

# Development and Evaluation of QconCAT Constructs for Absolute Quantification of Human Plasma Glycoproteins

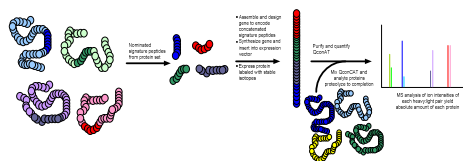
Hamid Mirzaei, Josh McBee, Simon Letarte, Julian Watts, Ruedi Aebersold  
Institute for Systems Biology, Seattle, WA

## Overview

- Establishing a framework for absolute quantification using heavy isotope coded peptides
- Biomarker validation
- Target peptides are concatenated and expressed in E.coli
- The construct is digested and mixed with target digest
- Quantification is done by measuring area under the peak of peptides in MS and MRM

## Introduction

Up and down regulation of proteins is one of indicative changes that occur as a result of disease progression. Measuring these changes in protein concentration is of great interest for biomarker discovery. Relative quantification of proteins (diseased vs. "normal") measures fold-change while absolute quantification determines protein abundance in terms of tangible concentrations. QconCATs are concatenated constructs (concatamers) of pure peptides with a sequence unique to the protein of interest. They are genetically engineered and expressed in E-coli grown in heavy isotope culture media. Tryptic digest of heavy QconCATs in known concentrations is used to absolutely quantify endogenous peptides (1). Here we report purification of QconCAT constructs (expressed by Entellecton Co.) designed to produce peptides that can be used for quantification of deglycosylated peptides isolated from human plasma.



## List of candidate peptides

- Leigh Anderson list of potential cancer biomarkers was used as the initial pool (~1500 proteins) (2)
- This list was refined by finding the common proteins between this list and Unipep database (150 proteins) (3)
- Since multiplexing requires knowledge of proteins concentration, only proteins of known concentrations were considered (40 proteins)
- 8 proteins were selected from the list of 40 proteins based on Peptide Atlas data
- Two tryptic peptides from each protein was selected (one with N-glyco motif and one without)

## List of selected proteins and peptides

Protein Name	pg/ml in normal	Peptide sequence 1 (ex-glyco)	Monoisotopic	Peptide sequence 2 (tryptic)	Monoisotopic
Transferrin	2.95E+08	ALGISPFHEAEVFTADDSGPR	2451.17	TSESGELHGLTTEEFVEGIYK	2454.13
Ceruloplasmin	2.80E+08	EHEGAIYPDITDFQR	1892.80	TTIEKPWLVGLGPIK	1911.11
Clusterin	1.01E+08	LADLTQGEDQYYLR	1683.79	LFDSDPITVTPVEVSR	1872.97
Paraoxonase 1	5.93E+07	HADWLTPLK	1180.61	YVYIAELLAHK	1318.71
Vitronectin	3.40E+05	DGSLFAFR	911.43	FEDGVLDPDYPYR	1421.63
Mast/stem cell growth factor receptor	3.33E+05	SEDESINR	948.40	YVSELHLTR	1116.58
ICAM-1	2.14E+05	LDPTVYGDSDFSAK	1614.73	ASVSVTAEDGQTR	1448.66
Thrombospondin 1	2.00E+05	VVDSTTGPGEHLR	1366.67	IPESGGDMSVDFIDELTGAAR	2194.04

Table 1. List of proteins that were selected for absolute quantification. These proteins vary 1000 fold in serum concentration. One peptide with N-glycosylation motif and one regular tryptic peptide were selected from each protein. These peptides were concatenated and expressed in E-Coli by Entellecton Co.

## Purification and digestion of expressed QconCAT

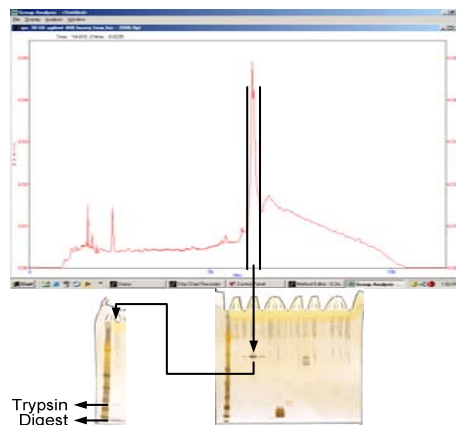


Figure 1. RPC purification of expressed QconCAT using a C5 Reversed phase column and a Vision (Bioac) HPLC instrument. A 60 min gradient was used and fractions were collected manually. Collected fractions were then analyzed by gel electrophoresis. A fraction containing pure construct was then re-suspended in digestion buffer. To avoid complications caused by buffer exchange the concentration was measured in digestion buffer. Fraction was then digested by trypsin. Fraction was then digested by trypsin. Fraction was then digested by trypsin.

## LC/MS analysis of the digest

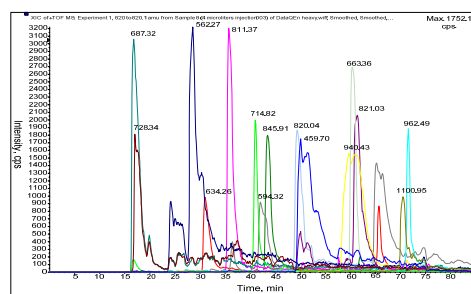


Figure 2. Extracted ion chromatogram of the QconCAT digest. QconCAT digest was subjected to nanoLC/MSMS analysis using Agilent 1100 HPLC interfaced with a PE-Sciex QSTAR mass spectrometer. All peptides except one were found

## Testing digestion reproducibility

#	Peptide	Observed ratio after 1:1 mix	Averaged ratio (MS only)	MRM ratios
1	ALGISPFHEAEVFTADDSGPR	1.1234	1.06	1.19
2	EHEGAIYPDITDFQR	1.14119	1.21	1.02
3	LADLTQGEDQYYLR	1.01321	1.01321	1.07
4	HADWLTPLK	1.1631	1.1631	2.99
5	DGSLFAFR	1.27423	1.27423	1.20
6	LDPTVYGDSDFSAK	1.20779	1.20779	1.18
7	VVDSTTGPGEHLR	1.20867	1.28	1.10
8	TSESGELHGLTTEEFVEGIYK	1.12749	1.1395	1.56
9	TTIEKPWLVGLGPIK	2.1141	2.1141	2.3
10	LFDSDPITVTPVEVSR	0.69	0.69	0.72
11	YVYIAELLAHK	0.45	0.45	0.32
12	FEDGVLDPDYPYR	0.7521	0.7521	0.5
13	YVSELHLTR	1.11043	1.11043	0.98
14	ASVSVTAEDGQTR	1.10712	1.1123	1.33
15	IPESGGDMSVDFIDELTGAAR	1.13521	1.13521	1.21

Table 2. One of crucial steps in application of QconCATs is the digestion. To test the digestion reproducibility light and heavy forms of the QconCAT were digested separately and then mixed in 1:1 ratio. Both LC/MS(QTSAR) and LC/MRM (4000QTRAP) were used to measure the ratios. In some cases variation in digestion was observed.

## Absolute quantification of target proteins

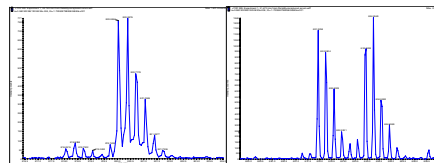


Figure 3. Tryptic digest of the heavy isotope coded QconCAT was mixed with glycoprotein serum. Peptide concentration was measured via MS ratio using the formula below

$$\left[ \frac{\text{Qconpeptide} \times \text{MW}}{\text{QconMW}} \times \text{QconConc} \right] \times 0.25 \times \left( \frac{1}{H/L \text{ ratio}} \right) \times \left( \frac{8}{5} \right) = \text{serumpeptide} \times \text{Conc}$$

## Results

Glycoprotein	Deglyc. peptide	Deglyc. peptide Conc. (pg/mL)	Tryptic peptide	Tryptic peptide Conc. (pg/mL)
Transferrin	ALGISPFHEAEVFTADDSGPR	0.041391891	TSESGELHGLTTEEFVEGIYK	0.039793083
Ceruloplasmin	EHEGAIYPDITDFQR	0.318991501	TTIEKPWLVGLGPIK	0.029277233
Clusterin	LADLTQGEDQYYLR	0.267511808	LFDSDPITVTPVEVSR	0.05361341
Paraoxonase 1	HADWLTPLK	0.085012797	YVYIAELLAHK	0.021718498
Vitronectin	DGSLFAFR	0.2817512904	FEDGVLDPDYPYR	0.040057189
Mast/stem cell growth factor receptor			YVSELHLTR	0.018404642
ICAM-1	LDPTVYGDSDFSAK	0.039499034	ASVSVTAEDGQTR	0.050710628
Thrombospondin 1	VVDSTTGPGEHLR	0.052084804	IPESGGDMSVDFIDELTGAAR	0.052099341

Table 3. The concentration reported in this table are based on assumption that QconCAT digestion was complete without any miss-cleavages or PTMs.

## Absolute quantification of target proteins using MRM

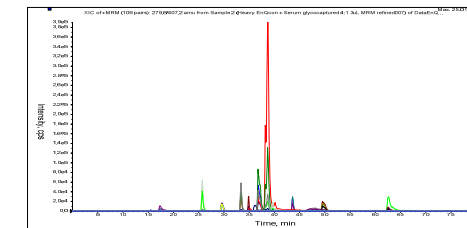


Figure 4. Large scale quantifications are done via MRM analysis of heavy and light peptides. In this experiment an MRM platform was tested for QconCAT method. Selected transitions for each peptide were monitored and the area under the peak was compared for quantification.

## Results

#	Peptide	T <sub>R</sub>	M/Z	Calculated Conc. (MS only)	Calculated Conc. MRM
1	ALGISPFHEAEVFTADDSGPR	51.406	(1229.5716)2+ (615.2863)4+ (820.0432)3+	0.0414	0.0395/0.060
2	EHEGAIYPDITDFQR	30.87	(863.3120)2+ (633.9396)3+	0.319	0.215/0.221
3	LADLTQGEDQYYLR	44.92	(845.9170)2+	0.267	0.321/0.326
4	HADWLTPLK	41.92	(594.3264)2+	0.0850	0.0590/0.059
5	DGSLFAFR	49.097	(459.7068)2+	0.0275	0.0765/0.077
6	LDPTVYGDSDFSAK	37.272	(811.3754)2+	0.039	0.079/0.079
7	VVDSTTGPGEHLR	16.212	(687.3299)2+ (458.2204)3+	0.052	0.017/0.23
8	TSESGELHGLTTEEFVEGIYK	60.51	(1231.0584)2+ (621.0303)3+	0.038	0.046/0.051
9	TTIEKPWLVGLGPIK	71.969	(962.4979)2+ (642.0538)3+	0.029	0.025/0.027
10	LFDSDPITVTPVEVSR	63.399	(940.4325)2+	0.053	0.070/0.073
11	YVYIAELLAHK	59.892	(663.3696)2+	0.022	0.022/0.022
12	FEDGVLDPDYPYR	42.751	(714.8234)2+	0.040	0.016/0.016
13	YVSELHLTR	28.041	(862.2765)2+	0.018	0.100/0.100
14	ASVSVTAEDGQTR	18.786	(485.8442)3+ (728.3410)2+	0.050	0.23/0.30
15	IPESGGDMSVDFIDELTGAAR	72.16	(1100.9548)2+	0.052	0.0450/0.045

Table 4. In most cases results are compatible between the two methods (MS and MRM). Some differences are observed.

## Conclusion

- QconCAT construct was expressed successfully
- Digestion was complete in most cases and fairly reproducible (other digestion protocols need to be explored for optimum digestion condition).
- Absolute concentration of target proteins was determined successfully.
- Software are required to design such constructs from MS data such as data accumulated in Peptide Atlas database.

## Acknowledgments

This work was supported in part with federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, under contract N01-HV-28179 (to R.A.) and with federal funds from the National Cancer Institute, National Institutes of Health, under contract No. N01-CO-12400 (to J.W.).

## References

- Pratt JM, Simpson DM, Doherty MK, Rivers J, Gaskell SJ, Beynon RJ Nat Protoc 2006;1(2):1029-43
- Malu Polanski and N. Leigh Anderson, Biomarker Insights, 2006:2
- Zhang, H., Loriaux, P., Eng, J., Campbell, D., Keller, A., Moss, P., Bonneau, R., Zhang, N., Zhou, Y., Wollschlaedl, B., Cooke, K., Yi, E.C., Lee, H., Peskind, E.R., Zhang, J., Smith, R.D. and Aebersold, R. 2006; Genome Biol., 7, R73.