

Epitope Mapping report

Date: 19.09.2016

Target: anti-actin mAb Ab-5 (BD, Nr. 612656)

Antigen: chicken gizzard muscle actin

Panning conditions: Immobilized on M270 carboxy beads

NGS library statistics:

| | |
|----------------------------|---------------------------------|
| Dataset: | anti_actin_ab-1pr (1st Panning) |
| Number of sequences: | 466,434 |
| Number of valid sequences: | 310,946 |
| Pattern/Library: | ENTE-1 |
| 3-mer motifs: | 139,270 |
| 4-mer motifs: | 7,253 |

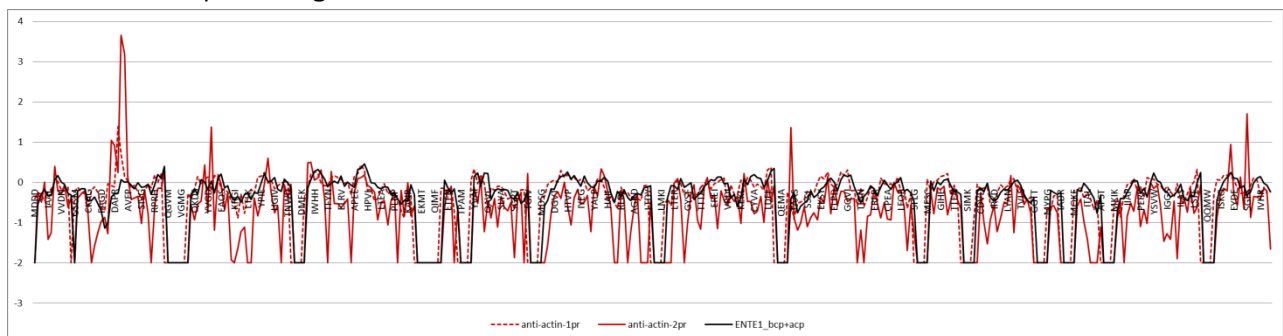
Number of sequences is the output from NGS

Valid sequences are those finally fitting all QC standards and used for the analysis

| | |
|----------------------------|---------------------------------|
| Dataset: | anti_actin_ab-2pr (2nd Panning) |
| Number of sequences: | 665,029 |
| Number of valid sequences: | 461,105 |
| Pattern/Library: | ENTE-1 |
| 3-mer motifs: | 126,297 |
| 4-mer motifs: | 7,249 |

The number of *motifs* in the first round should be close to theory, i.e. ca 7,200 resp. 132,000 for medium size data sets of up to 500,000 sequences

Motif statistics plotted against chicken actin:



Statistic of 4mer motifs in the data sets. The reference data set (black) contains almost 2 Mio Sequences from the naïve library.

X-axis: All possible 4mer fragments of the antigen starting from 1-4, 2-5 ...to... end

Y-axis: log of enrichment vs. calculated probability in the data set, e.g. 3 is 10^3 resp. 1000-fold enrichment. Control ENTE1 naïve starting library has max 0.5 deviation from 0. After the first panning round the library is depleted of motifs, many approach a cut off of <-2 , which should not be taken as a save value with respect to the number of observed sequences. Because this would mean that the motif is less than 1/100 of the expected average, which usually is less than 1 sequence in a million.

Epitope

Most likely epitope sequence is:

D(e/d)xPRAVF

Presentation as web logo first selection round



This epitope is based on the statistical analysis based on the protein sequence below. Genetic variants may lead to other results.

The web-logo is available online:
<http://weblogo.berkeley.edu/>

The Weblogo is based on each sequence counted once, the following table for the fingerprint counts each found sequence, i.e. enriched sequences have stronger weight, confirming D preceding and F following PRAV.

| | A5 | A4 | A3 | A2 | A1 | M | B1 | B2 | B3 | B4 | B5 |
|---|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|-----|
| C | 0 | 0 | 11 | 15 | 3 | PRAV | 58 | 14 | 18 | 7 | 8 |
| P | 10 | 6 | 7 | 9 | 0 | PRAV | 20 | 19 | 13 | 0 | 9 |
| G | 8 | 3 | 25 | 28 | 15 | PRAV | 22 | 8 | 1 | 3 | 6 |
| A | 4 | 123 | 16 | 23 | 0 | PRAV | 24 | 10 | 10 | 11 | 5 |
| V | 20 | 40 | 0 | 6 | 2 | PRAV | 64 | 5 | 1 | 1 | 15 |
| I | 11 | 15 | 85 | 3 | 3 | PRAV | 26 | 6 | 11 | 7 | 0 |
| L | 14 | 10 | 17 | 10 | 0 | PRAV | 5 | 11 | 7 | 7 | 9 |
| M | 0 | 37 | 0 | 0 | 0 | PRAV | 0 | 0 | 0 | 0 | 0 |
| F | 35 | 38 | 2 | 39 | 1 | PRAV | 105 | 65 | 14 | 5 | 1 |
| Y | 13 | 31 | 6 | 30 | 2 | PRAV | 63 | 53 | 32 | 21 | 6 |
| W | 20 | 40 | 8 | 3 | 2 | PRAV | 38 | 6 | 0 | 0 | 0 |
| T | 17 | 3 | 9 | 26 | 0 | PRAV | 12 | 5 | 1 | 2 | 2 |
| S | 8 | 20 | 65 | 33 | 36 | PRAV | 24 | 1 | 35 | 0 | 35 |
| N | 15 | 27 | 46 | 104 | 0 | PRAV | 7 | 18 | 5 | 39 | 2 |
| Q | 27 | 6 | 34 | 47 | 1 | PRAV | 47 | 69 | 7 | 6 | 8 |
| R | 46 | 34 | 14 | 4 | 0 | PRAV | 8 | 4 | 1 | 0 | 7 |
| K | 52 | 5 | 7 | 2 | 2 | PRAV | 3 | 14 | 0 | 1 | 5 |
| H | 12 | 23 | 24 | 19 | 3 | PRAV | 12 | 19 | 1 | 5 | 1 |
| E | 34 | 40 | 36 | 80 | 436 | PRAV | 4 | 6 | 20 | 6 | 5 |
| D | 174 | 45 | 134 | 65 | 40 | PRAV | 4 | 13 | 24 | 4 | 1 |
| | 520 | 546 | 546 | 546 | 546 | | 546 | 346 | 201 | 125 | 125 |

Fingerprint is the absolute or relative frequency of amino acids observed surrounding the epitope core sequence. Fingerprints may vary with the search sequence. Usually repeated runs with the ENTE-1 library result in similar fingerprints. The motif fingerprint of an antibody is therefore as reliable as a human fingerprint.

Epitope in Sequence

```
>gi|45382927|ref|NP_990849.1| actin, cytoplasmic 1 [Gallus gallus]
MDDDIAALVVDNGSGMCKAGFAGDDAPRAVFP SIVGRPRHQGVMVGMGQKDSYVGDEAQS KRGI LTLKYP I EHGIVTNWDD
MEKIWHHTFYNELRVAP E EHPVLLTEAPLNPKANREKMTQIMFETFNT PAMYVAIQAVLSLYASGR TTGIVMDSGDGV THT
VPIYEGYALPHAILRLDLAGRDLTDYLMKILTERGYSFTTTAEREIVRDIKEKLCYVALDFEQEMATAASSSSSLEKSYELP
DGQVITIGNERFRCPEALFQPSFLGMESCGIHETTFNSIMKCDVDIRKDL YANTVLSGGTTMYPGIADRMQKEITALAPST
MKIKIIAPPERKYSVWIGGSILASLSTFQQMWISKQEYDESGPSIVHRKCF
```

List of sequences and their frequencies from the first selection round with the motif PRAV

```
>seq1 | count: 4          -----DHQCEPRAVAWDYPAC-
-----VDNSPRAVQQCQYQTC
>seq2 | count: 5          -----DKTSEPRAVQQEYARC-
-----YDTSPRAVIKEALPTF
>seq3 | count: 4          -----KSGSDPRAVTPPYVQD-
-----VDNSPRAVQQCEKGVY
>seq4 | count: 4          -----VAQSHPRAVVCQG CDC-
-----YDTSPRAVIKELEAVI
>seq5 | count: 2          -----KSGSDPRAVTPPCVQD-
-----HIRSPRAVYDIILDGY
>seq9 | count: 2          -----IPPCDPRAVEQH YRIC-
-----YDTSPRAVIKEHQIY
>seq10 | count: 1         -----SDGCDPRAVVKEAQRI-
-----HETSPRAVYQGNLCKF
>seq15 | count: 1         -----VEFSHPRAVRQPNEPF-
-----VDNSPRAVQQCIFQVF
>seq16 | count: 1         -----VWHS DPRAVRGFFDAD-
-----RLVSPRAVPPCAVEYC
>seq21 | count: 1         -----FRMSEEPRAVWQSN SY--
-----YDTSPRAVIKEACPCI
>seq24 | count: 1         -----IWASGEPRAVQELCRY--
-----IQQCPRAVVTIALHNN
>seq27 | count: 2         -----HQV SIEPRAVQHLAGD--
-----RITSEPRAVAIEHGSY-
>seq28 | count: 1         -----QGV CEDPRAVRGDTPC--
-----EDACEPRAVHCNKCHC-
>seq29 | count: 6         -----IWASGEPRAVQPEVKN--
```

```

>seq71|count: 2
-----HYACLEPRAVCQYNCC--
>seq72|count: 2
-----NVWSQDPRAVQQYEQY--
>seq74|count: 1
-----HDACSEPRAVQQTFVY--
>seq77|count: 1
-----ESRSEEPRAVQPDQVY--
>seq79|count: 1
-----DARSQEPRAVWQAYHN--
>seq80|count: 3
-----ESRSEEPRAVQGFIAAY--
>seq83|count: 2
-----TYASAVPRAVQQDLTY--
>seq84|count: 3
-----RDRCTDPRAVQRQDGI--
>seq86|count: 1
-----IWASGEPRAVQQEICY--
>seq87|count: 1
-----IWASGEPRAVQGRYNF--
>seq93|count: 1
-----ANVCQEPRAVETVQVC--
>seq98|count: 1
-----KVASNSPRAVIVQDPY--
>seq100|count: 2
-----IWASGEPRAVQFNGCI--
>seq104|count: 2
-----HQMSESPRAVWQQFQY--
>seq105|count: 1
-----KESSHEPRAVETFLYY--
>seq111|count: 1
-----HTVCHGPRAVGFPFND--
>seq121|count: 1
-----IWASGEPRAVQQFHYY--
>seq126|count: 2
--YLSLGHGWEPRAVYNY-----
>seq127|count: 11
--DIRSTEDGEPRAVAYD-----
>seq128|count: 1
--EHQSLRIQEPRAVVA-----
>seq129|count: 18
--LWHSKRDDPRAVGHY-----
>seq130|count: 1
--PHNCIFAEDPRAVFQC-----
>seq131|count: 1
--NQVSHEDQEPRAVPYI-----
>seq134|count: 3
--IHSCVGGIPRAVCPA-----
>seq135|count: 1
--LEMCNVHEEPRAVVLC-----
>seq136|count: 4
--RQHSEHNNEEPRAVIQF-----
>seq139|count: 5
--NRQCNSHDDPRAVSLI-----
>seq140|count: 2
--VNPSTYDDEPRAVDPC-----
>seq141|count: 1
--FQGSRHNDPRAVQYF-----
>seq147|count: 1
--TQHCDSHQFPRAVHFN-----
>seq152|count: 2
--VDWSPWDAEPRAVHVI-----
>seq153|count: 2
--VYPCKDQAEPRAVSYY-----
>seq156|count: 2
--QVRSHHNPEPRAVQVD-----
>seq157|count: 1
--DKHSQFYEDPRAVYRC-----
>seq158|count: 1
--TRFSYEPHEPRAVITY-----
>seq159|count: 2
--DKNSEWAAEPRAVDLY-----
>seq165|count: 3
--SNSSQHALEPRAVVLC-----
>seq169|count: 4
--RVNCEEPNEPRAVPFA-----
>seq174|count: 1
--IHHSEDNWDPRAVPFF-----
>seq177|count: 1
--RVQCHWWEEPRAVCGN-----
>seq180|count: 2
--SVRCASRAEPRAVPQY-----
>seq189|count: 1
--HVFSYKWEEPRAVYTY-----
>seq196|count: 1
--EHQSLRIQEPRAVVA-----
>seq199|count: 1
--HVGSSLQEEPRAVVQF-----
>seq201|count: 1
--WSMSAYQEDPRAVPAF-----
>seq202|count: 7
--HEGCHRVDDEPRAVCD-----
>seq203|count: 19
--KRSSWDNDFEPRAVVF-----
>seq204|count: 3
--SNFCALNQDEPRAVGY-----
>seq205|count: 5
--DVNSYDWTQEPRAVIY-----
>seq206|count: 6
--GETSPQADDSPRAVHC-----
>seq207|count: 9
--GYPSQFDEDGPRAVVY-----
>seq208|count: 12
--LIGSNKEEFEPRAVVN-----

```

```

>seq209|count: 3
-HITSYWYEAEPRAVLN-----
>seq212|count: 16
-IYSCTDANQEPRAVSF-----
>seq213|count: 4
-LPHSTDWYPEPRAVPY-----
>seq214|count: 1
-YHNCQTINDDPRAVHC-----
>seq216|count: 1
-VETSVQEWEEPRAVTF-----
>seq219|count: 1
-WRPCDHPKCDPRAVVY-----
>seq222|count: 1
-HLACWFNREEPRAVEI-----
>seq225|count: 1
-AYQSKTYYNGPRAVFY-----
>seq227|count: 1
-QRACAQDEQEPRAVIN-----
>seq231|count: 1
-EPNSNWWGEDPRAVQY-----
>seq234|count: 1
-RPMCLKGETEPRAVYY-----
>seq235|count: 1
-RPTCHHFDNEPRAVPC-----
>seq236|count: 1
-VTHSENQDLSPRAVPY-----
>seq239|count: 2
-EQQCTLQQAEPRAVAY-----
>seq244|count: 1
-IYSCTDANHEPRAVSF-----
>seq249|count: 2
-TEPCSDYHQEPRAVQF-----
>seq252|count: 1
-VESCAQGWFSRAVCA-----
>seq255|count: 4
-GVNCFPDDTEPRAVQF-----
>seq258|count: 1
-NFWCWQYQNDPRAVCA-----
>seq262|count: 7
-EQSCEQWQAEPRAVCF-----
>seq263|count: 3
-QGQCVDDQHEPRAVRF-----
>seq266|count: 2
-KNGSLLTLEEPRAVVY-----
>seq270|count: 4
-VHVSEIFNSEPRAVVY-----
>seq276|count: 1
-NYACFITLLYPRAVLA-----
>seq277|count: 2
-ENTSIDVQYEPRAVCY-----
>seq280|count: 2
-FPNSWGYAVEPRAVVA-----
>seq281|count: 1
-HIVCSIENHDPRAVYY-----
>seq289|count: 1
-TQNSLYPPHYPRAVPF-----
>seq291|count: 1
-WQTCIDFEEEPRAVYI-----
>seq292|count: 3
-KPFCVYFWEEPRAVPC-----
>seq294|count: 1
-WQQSLEWADEPRAVRA-----
>seq299|count: 1
-QPVCLILLDHPRAVLF-----
>seq300|count: 2
-QGMSQVLEIEPRAVKI-----
>seq303|count: 1
-GYPSQVDEDGPRAVVY-----
>seq306|count: 2
-EQSCAFDDRDPRAVVA-----
>seq310|count: 1
-FFQCPSFSCEPRAVKD-----
>seq315|count: 1
-NLPSNKYNDEPRAVYD-----
>seq327|count: 1
-PRSCSKPQVEPRAVVC-----
>seq336|count: 1
-ASMSTNPRKDPRAVQD-----
>seq338|count: 1
-KQFSRFVWEDPRAVHD-----
>seq343|count: 1
-EQSCEQWQAKPRAVCF-----
>seq347|count: 5
HNMCVHKLHFEPRAVC-----
>seq348|count: 1
VNSCDALENDEPRAVC-----
>seq349|count: 21
INRSDVFFDYEPRAVY-----
>seq350|count: 75
KPSSNIDAINPRAVF-----
>seq358|count: 3
HPLSVSHRGQEPRAVY-----
>seq359|count: 10
EEQCDQDILNEPRAVY-----
>seq360|count: 1
VQWSVDTYHLEPRAVF-----
>seq362|count: 1
QRFCVLYLGQCPRAVC-----
>seq363|count: 6
DNQSHIEDKEEPRAVY-----
>seq365|count: 4
KRNSLNLNHEEPRAVF-----
>seq369|count: 10
QEPCFWEVNQEPRAVC-----

```

| | |
|--------------------------------|--------------------------------|
| >seq371 count: 3 | >seq476 count: 1 |
| IDQCHSDWAYEPRAVY----- | AELCEKVHHEEPRAVY----- |
| >seq372 count: 3 | >seq477 count: 2 |
| YRMCVHDFIDEPRAVY----- | QRQCTHEWELEPRAVC----- |
| >seq376 count: 9 | >seq483 count: 1 |
| SGPSVNDDDSE PRAVF ----- | RQMCFANYDVEPRAVC----- |
| >seq380 count: 6 | >seq496 count: 1 |
| EKMCHLDERHEPRAVC----- | VRPSVIAAFEPRAVI----- |
| >seq382 count: 1 | >seq512 count: 2 |
| EDASIWRWSNEPRAVC----- | TQHSHDWWLEEPRAVN----- |
| >seq383 count: 2 | >seq522 count: 1 |
| YPRSPSNYSEEPRAVN----- | HGVCNQEFQCQ PRAVF ----- |
| >seq389 count: 2 | >seq532 count: 1 |
| PTPCYEWYEHEPRAVY----- | FVHSGPVPNFEPRAVI----- |
| >seq391 count: 2 | |
| LTTSDDHAGHEPRAVA----- | |
| >seq398 count: 1 | |
| ESHCKGQVFSEPRAVY----- | |
| >seq403 count: 2 | |
| HWACPDNDRYEPRAVC----- | |
| >seq406 count: 4 | |
| PKTSNKVWDTE PRAVF ----- | |
| >seq409 count: 2 | |
| PTFCQWPDHDEPRAVN----- | |
| >seq410 count: 2 | |
| IYVSASVADPG PRAVF ----- | |
| >seq416 count: 2 | |
| NNQSFVPRNCE PRAVF ----- | |
| >seq417 count: 1 | |
| KEWSHYHHQTCPRAVC----- | |
| >seq421 count: 1 | |
| SGLSWYYGREE PRAVF ----- | |
| >seq427 count: 1 | |
| YRCENFEAGEPRAVN----- | |
| >seq435 count: 1 | |
| TAMSYLIQEDGPRAVY----- | |
| >seq438 count: 1 | |
| GTRCIHDWDVEPRAVC----- | |
| >seq446 count: 2 | |
| YWQCYKGQIDEPRAVC----- | |
| >seq452 count: 1 | |
| YYQSKAKIHYW PRAVF ----- | |
| >seq457 count: 1 | |
| YNWCQFEKAYKPRAVY----- | |
| >seq458 count: 1 | |
| HNGCWQNFCGE PRAVF ----- | |
| >seq459 count: 1 | |
| NPSSNIDAIN PRAVF ----- | |
| >seq461 count: 1 | |
| GRQSNQQVDKE PRAVF ----- | |
| >seq462 count: 1 | |
| NSQCIHHRPWPRAVA----- | |