

| fHbp, ID | Sequence | |
|----------|---|-----|
| 15 | CSSGGGGGGGVAADIGAGLADALTAPLDHKDKGLKSLTLEDNISQNGTLTSAQGAER | 60 |
| 14 | -----CSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEKLKLAQAQGAEK | 55 |
| 10 | -----CSSGGGGVAADIGAGLADALTAPLDHKDKGLQSLMLDQSVRKNEKLKLAQAQGAEK | 55 |
| 12 | -----CSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEKLKLAQAQGAEK | 55 |
| 215 | -----CSSGGGGVAADIGAGLADALTAPLDHKDKGLQSLTLDQSVRKNEKLKLAQAQGAEK | 55 |
| 37 | -----CSSGGGGVAADIGAGLADALTAPLDHKDKGLQSLMLDQSVRKNEKLKLAQAQGAEK | 55 |
| 4 | -----CSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEKLKLAQAQGAEK | 55 |
| 110 | -----CSSGGGGVAADIGAGLADALTAPLDHKDKGLQSLTLDQSVRKNEKLKLAQAQGAEK | 55 |
| 1 | -----CSSGGGGVAADIGAGLADALTAPLDHKDKGLQSLTLDQSVRKNEKLKLAQAQGAEK | 55 |
| 144 | -----CSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEKLKLAQAQGAEK | 55 |
| | *****.*:*:*:*:*:*:*:*:*:*: | |
| 15 | TFKAGDKDNSLNTGKLKNDKISRDFIRQIEVDGQLITLESGEFQVYKQSHSALTALQTE | 120 |
| 14 | TYGNGD---SLNTGKLKNDKISRDFIRQIEVDGQLITLESGEFQVYKQSHSALTALQTE | 112 |
| 10 | TYGNGD---SLNTGKLKNDKISRDFIRQIEVDGQLITLESGEFQVYKQSHSALTALQTE | 112 |
| 12 | TYGNGD---SLNTGKLKNDKISRDFIRQIEVDGQLITLESGEFQVYKQSHSALTALQTE | 112 |
| 215 | TYGNGD---SLNTGKLKNDKISRDFIRQIEVDGQLITLESGEFQVYKQSHSALTALQTE | 112 |
| 37 | TYGNGD---SLNTGKLKNDKISRDFIRQIEVDGQLITLESGEFQVYKQSHSALTALQTE | 112 |
| 4 | TYGNGD---SLNTGKLKNDKISRDFIRQIEVDGQLITLESGEFQVYKQSHSALTALQTE | 112 |
| 110 | TYGNGD---SLNTGKLKNDKISRDFIRQIEVDGQLITLESGEFQVYKQSHSALTALQTE | 112 |
| 1 | TYGNGD---SLNTGKLKNDKISRDFIRQIEVDGQLITLESGEFQVYKQSHSALTALQTE | 112 |
| 144 | TYGNGD---SLNTGKLKNDKISRDFIRQIEVDGQLITLESGEFQVYKQSHSALTALQTE | 112 |
| | *: ** *****.*:*:*:*:*:*:*:*:*:*: | |
| 15 | QVQDSEHSGKMVAKRFRIGDIAGEHTSFGKLPKDVMATYRGTAFGSDDAGGKLTYYTIDF | 180 |
| 14 | QVQDSEHSGKMVAKRFRIGDIAGEHTSFGKLPKDVMATYRGTAFGSDDAGGKLTYYTIDF | 172 |
| 10 | QVQDSEHSGKMVAKRFRIGDIAGEHTSFGKLPKGGSATYRGTAFGSDDAGGKLTYYTIDF | 172 |
| 12 | QVQDSEHSGKMVAKRFRIGDIAGEHTSFGKLPKGGSATYRGTAFGSDDAGGKLTYYTIDF | 172 |
| 215 | QVQDSEHSGKMVAKRFRIGDIAGEHTSFGKLPKGGSATYRGTAFGSDDAGGKLTYYTIDF | 172 |
| 37 | QVQDSEHSGKMVAKRFRIGDIAGEHTSFGKLPESDRATYRGTAFGSDDAGGKLTYYTIDF | 172 |
| 4 | QVQDSEHSGKMVAKRFRIGDIAGEHTSFGKLPESDRATYRGTAFGSDDAGGKLTYYTIDF | 172 |
| 110 | QVQDSEHSGKMVAKRFRIGDIAGEHTSFGKLPESDRATYRGTAFGSDDAGGKLTYYTIDF | 172 |
| 1 | QVQDSEHSGKMVAKRFRIGDIAGEHTSFGKLPESDRATYRGTAFGSDDAGGKLTYYTIDF | 172 |
| 144 | QVQDSEHSGKMVAKRFRIGDIAGEHTSFGKLPESDRATYRGTAFGSDDAGGKLTYYTIDF | 172 |
| | * ** *.* *****.*:*:*:*:*:*:*:*:*:*: | |
| 15 | AAKQGGKIEHLKSPELNVDLAAADIKPDEKRRHAVISGSVLYNQAEKGSYSLGIFGGQAQ | 240 |
| 14 | AAKQGGKIEHLKSPELNVDLAAADIKPDEKRRHAVISGSVLYNQAEKGSYSLGIFGGQAQ | 232 |
| 10 | AAKQGGKIEHLKSPELNVDLAAADIKPDEKRRHAVISGSVLYNQAEKGSYSLGIFGGQAQ | 232 |
| 12 | AAKQGGKIEHLKSPELNVDLAAADIKPDEKRRHAVISGSVLYNQAEKGSYSLGIFGGQAQ | 232 |
| 215 | AAKQGGKIEHLKSPELNVDLAAADIKPDEKRRHAVISGSVLYNQAEKGSYSLGIFGGQAQ | 232 |
| 37 | AAKQGGKIEHLKSPELNVDLAAADIKPDEKRRHAVISGSVLYNQAEKGSYSLGIFGGQAQ | 232 |
| 4 | AAKQGGKIEHLKSPELNVDLAAADIKPDEKRRHAVISGSVLYNQAEKGSYSLGIFGGQAQ | 232 |
| 110 | AAKQGGKIEHLKSPELNVDLAAADIKPDEKRRHAVISGSVLYNQAEKGSYSLGIFGGQAQ | 232 |
| 1 | AAKQGGKIEHLKSPELNVDLAAADIKPDEKRRHAVISGSVLYNQAEKGSYSLGIFGGQAQ | 232 |
| 144 | AAKQGGKIEHLKSPELNVDLAAADIKPDEKRRHAVISGSVLYNQAEKGSYSLGIFGGQAQ | 232 |
| | *****.*:*:*:*:*:*:*:*:*:*: | |
| 15 | EVAGSAEVTANGIRHIGLAQAQ--- | 263 |
| 14 | EVAGSAEVTANGIRHIGLAQAQ--- | 255 |
| 10 | EVAGSAEVTANGIRHIGLAQAQ--- | 255 |
| 12 | EVAGSAEVTANGIRHIGLAQAQ--- | 255 |
| 215 | EVAGSAEVTANGIRHIGLAQAQ--- | 255 |
| 37 | EVAGSAEVTANGIRHIGLAQAQ--- | 258 |
| 4 | EVAGSAEVTANGIRHIGLAQAQ--- | 255 |
| 110 | EVAGSAEVTANGIRHIGLAQAQ--- | 255 |
| 1 | EVAGSAEVTANGIRHIGLAQAQ--- | 255 |
| 144 | EVAGSAEVTANGIRHIGLAQAQ--- | 255 |
| | *****.*:*:*:*:*:*:*:*:*:*: | |

Figure S1. Alignment of amino acid sequences of cross-reactive variants of fHbp from isolates of MenB strains, 2013-2023, and fHbp peptide, ID 1, used in the 4CMenB vaccine (the sequence is marked with a green dot). Alignment was done in Clustal Omega service (<https://www.kelleybioinfo.org/algorithms/tutorial/TAl2.pdf>). Green color indicates polymorphic sites that occur more often than 30% of the sequences (i.e., at least 30% of the studied sequences of cross-reactive fHbp variants contain an amino acid substitution in this position). An asterisk “*” indicates fully conserved amino acid position, two dots “:” indicate positions with conservation between amino acids with similar properties and one dot “.” indicates positions with conservation between amino acids with weakly similar properties. The number of last amino acid residue in line is indicated on right.

Table S1. Polymorphic sites in the C-terminal region that occur at different frequencies in fHbp sequences from the “Cross-reactive” and “None” groups*.

| Frequency of substitution, % | Cross-reactive fHbp | None fHbp |
|------------------------------|---|---|
| 100 | - | S140A <u>D142N</u> K143Q G147- <u>R149K</u> T151E R153H <u>G158S</u> S186T <u>D197N</u> T198L P200A <u>G202E</u> R204S S209L S211D V212T I213R <u>N215G</u> <u>G229D</u> <u>K230R</u> G225A I226L S221T S223H T155K I246V <u>R247H</u> H248E L251I T242I N244E G245K A253G V234I E239T |
| 50-99 | <u>E146K</u> <u>N178H</u> <u>G202E</u> <u>R204H</u> <u>A217D</u> <u>K230Q</u> K241E <u>V243A</u> R247H | A162P G163N K165R T167H T169S <u>N178H</u> L189Q <u>D192E</u> A195S Q217S <u>V243R</u> |
| <50 | D142G <u>G147D</u> G148V <u>R149S</u> G158S <u>N178Y</u> <u>D192E</u> D197Y <u>D197N</u> R200Q <u>G229E</u> | <u>E146D</u> <u>E146S</u> <u>E146G</u> <u>G147D</u> A174K A174N A174S <u>N178Y</u> <u>K180R</u> Q217G <u>V243G</u> |

*Underlined polymorphisms are common for sequences of fHbp from “Cross-reactive” and “None” groups. Polymorphisms in red were mentioned in literature earlier as potentially important for antigen-antibody interaction [13, 25, 26]. The fHbp vaccine antigen (ID 1) was used as a control.

“.” indicates positions with conservation between amino acids with weakly similar properties. The number of last amino acid residue in line is indicated on the right.

| NHBA, ID | Sequence | |
|----------|--|-----|
| 197 | ARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEK | 354 |
| 43 | ARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEK | 353 |
| 47 | ARVRRSARSRRSLPAEMPLIPVNQADTLIVDGEVVSITGHSNIFAPEGNYRYLTYGAEK | 352 |
| 120 | ARVRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVILTEHPGNIFAPEGNYRYLTYGAEK | 337 |
| 6 | ARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEK | 292 |
| 9 | ARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEK | 292 |
| 17 | ARVRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEK | 291 |
| 18 | ARVRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEK | 291 |
| 31 | ---TRRSARSRRSLPAEIPLIPVNQADTLIVDGEAVSLTGYSNIFAPEGNYRYLTYGAEK | 285 |
| 63 | ---TRRSARSRRSLPAEIPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEK | 285 |
| 30 | ---RSARSRRSLPAEMPLIPVNQVDTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEK | 346 |
| 2 | ARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEK | 358 |
| 112 | -RSARSRRSLPAEMPLIPVNQADTLIVDGEAVILTEHPGNIFAPEGNYRYLTYGAEK | 351 |
| | ***:*****:****:***:***:*****:***** | |
| 197 | LSGGSYALSVQGEPAKGEMLAGTAVYNGEVLHFHFMENGRPSPPGGRFAAKVDFGSKSVGDG | 414 |
| 43 | LSGGSYALRVQGEPAKGEMLAGTAVYNGEVLHFHFMENGRPSPPGGRFAAKVDFGSKSVGDG | 413 |
| 47 | LSGGSYALRVQGEPAKGEMLAGTAVYNGEVLHFHTENGRSYPTKGRFAAKVDFGSKSVGDG | 412 |
| 120 | LSGGSYALRVQGEPAKGEMLAGTAVYNGEVLHFHTENGRSYPTKGRFAAKVDFGSKSVGDG | 397 |
| 6 | LSGGSYALSVQGEPAKGEMLAGTAVYNGEVLHFHTENGRSYPTKGRFAAKVDFGSKSVGDG | 352 |
| 9 | LSGGSYALSVQGEPAKGEMLAGTAVYNGEVLHFHTENGRSYPTKGRFAAKVDFGSKSVGDG | 352 |
| 17 | LSGGSYALSVQGEPAKGEMLAGTAVYNGEVLHFHTENGRSYPTKGRFAAKVDFGSKSVGDG | 351 |
| 18 | LSGGSYALSVQGEPAKGEMLAGTAVYNGEVLHFHTENGRSYPTKGRFAAKVDFGSKSVGDG | 351 |
| 31 | LSGGSYALSVQGEPAKGEMLAGTAVYNGEVLHFHTENGRSYPTKGRFAAKVDFGSKSVGDG | 345 |
| 63 | LSGGSYALRVQGEPAKGEMLAGTAVYNGEVLHFHFMENGRPSPPGGRFAAKVDFGSKSVGDG | 345 |
| 30 | LSGGSYALRVQGEPAKGEMLAGTAVYNGEVLHFHFMENGRPSPPGGRFAAKVDFGSKSVGDG | 406 |
| 2 | LPGGSYALRVQGEPSKGEMLAGTAVYNGEVLHFHTENGRPSPPSRGRFAAKVDFGSKSVGDG | 418 |
| 112 | LSGGSYALRVQGEPAKGEMLAGTAVYNGEVLHFHFMENGRPSPPGGRFAAKVDFGSKSVGDG | 411 |
| | *:*****:*****:*****:*****:*****:*****:***** | |
| 197 | IIDSGDDLHMGTKQKFAVIDGNGFKGTWTENGGDVSGRFYGPAGEEVAGKYSYRPTDAE | 474 |
| 43 | IIDSGDDLHMGTKQKFAVIDGNGFKGTWTENGGDVSGRFYGPAGEEVAGKYSYRPTDAE | 473 |
| 47 | IIDSGDDLHMGTKQKFAVIDGNGFKGNWTENGGDVSGRFYGPAGEEVAGKYSYRPTDAE | 472 |
| 120 | IIDSGDDLHMGTKQKFAAIDGNGFKGTWTENGGDVSGRFYGPAGEEVAGKYSYRPTDAE | 457 |
| 6 | IIDSGDDLHMGTKQKFAAIDGNGFKGTWTENGGDVSGRFYGPAGEEVAGKYSYRPTDAE | 412 |
| 9 | IIDSGDDLHMGTKQKFAAIDGNGFKGNWTENGGDVSGRFYGPAGEEVAGKYSYRPTDAE | 412 |
| 17 | IIDSGDDLHMGTKQKFAAIDGNGFKGTWTENGGDVSGRFYGPAGEEVAGKYSYRPTDAE | 411 |
| 18 | IIDSGDDLHMGTKQKFAAIDGNGFKGTWTENGGDVSGRFYGPAGEEVAGKYSYRPTDAE | 411 |
| 31 | IIDSGDDLHMGTKQKFAAIDGNGFKGTWTENGGDVSGRFYGPAGEEVAGKYSYRPTDAE | 405 |
| 63 | IIDSGDDLHMGTKQKFAVIDGNGFKGAWTENSAGDVSGRFYGPAGEEVAGKYSYRPTDAE | 405 |
| 30 | IIDSGDDLHMGTKQKFAAIDGNGFKGTWTENGGDVSGRFYGPAGEEVAGKYSYRPTDAE | 466 |
| 2 | IIDSGDDLHMGTKQKFAAIDGNGFKGTWTENGGDVSGRFYGPAGEEVAGKYSYRPTDAE | 478 |
| 112 | IIDSGDDLHMGTKQKFAAIDGNGFKGAWTENSAGDVSGRFYGPAGEEVAGKYSYRPTDAE | 471 |
| | *****:*****:*****:*****:*****:*****:*****:***** | |
| 197 | KGGFGVFAGKKEQD | 488 |
| 43 | KGGFGVFAGKKEQD | 487 |
| 47 | KGGFGVFAGKKEQD | 486 |
| 120 | KGGFGVFAGKKEQD | 471 |
| 6 | KGGFGVFAGKKEQD | 426 |
| 9 | KGGFGVFAGKKEQD | 426 |
| 17 | KGGFGVFAGKKEQD | 425 |
| 18 | KGGFGVFAGKKEQD | 425 |
| 31 | KGGFGVFAGKKEQD | 419 |
| 63 | KGGFGVFAGKKEQD | 419 |
| 30 | KGGFGVFAGKKEQD | 480 |
| 2 | KGGFGVFAGKKEQD | 492 |
| 112 | KGGFGVFAGKKEQD | 485 |
| | ***** | |

Figure S3. Comparison of the amino acid sequences of NHBA C-terminal domains from the “None” group with corresponding sequence of the NHBA peptide, ID 2, included in the 4CMenB vaccine. Red indicates the positions at which substitutions occur in the sequences of each NHBA of this group, and blue indicates the remaining polymorphic sites in the C-terminal domain. Alignment of amino acid sequences constructed in Clustal Omega (<https://www.kelleybioinfo.org/algorithms/tutorial/TAl2.pdf>). An asterisk “*” indicates fully conserved amino acid position, two dots “:” indicate positions with conservation between amino acids with similar properties and one dot

“.” indicates positions with conservation between amino acids with weakly similar properties. The number of last amino acid residue in line is indicated on the right.

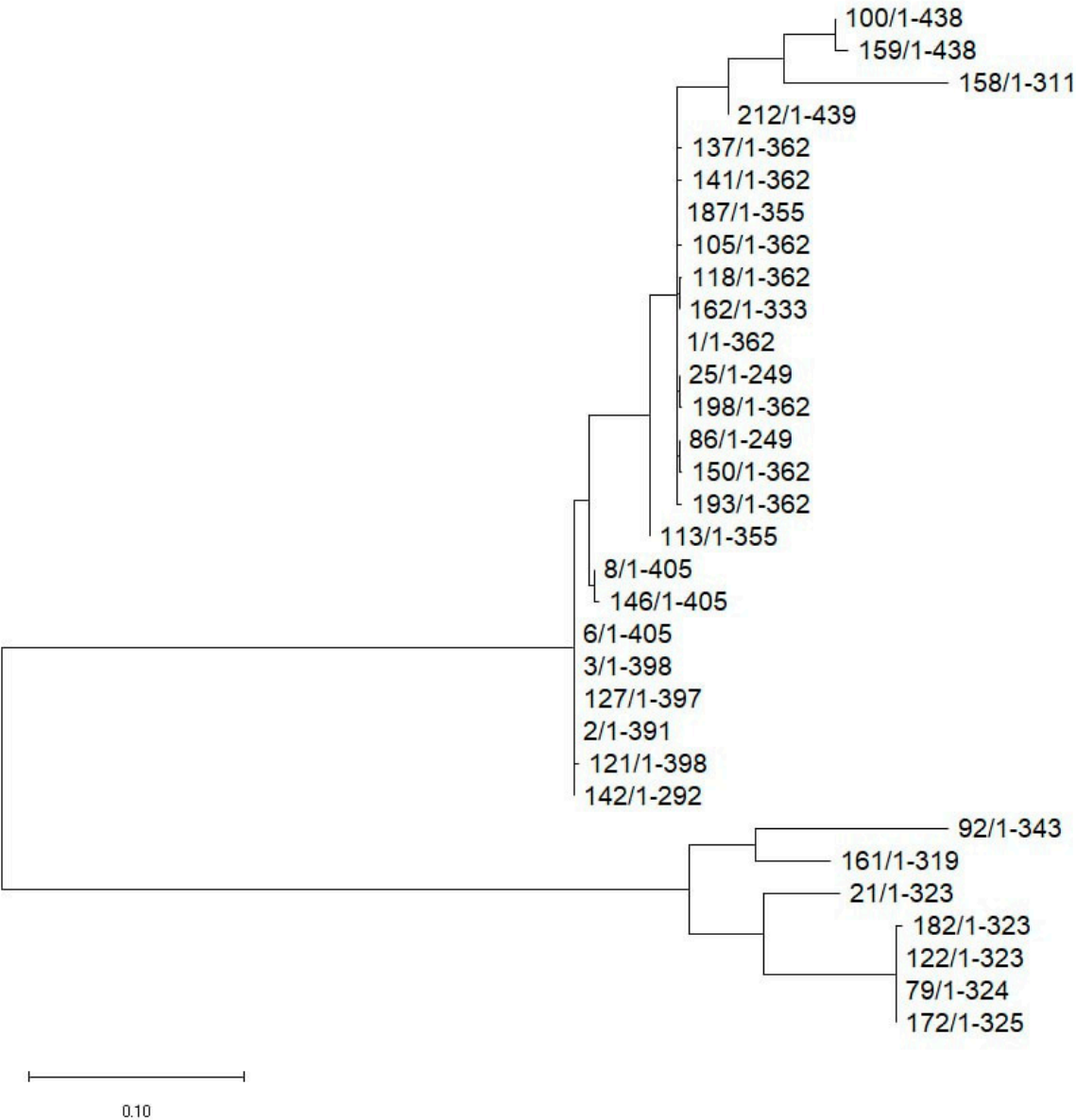


Figure S4. The tree of the NadA. Multiple alignments of amino acids sequences produced by the MUSCLE algorithm to construct phylogenetic trees of the NadA antigen. The trees were constructed using the Maximum Likelihood algorithm and visualized in MEGA software.