

## Epitope Mapping report

Date: 08.03.2016

(data from Fraunhofer IZI)

Target: Anti-human C-myc (BD Pharmingen, 551102)

Antigen: Synthetic peptide A(408) E E Q K L I S E E D L L R K R R E Q L K  
H K L E Q L R N S C A(438) of human c-Myc.

Panning conditions: Immunotubes protocol  
NGS library statistics:

Dataset:	it_c-myc_1pr (1 <sup>st</sup> panning round)
Number of sequences:	715 525
Number of valid sequences:	520 549
Pattern/Library:	ENTE1
3-mer motifs:	7 254
4-mer motifs:	132 168

Dataset:	it_c-myc_2pr (2 <sup>nd</sup> panning round)
Number of sequences:	1 061 440
Number of valid sequences:	750 183
Pattern/Library:	ENTE1
3-mer motifs:	7 246
4-mer motifs:	117 878

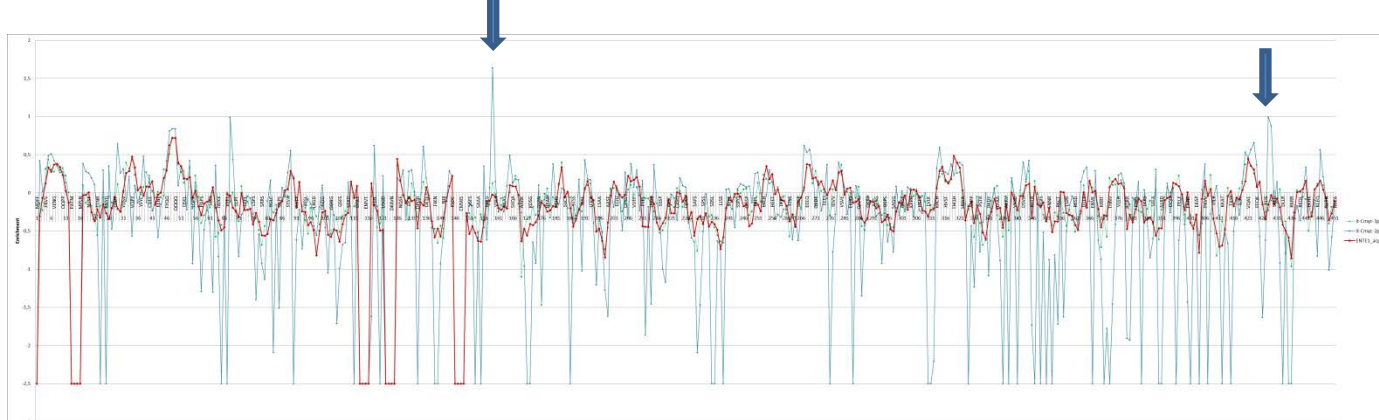
*Number of sequences* is the output from NGS

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

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 enrichment statistics

### 4mers



Data based on all selection rounds' 4mer motifs in the data sets plotted against the full length myc sequence. The reference data set (red) contains almost 2 Mio Sequences from the naïve library. The right arrow points at the sequence LISE of the peptide used in the immunisation, the left arrow points at the motif LVSE, which is ten times more enriched.

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.

### Potential Epitope:

LiSEed

Strongly enriched potential epitope variation:

LVSE

### Fingerprint

Fingerprint in Web-Logo style from all sequences containing LVSE and LISE motifs. A slight preference for the LISEEDL sequence can still be seen. The antibody is still likely to have a broad specificity regarding mimotopes as enriched and listed in the end of this data sheet.



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/>

ingerprint 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: myc proto-oncogene protein [Homo sapiens]

```
MDFFRVVENQPPATMPLNVSFTNRNYDLDYDSVQPYFYCDEEENFYQQQQSELQPPAPSEDIWKKFELLPTP
PLSPSRRSGLCSPSYVAVTPFSLRGDNDGGGSFSTADQLEMVTELLGGDMVNQSFICDPDETFIKNI I IQDC
MWSGFSAAAKLVSEKLASYQAARKDSGSPNPARGHSVCSTSSLYLQDLSAAASECIDPSVVFYPLNDSSSPKS
CASQDSSAFSPSSDLLSSTESSPQGSPEPLVLHEETPPTTSSDSEEEQEDEEIDVVSVEKRQAPGKRSESGS
PSAGGHSKPPHSPLVLRCHVSTHQHNYAAPSTRKDYPAAKRVKLDsvrvlRQISNNRKCTSPRSSDTEENVK
RRTHNVLERQRRNELKRSFFALRDQIPELENNEKAPKVVILKKATAYILSVQAEQKLISEEDLLRKRREQLKH
KLEQLRNCA
```

## List of found sequence examples:

## FASTA file contents, aligned LVSE and LISE sequences

```

>Cmyc_immunogen
-----AEEQKLISEEDL--
>ab anti-Cmyc-2pr - LVSE - seq130 | count: 12
-----FDHSQWLVSEEDCDQD-----
>ab:1 anti-Cmyc-2pr - LISE - seq16 | count: 45
-----EEPCRLISEFELLTEY-----
>ab:2 anti-Cmyc-2pr - LISE - seq56 | count: 1
-----EEPCRLISEIELLTEY-----
>ab:3 anti-Cmyc-2pr - LISE - seq62 | count: 36
-----EDHCQFLISEEQVYNI-----
>ab:4 anti-Cmyc-2pr - LISE - seq66 | count: 1
-----EDHCQFLISEEQVYNF-----
>ab:5 anti-Cmyc-2pr - LISE - seq75 | count: 2
-----WWPSQWLISEYEFFIF-----
>ab:6 anti-Cmyc-2pr - LISE - seq88 | count: 1
-----EDHCHFLISEGQVYNI-----
>ab:7 anti-Cmyc-2pr - LISE - seq12 | count: 13
----EQMSNEYLISEFHARF-----
>ab:8 anti-Cmyc-2pr - LISE - seq14 | count: 12
----VTHSWIRLISEQEFND-----
>ab:9 anti-Cmyc-2pr - LISE - seq16 | count: 4
----VQRSHAYLISENDILY-----
>ab:10 anti-Cmyc-2pr - LISE - seq115 | count: 4
----GSRCVLGLISEPTSTC-----
>ab:11 anti-Cmyc-2pr - LVSE - seq1 | count: 9
-----GLVSEVDQTGQHHQAC-
>ab:12 anti-Cmyc-2pr - LVSE - seq2 | count: 5
-----ELVSEKHIAEELDKPD-
>ab:13 anti-Cmyc-2pr - LVSE - seq3 | count: 11
-----HLVSELLFFSQQRIHA-
>ab:14 anti-Cmyc-2pr - LVSE - seq4 | count: 19
-----VLVSETDYHPWQLRSF-
>ab:15 anti-Cmyc-2pr - LVSE - seq5 | count: 24
-----GLVSEITRRCIYHAHC-
>ab:16 anti-Cmyc-2pr - LVSE - seq8 | count: 4
-----YLVSEQHYILEIADFY-
>ab:17 anti-Cmyc-2pr - LVSE - seq9 | count: 13
-----DLVSEHEKVCYNTYVF-
>ab:18 anti-Cmyc-2pr - LVSE - seq12 | count: 5
-----DLVSEPIQVSEGVHEF-
>ab:19 anti-Cmyc-2pr - LVSE - seq14 | count: 4
-----HLVSEGRSVRQVAIAF-
>ab:20 anti-Cmyc-2pr - LVSE - seq18 | count: 2
-----HLVSEERAHLPAESIIY-

```

```
>ab:21 anti-Cmyc-2pr - LVSE - seq19|count: 1
-----VLVSETDYHPWLLRSF-
>ab:22 anti-Cmyc-2pr - LVSE - seq29|count: 2
-----DLVSEFWIFAGAAQHC-
>ab:23 anti-Cmyc-2pr - LVSE - seq45|count: 7
-----GLVSEHQIHVITHCAC-
>ab:24 anti-Cmyc-2pr - LVSE - seq54|count: 2
-----VLVSEPYDPQG-ADGCI
>ab:25 anti-Cmyc-2pr - LVSE - seq67|count: 4
-----ILVSEQPFDPRGYKVF-
>ab:26 anti-Cmyc-2pr - LVSE - seq99|count: 1
-----QLVSEQVEQQCCTHPF-
>ab:27 anti-Cmyc-2pr - LVSE - seq102|count: 1
-----HLVSELLFYQQRIHA-
>ab:28 anti-Cmyc-2pr - LVSE - seq110|count: 1
-----ELVSEKHIAEEQDKPD-
>ab:29 anti-Cmyc-2pr - LVSE - seq116|count: 4
-----FQMCALVSENGVPVFD-----
>ab:30 anti-Cmyc-2pr - LVSE - seq119|count: 2
-----PFLSVLVSEFIRVYAY-----
>ab:31 anti-Cmyc-2pr - LVSE - seq122|count: 52
-----TLHSQRLVSEWYLLQF-----
>ab:32 anti-Cmyc-2pr - LVSE - seq150|count: 1
-----TLHSHRLVSEWYLLQF-----
>ab:33 anti-Cmyc-2pr - LVSE - seq187|count: 30
-----EPRCQGALVSEFEHLF-----
>ab:34 anti-Cmyc-2pr - LVSE - seq188|count: 6
-----IDSCHITLVSEFYLY-----
>ab:35 anti-Cmyc-2pr - LVSE - seq216|count: 2
-----EPHCQQWLVSEFRHLY-----
>ab:36 anti-Cmyc-2pr - LVSE - seq225|count: 2
FQRSDHEFECQLVSEI-----
>ab:37 anti-Cmyc-2pr - LVSE - seq226|count: 5
VIVSQKQWGTRLVSEN-----
>ab:38 anti-Cmyc-2pr - LISE - seq1|count: 15
-----YIVCLISESDCVYRND-----
```