Chapter 3

Modular Nucleic Acid Assembled p/MHC Microarrays for Multiplexed Sorting of Antigen-Specific T Cells

2.1 Introduction

T cells constitute an important part of the acquired immune system. They recognize a diversity of antigens through the highly variable, hetero-dimeric T cell receptor protein (TCR), with approximately $10^7$ different antigen specificities (1). The initiation of the T cell immune response is triggered by the engagement of the TCR with processed antigenic peptides (e.g. from a bacterial pathogen) that are bound to Major Histocompatibility Complex (p/MHC) molecules presented on the surface of antigen presenting cells (APCs), leading to downstream T cell proliferation and maturation into effector populations. After pathogen clearance, a subset of the activated T cells transition into memory cells, providing the immune system with the capacity for rapid response
towards previously encountered pathogens. As a consequence, an individual’s collection of T cells and their antigen specificities, collectively called the T cell repertoire, is an evolving, extensive repository of cellular immune responses against self and foreign antigens. It is of fundamental and therapeutic importance to detect and survey these T cell populations.

The development of soluble p/MHC tetramers for labeling antigen-specific T cells has enabled the direct phenotypic analysis of antigen-specific T cell populations with flow cytometry (2). Conventionally, p/MHC tetramers are prepared by mixing enzymatically biotinylated p/MHC molecules with preparations of streptavidin (SA)-fluorophore conjugates. While p/MHC monomers have low affinities (2, 3), their tetramer counterparts exhibit much higher avidity, permitting T cell detection via flow cytometry to become a standard assay. However, because p/MHC tetramer-stained T cell populations are encoded optically (i.e. one unique fluorophore required per p/MHC specificity), the number of antigen-specificities that can be interrogated simultaneously within a population is limited by spectral overlap. In addition, serial flow cytometry detection of distinct antigen-specific T cells is generally restricted by sample size. Efforts to increase the degree of multiplexing generally revolve around polychromatic flow cytometry utilizing quantum dots (4, 5). However, cost, sample preparation time, and color compensation complexity also increase correspondingly.

As an alternative to flow cytometry, several groups have reported microarray-based T cell detection schemes, in which collections of p/MHC complexes are printed on a supporting substrate (6–9). A population of cells is applied directly to the p/MHC array where target antigen-specific T cells bind to regions spotted with the cognate p/MHC and
are detected optically. Analogous to DNA and protein microarrays, the readout of such assays is dependent on location rather than distinct fluorescent signals, thus potentially increasing the degree of multiplexing.

A factor that needs to be addressed before p/MHC arrays are used for broader studies and applications concerns the reproducibility and robustness of p/MHC arrays produced by spotting onto treated and/or derivatized surfaces. Ideally p/MHC complexes should be immobilized such that their functional conformations are preserved. Analogous protein arrays produced via antibody adsorption to unmodified and derivatized surfaces can suffer from surface induced effects including protein denaturation and protein adsorption in inactive orientations (10–12). To circumvent such problems, customized surfaces and relatively mild chemistries for protein immobilization have been developed (13–18). However, often the surface that meets the demands of the application requires a high level of technical expertise and/or is limited in accessibility (19).

We report here on the method of Nucleic Acid Cell Sorting (NACS), which is based upon the design and application of nucleic acid assembled p/MHC tetramer arrays for multiplexed sorting of antigen-specific T cells. For NACS, p/MHC tetramers of distinct specificities are conjugated to unique sequences of ssDNA in a site-specific fashion. A collection of ssDNA-tagged p/MHC complexes is then self-assembled by DNA hybridization onto a glass slide printed with the complementary DNA sequences. Fully assembled p/MHC tetramer arrays are used to sort mixed populations of antigen-specific T cells (Figure 3.1). This strategy of using DNA pendants as molecular linkages (20-25) is simple and highly modular. Most importantly, T cell array binding is
optimized by utilizing cysteine-engineered streptavidin (SAC) for ssDNA-p/MHC tetramer production, resulting in NACS p/MHC arrays that outperform conventional spotted arrays assessed by performance criteria such as reproducibility and homogeneity. The versatility of using DNA tags is also exploited to enable selective detachment of T cells with restriction endonucleases. Demonstrative experiments regarding NACS sensitivity, multiplexing and limit of detection are performed with cell lines and finally with T cells isolated from cancer patients.

Figure 3.1 Self-assembled ssDNA-p/MHC tetramer arrays for multiplexed sorting of antigen-specific cells ssDNA-tagged p/MHC tetramers are produced by coupling ssDNA site-specifically to SAC prior to exposure to molar excess of biotinylated p/MHC monomers. p/MHC tetramer arrays are formed by pooling ssDNA-p/MHC tetramers of select specificity and hybridization to a complementary printed ssDNA microarray. T cells expressing the cognate TCR are detected by binding to the surface confined tetramer.
3.2 Experimental Methods

3.2.1 Microarray Fabrication

All DNA strands were purchased from IDT with HPLC purification. DNA microarrays were printed by the microarray facility at the Institute for Systems Biology (ISB—Seattle, WA) on amine-coated glass slides (GAPS II, Corning) in identical triplicate 12x12 arrays containing alternative rows of A, B and C spots, or $A_{\text{EcoRI}}$ and $B_{\text{BamHI}}$ with a SMPXB15 pin (Arrayit). Sequences for all strands can be found at Table 3.1.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5' - AAA AAA AAA AAA AAT CCT GGA GCT AAG TCC GTA AAA AAA AAA AAT CCT</td>
</tr>
<tr>
<td></td>
<td>GGA GCT AAG TCC GTA AAA AAA AAA A</td>
</tr>
<tr>
<td>A'</td>
<td>5' - NH$_2$- AAA AAA AAA ATA CGG ACT TAG CTC CAG GAT</td>
</tr>
<tr>
<td>B</td>
<td>5' - AAA AAA AAA AAA AGC CTC ATT GAA TCA TGC CTA AAA AAA AAA AGC CTC</td>
</tr>
<tr>
<td></td>
<td>ATT GAA TCA TGC CTA AAA AAA AAA A</td>
</tr>
<tr>
<td>B'</td>
<td>5' - NH$_2$- AAA AAA AAA ATA GGC ATG ATT CAA TGA GCC</td>
</tr>
<tr>
<td>C</td>
<td>5' - AAA AAA AAA AAA AGC ACT CGT CTA TCG CTA AAA AAA AAA AGC ACT</td>
</tr>
<tr>
<td></td>
<td>CGT CTA CTA TCG CTA AAA AAA AAA A</td>
</tr>
<tr>
<td>C'</td>
<td>5' - NH$_2$- AAA AAA AAA ATA GCG ATA GTA GAC GAG TGC</td>
</tr>
<tr>
<td>$A_{\text{EcoRI}}$</td>
<td>5' - AAA AAA AAA AAA GAG CTA AGT CCG TAG AAT TCA AAA AAA AAA GAG CTA</td>
</tr>
<tr>
<td></td>
<td>AGT CCG TAG AAT TCA AAA AAA AAA A</td>
</tr>
<tr>
<td>$A_{\text{EcoRI}'}$</td>
<td>5' - NH$_2$- AAA AAA AAA AGA ATT CTA CGG ACT TAG CTC CAG GAT</td>
</tr>
<tr>
<td>$B_{\text{BamHI}}$</td>
<td>5' - AAA AAA AAA AAA TTG AAT CAT GCC TAG GAT CCA AAA AAA AAA TTG</td>
</tr>
<tr>
<td></td>
<td>AAT CAT GCC TAG GAT CCA AAA AAA AAA A</td>
</tr>
<tr>
<td>$B_{\text{BamHI}'}$</td>
<td>5' - NH$_2$- AAA AAA AAA AGG ATC CTA GCC ATG ATT CAA TGA GCC</td>
</tr>
</tbody>
</table>
* All sequences to be conjugated to SAC (A’, B’, C’, A_EcoRI’, and B_BamHI’) were designed with a polyA linker followed by a 20mer hybridization region. The 5’ amine is required for the attachment of the hetero-bifunctional maleimide derivative MHPH. Sequences printed on glass substrates (A, B, C, A_EcoRI, and B_BamHI) were designed with two hybridization regions separated by polyAs. This was designed to facilitate electrostatic adsorption to amine glass substrates.

3.2.2 Production of ssDNA-SAC conjugates

The pET-3a plasmid containing the SAC gene was a kind gift from Takeshi Sano (Harvard Medical School). The expression of SAC was performed according to previously published protocols (26). Briefly, transformed BL21(DE3)-pLysE cells were grown at 37°C with shaking in LB medium and selection antibiotics ampicilin and chlorophenicol. The cells were induced at OD600 = 0.6 with IPTG and kept spinning for another 4 hours. The culture was then centrifuged at 1600g for 10 min and lysed with lysis buffer (2 mM EDTA, 30 mM Tris-HCl, 0.1% Triton X-100, pH 8.0). The insoluble inclusion bodies were then separated from the lysate by centrifugation at 39,000g for 15 min and dissolved in 6 M guanidine-HCl, pH 1.5 to the original culture volume. The SAC lysate was then refolded by dialysis in 0.2 M Sodium acetate, 10 mM β-mercaptoethanol (β-ME) pH 6.0 overnight before dialyzed against 50mM Sodium bicarbonate, 500 mM NaCl, 10 mM β-ME pH 11 in preparation for column purification. Refolded volumes of SAC were mixed 1:1 with binding buffer (50 mM Sodium bicarbonate, 500 mM NaCl, 10mM β-ME, pH 11). A gravity column packed with 1.5 ml of iminobiotin agarose resin (Pierce) was washed with 10 ml of binding buffer. The refolded mixture was then applied to the column and the eluted fractions were collected and reapplied to the column again, to maximize SAC recovery. After washing the
column with 20 ml binding buffer, SAC was eluted with pH 4 elution buffer (50 mM Sodium acetate, 10mM β-ME). Fractions containing SAC, as monitored by OD280, were collected, buffer exchanged to PBS containing 10 mM β-ME, and concentrated to 1 mg/ml final concentration using 10K mwco filters (Millipore). Immediately prior to conjugation, stock SAC was buffer exchanged to PBS containing 5mM Tris(2-Carboxyethyl) phosphine Hydrochloride (TCEP) using zeba desalting columns (Pierce). MHPH (3-N-Maleimido-6-hydraziniumpyridine hydrochloride, Solulink) in DMF was added to SAC at a molar excess of 300:1. In parallel, SFB in DMF (succinimidy l 4-formylbenzoate, Solulink) was added in a 40:1 molar excess to 5’aminated oligos. The mixtures were reacted at room temperature (RT) for 3–4 hours before buffer exchanged to citrate (50mM sodium citrate, 150 mM NaCl, pH 6.0) using zeba columns. The SFB-labeled oligos were combined in a 20:1 molar excess with the derivatized SAC and incubated overnight at RT. Unreacted oligos were removed using a Pharmacia Superdex 200 gel filtration column at 0.5 ml/min isocratic flow of PBS. Fractions containing the SAC-oligo conjugates were concentrated using 10K mwco concentration filters (Millipore).

### 3.2.3 Preparation of T cells

The cDNA from the alpha and beta chains of a TCR specific for tyrosinase368-376 was a kind gift from Michael I. Nishimura (Medical University of South Carolina, Charleston, SC). The alpha and beta chains were cloned into a lentiviral vector where both transgenes were linked by a 2A self-cleaving sequence as described (27). Concentrated supernatant from this lentiviral vector was used to infect Jurkat cells to generate Jurkatα-Tyro cells. The MSGV1-F5AfT2AB retroviral vector expressing the F5
MART-1 TCR was a kind gift from Steven A. Rosenberg and Richard Morgan (Surgery Branch, National Cancer Institute Bethesda, MD). The MSGV1-F5AfT2AB retroviral supernatant was used to infect Jurkat cells to generate the Jurkat$^{\alpha\text{-MART-1}}$ cell line. The Jurkat$^{\alpha\text{-NY-ESO-1}}$ cell line was a generous gift from Robert Prins (UCLA). To generate primary human T cell cultures expressing the F5 MART-1 TCR, PBMCs obtained from leukapheresis were activated for 48 hours with 50 ng/ml of OKT3 (muromonab anti-human CD3 antibody, Ortho-Biotech, Bridgewater, NJ) and 300 U/ml of IL-2 (adesleukin, Novartis, Emeryville, CA). MSGV1-F5AfT2AB retrovirus supernatant was applied to retronectin-coated wells (Takara Bio Inc., Japan). Then activated PBMC in RPMI plus 5% human AB serum supplemented by 300 IU of IL-2 were added to these wells and incubated at 37°C overnight at 5% CO2. On the following day, PBMC are transferred to a second set of pre-coated retronectin retroviral vector tissue culture plates and incubated at 37°C overnight at 5% CO2. Cells were subsequently washed and re-suspended in culture media described above. Frozen leukapheresis fractions from patients NRA11, NRA 13 (UCLA IRB#03-12-023 ) and F5-1 (UCLA IRB #08-02-020-02A) were thawed and incubated overnight in RPMI supplemented with 10% human AB serum and 1% penicillin, streptomycin, and amphotericin (Omega Scientific). F5-1 cells were used immediately following incubation. NRA11 and NRA13 samples were CD8+ enriched (anti-CD8 microbeads, Miltenyi Biotech) using an AutoMACS machine according to the manufacturer’s instructions. Following separation, the cells were kept at in RPMI-humanAB media containing 30 U IL2/mL.
3.2.4 T Cell Sorting Methods

The HLA-A*0201 restricted MHC class I monomers loaded with tyrosinase\textsubscript{369-377} (YMDGTMSQV), MART-1\textsubscript{26-35} (ELAGIGILTV) and NY-ESO-1\textsubscript{1157-165} (SLLMWITQC) were produced in house according to previous published protocols (28). A2.1-restricted EBV BMLF1\textsubscript{259-267} (GLCTLVAML), CMV pp65\textsubscript{495-503} (NLVPMVATV), murine H-2Kb/-OVA\textsubscript{257-264} (SIINFEKL), and murine H-2Db/-gp100\textsubscript{25-33} (KVPRNQDWL) as well as all fluorescent HLA-A*0201 tetramers were purchased from Beckman Coulter. Lipophilic cell membrane staining dyes DiO, DiD, and DiL were purchased from Invitrogen.

Prior to experiments, microarray slides were blocked to prevent non-specific cell binding with 1 mg/ml PEG-NHS ester (Sunbio) in PBS for 2 hours at RT. Four-fold molar excess of p/MHC monomers were combined with ssDNA-SAC at 37°C for 20 min. ssDNA-p/MHC tetramers were hybridized to DNA arrays for 1 hour at 37°C in 200 μl media and rinsed with 3% FBS in PBS. T cells (10\textsuperscript{6} /100 μl media) were incubated on the array at 37°C for 30 min. The arrays were rinsed with 3% FBS in PBS and cell capture visualized via brightfield (Nikon Eclipse TE2000) and/or confocal microscopy (Nikon E800). Post T cell capture p/MHC tetramer staining was done by incubating the array with 200 μl of media containing fluorescent p/MHC tetramer along with fluorescent cDNA (Cy5-A’ and/or Cy3-B’). The arrays were rinsed with 3% FBS in PBS prior to imaging. For selective T cell release experiments, three identical arrays were used to immobilize cells. Treatment with EcoRI, BamHI, or DNase was in RPMI media for 1–2 hours at 37°C. DNase was purchased from Sigma, all other enzymes from NEbiolabs.
For p/MHC comparative studies, SuperEpoxy and SuperProtein (representing covalent and hydrophobic surfaces respectively) were purchased from Arrayit (Sunnyvale, CA). Amine GAPS II slides (electrostatic) were purchased from Corning. Polycarboxylate hydrogel (hydrophilic) slides were purchased from XanTec (Germany). Fluorescent MART-1 tetramers were printed according to manufacturer’s instructions for each slide. Cell sorting images were quantified with ImageJ (NIH) and fitted to the Hill Function (NACS n=2, R2=0.95, Covalent n=2.1, R2 =0.97) with Origin (OriginLab, MA).

3.3 Results and Discussion

3.3.1 Rational design of ssDNA-encoded p/MHC tetramers

The standard scaffold most frequently used to assemble p/MHC monomers into tetramers is SA-phycobiliprotein (using the protein fluorophores phycoerythrin (PE) or allophycocyanain (APC)) conjugates. Because SA-phycobiliprotein conjugates are produced via chemical cross-linking, most functional groups are exhausted and/or modified, prohibiting the conjugation of ssDNA. In addition, the attachment of molecular fluorophores to native SA reduces the binding capacity for biotin, an effect attributed to the modification of lysine121 that occurs with amide coupling strategies. This residue is in close proximity to the ligand binding pockets (29, 30). To circumvent this, Altman and co-workers (29) employed a recombinant mutant of SA for fluorescent p/MHC tetramer preparations. This variant incorporates a cysteine residue at the carboxy-terminus (31), a site removed from the biotin binding pocket. The conjugation
of cysteine-reactive maleimide derivatives is restricted to the C-terminus because cysteine residues are absent in native SA.

We expressed SAC, coupled the protein with 5’-maleimide ssDNA, and verified the formation of conjugates with mobility shift assays (Figure 3.2). In parallel, ssDNA was coupled to native SA for direct comparison. To test biotin binding capacity, SAC-oligo conjugates were probed with 2-(4′-Hydroxyazobenzene) benzoic acid (HABA) (32), a molecular mimic of biotin with distinct optical density coefficients dependent on whether biotin is bound to SA or not. A biotin:SA molar ratio of association significantly below 4 in the assay would indicate a reduction in biotin binding capacity. Conjugates derived from native SA were greater than one full unit below the expected value (2.86 versus 4.0), while conjugates formed with SAC maintained near optimal (3.7) binding capacity (Figure 3.2C). These conjugates were then tested across 4 different monoclonal T cell populations (2 human TCR-transduced cell lines and 2 murine TCR-transgenic splenocyte cell suspensions). ssDNA-tagged SAC constructs had markedly higher cell capture efficiencies (Figure 3.3B) when compared with p/MHC tetramers prepared with native SA (Figure 3.3A). All subsequent NACS tetramers were prepared with the SAC variant.
Figure 3.2  An engineered variant of streptavidin expressing C-terminal cysteine residues has superior biotin binding capacity compared to native streptavidin post conjugation with ssDNA. (A) The various stages of SAC expression, refolding, and purification were analyzed on a denaturing PAGE gel. The molecular weight of a SAC monomer is ~12kDa. (B) A gel mobility shift assay to verify the formation of ssDNA-SAC conjugates. Individual bands representing SAC-oligo conjugates differing by one DNA strand can be resolved. Lower order SAC-oligo conjugates (1–2 oligos per protein) run “lighter” when compared to unmodified SAC because of the difference in charge/mass density of nucleic acids. Higher order SAC-oligo conjugates corresponding to 3–4 DNA strands per SA were favored. (C) The molecule 2-(4’-Hydroxyazobenzene) benzoic acid (HABA) was used to determine the molar ratio of association of biotin to SA. Native SA-oligo conjugates bound ~2.9 moles of biotin per mole of SA, a significant decline when compared to the 4:1 ratio of unmodified SAC. SAC-oligo conjugates maintained near optimum binding capacity (3.7:1).
Figure 3.3  T cell capture efficiency is optimal when utilizing ssDNA-SAC conjugates to generate NACS p/MHC tetramers (A) ssDNA-p/MHC constructs derived from native SA were used to sort 4 different human/murine transgenic T cell populations. The T cell capture efficiencies were highly varied amongst the four T cell populations. (B) ssDNA-p/MHC tetramers derived from ssDNA-SAC conjugates were used to sort the four T cell populations. The resulting cell capture efficiencies were markedly improved over native SA-oligo conjugates,
demonstrating that SAC is necessary for the production of high affinity ssDNA-p/MHC tetramers.

3.3.2 Performance of p/MHC arrays produced via DNA immobilization and direct spotting

We directly compared the performance of NACS with conventional direct spotting strategies on various model substrates. The substrates were selected to represent the spectrum of surface chemistries typically used to immobilize proteins (covalent, electrostatic, hydrophobic, and hydrophilic adsorption). Serial dilutions of a fluorescent p/MHC tetramer, MART-1, were spotted on the substrates according to manufacturer’s instructions. Jurkat$^{\alpha$-MART-1} T cells (the human T leukemia cell line Jurkat transduced with the F5 MART-1 TCR (33) specific for MART-1) were applied to the array and representative images collected (Figure 3.4A) and quantified (Figure 3.4B). We observed little to no T cell capture (electrostatic, hydrophilic) or significant noise (hydrophobic) on the majority of the surfaces investigated compared to NACS arrays immobilized with identical concentrations of p/MHC tetramers. T cell binding was observed on one surface (covalent) but cell capture was highly variable as evidenced by both intra-spot and inter-spot heterogeneity and cross experimental variation (Figure 3.5). Moreover, to achieve equivalent T cell capture densities, NACS p/MHC arrays required $>5$ times less material than covalent immobilization. (p/MHC monomer at half max $\equiv K_{1/2} = 1.1$ ng for NACS and 5.7 ng for covalent immobilization).
The performance and reproducibility of NACS p/MHC arrays is markedly improved and represents an integral step towards expanding array-based T cell detection schemes for broader applications. This likely has a few causes. First, surface-tethered ssDNA-p/MHC tetramers may enjoy greater orientational freedom at the surface/solution interface compared with adsorbed proteins which are required to conform to the surface. This effect may increase the density of functional protein and consequently reduce the amount of material required for array production. Second, the hydration state of the environment during the production and subsequent storage of protein arrays is an important factor for array reproducibility (10, 13, 18). This effect is minimized with NACS because DNA chips can be printed and stored dry for extended periods of time and ssDNA-tagged p/MHC tetramer arrays are self-assembled in solution immediately prior to an experiment.
Figure 3.5  Comparison of the performance of p/MHC arrays produced by NACs and by spotting The consistency and robustness of T cell immobilization with NACS is evident when compared directly with spotted arrays, which suffers from significant levels of inter-spot, intra-spot, and inter-experimental heterogeneity. Each row represents a separate experiment performed on a different slide.

3.3.3 NACS specificity and limit of detection

To evaluate the specificity of p/MHC array assembly and T cell sorting, a ssDNA-p/MHC tetramer, tyrosinase, with pendant DNA sequence A’ was hybridized to a DNA microarray printed with the complementary strand (A) along with two additional distinct sequences (designated B and C). A homogeneous population of Jurkatα-Tyr cells (Jurkat cells transduced with a TCR specific for tyrosinase) (34) was then applied to the array. Jurkatα-Tyr T cells localized to the complementary spots (A) containing the hybridized
cognate p/MHC but not to spots printed with the non-complementary sequences B and C (Figure 3.6A). The mean binding capacity calculated from three spots (~600 μm) was ~1486 ± 62 Jurkatα-Tyr T cells.

**Figure 3.6** Mulitplexed nucleic acid cell sorting of antigen-specific T cells (A) Tyrosinase p/MHC tetramer conjugated to ssDNA sequence A’ was hybridized to an array printed with DNA complement strand A and non-complement strands B and C (dashed circles). Jurkatα-Tyr cells were localized to spot A only. (B) A 1:1:1 mixture of Jurkatα-MART-1 (red), Jurkatα-Tyr (green) and Jurkatα-NY-ESO-1 cells was selectively sorted on an array hybridized with MART-1, tyrosinase, and NY-ESO-1 p/MHC ssDNA-tetramers (top left). The remaining three panels are representative images of spots A, B and C.
To illustrate the multiplexing capability of NACS, MART-1, tyrosinase, and NY-ESO-1 ssDNA-p/MHC tetramers encoded to DNA sequences A, B and C respectively were combined and assembled simultaneously to a three element DNA microarray (strands A, B, and C). A 1:1:1 mixed population of Jurkat$^{\alpha}$-MART-1, Jurkat$^{\alpha}$-Tyr and Jurkat$^{\alpha}$-NY-ESO-1 (Jurkat human T cells transduced with the TCR (35) specific for NY-ESO-1) cells prestained with lipophilic dyes (red, green and blue respectively) was applied to the array and localized into alternating columns (Figure 3.6B). Minimal cross-reactivity was observed. The average density of spots was about a factor of three less than homogeneous sorting (440 ± 28 T cells/spot).

To determine the limit of detection, target populations of Jurkat$^{\alpha}$-Tyr cells were spiked in at 10%, 1% and 0.1% into wild type (w.t.) Jurkat cells and sorted (Figure 3.7). The T cell capture density per spot per species for each mixture was enumerated and averaged (right panel). The number of non-specific w.t. Jurkat cells that adhered to the array was constant throughout all dilutions while the number of Jurkat$^{\alpha}$-Tyr T cells captured per spot decreased linearly in relation to the fractional composition of Jurkat$^{\alpha}$-Tyr cells with a detection limit that was ~ 1 in 1000 cells—a limit that corresponds well to the total number of cells that can be captured per spot. Thus, the sensitivity of this approach is strictly a geometric constraint since antigen-specific T cells that settle on inert areas cannot sample and bind to their cognate p/MHC tetramer. The sensitivity can be improved by increasing the size of the capture region (i.e. increase spot diameter and/or incorporate spot redundancy) or by reducing inert regions (i.e. increase printing density).
It should be noted, however, that the sensitivity of this approach cannot be increased without limiting scalability. Since spot diameters are required to be sufficiently large to detect a population at low frequency, this sets an upper bound on the number of distinct spots that can be patterned on a substrate. In the current instance, 600 μm spots are printed in 12 by 12 grids (~1 in²), enabling the potential identification of 144 distinct antigen-specificities ≥ 0.1% from $10^6$ T cells ($10^6$ cells is typically required to cover a 1 in² region). In order to detect target populations below 0.1% without decreasing multiplexing, p/MHC-specific enrichment strategies (e.g. magnetic bead-based schemes (36)) can be implemented prior to cell sorting. Lastly, the recovery of this technique—defined as the total number of antigen-specific T cells captured as a fraction of the number of T cells applied to the array—is low relative to other sorting technologies like FACS and bead-based schemes. Typical T cell recovery for NACS is ≤ 20% (37). In comparison, fluorescent-based sorting techniques like FACS have high recoveries, since
each stained T cell can be identified individually by the cytometer and sorted from the null population. The recovery (as well as sensitivity) of NACS will likely be improved by circulating T cells over the p/MHC array with agitation or with integrated microfluidic devices to allow T cells to sample the entire array.

### 3.3.4 Selective Release of Immobilized T cells with Restriction Endonucleases

Antigen-specific T cells immobilized onto glass are immediately available for secondary assays, since many such as immunohistochemistry (IHC), fluorescent in situ hybridization (FISH) and cytokine secretion assays (6, 8) are traditionally performed or are compatible with cells localized to a substrate. However, several other relevant assays, such as those designed to assess T cell phenotype or functional status like genomic/mRNA analysis or simply further culture for phenotypic enrichment would require a method for releasing the captured cells. Any release scheme should ideally be selective for given cell types. For NACS, we explored whether the DNA tethers could be selectively cleaved by exploiting the sequence specificity of restriction endonucleases. We integrated unique restriction sites to each DNA sequence employed for cell sorting, and found that the adhesion of different populations of antigen-specific T cells could, in fact, be independently controlled (Figure 3.8). Jurkat$^{\alpha\text{-MART-1}}$ and Jurkat$^{\alpha\text{-Tyr}}$ cells prestained with lipophilic dyes (red and green respectively) were sorted on an array printed with DNA sequences $A_{\text{EcoRI}}$ and $B_{\text{BamHI}}$ (Figure 3.8i). These oligonucleotides were modified by incorporating 6 bp restriction sites specific for endonucleases EcoRI and BamHI respectively. After T cell immobilization, the array was treated with BamHI which cleaved the $B_{\text{BamHI}}$ spots and selectively released the bound Jurkat$^{\alpha\text{-Tyr}}$ cells (Figure 3.8ii). Conversely, on a separate but identically cell sorted array, Jurkat$^{\alpha\text{-MART-1}}$
cells were released after treatment with EcoRI (Figure 3.8iii). A second round of enzymatic treatment with the complementary endonuclease (EcoRI to state (ii) or BamHI to state (iii)) removed the remaining adherent cells (iv, v). Alternatively all captured cells (i) could be released non-selectively in a single step with the addition of DNase (vi).

Figure 3.8 Programmed release of sorted T cells by endonuclease cleavage (A) DNA sequences containing EcoRI and BamHI restriction sites were used for T cell sorting. (B) Fluorescent images of Jurkatα-MART-1 (red) and Jurkatα-Tyro (green) cells captured on p/MHC array (i) and after treatment with BamHI (ii) or EcoRI (iii). Only cells localized to DNA spots containing the target restriction sequence were released. A second round of enzymatic treatment released the remaining bound cells (iv, v). Panel (vi) illustrates the array after treatment of state (i) with DNase.
3.3.5 NACS sorting of endogenous primary human T cells

Detection of primary human T cells isolated from peripheral blood is generally more demanding than cultured cell lines because a single population of antigen-specific T cells is present within a large background of differing blood cells and of T cells expressing monoclonal and polyclonal TCRs of diverse specificities. In addition, these T cells would be expressing endogenous levels of TCR. We explored whether the same attributes of NACS that were found in the above examples would apply equally to endogenous primary human T cells. Frozen leukapheresis samples from patient NRA13 were CD8+ enriched and applied to a CMV and Epstein-barr virus (EBV BMLF1 \textsubscript{259-267}/HLA-A2.1) p/MHC array. T cells were captured only within the EBV regions only (Figure 3.9A). NRA13 CD8+ cells were verified by flow cytometry to be \( \sim 5\% \) EBV-specific and \( \sim 0\% \) CMV-specific (Figure 3.10A). The registry of the array was determined with the addition of A’-cy3 (red) and B’-cy5 (blue) conjugates.

For multiplexed detection, a 1:1 mixture of EBV-specific and CMV-specific CD8+ T cells was produced by combining NRA13 lymphocytes with CMV-specific T cells from patient NRA11 (Figure 3.10B). Following cell sorting and fluorescent p/MHC tetramer staining, the populations were complementary stained for the appropriate antigen-specificity (Figure 3.9B). The detection limit of antigen-specific T cells was evaluated from using serially diluted mixtures of EBV-specific T cells (\( \sim 0.4\% \), \( 0.2\% \), and \( 0.1\% \) by FACS) that were probed on an array. Isolated hits were resolved in frequencies as low as \( \sim 0.1\% \) (Figure 3.11, red arrows). The number of unstained cells within the capture regions (black arrow) was constant throughout all dilutions (\( \sim 1–2 \) cells/spot) and likely represents the level of background from non-specific interactions.
It should be noted that while we incorporated fluorescent p/MHC tetramer staining after T cell immobilization for illustrative purposes, the specificity of the captured cells could be determined solely from the registry of the array.

Figure 3.9 NACS sorting of endogenous primary human T cells specific for Epstein-Barr virus and Cytomegalovirus (A) CD8+ T cells from patient NRA 13 were captured on EBV p/MHC spots and no T cells were captured on CMV spots (left panel). The right two panels are representative images after the cells were stained with fluorescent EBV (blue) and CMV p/MHC tetramers (red). (B) T cells detected from a 1:1 mixture of NRA11 and NRA 13 (left panel) were verified to be specific for EBV and CMV (right panels).
Figure 3.10 The T cell specificities of PBMCs from patients NRA11 and NRA13 (A) Lymphocytes isolated from NRA13 contained significant levels of EBV specific T cells (4.9%) with minimal CMV specific T cells. (B) Lymphocytes isolated from NRA11 contained high levels of CMV specific T cells (9%) with a low population of EBV-specific cells (0.12%). Leukapheresis fractions from both patients were kept as frozen aliquots before thawing, CD8+ enrichment and NACS sorting.
3.3.6 Persistence of MART-1 specific, TCR-engineered human T cells in vivo

One important potential application of NACS is to utilize the technique for monitoring cancer patients that are undergoing a particular type of immunotherapy that involves TCR engineering of peripheral blood mononuclear cells (PBMCs). This is an emerging clinical approach to rapidly generate large numbers of tumor antigen-specific T cells. In particular, a cancer patient’s T cells are collected, genetically modified to express a TCR specific for a desired cancer antigen, and then they are introduced back into the patient as a cancer therapy. After re-infusion into the patient, the modified T cells can subsequently traffic to and engage with cancer cells, promoting tumor regression in a subset of patients with metastatic cancers. This type of directed cell
therapy is utilized primarily in patients with melanoma (skin cancer), but also for other cancers (38, 39). This approach is called adoptive cell transfer (ACT). We monitored the presence and abundance of TCR-engineered T cells in peripheral blood of a patient undergoing ACT. White blood cells were collected from patient F5-1 with metastatic melanoma. After in vitro expansion and TCR transduction, the cells were > 80% specific for the MART-1 p/MHC tetramer (Figure 12A). These cells were subsequently infused into patient F5-1 and the presence and persistence of MART-1-specific T cells in the peripheral blood was monitored by NACS at days 0 (prior to infusion), 9, 14 and 30 (Figure 12B). MART-1-specific T cells were not detectable in pre-infusion samples from the patient but they were detectable at all subsequent time points. This is represented as a gradual rise in the abundance of MART-1-specific T cells until day 30 (left axis). In comparison, the frequency of MART-1-specific T cells detected by flow cytometry spiked initially before decreasing and finally increasing by day 30 (right axis).

The abundance of TCR-engineered T cells, as measured by NACS, was generally correlated with parallel measurements using flow cytometry. The prospect of array-based T cell detection schemes for experimental cancer immuno-therapies such as ACT will likely increase as more cancer associated antigens are identified and targeted. For example, more than 50 different melanoma associated antigens have thus far been characterized (40, 41), and this means that more T cell types are likely to be employed for future therapies (39, 42, 43). However, the desire to probe larger sets of cancer associated antigens will be compounded by the amount of sample that can be practically collected from a patient. NACS appears to provide a feasible approach towards carrying
out the highly multiplexed cellular measurements that will eventually be required as ACT-like therapies move forward.

Figure 3.12 Monitoring the presence of infused MART-1-specific, TCR-engineered T cells
(A) Flow cytometry plots (top panels) and NACS (bottom panels) monitoring the transduction of T cells with a MART-1-specific TCR. (B) Blood cells from patient F5-1 were collected at days 0, 9, 14 and 30 after infusion. The abundance of MART-1-specific T cells was monitored by NACS (left axis) and flow cytometry (right axis).

3.3.7 Homogeneous platform for cell sorting and functional analysis

Antigen-specific T cells that are sorted by NACS are immobilized to a solid surface and are immediately available for further studies. Traditional surface bound assays such as IHC, FISH and ELISPOT should integrate seamlessly with NACS. In addition, antigen-specific T cells have been shown to be activated upon binding to p/MHC arrays (6–8) and secrete various immunomodulatory cytokines. The cytokine secretion profiles of CD8+ T cells provide valuable information concerning the phenotype of the T cell (e.g. effector and anergic phenotypes can be differentiated by the
production of IFN-γ). Thus integrating on-chip cytokine measurements with antigen-specific T cell capture would greatly expand the diagnostic capability of NACS p/MHC tetramer arrays.

We proceeded by integrating an ELISPOT-type sandwich assay with p/MHC NACS to detect cytokines produced by captured murine TCR transgenic splenocytes “on-the-spot” (Figure 3.13). Three murine anti-cytokine antibodies (IL-2, IFN-γ and TNF-α) were encoded with DNA strands A’, B’, and C’ respectively. H-2Kb-OVA257-264 ssDNA-p/MHC tetramers were encoded to all three strands. The ssDNA-p/MHC tetramers and antibody conjugates were pooled and assembled to a microarray printed with the complementary strands A, B and C. Murine OT1 lymphocytes (derived from TCR transgenic mice in which most splenocytes are specific for the model antigen OVA257-264), were then seeded on the array. Following incubation periods of 2, 5, or 18 hours, pooled cytokine detection antibodies were added and the slide imaged by confocal
microscopy (Figure 3.14A). The inflammatory cytokine IFN-γ was detected at time points 5 and 18, manifest as discrete diffusive clusters (~50–100 μm in diameter at 5 hrs) that increased in average diameter temporally, attributable to molecular diffusion and sustained secretion. Examination of the vicinity of each burst showed that underlying each fluorescent cluster was a single cell while neighboring cells appeared to be non-responders (Figure 3.14B), suggestive that each IFN-γ burst was derived from a single cell. The number of IFN-γ clusters remained constant at ~3 between hours 5 and 18, indicating no increase in the number of activated T cells between those hours. No significant levels of murine IL-2 and TNF-α were detected at these time points.

Figure 3.14 Dynamic cytokine profiling of NACS sorted murine OT1 lymphocytes (A) Discrete IFN-γ clusters were detected at 5 and 18 hours with no detectable presence of IL-2 and TNF-α. No increase in the average number of IFN-γ clusters between time points 5 and 18 hours was observed. (B) Representative confocal and brightfield images of an IFN-γ burst. Isolated T cells were detected (arrow) within the diffusive IFN-γ regions surrounded by non-responsive T cells.
3.4 Conclusions

We have described a method for generating robust and modular p/MHC arrays for high efficiency T cell sorting. The inclusion of a larger set of orthogonal DNA sequences (20, 25) will enable the modular assembly of higher order p/MHC arrays for T cell screening experiments (e.g. one working set of DNA sequences can be used interchangeably to generate any combination of p/MHC arrays). This would find immediate utility in the field of TCR peptide epitope discovery where recently, novel antigen peptides were discovered via high-throughput CD8+ screening experiments utilizing multi-color flow cytometry in mice and humans (44, 45) (as many as 2,000 distinct p/MHC tetramers were prepared and tested). NACS arrays have the potential to streamline such experiments. Although traditional methods of producing single p/MHC monomers are time and labor intensive, recent reports using conditional peptide exchange technology enables the relatively straightforward construction of 1000 element p/MHC libraries rapidly (44–46). The integration of NACS with these peptide exchange technologies is a realistic option.

We have also demonstrated a number of advantages of the NACS platform. It significantly outperforms literature approaches that utilize surface-bound p/MHC tetramers to capture cells. It is simple and inexpensive to implement since cell sorting is performed on glass substrates prepared via traditional DNA printing technologies. In addition, sorted cells may be selectively released, which should permit for the deployment of a host of bioanalytical methods on NACS sorted cells. We envision that NACS will find uses beyond multiplexed sorting of T cells based on TCR specificity. The principal components of this platform—streptavidin-cysteine core and orthogonal
single stranded DNA sequences—were rationally developed to enable oriented coupling and spatial addressing. Thus this platform is amenable to any family of binding proteins or small molecule binders labeled with biotin. The increase in avidity of p/MHC tetramers over monomers as a consequence of the valency of SA should likewise extend to other capture agents, making it feasible to generate cellular arrays with probes ranging from high to moderate affinities like antibodies, aptamers or peptides.
3.5 References


19. For example, the polyacrylamide surface employed by Davis and co-workers has been discontinued from the manufacturer (Perkin Elmer).


37. For a typical homogeneous sorting experiment, 10^6 T cells are applied to an array containing 144 spots. Each spot can bind to ~1486 ± 62 T cells. The recovery, R, is calculated as R = 100% x (144 x 1486)/10^6 = 20%. The recovery is valid even with heterogeneous samples (e.g. 1:1 mixture of two antigen-specific T cell populations) when the entire array is used to sort a single target population (R = 100% x (144 spots x 661)/(5x10^5) ≈ 20%). For multiplexed experiments, R is proportional to the fraction of the 144 spots dedicated for cell capture. For example, sorting 2 populations from a 1:1 mixture, R = 100% x (72 spots x 661)/(5x10^5) ≈ 10%.


3.6 Appendix A: Protein Sequences

Core streptavidin + Cys in pET-3a vector (NdeI/BamHI)

NdeI
1 CATATGGGCA TCACCGGCAC CTGGTACAAC CAGCTCGGCT CGACCTTCAT
   H M I T G T W Y N Q L G S T F I
51 CGTGACCCGC GGCGCCGACG GCGCCCTGAC CGGAACCTAC GAGTCGGCCG
   V T A G A D G A L T G T Y E S A V
101 TCGGCAACGC CGAGAGCCGC TACGTCTCTGA CCGTCTGTTA CGACACGCGCC
   G N A E S R Y V L T G R Y D S A
151 CCGGCCACCG ACGGCAGCGG CACCGCCCTC GGGACCTGAC GAGTCGGCCG
   P A T D G S G T A L G W T V A W K
201 GAATAACTAC CGCAACCGCC AACTCCGCACG CAGTGGAGAC CGACCAGTACG
   N N Y R N A H S A T T W S G Q Y V
251 TCGGCACGGC CGAGCAGGCAG ATCAACACCC AGTGGGCTGCT GGGCCCTGAC
   G G A E A R I N T Q W L T S G
301 ACCACCGAGG CCAACGCCTG GAAGTCCACG CTGGTCGGCC ACGACACGTT
   T T E A N A W K S T L V G H D T F
BamHI
351 CACCAAGGTG GGTGGTTCTCT GTGGCCGCTAG GGATCC
   T K V G G S G C P *

Human \(\beta_2m\) in pET-3a vector NdeI/BamHI

NdeI
1 CATATGATCC AGCGTACTCC AAGATTCAG GTTTACTCAC GTCATCCACG
   H M I Q R T P K I Q V Y S R H P A
51 AGAGAATGGGA AAGTCAAATT TCCTGAATTG CTATGTGTCT GGGTTTCATC
   E N G K S N F L N C Y V S G F H P
101 CATCCGACAT TGAAGTTGAC TTACTGAAGA ATGGAGAGAG AATTGAAAAA
   S D I E V D L L K N G E R I E K
151 GTGGAGGATT CAGACTTCTC TTTCCAGCAAG GACTGGTCTT TCTATCTCTT
   V E H S D L S F S K D W S F Y L L
201 GTATTTACTG GAATTCGCC AACTGAAAGA AGATGAGATAG GCCCTGCGCTG
   Y Y T E F T P T E K D E Y A C R V
99

251 TGAACCACGT GACTTTGTCA CAGCCCAAGA TAGTTAAGTG GGATCGAGAC

                N H V T L S Q P K I V K W D R D

BamHI

301 ATGTAA GGATCC

M *

Human HLA-A2.1 (biotin tag)

1 ATGGGCTCTC ACTCCATGAG GTATTTTCCTC ACATCCGTGT CCCGGCCCGG

                M G S H S M R Y F F T S V S R P G

51 CCGCGGGGAG CCCCGCTTCA TCGCAGTGGG CTACGTGGAC GACACGCAGT

                R G E P R F I A V G Y V D D T Q F

101 TCGTGCGGTT CGCACAGCCGAG CCCCGCGAGCC AGAGGATGGA GCCCGGCGCGG

                V R F D S D A A S Q R M E P R A

151 CCGTGGATAG AGCAGAGAGG TCCGGAGTAT TGGGACCGGG AGACACCGAA

                P W I E Q E G P E Y W D G E T R K

201 AGTGAAGGCC CACTCACAGA CTCACCGAGT GGACCTGGGG ACCCTGCGCGG

                V K A H S Q T H R V D L G T L R G

251 GCTACTACAA CCAGAGCGAG GCCGTTTCCTC ACACCGTCCA GAGGATGTAT

                Y Y N Q S E A G S H T V Q R M Y

301 GGCTGCGGACG TGGGGTGCGGA CTGGCGCGCTTC ATCCCGCGGGT ACCACCGAGA

                G C D V G S D W R F L R G Y H Q Y

351 CGCTACGAGC GGAAGAGGATT ACATCAGCCTC GAAAGAGGAC CTGGCGGCTTT

                A Y D G K D Y I A L K E D L R S W

401 GGACCCGGGC GGACATGCGCA GTCTAGACCA CCAAGCAGCAA GTGGGAGGGCG

                T A A D M A Q T T K H K W E A

451 GGGCATGTGG CCGAGCAGTT GAGAGCCTAC CTGGAGGCGA CTGGCGGTGGA

                A H V A E Q L R A Y L E G T C V E

501 GTGGCTTCCGC AGATACTCGG AGAAAGGGGAA GGAGACGCTG CAGCGCAGCGG

                W L R R Y L E N G K E T L Q R T D

551 ACAGCCCCCCA AAGCATGATG ACTCACCACG CTGTCTCTGA CCATGAAGCC

                A P K T H M T H H A V S D H E A

601 ACCCTGAGGT GCTGGGCCCT GAGCTTCTAC CCTGCGGAGA TCACACTGAC

                T L R C W A L S F Y P A E I T L T

651 CTGGCAGCGG GATGGGAGAG ACCAGACCCA GGACACGGGAG CTGGTGGAGA
BirA ligase biotin sequence: GGGLNDIFEAQKIEWH

Murine β₂m

1 ATGATCCAGA AAACCCCTCA AATTCAAGTA TACTCACGCC ACCCACCAGA M I Q K T P Q I Q V Y S R H P P E
51 GAATGGGAAG CCGAACATAC TGAAGCTGTA CGTAACACAG TTCCACCAGC N G K P N I L N C Y V T Q F H P P
101 CTCACATTGA AATCCAAATG CTGAAGACCG GGAAAATAT TCCTAAGAT A H I E I Q M L K N G K K I P K V
151 GAGATGTCAG ATATGTCCTT CAGCAAGGAC TGGCTTTTCT ATATCCGTGC E M S D M S F S K D W S F Y I L A
201 TCACACTGAA TTCACCCCCA CTGAGACTGA TACATACGCC TGCAGAGTTA H T E F T P T E T D T Y A C R V K
251 AGCATGACAG TATGGCCAG CCAAGGACCG TCTACTGGGA TCGAGACCTA H D S M A E P K T V Y W D R D M
301 TGA *

Murine Kb (biotin tag)

NdeI

1 CATATGGGTC CACACTCTCT CGGCTATTTC GTTACGGCTG TTAGCCGTCC H M G P H S L R Y F V T A V S R P
51 GGGTCTGGGT GAGCCGCGCT ACATGGAAGT CGGTTACGTC GACGACACCG
<table>
<thead>
<tr>
<th>Start Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>101</td>
<td>AATTCGTTGC TTTCGACAGC GACGCCGGAGA ACCCGCGTGA TGAGCCGGCTFVRFDSDAENPREWERTQ</td>
</tr>
<tr>
<td>151</td>
<td>GCGCGTTTGA TGGAGCAGGA AGGTCCCGAG TGCTTGGATCTG CGCACTCTGC KAGNEQSFVRDLRTL</td>
</tr>
<tr>
<td>201</td>
<td>AAAGGCAGAG GGCAATGAAC AGAGCTTTCG TGTTGATCTG CGCACTCTGC KAKGNEQSFVRDLRTL</td>
</tr>
<tr>
<td>251</td>
<td>TGGTTTACTA CAACCAAGGC AAAGGTGGCA GCCATACCAT TCAAGTGATTGYYNQSKGSHTIQVI</td>
</tr>
<tr>
<td>301</td>
<td>AGCGGTTGTA AAGTGGGACTC TGATGGCCGC CTGTTGGGCTG GTTATCAGCA SGCEVGSDRLLG YQQ</td>
</tr>
<tr>
<td>351</td>
<td>ATATGCACTAC GACGGTTGCG ACTACATTGC GCTGAATGAA GATCTGAAAYAYDGCDYIAALNEDDLK</td>
</tr>
<tr>
<td>401</td>
<td>CGTGGACTGC GCCGGGACATG GCCGCACCTGA TTACAAACCA CAAAGTGGGAG WATADMAALITKHKE</td>
</tr>
<tr>
<td>451</td>
<td>CAAGCCGGGC CAGCCGGACGC CCTGCGTGCG TATCTGGGAG GCACCTGCTGTQAGEAERLRAYLEGTCV</td>
</tr>
<tr>
<td>501</td>
<td>GGAATGGCGTA CGCCGGCTATC TGAAGATGGG CAATGCCACC TGCTGGCTGTA EWLRRLRYLKNGNATL LRT</td>
</tr>
<tr>
<td>551</td>
<td>CGGATCCCCC GAAAGGCGCAC GTGACCGCACCG ATAGCCGTCC TCGGATAAAA DSPKAKHVTHHSRPE DK</td>
</tr>
<tr>
<td>601</td>
<td>GTTACCGCTGC GTTGCTTGGGC ACTGGGCTTT TACCCGCCAG ATATCACCTTT VTLRCWALGFYPADITL</td>
</tr>
<tr>
<td>651</td>
<td>GACGGTGGCA AAGGTGCTGAT TCAGATGGAA GAACTGGGTGG TWQLNIGEQDMELVE</td>
</tr>
<tr>
<td>701</td>
<td>AGACTCGTCC GCCGGACGCG ACGTCACCTCC AGAAATGGGC ATCGGGTCGTG TRPAGDGFQKWASVV</td>
</tr>
<tr>
<td>751</td>
<td>GTCCTCTCGG GTAAAGAGCA ATACTATACC TGCCACGTTTT ACCACCAAGGVPLGKEQYYTCHVYHQG</td>
</tr>
<tr>
<td>801</td>
<td>TCTGGCGGAG CGGTCACTCT TGGCGTGGGA GCCACGCGGAG ACCACCGGCA LPETPLTLRWEPPSTGS</td>
</tr>
<tr>
<td>851</td>
<td>GCAGTTGTTAG CGCGGTTTCC CGCGGTGGCG GTCTGAACGA CATCTTGTAG GGS GGG SAGGLNDIFE</td>
</tr>
</tbody>
</table>

**BamHI**

<table>
<thead>
<tr>
<th>Start Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>901</td>
<td>GCCCGAAGAGA TCGAGTGGCA TTAAGGATCC</td>
</tr>
</tbody>
</table>
A Q K I E W H *

BirA ligase biotin sequence: GGGLNDIFEAQKIEWH

Murine Db (biotin tag)

1 ATGGGGCCAC ACTCGATGCG GTATTTTCGAG ACCGCCGTGT CCCGGCCCGG
   M G P H S M R Y F E T A V S R P G
51 CCTCGAGGAG CCCCGGTACA TCTCTGTCGG CTATGTGGAC AACAAGGAGT
   L E E P R Y I S V G Y V D N K E F
101 TCGTGCGCTTT CGACAGCGAC GCGGAGAATC CGAGATATGA GCCGCGGGCG
   V R F D S D A E N P R Y E P R A
151 CGGTGGATGG AGGCAGGAGG GCCGCGGTAT TGGGACGGGG AAACACAGAA
   P W M E Q E G P E Y W E R E T Q K
201 AGCCAAAGGGC CAAGAGCGATG GTGTCGGAGT GAGCGGTCGG AACCTGCTCG
   A K G Q E Q W F R V S L R N L L G
251 GCTACTACAA CCAGAGCGCG GCGGCTCTTC ACACACTCCA GCAGATGTCT
   Y Y N Q S A G G S H T L Q Q M S
301 GGCTGTGACT TTGGGCGCGCA CTGCGCGCTTC CTCCGCGGGT ACCTGCAGTT
   G C D L G S D W R L R G Y L Q F
351 CGCCCTAGAA GGGCCCGGATT ACACTCGGCTT GAGCAGAAGAC CTGAAAACGT
   A Y E G R D Y I A L N E D L K T W
401 GGACGCGGGGC GGACATGGCG GCAGGGCTCAG CCCAGCGCAA GTGGGAGCAG
   T A A D M A A Q I T R R K W E Q
451 AGTGGTGCTT CAGAGCATTG CAGACCCGCTT CAGGCGGGCG AGTGCGTGGAA
   S G A A E H Y K A Y L E G C V E
501 GTGCTGTCAC AGATACCTGA AGAACGGGGA CGCGACCAGTG CTGCACACAG
   W L H R Y L K N G N A T L L R T D
551 ATTCACAAAA GCACATGGTG ACCCATCACC CGGACGTCTAA AGGTGGAACTC
   S P K A H V T H H P R S K G E V
601 ACCTCGAGGG TGTCGGCCCT GGCTTCTACTAC TCGTGACCA TCACCTGAC
   T L R C W A L G F Y P A D I T L T
651 CTGGCAGTTT AAATGGGAGG AGCTGACCCA GGACATGGAG CTTGAGGGAG
   W Q L N G E E L T Q D M E L V E T
701  CCAGGCCTGC  AGGGGATGGA  ACCTTCCAGA  AGTGGGCA[T]C  TGTGGTGGTG
     R  P  A  G  D  G  T  F  Q  K  W  A  S  V  V  V
751  CCTCTTGGGA  AGGAGCAGAA  TTACACATGC  CGTGTGACC  ATGAGGGGCT
     P  L  G  K  E  Q  N  Y  T  C  R  V  Y  H  E  G  L
801  GCCTGAGCCC  T[TC]CACCCTGA  GATGGGAGCC  TCCTCCATCC  ACTGGATCCG
     P  E  P  L  T  L  R  W  E  P  P  S  T  G  S  G
851  GTGGTTCCGG  TGGTTCCCGCG  GGTGTTGGTT  TGAACGACAT  CTTCGAAGCT
     G  S  G  G  S  A  G  G  G  L  N  D  I  F  E  A
901  CAGAAAATCG  AATGGC[ACTA]  A
     Q  K  I  E  W  H  *
3.7 Appendix B: Chromatography

3.7.1 Iminobiotin Purification of SAC

Figure 3.15 SAC purification with an iminobiotin agarose support column. (A) The structure of biotin. (B) Iminobiotin, the structural analog of biotin, has two distinct states; (1) a neutral, high affinity SAC binding state at pH 11 and a cationic state, low affinity SAC binding state at pH 4. (C) Elution profile of SA compared to SAC refolded in various buffers.
3.7.2 FPLC of SAC-ssDNA conjugates

*0.5ml/min isocratic flow of PBS

Figure 3.16 Fast protein liquid chromatography of ssDNA-SAC conjugates. A typical successful conjugation reaction will yield three distinct peaks, corresponding to the ssDNA-SAC conjugates, excess unreacted ssDNA, and excess small molecules. It is possible to resolve shoulders in the ssDNA-SAC peak, corresponding to extent of ssDNA modification.