Affinity maturation of a cyclic peptide handle for therapeutic antibodies using deep mutational scanning*

Authors: Martijn van Rosmalen, Brian M.G. Janssen, Natalie M. Hendrikse, Ardjan J. van der Linden, Pascal A. Pieters, Dave Wanders, Tom F.A. de Greef and Maarten Merkx

Author affiliation: Laboratory of Chemical Biology and Institute for Complex Molecular Systems (ICMS), Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

*Running title: Affinity maturation of a cyclic peptide handle for therapeutic antibodies

To whom correspondence should be addressed: Maarten Merkx, Laboratory of Chemical Biology, Eindhoven University of Technology, PO Box 513 (STO 3.22), 5600 MB Eindhoven, The Netherlands; m.merkx@tue.nl.

Keywords: directed evolution, cyclic peptide, yeast display, cetuximab, deep sequencing, antibody, protein engineering, Epidermal Growth Factor Receptor (EGFR)

ABSTRACT

Meditopes are cyclic peptides that bind in a specific pocket in the antigen-binding fragment (Fab) of a therapeutic antibody such as cetuximab. Provided their moderate affinity can be enhanced, meditope peptides could be used as a specific non-covalent and paratope-independent handles in targeted drug delivery, molecular imaging and therapeutic drug monitoring. Here we show that the affinity of a recently reported meditope for cetuximab can be substantially enhanced using a combination of yeast display and deep mutational scanning. Deep sequencing was used to construct a fitness landscape of this protein-peptide interaction and four mutations were identified that together improved the affinity for cetuximab 10-fold to 15 nM. Importantly, the increased affinity translated into enhanced cetuximab-mediated recruitment to EGFR-overexpressing cancer cells. While in silico Rosetta simulations correctly identified positions that were tolerant to mutation, modeling did not accurately predict the affinity enhancing mutations. The experimental approach reported here should be generally applicable and could be used to develop meditope peptides with low nM affinity for other therapeutic antibodies.

Introduction

The development of new antibody conjugation strategies is a fast expanding field with applications in targeted drug delivery and tumor imaging. Classical, non-site-specific coupling of drugs or dyes to primary amines or thiols carries the disadvantage of generating a heterogeneous mixture of antibodies with different numbers of attached molecules as well as different sites of attachment. Therefore, in recent years, several interesting new types of bioorthogonal chemistry have been explored for antibody conjugation (1–3). These include the ribosomal incorporation of non-canonical amino acids, chemo-enzymatic conjugation of payloads to the conserved N-glycan of native mAbs (4), the use of SNAP-tags and other self-labeling protein and peptide tags (5, 6), or light-controlled covalent attachment of protein G (7, 8). In addition to these covalent antibody conjugation strategies, Donaldson et al developed a non-covalent affinity tag for therapeutic antibodies, which they coined a ‘meditope’ (9, 10). A meditope is a peptide that binds in a large pocket between the four immunoglobulin domains...
in the antigen-binding fragment (Fab) of antibodies (Figure 1a). They identified a disulfide-constrained cyclic peptide (sequence: CQFDLSTRRLKC) that binds in this manner to the therapeutic anti-Epidermal Growth Factor Receptor (EGFR) antibody cetuximab (trade name: Erbitux). Since the binding site that this meditope recognizes is distinct from the pockets found in natural human antibodies, the tag serves as a specific non-covalent and paratope-independent handle that can be used to specifically bind cetuximab. Intriguingly, the site could be grafted onto another, humanized therapeutic antibody, (trastuzumab) by mutating 13 key interacting residues into their cetuximab counterparts.

At present, many of the potential applications for the cetuximab meditope peptide are limited by its moderate affinity. The original meditope was identified in a phage display screen of random 10-mer peptides flanked by cysteines (11). As the sequence space of such a library is $10^6$ times larger than the library that was used (12), systematic exploration of the local sequence space is likely to reveal affinity enhancing mutations. Yeast surface display (Figure 1b) is particularly suitable for such affinity maturation (13). In contrast to in vitro molecular display technologies such as phage- or ribosome display, which require panning as a means of selection, Fluorescence Activated Cell Sorting (FACS) is used in yeast display, which offers more refined control of the selection pressure (14). Yeast display is also less vulnerable to unintended selection of the most infectious or fastest-replicating clones (15). Finally, the yeast secretory pathway acts as a filter, ensuring that only well folded full-length proteins and peptides will be displayed (16).

Because subtle affinity enhancements may escape detection in classical screens for a select number of ‘hit’ sequences, recently high-throughput protein display technologies have been used in conjunction with deep sequencing before and after selection, a combination loosely referred to as ‘deep mutational scanning’ (17–19). This strategy allows the construction of fitness landscapes of the selected property (i.e. affinity) in which subtle effects of individual mutations can be discerned by the degree to which they are enriched or depleted by the selection. Deep mutational scanning has been applied successfully in combination with computational protein design for the engineering of enzyme inhibitors (20) and antibody binding proteins (21), for the evolution of antibodies (22–24) and T-cell receptors (25, 26), for epitope mapping (27, 28), engineering immune co-receptors (29) as well as for establishing protein structure-function relationships (30, 31).

In this work we used yeast display and deep mutational scanning to improve the affinity of the meditope peptide for the therapeutic antibody cetuximab by screening all single, and a select set of double amino acid substitutions. The effects of these mutations were quantified using fluorescence anisotropy titration experiments and the best variant was tested for improved cancer cell targeting. The experimental approach was also compared to in silico mutagenesis with Rosetta using the crystal structure of the meditope-cetuximab complex, to determine whether subtle affinity enhancing mutations could also be identified computationally.

**Results**

**Display of the cetuximab meditope on the yeast cell surface**

The phage display screen that identified the cetuximab meditope sequence CQFDLSTRRLKC (hereafter called Md1) also reported three other binding sequences, CVWQRWQKSYVC (Md2), CQYNLSSRALKC (Md3), and CMWDRFSRWYKC (Md4) (11). Since the cetuximab affinities of these other three peptides were not available, we synthesized fluorescently labeled versions of all 4 peptides and determined their affinity for cetuximab in a fluorescence polarization assay, yielding $K_d$ values of $134.6 \pm 5.6$ nM for Md1, $371 \pm 58$ nM for Md2 and $739 \pm 115$ nM for Md3 (Figure 2b). For Md4 no significant binding was observed up to 10 µM of cetuximab. Next, we tested whether these disulfide-constrained peptides could be efficiently displayed and cyclized by the yeast secretory machinery. Correct folding and display was analyzed for each of the four peptide sequences by using FACS analysis to study binding of fluorescently labeled cetuximab at various concentrations. Remarkably, cetuximab binding was only observed for the original meditope sequence, but not for any of the other three
sequences (Figure 2b). The affinity of the interaction between cetuximab and the meditope peptide displayed on the yeast surface seemed to be slightly weaker than that determined in the fluorescence polarization assay, but an accurate determination of $K_d$ was precluded by the lack of a plateau. Especially for weaker interactions, dissociation of the bound cetuximab during washing may result in an underestimation of the affinity. The lack of cetuximab binding observed for Md2 and Md3, which did respond in the fluorescence polarization assay, suggests that these peptides were not cyclized efficiently by the yeast secretory machinery. Yeast display has been used extensively on different types of proteins (25, 32–36) as well as cysteine-knot cyclic peptides (37), but those molecules possess a well-defined tertiary structure that will favor correct disulfide pairing during oxidative folding in the endoplasmic reticulum. Simple disulfide-constrained cyclic peptides lack such a stabilizing scaffold and they may therefore form incorrect disulfide bridges with other proteins.

**Systematic screen of single amino acid substitutions**

Having established that Md1 has the highest cetuximab affinity and is correctly displayed on the yeast surface, we made ten libraries of Md1 each harboring a single degenerate NNK codon (the cysteines were not mutated). The pooled single NNK libraries were sorted in 3 consecutive rounds for cetuximab binding. Two different selection stringencies were used in parallel using either 100 nM or 30 nM cetuximab (Figure 3a). Sequence analysis of 12 single clones obtained after the third sorting round revealed three mutations that were enriched: Q1V, S5G and K10R. These three mutant clones were cultured individually and analyzed by FACS, showing a modest increase in cetuximab binding (Figure 3b). To quantify the effect of each mutation, we synthesized fluorescently labeled meditope peptides containing either a single mutation or harboring combinations of two or three of these mutations. After verification of correct cyclization by mass spectrometry, the dissociation constants for cetuximab-peptide complexes were determined in a fluorescence polarization assay (Figure 3c, table I). The effects of the individual mutations on cetuximab binding were modest, with affinity increases of 1.2 – 1.8 fold. Fortunately, these subtle effects were fully additive, with the Q1V/S5G/K10R triple mutant showing a four-fold lower dissociation constant of $30.3 \pm 0.9$ nM.

**Constructing a local fitness landscape with next-generation sequencing**

Since the three identified mutations showed only minor improvements in affinity, we wondered whether other subtle mutations had evaded detection because of the limited number of analyzed clones. To determine the optimal amino acid at every position we applied deep mutational scanning, in which the amino acid frequencies at every position are determined by deep sequencing both before and after selection. The degree of enrichment or depletion of an amino acid reflects its positive or negative effect on the binding affinity. Using this information a ‘fitness landscape’ can also be created to identify positions that tolerate diversity better than others, which is useful for designing double mutation libraries.

Figure 4a shows the deep mutational scanning strategy used in this study (see Materials & Methods for a detailed description). Briefly, the 10 single NNK libraries were sorted separately for two consecutive rounds using 100 nM cetuximab. In the first round, a rectangular sorting gate was drawn to enrich all functional binders, regardless of their specific affinities (Figure S1a). In the second sorting round, a triangular sorting gate was drawn such that half of the expressing population of a control sample displaying Md1 fell within this gate (Figure S1b). In this way, any favorable mutants would be enriched more than the original, as a larger fraction of their expressing population would fall within the sorting gate.

DNA was extracted from samples of each unsorted and each twice-sorted population. The cyclic peptide-encoding regions were amplified by PCR, during which Multiplex IDentifier (MID) barcode sequences were incorporated. Using six different forward and six different reverse primers containing unique MIDs, each sub-library (e.g. ‘NNK10 unsorted’) was tagged with an unambiguous combination of MIDs (Figure 4a). Only after each sub-library was uniquely tagged they were combined into a single sequencing sample. Sequencing was done using the Illumina MiSeq platform. Paired-end sequencing was used, since the read-length of the Illumina method was
too short to capture the cyclic peptide sequence and both MIDs in a single read (see Materials & Methods).

From the sequence analysis, enrichment ratios were calculated. In order to minimize the effect of sampling error on the enrichment ratios, we defined the ‘robust enrichment ratio’ (RER) as a metric, which is in effect equal to the lower 95% confidence limit of the enrichment ratio (see Materials & Methods for the definition). The robust enrichment ratios were represented as a heat map (figure 4b). The only mutations with higher RERs than the original residues at their respective positions were the three previously identified beneficial mutations, Q1V, S5G and K10R. This confirmed that our deep mutational scanning strategy was able to correctly identify mutations with enhanced affinities, and suggested that no other beneficial single mutations had been previously overlooked. In addition to identifying beneficial mutations, the analysis also revealed that some positions were more tolerant to mutation than others. Notably, at positions 1, 3 and to some extent 5 and 6 many residues had RERs above 1, meaning that these mutants retained at least significant binding. In contrast, substitutions at positions 2, 7, 8 and 9 abrogated binding completely. At position 4, only one mutation (L4M) had an RER above 1, but otherwise this position was also intolerant to mutation.

**Affinity maturation using double amino acid substitution libraries**

The effects of the three mutations that had been identified thus far were largely additive, but screening multi amino acid substitution libraries could identify mutations that are cooperative, i.e. the effect of one mutation is enhanced by the presence of another mutation. Deep mutational scanning of the single mutant libraries revealed that positions 1, 3, 5 and 6 tolerated diversity. We therefore created five double NNK libraries, harboring degenerate codons at positions 1+3, at 1+5, 1+6, 3+5, and 5+6. These double amino acid substitution libraries were prepared, sorted and sequenced in the same way as the single NNK libraries (Figure S1). Analysis of the sequencing data confirmed that mutations Q1V and S5G were more strongly enriched compared to the respective original amino acids in the libraries NNK 1+3, 1+5, 1+6 and 3+5 (Figure 5a). Only in NNK 5+6 the original serine at position 5 was more enriched than glycine. Two additional mutations, D3N and T6I were found that had high RERs. In both the NNK 1+3 and NNK 3+5 libraries an aspartate was more enriched at position 3 than the original asparagine, which suggested that the D3N mutation is only favorable for binding in the context of other mutations at positions 1 and/or 5. Other potentially beneficial mutations that only appeared in the double amino acid substitution libraries are T6I and T6V. These both had higher RERs (2.5 and 2.4 respectively) than the original threonine (RER 1.4) in the NNK 1+6 library. In the NNK 5+6 library, the threonine at position 6 had a higher enrichment ratio (8.5) than isoleucine (3.5) or valine (2.4). This suggests that mutating threonine 6 to a hydrophobic valine or isoleucine is favorable only in the context of a second mutation at position 1, but is disfavored to some extent in the context of another mutation at position 5. In addition to the enrichment of amino acids, we also assessed the enrichments of specific pairs of mutations, to determine which residues were frequently selected together (Supplementary Figure S2 – S6). From these analyses, mutations D3N and T6I also emerged as likely candidates that could further improve the affinity.

To assess whether the D3N and T6I could further enhance the affinity of the Q1V/S5G/K10R meditope variant, three fluorescently labeled peptides were synthesized, carrying D3N, or T6I, or both in addition to the Q1V, S5G and K10R mutations. Fluorescence polarization assays showed that each of the single mutations enhanced cetuximab binding by a factor of 2, yielding $K_d$ values of 14.5 ± 0.8 nM and 16.8 ± 2.1 nM for the tetramutants Q1V/D3N/S5G/K10R and Q1V/S5G/T6I/K10R, respectively. A pentamutant peptide carrying Q1V/D3N/S5G/T6I/K10R had a dissociation constant of 15.8 ± 3.9 nM. Apparently, the two additional mutations indeed improved the affinity further by a factor 2, but their effects were not additive. These combined mutations thus yielded an almost 10-fold increase in affinity.

**Comparing the local fitness landscape with in silico Rosetta simulations**

The availability of the crystal structure of the complex between the original meditope peptide (Md1) and cetuximab (9) provided an opportunity...
to compare the fitness landscape determined using deep mutational sequencing with one obtained based on in silico mutagenesis. Such a comparison would serve as a rigorous test to assess whether in silico modeling could substitute for experimental screening. The well-established Rosetta protein design software was used for in silico modeling, which employs an empirical score function to find the lowest energy conformation of mutant proteins or peptides (38, 39). Using the published crystal structure as a starting point, all single amino acid substitutions were introduced in the peptide, and the binding of each mutant was compared to the original meditope. This strategy is effectively a computational counterpart of the single NNK screen carried out with deep mutational scanning.

Figure 6a shows the results of the computational screen as a heat-map depicting the difference in binding scores between the mutants and the original peptide (Figure 6a). In this heat map, negative values (black or dark gray areas) represent binding scores that were larger than that of the non-mutated peptide, suggesting affinity-improving mutations. Highly positive values (white or light gray areas) on the other hand indicated a decreased affinity. The simulations globally reflected the experimental observations that positions 1, 3, 5 and 6 are most tolerant to mutation. The computational modelling suggested that especially positions 4, 7 and 8 and to a lesser extent positions 2 and 9 were intolerant to mutation, which is also in agreement with experimental results. In the model position 10 was more tolerant to mutation than observed experimentally.

Of the three beneficial mutations that were experimentally selected, only K10R was correctly identified as an improving mutation by Rosetta, while Q1V and S5G did not have higher interface scores than the original peptide. The model suggested several other mutations with high interface scores, in particular mutations of Serine 5 into any of the three aromatic residues. Given the fact that positions 5 and 6 are both tolerant to mutation, adjacent, and situated deeply within the pocket where there is some empty space, mutations at position 6 in combination with aromatics at position 5 were also explored (Figure 6b). This analysis suggested that methionine at position 6 could improve the affinity in the context of an additional S5Y or S5F mutation, but not in the original background. Aromatic residues at position 5, especially in combination with the methionine at position 6, would better fill the empty space in the back of the binding pocket and pack against the constant domain of cetuximab (Figure 6c). The side chain of a tyrosine at position 5 can even come into hydrogen-bonding distance of the backbone carbonyl of valine 152 (Kabat numbering) (40) of the C_{H1} domain (Figure 6d). Besides aromatics at position 5 the model also predicted favorable binding for mutation D3R. The side chain of an arginine mutation at position 3 seemed to occupy a considerable amount of empty space and pack against the constant domains (Figure 6e), forming hydrogen bonds with the side chain of glutamine 150 and with the carbonyl of alamine 176, both in the C_{H1} domain of cetuximab (Figure 6f).

Peptides carrying these single mutations or combinations of two or three mutations (in the background of the Q1V/S5G/K10R triple mutation) were made and their affinities were measured by fluorescence polarization (Figure 6g). None of these peptides had improved affinities, however. The peptide carrying S5Y had a $K_d$ of 86.3 ± 6.9 nM. Compared to 28.2 ± 1.8 nM for the Q1V/S5G/K10R peptide or to 63.2 ± 1.2 nM for the peptide carrying only Q1V and K10R (Figure 3d) this was clearly not the improvement that the Rosetta simulations suggested. Similarly, the T6M mutant had a $K_d$ of 162 ± 7 nM and the combination of these two mutations – which the modeling predicted to be highly cooperative – resulted in a $K_d$ of 487 ± 149 nM. The D3R mutation had a reduced affinity ($K_d$ of 186 ± 29 nM) and for a peptide with Q1V/D3R/S5Y/T6M/K10R) no binding was observed. Taken together, the in silico modeling was able to identify positions that were amenable to mutation, but was much less successful in predicting mutations that improved the affinity. Several other studies also found that while in silico design methods can be used effectively to provide a global insight into mutable residues and even provide the first important step in the de novo design of interaction interfaces, these designs invariably require experimental affinity maturation to fully optimize the binding strength (20, 21). Our results show that this need for experimental affinity maturation persists even for concave
binding pockets with structurally well characterized peptide ligands.

**Meditope targeting to EGFR overexpressing cancer cells in a cetuximab-dependent manner**

Having improved the affinity of the cetuximab meditope peptide from 130 nM to 15 nM, we assessed whether the increased affinity also translates into improved cetuximab-mediated EGFR targeting on cancer cells. A431 skin epidermoid carcinoma cells, which show highly elevated EGFR expression levels, were incubated with cetuximab and FITC-conjugated meditope peptides (Figure 7). Cells stained with the tetramer Q1V/D3N/S5G/K10R were almost an order of magnitude more fluorescent than those stained with the original meditope (Figure 7a). In agreement with observations by Donaldson *et al.* (9), significant binding of the original meditope peptide was only observed at a peptide concentration of 500 nM. In contrast, the mutant meditope showed significant binding at 50 nM and the fluorescence intensity at 100 and 500 nM was consistently higher than for the original peptide (Figure 7b). The substantially lower concentration of meditope peptide that is required enhances the feasibility of the non-covalent cetuximab ‘labeling’ approach for applications in molecular imaging and targeted drug therapies.

**Discussion**

In this study we showed that the affinity of a cyclic peptide meditope for the therapeutic antibody cetuximab, originally obtained from phage display screening, could be improved 10-fold using a combination of yeast display and deep mutational scanning. Construction of an accurate map of the fitness landscape of the meditope-antibody interaction allowed identification of 4 mutations that by themselves increased the affinity only subtly, but together increased the affinity from $K_d$ 130 nM to 15 nM. This affinity increase also translated into an order of magnitude more efficient cetuximab-mediated targeting of EGFR-overexpressing tumor cells. The strategy developed in this work can be more generally applied for affinity maturation of protein-protein or protein-peptide interactions, for example to develop peptide meditopes targeting other antibodies.

Most previously described studies have used targeted library screening and deep sequencing for epitope mapping or to identify which positions in a protein are functionally important (27, 28, 30). Some also used deep mutational scanning with the explicit goal of improving the affinity, however. Forsyth *et al.* used mammalian display and Roche 454 sequencing to improve the affinity of a humanized version of cetuximab for EGFR, realizing a 4-fold improvement and correctly identifying many affinity-improving mutations (22). Reich *et al.* successfully used yeast display and NGS to rank over a thousand BH3 peptides for their affinity to the anti-apoptotic protein Bcl-Xl (14). Koenig *et al.* have used phage display with deep mutational scanning of single and triple mutation libraries to generate a Fab fragment with sub-nanomolar affinities to two unrelated antigens (24).

In previous studies that use systematic single site saturation mutagenesis, sub-libraries were typically pooled before the selection (17–19, 22, 23, 28, 29). This approach causes a problem in identifying which amino acid is the best choice at each position, however. Sequences harboring mutations at positions that tolerate diversity will be more strongly enriched than the corresponding wild-type residues simply because they will not harbor deleterious mutations at intolerant positions. Therefore, if libraries are pooled before selections, mutations with higher enrichment ratios than the wild-type residue do not necessarily confer increased affinity. This effect becomes especially pronounced if the number of randomized positions in the library is small. To work around this problem we performed the selections of the different single substitution libraries separately and employed a DNA barcoding scheme for identification of sequencing reads. This strategy carried an additional advantage as it allowed us to quickly determine which positions tolerated diversity better than others by a simple FACS analysis. This information is important for designing multi amino acid substitution libraries. Separate selections are admittedly more labor intensive than pooled ones, especially if the number of single-site libraries increases, but adjustment of the barcoding strategy could help to remedy this issue. A possible
strategy for larger proteins could be to incorporate unique barcode sequences within the plasmid to identify post-sequencing from which library each read was derived.

Another methodological issue that we addressed is how to deal with sampling error. Enrichment Ratios are usually calculated by dividing the frequency of a particular mutation in the selected population by the corresponding frequency in the unselected library (17–19). However, it is easier to enrich rare amino acids than frequent ones. If a mutation of interest is rare in the unselected library (e.g. counted only 30 times), sampling error can lead to relatively large overestimations of the true Enrichment Ratio. We found the effect of sampling error to be particularly strong when analyzing the enrichment of pairs of amino acids in the double NNK libraries. Many pairs occurred less than 100 times and, although the normal Enrichment Ratio gave a qualitative indication of which pairs were enriched, quantitatively it was unreliable. Some pairs had absurdly high enrichments of over 100-fold. We therefore defined a new parameter called the 'Robust Enrichment Ratio', which corrects for sampling errors (see Materials & Methods). In effect this metric is the Enrichment Ratio's lower 95% confidence limit. The Robust Enrichment Ratio is a useful metric in all cases where sampling error plays an important role (i.e. when individual clones are read less than 100 times). This is the case 1) when studying the effects of pairs of mutations in double NNK libraries, 2) when scanning large proteins and 3) when different deep sequencing technologies are used (e.g. 454 sequencing, which yields longer reads, but substantially lower coverage).

Several interesting applications could be explored using the meditope technology. Conjugation of the peptide to a dye or MRI contrast agent could be used for in vivo (tumor) imaging in patients undergoing antibody treatment. Non-covalent antibody conjugates have been explored by several groups (41–43), mostly making use of generic Fc-binding protein domains (e.g. S. aureus protein A). Contrary to covalently attached agents, a non-covalent agent would exhibit fast pharmacokinetic clearance rate in the free form, while the rate of clearance of the complex would essentially be dictated by the antibody. Thus, the conjugated form would remain in the body for a more extended time period, which would be desirable for imaging. Using a peptide handle specific for the therapeutic antibody, rather than protein A would prevent the dye or contrast agent from binding to endogenous IgG1 in the bloodstream before it can reach the target site.

Another application that could benefit from an antibody handle is testing patients for the presence of Anti-Drug Antibodies (ADAs). The presence of ADAs indicates an immunogenic response to the drug and usually results in the drug being ineffective. To detect ADAs, the drug-ADA complexes must be captured intact from plasma, which is rich in background IgGs. While the target protein of the drug can be used as a bait for this purpose, many ADAs are raised against the antigen binding site. These will dissociate during the capture of the drug and evade detection. Avery et al recently reported the construction of a very potent non-covalent Cetuximab-ligand by fusing the meditope peptide Md1 to a fragment of Protein L, a domain originating from Finegoldia Magna (44) that interacts specifically with immunoglobulin κ-light-chains (45). Fusion to protein L increased the affinity for cetuximab by three orders of magnitude. Although the affinity increase realized in this study was smaller, in our case the affinity enhancement did not rely on fusions with bacterial proteins. In addition to the fact that protein L fusions will also bind to endogenous IgGs, Protein L is highly immunogenic (46–48), and fusing it to the meditope will probably be detrimental for in vivo applications. The meditope technology could be a general strategy for non-covalent site-specific antibody binding. It has already been shown that meditope binding site of cetuximab can be grafted on other therapeutic antibodies (9). Moreover, pockets of similar dimensions are present in the Fab regions of all IgGs. Having shown that cyclic peptides with low nM affinity can be developed for the cetuximab binding site, it is likely that meditope sequences specific for other therapeutic antibodies can be developed using the approach presented here.

Materials & Methods
Molecular biology: The vector pCTCON2 was a kind gift from prof. dr. Dane Wittrup (49). A double stop codon was introduced between the NheI and BamHI sites within this vector using the QuickChange II Site-directed mutagenesis kit (Agilent Genomics) and the oligonucleotides pCT-Stop-F and pCT-Stop-R (exact sequences in Supplementary Table S1) according to the manufacturer’s instructions. The vector was linearized by digesting 3 µg of DNA with 60 units NheI (New England Biolabs; NEB) and 60 units BamHI (NEB) in 60 µL CutSmart buffer (NEB) at 37 °C for 1 hour. Linear vector DNA was purified by electrophoresis using a 1% agarose/TAE gel and extracted from the gel using the QIAquick Gel Extraction kit (QIAGEN).

Inserts for the Md1, Md2, Md3 and Md4 peptides were made as follows: 200 nM (final concentration) of each of four primers was mixed in a total volume of 50 µL, containing 25 µL 2× Phusion High Fidelity PCR Master Mix with HF buffer (NEB). Two of the primers, Constant-F and Constant-R, were used for all inserts and two primers were insert-specific (e.g. ‘Md1-F’ and ‘Md1-R’). The reaction was performed by initially heating at 98 °C for 2 minutes, 5 subsequent cycles of denaturation at 98 °C for 10 seconds and annealing / extension at 72 °C for 30 seconds and a final extension at 72 °C for 10 minutes. 1 µL of the product was used as a template in a 25 µL reaction containing 25 µL 2× Phusion High Fidelity PCR Master Mix with HF buffer (NEB) and 500 nM Constant-F and Constant-R primers. Thermo-cycling was done as described above, but 35 cycles were used instead of 5. The products were purified using the QIAquick PCR Purification kit (QIAGEN). Inserts of libraries were prepared in a 2 step manner: 1 µM of primers Constant-F, Constant-R and one library specific primer (e.g. ‘NNK1’) were mixed in a 50 µL reaction containing 25 µL 2× Phusion High Fidelity PCR Master Mix with HF buffer (NEB). The reaction was done in 10 cycles of denaturation at 98 °C for 10 seconds, annealing at 60 °C for 30 seconds and extension at 72 °C for 10 seconds, and a final elongation step at 72 °C for 1 minute. 1 µL of the reaction was used as a template for a second reaction containing 1 µM Constant-F and Constant-R primers and 50 µL 2× Phusion High Fidelity PCR Master Mix with HF buffer (NEB) in a total volume of 100 µL. Cycling was done as for the first reaction, but 35 cycles were used instead of 10. These inserts were purified using the QIAquick PCR Purification kit (QIAGEN).

Yeast Transformation: pCTCON2 plasmids and inserts were transformed into yeast strain EBY100 using the LiAc/PEG/ssDNA method as described (50–52). For clonal inserts and single NNK libraries the protocol for transformation of frozen yeast was followed (50, 51). EBY100 was cultured in YPD medium, 2% D-(+)-Glucose (Sigma Aldrich) 2% Peptone from Casein (Merck Millipore) and 1% Yeast Extract (Merck Millipore) at 30 °C for 16 hours. The pre-culture was diluted in 400 mL fresh 2×YPD medium at an Optical Density (OD600) of 0.5 and cultured at 30 °C shaking at 220 rpm until the OD600 reached 2.0. Cells were then pelleted for 30 min. at 3000 ×g, washed twice, first with 200 mL ice-cold sterile water and then with 80 mL ice-cold sterile water, pelleting at 3000 ×g each time. The cells were re-suspended in 10% dimethyl-sulfoxide (DMSO), 5% glycerol in sterile water. 50 µL aliquots (containing ca. 3×1010 cells) were frozen at -80 °C for long term storage. For each transformation reaction, an aliquot was thawed and centrifuged at 13400 ×g for 30 seconds and the supernatant was replaced with 346 µL of ice-cold Transformation Buffer and yeast was kept on ice until heat-shock. Transformation Buffer was always prepared on the day of use by mixing (per reaction) 260 µL of a 50% (w/v) sterile aqueous solution of Polyethylene Glycol (PEG) Mw 3350 (Sigma Aldrich), 36 µL sterile 1 M LiAc (Sigma Aldrich) and 50 µL of 2 mg mL⁻¹ single-stranded DNA solution. DNA from salmon testes (Sigma Aldrich) was dissolved in 10 mM TrisHCl (Sigma Aldrich), 1 mM Ethylene-diaminetetraacetic Acid (EDTA; Sigma Aldrich) pH 8.0, filter-sterilized, denatured at 95 °C for 10 minutes, chilled on ice and immediately used or stored at -20 °C to prevent re-anneling. 100 ng of linearized vector DNA and 30 ng insert DNA were mixed with the yeast and Transformation Buffer and water was added to make a total volume of 360 µL. Yeast were then heat-shocked at 42 °C for 50 minutes with regular inversion to prevent settling of the cells. After the heat-shock, the cells were pelleted at 13400 × g for 2 minutes and re-suspended in 200 µL sterile water and plated on SD-CAA agar (2 g L⁻¹ D-(+)-Glucose (Sigma Aldrich), 6.7 g L⁻¹ Difco Yeast Nitrogen Base without Amino Acids
Running title: Affinity maturation of a cyclic peptide handle for therapeutic antibodies

(1) Bacto Casamino Acids (BD Biosciences), 5.4 g L\(^{-1}\) \(\text{Na}_2\text{HPO}_4\) (Calbiochem), 8.56 g L\(^{-1}\) \(\text{NaH}_2\text{PO}_4\cdot\text{H}_2\text{O}\) (Calbiochem), 15 g L\(^{-1}\) Agar (Merck Millipore). Colonies were counted after 3 days of incubation at 30 °C. For clonal cultures, a single colony was picked, while for single NNK libraries, colonies were gently dislodged with a sterile Drigalski spatula and the complete libraries were kept in liquid culture. For double NNK libraries the high-efficiency transformation protocol was followed (52). EBY100 was cultured and washed as described before, but were not frozen. Instead 3×10\(^{10}\) cells were immediately re-suspended in 346 µL Transformation Buffer. 1 µg of linear vector DNA and 30 ng of the respective inserts was added and the volume was made up to 360 µL with sterile water. From there, the transformation was done as described above. Double NNK libraries were not plated after transformation, but were re-suspended in 5 mL of liquid SDCAA medium pH 4.5 (as the SDCAA plates described above, but without agar and with 3.92 g anhydrous citric acid (Amresco) and 6.07 g trisodium citrate dehydrate (Calbiochem) instead of the phosphate salts). 0.1% of each library was plated onto SDCAA agar to determine the library size and colonies were counted after 3 days at 30 °C.

Yeast Display and FACS: For yeast display, the protocol by Chao et al (49) was used as a reference. Clonal yeast cultures and libraries were inoculated at an initial OD\(_{600}\) of 0.01 in 5 mL SDCAA medium in sterile 14 mL Falcon culture tubes with snap-caps (BD Biosciences) and cultured for 16-20 hours at 30 °C and shaking at 220 rpm. Cultures were then diluted to an OD\(_{600}\) of 0.1 in fresh SDCAA medium and cultured until the OD\(_{600}\) reached 0.4. Then, cells were centrifuged at 3000 × g and re-suspended in 5 mL SGCAA medium, which was prepared as SDCAA, but with the glucose replaced by 2% D-(+)-Galactose (Sigma Aldrich). The yeast was then cultured for another 2 days at 20 °C shaking at 220 rpm to induce peptide display. Cetuximab (Erbitux, Merck) was obtained from the Catharina hospital (Eindhoven, the Netherlands) and conjugated with Alexa Fluor 647 using an amine-reactive N-Hydroxy-Succinimide (NHS) ester. The coupling reaction was done in PBS at room temperature using a 20 fold molar excess of the dye. Excess of dye was removed using a PD-10 buffer exchange column (GE Healthcare) and the conjugate was concentrated using an Amicon centrifuge concentrator with a molecular weight cutoff of 50 kDa. The concentration and dye-to-antibody ratio were determined by UV-VIS absorbance at 280 and 633 nm using a NanoDrop ND1000 spectrophotometer and the molar extinction coefficients of 210,000 M\(^{-1}\)cm\(^{-1}\) at 280 nm and 264,000 M\(^{-1}\)cm\(^{-1}\) at 651 nm. The dye-to-antibody ratio was ca. 4. Yeast displaying the cyclic peptides were washed with 1 mL PBSF; Phosphate Buffered Saline (PBS) pH 7.4 with 1 mg mL\(^{-1}\) Bovine Serum Albumin (BSA) and incubated with 1:250 mouse anti-HA monoclonal antibody (Thermo Fisher Scientific) diluted in PBSF for 30 minutes at room temperature. The cells were then washed with 1 mL ice-cold PBSF and incubated for 20 minutes on ice with 30 or 100 nM Alexa Fluor 647 conjugated cetuximab and 1:100 Alexa Fluor 488 conjugated goat anti mouse monoclonal antibody (Thermo Fisher Scientific) in PBSF. Cells were washed once more with 1 mL ice-cold PBSF and kept on ice until FACS analysis or sorting. Cells were analyzed and/or sorted using a FACS Aria III Fluorescence Activated Cell Sorter (BD Biosciences) with a 70 µm nozzle and FACS Flow sheath fluid (BD Biosciences) at a pressure of 70 psi. Forward and Side Scatter (FSC and SSC) as well as Alexa Fluor 488 were excited using a 488 nm laser. SSC was detected using a 488/10 BP filter and Alexa Fluor 488 was detected using a 530/30 BP filter. Alexa Fluor 633 was excited using a 633 nm laser and detected using a 660/20 BP filter. Yeast cells were selected in a scatter plot and doublet discrimination was done in both an FSC and SSC plot of signal width vs. height. First round sorting was done using the purity mode ‘yield’, while subsequent sorting was done in ‘4-way purity’ mode. For the separate selections of single and double NNK libraries described in Figures 4 and 6, at least 500,000 cells were selected in each round for libraries NNK 2, 4, 7-10 and at least one million cells were selected each round for all other libraries. Selected cells were centrifuged at 3,000 × g for 5 minutes to remove the sheath fluid and then cultured at 30 °C for 1 or 2 days in SD-CAA, after which DNA was extracted for deep sequencing (see below). For the selections of the pooled libraries described in Figure 3 5×10\(^5\) and 3×10\(^5\) cells were selected respectively in the first and second round, both in
‘4-way purity’ mode. A third selection was done with the purity mode set to ‘single cell’, in which 100 cells were spotted directly on an SD-CAA agar plate, with only a single cell spotted per position on the plate. Colonies were allowed to grow at 30 °C for 3 days. 12 colonies were picked and grown in 5 mL liquid SD-CAA medium for 1 day. The cells were lysed and plasmid DNA was extracted by using the Zymoprep Yeast Plasmid DNA Miniprep II (Zymo Research). To improve the quality of the DNA, it was transformed into NovaBlue (Merck Novagen) bacteria, plated on LB agar with 100 mg L⁻¹ Ampicillin. DNA was isolated from the bacteria using the QIAprep Miniprep kit and the DNA was sequenced by StarSEQ GmbH (Mainz, Germany) using the primer Seq-pCON-F.

**Deep Sequencing Library Preparation:** From each library between 10⁸ and 10⁹ cells were collected and DNA was extracted using the Zymoprep Yeast Plasmid DNA Miniprep II kit (Zymo Research). The eluates were diluted 20-fold in autoclaved MilliQ water and 1 µL of each were used as templates in separate 50 µL PCR reactions containing 500 nM of YD-tag-F and YD-tag-R primers, 200 µM deoxynucleotides (dNTPs; NEB) 3 units of Phusion High Fidelity DNA polymerase (NEB) in 1× Phusion HF Buffer (NEB). Reactions were initially heated to 98 °C for 30 seconds, then cycled 24 times between 98 °C for 10 seconds, 64 °C for 30 seconds and 72 °C for 10 seconds, and finally extended at 72 °C for 30 seconds. 1 µL of each of the products of these reactions was used as template in a second 50 µL PCR in which each library was amplified with a unique pair of primers (see supplementary table S3). The conditions of this reaction were 500 nM forward and reverse primers, 200 µM dNTPs (NEB) and 3 units of Phusion High Fidelity DNA polymerase in 1× Phusion HF buffer. These reactions were initially denatured at 98 °C for 30 seconds, then cycled 25 times between heating at 98 °C for 10 seconds and annealing / extension at 72 °C for 40 seconds and finally extended for 1 minute at 72 °C. The products of these reactions were separately purified by electrophoresis through a 3 % agarose/TAE gel and extracted from the gel using the QIAquick Gel Extraction kit (QIAGEN). 1 µL of each PCR product was used as a template in a third 50 µL PCR containing 500 nM Amp-F and Amp-R primers, 200 µM dNTPs, and 3 units of Phusion High Fidelity DNA polymerase (NEB) in 1× Phusion HF Buffer (NEB). This reaction was heated at 98 °C for 30 seconds, then cycled 25 times between denaturation at 98 °C for 10 seconds, annealing at 59 °C for 30 seconds and extension at 72 °C for 10 seconds and finally extended at 72 °C for 30 seconds. The products of these reactions were purified using the QIAquick PCR Purification kit (QIAGEN) and eluted in TE buffer supplied in the kit. The concentrations of the eluates were measured by UV absorbance spectroscopy using a NanoDrop ND100 spectrophotometer and the size and purity of the final products was confirmed by electrophoresis through a 3% agarose/TAE gel. Equimolar amounts of each library were combined into a single sequencing sample. 2×150 Paired End sequencing was performed by Eurofins Genomics GmbH (Ebersberg, Germany) using the Illumina MiSeq platform and V3 chemistry. Due to the low complexity of the sample, up to 30% PhiX control had to be spiked in. Sequencing returned a total of 19,750,002 paired-end reads.

**Deep Sequencing Data Analysis:** Data was processed in several consecutive steps. First, the reads were clustered by their combination of barcode sequences using a dedicated Python script (Supplementary file ‘NGSRead.py’). Then, sequences with insertions, deletions, Phred scores below 30 (within the stretch encoding the cyclic peptide) or with unintended additional mutations were discarded and the sequences that passed these filters were translated, using two Matlab scripts (Supplementary files ‘ScoringScript.m’ and ‘removeNonWildTypes.m’). After subtraction of reads that did not fulfill the criteria, each sorted and unsorted sub-library still contained over 100,000 reads (Supplementary table S3). Subsequently, the amino acids at the mutated position in each library were counted. The amino acid frequency distributions in the unsorted single NNK libraries was compared to the theoretically expected distribution (Supplementary table S4). Certain amino acids were significantly over- or underrepresented, which may be caused by bias in the PCR step on the Illumina sequencing chip (53–55). Nevertheless, each individual residue was counted at least 1000 times at every position in the unsorted single NNK libraries (one exception being histidine at position 9, which was counted 738 times). For the double NNK libraries a more...
detailed analysis was made where the enrichment of each pair of amino acids was investigated. Amino acid pairs that were counted less than 10 times in the unsorted libraries were excluded from this analysis.

Robust Enrichment Ratio definition: The enrichment Ratio is commonly used as a metric for the degree to which a mutation contributes to the property that is selected for. It is defined as

\[ ER_{r,x} = \frac{f_{r,x}^s}{f_{r,x}^u} \]  \hspace{1cm} (Eq. 1)

Where \( f_{r,x}^s \) is the frequency of amino acid \( r \) at position \( x \) in the selected population and \( f_{r,x}^u \) is the corresponding frequency in the unselected population. The frequencies are calculated by dividing the number of times that amino acid \( r \) is counted at position \( x \) (\( n_{r,x} \)) by the total number of counts at that position.

\[ f_{r,x} = \frac{n_{r,x}}{\sum n_{r,x}} \]  \hspace{1cm} (Eq. 2)

This definition can also apply to a pair of amino acids at two positions. The frequency in the unselected population is usually a very small number, particularly if pairs of residues are investigated. When we calculated enrichment ratios of amino acid pairs in this way, sampling error lead to quite drastic overestimation of the true enrichment ratios of some pairs. To determine a relation between the magnitude of the sampling error and the number of counts in the distribution (\( n_{r,x} \)), we performed 60 Bootstrap analyses on a set of 20 simulated frequency distributions with varying number of counts \( n_{r,x} \) between 1 and 100 and a total number of reads of 10,000 or 100,000 (Supplementary Figure S7). We determined that the Coefficient of Variation (CV) of the Bootstrap analyses depended on \( n_{r,x} \) as:

\[ CV_{r,x} = \frac{1}{\sqrt{n_{r,x}}} \]  \hspace{1cm} (Eq. 3)

Therefore, we defined the Robust Enrichment Ratio (RER) as:

\[ RER_{r,x} = \frac{f_{r,x}^s - 2CV_{r,x}^s f_{r,x}^s}{f_{r,x}^u + 2CV_{r,x}^u f_{r,x}^u} \]  \hspace{1cm} (Eq. 4)

The RER can be thus regarded as the lower 95% confidence limit of the enrichment ratio under the assumption that sampling is unbiased.

**Peptide synthesis:** All peptides were synthesized by standard Fmoc Solid phase peptide synthesis (SPPS) on a 50 or 100 µmol scale on an automated peptide synthesizer (Intavis Multiprep RS) using a NovaSyn TGR resin (Novabiochem) as a solid support. Each coupling was done 2 times for 30 minutes each. Peptide 1 (see Supplementary table S2 for numbering) was fluorescently labeled on resin at the N-terminus using a 5-fold molar excess of 5-(6)-carboxyfluorescein (FAM) for 2×8 hours, essentially as described before (56). The other peptides were manually Fmoc-deprotected 2×3 minutes in 20% piperidine in N-Methyl-2-Pyrolidone (NMP) and washed 4 times with NMP. An Fmoc-protected Pentanoic acid (O1Pen) linker (Iris Biotec) was activated by mixing 2 volumes 0.2 M linker in NMP with 1 volume 0.33 M 2-(1H-Benzotriazol-1-yl)-1,1,3,3-Tetramethyluronium Hexafluorophosphate (HBTU) in NMP, incubating for 1 minute and then adding 1 volume 1.6 M N,N-Diisopropylethylamine (DIPEA) in NMP. An 8-fold molar excess of activated O1Pen linker was coupled on resin to the free amine of the peptides for 60 minutes. Peptides were then washed and Fmoc-deprotected as described above. Fluorescein Isothiocyanate (FITC) was coupled to the O1Pen linker using a 7-fold molar excess of FITC (Sigma Aldrich) and a 14-fold molar excess of DIPEA and reacting overnight. Beads were washed 4 times with NMP, then 5 times alternating between Diethylether and Dichloromethane (DCM) and then swollen in DCM. Cleavage from the resin was done for 2 hours in 95% Trifluoroacetic Acid (TFA), 2.5% Triisopropylsilane (TIS), 2.5% Ethanedithiol (EDT), 2.5% water. The peptides were precipitated in ice-cold diethylether, repeatedly washed and then dissolved in a water acetonitrile (ACN) mixture. Water / ACN ratio was varied between 30 and 50 % depending on peptide solubility. The peptides were then lyophilized. Cyclization of peptides 1-12 (Supplementary table S2) was done in 1-3 M Guanidinium Chloride (GdmCl), 100 mM Tris-HCl pH 8.0 at a peptide concentration of 1 mg mL\(^{-1}\) for several days stirring in an open tube at 4 °C. Excessive salts were removed using a solid phase extraction column (Strata XL, Phenomenex), from
which the peptides were eluted in 70% (v/v) water / 30% (v/v) acetonitrile. Peptides 13-22 (supplementary table S2) were cyclized in 69% (v/v) Phosphate Buffered Saline pH 7.4 / 30% (v/v) Acetonitrile, containing 1% (v/v) dimethylsulfoxide (DMSO) as a mild oxidizing agent (57) for several days in an open tube at 4°C. Reversed phase HPLC was performed on a Shimadzu LC-8A HPLC system by using a VYDAC protein and peptide C18 column.

**Fluorescence Polarization Assays:** Cetuximab was obtained from the Catharina hospital pharmacy, Eindhoven, The Netherlands, and was titrated to 10 nM FITC-labeled peptides in Phosphate Buffered Saline pH 7.4 (PBS, Sigma Aldrich) containing 1 mg mL⁻¹ Bovine Serum Albumin (BSA, Sigma Aldrich) in total sample volumes of 50 µL in a black 384 wells OptiPlate (Perkin Elmer). Polarization was recorded using a TECAN Infinite F500 Plate Reader. Experiments were performed in duplo or in triplo as indicated in the respective figure legends. Hill equations of the form of Eq. 5 were fit through the data using Origin 2015 (Originlab), using statistical weights for the fit.

\[
P = P_{\text{min}} + \frac{(P_{\text{max}}-P_{\text{min}}) \times [\text{cetuximab}]^n}{(K_d^n + [\text{cetuximab}]^n)} \quad (\text{Eq}. \ 5)
\]

Where P is the polarization and n is the Hill coefficient. Hill coefficients did not deviate much from one, indicating no cooperativity, which was also expected.

**Rosetta modeling:** The crystal structure of cetuximab in complex with the Md1 peptide was obtained from the Protein Data Bank (PDB ID: 4GW1) (9). The structure was refined using the Rosetta relax application using the score12 score function. A scan of the influence of mutations was executed utilizing the Rosetta backrub application, which allows optimization using a flexible peptide backbone. (39) Mutations were systematically introduced at the desired positions and the conformation was optimized using the Rosetta backrub application. All residues near the interface were taken as pivot residues. For each mutation, the minimal energy structure of an ensemble of 100 structures was taken and its interface score compared to the wild type variant. Structure visualization was done using PyMol (Schrödinger Software).

**Cell experiments:** A431 skin epidermoid carcinoma cells (ATCC number: CRL-1555) were cultured in RPMI 1640 medium + Glutamine (Fisher Scientific) supplemented with 10% (v/v) Fetal Bovine Serum (Gibco) and 1% (v/v) Penicillin / Streptomycin (Gibco) in Falcon polystyrene cell culture flasks (BD Biosciences). Cells were passaged twice per week by washing 3 times with sterile PBS, incubating with Trypsin / EDTA (Gibco) for 5 minutes at 37 °C, briefly slapping the flask to dislodge cells and neutralizing the trypsin by adding fresh medium to the original volume. Each passage, cells were diluted 6-fold. Experiments were done only with cells at passage numbers between 10 and 20. Cells were trypsinsized as described above and washed with 1 mL PBS pH 7.4, 1 mg mL⁻¹ BSA (PBSF). Cyclic peptides Md1 and Q1V/D3N/S5G/K10R (Peptides 13 and 15, Supplementary table S2) were dissolved at 1 µM concentration in 10% DMSO in PBSF. UV-VIS absorbance was measured on a NanoDrop ND1000 spectrophotometer and the fluorescein molar extinction coefficient of 75,000 M⁻¹cm⁻¹ at 496 nm (58) was used for determining the peptide concentrations. 2.5×10⁵ Cells were incubated for 30 minutes at room temperature in 1 mL PBSF with 10 nM cetuximab and 0, 50, 100 or 500 nM FITC-labeled cyclic peptides, which had been pre-incubated for 1 hour at room temperature before the cells were added. Control samples contained 500 nM peptide, but no cetuximab. After incubation the cells were centrifuged at 500 ×g for 5 minutes, re-suspended in PBSF and analyzed by flow cytometry using a FACS Aria III. FITC was excited by the 488 nm laser and detected through a 530/30 BP filter. Side scatter (SSC) was excited by the 488 nm laser and detected through a 488/10 BP filter. The cells were gated in an FSC/SSC plot. Doublet discrimination was done using both the Forward and Side scatter signal width vs height plots. Measurements were performed in duplicate.

**Acknowledgements:**

The work was supported by an ERC starting grant (ERC-2011-StG 280255), an ERC Proof-of-Concept Grant (ERC-2013-PoC 632274) and NanoNext, a micro and nanotechnology consortium of the Government of the Netherlands and 130 partners.
Conflict of interest:
The authors declare that they have no conflicts of interest with the contents of this article.

Author contributions:
MvR conceived and coordinated the study, carried out the experiments shown in figure 4, 5b, 6g, 7 and wrote the manuscript. BMGJ carried out the experiments shown in figure 3c. NMH carried out the experiments in figure 3a and b, AJvdL wrote the python and matlab scripts for and performed the bioinformatics on the deep sequencing data. DW carried out the experiments in figure 2, PAP and TFAdG performed the Rosetta screen described in figure 6a and b, MM conceived and coordinated the study and revised the manuscript.

References
Yeast polypeptide fusion surface display levels predict thermal stability and soluble secretion efficiency. J. Mol. Biol. 292, 949–956
Running title: Affinity maturation of a cyclic peptide handle for therapeutic antibodies


Table 1. Dissociation constants of mutant peptides

<table>
<thead>
<tr>
<th>Sequence</th>
<th>$K_d$ (nM) $^\dagger$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meditope 1</td>
<td>129.4 ± 4.7</td>
</tr>
<tr>
<td>Q1V</td>
<td>111.5 ± 3.6</td>
</tr>
<tr>
<td>S5G</td>
<td>113.9 ± 3.1</td>
</tr>
<tr>
<td>K10R</td>
<td>81.3 ± 1.9</td>
</tr>
<tr>
<td>Q1V/S5G</td>
<td>69.4 ± 2.2</td>
</tr>
<tr>
<td>Q1V/K10R</td>
<td>63.2 ± 1.2</td>
</tr>
<tr>
<td>S5G/K10R</td>
<td>41.1 ± 1.2</td>
</tr>
<tr>
<td>Q1V/S5G/K10R</td>
<td>30.3 ± 0.9</td>
</tr>
</tbody>
</table>

$^*$ Exact peptide sequences (including labels, linkers and flanking residues) in Supplementary Table S2, numbers 5-12.

$^\dagger$ $K_d$ values were calculated with fluorescence polarization measurements (Figure 3d). Values are mean ± SD from the fit of Hill curves through the data presented in figure 3c.
Figure 1: Improving cyclic peptide binding to cetuximab using yeast display. A) Structure of the Fab fragment of cetuximab indicating the position of the meditope in the central cavity. B) Schematic representation of the yeast display system. Simultaneous labeling of the HA tag and cetuximab with different fluorophores allows selection for binding to be normalized for variations in expression.
Figure 2 Display of cyclic meditope peptide on yeast cells. A) Cetuximab titration to 10 nM FITC-labeled peptides (Supplementary table S2, numbers 1-4) in PBS solution, pH 7.4 with 1 mg mL⁻¹ Bovine Serum Albumin (BSA). Fluorescence polarization was used as a readout. Data represent mean ± SD from duplicate experiments. B) Titration of cetuximab to yeast displaying the indicated cyclic peptides. Cells were labeled with the indicated concentrations of cetuximab conjugated to Alexa Fluor 647 as well as mouse anti HA and an AlexaFluor 488 conjugated goat anti-mouse secondary antibody. The flow-cytometric Alexa Fluor 647 mean fluorescence of the Alexa Fluor 488 positive cells was used as a readout.
Figure 3 Selection of improved mutants from single amino acid substitution libraries

A) FACS sorting of pooled libraries. Libraries were pooled prior to sorting and incubated with either 30 or 100 nM Alexa Fluor 647 conjugated cetuximab, mouse anti-HA and an Alexa Fluor 488 conjugated goat anti mouse secondary antibody. Plots show the library and sorting gate used in the first (left), second (middle) and third (right) sorting rounds. B) FACS histograms of cells expressing either the original meditope or mutants Q1V, S5G or K10R. Cells were labeled with 30 nM Alexa Fluor 647 conjugated cetuximab. Populations on the left are non-expressing cells, while populations on the right are expressing cells. D) Fluorescence polarization assay. Cetuximab was titrated to 10 nM of the indicated synthetic peptides in 50 mM sodium phosphate pH 7.0; 100 mM sodium chloride; 1 mg mL⁻¹ BSA. Indicated mutations are with respect to Md1. For exact sequences with linkers and flanking residues see Supplementary table S2 peptides 5-12. Error bars represent mean ± standard deviation (SD) from duplicate measurements.
Figure 4 Deep sequencing analysis of separately sorted single NNK libraries. A) Schematic overview of the deep mutational scanning strategy. B) Heat map of Robust Enrichment Ratios (see Materials & Methods) of all single amino acid substitutions of the original cetuximab meditope. All RERs above 1 are indicated in the respective squares. At each position the original residue is encircled.
Figure 5 Deep sequencing analysis of separately sorted double NNK libraries. A) Heat map of Robust Enrichment Ratios (see Materials & Methods) of double amino acid substitutions of the original cetuximab meditope. All RERs above 1 are indicated in the respective squares. At each position the original residue is encircled. B) Cetuximab titration to 10 nM of FITC-conjugated synthetic peptides in PBS pH 7.4 1 mg mL⁻¹. Fluorescence polarization was used as a readout. Data points represent mean ± standard deviation of triplicate measurements. The indicated mutations are with respect to Md1. For exact peptide sequences see supplementary table S2 peptides 13-16.
Figure 6 In silico modeling of meditope mutations. A) Heat map of interaction free energies of cetuximab with single amino acid substitutions of the meditope generated by Rosetta modeling. Negative values (improved binding) is represented with dark shades of gray, while positive values (weaker binding than the original meditope) is represented with light shades of gray. Positions are shown on the horizontal axis. B) Heat maps of interaction free energies of cetuximab with amino acid substitutions at position 6 of the meditope where position 5 is mutated to tryptophan (left), phenyl alanine (middle) or tyrosine (right), generated by Rosetta modeling. Values are normalized to the respective single amino acid substitution at position 5. C) Model of double mutant S5Y/T6M showing the position of the side chains that fill the empty space in the binding pocket. D) Model of double mutant S5Y/T6M showing the possible hydrogen bond between tyrosine 5 of the meditope and valine 152 of the antibody. E) Model of mutant D3R showing possible hydrogen bonds between arginine 3 of the meditope and glutamine 150 and alanine 176 of the antibody. F) Fluorescence polarization assay. Cetuximab was titrated to 10 nM of the indicated mutant peptides in PBS pH 7.4 1 mg mL⁻¹ BSA. Error bars represent mean ± SD of triplicate measurements. The indicated mutations are with respect to Md1. For exact sequences see Supplementary table S2 peptides 13, 14 and 18-21.
Figure 7 Cetuximab mediated targeting of meditopes to EGFR overexpressing cancer cells. A) A431 cells were incubated with 5 nM unlabeled cetuximab and 50 nM FITC conjugated peptide (Supplementary table S2, numbers 13 and 15) and analyzed by flow-cytometry. Mutations are with respect to Md1. B) Concentration dependent meditope binding. A431 cells were incubated with cetuximab and the indicated concentrations of cyclic peptides as in (A) and analyzed by flow-cytometry. Data represent mean ± SD from duplicate experiments. The control consisted of 500 nM of the respective meditope peptides without cetuximab and represents background binding to the cells due to the hydrophobic nature of the peptide.
Affinity maturation of a cyclic peptide handle for therapeutic antibodies – Supporting information

Affinity maturation of a cyclic peptide handle for therapeutic antibodies using deep mutational scanning – Supporting information

Martijn van Rosmalen, Brian M.G. Janssen, Natalie, M. Hendrikse, Ardjan J. van der Linden, Pascal A. Pieters, Dave Wanders, Tom F.A. de Greef and Maarten Merkx

Contents:

Supplementary table S1: Oligonucleotides used in this study  S2
Supplementary table S2: Peptides used in this study  S4
Supplementary table S3: Next-generation sequencing data analysis  S5
Supplementary figure S1: Separate selection of single and double amino acid substitution libraries  S6
Supplementary Figure S2: Fitness landscape of double amino acid substitutions at positions 1 and 3  S7
Supplementary Figure S3: Fitness landscape of double amino acid substitutions at positions 1 and 5  S8
Supplementary Figure S4: Fitness landscape of double amino acid substitutions at positions 1 and 6  S9
Supplementary Figure S5: Fitness landscape of double amino acid substitutions at positions 3 and 5  S10
Supplementary Figure S6: Fitness landscape of double amino acid substitutions at positions 5 and 6  S11
Supplementary Figure S7: Relation between the number of counts and the coefficient of variation  S12

S1
**Supplementary table S1: Oligonucleotides used in this study**

<table>
<thead>
<tr>
<th>#</th>
<th>Name:</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pCT-Stop-F</td>
<td>ggttcgctagctaatgagcggtggcggc</td>
</tr>
<tr>
<td>2</td>
<td>pCT-Stop-R</td>
<td>gcccacccacactatctctcgaccc</td>
</tr>
<tr>
<td>3</td>
<td>Constant-F</td>
<td>ctagtcttggtgctagctaatgagcggtggcggc</td>
</tr>
<tr>
<td>4</td>
<td>Constant-R</td>
<td>tatacgatcetccatcataagctctctctctctcaaatgcctgctaatgcctcaccaca</td>
</tr>
<tr>
<td>5</td>
<td>Md1-F</td>
<td>ctgtctaatcttatagctggaattttgtcggatcgaacaaagc</td>
</tr>
<tr>
<td>6</td>
<td>Md1-R</td>
<td>ccaacacactacactatagcaggacagcactccactacccacactacccacactacccacactaccc</td>
</tr>
<tr>
<td>7</td>
<td>Md3-F</td>
<td>ggttcagcttttaaaatgtctctctctctgctaatgagcggtggcggc</td>
</tr>
<tr>
<td>8</td>
<td>Md3-R</td>
<td>ccaacacactacactacactacactacactacactacactacactacactacactacactacactaca</td>
</tr>
<tr>
<td>9</td>
<td>Md2-F</td>
<td>gttgcagacagagctaatgcctttggtgctaatgagcggtggcggc</td>
</tr>
<tr>
<td>10</td>
<td>Md2-R</td>
<td>ccaacacactacactacactacactacactacactacactacactacactacactacactacactaca</td>
</tr>
<tr>
<td>11</td>
<td>Md4-F</td>
<td>ctgtctaatcttatagctggaattttgtcggatcgaacaaagc</td>
</tr>
<tr>
<td>12</td>
<td>Md4-R</td>
<td>ccaacacactacactatagcaggacagcactccactacccacactacccacactacccacactaccc</td>
</tr>
<tr>
<td>13</td>
<td>NNK1</td>
<td>ggttcgctagctgtnnktttgactatctacgagagaattgttggttgctaatgcctcaccaca</td>
</tr>
<tr>
<td>14</td>
<td>NNK2</td>
<td>ggttcgctagctgtnnktttgactatctacgagagaattgttggttgctaatgcctcaccaca</td>
</tr>
<tr>
<td>15</td>
<td>NNK3</td>
<td>ggttcgctagctgtnnktttgactatctacgagagaattgttggttgctaatgcctcaccaca</td>
</tr>
<tr>
<td>16</td>
<td>NNK4</td>
<td>ggttcgctagctgtnnktttgactatctacgagagaattgttggttgctaatgcctcaccaca</td>
</tr>
<tr>
<td>17</td>
<td>NNK5</td>
<td>ggttcgctagctgtnnktttgactatctacgagagaattgttggttgctaatgcctcaccaca</td>
</tr>
<tr>
<td>18</td>
<td>NNK6</td>
<td>ggttcgctagctgtnnktttgactatctacgagagaattgttggttgctaatgcctcaccaca</td>
</tr>
<tr>
<td>19</td>
<td>NNK7</td>
<td>ggttcgctagctgtnnktttgactatctacgagagaattgttggttgctaatgcctcaccaca</td>
</tr>
<tr>
<td>20</td>
<td>NNK8</td>
<td>ggttcgctagctgtnnktttgactatctacgagagaattgttggttgctaatgcctcaccaca</td>
</tr>
<tr>
<td>21</td>
<td>NNK9</td>
<td>ggttcgctagctgtnnktttgactatctacgagagaattgttggttgctaatgcctcaccaca</td>
</tr>
<tr>
<td>22</td>
<td>NNK10</td>
<td>ggttcgctagctgtnnktttgactatctacgagagaattgttggttgctaatgcctcaccaca</td>
</tr>
<tr>
<td>23</td>
<td>NNK1+3</td>
<td>ggttcgctagctgtnnktttgactatctacgagagaattgttggttgctaatgcctcaccaca</td>
</tr>
<tr>
<td>24</td>
<td>NNK1+5</td>
<td>ggttcgctagctgtnnktttgactatctacgagagaattgttggttgctaatgcctcaccaca</td>
</tr>
<tr>
<td>25</td>
<td>NNK1+6</td>
<td>ggttcgctagctgtnnktttgactatctacgagagaattgttggttgctaatgcctcaccaca</td>
</tr>
<tr>
<td>26</td>
<td>NNK3+5</td>
<td>ggttcgctagctgtnnktttgactatctacgagagaattgttggttgctaatgcctcaccaca</td>
</tr>
<tr>
<td>27</td>
<td>NNK5+6</td>
<td>ggttcgctagctgtnnktttgactatctacgagagaattgttggttgctaatgcctcaccaca</td>
</tr>
<tr>
<td>28</td>
<td>YD-tag-F</td>
<td>cgaagtaacacacagacgacgacgactctctctctctctctctctctctctctctctctctctctctc</td>
</tr>
<tr>
<td>29</td>
<td>YD-tag-R</td>
<td>ccaacacagacacacagacagagacagacacacagacacacagacacagacacacagacacacagac</td>
</tr>
<tr>
<td>30</td>
<td>NGS-F1</td>
<td>aatgacacgagcaccacaccagacagacagacagacagacagacagacagacagacagacagacagac</td>
</tr>
<tr>
<td>31</td>
<td>NGS-F2</td>
<td>aatgacacgagcaccacaccagacagacagacagacagacagacagacagacagacagacagacagac</td>
</tr>
<tr>
<td>32</td>
<td>NGS-F3</td>
<td>aatgacacgagcaccacaccagacagacagacagacagacagacagacagacagacagacagacagac</td>
</tr>
<tr>
<td>33</td>
<td>NGS-F4</td>
<td>aatgacacgagcaccacaccagacagacagacagacagacagacagacagacagacagacagacagac</td>
</tr>
<tr>
<td>34</td>
<td>NGS-F5</td>
<td>aatgacacgagcaccacaccagacagacagacagacagacagacagacagacagacagacagacagac</td>
</tr>
<tr>
<td>35</td>
<td>NGS-F6</td>
<td>aatgacacgagcaccacaccagacagacagacagacagacagacagacagacagacagacagacagac</td>
</tr>
<tr>
<td>36</td>
<td>NGS-R1</td>
<td>caacagacacagacagacagacagacagacagacagacagacagacagacagacagacagacagacagac</td>
</tr>
<tr>
<td>37</td>
<td>NGS-R2</td>
<td>caacagacacagacagacagacagacagacagacagacagacagacagacagacagacagacagacagac</td>
</tr>
<tr>
<td>38</td>
<td>NGS-R3</td>
<td>caacagacacagacagacagacagacagacagacagacagacagacagacagacagacagacagacagac</td>
</tr>
</tbody>
</table>
Supplementary table S1 (Continued): Oligonucleotides used in this study

<table>
<thead>
<tr>
<th>#</th>
<th>Name:</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>39</td>
<td>NGS-R2</td>
<td>caacagacacgccatacagatgactggaggtgcttcagactgtgctctctcagatctnnntaaggcagagtcacacagagacagctgacatg</td>
</tr>
<tr>
<td>40</td>
<td>NGS-R3</td>
<td>caacagacacgccatacagatgactggaggtgcttcagactgtgctctctcagatctnnntaaggcagagtcacacagagacagctgacatg</td>
</tr>
<tr>
<td>41</td>
<td>NGS-R4</td>
<td>caacagacacgccatacagatgactggaggtgcttcagactgtgctctctcagatctnnntaaggcagagtcacacagagacagctgacatg</td>
</tr>
<tr>
<td>42</td>
<td>NGS-R5</td>
<td>caacagacacgccatacagatgactggaggtgcttcagactgtgctctctcagatctnnntaaggcagagtcacacagagacagctgacatg</td>
</tr>
<tr>
<td>43</td>
<td>NGS-R6</td>
<td>caacagacacgccatacagatgactggaggtgcttcagactgtgctctctcagatctnnntaaggcagagtcacacagagacagctgacatg</td>
</tr>
<tr>
<td>44</td>
<td>Amp-F</td>
<td>aatgatacgcacacac</td>
</tr>
<tr>
<td>45</td>
<td>Amp-R</td>
<td>caacagacacgccac</td>
</tr>
<tr>
<td>46</td>
<td>Seq-pCON-F</td>
<td>cgacagtgaaggtgatcatacctgacgttcecaagactacgctctgagctagtcgagctagt</td>
</tr>
</tbody>
</table>
Affinity maturation of a cyclic peptide handle for therapeutic antibodies – Supporting information

Supplementary table S2: Peptides used in this study

<table>
<thead>
<tr>
<th>#</th>
<th>Meditope / Mutations</th>
<th>Sequence</th>
<th>Mw&lt;sub&gt;calc&lt;/sub&gt; (Da)</th>
<th>Mw&lt;sub&gt;obs&lt;/sub&gt; (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Md1</td>
<td>FAM-GGSGGSPGLAGGGSGGSCQFDLSTRLKCG</td>
<td>3138.0</td>
<td>3138.9</td>
</tr>
<tr>
<td>2</td>
<td>Md3</td>
<td>FITC-O1PEN-GPLGLAGCQVYNLSSRALKCG</td>
<td>2495.8</td>
<td>2495.9</td>
</tr>
<tr>
<td>3</td>
<td>Md2</td>
<td>FITC-O1PEN-GPLGLAGCVWQRWQKSYVCG</td>
<td>2696.1</td>
<td>2696.1</td>
</tr>
<tr>
<td>4</td>
<td>Md4</td>
<td>FITC-O1PEN-GPLGLAGCMWDREPHWYKCG</td>
<td>2789.2</td>
<td>2789.4</td>
</tr>
<tr>
<td>5</td>
<td>Md1</td>
<td>FITC-O1PEN-GPLGLAGCQFDLSTRLKCG</td>
<td>2580.5</td>
<td>2580.4</td>
</tr>
<tr>
<td>6</td>
<td>Q1V</td>
<td>FITC-O1PEN-GPLGLAGCVFDLSTRLKCG</td>
<td>2549.2</td>
<td>2549.1</td>
</tr>
<tr>
<td>7</td>
<td>S5G</td>
<td>FITC-O1PEN-GPLGLAGCQFDLGTRRLKCG</td>
<td>2548.2</td>
<td>2548.2</td>
</tr>
<tr>
<td>8</td>
<td>K10R</td>
<td>FITC-O1PEN-GPLGLAGCQFDLSTRLRKC</td>
<td>2606.2</td>
<td>2606.4</td>
</tr>
<tr>
<td>9</td>
<td>Q1V/S5G</td>
<td>FITC-O1PEN-GPLGLAGCVFDLGTRRLKCG</td>
<td>2519.2</td>
<td>2519.6</td>
</tr>
<tr>
<td>10</td>
<td>Q1V/K10R</td>
<td>FITC-O1PEN-GPLGLAGCVFDLSTRLRKC</td>
<td>2577.2</td>
<td>2577.5</td>
</tr>
<tr>
<td>11</td>
<td>S5G/K10R</td>
<td>FITC-O1PEN-GPLGLAGCQFDLGTRRLKCG</td>
<td>2576.2</td>
<td>2576.2</td>
</tr>
<tr>
<td>12</td>
<td>Q1V/S5G/K10R</td>
<td>FITC-O1PEN-GPLGLAGCVFDLSTRLRKC</td>
<td>2549.5</td>
<td>2549.7</td>
</tr>
<tr>
<td>13</td>
<td>Md1</td>
<td>FITC-O1PEN-GCQFDLSTRLKCG</td>
<td>2069.9</td>
<td>2070.2</td>
</tr>
<tr>
<td>14</td>
<td>Q1V/S5G/K10R</td>
<td>FITC-O1PEN-GCVFDLGTRRLRKC</td>
<td>2038.9</td>
<td>2039.0</td>
</tr>
<tr>
<td>15</td>
<td>Q1V/D3N/S5G/K10R</td>
<td>FITC-O1PEN-GCVFNLGTRRLRKC</td>
<td>2037.9</td>
<td>2039.0</td>
</tr>
<tr>
<td>16</td>
<td>Q1V/S5G/T6l/K10R</td>
<td>FITC-O1PEN-GCVFDLGIIRLRCG</td>
<td>2050.9</td>
<td>2052.0</td>
</tr>
<tr>
<td>17</td>
<td>Q1V/D3N/S5G/T6l/K10R</td>
<td>FITC-O1PEN-GCVFNLGIIRLRCG</td>
<td>2049.9</td>
<td>2051.0</td>
</tr>
<tr>
<td>18</td>
<td>Q1V/S5Y/K10R</td>
<td>FITC-O1PEN-GCVFDLTYTRRLRGC</td>
<td>2144.9</td>
<td>2145.6</td>
</tr>
<tr>
<td>19</td>
<td>Q1V/S5G/T6M/K10R</td>
<td>FITC-O1PEN-GCVFDLMRRLRGC</td>
<td>2068.9</td>
<td>2069.0</td>
</tr>
<tr>
<td>20</td>
<td>Q1V/S5Y/T6M/K10R</td>
<td>FITC-O1PEN-GCVFDLYMRRLRGC</td>
<td>2174.9</td>
<td>2177.0</td>
</tr>
<tr>
<td>21</td>
<td>Q1V/D3R/S5G/K10R</td>
<td>FITC-O1PEN-GCVFRLGTRRLRGC</td>
<td>2079.9</td>
<td>2080.8</td>
</tr>
<tr>
<td>22</td>
<td>Q1V/D3R/S5Y/T6M/K10R</td>
<td>FITC-O1PEN-GCVFRLYMRRRLRGC</td>
<td>2216.0</td>
<td>2217.8</td>
</tr>
</tbody>
</table>
### Supplementary table S3: Next-generation sequencing data analysis

<table>
<thead>
<tr>
<th>Library</th>
<th>Selection</th>
<th>Barcode pair</th>
<th>Total</th>
<th>In frame</th>
<th>Phred $\geq 40$</th>
<th>Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>NGS-F1 &amp; NGS-R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NNK 1</td>
<td>Unsorted</td>
<td>NGS-F3 &amp; NGS-R5</td>
<td>320914</td>
<td>214740</td>
<td>162338</td>
<td>149729</td>
</tr>
<tr>
<td></td>
<td>2×Sorted</td>
<td>NGS-F3 &amp; NGS-R5</td>
<td>324141</td>
<td>261482</td>
<td>177569</td>
<td>173774</td>
</tr>
<tr>
<td>NNK 2</td>
<td>Unsorted</td>
<td>NGS-F1 &amp; NGS-R2</td>
<td>426825</td>
<td>242874</td>
<td>169205</td>
<td>156610</td>
</tr>
<tr>
<td></td>
<td>2×Sorted</td>
<td>NGS-F3 &amp; NGS-R6</td>
<td>573816</td>
<td>436730</td>
<td>313198</td>
<td>291080</td>
</tr>
<tr>
<td>NNK 3</td>
<td>Unsorted</td>
<td>NGS-F1 &amp; NGS-R3</td>
<td>557810</td>
<td>439714</td>
<td>332404</td>
<td>316464</td>
</tr>
<tr>
<td></td>
<td>2×Sorted</td>
<td>NGS-F2 &amp; NGS-R1</td>
<td>579336</td>
<td>440080</td>
<td>330250</td>
<td>326072</td>
</tr>
<tr>
<td>NNK 4</td>
<td>Unsorted</td>
<td>NGS-F1 &amp; NGS-R4</td>
<td>598847</td>
<td>389158</td>
<td>300678</td>
<td>288976</td>
</tr>
<tr>
<td></td>
<td>2×Sorted</td>
<td>NGS-F2 &amp; NGS-R2</td>
<td>527756</td>
<td>395920</td>
<td>302213</td>
<td>295208</td>
</tr>
<tr>
<td>NNK 5</td>
<td>Unsorted</td>
<td>NGS-F1 &amp; NGS-R5</td>
<td>575770</td>
<td>299262</td>
<td>228258</td>
<td>216273</td>
</tr>
<tr>
<td></td>
<td>2×Sorted</td>
<td>NGS-F2 &amp; NGS-R3</td>
<td>553367</td>
<td>422883</td>
<td>309737</td>
<td>306061</td>
</tr>
<tr>
<td>NNK 6</td>
<td>Unsorted</td>
<td>NGS-F1 &amp; NGS-R6</td>
<td>586155</td>
<td>280257</td>
<td>212621</td>
<td>201358</td>
</tr>
<tr>
<td></td>
<td>2×Sorted</td>
<td>NGS-F2 &amp; NGS-R4</td>
<td>380663</td>
<td>292548</td>
<td>219974</td>
<td>217869</td>
</tr>
<tr>
<td>NNK 7</td>
<td>Unsorted</td>
<td>NGS-F4 &amp; NGS-R1</td>
<td>572134</td>
<td>285791</td>
<td>207503</td>
<td>191531</td>
</tr>
<tr>
<td></td>
<td>2×Sorted</td>
<td>NGS-F2 &amp; NGS-R5</td>
<td>657865</td>
<td>507973</td>
<td>384340</td>
<td>305451</td>
</tr>
<tr>
<td>NNK 8</td>
<td>Unsorted</td>
<td>NGS-F4 &amp; NGS-R2</td>
<td>556224</td>
<td>395678</td>
<td>292007</td>
<td>270406</td>
</tr>
<tr>
<td></td>
<td>2×Sorted</td>
<td>NGS-F2 &amp; NGS-R6</td>
<td>594070</td>
<td>497696</td>
<td>376900</td>
<td>366672</td>
</tr>
<tr>
<td>NNK 9</td>
<td>Unsorted</td>
<td>NGS-F4 &amp; NGS-R3</td>
<td>578099</td>
<td>215278</td>
<td>153364</td>
<td>145820</td>
</tr>
<tr>
<td></td>
<td>2×Sorted</td>
<td>NGS-F5 &amp; NGS-R1</td>
<td>524878</td>
<td>447741</td>
<td>348441</td>
<td>340554</td>
</tr>
<tr>
<td>NNK 10</td>
<td>Unsorted</td>
<td>NGS-F4 &amp; NGS-R4</td>
<td>569154</td>
<td>344239</td>
<td>245380</td>
<td>221456</td>
</tr>
<tr>
<td></td>
<td>2×Sorted</td>
<td>NGS-F5 &amp; NGS-R2</td>
<td>599217</td>
<td>514084</td>
<td>386256</td>
<td>380678</td>
</tr>
<tr>
<td>NNK 1+3</td>
<td>Unsorted</td>
<td>NGS-F4 &amp; NGS-R5</td>
<td>480808</td>
<td>249144</td>
<td>171231</td>
<td>154901</td>
</tr>
<tr>
<td></td>
<td>2×Sorted</td>
<td>NGS-F5 &amp; NGS-R3</td>
<td>561375</td>
<td>425472</td>
<td>322978</td>
<td>316870</td>
</tr>
<tr>
<td>NNK 1+5</td>
<td>Unsorted</td>
<td>NGS-F4 &amp; NGS-R6</td>
<td>517949</td>
<td>379512</td>
<td>255202</td>
<td>230070</td>
</tr>
<tr>
<td></td>
<td>2×Sorted</td>
<td>NGS-F5 &amp; NGS-R4</td>
<td>551138</td>
<td>408722</td>
<td>306911</td>
<td>299525</td>
</tr>
<tr>
<td>NNK 1+6</td>
<td>Unsorted</td>
<td>NGS-F3 &amp; NGS-R1</td>
<td>580134</td>
<td>382492</td>
<td>257135</td>
<td>233062</td>
</tr>
<tr>
<td></td>
<td>2×Sorted</td>
<td>NGS-F5 &amp; NGS-R5</td>
<td>439989</td>
<td>369595</td>
<td>285171</td>
<td>279680</td>
</tr>
<tr>
<td>NNK 3+5</td>
<td>Unsorted</td>
<td>NGS-F3 &amp; NGS-R2</td>
<td>469389</td>
<td>303259</td>
<td>216599</td>
<td>199507</td>
</tr>
<tr>
<td></td>
<td>2×Sorted</td>
<td>NGS-F5 &amp; NGS-R6</td>
<td>591974</td>
<td>437146</td>
<td>332563</td>
<td>326602</td>
</tr>
<tr>
<td>NNK 5+6</td>
<td>Unsorted</td>
<td>NGS-F3 &amp; NGS-R4</td>
<td>515046</td>
<td>330584</td>
<td>230697</td>
<td>209744</td>
</tr>
<tr>
<td></td>
<td>2×Sorted</td>
<td>NGS-F6 &amp; NGS R2</td>
<td>492833</td>
<td>349569</td>
<td>259572</td>
<td>257428</td>
</tr>
</tbody>
</table>
Supplementary Figure S1 Separate selection of single and double amino acid substitution libraries: FACS density plots of libraries during the first and second sorting round. Yeast were separately labeled with 100 nM Alexa Fluor 647 conjugated cetuximab as well as Mouse anti-HA and Alexa Fluor 488 Goat anti-Mouse secondary antibody. Alexa Fluor 488 fluorescence intensity is depicted on the horizontal axis of each plot and Alexa Fluor 647 intensity on the vertical axis. The indicated gates were used for sorting.
Affinity maturation of a cyclic peptide handle for therapeutic antibodies – Supporting information

Supplementary Figure S2: Fitness landscape of double amino acid substitutions at positions 1 and 3. Heat-map of pairwise Robust Enrichment Ratios of combinations of amino acids from the double NNK 1+3 library. Empty squares are combinations that were counted less than 10 times in the unselected library. The ‘wildtype’ combination is underlined.
**Supplementary Figure S3: Fitness landscape of double amino acid substitutions at positions 1 and 5.**

Heat-map of pairwise Robust Enrichment Ratios of combinations of amino acids from the double NNK 1+5 library. Empty squares are combinations that were counted less than 10 times in the unselected library. The ‘wildtype’ combination is underlined.
Supplementary Figure S4: Fitness landscape of double amino acid substitutions at positions 1 and 6.

Heat-map of pairwise Robust Enrichment Ratios of combinations of amino acids from the double NNK 1+6 library. Empty squares are combinations that were counted less than 10 times in the unselected library. The ‘wildtype’ combination is underlined.
### Supplementary Figure S5: Fitness landscape of double amino acid substitutions at positions 3 and 5.

Heat-map of pairwise Robust Enrichment Ratios of combinations of amino acids from the double NNK 3+5 library. Empty squares are combinations that were counted less than 10 times in the unselected library. The ‘wildtype’ combination is underlined.
Supplementary Figure S6: Fitness landscape of double amino acid substitutions at positions 5 and 6. Heat-map of pairwise Robust Enrichment Ratios of combinations of amino acids from the double NNK 5+6 library. Empty squares are combinations that were counted less than 10 times in the unselected library. The ‘wildtype’ combination is underlined.
Supplementary Figure S7: Relation between the number of counts and the coefficient of variation. Coefficients of variation were calculated from Bootstrap samples with different number of counts of a particular amino acid in a total number of 10,000 reads. Error bars represent mean ± SD of three replicate analyses, where each single analysis consisted of 100 Bootstraps. Bootstraps with 1, 2 and 3 counts were repeated with a total number of 100,000 reads to show that the CV is dependent on the number of counts instead of the frequency. The curve with the indicated formula was fit through the data. The inset presents the same data, but on a log log scale.
Affinity maturation of a cyclic peptide handle for therapeutic antibodies using deep mutational scanning
Martijn van Rosmalen, Brian M.G. Janssen, Nathalie M. Hendrikse, Ardjan J. van der Linden, Pascal A. Pieters, Dave Wanders, Tom F.A. de Greef and Maarten Merkx

J. Biol. Chem. published online December 14, 2016

Access the most updated version of this article at doi: 10.1074/jbc.M116.764225

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2016/12/14/M116.764225.DC1.html

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/early/2016/12/14/jbc.M116.764225.full.html#ref-list-1