Data Article

Peptide data on the disulfide bond analysis of baculovirus produced Pfs25 by LC-MSMS

Shwu-Maan Lee *, Jordan L. Plieskatt, C. Richter King

PATH Malaria Vaccine Initiative (MVI), 455 Massachusetts Avenue NW, Suite 1000, Washington, DC 20001-2621, USA

A R T I C L E   I N F O

Article history:
Received 17 November 2017
Received in revised form
29 January 2018
Accepted 6 March 2018
Available online 12 March 2018

Keywords:
Pfs25
Disulfide
Mass spectrometry
Malaria
LC-MSMS

A B S T R A C T

This article contains the peptide data obtained while performing disulfide bond mapping of the recombinant Plasmodium falciparum protein, Pfs25, produced from the baculovirus expression system. Pfs25 is a malaria transmission-blocking vaccine candidate, with a compact and complex structure including 22 cysteines. This supplementary data is related to the research “Disulfide bond mapping of Pfs25, a recombinant malaria transmission blocking vaccine candidate” (Lee et al., 2018) [1]. In brief, Pfs25 was digested with trypsin/Lys-C and derived peptides separated by High Performance Liquid Chromatography (HPLC) and analyzed by mass spectrometry (MS) by MS E fragmentation. The theoretical peptides and their respective masses along with disulfide bond locations with linked peptides are presented here alongside the mass spectrometry analysis. The raw mass spectrometry data is made available through the Mass Spectrometry Interactive Virtual Environment (MassIVE) with identifier: MSV000081982.

© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Specifications Table

| Subject area       | Chemistry, Biology. |
|--------------------|---------------------|
| More specific subject area | Disulfide bond analysis by liquid chromatography and mass spectrometric analysis |
| Type of data       | Tables, figures     |
| How data was acquired | Data was generated using liquid chromatography (Waters 2695 Separations Module and Waters 2489 UV/Vis Detector) and mass spectroscopy (Waters QTOF Premier mass spectrometer) |
| Data format        | Collated data from analysis with Waters BiopharmaLynx 1.3 and MassLynx |
| Experimental factors | Recombinant Pfs25 digested with 20 µg of trypsin/Lys-C at 37 °C overnight and subsequent further digestion by additional 20 µg of trypsin/Lys-C for 3–4 hours at 37 °C |
| Experimental features | Identification of the proper pairing of 11 disulfide bonds in Pfs25 through digestion of peptides and LC-MS/MS |
| Data source location | Mass spectrometry data acquired in Middleton, WI, USA |
| Data accessibility | Data is provided within this article and RAW MS files have been deposited in the Mass Spectrometry Interactive Virtual Environment (MassIVE) with identifier: MSV000081982 (ftp://massive.ucsd.edu/MSV000081982). MassIVE is a member of the ProteomeXchange Consortium |

Value of the data

- The derived peptides and mass spectrometry data is provided here for further details from the disulfide bond analysis of Pfs25.
- The disulfide bond locations and linked peptides are discussed alongside the mass spectrometry analysis here and data made accessible to the scientific community.
- Pfs25 is a compact and complex 17.9 kDa protein with 22 cysteines (11 disulfide bonds) that has presented difficulty in prior disulfide bond analysis
- A method was developed to map the disulfide bonds of a complex and compact protein, which may be applicable to other proteins, an important step in recombinant protein development for vaccines.

1. Data

The Pfs25 disulfide bond mapping peptides are discussed in further detail in this manuscript to further support the elucidation of disulfide bonds of Pfs25 as discussed in [1]. Further, the mass spectrometry RAW files have been deposited in the Mass Spectrometry Interactive Virtual Environment (MassIVE). Theoretical peptides, produced from Trypsin/Lys-C digestion of Pfs25, are presented in Table 1.

Utilizing BiopharmaLynx 1.3 the mass spectral data was analyzed and compared to the theoretical peptides to obtain the localization of the 11 disulfide bonds present in the recombinant Pfs25. The disulfide bond locations and linked peptides (including theoretical and observed) masses are presented in Table 2. Each disulfide bond (referenced by nomenclature SS#) is further presented with the peptide information and mass spectrometry (MS) and MSMS data obtained during the analysis in the subsequent figures and tables presented in this manuscript.
2. Disulfide bond SS1

A total of 30 fragments were observed, with 20 fragment ions of this peptide consistent with the linkage of Cys10 and Cys24. The remaining ten fragment ions were consistent with constituent peptides (Table 3, Fig. 1).

3. Disulfide bond SS2

A total of 39 fragment ions of this peptide were observed with 36 fragment ions consistent with the linkage of Cys26 and Cys38. Three additional fragments were consistent with constituent peptides (Table 4, Fig. 2).

Table 1
Theoretical Fragments for Trypsin/Lys-C Digestion of Pfs25.

| Peptides |  | Peptide Label | Theoretical Mass (Da) |
|----------|---------------------------------|----------------------|
| DAK      | Position 1–3                     | T1                   | 332.17                |
| VTVDTVK  | Position 4–11                    | T2                   | 863.44                |
| R        | Position 12–12                   | T3                   | 174.11                |
| GIFLQMSGHLECK | Position 13–25             | T4                   | 1461.71               |
| CENDIVLNEETCEEK | Position 26–41             | T5                   | 1865.80               |
| VLK      | Position 42–44                   | T6                   | 358.26                |
| CDEK     | Position 45–48                   | T7                   | 493.18                |
| TVNKPGDFS | Position 49–59                  | T8                   | 1194.57               |
| CIK      | Position 60–62                   | T9                   | 362.20                |
| IDGNPVSYACK  | Position 63–73                | T10                  | 1165.54               |
| CNLGYDMVNVNCIPNECK | Position 74–91 | T11                  | 2027.86               |
| QVTCGNGK | Position 92–99                   | T12                  | 805.38                |
| CILDTSNPV | Position 100–109                | T13                  | 1088.55               |
| TGVCSCNIGK | Position 110–119                | T14                  | 980.44                |
| VPNVQDQNK | Position 120–128                | T15                  | 1040.53               |
| CSK      | Position 129–131                 | T16                  | 336.15                |
| DGETK    | Position 132–136                 | T17                  | 548.24                |
| CSLK     | Position 137–140                 | T18                  | 449.23                |
| CLK      | Position 141–143                 | T19                  | 362.20                |
| ENETCK   | Position 144–149                 | T20                  | 722.29                |
| AVDGIYK  | Position 150–156                 | T21                  | 764.41                |
| CDCK     | Position 157–160                 | T22                  | 467.15                |
| DGFIHQESSICTTHHHHHH | Position 161–179 | T23                  | 2248.98               |

Table 2
Disulfide bond locations for Pfs25 including theoretical and observed masses of fragments.

| Disulfide Bond | Linked Cysteines | Tryptic Peptide | Theoretical Mass (Da) | Observed Mass (Da) | Mass Error (ppm) |
|----------------|------------------|-----------------|-----------------------|-------------------|------------------|
| SS1            | Cys10-Cys24      | T2 + T4         | 2323.1375             | 2323.1152         | 9.6              |
| SS2            | Cys26-Cys38      | T5              | 1863.7866             | 1863.7752         | 6.1              |
| SS3            | Cys45-Cys56      | T7 + T9         | 853.3674              | 853.3656          | 2.1              |
| SS4            | Cys54-Cys72      | T8 + T10        | 2358.0670             | 2358.0799         | 5.5              |
| SS5            | Cys24-Cys85      | T11 + T13       | 3112.3796             | 3112.3534         | 8.4              |
| SS6            | Cys90-Cys100     | T12 + T14 + T16 | 2117.9326             | 2117.9182         | 6.8              |
| SS7            | Cys95-Cys113     | T18 + T20       | 1169.5056             | 1169.4970         | 7.3              |
| SS8            | Cys115-Cys129    | T19 + T22 + T23 | 3074.3003             | 3074.2795         | 6.8              |
4. Disulfide bond SS3

A total of nine fragments were observed, with six fragment ions of this peptide consistent with the linkage of Cys45 and Cys60. Three fragment ions were consistent with constituent peptides (Table 5, Fig. 3).

5. Disulfide bond SS4

A total of 68 fragments were observed, with 45 fragment ions of this peptide consistent with the linkage of T8 to T10 through Cys54 and Cys72. The remaining 23 fragments were consistent with constituent peptides (Table 6, Fig. 4).

6. Disulfide bonds SS5 and SS6

A total of 90 fragments were observed and four fragment ions (1/b12, 1/b13, 1/b15, and 1/b16) were consistent with an internal disulfide bond linkage between Cys24 and Cys30. Thirty-three fragment ions were consistent with the linkage of Cys90 and Cys100. An additional 44 fragments of this
Fig. 1. Disulfide bond SS1.
peptide were consistent with the combined linkages of Cys74 to Cys85 and Cys90 to Cys100 and nine fragments consistent with constituent peptides (Table 7, Fig. 5).

7. Disulfide bonds SS7 and SS8

A total of 39 fragments were observed. Three fragment ions (1/b7-2/b5, 1/y6-2/a4, and 1/y6-2/b4) were specific to the linkage between T12 and T14 and confirmed the Cys85 to Cys113 linkage. Four fragment ions (2/y5-3/b2, 2/y5-3/y3, 2/y6-3/a2, and 2/y6-3/b2) were specific to the linkage between T14 and T16 and confirmed the linkage of Cys115 to Cys129. A further 21 fragment ions were consistent with the linkage of T12, T14, and T16 and remaining 11 fragments consistent with constituent peptides (Table 8, Fig. 6).
8. Disulfide bond SS9

A total of 35 fragment ions were observed with 25 fragment ions of this peptide consistent with the linkage of T18 to T20 through Cys137 and Cys148. Ten fragments were consistent with constituent peptides (Table 9, Fig. 7).
Table 5
Disulfide bond SS3 (Cys45-Cys60) peptides.

| Assignment | Theoretical Mass (Da) | Observed Mass (Da) | Mass Error (Da) | Intensity (counts) | Identification |
|------------|-----------------------|--------------------|-----------------|--------------------|----------------|
| 9 Fragments |                       |                    |                 |                    |                |
| Const. Pep.3 |                      |                    |                 |                    |                |
| 1/y1       | 147.1133              | 147.115            | −0.0017         | 163                | K              |
| 1/y2       | 276.1559              | 276.158            | −0.0002         | 84                 | EK             |
| 1/y3       | 391.1829              | 391.2075           | −0.0246         | 47                 | DEK            |
| Cys45 and Cys60 |                |                    |                 |                    |                |
| 1/a1−2/a2  | 262.1048              | 262.0183           | 0.0865          | 22                 | C=Cl           |
| 1/b1−2/b3  | 421.0852              | 421.1291           | −0.0439         | 117                | C=CIK          |
| 1/b2−2/y3  | 579.2271              | 579.2261           | 0.001           | 67                 | CD=CIK         |
| 1/y4−2/a2  | 680.2748              | 680.2768           | −0.002          | 23                 | CDEK=CIK       |
| 1/y4−2/b2  | 708.2697              | 708.2832           | −0.0135         | 126                | CDEK=Cl        |
| 1/y4−2/y3  | 854.3752              | 854.3748           | 0.0004          | 3305               | CDEK=CIK       |

Fig. 3. Disulfide bond SS3.
Table 6

Disulfide bond SS4 (Cys54-Cys72) peptides.

| Assignment | Theoretical Mass (Da) | Observed Mass (Da) | Mass Error (Da) | Intensity (counts) | Identification |
|------------|-----------------------|--------------------|----------------|-------------------|----------------|
| 68 Fragments |                       |                    |                |                   |                |
| Constituent Peptides | 23 Fragments |                   |                |                   |                |
| 1/a2       | 173.129               | 173.129            | −0.0007        | 4507              | TV             |
| 1/a3       | 287.179               | 287.0969           | 0.075          | 897               | TVN            |
| 1/a4       | 415.2669              | 415.2224           | 0.0445         | 28                | TVN            |
| 1/b2       | 201.1239              | 201.1235           | 0.0004         | 2754              | TVN            |
| 1/b3       | 315.1668              | 315.0896           | 0.0722         | 51                | TVN            |
| 1/b4       | 443.2618              | 443.2626           | −0.0008        | 248               | TVN            |
| 1/y1       | 147.1133              | 147.0623           | 0.0124         | 1146              | TVN            |
| 1/y2       | 234.1454              | 234.1455           | −0.0001        | 700               | SK             |
| 1/y3       | 381.2138              | 381.2014           | 0.0124         | 1146              | TVN            |
| 1/y4       | 496.2407              | 496.2404           | 0.0003         | 153               | DFK            |
| 1/y5       | 553.2622              | 553.2604           | 0.018          | 564               | GDFSK          |
| 2/a3       | 258.1454              | 258.1545           | −0.0092        | 78                | TVN            |
| 2/a4       | 327.1883              | 327.1927           | −0.0044        | 300               | IDGN           |
| 2/a5       | 469.2411              | 469.2278           | 0.0132         | 76                | IDGN           |
| 2/a6       | 568.3094              | 568.3065           | 0.0029         | 91                | IDGNPV         |
| 2/a7       | 818.4048              | 818.3653           | 0.0396         | 129               | IDGNPVSY       |
| 2/b1       | 114.0919              | 114.041            | 0.0509         | 18                | I              |
| 2/b2       | 229.1188              | 229.119            | −0.0002        | 1962              | TVN            |
| 2/b3       | 286.1403              | 286.1415           | −0.0012        | 727               | IDN            |
| 2/b4       | 400.1832              | 400.1889           | −0.0057        | 1203              | TVN            |
| Cys54 and Cys72  | 45 Fragments |                   |                |                   |                |
| 1/a6-2/y5  | 1183.5604             | 1183.541           | 0.0198         | 1608              | TVN            |
| 1/a6-2/y7  | 1379.6815             | 1379.694           | 0.0073         | 143               | TVN            |
| 1/a8-2/y7  | 1531.73               | 1531.666           | 0.0636         | 108               | TVN            |
| 1/a9-2/b10 | 1951.8682             | 1951.823           | 0.0454         | 93                | TVN            |
| 1/b10-2/a10| 2038.9003             | 2038.993           | 0.0012         | 727               | TVN            |
| 1/b6-2/y2  | 497.236               | 497.2373           | −0.0013        | 252               | TVN            |
| 1/b6-2/y8  | 961.4599              | 961.4065           | 0.0534         | 124               | TVN            |
| 1/b8-2/y3  | 1521.1719             | 1521.674           | 0.0449         | 161               | TVN            |
| 1/b8-2/y7  | 1464.6979             | 1464.675           | 0.0228         | 47                | TVN            |
| 1/b8-2/y10 | 1865.8162             | 1865.827           | −0.0106        | 68                | TVN            |
| 1/b8-2/y11 | 1978.9003             | 1978.891           | 0.0092         | 579               | TVN            |
| 1/b8-2/y2  | 1062.4712             | 1062.441           | 0.0303         | 364               | TVN            |
| 1/b8-2/y6  | 1482.6721             | 1482.651           | 0.021          | 137               | TVN            |
| 1/b8-2/y7  | 1579.7249             | 1579.69            | 0.0347         | 456               | TVN            |
| 1/b9-2/y5  | 1530.6721             | 1530.667           | 0.0054         | 180               | TVN            |
| 1/b9-2/y7  | 1726.7932             | 1726.741           | 0.0527         | 66                | TVN            |
| 1/b9-2/y10 | 2246.0222             | 2246.022           | 0.0005         | 2740              | TVN            |
| 1/b9-2/y11 | 2359.1062             | 2359.109           | −0.0027        | 57983              | TVN            |
| 1/b10-2/a10| 1645.699              | 1645.775           | −0.0062        | 368               | TVN            |
| 1/b10-2/y11| 1819.7994             | 1819.795           | 0.0044         | 57                | TVN            |
| 1/b10-2/y4 | 1137.4708             | 1137.472           | −0.0007        | 738               | TVN            |
| 1/b10-2/y6 | 1420.624              | 1420.613           | 0.0114         | 404               | TVN            |
| 1/b10-2/y7 | 1770.7467             | 1770.722           | 0.0243         | 43                | TVN            |
| 1/b11-2/y1 | 1916.8523             | 1916.848           | 0.0339         | 3031              | TVN            |
| 1/b11-2/y2 | 1000.4232             | 1000.424           | −0.0004        | 627               | TVN            |
| 1/b11-2/y3 | 1071.4603             | 1071.461           | −0.0007        | 825               | TVN            |

S.-M. Lee et al. / Data in Brief 18 (2018) 209–233 217
9. Disulfide bonds SS10 and SS11

A total of 65 fragments were observed. Eleven fragment ions were specific to the linkage of T19 and T22, and confirmed the linkage of Cys141 to Cys157. Three fragment ions were specific to the linkage between peptides T22 and T23 and confirmed the linkage of Cys159 to Cys172. Thirty-two fragment ions were consistent with the linkage of T12, T14, and T16 and an additional 19 fragments were consistent with constituent peptides (Table 10, Fig. 8).

10. Experimental design, materials and methods

10.1. Sample preparation

Baculovirus Pfs25 [2] was denatured and digested as described in [1].

10.2. Chromatography

Digested peptides were separated with a 2695 Separations Module (Waters Corporation; Milford MA) and a 2489 UV/Vis Detector (Waters Corporation; Milford, MA) set at 214 nm. An XBridge (Waters Corporation; Milford, MA) BEH 300 C18 (2.1×250 mm, 5 µm) was used at a column temperature of 37 °C and gradient with 0.1% Trifluoroacetic acid (TFA) in purified water (Mobile Phase A) and 0.1% TFA in acetonitrile (Mobile Phase B) as described in [1].

10.3. Mass spectrometry

MS analysis was done with a QTOF Premier mass spectrometer (Waters Corporation; Milford, MA) equipped with an electrospray source as described in [1]. MS data was acquired in MS² mode using MassLynx v4.1 (Waters Corporation; Milford, MA). RAW MS files have been deposited in the Mass Spectrometry Interactive Virtual Environment (MassIVE) with identifier: MSV000081982.

10.4. Analysis of mass spectra

The mass spectral data was analyzed using BiopharmaLynx 1.3 (Waters Corporation; Milford, MA) as described in [1].
Fig. 4. Disulfide bond SS4.
Table 7
Disulfide bond SS5 (Cys74-Cys85) and SS6 (Cys90-Cys100) peptides.

| Assignment | Theoretical Mass (Da) | Observed Mass (Da) | Mass Error (Da) | Intensity (counts) | Identification |
|------------|-----------------------|--------------------|-----------------|-------------------|----------------|
| **90 Fragments** | | | | | |
| Cys74 and Cys85 4 Fragments | | | | | |
| 1/b12 | 1324.513 | 1324.516 | −0.0031 | 1267 | CNLGYDMVNNVC (Internal) |
| 1/b13 | 1437.597 | 1437.596 | 0.0006 | 3004 | CNLGYDMVNNVC1 (Internal) |
| 1/b15 | 1648.692 | 1648.71 | −0.0175 | 311 | CNLGYDMVNNVCIPN (Internal) |
| 1/b16 | 1777.735 | 1777.728 | 0.007 | 105 | CNLGYDMVNNVCIPNE (Internal) |
| **Constituent Peptides** | | | | | |
| 9 Fragments | | | | | |
| 1/y1 | 147.1133 | 147.1123 | 0.001 | 2902 | K |
| 2/y2 | 246.1818 | 246.1842 | −0.0024 | 789 | VK |
| 2/y3 | 343.2345 | 343.1509 | 0.0836 | 258 | PVK |
| 2/y4 | 457.2774 | 457.2771 | 0.0003 | 855 | NPVK |
| 2/y5 | 544.3095 | 544.3134 | −0.0039 | 711 | SNPVK |
| 2/y6 | 645.3572 | 645.3581 | −0.0009 | 3696 | TSNPVK |
| 2/y7 | 760.3841 | 760.3795 | 0.0046 | 2330 | DTSNPVK |
| 2/y8 | 873.4681 | 873.4631 | 0.0051 | 2651 | LDTSNPVK |
| 2/y9 | 986.5522 | 986.5494 | 0.0029 | 381 | ILDTSNPVK |
| **Cys90-Cys100 33 fragments** | | | | | |
| 1/a17-2/a7 | 2569.065 | 2569.078 | −0.0127 | 433 | CNLGYDMVNNVCIPNEC=CILDTSN |
| 1/y18-2/b2 | 2240.927 | 2240.983 | −0.0564 | 114 | CNLGYDMVNNVCIPNECK=CID |
| 1/y18-2/b4 | 2469.038 | 2469.068 | −0.03 | 1446 | CNLGYDMVNNVCIPNECK=CILDTS |
| 1/y18-2/b6 | 2657.118 | 2657.106 | 0.012 | 90 | CNLGYDMVNNVCIPNECK=CILDTS |
| 1/y2-2/a8 | 1063.492 | 1063.406 | 0.0859 | 167 | CK=CILDTSNP |
| 1/y2-2/b1 | 351.1161 | 351.204 | −0.0879 | 78 | CK=C |
| 1/y2-2/b2 | 464.2001 | 464.2095 | −0.0094 | 359 | CK=CIL |
| 1/y2-2/b3 | 577.2842 | 577.284 | 0.0002 | 142 | CK=CIL |
| 1/y2-2/b4 | 692.3112 | 692.2997 | 0.0115 | 322 | CK=CILDTS |
| 1/y2-2/b7 | 994.4338 | 994.4278 | 0.006 | 837 | CK=CILDTSN |
| 1/y2-2/y10 | 1336.661 | 1336.634 | 0.027 | 664 | CK=CILDTSNPVK |
| 1/y3-2/b1 | 480.1587 | 480.1505 | 0.0081 | 107 | ECK=C |
| 1/y3-2/b4 | 821.3538 | 821.3419 | 0.0118 | 460 | ECK=CILDTS |
| 1/y3-2/b5 | 922.4014 | 922.3771 | 0.0244 | 283 | ECK=CILDTSN |
| 1/y3-2/b8 | 1220.524 | 1220.519 | 0.0098 | 340 | ECK=CILDTSNP |
| 1/y3-2/b9 | 1319.598 | 1319.584 | 0.0138 | 987 | ECK=CILDTSNPV |
| 1/y4-2/a1 | 566.2067 | 566.2282 | −0.0215 | 232 | NECK=C |
| 1/y4-2/a2 | 679.2908 | 679.3111 | −0.0203 | 202 | NECK=C|
| 1/y4-2/a7 | 1209.524 | 1209.502 | 0.0221 | 927 | NECK=CILDTSN |
| 1/y4-2/b3 | 820.3697 | 820.3667 | 0.003 | 154 | NECK=CIL |
| 1/y4-2/b4 | 935.3967 | 935.3924 | 0.0043 | 481 | NECK=CILD |
| Fragment | m/z (Da) | Matched m/z (Da) | Mass Error (ppm) | Intensity (arb. units) | Comment |
|----------|---------|-----------------|-----------------|-----------------------|---------|
| Cys74 to Cys85 and Cys90 to Cys100 | 1732.59 | 1732.579 | 0.0105 | 789 | NVCIPNECK = C = CI |
| | 2594.151 | 2594.122 | 0.0295 | 85 | NVCIPNECK = CILDTSNPV = CNLGY |
| | 2003.83 | 2003.83 | 0.0873 | 148 | VNNVCIPNECK = CNL = CILD |
| | 2525.993 | 2525.993 | 0.084 | 186 | VNNVCIPNECK = CNLGY = CILDTS |
| | 2093.895 | 2093.895 | 0.0625 | 298 | MVNVNCIPNECK = CI = CNLGY |
| | 2492.059 | 2492.059 | 0.0166 | 490 | DMVNVNCIPNECK = C = CILDTSNPV |
| | 1992.847 | 1992.854 | -0.0063 | 826 | DMVNVNCIPNECK = CN = CIL |
| | 2190.948 | 2190.909 | 0.0388 | 313 | DMVNVNCIPNECK = CIL = CNL |
| | 2208.922 | 2208.922 | -0.0498 | 102 | DMVNVNCIPNECK = CILD = CN |
| | 2525.986 | 2525.986 | 0.0891 | 573 | YDMVNNVCIPNECK = C = CILDTSNPV |
| | 2882.266 | 2882.266 | 0.0254 | 1679 | YDMVNNVCIPNECK = CILDTSNPV = CNL |
| | 2684.165 | 2684.074 | 0.0908 | 230 | YDMVNNVCIPNECK = C = CILDTSNPV |
| | 2854.198 | 2854.166 | 0.0322 | 311 | YDMVNNVCIPNECK = CN = CILDTSNPV |
| | 1121.425 | 1121.524 | -0.0988 | 3609 | CIPNECK = C = CIL |

C.-M. Lee et al. / Data in Brief 18 (2018) 209–233
| Assignment                  | Theoretical Mass (Da) | Observed Mass (Da) | Mass Error (Da) | Intensity (counts) | Identification              |
|-----------------------------|-----------------------|--------------------|-----------------|--------------------|-----------------------------|
| 1/y7-2/a1-1/a2              | 1066.394              | 1066.391           | 0.0034          | 286                | CIPNECK = C = CN           |
| 1/y7-2/a1-1/a8              | 1744.699              | 1744.779           | −0.0803         | 1284               | CIPNECK = C = CNLGYDMV     |
| 1/y7-2/a2-1/a7              | 1758.715              | 1758.619           | 0.0957          | 156                | CIPNECK = CI = CNLGYDM     |
| 1/y7-2/a3-1/a9              | 2084.91               | 2084.817           | 0.0925          | 164                | CIPNECK = CIL = CNLGYDMVN  |
| 1/y7-2/a5-1/a3              | 1621.721              | 1621.642           | 0.0795          | 207                | CIPNECK = CILDT = CNL      |
| 1/y7-2/a9-1/a11             | 2911.292              | 2911.285           | 0.0068          | 205                | CIPNECK = CILDTSNPV = CNLGYDMVNN |
| 1/y7-2/a9-1/a8              | 2584.138              | 2584.076           | 0.062           | 990                | CIPNECK = CILDTSNPV = CNLGYDMV |
| 1/y7-2/b1-1/a3              | 1207.473              | 1207.5             | −0.027          | 508                | CIPNECK = C = CNL          |
| 1/y7-2/b1-1/a6              | 1542.585              | 1542.666           | −0.0806         | 928                | CIPNECK = C = CNLGYD       |
| 1/y7-2/b2-1/b4              | 1405.574              | 1405.526           | 0.0474          | 125                | CIPNECK = CI = CNLG        |
| 1/y7-2/b3-1/b4              | 1518.658              | 1518.643           | 0.015           | 197                | CIPNECK = CIL = CNLG       |
| 1/y7-2/b4-1/a10             | 2341.975              | 2341.988           | −0.0132         | 340                | CIPNECK = CILD = CNLGYDMVNN |
| 1/y7-2/b6-1/a10             | 2530.054              | 2529.985           | 0.0696          | 102                | CIPNECK = CILDTS = CNLGYDMVNN |
| 1/y7-2/b8-1/b3              | 1975.839              | 1975.88            | −0.0417         | 245                | CIPNECK = CILDTSNP = CNL   |
| 1/y8-1/a1-2/a5              | 1493.663              | 1493.629           | 0.0336          | 182                | VCIPNECK = C = CILD        |
| 1/y8-2/b9-1/a3              | 2145.981              | 2145.931           | 0.0496          | 164                | VCIPNECK = CILDTSNPV = CNL |
| 1/y9-1/b1-2/b5              | 1663.695              | 1663.662           | 0.0337          | 104                | NVCPNECK = C = CILD        |
| 1/y9-2/b1-1/a6              | 1755.696              | 1755.718           | −0.0212         | 658                | NVCPNECK = C = CNLGYD      |
| 1/y9-2/b1-1/b1              | 1221.453              | 1221.521           | −0.068          | 782                | NVCPNECK = C = C           |
|                             | 2496.067              | 2496.051           | 0.0159          | 634                | NVCPNECK = CILDTSNP = CNLGYD |
| Pair  | Score1 | Score2 | p | N   | Mutation |
|-------|--------|--------|---|-----|----------|
| 1/y9-2/b8-1| 2260.023 | 2259.99 | 0.0334 | 668 | NVCIPNECK=CILDTSNPV=CNL |
| 2/y10-1/y12-1/b2 | 2665.192 | 2665.097 | 0.0942 | 258 | CILDTSNPV=MVNVCIPNECK=CN |
| 2/y10-1/y7-1/b10 | 3014.319 | 3014.31 | 0.0088 | 1407 | CILDTSNPV=CIPNECK=CNLGYDMVNN |
| 2/y10-1/y7-1/b11 | 3113.388 | 3113.393 | −0.0056 | 120269 | CILDTSNPV=CIPNECK=CNLGYDMVNNV |
| 3/y10-1/y9-1/b6 | 2556.124 | 2556.095 | 0.0291 | 1309 | CILDTSNPV=CIPNECK=CNLGYD |
| 2/y10-1/y9-1/a4 | 2463.15 | 2463.093 | 0.0574 | 609 | CILDTSNPV=NVCIPNECK=CNLG |
| 2/y10-1/y9-1/a5 | 2626.214 | 2626.114 | 0.0999 | 654 | CILDTSNPV=NVCIPNECK=CNLGY |
| 2/y10-1/y9-1/b3 | 2434.124 | 2434.059 | 0.0645 | 571 | CILDTSNPV=NVCIPNECK=CNL |
Fig. 5. Disulfide bonds SS5 and SS6.
Table 8
Disulfide bond SS7 (Cys95-Cys113) and SS8 (Cys115-Cys129) peptides.

| Assignment | Theoretical Mass (Da) | Observed Mass (Da) | Mass Error (Da) | Intensity (counts) | Identification |
|------------|-----------------------|--------------------|----------------|-------------------|----------------|
| **39 Fragments** | | | | | |
| **Constituent Peptides 11** | | | | | |
| 1/a2 | 200.1399 | 200.141 | −0.0011 | 247 | QV |
| 1/a3 | 301.1876 | 301.1464 | 0.0412 | 100 | QVT |
| 1/b1 | 129.0664 | 129.0653 | 0.0011 | 33 | Q |
| 1/y1 | 147.1133 | 147.1145 | −0.0012 | 1258 | K |
| 1/y3 | 318.1777 | 318.1789 | −0.0012 | 47 | NGK |
| 1/y4 | 375.1992 | 375.2 | −0.0008 | 52 | GNGK |
| 2/a2 | 131.082 | 131.0274 | 0.0546 | 30 | T |
| 2/b1 | 102.0555 | 102.0558 | −0.0003 | 107 | T |
| 2/b2 | 159.077 | 159.0768 | 0.0002 | 140 | TG |
| 2/y2 | 204.1348 | 204.0705 | 0.0643 | 30 | GK |
| 2/y4 | 431.2618 | 431.2624 | −0.0006 | 72 | NGK |

Cys95 to Cys113 3 Fragments
| 1/b7-2/b5 | 1105.441 | 1105.43 | 0.0109 | 157 | QVTGNG = TGVC |
| 1/y6-2/a4 | 909.3923 | 909.4301 | −0.0378 | 58 | TCGNGK = TGVC |
| 1/y6-2/b4 | 937.3871 | 937.3987 | −0.0115 | 257 | TCGNGK = TGVC |

Cys115 to Cys129 4 Fragments
| 2/y5-3/b2 | 722.2966 | 722.3025 | 0.0059 | 45 | CNIGK = CS |
| 2/y5-3/y3 | 868.4021 | 868.3544 | 0.0477 | 103 | CNIGK = CSK |
| 2/y6-3/a2 | 781.3337 | 781.4172 | −0.0835 | 21 | SCNIGK = CS |
| 2/y6-3/b2 | 809.3286 | 809.3259 | 0.0027 | 90 | SCNIGK = CS |

T12, T14, and T16 linkage 21 Fragments
| 1/a4-2/a6-3/a1 | 997.3728 | 997.3982 | −0.0254 | 39 | QVTGVCSC = C |
| 1/a4-2/a6-3/a2 | 1084.405 | 1084.434 | −0.0294 | 146 | QVTGVCSC = CS |
| 1/a4-2/a8-3/b2 | 1339.527 | 1339.585 | −0.0582 | 210 | QVTGVCSCNI = CS |
| 1/a4-2/y8-3/y3 | 1558.685 | 1558.64 | 0.0449 | 111 | QVTGVCSCNIK = CSK |
| 1/a5-2/a6-3/a1 | 1054.394 | 1054.432 | −0.0376 | 60 | QVTGVCSC = C |
| 1/a6-2/y8-3/y3 | 1729.749 | 1729.66 | 0.089 | 51 | QVTGVCSCNIK = CSK |
| 1/a7-2/a7-3/b1 | 1367.497 | 1367.472 | 0.0244 | 33 | QVTGVCSCNIK = CSK |
| 1/b5-2/a9-3/b2 | 1481.565 | 1481.493 | 0.0713 | 35 | QVTGVCSCNIK = CSK |
| 1/b6-2/a7-3/b1 | 1338.47 | 1338.53 | −0.0596 | 72 | QVTGVCSCNIK = CSK |
| 1/b6-2/b9-3/y3 | 1769.708 | 1769.651 | 0.0566 | 1155 | QVTGVCSCNIK = CSK |
| 1/y5-2/b6-3/y3 | 1360.512 | 1360.561 | −0.0047 | 249 | CNGK = TGVCSC = CSK |
| 1/y5-2/b9-3/b1 | 1411.523 | 1411.599 | −0.0076 | 348 | CNGK = TGVCSCNIK = CS |
| 1/y5-2/y8-3/a1 | 1371.564 | 1371.529 | 0.035 | 254 | CNGK = VCSCNIK = CS |
| 1/y5-2/y8-3/b2 | 1486.591 | 1486.645 | −0.0536 | 270 | CNGK = VCSCNIK = CS |
| 1/y6-2/y10-3/y3 | 1891.814 | 1891.761 | 0.0521 | 105 | TCGNGK = TGVCSCNIK = CSK |
| Assignment       | Theoretical Mass (Da) | Observed Mass (Da) | Mass Error (Da) | Intensity (counts) | Identification          |
|------------------|-----------------------|--------------------|-----------------|-------------------|--------------------------|
| 1/y7-2/b8-3/b2   | 1641.649              | 1641.631           | 0.0188          | 78                | VTGNK = TGVCSCNI = CS    |
| 1/y7-2/y10-3/y3  | 1990.882              | 1990.798           | 0.0837          | 124               | VTGNK = TGVCSCNIGK = CSK |
| 1/y7-2/y8-3/b2   | 1686.707              | 1686.75            | −0.0424         | 6211              | VTGNK = VSCNIGK = CS     |
| 1/y8-2/b8-3/b1   | 1682.676              | 1682.723           | −0.0466         | 32                | QVTGNK = TGVCSCNI = C    |
| 1/y8-2/b9-3/y3   | 1972.835              | 1972.819           | 0.0156          | 407               | QVTGNK = TGVCSCNIG = CSK |
| 1/y8-2/y10-3/y3  | 2118.94               | 2118.932           | 0.0085          | 4955              | QVTGNK = TGVCSCNIG = CSK |
Fig. 6. Disulfide bonds SS7 and SS8.
Table 9
Disulfide bond SS9 (Cys137-Cys148) peptides.

| Assignment | Theoretical Mass (Da) | Observed Mass (Da) | Mass Error (Da) |Intensity (counts) | Identification |
|------------|-----------------------|--------------------|-----------------|-------------------|----------------|
| 35 Fragments | constituent peptides 10 Fragments | | | | |
| 1/y2 | 260.1974 | 260.2041 | -0.0067 | 262 | LK |
| 1/y3 | 347.2294 | 347.2299 | -0.0005 | 129 | SLK |
| 2/a1 | 102.0555 | 102.0559 | -0.0004 | 754 | E |
| 2/a2 | 216.0984 | 216.1371 | -0.0387 | 65 | EN |
| 2/a4 | 446.1887 | 446.2249 | -0.0362 | 20 | ENET |
| 2/b1 | 130.0504 | 130.0469 | 0.0035 | 17 | E |
| 2/b2 | 244.0933 | 244.0929 | 0.0004 | 367 | EN |
| 2/b3 | 373.1359 | 373.1376 | -0.0017 | 171 | EN |
| 2/b4 | 474.1836 | 474.1958 | -0.0122 | 88 | ENET |
| 2/y1 | 147.1133 | 147.1133 | 0 | 1131 | K |

Cys137 and Cys148 25 Fragments

| Assignment | Theoretical Mass (Da) | Observed Mass (Da) | Mass Error (Da) |Intensity (counts) | Identification |
|------------|-----------------------|--------------------|-----------------|-------------------|----------------|
| 1/a1-2/y2 | 323.1212 | 323.1008 | 0.0204 | 15 | C=CK |
| 1/a1-2/y4 | 553.2114 | 553.2058 | 0.0056 | 53 | C=ETCK |
| 1/a1-2/y5 | 667.2544 | 667.2748 | -0.0204 | 59 | C=NETCK |
| 1/a2-2/y2 | 410.1532 | 410.1623 | -0.0091 | 36 | CS=CK |
| 1/a2-2/y4 | 640.2435 | 640.2652 | -0.0217 | 34 | CS=ETCK |
| 1/a3-2/y3 | 624.2849 | 624.2893 | -0.0044 | 27 | CSL=CKT |
| 1/a3-2/y5 | 867.3705 | 867.3508 | 0.0197 | 89 | CSL=NETCK |
| 1/b1-2/a5 | 650.1914 | 650.2377 | -0.0463 | 130 | C=ENETC |
| 1/b2-2/a5 | 737.2234 | 737.2545 | -0.0311 | 139 | CS=ENETC |
| 1/b1-2/y2 | 351.1161 | 351.1717 | -0.0556 | 207 | C=CK |
| 1/b1-2/y3 | 452.1638 | 452.1826 | -0.0188 | 23 | C=CKT |
| 1/b2-2/y2 | 438.1481 | 438.1512 | -0.0031 | 139 | CS=CK |
| 1/b2-2/y3 | 539.1958 | 539.2008 | -0.005 | 166 | CS=CKT |
| 1/b2-2/y5 | 782.2813 | 782.3434 | -0.0621 | 180 | CS=NETCK |
| 1/b3-2/y2 | 551.2322 | 551.2324 | -0.0002 | 219 | CSL=CK |
| 1/y4-2/y2 | 697.3377 | 697.3329 | 0.0048 | 465 | CSLK=CK |
| 1/y4-2/a5 | 961.413 | 961.4069 | 0.0061 | 130 | CSLK=ENETC |
| 1/y4-2/b5 | 1024.408 | 1024.4126 | -0.0046 | 502 | CSLK=ENETC |
| 1/y4-2/y3 | 798.3854 | 798.3802 | 0.0052 | 553 | CSLK=CKT |
| 1/y4-2/y4 | 927.428 | 927.4216 | 0.0064 | 223 | CSLK=ETCK |
| 1/y4-2/y5 | 1041.4709 | 1041.4377 | 0.0332 | 94 | CSLK=NETCK |
| 1/y4-2/y6 | 1170.5134 | 1170.511 | 0.0024 | 25002 | CSLK=ENETCK |
| 1/a1-2/y6 | 796.2969 | 796.3787 | -0.0818 | 646 | C=ENETC |
| 1/b2-2/y6 | 911.3239 | 911.2963 | 0.0276 | 155 | CS=ENETC |
| 1/y4-2/z3 | 781.3588 | 781.3229 | 0.0359 | 77 | CSLK=CKT |
Fig. 7. Disulfide bond SS9.
Table 10
Disulfide bond SS10 (Cys141-Cys157) and SS11 (Cys159-Cys172) peptides.

| Assignment | Theoretical Mass (Da) | Observed Mass (Da) | Mass Error (Da) | Intensity (counts) | Identification |
|------------|-----------------------|--------------------|-----------------|-------------------|----------------|
| 65 Fragments |                       |                    |                 |                   |                |
|            | Constituent Peptides 19 fragments |       |                 |                   |                |
| 1/y1       | 147.1133              | 147.1126           | 0.0007          | 1153              | K              |
| 1/y2       | 260.1974              | 260.2021           | -0.0047         | 167               | LK             |
| 3/a2       | 145.0613              | 145.0646           | -0.0033         | 400               | DG             |
| 3/a5       | 518.2979              | 518.296            | 0.0018          | 109               | DGFII          |
| 3/a8       | 890.426               | 890.3921           | 0.0339          | 84                | DGFIIIDQEQ     |
| 3/a9       | 977.458               | 977.3666           | 0.0914          | 120               | DGFIIIDQES     |
| 3/b1       | 116.0348              | 116.0532           | -0.0184         | 28                | D              |
| 3/b2       | 173.0562              | 173.0561           | 0.0001          | 645               | DG             |
| 3/b3       | 320.1246              | 320.1256           | -0.0010         | 6533              | DGF            |
| 3/b4       | 433.2087              | 433.2064           | 0.0023          | 1906              | DGFII          |
| 3/b5       | 546.2927              | 546.2884           | 0.0043          | 325               | DGFII          |
| 3/b6       | 661.3197              | 661.3064           | 0.0133          | 295               | DGFIIID        |
| 3/b9       | 1005.453              | 1005.383           | 0.0699          | 237               | DGFIIIDQES     |
| 3/y1       | 156.0773              | 156.0769           | 0.0004          | 4578              | H              |
| 3/y2       | 293.1362              | 293.1368           | -0.0006         | 1684              | HH             |
| 3/y3       | 430.1951              | 430.1913           | 0.0038          | 1430              | HHH            |
| 3/y4       | 567.254               | 567.2552           | -0.0012         | 1460              | HHHH           |
| 3/y5       | 704.3129              | 704.3094           | 0.0035          | 1025              | HHHHH          |
| 3/y7       | 942.4196              | 942.4053           | 0.0143          | 588               | THHHHHH        |

Cys141 to Cys157 11 Fragments

| Assignment | Theoretical Mass (Da) | Observed Mass (Da) | Mass Error (Da) | Intensity (counts) | Identification |
|------------|-----------------------|--------------------|-----------------|-------------------|----------------|
| 1/a1-2/a1  | 149.0207              | 149.1165           | -0.0958         | 16                | C=C            |
| 1/a2-2/a1  | 262.1048              | 262.1324           | -0.0276         | 39                | CL=C           |
| 1/a2-2/a2  | 377.1317              | 377.1469           | -0.0152         | 258               | CL=CD          |
| 1/a2-2/b1  | 290.0997              | 290.1243           | -0.0246         | 45                | CL=C           |
| 1/a2-2/b2  | 405.1266              | 405.1262           | 0.0005          | 230               | CL=CD          |
| 1/y3-2/a12 | 551.2322              | 551.2063           | 0.0259          | 62                | CL=C           |
| 1/b1-2/a2  | 292.0426              | 292.0312           | 0.0114          | 82                | C=CD           |
| 1/b1-2/b1  | 205.0106              | 205.0966           | -0.086          | 154               | C=C            |
| 1/b2-2/b1  | 318.0946              | 318.0872           | 0.0074          | 36                | CL=C           |
| 1/y3-2/b1  | 464.2001              | 464.214            | -0.0139         | 41                | CLK=C          |
| 1/y3-2/b2  | 579.2271              | 579.2289           | -0.0018         | 1027              | CLK=CD         |

Cys159 to Cys172 3 fragments

| Assignment | Theoretical Mass (Da) | Observed Mass (Da) | Mass Error (Da) | Intensity (counts) | Identification |
|------------|-----------------------|--------------------|-----------------|-------------------|----------------|
| 2/y2-3/a12 | 1527.682              | 1527.653           | 0.0291          | 242               | CK=DGFIIIDQESSIC |
| 2/y2-3/y19 | 2497.089              | 2497.066           | 0.0227          | 994               | CK=DGFIIIDQESSICTHHHHHHH |
| 2/y3-3/a12 | 1642.709              | 1642.633           | 0.0763          | 69                | DCK=DGFIIIDQESSIC |
| Linkage Fragments | Mass (Da) | Delta Mass (Da) | p-Value | CID |
|-------------------|----------|----------------|---------|-----|
| T12, T14, and T16 |          |                |         |     |
| 1/a1–2/a3-3/a16  | 2156.8411| 0.0266         | 37      | CDC |
| 1/a1–2/a3-3/y10  | 1609.5782| 0.0193         | 132     | CDC |
| 1/a1–2/y4-3/y18  | 2673.0967| -0.0608        | 74      | CDC |
| 1/a1–2/y4-3/y8   | 1583.5625| -0.0533        | 363     | CDC |
| 1/a2-2/a3-3/a16  | 2269.9225| 0.0051         | 921     | CDC |
| 1/a2-2/a3-3/a18  | 2544.043 | 0.0193         | 132     | CDC |
| 1/a2-2/y4-3/a15  | 2306.9666| 0.0623         | 74      | CDC |
| 1/b1–2/b3-3/a12  | 1700.6064| -0.0505        | 100     | CDC |
| 1/b1–2/b3-3/y14  | 2124.7283| -0.0476        | 72      | CDC |
| 1/b1–2/y4-3/a14  | 2054.8186| -0.0437        | 85      | CDC |
| 1/b2-2/b3-3/b13  | 1942.7332| -0.0597        | 95      | CDC |
| 1/b2-2/b3-3/y8   | 1578.536 | -0.0721        | 228     | CDC |
| 1/b2-2/y4-3/a14  | 2197.9026| -0.0066        | 671     | CDC |
| 1/b2-2/y4-3/b17  | 2637.042 | 0.0576         | 233     | CDC |
| 1/b2-2/y4-3/y10  | 1924.7576| -0.0524        | 84      | CDC |
| 1/b2-2/y4-3/y16  | 2610.0859| -0.0017        | 407     | CDC |
| 1/y3-2/y4-3/a18  | 2892.2439| 0.0803         | 403     | CDK |
| 1/y3-2/y4-3/b14  | 2372.0332| -0.0352        | 341     | CDK |
| 1/y3-2/y4-3/b18  | 2920.2388| -0.0088        | 1160    | CDK |
| 1/y3-2/y4-3/y11  | 2157.8953| 0.0024         | 922     | CDK |
| 1/y3-2/y4-3/y13  | 2414.9963| -0.011         | 897     | CDK |
| 1/y3-2/y4-3/y14  | 2530.0232| -0.0046        | 2415    | CDK |
| 1/y3-2/y4-3/y16  | 2756.1914| -0.0088        | 1421    | CDK |
| 1/y3-2/y4-3/y18  | 2960.2812| 0.043          | 911     | CDK |
| 1/y3-2/y4-3/y19  | 3075.3081| 0.0005         | 42471   | CDK |
| 1/y3-2/y4-3/y8   | 1870.7471| -0.0212        | 131     | CDK |
| 1/y3-2/y4-3/z14  | 2512.9968| -0.0657        | 102     | CDK |
| 1/y3-2/y4-3/z8   | 1853.7205| -0.0724        | 66      | CDK |
Acknowledgements

The authors thank Ashley Birkett and Merribeth Morin of PATH’s Malaria Vaccine Initiative (MVI) for their support and assistance. The authors thank Steven Becht, Ying-Hua Chang and Jie Ding from PPD GMP Lab in Middleton, WI for work on the LC-MS/MS method for peptide mapping analysis. The authors thank the project team at Syngene International, a Biocon company, in Bangalore, India for cloning, production, and preliminary characterization of the baculovirus-expressed Pfs25. The work

Fig. 8. Disulfide bonds SS10 and SS11.
was supported by the Bill & Melinda Gates Foundation (OPP1108403). The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Foundation.

Transparency document.  Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2018.03.034.

References

[1] S.M. Lee, J.L. Plieskatt, C.R. King, Disulfide bond mapping of Pfs25, a recombinant malaria transmission blocking vaccine candidate, Anal. Biochem. 542 (2018) 20–23.
[2] S.M. Lee, C.K. Wu, J. Plieskatt, D.H. McAdams, K. Miura, C. Ockenhouse, C.R. King, Assessment of Pfs25 expressed from multiple soluble expression platforms for use as transmission-blocking vaccine candidates, Malar. J. 15 (2016) 405.