Herein we present the data necessary for generation of alternative means to produce equimolar mixtures of peptides ("Design and Expression of a QconCAT Protein to Validate Hi3 Protein Quantification of Influenza Vaccine Antigens" [1]), such as QConCAT ("Trends in QconCATs for targeted proteomics" [2], "Natural flanking sequences for peptides included in a quantification concatamer internal standard" [3]) and SpikeTides versus the label free Hi3 approach. The experimental design and the interpretation of results are discussed in the original article [1].

© 2016 Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Experimental factors

The samples were digested with a higher than usual ratios of trypsin/substrate, source voltages optimized to reduce in-source fragmentation.

Experimental features

The peptides were analyzed on a using a reversed phase (BEH130 C18) column with data independent MSMS then data analysis using PLGS 3.0 Waters Ltd.

Data accessibility

Data is within this article

Value of the data

- Excellent reproducibility of peptide concentration data produced by the analysis of tryptic digestions is the necessary for method comparisons. The costs associated with a particular approach in terms of reagents and time may be a pivotal factor.
- The Hi3 method is a very valuable approach in cases of relatively simple proteomes. The speed of method development for core laboratories can be critical. The Hi3 method is relatively straightforward, once reproducible results are obtained.
- The QConCAT design is very elegant but there are critical issues, such as the protein construct solubility and stability, which require careful attention to obtain optimal results. The database attached is necessary for analysis of the mass spectral results.

1. Data

The data are compiled in the bar graphs for the quantitative proteome analyses of trivalent influenza vaccines and standards using three methods: Hi3, QconCAT and synthetic peptides.

Databases contained either the customized QconCAT protein sequence or the full sequences of the proteins represented therein, along with trypsin, several human keratins (cRAP from http://www.thegpm.org/crap/) and, since the viruses are grown in chicken eggs, the entire chicken (Gallus gallus) proteome (ftp://ftp.ensembl.org/pub/current_fasta/gallus_gallus.pep/) as well as selected full length influenza proteins (mostly from GISAID http://platform.gisaid.org/epi3/frontend#4f5b25, and the World Health Organization http://www.who.int/influenza/vaccines/virus/en/).

2. Experimental design, materials and methods

Data was obtained by the comparison of mass spectral signal from the three most intense fully tryptic peptides identified from the samples versus the internal standard protein. Experiments were designed to compare the mass spectral signal strengths from equimolar tryptic peptides identified by the Hi3 method by five different QconCAT designs as well as from synthetic peptides. The samples analyzed were commercial trivalent influenza vaccines as well as monovalent influenza reference standards. The mass spectra were obtained by reversed phase separation using a C18 UPLC column, Waters nanoAcquity UPLC, directly coupled to a Waters Synapt HDMS mass spectrometer. The mass spectrometer was programmed to carry out data-independent MSMS and incorporated a lock spray of glu-fibrinopeptide. The data was processed using Protein Lynx Global Server 3.0 for the identification of the three most intense peptides from a given protein identified in a custom database which has been attached. The intensities from all charge states of a peptide are included in the software calculations; however, the intensities resulting from in-source fragmentation and modified peptides must be added manually.
2.1. Codon optimized gene sequences

2.1.1. QconCAT 1

```
CATATGGATGACGATGATAAACTGGTGAGACGATGATAAACTGGTGAACGAACTGACCGAATTCGCGAAACTGGGCGAATACGGCTTCCA-
GAATGCACTGATCGTTCGTCATCTGGTTGACGAACCACAAAACCTGATTAAAGACCGATCGAGGATTTCAAAACTGCCAATCGTGTGGGTG-
CATGAAGGGCAGCGCTGACTTGAACGGGAAATCTTGAGCTGATCGAAGGGCAGCGCTGACTTGAACGGGAAATCTTGAGCTGATCGAAG
```

2.1.2. QconCAT 2

```
CATATGGACGATGACGATAAACTGGTGAGACGATGATAAACTGGTGAACGAACTGACCGAATTCGCGAAACTGGGCGAATACGGCTTCCA-
GAATGCACTGATCGTTCGTCATCTGGTTGACGAACCACAAAACCTGATTAAAGACCGATCGAGGATTTCAAAACTGCCAATCGTGTGGGTG-
CATGAAGGGCAGCGCTGACTTGAACGGGAAATCTTGAGCTGATCGAAGGGCAGCGCTGACTTGAACGGGAAATCTTGAGCTGATCGAAG
```

D.G.S. Smith et al. / Data in Brief 9 (2016) 169–176
2.1.3. QconCAT 3 and 4
CATATGGATGACGACGAT
CAAAGCAAGCGGTAAACTGGTGAACGAGCTGACTGAGTTCGCTAAAGCGAGCGGTAAGCTGGGTGAATACGGTTTTCCAGAACGCACTGATCGTACGTGCTTCTGGTAAACACCTGGTCGACGAGCCTCAAAACCTGATTAAAGCTAAGCGGCAAAGACGCCTTCCTGGGTAGCTTCCTGTACGAGTACAGCCGTGCTTCTGGCAAAGTAGTAGGTCTGAGCACTCTGCCAGAAATCTACGAGAAAGCGTCTGGTAAGCTGCCTCTGGTCGGTGGTCATGAAGGTGCTGGTGTAAGTAGTAGGTATGGGTGAGAACGTGAAAGCATCCGGTAAAAGCATCAGCATCGTCGGTTCCTACGTCGGTAACC
GTGCCAAGGCTTAAACTGGATGACGACGAT
CAAAGCAAGCGGTAAACTGGTGAACGAGCTGACTGAGTTCGCTAAAGCGAGCGGTAAGCTGGGTGAATACGGTTTTCCAGAACGCACTGATCGTACGTGCTTCTGGTAAACACCTGGTCGACGAGCCTCAAAACCTGATTAAAGCTAAGCGGCAAAGACGCCTTCCTGGGTAGCTTCCTGTACGAGTACAGCCGTGCTTCTGGCAAAGTAGTAGGTCTGAGCACTCTGCCAGAAATCTACGAGAAAGCGTCTGGTAAGCTGCCTCTGGTCGGTGGTCATGAAGGTGCTGGTGTAAGTAGTAGGTATGGGTGAGAACGTGAAAGCATCCGGTAAAAGCATCAGCATCGTCGGTTCCTACGTCGGTAACC

2.1.4. QconCAT 5
CATATGGCTGGTCGTGCGTCTGGTAAACTGGGTGAGTACGGTTTTCCAGAACGCGCTGATCGTACGTGCGTCTGGTAAAGTAGTCGGTCTGTCTACCCTCCCGGAGATCTACGAAAAAGCGTCCGGTAAAGAGGTCCTGGTTCTGTGGGTATTCACCACCCGTCTACTTCTGCAGATCAGCAGTCTCTGTACCAGAACGCAGACGCTTACGTATTCGTTGGCTCTTCTCGTGCGTCTGGCAAAATCGACCTGTGGTCTTACAACGCCGAACTGCTGGTTGCCCTGGAAAACCAGCAGACTATCGACCTGACCGACTCCGAAATGAACAAAGCGTCCGGTAAACTGTCTGGTGCGATGGACGAACTGCAACAGAAATCCTGGAACTGGACGAGAAAGCCAGCGGTAAAACCTTCTTCCTGACTCAGGGTGCGCTGCTGAACGACAAAGCTTCTGGCAAAACCCTGCTGATGAACGAACTGGGTGTTCCGTTTCACCTCGGTACCAAAGCGTCTGGTAAAGGTGTTACCCTGCTGCTGCCGGAACCGGAATGGACTTATCCACGTGCCTCTGGTAAAGGTGGTCTGGAACCGATCAACTTTCAGACGGCCGCAGATCAGGCACGTGCTTCTGGTAAAATCTCTCAGGCTGTTCACGCCGCGCAACGAAATCAACGAAGCAGGTCGTGCTTCTGGCAAACTGAACGTTGAAACCGACACCGCGGAAATCCGTGCCGTCTGGTAAACTGGTTGACAGCGTTGTTTCTTGGTCCAAAGCGTCCGGTAAATACAACGGTATCATCACCGACACCATCAAAGCCTCTGGTAAATTCACCTCCTCTGCCAACGGTGTTACGACCCACTACGTATCTCAGATCGGTGGTTTCGGGATCAGCGAAGACGGTGGTCTGCCGCAGTCTGGTCGTGCTTCTGGTAAATCTACCCAGGCGGCGATTGACCAGATCAACGGTAAAGCGTCCGGCAAATCTACGCAGAACGCGATCGACGAGATCACCAACAAAGCCTCTGG

D.G.S. Smith et al. / Data in Brief 9 (2016) 169–176
AGCAGGCGGTTATCCGAACATCGGTTTTCGTCCACGTGCTTCTGGCAAACTGAACTGGCTGACCCACCTGAAC
TTCAAAGCTCCCGGTAAAATGAACACCCAGTTCACGGCGGTTGGTAAAGCGTCTGGTAAAGCCAACGAACTGC
TGATCAACGTGAAAGCCTCCGGTAAAACACCTCGTAGACGAACCGCAGAACCTGATCAAAGCGTCTGGCAA
ACGTACTCCACGGCTCTTCTGTTGACTCTCAGACCGCTATGGTTCTGGTCAACGCGATCGTATTCAAACGCA
GGTAAAGGTAACTCTGCCCCGCTGATCATTCGTGCCTCTGGTAAAGGTTGGGCTTTTGACGACGGTAACGACGT
TTGGATGGGTCGTGCGTCTGGTAAAGGCGACGTATTCGTTATCCGTGCGTCTGGTAAAGCGGACACCATCTCT
CTCAGATCGAACCCTGCTGCTGTCTAACGAGGGTATCATCAACTCCGAAGACGAGCACCTGCTGGCTCTGC
GAACGTGCCTCTGGTAAACTGGCAGCGGCCCTGGAACACCACCACCACCACCACCACCTAATAGGGATCC.

2.2. Peptide Sequences

2.2.1. QconCAT 1-4

| Protein | Peptides (*sequence omitted in QconCAT 2) |
|---------|------------------------------------------|
| BSA – Bovine Serum Albumin | LVNELTEFAK LGEYGFQNALIVR HLVDEPQLNLK DAFLGSLVEYSR* |
| ADH – Alcohol Dehydrogenase (*Saccharomyces cerevisiae*) | VVGLSTLPEYIEK LPLVGGHEAGAGVVMGENVK SISIVGSVGNR ANELLINVK* |
| H1 – Hemagglutinin (H1N1) | STQNAIDEITNK MNYYWTILVEPKD EQLSVSFFER MNTQFTAVGK* |
| N1 – Neuraminidase A/California (H1N1) | TFLTQGALLNDK YNGIITDTIK YNGGVWGR GDVFVIR* |
| H3 – Hemagglutinin A/Victoria (H3N2) | STQFAQIDINGK IDLWSYNAELLVALENQHTIDLTDSEMK EFSEVEGR WDLFVER* |
| N2 – Neuraminidase A/Victoria (H3N2) | TLLMNELGVPFHGLTK LVDSVVSWSK VIEGWSNPK ILFIEEGK* |
| HB – Hemagglutinin B/Brisbane | LSGAMDELHNEILELDEK LSTHNVINAENAPGGPYK IVVDMVQK NLNSLSELEVK* |
| NB – Neuraminidase B/Brisbane | GVTLLLPPEWTPYPR LNVETDTAER YGEAYTDTHSYANK EWTYIGVDGDPNNALK* |
| Oval – Ovalbumin (*Gallus gallus*) | GGLPIFQTAADQAR ISQAVHAAHAEEINEAGR LTEWTSSNVMEER NVLQPSVSDQATAMLVNAIVFK* |
### 2.2.2. QconCAT 5 and SpikeTides

| Protein                          | Peptides *(sequence unique to QconCAT 5) (*sequences obtained as SpikeTides) |
|---------------------------------|---------------------------------------------------------------------------------|
| BSA – Bovine Serum Albumin      | LGEYGFQNALIVR*                                                                 |
|                                 | LVNELTEFAK*                                                                   |
|                                 | DAFLSFLYEYSR*                                                                  |
|                                 | HLVDEPQNLIK                                                                   |
| ADH – Alcohol Dehydrogenase     | VVGLSTLPEIYEK                                                                  |
| (Saccharomyces cerevisiae)       | LPLVGGHEGAGVVGMGENVK                                                             |
|                                 | SISIVGSYVGNR                                                                  |
|                                 | ANELLINVK                                                                     |
| H1 – Hemagglutinin (H1N1)       | EVLVLLWGIHHPSTSDQQSLYQNADAYVFVGVSSR*                                           |
|                                 | STQNAIDEITNK*                                                                  |
|                                 | MNYYWTLVEPGDK*                                                                 |
|                                 | MNTQFTAVGK                                                                    |
| N1 – Neuraminidase A/California  | TFFLTQGALLNDK                                                                  |
| (H1N1)                          | YNGIITDTIK                                                                    |
|                                 | YGNGVWIGR                                                                     |
|                                 | GDVFVIR                                                                        |
| H3 – Hemagglutinin A/Victoria   | IDLWSYNAELVALLENQHTIDLDSEMNK*                                                  |
| (H3N2)                          | STQAACDQINGK*                                                                  |
|                                 | SQQAVIPNGFRPR*                                                                 |
|                                 | LNWLTHLNFK                                                                    |
| N2 – Neuraminidase A/Victoria   | TLLMNELGVFPHLGTK                                                                |
| (H3N2)                          | LVDSVVSWSK                                                                    |
|                                 | SGYSGIFSVEGK                                                                  |
|                                 | GWAFDDGNDVWMGR                                                                 |
| HB – Hemagglutinin B/Brisbane   | LSGAMDELEHNEILELEDKE*                                                          |
|                                 | FTSSANGVTTHVFSQIGGFPDQTEDDGLPQSGR*                                            |
|                                 | NLNSLESEVK*                                                                   |
|                                 | ADTISSQIELAVLSNEGIINSEDEHLLALER                                               |
| NB – Neuraminidase B/Brisbane   | GVTLLLPEPEWTPYR                                                                 |
|                                 | LNVETDTAERIR                                                                  |
|                                 | YGEAYTDHYSYAK                                                                 |
|                                 | GNSAPLIR                                                                       |
| Oval – Ovalbumin (Gallus gallus)| GGLEPINFQTAADQQR                                                                |
|                                 | ISQAVHAAHAEEINEAGR                                                             |
|                                 | LTEWTSSNVMEER                                                                  |
|                                 | NVLQPSSVDSQTMVLMNAIVFK                                                           |
QconCAT final sequence where using the spacer peptide ASGK:

MAGRA ASGKBSA-1 ASGKADH-1 ASGKH3-1 ASGKH1-1 ASGKHB-1 ASGKN1-1 ASGKN2-1 ASGKNB-1 ASGKOV-1 ASGKOV-2 ASGK NB-2 ASGKN2-2 ASGKN1-2 ASGKH2-2 ASGKHB-2 ASGKH1-2 ASGKADH-2 ASGKBSA-2 ASGKHB-3 ASGKN1-3 ASGKN2-3 ASGKNB-3 ASGKOV-3 ASGKBSA-3 ASGKADH-3 ASGKH1-3 ASGKH3-3 ASGKH3-4 ASGKH1-4 ASGKADH-4 ASGKBSA-4 ASGKOV-4 ASGKNB-4 ASGKN2-4 ASGKN1-4 ASGKHB-4 ASGKAAALEHHHHHH

Hemagglutinin (μg/ml) in trivalent influenza vaccine by quantitative method

Ratio of average "Hi3" values from all proteins represented in QconCat relative to BSA standard.

FASTA Protein Database used: 2015_2016_FluQuant.fasta
Acknowledgments

The authors would like to thank Marybeth Creskey and Lisa Walrond for valuable contributions.

Transparency document. Supplementary material

Transparency data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2016.08.035.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2016.08.035.

References

[1] D.G.S. Smith, G. Gingras, Y. Aubin, T.D. Cyr, Design and expression of a QconCAT protein to validate Hi3 protein quantification of influenza vaccine antigens, J. Proteom. 146 (2016) 133–140.
[2] J. Chen, I.V. Turko, Trends in QconCATs for targeted proteomics, TrAC Trends Anal. Chem. 57 (2014) 1–5.
[3] C.S. Cheung, K.W. Anderson, M. Wang, I.V. Turko, Natural flanking sequences for peptides included in a quantification concatamer internal standard, Anal. Chem. 87 (2) (2015) 1097–1102.