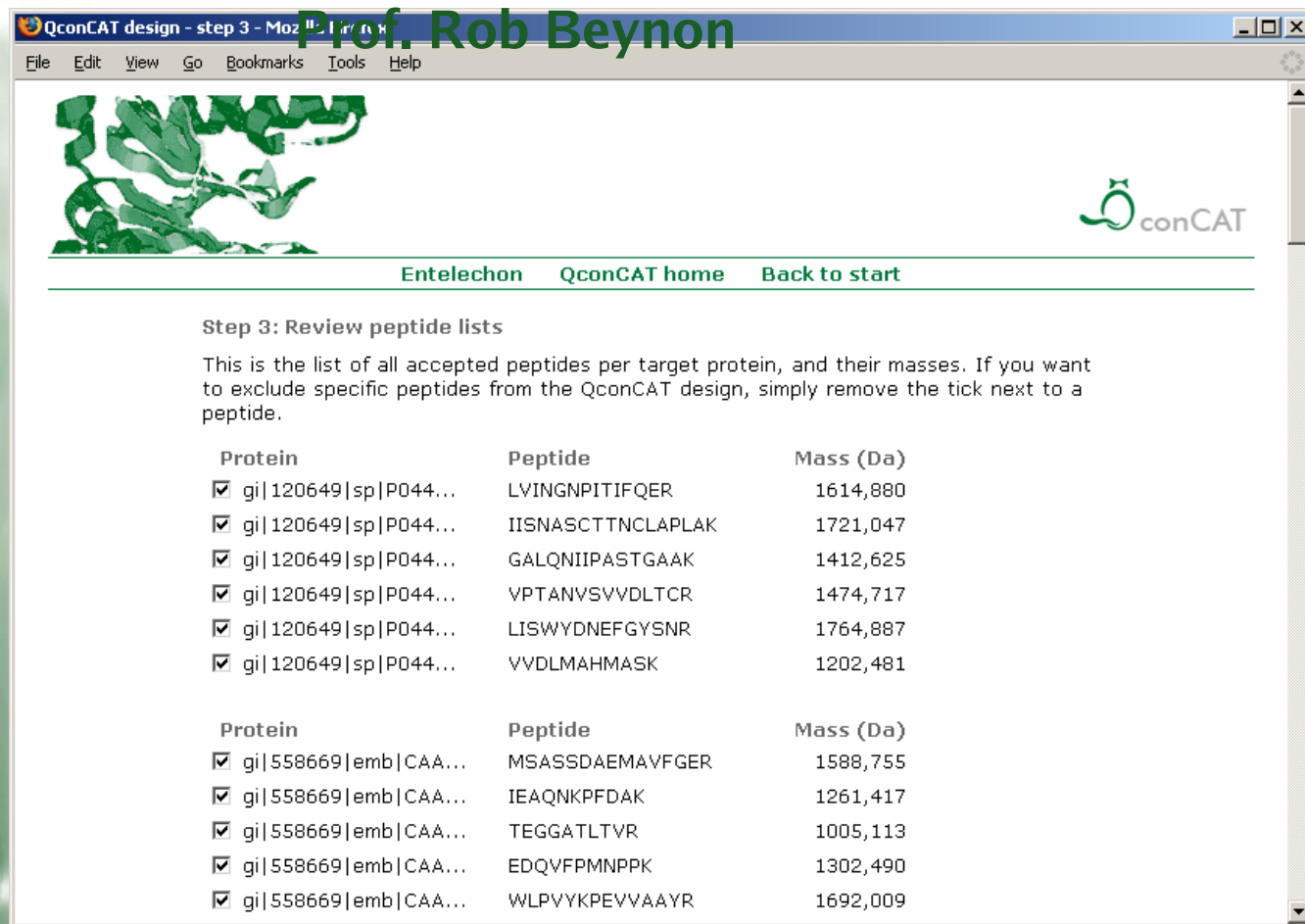
- **Due to Non-uniform Instrument Response, Mass Spectrometric Quantitation Must Be Approached Very Carefully:**
 - Only chemically identical peptides may be quantitatively compared.
 - Stable isotopes are introduced into one sample, then:
 - The same protein(s)/peptide(s) in the control and test samples are mass spectrometrically compared, i.e., one “native” and one labeled with stable isotopes.



Web application for QconCAT design

Based on the QconCAT design helper tool
by

Prof. Rob Beynon



The screenshot shows a web browser window titled "QconCAT design - step 3 - Mozilla Firefox". The browser's address bar is partially visible. The page content includes a navigation menu with links for "Entelechon", "QconCAT home", and "Back to start". Below the navigation menu, the page is titled "Step 3: Review peptide lists". A paragraph explains that this is a list of accepted peptides per target protein, and users can exclude specific peptides by removing the tick next to a peptide. The main content consists of two tables, each with three columns: "Protein", "Peptide", and "Mass (Da)".

Protein	Peptide	Mass (Da)
<input checked="" type="checkbox"/> gi 120649 sp P044...	LVINGNPITIFQER	1614,880
<input checked="" type="checkbox"/> gi 120649 sp P044...	IISNASCTTNCLAPLAK	1721,047
<input checked="" type="checkbox"/> gi 120649 sp P044...	GALQNIIPASTGAAK	1412,625
<input checked="" type="checkbox"/> gi 120649 sp P044...	VPTANVSVDLTCR	1474,717
<input checked="" type="checkbox"/> gi 120649 sp P044...	LISWYDNEFGYSNR	1764,887
<input checked="" type="checkbox"/> gi 120649 sp P044...	VVDLMAHMASK	1202,481

Protein	Peptide	Mass (Da)
<input checked="" type="checkbox"/> gi 558669 emb CAA...	MSASSDAEMAVFGER	1588,755
<input checked="" type="checkbox"/> gi 558669 emb CAA...	IEAQNKPFDAK	1261,417
<input checked="" type="checkbox"/> gi 558669 emb CAA...	TEGGATLTVR	1005,113
<input checked="" type="checkbox"/> gi 558669 emb CAA...	EDQVFPMPNPPK	1302,490
<input checked="" type="checkbox"/> gi 558669 emb CAA...	WLPVYKPEVVAAYR	1692,009

Web application for QconCAT design

QconCAT - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

gi|416950|sp|P139... VVIGMDVAASEFYR 1557,806

Peptide set #05:

Select this peptide set

Prefix:

Suffix:

Protein	Peptide	Mass (Da)
gi 120649 sp P044...	LISWYDNEFGYSNR	1764,887
gi 558669 emb CAA...	AITDAAMMAEELK	1394,649
gi 178322 gb AAA5...	IIFVVGGPGSGK	1131,357
gi 4507645 ref NP...	QSLGELIGTLNAAK	1415,626
gi 4960066 gb AAD...	LSPLGEEMR	1032,203
gi 4557032 ref NP...	GEMMDLQHGSFLQTPK	1933,254
gi 416950 sp P139...	LAMQEFMILPVGASSFK	1870,282

Peptide set #06:

Select this peptide set

Prefix:

Suffix:

Protein	Peptide	Mass (Da)
gi 120649 sp P044...	VVDLMAHMASK	1202,481
gi 558669 emb CAA...	HADSMaelGEQIDNLQR	1928,085
gi 178322 gb AAA5...	EVQQGEEFER	1251,292
gi 4507645 ref NP...	VTNGAFTGEISPGMIK	1622,878

Frequently asked questions

- **Efficiency of the tryptic digest**
 - How efficient is the digest?
 - What if target proteins and QconCAT are digested at different rates?
- **Quantification of the concatamer**
 - How accurate is it?
- **Isotope-labelling**
 - Available options
 - Accuracy
- **Turnaround time**
- **Separation of peptides**

The method 4 - typical QconCAT gene

CATAT GGCGG GTAAA GTTAT CCGTG GCTTC CTGAT CGACG GTTAC CCGCGT GTAGT TCTGG CITAT GAACC GGTGT GGGCC ATCGG CACCG GTAAA AAC CTG GCGCC ATACT CCGAT GAA
 M A G K V I R G F L I D G Y P R V V L A Y E P V W A I G T G K N L A P Y S D E

Quantification

Purification

Adenylate kinase

Triosephosphate isomerase

Apolipoprotein AI

CTGCGCGC GATCA GCTGT TTACCG CGACT GAAGGT CGTAG CTATG AACTGC CGGAT GGTTCAGGTGATTACTAT TGGTA ACGAA CGTCAG GTTGT GGAAA GC CCT ACGAA GTTATC CGT
 L R G D Q L T T A T E G R S Y E L P D G Q V I T I G N E R Q V V E S A Y E V I R

Myosin binding C

Alpha actin

Lactate dehydrogenase B

CTGAT TACCG GCGAA CAGCT GGTGT AAATT TATCG TGCTA CCGAC GCAGAA TCTGA AGTTG CCAGC CTGAA CCGTT CCCTG GAAGA TCAGC TGTCC GAAATC AAGT GCTGT ATCCA AAC
 L I I G E Q L G E I Y R A I D A E S E V A S L N R S L E D Q L S E I K V L Y P N

Beta enolase

Tropomyosin A

Embryonic myosin

GATAACTTCTTTGAAGGCAAAGGCATCCTGGCAGCGGACGAATCCGTTGGCACCATGGGTAACCGCGCCACCGACGCTGAAGCTGAAGTTGCGTCCCTGAACCGTCTGCAGAACGAAGTT
 L N F F E G K G I L A A D E S V G I M G N R A I D A E A E V A S L N R L Q N E V

Glycogen phosphorylase

Aldolase B

Tropomyosin B

GAAGA CCTGA TGGTT GATGT TGAAC GCCTG GTTTC TTGGT ACGAC AACGAG TTCGG TTA CT CCAAC CGTGC TCTGG AATCT CCGGA ACGCC CGTTC CTG GCG ATCCT GGTG CCGCT AAA
 E D L M V D V E R L V S W Y D N E F G Y S N R A L E S P E R P F L A I L G G A K

Adult myosin HC

GAPDH

Phosphoglycerate kinase

CAGGT AGTGG ACTCT GCCTA CGAAG TTATC AAAGC AGCTG TACCG AGCGGC CCGTC TACCG GTATT TACGA AGCAC TGGA A CTGCG TCTGC TGCCG TCTGAA TCTGC GCTGC TGCCG GCA
 C V V D S A Y E V I K A A V P S G A S T G I Y E A L E L R L L P S E S A L L P A

Lactate dehydrogenase A

Alpha enolase

Actin polymerisation inhibitor

CCGGG TTCTC CGTAC GGCCG CTTTG GTGTG GAACA GAACG TTGAT ATGGTG TTCGC GTCTT TCATC CGTGG TACTG GTGGC GTTGA TACCG CTGCA GTAAGT GCGGT ATTCTG ACATC AGC
 F G S P Y G R F G V E Q N V D M V F A S F I R G T G G V D T A A V G A V F D I S

Pyruvate kinase

Creatinine kinase

AACGC AGACC GTGCG GGTAA AGTTA TCTGC TCTGC GGAAG GATCC AAGCTT GCGGC CGCAC TCGAG CACCA CCACC ACCAC CACTG A
 N A D R A G K V I C S A E G S K L A A A L E H H H H H H *

Quantification

Purification